MAST10007 Linear Algebra

January, 2026

Arun Ram

These slides have been made by Arun Ram, for teaching of the summer session of MAST10007 Linear Algebra at University of Melbourne in 2026. The template is from the University of Melbourne School of Mathematics and Statistics slide deck produced by members of the School. Some examples from that slide deck have been retained, All the solutions and writing has been reworked, as well as the exposition and ordering of the material.

I found slides to be an unusual medium. Each slide has very little space (compared to a page of ordinary LaTeX) and there is some necessity to assume that the reader has little specific recall of other slides. I found it important to repeat slides often and to continually insert cues and portions of material that had appeared in previous slides.

Slides form some unusual medium between a book and a lecture: there is an impetus for completeness and linearity that one often strives for in a book format, but it is not appropriate for the storytelling framework of a lecture situation. At the same time, one cannot revert to a story telling framework, as the slide need to hold together in broader arcs and structure because they will certainly be being used as a book type resource by students.

As a result there were many places that choices were made that are absolutely not appropriate for engaging lectures and other places that choices were made that are absolutely not appropriate for a coherent book type resource. Slide decks sit in a strange medium between lecturing and printed resource materials. Having done this exercise I am even more convinced that reading from slides is not an optimally healthy or effective way to deliver quality mathematics lectures.

The Hilbert space $\mathbb{R}^n$

(a) $\mathbb{R}^n$, basis, inner product, length, distance, angles, projection

(b) Equations of lines and planes in $\mathbb{R}^3$

(c) Cross products (are only available in $\mathbb{R}^3$

Matrices

(a) addition, scalar multiplication, basis, multiplication, inverses

(b) Factoring for matrices

(c) The rank theorem

Linear systems and kernels

(a) Finding inverses

(b) Solutions of linear systems

(c) Kernels and images

Eigenvalues, eigenvectors and diagonalization

(a) Bases of kernels and images
(b) Eigenvectors, eigenvalues and diagonlization
(c) Symmetric, Hermitian and orthogonal matrices

Vector spaces and linear transformations

(a) Definitions, examples and bases
(b) Linear transformations
(c) span, linear independence and bases

Bases and matrices

(a) The minimax basis theorem
(b) Kernels and images of linear transformations
(c) With respect to a basis

Inner product spaces

(a) Definitions, examples and Gram matrices
(b) Gram-Schmidt orthogonalization
(c) Projections using an orthonormal basis

Applications

(a) Graphs and networks
(b) Application of diagonalization to dynamics
(c) Data correlation and line of best fit

Additional topics

(a) Traces and determinants
(b) Singular value decomposition
(c) Learning to do proofs

Lecture 1: The Hilbert space $\mathbb{R}^n$

Definition (The vector space $\mathbb{R}^n$)

Let $n \in \mathbb{Z}_{>0}$. The *$\mathbb{R}$-vector space $\mathbb{R}^n$* is

$$\mathbb{R}^n = M_{n\times 1}(\mathbb{R}) = \{|x_1, \ldots, x_n\rangle \mid x_i \in \mathbb{R}\} \quad \text{where} \quad |x_1, \ldots, x_n\rangle = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

The *addition and scalar multiplication* are given by

$$|x_1, x_2, \ldots, x_n\rangle + |y_1, y_2, \ldots, y_n\rangle = |x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n\rangle$$

and

$$c|x_1, x_2, \ldots, x_n\rangle = |cx_1, cx_2, \ldots, cx_n\rangle \quad \text{for } c \in \mathbb{R}.$$

The notation $|x_1, \ldots, x_n\rangle$ is *Dirac's ket notation* for the column vector with entries $x_1, \ldots, x_n$.

Definition (The favorite basis of $\mathbb{R}^n$)

Let $e_1, \ldots, e_n$ be the length n column vectors given by

$$e_i \text{ has } 1 \text{ in the } i\text{th spot and } 0 \text{ elsewhere.}$$

Every vector in $\mathbb{R}^n$ is a (unique) linear combination of $e_1, \ldots, e_n$.

('linear' means using scalar multiplication and addition).

For example, if $n = 4$ then

$$e_1 = \begin{pmatrix}1\\0\\0\\0\end{pmatrix}, \quad e_2 = \begin{pmatrix}0\\1\\0\\0\end{pmatrix}, \quad e_3 = \begin{pmatrix}0\\0\\1\\0\end{pmatrix}, \quad e_4 = \begin{pmatrix}0\\0\\0\\1\end{pmatrix}$$

$$\text{and} \quad \begin{pmatrix}3\\5\\-2\\0\end{pmatrix} = 3\begin{pmatrix}1\\0\\0\\0\end{pmatrix} + 5\begin{pmatrix}0\\1\\0\\0\end{pmatrix} + (-2)\begin{pmatrix}0\\0\\1\\0\end{pmatrix} + 0\begin{pmatrix}0\\0\\0\\1\end{pmatrix}$$

$$= 3e_1 + 5e_2 + (-2)e_3 + 0e_4.$$

Definition (Inner product, length function and distance function)

The *standard inner product on $\mathbb{R}^n$* is $\langle \ | \ \rangle\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ given by

$$\langle x_1, \ldots, x_n | y_1, \ldots, y_n \rangle = \begin{pmatrix} x_1 & x_2 & \cdots & x_n \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = x_1 y_1 + \cdots + x_n y_n.$$

The *length function* is $\| \ \|\colon \mathbb{R}^n \to \mathbb{R}_{\geqslant 0}$ given by

$$\| \, |x_1, \ldots, x_n\rangle \| = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}.$$

The *distance function* is $d\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}_{\geqslant 0}$ given by

$$d(|x_1, \ldots, x_n\rangle, |y_1, \ldots, y_n\rangle) = \| \, |x_1, \ldots, x_n\rangle - |y_1, \ldots, y_n\rangle \|.$$

Theorem (Cauchy-Schwarz and the triangle inequality)

Let $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$. *Then*

$$|\langle \mathbf{u}, \mathbf{v} \rangle| \leqslant \|\mathbf{u}\| \cdot \|\mathbf{v}\| \quad \textit{and} \quad \|\mathbf{u} + \mathbf{v}\| \leqslant \|\mathbf{u}\| + \|\mathbf{v}\|$$

If

$$\mathbf{x} = |x_1, x_2, \ldots, x_n\rangle \quad \text{and} \quad \mathbf{y} = |y_1, y_2, \ldots, y_n\rangle$$

then

$$\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^T \mathbf{y} = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n$$

and

$$\|\mathbf{x}\| = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle} \quad \text{and} \quad \|\mathbf{x}\|^2 = \langle \mathbf{x}, \mathbf{x} \rangle \qquad \text{and}$$

$$\begin{aligned} d(\mathbf{x}, \mathbf{y}) = \|\mathbf{y} - \mathbf{x}\| &= \sqrt{\langle \mathbf{y} - \mathbf{x}, \mathbf{y} - \mathbf{x} \rangle} \\ &= \sqrt{(y_1 - x_1)^2 + \cdots + (y_n - x_n)^2}. \end{aligned}$$

Easy to establish properties that are used VERY often.

Let $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^n$ and let $c \in \mathbb{R}$.

$$\begin{aligned}\langle \mathbf{y}, \mathbf{x} \rangle &= y_1 x_1 + y_2 x_2 + \cdots + y_n x_n \\ &= x_1 y_1 + x_2 y_2 + \cdots x_n y_n = \langle \mathbf{x}, \mathbf{y} \rangle,\end{aligned}$$

$$\langle \mathbf{x}, \mathbf{y} + \mathbf{z} \rangle = \mathbf{x}^T (\mathbf{y} + \mathbf{z}) = \mathbf{x}^T \mathbf{y} + \mathbf{x}^T \mathbf{z} = \langle \mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{x}, \mathbf{z} \rangle,$$

$$\langle \mathbf{x} + \mathbf{y}, \mathbf{z} \rangle = \langle \mathbf{z}, \mathbf{x} + \mathbf{y} \rangle = \langle \mathbf{z}, \mathbf{x} \rangle + \langle \mathbf{z}, \mathbf{y} \rangle = \langle \mathbf{x}, \mathbf{z} \rangle + \langle \mathbf{y}, \mathbf{z} \rangle,$$

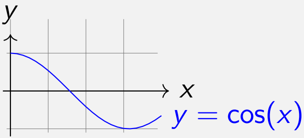
$$\langle \mathbf{x}, c\mathbf{y} \rangle = \mathbf{x}^T c\mathbf{y} = c\mathbf{x}^T \mathbf{y} = c\langle \mathbf{x}, \mathbf{y} \rangle, \langle c\mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{y}, c\mathbf{x} \rangle = c\langle \mathbf{y}, \mathbf{x} \rangle = c\langle \mathbf{x}, \mathbf{y} \rangle,$$

$$\|c\mathbf{x}\| = \sqrt{\langle c\mathbf{x}, c\mathbf{x} \rangle} = \sqrt{c^2 \langle \mathbf{x}, \mathbf{x} \rangle} = \sqrt{c^2} \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle} = |c| \cdot \|\mathbf{x}\|.$$

Definition (Angle and projection)

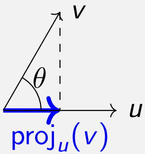
Let $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ with $\mathbf{u} \neq 0$ and $\mathbf{v} \neq 0$. The *angle between* $\mathbf{u}$ *and* $\mathbf{v}$ is $\theta(\mathbf{u}, \mathbf{v})$ given by

$$\cos(\theta(\mathbf{u}, \mathbf{v})) = \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\|\mathbf{u}\| \cdot \|\mathbf{v}\|}.$$



The *projection of* $\mathbf{v}$ *onto* $\mathbf{u}$ is

$$\operatorname{proj}_{\mathbf{u}} \mathbf{v} = \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\langle \mathbf{u}, \mathbf{u} \rangle} \mathbf{u}.$$



Let $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$. The vectors $\mathbf{u}$ and $\mathbf{v}$ are *perpendicular* if $\langle \mathbf{u}, \mathbf{v} \rangle = 0$. The vectors $\mathbf{u}$ and $\mathbf{v}$ are *parallel* if $\langle \mathbf{u}, \mathbf{v} \rangle \in \{1, -1\}$.

Example E1. If $\mathbf{u} = |1,3,1,2\rangle$ and $\mathbf{v} = |2,1,-1,3\rangle$ in $\mathbb{R}^4$ then

$$\mathbf{u} - \mathbf{v} = |1,-2,0,1\rangle$$

and the distance between the points $(1,3,1,2)$ and $(2,1,-1,3)$ is

$$\begin{aligned} d(\mathbf{u},\mathbf{v}) &= \|\ |1,-2,0,1\rangle\ \| = \sqrt{1^2 + (-2)^2 + 0^2 + 1^1} \\ &= \sqrt{1+4+0+1} = \sqrt{6}. \end{aligned}$$

Example E2. If $\mathbf{u} = |0,2,2,-1\rangle$ and $\mathbf{v} = |-1,1,1,-1\rangle$ in R^4 then

$$\begin{aligned} \langle \mathbf{u},\mathbf{v}\rangle &= \langle 0,2,2,-1|-1,1,1,-1\rangle \\ &= 0\cdot(-1) + 2\cdot 1 + 2\cdot 1 + (-1)\cdot(-1) \\ &= 0+2+1+1 = 5 \end{aligned}$$

and

$$\|\mathbf{u}\| = \sqrt{0+4+4+1} = \sqrt{9} = 3$$

and

$$\|\mathbf{v}\| = \sqrt{1+1+1+1} = \sqrt{4} = 2.$$

Since $|5| \leqslant 3 \cdot 2$ we observe that, in this case,

$$|\langle \mathbf{u}, \mathbf{v} \rangle| \leqslant \|\mathbf{u}\| \cdot \|\mathbf{v}\|.$$

Example E4. Let $\mathbf{u} = (2, -1, -2)$ and $\mathbf{v} = (2, 1, 3)$. Find vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ such that

$$\mathbf{v} = \mathbf{v}_1 + \mathbf{v}_2$$

where $\mathbf{v}_1$ is parallel to $\mathbf{u}$ and $\mathbf{v}_2$ is perpendicular to $\mathbf{u}$.

Solution: Since the projection of $\mathbf{v}$ onto $\mathbf{u}$ is parallel to $\mathbf{u}$ then let

$$\begin{aligned} \mathbf{v}_1 &= \operatorname{proj}_{\mathbf{u}} \mathbf{v} = \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\langle \mathbf{u}, \mathbf{u} \rangle} \mathbf{u} = \frac{-3}{9} \mathbf{u} \\ &= \tfrac{-1}{3} |2, -1, -2\rangle = |\tfrac{-2}{3}, \tfrac{1}{3}, \tfrac{2}{3}\rangle \end{aligned}$$

and

$$\mathbf{v}_2 = \mathbf{u} - \mathbf{v}_1 = |2, -1, -2\rangle - |\tfrac{-2}{3}, \tfrac{1}{3}, \tfrac{2}{3}\rangle = |\tfrac{8}{3}, \tfrac{2}{3}, \tfrac{7}{3}\rangle.$$

Then $\mathbf{u} = \mathbf{v}_1 + \mathbf{v}_2$ and $\mathbf{v}_1$ is parallel to $\mathbf{u}$ and $\mathbf{v}_2$ is perpendicular to $\mathbf{u}$.

Lecture 2: Equations of lines and planes in $\mathbb{R}^3$

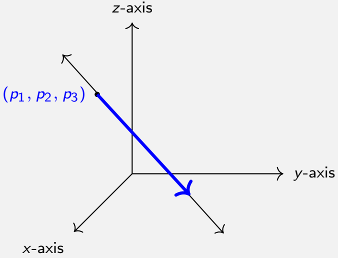
Let $\mathbf{u}, \mathbf{v} \in \mathbb{R}^3$. *The set of linear combinations of* $\mathbf{v}$ is

$$\mathbb{R}\text{-span}\{\mathbf{v}\} = \{t\mathbf{v} \mid t \in \mathbb{R}\}.$$

Definition

The *line in* $\mathbb{R}^3$ *with direction* $\mathbf{v} = |v_1, v_2, v_3\rangle$ *going through the point* $p = |p_1, p_2, p_3\rangle$ is

$$p + \mathbb{R}\mathbf{v} = \{p + t\mathbf{v} \mid t \in \mathbb{R}\}.$$



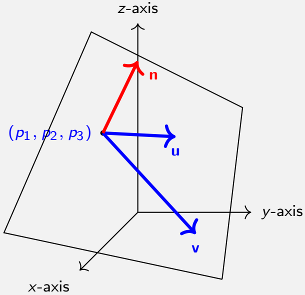
The set of linear combinations of $\{\mathbf{u}, \mathbf{v}\}$ is

$$\mathbb{R}\text{-span}\{\mathbf{u}, \mathbf{v}\} = \{s\mathbf{u} + t\mathbf{v} \mid s, t \in \mathbb{R}\},$$

Definition

The *plane in* $\mathbb{R}^3$ *spanned in directions* $\mathbf{u} = |u_1, u_2, u_3\rangle$ *and* $\mathbf{v} = |v_1, v_2, v_3\rangle$ *going through the point* $p = |p_1, p_2, p_3\rangle$ is

$$p + \mathbb{R}\mathbf{u} + \mathbb{R}\mathbf{v} = \{p + s\mathbf{u} + t\mathbf{v} \mid s, t \in \mathbb{R}\}.$$



Equations of lines in $\mathbb{R}^3$

Definition

The *line in $\mathbb{R}^3$ with direction $\mathbf{v} = |v_1, v_2, v_3\rangle$ going through the point $p = |p_1, p_2, p_3\rangle$* is

$$p + \mathbb{R}\mathbf{v} = \{p + t\mathbf{v} \mid t \in \mathbb{R}\}.$$

The points in the line are the $|x, y, z\rangle$ in $\mathbb{R}^3$ such that

$$(x, y, z) = (p_1, p_2, p_3) + t(v_1, v_2, v_3), \quad \text{with } t \in \mathbb{R}, \qquad \textit{(vector equation)}$$

or

$$\begin{aligned} x &= p_1 + tv_1, \\ y &= p_2 + tv_2, \\ z &= p_3 + tv_3, \end{aligned} \qquad \text{with } t \in \mathbb{R}, \qquad \textit{(parametric equation)}$$

Solving for t gives that the points on the line are the $|x, y, z\rangle$ in $\mathbb{R}^3$ which satisfy the equations

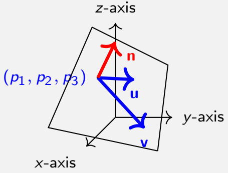
$$\frac{x - p_1}{v_1} = \frac{y - p_2}{v_2} = \frac{z - p_3}{v_3}. \qquad \textit{(Cartesian form)}$$

Equations of planes in $\mathbb{R}^3$

Definition

The *plane in $\mathbb{R}^3$ spanned in directions $\mathbf{u} = |u_1, u_2, u_3\rangle$ and $\mathbf{v} = |v_1, v_2, v_3\rangle$ going through the point $p = |p_1, p_2, p_3\rangle$* is

$$p + \mathbb{R}\mathbf{u} + \mathbb{R}\mathbf{v} = \{p + s\mathbf{u} + t\mathbf{v} \mid s, t \in \mathbb{R}\}.$$



The points in the line are the $|x, y, z\rangle$ in $\mathbb{R}^3$ such that

$$\begin{aligned} x &= p_1 + su_1 + tv_1, \\ y &= p_2 + su_2 + tv_2, \\ z &= p_3 + su_3 + tv_3, \end{aligned} \qquad \text{with } s, t \in \mathbb{R}. \qquad \textit{(parametric equation)}$$

The *vector equation* is

$$(x, y, z) = (p_1, p_2, p_3) + s(u_1, u_2, u_3) + t(v_1, v_2, v_3), \quad \text{with } s, t \in \mathbb{R}.$$

Let $\mathbf{n} = |a, b, c\rangle$ be such that n is perpendicular to both $\mathbf{u}$ and $\mathbf{v}$. In other words, $\mathbf{n}$ is a vector perpendicular to the plane. Then

$$\begin{aligned} \langle \mathbf{n} | x, y, z\rangle &= \langle n, p + s\mathbf{u} + t\mathbf{v}\rangle = \langle \mathbf{n}, p\rangle + s\langle \mathbf{n}, \mathbf{u}\rangle + t\langle \mathbf{n}, \mathbf{v}\rangle \\ &= \langle \mathbf{n}, p\rangle + s \cdot 0 + t \cdot 0 = \langle \mathbf{n}, \mathbf{r}_0\rangle, \end{aligned}$$

and since $\langle \mathbf{n} | x, y, z\rangle = \langle a, b, c | x, y, z\rangle = ax + by + cz$ then the plane is the set of $|x, y, z \in \mathbb{R}^3$ such that

$$ax + by + cz = \langle p, \mathbf{n}\rangle. \qquad \textit{(Cartesian form)}$$

Example E8. Determine the vector, parametric and Cartesian equations of the line through the points $P = (-1, 2, 3)$ and $Q = (4, -2, 5)$.

Since the direction of the line is

$$Q - P = |4, -2, 5\rangle - | - 1, 2, 3\rangle = |5, -4, 2\rangle$$

and

$$P = |-1, 2, 3\rangle \text{ is a point on the line}$$

then the line is the set of points in $\mathbb{R}^3$ given by

$$\{\, |-1, 2, 3\rangle + \ t \cdot |5, -4, 2\rangle \ | \ t \in \mathbb{R}\}.$$

Parametric equations for the line are

$$\begin{aligned} x &= -1 + 5t, \\ y &= 2 - 4t, \\ z &= 3 + 2t, \end{aligned} \qquad \text{with } t \in \mathbb{R}.$$

Solving for t, the Cartesian equation of the line is

$$\frac{x+1}{5} = \frac{y-2}{-4} = \frac{z-3}{2}.$$

Example E9. Find a vector equation of the 'friendly' line through the point $(2, 0, 1)$ that is parallel to the 'enemy' line

$$\frac{x-1}{1} = \frac{y+2}{-2} = \frac{z-6}{2}.$$

Does the point $(0, 4, -3)$ lie on the 'friendly' line?

Letting

$$t = \frac{x-1}{1} = \frac{y+2}{-2} = \frac{z-6}{2}$$

gives

$$\begin{aligned} x &= 1 + t, \\ y &= -2 - 2t, \quad \text{with } t \in \mathbb{R}, \\ z &= 6 + 2t \end{aligned}$$

and

$$\{|1, -2, 6\rangle + t|1, -2, 2\rangle \mid t \in \mathbb{R}\}$$

is the set of points in $\mathbb{R}^3$ that lie on the 'enemy' line.

The 'friendly' line we want is parallel to the 'enemy' line and goes through the point $|2,0,1\rangle$.

So the 'friendly' line consists of the points

$$\{\,|2,0,1\rangle + t\,|1,-2,2\rangle \mid t \in \mathbb{R}\}.$$

Since

$$|2,0,1\rangle + (-2)\cdot|1,-2,2\rangle = |0,4,-3\rangle$$

then $|0,4,-3\rangle$ is on the 'friendly' line.

Example E11. Find the vector equation for the plane in $\mathbb{R}^3$ containing the points $P = |1,0,2\rangle$ and $Q = |1,2,3\rangle$ and $R = |4,5,6\rangle$.

The point $|1,0,2\rangle$ is in the plane and two vectors in the plane are

$$Q - P = |0,2,1\rangle \quad \text{and} \quad R - P = |3,5,4\rangle.$$

So the points in the plane are the points $|x,y,z\rangle$ in $\mathbb{R}^3$ which satisfy

$$|x,y,z\rangle = |1,0,2\rangle + s|0,2,1\rangle + t|3,5,4\rangle \quad \text{with } s,t \in \mathbb{R}.$$

Example E12. Where does the line

$$\frac{x-1}{1} = \frac{y-2}{2} = \frac{z-3}{3}$$

intersect the plane $3x + 2y + z = 20$?

The line in parametric form is

$$\begin{aligned} x &= 1 + t, \\ y &= 2 + 2t, \qquad \text{with } t \in \mathbb{R}, \\ z &= 3 + 3t, \end{aligned}$$

and plugging into the equation of the plane gives

$$20 = 3(t+1) + 2(2t+2) + (3t+3) = 10t + 10 \text{ so that } t = 1.$$

Thus the point $|x, y, z\rangle$ with $x = 1 + 1 = 2$, $y = 2 + 2 = 4$ and $z = 3 + 3$ is on both the line and the plane.

Example E13. Find a vector form for the line of intersection of the two planes $x + 3y + 2z = 6$ and $3x + 2y + z = 11$.

The points on the intersection of the two planes are the points $|x, y, z\rangle$ that satisfy the system of equations

$$\begin{aligned} 3x + 2y - z &= 11, \\ x + 3y + 2z &= 6. \end{aligned}$$

One of the main points of this course is to learn how to use matrices as an efficient and organized mechanics for solving systems of equations of this type. For now, let's proceed ad hoc. The second equation gives

$$x = 6 - 3y - 2z, \quad \text{and plugging back into} \quad 3x + 2y - z = 11$$

gives

$$\begin{aligned} 11 &= 3(6 - 3y - 2z) + 2y - z = 18 - 9y - 6z + 2y - z \\ &= 18 - 7y - 7z. \end{aligned}$$

So $7y = 7 - 7z$ and $y = 1 - z$. So $x = 6 - 3y - z = 6 - 3(1-z) - 2z$ and

$$\begin{aligned} x &= 3 + z, \\ y &= 1 - z, \qquad \text{where } z \text{ can be any number.} \\ z &= 0 + z, \end{aligned}$$

So the line is the set of points $|x, y, z\rangle$ such that

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \\ 0 \end{pmatrix} + z \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}, \quad \text{with } z \in \mathbb{R}.$$

So the line is

$$p + \mathbb{R}\mathbf{v}, \quad \text{where } p = |3, 1, 0\rangle \text{ and } \mathbf{v} = |1, -1, 1\rangle.$$

Let $\mathbf{i}, \mathbf{j}, \mathbf{k} \in \mathbb{R}^3$ be given by

$$\mathbf{i} = \langle 1, 0, 0 \rangle, \quad \mathbf{j} = \langle 0, 1, 0 \rangle, \quad \mathbf{k} = \langle 0, 0, 1 \rangle.$$

Proposition (Standard basis of $\mathbb{R}^3$)

Let $\mathbf{v} \in \mathbb{R}^3$.

(a) *If $\mathbf{v} = \langle a_1, a_2, a_3 \rangle$ then $\mathbf{v} = a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}$.*

(b) *If $a_1, a_2, a_3 \in \mathbb{R}$ and $a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k} = 0$ then $a_1 = 0$ and $a_2 = 0$ and $a_3 = 0$.*

Every vector in $\mathbb{R}^3$ is a (unique) linear combination of $\mathbf{i}, \mathbf{j}$ and $\mathbf{k}$ ('linear' means using scalar multiplication and addition).

For example, $\langle 5, -1, -4 \rangle = 5\mathbf{i} + (-1)\mathbf{j} + (-4)\mathbf{k}$.

The *determinant* is a shorthand for specific expressions. It will soon become evident that these, perhaps initially complicated looking, expressions have rather amazing properties.

The *determinant of a 1×1 matrix* is $\det(a_{11}) = a_{11}$.
The *determinant of a 2×2 matrix* is

$$\det\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = a_{11}a_{22} - a_{12}a_{21}.$$

The *determinant of a 3×3 matrix* is

$$\det\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{matrix} a_{11}a_{22}a_{33} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31} \\ -a_{11}a_{23}a_{32} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32}. \end{matrix}$$

Note that

$$\det\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{matrix} a_{11}\det\begin{pmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{pmatrix} - a_{12}\det\begin{pmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{pmatrix} \\ +a_{13}\det\begin{pmatrix} a_{21} & a_{21} \\ a_{31} & a_{32} \end{pmatrix}. \end{matrix}$$

Definition (Cross product)

Let $\mathbf{u} = |u_1, u_2, u_3\rangle \in \mathbb{R}^3$ and let $\mathbf{v} = |v_1, v_2, v_3\rangle \in \mathbb{R}^3$. The *cross product* of $\mathbf{u}$ and $\mathbf{v}$ is given by

$$\mathbf{u} \times \mathbf{v} = (u_2 v_3 - u_3 v_2)\mathbf{i} + (u_3 v_1 - u_1 v_3)\mathbf{j} + (u_1 v_2 - u_2 v_1)\mathbf{k}.$$

In terms of determinants $\mathbf{u} \times \mathbf{v}$ is

$$\begin{aligned}\mathbf{u} \times \mathbf{v} &= \det\begin{pmatrix} u_2 & u_3 \\ v_2 & v_3 \end{pmatrix}\mathbf{i} - \det\begin{pmatrix} u_1 & u_3 \\ v_1 & v_3 \end{pmatrix}\mathbf{j} + \det\begin{pmatrix} u_1 & u_2 \\ v_1 & v_2 \end{pmatrix}\mathbf{k} \\ &\text{“}=\text{”} \ \det\begin{pmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{pmatrix},\end{aligned}$$

where the last 3×3 determinant on the right hand side doesn't really make sense (because $\mathbf{i}, \mathbf{j}, \mathbf{k}$ are not numbers); but this “determinant” is a very useful mnemonic.

If $\mathbf{u} = |u_1, u_2, u_3\rangle, \mathbf{v} = |v_1, v_2, v_3\rangle, \mathbf{w} = |w_1, w_2, w_3\rangle$ then

$$\begin{aligned}
\langle \mathbf{u}, \mathbf{v} \times \mathbf{w} \rangle &= \langle u_1, u_2, u_3 | (v_2 w_3 - w_3 v_2, -(v_1 w_3 - v_3 w_1), v_1 w_2 - v_2 w_1 \rangle \\
&= u_1(v_2 w_3 - v_3 w_2) - u_2(v_1 w_3 - v_3 w_1) + u_3(v_1 w_2 - v_2 w_1) \\
&= \det \begin{pmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{pmatrix} .
\end{aligned}$$

Since

$$\langle \mathbf{v}, \mathbf{v} \times \mathbf{w} \rangle = \det \begin{pmatrix} v_1 & v_2 & v_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{pmatrix} = 0$$

and

$$\langle \mathbf{w}, \mathbf{v} \times \mathbf{w} \rangle = \det \begin{pmatrix} w_1 & w_2 & w_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{pmatrix} = 0$$

then

$\mathbf{v} \times \mathbf{w}$ is perpendicular to both $\mathbf{v}$ and $\mathbf{w}$.

Example E5. Find a vector perpendicular to both $|1,1,1\rangle$ and $|1,-1,-2\rangle$.

Solution: By definition of the cross product

$$\begin{aligned} &|1,1,1\rangle \times |1,-1,-2\rangle \\ &\quad = |1\cdot(-2) - 1\cdot(-1), -(1\cdot(-2) - 1\cdot 1), 1\cdot(-1) - 1\cdot 1\rangle \\ &\quad = |-1,3,-2\rangle. \end{aligned}$$

The vector $|-1,3,-2\rangle$ is perpendicular to both $|1,1,1\rangle$ and $|1,-1,-2\rangle$ since

$$\langle -1,3,-2 \mid 1,1,1\rangle = -1+3-2 = 0$$

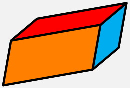
and

$$\langle -1,3,-2 \mid 1,-1,-2\rangle = -1-3+4 = 0.$$

Theorem (Volumes of parallelipipeds)

(3) *Let* $\mathbf{u} = |u_1, u_2, u_3\rangle \in \mathbb{R}^3$ *and* $\mathbf{v} = |v_1, v_2, v_3\rangle \in \mathbb{R}^3$ *and* $\mathbf{w} = |w_1, w_2, w_3\rangle \in \mathbb{R}^3$. *The volume of the parallelipiped with vertices* $0, \mathbf{u}, \mathbf{v}, \mathbf{w}, \mathbf{u}+\mathbf{v}, \mathbf{u}+\mathbf{w}, \mathbf{v}+\mathbf{w}, \mathbf{u}+\mathbf{v}+\mathbf{w}$ *is*

$$\left|\det\begin{pmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{pmatrix}\right|.$$



(2) *Let* $\mathbf{u} = |u_1, u_2\rangle \in \mathbb{R}^2$ *and* $\mathbf{v} = |v_1, v_2\rangle \in \mathbb{R}^2$. *The area of the paralellogram with vertices* $0, \mathbf{u}, \mathbf{v}, \mathbf{u}+\mathbf{v}$ *is*

$$\left|\det\begin{pmatrix} u_1 & u_2 \\ v_1 & v_2 \end{pmatrix}\right|.$$

(1) *Let* $\mathbf{u} = |u_1\rangle \in \mathbb{R}^1$. *The length of the segment with endpoints* 0 *to* $\mathbf{u}$ *is*

$$|\det(u_1)|.$$

Example E7. Find the volume of the parallelipiped with adjacent edges $\overrightarrow{PQ}$, $\overrightarrow{PR}$, $\overrightarrow{PS}$, where

$$P = |2, 0, -1\rangle, \;\; Q = |4, 1, 0\rangle, \;\; R = |3, -1, 1\rangle \text{ and } S = |2, -2, 2\rangle.$$

Since the edges of the parallelipiped are

$$\overrightarrow{PQ} = P - Q = |2, 1, 1\rangle, \quad \overrightarrow{PR} = P - R = |1, -1, 2\rangle,$$

$$\overrightarrow{PS} = P - S = |0, -2, 3\rangle,$$

then

$$\begin{aligned}(\text{Volume of parallelipiped}) &= |\langle \overrightarrow{PQ}, \overrightarrow{PR} \times \overrightarrow{PS}\rangle| \\ &= \left| \det \begin{pmatrix} 2 & 1 & 1 \\ 1 & -1 & 2 \\ 0 & -2 & 3 \end{pmatrix} \right| = \left| 2 \cdot \det \begin{pmatrix} -1 & 2 \\ -2 & 3 \end{pmatrix} - \det \begin{pmatrix} 1 & 1 \\ -2 & 3 \end{pmatrix} \right| \\ &= |2(-3+4) - (3+2)| = |-3| = 3.\end{aligned}$$

Example E6. Find the area of the triangle in $\mathbb{R}^3$ with vertices $|2,-5,4\rangle$, $|3,-4,5\rangle$ and $|3,-6,2\rangle$.

$$\text{Letting}\quad \mathbf{u} = |3,-4,5\rangle - |2,-5,4\rangle = |1,1,1\rangle \qquad \text{and}$$
$$\mathbf{v} = |3,-6,2\rangle - |2,-5,4\rangle = |1,-1,-2\rangle, \quad \text{then}$$

$$\begin{aligned}\mathbf{u}\times\mathbf{v} &= |1,1,1\rangle \times |1,-1,-2\rangle \\ &= |1\cdot(-2) - 1\cdot(-1), -(1\cdot(-2) - 1\cdot 1), 1\cdot(-1) - 1\cdot 1\rangle \\ &= |-1,3,-2\rangle.\end{aligned}$$

Then

$$\begin{aligned}(\text{Area of triangle}) &= \tfrac{1}{2}(\text{area of rectangle with edges } \mathbf{u} \text{ and } \mathbf{v}) \\ &= \tfrac{1}{2}\frac{1}{\|\mathbf{u}\times\mathbf{v}\|}\begin{pmatrix}\text{volume of parallelipiped} \\ \text{with edges } \mathbf{u}, \mathbf{v} \text{ and } \mathbf{u}\times\mathbf{v})\end{pmatrix} \\ &= \tfrac{1}{2}\frac{1}{\|\mathbf{u}\times\mathbf{v}\|}\langle \mathbf{u}\times\mathbf{v}, \mathbf{u}\times\mathbf{v}\rangle \\ &= \tfrac{1}{2}\|\mathbf{u}\times\mathbf{v}\| = \tfrac{1}{2}\|\,|-1,3,-2\rangle\,\| \\ &= \tfrac{1}{2}\sqrt{(-1)^2 + 3^2 + (-2)^2} = \frac{\sqrt{14}}{2}.\end{aligned}$$

Example E10. Find the Cartesian equation of the plane with vector form

$$|x, y, z\rangle = s\,|1, -1, 0\rangle + t\,|2, 0, 1\rangle + |-1, 1, 1\rangle, \text{ with } s, t \in \mathbb{R}.$$

A normal vector to this plane is

$$\mathbf{n} = \mathbf{u} \times \mathbf{v}, \qquad \text{where } \mathbf{u} = |1, -1, 0\rangle \text{ and } \mathbf{v} = |2, 0, 1\rangle.$$

Then $\mathbf{n} = \mathbf{u} \times \mathbf{v} = |-1-0, -(1-0), 0-(-2)\rangle = |-1, -1, 2\rangle$.
Then $|-1, 1, 1\rangle$ is a point in the plane, and

$$\langle -1, 1, 1 \,|\, \mathbf{u} \times \mathbf{v}\rangle = \langle -1, 1, 1 \,|\, -1, -1, 2\rangle = 1 - 1 + 2 = 2.$$

Since the plane is

$$|-1, 1, 1\rangle + \{|x, y, z\rangle \in \mathbb{R}^3 \mid \langle x, y, z \,|\, -1, -1, 2\rangle = 0\}$$

then the Cartesian equation of the plane is

$$-x - y + 2z = 2.$$

Lecture 4: Matrices

A matrix is a table of numbers.

$$A = \begin{pmatrix} 78 & 62 & 91 & 85 \\ 32 & 41 & 24 & 39 \\ 6 & 99 & 29 & 81 \end{pmatrix}$$

Some applications of matrices are

1. Solving systems of linear equations
2. lengths, distances, angles, projections
3. Equations of lines and planes, volumes of parallelipipeds
4. graphs and networks
5. Data processing and analysis of data
6. Dynamics
7. Symmetry
8. Quantum mechanics
9. ... and many many more ...

Addition

$$\begin{pmatrix} 78 & 62 & 91 & 85 \\ 32 & 41 & 24 & 39 \\ 6 & 99 & 29 & 81 \end{pmatrix} + \begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ -1 & -2 & -3 & -4 \end{pmatrix} = \begin{pmatrix} 79 & 64 & 94 & 89 \\ 37 & 48 & 31 & 47 \\ 5 & 97 & 26 & 77 \end{pmatrix}$$

Scalar multiplication

$$\frac{1}{3}\begin{pmatrix} 78 & 62 & 91 & 85 \\ 32 & 41 & 24 & 39 \\ 6 & 99 & 29 & 81 \end{pmatrix} = \begin{pmatrix} 26 & \frac{62}{3} & 27 & \frac{85}{3} \\ 10\frac{2}{3} & \frac{41}{3} & 8 & 13 \\ 2 & 33 & \frac{29}{3} & 27 \end{pmatrix}$$

Definition (Matrix units)

Let $t, s \in \mathbb{Z}_{>0}$ and let $i \in \{1, \ldots, t\}$ and $j \in \{1, \ldots, s\}$. The *matrix unit E_{ij}* is the matrix

$$E_{ij} \in M_{t\times s}(\mathbb{Q}) \quad \text{which has} \quad \begin{array}{l}\text{1 in the } (i,j)\text{-entry} \\ \text{and 0 elsewhere,}\end{array}$$

The favourite basis of $M_{t\times s}(\mathbb{Q})$

If $t = 2$ and $s = 3$ then

$$E_{11} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad E_{12} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad E_{13} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

$$E_{21} = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \qquad E_{22} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \quad E_{23} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Every matrix is a (unique) linear combination of E_{ij}

('linear' means using scalar multiplication and addition).

$$\begin{pmatrix} 78 & 62 & 91 \\ 32 & 41 & 24 \end{pmatrix} = \begin{matrix} 78E_{11} + 62E_{12} + 91E_{13} \\ +32E_{21} + 41E_{22} + 24E_{23} \end{matrix}$$

Multiplication

In English. The (i,j) entry of AB is the ith row of A times the jth column of B.

In Math.

$$E_{ij}E_{k\ell} = \delta_{jk}E_{i\ell}, \qquad \text{where} \quad \delta_{jk} = \begin{cases} 1, & \text{if } j=k, \\ 0 & \text{otherwise.} \end{cases}$$

Examples:

$$\begin{pmatrix} 2 & 5 & 11 & 13 \end{pmatrix} \begin{pmatrix} 4 \\ 0 \\ 3 \\ -2 \end{pmatrix} \begin{aligned} &= 2\cdot 4 + 5\cdot 0 + 11\cdot 3 + 13\cdot(-2) \\ &= 8 + 33 - 26 = 15. \end{aligned}$$

$$\begin{pmatrix} 78 & 62 & 91 & 85 \\ 32 & 41 & 24 & 39 \\ 6 & 99 & 29 & 81 \end{pmatrix} \begin{pmatrix} \frac{2}{100} \\ \frac{85}{100} \\ \frac{1}{100} \\ \frac{12}{100} \end{pmatrix} = \begin{pmatrix} \frac{78\cdot 2+62\cdot 85+91\cdot 1+85\cdot 12}{100} \\ \frac{32\cdot 2+41\cdot 85+24\cdot 1+39\cdot 12}{100} \\ \frac{6\cdot 2+99\cdot 85+29\cdot 1+81\cdot 12}{100} \end{pmatrix} = \begin{pmatrix} 65.37 \\ 40.41 \\ 94.28 \end{pmatrix}$$

Theorem (Properties of matrix operations)

Let $t, s \in \mathbb{Z}_{>0}$ and let $M_{t\times s}(\mathbb{Q})$ denote the set of $t \times s$ matrices with entries in $\mathbb{Q}$.

