Mathematical Methods of Economic Analysis — ECO 6405/7405

Professor Boyd   **Aug. 23, 2022**

Basic Course Information

Syllabus: The full syllabus is available indirectly via Canvas. There's a link to it at the end of of the Canvas syllabus page. Alternatively, try: `http://faculty.fiu.edu/~boydj/MathEcon.html`

Office Hours: If you have questions, you may ask immediately after class, or come to my office. Regular office hours are 12:45-1:45pm and 3:30-4:45pm on Tuesdays and Thursdays. I will be happy to make an appointment for another time if that is more convenient. My office is DM-311A, my office phone number is 305-348-3287, and my email is **John.Boyd@fiu.edu** or **boydj@fiu.edu**. Sending it directly is preferred. Email sent via Canvas is subject to delays.

Course Description

This course focuses on mathematical methods used in modern economics. These include:

1. linear methods often used in mathematical modelling,

2. the portions of mathematical analysis relevant for studying optimization problems,

3. the construction and solution of optimization problems, and

4. the solution of difference and differential equations characteristic of modern intertemporal models, as used in both microeconomics and macroeconomics.

The first part of the course develops some basic mathematical tools of analysis which we will use to solve optimization problems. This covers roughly parts II and III of the text, and may include excerpts from parts VI and VII. The second part (part IV of the textbook) covers classical, calculus-based methods of optimization including Lagrange multipliers and the Kuhn-Tucker theorem. The methods of Lagrange and Kuhn-Tucker have been invaluable in solving many of the problems you will typically encounter in economics (consumer and producer choice, social welfare max, etc.). We then cover the solution of difference and differential equations, and their stability properties (part V). If time permits, we will look at dynamic optimization and the Maximum Principle.

Course Objectives

By the end of the course, at a minimum, you should be able to:

- Determine whether a linear system has a solution, and if so, how many.
- Solve linear systems using both determinants and the Gauss-Jordan method.
- Find eigenvalues and eigenvectors.
- Use the functional calculus.
- Determine whether an optimization problem has a solution.
- Characterize the solutions of optimization problems via the first order conditions.
- Solve unconstrained optimization problems using first and second order conditions.
- Solve constrained optimization problems using the Kuhn-Tucker Theorem.
- Exploit special features such as homogeneity or convexity when solving optimization problems.
- Solve linear difference and differential systems.
- Characterize the long-run behavior of difference and differential systems using eigenvalues.

Textbook

Carl Simon and Lawrence Blume, *Mathematics for Economists*, W. W. Norton, New York, 1994.

Simon and Blume's book is the main text. I plan to cover Parts II-IV and VII of Simon and Blume, with some excerpts from Part VI. Time permitting, we will then turn our attention to Part V and dynamic models.

Some additional books you may find helpful to consult are listed in the online syllabus:

`http://faculty.fiu.edu/~boydj/MathEcon.html`

Constrained Optimization Survival Guide. You may find this handout on basic optimization helpful, particularly in your micro course:

`http://faculty.fiu.edu/~boydj/mathecon/survival.pdf`

Exams and Homework

Grades will be based on two midterm exams (worth 25% each), a final exam (40%), and homework assignments (10%). In addition to being announced in class, homework assignments will be posted on the online syllabus.

Homework will be submitted in person or by emailing it to me. If you email it, it may be easiest to write it out and then photograph it with your phone. If so, please combine the pages into a single pdf. I will not be happy if I see 10 separate files for one assignment.

Homework is graded as follows: ✓+ (3 pts) means that it is mostly correct, no major errors. ✓ (2 points) indicates you've missed at least one problem. ✓− (1 point) means that at least two problems or equivalent are mostly incorrect. On difficult assignments three misses may be required for a ✓−. A zero is also possible, and usually means the homework wasn't turned in.

Exams

There will be two midterm exams, each worth 25% of your grade, and a final, worth 40% of your grade.

- The first midterm is scheduled for **Thursday, September 22**.

- The second midterm is scheduled for **Thursday, October 27**.

- The final will be at the officially scheduled time, **5pm on Tuesday, December 6, 2022**, in our regular classroom, DM-164.

Previous exams are available from the online syllabus. The material covered varies from year to year and some of the questions on previous exams may not be relevant for the material we cover this year. A few of the answers contain minor errors.

Tentative Course Outline

The online syllabus includes a tentative course outline. It is subject to change, especially if there are hurricanes. It is probably a bit over ambitious, but we'll see. There have been years when we've made it all the way to the end. Once the semester is underway, the lecture notes will show you what was actually covered.

Next Up: **Chapter Six**

6. Intro to Linear Equations and Systems

6.1 Sample Linear Equations

We start with some sample linear equations.

$$x = 2 \qquad \text{a point in } \mathbb{R}$$
$$2x = 3y + 7 \qquad \text{a line in } \mathbb{R}^2$$
$$10 = x + 3y + 4z \qquad \text{a plane in } \mathbb{R}^3$$
$$0 = 0 \qquad \text{anything is a solution!}$$
$$x = x + 2 \qquad \text{there are no solutions!!}$$

Here $\mathbb{R}$ denotes the set of real numbers and $\mathbb{R}^n$ is the set of all n-tuples of real numbers, $(x_1, \dots, x_n)$, with each $x_i \in \mathbb{R}$. An n-*tuple* is an ordered list of n objects (numbers in this case). Thus $(2, 2, 1, 2)$ and $(2, 2, 2, 1)$ are different n-tuples.

6.2 Equations that are Not Linear

Here are some equations that are not linear equations.

$$y = e^x$$
$$x_3 = \cos x_1 + 3 \sin^2 x_2$$
$$x_2 = 2x_2^2 + 3x_2 + 12$$
$$z = 7xy + 5y + 2$$
$$z = 3 + \sqrt{x + xy + \sqrt{y}}$$
$$y = \sqrt{x^2}$$

For the last equation, keep in mind that it is the positive square root. That equation can also be written $y = |x|$.

6.3 Linear Equations

The key thing about linear equations is that they are linear in each variable. You can't take powers or other functions of variables, nor multiply them together. The only thing you can multiply a variable by is a constant. More formally, a *linear equation* in n variables $x_1, x_2, \ldots, x_n$ is an equation that can be written in the form

$$a_1 x_1 + a_2 x_2 + \cdots + a_n x_n = b \tag{6.3.1}$$

for some real numbers a_i, $i = 1, \ldots, n$ and b. In linear equations you can multiply variables by real numbers (*scalars*) and add terms together. That's it. No squares, cosines, or any other functions are allows.

We can use summation notation to write equation (6.3.1) in the more compact form

$$\sum_{i=1}^{n} a_i x_i = b$$

You'll notice that there is only one variable in each term, and that it always appears as x_i itself, never as a non-trivial function of x_i.

6.4 Solutions to Linear Equations

Suppose we have a linear equation in n variables:

$$a_1 x_1 + a_2 x_2 + \cdots + a_n x_n = b \qquad (6.3.1)$$

Consider the solution set,

$$\left\{ (x_1, \ldots, x_n) \text{ that solve equation (6.3.1)} \right\}$$

Before going on, let's unpack the notation a bit. The term $(x_1, \ldots, x_n)$ is an n-*tuple of numbers*, $x_1, x_2, \ldots, x_n$. By n-tuple, we mean that there are n numbers and that they are **in a particular order**, with the first one first and last one last. For small n, these have semi-latin names: single, double, triple, quadruple, quintuple, sextuple (or hextuple), septuple, octuple, and we usually don't use their names much beyond that. We often use a single subscript, $i = 1, \ldots, n$ to denote distinct entries in the n-tuple.

6.5 Dummy Variables

The variable i that we use to denote the entries in the n-tuple is known as a *dummy variable*. The point is that it doesn't matter what we call it. I could call it i, j, or even v and it would make no difference for anything. It's a bad idea to call it n, because n is the number of variables and actually means something. It's also bad form to give the name of any non-dummy variables you're using.

If you are doing calculation involving 2 or more dummy variables, it is sometimes a problem if you use the same letter for both variables. Sometimes its safe (and you may see me doing such a thing when it is), but sometimes it can lead to confusion. If you are unsure if its safe, do yourself a favor and use two different letters.

6.6 What are the Solutions to Linear Equations?

What does the solution set to equation (6.3.1) look like?

$$a_1 x_1 + a_2 x_2 + \cdots + a_n x_n = b \qquad (6.3.1)$$

If we have 2 variables, a linear equation is the equation of a straight line in two-dimensional space, $\mathbb{R}^2$. With 3 variables, equation (6.3.1) describes a plane in 3-space, $\mathbb{R}^3$. In general, with n variables, the solutions to equation (6.3.1) form a *hyperplane*, an $(n-1)$-dimensional vector subspace of $\mathbb{R}^n$. (I.e., one dimension smaller than the ambient space that contains it.) For all $n \geq 2$, if you have two distinct points in the solution set, the line they generate will also be in the solution set.

