A First Course in Linear Algebra

by
Robert A. Beezer
Department of Mathematics and Computer Science
University of Puget Sound

Version 0.1
July 27, 2004

Preface

This textbook is designed to teach the university mathematics student the basics of the subject of linear algebra. There are no prerequisites other than ordinary algebra, but it is probably best used by a student who has the "mathematical maturity" of a sophomore or junior.

The text has two goals: to teach the fundamental concepts and techniques of matrix algebra and abstract vector spaces, and to teach the techniques of developing the definitions and theorems of a coherent area of mathematics. So there is an emphasis on worked examples of non-trivial size and on proving theorems carefully.

This book is free. That means you may use it at no cost, other than downloading it, and perhaps choosing to print it. But it is also free in another sense of the word, it has freedom. While copyrighted, this is only to allow some modicum of editorial control over the central portions of the text. Otherwise, it has been designed to be expanded through contributions from others, and so that it can be easily customized for different uses. It will not ever go "out of print" nor will updates be designed to frustrate the used book market. You may make as many copies as you want, and use them however you see fit.

Topics The first half of this text (through Chapter M [97]) is basically a course in matrix algebra, though the foundation of some more advanced ideas is also being laid in these early sections. Vectors are presented exclusively as column vectors (since we also have the typographic freedom to avoid the cost-cutting move of displaying column vectors inline as the transpose of row vectors), and linear combinations are presented very early. Spans, null spaces and ranges are also presented very early, simply as sets, saving most of their vector space properties for later, so they are familiar objects before being scrutinized carefully.

You cannot do *everything* early, so in particular matrix multiplication comes late. However, with a definition built on linear combinations of column vectors, it should seem more natural than the usual definition using dot products of rows with columns. And this delay emphasizes that linear algebra is built upon vector addition and scalar amultiplication. Of course, matrix inverses must wait for matrix multiplication, but this doesn't prevent nonsingular matrices from occurring sooner. Vector space properties are hinted at when vectors and matrices are first defined, but the notion of a vector space is saved for a more axiomatic treatment later. Once bases and dimension have been explored in the context of vector spaces, linear transformations and their matrix representations follow. The goal of the book is to go as far as canonical forms and matrix decompositions.

Linear algebra is an ideal subject for the novice mathematics student to learn how to develop a topic precisely, with all the rigor mathematics requires. Unfortunately, much of this rigor seems to have escaped the standard calculus curriculum, so for many university students this is their first exposure to careful definitions and theorems, and the expectation that they fully understand them, to say nothing of the expectation that they become proficient in formulating their own proofs. We have tried to make this text as helpful as possible with this transistion. Every definition is stated carefully, set apart from the text. Likewise, every theorem is carefully stated, and almost every one has a complete proof. Definitions and theorems are catalogued in order of their appearance in the front of the book, and in the index at the back. Along the way, there are discussions of some of the more important ideas relating to formulating proofs (Proof Techniques), which is advice mostly.

Freedom This book is freely-distributable. When it is ready, I plan to release the complete edition under something like a Creative Commons license that will formalize this arrangement. Once it settles down, I also plan to release the TEX source as well. This arrangement will provide many benefits unavailable with traditional texts, such as:

- No cost, or low cost, to students. With no physical vessel (i.e. paper, binding), no transportation costs (Internet bandwidth being a negligible cost) and no marketing costs (evaluation and desk copies are free to all), it can be obtained by anyone with a computer and an Internet connection, and a teacher can make available paper copies in sufficient quantities for a class. The cost to print a copy is not insignificant, but is just a fraction of the cost of a regular textbook. Students will not feel the need to sell back their book, and in future years can even pick up a newer edition.

- The book will not go out of print. No matter what, a teacher will be able to maintain their own copy and use the book for as many years as they desire. Further, the naming schemes for chapters, sections, theorems, etc. is designed so that the addition of new material will not break any course syllabi.

- With many eyes reading the book and with frequent postings of updates, the reliability should become very high. Please report any errors you find that persist into the latest version.

- For those with a working installation of the popular typesetting program TEX, the book has been designed so that it can be customized. Page layout, presence of exercises and/or solutions, presence of sections or chapters can all be easily controlled. Furthermore, various pieces of mathematical notation are achieved via TEX macros. So by changing a single macro, one's favorite notation can be reflected throughout the text. For example, every transpose of a matrix is coded in the source as `\transpose{A}`, which when printed will yield A^t. However by changing the definition of `\transpose{ }`, any desired alternative notation will then appear throughout the text instead.

- The book has also been designed to make it easy for others to contribute material. Would you like to see a section on symmetric bilinear forms? Consider writing one and contributing it. Does there need to be more exercises about the null space of a matrix? Send me some. Historical Notes? Contact me, and we'll see about adding those in also.

I can think of just one drawback. The author is not guaranteed any royalties.

This project is the result of the confluence of several events:

- Having taught an introductory linear algebra course eighteen times prior, I decided in January 2003 to type up a good set of course notes. Then I decided to make copies for the students, necessitating just a bit more quality control. They became more and more complete as the semester wore on. I used them again for the Fall 2003 semester, and spent a sabbatical during the Spring 2004 semester doing a total rewrite with an eye toward book form, and incorporating many new ideas I had for how best to present the material.

- I've used TEX and the Internet for many years, so there is little to stand in the way of typesetting, distributing and "marketing" a free book.

- In September 2003, both of the textbooks I was using for my courses came out in new editions. Trivial changes required complete rewrites of my syllabi and web pages, and substantial reorganization of the notes that preceded.

- With recreational and professional interests in software developement, I've long been fascinated by the open-source software movement, as exemplified by the success of GNU and Linux, though public-domain TEX might also deserve mention. It should be obvious that this is an attempt to carryover that model of creative endeavor to textbook publishing.

However, much of my motivation for writing this book is captured by H.M. Cundy and A.P. Rollet in their Preface to the First Edition of *Mathematical Models* (1952), especially the final sentence,

> This book was born in the classroom, and arose from the spontaneous interest of a Mathematical Sixth in the construction of simple models. A desire to show that even in mathematics one could have fun led to an exhibition of the results and attracted considerable attention throughout the school. Since then the Sherborne collection has grown, ideas have come from many sources, and widespread interest has been shown. It seems therefore desirable to give permanent form to the lessons of experience so that others can benefit by them and be encouraged to undertake similar work.

How To Use This Book Chapter, Theorems, etc. are not numbered in this book, but are instead referenced by acronyms. This means that Theorem XYZ will always be Theorem XYZ, no matter if new sections are added, or if an individual decides to remove certain other sections. Within sections, the subsections and examples are acronyms that begin with the acronym of the section. So Example XYZ.AB will be found within Section XYZ. At first, all the letters flying around may be confusing, but with time, you will begin to recognize the more important ones on sight. Furthermore, there are lists of theorems, examples, etc. in the front of the book, and an index that contains every acronym. If you are reading this in an electronic PDF version, you will see that all of the cross-references are hyperlinks, allowing you to click to a definition or example, and then use the back button to return. In printed versions, you will have to rely on the page numbers. However, note that page numbers are not permanent! Different editions, different margins, or different sized paper will affect what content is on each page. And in time, the addition of new material will affect the page numbering.

Robert A. Beezer
Tacoma, Washington
August, 2004

Contents

Contributors

Beezer, Robert. University of Puget Sound. HTTP://BUZZARD.UPS.EDU/

Definitions

Notation

Theorems and Corollaries

Examples

Proof Techniques

SLE: Systems of Linear Equations

Section WILA
What is Linear Algebra?

Subsection LA
"Linear" + "Algebra"

The subject of linear algebra can be partially explained by the meaning of the two terms comprising the title. "Linear" is a term you will appreciate better at the end of this course, and indeed, attaining this appreciation could be taken as one of the primary goals of this course. However for now, you can understand it to mean anything that is "straight" or "flat." For example in the xy-plane you might be accustomed to describing straight lines (is there any other kind?) as the set of solutions to an equation of the form $y = mx + b$, where the slope m and the y-intercept b are constants that together describe the line. In multivariate calculus, you may have discussed planes. Living in three dimensions, with coordinates described by triples (x, y, z), they can be described as the set of solutions to equations of the form $ax + by + cz = d$, where a, b, c, d are constants that together determine the plane. While we might describe planes as "flat," lines in three dimensions might be described as "straight." From a multivariate calculus course you will recall that lines are sets of points described by equations such as $x = 3t - 4$, $y = -7t + 2$, $z = 9t$, where t is a parameter that can take on any value.

Another view of this notion of "flatness" is to recognize that the sets of points just described are solutions to equations of a relatively simple form. These equations involve addition and multiplication only. We will have a need for subtraction, and occasionally we will divide, but mostly you can describe "linear" equations as involving only addition and multiplication. Here are some examples of typical equations we will see in the next few sections:

$$2x + 3y - 4z = 13 \qquad 4x_1 + 5x_2 - x_3 + x_4 + x_5 = 0 \qquad 9a - 2b + 7c + 2d = -7$$

What we will not see are equations like:

$$xy + 5yz = 13 \qquad x_1 + x_2^3/x_4 - x_3x_4x_5^2 = 0 \qquad \tan(ab) + \log(c - d) = -7$$

The exception will be that we will on occassion need to take a square root.

You have probably heard the word "algebra" frequently in your mathematical preparation for this course. Most likely, you have spent a good ten to fifteen years learning the algebra of the real numbers, along with some introduction to the very similar algebra of complex numbers. However, there are many new algebras to learn and use, and likely linear algebra will be your second algebra. Like learning a second language, the necessary adjustments can be challenging at times, but the rewards are many. And it will make learning your third and fourth algebras even easier. Perhaps you have heard of "groups" and "rings" (or maybe you have studied them already), which are excellent examples of other algebras with very interesting properties and applications. In any event, prepare yourself to learn a new algebra and realize that some of the old rules you used for the real numbers may no longer apply to this *new* algebra you will be learning!

The brief discussion above about lines and planes suggests that linear algebra has an inherently geometric nature, and this is true. Examples in two and three dimensions can be used to provide valuable insight into important concepts of this course. However, much of the power of linear algebra will be the ability to work with "flat" or "straight" objects in higher dimensions, without concerning ourselves with visualizing the situation. While much of our intuition will come from examples in two and three dimensions, we will maintain an *algebraic* approach to the subject, with the geometry being secondary. Others may wish to switch this emphasis around, and that can lead to a very fruitful and beneficial course, but here and now we are laying our bias bare.

Subsection A
An application: packaging trail mix

We finish this section with a rather involved example that will highlight some of the power and techniques of linear algebra. Work through all of the details with pencil and paper, until you believe all the assertions made. However, in this introductory example, do not concern yourself with how some of the results are obtained or how you might be expected to solve a similar problem. We will come back to this example later and expose some of the techniques used and properties exploited. For now, use your background in mathematics to convince yourself that everthing said here really is correct.

Example WILA.TM
Trail Mix Packaging
Suppose you are the production manager at a food-packaging plant and one of your product lines is trail mix, a healthy snack popular with hikers and backpackers, containing raisins, peanuts and hard-shelled chocolate pieces. By adjusting the mix of these three ingredients, you are able to sell three varieties of this item. The fancy version is sold

in half-kilogram packages at outdoor supply stores and has more chocolate and fewer raisins, thus commanding a higher price. The standard version is sold in one kilogram packages in grocery stores and gas station mini-markets. Since the standard version has roughly equal amounts of each ingredient, it is not as expensive as the fancy version. Finally, a bulk version is sold in bins at grocery stores for consumers to load into plastic bags in amounts of their choosing. To appeal to the shoppers that like bulk items for their economy and healthfullness, this mix has much more raisins (at the expense of chocolate) and therefore sells for less.

Your production facilities have limited storage space and early each morning you are able to receive and store 380 kilograms of raisins, 500 kilograms of peanuts and 620 kilograms of chocolate pieces. As production manager, one of your most important duties is to decide how much of each version of trail mix to make every day. Clearly, you can have up to 1500 kilograms of raw ingredients available each day, so to be the most productive you will likely produce 1500 kilograms of trail mix each day. But how should these ingredients be allocated to the mixing of the bulk, standard and fancy versions?

First, we need a little more information about the mixes. Workers mix the ingredients in 15 kilogram batches, and each row of the table below gives a recipe for a 15 kilogram batch. There is some additional information on the costs of the ingredients and the price the manufacturer can charge for the different versions of the trail mix.

	Raisins (kg/batch)	Peanuts (kg/batch)	Chocolate (kg/batch)	Cost (\$/kg)	Sale Price (\$/kg)
Bulk	7	6	2	3.69	4.99
Standard	6	4	5	3.86	5.50
Fancy	2	5	8	4.45	6.50
Storage (kg)	380	500	620		
Cost (\$/kg)	2.55	4.65	4.80		

As production manager, it is important to realize that you only have three decisions to make — the amount of bulk mix to make, the amount of standard mix to make and the amount of fancy mix to make. Everything else is beyond your control or is handled by another department within the company. Principally, you are also limited by the amount of raw ingredients you can store each day. Let us denote the amount of each mix to produce each day, measured in kilograms, by the variable quantities b, s and f. Your production schedule can be described as values of b, s and f that do several things. First, we cannot make negative quantities of each mix, so

$$b \geq 0 \qquad\qquad s \geq 0 \qquad\qquad f \geq 0.$$

Second, the storage capacity for the raw ingredients leads to three (linear) equations, one for each ingredient,

$$\begin{aligned}
\frac{7}{15}b + \frac{6}{15}s + \frac{2}{15}f &= 380 \quad \text{(raisins)}\\
\frac{6}{15}b + \frac{4}{15}s + \frac{5}{15}f &= 500 \quad \text{(peanuts)}\\
\frac{2}{15}b + \frac{5}{15}s + \frac{8}{15}f &= 620 \quad \text{(chocolate)}
\end{aligned}$$

It happens that this system of three equations has just one solution. In other words, as production manager, your job is easy, since there is but one way to use up all of your raw ingredients making trail mix. This single solution is

$$b = 300\text{ kg} \qquad s = 300\text{ kg} \qquad f = 900\text{ kg}.$$

We do not yet have the tools to explain why this solution is the only one, but it should be simple for you to verify that this is indeed a solution. (Go ahead, we will wait.) Determining solutions such as this, and establishing that they are unique, will be the main motivation for our initial study of linear algebra.

So we have solved the problem of making sure that we make the best use of our limited storage space, and each day use up all of the raw ingredients that are shipped to us. Additionally, as production manager, you must report weekly to the CEO of the company, and you know he will be more interested in the profit derived from your decisions than in the actual production levels. So you compute,

$$300(4.99 - 3.69) + 300(5.50 - 3.86) + 900(6.50 - 4.45) = 2727$$

for a daily profit of \$2,727 from this production schedule. The computation of the daily profit is also beyond our control, though it is definitely of interest, and it too looks like a “linear” computation.

As often happens, things do not stay the same for long, and now the marketing department has suggested that your company’s trail mix products standardize on every mix being one-third peanuts. Adjusting the peanut portion of each recipe by also adjusting the chocolate portion, leads to revised recipes, and slightly different costs for the bulk and standard mixes, as given in the following table.

	Raisins (kg/batch)	Peanuts (kg/batch)	Chocolate (kg/batch)	Cost (\$/kg)	Sale Price (\$/kg)
Bulk	7	5	3	3.70	4.99
Standard	6	5	4	3.85	5.50
Fancy	2	5	8	4.45	6.50
Storage (kg)	380	500	620		
Cost (\$/kg)	2.55	4.65	4.80		

In a similar fashion as before, we desire values of b, s and f so that

$$b \geq 0, \qquad s \geq 0, \qquad f \geq 0$$

and

$$\begin{aligned}
\frac{7}{15}b + \frac{6}{15}s + \frac{2}{15}f &= 380 \quad \text{(raisins)}\\
\frac{5}{15}b + \frac{5}{15}s + \frac{5}{15}f &= 500 \quad \text{(peanuts)}\\
\frac{3}{15}b + \frac{4}{15}s + \frac{8}{15}f &= 620 \quad \text{(chocolate)}
\end{aligned}$$

It now happens that this system of equations has *infinitely* many solutions, as we will now demonstrate. Let f remain a variable quantity. Then if we make f kilograms of the fancy mix, we will make $4f - 3300$ kilograms of the bulk mix and $-5f + 4800$ kilograms of the standard mix. let us now verify that, for any choice of f, the values of $b = 4f - 3300$ and $s = -5f + 4800$ will yield a production schedule that exhausts all of the day's supply of raw ingredients. Grab your pencil and paper and play along.

$$\begin{aligned}
\frac{7}{15}(4f - 3300) + \frac{6}{15}(-5f + 4800) + \frac{2}{15}f = 0f + \frac{5700}{15} = 380\\
\frac{5}{15}(4f - 3300) + \frac{5}{15}(-5f + 4800) + \frac{5}{15}f = 0f + \frac{7500}{15} = 500\\
\frac{3}{15}(4f - 3300) + \frac{4}{15}(-5f + 4800) + \frac{8}{15}f = 0f + \frac{9300}{15} = 620
\end{aligned}$$

Again, right now, do not be concerned about how you might derive expressions like those for b and s that fit so nicely into this system of equations. But do convince yourself that they lead to an infinite number of possibilities for solutions to the three equations that describe our storage capacities. As a practical matter, there really are not an infinite number of solutions, since we are unlikely to want to end the day with a fractional number of bags of fancy mix, so our allowable values of f should probably be integers. More importantly, we need to remember that we cannot make negative amounts of each mix! Where does this lead us? Positive quantities of the bulk mix requires that

$$b \geq 0 \qquad \Rightarrow \qquad 4f - 3300 \geq 0 \qquad \Rightarrow \qquad f \geq 825.$$

Similarly for the standard mix,

$$s \geq 0 \qquad \Rightarrow \qquad -5f + 4800 \geq 0 \qquad \Rightarrow \qquad f \leq 960.$$

So, as production manager, you really have to choose a value of f from the set

$$\{825,\ 826,\ \ldots,\ 960\}$$

leaving you with 136 choices, each of which will exhaust the day's supply of raw ingredients. Pause now and think about which will *you* would choose.

Recalling your weekly meeting with the CEO suggests that you might want to choose a production schedule that yields the biggest possible profit for the company. So you compute an expression for the profit based on your as yet undetermined decision for the value of f,

$$(4f - 3300)(4.99 - 3.70) + (-5f + 4800)(5.50 - 3.85) + (f)(6.50 - 4.45) = -1.04f + 3663.$$

Since f has a negative coefficient it would appear that mixing fancy mix is detrimental to your profit and should be avoided. So you will make the decision to set daily fancy mix production at $f = 825$. This has the effect of setting $b = 4(825) - 3300 = 0$ and we stop producing bulk mix entirely. So the remainder of your daily production is standard mix at the level of $s = -5(825) + 4800 = 675$ kilograms and the resulting daily profit is $(-1.04)(825) + 3663 = 2805$. It is a pleasant surprise that daily profit has risen to \$2,805, but this is not the most important part of the story. What is important here is that there are a large number of ways to produce trail mix that use all of the day's worth of raw ingredients *and* you were able to easily choose the one that netted the largest profit. Notice too how all of the above computations look "linear."

In the food industry, things do not stay the same for long, and now the sales department says that increased competition has lead to the decision to stay competitive and charge just \$5.25 for a kilogram of the standard mix, rather than the previous \$5.50 per kilogram. This decision has no effect on the possibilities for the production schedule, but will affect the decision based on profit considerations. So you revisit just the profit computation, suitably adjusted for the new selling price of standard mix,

$$(4f - 3300)(4.99 - 3.70) + (-5f + 4800)(5.25 - 3.85) + (f)(6.50 - 4.45) = 0.21f + 2463.$$

Now it would appear that fancy mix is beneficial to the company's profit since the value of f has a positive coefficient. So you take the decision to make as much fancy mix as possible, setting $f = 960$. This leads to $s = -5(960) + 4800 = 0$ and the increased competition has driven you out of the standard mix market all together. The remainder of production is therefore bulk mix at a daily level of $b = 4(960) - 3300 = 540$ kilograms and the resulting daily profit is $0.21(960) + 2463 = 2664.6$. A daily profit of \$2,664.60 is less than it used to be, but as production manager, you have made the best of a difficult situation and shown the sales department that the best course is to pull out of the highly competitive standard mix market completely. △

This example is taken from a field of mathematics variously known by names such as operations research, system science or management science. More specifically, this is an perfect example of problems that are solved by the techniques of "linear programming."

There is a lot going on under the hood in this example. The heart of the matter is the solution to simultaneous sytems of linear equations, which is the topic of the next few sections, and a recurrent theme throughout this course. We will return to this example on several occassions to reveal some of the reasons for its behavior.

Section SSSLE
Solving Systems of Simultaneous Linear Equations

We will motivate our study of linear algebra by considering the problem of solving several linear equations simultaneously. The word "solve" tends to get abused somewhat, as in "solve this problem." When talking about equations we understand a more precise meaning: find *all* of the values of some variable quantities that make an equation, or several equations, true.

Example SSSLE.STNE
Solving two (nonlinear) equations
Suppose we desire the simultaneous solutions of the two equations,

$$\begin{aligned} x^2 + y^2 &= 1 \\ -x + \sqrt{3}y &= 0. \end{aligned}$$

You can easily check by substitution that $x = \frac{\sqrt{3}}{2},\ y = \frac{1}{2}$ and $x = -\frac{\sqrt{3}}{2},\ y = -\frac{1}{2}$ are both solutions. We need to also convince ourselves that these are the *only* solutions. To see this, plot each equation on the xy-plane, which means to plot $(x,\ y)$ pairs that make an individual equation true. In this case we get a circle centered at the origin with radius 1 and a straight line through the origin with slope $\frac{1}{\sqrt{3}}$. The intersections of these two curves are our desired simultaneous solutions, and so we believe from our plot that the two solutions we know already are the only ones. We like to write solutions as sets, so in this case we write the set of solutions as

$$S = \{(\tfrac{\sqrt{3}}{2},\ \tfrac{1}{2}),\quad (-\tfrac{\sqrt{3}}{2},\ -\tfrac{1}{2})\}$$ △

In order to discuss systems of linear equations carefully, we need a precise definition. And before we do that, we will introduce our periodic discussions about "proof techniques." Linear algebra is an excellent setting for learning how to read, understand and formulate proofs. To help you in this process, we will digress, at irregular intervals, about some important aspect of working with proofs.

Proof Technique D
Definitions
A definition is a made-up term, used as a kind of shortcut for some typically more complicated idea. For example, we say a whole number is **even** as a shortcut for saying that when we divide the number by two we get a remainder of zero. With a precise definition, we can answer certain questions unambiguously. For example, did you ever wonder if zero was an even number? Now the answer should be clear since we have a precise definition of what we mean by the term even.

A single term might have several possible definitions. For example, we could say that the whole number n is even if there is another whole number k such that $n = 2k$. We say this is an equivalent definition since it categorizes even numbers the same way our first definition does.

Definitions are like two-way streets — we can use a definition to replace something rather complicated by its definition (if it fits) *and* we can replace a definition by its more complicated description. A definition is usually written as some form of an implication, such as "If something-nice-happens, then **blatzo**." However, this also means that "If blatzo, then something-nice-happens," even though this may not be formally stated. This is what we mean when we say a definition is a two-way street — it is really two implications, going in opposite "'directions."

Anybody (including you) can make up a definition, so long as it is unambiguous, but the real test of a definition's utility is whether or not it is useful for describing interesting or frequent situations.

We will talk about theorems later (and especially equivalences). For now, be sure not to confuse the notion of a definition with that of a theorem.

In this book, we will display every new definition carefully set-off from the text, and the term being defined will be written thus: **definition**. Additionally, there is a full list of all the definitions, in order of their appearance located at the front of the book (Definitions). Finally, the acroynm for each definition can be found in the index (Index). Definitions are critical to doing mathematics and proving theorems, so we've given you lots of ways to locate a definition should you forget its. . . uh, uh, well, . . . definition.

Can you formulate a precise definition for what it means for a number to be **odd**? (Don't just say it is the opposite of even. Act as if you don't have a definition for even yet.) Can you formulate your definition a second, equivalent, way? Can you employ your definition to test an odd and an even number for "odd-ness"? ♢

Definition SSLE
System of Simultaneous Linear Equations
A **system of simultaneous linear equations** is a collection of m equations in the variable quantities $x_1, x_2, x_3, \ldots, x_n$ of the form,

$$\begin{aligned}
a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \cdots + a_{1n}x_n &= b_1\\
a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \cdots + a_{2n}x_n &= b_2\\
a_{31}x_1 + a_{32}x_2 + a_{33}x_3 + \cdots + a_{3n}x_n &= b_3\\
&\vdots\\
a_{m1}x_1 + a_{m2}x_2 + a_{m3}x_3 + \cdots + a_{mn}x_n &= b_m
\end{aligned}$$

where the values of a_{ij}, b_i and x_j are from the set of complex numbers, $\mathbb{C}$. ⊙

Don't let the mention of the complex numbers, $\mathbb{C}$, rattle you. We will stick with real numbers exclusively for many more sections, and it will sometimes seem like we only work with integers! However, we want to leave the possibility of complex numbers open, and there will be ocassions in subsequent sections where they are necessary. For now, here is an example to illustrate using the notation introduced in Definition SSLE [8].

Example SSSLE.NSE
Notation for a system of equations
Given the system of simultaneous linear equations,

$$\begin{aligned} x_1 + 2x_2 + \quad x_4 &= 7 \\ x_1 + x_2 + x_3 - x_4 &= 3 \\ 3x_1 + x_2 + 5x_3 - 7x_4 &= 1 \end{aligned}$$

we have $n = 4$ variables and $m = 3$ equations. Also,

$$\begin{array}{llll} a_{11} = 1 & a_{12} = 2 & a_{13} = 0 & a_{14} = 1 \\ a_{21} = 1 & a_{22} = 1 & a_{23} = 1 & a_{24} = -1 \\ a_{31} = 3 & a_{32} = 1 & a_{33} = 5 & a_{34} = -7 \\ b_1 = 7 & b_2 = 3 & b_3 = 1. & \end{array}$$

Additionally, convince yourself that $x_1 = -2$, $x_2 = 4$, $x_3 = 2$, $x_4 = 1$ is one solution (but it is not the only one!). △

We will often shorten the term "system of simultaneous linear equations" to "system of linear equations" or just "system of equations" leaving the linear aspect implied.

Subsection PSS
Possibilities for solution sets

The next example illustrates the possibilities for the solution set of a system of linear equations. We will not be too formal here, and the necessary theorems to back up our claims will come in subsequent sections. So read for feeling and come back later to revisit this example.

Example SSSLE.TTS
Three typical systems
Consider the system of two equations with two variables,

$$\begin{aligned} 2x_1 + 3x_2 &= 3 \\ x_1 - x_2 &= 4. \end{aligned}$$

If we plot the solutions to each of these equations separately on the x_1x_2-plane, we get two lines, one with negative slope, the other with positive slope. They have exactly one point in common, $(x_1, x_2) = (3, -1)$, which is the solution $x_1 = 3$, $x_2 = -1$. From the geometry, we believe that this is the only solution to the system of equations, and so we say it is unique.

Now adjust the system with a different second equation,

$$\begin{aligned} 2x_1 + 3x_2 &= 3 \\ 4x_1 + 6x_2 &= 6. \end{aligned}$$

A plot of the solutions to these equations individually results in two lines, one on top of the other! There are infinitely many pairs of points that make both equations true. We will learn shortly how to describe this infinite solution set precisely. Notice now how the second equation is just a multiple of the first.

One more minor adjustment

$$\begin{aligned} 2x_1 + 3x_2 &= 3 \\ 4x_1 + 6x_2 &= 10. \end{aligned}$$

A plot now reveals two lines with identical slopes, i.e. parallel lines. They have no points in common, and so the system has a solution set that is empty, $S = \emptyset$. △

This example exhibits all of the typical behaviors of a system of equations. A subsequent theorem will tell us that every system of simultaneous linear equations has a solution set that is empty, contains a single solution or contains infinitley many solutions. Example SSSLE.STNE [7] yielded exactly two solutions, but this does not contradict the forthcoming theorem, since the equations are not linear and do not match the form of Definition SSLE [8].

Subsection ESEO
Equivalent systems and equation operations

With all this talk about finding solution sets for systems of linear equations, you might be ready to begin learning how to find these solution sets yourself. We begin with our first definition that takes a common word and gives it a very precise meaning in the context of systems of linear equations.

Definition ES
Equivalent Systems
Two systems of simultaneous linear equations are **equivalent** if their solution sets are equal. ⊙

Notice here that the two systems of equations could *look* very different (i.e. not be equal), but still have equal solution sets, and we would then call the systems equivalent. Two linear equations in two variables might be plotted as two lines that intersect in a single point. A different system, with three equations in two variables might have a plot that is three lines, all intersecting at a common point, with this common point identical to the intersection point for the first system. By our definition, we could then say these two very different looking systems of equations are equivalent, since they have identical solution sets. It is really like a weaker form of equality, where we allow the systems to be different in some respects, but use the term equivalent to highlight the situation when their solution sets are equal.

With this definition, we can begin to describe our strategy for solving linear systems. Given a system of linear equations that looks difficult to solve, we would like to have an *equivalent* system that is easy to solve. Since the systems will have equal solution sets, we can solve the "easy" system and get the solution set to the "difficult" system. Here come the tools for making this strategy viable.

Definition EO
Equation Operations
Given a system of simultaneous linear equations, the following three operations will transform the system into a different one, and each is known as an **equation operation**.

1. Swap the locations of two equations in the list.
2. Multiply each term of an equation by a nonzero quantity.
3. Multiply each term of one equation by some quantity, and add these terms to a second equation, on both sides of the equality. Leave the first equation the same after this operation, but replace the second equation by the new one. ⊙

These descriptions might seem a bit vague, but the proof or the examples that follow should make it clear what is meant by each.

Proof Technique T
Theorems
Higher mathematics is about understanding theorems. Reading them, understanding them, applying them, proving them. We are ready to prove our first momentarily. Every theorem is a shortcut — we prove something in general, and then whenever we find a specific instance covered by the theorem we can immediately say that we know something else about the situation by applying the theorem. In many cases, this new information can be gained with much less effort than if we did not know the theorem.

The first step in understanding a theorem is to realize that the statement of every theorem can be rewrritten using statements of the form "If something-happens, then something-else-happens." The "something-happens" part is the **hypothesis** and the "something-else-happens" is the **conclusion**. To understand a theorem, it helps to rewrite its statement using this construction. To apply a theorem, we verify that "something-happens" in a particular instance and immediately conclude that "something-else-happens." To prove a theorem, we must argue based on the assumption that the hypothesis is true, and arrive through the process of logic that the conclusion must then also be true. ◊

Theorem EOPSS
Equation Operations Preserve Solution Sets
Suppose we apply one of the three equation operations to the system of simultaneous linear equations

$$
\begin{aligned}
a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \cdots + a_{1n}x_n &= b_1 \\
a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \cdots + a_{2n}x_n &= b_2 \\
a_{31}x_1 + a_{32}x_2 + a_{33}x_3 + \cdots + a_{3n}x_n &= b_3 \\
&\vdots \\
a_{m1}x_1 + a_{m2}x_2 + a_{m3}x_3 + \cdots + a_{mn}x_n &= b_m.
\end{aligned}
$$

Then the original system and the transformed system are equivalent systems. □

Proof Before we begin the proof, we make two comments about proof techniques.

Proof Technique GS
Getting Started
"I don't know how to get started!" is often the lament of the novice proof-builder. Here are a few pieces of advice.

1. As mentioned in Technique T [11], rewrite the statement of the theorem in an "if-then" form. This will simplify identifying the hypothesis and conclusion, which are referenced in the next few items.

2. Ask yourself what it is you are trying to prove. This is always part of your conclusion. Are you being asked to conclude that two numbers are equal, that a function is differentiable or a set is a subset of another? You cannot bring other techniques to bear if you do not know what *type* of conclusion you have.

3. Write down reformulations of your hypotheses. Interpret and translate each definition properly.

4. Write your hypothesis at the top of a sheet of paper and your conclusion at the bottom. See if you can formulate a statement that precedes the conclusion and also implies it. Work down from your hypothesis, and up from your conclusion, and see if you can meet in the middle. When you are finished, rewrite the proof nicely, from hypothesis to conclusion, with verifiable implications giving each subsequent statement. ◊

Proof Technique SE
Set Equality
In the theorem we are trying to prove, the conclusion is that two systems are equivalent. By Definition ES [10] this translates to requiring that solution sets be equal for the

two systems. So we are being asked to show *that two sets are equal.* How do we do this? Well, there is a very standard technique, and we will use it repeatedly through the course. So lets add it to our toolbox now.

A **set** is just a collection of items, which we refer to generically as **elements**. If A is a set, and a is one of its elements, we write that piece of information as $a \in A$. Similarly, if b is not in A, we write $b \notin A$. Given two sets, A and B, we say that A is a **subset** of B if all the elements of A are also in B. More formally (and much easier to work with) we describe this situation as follows: A is a subset of B if whenever $x \in A$, then $x \in B$. Notice the use of the "if-then" construction here. The notation for this is $A \subseteq B$. (If we want to diallow the possibility that A is the same as B, we use $A \subset B$.)

But what does it mean for two sets to be **equal**? They must be the same. Well, that explanation is not really too helpful, is it? How about: If $A \subseteq B$ and $B \subseteq A$, then A equals B. This gives us something to work with, if A is a subset of B, and *vice versa*, then they must really be the same set. We will now make the symbol "=" do double-duty and extend its use to statements like $A = B$, where A and B are sets. ♢

Now we can take each equation operation in turn and show that the solution sets of the two systems are equal, using the technique just outlined.

1. I will not be our habit in proofs to resort to saying statements are "obvious," but in this case, it should be. There is nothing about the *order* in which we write linear equations that affects their solutions, so the solution set will be equal if the systems only differ by a rearrangement of the order of the equations.

2. Suppose $\alpha \neq 0$ is a number. Let's choose to multiply the terms of equation i by α to build the new system of equations,

$$\begin{aligned}
a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \cdots + a_{1n}x_n &= b_1\\
a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \cdots + a_{2n}x_n &= b_2\\
a_{31}x_1 + a_{32}x_2 + a_{33}x_3 + \cdots + a_{3n}x_n &= b_3\\
&\vdots\\
\alpha a_{i1}x_1 + \alpha a_{i2}x_2 + \alpha a_{i3}x_3 + \cdots + \alpha a_{in}x_n &= \alpha b_i\\
&\vdots\\
a_{m1}x_1 + a_{m2}x_2 + a_{m3}x_3 + \cdots + a_{mn}x_n &= b_m.
\end{aligned}$$

Let S denote the solutions to the system in the statement of the theorem, and let T denote the solutions to the transformed system.

(a) Show $S \subseteq T$. Suppose $(x_1,\, x_2,\, x_3,\, \ldots, x_n) = (\beta_1,\, \beta_2,\, \beta_3,\, \ldots, \beta_n) \in S$ is a solution to the original system. Ignoring the i-th equation for a moment, we know it makes every other equation of the transformed system true. We also know that

$$a_{i1}\beta_1 + a_{i2}\beta_2 + a_{i3}\beta_3 + \cdots + a_{in}\beta_n = b_i$$

which we can multiply by α to get

$$\alpha a_{i1}\beta_1 + \alpha a_{i2}\beta_2 + \alpha a_{i3}\beta_3 + \cdots + \alpha a_{in}\beta_n = \alpha b_i.$$

This says that the i-th equation of the transformed system is also true, so we have established that $(\beta_1,\, \beta_2,\, \beta_3,\, \ldots, \beta_n) \in T$, and therefore $S \subseteq T$.

(b) Now show $T \subseteq S$. Suppose $(x_1,\, x_2,\, x_3,\, \ldots, x_n) = (\beta_1,\, \beta_2,\, \beta_3,\, \ldots, \beta_n) \in S$ is a solution to the transformed system. Ignoring the i-th equation for a moment, we know it makes every other equation of the original system true. We also know that

$$\alpha a_{i1}\beta_1 + \alpha a_{i2}\beta_2 + \alpha a_{i3}\beta_3 + \cdots + \alpha a_{in}\beta_n = \alpha b_i$$

which we can multiply by $\frac{1}{\alpha}$, since $\alpha \neq 0$, to get

$$a_{i1}\beta_1 + a_{i2}\beta_2 + a_{i3}\beta_3 + \cdots + a_{in}\beta_n = b_i$$

This says that the i-th equation of the original system is also true, so we have established that $(\beta_1,\, \beta_2,\, \beta_3,\, \ldots, \beta_n) \in S$, and therefore $T \subseteq S$. Locate the key point where we required that $\alpha \neq 0$, and consider what would happen if $\alpha = 0$.

3. Suppose α is a number. Let's choose to multiply the terms of equation i by α and add them to equation j in order to build the new system of equations,

$$\begin{aligned}
&a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1\\
&a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2\\
&a_{31}x_1 + a_{32}x_2 + \cdots + a_{3n}x_n = b_3\\
&\vdots\\
&(\alpha a_{i1} + a_{j1})x_1 + (\alpha a_{i2} + a_{j2})x_2 + \cdots + (\alpha a_{in} + a_{jn})x_n = \alpha b_i + b_j\\
&\vdots\\
&a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m.
\end{aligned}$$

Let S denote the solutions to the system in the statement of the theorem, and let T denote the solutions to the transformed system.

(a) Show $S \subseteq T$. Suppose $(x_1,\, x_2,\, x_3,\, \ldots, x_n) = (\beta_1,\, \beta_2,\, \beta_3,\, \ldots, \beta_n) \in S$ is a solution to the original system. Ignoring the j-th equation for a moment, we know it makes every other equation of the transformed system true. Using the fact that it makes the i-th and j-th equations true, we find

$$\begin{aligned}
&(\alpha a_{i1} + a_{j1})\beta_1 + (\alpha a_{i2} + a_{j2})\beta_2 + \cdots + (\alpha a_{in} + a_{jn})\beta_n =\\
&(\alpha a_{i1}\beta_1 + \alpha a_{i2}\beta_2 + \cdots + \alpha a_{in}\beta_n) + (a_{j1}\beta_1 + a_{j2}\beta_2 + \cdots + a_{jn}\beta_n) =\\
&\alpha(a_{i1}\beta_1 + a_{i2}\beta_2 + \cdots + a_{in}\beta_n) + (a_{j1}\beta_1 + a_{j2}\beta_2 + \cdots + a_{jn}\beta_n) = \alpha b_i + b_j.
\end{aligned}$$

This says that the j-th equation of the transformed system is also true, so we have established that $(\beta_1,\, \beta_2,\, \beta_3,\, \ldots, \beta_n) \in T$, and therefore $S \subseteq T$.

(b) Now show $T \subseteq S$. Suppose $(x_1,\,x_2,\,x_3,\,\ldots,x_n) = (\beta_1,\,\beta_2,\,\beta_3,\,\ldots,\beta_n) \in S$ is a solution to the transformed system. Ignoring the j-th equation for a moment, we know it makes every other equation of the original system true. We then find

$$\begin{aligned}
&a_{j1}\beta_1 + a_{j2}\beta_2 + \cdots + a_{jn}\beta_n = \\
&a_{j1}\beta_1 + a_{j2}\beta_2 + \cdots + a_{jn}\beta_n + \alpha b_i - \alpha b_i = \\
&a_{j1}\beta_1 + a_{j2}\beta_2 + \cdots + a_{jn}\beta_n + (\alpha a_{i1}\beta_1 + \alpha a_{i2}\beta_2 + \cdots + \alpha a_{in}\beta_n) - \alpha b_i = \\
&a_{j1}\beta_1 + a_{j2}\beta_2 + \cdots + a_{jn}\beta_n + (\alpha a_{i1}\beta_1 + \alpha a_{i2}\beta_2 + \cdots + \alpha a_{in}\beta_n) - \alpha b_i = \\
&(\alpha a_{i1} + a_{j1})\beta_1 + (\alpha a_{i2} + a_{j2})\beta_2 + \cdots + (\alpha a_{in} + a_{jn})\beta_n - \alpha b_i = \\
&\alpha b_i + b_j - \alpha b_i = b_j
\end{aligned}$$

This says that the j-th equation of the original system is also true, so we have established that $(\beta_1,\,\beta_2,\,\beta_3,\,\ldots,\beta_n) \in S$, and therefore $T \subseteq S$.

Why didn't we need to require that $\alpha \neq 0$ for this row operation? In other words, how does the third statement of the theorem read when $\alpha = 0$? Does our proof require some extra care when $\alpha = 0$? Compare your answers with the similar situation for the second row operation. ■

Theorem EOPSS [12] is the necessary tool to complete our strategy for solving systems of equations. We will use equation operations to move from one system to another, all the while keeping the solution set the same. With the right sequence of operations, we will arrive at a simpler equation to solve. The next two examples illustrate this idea, while saving some of the details for later.

Example SSSLE.US
Three equations, one solution
We solve the following system by a sequence of equation operations.

$$\begin{aligned}
x_1 + 2x_2 + 2x_3 &= 4\\
x_1 + 3x_2 + 3x_3 &= 5\\
2x_1 + 6x_2 + 5x_3 &= 6
\end{aligned}$$

$\alpha = -1$ times equation 1, add to equation 2:

$$\begin{aligned}
x_1 + 2x_2 + 2x_3 &= 4\\
0x_1 + 1x_2 + 1x_3 &= 1\\
2x_1 + 6x_2 + 5x_3 &= 6
\end{aligned}$$

$\alpha = -2$ times equation 1, add to equation 3:

$$\begin{aligned}
x_1 + 2x_2 + 2x_3 &= 4\\
0x_1 + 1x_2 + 1x_3 &= 1\\
0x_1 + 2x_2 + 1x_3 &= -2
\end{aligned}$$

$\alpha = -2$ times equation 2, add to equation 3:

$$\begin{aligned} x_1 + 2x_2 + 2x_3 &= 4 \\ 0x_1 + 1x_2 + 1x_3 &= 1 \\ 0x_1 + 0x_2 - 1x_3 &= -4 \end{aligned}$$

$\alpha = -1$ times equation 3:

$$\begin{aligned} x_1 + 2x_2 + 2x_3 &= 4 \\ 0x_1 + 1x_2 + 1x_3 &= 1 \\ 0x_1 + 0x_2 + 1x_3 &= 4 \end{aligned}$$

which can be written more clearly as

$$\begin{aligned} x_1 + 2x_2 + 2x_3 &= 4 \\ x_2 + x_3 &= 1 \\ x_3 &= 4 \end{aligned}$$

This is now a very easy system of equations to solve. The third equation requires that $x_3 = 4$ to be true. Making this substitution into equation 2 we arrive at $x_2 = -3$, and finally, substituting these values of x_2 and x_3 into the first equation, we find that $x_1 = 2$. Note too that this is the only solution to this final system of equations, since we were forced to choose these values to make the equations true. Since we performed equation operations on each system to obtain the next one in the list, all of the systems listed here are all equivalent to each other by Theorem EOPSS [12]. Thus $(x_1,\, x_2,\, x_3) = (2, -3, 4)$ is the unique solution to the *orginal* system of equations (and all of the other systems of equations). △

Example SSSLE.IS
Three equations, infinitely many solutions
The following system of equations made an appearance earlier in this section (Example SSSLE.NSE [9]), where we listed *one* of its solutions. Now, we will try to find all of the solutions to this system.

$$\begin{aligned} x_1 + 2x_2 + 0x_3 + x_4 &= 7 \\ x_1 + x_2 + x_3 - x_4 &= 3 \\ 3x_1 + x_2 + 5x_3 - 7x_4 &= 1 \end{aligned}$$

$\alpha = -1$ times equation 1, add to equation 2:

$$\begin{aligned} x_1 + 2x_2 + 0x_3 + x_4 &= 7 \\ x_1 - x_2 + x_3 - 2x_4 &= -4 \\ 3x_1 + x_2 + 5x_3 - 7x_4 &= 1 \end{aligned}$$

$\alpha = -3$ times equation 1, add to equation 3:

$$\begin{aligned} x_1 + 2x_2 + 0x_3 + x_4 &= 7 \\ 0x_1 - x_2 + x_3 - 2x_4 &= -4 \\ 0x_1 - 5x_2 + 5x_3 - 10x_4 &= -20 \end{aligned}$$

$\alpha = -5$ times equation 2, add to equation 3:

$$\begin{aligned} x_1 + 2x_2 + 0x_3 + x_4 &= 7 \\ 0x_1 - x_2 + x_3 - 2x_4 &= -4 \\ 0x_1 + 0x_2 + 0x_3 + 0x_4 &= 0 \end{aligned}$$

$\alpha = -1$ times equation 2:

$$\begin{aligned} x_1 + 2x_2 + 0x_3 + x_4 &= 7 \\ 0x_1 + x_2 - x_3 + 2x_4 &= 4 \\ 0x_1 + 0x_2 + 0x_3 + 0x_4 &= 0 \end{aligned}$$

which can be written more clearly as

$$\begin{aligned} x_1 + 2x_2 + x_4 &= 7 \\ x_2 - x_3 + 2x_4 &= 4 \\ 0 &= 0 \end{aligned}$$

What does the equation $0 = 0$ mean? We can choose *any* values for x_1, x_2, x_3, x_4 and this equation will be true, so we just need to only consider further the first two equations since the third is true no matter what. We can analyze the second equation without consideration of the variable x_1. It would appear that there is considerable latitude in how we can choose x_2, x_3, x_4 and make this equation true. Lets choose x_3 and x_4 to be *anything* we please, say $x_3 = \beta_3$ and $x_4 = \beta_4$. Then equation 2 becomes

$$x_2 - \beta_3 + 2\beta_4 = 4 \qquad \Rightarrow \qquad x_2 = 4 + \beta_3 - 2\beta_4$$

Now we can take these arbitrary values for x_3 and x_4, and this expression for x_2 and employ them in equation 1,

$$x_1 + 2(4 + \beta_3 - 2\beta_4) + \beta_4 = 7 \qquad \Rightarrow \qquad x_1 = -1 - 2\beta_3 + 3\beta_4$$

So our arbitrary choices of values for x_3 and x_4 (β_3 and β_4) translate into specific values of x_1 and x_2. The lone solution given in Example SSSLE.NSE [9] was obtained by choosing $\beta_3 = 2$ and $\beta_4 = 1$. Now we can easily and quickly find many more (infinitely more). Suppose we choose $\beta_3 = 5$ and $\beta_4 = -2$, then we compute

$$\begin{aligned} x_1 &= -1 - 2(5) + 3(-2) = -17 \\ x_2 &= 4 + 5 - 2(-2) = 13 \end{aligned}$$

and you can verify that $(x_1,\, x_2,\, x_3,\, x_4) = (-17,\, 13,\, 5,\, -2)$ makes all three equations true. The entire solution set is written as

$$S = \{(-1 - 2\beta_3 + 3\beta_4,\, 4 + \beta_3 - 2\beta_4,\, \beta_3,\, \beta_4) \mid \beta_3 \in \mathbb{C},\, \beta_4 \in \mathbb{C}\}$$

It would be instructive to finish off your study of this example by taking the general form of the solutions given in this set and substituting them into each of the three equations and verify that they are true in each case. △

In the next section we will describe how to use equation operations to systematically solve any system of simultaneous linear equations. But first, one of our more important pieces of advice about doing mathematics.

Proof Technique L
Language
Mathematics is a language. It is a way to express complicated ideas clearly, precisely, and unambiguously. Because of this, it can be difficult to read. Read slowly, and have pencil and paper at hand. It will usually be necessary to read something several times. While reading can be difficult, its even hard to speak mathematics, and so that is the topic of this technique.

I am going to suggest a simple modification to the way you use language that will make it much, much easier to become proficient at speaking mathematics and eventually it will become second nature. Think of it as a training aid or practice drill you might use when learning to become skilled at a sport.

First, eliminate pronouns from your vocabulary when discussing linear algebra, in class or with your colleagues. Do not use: it, that, those, their or similar sources of confusion. This is the single easiest step you can take to make your oral expression of mathematics clearer to others, and in turn, it will greatly help your own understanding.

Now rid yourself of the word "thing" (or variants like "something"). When you are tempted to use this word realize that there is some object you want to discuss, and we likely have a definition for that object (see the discussion at Technique D [7]). Always "think about your objects" and many aspects of the study of mathematics will get easier. Ask yourself: "Am I working with a set, a number, a function, an operation, or what?" Knowing what an object *is* will allow you to narrow down the procedures you may apply to **it**. If you have studied an object-oriented computer programming language, then perhaps this advice will be even clearer, since you know that a compiler will often complain with an error message if you confuse your objects.

Third, eliminate the verb "works" (as in "the equation works") from your vocabulary. This term is used as a substitute when we are not sure just what we are trying to accomplish. Usually we are trying to say that some object fulfills some condition. The condition might even have a definition associated with it, making it even easier to describe.

Last, speak slooooowly and thoughtfully as you try to get by without all these lazy words. It is hard at first, but you will get better with practice. Especially in class, when the pressure is on and all eyes are on you, don't succumb to the temptation to use these weak words. Slow down, we'd all rather wait for a slow, well-formed question or answer than a fast, sloppy, incomprehensible one.

When studying with friends, you might make a game of catching one another using pronouns, "thing," or "works." I know I'll be calling you on it! $\lozenge$

Section RREF
Reduced Row-Echelon Form

After solving a few systems of equations, you will recognize that it doesn't matter so much *what* we call our variables, as opposed to what numbers act as their coefficients. A system in the variables x_1, x_2, x_3 would behave the same if we changed the names of the variables to a, b, c and kept all the constants the same and in the same places. In this section, we will isolate the key bits of information about a system of equations into something called a matrix, and then use this matrix to systematically solve the equations. Along the way we will obtain one of our most important and useful computational tools.

Definition M
Matrix
An $m \times n$ **matrix** is a rectangular layout of numbers from $\mathbb{C}$ having m rows and n columns. $\odot$

Notation MN
Matrix Notation
We will use upper-case Latin letters from the start of the alphabet (A, B, $C, \ldots$) to denote matrices and squared-off brackets to delimit the layout. Many use large parentheses instead of brackets — the distinction is not important. $\triangledown$

Example RREF.AM
A matrix

$$B = \begin{bmatrix} -1 & 2 & 5 & 3 \\ 1 & 0 & -6 & 1 \\ -4 & 2 & 2 & -2 \end{bmatrix}$$

is a matrix with $m = 3$ rows and $n = 4$ columns. $\triangle$

A calculator or computer language can be a convenient way to perform calculations with matrices. But first you have to enter the matrix. Here's how it is done on various computing platforms.

Definition AM
Augmented Matrix

Suppose we have a system of m equations in the n variables $x_1, x_2, x_3, \ldots, x_n$ written as

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \cdots + a_{2n}x_n &= b_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 + \cdots + a_{3n}x_n &= b_3 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + a_{m3}x_3 + \cdots + a_{mn}x_n &= b_m \end{aligned}$$

then the **augmented matrix** of the system of equations is the $m \times (n+1)$ matrix

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & \ldots & a_{1n} & b_1 \\ a_{21} & a_{22} & a_{23} & \ldots & a_{2n} & b_2 \\ a_{31} & a_{32} & a_{33} & \ldots & a_{3n} & b_3 \\ \vdots & & & & & \\ a_{m1} & a_{m2} & a_{m3} & \ldots & a_{mn} & b_m \end{bmatrix}$$ ⊙

The augmented matrix *represents* all the important information in the system of equations, since the names of the variables have been ignored, and the only connection with the variables is the location of their coefficients in the matrix. It is important to realize that the augmented matrix is just that, a matrix, and *not* a system of equations. In particular, the augmented matrix does not have any "solutions," though it will be useful for finding solutions to the system of equations that it is associated with. Notice too that an augmented matrix always belongs to some system of equations, and vice versa. Here's a quick example.

Example RREF.AMAA
Augmented matrix for Archetype A
Archetype A is the following system of 3 equations in 3 variables.

$$\begin{aligned} x_1 - x_2 + 2x_3 &= 1 \\ 2x_1 + x_2 + x_3 &= 8 \\ x_1 + x_2 &= 5 \end{aligned}$$

Here is its augmented matrix.

$$\begin{bmatrix} 1 & -1 & 2 & 1 \\ 2 & 1 & 1 & 8 \\ 1 & 1 & 0 & 5 \end{bmatrix}$$ △

An augmented matrix for a system of equations will save us the tedium of continually writing down the names of the variables as we solve the system. It will also release us from any dependence on the actual names of the variables. We have seen how certain operations we can perform on equations (Definition EO [11]) will preserve their solutions (Theorem EOPSS [12]). The next two definitions and the following theorem carry over these ideas to augmented matrices.

Definition RO
Row Operations
The following three operations will transform an $m \times n$ matrix into a different matrix of the same size, and each is known as an **row operation**.

1. Swap the locations of two rows.
2. Multiply each entry of a single row by a nonzero quantity.
3. Multiply each entry of one row by some quantity, and add these values to the entry in the same column of a second row. Leave the first row the same after this operation, but replace the second row by the new values. ⊙

We will use a kind of shorthand to describe these operations:

1. $R_i \leftrightarrow R_j$: Swap the location of rows i and j.
2. αR_j: Multiply row i by the nonzero scalar α.
3. $\alpha R_i + R_j$: Multiply row i by the scalar α and add to row j.

Definition REM
Row-Equivalent Matrices
Two matrices, A and B, are **row-equivalent** if one can be obtained from the other by a sequence of row operations. ⊙

Example RREF.TREM
Two row-equivalent matrices
The matrices

$$A = \begin{bmatrix} 2 & -1 & 3 & 4 \\ 5 & 2 & -2 & 3 \\ 1 & 1 & 0 & 6 \end{bmatrix} \qquad B = \begin{bmatrix} 1 & 1 & 0 & 6 \\ 3 & 0 & -2 & -9 \\ 2 & -1 & 3 & 4 \end{bmatrix}$$

are row-equivalent as can be seen from

$$\begin{bmatrix} 2 & -1 & 3 & 4 \\ 5 & 2 & -2 & 3 \\ 1 & 1 & 0 & 6 \end{bmatrix} \xrightarrow{R_1 \leftrightarrow R_3} \begin{bmatrix} 1 & 1 & 0 & 6 \\ 5 & 2 & -2 & 3 \\ 2 & -1 & 3 & 4 \end{bmatrix} \xrightarrow{-2R_1 + R_2} \begin{bmatrix} 1 & 1 & 0 & 6 \\ 3 & 0 & -2 & -9 \\ 2 & -1 & 3 & 4 \end{bmatrix}$$

We can also say that any pair of these three matrices are row-equivalent. △

Notice that each of the three row operations is reversible, so we do not have to be careful about the distinction between "A is row-equivalent to B" and "B is row-equivalent to A." The preceding definitions are designed to make the following theorem possible. It says that row-equivalent matrices represent systems of linear equations that have identical solution sets.

Theorem REMES
Row-Equivalent Matrices represent Equivalent Systems
Suppose that A and B are row-equivalent augmented matrices. Then the systems of linear equations that they represent are equivalent systems. □

Proof If we perform a single row operation on an augmented matrix, it will have the same effect as if we did the analagous equation operation on the corresponding system of equations. By exactly the same methods as we used in the proof of Theorem EOPSS [12] we can see that each of these row operations will preserve the set of solutions for the corresponding system of equations. ■

So at this point, our strategy is to begin with a system of equations, represent it by an augmented matrix, perform row operations (which will preserve solutions for the corresponding systems) to get a "simpler" augmented matrix, convert back to a "simpler" system of equations and then solve that system, knowing that its solutions are those of the original system. Here's a rehash of Example SSSLE.US [15] as an exercise in using our new tools.