1. *If $A, B \in M_{t\times s}(\mathbb{Q})$ then $A + B = B + A$.*
2. *If $A, B, C \in M_{t\times s}(\mathbb{Q})$ then $A + (B + C) = (A + B) + C$.*
3. *If $A \in M_{t\times s}(\mathbb{Q})$, $B \in M_{s\times r}(\mathbb{Q})$ and $C \in M_{r\times q}(\mathbb{Q})$ then*
$$A(BC) = (AB)C.$$
4. *If $A, B \in M_{t\times s}(\mathbb{Q})$ and $C, D \in M_{s\times r}(\mathbb{Q})$ then*
$$A(C + D) = AC + AD \quad \text{and} \quad (A + B)C = AC + BC.$$
5. *If $A \in M_{t\times s}(\mathbb{Q})$, $B \in M_{s\times r}(\mathbb{Q})$ and $c \in \mathbb{Q}$ then $A(cB) = c(AB)$.*
6. *If $A \in M_{t\times s}(\mathbb{Q})$ and 1 is the identity in $M_{s\times s}(\mathbb{Q})$ then $A \cdot 1 = A$.*
7. *If $A \in M_{t\times s}(\mathbb{Q})$ and 1 is the identity in $M_{t\times t}(\mathbb{Q})$ then $1 \cdot A = A$.*
8. *If $A \in M_{t\times s}(\mathbb{Q})$ then $A + 0 = A$ and $0 + A = A$.*

Warning. The list of properties of matrix operations says that for the most part the matrix number system works much like the ordinary integer number system. But be careful.

$$\text{If} \quad A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$$

then

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

and

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

and so AB is *not the same* as BA. For most matrices A and B, the product AB is not the same as BA. When it does happen, that should be viewed as very special and very lucky. Don't push your luck.

Definition (Invertible matrices)

Let $n \in \mathbb{Z}_{>0}$. Let E_{ij} be the $n \times n$ matrix with 1 in the (i,j) entry and 0 elswhere. The *identity matrix* is

$$1 = E_{11} + \cdots + E_{nn} \quad \text{in } M_{n\times n}(\mathbb{Q}).$$

The set of invertible $n \times n$ matrices is

$$GL_n(\mathbb{Q}) = \left\{ A \in M_{n\times n}(\mathbb{Q}) \;\middle|\; \begin{matrix} \text{there exists } A^{-1} \in M_{n\times n}(\mathbb{Q}) \\ \text{such that } AA^{-1} = 1 \text{ and } A^{-1}A = 1. \end{matrix} \right\}$$

If $n = 2$ then

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

and

$$\begin{pmatrix} 0 & 1 \\ 1 & -c \end{pmatrix} \begin{pmatrix} c & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = 1.$$

Example A1. (Root matrices and their inverses) If $c \in \mathbb{Q}$ and

$$x_{12}(c) = \begin{pmatrix} 1 & c & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad x_{13}(c) = \begin{pmatrix} 1 & 0 & c \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad x_{23}(c) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & c \\ 0 & 0 & 1 \end{pmatrix}$$

then

$$x_{12}(c)^{-1} = \begin{pmatrix} 1 & -c & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad x_{13}(c)^{-1} = \begin{pmatrix} 1 & 0 & -c \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

and

$$x_{23}(c)^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -c \\ 0 & 0 & 1 \end{pmatrix}.$$

Check these claims by multiplying the matrices.

Example A2. (Diagonal generators and their inverses) If $d \in \mathbb{Q}$ and $d \neq 0$ and

$$h_1(d) = \begin{pmatrix} d & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \; h_2(d) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & d & 0 \\ 0 & 0 & 1 \end{pmatrix}, \; h_3(d) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & d \end{pmatrix}$$

then

$$h_1(d)^{-1} = \begin{pmatrix} \frac{1}{d} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad h_2(d)^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{d} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

and

$$h_3(d)^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{1}{d} \end{pmatrix}.$$

Check these claims by multiplying the matrices.

Example A3. (Row reducers and their inverses.) If $c \in \mathbb{Q}$ then

$$s_1(c) = \begin{pmatrix} c & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{and} \quad s_2(c) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & c & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

then

$$s_1(c)^{-1} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & -c & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{and} \quad s_2(c)^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & -c \end{pmatrix}$$

Check these claims by multiplying the matrices.

Let $n \in \mathbb{Z}_{>0}$. Let E_{ij} be the $n \times n$ matrix with 1 in the (i, j) entry and 0 elswhere.

Definition (root matrices, diagonal generators and row reducers)

Let $i, j \in \{1, \ldots, n\}$ with $i \neq j$. Let $c \in \mathbb{Q}$. The *root matrix $x_{ij}(c)$* is

$$x_{ij}(c) \in M_{n\times n}(\mathbb{Q}) \quad \text{given by} \quad x_{ij}(c) = 1 + cE_{ij}.$$

Let $i \in \{1, \ldots, n\}$. Let $d \in \mathbb{Q}$ with $d \neq 0$. The *diagonal generator $h_i(d)$* is

$$h_i(d) = 1 + (d-1)E_{ii}.$$

Let $i \in \{1, \ldots, n-1\}$ and let $c \in \mathbb{Q}$. The *row reducer $s_i(c)$* is

$$s_i(c) = 1 - E_{ii} - E_{i+1,i+1} + E_{i,i+1} + E_{i+1,i} + cE_{ii}.$$

Theorem (Generators for GL_n)

Let $A \in GL_n(\mathbb{Q})$. Then A can be written as a product of row reducers, diagonal generators and root matrices.

Tutorial: Row operations

Let $n \in \mathbb{Z}_{>0}$. Let E_{ij} be the $n \times n$ matrix with 1 in the (i,j) entry and 0 elsewhere.

Definition (root matrices, diagonal generators and row reducers)

Let $i, j \in \{1, \ldots, n\}$ with $i \neq j$. Let $c \in \mathbb{Q}$. The *root matrix $x_{ij}(c)$* is

$$x_{ij}(c) \in M_{n\times n}(\mathbb{Q}) \quad \text{given by} \quad x_{ij}(c) = 1 + cE_{ij}.$$

Let $i \in \{1, \ldots, n\}$. Let $d \in \mathbb{Q}$ with $d \neq 0$. The *diagonal generator $h_i(d)$* is

$$h_i(d) = 1 + (d-1)E_{ii}.$$

Let $i \in \{1, \ldots, n-1\}$ and let $c \in \mathbb{Q}$. The *row reducer $s_i(c)$* is

$$s_i(c) = 1 - E_{ii} - E_{i+1,i+1} + E_{i,i+1} + E_{i+1,i} + cE_{ii}.$$

Row operations

Let

$$A = \begin{pmatrix} 3 & -9 & 7 \\ 13 & -21 & 35 \\ 300 & -100 & 200 \end{pmatrix} \quad \text{and} \quad x_{13}(54) = \begin{pmatrix} 1 & 0 & 54 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Left multiplying by $x_{13}(54)$ adds $54 \cdot (\text{row } 3)$ to row 1:

$$\begin{aligned} x_{13}(54)A &= \begin{pmatrix} 1 & 0 & 54 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 3 & -9 & 7 \\ 13 & -21 & 35 \\ 300 & -100 & 200 \end{pmatrix} \\ &= \begin{pmatrix} 16203 & -5409 & 10807 \\ 13 & -21 & 35 \\ 300 & -100 & 200 \end{pmatrix}. \end{aligned}$$

Row operations

Let

$$A = \begin{pmatrix} 3 & -9 & 7 \\ 13 & -21 & 35 \\ 300 & -100 & 200 \end{pmatrix} \quad \text{and} \quad h_3(6) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 6 \end{pmatrix}.$$

Left multiplying by $h_3(6)$ multiplies row 3 by 6:

$$\begin{aligned} h_3(6)A &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 6 \end{pmatrix} \begin{pmatrix} 3 & -9 & 7 \\ 13 & -21 & 35 \\ 300 & -100 & 200 \end{pmatrix} \\ &= \begin{pmatrix} 3 & -9 & 7 \\ 13 & -21 & 35 \\ 1800 & -600 & 1200 \end{pmatrix}. \end{aligned}$$

Row operations

Let

$$A = \begin{pmatrix} 3 & -9 & 7 \\ 13 & -21 & 35 \\ 300 & -100 & 200 \end{pmatrix} \quad \text{and} \quad s_2(-5) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -5 & 1 \\ 0 & 1 & 0 \end{pmatrix}.$$

Left multiplying by $s_2(-5)$ moves row 2 to be row 3 and makes row 2 equal to $(-5) \cdot (\text{row } 2) + (\text{row } 3)$:

$$\begin{aligned} s_2(-5) \cdot A &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & -5 & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 3 & -9 & 7 \\ 13 & -21 & 35 \\ 300 & -100 & 200 \end{pmatrix} \\ &= \begin{pmatrix} 3 & -9 & 7 \\ 235 & 5 & 25 \\ 13 & -21 & 35 \end{pmatrix}. \end{aligned}$$

Lecture 5: Finding inverses

Definition (Invertible matrices)

Let $n \in \mathbb{Z}_{>0}$. *The set of invertible $n \times n$ matrices* is

$$GL_n(\mathbb{Q}) = \left\{ A \in M_{n\times n}(\mathbb{Q}) \;\middle|\; \begin{array}{l} \text{there exists } A^{-1} \in M_{n\times n}(\mathbb{Q}) \\ \text{such that } AA^{-1} = 1 \text{ and } A^{-1}A = 1. \end{array} \right\}$$

Definition (root matrices, diagonal generators and row reducers)

Let $n \in \mathbb{Z}_{>0}$. Let E_{ij} be the $n \times n$ matrix with 1 in the (i,j) entry and 0 elsewhere. Let $i, j \in \{1, \ldots, n\}$ with $i \neq j$. Let $c \in \mathbb{Q}$. The *root matrices $x_{k\ell}(c)$*, the *diagonal generators* and the *row reducers* are given by

$$x_{k\ell}(c) = 1 + cE_{k\ell}, \quad h_k(d) = 1 + (d-1)E_{kk} \quad \text{and}$$

$$s_i(c) = 1 - E_{ii} - E_{i+1,i+1} + E_{i,i+1} + E_{i+1,i} + cE_{ii}.$$

for $c, d \in \mathbb{Q}$ with $d \neq 0$, $k, \ell \in \{1, \ldots, n\}$ with $k \neq \ell$ and $i \in \{1, \ldots, n-1\}$.

Example M6 Find the inverse of $A = \begin{pmatrix} 1 & 2 & 1 \\ -1 & -1 & 1 \\ 0 & 1 & 3 \end{pmatrix}$.

Start with $AA^{-1} = 1$ which is

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

Left multiply both sides by $s_1(-1)^{-1}$, which is the matrix

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \text{to get} \quad \begin{pmatrix} -1 & -1 & 1 \\ 0 & 1 & 2 \\ 0 & 1 & 3 \end{pmatrix} A^{-1} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Left multiply both sides by $s_2(1)^{-1}$, which is the matrix

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & -1 \end{pmatrix}, \quad \text{to get} \quad \begin{pmatrix} -1 & -1 & 1 \\ 0 & 1 & 3 \\ 0 & 0 & -1 \end{pmatrix} A^{-1} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & -1 \end{pmatrix}.$$

Left multiply both sides by $h_3(-1)^{-1}h_1(-1)^{-1}$, which is the matrix

$$\begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \quad \text{to get} \quad \begin{pmatrix} 1 & 1 & -1 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{pmatrix} A^{-1} = \begin{pmatrix} 0 & -1 & 0 \\ 0 & 0 & 1 \\ -1 & -1 & 1 \end{pmatrix}.$$

Left multiply both sides by $x_{23}(3)^{-1}$, which is the matrix

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -3 \\ 0 & 0 & 1 \end{pmatrix}, \quad \text{to get} \quad \begin{pmatrix} 1 & 1 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} A^{-1} = \begin{pmatrix} 0 & -1 & 0 \\ 3 & 3 & -2 \\ -1 & -1 & 1 \end{pmatrix}.$$

Left multiply both sides by $x_{13}(-1)^{-1}$, which is the matrix

$$\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \text{to get} \quad \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} A^{-1} = \begin{pmatrix} -1 & -2 & 1 \\ 3 & 3 & -2 \\ -1 & -1 & 1 \end{pmatrix}.$$

Left multiply both sides by $x_{12}(1)^{-1}$, which is the matrix

$$\begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \text{to get} \quad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} A^{-1} = \begin{pmatrix} -4 & -5 & 3 \\ 3 & 3 & -2 \\ -1 & -1 & 1 \end{pmatrix}.$$

Check:

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

In summary,

$$A^{-1} = \begin{array}{l} x_{12}(1)^{-1} x_{13}(-1)^{-1} x_{23}(3)^{-1} \\ \quad \cdot h_3(-1)^{-1} h_1(-1)^{-1} \cdot s_2(1)^{-1} s_1(-1)^{-1} \end{array}$$

and

$$A = s_1(-1) s_2(1) \cdot h_1(-1) h_3(-1) \cdot x_{23}(3) x_{13}(-1) x_{12}(1).$$

Example M6 Find the inverse of $A = \begin{pmatrix} 1 & 2 & 1 \\ -1 & -1 & 1 \\ 0 & 1 & 3 \end{pmatrix}$.

Start with $AA^{-1} = 1$ which is

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

Left multiply both sides by

$$\begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{to get} \quad \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 2 \\ 0 & 1 & 3 \end{pmatrix} A^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Left multiply both sides by

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix} \quad \text{to get} \quad \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix} A^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ -1 & -1 & 1 \end{pmatrix}$$

Left multiply both sides by

$$\begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{to get} \quad \begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} A^{-1} = \begin{pmatrix} 2 & 1 & -1 \\ 3 & 3 & -2 \\ -1 & -1 & 1 \end{pmatrix}$$

Left multiply both sides by

$$\begin{pmatrix} 1 & -2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{to get} \quad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} A^{-1} = \begin{pmatrix} -4 & -5 & -3 \\ 3 & 3 & -2 \\ -1 & -1 & 1 \end{pmatrix}$$

Example M8 Find the inverse of $\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$.

Start with $AA^{-1} = 1$ which is $\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} A^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

Left multiply both sides by

$$\begin{pmatrix} 0 & 1 \\ 1 & -\frac{1}{3} \end{pmatrix} \quad \text{to get} \quad \begin{pmatrix} 3 & 4 \\ 0 & \frac{2}{3} \end{pmatrix} A^{-1} = \begin{pmatrix} 0 & 1 \\ 1 & -\frac{1}{3} \end{pmatrix}.$$

Left multiply both sides by

$$\begin{pmatrix} \frac{1}{3} & 0 \\ 0 & \frac{3}{2} \end{pmatrix} \quad \text{to get} \quad \begin{pmatrix} 1 & \frac{4}{3} \\ 0 & 1 \end{pmatrix} A^{-1} = \begin{pmatrix} 0 & \frac{1}{3} \\ \frac{3}{2} & -2 \end{pmatrix}.$$

Left multiply both sides by

$$\begin{pmatrix} 1 & -\frac{4}{3} \\ 0 & 1 \end{pmatrix} \quad \text{to get} \quad \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} A^{-1} = \begin{pmatrix} -2 & 1 \\ \frac{3}{2} & -\frac{1}{2} \end{pmatrix}.$$

Theorem (Inverses of products)

Let $A, B \in GL_n(\mathbb{Q})$. Then $AB \in GL_n(\mathbb{Q})$ and

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

This is because, by associativity,

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

and

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

The theorem tells us that if we want to find A^{-1} we can factor A into a product of row reducers, diagonal generators and root matrices and then multiply the inverses of the factors (in reverse order) to get the inverse of A.

Theorem (Generators for GL_n)

Let $A \in GL_n(\mathbb{Q})$. Then A can be written as a product of row reducers, diagonal generators and root matrices.

Lecture 6: Factoring and the rank theorem

Root matrices.

$$x_{12}(c) = \begin{pmatrix} 1 & c & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad x_{13}(c) = \begin{pmatrix} 1 & 0 & c \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad x_{23}(c) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & c \\ 0 & 0 & 1 \end{pmatrix}$$

Diagonal generators.

$$h_1(d) = \begin{pmatrix} d & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \ h_2(d) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & d & 0 \\ 0 & 0 & 1 \end{pmatrix}, \ h_3(d) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & d \end{pmatrix}$$

Row reducers.

$$s_1(c) = \begin{pmatrix} c & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{and} \quad s_2(c) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & c & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

Inverses of Root matrices. $x_{12}(c)^{-1} = \begin{pmatrix} 1 & -c & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$,

$$x_{13}(c)^{-1} = \begin{pmatrix} 1 & 0 & -c \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad x_{23}(c)^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -c \\ 0 & 0 & 1 \end{pmatrix}$$

Inverses of Diagonal generators. $h_1(d)^{-1} = \begin{pmatrix} d^{-1} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$,

$$h_2(d)^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & d^{-1} & 0 \\ 0 & 0 & 1 \end{pmatrix}, \ h_3(d)^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & d^{-1} \end{pmatrix}$$

Inverses of Row reducers.

$$s_1(c)^{-1} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & -c & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{and} \quad s_2(c)^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & -c \end{pmatrix}$$

Example A8. By multiplying out the matrices on the left hand side check that

$$\begin{array}{r} s_4(1)s_3(2)s_2(3)s_1(4) \\ \cdot s_4(5)s_3(6)s_2(7) \\ \cdots s_4(8)s_3(9) \\ \cdot s_4(10) \end{array} = \begin{pmatrix} 4 & 7 & 9 & 10 & 1 \\ 3 & 6 & 8 & 1 & 0 \\ 2 & 5 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Example A9. By multiplying out the matrices on the left hand side check that

$$h_1(15)h_2(-3)h_3(76)h_4(-19)h_5(2) = \begin{pmatrix} 15 & 0 & 0 & 0 & 0 \\ 0 & -3 & 0 & 0 & 0 \\ 0 & 0 & 76 & 0 & 0 \\ 0 & 0 & 0 & -19 & 0 \\ 0 & 0 & 0 & 0 & 2 \end{pmatrix}.$$

Example A5. By multiplying out the matrices on the left hand side check that

$$\begin{aligned}&x_{45}(1)x_{35}(2)x_{25}(3)\\ &\cdot x_{15}(4)x_{34}(5)x_{24}(6)\\ &\cdot x_{14}(7)x_{23}(8)x_{13}(9)\\ &\cdot x_{12}(10)\end{aligned} = \begin{pmatrix} 1 & 10 & 9 & 7 & 4 \\ 0 & 1 & 8 & 6 & 3 \\ 0 & 0 & 1 & 5 & 2 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

Example A6. By multiplying out the matrices on the left hand side check that

$$\begin{aligned}&x_{12}(10)^{-1}x_{13}(9)^{-1}x_{23}(8)^{-1}\\ &\cdot x_{14}(7)^{-1}x_{24}(6)^{-1}x_{34}(5)^{-1}\\ &\cdot x_{15}(4)^{-1}x_{25}(3)^{-1}x_{35}(2)^{-1}\\ &\cdot x_{45}(1)^{-1}\end{aligned} = \begin{pmatrix} 1 & -10 & 71 & -302 & 186 \\ 0 & 1 & -8 & 34 & -21 \\ 0 & 0 & 1 & -5 & 3 \\ 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Then check that

$$\begin{pmatrix} 1 & -10 & 71 & -302 & 186 \\ 0 & 1 & -8 & 34 & -21 \\ 0 & 0 & 1 & -5 & 3 \\ 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 10 & 9 & 7 & 4 \\ 0 & 1 & 8 & 6 & 3 \\ 0 & 0 & 1 & 5 & 2 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

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

Definition (Invertible matrices)

Let $n \in \mathbb{Z}_{>0}$. *The set of invertible $n \times n$ matrices* is

$$GL_n(\mathbb{Q}) = \left\{ A \in M_{n\times n}(\mathbb{Q}) \;\middle|\; \begin{array}{l} \text{there exists } A^{-1} \in M_{n\times n}(\mathbb{Q}) \\ \text{such that } AA^{-1} = 1 \text{ and } A^{-1}A = 1. \end{array} \right\}$$

Example A1. By multiplying out the matrices on the left hand side check that

$$s_1(\tfrac{1}{3})h_1(3)h_2(\tfrac{2}{3})x_{12}(\tfrac{4}{3}) = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}.$$

Example A2. By multiplying out the matrices on the left hand side check that

$$x_{12}(\tfrac{4}{3})^{-1}h_2(\tfrac{2}{3})^{-1}h_1(3)^{-1}s_1(\tfrac{1}{3})^{-1} = \begin{pmatrix} -2 & 1 \\ \frac{3}{2} & -\frac{1}{2} \end{pmatrix}.$$

Then check that

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

Example A3. By multiplying out the matrices on the left hand side. check that

$$s_1(-1)s_2(1)h_1(-1)h_3(-1)x_{23}(3)x_{13}(-1)x_{12}(1) = \begin{pmatrix} 1 & 2 & 1 \\ -1 & -1 & 1 \\ 0 & 1 & 3 \end{pmatrix}$$

Example A4. By multiplying out the matrices on the left hand side check that

$$\begin{matrix} x_{12}(1)^{-1}x_{13}(-1)^{-1}x_{23}(3)^{-1} \\ \cdot h_3(-1)^{-1}h_1(-1)^{-1}s_2(1)^{-1}s_1(-1)^{-1} \end{matrix} = \begin{pmatrix} -4 & -5 & 3 \\ 3 & 3 & -2 \\ -1 & -1 & 1 \end{pmatrix}.$$

Then check that

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

An upcoming lecture will specify a specific factoring algorithm that can factor any matrix. The output of the factoring algorithm will give us the following theorems.

Theorem (Factoring for invertible matrices)

Let $n \in \mathbb{Z}_{>0}$. Let $A \in GL_n(\mathbb{Q})$. The factoring algorithm gives

$$A = (\text{product of } s_i(c)\text{s}) \cdot (\text{product of } h_i(d)\text{s}) \cdot (\text{product of } x_{ij}(c)\text{s})$$

This last theorem tells us that we can factor any invertible matrix as a product of $s_i(c)$s, $h_i(d)$s and $x_{ij}(c)$s. The next theorem deals with matrices that don't have to be invertible.

Theorem (Factoring for all matrices)

Let $s, t \in \mathbb{Z}_{>0}$. Let E_{ij} be the $t \times s$ matrix with 1 in the (i, j) entry and 0 elswhere. For $r \in \{1, \ldots, \min(s, t)\}$ let

$$1_r = E_{11} + \cdots + E_{rr}.$$

Let $A \in M_{t \times s}(\mathbb{Q})$. The factoring algorithm gives

$$A = (\textit{product of } s_i(c)s) \cdot (\textit{product of } h_i(d)s) \cdot (\textit{product of } x_{ij}(c)s) \\ \cdot 1_r \cdot (\textit{product of } s_i(c)s) \cdot (\textit{product of } x_{ij}(c)s).$$

The number r that comes out of the factoring algorithm is the *rank* of A. Later the rank of A will be realised as the dimension of the image of A,

$$r = \dim(\mathrm{im}(A)) = \mathrm{rank}(A) \quad \text{is the } \textit{rank} \text{ of } A.$$

Example M10. Let $A = \begin{pmatrix} 1 & -1 & 2 & 1 \\ 0 & 1 & 1 & -2 \\ 1 & -3 & 0 & 5 \end{pmatrix} \in M_{3\times 4}(\mathbb{Q})$. Then

$$A = s_2(0) \begin{pmatrix} 1 & -1 & 2 & 1 \\ 1 & -3 & 0 & 5 \\ 0 & 1 & 1 & -2 \end{pmatrix} = s_2(0)s_1(1) \begin{pmatrix} 1 & -3 & 0 & 5 \\ 0 & 2 & 2 & -4 \\ 0 & 1 & 1 & -2 \end{pmatrix}$$

$$= s_2(0)s_1(1)s_2(2) \begin{pmatrix} 1 & -3 & 0 & 5 \\ 0 & 1 & 1 & -2 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$= s_2(0)s_1(1)s_2(2)x_{12}(-3) \begin{pmatrix} 1 & 0 & 0 & 5 \\ 0 & 1 & 1 & -2 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Since this last right hand factor has a row of 0s then A is not invertible.

$$A = s_2(0)s_1(1)s_2(2)x_{12}(-3) \cdot 1_2 \cdot x_{23}(1)x_{14}(5)x_{24}(-2).$$

So $\mathrm{rank}(A) = 2$.

Tutorial: Inverses of an arbitrary 2×2 matrix

Example A1 Let $a, b, c, d \in \mathbb{Q}$ and find the inverse of $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$.

Start with $AA^{-1} = 1$, which is

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} A^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Case 1: $c \neq 0$. Left multiply by $s_1(\frac{a}{c})^{-1}$, which is the matrix

$$\begin{pmatrix} 0 & 1 \\ 1 & -\frac{a}{c} \end{pmatrix}, \quad \text{to get} \begin{pmatrix} c & d \\ 0 & b - \frac{a}{c}d \end{pmatrix} A^{-1} = \begin{pmatrix} 0 & 1 \\ 1 & -\frac{a}{c} \end{pmatrix}.$$

Case 1a: $c \neq 0$ and $ad - bc \neq 0$. Left multiply by $h_1(c)^{-1} h_2(\frac{bc-ad}{c})^{-1}$, which is the matrix

$$\begin{pmatrix} \frac{1}{c} & 0 \\ 0 & \frac{c}{bc-ad} \end{pmatrix}, \quad \text{to get} \begin{pmatrix} 1 & \frac{d}{c} \\ 0 & 1 \end{pmatrix} A^{-1} = \begin{pmatrix} 0 & \frac{1}{c} \\ \frac{c}{bc-ad} & -\frac{a}{bc-ad} \end{pmatrix}.$$

Left multiply by $x_{12}(\frac{d}{c})^{-1}$, which is the matrix

$$\begin{pmatrix} 1 & -\frac{d}{c} \\ 0 & 1 \end{pmatrix}, \quad \text{to get} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} A^{-1} = \begin{pmatrix} -\frac{d}{bc-ad} & \frac{1}{c} + \frac{ad}{c(bc-ad)} \\ \frac{c}{bc-ad} & -\frac{a}{bc-ad} \end{pmatrix}.$$

So

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

Case 1b: $c \neq 0$ and $ad - bc = 0$. Then

$$\begin{pmatrix} c & d \\ 0 & 0 \end{pmatrix} A^{-1} = \begin{pmatrix} 0 & 1 \\ 1 & -\frac{a}{c} \end{pmatrix}$$

and there does not exist any matrix A^{-1} that makes this equation true.

Case 2: $c = 0$. Then

$$\begin{pmatrix} a & b \\ 0 & d \end{pmatrix} A^{-1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Case 2a: $c = 0$ and $d \neq 0$ and $a \neq 0$.

Left multiply by $h_2(d)^{-1}h_1(a)^{-1}$, which is the matrix

$$\begin{pmatrix} \frac{1}{a} & 0 \\ 0 & \frac{1}{d} \end{pmatrix}, \quad \text{to get} \quad \begin{pmatrix} 1 & \frac{b}{a} \\ 0 & 1 \end{pmatrix} A^{-1} = \begin{pmatrix} \frac{1}{a} & 0 \\ 0 & \frac{1}{d} \end{pmatrix}.$$

Left mulltiply by $x_{12}(\frac{b}{a})^{-1}$, which is the matrix

$$\begin{pmatrix} \frac{1}{a} & 0 \\ 0 & \frac{1}{d} \end{pmatrix}, \quad \text{to get} \quad \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} A^{-1} = \begin{pmatrix} \frac{1}{a} & -\frac{b}{ad} \\ 0 & \frac{1}{d} \end{pmatrix} = \begin{pmatrix} \frac{d}{ad} & \frac{-b}{ad} \\ 0 & \frac{a}{ad} \end{pmatrix}.$$

Recalling that $c = 0$ then

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

Theorem (Inverse of a 2×2 matrix)

Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in M_{2\times 2}(\mathbb{Q})$. *Then*

1. *If* $ad - bc \neq 0$ *then* $A^{-1} = \dfrac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$.
2. *If* $ad - bc = 0$ *then* A^{-1} *does not exist.*

Example M5. Let $A = \begin{bmatrix} 2 & -1 \\ 1 & 1 \end{bmatrix}$. Then

$$A^{-1} = \frac{1}{(2\cdot 1 - 1\cdot(-1))}\begin{pmatrix} 1 & 1 \\ -1 & 2 \end{pmatrix} = \frac{1}{3}\begin{pmatrix} 1 & 1 \\ -1 & 2 \end{pmatrix} = \begin{pmatrix} \frac{1}{3} & \frac{1}{3} \\ \frac{-1}{3} & \frac{2}{3} \end{pmatrix}.$$

Check:

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

Suggestion: Figure out the formulas for the inverse of an arbitrary 3×3 matrix.

What are the $s_i(c)$ matrices?

In math: Let $n \in \mathbb{Z}_{>0}$ and let E_{ij} be the $n \times n$ matrix with 1 in the (i, j) entry and 0 elsewhere. For $i \in \{1, \ldots, n-1\}$ and $p, q \in \mathbb{Z}$ with $q \neq 0$ define

$$s_i\left(\tfrac{p}{q}\right) = 1 - E_{ii} - E_{i+1,i+1} + E_{i,i+1} - E_{i+1,i} + \tfrac{p}{q} E_{i,i}.$$

Note:

$$s_i\left(\tfrac{p}{q}\right)^{-1} = 1 - E_{ii} - E_{i+1,i+1} + E_{i,i+1} - E_{i+1,i} - \tfrac{p}{q} E_{i+1,i+1}.$$

In English: $s_i\left(\frac{p}{q}\right)$ is the $n \times n$ matrix with

(a) 1s on the diagonal except that the (i, i) enrty is c and the $(i+1, i+1)$ entry is 0, and

(b) all other entries are 0 except that the $(i, i+1)$ entry is 1 and the $(i+1, i)$ entry is 0.

What are the $s_i(c)$ matrices?

By Cartoon: If $n = 8$ and $\frac{p}{q} = \frac{7}{12}$ then

$$s_6\left(\tfrac{7}{12}\right) = \begin{pmatrix} 1 &&&&&&& \\ & 1 &&&&&& \\ && 1 &&&&& \\ &&& 1 &&&& \\ &&&& 1 &&& \\ &&&&& \frac{7}{12} & 1 & \\ &&&&& 1 & 0 & \\ &&&&&&& 1 \end{pmatrix}$$

Note

$$s_6\left(\tfrac{7}{12}\right)^{-1} = \begin{pmatrix} 1 &&&&&&& \\ & 1 &&&&&& \\ && 1 &&&&& \\ &&& 1 &&&& \\ &&&& 1 &&& \\ &&&&& 0 & 1 & \\ &&&&& 1 & -\frac{7}{12} & \\ &&&&&&& 1 \end{pmatrix}$$

We will factor off $s_i(\frac{p}{q})$ matrices,

step by step, to make more and more lower triangular entries 0.

Make lower triangular entries 0 in this order:

$$\begin{pmatrix} * & * & * & * & * \\ 4 & * & * & * & * \\ 3 & 7 & * & * & * \\ 2 & 6 & 9 & * & * \\ 1 & 5 & 8 & 10 & * \end{pmatrix}.$$

To make the (nonzero) (i, j) entry of the matrix A into 0:

In Math: Let q be the (i, j)-entry of A and let p be the $(i-1, j)$ entry of A. Assume $q \neq 0$. Then

$$A = s_{i-1}\left(\tfrac{p}{q}\right)B, \qquad \text{where} \quad B = s_{i-1}\left(\tfrac{p}{q}\right)^{-1}A,$$

and B has 0 in the (i, j)-entry.

In English: Let q be the (i,j) entry of A. If $q \neq 0$ then make the (i,j) into 0 as follows. Let p be the $(i-1,j)$ entry of A Then write

$$A = s_{i-1}\left(\tfrac{p}{q}\right)B, \qquad \text{where}$$

(a) The ith row of A moves up one row
to become the $(i-1)$st row of B,

(b) The ith row of B is ((the $(i-1)$st row of A)-$\frac{p}{q}$(ith row of A)), and

(c) all other rows of B are the same as the corresponding rows fo A.

In hybrid Math-English:

$$A = s_{i-1}\left(\tfrac{p}{q}\right)B, \qquad \text{where}$$

(a) $\mathrm{row}_{i-1}(B) = \mathrm{row}_i(A)$,

(b) $\mathrm{row}_i(B) = \mathrm{row}_{i-1}(A) - \frac{p}{q}\mathrm{row}_i(A)$,

(c) if $j \notin \{i-1, i\}$ then $\mathrm{row}_j(B) = \mathrm{row}_(A)$, *In Cartoon:* Suppose

$$A = \begin{matrix} \\ i-1 \\ i \\ \\ \end{matrix} \left(\begin{array}{ccccccc} & & & \text{STUFF} & & & \\ 0 & 0 & p & r & t & v & x \\ 0 & 0 & q & s & u & w & y \\ 0 & 0 & 0 & z & e & f & g \\ & & & & & & \end{array} \right), \quad \text{with } q \neq 0.$$

Then

$$A = s_{i-1}\left(\tfrac{p}{q}\right) B,$$

where

$$B = \begin{matrix} \\ i-1 \\ i \\ \\ \end{matrix} \left(\begin{array}{ccccccc} & & & \text{STUFF} & & & \\ 0 & 0 & q & s & u & w & y \\ 0 & 0 & 0 & r - \frac{p}{q}s & t - \frac{p}{q}u & v - \frac{p}{q}w & x - \frac{p}{q}y \\ 0 & 0 & 0 & z & e & f & g \\ & & & & & & \end{array} \right).$$

In English: Let q be the (i,j) entry of A. If $q \neq 0$ then make the (i,j) into 0 as follows. Let p be the $(i-1,j)$ entry of A Then write

$$A = s_{i-1}\left(\tfrac{p}{q}\right)B, \qquad \text{where}$$

(a) The ith row of A moves up one row to become the $(i-1)$st row of B,

(b) The ith row of B is ((the $(i-1)$st row of A)-$\frac{p}{q}$(ith row of A)), and

(c) all other rows of B are the same as the corresponding rows fo A.

In hybrid Math-English:

$$A = s_{i-1}\left(\tfrac{p}{q}\right)B, \qquad \text{where}$$

$$\begin{array}{ll}
\text{(a)} & \mathrm{row}_{i-1}(B) = \mathrm{row}_i(A), \\
\text{(b)} & \mathrm{row}_i(B) = \mathrm{row}_{i-1}(A) - \frac{p}{q}\mathrm{row}_i(A), \\
\text{(c)} & \text{if } j \notin \{i-1, i\} \text{ then } \quad \mathrm{row}_j(B) = \mathrm{row}_(A),
\end{array}$$

Full row reduction.

Let $s, t \in \mathbb{Z}_{>0}$ and let $A \in M_{t\times s}(\mathbb{Q})$.

Let j_1 be minimal such that

$$\text{column } j_1 \text{ of } A \text{ has a nonzero entry.}$$

Let i_1 be maximal such that $A(i_1, j_1) \neq 0$. Let

$$A^{(1)} = s_1\left(\frac{A(1,j_1)}{A(i_1,j_1)}\right)^{-1} s_2\left(\frac{A(2,j_1)}{A(i_1,j_1)}\right)^{-1} \cdots s_{i_1-1}\left(\frac{A(i_1-1,j_1)}{A(i_1,j_1)}\right)^{-1} A.$$

Let j_2 be minimal such that

$$\text{column } j_2 \text{ of } A^{(1)} \text{ has a nonzero entry below row 1.}$$

Let $i_2 > 1$ be maximal such that $A^{(1)}(i_2, j_2) \neq 0$. Let

$$A^{(2)} = s_2\left(\frac{A^{(1)}(2,j_2)}{A^{(1)}(i_2,j_2)}\right)^{-1} s_3\left(\frac{A^{(1)}(3,j_2)}{A^{(1)}(i_2,j_2)}\right)^{-1} \cdots s_{i_2-1}\left(\frac{A^{(1)}(i_2-1,j_2)}{A^{(1)}(i_2,j_2)}\right)^{-1} A^{(1)}.$$

Let j_3 be minimal such that

$$\text{column } j_3 \text{ of } A^{(2)} \text{ has a nonzero entry below row 2.}$$

Let $i_3 > 2$ be maximal such that $A^{(2)}(i_3, j_3) \neq 0$. Let

$$A^{(3)} = s_3\left(\tfrac{A^{(2)}(3,j_3)}{A^{(2)}(i_3,j_3)}\right)^{-1} s_4\left(\tfrac{A^{(2)}(4,j_3)}{A^{(2)}(i_3,j_3)}\right)^{-1} \cdots s_{i_3-1}\left(\tfrac{A^{(2)}(i_2-1,j_2)}{A^{(2)}(i_3,j_3)}\right)^{-1} A^{(2)}.$$

Continue this process until it happens that there does not exist j_{r+1} such that column j_{r+1} of $A^{(r)}$ has a nonzero entry below row r.

