Mathematics for Economists: Basic Linear Algebra

Juuso Välimäki

Spring 2022

In these notes, we will review the following topics from Matrix Algebra:

- Matrices and vectors
- Systems of linear equations
- Square matrices and elementary row operations
- Rank of a matrix
- Singular and non-singular matrices
- Linear economic models
- Matrix multiplication as a linear function: injectivity, surjectivity, bijectivity.

Vectors and matrices

Recall from high school the notion of vectors in the plane and in three-dimensional space. A vector is an object with a length and a direction. It is convenient to regard vectors as arrows that start at the origin and end at some point (in the plane or in space depending on the context). With this way of thinking, each vector can be identified with the coordinates of its endpoint.

In the plane $\mathbb{R}^2$, a vector $\boldsymbol{x}$ is an ordered pair of its coordinates along the two axis $\boldsymbol{x} = (x_1, x_2)$ where x_i is a real number for $i \in \{1, 2\}$. Notice that the order matters and e.g. $(1, 2) \neq (2, 1)$. In three dimensions, $\boldsymbol{x} = (x_1, x_2, x_3)$. In the plane, vector $\boldsymbol{y} = (y_1, y_2)$ is equal $\boldsymbol{x} = (x_1, x_2)$ if $y_1 = x_1$ and $y_2 = x_2$, and similarly for the three dimensional case.

We define the addition of vectors by $\boldsymbol{x} + \boldsymbol{y} = (x_1 + y_1, x_2 + y_2)$ and the multiplication of vectors by a real numbers $a \in \mathbb{R}$ by $a\boldsymbol{x} = (ax_1, ax_2)$. With this definition, sums and real multiples of vectors are again vectors.

We can define vectors in a similar manner for any dimension k. A vector $\boldsymbol{x}$ is a k -dimensional vector if $\boldsymbol{x} = (x_1, x_2, ..., x_k)$, where x_i is a real number for all $i \in \{1, ..., k\}$. In this case, we write $\boldsymbol{x} \in \mathbb{R}^k$. For any $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^k$, define $\boldsymbol{x} + \boldsymbol{y} = (x_1 + y_1, x_2 + y_2, ..., x_k + y_k)$ and for $a \in \mathbb{R}$, $a\boldsymbol{x} = (ax_1, ax_2, ..., ax_k)$. A fancy way of saying that sums and scalar multiples of k -dimensional vectors are also k -dimensional 'vectors is that: $\mathbb{R}^k$ is a real vector space for all k.

A matrix is an array of real numbers into rows and columns. An $m \times n$ -matrix is a matrix with m rows and with n columns. For now, think of matrices as just being arrays. (Later in these notes, it will become apparent that matrices represent linear functions on vectors.)

A matrix $\boldsymbol{A}$ is then an array of the following form:

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

We write then the product of an $m \times n$ -matrix $\boldsymbol{A}$ and a vector $\boldsymbol{x} \in \mathbb{R}^n$ as:

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{11}x_1 & +a_{12}x_2 & \cdots & +a_{1n}x_n \\ a_{21}x_1 & +a_{22}x_2 & \cdots & +a_{2n}x_n \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}x_1 & +a_{m2}x_2 & \cdots & +a_{mn}x_n \end{pmatrix}.$$

At this point, it should be somewhat mysterious why we define multiplication in this way. Hopefully this becomes clear when we talk about matrices as representing linear functions. Just note that the end result of the multiplying an $m \times n$ -matrix $\boldsymbol{A}$ and a vector $\boldsymbol{x} \in \mathbb{R}^n$ is an m -dimensional vector.

We can view vectors as special matrices. A row vector is an (1 $\times$ n) -matrix and a column vector is an (m $\times$ n) -matrix. Whenever we write $\boldsymbol{x} \in \mathbb{R}^k$, we take $\boldsymbol{x}$ to be a column vector.

Gaussian Elimination

Systems of linear equations

A single equation of the form:

$$ax = b,$$

has a solution for all b if $a \neq 0$. The situation is not as obvious if we have many such equations in many real variables. Consider as a first example the following pair of equations:

$$\begin{array}{llll} 1x & +2y & +3z & = 4, \\ 2x & +4y & +6z & = 6. \end{array}$$

The equality of the left-hand side and the right-hand side of an equation is maintained if both sides are multiplied by the same number. Multiplying the first equation by 2, we get:

$$2x + 4y + 6z = 8,$$

and this is inconsistent with the second equation. Hence we see that this pair of equation has no solutions. If the constant on the right hand side of the first equation is 3, the first equation holds if and only if the second equation hold. As a result, and triple $(x, y, z) = (3 - 2y - 3z, y, z)$ gives a solution to the system.

Gaussian elimination provides a systematic approach to the number of solutions to linear systems of equations. A system of m linear equations in n real variables $(x_1, ..., x_n)$ is written as:

$$\begin{array}{cccccc} a_{11}x_1 & +a_{12}x_2 & \cdots & +a_{1n}x_n & = & b_1, \\ a_{21}x_1 & +a_{22}x_2 & \cdots & +a_{2n}x_n & = & b_2, \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ a_{m1}x_1 & +a_{m2}x_2 & \cdots & +a_{mn}x_n & = & b_m. \end{array} \tag{1}$$

In matrix form this is:

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b},$$

where

$$\boldsymbol{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}, \ \boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \ \boldsymbol{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}.$$

The key to Gaussian elimination process relies on two basic arithmetic facts.

1. The solution to an equation is unchanged if both sides of the equation are multiplied by the same non-zero number.

2. If $(x_1, ..., x_n)$ satisfies

   $$\begin{array}{ccccc} a_{11}x_1 & +a_{12}x_2 & \cdots & +a_{1n}x_n & = & b_1, \end{array}$$

   and

   $$\begin{array}{ccccc} a_{21}x_1 & +a_{22}x_2 & \cdots & +a_{2n}x_n & = & b_2, \end{array}$$

   then $(x_1, ..., x_n)$ satisfies:

   $$(a_{11} + a_{21})\, x_1 + ... + (a_{1n} + a_{2n})\, x_n = b_1 + b_2.$$

We call the two fundamental steps in this elimination process *elementary row operations*. They are:

i) Swapping equations in the system (i.e. swapping rows in the associated matrix).

ii) Summing scalar multiples of one equation to another equation (adding a scalar multiple of a row in the augmented matrix to another row).

Solving systems of equations via elementary row operations

Homogenous systems

Consider the system of equations in matrix form

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0},$$

where

$$\boldsymbol{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}, \ \boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix}.$$

Since the right hand side of the equation is zero, this is called a *homogenous system*. It has always a trivial solution $\boldsymbol{x} = (0, ..., 0)$, but we want to know if it has other solutions.

If column k of $\boldsymbol{A}$ has only zeroes, then any vector $\boldsymbol{x}$ such that $x_i = 0$ for $i \neq k$ satisfies the equation system and there are infinitely many solutions (and x_k is not really a variable in the system). Therefore, assume that all columns of $\boldsymbol{A}$ have a non-zero element.

Gaussian elimination gives a systematic way to follow the process of using one of the equations for solving a variable in terms of the other variables and then substituting the result into the other equations. This corresponds to the second type of elementary row operations. The first operation is used to keep the matrix in a format that is easy to read.

