LINEAR ALGEBRA

Properties of Vectors

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

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

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

Any pair $\underline{u}$, $\underline{v}$ $\in \mathbb{R}^n$ are equal iff

1. the corresponding components u_i's and v_i's are equal

2. dimensions of both vectors are the same

and we write $\underline{u} = \underline{v}$.

Examples:

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

$$\underline{u}_1\ ,\ \underline{u}_3 \in \mathbb{R}^2$$

$$\underline{u}_2\ ,\ \underline{u}_4 \in \mathbb{R}^4$$

2. $(x+y,\ x-z,\ 2z-1) = (3,\ -2,\ 5)$.

For equality to hold corresponding components are equal, so

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

Vector Arithmetic

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

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

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

$$k\underline{u} = \left(ku_1,\ ku_2,\ \ldots\ldots\ldots\ldots,\ ku_n\right)$$

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

Examples:

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

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

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

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

$\underline{1} \in \mathbb{R}^n$ is the *unit vector* given by $\left(1,\ 1, \ldots\ldots,\ 1\right)$. Similarly $\underline{0} = \left(0,\ 0,\ \ldots\ldots,\ 0\right)$ 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

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

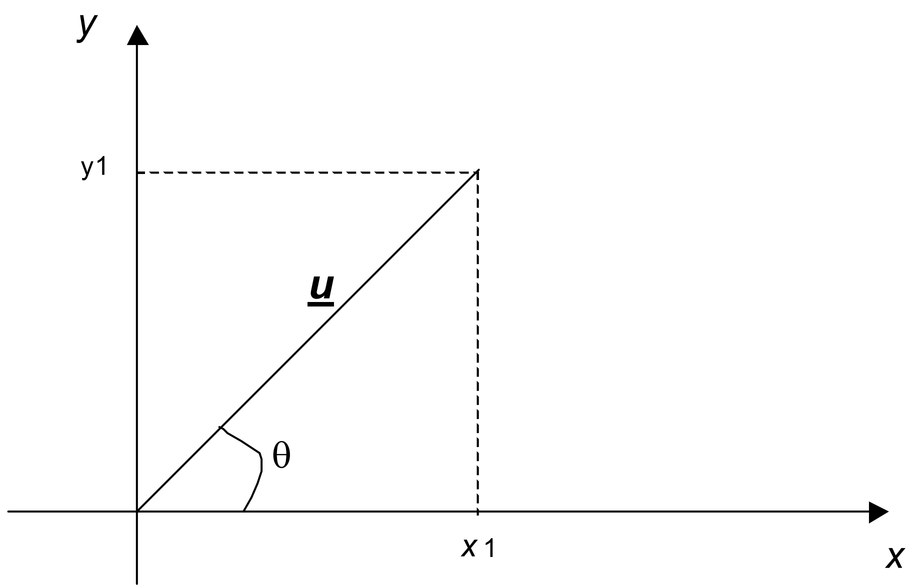
which is clearly a scalar quantity. The operation is commutative , i.e.

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

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.

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

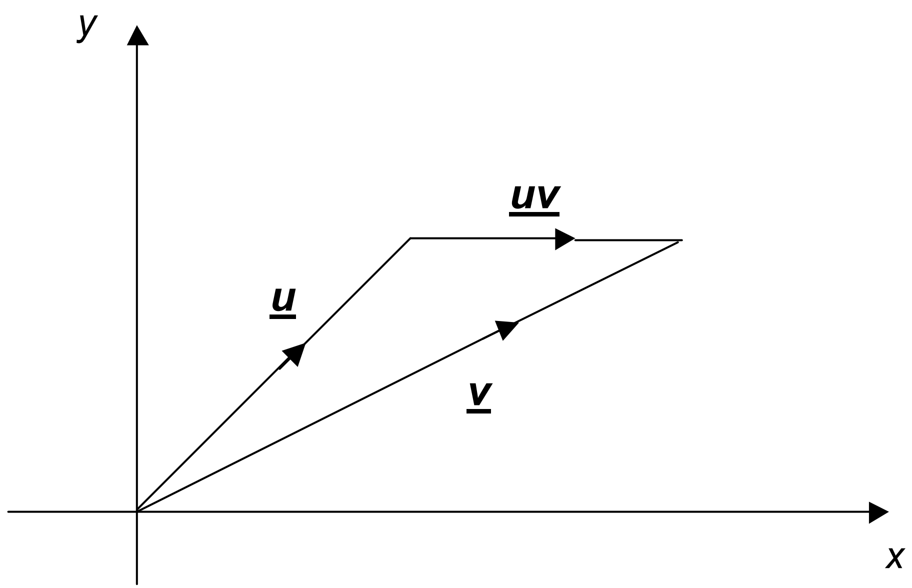
$$\theta = \arctan \tfrac{y_1}{x_1}.$$

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

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

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

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



$$|\underline{v} - \underline{u}| = \left|\left(v_1 , v_2\right) - \left(u_1 , u_2\right)\right| = \sqrt{(v_1 - u_1)^2 + (v_2 - u_2)^2}$$

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

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

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

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

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

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

$\boxed{\text{Examples:}}$ Consider the following vectors

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

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

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

$$\cos\theta = \frac{\left(2,\ -1,\ 0,\ -3\right).\left(1,\ -1,\ -1,\ 3\right)}{2\sqrt{3}\,\sqrt{14}} = -\sqrt{\frac{3}{14}} \rightarrow$$

$$\theta = \cos^{-1}\left(-\sqrt{\tfrac{3}{14}}\right)$$

Matrices

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

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

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.
A vector is an example of a $(m \times 1)$ matrix, i.e.

$$\underline{X} = \begin{pmatrix} x_1 \\ x_2 \\ : \\ : \\ : \\ x_m \end{pmatrix}$$

We define the set of all $(m \times n)$ real matrices as ${}^m\mathbb{R}^n$.

Matrix Arithmetic

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

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

and the corresponding elements are added to give

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

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

$\boxed{\text{Examples:}}$

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

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

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

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.