Then $A^{(r)}$ has the property that

the first nonzero entry in row $j+1$
is to the right of the first nonzero entry in row j

and

$$
\begin{aligned}
A = & \left(s_{i_1-1}\left(\tfrac{A(i_1-1,j_1 1)}{A(i_1,j_1)}\right)\cdots s_2\left(\tfrac{A(2,j_1)}{A(i_1,j_1)}\right) s_1\left(\tfrac{A(1,j_1)}{A(i_1,j_1)}\right)\right) \\
& \cdot \left(s_{i_2-1}\left(\tfrac{A^{(1)}(i_2-1,j_2)}{A^{(1)}(i_2,j_2)}\right)\cdots s_3\left(\tfrac{A^{(1)}(3,j_2)}{A^{(1)}(i_2,j_2)}\right) s_2\left(\tfrac{A^{(1)}(2,j_2)}{A^{(1)}(i_2,j_2)}\right)\right) \\
& \cdots \\
& \cdot \left(s_{i_r-1}\left(\tfrac{A^{(r-1)}(j_r-1,j_r)}{A^{(r-1)}(i_r,j_r)}\right)\cdots s_{r+1}\left(\tfrac{A^{(r-1)}(r+1,j_r)}{A^{(r-1)}(i_r,j_r)}\right) s_r\left(\tfrac{A^{(r-1)}(r,j_r)}{A^{(r-1)}(i_r,j_r)}\right)\right) \\
& \cdot A^{(r)}.
\end{aligned}
$$

Then

$$
\begin{aligned}
A^{(r)} = & \big(h_1(A^{(r)}(1,j_1)) \cdots h_r(A^{(r)}(r,j_r))\big) \\
& \cdot \left(x_{r-1,j_r}\left(\frac{A^{(r)}(r-1,j_r)}{A^{(r)}(r-1,j_{r-1})}\right) \cdots x_{1,j_r}\left(\frac{A^{(r)}(1,j_r)}{A^{(r)}(1,j_1)}\right)\right) \\
& \cdots \\
& \cdot \left(x_{2,j_3}\left(\frac{A^{(r)}(2,j_3)}{A^{(r)}(2,j_2)}\right) x_{1,j_3}\left(\frac{A^{(r)}(1,j_3)}{A^{(r)}(1,j_1)}\right)\right) \\
& \cdot x_{1,j_2}\left(\frac{A^{(r)}(1,j_2)}{A^{(r)}(1,j_1)}\right) \\
& \cdot R,
\end{aligned}
$$

where R is given by

$$
R(k,j) = \begin{cases} \frac{A^{(r)}(k,j)}{A^{(r)}(k,j_k)}, & \text{if } k \in \{1,\dots,r\} \text{ and } j \in \{j_k, j_k+1, \dots, s\} \\ & \text{and } j \notin \{j_{k+1},\dots,j_r\}, \\ 0, & \text{otherwise.} \end{cases}
$$

Let $c_{r+1} < \cdots < c_{s-1} < c_s$ be such that $\{j_1, \ldots, j_r, c_{r+1}, \ldots, c_s\} = \{1, \ldots, s\}$. Then

$$R = 1_r \cdot Q, \quad \text{where} \quad Q \in GL_s(\mathbb{Q}) \quad \text{is given by}$$

$$\begin{aligned}
Q = {} & \Big(x_{r,s}\Big(\frac{A^{(r)}(r, c_s)}{A^{(r)}(r, j_r)}\Big) \cdots x_{1,s}\Big(\frac{A^{(r)}(1, c_s)}{A^{(r)}(1, j_1)}\Big)\Big) \\
& \cdot \Big(x_{r,s-1}\Big(\frac{A^{(r)}(r, c_{s-1})}{A^{(r)}(r, j_r)}\Big) \cdots x_{1,s}\Big(\frac{A^{(r)}(1, c_{s-1})}{A^{(r)}1, j_1}\Big)\Big) \\
& \cdots \\
& \cdot \Big(x_{r,r+1}\Big(\frac{A^{(r)}(r, c_{r+1})}{A^{(r)}(r, j_r)}\Big) \cdots x_{1,r+1}\Big(\frac{A^{(r)}(1, c_{r+1})}{A^{(r)}(1, j_1)}\Big)\Big) \\
& \cdot (s_r \cdots s_{j_r-1}) \cdots (s_2 \cdots s_{j_2-1}) \cdot (s_1 \cdots s_{j_1-1}).
\end{aligned}$$

Summary. In summary, $A = P1_rQ$ where $P \in GL_t(\mathbb{Q})$ and $Q \in GL_s(\mathbb{Q})$ are given by

$$
\begin{aligned}
P =& \left(s_{i_1-1}\left(\tfrac{A(i_1-1,j_11)}{A(i_1,j_1)}\right)\cdots s_2\left(\tfrac{A(2,j_1)}{A(i_1,j_1)}\right) s_1\left(\tfrac{A(1,j_1)}{A(i_1,j_1)}\right)\right)\\
&\cdot\left(s_{i_2-1}\left(\tfrac{A^{(1)}(i_2-1,j_2)}{A^{(1)}(i_2,j_2)}\right)\cdots\ s_3\left(\tfrac{A^{(1)}(3,j_2)}{A^{(1)}(i_2,j_2)}\right) s_2\left(\tfrac{A^{(1)}(2,j_2)}{A^{(1)}(i_2,j_2)}\right)\right)\\
&\cdots\\
&\cdot\left(s_{i_r-1}\left(\tfrac{A^{(r-1)}(j_r-1,j_r)}{A^{(r-1)}(i_r,j_r)}\right)\cdots s_{r+1}\left(\tfrac{A^{(r-1)}(r+1,j_r)}{A^{(r-1)}(i_r,j_r)}\right) s_r\left(\tfrac{A^{(r-1)}(r,j_r)}{A^{(r-1)}(i_r,j_r)}\right)\right)\\
&\left(h_1(A^{(r)}(1,j_1))\cdots h_r(A^{(r)}(r,j_r))\right)\\
&\cdot\left(x_{r-1,j_r}\left(\frac{A^{(r)}(r-1,j_r)}{A^{(r)}(r-1,j_{r-1})}\right)\cdots x_{1,j_r}\left(\frac{A^{(r)}(1,j_r)}{A^{(r)}(1,j_1)}\right)\right)\\
&\cdots\\
&\cdot\left(x_{2,j_3}\left(\frac{A^{(r)}(2,j_3)}{A^{(r)}(2,j_2)}\right)x_{1,j_3}\left(\frac{A^{(r)}(1,j_3)}{A^{(r)}(1,j_1)}\right)\right)\cdot x_{1,j_2}\left(\frac{A^{(r)}(1,j_2)}{A^{(r)}(1,j_1)}\right)
\end{aligned}
$$

and

$$Q = \left(x_{r,s}\left(\frac{A^{(r)}(r,c_s)}{A^{(r)}(r,j_r)}\right)\cdots x_{1,s}\left(\frac{A^{(r)}(1,c_s)}{A^{(r)}(1,j_1)}\right)\right)$$
$$\cdot \left(x_{r,s-1}\left(\frac{A^{(r)}(r,c_{s-1})}{A^{(r)}(r,j_r)}\right)\cdots x_{1,s}\left(\frac{A^{(r)}(1,c_{s-1})}{A^{(r)}1,j_1}\right)\right)$$
$$\cdots$$
$$\cdot \left(x_{r,r+1}\left(\frac{A^{(r)}(r,c_{r+1})}{A^{(r)}(r,j_r)}\right)\cdots x_{1,r+1}\left(\frac{A^{(r)}(1,c_{r+1})}{A^{(r)}(1,j_1)}\right)\right)$$
$$\cdot (s_r(0)\cdots s_{j_r-1}(0))\cdots(s_2(0)\cdots s_{j_2-1}(0))\cdot(s_1(0)\cdots s_{j_1-1}(0)).$$

Theorem (The rank theorem)

Let $A \in M_{t\times s}(\mathbb{Q})$. Then there exist $P \in GL_t(\mathbb{Q})$ and $Q \in GL_s(\mathbb{Q})$ and $r \in \{1,\dots,\min(s,t)\}$ such that

$$A = P1_rQ, \qquad \textit{where} \quad 1_r = E_{11}+E_{22}+\cdots+E_{rr}.$$

If

$$A = \begin{pmatrix} 3 & 1 \\ -1 & 4 \end{pmatrix} \quad \text{and} \quad \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \quad \text{and} \quad \mathbf{b} = \begin{pmatrix} 7 \\ 2 \end{pmatrix}$$

then

$$A\mathbf{x} = \mathbf{b} \quad \text{is the same as} \quad \begin{pmatrix} 3 & 1 \\ -1 & 4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 7 \\ 2 \end{pmatrix}$$

is the same as

$$\begin{pmatrix} 3x_1 + \ \ x_2 \\ -x_1 + 4x_2 \end{pmatrix} = \begin{pmatrix} 7 \\ 2 \end{pmatrix}$$

which is the same as

$$\begin{aligned} 3x_1 + \ \ x_2 &= 7 \\ -x_1 + 4x_2 &= 2 \end{aligned}$$

In general $A\mathbf{x} = \mathbf{b}$ looks like

$$\begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ \vdots & \vdots & & \vdots \\ A_{m1} & A_{m2} & \cdots & A_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}.$$

Definition (Solutions of a linear system)

Let $A \in M_{m\times n}(\mathbb{Q})$ and $b \in M_{n\times 1}(\mathbb{Q})$. The set of solutions of $A\mathbf{x} = \mathbf{b}$ is

$$\mathrm{Sol}(A\mathbf{x} = \mathbf{b}) = \left\{ \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in M_{n\times 1}(\mathbb{Q}) \;\middle|\; A\mathbf{x} = \mathbf{b} \right\}.$$

Example A. If $A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ and $\mathbf{b} = \begin{pmatrix} 7 \\ 2 \end{pmatrix}$ then $\mathrm{Sol}(A\mathbf{x} = \mathbf{b}) = \left\{ \begin{pmatrix} 7 \\ 2 \end{pmatrix} \right\}$

and $\begin{aligned} x_1 + 0x_2 &= 7, \\ 0x_1 + x_2 &= 2. \end{aligned}$ has exactly one solution $\begin{aligned} x_1 &= 7, \\ x_2 &= 2. \end{aligned}$

If $A = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}$ and $\mathbf{b} = \begin{pmatrix} 7 \\ 2 \end{pmatrix}$ then $\mathrm{Sol}(A\mathbf{x} = \mathbf{b}) = \emptyset$ and

$\begin{aligned} x_1 + 0x_2 &= 7, \\ x_1 + 0x_2 &= 2. \end{aligned}$ has no solutions.

If $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ and $\mathbf{b} = \begin{pmatrix} 7 \\ 0 \end{pmatrix}$ then $\mathrm{Sol}(A\mathbf{x} = \mathbf{b}) = \left\{ \begin{pmatrix} 7 \\ c \end{pmatrix} \;\middle|\; c \in \mathbb{Q} \right\}$,

$\begin{aligned} x_1 + 0x_2 &= 7, \\ 0x_1 + 0x_2 &= 0. \end{aligned}$ has infinitely many solutions $\begin{aligned} x_1 &= 7, \\ x_2 &= c, \end{aligned}$ for any $c \in \mathbb{Q}$.

Example LS2,3&4. If $A = \begin{pmatrix} 2 & -1 \\ 1 & 1 \end{pmatrix}$, $\mathbf{x} = \begin{pmatrix} x \\ y \end{pmatrix}$ and $\mathbf{b} = \begin{pmatrix} 3 \\ 0 \end{pmatrix}$ then $A\mathbf{x} = \mathbf{b}$ is

$$\begin{pmatrix} 2 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 3 \\ 0 \end{pmatrix} \quad \text{which is} \quad \begin{aligned} 2x - y &= 3, \\ x + y &= 0. \end{aligned}$$

Start with

$$\begin{pmatrix} 2 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 3 \\ 0 \end{pmatrix}.$$

Left multiply both sides by

$$\begin{pmatrix} 0 & 1 \\ 1 & -2 \end{pmatrix} \quad \text{to get} \quad \begin{pmatrix} 1 & 1 \\ 0 & -3 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 3 \end{pmatrix}.$$

Left multiply both sides by

$$\begin{pmatrix} 0 & 1 \\ 1 & -\frac{1}{3} \end{pmatrix} \quad \text{to get} \quad \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \end{pmatrix}.$$

Left multiply both sides by

$$\begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} \quad \text{to get} \quad \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \quad \text{which is} \quad \begin{matrix} x = 1, \\ y = -1. \end{matrix}$$

So $\quad \mathrm{Sol}(A\mathbf{x} = \mathbf{b}) = \left\{ \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\}$ (exactly *one* solution).

Example LS5&6. Solve the following system of linear equations.

$$
\begin{aligned}
4x - 2y + 5z &= 31, \\
2x - 3y - 2z &= 13, \\
x - 3y + 2z &= 11.
\end{aligned}
$$

In matrix form, this is $A\mathbf{x} = \mathbf{b}$, where

$$
A = \begin{pmatrix} 4 & -2 & 5 \\ 2 & -3 & -2 \\ 1 & -3 & 2 \end{pmatrix}, \quad \mathbf{x} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 31 \\ 13 \\ 11 \end{pmatrix}.
$$

Start with

$$
\begin{pmatrix} 4 & -2 & 5 \\ 2 & -3 & -2 \\ 1 & -3 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 31 \\ 13 \\ 11 \end{pmatrix}.
$$

Left multiply both sides by

$$\begin{pmatrix}1 & 0 & 0\\ 0 & 0 & 1\\ 0 & 1 & -2\end{pmatrix} \quad \text{to get} \quad \begin{pmatrix}4 & -2 & 5\\ 1 & -3 & 2\\ 0 & 3 & -6\end{pmatrix}\begin{pmatrix}x\\ y\\ z\end{pmatrix} = \begin{pmatrix}31\\ 11\\ -9\end{pmatrix}.$$

Left multiply both sides by

$$\begin{pmatrix}0 & 1 & 0\\ 1 & -4 & 1\\ 0 & 0 & 1\end{pmatrix} \quad \text{to get} \quad \begin{pmatrix}1 & -3 & 2\\ 0 & 10 & -3\\ 0 & 3 & -6\end{pmatrix}\begin{pmatrix}x\\ y\\ z\end{pmatrix} = \begin{pmatrix}11\\ -13\\ -9\end{pmatrix}.$$

Left multiply both sides by

$$\begin{pmatrix}1 & 0 & 0\\ 0 & 0 & 1\\ 0 & 1 & -\frac{10}{3}\end{pmatrix} \quad \text{to get} \quad \begin{pmatrix}1 & -3 & 2\\ 0 & 3 & -6\\ 0 & 0 & 17\end{pmatrix}\begin{pmatrix}x\\ y\\ z\end{pmatrix} = \begin{pmatrix}11\\ -9\\ 17\end{pmatrix}.$$

Left multiply both sides by

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{3} & 0 \\ 0 & 0 & \frac{1}{17} \end{pmatrix} \quad \text{to get} \quad \begin{pmatrix} 1 & -3 & 2 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 11 \\ -3 \\ 1 \end{pmatrix}.$$

Left multiply both sides by

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{to get} \quad \begin{pmatrix} 1 & -3 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 11 \\ -1 \\ 1 \end{pmatrix}.$$

Left multiply both sides by

$$\begin{pmatrix} 1 & 0 & -2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{to get} \quad \begin{pmatrix} 1 & -3 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 9 \\ -1 \\ 1 \end{pmatrix}.$$

Left multiply both sides by

$$\begin{pmatrix} 1 & 3 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{to get} \quad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 6 \\ -1 \\ 1 \end{pmatrix},$$

So

$$\begin{aligned} x &= 6, \\ y &= -1, \\ z &= 1, \end{aligned} \qquad \text{or, equivalently,} \quad \mathrm{Sol}(A\mathbf{x} = \mathbf{b}) = \left\{ \begin{pmatrix} 6 \\ -1 \\ 1 \end{pmatrix} \right\}$$

(exactly *one* solution).

Solving problems with an unknown parameter.

Example L11. Find the values of $a, b \in \mathbb{Q}$ for which the system

$$\begin{aligned} x - 2y + z &= 4, \\ 2x - 3y + z &= 7, \\ 3x - 6y + az &= b, \end{aligned} \qquad \text{has} \qquad \begin{array}{l} \text{(a) no solution,} \\ \text{(b) a unique solution,} \\ \text{(c) LOTS of solutions.} \end{array}$$

In matrix form this system is

$$\begin{pmatrix} 3 & -6 & a \\ 2 & -3 & 1 \\ 1 & -2 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} b \\ 7 \\ 4 \end{pmatrix}.$$

Multiply both sides by

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & -2 \end{pmatrix} \quad \text{to get} \quad \begin{pmatrix} 3 & -6 & a \\ 1 & -2 & 1 \\ 0 & 1 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} b \\ 4 \\ -1 \end{pmatrix}.$$

Multiply both sides by

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & -3 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{to get} \quad \begin{pmatrix} 1 & -2 & 1 \\ 0 & 0 & a-3 \\ 0 & 1 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 4 \\ b-12 \\ -1 \end{pmatrix}.$$

Multiply both sides by

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \quad \text{to get} \quad \begin{pmatrix} 1 & -2 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & a-3 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 4 \\ -1 \\ b-12 \end{pmatrix}.$$

Multiply both sides by

$$\begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{to get} \quad \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & a-3 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2 \\ -1 \\ b-12 \end{pmatrix}.$$

Case 1: $a-3\neq 0$. Multiply both sides by

$$\begin{pmatrix}1&0&0\\0&1&0\\0&0&\frac{1}{a-3}\end{pmatrix} \quad \text{to get} \quad \begin{pmatrix}1&0&-1\\0&1&-1\\0&0&1\end{pmatrix}\begin{pmatrix}x\\y\\z\end{pmatrix}=\begin{pmatrix}2\\-1\\\frac{b-12}{a-3}\end{pmatrix}.$$

Multiply both sides by

$$\begin{pmatrix}1&0&0\\0&1&1\\0&0&1\end{pmatrix} \quad \text{to get} \quad \begin{pmatrix}1&0&-1\\0&1&0\\0&0&1\end{pmatrix}\begin{pmatrix}x\\y\\z\end{pmatrix}=\begin{pmatrix}2\\-1+\frac{b-12}{a-3}\\\frac{b-12}{a-3}\end{pmatrix}.$$

Multiply both sides by

$$\begin{pmatrix}1&0&1\\0&1&0\\0&0&1\end{pmatrix} \quad \text{to get} \quad \begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix}\begin{pmatrix}x\\y\\z\end{pmatrix}=\begin{pmatrix}2+\frac{b-12}{a-3}\\-1+\frac{b-12}{a-3}\\\frac{b-12}{a-3}\end{pmatrix}.$$

So

$$\begin{aligned} x &= 2 + \tfrac{b-12}{a-3}, \\ y &= -1 + \tfrac{b-12}{a-3}, \\ z &= \tfrac{b-12}{a-3}, \end{aligned}$$

or, equivalently,

$$\mathrm{Sol}(A\mathbf{x} = \mathbf{b}) = \left\{ \begin{pmatrix} 2 + \frac{b-12}{a-3} \\ -1 + \frac{b-12}{a-3} \\ \frac{b-12}{a-3} \end{pmatrix} \right\} \qquad \text{(exactly } \textit{one} \text{ solution).}$$

Case 2: $a - 3 = 0$. Then

$$\begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2 \\ -1 \\ b-12 \end{pmatrix}.$$

Case 2a: $b - 12 \neq 0$. If $b - 12 \neq 0$ then this system has *no solution*.

Case 2b: $b - 12 = 0$. If $b - 12 = 0$ then this system is

$$\begin{aligned} x - z &= 2, \\ y - z &= -1, \end{aligned} \quad \text{which is} \quad \begin{aligned} x &= 2 + z, \\ y &= -1 + z, \\ z &= 0 + z, \end{aligned}$$

where z can be any number. So

$$\mathrm{Sol}(A\mathbf{x} = \mathbf{b}) = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} + \mathrm{span}\{ \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \},$$

and there are *LOTS of solutions*.

Theorem

If $A \in GL_n(\mathbb{Q})$ then every linear system of the form $A\mathbf{x} = \mathbf{b}$ has a unique solution, given by

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

So, if $A \in GL_n\mathbb{Q})$ then

$$\mathrm{Sol}(A\mathbf{x} = \mathbf{b}) = \{A^{-1}\mathbf{b}\}, \quad \textit{which contains exactly one element.}$$

This is because,

$$\text{left multiplying both sides of} \quad A\mathbf{x} = \mathbf{b} \quad \text{by } A^{-1},$$

gives

$$A^{-1}A\mathbf{x} = A^{-1}\mathbf{b}, \qquad \text{which says} \quad \mathbf{x} = A^{-1}\mathbf{b}.$$

The set of $s \times 1$ matrices with entries in $\mathbb{Q}$ is

$$\mathbb{Q}^s = M_{s\times 1}(\mathbb{Q}).$$

Definition (Kernel and image of a matrix)

Let $A \in M_{t\times s}(\mathbb{Q})$. The *kernel of A* is

$$\ker(A) = \{x \in \mathbb{Q}^s \mid Ax = 0\}$$

and the *image of A* is

$$\mathrm{im}(A) = \{Ax \mid s \in \mathbb{Q}^s\}.$$

Definition (Solutions of a linear system)

Let $A \in M_{t\times s}(\mathbb{Q})$ and let $b \in \mathbb{Q}^s$. The set of *solutions of the linear system $Ax = b$* is

$$\mathrm{Sol}(Ax = b) = \{x \in \mathbb{Q}^s \mid Ax = b\}.$$

Example LS7. If $A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 2 & 2 \\ 0 & 0 & 0 \end{pmatrix}$, $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ and $\mathbf{b} = \begin{pmatrix} -2 \\ 4 \\ -3 \end{pmatrix}$ then $A\mathbf{x} = \mathbf{b}$ is the system

$$\begin{aligned} x_1 + 0x_2 + x_3 &= -2, \\ 0x_1 + 2x_2 + 2x_3 &= 4, \\ 0x_1 + 0x_2 + 0x_3 &= -3, \end{aligned} \qquad \text{which has no solutions}$$

(no choice of $x_1, x_2 x_3 \in \mathbb{Q}$ will satisfy the third equation). So

$$\mathrm{Sol}(A\mathbf{x} = \mathbf{b}) = \emptyset.$$

Then $A\mathbf{x} = 0$ is the system

$$\begin{aligned} x_1 + 0x_2 + x_3 &= 0, \\ 0x_1 + 2x_2 + 2x_3 &= 0, \\ 0x_1 + 0x_2 + 0x_3 &= 0, \end{aligned} \qquad \text{which is} \qquad \begin{aligned} x_1 &= -x_3, \\ x_2 &= x_3, \\ x_3 &= x_3, \end{aligned}$$

where x_3 can be any number.

So

$$\ker(A) = \left\{ x_3 \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix} \;\middle|\; x_3 \in \mathbb{Q} \right\} = \mathbb{Q}\text{-span} \left\{ \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix} \right\}.$$

Then

$$\begin{aligned} \mathrm{im}(A) &= \left\{ \begin{pmatrix} 1 & 0 & 1 \\ 0 & 2 & 2 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \;\middle|\; x_1, x_2, x_3 \in \mathbb{Q} \right\} \\ &= \left\{ x_1 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} \;\middle|\; x_1, x_2, x_3 \in \mathbb{Q} \right\} \\ &= \mathbb{Q}\text{-span} \left\{ \text{columns of } \begin{pmatrix} 1 & 0 & 1 \\ 0 & 2 & 2 \\ 0 & 0 & 0 \end{pmatrix} \right\} \end{aligned}$$

Example LS8. If $A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ and $\mathbf{b} = \begin{pmatrix} -2 \\ 4 \\ 15 \end{pmatrix}$ then

$$\begin{aligned} x_1 + 0x_2 + 0x_3 &= 2, \\ 0x_1 + 1x_2 + 0x_3 &= -4, \\ 0x_1 + 0x_2 + x_3 &= 15, \end{aligned}$$

which has exactly one solution

$$\begin{aligned} x_1 &= 2, \\ x_2 &= -4, \\ x_3 &= 15. \end{aligned}$$

$$\mathrm{Sol}(A\mathbf{x} = \mathbf{b}) = \left\{ \begin{pmatrix} 2 \\ -4 \\ 15 \end{pmatrix} \right\}.$$

Then $A\mathbf{x} = 0$ is the system

$$\begin{aligned} x_1 + 0x_2 + 0x_3 &= 0, \\ 0x_1 + 1x_2 + 0x_3 &= 0, \\ 0x_1 + 0x_2 + x_3 &= 0, \end{aligned}$$

which has exactly one solution

$$\begin{aligned} x_1 &= 0, \\ x_2 &= 0, \\ x_3 &= 0. \end{aligned}$$

So

$$\ker(A) = \left\{ \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \right\}.$$

Then

$$\begin{aligned} \mathrm{im}(A) &= \left\{ \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \;\middle|\; x_1, x_2, x_3 \in \mathbb{Q} \right\} \\ &= \left\{ x_1 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \;\middle|\; x_1, x_2, x_3 \in \mathbb{Q} \right\} \\ &= \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \;\middle|\; x_1, x_2, x_3 \in \mathbb{Q} \right\} = \mathbb{Q}^3. \end{aligned}$$

Example LS9. If

$$A = \begin{pmatrix} 1 & 2 & 0 & 0 & 5 \\ 0 & 0 & 1 & 0 & 6 \\ 0 & 0 & 0 & 1 & 7 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad \text{and} \quad \mathbf{b} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \quad \text{then}$$

$$\begin{aligned} x_1 + 2x_2 + 0x_3 + 0x_4 + 5x_5 &= 1, \\ 0x_1 + 0x_2 + 1x_3 + 0x_4 + 6x_5 &= 2, \\ 0x_1 + 0x_2 + 0x_3 + x_4 + 7x_5 &= 3, \\ 0x_1 + 0x_2 + 0x_3 + 0x_4 + 0x_5 &= 0, \end{aligned}$$

which has an infinite number of solutions.

More specifically,

$$\mathrm{Sol}(A\mathbf{x} = \mathbf{b}) = \left\{ \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} \;\middle|\; \begin{array}{l} x_1 = 1 - 2x_2 - 5x_5, \\ x_2 \in \mathbb{Q}, \\ x_3 = 2 - 6x_5, \\ x_4 = 3 - 7x_5, \\ x_5 \in \mathbb{Q} \end{array} \right\}$$

Equivalently,

$$
\begin{aligned}
\mathrm{Sol}(A\mathbf{x}=\mathbf{b}) &= \left\{ \mathbf{x} = \begin{pmatrix}1\\0\\2\\3\\0\end{pmatrix} + \begin{pmatrix}-2x_2\\x_2\\0\\0\end{pmatrix} + \begin{pmatrix}-5x_5\\0\\-6x_5\\-7x_5\\x_5\end{pmatrix} \;\middle|\; x_2, x_5 \in \mathbb{Q} \right\} \\
&= \begin{pmatrix}1\\0\\2\\3\\0\end{pmatrix} + \left\{ x_2 \begin{pmatrix}-2\\1\\0\\0\\0\end{pmatrix} + x_5 \begin{pmatrix}-5\\0\\-6\\-7\\1\end{pmatrix} \;\middle|\; x_2, x_5 \in \mathbb{Q} \right\} \\
&= \begin{pmatrix}1\\0\\2\\3\\0\end{pmatrix} + \mathbb{Q}\text{-span}\left\{ \begin{pmatrix}-2\\1\\0\\0\\0\end{pmatrix}, \begin{pmatrix}-5\\0\\-6\\-7\\1\end{pmatrix} \right\} \\
&= \mathbf{p} + \ker(A).
\end{aligned}
$$

Theorem (Computing solutions of linear systems)

Let $A \in M_{t\times s}(\mathbb{Q})$ *and let* $b \in \mathbb{Q}^t$*. Then there exist* $P \in GL_t(\mathbb{Q})$ *and* $Q \in GL_s(\mathbb{Q})$ *and* $r \in \{1, \ldots, \min(s,t)\}$ *such that*

$$A = P 1_r Q, \qquad \text{where} \quad 1_r = E_{11} + E_{22} + \cdots + E_{rr}$$

and

$$Sol(Ax = b) = \begin{cases} Q^{-1} \begin{pmatrix} (P^{-1}b)_1 \\ \vdots \\ (P^{-1}b)_r \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \ker(A), & \text{if entries } r{+}1, \ldots, t \text{ of } P^{-1}b \text{ are all } 0, \\ \emptyset, & \text{otherwise.} \end{cases}$$

Lecture 10: Kernel and image of a matrix

Definition (Kernel and image of a matrix)

Let $A \in M_{t\times s}(\mathbb{Q})$.

The *kernel of A* is $\ker(A) = \{x \in \mathbb{Q}^s \mid Ax = 0\}$.

The *image of A* is $\mathrm{im}(A) = \{y \in \mathbb{Q}^t \mid \text{there exists } x \text{ such that } y = Ax\}$.

$$\mathrm{im}(A) = \{Ax \mid x \in \mathbb{R}^s\} = \left\{ \begin{pmatrix} | & & | \\ a_1 & \cdots & a_s \\ | & & | \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_s \end{pmatrix} \;\middle|\; x_1, \ldots, x_s \in \mathbb{R} \right\}$$

$$= \left\{ x_1 \begin{pmatrix} | \\ a_1 \\ | \end{pmatrix} + \cdots + x_s \begin{pmatrix} | \\ a_s \\ | \end{pmatrix} \;\middle|\; x_1, \ldots, x_s \in \mathbb{R} \right\}$$

$$= \mathbb{R}\text{-span}\{\text{columns of } A\}.$$

So $\mathrm{im}(A)$ is the set of linear combinations of the columns of A.

The following Proposition specifies how the kernel and image change if A is multiplied (on the left or the right) by an invertible matrix.

Proposition (How kernel and image change)

Let $A \in M_{t\times s}(\mathbb{Q})$ and $P \in GL_t(\mathbb{Q})$ and $Q \in GL_s(\mathbb{Q})$. Then

$$\ker(PA) = \ker(A), \qquad \ker(AQ) = Q^{-1}\cdot\ker(A),$$
$$im(PA) = P\cdot im(A), \qquad im(AQ) = im(A).$$

$$\begin{aligned}
\ker(PA) &= \{x \in \mathbb{Q}^s \mid PAx = 0\} = \{x \in \mathbb{Q}^s \mid P^{-1}PAx = P^{-1}0\} \\
&= \{x \in \mathbb{Q}^s \mid Ax = 0\} = \ker(A), \\
\ker(AQ) &= \{x \in \mathbb{Q}^s \mid AQx = 0\} = \{Q^{-1}Qx \in \mathbb{Q}^s \mid AQx = 0\} \\
&= Q^{-1}\cdot\{Qx \in \mathbb{Q}^s \mid AQx = 0\} \\
&= Q^{-1}\{y \in \mathbb{Q}^s \mid Ay = 0\} = Q^{-1}\cdot\ker(A), \\
\mathrm{im}(PA) &= \{PAx \mid x \in \mathbb{Q}^s\} = P\{Ax \mid x \in \mathbb{Q}^s\} = P\cdot\mathrm{im}(A), \\
\mathrm{im}(AQ) &= \{AQx \mid x \in \mathbb{Q}^s\} = \{Ay \mid Q^{-1}y \in \mathbb{Q}^s\} \\
&= \{Ay \mid y \in \mathbb{Q}^s\} = \mathrm{im}(A). \qquad \square
\end{aligned}$$

A *subspace of* $\mathbb{Q}^s$ is a subset $W \subseteq \mathbb{Q}^s$ such that

(a) $0 \in W$,

(b) If $w_1, w_2 \in W$ then $w_1 + w_2 \in W$,

(c) If $w \in W$ and $c \in \mathbb{Q}$ then $cw \in W$.

Proposition

Let $A \in M_{t\times s}(\mathbb{Q})$. Then $\ker(A)$ is a subspace of $\mathbb{Q}^s$.

Proof. (a) Since $A0 = 0$ then $0 \in \ker(A)$.

(b) Assume $w_1, w_2 \in \ker(A)$. Then $Aw_1 = 0$ and $Aw_2 = 0$. So

$$A(w_1 + w_2) = Aw_1 + Aw_2 = 0 + 0 = 0. \quad \text{So } w_1 + w_2 \in \ker(A).$$

(c) Assume $w \in \ker(A)$ and $c \in \mathbb{Q}$. Then $Aw = 0$ and

$$A(cw) = cAw = c0 = 0. \quad \text{So } cw \in \ker(A).$$

So $\ker(A)$ is a subspace of $\mathbb{Q}^s$. □

A *subspace* of $\mathbb{Q}^t$ is a subset $Y \subseteq \mathbb{Q}^t$ such that

(a) $0 \in Y$,

(b) If $y_1, y_2 \in Y$ then $y_1 + y_2 \in Y$,

(c) If $y \in Y$ and $c \in \mathbb{Q}$ then $cy \in Y$.

Proposition

Let $A \in M_{t\times s}(\mathbb{Q})$. Then $im(A)$ is a subspace of $\mathbb{Q}^t$.

Proof. (a) Since $0 = A0$ then $0 \in \text{im}(A)$.

(b)Assume $y_1, y_2 \in \text{im}(A)$. Then there exist $x_1, x_2 \in \mathbb{Q}^s$ such that $y_1 = Ax_1$ and $y_2 = Ax_2$. Then

$$y_1 + y_2 = Ax_1 + Ax_2 = A(x_1 + x_2). \quad \text{So } y_1 + y_2 \in \text{im}(A).$$

(c) Assume $y \in \text{im}(A)$ and $c \in \mathbb{Q}$. Then there exists $x \in \mathbb{Q}^s$ such that $y = Ax$. Then

$$cy = cAx = A(cx). \quad \text{So } cy \in \text{im}(A).$$

So $\text{im}(A)$ is a subspace of $\mathbb{Q}^t$. □

Let W be a subspace of $\mathbb{Q}^s$. A set $B = \{b_1, \ldots, b_k\}$ is *a basis of W* if every element of W is

a unique linear combination of $b_1, \ldots, b_k$.

Let Y be a subspace of $\mathbb{Q}^t$. A set $D = \{d_1, \ldots, b_\ell\}$ is *a basis of Y* if every element of Y is

a unique linear combination of $d_1, \ldots, d_\ell$.

A set $B = \{b_1, \ldots, b_k\}$ is *a basis of* $\ker(A)$ if every element of $\ker(A)$ is

a unique linear combination of $b_1, \ldots, b_k$.

A set $D = \{d_1, \ldots, d_\ell\}$ is *a basis of im*(A) if every element of $\mathrm{im}(A)$ is

a unique linear combination of $d_1, \ldots, d_\ell$.

Let $t, s \in \mathbb{Z}_{>0}$ and let E_{ij} be the $t \times s$ matrix with 1 in the (i,j) entry and 0 elsewhere. Let $r \in \{1, \ldots, \min(s,t)\}$ and let

$$1_r = E_{11} + \cdots + E_{rr}.$$

Let $e_1, \ldots, e_s$ be the standard basis of $\mathbb{Q}^s$. Then

$$\{e_{r+1}, \ldots, e_s\} \quad \text{is a basis of } \ker(1_r).$$

If $Q \in GL_s(\mathbb{Q})$ is invertible then

$$\{Q^{-1}e_{r+1}, \ldots, Q^{-1}e_s\} \quad \text{is a basis of} \quad Q^{-1}\ker(1_r).$$

For example, if $s = 5$ and $t = 6$ and $r = 2$ then

$$1_2 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad \text{and} \quad \ker(1_2) = \text{span}\left\{ \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right\}.$$

Let $t, s \in \mathbb{Z}_{>0}$ and let E_{ij} be the $t \times s$ matrix with 1 in the (i, j) entry and 0 elsewhere. Let $r \in \{1, \ldots, \min(s, t)\}$ and let

$$1_r = E_{11} + \cdots + E_{rr}.$$

Let $e_1, \ldots, e_t$ be the standard basis of $\mathbb{Q}^t$. Then

$$\{e_1, \ldots, e_r\} \quad \text{is a basis of} \quad \text{im}(1_r).$$

If $P \in GL_t(\mathbb{Q})$ is invertible then

$$\{Pe_1, \ldots, Pe_r\} \quad \text{is a basis of} \quad P\text{im}(1_r).$$

For example, if $s = 5$ and $t = 6$ and $r = 2$ then

$$1_2 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad \text{and} \quad \text{im}(1_2) = \text{span}\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \right\}.$$

Theorem (Computing kernels and images)

Let $P \in GL_t(\mathbb{Q})$, $Q \in GL_s(\mathbb{Q})$. *Let* $r \in \{1, \ldots, \min(s,t)\}$ *and let*

$$1_r = E_{11} + E_{22} + \cdots + E_{rr} \quad \text{in} \quad M_{t\times s}(\mathbb{Q}).$$

Let

$$A = P1_rQ.$$

Then

$$\ker(A) \text{ has basis } \{\text{last } s-r \text{ columns of } Q^{-1}\},$$

$$im(A) \text{ has basis } \{\text{first } r \text{ columns of } P\}.$$

Proof. By the How Kernel and Image Change Proposition

$$\ker(A) = \ker(P1_rQ) = \ker(1_rQ) = Q^{-1}\ker(1_r) \quad \text{and}$$
$$\mathrm{im}(A) = \mathrm{im}(P1_rQ) = \mathrm{im}(P1_r) = P\mathrm{im}(1_r)$$

Since $\{Q^{-1}e_{r+1}, \ldots, Q^{-1}e_s\}$ is a basis of $Q^{-1}\ker(1_r)$ then

$$\ker(A) \text{ has basis } \{\text{last } s-r \text{ columns of } Q^{-1}\},$$

Since $\{Pe_1, \ldots, Pe_r\}$ is a basis of $P\text{im}(1_r)$ then

$$\text{im}(A) \text{ has basis } \{\text{first } r \text{ columns of } P\}. \qquad \square$$

Let $A \in M_{t\times s}(\mathbb{Q})$. By definition,

$\dim(\ker(A))$ is the number of elements in a basis of $\ker(A)$, and
$\dim(\text{im}(A))$ is the number of elements in a basis of $\text{im}(A)$.

From the Computing Kernels and Images Theorem,

$$\dim(\ker(A)) = s - r \quad \text{and} \quad \dim(\text{im}(A)) = r = \text{rank}(A).$$

Corollary (rank-nullity theorem)

Let $A \in M_{t\times s}(\mathbb{Q})$. Since $(s-r)+r = s$ then

$$\dim(\mathit{im}(A)) + \dim(\ker(A)) = (\textit{number of columns of } A).$$

The following proposition shows that every invertible matrix is square and has kernel equal to $\{0\}$.

Proposition (Invertible matrices are square)

Let $A \in M_{t\times s}(\mathbb{Q})$ and let $r = rank(A)$. Assume there exists $B \in M_{s\times t}(\mathbb{Q})$ such that $AB = 1$ and $BA = 1$. Then

$$\ker(A) = 0, \quad im(A) = \mathbb{Q}^t \quad and \quad r = s = t.$$

Proof. (a) Assume $Ax = 0$. Then

$$x = BAx = B0 = 0. \qquad \text{So } \ker(A) = \{0\}.$$

(b) If $y \in \mathbb{Q}^t$ then $y = ABy = A(By)$ and $y \in \mathrm{im}(A)$. So

$$\mathrm{Im}(A) = \mathbb{Q}^t.$$

(c) Let $P \in GL_t(\mathbb{Q})$ and $Q \in GL_s(\mathbb{Q})$ be such that $A = P1_rQ$.
Since $\ker(A) = 0$ and $\ker(A) = \mathrm{span}\{Q^{-1}e_{r+1}, \ldots, Q^{-1}s\}$ then $r = s$.
Since $r = s$ and $\mathbb{Q}^t = \mathrm{Im}(A)$ has basis $\{Pe_1, \ldots, Pe_s\}$ then

$$\mathbb{Q}^t = P^{-1}\mathbb{Q}^t \text{ has basis } \{e_1, \ldots, e_s\}. \quad \text{So } r = s = t. \qquad \square$$

Example V27&28. Let

$$S = \{|1,3,-1,1\rangle,\ |2,6,0,4\rangle,\ |3,9,-2,4\rangle\ \}.$$

Then

$$S = \left\{ \begin{pmatrix} 1 \\ 3 \\ -1 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 6 \\ 0 \\ 4 \end{pmatrix}, \begin{pmatrix} 3 \\ 9 \\ -2 \\ 4 \end{pmatrix} \right\}$$

and

$$\mathbb{R}\text{-span}(S) = \mathrm{im}(A),\ \text{where } A = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 6 & 0 \\ -1 & 0 & -2 \\ 1 & 4 & 4 \end{pmatrix}.$$

Now

$$A = \begin{pmatrix} 1 & 2 & 0 & 0 \\ 3 & 6 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ 1 & 4 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & \frac{1}{2} \\ 0 & 0 & 1 \end{pmatrix} = P 1_2 Q,$$

where

$$P = \begin{pmatrix} 1 & 2 & 0 & 0 \\ 3 & 6 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ 1 & 4 & 0 & 1 \end{pmatrix}, \quad 1_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad Q = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & \frac{1}{2} \\ 0 & 0 & 1 \end{pmatrix}.$$

Then

$$\mathrm{im}(A) = \mathrm{im}(P 1_2 Q) = P\mathrm{im}(1_2) = \mathrm{span}\left\{ P \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, P \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \right\}.$$

Thus

$$\mathrm{im}(A) \quad \text{has basis} \quad \left\{ \begin{pmatrix} 1 \\ 3 \\ -1 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 6 \\ 0 \\ 4 \end{pmatrix} \right\}.$$

Since

$$\ker(A) = \ker(P 1_2 Q) = \ker(1_2 Q) = Q^{-1} \ker(1_2)$$

$$\text{and} \quad Q^{-1} = \begin{pmatrix} 1 & 0 & -2 \\ 0 & 1 & -\frac{1}{2} \\ 0 & 0 & 1 \end{pmatrix} \quad \text{then} \quad \ker(A) = \mathbb{R}\text{-span}\left\{ \begin{pmatrix} -2 \\ \frac{1}{2} \\ 1 \end{pmatrix} \right\}.$$

Thus $\ker(A)$ has basis $\{|-2, \frac{1}{2}, 1\rangle\}$.

Lecture 11: Eigenvalues and eigenvectors

Definition (Eigenvectors and eigenvalues.)

Let $A \in M_n(\mathbb{Q})$.

- An *eigenvalue of A* is an element $\lambda \in \mathbb{Q}$ such that $\ker(A - \lambda) \neq 0$.

Let $A \in M_n(\mathbb{Q})$ and $\lambda \in \mathbb{Q}$.

- An *eigenvector of A of eigenvalue λ* is a nonzero element of $\ker(A - \lambda)$.

If v is an eigenvector of A of eigenvalue λ then

$$(A - \lambda)v = 0 \quad \text{and} \quad Av = \lambda v.$$

Definition (Linearly independent eigenvectors)

Let $A \in M_n(\mathbb{Q})$ and let $p_1, \dots, p_k$ be eigenvectors of A. The set $\{p_1, \dots, p_k\}$ is *linearly independent* if $p_1, \dots, p_k$ satisfy the condition

$$\text{if } c_1, \dots, c_k \in \mathbb{Q} \text{ and } c_1 p_1 + \cdots + c_k p_k = 0$$

$$\text{then } c_1 = 0 \text{ and } c_2 = 0 \text{ and } \dots \text{ and } c_k = 0.$$

Theorem (Diagonalization.)

Let $A \in M_{n\times n}(\mathbb{F})$. The matrix A has n linearly independent eigenvectors $\mathbf{p}_1, \dots, \mathbf{p}_n \in \mathbb{F}^n$ with eigenvalues $\lambda_1, \dots, \lambda_n$ if and only if $A = PDP^{-1}$ where,

$$P = \begin{pmatrix} | & & | \\ \mathbf{p}_1 & \cdots & \mathbf{p}_n \\ | & & | \end{pmatrix} \quad \textit{and} \quad D = \mathrm{diag}(\lambda_1, \dots, \lambda_n) = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$$

so that $\mathbf{p}_1, \dots, \mathbf{p}_n$ are the columns of P and D is the diagonal matrix with diagonal entries $\lambda_1, \dots, \lambda_n$.