If necessary, swap rows of $\boldsymbol{A}$ so that you can eliminate x_1 from the equation given by the top row, i.e. $a_{11} \neq 0$. Add first row multiplied by $-\frac{a_{k1}}{a_{11}}$ to row k for all $k > 1$. This eliminates (makes zero) the elements in the first column of the matrix on rows $k > 1$. We get the following new matrix

$$\boldsymbol{A}^{(1)} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} - a_{21}\frac{a_{12}}{a_{11}} & \cdots & a_{2n} - a_{21}\frac{a_{1n}}{a_{11}} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{m2} - a_{m1}\frac{a_{12}}{a_{11}} & \cdots & a_{mn} - a_{m1}\frac{a_{1n}}{a_{11}} \end{pmatrix}$$
$$=: \begin{pmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1n}^{(1)} \\ 0 & a_{22}^{(1)} & \cdots & a_{2n}^{(1)} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{m2}^{(1)} & \cdots & a_{mn}^{(1)} \end{pmatrix}.$$

Make sure that you understand how this corresponds to the elimination of x_1 from equations given by row $k > 1$ by using the equation on row 1.

Continue by eliminating x_2. If $a_{22}^{(1)} \neq 0$, add the second row multiplied by $-\frac{a_{k2}^{(1)}}{a_{22}^{(1)}}$ to each row $k > 2$.

If $a_{22}^{(1)} = 0$, swap row 2 with k' such that $a_{k'2}^{(1)} \neq 0$ and proceed as before. If $a_{k2}^{(1)} = 0$ for all $k \geq 2$, the elimination of x_1 also eliminated x_2. In this case, continue by eliminating x_3, i.e. multiply the second row of $A^{(1)}$ by $\frac{a_{k3}^{(1)}=0}{a_{23}^{(1)}}$ and add to all rows $k > 2$. If $a_{23}^{(1)} = 0$, swap rows if needed and proceed as before.

This results in a new matrix

$$\boldsymbol{A}^{(2)} = \begin{pmatrix} a_{11}^{(2)} & a_{12}^{(2)} & \cdots & a_{1n}^{(2)} \\ 0 & a_{22}^{(2)} & \cdots & a_{2n}^{(2)} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn}^{(2)} \end{pmatrix}.$$

By repeating the above steps, we get matrices $\boldsymbol{A}^{(3)}, \boldsymbol{A}^{(4)}$ etc. until after k eliminations, we get e.g. for $m = n = 5$,

$$\begin{pmatrix} a_{11}^{(5)} & a_{12}^{(5)} & \cdot & \cdot & a_{15}^{(5)} \\ 0 & a_{22}^{(5)} & a_{23}^{(5)} & \cdot & a_{25}^{(5)} \\ 0 & 0 & a_{33}^{(5)} & a_{34}^{(5)} & a_{35}^{(5)} \\ 0 & 0 & 0 & a_{44}^{(5)} & a_{45}^{(5)} \\ 0 & 0 & 0 & 0 & a_{55}^{(5)} \end{pmatrix},$$

or

$$\begin{pmatrix} a_{11}^{(4)} & a_{12}^{(4)} & \cdot & \cdot & a_{15}^{(4)} \\ 0 & a_{22}^{(4)} & a_{23}^{(4)} & \cdot & a_{25}^{(4)} \\ 0 & 0 & 0 & a_{34}^{(4)} & a_{35}^{(4)} \\ 0 & 0 & 0 & 0 & a_{45}^{(4)} \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

We say that a matrix $\boldsymbol{A}$ is in *row echelon form* if each row k has a larger number of initial zero elements than row $k-1$. Both of the matrices above are in row echelon form. All matrices can be transformed into row echelon form by elementary row operations. The first non-zero element on each row is called a pivot of the matrix.

The number of non-zero rows (or pivots) is the called the *row rank* of a matrix in row echelon form. The top matrix above has row rank 5 and the one below it has row rank 4.

Since each row in the row echelon form starts with more zeros than the previous row, the row rank is always less than or equal to the number of columns. If the row rank is equal to the number of columns, the only solution is the trivial solution $\boldsymbol{x} = 0$. If row rank is less than the number of columns, the system has infinitely many solutions.

In the first case above, the trivial solution is the only solution to the system. This can be seen as follows. The last row in the row echelon form implies that $x_5 = 0$ in any solution. Using this, the second to last row implies that $x_4 = 0$ etc.

In the second case above, x_3 can be chosen freely. For each choice of x_3, the other variables are uniquely determined.

Example 1. Find the row echelon form of the following matrix:

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

i9 Multiply first row by $-\frac{1}{2}$ and add the the second and third row:

$$\begin{pmatrix} 2 & 1 & -1 \\ 1-1 & 2-\frac{1}{2} & 2+\frac{1}{2} \\ 1-1 & 0-\frac{1}{2} & 1+\frac{1}{2} \end{pmatrix} = \begin{pmatrix} 2 & 1 & -1 \\ 0 & \frac{3}{2} & \frac{5}{2} \\ 0 & -\frac{1}{2} & \frac{3}{2} \end{pmatrix}$$

ii) Multiply second row by $\frac{1}{3}$ and add to third row:

$$\begin{pmatrix} 2 & 1 & -1 \\ 0 & \frac{3}{2} & \frac{5}{2} \\ 0 & 0 & \frac{3}{2}+\frac{5}{6} \end{pmatrix} = \begin{pmatrix} 2 & 1 & -1 \\ 0 & \frac{3}{2} & \frac{5}{2} \\ 0 & 0 & \frac{7}{3} \end{pmatrix}.$$

Since the row echelon form has three pivots, it has rank 3, and we know that the system

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}$$

has a unique solution $\boldsymbol{x} = 0$.

Example 2. The following matrix $\boldsymbol{A}$ does not have full rank:

$$\boldsymbol{A} = \begin{pmatrix} 2 & 0 & 4 \\ 1 & 1 & 3 \\ 2 & 1 & 5 \end{pmatrix}.$$

To see this, eliminate the first entry in the second and the third row by using the first row:

$$\begin{pmatrix} 2 & 0 & 4 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix}.$$

When eliminating the second entry on the third row by using the second, we get row echelon form:

$$\begin{pmatrix} 2 & 0 & 4 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

Hence any $\boldsymbol{x} = (-2x_3, -x_3, x_3)$ solves the system

$$\boldsymbol{Ax} = \mathbf{0}.$$

Non-homogenous systems

Consider next the system of n equations in n variables.

$$\boldsymbol{Ax} = \boldsymbol{b}.$$

We will perform elementary row operations to transform $\boldsymbol{A}$ to its row echelon form. It is now useful to consider the *augmented matrix*:

$$(\boldsymbol{A}\vdots\boldsymbol{b}) = \left(\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} & b_n \end{array}\right).$$

We perform the elementary row operations on the entire matrix $(\boldsymbol{A}\vdots\boldsymbol{b})$ to keep track of the right hand side. Obviously this was not necessary in the homogenous case where the right hand side is zero.

$$\left(\begin{array}{cccc|c} a_{11}^{(2)} & a_{12}^{(k)} & \cdots & a_{1n}^{(k)} & b_1^{(k)} \\ 0 & a_{22}^{(k)} & \cdots & a_{2n}^{(k)} & b_2^{(k)} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & a_{nn}^{(k)} & b_n^{(k)} \end{array}\right).$$

i) If $\boldsymbol{A}$ and $(\boldsymbol{A}\vdots\boldsymbol{b})$ have the same rank, then the system has a solution.

ii) If rank $(\boldsymbol{A}\vdots\boldsymbol{b})$) = rank $(\boldsymbol{A}) = n$, the solution is unique.

iii) If rank $(\boldsymbol{A}\vdots\boldsymbol{b})$ = rank $(\boldsymbol{A}) < n$, the system has infinitely many solutions.

iv) If rank $(\boldsymbol{A}\vdots\boldsymbol{b}) >$ rank $(\boldsymbol{A})$, then it has no solution.

v) If $\boldsymbol{x}^0$ is a solution to the homogenous system and $\boldsymbol{x}^1$ is a solution of the non-homogenous system, then $\boldsymbol{x}^0 + \boldsymbol{x}^1$ is also a solution to the non-homogenous system.

vi) If $\boldsymbol{x}'$ and $\boldsymbol{x}''$ solve the non-homogenous system, then $(\boldsymbol{x}' - \boldsymbol{x}'')$ solves the homogenous system.

