CERTIFICATE IN QUANTITATIVE FINANCE

CQF

Certificate in Quantitative Finance

GLOBAL STANDARD IN FINANCIAL ENGINEERING

2 Introduction to Linear Algebra

2.1 Properties of Vectors

We consider real $n-$dimensional vectors belonging to the set $\mathbb{R}^n$. An $n-$tuple

$$\underline{v} = (v_1,\ v_2,\ ....,\ v_n) \in \mathbb{R}^n$$

is a vector of dimension n. The elements v_i $(i = 1, ..., n)$ are called components of $\underline{v}$.

Any pair $\underline{u},\ \underline{v} \in \mathbb{R}^n$ are equal iff the corresponding components u_i's and v_i's are equal

Examples:

$$\underline{u}_1 = (1,\ 0)\,,\quad \underline{u}_2 = \left(1,\ e,\ \sqrt{3},\ 6\right),\quad \underline{u}_3 = (3,\ 4)\,,\quad \underline{u}_4 = (\pi,\ \ln 3,\ 2,\ 1)$$

1. $\underline{u}_1$, $\underline{u}_3 \in \mathbb{R}^2$ and $\underline{u}_2$, $\underline{u}_4 \in \mathbb{R}^4$

2. $(x + y,\ x - z,\ 2z - 1) = (3,\ -2,\ 5)$ .For equality to hold corresponding components are equal, so

$$\left.\begin{array}{c} x + y = 3 \\ x - z = -2 \\ 2z - 1 = 5 \end{array}\right\} \Rightarrow x = 1;\ y = 2;\ z = 3$$

2.1.1 Vector Arithmetic

Let $\underline{u},\ \underline{v} \in \mathbb{R}^n$. Then *vector addition* is defined as

$$\underline{u} + \underline{v} = (u_1 + v_1,\ u_2 + v_2,\ ....,\ u_n + v_n)$$

If $k \in \mathbb{R}$ is any scalar then

$$k\underline{u} = (ku_1,\ ku_2,\ ....,\ ku_n)$$

Note: vector addition only holds if the dimensions of each are identical.

Examples:

$\underline{u} = (3,\ 1,\ -2,\ 0)\,,\quad \underline{v} = (5,\ -5,\ 1,\ 2)\,,\quad \underline{w} = (0,\ -5,\ 3,\ 1)$

1. $\underline{u} + \underline{v} = (3 + 5,\ 1 - 5,\ -2 + 1,\ 0 + 2) = (8,\ -4,\ -1,\ 2)$

2. $2\underline{w} = (2.0,\ 2.(-5),\ 2.3,\ 2.1) = (0,\ -10,\ 6,\ 2)$

3. $\underline{u} + \underline{v} - 2\underline{w} = (8,\ -4,\ -1,\ 2) - (0,\ -10,\ 6,\ 2) = (8,\ 6,\ -7,\ 0)$

$\underline{0} = (0,\ 0,\ .....,\ 0)$ is the *zero vector*.

Vectors can also be multiplied together using the *dot product* . If $\underline{u},\ \underline{v} \in \mathbb{R}^n$ then the dot product denoted by $\underline{u}.\underline{v}$ is

$$\underline{u}.\underline{v} = u_1 v_1 + u_2 v_2 + ... + u_n v_n \in \mathbb{R}$$

which is clearly a scalar quantity. The operation is commutative , i.e.

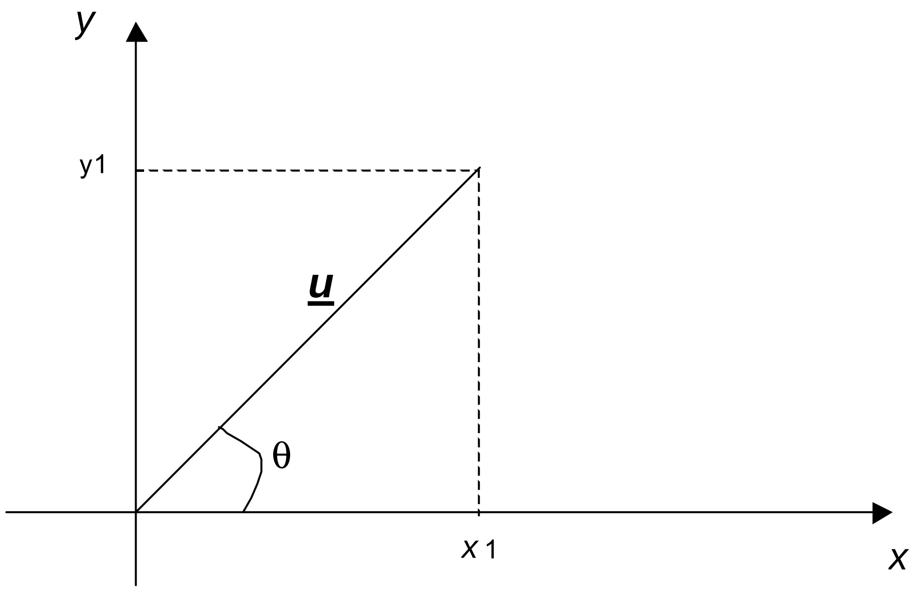
$$\underline{u}.\underline{v} = \underline{v}.\underline{u}$$

If a pair of vectors have a scalar product which is zero, they are said to be *orthogonal.*

Geometrically this means that the two vectors are perpendicular to each other.