Example RREF.US
Three equations, one solution
We solve the following system using augmented matrices and row operations.

$$\begin{aligned} x_1 + 2x_2 + 2x_3 &= 4 \\ x_1 + 3x_2 + 3x_3 &= 5 \\ 2x_1 + 6x_2 + 5x_3 &= 6 \end{aligned}$$

Form the augmented matrix,

$$A = \begin{bmatrix} 1 & 2 & 2 & 4 \\ 1 & 3 & 3 & 5 \\ 2 & 6 & 5 & 6 \end{bmatrix}$$

and apply row operations,

$$\xrightarrow{-1R_1+R_2} \begin{bmatrix} 1 & 2 & 2 & 4 \\ 0 & 1 & 1 & 1 \\ 2 & 6 & 5 & 6 \end{bmatrix}$$

$$\xrightarrow{-2R_1+R_3} \begin{bmatrix} 1 & 2 & 2 & 4 \\ 0 & 1 & 1 & 1 \\ 0 & 2 & 1 & -2 \end{bmatrix}$$

$$\xrightarrow{-2R_2+R_3} \begin{bmatrix} 1 & 2 & 2 & 4 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & -1 & -4 \end{bmatrix}$$

$$\xrightarrow{-1R_3} \begin{bmatrix} 1 & 2 & 2 & 4 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 4 \end{bmatrix}$$

So the matrix

$$B = \begin{bmatrix} 1 & 2 & 2 & 4 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 4 \end{bmatrix}$$

is row equivalent to A and by Theorem REMES [23] the system of equations below has the same solution set as the original system of equations.

$$\begin{aligned} x_1 + 2x_2 + 2x_3 &= 4 \\ x_2 + x_3 &= 1 \\ x_3 &= 4 \end{aligned}$$

Solving this "simpler" system is straightforward and is identical to the process in Example SSSLE.US [15]. △

The preceding example amply illustrates the definitions and theorems we have seen so far. But it still leaves two questions unanswered. Exactly what is this "simpler" form for a matrix, and just how do we get it? Here's the answer to the first question, a definition of reduced row-echelon form.

Definition RREF
Reduced Row-Echelon Form
A matrix is in **reduced row-echelon form** if it meets all of the following conditions:

1. A row where every entry is zero is below any row containing a nonzero entry.
2. The leftmost nonzero entry of a row is equal to 1.
3. The leftmost nonzero entry of a row is the only nonzero entry in its column.
4. Suppose rows i and s are any two rows that both have nonzero entries, where $s > i$. If the leftmost nonzero entry of row i is in column j, and the leftmost nonzero entry of row s is in column t, then $t > j$. ⊙

Because we will make frequent reference to reduced row-echelon form, we make precise definitions of two terms.

Definition ZRM
Zero Row of a Matrix
A row of a matrix where every entry is zero is called a **zero row**. ⊙

Definition LO
Leading Ones
For a matrix in reduced row-echelon form, the leftmost nonzero entry of any row that is not a zero row will be called a **leading 1**. ⊙

Example RREF.RREF
A matrix in reduced row-echelon form
The matrix C is in reduced row-echelon form.

$$\begin{bmatrix} 1 & -3 & 0 & 6 & 0 & 0 & -5 & 9 \\ 0 & 0 & 0 & 0 & 1 & 0 & 3 & -7 \\ 0 & 0 & 0 & 0 & 0 & 1 & 7 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

△

Example RREF.NRREF
A matrix not in reduced row-echelon form
The matrix D is not in reduced row-echelon form, as it fails each of the four requirements once.

$$\begin{bmatrix} 1 & 0 & -3 & 0 & 6 & 0 & 7 & -5 & 9 \\ 0 & 0 & 0 & 5 & 0 & 1 & 0 & 3 & -7 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 7 & 3 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & -4 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

△

Proof Technique C
Constructive Proofs
Conclusions of proofs come in a variety of types. Often a theorem will simply *assert* that something exists. The best way, but not the only way, to show something exists is to actually build it. Such a proof is called **constructive**. The thing to realize about constructive proofs is that the proof itself will contain a procedure that might be used computationally to construct the desired object. If the procedure is not too cumbersome, then the proof itself is as useful as the statement of the theorem. Such is the case with our next theorem. ◊

Theorem REMEF
Row-Equivalent Matrix in Echelon Form
Suppose A is a matrix. Then there is a (unique!) matrix B so that

1. A and B are row-equivalent.
2. B is in reduced row-echelon form. □

Proof Suppose that A has m rows. We will describe a process for converting A into B via row operations.

Set $i = 1$.

1. If $i = m + 1$, then stop converting the matrix.

2. Among all of the entries in rows i through m locate the leftmost nonzero entry (there may be several choices at this step). Denote the column of this entry by j. If this is not possible, then stop converting the matrix.

3. Swap the row containing the chosen leftmost nonzero entry with row i.

4. Use the second row operation to multiply row i by the reciprocal of the value in column j, thereby creating a leading 1 in row i at column j.

5. Use row i and the third row operation to convert every other entry in column j into a zero.

6. Increase i by one and return to step 1.

The result of this procedure is the matrix B. We need to establish that it has the requisite properties. First, the steps of the process only use row operations to convert the matrix, so A and B are row-equivalent.

Its a bit more work to be certain that B is in reduced row-echelon form. At the conclusion of the i-th trip through the steps, we claim the first i rows form a matrix in reduced row-echelon form, and the entries in rows $i + 1$ through m in columns 1 through j are all zero. To see this, notice that

1. The definition of j insures that the entries of rows $i + 1$ through m, in columns 1 through $j - 1$ are all zero.

2. Row i has a leading nonzero entry equal to 1 by the result of step 4.

3. The employment of the leading 1 of row i in step 5 will make every element of column j zero in rows 1 through $i + 1$, as well as in rows $i + 1$ through m.

4. Rows 1 through $i - 1$ are only affected by step 5. The zeros in columns 1 through $j - 1$ of row i mean that none of the entries in columns 1 through $j - 1$ for rows 1 through $i - 1$ will change by the row operations employed in step 5.

5. Since columns 1 through j are are all zero for rows $i + 1$ through m, any nonzero entry found on the next pass will be in a column to the right of column j, ensuring that the fourth condition of reduced row-echelon form is met.

6. If the procedure halts with $i = m + 1$, then every row of B has a leading 1, and hence has no zero rows. If the procedure halts because step 2 fails to find a nonzero entry, then rows i through m are all zero rows, and they are all at the bottom of the matrix. ■

So now we can put it all together. Begin with a system of linear equations (Definition SSLE [8]), and represent it by its augmented matrix (Definition AM [20]). Use row operations (Definition RO [22]) to convert this matrix into reduced row-echelon form (Definition RREF [24]), using the procedure outlined in the proof of Theorem REMEF [25]. Theorem REMEF [25] also tells us we can always accomplish this, and that the result is row-equivalent (Definition REM [22]) to the original augmented matrix. Since the matrix in reduced-row echelon form has the same solution set, we can analyze it instead of the original matrix, viewing it as the augmented matrix of a different system of equations. The beauty of augmented matrices in reduced row-echelon form is that the solution sets to their corresponding systems can be easily determined, as we will see in the next few examples and in the next section.

We will now run through some examples of using these definitions and theorems to solve some systems of equations. From now on, when we have a matrix in reduced row-echelon form, we will mark the leading 1's with a small box. In your work, you can box them, circle them or write them in a different color. This device will prove very useful later and is a very good habit to develop now.

Example RREF.SAB
Solutions for Archetype B
Solve the following system of equations.

$$\begin{aligned} -7x_1 - 6x_2 - 12x_3 &= -33 \\ 5x_1 + 5x_2 + 7x_3 &= 24 \\ x_1 + 4x_3 &= 5 \end{aligned}$$

Form the augmented matrix for starters,

$$\begin{bmatrix} -7 & -6 & -12 & -33 \\ 5 & 5 & 7 & 24 \\ 1 & 0 & 4 & 5 \end{bmatrix}$$

and work to reduced row-echelon form, first with $i = 1$,

$$\xrightarrow{R_1 \leftrightarrow R_3} \begin{bmatrix} 1 & 0 & 4 & 5 \\ 5 & 5 & 7 & 24 \\ -7 & -6 & -12 & -33 \end{bmatrix} \quad \xrightarrow{-5R_1+R_2} \begin{bmatrix} 1 & 0 & 4 & 5 \\ 0 & 5 & -13 & -1 \\ -7 & -6 & -12 & -33 \end{bmatrix}$$

$$\xrightarrow{7R_1+R_3} \begin{bmatrix} \boxed{1} & 0 & 4 & 5 \\ 0 & 5 & -13 & -1 \\ 0 & -6 & 16 & 2 \end{bmatrix}$$

Now, with $i = 2$,

$$\xrightarrow{\frac{1}{5}R_2} \begin{bmatrix} \boxed{1} & 0 & 4 & 5 \\ 0 & 1 & \frac{-13}{5} & \frac{-1}{5} \\ 0 & -6 & 16 & 2 \end{bmatrix} \quad \xrightarrow{6R_2+R_3} \begin{bmatrix} \boxed{1} & 0 & 4 & 5 \\ 0 & \boxed{1} & \frac{-13}{5} & \frac{-1}{5} \\ 0 & 0 & \frac{2}{5} & \frac{4}{5} \end{bmatrix}$$

And finally, with $i = 3$,

$$\xrightarrow{\frac{5}{2}R_3} \begin{bmatrix} \boxed{1} & 0 & 4 & 5 \\ 0 & \boxed{1} & \frac{-13}{5} & \frac{-1}{5} \\ 0 & 0 & 1 & 2 \end{bmatrix} \quad \xrightarrow{\frac{13}{5}R_3+R_2} \begin{bmatrix} \boxed{1} & 0 & 4 & 5 \\ 0 & \boxed{1} & 0 & 5 \\ 0 & 0 & 1 & 2 \end{bmatrix} \quad \xrightarrow{-4R_3+R_1} \begin{bmatrix} \boxed{1} & 0 & 0 & -3 \\ 0 & \boxed{1} & 0 & 5 \\ 0 & 0 & \boxed{1} & 2 \end{bmatrix}$$

This is now the augmented matrix of a very simple system of equations, namely $x_1 = -3$, $x_2 = 5$, $x_3 = 2$, which has an obvious solution. Furthermore, we can see that this is the *only* solution to this system, so we have determined the entire solution set. You might compare this example with the procedure we used in Example SSSLE.US [15]. △

Archetypes A and B are meant to contrast each other in many respects. So lets solve Archetype A now.

Example RREF.SAA
Solutions for Archetype A
Solve the following system of equations.

$$\begin{aligned} x_1 - x_2 + 2x_3 &= 1 \\ 2x_1 + x_2 + x_3 &= 8 \\ x_1 + x_2 &= 5 \end{aligned}$$

Form the augmented matrix for starters,

$$\begin{bmatrix} 1 & -1 & 2 & 1 \\ 2 & 1 & 1 & 8 \\ 1 & 1 & 0 & 5 \end{bmatrix}$$

and work to reduced row-echelon form, first with $i = 1$,

$$\xrightarrow{-2R_1+R_2} \begin{bmatrix} 1 & -1 & 2 & 1 \\ 0 & 3 & -3 & 6 \\ 1 & 1 & 0 & 5 \end{bmatrix} \xrightarrow{-1R_1+R_3} \begin{bmatrix} \boxed{1} & -1 & 2 & 1 \\ 0 & 3 & -3 & 6 \\ 0 & 2 & -2 & 4 \end{bmatrix}$$

Now, with $i = 2$,

$$\xrightarrow{\frac{1}{3}R_2} \begin{bmatrix} \boxed{1} & -1 & 2 & 1 \\ 0 & 1 & -1 & 2 \\ 0 & 2 & -2 & 4 \end{bmatrix} \xrightarrow{1R_2+R_1} \begin{bmatrix} \boxed{1} & 0 & 1 & 3 \\ 0 & 1 & -1 & 2 \\ 0 & 2 & -2 & 4 \end{bmatrix} \xrightarrow{-2R_2+R_3} \begin{bmatrix} \boxed{1} & 0 & 1 & 3 \\ 0 & \boxed{1} & -1 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

The system of equations represented by this augmented matrix needs to be considered a bit differently than that for Archetype B. First, the last row of the matrix is the equation $0 = 0$, which is *always* true, so we can safely ignore it as we analyze the other two equations. These equations are,

$$\begin{aligned} x_1 + x_3 &= 3 \\ x_2 - x_3 &= 2. \end{aligned}$$

While this sytem is fairly easy to solve, it also appears to have a multitude of solutions. For example, choose $x_3 = 1$ and see that then $x_1 = 2$ and $x_2 = 3$ will together form a solution. Or choose $x_3 = 0$, and then discover that $x_1 = 3$ and $x_2 = 2$ lead to a solution. Try it yourself: pick *any* value of x_3 you please, and figure out what x_1 and x_2 should be to make the first and second equations (respectively) true. We'll wait while you do that. Because of this behavior, we say that x_3 is a "free" or "independent" variable. But why do we vary x_3 and not some other variable? For now, notice that the third column of the augmented matrix does not have any leading 1's in its column. With this idea, we can rearrange the two equations, solving each for the variable that corresponds to the leading 1 in that row.

$$\begin{aligned} x_1 &= 3 - x_3 \\ x_2 &= 2 + x_3 \end{aligned}$$

To write the solutions in set notation, we have

$$S = \{(3 - x_3, 2 + x_3,\, x_3) \mid x_3 \in \mathbb{C}\}$$

We'll learn more in the next section about systems with infinitely many solutions and how to express their solution sets. Right now, you might look back at EXAMPLE SSSLE.IS [16].△

Example RREF.SAE
Solutions for Archetype E
Solve the following system of equations.

$$\begin{aligned} 2x_1 + x_2 + 7x_3 - 7x_4 &= 2 \\ -3x_1 + 4x_2 - 5x_3 - 6x_4 &= 3 \\ x_1 + x_2 + 4x_3 - 5x_4 &= 2 \end{aligned}$$

Form the augmented matrix for starters,

$$\begin{bmatrix} 2 & 1 & 7 & -7 & 2 \\ -3 & 4 & -5 & -6 & 3 \\ 1 & 1 & 4 & -5 & 2 \end{bmatrix}$$

and work to reduced row-echelon form, first with $i = 1$,

$$\xrightarrow{R_1 \leftrightarrow R_3} \begin{bmatrix} 1 & 1 & 4 & -5 & 2 \\ -3 & 4 & -5 & -6 & 3 \\ 2 & 1 & 7 & -7 & 2 \end{bmatrix}$$

$$\xrightarrow{3R_1 + R_2} \begin{bmatrix} 1 & 1 & 4 & -5 & 2 \\ 0 & 7 & 7 & -21 & 9 \\ 2 & 1 & 7 & -7 & 2 \end{bmatrix}$$

$$\xrightarrow{-2R_1 + R_3} \begin{bmatrix} \boxed{1} & 1 & 4 & -5 & 2 \\ 0 & 7 & 7 & -21 & 9 \\ 0 & -1 & -1 & 3 & -2 \end{bmatrix}$$

Now, with $i = 2$,

$$\xrightarrow{R_2 \leftrightarrow R_3} \begin{bmatrix} \boxed{1} & 1 & 4 & -5 & 2 \\ 0 & -1 & -1 & 3 & -2 \\ 0 & 7 & 7 & -21 & 9 \end{bmatrix}$$

$$\xrightarrow{-1R_2} \begin{bmatrix} \boxed{1} & 1 & 4 & -5 & 2 \\ 0 & 1 & 1 & -3 & 2 \\ 0 & 7 & 7 & -21 & 9 \end{bmatrix}$$

$$\xrightarrow{-1R_2+R_1} \begin{bmatrix} \boxed{1} & 0 & 3 & -2 & 0 \\ 0 & 1 & 1 & -3 & 2 \\ 0 & 7 & 7 & -21 & 9 \end{bmatrix}$$

$$\xrightarrow{-7R_2+R_3} \begin{bmatrix} \boxed{1} & 1 & 4 & -5 & 2 \\ 0 & \boxed{1} & 1 & -3 & 2 \\ 0 & 0 & 0 & 0 & -5 \end{bmatrix}$$

And finally, with $i = 3$,

$$\xrightarrow{-\frac{1}{5}R_3} \begin{bmatrix} \boxed{1} & 1 & 4 & -5 & 2 \\ 0 & \boxed{1} & 1 & -3 & 2 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

$$\xrightarrow{-2R_3+R_2} \begin{bmatrix} \boxed{1} & 1 & 4 & -5 & 2 \\ 0 & \boxed{1} & 1 & -3 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

$$\xrightarrow{-2R_3+R_1} \begin{bmatrix} \boxed{1} & 1 & 4 & -5 & 0 \\ 0 & \boxed{1} & 1 & -3 & 0 \\ 0 & 0 & 0 & 0 & \boxed{1} \end{bmatrix}$$

Lets analyze the equations in the system represented by this augmented matrix. The third equation will read $0 = 1$. This is patently false, all the time. No choice of values for our variables will ever make it true. We're done. Since we cannot even make the last equation true, we have no hope of making all of the equations simultaneously true. So this sytem has no solutions, and its solution set is the empty set, $\emptyset = \{\ \}$.

Notice that we could have reached this conclusion sooner. After performing the row operation $-7R_2 + R_3$, we can see that the third equation reads $0 = -5$, a false statement. Since the sytem represented by this matrix has no solutions, none of the systems represented has any solutions. However, for this example, we have chosen to bring the matrix fully to reduced row-echelon form for the practice. $\triangle$

These three examples (Example RREF.SAB [27], Example RREF.SAA [28], Example RREF.SAE [29]) illustrate the full range of possibilities for a system of linear equations — no solutions, one solution, or infinitely many solutions. In the next section we'll examine these three scenarios more closely.

Section TSS
Types of Solution Sets

We will now be more careful about analyzing the reduced row-echelon form derived from the augmented matrix of a system of linear equations. In particular, we will see how to systematically handle the situation when we have infinitely many solutions to a system, and we will prove that every system of linear equations has either zero, one or infinitely many solutions. With these tools, we will be able to solve any system by a well-described method.

The computer scientist Donald Knuth said, "Science is what we understand well enough to explain to a computer. Art is everything else." In this section we'll remove solving systems of equations from the realm of art, and into the realm of science. We begin with a definition.

Definition CS
Consistent System
A system of linear equations is **consistent** if it has at least one solution. Otherwise, the system is called **inconsistent**. $\odot$

We will want to first recognize when a system is inconsistent or consistent, and in the case of consistent sytems we will be able to further refine the types of solutions possible. We will do this by analyzing the reduced row-echelon form of a matrix, so we now initiate some notation that will help us talk about this form of a matrix.

Notation RREFA
Reduced Row-Echelon Form Analysis
Suppose that B is an $m \times n$ matrix that is in reduced row-echelon form. Let r equal the number of rows of B that are not zero rows. Each of these r rows then contains a leading 1, so let d_i equal the column number where row i's leading 1 is located. For columns without a leading 1, let f_i be the column number of the i-th column (reading from left to right) that does not contain a leading 1. Let

$$D = \{d_1,\, d_2,\, d_3,\, \ldots,\, d_r\} \qquad\qquad F = \{f_1,\, f_2,\, f_3,\, \ldots,\, f_{n-r}\} \qquad \triangledown$$

This notation can be a bit confusing, since we have subscripted variables that are in turn equal to subscripts used to index the matrix. However, many questions about matrices and systems of equations can be answered once we know r, D and F. An example may help.

Example TSS.RREFN
Reduced row-echelon form notation

For the 5×8 matrix

$$B = \begin{bmatrix} \boxed{1} & 5 & 0 & 0 & 2 & 8 & 0 & 5 \\ 0 & 0 & \boxed{1} & 0 & 4 & 7 & 0 & 2 \\ 0 & 0 & 0 & \boxed{1} & 3 & 9 & 0 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 & \boxed{1} & 4 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

in reduced row-echelon form we have

$$\begin{aligned} r &= 4 \\ d_1 &= 1 & d_2 &= 3 & d_3 &= 4 & d_4 &= 7 \\ f_1 &= 2 & f_2 &= 5 & f_3 &= 6 & f_4 &= 8. \end{aligned}$$

Notice that the sets $D = \{d_1, d_2, d_3, d_4\} = \{1, 3, 4, 7\}$ and $F = \{f_1, f_2, f_3, f_4\} = \{2, 5, 6, 8\}$ have nothing in common and together account for all of the columns of B (we say its a **partition** of the set of column indices). It is only a coincidence in this example that the two sets have the same size. △

Before proving some theorems about the possibilities for solution sets to systems of equations, lets analyze one particular system with an infinite solution set very carefully as an example. We'll use this technique frequently, and shortly we'll refine it slightly.

Archetypes I and J are both fairly large for doing computations by hand (though not impossibly large). Their properties are very similar, so we will frequently analyze the situation in Archetype I, and leave you the joy of analyzing Archetype J yourself. So work through Archetype I with the text, by hand and/or with a computer, and then tackle Archetype J yourself (and check your results with those listed). Notice too that the archetypes describing systems of equations each lists the values of r, D and F. Here we go…

Example TSS.ISS
Describing infinite solution sets
The system of $m = 4$ equations in $n = 7$ variables

$$\begin{aligned} x_1 + 4x_2 - x_4 + 7x_6 - 9x_7 &= 3 \\ 2x_1 + 8x_2 - x_3 + 3x_4 + 9x_5 - 13x_6 + 7x_7 &= 9 \\ 2x_3 - 3x_4 - 4x_5 + 12x_6 - 8x_7 &= 1 \\ -x_1 - 4x_2 + 2x_3 + 4x_4 + 8x_5 - 31x_6 + 37x_7 &= 4 \end{aligned}$$

has a 4×8 augmented matrix that is row-equivalent to the following matrix (check this!), and which is in reduced row-echelon form (the existence of this matrix is guaranteed by Theorem REMEF [25]),

$$\begin{bmatrix} \boxed{1} & 4 & 0 & 0 & 2 & 1 & -3 & 4 \\ 0 & 0 & \boxed{1} & 0 & 1 & -3 & 5 & 2 \\ 0 & 0 & 0 & \boxed{1} & 2 & -6 & 6 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

So we find that $r = 3$ and

$$D = \{d_1,\, d_2,\, d_3\} = \{1,\, 3,\, 4\} \qquad F = \{f_1,\, f_2,\, f_3,\, f_4,\, f_5\} = \{2,\, 5,\, 6,\, 7,\, 8\}\,.$$

Let i denote one of the $r = 3$ non-zero rows, and then we see that we can solve the corresponding equation represented by this row for the variable x_{d_i} and write it as a linear function of the variables x_{f_1}, x_{f_2}, x_{f_3}, x_{f_4} (notice that $f_5 = 8$ does not reference a variable). We'll do this now, but you can already see how the subscripts upon subscripts takes some getting used to.

$$\begin{aligned} &(i = 1) && x_{d_1} = x_1 = 4 - 4x_2 - 2x_5 - x_6 + 3x_7\\ &(i = 2) && x_{d_2} = x_3 = 2 - x_5 + 3x_6 - 5x_7\\ &(i = 3) && x_{d_3} = x_4 = 1 - 2x_5 + 6x_6 - 6x_7 \end{aligned}$$

Each element of the set $F = \{f_1,\, f_2,\, f_3,\, f_4,\, f_5\} = \{2,\, 5,\, 6,\, 7,\, 8\}$ is the index of a variable, except for $f_5 = 8$. We refer to $x_{f_1} = x_2$, $x_{f_2} = x_5$, $x_{f_3} = x_6$ and $x_{f_4} = x_7$ as "free" (or "independent") variables since they are allowed to assume any possible combination of values that we can imagine and we can continue on to build a solution to the system by solving individual equations for the values of the other ("dependent") variables.

Each element of the set $D = \{d_1,\, d_2,\, d_3\} = \{1,\, 3,\, 4\}$ is the index of a variable. We refer to the variables $x_{d_1} = x_1$, $x_{d_2} = x_3$ and $x_{d_3} = x_4$ as "dependent" variables since they *depend* on the *independent* variables. More precisely, for each possible choice of values for the independent variables we get *exactly one* set of values for the dependent variables that combine to form a solution of the system.

To express the solutions as a set with elements that are 7-tuples, we write

$$\{(4-4x_2-2x_5-x_6+3x_7,\, x_2, 2-x_5+3x_6-5x_7,\, 1-2x_5+6x_6-6x_7,\, x_5,\, x_6\, x_7) \mid x_2,\, x_5,\, x_6,\, x_7 \in \mathbb{C}\}$$

The condition that $x_2,\, x_5,\, x_6,\, x_7 \in \mathbb{C}$ is how we specify that the variables $x_2,\, x_5,\, x_6,\, x_7$ are "free" to assume any possible values.

This systematic approach to solving a system of equations will allow us to create a precise description of the solution set for any consistent system once we have found the reduced row-echelon form of the augmented matrix. It will work just as well when the set of free variables is empty and we get just a single solution. And we could program a computer to do it! Now have a whack at Archetype J, mimicking the discussion in this example. We'll still be here when you get back. △

Sets are an important part of algebra, and we've seen a few already. Being comfortable with sets is important for undewrstanding and writing proofs. So here's another proof technique.

Proof Technique SN
Set Notation
Sets are typically written inside of braces, as { }, and have two components. The first

is a description of the the type of objects contained in a set, while the second is some sort of restriction on the properties the objects have. Every object in the set must be of the type described in the first part and it must satisfy the restrictions in the second part. Conversely, any object of the proper type for the first part, that also meets the conditions of the second part, will be in the set. These two parts are set off from each other somehow, often with a vertical bar (|) or a colon (:). Membership of an element in a set is denoted with the symbol $\in$.

I like to think of sets as clubs. The first part is some description of the type of people who *might* belong to the club, the basic objects. For example, a bicycle club would describe its members as being people who like to ride bicycles. The second part is like a membership committe, it restricts the people who are allowed in the club. Continuing with our bicycle club, we might decide to limit ourselves to "serious" riders and only have members who can document having ridden 100 kilometers or more in a single day at least one time.

The restrictions on membership can migrate around some between the first and second part, and there may be several ways to describe the same set of objects. Here's a more mathematical example, employing the set of all integers, $\mathbb{Z}$, to describe the set of even integers.

$$E = \{x \in \mathbb{Z} \mid x \text{ is an even number}\} = \{x \in \mathbb{Z} \mid 2 \text{ divides } x \text{ evenly}\} = \{2k \mid k \in \mathbb{Z}\}$$

Notice how this set tells us that its objects are integer numbers (not, say, matrices or functions, for example) and just those that are even. So we can write that $10 \in E$, while $17 \notin E$ once we check the membership criteria. We also recognize the question

$$\begin{bmatrix} 1 & -3 & 5 \\ 2 & 0 & 3 \end{bmatrix} \in E?$$

as being ridiculous. ◊

We mix our metaphors a bit when we call variables free versus dependent. Maybe we should call dependent variables "enslaved"? Here's the definition.

Definition IDV
Independent and Dependent Variables
Suppose A is the augmented matrix of a system of linear equations and B is a row-equivalent matrix in reduced row-echelon form. Suppose j is the number of a column of B that contains the leading 1 of some row, and it is not the last column. Then the variable j is **dependent**. A variable that is not dependent is called **independent** or **free**. ⊙

We can now use the values of m, n, r, and the independent and dependent variables to categorize the solutions sets to linear systems through a sequence of theorems. First the distinction between consistent and inconsistent systems, after two explanations of some proof techniques we will be using.

Proof Technique E
Equivalences
When a theorem uses the phrase “if and only if” (or the abbreviation “iff”) it is a shorthand way of saying that two if-then statements are true. So if a theorem says “A if and only if B,” then it is true that “if A, then B” while it is also true that “if B, then A.” For example, it may be a theorem that “I wear bright yellow knee-high plastic boots if and only if it is raining.” This means that I *never* forget to wear my super-duper yellow boots when it is raining *and* I wouldn’t be seen in such silly boots *unless* it was raining. You never have one without the other. I’ve got my boots on and it is raining *or* I don’t have my boots on and its dry.

The upshot for proving such theorems is that its like a 2-for-1 sale, we get to do *two* proofs. Assume A and conclude B, then start over and assume B and conclude A. For this reason, “if and only if” is sometimes abbreviated by $\iff$, while proofs indicate which of the two implications is being proved by prefacing each with $\Rightarrow$ or $\Leftarrow$. A carefully written proof will remind the reader which statement is being used as the hypothesis, a quicker version will let the reader deduce it from the direction of the arrow. Tradition dictates we do the “easy” half first, but that’s hard for a student to know until you’ve finished doing both halves! Oh well, if you rewrite your proofs (a good habit), you can then choose to put the easy half first.

Theorems of this type are called equivalences or characterizations, and they are some of the most pleasing results in mathematics. They say that two objects, or two situations, are really the same. You don’t have one without the other, like rain and my yellow boots. The more different A and B seem to be, the more pleasing it is to discover they are really equivalent. And if A describes a very mysterious solution or involves a tough computation, while B is transparent or involves easy computations, then we’ve found a great shortcut for better understanding or faster computation. Remember that every theorem really is a shortcut in some form. You will also discover that if proving $A \Rightarrow B$ is very easy, then proving $B \Rightarrow A$ is likely to be proportionately harder. Sometimes the two halves are about equally hard. And in rare cases, you can string together a whole sequence of other equivalences to form the one you’re after and you don’t even need to do two halves. In this case, the argument of one half is just the argument of the other half, but in reverse.

One last thing about equivalences. If you see a statement of a theorem that says two things are “equivalent,” translate it first into an “if and only if” statement. $\Diamond$

Proof Technique CP
Contrapositives
The **contrapositive** of an implication $A \Rightarrow B$ is the implication $\text{not}(B) \Rightarrow \text{not}(A)$, where “not” means the logical negation, or opposite. An implication is true if and only if its contrapositive is true. In symbols, $(A \Rightarrow B) \iff (\text{not}(B) \Rightarrow \text{not}(A))$ is a theorem. Such statements about logic, that are always true, are known as **tautologies**.

For example, its a theorem that “if a vehicle is a fire truck, then it has big tires and has a siren.” (Yes, I’m sure you can conjure up a counterexample, but play along with me anyway.) The contrapositive is "if a vehicle does not have big tires or does not have a siren, then it is not a fire truck." Notice how the "and" became an "or" when we negated the conclusion of the original theorem.

It will frequently happen that it is easier to construct a proof of the contrapositive than of the original implication. If you are having difficulty formulating a proof of some implication, see if the contrapositive is easier for you. The trick is to construct the negation of complicated statements accurately. More on that later. ◊

Theorem RCLS
Recognizing Consistency of a Linear System
Suppose A is the augmented matrix of a system of linear equations with m equations in n variables. Suppose also that B is a row-equivalent matrix in reduced row-echelon form with r rows that are not zero rows. Then the system of equations is inconsistent if and only if the leading 1 of row r is located in column $n+1$ of B. □

Proof ($\Leftarrow$) The first half of the proof begins with the assumption that the leading 1 of row r is located in column $n+1$ of B. Then row r of B begins with n consecutive zeros, finishing with the leading 1. This is a representation of the equation $0 = 1$, which is false. Since this equation is false for any collection of values we might choose for the variables, there are no solutions for the system of equations, and it is inconsistent.

($\Rightarrow$) For the second half of the proof, we wish to show that if we assume the system is inconsistent, then the final leading 1 is located in the last column. But instead of proving this directly, we'll form the logically equivalent statement that is the contrapositive, and prove that instead (see Technique CP [35]). Turning the implication around, and negating each portion, we arrive at the equivalent statement: If the leading 1 of row r is not in column $n+1$, then the system of equations is consistent.

If the leading 1 for row i is located somewhere in columns 1 through n, then *every* preceding row's leading 1 is also located in columns 1 through n. In other words, since the last leading 1 is not in the last column, no leading 1 for any row is in the last column, due to the echelon layout of the leading 1's. Let $b_{i,n+1}$, $1 \leq i \leq r$ denote the entries of the last column of B for the first r rows. Employ our notation for columns of the reduced row-echelon form of a matrix (see Notation RREFA [31]) to B and set $x_{f_i} = 0$, $1 \leq i \leq n-r$ and then set $x_{d_i} = b_{i,n+1}$, $1 \leq i \leq r$. These values for the variables make the equations represented by the first r rows all true (convince yourself of this). Rows $r+1$ through m (if any) are all zero rows, hence represent the equation $0 = 0$ and are also all true. We have now identified one solution to the system, so we can say it is consistent. ■

The beauty of this theorem being an equivalence is that we can unequivocally test to see if a system is consistent or inconsistent by looking at just a single entry of the reduced row-echelon form matrix. We could program a computer to do it!

Notice that for a consistent system the row-reduced augmented matrix has $n+1 \in F$, so the largest element of F does not refer to a variable. Also, for an inconsistent system, $n+1 \in D$, and it then does not make much sense to discuss whether or not variables are free or dependent since there is no solution.

Theorem ICRN
Inconsistent Systems, r and n
Suppose A is the augmented matrix of a system of linear equations with m equations in n variables. Suppose also that B is a row-equivalent matrix in reduced row-echelon form with r rows that are not completely zeros. If $r = n + 1$, then the system of equations is inconsistent. □

Proof If $r = n + 1$, then $D = \{1, 2, 3, \ldots, n, n + 1\}$ and every column of B contains a leading 1. In particular, the entry of column $n + 1$ for row $r = n + 1$ is a leading 1. Theorem RCLS [36] then says that the system is inconsistent. ■

Theorem CSRN
Consistent Systems, r and n
Suppose A is the augmented matrix of a *consistent* system of linear equations with m equations in n variables. Suppose also that B is a row-equivalent matrix in reduced row-echelon form with r rows that are not zero rows. Then $r \leq n$. If $r = n$, then the system has a unique solution, and if $r < n$, then the system has infinitely many solutions. □

Proof This theorem contains three implications that we must establish. Notice first that the echelon layout of the leading 1's means that there are at most $n + 1$ leading 1's and therefore $r \leq n + 1$ for any system. We are assuming this system is consistent, so we know by Theorem ICRN [37] that $r \neq n + 1$. Together these two observations leave us with $r \leq n$.

When $r = n$, we find $n - r = 0$ free variables (i.e. $F = \{n + 1\}$) and any solution must equal the unique solution given by the first n entries of column $n + 1$ of B.

When $r < n$, we have $n - r > 0$ free variables, corresponding to columns of B without a leading 1, excepting the final column, which also does not contain a leading 1 by Theorem RCLS [36]. By varying the values of the free variables suitably, we can demonstrate infinitely many solutions. ■

Corollary FVCS
Free Variables for Consistent Systems
Suppose A is the augmented matrix of a *consistent* system of linear equations with m equations in n variables. Suppose also that B is a row-equivalent matrix in reduced row-echelon form with r rows that are not completely zeros. Then the solution set can be described with $n - r$ free variables. □

Proof Technique CV
Converses
The **converse** of the implication $A \Rightarrow B$ is the implication $B \Rightarrow A$. There is no guarantee that the truth of these two statements are related. In particular, if an implication has been proven to be a theorem, then do not try to use its converse too as if it were a theorem.

Sometimes the converse is true (and we have an equivalence, see Technique E [35]). But more likely the converse is false, especially if it wasn't included in the statement of the original theorem.

For example, we have the theorem, "if a vehicle is a fire truck, then it is has big tires and has a siren." The converse is false. The statement that "if a vehicle has big tires and a siren, then it is a fire truck" is false. A police vehicle for use on a sandy public beach would have big tires and a siren, yet is not equipped to fight fires.

We bring this up now, because Theorem CSRN [37] has a tempting converse. Does this theorem say that if $r < n$, then the system is consistent? Definitely not, as Archetype E [183] has $r = 2$ and $n = 4$ but is inconsistent. This example is then said to be a counterexample to the converse. Whenever you think a theorem that is an implication might actually be an equivalence, its good to hunt around for a counterexample that shows the converse to be false. ♢

Example TSS.CFV
Counting free variables
For each archetype that is a system of equations, the values of n and r are listed. Many also contain a few sample solutions. We can use this information profitably, as illustrated by four examples.

1. Archetype A [167] has $n = 3$ and $r = 2$. It can be seen to be consistent by the sample solutions given. Its solution set then has $n - r = 1$ free variables, and therefore will be infinite.

2. Archetype B [171] has $n = 3$ and $r = 3$. It can be seen to be consistent by the single sample solution given. Its solution set then has $n - r = 0$ free variables, and therefore will have just the single solution.

3. Archetype H [196] has $n = 2$ and $r = 3$. In this case, $r = n + 1$, so Theorem ICRN [37] says the system is inconsistent. We should not try to apply Corollary FVCS [37] to count free variables, since the theorem only applies to consistent systems. (What would happen if you did?)

4. Archetype E [183] has $n = 4$ and $r = 3$. However, by looking at the reduced row-echelon form of the augmented matrix, we find a leading 1 in row 3, column 4. By Theorem RCLS [36] we recognize the system is then inconsistent. (Why doesn't this example contradict Theorem ICRN [37]?) △

We have accomplished a lot so far, but our main goal has been the following theorem, which is now very simple to prove. The proof is so simple that we ought to call it a corollary, but the result is important enough that it deserves to be called a theorem. Notice that this theorem was presaged first by Example SSSLE.TTS [9] and further foreshadowed by other examples.

Theorem PSSLS
Possible Solution Sets for Linear Systems
A simultaneous system of linear equations has no solutions, a unique solution or infinitely many solutions. □

Proof By definition, a system is either inconsistent or consistent. The first case describes systems with no solutions. For consistent systems, we have the remaining two possibilites as guranteed by, and described in, Theorem CSRN [37]. ■

We have one more theorem to round out our set of tools for determining solution sets to systems of linear equations.

Theorem CMVEI
Consistent, More Variables than Equations implies Infinite solutions
Suppose A is the augmented matrix of a consistent system of linear equations with m equations in n variables. If $n > m$, then the system has infinitely many solutions. □

Proof Suppose that B is a row-equivalent matrix in reduced row-echelon form with r rows that are not completely zeros. Because B has m rows in total the number that are not zero rows is fewer, that is, $r \leq m$. Follow this with the hypothesis that $n > m$ and we find that the system has a solution set described by at least one free variable because

$$n - r \geq n - m > 0.$$

A consistent system with free variables will have an infinite number of solutions, as given by Theorem CSRN [37]. ■

Notice that to use this theorem we need only know that the system is consistent, together with the values of m and n. We do not necessarily have to compute a row-equivalent reduced row-echelon form matrix, even though we discussed such a matrix in the proof. This is the substance of the following example.

Example TSS.OSGM
One solution gives many
Archetype D is the system of $m = 3$ equations in $n = 4$ variables,

$$\begin{aligned} 2x_1 + x_2 + 7x_3 - 7x_4 &= 8 \\ -3x_1 + 4x_2 - 5x_3 - 6x_4 &= -12 \\ x_1 + x_2 + 4x_3 - 5x_4 &= 4 \end{aligned}$$

and the solution $x_1 = 0$, $x_2 = 1$, $x_3 = 2$, $x_4 = 1$ can be checked easily by substitution. Having been *handed* this solution, we know the system is consistent. This, together with $n > m$, allows us to apply Theorem CMVEI [39] and conclude that the system has infinitely many solutions. △

These theorems give us the procedures and implications that allow us to completely solve any simultaneous system of linear equations. The main computational tool is using row operations to convert an augmented matrix into reduced row-echelon form. Here's a broad outline of how we would instruct a computer to solve a system of linear equations.

1. Repesent a system of linear equations by an augmented matrix (an array is the appropriate data structure in most computer languages).

2. Convert the matrix to a row-equivalent matrix in reduced row-echelon form using the procedure from the proof of Theorem REMEF [25].

3. Determine r and locate the leading 1 of row r. If it is in column $n+1$, output the statement that the system is inconsistent and halt.

4. With the leading 1 of row r not in column $n+1$, there are two possibilities:

   (a) $r = n$ and the solution is unique. It can be read off directly from the entries in rows 1 through n of column $n+1$.

   (b) $r < n$ and there are infinitely many solutions. If only a single solution is needed, set all the free variables to zero and read off the dependent variable values from column $n+1$, as in the second half of the proof of Theorem RCLS [36]. If the entire solution set is required, figure out some nice compact way to describe it, since your finite computer is not big enough to hold all the solutions (we'll have such a way soon).

The above makes it all sound a bit simpler than it really is. In practice, row operations employ division (usually to get a leading entry of a row to convert to a leading 1) and that will introduce round-off errors. Entries that should be zero sometimes end up being very, very small nonzero entries, or small entries lead to overflow errors when used as divisors. A variety of strategies can be employed to minimize these sorts of errors, and this is one of the main topics in the important subject known as numerical linear algebra.

In this section we've gained a foolproof procedure for solving any system of linear equations, no matter how many equations or variables. We also have a handful of theorems that allow us to determine partial information about a solution set without actually constructing the whole set itself.

Section HSE
Homogenous Systems of Equations

In this section we specialize to systems of linear equations where every equation has a zero as its constant term. Along the way, we will begin to express more and more ideas in the language of matrices and begin a move away from writing out whole systems of equations. The ideas initiated in this section will carry through the remainder of the course.

Subsection SHS
Solutions of Homogenous Systems

As usual, we begin with a definition.

Definition HS
Homogenous System
A system of linear equations is **homogenous** if each equation has a 0 for its constant term. Such a system then has the form,

$$\begin{aligned}
a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \cdots + a_{1n}x_n &= 0\\
a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \cdots + a_{2n}x_n &= 0\\
a_{31}x_1 + a_{32}x_2 + a_{33}x_3 + \cdots + a_{3n}x_n &= 0\\
&\vdots\\
a_{m1}x_1 + a_{m2}x_2 + a_{m3}x_3 + \cdots + a_{mn}x_n &= 0
\end{aligned}$$

⊙

Example HSE.AHSAC
Archetype C as a homogenous system
For each archetype that is a system of equations, we have formulated a similar, yet different, homogenous system of equations by replacing each equation's constant term with a zero. To wit, for Archetype C, we can convert the orginal system of equations into the homogenous system,

$$\begin{aligned}
2x_1 - 3x_2 + x_3 - 6x_4 &= 0\\
4x_1 + x_2 + 2x_3 + 9x_4 &= 0\\
3x_1 + x_2 + x_3 + 8x_4 &= 0
\end{aligned}$$

Can you quickly find a solution to this system without row-reducing the augmented matrix? △

As you might have discovered by studying Example HSE.AHSAC [41], setting each variable to zero will *always* be a solution of a homogenous system. This is the substance of the following theorem.

Theorem HSC
Homogenous Systems are Consistent
Suppose that a system of linear equations is homogenous. Then it is consistent. □

Proof Set each variable of the system to zero. When substituting these values into each equation, the left-hand side evaluates to zero, no matter what the coefficients are. Since a homogenous system has zero on the right-hand side of each equation as the constant term, each equation is true. With one demonstrated solution, we can call the system consistent. ■

Since this solution is so obvious, we now define it as the trivial solution.

Definition TS
Trivial Solution
Suppose a homogenous system of linear equations has n variables. The solution $x_1 = 0$, $x_2 = 0, \ldots, x_n = 0$ is called the **trivial solution**. ⊙

Here are three typical examples, which we will reference throughout this section. Work through the row operations as we bring each to reduced row-echelon form. Also notice what is similar in each example, and what differs.

Example HSE.HUSAB
Homogenous, unique solution, Archetype B
Archetype B can be converted to the homogenous system,

$$\begin{aligned} -11x_1 + 2x_2 - 14x_3 &= 0 \\ 23x_1 - 6x_2 + 33x_3 &= 0 \\ 14x_1 - 2x_2 + 17x_3 &= 0 \end{aligned}$$

whose augmented matrix row-reduces to

$$\begin{bmatrix} \boxed{1} & 0 & 0 & 0 \\ 0 & \boxed{1} & 0 & 0 \\ 0 & 0 & \boxed{1} & 0 \end{bmatrix}$$

By Theorem HSC [42], the system is consistent, and so the computation $n - r = 3 - 3 = 0$ means the solution set contains just a single solution. Then, this lone solution must be the trivial solution. △

Example HSE.HISAA
Homogenous, infinite solutions, Archetype A
Archetype A [167] can be converted to the homogenous system,

$$\begin{aligned} x_1 - x_2 + 2x_3 &= 0 \\ 2x_1 + x_2 + x_3 &= 0 \\ x_1 + x_2 \qquad &= 0 \end{aligned}$$

whose augmented matrix row-reduces to

$$\begin{bmatrix} \boxed{1} & 0 & 1 & 0 \\ 0 & \boxed{1} & -1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

By Theorem HSC [42], the system is consistent, and so the computation $n - r = 3 - 2 = 1$ means the solution set contains one free variable by Corollary FVCS [37], and hence has infinitely many solutions. We can describe this solution set using the free variable x_3,

$$S = \{(x_1,\, x_2,\, x_3) \mid x_1 = -x_3,\, x_2 = x_3\} = \{(-x_3,\, x_3,\, x_3) \mid x_3 \in \mathbb{C}\}$$

Geometrically, these are points in three dimensions that lie on a line through the origin.△

Example HSE.HISAD
Homogenous, infinite solutions, Archetype D
Archetype D [179] (and identically, Archetype E [183]) can be converted to the homogenous system,

$$\begin{aligned} 2x_1 + x_2 + 7x_3 - 7x_4 &= 0 \\ -3x_1 + 4x_2 - 5x_3 - 6x_4 &= 0 \\ x_1 + x_2 + 4x_3 - 5x_4 &= 0 \end{aligned}$$

whose augmented matrix row-reduces to

$$\begin{bmatrix} \boxed{1} & 0 & 3 & -2 & 0 \\ 0 & \boxed{1} & 1 & -3 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

By Theorem HSC [42], the system is consistent, and so the computation $n - r = 4 - 2 = 2$ means the solution set contains two free variables by Corollary FVCS [37], and hence has infinitely many solutions. We can describe this solution set using the free variables x_3 and x_4,

$$\begin{aligned} S &= \{(x_1,\, x_2,\, x_3,\, x_4) \mid x_1 = -3x_3 + 2x_4,\, x_2 = -x_3 + 3x_4\} \\ &= \{(-3x_3 + 2x_4,\, -x_3 + 3x_4,\, x_3,\, x_4) \mid x_3,\, x_4 \in \mathbb{C}\} \end{aligned}$$

△

After working through these examples, you might perform the same computations for the slightly larger example, Archetype J [206].

Example HSE.HISAD [43] suggests the following theorem.

Theorem HMVEI
Homogenous, More Variables than Equations implies Infinite solutions
Suppose that a homogenous system of linear equations has m equations and n variables with $n > m$. Then the system has infinitely many solutions. □

Proof We are assuming the system is homogenous, so Theorem HSC [42] says it is consistent. Then the hypothesis that $n > m$, together with Theorem CMVEI [39], gives infinitely many solutions. ■

Example HSE.HUSAB [42] and Example HSE.HISAA [43] are concerned with homogenous systems where $n = m$ and expose a fundamental distinction between the two examples. One has a unique solution, while the other has infinitely many. These are exactly the only two possibilities for a homogenous system and illustrate that each is possible (unlike the case when $n > m$ where Theorem HMVEI [44] tells us that there is only one possibility for a homogenous system).

Subsection MVNSE
Matrix and Vector Notation for Systems of Equations

Notice that when we do row operations on the augmented matrix of a homogenous system of linear equations the last colum of the matrix is all zeros. Any one of the three allowable row operations will convert zeros to zeros and thus, the final column of the matrix in reduced row-echelon form will also be all zeros. This observation might suffice as a first explanation of the reason for some of the following definitions.

Definition CV
Column Vector
A **column vector** of **size** m is an $m \times 1$ matrix. We will frequently refer to a column vector as simply a **vector**. ⊙

Notation VN
Vector (u)
Vectors will be written in bold, usually with lower case letters **u**, **v**, **w**, **x**, **y**, **z**. Some books like to write vectors with arrows, such as $\vec{u}$. Writing by hand, some like to put arrows on top of the sysmbol, or a tilde underneath the symbol, as in $\underset{\sim}{u}$. ▽

Definition ZV
Zero Vector
The **zero vector** is the $m \times 1$ matrix

$$\mathbf{0} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

⊙

Notation ZVN
Zero Vector (0)
The zero vector will be written as $\mathbf{0}$. ▽

Definition CM
Coefficient Matrix
For a system of linear equations,

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \cdots + a_{2n}x_n &= b_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 + \cdots + a_{3n}x_n &= b_3 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + a_{m3}x_3 + \cdots + a_{mn}x_n &= b_m \end{aligned}$$

the **coefficient matrix** is the $m \times n$ matrix

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \dots & a_{3n} \\ \vdots & & & & \\ a_{m1} & a_{m2} & a_{m3} & \dots & a_{mn} \end{bmatrix}$$

⊙

Definition VOC
Vector of Constants
For a system of linear equations,

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \cdots + a_{2n}x_n &= b_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 + \cdots + a_{3n}x_n &= b_3 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + a_{m3}x_3 + \cdots + a_{mn}x_n &= b_m \end{aligned}$$

the **vector of constants** is the column vector of size m

$$\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_m \end{bmatrix}$$ ⊙

Definition SV
Solution Vector
For a system of linear equations,

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \cdots + a_{2n}x_n &= b_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 + \cdots + a_{3n}x_n &= b_3 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + a_{m3}x_3 + \cdots + a_{mn}x_n &= b_m \end{aligned}$$

the **solution vector** is the column vector of size m

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_m \end{bmatrix}$$ ⊙

The solution vector may do double-duty on occassion. It might refer to a list of variable quantities at one point, and subsequently refer to values of those variables that actually form a particular solution to that system.

Notation AMN
Augmented Matrix ($[A|\mathbf{b}]$)
With these definitions, we will write the augmented matrix of system of linear equations in the form $[A|\mathbf{b}]$ in order to identify and distinguish the coefficients and the constants. ▽

Notation LSN
Linear System ($\mathcal{LS}(A,\,\mathbf{b})$)
We will write $\mathcal{LS}(A,\,\mathbf{b})$ to denote the system of linear equations with A as a coefficient matrix and $\mathbf{b}$ as the vector of constants. ▽

Example HSE.NSLE
Notation for systems of linear equations
The system of linear equations

$$\begin{aligned} 2x_1 + 4x_2 - 3x_3 + 5x_4 + x_5 &= 9 \\ 3x_1 + x_2 + \quad + x_4 - 3x_5 &= 0 \\ -2x_1 + 7x_2 - 5x_3 + 2x_4 + 2x_5 &= -3 \end{aligned}$$

has coefficient matrix

$$A = \begin{bmatrix} 2 & 4 & -3 & 5 & 1 \\ 3 & 1 & 0 & 1 & -3 \\ -2 & 7 & -5 & 2 & 2 \end{bmatrix}$$

and vector of constants

$$\mathbf{b} = \begin{bmatrix} 9 \\ 0 \\ -3 \end{bmatrix}$$

and so will be referenced as $\mathcal{LS}(A, \mathbf{b})$. △

With these definitions and notation a homogenous system will be notated as $[A|\mathbf{0}]$, and when converted to reduced row-echelon form it will still have the final column of zeros. So in this case, we may be as likely to just reference only the coefficient matrix.

Subsection NSM
Null Space of a Matrix

The set of solutions to a homogenous system (which by THEOREM HSC [42] is never empty) is of enough interest to warrant its own name. However, we define it as a property of the coefficient matrix, not as a property of some system of equations.

Definition NSM
Null Space of a Matrix
The **null space** of a matrix A, denoted $\mathrm{N}(A)$, is the set of all the vectors that are solutions to the homogenous system $\mathcal{LS}(A, \mathbf{0})$. ⊙

In the Archetypes (CHAPTER A [163]) each example that is a system of equations also has a corresponding homogenous system of equations listed, and several sample solutions are given. These solutions will be elements of the null space of the coefficient matrix. We'll look at one example.

Example HSE.NSEAI
Null space elements of Archetype I

The write-up for Archetype I [201] lists several solutions of the corresponding homogenous system. Here are two, written as solution vectors. We can say that they are in the null space of the coefficient matrix for the system of equations in Archetype I [201].

$$\mathbf{x} = \begin{bmatrix} 3 \\ 0 \\ -5 \\ -6 \\ 0 \\ 0 \\ 1 \end{bmatrix} \qquad \mathbf{y} = \begin{bmatrix} -4 \\ 1 \\ -3 \\ -2 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

However, the vector

$$\mathbf{z} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 2 \end{bmatrix}$$

is not in the null space, since it is not a solution to the homogenous sytem. For example, it fails to even make the first equation true. △

Section NSM
NonSingular Matrices

In this section we specialize to systems with equal numbers of equations and variables, which will prove to be a case of special interest.

Subsection NSM
NonSingular Matrices

Our theorems will now establish connections between systems of equations (homogenous or otherwise), augmented matrices representing those systems, coefficient matrices, constant vectors, the reduced row-echelon form of matrices (augmented and coefficient) and solution sets. Be very careul in your reading, writing and speaking about systems of equations, matrices and sets of vectors. Now would be a good time to review the discussion about speaking and writing mathematics in TECHNIQUE L [18].

Definition SQM
Square Matrix
A matrix with m rows and n columns is **square** if $m = n$. In this case, we say the matrix has **size** n. To emphasize the situation when a matrix is not square, we will call it **rectangular**. $\odot$

We can now present one of the central definitions of linear algebra.

Definition NM
Nonsingular Matrix
Suppose A is a square matrix. And suppose the homogenous linear system of equations $\mathcal{LS}(A, \mathbf{0})$ has *only* the trivial solution. Then we say that A is a **nonsingular** matrix. Otherwise we say A is a **singular** matrix. $\odot$

While the definition of a nonsingular matrix involves a system of equations, it is a property only of matrices, and just square matrices at that. So it makes no sense to call a system of equations nonsingular, nor does it make any sense to call a 5×7 matrix singular.

Example NSM.S
A singular matrix, Archetype A

Example HSE.HISAA [43] shows that the coefficient matrix derived from Archetype A [167], specifically the 3×3 matrix,

$$A = \begin{bmatrix} 1 & -1 & 2 \\ 2 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix}$$

is a singular matrix since there are nontrivial solutions to the homogenous system $\mathcal{LS}(A, \mathbf{0})$. △

Example NSM.NS
A nonsingular matrix, Archetype B
Example HSE.HUSAB [42] shows that the coefficient matrix derived from Archetype B [171], specifically the 3×3 matrix,

$$B = \begin{bmatrix} -7 & -6 & -12 \\ 5 & 5 & 7 \\ 1 & 0 & 4 \end{bmatrix}$$

is a nonsingular matrix since the homogenous system, $\mathcal{LS}(B, \mathbf{0})$, has only the trivial solution. △

Notice that we will not discuss Example HSE.HISAD [43] as being a singular or nonsingular coefficient matrix since the matrix is not square.

The next theorem combines with our main computational technique (row-reducing a matrix) to make it easy to recognize a nonsingular matrix. But first a definition.

Definition IM
Identity Matrix
The $n \times n$ **identity matrix**, $I_n = (a_{ij})$ has $a_{ij} = 1$ whenever $i = j$, and $a_{ij} = 0$ whenever $i \neq j$. ⊙

Example NSM.IM
An identity matrix
The 4×4 identity matrix is

$$I_4 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

△

Notice that an identity matrix is square, and in reduced row-echelon form.

Theorem NSRRI
NonSingular matrices Row Reduce to the Identity matrix
Suppose that A is a square matrix and B is a row-equivalent matrix in reduced row-echelon form. Then A is nonsingular if and only if B is the identity matrix. □

Proof ($\Leftarrow$) Suppose B is the identity matrix. When the augmented matrix $[A|\mathbf{0}]$ is row-reduced, the result is $[B|\mathbf{0}] = [I_n|\mathbf{0}]$. The number of nonzero rows is equal to the number of variables in the linear system of equations $\mathcal{LS}(A,\,\mathbf{0})$, so $n = r$ and Corollary FVCS [37] gives $n - r = 0$ free variables. Thus, the homogenous system $\mathcal{LS}(A,\,\mathbf{0})$ has just one solution, which must be the trivial solution. This is exactly the definition of a nonsingular matrix.