When n is larger than 2 or 3, don't ask me to draw them!

6.7 Linear Systems

A *linear system* in n variables $x_1, \ldots, x_n$ is a collection of linear equations. Let a_{ij}, b_i be real numbers with $i = 1, \ldots, m$ and $j = 1, \ldots, n$. We can write the system as

$$
\begin{aligned}
a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\
a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\
a_{31}x_1 + a_{32}x_2 + \cdots + a_{3n}x_n &= b_3 \\
&\vdots \\
a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m
\end{aligned}
\tag{6.7.2}
$$

   Notice that we have two subscripts on each of the a's. We can write a_{ij} where i denotes the equation, $i = 1, \ldots, m$ and j denotes the variable, $j = 1, \ldots, n$. There don't have to be the same number of variables (n) and equations (m).

6.8 Homogeneous Linear Systems

One consequence of linearity is that if $(x_1, \dots, x_n)$ and $(x'_1, \dots, x'_n)$ both solve the linear system (6.7.2), their difference solves the linear system with the same coefficients but with each $b_i = 0$. This is referred to as the associated *homogeneous system*.

NB: That's homog**ene**ous with two e's, not homogenous with one.

$$a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0$$
$$a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = 0$$
$$a_{31}x_1 + a_{32}x_2 + \cdots + a_{3n}x_n = 0$$
$$\vdots \qquad\qquad\qquad \vdots$$
$$a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = 0$$

If $(x_1, \dots, x_n)$ and $(x'_1, \dots, x'_n)$ solve the homogeneous system, so does any linear combination of them.

Also, the zero n-tuple $(0, \dots, 0)$ always solves any **homogeneous system** in n variables.

6.9 Consequences of Linearity

The system is **linear**. If we add two solution n-tuples together, coordinate by coordinate, we get another solution. Also, if we multiply every coordinate by the same real number, we also get a solution. In other words, the two linear operations, addition of n-tuples and multiplication of an n-tuple by a real number (a *scalar*) both give us another n-tuple.

By the linearity of the system, if we have one solution $(x_1^*, \ldots, x_n^*)$ to equation (6.7.2), any other solution can be written as the sum of it and a solution to the associated homogeneous system. This property is shared with other types of linear systems, such as systems of linear differential equations.

6.10 What are the Solutions to Linear Systems?

As with linear equations, we ask what the solution set of a linear system is. What does $\{(x_1, \dots, x_n) : (x_1, \dots, x_n)$ solves equation (6.7.2)$\}$ look like?

To help build intuition, consider the case $n = 2$. If they are non-trivial, the linear equations that make up the system (6.7.2) each describe lines in $\mathbb{R}^2$. If it satisfies each equation, a point (x_1, x_2) must be on every line, meaning that it is in the intersection of a bunch of lines.

The intersection might be empty (no solution), it might be a single point (a unique solution), or it might contain at least two points. If it contains two points, these points determine a straight line, and that line must be the straight line described by each of the linear equations in system (6.7.2). That whole line must be the intersection. There are infinitely many solutions!

There is one more case to consider. What if the equations are trivial? In that case, they impose no restrictions, as in the system

$$0 = 0$$
$$42 = 42$$
$$137 = 137.$$

If this is a system in n variables, anything in $\mathbb{R}^n$ solves it.[1]

---

[1] The numbers come from Brahmagupta (earliest known use of the number zero), Douglas Adams (the ultimate answer), and Arnold Sommerfield (fine structure constant).

6.11 Illustrating the Solutions to Linear Systems

It follows that in $\mathbb{R}^2$ the intersection that is the solution set is either empty, a single point, a straight line, or the entirety of $\mathbb{R}^2$.

In $\mathbb{R}^3$, the solution set is the intersection of planes. It can be empty, a point, a straight line, a plane, or the whole space.

The possibilities are similar in higher dimensions. They are always the intersections of hyperplanes, or else the whole space.
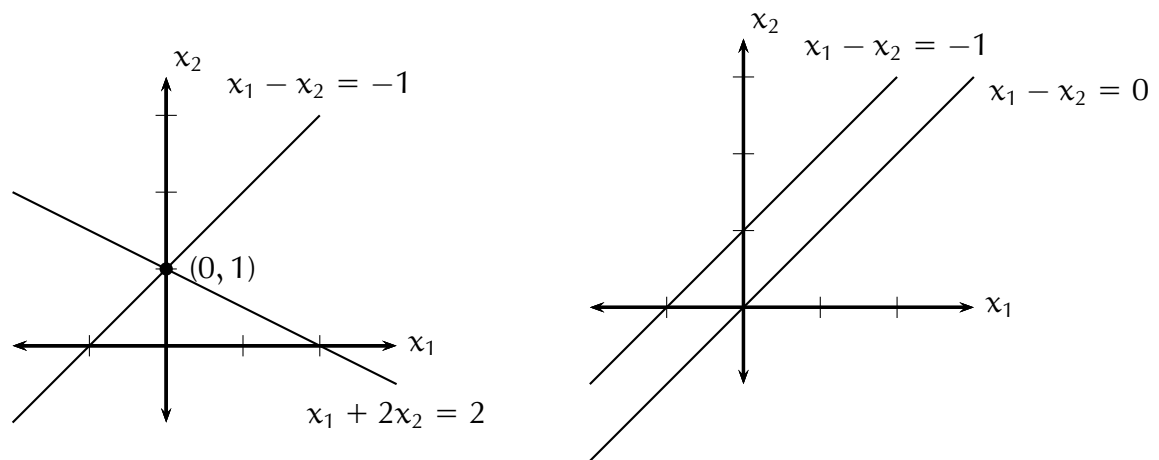
Figure 6.11.1: The left diagram has one solution. It is the unique intersection of the two linelines at $(0, 1)$.

The right diagram has parallel lines and no solutions.

6.12 Example: Taxes and Charitable Deductions

Let's examine a simple linear system.

Suppose a company has before-tax profits of $100,000. It will contribute 10% of its after-tax profits to the Red Cross. It pays a state tax equal to 5% of its post-contribution profit and a federal tax of 40% of its profit post contribution and state tax. How much does the company pay in federal taxes.

We can set this up as a linear system. Let C be the charitable contribution, S be the state tax, and F the federal tax. After-tax profits are $100,000 - (S + F)$, so $C = 0.1\big(100,000 - (S + F)\big)$. We rewrite this as

$$C + 0.1S + 0.1F = 10,000$$

The state tax is 5% of the profit net of the donation. Then $S = 0.05(100,000 - C)$ or

$$0.05C + S = 5,000.$$

Federal taxes are 40% of the remaining profit, $F = 0.40\big(100,000 - (C + S)\big)$. In other words $0.4C + 0.4S + F = 40,000$.

We put these three equations together to form our linear system:

$$C + 0.1S + 0.1F = 10,000$$
$$0.05C + S = 5,000$$
$$0.4C + 0.4S + F = 40,000.$$

By solving this system of three equations in three unknowns, we can calculate the charitable contribution (C) as well as the state (S) and federal (F) taxes.

6.13 Example: Input-Output Model I

Input-output models provide another example of a linear system. Similar models for the entire US economy were studied by Wassily Leontief in the 1940's and 1950's.[2]

The basic model includes a primary good, which we will refer to as labor and denote by index 0, and n produced goods. The produced goods can either be used as inputs to production or consumed and are labeled $1, \ldots, n$.

Production involves both *fixed proportions* and *constant returns to scale*. Fixed proportions means that the inputs are always combined in the same ratio, as hydrogen and oxygen are always combined in a 2:1 ratio to make water, H_2O (or a 1:1 ratio to make hydrogen peroxide H_2O_2). Constant returns to scale means that scaling all inputs by a factor $\alpha > 0$ scales output by that same factor α.

For example, suppose that the production of one 8 ounce glass of chocolate milk requires one 8 ounce glass of milk, two teaspoons of chocolate mix, and two minutes of labor, to make it well-stirred.

Then two glasses of chocolate milk can be produced using two glasses of ordinary milk, four teaspoons of chocolate mix, and four minutes of labor (doubling all inputs doubles the output).

Let a_{ij} be the amount of input i required to produce one unit of good j. These are referred to as the *input-output coefficients*. In our chocolate milk example, we let good 1 be ordinary milk, good 2 be chocolate mix, and good 3 be chocolate milk. The input coefficients are $a_{03} = 1/30$, $a_{13} = 1$, $a_{23} = 2$, and $a_{33} = 0$, where labor is measured in hours, both types of milk in 8 ounce glasses, and chocolate mix in teaspoons.