2.1.2 Concept of Length in $\mathbb{R}^n$

Recall in 2-D $\underline{u} = (x_1, y_1)$



The length or *magnitude* of $\underline{u}$, written $|u|$ is given by Pythagoras

$$|u| = \sqrt{(x_1)^2 + (y_1)^2}$$

and the angle θ the vector makes with the horizontal is

$$\theta = \arctan \frac{y_1}{x_1}.$$

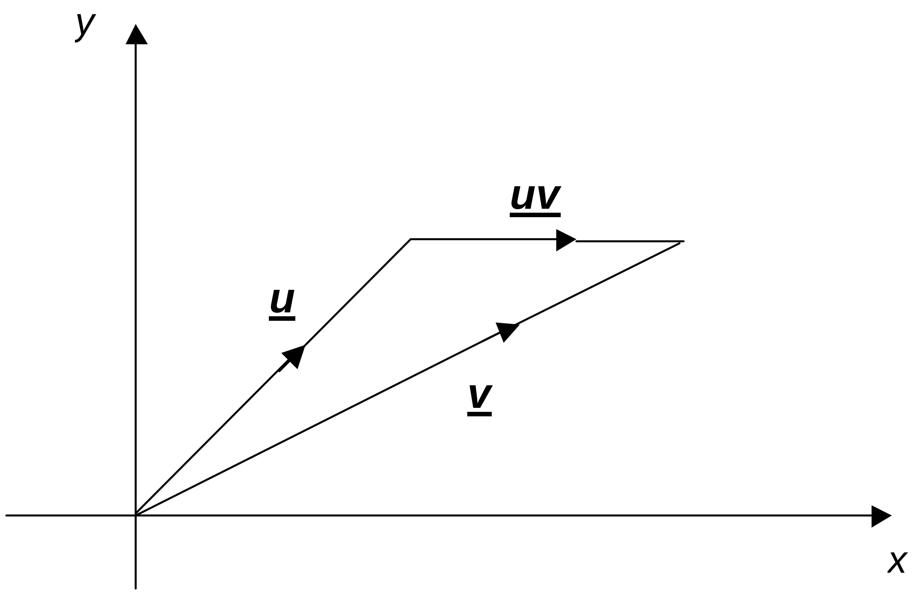
Any vector $\underline{u}$ can be expressed as

$$\underline{u} = |u| \, \underline{\widehat{u}}$$

where $\underline{\widehat{u}}$ is the *unit vector* because $|\underline{\widehat{u}}| = 1$.

Given any two vectors $\underline{u}, \ \underline{v} \in \mathbb{R}^2$, we can calculate the distance between them

$$\begin{aligned} |\underline{v} - \underline{u}| &= |(v_1 \ , \ v_2) - (u_1 \ , \ u_2)| \\ &= \sqrt{(v_1 - u_1)^2 + (v_2 - u_2)^2} \end{aligned}$$



In 3D (or $\mathbb{R}^3$) a vector $\underline{v} = (x_1,\ y_1,\ z_1)$ has length/magnitude

$$|v| = \sqrt{(x_1)^2 + (y_1)^2 + (z_1)^2}.$$

To extend this to $\mathbb{R}^n$, is similar.

Consider $\underline{v} = (v_1,\ v_2,\ \ldots\ldots\ldots,\ v_n) \in \mathbb{R}^n$. The length of $\underline{v}$ is called the *norm* and denoted $||\underline{v}||$, where

$$||\underline{v}|| = \sqrt{(v_1)^2 + (v_2)^2 + ... + (v_n)^2}$$

If $\underline{u},\ \underline{v} \in \mathbb{R}^n$ then the distance between $\underline{u}$ and $\underline{v}$ is can be obtained in a similar fashion

$$||\underline{v} - \underline{u}|| = \sqrt{(v_1 - u_1)^2 + (v_2 - u_2)^2 + ... + (v_n - u_n)^2}$$

We mentioned earlier that two vectors $\underline{u}$ and $\underline{v}$ in two dimension are orthogonal if $\underline{u}.\underline{v} = 0$.

The idea comes from the definition

$$\underline{u}.\underline{v} = |u|\ .\ |v| \cos\theta.$$

Re-arranging gives the angle between the two vectors. Note when $\theta = \pi/2$, $\underline{u}.\underline{v} = 0$.