($\Rightarrow$) We will prove the contrapositive. Suppose B is not the identity matrix. When the augmented matrix $[A|\mathbf{0}]$ is row-reduced, the result is $[B|\mathbf{0}]$. The number of rows not completely zero is less than the number of variables for the system of equations, so Corollary FVCS [37] gives $n - r > 0$ free variables. Thus, the homogenous system $\mathcal{LS}(A,\,\mathbf{0})$ has infinitely many solutions, so the system has more solutions than just the trivial solution. Thus the matrix is not nonsingular, the desired conclusion. ■

Notice that since this theorem is an equivalence it will always allow us to determine if a matrix is either nonsingular or singular. Here are two examples of this, continuing our study of Archetype A and Archetype B.

Example NSM.SRR
Singular matrix, row-reduced
The coefficient matrix for Archetype A [167] is

$$A = \begin{bmatrix} 1 & -1 & 2 \\ 2 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix}$$

which when row-reduced becomes the row-equivalent matrix

$$B = \begin{bmatrix} \boxed{1} & 0 & 1 \\ 0 & \boxed{1} & -1 \\ 0 & 0 & 0 \end{bmatrix}.$$

Since this matrix is not the 3×3 identity matrix, Theorem NSRRI [50] tells us that A is a singular matrix. △

Example NSM.NSRR
NonSingular matrix, row-reduced
The coefficient matrix for Archetype B [171] is

$$A = \begin{bmatrix} -7 & -6 & -12 \\ 5 & 5 & 7 \\ 1 & 0 & 4 \end{bmatrix}$$

which when row-reduced becomes the row-equivalent matrix

$$B = \begin{bmatrix} \boxed{1} & 0 & 0 \\ 0 & \boxed{1} & 0 \\ 0 & 0 & \boxed{1} \end{bmatrix}.$$

Since this matrix is the 3×3 identity matrix, Theorem NSRRI [50] tells us that A is a nonsingular matrix. △

Example NSM.NSS
Null space of a singular matrix
Given the coefficient matrix from Archetype A [167],

$$A = \begin{bmatrix} 1 & -1 & 2 \\ 2 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix}$$

the null space is the set of solutions to the homogenous system of equations $\mathcal{LS}(A,\,\mathbf{0})$ has a solution set and null space constructed in Example HSE.HISAA [43] as

$$\mathrm{N}(A) = \left\{ \begin{bmatrix} -x_3 \\ x_3 \\ x_3 \end{bmatrix} \mid x_3 \in \mathbb{C} \right\}$$ △

Example NSM.NSNS
Null space of a nonsingular matrix
Given the coefficient matrix from Archetype B,

$$A = \begin{bmatrix} -7 & -6 & -12 \\ 5 & 5 & 7 \\ 1 & 0 & 4 \end{bmatrix}$$

the homogenous system $\mathcal{LS}(A,\,\mathbf{0})$ has a solution set constructed in Example HSE.HUSAB [42] that contains only the trivial solution, so the null space has only a single element,

$$\mathrm{N}(A) = \left\{ \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \right\}$$ △

These two examples illustrate the next theorem, which is another equivalence.

Theorem NSTNS
NonSingular matrices have Trivial Null Spaces
Suppose that A is a square matrix. Then A is nonsingular if and only if the null space of A, $\mathrm{N}(A)$, contains only the trivial solution to the system $\mathcal{LS}(A,\,\mathbf{0})$, i.e. $\mathrm{N}(A) = \{\mathbf{0}\}$. □

Proof The null space of a square *matrix*, A, is the set of solutions to the homogenous *system*, $\mathcal{LS}(A,\,\mathbf{0})\mathbf{0}$. A *matrix* is nonsingular if and only if the set of solutions to the homogenous *system*, $\mathcal{LS}(A,\,\mathbf{0})$, has only a trivial solution. These two observations may be chained together to construct the two proofs necessary for each of half of this theorem. ■

Proof Technique U
Uniqueness
A theorem will sometimes claim that some object, having some desirable property, is unique. In other words, there should be only one such object. To prove this, a standard technique is to assume there are two such objects and proceed to analyze the consequences. The end result may be a contradiction, or the conclusion that the two allegedly different objects really are equal. ◊

The next theorem pulls a lot of ideas together. It tells us that we can learn a lot about solutions to a system of linear equations with a square coefficient matrix by examining a similar homogenous system.

Theorem NSMUS
NonSingular Matrices and Unique Solutions
Suppose that A is a square matrix. A is a nonsingular matrix if and only if the system $\mathcal{LS}(A,\,\mathbf{b})$ has a unique solution for every choice of the constant vector $\mathbf{b}$. □

Proof (⇐) The hypothesis for this half of the proof is that the system $\mathcal{LS}(A,\,\mathbf{b})$ has a unique solution for *every* choice of the constant vector $\mathbf{b}$. We will make a very specific choice for $\mathbf{b}$: $\mathbf{b}=\mathbf{0}$. Then we know that the system $\mathcal{LS}(A,\,\mathbf{0})$ has a unique solution. But this is precisely the definition of what it means for A to be nonsingular. If this half of the proof seemed too easy, perhaps we'll have to work a bit harder to get the implication in the opposite direction.

(⇒) We will assume A is nonsingular, and try to solve the system $\mathcal{LS}(A,\,\mathbf{b})$ without making any assumptions about $\mathbf{b}$. To do this we will begin by constructing a new homogenous linear system of equations that looks very much like the original. Suppose A has size n (why must it be square?) and write the original system as,

$$
\begin{aligned}
a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \cdots + a_{1n}x_n &= b_1\\
a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \cdots + a_{2n}x_n &= b_2\\
a_{31}x_1 + a_{32}x_2 + a_{33}x_3 + \cdots + a_{3n}x_n &= b_3\\
&\vdots\\
a_{n1}x_1 + a_{n2}x_2 + a_{n3}x_3 + \cdots + a_{nn}x_n &= b_n
\end{aligned}
\tag{$*$}
$$

form the new, homogenous system in n equations with $n+1$ variables, by adding a new variable y, whose coefficients are the negatives of the constant terms,

$$
\begin{aligned}
a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \cdots + a_{1n}x_n - b_1y &= 0\\
a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \cdots + a_{2n}x_n - b_2y &= 0\\
a_{31}x_1 + a_{32}x_2 + a_{33}x_3 + \cdots + a_{3n}x_n - b_3y &= 0\\
&\vdots\\
a_{n1}x_1 + a_{n2}x_2 + a_{n3}x_3 + \cdots + a_{nn}x_n - b_ny &= 0
\end{aligned}
\tag{$**$}
$$

Since this is a homogenous system with more variables than equations ($m = n + 1 > n$), Theorem HMVEI [44] says that the system has infinitely many solutions. We will choose one of these solutions, *any* one of these solutions, so long as it is *not* the trivial solution. Write this solution as

$$x_1 = c_1 \qquad x_2 = c_2 \qquad x_3 = c_3 \qquad \dots \qquad x_n = c_n \qquad y = c_{n+1}$$

We know that at least one value of the c_i is nonzero, but we will now show that in particular $c_{n+1} \neq 0$. We do this using a proof by contradiction. So suppose the c_i form a solution as described, and in addition that $c_{n+1} = 0$. Then we can write the i-th equation of system $(**)$ as,

$$a_{i1}c_1 + a_{i2}c_2 + a_{i3}c_3 + \cdots + a_{in}c_n - b_i(0) = 0$$

which becomes

$$a_{i1}c_1 + a_{i2}c_2 + a_{i3}c_3 + \cdots + a_{in}c_n = 0$$

Since this is true for each i, we have that $x_1 = c_1$, $x_2 = c_2$, $x_3 = c_3, \dots, x_n = c_n$ is a solution to the homogenous system $\mathcal{LS}(A, \mathbf{0})$ formed with a nonsingular coefficient matrix. This means that the only possible solution is the trivial solution, so $c_1 = 0$, $c_2 = 0$, $c_3 = 0$, $\dots$, $c_n = 0$. So, assuming simply that $c_{n+1} = 0$, we conclude that *all* of the c_i are zero. But this contradicts our choice of the c_i as not being the trivial solution to the system $(**)$. So $c_{n+1} \neq 0$.

We now propose and verify a solution to the original system $(*)$. Set

$$x_1 = \frac{c_1}{c_{n+1}} \qquad x_2 = \frac{c_2}{c_{n+1}} \qquad x_3 = \frac{c_3}{c_{n+1}} \qquad \dots \qquad x_n = \frac{c_n}{c_{n+1}}$$

Notice how it was necessary that we know that $c_{n+1} \neq 0$ for this step to succeed. Now, evaluate the i-th equation of system $(*)$ with this proposed solution,

$$\begin{aligned}
&a_{i1}\frac{c_1}{c_{n+1}} + a_{i2}\frac{c_2}{c_{n+1}} + a_{i3}\frac{c_3}{c_{n+1}} + \cdots + a_{in}\frac{c_n}{c_{n+1}} = \\
&\frac{1}{c_{n+1}}\left(a_{i1}c_1 + a_{i2}c_2 + a_{i3}c_3 + \cdots + a_{in}c_n\right) = \\
&\frac{1}{c_{n+1}}\left(c_{n+1}b_i\right) = b_i
\end{aligned}$$

Since this equation is true for every i, we have found a solution to the system. To finish, we still need to establish that this solution is *unique*.

With one solution in hand, we will entertain the possibility of a second solution. So assume system $(*)$ has two solutions,

$$\begin{aligned}
x_1 &= d_1 & x_2 &= d_2 & x_3 &= d_3 & &\dots & x_n &= d_n \\
x_1 &= e_1 & x_2 &= e_2 & x_3 &= e_3 & &\dots & x_n &= e_n
\end{aligned}$$

Then,

$$\begin{aligned}
&(a_{i1}(d_1 - e_1) + a_{i2}(d_2 - e_2) + a_{i3}(d_3 - e_3) + \cdots + a_{in}(d_n - e_n)) = \\
&(a_{i1}d_1 + a_{i2}d_2 + a_{i3}d_3 + \cdots + a_{in}d_n) - (a_{i1}e_1 + a_{i2}e_2 + a_{i3}e_3 + \cdots + a_{in}e_n) = \\
&b_i - b_i = 0
\end{aligned}$$

This is the i-th equation of the homogenous system $\mathcal{LS}(A,\,\mathbf{0})$ evaluted with $x_j = d_j - e_j$, $1 \leq j \leq n$. Since A is nonsingular, we must conclude that this solution is the trivial solution, and so $0 = d_j - e_j$, $1 \leq j \leq n$. That is, $d_j = e_j$ for all j and the two solutions are identical, meaning any solution to $(*)$ is unique. ■

This important theorem deserves several comments. First, notice that the proposed solution ($x_i = \frac{c_i}{c_{n+1}}$) appeared in the proof with no motivation whatsoever. This is just fine in a proof. A proof should *convince* you that a theorem is *true*. It is your job to *read* the proof and be convinced of every assertion. Questions like "Where did that come from?" or "How would I think of that?" have no bearing on the validity of the proof.

Second, this theorem helps to explain part of our interest in nonsingular matrices. If a matrix is nonsingular, then no matter what vector of constants we pair it with, using the matrix as the coefficient matrix will *always* yield a linear system of equations with a solution, and the solution is unique. To determine if a matrix has this property (nonsingularity) it is enough to just solve one linear system, the homogenous system with the matrix as coefficient matrix and the zero vector as the vector of constants.

Finally, formulating the negation of the second part of this theorem is a good exercise. A singular matrix has the property that for *some* value of the vector $\mathbf{b}$, the system $\mathcal{LS}(A,\,\mathbf{b})$ does not have a unique solution (which means that it has no solution or infinitely many solutions). We will be able to say more about this case later.

Proof Technique ME
Multiple Equivalences

A very specialized form of a theorem begins with the statement "The following are equivalent…" and then follows a list of statements. Informally, this lead-in sometimes gets abbreviated by "TFAE." This formulation means that any two of the statements on the list can be connected with an "if and only if" to form a theorem. So if the list has n statements then there are $\frac{n(n-1)}{2}$ possible equivalences that can be constructed (and are claimed to be true).

Suppose a theorem of this form has statements denoted as A, B, C,… Z. To prove the entire theorem, we can prove $A \Rightarrow B$, $B \Rightarrow C$, $C \Rightarrow D$,…, $Y \Rightarrow Z$ and finally, $Z \Rightarrow A$. This circular chain of n equivalences would allow us, logically, if not practically, to form any one of the $\frac{n(n-1)}{2}$ possible equivalences by chasing the equivalences around the circle as far as required. ◊

Square matrices that are nonsingular have a long list of interesting properties, which we will start to catalog in the following, recurring, theorem. Of course, singular matrices will have all of the opposite properties.

Theorem NSME1
NonSingular Matrix Equivalences, Round 1
Suppose that A is a square matrix. The following are equivalent.

1. A is nonsingular.

2. A row-reduces to the identity matrix.

3. The null space of A contains only the trivial solution, $\mathcal{N}(A) = \{\mathbf{0}\}$.

4. The linear system $\mathcal{LS}(A,\,\mathbf{b})$ has a unique solution for every possible choice of $\mathbf{b}$. □

Proof That A is nonsingular is equivalent to each of the subsequent statements by, in turn, Theorem NSRRI [50], Theorem NSTNS [52] and Theorem NSMUS [53]. So the statement of this theorem is just a convenient way to organize all these results. ■

V: Vectors

Section VO
Vector Operations

We have worked extensively in the last chapter with matrices, and some with vectors. In this chapter we will develop the properties of vectors, making stronger connections with systems of equations, while preparing to study vector spaces. In this section we define some new operations involving vectors, and collect some basic properties of these operations. Begin by recalling our definition of a column vector as a matrix with just one column (Definition CV [44]). The collection of all possible vectors of a fixed size is a commonly used set, so we start with its definition.

Definition VSCM
Vector Space $\mathbb{C}^m$
The vector space $\mathbb{C}^m$ is the set of all column vectors of size m with entries from the set of complex numbers. $\odot$

When this set is defined using only entries from the real numbers, it is written as $\mathbb{R}^m$ and is known as **Euclidean m-space**.

The term "vector" is used in a variety of different ways. We have defined it as a matrix with a single column. It could simply be an ordered list of numbers, and written like $(2, 3, -1, 6)$. Or it could be interpreted as a point in m dimensions, such as $(3, 4, -2)$ representing a point in three dimensions relative to x, y and z axes. With an interpretation as a point, we can construct an arrow from the origin to the point which is consistent with the notion that a vector has direction and magnitude.

All of these ideas can be shown to be related and equivalent, so keep that in mind as you connect the ideas of this course with ideas from other disciplines. For now, we'll stick with the idea that a vector is a matrix with just one column, or even more simply, just a list of numbers, in some order.

Subsection VEASM
Vector equality, addition, scalar multiplication

We start our study of this set by first defining what it means for two vectors to be the same.

Definition CVE
Column Vector Equality
The vectors

$$\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ \vdots \\ u_m \end{bmatrix} \qquad \mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ \vdots \\ v_m \end{bmatrix}$$

are **equal**, written $\mathbf{u} = \mathbf{v}$ provided that $u_i = v_i$ for all $1 \leq i \leq m$. ⊙

Now this may seem like a silly (or even stupid) thing to say so carefully. Of course two vectors are equal if they are equal for each corresponding entry! Well, this is not as silly as it appears. We will see a few ocassions later where the obvious definition is *not* the right one. And besides, in doing mathematics we need to be very careful about making all the necessary definitions and making them unambiguous. And we've done that here.

Notice now that the symbol '=' is now doing triple-duty. We know from our earlier education what it means for two numbers (real or complex) to be equal, and we take this for granted. Earlier, in Technique SE [13] we discussed at some length what it meant for two sets to be equal. Now we have defined what it means for two vectors to be equal, and that definition builds on our definition for when two numbers are equal when we use the condition $u_i = v_i$ for all $1 \leq i \leq m$. So think carefully about your objects when you see an equal sign and think about just which notion of equality you have encountered. This will be especially important when you are asked to construct proofs whose conclusion states that two objects are equal.

OK, lets do an example of vector equality that begins to hint at the utility of this definition.

Example VO.VESE
Vector equality for a system of equations
Consider the system of simultaneous linear equations in Archetype B [171],

$$\begin{aligned} -7x_1 - 6x_2 - 12x_3 &= -33 \\ 5x_1 + 5x_2 + 7x_3 &= 24 \\ x_1 + 4x_3 &= 5 \end{aligned}$$

Note the use of three equals signs — each indicates an equality of numbers (the linear expressions are numbers when we evaluate them with fixed values of the variable quantities). Now write the vector equality,

$$\begin{bmatrix}-7x_1 - 6x_2 - 12x_3\\ 5x_1 + 5x_2 + 7x_3\\ x_1 + 4x_3\end{bmatrix} = \begin{bmatrix}-33\\ 24\\ 5\end{bmatrix}.$$

By Definition CVE [58], this *single* equality (of two column vectors) translates into *three* simultaneous equalities of numbers that form the system of equations. So with this new notion of vector equality we can become less reliant on referring to *systems* of *simultaneous* equations. There's more to vector equality than just this, but this is a good example for starters and we will develop it further. △

We will now define two operations on the set $\mathbb{C}^m$. By this we mean well-defined procedures that somehow convert vectors into other vectors. Here are two of the most basic definitions of the entire course.

Definition CVA
Column Vector Addition
Given the vectors

$$\mathbf{u} = \begin{bmatrix}u_1\\ u_2\\ u_3\\ \vdots\\ u_m\end{bmatrix} \qquad \mathbf{v} = \begin{bmatrix}v_1\\ v_2\\ v_3\\ \vdots\\ v_m\end{bmatrix}$$

the **sum** of **u** and **v** is the vector

$$\mathbf{u} + \mathbf{v} = \begin{bmatrix}u_1\\ u_2\\ u_3\\ \vdots\\ u_m\end{bmatrix} + \begin{bmatrix}v_1\\ v_2\\ v_3\\ \vdots\\ v_m\end{bmatrix} = \begin{bmatrix}u_1 + v_1\\ u_2 + v_2\\ u_3 + v_3\\ \vdots\\ u_m + v_m\end{bmatrix}.$$ ⊙

So vector addition takes two vectors of the same size and combines them (in a natural way!) to create a new vector of the same size. Notice that this definition is required, even if we agree that this is the obvious, right, natural or correct way to do it. Notice too that the symbol '+' is being recycled. We all know how to add *numbers*, but now we have the same symbol extended to double-duty and we use it to indicate how to add two new objects, vectors. And this definition of our new meaning is built on our previous meaning of addition via the expressions $u_i + v_i$. Think about your objects, especially when doing proofs. Vector addition is easy, here's an example from $\mathbb{C}^4$.

Example VO.VA
Addition of two vectors in $\mathbb{C}^4$
If

$$\mathbf{u} = \begin{bmatrix} 2 \\ -3 \\ 4 \\ 2 \end{bmatrix} \qquad \mathbf{v} = \begin{bmatrix} -1 \\ 5 \\ 2 \\ -7 \end{bmatrix}$$

then

$$\mathbf{u} + \mathbf{v} = \begin{bmatrix} 2 \\ -3 \\ 4 \\ 2 \end{bmatrix} + \begin{bmatrix} -1 \\ 5 \\ 2 \\ -7 \end{bmatrix} = \begin{bmatrix} 2 + (-1) \\ -3 + 5 \\ 4 + 2 \\ 2 + (-7) \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 6 \\ -5 \end{bmatrix}. \qquad \triangle$$

Our second operation takes two objects of different types, specifically a number and a vector, and combines them to create another vector. In this context we call a number a **scalar** in order to emphasize that it is not a vector.

Definition CVSM
Column Vector Scalar Multiplication
Given the vector

$$\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ \vdots \\ u_m \end{bmatrix}$$

and the scalar $\alpha \in \mathbb{C}$, the **scalar multiple** of $\mathbf{u}$ by α is

$$\alpha\mathbf{u} = \alpha \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ \vdots \\ u_m \end{bmatrix} = \begin{bmatrix} \alpha u_1 \\ \alpha u_2 \\ \alpha u_3 \\ \vdots \\ \alpha u_m \end{bmatrix}. \qquad \odot$$

Notice that we are doing a kind of multiplication here, but we are *defining* a new type, perhaps in what appears to be a natural way. We use concatenation (smashing two symbols together side-by-side) to denote this operation rather than using a symbol like we did with vector addition. So this can be another source of confusion. When two symbols are next to each other, are we doing regular old multiplication, the kind we've done for years, or are we doing scalar vector multiplication, the operation we just defined? Think about your objects — if the first object is a scalar, and the second is a vector, then it *must* be that we are doing our new operation, and the *result* of this operation will be another vector.

Notice how consistentcy in notation can be an aid here. If we write scalars as lower case Greek letters from the start of the alphabet (such as α, β, ...) and write vectors in bold Latin letters from the end of the alphabet ($\mathbf{u}$, $\mathbf{v}$, $\ldots$), then we have some hints about what type of objects we are working with. This can be a blessing *and* a curse, since when we go read another book about linear algebra, or read an application in another discipline (physics, economics, $\ldots$) the types of notation employed may be very different and hence unfamiliar.

Again, computationally, vector scalar multiplication is very easy.

Example VO.VSM
Scalar multiplication in $\mathbb{C}^5$
If

$$\mathbf{u} = \begin{bmatrix} 3 \\ 1 \\ -2 \\ 4 \\ -1 \end{bmatrix}$$

and $\alpha = 6$, then

$$\alpha\mathbf{u} = 6 \begin{bmatrix} 3 \\ 1 \\ -2 \\ 4 \\ -1 \end{bmatrix} = \begin{bmatrix} 6(3) \\ 6(1) \\ 6(-2) \\ 6(4) \\ 6(-1 \end{bmatrix} = \begin{bmatrix} 18 \\ 6 \\ -12 \\ 24 \\ -6 \end{bmatrix}.$$

△

Subsection VSP
Vector Space Properties

With definitions of vector addition and scalar multiplication we can state, and prove, several properties of each operation, and some properties that involve their interplay. We now collect ten of them here for later reference.

Theorem VSPCM
Vector Space Properties of $\mathbb{C}^m$
Supppose that $\mathbf{u}$, $\mathbf{v}$ and $\mathbf{w}$ are vectors in $\mathbb{C}^m$ and α and β are scalars. Then

1. $\mathbf{u} + \mathbf{v} \in \mathbb{C}^m$ (Additive closure)
2. $\alpha\mathbf{u} \in \mathbb{C}^m$ (Scalar closure)
3. $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$ (Commutativity)
4. $\mathbf{u} + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w}$ (Associativity)
5. There is a vector, $\mathbf{0}$, called the **zero vector**, such that $\mathbf{u} + \mathbf{0} = \mathbf{u}$ for all $\mathbf{u} \in \mathbb{C}^m$. (Additive identity)

6. For each vector $\mathbf{u} \in \mathbb{C}^m$, there exists a vector $-\mathbf{u} \in \mathbb{C}^m$ so that $\mathbf{u} + (-\mathbf{u}) = \mathbf{0}$. (Additive inverses)
7. $\alpha(\beta\mathbf{u}) = (\alpha\beta)\mathbf{u}$ (Associativity)
8. $\alpha(\mathbf{u} + \mathbf{v}) = \alpha\mathbf{u} + \alpha\mathbf{v}$ (Distributivity)
9. $(\alpha + \beta)\mathbf{u} = \alpha\mathbf{u} + \beta\mathbf{u}$ (Distributivity)
10. $1\mathbf{u} = \mathbf{u}$ □

Proof While some of these properties seem very obvious, they all require proof. However, the proofs are not very interesting, and border on tedious. We'll prove one version of distributivity very carefully, and you can test your proof-building skills on some of the others.

$$\begin{aligned}
(\alpha+\beta)\mathbf{u} &= (\alpha+\beta)\begin{bmatrix}u_1\\u_2\\u_3\\\vdots\\u_m\end{bmatrix} = \begin{bmatrix}(\alpha+\beta)u_1\\(\alpha+\beta)u_2\\(\alpha+\beta)u_3\\\vdots\\(\alpha+\beta)u_m\end{bmatrix} = \begin{bmatrix}\alpha u_1+\beta u_1\\\alpha u_2+\beta u_2\\\alpha u_3+\beta u_3\\\vdots\\\alpha u_m+\beta u_m\end{bmatrix}\\
&= \begin{bmatrix}\alpha u_1\\\alpha u_2\\\alpha u_3\\\vdots\\\alpha u_m\end{bmatrix} + \begin{bmatrix}\beta u_1\\u\beta u_2\\\beta u_3\\\vdots\\\beta u_m\end{bmatrix} = \alpha\begin{bmatrix}u_1\\u_2\\u_3\\\vdots\\u_m\end{bmatrix} + \beta\begin{bmatrix}u_1\\u_2\\u_3\\\vdots\\u_m\end{bmatrix}\\
&= \alpha\mathbf{u} + \beta\mathbf{u}
\end{aligned}$$

■

Be careful with the notion of the vector $-\mathbf{u}$. This is a vector that we add to $\mathbf{u}$ so that the result is the particular vector $\mathbf{0}$. This is basically a property of vector addition. It happens that we can compute $-\mathbf{u}$ using the *other* operation, scalar multiplication. We can prove this directly by writing that

$$-\mathbf{u} = \begin{bmatrix}-u_1\\-u_2\\-u_3\\\vdots\\-u_m\end{bmatrix} = (-1)\begin{bmatrix}u_1\\u_2\\u_3\\\vdots\\u_m\end{bmatrix} = (-1)\mathbf{u}$$

We will see later how to derive this property as a *consequence* of several of the ten properties listed in Theorem VSPCM [61].

Section LC
Linear Combinations

Subsection LC
Linear Combinations

In Section VO [57] we defined vector addition and scalar multiplication. These two operations combine nicely to give us a construction known as a linear combination, a construct that we will work with throughout this course.

Definition LCCV
Linear Combination of Column Vectors
Given n vectors $\mathbf{u}_1$, $\mathbf{u}_2$, $\mathbf{u}_3$, $\ldots$, $\mathbf{u}_n$ and n scalars α_1, α_2, α_3, $\ldots$, α_n, their **linear combination** is the vector

$$\alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \alpha_3\mathbf{u}_3 + \cdots + \alpha_n\mathbf{u}_n. \qquad \odot$$

So this definition takes an equal number of scalars and vectors, combines them using our two new operations (scalar multiplication and vector addition) and creates a single brand-new vector, of the same size as the original vectors. When a definition or theorem employs a linear combination, think about the nature of the objects that go into its creation (lists of scalars and vectors), and the type of object that results (a single vector). Computationally, a linear combination is pretty easy.

Example LC.TLC
Two linear combinations in $\mathbb{C}^6$
Suppose that

$$\alpha_1 = 3 \qquad \alpha_2 = -4 \qquad \alpha_3 = 2 \qquad \alpha_4 = -1$$

and

$$\mathbf{u}_1 = \begin{bmatrix} 2 \\ 4 \\ -3 \\ 1 \\ 2 \\ 9 \end{bmatrix} \qquad \mathbf{u}_2 = \begin{bmatrix} 6 \\ 3 \\ 0 \\ -2 \\ 1 \\ 4 \end{bmatrix} \qquad \mathbf{u}_3 = \begin{bmatrix} -5 \\ 2 \\ 1 \\ 1 \\ -3 \\ 0 \end{bmatrix} \qquad \mathbf{u}_4 = \begin{bmatrix} 3 \\ 2 \\ -5 \\ 7 \\ 1 \\ 3 \end{bmatrix}$$

then their linear combination is

$$\alpha_1\mathbf{u_1} + \alpha_2\mathbf{u_2} + \alpha_3\mathbf{u_3} + \alpha_4\mathbf{u_4} = (1)\begin{bmatrix}2\\4\\-3\\1\\2\\9\end{bmatrix} + (-4)\begin{bmatrix}6\\3\\0\\-2\\1\\4\end{bmatrix} + (2)\begin{bmatrix}-5\\2\\1\\1\\-3\\0\end{bmatrix} + (-1)\begin{bmatrix}3\\2\\-5\\7\\1\\3\end{bmatrix}$$

$$= \begin{bmatrix}2\\4\\-3\\1\\2\\9\end{bmatrix} + \begin{bmatrix}-24\\-12\\0\\8\\-4\\-16\end{bmatrix} + \begin{bmatrix}-10\\4\\2\\2\\-6\\0\end{bmatrix} + \begin{bmatrix}-3\\-2\\5\\-7\\-1\\-3\end{bmatrix} = \begin{bmatrix}-35\\-6\\4\\4\\-9\\-8\end{bmatrix}.$$

A different linear combination, of the same set of vectors, can be formed with different scalars. Take

$$\beta_1 = 3 \qquad \beta_2 = 0 \qquad \beta_3 = 5 \qquad \beta_4 = -6$$

and form the linear combination

$$\beta_1\mathbf{u_1} + \beta_2\mathbf{u_2} + \beta_3\mathbf{u_3} + \beta_4\mathbf{u_4} = (3)\begin{bmatrix}2\\4\\-3\\1\\2\\9\end{bmatrix} + (0)\begin{bmatrix}6\\3\\0\\-2\\1\\4\end{bmatrix} + (5)\begin{bmatrix}-5\\2\\1\\1\\-3\\0\end{bmatrix} + (-6)\begin{bmatrix}3\\2\\-5\\7\\1\\3\end{bmatrix}$$

$$= \begin{bmatrix}6\\12\\-9\\3\\6\\27\end{bmatrix} + \begin{bmatrix}0\\0\\0\\0\\0\\0\end{bmatrix} + \begin{bmatrix}-25\\10\\5\\5\\-15\\0\end{bmatrix} + \begin{bmatrix}-16\\-2\\5\\-7\\-1\\-3\end{bmatrix} = \begin{bmatrix}-35\\20\\1\\1\\-10\\24\end{bmatrix}.$$

Notice how we could keep our set of vectors fixed, and use different sets of scalars to construct different vectors. You might build a few new linear combinations of $\mathbf{u}_1$, $\mathbf{u}_2$, $\mathbf{u}_3$, $\mathbf{u}_4$ right now. We'll be right here when you get back. What vectors were you able to create? Do you think you could create the vector

$$\mathbf{w} = \begin{bmatrix}13\\15\\5\\-17\\2\\25\end{bmatrix}$$

with a "suitable" choice of four scalars? Do you think you could create *any* possible vector from $\mathbb{C}^6$ by choosing the proper scalars? These last two questions are very fundamental, and time spent considering them *now* will prove beneficial later. △

Proof Technique DC
Decompositions
Much of your mathematical upbringing, especially once you began a study of algebra, revolved around simplifying expressions — combining like terms, obtaining common denominators so as to add fractions, factoring in order to solve polynomial equations. However, as often as not, we will do the opposite. Many theorems and techniques will revolve around taking some object and "decomposing" it into some combination of other objects, ostensibly in a more complicated fashion. When we say something can "be written as" something else, we mean that the one object can be decomposed into some combination of other objects. This may seem unnatural at first, but results of this type will give us insight into the structure of the original object by exposing its building blocks. ◊

Example LC.ABLC
Archetype B as a linear combination
In this example we will rewrite Archetype B [171] in the language of vectors, vector equality and linear combinations. In Example VO.VESE [58] we wrote the simultaneous system of $m = 3$ equations as the vector equality

$$\begin{bmatrix} -7x_1 - 6x_2 - 12x_3 \\ 5x_1 + 5x_2 + 7x_3 \\ x_1 + 4x_3 \end{bmatrix} = \begin{bmatrix} -33 \\ 24 \\ 5 \end{bmatrix}.$$

Now we will bust up the linear expressions on the left, first using vector addition,

$$\begin{bmatrix} -7x_1 \\ 5x_1 \\ x_1 \end{bmatrix} + \begin{bmatrix} -6x_2 \\ 5x_2 \\ 0x_2 \end{bmatrix} + \begin{bmatrix} -12x_3 \\ 7x_3 \\ 4x_3 \end{bmatrix} = \begin{bmatrix} -33 \\ 24 \\ 5 \end{bmatrix}.$$

Now we can rewrite each of these $n = 3$ vectors as a scalar multiple of a fixed vector, where the scalar is one of the unknown variables, converting the left-hand side into a linear combination

$$x_1 \begin{bmatrix} -7 \\ 5 \\ 1 \end{bmatrix} + x_2 \begin{bmatrix} -6 \\ 5 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -12 \\ 7 \\ 4 \end{bmatrix} = \begin{bmatrix} -33 \\ 24 \\ 5 \end{bmatrix}.$$

We can now interpret the problem of solving the system of equations as determining values for the scalar multiples that make the vector equation true. In the analysis of Archetype A [167], we were able to determine that it had only one solution. A quick way to see this is to row-reduce the coefficient matrix to the 3×3 identity matrix and apply Theorem NSRRI [50] to determine that the coefficient matrix is nonsingular. Then Theorem NSMUS [53] tells us that the system of equations has a unique solution.

This solution is

$$x_1 = -3 \qquad x_2 = 5 \qquad x_3 = 2.$$

So, in the context of this example, we can express the fact that these values of the variables are a solution by writing the linear combination,

$$(-3)\begin{bmatrix}-7\\5\\1\end{bmatrix} + (5)\begin{bmatrix}-6\\5\\0\end{bmatrix} + (2)\begin{bmatrix}-12\\7\\4\end{bmatrix} = \begin{bmatrix}-33\\24\\5\end{bmatrix}.$$

Furthermore, these are the only three scalars that will accomplish this equality, since they come from a unique solution.

Notice how the three vectors in this example are the columns of the coefficient matrix of the system of equations. This is our first hint of the important interplay between the vectors that form the columns of a matrix, and the matrix itself. △

With any discussion of Archetype A [167] or Archetype B [171] we should be sure to contrast with the other.

Example LC.AALC
Archetype A as a linear combination
As a vector equality, Archetype A [167] can be written as

$$\begin{bmatrix}x_1 - x_2 + 2x_3\\2x_1 + x_2 + x_3\\x_1 + x_2\end{bmatrix} = \begin{bmatrix}1\\8\\5\end{bmatrix}.$$

Now bust up the linear expressions on the left, first using vector addition,

$$\begin{bmatrix}x_1\\2x_1\\x_1\end{bmatrix} + \begin{bmatrix}-x_2\\x_2\\x_2\end{bmatrix} + \begin{bmatrix}2x_3\\x_3\\0x_3\end{bmatrix} = \begin{bmatrix}1\\8\\5\end{bmatrix}.$$

Rewrite each of these $n = 3$ vectors as a scalar multiple of a fixed vector, where the scalar is one of the unknown variables, converting the left-hand side into a linear combination

$$x_1\begin{bmatrix}1\\2\\1\end{bmatrix} + x_2\begin{bmatrix}-1\\1\\1\end{bmatrix} + x_3\begin{bmatrix}2\\1\\0\end{bmatrix} = \begin{bmatrix}1\\8\\5\end{bmatrix}.$$

Row-reducing the augmented matrix for Archetype A [167] leads to the conclusion that the system is consistent and has free variables, hence infinitely many solutions. So for example, the two solutions

$$x_1 = 2 \qquad x_2 = 3 \qquad x_3 = 1$$
$$x_1 = 3 \qquad x_2 = 2 \qquad x_3 = 0$$

can be used together to say that,

$$(2)\begin{bmatrix}1\\2\\1\end{bmatrix}+(3)\begin{bmatrix}-1\\1\\1\end{bmatrix}+(1)\begin{bmatrix}2\\1\\0\end{bmatrix}=\begin{bmatrix}1\\8\\5\end{bmatrix}=(3)\begin{bmatrix}1\\2\\1\end{bmatrix}+(2)\begin{bmatrix}-1\\1\\1\end{bmatrix}+(0)\begin{bmatrix}2\\1\\0\end{bmatrix}$$

Ignore the middle of this equation, and move all the terms to the left-hand side,

$$(2)\begin{bmatrix}1\\2\\1\end{bmatrix}+(3)\begin{bmatrix}-1\\1\\1\end{bmatrix}+(1)\begin{bmatrix}2\\1\\0\end{bmatrix}+(-3)\begin{bmatrix}1\\2\\1\end{bmatrix}+(-2)\begin{bmatrix}-1\\1\\1\end{bmatrix}+(-0)\begin{bmatrix}2\\1\\0\end{bmatrix}=\begin{bmatrix}0\\0\\0\end{bmatrix}.$$

Regrouping gives

$$(-1)\begin{bmatrix}1\\2\\1\end{bmatrix}+(1)\begin{bmatrix}-1\\1\\1\end{bmatrix}+(1)\begin{bmatrix}2\\1\\0\end{bmatrix}=\begin{bmatrix}0\\0\\0\end{bmatrix}.$$

Notice that these three vectors are the columns of the coefficient matrix for the system of equations in Archetype A [167]. This equality says there is a linear combination of those columns that equals the vector of all zeros. Give it some thought, but this says that

$$x_1 = -1 \qquad x_2 = 1 \qquad x_3 = 1$$

is a nontrivial solution to the homogenous system of equations with the coefficient matrix for the original system in Archetype A [167]. In particular, this demonstrates that this coefficient matrix is singular. △

There's a lot going on in the last two examples. Come back to them in a while and make some connections with the intervening material. For now, we will summarize and explain some of this behavior with a theorem.

Theorem SLSLC
Solutions to Linear Systems are Linear Combinations
Denote the columns of the $m \times n$ matrix A as the vectors $\mathbf{A}_1, \mathbf{A}_2, \mathbf{A}_3, \ldots, \mathbf{A}_n$. Then $\mathbf{x} = \begin{bmatrix}\alpha_1\\\alpha_2\\\alpha_3\\\vdots\\\alpha_n\end{bmatrix}$ is a solution to the linear system of equations $\mathcal{LS}(A, \mathbf{b})$ if and only if

$$\alpha_1\mathbf{A}_1 + \alpha_2\mathbf{A}_2 + \alpha_3\mathbf{A}_3 + \cdots + \alpha_n\mathbf{A}_n = \mathbf{b}$$

□

Proof Write the system of equations $\mathcal{LS}(A, b)$ as

$$\begin{aligned}
a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \cdots + a_{1n}x_n &= b_1\\
a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \cdots + a_{2n}x_n &= b_2\\
a_{31}x_1 + a_{32}x_2 + a_{33}x_3 + \cdots + a_{3n}x_n &= b_3\\
&\vdots\\
a_{m1}x_1 + a_{m2}x_2 + a_{m3}x_3 + \cdots + a_{mn}x_n &= b_m.
\end{aligned}$$

Now use vector equality (DEFINITION CVE [58]) to replace the m simultaneous equalities by one vector equality,

$$\begin{bmatrix} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \cdots + a_{2n}x_n \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 + \cdots + a_{3n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + a_{m3}x_3 + \cdots + a_{mn}x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_m \end{bmatrix}.$$

Use vector addition (DEFINITION CVA [59]) to rewrite,

$$\begin{bmatrix} a_{11}x_1 \\ a_{21}x_1 \\ a_{31}x_1 \\ \vdots \\ a_{m1}x_1 \end{bmatrix} + \begin{bmatrix} a_{12}x_2 \\ a_{22}x_2 \\ a_{32}x_2 \\ \vdots \\ a_{m2}x_2 \end{bmatrix} + \begin{bmatrix} a_{13}x_3 \\ a_{23}x_3 \\ a_{33}x_3 \\ \vdots \\ a_{m3}x_3 \end{bmatrix} + \cdots + \begin{bmatrix} a_{1n}x_n \\ a_{2n}x_n \\ a_{3n}x_n \\ \vdots \\ a_{mn}x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_m \end{bmatrix}.$$

And finally, use the definition of vector scalar multiplication DEFINITION CVSM [60],

$$x_1 \begin{bmatrix} a_{11} \\ a_{21} \\ a_{31} \\ \vdots \\ a_{m1} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ a_{22} \\ a_{32} \\ \vdots \\ a_{m2} \end{bmatrix} + x_3 \begin{bmatrix} a_{13} \\ a_{23} \\ a_{33} \\ \vdots \\ a_{m3} \end{bmatrix} + \cdots + x_m \begin{bmatrix} a_{1n} \\ a_{2n} \\ a_{3n} \\ \vdots \\ a_{mn} \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_m \end{bmatrix}.$$

and use notation for the various column vectors,

$$x_1\mathbf{A}_1 + x_2\mathbf{A}_2 + x_3\mathbf{A}_3 + \cdots + x_n\mathbf{A}_n = \mathbf{b}.$$

Each of the expressions above is just a rewrite of another one. So if we begin with a solution to the system of equations, substituting its values into the original system will make the equations simultaneously true. But then these same values will also make the final expression with the linear combination true. Reversing the argument, and employing the equations in reverse, will give the other half of the proof. ■

In other words, this theorem tells us that solutions to systems of equations are linear combinations of the column vectors of the coefficient matrix ($\mathbf{A}_i$) which yield the constant vector $\mathbf{b}$. Look through the archetypes that are systems of equations and examine a few of the advertised solutions. In each case use the solution to form a linear combination of the columns of the coefficient matrix and verify that the result equals the constant vector.

Section SS
Spanning Sets

In this section we will describe a compact way to indicate the elements of an infinite set of vectors, making use of linear combinations. This will give us a nice way to describe the elements of a set of solutions to a linear system, or the elements of the null space of a matrix.

Subsection VFSS
Vector Form of Solution Sets

We have recently begun writing solutions to systems of equations as column vectors. For example Archetype B [171] has the solution $x_1 = -3$, $x_2 = 5$, $x_3 = 2$ which we now write as

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -3 \\ 5 \\ 2 \end{bmatrix}.$$

Now, we will use column vectors and linear combinations to express *all* of the solutions to a linear system of equations in a compact and understandable way. First, here's an example that will motivate our next theorem. This is a valuable technique, almost the equal of row-reducing a matrix, so be sure you get comfortable with it over the course of this section.

Example SS.VFSAD
Vector form of solutions for Archetype D
Archetype D [179] is a linear system of 3 equations in 4 variables. Row-reducing the augmented matrix yields

$$\begin{bmatrix} \boxed{1} & 0 & 3 & -2 & 4 \\ 0 & \boxed{1} & 1 & -3 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

and we see $r = 2$ nonzero rows. Also, $D = \{1,\, 2\}$ so the dependent variables are then x_1 and x_2. $F = \{3,\, 4,\, 5\}$ so the two free variables are x_3 and x_4. We will develop a linear combination that expresses a typical solution, in three steps.

Step 1. Write the vector of variables as a fixed vector, plus a linear combination of

$n - r$ vectors, using the free variables as the scalars.

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} \\ \\ \\ \\ \end{bmatrix} + x_3 \begin{bmatrix} \\ \\ \\ \\ \end{bmatrix} + x_4 \begin{bmatrix} \\ \\ \\ \\ \end{bmatrix}$$

Step 2. For each free variable, use 0's and 1's to ensure equality for the corresponding entry of the the vectors.

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} \\ \\ 0 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} \\ \\ 1 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} \\ \\ 0 \\ 1 \end{bmatrix}$$

Step 3. For each dependent variable, use the augmented matrix to formulate an equation expressing the dependent variable as a constant plus multiples of the free variables. Convert this equation into entries of the vectors that ensure equality for each dependent variable, one at a time.

$$x_1 = 4 - 3x_3 + 2x_4 \quad \Rightarrow \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 4 \\ \\ 0 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -3 \\ \\ 1 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} 2 \\ \\ 0 \\ 1 \end{bmatrix}$$

$$x_2 = 0 - 1x_3 + 3x_4 \quad \Rightarrow \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 4 \\ 0 \\ 0 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -3 \\ -1 \\ 1 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} 2 \\ 3 \\ 0 \\ 1 \end{bmatrix}$$

This final *form* of a typical solution is especially pleasing and useful. For example, we can build solutions quickly by choosing values for our free variables, and then compute a linear combination. Such as

$$x_3 = 2,\ x_4 = -5 \quad \Rightarrow \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 4 \\ 0 \\ 0 \\ 0 \end{bmatrix} + (2) \begin{bmatrix} -3 \\ -1 \\ 1 \\ 0 \end{bmatrix} + (-5) \begin{bmatrix} 2 \\ 3 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -12 \\ -17 \\ 2 \\ -5 \end{bmatrix}$$

or,

$$x_3 = 1,\ x_4 = 3 \quad \Rightarrow \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 4 \\ 0 \\ 0 \\ 0 \end{bmatrix} + (1) \begin{bmatrix} -3 \\ -1 \\ 1 \\ 0 \end{bmatrix} + (3) \begin{bmatrix} 2 \\ 3 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 7 \\ 8 \\ 1 \\ 3 \end{bmatrix}.$$

You'll find the second solution listed in the write-up for Archetype D [179], and you might check the first solution by substituting it back into the original equations.

While this form is useful for quickly creating solutions, its even better because it tells us *exactly* what every solution looks like. We know the solution set is infinite, which is pretty big, but now we can say that a solution is some multiple of $\begin{bmatrix}-3\\-1\\1\\0\end{bmatrix}$ plus a multiple of $\begin{bmatrix}2\\3\\0\\1\end{bmatrix}$ plus the fixed vector $\begin{bmatrix}4\\0\\0\\0\end{bmatrix}$. Period. So it only takes us *three* vectors to describe the entire infinite solution set, provided we also agree on how to combine the three vectors into a linear combination. △

We'll now formalize the last example as a theorem.

Theorem VFSLS
Vector Form of Solutions to Linear Systems
Suppose that $[A|\mathbf{b}]$ is the augmented matrix for a consistent linear system $\mathcal{LS}(A,\,\mathbf{b})$ of m equations in n variables. Denote the vector of variables as $\mathbf{x} = (x_i)$. Let $B = (b_{ij})$ be a row-equivalent $m\times(n+1)$ matrix in reduced row-echelon form. Suppose that B has r nonzero rows, columns without leading 1's having indices $F = \{f_1,\, f_2,\, f_3,\, \ldots,\, f_{n-r},\, n+1\}$, and columns with leading 1's having indices $D = \{d_1,\, d_2,\, d_3,\, \ldots,\, d_r\}$. Define vectors $\mathbf{c} = (c_i)$, $\mathbf{u}_j = (u_{ij})$, $1 \leq j \leq n-r$ of size n by

$$c_i = \begin{cases} 0 & \text{if } i \in F \\ b_{k,n+1} & \text{if } i \in D,\ i = d_k \end{cases}$$

$$u_{ij} = \begin{cases} 1 & \text{if } i \in F,\ i = f_j \\ 0 & \text{if } i \in F,\ i \neq f_j \\ -b_{k,f_j} & \text{if } i \in D,\ i = d_k \end{cases}.$$

Then the system of equations represented by the vector equation

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix} = \mathbf{c} + x_{f_1}\mathbf{u}_1 + x_{f_2}\mathbf{u}_2 + x_{f_3}\mathbf{u}_3 + \cdots + x_{f_{n-r}}\mathbf{u}_{n-r}$$

is equivalent to $\mathcal{LS}(A,\,\mathbf{b})$. □

Proof We are being asked to prove that two systems of equations are equivalent, that is, they have identical solution sets. First, $\mathcal{LS}(A,\,\mathbf{b})$ is equivalent to the linear system of equations that has the matrix B as its augmented matrix (Theorem REMES [23]). We will now show that the equations in the conclusion of the proof are either always true, or are simple rearrangements of the equations in the system with B as its augmented matrix. This will then establish that all three systems are equivalent to each other.

Suppose that $i \in F$, so $i = f_k$ for a particular choice of k, $1 \leq k \leq n - r$. Consider the equation given by entry i of the vector equality.

$$\begin{aligned}
x_i =& c_i + x_{f_1}u_{i1} + x_{f_2}u_{i2} + x_{f_3}u_{i3} + \cdots + x_{f_{n-r}}u_{i,n-r}\\
x_i =& c_{f_k} + x_{f_1}u_{f_k1} + x_{f_2}u_{f_k2} + x_{f_3}u_{f_k3} + \cdots + x_{f_k}u_{f_kf_k} + \cdots + x_{f_{n-r}}u_{f_k,n-r}\\
x_i =& 0 + x_{f_1}(0) + x_{f_2}(0) + x_{f_3}(0) + \cdots + x_{f_k}(1) + \cdots + x_{f_{n-r}}(0) = x_{f_k}\\
x_i =& x_i.
\end{aligned}$$

This means that equality of the two vectors in entry i represents the equation $x_i = x_i$ when $i \in F$. Since this equation is always true, it does not restrict the possibilities for the solution set.

Now consider the i-th entry, when $i \in D$, and suppose that $i = d_k$, for some particular choice of k, $1 \leq k \leq r$. Consider the equation given by entry i of the vector equality.

$$\begin{aligned}
x_i =& c_i + x_{f_1}u_{i1} + x_{f_2}u_{i2} + x_{f_3}u_{i3} + \cdots + x_{f_{n-r}}u_{i,n-r}\\
x_i =& c_{d_k} + x_{f_1}u_{d_k1} + x_{f_2}u_{d_k2} + x_{f_3}u_{d_k3} + \cdots + x_{f_{n-r}}u_{d_k,n-r}\\
x_i =& b_{k,n+1} + x_{f_1}(-b_{k,f_1}) + x_{f_2}(-b_{k,f_2}) + x_{f_3}(-b_{k,f_3}) + \cdots + x_{f_{n-r}}(-b_{k,f_{n-r}})\\
x_i =& b_{k,n+1} - \left(b_{k,f_1}x_{f_1} + b_{k,f_2}x_{f_2} + b_{k,f_3}x_{f_3} + \cdots + b_{k,f_{n-r}}x_{f_{n-r}}\right)
\end{aligned}$$

Rearranging, this becomes,

$$x_i + b_{k,f_1}x_{f_1} + b_{k,f_2}x_{f_2} + b_{k,f_3}x_{f_3} + \cdots + b_{k,f_{n-r}}x_{f_{n-r}} = b_{k,n+1}.$$

This is exactly the equation represented by row k of the matrix B. So the equations represented by the vector equality in the conclusion are exactly the equations represented by the matrix B, along with additional equations that are always true. So the solution sets will be identical. ■

Theorem VFSLS [72] formalizes what happened in the three steps of Example SS.VFSAD [70]. The theorem will be useful in proving other theorems, and it it is useful since it tells us an exact procedure for simply describing an infinite solution set. We could program a computer to implement it, once we have the augmented matrix row-reduced and have checked that the system is consistent. By Knuth's definition, this completes our conversion of linear equation solving from art into science. Notice that it even applies (but is overkill) in the case of a unique solution. However, as a practical matter, I prefer the three-step process of Example SS.VFSAD [70] when I need to describe an infinite solution set. So lets practice some more, but with a bigger example.

Example SS.VFSAI
Vector form of solutions for Archetype I

Archetype I [201] is a linear system of $m = 4$ equations in $n = 7$ variables. Row-reducing the augmented matrix yields

$$\begin{bmatrix} \boxed{1} & 4 & 0 & 0 & 2 & 1 & -3 & 4 \\ 0 & 0 & \boxed{1} & 0 & 1 & -3 & 5 & 2 \\ 0 & 0 & 0 & \boxed{1} & 2 & -6 & 6 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

and we see $r = 3$ nonzero rows. The columns with leading 1's are $D = \{1, 3, 4\}$ so the r dependent variables are x_1, x_3, x_4. The columns without leading 1's are $F = \{2, 5, 6, 7, 8\}$, so the $n - r = 4$ free variables are x_2, x_5, x_6, x_7.

Step 1. Write the vector of variables ($\mathbf{x}$) as a fixed vector ($\mathbf{c}$), plus a linear combination of $n - r = 4$ vectors ($\mathbf{u}_1$, $\mathbf{u}_2$, $\mathbf{u}_3$, $\mathbf{u}_4$), using the free variables as the scalars.

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \end{bmatrix} = \begin{bmatrix} \\ \\ \\ \\ \\ \\ \\ \end{bmatrix} + x_2 \begin{bmatrix} \\ \\ \\ \\ \\ \\ \\ \end{bmatrix} + x_5 \begin{bmatrix} \\ \\ \\ \\ \\ \\ \\ \end{bmatrix} + x_6 \begin{bmatrix} \\ \\ \\ \\ \\ \\ \\ \end{bmatrix} + x_7 \begin{bmatrix} \\ \\ \\ \\ \\ \\ \\ \end{bmatrix}$$

Step 2. For each free variable, use 0's and 1's to ensure equality for the corresponding entry of the the vectors. Take note of the pattern of 0's and 1's at this stage, because this is the best look you'll have at it. We'll state an important theorem in the next section and the proof will essentially rely on this observation.

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \end{bmatrix} = \begin{bmatrix} \\ 0 \\ \\ \\ 0 \\ 0 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} \\ 1 \\ \\ \\ 0 \\ 0 \\ 0 \end{bmatrix} + x_5 \begin{bmatrix} \\ 0 \\ \\ \\ 1 \\ 0 \\ 0 \end{bmatrix} + x_6 \begin{bmatrix} \\ 0 \\ \\ \\ 0 \\ 1 \\ 0 \end{bmatrix} + x_7 \begin{bmatrix} \\ 0 \\ \\ \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

Step 3. For each dependent variable, use the augmented matrix to formulate an equation expressing the dependent variable as a constant plus multiples of the free variables. Convert this equation into entries of the vectors that ensure equality for each dependent variable, one at a time.

$$x_1 = 2 - 4x_2 - 2x_5 - 1x_6 + 3x_7 \quad \Rightarrow$$

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \end{bmatrix} = \begin{bmatrix} 4 \\ 0 \\ \\ \\ 0 \\ 0 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} -4 \\ 1 \\ \\ \\ 0 \\ 0 \\ 0 \end{bmatrix} + x_5 \begin{bmatrix} -2 \\ 0 \\ \\ \\ 1 \\ 0 \\ 0 \end{bmatrix} + x_6 \begin{bmatrix} -1 \\ 0 \\ \\ \\ 0 \\ 1 \\ 0 \end{bmatrix} + x_7 \begin{bmatrix} 3 \\ 0 \\ \\ \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

$$x_3 = 2 + 0x_2 + 3x_5 - 5x_6 - 2x_7 \quad \Rightarrow$$

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \end{bmatrix} = \begin{bmatrix} 4 \\ 0 \\ 2 \\ \\ 0 \\ 0 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} -4 \\ 1 \\ 0 \\ \\ 0 \\ 0 \\ 0 \end{bmatrix} + x_5 \begin{bmatrix} -2 \\ 0 \\ 3 \\ \\ 1 \\ 0 \\ 0 \end{bmatrix} + x_6 \begin{bmatrix} -1 \\ 0 \\ -5 \\ \\ 0 \\ 1 \\ 0 \end{bmatrix} + x_7 \begin{bmatrix} 3 \\ 0 \\ -2 \\ \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

$$x_4 = 1 + 0x_2 - 2x_5 + 6x_6 - 6x_7 \quad \Rightarrow$$

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \end{bmatrix} = \begin{bmatrix} 4 \\ 0 \\ 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} -4 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} + x_5 \begin{bmatrix} -2 \\ 0 \\ 3 \\ -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + x_6 \begin{bmatrix} -1 \\ 0 \\ -5 \\ 6 \\ 0 \\ 1 \\ 0 \end{bmatrix} + x_7 \begin{bmatrix} 3 \\ 0 \\ -2 \\ -6 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

We can now use this final expression to quickly build solutions to the system. You might try to recreate each of the solutions listed in the write-up for ARCHETYPE I [201]. (Hint: look at the values of the free variables in each solution, and notice that the vector **c** has 0's in these locations.)

Even better, we have a description of the infinite solution set, based on just 5 vectors, which we combine in linear combinations to produce solutions.

Whenever we discuss ARCHETYPE I [201] you know that's your cue to go work through ARCHETYPE J [206] by yourself. Remember to take note of the 0/1 pattern at the conclusion of Step 2. Have fun — we won't go anywhere while you're away. △

This technique is so important, that we'll do one more example. However, an important distinction will be that this system is homogenous.