Example EV2,6&9. Find the eigenvalues of $A = \begin{pmatrix} 1 & 4 \\ 1 & 1 \end{pmatrix}$.

First,

$$A - t = \begin{pmatrix} 1-t & 4 \\ 1 & 1-t \end{pmatrix} = \begin{pmatrix} 1-t & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1-t \\ 0 & 4-(1-t)^2 \end{pmatrix}$$
$$= \begin{pmatrix} 1-t & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1-t \\ 0 & -(t+1)(t-3) \end{pmatrix}$$

Case 1: $t+1=0$. Then

$$A+1 = \begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad \ker(A+1) = \text{span}\left\{ \begin{pmatrix} -2 \\ 1 \end{pmatrix} \right\}$$

Case 1: $t-3=0$. Then

$$A-3 = \begin{pmatrix} -2 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & -2 \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad \ker(A-3) = \text{span}\left\{ \begin{pmatrix} 2 \\ 1 \end{pmatrix} \right\}.$$

$$\text{If} \quad P = \begin{pmatrix} -2 & 2 \\ 1 & 1 \end{pmatrix} \text{ and } D = \begin{pmatrix} -1 & 0 \\ 0 & 3 \end{pmatrix} \text{ then } P^{-1} = \frac{1}{4} \begin{pmatrix} -1 & 2 \\ 1 & 2 \end{pmatrix}$$

$$\begin{aligned} \text{and} \quad PDP^{-1} &= \begin{pmatrix} -2 & 2 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} -1 & 0 \\ 0 & 3 \end{pmatrix} \cdot \tfrac{1}{4} \begin{pmatrix} -1 & 2 \\ 1 & 2 \end{pmatrix} \\ &= \tfrac{1}{4} \begin{pmatrix} -2 & 2 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & -2 \\ 3 & 6 \end{pmatrix} \\ &= \tfrac{1}{4} \begin{pmatrix} 4 & 16 \\ 4 & 4 \end{pmatrix} = \begin{pmatrix} 1 & 4 \\ 1 & 1 \end{pmatrix} = A. \end{aligned}$$

The characteristic polynomial of A is

$$\det(A - t) = \det(D - t) = (-1 - t)(3 - t).$$

Example EV3,4&10. Find the eigenvalues of $A = \begin{pmatrix} 2 & -3 & 6 \\ 0 & 5 & -6 \\ 0 & 1 & 0 \end{pmatrix}$.

Find $\ker(A - t)$ by row reduction:

$$\begin{aligned}
A - t &= \begin{pmatrix} 2-t & -3 & 6 \\ 0 & 5-t & -6 \\ 0 & 1 & 0-t \end{pmatrix} \\
&= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 5-t & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 2-t & -3 & 6 \\ 0 & 1 & -t \\ 0 & 0 & -6-(5-t)(-t) \end{pmatrix} \\
&= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 5-t & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 3 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 2-t & 0 & 3(2-t) \\ 0 & 1 & -t \\ 0 & 0 & -(t^2-5t+6) \end{pmatrix} \\
&= \begin{pmatrix} 1 & 3 & 0 \\ 0 & 5-t & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 2-t & 0 & 3(2-t) \\ 0 & 1 & -t \\ 0 & 0 & -(t-2)(t-3) \end{pmatrix}.
\end{aligned}$$

Case 1: $t - 2 = 0$. Then

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

and

$$\ker(A-2) = \ker \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \end{pmatrix} = \operatorname{span}\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix} \right\}.$$

Case 2: $t - 3 = 0$. Then

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

and

$$\ker(A-3) = \ker \begin{pmatrix} -1 & 0 & -3 \\ 0 & 1 & -3 \\ 0 & 0 & 0 \end{pmatrix} = \operatorname{span}\left\{ \begin{pmatrix} -3 \\ 3 \\ 1 \end{pmatrix} \right\}.$$

Then $A = PDP^{-1}$ where

$$P = \begin{pmatrix} 1 & 0 & -3 \\ 0 & 2 & 3 \\ 0 & 1 & 1 \end{pmatrix}, \quad D = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix},$$

$$\text{and} \quad P^{-1} = -\begin{pmatrix} -1 & 0 & 0 \\ -3 & 1 & -1 \\ 6 & -3 & 2 \end{pmatrix}^t = \begin{pmatrix} 1 & 3 & -6 \\ 0 & -1 & 3 \\ 0 & 1 & -2 \end{pmatrix}.$$

The characteristic polynomial of A is

$$\det(A-t) = \det(D-t) = (2-t)^2(3-t).$$

Example EV5,8&12. As an element of $M_{2\times 2}(\mathbb{R})$, the matrix

$$A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \quad \text{has no eigenvalues and no eigenvectors.}$$

The linear transformation

$$\begin{array}{rccc} T: & \mathbb{R}^2 & \to & \mathbb{R}^2 \\ & v & \mapsto & Av \end{array} \quad \text{is a rotation of } \tfrac{3\pi}{2} \text{ about } (0,0).$$

As an element of $M_{2\times 2}(\mathbb{C})$, the matrix

$$A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \quad \text{has two eigenvalues, } i \text{ and } -i.$$

$$\ker(A - i) = \mathrm{span}\left\{ \begin{pmatrix} i \\ 1 \end{pmatrix} \right\} \qquad \text{and} \qquad \ker(A + i) = \mathrm{span}\left\{ \begin{pmatrix} -i \\ 1 \end{pmatrix} \right\}.$$

$\begin{pmatrix} i \\ 1 \end{pmatrix}$ is an eigenvector of eigenvalue i and
$\begin{pmatrix} -i \\ 1 \end{pmatrix}$ is an eigenvector of eigenvalue $-i$.

If

$$P = \begin{pmatrix} -i & i \\ 1 & 1 \end{pmatrix} \text{ and } D = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} \text{ then } P^{-1} = \frac{1}{-2i} \begin{pmatrix} 1 & -i \\ -1 & -i \end{pmatrix}$$

and

$$\begin{aligned} PDP^{-1} &= \begin{pmatrix} -i & i \\ 1 & 1 \end{pmatrix} \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} \cdot \frac{1}{-2i} \begin{pmatrix} 1 & -i \\ -1 & -i \end{pmatrix} \\ &= \tfrac{1}{2} i \begin{pmatrix} -i & i \\ 1 & 1 \end{pmatrix} \begin{pmatrix} i & 1 \\ i & -1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = A. \end{aligned}$$

The characteristic polynomial of A is

$$\det(A - t) = \det(D - t) = (i - t)(-i - t) = t^2 + 1.$$

Example EV11. If $PDP^{-1} = A$ then the columns of P are linearly independent eigenvectors of A. Here is an example where A does not have n linearly independent eigenvectors.

$$\text{If} \qquad A = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} \qquad \text{then} \qquad A - t = \begin{pmatrix} 1-t & 2 \\ 0 & 1-t \end{pmatrix}$$

which has a single row of 0s when $t = 1$.
(The characteristic polynomial of A is $\det(A - t) = (1-t)^2$.)

$$\ker(A - 1) = \text{span}\left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\}.$$

Since A does not have two linearly independent eigenvectors then

A is not diagonalizable.

Lecture 12: Symmetric, Hermitian, unitary and orthogonal matrices

Definition (Transpose of a matrix)

Let $A \in M_{t\times s}(\mathbb{Q})$. The *transpose of A* is $A^T \in M_{s\times t}(\mathbb{Q})$ given by

$$(A^T)_{ij} = A_{ji}, \quad \text{for } i \in \{1,\dots,s\} \text{ and } j \in \{1,\dots,t\}.$$

Example M4. If $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$ then $A^T = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}$.

Definition (Symmetric, Hermitian, Unitary, Orthogonal matrices.)

A *symmetric matrix* is $A \in M_{n\times n}(\mathbb{C})$ such that $A = A^T$.

An *orthogonal matrix* is $A \in M_{n\times n}(\mathbb{C})$ such that $AA^T = 1$.

A *Hermitian matrix* is $A \in M_{n\times n}(\mathbb{C})$ such that $A = \overline{A}^T$.

A *unitary matrix* is $A \in M_{n\times n}(\mathbb{C})$ such that $A\overline{A}^T = 1$.

Example IP22. Let $A \begin{pmatrix} 1 & i \\ -i & 1 \end{pmatrix}$ and $B = \begin{pmatrix} -i & 0 \\ 0 & -i \end{pmatrix}$. Since

$$A^* = \bar{A}^T = \begin{pmatrix} 1 & i \\ -i & 1 \end{pmatrix} = A \quad \text{and} \quad B^* = \bar{B}^T = \begin{pmatrix} -i & 0 \\ 0 & -i \end{pmatrix} \neq B$$

then A is Hermitian and B is not Hermitian.

Example IP21. The matrix $U = \frac{1}{\sqrt{2}} \begin{pmatrix} -i & i \\ 1 & 1 \end{pmatrix}$ is unitary since

$$UU^* = \frac{1}{\sqrt{2}} \begin{pmatrix} -i & i \\ 1 & 1 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} i & 1 \\ -i & 1 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Example IP15. $Q = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$ is orthogonal since

$$\begin{aligned} QQ^T &= \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \\ &= \begin{pmatrix} \cos^2\theta + \sin^2\theta & 0 \\ 0 & \cos^2\theta + \sin^2\theta \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \end{aligned}$$

Definition (The general linear group)

The *general linear group* $GL_n(\mathbb{R})$ is the set

$$GL_n(\mathbb{R}) = \left\{ A \in M_{n\times n}(\mathbb{R}) \;\middle|\; \begin{array}{l} \text{there exists } A^{-1} \in M_{n\times n}(\mathbb{R}) \\ \text{such that } AA^{-1} = 1 \text{ and } A^{-1}A = 1 \end{array} \right\}$$

Definition (The orthogonal and unitary groups.)

The *orthogonal group* $O_n(\mathbb{R})$ is the set

$$O_n(\mathbb{R}) = \{A \in M_{n\times n}(\mathbb{R}) \mid AA^T = 1\}.$$

The *unitary group* $U_n(\mathbb{C})$ is the set

$$U_n(\mathbb{C}) = \{A \in M_{n\times n}(\mathbb{C}) \mid A\overline{A}^T = 1\}.$$

Example IP17. Assume $Q \in O_n(\mathbb{R})$. Then $1 = QQ^T$ and

$$1 = \det(1) = \det(QQ^T) = \det(Q)\det(Q^T) = \det(Q)\det(Q) = \det(Q)^2.$$

So $\det(Q) \in \{1, -1\}$.

Definition (Standard inner products on $\mathbb{R}^n$ and $\mathbb{C}^n$)

(a) The *standard inner product on $\mathbb{R}^n$* is $\langle , \rangle \colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ given by

$$\langle \mathbf{x}, \mathbf{y} \rangle = x_1 y_1 + \cdots + x_n y_n,$$

if $\mathbf{x} = |x_1, \ldots, x_n\rangle$ and $\mathbf{y} = |y_1, \ldots, y_n\rangle$.

(b) The *standard inner product on $\mathbb{C}^n$* is $\langle , \rangle \colon \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C}^n$ given by

$$\langle \mathbf{x}, \mathbf{y} \rangle = x_1 \overline{y_1} + \cdots + x_n \overline{y_n},$$

if $\mathbf{x} = |x_1, \ldots, x_n\rangle$ and $\mathbf{y} = |y_1, \ldots, y_n\rangle$.

Example IP16. Let $u, v \in \mathbb{R}^n$ and let $Q \in O_n(\mathbb{R})$. Then

$$\langle u|v\rangle = u^T v \quad \text{and}$$
$$\langle Qu|Qv\rangle = (Qu)^T Qv = u^T Q^T Qv = u^T \cdot 1 \cdot v = u^T v.$$

So $\langle Qu|Qv\rangle = \langle u|v\rangle$.

Definition (Orthonormal basis of $\mathbb{R}^n$ and of $\mathbb{C}^n$)

A *basis of* $\mathbb{R}^n$ is a subset $\{b_1, \ldots, b_n\}$ of $\mathbb{R}^n$ such that

every vector in $\mathbb{R}^n$ is a unique $\mathbb{R}$-linear combination of $b_1, \ldots, b_n$.

A *basis of* $\mathbb{C}^n$ is a subset $\{b_1, \ldots, b_n\}$ of $\mathbb{C}^n$ such that

every vector in $\mathbb{C}^n$ is a unique $\mathbb{C}$-linear combination of $b_1, \ldots, b_n$.

An *orthonormal basis of* $\mathbb{R}^n$ is a basis of $\{b_1, \ldots, b_n\}$ of $\mathbb{R}^n$ such that

$$\text{if } i, j \in \{1, \ldots, n\} \quad \text{then} \quad \langle b_i, b_j \rangle = \delta_{ij},$$

where $\langle , \rangle$ is the standard inner product on $\mathbb{R}^n$.

An *orthonormal basis of* $\mathbb{C}^n$ is a basis of $\{b_1, \ldots, b_n\}$ of $\mathbb{C}^n$ such that

$$\text{if } i, j \in \{1, \ldots, n\} \quad \text{then} \quad \langle b_i, b_j \rangle = \delta_{ij},$$

where $\langle , \rangle$ is the standard inner product on $\mathbb{C}^n$.

Theorem

Let $A \in M_{n\times n}(\mathbb{R})$. Then $A \in GL_n(\mathbb{R})$ if and only if the columns of A form a basis of $\mathbb{R}^n$.

Theorem (Diagonalization)

Let $A \in M_{n\times n}(\mathbb{F})$. The matrix A has n linearly independent eigenvectors $\mathbf{p}_1, \ldots, \mathbf{p}_n \in \mathbb{F}^n$ with eigenvalues $\lambda_1, \ldots, \lambda_n$ if and only if $A = PDP^{-1}$ where,

$$P = \begin{pmatrix} | & & | \\ \mathbf{p}_1 & \cdots & \mathbf{p}_n \\ | & & | \end{pmatrix} \quad \textit{and} \quad D = \operatorname{diag}(\lambda_1, \ldots, \lambda_n) = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$$

so that $\mathbf{p}_1, \ldots, \mathbf{p}_n$ are the columns of P and D is the diagonal matrix with diagonal entries $\lambda_1, \ldots, \lambda_n$.

Theorem

Let $A \in M_{n\times n}(\mathbb{C})$*. Then* $A \in U_n(\mathbb{C})$ *if and only if the columns of* A *form an orthonormal basis of* $\mathbb{C}^n$ *with respect to the standard inner product on* $\mathbb{C}^n$*.*

Theorem (Hermitian diagonalization)

Let $A \in M_{n\times n}(\mathbb{C})$ *be a Hermitian matrix. If* $\mathbf{p}_1, \ldots, \mathbf{p}_n \in \mathbb{C}^n$ *are orthonormal eigenvectors for* A *with eigenvalues* $\lambda_1, \ldots, \lambda_n$ *and*

$$P = \begin{pmatrix} | & & | \\ \mathbf{p}_1 & \cdots & \mathbf{p}_n \\ | & & | \end{pmatrix} \quad \text{and} \quad D = \operatorname{diag}(\lambda_1, \ldots, \lambda_n) = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$$

then P *is unitary and* $A = PD\overline{P}^T$*.*

Theorem

Let $A \in M_{n\times n}(\mathbb{R})$. Then $A \in O_n(\mathbb{R})$ if and only if the columns of A form an orthonormal basis of $\mathbb{R}^n$ with respect to the standard inner product on $\mathbb{R}^n$.

Theorem (Real symmetric diagonalization)

Let $A \in M_{n\times n}(\mathbb{R})$ be a symmetric matrix. If $\mathbf{p}_1, \ldots, \mathbf{p}_n \in \mathbb{R}^n$ are orthonormal eigenvectors for A with eigenvalues $\lambda_1, \ldots, \lambda_n$ and

$$P = \begin{pmatrix} | & & | \\ \mathbf{p}_1 & \cdots & \mathbf{p}_n \\ | & & | \end{pmatrix} \quad \textit{and} \quad D = \operatorname{diag}(\lambda_1, \ldots, \lambda_n) = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$$

then P is orthogonal and $A = PD\overline{P}^T$.

Example IP18. The characteristic polynomial of the symmetric matrix

$$A = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \quad \text{is} \quad \begin{aligned} \det(A-t) &= (1-t)^2 - 1 \\ &= 1 - 2t + t^2 - 1 = t^2 - 2t \\ &= (t-0)(t-2). \end{aligned}$$

Then

$$\tfrac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \text{and} \quad \tfrac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

are eigenvectors of length 1 with eigenvalues 0 and 2, respectively. Then

$$Q = \tfrac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \quad \text{is orthogonal}$$

and

$$A = Q\begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix}Q^T.$$

Example IP23. The characteristic polynomial of the Hermitian matrix

$$A = \begin{pmatrix} 1 & i \\ -i & 1 \end{pmatrix} \quad \text{is} \quad \begin{aligned} \det(A - t) &= (1 - t)^2 - (-i) \cdot i \\ &= 1 - 2t + t^2 - 1 = t^2 - 2t \\ &= (t - 0)(t - 2). \end{aligned}$$

Then

$$\tfrac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \end{pmatrix} \quad \text{and} \quad \tfrac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -i \end{pmatrix}$$

are eigenvectors of A of length 1 with eigenvalues 0 and 2, respectively. Then

$$U = \tfrac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix} \quad \text{is unitary}$$

and

$$A = U \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix} \bar{U}^T.$$

Let $t, s \in \mathbb{Z}_{>0}$ and let E_{ij} be the $t \times s$ matrix with 1 in the (i, j) entry and 0 elsewhere.
Let $A \in M_{t\times s}(\mathbb{R})$. Find orthonormal eigenvectors $v_1, \ldots, v_s$ of $A^t A$ with eigenvalues $\lambda_1, \ldots, \lambda_s$ and let

$$V = \begin{pmatrix} | & & | \\ v_1 & \cdots & v_s \\ | & & | \end{pmatrix} \quad \text{and} \quad S = \sqrt{\lambda_1}E_{11} + \cdots + \sqrt{\lambda_s}E_{ss}.$$

$$\text{If } \lambda_i \neq 0 \text{ let } u_i = \frac{1}{\sqrt{\lambda_i}} A v_i.$$

Extend $u_1, \ldots, u_k$ to an orthonormal basis of $\mathbb{R}^t$ and let

$$U = \begin{pmatrix} | & & | \\ u_1 & \cdots & u_t \\ | & & | \end{pmatrix}.$$

Then $U \in O_t(\mathbb{R})$ and $V \in O_s(\mathbb{R})$ and $S \in M_{t\times s}(\mathbb{R})$ is 'diagonal' and

$$A = USV^T.$$

Example IP20. If

$$A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \quad \text{then} \quad A^T A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}$$

The columns v_1, v_2, v_3 of

$$V = \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{-1}{\sqrt{2}} \\ 0 & 1 & 0 \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{pmatrix} \quad \text{are orthonormal eigenvectors of } A^T A$$

with eigenvalues $\lambda = 2$, $\lambda_2 = 1$, $\lambda_3 = 0$. Let

$$u_1 = \frac{1}{\sqrt{2}} A v_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{and} \quad u_2 = \frac{1}{\sqrt{1}} A v_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Then $\{u_1, u_2\}$ is already an orthonormal basis of $\mathbb{R}^2$, Let

$$U = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad S = \begin{pmatrix} \sqrt{2} & 0 & 0 \\ 0 & \sqrt{1} & 0 \end{pmatrix}$$

Then $U \in O_2(\mathbb{R})$, $V \in O_3(\mathbb{R})$ and $S \in M_{2\times 3}(\mathbb{R})$ and $A = USV^T$.

Example IP19. If

$$A = \begin{pmatrix} 0 & -1 \\ 0 & 0 \end{pmatrix} \quad \text{then} \quad A^T A = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

The columns v_1, v_2 of

$$V = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \text{are orthonormal eigenvectors of } A^T A$$

with eigenvalues $\lambda_1 = 1$, $\lambda_2 = 0$. Let

$$u_1 = \frac{1}{\sqrt{1}} A v_1 = \begin{pmatrix} -1 \\ 0 \end{pmatrix} \quad \text{and let} \quad u_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

so that $\{u_1, u_2\}$ is an orthonormal basis of $\mathbb{R}^2$, Let

$$U = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad S = \begin{pmatrix} \sqrt{1} & 0 \\ 0 & \sqrt{0} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

Then $U \in O_2(\mathbb{R})$, $V \in O_2(\mathbb{R})$ and $S \in M_{2\times 2}(\mathbb{R})$ and $A = USV^T$.

Lecture 14: Traces and Determinants

Let $n \in \mathbb{Z}_{>0}$. Let E_{ij} be the $n \times n$ matrix with 1 in the (i,j) entry and 0 elsewhere. For $i \in \{1, \ldots, n-1\}$ and $c \in \mathbb{Q}$ define

$$s_i(c) = 1 - E_{ii} - E_{i+1,i+1} + E_{i,i+1} + E_{i+1,i} + cE_{ii}.$$

For $i \in \{1, \ldots, n\}$ and $d \in \mathbb{Q}$ with $d \neq 0$ define

$$h_i(d) = 1 + (d-1)E_{ii},$$

For $i, j \in \{1, \ldots, n\}$ with $i \neq j$ and $c \in \mathbb{Q}$ define

$$x_{ij}(c) = 1 + cE_{ij},$$

For $r \in \{1, \ldots, n\}$ define

$$1_r = E_{11} + \cdots + E_{rr}.$$

Definition (Determinant)

The *deteminant* is the function $\det\colon M_{n\times n}(\mathbb{Q}) \to \mathbb{Q}$ determined by

$$\text{if } A, B \in M_{n\times n}(\mathbb{Q}) \text{ then} \quad \det(AB) = \det(A)\det(B),$$

and if $i, j \in \{1, \dots, n\}$ with $i \neq j$ and $k \in \{1, \dots, n-1\}$ and $c, d \in \mathbb{Q}$ with $d \neq 0$ then

$$\det(x_{ij}(c)) = 1, \quad \det(h_i(d)) = d, \quad \det(s_k(c)) = -1.$$

Theorem (Factoring)

Let $n \in \mathbb{Z}_{>0}$. Let

$$1_r = E_{11} + \cdots + E_{rr} \qquad \text{in } M_{n\times n}(\mathbb{Q}).$$

Let $A \in M_{n\times n}(\mathbb{Q})$. The factoring algorithm gives

$$A = (\text{product of } s_i(c)\text{s}) \cdot (\text{product of } h_i(d)\text{s}) \cdot (\text{product of } x_{ij}(c)\text{s}) \cdot 1_r \cdot (\text{product of } s_i(c)\text{s}) \cdot (\text{product of } x_{ij}(c)\text{s}).$$

Example M12. Let $A = \begin{pmatrix} 2 & 6 & 9 \\ 0 & 3 & 8 \\ 0 & 0 & -1 \end{pmatrix}$. Then

$$\begin{aligned} A &= h_1(2)h_2(3)h_3(-1) \begin{pmatrix} 1 & 3 & \frac{9}{2} \\ 0 & 1 & \frac{8}{3} \\ 0 & 0 & 1 \end{pmatrix} \\ &= h_1(2)h_2(3)h_3(-1)x_{12}(3)x_{13}(\tfrac{9}{2})x_{23}(\tfrac{8}{3}), \end{aligned}$$

So

$$\begin{aligned} \det(A) &= \det\left(h_1(2)h_2(3)h_3(-1)x_{12}(3))x_{13}(\tfrac{9}{2})x_{23}(\tfrac{8}{3}\right) \\ &= \det(h_1(2))\det(h_2(3))\det(h_3(-1)) \\ &\qquad \cdot \det(x_{12}(3)))\det(x_{13}(\tfrac{9}{2}))\det(x_{23}(\tfrac{8}{3})) \\ &= 2 \cdot 3 \cdot (-1) \cdot 1 \cdot 1 \cdot 1 = -6. \end{aligned}$$

Example M13. Let $A = \begin{pmatrix} 1 & 2 & 1 \\ -1 & 1 & 1 \\ 0 & 1 & 3 \end{pmatrix}$. Then

$$\begin{aligned} A &= \begin{pmatrix} -1 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} -1 & 1 & 1 \\ 0 & 3 & 2 \\ 0 & 1 & 3 \end{pmatrix} \\ &= \begin{pmatrix} -1 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} -1 & 1 & 1 \\ 0 & 1 & 3 \\ 0 & 0 & -7 \end{pmatrix} \\ &= s_1(-1)s_2(3) \begin{pmatrix} -1 & 1 & 1 \\ 0 & 1 & 3 \\ 0 & 0 & -7 \end{pmatrix}. \end{aligned}$$

So $\det(A) = (-1) \cdot (-1) \cdot (-1) \cdot 1 \cdot (-7) = 7$.

Definition (Trace)

Let $A \in M_{n\times n}(\mathbb{Q})$. The *trace of A* is

$$\operatorname{Tr}(A) = A_{11} + \cdots + A_{nn} \quad \text{where} \quad A = \begin{pmatrix} A_{11} & \cdots & A_{1n} \\ \vdots & & \vdots \\ A_{n1} & \cdots & A_{nn} \end{pmatrix}$$

Example M3.

$$\operatorname{Tr}\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} = 1 + 5 + 9 = 15.$$

Proposition (Properties of trace)

Let $A, B \in M_{n\times n}(\mathbb{Q})$ and let $c \in \mathbb{Q}$. Then

$$\operatorname{Tr}(A+B) = \operatorname{Tr}(A) + \operatorname{Tr}(B), \qquad \operatorname{Tr}(cA) = c\operatorname{Tr}(A)$$

and

$$\operatorname{Tr}(AB) = \operatorname{Tr}(BA).$$

Theorem (Determinant and trace are conjugacy invariants)

Let $n \in \mathbb{Z}_{>0}$. Let $A \in M_{n\times n}(\mathbb{Q})$ and let $P \in GL_n(\mathbb{Q})$ so that P is an invertible $n \times n$ matrix. Then

$$\det(PAP^{-1}) = \det(A) \quad \textit{and} \quad \mathrm{Tr}(PAP^{-1}) = \mathrm{Tr}(A).$$

Example M14. Let $A \in M_{n\times n}(\mathbb{Q})$. Since $1 = AA^{-1}$ then $\det(1) = \det(A)\det(A^{-1})$. So

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

Let $P \in GL_n(\mathbb{Q})$. Then

$$\begin{aligned}\det(PAP^{-1}) &= \det(P)\det(A)\det(P^{-1}) = \det(P)\det(A)\frac{1}{\det(P)} \\ &= \det(P)\frac{1}{\det(P)}\det(A) = \det(A).\end{aligned}$$

Tutorial: Determinants by cofactor expansion

Definition (the (i, j)-cofactor)

Let $A \in M_{n\times n}(\mathbb{Q})$ and let $i, j \in \{1, \ldots, n\}$.

Let $A^{(i,j)}$ be the matrix A with the ith row removed and the jth column removed.

The *(i, j)-cofactor of A* is

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

Theorem (cofactor expansion)

Let $A \in M_{n\times n}(\mathbb{Q})$ *and let* $i, j \in \{1, \ldots, n\}$. *Then*

$$\det(A) = A_{i1}C_{i1} + A_{i2}C_{i2} + \cdots + A_{in}C_{in}$$

cofactor expansion across the i*th row*

and

$$\det(A) = A_{1j}C_{1j} + A_{2j}C_{2j} + \cdots + A_{nj}C_{nj}.$$

cofactor expansion down the j*th column*

Example M15 and 16. If $A = \begin{pmatrix} 1 & 2 & 1 \\ -1 & 1 & 1 \\ 0 & 1 & 3 \end{pmatrix}$ then the (2,3)-cofactor is

$$C_{23} = (-1)^{2+3} \det \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} = (-1)^5(1 \cdot 1 - 0 \cdot 2) = -1.$$

Using cofactor expansion along the third row,

$$\begin{aligned} \det(A) &= (-1)^{3+1} \cdot 0 \cdot \det \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} + (-1)^{3+2} \cdot 1 \cdot \det \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \\ &\quad + (-1)^{3+3} \cdot 3 \cdot \det \begin{pmatrix} 1 & 2 \\ -1 & 1 \end{pmatrix} \\ &= 0 - (1 \cdot 1 - (-1) \cdot 1) + 3(1 \cdot 1 - (-1) \cdot 2) \\ &= 0 - 2 + 9 = 7. \end{aligned}$$

Example M17. Let $A = \begin{pmatrix} 1 & -2 & 0 & 1 \\ 3 & 2 & 2 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & -4 & 2 & 4 \end{pmatrix}$.

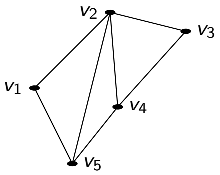
Using cofactor expansion down the fourth column,

$$\begin{aligned}
\det(A) &= (-1)^{1+4} \cdot 1 \cdot \det \begin{pmatrix} 3 & 2 & 2 \\ 1 & 0 & 1 \\ 0 & -4 & 2 \end{pmatrix} + 0 + 0 \\
&\qquad + (-1)^{4+4} \cdot 4 \cdot \det \begin{pmatrix} 1 & -2 & 0 \\ 3 & 2 & 2 \\ 1 & 0 & 1 \end{pmatrix} \\
&= -\Big((-1)^{1+1} \cdot 3 \cdot \det \begin{pmatrix} 0 & 1 \\ -4 & 2 \end{pmatrix} + (-1)^{2+1} \cdot 1 \cdot \det \begin{pmatrix} 2 & 2 \\ -4 & 2 \end{pmatrix} + 0 \Big) \\
&\qquad + 4 \Big((-1)^{1+1} \cdot 1 \cdot \det \begin{pmatrix} 2 & 2 \\ 0 & 1 \end{pmatrix} + (-1)^{1+2} \cdot (-2) \cdot \det \begin{pmatrix} 3 & 2 \\ 1 & 1 \end{pmatrix} + 0 \Big) \\
&= -3(0+4) + (4+8) + 4\big((2-0) + 2(2-3)\big) \\
&= -12 + 12 + 8 + 8 = 16.
\end{aligned}$$

Lecture 15: Applications to graphs and networks

Square matrices with 0, 1 entries are equivalent to graphs.

Example M1&2. The graph



has adjacency matrix $A = \begin{pmatrix} 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 \end{pmatrix}$

There is a 1 in the (i, j) entry if there is an edge connecting vertex i and vertex j.

Then

$$A^3 = A(A^2) = \begin{pmatrix} 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 2 & 1 & 1 & 2 & 1 \\ 1 & 4 & 1 & 2 & 2 \\ 1 & 1 & 2 & 1 & 2 \\ 2 & 2 & 1 & 3 & 1 \\ 1 & 2 & 2 & 1 & 3 \end{pmatrix}$$

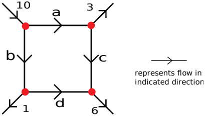
$$= \begin{pmatrix} 2 & 6 & 3 & 3 & 5 \\ 6 & 6 & 6 & 7 & 7 \\ 3 & 6 & 2 & 5 & 3 \\ 3 & 7 & 5 & 4 & 7 \\ 5 & 7 & 3 & 7 & 4 \end{pmatrix}.$$

The (i, j) entry of A^3 gives the number of paths of length three from vertex i to vertex j.

Example LS10. Calculating flows in networks

At each node • require (sum of flows in) = (sum of flows out).

Determine a, b, c and d in the network



Then

$$\begin{aligned} \text{Node 1:}\quad 10 &= a + b, \\ \text{Node 2:}\quad a &= 3 + c, \\ \text{Node 3:}\quad c + d &= 6, \\ \text{Node 4:}\quad b &= 1 + d. \end{aligned}$$

So

$$\begin{aligned} a + b + 0c + 0d &= 10, \\ a + 0b - c + 0d &= 2, \\ 0a + 0b + c + d &= 6, \\ 0a + b + 0c - d &= 1 \end{aligned} \quad \text{which is} \quad \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & -1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & -1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = \begin{pmatrix} 10 \\ 3 \\ 6 \\ 1 \end{pmatrix}.$$

Start with

$$\begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & -1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & -1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = \begin{pmatrix} 10 \\ 3 \\ 6 \\ 1 \end{pmatrix}.$$

Left multiply both sides by

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad \text{to get} \quad \begin{pmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & -1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = \begin{pmatrix} 3 \\ 7 \\ 6 \\ 1 \end{pmatrix}.$$

Left multiply both sides by

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \quad \text{to get} \quad \begin{pmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = \begin{pmatrix} 3 \\ 7 \\ 1 \\ 6 \end{pmatrix}.$$

Left multiply both sides by

$$\begin{pmatrix}1 & 0 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 1 & -1 & 0\\ 0 & 0 & 0 & 1\end{pmatrix} \quad \text{to get} \quad \begin{pmatrix}1 & 0 & -1 & 0\\ 0 & 1 & 0 & -1\\ 0 & 0 & 1 & 1\\ 0 & 0 & 1 & 1\end{pmatrix}\begin{pmatrix}a\\ b\\ c\\ d\end{pmatrix} = \begin{pmatrix}3\\ 1\\ 6\\ 6\end{pmatrix}.$$

Left multiply both sides by

$$\begin{pmatrix}1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 0 & 1\\ 0 & 0 & 1 & -1\end{pmatrix} \quad \text{to get} \quad \begin{pmatrix}1 & 0 & -1 & 0\\ 0 & 1 & 0 & -1\\ 0 & 0 & 1 & 1\\ 0 & 0 & 0 & 0\end{pmatrix}\begin{pmatrix}a\\ b\\ c\\ d\end{pmatrix} = \begin{pmatrix}3\\ 1\\ 6\\ 0\end{pmatrix}.$$

Left multiply both sides by

$$\begin{pmatrix}1 & 0 & 1 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\end{pmatrix} \quad \text{to get} \quad \begin{pmatrix}1 & 0 & 0 & 0\\ 0 & 1 & 0 & -1\\ 0 & 0 & 1 & 1\\ 0 & 0 & 0 & 0\end{pmatrix}\begin{pmatrix}a\\ b\\ c\\ d\end{pmatrix} = \begin{pmatrix}9\\ 1\\ 6\\ 0\end{pmatrix}.$$

This is the system

$$\begin{aligned} a &= 9, \\ b - d &= 1, \\ c + d &= 6, \\ 0 &= 0, \end{aligned} \qquad \text{which is} \qquad \begin{aligned} a &= 9 + 0d, \\ b &= 1 + 1d, \\ c &= 6 + (-1)d, \\ d &= 0 + 1d. \end{aligned}$$

where d can be any number. So

$$\begin{aligned} \mathrm{Sol}(A\mathbf{x} = \mathbf{b}) &= \left\{ \begin{pmatrix} 9 \\ 1 \\ 6 \\ 0 \end{pmatrix} + d \begin{pmatrix} 0 \\ 1 \\ -1 \\ 1 \end{pmatrix} \;\middle|\; d \in \mathbb{Q} \right\} \\ &= \begin{pmatrix} 9 \\ 1 \\ 6 \\ 0 \end{pmatrix} + \mathbb{Q}\text{-span} \left\{ \begin{pmatrix} 0 \\ 1 \\ -1 \\ 1 \end{pmatrix} \right\}. \end{aligned}$$

Lecture 16: Application of diagonalization to dynamics

Theorem (Diagonalization.)

Let $A \in M_n(\mathbb{F})$. The matrix A has n linearly independent eigenvectors $\mathbf{p}_1, \ldots, \mathbf{p}_n \in \mathbb{F}^n$ with eigenvalues $\lambda_1, \ldots, \lambda_n$ if and only if

$$A = PDP^{-1}$$

where,

$$P = \begin{pmatrix} | & & | \\ \mathbf{p}_1 & \cdots & \mathbf{p}_n \\ | & & | \end{pmatrix} \quad \textit{and} \quad D = \operatorname{diag}(\lambda_1, \ldots, \lambda_n) = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$$

so that $\mathbf{p}_1, \ldots, \mathbf{p}_n$ are the columns of P and D is the diagonal matrix with diagonal entries $\lambda_1, \ldots, \lambda_n$.

Example EV13.

$$\text{If} \qquad D = \begin{pmatrix} -4 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 2 \end{pmatrix} \qquad \text{then} \qquad D^{10} = \begin{pmatrix} (-4)^{10} & 0 & 0 \\ 0 & 3^{10} & 0 \\ 0 & 0 & 2^{10} \end{pmatrix}.$$

Example EV14. If $A = PDP^{-1}$ then

$$\begin{aligned} A^3 &= A \cdot A \cdot A = (PDP^{-1})(PDP^{-1})(PDP^{-1}) = PDP^{-1}PDP^{-1}PDP^{-1} \\ &= PD \cdot D \cdot DP^{-1} = PD^3P^{-1}, \end{aligned}$$

and, similarly, if $k \in \mathbb{Z}$ then

$$A^k = PD^kP^{-1}.$$

Let

$$A = \begin{pmatrix} 1 & 4 \\ 1 & 1 \end{pmatrix}, \qquad D = \begin{pmatrix} -1 & 0 \\ 0 & 3 \end{pmatrix}, \qquad P = \begin{pmatrix} -2 & 2 \\ 1 & 1 \end{pmatrix}.$$

Then

$$P^{-1} = \tfrac{1}{4}\begin{pmatrix} -1 & 2 \\ 1 & 2 \end{pmatrix} \qquad \text{and} \qquad A = PDP^{-1}.$$

So

$$\begin{aligned} A^k = PD^kP^{-1} &= \begin{pmatrix} -2 & 2 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} (-1)^k & 0 \\ 0 & 3^k \end{pmatrix} \tfrac{1}{4} \begin{pmatrix} -1 & 2 \\ 1 & 2 \end{pmatrix} \\ &= \begin{pmatrix} (-1)^k \cdot 2 & 3^k \cdot 2 \\ (-1)^k & 3^k \end{pmatrix} \tfrac{1}{4} \begin{pmatrix} -1 & 2 \\ 1 & 2 \end{pmatrix} \\ &= \tfrac{1}{4} \begin{pmatrix} 2((-1)^k + 3^k) & 4((-1)^{k+1} + 3^k) \\ (-1)^{k+1} + 3^k & 2((-1)^k + 3^k) \end{pmatrix} \end{aligned}$$

Example EV15. Let

$$x_n = \begin{pmatrix} r_n \\ p_n \\ w_n \end{pmatrix} \qquad \text{and} \qquad T = \begin{pmatrix} \frac{1}{2} & \frac{1}{4} & 0 \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ 0 & \frac{1}{4} & \frac{1}{2} \end{pmatrix}$$

and define an evolution process by

$$x_{n+1} = Tx_n.$$

This is the *Markov chain* defined by T. Since $T = PDP^{-1}$, where

$$P = \begin{pmatrix} 1 & -1 & 1 \\ 2 & 0 & -2 \\ 1 & 1 & 1 \end{pmatrix}, \; D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 \end{pmatrix}, \; P^{-1} = \begin{pmatrix} \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ -\frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{4} & -\frac{1}{4} & \frac{1}{4} \end{pmatrix}$$

then *the stationary state of the process on $\mathbb{R}^3$ defined by T* is

$$\lim_{n\to\infty} T^n x_0 = \lim_{n\to\infty} PD^nP^{-1}x_0 = \lim_{n\to\infty} P \begin{pmatrix} 1^n & 0 & 0 \\ 0 & \left(\frac{1}{2}\right)^n & 0 \\ 0 & 0 & 0 \end{pmatrix} P^{-1}$$

$$= P \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} P^{-1}x_0 = \begin{pmatrix} 1 & -1 & 1 \\ 2 & 0 & -2 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} x_0$$

$$= \begin{pmatrix} \frac{1}{4}(r_0 + p_0 + w_0) \\ \frac{1}{2}(r_0 + p_0 + w_0) \\ \frac{1}{4}(r_0 + p_0 + w_0) \end{pmatrix} = \begin{pmatrix} \frac{1}{4} \\ \frac{1}{2} \\ \frac{1}{4} \end{pmatrix}.$$

Lecture 17: Vector spaces and linear transformations

A field is a number system $\mathbb{F}$ that is similar to $\mathbb{Q}$, $\mathbb{R}$ and $\mathbb{C}$ (the precise definition is given on slide 139-140).

The number systems $\mathbb{Q}$, $\mathbb{R}$ and $\mathbb{C}$ are all fields. There are some 'more exotic' fields like *finite fields*. For example, if p is a prime number then the p-clock number system $\mathbb{F}_p$ is a finite field.

