# Mathematics for Economists: Basic Linear Algebra

Juuso Välimäki

Spring 2020

In these notes, we will cover the following topics:

- 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.

## 1 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, a vector $x$ is given by its coordinates along the two axis $x = (x_1, x_2)$ where $x_i$ is a real number for $i \in \{1, 2\}$. in three dimensional setting, $x = (x_1, x_2, x_3)$. In the plane, vector $y = (y_1, y_2)$ is the same as vector $x = (x_1, x_2)$ if $y_1 = x_1$ and $y_2 = x_2$, and similarly for the three dimensional case.

We can define the addition of vectors by $x + y = (x_1 + y_1, x_2 + y_2)$ and the multiplication of vectors by real numbers $a \in \mathbb{R}$ by $ax = (ax_1, ax_2)$. Hence we see that sums of vectors and multiples of vectors are again vectors.

We can define vectors in a similar manner for any dimension $k$. A vector $x$ is a $k$ -dimensional vector if $x = (x_1, x_2, ..., x_k)$, where $x_i$ is a real number for all $i \in \{1, ..., k\}$. In this case, we write $x \in \mathbb{R}^k$. For any $x, y \in \mathbb{R}^k$, define $x + y = (x_1 + y_1, x_2 + y_2, ..., x_k + y_k)$ and for $a \in \mathbb{R}$, $ax = (ax_1, ax_2, ..., ax_k)$.

A matrix is an array of real numbers into rows and columns. A $m \times n$ -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 $A$ 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 $A$ and a vector $x \in \mathbb{R}^n$ 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 $A$ and a vector $x \in \mathbb{R}^n$ 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 $x \in \mathbb{R}^k$, we take $x$ to be a column vector. With this notation, we can start our analysis.

## 2 Example: 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 $\alpha_{ij}$ ja $\beta_i$? 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 $\gamma_i$ 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 $b_i = m_i - k_i - \beta_i y$.

In these notes, we analyze linear models. We shall see some economic examples of linear systems and we will find ways of solving linear systems of equations.

Let's return to our example and solve for $p_1$ from the top equation:

$$p_1 = \frac{m_1 - k_1 - \beta_1 y - \alpha_{12} p_2}{(\alpha_{11} - \gamma_1)}.$$

Substituting into the second equation gives:

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

Multiplying both sides by $(\alpha_{22} - \gamma_2)(\alpha_{11} - \gamma_1)$ gives:

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

and therefore:

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

Plugging into the first equation gives:

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

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

Our next task is to come up with a nice algorithmic way of handling linear equation systems with $n$ variables. As you can imagine, using the above method of direct substitutions is tedious and prone to errors. In the first part of these notes, we see how to go about essentially the same idea in a systematic way.

We start with a system of $m$ equations in $n$ variables:

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

We write this in matrix form as:

$$Ax = b, \tag{1}$$

where we denote:

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

A key question is when the system of equations admits a solution and how many solutions it may have. At this point, you should spend a minute to think about how the answer to this depends on $m$ and $n$. What if $m = n$? Think about some examples with $m = n = 2$.

# 3 Systems of linear equations

We start with the analysis of systems of $n$ equations in $n$ variables:

$$\begin{array}{ccccccc} a_{11}x_1 & +a_{12}x_2 & \cdots & +a_{1n}x_n & = & b_1, \\ a_{21}x_1 & +a_{22}x_2 & \cdots & +a_{2n}x_n & = & b_2, \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ a_{n1}x_1 & +a_{n2}x_2 & \cdots & +a_{nn}x_n & = & b_n. \end{array} \tag{2}$$

In Matrix form:

$$Ax = b,$$

where

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}, \; x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \; b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}.$$

The first equation of (2) gives:

$$x_1 = \frac{b_1}{a_{11}} - \frac{a_{12}x_2}{a_{11}} - ... - \frac{a_{1n}x_n}{a_{11}}.$$

