Vectors and linear systems of equations

Geometry & Algebra

Vectors

- we use either bold or over arrow when noted by typing, the underline notation when written by hand
- The set of all ordered pairs of real numbers is denoted by $\mathbb{R}^2$, $\mathbb{R}^2 = \{(x, y) : x, y \in \mathbb{R}\}$. This is so called Cartesian setup.
- By vector, we mean a displacement between two or more points. It has a magnitude and a direction. Such as $\overrightarrow{AB}$ is the vector from A to B.
- Vector can be noted by:
  - $\mathbf{v} = \overrightarrow{v}$ = $\overrightarrow{AB} = [x, y]$
- Zero vector is a vector of zero displacement:
  - $\mathbf{0} = \overrightarrow{0} = \overrightarrow{AA}$

Vector addition

- In general if $\mathbf{u} = [u1, u2]$ and $\mathbf{v} = [v1, v2]$ then we define

$\mathbf{u} + \mathbf{v} = [u1 + v1, u2 + v2]$

- Ex.: We calculate $\mathbf{u} - \mathbf{v}$ where $\mathbf{u} = [2, 1]$ and $\mathbf{v} = [2, 3]$.

$\mathbf{u} - \mathbf{v} = [2 - 2, 1 - 3] = [0, -2]$

Scalar multiplication

- we scale up a vector
- Given vecotr $\mathbf{v} = [v_1, v_2]$ and $c \in \mathbb{R}$, the scalar multiple $c\mathbf{v}$ is

$$c\mathbf{v} = [cv_1, cv_2]$$

- Ex.: We calculate $c\mathbf{v}$ where $c = 3$ and $\mathbf{v} = [1, 2]$.

$$c\mathbf{v} = 3[1, 2] = [3 * 1, 3 * 2] = [3, 6]$$

Vector subtraction

- We write $-\mathbf{v}$ for $(-1)\mathbf{v}$, and then use this to define subtraction by

$$\mathbf{u} - \mathbf{v} = \mathbf{u} + (-\mathbf{v}) = [u_1 - v_1, u_2 - v_2]$$

- We add a vector with the same size, but the opposite direction.

Vectors in $\mathbb{R}^3$ and $\mathbb{R}^n$

- For $n \in \mathbb{N}, \mathbb{R}^n$ is the set of all n-tuples of real numbers,

$$\mathbb{R}^n = [x_1, x_2, ..., x_n : x_1, ..., x_n \in \mathbb{R}]$$

- A vector $\mathbf{v}$ in $\mathbb{R}^n$ can be written as a row or a column vector

$$[v_1, v_2, \ldots, v_n] = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

$$\textbf{Theorem 1.1 Algebraic properties of addition and scalar multiplication}$$
$$\text{(a) } \mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u} \text{ (commutativity of vector addition)}$$
$$\text{(b) } (\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w}) \text{ (associativity of vector addition)}$$
$$\text{(c) } \mathbf{u} + \mathbf{0} = \mathbf{u} \text{ (} \mathbf{0} \text{ is additive identity)}$$
$$\text{(d) } \mathbf{u} + (-\mathbf{u}) = \mathbf{0} \text{ (Every element has an inverse under addtion)}$$
$$\text{(e) } c(\mathbf{u} + \mathbf{v}) = c\mathbf{u} + c\mathbf{v} \text{ (Multiplication distributes across addition)}$$
$$\text{(f) } (c - d)\mathbf{u} = c\mathbf{u} - d\mathbf{u} \text{ (Note: these two +'s are different!)}$$
$$\text{(g) } c(d\mathbf{u}) = d(c\mathbf{u})$$
$$\text{(h) } 1\mathbf{u} = \mathbf{u}$$

Linear Combinations

- **Definition:** A vector $\mathbf{v}$ is a linear combination of vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$ if and only if there are scalars $c_1, c_2, \ldots, c_k$ such that

$$\mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k$$

The scalars $c_1, c_2, \ldots, c_k$ are called the coefficients of the linear combination

Linear equation

- **Definition:** A linear equation in the n variables $x_1, x_2, \ldots, x_n$ is an equation that can be written in the form

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b.$$

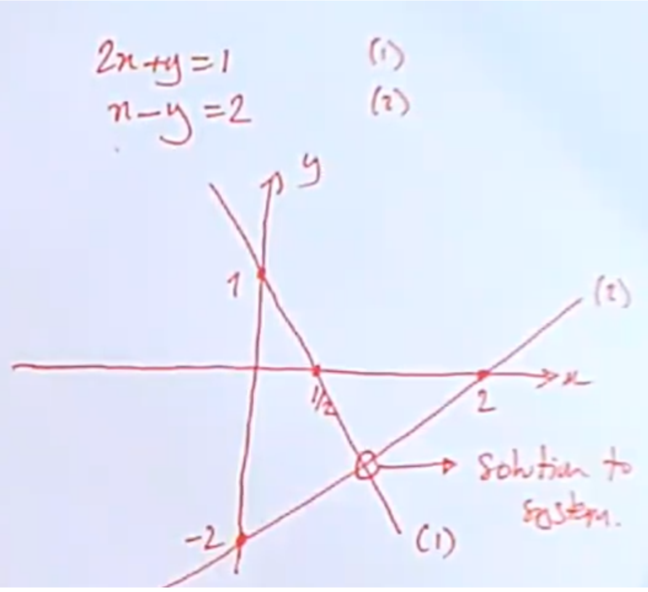
A solution of this linear equation is a vector $[s_1, s_2, \ldots, s_n]$ so that

$$a_1s_1 + a_2s_2 + \cdots + a_ns_n = b.$$

System of linear equations

A system of linear equations is a finite set of linear equations and a solition to a system of linear equations is a vector that is simultaneously a solution to all equations in the system.

- Ex.: We find all solutions to the system $2x + y = 1$, $x - y = 2$, including a useful sketch.



$$\begin{gathered} 2x + y = 1 \\ x - y = 2 \\ \Rightarrow x = 2 + y \\ \Rightarrow 2 * (2 + y) + y = 1 \\ \Rightarrow 4 + 3y = 1 \\ \Rightarrow y = -1 \\ \Rightarrow x = 1 \end{gathered}$$

- A system can either have infinitely many solutions, a unique solution or no solutions at all

Solving Linear Systems

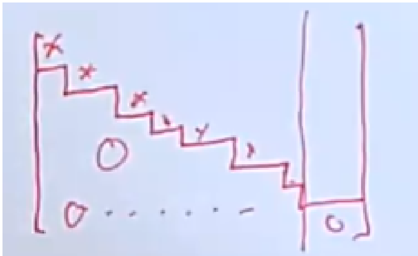
- A linear system

$$\begin{gathered} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1, \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m, \end{gathered}$$

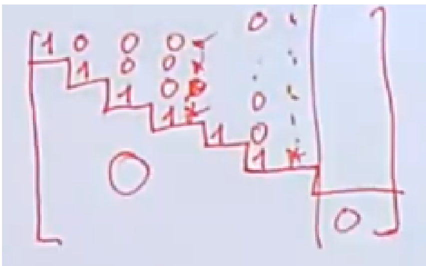
has augmented matrix

$$[A|\mathbf{b}] = \left[\begin{array}{cccc|c} a_{11} & a_{12} & \ldots & a_{1n} & b_1 \\ a_{21} & a_{22} & \ldots & a_{2n} & b_2 \\ & & \vdots & & \\ a_{m1} & a_{m2} & \ldots & a_{mn} & b_m \end{array}\right].$$

- **Definition:** A matrix is in row echelon form if and only if:
    - any all-zero rows are at the bottom
    - in each non-zeor row, the first non-zero entry (the leading entry) is to the left of any entries below it.



- Row reduction is the process for reducing any matrix to row-echelon form, using elementary row operations (EROs). It should be familiar from last year
- **Definition:** A matrix is in reduced row echelon form if and only if:
    - it is in row echelon form
    - the leading entry in eachh nonzero row is 1
    - each column containing a leading 1 has zeros everywhere else.



- Definition: A system of linear equations is homogeneous if and only if the constant term (i.e. b_i above) in each equation is zero.

Spanning sets

- Any 3-dimensional vector can be written as a combination of **i**, **j**, **k**
- Definition: If $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ is a set of vectors in $\mathbb{R}^n$, then the set of all linear combinations of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ is called the span of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ and is denoted $span\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ or span S. Symbolically,

$$span\, S = \{\sum_{i=1}^{k} \mathbf{v}_i c_i : c_1, c_2, \ldots, c_k \in \mathbb{R}\}.$$

- Ex.: Is $\begin{bmatrix} 8 \\ -5 \\ -15 \end{bmatrix}$ in the span of $\{\begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} -2 \\ 3 \\ 5 \end{bmatrix}\}$? Equivalently, is $\begin{bmatrix} 8 \\ -5 \\ -15 \end{bmatrix}$ a linear combination of $\begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} -2 \\ 3 \\ 5 \end{bmatrix}$?

$$\begin{bmatrix} 8 \\ -5 \\ -15 \end{bmatrix} = c_1 \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} + c_2 \begin{bmatrix} -2 \\ 3 \\ 5 \end{bmatrix}$$
$$-15 = 5c_2 \Rightarrow c_2 = -3$$
$$8 = c_1 + 6 \Rightarrow c_1 = 2$$
$$-5 = 2 * 2 + 3 * (-3)$$
$$\Rightarrow -5 = 4 - 9 = -5$$

The given vector belongs to the span.

- **Theorem 2.4** A system of linear eqautions with augmented matrix $[A|\mathbf{b}]$ is consistent if and only if $\mathbf{b}$ is a linear combination of the columns of A.

$$a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1$$
$$a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2$$
$$\vdots$$
$$a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m$$
$$x_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} + \cdots + x_n \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

$Proof.$ Not examinable.

- Ex.: Is the system $x + y = 1,\ x - y = 1$ consistent?

$$\begin{bmatrix} 1 & 1 \,|\, 1 \\ 1 & -1 \,|\, 1 \end{bmatrix}$$
$$\begin{bmatrix} 1 \\ 1 \end{bmatrix} = c_1 \begin{bmatrix} 1 \\ 1 \end{bmatrix} + c_2 \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$
$$1 = c_1 + c_2 \qquad c_1 = 1$$
$$1 = c_1 - c_2 \qquad c_2 = 0$$
$$\Rightarrow \text{The system is consistent.}$$

- Ex.: Is the system $x+y=1,\ 2x+2y=2$ consistent?

$$\begin{bmatrix} 1 & 1 & | & 1 \\ 2 & 2 & | & 2 \end{bmatrix}$$
$$\begin{bmatrix} 1 \\ 2 \end{bmatrix} = c_1 \begin{bmatrix} 1 \\ 2 \end{bmatrix} + c_2 \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$
$$1 = c_1 + c_2 \qquad c_1 = t$$
$$2 = 2c_1 + 2c_2 \qquad c_2 = 1 - t$$
$$\Rightarrow \text{The system is consistent.}$$

- Ex.: Is the system $x+y=1,\ x+y=0$ consistent?

$$\begin{bmatrix} 1 & 1 & | & 1 \\ 1 & 1 & | & 0 \end{bmatrix}$$
$$\begin{bmatrix} 1 \\ 0 \end{bmatrix} = c_1 \begin{bmatrix} 1 \\ 1 \end{bmatrix} + c_2 \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$
$$1 = c_1 + c_2$$
$$0 = c_1 + c_2$$
$$\Rightarrow \text{The system is inconsistent.}$$

- **Definition:** Let $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ be a set of vectors in $\mathbb{R}^n$. S is a *spanning set* for $\mathbb{R}^n$ if and only if $span\ S = \mathbb{R}^n$.
  - That is, S is a spanning set for $\mathbb{R}^n$ if and only if every vector in $\mathbb{R}^n$ can be written as a linear combination of elemnets of S. This combination might not be unique.
  - Ex.: Show that $\mathbb{R}^2 = span \left\{ \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right\}$

$$x = 2c_1 + c_2$$
$$y = c_1 - c_2$$
$$\text{We add these two equations to get:}$$
$$x + y = 3c_1$$
$$c_1 = \frac{1}{3}(x+y)$$
$$c_2 = c_1 - y = \frac{1}{3}x + \frac{1}{3}y - y = \frac{1}{3}x - \frac{2}{3}y$$

  - Validation:

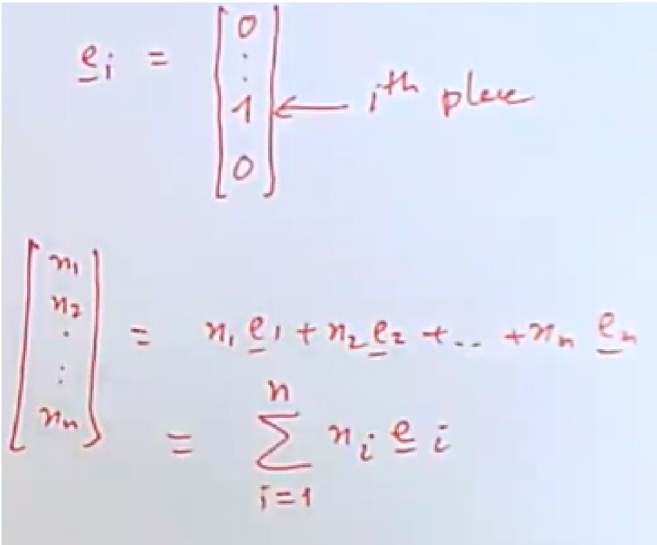
$$2c_1 + c_2 = \frac{2}{3}(x+y) + \frac{1}{3}x - \frac{2}{3} = x$$

$\Rightarrow \mathbb{R}^2 = span \left\{ \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right\}$

- Ex.: In $\mathbb{R}^2$ let $\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$. Show that $\mathbb{R}^2 = span\{\mathbf{e}_1, \mathbf{e}_2\}$.

$$\begin{bmatrix} x \\ y \end{bmatrix} = x\mathbf{e}_1 + y\mathbf{e}_2$$

- Ex.: In $\mathbb{R}^n$ let $\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$, $\mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}$, ..., $\mathbf{e}_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$. Show that $\mathbb{R}^n = span\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$.



Linear independence

Definition A set $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ of vectors in $\mathbb{R}^n$ is linearly independent if and only if the only solution to the equation

$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k = 0, \qquad c_i \in \mathbb{R}$

is $c_1 = c_2 = \cdots = c_k = 0$. A set of vectors is linearly dependent if it is not linearly independent.

Ex.: Are the vectors $\mathbf{u} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ and $\mathbf{w} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ linearly independent?

$$c_1 \begin{bmatrix} 2 \\ 1 \end{bmatrix} + c_2 \begin{bmatrix} 1 \\ -1 \end{bmatrix} + c_3 \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$
$$2c_1 + c_2 + c_3 = 0$$
$$c_1 - c_2 + 2c_3 = 0$$

$$\begin{bmatrix} 2 & 1 & 1 & | & 0 \\ 1 & -1 & 2 & | & 0 \end{bmatrix} \rightarrow^{R_1 \leftrightarrow R_2} \begin{bmatrix} 1 & -1 & 2 & | & 0 \\ 2 & 1 & 1 & | & 0 \end{bmatrix} \rightarrow^{R_2 \rightarrow R_2 - 2R_1} \begin{bmatrix} 1 & -1 & 2 & | & 0 \\ 0 & 3 & -3 & | & 0 \end{bmatrix}$$
$$\rightarrow^{R_2 \rightarrow R_2/3} \begin{bmatrix} 1 & -1 & 2 & | & 0 \\ 0 & 1 & -1 & | & 0 \end{bmatrix} \rightarrow^{R_1 \rightarrow 3R_1 + R_2} \begin{bmatrix} 1 & 0 & 1 & | & 0 \\ 0 & 1 & -1 & | & 0 \end{bmatrix}$$

$$\begin{aligned} c_1 + c_3 &= 0 \\ c_2 - c_3 &= 0 \end{aligned}$$

- We can choose c_3 to be whatever we like.

$$c_3 = 1 \Rightarrow c_1 = -1,\ c_2 = 1.$$

Hence, $\mathbf{u}, \mathbf{v}, \mathbf{w}$ are linearly dependent.

Theorem 2.5 Vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ in $\mathbb{R}^n$ are linearly dependent if and only if at least one of them can be expressed as a linear combination of the others.

$Proof$:

Assume $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ are linearly dependent.
There exists $c_1, c_2, \ldots, c_k$ not all zero such that $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k = 0$.

Assume $c_1 \neq 0$. So $\mathbf{v}_1 + \frac{c_2}{c_1}\mathbf{v}_2 + \cdots + \frac{c_k}{c_1}\mathbf{v}_k = 0$.

$$\mathbf{v}_1 = -\frac{c_2}{c_1}\mathbf{v}_2 - \cdots - \frac{c_k}{c_1}\mathbf{v}_k$$

Conversely, assume $\mathbf{v}_1$, say, can be expressed:

$\mathbf{v}_1 = c_2\mathbf{v}_2 + c_3\mathbf{v}_3 + \cdots + c_k\mathbf{v}_k$

Then

$$\begin{gathered} c_1'\mathbf{v}_1 + c_2'\mathbf{v}_2 + \cdots + c_k'\mathbf{v}_k = \mathbf{0} \\ c_1' = 1, c_i' = -c_i \quad i = 2, \ldots, k \end{gathered}$$

Any set of vectors containing the zero vector is linearly dependent. Suppose $\mathbf{v}_1 = 0$.
Then we just choose $c_1 \neq 0$.

$$1\mathbf{0} + 0\mathbf{v}_2 + 0\mathbf{v}_m = \mathbf{0}.$$

Lemma Two vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ are linearly dependent if and only if they are scalar multiples of each other.

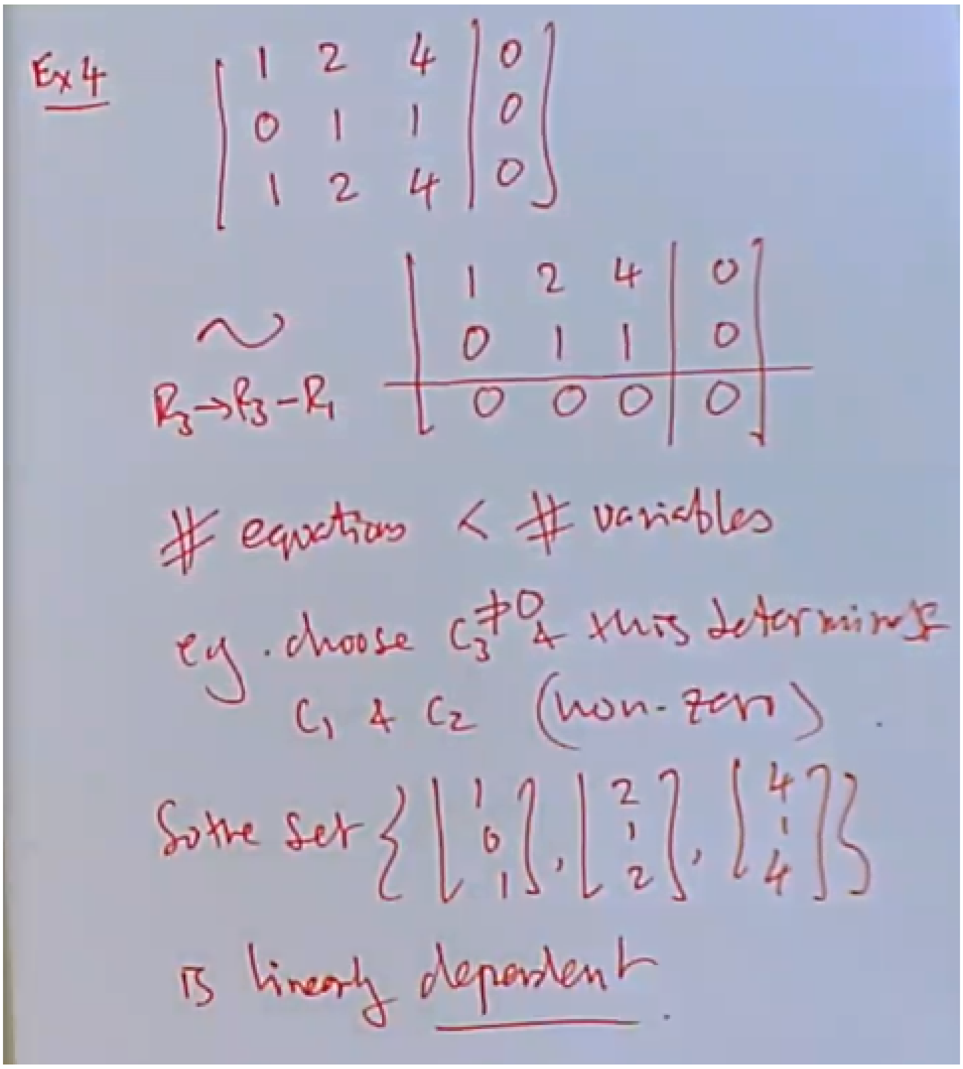
$$\begin{gathered} c_1\mathbf{v}_1 + c_2\mathbf{v}_2 = 0 \qquad c_1, c_2 \text{ not both } 0 \\ c_1 \neq 0 \Rightarrow \mathbf{v}_1 = -\frac{c_2}{c_1}\mathbf{v}_2 \\ \text{Conversely if } \mathbf{v}_1 = \lambda\mathbf{v}_2 \text{ then } \mathbf{v}_1 + (-\lambda)\mathbf{v}_2 = 0,\ \lambda \neq 0. \end{gathered}$$

Ex.: Determine whether the set $\left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \right\}$ is linearly independent.

$$\begin{bmatrix} 1 & 0 & 1 & | & 0 \\ 1 & 1 & 0 & | & 0 \\ 0 & 1 & 1 & | & 0 \end{bmatrix} \to c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3 = 0$$
$$\to^{R_2 \to R_2 - R_1} \begin{bmatrix} 1 & 0 & 1 & | & 0 \\ 0 & 1 & -1 & | & 0 \\ 0 & 1 & 1 & | & 0 \end{bmatrix}$$
$$\to^{R_3 \to R_3 - R_2} \begin{bmatrix} 1 & 0 & 1 & | & 0 \\ 0 & 1 & -1 & | & 0 \\ 0 & 0 & -2 & | & 0 \end{bmatrix}$$
$$\begin{aligned} c_1 \quad + c_3 &= 0 \\ c_2 - c_3 &= 0 \\ -2c_3 &= 0 \\ \Rightarrow c_1 = 0, c_2 = 0, c_3 &= 0. \end{aligned}$$

The set is linearly independent.

Example 4 Determine whether the set $\left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 4 \\ 1 \\ 4 \end{bmatrix} \right\}$ is linearly independent.



Theorem 2.6 Let $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$ be (column) vectors in $\mathbb{R}^n$ and let

$$A = [\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m]$$

be the $n \times m$ matrix with columns $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$. Then $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$ are linearly dependent if and only if the homogenenous linear system with augmented matrix $[A|\mathbf{b}]$ has non-trivial solution.

$Proof$:

- From Theorem 2.4:
  - $[A|\mathbf{b}]$ has a solution iff $\mathbf{b} \in span\{\text{columns of A}\}$
  - $\mathbf{b} = \mathbf{0}$
  - $[A|\mathbf{0}]$ has a solution iff $\mathbf{0} \in span\{\text{columns of A}\}$. If this a non-trivial solution (not all the $c_i = 0$) then the columns are linearly dependent.

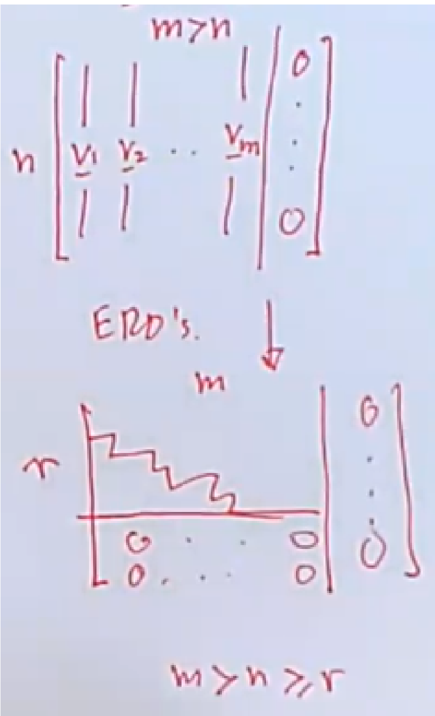
Theorem 2.8 If $m > n$ then any set of m vectors in $\mathbb{R}^n$ is linearly dependent.

$Proof$:

$$\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m \in \mathbb{R}^n$$

Consider

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_m\mathbf{v}_m = \mathbf{0}$$



There must be a non-zero solution to this set since only r (number of non-zero rows) of the c_i are determined by the other $m - r > 0$.

Ex.: Show that the set $\begin{bmatrix}1\\1\end{bmatrix}$, $\begin{bmatrix}5\\7\end{bmatrix}$ and $\begin{bmatrix}216\\16\end{bmatrix}$ are linearly dependent in $\mathbb{R}^2$.

$$m = 3,\ n = 2 \to m > n$$

Matrix Operations

Matrices

A matrix is a rectangular array of numbers called the entries or elements of the matrix

For a matrix A we write the i, j^{th} entry as a_{ij}.

For the matrix $A = \begin{bmatrix} 2 & 0 & 5 \\ 1 & 4 & -1 \end{bmatrix}$1 then $a_{11} = 2,\ a_{23} = -1,\ a_{21} = 1 \text{ and } a_{12} = 0.$

Symbolically this is expressed by $A = (a_{ij})$. So if A

is an $m \times n$ (m rows, n columns) matrix then

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

If the columns of A are column vectors $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ belonging to $\mathbb{R}^m$ we write

$$A = [\mathbf{a}_1 \dots \mathbf{a}_n],$$

and if the rows of A are row vectors $\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_m$ belonging to $\mathbb{R}^n$ we write

$$A = \begin{bmatrix} \mathbf{b}_1 \\ \vdots \\ \mathbf{b}_m \end{bmatrix}.$$

A square matrix is $square$ if $m = n$ (the number of rows is equal to the number of columns) and is $diagonal$ iff its off-diagonal elements are zero (for $i \neq j$, $a_{ij} = 0$). The $n \times n$ identity matrix $\mathbb{I}_n$ is the diagonal matrix with all diagonal entries equal to 1.

<u>Example 1</u> Of the matrices $A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 4 & 4 \\ 0 & 0 & 8 \end{bmatrix}$, $B = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$ and $C = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, B and C are diagonal and $C = \mathbb{I}_3$.

Two matrices are $equal$ if and only if they have the same number of rows and columns and the corresponding entries are equal.