The world of $\mathbb{F}$-vector spaces works for any field $\mathbb{F}$. But, the properties *depend* on $\mathbb{F}$. For example, with dimension of a vector space

The $\mathbb{R}$-dimension of $\mathbb{R}^3$ is 3.
The $\mathbb{C}$-dimension of $\mathbb{C}^3$ is 3.
The $\mathbb{R}$-dimension of $\mathbb{C}^3$ is 6.
The $\mathbb{Q}$-dimension of $\mathbb{R}^3$ is ∞.

We often write 'Let $\mathbb{F}$ be a field'. You are encouraged to think of $\mathbb{F}$ as $\mathbb{R}$ or $\mathbb{Q}$ (or whatever your favourite field is).

Later we may explore some cool applications of vector spaces that use finite fields (codes, fast Fourier transform, etc.).

Definition ($\mathbb{F}$-vector space)

Let $\mathbb{F}$ be a field. A *$\mathbb{F}$-vector space*, or $\mathbb{F}$-*module*, is a set V with functions

$$\begin{array}{rcl} V \times V & \to & V \\ (v_1, v_2) & \mapsto & v_1 + v_2 \end{array} \quad \text{and} \quad \begin{array}{rcl} \mathbb{F} \times V & \to & V \\ (c, v) & \mapsto & cv \end{array}$$

(*addition* and *scalar multiplication*) such that

(a) If $v_1, v_2, v_3 \in V$ then $(v_1 + v_2) + v_3 = v_1 + (v_2 + v_3)$,

(b) There exists $0 \in V$ such that if $v \in V$ then $0 + v = v$.

(c) If $v \in V$ then there exists $-v \in V$ such that $v + (-v) = 0$.

(d) If $v_1, v_2 \in V$ then $v_1 + v_2 = v_2 + v_1$,

(e) If $c \in \mathbb{F}$ and $v_1, v_2 \in V$ then $c(v_1 + v_2) = cv_1 + cv_2$,

(f) If $c_1, c_2 \in \mathbb{F}$ and $v \in V$ then $(c_1 + c_2)v = c_1 v + c_2 v$,

(g) If $c_1, c_2 \in \mathbb{F}$ and $v \in V$ then $c_1(c_2 v) = (c_1 c_2)v$,

(h) If $v \in V$ then $1v = v$.

Linear transformations are for comparing vector spaces.

Definition

Let $\mathbb{F}$ be a field and let V and W be $\mathbb{F}$-vector spaces. An *$\mathbb{F}$-linear transformation from V to W* is a function $f: V \to W$ such that

(a) If $v_1, v_2 \in V$ then $f(v_1 + v_2) = f(v_1) + f(v_2)$,

(b) If $c \in \mathbb{F}$ and $v \in V$ then $f(cv) = cf(v)$.

One vector space can be a subspace of another.

Definition (Subspace)

Let V be an $\mathbb{F}$-vector space. A *subspace of W* is a subset $W \subseteq V$ such that

(a) $0 \in W$,

(b) If $w_1, w_2 \in W$ then $w_1 + w_2 \in W$,

(c) If $w \in W$ and $c \in \mathbb{F}$ then $cw \in W$.

Definition (Basis and dimension)

Let $\mathbb{F}$ be a field and let V be an $\mathbb{F}$-vector space.
Let $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ be a subset of V.
An *$\mathbb{F}$-linear combination* of $\mathbf{v}_1, \ldots, \mathbf{v}_k$ is an element of the set

$$\mathbb{F}\text{-span}\{\mathbf{v}_1, \ldots, \mathbf{v}_k\} = \{c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k \mid c_1, c_2, \ldots, c_k \in \mathbb{F}\}.$$

The set $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ is *linearly independent over $\mathbb{F}$* if it satisfies:

$$\text{if } c_1, \ldots, c_k \in \mathbb{F} \text{ and } c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k = 0$$

$$\text{then } c_1 = 0,\ c_2 = 0,\ \ldots,\ c_k = 0.$$

An *$\mathbb{F}$-basis* of V is a subset $B \subseteq V$ such that

(a) $\mathbb{F}\text{-span}(B) = V$,

(b) B is linearly independent.

The *$\mathbb{F}$-dimension* of V is the number of elements of a $\mathbb{F}$-basis B of V.

Favourite vector spaces and favourite bases

1. $\mathbb{R}^n = \{|a_1, a_2, \ldots, a_n\rangle \mid a_1, a_2, \ldots, a_n \in \mathbb{R}\} = M_{n\times 1}(\mathbb{R})$ has basis

$$\{e_1, e_2, \ldots, e_n\}, \quad \text{where} \quad e_i = |0, \ldots, 0, 1, 0, \ldots, 0\rangle,$$

has 1 in the ith entry and 0 elsewhere.

2. $M_{m\times n}(\mathbb{R})$ has basis

$$\{E_{ij} \mid i \in \{1, \ldots, m\},\ j \in \{1, \ldots, n\}\},$$

where E_{ij} is the matrix with 1 in the (i, j) entry and 0 elsewhere.

3. $\mathbb{R}[x]_{\leqslant n} = \{a_0 + a_1 x + \cdots + a_n x^n \mid a_0, a_1, \ldots, a_n \in \mathbb{R}\}$

$$\text{has basis} \quad \{1, x, x^2, \ldots, x^n\}.$$

4. The vector space of polynomials with coefficients in $\mathbb{R}$ is

$\mathbb{R}[x] = \mathbb{R}\text{-span}\{1, x, x^2, x^3, \ldots\}$ which has basis $\{1, x, x^2, x^3, \ldots\}$.

Example V22. Let $v_1, v_2, v_3, v_4 \in \mathbb{R}^3$ be given by

$$v_1 = |1,2,3\rangle, \quad v_2 = |3,6,9\rangle, \quad v_3 = |-1,0,-2\rangle, \quad v_4 = |1,4,4\rangle.$$

(a) Is $\{v_1, v_2, v_3, v_4\}$ linearly independent?

(b) Express v_2 and v_4 as linear combinations of v_1 and v_3.

(c) Is $\{v_1, v_3\}$ linearly independent?

(a) Since $v_2 = 3v_1$ then $0 = 3v_1 - v_2 = 3v_1 - v_2 + 0v_3 + -v_4$.
So $c_1 = 3$, $c_2 = -1$, $c_3 = 0$, $c_4 = 0$ is a case that gives
$c_1v_1 + c_2v_2 + c_3v_3 + c_4v_4 = 0$.
So $\{v_1, v_2, v_3, v_4\}$ is not linearly independent.

(b) Since $v_2 = 3v_1 + 0v_3$ then $v_2 \in \mathbb{R}\text{-span}\{v_1, v_3\}$.
Since $v_1 + v_3 = (0,2,1)$ and $v_1 + |0,2,1\rangle = |1,4,4\rangle$.
So $v_4 = 2v_1 + v_3$. So $v_4 \in \mathbb{R}\text{-span}\{v_1, v_3\}$.

(c) To show: If $c_1, c_2 \in \mathbb{R}$ and $c_1|1,2,3\rangle + c_2|-1,0,2\rangle = |0,0,0\rangle$ then $c_1 = 0$ and $c_2 = 0$.

Assume $c_1, c_2 \in \mathbb{R}$ and $c_1|1,2,3\rangle + c_2|-1,0,2\rangle = |0,0,0\rangle$.
Then

$$\begin{aligned} c_1 - c_2 &= 0, \\ 2c_1 + 0c_2 &= 0, \\ 3c_1 + 2c_2 &= 0. \end{aligned}$$

The first equation gives $c_1 = c_2$ and the second equation gives $2c_1 = 0$ so that $c_2 = c_1 = 0$. This system has

$$\text{only one solution:} \qquad c_1 = 0, \quad c_2 = 0.$$

So $\{v_1, v_3\}$ is linearly independent. □

Example V7. Is the line

$$L = \{|x, y\rangle \in \mathbb{R}^2 \mid y = 2x + 1\} \qquad \text{a subspace of } \mathbb{R}^2?$$

Since $0 = |0, 0\rangle$ and $0 \neq 2 \cdot 0 + 1$ then $0 \notin L$.
So L is not a subspace of $\mathbb{R}^2$.

Example A8. Is the line

$$L = \{|x, y\rangle \in \mathbb{R}^2 \mid y = 2x\} \qquad \text{a subspace of } \mathbb{R}^2?$$

Since $|0, 0\rangle = |0, 2 \cdot 0\rangle$ then $|0, 0\rangle \in L$.
Assume $|a, 2a\rangle, |b, 2b\rangle \in L$. Then

$$|a, 2a\rangle + |b, 2b\rangle = |(a + b), 2(a + b)\rangle \in L.$$

Assume $|a, 2a\rangle \in L$ and $c \in \mathbb{R}$. Then

$$c \cdot |a, 2a\rangle = |(ca), 2(ca)\rangle \in L.$$

So L is a subspace of $\mathbb{R}^2$.

Definition (Field)

A *field* is a set $\mathbb{F}$ with functions

$$\begin{array}{ccc} \mathbb{F}\times\mathbb{F} & \longrightarrow & \mathbb{F} \\ (a,b) & \longmapsto & a+b \end{array} \quad \text{and} \quad \begin{array}{ccc} \mathbb{F}\times\mathbb{F} & \longrightarrow & \mathbb{F} \\ (a,b) & \longmapsto & ab \end{array}$$

such that

(Fa) If $a, b, c \in \mathbb{F}$ then $(a+b)+c = a+(b+c)$,

(Fb) If $a, b \in \mathbb{F}$ then $a+b = b+a$,

(Fc) There exists $0 \in \mathbb{F}$ such that

$$\text{if } a \in \mathbb{F} \quad \text{then} \quad 0+a = a \text{ and } a+0 = a,$$

(Fd) If $a \in \mathbb{F}$ then there exists $-a \in \mathbb{F}$ such that $a+(-a) = 0$ and $(-a)+a = 0$,

(Fe) If $a, b, c \in \mathbb{F}$ then $(ab)c = a(bc)$,

Definition (Field continued)

(Ff) If $a, b, c \in \mathbb{F}$ then

$$(a+b)c = ac + bc \qquad \text{and} \qquad c(a+b) = ca + cb,$$

(Fg) There exists $1 \in \mathbb{F}$ such that

$$\text{if } a \in \mathbb{F} \quad \text{then} \quad 1 \cdot a = a \text{ and } a \cdot 1 = a,$$

(Fh) If $a \in \mathbb{F}$ and $a \neq 0$ then there exists $a^{-1} \in \mathbb{F}$ such that $aa^{-1} = 1$ and $a^{-1}a = 1$,

(Fi) If $a, b \in \mathbb{F}$ then $ab = ba$.

Lecture 18: Linear transformations

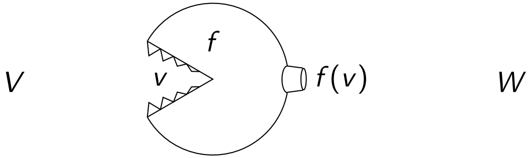
Linear transformations are for comparing vector spaces.

Definition

Let $\mathbb{F}$ be a field and let V and W be $\mathbb{F}$-vector spaces. An *$\mathbb{F}$-linear transformation from V to W* is a function $f: V \to W$ such that

(a) If $v_1, v_2 \in V$ then $f(v_1 + v_2) = f(v_1) + f(v_2)$,

(b) If $c \in \mathbb{F}$ and $v \in V$ then $f(cv) = cf(v)$.



Example A1. Let $A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \end{pmatrix}$.

Let $T\colon \mathbb{R}^4 \to \mathbb{R}^2$ be the function given by $T(x) = Ax$ so that

$$T\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} x_1 + 2x_2 + 3x_3 + 4x_4 \\ 5x_5 + 6x_6 + 7x_7 + 8x_8 \end{pmatrix}.$$

Show that T is a linear transformation.

Let $u, v \in \mathbb{R}^4$. Then, by the distributive property of matrix multiplication,

$$T(u+v) = A(u+v) = Au + Av = T(u) + T(v).$$

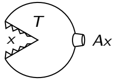
Let $u \in R^4$ and $c \in \mathbb{R}$. Then, by the associative property of scalar multiplication for matrices,

$$T(cu) = Acu = cAu = cT(u).$$

So T is a linear transformation.

Example A2. Let $t, s \in \mathbb{Z}_{>0}$ and $A \in M_{t\times s}(\mathbb{R})$. Let $T\colon \mathbb{R}^s \to \mathbb{R}^t$ be the function given by

$$T(x) = Ax.$$



Show that T is a linear transformation.
Let $u, v \in \mathbb{R}^s$. Then, by the distributive property of matrix multiplication for matrices,

$$T(u + v) = A(u + v) = Au + Av = T(u) + T(v).$$

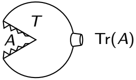
Let $u \in \mathbb{R}^s$ and $c \in \mathbb{R}$. Then, by the associative property of scalar multiplication for matrices,

$$T(cu) = Acu = cAu = cT(u).$$

So T is a linear transformation.

Example A3. Let $n \in \mathbb{Z}_{>0}$ and let $T\colon M_n(\mathbb{Q}) \to \mathbb{Q}$ by the function given by

$$T(A) = \mathrm{Tr}(A).$$



Show that T is a $\mathbb{Q}$-linear transformation.
Let $A, B \in M_{n\times n}(\mathbb{Q})$. Then

$$\begin{aligned} T(A+B) = \mathrm{Tr}(A+B) &= (A+B)_{11} + \cdots + (A+B)_{nn} \\ &= A_{11} + B_{11} + \cdots + A_{nn} + B_{nn} \\ &= A_{11} + \cdots + A_{nn} + + B_{11} + \cdots B_{nn} \\ &= \mathrm{Tr}(A) + \mathrm{Tr}(B) = T(A) + T(B). \end{aligned}$$

Let $A \in M_{n\times n}(\mathbb{Q})$ and $c \in \mathbb{Q}$. Then

$$\begin{aligned} T(cA) = \mathrm{Tr}(cA) &= (cA)_{11} + \cdots + (cA)_{nn} \\ &= cA_{11} + \cdots + cA_{nn} \\ &= c(A_{11} + \cdots + A_{nn}) = c\,\mathrm{Tr}(A) = cT(A). \end{aligned}$$

So T is a linear transformation.

Example A4. Let $T\colon M_{2\times 2}(\mathbb{Q}) \to \mathbb{Q}$ be the function given by

$$T\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \det\begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc.$$

Show that T is a linear transformation.
Let $A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ and let $c = 2$. Then

$$T(cA) = \det(2A) = \det\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} = 2\cdot 2 - 0\cdot 0 = 4$$

and

$$cT(A) = 2\det\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = 2\cdot(1\cdot 1 - 0\cdot 0) = 2.$$

So this gives an example where $T(cA) \neq cT(A)$.
So T *cannot possibly be a linear transformation*.

Lecture 19: span, linear independence and bases

Definition (Basis and dimension)

Let $\mathbb{F}$ be a field and let V be an $\mathbb{F}$-vector space.
Let $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ be a subset of V.
An *$\mathbb{F}$-linear combination* of $\mathbf{v}_1, \ldots, \mathbf{v}_k$ is an element of the set

$$\mathbb{F}\text{-span}\{\mathbf{v}_1, \ldots, \mathbf{v}_k\} = \{c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k \mid c_1, c_2, \ldots, c_k \in \mathbb{F}\}.$$

The set $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ is *linearly independent over $\mathbb{F}$* if it satisfies:

$$\text{if } c_1, \ldots, c_k \in \mathbb{F} \text{ and } c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k = 0$$

$$\text{then } c_1 = 0,\ c_2 = 0,\ \ldots,\ c_k = 0.$$

An *$\mathbb{F}$-basis* of V is a subset $B \subseteq V$ such that

(a) $\mathbb{F}\text{-span}(B) = V$,

(b) B is linearly independent.

The *$\mathbb{F}$-dimension* of V is the number of elements of a $\mathbb{F}$-basis B of V.

Example A9. Let V be a $\mathbb{Q}$-vector space and let $v_1, v_2, v_3, v_4, v_5 \in V$. Let $S = \{v_1, v_2, v_3, v_4, v_5\}$. Show that $\mathbb{Q}$-span(S) is a subspace of V.

(a) Since $0 = 0v_1 + 0v_2 + 0v_3 + 0v_4 + 0v_5$ then $0 \in \mathbb{Q}$-span(S).

(b) Assume $a = a_1v_1 + a_2v_2 + a_3v_3 + a_4v_4 + a_5v_5 \in \mathbb{Q}$-span$(S)$ and $b = b_1v_1 + b_2v_2 + b_3v_3 + b_4v_4 + b_5v_5 \in \mathbb{Q}$-span$(S)$. Then

$$\begin{aligned} a + b &= a_1v_1 + a_2v_2 + a_3v_3 + a_4v_4 + a_5v_5 \\ &\quad + b_1v_1 + b_2v_2 + b_3v_3 + b_4v_4 + b_5v_5 \\ &= (a_1 + b_1)v_1 + (a_2 + b_2)v_2 + (a_3 + b_3)v_3 \\ &\quad + (a_4 + b_4)v_4 + (a_5 + b_5)v_5. \end{aligned}$$

So $a + b \in \mathbb{Q}$-span(S).

(c) Assume $a = a_1v_1 + a_2v_2 + a_3v_3 + a_4v_4 + a_5v_5 \in \mathbb{Q}$-span$(S)$ and $c \in \mathbb{Q}$. Then

$$\begin{aligned} ca &= c(a_1v_1 + a_2v_2 + a_3v_3 + a_4v_4 + a_5v_5) \\ &= ca_1v_1 + ca_2v_2 + ca_3v_3 + ca_4v_4 + ca_5v_5 \in \mathbb{Q}\text{-span}(S). \end{aligned}$$

So $\mathbb{Q}$-span(S) is a subspace of V.

Example V13. In $\mathbb{R}[x]_{\leqslant 2}$, is $1-2x-x^2 \in \mathbb{R}\text{-span}\{1+x+x^2, 3+x^2\}$?

By definition $\mathbb{R}\text{-span}\{1+x+x^2, 3+x^2\}$

$$= \{c_1(1+x+x^2)+c_2(3+x^2) \mid c_1c_2 \in \mathbb{R}\}.$$

So we need to show that there exist $c_1, c_2 \in \mathbb{R}$ such that

$$c_1(1+x+x^2)+c_2(3+x^2) = 1-2x-x^2.$$

So we need to show that the system $\begin{aligned} c_1+3c_2 &= 1, \\ c_1+0c_2 &= -2, \\ c_1+c_2 &= -1, \end{aligned}$ has a solution.

The second equation gives $c_1 = -2$ and then $c_2 = -1-c_1 = 1+2 = 3$. Since the equation $c_1+3c_2 = 1$ also works when $c_1 = -2$ and $c_2 = 3$ then $c_1 = -2$, $c_2 = 1$ is a solution to this system.

Alternatively, the solution can be found by row reduction as follows. In matrix form the equations are

$$\begin{pmatrix} 1 & 3 \\ 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} 1 \\ -2 \\ -1 \end{pmatrix}.$$

Left multiply both sides by

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & -1 \end{pmatrix} \quad \text{to get} \quad \begin{pmatrix} 1 & 3 \\ 1 & 1 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix}.$$

Left multiply both sides by

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{to get} \quad \begin{pmatrix} 1 & 1 \\ 0 & 2 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} -1 \\ 2 \\ -1 \end{pmatrix}.$$

Left multiply both sides by

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & -2 \end{pmatrix} \quad \text{to get} \quad \begin{pmatrix} 1 & 1 \\ 0 & -1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} -1 \\ -1 \\ 0 \end{pmatrix}.$$

Left multiply both sides by

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{to get} \quad \begin{pmatrix} 1 & 1 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}.$$

Left multiply both sides by

$$\begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{to get} \quad \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix}.$$

So $c_1 = -2$ and $c_2 = 1$ is a solution.

So $-2(1 + x + x^2) + (3 + x^2) = 1 - 2x - x^2$.

So $1 - 2x - x^2 \in \mathbb{R}\text{-span}\{1 + x + x^2, 3 + x^2\}$.

So $1 - 2x - x^2$ is a linear combination of $1 + x + x^2$ and $3 + x^2$ and

$$1 - 2x - x^2 \in \mathbb{R}\text{-span}\{1 + x + x^2, 3 + x^2\}.$$

$\square$

Example V16new. Let S be the subset of $\mathbb{R}[x]_{\leqslant 2}$ given by

$S = \{1+2x, 1+5x+3x^2, x+x^2\}$. Show that $\text{span}(S) = \mathbb{R}[x]_{\leqslant 2}$.

Proof. By definition

$$\mathbb{R}\text{-span}(S) = \{c_1(1+2x)+c_2(1+5x+3x^2)+c_3(x+x^2) \mid c_1, c_2, c_3 \in \mathbb{R}\}.$$

To show: (a) $\mathbb{R}\text{-span}(S) \subseteq \mathbb{R}^3$
(b) $\mathbb{R}^3 \subseteq \mathbb{R}\text{-span}(S)$.

(a) Since $S \subseteq \mathbb{R}^3$ and $\mathbb{R}^3$ is closed under addition and scalar multiplication then $\mathbb{R}\text{-span}(S) \subseteq \mathbb{R}[x]_{\leqslant 2}$.

(b) To show: $\mathbb{R}[x]_{\leqslant 2} \subseteq \text{span}(S)$.

To show: $\mathbb{R}\text{-span}\{1, x, x^2\} \subseteq \text{span}(S)$.

Since $\mathbb{R}\text{-span}(S)$ is closed under addition and scalar multiplication,

To show: $\{1, x, x^2\} \subseteq \mathbb{R}\text{-span}(S)$.

To show: There exist $c_1, c_2, c_3, d_1, d_2, d_3, r_1, r_2, r_3 \in \mathbb{R}$ such that

$$\begin{aligned} c_1(1+2x) + c_2(1+5x+3x^2) + c_3(x+x^2) &= 1 + 0x + x^2, \\ d_1(1+2x) + d_2(1+5x+3x^2) + d_3(x+x^2) &= 0 + x + 0x^2, \\ r_1(1+2x) + r_2(1+5x+3x^2) + r_3(x+x^2) &= 0 + 0x + x^2. \end{aligned}$$

To show: There exist $c_1, c_2, c_3, d_1, d_2, d_3, r_1, r_2, r_3 \in \mathbb{R}$ such that

$$\begin{pmatrix} 1 & 1 & 0 \\ 2 & 5 & 1 \\ 0 & 3 & 1 \end{pmatrix} \begin{pmatrix} c_1 & d_1 & r_1 \\ c_2 & d_2 & r_2 \\ c_3 & d_3 & r_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Multiply both sides by

$$\begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \text{ to get } \begin{pmatrix} 1 & 1 & 0 \\ 0 & 3 & 1 \\ 0 & 3 & 1 \end{pmatrix} \begin{pmatrix} c_1 & d_1 & r_1 \\ c_2 & d_2 & r_2 \\ c_3 & d_3 & r_3 \end{pmatrix} = \begin{pmatrix} 1 & -2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Multiply both sides by

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix} \text{ to get } \begin{pmatrix} 1 & 1 & 0 \\ 0 & 3 & 1 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} c_1 & d_1 & r_1 \\ c_2 & d_2 & r_2 \\ c_3 & d_3 & r_3 \end{pmatrix} = \begin{pmatrix} 1 & -2 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix}.$$

Since the bottom row on the left hand side is all 0 and the bottom row on the right hand sides is not all 0 then there *do not exist* $c_1, c_2, c_3, d_1, d_2, d_3, r_1, r_2, r_3 \in \mathbb{R}$ such that

$$\begin{pmatrix} 1 & 1 & 0 \\ 2 & 5 & 1 \\ 0 & 3 & 1 \end{pmatrix} \begin{pmatrix} c_1 & d_1 & r_1 \\ c_2 & d_2 & r_2 \\ c_3 & d_3 & r_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

So $\{1, x, x^2\} \not\subseteq \mathbb{R}\text{-span}(S)$.

So $\text{span}(S) \neq \mathbb{R}[x]_{\leqslant 2}$. $\square$

Example V23. Is $S = \{(1,-1),(2,4)\}$ a basis of $\mathbb{R}^2$?
Let

$$A = \begin{pmatrix} 1 & 2 \\ -1 & 4 \end{pmatrix}. \qquad \text{Then} \qquad A^{-1} = \frac{1}{6}\begin{pmatrix} 4 & -2 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} \frac{2}{3} & -\frac{1}{3} \\ \frac{1}{6} & \frac{1}{6} \end{pmatrix}.$$

So

$$\begin{pmatrix} 1 & 2 \\ -1 & 4 \end{pmatrix}\begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \qquad \text{gives} \qquad \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

So S is linearly independent.

If $|a,b\rangle \in \mathbb{R}^2$ then $|a,b\rangle = c_1|1,-1\rangle + c_2|2,4\rangle$, where

$$\begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} \frac{2}{3} & -\frac{1}{3} \\ \frac{1}{6} & \frac{1}{6} \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \frac{2}{3}a - \frac{1}{3}b \\ \frac{1}{6}a + \frac{1}{6}b \end{pmatrix}.$$

So $\mathbb{R}^2 \subseteq \mathbb{R}\text{-span}(S)$. Since $S \subseteq \mathbb{R}^2$ and $\mathbb{R}^2$ is closed under addition and scalar multiplication then $\mathbb{R}\text{-span}(S) \subseteq \mathbb{R}^2$. So $\mathbb{R}\text{-span}(S) = \mathbb{R}^2$.

So S is a basis of $\mathbb{R}^2$.

Example V21. Let S be the subset of $M_2(\mathbb{R})$ given by

$$S = \left\{ \begin{pmatrix} 1 & 3 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} -2 & 1 \\ 1 & -1 \end{pmatrix}, \begin{pmatrix} 1 & 10 \\ 4 & 2 \end{pmatrix} \right\}.$$

Is S linearly independent?

To show: If $c_1, c_2, c_3 \in \mathbb{R}$ and

$$c_1 \begin{pmatrix} 1 & 3 \\ 1 & 1 \end{pmatrix} + c_2 \begin{pmatrix} -2 & 1 \\ 1 & -1 \end{pmatrix} + c_3 \begin{pmatrix} 1 & 10 \\ 4 & 2 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

then $c_1 = 0$, $c_2 = 0$, $c_3 = 0$.

Suppose an oracle tells you to try (or you guess) $c_1 = -3$, $c_2 = 1$, $c_3 = -1$ and then you verify that

$$\begin{aligned} -3 \begin{pmatrix} 1 & 3 \\ 1 & 1 \end{pmatrix} + \begin{pmatrix} -2 & 1 \\ 1 & -1 \end{pmatrix} - \begin{pmatrix} 1 & 10 \\ 4 & 2 \end{pmatrix} &= -3 \begin{pmatrix} 1 & 3 \\ 1 & 1 \end{pmatrix} + \begin{pmatrix} -3 & -9 \\ -3 & -3 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}. \end{aligned}$$

This means that you don't have to have c_1, c_2 and c_3 all 0 to get a zero linear combination.
So S is not linearly independent.

If you have no oracle, or are not a good guesser, then proceed as follows. Assume $c_1, c_2, c_3 \in \mathbb{R}$ and

$$c_1 \begin{pmatrix} 1 & 3 \\ 1 & 1 \end{pmatrix} + c_2 \begin{pmatrix} -2 & 1 \\ 1 & -1 \end{pmatrix} + c_3 \begin{pmatrix} 1 & 10 \\ 4 & 2 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Then

$$\begin{aligned} c_1 - 2c_2 + c_3 &= 0, \\ 3c_1 + c_2 + 10c_3 &= 0, \\ c_1 + c_2 + 4c_3 &= 0, \\ c_1 - c_2 + 2c_3 &= 0, \end{aligned} \quad \text{or, equivalently,} \quad \begin{pmatrix} 1 & -2 & 1 \\ 3 & 1 & 10 \\ 1 & 1 & 4 \\ 1 & -1 & 2 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Left multiply both sides by

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & -1 \end{pmatrix} \quad \text{to get} \quad \begin{pmatrix} 1 & -2 & 1 \\ 3 & 1 & 10 \\ 1 & -1 & 2 \\ 0 & 2 & 2 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Left multiply both sides by

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & -3 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \text{to get} \quad \begin{pmatrix} 1 & -2 & 1 \\ 1 & -1 & 2 \\ 0 & 4 & 4 \\ 0 & 2 & 2 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Left multiply both sides by

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \text{to get} \quad \begin{pmatrix} 1 & -1 & 2 \\ 0 & -1 & -1 \\ 0 & 4 & 4 \\ 0 & 2 & 2 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Left multiply both sides by

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & -2 \end{pmatrix} \quad \text{to get} \quad \begin{pmatrix} 1 & -1 & 2 \\ 0 & -1 & -1 \\ 0 & 2 & 2 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Left multiply both sides by

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \text{to get} \quad \begin{pmatrix} 1 & -1 & 2 \\ 0 & 2 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Left multiply both sides by

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \text{to get} \quad \begin{pmatrix} 1 & -1 & 2 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Left multiply both sides by

$$\begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \text{to get} \quad \begin{pmatrix} 1 & 0 & 3 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

This gives the system

$$c_1 + 3c_3 = 0, \quad c_2 + c_3 = 0, \qquad \text{which is} \qquad \begin{aligned} c_1 &= -3c_3 \\ c_2 &= -c_3, \\ c_3 &= c_3, \end{aligned}$$

which has solutions

$$\left\{ c_3 \begin{pmatrix} -3 \\ -1 \\ 1 \end{pmatrix} \;\middle|\; c_3 \in \mathbb{R} \right\} = \mathbb{R}\text{-span} \left\{ \begin{pmatrix} -3 \\ -1 \\ 1 \end{pmatrix} \right\}.$$

So $c_1 = 0$, $c_2 = 0$, $c_3 = 0$ is not the only solution.
This means that you don't have to have c_1, c_2 and c_3 all 0 to get a zero linear combination.
So S is not linearly independent.

Lecture 20: Kernel and image of a linear transformation

Definition (Kernel and image of a linear transformation)

The *kernel* of an $\mathbb{F}$-linear transformation $f\colon V \to W$ is the set

$$\ker(f) = \{v \in V \mid f(v) = 0\}.$$

The *image* of an $\mathbb{F}$-linear transformation $f\colon V \to W$ is the set

$$\operatorname{im}(f) = \{f(v) \mid v \in V\}.$$

Definition (Kernel and image of a matrix)

Let $A \in M_{t\times s}(\mathbb{Q})$. The *kernel of A* is

$$\ker(A) = \{x \in \mathbb{Q}^s \mid Ax = 0\}$$

and the *image of A* is

$$\operatorname{im}(A) = \{Ax \mid s \in \mathbb{Q}^s\}.$$

Example A5. Let $T\colon V \to W$ be an $\mathbb{R}$-linear transformation. Show that $\ker(T) = \{v \in V \mid T(v) = 0\}$ is a subspace of V.

Let $v_1, v_2 \in \ker(T)$. Then

$$T(v_1 + v_2) = T(v_1) + T(v_2) = 0 + 0 = 0. \quad \text{So } v_1 + v_2 \in \ker(T).$$

Subtracting $T(0)$ from each side of the equation $T(0) = T(0+0) = T(0) + T(0)$ gives

$$0 = T(0), \quad \text{and so} \quad 0 \in \ker(T).$$

Let $v \in \ker(T)$ and let $c \in \mathbb{R}$. Then

$$T(cv) = cT(v) = c \cdot 0 = 0 \quad \text{and so} \quad cv \in \ker(T).$$

So $\ker(T)$ is a subspace of V.

Example A6. Let $T\colon V \to W$ be an $\mathbb{R}$-linear transformation. Show that $\operatorname{im}(T) = \{T(v) \mid v \in V\}$ is a subspace of W.
Subtracting $T(0)$ from each side of the equation $T(0) = T(0+0) = T(0) + T(0)$ gives

$$0 = T(0), \quad \text{and so} \quad 0 \in \operatorname{im}(T).$$

Let $w_1, w_2 \in W$. Then there exist $v_1, v_2 \in V$ such that

$$T(v_1) = w_1 \quad \text{and} \quad T(v_2) = w_2.$$

Then $w_1 + w_2 = T(v_1) + T(v_2) = T(v_1 + v_2)$,

$$\text{and so} \quad w_1 + w_2 \in \operatorname{im}(T).$$

Let $w \in W$ and let $c \in \mathbb{R}$. Then there exists $v \in V$ such that

$$T(v) = w.$$

Then $cw = cT(v) = T(cv)$

$$\text{and so} \quad cw \in \operatorname{im}(T).$$

So $\operatorname{im}(T)$ is a subspace of W.

Definition (Injective, surjective, bijective, invertible)

Let S and T be sets and let $f\colon S \to T$ be a function from S to T.

(a) The function $f\colon S \to T$ is *injective* if f satistifes

$$\text{if } s_1, s_2 \in S \text{ and } f(s_1) = f(s_2) \quad \text{then } s_1 = s_2.$$

(b) The function $f\colon S \to T$ is *surjective* if f satisfies

$$\text{if } t \in T \quad \text{then there exists } s \in S \text{ such that } f(s) = t.$$

(c) The function $f\colon S \to T$ is *bijective* if f is

both injective and surjective.

(d) The function $f\colon S \to T$ is *invertible* if there exists a function $g\colon T \to S$ such that

$$g \circ f = \mathrm{Id}_S \quad \text{and} \quad f \circ g = \mathrm{Id}_T\,.$$

Definition

Let V be a vector space. The *dimension* of V is

$$\dim(V) = (\text{number of elements in a basis } B \text{ of } V).$$

Theorem (The rank-nullity theorem)

Let $f\colon V \to W$ be an $\mathbb{F}$-linear transformation. Then

(a) $\ker(f)$ *is a subspace of V.*

(b) $\mathrm{im}(f)$ *is a subspace of W.*

(c) $\dim(\ker(f)) + \dim(im(f)) = \dim(V)$.

Theorem

Let $f\colon V \to W$ be an $\mathbb{F}$-linear transformation. Then

(a) *f is injective if and only if $\ker(f) = \{0\}$.*

(b) *f is surjective if and only if $im(f) = W$.*

(c) *f is invertible if and only if f is both injective and surjective.*

Example LT15. Let $T\colon \mathbb{R}^3 \to \mathbb{R}^2$ be the linear transformation given by

$$T(x,y,z) = (2x - y, y + z).$$

Find bases for $\ker(T)$ and $\operatorname{Im}(T)$ and verify the rank-nullity theorem.

$$\begin{aligned}
\ker(T) &= \{|x,y,z\rangle \in \mathbb{R}^3 \mid T(x,y,z) = |0,0\rangle\} \\
&= \{|x,y,z\rangle \in \mathbb{R}^3 \mid |2x-y, y+z\rangle = |0,0\rangle\} \\
&= \left\{ |x,y,z\rangle \in \mathbb{R}^3 \;\middle|\; \begin{array}{r} 2x - y = 0, \\ y + z = 0 \end{array} \right\} \\
&= \left\{ |x,y,z\rangle \in \mathbb{R}^3 \;\middle|\; \begin{array}{l} x = \frac{1}{2}y, \\ y = y, \\ z = -y \end{array} \right\} \\
&= \{y \cdot |\tfrac{1}{2}, 1, 1\rangle \in \mathbb{R}^3 \mid y \in \mathbb{R}\} = \mathbb{R}\text{-span}\{|\tfrac{1}{2}, 1, -1\rangle\}
\end{aligned}$$

and $\{|\frac{1}{2}, 1, -1\rangle\}$ is a basis of $\ker(T)$. So $\dim(\ker(T)) = 1$.

Since

$$T(\tfrac{1}{2}, 0, 0) = |1, 0\rangle \quad \text{and} \quad T(0, 0, 1) = |0, 1\rangle$$

then

$$|1, 0\rangle \text{ and } |0, 1\rangle \text{ are elements of } \mathrm{im}(T).$$

Since $\mathrm{im}(T)$ is a subspace of $\mathbb{R}^2$ then $\mathbb{R}\text{-span}\{|1, 0\rangle, |0, 1\rangle\}$ is a subset of $\mathrm{im}(T)$. So

$$\mathrm{im}(T) = \mathbb{R}^2 \quad \text{and} \quad \{|1, 0\rangle, |0, 1\rangle\} \text{ is a basis of } \mathrm{im}(T).$$

So $\dim(\mathrm{im}(T)) = 2$ and

$$\dim(\ker(T)) + \dim(\mathrm{im}(T)) = 2 + 1 = 3 \quad \text{and} \quad 3 = \dim(\mathbb{R}^3)$$

is the dimension of the source of the linear transformation $T\colon \mathbb{R}^3 \to \mathbb{R}^2$.

Example LT16&17. Let $T\colon \mathbb{R}[x]_{\leqslant 2} \to \mathbb{R}[x]_{\leqslant 1}$ be the linear transformation given by

$$T(a_0 + a_1x + a_2x^2) = (a_0 - a_1 + a_2)(1 + 2x).$$

(a) Find bases for $\ker(T)$ and $\operatorname{Im}(T)$.

(b) Is T injective?

(c) Is T surjective?

$$\begin{aligned}
\ker(T) &= \{a_0 + a_1x + a_2x^2 \in \mathbb{R}[x]_{\leqslant 2} \mid T(a_0 + a_1x + a_2x^2) = 0 + 0x\} \\
&= \{a_0 + a_1x + a_2x^2 \in \mathbb{R}[x]_{\leqslant 2} \mid (a_0 - a_1 + a_2)(1 + 2x) = 0 + 0x\} \\
&= \left\{ a_0 + a_1x + a_2x^2 \in \mathbb{R}[x]_{\leqslant 2} \;\middle|\; \begin{array}{c} a_0 - a_1 + a_2 = 0, \\ 2(a_0 - a_1 + a_2) = 0 \end{array} \right\} \\
&= \{a_0 + a_1x + a_2x^2 \in \mathbb{R}[x]_{\leqslant 2} \mid a_0 = a_1 - a_2\} \\
&= \{(a_1 - a_2) + a_1x + a_2x^2 \mid a_1, a_2 \in \mathbb{R}\} \\
&= \{a_1(1 + x) + a_2(-1 + x^2) \mid a_1, a_2 \in \mathbb{R}\} \\
&= \mathbb{R}\text{-span}\{1 + x, -1 + x^2\}
\end{aligned}$$

and $\{1+x, -1+x^2\}$ is a basis of $\ker(T)$.

$$\begin{aligned}\operatorname{im}(T) &= \{T(a_0+a_1x+a_2x^2) \mid a_0, a_1, a_2 \in \mathbb{R}\} \\ &= \{(a_0-a_1+a_2)(1+2x) \mid a_0, a_1, a_2 \in \mathbb{R}\} \\ &= \{a(1+2x) \mid a \in \mathbb{R}\} = \mathbb{R}\text{-span}\{1+2x\}\end{aligned}$$

and $\{1+2x\}$ is a basis of $\operatorname{im}(T)$. So $\dim(\operatorname{im}(T)) = 1$.

Since $\ker(T) \neq \{0\}$ then T is not injective.

Since $\mathbb{R}[x]_{\leqslant 1} = \{c_0 + c_1x \mid c_1, c_2 \in \mathbb{R}\}$ then

$$\operatorname{im}(T) \neq \mathbb{R}[x]_{\leqslant 1} \quad \text{and} \quad T \text{ is not surjective.}$$

Lecture 21: With respect to a basis

Even in an arbitrary vector space, vectors and linear transformations can be converted to matrices, *provided that the corresponding column vectors and matrices are constructed with respect to a basis*.

Definition (Basis)

Let V be an $\mathbb{F}$-vector space. A *basis* of V is a set $S = \{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n\}$ such that every vector in V is a unique linear combination of $\mathbf{b}_1, \ldots, \mathbf{b}_n$.