If $\underline{u},\ \underline{v} \in \mathbb{R}^n$ we write

$$\underline{u}.\underline{v} = \|u\|.\|v\| \cos\theta$$

Examples: Consider the following vectors

$$\begin{aligned} \underline{u} &= (2, -1, 0, -3)\,,\ \underline{v} = (1, -1, -1, 3)\,, \\ \underline{w} &= (1, 3, -2, 2) \end{aligned}$$

$$\|\underline{u}\| = \sqrt{(2)^2 + (-1)^2 + (0)^2 + (-3)^2} = \sqrt{14}$$

Distance between $\underline{v}$ & $\underline{w} = \|\underline{w} - \underline{v}\| =$

$$\sqrt{(1-1)^2 + (3-(-1))^2 + (-2-(-1))^2 + (2-3)^2} = 3\sqrt{2}$$

The angle between $\underline{u}$ & $\underline{w}$ can be obtained from

$$\cos\theta = \frac{\underline{u}.\underline{v}}{|\,|u|\,|\;|\,|v|\,|}.$$

Hence

$$\begin{aligned}\cos\theta &= \frac{(2,-1,0,-3)\,.\,(1,-1,-1,3)}{2\sqrt{3}\sqrt{14}} = -\sqrt{\frac{3}{14}} \longrightarrow \\ \theta &= \cos^{-1}\left(-\sqrt{\tfrac{3}{14}}\right)\end{aligned}$$

2.2 Matrices

A *matrix* is a rectangular array $A = \left(a_{i\ j}\right)$ for $i = 1, ..., m$; $j = 1, ..., n$ written

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

and is an $(m \times n)$ matrix, i.e. m rows and n columns.

If $m = n$ the matrix is called *square*. The product mn gives the number of elements in the matrix.

2.2.1 Matrix Arithmetic

Let $A,\ B \in {}^m\mathbb{R}^n$

$$A + B = \begin{pmatrix} a_{11} & a_{12} & \cdots\cdot & a_{1n} \\ a_{21} & a_{22} & \cdots\cdot & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & .. & \cdots\cdot & \vdots \\ a_{m1} & a_{m2} & \cdots\cdot & a_{mn} \end{pmatrix} + \begin{pmatrix} b_{11} & b_{12} & .. & .. & b_{1n} \\ b_{21} & b_{22} & .. & .. & b_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & .. & .. & .. & \vdots \\ b_{m1} & b_{m2} & .. & .. & b_{mn} \end{pmatrix}$$

and the corresponding elements are added to give

$$\begin{pmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \cdots\cdot & a_{1n} + b_{1n} \\ a_{21} + b_{21} & a_{22} + +b_{22} & \cdots\cdot & a_{2n} + b_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & .. & \cdots\cdot & \vdots \\ a_{m1} + b_{m1} & a_{m2} + b_{m2} & \cdots\cdot & a_{mn} + b_{mn} \end{pmatrix} = B + A$$

Matrices can only added if they are of the same form.

Examples:

$$
\begin{aligned}
A &= \begin{pmatrix} 1 & -1 & 2 \\ 0 & 3 & 4 \end{pmatrix}, \quad B = \begin{pmatrix} 4 & 0 & -3 \\ -1 & -2 & 3 \end{pmatrix}, \\
C &= \begin{pmatrix} 2 & -3 & 1 \\ 5 & -1 & 2 \\ -1 & 0 & 3 \end{pmatrix}, \quad D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}
\end{aligned}
$$

$$
A + B = \begin{pmatrix} 5 & -1 & -1 \\ -1 & 1 & 7 \end{pmatrix}; \quad C + D = \begin{pmatrix} 3 & -3 & 1 \\ 5 & 0 & 2 \\ -1 & 0 & 4 \end{pmatrix}
$$

We cannot perform any other combination of addition as A and B are (2×3) and C and D are (3×3).

2.2.2 Matrix Multiplication

To multiply two square matrices $\mathbf{A}$ and $\mathbf{B}$, so that $\mathbf{C} = \mathbf{AB}$, the elements of $\mathbf{C}$ are found from the recipe

$$C_{ij} = \sum_{k=1}^{N} A_{ik} B_{kj}.$$

That is, the i^{th} row of $\mathbf{A}$ is dotted with the j^{th} column of $\mathbf{B}$. For example,

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} e & f \\ g & h \end{pmatrix} = \begin{pmatrix} ae + bg & af + bh \\ ce + dg & cf + dh \end{pmatrix}.$$

Note that in general $\mathbf{AB} \neq \mathbf{BA}$. The general rule for multiplication is

$$A_{pn}\ B_{nm} \rightarrow C_{pm}$$

Example:

$$
\begin{aligned}
& \begin{pmatrix} 2 & 1 & 0 \\ 2 & 0 & 2 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 0 & 3 \\ 1 & 2 \end{pmatrix} \\
= \; & \begin{pmatrix} 2.1 + 1.0 + 0.1 & 2.2 + 1.3 + 0.2 \\ 2.1 + 0.0 + 2.1 & 2.2 + 0.3 + 2.2 \end{pmatrix} \\
= \; & \begin{pmatrix} 2 & 7 \\ 4 & 8 \end{pmatrix}
\end{aligned}
$$