Matrix addition and scalar multiplication

IF $A=(a_{ij})$ and $B=(b_{ij})$ are $m \times n$ matrices and c is a scalar then we define new $m \times n$ matrices $A+B$ and cA componentwise:

$$\begin{aligned} A+B &= (a_{ij})+(b_{ij}) \\ &= (a_{ij}+b_{ij}) \\ cA &= c(a_{ij}) \\ &= (ca_{ij}). \end{aligned}$$

Matrix multiplication

If A is an $m \times n$ matrix and B is an $n \times r$ matrix then $C=AB$ is the $m \times r$ matrix with $(i,j)^{th}$ entry given by

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj} = \sum_{k=1}^{n} a_{ik}b_{kj}.$$

Theorem 3.1 Let A be an $m \times n$ matrix, and

$$\mathbf{e}_i = [0, \cdots, 0, 1, 0, \cdots, 0],$$

with the 1 in the i^{th} position, and

$$\mathbf{f}_j = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix},$$

with the 1 in the j^{th} position. Then

(a) $\mathbf{e}_i A$ is the i^{th} row of A,

(b) $A\mathbf{f}_j$ is the j^{th} column of A .

Matrix powers

If A is an $n \times n$ matrix and k is a positive integer then

$A^0 = \mathbb{I}_n$, $A^2 = AA$, $A^k = AA \dots A \leftarrow k$-times

Example 2 For $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ we compute A^2, A^3 and A^4. The most obvious conjecture for A^n is, in fact, correct.

$$
\begin{aligned}
A &= \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \\
A^2 &= \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} \\
A^3 &= \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix} \\
A^4 &= \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 4 \\ 0 & 1 \end{bmatrix}
\end{aligned}
$$

Conjecture:
$A^k = \begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix}$

Transpose of a matrix

The transpose of an $m \times n$ matrix A is the $n \times m$ matrix A^T obtained by interchanging the rows and columns of A.

Ex.: We write down the transposes of the matrices

$$A = \begin{bmatrix} 2 & 3 & 1 \\ 1 & 2 & 4 \end{bmatrix} \text{ and } B = \begin{bmatrix} 2 & 0 \\ 1 & 3 \end{bmatrix}.$$

$$
\begin{aligned}
A^T &= \begin{bmatrix} 2 & 1 \\ 3 & 2 \\ 1 & 4 \end{bmatrix} \\
B^T &= \begin{bmatrix} 2 & 1 \\ 0 & 3 \end{bmatrix}
\end{aligned}
$$

Symmetric matrix

A square matrix is *symmetric* if $A^T = A$. That is, A is equal to its own transpose.

Ex.: Only one of the matrices

$$A = \begin{bmatrix} 2 & 0 \\ 1 & 3 \end{bmatrix} \text{ and } B = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 3 & 7 \\ 1 & 7 & 5 \end{bmatrix}$$

is symmetric.

We can see that

$$A^T = \begin{bmatrix} 2 & 1 \\ 0 & 3 \end{bmatrix} \neq \begin{bmatrix} 2 & 0 \\ 1 & 3 \end{bmatrix} = A$$

so A is not symmetric and B is as

$$B^T = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 3 & 7 \\ 1 & 7 & 5 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 3 & 7 \\ 1 & 7 & 5 \end{bmatrix} = B.$$

$A = A^T \Leftrightarrow a_{ij} = a_{ji}$

Matrix Algebra

The addition and scalar multiplication rules for $m \times n$matrices obey the following algrebraic properties.

Theorem 3.2 Let A, B and C be $m \times n$ matrices and c and d be scalars. Then

1. $A + B = B + A,$
2. $(A + B) + C = A + (B + C),$
3. $A + 0 = A,$
4. $A + (-A) = 0,$
5. $c(A + B) = cA + cB,$
6. $(c + d)A = cA + dA,$
7. $c(dA) = (cd)A$
8. $1A = A.$

Spans & Linear Independence

Inner product of vectors

Let A be $m \times n$ matrix and B an $r \times n$ matrix. The product of these matrices is $m \times r$ matrix.

But what if $m = r = 1$, the product would be just a single number.

We make these definitions in exactly the same way for matrices as we did for vectors in $\mathbb{R}^n$.

Definition If $A_1, \ldots, A_k$ are $m \times n$ matrices and $c_1, \ldots, c_k$ are scalars, we may form the *linear combination*

$$c_1 A_1 + c_2 A_2 + \cdots + c_k A_k$$

Example 1 The matrix $B = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ is not a linear combination of $A_1 = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$, $A_2 = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ and $A_3 = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$?

$$\begin{bmatrix}1 & 2\\3 & 4\end{bmatrix} = c_1\begin{bmatrix}1 & 1\\0 & 1\end{bmatrix} + c_2\begin{bmatrix}0 & 1\\-1 & 0\end{bmatrix} + c_3\begin{bmatrix}1 & 0\\1 & 1\end{bmatrix}$$
$$= \begin{bmatrix}c_1 + c_3 & c_1 + c_2\\c_3 - c_2 & c_1 + c_3\end{bmatrix}$$

$$\begin{matrix}1 = c_1 + c_3 & 2 = c_1 + c_2\\3 = c_3 - c_2 & 4 = c_1 + c_3\end{matrix}$$

Is there a solution?

$$\begin{bmatrix}1 & 0 & 1 & 0 & | & 1\\1 & 1 & 0 & 0 & | & 2\\0 & -1 & 1 & 0 & | & 3\\1 & 0 & 1 & 0 & | & 4\end{bmatrix}$$

As first and last row are in contradiction, this augmented matrix does not have a solution. ($1 \neq 4$)

Span of a set

If $S = \{A_1, \ldots, A_k\}$ is a set of $m \times n$ matrices, then the *span* of S is the set of all linear combinations of the elements of S.

That is,

$$Span\ S = \{c_1A_1 + \cdots + c_kA_k : c_1, \ldots, c_k \text{ are scalars}\}$$
$$= \left\{\sum_{i=1}^{k} c_iA_i : c_1 \ldots, c_k \text{ are scalars}\right\}.$$

Ex.: In the previous example does B belong to the span of A_1, A_2, A_3?

→ No

Ex.: Describe $Span\{A_1, A_3\}$ in the previous two examples.

$$Span\{A_1, A_3\} = \{c_1A_1 + c_2A_3 | c_1, c_2 \in \mathbb{R}\}$$

$$c_1A_1 + c_2A_3 = c_1\begin{bmatrix}1 & 1\\0 & 1\end{bmatrix} + c_2\begin{bmatrix}1 & 0\\1 & 1\end{bmatrix} = \begin{bmatrix}c_1 + c_2 & c_1\\c_2 & c_1 + c_2\end{bmatrix}$$

$$A = \begin{bmatrix}a_{11} & a_{12}\\a_{21} & a_{22}\end{bmatrix}$$
$$a_{11} = a_{22} = a_{12} + a_{21}$$
$$Span\{A_1, A_3\} = \left\{\begin{bmatrix}a_{11} & a_{12}\\a_{21} & a_{22}\end{bmatrix} | a_{11} = a_{22} = a_{12} + a_{21}\right\}$$

Linear independence

A collection $\{A_1, \ldots, A_k\}$ of $m \times n$ matrices is *linearly independent* if the only solution to the equation

$$c_1A_1 + c_2A_2 + \cdots + c_kA_k = 0$$

for scalars $c_1, \ldots, c_k$ is $c_1 = c_2 = \cdots = c_k = 0$. If there are non-trivial coefficients which satisfy this equation then the set $\{A_1, \ldots, A_k\}$ is *linearly dependent*.

Ex.: Determine whether the set $\{A_1, A_2, A_3\}$ from the above examples, is linearly independent.

$$c_1\begin{bmatrix}1 & 1\\ 0 & 1\end{bmatrix} + c_2\begin{bmatrix}0 & 1\\ -1 & 0\end{bmatrix} + c_3\begin{bmatrix}1 & 0\\ 1 & 1\end{bmatrix} = 0$$

$$\begin{bmatrix}c_1 + c_3 & c_1 + c_2\\ c_3 - c_2 & c_1 + c_3\end{bmatrix} = \begin{bmatrix}0 & 0\\ 0 & 0\end{bmatrix}$$

$$\begin{bmatrix}1 & 0 & 1 & 0 & | & 0\\ 1 & 1 & 0 & 0 & | & 0\\ 0 & -1 & 1 & 0 & | & 0\\ 1 & 0 & 1 & 0 & | & 0\end{bmatrix} R_2 \to R_2 - R_1,\ R_4 \to R_4 - R_1$$

$$\begin{bmatrix}1 & 0 & 1 & 0 & | & 0\\ 0 & 1 & -1 & 0 & | & 0\\ 0 & -1 & 1 & 0 & | & 0\\ 0 & 0 & 0 & 0 & | & 0\end{bmatrix} R_3 \to R_3 + R_2$$

$$\begin{bmatrix}1 & 0 & 1 & 0 & | & 0\\ 0 & 1 & -1 & 0 & | & 0\\ 0 & 0 & 0 & 0 & | & 0\\ 0 & 0 & 0 & 0 & | & 0\end{bmatrix} R_3 \to R_3 + R_2$$

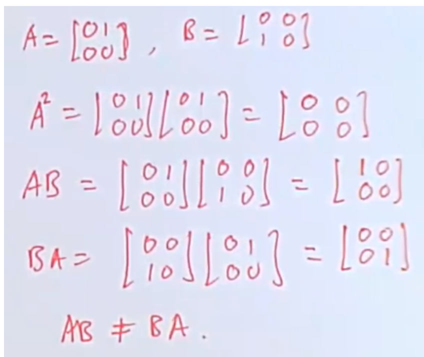
$$c_1 + c_3 = 0$$
$$c_2 - c_3 = 0$$

Choose $c_3 = 1$, say. Then $c_1 = -1$ and $c_2 = 1$. The set $\{A_1, A_2, A_3\}$ is linearly dependent.

Multiplication and Inverse Matrices

Properties of matrix multiplication

Matrix multiplication behaves differently from multiplication of numbers. In general, multiplication is not commutative. Also, we could have $A^2 = 0$ even if $A \neq 0$.



Theorem 3.3 Let A, B and C be matrices and k be a scalar. The following identities hold whenever the operations involved can be performed.

1. $A(BC) = (AB)C$, (associativity of matrix multiplication)
2. $A(B + C) = AB + AC$, (left multiplication disctributes across addition)
3. $(A + B)C = AC + BC$, (right multiplication distributes across addition)
4. $k(AB) = (kA)B = A(kB)$, (scalar multiplication commutes with matrix multiplication)
5. $\mathbb{I}_m A = A = A\mathbb{I}_n$ if A is $m \times n$ (left/right multiplicative identities).

Proof: 1)

$$(A(BC))_{ij} = \sum_k A_{ik}(BC)_{kj} = \sum_k \sum_l A_{ik} B_{kl} C_{lj}$$
$$((AB)C)_{ij} = \sum_l (AB)_{il} C_{lj} = \sum_l \sum_k A_{ik} B_{kl} C_{lj}$$

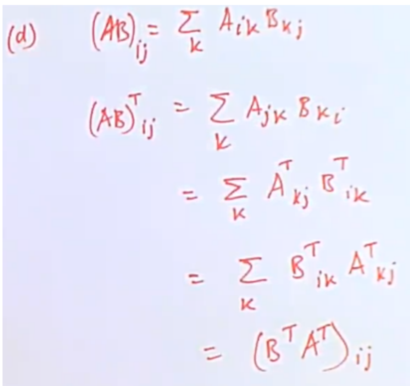
Interchange order of summation:

$$A(BC) = (AB)C$$

Properties of transpose matrices

Similarly we have algebraic rules for transpose.
Theorem 3.4 Let A and B be matrices. The following identities hold whenever the operation involved can be performed.

1. $(A^T)^T = A$,
2. $(A + B)^T = A^T + B^T$,
3. $(kA)^T = k(A^T)$,
4. $(AB)^T = B^T A^T$,



5. $(A^m)^T = (A^T)^m$ for all integer $m \geq 0$.

Theorem 3.5

(a) If A is a square matrix then $A + A^T$ is a symmetric matrix,

(b) For any matrix A, AA^T and $A^T A$ are symmetric matrices.

Proof: of (b)

$$(AA^T)^T = (A^T)^T A^T = AA^T$$

AA^T is symmetric.

The inverse of a matrix

Definition If A is an $n \times n$ matrix, the *inverse* of A is an $n \times n$ matrix A' such that

$$AA' = \mathbb{I}_n, \qquad \text{and} \qquad A'A = \mathbb{I}_n.$$

If A' exists we say A is *invertible*. If no inverse exists, then we say that A is not invertible.

Ex.: Consider the matrices $A = \begin{bmatrix} 1 & 2 \\ 3 & 7 \end{bmatrix}$ and $B = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$. We write down an inverse for A and see what goes wrong if we try to find an inverse for B.

$$\begin{gathered} \begin{bmatrix} 1 & 2 \\ 3 & 7 \end{bmatrix} \begin{bmatrix} a' & b' \\ c' & d' \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\ a' + 2c' = 1 \quad\quad b' + 2d' = 0 \\ 3a' + 7c' = 0 \quad\quad 3b' + 7d' = 1 \end{gathered}$$

$\rightarrow a' = 7, d' = 1, b' = -2, c' = -3.$

$$A' = \begin{bmatrix} 7 & -2 \\ -3 & 1 \end{bmatrix}$$

$$\begin{gathered} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} a' & b' \\ c' & d' \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\ c' = 1 \quad\quad d' = 0 \quad\quad 0 = 0 \\ 0 = 1 \rightarrow \text{contradiction} \end{gathered}$$

Theorem 3.6 If an $n \times n$ matrix is invertible then its inverse is unique.

Proof: Suppose A', A'' are both inverses of A.

$$\begin{gathered} A'A = AA' = \mathbb{I}_n \\ A''A = AA'' = \mathbb{I}_n \\ A'' = \mathbb{I}_n A'' = A'AA'' = A'\mathbb{I}_n = A' \\ A'' = A' \end{gathered}$$

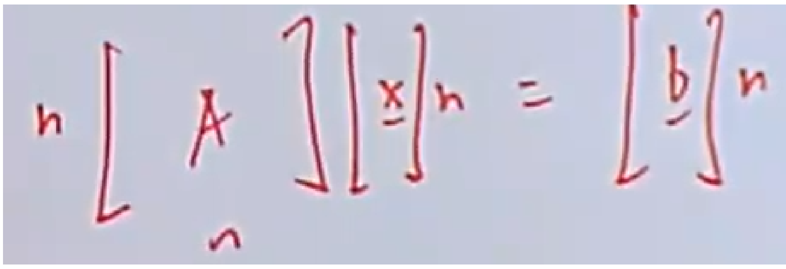
Inverse is unique.

Notation If A is invertible we write A^{-1} for its inverse.

Important Warning. We **do not** write $\frac{1}{A}$ for the inverse of A. That is a notation reserved for non-zero numbers and matrices are not numbers.

Theorem 3.7 If A is an invertible $n \times n$ matrix then the system of linear equations given by $A\mathbf{x} = \mathbf{b}$ has the unique solution given by $\mathbf{x} = A^{-1}\mathbf{b}$.

Proof:



A invertible $\Rightarrow \mathbf{x} = A^{-1}\mathbf{b}$ unique

$$\begin{gathered} A\mathbf{x} = \mathbf{b} \\ A^{-1}A\mathbf{x} = A^{-1}\mathbf{b} \\ \mathbf{x}A^{-1}\mathbf{b} \end{gathered}$$

Suppose not unique:

$$\begin{gathered} A\mathbf{x} = \mathbf{b} \\ A\mathbf{x}' = \mathbf{b} \\ \Rightarrow A(\mathbf{x} - \mathbf{x}') = \mathbf{b} - \mathbf{b}' = \mathbf{0} \\ A^{-1}A(\mathbf{x} - \mathbf{x}') = A^{-1}\mathbf{0} = \mathbf{0} \\ \mathbf{x} - \mathbf{x}' = \mathbf{0} \\ \mathbf{x} = \mathbf{x}' \end{gathered}$$

Theorem 3.8 If $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ then A is invertible if $ad - bc \neq 0$, in which case

$$A^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

If $ad - bc = 0$ then A is not invertible.

Proof:

$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} a' & b' \\ c' & d' \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$
$(11)\ aa' + bc' = 1 \quad ab' + bd' = 0\ (21)$
$(12)\ ca' + dc' = 0 \quad cb' + dd' = 1\ (22)$

$$(11) \times d - (12) \times b = ada' - bca' = d$$
$$(ad - bc)a' = d$$
$$(11) \times c - (12) \times a = c$$
$$((bc - ad)c' = c$$
$$a' = \frac{d}{\Delta} \quad \Delta = ad - bc$$
$$\Delta \text{ is called Determinant}$$
$$c' = -\frac{c}{\Delta}$$
$$d' = \frac{a}{\Delta}$$
$$b' = -\frac{b}{\Delta}$$

$\Delta \neq 0$ allows us to construct a', b', c', d'

$\Delta = 0$: no solution & no inverse

For general $n \times n$ matrices, there are algorithms (using row reduction) for finding the inverse, but not convenient general fomula as we have in the 2×2 case.

Definition For a 2×2 matrix

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix},$$

we call $ad - bc$ the determinant of A, so that

$$det(A) = ad - bc.$$

Ex.: We find the inverses, if they exist, of $A = \begin{bmatrix} 1 & 5 \\ 2 & 4 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 5 \\ 2 & 10 \end{bmatrix}$.

$det(A) = 1 \times 4 - 2 \times 5 = 4 - 10 = -6 \neq 0$
$det(B) = 1 \times 10 - 5 \times 2 = 10 - 10 = 0$
$\Rightarrow \text{ B is not inversible.}$

$$A' = \frac{1}{-6} \begin{bmatrix} 4 & -5 \\ -2 & 1 \end{bmatrix} = \frac{1}{6} \begin{bmatrix} -4 & 5 \\ 2 & -1 \end{bmatrix}$$

Ex.: Solve the system

$$x + 5y = 3, \ 2x + 4y = 1$$

using the inverse of the coefficient matrix.

$$\begin{bmatrix} 1 & 5 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$$
$$\begin{bmatrix} x \\ y \end{bmatrix} = \frac{1}{6} \begin{bmatrix} -4 & 5 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} 3 \\ 1 \end{bmatrix}$$
$$= \frac{1}{6} \begin{bmatrix} -12+5 \\ 6-1 \end{bmatrix}$$
$$= \begin{bmatrix} -7/6 \\ 5/6 \end{bmatrix}$$

Theorem 3.9

1. If A is an invertible matrix, then A^{-1} is invertible and $(A^{-1})^{-1} = A$.

   Proof: $AA^{-1} = A^{-1}A = \mathbb{I}_n$

   $A^{-1} = B$

   $AB = BA = \mathbb{I}_n$

   So $A = B^{-1}$ (Definition)

   $A = (A^{-1})^{-1}$

2. If A is an invertible matrix and $c \neq 0$ is a scalar then cA is invertible and $(cA)^{-1} = \frac{1}{c}A^{-1}$.

3. If A and B are invertible matrices of the same size, then AB is invertible and $(AB)^{-1} = B^{-1}A^{-1}$.

   Proof: $AB.(AB)^{-1} = \mathbb{I}_n$

   $AB.B^{-1}A^{-1} = A(BB^{-1})A^{-1} = A\mathbb{I}_n A^{-1} = AA^{-1} = \mathbb{I}_n$

   By uniqueness of the inverse:

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

   $A^n \;\; n \in \mathbb{Z}_+$

   $A^n \;\; n \in \mathbb{Z}_- \quad A^n = (A^{-1})^{-n}, -n \in \mathbb{Z}_+$

4. If A is an invertible matrix, then A^T is invertible and $(A^T)^{-1} = (A^{-1})^T$.

5. If A is invertible matrix then A^n is invertible for all integers $n \geq 0$ and $(A^n)^{-1} = (A^{-1})^n$.

Elementary Matrices

Definition An elementary matrix is a matrix which can be obtained by performing one elementary row operation on an identity matrix.

Example: We describe EROs corresponding to the following elementary matrices:

$$\begin{bmatrix}1 & 0 & 0\\0 & 1 & 0\\0 & 0 & 4\end{bmatrix} = E_1 \begin{bmatrix}1 & 0 & 0\\0 & 1 & 0\\0 & 0 & 1\end{bmatrix} R_3 \to 4R_3$$

$$\begin{bmatrix}0 & 1 & 0\\1 & 0 & 0\\0 & 0 & 1\end{bmatrix} = E_2 \begin{bmatrix}1 & 0 & 0\\0 & 1 & 0\\0 & 0 & 1\end{bmatrix} R_1 \leftrightarrow R_2$$

$$\begin{bmatrix}1 & 0 & 3\\0 & 1 & 0\\0 & 0 & 1\end{bmatrix} = E_3 \begin{bmatrix}1 & 0 & 0\\0 & 1 & 0\\0 & 0 & 1\end{bmatrix} R_1 \to R_1 + 3R_2$$

Performing EROs is entirely equivalent to left multiplication by elementrary matrices.

Theorem 3.10 Let E be the elementary matrix obtained by performing an ERO on $\mathbb{I}_n$. If the same ERO is performed on an $n \times r$ matrix A, then the result is the matrix EA.

$$\begin{bmatrix}0 & 1 & 0\\1 & 0 & 0\\0 & 0 & 1\end{bmatrix}\begin{bmatrix}a_{11} & a_{12} & a_{13}\\a_{21} & a_{22} & a_{23}\\a_{31} & a_{32} & a_{33}\end{bmatrix} = \begin{bmatrix}a_{21} & a_{22} & a_{23}\\a_{11} & a_{12} & a_{13}\\a_{31} & a_{32} & a_{33}\end{bmatrix} \text{ which is } \begin{bmatrix}a_{11} & a_{12} & a_{13}\\a_{21} & a_{22} & a_{23}\\a_{31} & a_{32} & a_{33}\end{bmatrix} R_1 \leftrightarrow R_2$$

Theorem 3.11 Each elementary matrix is invertible, and its inverse is an elementary matrix of the same type.

Example Construct the elementary matrices for the reverse ERO corresponding to the following elementary matrices

$$\begin{array}{cc} R_3 \to 4R_3 & R_3 \to \frac{1}{4}R_3 \\ E_1 = \begin{bmatrix}1 & 0 & 0\\0 & 1 & 0\\0 & 0 & 4\end{bmatrix} & E_1^{-1} = \begin{bmatrix}1 & 0 & 0\\0 & 1 & 0\\0 & 0 & \frac{1}{4}\end{bmatrix} \\ R_1 \to R_2 & R_2 \to R_1 \\ E_2 = \begin{bmatrix}0 & 1 & 0\\1 & 0 & 0\\0 & 0 & 1\end{bmatrix} & E_2^{-1} = \begin{bmatrix}0 & 1 & 0\\1 & 0 & 0\\0 & 0 & 1\end{bmatrix} \\ R_1 \to R_1 + 3R_3 & R_1 \to R_1 - 3R_3 \\ E_3 = \begin{bmatrix}1 & 0 & 3\\0 & 1 & 0\\0 & 0 & 1\end{bmatrix} & E_3^{-1} = \begin{bmatrix}1 & 0 & -3\\0 & 1 & 0\\0 & 0 & 1\end{bmatrix} \end{array}$$

Invertible Matrices

Theorem 3.12 The Fundamental theorem of invertible matrices

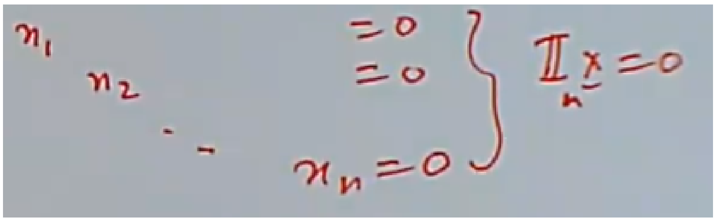
Let A be an $n \times n$ matrix. The following statements are equivalent:

1. A is invertible

2. $A\mathbf{x} = \mathbf{b}$ has a unique solution for every $\mathbf{b}$ in $\mathbb{R}^n$
3. $A\mathbf{x} = \mathbf{0}$ has only the trivial solution
4. The reduced echelon form of A is $\mathbb{I}_n$
5. A is a product of elementary matrices

$Proof$:

- 1 $\Rightarrow$ 2   Theorem 3.7
- 2 $\Rightarrow$ 3 $\mathbf{x} = 0$ is a solution. By 2 it is unique.
- 3 $\Rightarrow$ 4   Write $A = E_1 \dots E_p R$ where R is the reduced echelon form of A. If $\mathbf{x} = 0$ is the only solution



  So $\mathbb{R} = \mathbb{I}_n$

- 4 $\Rightarrow$ 5 $A = E_1 E_2 \dots E_p \mathbb{I}_n = E_1 E_2 \dots E_p$
- 5 $\Rightarrow$ 1   Note that $A' = E_p^{-1} \dots E_1^{-1}$. Then $AA' = A'A = \mathbb{I}_n$. So $A' = A^{-1}$.

Example Express $A = \begin{bmatrix} 1 & 2 \\ 3 & 7 \end{bmatrix}$ as a product of elementary matrices.

$$\begin{bmatrix} 1 & 2 \\ 3 & 7 \end{bmatrix} \xrightarrow{R_2 \to R_3 - 3R_1} \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} \xrightarrow{R_1 \to R_1 - 2R_2} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ -3 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 7 \end{bmatrix}$$
$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} 1 & 2 \\ 3 & 7 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$$
$$\Rightarrow \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$$

Theorem 3.13 A one-sided inverse is a two-sided inverse Let A be a square matrix. If B is a square matrix such that either $AB = \mathbb{I}_n$ or $BA = \mathbb{I}_n$, then A is invertible and $A^{-1} = B$.

$Proof$: Suppose $AB = \mathbb{I}_n$ $(*)$

Consider $B\mathbf{x} = \mathbf{0} \overset{(*)}{\Longrightarrow} AB\mathbf{x} = \mathbf{0} = \mathbf{x} = \mathbf{0}$.

$\mathbf{x} = \mathbf{0}$ is the unique solution to $B\mathbf{x} = \mathbf{0}$.

By Theorem 3.12, B is invertible. ($B^{-1} = BB^{-1} = \mathbb{I}_n$)

$(*) \quad AB = \mathbb{I}_n$

$ABB^{-1} = B^{-1} \Rightarrow A = B^{-1}$.

Theorem 3.14 Let A be a square matrix. If a sequence of elementary row operations reduces A to $\mathbb{I}$, then the same sequence reduces $\mathbb{I}$ to A^{-1}.