Definition (Coordinates)

Let $B = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ be an ordered basis for an $\mathbb{F}$-vector space V and let $v \in V$. The *coordinate vector of v with respect to $\mathcal{B}$* is $[v]_B \in \mathbb{F}^n$ given by

$$[v]_B = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} \quad \text{if} \quad v = c_1\mathbf{b}_2 + \cdots + c_n\mathbf{b}_n.$$

Example V33. The coordinate vector of $v = (1, 5)$ with respect to the basis $S = \{(1, 0), (0, 1)\}$ of $\mathbb{R}^2$ is

$$[v]_S = \begin{pmatrix} 1 \\ 5 \end{pmatrix} \qquad \text{since} \quad (1, 5) = 1 \cdot (1, 0) + 5 \cdot (0.1).$$

The coordinate vector of $v = (1, 5)$ with respect to the basis $B = \{(2, 1), (-1, 1)\}$ of $\mathbb{R}^2$ is

$$[v]_B = \begin{pmatrix} 2 \\ 3 \end{pmatrix} \qquad \text{since} \quad (1, 5) = 2 \cdot (2, 1) + 3 \cdot (-1.1).$$

Example V34. The coordinate vector of $p = 2 + 7x - 9x^2$ with respect to the basis $B = \{2, \frac{1}{2}x, -3x^2\}$ of $\mathbb{Q}[x]_{\leqslant 2}$ is

$$[p]_B = \begin{pmatrix} 1 \\ 14 \\ 3 \end{pmatrix} \quad \text{since} \quad 2 + 7x - 9x^2 = 1 \cdot 2 + 14 \cdot (\tfrac{1}{2}x) + 3 \cdot (-3x^2).$$

Definition

Let $f\colon V \to W$ be an $\mathbb{F}$-linear transformation. Let $B = \{\mathbf{b}_1, \ldots, \mathbf{b}_s\}$ be a basis of V and let $C = \{\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_t\}$ be a basis of W. Suppose that

$$\begin{aligned} f(\mathbf{b}_1) &= A_{11}\mathbf{c}_1 + A_{21}\mathbf{c}_2 + \cdots + A_{n1}\mathbf{c}_n, \\ f(\mathbf{b}_2) &= A_{12}\mathbf{c}_1 + A_{22}\mathbf{c}_2 + \cdots + A_{n2}\mathbf{c}_n, \\ &\vdots \\ f(\mathbf{b}_n) &= A_{1n}\mathbf{c}_1 + A_{2n}\mathbf{c}_2 + \cdots + A_{nn}\mathbf{c}_n, \end{aligned}$$

The *matrix of f with respect to bases B and C* is the matrix

$$[f]_{CB} = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{pmatrix}$$

Definition (Change of basis matrix)

Let V be and $\mathbb{F}$-vector space. Let $B = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ be a basis of V and let $C = \{\mathbf{c}_1, \ldots, \mathbf{c}_n\}$ be another basis of V. Let

$$\begin{aligned}
\mathbf{b}_1 &= A_{11}\mathbf{c}_1 + A_{21}\mathbf{c}_2 + \cdots + A_{n1}\mathbf{c}_n, \\
\mathbf{b}_2 &= A_{12}\mathbf{c}_1 + A_{22}\mathbf{c}_2 + \cdots + A_{n2}\mathbf{c}_n, \\
&\vdots \\
\mathbf{b}_n &= A_{1n}\mathbf{c}_1 + A_{2n}\mathbf{c}_2 + \cdots + A_{nn}\mathbf{c}_n,
\end{aligned}$$

The *change of basis matrix from B to C* is

$$[I]_{CB} = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{pmatrix}$$

The change of basis matrix is the matrix of the identity transformation I with respect to the basis B and C.

Let $T: U \to V$ and $f: V \to W$ be linear transformations. Let

B be a basis of U, C a basis of V D a basis of W.

Then

$$[f \circ T]_{DB} = [f]_{DC}[T]_{CB}.$$

Let $T: V \to W$ be a linear transformation.

S be a basis of V.	C be a basis of W,
B be another basis of V,	D be another basis of W.

Then

$$[I]_{DC}[T]_{CB}[I]_{BS} = [T]_{DS} \quad \text{and} \quad [I]_{SB}[I]_{BS} = [I]_{SS} = 1.$$

This last equation tells us that $[I]_{SB}$ is invertible. Since invertible matrices must be square then B and S have the same number of elements.

Theorem

Let V be an $\mathbb{F}$-vector space. Any two bases of V have the same number of elements.

Example LT2&14. The derivative with respect to x is the linear transformation $T\colon \mathbb{R}[x]_{\leqslant 3} \to \mathbb{R}[x]_{\leqslant 2}$ given by

$$T(a_0 + a_1x + a_2x^2 + a_3x^3) = a_1 + 2a_2x + 3a_3x^2.$$

Since

$$\begin{aligned}
T(1) &= T(1 + 0x + 0x^2 + 0x^3) = 0 + 0x + 0x^2,\\
T(x) &= T(0 + 1x + 0x^2 + 0x^3) = 1 + 0x + 0x^2,\\
T(x^2) &= T(0 + 0x + 1x^2 + 0x^3) = 0 + 2x + 0x^2,\\
T(x^3) &= T(0 + 0x + 0x^2 + 1x^3) = 0 + 0x + 3x^2,
\end{aligned}$$

then the matrix of T with respect to the basis $S = \{1, x, x^2, x^3\}$ of $\mathbb{R}[x]_{\leqslant 3}$ and the basis $B = \{1, x, x^2\}$ of $\mathbb{R}[x]_{\leqslant 2}$ is

$$[T]_{BS} = \begin{pmatrix} 0 & 1 & 0 & 0\\ 0 & 0 & 2 & 0\\ 0 & 0 & 0 & 3 \end{pmatrix}.$$

Then

$$\begin{aligned}
\ker(T) &= \{p \in \mathbb{R}[x]_{\leqslant 3} \mid T(p) = 0\} \\
&= \left\{ a_0 + a_1x + a_2x^2 + a_3x^3 \;\middle|\; \begin{array}{c} a_1 + 2a_2x + 3a_3x^2 \\ = 0 + 0x + 0x^2 + 0x^3 \end{array} \right\} \\
&= \{a_0 + a_1x + a_2x^2 + a_3x^3 \mid a_1 = 0 \text{ and } a_2 = 0 \text{ and } a_3 = 0\} \\
&= \{a_0 + 0x + 0x^2 + 0x^3 \mid a_0 \in \mathbb{R}\} \\
&= \{a_0 \mid a_0 \in \mathbb{R}\} = \mathbb{R}\text{-span}\{1\}
\end{aligned}$$

and

$$\begin{aligned}
\operatorname{im}(T) &= \{T(p) \mid p \in \mathbb{R}[x]_{\leqslant 3}\} \\
&= \{T(a_0 + a_1x + a_2x^2 + a_3x^3) \mid a_0, a_1, a_2, a_3 \in \mathbb{R}\} \\
&= \{a_1 + 2a_2x + 3a_3x^2 \mid a_1, a_2, a_3 \in \mathbb{R}\} = \mathbb{R}[x]_{\leqslant 2},
\end{aligned}$$

Example LT4. Let $T\colon \mathbb{R}^3 \to \mathbb{R}^2$ be the function given by

$$T(x_1, x_2, x_3) = |x_2 - 2x_3, 3x_1 + x_3\rangle = \begin{pmatrix} 0 & 1 & -2 \\ 3 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

With respect to the basis $S = \{|1,0,0\rangle, |0,1,0\rangle, |0,0,1\rangle\}$ of $\mathbb{R}^3$ and the basis $B = \{|1,0\rangle, |0,1\rangle\}$ of $\mathbb{R}^2$ the matrix of T is

$$[T]_{BS} = \begin{pmatrix} 0 & 1 & -2 \\ 3 & 0 & 1 \end{pmatrix}.$$

Example LT11. Let $T\colon M_{2\times 2}(\mathbb{R}) \to M_{2\times 2}(\mathbb{R})$ be the linear transformation given by

$$T(Q) = Q^t.$$

Find the matrix of T with respect to the basis $B = \{E_{11}, E_{12}, E_{21}, E_{22}\}$, where E_{ij} is the matrix with 1 in the (i,j) entry and 0 elsewhere.
Since

$$\begin{aligned}
T(E_{11}) &= E_{11} = 1\cdot E_{11} + 0\cdot E_{12} + 0\cdot E_{21} + 0\cdot E_{22},\\
T(E_{12}) &= E_{21} = 0\cdot E_{11} + 0\cdot E_{12} + 1\cdot E_{21} + 0\cdot E_{22},\\
T(E_{21}) &= E_{12} = 0\cdot E_{11} + 1\cdot E_{12} + 0\cdot E_{21} + 0\cdot E_{22},\\
T(E_{22}) &= E_{12} = 0\cdot E_{11} + 0\cdot E_{12} + 0\cdot E_{21} + 1\cdot E_{22},
\end{aligned}$$

then the matrix of T with respect to the basis $B = \{E_{11}, E_{12}, E_{21}, E_{22}\}$ is

$$[T]_{BB} = \begin{pmatrix} 1 & 0 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Example LT12. Let $T\colon \mathbb{R}[x]_{\leqslant 2} \to \mathbb{R}[x]_{\leqslant 1}$ be the linear transformation given by

$$T(a_0 + a_1x + a_2x^2) = (a_0 + a_2) + a_0x.$$

(a) Find the matrix of T with respect to the basis $B = \{1, x, x^2\}$ of $\mathbb{R}[x]_{\leqslant 2}$ and the basis $C = \{1, x\}$ of $\mathbb{R}[x]_{\leqslant 1}$.

(b) Find the matrix of T with respect to the basis $B = \{1, x, x^2\}$ of $\mathbb{R}[x]_{\leqslant 2}$ and the basis $D = \{2, 3x\}$ of $\mathbb{R}[x]_{\leqslant 1}$.

Let $b_1 = 1$, $b_2 = x$, $b_3 = x^2$ and $c_1 = 2$, $c_2 = 3x$ and $d_1 = 1$, $d_2 = x$. Since

$$\begin{aligned}
T(1) &= 1 + x && = 1 \cdot 1 + 1 \cdot x && = \tfrac{1}{2} \cdot 2 + \tfrac{1}{3} \cdot (3x), \\
T(x) &= 0 && = 0 \cdot 1 + 0 \cdot x && = 0 \cdot 2 + 0 \cdot (3x), \\
T(x^2) &= 1 && = 1 \cdot 1 + 0 \cdot x && = \tfrac{1}{2} \cdot 2 + 0 \cdot (3x),
\end{aligned}$$

then

$$\begin{aligned}
T(b_1) &= 1 \cdot d_1 + 1 \cdot d_2, \\
T(b_2) &= 0 \cdot d_1 + 0 \cdot d_2, \\
T(b_3) &= 1 \cdot d_1 + 0 \cdot d_2,
\end{aligned}
\qquad \text{and} \qquad
\begin{aligned}
T(b_1) &= \tfrac{1}{2} c_1 + \tfrac{1}{3} c_2, \\
T(b_2) &= 0 \cdot c_1 + 0 \cdot c_2, \\
T(b_3) &= \tfrac{1}{2} \cdot c_1 + 0 \cdot c_2,
\end{aligned}$$

and

$$[T]_{DB} = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \qquad \text{and} \qquad [T]_{CB} = \begin{pmatrix} \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{3} & 0 & 0 \end{pmatrix}.$$

Example LT13. Suppose that $T\colon \mathbb{R}^3 \to \mathbb{R}^2$ is a linear transformation and that the matrix of T with respect to
the basis $A = \{|1,0,0\rangle, |0,1,0\rangle, |0,0,1\rangle\}$ of $\mathbb{R}^3$ and
the basis $S = \{|1,0\rangle, |0,1\rangle\}$ of $\mathbb{R}^2$ is

$$[T]_{SA} = \begin{pmatrix} 5 & 1 & 0 \\ 1 & 5 & -2 \end{pmatrix}.$$

Find the matrix of T with respect to
the basis $B = \{|1,1,0\rangle, |1,-1,0\rangle, |1,-1,-2\rangle\}$ of $\mathbb{R}^3$ and
the basis $C = \{|1,1\rangle, |1,-1\rangle\}$ of $\mathbb{R}^2$.

The answer is

$$[T]_{CB} = \begin{pmatrix} 6 & 0 & 2 \\ 0 & 4 & 2 \end{pmatrix}$$

since

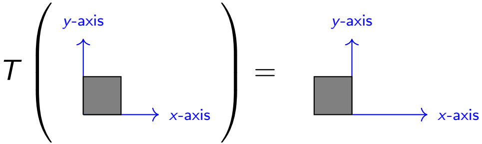
$$T(1,1,0) = \begin{pmatrix} 5 & 1 & 0 \\ 1 & 5 & -2 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 6 \\ 6 \end{pmatrix} = 6 \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} + 0 \cdot \begin{pmatrix} 1 \\ -1 \end{pmatrix},$$

$$T(1,-1,0) = \begin{pmatrix} 5 & 1 & 0 \\ 1 & 5 & -2 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} = \begin{pmatrix} 4 \\ -4 \end{pmatrix} = 0 \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} + 4 \cdot \begin{pmatrix} 1 \\ -1 \end{pmatrix},$$

$$T(1,-1,-2) = \begin{pmatrix} 5 & 1 & 0 \\ 1 & 5 & -2 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \\ -2 \end{pmatrix} = \begin{pmatrix} 4 \\ 0 \end{pmatrix} = 2 \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} + 2 \cdot \begin{pmatrix} 1 \\ -1 \end{pmatrix},$$

Lecture 22: Picturing linear transformations $T\colon \mathbb{R}^2 \to \mathbb{R}^2$

Example LT5. Let $T\colon \mathbb{R}^2 \to \mathbb{R}^2$ be the transformation which is reflection across the y-axis.



Then

$$T\begin{pmatrix}1\\0\end{pmatrix} = \begin{pmatrix}-1\\0\end{pmatrix} = (-1)\cdot\begin{pmatrix}1\\0\end{pmatrix} + 0\cdot\begin{pmatrix}0\\1\end{pmatrix} \quad \text{and}$$

$$T\begin{pmatrix}0\\1\end{pmatrix} = \begin{pmatrix}0\\1\end{pmatrix} = 0\cdot\begin{pmatrix}1\\0\end{pmatrix} + 1\cdot\begin{pmatrix}0\\1\end{pmatrix}$$

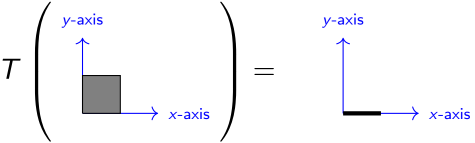
then the matrix of T with respect to the basis $S = \{|1,0\rangle,\ |0,1\rangle\}$ is

$$[T]_{SS} = \begin{pmatrix}-1 & 0\\ 0 & 1\end{pmatrix}.$$

Example LT19. Find the matrix of the linear transformation

$$T\colon \mathbb{R}^2 \to \mathbb{R}^2 \quad \text{which is projection onto the } x \text{ axis.}$$

Is T injective? Is T surjective. Is T invertible?



Since

$$T\begin{pmatrix}1\\0\end{pmatrix} = \begin{pmatrix}1\\0\end{pmatrix} = 1\cdot\begin{pmatrix}1\\0\end{pmatrix} + 0\cdot\begin{pmatrix}0\\1\end{pmatrix}$$

and

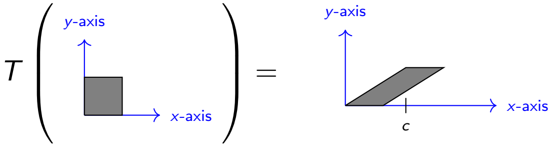
$$T\begin{pmatrix}0\\1\end{pmatrix} = \begin{pmatrix}0\\0\end{pmatrix} = 0\cdot\begin{pmatrix}1\\0\end{pmatrix} + 0\cdot\begin{pmatrix}0\\1\end{pmatrix}$$

then the matrix of T with respect to the basis $S = \{|1,0\rangle, |0,1\rangle\}$ is

$$[T]_{SS} = \begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix}.$$

The linear transformation T is not injective, not surjective, not invertible.

Example LT9. Let $c \in \mathbb{R}_{>0}$ and let $T\colon \mathbb{R}^2 \to \mathbb{R}^2$ be the linear transformation which is the shear by a factor of c along the x-axis.



Since

$$T\begin{pmatrix}1\\0\end{pmatrix} = \begin{pmatrix}1\\0\end{pmatrix} = 1\cdot\begin{pmatrix}1\\0\end{pmatrix} + 0\cdot\begin{pmatrix}0\\1\end{pmatrix}$$

and

$$T\begin{pmatrix}0\\1\end{pmatrix} = \begin{pmatrix}c\\1\end{pmatrix} = c\cdot\begin{pmatrix}1\\0\end{pmatrix} + 1\cdot\begin{pmatrix}0\\1\end{pmatrix}$$

then the matrix of T with respect to the basis $S = \{|1,0\rangle, |0,1\rangle\}$ is

$$[T]_{SS} = \begin{pmatrix}1 & c\\0 & 1\end{pmatrix}.$$

The linear transformation T is invertible and

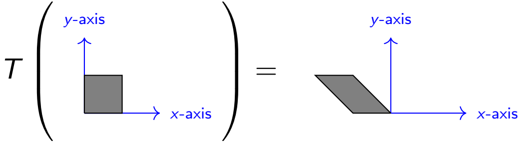
$$\ker(T) = \{0\} \quad \text{and} \quad \operatorname{im}(T) = \mathbb{R}^2.$$

Example LT10. Find the image of $(x, y) \in \mathbb{R}^2$ after a shear along the x-axis with $c = 1$ followed by a reflection across the y-axis.

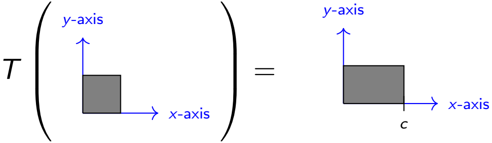
$$T\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -1 & -1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -x - y \\ y \end{pmatrix}.$$

So

$$\begin{aligned} \operatorname{im}(T) &= \left\{ x\begin{pmatrix} -1 \\ 0 \end{pmatrix} + y\begin{pmatrix} -1 \\ 1 \end{pmatrix} \;\middle|\; x, y \in \mathbb{R} \right\} \\ &= \mathbb{R}\text{-span}\left\{ \begin{pmatrix} -1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \end{pmatrix} \right\} = \mathbb{R}^2. \end{aligned}$$



Example LT8. Let $c \in \mathbb{R}_{>0}$ and let $T\colon \mathbb{R}^2 \to \mathbb{R}^2$ be the linear transformation which is stretching of the x-axis by a factor of c.



Since

$$T\begin{pmatrix}1\\0\end{pmatrix} = \begin{pmatrix}c\\0\end{pmatrix} = c\cdot\begin{pmatrix}1\\0\end{pmatrix} + 0\cdot\begin{pmatrix}0\\1\end{pmatrix}$$

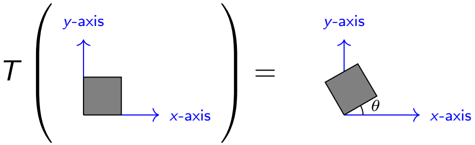
and

$$T\begin{pmatrix}0\\1\end{pmatrix} = \begin{pmatrix}0\\1\end{pmatrix} = 0\cdot\begin{pmatrix}1\\0\end{pmatrix} + 1\cdot\begin{pmatrix}0\\1\end{pmatrix}$$

then the matrix of T with respect to the basis $B = \{(1,0),(0,1)\}$ is

$$[T]_{BB} = \begin{pmatrix}c & 0\\ 0 & 1\end{pmatrix}.$$

Example LT7&18. Let $T\colon \mathbb{R}^2 \to \mathbb{R}^2$ be the linear transformation which is rotation by θ (about the origin counterclockwise).



Since

$$T\begin{pmatrix}1\\0\end{pmatrix} = \begin{pmatrix}\cos\theta\\ \sin\theta\end{pmatrix} = \cos\theta\cdot\begin{pmatrix}1\\0\end{pmatrix} + \sin\theta\cdot\begin{pmatrix}0\\1\end{pmatrix} \qquad \text{and}$$

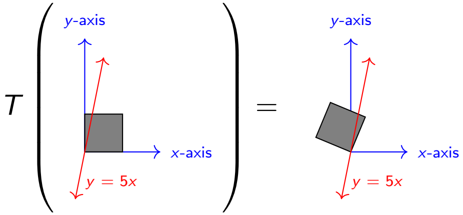
$$T\begin{pmatrix}0\\1\end{pmatrix} = \begin{pmatrix}-\sin\theta\\ \cos\theta\end{pmatrix} = -\sin\theta\cdot\begin{pmatrix}1\\0\end{pmatrix} + \cos\theta\cdot\begin{pmatrix}0\\1\end{pmatrix}$$

then the matrix of T with respect to the basis $S = \{(1,0),(0,1)\}$ is

$$[T]_{SS} = \begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix} \qquad \text{and} \qquad [T^{-1}]_{SS} = \begin{pmatrix}\cos\theta & \sin\theta\\ -\sin\theta & \cos\theta\end{pmatrix}$$

is the matrix of the rotation by $-\theta$ with respect to the basis S.

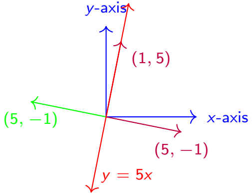
Example EV1. Let $T\colon \mathbb{R}^2 \to \mathbb{R}^2$ be the reflection in the line $y = 5x$.



Identify two lines through the origin that are invariant under T and find the image of the direction vectors for each of these lines.

Let

$$B = \{(1,5),(5,-1)\},$$



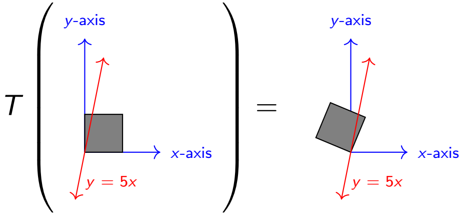
One line is the line $y = 5x$ and the other line is the line orthogonal to $y = 5x$. The line $y = 5x$ has slope 5 and the line orthogonal to $y = 5x$ has slope $-\frac{1}{5}$ and equation $y = -\frac{1}{5}x$. The corresponding direction vectors of these lines are $(1,5)$ and $(1,-\frac{1}{5})$ and

$$T(1,5) = (1,5) \qquad \text{and} \qquad T(1,-\tfrac{1}{5}) = -(1,-\tfrac{1}{5}) = (-1,\tfrac{1}{5}).$$

If $\mathbf{v}_1 = (1,5)$ and $v_2 = (1,-\frac{1}{5})$ then

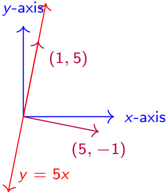
$$Tv_1 = 1 \cdot v_1 \quad \text{and} \quad Tv_2 = (-1) \cdot v_2.$$

Example LT6&23. Let $T: \mathbb{R}^2 \to \mathbb{R}^2$ be the linear transformation which is reflection in the line $y = 5x$.



Let

$$B = \{(1,5),(5,-1)\},$$



so that the first vector in B is a vector in the direction of the line $y = 5x$ and the second vector in B is a vector perpendicular to $y = 5x$.

Since

$$[T]_{BB} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \qquad [I]_{SB} = \begin{pmatrix} 1 & 5 \\ 5 & -1 \end{pmatrix}$$

and

$$[I]_{BS} = ([I]_{SB})^{-1} = -\tfrac{1}{26} \begin{pmatrix} -1 & -5 \\ -5 & 1 \end{pmatrix}$$

then the matrix of T with respect to the basis $S = \{(1,0),(0,1)\}$ is

$$[T]_{SS} = [I[_{SB}[T]_{BB}[I]_{BS} = -\tfrac{1}{26} \begin{pmatrix} 1 & 5 \\ 5 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} -1 & -5 \\ -5 & 1 \end{pmatrix}$$

$$= -\tfrac{1}{26} \begin{pmatrix} 1 & 5 \\ 5 & -1 \end{pmatrix} \begin{pmatrix} -1 & -5 \\ 5 & -1 \end{pmatrix} = -\tfrac{1}{26} \begin{pmatrix} 24 & -10 \\ -10 & -24 \end{pmatrix} = \begin{pmatrix} -\frac{12}{13} & \frac{5}{13} \\ \frac{5}{13} & \frac{12}{13} \end{pmatrix}$$

In other words,

$$T \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} -\frac{12}{13} \\ \frac{5}{13} \end{pmatrix} = -\frac{12}{13} \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \frac{5}{13} \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

and

$$T \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{5}{13} \\ \frac{12}{13} \end{pmatrix} = \frac{5}{13} \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \frac{12}{13} \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Example LT20,21&24. Let $T\colon \mathbb{R}^2 \to \mathbb{R}^2$ be the linear transformation given by

$$T(x, y) = (3x - y, -x + 3y).$$

and let B and S be the bases of $\mathbb{R}^2$ given by

$$B = \{(1, 1), (-1, 1)\} \qquad \text{and} \qquad S = \{(1, 0), (0, 1)\}.$$

Let u and v be the vectors in $\mathbb{R}^2$ determined by

$$[u]_B = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \qquad \text{and} \qquad [v]_S = \begin{pmatrix} 0 \\ 2 \end{pmatrix}.$$

The change of basis matrices between B and S are

$$[I]_{BS} = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \qquad \text{and} \qquad [I]_{SB} = ([I]_{BS})^{-1} = \tfrac{1}{2}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}.$$

Then

$$[u]_S = \begin{pmatrix} 0 \\ 2 \end{pmatrix} \qquad \text{and} \qquad [v]_B = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

The matrices of T with respect to S and B are

$$[T]_{SS} = \begin{pmatrix} 3 & -1 \\ -1 & 3 \end{pmatrix} \qquad \text{and} \qquad [T]_{BB} = \begin{pmatrix} 2 & 0 \\ 0 & 4 \end{pmatrix}$$

so that T stretches by a factor of 2 in the direction $|1, 1\rangle$ and T stretches by a factor of 4 in the direction $|1, -1\rangle$.

Lecture 23: Inner product spaces

Definition (Inner product)

Let $\mathbb{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{C}$. Let

$$\begin{aligned} \overline{}: \mathbb{C} \to \mathbb{C} &\quad \text{be given by} \quad \overline{a+bi} = a - bi, \\ \overline{}: \mathbb{R} \to \mathbb{R} &\quad \text{be given by} \quad \overline{a} = a. \end{aligned}$$

Let V be an $\mathbb{F}$-vector space. An *inner product on V* is a function

$$\begin{array}{rccc} \langle , \rangle : & V \times V & \to & \mathbb{F} \\ & (u, v) & \mapsto & \langle u, v \rangle \end{array} \quad \text{such that}$$

(1) If $u, v \in V$ then $\langle u, v \rangle = \overline{\langle v, u \rangle}$,

(2) If $u, v \in V$ and $\alpha \in \mathbb{F}$ then $\langle \alpha u, v \rangle = \alpha \langle u, v \rangle$,

(3) if $u, v, w \in V$ then $\langle u + v, w \rangle = \langle u, w \rangle + \langle v, w \rangle$,

(4) (Positive semi-definite) If $u \in V$ then $\langle u, u \rangle \in \mathbb{R}_{\geqslant 0}$.

(5) (Positive definite) If $u \in V$ and $\langle u, u \rangle = 0$ then $u = 0$.

Definition (Length, distance, angles.)

Let V be an $\mathbb{F}$-vector space with an inner product.
Length is the function $\| \ \| \colon V \to \mathbb{R}_{\geqslant 0}$ given by

$$\|u\|^2 = \langle u, u \rangle.$$

Distance is the function $d \colon V \times V \to \mathbb{R}_{\geqslant 0}$ given by

$$d(u, v) = \|v - u\|.$$

Angle is the function $\theta \colon V \times V \to \mathbb{R}_{[0,\pi]}$ given by

$$\cos(\theta(u, v)) = \frac{\mathrm{Re}(\langle u, v \rangle)}{\|u\| \cdot \|v\|}.$$

Vecotrs u, v are *orthogonal* if $\langle u, v \rangle = 0$.

Definition (Standard inner products.)

(1) $$\begin{array}{rccc} \langle,\rangle: & \mathbb{R}^n \times \mathbb{R}^n & \to & \mathbb{R} \\ & (u,v) & \mapsto & \langle u|v\rangle \end{array}$$ given by

$$\langle u_1, u_2, \ldots, u_n | v_1, v_2, \ldots, v_n\rangle = u_1 v_1 + \cdots + u_n v_n,$$

(2) $$\begin{array}{rccc} \langle,\rangle: & \mathbb{C}^n \times \mathbb{C}^n & \to & \mathbb{C} \\ & (u,v) & \mapsto & \langle u|v\rangle \end{array}$$ given by

$$\langle u_1, u_2, \ldots, u_n | v_1, v_2, \ldots, v_n\rangle = u_1 \overline{v_1} + \cdots + u_n \overline{v_n},$$

(3) $\mathbb{F}[x]_{\leqslant n} \times \mathbb{F}[x]_{\leqslant n} \to \mathbb{F}$ given by

$$\langle a_0 + a_1 x + \cdots + a_n x^n \mid c_1 + c_1 x + \cdots + c_n x^n\rangle$$
$$= \int_0^1 (a_0 + a_1 x + \cdots a_n x^n)(\bar{c}_0 + \bar{c}_1 x + \cdots + \bar{c}_n x^n)\, dx.$$

Example IP5. Let $u = |1+i, 1 = i\rangle$ and $v = |i, 1\rangle$ in $\mathbb{C}^2$ with the standard inner product. Then

$$\begin{aligned}
\langle u|v\rangle &= \langle 1+i, 1-i|i, 1\rangle = (1+i)\bar{i} + (1-i)\bar{1} \\
&= (1-i)(-i) + 1 - i = -i + 1 + 1 - i = 2 - 2i, \\
\langle v|u\rangle &= \langle i, 1|1+i, 1-i\rangle = i\overline{(1+i)} + 1 \cdot \overline{(1-i)} \\
&= i(1-i) + 1 + i = i + 1 + 1 + i = 2 + 2i, \\
\langle u|u\rangle &= \langle 1+i, 1-i|1+i, 1-i\rangle = (1+i)\overline{(1+i)} + (1-i)\overline{(1-i)} \\
&= (1+i)(1-i) + (1-i)(1+i) = 1 + 1 + 1 + 1 = 4, \\
\langle u|u\rangle &= \langle i, 1 \mid i, 1\rangle = i \cdot \bar{i} + 1 \cdot \bar{1} = i(-i) + 1 = 1 + 1 = 2, \\
d(u,v) &= \sqrt{\langle 1, -i|1, -i\rangle} = \sqrt{1 \cdot \bar{1} + (-i)\overline{(-i)}} = \sqrt{1+1} = \sqrt{2}, \\
\cos(\theta(u,v)) &= \frac{\mathrm{Re}(\langle u|v\rangle)}{\|u\| \cdot \|v\|} = \frac{\mathrm{Re}(2-2i)}{2\sqrt{2}} = \frac{2}{2\sqrt{2}} = \frac{1}{\sqrt{2}},
\end{aligned}$$

So $\theta(u,v) = \frac{\pi}{4}$.

Example IP7. Let $u = 1$ and $v = x$ in $\mathbb{R}[x]_{\leqslant 2}$ with the standard inner product. Then

$$\langle u|u\rangle = \langle 1|1\rangle = \int_0^1 dx = x\Big]_0^1 = 1,$$

$$\langle v|v\rangle = \langle x|x\rangle = \int_0^1 x^2\,dx = \tfrac{1}{3}x^3\Big]_0^1 = \tfrac{1}{3} - 0 = \tfrac{1}{3},$$

$$\langle u|v\rangle = \langle 1|x\rangle = \int_0^1 x\,dx = \tfrac{1}{2}x^2\Big]_0^1 = \tfrac{1}{2} - 0 = \tfrac{1}{2},$$

$$d(u,v) = \sqrt{\langle x-1|x-1\rangle} = \sqrt{\int_0^1 (x-1)^2\,dx}$$

$$= \sqrt{\tfrac{1}{3}(x-1)^3\Big]_0^1} = \sqrt{0 - \tfrac{1}{3}(-1)^3} = \sqrt{\tfrac{1}{3}} = \tfrac{1}{\sqrt{3}},$$

$$\cos(\theta(u,v)) = \frac{\mathrm{Re}(\langle u|v\rangle)}{\|u\|\cdot\|v\|} = \frac{\frac{1}{2}}{1\cdot\frac{1}{\sqrt{3}}} = \frac{\sqrt{3}}{2}.$$

So $\theta(u,v) = \frac{\pi}{6}$.

Example IP2. The map $\langle , \rangle \colon \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}$ given by

$$\begin{aligned}\langle (u_1, u_2, u_3), (v_1, v_2, v_3) \rangle &= u_1 v_1 - u_2 v_2 + u_3 v_3 \\ &= \begin{pmatrix} v_1 & v_2 & v_3 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix}\end{aligned}$$

has

$$\langle (0,1,0), (0,1,0) \rangle = 0 - 1 + 0 = -1 \notin \mathbb{R}_{\geqslant 0}.$$

So $\langle , \rangle$ is not positive definite.

Example IP6. The map $\langle , \rangle \colon \mathbb{C}^2 \times \mathbb{C}^2 \to \mathbb{R}$ given by

$$\langle (u_1, u_2), (v_1, v_2) \rangle = i u_1 \overline{v_1} - i u_2 \overline{v_2} = \begin{pmatrix} \overline{v_1} & \overline{v_2} \end{pmatrix} \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$$

has

$$\langle (1,0), (1,0) \rangle = i \cdot 1 \cdot \bar{1} - i \cdot 0 \cdot \bar{0} = i \notin \mathbb{R}_{\geqslant 0}.$$

So $\langle , \rangle$ is not positive definite.

Example IP3. The map $\langle , \rangle \colon \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ given by

$$\begin{aligned}\langle (u_1, u_2), (v_1, v_2) \rangle &= 2u_1v_1 - 2u_1v_2 - 2u_2v_1 + u_2v_2 \\ &= \begin{pmatrix} v_1 & v_2 \end{pmatrix} \begin{pmatrix} 2 & -2 \\ -2 & 3 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}\end{aligned}$$

has

$$\begin{aligned}\langle (u_1, u_2), (u_1, u_2) \rangle &= 2u_1^2 - 2u_2u_2 - 2u_2u_1 + 3u_2^2 \\ &= 2u_1^2 - 4u_1u_2 + 3u_2^2 \\ &= 2(u_1^2 - 2u_1u_2 + u_2^2) + u_2^2 \\ &= (u_1 - u_2)^2 + u_2^2 \in \mathbb{R}_{\geqslant 0}.\end{aligned}$$

Assume $\langle (u_1, u_2), (u_1, u_2) \rangle = 0$.
Then $2(u_1 - u_2)^2 + u_2^0 = 0$ then $u_2^2 = 0$ and $(u_1 - u_2)^2 = 0$.
So $u_2 = 0$ and $u_1 = u_2 = 0$. So $(u_1, u_2) = 0$.

So $\langle , \rangle$ is positive definite.

Example IP1. The map $\langle , \rangle \colon \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ given by

$$\langle (u_1, u_2), (v_1, v_2) \rangle = u_1 v_1 + 2u_2 v_2 = \begin{pmatrix} v_1 & v_2 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$$

has $\langle (u_1, u_2), (u_1, u_2) \rangle = u_1^2 + 2u_2^2 \in \mathbb{R}_{\geqslant 0}$.

Assume $\langle (u_1, u_2), (u_1, u_2) \rangle = 0$.

Then $u_1^2 + 2u_2^0 = 0$ so that $u_1^2 = 0$ and $2u_2^2 = 0$.

So $(u_1, u_2) = 0$.

So $\langle , \rangle$ is positive definite.

Example IP8. Let V be an $\mathbb{F}$-vector space with an inner product. Let $u, v \in V$ and suppose that u and v are orthogonal. Then

$$\begin{aligned} \|u + v\|^2 = \langle u + v, u + v \rangle &= \langle u, u + v \rangle + \langle v, u + v \rangle \\ &= \langle u, u \rangle + \langle u, v \rangle + \langle v, u \rangle + \langle v, v \rangle \\ &= \|u\|^2 + 0 + 0 + \|v\|^2 \\ &= \|u\|^2 + \|v\|^2. \end{aligned}$$

This is the Pythagorean theorem.

Example IPA1. Let $A \in M_{n\times n}(\mathbb{C})$ satisifying $A = \bar{A}^t$ and let $\langle , \rangle \colon \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C}$ be given by

$$\langle (u_1, \dots, u_n), (v_1, \dots, v_n) \rangle = (u_1, \dots, u_n) A \begin{pmatrix} \overline{v_1} \\ \vdots \\ \overline{v_n} \end{pmatrix} = u^t A \bar{v},$$

If $u, v \in C^n$ then

$$\begin{aligned} \overline{\langle v, u \rangle} = \overline{(v^t A \bar{u})} = \overline{(\bar{u}^t A^t v)}^t = (u^t \bar{A}^t \bar{v})^t \\ = (u^t A \bar{v})^t = \langle u, v \rangle, \end{aligned}$$

and if $\alpha \in \mathbb{C}$ and $u, v \in \mathbb{C}^n$ then

$$\langle \alpha u, v \rangle = (\alpha u)^t A \bar{v} = \alpha u^t A v = \alpha \langle u, v \rangle$$

and, if $u, v, w \in \mathbb{C}^n$ then

$$\begin{aligned} \langle u + v, w \rangle &= (u + v)^t A \bar{w} = (u^t + v^t) A \bar{w} \\ &= u^t A \bar{w} + v^t A \bar{w} = \langle u, w \rangle + \langle v, w \rangle. \end{aligned}$$

So $\langle , \rangle$ satisfies all the properties of an inner product, except perhaps the positive definiteness.

Lecture 24: Gram matrices, orthogonality and projections

Definition (Gram matrix.)

Let V be an $\mathbb{F}$-vector space with inner product $\langle , \rangle V \times V \to \mathbb{F}$.
Let $B = \{b_1, \dots, b_n\}$ be a basis of V.
The *Gram matrix of $\langle , \rangle$ with respect to B* is the matrix

$$A = \begin{pmatrix} \langle b_1, b_1 \rangle & \langle b_1, b_2 \rangle & \cdots & \langle b_1, b_n \rangle \\ \langle b_2, b_1 \rangle & \langle b_2, b_2 \rangle & \cdots & \langle b_2, b_n \rangle \\ \vdots & & & \vdots \\ \langle b_n, b_1 \rangle & \langle b_n, b_2 \rangle & \cdots & \langle b_n, b_n \rangle \end{pmatrix}$$

In other words, the (i,j) entry of the Gram matrix A of $\langle , \rangle$ with respect to the basis B is

$$A_{ij} = \langle b_i, b_j \rangle.$$

Example IPA2. Let V be a $\mathbb{F}$-vector space with an inner product $\langle , \rangle \colon V \times V \to \mathbb{F}$. Let $B = \{b_1, \ldots, b_n\}$ be a basis of V and let

A be the Gram matrix of $\langle , \rangle$ with respect to the basis B.