By substituting into the other $(n-1)$ equations, we get:

$$\begin{array}{cccccc} (a_{22} - a_{21}\frac{a_{12}}{a_{11}})x_2 & \cdots & +\left(a_{2n} - a_{21}\frac{a_{1n}}{a_{11}}\right) x_n & = & b_2 - \frac{b_1}{a_{11}}, \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ (a_{n2} - a_{n1}\frac{a_{12}}{a_{11}})x_2 & \cdots & +\left(a_{nn} - a_{n1}\frac{a_{1n}}{a_{11}}\right) x_n & = & b_n - \frac{b_1}{a_{11}}. \end{array} \tag{3}$$

We can repeat the previous step for system (3), i.e. solve $x_2$ and substitute into the other equations. Since after each such step we are left with a system that has one fewer equation than in the previous step we have after $(n-1)$ steps a linear equation in a single variable:

$$cx_n = d.$$

This equation has a (unique) solution if $c \neq 0$. It has infinitely many solutions if $c = d = 0$, and no solutions if $c = 0$ and $d \neq 0$. Our task is therefore to examine when the last step has a solution.

Recall a few rules of basic arithmetics:

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:

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

3. By combining the first two items we see that the set of solutions to a system of equations does not change if a multiple of one equation is added to another equation.

4. The set of solutions to a system of equations does not depend on the order n which the equations are written

Elementary row operations are applications of these rules.

## 3.1 Solving systems of equations via elementary row operations

Consider the system of equations in matrix form

$$Ax = 0,$$

where

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix},\ x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

This system has always a trivial solution $x = (0, ..., 0)$, but we can ask if it has other solutions.

Start by considering matrix $A$. If $a_{11} = 0$, swap row 1 with row $k$, where $a_{k1} \neq 0$. if no such row exists,the vector $(x, 0, ..., 0)$ satisfies the system of equations for all $x$ and therefore the solution is not unique.

Assume next that for some $k$, $a_{k1} \neq 0$ and swap rows 1 and $k = 1$ if $a_{11} = 0$. Multiply the first row by$\frac{1}{a_{11}}$ :and add the multiplied first row to each row $k'$ multiplied by $-\frac{a_{k1}}{a_{11}}$. We get the following new matrix

$$A^{(1)} = \begin{pmatrix} 1 & \frac{a_{12}}{a_{11}} & \cdots & \frac{a_{1n}}{a_{11}} \\ 0 & a_{22} - a_{21}\frac{a_{12}}{a_{11}} & \cdots & a_{2n} - a_{21}\frac{a_{1n}}{a_{11}} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n2} - a_{n1}\frac{a_{12}}{a_{11}} & \cdots & a_{nn} - a_{n1}\frac{a_{1n}}{a_{11}} \end{pmatrix}$$

$$= \begin{pmatrix} 1 & a_{12}^{(1)} & \cdots & a_{1n}^{(1)} \\ 0 & a_{22}^{(1)} & \cdots & a_{2n}^{(1)} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n2}^{(1)} & \cdots & a_{nn}^{(1)} \end{pmatrix}.$$

If

$$a_{22} - a_{21}\frac{a_{12}}{a_{11}} \neq 0,$$

multiply the second row by

$$\frac{1}{a_{22} - a_{21}\frac{a_{12}}{a_{11}}},$$

and add the resulting second row multiplied by

$$-\frac{a_{k2} - a_{k1}\frac{a_{12}}{a_{11}}}{a_{22} - a_{21}\frac{a_{12}}{a_{11}}}$$

to each row $k > 2$. If

$$a_{22} - a_{21}\frac{a_{12}}{a_{11}} = 0,$$

swap row 2 and $k''$ such that

$$a_{k''2} - a_{k''1}\frac{a_{12}}{a_{11}} \neq 0$$

and proceed as before. If $a_{k2}^{(1)} = 0$ for all $k \geq 2$, multiply the second row of $A^{(1)}$ by

$$\frac{1}{a_{23}^{(1)}}.$$

(or swap the rows if $a_{23}^{(1)} = 0$) and proceed as before.