$Proof$:

$$\begin{gathered} E_1 \ldots E_p A = \mathbb{I}_n \\ (\text{Thm } 3.12 \ \Rightarrow A \text{ is invertible}) \\ A = E_p^{-1} E_{p-1}^{-1} \ldots E_1^{-1} \mathbb{I}_n \\ A^{-1} = E_1 E_2 \ldots E_p \mathbb{I}_n \end{gathered}$$

Gauss-Jordan method for computing the inverse

A square matrix is invertible if and only if there are EROs transforming A to $\mathbb{I}$. In this case, these EROs compute A^{-1} as

$$[A|\mathbb{I}] \xrightarrow{EROs} [\mathbb{I}|A^{-1}].$$

Example Find the inverses of

$$B = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 3 \end{bmatrix}, C = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$$

if they exist.

$$(B)$$

$$\begin{bmatrix} 1 & 1 & 1|1 & 0 & 0 \\ 0 & 2 & 1|0 & 1 & 0 \\ 0 & 0 & 3|0 & 0 & 1 \end{bmatrix} \xrightarrow[R_3 \to \frac{1}{3}R_3]{R_2 \to \frac{1}{2}R_2} \begin{bmatrix} 1 & 1 & 1|1 & 0 & 0 \\ 0 & 1 & \frac{1}{2}|0 & \frac{1}{2} & 0 \\ 0 & 0 & 1|0 & 0 & \frac{1}{3} \end{bmatrix}$$
$$\xrightarrow{R_1 \to R_1 - R_2} \begin{bmatrix} 1 & 0 & \frac{1}{2}|1 & -\frac{1}{2} & 0 \\ 0 & 1 & \frac{1}{2}|0 & \frac{1}{2} & 0 \\ 0 & 0 & 1|0 & 0 & \frac{1}{3} \end{bmatrix}$$
$$\xrightarrow[R_2 \to R_2 - \frac{1}{2}R_3]{R_1 \to R_1 - \frac{1}{2}R_3} \begin{bmatrix} 1 & 0 & 0|1 & -\frac{1}{2} & -\frac{1}{6} \\ 0 & 1 & 0|0 & \frac{1}{2} & -\frac{1}{6} \\ 0 & 0 & 1|0 & 0 & \frac{1}{3} \end{bmatrix}$$

$$(C)$$

$$\left[\begin{array}{ccc|ccc}1&2&3&1&0&0\\4&5&6&0&1&0\\7&8&9&0&0&1\end{array}\right]\xrightarrow[R_3\to R_3-5R_1]{R_2\to R_2-4R_1}\left[\begin{array}{ccc|ccc}1&2&3&1&0&0\\0&-3&-6&-4&1&0\\0&-3&-6&-5&0&1\end{array}\right]$$

$$\xrightarrow{R_3\to R_3-R_2}\left[\begin{array}{ccc|ccc}1&2&3&1&0&0\\0&-3&-6&-4&1&0\\0&0&0&1&-1&1\end{array}\right]$$

$$\Rightarrow C \text{ is not reducible to } \mathbb{I}_3$$

$$\Rightarrow C \text{ is not an inverse \& Gauss-Jordan method fails.}$$

Subspaces

The idea is that we note that a plane through the origin in $\mathbb{R}^3$ (for example x, y-plane), looks like a copy of $\mathbb{R}^2$ in its own right. We want to generalise this.

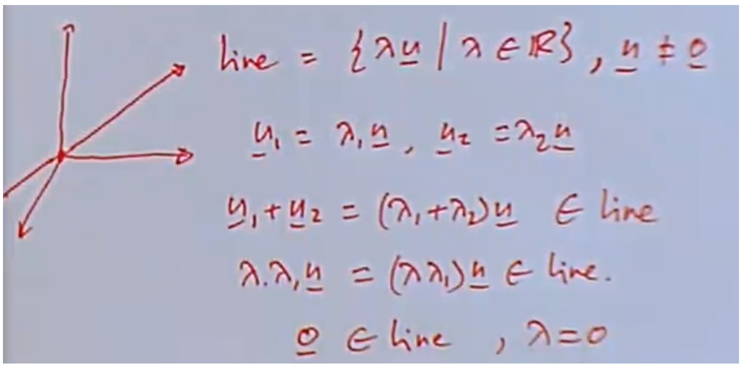
Definition A subspace of $\mathbb{R}^n$ is a collection S of vectors in $\mathbb{R}^n$ such that:

1. The zero vector $\mathbf{0}$ is in S.
2. If $\mathbf{u}$ and $\mathbf{v}$ are in S then $\mathbf{u} + \mathbf{v}$ is in S.
3. If $\mathbf{u}$ is in S and $c \in \mathbb{R}$ is a scalar, then $c\mathbf{u}$ is in S.

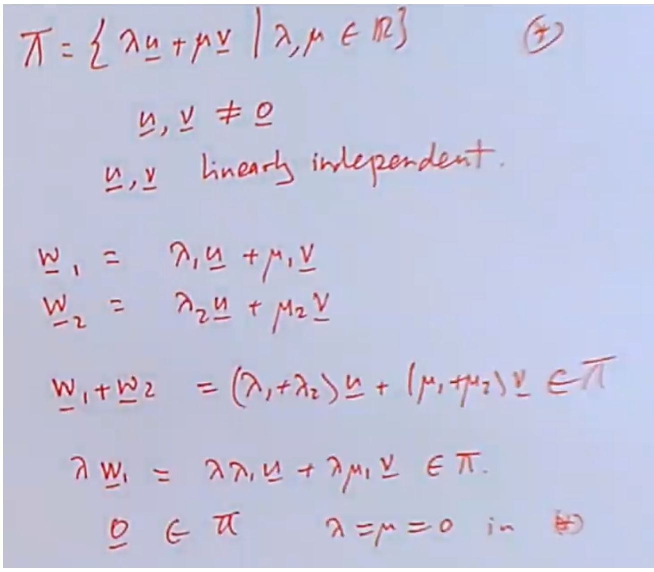
It is possible to combine properties (2) and (3) into the single condition that 'S is closed under taking linear combinations', i.e. if $\mathbf{u}_1, \ldots, \mathbf{u}_n \in S$ and $c_1, \ldots, c_n \in \mathbb{R}$, then $c_1\mathbf{u}_1 + \cdots + c_n\mathbf{u}_n \in S$.

Key examples include:

- $\mathbb{R}^n$ is a subspace of itself; $\{\mathbf{0}\}$ is a subspace of $\mathbb{R}^n$.
- A line through the origin in $\mathbb{R}^3$ is a subspace of $\mathbb{R}^3$.



- A plane through the origin in $\mathbb{R}^3$ is a subspace of $\mathbb{R}^3$.



Example Is the set

$$S = \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} \in \mathbb{R}^3 | x = 1 + 2y, y = 4z \right\}$$

a subspace of $\mathbb{R}^3$?

Does $\mathbf{0}$ belongs to it?

Is so then $x = y = z = 0$ should solve. But $0 = 1 + 2 \cdot 0 \text{ which is false}, 0 = 0 \text{ which is true}$

S is not a vector subspace.

Example Is the set

$$S = \left\{ \begin{bmatrix} x \\ x^2 \end{bmatrix} : x \in \mathbb{R} \right\}$$

a subspace of $\mathbb{R}^2$?

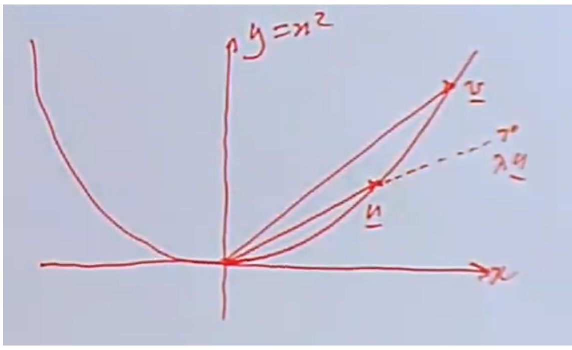
$\mathbf{0} \in S$ ie $n = 0$

$$\lambda \begin{bmatrix} x \\ x^2 \end{bmatrix} = \begin{bmatrix} \lambda x \\ \lambda x^2 \end{bmatrix} ? = \begin{bmatrix} x' \\ x'^2 \end{bmatrix} ?$$

Choose $x = 1$ $\begin{bmatrix} \lambda \\ \lambda \end{bmatrix}$

Choose $x = 2$ $\begin{bmatrix} 2 \\ 2 \end{bmatrix}$ but $2 \neq 2^2$

So S is not a vector subspace because $\mathbf{u} \in S \not\Rightarrow \lambda \mathbf{u} \in S, \lambda \in \mathbb{R}$



Graphical representation. $\lambda \mathbf{u}$ is not in the vector space.

Theorem 3.19 A span is a subspace Let $\mathbf{v}_1, \ldots, \mathbf{v}_k$ be vectors in $\mathbb{R}^n$. The $\text{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ is a subspace of $\mathbb{R}^n$.

$$S = \text{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$$
$$\mathbf{u}, \mathbf{w} \in S$$
$$\mathbf{u} = \sum_{i=1}^{k} \lambda_i \mathbf{v}_i$$
$$\mathbf{w} = \sum_{i=1}^{k} \mu_i \mathbf{v}_i$$
$$\circ\ \mathbf{u} + \mathbf{w} = \lambda_1 \mathbf{v}_1 + \cdots + \lambda_k \mathbf{v}_k + \mu_1 \mathbf{v}_1 + \cdots + \mu_k \mathbf{v}_k$$
$$= (\lambda_1 + \mu_1)\mathbf{v}_1 + \cdots + (\lambda_k + \mu_k)\mathbf{v}_k \in S$$
$$\circ\ \lambda \mathbf{u} = \lambda(\lambda_1 \mathbf{u}_1 + \cdots + \lambda_k \mathbf{u}_k) =$$
$$= (\lambda\lambda_1)\mathbf{u}_1 + \cdots + (\lambda\lambda_k)\mathbf{u}_k \in S$$
$$\circ\ \mathbf{0} = \mathbf{0} \cdot \mathbf{u}_1 + \cdots + \mathbf{0} \cdot \mathbf{u}_k \in S$$

S is a vector subspace of $\mathbb{R}^n$.

Subspaces associated with Matrices

Definition Let A be an $m \times n$ matrix.

1. The row space of A, written $row(A)$ is the subspace of $\mathbb{R}^n$ spanned by the rows of A.
2. The column space of A, written $col(A)$ is the subspace of $\mathbb{R}^n$ spanned by the columns of A.

Example Construct the row and column spaces of the matrix

$$A = \begin{bmatrix} 4 & 7 & 2 \\ 0 & 3 & -1 \end{bmatrix}.$$

$$row(A)\{\lambda[4, 7, 2] + \mu[0, 3, -1] | \lambda, \mu \in \mathbb{R}\} = \{[4\lambda, 7\lambda + 3\mu, 2\lambda - \mu] | \lambda, \mu \in \mathbb{R}\}$$

$$col(A)\left\{\lambda \begin{bmatrix} 4 \\ 0 \end{bmatrix} + \mu \begin{bmatrix} 7 \\ 3 \end{bmatrix} + \nu \begin{bmatrix} 2 \\ -1 \end{bmatrix} | \lambda, \mu, \nu \in \mathbb{R}\right\} = \left\{\begin{bmatrix} 4\lambda + 7\mu + 2\nu \\ 3\mu - \nu \end{bmatrix} | \lambda, \mu, \nu \in \mathbb{R}\right\}$$

Could write

$$\begin{bmatrix} 7 \\ 3 \end{bmatrix} = \frac{13}{4} \begin{bmatrix} 4 \\ 0 \end{bmatrix} - 3 \begin{bmatrix} 2 \\ -1 \end{bmatrix}$$

$$col(A) = \left\{ \lambda \begin{bmatrix} 4 \\ 0 \end{bmatrix} + \nu \begin{bmatrix} 2 \\ -1 \end{bmatrix} + \mu \left(\frac{13}{4} \begin{bmatrix} 4 \\ 0 \end{bmatrix} - 3 \begin{bmatrix} 2 \\ -1 \end{bmatrix} \right) | \lambda, \mu, \nu \in \mathbb{R} \right\}$$

Note that $row(A) = col(A^T)$ and $col(A) = row(A^T)$

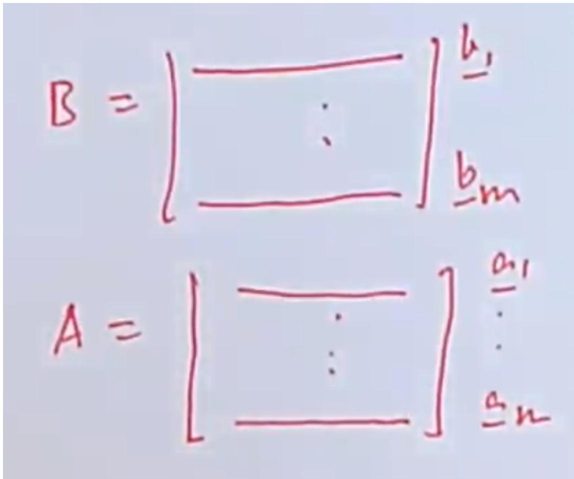
$$A^T = \begin{bmatrix} 4 & 0 \\ 7 & 3 \\ 2 & -1 \end{bmatrix}$$

Theorem 3.20 Let B be any matrix which is row equivalent to A. Then $row(B) = row(A)$.

$Proof : B \sim A$

A & B related by EROs.

$row(B) = \lambda_1 \mathbf{b}_1 + \cdots + \lambda_m \mathbf{b}_m$



Scalar multiplication of a row

$\mathbf{b}_i \to \mathbf{a}_i = \lambda \mathbf{b}_i$

$$\lambda_1 \mathbf{b}_1 + \cdots + \lambda_m \mathbf{b}_m \to \lambda_1 \mathbf{b}_1 + \cdots + \lambda \lambda_i \mathbf{b}_i + \cdots + \lambda_m \mathbf{b}_m \in row(B)$$

Interchange of rows $i \leftrightarrow j$

$$\lambda_1 \mathbf{b}_1 + \cdots + \lambda_i \mathbf{b}_i + \cdots + \lambda_j \mathbf{b}_j + \cdots + \lambda_m \mathbf{b}_m \to$$
$$\to \lambda_1 \mathbf{b}_1 + \cdots + \lambda_i \mathbf{b}_j + \cdots + \lambda_j \mathbf{b}_i + \cdots + \lambda_m \mathbf{b}_m \in row(B)$$

Addition of rows $\mathbf{b}_i \to \mathbf{b}_i + \mathbf{b}_j$

$$\lambda_1\mathbf{b}_1 + \cdots + \lambda_m\mathbf{b}_m \to \lambda_1\mathbf{b}_1 + \cdots + \lambda_i(\mathbf{b}_i + \mathbf{b}_j) + \cdots + \lambda_m\mathbf{b}_m =$$
$$= \lambda_1\mathbf{b}_1 + \cdots + \lambda_i\mathbf{b}_i + \cdots + (\lambda_i + \lambda_j)\mathbf{b}_j + \lambda_m\mathbf{b}_m$$

$$B \to B' \to \cdots \to A$$
$$row(B') \subseteq row(B)\ldots$$
$$row(A) \subseteq row(B)$$

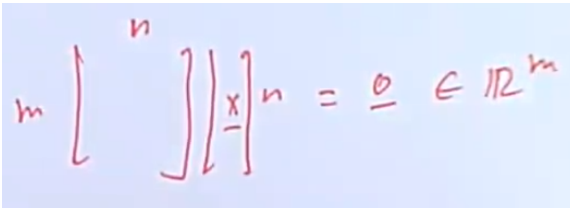
We reverse the argument $A \to B$. $row(B) \subseteq row(A)$, so row space of A and B are equal.

$$row(A) = row(B)$$

Theorem 3.21 Let A be an $m \times n$ matrix. Let N be the set of solutions to the homogeneous linear sustem $A\mathbf{x} = \mathbf{0}$. Then N is a subspace of $\mathbb{R}^n$.

$Proof$:

$$N = \{\mathbf{x} \in \mathbb{R}^n | A\mathbf{x} = \mathbf{0}\}$$



1. $\mathbf{0} \in N$ (N is not empty)
   $A\mathbf{0} = \mathbf{0}$, so $\mathbf{0} \in N$
2. $\mathbf{x}, \mathbf{x'} \in N \Rightarrow \mathbf{x} + \mathbf{x'} \in N$
   $A(\mathbf{x} + \mathbf{x'}) = A\mathbf{x} + A\mathbf{x'} = \mathbf{0} + \mathbf{0} = \mathbf{0}$. So $\mathbf{x}, \mathbf{x'} \in N$
3. $\lambda\mathbf{x} \in N,\ \lambda \in \mathbb{R}$
   $A(\lambda\mathbf{x}) = \lambda A\mathbf{x} = \lambda\mathbf{0} = \mathbf{0}$. So $\lambda\mathbf{x} \in N$.

N is a subspace of $\mathbb{R}^n$.

Example Let $A = \begin{bmatrix} 3 & -1 & 2 \end{bmatrix}$. Contruct the set of solutions to $A\mathbf{x} = \mathbf{0}$ and show that it is a subspace of $\mathbb{R}^3$.

$$\begin{bmatrix} 3 & -1 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = [0]$$
$$3x - y + 2z = 0 \leftarrow \text{This is an equation for a plane.}$$
$$N : y = t,\ z = s,\ x = \frac{1}{3}(t - 2s),\ t, s \in \mathbb{R}$$

- $t = s = 0 \Rightarrow x = y = z = 0$

  so $\mathbf{0} \in N$

- $t \to \lambda t,\ s \to \lambda s$

  so $y \to \lambda y,\ z \to \lambda z$

  What happens to x? $x \to \frac{1}{3}(\lambda t - 2\lambda s) = \lambda \frac{1}{3}(t + 2s) = \lambda x$

  So N is closed under scalar multiplication.

- $\mathbf{x} \in N \quad s, t$
  $\mathbf{x'} \in N \quad s', t'$
  $y = t,\ z = s,\ x = \frac{1}{3}(t - 2s)$
  $y' = t',\ z' = s',\ x' = \frac{1}{3}(t' - 2s')$

$$\left[\tfrac{1}{3}(t + t' - 2(s + s')) \quad t + t' \quad s + s'\right] = \left[\tfrac{1}{3}(t'' - 2s'') \quad t'' \quad s''\right] \in N.$$
$$t'',\ s'' \in \mathbb{R}$$

Definition Let A be an $m \times n$ matrix. The *null space* of A, written $\text{null}(A)$, is the subspace of solutions to the homogeneouus equation $A\mathbf{x} = \mathbf{0}$. Thus

$$\text{null}(A) = \{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x} = \mathbf{0}\}.$$

Theorem 3.22 Let A be an $m \times n$ real matrix. Then for any system $A\mathbf{x} = \mathbf{b}$ of linear equations exactly one of the following is true:

1. There is no solution;
2. There is a unique solution;
3. There are infinitely many solutions.

$Proof$: Assume neither (1) nor (2) holds. We show (3) must then hold.

There must exist at least two distinct solutions: $\mathbf{x}$, $\mathbf{x'}\ \ \mathbf{x} \neq \mathbf{x'}$

Then $A\mathbf{x} = \mathbf{b}$

$\quad A\mathbf{x'} = \mathbf{b}$

$$A(\mathbf{x} - \mathbf{x'}) = \mathbf{b} - \mathbf{b} = \mathbf{0}$$
$$\mathbf{x} - \mathbf{x'} \in \text{null}(A)\ \&\ \mathbf{x} - \mathbf{x'} \neq \mathbf{0}$$

Call $\mathbf{x} - \mathbf{x'} = \mathbf{y}$.

Then $\mathbf{y} \in \text{null}(A)$ & $t\mathbf{y} \in \text{null}(A)\ \forall t \in \mathbb{R}$.

Consider $\{\mathbf{z} = \mathbf{x} + t\mathbf{y} | t\}$ an infinite set of vectors.

Then $A\mathbf{z} = A\mathbf{x} + At\mathbf{y} = A\mathbf{x} + tA\mathbf{y} = \mathbf{b} + \mathbf{0}$

So $A\mathbf{x} = \mathbf{b}$ has infinite family of solutions.

Basis, dimension and rank

Basis

Definition A basis for a subspace S of $\mathbb{R}^n$ is a set of vectors in S that

- a) span S, and
- b) is linearly independent.

For example, the satandard unit vectors $\mathbf{e}_1, \ldots, \mathbf{e}_n$ are a basis for $\mathbb{R}^n$. This basis is called the standard basis.

A non-standard basis for $\mathbb{R}^2$ is

$$\left\{ \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}.$$

Example: Find a basis for the row space of

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 5 & 7 & 9 \end{bmatrix}$$

Row echelon for of A is

$$A \sim \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix} = R$$

By theorem 3.20 $\operatorname{row}(A) = \operatorname{Row}(R)$

$\operatorname{row}(R)$ is spanned by the non-zero rows of R

The stair case structure of R tells is the non-zero rows are linearly independent.

So a basis for $\operatorname{row}(A)$ is: $\{[1, 1, 1], [0, 1, 2]\}$

Example: Find a basis for $\operatorname{Span}(\mathbf{u}, \mathbf{v}, \mathbf{w})$, where $\mathbf{u} = \begin{bmatrix} 1 & 0 & 3 \end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} 4 & 2 & 1 \end{bmatrix}$, $\mathbf{w} = \begin{bmatrix} 5 & 1 & 2 \end{bmatrix}$.

Let $B = \begin{bmatrix} 1 & 0 & 3 \\ 4 & 2 & 1 \\ 5 & 1 & 2 \end{bmatrix}$, where the rows of B are $\mathbf{u}, \mathbf{v}, \mathbf{w}$. $\text{Span}(\mathbf{u}, \mathbf{v}, \mathbf{w}) = \text{Row}(B)$

$$B \sim \underset{\text{row echelon form}}{\begin{bmatrix} 1 & 0 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}}$$

So basis for $\text{Span}(\mathbf{u}, \mathbf{v}, \mathbf{w})$ is $\{\begin{bmatrix} 1 & 0 & 3 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 2 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}\}$.

Example: Find a basis for the $\text{col}(A)$, where A is given in the first example.

$$\text{col}(A) = \text{row}(A^T)$$

Consider

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 5 & 7 & 9 \end{bmatrix} \Rightarrow A^T = \begin{bmatrix} 1 & 4 & 5 \\ 2 & 5 & 7 \\ 3 & 6 & 9 \end{bmatrix} \sim \begin{bmatrix} 1 & 4 & 5 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix} \text{ in row echelon form}$$

The non-zero rows of the row-echelon form of A^T, are linearly independent, $\text{span row}(A^T)$.

So the basis for $\text{col}(A)$ is: $\left\{ \begin{bmatrix} 1 \\ 4 \\ 5 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \right\}$.

Example: Find a basis for $\text{null}(A)$, where A is given in the previous example.

Recall $\text{null}(A) = \{\mathbf{x} \in \mathbb{R}^3 | A\mathbf{x} = \mathbf{0}\}$

$$[A|\mathbf{0}] = A = \begin{bmatrix} 1 & 2 & 3 & | & 0 \\ 4 & 5 & 6 & | & 0 \\ 5 & 7 & 9 & | & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & -1 & | & 0 \\ 0 & 1 & 2 & | & 0 \\ 0 & 0 & 0 & | & 0 \end{bmatrix} \text{ in row echelon form}$$

Write $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ then the general solution of $A\mathbf{x} = \mathbf{0}$ is:

$$\begin{array}{cccccc} x_1 & & -x_3 & = & 0 & \text{i.e.} \quad x_1 = x_3 \\ & x_2 & +2x_3 & = & 0 & \text{i.e.} \quad x_2 = -2x_3 \end{array}$$

for all $x_3 \in \mathbb{R}$.

So $\text{null}(A) = \left\{ x_3 \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} : x_3 \in \mathbb{R} \right\} = \text{span} \left\{ \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} \right\}$.

Since a single non-zero vector is linearly independent, A basis for $\text{null}(A)$ is: $\left\{ \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} \right\}$.

Theorem 3.23 The basis theorem Let S be a subspace of $\mathbb{R}^n$. Then any two bases for S have the same number of vectors.

$Proof$: We aim to prove this theorem by contradiction.

Let B be a basis for S with r vectors $\mathbf{v}_1, \ldots, \mathbf{v}_r \in S$ and let C be a basis of S with s vectors $\mathbf{w}_1, \ldots, \mathbf{w}_s \in S$.

Assume $r < s$.

Consider $c_1 \mathbf{w}_1 + \cdots + c_s \mathbf{w}_s = \mathbf{0} \qquad (*)$

for $c_!, \ldots, c_s \in \mathbb{R}$.

Since B is a basis B spans S.

So for each $i = 1, \ldots, s$ there exists scalars $a_{ij} \in \mathbb{R}$ such that

$$\mathbf{w}_i = a_{i1}\mathbf{v}_1 + \cdots + a_{ir}\mathbf{v}_r$$

Substituting into $(*)$ gives

$$\begin{aligned} c_1(a_{11}\mathbf{v}_1 + \cdots + a_{1r}\mathbf{v}_r) + c_2(a_{21}\mathbf{v}_1 + \cdots + a_{2r}\mathbf{v}_r) \\ + \cdots + c_s(a_{s1}\mathbf{v}_1 + \cdots + a_{sr}\mathbf{v}_r) = \mathbf{0} \\ \iff (c_1 a_{11} + c_2 a_{21} + \cdots + c_s a_{s1})\mathbf{v}_1 \\ + \cdots + (c_1 a_{1r} + c_2 a_{2r} + \cdots + c_s a_{sr})\mathbf{v}_r = \mathbf{0} \end{aligned}$$

Since B is linearly independent then $(*)$ holds if and only if

$$\begin{aligned} c_1 a_{11} + c_2 a_{21} + \cdots + c_s a_{s1} = 0 \\ c_1 a_{12} + c_2 a_{22} + \cdots + c_s a_{s2} = 0 \\ \vdots \\ c_1 a_{1r} + c_2 a_{2r} + \cdots + c_s a_{sr} = 0 \end{aligned}$$

We have r homogenneous equations in s variables, we have $r < s$, so we have fewer equations than variables.

Therefore $(*)$ has infinitely many solutions, and in particulara non-trivial solution, which means that ws have to be linearly dependent, so C is linearly dependent, which contradicts the statement.

Therefore $r \geq s$. If $r > s$ intechange the rows of B and C in the argument which gives a contradiction.

So $r = s$. $\square$

Dimension

If S is a nontrivial subspace of $\mathbb{R}^n$, then the number of vectors in a basis for S is called the *dimension* of S, denoted $\dim(S)$. If $S - \{\mathbf{0}\}$ then we define $\dim(S) - 0$.

$\mathbb{R}^n$ forms a key example. Since the standard basis $\mathbf{e}_1, \ldots, \mathbf{e}_n$ for $\mathbb{R}^n$ has n-elements, the dimension of $\mathbb{R}^n$ is n.

Example: $S = \{\mathbf{0}\}$. dimension $\dim(S) = 0$.

Example: Find the dimension of a plane through the origin in $\mathbb{R}^3$.