2.2.3 Transpose

The **transpose** of a matrix with entries A_{ij} is the matrix with entries A_{ji}; the entries are 'reflected' across the leading diagonal, i.e. rows become columns. The transpose of $\mathbf{A}$ is written $\mathbf{A}^{\mathsf{T}}$. If $\mathbf{A} = \mathbf{A}^{\mathsf{T}}$ then $\mathbf{A}$ is **symmetric**. For example, of the matrices

$$\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, \qquad \mathbf{B} = \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix}, \qquad \mathbf{C} = \mathbf{ix} \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix},$$

we have $\mathbf{B} = \mathbf{A}^T$ and $\mathbf{C} = \mathbf{C}^T$. Note that for any matrix $\mathbf{A}$ and B

(i) $(\mathbf{A} + B)^T = \mathbf{A}^T + B^T$

(ii) $\left(\mathbf{A}^T\right)^T = \mathbf{A}$

(iii) $(kA)^T = kA^T, \quad k$ is a scalar

(iv) $(AB)^T = B^T A^T$

Example:

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \\ 2 & 2 \end{pmatrix} \rightarrow A^T = \begin{pmatrix} 2 & 1 & 2 \\ 1 & 2 & 2 \end{pmatrix}$$

A *skew-symmetric* matrix has the property $a_{ij} = -a_{ji}$ with $a_{ii} = 0$. For example

$$\begin{pmatrix} 0 & 3 & -4 \\ -3 & 0 & 1 \\ 4 & -1 & 0 \end{pmatrix}$$

2.2.4 Matrix Representation of Linear Equations

We begin by considering a two-by-two set of equations for the unknowns x and y :

$$\begin{aligned} ax + by &= p \\ cx + dy &= q \end{aligned}$$

The solution is easily found. To get x, multiply the first equation by d, the second by b, and subtract to eliminate y :

$$(ad - bc)\, x = dp - bq.$$

Then find y :

$$(ad - bc)\, y = aq - cp.$$

This works and gives a unique solution *as long as* $ad - bc \neq 0$.

If $ad - bc = 0$, the situation is more complicated: there may be no solution at all, or there may be many.

Examples:

Here is a system with a unique solution:

$$\begin{aligned} x - y &= 0 \\ x + y &= 2 \end{aligned}$$

The solution is $x = y = 1$.

Now try

$$\begin{aligned} x - y &= 0 \\ 2x - 2y &= 2 \end{aligned}$$

Obviously there is no solution: from the first equation $x = y$, and putting this into the second gives $0 = 2$. Here $ad - bc = 1\,(-2) - (1-)\,2 = 0$.

Also note what is being said:

$$\left.\begin{array}{r} x = y \\ x = 1 + y \end{array}\right\} \text{ Impossible.}$$

Lastly try

$$\begin{aligned} x - y &= 1 \\ 2x - 2y &= 2. \end{aligned}$$

The second equation is twice the first so gives no new information. Any x and y satisfying the first equation satisfy the second. This system has many solutions.

Note: If we have one equation for two unknowns the system is undetermined and has many solutions. If we have *three* equations for two unknowns, it is over-determined and in general has no solutions at all.

Then the general (2×2) system is written

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} p \\ q \end{pmatrix}$$

or

$$\mathbf{A}\underline{\mathbf{x}} = \underline{\mathbf{p}}.$$

The equations can be solved if the matrix $\mathbf{A}$ is **invertible.** This is the same as saying that its **determinant**

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$$

is not zero.

These concepts generalise to systems of N equations in N unknowns. Now the matrix $\mathbf{A}$ is $N \times N$ and the vectors $\mathbf{x}$ and $\mathbf{p}$ have N entries.

Here are two special forms for $\mathbf{A}$. One is the $n \times n$ identity matrix,

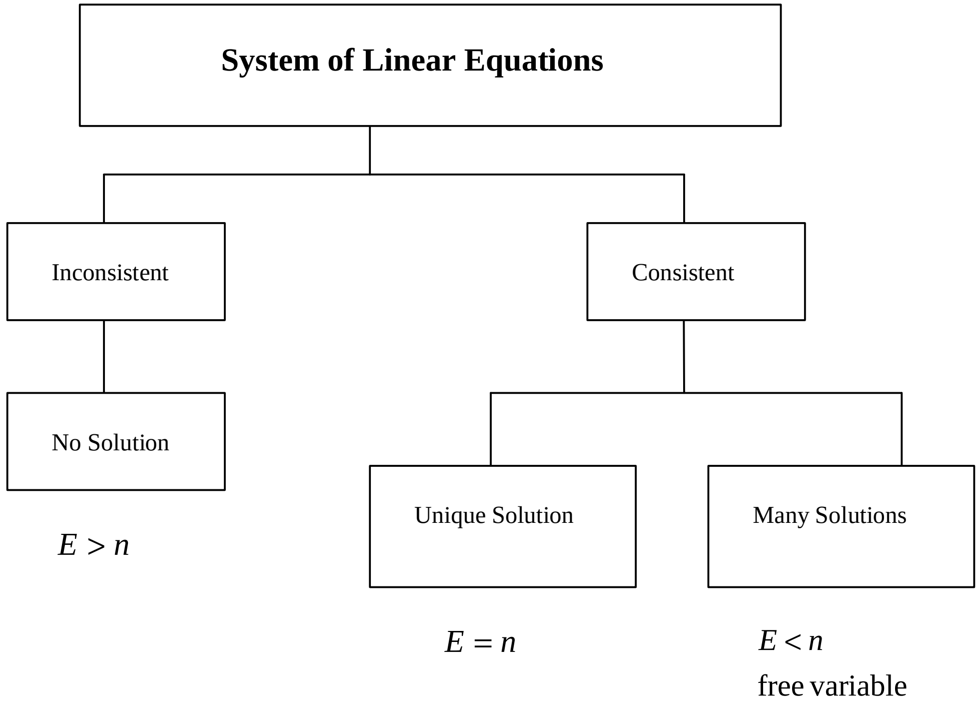
$$\mathsf{I} = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & \\ 0 & 0 & 1 & \dots & \vdots \\ \vdots & & \ddots & & 0 \\ 0 & & \dots & 0 & 1 \end{pmatrix}.$$