Let $u = u_1 b_1 + \cdots + u_n b_n \in V$ and let $v = v_1 b_1 + \cdots + v_n b_n \in V$ so that

$$[u]_B = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix} \quad \text{and} \quad [v]_B = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}.$$

Then

$$\begin{aligned} \langle u, v \rangle &= \langle u_1 b_1 + \cdots + u_n b_n, v_1 b_1 + \cdots + v_n b_n \rangle \\ &= \sum_{i,j=1}^{n} u_i \overline{v_j} \langle b_i, b_j \rangle = \sum_{i,j=1}^{n} u_i A_{ij} \overline{v_j} \\ &= [u]_B^t A [\overline{v}]_B. \end{aligned}$$

Example IP4. The $\mathbb{R}$-vector space $\mathbb{R}[x]_{\leqslant 2}$ has basis $B = \{1, x, x^2\}$. Since the standard inner product $\langle \mid \rangle \colon \mathbb{R}[x]_{\leqslant 2} \times \mathbb{R}[x]_{\leqslant 2} \to \mathbb{R}$ has

$$\langle p|q\rangle = \int_0^1 pq\, dx$$

then

$$\begin{array}{ccc} \langle 1|1\rangle = 1, & \langle 1|x\rangle = \frac{1}{2}, & \langle 1|x^2\rangle = \frac{1}{3}, \\ \langle x|1\rangle = \frac{1}{2}, & \langle x|x\rangle = \frac{1}{3}, & \langle x|x^2\rangle = \frac{1}{4}, \\ \langle x^2|1\rangle = \frac{1}{3}, & \langle x^2|x\rangle = \frac{1}{4}, & \langle x^2|x^2\rangle = \frac{1}{5}, \end{array}$$

then the Gram matrix of $\langle \mid \rangle$ with respect fo B is

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

If $u = 7 + 3x + 2x^2$ and $v = 5 + x^2$ then

$$[u]_B = \begin{pmatrix} 7 \\ 3 \\ 2 \end{pmatrix} \quad \text{and} \quad [v]_B = \begin{pmatrix} 5 \\ 0 \\ 1 \end{pmatrix}.$$

Then

$$\begin{aligned}\langle u|v\rangle &= \int_0^1 (7+3x+2x^2)(5+x^2)\,dx \\ &= \int_0^1 (35+7x^2+15x+3x^3+10x^2+2x^4)\,dx \\ &= 35+\tfrac{7}{3}+\tfrac{15}{2}+\tfrac{3}{4}+\tfrac{10}{3}+\tfrac{2}{5}\end{aligned}$$

and

$$\begin{aligned}[u]_B^t A[v]_B &= \begin{pmatrix}7 & 3 & 2\end{pmatrix}\begin{pmatrix}1 & \frac{1}{2} & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5}\end{pmatrix}\begin{pmatrix}5 \\ 0 \\ 1\end{pmatrix} \\ &= \begin{pmatrix}7 & 3 & 2\end{pmatrix}\begin{pmatrix}5+\frac{1}{3} \\ \frac{5}{2}+\frac{1}{4} \\ \frac{5}{3}+\frac{1}{5}\end{pmatrix} \\ &= 35+\tfrac{7}{3}+\tfrac{15}{2}+\tfrac{3}{4}+\tfrac{10}{3}+\tfrac{2}{5}.\end{aligned}$$

So $\langle u|v\rangle = [u]_B^t A[v]_B$.

Lecture 25: Projections and orthogonalisation

Definition (Orthogonal and orthonormal sequences.)

Let V be an $\mathbb{F}$-vector space with an inner product $\langle , \rangle \colon V \times V \to \mathbb{F}$.

Let $u, v \in V$. The vectors u and v are

$$\textit{orthogonal} \quad \text{if } \langle u, v \rangle = 0.$$

An *orthogonal sequence* is a sequence $(b_1, \ldots, b_k)$ of vectors in V such that

$$\text{if } i, j \in \{1, \ldots, k\} \text{ and } i \neq j \quad \text{then} \quad \langle b_i, b_j \rangle = 0.$$

An *orthonormal sequence* is an orthogonal sequence $(b_1, \ldots, b_k)$ such that

$$\text{if } i \in \{1, \ldots, k\} \text{ then } \langle b_i, b_i \rangle = 1.$$

An *ordered orthonormal basis of V* is an orthonormal sequence $(b_1, \ldots, b_k)$ in V such that B is a basis of V.

Proposition

Assume $B = (b_1, \ldots, b_n)$ *is an ordered orthonormal basis of* V *and* $x \in V$*. Then*

$$x = \langle x, b_1 \rangle b_1 + \cdots + \langle x, b_n \rangle b_n.$$

Definition (Orthogonal projections.)

Let W be a subspace of V. Let $\{b_1, \ldots, b_k\}$ be an orthonormal basis of W. Let $x \in V$. The *orthogonal projection of x onto W* is

$$\text{proj}_W(x) = \langle x, b_1 \rangle b_1 + \cdots + \langle x, b_k \rangle b_k.$$

Example IP9,10&11 Let $\langle , \rangle \colon \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}$ be given by

$$\langle (u_1, u_2, u_3), (v_1, v_2, v_3) \rangle = u_1 v_1 + 2u_2 v_2 + u_3 v_3.$$

Let

$$S = \{(1, 1, 1), (1 - 1, 1), (1, 0 - 1)\}.$$

Then

$$\begin{aligned} \langle (1, 1, 1), (1, -1, 1) \rangle &= 1 - 2 + 1 = 0, \\ \langle (1, 1, 1), (1, 0, -1) \rangle &= 1 + 0 - 1 = 0, \\ \langle (1, -1, 1), (1, 0, -1) \rangle &= 1 + 0 - 1 = 0, \end{aligned}$$

So S is an orthogonal sequence in $\mathbb{R}^3$ with respect to $\langle , \rangle$.

Let

$$b_1 = \frac{1}{\|u\|}u, \qquad \text{where} \quad u = (1,1,1),$$
$$b_2 = \frac{1}{\|v\|}v, \qquad \text{where} \quad v = (1,-1,1),$$
$$b_3 = \frac{1}{\|w\|}w, \qquad \text{where} \quad w = (1,0,01).$$

Then

$$b_1 = \tfrac{1}{2}(1,1,1) \quad \text{since} \quad \langle (1,1,1),(1,1,1)\rangle = 4,$$
$$b_2 = \tfrac{1}{2}(1,-1,1) \quad \text{since} \quad \langle (1,-1,1),(1,-1,1)\rangle = 4,$$
$$b_2 = \tfrac{1}{\sqrt{2}}(1,0,-1) \quad \text{since} \quad \langle (1,0,-1),(1,0,-1)\rangle = 2,$$

and $\{b_1, b_2, b_3\}$ is an orthonormal sequence in $\mathbb{R}^3$ with respect to $\langle , \rangle$.

Let $x = |1, 1, -1\rangle$ and $c_2, c_2, c_3 \in \mathbb{R}$ such that

$$x = c_1 b_1 + c_2 b_2 + c_3 b_3.$$

Then

$$\begin{aligned} c_1 &= c_1 \langle b_1, b_1 \rangle + 0 + 0 = \langle c_1 b_1 + c_2 b_2 + c_3 b_3, b_1 \rangle = \langle x, b_1 \rangle \\ &= \langle (1, 1, -1), \tfrac{1}{2}(1, 1, 1) \rangle = \tfrac{1}{2}(1 + 2 - 1), \\ c_2 &= c_2 \langle b_2, b_2 \rangle = \langle c_1 b_1 + c_2 b_2 + c_3 b_3, b_2 \rangle = \langle x, b_2 \rangle \\ &= \langle (1, 1, -1), \tfrac{1}{2}(1, -1, 1) \rangle = \tfrac{1}{2}(1 - 2 - 1), \\ c_3 &= \langle x, b_3 \rangle = \langle (1, 1, -1), \tfrac{1}{\sqrt{2}}(1, 0, -1) \rangle = \tfrac{1}{\sqrt{2}}(1 + 0 + 1) = \tfrac{2}{\sqrt{2}} = \sqrt{2}. \end{aligned}$$

So x is written as a linear combination of the basis elements in the form

$$\begin{aligned} x = (1, 1, -1) &= \langle x, b_1 \rangle b_1 + \langle x, b_2 \rangle b_2 + \langle x, b_3 \rangle b_3 \\ &= 1 \cdot b_1 + (-1) \cdot b_2 + \sqrt{2} b_3 \\ &= (1, 1, 1) - (1, -1, 1) + \sqrt{2}(1, 0, -1). \end{aligned}$$

Example IPA3. Let $V = \mathbb{R}^n$ and let $u \in V$ with $u \neq 0$. Let

$$W = \mathbb{R}\text{-span}\{u\} = \{au \mid a \in \mathbb{R}\}.$$

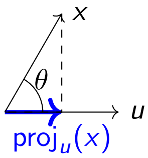
Then W is a 1-dimensional subspace of V. Let

$$b_1 = \frac{1}{\|u\|} u.$$

Then $\{b_1\}$ is an orthonormal basis of W.
Let $x \in V$. Then

$$\text{proj}_W(x) = \langle x, b_1 \rangle b_1 = \langle x, \frac{1}{\|u\|} u \rangle \frac{1}{\|u\|} u$$

$$= \frac{\langle x, u \rangle}{\|u\|^2} u = \text{proj}_u(x).$$



Example IP12. Let $V = \mathbb{R}^3$ and let

$$W = \{|x, y, z\rangle \in \mathbb{R}^3 \mid x + y + z = 0\}.$$

The set

$$\{b_1, b_2\} = \left\{\tfrac{1}{\sqrt{2}}|1, -1, 0\rangle, \tfrac{1}{\sqrt{6}}|1, 1, -2\rangle\right\}$$

is an orthonormal basis of W with respect to the standard inner product on $\mathbb{R}^3$.

Let $x = |1, 2, 3\rangle$. Then

$$\begin{aligned}
\mathrm{proj}_W(x) &= \langle x|b_1\rangle b_1 + \langle x|b_2\rangle b_2 \\
&= \langle 1, 2, 3|\tfrac{1}{\sqrt{2}}, \tfrac{-1}{\sqrt{2}}, 0\rangle \cdot \tfrac{1}{\sqrt{2}}|1, -1, 0\rangle \\
&\quad + \langle 1, 2, 3|\tfrac{1}{\sqrt{6}}, \tfrac{1}{\sqrt{6}}, \tfrac{-2}{\sqrt{6}}\rangle \cdot \tfrac{1}{\sqrt{6}}|1, 1, -2\rangle \\
&= \left(\tfrac{1}{\sqrt{2}} - \tfrac{2}{\sqrt{2}}\right)\tfrac{1}{\sqrt{2}}|1, -1, 0\rangle + \tfrac{1}{6}(1 + 2 - 6)|1, 1, -2)\rangle \\
&= |\tfrac{-1}{2}, \tfrac{1}{2}, 0\rangle + |\tfrac{-1}{2}, \tfrac{-1}{2}, 1) = |-1, 0, 1\rangle.
\end{aligned}$$

The shortest distance from x to W is

$$\begin{aligned}
\|x - \mathrm{proj}_W(x)\| &= \|\, |1, 2, 3\rangle - |-1, 0, 1\rangle \,\| \\
&= \|\, |2, 2, 2\rangle \,\| = \sqrt{4 + 4 + 4} = 2\sqrt{3}.
\end{aligned}$$

Example IP13. *(The Gram-Schmidt process of orthogonalization)*
Let $V = \mathbb{R}^3$ with the standard inner product. Let $S = \{v_1, v_2, v_3\}$ with

$$v_1 = |1, 1, 1\rangle, \quad v_2 = |0, 1, 1\rangle, \quad v_1 = |0, 0, 1\rangle.$$

Convert S into an orthonormal basis B.

Step 1. Make v_1 into a unit vector. Let

$$b_1 = \frac{1}{\|v_1\|} v_1 = \tfrac{1}{\sqrt{3}}|1, 1, 1\rangle = |\tfrac{1}{\sqrt{3}}, \tfrac{1}{\sqrt{3}}, \tfrac{1}{\sqrt{3}}\rangle$$

and let $S = \{b_1, v_2, v_3\}$.

Step 2. Make v_2 orthogonal to b_1. Let

$$\begin{aligned} u_2 &= v_2 - \langle v, b_1\rangle b_1 \\ &= |0, 1, 1\rangle - \tfrac{2}{\sqrt{3}}|\tfrac{1}{\sqrt{3}}, \tfrac{1}{\sqrt{3}}, \tfrac{1}{\sqrt{3}}\rangle \\ &= |\tfrac{-2}{\sqrt{3}}, \tfrac{1}{\sqrt{3}}, \tfrac{1}{\sqrt{3}}\rangle \end{aligned}$$

and let $S_2 = \{b_1, u_2, v_3\}$.

Step 3. Make u_2 into a unit vector. Let

$$b_2 = \frac{1}{\|u_2\|}u_2 = \frac{1}{\sqrt{6}/3}|\tfrac{-2}{3}, \tfrac{1}{3}, \tfrac{1}{3}\rangle = |\tfrac{-2}{\sqrt{6}}, \tfrac{1}{\sqrt{6}}, \tfrac{1}{\sqrt{6}}\rangle$$

and let $S_3 = \{b_1, b_2, v_3\}$.

Step 4. Make v_3 orthogonal to b_1 and b_2. Let

$$\begin{aligned} u_3 &= v_3 - \langle v_3, b_1\rangle b_1 - \langle v_3, b_2\rangle b_2 \\ &= |0, 0, 1\rangle - \tfrac{1}{\sqrt{3}}|\tfrac{1}{\sqrt{3}}, \tfrac{1}{\sqrt{3}}, \tfrac{1}{\sqrt{3}}\rangle - \tfrac{1}{\sqrt{6}}|\tfrac{-2}{\sqrt{6}}, \tfrac{1}{\sqrt{6}}, \tfrac{1}{\sqrt{6}}\rangle \\ &= |\tfrac{-1}{3} + \tfrac{2}{6}, \tfrac{-1}{3} - \tfrac{1}{6}, 1 - \tfrac{1}{3} - \tfrac{1}{6}\rangle = |0, \tfrac{-1}{2}, \tfrac{1}{2}\rangle. \end{aligned}$$

Step 5. Make u_3 into a unit vector. Let

$$b_3 = \frac{1}{\|u_3\|}u_3 = \frac{1}{\sqrt{2/4}}|0, \tfrac{-1}{2}, \tfrac{1}{2}\rangle = |0, \tfrac{-1}{\sqrt{2}}, \tfrac{1}{\sqrt{2}}\rangle.$$

Then

$$B = \{b_1, b_2, b_3\} \quad \text{is an orthonormal set.}$$

Lecture 26: Learning to do proofs – Orthogonality and linear independence

Definition (Orthogonal and orthonormal sequences.)

Let V be an $\mathbb{F}$-vector space with an inner product $\langle , \rangle \colon V \times V \to \mathbb{F}$.

Let $u, v \in V$. The vectors u and v are

$$\textit{orthogonal} \quad \text{if } \langle u, v \rangle = 0.$$

An *orthogonal sequence* is a sequence $(b_1, \ldots, b_k)$ of vectors in V such that

$$\text{if } i, j \in \{1, \ldots, k\} \text{ and } i \neq j \quad \text{then} \quad \langle b_i, b_j \rangle = 0.$$

An *orthonormal sequence* is an orthogonal sequence $(b_1, \ldots, b_k)$ such that

$$\text{if } i \in \{1, \ldots, k\} \text{ then } \langle b_i, b_i \rangle = 1.$$

An *ordered orthonormal basis of V* is an orthonormal sequence $(b_1, \ldots, b_k)$ in V such that B is a basis of V.

Theorem (Pythagorean Theorem)

Let V be a $\mathbb{C}$-vector space with an inner product $\langle , \rangle : V \times V \to \mathbb{C}$. Let $u, v \in V$. If $\langle u, v \rangle = 0$ then

$$\|u + v\|^2 = \|u\|^2 + \|v\|^2.$$

Proof. Assume $\langle u, v \rangle = 0$.
To show $\|u + v\|^2 = \|u\|^2 + \|v\|^2$.

$$\begin{aligned} \|u + v\|^2 &= \langle u + v, u + v \rangle \\ &= \langle u, u \rangle + \langle u, v \rangle + \langle v, u \rangle + \langle v, v \rangle \\ &= \langle u, u \rangle + \langle u, v \rangle + \overline{\langle u, v \rangle} + \langle v, v \rangle \\ &= \|u\|^2 + 0 + \overline{0} + \|v\|^2 \\ &= \|u\|^2 + \|v\|^2. \end{aligned}$$

□

Proposition (Orthogonal sets are linearly independent)

Let V be a vector space with inner product $\langle , \rangle \colon V \times V \to \mathbb{C}$.
Let $B = \{b_1, \ldots, b_k\}$ be an orthogonal set in V.
Then B is linearly independent.

Proof. Assume B is an orthogonal set in V.
To show: B is linearly independent.
To show: If $c_1, \ldots, c_k \in \mathbb{C}$ and $c_1 b_1 + \cdots + c_k b_k = 0$

$$\text{then } c_1 = 0, c_2 = 0, \ldots, c_k = 0.$$

Assume $c_1, \ldots, c_k \in \mathbb{C}$ and $c_1 b_1 + \cdots + c_k b_k = 0$.

To show: $c_1 = 0, c_2 = 0, \ldots, c_k = 0$.

To show: If $i \in \{1, \ldots, k\}$ then $c_i = 0$.

Assume $i \in \{1, \ldots, k\}$. To show: $c_i = 0$.

$$\begin{aligned} 0 &= \langle c_1 b_1 + \cdots + c_k b_k, b_i \rangle \\ &= c_1 \langle b_1, b_i \rangle + \cdots + c_{i-1} \langle b_{i-1}, b_i \rangle + c_i \langle b_i, b_i \rangle \\ &\qquad + c_{i+1} \langle b_{i+1}, b_i \rangle + \cdots + c_k \langle b_k, b_i \rangle \\ &= c_1 \cdot 0 + \cdots + c_{i-1} \cdot 0 + c_i \langle b_i, b_i \rangle \\ &\qquad + c_{i+1} \cdot 0 + \cdots + c_k \cdot 0 \\ &= c_i \langle b_i, b_i \rangle. \end{aligned}$$

Since $\langle , \rangle$ is an inner product and $b_i \neq 0$ then $\langle b_i, b_i \rangle \neq 0$. So

$$c_i = \frac{1}{\langle b_i, b_i \rangle} \cdot 0 = 0.$$

So B is linearly independent. □

Lecture 27: Learning to do proofs – Linear transformations

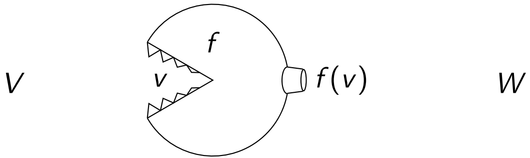
Linear transformations are for comparing vector spaces.

Definition (Linear transformation)

Let $\mathbb{F}$ be a field and let V and W be $\mathbb{F}$-vector spaces. An *$\mathbb{F}$-linear transformation from V to W* is a function $f: V \to W$ such that

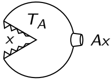
(a) If $v_1, v_2 \in V$ then $f(v_1 + v_2) = f(v_1) + f(v_2)$,

(b) If $c \in \mathbb{F}$ and $v \in V$ then $f(cv) = cf(v)$.



Example A2. Let $t, s \in \mathbb{Z}_{>0}$ and $A \in M_{t\times s}(\mathbb{R})$. Let $T_A\colon \mathbb{R}^s \to \mathbb{R}^t$ be the function given by

$$T_A(x) = Ax.$$



Show that T_A is a linear transformation.
Let $u, v \in \mathbb{R}^s$. Then, by the distributive property of matrix multiplication for matrices,

$$T_A(u+v) = A(u+v) = Au + Av = T_A(u) + T_A(v).$$

Let $u \in \mathbb{R}^s$ and $c \in \mathbb{R}$. Then, by the associative property of scalar multiplication for matrices,

$$T_A(cu) = Acu = cAu = cT_A(u).$$

So T_A is a linear transformation.

Let $T\colon V \to W$ be a linear transformation. Assume that T has an inverse function $T^{-1}\colon W \to V$. Show that T^{-1} is a linear transformation.
Assume $w_1, w_2 \in W$. Then

$$\begin{aligned} T^{-1}(w_1 + w_2) &= T^{-1}(T(T^{-1}(w_1)) + T(T^{-1}(w_2))) \\ &= T^{-1}(T(T^{-1}(w_1) + T^{-1}(w_2)) \\ &= T^{-1}(w_1) + T^{-1}(w_2), \end{aligned}$$

where the first equality is because $T \circ T^{-1} = \mathrm{Id}$), the second equality is because T is a linear transformation) and the third equality is because $T^{-1} \circ T = \mathrm{Id}$. Assume $w \in W$ and $c \in \mathbb{R}$. Then

$$T^{-1}(cw) = T^{-1}(c \cdot T(T^{-1}(w))) = T^{-1}T(c \cdot T^{-1}(w)) = c \cdot T^{-1}(w).$$

So T^{-1} is a linear transformation.

Definition (Kernel and image of a linear transformation)

The *kernel* of an $\mathbb{F}$-linear transformation $f : V \to W$ is the set

$$\ker(f) = \{v \in V \mid f(v) = 0\}.$$

The *image* of an $\mathbb{F}$-linear transformation $f : V \to W$ is the set

$$\operatorname{im}(f) = \{f(v) \mid v \in V\}.$$

Definition (Kernel and image of a matrix)

Let $A \in M_{t\times s}(\mathbb{Q})$. The *kernel of A* is

$$\ker(A) = \{x \in \mathbb{Q}^s \mid Ax = 0\}$$

and the *image of A* is

$$\operatorname{im}(A) = \{Ax \mid s \in \mathbb{Q}^s\}.$$

A *subspace of* $\mathbb{Q}^s$ is a subset $W \subseteq \mathbb{Q}^s$ such that

(a) $0 \in W$,

(b) If $w_1, w_2 \in W$ then $w_1 + w_2 \in W$,

(c) If $w \in W$ and $c \in \mathbb{Q}$ then $cw \in W$.

Proposition

Let $A \in M_{t\times s}(\mathbb{Q})$. Then $\ker(A)$ *is a subspace of* $\mathbb{Q}^s$*.*

Proof. (a) Since $A0 = 0$ then $0 \in \ker(A)$.

(b) Assume $w_1, w_2 \in \ker(A)$. Then $Aw_1 = 0$ and $Aw_2 = 0$. So

$$A(w_1 + w_2) = Aw_1 + Aw_2 = 0 + 0 = 0. \quad \text{So } w_1 + w_2 \in \ker(A).$$

(c) Assume $w \in \ker(A)$ and $c \in \mathbb{Q}$. Then $Aw = 0$ and

$$A(cw) = cAw = c0 = 0. \quad \text{So } cw \in \ker(A).$$

So $\ker(A)$ is a subspace of $\mathbb{Q}^s$. □

A *subspace* of $\mathbb{Q}^t$ is a subset $Y \subseteq \mathbb{Q}^t$ such that

(a) $0 \in Y$,

(b) If $y_1, y_2 \in Y$ then $y_1 + y_2 \in Y$,

(c) If $y \in Y$ and $c \in \mathbb{Q}$ then $cy \in Y$.

Proposition

Let $A \in M_{t\times s}(\mathbb{Q})$. Then $im(A)$ is a subspace of $\mathbb{Q}^t$.

Proof. (a) Since $0 = A0$ then $0 \in \mathrm{im}(A)$.

(b)Assume $y_1, y_2 \in \mathrm{im}(A)$. Then there exist $x_1, x_2 \in \mathbb{Q}^s$ such that $y_1 = Ax_1$ and $y_2 = Ax_2$. Then

$$y_1 + y_2 = Ax_1 + Ax_2 = A(x_1 + x_2). \quad \text{So } y_1 + y_2 \in \mathrm{im}(A).$$

(c) Assume $y \in \mathrm{im}(A)$ and $c \in \mathbb{Q}$. Then there exists $x \in \mathbb{Q}^s$ such that $y = Ax$. Then

$$cy = cAx = A(cx). \quad \text{So } cy \in \mathrm{im}(A).$$

So $\mathrm{im}(A)$ is a subspace of $\mathbb{Q}^t$. □

Example A5. Let $T\colon V \to W$ be an $\mathbb{R}$-linear transformation. Show that $\ker(T) = \{v \in V \mid T(v) = 0\}$ is a subspace of V.

Let $v_1, v_2 \in \ker(T)$. Then

$$T(v_1 + v_2) = T(v_1) + T(v_2) = 0 + 0 = 0. \quad \text{So } v_1 + v_2 \in \ker(T).$$

Subtracting $T(0)$ from each side of the equation $T(0) = T(0+0) = T(0) + T(0)$ gives

$$0 = T(0), \quad \text{and so} \quad 0 \in \ker(T).$$

Let $v \in \ker(T)$ and let $c \in \mathbb{R}$. Then

$$T(cv) = cT(v) = c \cdot 0 = 0 \quad \text{and so} \quad cv \in \ker(T).$$

So $\ker(T)$ is a subspace of V.

Example A6. Let $T\colon V \to W$ be an $\mathbb{R}$-linear transformation. Show that $\text{im}(T) = \{T(v) \mid v \in V\}$ is a subspace of W.
Subtracting $T(0)$ from each side of the equation $T(0) = T(0+0) = T(0) + T(0)$ gives

$$0 = T(0), \quad \text{and so} \quad 0 \in \text{im}(T).$$

Let $w_1, w_2 \in W$. Then there exist $v_1, v_2 \in V$ such that

$$T(v_1) = w_1 \quad \text{and} \quad T(v_2) = w_2.$$

Then $w_1 + w_2 = T(v_1) + T(v_2) = T(v_1 + v_2)$,

$$\text{and so} \quad w_1 + w_2 \in \text{im}(T).$$

Let $w \in W$ and let $c \in \mathbb{R}$. Then there exists $v \in V$ such that

$$T(v) = w.$$

Then $cw = cT(v) = T(cv)$

$$\text{and so} \quad cw \in \text{im}(T).$$

So $\text{im}(T)$ is a subspace of W.

Example V27&28. Let

$$S = \{|1,3,-1,1\rangle,\ \ |2,6,0,4\rangle,\ \ |3,9,-2,4\rangle\ \}.$$

Then

$$S = \left\{ \begin{pmatrix} 1 \\ 3 \\ -1 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 6 \\ 0 \\ 4 \end{pmatrix}, \begin{pmatrix} 3 \\ 9 \\ -2 \\ 4 \end{pmatrix} \right\}$$

and

$$\mathbb{R}\text{-span}(S) = \mathrm{im}(A),\ \ \text{where } A = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 6 & 0 \\ -1 & 0 & -2 \\ 1 & 4 & 4 \end{pmatrix}.$$

Lecture 29: Learning to do proofs – The minimax basis theorem

Definition (Spanning set, linearly independent set, basis)

Let V be an $\mathbb{F}$-vector space and let $B = \{v_1, \ldots, v_k\}$ be a subset of V. The subset B is a *spanning set of V* if B satisfies

$$\{c_1 v_1 + \cdots + c_k v_k \mid c_1, \ldots, c_k \in \mathbb{F}\} = V.$$

The subset B is a *linearly independent set in V* if B satisfies

$$\text{if } c_1, \ldots, c_k \in \mathbb{F} \text{ and } c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k = 0$$

$$\text{then } c_1 = 0, \ldots, c_k = 0.$$

The subset B is a *basis of V* if B satisfies:

$$B \text{ is a spanning set of } V \text{ and } B \text{ is a linearly independent set in } V.$$

Theorem (Basis Minimax Theorem)

Let V be an $\mathbb{F}$-vector space and let B be a subset of V. The following are equivalent:

(a) *B is a basis of V.*

(b) *B is a minimal spanning set of V.*

(c) *B is a maximal linearly independent set of V.*

Theorem (Exchange Theorem)

Let V be an $\mathbb{F}$-vector space. Let $B = \{v_1, \dots, v_k\}$ be a basis of V and let $D = \{d_1, \dots, d_\ell\}$ be another basis of V. Then there exists $d_{i_1} \in D$ such that

$$\{d_{i_1}, b_2, b_3, \dots, b_k\} \quad \text{is a basis of } V.$$

Theorem (Dimension Theorem)

Let V be an $\mathbb{F}$-vector space. Any two bases of V have the same number of elements.

The Dimension Theorem is the reason that

$$\dim(V) \text{ makes sense to consider.}$$

Definition (Dimension)

Let V be a vector space. The *dimension* of V is

$$\dim(V) = (\text{number of elements in a basis } B \text{ of } V).$$

The following provides an example of a spanning set that is not minimal, and another spanning set for the same subspace that is minimal.

Example V14. Let S be the subset of $\mathbb{R}^3$ given by

$$S = \{(1,1,1),(2,2,2),(3,3,3)\}. \qquad \text{Determine } \mathbb{R}\text{-span}(S).$$

In this case

$$\begin{aligned}
\mathbb{R}\text{-span}(S) &= \{c_1\,|1,1,1\rangle + c_2\,|2,2,2\rangle + c_3\,|3,3,3\rangle \mid c_1,c_2,c_3 \in \mathbb{R}\} \\
&= \{c_1\,|1,1,1\rangle + 2c_2\,|1,1,1\rangle + 3c_3\,|1,1,1\rangle \mid c_1,c_2,c_3 \in \mathbb{R}\} \\
&= \{(c_1 + 2c_2 + 3c_3)\,|1,1,1\rangle \mid c_1,c_2,c_3 \in \mathbb{R}\} \\
&= \{t\,|1,1,1\rangle \mid t \in \mathbb{R}\} = \mathbb{R}\text{-span}\{|1,1,1\rangle\} \\
&= \{|t,t,t\rangle \mid t \in \mathbb{R}\}.
\end{aligned}$$

Here $\{\ |1,1,1\rangle\ \}$ is a basis of $\mathbb{R}$-span(S) and

$$\dim(\mathbb{R}\text{-span}(S)) = 1 \quad \text{(even though } S \text{ has 3 elements).}$$

□

Proposition (Span is a subspace)

Let V be a vector space. Let $B = \{b_1, \dots, b_k\}$ be a subset of V. Then $\text{span}(B)$ is a subspace of V.

Proof.

To show: (1) $0 \in \text{span}(B)$.
(2) If $v_1, v_2 \in \text{span}(B)$ then $v_1 + v_2 \in \text{span}(B)$.
(3) If $v \in \text{span}(B)$ and $c \in \mathbb{R}$ then $cv \in \text{span}(B)$.

(1) Since $0 = 0b_1 + \cdots 0b_k$ then $0 \in \text{span}\{b_1, \dots, b_k\} = \text{span}(B)$.

(2) Assume $v_1, v_2 \in \text{span}(B)$. To show $v_1 + v_2 \in \text{span}(B)$.
Since $v_1, v_2 \in \text{span}(B)$
then there exist $a_1, \dots, a_k, c_1, \dots, c_k \in \mathbb{R}$ such that

$$v_1 = a_1 b_1 + \cdots + a_k b_k \quad \text{and} \quad v_2 = c_1 b_1 + \cdots + c_k b_k.$$

Then

$$
\begin{aligned}
v_1 + v_2 &= (a_1 b_1 + \cdots + a_k b_k) + (c_1 b_1 + \cdots + c_k b_k) \\
&= (a_1 + c_1) b_1 + \cdots + (a_k + c_k) b_k.
\end{aligned}
$$

So $v_1 + v_2 \in \operatorname{span}\{b_1, \ldots, b_k\} = \operatorname{span}(B)$.

(3) Assume $v \in \operatorname{span}(B)$ and $c \in \mathbb{R}$.
To show $cv \in \operatorname{span}(B)$.
Since $v \in \operatorname{span}(B)$ then there exist $a_1, \ldots, a_k \in \mathbb{R}$ such that

$$
v = a_1 b_1 + \cdots + a_k b_k.
$$

Then

$$
cv = c(a_1 b_1 + \cdots + a_k b_k) = (ca_1) b_1 + \cdots + (ca_k) b_k.
$$

So $cv \in \operatorname{span}\{b_1, \ldots, b_k\}$. So $cv \in \operatorname{span}(B)$. $\square$

Proof of the Dimension Theorem

Assume

$$B = \{b_1, \dots, b_k\} \quad \text{is a basis of } V \text{ and}$$
$$D = \{d_1, \dots, d_\ell\} \quad \text{is another basis of } V.$$

Using the Exchange theorem, there exists $d_{i_1} \in D$ such that $d_{i_1} \notin \operatorname{span}(B - b_1)$. Then

$$B_1 = \{d_{i_1}, b_2, b_3, b_4, \dots b_k\} \quad \text{is a basis of } V.$$

Using the Exchange theorem, there exists $d_{i_2} \in D$ such that $d_{i_2} \notin \operatorname{span}(B_1 - b_2)$. Then

$$B_2 = \{d_{i_1}, d_{i_2}, b_3, b_4, \dots b_k\} \quad \text{is a basis of } V.$$

Continue this replacement process to obtain

$$B' = \{d_{i_1}, \dots, d_{i_k}\} \subseteq D, \quad \text{such that } B' \text{ is a basis of } V.$$

By the Minimax Theorem D is a minimal spanning set.
So $B' = D$ and $k = \ell$. $\square$

Proof of the Exchange Theorem

Assume

$$\begin{aligned} B &= \{b_1, \ldots, b_k\} \quad \text{is a basis of } V \text{ and} \\ D &= \{d_1, \ldots, d_\ell\} \quad \text{is another basis of } V. \end{aligned}$$

If $d_1, \ldots, d_\ell \in \operatorname{span}(B - \{b_1\})$ then

$$V = \operatorname{span}(d_1, \ldots, d_\ell) \subseteq \operatorname{span}(B - \{b_1\}) \subseteq V$$

giving $V = \operatorname{span}(B - \{b_1\})$.

But since B is a minimal spanning set then $V \neq \operatorname{span}(B - \{b_1\})$ and so

$$\text{there exists } d_{i_1} \in D \text{ such that } d_{i_1} \notin \operatorname{span}(B - \{b_1\}).$$

$$d_{i_1} = c_1 b_1 + c_2 b_2 + \cdots + c_k b_k, \quad \text{with } c_1 \neq 0.$$

To show: $B_1 = \{d_{i_1}, b_2, \ldots, b_k\}$ is a basis of V.

To show: (1) $\text{span}\{d_{i_1}, b_2, \ldots, b_k\} = V$.
(2) $\{d_{i_1}, b_2, \ldots, b_k\}$ is linearly independent.

(1) Since

$$b_1 = c_1^{-1}(-d_{i_1} + c_2 b_2 + \cdots + c_k v_k)$$

then $b_1, b_2, \ldots, b_k \in \text{span}\{d_{i_1}, b_2, \ldots, b_k\}$. So
$V = \text{span}\{b_1, \ldots, b_k\} \subseteq \text{span}\{d_{i_1}, b_2, \ldots, b_k\} \subseteq V$. So

$$V = \text{span}(d_{i_1}, b_2, \ldots, b_k\}.$$

(2) If $a_1 d_{i_1} + a_2 b_2 + \cdots + a_k b_k = 0$ then

$$a_1(c_1 b_1 + c_2 b_2 + \cdots + c_k b_k) + a_2 b_2 + \cdots + a_k b_k = 0.$$

Since B is linearly independent then $a_1 c_1 = 0$.
Since $c_1 \neq 0$ then $a_1 = 0$ and $a_2 b_2 + \cdots + a_k b_k = 0$.
Since B is linearly independent then $a_2 = 0, \ldots, a_k = 0$.
So $\{d_{i_1}, b_2, \ldots, b_k\}$ is linearly independent.
So $\{d_{i_1}, b_2, \ldots, b_k\}$ is a basis of V. □

Proof of the Minimax Basis Theorem

(a) $\Rightarrow$ (b): Assume $B = \{b_1, \ldots, b_k\}$ is a basis of B.
To show: B is a minimal spanning set of V.
To show: (1) B Is a spanning set.
(2) If $i \in \{1, \ldots, k\}$ then $B - \{b_i\}$ is not a spanning set.

(1) Since B is a basis then B is a spanning set.

(2) To show: If $i \in \{1, \ldots, k\}$ and $B - \{b_i\}$ is a spanning set then B is not a basis.
Assume $i \in \{1, \ldots, k\}$ and $B - \{b_i\}$ is a spanning set.
Then there exist $c_1, \ldots, c_{i-1}, c_{i+1}, \ldots, c_k \in \mathbb{R}$ such that

$$b_i = c_1 b_1 + \cdots + c_{i-1} b_{i-1} + c_{i+1} b_{i+1} + \cdots + c_k b_k.$$

Then

$$0 = c_1 b_1 + \cdots + c_{i-1} b_{i-1} - b_i + c_{i+1} b_{i+1} + \cdots + c_k b_k.$$

So $\{b_1, \ldots, b_k\}$ is not linearly independent.
So B is not a basis.
So if $i \in \{1, \ldots, k\}$ then $B - \{b_i\}$ is not a spanning set.

(b) $\Rightarrow$ (c): Assume B is a minimal spanning set.
To show: B is a maximal linearly independent set in V.
To show: (1) B is a linearly independent set in V.
(2) If $v \in V$ then $B \cup \{v\}$ is not linearly independent.

(1) To show: If B is a spanning set and B is not linearly independent then B is not a minimal spanning set.
Assume B is a spanning set and B is not linearly independent.
Then there exist $c_1, \ldots, c_k \in \mathbb{R}$ and $i \in \{1, \ldots, k\}$ such that

$$c_1 b_1 + \cdots + c_k b_k = 0 \quad \text{and} \quad c_i \neq 0.$$

Then $b_i = -c_i^{-1}(c_1 b_1 + \cdots + c_{i-1} b_{i-1} + c_{i+1} b_{i+1} + \cdots + c_k b_k)$.
So $\text{span}(B - \{b_i\}) \supseteq \text{span}(b_1, \ldots, b_k) = V$.
So $\text{span}(B - \{b_i\}) = V$ and B is not a minimal spanning set of B.
So if B is a minimal spanning set then B is linearly independent.

(2) To show: If $v \in V$ then $B \cup \{v\}$ is not linearly independent.
Assume $v \in V$. To show: $B \cup \{v\}$ is not linearly independent.

Since $\text{span}(B) = V$ then there exist $c_1, \ldots, c_k \in \mathbb{R}$ such that

$$v = c_1 b_1 + \cdots c_k b_k.$$

So $0 = c_1 b_1 + \cdots + c_k b_k - v$.
So $B \cup \{v\} = \{b_1, \ldots, b_k, v\}$ is not linearly independent.

(c) $\Rightarrow$ (a): Assume B is a maximal linearly independent set.
To show: B is a basis.
To show: $\text{span}(B) = V$.
Assume $v \in V$. To show $v \in \text{span}(B)$.
Since B is a maximal linearly independent set then $B \cup \{v\}$ is not linearly independent.
So there exist $c_1, \ldots, c_k, c_{k+1} \in \mathbb{R}$ and $i \in \{1, \ldots, k+1\}$ such that

$$c_1 b_1 + \cdots + c_k b_k + c_{k+1} v = 0 \quad \text{and} \quad c_i \neq 0.$$

The case $c_{k+1} = 0$ cannot occur since B is linearly independent.
So $c_{k+1} \neq 0$ and $v = -c_{k+1}^{-1}(c_1 b_1 + \cdots c_k b_k)$.
So $v \in \text{span}\{b_1, \ldots, b_k\} = \text{span}(B)$.
So $V = \text{span}(B)$. So B is a basis of V. $\square$

Lecture 30: Learning to do proofs – Invertible matrices are square

Definition (Kernel and image of a matrix)

Let $A \in M_{t \times s}(\mathbb{Q})$. The *kernel of A* is

$$\ker(A) = \{x \in \mathbb{Q}^s \mid Ax = 0\}$$

and the *image of A* is

$$\mathrm{im}(A) = \{Ax \mid s \in \mathbb{Q}^s\}.$$

Proposition

Let $s, t \in \mathbb{Z}_{>0}$ *and let* $A \in M_{t\times s}(\mathbb{R})$.
If $\ker(A) = 0$ *then the columns of* A *are linearly independent.*

Proof. Let $a_1, \ldots, a_s$ be the columns of A.
Assume $\ker(A) = 0$.
To show: $a_1, \ldots, a_s$ are linearly independent.
To show: If $c_1, \ldots, c_s \in \mathbb{R}$ and $c_1 a_1 + \cdots + c_s a_s = 0$

$$\text{then } c_1 = 0, c_2 = 0, \ldots, c_s = 0.$$

Assume $c_1, \ldots, c_s \in \mathbb{R}$ and $c_1 a_1 + \cdots + c_s a_s = 0$.
Then

$$A\begin{pmatrix} c_1 \\ \vdots \\ c_s \end{pmatrix} = 0. \qquad \text{So } \begin{pmatrix} c_1 \\ \vdots \\ c_s \end{pmatrix} \in \ker(A). \qquad \text{So } \begin{pmatrix} c_1 \\ \vdots \\ c_s \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}.$$

So $c_1 = 0, c_2 = 0, \ldots, c_s = 0$.
So $\{a_1, \ldots, a_s\}$ is linearly independent. □

Proposition

Let $s, t \in \mathbb{Z}_{>0}$ and let $A \in M_{t\times s}(\mathbb{R})$. Let $a_1, \dots, a_s$ be the columns of A. Then

$$im(A) = span\{a_1, \dots, a_s\}.$$

Proof.

$$\text{im}(A) = \{Ax \mid x \in \mathbb{R}^s\} = \left\{ \begin{pmatrix} | & & | \\ a_1 & \cdots & a_s \\ | & & | \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_s \end{pmatrix} \;\middle|\; x_1, \dots, x_s \in \mathbb{R} \right\}$$

$$= \left\{ x_1 \begin{pmatrix} | \\ a_1 \\ | \end{pmatrix} + \cdots + x_s \begin{pmatrix} | \\ a_s \\ | \end{pmatrix} \;\middle|\; x_1, \dots, x_s \in \mathbb{R} \right\}$$

$$= \mathbb{R}\text{-span}\{\text{columns of } A\}.$$

So im(A) is the set of linear combinations of the columns of A. □

Theorem (Invertible matrices are square)

Let $s, t \in \mathbb{Z}_{>0}$ *and let* $A \in M_{t\times s}(\mathbb{R})$*. Suppose there exists*

$$P \in M_{s\times t}(\mathbb{R}) \quad \textit{be such that} \quad PA = 1.$$

Suppose there exists

$$Q \in M_{s\times t}(\mathbb{R}) \quad \textit{be such that} \quad AQ = 1.$$

Then

(a) $\ker(A) = 0$.