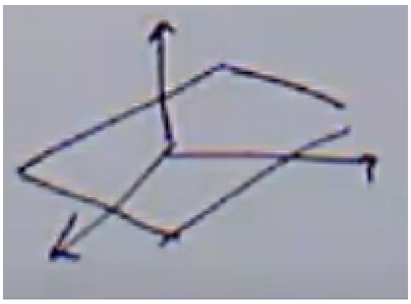
Let a plane in $\mathbb{R}^3$ be given by

$$\{\lambda\mathbf{u} + \mu\mathbf{v} : \lambda, \mu \in \mathbb{R}\}$$

such that $\mathbf{u}, \mathbf{v} \in \mathbb{R}^3$, $\mathbf{u}, \mathbf{v} \neq \mathbf{0}$ and $\mathbf{u}, \mathbf{v}$ are linearly independent.

Basis for the palne is $\{\mathbf{u}, \mathbf{v}\}$.

Dimension of a plane through the origin is 2.



This corresponds to our intuition of what dimension should be.

Example: For the matrix

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 5 & 7 & 9 \end{bmatrix}$$

find the dimension of $\mathrm{row}(A)$ and $\mathrm{col}(A)$.

From previous examples, we know that basis for $\mathrm{row}(A)$ is: $\{[1,1,1],[0,1,2]\}$ and basis for $\mathrm{col}(A)$ is: $\left\{ \begin{bmatrix} 1 \\ 4 \\ 5 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \right\}$.

So $\dim(\mathrm{row}(A)) = 2$ and $\dim(\mathrm{col}(A)) = 2$ because both of these sets containt 2 vectors.

Theorem 3.24 The row and column spaces of a matrix A have the same dimension.

Proof Let R be the reduced row echelon form of A. By Theorem 3.20, $\mathrm{row}(A) = \mathrm{row}(R)$, so

$$\begin{aligned} \dim(\mathrm{row}(A)) &= \dim(\mathrm{row}(R)) \\ &= \text{number of nonzero rows of } R \\ &= \text{number of leading 1s of } R \end{aligned}$$

Let this number be called r.

Now $\mathrm{col}(A) \neq \mathrm{col}(R)$, but the columns of A and R have the same dependence relationships. Therefore, $\dim(\mathrm{col}(A)) = \dim(\mathrm{col}(R))$. Since there are r leading 1s, R has r columns that are standard unit vectors, $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_r$. (These will be vectors in $\mathbb{R}^m$ if A and R are $m \times n$ matrices.) These r vectors are linearly independent, and the remaining columns of R are linear combinations of them. Thus, $\dim(\mathrm{col}(R)) = r$. It follows that $\dim(\mathrm{row}(A)) = r = \dim(\mathrm{col}(A))$, as we wished to prove.

This theorem was referred to during a lecture, but was not covered in one.

Definition The *rank* of a matrix A is the dimension of its row and column spaces and is denoted $\mathrm{rank}(A)$.

Theorem 3.25 For any matrix A, $\mathrm{rank}(A) = \mathrm{rank}(A^T)$.

$Proof$: This follows form the previous theorem, as the row space of A^T is the same as the column space of A.

$$\begin{aligned} \mathrm{rank}(A) = \dim(\mathrm{row}(A)) &= \dim(\mathrm{col}(A)) \text{ by Theorem 3.24} \\ &= \dim(\mathrm{row}(A^T)) = \mathrm{rank}(A^T). \ \square \end{aligned}$$

Definition The *nullity* of a matrix A is the dimension of its null space $\mathrm{null}(A)$, and is denoted $\mathrm{nullity}(A)$.

The rank and nullity are related by the following result.

Theorem 3.26 The rank theorem If A is an $m \times n$ matrix then

$$\operatorname{rank}(A) + \operatorname{nullity}(A) = n,$$

where n is the number of columns of A.

$Proof$: Let R be the reduced row echelon form of A. Suppose $\operatorname{rank}(A) = r$. Then the $\operatorname{rank}(A) = \operatorname{rank}(R)$ following from Theorem 3.20.

R has r leading 1's, then there r leading variables and $n - r$ free variables in the solution of $A\mathbf{x} = \mathbf{0}$.

So $\operatorname{nullity}(A) = n - r$.

Then $\operatorname{rank}(A) + \operatorname{nullity}(A) = r + n - r = n. \square$

Example: Let C be a 3×4 matrix. If $\operatorname{rank}(C) = 3$, what is the dimension of $\operatorname{null}(C)$?

C has 4 columns.

By the rank theorem

$$\operatorname{rank}(C) + \operatorname{nullity}(C) = 4$$

Since $\operatorname{rank}(C) = 3$ we have

$$3 + \operatorname{nullity}(C) = 4$$

so $\operatorname{nullity}(C) = 1 = \dim(\operatorname{null}(C))$.

Example: Find the nullity of the following matrix

$$A = \begin{bmatrix} 1 & 2 & 0 & 1 & 8 \\ 56 & 2 & 101 & -8 & -38 \end{bmatrix}.$$

We note the rows of A are linearly independent. So a basis for the $\operatorname{row}(A)$ is the two rows of A. So $\operatorname{rank}(A) = 2$.

By the rank theorem:

$$2 + \operatorname{nullity}(A) = 5.$$

So $\operatorname{nullity}(A) = 3$.

Coordinates

We now work with an *ordered basis* B

$$B : \mathbf{v}_1 \dots, \mathbf{v}_k,$$

i.e. the order matters, B is not just a set of vectors.

Theorem 3.29 Let S be a subspace of $\mathbb{R}^n$ and let $B : \mathbf{v}_1 \dots, \mathbf{v}_k$ be an ordered basis for S. Then for every $\mathbf{v} \in S$, there is exactly one way to write $\mathbf{v}$ as an ordered linear combination of the basis vector in B.

Ex.:

$\mathbb{R}^2 : \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \end{bmatrix} \right\} = \{\mathbf{v}_1, \mathbf{v}_2\}$

$\mathbf{v} = \begin{bmatrix} 0 \\ -1 \end{bmatrix} = 1 \times \mathbf{v}_1 - 1 \times \mathbf{v}_2$ and this combination is unique

$Proof:$

Since B is a basis for S, B spans S. So for $\mathbf{v} \in S$ there exists $c_1, \dots, c_k \in \mathbb{R}$ such that:

$$\mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k$$

So there is at least one way to write $\mathbf{v}$ as an ordered linear combination of B.

Let's assume $\mathbf{v} = d_1\mathbf{v}_1 + d_2\mathbf{v}_2 + \cdots + d_k\mathbf{v}_k$ for $d_1, \dots, d_k \in \mathbb{R}$ and also $\mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k$ for $c_1, \dots, c_k \in \mathbb{R}$.

So $\mathbf{0} = \mathbf{v} - \mathbf{v} = (d_1\mathbf{v}_1 + d_2\mathbf{v}_2 + \cdots + d_k\mathbf{v}_k) - (c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k)$
$= (d_1 - c_1)\mathbf{v}_1 + (d_2 - c_2)\mathbf{v}_2 + \cdots + (d_k - c_k)\mathbf{v}_k$

Since B is linearly independent then it follows that

$$d_1 - c_1 = 0,\ d_2 - c_2 = 0,\ \dots,\ d_k - c_k = 0$$
$$\Rightarrow d_1 = c_1,\ d_2 = c_2,\ \dots,\ d_k = c_k$$

So there is exactly one way to write $\mathbf{v}$ as an ordered linear combination of B. $\square$

Definition Let S be a subspace of $\mathbb{R}^n$ and let $B : \mathbf{v}_1, \dots, \mathbf{v}_k$ be an ordered basis for S. Let $\mathbf{v} \in S$ and write $\mathbf{v} = c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k$. Then $c_1, c_2, \dots, c_k$ are called the *coordinates of* $\mathbf{v}$ *with respect to* B and the column vector

$$[\mathbf{v}]_B = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_k \end{bmatrix}$$

is called the *coordinate vector of* $\mathbf{v}$ *with respect to* B.

Example: Let $\varepsilon : \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ be the standard basis for $\mathbb{R}^3$.

Find the coordinate of $\mathbf{v} = \begin{bmatrix} 5 \\ 2 \\ 4 \end{bmatrix}$ with respect to ε.

$$\mathbf{v} = \mathbf{5e}_1 + \mathbf{2e}_2 + \mathbf{4e}_3$$
$$[\mathbf{v}]_\varepsilon = \begin{bmatrix} 5 \\ 2 \\ 4 \end{bmatrix}$$

In general, if $\mathbf{v} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$ then $[\mathbf{v}]_\varepsilon = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$. That is, the coordinate vector of any $\mathbf{v} \in \mathbb{R}^3$ with respect to the standard basis is just $\mathbf{v}$ itself (as a column vector).

Example:

Let P be the plane through the roigin in $\mathbb{R}^3$ spanned by $\mathbf{u} = \begin{bmatrix} 1 \\ 3 \\ 0 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}$. Then $B : \mathbf{u}, \mathbf{v}$ is an ordered basis for P. Construct the coordinate vector of $\mathbf{w} = \begin{bmatrix} 1 \\ 5 \\ 4 \end{bmatrix}$ with respect to B.

$\mathbf{w} = c_1\mathbf{u} + c_2\mathbf{v}$, where $c_1, c_2 \in \mathbb{R}$

We want to find c_1, c_2

By inspection: $\mathbf{w} = \mathbf{1u} + \mathbf{2v}$

So $[\mathbf{w}]_B = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$

Plane is two dimensional, so both $\mathbf{u}, \mathbf{v}$ belong to 2 dimensional surface

That's why our coordinate vector has just two entries

Linear Transformations

Change of basis

Definition Let $\mathcal{B} : \mathbf{u}_1, \ldots, \mathbf{u}_n$ and $\mathcal{C} : \mathbf{v}_1, \ldots, \mathbf{v}_n$ be ordered bases for a $\mathbb{R}^n$. The $n \times n$ matrix whose columns are the coordinate vectors

$$[\mathbf{u}_1]\mathcal{C}, \ldots, [\mathbf{u}_n]\mathcal{C}$$

of the vectors in $\mathcal{B}$ with respect to the basis $\mathcal{C}$ is denoted $P_{\mathcal{C}\leftarrow\mathcal{B}}$ and called the change of basis matrix from $\mathcal{B}$ to $\mathcal{C}$.

Example: For $\mathbb{R}^2$, let $\mathcal{B} = \{\mathbf{u}_1, \mathbf{u}_2\}$, $\mathcal{C} = \{\mathbf{v}_1, \mathbf{v}_2\}$, where

$$\mathbf{u}_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

What is the change of basis matrix $P_{\mathcal{C}\leftarrow\mathcal{B}}$?

$$\begin{aligned} \mathbf{u}_1 &= \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \mathbf{v}_1 + \mathbf{v}_2 \\ \mathbf{u}_2 &= \begin{bmatrix} 1 \\ -1 \end{bmatrix} = 2\begin{bmatrix} 1 \\ 0 \end{bmatrix} - \begin{bmatrix} 1 \\ 1 \end{bmatrix} = 2\mathbf{v}_1 - \mathbf{v}_2 \\ P_{\mathcal{C}\leftarrow\mathcal{B}} &= \begin{bmatrix} 1 & 2 \\ 1 & -1 \end{bmatrix} \end{aligned}$$

Theorem 6.12 Let $\mathcal{B} : \mathbf{u}_1, \ldots, \mathbf{u}_n$ and $\mathcal{C} : \mathbf{v}_1, \ldots, \mathbf{v}_n$ be ordered bases for $\mathbb{R}^n$ and let $P_{\mathcal{C}\leftarrow\mathcal{B}}$ be the change of basis matrix from $\mathcal{B}$ to $\mathcal{C}$. Then

1. For all $\mathbf{x} \in \mathcal{C}$, $P_{\mathcal{C}\leftarrow\mathcal{B}}[\mathbf{x}]_{\mathcal{B}} = [\mathbf{x}]_{\mathcal{C}}$;
2. $P_{\mathcal{C}\leftarrow\mathcal{B}}$ is the unique $n \times n$ matrix P such that $P[\mathbf{x}]_{\mathcal{B}} = [\mathbf{x}]_{\mathcal{C}}$;
3. $P_{\mathcal{C}\leftarrow\mathcal{B}}$ is invertible and $P^{-1}_{\mathcal{C}\leftarrow\mathcal{B}} = P_{\mathcal{B}\leftarrow\mathcal{C}}$

Note that the statement in part 1) of the theorem mirrors the pattern of indices in matrix multiplication. In fact the definition of the change of basis matrix has been cunningly arranged to make this happen. Suppose we have a general vector $x \in \mathbb{R}^n$. We may write it as a column vector of coordinates with respect to each of the bases $\mathcal{B}$ and $\mathcal{C}$.

$$\sum_{i=1}^{n} [\mathbf{x}]_{\mathcal{B},i} \mathbf{u}_i = \sum_{j=1}^{n} [\mathbf{x}]_{\mathcal{C},j} \mathbf{v}_j .$$

From the definition of the change of basis matrix:

$$\mathbf{u}_i = \sum_{j=1}^{n} (P_{\mathcal{C}\leftarrow\mathcal{B}})_{ji} \mathbf{v}_j .$$

(Note that i labels the columns.)

Putting these formulae together yields,

$$\sum_{i=1}^{n} [\mathbf{x}]_{\mathcal{B},i} \sum_{j=1}^{n} (P_{\mathcal{C}\leftarrow\mathcal{B}})_{ji} \mathbf{v}_j = \sum_{j=1}^{n} [\mathbf{x}]_{\mathcal{C},j} \mathbf{v}_j ,$$

and since the $\mathbf{v}_j$ are a basis,

$$\sum_{i=1}^{n} [\mathbf{x}]_{\mathcal{B},i} \left(P_{\mathcal{C}\leftarrow\mathcal{B}}\right)_{ji} = [\mathbf{x}]_{\mathcal{C},j}.$$

Looking at where the index i sits on the LHS means that as a product of matrices we have:

$$P_{\mathcal{C}\leftarrow\mathcal{B}}[\mathbf{x}]_{\mathcal{B}} = [\mathbf{x}]_{\mathcal{C}}$$

the vectors being column vectors.

Example: For $\mathcal{B}$ and $\mathcal{C}$ as in the previous example:

1. Find $[\mathbf{x}]_{\mathcal{C}}$, where $[\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix} 3 \\ 5 \end{bmatrix}$.
2. Find the change of basis matrix $P_{\mathcal{B}\leftarrow\mathcal{C}}$

1.

$$\begin{gathered} P_{\mathcal{C}\leftarrow\mathcal{B}}[\mathbf{x}]_{\mathcal{B}} = [\mathbf{x}]_{\mathcal{C}} \\ \begin{bmatrix} 1 & 2 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 3 \\ 5 \end{bmatrix} = \begin{bmatrix} 13 \\ -2 \end{bmatrix} \end{gathered}$$

2.

$$P_{\mathcal{B}\leftarrow\mathcal{C}} = \begin{bmatrix} 1 & 2 \\ 1 & -1 \end{bmatrix}^{-1}$$
$$\Rightarrow \det = -1 - 2 = -3$$
$$\Rightarrow \frac{1}{-3}\begin{bmatrix} -1 & -2 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 1/3 & 2/3 \\ 1/3 & -1/3 \end{bmatrix}$$

Example: Let $V = P_2(\mathbb{R})$ and let $B : 1, x, x^2$ be the standard basis for V. Let $\mathcal{C}$ be the basis $2 - x, 2 + x^2, 2x + x^2$. Find the $P_{\mathcal{C}\leftarrow\mathcal{B}}$ and $P_{\mathcal{B}\leftarrow\mathcal{C}}$. Hence find the coordinate vector of $p(x) = 3 + 4x + 7x^2$ with respect to $\mathcal{C}$.

$P_2(\mathbb{R})$ is the set of all polynomials of degree at most 2

$P_2(\mathbb{R}) = \{a_0 + a_1x + a_2x^2 | a_0, a_1, a_2 \in \mathbb{R}\}$

$$[a_0, a_1, a_2] \longleftrightarrow a_0 + a_1x + a_2x^2$$
$$[b_0, b_1, b_2] \longleftrightarrow b_0 + b_1x + b_2x^2$$

We can represent the row vectors as column vectors and identify $P_2(\mathbb{R})$ with $\mathbb{R}^3$

$$\begin{array}{cccc} & \mathbf{u}_1 & \mathbf{u}_2 & \mathbf{u}_3 \\ B & \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} & \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} & \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \end{array}$$

Where $\mathbf{u}_1$ corresponds to x^0, $\mathbf{u}_2$ corresponds to x^1 and $\mathbf{u}_3$ corresponds to x^2

$$\begin{array}{cccc} & \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 \\ C & \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix} & \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix} & \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix} \\ \\ & 2 - x & 2 + x^2 & 2x + x^2 \end{array}$$

In order to find $P_{C\leftarrow B}$, we would have to find a solution to

$$\mathbf{u}_1 = x\mathbf{v}_1 + y\mathbf{v}_2 + z\mathbf{v}_3$$
$$\vdots$$

this would be quite complicated.

We can focus on $P_{B\leftarrow C}$

$$\begin{aligned}\mathbf{v}_1 &= 2\mathbf{u}_1 - \mathbf{u}_2 \\ \mathbf{v}_2 &= 2\mathbf{u}_1 + \mathbf{u}_3 \\ \mathbf{v}_2 &= 2\mathbf{u}_2 + \mathbf{u}_3\end{aligned}$$
$$\Rightarrow P_{B\leftarrow C} = \begin{bmatrix} 2 & 2 & 0 \\ -1 & 0 & 2 \\ 0 & 1 & 1 \end{bmatrix} \Rightarrow P_{C\leftarrow B} = P_{B\leftarrow C}^{-1} = \begin{bmatrix} 1 & 1 & -2 \\ -1/2 & -1 & 2 \\ 1/2 & 1 & -1 \end{bmatrix}$$

Now let's find the coordinate vector of $p(x) = 3 + 4x + 7x^2$ with respect to $\mathcal{C}$
.

$$[\mathbf{p}]_B = \begin{bmatrix} 3 \\ 4 \\ 7 \end{bmatrix}_B$$
$$[\mathbf{p}]_C = P_{C\leftarrow B}\begin{bmatrix} 3 \\ 4 \\ 7 \end{bmatrix} = \begin{bmatrix} 1 & 1 & -2 \\ -1/2 & -1 & 2 \\ 1/2 & 1 & -1 \end{bmatrix}\begin{bmatrix} 3 \\ 4 \\ 7 \end{bmatrix} = \begin{bmatrix} -7 \\ 23/2 \\ -3/2 \end{bmatrix}$$

Theorem 6.13 Gauss Jordan method for computing a change of basis matrix
Let $B : \mathbf{u}_1, \ldots, \mathbf{u}_n$ and $C : \mathbf{v}_1, \ldots, \mathbf{v}_n$ be ordered bases for a vector space V. For any basis ε for V (such as the standard basis if $V = \mathbb{R}^n$), let B be the matrix with columns $[\mathbf{u}_1]_\varepsilon, \ldots, [\mathbf{u}_n]_\varepsilon$ and C the matrix with columns $[\mathbf{v}_1]_\varepsilon, \ldots, [\mathbf{v}_n]_\varepsilon$. Then applying row reduction to the $n \times 2n$ augmented matrix produces

$$[C|B] \to [\mathbb{I}_n|P_{C\leftarrow B}]$$

$$\begin{gathered} C : \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\} \\ B : \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\} \\ \varepsilon : \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\} \end{gathered}$$
$$\begin{bmatrix} \vdots & \vdots & \vdots & | & \vdots & \vdots & \vdots \\ \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 & | & \mathbf{u}_1 & \mathbf{u}_2 & \mathbf{u}_3 \\ \vdots & \vdots & \vdots & | & \vdots & \vdots & \vdots \end{bmatrix}$$
$$\sim \begin{bmatrix} P_{B\leftarrow C} & | & \mathbb{I}_3 \end{bmatrix}$$
$$\sim \begin{bmatrix} \mathbb{I}_3 & | & P_{C\leftarrow B} \end{bmatrix}$$

Example: Return to the previous example, and use the theorem to compute $P_{C\leftarrow B}$.

$$\left[\begin{array}{ccc|ccc} 2 & 2 & 0 & 1 & 0 & 0 \\ -1 & 0 & 2 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \end{array}\right] \xrightarrow[(R_1\to -R_1)]{R_1\leftrightarrow R_2} \left[\begin{array}{ccc|ccc} 1 & 0 & -2 & 0 & -1 & 0 \\ 2 & 2 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \end{array}\right]$$
$$\xrightarrow{R_2\to R_2-2R_1} \left[\begin{array}{ccc|ccc} 1 & 0 & -2 & 0 & -1 & 0 \\ 0 & 2 & 4 & 1 & 2 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \end{array}\right]$$
$$\xrightarrow{R_2\leftrightarrow R_3} \left[\begin{array}{ccc|ccc} 1 & 0 & -2 & 0 & -1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 2 & 4 & 1 & 2 & 0 \end{array}\right]$$
$$\xrightarrow{R_3\to R_3-2R_2} \left[\begin{array}{ccc|ccc} 1 & 0 & -2 & 0 & -1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 2 & 1 & 2 & -2 \end{array}\right]$$
$$\xrightarrow{R_3\to \frac{1}{2}R_3} \left[\begin{array}{ccc|ccc} 1 & 0 & -2 & 0 & -1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1/2 & 1 & -1 \end{array}\right]$$
$$\xrightarrow[R_2\to R_2-R_3]{R_1\to R_1+2R_3} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1/2 & 1 & -2 \\ 0 & 1 & 0 & -1/2 & -1 & 2 \\ 0 & 0 & 1 & 1/2 & 1 & -1 \end{array}\right]$$

Linear maps

Definition A linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$ is a mapping $T : \mathbb{R}^n \to \mathbb{R}^m$ such that for all $\mathbf{v}_1$, $\mathbf{v}_2 \in \mathbb{R}^n$ and scalars $c \in \mathbb{R}$,

1. $T(\mathbf{v}_1 + \mathbf{v}_2) = T(\mathbf{v}_1) + T(\mathbf{v}_2)$;
2. $T(c\mathbf{v}_1) = cT(\mathbf{v}_1)$.

Example: Consider the transformation $T : \mathbb{R}^2 \to \mathbb{R}^3$ given by

$$T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ 2x-y \\ 3x+4y \end{bmatrix}.$$

We show this to be a linear transformation.

$$T\left(\begin{bmatrix}x\\y\end{bmatrix}+\begin{bmatrix}x'\\y'\end{bmatrix}\right)=T\left(\begin{bmatrix}x+x'\\y+y'\end{bmatrix}\right)=\begin{bmatrix}x+x'\\2(x+x')-(y+y')\\3(x+x')+4(y+y')\end{bmatrix}$$
$$=\begin{bmatrix}x\\2x-y\\3x+4y\end{bmatrix}+\begin{bmatrix}x'\\2x'-y'\\3x'+4y'\end{bmatrix}=T\left(\begin{bmatrix}x\\y\end{bmatrix}\right)+T\left(\begin{bmatrix}x'\\y'\end{bmatrix}\right)$$

$$T\left(c\begin{bmatrix}x\\y\end{bmatrix}\right)=T\left(\begin{bmatrix}cx\\cy\end{bmatrix}\right)=\begin{bmatrix}cx\\c2x-cy\\c3x+c4y\end{bmatrix}=c\begin{bmatrix}x\\2x-y\\3x+4y\end{bmatrix}=cT\left(\begin{bmatrix}x\\y\end{bmatrix}\right)$$

Example: Every matrix transformation is a linear transfromation, i.e. if A is an $m\times n$ matrix, then $T_A:\mathbb{R}^n\to\mathbb{R}^m$ given by $T_A(\mathbf{x})=A\mathbf{x}$ is linear.

$$A(\mathbf{x}+\mathbf{x}')=A\mathbf{x}+A\mathbf{x}'$$
$$A(c\mathbf{x})=cA\mathbf{x}$$

Example: The set of polynomials of degree less than or equal to n in x with real coefficients is denoted $P_n(\mathbb{R})$. We can identify this set with $\mathbb{R}^{n+1}$ by the correspondence,

$$a_0+a_1x+a_2x^2+\cdots+a_nx^n\leftrightarrow\begin{bmatrix}a_0\\a_1\\a_2\\\vdots\\a_n\end{bmatrix}$$

and this makes $P_n(\mathbb{R})$ into an example of what we later think of as a vector space in a more abstract sense. Clearly the usual operation of addition and scalar multiplication on $P_n(\mathbb{R})$ correspond precisely to those on $\mathbb{R}^{n+1}$.

Now define the differentiation operator for, say, $n=3$, by $D:P_3(\mathbb{R})\to P_2(\mathbb{R})$ where $D(p(x))=p'(x)$. There is a corresponding map $\tilde{D}:\mathbb{R}^4\to\mathbb{R}^3$.

We write down $\tilde{D}$ as a map and show that it is linear.

$$D : P_3(\mathbb{R}) \to P_2(\mathbb{R})$$
$$p(n) \to p'(n)$$
$$a_0 + a_1 n + a_2 n^2 + a_3 n^3 \to a_1 + 2a_2 x + 3a_3 x^2$$
$$\tilde{D} : \mathbb{R}^4 \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{bmatrix} \to \begin{bmatrix} a_1 \\ 2a_2 \\ 3a_3 \end{bmatrix} \mathbb{R}^3$$
$$\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} a_1 \\ 2a_2 \\ 3a_3 \end{bmatrix}$$
$$\tilde{D} \text{ is a } 3 \times 4 \text{ matrix} : \mathbb{R}^4 \to \mathbb{R}^3 \text{and hence is linear}$$

We could have taken a different approach thinking about D as a linear map on two polynomials:

$$p(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3$$
$$q(x) = b_0 + b_1 x + b_2 x^2 + b_3 x^3$$
$$(p+q)(x) = p(x) + q(x)$$
$$= a_0 + b_0 + (a_1 + b_1)x + (a_2 + b_2)x^2 + (a_3 + b_3)x^3$$
$$(p+q)'(x) = a_1 + b_1 + 2(a_2 + b_2)x + 3(a_3 + b_3)x^2 = p'(x) + q'(x)$$
$$D(p+q) = Dp + Dq$$
$$D(cp) = cD(p)$$

Example: Thinking of $M_{2,2}(\mathbb{R})$ as a copy of $\mathbb{R}^4$, is $\det : M_{2,2}(\mathbb{R}) \to \mathbb{R}$ linear?

$$M_{2,2} \quad \begin{bmatrix} a & b \\ c & d \end{bmatrix} \leftrightarrow \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}$$

$$\tilde{\det} : \mathbb{R}^4 \to \mathbb{R}$$
$$\mathbf{m} \to ad - bc$$

Is $\tilde{\det}$ linear?

$$\begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} \to ad - bc$$

But

$$\lambda \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = \begin{bmatrix} \lambda a \\ \lambda b \\ \lambda c \\ \lambda d \end{bmatrix} \to \lambda a \lambda d - \lambda b \lambda c = \lambda^2 ad - \lambda^2 bc = \lambda^2 (ad - bc) \neq \lambda(ad - bc)$$

Therefore $\tilde{\det}$ is not a linear map.