The other is the **tridiagonal form.** This is common in finite difference numerical schemes.

$$A = \begin{pmatrix} * & * & 0 & \cdots & \cdots & 0 \\ * & \ddots & \ddots & \ddots & & \vdots \\ 0 & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & \ddots & * \\ 0 & \cdots & \cdots & 0 & * & * \end{pmatrix}$$

There is a *main diagonal*, one above and below called the *super diagonal* and *sub-diagonal* in turn.

To conclude:



where $E =$ number of equations and $n =$ unknowns.

The theory and numerical analysis of linear systems accounts for quite a large branch of mathematics.

2.3 Using Matrix Notation For Solving Linear Systems

The usual notation for systems of linear equations is that of matrices and vectors. Consider the system

$$\begin{aligned} ax + by + cz &= p \\ dx + ey + fz &= q \\ gx + hy + iz &= r \end{aligned} \qquad (*)$$

for the unknown variables $x,\ y,\ z$. We gather the unknowns x, y and z and the given $p,\ q$ and r into vectors:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix}, \qquad \begin{pmatrix} p \\ q \\ r \end{pmatrix}$$

and put the coefficients into a matrix

$$A = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}.$$

A is called the *coefficient matrix* of the linear system $(*)$ and the special matrix formed by

$$\left(\begin{array}{ccc|c} a & b & c & p \\ d & e & f & q \\ g & h & i & r \end{array}\right)$$

is called the *augmented matrix.*

Now consider a general linear system consisting of n equations in n unknowns which can be written in augmented form as

$$\left(\begin{array}{cccccc|c} a_{11} & a_{12} & .. & .. & .. & a_{1n} & b_1 \\ a_{21} & a_{22} & .. & .. & .. & a_{2n} & b_2 \\ : & & & & & : & : \\ : & & & & & : & : \\ : & .. & .. & .. & .. & : & : \\ a_{n1} & a_{n2} & - & - & - & a_{nn} & b_n \end{array}\right).$$

We can perform a series of row operations on this matrix and reduce it to a simplified matrix of the form

$$\left(\begin{array}{cccccc|c} a_{11} & a_{12} & .. & .. & .. & a_{1n} & b_1 \\ 0 & a_{22} & .. & .. & .. & a_{2n} & b_2 \\ 0 & 0 & & & & : & : \\ 0 & 0 & 0 & & & : & : \\ : & .. & .. & .. & .. & : & : \\ 0 & 0 & - & - & 0 & a_{nn} & b_n \end{array}\right).$$

Such a matrix is said to be of *echelon form* if the number of zeros preceding the first non-zero entry of each row increases row by row.

A matrix A is said to be *row equivalent* to a matrix B, written $A \sim B$ if B can be obtained from A from a finite sequence of operations called *elementary row operations* of the form:

[ER_1]: Interchange the i th and j th rows: $R_i \leftrightarrow R_j$

[ER_2]:Replace the i th row by itself multiplied by a non-zero constant k : $R_i \longrightarrow kR_i$

[ER_3]:Replace the i th row by itself plus k times the j th row: $R_i \longrightarrow R_i + kR_j$

These have no affect on the solution of the of the linear system which gives the augmented matrix.