This results in a new matrix

$$A^{(2)} = \begin{pmatrix} 1 & \frac{a_{12}}{a_{11}} & \cdots & \frac{a_{1n}}{a_{11}} \\ 0 & 1 & \cdots & a_{2n} - a_{21}\frac{a_{1n}}{a_{11}} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} - a_{n1}\frac{a_{1n}}{a_{11}} \end{pmatrix}$$

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

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

$$\begin{pmatrix} 1 & a_{12}^{(1)} & \cdot & \cdot & a_{15}^{(1)} \\ 0 & 1 & a_{23}^{(2)} & \cdot & a_{25}^{(2)} \\ 0 & 0 & 1 & a_{34} & a_{35}^{(3)} \\ 0 & 0 & 0 & 1 & a_{45}^{(4)} \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

or

$$\begin{pmatrix} 1 & a_{12}^{(1)} & \cdot & \cdot & a_{15}^{(1)} \\ 0 & 1 & a_{23}^{(2)} & \cdot & a_{25}^{(2)} \\ 0 & 0 & 0 & 1 & a_{35}^{(3)} \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

We say that a matrix $A$ 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 number of non-zero rows 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. If the row rank of the matrix is less than its total number of rows, the system of equations has an infinite number of solutions. In the second case above, the variable $x_3$ can be chosen freely. For each choice of $x_3$, the other variables are uniquely determined. If row rank equals the number or rows, then $x = (0, ..., 0)$ is the only solution to the system of equations.

Consider next the system

$$Ax = b.$$

We will perform elementary row operations for the *augmented matrix* below:

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

We aim to transform matrix $A$ to its row echelon form using elementary row operations:

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

If

$$\operatorname{rank}\left(A \vdots b\right) = \operatorname{rank}(A),$$

then the system has a solution. If

$$\operatorname{rank}\left(A \vdots b\right) > \operatorname{rank}(A),$$

it has no solutions. If

$$\operatorname{rank}\left(A \vdots b\right) = \operatorname{rank}(A) = n,$$

the solution is unique. If

$$\operatorname{rank}\left(A \vdots b\right) = \operatorname{rank}(A) < n,$$

then the system has infinitely many solutions.

**Examples of elementary row operations**

- **Finding the row echelon form**
  - Let
  $$A = \begin{pmatrix} 2 & 1 & -1 \\ 1 & 2 & 2 \\ 1 & 0 & 1 \end{pmatrix}$$
  Multiply first row by $-\frac{1}{2}$ and add to second and third row:
  $$\begin{pmatrix} 2 & 1 & -1 \\ 0 & 2-\frac{1}{2} & 2+\frac{1}{2} \\ 0 & 0-\frac{1}{2} & 1+\frac{1}{2} \end{pmatrix} = \begin{pmatrix} 2 & 1 & -1 \\ 0 & \frac{3}{2} & \frac{5}{2} \\ 0 & -\frac{1}{2} & \frac{3}{2} \end{pmatrix}$$
  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 rank 3, we know that the system
  $$Ax = b$$
  has a unique solution for all $b$.

- **Solving a system of equations**

  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:

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

eli

$$x_1 = \frac{3}{7}.$$

- **Matrix without full rank**

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

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:

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

Matrix $A$ has $rank(A) = 2$ and the system of equations has either zero or infinitely many solutions.

$$Ax = b$$

I leave it as an exercise using the augmented matrix to show that the system has a solution only if

$$b_3 = b_2 + \frac{1}{2}b_1.$$

- **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:

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

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

$$rank\,(A) < n.$$

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

# 4 Matrix Algebra

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

- Matrix equality:

  $$A = B$$

  if for all $i, j$:

  $$a_{ij} = b_{ij}.$$

- Scalar multiplication:

  Let$r \in \mathbb{R}$. Define

  $$rA = C,$$

  where for all $i, j$:

  $$c_{ij} = ra_{ij}.$$

- Addition:

  $$A + B = C,$$

  where for all $i, j$:

  $$c_{ij} = a_{ij} + b_{ij}.$$

  This is defined only if the matrices are of the same dimension (same number of rows and columns).