$(*)$ **Example:** Note that $\mathbb{R} = \mathbb{R}^1$. Consider $T : \mathbb{R} \to \mathbb{R}$ given by $T(x) = 2x + 1$. Is T linear?

$$\begin{gathered} T : \mathbb{R}' \to \mathbb{R}' \\ x \to 2x + 1 \\ \lambda x \to 2\lambda x + 1 \neq \lambda(2x + 1) \end{gathered}$$

T is not linear.

Example: The zero map $T_0 : \mathbb{R}^n \to \mathbb{R}^m$ given by $T_0(\mathbf{v}) = \mathbf{0}_{\mathbb{R}^m}$ for all $\mathbf{v} \in V$ is linear.

Example: The identity transoformation $\mathbb{I}_n : \mathbb{R}^n \to \mathbb{R}^n$ given by $\mathbb{I}_n(\mathbf{v}) = \mathbf{v}$ for all $\mathbf{v} \in V$ is linear.

Example: It can be useful to think of changing the basis as a linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^n$: namely the transformation corresponding to $P_{C \leftarrow B}$ which takes $[\mathbf{x}]_B$ to $[\mathbf{x}]_C$.

Theorem 6.14 Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. Then

1. $T(\mathbf{0}_{\mathbb{R}^n}) = \mathbf{0}_{\mathbb{R}^m}$;
2. $T(-\mathbf{v}) = -T(\mathbf{v})$ for $\mathbf{v} \in \mathbb{R}^n$;
3. $T(\mathbf{u} - \mathbf{v}) = T(\mathbf{u}) - T(\mathbf{v})$ for $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$.

   1.) $T(\mathbf{0}) = T(0 \cdot \mathbf{n}) = 0 \cdot T(\mathbf{n}) = \mathbf{0}$

   2.) $T(-\mathbf{v}) = T(-1 \cdot \mathbf{v}) = -T(\mathbf{v})$

   3.) $T(\mathbf{u} - \mathbf{v}) = T(\mathbf{u}) + T(-\mathbf{v}) = T(\mathbf{u}) - T(\mathbf{v})$

Example: Revisit example $(*)$ above, using part 1. of Theorem 6.14.

$$T : \mathbb{R}' \to \mathbb{R}'$$
$$x \to 2x + 1$$

If T were linear then $T(0) = 0$, but $T(0) = 1 \neq 0$. Hence, T not linear.

Theorem 6.15 Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation and let B : $\mathbf{v}_1, \ldots, \mathbf{v}_n$ br a basis for $\mathbb{R}^n$. Then T is completely determined by its effect on B. More precisely, if $\mathbf{v} \in \mathbb{R}^n$ has

$$\mathbf{v} = c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n$$

for scalars $c_1, \ldots, c_n$, then

$$T(\mathbf{v}) = c_1T(\mathbf{v}_1) + \cdots + c_nT(\mathbf{v}_n).$$

$Proof$: This follows from the fact that a linear map must preserve linear combinations.

$$\begin{aligned} T(c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n) &= T(c_1\mathbf{v}_1) + T(c_2\mathbf{v}_2) + \cdots + T(c_n\mathbf{v}_n) \\ &= c_1T(\mathbf{v}_1) + \cdots + c_nT(\mathbf{v}_n) \end{aligned}$$

Note that there is no assumption that the set $\{T(\mathbf{v}_1), \ldots T(\mathbf{v}_n)\}$ be a basis of $\mathbb{R}^m$.

Example: Theorem 6.15 can be used to define linear maps by specifying their behaviour on basis elements. For instance, we can let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be the (necessarily unique) linear map with

$$T\left(\begin{bmatrix}1\\0\end{bmatrix}\right) = \begin{bmatrix}3\\0\end{bmatrix} \text{ and } T\left(\begin{bmatrix}0\\1\end{bmatrix}\right) = \begin{bmatrix}4\\2\end{bmatrix}.$$

$$T\begin{bmatrix}x\\y\end{bmatrix} = T\left(x\begin{bmatrix}1\\0\end{bmatrix} + y\begin{bmatrix}0\\1\end{bmatrix}\right) = x\begin{bmatrix}3\\1\end{bmatrix} + y\begin{bmatrix}4\\2\end{bmatrix} = \begin{bmatrix}3x+4y\\x+2y\end{bmatrix} = \begin{bmatrix}3 & 4\\1 & 2\end{bmatrix}\begin{bmatrix}x\\y\end{bmatrix}$$

$\begin{bmatrix}3 & 4\\1 & 2\end{bmatrix}$ is a matrix associated with the linear map T.

Definition Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. Let $B : \mathbf{v}_1, \ldots, \mathbf{v}_n$ be a basis for $\mathbb{R}^n$ and $C : \mathbf{u}_1, \ldots, \mathbf{u}_m$ be a basis for $\mathbb{R}^m$. Then the matrix of T with respect to these bases is defined to be

$$[T] = (T_{ij}) \text{ where } T(\mathbf{v}_i) = \sum_{j=1}^{m} T_{ji}\mathbf{u}_j.$$

Note carefully the order of the indices in the right hand expression.

That this definition is correct follows from the following calculation

$$\begin{aligned} T\left(\sum_{i=1}^{n} x_i \boldsymbol{v}_i\right) &= \sum_{i=1}^{n} x_i T(\boldsymbol{v}_i) \\ &= \sum_{i=1}^{n} x_i \sum_{j=1}^{m} T_{ji}(\boldsymbol{u}_j) \\ &= \sum_{j=1}^{m} \left(\sum_{i=1}^{n} T_{ji} x_i\right) \boldsymbol{u}_j \end{aligned}$$

so that the action of T on the coordinates of $\mathbf{x} \in \mathbb{R}^n$ with respect to B and $T(\mathbf{x}) \in \mathbb{R}^m$ with respect to C is

$$[T] : \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \mapsto \begin{bmatrix} T_{11} & T_{12} & \dots & T_{1n} \\ T_{21} & T_{22} & \dots & T_{2n} \\ \vdots & & & \\ T_{m1} & T_{m2} & \dots & T_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

Composition of linear transformations

Definition Let $T : \mathbb{R}^n \to \mathbb{R}^m$ and $S : \mathbb{R}^m \to R^p$ be transformations. The composition $S \circ T$ is the mapping $S \circ T : \mathbb{R}^n \to \mathbb{R}^p$ defined by

$$(S \circ T)(\mathbf{u}) = S(T(\mathbf{u})) \in \mathbb{R}^p$$

for $\mathbf{u} \in \mathbb{R}^n$.

Theorem 6.16 If $Y : \mathbb{R}^n \to \mathbb{R}^m$ and $S : \mathbb{R}^m \to \mathbb{R}^p$ are linear transformations, then $S \circ T : \mathbb{R}^n \to \mathbb{R}^p$ is also linear.

$$
\begin{aligned}
(S \circ T)(\mathbf{u}+\mathbf{v}) &= S(T(\mathbf{u}+\mathbf{v})) \\
&= S(T(\mathbf{u})+T(\mathbf{v})) \\
&= S \circ T(\mathbf{u}) + S \circ T(\mathbf{v})
\end{aligned}
$$

$$
\begin{aligned}
(S \circ T)(\lambda \mathbf{u}) &= S(T(\lambda \mathbf{u})) \\
&= S(\lambda T(\mathbf{u})) \\
&= \lambda S(T(\mathbf{u})) \\
&= \lambda (S \circ T)(\mathbf{u})
\end{aligned}
$$

Inverses of linear transformations

We now turn to invertibility of linear transformations from $\mathbb{R}^n$ to itself.

Given a vector space V, let $I_n : \mathbb{R}^n \to \mathbb{R}^n$ be the identity linear transformation given by $I_n(\mathbf{v}) = \mathbf{v}$ for all $\mathbf{v} \in \mathbb{R}^n$. This is called the identity map on $\mathbb{R}^n$.

Definition A linear transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ is invertible if there exists a linear transformation $T' : \mathbb{R}^n \to \mathbb{R}^n$ with $T' \circ T = I_n$ and $T \circ T' = I_n$. Such a map T' is called an inverse of T.

Theorem 6.17 If $T : \mathbb{R}^n \to \mathbb{R}^n$ is an invertible linear transformation then its inverse is unique.

$Proof$: Suppose T' and T'' are both inverses to T:

$T \circ T' = T' \circ T = I_n$ and $T \circ T'' = T'' \circ T = I_n$, take

$$T'' = T' \circ T \circ T'' = T'$$

therefore the inverse is unique.

Given an invertible map $T : V \to W$ we write T^{-1} for the inverse of T. Then T^{-1} is also invertible and $(T^{-1})^{-1} = T$.

Theorem 3.33 Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be an invertible linear transformation. Then the matrices od T and T^{-1} with respect to any basis of $\mathbb{R}^n$ are also in inverse:

$$[T^{-1}] = [T]^{-1}.$$

Determinants and eigenvalues

Determinants

Recall that if

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

then

$$\det(A) = |A| = a_{11}a_{22} - a_{12}a_{21}.$$

We now extend this definition to determinants of $n \times n$ matrices.

$n = 1$: if $A = (a_{11})$ define $\det(A) = a_{11}$.

$n = 3$: if

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

then we define

$\det(A) = |A| = a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}.$

We write A_{ij} for a submatrix of A obtained by deleting row i and column j. Then

$$\det(A) = a_{11}\det(A_{11}) - a_{12}\det(A_{12}) + a_{13}\det(A_{13}) = \sum_{j=1}^{3}(-1)^{j+1}a_{1j}\det(A_{1j})$$

Example: Find $\det(A)$ where

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 1 & 0 \\ 7 & 3 & 9 \end{bmatrix}.$$

$$\begin{aligned} A = \begin{vmatrix} 1 & 2 & 3 \\ 4 & 1 & 0 \\ 7 & 3 & 9 \end{vmatrix} &= 1\begin{vmatrix} 1 & 0 \\ 3 & 9 \end{vmatrix} - 2\begin{vmatrix} 4 & 0 \\ 7 & 9 \end{vmatrix} + 3\begin{vmatrix} 4 & 1 \\ 7 & 3 \end{vmatrix} \\ &= 1(9-0) - 2(36-0) + 3(12-7) \\ &= 9 - 72 + 15 = -48 \end{aligned}$$

Definition Let $A = (a_{ij})$ be an $n \times n$ matrix, $n \geq 2$. Then the *determinant* of A, $\det(A)$, is the scalar

$\det(A) = a_{11}\det(A_{11}) - a_{12}\det(A_{12}) + \cdots + (-1)^{1+n}a_{1n}\det(A_{1n})$
$= \sum_{j=1}^{n}(-1)^{1+j}a_{1j}\det(A_{1j}).$

In this definition we are expanding along the first row.

To compute determinant, we can expand along any row or any column. More precisely:

Theorem 4.1 Let A be an $n \times n$ matrix. Then for any i we can expand along the ith row:

$\det(A) = (-1)^{i+1}a_{i1}\det(A_{i1}) + (-1)^{i+2}a_{i2}\det(A_{i2}) + \cdots + (-1)^{i+n}a_{in}\det(A_{in})$
$= \sum_{j=1}^{n}(-1)^{i+j}a_{ij}\det(A_{ij})$

and for any j we can expand along the jth column:

$\det(A) = (-1)^{1+j}a_{1j}\det(A_{1j}) + (-1)^{2+j}a_{2j}\det(A_{2j}) + \cdots + (-1)^{n+j}a_{nj}\det(A_{nj})$
$= \sum_{i=1}^{n}(-1)^{i+j}a_{ij}\det(A_{ij}).$

Example: Find

$$\det \begin{bmatrix} 1 & 2 & 3 \\ 4 & 1 & 0 \\ 7 & 3 & 9 \end{bmatrix}$$

using column 3.

$\det \begin{bmatrix} 1 & 2 & 3 \\ 4 & 1 & 0 \\ 7 & 3 & 9 \end{bmatrix} = 3 \det \begin{bmatrix} 4 & 1 \\ 7 & 3 \end{bmatrix} - 0 \det \begin{bmatrix} 1 & 2 \\ 7 & 3 \end{bmatrix} + 9 \det \begin{bmatrix} 1 & 2 \\ 4 & 1 \end{bmatrix}$
$= 3(12-7) - 0 + 9(1-8) = 15 - 63 = 48$

Example: Find $\det(B)$, where

$$B = \begin{bmatrix} 5 & 1 & 0 & 1 \\ 7 & 0 & 3 & 1 \\ 2 & 0 & 0 & 0 \\ 1 & 2 & 4 & 9 \end{bmatrix}.$$

We can see that $\text{row } 3$ has many zeros, so let's choose that one.

$\det(B) = 2 \begin{vmatrix} 1 & 0 & 1 \\ 0 & 3 & 1 \\ 2 & 4 & 9 \end{vmatrix} + 0$
$= 2 \left(\begin{vmatrix} 3 & 1 \\ 4 & 9 \end{vmatrix} - 0 \begin{vmatrix} 0 & 2 \\ 1 & 9 \end{vmatrix} + \begin{vmatrix} 0 & 3 \\ 2 & 4 \end{vmatrix} \right) = 2(23 - 0 + (-6)) = 34.$

Example: Compute $\det(C)$, where

$$B = \begin{bmatrix} 5 & 1 & 0 & 1 \\ 0 & 2 & 3 & 1 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 2 \end{bmatrix}.$$

$\det(B) = 5 \begin{vmatrix} 2 & 3 & 1 \\ 0 & 1 & 4 \\ 0 & 0 & 2 \end{vmatrix} + 0$
$= 5 \cdot 2 \begin{vmatrix} 1 & 4 \\ 0 & 2 \end{vmatrix} + 0 = 5 \cdot 2 \cdot 2 = 20.$

Theorem 4.2 The determinant of an upper or lower triangular matrix $A = (a_{ij})$ is the product of the entries $a_{11}, a_{22}, \ldots, a_{nn}$ on its main diagonal.

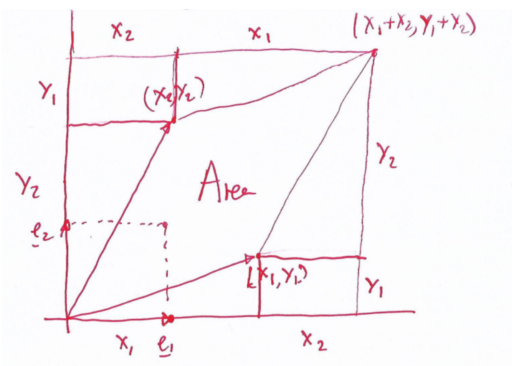
Note that (a_{ij}) is upper triangular matrix when $a_{ij} = 0$ for $i > j$ and lower triangular matrix when $a_{ij} = 0$ for $i < j$.

Example: $\det(\mathbb{I}_n) = 1$. As $\det(\mathbb{I}_n) = 1 \times 1 \times \cdots \times 1 = 1$.

Example: The geometrical interpretantion of the 2×2 determinant $\begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix}$,

follows from consideration of the area of a parallelogram with vertices

$$(0,0),\ (x_1, y_1),\ (x_2, y_2),\ (x_1 + x_2, y_1 + y_2)$$



$$
\begin{aligned}
&Area = (x_1 + x_2)(y_1 + y_2) - 2y_1x_2 - y_2x_2 - y_1x_1 \\
&= \cancel{x_1y_1} + x_2y_1 + x_1y_2 + \cancel{x_2y_2} - 2y_1x_2 - \cancel{y_2x_2} - \cancel{y_1x_1} \\
&= x_1y_2 - x_2y_1 \\
&= \begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix}
\end{aligned}
$$

Properties of determinants

We now want to produce methods for efficiently computing the determinant. We start by relating the determinant to row operations.

Theorem 4.3 Let $A = (a_{ij})$ be a square matrix.

a) If B is obtained from A by swapping any two adjacent rows then

$\det(B) = -\det(A)$.

b) If B is obtained from A by multiplying a row by k, then $\det(B) = k\det(A)$.

c) If B is obtained from A by adding a multiple of one row of A to another row of A, then $\det(B) = \det(A)$.

It can be shown that statement (a) can be extended to apply to interchange of any two rows of a matrix. Hence, if B is obtained from A by swapping any two rows then $\det(B) = -\det(A)$.

The point of the previous theorem is that you can use EROs to compute determinants efficiently. Care is needed to track the effect on the determinant of the ERO.

Example: Find $\det(A)$, where

$$A = \begin{bmatrix} 2 & 5 & 7 \\ 1 & 4 & 5 \\ 3 & 6 & 12 \end{bmatrix}.$$

$$\begin{bmatrix} 2 & 5 & 7 \\ 1 & 4 & 5 \\ 3 & 6 & 12 \end{bmatrix} \xrightarrow{R_1 \leftrightarrow R_2} \begin{bmatrix} 1 & 4 & 5 \\ 2 & 5 & 7 \\ 3 & 6 & 12 \end{bmatrix} \xrightarrow[R_2 \to R_3 - 3R_1]{R_2 \to R_2 - 2R_1} \begin{bmatrix} 1 & 4 & 5 \\ 0 & -3 & -3 \\ 0 & -6 & -3 \end{bmatrix} \xrightarrow[R_2 \to -\frac{1}{3}R_3]{R_2 \to -\frac{1}{3}R_2} \begin{bmatrix} 1 & 4 & 5 \\ 0 & 1 & 1 \\ 0 & 2 & 1 \end{bmatrix} \xrightarrow{R_3 \to R_3 - 2R_2} \begin{bmatrix} 1 & 4 & 5 \\ 0 & 1 & 1 \\ 0 & 0 & -1 \end{bmatrix}$$

The determinant was changing as

$$D \xrightarrow{R_1 \leftrightarrow R_2} -D \xrightarrow[R_2 \to R_3 - 3R_1]{R_2 \to R_2 - 2R_1} -D \xrightarrow[R_2 \to -\frac{1}{3}R_3]{R_2 \to -\frac{1}{3}R_2} -D\left(-\frac{1}{3}\right)\left(-\frac{1}{3}\right) = -\frac{1}{9}D \xrightarrow{R_3 \to R_3 - 2R_2} -\frac{1}{9}D \xrightarrow{R_3 \to -R_3} \frac{1}{9}D$$

The last matrix is upper triangular, so $\frac{1}{9}D = 1 \cdot 1 \cdot 1 = 1$ and hence $D = 9$.

Theorem 4.4 Let E be an $n \times n$ elementary matrix

a) If E results from swapping two rows of $\mathbb{I}_n$ then $\det(E) = -1$.

b) If E results from multiplying one row of $\mathbb{I}_n$ by $k \neq 0$ then $\det(E) = k$.

c) If E results from adding a multiple of one row of $\mathbb{I}_n$ to another row then $\det(E) = 1$.

$Proof$: This follows from Theorem 4.3, since $\det(\mathbb{I}_n) = 1$. $\square$

The next result combines results of Theorem 4.3 and 4.4.

Lemma 4.5 Let B be an $n \times n$ matrix and E be an $n \times n$ elementary matrix. Then

$$\det(EB) = \det(E)\det(B).$$

From this result it follows that if a matrix A is expressed as a product of elementary matrices, $E_1, E_2, \ldots, E_p$ and the reduced riw echelon form R,

$$A = E_1E_2\ldots E_pR$$

then

$\det A = \det E_1 \det E_2 \ldots \det E_p \det R.$

Using the previous lemma (and the fundamental theorem of invertible matrices) we obtain the key result that we can test for invertibility using the determinant.

Theorem 4.6 A square matrix A is invertible if and only if $\det(A) \neq 0$.

$Proof$: $\det R \neq 0 \iff$ a leading entry on every row $\iff R = \mathbb{I}_n$.

$\det A \neq 0 \iff \det R \neq 0 \iff R = \mathbb{I}_n \iff A$ is invertible

The following key properties of determinants can be verified using the tools we have so far.

Theorem 4.7-4.10 Let A and B be $n \times n$ matrices.

1 $\det(kA) = k^n \det(A)$ for all scalars k.

2 $\det(AB) = \det(A)\det(B)$.

3 If A is invertible then

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

4 $\det(A) = \det(A^T)$.

$Proof$: (2) $A = E_1 E_2 \ldots E_p R_A \; B = E'_1 \ldots E'_G R_B$

$$AB = E_1 \ldots E_p R_A E'_1 \ldots E'_G R_B$$
$$\det(AB) = \det(E_1 \ldots E_p R_A \det(E'_1 \ldots E_G R_B)$$

(3) $AA^{-1} = \mathbb{I}_n$

$\det(AA^{-1}) = 1$

$\det(A)\det(A^{-1}) = 1$

$\det(A^{-1}) = \frac{1}{\det(A)}$

(4) row expansion of $A =$ column expansion of A^T

Eigen values and eigenvectors of $n \times n$ matrices

As a matrix is invertible if and only if it has non-zero determinant, we can use this to compute eigenvalues.

The eigenvalues of a square matrix A are the solutions λ of the equation

$$\det(A - \lambda\mathbb{I}) = 0.$$

$$A\mathbf{x} = \lambda\mathbf{x}$$
$$\lambda : \text{eigenvalue}$$
$$\mathbf{x} : \text{eigenvector}, \mathbf{x} \neq \mathbf{0}$$
$$(A - \lambda\mathbb{I}_n)\mathbf{x} = \mathbf{0}$$
$$\text{If an } \mathbf{x} \neq \mathbf{0} \text{ is to exists then } |A - \lambda\mathbb{I}_n| = 0$$
$$\text{i.e. } A - \lambda\mathbb{I}_n \text{ is not invertible.}$$

Definition We call the polynomial $\det(A - \lambda\mathbb{I})$ the *characteristic polynomial* of A.

$$\begin{vmatrix} a_{11}-\lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22}-\lambda & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn}-\lambda \end{vmatrix} \texttt{Characteristic polynomial}$$

If A is an $n \times n$ matrix then its characteristic polynomial has degree n.

- Product of all diagonal elements is always a term which appears inside the determinant.

A degree n polynomial has at most n distinct roots (in $\mathbb{R}$ or $\mathbb{C}$), so an $n \times n$ matrix A has at most n distinct eigenvalues.

Here is a procedure for finding eigenvalues and eigenvectors of an $n \times n$ matrix A:

a) Compute the characteristic polynomial, $\det(A - \lambda\mathbb{I})$.

b) Find the eigenvalues of A by solvingthe characteristic equation $\det(A - \lambda\mathbb{I}) = 0$.

c) For each eigenvalue λ, find the null space of the matrix $A - \lambda\mathbb{I}$. This is the eigenspace E_λ. The λ-eigenvectors of A are the non-zero vectors in E_λ.

E_λ= Null space of $A - \lambda\mathbb{I}_n$

Note: $\mathbf{0} \in E_\lambda$ but $\mathbf{0}$ is not an eigenvector

We might also want to find bases for these eigenspaces.

Example: Find the eigenvalues and a basis for the corresponding eigenspaces of the matrix

$$A = \begin{bmatrix} 1 & 2 & 0 \\ -1 & -1 & 1 \\ 0 & 1 & 1 \end{bmatrix}.$$

$$\begin{gathered} A - \lambda\mathbb{I}_3 = \begin{bmatrix} 1-\lambda & 2 & 0 \\ -1 & -1-\lambda & 1 \\ 0 & 1 & 1-\lambda \end{bmatrix} \\ |A - \lambda\mathbb{I}_3| = (1-\lambda)\begin{vmatrix} -1-\lambda & 1 \\ 1 & 1-\lambda \end{vmatrix} - 2\begin{vmatrix} -1 & 1 \\ 0 & 1-\lambda \end{vmatrix} + 0 \\ = (1-\lambda)(\lambda^2 - 1 - 1) - 2(\lambda - 1) = (1-\lambda)(\lambda^2 - 2 + 2) = \lambda^2(1-\lambda) \end{gathered}$$

Hence eigenvaules are $\lambda = 0$ and $\lambda = 1$.

- $E_0 = \left\{ \mathbf{x} | \begin{bmatrix} 1 & 2 & 0 \\ -1 & -1 & 1 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \right\}$

$$\begin{gathered} x + 2y = 0 \\ -x - y + z = 0 \\ y + z = 0 \Rightarrow y = -z \\ \Rightarrow x = 2z \\ -2z + z + z = 0,\ z \text{ is arbitrary} \end{gathered}$$

$E_0 = \left\{ \begin{bmatrix} 2z \\ -z \\ z \end{bmatrix} | z \in \mathbb{R} \right\}$ make a choice of e-vector: $\begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix} \equiv \mathbf{e}_0$

- $E_1 = \left\{ \mathbf{x} | \begin{bmatrix} 1-1 & 2 & 0 \\ -1 & -1-1 & 1 \\ 0 & 1 & 1-1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \right\}$

  $= \left\{ \mathbf{x} | \begin{bmatrix} 0 & 2 & 0 \\ -1 & -2 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \right\}$

$$2y = 0 \Rightarrow y = 0$$
$$-x - 2y + z = 0 \Rightarrow x = z$$
$$y = 0$$

$$E_1 = \left\{ \begin{bmatrix} z \\ 0 \\ z \end{bmatrix} | z \in \mathbb{R} \right\} \mathbf{e}_1 \equiv \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

Definition The *algebraic multiplicity* of an eigenvalue is its multiplicity as a root of the characteristic equation. The *geometric multiplicity* of an eigenvalue λ is the dimension of its eigenspace E_λ.

The geometric multiplicity will always be less than or equal the algebraic multiplicity and they really can be different.

Consider

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}.$$
$$|A - \lambda \mathbb{I}_2| = \begin{vmatrix} 1-\lambda & 1 \\ 0 & 1-\lambda \end{vmatrix} = (1-\lambda)^2$$

algebraic multiplicity of $\lambda = 1$ is 2

$E_1 = \left\{ \mathbf{x} | \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \right\}$

$y = 0$

$E_1 = \left\{ \begin{bmatrix} x \\ 0 \end{bmatrix} | x \in \mathbb{R} \right\} \mathbf{e}_1 \equiv \begin{bmatrix} 1 \\ 0 \end{bmatrix}$

geometric multiplicity of $\lambda = 1$ is 1.