$\boxed{\text{Examples:}}$

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

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

Lastly try

$$x - y = 1$$
$$2x - 2y = 2.$$

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 **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 identity matrix

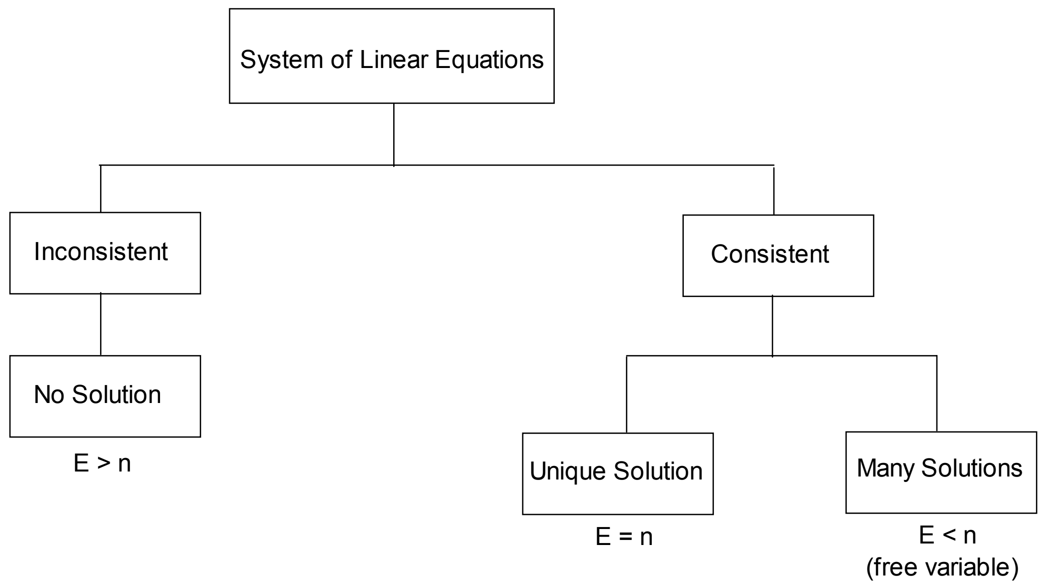
$$\mathrm{I} = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & \\ 0 & 0 & 1 & \dots & \vdots \\ \vdots & & \ddots & & 0 \\ 0 & & \dots & 0 & 1 \end{pmatrix}.$$

This is its own inverse and for any $\mathbf{x}$, $\mathrm{I}\mathbf{x} = \mathbf{x}$. The other is the **tridiagonal form.**

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

This is common in finite difference numerical schemes. There is a main diagonal, and one above called the *super diagonal* and one below called the *sub-diagonal.*

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.

Using Matrix Notation For Solution

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

$$\begin{aligned} ax + by + cz &= p \\ dx + ey + fz &= q \\ gx + hy + iz &= r \end{aligned} \tag{§}$$

We gather the unknowns x, y and z and the given p, q and r into vectors:

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

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*.

So as an example the augmented matrix for the system

$$2x - 3y + 6z = -1$$
$$3x + 6y + z = 4$$

is

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

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

$$\left(\begin{array}{cccccc|c} a_{11} & a_{12} & .. & .. & .. & a_{1n} & b_1 \\ a_{21} & a_{22} & .. & .. & .. & a_{2n} & b_2 \\ : & & & & & : & : \\ : & & & & & : & : \\ : & .. & .. & .. & .. & : & : \\ a_{m1} & a_{m2} & - & - & - & a_{mn} & b_m \end{array}\right).$$

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

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

Such a matrix is said to be of *echelon form* if the number of zeros preceding the first nonzero 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 nonzero constant k : $R_i \to kR_i$
[ER_3]:Replace the i^{th} row by itself plus k times the j^{th} row: $R_i \to R_i + kR_j$

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

$$\left(\begin{array}{ccc|c} 2 & 1 & -2 & 10 \\ 3 & 2 & 2 & 1 \\ 5 & 4 & 3 & 4 \end{array}\right) \underset{R_3 \to 2R_3 - 5R_1}{\overset{R_2 \to 2R_2 - 3R_1}{\sim}} \left(\begin{array}{ccc|c} 2 & 1 & -2 & 10 \\ 0 & 1 & 10 & -28 \\ 0 & 3 & 16 & -42 \end{array}\right) \underset{R_1 \to R_1 - R_2}{\overset{R_3 \to R_3 - 3R_2}{\sim}} \left(\begin{array}{ccc|c} 2 & 0 & -12 & 38 \\ 0 & 1 & 10 & -28 \\ 0 & 0 & -14 & 42 \end{array}\right)$$

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

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.

$$\begin{aligned} y - 2z &= 2 \ \text{ so } \ z = a \ \text{ is a free variable} \Rightarrow y = 2(1 + a) \\ x + 2y - 3z &= 6 \to x = 6 - 2y + 3z = 2 - a \\ &\Rightarrow x = 2 - a; \ \ y = 2(1 + a); \ \ z = a \end{aligned}$$

Therefore there are many solutions

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

3.

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

$$\left(\begin{array}{ccc|c} 1 & 2 & -3 & -1 \\ 3 & -1 & 2 & 7 \\ 5 & 3 & -4 & 2 \end{array}\right) \underset{R_3 \to R_3 - 5R_1}{\overset{R_2 \to R_2 - 3R_1}{\sim}} \left(\begin{array}{ccc|c} 1 & 2 & -3 & -1 \\ 0 & -7 & 11 & 10 \\ 0 & -7 & 11 & 7 \end{array}\right) \overset{R_3 \to R_3 - R_2}{\sim} \left(\begin{array}{ccc|c} 1 & 2 & -3 & -1 \\ 0 & -7 & 11 & 10 \\ 0 & 0 & 0 & -3 \end{array}\right)$$

The last line reads $0 = -3$. Also middle iteration shows that the second and third equations are inconsistent. Hence no solution exists.

Matrix Multiplication

To multiply two square matrices **A** and **B**, so that **C** = **AB**, the elements of **C** are found from the recipe

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

That is, the i th row of **A** is dotted with the j th column of **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 **AB** ≠ **BA**. The general rule for multiplication is

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

$\boxed{\text{Example:}}$

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

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 $\mathbf{A}^{\mathrm{T}}$. If $\mathbf{A} = \mathbf{A}^{\mathrm{T}}$ then $\mathbf{A}$ is **symmetric**. For example, of the matrices

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

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)** $(\mathbf{A}^T)^T = \mathbf{A}$
- **(iii)** $(kA)^T = kA^T$, k is a scalar
- **(iv)** $(AB)^T = B^T A^T$

$\boxed{\text{Example:}}$

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

Inverse

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

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

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

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

is

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

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

M_{ij} is the matrix obtained by "covering the ith row and jth column", and is called the **Minor**.

Consider the following example with

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

So the determinant is given by

$$\begin{aligned} |A| &= (-1)^{1+1}A_{11}M_{11} + (-1)^{1+2}A_{12}M_{12} + (-1)^{1+3}A_{13}M_{13} \\ &= 2\begin{vmatrix} -4 & 2 \\ -1 & 5 \end{vmatrix} - 3\begin{vmatrix} 0 & 2 \\ 1 & 5 \end{vmatrix} + (-4)\begin{vmatrix} 0 & -4 \\ 1 & -1 \end{vmatrix} \\ &= 2(-4 \times 5 + 2 \times 1) - 3(0 \times 5 - 2 \times 1) - 4(0 \times -1 + 4 \times 1) = 5 - 3 \\ &= 2 \end{aligned}$$

Here we have expanded about the 1st row - we can do this about any row. If we expand about the 2nd row - we should still get $|A| = -46$.