- Matrix difference: (combining the two previous ones).

$$A - B = A + (-B) = C,$$

where for all $i, j$:

$$c_{ij} = a_{ij} - b_{ij}.$$

- Matrix multiplication: Let $A$ be an $m \times n$ -matrix and $B$ an $n \times k$ -matrix. The product of $A$ and $B$ is defined as:

$$AB = C,$$

where

$$c_{ij} = \sum_{h=1}^{n} a_{ih} b_{hj}.$$

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

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

- Why define multiplication like this? Why not element by element? Let

$$y = Ax \in \mathbb{R}^m.$$

This is a particular form of a function $y = f(x)$. Consider then

$$z = By \in \mathbb{R}^k.$$

This is a particular form of $z = g(y)$ If you write the composite function $z = g(f(x))$ for this case

$$z = Cx = BAx,$$

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.

- Some rules:

$$\begin{aligned}(A+B)+C &= A+(B+C),\\ (AB)C &= A(BC),\\ A+B &= B+A,\\ A(B+C) &= AB+AC,\\ (A+B)C &= AC+BC.\end{aligned}$$

- Note:

$$AB \neq BA.$$

- Can you find easy examples of this?

- Transpose:

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

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

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

  Rules for transpose:

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

**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 $\Lambda$ is a square matrix such that $\lambda_{ij} = 0$ if $i \neq j$.

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

- Unit matrix $I$ is a diagonal matrix with $\lambda_{ii} = 1$ :

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

- Upper triangular matrix $A$ 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 $A$ 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 symmetic matrix is $A$ a square matrix such that

$$A = A^\top.$$

- 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

$$E_{23} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

by permuting the last two rows of the identity matrix.

- Elementary row operations can be represented as results of matrix multiplication as follows. Let $E_{ij}$ 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:
$$E_{ij}A.$$
Let $E_i(r)$ be the matrix obtained by multiplying row $i$ of the unit matrix by scalar $r$.
$$E_2(r) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & r & 0 \\ 0 & 0 & 1 \end{pmatrix}$$
Multiplying the $i^{th}$ row of $A$ corresponds to the product
$$E_i(r)A.$$
Let $E_{ij}(r)$ be a matrix obtained by adding to the unit matrix a matrix whose element $ji$ is $r$ and all other elements are zeros.
$$E_{23}(r) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & r & 1 \end{pmatrix}.$$
Adding row $i$ multiplied by $r$ to row $j$ mis obtained by:
$$E_{ij}(r)A.$$
Hence we have shown that the elementary operations can be performed as matrix multiplications by elementary matrices $E_{ij}, E_i(r), E_{ij}(r)$.

**Inverting a matrix**

Consider square matrices $A$ with $n$ columns and rows. The inverse matrix of $A$ is denoted by $A^{-1}$. For the inverse matrix, we have:
$$AA^{-1} = I.$$
Recall from the previous section that he system of equations
$$Ax = b$$

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

Solve the systems of equations

$$Ax = \mathbf{e}_i$$

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

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

for all $i$.

By the definition of matrix multiplication, we have:

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

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

$$(A\,|I\,).$$

Example:

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

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

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:

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

To check the result:

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

Rules for inverse matrices:

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

If $A$ and $B$ have inverse matrices:

$$(AB)^{-1} = B^{-1}A^{-1}.$$

**Determinant**

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

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

$$M_{ij} = \det A_{ij}.$$

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

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

The determinant of $A$ is defined recursively as:

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

(Where is the recursion in the previous formula?)