---

[2] Wassily Leontief (1905–1999) was a Soviet-American economist who received the Nobel Prize in 1973. Starting in 1949, he used early computers to build a 500-sector model of the US economy. He developed his input-output model to study the US economy.

6.14 Example: Input-Output Model II

Suppose we have input-output coefficients a_{ij} and want to produce x_j units of good j for each $j = 1, \ldots, n$. This requires input of

$$\sum_{j=1}^{n} a_{ij} x_j$$

units of good i for $i = 0, \ldots, n$. We can think of this expression as the demand for good i by producers. Note that labor is included ($i = 0$).

Consumers consume goods $i = 1, \ldots, n$. The quantity demanded of good i is c_i for $i = 1, \ldots, n$. The consumers do not demand labor. They supply it.

Finally, suppose the economy has L units of labor which is supplied by consumers. If we want to produce enough so that consumers can consume $(c_1, \ldots, c_n)$, we must find non-negative values of the x_i so that

$$\text{Supply} = \text{Producer Demand} + \text{Consumer Demand}$$
$$L = a_{01}x_1 + a_{02}x_2 + \cdots + a_{0n}x_n$$
$$x_1 = a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n + c_1$$
$$x_2 = a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n + c_2$$
$$\vdots \qquad\qquad\qquad \vdots$$
$$x_n = a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n + c_n.$$

This linear system is the *Leontief system*. Basically, it says that supply is equal to the sum of producer and consumer demand in every market. This means it is a primitive general equilibrium model, where all markets must simultaneously clear. (Primitive, because it lacks prices.)

6.15 Example: Markovian Employment Model I

This model has two possible states, employed and unemployed. We don't distinguish how long someone has been either employed or unemployed.

If someone is unemployed, they have a probability p, $0 \leq p \leq 1$ of finding a job this week. There is a $(1 - p)$ chance they don't find a job and remain unemployed.

If someone is employed, they have a probability q, $0 \leq q \leq 1$ of losing that job this week and becoming unemployed. There is a probability $(1 - q)$ of remaining employed.

Models where the probability of moving from one state to another depends only on the current state are called *Markovian*. The probabilities of moving between the various states are called *transition probabilities*.[3]

Now suppose x people are currently employed and y are unemployed. Next week, $qx + (1-p)y$ will be unemployed next week, and $(1-q)x + py$ will be employed next week. These quantities always sum to $x + y$, the total labor force.

[3] Andrei Markov (1856–1922) was a Russian mathematician. He is best known for his work in probability and statistics, particularly for Markov chains and processes, which are often used in economics. He also made significant contributions to other areas, such as the Riesz-Markov Representation Theorem.

6.16 Example: Markovian Employment Model II

We now consider a sequence of weeks, indexed by time t. Then our system becomes

$$x_{t+1} = (1 - q)x_t + py_t$$
$$y_{t+1} = qx_t + (1 - p)y_t.$$

This is a case where we have a linear system of *difference equations*. These equations can be rewritten so the differences $x_{t+1} - x_t$ and $y_{t+1} - y_t$ are functions of (x_t, y_t), as in

$$x_{t+1} - x_t = -qx_t + py_t$$
$$y_{t+1} - y_t = +qx_t - py_t.$$

Once again, $x_t + y_t$ remains constant. We say that (x, y) is a *stationary distribution* if $(x_t, y_t) = (x, y)$ implies that $(x_{t+1}, y_{t+1}) = (x, y)$. In other words, once we reach a stationary distribution, the system no longer evolves over time. It remains at (x, y).

6.17 Finding the Stationary Distribution

Let's find the stationary distribution under the assumption that $x + y = N$. We will also require that $p > 0$. The stationary distribution must solve the following equations.

$$x = (1 - q)x + py \tag{6.17.3}$$

$$y = qx + (1 - p)y \tag{6.17.4}$$

$$x + y = N. \tag{6.17.5}$$

We will solve for (x, y). Before proceeding, note that equations (6.17.3) and (6.17.4) are different forms of the same equation. Both simplify to $qx = py$. This means that it is enough to satisfy (6.17.3) and (6.17.5) as (6.17.4) automatically follows.

Now we rewrite equation (6.17.3) so that $qx = py$. Since $p > 0$,

$$y = \frac{q}{p}x.$$

We plug this into (6.17.5), obtaining $x + qx/p = N$. Then

$$x\left(\frac{p + q}{p}\right) = N$$

so

$$x = \frac{p}{p + q}N \quad \text{and} \quad y = \frac{q}{p + q}N.$$

We have found the stationary distribution by solving the system (6.17.3)–(6.17.5).

6.18 Example: IS-LM Model I

A simple IS-LM model without trade provides another example of a linear system. Let Y denote GDP, C consumption, I investment by firms, and G government spending. With all of GDP accounted for,[4]

$$Y = C + I + G.$$

We use a simplified Keynesian consumption function, where $C = bY$ with $0 < b < 1$. Here b is the *marginal propensity to consume* and $s = (1 - b)$ is the *marginal propensity to save*. Firms' investment depends on the interest rate r. We expect it to be decreasing in r because higher interest rates decrease the present value of future income. We write

$$I = I^0 - ar.$$

The parameter a is called the *marginal efficiency of capital*. We can combine these equations to obtain the IS (investment-saving) curve. Here

$$Y = bY + (I^0 - ar) + G,$$

so

$$sY + ar = I^0 + G \tag{IS}$$

---

[4] Of course, Lord John Maynard Keynes (1883–1946) is well-known to all economists. I'll just mention *The Economic Consequences of the Peace* (1919), *A Treatise on Probability* (1921), the *Treatise on Money* (1930), and the *The General Theory of Employment, Interest and Money* (1936). Many of his ideas on international economic institutions were incorporated into the International Monetary Fund and International Bank for Reconstruction and Development (now part of the World Bank). These were created by the 1944 Bretton Woods conference in New Hampshire, where Keynes played a leading role.

It's somewhat less known that the IS-LM model was created by Sir John R. Hicks (1904–1989) soon after Keynes published the *General Theory*. See " 'Mr. Keynes and the Classics': A Suggested Interpretation' ", *Econometrica* **5**, 147–159. Hicks won the Economics Nobel in 1972 and has made numerous contributions to both micro and macro. The first part of his 1939 book *Value and Capital* based consumer theory on ordinal utility and stressed the importance of the income and substitution effects. The second part focused on dynamic economics. With R.G.D. Allen, he developed various measures of the elasticity of substitution.

6.19 Example: IS-LM Model II

The other half of the model is the LM (liquidity-money) equation, which characterizes money market clearing. Money demand comes in two parts. One is from use in transactions, and is proportional to GDP. The other part is speculative demand. Investors must decide whether to hold bonds or cash. The interest rate negatively affects the speculative demand. Thus money demand is $M_d = mY + (M^0 - hr)$ where m, h, and M^0 are all positive.

Money supply is fixed at M_s in this model. We set supply equal to demand

$$M_s = mY + M^0 - hr$$

and rearrange to obtain the LM equation

$$mY - hr = M_s - M^0 \tag{LM}$$

The IS-LM model is formed by combining the IS and LM equations.

$$sY + ar = I^0 + G$$
$$mY - hr = M_s - M^0.$$

We have two equations in the two unknowns Y and r. Here $I^0 + G$ and $M_s - M^0$ are the constant terms and s, a, m, and $-h$ the coefficients.

7. Solving Linear Systems

There are three key questions concerning the solution of linear systems:

1. **Existence:** Is there a solution?
2. **Uniqueness:** Is there one solution or many?
3. **Characterization:** How can we recognize the solutions?

7.1 Methods of Solution of Linear Systems

Let's consider a linear system with m equations in n variables:

$$\sum_{j=1}^{n} a_{ij}x_j = b_i, \quad i = 1,\ldots,m.$$

Three methods have proven useful to solve linear systems.

1. **Substitution**. As used in the Markovian employment model.
2. **Elimination of Variables**. The Gauss and Gauss-Jordan methods
3. **Cramer's Rule**.

We already saw substitution at work in the Markovian employment model. It can be a quick way to solve linear systems when they have the right kind of structure, but in general, it may be difficult to make progress using substitution.