We now calculate the adjoint:

$$(-1)^{1+1}M_{11} = +\begin{vmatrix} -4 & 2 \\ -1 & 5 \end{vmatrix} \quad (-1)^{1+2}M_{12} = -\begin{vmatrix} 0 & 2 \\ 1 & 5 \end{vmatrix} \quad (-1)^{1+3}M_{13} = +\begin{vmatrix} 0 & -4 \\ 1 & -1 \end{vmatrix}$$

$$(-1)^{2+1}M_{21} = -\begin{vmatrix} 3 & -4 \\ -1 & 5 \end{vmatrix} \quad (-1)^{2+2}M_{22} = +\begin{vmatrix} 2 & -4 \\ 1 & 5 \end{vmatrix} \quad (-1)^{2+3}M_{23} = -\begin{vmatrix} 2 & 3 \\ 1 & -1 \end{vmatrix}$$

$$(-1)^{3+1}M_{31} = +\begin{vmatrix} 3 & -4 \\ -4 & 2 \end{vmatrix} \quad (-1)^{3+2}M_{32} = -\begin{vmatrix} 2 & -4 \\ 0 & 2 \end{vmatrix} \quad (-1)^{3+3}M_{33} = +\begin{vmatrix} 2 & 3 \\ 0 & -4 \end{vmatrix}$$

$$\text{adj}\, A = \begin{pmatrix} -18 & 2 & 4 \\ -11 & 14 & 5 \\ -10 & -4 & -8 \end{pmatrix}^{\mathrm{T}}$$

We can now write the inverse of A

$$A^{-1} = \frac{1}{46}\begin{pmatrix} -18 & -11 & -10 \\ 2 & 14 & -4 \\ 4 & 5 & -8 \end{pmatrix}$$

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$_1$]: $R_i \leftrightarrow R_j \Rightarrow |A| \to -|A|$

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

Orthogonal Matrices

A matrix **P** is **orthogonal** if

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

This means that the rows and columns of **P** are orthogonal and have unit length. It also means that

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

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•column $= 0$, i.e. they are perpendicular to each other.

$$\left(\cos\theta,\ \sin\theta\right).\left(-\sin\theta,\ \cos\theta\right) = -\cos\theta\sin\theta + \sin\theta\cos\theta = 0$$

$$\left(\cos\theta,\ \sin\theta\right).\left(\sin\theta,\ -\cos\theta\right) = \cos\theta\sin\theta - \sin\theta\cos\theta = 0$$

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

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

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

Eigenvalues and Eigenvectors

If $\mathbf{A}$ is a square matrix, $\underline{\mathbf{v}}$ is an **eigenvector** of $\mathbf{A}$ with **eigenvalue** λ if

$$\mathbf{A}\underline{\mathbf{v}} = \boldsymbol{\lambda}\underline{\mathbf{v}} \qquad \text{or} \qquad (\mathbf{A} - \boldsymbol{\lambda}\mathrm{I})\underline{\mathbf{v}} = \mathbf{0}.$$

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

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

and each solution has a corresponding eigenvector $\underline{\mathbf{v}}$. $\mathbf{A} - \boldsymbol{\lambda}\mathrm{I}$ is the *characteristic polynomial.*

The eigenvectors are in some sense 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 $\mathbf{A} = \mathbf{A}^{\mathrm{T}}$. 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

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

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

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

Effectively, all this says that if we change coordinates (basis) so that we use the eigenvectors instead of the original coordinate axes, the matrix $\mathbf{A}$ is in its simplest possible form.

For example, if

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

then

$$\det(\mathbf{A} - \boldsymbol{\lambda}\mathrm{I}) = \begin{vmatrix} 1-\lambda & 2 \\ 2 & 1-\lambda \end{vmatrix} = (1-\lambda)^2 - 4,$$

so that the eigenvalues, the roots of this equation, are $\lambda_1 = -1$ and $\lambda_2 = 3$.

For $\lambda_1 = -1$: $\begin{pmatrix} 1-\lambda_1 & 2 \\ 2 & 1-\lambda_1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \rightarrow$

$\begin{pmatrix} 2 & 2 & 0 \\ 2 & 2 & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \rightarrow x = -y$, so put $y = \alpha \therefore \underline{\mathbf{v}}_1 = \alpha\begin{pmatrix} 1 \\ -1 \end{pmatrix}$

Similarly for $\lambda_2 = 3$: $\begin{pmatrix} 1-\lambda_2 & 2 \\ 2 & 1-\lambda_2 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \rightarrow$

$\begin{pmatrix} -2 & 2 & 0 \\ 2 & -2 & 0 \end{pmatrix} \sim \begin{pmatrix} -1 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \rightarrow x = y$, so put $y = \beta \therefore \underline{\mathbf{v}}_2 = \beta\begin{pmatrix} 1 \\ 1 \end{pmatrix}$

If we take α and $\beta = 1$the corresponding eigenvectors are

$$\underline{\mathbf{v}}_1 = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \qquad \text{and} \qquad \underline{\mathbf{v}}_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix},$$

Now normalise these, i.e $|\underline{\mathbf{v}}| = 1$. We know $\underline{\mathbf{v}} = |\underline{\mathbf{v}}|\underline{\hat{\mathbf{v}}}$, therefore

and we have normalised eigenvectors

$$\underline{\mathbf{v}}_1 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix} \quad \text{and} \quad \underline{\mathbf{v}}_2 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix},$$

Hence

$$\mathbf{P} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \Rightarrow \mathbf{P}^{\mathrm{T}} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$$

so that

$$\begin{aligned}
\mathbf{P}^{\mathrm{T}}\mathbf{AP} &= \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}\frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \\
&= \frac{1}{2}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} -1 & 3 \\ 1 & 3 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & 3 \end{pmatrix} \\
&= D.
\end{aligned}$$

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 nonzero;
2. If and only if all the eigenvalues are nonzero;
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.

(Note: there are many other criteria for invertibility.)

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

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

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.

Examples:

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

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

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

Vector Spaces

We are interested in the non-abstract treatment of this subject. Throughout we will use the term *field* denoted F to refer to a set of scalars.