Example 3. Consider a numerical example for the previous system:

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

$$\begin{pmatrix} 2 & 1 & -1 \\ 1 & 2 & 2 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}.$$

The augmented matrix is now:

$$\left(\begin{array}{ccc|c} 2 & 1 & -1 & 2 \\ 1 & 2 & 2 & 1 \\ 1 & 0 & 1 & 0 \end{array}\right).$$

Repeat the elementary row operations:

$$\left(\begin{array}{ccc|c} 2 & 1 & -1 & 2 \\ 0 & \frac{3}{2} & \frac{5}{2} & 0 \\ 0 & -\frac{1}{2} & \frac{3}{2} & -1 \end{array}\right)$$

and

$$\left(\begin{array}{ccc|c} 2 & 1 & -1 & 2 \\ 0 & \frac{3}{2} & \frac{5}{2} & 0 \\ 0 & 0 & \frac{7}{3} & -1 \end{array}\right).$$

We get:

$$x_3 = \frac{-3}{7}.$$

Substituting into the second row:

$$\frac{3}{2}x_2 + \frac{5}{2}\left(\frac{-3}{7}\right) = 0.$$

Hence:

$$x_2 = \frac{5}{7}.$$

The first row gives:

$$2x_1 + \frac{5}{7} - \frac{-3}{7} = 2 \iff x_1 = \frac{3}{7}.$$

Linear dependence

Consider a set of n column vectors $\mathbf{a}_1, \mathbf{a}_2, ..., \mathbf{a}_n \in \mathbb{R}^m$, where

$$\mathbf{a}_i = \begin{pmatrix} a_{1i} \\ \vdots \\ a_{mi} \end{pmatrix}$$

We say that $\{\mathbf{a}_1, \mathbf{a}_2, ..., \mathbf{a}_n\}$ are linearly dependent if there exists $\lambda \neq 0$, $\lambda = (\lambda_1, \lambda_2, ..., \lambda_n)$ such that

$$\lambda_1\mathbf{a}_1 + \lambda_2\mathbf{a}_2 + ... + \lambda_n\mathbf{a}_n = 0.$$

Write the vectors as a matrix:

$$\boldsymbol{A} = (\mathbf{a}_1, \mathbf{a}_2, ..., \mathbf{a}_n) = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}.$$

The vectors are linearly dependent of there is a $\lambda \neq 0$, such that

$$A\lambda = \mathbf{0}.$$

By using the rank criterion, $\{\mathbf{a}_1, \mathbf{a}_2, ..., \mathbf{a}_n\}$ are linearly dependent if and only if

$$rank\,(\boldsymbol{A}) < n.$$

We see immediately that $\{\mathbf{a}_1, \mathbf{a}_2, ..., \mathbf{a}_n\}$ are linearly dependent if $m < n$.

Matrix Algebra

Let $\boldsymbol{A}$ be a $m \times n$ -matrix. The element of $\boldsymbol{A}$ on the i^{th} row and j^{th} column is denoted by a_{ij}. Similarly for $\boldsymbol{B}, \boldsymbol{C}$, etc. we write the typical element as $b_{ij},\ c_{ij}$, etc.

- Matrix equality:
$$\boldsymbol{A} = \boldsymbol{B}$$
if for all i, j:
$$a_{ij} = b_{ij}.$$

- Scalar multiplication:
For $r \in \mathbb{R}$.
$$r\boldsymbol{A} = \boldsymbol{C},$$
where for all i, j:
$$c_{ij} = ra_{ij}.$$

- Addition (defined only for matrices of same size):
$$\boldsymbol{A} + \boldsymbol{B} = \boldsymbol{C},$$
where for all i, j:
$$c_{ij} = a_{ij} + b_{ij}.$$

- Matrix difference: (combining the two previous ones).
$$\boldsymbol{A} - \boldsymbol{B} = \boldsymbol{A} + (-\boldsymbol{B}) = \boldsymbol{C},$$
where for all i, j:
$$c_{ij} = a_{ij} - b_{ij}.$$

- Matrix multiplication: Let $\boldsymbol{A}$ be an $m \times n$ -matrix and $\boldsymbol{B}$ an $n \times k$ -matrix. The product of $\boldsymbol{A}$ and $\boldsymbol{B}$ is defined as:
$$\boldsymbol{AB} = \boldsymbol{C},$$
where
$$c_{ij} = \sum_{h=1}^{n} a_{ih} b_{hj}.$$

In other words, the element c_{ij} of the product matrix $\boldsymbol{C}$ is the dot product of the $i\ th$ row of $\boldsymbol{A}$ and the j^{th} column of $\boldsymbol{B}$.

Note that $\boldsymbol{A}$ must have the same number of columns as $\boldsymbol{B}$ has rows for multiplication to be defined.

- Why define multiplication like this? Why not element by element? Let
$$\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x} \in \mathbb{R}^m.$$
This is a particular form of a function $\boldsymbol{y} = f(\boldsymbol{x})$. Consider then
$$\boldsymbol{z} = \boldsymbol{B}\boldsymbol{y} \in \mathbb{R}^k.$$
This is a particular form of $\boldsymbol{z} = g(\boldsymbol{y})$ If you write the composite function $\boldsymbol{z} = g(f(\boldsymbol{x}))$ for this case
$$\boldsymbol{z} = \boldsymbol{C}\boldsymbol{x} = \boldsymbol{B}\boldsymbol{A}\boldsymbol{x},$$
and follow the rules for multiplying a matrix and a vector, you get the above definition for matrix multiplication. In other words, matrix multiplication corresponds to the composition of the functions represented by the matrix.

- A truly wonderful webpage for a first course in linear algebra can be found on the (amazing) 3blue1brown channel at Linear Algebra. Chapters 3 and 4 cover this material, but the whole package is strongly recommended.

- Some rules:
$$\begin{aligned}
(\boldsymbol{A}+\boldsymbol{B})+\boldsymbol{C} &= \boldsymbol{A}+(\boldsymbol{B}+\boldsymbol{C}), \\
(\boldsymbol{A}\boldsymbol{B})\,\boldsymbol{C} &= \boldsymbol{A}\,(\boldsymbol{B}\boldsymbol{C}), \\
\boldsymbol{A}+\boldsymbol{B} &= \boldsymbol{B}+\boldsymbol{A}, \\
\boldsymbol{A}\,(\boldsymbol{B}+\boldsymbol{C}) &= \boldsymbol{A}\boldsymbol{B}+\boldsymbol{A}\boldsymbol{C}, \\
(\boldsymbol{A}+\boldsymbol{B})\,\boldsymbol{C} &= \boldsymbol{A}\boldsymbol{C}+\boldsymbol{B}\boldsymbol{C}.
\end{aligned}$$

- Note:
$$\boldsymbol{A}\boldsymbol{B} \neq \boldsymbol{B}\boldsymbol{A}.$$

- Can you find easy examples of this?

- Transpose:

  The transpose of $\boldsymbol{A}$ denoted by $\boldsymbol{A}^\top$ is defined as:

$$a_{ij}^T = a_{ji}.$$

  In other words, we obtain $\boldsymbol{A}^\top$ from $\boldsymbol{A}$ by turning row i into column i (and therefore column j into row j).