Examples:

Solve the following linear systems

1.

$$\left.\begin{aligned} 2x + y - 2z &= 10 \\ 3x + 2y + 2z &= 1 \\ 5x + 4y + 3z &= 4 \end{aligned}\right\} \equiv A\underline{x} = \underline{b} \text{ with}$$

$$A = \begin{pmatrix} 2 & 1 & -2 \\ 3 & 2 & 2 \\ 5 & 4 & 3 \end{pmatrix} \text{ and } \underline{b} = \begin{pmatrix} 10 \\ 1 \\ 4 \end{pmatrix}$$

The augmented matrix for this system is

$$\left(\begin{array}{ccc|c} 2 & 1 & -2 & 10 \\ 3 & 2 & 2 & 1 \\ 5 & 4 & 3 & 4 \end{array}\right) \underset{R_3 \to 2R_3 - 5R_1}{\overset{R_2 \to 2R_2 - 3R_1}{\sim}} \left(\begin{array}{ccc|c} 2 & 1 & -2 & 10 \\ 0 & 1 & 10 & -28 \\ 0 & 3 & 16 & -42 \end{array}\right)$$

$$\underset{R_1 \to R_1 - R_2}{\overset{R_3 \to R_3 - 3R_2}{\sim}} \left(\begin{array}{ccc|c} 2 & 0 & -12 & 38 \\ 0 & 1 & 10 & -28 \\ 0 & 0 & -14 & 42 \end{array}\right)$$

$$\begin{aligned} -14z &= 42 \to z = -3 \\ y + 10z &= -28 \to y = -28 + 30 = 2 \\ x - 6z &= 19 \to x = 19 - 18 = 1 \end{aligned}$$

Therefore solution is unique with

$$\underline{x} = \begin{pmatrix} 1 \\ 2 \\ -3 \end{pmatrix}$$

2.

$$\left.\begin{aligned} x + 2y - 3z &= 6 \\ 2x - y + 4z &= 2 \\ 4x + 3y - 2z &= 14 \end{aligned}\right\}$$

$$\left(\begin{array}{ccc|c} 1 & 2 & -3 & 6 \\ 2 & -1 & 4 & 2 \\ 4 & 3 & -2 & 14 \end{array}\right) \underset{R_3 \to R_3 - 4R_1}{\overset{R_2 \to R_2 - 2R_1}{\sim}}$$

$$\left(\begin{array}{ccc|c} 1 & 2 & -3 & 6 \\ 0 & -5 & 10 & -10 \\ 0 & -5 & 10 & -10 \end{array}\right) \underset{R_2 \to 0.5R_2}{\overset{R_3 \to R_3 - R_2}{\sim}}$$

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

Number of equations is less than number of unknowns.

$y - 2z = 2$ so $z = a$ is a free variable$\Rightarrow$ $y = 2\,(1 + a)$

$x + 2y - 3z = 6 \to x = 6 - 2y + 3z = 2 - a$

$$\Rightarrow x = 2 - a; \quad y = 2\,(1 + a)\,; \quad z = a$$

Therefore there are many solutions

$$\underline{x} = \begin{pmatrix} 2 - a \\ 2\,(1 + a) \\ a \end{pmatrix}$$

$$\left.\begin{array}{c} x+2y-3z=-1 \\ 3x-y+2z=7 \\ 5x+3y-4z=2 \end{array}\right\}$$

$$\left(\begin{array}{ccc|c} 1 & 2 & -3 & -1 \\ 3 & -1 & 2 & 7 \\ 5 & 3 & -4 & 2 \end{array}\right) \underset{R_3 \to R_3 - 5R_1}{\overset{R_2 \to R_2 - 3R_1}{\sim}}$$

$$\left(\begin{array}{ccc|c} 1 & 2 & -3 & -1 \\ 0 & -7 & 11 & 10 \\ 0 & -7 & 11 & 7 \end{array}\right) \overset{R_3 \to R_3 - R_2}{\sim}$$

$$\left(\begin{array}{ccc|c} 1 & 2 & -3 & -1 \\ 0 & -7 & 11 & 10 \\ 0 & 0 & 0 & -3 \end{array}\right)$$

The last line reads $0=-3$. Also middle iteration shows that the second and third equations are inconsistent.

Hence no solution exists.

2.4 Matrix Inverse

The **inverse** of a matrix $\mathbf{A}$, written $\mathbf{A}^{-1}$, satisfies

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

It may not always exist, but if it does, the solution of the system

$$\mathbf{A}\mathbf{x} = \mathbf{p}$$

is

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{p}.$$

The inverse of the matrix for the special case of a 2×2 matrix

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \frac{1}{ad - bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

provided that $ad - bc \neq 0$.

The inverse of any $n \times n$ matrix A is defined as

$$A^{-1} = \frac{1}{|A|}\text{adj } A$$

where adj $A = \left[(-1)^{\text{i+j}} \left|M_{\text{ij}}\right|\right]^{\text{T}}$ is the adjoint, i.e. we form the matrix of A's cofactors and transpose it.

M_{ij} is the square sub-matrix obtained by "covering the i$^{\text{th}}$ row and j$^{\text{th}}$ column", and its determinant is called the **Minor** of the element a_{ij}. The term $A_{\text{ij}} = (-1)^{\text{i+j}} \left|M_{\text{ij}}\right|$ is then called the **cofactor** of a_{ij}.