Definition:

A *vector space* V over F is a set with a binary operation called *Vector Addition*, denoted $V \times V \to V$, $\left(x, y\right) \mapsto x + y$, and a function $F \times V \to V$, $\left(c, x\right) \mapsto cx$, $(c \in F)$, called *Scalar Multiplication* such that the following eight rules hold:

+1) + is associative $\forall x, y, z \in V \;\; (x + y) + z = x + (y + z)$

+2) + is commutative $\forall x, y \in V \;\; x + y = y + x$

+3) + has a neutral $\exists\, 0 \in V \;\; \forall x \in V \quad x + 0 = x$

+4) + has inverse $\forall x \in V \;\; \exists\, y \in V \;\; x + y = 0 \;\; y$ is denoted $(-x)$

•1) • is associative $\forall c, d \in F, \;\; \forall x \in V \;\; c(dx) = (cd)x$

•2) • is commutative $\forall x, y \in V \;\; xy = yx$

•3) • has a neutral $\forall x \in V \;\; 1 \bullet x = x \quad (1 \neq 0)$

•4) • has an inverse $\forall x \in V \;\; (x \neq 0) \Rightarrow \;\; \left(\exists\, y \in V \;\; xy = 1\right) \;\; y$ is denoted (x^{-1})

+•1) Right distributive $\forall c \in F \;\; \forall x, y \in V \;\; c(x + y) = cx + cy$ scalar multiplication is distributive over vector addition

+•2) Left distributive $(c + d)x = cx + dx$

Remarks:

1. Elements of F are called <u>SCALARS</u> and elements of V are called <u>VECTORS</u>.

2. If $F = \mathbb{R}$ we say V is a *real* vector space
$= \mathbb{C}$ *complex*
$= \mathbb{Q}$ *rational*

3. At this stage we have

2	$+$'s	addition
2	$\cdot$'s	multiplication
2	0's	neutrals
2	$-$'s	inverses

(and things usually get a lot worse)

4. The axioms can be used to deduce various rules.

Examples:

1. Let $m, n \in \mathbb{N}^+$ then ${}^mF^n$ is a vector space over F with respect to operations of matrix addition and scalar multiplication.

2. Let $V = \mathbb{R}[x]$ denote the set of all polynomials

$$\sum_{n=0}^{N} a_n x^n \quad n \in \mathbb{N}, \;\; a_i(i = 1, \ldots, N) \in \mathbb{R}$$

Then V is a vector space over $\mathbb{R}$ w.r.t. addition of polynomials and multiplication by a constant.

3. Let F be an arbitrary field and V the set of all n dimensional vectors with vector addition

$$(a_1, \ldots\ldots, a_n) + (b_1, \ldots\ldots, b_n) = (a_1 + b_1, \ldots\ldots\ldots\ldots, a_n + b_n)$$

and scalar multiplication

$$k(a_1, a_2, \ldots\ldots, a_n) = (ka_1, ka_2, \ldots\ldots, ka_n)$$

where $a_i, b_i, k \in F$. Then V is a vector space over F.

Subspaces

Definition:

A subspace U of a vector space V (over F) is a subset , i.e. $U \subset V$ such that

(i) $0 \in U$
(ii) $\forall x, y \in U \quad x + y \in U$
(iii) $\forall c \in F \quad \forall x \in U \quad cx \in U$

Note:
U is then a vector space with vector addition and scalar multiplication calculated in V.

$U \times U \to U$ is a binary operation on U

$$(x, y) \mapsto x + y$$

$F \times U \to U,$ is a function

$$(c, x) \mapsto cx$$

The eight rules obviously hold because they are satisfied in the larger space and will automatically be satisfied in every subspace.

Examples:

(1) Let $V =$ set of all 3×3 matrices and $U, W \subset V$ such that

$U =$ set of lower triangular matrices

$W =$ set of symmetric matrices.

Suppose $A, B \in U$; $C, D \in W$, then $A + B \in U$ and $cA \in U$ where $c \in \mathbb{R}$. The sums $A + B$ and cA inherit properties of A and B; and similarly $C + D \in W$ and $kC \in W$ where $k \in \mathbb{R}$. 0 is in both spaces. Hence U and W are subspaces of V.

(2) Consider the vector space $V = \mathbb{R}^2$ over $F = \mathbb{R}$. $U = \{(x, y) \,/\, x, y \in [0, \infty)\}$ is the subset consisting of vectors whose components are ≥ 0, i.e the first quadrant, all co-ordinates $(x, y) \geq 0$. Now $\underline{0} \in U$ and U is closed under vector addition $(x + y) \in U$. What about closure under scalar multiplication? Suppose $\underline{u} = (1, 1) \in U$ and $c = -1 \in \mathbb{R}$, then $c\underline{u} = (-1, -1) \notin U$ $\therefore$ scalar multiplication fails. Hence U is not a subspace of $\mathbb{R}^2$.

Now suppose $W = \{(-a, -b) \,/\, a, b \in [0, \infty)\} \subset V$, i.e. the third quadrant. Define a new subset of V such that $S = U + W \subset V$. We see that for any vector $\underline{w}$ in S, $k\underline{w} \in S$ where $k \in \mathbb{R}$, so closure under scalar multiplication. However now addition fails because $(2, 1) + (-1, -3) = (1, -2)$ which is in neither quadrant.

So the smallest subspace containing the 1st quadrant is the whole space $\mathbb{R}^2$.

Definition:

Let V be a vector space over a field F. Suppose the vectors $v_1, v_2, \ldots\ldots, v_n \in V$ are a finite sequence of elements. We say that $v_1, \ldots\ldots, v_n$ are linearly dependent if $\exists$ scalars $\lambda_1, \ldots, \lambda_n \in F$ (not all zero) such that

$$\lambda_1 v_1 + \lambda_2 v_2 + \ldots\ldots + \lambda_n v_n = 0.$$

$\boxed{\text{Examples:}}$

1. Let $F = \mathbb{R}, \ \ V = \mathbb{R}^3$

$$\underline{v}_1 = \begin{pmatrix} 1 \\ 4 \\ 6 \end{pmatrix}, \quad \underline{v}_2 = \begin{pmatrix} 2 \\ 8 \\ 5 \end{pmatrix}, \quad \underline{v}_3 = \begin{pmatrix} 3 \\ 12 \\ 0 \end{pmatrix},$$

are $\underline{v}_1, \underline{v}_2, \underline{v}_3$ linearly dependent $\left(\text{over } \mathbb{R}\right)$?