  Rules for transpose:

$$\begin{aligned}(\boldsymbol{A}+\boldsymbol{B})^\top &= \boldsymbol{A}^\top + \boldsymbol{B}^\top, \\ \left(\boldsymbol{A}^\top\right)^\top &= \boldsymbol{A}, \\ (\boldsymbol{A}\boldsymbol{B})^\top &= \boldsymbol{B}^\top \boldsymbol{A}^\top.\end{aligned}$$

Special matrices

- A square matrix has the same number of rows and columns.
- Column matrix is a column vector, i.e. it has m rows and a single column.
- Unit column vector $\mathbf{e}_i$: $e_j = 0$ if $j \neq i$ ja $e_j = 1$ if $j = i$.
- Row matrix is a row vector. It has a single row and n columns.
- Diagonal matrix Λ is a square matrix such that $\lambda_{ij} = 0$ if $i \neq j$.

$$\begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \lambda_n \end{pmatrix}.$$

- Identity matrix $\boldsymbol{I}$ (the multiplicative unit) is a diagonal matrix with $\lambda_{ii} = 1$:

$$\boldsymbol{I} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

- Upper triangular matrix $\boldsymbol{A}$ is a square matrix such that $a_{ij} = 0$ if $i > j$:

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ 0 & \ddots & \vdots \\ 0 & 0 & a_{nn} \end{pmatrix}.$$

- Lower triangular matrix $\boldsymbol{A}$ is a square matrix such that $a_{ij} = 0$ if $i < j$:

$$\begin{pmatrix} a_{11} & 0 & 0 \\ \vdots & \ddots & 0 \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}.$$

- A symmetric matrix is $\boldsymbol{A}$ a square matrix such that

$$\boldsymbol{A} = \boldsymbol{A}^{\top}.$$

- Permutation matrix is a matrix with zeros and ones as elements. Each row and each column has a single one. Permutation matrices are obtained from the unit matrix by interchanging (permuting) rows. For example with $n = 3$ we get

$$\boldsymbol{E}_{23} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

by permuting the last two rows of the identity matrix.

- Elementary row operations can be represented as results of matrix multiplication as follows. Let $\boldsymbol{E}_{ij}$ be a permutation matrix where rows i and j have been permuted. Permuting the rows i and j of A can be written as matrix product:

$$\boldsymbol{E}_{ij}\boldsymbol{A}.$$

Let $\boldsymbol{E}_i\,(r)$ be the matrix obtained by multiplying row i of the identity matrix by scalar r.

$$\boldsymbol{E}_2\,(r) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & r & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Multiplying the i^{th} row of $\boldsymbol{A}$ corresponds to the product

$$\boldsymbol{E}_i\left(r\right)\boldsymbol{A}.$$

Let $\boldsymbol{E}_{ij}\left(r\right)$ be a matrix obtained by adding to the identity matrix a matrix whose element ji is r and all other elements are zeros.

$$\boldsymbol{E}_{23}\left(r\right)=\left(\begin{array}{ccc}1&0&0\\0&1&0\\0&r&1\end{array}\right).$$

Adding row i of $\boldsymbol{A}$ multiplied by r to row j is obtained by:

$$\boldsymbol{E}_{ij}\left(r\right)\boldsymbol{A}.$$

Hence we have shown that the elementary operations can be performed as matrix multiplications by elementary matrices $\boldsymbol{E}_{ij}, \boldsymbol{E}_i(r), \boldsymbol{E}_{ij}(r)$.

Inverting a matrix

Consider square matrices $\boldsymbol{A}$ with n columns and rows. The inverse matrix of $\boldsymbol{A}$ is denoted by $\boldsymbol{A}^{-1}$. For the inverse matrix, we have:

$$\boldsymbol{A}\boldsymbol{A}^{-1}=\boldsymbol{I}.$$

Recall from the previous section that he system of equations

$$\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$$

has a unique solution for all $\boldsymbol{b}$ if $rank(\boldsymbol{A})=n.$

Solve the systems of equations

$$\boldsymbol{A}\boldsymbol{x}=\mathbf{e}_i$$

for all$i=1,...,n,$ and denote the solutions by $\mathbf{x}_i$. In other words,

$$\boldsymbol{A}\mathbf{x}_i=\mathbf{e}_i$$

for all $i.$

By the definition of matrix multiplication, we have:

$$\boldsymbol{A}^{-1}=\left(\mathbf{x}_1,...,\mathbf{x}_n\right).$$

As a result, we see that we can find the inverse matrix via elementary row operations for the augmented matrix.

$$(\boldsymbol{A}\,|\boldsymbol{I}\,)\,.$$

Example:

$$\boldsymbol{A} = \begin{pmatrix} 2 & 3 & 1 \\ 0 & 2 & 1 \\ 1 & 0 & 1 \end{pmatrix}$$

$$(\boldsymbol{A}|\boldsymbol{I}) = \left(\begin{array}{ccc|ccc} 2 & 3 & 1 & 1 & 0 & 0 \\ 0 & 2 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \end{array}\right).$$

Eliminate the first element on the third row with the first row:

$$\left(\begin{array}{ccc|ccc} 2 & 3 & 1 & 1 & 0 & 0 \\ 0 & 2 & 1 & 0 & 1 & 0 \\ 0 & -\frac{3}{2} & \frac{1}{2} & -\frac{1}{2} & 0 & 1 \end{array}\right),$$

and the second element using the second row:

$$\left(\begin{array}{ccc|ccc} 2 & 3 & 1 & 1 & 0 & 0 \\ 0 & 2 & 1 & 0 & 1 & 0 \\ 0 & 0 & \frac{5}{4} & -\frac{1}{2} & \frac{3}{4} & 1 \end{array}\right).$$

Multiply the third row by$\frac{4}{5}$

$$\left(\begin{array}{ccc|ccc} 2 & 3 & 1 & 1 & 0 & 0 \\ 0 & 2 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & -\frac{2}{5} & \frac{3}{5} & \frac{4}{5} \end{array}\right),$$

Add third row multiplied by -1 to second and first row:

$$\left(\begin{array}{ccc|ccc} 2 & 3 & 0 & \frac{7}{5} & -\frac{3}{5} & -\frac{4}{5} \\ 0 & 2 & 0 & \frac{2}{5} & \frac{2}{5} & -\frac{4}{5} \\ 0 & 0 & 1 & -\frac{2}{5} & \frac{3}{5} & \frac{4}{5} \end{array}\right)$$

Divide second row by 2:

$$\left(\begin{array}{ccc|ccc} 2 & 3 & 0 & \frac{7}{5} & -\frac{3}{5} & -\frac{4}{5} \\ 0 & 1 & 0 & \frac{1}{5} & \frac{1}{5} & -\frac{2}{5} \\ 0 & 0 & 1 & -\frac{2}{5} & \frac{3}{5} & \frac{4}{5} \end{array}\right),$$

Multiply second row by -3 and add to first:

$$\left(\begin{array}{ccc|ccc} 2 & 0 & 0 & \frac{4}{5} & -\frac{6}{5} & \frac{2}{5} \\ 0 & 1 & 0 & \frac{1}{5} & \frac{1}{5} & -\frac{2}{5} \\ 0 & 0 & 1 & -\frac{2}{5} & \frac{3}{5} & \frac{4}{5} \end{array}\right).$$

Finally divide first row by 2:

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{2}{5} & -\frac{3}{5} & \frac{1}{5} \\ 0 & 1 & 0 & \frac{1}{5} & \frac{1}{5} & -\frac{2}{5} \\ 0 & 0 & 1 & -\frac{2}{5} & \frac{3}{5} & \frac{4}{5} \end{array}\right).$$

We obtain:

$$\boldsymbol{A}^{-1} = \begin{pmatrix} \frac{2}{5} & -\frac{3}{5} & \frac{1}{5} \\ \frac{1}{5} & \frac{1}{5} & -\frac{2}{5} \\ -\frac{2}{5} & \frac{3}{5} & \frac{4}{5} \end{pmatrix}.$$