7.2 Rewriting the Equation System

Gauss and Gauss-Jordan elimination is based on the fact that the equations of a linear system may be manipulated in certain ways without affecting the solutions of the system.[1]

Consider the linear system

$$
\begin{align}
a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \tag{E_1}\\
a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \tag{E_2}\\
a_{31}x_1 + a_{32}x_2 + \cdots + a_{3n}x_n &= b_3 \tag{E_3}\\
\vdots \qquad \vdots \qquad \vdots \qquad \qquad \vdots \\
a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \tag{E_m}
\end{align}
$$

We refer to the ith equation as E_i and the whole system (E_1)–(E_m) as E. We will be interested in ways of altering the equations that do not change the solutions to the system.

---

[1] Gaussian elimination predates both Gauss and Jordan. It was used by Sir Isaac Newton, and first published by Michel Rolle (1652–1719) of Rolle's Theorem. Some special cases were already known to Chinese mathematicians by 179 AD.

Sir Isaac Newton (1643-1727 NS) was one of the most influential scientists of all time. He was famous for developing the calculus, Newton's Laws of Motion and Gravitation, classical mechanics, his theory of optics and color, the first practical reflecting telescope, forms of which are still made today, and much, much more. He also was an alchemist and theologian. He even was Master of the Mint (1699–1727). He took the job very seriously, overseeing the Great Recoining and punishing counterfeiters and clippers (who cut pieces of metal from the edges of coins).

7.3 Rewrites that Don't Change the Solution Set

The following theorem forms the basis for the Gauss-Jordan method.[23] It tells us that certain changes to the system of equations **do not change the solution set**. We can reorder equations or use either of two non-destructive linear operations: adding equations and multiplying by a non-zero number. Multiplying by zero destroys information, and is not allowed.

Theorem 7.3.1. *Suppose a linear system* $E = (E_1)–(E_m)$ *is altered by either*

1. *changing the order of the equations,*
2. *multiplying an equation by a non-zero constant, or*
3. *adding one equation to another.*

Then the set of solutions to the system is not changed.

In other words, none of our three basic equation operations affect the solution set.

---

[2] Carl Friedrich Gauss (1777-1855) was a German who ranks among the greatest mathematicians of all time. He also made important contributions to various scientific fields. In physics, he discovered Gauss's Law, which relates the distribution of electric charges to the electric field they generate. His first published work in mathematics was a text number theory at age 24. The book developed many new results, made the field rigorous and laid the groundwork for its further development. As economists, we know him for the Gaussian (normal) distribution and ordinary least squares. He was perhaps most influential in geometry. He was to the first to discover non-Euclidean geometries. He is also known for the Gaussian curvature and the Gauss-Bonnet theorem, an important result in differential geometry.

[3] Camille Jordan (1838–1922) was a French mathematician who worked in both algebra and analysis. Much later, we will later the Jordan normal form for matrices. In algebra, he worked on group theory, including the Jordan-Hölder Theorem. In analysis, he's known for the Jordan Curve Theorem used in complex analysis, and the Jordan content, a method of measuring area.

7.4 Proof of Theorem 7.3.1

Theorem 7.3.1. *Suppose a linear system* $E = (E_1)–(E_m)$ *is altered by either*

1. *changing the order of the equations,*
2. *multiplying an equation by a non-zero constant, or*
3. *adding one equation to another.*

Then the set of solutions to the system is not changed.

Proof. (1) Consider the set of solutions of equations $E = (E_1)–(E_m)$. If we change the order of equations, we still must satisfy $(E_1)–(E_m)$, which means the solutions have not changed.

(2) Suppose we multiply equation i by $\alpha \neq 0$. Then we have a new linear equation

$$\alpha a_{i1}x_1 + \alpha a_{i2}x_2 + \cdots \alpha a_{in}x_n = \alpha b_i \qquad (E_i')$$

Consider the new system E' which is the same as E, except that E_i has been replaced by E_i'. If the numbers $(x_1, \ldots, x_n)$ satisfy the old system, they obey each of the new equations, including E_i'. And if $(x_1, \ldots, x_n)$ satisfies the new system, the fact that $\alpha \neq 0$ allows us to divide by α, showing that E_i is also satisfied. Since the other equations are the same, the old system E is also satisfied.

(3) Suppose we replace equation (E_i) with the sum of (E_i) and (E_j). Call this equation (E_i''). Define a new system E'' starting from E by replacing E_i with E_i'' and leaving the remaining equations of E unchanged. If $(x_1, \ldots, x_n)$ solves E, it must also solve E_i'', meaning the system E'' is satisfied. Conversely, if E'' is satisfied by $(x_1, \ldots, x_n)$, so is the equation formed by subtracting $E_j'' = E_j$ from E_i''. But that is just E_i, so E is also satisfied. ∎

7.5 Corollary to Theorem 7.3.1

The result is sometimes stated in a slightly different form, where we add a multiple of one equation to another. But we can get there from Theorem (7.3.1) in a couple of steps. First, we multiply the one equation by $\alpha \neq 0$, in step two we add it to the other, in step three, we divide the first equation by α, leaving us with a system where the only change to replace the second equation by the sum of α times the first equation and the second equation. We restrict our attention to non-zero α because the $\alpha = 0$ case leaves everything unchanged.

That gives us the following corollary.

Corollary 7.5.1. *Suppose a linear system* $E = (E_1)–(E_m)$ *is altered by either*

1. *interchanging two equations,*

2. *multiplying an equation by a non-zero constant, or*

3. *replacing one equation by its sum with a non-zero multiple of another.*

Then the set of solutions to the system is not changed.

The three equation operations can be referred to as *elementary equation operations*. When we later put them into a matrix context, they will become the elementary row operations. In (1), changing the order of equations was replaced by swapping any two equations. That basic operation can be used to generate any change in their order.

7.6 Gauss-Jordan Elimination

Before starting the Gauss-Jordan procedure, we make sure that each equation has the variables written in order. Thus we write

$$3x_1 + 5x_2 + 16x_3 - 8x_4 = 0$$

rather than something like

$$5x_2 - 8x_4 + 16x_3 + 3x_1 = 0.$$

   The Gauss-Jordan procedure proper starts (Step 1) by arranging the equations so that the first non-zero term in the top equation is as far to the left as possible.

   Step 2 is to divide that equation by the first non-zero coefficient, resulting in an equation where the leading non-zero coefficient is one.

   Step 3 is referred to as the *pivot* step. Multiples of the top equation are subtracted from each of the other equations (as needed) to zero out the coefficient on the pivot variable in those equations, leaving just one equation (the top one) where the pivot variable is non-zero (in fact, one).

   Finally, we keep repeating this procedure on the remaining equations until none are left. We do use the pivot to zero out both the terms above and below the row we are working with. In the end, we will have a simplified system of equations telling us the value of various variables, sometimes in terms of other variables.

7.7 Gauss-Jordan in Action

Consider the system

$$2x_1 - 2x_2 = 10 \tag{1}$$
$$3x_1 + 2x_2 = 3 \tag{2}$$

We start by making sure the top equation has a leading non-zero coefficient as far left as possible. Since x_1 has coefficient 2, this is true, and we do not have to rearrange the equations. Next, we divide equation (1) by 2 obtaining

$$x_1 - x_2 = 5 \tag{1'}$$
$$3x_1 + 2x_2 = 3 \tag{2}$$

where equation $(1') = \frac{1}{2} \times (1)$.

Next comes the pivot step, where we subtract 3 times equation $(1')$ from equation (2).

$$x_1 - x_2 = 5 \tag{1'}$$
$$0x_1 + 5x_2 = -12. \tag{2'}$$

Notice that $(2') = (2) - 3 \times (1')$.

7.8 Repeat Until Done

We now repeat the procedure on the remaining equations. In this case, there is only one such. Since the coefficients on x_1 are zero in the remaining equations due to the pivot step, we look for a non-zero coefficient on x_2, and there it is!

We then divide (2') by that coefficient, 5, obtaining (2") = (2')/5:

$$x_1 - x_2 = 5 \tag{1'}$$
$$x_2 = -12/5. \tag{2''}$$

We finish the problem off by pivoting on x_2, (1") = (1') + (2"):

$$x_1 = 13/5 \tag{1''}$$
$$x_2 = -12/5. \tag{2''}$$

This system is now in a form where we can read off the solution, $(x_1, x_2) = (13/5, -12/5)$.

It is often a good idea to plug your solutions back in the original system to make sure you have not made any calculation errors, something that is far too easy to do.

$$2x_1 - 2x_2 = 26/5 + 24/5 = 10$$
$$3x_1 + 2x_2 = 39/5 - 24/5 = 3$$

Our answers passed the test. We made no mistake in deriving them.