$\exists ?\ \lambda_1, \ \lambda_2, \ \lambda_3 \in \mathbb{R}$ not all zero / $\lambda_1 v_1 + \lambda_2 v_2 + \lambda_3 v_3 = 0$. That is

$\exists ?\ \lambda_1, \ \lambda_2, \ \lambda_3 \in \mathbb{R}$ not all zero / $\lambda_1 \begin{pmatrix} 1 \\ 4 \\ 6 \end{pmatrix} + \lambda_2 \begin{pmatrix} 2 \\ 8 \\ 5 \end{pmatrix} + \lambda_3 \begin{pmatrix} 3 \\ 12 \\ 0 \end{pmatrix} = 0.$

So we solve $\begin{pmatrix} 1 & 2 & 3 \\ 4 & 8 & 12 \\ 6 & 5 & 0 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$, by writing in augmented form:

$$\left(\begin{array}{ccc|c} 1 & 2 & 3 & 0 \\ 4 & 8 & 12 & 0 \\ 6 & 5 & 0 & 0 \end{array}\right)$$

and row reducing. We will ignore the right hand side as it remains unchanged through out.

$$\begin{array}{r} R_2 \to R_2 - 4R_1 \\ R_3 \to R_3 - 6R_1 \\ R_2 \leftrightarrow R_3 \end{array} \to \begin{pmatrix} 1 & 2 & 3 \\ 0 & -7 & -18 \\ 0 & 0 & 0 \end{pmatrix}$$

$$R_2 \to -\frac{1}{7}R_2 \to \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & \frac{18}{7} \\ 0 & 0 & 0 \end{pmatrix}$$

$$R_1 \to R_1 - 2R_2 \to \begin{pmatrix} 1 & 0 & -\frac{15}{7} \\ 0 & 1 & \frac{18}{7} \\ 0 & 0 & 0 \end{pmatrix}$$

λ_3 – free variable; $\lambda_1 = \frac{15}{7}\lambda_3$; $\lambda_2 = -\frac{18}{7}\lambda_3$

Put $\lambda_3 = 7$, then $\lambda_1 = 15$; $\lambda_2 = -18$. So λ's not all zero

$$15\underline{v}_1 - 18\underline{v}_2 + 7\underline{v}_3 = 0.$$

Hence each vector can be expressed as a combination of the others, by re-arranging.

2. Are $\underline{v}_1 = \begin{pmatrix} 1 \\ 0 \\ 3 \end{pmatrix}$, $\underline{v}_2 = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$, $\underline{v}_3 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$ linearly dependent?

$0\underline{v}_1 + 15\underline{v}_2 + 0\underline{v}_3 = 0$
Scalars not all zero.

$\boxed{\textbf{Moral:}}$ If any of $v_1, v_2, \ldots\ldots, v_n$ is $\underline{0}$ then the set of vectors is linearly dependent.

3. Are $\underline{v}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$, $\underline{v}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$, $\underline{v}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ linearly dependent?

Augmented matrix is already in row reduced echelon form $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ and there is only one solution $\lambda_1 = \lambda_2 = \lambda_3 = 0$. So vectors are **NOT** linearly dependent.

Definition:

If $v_1, v_2, \ldots\ldots, v_n \in V$ are not linearly dependent then we say that they are *linearly independent*. That is if $\forall$ scalars $\lambda_1, \ldots., \lambda_n \in F$

$$\sum_{i=1}^{n} \lambda_i v_i = 0$$

$$\Rightarrow \lambda_1 = \ldots\ldots = \lambda_n = 0$$

So the solution to $\sum_{i=1}^{n} \lambda_i v_i = 0$ is the trivial solution $\lambda_1 = \ldots\ldots = \lambda_n = 0$.

The sequence of vectors in Examples 3 was linearly independent.

Definition:

Let $\underline{v}_1, \underline{v}_2, \ldots\ldots, \underline{v}_n$ $(n \in \mathbb{N})$ be a finite sequence in V. A *linear combination* of $\underline{v}_1, \underline{v}_2, \ldots\ldots, \underline{v}_n$ is a vector $\underline{v} \in V$ which can be expressed in the form

$$\underline{v} = \lambda_1 v_1 + \lambda_2 v_2 + \ \ldots\ldots + \lambda_n v_n$$

for some $\lambda_1, \ldots, \lambda_n \in F$.

i.e. $\exists\ \lambda_1, \ldots, \lambda_n \in F \,/\ \underline{v} = \lambda_1 v_1 + \lambda_2 v_2 + \ \ldots\ldots + \lambda_n v_n$.

Example:

Is $(3,\ 12,\ 0)$ a linear combination of $(1,\ 4,\ 6)$ and $(2,\ 8,\ 5)$?
(Geometrically, does $(3,\ 12,\ 0)$ lie in the same plane as $(1,\ 4,\ 6)$ and $(2,\ 8,\ 5)$?)

So $\exists ?\ \lambda_1, \lambda_2 \in F\,/$

$$\begin{aligned}(3,\ 12,\ 0) &= \lambda_1(1,\ 4,\ 6) + \lambda_2(2,\ 8,\ 5) \\ &= (\lambda_1 + 2\lambda_2,\ 4\lambda_1 + 8\lambda_2,\ 6\lambda_1 + 5\lambda_2)\end{aligned}$$

which gives the linear system

$$\begin{aligned}\lambda_1 + 2\lambda_2 &= 3 \\ 4\lambda_1 + 8\lambda_2 &= 12 \\ 6\lambda_1 + 5\lambda_2 &= 0\end{aligned}$$

which in augmented form is

$$\left(\begin{array}{cc|c} 1 & 2 & 3 \\ 4 & 8 & 12 \\ 6 & 5 & 0 \end{array}\right).$$

We can row reduce to obtain

$$\left(\begin{array}{cc|c} 1 & 0 & -\frac{15}{7} \\ 0 & 1 & \frac{18}{7} \\ 0 & 0 & 0 \end{array}\right) \rightarrow \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} = \begin{pmatrix} -\frac{15}{7} \\ \frac{18}{7} \end{pmatrix}.$$

So answer is Yes, $\begin{pmatrix} 3 \\ 12 \\ 0 \end{pmatrix} = -\frac{15}{7}\begin{pmatrix} 1 \\ 4 \\ 6 \end{pmatrix} + \frac{18}{7}\begin{pmatrix} 2 \\ 8 \\ 5 \end{pmatrix}$ is a linear combination of $\begin{pmatrix} 1 \\ 4 \\ 6 \end{pmatrix}$ and $\begin{pmatrix} 2 \\ 8 \\ 5 \end{pmatrix}$.

This could also have been deduced from a previous example where we showed

$$-15\begin{pmatrix} 1 \\ 4 \\ 6 \end{pmatrix} + 18\begin{pmatrix} 2 \\ 8 \\ 5 \end{pmatrix} - 7\begin{pmatrix} 3 \\ 12 \\ 0 \end{pmatrix} = 0.$$

Linear dependence here gives linear combination.

Linear ODE's of Order at least 2

General n^th order linear ode is of form:

$$a_n(x)y^{(n)} + a_{n-1}(x)y^{(n-1)} + \ldots\ldots\ldots + a_1(x)y' + a_0(x)y = g(x)$$

Use symbolic notation:

$$D \equiv \frac{d}{dx} \; ; \; D^r \equiv \frac{d^r}{dx^r} \quad \text{so } D^r y \equiv \frac{d^r y}{dx^r}$$

$\therefore \; a_r D^r \equiv a_r(x)\dfrac{d^r}{dx^r}$ so

$$\boxed{a_r D^r y = a_r(x)\frac{d^r y}{dx^r}}$$

Now introduce

$$L = a_n D^n + a_{n-1}D^{n-1} + a_{n-2}D^{n-2} + \ldots\ldots\ldots\ldots + a_1 D + a_0$$

so we can write a linear ode in the form

$$\boxed{L\,y = g}$$

L – Linear Differential Operator of order n and its definition will be used throughout.