Example SS.VFSAL
Vector form of solutions for Archetype L
Archetype L [214] is presented simply as the 5×5 matrix

$$L = \begin{bmatrix} -2 & -1 & -2 & -4 & 4 \\ -6 & -5 & -4 & -4 & 6 \\ 10 & 7 & 7 & 10 & -13 \\ -7 & -5 & -6 & -9 & 10 \\ -4 & -3 & -4 & -6 & 6 \end{bmatrix}$$

We'll interpret it here as the coefficient matrix of a homogenous system and reference this matrix as L. So we are solving the homogenous system $\mathcal{LS}(L, \mathbf{0})$ having $m = 5$ equations in $n = 5$ variables. If we built the augmented matrix, we would a sixth column to L containing all zeros. As we did row operations, this sixth column would remain all zeros. So instead we will row-reduce the coefficient matrix, and mentally remember the missing sixth column of zeros. This row-reduced matrix is

$$\begin{bmatrix} \boxed{1} & 0 & 0 & 1 & -2 \\ 0 & \boxed{1} & 0 & -2 & 2 \\ 0 & 0 & \boxed{1} & 2 & -1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

and we see $r = 3$ nonzero rows. The columns with leading 1's are $D = \{1,\, 2,\, 3\}$ so the r dependent variables are $x_1,\, x_2,\, x_3$. The columns without leading 1's are $F = \{4,\, 5\}$, so the $n - r = 2$ free variables are $x_4,\, x_5$. Notice that if we had included the all-zero vector of constants to form the augmented matrix for the system, then the index 6 would have appeared in the set F, and subsequently would have been ignored when listing the free variables.

Step 1. Write the vector of variables ($\mathbf{x}$) as a fixed vector ($\mathbf{c}$), plus a linear combination of $n - r = 2$ vectors ($\mathbf{u}_1,\, \mathbf{u}_2$), using the free variables as the scalars.

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} \\ \\ \\ \\ \\ \end{bmatrix} + x_4 \begin{bmatrix} \\ \\ \\ \\ \\ \end{bmatrix} + x_5 \begin{bmatrix} \\ \\ \\ \\ \\ \end{bmatrix}$$

Step 2. For each free variable, use 0's and 1's to ensure equality for the corresponding entry of the the vectors. Take note of the pattern of 0's and 1's at this stage, even if it is not as illuminating as in other examples.

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} \\ \\ \\ 0 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} \\ \\ \\ 1 \\ 0 \end{bmatrix} + x_5 \begin{bmatrix} \\ \\ \\ 0 \\ 1 \end{bmatrix}$$

Step 3. For each dependent variable, use the augmented matrix to formulate an equation expressing the dependent variable as a constant plus multiples of the free variables. Don't forget about the "missing" sixth column being full of zeros. Convert this equation into entries of the vectors that ensure equality for each dependent variable, one at a time.

$$x_1 = 0 - 1x_4 + 2x_5 \quad \Rightarrow \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 0 \\ \\ \\ 0 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} -1 \\ \\ \\ 1 \\ 0 \end{bmatrix} + x_5 \begin{bmatrix} 2 \\ \\ \\ 0 \\ 1 \end{bmatrix}$$

$$x_2 = 0 + 2x_4 - 2x_5 \quad \Rightarrow \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \\ 0 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} -1 \\ 2 \\ \\ 1 \\ 0 \end{bmatrix} + x_5 \begin{bmatrix} 2 \\ -2 \\ \\ 0 \\ 1 \end{bmatrix}$$

$$x_3 = 0 - 2x_4 + 1x_5 \quad \Rightarrow \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} -1 \\ 2 \\ -2 \\ 1 \\ 0 \end{bmatrix} + x_5 \begin{bmatrix} 2 \\ -2 \\ 1 \\ 0 \\ 1 \end{bmatrix}$$

The vector $\mathbf{c}$ will always have 0's in the entries corresponding to free variables. However, since we are solving a homogenous system, the row-reduced augmented matrix has zeros in column $n + 1 = 6$, and hence *all* the entries of $\mathbf{c}$ are zero. So we can write

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \mathbf{0} + x_4 \begin{bmatrix} -1 \\ 2 \\ -2 \\ 1 \\ 0 \end{bmatrix} + x_5 \begin{bmatrix} 2 \\ -2 \\ 1 \\ 0 \\ 1 \end{bmatrix} = x_4 \begin{bmatrix} -1 \\ 2 \\ -2 \\ 1 \\ 0 \end{bmatrix} + x_5 \begin{bmatrix} 2 \\ -2 \\ 1 \\ 0 \\ 1 \end{bmatrix}$$

It will always happen that the solutions to a homogenous system has $\mathbf{c} = \mathbf{0}$ (even in the case of a unique solution?). So our expression for the solutions is a bit more pleasing. In this example it says that the solutions are *all possible* linear combinations of the two vectors $\mathbf{u}_1 = \begin{bmatrix} -1 \\ 2 \\ -2 \\ 1 \\ 0 \end{bmatrix}$ and $\mathbf{u}_2 = \begin{bmatrix} 2 \\ -2 \\ 1 \\ 0 \\ 1 \end{bmatrix}$, with no mention of any fixed vector entering into the linear combination.

This observation will motivate our next subsection and definition, and after that we'll conclude the section by formalizing this situation. △

Subsection SSV
Span of a Set of Vectors

In Example SS.VFSAL [75] we saw the solution set of a homogenous system described as all possible linear combinations of two particular vectors. This happens to be a useful way to construct or describe infinite sets of vectors, so we encapsulate this idea in a definition.

Definition SSV
Span of a set of Vectors
Given a set of vectors $S = \{\mathbf{u}_1,\,\mathbf{u}_2,\,\mathbf{u}_3,\,\ldots,\,\mathbf{u}_t\}$, their **span**, $\mathrm{Sp}\,(S)$, is the set of all possible linear combinations of $\mathbf{u}_1,\,\mathbf{u}_2,\,\mathbf{u}_3,\,\ldots,\,\mathbf{u}_t$. Symbolically,

$$\begin{aligned}\mathrm{Sp}\,(S) &= \{\alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \alpha_3\mathbf{u}_3 + \cdots + \alpha_t\mathbf{u}_t \mid \alpha_i \in \mathbb{C},\, 1 \leq i \leq t\} \\ &= \left\{\sum_{i=1}^{t} \alpha_i\mathbf{u}_i \mid \alpha_i \in \mathbb{C},\, 1 \leq i \leq t\right\}\end{aligned}$$ ⊙

The span is just a set of vectors, though in all but one situation it is an infinite set. (Just when is it not infinite?) So we start with a finite collection of vectors (t of them to be precise), and use this finite set to describe an infinite set of vectors. We will see this construction repeatedly, so lets work through some examples to get comfortable with it. The most obvious question about a set is if a particular item of the correct type is in the set, or not.

Example SS.SCAA
Span of the columns of Archetype A
Begin with the finite set of three vectors of size 3

$$S = \{\mathbf{u}_1,\,\mathbf{u}_2,\,\mathbf{u}_3\} = \left\{\begin{bmatrix}1\\2\\1\end{bmatrix},\,\begin{bmatrix}-1\\1\\1\end{bmatrix},\,\begin{bmatrix}2\\1\\0\end{bmatrix}\right\}$$

and consider the infinite set $U = \mathrm{Sp}\,(S)$. The vectors of S could have been chosen to be anything, but for reasons that will become clear later, we have chosen the three columns of the coefficient matrix in Archetype A [167]. First, as an example, note that

$$\mathbf{v} = (5)\begin{bmatrix}1\\2\\1\end{bmatrix} + (-3)\begin{bmatrix}-1\\1\\1\end{bmatrix} + (7)\begin{bmatrix}2\\1\\0\end{bmatrix} = \begin{bmatrix}22\\14\\2\end{bmatrix}$$

is in $\mathrm{Sp}\,(S)$, since it is a linear combination of $\mathbf{u}_1,\,\mathbf{u}_2,\,\mathbf{u}_3$. We write this succinctly as $\mathbf{v} \in \mathrm{Sp}\,(S)$. There is nothing magical about the scalars $\alpha_1 = 5$, $\alpha_2 = -3$, $\alpha_3 = 7$, they could have been chosen to be anything. So repeat this part of the example yourself, using different values of α_1, α_2, α_3. What happens if you choose all three scalars to be zero?

So we know how to quickly construct sample elements of the set $\mathrm{Sp}(S)$. A slightly different question arises when you are handed a vector of the correct size and asked if it is an element of $\mathrm{Sp}(S)$. For example, is $\mathbf{w} = \begin{bmatrix} 1 \\ 8 \\ 5 \end{bmatrix}$ in $\mathrm{Sp}(S)$? More succinctly, $\mathbf{w} \in \mathrm{Sp}(S)$?

To answer this question, we will look for scalars α_1, α_2, α_3 so that

$$\alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \alpha_3\mathbf{u}_3 = \mathbf{w}.$$

By Theorem SLSLC [67] this linear combination becomes the system of equations

$$\begin{aligned} \alpha_1 - \alpha_2 + 2\alpha_3 &= 1 \\ 2\alpha_1 + \alpha_2 + \alpha_3 &= 8 \\ \alpha_1 + \alpha_2 &= 5. \end{aligned}$$

Building the augmented matrix for this linear system, and row-reducing, gives

$$\begin{bmatrix} \boxed{1} & 0 & 1 & 3 \\ 0 & \boxed{1} & -1 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

This system has infinitely many solutions (there's a free variable), but all we need is one. The solution,

$$\alpha_1 = 2 \qquad \alpha_2 = 3 \qquad \alpha_3 = 1$$

tells us that

$$(2)\mathbf{u}_1 + (3)\mathbf{u}_2 + (1)\mathbf{u}_3 = \mathbf{w}$$

so we are convinced that $\mathbf{w}$ really is in $\mathrm{Sp}(S)$. Lets ask the same question again, but this time with $\mathbf{y} = \begin{bmatrix} 2 \\ 4 \\ 3 \end{bmatrix}$, i.e. is $\mathbf{y} \in \mathrm{Sp}(S)$?

So we'll look for scalars α_1, α_2, α_3 so that

$$\alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \alpha_3\mathbf{u}_3 = \mathbf{y}.$$

By Theorem SLSLC [67] this linear combination becomes the system of equations

$$\begin{aligned} \alpha_1 - \alpha_2 + 2\alpha_3 &= 2 \\ 2\alpha_1 + \alpha_2 + \alpha_3 &= 4 \\ \alpha_1 + \alpha_2 &= 3. \end{aligned}$$

Building the augmented matrix for this linear system, and row-reducing, gives

$$\begin{bmatrix} \boxed{1} & 0 & 1 & 0 \\ 0 & \boxed{1} & -1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

This system inconsistent (there's a leading 1 in the last column), so there are no scalars α_1, α_2, α_3 that will create a linear combination of $\mathbf{u}_1$, $\mathbf{u}_2$, $\mathbf{u}_3$ that equals $\mathbf{y}$. More precisely, $\mathbf{y} \notin \mathrm{Sp}\,(S)$.

There are three things to observe in this example. (1) It is easy to construct vectors in $\mathrm{Sp}\,(S)$. (2) It is possible that some vectors are in $\mathrm{Sp}\,(S)$ (e.g. $\mathbf{w}$), while others are not (e.g. $\mathbf{y}$). (3) Deciding if a given vector is in $\mathrm{Sp}\,(S)$ leads to solving a linear system of equations.

With a computer progam in hand to solve systems of linear equations, could you create a program to decide if a vector was, or wasn't, in the span of a given set of vectors? Is this art or science?

This example was built on vectors from the columns of the coefficient matrix of Archetype A [167]. Study the determination that $\mathbf{v} \in \mathrm{Sp}\,(S)$ and see if you can connect it with some of the other properties of Archetype A [167]. △

Lets do a similar example to Example SS.SCAA [78], only now with Archetype B [171].

Example SS.SCAB
Span of the columns of Archetype B

Begin with the finite set of three vectors of size 3

$$R = \{\mathbf{v}_1,\, \mathbf{v}_2,\, \mathbf{v}_3\} = \left\{ \begin{bmatrix} -7 \\ 5 \\ 1 \end{bmatrix}, \begin{bmatrix} -6 \\ 5 \\ 0 \end{bmatrix}, \begin{bmatrix} -12 \\ 7 \\ 4 \end{bmatrix} \right\}$$

and consider the infinite set $V = \mathrm{Sp}\,(R)$. The vectors of R have been chosen as the three columns of the coefficient matrix in Archetype B [171]. First, as an example, note that

$$\mathbf{x} = (2) \begin{bmatrix} -7 \\ 5 \\ 1 \end{bmatrix} + (4) \begin{bmatrix} -6 \\ 5 \\ 0 \end{bmatrix} + (-3) \begin{bmatrix} -12 \\ 7 \\ 4 \end{bmatrix} = \begin{bmatrix} -2 \\ 9 \\ -10 \end{bmatrix}$$

is in $\mathrm{Sp}\,(R)$, since it is a linear combination of $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$. In other words, $\mathbf{x} \in \mathrm{Sp}\,(R)$. Try some different values of α_1, α_2, α_3 yourself, and see what vectors you can create as elements of $\mathrm{Sp}\,(R)$.

Now ask if a given vector is an element of $\mathrm{Sp}\,(R)$. For example, is $\mathbf{z} = \begin{bmatrix} -33 \\ 24 \\ 5 \end{bmatrix}$ in $\mathrm{Sp}\,(S)$? Is $\mathbf{z} \in \mathrm{Sp}\,(R)$?

To answer this question, we will look for scalars α_1, α_2, α_3 so that

$$\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \alpha_3\mathbf{v}_3 = \mathbf{z}.$$

By Theorem SLSLC [67] this linear combination becomes the system of equations

$$\begin{aligned} -7\alpha_1 - 6\alpha_2 - 12\alpha_3 &= -33 \\ 5\alpha_1 + 5\alpha_2 + 7\alpha_3 &= 24 \\ \alpha_1 + 4\alpha_3 &= 5. \end{aligned}$$

Building the augmented matrix for this linear system, and row-reducing, gives

$$\begin{bmatrix} \boxed{1} & 0 & 0 & -3 \\ 0 & \boxed{1} & 0 & 5 \\ 0 & 0 & \boxed{1} & 2 \end{bmatrix}.$$

This system has a unique solution,

$$\alpha_1 = -3 \qquad \alpha_2 = 5 \qquad \alpha_3 = 2$$

telling us that

$$(-3)\mathbf{v}_1 + (5)\mathbf{v}_2 + (2)\mathbf{v}_3 = \mathbf{z}$$

so we are convinced that $\mathbf{z}$ really is in $\mathrm{Sp}\left(R\right)$.

There is no point in replacing $\mathbf{z}$ with another vector and doing this again. A question about membership in $\mathrm{Sp}\left(R\right)$ inevitably leads to a system of three equations in the three variables α_1, α_2, α_3 with a coefficient matrix whose columns are the vectors $\mathbf{v}_1$, $\mathbf{v}_2$, $\mathbf{v}_3$. This particular coefficient matrix is nonsingular, so by Theorem NSMUS [53], it is guaranteed to have a solution. (This solution is unique, but that's not important here.) So *no matter* which vector we might have chosen for $\mathbf{z}$, we would have been *certain* to discover that it was an element of $\mathrm{Sp}\left(R\right)$. Stated differently, every vector of size 3 is in $\mathrm{Sp}\left(R\right)$, or $\mathrm{Sp}\left(R\right) = \mathbb{C}^3$.

Compare this example with Example SS.SCAA [78], and see if you can connect $\mathbf{z}$ with some aspects of the write-up for Archetype B [171]. △

Subsection SSNS
Spanning Sets of Null Spaces

We saw in Example SS.VFSAL [75] that when a system of equations is homogenous the solution set can be expressed in the form described by Theorem VFSLS [72] where the vector $\mathbf{c}$ is the zero vector. We can essentially ignore this vector, so that the remainder of the typical expression for a solution looks like an arbitrary linear combination, where the scalars are the free variables and the vectors are $\mathbf{u}_1$, $\mathbf{u}_2$, $\mathbf{u}_3$, $\ldots$, $\mathbf{u}_{n-r}$. Which sounds a lot like a span. This is the substance of the next theorem.

Theorem SSNS
Spanning Sets for Null Spaces
Suppose that A is an $m \times n$ matrix, and B is a row-equivalent matrix in reduced row-echelon form with r nonzero rows. Let $D = \{d_1, d_2, d_3, \ldots, d_r\}$ and $F = \{f_1, f_2, f_3, \ldots, f_{n-r}\}$ be the sets of column indices where B does and does not (respectively) have leading 1's.

Construct the $n-r$ vectors $\mathbf{u}_j = (u_{ij})$, $1 \leq j \leq n-r$ of size n as

$$u_{ij} = \begin{cases} 1 & \text{if } i \in F,\ i = f_j \\ 0 & \text{if } i \in F,\ i \neq f_j \\ -b_{k,f_j} & \text{if } i \in D,\ i = d_k \end{cases} .$$

Then the null space of A is given by

$$\mathrm{N}(A) = \mathrm{Sp}\left(\{\mathbf{u}_1,\, \mathbf{u}_2,\, \mathbf{u}_3,\, \ldots,\, \mathbf{u}_{n-r}\}\right). \qquad \square$$

Proof Consider the homogenous system with A as a coefficient matrix, LS(A, $\mathbf{0}$)$\mathbf{0}$. Its set of solutions is, by definition, the null space of A, $N(A)$. Row-reducing the augmented matrix of this homogenous system will create the row-equivalent matrix B'. Row-reducing the augmented matrix that has a final column of all zeros, yields B', which is the matrix B, along with an additional column (index $n+1$) that is still totally zero.

Now apply Theorem VFSLS [72], noting that our homogenous system is consistent (Theorem HSC [42]). The vector $\mathbf{c}$ has zeros for each entry that corresponds to an index in F. For entries that correspond to an index in D, the value is $-b'_{k,n+1}$, but for B' these entries in the final column are all zero. So $\mathbf{c} = \mathbf{0}$. This says that a solution of the homogenous system is of the form

$$\mathbf{x} = \mathbf{c} + x_{f_1}\mathbf{u}_1 + x_{f_2}\mathbf{u}_2 + x_{f_3}\mathbf{u}_3 + \cdots + x_{f_{n-r}}\mathbf{u}_{n-r} = x_{f_1}\mathbf{u}_1 + x_{f_2}\mathbf{u}_2 + x_{f_3}\mathbf{u}_3 + \cdots + x_{f_{n-r}}\mathbf{u}_{n-r}$$

where the free variables x_{f_j} can each take on any value. Rephrased this says

$$\begin{aligned} \mathrm{N}(A) &= \left\{x_{f_1}\mathbf{u}_1 + x_{f_2}\mathbf{u}_2 + x_{f_3}\mathbf{u}_3 + \cdots + x_{f_{n-r}}\mathbf{u}_{n-r} \mid x_{f_1},\, x_{f_2},\, x_{f_3},\, \ldots,\, x_{f_{n-r}} \in \mathbb{C}\right\} \\ &= \mathrm{Sp}\left(\{\mathbf{u}_1,\, \mathbf{u}_2,\, \mathbf{u}_3,\, \ldots,\, \mathbf{u}_{n-r}\}\right). \end{aligned} \qquad \blacksquare$$

Here's an example that will exercise the span construction and Theorem SSNS [81], while also pointing the way to the next section.

Example SS.SCAD
Span of the columns of Archetype D
Begin with the set of four vectors of size 3

$$T = \{\mathbf{w}_1,\, \mathbf{w}_2,\, \mathbf{w}_3,\, \mathbf{w}_4\} = \left\{ \begin{bmatrix} 2 \\ -3 \\ 1 \end{bmatrix},\, \begin{bmatrix} 1 \\ 4 \\ 1 \end{bmatrix},\, \begin{bmatrix} 7 \\ -5 \\ 4 \end{bmatrix},\, \begin{bmatrix} -7 \\ -6 \\ -5 \end{bmatrix} \right\}$$

and consider the infinite set $W = \mathrm{Sp}(T)$. The vectors of T have been chosen as the four columns of the coefficient matrix in Archetype D [179]. Check that the vector

$$\mathbf{u}_2 = \begin{bmatrix} 2 \\ 3 \\ 0 \\ 1 \end{bmatrix}$$

is a solution to the homogenous system $\mathcal{LS}(D,\,\mathbf{0})$ (it is the second vector of the spanning set for the null space of the coefficient matrix D, as described in Theorem SSNS [81]). Applying Theorem SLSLC [67], we can write the linear combination,

$$2\mathbf{w}_1 + 3\mathbf{w}_2 + 0\mathbf{w}_3 + 1\mathbf{w}_4 = \mathbf{0}$$

which we can solve for $\mathbf{w}_4$,

$$\mathbf{w}_4 = (-2)\mathbf{w}_1 + (-3)\mathbf{w}_2.$$

This equation says that whenever we encounter the vector $\mathbf{w}_4$, we can replace it with a specific linear combination of the vectors $\mathbf{w}_1$ and $\mathbf{w}_2$. So using $\mathbf{w}_4$ in the set T, along with $\mathbf{w}_1$ and $\mathbf{w}_2$, is excessive. An example of what we mean here can be illustrated by the computation,

$$\begin{aligned}
5\mathbf{w}_1 + (-4)\mathbf{w}_2 + 6\mathbf{w}_3 + (-3)\mathbf{w}_4 &= 5\mathbf{w}_1 + (-4)\mathbf{w}_2 + 6\mathbf{w}_3 + (-3)\left((-2)\mathbf{w}_1 + (-3)\mathbf{w}_2\right)\\
&= 5\mathbf{w}_1 + (-4)\mathbf{w}_2 + 6\mathbf{w}_3 + (6\mathbf{w}_1 + 9\mathbf{w}_2)\\
&= 11\mathbf{w}_1 + 5\mathbf{w}_2 + 6\mathbf{w}_3.
\end{aligned}$$

So what began as a linear combination of the vectors $\mathbf{w}_1,\,\mathbf{w}_2,\,\mathbf{w}_3,\,\mathbf{w}_4$ has been reduced to a linear combination of the vectors $\mathbf{w}_1,\,\mathbf{w}_2,\,\mathbf{w}_3$. A careful proof using our definition of set equality (Technique SE [13]) would now allow us to conclude that this reduction is possible for any vector in W, so

$$W = \mathrm{Sp}\left(\{\mathbf{w}_1,\,\mathbf{w}_2,\,\mathbf{w}_3\}\right).$$

So the span of our set of vectors, W, has not changed, but we have *described* it by the span of a set of *three* vectors, rather than *four*. Furthermore, we can achieve yet another, similar, reduction.

Check that the vector

$$\mathbf{u}_1 = \begin{bmatrix}-3\\-1\\1\\0\end{bmatrix}$$

is a solution to the homogenous system $\mathcal{LS}(D,\,\mathbf{0})$ (it is the first vector of the spanning set for the null space of the coefficient matrix D, as described in Theorem SSNS [81]). Applying Theorem SLSLC [67], we can write the linear combination,

$$(-3)\mathbf{w}_1 + (-1)\mathbf{w}_2 + 1\mathbf{w}_3 = \mathbf{0}$$

which we can solve for $\mathbf{w}_3$,

$$\mathbf{w}_3 = 3\mathbf{w}_1 + 1\mathbf{w}_2.$$

This equation says that whenever we encounter the vector $\mathbf{w}_3$, we can replace it with a specific linear combination of the vectors $\mathbf{w}_1$ and $\mathbf{w}_2$. So, as before, the vector $\mathbf{w}_3$ is not needed in the description of W, provided we have $\mathbf{w}_1$ and $\mathbf{w}_2$ available. In particular, a careful proof would show that

$$W = \mathrm{Sp}\left(\{\mathbf{w}_1,\, \mathbf{w}_2\}\right).$$

So W began life as the span of a set of four vectors, and we have now shown (utilizing solutions to a homogenous system) that W can also be described as the span of a set of just two vectors. Convince yourself that we cannot go any further. In other words, it is not possible to dismiss either $\mathbf{w}_1$ or $\mathbf{w}_2$ in a similar fashion and winnow the set down to just one vector.

What was it about the original set of four vectors that allowed us to declare certain vectors as surplus? And just which vectors were we able to dismiss? And why did we have to stop once we had two vectors remaining? The answers to these questions motivate "linear independence," our next section and next definition, and so are worth considering carefully *now*. $\triangle$

Section LI
Linear Independence

Subsection LIV
Linearly Independent Vectors

Theorem SLSLC [67] tells us that a solution to a homogenous system of equations is a linear combination of the columns of the coefficient matrix that equals the zero vector. We used just this situation to our advantage (twice!) in Example SS.SCAD [82] where we reduced the set of vectors used in a span construction from four down to two, by declaring certain vectors as surplus. The next two definitions will allow us to formalize this situation.

Definition RLDCV
Relation of Linear Dependence for Column Vectors
Given a set of vectors $S = \{\mathbf{u}_1,\,\mathbf{u}_2,\,\mathbf{u}_3,\,\ldots,\,\mathbf{u}_n\}$, an equation of the form

$$\alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \alpha_3\mathbf{u}_3 + \cdots + \alpha_n\mathbf{u}_n = \mathbf{0}$$

is a **relation of linear dependence** on S. If this equation is formed in a trivial fashion, i.e. $\alpha_i = 0$, $1 \le i \le n$, then we say it is a **trivial relation of linear dependence** on S. ⊙

Definition LICV
Linear Independence of Column Vectors
The set of vectors $S = \{\mathbf{u}_1,\,\mathbf{u}_2,\,\mathbf{u}_3,\,\ldots,\,\mathbf{u}_n\}$ is **linearly dependent** if there is a relation of linear dependence on S that is not trivial. In the case where the *only* relation of linear dependence on S is the trivial one, then S is a **linearly independent** set of vectors. ⊙

Notice that a relation of linear dependence is an *equation*. Though most of it is a linear combination, it is not a linear combination (that would be a vector). Linear independence is a property of a *set* of vectors. It is easy to take a set of vectors, and an equal number of scalars, *all zero*, and form a linear combination that equals the zero vector. When the easy way is the only way, then we say the set is linearly independent. Here's a couple of examples.

Example LI.LDS
Linearly dependent set in $\mathbb{C}^5$

Consider the set of $n = 4$ vectors from $\mathbb{C}^5$,

$$S = \left\{ \begin{bmatrix} 2 \\ -1 \\ 3 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ -1 \\ 5 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ -3 \\ 6 \\ 1 \end{bmatrix}, \begin{bmatrix} -6 \\ 7 \\ -1 \\ 0 \\ 1 \end{bmatrix} \right\}.$$

To determine linear independence we first form a relation of linear dependence,

$$\alpha_1 \begin{bmatrix} 2 \\ -1 \\ 3 \\ 1 \\ 2 \end{bmatrix} + \alpha_2 \begin{bmatrix} 1 \\ 2 \\ -1 \\ 5 \\ 2 \end{bmatrix} + \alpha_3 \begin{bmatrix} 2 \\ 1 \\ -3 \\ 6 \\ 1 \end{bmatrix} + \alpha_4 \begin{bmatrix} -6 \\ 7 \\ -1 \\ 0 \\ 1 \end{bmatrix} = \mathbf{0}.$$

We know that $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = 0$ is a solution to this equation, but that is of no interest whatsoever. That is *always* the case, no matter what four vectors we might have chosen. We are curious to know if there are other, nontrivial, solutions. Theorem SLSLC [67] tells us that we can find such solutions as solutions to the homogenous system $\mathcal{LS}(A, \mathbf{0})$ where the coefficient matrix has these four vectors as columns,

$$A = \begin{bmatrix} 2 & 1 & 2 & -6 \\ -1 & 2 & 1 & 7 \\ 3 & -1 & -3 & -1 \\ 1 & 5 & 6 & 0 \\ 2 & 2 & 1 & 1 \end{bmatrix}.$$

Row-reducing this coefficient matrix yields,

$$\begin{bmatrix} \boxed{1} & 0 & 0 & -2 \\ 0 & \boxed{1} & 0 & 4 \\ 0 & 0 & \boxed{1} & -3 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

We could solve this homogenous system completely, but for this example all we need is one nontrivial solution. Setting the lone free variable to any nonzero value, such as $x_4 = 1$, yields the nontrivial solution

$$\mathbf{x} = \begin{bmatrix} 2 \\ -4 \\ 3 \\ 1 \end{bmatrix}.$$

completing our application of Theorem SLSLC [67], we have

$$2\begin{bmatrix}2\\-1\\3\\1\\2\end{bmatrix}+(-4)\begin{bmatrix}1\\2\\-1\\5\\2\end{bmatrix}+3\begin{bmatrix}2\\1\\-3\\6\\1\end{bmatrix}+1\begin{bmatrix}-6\\7\\-1\\0\\1\end{bmatrix}=\mathbf{0}.$$

This is a relation of linear dependence on S that is not trivial, so we conclude that S is linearly dependent. △

Example LI.LIS
Linearly independent set in $\mathbb{C}^5$

Consider the set of $n = 4$ vectors from $\mathbb{C}^5$,

$$T=\left\{\begin{bmatrix}2\\-1\\3\\1\\2\end{bmatrix},\begin{bmatrix}1\\2\\-1\\5\\2\end{bmatrix},\begin{bmatrix}2\\1\\-3\\6\\1\end{bmatrix},\begin{bmatrix}-6\\7\\-1\\1\\1\end{bmatrix}\right\}.$$

To determine linear independence we first form a relation of linear dependence,

$$\alpha_1\begin{bmatrix}2\\-1\\3\\1\\2\end{bmatrix}+\alpha_2\begin{bmatrix}1\\2\\-1\\5\\2\end{bmatrix}+\alpha_3\begin{bmatrix}2\\1\\-3\\6\\1\end{bmatrix}+\alpha_4\begin{bmatrix}-6\\7\\-1\\1\\1\end{bmatrix}=\mathbf{0}.$$

We know that $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = 0$ is a solution to this equation, but that is of no interest whatsoever. That is *always* the case, no matter what four vectors we might have chosen. We are curious to know if there are other, nontrivial, solutions. Theorem SLSLC [67] tells us that we can find such solutions as solution to the homogenous system $\mathcal{LS}(B,\,\mathbf{0})$ where the coefficient matrix has these four vectors as columns,

$$B=\begin{bmatrix}2&1&2&-6\\-1&2&1&7\\3&-1&-3&-1\\1&5&6&1\\2&2&1&1\end{bmatrix}.$$

Row-reducing this coefficient matrix yields,

$$\begin{bmatrix}\boxed{1}&0&0&0\\0&\boxed{1}&0&0\\0&0&\boxed{1}&0\\0&0&0&\boxed{1}\\0&0&0&0\end{bmatrix}.$$

From the form of this matrix, we see that there are no free variables, so the solution is unique, and because the system is homogenous, this unique solution is the trivial solution. So we now know that there is but one way to combine the four vectors of T into a relation of linear dependence, and that one way is the easy and obvious way. In this situation we say that the set, T, is linearly independent. △

Example LI.LDS [85] and Example LI.LIS [87] relied on solving a homogenous system of equations to determine linear independence. We can codify this process in a time-saving theorem.

Theorem LIVHS
Linearly Independent Vectors and Homogenous Systems
Suppose that A is an $n \times m$ matrix and $S = \{\mathbf{A}_1,\, \mathbf{A}_2,\, \mathbf{A}_3,\, \ldots,\, \mathbf{A}_n\}$ is the set of vectors in $\mathbb{C}^m$ that are the columns of A. Then S is a linearly independent set if and only if the homogenous sytem $\mathcal{LS}(A,\, \mathbf{0})$ has a unique solution. □

Proof (⇐) Suppose that $\mathcal{LS}(A,\, \mathbf{0})$ has a unique solution. Since it is a homogenous system, this solution must be the trivial solution $\mathbf{x} = \mathbf{0}$. By Theorem SLSLC [67], this means that the only relation of linear dependence on S is the trivial one. So S is linearly independent.

(⇒) We will prove the contrapositive. Suppose that $\mathcal{LS}(A,\, \mathbf{0})$ does not have a unique solution. Since it is a homogenous system, it is consistent (Theorem HSC [42]), and so must have infinitely many solutions (Theorem PSSLS [38]). One of these infinitely many solutions must be nontrivial (in fact, almost all of them are), so choose one. By Theorem SLSLC [67] this nontrivial solution will give a nontrivial relation of linear dependence on S, so we can conclude that S is a linearly dependent set. ■

Since Theorem LIVHS [88] is an equivalence, we can use it to determine the linear independence or dependence of any set of column vectors, just by creating a corresponding matrix and analyzing the row-reduced form. Let's illustrate this with another example.

Example LI.LLDS
Large linearly dependent set in $\mathbb{C}^4$
Consider the set of $n = 9$ vectors from $\mathbb{C}^4$,

$$R = \left\{ \begin{bmatrix}-1\\3\\1\\2\end{bmatrix}, \begin{bmatrix}7\\1\\-3\\6\end{bmatrix}, \begin{bmatrix}1\\2\\-1\\-2\end{bmatrix}, \begin{bmatrix}0\\4\\2\\9\end{bmatrix}, \begin{bmatrix}5\\-2\\4\\3\end{bmatrix}, \begin{bmatrix}2\\1\\-6\\4\end{bmatrix}, \begin{bmatrix}3\\0\\-3\\1\end{bmatrix}, \begin{bmatrix}1\\1\\5\\3\end{bmatrix}, \begin{bmatrix}-6\\-1\\1\\1\end{bmatrix} \right\}.$$

To employ Theorem LIVHS [88], we form a 4×9 coefficient matrix, C,

$$C = \begin{bmatrix} -1 & 7 & 1 & 0 & 5 & 2 & 3 & 1 & -6\\ 3 & 1 & 2 & 4 & -2 & 1 & 0 & 1 & -1\\ 1 & -3 & -1 & 2 & 4 & -6 & -3 & 5 & 1\\ 2 & 6 & -2 & 9 & 3 & 4 & 1 & 3 & 1 \end{bmatrix}.$$

To determine if the homogenous system $\mathcal{LS}(C, \mathbf{0})$ has a unique solution or not, we would normally row-reduce this matrix. But in this particular example, we can do better. Theorem HMVEI [44] tells us that since the system is homogenous with $n = 9$ variables in $m = 4$ equations, and $n > m$, there must be infinitely many solutions. Since there is not a unique solution, Theorem LIVHS [88] says the set is linearly dependent. △

The situation in Example LI.LLDS [88] is slick enough to warrant formulating as a theorem.

Theorem MVSLD
More Vectors than Size implies Linear Dependence
Suppose that $S = \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \ldots, \mathbf{u}_n\}$ is the set of vectors in $\mathbb{C}^m$, and that $n > m$. Then S is a linearly dependent set. □

Proof Form the $m \times n$ coefficient matrix A that has the column vectors $\mathbf{u}_i$, $1 \leq i \leq n$ as its columns. Consider the homogenous system $\mathcal{LS}(A, \mathbf{0})$. By Theorem HMVEI [44] this system has infinitely many solutions. Since the system does not have a unique solution, Theorem LIVHS [88] says the columns of A form a linearly dependent set, which is the desired conclusion. ■

Subsection LDSS
Linearly Dependent Sets and Spans

In any linearly dependent set there is always one vector that can be written as a linear combination of the others. This is the substance of the upcoming Theorem DLDS [89]. Perhaps this will explain the use of the word "dependent." In a linearly dependent set, at least one vector "depends" on the others (via a linear combination).

If we use a linearly dependent set to construct a span, then we can *always* create the same infinite set with a starting set that is one vector smaller in size. We will illustrate this behavior in Example LI.RS [90]. However, this will not be possible if we build a span from a linearly independent set. So in a certain sense, using a linearly independent set to formulate a span is the best possible way to go about it — there aren't any extra vectors being used to build up all the necessary linear combinations. OK, here's the theorem, and then the example.

Theorem DLDS
Dependency in Linearly Dependent Sets
Suppose that $S = \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \ldots, \mathbf{u}_n\}$ is a set of vectors. Then S is a linearly dependent set if and only if there is an index t, $1 \leq t \leq n$ such that $\mathbf{u_t}$ is a linear combination of the vectors $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \ldots, \mathbf{u}_{t-1}, \mathbf{u}_{t+1}, \ldots, \mathbf{u}_n$. □

Proof (⇒) Suppose that S is linearly dependent, so there is a nontrivial relation of linear dependence,

$$\alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \alpha_3\mathbf{u}_3 + \cdots + \alpha_n\mathbf{u}_n = \mathbf{0}.$$

Since the α_i cannot all be zero, choose one, say α_t, that is nonzero. Then,

$$-\alpha_t\mathbf{u}_t = \alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \alpha_3\mathbf{u}_3 + \cdots + \alpha_{t-1}\mathbf{u}_{t-1} + \alpha_{t+1}\mathbf{u}_{t+1} + \cdots + \alpha_n\mathbf{u}_n$$

and we can multiply by $\frac{-1}{\alpha_t}$ since $\alpha_t \neq 0$,

$$\mathbf{u}_t = \frac{-\alpha_1}{\alpha_t}\mathbf{u}_1 + \frac{-\alpha_2}{\alpha_t}\mathbf{u}_2 + \frac{-\alpha_3}{\alpha_t}\mathbf{u}_3 + \cdots + \frac{-\alpha_{t-1}}{\alpha_t}\mathbf{u}_{t-1} + \frac{-\alpha_{t+1}}{\alpha_t}\mathbf{u}_{t+1} + \cdots + \frac{-\alpha_n}{\alpha_t}\mathbf{u}_n.$$

Since the values of $\frac{\alpha_i}{\alpha_t}$ are again scalars, we have expressed $\mathbf{u}_t$ as the desired linear combination.

(⇐) Suppose that the vector $\mathbf{u}_t$ is a linear combination of the other vectors in S. Write this linear combination as

$$\beta_1\mathbf{u}_1 + \beta_2\mathbf{u}_2 + \beta_3\mathbf{u}_3 + \cdots + \beta_{t-1}\mathbf{u}_{t-1} + \beta_{t+1}\mathbf{u}_{t+1} + \cdots + \beta_n\mathbf{u}_n = \mathbf{u_t}$$

and move $\mathbf{u}_t$ to the other side of the equality

$$\beta_1\mathbf{u}_1 + \beta_2\mathbf{u}_2 + \beta_3\mathbf{u}_3 + \cdots + \beta_{t-1}\mathbf{u}_{t-1} + (-1)\mathbf{u}_t + \beta_{t+1}\mathbf{u}_{t+1} + \cdots + \beta_n\mathbf{u}_n = \mathbf{0}.$$

Then the scalars $\beta_1,\,\beta_2,\,\beta_3,\,\ldots,\,\beta_t = -1,\,\ldots,\,\beta_n$ provide a *nontrivial* linear combination of the vectors in S, thus establishing that S is a linearly independent set. ■

This theorem can be used, sometimes repeatedly, to whittle down the size of a set of vectors used in a span construction. We have seen some of this already in Example SS.SCAD [82], but in the next example we will detail some of the subtleties.

Example LI.RS
Reducing a span in $\mathbb{C}^5$

Consider the set of $n = 4$ vectors from $\mathbb{C}^5$,

$$R = \{\mathbf{v}_1,\,\mathbf{v}_2,\,\mathbf{v}_3,\,\mathbf{v}_4\} = \left\{ \begin{bmatrix}1\\2\\-1\\3\\2\end{bmatrix},\ \begin{bmatrix}2\\1\\3\\1\\2\end{bmatrix},\ \begin{bmatrix}0\\-7\\6\\-11\\-2\end{bmatrix},\ \begin{bmatrix}4\\1\\2\\1\\6\end{bmatrix} \right\}$$

and define $V = \mathrm{Sp}(R)$.

To employ Theorem LIVHS [88], we form a 5×4 coefficient matrix, D,

$$D = \begin{bmatrix}1 & 2 & 0 & 4\\ 2 & 1 & -7 & 1\\ -1 & 3 & 6 & 2\\ 3 & 1 & -11 & 1\\ 2 & 2 & -2 & 6\end{bmatrix}$$

and row-reduce to understand solutions to the homogenous system $\mathcal{LS}(D,\,\mathbf{0})$,

$$\begin{bmatrix} \boxed{1} & 0 & 0 & 4 \\ 0 & \boxed{1} & 0 & 0 \\ 0 & 0 & \boxed{1} & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

We can find infinitely many solutions to this system, most of them nontrivial, and we choose anyone we like to build a relation of linear dependence on R. Lets begin with $x_4 = 1$, to find the solution

$$\begin{bmatrix} -4 \\ 0 \\ -1 \\ 1 \end{bmatrix}.$$

So we can write the relation of linear dependence,

$$(-4)\mathbf{v}_1 + 0\mathbf{v}_2 + (-1)\mathbf{v}_3 + 1\mathbf{v}_4 = \mathbf{0}.$$

Theorem DLDS [89] guarantees that we can solve this relation of linear dependence for *some* vector in R, but the choice of which one is up to us. Notice however that $\mathbf{v}_2$ has a zero coefficient. In this case, we cannot choose to solve for $\mathbf{v}_2$. Maybe some other relation of linear dependence would produce a nonzero coefficient for $\mathbf{v}_2$ if we just had to solve for this vector. Unfortunately, this example has been engineered to *always* produce a zero coefficient here, as you can see from solving the homogenous system. Every solution has $x_2 = 0$!

OK, if we are convinced that we cannot solve for $\mathbf{v}_2$, lets instead solve for $\mathbf{v}_3$,

$$\mathbf{v}_3 = (-4)\mathbf{v}_1 + 0\mathbf{v}_2 + 1\mathbf{v}_4 = (-4)\mathbf{v}_1 + 1\mathbf{v}_4.$$

We now claim that this particular equation will allow us to write

$$V = \mathrm{Sp}\left(R\right) = \mathrm{Sp}\left(\{\mathbf{v}_1,\,\mathbf{v}_2,\,\mathbf{v}_3,\,\mathbf{v}_4\}\right) = \mathrm{Sp}\left(\{\mathbf{v}_1,\,\mathbf{v}_2,\,\mathbf{v}_4\}\right)$$

in essence declaring $\mathbf{v}_3$ as surplus for the task of building V as a span. This claim is an equality of two sets, so we will use Technique SE [13] to establish it carefully. Let $R' = \{\mathbf{v}_1,\,\mathbf{v}_2,\,\mathbf{v}_4\}$ and $V' = \mathrm{Sp}\left(R'\right)$. We want to show that $V = V'$.

First show that $V' \subseteq V$. Since every vector of R' is in R, any vector we can construct in V' as a linear combination of vectors from R' can also be constructed as a vector in V by the same linear combination of the same vectors in R. That was easy, now turn it around.

Next show that $V \subseteq V'$. Choose any $\mathbf{v}$ from V. Then there are scalars $\alpha_1,\,\alpha_2,\,\alpha_3,\,\alpha_4$

so that

$$\begin{aligned}
\mathbf{v} &= \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \alpha_3\mathbf{v}_3 + \alpha_4\mathbf{v}_4 \\
&= \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \alpha_3\left((-4)\mathbf{v}_1 + 1\mathbf{v}_4\right) + \alpha_4\mathbf{v}_4 \\
&= \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \left((-4\alpha_3)\mathbf{v}_1 + \alpha_3\mathbf{v}_4\right) + \alpha_4\mathbf{v}_4 \\
&= \left(\alpha_1 - 4\alpha_3\right)\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \left(\alpha_3 + \alpha_4\right)\mathbf{v}_4.
\end{aligned}$$

This equation says that $\mathbf{v}$ can then be writen as a linear combination of the vectors in R' and hence qualifies for membership in V'. So $V \subseteq V'$ and we have established that $V = V'$.

If R' was also linearly dependent (its not), we could reduce the set even further. Notice that we could have chosen to eliminate any one of $\mathbf{v}_1$, $\mathbf{v}_3$ or $\mathbf{v}_4$, but somehow $\mathbf{v}_2$ is essential to the creation of V since it cannot be replaced by any linear combination of $\mathbf{v}_1$, $\mathbf{v}_3$ or $\mathbf{v}_4$. △

Subsection LINSM Linear Independence and NonSingular Matrices

We will now specialize to sets of n vectors from $\mathbb{C}^n$. This will put Theorem MVSLD [89] off-limits, while Theorem LIVHS [88] will involve square matrices. Lets begin by contrasting Archetype A [167] and Archetype B [171].

Example LI.LDCAA
Linearly dependent columns in Archetype A
Archetype A [167] is a system of linear equations with coefficient matrix,

$$A = \begin{bmatrix} 1 & -1 & 2 \\ 2 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix}.$$

Do the columns of this matrix form a linearly independent or dependent set? By Example NSM.S [49] we know that A is singular. According to the definition of nonsingular matrices, Definition NM [49], the homogenous system $LS(A,\, \mathbf{0})\mathbf{0}$ has infinitely many solutions. So by Theorem LIVHS [88], the columns of A form a linearly dependent set. △

Example LI.LICAB
Linearly independent columns in Archetype B
Archetype B [171] is a system of linear equations with coefficient matrix,

$$B = \begin{bmatrix} -7 & -6 & -12 \\ 5 & 5 & 7 \\ 1 & 0 & 4 \end{bmatrix}.$$

Do the columns of this matrix form a linearly independent or dependent set? By Example NSM.NS [50] we know that B is nonsingular. According to the definition of nonsingular matrices, Definition NM [49], the homogenous system $\mathcal{LS}(A,\,\mathbf{0})\mathbf{0}$ has a unique solution. So by Theorem LIVHS [88], the columns of A form a linearly independent set. △

That Archetype A [167] and Archetype B [171] have opposite properties for the columns of their coefficient matrices is no accident. Here's the theorem, and then we will update our equivalences for nonsingular matrices, Theorem NSME1 [56].

Theorem NSLIC
NonSingular matrices have Linearly Independent Columns
Suppose that A is a square matrix. Then A is nonsingular if and only if the columns of A form a linearly independent set. □

Proof This is a proof where we can chain together equivalences, rather than proving the two halves separately. By definition, A is nonsingular if and only if the homogenous system $\mathcal{LS}(A,\,\mathbf{0})\mathbf{0}$ has a unique solution. Theorem LIVHS [88] then says that the system $\mathcal{LS}(A,\,\mathbf{0})$ has a unique solution if and only if the columns of A are a linearly independent set. ■

Here's an update to Theorem NSME1 [56].

Theorem NSME2
NonSingular Matrix Equivalences, Round 2
Suppose that A is a square matrix. The following are equivalent.

1. A is nonsingular.
2. A row-reduces to the identity matrix.
3. The null space of A contains only the zero vector, $\mathcal{N}(A) = \{\mathbf{0}\}$.
4. The linear system $\mathcal{LS}(A,\,\mathbf{b})$ has a unique solution for every possible choice of $\mathbf{b}$.
5. The columns of A form a linearly independent set. □

Proof Theorem NSLIC [93] is yet another equivalence for a nonsingular matrix, so we can add it to the list in Theorem NSME1 [56]. ■

Subsection NSSLI
Null Spaces, Spans, Linear Independence

In Subsection SS.SSNS [81] we proved Theorem SSNS [81] which provided $n-r$ vectors that could be used with the span construction to build the entire null space of a matrix. As we have seen in Theorem DLDS [89] and Example LI.RS [90], linearly dependent sets carry redundant vectors with them when used in building a set as a span. Our aim now is to show that the vectors provided by Theorem SSNS [81] form a linearly independent set, so in one sense they are as efficient as possible a way to describe the null space. The proof will also be a good exercise in showing how to prove a conclusion that states a set is linearly independent.

The proof is really quite straightforward, and relies on the "pattern" of zeros and ones that arise in the vectors $\mathbf{u}_i$, $1 \leq i \leq n-r$ in the entries that correspond to the free variables. So take a look at Example SS.VFSAD [70], Example SS.VFSAI [73] and Example SS.VFSAL [75], especially during the conclusion of Step 2 (temporarily ignore the construction of the constant vector, $\mathbf{c}$).

Theorem BNS
Basis for Null Spaces
Suppose that A is an $m \times n$ matrix, and B is a row-equivalent matrix in reduced row-echelon form with r nonzero rows. Let $D = \{d_1,\, d_2,\, d_3,\, \ldots,\, d_r\}$ and $F = \{f_1,\, f_2,\, f_3,\, \ldots,\, f_{n-r}\}$ be the sets of column indices where B does and does not (respectively) have leading 1's. Construct the $n-r$ vectors $\mathbf{u}_j = (u_{ij})$, $1 \leq j \leq n-r$ of size n as

$$u_{ij} = \begin{cases} 1 & \text{if } i \in F,\ i = f_j \\ 0 & \text{if } i \in F,\ i \neq f_j \\ -b_{k,f_j} & \text{if } i \in D,\ i = d_k \end{cases}.$$

Then the set $\{\mathbf{u}_1,\, \mathbf{u}_2,\, \mathbf{u}_3,\, \ldots,\, \mathbf{u}_{n-r}\}$ is linearly independent. □

Proof To prove the linear independence of a set, we need to start with a relation of linear dependence and somehow conclude that the scalars involved *must all be zero*, i.e. that the relation of linear dependence only happens in the trivial fashion. To this end, we start with

$$\alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \alpha_3\mathbf{u}_3 + \cdots + \alpha_{n-r}\mathbf{u}_{n-r} = \mathbf{0}.$$

For each j, $1 \leq j \leq n-r$, consider the entry of the vectors on both sides of this equality in position f_j. On the right, its easy since the zero vector has a zero in each entry. On the left we find,

$$\begin{aligned} &\alpha_1\mathbf{u}_{f_j1} + \alpha_2\mathbf{u}_{f_j2} + \alpha_3\mathbf{u}_{f_j3} + \cdots + \alpha_j\mathbf{u}_{f_j,j} + \cdots + \alpha_{n-r}\mathbf{u}_{f_j,n-r} = \\ &\alpha_1(0) + \alpha_2(0) + \alpha_3(0) + \cdots + \alpha_j(1) + \cdots + \alpha_{n-r}(0) = \\ &\alpha_j \end{aligned}$$

So for all j, $1 \leq j \leq n-r$, we have $\alpha_j = 0$, which is the conclusion that tells us that the *only* relation of linear dependence on $\{\mathbf{u}_1,\, \mathbf{u}_2,\, \mathbf{u}_3,\, \ldots,\, \mathbf{u}_{n-r}\}$ is the trivial one, hence the set is linearly independent, as desired. ■

Example LI.NSLIL
Null space spanned by linearly independent set, Archetype L
In Example SS.VFSAL [75] we previewed Theorem SSNS [81] by finding a set of two vectors such that their span was the null space for the matrix in Archetype L [214]. Writing the matrix as L, we have

$$N(L) = \mathrm{Sp}\left(\left\{\begin{bmatrix}-1\\2\\-2\\1\\0\end{bmatrix},\,\begin{bmatrix}2\\-2\\1\\0\\1\end{bmatrix}\right\}\right).$$

Solving the homogenous system $\mathcal{LS}(L,\,\mathbf{0})$ resulted in recognizing x_4 and x_5 as the free variables. So look in entries 4 and 5 of the two vectors above and notice the pattern of zeros and ones that provides the linear independence of the set. △

M: Matrices

Section MO
Matrix Operations

We have made frequent use of matrices for solving systems of equations, and we have begun to investigate a few of their properties, such as the null space and nonsingularity. In this Chapter, we will take a more systematic approach to the study of matrices, and in this section we will backup and start simple. We start with the definition of an important set.

Definition VSM
Vector Space of $m \times n$ Matrices
The vector space M_{mn} is the set of all $m \times n$ matrices with entries from the set of complex numbers. ⊙

Subsection MEASM
Matrix equality, addition, scalar multiplication

Just as we made, and used, a careful definition of equality for column vectors, so too, we have precise definitions for matrices.

Definition ME
Matrix Equality
The $m \times n$ matrices

$$A = (a_{ij}) \qquad\qquad B = (b_{ij})$$

are **equal**, written $A = B$ provided $a_{ij} = b_{ij}$ for all $1 \leq i \leq m$, $1 \leq j \leq n$. ⊙

So equality of matrices translates to the equality of complex numbers, on an entry-by-entry basis. Notice that we now have our *fourth* definition that uses the symbol '=' for shorthand. Whenever a theorem has a conclusion saying two matrices are equal (think about your objects), we will consider appealing to this definition as a way of formulating the top-level structure of the proof. We will now define two operations on the set M_{mn}. Again, we will overload a symbol ('+') and a convention (concatenation for multiplication).

Definition MA
Matrix Addition
Given the $m \times n$ matrices

$$A = (a_{ij}) \qquad\qquad B = (b_{ij})$$

define the **sum** of A and B to be $A + B = C = (c_{ij})$, where

$$c_{ij} = a_{ij} + b_{ij}, \quad 1 \leq i \leq m,\, 1 \leq j \leq n. \qquad \odot$$

So matrix addition takes two matrices of the same size and combines them (in a natural way!) to create a new matrix of the same size. Perhaps this is the "obvious" thing to do, but it doesn't relieve us from the obligation to state it carefully.

Example MO.MA
Addition of two matrices in M_{23}
If

$$A = \begin{bmatrix} 2 & -3 & 4 \\ 1 & 0 & -7 \end{bmatrix} \qquad\qquad B = \begin{bmatrix} 6 & 2 & -4 \\ 3 & 5 & 2 \end{bmatrix}$$

then

$$A + B = \begin{bmatrix} 2 & -3 & 4 \\ 1 & 0 & -7 \end{bmatrix} + \begin{bmatrix} 6 & 2 & -4 \\ 3 & 5 & 2 \end{bmatrix} = \begin{bmatrix} 2+6 & -3+2 & 4+(-4) \\ 1+3 & 0+5 & -7+2 \end{bmatrix} = \begin{bmatrix} 8 & -1 & 0 \\ 4 & 5 & -5 \end{bmatrix} \quad \triangle$$

Our second operation takes two objects of different types, specifically a number and a matrix, and combines them to create another matrix. As with vectors, in this context we call a number a **scalar** in order to emphasize that it is not a matrix.

Definition SMM
Scalar Matrix Multiplication
Given the $m \times n$ matrix $A = (a_{ij})$ and the scalar $\alpha \in \mathbb{C}$, the **scalar multiple** of A by α is the matrix $\alpha A = C = (c_{ij})$, where

$$c_{ij} = \alpha a_{ij}, \quad 1 \leq i \leq m,\, 1 \leq j \leq n. \qquad \odot$$

Notice again that we have yet another kind of multiplication, and it is again written putting two symbols side-by-side. Computationally, scalar matrix multiplication is very easy.

Example MO.MSM
Scalar multiplication in M_{32}
If

$$A = \begin{bmatrix} 2 & 8 \\ -3 & 5 \\ 0 & 1 \end{bmatrix}$$

and $\alpha = 7$, then

$$\alpha A = 7 \begin{bmatrix} 2 & 8 \\ -3 & 5 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 7(2) & 7(8) \\ 7(-3) & 7(5) \\ 7(0) & 7(1) \end{bmatrix} = \begin{bmatrix} 14 & 56 \\ -21 & 35 \\ 0 & 7 \end{bmatrix}$$

△

Subsection VSP
Vector Space Properties

To refer to matrices abstractly, we have used notation like $A = (a_{ij})$ to connect the name of a matrix with names for its individual entries. As expressions for matrices become more complicated, we will find this notation more cumbersome. So here's some notation that will help us along the way.

Notation MEN
Matrix Entries ($[A]_{ij}$)
For a matrix A, the notation $[A]_{ij}$ will refer to the complex number in row i and column j of A. ▽

As an example, we could rewrite the defining property for matrix addition as

$$[A + B]_{ij} = [A]_{ij} + [B]_{ij}\,.$$

Be careful with this notation, since it is easy to think that $[A]_{ij}$ refers to the *whole* matrix. It does not. It is just a *number*, but is a convenient way to talk about all the entries at once. You might see some of the motivation for this notation in the definition of matrix equality, Definition ME [97].