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

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

Examples:

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

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

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

2.
$$\det \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ 0 & \ddots & \vdots \\ 0 & 0 & a_{nn} \end{pmatrix} = a_{11} \cdot a_{22} \cdot \ldots \cdot a_{nn}.$$

3. 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.

   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. For the third, compute the determinant for

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

   Expand via row $j$ to get:

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

   where

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

   Matrix $B$ 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 $2 \times 2$ matrices one sees immediately that the determinant is zero if the rows are identical.

4. Rules for computing the determinant:

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

**Cramer's rule**

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

$$Ax = b$$

has then a unique solution

$$x_i = \frac{\det B_i}{\det A},$$

where $B_i$ is the matrix obtained by replacing the $i^{th}$ column of $A$ by the column vector $b$.

Example:

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

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

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

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

**Inverting a matrix**
Cofactorimatrix of $A$ is given by:

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

where the cofactors are as above. The transpose of $C^\top$ is called the adjoint of $A$ $adj\,(A)$ :

$$adj\,(A) = C^\top.$$

Then:

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

Example: compute $adj(A)$, for

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

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

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

Therefore

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

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

**Extra material: Dominant diagonal matrices**
A matrix $B$

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

is a dominant diagonal matrix if

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

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

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

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 $\widehat{B}$ :

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

Claim: If $B$ is a dominant diagonal matrix, then $\widehat{B}$ is also dominant diagonal matrix.

Proof:

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

3.

$$\begin{aligned}\widehat{b}_{2j} + \widehat{b}_{3j} + ... + \widehat{b}_{nj} &= b_{jj} - \frac{b_{j1}}{b_{11}} b_{1j} - \Sigma_{i \neq 1,j} \left( b_{ij} + \frac{b_{i1}}{b_{11}} b_{1j} \right) \\ &= b_{jj} - \Sigma_{i \neq 1,j} b_{ij} - \frac{\Sigma_{i \neq 1} b_{i1}}{b_{11}} b_{1j} \\ &> b_{jj} - \Sigma_{i \neq 1} b_{ij} \\ &> 0.\end{aligned}$$

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

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

The same proof as above gives inductively that

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

is dominant diagonal.

This shows that for dominant diagonal $B$,

$$Bx = y$$

has a unique solution for all $y$ (since $rank\,(A) = n$) and also if $y \geq 0$, the solutions

$$x \geq 0,$$

since after the elementary row operations,

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

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 prove the result.

## 4.1 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 $y = (y_1, ..., y_n)$?

By the previous section, we see that if $(I - A)$ 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 $x \geq 0$ such that

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

2. **Equilirium 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. The optimal output depends on own marginal costs $c_i$, and the output of others:

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

   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

$$\beta q_j = \beta q_1 + c_1 - c_j. \tag{4}$$

The first equation gives:

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

or

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

Combining these gives:

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

The other outputs are easily computed from (4).

3. **Portfolio choice: teaser into finance**

   Consider a world of uncertainty. Let $s \in S$ be a description of the realized state of the world. Economic decision makers do not know which

   $$s_i \in S = \{s_1, ..., s_m\}$$

   is realized. Each $s_i$ is a full description of all the relevant aspects of the situation for the decision maker.

   Concrete examples of states of the world:

   - Fire insurance: $S = \{s_1, s_2\}$, where $s_1 =$ "house burns", $s_2 =$ "house does not burn".
   - Preparing for world after corona virus: $S = \{s_1, s_2, s_3\}$, where $s_1 =$ "depression", $s_2 =$ "recession", $s_3 =$ "return to normal".
   - Umbrella?: $S = \{s_1, s_2\}$, where $s_1 =$ "rains", $s_2 =$ "shines".

   Financial instruments give a way to protect against states of the world. Assume we have $n$ instruments. Before the uncertainty is resolved, an economic decision maker can buy financial instruments (or assets) of different types. Suppose that there are $k$ different instruments $j$ for

$j \in \{1, ..., k\}$. Denote the price of asset $j$ by $p_j$. Asset $j$ has a return of $y_{ij}$ in state $i$.