If $g(x) = 0\ \forall x$, then $L\,y = 0$ is said to be **HOMOGENEOUS**.
$L\,y = 0$ is said to be the homogeneous part of $L\,y = g$.

L is a *linear operator* because as is trivially verified:
(1) $L\,(y_1 + y_2) = L\,(y_1) + L(y_2)$
(2) $L\,(cy) = cL(y) \qquad c \in \mathbb{R}$

GS of $Ly = g$ is given by

$$\boxed{y = y_c + y_p}$$

where y_c – Complimentary Function & y_p – Particular Integral (or Particular Solution)

$$\left.\begin{array}{l} y_c \text{ is solution of } Ly = 0 \\ y_p \text{ is solution of } Ly = g \end{array}\right\} \therefore \text{ GS } y = y_c + y_p$$

Look at homogeneous case $Ly = 0$. Put Ⓢ = all solutions of $Ly = 0$. Then Ⓢ forms a vector space of dimension n. Functions $y_1(x), \ldots\ldots\ldots, y_n(x)$ are LINEARLY DEPENDENT if $\exists\, \lambda_1, \ldots\ldots\ldots, \lambda_n \in \mathbb{R}$ (not all zero) s.t

$$\lambda_1 y_1(x) + \lambda_2 y_2(x) + \ldots\ldots\ldots\ldots + \lambda_n y_n(x) = 0$$

Otherwise y_i's $\left(i = 1, \ldots\ldots, n\right)$ are said to be LINEARLY INDEPENDENT (Lin. Indep.) $\Rightarrow$ whenever

$$\lambda_1 y_1(x) + \lambda_2 y_2(x) + \ldots\ldots\ldots\ldots + \lambda_n y_n(x) = 0 \ \forall x \text{ then } \lambda_1 = \lambda_2 = \ldots\ldots\ldots = \lambda_n = 0$$

FACT:

(1) $L - n^{\text{th}}$ order linear operator, then $\exists\, n$ Lin. Indep. solutions $y_1, \ldots\ldots, y_n$ of $Ly = 0$ s.t. G.S. of $Ly = 0$ is given by

$$y = \lambda_1 y_1 + \lambda_2 y_2 + \ldots\ldots\ldots\ldots + \lambda_n y_n \quad \underset{1 \le i \le n}{\lambda_i \in \mathbb{R}} \;.$$

(2) Any n Lin. Indep. solutions of $Ly = 0$ have this property. To solve $Ly = 0$ we need only find by "hook or by crook" n Lin. Indep. solutions.

All basic features appear for the case $n = 2$, so we analyse this.

$$L\,y = a\frac{d^2y}{dx^2} + b\frac{dy}{dx} + cy = 0 \quad a,\, b,\, c \in \mathbb{R}$$

Recall - For a constant coefficient problem , we seek the existence of a solution of the form $y = \exp(\lambda x)$, which gives

$$L(e^{\lambda x}) = (aD^2 + bD + c)e^{\lambda x}$$

hence $a\lambda^2 + b\lambda + c = 0$ and so λ is a root of the quadratic equation

$$a\lambda^2 + b\lambda + c = 0 \qquad \textbf{AUXILLIARY EQUATION (A.E)}$$

There are three cases to consider for $b^2 - 4ac$. Let us look for equality: $b^2 - 4ac = 0$

So $\lambda = \lambda_1 = \lambda_2 = -\dfrac{b}{2a}$

Clearly $e^{\lambda x}$ is a solution of $L\,y = 0$ - but theory tells us there exist two solutions for a 2nd order ode. So now try $y = x\exp(\lambda x)$

$$\begin{aligned} L(xe^{\lambda x}) &= (aD^2 + bD + c)(xe^{\lambda x}) \\ &= \underbrace{(a\lambda^2 + b\lambda + c)}_{=0}(xe^{\lambda x}) + \underbrace{(2a\lambda + b)}_{=0}(e^{\lambda x}) \\ &= 0 \end{aligned}$$

This gives a 2nd solution $\therefore$ GS is $y = c_1\exp(\lambda x) + c_2 x\exp(\lambda x)$, hence

$$\boxed{y = (c_1 + c_2 x)\exp(\lambda x)}$$

General n^{th} Order Equation

Consider

$$L\,y = a_n y^{(n)} + a_{n-1} y^{(n-1)} + \ldots\ldots\ldots + a_1 y' + a_0 y = 0$$

then

$$L \equiv D^n + \widehat{a}_{n-1} D^{n-1} + \widehat{a}_{n-2} D^{n-2} + \ldots\ldots\ldots\ldots\ldots + \widehat{a}_1 D + \widehat{a}_0 \quad \widehat{a}_i \in \mathbb{R}\ (0 \le i \le n-1)$$

$\left(\text{we have divided through by } a_n, \text{ i.e. } \widehat{a}_i = \frac{a_i}{a_n}\right)$ so $L\,y = 0$

A.E becomes $\boxed{\lambda^n + \widehat{a}_{n-1}\lambda^{n-1} + \ldots\ldots\ldots\ldots + \widehat{a}_1\lambda + \widehat{a}_0 = 0}$

Case 1 (Basic)

n distinct roots $\lambda_1, \ldots\ldots\ldots, \lambda_n$ then $e^{\lambda_1 x}, e^{\lambda_2 x}, \ldots\ldots.., e^{\lambda_n x}$ are n Lin. Indep. solutions giving a GS

$$y = \beta_1 e^{\lambda_1 x} + \beta_2 e^{\lambda_2 x} + \ldots\ldots\ldots + \beta_n e^{\lambda_n x}$$

β_i – arb.

$\boxed{\textbf{Case 2}}$

If λ is a real r – fold root of the A.E then $e^{\lambda x}, xe^{\lambda x}, x^2 e^{\lambda x}, \ldots\ldots\ldots\ldots, x^{r-1}e^{\lambda x}$ are r Lin. Indep. solutions of $L\,y = 0$, i.e.

$$y = e^{\lambda x}\left(\alpha_1 + \alpha_2 x + \alpha_3 x^2 \ldots\ldots\ldots + \alpha_r\, x^{r-1}\right)$$

α_i – arb.