If we consider

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$
$$|A - \lambda \mathbb{I}_2| = \begin{vmatrix} 1-\lambda & 0 \\ 0 & 1-\lambda \end{vmatrix} = (1-\lambda)^2$$
$$A\mathbf{x} = \mathbf{x} \text{ for all } \mathbf{x}.$$
$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, e_1' = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$
$$\text{are a basis for } E_1.$$

Example What are the algebraic and geometric multiplicities of the eigenvalues of the matrix

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

from the previous example?

$$\lambda = 0 \quad E_0 = \text{span} \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix}$$
$$\lambda = 1 \quad E_1 = \text{span} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$
$$|A - \lambda \mathbb{I}_3| = \lambda^2(1-\lambda)$$

so $\lambda = 0$ has algebraic multiplicity 2 and geometric 1

$\lambda = 1$ has algebraic multiplicity 1 and geometric 1 (as its eigenspace is spanned by a single vector)

Theorem 4.15 The eigenvalues of an upper- or lower-triangular matrix A are the entries on its main diagonal.

$$\begin{vmatrix} a_{11}-\lambda & a_{12} & \dots & a_{1n} \\ 0 & a_{22}-\lambda & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_{nn}-\lambda \end{vmatrix} = (a_{11}-\lambda)(a_{22}-\lambda)\cdots(a_{nn}-\lambda)$$

Example What are the eigenvalues of the matrices

$$A = \begin{bmatrix} 5 & 2 & 4 \\ 0 & 3 & 1 \\ 0 & 0 & -2 \end{bmatrix} \text{ and } B = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 17 \end{bmatrix}?$$

$A : \lambda = 5, 3, -2$ and $B : \lambda = 2, 1, 17$

Theorem 4.16 A square matrix A is invertible if and only if zero is **not** an eigenvalue of A.

Recall that the square matrix A is invertible if and only if the equation $A\mathbf{x} = \mathbf{0}$ has the unique solution $\mathbf{x} = \mathbf{0}$.

0 is not an eigenvalue $\iff$ $A\mathbf{x} = \mathbf{0}$ for no $\mathbf{x} \neq \mathbf{0}$

Theorem 4.20 Let A be an $n \times n$ matrix and let $\lambda_1, \lambda_2, \cdots, \lambda_m$ be **distinct** eigenvalues of A with corresponding eigenvectors $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_m$. Then $\{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_m\}$ is linearly independent.

Proof The proof is indirect. We will assume that $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$ are linearly *dependent* and show that this assumption leads to a contradiction.

If $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$ are linearly dependent, then one of these vectors must be expressible as a linear combination of the previous ones. Let $\mathbf{v}_{k+1}$ be the first of the vectors $\mathbf{v}_i$ that can be so expressed. In other words, $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ are linearly independent, but there are scalars $c_1, c_2, \ldots, c_k$ such that

$$\mathbf{v}_{k+1} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k \qquad (1)$$

Multiplying both sides of Equation (1) by A from the left and using the fact that $A\mathbf{v}_i = \lambda_i\mathbf{v}_i$ for each i, we have

$$\begin{aligned} \lambda_{k+1}\mathbf{v}_{k+1} = A\mathbf{v}_{k+1} &= A(c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k) \\ &= c_1A\mathbf{v}_1 + c_2A\mathbf{v}_2 + \cdots + c_kA\mathbf{v}_k \qquad (2) \\ &= c_1\lambda_1\mathbf{v}_1 + c_2\lambda_2\mathbf{v}_2 + \cdots + c_k\lambda_k\mathbf{v}_k \end{aligned}$$

Now we multiply both sides of Equation (1) by λ_{k+1} to get

$$\lambda_{k+1}\mathbf{v}_{k+1} = c_1\lambda_{k+1}\mathbf{v}_1 + c_2\lambda_{k+1}\mathbf{v}_2 + \cdots + c_k\lambda_{k+1}\mathbf{v}_k \qquad (3)$$

When we subtract Equation (3) from Equation (2), we obtain

$$0 = c_1(\lambda_1 - \lambda_{k+1})\mathbf{v}_1 + c_2(\lambda_2 - \lambda_{k+1})\mathbf{v}_2 + \cdots + c_k(\lambda_k - \lambda_{k+1})\mathbf{v}_k$$

The linear independence of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ implies that

$$c_1(\lambda_1 - \lambda_{k+1}) = c_2(\lambda_2 - \lambda_{k+1}) = \cdots = c_k(\lambda_k - \lambda_{k+1}) = 0$$

Since the eigenvalues λ_i are all distinct, the terms in parentheses $(\lambda_i - \lambda_{k+1})$, $i = 1, \ldots, k$, are all nonzero. Hence, $c_1 = c_2 = \cdots = c_k = 0$. This implies that

$$\mathbf{v}_{k+1} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k = 0\mathbf{v}_1 + 0\mathbf{v}_2 + \cdots + 0\mathbf{v}_k = \mathbf{0}$$

which is impossible, since the eigenvector $\mathbf{v}_{k+1}$ cannot be zero. Thus, we have a contradiction, which means that our assumption that $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$ are linearly dependent is false. It follows that $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$ must be linearly independent.

$Proof$: Let's assume $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_m$ are linearly dependent.

By considering increasing subsets $\{\mathbf{v}_1\}$, $\{\mathbf{v}_1, \mathbf{v}_2\}$, $\{\mathbf{v}_1, \mathbf{v}_2, \cdots\}$, $\ldots$

We find a k s.t.

.$\mathbf{v}_{k+1} \in \text{span}\{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k\}$

$\mathbf{v}_{k+1} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k \quad (*)$

where $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k$ are linearly independent

$A(*): \ \lambda_{k+1}\mathbf{v}_{k+1} = \lambda_1c_1\mathbf{v}_1 + \lambda_2c_2\mathbf{v}_2 + \cdots + \lambda_kc_k\mathbf{v}_k \quad (1)$

But also take $\lambda_{k+1}(*) : \lambda_{k+1}\mathbf{v}_{k+1} = \lambda_{k+1}c_1\mathbf{v}_1 + \lambda_{k+1}c_2\mathbf{v}_2 + \cdots + \lambda_{k+1}c_k\mathbf{v}_k \quad (2)$

Now subtract (1) and (2)

$$\mathbf{0} = (\lambda_1 - \lambda_{k+1})c_1\mathbf{v}_1 + (\lambda_2 - \lambda_{k+1})c_2\mathbf{v}_2 + \cdots + (\lambda_k - \lambda_{k+1})c_k\mathbf{v}_k \quad (2)$$

Since $\lambda_i - \lambda_{k+1} \neq 0 \ \forall i \neq k+1$ and not all c_i are zero.

Therefore $\{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_m\}$ is linearly dependent.

Similarity and Diagonalisation

Similarity

Definition Let A and B be $n \times n$ matrices. We say that A is similar to B if there exists an invertible $n \times n$ matrix P such that $P^{-1}AP = B$.

If A is similar to be we write $A \sim B$.

Note that for P invertible,

$$P^{-1}AP = B \iff A = PBP^{-1} \iff AP = PB.$$

Example Let $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$ and $P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. Use P to show that A is similar to B.

First, check $\det(P)$. $\det(P) = 0 - 1 = -1 \neq 0$, therefore P is invertible.

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

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

So $AP = PB$ and hence $P^{-1}AP = B$ so $A \sim B$. $\square$

Theorem 4.21 Let A, B and C be $n \times n$ matrices. Then

1. $A \sim A$
2. If $A \sim B$, then $B \sim A$
3. If $A \sim B$ and $B \sim C$, then $A \sim C$

$Proof$: (1) Let $P = \mathbb{I}$. Then $P^{-1}AP = \mathbb{I}A\mathbb{I} = A$.

(2) Suppose $A \sim B$. Then there exists an invertible matrix P such that $P^{-1}AP = B$. Let $Q = P^{-1}$, which is invertible. Then $QAQ^{-1} = B$, s o $AQ^{-1} = Q^{-1}B$ then $A = Q^{-1}BQ$. So $B \sim A$.

Theorem 4.22 Let A and B be $n \times n$ matrices with $A \sim B$. The:

1. $\det(A) = \det(B)$
2. A is invertible if and only if B is invertible
3. A and B have the same rank
4. A and B have the same characteristic polynomial
5. A and B have the same eigenvalues

$Proof$:

(1) Let P be an invertible $n \times n$ matrix such that $P^{-1}AP = B$.

$$\begin{aligned}&\det(B) = \det(P^{-1}AP)\\ &= \det(P^{-1})\det(A)\det(P) \quad \text{by Theorem 4.8}\\ &= \tfrac{1}{\det(P)}\det(A)\det(P) \qquad \text{by Theorem 4.9}\\ &= \det(A)\end{aligned}$$

(4) Suppose $\lambda \in \mathbb{R}$, and P is an $n \times n$ invertible matrix such that $P^{-1}AP = B$. Then,

$$\begin{aligned}&P^{-1}(A - \lambda\mathbb{I})P = P^{-1}AP - \lambda P^{-1}P\\ &= B - \lambda\mathbb{I}\end{aligned}$$

So $A - \lambda\mathbb{I} \sim B - \lambda\mathbb{I}$.

Now take $\det(A - \lambda\mathbb{I}) = \det(B - \lambda\mathbb{I})$ by (1).

Example Show that $A = \begin{pmatrix} 2 & 3 \\ 1 & 4 \end{pmatrix}$ and $B = \begin{pmatrix} 2 & 4 \\ 1 & 3 \end{pmatrix}$ are not similar.

$\det(A) = 2 \cdot 4 - 3 \cdot 1 = 5$

$\det(B) = 2 \cdot 3 - 4 \cdot 1 = 2$

so $\det(A) \neq \det(B)$ and A is not similar to B by Theorem 4.22 (1).

Note: Theorem 4.22 is an implication, not an equivalence, so two matrices satisfying the conditions might not be similar.

$\det(A) = \det(B)$ does not gurantee $A \sim B$.

Example

$A = \begin{bmatrix} 1 & 3 \\ 2 & 2 \end{bmatrix} \det(A) = 1 \cdot 2 - 3 \cdot 2 = -4$

$B = \begin{bmatrix} 1 & 1 \\ 3 & -1 \end{bmatrix}$ $\det(B) = 1 \cdot (-1) - 1 \cdot 3 = -4$

$\det(A) = \det(B)$, but A is not similar to B because Thm 4.22 (4) is not satisfied.

Diagonalisation

Definition An $n \times n$ matrix A is *diagonalisable* if it is similar to a diagonal matrix.

That is A is diagonalisable if there exists an invertible matrix P and a diagonal matrix D such that $P^{-1}AP = D$.

Theorem 4.23 Let A be an $n \times n$ matrix. Then A is diagonalisable if and only if it has n linearly independent eigenvectors.

Moreover, there exists an invertible matrix P and a diagonal matrix D such that $P^{-1}AP = D$ if and only if:

- The columns of P are n linearly independent eigenvectors of A
- The diagonal entries of D are the eigenvalues of A corresponding to the eigenvectors in P (in the same order)

$Proof$: Suppose A is $n \times n$ diagonisable and so there exists an invertible matrix P such that $P^{-1}AP = D$, equivalently $AP = PD$.

Let $P = [\underline{p}_1, \underline{p}_2, \ldots, \underline{p}_n]$, where $\underline{p}_i$ is the ith column of P, and

$D = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}$, where $\lambda_i \in \mathbb{R}$.

Then $AP = PD \Rightarrow A[\underline{p}_1, \underline{p}_2, \ldots, \underline{p}_n] = [\underline{p}_1, \underline{p}_2, \ldots, \underline{p}_n] \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}$.

Note: From my understanding, we do not handle $\underline{p}_i$s as vectors. Just elements of a row vector.

$\Rightarrow [A\underline{p}_1, A\underline{p}_2, \ldots, A\underline{p}_n] = [\lambda_1\underline{p}_1, \ldots, \lambda_n\underline{p}_n]$

Equating columns gives:

$A\underline{p}_1 = \lambda_1 \underline{p}_1, \dots, A\underline{p}_n = \lambda_n \underline{p}_n$ (which is exactly what happens when λ_1 being an eigenvalue of A, with $\underline{p}_1$ being an eigenvector)

The columns of P are eigenvectors of A wuth corresponding eigenvalues being the entries in D, written in the same order.

The columns of P are linearly independent since P is invertible. So A has n linearly independent eigenvectors.

Now, we have to prove the converse.

Suppose A has n linearly independent eigenvectors, $\underline{p}_1, \dots, \underline{p}_n \in \mathbb{R}^n$, with corresponding eigenvalues $\lambda_1, \dots, \lambda_n \in \mathbb{R}$.

Then $A\underline{p}_1 = \lambda_1 \underline{p}_1, \dots, A\underline{p}_n = \lambda_n \underline{p}_n$.

Let P be the matrix with columns $\underline{p}_1, \dots, \underline{p}_n$.

Then we reverse the steps from the first part of the proof, to get $AP = PD$, where D is the diagonal matrix with entries $\lambda_1, \dots, \lambda_n$.

Since columns of P are linearly independent, P is invertible. So $P^{-1}AP = D$.

So A is diagonalisable as it is similar to a diagonal matrix. $\square$

Example Is

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

diagonalisable?

$$A - \lambda \mathbb{I}_3 = \begin{bmatrix} 1-\lambda & 2 & 0 \\ -1 & -1-\lambda & 1 \\ 0 & 1 & 1-\lambda \end{bmatrix}$$
$$|A - \lambda \mathbb{I}_3| = (1-\lambda)\begin{vmatrix} -1-\lambda & 1 \\ 1 & 1-\lambda \end{vmatrix} - 2\begin{vmatrix} -1 & 1 \\ 0 & 1-\lambda \end{vmatrix} + 0$$
$$= (1-\lambda)(\lambda^2 - 1 - 1) - 2(\lambda - 1) = (1-\lambda)(\lambda^2 - 2 + 2) = \lambda^2(1-\lambda)$$

So A has eigenvalues $\lambda_1 = 1$ and $\lambda_2 = 0$ and eigenspaces

$E_1 = \text{span}\left(\begin{bmatrix}1\\0\\1\end{bmatrix}\right)$, $E_0 = \text{span}\left(\begin{bmatrix}2\\-1\\1\end{bmatrix}\right)$.

Since there are only 2 linearly independent eigenvectors, A is not diagonalisable.

Example If possible find a matrix P that diagonalises $A = \begin{pmatrix}2 & 1\\1 & 2\end{pmatrix}$.

We must find all λ first such that $\det(A - \lambda\mathbb{I}) = 0$.

$\det(A - \lambda\mathbb{I}) = \det\begin{pmatrix}2-\lambda & 1\\1 & 2-\lambda\end{pmatrix} = \lambda^2 - 4\lambda + 3 = (\lambda - 3)(\lambda - 1)$

Setting $\det(A - \lambda\mathbb{I}) = 0$ and solving for λ gives the eigenvalues of A as

$\lambda_1 = 3,\ \lambda_2 = 1.$

Eigenspace for $\lambda = 1$:

Consider $(A - 1\cdot\mathbb{I}|0) = \begin{pmatrix}2-\lambda & 1 & | & 0\\1 & 2-\lambda & | & 0\end{pmatrix} = \begin{pmatrix}1 & 1 & | & 0\\1 & 1 & | & 0\end{pmatrix} = \begin{pmatrix}1 & 1 & | & 0\\0 & 0 & | & 0\end{pmatrix}$

The general solution for vector $\underline{x} = \begin{bmatrix}x_1\\x_2\end{bmatrix} \in \text{null}(A - \mathbb{I})$ is $x_1 + x_2 = 0 \Rightarrow$

$x_1 = -x_2.$

Hence $E_1 = \left\{x_2\begin{bmatrix}-1\\1\end{bmatrix} : x_2 \in \mathbb{R}\right\} = \text{span}\left(\begin{bmatrix}-1\\1\end{bmatrix}\right).$

and similarly, it can be shown that $E_3 = \text{span}\left(\begin{bmatrix}1\\1\end{bmatrix}\right).$

Since A has two linearly independent eigenvectors, A is diagonalisable, where

$P = \begin{bmatrix}1 & -1\\1 & 1\end{bmatrix}$, $D = \begin{bmatrix}3 & 0\\0 & 1\end{bmatrix}$. (Note, order matters)

$\begin{bmatrix}-1\\1\end{bmatrix}$ is the second column, so $\lambda_2 = 1$ is in the second column

$\begin{bmatrix}1\\1\end{bmatrix}$ is the first column, so $\lambda_1 = 3$ is in the first column

and $P^{-1}AP = D$.

Combining the previous result with Theorem 4.20, we get the following criteria for diagonalisability.

Theorem 4.25 Let A be an $n \times n$ matrix with n distinct eigenvalues. Then A is diagonalisable.

$Proof$: The eigenvectors corresponding to the n-distinct eigenvalues are linearly independent (by Theorem 4.20), so the result follows from Theorem 4.23.

Theorem 4.27 (The diagonalisation theorem) Let A be an $n \times n$ matrix whose distinct eigenvalues are $\lambda_1, \ldots, \lambda_k$. The following are equivalent:

a) A is diagonalisable

b) The union of a basis for each of the eaigenspaces of A contains n vectors

c) The algebraic multiplicity of each eigenvalue is equal to its geometric multiplicity and the sum of these multiplicites across all eigenvalues

Computing powers of a square matrix

If D is an $n \times n$ diagonal matrix, say $D = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$, then the k-th power of D can be computed as

$$D^k = \operatorname{diag}(\lambda_1^k, \ldots, \lambda_n^k).$$

Now if A is an $n \times n$ diagonalisable matrix, there exists an invertible matrix P such that $P^{-1}AP$ is a diagonal matrix. Let $D := P^{-1}AP$ be this diagonal matrix. Then we clearly have $A = PDP^{-1}$. It follows that

$$A^k = (PDP^{-1})(PDP^{-1}) \ldots (PDP^{-1}) = PD^kP^{-1}.$$

Example Let $A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$ find A^k for any $k \in \mathbb{N}$.

In previous example we showed $P^{-1}AP = \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix}$, where $P = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$.

Let $D = \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix}$, then for any $k \in \mathbb{N}$, we have

$$A^k = (PDP^{-1})^k = PD^kP^{-1} = P \begin{bmatrix} 3^k & 0 \\ 0 & 1^k \end{bmatrix} P^{-1}$$

Computing $P^{-1} = 1/2 \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$ we conclude

$$A^k = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 3^k & 0 \\ 0 & 1 \end{bmatrix} \cdot 1/2 \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} = 1/2 \begin{bmatrix} 3^k + 1 & 3^k - 1 \\ 3^k - 1 & 3^k + 1 \end{bmatrix}.$$

Orthogonality and Gram-Schmidt

Inner product

Definition Let $\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$ be vectors in $\mathbb{R}^n$ then *the dot product* $\mathbf{u} \cdot \mathbf{v}$ is defined by

$$\mathbf{u} \cdot \mathbf{v} = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n.$$

The dot product of vectors in $\mathbb{R}^n$ is a special case of a more general notion of *inner product*.

Example Consider the vectors $\mathbf{u} = \begin{bmatrix} 7 & -2 & 1 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} -1 & 1 & 2 \end{bmatrix}$. Compute $\mathbf{u} \cdot \mathbf{v}$.

$\mathbf{u} \cdot \mathbf{v} = -7 - 2 + 2 = -7$

The fundamental properties of inner products in $\mathbb{R}^n$ are given by:

Theorem 1.2 Properties of inner products Let $\mathbf{u}$, $\mathbf{v}$, $\mathbf{w} \in \mathbb{R}^n$ and $c \in \mathbb{R}$. Then

1. $\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$ (commutativity of inner product)
2. $\mathbf{u} \cdot (\mathbf{v} + \mathbf{w}) = \mathbf{u} \cdot \mathbf{v} + \mathbf{u} \cdot \mathbf{w}$ (distributivity of inner product)
3. $(c\mathbf{u}) \cdot \mathbf{v} = c(\mathbf{u} \cdot \mathbf{v})$
4. $\mathbf{u} \cdot \mathbf{u} \geq 0$ and $\mathbf{u} \cdot \mathbf{u} = 0$ if and only if $\mathbf{u} = \mathbf{0}$.

$Proof$: of (1) Applying the definition of $\mathbf{u} \cdot \mathbf{v}$ and $\mathbf{v} \cdot \mathbf{u}$

$\mathbf{u} \cdot \mathbf{v} = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n = v_1 u_1 + v_2 u_2 + \cdots + v_n u_n = \mathbf{v} \cdot \mathbf{u}$,

since $u_i,\ v_i \in \mathbb{R}$

Length or norm of a vector

Definition Let $\mathbf{v} = (v_1, v_2, \ldots, v_n)$ be a vector in $\mathbb{R}^n$ then the *lenght* or *norm* of $\mathbf{v}$ is non-negative scalar $||\mathbf{v}||$ defined by

$$||\mathbf{v}|| = \sqrt{\mathbf{v} \cdot \mathbf{v}} = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2}.$$

The vector with norm 1 is called a unit vector. We can always find a unit vector in the same direction as $\mathbf{v}$ by dividing by $||\mathbf{v}||$.

The fundamental properties of norms in $\mathbb{R}^n$ are given by:

Theorem 1.3 Properties of norms Let $\mathbf{v} \in \mathbb{R}^n$ and $c \in \mathbb{R}$. Then

1. $||\mathbf{v}|| = 0$ if and only if $\mathbf{v} = \mathbf{0}$.
2. $||c\mathbf{v}|| = |c| \cdot ||\mathbf{v}||$

$Proof$: of (2)

$||c\mathbf{v}||^2 = (c\mathbf{v}) \cdot (c\mathbf{v}) = c^2(\mathbf{v} \cdot \mathbf{v}) = c^2||\mathbf{v}||^2$

Taking square root of both sides and using $\sqrt{c^2} = |c|$, then we get $||c\mathbf{v}|| = |c| \cdot ||\mathbf{v}||$.

Theorem 1.4 Cauchy-Schwartz Inequality Let $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$. Then

$$|\mathbf{u} \cdot \mathbf{v}| \leq ||\mathbf{u}|| \, ||\mathbf{v}||$$

Moreover, we have equality if and only if $\mathbf{u}$ and $\mathbf{v}$ are linearly dependent.

$Proof$:

If $\mathbf{u} = \mathbf{0}$ then $\mathbf{u} \cdot \mathbf{v} = 0$ and $||\mathbf{u}|| = 0$ and so both sides of the inequality are zero.

So we only need to consider the case $\mathbf{u} \neq 0$.

For any $t \in \mathbb{R}$ we have $0 \leq (\mathbf{u} + t\mathbf{v}) \cdot (\mathbf{u} + t\mathbf{v}) = \mathbf{u} \cdot \mathbf{u} + (t\mathbf{v}) \cdot \mathbf{u} + \mathbf{u} \cdot (t\mathbf{v}) + (t\mathbf{v}) \cdot (t\mathbf{v}) = \mathbf{u} \cdot \mathbf{u} + 2t(\mathbf{u} \cdot \mathbf{v}) + t^2(\mathbf{v} \cdot \mathbf{v}) = c + bt + at^2$

The inequlity implies that $at^2 + bt + c$ has either no real roots or a repeated real root.

So we require $b^2 - 4ac \leq 0$ or equivalently $(2\mathbf{u} \cdot \mathbf{v})^2 - 4(\mathbf{v} \cdot \mathbf{v})(\mathbf{u} \cdot \mathbf{u}) \leq 0$

Hence we require $(\mathbf{u} \cdot \mathbf{v})^2 \leq (\mathbf{v} \cdot \mathbf{v})(\mathbf{u} \cdot \mathbf{u})$.

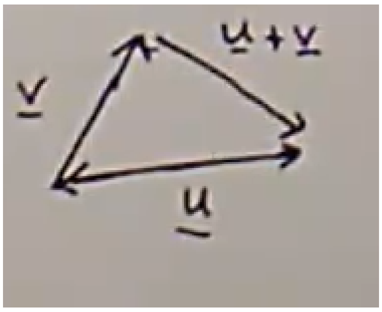
Taking the square roots of both sides and using the fact that $\mathbf{u} \cdot \mathbf{u} \geq 0$, $\mathbf{v} \cdot \mathbf{v} \geq 0$

we have $|\mathbf{u} \cdot \mathbf{v}| \leq ||\mathbf{u}|| \, ||\mathbf{v}||$. $\square$

Theorem 1.5 The Triangle Inequality Let $\mathbf{u}, \ \mathbf{v} \in \mathbb{R}^n$. Then

$$||\mathbf{u} + \mathbf{v}|| \leq ||\mathbf{u}|| + ||\mathbf{v}||$$

$Proof$:



Orthogonality

Now we have talked about the lenght of a vector it is natural to talk about the angle between two non-zero vectors $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$.

The angle θ is given by

$$\theta = \cos^{-1}\left(\frac{\mathbf{u} \cdot \mathbf{v}}{||\mathbf{u}|| \, ||\mathbf{v}||}\right), \quad \text{where } 0 \leq \theta \leq \pi$$

Definition Let $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ be a set of vectors in $\mathbb{R}^n$. S is an *orthogonal set* if

$$\mathbf{v}_i \cdot \mathbf{v}_j = 0 \text{ whenever } i \neq j, \forall i, j = 1, 2, \ldots, k.$$

That is, S is an orthogonal set of and only if every distinct pair of vectors in S are ortogonal.

The standard basis $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ of $\mathbb{R}^n$ is an orthogonal set.

Theorem 5.1 Let $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ be an orthogonal set of non-zero vectors in $\mathbb{R}^n$. Then S is linearly independent set.

$Proof$: Let $c_1, c_2, \ldots, c_k \in \mathbb{R}$ such that $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k = \mathbf{0}$.

So $\mathbf{v}_i \cdot (c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k) = \mathbf{v}_i \cdot 0 = \mathbf{0}$

then $c_1\mathbf{v}_i \cdot \mathbf{v}_1 + c_2\mathbf{v}_i \cdot \mathbf{v}_2 + \cdots + c_k\mathbf{v}_i \cdot \mathbf{v}_i + \cdots + c_k\mathbf{v}_i \cdot \mathbf{v}_k = \mathbf{0}\ (1)$

Since S is an orthogonal set all the dot products are xero except $\mathbf{v}_i \cdot \mathbf{v}_i$, so (1) becomes

$$c_i(\mathbf{v}_i \cdot \mathbf{v}_i) = \mathbf{0}$$

since $\mathbf{v}_i \neq \mathbf{0}$, we have $c_i = 0$. This is true for all $i = 1, \ldots, k$. So S us a linearly independent set. $\square$

Definition A basis $\mathcal{B}$ for a subspace V of $\mathbb{R}^n$ is an *orthogonal basis* for V if and only if $\mathcal{B}$ is an orthogonal set.

Coordinates relative to an orthogonal basis

If a basis is orthogonal the following theorem gives a simple method of finding the coordinate vectors with respect to this basis.

Theorem 5.2 Let $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ be an orthogonal basis for subspace V of $\mathbb{R}^n$. For any $\mathbf{v} \in V$ then there are $c_!, c_2, \ldots, c_k \in \mathbb{R}$ sych that

$$\mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k,$$

where $c_i = \frac{\mathbf{v} \cdot \mathbf{v}_i}{\mathbf{v}_i \cdot \mathbf{v}_i}$, for $i = 1, \ldots, k$.