Hence buying $n_j$ units of asset $j$ costs $p_j n_j$ and has a return of $n_j y_{ij}$ in state $i$. Write $r_{ij} := \frac{y_{ij}}{p_j}$ for the return normalized by the price of the asset. This tells simply how much a Euro invested in $j$ returns in state $j$. Write all the normalized returns as a matrix

$$R = \begin{pmatrix} r_{11} & \cdots & r_{1k} \\ \vdots & \ddots & \vdots \\ r_{m1} & \cdots & r_{mk} \end{pmatrix}.$$

Let $x = (x_1, ..., x_k)$ denote an investment portfolio where $x_j := p_j n_j$ is the monetary amount invested in instrument $j$. Total return for the portfolio is then:

$$Rx = \begin{pmatrix} r_{11} & \cdots & r_{1k} \\ \vdots & \ddots & \vdots \\ r_{m1} & \cdots & r_{mk} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_k \end{pmatrix}.$$

In perfect financial markets all instruments can be sold and bought. This implies that each $x_j$ can be positive or negative.

Portfolio $x$ is risk free if

$$\Sigma_j r_{ij} x_j = \Sigma_j r_{i'j} x_j \text{ for all } i, i',$$

i.e. its return is the same across all states. Portfolio $x$ is an arbitrage portfolio if:

$$\Sigma_j x_j = 0.$$

Since we have normalized the prices of the instruments, this simply means that the price of the portfolio is zero before uncertainty is resolved. Therefore arbitrage portfolios can be bought without own initial wealth.

Arbitrage means the possibility of making a sure gain using an arbitrage portfolio. Hence the model allows for arbitrage if there is an arbitrage portfolio $x$, such that:

$$Rx \geq 0.$$

We say that the model has state prices $p = (q_1, ..., q_m)$ if

$$\Sigma_i q_i r_{ij} = 1 \text{ for all } j.$$

The price $q_i$ can be thought of as the price of a unit of wealth in state $i$. Recall that $r_{ij} 0 \frac{y_{ij}}{p_j}$ so an alternative way of writing the above equation is:

$$\Sigma_i q_i y_{ij} = p_j \text{ for all } j.$$

In matrix form, state prices $q$ satisfy:

$$q^\top R = (1, ..., 1)^\top R^\top q = (1, ..., 1).$$

State prices exist if $m \geq k$ and $rank\,(R) = k.$ They are not uniquely determined if $m > k$. We say that markets are complete if $m = rank(R) = k$

The fundamental theorem of finance relates the existence of positive state prices to the lack of arbitrage.

**Theorem 1** *Asset prices $p = (p_1, ..., p_k) \in \mathbb{R}^k$ admit no arbitrage if and only if the market has a strictly positive state price $q = (q_1, ..., q_m)$.*

Complete markets with no arbitrage have therefore a unique strictly positive state price.

# 5 Linear functions

The last topic in these notes concerns matrices as representations of linear functions. This is of fundamental importance for the analysis of functions of $n$ real variables. In general, functions are very complicated. Linear functions, on the other hand are much easier to analyze. Our strategy will be to reduce the general case as much as we can to the linear case and this will be done with the help of local analysis via calculus.

Recall that a function is a relation defined on the cartesian product of two sets $X$ and $Y$. We call $X$ the domain for a function $f$ and we call $Y$ the codomain. A function associates for each element $x \in X$ a single element $y \in Y$. The common notation for this is that $y = f(x) \in Y$.