$\boxed{\textbf{Case 3}}$

If $\lambda = p + iq$ is a r - fold root of the A.E then so is $p - iq$

$$\left.\begin{array}{c} e^{px}\cos qx,\ xe^{px}\cos qx,\ \ldots\ldots\ldots\ldots\ldots\ldots,\ x^{r-1}e^{px}\cos qx \\ e^{px}\sin qx,\ xe^{px}\sin qx,\ \ldots\ldots\ldots\ldots\ldots\ldots,\ x^{r-1}e^{px}\sin qx \end{array}\right\}$$

$\rightarrow$ $2r$ Lin. Indep. solutions of $L\,y = 0$

GS $y = e^{px}(c_1 + c_2x + c_3x^2 + \ldots\ldots\ldots\ldots)\cos qx + e^{px}(C_1 + C_2x + C_3x^2 + \ldots\ldots\ldots\ldots)\sin qx$

Portfolio Management & Constrained Optimisation - Method of Lagrange Multipliers (Revisited)

To conclude we set out the working presented in class for the application of Lagrange Multipliers to optimise a portfolio of stocks. Although this problem has previously been discussed, it is included as a way to illustrate a finance related application of the subject matter of this lecture. The simple case of 3 stocks is considered.

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

$$|A| = 2$$

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

We note A is symmetric, hence $A^{\mathrm{T}} = A$. We can now write the inverse of A

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

Problem:

Consider a Markowitz world with a three asset risky economy where the covariance matrix of expected returns is given by A (as above)

$$\begin{pmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} \\ \sigma_{12} & \sigma_2^2 & \sigma_{23} \\ \sigma_{13} & \sigma_{23} & \sigma_3^2 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 3 \end{pmatrix}$$

Firstly show that the covariance matrix is strictly positive definite in the sense that

$$\begin{pmatrix} x & y & z \end{pmatrix} \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} \\ \sigma_{12} & \sigma_2^2 & \sigma_{23} \\ \sigma_{13} & \sigma_{23} & \sigma_3^2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} > 0$$

if $\begin{pmatrix} x & y & z \end{pmatrix} \neq \begin{pmatrix} 0 & 0 & 0 \end{pmatrix}$. [It is sufficient to show that the resulting quadratic form is a sum of perfect squares.]

Deduce that the inverse of the covariance matrix is

$$\begin{pmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} \\ \sigma_{12} & \sigma_2^2 & \sigma_{23} \\ \sigma_{13} & \sigma_{23} & \sigma_3^2 \end{pmatrix}^{-1} = \frac{1}{2}\begin{pmatrix} 5 & -3 & 1 \\ -3 & 3 & -1 \\ 1 & -1 & 1 \end{pmatrix}$$

Assume that the expected returns on the risky assets are respectively,

$$\begin{pmatrix} R_1 \\ R_2 \\ R_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}.$$

Deduce that the boundary of opportunity set is given by

$$\sigma_{\Pi}^{2}(r) = \frac{1}{4}(3r^2 - 8r + 8),$$

where r is the prescribed level of expected return and $\sigma(r)$ is the minimal level of risk corresponding to the level of expected return. Further, deduce that the boundary of the opportunity set which can be achieved without short sales corresponds to $2 \leq r \leq \frac{8}{3}$.

Solution:

A symmetric matrix A is positive definite if for all $\underline{x} \neq \underline{0}$, $\underline{x}^{\mathrm{T}} A \underline{x} > 0$. Therefore using the covariance matrix given

$$(x,\ y,\ z)\begin{pmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 3 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = (x,\ y,\ z)\begin{pmatrix} x+y \\ x+2y+z \\ y+3z \end{pmatrix}$$

$= x^2 + 2y^2 + 3z^2 + 2xy + 2yz$
$= (x+y)^2 + (y+z)^2 + 2z^2$ (sum of squares)
$\Rightarrow$ for $(x,\ y,\ z) \neq \underline{0}$, the quadratic form above is always positive, hence the covariance matrix is strictly positive definite.

The inverse of the covariance matrix A is given above

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

Note that such a matrix has to be symmetric as $\sigma_{ij} = \sigma_{ji}$ and $\rho_{ij} = \rho_{ji}$.

Boundary of the opportunity set:

Create a portfolio $\Pi = \sum_{i=1}^{3} \lambda_i S_i$

$S_i = \begin{pmatrix} S_1 \\ S_2 \\ S_3 \end{pmatrix}$ are risky assets and $\lambda_i = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{pmatrix}$ is the allocation vector, i.e fraction of our total wealth being invested in each asset. So

$\sum_{i=1}^{3} \lambda_i = 1.$

The expected return is

$$\begin{aligned} R_\Pi &= \text{Expected return} = \underline{\lambda} . \underline{R} \\ &= (\lambda_1, \lambda_2, \lambda_3).(1, 2, 3) \\ &= \lambda_1 + 2\lambda_2 + 3\lambda_3 \\ &= r \end{aligned}$$

.

Now define the risk: $\sigma_\Pi^2 = \underline{\lambda}^T A \underline{\lambda} \Rightarrow \sigma_\Pi^2 = \lambda_1^2 + 2\lambda_2^2 + 3\lambda_3^2 + 2\lambda_1\lambda_2 + 2\lambda_2\lambda_3$ (we have used the working earlier, when we showed that the covariance matrix is +ve definite - with the vector $\underline{x}$ now replaced by the allocation vector $\underline{\lambda}$).

The problem here is to find an allocation vector $\underline{\lambda}$ (for a prescribed level of return r), so that σ_Π^2 is minimised under the restrictions

$$f(\underline{\lambda}) = \lambda_1 + 2\lambda_2 + 3\lambda_3 - r = 0 \qquad \text{which comes from } \underline{\lambda}.\underline{R} = r \text{ and}$$

$$g(\underline{\lambda}) = \lambda_1 + \lambda_2 + \lambda_3 - 1 = 0 \qquad \text{which is our total wealth } \sum_{i=1}^{3} \lambda_i = 1.$$

Introduce the Lagrange function L and Lagrange multipliers α, β:

$$L\left(\lambda_1,\ \lambda_2,\ \lambda_3,\ \alpha,\ \beta\right) = \sigma_{\Pi}^2 + \alpha g\,(\underline{\lambda}) + \beta f\,(\underline{\lambda}) \quad \Rightarrow$$

$$L = \underline{\lambda}^T A \underline{\lambda} + \beta(r - \underline{\lambda}.\underline{R}) + \alpha(1 - \underline{\lambda}.\underline{1})$$

where $\underline{1}$ is the unit vector.