$Proof$: Since S is a basis for V every $\mathbf{v} \in\ V$ can be expressed as

$\mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k$

Taking the dot product with $\mathbf{v}_i$ gives

$\mathbf{v}_i \cdot \mathbf{v} = c_1\mathbf{v}_i \cdot \mathbf{v}_1 + c_2\mathbf{v}_i \cdot \mathbf{v}_2 + \cdots + c_k\mathbf{v}_i \cdot \mathbf{v}_k,$

Since S is an orthogonal set, all the dot products are zero except the one with $\mathbf{v}_i \cdot \mathbf{v}_i$, so

$$\mathbf{v} \cdot \mathbf{v}_i = c_i\mathbf{v}_i \cdot \mathbf{v}_i$$

Hence $c_i = \frac{\mathbf{v} \cdot \mathbf{v}_i}{\mathbf{v}_i \cdot \mathbf{v}_i}$. $\square$

Example Find the coordinates of $\mathbf{u} = [1, 1, 1]$ with respect to the orthogonal basis $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ where $\mathbf{v}_1 = [0, 1, 0]$, $\mathbf{v}_2 = [-4, 0, 3]$, $\mathbf{v}_3 = [3, 0, 4]$.

Let $\mathbf{u} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3$, where

$c_1 = \frac{0+1+0}{0+1+0} = 1$

$c_2 = \frac{\mathbf{u}\cdot\mathbf{v}_2}{\mathbf{v}_2\cdot\mathbf{v}_2} = \frac{-4+3}{16+9} = -\frac{1}{25}$

$c_3 = \frac{\mathbf{u}\cdot\mathbf{v}_3}{\mathbf{v}_3\cdot\mathbf{v}_3} = \frac{3+4}{9+16} = \frac{7}{25}$

Thus $\mathbf{u} = \mathbf{v}_1 - \frac{1}{25}\mathbf{v}_2 + \frac{7}{25}\mathbf{v}_3$

Hence $[\mathbf{u}]_{\mathcal{B}} = \begin{bmatrix} 1 \\ -\frac{1}{25} \\ \frac{7}{25} \end{bmatrix}$

Definition Let $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ be a set of vectors in $\mathbb{R}^n$. S is an *orthonormal set* if S is an orthogonal set of unit vectors (i.e. S is an orthogonal set and $||\mathbf{v}_i|| = 1$, for each $\mathbf{v}_i \in S$). Set S is an *orthobormal basis* for subspace V of $\mathbb{R}^n$ if it is a basis for V and is an orthonormal set.

Gram-Schmidt process

We can always find an orthonormal basis for a subspace $V$of $\mathbb{R}^n$.

Theorem 5.15 The Gram-Schmidt Process Let $\{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_n\}$ be a basis for subspace W of $\mathbb{R}^n$. Define:

$$\begin{aligned} \mathbf{v}_1 &= \mathbf{w}_1 \\ \mathbf{v}_2 &= \mathbf{w}_2 - \left(\frac{\mathbf{w}_2 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\right)\mathbf{v}_1 \\ \mathbf{v}_3 &= \mathbf{w}_3 - \left(\frac{\mathbf{w}_3 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\right)\mathbf{v}_1 - \left(\frac{\mathbf{w}_3 \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2}\right)\mathbf{v}_2 \\ &\vdots \\ \mathbf{v}_n &= \mathbf{w}_n - \sum_{i=1}^{n-1}\left(\frac{\mathbf{w}_n \cdot \mathbf{v}_i}{\mathbf{v}_i \cdot \mathbf{v}_i}\right)\mathbf{v}_i \end{aligned}$$

The vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ are an orthogonal basis for W. If we set $\mathbf{u}_i = \frac{\mathbf{v}_i}{||\mathbf{v}_i||}$ then $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n$ are an orthonormal basis for W.

The above process does not affect the subspace generated by these vectors, i.e. $\text{Span}\{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_n\} = \text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$.

Example Construct an orthonormal basis for the subspace $W = \text{Span}\{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_n\}$ of $\mathbb{R}^4$, where

$\mathbf{w}_1 = (1, 0, 0, -1), \mathbf{w}_2 = (0, 1, 0, 2), \mathbf{w}_3 = (0, 0, 1, 2)$.

We first note W is a linearly independent set.

We can now apply the Gram-Schmidt process.

$\mathbf{v}_1 = \mathbf{w}_1$

$\mathbf{v}_2 = (0, 1, 0, 2) - \left(\frac{0+0+0-2}{1+0+0+1}\right)(1, 0, 0, -1) = (0, 1, 0, 2) + (1, 0, 0, -1) = (1, 1, 0, 1)$

$\mathbf{v}_3 = (0, 0, 1, 2) - \frac{0+0+0-2}{2}(1, 0, 0, -1) - \frac{0+0+0-2}{1+1+0+1}(1, 1, 0, 1)$
$= (0, 0, 1, 2) + (1, 0, 0, -1) - \frac{2}{3}(1, 1, 0, 1) = (\frac{1}{3}, -\frac{2}{3}, 1, \frac{1}{3})$

$\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ are an orthogonal basis

We normalize the vectors, $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$

$\mathbf{u}_1 = \frac{\mathbf{v}_1}{||\mathbf{v}_1||} = \frac{(1,0,0,-1)}{\sqrt{2}}$

$\mathbf{u}_2 = \frac{\mathbf{v}_2}{||\mathbf{v}_2||} = \frac{1}{\sqrt{3}}(1, 1, 0, 1)$

$\mathbf{u}_3 = \frac{\mathbf{v}_3}{||\mathbf{v}_3||} = \frac{(\frac{1}{3}, -\frac{2}{3}, 1, \frac{1}{3})}{\sqrt{\frac{5}{3}}} = \frac{1}{\sqrt{15}}(1, -2, 3, 1)$

So $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$ is an orthonormal basis for W.

Example Find an orthogonal basis for $\mathbb{R}^3$ that contains $\mathbf{w} = (2, 1, 0)$.

We first find a basis for $\mathbb{R}^3$ containing $\mathbf{w}$.

We can extend $\mathbf{w}$ to a basis by considering $\mathbf{w}$ together with the standard basis vectors

$\{\mathbf{w}, \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ which spans $\mathbb{R}^3$.

Construct a matrix A where $\mathbf{w}, \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ are the rows. We know $\text{row}(A) = \mathbb{R}^3$, but to find a basis we row reduce A.

$$A = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

The row of zeros tells us the second vector is not necessary. So a basis for $\mathbb{R}^3$ is $\{\mathbf{w}, \mathbf{e}_2, \mathbf{e}_3\}$

We apply Gram-Schmidt process:

Let $\mathbf{v}_1 = \mathbf{w}$

$\mathbf{v}_2 = \mathbf{e}_2 - \frac{\mathbf{e}_2 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 = (-\frac{2}{5}, \frac{4}{5}, 0)$ (as $\mathbf{v}_1 = \mathbf{w}$)

$\mathbf{v}_3 = \mathbf{e}_3 - \frac{\mathbf{e}_3 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 - \frac{\mathbf{e}_3 \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2}\mathbf{v}_2 = (0,0,1) - 0\mathbf{v}_1 - 0\mathbf{v}_2 = (0,0,1)$

The orthogonal basis for $\mathbb{R}^3$ containing $\mathbf{w}$ is:

$$\begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -\frac{2}{5} \\ \frac{4}{5} \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

Orthogonal matrices

Definition An orthogonal matrix is an $n \times n$ matrix Q whose columns form an orthonormal set.

Theorem 5.5 A square matrix Q is orthogonal if and only if $Q^{-1} = Q^T$.

$Proof$: Q is invertible and $Q^{-1} = Q^T$ if and only if $Q^TQ = \mathbb{I}$ by Theorem 3.13.

So we need to show Q is orthogonal if and only if $Q^TQ = \mathbb{I}$.

Let $\mathbf{q}_i$ be the ith column of Q, hence the ith row of Q^T.

The (i,j) entry of Q^TQ is the dot product of the ith row of Q^T and the jth column of Q. i.e.

$$(Q^TQ)_{ij} = \mathbf{q}_i \cdot \mathbf{q}_j$$

Since the columns of Q form an orthonormal set if and only if

$$\mathbf{q}_i \cdot \mathbf{q}_j = 0 \text{ if } i \neq j \wedge 1 \text{ if } i = j$$

or equivalently if and only if

$$(Q^TQ)_{ij} = 0 \text{ if } i \neq j \wedge 1 \text{ if } i = j$$

which is exactly the entries of the identity matrix. $\square$

Example Show the rotation matrix $Q = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$ is an orthogonal matrix for every $\theta \in \mathbb{R}$.

$\begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix} \cdot \begin{bmatrix} -\sin\theta \\ cos\theta \end{bmatrix} = -\cos\theta\sin\theta + \sin\theta\cos\theta = 0$

$\left|\left|\begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}\right|\right| = \sqrt{\cos^2\theta + \sin\theta^2} = 1$

Similarly, we can show that the norm of second column is 1.

Hwence the columns of Q form an orthonormal set.

Check: $Q^{-1} = Q^T$.

Theorem 5.6 Let Q be an $n \times n$ matrix. The following are equivalent,

a) Q is orthogonal

b) $||Q\mathbf{x}|| = ||\mathbf{x}||$ for all $\mathbf{x} \in \mathbb{R}^n$

c) $Q\mathbf{x} \cdot Q\mathbf{y} = \mathbf{x} \cdot \mathbf{y}$ for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$

This says every orthogonal matrix is an isometry, that is the matrix transformation preserves length.

$Proof$:

Show $(a) \Rightarrow (c) \Rightarrow (b) \Rightarrow (a)$.

Note since $\mathbf{x}, \mathbf{y}$ are column vectors in $\mathbb{R}^n$ then $\mathbf{x} \cdot \mathbf{y} = \mathbf{x}^T\mathbf{y}$

$(a) \Rightarrow (c)$ Assume Q is orthogonal so $Q^TQ = \mathbb{I}$.

$Q\mathbf{x} \cdot Q\mathbf{y} = (Q\mathbf{x})^T(Q\mathbf{y}) = \mathbf{x}^TQ^TQ\mathbf{y} = \mathbf{x}^T\mathbf{y} = \mathbf{x} \cdot \mathbf{y}$

$(c) \Rightarrow (b)$ Assume $Q\mathbf{x} \cdot Q\mathbf{y} = \mathbf{x} \cdot \mathbf{y}$ for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$.

$Q\mathbf{x} \cdot Q\mathbf{x} = \mathbf{x} \cdot \mathbf{x}$. Hence $||Q\mathbf{x}||^2 = ||\mathbf{x}||^2$ so $||Q\mathbf{x}|| = ||\mathbf{x}||$.

$(b) \Rightarrow (a)$ Assume $||Q\mathbf{x}|| = ||\mathbf{x}||$ for all $\mathbf{x} \in \mathbb{R}^n$. Then

$Q\mathbf{x} \cdot Q\mathbf{y} = \frac{1}{4}||Q\mathbf{x} + Q\mathbf{y}||^2 - \frac{1}{4}||Q\mathbf{x} - Q\mathbf{y}||^2$
$= \frac{1}{4}||Q(\mathbf{x} + \mathbf{y})||^2 - \frac{1}{4}||Q(\mathbf{x} - \mathbf{y})||^2$

by (b) $||Q\mathbf{x}|| = ||\mathbf{x}||$, so

$\frac{1}{4}||Q(\mathbf{x} + \mathbf{y})||^2 - \frac{1}{4}||Q(\mathbf{x} - \mathbf{y})||^2 = \frac{1}{4}||\mathbf{x} + \mathbf{y}||^2 - \frac{1}{4}||\mathbf{x} - \mathbf{y}||^2$
$= \mathbf{x} \cdot \mathbf{y}$ for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$

(Now we have shown that (b) implies (c), but we have to now use it to show (b) implies (a))

If $\mathbf{e}_i$ is the ith standard basis vector for $\mathbb{R}^n$ then $\mathbf{q}_i = Q\mathbf{e}_i$ (the ith column of Q).

Consequently,

$$\mathbf{q}_i \cdot \mathbf{q}_j = (Q\mathbf{e}_i) \cdot (Q\mathbf{e}_j) = \mathbf{e}_i \cdot \mathbf{e}_j \text{ by } (c)$$
$$= 0 \text{ if } i \neq j \wedge 1 \text{ if } i = j$$

Thus the columns of Q form an orthonormal set, so Q is an orthogonal matrix. $\square$

Theorem 5.8 Let Q be an $n \times n$ orthogonal matrix. Then,

(a) Q^{-1} is orthogonal

(b) $\det Q = \pm 1$

(c) If λ is an eigenvalue of Q, then $|\lambda| = 1$

(d) If Q_1 and Q_2 are orthogonal $n \times n$ matrices then so is $Q_1 Q_2$

$Proof$: of (b)

$Q^{-1} = Q^T$ since Q is an orthogonal matrix.

Taking determinants: $\det(Q^{-1}) = \det(Q^T)$

But, $\det(Q^{-1} = \frac{1}{\det(Q)}$ and $\det(Q^T) = \det(Q)$

Hence, $\frac{1}{\det(Q)} = \det(Q)$, and $1 = (\det(Q))^2$ since $\det(Q) \in \mathbb{R}$

we must have $\det(Q) = \pm 1$.

$Proof$: of (c) Let λ be an eigenvalue of Q with eigenvector $\mathbf{v}$. Then $Q\mathbf{v} = \lambda\mathbf{v}$.

By theorem 5.6 (b) we have

$||\mathbf{v}|| = ||Q\mathbf{v}|| = ||\lambda\mathbf{v}|| = |\lambda|\,||\mathbf{v}||$

Since $||\mathbf{v}|| \neq 0$, we have $|\lambda| = 1$.

Orthogonal diagonalisation and quadratic forms

Orthogonal diagonalisation of symmetric matrices

Square real valued matrices do not necessarily have real eigenvalues and not all square matrices are diagonalisable. Are there any matrices that do behave nicely?

Example If possible diagonalise $\begin{bmatrix} 2 & 6 \\ 6 & 18 \end{bmatrix}$.

The chracteristic polynomial for A is $\lambda^2 - 20\lambda = \lambda(\lambda - 20)$.

Hence the eigenvalues of A are $\lambda_1 - 0$ and $\lambda_2 = 20$.

Solving $A\underline{v}_i = \lambda_i \underline{v}_i$, $i = 1, 2$ to find the corresponding eigenvectors $\underline{v}_1$, $\underline{v}_2$.

$\lambda_1 = 0$: $A\underline{v}_1 = \underline{0}$, hence $[A|\underline{0}] \sim \begin{bmatrix} 1 & 3 & | & 0 \\ 0 & 0 & | & 0 \end{bmatrix}$

Let $\underline{v}_1 = \begin{pmatrix} x \\ y \end{pmatrix}$ then $x + 3y = 0$.

So $E_0 = \left\{ y \begin{pmatrix} -3 \\ 1 \end{pmatrix} : y \in \mathbb{R} \right\} = \operatorname{span}\left(\left\{ \begin{pmatrix} -3 \\ 1 \end{pmatrix} \right\} \right)$

$\lambda_2 = 20$: $A\underline{v}_2 = \underline{0}$, hence $[A|\underline{0}] \sim \begin{bmatrix} -3 & 1 & | & 0 \\ 0 & 0 & | & 0 \end{bmatrix}$

so let $\underline{v}_2 = \begin{pmatrix} x \\ y \end{pmatrix}$ then $-3x + y = 0$.

So $E_{20} = \left\{ y \begin{pmatrix} \frac{1}{3} \\ 1 \end{pmatrix} : y \in \mathbb{R} \right\} = \operatorname{span}\left(\left\{ \begin{pmatrix} \frac{1}{3} \\ 1 \end{pmatrix} \right\} \right)$

So A is diagonalisable and we set $P = \begin{bmatrix} -3 & \frac{1}{3} \\ 1 & 1 \end{bmatrix}$

and we have $P^{-1}AP = D = \begin{bmatrix} 0 & 0 \\ 0 & 20 \end{bmatrix}$.

Notice $\begin{bmatrix} -3 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} \frac{1}{3} \\ 1 \end{bmatrix}$ are orthogonal. (They are perpendicular to each other.)

If we normalise them let

$\underline{u}_1 = \begin{bmatrix} \frac{-3}{\sqrt{10}} \\ \frac{1}{\sqrt{10}} \end{bmatrix}$, $\underline{u}_2 = \begin{bmatrix} \frac{1}{\sqrt{10}} \\ \frac{3}{\sqrt{10}} \end{bmatrix}$

Take $Q = [\underline{u}_1 \ \underline{u}_2] = \begin{bmatrix} \frac{-3}{\sqrt{10}} & \frac{1}{\sqrt{10}} \\ \frac{1}{\sqrt{10}} & \frac{3}{\sqrt{10}} \end{bmatrix}$

$Q^{-1}AQ = D$. But Q is an orthogonal matrix. So $Q^{-1} = Q^T$, so $Q^T AQ = D$.

Definition Let A be an $n \times n$ matrix. A is *orthogonally diagonasible* if there exists an orthogonal matrix Q and a diagonal matrix D such that $Q^T AQ = D$
.

Theorem 5.17 If A is orthogonally diagonalisable, then A is symmetric.

Note: If M is a symmetric matrix, then $M = M^T$.

$Proof$: There exists an orthogonal matrix Q and a diagonal matrix D such that $Q^T AQ = D$. We also have $Q^{-1} = Q^T$ and so $Q^T Q = \mathbb{I} = QQ^T$.

$QDQ^T = QQ^T AQQ^T = \mathbb{I}A\mathbb{I} = A$.

Then $A^T = (QDQ^T)^T = (Q^T)^T D^T Q^T = QDQ^T = A$.

Since $D^T = D$. Hence, we've shown that A is symmetric.

Theorem 5.18 and 5.19 If A is a real symmetric matrix then

a) The eigenvalues of A are all real.

b) Eigenvector from different eigenspaces are orthogonal.

$Proof$:

Remarks:

IF $z = a + bi$, $a, b \in \mathbb{R}$ the $\overline{z} = a - bi$

z is real if $b = 0$, and so $z = \overline{z}$.

(i) If $A = [a_{ij}]$ is an $n \times n$ matrix then the complex conjugate of A is $\overline{A} = [\overline{a}_{ij}]$.

(ii) The rules for complex conjugates extend to matrices so $\overline{AB} = \overline{A}\,\overline{B}$.

Proof of $a)$ Suppose λ os an eigenvlaue of A with corresponding eigenvector $\underline{v}$, so $A\underline{v} = \lambda\underline{v}$. Taking complex conjugates gives $\overline{A\underline{v}} = \overline{\lambda\underline{v}}$, since A is real we have:

$A\overline{\underline{v}} = \overline{\lambda}\overline{\underline{v}}$.

Taking transposes and using the fact that $A = A^T$:

$$\overline{\underline{v}}^T A = \overline{\underline{v}}^T A^T \Leftarrow (A\overline{\underline{v}})^T = (\overline{\lambda}\overline{\underline{v}})^T \Rightarrow \overline{\lambda}\overline{\underline{v}}^T$$

Multiplying on the right by $\underline{v}$.

$\lambda\overline{\underline{v}}^T\underline{v} = \overline{\underline{v}}^T(\lambda\underline{v}) = \overline{\underline{v}}^T A\underline{v} = \overline{\lambda}\overline{\underline{v}}^T\underline{v}$

Then:

$(\lambda - \overline{\lambda})\overline{\underline{v}}^T\underline{v} = \underline{0}$

If $\underline{v} = \begin{bmatrix} a_1 + b_1 i \\ \vdots \\ a_n + b_n i \end{bmatrix}$, $a_i, b_i \in \mathbb{R}$ then $\overline{\underline{v}} = \begin{bmatrix} a_1 - b_1 i \\ \vdots \\ a_n - b_n i \end{bmatrix}$.

So $\overline{\underline{v}}^T\underline{v} = (a_1^2 + b_1^2) + \cdots + (a_n^2 + b_n^2) \neq 0$ since $\underline{v} \neq 0$.

So $\lambda = \overline{\lambda}$, so λ is real.

Proof of $b)$ Let $\underline{v}_1$ and $\underline{v}_2$ be eigenvectors corresponding to distinct eigenvalues λ_1, λ_2 of A.

So $A\underline{v}_1 = \lambda_1\underline{v}_1$. Taking the dot product with $\underline{v}_2$:

$$\underline{v}_1^T A^T\underline{v}_2 = (A\underline{v}_1)^T = \underline{v}_2 \Leftarrow A\underline{v}_1 \cdot \underline{v}_2 = \lambda_1\underline{v}_1 \cdot \underline{v}_2$$

Then: $\lambda_1\underline{v}_1 \cdot \underline{v}_2 = \underline{v}_1^T A^T\underline{v}_2 = \underline{v}_1^T A\underline{v}_2 = \underline{v}^T\lambda_2\underline{v}_2 = \lambda_2\underline{v}_1^T\underline{v}_2 = \lambda_2\underline{v}_1 \cdot \underline{v}_2$

So $(\lambda_1 - \lambda_2)(\underline{v}_1 \cdot \underline{v}_2) = \underline{0}$, but $\lambda_1 \neq \lambda_2$ (they are distinct), so $\underline{v}_1 \cdot \underline{v}_2 = 0$ as required. $\square$

Theorem 5.20 The Spectral Theorem Let A be an $n \times n$ matrix. Then A is symmetric if and only if A is orthogonally diagonalisable.

$Proof$: We proved " $\Rightarrow$ " in Theorem 5.17.

So we now show that if $A = A^T$ then there exists an orthogonal matrix Q and a diagonal matrix D such that $Q^T AQ = D$.

We prove this by induction on the size n of the matrix.

- $n = 1$: Nothing to prove since $|x|$ matrices are already diagonal.
- Assume true for $n = k$: Assume every $k \times k$ matrix A that is real and symmetric is also orthogonally diagonalisable.
- $n = k + 1$
  Let λ_1 be an eigenvalue of A and let $\underline{v}_1$ be the corresponding eigenvector.
  So $A\underline{v}_1 = \lambda_1 \underline{v}_1$. By Thm 5.18 λ_1 is real and so is $\underline{v}_1$.
  We can assume $\underline{v}_1$ is a unit vector because we can normalise $\underline{v}_1$.
  Using the Gram-Schmidt process we extend $\underline{v}_1$ to an orthonormal basis $\{\underline{v}_1, \ldots, \underline{v}_n\}$ of $\mathbb{R}^n$. Let $Q_1 = [\underline{v}_1, \underline{v}_2, \ldots, \underline{v}_n]$ which is an orthogonal matrix.
  $$Q_1^T AQ_1 = \begin{bmatrix} \underline{v}_1^T \\ \vdots \\ \underline{v}_n^T \end{bmatrix} A \begin{bmatrix} \underline{v}_1 & \ldots & \underline{v}_n \end{bmatrix} = \begin{bmatrix} \underline{v}_1^T \\ \vdots \\ \underline{v}_n^T \end{bmatrix} \begin{bmatrix} A\underline{v}_1 & A\underline{v}_2 & \ldots & A\underline{v}_n \end{bmatrix}$$
  $$= \begin{bmatrix} \underline{v}_1^T \\ \vdots \\ \underline{v}_n^T \end{bmatrix} \begin{bmatrix} \lambda_1\underline{v}_1 & A\underline{v}_2 & \ldots & A\underline{v}_n \end{bmatrix}$$
  $$= \begin{bmatrix} \lambda_1 & | & * \\ - & - & - \\ \underline{0} & | & A_1 \end{bmatrix} = B.$$
  Since $\underline{v}_1^T \underline{v}_1 = \underline{v}_1 \cdot \underline{v}_1 = 1$ and $\underline{v}_i^T \underline{v}_1 = \underline{v}_i \cdot \underline{v}_1 = 0$, $i \neq 1$.
  Moreover $B^T = (Q_1^T AQ_1)^T = Q_1^T A^T (Q_1^T)^T = Q_1^T AQ_1 = B$. So B is symmetric, and $B = \begin{bmatrix} \lambda_1 & | & \underline{0} \\ - & - & - \\ \underline{0} & | & A_1 \end{bmatrix}$ where A_1 is symmetric, and it is a $k \times k$ matrix and by induction hypothesis there is an orthogonal matrix P_2 such that $P_2^T A_1 P_2 = D_1$ is a real diagonal matrix.
  Let Q_2 be the block matrix $Q_2 = \begin{bmatrix} 1 & | & \underline{0} \\ - & - & - \\ \underline{0} & | & P_2 \end{bmatrix}$.
  Q_2 is an orthogonal $(k + 1) \times (k + 1)$ matrix and so is $Q = Q_1 Q_2$.

$Q^T AQ = (Q_1Q_2)^T A(Q_1Q_2) = Q_2^T Q_1^T AQ_1Q_2 = Q_2^T (Q_1^T AQ_1)Q_2 = Q_2^T BQ_2$.

$$= \left[\begin{array}{c|c} 1 & \underline{0} \\ \hline \underline{0} & P_2^T \end{array}\right] \left[\begin{array}{c|c} \lambda_1 & \underline{0} \\ \hline \underline{0} & A_1 \end{array}\right] \left[\begin{array}{c|c} 1 & \underline{0} \\ \hline \underline{0} & P_2 \end{array}\right]$$

$$= \begin{bmatrix} \lambda_1 & | & \underline{0} \\ - & -- & ----- \\ \underline{0} & | & P_2^T A_1 P_2 \end{bmatrix} = \begin{bmatrix} \lambda_1 & | & \underline{0} \\ - & -- & - \\ \underline{0} & | & D_1 \end{bmatrix}$$

which is diagonal. This completes the induction step.

So for all $n \geq 1$, $n \in \mathbb{N}$,

$n \times n$ real symmetric matrices are orthogonally diagonalisable. $\square$

Example Orthogonally diagonalise the matrix,

$$A = \begin{bmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \\ 2 & -2 & 3 \end{bmatrix}.$$

$A = A^T$, so we need to determine Q such that $Q^T AQ = D$, a diagonal matrix.

We first find the eigenvalues λ that satisfy $\det(a - \lambda \mathbb{I}) = 0$ where,

$$\det(A - \lambda \mathbb{I}) = \det \begin{pmatrix} 3-\lambda & 2 & 2 \\ 2 & 3-\lambda & -2 \\ 2 & -2 & 3-\lambda \end{pmatrix}$$
$$= (3-\lambda) \begin{vmatrix} 3-\lambda & -2 \\ -2 & 3-\lambda \end{vmatrix} - 2 \begin{vmatrix} 2 & -2 \\ 2 & 3-\lambda \end{vmatrix} + 2 \begin{vmatrix} 2 & 3-\lambda \\ 2 & -2 \end{vmatrix}$$
$$= -(\lambda - 5)^2(\lambda + 1)$$

It follows that $\lambda_1 = -1$ and $\lambda_2 = 5$ are the eigenvalues.