To check the result:

$$\boldsymbol{A}\boldsymbol{A}^{-1} = \begin{pmatrix} 2 & 3 & 1 \\ 0 & 2 & 1 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} \frac{2}{5} & -\frac{3}{5} & \frac{1}{5} \\ \frac{1}{5} & \frac{1}{5} & -\frac{2}{5} \\ -\frac{2}{5} & \frac{3}{5} & \frac{4}{5} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Rules for inverse matrices:

$$\left(\boldsymbol{A}^{-1}\right)^{-1} = \boldsymbol{A},$$
$$\left(\boldsymbol{A}^{\top}\right)^{-1} = \left(\boldsymbol{A}^{-1}\right)^{\top},$$

If $\boldsymbol{A}$ and $\boldsymbol{B}$ have inverse matrices:

$$\left(\boldsymbol{A}\boldsymbol{B}\right)^{-1} = \boldsymbol{B}^{-1}\boldsymbol{A}^{-1}.$$

Determinant

Consider $n \times n$ square matrix $\boldsymbol{A}$. For $n = 1$ define the determinant as $\det \boldsymbol{A} = a_{11}$.

For a general $n \times n$ matrix $\boldsymbol{A}$ and remove row i and and column j to get an $(n-1) \times (n-1)$ matrix $\boldsymbol{A}_{ij}$. Let

$$M_{ij} = \det \boldsymbol{A}_{ij}.$$

Matrix $\boldsymbol{A}$ (i, j) has a cofactor C_{ij} defined as:

$$C_{ij} \equiv (-1)^{i+j} M_{ij}.$$

The determinant of $\boldsymbol{A}$ is defined recursively as:

$$\det \boldsymbol{A} = \Sigma_{j=1}^{n} (-1)^{(i+j)} a_{ij} C_{ij}.$$

(Where is the recursion in the previous formula?)

The determinant can also be computed by expanding similarly along a column:

$$\det \boldsymbol{A} = \Sigma_{j=1}^{n} (-1)^{(i+j)} a_{ij} C_{ij}.$$

Example 4.

$$\boldsymbol{A} = \begin{pmatrix} 2 & 3 & 1 \\ 0 & 2 & 1 \\ 1 & 0 & 1 \end{pmatrix}.$$

$$\det \boldsymbol{A} = 2 \det \begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix} - 0 \det \begin{pmatrix} 3 & 1 \\ 0 & 1 \end{pmatrix}$$

$$+1 \det \begin{pmatrix} 3 & 1 \\ 2 & 1 \end{pmatrix} = 4 + 1 = 5.$$

Example 5.

$$\det \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ 0 & \ddots & \vdots \\ 0 & 0 & a_{nn} \end{pmatrix} = a_{11} \cdot a_{22} \cdot \ldots \cdot a_{nn}.$$

Proposition 1. i) The determinant is zero if and only if the matrix does not have full rank.

ii) Swapping rows changes the sign of the determinant.

iii) Adding (scalar multiples) of rows does not change the determinant.

Proof. (Sketch) The first point results from ii) and iii) since elementary operations can only change the sign of the determinant. To see the second point, show that this is true for 2x2 matrices and therefore for all matrices (by expanding along rows that were not swapped). For the third, compute the determinant for

$$\boldsymbol{A}' = \begin{pmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{j1} + ra_{i1} & \cdots & a_{jj} + ra_{ij} & \cdots & a_{jn} + ra_{in} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nj} & \cdots & a_{nn} \end{pmatrix}$$

Expand via row j to get:

$$\begin{aligned}\det \boldsymbol{A}' &= r\Sigma_{k=1}^{n} a_{jk}C_{jk} + r\Sigma_{k=1}^{n} a_{ik}C_{ik} \\ &= \det \boldsymbol{A} + r \det \boldsymbol{B},\end{aligned}$$

where

$$\boldsymbol{B} = \begin{pmatrix} a_{11} & \cdots & a_{ij} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{i1} & & a_{ij} & & a_{in} \\ \vdots & & \vdots & & \vdots \\ a_{i1} & \cdots & a_{ij} & & a_{in} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & a_{nj} & \cdots & a_{nn} \end{pmatrix}.$$

Matrix $\boldsymbol{B}$ has the i^{th} row of A as both row i and row j. Since the determinant can be developed along any row, i and j can always be left as the last two to be eliminated. For 2×2 matrices one sees immediately that the determinant is zero if the rows are identical. $\square$

Rules for computing the determinant:

$$\begin{aligned}\det \boldsymbol{A}^{\top} &= \det \boldsymbol{A}. \\ \det \boldsymbol{AB} &= \det \boldsymbol{A} \det \boldsymbol{B}, \\ \det \boldsymbol{A}^{-1} &= \frac{1}{\det \boldsymbol{A}}, \\ \det \boldsymbol{A} + \boldsymbol{B} &\neq \det \boldsymbol{A} + \det \boldsymbol{B} \text{ in general.}\end{aligned}$$

Cramer's rule

Assume that $\boldsymbol{A}$ has full rank and therefore $\det A \neq 0$). The system of equations

$$\boldsymbol{Ax} = \boldsymbol{b}$$

has then a unique solution

$$\boldsymbol{x}_i = \frac{\det \boldsymbol{B}_i}{\det \boldsymbol{A}},$$

where $\boldsymbol{B}_i$ is the matrix obtained by replacing the i^{th} column of $\boldsymbol{A}$ by the column vector $\boldsymbol{b}$.

Example 6.

$$\begin{pmatrix} 2 & 3 & 1 \\ 0 & 2 & 1 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}.$$

$$x_1 = \frac{\det \begin{pmatrix} 2 & 3 & 1 \\ 1 & 2 & 1 \\ 0 & 0 & 1 \end{pmatrix}}{\det \begin{pmatrix} 2 & 3 & 1 \\ 0 & 2 & 1 \\ 1 & 0 & 1 \end{pmatrix}} = \frac{1}{5},$$

$$x_2 = \frac{\det \begin{pmatrix} 2 & 2 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}}{5} = \frac{3}{5},$$

$$x_3 = \frac{\det \begin{pmatrix} 2 & 3 & 2 \\ 0 & 2 & 1 \\ 1 & 0 & 0 \end{pmatrix}}{5} = \frac{-1}{5},$$

Inverting a matrix with determinants

Cofactor matrix of $\boldsymbol{A}$ is given by:

$$\boldsymbol{C} = (C_{ij}),$$

where the cofactors are as above. Its transpose $\boldsymbol{C}^\top$ is called the adjungated matrix of $\boldsymbol{A}$ (or the classic adjoint) $adj(\boldsymbol{A})$:

$$adj(\boldsymbol{A}) = \boldsymbol{C}^\top.$$

Then:

$$\boldsymbol{A}^{-1} = \frac{1}{\det \boldsymbol{A}} \cdot adj(\boldsymbol{A}).$$

Example 7. Compute $adj(\boldsymbol{A})$, for

$$\boldsymbol{A} = \begin{pmatrix} 2 & 3 & 1 \\ 0 & 2 & 1 \\ 1 & 0 & 1 \end{pmatrix}.$$

$$\begin{aligned} C_{11} &= 2,\ C_{12} = 1\ , C_{13} = -2, \\ C_{21} &= -3,\ C_{22} = 1,\ C_{23} = 3, \\ C_{31} &= 1,\ C_{32} = -1,\ C_{33} = 4. \end{aligned}$$

$$adj\,(\boldsymbol{A}) = \begin{pmatrix} 2 & -3 & 1 \\ 1 & 1 & -2 \\ -2 & 3 & 4 \end{pmatrix},$$

Therefore

$$\boldsymbol{A}^{-1} = \frac{1}{\det \boldsymbol{A}} \cdot \begin{pmatrix} 2 & -3 & 1 \\ 1 & 1 & -2 \\ -2 & 3 & 4 \end{pmatrix},$$

which corresponds to what we computed before since $\det \boldsymbol{A} = 5$.