We write also

$$f : X \to Y.$$

Functions are sometimes called mappings, maps or transformations. The sets $X, Y$ can be very general and the function can take many forms. Here are some examples:

- $X$ is the set of all humans, dead and alive and $Y = X$. For each $x$, $f(x)$ is the mother of $x$.
- $X$ is the population of Finland, $Y = \{0, 1\}$. $f(x) = 1$ if $x$ has Covid -antibodies in her blood and $f(x) = 0$ otherwise.
- $X$ is the set of feasible portfolios, $Y = \mathbb{R}$. $f(x)$ is the expected return on portfolio $x$.
- $X$ is the set of all humans, dead and alive and $Y = X$. The following relation $y = f(x)$ if $y$ is a child of $x$ is not a function since some $x$ have many children and some have none.

When are functions easy to understand? Intuitively, one would require that this is the case when the function can be understood based on a few representative cases and then extrapolated to the entire population.

This is exactly the reason why linear functions are so nice to work with. In order to get simplicity of the type desired above, we need first of all an easy way to describe the domain $X$ in terms of a few typical cases. We get this by assuming that $X$ is a vector space. This is a big word for a simple idea. It just says that scalar multiples of elements of $X$ are also elements of $X$. More formally, we require that if $a \in \mathbb{R}$ and $x \in X$, then $ax \in X$. Verify that the real line and k- dimensional vectors as defined at the beginning of these notes satisfy this requirement. The other requirement for a vector space is that whenever both $x$ and $x'$ are elements of $X$, we can define an operation $x + x'$ so that $x + x' \in X$. Again verify that the coordinate-wise addition for vectors (and real numbers) satisfies this.

For vectors of real numbers $X = \mathbb{R}^k$, we can make the useful observation based on the definitions above that with our definition of addition and scalar multiplication, we can write

$$x = (x_1, ..., x_n) = x_1\mathbf{e_1} + x_2\mathbf{e_2} + ... + x_n\mathbf{e_n}$$

$$= \sum_{i=1}^{n} x_i\mathbf{e_i},$$

where $\mathbf{e_i}$ is the $i^{th}$ unit coordinate vector.

A function $f : X \to Y$ between two vector spaces $X$ and $Y$ is called linear if for all $x, x' \in X$ and all $a \in \mathbb{R}$, i) $f(ax) = af(x)$ and ii) $f(x + x') = f(x) + f(x')$.

Here is the great thing about linear functions. If we know $f(\mathbf{e_i})$ for $i \in \{1, ..., n\}$, then we know $f(x)$ for all $x \in X$.

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

Suppose now that $Y = \mathbb{R}^m$. Define $\mathbf{a_i} = f(\mathbf{e_i}) \in \mathbb{R}^m$. Then we can represent the linear function $f$ by an $(m \times n)$ -matrix $A$:

$$A = [\mathbf{a_1}, ..., \mathbf{a_n}].$$

and

$$f(x) = Ax.$$

Similarly multiplication by an arbitrary $(m \times n)$ -matrix $A$ gives rise to a linear function from $\mathbb{R}^n$ to $\mathbb{R}^m$.

We say that a function $f : X \to Y$ is surjective or onto if for all $y \in Y$, there is (at least one) $x \in X$ such that $y = f(x)$. $f$ is called injective or one-to-one if for each $yinY$, there is at most one $x \in X$ such that $y = f(x)$. Finally, it is called bijective if it is both injective and surjective.

Note that for a bijective function, we can define an inverse function $f^{-1} : Y \to X$ so that we have for all $x \in X$ and all $y \in Y$:

$$x = f^{-1}(f(x)), \quad y = f(f^{-1}(y)).$$

Hence for a linear function $f$ represented by an $(m \times n$ -matrix A, we can use our results from the systems of equations to classify $f$. $f$ is surjective if for all $y \in Y$, there is an $x \in X$ such that:

$$Ax = y.$$

Recall that this is true if $rank(A) = m$.

It is injective if $rank(A) = n$. It is bijective if $m = rank(A) = n$. In this case, the inverse function $f^{-1}$ is represented by the inverse matrix $A^{-1}$.

We will make repeated use of these ideas in our analysis of non-linear functions. Later in this course, we will tackle some more advanced topics in linear algebra, but they will be covered in a self-contained manner.