Eigenspace $E_{-1} = \left\{ t \begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix} : t \in \mathbb{R} \right\}$ and so $\left\{ \begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix} \right\}$ is basis for E_{-1}

Eigenspace $E_5 = \left\{ t \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + s \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} : t, s \in \mathbb{R} \right\}$ and so $\left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \right\}$ is basis for E_5

We apply Gram-Schmidt to find an orthonormal basis. This gives:

$\begin{pmatrix} \frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{3}} \end{pmatrix}, \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{pmatrix}, \begin{pmatrix} \frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{6}} \\ \frac{2}{\sqrt{6}} \end{pmatrix}$

We normalized the first vector and then as vectors from different eigenspaces are already orthogonal, we don't apply Gram-Schmidt to all three of them but just to the second eigenspace ones.

It follows that $Q = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{3}} & 0 & \frac{2}{\sqrt{6}} \end{pmatrix}$ is an orthogonal matrix, such that

$Q^T A Q = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 5 \end{pmatrix}$

Quadratic forms

Definition A quadratic form in n variables is a function $f : \mathbb{R}^n \to \mathbb{R}$ of the form

$f(\mathbf{x}) = \mathbf{x}^T A \mathbf{x} = a_{11}x_1^2 + a_{22}x_2^2 + \cdots + a_{nn}x_n^2 + \sum_{i<j} 2a_{ij}x_i x_j,$

where A is symmetric $n \times n$ matrix and $\mathbf{x} \in \mathbb{R}^n$. A is the matrix associated with f. A be an $n \times n$ matrix.

Remarks

- Every quadratic form in 1 variable has the form $q(x) = ax^2$, $a \in \mathbb{R}$.
- An example of a quadratic form in 2 variables is $q(x_1, x_2) = x_1^2 + x_2^2$.
- A is not unique, two different A's can lead to the same function f, so $A \to f$ is not injective.
- Every A defines a quadratic form f, so $A \to f$ is surjective.

Example What is the quadratic form assiciated with

$$A = \begin{bmatrix} 17 & 4 \\ 4 & -9 \end{bmatrix}.$$

Let $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$, $x_i \in \mathbb{R}$. Then

$f(\mathbf{x}) = \mathbf{x}^T A\mathbf{x} = [x_1\ x_2] \begin{bmatrix} 17 & 4 \\ 4 & -9 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = [x_1\ x_2] \begin{bmatrix} 17x_1 + 4x_2 \\ 4x_1 - 9x_2 \end{bmatrix}$
$= 17x_1^2 + 4x_2x_1 + 4x_1x_2 - 9x_2^2 = 17x_1^2 + 8x_1x_2 - 9x_2^2$

Example Find the matrix associated with the quadratic form

$f(x_1, x_2, x_3) = 7x_1^2 + 5x_1x_2 - 6x_2^2 - 9x_2x_3 + 14x_3^2.$

$a_{11} = 7$, $a_{22} = -6$, $a_{33} = 14$.

The coefficients of the squared terms x_1^2 appear on tje diagonal of A.

Also $a_{12} = a_{21} = \frac{5}{2}$, $a_{23} = a_{32} = -\frac{9}{2}$ and $a_{13} = a_{31} = 0$. The coefficients of the cross product terms x_ix_j are a_{ij} and a_{ji}. Hence,

$A = \begin{bmatrix} 7 & \frac{5}{2} & 0 \\ \frac{5}{2} & -6 & -\frac{9}{2} \\ 0 & -\frac{9}{2} & 14 \end{bmatrix}.$

Definition Two quadratic forms q_1 and q_2 in n variables are *equaivalent*, if there exists an $n \times n$ matrix P such that $q_1(P\mathbf{x}) = q_2(\mathbf{x})$, for all $\mathbf{x} \in \mathbb{R}^n$.

Theorem 5.21 Principal Axes Theorem Every quadratic form can be diagonalised. If A is an $n \times n$ symmetric matrix such that there exists a quadratic form $q(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$, $\mathbf{x} \in \mathbb{R}^n$, and if Q is an orthogonal matrix such that $Q^T AQ = D$ is diagonal matrix, then the change of variables $\mathbf{x} = Q\mathbf{y}$ transforms the quadratic form q into $\mathbf{y}^T D\mathbf{y}$, which has no cross product terms. If the eigenvalues of A are $\lambda_1, \ldots, \lambda_n$ and $\mathbf{y} = [y_1, \ldots, y_n]^T$ then

$$\mathbf{x}^T A\mathbf{x} = \mathbf{y}^T D\mathbf{y} = \lambda_1 y_1^2 | \cdots | y_n^2$$

$Proof$: Let A be an $n \times n$ symmetric matrix describing $q(\mathbf{x}) = \mathbf{x}^T A\mathbf{x},\ \mathbf{x} \in \mathbb{R}^n$.

So there exists an orthogonal $n \times n$ matrix Q such that $Q^T A Q = D$, where D is diagonal.

Let $\mathbf{x} = Q\mathbf{y}$ or $\mathbf{y} = Q^{-1}\mathbf{x} = Q^T\mathbf{x}$.

$\mathbf{x}^T A\mathbf{x} = (Q\mathbf{y})^T A(Q\mathbf{y}) = \mathbf{y}^T Q^T AQ\mathbf{y} = \mathbf{y}^T D\mathbf{y}$

which is a quadratic form with no cross term, since D is diagonal. $\square$

$Note$: We could reorder the columns of Q so that $Q^T AQ = \mathrm{diag}(\lambda_1, \ldots, \lambda_p, \lambda_{p+1}, \ldots, \lambda_{p+r}, 0, \ldots 0)$ where $\lambda_1, \ldots, \lambda_p > 0$ and $\lambda_{p+1}, \ldots, \lambda_{p+r} < 0$.

Definition Let Q be a non-zero quadratic form in n variables that transforms such that $q(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$, $\mathbf{x} \in \mathbb{R}^n$, A is a symmetric $n \times n$ matrix with p positive eigenvalues and r negative eigenvalues. Then $p + r$ is the *rank* of q and $p - r$ is the *signature* of q.

Example Find a change of variables that transforms the quadratic form

$$f(x, y, z) = \frac{1}{121}(183x^2 + 266y^2 + 35z^2 + 12xy + 408xz + 180yz)$$

into one with no cross-product terms.

The matrix associated to f is

$A = \frac{1}{121}\begin{bmatrix} 183 & 6 & 204 \\ 6 & 266 & 90 \\ 204 & 90 & 35 \end{bmatrix}$ which has eigenvalues $\lambda_1 = 3$, $\lambda_2 = 2$, $\lambda_3 = -1$.

The corrsponding orthonormal eigenvectors are

$\mathbf{q}_1 = \frac{1}{11}\begin{pmatrix} 7 \\ 6 \\ 6 \end{pmatrix}$, $\mathbf{q}_2 = \frac{1}{11}\begin{pmatrix} 6 \\ -9 \\ 2 \end{pmatrix}$, $\mathbf{q}_3 = \frac{1}{11}\begin{pmatrix} 6 \\ 2 \\ -9 \end{pmatrix}$.

So $Q = \begin{bmatrix} \frac{7}{11} & \frac{6}{11} & \frac{6}{11} \\ \frac{6}{11} & \frac{-9}{11} & \frac{2}{11} \\ \frac{6}{11} & \frac{2}{11} & \frac{-9}{11} \end{bmatrix}$ and $D = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{bmatrix}$ then $Q^T AQ = D$. The change of variables $\mathbf{x} = Q\mathbf{y}$ where $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$, $\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$ converts f into

$$f(\mathbf{y}) = [y_1, y_2, y_3] \begin{bmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = 2y_1^2 + 2y_2^2 - y_3^2.$$

$\mathrm{rank}(f) = 3$ and the $\mathrm{signature}(f) = 1$.

Complex matrices

Quadratic forms - continued

Definition A quadratic form $f(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$ in n-variables is classified as one of the following:

a) *positive definite* if $f(\mathbf{x}) > 0$ for all $\mathbf{x} \neq 0$

b) *positive semidefinite* if $f(\mathbf{x}) \geq 0$ for all $\mathbf{x}$

c) *negative definite* if $f(\mathbf{x}) < 0$ for all $\mathbf{x} \neq 0$

d) *negative semidefinite* if $f(\mathbf{x}) \leq 0$ for all $\mathbf{x}$

e) *indefinite* if $f(\mathbf{x})$ takes on both positive and negative values.

Recall:

Rank: # of positive eigenvalues of A + # of negative eigenvalues of A

Signature: # of positive eigenvalues of A - # of negative eigenvalues of A

A symmetric matrix A is called *positive definite*, *positive semidefinite*, *negative definite*, *negative semidefinite*, *indefinite* if the associated quadratic form $f(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$ has the corresponding property.

Theorem 5.22 Let A be an $n \times n$ symmetric matrix. The quadratic form $f(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$ is
a. positive definite if and only if all of the eigenvalues of A are positive. (signature is n)

b. positive semidefinite if and only if all of the eigenvalues of A are nonnegative. (signature = rank)

c. negative definite if and only if all of the eigenvalues of A are negative. (signature is $-n$)

d. negative semidefinite if and only if all of the eigenvalues of A are non positive. (signature = - rank)

e. indefinite if and only if A has both positive and negative eigenvalues.
(-rank < signature < rank)

Example Classify $f(x_1, x_2, x_3) = 3x_1^2 + 2x_1x_2 + 2x_1x_3 + 4x_2x_3$ as positive definite, negative definite, indefinite, or none of these.

The matrix associated to f is:

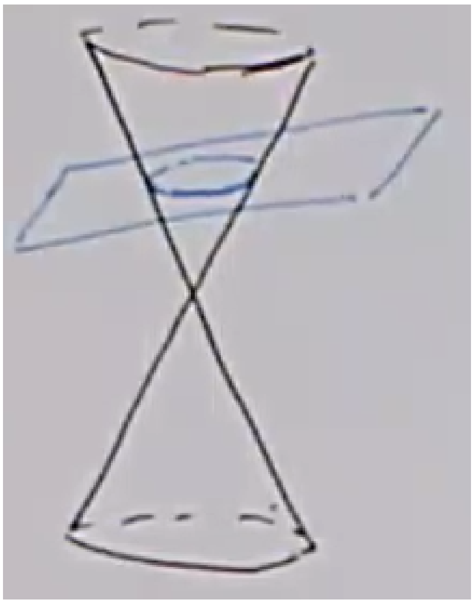
$$A = \begin{bmatrix} 3 & 1 & 1 \\ 1 & 0 & 2 \\ 1 & 2 & 0 \end{bmatrix}$$

The characteristic polynomial is $-\lambda^3 + 3\lambda^2 + 6\lambda - 8$

and the eigenvalues of A are $\lambda = 1, 4, -2$. Therefore, f is indefinite.

Quadratic and conic sections

A conic section or conic is a curve that results from cutting a double-napped cone with a plane.



Ellipse

Parabola

Hyperbola

Quadratic forms in 2-variables naturally give the equation of a conic $\mathbf{x}^T A\mathbf{x} = k$, where $k \neq 0$ is in $\mathbb{R}$. Applying the *principal axes theorem* the conic becomes

$$\lambda_1(x')^2 + \lambda_2(y')^2 = k, \text{ where } \mathbf{x}' = \begin{bmatrix} x' \\ y' \end{bmatrix},$$

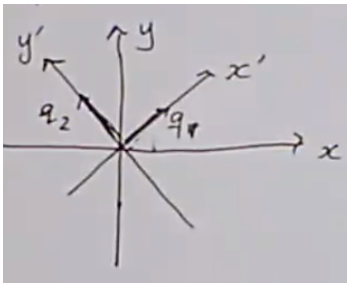
λ_1, λ_2 are the eigenvalues of A, $\mathbf{x} = Q\mathbf{x}'$ and $Q^T AQ = D \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$.

Explanation: $\mathbf{x}^T A\mathbf{x} = k$, $k \in \mathbb{R}$, $\mathbf{x} \in \mathbb{R}^2$.

Then there is orthogonal matrix Q such that $Q^T AQ = D$, a diagonal matrix. Then

$\mathbf{x}^T A\mathbf{x} = \mathbf{x}^T Q(Q^T AQ)Q^T\mathbf{x} = (\mathbf{x}^T Q)D(Q^T\mathbf{x}) = (Q^T\mathbf{x})^T D(Q^T\mathbf{x})$
$\text{Now let } \mathbf{x}' = Q^T\mathbf{x}, \text{ then}$
$(Q^T\mathbf{x})^T D(Q^T\mathbf{x}) = (\mathbf{x}')^T D\mathbf{x}' = \lambda_1(x')^2 + \lambda_2(y')^2 \text{ where } \mathbf{x}' = \begin{bmatrix} x' \\ y' \end{bmatrix}$

Notice $\lambda_1(x')^2 + \lambda_2(y')^2 = k$ is an ellipse (λ_1, λ_2 positive) or hyperbola (λ_1, λ_2 have opposite signs) in the $x'y'$-coordinate system. We relate this to the original $x'y'$-coordinate system by determining where the standard basis $\mathbf{e}_1' = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\mathbf{e}_2' = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ is sent by the coordinate change $\mathbf{x} = Q\mathbf{x}'$. We have



$$Q\mathbf{e}_1' = \mathbf{q}_1 \text{ and } Q\mathbf{e}_2' = \mathbf{q}_2, \text{ where } Q = [\mathbf{q}_1\ \mathbf{q}_2]$$

$\mathbf{q}_1$, $\mathbf{q}_2$ are the eigenvectors of A and are hence the principal axis of the ellipse or hyperbola. We can scale this idea up to higher dimensions and ellipsoid and hyperboloid. So

- $\mathbf{x}^T A\mathbf{x} = k$ represents an ellipse if $\lambda_1 > 0$ and $\lambda_2 > 0$

- $\mathbf{x}^T A\mathbf{x} = k$ has no graph if $\lambda_1 < 0$ and $\lambda_2 < 0$
- $\mathbf{x}^T A\mathbf{x} = k$ represents a hyperbola if λ_1 and λ_2 have opposite signs.

Example Identify and graph the conic whose equation is

$$5x^2 - 4xy + 8y^2 = 36$$

as positive definite, negative definite, indefinite, or none of these.

The LHS of the equation is a quadratic form so we can write it as:

$\mathbf{x}^T A\mathbf{x} = 36$

$A = \begin{bmatrix} 5 & -2 \\ -2 & 8 \end{bmatrix}$

A has eigenvalues : $\lambda = 4, 9$

The orthonormal bases for the eigenspaces are:

$\lambda = 4$: $\begin{bmatrix} \frac{2}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} \end{bmatrix}$, $\lambda = 9$: $\begin{bmatrix} -\frac{1}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} \end{bmatrix}$

The matrix Q that orthogonally diagonalises A is:

$Q = \begin{bmatrix} \frac{2}{\sqrt{5}} & -\frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{bmatrix}$ where $Q^T AQ = D = \begin{bmatrix} 4 & 0 \\ 0 & 9 \end{bmatrix}$.

The change of variables $\mathbf{x} = Q\mathbf{x}'$ converts $\mathbf{x}^T A\mathbf{x} = 36$ into $(\mathbf{x}')^T D\mathbf{x}' = 36$ by means of a rotation.

$Q = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$ which implies $\cos\theta = \frac{2}{\sqrt{5}}$, $\sin\theta = \frac{1}{\sqrt{5}}$, $\tan\theta = \frac{\sin\theta}{\cos\theta} = \frac{1}{2}$,

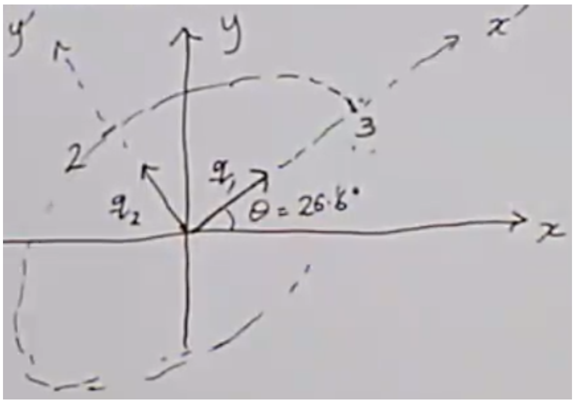
so $\theta \approx 26.6^\circ$

If $\mathbf{x}' = \begin{bmatrix} x' \\ y' \end{bmatrix}$ then $(\mathbf{x}')^T D\mathbf{x} = 36$ is $4(x')^2 + 9(y')^2 = 36$ or $\frac{(x')^2}{9} + \frac{(y')^2}{4} = 1$ which is an ellipse in $x'y'$-coordinates.

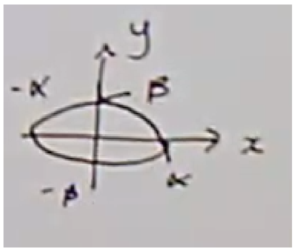
To graph the ellipse, we look at the vector $\mathbf{e}_1' = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $\mathbf{e}_2' = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ in the $x'y'$ coordinates this corresponds to

$$Qe_1' = \begin{bmatrix} \frac{2}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} \end{bmatrix} = \mathbf{q}_1,\ Qe_2' = \begin{bmatrix} -\frac{1}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} \end{bmatrix} = \mathbf{q}_2$$

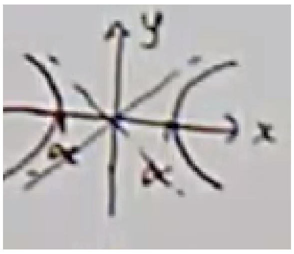
in the xy-coordinates these are the eigenvectors of A.



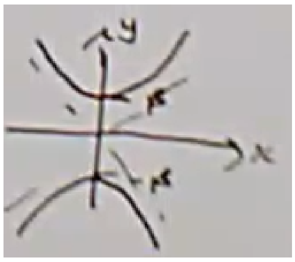
Recall: $\frac{x^2}{\alpha^2} + \frac{y^2}{\beta^2} = 1$ is an ellipse



$\frac{x^2}{\alpha^2} - \frac{y^2}{\beta^2} = 1$ is a hyperbola of form



$-\frac{x^2}{\alpha^2} + \frac{y^2}{\beta^2} = 1$ is a hyperbola of form



Unitary and Hermitian matrices

Definition Let A be a complex $n \times n$ matrix with $A = [a_{ij}]$ we define *complex conjugate* by $\overline{A} = [\overline{a}_{ij}]$. The *conjugate transpose* of A is the matrix $A^* = \overline{A}^T$.

Example Let $A = \begin{bmatrix} i & 3i+5 & 6 \\ 0 & 1 & 2i \\ 1-i & 3 & 0 \end{bmatrix}$ find the conjugate transpose A^*.

$$\overline{A} = \begin{bmatrix} -i & -3i+5 & 6 \\ 0 & 1 & -2i \\ 1+i & 3 & 0 \end{bmatrix}$$

$A^* = \overline{A}^T = \begin{bmatrix} -i & 0 & 1+i \\ -3i+5 & 1 & 3 \\ 6 & -2i & 0 \end{bmatrix}$

Theorem If $A, B \in M_{n\times n}(\mathbb{C})$ be matrices and $c \in \mathbb{C}$. Then

a) $(A^*)^* = A$

b) $(A+B)^* = A^* + B^*$

c) $(CA)^* = \overline{c}A^*$

d) $(AB)^* = B^*A^*$

$Proof\ of\ d)$ write $A = [a_{ij}]$ and $B = [b_{ij}]$. Let c_{ij} be the $(ij)^{\text{th}}$ entry of the product AB i.e. $c_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj}$

Hence the $(ij)^{\text{th}}$ entry of $(AB)^*$ is

$\overline{c}_{ji} = \sum_{k=1}^{n} \overline{a_{jk}b_{ki}}$

On the other hand, let d_{ij} be the $(ij)^{\text{th}}$ entry of the product B^*A^*. Then

$d_{ij} = \sum_{k=1}^{n} \overline{b}_{ki}\overline{a}_{jk}$.

Hence we have $d_{ij} = \overline{c}_{ji}$.

Definition Let $\mathbf{u} = \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}$ be vectors in $\mathbb{C}^n$, then the *complex dot product* of $\mathbf{u}$ and $\mathbf{v}$ is defined by

$$\mathbf{u}\cdot\mathbf{v} = \overline{u}_1 v_1 + \cdots + \overline{u}_n v_n.$$

Theorem For every $\mathbf{u}, \mathbf{v} \in \mathbb{C}^n$ we have $\mathbf{u}\cdot\mathbf{v} = \mathbf{u}^*\mathbf{v}$

$Proof: \mathbf{u}\cdot\mathbf{v} = \overline{u}_1 v_1 + \cdots + \overline{u}_n v_n = \overline{\mathbf{u}}^T\mathbf{v} = \mathbf{u}^*\mathbf{v}. \square$

Definition A square complex matrix A is called *Hermitian (self-adjoint)* if $A^* = A$.

Example Show that $A = \begin{bmatrix} 1 & 1+i & 3 \\ 1-i & 5 & e^{i\frac{\pi}{2}} \\ 3 & e^{-i\frac{\pi}{2}} & -1 \end{bmatrix}$ is Hermitian.

$$A^* = \overline{A}^T = \begin{bmatrix} 1 & 1-i & 3 \\ 1+i & 5 & e^{-i\frac{\pi}{2}} \\ 3 & e^{i\frac{\pi}{2}} & -1 \end{bmatrix}^T = \begin{bmatrix} 1 & 1+i & 3 \\ 1-i & 5 & e^{i\frac{\pi}{2}} \\ 3 & e^{-i\frac{\pi}{2}} & -1 \end{bmatrix} = A.$$

Diagonal entries have to be real. The off-diagonal entries have to form complex conjugate pairs.

Many results for Hermitian matrices are analogous to those for symmetric matrices. Let $A \in M_{n\times n}(\mathbb{C})$ be a Hermitian matrix.

a) Every eigenvalue of a Hermitian matrix A is a real number.

b) The eigenvectors corresponding to distinct eigenvalues of A are orthogonal.

Definition A square complex matrix U is called *unitary* if $U^{-1} = U^*$.

Many results for unitary matrices are analogous to those for orthogonal matrices. Let $U \in M_{n\times n}(\mathbb{C})$. The following are equivalent:

a) U is unitary matrix

b) The columns of U form an orthonormal basis for $\mathbb{C}^n$ with respect to the complex dot product

c) $||U\mathbf{x}|| = ||\mathbf{x}||$ for every $\mathbf{x} \in \mathbb{C}^n$

d) $U\mathbf{x} \cdot U\mathbf{y} = \mathbf{x} \cdot \mathbf{y}$ for every $\mathbf{x}, \mathbf{y} \in \mathbb{C}^n$

Definition A square complex matrix A is called *unitrary diagonalisable* if there exists a unitary matrix U and diagonal matrix D such that $U^*AU = D$.

Example Unitarily diagonalise the matrix

$$A = \begin{bmatrix} 2 & 1+i \\ 1-i & 3 \end{bmatrix}.$$

The characteristic polynomial of A is $\det(A - \lambda\mathbb{I}) = \begin{vmatrix} 2-\lambda & 1+i \\ 1-i & 3-\lambda \end{vmatrix} = (2 - \lambda)(3-\lambda) - (1+i)(1-i)$
$= 6 - 5\lambda + \lambda^2 - 2 = \lambda^2 - 5\lambda + 4 = (\lambda - 4)(\lambda - 1)$

So the eigenvalues of A are $\lambda = 1, 4$.

Note A is Hermitian.

We next calculate bases for the corresponding eigenspaces. We solve $(A - \lambda\mathbb{I})\mathbf{v} = \mathbf{0}$ where $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$.

$\lambda_1 = 1$: $(A - \mathbb{I})\mathbf{v} = \mathbf{0}$ has corresponding augemented matrix

$$\begin{bmatrix} 1 & 1+i & | & 0 \\ 1-i & 2 & | & 0 \end{bmatrix} \xrightarrow{R_2 \to R_2 - (1-i)R_1} \begin{bmatrix} 1 & 1+i & | & 0 \\ 0 & 2-(1-i)(1+i) & | & 0 \end{bmatrix} = \begin{bmatrix} 1 & 1+i & | & 0 \\ 0 & 0 & | & 0 \end{bmatrix}$$

so $v_1 + (1+i)v_2 = 0$. Hence

$E_1 = \left\{ t \begin{bmatrix} -1-i \\ 1 \end{bmatrix} : t \in \mathbb{R} \right\}$, so the bases is $\mathbf{u}_1 = \begin{bmatrix} -1-i \\ 1 \end{bmatrix}$

For the other eigenvalue the steps are similar.

E_4 has bases $\mathbf{u}_2 = \begin{bmatrix} \frac{1}{2}(1+i) \\ 1 \end{bmatrix}$.

$\mathbf{u}_1$ and $\mathbf{u}_2$ are orthogonal.

We normalise the basis vectors

$\mathbf{q}_1 = \frac{\mathbf{u}_1}{||\mathbf{u}_1||} = \begin{bmatrix} \frac{-1-i}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{bmatrix}$, $\mathbf{q}_2 = \frac{\mathbf{u}_2}{||\mathbf{u}_2||} = \begin{bmatrix} \frac{(1+i)}{\sqrt{6}} \\ \frac{2}{\sqrt{6}} \end{bmatrix}$.

Thus A is unitarily diagonalised by the matrix

$Q = [\mathbf{q}_1 \ \mathbf{q}_2] = \begin{bmatrix} \frac{-1-i}{\sqrt{3}} & \frac{(1+i)}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{2}{\sqrt{6}} \end{bmatrix}$

and $Q^* A Q = D = \begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix}$.

Theorem Every Hermitian matrix A is *unitarily diagonalisable*.

Note: Note the above is the complex Spectral Theorem, analogous to 5.20, but **not** all Unitarily diagonalisable matrices are Hermitian.

For an example see $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$, which is **not** Hermitian, but is unitarily diagonalisable.

Theorem A square complex matrix A is *unitarily diagonalisable* if and only if

$$A^* A = AA^*.$$

Definition A square complex matrix A is called *normal* if $A^* A = AA^*$.

Theorem Every Hermitian matrix, every unitary matrix, and every *skew Hermitian matrix* ($A^* = -A$) is normal.

(Note that in the real case, this result refers to symmetric, orthogonal and skew-symmetric matrices respectively.)

$Proof$:

(i) Let A be a square Hermitian matrix so $A^* = A$.

$A * A = AA = AA^*$.

(ii) Let A be a unitary matrix, so $A^{-1} = A^*$

So $A^{-1}A = \mathbb{I} = AA^{-1}$, but $A^{-1} = A^*$ so $A^* A = \mathbb{I} = AA^*$

(iii) Let A be Hermitian symmetric, so $A^* = -A$.

$AA^* = -AA = A^* A$. $\square$