7.9 Equations as Number Arrays   08/25/22

As you have seen, even a simple Gauss-Jordan calculation involves a lot of writing. Things can be simplified by converting the equations to a rectangular array of numbers, a *matrix*.[4]

Suppose we start with the system

$$\begin{aligned} 2x_1 + 3x_2 + x_3 &= 10 \\ 5x_1 + 2x_2 + 2x_3 &= 5 \end{aligned} \tag{7.9.1}$$

The x_i's are repeated in both equations. **As long as we write them in the same order**, it is enough to keep track of the coefficients. The first coefficient then belongs to x_1, the second to x_2, etc. We don't need to explicitly track the x_i's.

Replace the system by

$$\left(\begin{array}{ccc|c} 2 & 3 & 1 & 10 \\ 5 & 2 & 2 & 5 \end{array} \right),$$

where we have used a vertical bar to separate the constant terms b_i from the coefficients of the x_i. Column one contains the x_1 coefficients, column two the x_2 coefficients, etc., through the next-to-last column. The last column contains the constant terms.

---

[4] The German mathematician Gottfried Leibniz (1646–1716) seems to have been the first to write the coefficients as an array of number. He was a true polymath, working in a wide variety of fields. He was even a diplomat! He is best known as the co-inventor of calculus, which he invented independently of Newton. He used infinitesimals. Interestingly, some of his rules for the use were justified in the mid-20th century by the development of non-standard analysis and the hyperreal numbers. He expended considerable effort in developing the system of calculus, derivatives, integrals, the Fundamental Theorem of Calculus, and used these results to solve a variety of problems (as did Newton). The roughly simultaneous discovery of calculus led to a bitter feud, not only between Newton and Leibniz, but also between their supporters.

7.10 Coefficient and Augmented Matrices: First Look

We refer to

$$A = \begin{pmatrix} 2 & 3 & 1 \\ 5 & 2 & 2 \end{pmatrix}$$

as the *coefficient matrix* and

$$\hat{A} = \begin{pmatrix} 2 & 3 & 1 & 10 \\ 5 & 2 & 2 & 5 \end{pmatrix}$$

as the *augmented matrix*. Here the term *matrix* means that we have a rectangular array of numbers. Since there are two equations, the matrix A has two rows, while the three variables mean there are three columns, giving us a 2×3 matrix. If A is the matrix of coefficients, we denote the corresponding augmented matrix by $\hat{A}$. It has an extra column for the constant terms, meaning it is a 2×4 matrix.

7.11 Coefficient and Augmented Matrices: General Case

In the general case, we have a linear system of m equations in n variables:

$$a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1$$

$$a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2$$

$$a_{31}x_1 + a_{32}x_2 + \cdots + a_{3n}x_n = b_3$$

$$\vdots \qquad\qquad\qquad \vdots$$

$$a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m.$$

The *coefficient matrix* is the $m \times n$ matrix formed by the coefficients a_{mn}.

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

and the *augmented matrix* is the $m \times (n+1)$ matrix including the constant terms

$$\hat{A} = \left(\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array} \right).$$

I've put in a vertical bar to emphasize that we're adding an extra column to form the augmented matrix $\hat{A}$. In calculations, I will omit it for simplicity, as I've done on the following page.

7.12 Simplifying the Calculations

We can use the augmented matrix to simplify the calculations of the
Gauss-Jordan method by keeping track of the variables by position (col-
umn) rather than continually writing down the variables. When we add
or multiply equations we are adding or multiplying the coefficients, in-
cluding the constant term.

The elementary equation operations become the *elementary row op-
erations* in the augmented matrix: swapping any two rows, multiplying a
row by a non-zero constant, and adding a non-zero multiple of one row
to another. The Gauss-Jordan procedure applied to matrices is called
row reduction.

Let's try this with the augmented matrix $\hat{\mathbf{A}}$ given above.

$$
\hat{\mathbf{A}} = \begin{pmatrix} 2 & 3 & 1 & 10 \\ 5 & 2 & 2 & 5 \end{pmatrix} \xrightarrow{(1)/2} \begin{pmatrix} 1 & 3/2 & 1/2 & 5 \\ 5 & 2 & 2 & 5 \end{pmatrix}
$$

$$
\xrightarrow{(2)-5(1)} \begin{pmatrix} 1 & 3/2 & 1/2 & 5 \\ 0 & -11/2 & -1/2 & -20 \end{pmatrix}
$$

$$
\xrightarrow{-\frac{2}{11}*(2)} \begin{pmatrix} 1 & 3/2 & 1/2 & 5 \\ 0 & 1 & 1/11 & 40/11 \end{pmatrix}
$$

$$
\xrightarrow{(1)-(3/2)(2)} \begin{pmatrix} 1 & 0 & 4/11 & -5/11 \\ 0 & 1 & 1/11 & 40/11 \end{pmatrix}.
$$

7.13 Interpreting the Solution

At this point, we have run out of rows. Our fully reduced augmented matrix is

$$\begin{pmatrix} 1 & 0 & 4/11 & -5/11 \\ 0 & 1 & 1/11 & 40/11 \end{pmatrix}.$$

To see what we've got, let's write down the corresponding equations.

$$x_1 + \frac{4}{11}x_3 = -\frac{5}{11}$$
$$x_2 + \frac{1}{11}x_3 = \frac{40}{11}.$$

We have two equations with three unknowns. With this system, every choice of x_3 (or x_1 or x_2) will give us a different solution.

Any combination of the form

$$(x_1, x_2, x_3) = \left(-\frac{5 + 4x_3}{11}, \frac{40 - x_3}{11}, x_3 \right)$$

is a solution.

By plugging back into the original system (7.9.1), we can verify that they all work. This system not only has a solution, it has infinitely many solutions!

Since the solutions are linear functions of a single parameter, the solution set is a line in $\mathbb{R}^3$.

7.14 A System without a Solution

The linear system

$$x_1 + 3x_2 = 6$$
$$2x_1 + 6x_2 = 10$$

has no solution.

To see this, we attempt to find a solution by forming the augmented matrix and row-reducing.

$$\begin{pmatrix} 1 & 3 & 6 \\ 2 & 6 & 10 \end{pmatrix} \xrightarrow{(2)-2(1)} \begin{pmatrix} 1 & 3 & 6 \\ 0 & 0 & -2 \end{pmatrix} \xrightarrow{-\frac{1}{2}\times(2)} \begin{pmatrix} 1 & 3 & 6 \\ 0 & 0 & 1 \end{pmatrix} \xrightarrow{(1)-6(2)} \begin{pmatrix} 1 & 3 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

This matrix is a problem. To see the problem, write out the corresponding system of equations

$$x_1 + 3x_2 = 0$$
$$0 = 0x_1 + 0x_2 = 1.$$

The second equation cannot possibly be true. That means the system has no solutions.

In fact, there's a problem with the first equation too. It contradicts one of our original equations (the first one).

7.15 Row-Echelon Form

A row of a matrix has k *leading zeros* if the first k elements of the row are zero and the $(k + 1)^{\text{st}}$ element is not zero. If a row starts with a non-zero element, that means there are zero leading zeros.

We say a matrix is in *row echelon form* if each row has more leading zeros than the row before it. The number of leading zeros increases from row to row. Thus

$$A = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix}$$

is **not** in row-echelon form. The first row has two leading zeros, the second row has none, and the bottom row has one leading zero. The order of the rows needs to be changed to put the matrix in row-echelon form. In this case, we need to interchange rows one and two, and then interchange the new row two (old row one) with row three to put it in row-echelon form.

In contrast, the matrices

$$B = \begin{pmatrix} 1 & 3/2 & 1/2 & 5 \\ 0 & 0 & -1/2 & -20 \end{pmatrix} \quad \text{and} \quad C = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}$$

are in row-echelon form. Matrix **B** has no leading zeros in row one, and two leading zeros in row two. The rows are arranged so that the number of leading zeros is increasing, the hallmark of row-echelon form.

Matrix **C** has no leading zeros in the first row, one in the second row, and two leading zeros in the third row. Again, the number of leading zeros is increasing, so **C** is in row-echelon form.

7.16 Reduced Row-Echelon Form

A matrix is in *reduced row-echelon form* if (1) it is in row-echelon form, (2) each leading non-zero term is 1, and (3) there are no other non-zero terms in each column where there is a leading 1 (we have pivoted using each leading 1).

The matrix $\mathbf{B}$ above is in row-echelon form, but not reduced row-echelon form because the leading non-zero term in the second row is $-1/2$ rather than 1. Even if we change it to 1, it is still not in reduced row-echelon form because there is another non-zero term in the second column.

The matrix $\mathbf{C}$ is also not in reduced row-echelon form even though every row has leading one. The leading ones in rows one and two are in otherwise zero columns. The problem is in row three. The leading one is in column three and the rest of the column is not zero. Further row reduction is possible by pivoting on the entry in row three, column three.