With definitions of vector addition and scalar multiplication we can state, and prove, several properties of each operation, and some properties that involve their interplay. We now collect ten of them here for later reference.

Theorem VSPM
Vector Space Properties of M_{mn}
Supppose that A, B and C are $m \times n$ matrices in M_{mn} and α and β are scalars. Then

1. $A + B \in M_{mn}$ (Additive closure)
2. $\alpha A \in M_{mn}$ (Scalar closure)
3. $A + B = B + A$ (Commutativity)
4. $A + (B + C) = (A + B) + C$ (Associativity)
5. There is a matrix, $\mathcal{O}$, called the **zero matrix**, such that $A+\mathcal{O} = A$ for all $A \in M_{mn}$. (Additive identity)
6. For each matrix $A \in M_{mn}$, there exists a matrix $-A \in M_{mn}$ so that $A+(-A) = \mathcal{O}$. (Additive inverses)
7. $\alpha(\beta A) = (\alpha\beta)A$ (Associativity)
8. $\alpha(A + B) = \alpha A + \alpha B$ (Distributivity)
9. $(\alpha + \beta)A = \alpha A + \beta A$ (Distributivity)
10. $1A = A$ □

Proof While some of these properties seem very obvious, they all require proof. However, the proofs are not very interesting, and border on tedious. We'll prove one version of distributivity very carefully, and you can test your proof-building skills on some of the others. We'll give our new notation for matrix entries a workout here. Compare the style of the proofs here with those given for vectors in Theorem VSPCM [61] — while the objects here are more complicated, our notation makes the proofs cleaner.

To prove that $(\alpha + \beta)A = \alpha A + \beta A$, we need to establish the equality of two matrices (see Technique GS [12]). Definition ME [97] says we need to establish the equality of their entries, one-by-one. How do we do this, when we do not even know how many entries the two matrices might have? This is where Notation MEN [99] comes into play. Ready? Here we go.

For *any* i and j, $1 \leq i \leq m$, $1 \leq j \leq n$,

$$[(\alpha + \beta)A]_{ij} = (\alpha + \beta)\,[A]_{ij} = \alpha\,[A]_{ij} + \beta\,[A]_{ij} = [\alpha A]_{ij} + [\beta A]_{ij} = [\alpha A + \beta A]_{ij}\,.$$

A one-liner! There are several things to notice here. (1) Each equals sign is an equality of numbers. (2) The two ends of the equation, being true for any i and j, allow us to conclude the equality of the matrices. (3) There are several plus signs, and several instances of concatenation. Identify each one, and state exactly what operation is being represented by each. (4) State the definition or theorem that makes each step of the proof possible. ■

For now, note the similarities between Theorem VSPM [100] about matrices and Theorem VSPCM [61] about vectors.

The zero matrix described in this theorem, $\mathcal{O}$, is what you would expect — a matrix full of zeros.

Definition ZM
Zero Matrix
The $m \times n$ **zero matrix** is written as $\mathcal{O} = \mathcal{O}_{m\times n} = (z_{ij})$ and defined by $z_{ij} = 0$ for all $1 \leq i \leq m$, $1 \leq j \leq n$. Or, equivalently, $[\mathcal{O}]_{ij} = 0$, for all $1 \leq i \leq m$, $1 \leq j \leq n$. ⊙

Subsection TSM Transposes and Symmetric Matrices

We describe one more common operation we can perform on matrices. Informally, to transpose a matrix is to build a new matrix by swapping its rows and columns.

Definition TM
Transpose of a Matrix
Given an $m \times n$ matrix A, its **transpose** is the $n \times m$ matrix A^t given by

$$\left[A^t\right]_{ij} = [A]_{ji}\,, \quad 1 \leq i \leq n,\, 1 \leq j \leq m.$$ ⊙

Example MO.TM
Transpose of a 3×4 matrix
Suppose

$$D = \begin{bmatrix} 3 & 7 & 2 & -3 \\ -1 & 4 & 2 & 8 \\ 0 & 3 & -2 & 5 \end{bmatrix}.$$

We could formulate the transpose, entry-by-entry, using the definition. But it is easier to just systematically rewrite rows as columns (or vice-versa). The form of the definition given will be more useful in proofs. So we have

$$D^t = \begin{bmatrix} 3 & -1 & 0 \\ 7 & 4 & 3 \\ 2 & 2 & -2 \\ -3 & 8 & 5 \end{bmatrix}.$$ △

It will sometimes happen that a matrix is equal to its transpose. In this case, we will call a matrix symmetric. These matrices occur naturally in certain situations, and also have some nice properties, so it is worth stating the definition carefully. Informally a matrix is symmetric if we can "flip" it about the main diagonal (upper-left corner, running down to the lower-right corner) and have it look unchanged.

Definition SM
Symmetric Matrix
The matrix A is **symmetric** if $A = A^t$. ⊙

Example MO.SM
A symmetric 5×5 matrix
The matrix

$$E = \begin{bmatrix} 2 & 3 & -9 & 5 & 7 \\ 3 & 1 & 6 & -2 & -3 \\ -9 & 6 & 0 & -1 & 9 \\ 5 & -2 & -1 & 4 & -8 \\ 7 & -3 & 9 & -8 & -3 \end{bmatrix}$$

is symmetric. △

You might have noticed that Definition SM [102] did not specify the size of the matrix A, as has been our custom. That's because it wasn't necessary. An alternative would have been to state the definition just for square matrices, but this is the substance of the next proof. But first, a bit more advice about constructing proofs.

Proof Technique P
Practice
Here is a technique used by many practicing mathematicians when they are teaching themselves new mathematics. As they read a textbook, monograph or research article, they attempt to prove each new theorem themselves, *before* reading the proof. Often the proofs can be very difficult, so it is wise not to spend too much time on each. Maybe limit your losses and try each proof for 10 or 15 minutes. Even if the proof is not found, it is time well-spent. You become more familar with the definitions involved, and the hypothesis and conclusion of the theorem. When you do work through the proof, it might make more sense, and you will gain added insight about just how to construct a proof.

The next theorem is a great place to try this technique. ♢

Theorem SMS
Symmetric Matrices are Square
Supppose that A is a symmetric matrix. Then A is square. □

Proof We start by specifying A's size, without assuming it is square, since we are tyring to *prove* that, so we can't also assume it. Suppose A is an $m \times n$ matrix. Because A is symmetric, we know by Definition SM [102] that $A = A^t$. So, in particular, A and A^t have the same size. The size of A^t is $n \times m$, so from $m \times n = n \times m$, we conclude that $m = n$, and hence A must be square. ■

We finish this section with another easy theorem, but it illustrates the interplay of our three new operations, our new notation, and the techniques used to prove matrix equalities.

Theorem TASM
Transposes, Addition, Scalar Multiplication
Supppose that A and B are $m \times n$ matrices. Then

1. $(A+B)^t = A^t + B^t$
2. $(\alpha A)^t = \alpha A^t$
3. $(A^t)^t = A$ □

Proof Each statement to be proved is an equality of matrices, so we work entry-by-entry and use Definition ME [97]. Think carefully about the objects involved here, the many uses of the plus sign and concatenation, and the justification for each step.

$$\left[(A+B)^t\right]_{ij} = [A+B]_{ji} = [A]_{ji} + [B]_{ji} = \left[A^t\right]_{ij} + \left[B^t\right]_{ij} = \left[A^t + B^t\right]_{ij}$$

and

$$\left[(\alpha A)^t\right]_{ij} = [\alpha A]_{ji} = \alpha\,[A]_{ji} = \alpha\left[A^t\right]_{ij} = \left[\alpha A^t\right]_{ij}$$

and

$$\left[\left(A^t\right)^t\right]_{ij} = \left[A^t\right]_{ji} = [A]_{ij}$$ ■

Section ROM
Range of a Matrix

Theorem SLSLC [67] showed us that there is a natural correspondence between solutions to linear systems and linear combinations of the columns of the coefficient matrix. This idea motivates the following important definition.

Definition RM
Range of a Matrix
Suppose that A is an $m \times n$ matrix with columns $\{\mathbf{A}_1,\, \mathbf{A}_2,\, \mathbf{A}_3,\, \ldots,\, \mathbf{A}_n\}$. Then the **range** of A, written $\mathrm{R}\left(A\right)$, is the subset of $\mathbb{C}^m$ containing all linear combinations of the columns of A,

$$\mathrm{R}\left(A\right) = \mathrm{Sp}\left(\{\mathbf{A}_1,\, \mathbf{A}_2,\, \mathbf{A}_3,\, \ldots,\, \mathbf{A}_n\}\right)$$ ⊙

Subsection RSE
Range and systems of equations

Upon encountering any new set, the first question we ask is what objects are in the set, and which objects are not? Here's an example of one way to answer this question, and it will motivate a theorem that will then answer the question precisely.

Example ROM.RMCS
Range of a matrix and consistent systems
Archetype D [179] and Archetype E [183] are linear system of equations, with an identical 3×4 coefficient matrix, which we call A here. However, Archetype D [179] is consistent, while Archetype E [183] is not. We can explain this distinction with the range of the matrix A.

The column vector of constants, $\mathbf{b}$, in Archetype D [179] is

$$\mathbf{b} = \begin{bmatrix} 8 \\ -12 \\ 4 \end{bmatrix}.$$

One solution to $LS(A,\, \mathbf{b})$, as listed, is

$$\mathbf{x} = \begin{bmatrix} 7 \\ 8 \\ 1 \\ 3 \end{bmatrix}.$$

By Theorem SLSLC [67], we can summarize this solution as a linear combination of the columns of A that equals $\mathbf{b}$,

$$7\begin{bmatrix}2\\-3\\1\end{bmatrix}+8\begin{bmatrix}1\\4\\1\end{bmatrix}+1\begin{bmatrix}7\\-5\\4\end{bmatrix}+3\begin{bmatrix}-7\\-6\\-5\end{bmatrix}=\begin{bmatrix}8\\-12\\4\end{bmatrix}=\mathbf{b}.$$

This equation says that $\mathbf{b}$ is a linear combination of the columns of A, and then by Definition RM [104], we can say that $\mathbf{b}\in\mathcal{R}(A)$.

On the other hand, Archetype E [183] is the linear system $\mathcal{LS}(A,\,\mathbf{c})$, where the vector of constants is

$$\mathbf{c}=\begin{bmatrix}2\\3\\2\end{bmatrix}$$

and this sytem of equations is inconsistent. This means $\mathbf{c}\not\in\mathcal{R}(A)$, for if it were, then it would equal a linear combination of the columns of A and Theorem SLSLC [67] would lead us to a solution of the system $\mathcal{LS}(A,\,\mathbf{c})$. △

So if we fix the coefficient matrix, and vary the vector of constants, we can sometimes find consistent systems, and sometimes inconsistent systems. The vectors of constants that lead to consistent systems are exactly the elements of the range. This is the content of the next theorem, and since it is an equivalence, it provides an alternate view of the range.

Theorem RCS
Range and Consistent Systems
Suppose A is an $m\times n$ matrix and $\mathbf{b}$ is a vector of size m. Then $\mathbf{b}\in\mathcal{R}(A)$ if and only if $\mathcal{LS}(A,\,\mathbf{b})$ is consistent. □

Proof ($\Rightarrow$) Suppose $\mathbf{b}\in\mathcal{R}(A)$. Then we can write $\mathbf{b}$ as some linear combination of the columns of A. By Theorem SLSLC [67] we can use the scalars from this linear combination to form a solution to $\mathcal{LS}(A,\,\mathbf{b})$, so this system is consistent.

($\Leftarrow$) If $\mathcal{LS}(A,\,\mathbf{b})$ is consistent, a solution may be used to write $\mathbf{b}$ as a linear combination of the columns of A. This qualifies $\mathbf{b}$ for membership in $\mathcal{R}(A)$. ■

Given a vector $\mathbf{b}$ and a matrix A it is now very mechanical to test if $\mathbf{b}\in\mathcal{R}(A)$. Form the linear system $\mathcal{LS}(A,\,\mathbf{b})$, row-reduce the augmented matrix, [A—$\mathbf{b}$], and test for consistency with Theorem RCLS [36].

Subsection RSOC
Range spanned by original columns

So we have a foolproof, automated procedure for determining membership in $\mathcal{R}(A)$. While this works just fine a vector at a time, we would like to have a more useful description of the set $\mathrm{R}(A)$ as a whole. The next example will preview the first of two fundamental results about the range of a matrix.

Example ROM.COC
Casting out columns, Archetype I

Archetype I [201] is a system of linear equations with $m = 4$ equations in $n = 7$ variables. Let I denote the 4×7 coefficient matrix from this system, and consider the range of I, $\mathrm{R}(I)$. By the definition, we have

$$\begin{aligned}\mathrm{R}(I) =& \mathrm{Sp}\left(\{\mathbf{I}_1,\,\mathbf{I}_2,\,\mathbf{I}_3,\,\mathbf{I}_4,\,\mathbf{I}_5,\,\mathbf{I}_6,\,\mathbf{I}_7\}\right)\\ =& \mathrm{Sp}\left(\left\{\begin{bmatrix}1\\2\\0\\-1\end{bmatrix},\,\begin{bmatrix}4\\8\\0\\-4\end{bmatrix},\,\begin{bmatrix}0\\-1\\2\\2\end{bmatrix},\,\begin{bmatrix}-1\\3\\-3\\4\end{bmatrix},\,\begin{bmatrix}0\\9\\-4\\8\end{bmatrix},\,\begin{bmatrix}7\\-13\\12\\-31\end{bmatrix},\,\begin{bmatrix}-9\\7\\-8\\37\end{bmatrix}\right\}\right)\end{aligned}$$

The set of columns of I is obviously linearly dependent, since we have $n = 7$ vectors from $\mathbb{C}^4$ (see Theorem MVSLD [89]). So we can slim down this set some, and still create the range as the span of a set. The row-reduced form for I is the matrix

$$\begin{bmatrix}\boxed{1} & 4 & 0 & 0 & 2 & 1 & -3\\ 0 & 0 & \boxed{1} & 0 & 1 & -3 & 5\\ 0 & 0 & 0 & \boxed{1} & 2 & -6 & 6\\ 0 & 0 & 0 & 0 & 0 & 0 & 0\end{bmatrix}$$

so we can easily create solutions to the homogenous sytem $LS(I,\,\mathbf{0})$ using the free variables x_2, x_5, x_6, x_7. Any such solution will correspond to a relation of linear dependence on the columns of I. These will allow us to solve for one column vector as a linear combination of some others, in the spirit of Theorem DLDS [89], and remove that vector from the set. We'll set about this task methodically. Set the free variable x_2 to one, and set the other free variables to zero. Then a solution to $LS(I,\,\mathbf{0})$ is

$$\mathbf{x} = \begin{bmatrix}-4\\1\\0\\0\\0\\0\\0\end{bmatrix}$$

which can be used to create the linear combination

$$(-4)\mathbf{I}_1 + 1\mathbf{I}_2 + 0\mathbf{I}_3 + 0\mathbf{I}_4 + 0\mathbf{I}_5 + 0\mathbf{I}_6 + 0\mathbf{I}_7 = \mathbf{0}$$

This can then be arranged and solved for $\mathbf{I}_2$, resulting in $\mathbf{I}_2$ expressed as a linear combination of $\{\mathbf{I}_1,\,\mathbf{I}_3,\,\mathbf{I}_4\}$,

$$\mathbf{I}_2 = 4\mathbf{I}_1 + 0\mathbf{I}_3 + 0\mathbf{I}_4$$

This means that $\mathbf{I}_2$ is surplus, and we can create $\mathcal{R}(I)$ just as well with a smaller set with this vector removed,

$$\mathcal{R}(I) = \mathrm{Sp}(\{\mathbf{I}_1, \mathbf{I}_3, \mathbf{I}_4, \mathbf{I}_5, \mathbf{I}_6, \mathbf{I}_7\})$$

Set the free variable x_5 to one, and set the other free variables to zero. Then a solution to $\mathcal{LS}(I, \mathbf{0})$ is

$$\mathbf{x} = \begin{bmatrix} -2 \\ 0 \\ -1 \\ -2 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

which can be used to create the linear combination

$$(-2)\mathbf{I}_1 + 0\mathbf{I}_2 + (-1)\mathbf{I}_3 + (-2)\mathbf{I}_4 + 1\mathbf{I}_5 + 0\mathbf{I}_6 + 0\mathbf{I}_7 = \mathbf{0}$$

This can then be arranged and solved for $\mathbf{I}_5$, resulting in $\mathbf{I}_5$ expressed as a linear combination of $\{\mathbf{I}_1, \mathbf{I}_3, \mathbf{I}_4\}$,

$$\mathbf{I}_5 = 2\mathbf{I}_1 + 1\mathbf{I}_3 + 2\mathbf{I}_4$$

This means that $\mathbf{I}_5$ is surplus, and we can create $\mathcal{R}(I)$ just as well with a smaller set with this vector removed,

$$\mathcal{R}(I) = \mathrm{Sp}(\{\mathbf{I}_1, \mathbf{I}_3, \mathbf{I}_4, \mathbf{I}_6, \mathbf{I}_7\})$$

Do it again, set the free variable x_6 to one, and set the other free variables to zero. Then a solution to $\mathcal{LS}(I, \mathbf{0})$ is

$$\mathbf{x} = \begin{bmatrix} -1 \\ 0 \\ 3 \\ 6 \\ 0 \\ 1 \\ 0 \end{bmatrix}$$

which can be used to create the linear combination

$$(-1)\mathbf{I}_1 + 0\mathbf{I}_2 + 3\mathbf{I}_3 + 6\mathbf{I}_4 + 0\mathbf{I}_5 + 1\mathbf{I}_6 + 0\mathbf{I}_7 = \mathbf{0}$$

This can then be arranged and solved for $\mathbf{I}_6$, resulting in $\mathbf{I}_6$ expressed as a linear combination of $\{\mathbf{I}_1, \mathbf{I}_3, \mathbf{I}_4\}$,

$$\mathbf{I}_6 = 1\mathbf{I}_1 + (-3)\mathbf{I}_3 + (-6)\mathbf{I}_4$$

This means that $\mathbf{I}_6$ is surplus, and we can create $\mathrm{R}(I)$ just as well with a smaller set with this vector removed,

$$\mathrm{R}(I) = \mathrm{Sp}(\{\mathbf{I}_1,\, \mathbf{I}_3,\, \mathbf{I}_4,\, \mathbf{I}_7\})$$

Set the free variable x_7 to one, and set the other free variables to zero. Then a solution to $\mathcal{LS}(I,\, \mathbf{0})$ is

$$\mathbf{x} = \begin{bmatrix} 3 \\ 0 \\ -5 \\ -6 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

which can be used to create the linear combination

$$3\mathbf{I}_1 + 0\mathbf{I}_2 + (-5)\mathbf{I}_3 + (-6)\mathbf{I}_4 + 0\mathbf{I}_5 + 0\mathbf{I}_6 + 1\mathbf{I}_7 = \mathbf{0}$$

This can then be arranged and solved for $\mathbf{I}_7$, resulting in $\mathbf{I}_7$ expressed as a linear combination of $\{\mathbf{I}_1,\, \mathbf{I}_3,\, \mathbf{I}_4\}$,

$$\mathbf{I}_7 = (-3)\mathbf{I}_1 + 5\mathbf{I}_3 + 6\mathbf{I}_4$$

This means that $\mathbf{I}_7$ is surplus, and we can create $\mathrm{R}(I)$ just as well with a smaller set with this vector removed,

$$\mathrm{R}(I) = \mathrm{Sp}(\{\mathbf{I}_1,\, \mathbf{I}_3,\, \mathbf{I}_4\})$$

You might think we could keep this up, but we have run out of free variables. And not coincidentally, the set $\{\mathbf{I}_1,\, \mathbf{I}_3,\, \mathbf{I}_4\}$ is linearly independent (check this!). Hopefully it is clear how each free variable was used to eliminate the corresponding column from the set used to span the range for this will be the essence of the proof of the next theorem. See if you can mimic this example using Archetype J [206]. Go ahead, we'll go grab a cup of coffee and be back before you finish up.

For extra credit, notice that the vector

$$\mathbf{b} = \begin{bmatrix} 3 \\ 9 \\ 1 \\ 4 \end{bmatrix}$$

is the vector of constants in the definition of Archetype I [201]. Since the system $\mathcal{LS}(I,\, \mathbf{b})$ is consistent, we know by Theorem RCS [105] that $\mathbf{b} \in \mathrm{R}(I)$. This means that $\mathbf{b}$ must be a linear combination of just $\mathbf{I}_1$, $\mathbf{I}_3$, $\mathbf{I}_4$. Can you find this linear combination? Did you notice that there is just a single (unique) answer? Hmmmm. △

We will now formalize the previous example, which will make it trivial to determine a linearly independent set of vectors that will span the range of a matrix. However, the connections made in the last example are worth working through the example (and Archetype J [206]) carefully before employing the theorem.

Theorem BROC
Basis of the Range with Original Columns
Suppose that A is an $m \times n$ matrix with columns $\mathbf{A}_1,\,\mathbf{A}_2,\,\mathbf{A}_3,\,\ldots,\,\mathbf{A}_n$, and B is a row-equivalent matrix in reduced row-echelon form with r nonzero rows. Let $D = \{d_1,\,d_2,\,d_3,\,\ldots,\,d_r\}$ be the set of column indices where B has leading 1's. Let $S = \{\mathbf{A}_{d_1},\,\mathbf{A}_{d_2},\,\mathbf{A}_{d_3},\,\ldots,\,\mathbf{A}_{d_r}\}$. Then

1. $\mathrm{R}\left(A\right) = \mathrm{Sp}\left(S\right)$.
2. S is a linearly independent set. □

Proof We have two conclusions stemming from the same hypothesis. We'll prove the first conclusion first. By definition

$$\mathrm{R}\left(A\right) = \mathrm{Sp}\left(\{\mathbf{A}_1,\,\mathbf{A}_2,\,\mathbf{A}_3,\,\ldots,\,\mathbf{A}_n\}\right).$$

We will find relations of linear dependence on this set of column vectors, each one involving a column corresponding to a free variable along with the columns corresponding to the dependent variables. By expressing the free variable column as a linear combination of the dependent variable columns, we will be able to reduce the set S down to only the set of dependent variable columns while preserving the span.

Let $F = \{f_1,\,f_2,\,f_3,\,\ldots,\,f_{n-r}\}$ be the sets of column indices where B does not have any leading 1's. For each j, $1 \leq j \leq n-r$ construct the specific solution to the homogenous system $\mathcal{LS}(A,\,\mathbf{0})$ given by Theorem VFSLS [72] where the free variables are chosen by the rule

$$\mathbf{x}_{f_i} = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j \end{cases}.$$

This leads to a solution that is exactly the vector $\mathbf{u}_j$ as defined in Theorem SSNS [81],

$$u_{ij} = \begin{cases} 1 & \text{if } i \in F,\, i = f_j \\ 0 & \text{if } i \in F,\, i \neq f_j \\ -b_{k,f_j} & \text{if } i \in D,\, i = d_k \end{cases}.$$

Then Theorem SLSLC [67] says this solution corresponds to the following relation of linear dependence on the columns of A,

$$(-b_{1f_j})\mathbf{A}_{d_1} + (-b_{2f_j})\mathbf{A}_{d_2} + (-b_{3f_j})\mathbf{A}_{d_3} + \cdots + (-b_{rf_j})\mathbf{A}_{d_r} + (1)\mathbf{A}_{f_j} = \mathbf{0}.$$

This can be rearranged as

$$\mathbf{A}_{f_j} = b_{1f_j}\mathbf{A}_{d_1} + b_{2f_j}\mathbf{A}_{d_2} + b_{3f_j}\mathbf{A}_{d_3} + \cdots + b_{rf_j}\mathbf{A}_{d_r}.$$

This equation can be interpreted to tell us that $\mathbf{A}_{f_j} \in \mathrm{Sp}\left(S\right)$ for all $1 \leq j \leq n-r$, so $\mathrm{Sp}\left(\{\mathbf{A}_1,\,\mathbf{A}_2,\,\mathbf{A}_3,\,\ldots,\,\mathbf{A}_n\}\right) \subseteq \mathrm{Sp}\left(S\right)$. It should be easy to convince yourself (so go ahead and do it!) that the opposite is true, i.e. $\mathrm{Sp}(S) \subseteq \mathrm{Sp}(\{\mathbf{A}_1, \mathbf{A}_2, \mathbf{A}_3, \ldots, \mathbf{A}_n\})$. These two statements about subsets combine (see Technique SE [13]) to give the desired set equality as our conclusion.

Our second conclusion is that S is a linearly independent set. To prove this, we will begin with a relation of linear dependence on S. So suppose there are scalars $\alpha_1, \alpha_2, \alpha_3, \ldots, \alpha_r$ so that

$$\alpha_1 \mathbf{A}_{d_1} + \alpha_2 \mathbf{A}_{d_2} + \alpha_3 \mathbf{A}_{d_3} + \cdots + \alpha_r \mathbf{A}_{d_r} = \mathbf{0}.$$

To establish linear independence, we wish to deduce that $\alpha_i = 0$, $1 \leq i \leq r$. This relation of linear dependence allows us to use Theorem SLSLC [67] and construct a solution, $\mathbf{x}$, to the homogenous system $\mathcal{LS}(A, \mathbf{0})$. This vector, $\mathbf{x}$, has α_i in entry d_i, and zeros in every entry at an index in F. This is equivalent then to a solution, $\mathbf{x}$, where each free variable equals zero. What would such a solution look like?

By Theorem VFSLS [72], or from details contained in the proof of Theorem SSNS [81], we see that the only solution to a homogenous system with the free variables all chosen to be zero is precisely the trivial solution, so $\mathbf{x} = \mathbf{0}$. Since each α_i occurs somewhere as an entry of $\mathbf{x} = \mathbf{0}$, we conclude, as desired, that $\alpha_i = 0$, $1 \leq i \leq r$, and hence the set S is linearly independent. ■

This is a very pleasing result since it gives us a handful of vectors that describe the entire range (through the span), and we believe this set is as small as possible because we cannot create any more relations of linear dependence to trim it down further. Furthermore, we defined the range (Definition RM [104]) as all linear combinations of the columns of the matrix, and the elements of the set S are still columns of the matrix (we won't be so lucky in the next two constructions of the range).

Procedurally this theorem is very easy to apply. Row-reduce the original matrix, identify r columns with leading 1's in this reduced matrix, and grab the corresponding columns of the original matrix. Its still important to study the proof of Theorem BROC [109] and its motivation in Example ROM.COC [106]. We'll trot through an example all the same.

Example ROM.ROC
Range with original columns, Archetype D

Let's determine a compact expression for the entire range of the coefficient matrix of the system of equations that is Archetype D [179]. Notice that in Example ROM.RMCS [104] we were only determining if individual vectors were in the range or not.

To start with the application of Theorem BROC [109], call the coefficient matrix A

$$A = \begin{bmatrix} 2 & 1 & 7 & -7 \\ -3 & 4 & -5 & -6 \\ 1 & 1 & 4 & -5 \end{bmatrix}.$$

and row-reduce it to reduced row-echelon form,

$$B = \begin{bmatrix} \boxed{1} & 0 & 3 & -2 \\ 0 & \boxed{1} & 1 & -3 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

There are leading 1's in columns 1 and 2, so $D = \{1, 2\}$. To construct a set that spans $\mathrm{R}(A)$, just grab the columns of A indicated by the set D, so

$$\mathrm{R}(A) = \mathrm{Sp}\left(\left\{\begin{bmatrix} 2 \\ -3 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 4 \\ 1 \end{bmatrix}\right\}\right).$$

That's it.

In Example ROM.RMCS [104] we determined that the vector

$$\mathbf{c} = \begin{bmatrix} 2 \\ 3 \\ 2 \end{bmatrix}$$

was not in the range of A. Try to write $\mathbf{c}$ as a linear combination of the first two columns of A. What happens?

Also in Example ROM.RMCS [104] we determined that the vector

$$\mathbf{b} = \begin{bmatrix} 8 \\ -12 \\ 4 \end{bmatrix}$$

was in the range of A. Try to write $\mathbf{b}$ as a linear combination of the first two columns of A. What happens? Did you find a unique solution to this question? Hmmmm. △

Subsection RNS
The range as null space

We've come to know null spaces quite well, since they are the sets of solutions to homogenous systems of equations. In this subsection, we will see how to describe the range of a matrix as the null space of a different matrix. Then all of our techniques for studying null spaces can be brought to bear on ranges. As usual, we will begin with an example, and then genealize to a theorem.

Example ROM.RNSAD
Range as null space, Archetype D
Begin again with the coefficient matrix of Archetype D [179],

$$A = \begin{bmatrix} 2 & 1 & 7 & -7 \\ -3 & 4 & -5 & -6 \\ 1 & 1 & 4 & -5 \end{bmatrix}$$

and we will describe another approach to finding the range of A.

Theorem RCS [105] says a vector is in the range only if it can be used as the vector of constants for a system of equations with coefficient matrix A and result in a consistent system. So suppose we have an arbitrary vector in the range,

$$\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} \in \mathrm{R}\left(A\right).$$

Then the linear system $\mathcal{LS}(A,\,b)$ will be consistent by Theorem RCS [105]. Let's consider solutions to this system by first creating the augmented matrix

$$[A|\mathbf{b}] = \begin{bmatrix} 2 & 1 & 7 & -7 & b_1 \\ -3 & 4 & -5 & -6 & b_2 \\ 1 & 1 & 4 & -5 & b_3 \end{bmatrix}.$$

To locate solutions we would row-reduce this matrix and bring it to reduced row-echelon form. Despite the presence of variables in the last column, there is nothing to stop us from doing this. Except our numerical routines on calculators can't be used, and even some of the symbolic algebra routines do some unexpected maneuvers with this computation. So do it by hand. Notice along the way that the row operations are *exactly* the same ones you would do if you were just row-reducing the coefficient matrix alone, say in connection with a homogenous system of equations. The column with the b_i acts as a sort of bookkeeping device. Here's what you should get:

$$[A|\mathbf{b}] = \begin{bmatrix} \boxed{1} & 0 & 3 & -2 & -\frac{1}{7}b_2 + \frac{4}{7}b_3 \\ 0 & \boxed{1} & 1 & -3 & \frac{1}{7}b_2 + \frac{3}{7}b_3 \\ 0 & 0 & 0 & 0 & b_1 + \frac{1}{7}b_2 - \frac{11}{7}b_3 \end{bmatrix}.$$

Is this a consistent system or not? There's an expression in the last column of the third row, preceded by zeros. Theorem RCLS [36] tells us to look at the leading 1 of the last nonzero row, and see if it is in the final column. Could the expression at the end of the third row be a leading 1 in the last column? The answer is: maybe. It depends on $\mathbf{b}$. Some vectors are in the range, some are not. For $\mathbf{b}$ to be in the range, the system $\mathcal{LS}(A,\,\mathbf{b})$ must be consistent, and the expression in question must not be a leading 1. The only way to prevent it from being a leading 1 is if it is zero, since any nonzero value could be scaled to equal 1 by a row operation. So we have

$$\mathbf{b} \in \mathrm{R}\left(A\right) \quad\iff\quad \mathcal{LS}(A,\,\mathbf{b}) \text{ is consistent} \quad\iff\quad b_1 + \frac{1}{7}b_2 - \frac{11}{7}b_3 = 0.$$

So we have an algebraic description of vectors that are, or are not, in the range. And this description looks like a single linear homogenous equation in the variables b_1, b_2, b_3. The coefficient matrix of this (simple) homogenous system has the following coefficient matrix

$$K = \begin{bmatrix} 1 & \frac{1}{7} & -\frac{11}{7} \end{bmatrix}.$$

So we can write that $\mathrm{R}\left(A\right) = \mathrm{N}\left(K\right)$! Example ROM.RMCS [104] has a vector $\mathbf{b} \in \mathrm{R}\left(A\right)$ and a vector $\mathbf{c} \notin \mathrm{R}\left(A\right)$. Test each of these vectors for membership in $\mathrm{N}\left(K\right)$.

The four columns of the matrix A are definitely in the range, are they also in $\mathrm{N}(K)$? (The work above tells us these answers shouldn't be surprising, but perhaps doing the computations makes it feel a bit remarkable?).

Theorem SSNS [81] tells us how to find a set that spans a null space, and Theorem BNS [94] tells us that the same set is linearly indepenent. If we compute this set for K in this example, we find

$$\mathrm{R}(A) = \mathrm{N}(K) = \mathrm{Sp}\left(\{\mathbf{u}_1,\, \mathbf{u}_2\}\right) = \mathrm{Sp}\left(\left\{\begin{bmatrix}-\frac{1}{7}\\1\\0\end{bmatrix},\,\begin{bmatrix}\frac{11}{7}\\0\\1\end{bmatrix}\right\}\right).$$

Can you write these two new vectors $(\mathbf{u}_1,\, \mathbf{u}_2)$ each as linear combinations of the columns of A? Uniquely? Can you write each of them as linear cominations of just the first *two* columns of A? Uniquely? Hmmmm.

Doing row operations by hand with variables can be a bit error prone, so lets continue with this example and see if we can impove on it some. Rather than having b_1, b_2, b_3 all moving around in the same column, lets put each in its own column. So if we instead row-reduce

$$\begin{bmatrix}2 & 1 & 7 & -7 & b_1 & 0 & 0\\ -3 & 4 & -5 & -6 & 0 & b_2 & 0\\ 1 & 1 & 4 & -5 & 0 & 0 & b_3\end{bmatrix}$$

we find

$$\begin{bmatrix}\boxed{1} & 0 & 3 & -2 & 0 & -\frac{1}{7}b_2 & \frac{4}{7}b_3\\ 0 & \boxed{1} & 1 & -3 & 0 & \frac{1}{7}b_2 & \frac{3}{7}b_3\\ 0 & 0 & 0 & 0 & b_1 & \frac{1}{7}b_2 & -\frac{11}{7}b_3\end{bmatrix}.$$

If we sum the entries of the third row in columns 4, 5 and 6, and set it equal to zero, we get the equation

$$b_1 + \frac{1}{7}b_2 - \frac{11}{7}b_3 = 0$$

which we recognize as the previous condition for membership in the range. Perhaps you can see the row operations reflected in the revised form of the matrix involving the variables. You might also notice that the variables are acting more and more like placeholders (and just getting in the way). Lets try removing them. One more time. Now row-reduce, using a calculator if you like since there are no symbols,

$$\begin{bmatrix}2 & 1 & 7 & -7 & 1 & 0 & 0\\ -3 & 4 & -5 & -6 & 0 & 1 & 0\\ 1 & 1 & 4 & -5 & 0 & 0 & 1\end{bmatrix}$$

to obtain

$$\begin{bmatrix}\boxed{1} & 0 & 3 & -2 & 0 & -\frac{1}{7} & \frac{4}{7}\\ 0 & \boxed{1} & 1 & -3 & 0 & \frac{1}{7} & \frac{3}{7}\\ 0 & 0 & 0 & 0 & \boxed{1} & \frac{1}{7} & -\frac{11}{7}\end{bmatrix}.$$

The matrix K from above is now sitting down in the last row, in our "new" columns (5,6,7), and can be read off quicky. Why, we could even program a computer could do it! △

ARCHETYPE D [179] is just a tad on the small size for fully motivating the following theorem. In practice, the original matrix may row-reduce to a matrix with several nonzero rows. If we row-reduced an augmented matrix having a variable vector of constants ($\mathbf{b}$), we might find several expressions that need to be zero for the system to be consistent. Notice that it is not enough to make just the last expression zero, as then the one above it would also have to be zero, etc. In this way we typically end up with *several* homogenous equations prescribing elements of the range, and the coefficient matrix of this system (K) will have several rows.

Here's a theorem based on the preceding example, which will give us another procedure for describing the range of a matrix.

Theorem RNS
Range as a Null Space
Suppose that A is an $m \times n$ matrix. Create the $m \times (n+m)$ matrix M by placing the $m \times m$ identity matrix I_m to the right of the matrix A. Symbolically, $M = [A|I_m]$. Let N be a matrix that is row-equivalent to M and in reduced row-echelon form. Suppose there are r leading 1's of N in the first n columns. If $r = m$, then $\mathrm{R}\left(A\right) = \mathbb{C}^m$. Otherwise, $r < m$ and let K be the $(m-r) \times m$ matrix formed from the entries of N in the last $m - r$ rows and last m columns. Then

1. K is in reduced row-echelon form.

2. K has no zero rows, or equivalently, K has $m - r$ leading 1's.

3. $\mathrm{R}\left(A\right) = \mathrm{N}\left(K\right)$. □

Proof Let B denote the $m \times n$ matrix that is the first n columns of N, and let J denote the $m \times m$ matrix that is the final m columns of N. Then the sequence of row operations that convert M into N, will also convert A into B and I_m into J.

When $r = m$, there are m leading 1's in N that occur in the first n columns, so B has no zero rows. Thus, the linear system $\mathcal{LS}(A, \mathbf{b})$ is never inconsistent, no matter which vector is chosen for $\mathbf{b}$. So by THEOREM RCS [105], every $\mathbf{b} \in \mathbb{C}^m$ is in $\mathrm{R}\left(A\right)$.

Now consider the case when $r < m$. The final $m-r$ rows of B are zero rows since the leading 1's of these rows for N are located in columns $n+1$ or higher. The final $m-r$ rows of J form the matrix K.

Since N is in reduced row-echelon form, and the first n entries of each of the final $m-r$ rows are zero, K will have leading 1's in an echelon pattern, any zero rows are at the bottom (but we'll soon see that there aren't any), and columns with leading 1's will be otherwise zero. In other words, K is in reduced row-echelon form.

THEOREM NSRRI [50] tells us that the matrix I_m is nonsingular, since it is row-equivalent to the identity matrix (by an empty sequence of row operations!). Therefore, it cannot be row-equivalent to a matrix with a zero row. Why not? A square matrix with a zero row is the coefficient matrix of a homogenous system that has more variables than equations, if we consider the zero row as a "nonexistent" equation. THEOREM HMVEI [44] then says the system has infinitely many solutions. In turn this implies that the homogenous linear system $\mathcal{LS}(I_m, \mathbf{0})\mathbf{0}$ has infinitely many solutions, implying that I_m is singular, a contradiction. Since K is part of J, and J is row-equivalent to I_m, there can be no zero rows in K. If K has no zero rows, then it must have a leading 1 in each of its $m-r$ rows.

For our third conclusion, begin by supposing that $\mathbf{b}$ is an arbitrary vector in $\mathbb{C}^m$. To the vector $\mathbf{b}$ apply the row operations that convert M to N and call the resulting vector $\mathbf{c}$. Then the linear systems $\mathcal{LS}(A, \mathbf{b})$ and $\mathcal{LS}(B, \mathbf{c})$ are equivalent. Also, the linear systems $\mathcal{LS}(I_m, \mathbf{b})$ and $\mathcal{LS}(J, \mathbf{c})$ are equivalent and the unique solution to each is simply $\mathbf{b}$. Finally, let $\mathbf{c}^*$ be the vector of length $m-r$ containing the final $m-r$ entries of $\mathbf{c}$. Then we have the following,

$$\begin{aligned}
\mathbf{b}\in \mathcal{R}(A) &\iff \mathcal{LS}(A, \mathbf{b}) \text{ is consistent} && \text{Theorem RCS [105]}\\
&\iff \mathcal{LS}(B, \mathbf{c}) \text{ is consistent} && \text{Definition ES [10]}\\
&\iff \mathbf{c}^* = \mathbf{0} && \text{Theorem RCLS [36]}\\
&\iff \mathbf{b} \text{ is a solution to } \mathcal{LS}(K, \mathbf{0}) && \mathbf{b} \text{ is unique solution to } \mathcal{LS}(J, \mathbf{c})\\
&\iff \mathbf{b}\in \mathcal{N}(K). && \text{Definition NSM [47]}
\end{aligned}$$

Running these equivalences in the two different directions will establish the subset inclusions needed by Technique SE [13] and so we can conclude that $\mathcal{R}(A) = \mathcal{N}(K)$. ■

We've commented that Archetype D [179] was a tad small to fully appreciate this theorem. Let's apply it now to an Archetype where there's a bit more action.

Example ROM.RNSAG
Range as null space, Archetype G

Archetype G [192] and Archetype H [196] are both systems of $m = 5$ equations in $n = 2$ variables. They have identical coefficient matrices, which we will denote here as the matrix G,

$$G = \begin{bmatrix} 2 & 3\\ -1 & 4\\ 3 & 10\\ 3 & -1\\ 6 & 9 \end{bmatrix}.$$

Adjoin the 5×5 identity matrix, I_5, to form

$$M = \begin{bmatrix} 2 & 3 & 1 & 0 & 0 & 0 & 0\\ -1 & 4 & 0 & 1 & 0 & 0 & 0\\ 3 & 10 & 0 & 0 & 1 & 0 & 0\\ 3 & -1 & 0 & 0 & 0 & 1 & 0\\ 6 & 9 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

This row-reduces to

$$N = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & \frac{3}{11} & \frac{1}{33} \\ 0 & 1 & 0 & 0 & 0 & -\frac{2}{11} & \frac{1}{11} \\ 0 & 0 & 1 & 0 & 0 & 0 & -\frac{1}{3} \\ 0 & 0 & 0 & 1 & 0 & 1 & -\frac{1}{3} \\ 0 & 0 & 0 & 0 & 1 & 1 & -1 \end{bmatrix}.$$

The first $n = 2$ columns contain $r = 2$ leading 1's, so we extract K from the final $m - r = 3$ rows in the final $m = 5$ columns. Since this matrix is guaranteed to be in reduced row-echelon form, we mark the leading 1's.

$$K = \begin{bmatrix} \boxed{1} & 0 & 0 & 0 & -\frac{1}{3} \\ 0 & \boxed{1} & 0 & 1 & -\frac{1}{3} \\ 0 & 0 & \boxed{1} & 1 & -1 \end{bmatrix}.$$

Theorem RNS [114] now allows us to conclude that $\mathrm{R}(G) = \mathrm{N}(K)$. But we can do better. Theorem SSNS [81] tells us how to find a set that spans a null space, and Theorem BNS [94] tells us that the same set is linearly indepenent. The matrix K has 3 nonzero rows and 5 columns, so the homogenous system $\mathcal{LS}(K, \mathbf{0})$ will have solutions described by two free variables x_4 and x_5 in this case. Applying these results in this example yields,

$$\mathrm{R}(G) = \mathrm{N}(K) = \mathrm{Sp}(\{\mathbf{u}_1, \mathbf{u}_2\}) = \mathrm{Sp}\left(\left\{\begin{bmatrix} 0 \\ -1 \\ -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} \frac{1}{3} \\ \frac{1}{3} \\ 1 \\ 0 \\ 1 \end{bmatrix}\right\}\right).$$

As mentioned earlier, Archetype G [192] is consistent, while Archetype H [196] is inconsistent. See if you can write the two different vectors of constants as linear combinations of $\mathbf{u}_1$ and $\mathbf{u}_2$. How about the two columns of G? They must be in the range of G also. Are your answers unique? Do you notice anything about the scalars that appear in the linear combinations you are forming? △

Subsection RNSM
Range of a Nonsingular Matrix

Lets specialize to square matrices and contrast the ranges of the coefficient matrices in Archetype A [167] and Archetype B [171].

Example ROM.RAA
Range of Archetype A

The coefficient matrix in Archetype A [167] is

$$A = \begin{bmatrix} 1 & -1 & 2 \\ 2 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix}$$

which row-reduces to

$$\begin{bmatrix} \boxed{1} & 0 & 1 \\ 0 & \boxed{1} & -1 \\ 0 & 0 & 0 \end{bmatrix}.$$

Columns 1 and 2 have leading 1's, so by Theorem BROC [109] we can write

$$\mathrm{R}\left(A\right) = \mathrm{Sp}\left(\{\mathbf{A}_1,\ \mathbf{A}_2\}\right) = \mathrm{Sp}\left(\left\{\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix},\ \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}\right\}\right).$$

We want to show in this example that $\mathrm{R}\left(A\right) \neq \mathbb{C}^3$. So take, for example, the vector $\mathbf{b} = \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix}$. Then there is no solution to the system $\mathcal{LS}(A,\ \mathbf{b})$, or equivalently, it is not possible to write $\mathbf{b}$ as a linear combination of $\mathbf{A}_1$ and $\mathbf{A}_2$. Try one of these two computations yourself. (Or try both!). Since $\mathbf{b} \notin \mathrm{R}\left(A\right)$, the range of A cannot be all of $\mathbb{C}^3$. So by varying the vector of constants, it is possible to create inconsistent systems of equations with this coefficient matrix (the vector $\mathbf{b}$ being one such example). △

Example ROM.RAB
Range of Archetype B

The coefficient matrix in Archetype B [171], call it B here, is known to be nonsingular (see Example NSM.NS [50]). By Theorem NSMUS [53], the linear system $\mathcal{LS}(B,\ \mathbf{b})$ has a (unique) solution for every choice of $\mathbf{b}$. Theorem RCS [105] then says that $\mathbf{b} \in \mathrm{R}\left(B\right)$ for all $\mathbf{b} \in \mathbb{C}^3$. Stated differently, there is no way to build an inconsistent system with the coefficient matrix B, but then we knew that already from Theorem NSMUS [53]. △

Theorem RNSM
Range of a NonSingular Matrix

Suppose A is a square matrix of size n. Then A is nonsingular if and only if $\mathrm{R}\left(A\right) = \mathbb{C}^n$. □

Proof ($\Leftarrow$) Suppose A is nonsingular. By Theorem NSMUS [53], the linear system $\mathcal{LS}(A,\ \mathbf{b})$ has a (unique) solution for every choice of $\mathbf{b}$. Theorem RCS [105] then says that $\mathbf{b} \in \mathrm{R}\left(A\right)$ for all $\mathbf{b} \in \mathbb{C}^n$. In other words, $\mathrm{R}\left(A\right) = \mathbb{C}^n$.

($\Rightarrow$) We'll prove the contrapositive (see Technique CP [35]). Suppose that A is singular. By Theorem NSRRI [50], A will not row-reduce to the identity matrix I_n. So the row-equivalent matrix B of Theorem RNS [114] has $r < n$ nonzero rows and then the matrix K is a nonzero matrix (it has at least one leading 1 in it). By Theorem NSRRI [50], $\mathrm{R}(A) = \mathrm{N}(K)$. If we can find one vector of $\mathbb{C}^n$ that is not in $\mathrm{N}(K)$, then we can conclude that $\mathrm{R}(A) \neq \mathbb{C}^n$, the desired conclusion for the contrapositive.

The matrix K has at least one nonzero entry, suppose it is located in column t. Let $\mathbf{x} \in \mathbb{C}^n$ be a vector of all zeros, except a 1 in entry t. Use this vector to form a linear combination of the columns of K, and the result will be just column t of K, which is nonzero. So by Theorem SLSLC [67], the vector $\mathbf{x}$ cannot be a solution to the homogenous system $\mathcal{LS}(K, \mathbf{0})$, so $\mathbf{x} \notin \mathrm{N}(K)$. ■

With this equivalence for nonsingular matrices we can update our list, Theorem NSME2 [93].

Theorem NSME3
NonSingular Matrix Equivalences, Round 3
Suppose that A is a square matrix of size n. The following are equivalent.

1. A is nonsingular.
2. A row-reduces to the identity matrix.
3. The null space of A contains only the zero vector, $\mathrm{N}(A) = \{\mathbf{0}\}$.
4. The linear system $\mathcal{LS}(A, \mathbf{b})$ has a unique solution for every possible choice of $\mathbf{b}$.
5. The columns of A are a linearly independent set.
6. The range of A is $\mathbb{C}^n$. □

Proof Since Theorem RNSM [117] is an equivalence, we can add it to the list in Theorem NSME2 [93]. ■

Section RSOM
Row Space Of a Matrix

The range of a matrix is sometimes called the **column space** since it is formed by taking all possible linear combinations of the columns of the matrix. We can do a smilar construction with the rows of a matrix, and that is the topic of this section. This will provide us with even more connections with row operations. However, we are also going to try to parlay our knowledge of the range, so we'll get at the rows of a matrix by working with the columns of the transpose. A side-benefit will be a third way to describe the range of a matrix.

Subsection RSM
Row Space of a Matrix

Definition RS
Row Space of a Matrix
Suppose A is an $m \times n$ matrix. Then the **row space** of A, $\operatorname{rs}(A)$, is the range of A^t, i.e. $\operatorname{rs}(A) = \mathcal{R}(A^t)$. $\odot$

Informally, the row space is the set of all linear combinations of the rows of A. However, we write the rows as column vectors, thus the necessity of using the transpose to make the rows into columns. With the row space defined in terms of the range, all of the results of Section ROM [104] can be applied to row spaces.

Notice that if A is a rectangular $m \times n$ matrix, then $\mathcal{R}(A) \subseteq \mathbb{C}^m$, while $\operatorname{rs}(A) \subseteq \mathbb{C}^n$ and the two sets are not comparable since they do not even hold objects of the same type. However, when A is square of size n, both $\mathcal{R}(A)$ and $\operatorname{rs}(A)$ are subsets of $\mathbb{C}^n$, though usually the sets will not be equal.

Example RSOM.RSAI
Row space of Archetype I
The coefficient matrix in Archetype I [201] is

$$I = \begin{bmatrix} 1 & 4 & 0 & -1 & 0 & 7 & -9 \\ 2 & 8 & -1 & 3 & 9 & -13 & 7 \\ 0 & 0 & 2 & -3 & -4 & 12 & -8 \\ -1 & -4 & 2 & 4 & 8 & -31 & 37 \end{bmatrix}.$$

To build the row space, we transpose the matrix,

$$I^t = \begin{bmatrix} 1 & 2 & 0 & -1 \\ 4 & 8 & 0 & -4 \\ 0 & -1 & 2 & 2 \\ -1 & 3 & -3 & 4 \\ 0 & 9 & -4 & 8 \\ 7 & -13 & 12 & -31 \\ -9 & 7 & -8 & 37 \end{bmatrix}$$

Then the columns of this matrix are used in a span to build the row space,

$$\operatorname{rs}(I) = \mathcal{R}\left(I^t\right) = \operatorname{Sp}\left(\left\{ \begin{bmatrix} 1 \\ 4 \\ 0 \\ -1 \\ 0 \\ 7 \\ -9 \end{bmatrix}, \begin{bmatrix} 2 \\ 8 \\ -1 \\ 3 \\ 9 \\ -13 \\ 7 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 2 \\ -3 \\ -4 \\ 12 \\ -8 \end{bmatrix}, \begin{bmatrix} -1 \\ -4 \\ 2 \\ 4 \\ 8 \\ -31 \\ 37 \end{bmatrix} \right\}\right).$$

However, we can use Theorem BROC [109] to get a slightly better description. First, row-reduce I^t,

$$\begin{bmatrix} \boxed{1} & 0 & 0 & -\frac{31}{7} \\ 0 & \boxed{1} & 0 & \frac{12}{7} \\ 0 & 0 & \boxed{1} & \frac{13}{7} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Since there are leading 1's in columns with indices $D = \{1, 2, 3\}$, the range of I^t can be spanned by just the first three columns of I^t,

$$\operatorname{rs}(I) = \mathcal{R}\left(I^t\right) = \operatorname{Sp}\left(\left\{ \begin{bmatrix} 1 \\ 4 \\ 0 \\ -1 \\ 0 \\ 7 \\ -9 \end{bmatrix}, \begin{bmatrix} 2 \\ 8 \\ -1 \\ 3 \\ 9 \\ -13 \\ 7 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 2 \\ -3 \\ -4 \\ 12 \\ -8 \end{bmatrix} \right\}\right).$$

The technique of Theorem RNS [114] could also have been applied to the matrix I^t, by adjoining the 7×7 identity matrix, I_7 and row-reducing the resulting 11×7 matrix. The 4×7 matrix K that results is

$$K = \begin{bmatrix} 1 & 0 & 0 & 0 & -\frac{3}{7} & -\frac{1}{7} & 0 \\ 0 & 1 & 0 & 0 & -\frac{12}{7} & -\frac{4}{7} & 0 \\ 0 & 0 & 1 & 0 & -\frac{3}{7} & -\frac{9}{14} & -\frac{1}{2} \\ 0 & 0 & 0 & 1 & \frac{1}{7} & \frac{13}{28} & \frac{1}{4} \end{bmatrix}.$$

Then $\text{rs}(I) = \mathcal{R}(I^t) = \mathcal{N}(K)$, and we could use Theorem SSNS [81] to express the null space of K as the span of three vectors, one for each free variable in the homogenous system $\mathcal{LS}(K, \mathbf{0})$. △

The row space would not be too interesting if it was simply the range of the transpose. However, when we do row operations on a matrix we have no effect on the many linear combinations that can be formed with the rows of the matrix. This is stated more carefully in the following theorem.

Theorem REMRS
Row-Equivalent Matrices have equal Row Spaces
Suppose A and B are row-equivalent matrices. Then $\text{rs}(A) = \text{rs}(B)$. □

Proof Two matrices are row-equivalent (Definition REM [22]) if one can be obtained from another by a sequence of possibly many row operations. We will prove the theorem for two matrices that differ by a single row operation, and then this result can be applied repeatedly to get the full statement of the theorem. The row spaces of A and B are spans of the columns of their transposes. For each row operation we perform on a matrix, we can define an analagous operation on the columns. Perhaps we should call these **column operations**. Instead, we will still call them row operations, but we will apply them to the columns of the transposes.

Refer to the columns of A^t and B^t as $\mathbf{A}_i$ and $\mathbf{B}_i$, $1 \leq i \leq m$. The row operation that switches rows will just switch columns of the transposed matrices. This will have no effect on the possible linear combinations formed by the columns.

Suppose that B^t is formed from A^t by multiplying column $\mathbf{A}_t$ by $\alpha \neq 0$. In other words, $\mathbf{B}_t = \alpha\mathbf{A}_t$, and $\mathbf{B}_i = \mathbf{A}_i$ for all $i \neq t$. We need to establish that two sets are equal, $\mathcal{R}(A^t) = \mathcal{R}(B^t)$. We will take a generic element of one and show that it is contained in the other.

$$\begin{aligned}
&\beta_1\mathbf{B}_1 + \beta_2\mathbf{B}_2 + \beta_3\mathbf{B}_3 + \cdots + \beta_t\mathbf{B}_t + \cdots + \beta_m\mathbf{B}_m = \\
&\beta_1\mathbf{A}_1 + \beta_2\mathbf{A}_2 + \beta_3\mathbf{A}_3 + \cdots + \beta_t(\alpha\mathbf{A}_t) + \cdots + \beta_m\mathbf{A}_m = \\
&\beta_1\mathbf{A}_1 + \beta_2\mathbf{A}_2 + \beta_3\mathbf{A}_3 + \cdots + (\alpha\beta_t)\mathbf{A}_t + \cdots + \beta_m\mathbf{A}_m
\end{aligned}$$

says that $\mathcal{R}(B^t) \subseteq \mathcal{R}(A^t)$. Similarly,

$$\begin{aligned}
&\gamma_1\mathbf{A}_1 + \gamma_2\mathbf{A}_2 + \gamma_3\mathbf{A}_3 + \cdots + \gamma_t\mathbf{A}_t + \cdots + \gamma_m\mathbf{A}_m = \\
&\gamma_1\mathbf{A}_1 + \gamma_2\mathbf{A}_2 + \gamma_3\mathbf{A}_3 + \cdots + \left(\frac{\gamma_t}{\alpha}\alpha\right)\mathbf{A}_t + \cdots + \gamma_m\mathbf{A}_m = \\
&\gamma_1\mathbf{A}_1 + \gamma_2\mathbf{A}_2 + \gamma_3\mathbf{A}_3 + \cdots + \frac{\gamma_t}{\alpha}(\alpha\mathbf{A}_t) + \cdots + \gamma_m\mathbf{A}_m = \\
&\gamma_1\mathbf{B}_1 + \gamma_2\mathbf{B}_2 + \gamma_3\mathbf{B}_3 + \cdots + \frac{\gamma_t}{\alpha}\mathbf{B}_t + \cdots + \gamma_m\mathbf{B}_m
\end{aligned}$$

says that $\mathcal{R}(A^t) \subseteq \mathcal{R}(B^t)$. So $\text{rs}(A) = \mathcal{R}(A^t) = \mathcal{R}(B^t) = \text{rs}(B)$ when a single row operation of the second type is performed.