Minimise L **with respect to** $\underline{\lambda}$, by differentiating L with respect to a vector quantity, and setting it to zero. We do this like any other differentiation, and note that $\underline{\lambda}^T A \underline{\lambda}$ behaves like $\underline{\lambda}^2 A$. Alternatively we can use the product rule on $\underline{\lambda}^T A \underline{\lambda}$.

$$\frac{\partial L}{\partial \underline{\lambda}} = 2\underline{\lambda} A - \beta \underline{R} - \alpha \underline{1} = 0 = (\underline{\lambda}^T A + A\underline{\lambda}) - \beta \underline{R} - \alpha \underline{1}$$

$$\rightarrow 2\underline{\lambda} A = \beta \underline{R} + \alpha \underline{1}$$

$$\rightarrow \underline{\lambda} = \frac{1}{2} A^{-1}(\alpha \underline{1} + \beta \underline{R})$$

$$\therefore \underline{\lambda} = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{pmatrix} = \frac{1}{2} A^{-1}(\alpha \underline{1} + \beta \underline{R})$$

$$= \frac{1}{4} \begin{pmatrix} 5 & -3 & 1 \\ -3 & 3 & -1 \\ 1 & -1 & 1 \end{pmatrix} \left(\alpha \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + \beta \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \right)$$

$$= \frac{1}{4} \begin{pmatrix} 3\alpha + 2\beta \\ -\alpha \\ \alpha + 2\beta \end{pmatrix}$$

OR

Calculate

$$\frac{\partial L}{\partial \lambda_1} = \frac{\partial L}{\partial \lambda_2} = \frac{\partial L}{\partial \lambda_3} = 0$$

for which we know

$$L = \lambda_1^2 + 2\lambda_2^2 + 3\lambda_3^2 + 2\lambda_1\lambda_2 + 2\lambda_2\lambda_3 - \alpha(\lambda_1 + \lambda_2 + \lambda_3 - 1) - \beta(\lambda_1 + 2\lambda_2 + 3\lambda_3 - r)$$

so

$$\left.\begin{aligned} \frac{\partial L}{\partial \lambda_1} &= 2\lambda_1 + 2\lambda_2 - \alpha - \beta = 0 \\ \frac{\partial L}{\partial \lambda_2} &= 4\lambda_2 + 2\lambda_1 + 2\lambda_3 - \alpha - 2\beta = 0 \\ \frac{\partial L}{\partial \lambda_3} &= 6\lambda_3 + 2\lambda_2 - \alpha - 3\beta = 0 \end{aligned}\right\} \text{Solve for } \lambda\text{'s in terms of } \alpha \text{ and } \beta$$

which gives

$$\underline{\lambda} = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{pmatrix} = \frac{1}{4}\left(\alpha\begin{pmatrix} 3 \\ -1 \\ 1 \end{pmatrix} + \beta\begin{pmatrix} 2 \\ 0 \\ 2 \end{pmatrix}\right) = \frac{1}{4}\begin{pmatrix} 3\alpha + 2\beta \\ -\alpha \\ \alpha + 2\beta \end{pmatrix}$$

and using these values of $\lambda_1, \lambda_2, \lambda_3$ allows us to calculate $\sigma_{\Pi}^2(\alpha, \beta)$.

We now minimize L with respect to α & β in turn so

$$\left.\begin{aligned}\frac{\partial L}{\partial \alpha} &= 0 \Rightarrow \underline{\lambda}.\underline{1} = \lambda_1 + \lambda_2 + \lambda_3 = 1 \\ \frac{\partial L}{\partial \beta} &= 0 \Rightarrow \underline{\lambda}.\underline{R} = \lambda_1 + 2\lambda_2 + 3\lambda_3 = r\end{aligned}\right\}.$$

So we have

$$\begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{pmatrix} = \frac{1}{4}\begin{pmatrix} 3\alpha + 2\beta \\ -\alpha \\ \alpha + 2\beta \end{pmatrix} \quad \text{and} \quad \left\{\begin{aligned} \lambda_1 + \lambda_2 + \lambda_3 &= 1 \\ \lambda_1 + 2\lambda_2 + 3\lambda_3 &= r \end{aligned}\right.$$

which gives us the Lagrange multipliers α and β in terms of r

$$\left.\begin{aligned}\alpha &= 4 - 2r \\ \beta &= \tfrac{3}{2}r - 2\end{aligned}\right\}.$$

Substituting the values of these two parameters in $\underline{\lambda}$ gives

$$\begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{pmatrix} = \frac{1}{4}\begin{pmatrix} 8 - 3r \\ 2r - 4 \\ r \end{pmatrix},$$

and we can obtain our expression for risk σ_{Π}^2 by substituting in these λ's and simplifying to give

$$\sigma_{\Pi}^2(r) = \frac{1}{4}(3r^2 - 8r + 8).$$

So for varying amounts of return r we can calculate the minimum risk. This is a hyperbola parameterised by r and is the **Boundary of the Opportunity Set.**

Excluding short sales means $\lambda_i \geq 0, \; i = 1, 2, 3$
So

$$8 - 3r \geq 0 \Rightarrow r \leq \frac{8}{3}$$

$$2r - 4 \geq 0 \Rightarrow r \geq 2$$

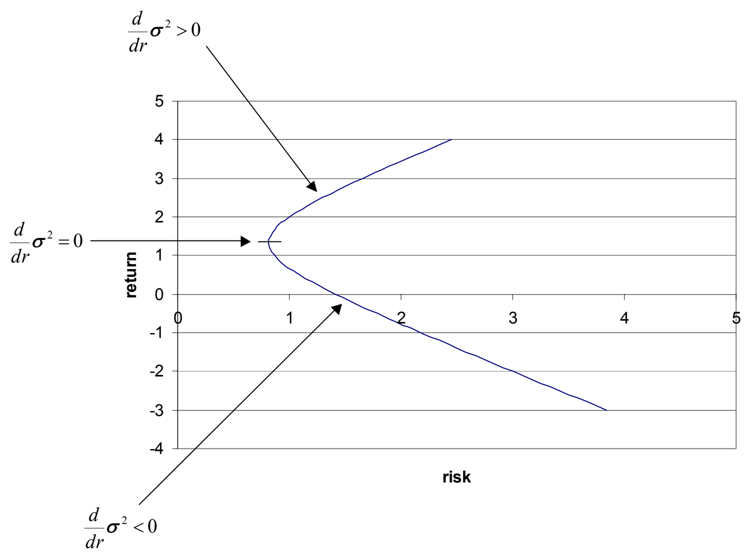
$$r \geq 0$$

and these three conditions are only satisfied when $2 \leq r \leq \frac{8}{3}$. Thus we can only achieve minimal risk without short sales if

$$2 \leq r \leq \frac{8}{3}.$$

Having obtained $\sigma_{\Pi}^2(r)$ we can calculate the **efficient frontier**.
For this we require

$$\frac{d}{dr}\sigma_{\Pi}^2 \geq 0$$



So

$$\frac{d}{dr}\sigma_\Pi^2 = \frac{1}{4}(6r - 8) = 0$$
$$\rightarrow r = \frac{4}{3}.$$

The minimum level of risk corresponding to $r = \frac{4}{3}$ is $\sigma_\Pi^2(4/3) = 2/3$.

We have the efficient frontier

$$\sigma_\Pi^2 = \frac{1}{4}(3r^2 - 8r + 8) \quad \text{for} \quad r \geq \frac{4}{3}.$$