Here is a matrix in reduced row-echelon form:

$$\begin{pmatrix} 0 & 1 & 2 & 0 & 5 \\ 0 & 0 & 0 & 1 & 2 \end{pmatrix}.$$

Row one has one leading zero while row two has three leading zeros. Moreover, both columns two and four are pivot columns. Only row one has a non-zero entry in column two and only row two has a non-zero entry in column four.

7.17 Gauss-Jordan and Reduced Row-Echelon Form

The matrix
$$\begin{pmatrix} 1 & 0 & 4/11 & -5/11 \\ 0 & 1 & 1/11 & 40/11 \end{pmatrix},$$

which we found using the Gauss-Jordan procedure, is in reduced row-echelon form, as is

$$\mathbf{R} = \begin{pmatrix} 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

The Gauss-Jordan method always yields a matrix in reduced row-echelon form due to the pivot steps. The simpler method of Gaussian elimination, which omits the pivot step, puts the matrix in row-echelon form, but it will usually not be reduced row-echelon form. Gaussian elimination is more efficient for calculation, but Gauss-Jordan elimination is more useful for understanding the solution.

7.18 Gauss–Jordan Always Yields a Reduced Matrix

All matrices produced by the Gauss-Jordan procedure must be in reduced row-echelon form.

Theorem 7.18.1. *If* **A** *is a matrix produced by the Gauss-Jordan procedure, it is in reduced row-echelon form.*

Proof. In each cycle of the Gauss-Jordan procedure, first find a row with a non-zero term as far left as possible, then divide to make the leading term 1. Following that, we pivot to eliminate any non-zero terms in that column. At that point the rows below the pivot row have more leading zeros than the pivot. As we repeat the Gauss-Jordan cycle, the number of leading zeros increases.

When the Gauss-Jordan procedure is done, the higher the row, the farther left the leading 1 will be. Moreover, each leading 1 is in a column that is otherwise zero due to the pivot step. ∎

7.19 Basic and Free Variables

If an $m \times n$ matrix is in reduced row-echelon form, a variable is called *basic* if its column contains a pivot, and called *free* if its column contains no pivot. Notice that every variable is either basic or free, but never both.

In the reduced coefficient matrix

$$\mathbf{R} = \begin{pmatrix} 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

variables x_2 and x_5 have pivots, so they are basic. The other variables, (x_1, x_3, x_4), are free.

Because each of the n variables is either basic or free, but never both, we have the following equation.

$$\#\text{basic vars} + \#\text{free vars} = n. \tag{7.19.2}$$

7.20 Free Variables determine the Basic Variables

As we saw earlier we can reduce the system

$$2x_1 + 3x_2 + x_3 = 10$$
$$5x_1 + 2x_2 + 2x_3 = 5$$

$$(7.9.1)$$

to

$$x_1 + \frac{4}{11}x_3 = -\frac{5}{11}$$
$$x_2 + \frac{1}{11}x_3 = \frac{40}{11}.$$

When a system has a solution and one or more free variables, we can choose the values of the free variables (x_3 here) to be anything we wish. After the free variables have been chosen, the values of the basic variables are fully determined.

The set of solutions in equation (7.9.1) depends linearly on a single parameter, x_3. That means the solution set is a line, an infinite straight line. If we had two free variables, we would get a plane. Three free variables would give us a 3-dimensional solution set. This solution set contains all lines generated by any two distinct points in it, and all planes generated by three non-collinear points in it. The situation in higher dimensions is similar.

7.21 Basic Variables, Free Variables, and Rank

Because the number of free variables is tied to the solution set, it remains unchanged under any of the elementary equation/row operations.[5] Then we can use equation (7.19.2) to see that the number of basic variables is also unaffected by elementary row operations.

We now define the *rank* of an $m \times n$ matrix as the number of basic variables it possesses. The rank obeys

$$\text{rank}\,\mathbf{A} + \#\text{free vars} = n. \qquad (7.21.3)$$

Since the rank can be derived from number of variables and type of the solution set (point, line, plane, etc.), **the rank remains unchanged** whenever we use elementary row or equation operations.

---

[5] This is a argument in favor of the statement, but is not yet a proof. Chapter 27 of Simon and Blume has more on this. See section 27.3 of the notes.

7.22 More on the Rank of a Matrix

If a matrix is in row echelon form, the number of basic variables will be the number of non-zero rows in the matrix. It follows that the *rank* of a matrix can equivalently be defined as the number of non-zero rows in row echelon form. It does not depend on which row echelon form we use as they all have the same solution set and so the same number of basic and free variables. In fact, the solution set is a flat set Z in $\mathbb{R}^n$ with dimension equal to the number of free variables.[6] The rank is then

$$\operatorname{rank} \mathbf{A} = n - \dim Z = n - \#\text{free vars.}$$

---

[6] We will not show this until after we have defined dimension in Chapter 11.

7.23 Some Properties of the Rank

We start with a few properties of the rank.

Theorem 7.23.1. *The rank of a matrix has the following properties.*

1. *If $\hat{\mathbf{A}}$ is an augmented matrix, then either $\operatorname{rank}\hat{\mathbf{A}} = \operatorname{rank}\mathbf{A}$ or $\operatorname{rank}\hat{\mathbf{A}} = 1 + \operatorname{rank}\mathbf{A}$.*

2. $\operatorname{rank}\mathbf{A} \leq$ *number of rows of* $\mathbf{A}$.

3. $\operatorname{rank}\mathbf{A} \leq$ *number of columns of* $\mathbf{A}$.

Proof. When we row-reduce $\hat{\mathbf{A}}$, we obtain a matrix of the form $(\mathbf{A}^* \mid \mathbf{B}^*)$ where $\mathbf{A}^*$ is a row reduction of $\mathbf{A}$ and $\mathbf{B}^*$ is an $m \times 1$ matrix obtained by row reducing the $\mathbf{B}$ portion of the augmented matrix. Either the column $\mathbf{B}^*$ adds a single non-zero row to $(\mathbf{A}^* \mid \mathbf{B}^*)$, or it does not, showing that $\operatorname{rank}\hat{\mathbf{A}}$ is either $\operatorname{rank}\mathbf{A}$ or $\operatorname{rank}\mathbf{A} + 1$.

Since the number of non-zero rows in the row reduction is no more than the number of rows in $\mathbf{A}$, (2) follows.

For (3), every non-zero row in the row reduction has a leading non-zero entry is in a different column, so there are no more leading non-zero entries than there are columns, establishing the result. ∎

7.24 Rank and Solutions

A key result relates the ranks of a matrix and of its augmented matrix to the existence of a solution to the linear system described by the augmented matrix.

Augmented Rank Theorem. *A linear system with coefficient matrix $\mathbf{A}$ and augmented matrix $\hat{\mathbf{A}}$ has a solution if and only if* $\operatorname{rank} \mathbf{A} = \operatorname{rank} \hat{\mathbf{A}}$.

   Proof. As in the proof of Theorem (7.23.1), we can write the row-reduced matrix for $\hat{\mathbf{A}}$ in the form $(\mathbf{A}^* \,|\, \mathbf{B}^*)$.

   If case ($\Longleftarrow$): $\operatorname{rank} \mathbf{A} = \operatorname{rank} \hat{\mathbf{A}}$, then each non-zero row of $\mathbf{A}^*$ contains a basic variable and we can find a solution by setting all of the free variables to zero.

   Only if case ($\Longrightarrow$): If the system has a solution, then $\mathbf{B}^*$ cannot have a non-zero entry corresponding to a zero row of $\mathbf{A}^*$ as it would contradict the fact that we have a solution. It follows that $\operatorname{rank} \mathbf{A} = \operatorname{rank} \hat{\mathbf{A}}$. ∎

7.25 The Augmented Rank Theorem: Two Corollaries

In the following corollaries, $\mathbf{A}$ is an $m \times n$ coefficient matrix.

Corollary 7.25.1. *A system of linear equations either has no solution, one solution or infinitely many solutions.*

Proof. By the Augmented Rank Theorem, there is no solution if and only if rank $\mathbf{A} <$ rank $\hat{\mathbf{A}}$. Furthermore, there will be solutions if and only rank $\mathbf{A} =$ rank $\hat{\mathbf{A}}$. In this case, if there are any free variables there will be infinitely many solutions. That leaves the case of no free variables where the solution is unique. ∎

Corollary 7.25.2. *If a linear system has a unique solution, then $m \geq n$, meaning it has at least as many equations as variables.*

Proof. If there is a unique solution, there are no free variables. Every unknown must be basic, and that requires at least n equations, so $m \geq n$. ∎

7.26 The Augmented Rank Theorem: A Third Corollary

Corollary 7.26.1. *If a linear system has more variables than its rank, either it has no solution or infinitely many solutions. In particular, this holds if it has more variables than equations, if $n > m$.*

Proof. We know that for an $m \times n$ matrix A, $n = \text{rank}\, A + \#\text{free vars}$ since the rank is the number of basic variables. If $n > \text{rank}\, A$, there must be free variables and so either no solution, or infinitely many solutions.