Consider the following example with

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

So the determinant is given by $|A| =$

$$(-1)^{1+1}\,A_{11}\,|M_{11}| + (-1)^{1+2}\,A_{12}\,|M_{12}| + (-1)^{1+3}\,A_{13}\,|M_{13}|$$

$$\begin{aligned} &= 1\begin{vmatrix} 2 & 1 \\ 1 & 3 \end{vmatrix} - 1\begin{vmatrix} 1 & 1 \\ 0 & 3 \end{vmatrix} + 0\begin{vmatrix} 1 & 2 \\ 0 & 1 \end{vmatrix} \\ &= (2 \times 3 - 1 \times 1) - (1 \times 3 - 1 \times 0) + 0 = 5 - 3 \\ &= 2 \end{aligned}$$

Here we have expanded about the 1$^{\text{st}}$ row - we can do this about any row. If we expand about the 2$^{\text{nd}}$ row - we should still get $|A| = 2$.

We now calculate the adjoint:

$$(-1)^{1+1}\,M_{11} = +\begin{vmatrix} 2 & 1 \\ 1 & 3 \end{vmatrix} \quad (-1)^{1+2}\,M_{12} = -\begin{vmatrix} 1 & 1 \\ 0 & 3 \end{vmatrix} \quad (-1)^{1+3}\,M_{13} = +\begin{vmatrix} 1 & 2 \\ 0 & 1 \end{vmatrix}$$

$$(-1)^{2+1}\,M_{21} = -\begin{vmatrix} 1 & 0 \\ 1 & 3 \end{vmatrix} \quad (-1)^{2+2}\,M_{22} = +\begin{vmatrix} 1 & 0 \\ 0 & 3 \end{vmatrix} \quad (-1)^{2+3}\,M_{23} = -\begin{vmatrix} 1 & 1 \\ 0 & 1 \end{vmatrix}$$

$$(-1)^{3+1}\,M_{31} = +\begin{vmatrix} 1 & 0 \\ 2 & 1 \end{vmatrix} \quad (-1)^{3+2}\,M_{32} = -\begin{vmatrix} 1 & 0 \\ 1 & 1 \end{vmatrix} \quad (-1)^{3+3}\,M_{33} = +\begin{vmatrix} 1 & 1 \\ 1 & 2 \end{vmatrix}$$

$$\text{adj } A = \begin{pmatrix} 5 & -3 & 1 \\ -3 & 3 & -1 \\ 1 & -1 & 1 \end{pmatrix}^{\mathsf{T}}$$

We can now write the inverse of A (which is symmetric)

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

Elementary row operations (as mentioned above) can be used to simplify a determinant, as increased numbers of zero entries present, requires less calculation. There are two important points, however. Suppose the value of the determinant is $|A|$, then:

[ER$_1$]: $R_i \leftrightarrow R_j \Rightarrow |A| \rightarrow -|A|$

[ER$_2$]: $R_i \rightarrow kR_i \Rightarrow |A| \rightarrow k|A|$

2.5 Orthogonal Matrices

A matrix $\mathbf{P}$ is **orthogonal** if

$$\mathbf{P}\mathbf{P}^{\mathsf{T}} = \mathbf{P}^{\mathsf{T}}\mathbf{P} = \mathsf{I}.$$

This means that the rows and columns of $\mathbf{P}$ are orthogonal and have unit length. It also means that

$$\mathbf{P}^{-1} = \mathbf{P}^{\mathsf{T}}.$$

In two dimensions, orthogonal matrices have the form

$$\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix}$$

for some angle θ and they correspond to rotations or reflections.

So rows and columns being orthogonal means row i·row $j = 0$, i.e. they are perpendicular to each other.

$$\begin{aligned}
(\cos\theta,\ \sin\theta)\cdot(-\sin\theta,\ \cos\theta) &= \\
-\cos\theta\sin\theta+\sin\theta\cos\theta &= 0 \\
(\cos\theta,\ \sin\theta)\cdot(\sin\theta,\ -\cos\theta) &= \\
\cos\theta\sin\theta-\sin\theta\cos\theta &= 0
\end{aligned}$$

$$\underline{v}=(\cos\theta,\ -\sin\theta)^T \to |\underline{v}| = \cos^2\theta+(-\sin\theta)^2 = 1$$

Finally, if $P = \begin{pmatrix}\cos\theta & \sin\theta\\ -\sin\theta & \cos\theta\end{pmatrix}$ then

$$P^{-1} = \frac{1}{\underbrace{\cos^2\theta-\left(-\sin^2\theta\right)}_{=1}}\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix} = P^T.$$

2.6 Eigenvalues and Eigenvectors

If $\mathbf{A}$ is a square matrix, the problem is to find values of λ (**eigenvalue**) for which

$$\mathbf{A}\underline{\mathbf{v}} = \boldsymbol{\lambda}\underline{\mathbf{v}}$$

has a non-trivial vector solution $\underline{\mathbf{v}}$ (**eigenvector**). We can write the above as

$$(\mathbf{A} - \boldsymbol{\lambda}\mathsf{I})\,\underline{\mathbf{v}} = \mathbf{0}.$$

An $N \times N$ matrix has exactly N eigenvalues, not all necessarily real or distinct; they are the roots of the *characteristic equation*

$$\det(\mathbf{A} - \boldsymbol{\lambda}\mathsf{I}) = 0.$$

Each solution has a corresponding eigenvector $\underline{\mathbf{v}}$. $\det(\mathbf{A} - \boldsymbol{\lambda}\mathsf{I})$ is the *characteristic polynomial.*

The eigenvectors can be regarded as special directions for the matrix $\mathbf{A}$. In complete generality this is a vast topic. Many Boundary-Value Problems can be reduced to eigenvalue problems.