Suppose now that B^t is formed from A^t by replacing $\mathbf{A}_t$ with $\alpha\mathbf{A}_s + \mathbf{A}_t$ for some $\alpha \in \mathbb{C}$ and $s \neq t$. In other words, $\mathbf{B}_t = \alpha\mathbf{A}_s + \mathbf{A}_t$, and $\mathbf{B}_i = \alpha\mathbf{A}_i$ for $i \neq t$.

$$
\begin{aligned}
&\beta_1\mathbf{B}_1 + \beta_2\mathbf{B}_2 + \beta_3\mathbf{B}_3 + \cdots + \beta_s\mathbf{B}_s + \cdots + \beta_t\mathbf{B}_t + \cdots + \beta_m\mathbf{B}_m = \\
&\beta_1\mathbf{A}_1 + \beta_2\mathbf{A}_2 + \beta_3\mathbf{A}_3 + \cdots + \beta_s\mathbf{A}_s + \cdots + \beta_t\left(\alpha\mathbf{A}_s + \mathbf{A}_t\right) + \cdots + \beta_m\mathbf{A}_m = \\
&\beta_1\mathbf{A}_1 + \beta_2\mathbf{A}_2 + \beta_3\mathbf{A}_3 + \cdots + \beta_s\mathbf{A}_s + \cdots + \left(\beta_t\alpha\right)\mathbf{A}_s + \beta_t\mathbf{A}_t + \cdots + \beta_m\mathbf{A}_m = \\
&\beta_1\mathbf{A}_1 + \beta_2\mathbf{A}_2 + \beta_3\mathbf{A}_3 + \cdots + \beta_s\mathbf{A}_s + \left(\beta_t\alpha\right)\mathbf{A}_s + \cdots + \beta_t\mathbf{A}_t + \cdots + \beta_m\mathbf{A}_m = \\
&\beta_1\mathbf{A}_1 + \beta_2\mathbf{A}_2 + \beta_3\mathbf{A}_3 + \cdots + \left(\beta_s + \beta_t\alpha\right)\mathbf{A}_s + \cdots + \beta_t\mathbf{A}_t + \cdots + \beta_m\mathbf{A}_m
\end{aligned}
$$

says that $\mathrm{R}\left(B^t\right) \subseteq \mathrm{R}\left(A^t\right)$. Similarly,

$$
\begin{aligned}
&\gamma_1\mathbf{A}_1 + \gamma_2\mathbf{A}_2 + \gamma_3\mathbf{A}_3 + \cdots + \gamma_s\mathbf{A}_s + \cdots + \gamma_t\mathbf{A}_t + \cdots + \gamma_m\mathbf{A}_m = \\
&\gamma_1\mathbf{A}_1 + \gamma_2\mathbf{A}_2 + \gamma_3\mathbf{A}_3 + \cdots + \gamma_s\mathbf{A}_s + \cdots + \left(-\alpha\gamma_t\mathbf{A}_s + \alpha\gamma_t\mathbf{A}_s\right) + \gamma_t\mathbf{A}_t + \cdots + \gamma_m\mathbf{A}_m = \\
&\gamma_1\mathbf{A}_1 + \gamma_2\mathbf{A}_2 + \gamma_3\mathbf{A}_3 + \cdots + \left(-\alpha\gamma_t\mathbf{A}_s\right) + \gamma_s\mathbf{A}_s + \cdots + \left(\alpha\gamma_t\mathbf{A}_s + \gamma_t\mathbf{A}_t\right) + \cdots + \gamma_m\mathbf{A}_m = \\
&\gamma_1\mathbf{A}_1 + \gamma_2\mathbf{A}_2 + \gamma_3\mathbf{A}_3 + \cdots + \left(-\alpha\gamma_t + \gamma_s\right)\mathbf{A}_s + \cdots + \gamma_t\left(\alpha\mathbf{A}_s + \mathbf{A}_t\right) + \cdots + \gamma_m\mathbf{A}_m = \\
&\gamma_1\mathbf{B}_1 + \gamma_2\mathbf{B}_2 + \gamma_3\mathbf{B}_3 + \cdots + \left(-\alpha\gamma_t + \gamma_s\right)\mathbf{B}_s + \cdots + \gamma_t\mathbf{B}_t + \cdots + \gamma_m\mathbf{B}_m
\end{aligned}
$$

says that $\mathrm{R}\left(A^t\right) \subseteq \mathrm{R}\left(B^t\right)$. So $\mathrm{rs}\left(A\right) = \mathrm{R}\left(A^t\right) = \mathrm{R}\left(B^t\right) = \mathrm{rs}\left(B\right)$ when a single row operation of the third type is performed.

So the row space is preserved by each row operation, and hence row spaces of row-equivalent matrices are equal. ■

Example RSOM.RSREM
Row spaces of two row-equivalent matrices
In Example RREF.TREM [22] we saw that the matrices

$$
A = \begin{bmatrix} 2 & -1 & 3 & 4 \\ 5 & 2 & -2 & 3 \\ 1 & 1 & 0 & 6 \end{bmatrix} \qquad B = \begin{bmatrix} 1 & 1 & 0 & 6 \\ 3 & 0 & -2 & -9 \\ 2 & -1 & 3 & 4 \end{bmatrix}
$$

are row-equivalent by demonstrating a sequence of two row operations that converted A into B. Applying Theorem REMRS [121] we can say

$$
\mathrm{rs}\left(A\right) = \mathrm{Sp}\left(\left\{\begin{bmatrix} 2 \\ -1 \\ 3 \\ 4 \end{bmatrix}, \begin{bmatrix} 5 \\ 2 \\ -2 \\ 3 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \\ 6 \end{bmatrix}\right\}\right) = \mathrm{Sp}\left(\left\{\begin{bmatrix} 1 \\ 1 \\ 0 \\ 6 \end{bmatrix}, \begin{bmatrix} 3 \\ 0 \\ -2 \\ -9 \end{bmatrix}, \begin{bmatrix} 2 \\ -1 \\ 3 \\ 4 \end{bmatrix}\right\}\right) = \mathrm{rs}\left(B\right) \quad \triangle
$$

Theorem REMRS [121] is at its best when one of the row-equivalent matrices is in reduced row-echelon form. The vectors that correspond to the zero rows can be ignored (who needs the zero vector when building a span?). The echelon pattern insures that the nonzero rows yield vectors that are linearly independent. Here's the theorem.

Theorem BRS
Basis for the Row Space
Suppose that A is a matrix and B is a row-equivalent matrix in reduced row-echelon form. Let S be the set of nonzero columns of B^t. Then

1. $\mathrm{rs}\left(A\right) = \mathrm{Sp}\left(S\right)$.
2. S is a linearly independent set. □

Proof From Theorem REMRS [121] we know that $\mathrm{R}\left(A\right) = \mathrm{R}\left(B\right)$. If B has any zero rows, these correspond to columns of B^t that are the zero vector. We can safely toss out the zero vector in the span construction, since it can be recreated from the nonzero vectors by a linear combination where all the scalars are zero.

Suppose B has r nonzero rows and let $D = \{d_1, d_2, d_3, \ldots, d_r\}$ denote the column indices of B that have a leading one in them. Denote the r column vectors of B^t, the vectors in S, as $\mathbf{B}_1, \mathbf{B}_2, \mathbf{B}_3, \ldots, \mathbf{B}_r$. To show that S is linearly independent, start with a relation of linear dependence

$$\alpha_1\mathbf{B}_1 + \alpha_2\mathbf{B}_2 + \alpha_3\mathbf{B}_3 + \cdots + \alpha_r\mathbf{B}_r = \mathbf{0}$$

Now consider this equation across entries of the vectors in location d_i, $1 \leq i \leq r$. Since B is in reduced row-echelon form, the entries of column d_i are all zero, except for a (leading) 1 in row i. Considering the column vectors of B^t, the linear combination for entry d_i is

$$\alpha_1(0) + \alpha_2(0) + \alpha_3(0) + \cdots + \alpha_i(1) + \cdots + \alpha_r(0) = 0$$

and from this we conclude that $\alpha_i = 0$ for all $1 \leq i \leq r$, establishing the linear independence of S. ■

Example RSOM.IS
Improving a span
Suppose in the course of analyzing a matrix (its range, its null space, its. . .) we encounter the following set of vectors, described by a span

$$X = \mathrm{Sp}\left(\left\{\begin{bmatrix}1\\2\\1\\6\\6\end{bmatrix}, \begin{bmatrix}3\\-1\\2\\-1\\6\end{bmatrix}, \begin{bmatrix}1\\-1\\0\\-1\\-2\end{bmatrix}, \begin{bmatrix}-3\\2\\-3\\6\\-10\end{bmatrix}\right\}\right)$$

Let A be the matrix whose rows are the vectors in X, so by design $\mathrm{rs}\left(A\right) = X$,

$$A = \begin{bmatrix}1 & 2 & 1 & 6 & 6\\3 & -1 & 2 & -1 & 6\\1 & -1 & 0 & -1 & -2\\-3 & 2 & -3 & 6 & -10\end{bmatrix}$$

Row-reduce A to form a row-equivalent matrix in reduced row-echelon form,

$$B = \begin{bmatrix} \boxed{1} & 0 & 0 & 2 & -1 \\ 0 & \boxed{1} & 0 & 3 & 1 \\ 0 & 0 & \boxed{1} & -2 & 5 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Then Theorem BRS [123] says we can grab the nonzero columns of B^t and write

$$X = \mathrm{rs}\left(A\right) = \mathrm{rs}\left(B\right) = \mathrm{Sp}\left(\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \\ 2 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ 3 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ -2 \\ 5 \end{bmatrix} \right\}\right)$$

These three vectors provide a much-improved description of X. There are fewer vectors, and the pattern of zeros and ones in the first three entries makes it easier to determine membership in X. And all we had to do was row-reduce the right matrix and toss out a zero row. Next to row operations themselves, this is probably the most powerful computational technique at your disposal. △

Theorem BRS [123] and the techniques of Example RSOM.IS [123] will provide yet another description of the range of a matrix. First we state a triviality as a theorem, so we can reference it later

Theorem RMSRT
Range of a Matrix is Row Space of Transpose
Suppose A is a matrix. Then $\mathrm{R}\left(A\right) = \mathrm{rs}\left(A^t\right)$. □

Proof Apply Theorem TASM [103] with Definition RS [119],

$$\mathrm{rs}\left(A^t\right) = \mathrm{R}\left(\left(A^t\right)^t\right) = \mathrm{R}\left(A\right).$$ ■

So to find yet another expression for the range of a matrix, build its transpose, row-reduce it, toss out the zero rows, and convert the nonzero rows to column vectors to yield an improved spanning set. We'll do Archetype I [201], then you do Archetype J [206].

Example RSOM.RROI
Range from row operations, Archetype I
To find the range of the coefficient matrix of Archetype I [201], we proceed as follows. The matrix is

$$I = \begin{bmatrix} 1 & 4 & 0 & -1 & 0 & 7 & -9 \\ 2 & 8 & -1 & 3 & 9 & -13 & 7 \\ 0 & 0 & 2 & -3 & -4 & 12 & -8 \\ -1 & -4 & 2 & 4 & 8 & -31 & 37 \end{bmatrix}.$$

The transpose is

$$\begin{bmatrix} 1 & 2 & 0 & -1 \\ 4 & 8 & 0 & -4 \\ 0 & -1 & 2 & 2 \\ -1 & 3 & -3 & 4 \\ 0 & 9 & -4 & 8 \\ 7 & -13 & 12 & -31 \\ -9 & 7 & -8 & 37 \end{bmatrix}.$$

Row-reduced this becomes,

$$\begin{bmatrix} 1 & 0 & 0 & -\frac{31}{7} \\ 0 & 1 & 0 & \frac{12}{7} \\ 0 & 0 & 1 & \frac{13}{7} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Now, using Theorem RMSRT [124] and Theorem BRS [123]

$$\mathcal{R}(I) = \mathrm{rs}\left(I^t\right) = \mathrm{Sp}\left(\left\{\begin{bmatrix} 1 \\ 0 \\ 0 \\ -\frac{31}{7} \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ \frac{12}{7} \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ \frac{13}{7} \end{bmatrix}\right\}\right).$$

This is a very nice description of the range. Fewer vectors than the 7 involved in the definition, and the structure of the zeros and ones in the first 3 slots can be use to advantage. For example, Archetype I [201] is presented as consistent system of equations with a vector of constants

$$\mathbf{b} = \begin{bmatrix} 3 \\ 9 \\ 1 \\ 4 \end{bmatrix}.$$

Since $\mathcal{LS}(I, \mathbf{b})$ is consistent, Theorem RCS [105] tells us that $\mathbf{b} \in \mathcal{R}(I)$. But we could see this quickly with the following computation, which really only involves any work in the 4th entry of the vectors as the scalars in the linear combination are *dictated* by the first three entries of $\mathbf{b}$.

$$\mathbf{b} = \begin{bmatrix} 3 \\ 9 \\ 1 \\ 4 \end{bmatrix} = 3\begin{bmatrix} 1 \\ 0 \\ 0 \\ -\frac{31}{7} \end{bmatrix} + 9\begin{bmatrix} 0 \\ 1 \\ 0 \\ \frac{12}{7} \end{bmatrix} + 1\begin{bmatrix} 0 \\ 0 \\ 1 \\ \frac{13}{7} \end{bmatrix}$$

Can you now rapidly construct several vectors, $\mathbf{b}$, so that $\mathcal{LS}(I, \mathbf{b})$ is consistent, and several more so that the system is inconsistent? △

Example ROM.COC [106] and Example RSOM.RROI [124] each describes the range of the coefficient matrix from Archetype I [201] as the span of a set of $r = 3$

linearly independent vectors. It is no accident that these two different sets both have the same size. If we (you?) were to calculate the range of this matrix using the null space of the matrix K from Theorem RNS [114] then we would again find a set of 3 linearly independent vectors that span the range. More on this later.

So we have three different methods to obtain a description of the range of a matrix as the span of a linearly independent set. Theorem BROC [109] is sometimes useful since the vectors it specifies are equal to actual columns of the matrix. Theorem RNS [114] tends to create vectors with lots of zeros, and strategically placed 1's, near the end of the vector. Finally, Theorem BRS [123] and Theorem RMSRT [124] combine to create vectors with lots of zeros, and strategically placed 1's, near the front of the vector.

Section MOM
Multiplication of Matrices

We know how to add vectors and how to multiply them by scalars. Together, these operations give us the possibility of making linear combinations. Similarly, we know how to add matrices and how to multiply matrices by scalars. In this section we mix all these ideas together and produce an operation known as matrix multiplication. This will lead to some results that are both surprising and central. We begin with a definition of how to multiply a vector by a matrix.

Subsection MVP
Matrix-Vector Product

We have repeatedly seen the importance of forming linear combinations of the columns of a matrix. As one example of this, Theorem SLSLC [67] said that every solution to a system of linear equations gives rise to a linear combination of the column vectors of the coefficient matrix that equals the vector of constants. This theorem, and others, motivates the following central definition.

Definition MVP
Matrix-Vector Product
Suppose A is an $m \times n$ matrix with columns $\mathbf{A}_1, \mathbf{A}_2, \mathbf{A}_3, \ldots, \mathbf{A}_n$ and $\mathbf{u}$ is a vector of size n. Then the **matrix-vector product** of A with $\mathbf{u}$ is

$$A\mathbf{u} = [\mathbf{A}_1|\mathbf{A}_2|\mathbf{A}_3|\ldots|\mathbf{A}_n] \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ \vdots \\ u_n \end{bmatrix} = u_1\mathbf{A}_1 + u_2\mathbf{A}_2 + u_3\mathbf{A}_3 + \cdots + u_n\mathbf{A}_n \qquad \odot$$

So, the matrix-vector product is yet another version of "multiplication," at least in the sense that we have yet again overloaded concatenation of two symbols. Remember your objects, an $m \times n$ matrix times a vector of size n will create a vector of size m. So if A is rectangular, then the size of the vector changes. With all the linear combinations we have performed so far, this computation should now seem second nature.

Example MOM.MTV
A matrix times a vector

Consider

$$A = \begin{bmatrix} 1 & 4 & 2 & 3 & 4 \\ -3 & 2 & 0 & 1 & -2 \\ 1 & 6 & -3 & -1 & 5 \end{bmatrix} \qquad \mathbf{u} = \begin{bmatrix} 2 \\ 1 \\ -2 \\ 3 \\ -1 \end{bmatrix}$$

Then

$$A\mathbf{u} = 2\begin{bmatrix} 1 \\ -3 \\ 1 \end{bmatrix} + 1\begin{bmatrix} 4 \\ 2 \\ 6 \end{bmatrix} + (-2)\begin{bmatrix} 2 \\ 0 \\ -3 \end{bmatrix} + 3\begin{bmatrix} 3 \\ 1 \\ -1 \end{bmatrix} + (-1)\begin{bmatrix} 4 \\ -2 \\ 5 \end{bmatrix} = \begin{bmatrix} 7 \\ 1 \\ 6 \end{bmatrix}. \qquad \triangle$$

This definition now makes it possible to represent systems of linear equations compactly in terms of an operation.

Theorem SLEMM
Systems of Linear Equations as Matrix Multiplication
Solutions to the linear system $\mathcal{LS}(A, \mathbf{b})$ are the solutions for $\mathbf{x}$ in the vector equation $A\mathbf{x} = \mathbf{b}$. □

Proof This theorem says (not very clearly) that two sets (of solutions) are equal. So we need to show that one set of solutions is a subset of the other, and vice versa (recall Technique SE [13]). Both of these inclusions are easy with the following chain of equivalences,

$$\begin{aligned} &\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_i \end{bmatrix} \text{ is a solution to } \mathcal{LS}(A, \mathbf{b}) \\ \iff &x_1\mathbf{A}_1 + x_2\mathbf{A}_2 + x_3\mathbf{A}_3 + \cdots + x_n\mathbf{A}_n = \mathbf{b} && \text{Theorem SLSLC [67]} \\ \iff &\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_i \end{bmatrix} \text{ is a solution to } A\mathbf{x} = \mathbf{b} && \text{Definition MVP [127].} \end{aligned}$$

■

Example MOM.NSLE
Notation for systems of linear equations

Consider the system of linear equations from Example HSE.NSLE [47].

$$\begin{aligned} 2x_1 + 4x_2 - 3x_3 + 5x_4 + x_5 &= 9 \\ 3x_1 + x_2 + \quad + x_4 - 3x_5 &= 0 \\ -2x_1 + 7x_2 - 5x_3 + 2x_4 + 2x_5 &= -3 \end{aligned}$$

has coefficient matrix

$$A = \begin{bmatrix} 2 & 4 & -3 & 5 & 1 \\ 3 & 1 & 0 & 1 & -3 \\ -2 & 7 & -5 & 2 & 2 \end{bmatrix}$$

and vector of constants

$$\mathbf{b} = \begin{bmatrix} 9 \\ 0 \\ -3 \end{bmatrix}$$

and so will be described compactly by the equation $A\mathbf{x} = \mathbf{b}$. △

Subsection MM
Matrix Multiplication

We now define how to multiply two matrices together. Stop for a minute and think about how you might define this new operation. Many books would present this definition much earlier in the course. However, we have taken great care to delay it as long as possible and to present as many ideas as practical based mostly on the notion of linear combinations. Towards the conclusion of the course, or when you perhaps take a second course in linear algebra, you may be in a position to appreciate the reasons for this. For now, understand that matrix multiplication is a central definition and perhaps you will appreciate its importance more by having saved it for later.

Definition MM
Matrix Multiplication
Suppose A is an $m \times n$ matrix and B is an $n \times p$ matrix with columns $\mathbf{B}_1, \mathbf{B}_2, \mathbf{B}_3, \ldots, \mathbf{B}_p$. Then the **matrix product** of A with B is the $m \times p$ matrix where column i is the matrix-vector product $A\mathbf{B}_i$. Symbolically,

$$AB = A\left[\mathbf{B}_1|\mathbf{B}_2|\mathbf{B}_3|\ldots|\mathbf{B}_p\right] = \left[A\mathbf{B}_1|A\mathbf{B}_2|A\mathbf{B}_3|\ldots|A\mathbf{B}_p\right].$$ ⊙

Example MOM.PTM
Product of two matrices

Set

$$A = \begin{bmatrix} 1 & 2 & -1 & 4 & 6 \\ 0 & -4 & 1 & 2 & 3 \\ -5 & 1 & 2 & -3 & 4 \end{bmatrix} \qquad B = \begin{bmatrix} 1 & 6 & 2 & 1 \\ -1 & 4 & 3 & 2 \\ 1 & 1 & 2 & 3 \\ 6 & 4 & -1 & 2 \\ 1 & -2 & 3 & 0 \end{bmatrix}$$

Then

$$AB = \left[A\begin{bmatrix} 1 \\ -1 \\ 1 \\ 6 \\ 1 \end{bmatrix} \middle| A\begin{bmatrix} 6 \\ 4 \\ 1 \\ 4 \\ -2 \end{bmatrix} \middle| A\begin{bmatrix} 2 \\ 3 \\ 2 \\ -1 \\ 3 \end{bmatrix} \middle| A\begin{bmatrix} 1 \\ 2 \\ 3 \\ 2 \\ 0 \end{bmatrix} \right] = \begin{bmatrix} 28 & 17 & 20 & 10 \\ 20 & -13 & -3 & -1 \\ -18 & -44 & 12 & -3 \end{bmatrix}. \qquad \triangle$$

Is this the definition of matrix multiplication you expected? Perhaps our previous operations for matrices caused you to think that we might multiply two matrices of the *same* size, *entry-by-entry*? Notice that our current definition uses matrices of different sizes (though the number of columns in the first must equal the number of rows in the second), and the result is of a third size. Notice too in the previous example that we cannot even consider the product BA, since the sizes of the two matrices in this order aren't right.

But it gets weirder than that. Many of your old ideas about "multiplication" won't apply to matrix multiplication, but some still will. So make no assumptions, and don't do anything until you have a theorem that says you can. Even if the sizes are right, matrix multiplication is not commutative — order matters.

Example MOM.MMNC
Matrix Multiplication is not commutative
Set

$$A = \begin{bmatrix} 1 & 3 \\ -1 & 2 \end{bmatrix} \qquad B = \begin{bmatrix} 4 & 0 \\ 5 & 1 \end{bmatrix}.$$

Then we have two square, 2×2 matrices, so Definition MM [129] allows us to multiply them in either order. We find

$$AB = \begin{bmatrix} 19 & 3 \\ 6 & 2 \end{bmatrix} \qquad BA = \begin{bmatrix} 4 & 12 \\ 4 & 17 \end{bmatrix}$$

and $AB \neq BA$. Not even close. It should not be hard for you to construct other pairs of matrices that do not commute (try a couple of 3×3's). Can you find a pair of non-identical matrices that *do* commute? $\triangle$

Subsection MMEE
Matrix Multiplication, Entry-by-Entry

While certain "natural" properties of multiplication don't hold, many more do. In the next subsection, we'll state and prove the relevant theorems. But first, we need a theorem that provides an alternate means of multiplying two matrices. In many texts, this would be given as the *definition* of matrix multiplication. We prefer to turn it around and have the following formula as a consequence of the definition. It will prove useful for proofs of matrix equality, where we need to examine products of matrices, entry-by-entry.

Theorem EMP
Entries of matrix products
Suppose $A = (a_{ij})$ is an $m \times n$ matrix and $B = (b_{ij})$ is an $n \times p$ matrix. Then the entries of $AB = C = (c_{ij})$ are given by

$$[C]_{ij} = c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + a_{i3}b_{3j} + \cdots + a_{in}b_{nj} = \sum_{k=1}^{n} a_{ik}b_{kj} = \sum_{k=1}^{n} [A]_{ik} [B]_{kj} \qquad \square$$

Proof The value of c_{ij} lies in column j of the product of A and B, and so by Definition MM [129] is the value in location i of the matrix-vector product $A\mathbf{B}_j$. By Definition MVP [127] this matrix-vector product is a linear combination

$$\begin{aligned} A\mathbf{B}_j &= b_{1j}\mathbf{A}_1 + b_{2j}\mathbf{A}_2 + b_{3j}\mathbf{A}_3 + \cdots + b_{nj}\mathbf{A}_n \\ &= b_{1j}\begin{bmatrix} a_{11} \\ a_{21} \\ a_{31} \\ \vdots \\ a_{m1} \end{bmatrix} + b_{2j}\begin{bmatrix} a_{12} \\ a_{22} \\ a_{32} \\ \vdots \\ a_{m2} \end{bmatrix} + b_{3j}\begin{bmatrix} a_{13} \\ a_{23} \\ a_{33} \\ \vdots \\ a_{m3} \end{bmatrix} + \cdots + b_{nj}\begin{bmatrix} a_{1n} \\ a_{2n} \\ a_{3n} \\ \vdots \\ a_{mn} \end{bmatrix} \end{aligned}$$

We are after the value in location i of this linear combination. Using Definition CVA [59] and Definition CVSM [60] we course through this linear combination in location i to find

$$b_{1j}a_{i1} + b_{2j}a_{i2} + b_{3j}a_{i3} + \cdots + b_{nj}a_{in}.$$

Reversing the order of the products (regular old multiplication *is* commutative) yields the desired expression for c_{ij}. ■

Example MOM.PTMEE
Product of two matrices, entry-by-entry

Consider again the two matrices from EXAMPLE MOM.PTM [129]

$$A = \begin{bmatrix} 1 & 2 & -1 & 4 & 6 \\ 0 & -4 & 1 & 2 & 3 \\ -5 & 1 & 2 & -3 & 4 \end{bmatrix} \qquad B = \begin{bmatrix} 1 & 6 & 2 & 1 \\ -1 & 4 & 3 & 2 \\ 1 & 1 & 2 & 3 \\ 6 & 4 & -1 & 2 \\ 1 & -2 & 3 & 0 \end{bmatrix}$$

Then suppose we just wanted the entry of AB in the second row, third column:

$$\begin{aligned} [AB]_{23} =& a_{21}b_{13} + a_{22}b_{23} + a_{23}b_{33} + a_{24}b_{43} + a_{25}b_{53} \\ =& (0)(2) + (-4)(3) + (1)(2) + (2)(-1) + (3)(3) = -3 \end{aligned}$$

Notice how there are 5 terms in the sum, since 5 is the common dimension of the two matrices (column count for A, row count for B). In the conclusion of THEOREM EMP [131], it would be the index k that would run from 1 to 5 in this computation. Here's a bit more practice.

The entry of third row, first column:

$$\begin{aligned} [AB]_{31} =& a_{31}b_{11} + a_{32}b_{21} + a_{33}b_{31} + a_{34}b_{41} + a_{35}b_{51} \\ =& (-5)(1) + (1)(-1) + (2)(1) + (-3)(6) + (4)(1) = -18 \end{aligned}$$

To get some more practice on your own, complete the computation of the other 10 entries of this product. Construct some other pairs of matrices (of compatible sizes) and compute their product two ways. First use DEFINITION MM [129]. Since linear combinations are straightforward for you now, this should be easy to do and to do correctly. Then do it again, using THEOREM EMP [131]. Since this process may take some practice, use your first computation to check your work. △

THEOREM EMP [131] is the way most people compute matrix products by hand. It will also be very useful for the theorems we are going to prove shortly. However, the definition is frequently the most useful for its connections with deeper ideas like the nullspace and range. For example, an alternative (and popular) definition of the range of an $m \times n$ matrix A would be

$$\mathrm{R}\left(A\right) = \left\{A\mathbf{x} \mid \mathbf{x} \in \mathbb{C}^n\right\}.$$

We recognize this as saying take *all* the matrix vector products possible with the matrix A. By DEFINITION MVP [127] we see that this means take all possible linear combinations of the columns of A — precisely our version of the definition of the range (DEFINITION RM [104]).

Subsection PMM
Properties of Matrix Multiplication

In this subsection, we collect properties of matrix multiplication and its interaction with matrix addition (DEFINITION MA [98]), scalar matrix multiplication (DEFINItion SMM [98]), the identity matrix (Definition IM [50]), the zero matrix (Definition ZM [101]) and the transpose (Definition TM [101]). Whew! Here we go. These are great proofs to practice with, so try to concoct the proofs before reading them, they'll get progressively harder as we go.

Theorem MMZM
Matrix Multiplication and the Zero Matrix
Suppose A is an $m \times n$ matrix. Then
1. $A\mathcal{O}_{n\times p} = \mathcal{O}_{m\times p}$
2. $\mathcal{O}_{p\times m}A = \mathcal{O}_{p\times n}$ □

Proof We'll prove (1) and leave (2) to you. Entry-by-entry,

$$\begin{aligned}
\left[A\mathcal{O}_{n\times p}\right]_{ij} &= \sum_{k=1}^{n} [A]_{ik} \left[\mathcal{O}_{n\times p}\right]_{kj} && \text{Definition MM [129]}\\
&= \sum_{k=1}^{n} [A]_{ik}\, 0 && \text{Definition ZM [101]}\\
&= \sum_{k=1}^{n} 0 = 0.
\end{aligned}$$

So every entry of the product is the scalar zero, i.e. the result is the zero matrix. ■

Theorem MMIM
Matrix Multiplication and Identity Matrix
Suppose A is an $m \times n$ matrix. Then
1. $AI_n = A$
2. $I_mA = A$ □

Proof Again, we'll prove (1) and leave (2) to you. Entry-by-entry,

$$\begin{aligned}
\left[AI_n\right]_{ij} &= \sum_{k=1}^{n} [A]_{ik} \left[I_n\right]_{kj} && \text{Definition MM [129]}\\
&= [A]_{ij} \left[I_n\right]_{jj} + \sum_{k=1,k\neq j}^{n} [A]_{ik} \left[I_n\right]_{kj}\\
&= [A]_{ij} (1) + \sum_{k=1,k\neq j}^{n} [A]_{ik} (0) && \text{Definition IM [50]}\\
&= [A]_{ij} + \sum_{k=1,k\neq j}^{n} 0\\
&= [A]_{ij}
\end{aligned}$$

So the matrices A and AI_n are equal, entry-by-entry, and by the definition of matrix equality (Definition ME [97]) we can say they are equal matrices. ■

It is this theorem that gives the identity matrix its name. It is a matrix that behaves with matrix multiplication like the scalar 1 does with scalar multiplication. To multiply by the identity matrix is to have no effect on the other matrix.

Theorem MMDAA
Matrix Multiplication Distributes Across Addition
Suppose A is an $m \times n$ matrix and B and C are $n \times p$ matrices and D is a $p \times s$ matrix. Then
1. $A(B+C) = AB + AC$
2. $(B+C)D = BD + CD$ □

Proof We'll do (1), you do (2). Entry-by-entry,

$$\begin{aligned}
[A(B+C)]_{ij} &= \sum_{k=1}^{n} [A]_{ik} [B+C]_{kj} && \text{Definition MM [129]}\\
&= \sum_{k=1}^{n} [A]_{ik} ([B]_{kj} + [C]_{kj}) && \text{Definition MA [98]}\\
&= \sum_{k=1}^{n} [A]_{ik} [B]_{kj} + [A]_{ik} [C]_{kj}\\
&= \sum_{k=1}^{n} [A]_{ik} [B]_{kj} + \sum_{k=1}^{n} [A]_{ik} [C]_{kj}\\
&= [AB]_{ij} + [AC]_{ij} && \text{Definition MM [129]}\\
&= [AB + AC]_{ij} && \text{Definition MA [98]}
\end{aligned}$$

So the matrices $A(B+C)$ and $AB+AC$ are equal, entry-by-entry, and by the definition of matrix equality (Definition ME [97]) we can say they are equal matrices. ■

Theorem MMSMM
Matrix Multiplication and Scalar Matrix Multiplication
Suppose A is an $m \times n$ matrix and B is an $n \times p$ matrix. Let α be a scalar. Then $\alpha(AB) = (\alpha A)B = A(\alpha B)$. □

Proof These are equalities of matrices. We'll do the first one, the second is similar and will be good practice for you.

$$\begin{aligned}
[\alpha(AB)]_{ij} &= \alpha\,[AB]_{ij} && \text{Definition SMM [98]}\\
&= \alpha \sum_{k=1}^{n} [A]_{ik}\,[B]_{kj} && \text{Definition MM [129]}\\
&= \sum_{k=1}^{n} \alpha\,[A]_{ik}\,[B]_{kj} &&\\
&= \sum_{k=1}^{n} [\alpha A]_{ik}\,[B]_{kj} && \text{Definition SMM [98]}\\
&= [(\alpha A)B]_{ij} && \text{Definition MM [129]}
\end{aligned}$$

So the matrices $\alpha(AB)$ and $(\alpha A)B$ are equal, entry-by-entry, and by the definition of matrix equality (Definition ME [97]) we can say they are equal matrices. ■

Theorem MMA
Matrix Multiplication is Associative
Suppose A is an $m \times n$ matrix, B is an $n \times p$ matrix and D is a $p \times s$ matrix. Then $A(BD) = (AB)D$. □

Proof A matrix equality, so we'll go entry-by-entry, no surprise there.

$$\begin{aligned}
[A(BD)]_{ij} &= \sum_{k=1}^{n} [A]_{ik}\,[BD]_{kj} && \text{Definition MM [129]}\\
&= \sum_{k=1}^{n} [A]_{ik} \left(\sum_{\ell=1}^{p} [B]_{k\ell}\,[D]_{\ell j} \right) && \text{Definition MM [129]}\\
&= \sum_{k=1}^{n} \sum_{\ell=1}^{p} [A]_{ik}\,[B]_{k\ell}\,[D]_{\ell j} &&
\end{aligned}$$

We can switch the order of the summation since these are finite sums,

$$= \sum_{\ell=1}^{p} \sum_{k=1}^{n} [A]_{ik}\,[B]_{k\ell}\,[D]_{\ell j}$$

As $[D]_{\ell j}$ does not depend on the index k, we can factor it out of the inner sum,

$$\begin{aligned}
&= \sum_{\ell=1}^{p} [D]_{\ell j} \left(\sum_{k=1}^{n} [A]_{ik} [B]_{k\ell} \right) \\
&= \sum_{\ell=1}^{p} [D]_{\ell j} [AB]_{i\ell} && \text{Definition MM [129]} \\
&= \sum_{\ell=1}^{p} [AB]_{i\ell} [D]_{\ell j} \\
&= [(AB)D]_{ij} && \text{Definition MM [129]}
\end{aligned}$$

So the matrices $(AB)D$ and $A(BD)$ are equal, entry-by-entry, and by the definition of matrix equality (Definition ME [97]) we can say they are equal matrices. ■

One more theorem in this style, and its a good one. If you've been practicing with the previous proofs you should be able to do this one yourself.

Theorem MMT
Matrix Multiplication and Transposes
Suppose A is an $m \times n$ matrix and B is an $n \times p$ matrix. Then $(AB)^t = B^t A^t$. □

Proof This theorem may be surprising but if we check the sizes of the matrices involved, then maybe it will not seem so far-fetched. First, AB has size $m \times p$, so its transpose has size $p \times m$. The product of B^t with A^t is a $p \times n$ matrix times an $n \times m$ matrix, also resulting in a $p \times m$ matrix. So at least our objects are compatible for equality (and would not be, in general, if we didn't reverse the order of the operation).

Here we go again, entry-by-entry,

$$\begin{aligned}
\left[(AB)^t\right]_{ij} &= [AB]_{ji} && \text{Definition TM [101]} \\
&= \sum_{k=1}^{n} [A]_{jk} [B]_{ki} && \text{Definition MM [129]} \\
&= \sum_{k=1}^{n} [B]_{ki} [A]_{jk} \\
&= \sum_{k=1}^{n} \left[B^t\right]_{ik} \left[A^t\right]_{kj} && \text{Definition TM [101]} \\
&= \left[B^t A^t\right]_{ij} && \text{Definition MM [129]}
\end{aligned}$$

So the matrices $(AB)^t$ and $B^t A^t$ are equal, entry-by-entry, and by the definition of matrix equality (Definition ME [97]) we can say they are equal matrices. ■

This theorem seems odd at first glance, since we have to switch the order of A and B. But if we simply consider the sizes of the matrices involved, we can see that the switch is necessary for this reason alone. That the individual entries of the products then come along is a bonus.

Notice how none of these proofs above relied on writing out huge general matrices with lots of ellipses ("...") and trying to formulate the equalities a whole matrix at a time. This messy business is a "proof technique" to be avoided at all costs.

These theorems, along with Theorem VSPM [100], give you the "rules" for how matrices interact with the various operations we have defined. Use them and use them often. But don't try to do anything with a matrix that you don't have a rule for. Together, we would infomally call all these operations, and the attendant theorems, "the algebra of matrices." Notice, too, that every column vector is just a $n \times 1$ matrix, so these theorems apply to column vectors also. Finally, these results may make us feel that the definition of matrix multiplication is not so unnatural.

Subsection PSHS
Particular Solutions, Homogenous Solutions

Having delayed presenting matrix multiplication, we have one theorem we could have stated long ago, but its proof is much easier now that we know how to represent a system of linear equations with matrix multiplication and how to mix matrix multiplication with other operations.

The next theorem tells us that in order to find all of the solutions to a linear system of equations, it is sufficient to find just one solution, and then find all of the solutions to the corresponding homogenous system. This explains part of our interest in the null space, the set of all solutions to a homogenous system.

Theorem PSPHS
Particular Solution Plus Homogenous Solutions
Suppose that $\mathbf{z}$ is one solution to the linear system of equations $\mathcal{LS}(A,\,b)$. Then $\mathbf{y}$ is a solution to $\mathcal{LS}(A,\,b)$ if and only if $\mathbf{y} = \mathbf{z} + \mathbf{w}$ for some vector $\mathbf{w} \in N(A)$. □

Proof We will work with the vector equality representations of the relevant systems of equations, as described by Theorem SLEMM [128].

($\Leftarrow$) Suppose $\mathbf{y} = \mathbf{z} + \mathbf{w}$ and $\mathbf{w} \in N(A)$. Then

$$A\mathbf{y} = A(\mathbf{z} + \mathbf{w}) = A\mathbf{z} + A\mathbf{w} = \mathbf{b} + \mathbf{0} = \mathbf{b}$$

demonstrating that $\mathbf{y}$ is a solution.

($\Rightarrow$) Suppose $\mathbf{y}$ is a solution to $\mathcal{LS}(A,\,b)$. Then

$$A(\mathbf{y} - \mathbf{z}) = A\mathbf{y} - A\mathbf{z} = \mathbf{b} - \mathbf{b} = \mathbf{0}$$

which says that $\mathbf{y} - \mathbf{z} \in \mathrm{N}(A)$. In other words, $\mathbf{y} - \mathbf{z} = \mathbf{w}$ for some vector $\mathbf{w} \in N(A)$. Rewritten, this is $\mathbf{y} = \mathbf{z} + \mathbf{w}$, as desired. ■

Example MOM.PSNS
Particular solutions, homogenous solutions, Archetype D
Archetype D [179] is a consistent system of equations with a nontrivial null space. The write-up for this system begins with three solutions,

$$\mathbf{y}_1 = \begin{bmatrix}0\\1\\2\\1\end{bmatrix} \qquad \mathbf{y}_2 = \begin{bmatrix}4\\0\\0\\0\end{bmatrix} \qquad \mathbf{y}_3 = \begin{bmatrix}7\\8\\1\\3\end{bmatrix}$$

We will choose to have $\mathbf{y}_1$ play the role of $\mathbf{z}$ in the statement of Theorem PSPHS [137], any one of the three vectors listed here (or others) could have been chosen. To illustrate the theorem, we should be able to write each of these three solutions as the vector $\mathbf{z}$ plus a solution to the corresponding homogenous system of equations. Since $\mathbf{0}$ is always a solution to a homogenous system we can easily write

$$\mathbf{y}_1 = \mathbf{z} = \mathbf{z} + \mathbf{0}.$$

The vectors $\mathbf{y}_2$ and $\mathbf{y}_3$ will require a bit more effort. Solutions to the homogenous system are exactly the elements of the null space of the coefficient matrix, which is

$$\mathrm{Sp}\left(\left\{\begin{bmatrix}-3\\-1\\1\\0\end{bmatrix}, \begin{bmatrix}2\\3\\0\\1\end{bmatrix}\right\}\right)$$

Then

$$\mathbf{y}_2 = \begin{bmatrix}4\\0\\0\\0\end{bmatrix} = \begin{bmatrix}0\\1\\2\\1\end{bmatrix} + \begin{bmatrix}4\\-1\\-2\\-1\end{bmatrix} = \begin{bmatrix}0\\1\\2\\1\end{bmatrix} + \left((-2)\begin{bmatrix}-3\\-1\\1\\0\end{bmatrix} + (-1)\begin{bmatrix}2\\3\\0\\1\end{bmatrix}\right) = \mathbf{z} + \mathbf{w}_2$$

where

$$\mathbf{w_2} = \begin{bmatrix}4\\-1\\-2\\-1\end{bmatrix} = (-2)\begin{bmatrix}-3\\-1\\1\\0\end{bmatrix} + (-1)\begin{bmatrix}2\\3\\0\\1\end{bmatrix}$$

is obviously a solution of the homogenous system since it is written as a linear combination of the vectors describing the null space of the coefficient matrix as a span (or as a check, you could just evalute the equations in the homogenous system with $\mathbf{w_2}$).

Again

$$\mathbf{y}_3 = \begin{bmatrix}7\\8\\1\\3\end{bmatrix} = \begin{bmatrix}0\\1\\2\\1\end{bmatrix} + \begin{bmatrix}7\\7\\-1\\2\end{bmatrix} = \begin{bmatrix}0\\1\\2\\1\end{bmatrix} + \left((-1)\begin{bmatrix}-3\\-1\\1\\0\end{bmatrix} + 2\begin{bmatrix}2\\3\\0\\1\end{bmatrix}\right) = \mathbf{z} + \mathbf{w}_3$$

where

$$\mathbf{w_3} = \begin{bmatrix} 7 \\ 7 \\ -1 \\ 2 \end{bmatrix} = (-1) \begin{bmatrix} -3 \\ -1 \\ 1 \\ 0 \end{bmatrix} + 2 \begin{bmatrix} 2 \\ 3 \\ 0 \\ 1 \end{bmatrix}$$

is obviously a solution of the homogenous system since it is written as a linear combination of the vectors describing the null space of the coefficient matrix as a span (or as a check, you could just evalute the equations in the homogenous system with $\mathbf{w_2}$).

Here's another view of this theorem, in the context of this example. Grab two new solutions of the original system of equations, say

$$\mathbf{y}_4 = \begin{bmatrix} 11 \\ 0 \\ -3 \\ -1 \end{bmatrix} \qquad \mathbf{y}_5 = \begin{bmatrix} -4 \\ 2 \\ 4 \\ 2 \end{bmatrix}$$

and form their difference,

$$\mathbf{u} = \begin{bmatrix} 11 \\ 0 \\ -3 \\ -1 \end{bmatrix} - \begin{bmatrix} -4 \\ 2 \\ 4 \\ 2 \end{bmatrix} = \begin{bmatrix} 15 \\ -2 \\ -7 \\ -3 \end{bmatrix}.$$

It is no accident that **u** is a solution to the homogenous system (check this!). In other words, the difference between any two solutions to a linear system of equations is an element of the null space of the coefficient matrix. This is an equivalent way to state Theorem PSPHS [137]. If we let D denote the coefficient matrix then we can use the following application of Theorem PSPHS [137] as the basis of a formal proof of this assertion,

$$\begin{aligned} D(\mathbf{y}_4 - \mathbf{y}_5) &= D\left((z + \mathbf{w}_4) - (z + \mathbf{w}_5)\right) \\ &= D(\mathbf{w}_4 - \mathbf{w}_5) \\ &= D\mathbf{w}_4 - D\mathbf{w}_5 \\ &= \mathbf{0} - \mathbf{0} = \mathbf{0}. \end{aligned}$$

It would be very instructive to formulate the precise statement of a theorem and fill in the details and justifications of the proof. △

Section MISLE
Matrix Inverses and Systems of Linear Equations

We begin with a familiar example, performed in a novel way.

Example MISLE.SABMI
Solutions to Archetype B with a matrix inverse
Archetype B [171] is the system of $m = 3$ linear equations in $n = 3$ variables,

$$\begin{aligned} -7x_1 - 6x_2 - 12x_3 &= -33 \\ 5x_1 + 5x_2 + 7x_3 &= 24 \\ x_1 + 4x_3 &= 5 \end{aligned}$$

By Theorem SLEMM [128] we can represent this sytem of equations as

$$A\mathbf{x} = \mathbf{b}$$

where

$$A = \begin{bmatrix} -7 & -6 & -12 \\ 5 & 5 & 7 \\ 1 & 0 & 4 \end{bmatrix} \qquad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \qquad \mathbf{b} = \begin{bmatrix} -33 \\ 24 \\ 5 \end{bmatrix}.$$

We'll pull a rabbit out of our hat and present the 3×3 matrix B,

$$B = \begin{bmatrix} -10 & -12 & -9 \\ \frac{13}{2} & 8 & \frac{11}{2} \\ \frac{5}{2} & 3 & \frac{5}{2} \end{bmatrix}$$

and note that

$$BA = \begin{bmatrix} -10 & -12 & -9 \\ \frac{13}{2} & 8 & \frac{11}{2} \\ \frac{5}{2} & 3 & \frac{5}{2} \end{bmatrix} \begin{bmatrix} -7 & -6 & -12 \\ 5 & 5 & 7 \\ 1 & 0 & 4 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Now apply this computation to the problem of solving the system of equations,

$$\begin{aligned} A\mathbf{x} &= \mathbf{b} \\ B(A\mathbf{x}) &= B\mathbf{b} \\ (BA)\mathbf{x} &= B\mathbf{b} && \text{Theorem MMA [135]} \\ I_3\mathbf{x} &= B\mathbf{b} \\ \mathbf{x} &= B\mathbf{b} && \text{Theorem MMIM [133]} \end{aligned}$$

So we have

$$\mathbf{x} = B\mathbf{b} = \begin{bmatrix} -10 & -12 & -9 \\ \frac{13}{2} & 8 & \frac{11}{2} \\ \frac{5}{2} & 3 & \frac{5}{2} \end{bmatrix} \begin{bmatrix} -33 \\ 24 \\ 5 \end{bmatrix} = \begin{bmatrix} -3 \\ 5 \\ 2 \end{bmatrix}.$$

So with the help and assistance of B we have been able to determine a solution to the system represented by $A\mathbf{x} = \mathbf{b}$ through judicious use of matrix multiplication. We know by Theorem NSMUS [53] that since the coefficient matrix in this example is nonsingular, there would be a unique solution, no matter what the choice of $\mathbf{b}$. The derivation above amplifies this result, since we were *forced* to conclude that $\mathbf{x} = B\mathbf{b}$ and the solution couldn't be anything else. You should notice that this argument would hold for any particular value of $\mathbf{b}$. △

The matrix B of the previous example is called the inverse of A. When A and B are combined via matrix multiplication, the result is the identity matrix, which in this case left just the vector of unknowns, $\mathbf{x}$, on the left-side of the equation. This is entirely analagous to how we would solve a single linear equation like $3x = 12$. We would multiply both sides by $\frac{1}{3} = 3^{-1}$, the multiplicative inverse of 3. This works fine for any scalar multiple of x, except for zero, which does not have a multiplicative inverse. For matrices, it is more complicated. Some matrices have inverses, some do not. And when a matrix does have an inverse, just how would we compute it? In other words, just where did that matrix B in the last example come from? Are there other matrices that might have worked just as well?

Subsection IM
Inverse of a Matrix

Definition MI
Matrix Inverse
Suppose A and B are square matrices of size n such that $AB = I_n$ and $BA = I_n$. Then A is **invertible** and B is the **inverse** of A, and we write $B = A^{-1}$. ⊙

Notice that if B is the inverse of A, then we can just as easily say A is the inverse of B, or A and B are inverses of each other.

Not every square matrix has an inverse. In Example MISLE.SABMI [140] the matrix B is the inverse the coefficient matrix of Archetype B [171]. To see this it only remains to check that $AB = I_3$. What about Archetype A [167]? It is an example of a square matrix without an inverse.

Example MISLE.MWIAA
A matrix without an inverse, Archetype A

Consider the coefficient matrix from Archetype A [167],

$$A = \begin{bmatrix} 1 & -1 & 2 \\ 2 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix}$$

Suppose that A is invertible and does have an inverse, say B. Choose the vector of constants

$$\mathbf{b} = \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix}$$

and consider the system of equations $LS(A, \mathbf{b})$. Just as in Example MISLE.SABMI [140], this vector equation would have the unique solution $\mathbf{x} = B\mathbf{b}$.

However, this system is inconsistent. Form the augmented matrix $[A|\mathbf{b}]$ and row-reduce to

$$\begin{bmatrix} \boxed{1} & 0 & 1 & 0 \\ 0 & \boxed{1} & -1 & 0 \\ 0 & 0 & 0 & \boxed{1} \end{bmatrix}$$

which allows to recognize the inconsistency by Theorem RCLS [36].

So the assumption of A's inverse leads to a logical inconsistency (the system can't be both consistent and inconsistent), so our assumption is false. A is not invertible.

Its possible this example is less than satisfying. Just where did that particular choice of the vector $\mathbf{b}$ come from anyway? Turns out its not too mysterious. We wanted an inconsistent system, so Theorem RCS [105] suggested choosing a vector *outside* of the range of A (see Example ROM.RAA [116] for full disclosure). △

Lets look at one more matrix inverse before we embark on a more systematic study.

Example MISLE.MIAK
Matrix Inverse, Archetype K

Consider the matrix defined as Archetype K [211],

$$K = \begin{bmatrix} 10 & 18 & 24 & 24 & -12 \\ 12 & -2 & -6 & 0 & -18 \\ -30 & -21 & -23 & -30 & 39 \\ 27 & 30 & 36 & 37 & -30 \\ 18 & 24 & 30 & 30 & -20 \end{bmatrix}.$$

And the matrix

$$L = \begin{bmatrix} 1 & -\left(\frac{9}{4}\right) & -\left(\frac{3}{2}\right) & 3 & -6 \\ \frac{21}{2} & \frac{43}{4} & \frac{21}{2} & 9 & -9 \\ -15 & -\left(\frac{21}{2}\right) & -11 & -15 & \frac{39}{2} \\ 9 & \frac{15}{4} & \frac{9}{2} & 10 & -15 \\ \frac{9}{2} & \frac{3}{4} & \frac{3}{2} & 6 & -\left(\frac{19}{2}\right) \end{bmatrix}.$$

Then

$$KL = \begin{bmatrix} 10 & 18 & 24 & 24 & -12 \\ 12 & -2 & -6 & 0 & -18 \\ -30 & -21 & -23 & -30 & 39 \\ 27 & 30 & 36 & 37 & -30 \\ 18 & 24 & 30 & 30 & -20 \end{bmatrix} \begin{bmatrix} 1 & -\left(\frac{9}{4}\right) & -\left(\frac{3}{2}\right) & 3 & -6 \\ \frac{21}{2} & \frac{43}{4} & \frac{21}{2} & 9 & -9 \\ -15 & -\left(\frac{21}{2}\right) & -11 & -15 & \frac{39}{2} \\ 9 & \frac{15}{4} & \frac{9}{2} & 10 & -15 \\ \frac{9}{2} & \frac{3}{4} & \frac{3}{2} & 6 & -\left(\frac{19}{2}\right) \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

and

$$LK = \begin{bmatrix} 1 & -\left(\frac{9}{4}\right) & -\left(\frac{3}{2}\right) & 3 & -6 \\ \frac{21}{2} & \frac{43}{4} & \frac{21}{2} & 9 & -9 \\ -15 & -\left(\frac{21}{2}\right) & -11 & -15 & \frac{39}{2} \\ 9 & \frac{15}{4} & \frac{9}{2} & 10 & -15 \\ \frac{9}{2} & \frac{3}{4} & \frac{3}{2} & 6 & -\left(\frac{19}{2}\right) \end{bmatrix} \begin{bmatrix} 10 & 18 & 24 & 24 & -12 \\ 12 & -2 & -6 & 0 & -18 \\ -30 & -21 & -23 & -30 & 39 \\ 27 & 30 & 36 & 37 & -30 \\ 18 & 24 & 30 & 30 & -20 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

so by Definition MI [141], we can say that K is invertible and write $L = K^{-1}$. △

We will now concern ourselves less with whether or not an inverse of a matrix exists, but instead with how you can find one when it does exist. In Section MINSM [150] we will have some theorems that allow us to more quickly and easily determine when a matrix is invertible.

Subsection CIM Computing the Inverse of a Matrix

We will have occassion in this subsection (and later) to reference the following frequently used vectors, so we will make a useful definition now.

Definition SUV
Standard Unit Vectors
Let $\mathbf{e}_i$ denote column i of the identity matrix I_n. Then $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \ldots, \mathbf{e}_n\}$ is the set of **standard unit vectors**. ⊙

We've seen that the matrices from Archetype B [171] and Archetype K [211] both have inverses, but these inverse matrices have just dropped from the sky. How would we compute an inverse? And just when is a matrix invertible, and when is it not? Writing a putative inverse with n^2 unknowns and solving the resultant n^2 equations is one approach. Applying this approach to 2×2 matrices can get us somewhere, so just for fun, let's do it.

Theorem TTMI
Two-by-Two Matrix Inverse
Suppose

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

Then A is invertible if and only if $ad - bc \neq 0$. When A is invertible, we have

$$A^{-1} = \frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

□

Proof ($\Leftarrow$) If $ad - bc \neq 0$ then the displayed formula is legitimate (we are not dividing by zero), and it is a simple matter to actually check that $A^{-1}A = AA^{-1} = I_2$.

($\Rightarrow$) Assume that A is invertible, and proceed with a proof by contradiction, by assuming also that $ad - bc = 0$. This means that $ad = bc$. Let

$$B = \begin{bmatrix} e & f \\ g & h \end{bmatrix}$$

be a putative inverse of A. This means that

$$I_2 = AB = \begin{bmatrix} ae + bg & af + bh \\ ce + dg & cf + dh \end{bmatrix}$$

For the matrix on the right, multiply the top row by c and the bottom row by a. Since we are assuming that $ad = bc$, massage the bottom row by replacing ad by bc in two places. The result is that the two rows of the matrix are identical. Suppose we did the same to I_2, multiply the top row by c and the bottom row by a, and then arrived arrived at equal rows? Given the form of I_2 there is only one way this could happen: $a = 0$ and $c = 0$.

With this information, the product AB simplifies to

$$AB = \begin{bmatrix} bg & bh \\ dg & dh \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I_2$$

So $bg = dh = 1$ and thus b, g, d, h are all nonzero. But then bh and dg (the "other corners") must also be nonzero, so this is (finally) a contradiction. So our assumption was false and we see that $ad - bc \neq 0$ whenever A has an inverse. ■

There are several ways one could try to prove this theorem, but there is a continual temptation to divide by one of the eight entries involved (a through f), but we can never be sure if these numbers are zero or not. This could lead to an analysis by cases, which is messy,.... Note how the above proof never divides, but always multiplies, and how zero/nonzero considerations are handled. Pay attention to the expression $ad-bc$, we will see it again in a while.