Since $\text{rank}\, A \leq m$, the number of equations, the same holds if there are more variables than equations. ∎

It is possible to not have a solution in Corollary 7.26.1. Consider the case

$$x_1 + 0x_2 + 0x_3 = 1$$
$$3x_1 + 0x_2 + 0x_3 = 12$$

which has no solution even thought there are more variables (3) than equations (2).

7.27 Homogeneous Linear Systems

Recall that a system of linear equations is *homogeneous* if the right-hand side of every equation is zero. That is, it has form

$$
\begin{aligned}
a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= 0 \\
a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= 0 \\
a_{31}x_1 + a_{32}x_2 + \cdots + a_{3n}x_n &= 0 \\
\vdots \qquad\qquad\qquad \vdots \\
a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= 0
\end{aligned}
\qquad (7.27.4)
$$

7.28 Solution of Homogeneous Systems

One nice thing about homogeneous linear systems is that they always have a solution.

We can see that using the Augmented Rank Theorem. Since the right-hand side of a homogeneous system is zero, the constant terms cannot affect the rank of the augmented matrix and rank $\mathbf{A}$ = rank $\hat{\mathbf{A}}$. There will be at least one solution, leaving only the question is whether there are many solutions or just one.

Also, there's an obvious solution to every homogeneous system, the zero vector (n-tuple):

$$(x_1, \ldots , x_n) = (0, \ldots , 0).$$

Theorem 7.28.1. *Let* $\mathbf{A}$ *be the* $m \times n$ *coefficient matrix of a homogeneous linear system. The zero vector* $(x_1, \ldots , x_n) = (0, \ldots , 0)$ *solves this system.*

Proof. If you plug the zero value into equation (7.27.4), you will see they solve the system. ∎

7.29 Unique Solution to Homogeneous Systems

Now that we know the system can be solved by the zero vector, we still need to see whether that is the only solution, or whether there are others.

Corollary 7.29.1. *A homogeneous linear system has infinitely many solutions if and only if there are more unknowns than the rank of the $m \times n$ coefficient matrix, if $n > \text{rank}\,\mathbf{A}$. Equivalently, the solution $(0, \dots , 0)$ is unique if and only if $n = \text{rank}\,\mathbf{A}$.*

   Proof. If case ($\Longleftarrow$): This part is Corollary 7.26.1, taking account of the fact that our system has a solution.

   Only if case ($\Longrightarrow$): Suppose the system has infinitely many solutions. This means there are free variables, so $n > \text{rank}\,\mathbf{A}$.

   For the other part, note that $n \leq \text{rank}\,\mathbf{A}$, so either $n = \text{rank}\,\mathbf{A}$ or $n > \text{rank}\,\mathbf{A}$. The latter case is equivalent to having infinitely many solutions, so the form is equivalent to having a unique solution. ∎

7.30 Properties of Linear Systems I

We now take an $m \times n$ coefficient matrix $\mathbf{A}$ and consider the various systems of linear equations that may be formed that use this coefficient matrix. We examine what happens when we alter the right-hand side of the system (i.e., $(b_1, \ldots, b_n)$).

We list several corollaries to our previous results. In the proofs below, we will use m for the number of equations and n for the number of variables.

Corollary 7.30.1. *Let $\mathbf{A}$ be a matrix of coefficients for a linear system. The rank of $\mathbf{A}$ is equal to the number of rows of $\mathbf{A}$ if and only if the system has a solution for every right-hand side.*

Proof. If rank $\mathbf{A}$ is the number of rows of $\mathbf{A}$, it follows that rank $\hat{\mathbf{A}}$ is also the number of rows of $\mathbf{A}$ regardless of the right-hand side. Then the system has a solution by the Augmented Rank Theorem.

Now suppose the system has a solution for every right-hand side. If the rank of $\mathbf{A}$ where less than the number of rows, we may row reduce the matrix to obtain $\mathbf{A}^*$, which will have at least one zero row. Now let $\mathbf{B}^* = (b_1^*, \ldots, b_n^*) = (0, \ldots, 0, 1)$. This has no solution since the final 1 corresponds to the zero row.

We may now reverse the row reduction on the augmented matrix $\hat{\mathbf{A}}^*$ to obtain a system with coefficient matrix $\mathbf{A}$ and right-hand side $\mathbf{B}$ obtained by undoing the row reduction that has no solution, contradicting the fact that original the system has a solution for every right-hand side. This contradiction shows that the rank of $\mathbf{A}$ must equal the number of rows. ∎

7.31 Properties of Linear Systems II

Corollary 7.31.1. *If a system of linear equations has more equations than unknowns there is a right-hand side where there are no solutions.*

Proof. If there are more equations than unknowns, rank $\mathbf{A} \leq n < m$. We can then row-reduce $\mathbf{A}$ to obtain a matrix $\mathbf{A}^*$ with one or more zero rows. Now let $\mathbf{B}^* = (b_1^*, \ldots, b_n^*) = (0, \ldots, 0, 1)$. This has no solution. We may now reverse the row reduction on the augmented matrix $\hat{\mathbf{A}}^*$ to obtain a system with coefficient matrix $\mathbf{A}$ and right-hand side $\mathbf{B}$ obtained by undoing the row reduction that has no solution. ∎

Corollary 7.31.2. *A system of linear equations has at most one solution for every right-hand side if and only if the rank of the coefficient matrix $\mathbf{A}$ is equal to the number of columns of $\mathbf{A}$.*

Proof. Suppose the rank of the coefficient matrix is equal to the number of columns. In this case there are no free variables, and if a solution exists, it must be unique. ∎

7.32 Singular and Non-singular Matrices

We say a matrix $\mathbf{A}$ of coefficients is *non-singular* if there exists a unique solution to the system for every right-hand side. Conversely, a coefficient matrix $\mathbf{A}$ is *singular* if there is either a right-hand side with no solution or a right-hand side with infinitely many solutions. In the latter case, the homogeneous system will also have infinitely many solutions.

Corollary 7.32.1. *An $m \times n$ matrix $\mathbf{A}$ is non-singular if and only if the number of rows and columns of $\mathbf{A}$ are both equal to* rank $\mathbf{A}$, *that is* $m = n = \text{rank}\,\mathbf{A}$.

   Proof. By Corollary 7.30.1, a matrix $\mathbf{A}$ has a solution for every right-hand side if and only if rank $\mathbf{A}$ is equal to the number of rows of $\mathbf{A}$. By Corollary 7.31.2, $\mathbf{A}$ has at most one solution for every right-hand side if and only if rank $\mathbf{A}$ is the number of columns of $\mathbf{A}$.

   These two conditions together characterize non-singular matrices, proving the result. ∎

7.33 Reduction of Non-singular Matrices

If a matrix is non-singular, its reduced row-echelon form is special.

Lemma 7.33.1. *Suppose an $n \times n$ matrix $\mathbf{A}$ is non-singular. Then its reduced row-echelon form is*

$$
\begin{pmatrix}
1 & 0 & 0 & 0 & \cdots & 0 \\
0 & 1 & 0 & 0 & \cdots & 0 \\
0 & 0 & 1 & 0 & \cdots & 0 \\
\vdots & & & \ddots & & \vdots \\
\vdots & & & & \ddots & \vdots \\
0 & 0 & 0 & 0 & \cdots & 1
\end{pmatrix}. \tag{7.33.5}
$$

In other words, it has 1's on the diagonal and 0's everywhere else.

Proof. We know that $\mathbf{A}$ is square by Corollary 7.32.1. With n columns and n non-zero rows, the only way to write a matrix in reduced row-echelon form is as in equation (7.33.5).

The matrix in equation (7.33.5) is referred to as the *identity matrix*, $\mathbf{I}_n$. We will be properly introduced to it in section 8.18.

7.34 Endogenous Variables and Exogenous Variables

Consider the linear system

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 1 & 5 \\ 1 & 3 & 2 & 12 \end{pmatrix} \quad \text{and} \quad \mathbf{B} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Suppose x_3 and x_4 are *exogenous variables*, variables determined outside the system. Can we write the *endogenous variables*, the variables x_1 and x_2 that are determined by the system using the exogenous variables, as functions of x_3 and x_4?