Dominant diagonal matrices

For many economic models, it is important that the solutions are non-negative (for example, prices, consumptions etc. must be positive). The rank condition is not enough to tell us when the solution is positive. In this subsection, we see a sufficient condition for positive solutions. (In more advanced courses, you will see more sophisticated analysis of this using Farkas' Lemma and Separating Hyperplane Theorem.)

A matrix $\boldsymbol{B}$ given by:

$$\boldsymbol{B} = \begin{pmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & \ddots & \vdots \\ b_{n1} & \cdots & b_{nn} \end{pmatrix}.$$

is a dominant diagonal matrix if

1. $b_{ii} > 0$ for all i.

2. $b_{ij} \leq 0$ for all j.

3. For all j:
$$b_{jj} + \Sigma_{i \neq j} b_{ij} \geq 0.$$

Consider the row echelon form of dominant diagonal matrices. Eliminate the first elements on other rows by using the first row. This gives:

$$\begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1j} & \cdots & \cdots & b_{1n} \\ 0 & b_{22} - \frac{b_{21}}{b_{11}} b_{12} & \cdots & b_{2j} - \frac{b_{21}}{b_{11}} b_{1j} & \cdots & \cdots & b_{2n} - \frac{b_{21}}{b_{11}} b_{1n} \\ \vdots & & & \vdots & & & \vdots \\ \vdots & b_{j2} - \frac{b_{j1}}{b_{11}} b_{12} & \cdots & b_{jj} - \frac{b_{j1}}{b_{11}} b_{1j} & \cdots & \cdots & b_{jn} - \frac{b_{j1}}{b_{11}} b_{1n} \\ \vdots & \vdots & & \vdots & & & \vdots \\ 0 & b_{n2} - \frac{b_{n1}}{b_{11}} b_{12} & & b_{nj} - \frac{b_{n1}}{b_{11}} b_{1j} & & & b_{nn} - \frac{b_{n1}}{b_{11}} b_{1n} \end{pmatrix}.$$

Consider $(n-1) \times (n-1)$ partial matrix $\widehat{\boldsymbol{B}}$:

$$\begin{aligned} \widehat{\boldsymbol{B}} &= \begin{pmatrix} \widehat{b}_{22} & \cdots & \widehat{b}_{2n} \\ \vdots & \ddots & \vdots \\ \widehat{b}_{n2} & \cdots & \widehat{b}_{nn} \end{pmatrix} = \\ & \begin{pmatrix} b_{22} - \frac{b_{21}}{b_{11}} b_{12} & \cdots & b_{2j} - \frac{b_{21}}{b_{11}} b_{1j} & \cdots & \cdots & b_{2n} - \frac{b_{21}}{b_{11}} b_{1n} \\ & & \vdots & & & \vdots \\ b_{i2} - \frac{b_{j1}}{b_{11}} b_{12} & \cdots & b_{jj} - \frac{b_{j1}}{b_{11}} b_{1j} & \cdots & \cdots & b_{in} - \frac{b_{j1}}{b_{11}} b_{1n} \\ \vdots & & \vdots & & & \vdots \\ b_{n2} - \frac{b_{n1}}{b_{11}} b_{12} & & b_{nj} - \frac{b_{n1}}{b_{11}} b_{1j} & & & b_{nn} - \frac{b_{n1}}{b_{11}} b_{1n} \end{pmatrix}. \end{aligned}$$

Lemma 1. If $\boldsymbol{B}$ is a dominant diagonal matrix, then $\widehat{\boldsymbol{B}}$ is also dominant diagonal matrix.

Proof. We check that $\widehat{\boldsymbol{B}}$ satisfies the requirements for a dominant diagonal matrix.

1. $\widehat{b}_{jj} = b_{jj} - \frac{b_{j1}}{b_{11}} b_{1j} > b_{jj} + b_{1j} > 0$, since $b_{jj} > \Sigma_{i \neq j} b_{ij}$.

2. $\widehat{b}_{ij} = b_{ij} - \frac{b_{i1}}{b_{11}} b_{1j} \leq 0$ for $i \neq j$.

3.

$$\begin{aligned} \widehat{b}_{2j} + \widehat{b}_{3j} + ... + \widehat{b}_{nj} &= \Sigma_{i \neq 1} \left(b_{ij} - \frac{b_{i1}}{b_{11}} b_{1j} \right) \\ &= \Sigma_{i \neq 1} b_{ij} - \frac{\Sigma_{i \neq 1} b_{i1}}{b_{11}} b_{1j} \\ &> \Sigma_i b_{ij} \\ &> 0. \end{aligned}$$

Repeat this elimination step $j-1$ times to get

$$\begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1j} & \cdots & \cdots & b_{1n} \\ 0 & b_{22} - \frac{b_{21}}{b_{11}} b_{12} & \cdots & b_{2j} - \frac{b_{21}}{b_{11}} b_{1j} & \cdots & \cdots & b_{2n} - \frac{b_{21}}{b_{11}} b_{1n} \\ \vdots & 0 & + & \cdots & \cdots & \cdots & \cdots \\ \vdots & 0 & \cdots & \widehat{b}_{jj} & \cdots & \cdots & \widehat{b}_{jn} \\ \vdots & \vdots & 0 & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \widehat{b}_{nj} & \cdots & \cdots & \widehat{b}_{nn} \end{pmatrix}.$$

□

The same proof as above shows inductively that after $(j-1)$ eliminations, the matrix

$$\widehat{B}^{(j-1)} = \begin{pmatrix} \widehat{b}_{jj} & \cdots & \widehat{b}_{n2} \\ \vdots & \ddots & \vdots \\ \widehat{b}_{2n} & \cdots & \widehat{b}_{nn} \end{pmatrix}$$

is dominant diagonal.

With this lemma, we can prove the following proposition.

Proposition 2. If $\boldsymbol{B}$ is a dominant diagonal matrix, then

$$\boldsymbol{Bx} = \boldsymbol{y}$$

has a unique solution for all $\boldsymbol{y}$. If $\boldsymbol{y} \geq 0$, then the solution $\boldsymbol{x}$ is non-negative:

$$\boldsymbol{x} \geq 0.$$

Proof. For the first claim, notice that by the observation after the previous lemma, $\boldsymbol{B}$ has full rank and therefore, the equation has a unique solution.

For the second claim, note that after $(n-1)$ elementary row operations, we have a matrix with the following pattern of signs:

$$\begin{pmatrix} + & - & - \\ 0 & \ddots & - \\ 0 & 0 & + \end{pmatrix}.$$

Row n yields:

$$\lambda_{nn} x_n = y_n \Leftrightarrow x_n = \frac{y_n}{\lambda_{nn}}.$$

Since $\lambda_{nn} > 0$, $x_n > 0$ if $y_n > 0$.

Row $n-1$ gives:

$$\lambda_{n-1n-1}x_{n-1} + \lambda_{n-1n}x_n = y_{n-1}$$

or

$$x_{n-1} = \frac{y_{n-1} - \lambda_{n-1n}x_n}{\lambda_{n-1n-1}}.$$

Since $x_n \geq 0$, $\lambda_{n-1n} \leq 0$ and $\lambda_{n-1n-1} > 0$, we get $x_{n-1} \geq 0$ if $y_{n-1} \geq 0$.