We will just look at real symmetric matrices for which $\mathbf{A} = \mathbf{A}^{\mathsf{T}}$. For these matrices

- The eigenvalues are real;

- The eigenvectors corresponding to distinct eigenvalues are orthogonal;

- The matrix can be **diagonalised:** that is, there is an orthogonal matrix $\mathbf{P}$ such that

$$A = PDP^{\mathsf{T}} \qquad \text{or} \qquad P^{\mathsf{T}}AP = D$$

where $\mathbf{D}$ is **diagonal**, that is only the entries on the leading diagonal are non-zero, and these are equal to the eigenvalues of $\mathbf{A}$.

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

Example:

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

then so that the eigenvalues, i.e. the roots of this equation, are $\lambda_1 = -3$, $\lambda_2 = -2$ and $\lambda_3 = 6$.

Eigenvectors are now obtained from

$$\begin{pmatrix} 3-\lambda_i & 3 & 3 \\ 3 & -1-\lambda_i & 1 \\ 3 & 1 & -1-\lambda_i \end{pmatrix} \underline{\mathbf{v}}_i = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \quad i = 1,\ 2,\ 3$$

$$\lambda_1 = -3 : \qquad \begin{pmatrix} 6 & 3 & 3 \\ 3 & 2 & 1 \\ 3 & 1 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

Upon row reduction we have $\left(\begin{array}{ccc|c} 2 & 1 & 1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right) \rightarrow y = z$, so put $z = a$

and $2x = -y - z \rightarrow x = -\alpha \therefore \underline{\mathbf{v}}_1 = \alpha \begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix}$

Similarly

$$\lambda_2 = -2: \ \underline{\mathbf{v}}_2 = \beta \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}, \qquad \lambda_3 = 6: \ \underline{\mathbf{v}}_3 = \gamma \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}$$

If we take $\alpha = \beta = \gamma = 1$ the corresponding eigenvectors are

$$\underline{\mathbf{v}}_1 = \begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix}, \quad \underline{\mathbf{v}}_2 = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}, \quad \underline{\mathbf{v}}_3 = \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}$$

Now normalise these, i.e. $|\underline{\mathbf{v}}| = 1$. Use $\widehat{\underline{\mathbf{v}}} = \underline{\mathbf{v}}/\,|\underline{\mathbf{v}}|$ for normalised eigenvectors

$$\widehat{\underline{\mathbf{v}}}_1 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix}, \quad \widehat{\underline{\mathbf{v}}}_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}, \quad \widehat{\underline{\mathbf{v}}}_3 = \frac{1}{\sqrt{6}} \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}$$

Hence

$$\mathbf{P} = \begin{pmatrix} \frac{1}{\sqrt{3}} & 0 & \frac{2}{\sqrt{6}} \\ -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \end{pmatrix} \to \mathbf{P}^{\mathsf{T}} = \begin{pmatrix} \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{3}} \\ 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} \end{pmatrix}$$

so that

$$\begin{aligned}\mathbf{P}^{\mathsf{T}}\mathbf{A}\mathbf{P} &= \begin{pmatrix}-3 & 0 & 0\\ 0 & -2 & 0\\ 0 & 0 & 6\end{pmatrix}\\ &= D.\end{aligned}$$

2.6.1 Criteria for invertibility

A system of linear equations is uniquely solvable if and only if the matrix $\mathbf{A}$ is invertible. This in turn is true if any of the following is:

1. If and only if the determinant is non-zero;

2. If and only if all the eigenvalues are non-zero;

3. If (but not only if) it is **strictly diagonally dominant.**

In practise it takes far too long to work out the determinant. The second criterion is often useful though, and there are quite quick methods for working out the eigenvalues. The third method is explained on the next page.

A matrix $\mathbf{A}$ with entries A_{ij} is strictly diagonally dominant if

$$|A_{ii}| > \sum_{j \neq i} \left| A_{ij} \right| .$$

That is, the diagonal element in each row is bigger in modulus that the sum of the moduli of the off-diagonal elements in that row. Consider the following examples:

$$\begin{pmatrix} 2 & 0 & 1 \\ 1 & 4 & 2 \\ 1 & 3 & 6 \end{pmatrix} \quad \text{is s.d.d. and so invertible;}$$

$$\begin{pmatrix} 1 & 0 & 2 \\ 2 & 5 & 1 \\ 3 & 2 & 13 \end{pmatrix} \quad \text{is not s.d.d. but still invertible;}$$

$$\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \quad \text{is neither s.d.d. nor invertible.}$$