The augmented matrix is

$$\hat{\mathbf{A}} = \begin{pmatrix} 1 & 2 & 1 & 5 & 1 \\ 1 & 3 & 2 & 12 & 0 \end{pmatrix}$$

which row reduces to

$$\begin{pmatrix} 1 & 0 & -1 & -9 & 3 \\ 0 & 1 & 1 & 7 & -1 \end{pmatrix}$$

The corresponding equations are

$$x_1 = 3 + x_3 + 9x_4$$
$$x_2 = -1 - x_3 - 7x_4$$

and it is evident that for every choice of the exogenous variables (x_3, x_4), we get unique values of the endogenous variables (x_1, x_2).

7.35 Linear Implicit Function Theorem

The Linear Implicit Function Theorem gives conditions that let us write endogenous variables as functions of exogenous variables, as in the example in section 7.34. This implicit function theorem applies to linear systems. We will eventually extend this to include systems of differentiable functions in Chapter 15.

Linear Implicit Function Theorem. *Consider a system with coefficient matrix* **A** *and constant terms* **B**. *Suppose the variables are partitioned into* k *endogenous variables, labelled* $(x_1, \dots , x_k)$, *and* $(n-k)$ *exogenous variables, labelled* $(x_{k+1}, \dots , x_n)$. *Then for every choice of the exogenous variables* $(x_{k+1}^0, \dots , x_n^0)$ *and constant terms* **B**, *there exist unique values of the endogenous variables* $(x_1^0, \dots , x_k^0)$ *so that* $(x_1^0, \dots , x_n^0)$ *solve the system if and only if*

(a) $k = m$, *i.e., the number of endogenous variables, is the number of rows of* **A**.

(b) rank $\mathbf{A}_1 = k$, *where* $\mathbf{A}_1$ *is the matrix formed by the first* k *columns and rows of* **A**.

Compared to the version in Simon and Blume, I added the requirement that there is a solution for every set of constant terms **B**. Without that, the theorem can fail if the exogenous variables have no effect on the equations (i.e., their coefficients are all zero). See section 7.37 for an example.

7.36 Proof of Linear Implicit Function Theorem

Proof. **If case** ($\Longleftarrow$): Assuming (a) and (b), consider the reduced row-echelon form of $(\mathbf{A} \,|\, \mathbf{B})$. Corollary (7.32.1) tells us that $\mathbf{A}_1$ is non-singular, so in the reduced row-echelon form it is given by equation (7.33.5). That lets us write the reduced row-echelon form as $(\mathbf{A}_1^* \,|\, \mathbf{A}_2^* \,|\, \mathbf{B}^*) = (\mathbf{I}_k \,|\, \mathbf{A}_2^* \,|\, \mathbf{B}^*)$ where $\mathbf{A}_1^*$ is $k \times k$, $\mathbf{A}_2^*$ has $n-k$ columns, and $\mathbf{B}^*$ is derived from the constant terms. The basic variables $x_1, \ldots, x_k$ can be uniquely expressed in terms of the free variables $x_{k+1}, \ldots, x_n$ and the constant terms. For a given set of exogenous variables, there will be only one corresponding set of endogenous variables.

Only if case ($\Longrightarrow$): Since we can always solve for the k endogenous variables, Corollary 7.31.1 tells us that there are at most k equations (and hence k rows in $\mathbf{A}$). This establishes (a).

We can now write the coefficient matrix as $(\mathbf{A}_1 \,|\, \mathbf{A}_2)$ where $\mathbf{A}_1$ has k rows and k columns. The other variables are being treated as constants, and for this to always have a unique solution, the rank of $\mathbf{A}_1$ must be k. Thus (b) holds. ∎

The Linear Implicit Function Theorem gives conditions that allow us to write $(x_1, \ldots, x_k)$ as a function of $(x_{k+1}, \ldots, x_n)$.

7.37 Failure of the Uncorrected Theorem

Let's examine the following system

$$x_1 + x_3 + x_4 = 1$$
$$0x_1 + 0x_2 + 0x_3 + 0x_4 = 0.$$

Consider the case where there is one endogenous variable (x_1) and x_2, x_3, x_4 are all exogenous. For every choice of the exogenous variables, there is a unique value of $x_1 = 1 - x_3 - x_4$ that solves the system. Nonetheless, the theorem fails as $m = 2 > 1 = k$.

By requiring a unique solution for all right-hand sides and endogenous variables, we eliminate this possibility. This system fails that condition does not have a solution when $b_2 \neq 0$.

Yes, this example is pretty contrived, but the statement of the theorem in Simon and Blume allows it. My version of the theorem rules out this case.

7.38 LIFT Deficiencies: Choosing the Endogenous Variables

To make the Linear Implicit Function apply to the following case, we have to relabel the variables. Suppose

$$\mathbf{A} = \begin{pmatrix} 1 & 1 & 2 & 1 \\ 1 & 1 & 1 & 2 \end{pmatrix} \quad \text{and} \quad \mathbf{B} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

The augmented matrix is

$$\hat{\mathbf{A}} = \begin{pmatrix} 1 & 1 & 2 & 1 & 0 \\ 1 & 1 & 1 & 2 & 1 \end{pmatrix}$$

which row reduces to

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

with x_1 and x_2 intended as endogenous variables.

However, the only equation involving x_1 and x_2 is

$$x_1 + x_2 = 2 - 3x_4.$$

While this certainly has solutions for any values of exogenous variables x_3 and x_4, the solution fails to be unique. The problem here is that with $k = 2$ endogenous variables,

$$\mathbf{A}_1 = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

has rank one, which is less than $k = 2$.

7.39 Renaming the Variables

Here we must reconsider our choice of exogenous and endogenous variables.

Taking (x_2, x_4) as exogenous and (x_1, x_3) as endogenous would solve the problem as:

$$x_1 = 2 - x_2 - 3x_4$$
$$x_3 = -1 + x_4.$$

We can put this in the proper form for the Linear Implicit Function Theorem by relabelling the variables with $(y_1, y_2) = (x_1, x_3)$ and $(y_3, y_4) = (x_2, x_4)$. Then we have partitioned y into endogenous and exogenous variables as required by the Linear Implicit Function Theorem. The original matrix becomes

$$A' = \begin{pmatrix} 1 & 2 & 1 & 1 \\ 1 & 1 & 1 & 2 \end{pmatrix}.$$

Here A'_1, the first two columns of A', has rank 2. That means the Linear Implicit Function Theorem applies.

Row reducing the augmented matrix yields the system

$$y_1 = 2 - y_3 - 3y_4$$
$$y_2 = -1 + y_4,$$

which clearly shows the endogenous variables (y_1, y_2) are functions of the exogenous variables (y_3, y_4).

7.40 Rethinking the Linear Implicit Function Theorem

This is not the only renaming we could use. Setting $(y_1, y_2) = (x_1, x_4)$ and $(y_3, y_4) = (x_2, x_3)$ also works. The new coefficient matrix is

$$\begin{pmatrix} 1 & 1 & 1 & 2 \\ 1 & 2 & 1 & 1 \end{pmatrix}$$

Again, the first two columns have rank two and the Linear Implicit Function Theorem applies.

   More generally, suppose we partition the variable into k endogenous variables $\{j_1, \dots, j_k\}$ and $n-k$ exogenous variables $\{j_{k+1}, \dots, j_n\}$. We can use the j_i to rename the variables as $\{j_1, \dots, j_n\}$. We then form the matrix $\mathbf{A}'$ with $a'_{ij} = a_{ij_i}$. If the matrix $\mathbf{A}'$ satisfies the assumptions of the Linear Implicit Function Theorem, we can write the endogenous variables as functions of the exogenous variables.

7.41 Rethinking the Linear Implicit Function Theorem

This renaming of variables allows us to restate the Linear Implicit Function Theorem as follows.

Linear Implicit Function Theorem II. *Consider a system with coefficient matrix* **A** *and constant terms* **B***. Suppose the variables are partitioned into* k *endogenous variables,* $\{j_1, \dots, j_k\}$, *and* $(n-k)$ *exogenous variables,* $\{j_{k+1}, \dots, j_n\}$. *Then for every choice of the exogenous variables* $(x_{j_{k+1}}^0, \dots, x_{j_n}^0)$ *and constant terms* **B***, there exist unique values of the endogenous variables* $(x_{j_1}^0, \dots, x_{j_k}^0)$ *so that* $(x_{i_j}^0, \dots, x_{j_n}^0)$ *solve the system if and only if*

(a) $k = m$, *i.e., the number of endogenous variables, is the number of rows of* **A***.*

(b) *rank* **C** $= k$, *where* **C** *is the* $k \times k$ *matrix consisting of columns* $i_1, \dots, i_k$ *of* **A***.*

October 26, 2022