This theorem is cute, and its nice to have a formula for the inverse, and a condition that tells us when we can use it. However, this approach becomes impractical for larger matrices, even though it is possible to demonstrate that, in theory, there is a general formula. Instead, we will work column-by-column. Let's first work an example that will motivate the main theorem and remove some of the previous mystery.

Example MISLE.CMIAK
Computing a Matrix Inverse, Archetype K
Consider the matrix defined as Archetype K [211],

$$A = \begin{bmatrix} 10 & 18 & 24 & 24 & -12 \\ 12 & -2 & -6 & 0 & -18 \\ -30 & -21 & -23 & -30 & 39 \\ 27 & 30 & 36 & 37 & -30 \\ 18 & 24 & 30 & 30 & -20 \end{bmatrix}.$$

For its inverse, we desire a matrix B so that $AB = I_5$. Emphazing the structure of the columns and employing the definition of matrix multiplication Definition MM [129],

$$\begin{aligned} AB &= I_5 \\ A[\mathbf{B}_1|\mathbf{B}_2|\mathbf{B}_3|\mathbf{B}_4|\mathbf{B}_5] &= [\mathbf{e}_1|\mathbf{e}_2|\mathbf{e}_3|\mathbf{e}_4|\mathbf{e}_5] \\ [A\mathbf{B}_1|A\mathbf{B}_2|A\mathbf{B}_3|A\mathbf{B}_4|A\mathbf{B}_5] &= [\mathbf{e}_1|\mathbf{e}_2|\mathbf{e}_3|\mathbf{e}_4|\mathbf{e}_5]. \end{aligned}$$

Equating the matrices column-by-column we have

$$A\mathbf{B}_1 = \mathbf{e}_1 \qquad A\mathbf{B}_2 = \mathbf{e}_2 \qquad A\mathbf{B}_3 = \mathbf{e}_3 \qquad A\mathbf{B}_4 = \mathbf{e}_4 \qquad A\mathbf{B}_5 = \mathbf{e}_5.$$

Since the matrix B is what we are trying to compute, we can view each column, $\mathbf{B}_i$, as a column vector of unknowns. Then we have five systems of equations to solve, each with 5 equations in 5 variables. Notice that all 5 these systems has the same coefficient matrix. We'll now solve each system in turn,

Row-reduce the augmented matrix of the linear system $LS(A, \mathbf{e}_1)$,

$$\begin{bmatrix} 10 & 18 & 24 & 24 & -12 & 1 \\ 12 & -2 & -6 & 0 & -18 & 0 \\ -30 & -21 & -23 & -30 & 39 & 0 \\ 27 & 30 & 36 & 37 & -30 & 0 \\ 18 & 24 & 30 & 30 & -20 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} \boxed{1} & 0 & 0 & 0 & 0 & 1 \\ 0 & \boxed{1} & 0 & 0 & 0 & \frac{21}{2} \\ 0 & 0 & \boxed{1} & 0 & 0 & -15 \\ 0 & 0 & 0 & \boxed{1} & 0 & 9 \\ 0 & 0 & 0 & 0 & \boxed{1} & \frac{9}{2} \end{bmatrix} \rightarrow \mathbf{B}_1 = \begin{bmatrix} 1 \\ \frac{21}{2} \\ -15 \\ 9 \\ \frac{9}{2} \end{bmatrix}$$

Row-reduce the augmented matrix of the linear system $LS(A, \mathbf{e}_2)$,

$$\begin{bmatrix} 10 & 18 & 24 & 24 & -12 & 0 \\ 12 & -2 & -6 & 0 & -18 & 1 \\ -30 & -21 & -23 & -30 & 39 & 0 \\ 27 & 30 & 36 & 37 & -30 & 0 \\ 18 & 24 & 30 & 30 & -20 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} \boxed{1} & 0 & 0 & 0 & 0 & -\frac{9}{4} \\ 0 & \boxed{1} & 0 & 0 & 0 & \frac{43}{4} \\ 0 & 0 & \boxed{1} & 0 & 0 & -\frac{21}{2} \\ 0 & 0 & 0 & \boxed{1} & 0 & \frac{15}{4} \\ 0 & 0 & 0 & 0 & \boxed{1} & \frac{3}{4} \end{bmatrix} \rightarrow \mathbf{B}_2 = \begin{bmatrix} -\frac{9}{4} \\ \frac{43}{4} \\ -\frac{21}{2} \\ \frac{15}{4} \\ \frac{3}{4} \end{bmatrix}$$

Row-reduce the augmented matrix of the linear system $LS(A, \mathbf{e}_3)$,

$$\begin{bmatrix} 10 & 18 & 24 & 24 & -12 & 0 \\ 12 & -2 & -6 & 0 & -18 & 0 \\ -30 & -21 & -23 & -30 & 39 & 1 \\ 27 & 30 & 36 & 37 & -30 & 0 \\ 18 & 24 & 30 & 30 & -20 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} \boxed{1} & 0 & 0 & 0 & 0 & -\frac{3}{2} \\ 0 & \boxed{1} & 0 & 0 & 0 & \frac{21}{2} \\ 0 & 0 & \boxed{1} & 0 & 0 & -11 \\ 0 & 0 & 0 & \boxed{1} & 0 & \frac{9}{2} \\ 0 & 0 & 0 & 0 & \boxed{1} & \frac{3}{2} \end{bmatrix} \rightarrow \mathbf{B}_3 = \begin{bmatrix} -\frac{3}{2} \\ \frac{21}{2} \\ -11 \\ \frac{9}{2} \\ \frac{3}{2} \end{bmatrix}$$

Row-reduce the augmented matrix of the linear system $\mathcal{LS}(A,\,\mathbf{e}_4)$,

$$\begin{bmatrix} 10 & 18 & 24 & 24 & -12 & 0 \\ 12 & -2 & -6 & 0 & -18 & 0 \\ -30 & -21 & -23 & -30 & 39 & 0 \\ 27 & 30 & 36 & 37 & -30 & 1 \\ 18 & 24 & 30 & 30 & -20 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} \boxed{1} & 0 & 0 & 0 & 0 & 3 \\ 0 & \boxed{1} & 0 & 0 & 0 & 9 \\ 0 & 0 & \boxed{1} & 0 & 0 & -15 \\ 0 & 0 & 0 & \boxed{1} & 0 & 10 \\ 0 & 0 & 0 & 0 & \boxed{1} & 6 \end{bmatrix} \rightarrow \mathbf{B}_4 = \begin{bmatrix} 3 \\ 9 \\ -15 \\ 10 \\ 6 \end{bmatrix}$$

Row-reduce the augmented matrix of the linear system $\mathcal{LS}(A,\,\mathbf{e}_5)$,

$$\begin{bmatrix} 10 & 18 & 24 & 24 & -12 & 0 \\ 12 & -2 & -6 & 0 & -18 & 0 \\ -30 & -21 & -23 & -30 & 39 & 0 \\ 27 & 30 & 36 & 37 & -30 & 0 \\ 18 & 24 & 30 & 30 & -20 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} \boxed{1} & 0 & 0 & 0 & 0 & -6 \\ 0 & \boxed{1} & 0 & 0 & 0 & -9 \\ 0 & 0 & \boxed{1} & 0 & 0 & \frac{39}{2} \\ 0 & 0 & 0 & \boxed{1} & 0 & -15 \\ 0 & 0 & 0 & 0 & \boxed{1} & -\frac{19}{2} \end{bmatrix} \rightarrow \mathbf{B}_5 = \begin{bmatrix} -6 \\ -9 \\ \frac{39}{2} \\ -15 \\ -\frac{19}{2} \end{bmatrix}$$

We can now collect our 5 solution vectors into the matrix B,

$$\begin{aligned} B =& [\mathbf{B}_1|\mathbf{B}_2|\mathbf{B}_3|\mathbf{B}_4|\mathbf{B}_5] \\ =& \left[\begin{bmatrix} 1 \\ \frac{21}{2} \\ -15 \\ 9 \\ \frac{9}{2} \end{bmatrix}\middle|\begin{bmatrix} -\frac{9}{4} \\ \frac{43}{4} \\ -\frac{21}{2} \\ \frac{15}{4} \\ \frac{3}{4} \end{bmatrix}\middle|\begin{bmatrix} -\frac{3}{2} \\ \frac{21}{2} \\ -11 \\ \frac{9}{2} \\ \frac{3}{2} \end{bmatrix}\middle|\begin{bmatrix} 3 \\ 9 \\ -15 \\ 10 \\ 6 \end{bmatrix}\middle|\begin{bmatrix} -6 \\ -9 \\ \frac{39}{2} \\ -15 \\ -\frac{19}{2} \end{bmatrix}\right] = \begin{bmatrix} 1 & -\left(\frac{9}{4}\right) & -\left(\frac{3}{2}\right) & 3 & -6 \\ \frac{21}{2} & \frac{43}{4} & \frac{21}{2} & 9 & -9 \\ -15 & -\left(\frac{21}{2}\right) & -11 & -15 & \frac{39}{2} \\ 9 & \frac{15}{4} & \frac{9}{2} & 10 & -15 \\ \frac{9}{2} & \frac{3}{4} & \frac{3}{2} & 6 & -\left(\frac{19}{2}\right) \end{bmatrix} \end{aligned}$$

By this method, we know that $AB = I_5$. Check that $BA = I_5$, and then we will know that we have the inverse of A. △

Notice how the five systems of equations in the preceding example were all solved by *exactly* the same sequence of row operations. Wouldn't it be nice to avoid this obvious duplication of effort? Our main theorem for this section follows, and it mimics this previous example, while also avoiding all the overhead.

Theorem CINSM
Computing the Inverse of a NonSingular Matrix
Suppose A is a nonsingular square matrix of size n. Create the $n \times 2n$ matrix M by placing the $n \times n$ identity matrix I_n to the right of the matrix A. Let N be a matrix that is row-equivalent to M and in reduced row-echelon form. Finally, let B be the matrix formed from the final n columns of N. Then $AB = I_n$. □

Proof A is nonsingular, so by Theorem NSRRI [50] there is a sequence of row operations that will convert A into I_n. It is this same sequence of row operations that will convert M into N, since having the identity matrix in the first n columns of N is sufficient to guarantee that it is in reduced row-echelon form.

If we consider the systems of linear equations, $LS(A, \mathbf{e}_i)$, $1 \leq i \leq n$, we see that the aforementioned sequence of row operations will also bring the augmented matrix of each system into reduced row-echelon form. Furthermore, the unique solution to each of these systems appears in column $n+1$ of the row-reduced augmented matrix and is equal to column $n+i$ of N. Let $\mathbf{N}_1$, $\mathbf{N}_2$, $\mathbf{N}_3$, $\ldots$, $\mathbf{N}_{2n}$ denote the columns of N. So we find,

$$\begin{aligned} AB &= A[\mathbf{N}_{n+1}|\mathbf{N}_{n+2}|\mathbf{N}_{n+3}|\ldots|\mathbf{N}_{n+n}] \\ &= [A\mathbf{N}_{n+1}|A\mathbf{N}_{n+2}|A\mathbf{N}_{n+3}|\ldots|A\mathbf{N}_{n+n}] \\ &= [\mathbf{e}_1|\mathbf{e}_2|\mathbf{e}_3|\ldots|\mathbf{e}_n] \\ &= I_n \end{aligned}$$

as desired. ■

Does this theorem remind you of any others we've seen lately? (Hint: Theorem RNS [114].) We have to be just a bit careful here. This theorem only guarantees that $AB = I_n$, while the definition requires that $BA = I_n$ also. However, we'll soon see that this is *always* the case, in Theorem OSIS [151], so the title of this theorem is not inaccurate.

We'll finish by computing the inverse for the coefficient matrix of Archetype B [171], the one we just pulled from a hat in Example MISLE.SABMI [140]. There are more examples in the Archetypes (Chapter A [163]) to practice with, though notice that it is silly to ask for the inverse of a rectangular matrix (the sizes aren't right) and not every square matrix has an inverse (remember Example MISLE.MWIAA [141]?).

Example MISLE.CMIAB
Computing a Matrix Inverse, Archetype B

Archetype B [171] has a coefficient matrix given as

$$B = \begin{bmatrix} -7 & -6 & -12 \\ 5 & 5 & 7 \\ 1 & 0 & 4 \end{bmatrix}$$

Exercising Theorem CINSM [146] we set

$$M = \begin{bmatrix} -7 & -6 & -12 & 1 & 0 & 0 \\ 5 & 5 & 7 & 0 & 1 & 0 \\ 1 & 0 & 4 & 0 & 0 & 1 \end{bmatrix}.$$

which row reduces to

$$N = \begin{bmatrix} 1 & 0 & 0 & -10 & -12 & -9 \\ 0 & 1 & 0 & \frac{13}{2} & 8 & \frac{11}{2} \\ 0 & 0 & 1 & \frac{5}{2} & 3 & \frac{5}{2} \end{bmatrix}.$$

So

$$B^{-1} = \begin{bmatrix} -10 & -12 & -9 \\ \frac{13}{2} & 8 & \frac{11}{2} \\ \frac{5}{2} & 3 & \frac{5}{2} \end{bmatrix}$$

once we check that $B^{-1}B = I_3$ (the product in the opposite order is a consequence of the theorem). △

Subsection Properties of Matrix Inverses PMI

The inverse of a matrix enjoys some nice properties. We collect a few here. First, a matrix can have but one inverse.

Theorem MIU
Matrix Inverse is Unique
Suppose the square matrix A has an inverse. Then A^{-1} is unique. □

Proof As described in Technique U [53], we will assume that A has two inverses. The hypothesis tells there is at least one. Suppose then that B and C are both inverses for A. Then, repeated use of Definition MI [141] and Theorem MMIM [133] plus one application of Theorem MMA [135] gives

$$B = BI_n = B(AC) = (BA)C = I_nC = C$$

and we conclude that B and C cannot be different. So any matrix that acts like the inverse, must be *the* inverse. ■

When most of dress in the morning, we put on our socks first, followed by our shoes. In the evening we must then first remove our shoes, followed by our socks. Try to connect the conclusion of the following theorem with this everyday example.

Theorem SST
Socks and Shoes Theorem
Suppose A and B are invertible matrices of size n. Then $(AB)^{-1} = B^{-1}A^{-1}$. □

Proof At the risk of carrying our everyday analogies too far, the proof of this theorem is quite easy when we compare it to the workings of a dating service. We have a statement about the inverse of the matrix AB, which for all we know right now might not even exist. Suppose AB was to sign up for a dating service with two requirements for a compatible date. Upon multiplication on the left, and on the right, the result should be the identity matrix. In other words, AB's ideal date would be its inverse.

Now along comes the matrix $B^{-1}A^{-1}$ (which we know exists because our hypothesis says both A and B are invertible), also looking for a date. Lets see if $B^{-1}A^{-1}$ is a good match for AB,

$$\begin{aligned}(B^{-1}A^{-1})(AB) &= B^{-1}(A^{-1}A)B = B^{-1}I_nB = B^{-1}B = I_n\\ (AB)(B^{-1}A^{-1}) &= A(BB^{-1})A^{-1} = AI_nA^{-1} = AA^{-1} = I_n.\end{aligned}$$

So the matrix $B^{-1}A^{-1}$ has met all of the requirements to be AB's inverse (date) and we can write $(AB)^{-1} = B^{-1}A^{-1}$. ■

Theorem MIT
Matrix Inverse of a Transpose
Suppose A is an invertible matrix. Then $(A^t)^{-1} = (A^{-1})^t$. □

Proof As with the proof of Theorem SST [148], we see if $(A^{-1})^t$ is a suitable inverse for A^t. Apply Theorem MMT [136] to see that

$$\begin{aligned}(A^{-1})^t A^t &= (AA^{-1})^t = I_n^t = I_n \\ A^t(A^{-1})^t &= (A^{-1}A)^t = I_n^t = I_n\end{aligned}$$

The matrix $(A^{-1})^t$ has met all the requirements to be the inverse of A^t, so we can write $(A^t)^{-1} = (A^{-1})^t$. ■

Theorem MISM
Matrix Inverse of a Scalar Multiple
Suppose A is an invertible matrix and α is a nonzero scalar. Then $(\alpha A)^{-1} = \frac{1}{\alpha}A^{-1}$. □

Proof As with the proof of Theorem SST [148], we see if $\frac{1}{\alpha}A^{-1}$ is a suitable inverse for αA. Apply Theorem MMSMM [134] to see that

$$\begin{aligned}\left(\frac{1}{\alpha}A^{-1}\right)(\alpha A) &= \left(\frac{1}{\alpha}\alpha\right)\left(AA^{-1}\right) = 1I_n = I_n \\ (\alpha A)\left(\frac{1}{\alpha}A^{-1}\right) &= \left(\alpha\frac{1}{\alpha}\right)\left(A^{-1}A\right) = 1I_n = I_n\end{aligned}$$

The matrix $\frac{1}{\alpha}A^{-1}$ has met all the requirements to be the inverse of αA, so we can write $(\alpha A)^{-1} = \frac{1}{\alpha}A^{-1}$. ■

Notice that there are some likely theorems that are missing here. For example, it would be tempting to think that $(A+B)^{-1} = A^{-1} + B^{-1}$, but this is false. Can you find a counterexample?

Section MINSM
Matrix Inverses and NonSingular Matrices

We saw in Theorem CINSM [146] that if a square matrix A is nonsingular, then there is a matrix B so that $AB = I_n$. In other words, B is halfway to being an inverse of A. We will see in this section that B automatically fulfills the second condition ($BA = I_n$). Example MISLE.MWIAA [141] showed us that the coefficient matrix from Archetype A [167] had no inverse. Not coincidentally, this coefficient matrix is singular. We'll make all these connections precise now. Not many examples or definitions in this section, just theorems.

Subsection NSMI
NonSingular Matrices are Invertible

We need a couple of technical results for starters. Some books would call these minor, but essential, results "lemmas." We'll just call 'em theorems.

Theorem PWSMS
Product With a Singular Matrix is Singular
Suppose that A or B are matrices of size n, and one, or both, is singular. Then their product, AB, is singular. □

Proof We will use the vector equation representation of the relevant systems of equations throughout the proof (Theorem SLEMM [128]). We'll do the proof in two cases, and its interesting to notice how we break down the cases.

Case 1. Suppose B is singular. Then there is a nontrivial vector $\mathbf{z}$ so that $B\mathbf{z} = \mathbf{0}$. Then

$$(AB)\mathbf{z} = A(B\mathbf{z}) = A\mathbf{0} = \mathbf{0}$$

so we can conclude that AB is singular.

Case 2. Suppose B is nonsingular and A is singular. This is probably not the second case you were expecting. Why not just state the second case as "A is singular"? The best answer is that the proof is easier with the more restrictive assumption that A is singular *and* B is nonsingular. But before we see why, convince yourself that the two cases, as stated, will cover all the possibilities allowed by our hypothesis.

Since A is singular, there is a nontrivial vector $\mathbf{y}$ so that $A\mathbf{y} = \mathbf{0}$. Now consider the linear system $\mathcal{LS}(B,\,\mathbf{y})$. Since B is nonsingular, the system has a unique solution, which we will call $\mathbf{w}$. We claim $\mathbf{w}$ is not the zero vector. If $\mathbf{w} = \mathbf{0}$, then

$$\mathbf{y} = B\mathbf{w} = B\mathbf{0} = \mathbf{0}$$

contrary to $\mathbf{y}$ being nontrivial. So $\mathbf{w} \neq \mathbf{0}$. The pieces are in place, so here we go,

$$(AB)\mathbf{w} = A(B\mathbf{w}) = A\mathbf{y} = \mathbf{0}$$

which says, since $\mathbf{w}$ is nontrivial, that AB is singular. ■

Theorem OSIS
One-Sided Inverse is Sufficient
Suppose A is a nonsingular matrix of size n, and B is a square matrix of size n such that $AB = I_n$. Then $BA = I_n$. □

Proof The matrix I_n is nonsingular (since it row-reduces easily to I_n, Theorem NSRRI [50]). If B is singular, then Theorem PWSMS [150] would imply that I_n is singular, a contradiction. So B must be nonsingular also. Now that we know that B is nonsingular, we can apply Theorem CINSM [146] to assert the existence of a matrix C so that $BC = I_n$. This application of Theorem CINSM [146] could be a bit confusing, mostly because of the names of the matrices involved. B is nonsingular, so there must be a "right-inverse" for B, and we're calling it C.

Now

$$C = I_nC = (AB)C = A(BC) = AI_n = A.$$

So it happens that the matrix C we just found is really A in disguise. So we can write

$$I_n = BC = BA$$

which is the desired conclusion. ■

So Theorem OSIS [151] tells us that if A is nonsingular, then the matrix B guaranteed by Theorem CINSM [146] will be both a "right-inverse" and a "left-inverse" for A, so A is invertible and $A^{-1} = B$.

So if you have a nonsingular matrix, A, you can use the procedure described in Theorem CINSM [146] to find an inverse for A. If A is singular, then the procedure in Theorem CINSM [146] will fail as the first n columns of M will not row-reduce to the identity matrix.

This may feel like we are splitting hairs, but its important that we do not make unfounded assumptions. These observations form the next theorem.

Theorem NSI
NonSingularity is Invertibility
Suppose that A is a square matrix. Then A is nonsingular if and only if A is invertible.□

Proof ($\Leftarrow$) Suppose A is invertible, and consider the homogenous system represented by $A\mathbf{x} = \mathbf{0}$,

$$\begin{aligned} A\mathbf{x} &= \mathbf{0} \\ A^{-1}A\mathbf{x} &= A^{-1}\mathbf{0} \\ I_n\mathbf{x} &= \mathbf{0} \\ \mathbf{x} &= \mathbf{0} \end{aligned}$$

So A has a trivial null space, which is a fancy way of saying that A is nonsingular.

($\Rightarrow$) Suppose A is nonsingular. By Theorem CINSM [146] we find B so that $AB = I_n$. Then Theorem OSIS [151] tells us that $BA = I_n$. So B is A's inverse, and by construction, A is invertible. ■

So the properties of having an inverse and of having a trivial null space are one and the same. Can't have one without the other. Now we can update our list of equivalences for nonsingular matrices (Theorem NSME3 [118]).

Theorem NSME4
NonSingular Matrix Equivalences, Round 4
Suppose that A is a square matrix of size n. The following are equivalent.

1. A is nonsingular.
2. A row-reduces to the identity matrix.
3. The null space of A contains only the zero vector, $\mathcal{N}(A) = \{\mathbf{0}\}$.
4. The linear system $\mathcal{LS}(A, \mathbf{b})$ has a unique solution for every possible choice of $\mathbf{b}$.
5. The columns of A are a linearly independent set.
6. The range of A is $\mathbb{C}^n$.
7. A is invertible. □

In the case that A is a nonsingular coefficient matrix of a system of equations, the inverse allows us to very quickly compute the unique solution, for any vector of constants.

Theorem SNSCM
Solution with NonSingular Coefficient Matrix
Suppose that A is nonsingular. Then the unique solution to $\mathcal{LS}(A, \mathbf{b})$ is $A^{-1}\mathbf{b}$. □

Proof By Theorem NSMUS [53] we know already that $\mathcal{LS}(A, \mathbf{b})$ has a unique solution for every choice of $\mathbf{b}$. We need to show that the expression given is indeed a solution. That's easy, just "plug it in" to the corresponding vector equation representation,

$$A(A^{-1}\mathbf{b}) = (AA^{-1})\mathbf{b} = I_n\mathbf{b} = \mathbf{b}.$$ ■

151

VS: Vector Spaces

D: Determinants

E: Eigenvalues

LT: Linear Transformations

R: Representations

A: Archetypes

The American Heritage Dictionary of the English Language (Third Edition) gives two definitions of the word "archetype": 1. An original model or type after which other similar things are patterned; a prototype; and 2. An ideal example of a type; quintessence.

Either use might apply here. Our archetypes are typical examples of systems of equations, matrices and linear transformations. They have been designed to demonstrate the range of possibilities, allowing you to compare and contrast them. Several are of a size and complexity that is usually not presented in a textbook, but should do a better job of being "typical."

We have made frequent reference to many of these throughout the text, such as the frequent comparisons between Archetype A [167] and Archetype B [171]. Some we have left for you to investigate, such as Archetype J [206], which parallels Archetype I [201].

How should you use the archetypes? First, consult the description of each one as it is mentioned in the text. See how other facts about the example might illuminate whatever property or construction is being described in the example. Second, Each property has a short description that usually includes references to the relevant theorems. Perform the computations and understand the connections to the listed theorems. Third, each property has a small checkbox in front of it. Use the archetypes like a workbook and chart your progress by "checking-off" those properties that you understand.

The next page has a chart that summarizes some (but not all) of the properties described for each archetype. Notice that while there are several types of objects, there are fundamental connections between them. That some lines of the table do double-duty is meant to convey some of these connections. Consult this table when you wish to quickly find an example of a certain phenomenon.

	A	B	C	D	E	F	G	H	I	J		K	L
Type	S	S	S	S	S	S	S	S	S	S		M	M
Variables	3	3	4	4	4	4	2	2	7	9	Columns	5	5
Equations	3	3	3	3	3	4	5	5	4	6	Rows	5	5
Consistent	I	U	I	I	N	U	U	N	I	I			
Rank	2	3	3	2	2	4	2	2	3	4		5	3
Nullity	1	0	1	2	2	0	3	3	4	5		0	2
Full Rank	N	Y	Y	N	N	Y	Y	Y	N	N		Y	N
Nonsingular	N	Y				Y						Y	N
Invertible	N	Y				Y						Y	N
Determinant	0	6				-18						16	0

Archetype Facts
S=System of Equations, M=Matrix, LT=Linear Transformation
U=Unique solution, I=Infinitely many solutions, N=No solutions
Y=Yes, N=No, blank=Not Applicable

Archetype A

Summary Linear system of three equations, three unknowns. Singular coefficent matrix with dimension 1 null space. Integer eigenvalues and a degenerate eigenspace for coefficient matrix.

□ A system of linear equations (Definition SSLE [8]):

$$\begin{aligned} x_1 - x_2 + 2x_3 &= 1 \\ 2x_1 + x_2 + x_3 &= 8 \\ x_1 + x_2 &= 5 \end{aligned}$$

□ Some solutions to the system of linear equations (not necessarily exhaustive):

$x_1 = 2, \quad x_2 = 3, \quad x_3 = 1$

$x_1 = 3, \quad x_2 = 2, \quad x_3 = 0$

□ Augmented matrix of the linear system of equations (Definition AM [20]):

$$\begin{bmatrix} 1 & -1 & 2 & 1 \\ 2 & 1 & 1 & 8 \\ 1 & 1 & 0 & 5 \end{bmatrix}$$

□ Matrix in reduced row-echelon form, row-equivalent to augmented matrix:

$$\begin{bmatrix} \boxed{1} & 0 & 1 & 3 \\ 0 & \boxed{1} & -1 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

□ Analysis of the augmented matrix (Notation RREFA [31]):

$$r = 2 \qquad D = \{1,\, 2\} \qquad F = \{3,\, 4\}$$

□ Vector form of the solution set to the system of equations (Theorem VFSLS [72]). Notice the relationship between the free variables and the set F above. Also, notice the pattern of 0's and 1's in the entries of the vectors corresponding to elements of the set F for the larger examples.

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$$

☐ A system of equations can always be converted to a homogenous system (Definition HS [41]):

$$\begin{aligned} x_1 - x_2 + 2x_3 &= 0 \\ 2x_1 + x_2 + x_3 &= 0 \\ x_1 + x_2 \qquad &= 0 \end{aligned}$$

☐ Some solutions to the associated homogenous system of linear equations (not necessarily exhaustive):

$x_1 = 0, \quad x_2 = 0, \quad x_3 = 0$

$x_1 = -1, \quad x_2 = 1, \quad x_3 = 1$

$x_1 = -5, \quad x_2 = 5, \quad x_3 = 5$

☐ Form the augmented matrix of the homogenous linear system, and use row operations to convert to reduced row-echelon form. Notice how the entries of the final column remain zeros:

$$\begin{bmatrix} \boxed{1} & 0 & 1 & 0 \\ 0 & \boxed{1} & -1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

☐ Analysis of the augmented matrix for the homogenous system (Notation RREFA [31]). Notice the slight variation for the same analysis of the original system only when the original system was consistent:

$$r = 2 \qquad D = \{1, 2\} \qquad F = \{3, 4\}$$

☐ Coefficient matrix of original system of equations, and of associated homogenous system. This matrix will be the subject of further analysis, rather than the systems of equations.

$$\begin{bmatrix} 1 & -1 & 2 \\ 2 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix}$$

□ Matrix brought to reduced row-echelon form:

$$\begin{bmatrix} \boxed{1} & 0 & 1 \\ 0 & \boxed{1} & -1 \\ 0 & 0 & 0 \end{bmatrix}$$

□ Analysis of the row-reduced matrix (Notation RREFA [31]):

$$r = 2 \qquad D = \{1,\, 2\} \qquad F = \{3\}$$

□ Matrix (coefficient matrix) is nonsingular or singular? (Theorem NSRRI [50]) at the same time, examine the size of the set F above.Notice that this property does not apply to matrices that are not square.

Singular.

□ This is the null space of the matrix. The set of vectors used in the span construction is a linearly independent set of column vectors that spans the null space of the matrix (Theorem SSNS [81], Theorem BNS [94]). Solve the homogenous system with this matrix as the coefficient matrix and write the solutions in vector form (Theorem VFSLS [72]) to see these vectors arise.

$$\mathrm{Sp}\left(\left\{\begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}\right\}\right)$$

□ Range of the matrix, expressed as the span of a set of linearly independent vectors that are also columns of the matrix. These columns have indices that form the set D above. (Theorem BROC [109])

$$\mathrm{Sp}\left(\left\{\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix},\, \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}\right\}\right)$$

□ The range of the matrix, as it arises from row operations on an augmented matrix. The matrix K is computed as described in Theorem RNS [114]. This is followed by the range described by a set of linearly independent vectors that span the null space of K, computed as according to Theorem SSNS [81]. When $r = m$, the matrix K has no rows and the range is all of $\mathbb{C}^m$.

$K = \begin{bmatrix} 1 & -2 & 3 \end{bmatrix}$

$$\mathrm{Sp}\left(\left\{\begin{bmatrix}-3\\0\\1\end{bmatrix},\ \begin{bmatrix}2\\1\\0\end{bmatrix}\right\}\right)$$

☐ Range of the matrix, expressed as the span of a set of linearly independent vectors. These vectors are computed by row-reducing the transpose the matrix into reduced row-echelon form, tossing out the zero rows, and writing the remaining rows as column vectors. By Theorem RMSRT [124] and Theorem BRS [123], and in the style of Example RSOM.IS [123], this yields a linearly independent set of vectors that span the range.

$$\mathrm{Sp}\left(\left\{\begin{bmatrix}1\\0\\-\frac{1}{3}\end{bmatrix},\ \begin{bmatrix}0\\1\\\frac{2}{3}\end{bmatrix}\right\}\right)$$

☐ Row space of the matrix, expressed as a span of a set of linearly independent vectors, obtained from the nonzero rows of the equivalent matrix in reduced row-echelon form. (Theorem BRS [123])

$$\mathrm{Sp}\left(\left\{\begin{bmatrix}1\\0\\1\end{bmatrix},\ \begin{bmatrix}0\\1\\-1\end{bmatrix}\right\}\right)$$

☐ Inverse matrix, if it exists. The inverse is not defined for matrices that are not square, and if the matrix is square, then the matrix must be nonsingular. (Definition MI [141], Theorem NSI [151])

☐ Subspace dimensions associated with the matrix.

Matrix columns: 3 Rank: 2 Nullity: 1

☐ Determinant of the matrix, which is only defined for square matrices. (Zero/nonzero? Singular/nonsingular? Product of all eigenvalues?)

Determinant = 0

Archetype B

Summary System with three equations, three unknowns. Nonsingular coefficent matrix. Distinct integer eigenvalues for coefficient matrix.

□ A system of linear equations (Definition SSLE [8]):

$$\begin{aligned} -7x_1 - 6x_2 - 12x_3 &= -33 \\ 5x_1 + 5x_2 + 7x_3 &= 24 \\ x_1 + 4x_3 &= 5 \end{aligned}$$

□ Some solutions to the system of linear equations (not necessarily exhaustive):

$x_1 = -3, \quad x_2 = 5, \quad x_3 = 2$

□ Augmented matrix of the linear system of equations (Definition AM [20]):

$$\begin{bmatrix} -7 & -6 & -12 & -33 \\ 5 & 5 & 7 & 24 \\ 1 & 0 & 4 & 5 \end{bmatrix}$$

□ Matrix in reduced row-echelon form, row-equivalent to augmented matrix:

$$\begin{bmatrix} \boxed{1} & 0 & 0 & -3 \\ 0 & \boxed{1} & 0 & 5 \\ 0 & 0 & \boxed{1} & 2 \end{bmatrix}$$

□ Analysis of the augmented matrix (Notation RREFA [31]):

$$r = 3 \qquad D = \{1,\, 2,\, 3\} \qquad F = \{4\}$$

□ Vector form of the solution set to the system of equations (Theorem VFSLS [72]). Notice the relationship between the free variables and the set F above. Also, notice the pattern of 0's and 1's in the entries of the vectors corresponding to elements of the set F for the larger examples.

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -3 \\ 5 \\ 2 \end{bmatrix}$$

□ A system of equations can always be converted to a homogenous system (DEFINITION HS [41]):

$$\begin{aligned} -11x_1 + 2x_2 - 14x_3 &= 0 \\ 23x_1 - 6x_2 + 33x_3 &= 0 \\ 14x_1 - 2x_2 + 17x_3 &= 0 \end{aligned}$$

□ Some solutions to the associated homogenous system of linear equations (not necessarily exhaustive):
$x_1 = 0, \quad x_2 = 0, \quad x_3 = 0$

□ Form the augmented matrix of the homogenous linear system, and use row operations to convert to reduced row-echelon form. Notice how the entries of the final column remain zeros:

$$\begin{bmatrix} \boxed{1} & 0 & 0 & 0 \\ 0 & \boxed{1} & 0 & 0 \\ 0 & 0 & \boxed{1} & 0 \end{bmatrix}$$

□ Analysis of the augmented matrix for the homogenous system (NOTATION RREFA [31]). Notice the slight variation for the same analysis of the original system only when the original system was consistent:

$$r = 3 \qquad D = \{1,\, 2,\, 3\} \qquad F = \{4\}$$

□ Coefficient matrix of original system of equations, and of associated homogenous system. This matrix will be the subject of further analysis, rather than the systems of equations.

$$\begin{bmatrix} -7 & -6 & -12 \\ 5 & 5 & 7 \\ 1 & 0 & 4 \end{bmatrix}$$

□ Matrix brought to reduced row-echelon form:

$$\begin{bmatrix} \boxed{1} & 0 & 0 \\ 0 & \boxed{1} & 0 \\ 0 & 0 & \boxed{1} \end{bmatrix}$$

□ Analysis of the row-reduced matrix (Notation RREFA [31]):

$$r = 3 \qquad D = \{1,\, 2,\, 3\} \qquad F = \{\ \}$$

□ Matrix (coefficient matrix) is nonsingular or singular? (Theorem NSRRI [50]) at the same time, examine the size of the set F above.Notice that this property does not apply to matrices that are not square.

Nonsingular.

□ This is the null space of the matrix. The set of vectors used in the span construction is a linearly independent set of column vectors that spans the null space of the matrix (Theorem SSNS [81], Theorem BNS [94]). Solve the homogenous system with this matrix as the coefficient matrix and write the solutions in vector form (Theorem VFSLS [72]) to see these vectors arise.

$\mathrm{Sp}(\{\ \})$

□ Range of the matrix, expressed as the span of a set of linearly independent vectors that are also columns of the matrix. These columns have indices that form the set D above. (Theorem BROC [109])

$$\mathrm{Sp}\left(\left\{\begin{bmatrix} -7 \\ 5 \\ 1 \end{bmatrix},\, \begin{bmatrix} -6 \\ 5 \\ 0 \end{bmatrix},\, \begin{bmatrix} -12 \\ 7 \\ 4 \end{bmatrix}\right\}\right)$$

□ The range of the matrix, as it arises from row operations on an augmented matrix. The matrix K is computed as described in Theorem RNS [114]. This is followed by the range described by a set of linearly independent vectors that span the null space of K, computed as according to Theorem SSNS [81]. When $r = m$, the matrix K has no rows and the range is all of $\mathbb{C}^m$.

$K = [\,]$

$$\mathrm{Sp}\left(\left\{\begin{bmatrix}1\\0\\0\end{bmatrix},\,\begin{bmatrix}0\\1\\0\end{bmatrix},\,\begin{bmatrix}0\\0\\1\end{bmatrix}\right\}\right)$$

☐ Range of the matrix, expressed as the span of a set of linearly independent vectors. These vectors are computed by row-reducing the transpose the matrix into reduced row-echelon form, tossing out the zero rows, and writing the remaining rows as column vectors. By Theorem RMSRT [124] and Theorem BRS [123], and in the style of Example RSOM.IS [123], this yields a linearly independent set of vectors that span the range.

$$\mathrm{Sp}\left(\left\{\begin{bmatrix}1\\0\\0\end{bmatrix},\,\begin{bmatrix}0\\1\\0\end{bmatrix},\,\begin{bmatrix}0\\0\\1\end{bmatrix}\right\}\right)$$

☐ Row space of the matrix, expressed as a span of a set of linearly independent vectors, obtained from the nonzero rows of the equivalent matrix in reduced row-echelon form. (Theorem BRS [123])

$$\mathrm{Sp}\left(\left\{\begin{bmatrix}1\\0\\0\end{bmatrix},\,\begin{bmatrix}0\\1\\0\end{bmatrix},\,\begin{bmatrix}0\\0\\1\end{bmatrix}\right\}\right)$$

☐ Inverse matrix, if it exists. The inverse is not defined for matrices that are not square, and if the matrix is square, then the matrix must be nonsingular. (Definition MI [141], Theorem NSI [151])

$$\begin{bmatrix}-10 & -12 & -9\\ \frac{13}{2} & 8 & \frac{11}{2}\\ \frac{5}{2} & 3 & \frac{5}{2}\end{bmatrix}$$

☐ Subspace dimensions associated with the matrix.

Matrix columns: 3 Rank: 3 Nullity: 0

☐ Determinant of the matrix, which is only defined for square matrices. (Zero/nonzero? Singular/nonsingular? Product of all eigenvalues?)

Determinant $= 6$

Archetype C

Summary System with three equations, four variables. Consistent. Null space of coefficient matrix has dimension 1.

A system of linear equations (Definition SSLE [8]):

$$\begin{aligned} 2x_1 - 3x_2 + x_3 - 6x_4 &= -7 \\ 4x_1 + x_2 + 2x_3 + 9x_4 &= -7 \\ 3x_1 + x_2 + x_3 + 8x_4 &= -8 \end{aligned}$$

Some solutions to the system of linear equations (not necessarily exhaustive):

$x_1 = -7, \quad x_2 = -2, \quad x_3 = 7, \quad x_4 = 1$

$x_1 = -1, \quad x_2 = -7, \quad x_3 = 4, \quad x_4 = -2$

Augmented matrix of the linear system of equations (Definition AM [20]):

$$\begin{bmatrix} 2 & -3 & 1 & -6 & -7 \\ 4 & 1 & 2 & 9 & -7 \\ 3 & 1 & 1 & 8 & -8 \end{bmatrix}$$

Matrix in reduced row-echelon form, row-equivalent to augmented matrix:

$$\begin{bmatrix} \boxed{1} & 0 & 0 & 2 & -5 \\ 0 & \boxed{1} & 0 & 3 & 1 \\ 0 & 0 & \boxed{1} & -1 & 6 \end{bmatrix}$$

Analysis of the augmented matrix (Notation RREFA [31]):

$$r = 3 \qquad D = \{1,\, 2,\, 3\} \qquad F = \{4,\, 5\}$$

Vector form of the solution set to the system of equations (Theorem VFSLS [72]). Notice the relationship between the free variables and the set F above. Also, notice the pattern of 0's and 1's in the entries of the vectors corresponding to elements of the set F for the larger examples.

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} -5 \\ 1 \\ 6 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} -2 \\ -3 \\ 1 \\ 1 \end{bmatrix}$$

A system of equations can always be converted to a homogenous system (Definition HS [41]):

$$\begin{aligned} 2x_1 - 3x_2 + x_3 - 6x_4 &= 0 \\ 4x_1 + x_2 + 2x_3 + 9x_4 &= 0 \\ 3x_1 + x_2 + x_3 + 8x_4 &= 0 \end{aligned}$$

Some solutions to the associated homogenous system of linear equations (not necessarily exhaustive):

$x_1 = 0, \quad x_2 = 0, \quad x_3 = 0, \quad x_4 = 0$

$x_1 = -2, \quad x_2 = -3, \quad x_3 = 1, \quad x_4 = 1$

$x_1 = -4, \quad x_2 = -6, \quad x_3 = 2, \quad x_4 = 2$

Form the augmented matrix of the homogenous linear system, and use row operations to convert to reduced row-echelon form. Notice how the entries of the final column remain zeros:

$$\begin{bmatrix} \boxed{1} & 0 & 0 & 2 & 0 \\ 0 & \boxed{1} & 0 & 3 & 0 \\ 0 & 0 & \boxed{1} & -1 & 0 \end{bmatrix}$$

Analysis of the augmented matrix for the homogenous system (Notation RREFA [31]). Notice the slight variation for the same analysis of the original system only when the original system was consistent:

$$r = 3 \qquad D = \{1,\, 2,\, 3\} \qquad F = \{4,\, 5\}$$

Coefficient matrix of original system of equations, and of associated homogenous system. This matrix will be the subject of further analysis, rather than the systems of equations.

$$\begin{bmatrix} 2 & -3 & 1 & -6 \\ 4 & 1 & 2 & 9 \\ 3 & 1 & 1 & 8 \end{bmatrix}$$

□ Matrix brought to reduced row-echelon form:

$$\begin{bmatrix} \boxed{1} & 0 & 0 & 2 \\ 0 & \boxed{1} & 0 & 3 \\ 0 & 0 & \boxed{1} & -1 \end{bmatrix}$$

□ Analysis of the row-reduced matrix (Notation RREFA [31]):

$$r = 3 \qquad D = \{1,\, 2,\, 3\} \qquad F = \{4\}$$

□ This is the null space of the matrix. The set of vectors used in the span construction is a linearly independent set of column vectors that spans the null space of the matrix (Theorem SSNS [81], Theorem BNS [94]). Solve the homogenous system with this matrix as the coefficient matrix and write the solutions in vector form (Theorem VFSLS [72]) to see these vectors arise.

$$\mathrm{Sp}\left(\left\{\begin{bmatrix} -2 \\ -3 \\ 1 \\ 1 \end{bmatrix}\right\}\right)$$

□ Range of the matrix, expressed as the span of a set of linearly independent vectors that are also columns of the matrix. These columns have indices that form the set D above. (Theorem BROC [109])

$$\mathrm{Sp}\left(\left\{\begin{bmatrix} 2 \\ 4 \\ 3 \end{bmatrix},\ \begin{bmatrix} -3 \\ 1 \\ 1 \end{bmatrix},\ \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}\right\}\right)$$

□ The range of the matrix, as it arises from row operations on an augmented matrix. The matrix K is computed as described in Theorem RNS [114]. This is followed by the range described by a set of linearly independent vectors that span the null space of K, computed as according to Theorem SSNS [81]. When $r = m$, the matrix K has no rows and the range is all of $\mathbb{C}^m$.

$$K = [\ ]$$

$$\text{Sp}\left(\left\{\begin{bmatrix}1\\0\\0\end{bmatrix},\,\begin{bmatrix}0\\1\\0\end{bmatrix},\,\begin{bmatrix}0\\0\\1\end{bmatrix}\right\}\right)$$

□ Range of the matrix, expressed as the span of a set of linearly independent vectors. These vectors are computed by row-reducing the transpose the matrix into reduced row-echelon form, tossing out the zero rows, and writing the remaining rows as column vectors. By Theorem RMSRT [124] and Theorem BRS [123], and in the style of Example RSOM.IS [123], this yields a linearly independent set of vectors that span the range.

$$\text{Sp}\left(\left\{\begin{bmatrix}1\\0\\0\end{bmatrix},\,\begin{bmatrix}0\\1\\0\end{bmatrix},\,\begin{bmatrix}0\\0\\1\end{bmatrix}\right\}\right)$$

□ Row space of the matrix, expressed as a span of a set of linearly independent vectors, obtained from the nonzero rows of the equivalent matrix in reduced row-echelon form. (Theorem BRS [123])

$$\text{Sp}\left(\left\{\begin{bmatrix}1\\0\\0\\2\end{bmatrix},\,\begin{bmatrix}0\\1\\0\\3\end{bmatrix},\,\begin{bmatrix}0\\0\\1\\-1\end{bmatrix}\right\}\right)$$

□ Subspace dimensions associated with the matrix.

Matrix columns: 4 Rank: 3 Nullity: 1

Archetype D

Summary System with three equations, four variables. Consistent. Null space of coefficient matrix has dimension 2. Coefficient matrix identical to that of Archetype E, vector of constants is different.

□ A system of linear equations (Definition SSLE [8]):

$$\begin{aligned} 2x_1 + x_2 + 7x_3 - 7x_4 &= 8 \\ -3x_1 + 4x_2 - 5x_3 - 6x_4 &= -12 \\ x_1 + x_2 + 4x_3 - 5x_4 &= 4 \end{aligned}$$

□ Some solutions to the system of linear equations (not necessarily exhaustive):

$x_1 = 0, \quad x_2 = 1, \quad x_3 = 2, \quad x_4 = 1$

$x_1 = 4, \quad x_2 = 0, \quad x_3 = 0, \quad x_4 = 0$

$x_1 = 7, \quad x_2 = 8, \quad x_3 = 1, \quad x_4 = 3$

□ Augmented matrix of the linear system of equations (Definition AM [20]):

$$\begin{bmatrix} 2 & 1 & 7 & -7 & 8 \\ -3 & 4 & -5 & -6 & -12 \\ 1 & 1 & 4 & -5 & 4 \end{bmatrix}$$

□ Matrix in reduced row-echelon form, row-equivalent to augmented matrix:

$$\begin{bmatrix} \boxed{1} & 0 & 3 & -2 & 4 \\ 0 & \boxed{1} & 1 & -3 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

□ Analysis of the augmented matrix (Notation RREFA [31]):

$$r = 2 \qquad D = \{1,\, 2\} \qquad F = \{3,\, 4,\, 5\}$$

□ Vector form of the solution set to the system of equations (Theorem VFSLS [72]). Notice the relationship between the free variables and the set F above. Also, notice the

pattern of 0's and 1's in the entries of the vectors corresponding to elements of the set F for the larger examples.

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 4 \\ 0 \\ 0 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -3 \\ -1 \\ 1 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} 2 \\ 3 \\ 0 \\ 1 \end{bmatrix}$$

☐ A system of equations can always be converted to a homogenous system (Definition HS [41]):

$$\begin{aligned} 2x_1 + x_2 + 7x_3 - 7x_4 &= 0 \\ -3x_1 + 4x_2 - 5x_3 - 6x_4 &= 0 \\ x_1 + x_2 + 4x_3 - 5x_4 &= 0 \end{aligned}$$

☐ Some solutions to the associated homogenous system of linear equations (not necessarily exhaustive):

$x_1 = 0, \quad x_2 = 0, \quad x_3 = 0, \quad x_4 = 0$

$x_1 = -3, \quad x_2 = -1, \quad x_3 = 1, \quad x_4 = 0$

$x_1 = 2, \quad x_2 = 3, \quad x_3 = 0, \quad x_4 = 1$

$x_1 = -1, \quad x_2 = 2, \quad x_3 = 1, \quad x_4 = 1$

☐ Form the augmented matrix of the homogenous linear system, and use row operations to convert to reduced row-echelon form. Notice how the entries of the final column remain zeros:

$$\begin{bmatrix} \boxed{1} & 0 & 3 & -2 & 0 \\ 0 & \boxed{1} & 1 & -3 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

☐ Analysis of the augmented matrix for the homogenous system (Notation RREFA [31]). Notice the slight variation for the same analysis of the original system only when the original system was consistent:

$$r = 2 \qquad D = \{1,\,2\} \qquad F = \{3,\,4,\,5\}$$

☐ Coefficient matrix of original system of equations, and of associated homogenous system. This matrix will be the subject of further analysis, rather than the systems of

equations.

$$\begin{bmatrix} 2 & 1 & 7 & -7 \\ -3 & 4 & -5 & -6 \\ 1 & 1 & 4 & -5 \end{bmatrix}$$

☐ Matrix brought to reduced row-echelon form:

$$\begin{bmatrix} \boxed{1} & 0 & 3 & -2 \\ 0 & \boxed{1} & 1 & -3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

☐ Analysis of the row-reduced matrix (Notation RREFA [31]):

$$r = 2 \qquad D = \{1,\, 2\} \qquad F = \{3,\, 4\}$$

☐ This is the null space of the matrix. The set of vectors used in the span construction is a linearly independent set of column vectors that spans the null space of the matrix (Theorem SSNS [81], Theorem BNS [94]). Solve the homogenous system with this matrix as the coefficient matrix and write the solutions in vector form (Theorem VFSLS [72]) to see these vectors arise.

$$\mathrm{Sp}\left(\left\{\begin{bmatrix} -3 \\ -1 \\ 1 \\ 0 \end{bmatrix},\, \begin{bmatrix} 2 \\ 3 \\ 0 \\ 1 \end{bmatrix}\right\}\right)$$

☐ Range of the matrix, expressed as the span of a set of linearly independent vectors that are also columns of the matrix. These columns have indices that form the set D above. (Theorem BROC [109])

$$\mathrm{Sp}\left(\left\{\begin{bmatrix} 2 \\ -3 \\ 1 \end{bmatrix},\, \begin{bmatrix} 1 \\ 4 \\ 1 \end{bmatrix}\right\}\right)$$

☐ The range of the matrix, as it arises from row operations on an augmented matrix. The matrix K is computed as described in Theorem RNS [114]. This is followed by the range described by a set of linearly independent vectors that span the null space of K, computed as according to Theorem SSNS [81]. When $r = m$, the matrix K has no rows and the range is all of $\mathbb{C}^m$.

$K = \begin{bmatrix} 1 & \frac{1}{7} & -\frac{11}{7} \end{bmatrix}$

$$\mathrm{Sp}\left(\left\{\begin{bmatrix}\frac{11}{7}\\0\\1\end{bmatrix},\ \begin{bmatrix}-\frac{1}{7}\\1\\0\end{bmatrix}\right\}\right)$$

☐ Range of the matrix, expressed as the span of a set of linearly independent vectors. These vectors are computed by row-reducing the transpose the matrix into reduced row-echelon form, tossing out the zero rows, and writing the remaining rows as column vectors. By Theorem RMSRT [124] and Theorem BRS [123], and in the style of Example RSOM.IS [123], this yields a linearly independent set of vectors that span the range.

$$\mathrm{Sp}\left(\left\{\begin{bmatrix}1\\0\\\frac{7}{11}\end{bmatrix},\ \begin{bmatrix}0\\1\\\frac{1}{11}\end{bmatrix}\right\}\right)$$

☐ Row space of the matrix, expressed as a span of a set of linearly independent vectors, obtained from the nonzero rows of the equivalent matrix in reduced row-echelon form. (Theorem BRS [123])

$$\mathrm{Sp}\left(\left\{\begin{bmatrix}1\\0\\3\\-2\end{bmatrix},\ \begin{bmatrix}0\\1\\1\\-3\end{bmatrix}\right\}\right)$$

☐ Subspace dimensions associated with the matrix.

Matrix columns: 4 Rank: 2 Nullity: 2

Archetype E

Summary System with three equations, four variables. Inconsistent. Null space of coefficient matrix has dimension 2. Coefficient matrix identical to that of Archetype D, constant vector is different.

□ A system of linear equations (Definition SSLE [8]):

$$\begin{aligned} 2x_1 + x_2 + 7x_3 - 7x_4 &= 2 \\ -3x_1 + 4x_2 - 5x_3 - 6x_4 &= 3 \\ x_1 + x_2 + 4x_3 - 5x_4 &= 2 \end{aligned}$$

□ Some solutions to the system of linear equations (not necessarily exhaustive):

None. (Why?)

□ Augmented matrix of the linear system of equations (Definition AM [20]):

$$\begin{bmatrix} 2 & 1 & 7 & -7 & 2 \\ -3 & 4 & -5 & -6 & 3 \\ 1 & 1 & 4 & -5 & 2 \end{bmatrix}$$

□ Matrix in reduced row-echelon form, row-equivalent to augmented matrix:

$$\begin{bmatrix} \boxed{1} & 1 & 4 & -5 & 0 \\ 0 & \boxed{1} & 1 & -3 & 0 \\ 0 & 0 & 0 & 0 & \boxed{1} \end{bmatrix}$$

□ Analysis of the augmented matrix (Notation RREFA [31]):

$$r = 3 \qquad D = \{1,\,2,\,5\} \qquad F = \{3,\,4\}$$

□ Vector form of the solution set to the system of equations (Theorem VFSLS [72]). Notice the relationship between the free variables and the set F above. Also, notice the pattern of 0's and 1's in the entries of the vectors corresponding to elements of the set F for the larger examples.

Inconsistent system, no solutions exist.

□ A system of equations can always be converted to a homogenous system (Definition HS [41]):

$$\begin{aligned} 2x_1 + x_2 + 7x_3 - 7x_4 &= 0 \\ -3x_1 + 4x_2 - 5x_3 - 6x_4 &= 0 \\ x_1 + x_2 + 4x_3 - 5x_4 &= 0 \end{aligned}$$

□ Some solutions to the associated homogenous system of linear equations (not necessarily exhaustive):
$x_1 = 0, \quad x_2 = 0, \quad x_3 = 0, \quad x_4 = 0$

$x_1 = 4, \quad x_2 = 13, \quad x_3 = 2, \quad x_4 = 5$

□ Form the augmented matrix of the homogenous linear system, and use row operations to convert to reduced row-echelon form. Notice how the entries of the final column remain zeros:

$$\begin{bmatrix} \boxed{1} & 0 & 3 & -2 & 0 \\ 0 & \boxed{1} & 1 & -3 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

□ Analysis of the augmented matrix for the homogenous system (Notation RREFA [31]). Notice the slight variation for the same analysis of the original system only when the original system was consistent:

$$r = 2 \qquad D = \{1, 2\} \qquad F = \{3, 4, 5\}$$

□ Coefficient matrix of original system of equations, and of associated homogenous system. This matrix will be the subject of further analysis, rather than the systems of equations.

$$\begin{bmatrix} 2 & 1 & 7 & -7 \\ -3 & 4 & -5 & -6 \\ 1 & 1 & 4 & -5 \end{bmatrix}$$

□ Matrix brought to reduced row-echelon form:

$$\begin{bmatrix} \boxed{1} & 0 & 3 & -2 \\ 0 & \boxed{1} & 1 & -3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

□ Analysis of the row-reduced matrix (Notation RREFA [31]):

$$r = 2 \qquad D = \{1,\, 2\} \qquad F = \{3,\, 4\}$$

□ This is the null space of the matrix. The set of vectors used in the span construction is a linearly independent set of column vectors that spans the null space of the matrix (Theorem SSNS [81], Theorem BNS [94]). Solve the homogenous system with this matrix as the coefficient matrix and write the solutions in vector form (Theorem VFSLS [72]) to see these vectors arise.

$$\mathrm{Sp}\left(\left\{\begin{bmatrix}-3\\-1\\1\\0\end{bmatrix},\,\begin{bmatrix}2\\3\\0\\1\end{bmatrix}\right\}\right)$$

□ Range of the matrix, expressed as the span of a set of linearly independent vectors that are also columns of the matrix. These columns have indices that form the set D above. (Theorem BROC [109])

$$\mathrm{Sp}\left(\left\{\begin{bmatrix}2\\-3\\1\end{bmatrix},\,\begin{bmatrix}1\\4\\1\end{bmatrix}\right\}\right)$$

□ The range of the matrix, as it arises from row operations on an augmented matrix. The matrix K is computed as described in Theorem RNS [114]. This is followed by the range described by a set of linearly independent vectors that span the null space of K, computed as according to Theorem SSNS [81]. When $r = m$, the matrix K has no rows and the range is all of $\mathbb{C}^m$.

$$K = \begin{bmatrix} 1 & \frac{1}{7} & -\frac{11}{7} \end{bmatrix}$$

$$\text{Sp}\left(\left\{\begin{bmatrix}\frac{11}{7}\\0\\1\end{bmatrix},\begin{bmatrix}-\frac{1}{7}\\1\\0\end{bmatrix}\right\}\right)$$

□ Range of the matrix, expressed as the span of a set of linearly independent vectors. These vectors are computed by row-reducing the transpose the matrix into reduced row-echelon form, tossing out the zero rows, and writing the remaining rows as column vectors. By THEOREM RMSRT [124] and THEOREM BRS [123], and in the style of EXAMPLE RSOM.IS [123], this yields a linearly independent set of vectors that span the range.

$$\text{Sp}\left(\left\{\begin{bmatrix}1\\0\\\frac{7}{11}\end{bmatrix},\begin{bmatrix}0\\1\\\frac{1}{11}\end{bmatrix}\right\}\right)$$

□ Row space of the matrix, expressed as a span of a set of linearly independent vectors, obtained from the nonzero rows of the equivalent matrix in reduced row-echelon form. (THEOREM BRS [123])

$$\text{Sp}\left(\left\{\begin{bmatrix}1\\0\\3\\-2\end{bmatrix},\begin{bmatrix}0\\1\\1\\-3\end{bmatrix}\right\}\right)$$

□ Subspace dimensions associated with the matrix.