By substituting backwards, we get the result. $\square$

Linear models in economics

1. **Input-output -tables**

   Consider an economy producing n goods. All goods are final goods and potentially intermediate goods. The production of all goods happens simultaneously. Assume linear production in the sense that to produce x_i units of good i we need $a_{ji}x_i$ units of good j. If the economy produces a net output $(y_1, ..., y_n)$ the total output $(x_1, ..., x_n)$ can be computed as

$$\begin{aligned} x_1 - a_{11}x_1 - a_{12}x_2 - \ldots - a_{1n}x_n &= y_1, \\ x_2 - a_{21}x_1 - a_{22}x_2 - \ldots - a_{2n}x_n &= y_2, \\ &\vdots \\ x_n - a_{n1}x_1 - a_{n2}x_2 - \ldots - a_{nn}x_n &= y_n. \end{aligned}$$

   In vector notation:

$$\begin{pmatrix} 1-a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & 1-a_{22} & & -a_{2n} \\ \vdots & & \ddots & \vdots \\ -a_{n1} & -a_{n2} & \cdots & 1-a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}.$$

   What are the feasible net productions $\boldsymbol{y} = (y_1, ..., y_n)$?

   By the previous section, we see that if $(\boldsymbol{I} - \boldsymbol{A})$ is a dominant diagonal matrix, i.e. if for all i

$$\sum_j a_{ji} < 1,$$

then all net outputs are possible In other words, for all $y \geq 0$ there exists a $\boldsymbol{x} \geq 0$ such that

$$(\boldsymbol{I} - \boldsymbol{A})\,\boldsymbol{x} = \boldsymbol{y}.$$

2. **Equilibrium in oligopoly models**

   In intermediate microeconomics you will see the Cournot model of oligopolistic competition with constant marginal costs. In the model, n firms choose optimal level of production q_i taking into account their own impact on the price.

$$P(q_1, ..., q_n) = \alpha - \beta \sum_{i=1}^{n} q_i.$$

   The optimal output depends on own marginal costs c_i, and the output of others:

$$q_i = \frac{a - \beta \Sigma_{j \neq i} q_j - c_i}{2\beta}.$$

   Since the quantities should be positive, we see immediately that the model makes sense only if

$$c_i \leq a \text{ for all } i,$$

   and if the c_i are relatively close to each other.

   In matrix form:

$$\begin{pmatrix} 2\beta & \beta & \cdots & \beta \\ \beta & 2\beta & & \beta \\ \vdots & & \ddots & \vdots \\ \beta & \beta & \cdots & 2\beta \end{pmatrix} \begin{pmatrix} q_1 \\ q_2 \\ \vdots \\ q_n \end{pmatrix} = \begin{pmatrix} a - c_1 \\ a - c_2 \\ \vdots \\ a - c_n \end{pmatrix}.$$

   Perform elementary row operations on the augmented matrix

$$\left(\begin{array}{cccc|c} 2\beta & \beta & \cdots & \beta & a - c_1 \\ \beta & 2\beta & & \beta & a - c_2 \\ \vdots & & \ddots & \vdots & \vdots \\ \beta & \beta & \cdots & 2\beta & a - c_n \end{array}\right).$$

Subtracting the first row from all other rows we get:

$$\left(\begin{array}{cccc|c} 2\beta & \beta & \cdots & \beta & a-c_1 \\ -\beta & \beta & & 0 & c_1-c_2 \\ \vdots & & \ddots & \vdots & \vdots \\ -\beta & 0 & \cdots & \beta & c_1-c_n \end{array}\right).$$

We get

$$\beta q_j = \beta q_1 + c_1 - c_j. \tag{2}$$

The first equation gives:

$$2\beta q_1 + \beta \Sigma_{j \neq i} q_j = a - c_1.$$

Substituting from the first to the second,

$$2\beta q_1 + (n-1)\,\beta q_1 + (n-1)\,c_1 - \Sigma_{j \neq i} c_j = a - c_1.$$

Solving for q_1, we get:

$$\begin{aligned} (n+1)\,\beta q_1 &= a + \Sigma_{j \neq i} c_j - n c_1, \\ q_1 &= \frac{a - c_1 + \Sigma_{j \neq i}\,(c_j - c_1)}{(n+1)\,\beta}. \end{aligned}$$

The other outputs are computed from (2).

3. **Market Equilibrium**

We start by an analysis of the equilibrium determination of prices and quantities for two products. The demand Q_i^d for each good i depends on the prices of the two goods P_1 and P_2, on disposable income Y and on other factors K_i.

Assume that the demands take the following form:

$$Q_1^d = K_1 P_1^{\alpha_{11}} P_2^{\alpha_{12}} Y^{\beta_1},$$

$$Q_2^d = K_1 P_1^{\alpha_{21}} P_2^{\alpha_{22}} Y^{\beta_2}.$$

Exercise: How would you interpret the parameters α_{ij} ja β_i? How large is the percentage change in the demand for i if we have a small percentage change in P_i, P_j or Y? What do the signs of the parameters tell us?

Since we are writing the model to analyze price formation, we would take the Q_i^d and the P_i to be endogenous variables to be determined by the model and K_i would summarize the exogenous variables (i.e. ones not determined in the model).

The supplies Q_i^s for the two products are assumed to take the form:

$$Q_1^s = M_1 P_1^{\gamma_1},$$

$$Q_2^s = M_2 P_2^{\gamma_2}.$$

Again, we take the variables M_i to be exogenous to the model.

Exercise: What is the interpretation for γ_i and what do you think about their sign? Comment on the implicit assumption that Q_i^s does not depend on P_j.

In equilibrium, supply equals demand so that

$$Q_1^d = Q_1^s,$$

and

$$Q_i^d = Q_i^s.$$

So we have six equations for six endogenous variables $(Q_i^s, Q_i^d, P_i)_{i=1,2}$. Unfortunately this system seems rather complicated since the equations contain products and powers of endogenous variables.

A simple change of variables reduces the complexity. Define the following new variables:

$$q_i^d = \ln Q_i^d,\ q_i^s = \ln Q_i^s\, p_i = \ln P_i,\ y_i = \ln Y_i,\ m_i = \ln M_i,\ k_i = \ln K_i,\ i \in \{1,2\}$$

By taking logarithms on both sides of each equation, we can write the six equations for $i \in \{1,2\}$:

$$q_i^d = k_i + \alpha_{ii} p_i + \alpha_{ij} p_j + \beta_i y,$$

$$q_i^s = m_i + \gamma_i p_i,$$

$$q_i^s = q_i^d.$$

By the third equation, $q_i^d = q_i^s$ for $i \in \{1, 2\}$, and therefore the right hand sides in the first and the second equations are equalized:

$$k_i + \alpha_{ii} p_i + \alpha_{ij} p_j + \beta_i y = m_i + \gamma_i p_i, \; i \in \{1, 2\}.$$

In a partial equilibrium model, the income of the consumers is assumed to be determined outside the model, i.e. it is an exogenous variable. Therefore the only endogenous variables in this model are p_1 ja p_2.

Let's write the exogenous variables on the right-hand side and the endogenous variables on the left-hand side:

$$\begin{array}{ccl} (\alpha_{11} - \gamma_1)\, p_1 & +\alpha_{12} p_2 & = m_1 - k_1 - \beta_1 y, \\ \alpha_{21} p_1 & (\alpha_{22} - \gamma_2)\, p_2 & = m_2 - k_2 - \beta_2 y. \end{array}$$

Or in matrix form:

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} p_1 \\ p_2 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix},$$

where $a_{ii} = \alpha_{ii} - \gamma_i$, $a_{ij} = \alpha_{ij}$, ja $b_i = m_i - k_i - \beta_i y$.