Matrix columns: 4 Rank: 2 Nullity: 2

Archetype F

Summary System with four equations, four variables. Nonsingular coefficient matrix. Integer eigenvalues, one has "high" multiplicity.

□ A system of linear equations (Definition SSLE [8]):

$$\begin{aligned}33x_1 - 16x_2 + 10x_3 - 2x_4 &= -27\\ 99x_1 - 47x_2 + 27x_3 - 7x_4 &= -77\\ 78x_1 - 36x_2 + 17x_3 - 6x_4 &= -52\\ -9x_1 + 2x_2 + 3x_3 + 4x_4 &= 5\end{aligned}$$

□ Some solutions to the system of linear equations (not necessarily exhaustive):

$x_1 = 1, \quad x_2 = 2, \quad x_3 = -2, \quad x_4 = 4$

□ Augmented matrix of the linear system of equations (Definition AM [20]):

$$\begin{bmatrix} 33 & -16 & 10 & -2 & -27\\ 99 & -47 & 27 & -7 & -77\\ 78 & -36 & 17 & -6 & -52\\ -9 & 2 & 3 & 4 & 5\end{bmatrix}$$

□ Matrix in reduced row-echelon form, row-equivalent to augmented matrix:

$$\begin{bmatrix} \boxed{1} & 0 & 0 & 0 & 1\\ 0 & \boxed{1} & 0 & 0 & 2\\ 0 & 0 & \boxed{1} & 0 & -2\\ 0 & 0 & 0 & \boxed{1} & 4\end{bmatrix}$$

□ Analysis of the augmented matrix (Notation RREFA [31]):

$$r = 4 \qquad D = \{1,\, 2,\, 3,\, 4\} \qquad F = \{5\}$$

□ Vector form of the solution set to the system of equations (Theorem VFSLS [72]). Notice the relationship between the free variables and the set F above. Also, notice the pattern of 0's and 1's in the entries of the vectors corresponding to elements of the set F for the larger examples.

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ -2 \\ 4 \end{bmatrix}$$

□ A system of equations can always be converted to a homogenous system (Definition HS [41]):

$$\begin{aligned} 33x_1 - 16x_2 + 10x_3 - 2x_4 &= 0 \\ 99x_1 - 47x_2 + 27x_3 - 7x_4 &= 0 \\ 78x_1 - 36x_2 + 17x_3 - 6x_4 &= 0 \\ -9x_1 + 2x_2 + 3x_3 + 4x_4 &= 0 \end{aligned}$$

□ Some solutions to the associated homogenous system of linear equations (not necessarily exhaustive):
$x_1 = 0, \quad x_2 = 0, \quad x_3 = 0, \quad x_4 = 0$

□ Form the augmented matrix of the homogenous linear system, and use row operations to convert to reduced row-echelon form. Notice how the entries of the final column remain zeros:

$$\begin{bmatrix} \boxed{1} & 0 & 0 & 0 & 0 \\ 0 & \boxed{1} & 0 & 0 & 0 \\ 0 & 0 & \boxed{1} & 0 & 0 \\ 0 & 0 & 0 & \boxed{1} & 0 \end{bmatrix}$$

□ Analysis of the augmented matrix for the homogenous system (Notation RREFA [31]). Notice the slight variation for the same analysis of the original system only when the original system was consistent:

$$r = 4 \qquad D = \{1,\, 2,\, 3,\, 4\} \qquad F = \{5\}$$

□ Coefficient matrix of original system of equations, and of associated homogenous system. This matrix will be the subject of further analysis, rather than the systems of equations.

$$\begin{bmatrix} 33 & -16 & 10 & -2 \\ 99 & -47 & 27 & -7 \\ 78 & -36 & 17 & -6 \\ -9 & 2 & 3 & 4 \end{bmatrix}$$

☐ Matrix brought to reduced row-echelon form:

$$\begin{bmatrix} \boxed{1} & 0 & 0 & 0 \\ 0 & \boxed{1} & 0 & 0 \\ 0 & 0 & \boxed{1} & 0 \\ 0 & 0 & 0 & \boxed{1} \end{bmatrix}$$

☐ Analysis of the row-reduced matrix (Notation RREFA [31]):

$$r = 4 \qquad D = \{1,\, 2,\, 3,\, 4\} \qquad F = \{\ \}$$

☐ Matrix (coefficient matrix) is nonsingular or singular? (Theorem NSRRI [50]) at the same time, examine the size of the set F above.Notice that this property does not apply to matrices that are not square.

Nonsingular.

☐ This is the null space of the matrix. The set of vectors used in the span construction is a linearly independent set of column vectors that spans the null space of the matrix (Theorem SSNS [81], Theorem BNS [94]). Solve the homogenous system with this matrix as the coefficient matrix and write the solutions in vector form (Theorem VFSLS [72]) to see these vectors arise.

$\mathrm{Sp}\left(\{\ \}\right)$

☐ Range of the matrix, expressed as the span of a set of linearly independent vectors that are also columns of the matrix. These columns have indices that form the set D above. (Theorem BROC [109])

$$\mathrm{Sp}\left(\left\{\begin{bmatrix} 33 \\ 99 \\ 78 \\ -9 \end{bmatrix},\, \begin{bmatrix} -16 \\ -47 \\ -36 \\ 2 \end{bmatrix},\, \begin{bmatrix} 10 \\ 27 \\ 17 \\ 3 \end{bmatrix},\, \begin{bmatrix} -2 \\ -7 \\ -6 \\ 4 \end{bmatrix}\right\}\right)$$

☐ The range of the matrix, as it arises from row operations on an augmented matrix. The matrix K is computed as described in Theorem RNS [114]. This is followed by the range described by a set of linearly independent vectors that span the null space of K, computed as according to Theorem SSNS [81]. When $r = m$, the matrix K has no rows and the range is all of $\mathbb{C}^m$.

$K = []$

$$\text{Sp}\left(\left\{\begin{bmatrix}1\\0\\0\\0\end{bmatrix}, \begin{bmatrix}0\\1\\0\\0\end{bmatrix}, \begin{bmatrix}0\\0\\1\\0\end{bmatrix}, \begin{bmatrix}0\\0\\0\\1\end{bmatrix}\right\}\right)$$

□ Range of the matrix, expressed as the span of a set of linearly independent vectors. These vectors are computed by row-reducing the transpose the matrix into reduced row-echelon form, tossing out the zero rows, and writing the remaining rows as column vectors. By Theorem RMSRT [124] and Theorem BRS [123], and in the style of Example RSOM.IS [123], this yields a linearly independent set of vectors that span the range.

$$\text{Sp}\left(\left\{\begin{bmatrix}1\\0\\0\\0\end{bmatrix}, \begin{bmatrix}0\\1\\0\\0\end{bmatrix}, \begin{bmatrix}0\\0\\1\\0\end{bmatrix}, \begin{bmatrix}0\\0\\0\\1\end{bmatrix}\right\}\right)$$

□ Row space of the matrix, expressed as a span of a set of linearly independent vectors, obtained from the nonzero rows of the equivalent matrix in reduced row-echelon form. (Theorem BRS [123])

$$\text{Sp}\left(\left\{\begin{bmatrix}1\\0\\0\\0\end{bmatrix}, \begin{bmatrix}0\\1\\0\\0\end{bmatrix}, \begin{bmatrix}0\\0\\1\\0\end{bmatrix}, \begin{bmatrix}0\\0\\0\\1\end{bmatrix}\right\}\right)$$

□ Inverse matrix, if it exists. The inverse is not defined for matrices that are not square, and if the matrix is square, then the matrix must be nonsingular. (Definition MI [141], Theorem NSI [151])

$$\begin{bmatrix}-\left(\frac{86}{3}\right) & \frac{38}{3} & -\left(\frac{11}{3}\right) & \frac{7}{3}\\ -\left(\frac{129}{2}\right) & \frac{86}{3} & -\left(\frac{17}{2}\right) & \frac{31}{6}\\ -13 & 6 & -2 & 1\\ -\left(\frac{45}{2}\right) & \frac{29}{3} & -\left(\frac{5}{2}\right) & \frac{13}{6}\end{bmatrix}$$

□ Subspace dimensions associated with the matrix.

Matrix columns: 4 Rank: 4 Nullity: 0

☐ Determinant of the matrix, which is only defined for square matrices. (Zero/nonzero? Singular/nonsingular? Product of all eigenvalues?)

$$\text{Determinant} = -18$$

Archetype G

Summary System with five equations, two variables. Consistent. Null space of coefficient matrix has dimension 0. Coefficient matrix identical to that of Archetype H, constant vector is different.

□ A system of linear equations (Definition SSLE [8]):

$$\begin{aligned} 2x_1 + 3x_2 &= 6 \\ -x_1 + 4x_2 &= -14 \\ 3x_1 + 10x_2 &= -2 \\ 3x_1 - x_2 &= 20 \\ 6x_1 + 9x_2 &= 18 \end{aligned}$$

□ Some solutions to the system of linear equations (not necessarily exhaustive):

$x_1 = 6, \quad x_2 = -2$

□ Augmented matrix of the linear system of equations (Definition AM [20]):

$$\begin{bmatrix} 2 & 3 & 6 \\ -1 & 4 & -14 \\ 3 & 10 & -2 \\ 3 & -1 & 20 \\ 6 & 9 & 18 \end{bmatrix}$$

□ Matrix in reduced row-echelon form, row-equivalent to augmented matrix:

$$\begin{bmatrix} \boxed{1} & 0 & 6 \\ 0 & \boxed{1} & -2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

□ Analysis of the augmented matrix (Notation RREFA [31]):

$$r = 2 \qquad D = \{1,\, 2\} \qquad F = \{3\}$$

□ Vector form of the solution set to the system of equations (Theorem VFSLS [72]). Notice the relationship between the free variables and the set F above. Also, notice the

pattern of 0's and 1's in the entries of the vectors corresponding to elements of the set F for the larger examples.

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 6 \\ -2 \end{bmatrix}$$

☐ A system of equations can always be converted to a homogenous system (Definition HS [41]):

$$\begin{aligned} 2x_1 + 3x_2 &= 0 \\ -x_1 + 4x_2 &= 0 \\ 3x_1 + 10x_2 &= 0 \\ 3x_1 - x_2 &= 0 \\ 6x_1 + 9x_2 &= 0 \end{aligned}$$

☐ Some solutions to the associated homogenous system of linear equations (not necessarily exhaustive):
$x_1 = 0, \quad x_2 = 0$

☐ Form the augmented matrix of the homogenous linear system, and use row operations to convert to reduced row-echelon form. Notice how the entries of the final column remain zeros:

$$\begin{bmatrix} \boxed{1} & 0 & 0 \\ 0 & \boxed{1} & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

☐ Analysis of the augmented matrix for the homogenous system (Notation RREFA [31]). Notice the slight variation for the same analysis of the original system only when the original system was consistent:

$$r = 2 \qquad D = \{1,\, 2\} \qquad F = \{3\}$$

☐ Coefficient matrix of original system of equations, and of associated homogenous system. This matrix will be the subject of further analysis, rather than the systems of

equations.

$$\begin{bmatrix} 2 & 3 \\ -1 & 4 \\ 3 & 10 \\ 3 & -1 \\ 6 & 9 \end{bmatrix}$$

□ Matrix brought to reduced row-echelon form:

$$\begin{bmatrix} \boxed{1} & 0 \\ 0 & \boxed{1} \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

□ Analysis of the row-reduced matrix (Notation RREFA [31]):

$$r = 2 \qquad D = \{1,\, 2\} \qquad F = \{\ \}$$

□ This is the null space of the matrix. The set of vectors used in the span construction is a linearly independent set of column vectors that spans the null space of the matrix (Theorem SSNS [81], Theorem BNS [94]). Solve the homogenous system with this matrix as the coefficient matrix and write the solutions in vector form (Theorem VFSLS [72]) to see these vectors arise.

$$\mathrm{Sp}\left(\{\ \}\right)$$

□ Range of the matrix, expressed as the span of a set of linearly independent vectors that are also columns of the matrix. These columns have indices that form the set D above. (Theorem BROC [109])

$$\mathrm{Sp}\left(\left\{\begin{bmatrix} 2 \\ -1 \\ 3 \\ 3 \\ 6 \end{bmatrix},\, \begin{bmatrix} 3 \\ 4 \\ 10 \\ -1 \\ 9 \end{bmatrix}\right\}\right)$$

□ The range of the matrix, as it arises from row operations on an augmented matrix. The matrix K is computed as described in Theorem RNS [114]. This is followed by the range described by a set of linearly independent vectors that span the null space of K, computed as according to THEOREM SSNS [81]. When $r = m$, the matrix K has no rows and the range is all of $\mathbb{C}^m$.

$$K = \begin{bmatrix} 1 & 0 & 0 & 0 & -\frac{1}{3} \\ 0 & 1 & 0 & 1 - \frac{1}{3} & \\ 0 & 0 & 1 & 1 & -1 \end{bmatrix}$$

$$\mathrm{Sp}\left(\left\{\begin{bmatrix} \frac{1}{3} \\ \frac{1}{3} \\ 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \\ -1 \\ 1 \\ 0 \end{bmatrix}\right\}\right)$$

☐ Range of the matrix, expressed as the span of a set of linearly independent vectors. These vectors are computed by row-reducing the transpose the matrix into reduced row-echelon form, tossing out the zero rows, and writing the remaining rows as column vectors. By THEOREM RMSRT [124] and THEOREM BRS [123], and in the style of EXAMPLE RSOM.IS [123], this yields a linearly independent set of vectors that span the range.

$$\mathrm{Sp}\left(\left\{\begin{bmatrix} 1 \\ 0 \\ 2 \\ 1 \\ 3 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \\ -1 \\ 0 \end{bmatrix}\right\}\right)$$

☐ Row space of the matrix, expressed as a span of a set of linearly independent vectors, obtained from the nonzero rows of the equivalent matrix in reduced row-echelon form. (THEOREM BRS [123])

$$\mathrm{Sp}\left(\left\{\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}\right\}\right)$$

☐ Subspace dimensions associated with the matrix.

Matrix columns: 2 Rank: 2 Nullity: 0

Archetype H

Summary System with five equations, two variables. Inconsistent, overdetermined. Null space of coefficient matrix has dimension 0. Coefficient matrix identical to that of Archetype G, constant vector is different.

□ A system of linear equations (Definition SSLE [8]):

$$\begin{aligned} 2x_1 + 3x_2 &= 5 \\ -x_1 + 4x_2 &= 6 \\ 3x_1 + 10x_2 &= 2 \\ 3x_1 - x_2 &= -1 \\ 6x_1 + 9x_2 &= 3 \end{aligned}$$

□ Some solutions to the system of linear equations (not necessarily exhaustive):

None. (Why?)

□ Augmented matrix of the linear system of equations (Definition AM [20]):

$$\begin{bmatrix} 2 & 3 & 5 \\ -1 & 4 & 6 \\ 3 & 10 & 2 \\ 3 & -1 & -1 \\ 6 & 9 & 3 \end{bmatrix}$$

□ Matrix in reduced row-echelon form, row-equivalent to augmented matrix:

$$\begin{bmatrix} \boxed{1} & 0 & 0 \\ 0 & \boxed{1} & 0 \\ 0 & 0 & \boxed{1} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

□ Analysis of the augmented matrix (Notation RREFA [31]):

$$r = 3 \qquad D = \{1,\, 2,\, 3\} \qquad F = \{\ \}$$

□ Vector form of the solution set to the system of equations (Theorem VFSLS [72]). Notice the relationship between the free variables and the set F above. Also, notice the

pattern of 0's and 1's in the entries of the vectors corresponding to elements of the set F for the larger examples.

Inconsistent system, no solutions exist.

☐ A system of equations can always be converted to a homogenous system (Definition HS [41]):

$$\begin{aligned} 2x_1 + 3x_2 &= 0 \\ -x_1 + 4x_2 &= 0 \\ 3x_1 + 10x_2 &= 0 \\ 3x_1 - x_2 &= 0 \\ 6x_1 + 9x_2 &= 0 \end{aligned}$$

☐ Some solutions to the associated homogenous system of linear equations (not necessarily exhaustive):
$x_1 = 0, \quad x_2 = 0$

☐ Form the augmented matrix of the homogenous linear system, and use row operations to convert to reduced row-echelon form. Notice how the entries of the final column remain zeros:

$$\begin{bmatrix} \boxed{1} & 0 & 0 \\ 0 & \boxed{1} & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

☐ Analysis of the augmented matrix for the homogenous system (Notation RREFA [31]). Notice the slight variation for the same analysis of the original system only when the original system was consistent:

$$r = 2 \qquad D = \{1,\, 2\} \qquad F = \{3\}$$

☐ Coefficient matrix of original system of equations, and of associated homogenous system. This matrix will be the subject of further analysis, rather than the systems of

equations.

$$\begin{bmatrix} 2 & 3 \\ -1 & 4 \\ 3 & 10 \\ 3 & -1 \\ 6 & 9 \end{bmatrix}$$

☐ Matrix brought to reduced row-echelon form:

$$\begin{bmatrix} \boxed{1} & 0 \\ 0 & \boxed{1} \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

☐ Analysis of the row-reduced matrix (Notation RREFA [31]):

$$r = 2 \qquad D = \{1,\, 2\} \qquad F = \{\ \}$$

☐ This is the null space of the matrix. The set of vectors used in the span construction is a linearly independent set of column vectors that spans the null space of the matrix (Theorem SSNS [81], Theorem BNS [94]). Solve the homogenous system with this matrix as the coefficient matrix and write the solutions in vector form (Theorem VFSLS [72]) to see these vectors arise.

$\mathrm{Sp}\left(\{\ \}\right)$

☐ Range of the matrix, expressed as the span of a set of linearly independent vectors that are also columns of the matrix. These columns have indices that form the set D above. (Theorem BROC [109])

$$\mathrm{Sp}\left(\left\{\begin{bmatrix} 2 \\ -1 \\ 3 \\ 3 \\ 6 \end{bmatrix},\, \begin{bmatrix} 3 \\ 4 \\ 10 \\ -1 \\ 9 \end{bmatrix}\right\}\right)$$

☐ The range of the matrix, as it arises from row operations on an augmented matrix. The matrix K is computed as described in Theorem RNS [114]. This is followed by the range described by a set of linearly independent vectors that span the null space of K, computed as according to Theorem SSNS [81]. When $r = m$, the matrix K has no rows and the range is all of $\mathbb{C}^m$.

$$K = []$$

$$\mathrm{Sp}\left(\left\{\begin{bmatrix}\frac{1}{3}\\ \frac{1}{3}\\ 1\\ 0\\ 1\end{bmatrix},\ \begin{bmatrix}0\\ -1\\ -1\\ 1\\ 0\end{bmatrix}\right\}\right)$$

□ Range of the matrix, expressed as the span of a set of linearly independent vectors. These vectors are computed by row-reducing the transpose the matrix into reduced row-echelon form, tossing out the zero rows, and writing the remaining rows as column vectors. By Theorem RMSRT [124] and Theorem BRS [123], and in the style of Example RSOM.IS [123], this yields a linearly independent set of vectors that span the range.

$$\mathrm{Sp}\left(\left\{\begin{bmatrix}1\\ 0\\ 2\\ 1\\ 3\end{bmatrix},\ \begin{bmatrix}0\\ 1\\ 1\\ -1\\ 0\end{bmatrix}\right\}\right)$$

□ The range of the matrix, as it arises from row operations on an augmented matrix. The matrix K is computed as described in Theorem RNS [114]. This is followed by the range described by a set of linearly independent vectors that span the null space of K, computed as according to Theorem SSNS [81]. When $r = m$, the matrix K has no rows and the range is all of $\mathbb{C}^m$.

$$K = \begin{bmatrix}1 & 0 & 0 & 0 & -\frac{1}{3}\\ 0 & 1 & 0 & 1-\frac{1}{3} & \\ 0 & 0 & 1 & 1 & -1\end{bmatrix}$$

$$\mathrm{Sp}\left(\left\{\begin{bmatrix}\frac{1}{3}\\ \frac{1}{3}\\ 1\\ 0\\ 1\end{bmatrix},\ \begin{bmatrix}0\\ -1\\ -1\\ 1\\ 0\end{bmatrix}\right\}\right)$$

□ Row space of the matrix, expressed as a span of a set of linearly independent vectors, obtained from the nonzero rows of the equivalent matrix in reduced row-echelon form.

(Theorem BRS [123])

$$\mathrm{Sp}\left(\left\{\begin{bmatrix}1\\0\end{bmatrix},\,\begin{bmatrix}0\\1\end{bmatrix}\right\}\right)$$

☐ Subspace dimensions associated with the matrix.

Matrix columns: 2   Rank: 2   Nullity: 0

Archetype I

Summary System with four equations, seven variables. Consistent. Null space of coefficient matrix has dimension 4.

□ A system of linear equations (Definition SSLE [8]):

$$\begin{aligned} x_1 + 4x_2 - x_4 + 7x_6 - 9x_7 &= 3 \\ 2x_1 + 8x_2 - x_3 + 3x_4 + 9x_5 - 13x_6 + 7x_7 &= 9 \\ 2x_3 - 3x_4 - 4x_5 + 12x_6 - 8x_7 &= 1 \\ -x_1 - 4x_2 + 2x_3 + 4x_4 + 8x_5 - 31x_6 + 37x_7 &= 4 \end{aligned}$$

□ Some solutions to the system of linear equations (not necessarily exhaustive):

$x_1 = -25,\ x_2 = 4,\ x_3 = 22,\ x_4 = 29,\ x_5 = 1,\ x_6 = 2,\ x_7 = -3$

$x_1 = -7,\ x_2 = 5,\ x_3 = 7,\ x_4 = 15,\ x_5 = -4,\ x_6 = 2,\ x_7 = 1$

$x_1 = 4,\ x_2 = 0,\ x_3 = 2,\ x_4 = 1,\ x_5 = 0,\ x_6 = 0,\ x_7 = 0$

□ Augmented matrix of the linear system of equations (Definition AM [20]):

$$\begin{bmatrix} 1 & 4 & 0 & -1 & 0 & 7 & -9 & 3 \\ 2 & 8 & -1 & 3 & 9 & -13 & 7 & 9 \\ 0 & 0 & 2 & -3 & -4 & 12 & -8 & 1 \\ -1 & -4 & 2 & 4 & 8 & -31 & 37 & 4 \end{bmatrix}$$

□ Matrix in reduced row-echelon form, row-equivalent to augmented matrix:

$$\begin{bmatrix} \boxed{1} & 4 & 0 & 0 & 2 & 1 & -3 & 4 \\ 0 & 0 & \boxed{1} & 0 & 1 & -3 & 5 & 2 \\ 0 & 0 & 0 & \boxed{1} & 2 & -6 & 6 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

□ Analysis of the augmented matrix (Notation RREFA [31]):

$$r = 3 \qquad D = \{1,\, 3,\, 4\} \qquad F = \{2,\, 5,\, 6,\, 7,\, 8\}$$

□ Vector form of the solution set to the system of equations (Theorem VFSLS [72]). Notice the relationship between the free variables and the set F above. Also, notice the

pattern of 0's and 1's in the entries of the vectors corresponding to elements of the set F for the larger examples.

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \end{bmatrix} = \begin{bmatrix} 4 \\ 0 \\ 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} -4 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} + x_5 \begin{bmatrix} -2 \\ 0 \\ -1 \\ -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + x_6 \begin{bmatrix} -1 \\ 0 \\ 3 \\ 6 \\ 0 \\ 1 \\ 0 \end{bmatrix} + x_7 \begin{bmatrix} 3 \\ 0 \\ -5 \\ -6 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

□ A system of equations can always be converted to a homogenous system (Definition HS [41]):

$$\begin{aligned} x_1 + 4x_2 - x_4 + 7x_6 - 9x_7 &= 0 \\ 2x_1 + 8x_2 - x_3 + 3x_4 + 9x_5 - 13x_6 + 7x_7 &= 0 \\ 2x_3 - 3x_4 - 4x_5 + 12x_6 - 8x_7 &= 0 \\ -x_1 - 4x_2 + 2x_3 + 4x_4 + 8x_5 - 31x_6 + 37x_7 &= 0 \end{aligned}$$

□ Some solutions to the associated homogenous system of linear equations (not necessarily exhaustive):
$x_1 = 0$, $x_2 = 0$, $x_3 = 0$, $x_4 = 0$, $x_5 = 0$, $x_6 = 0$, $x_7 = 0$

$x_1 = 3$, $x_2 = 0$, $x_3 = -5$, $x_4 = -6$, $x_5 = 0$, $x_6 = 0$, $x_7 = 1$

$x_1 = -1$, $x_2 = 0$, $x_3 = 3$, $x_4 = 6$, $x_5 = 0$, $x_6 = 1$, $x_7 = 0$

$x_1 = -2$, $x_2 = 0$, $x_3 = -1$, $x_4 = -2$, $x_5 = 1$, $x_6 = 0$, $x_7 = 0$

$x_1 = -4$, $x_2 = 1$, $x_3 = 0$, $x_4 = 0$, $x_5 = 0$, $x_6 = 0$, $x_7 = 0$

$x_1 = -4$, $x_2 = 1$, $x_3 = -3$, $x_4 = -2$, $x_5 = 1$, $x_6 = 1$, $x_7 = 1$

□ Form the augmented matrix of the homogenous linear system, and use row operations to convert to reduced row-echelon form. Notice how the entries of the final column remain zeros:

$$\begin{bmatrix} \boxed{1} & 4 & 0 & 0 & 2 & 1 & -3 & 0 \\ 0 & 0 & \boxed{1} & 0 & 1 & -3 & 5 & 0 \\ 0 & 0 & 0 & \boxed{1} & 2 & -6 & 6 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

□ Analysis of the augmented matrix for the homogenous system (Notation RREFA [31]).

Notice the slight variation for the same analysis of the original system only when the original system was consistent:

$$r = 3 \qquad D = \{1,\,3,\,4\} \qquad F = \{2,\,5,\,6,\,7,\,8\}$$

□ Coefficient matrix of original system of equations, and of associated homogenous system. This matrix will be the subject of further analysis, rather than the systems of equations.

$$\begin{bmatrix} 1 & 4 & 0 & -1 & 0 & 7 & -9 \\ 2 & 8 & -1 & 3 & 9 & -13 & 7 \\ 0 & 0 & 2 & -3 & -4 & 12 & -8 \\ -1 & -4 & 2 & 4 & 8 & -31 & 37 \end{bmatrix}$$

□ Matrix brought to reduced row-echelon form:

$$\begin{bmatrix} \boxed{1} & 4 & 0 & 0 & 2 & 1 & -3 \\ 0 & 0 & \boxed{1} & 0 & 1 & -3 & 5 \\ 0 & 0 & 0 & \boxed{1} & 2 & -6 & 6 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

□ Analysis of the row-reduced matrix (Notation RREFA [31]):

$$r = 3 \qquad D = \{1,\,3,\,4\} \qquad F = \{2,\,5,\,6,\,7\}$$

□ This is the null space of the matrix. The set of vectors used in the span construction is a linearly independent set of column vectors that spans the null space of the matrix (Theorem SSNS [81], Theorem BNS [94]). Solve the homogenous system with this matrix as the coefficient matrix and write the solutions in vector form (Theorem VFSLS [72]) to see these vectors arise.

$$\mathrm{Sp}\left(\left\{\begin{bmatrix} -4 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix},\,\begin{bmatrix} -2 \\ 0 \\ -1 \\ -2 \\ 1 \\ 0 \\ 0 \end{bmatrix},\,\begin{bmatrix} -1 \\ 0 \\ 3 \\ 6 \\ 0 \\ 1 \\ 0 \end{bmatrix},\,\begin{bmatrix} 3 \\ 0 \\ -5 \\ -6 \\ 0 \\ 0 \\ 1 \end{bmatrix}\right\}\right)$$

□ Range of the matrix, expressed as the span of a set of linearly independent vectors that are also columns of the matrix. These columns have indices that form the set D

above. (Theorem BROC [109])

$$\mathrm{Sp}\left(\left\{\begin{bmatrix}1\\2\\0\\-1\end{bmatrix},\,\begin{bmatrix}0\\-1\\2\\2\end{bmatrix},\,\begin{bmatrix}-1\\3\\-3\\4\end{bmatrix}\right\}\right)$$

□ The range of the matrix, as it arises from row operations on an augmented matrix. The matrix K is computed as described in Theorem RNS [114]. This is followed by the range described by a set of linearly independent vectors that span the null space of K, computed as according to Theorem SSNS [81]. When $r = m$, the matrix K has no rows and the range is all of $\mathbb{C}^m$.

$$K = \begin{bmatrix}1 & -\frac{12}{31} & -\frac{13}{31} & \frac{7}{31}\end{bmatrix}$$

$$\mathrm{Sp}\left(\left\{\begin{bmatrix}-\frac{7}{31}\\0\\0\\1\end{bmatrix},\,\begin{bmatrix}\frac{13}{31}\\0\\1\\0\end{bmatrix},\,\begin{bmatrix}\frac{12}{31}\\1\\0\\0\end{bmatrix}\right\}\right)$$

□ Range of the matrix, expressed as the span of a set of linearly independent vectors. These vectors are computed by row-reducing the transpose the matrix into reduced row-echelon form, tossing out the zero rows, and writing the remaining rows as column vectors. By Theorem RMSRT [124] and Theorem BRS [123], and in the style of Example RSOM.IS [123], this yields a linearly independent set of vectors that span the range.

$$\mathrm{Sp}\left(\left\{\begin{bmatrix}1\\0\\0\\-\frac{31}{7}\end{bmatrix},\,\begin{bmatrix}0\\1\\0\\\frac{12}{7}\end{bmatrix},\,\begin{bmatrix}0\\0\\1\\\frac{13}{7}\end{bmatrix}\right\}\right)$$

□ Row space of the matrix, expressed as a span of a set of linearly independent vectors, obtained from the nonzero rows of the equivalent matrix in reduced row-echelon form. (Theorem BRS [123])

$$\mathrm{Sp}\left(\left\{\begin{bmatrix}1\\4\\0\\0\\2\\1\\-3\end{bmatrix},\ \begin{bmatrix}0\\0\\1\\0\\1\\-3\\5\end{bmatrix},\ \begin{bmatrix}0\\0\\0\\1\\2\\-6\\6\end{bmatrix}\right\}\right)$$

□ Subspace dimensions associated with the matrix.

Matrix columns: 7 Rank: 3 Nullity: 4

Archetype J

Summary System with six equations, nine variables. Consistent. Null space of coefficient matrix has dimension 5.

A system of linear equations (Definition SSLE [8]):

$$\begin{aligned}
x_1 + 2x_2 - 2x_3 + 9x_4 + 3x_5 - 5x_6 - 2x_7 + x_8 + 27x_9 &= -5\\
2x_1 + 4x_2 + 3x_3 + 4x_4 - x_5 + 4x_6 + 10x_7 + 2x_8 - 23x_9 &= 18\\
x_1 + 2x_2 + x_3 + 3x_4 + x_5 + x_6 + 5x_7 + 2x_8 - 7x_9 &= 6\\
2x_1 + 4x_2 + 3x_3 + 4x_4 - 7x_5 + 2x_6 + 4x_7 - 11x_9 &= 20\\
x_1 + 2x_2 + +5x_4 + 2x_5 - 4x_6 + 3x_7 + 8x_8 + 13x_9 &= -4\\
-3x_1 - 6x_2 - x_3 - 13x_4 + 2x_5 - 5x_6 - 4x_7 + 13x_8 + 10x_9 &= -29
\end{aligned}$$

Some solutions to the system of linear equations (not necessarily exhaustive):

$x_1 = 6,\ x_2 = 0,\ x_3 = -1,\ x_4 = 0,\ x_5 = -1,\ x_6 = 2,\ x_7 = 0,\ x_8 = 0,\ x_9 = 0$

$x_1 = 4,\ x_2 = 1,\ x_3 = -1,\ x_4 = 0,\ x_5 = -1,\ x_6 = 2,\ x_7 = 0,\ x_8 = 0,\ x_9 = 0$

$x_1 = -17,\ x_2 = 7,\ x_3 = 3,\ x_4 = 2,\ x_5 = -1,\ x_6 = 14,\ x_7 = -1,\ x_8 = 3,\ x_9 = 2$

$x_1 = -11,\ x_2 = -6,\ x_3 = 1,\ x_4 = 5,\ x_5 = -4,\ x_6 = 7,\ x_7 = 3,\ x_8 = 1,\ x_9 = 1$

Augmented matrix of the linear system of equations (Definition AM [20]):

$$\begin{bmatrix}
1 & 2 & -2 & 9 & 3 & -5 & -2 & 1 & 27 & -5\\
2 & 4 & 3 & 4 & -1 & 4 & 10 & 2 & -23 & 18\\
1 & 2 & 1 & 3 & 1 & 1 & 5 & 2 & -7 & 6\\
2 & 4 & 3 & 4 & -7 & 2 & 4 & 0 & -11 & 20\\
1 & 2 & 0 & 5 & 2 & -4 & 3 & 8 & 13 & -4\\
-3 & -6 & -1 & -13 & 2 & -5 & -4 & 13 & 10 & -29
\end{bmatrix}$$

Matrix in reduced row-echelon form, row-equivalent to augmented matrix:

$$\begin{bmatrix}
\boxed{1} & 2 & 0 & 5 & 0 & 0 & 1 & -2 & 3 & 6\\
0 & 0 & \boxed{1} & -2 & 0 & 0 & 3 & 5 & -6 & -1\\
0 & 0 & 0 & 0 & \boxed{1} & 0 & 1 & 1 & -1 & -1\\
0 & 0 & 0 & 0 & 0 & \boxed{1} & 0 & -2 & -3 & 2\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}$$

☐ Analysis of the augmented matrix (Notation RREFA [31]):

$$r = 4 \qquad D = \{1,\, 3,\, 5,\, 6\} \qquad F = \{2,\, 4,\, 7,\, 8,\, 9,\, 10\}$$

☐ Vector form of the solution set to the system of equations (Theorem VFSLS [72]). Notice the relationship between the free variables and the set F above. Also, notice the pattern of 0's and 1's in the entries of the vectors corresponding to elements of the set F for the larger examples.

$$\begin{bmatrix}x_1\\x_2\\x_3\\x_4\\x_5\\x_6\\x_7\\x_8\\x_9\end{bmatrix} = \begin{bmatrix}6\\0\\-1\\0\\-1\\2\\0\\0\\0\end{bmatrix} + x_2\begin{bmatrix}-2\\1\\0\\0\\0\\0\\0\\0\\0\end{bmatrix} + x_4\begin{bmatrix}-5\\0\\2\\1\\0\\0\\0\\0\\0\end{bmatrix} + x_7\begin{bmatrix}-1\\0\\-3\\0\\-1\\0\\1\\0\\0\end{bmatrix} + x_8\begin{bmatrix}2\\0\\-5\\0\\-1\\2\\0\\1\\0\end{bmatrix} + x_9\begin{bmatrix}-3\\0\\6\\0\\1\\3\\0\\0\\1\end{bmatrix}$$

☐ A system of equations can always be converted to a homogenous system (Definition HS [41]):

$$\begin{aligned}
x_1 + 2x_2 - 2x_3 + 9x_4 + 3x_5 - 5x_6 - 2x_7 + x_8 + 27x_9 &= 0\\
2x_1 + 4x_2 + 3x_3 + 4x_4 - x_5 + 4x_6 + 10x_7 + 2x_8 - 23x_9 &= 0\\
x_1 + 2x_2 + x_3 + 3x_4 + x_5 + x_6 + 5x_7 + 2x_8 - 7x_9 &= 0\\
2x_1 + 4x_2 + 3x_3 + 4x_4 - 7x_5 + 2x_6 + 4x_7 - 11x_9 &= 0\\
x_1 + 2x_2 + +5x_4 + 2x_5 - 4x_6 + 3x_7 + 8x_8 + 13x_9 &= 0\\
-3x_1 - 6x_2 - x_3 - 13x_4 + 2x_5 - 5x_6 - 4x_7 + 13x_8 + 10x_9 &= 0
\end{aligned}$$

☐ Some solutions to the associated homogenous system of linear equations (not necessarily exhaustive):

$x_1 = 0,\, x_2 = 0,\, x_3 = 0,\, x_4 = 0,\, x_5 = 0,\, x_6 = 0,\, x_7 = 0,\, x_8 = 0,\, x_9 = 0$

$x_1 = -2,\, x_2 = 1,\, x_3 = 0,\, x_4 = 0,\, x_5 = 0,\, x_6 = 0,\, x_7 = 0,\, x_8 = 0,\, x_9 = 0$

$x_1 = -23,\, x_2 = 7,\, x_3 = 4,\, x_4 = 2,\, x_5 = 0,\, x_6 = 12,\, x_7 = -1,\, x_8 = 3,\, x_9 = 2$

$x_1 = -17,\ x_2 = -6,\ x_3 = 2,\ x_4 = 5,\ x_5 = -3,\ x_6 = 5,\ x_7 = 3,\ x_8 = 1,\ x_9 = 1$

☐ Form the augmented matrix of the homogenous linear system, and use row operations to convert to reduced row-echelon form. Notice how the entries of the final column remain zeros:

$$\begin{bmatrix} \boxed{1} & 2 & 0 & 5 & 0 & 0 & 1 & -2 & 3 & 0 \\ 0 & 0 & \boxed{1} & -2 & 0 & 0 & 3 & 5 & -6 & 0 \\ 0 & 0 & 0 & 0 & \boxed{1} & 0 & 1 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & \boxed{1} & 0 & -2 & -3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

☐ Analysis of the augmented matrix for the homogenous system (Notation RREFA [31]). Notice the slight variation for the same analysis of the original system only when the original system was consistent:

$$r = 4 \qquad D = \{1,\ 3,\ 5,\ 6\} \qquad F = \{2,\ 4,\ 7,\ 8,\ 9,\ 10\}$$

☐ Coefficient matrix of original system of equations, and of associated homogenous system. This matrix will be the subject of further analysis, rather than the systems of equations.

$$\begin{bmatrix} 1 & 2 & -2 & 9 & 3 & -5 & -2 & 1 & 27 \\ 2 & 4 & 3 & 4 & -1 & 4 & 10 & 2 & -23 \\ 1 & 2 & 1 & 3 & 1 & 1 & 5 & 2 & -7 \\ 2 & 4 & 3 & 4 & -7 & 2 & 4 & 0 & -11 \\ 1 & 2 & 0 & 5 & 2 & -4 & 3 & 8 & 13 \\ -3 & -6 & -1 & -13 & 2 & -5 & -4 & 13 & 10 \end{bmatrix}$$

☐ Matrix brought to reduced row-echelon form:

$$\begin{bmatrix} \boxed{1} & 2 & 0 & 5 & 0 & 0 & 1 & -2 & 3 \\ 0 & 0 & \boxed{1} & -2 & 0 & 0 & 3 & 5 & -6 \\ 0 & 0 & 0 & 0 & \boxed{1} & 0 & 1 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 & \boxed{1} & 0 & -2 & -3 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

☐ Analysis of the row-reduced matrix (Notation RREFA [31]):

$$r = 4 \qquad D = \{1,\ 3,\ 5,\ 6\} \qquad F = \{2,\ 4,\ 7,\ 8,\ 9\}$$

□ This is the null space of the matrix. The set of vectors used in the span construction is a linearly independent set of column vectors that spans the null space of the matrix (Theorem SSNS [81], Theorem BNS [94]). Solve the homogenous system with this matrix as the coefficient matrix and write the solutions in vector form (Theorem VFSLS [72]) to see these vectors arise.

$$\mathrm{Sp}\left(\left\{\begin{bmatrix}-2\\1\\0\\0\\0\\0\\0\\0\\0\end{bmatrix},\begin{bmatrix}-5\\0\\2\\1\\0\\0\\0\\0\\0\end{bmatrix},\begin{bmatrix}-1\\0\\-3\\0\\-1\\0\\1\\0\\0\end{bmatrix},\begin{bmatrix}2\\0\\-5\\0\\-1\\2\\0\\1\\0\end{bmatrix},\begin{bmatrix}-3\\0\\6\\0\\1\\3\\0\\0\\1\end{bmatrix}\right\}\right)$$

□ Range of the matrix, expressed as the span of a set of linearly independent vectors that are also columns of the matrix. These columns have indices that form the set D above. (Theorem BROC [109])

$$\mathrm{Sp}\left(\left\{\begin{bmatrix}1\\2\\1\\2\\1\\-3\end{bmatrix},\begin{bmatrix}-2\\3\\1\\3\\0\\-1\end{bmatrix},\begin{bmatrix}3\\-1\\1\\-7\\2\\2\end{bmatrix},\begin{bmatrix}-5\\4\\1\\2\\-4\\-5\end{bmatrix}\right\}\right)$$

□ The range of the matrix, as it arises from row operations on an augmented matrix. The matrix K is computed as described in Theorem RNS [114]. This is followed by the range described by a set of linearly independent vectors that span the null space of K, computed as according to Theorem SSNS [81]. When $r = m$, the matrix K has no rows and the range is all of $\mathbb{C}^m$.

$$K = \begin{bmatrix}1 & 0 & \frac{186}{131} & \frac{51}{131} & -\frac{188}{131} & \frac{77}{131}\\ 0 & 1 & -\frac{272}{131} & -\frac{45}{131} & \frac{58}{131} & -\frac{14}{131}\end{bmatrix}$$

$$\text{Sp}\left(\left\{\begin{bmatrix}-\frac{77}{131}\\ \frac{14}{131}\\ 0\\ 0\\ 0\\ 1\end{bmatrix},\begin{bmatrix}\frac{188}{131}\\ -\frac{58}{131}\\ 0\\ 0\\ 1\\ 0\end{bmatrix},\begin{bmatrix}-\frac{51}{131}\\ \frac{45}{131}\\ 0\\ 1\\ 0\\ 0\end{bmatrix},\begin{bmatrix}-\frac{186}{131}\\ \frac{272}{131}\\ 1\\ 0\\ 0\\ 0\end{bmatrix}\right\}\right)$$

□ Range of the matrix, expressed as the span of a set of linearly independent vectors. These vectors are computed by row-reducing the transpose the matrix into reduced row-echelon form, tossing out the zero rows, and writing the remaining rows as column vectors. By Theorem RMSRT [124] and Theorem BRS [123], and in the style of Example RSOM.IS [123], this yields a linearly independent set of vectors that span the range.

$$\text{Sp}\left(\left\{\begin{bmatrix}1\\ 0\\ 0\\ 0\\ -1\\ -\frac{29}{7}\end{bmatrix},\begin{bmatrix}0\\ 1\\ 0\\ 0\\ -\frac{11}{2}\\ -\frac{94}{7}\end{bmatrix},\begin{bmatrix}0\\ 0\\ 1\\ 0\\ 10\\ 22\end{bmatrix},\begin{bmatrix}0\\ 0\\ 0\\ 1\\ \frac{3}{2}\\ 3\end{bmatrix}\right\}\right)$$

□ Row space of the matrix, expressed as a span of a set of linearly independent vectors, obtained from the nonzero rows of the equivalent matrix in reduced row-echelon form. (Theorem BRS [123])

$$\text{Sp}\left(\left\{\begin{bmatrix}1\\ 2\\ 0\\ 5\\ 0\\ 0\\ 1\\ -2\\ 3\end{bmatrix},\begin{bmatrix}0\\ 0\\ 1\\ -2\\ 0\\ 0\\ 3\\ 5\\ -6\end{bmatrix},\begin{bmatrix}0\\ 0\\ 0\\ 0\\ 1\\ 0\\ 1\\ 1\\ -1\end{bmatrix},\begin{bmatrix}0\\ 0\\ 0\\ 0\\ 0\\ 1\\ 0\\ -2\\ -3\end{bmatrix}\right\}\right)$$

□ Subspace dimensions associated with the matrix.

Matrix columns: 9 Rank: 4 Nullity: 5

Archetype K

Summary Square matrix of size 5. Nonsingular. 3 distinct eigenvalues, 2 of multiplicity 2.

☐ A matrix:

$$\begin{bmatrix} 10 & 18 & 24 & 24 & -12 \\ 12 & -2 & -6 & 0 & -18 \\ -30 & -21 & -23 & -30 & 39 \\ 27 & 30 & 36 & 37 & -30 \\ 18 & 24 & 30 & 30 & -20 \end{bmatrix}$$

☐ Matrix brought to reduced row-echelon form:

$$\begin{bmatrix} \boxed{1} & 0 & 0 & 0 & 0 \\ 0 & \boxed{1} & 0 & 0 & 0 \\ 0 & 0 & \boxed{1} & 0 & 0 \\ 0 & 0 & 0 & \boxed{1} & 0 \\ 0 & 0 & 0 & 0 & \boxed{1} \end{bmatrix}$$

☐ Analysis of the row-reduced matrix (Notation RREFA [31]):

$$r = 5 \qquad D = \{1,\,2,\,3,\,4,\,5\} \qquad F = \{\ \}$$

☐ Matrix (coefficient matrix) is nonsingular or singular? (Theorem NSRRI [50]) at the same time, examine the size of the set F above.Notice that this property does not apply to matrices that are not square.

Nonsingular.

☐ This is the null space of the matrix. The set of vectors used in the span construction is a linearly independent set of column vectors that spans the null space of the matrix (Theorem SSNS [81], Theorem BNS [94]). Solve the homogenous system with this matrix as the coefficient matrix and write the solutions in vector form (Theorem VFSLS [72]) to see these vectors arise.

$\mathrm{Sp}\left(\{\ \}\right)$

☐ Range of the matrix, expressed as the span of a set of linearly independent vectors that are also columns of the matrix. These columns have indices that form the set D

above. (Theorem BROC [109])

$$\mathrm{Sp}\left(\left\{\begin{bmatrix}10\\12\\-30\\27\\18\end{bmatrix},\begin{bmatrix}18\\-2\\-21\\30\\24\end{bmatrix},\begin{bmatrix}24\\-6\\-23\\36\\30\end{bmatrix},\begin{bmatrix}24\\0\\-30\\37\\30\end{bmatrix},\begin{bmatrix}-12\\-18\\39\\-30\\-20\end{bmatrix}\right\}\right)$$

□ The range of the matrix, as it arises from row operations on an augmented matrix. The matrix K is computed as described in Theorem RNS [114]. This is followed by the range described by a set of linearly independent vectors that span the null space of K, computed as according to Theorem SSNS [81]. When $r = m$, the matrix K has no rows and the range is all of $\mathbb{C}^m$.

$$K = []$$

$$\mathrm{Sp}\left(\left\{\begin{bmatrix}1\\0\\0\\0\\0\end{bmatrix},\begin{bmatrix}0\\1\\0\\0\\0\end{bmatrix},\begin{bmatrix}0\\0\\1\\0\\0\end{bmatrix},\begin{bmatrix}0\\0\\0\\1\\0\end{bmatrix},\begin{bmatrix}0\\0\\0\\0\\1\end{bmatrix}\right\}\right)$$

□ Range of the matrix, expressed as the span of a set of linearly independent vectors. These vectors are computed by row-reducing the transpose the matrix into reduced row-echelon form, tossing out the zero rows, and writing the remaining rows as column vectors. By Theorem RMSRT [124] and Theorem BRS [123], and in the style of Example RSOM.IS [123], this yields a linearly independent set of vectors that span the range.

$$\mathrm{Sp}\left(\left\{\begin{bmatrix}1\\0\\0\\0\\0\end{bmatrix},\begin{bmatrix}0\\1\\0\\0\\0\end{bmatrix},\begin{bmatrix}0\\0\\1\\0\\0\end{bmatrix},\begin{bmatrix}0\\0\\0\\1\\0\end{bmatrix},\begin{bmatrix}0\\0\\0\\0\\1\end{bmatrix}\right\}\right)$$

□ Row space of the matrix, expressed as a span of a set of linearly independent vectors, obtained from the nonzero rows of the equivalent matrix in reduced row-echelon form. (Theorem BRS [123])

$$\mathrm{Sp}\left(\left\{\begin{bmatrix}1&0&0&0&0\end{bmatrix},\begin{bmatrix}0&1&0&0&0\end{bmatrix},\begin{bmatrix}0&0&1&0&0\end{bmatrix},\begin{bmatrix}0&0&0&1&0\end{bmatrix},\begin{bmatrix}0&0&0&0&1\end{bmatrix}\right\}\right)$$

□ Inverse matrix, if it exists. The inverse is not defined for matrices that are not square, and if the matrix is square, then the matrix must be nonsingular. (Definition MI [141], Theorem NSI [151])

$$\begin{bmatrix} 1 & -\left(\frac{9}{4}\right) & -\left(\frac{3}{2}\right) & 3 & -6 \\ \frac{21}{2} & \frac{43}{4} & \frac{21}{2} & 9 & -9 \\ -15 & -\left(\frac{21}{2}\right) & -11 & -15 & \frac{39}{2} \\ 9 & \frac{15}{4} & \frac{9}{2} & 10 & -15 \\ \frac{9}{2} & \frac{3}{4} & \frac{3}{2} & 6 & -\left(\frac{19}{2}\right) \end{bmatrix}$$

☐ Subspace dimensions associated with the matrix.

Matrix columns: 5 Rank: 5 Nullity: 0

☐ Determinant of the matrix, which is only defined for square matrices. (Zero/nonzero? Singular/nonsingular? Product of all eigenvalues?)

Determinant = 16

Archetype L

Summary Square matrix of size 5. Singular, nullity 2. 2 distinct eigenvalues, each of "high" multiplicity.

A matrix:

$$\begin{bmatrix} -2 & -1 & -2 & -4 & 4 \\ -6 & -5 & -4 & -4 & 6 \\ 10 & 7 & 7 & 10 & -13 \\ -7 & -5 & -6 & -9 & 10 \\ -4 & -3 & -4 & -6 & 6 \end{bmatrix}$$

Matrix brought to reduced row-echelon form:

$$\begin{bmatrix} \boxed{1} & 0 & 0 & 1 & -2 \\ 0 & \boxed{1} & 0 & -2 & 2 \\ 0 & 0 & \boxed{1} & 2 & -1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Analysis of the row-reduced matrix (Notation RREFA [31]):

$$r = 5 \qquad D = \{1,\, 2,\, 3\} \qquad F = \{4,\, 5\}$$

Matrix (coefficient matrix) is nonsingular or singular? (Theorem NSRRI [50]) at the same time, examine the size of the set F above.Notice that this property does not apply to matrices that are not square.

Singular.

This is the null space of the matrix. The set of vectors used in the span construction is a linearly independent set of column vectors that spans the null space of the matrix (Theorem SSNS [81], Theorem BNS [94]). Solve the homogenous system with this matrix as the coefficient matrix and write the solutions in vector form (Theorem VFSLS [72]) to see these vectors arise.

$$\mathrm{Sp}\left(\left\{\begin{bmatrix}-1\\2\\-2\\1\\0\end{bmatrix},\ \begin{bmatrix}2\\-2\\1\\0\\1\end{bmatrix}\right\}\right)$$

□ Range of the matrix, expressed as the span of a set of linearly independent vectors that are also columns of the matrix. These columns have indices that form the set D above. (Theorem BROC [109])

$$\mathrm{Sp}\left(\left\{\begin{bmatrix}- -2\\-6\\10\\-7\\-4\end{bmatrix},\ \begin{bmatrix}-1\\-5\\7\\-5\\-3\end{bmatrix},\ \begin{bmatrix}-2\\-4\\7\\-6\\-4\end{bmatrix}\right\}\right)$$

□ The range of the matrix, as it arises from row operations on an augmented matrix. The matrix K is computed as described in Theorem RNS [114]. This is followed by the range described by a set of linearly independent vectors that span the null space of K, computed as according to Theorem SSNS [81]. When $r = m$, the matrix K has no rows and the range is all of $\mathbb{C}^m$.

$$K = \begin{bmatrix}1 & 0 & -2 & -6 & 5\\0 & 1 & 4 & 10 & -9\end{bmatrix}$$

$$\mathrm{Sp}\left(\left\{\begin{bmatrix}-5\\9\\0\\0\\1\end{bmatrix},\ \begin{bmatrix}6\\-10\\0\\1\\0\end{bmatrix},\ \begin{bmatrix}2\\-4\\1\\0\\0\end{bmatrix}\right\}\right)$$

□ Range of the matrix, expressed as the span of a set of linearly independent vectors. These vectors are computed by row-reducing the transpose the matrix into reduced row-echelon form, tossing out the zero rows, and writing the remaining rows as column vectors. By Theorem RMSRT [124] and Theorem BRS [123], and in the style of Example RSOM.IS [123], this yields a linearly independent set of vectors that span the range.

$$\mathrm{Sp}\left(\left\{\begin{bmatrix}1\\0\\0\\\frac{9}{4}\\\frac{5}{2}\end{bmatrix},\ \begin{bmatrix}0\\1\\0\\\frac{5}{4}\\\frac{3}{2}\end{bmatrix},\ \begin{bmatrix}0\\0\\1\\\frac{1}{2}\\1\end{bmatrix}\right\}\right)$$

☐ Row space of the matrix, expressed as a span of a set of linearly independent vectors, obtained from the nonzero rows of the equivalent matrix in reduced row-echelon form. (Theorem BRS [123])

$$\mathrm{Sp}\left(\left\{\begin{bmatrix}1\\0\\0\\1\\-2\end{bmatrix},\ \begin{bmatrix}0\\1\\0\\-2\\2\end{bmatrix},\ \begin{bmatrix}0\\0\\1\\2\\-1\end{bmatrix}\right\}\right)$$

☐ Inverse matrix, if it exists. The inverse is not defined for matrices that are not square, and if the matrix is square, then the matrix must be nonsingular. (Definition MI [141], Theorem NSI [151])

☐ Subspace dimensions associated with the matrix.

Matrix columns: 5 Rank: 3 Nullity: 2

☐ Determinant of the matrix, which is only defined for square matrices. (Zero/nonzero? Singular/nonsingular? Product of all eigenvalues?)

Determinant $= 0$

Index