It is now straightforward to compute the equilibrium $\boldsymbol{p} = (p_1, p_2)$. I leave it as an exercise to use Cramer's rule to get the result:

$$p_1 = \frac{(\alpha_{22} - \gamma_2)(m_1 - k_1 - \beta_1 y) - \alpha_{12}(m_2 - k_2 - \beta_2 y)}{(\alpha_{22} - \gamma_2)(\alpha_{11} - \gamma_1) - \alpha_{12}\alpha_{21}},$$

$$p_2 = \frac{(\alpha_{11} - \gamma_1)(m_2 - k_2 - \beta_2 y) - \alpha_{21}(m_1 - k_1 - \beta_1 y)}{(\alpha_{22} - \gamma_2)(\alpha_{11} - \gamma_1) - \alpha_{12}\alpha_{21}}.$$

The (logarithmic) equilibrium quantities are solved most easily from the supply curves. Finally P_i, Q_i are solved by exponentiating p_i, q_i.

What conditions do you need on the parameters of the model for positive solutions?

Norm and inner product (dot product)

There are two ways to think of the product of $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ if we treat vectors as special matrices. The first, called inner product or dot product is defined as:

$$\boldsymbol{x} \cdot \boldsymbol{y} := \boldsymbol{x}^\top \boldsymbol{x} = \sum_{i=1}^{n} x_i y_i,$$

and the result is a real number.

The second, called the cross product results in an $n \times n$ matrix $\boldsymbol{x}\boldsymbol{x}^\top$. We will use only the first in this course (and it is by far the more important).

With the help the inner product, we can define the length or the norm of a vector:

$$\|\boldsymbol{x}\| := \sqrt{\boldsymbol{x} \cdot \boldsymbol{x}}.$$

This is essentially just the equivalent of the Pythagorean Theorem in higher dimensions. The distance between vectors $\boldsymbol{x}, \boldsymbol{y}$ is just the norm of $(\boldsymbol{x} - \boldsymbol{y})$.

The projection of a vector $\boldsymbol{y}$ on $\boldsymbol{x}$ is defined as the point $t^*\boldsymbol{x}$ on the line $t\boldsymbol{x}$ for $t \in \mathbb{R}$ such that

$$(\boldsymbol{y} - t^*\boldsymbol{x}) \cdot \boldsymbol{x} = 0.$$

This gives an explicit formula for t^*:

$$t^* = \frac{\boldsymbol{x} \cdot \boldsymbol{y}}{\|\boldsymbol{x}\|}.$$

Hence the projection $P_{\boldsymbol{x}}(\boldsymbol{y})$ is given by:

$$P_{\boldsymbol{x}}(\boldsymbol{y}) = \boldsymbol{x}\frac{\boldsymbol{x} \cdot \boldsymbol{y}}{\|\boldsymbol{x}\|}.$$

By basic trigonometry, the angle θ between $\boldsymbol{x}$ and $\boldsymbol{y}$ satisfies:

$$\cos(\theta) = \frac{\|P_{\boldsymbol{x}}(\boldsymbol{y})\|}{\|\boldsymbol{y}\|} = \frac{\boldsymbol{x} \cdot \boldsymbol{y}}{\|\boldsymbol{x}\|\|\boldsymbol{y}\|}.$$

Since $-1 \leq \cos(\theta) \leq 1$ for all θ, we get Cauchy's inequality for all vectors $\boldsymbol{x}, \boldsymbol{y}$:

$$|\boldsymbol{x} \cdot \boldsymbol{y}| \leq \|\boldsymbol{x}\|\|\boldsymbol{y}\|.$$

Linear functions

A function $f : \mathbb{R}^n \to \mathbb{R}^m$ is said to be *linear* if the following two conditions are satisfied:

i) (Homogeneity) For all $\lambda \in \mathbb{R}$ and for all $\boldsymbol{x} \in \mathbb{R}^n$, $f(\lambda \boldsymbol{x}) = \lambda f(\boldsymbol{x})$,

ii) (Additivity) For all $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$, $f(\boldsymbol{x}+\boldsymbol{y}) = f(\boldsymbol{x}) + f(\boldsymbol{y})$.

By taking $\lambda = 0$ in i), we see that $f(0) = 0$ for all linear functions. In the case of $n = 1$, this rules out functions whose graphs are straight lines that do not go through the origin. In this simplest setting, i) actually implies ii), and fixing $f(1)$ determines the entire function.

For $n > 1$, requirement ii) has bite. Observe that we can write $\boldsymbol{x} = \sum_{i=1}^n x_i \boldsymbol{e}^i$. By i), $f(x_i \boldsymbol{e}^i) = x_i f(\boldsymbol{e}^i)$ for all i, x_i. By ii),

$$f(\boldsymbol{x}) = f(\sum_{i=1}^n x_i \boldsymbol{e}^i) = \sum_{i=1}^n x_i f(\boldsymbol{e}^i).$$

Hence a linear function is completely determined by n values $f(\boldsymbol{e}^i)$. If $m = 1$, then $f(\boldsymbol{e}^i) \in \mathbb{R}$ for all i and letting $f(\boldsymbol{e}^i) = a_i$ we see that all real linear functions from $\mathbb{R}^n$ are given by inner products $\boldsymbol{a} \cdot \boldsymbol{x} = \sum_{i=1}^n a_i x_i$.

If $m > 1$, then each $f(\boldsymbol{e}^i)$ is an m -dimensional vector. If we denote $\boldsymbol{a}^i = f(\boldsymbol{e}^i) \in \mathbb{R}^m$, we have as before $f(x) = \sum_{i=1}^n x_i \boldsymbol{a}^i$. Writing $\boldsymbol{A} = [\boldsymbol{a}^1, ..., \boldsymbol{a}^n]$ for the matrix consisting of columns $\boldsymbol{a}^i$. But this means that

$$f(x) = \boldsymbol{A}\boldsymbol{x}.$$

Many of the properties of linear functions also extend to *affine* functions of the form

$$f(\boldsymbol{x}) = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{b},$$

for some $\boldsymbol{b} \in \mathbb{R}^m$. Actually this is not so bad because by a linear change of origin, $\hat{f}((\boldsymbol{x}) := f(\boldsymbol{x}) - f(0) = \boldsymbol{A}\boldsymbol{x}$ is linear.

Why are linear functions so much simpler than non-linear?

i) A change in $\boldsymbol{x}$ has the same effect regardless of the starting point:

$$f(\boldsymbol{x}) - f(\hat{\boldsymbol{x}}) = A(\boldsymbol{x} - \hat{\boldsymbol{x}}).$$

ii) A function $f : \mathbb{R}^n \to \mathbb{R}^m$ is said to be *surjective* (or *onto*) if for all $\boldsymbol{b} \in \mathbb{R}^m$, there is an $\boldsymbol{x} \in \mathbb{R}^n$ such that $f(\boldsymbol{x}) = \boldsymbol{b}$. f is said to be *injective* or *one-to-one* if for all $\boldsymbol{x} \neq \boldsymbol{x}', \quad f(\boldsymbol{x}) \neq f(\boldsymbol{x}')$. f is said to be *bijective* if it is injective and surjective. Bijective functions f have an *inverse function* $f^{-1} : \mathbb{R}^m \to \mathbb{R}^n$ such that $f^{-1}(f(\boldsymbol{x})) = \boldsymbol{x}$ and $f(f^{-1}(\boldsymbol{y})) = \boldsymbol{y}$.

In matrix algebra, we saw that if $f(\boldsymbol{x}) = \boldsymbol{A}\boldsymbol{x}$, then f is bijective if and only if $\boldsymbol{A}$ has full rank. Gaussian elimination (or the determinant) gives an easy way of determining when linear functions are bijective and computing the inverse function $f^{-1}(\boldsymbol{x}) = A^{-1}\boldsymbol{x}$.