(b) $im(A) = \mathbb{R}^t$.

(c) *The set of columns of* A *is a basis of* $\mathbb{Q}^t$.

(d) $s = t$.

(e) $P = Q$.

Proof. (a) To show: $\ker(A) = \{0\}$.

To show: (1) $\{0\} \subseteq \ker(A)$.
(2) $\ker(A) \subseteq \{0\}$.

(1) Since $A \cdot 0 = 0$ then $0 \in \ker(A)$.
So $\{0\} \subseteq \ker(A)$.

(2) To show: If $x \in \ker(A)$ then $x \in \{0\}$.
Assume $x \in \ker(A)$. To show: $x = 0$.
Since $x \in \ker(A)$ then

$$Ax = 0. \quad \text{So } PAx = P0 = 0.$$

So $x = 1x = PAx = 0$. So $\ker(A) \subseteq \{0\}$.

So $\ker(A) = \{0\}$.

(b) To show: $\mathrm{im}(A) = \mathbb{R}^t$.

To show: (1) $\mathrm{im}(A) \subseteq \mathbb{R}^t$.
(2) $\mathbb{R}^t \subseteq \mathrm{im}(A)$.

(1) By definition of $(\mathrm{im}(A) = \{Ax \mid x \in \mathbb{R}^s\}$.
Since A is a $t \times s$ matrix then $\mathrm{im}(A) \subseteq \mathbb{R}^t$.

(2) To show: If $v \in \mathbb{R}^t$ then $v \in \mathrm{im}(A)$.
Assume $v \in \mathbb{R}^t$. To show: $v \in \mathrm{im}(A)$.

$$v = 1v = AQv \in \{Ax \mid x \in \mathbb{R}^t\} = \mathrm{im}(A).$$

So $v \in \mathrm{im}(A)$. So $\mathbb{R}^t \subseteq \mathrm{im}(A)$.

So $\mathbb{R}^t = \mathrm{im}(A)$.

(c) Since $\ker(A) = 0$ then

$$\text{the columns of } A \text{ are linearly independent.}$$

Since $\mathrm{im}(A) = \mathrm{span}\{\text{columns of } A\}$ and $\mathbb{R}^t = \mathrm{im}(A)$ then

$$\mathbb{R}^t = \mathrm{span}\{\text{columns of } A\}.$$

So $\{\text{columns of } A\}$ is a basis of $\mathbb{R}^t$.

(d) Let $a_1, \ldots, a_s$ be the columns of A. By part (c),

$$\{a_1, \ldots, a_s\} \quad \text{is a basis of } \mathbb{R}^t.$$

Let e_i be the $t \times 1$ matrix with 1 is in the ith entry and 0 elsewhere. Then

$$\{e_1, \ldots, e_t\} \quad \text{is a basis of } \mathbb{R}^t.$$

By the Dimension Theorem, any two bases of $\mathbb{R}^t$ have the same number of elements.
So $s = t$.

(e) To show: $P = Q$.

$$P = P \cdot 1 = P(AQ) = (PA)Q = 1 \cdot Q = Q. \qquad \square$$

Lecture 31: Application – Data Correlation

Correlation is a measure of how closely two variables are dependent.

Definition

The *mean* μ_X of a data set $X = \{x_1, \ldots, x_n\}$ is the average of the values in the data set.

$$\mu_X = \frac{1}{n}(x_1 + \cdots + x_n).$$

The *correlation of variables X and Y* is

$$\operatorname{corr}(X, Y) = \cos(\theta(X - \mu_X, Y - \mu_Y)), \quad \text{where}$$

$X - \mu_X = |x_1 - \mu_X, \ldots, x_n - \mu_X\rangle$ and $Y - \mu_Y = |y_1 - \mu_Y, \ldots, y_n - \mu_Y\rangle$.

Use $\quad \cos(\theta(\mathbf{u}, \mathbf{v})) = \dfrac{\langle \mathbf{u}, \mathbf{v}\rangle}{\|\mathbf{u}\| \cdot \|\mathbf{v}\|}$ to compute the correlation.

A value close to 1 indicates the values a highly correlated and a value close to -1 indicates the values are not at all correlated.

Example E3. Suppose the data set is assignment and exam marks for 7 students.

Student	Assignment Mark	Exam Mark
S1	99	100
S2	80	82.5
S3	79	79
S4	75.5	82.5
S5	87.5	91
S6	67	67.5
S7	76	68

The mean assignment mark is

$$\mu_A = \frac{1}{7}(99 + 80 + 79 + 75.5 + 87.5 + 67 + 76) = 80.5.$$

The mean exam mark is

$$\mu_E = \frac{1}{7}(100 + 82.5 + 79 + 82.5 + 91 + 67.5 + 68) = 81.5.$$

Then

$$A - \mu_A = |18.5, -0.5, -1.5, -5.5, 7, -13.5, -4.5\rangle,$$
$$E - \mu_E = |18.5, 1, -2.5, 1, 9.5, -14, -13.5\rangle$$

and the correlation between the assignment marks and the exam marks is

$$\begin{aligned} \operatorname{corr}(A, E) &= \cos(\theta(A - \mu_A, E - \mu_E)) \\ &= \frac{\langle A - \mu_A, E - \mu_E \rangle}{\|A - \mu_A\| \cdot \|E - \mu_E\|} = \frac{656.75}{(24.92)(28.62)} \approx 0.92. \end{aligned}$$

Lecture 32: Application – Data Line of best fit

Given a data set $D = \{(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)\}$ find

$$\text{the line of best fit} \quad y = a_{\text{best}}x + b_{\text{best}}.$$

Let

$$A = \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix}, \quad \mathbf{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}, \quad \mathbf{u} = \begin{pmatrix} a \\ b \end{pmatrix}$$

so that

$$\mathbf{y} - A\mathbf{u} = \begin{pmatrix} y_1 - (a + bx_1) \\ \vdots \\ y_n - (a + bx_n) \end{pmatrix} \quad \textit{PICTURE}$$

Goal: Minimize $\|\mathbf{y} - A\mathbf{u}\|$.

Let

$$W = \{A\mathbf{u} \mid \mathbf{u} = \begin{pmatrix} a \\ b \end{pmatrix} \in \mathbb{R}^2\} \text{ and}$$

PICTURE

$$\vec{s} \in \mathbb{R}^2 \quad \text{such that} \quad A\vec{s} = \text{proj}_W(\mathbf{y}).$$

Then $\|\mathbf{y} - A\vec{s}\|$ will be minimal if $\mathbf{y} - A\vec{s}$ is perpendicular to W.

So we want if $\mathbf{u} \in \mathbb{R}^2$ then $\langle \mathbf{y} - A\vec{s}, A\mathbf{u} \rangle = 0$.

So we want if $\mathbf{u} \in \mathbb{R}^2$ then $\mathbf{u}^t A^t (\mathbf{y} - A\vec{s}) = 0$.

So we want $A^t\mathbf{y} - A^t A\vec{s} = 0$.

So we want $A^t A\vec{s} = A^t\mathbf{y}$.

So we want

$$\vec{s} = (A^t A)^{-1} A^t \mathbf{y} = \begin{pmatrix} a_{\mathrm{best}} \\ b_{\mathrm{best}} \end{pmatrix}$$

and the line of best fit is $y = a_{\mathrm{best}} x + b_{\mathrm{best}}$.

Example IP14 Follow the above procedure. Given the data set $D = \{(-1,1),(1,1),(2,3)\}$ then

$$A = \begin{pmatrix} 1 & -1 \\ 1 & 1 \\ 1 & 2 \end{pmatrix} \quad \text{and} \quad \mathbf{y} = \begin{pmatrix} 1 \\ 1 \\ 3 \end{pmatrix}.$$

Then

$$A^t A = \begin{pmatrix} 1 & 1 & 1 \\ -1 & 1 & 2 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 1 & 1 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 3 & 2 \\ 2 & 6 \end{pmatrix} \quad \text{and}$$

$$A^t \mathbf{y} = \begin{pmatrix} 1 & 1 & 1 \\ -1 & 1 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 3 \end{pmatrix} = \begin{pmatrix} 5 \\ 6 \end{pmatrix} \quad \text{and} \quad (A^t A)^{-1} = \tfrac{1}{14} \begin{pmatrix} 6 & -2 \\ -2 & 3 \end{pmatrix}.$$

So

$$\vec{s} = \frac{1}{14} \begin{pmatrix} 6 & -2 \\ -2 & 3 \end{pmatrix} \begin{pmatrix} 5 \\ 6 \end{pmatrix} = \frac{1}{14} \begin{pmatrix} 18 \\ 8 \end{pmatrix} = \begin{pmatrix} \frac{9}{7} \\ \frac{4}{7} \end{pmatrix}.$$

So the line of best fit is

$$y = \tfrac{9}{7} + \tfrac{4}{7}x.$$

Example V6. Is $W = \{|x, y, z\rangle \in \mathbb{R}^3 \mid x + y + z = 0\}$ a $\mathbb{R}$-subspace of $\mathbb{R}^3$?

A $\mathbb{R}$-*subspace of* $\mathbb{R}^3$ is a subset $W \subseteq \mathbb{R}^3$ such that

(a) If $w_1, w_2 \in W$ then $w_1 + w_2 \in W$,

(b) $0 \in W$,

(c) If $w \in W$ then $-w \in W$,

(d) If $w \in W$ and $c \in \mathbb{R}$ then $cw \in W$.

Proof.

(a) Assume $w_1 = |a, b, c\rangle \in W$ and $w_2 = |x, y, z\rangle \in W$.
Then $a + b + c = 0$ and $x + y + z = 0$.
Then $w_1 + w_2 = |a + x, b + y, c + z\rangle$ and
$(a+x)+(b+y)+(c+z) = (a+b+c)+(x+y+z) = 0+0 = 0$.
So $w_1 + w_2 \in W$.

(b) $0 = |0, 0, 0\rangle$ satisfies $0 + 0 + 0 = 0$. So $0 \in W$.

(c) Assume $w = |x, y, z\rangle \in W$.

Then $x + y + z = 0$.

Then $-w = |-x, -y, -z\rangle$ and
$(-x) + (-y) + (-z) = -(x + y + z) = -0 = 0$.

So $-w \in W$.

(d) Assume $w = |x, y, z\rangle \in W$ and $c \in \mathbb{R}$.

Then $x + y + z = 0$.

Then $cw = |cx, cy, cz\rangle$ and
$cx + cy + cz = c(x + y + z) = c \cdot 0 = 0$.

So $cw \in W$.

So W is a subspace of $\mathbb{R}^3$. □

Example V7. Is the line $L = \{|x, y\rangle \in \mathbb{R}^2 \mid y = 2x + 1\}$ a subspace of $\mathbb{R}^2$?

A *subspace of* $\mathbb{R}^2$ is a subset $L \subseteq \mathbb{R}^2$ such that

(a) If $w_1, w_2 \in L$ then $w_1 + w_2 \in L$,

(b) $0 \in L$,

(c) If $w \in L$ then $-w \in L$,

(d) If $w \in L$ and $c \in \mathbb{R}$ then $cw \in L$.

Since $0 = |0, 0\rangle$ and $0 \neq 2 \cdot 0 + 1$ then $0 \notin L$.
So L is not a subspace of $\mathbb{R}^2$.

Example V8. Is $W = \{a_1x + a_2x^2 \mid a_1, a_2 \in \mathbb{R}\}$ a subspace of $\mathbb{R}[x]_{\leqslant 2}$?

A *subspace of* $\mathbb{R}[x]_{\leqslant 2}$ is a subset $W \subseteq \mathbb{R}[x]_{\leqslant 2}$ such that

(a) If $w_1, w_2 \in W$ then $w_1 + w_2 \in W$,

(b) $0 \in W$,

(c) If $w \in W$ then $-w \in W$,

(d) If $w \in W$ and $c \in \mathbb{R}$ then $cw \in W$.

Proof.

(a) Assume $w_1 = a_1x + a_2x^2 \in W$ and $w_2 = b_1x + b_2x^2 \in W$.

Then $a_1, a_2 \in \mathbb{R}$ and $b_1, b_2 \in \mathbb{R}$.

Then
$w_1 + w_2 = a_1x + a_2x^2 + b_1x + b_2x^2 = (a_1 + a_1)x + (b_1 + b_2)x^2$
and $a_1 + b_1 \in \mathbb{R}$ and $a_2 + b_2 \in \mathbb{R}$.

So $w_1 + w_2 \in W$.

(b) $0 = 0x + 0x^2$ satisfies $0 \in \mathbb{R}$ and $0 \in \mathbb{R}$. So $0 \in W$.

(c) Assume $w = a_1 x + a_2 x^2 \in W$.

Then $a_1, a_2 \in \mathbb{R}$.

Then $-w = -(a_1 x + a_2 x^2) = -a_1 x + (-a_2)x^2$ and $-a_1 \in \mathbb{R}$ and $-a_2 \in \mathbb{R}$.

So $-w \in W$.

(d) Assume $w = a_1 x + a_2 x^2 \in W$ and $c \in \mathbb{R}$.

Then $a_1, a_2 \in \mathbb{R}$.

Then $cw = c(a_1 x + a_2 x^2) = (ca_1)x + (ca_2)x^2$ and $ca_1 \in \mathbb{R}$ and $ca_2 \in \mathbb{R}$.

So $cw \in W$.

So W is a subspace of $\mathbb{R}[x]_{\leqslant 2}$. $\square$

Example V9. Is the set of real 2×2 matrices whose trace is equal to 0 a subspace of $M_{2\times 2}(\mathbb{R})$?

A *subspace of* $M_{2\times 2}(\mathbb{R})$ is a subset $W \subseteq M_{2\times 2}(\mathbb{R})$ such that

(a) If $w_1, w_2 \in W$ then $w_1 + w_2 \in W$,

(b) $0 \in W$,

(c) If $w \in W$ then $-w \in W$,

(d) If $w \in W$ and $c \in \mathbb{R}$ then $cw \in W$.

Proof. The set of real 2×2 matrices whose trace is equal to 0 is

$$W = \left\{ \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \;\middle|\; a_{11} + a_{22} = 0 \right\}.$$

(a) Assume $w_1 = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \in W$ and $w_2 = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \in W$.

Then $a_{11} + a_{22} = 0$ and $b_{11} + b_{22} = 0$.

Then $w_1 + w_2 = \begin{pmatrix} a_{11} + b_{11} & a_{12} + b_{12} \\ a_{21} + b_{21} & a_{22} + b_{22} \end{pmatrix}$ and

$$(a_{11} + b_{11}) + (a_{22} + b_{22}) = (a_{11} + a_{22}) + (b_{11} + b_{22}) = 0 + 0 = 0.$$

So $w_1 + w_2 \in W$.

(b) $0 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ $(0, 0, 0)$ satisfies $0 + 0 = 0$. So $0 \in W$.

(c) Assume $w = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \in W$.

Then $a_{11} + a_{22} = 0$.

Then $-w = -\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \begin{pmatrix} -a_{11} & -a_{12} \\ -a_{21} & -a_{22} \end{pmatrix}$ and

$(-a_{11}) + (-a_{22}) = -(a_{11} + a_{22}) = -0 = 0$.

So $-w \in W$.

(d) Assume $w = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \in W$ and $c \in \mathbb{R}$.

Then $a_{11} + a_{22} = 0$.

Then $cw = c\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \begin{pmatrix} ca_{11} & ca_{12} \\ ca_{21} & ca_{22} \end{pmatrix}$ and

$$ca_{11} + ca_{22} = c(a_{11} + a_{22}) = c \cdot 0 = 0.$$

So $cw \in W$.

So W is a subspace of $M_{2\times 2}(\mathbb{R})$. $\square$

Example V10. Is

$$S = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M_2(\mathbb{R}) \;\middle|\; ad - bc = 0 \right\} \text{ a subspace of } M_2(\mathbb{R})?.$$

A *subspace of* $M_{2\times 2}(\mathbb{R})$ is a subset $S \subseteq M_{2\times 2}(\mathbb{R})$ such that

(a) If $w_1, w_2 \in S$ then $w_1 + w_2 \in S$,

(b) $0 \in S$,

(c) If $w \in S$ then $-w \in S$,

(d) If $w \in S$ and $c \in \mathbb{R}$ then $cw \in S$.

Let $w_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$. Since $1 \cdot 0 - 0 \cdot 0 = 0 - 0 = 0$ then $w_1 \in S$.

Let $w_2 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$. Since $0 \cdot 1 - 0 \cdot 0 = 0 - 0 = 0$ then $w_2 \in S$.

Then

$$w_1 + w_2 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad \text{and} \quad 1 \cdot 1 - 0 \cdot 0 = 1.$$

So $w_1 + w_2 \notin S$.

So S is not a subspace of $M_{2\times 2}(\mathbb{R})$.

Lecture 34: Review – Linear transformation examples

Example LT3. Is the function $T\colon M_2(\mathbb{R}) \to \mathbb{R}$ given by

$$T\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \det\begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc \quad \text{a linear transformation?}$$

A *linear transformation from* $M_2(\mathbb{R})$ *to* $\mathbb{R}$ is a function $f\colon M_2(\mathbb{R}) \to \mathbb{R}$ such that

(a) If $v_1, v_2 \in M_2(\mathbb{R})$ then $f(v_1 + v_2) = f(v_1) + f(v_2)$,

(b) If $c \in \mathbb{R}$ and $v \in M_2(\mathbb{R})$ then $f(cv) = cf(v)$.

Since

$$1 = T\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = T\left(\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\right)$$

is not equal to

$$0 = 0 + 0 = T\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + T\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

then condition (a) does not hold and T is not a linear transformation.

Example LT4. Is the function $T\colon \mathbb{R}^3 \to \mathbb{R}^2$ given by

$$T(x_1, x_2, x_3) = (x_2 - 2x_3, 3x_1 + x_3) \qquad \text{a linear transformation?}$$

A *linear transformation from* $\mathbb{R}^3$ *to* $\mathbb{R}^2$ is a function $f\colon \mathbb{R}^3 \to \mathbb{R}^2$ such that

(a) If $u, v \in \mathbb{R}^3$ then $f(u+v) = f(u) + f(v)$,
(b) If $c \in \mathbb{R}$ and $v \in \mathbb{R}^3$ then $f(cv) = cf(v)$.

(a) Assume $u, v \in \mathbb{R}^3$ with $u = |u_1, u_2, u_3\rangle$ and $v = |v_1, v_2, v_3\rangle$. Then

$$\begin{aligned}
T(|u_1, u_2, u_3\rangle + |v_1, v_2, v_3\rangle &= T(|u_1 + v_1, u_2 + v_2, u_3 + v_3\rangle) \\
&= |(u_2 + v_2 - 2(u_3 + v_3), 3(u_1 + v_1) + (u_3 + v_3)\rangle \\
&= |u_2 - 2u_3 + v_2 - 2v_3, 3u_1 + u_3 + 3v_1 + v_3\rangle \\
&= |u_2 - 2u_3, 3u_1 + u_3\rangle + |v_2 - 2v_3, 3v_1 + v_3\rangle \\
&= T(|u_1, u_2, u_3\rangle) + T(|v_1, v_2, v_3\rangle)
\end{aligned}$$

(b) Assume $c \in \mathbb{R}$ and $u \in \mathbb{R}^3$ with $u = |u_1, u_2, u_3\rangle$. Then

$$\begin{aligned}
T(c \cdot |u_1, u_2, u_3\rangle) = T(|cu_1, cu_2, cu_3\rangle) &= |cu_2 - 2cu_3, 3cu_1 + cu_3\rangle \\
&= c|u_2 - 2u_3, 3u_1 + u_3\rangle = cT(|u_1, u_2, u_3\rangle).
\end{aligned}$$

So T is a linear transformation.

Example V12. In $\mathbb{R}^3$, is $|1,2,3\rangle \in \mathbb{R}\text{-span}\{|1,-1,2\rangle, |-1,1,2\rangle\}$?

By definition $\mathbb{R}\text{-span}\{|1,-1,2\rangle, |-1,1,2\rangle\}$
$$= \{c_1|1,-1,2\rangle + c_2|-1,1,2\rangle \mid c_1, c_2 \in \mathbb{R}\}.$$

So we need to show that there exist $c_1, c_2 \in \mathbb{R}$ such that

$$|1,2,3\rangle = c_1|1,-1,2\rangle + c_2|-1,1,2\rangle.$$

So we need to show that the system $\begin{matrix} c_1 - c_2 = 1, \\ -c_1 + c_2 = 2, \\ 2c_1 + 2c_2 = 3, \end{matrix}$ has a solution.

In matrix form the equations are $\begin{pmatrix} 2 & 2 \\ 1 & -1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \\ 2 \end{pmatrix}$.

Left multiply both sides by

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} \quad \text{to get} \quad \begin{pmatrix} 2 & 2 \\ -1 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} 3 \\ 2 \\ 3 \end{pmatrix}.$$

Already this gives an equation $0c_1 + 0c_2 = 3$, which has no solution.

So $|1, 2, 3\rangle \notin \mathbb{R}\text{-span}\{|1, -1, 2\rangle$ and $|-1, 1, 2\rangle\}$.

So $|1, 2, 3\rangle$ is not a linear combination of $|1, -1, 2\rangle$ and $|-1, 1, 2\rangle$.

So $|1, 2, 3\rangle \notin \mathbb{R}\text{-span}\{|1, -1, 2\rangle, |-1, 1, 2\rangle\}$. □

Example V13. In $\mathbb{R}[x]_{\leqslant 2}$, is $1-2x-x^2 \in \mathbb{R}\text{-span}\{1+x+x^2, 3+x^2\}$?

By definition $\mathbb{R}\text{-span}\{1+x+x^2, 3+x^2\}$

$$= \{c_1(1+x+x^2) + c_2(3+x^2) \mid c_1 c_2 \in \mathbb{R}\}.$$

So we need to show that there exist $c_1, c_2 \in \mathbb{R}$ such that

$$c_1(1+x+x^2) + c_2(3+x^2) = 1-2x-x^2.$$

So we need to show that the system $\begin{aligned} c_1 + 3c_2 &= 1, \\ c_1 + 0c_2 &= -2, \\ c_1 + c_2 &= -1, \end{aligned}$ has a solution.

In matrix form the equations are

$$\begin{pmatrix} 1 & 3 \\ 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} 1 \\ -2 \\ -1 \end{pmatrix}.$$

Left multiply both sides by

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & -1 \end{pmatrix} \quad \text{to get} \quad \begin{pmatrix} 1 & 3 \\ 1 & 1 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix}.$$

Left multiply both sides by

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{to get} \quad \begin{pmatrix} 1 & 1 \\ 0 & 2 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} -1 \\ 2 \\ -1 \end{pmatrix}.$$

Left multiply both sides by

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & -2 \end{pmatrix} \quad \text{to get} \quad \begin{pmatrix} 1 & 1 \\ 0 & -1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} -1 \\ -1 \\ 0 \end{pmatrix}.$$

Left multiply both sides by

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{to get} \quad \begin{pmatrix} 1 & 1 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}.$$

Left multiply both sides by

$$\begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{to get} \quad \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix}.$$

So $c_1 = -2$ and $c_2 = 1$ is a solution.

So $-2(1 + x + x^2) + (3 + x^2) = 1 - 2x - x^2$.

So $1 - 2x - x^2 \in \mathbb{R}\text{-span}\{1 + x + x^2, 3 + x^2\}$.

So $1 - 2x - x^2$ is a linear combination of $1 + x + x^2$ and $3 + x^2$. $\square$

Example V14. Let S be the subset of $\mathbb{R}^3$ given by

$$S = \{(1,1,1),(2,2,2),(3,3,3)\}. \qquad \text{Determine } \mathbb{R}\text{-span}(S).$$

In this case

$$\begin{aligned}
\mathbb{R}\text{-span}(S) &= \{c_1\,|1,1,1\rangle + c_2\,|2,2,2\rangle + c_3\,|3,3,3\rangle \mid c_1,c_2,c_3 \in \mathbb{R}\} \\
&= \{c_1\,|1,1,1\rangle + 2c_2\,|1,1,1\rangle + 3c_3\,|1,1,1\rangle \mid c_1,c_2,c_3 \in \mathbb{R}\} \\
&= \{(c_1 + 2c_2 + 3c_3)\,|1,1,1\rangle \mid c_1,c_2,c_3 \in \mathbb{R}\} \\
&= \{t\,|1,1,1\rangle \mid t \in \mathbb{R}\} \\
&= \{|t,t,t\rangle \mid t \in \mathbb{R}\}.
\end{aligned}$$

□

Example V15. Let S be the subset of $\mathbb{R}^2$ given by

$$S = \{|1,-1\rangle, |2,4\rangle\}. \qquad \text{Show that } \mathrm{span}(S) = \mathbb{R}^2.$$

Proof. By definition $\mathbb{R}\text{-span}(S) = \{c_1|1,-1\rangle + c_2|2,4\rangle \mid c_1, c_2 \in \mathbb{R}\}$.

To show: (a) $\mathbb{R}\text{-span}(S) \subseteq \mathbb{R}^2$
(b) $\mathbb{R}^2 \subseteq \mathbb{R}\text{-span}(S)$.

(a) Since $S \subseteq \mathbb{R}^2$ and $\mathbb{R}^2$ is closed under addition and scalar mutliplication then $\mathbb{R}\text{-span}(S) \subseteq \mathbb{R}^2$.

(b) To show: $\mathbb{R}^2 \subseteq \mathbb{R}\text{-span}(S)$.
To show: $\mathbb{R}\text{-span}\{|1,0\rangle, |0,1\rangle\} \subseteq \mathbb{R}\text{-span}(S)$.

Since $\mathbb{R}\text{-span}(S)$ is closed under addition and scalar multiplication, we can show $\{|1,0\rangle, |0,1\rangle\} \subseteq \mathbb{R}\text{-span}(S)$.

To show: There exist $c_1, c_2, d_1, d_2 \in \mathbb{R}$ such that

$$c_1|1,-1\rangle + c_2|2,4\rangle = |1,0\rangle \quad \text{and} \quad d_1|1,-1\rangle + d_2|2,4\rangle = |0,1\rangle.$$

To show: There exist $c_1, c_2, d_1, d_2 \in \mathbb{R}$ such that

$$\begin{pmatrix} 1 & 2 \\ -1 & 4 \end{pmatrix} \begin{pmatrix} c_1 & d_1 \\ c_2 & d_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Since

$$\begin{pmatrix} 1 & 2 \\ -1 & 4 \end{pmatrix} \begin{pmatrix} \frac{2}{3} & -\frac{1}{3} \\ \frac{1}{6} & \frac{1}{6} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

then

$$\tfrac{2}{3}|1, -1\rangle + \tfrac{1}{6}|2, 4\rangle = |1, 0\rangle, \text{ and}$$

$$-\tfrac{1}{3}|1, -1\rangle + \tfrac{1}{6}|2, 4\rangle = |0, 1\rangle.$$

So $|1, 0\rangle \in \mathbb{R}\text{-span}(S)$ and $|0, 1\rangle \in \mathbb{R}\text{-span}(S)$.
So $\mathbb{R}\text{-span}\{|1, 0\rangle, |0, 1\rangle\} \subseteq \mathbb{R}\text{-span}(S)$.
So $\mathbb{R}^2 \subseteq \mathbb{R}\text{-span}(S)$.
So $\mathbb{R}\text{-span}(S) = \mathbb{R}^2$. □

Example V16. Let S be the subset of $\mathbb{R}^3$ given by

$$S = \{|1,2,0\rangle, |1,5,3\rangle, |0,1,1\rangle\}. \qquad \text{Show that } \mathrm{span}(S) = \mathbb{R}^3.$$

Proof. By definition

$$\mathbb{R}\text{-span}(S) = \{c_1|1,2,0\rangle + c_2|1,5,3\rangle + c_3|0,1,1\rangle \mid c_1, c_2, c_3 \in \mathbb{R}\}.$$

To show: (a) $\mathbb{R}\text{-span}(S) \subseteq \mathbb{R}^3$
(b) $\mathbb{R}^3 \subseteq \mathbb{R}\text{-span}(S)$.

(a) Since $S \subseteq \mathbb{R}^3$ and $\mathbb{R}^3$ is closed under addition and scalar multiplication then $\mathbb{R}\text{-span}(S) \subseteq \mathbb{R}^3$.

(b) To show: $\mathbb{R}^3 \subseteq \mathrm{span}(S)$.

To show: $\mathbb{R}\text{-span}\{|1,0,0\rangle, |0,1,0\rangle, |0,0,1\rangle\} \subseteq \mathrm{span}(S)$.

Since $\mathbb{R}\text{-span}(S)$ is closed under addition and scalar multiplication,

To show: $\{|1,0,0\rangle, |0,1,0\rangle, |0,0,1\rangle\} \subseteq \mathbb{R}\text{-span}(S)$.

To show: There exist $c_1, c_2, c_3, d_1, d_2, d_3, r_1, r_2, r_3 \in \mathbb{R}$ such that

$$\begin{aligned} c_1|1,2,0\rangle + c_2|1,5,3\rangle + c_3|0,1,1\rangle &= |1,0,0\rangle, \\ d_1|1,2,0\rangle + d_2|1,5,3\rangle + d_3|0,1,1\rangle &= |0,1,0\rangle, \\ r_1|1,2,0\rangle + r_2|1,5,3\rangle + r_3|0,1,1\rangle &= |0,0,1\rangle, \end{aligned}$$

To show: There exist $c_1, c_2, c_3, d_1, d_2, d_3, r_1, r_2, r_3 \in \mathbb{R}$ such that

$$\begin{pmatrix} 1 & 1 & 0 \\ 2 & 5 & 1 \\ 0 & 3 & 1 \end{pmatrix} \begin{pmatrix} c_1 & d_1 & r_1 \\ c_2 & d_2 & r_2 \\ c_3 & d_3 & r_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Multiply both sides by

$$\begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \text{ to get } \begin{pmatrix} 1 & 1 & 0 \\ 0 & 3 & 1 \\ 0 & 3 & 1 \end{pmatrix} \begin{pmatrix} c_1 & d_1 & r_1 \\ c_2 & d_2 & r_2 \\ c_3 & d_3 & r_3 \end{pmatrix} = \begin{pmatrix} 1 & -2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Multiply both sides by

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix} \text{ to get } \begin{pmatrix} 1 & 1 & 0 \\ 0 & 3 & 1 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} c_1 & d_1 & r_1 \\ c_2 & d_2 & r_2 \\ c_3 & d_3 & r_3 \end{pmatrix} = \begin{pmatrix} 1 & -2 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix}.$$

Since the bottom row on the left hand side is all 0 and the bottom row on the right hand sides is not all 0 then there *do not exist* $c_1, c_2, c_3, d_1, d_2, d_3, r_1, r_2, r_3 \in \mathbb{R}$ such that

$$\begin{pmatrix} 1 & 1 & 0 \\ 2 & 5 & 1 \\ 0 & 3 & 1 \end{pmatrix} \begin{pmatrix} c_1 & d_1 & r_1 \\ c_2 & d_2 & r_2 \\ c_3 & d_3 & r_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

So $\{|1,0,0\rangle, |0,1,0\rangle, |0,0,1\rangle\} \not\subseteq \mathbb{R}\text{-span}(S)$.

So $\text{span}(S) \neq \mathbb{R}^2$. $\square$

Example V17. Let S be the subset of $\mathbb{R}[x]_{\leqslant 2}$ given by

$$S = \{1 + x + x^2, x^2\}. \qquad \text{Show that } \operatorname{span}(S) = \mathbb{R}[x]_{\leqslant 2}.$$

Proof. By definition

$$\mathbb{R}\text{-span}(S) = \{c_1(1 + x + x^2) + c_2 x^2 \mid c_1, c_2 \in \mathbb{R}\}.$$

To show: (a) $\operatorname{span}(S) \subseteq \mathbb{R}[x]_{\leqslant 2}$
(b) $\mathbb{R}[x]_{\leqslant 2} \subseteq \mathbb{R}\text{-span}(S)$.

(a) Since $S \subseteq \mathbb{R}[x]_{\leqslant 2}$ and $\mathbb{R}[x]_{\leqslant 2}$ is closed under addition and scalar multiplication then $\mathbb{R}\text{-span}(S) \subseteq \mathbb{R}[x]_{\leqslant 2}$.

(b) To show: $\mathbb{R}[x]_{\leqslant 2} \subseteq \mathbb{R}\text{-span}(S)$.

To show: $\mathbb{R}\text{-span}\{1, x, x^2\} \subseteq \mathbb{R}\text{-span}(S)$.

Since $\mathbb{R}\text{-span}(S)$ is closed under addition and scalar multiplication,

To show: $\{1, x, x^2\} \subseteq \mathbb{R}\text{-span}(S)$.

To show: There exist $c_1, c_2, d_1, d_2, r_1, r_2 \in \mathbb{R}$ such that

$$c_1(1+x+x^2) + c_2x^2 = 1, \quad d_1(1+x+x^2) + d_2x^2 = x,$$

and

$$r_1(1+x+x^2) + r_2x^2 = x^2.$$

To show: There exist $c_1, c_2, d_1, d_2, r_1, r_2 \in \mathbb{R}$ such that

$$\begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} c_1 & d_1 & r_1 \\ c_2 & d_2 & r_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Multiply both sides by

$$\begin{pmatrix} -1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \text{ to get } \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} c_1 & d_1 & r_1 \\ c_2 & d_2 & r_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Since the top row on the left hand side is all 0 and the top row on the right hand sides is not all 0 then there *do not exist* $c_1, c_2, d_1, d_2, r_1, r_2 \in \mathbb{R}$ such that

$$\begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} c_1 & d_1 & r_1 \\ c_2 & d_2 & r_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

So $\{1, x, x^2\} \not\subseteq \mathbb{R}\text{-span}(S)$.

So $\mathbb{R}\text{-span}\{1, x, x^2\} \not\subseteq \mathbb{R}\text{-span}(S)$.

So $\mathbb{R}[x]_{\leqslant 2} \not\subseteq \mathbb{R}\text{-span}(S)$.

So $\mathbb{R}\text{-span}(S) \neq \mathbb{R}[x]_{\leqslant 2}$. □

Example V18a Let S be the subset of $\mathbb{C}^3$ given by

$$S = \{|2i, -1, 1\rangle, |-6, -3i, 3i\rangle\}. \qquad \text{Is } S \ \mathbb{C}\text{-linearly independent?}$$

To show: If $c_1, c_2 \in \mathbb{C}$ and $c_1\,|2i, -1, 1\rangle + c_2|-6, -3i, 3i\rangle = |0, 0, 0\rangle$ then $c_1 = 0$, $c_2 = 0$.
Assume $c_1, c_2 \in \mathbb{C}$ and $c_1\,|2i, -1, 1\rangle + c_2|-6, -3i, 3i\rangle = |0, 0, 0\rangle$.
Then

$$\begin{aligned} 2ic_1 - 6c_2 &= 0, \\ -c_1 - 3ic_2 &= 0, \\ c_1 + 3ic_2 &= 0, \end{aligned} \qquad \text{or equivalently} \qquad \begin{pmatrix} 2i & -6 \\ -1 & -3i \\ 1 & 3i \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Skipping the row reduction steps (DON'T skip the row reduction steps on an exam or an assignment!), this system has solutions

$$\begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} + t \begin{pmatrix} -3i \\ 1 \end{pmatrix}, \qquad \text{with } t \in \mathbb{R}.$$

So $c_1 = 0$, $c_2 = 0$ is not the only solution.
So S is not linearly independent.

Example V18b. Let B be the subset of $\mathbb{R}^3$ given by

$$B = \{|2i, -1, 1\rangle, |4, 0, 2\rangle\}. \qquad \text{Is } B \text{ linearly independent?}$$

To show: If $c_1, c_2 \in \mathbb{C}$ and $c_1\,|2i, -1, 1\rangle + c_2|4, 0, 2\rangle = |0, 0, 0\rangle$ then $c_1 = 0$, $c_2 = 0$.
Assume $c_1, c_2 \in \mathbb{C}$ and $c_1\,|2i, -1, 1\rangle + c_2|4, 0, 2\rangle = |0, 0, 0\rangle$
Then

$$\begin{aligned} 2ic_1 + 4c_2 &= 0, \\ -c_1 + 0c_2 &= 0, \\ c_1 + 2c_2 &= 0, \end{aligned} \qquad \text{or equivalently} \qquad \begin{pmatrix} 2i & 4 \\ -1 & 0 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Skipping the row reduction steps (DON'T skip the row reduction steps on an exam or an assignment!), this system has only one solution $c_1 = 0$, $c_2 = 0$.
So S is linearly independent.

Example V19. Let S be the subset of $\mathbb{R}^3$ given by

$$S = \{(2,0,0),(6,1,7),(2,-1,2)\}. \qquad \text{Is } S \text{ linearly independent?}$$

To show:
If $c_1, c_2, c_3 \in \mathbb{R}$ and $c_1\,|2,0,0\rangle + c_2|6,1,7\rangle + c_3|2,-1,2\rangle = |0,0,0\rangle$
then $c_1 = 0$, $c_2 = 0$, $c_3 = 0$.
Assume $c_1, c_2, c_3 \in \mathbb{R}$ and
$c_1\,|2,0,0\rangle + c_2|6,1,7\rangle + c_3|2,-1,2\rangle = |0,0,0\rangle$.
Then

$$\begin{aligned} 2c_1 + 6c_2 + 2c_3 &= 0, \\ c_2 - c_3 &= 0, \\ 7c_2 + 2c_3 &= 0, \end{aligned} \quad \text{or equivalently} \quad \begin{pmatrix} 2 & 6 & 2 \\ 0 & 1 & -1 \\ 0 & 7 & 2 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Skipping the row reduction steps (DON'T skip the row reduction steps on an exam or an assignment!), this system has only one solution:
$c_1 = 0$, $c_2 = 0$, $c_3 = 0$.
So S is linearly independent.

Example V20&26. Let S be the subset of $\mathbb{R}[x]_{\leqslant 2}$ given by

$S = \{1 + 2x + 5x^2, 1 + x + x^2, 1 + 2x + 3x^2\}$. Is S a basis of $\mathbb{R}[x]_{\leqslant 2}$?

To show: If $c_1, c_2, c_3 \in \mathbb{R}$ and
$c_1(1 + 2x + 5x^2) + c_2(1 + x + x^2) + c_3(1 + 2x + 3x^2) = 0$
then $c_1 = 0$, $c_2 = 0$, $c_3 = 0$.

Assume $c_1, c_2, c_3 \in \mathbb{R}$ and
$c_1(1 + 2x + 5x^2) + c_2(1 + x + x^2) + c_3(1 + 2x + 3x^2) = 0$.
Then

$$\begin{aligned} c_1 + c_2 + c_3 &= 0, \\ 2c_1 + c_2 + 2c_3 &= 0, \\ 5c_1 + c_2 + 3c_2 &= 0, \end{aligned} \quad \text{or, equivalently,} \quad \begin{pmatrix} 1 & 1 & 1 \\ 2 & 1 & 2 \\ 5 & 1 & 3 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Skipping the row reduction steps (DON'T skip the row reduction steps on an exam or an assignment!), this system has only one solution: $c_1 = 0$, $c_2 = 0$, $c_3 = 0$.
So S is linearly independent.

Since $\dim(\mathbb{R}[x]_{\leqslant 2}) = 3$ and S contains 3 linearly independent elements then B is a basis for $\mathbb{R}[x]_{\leqslant 2}$.

Example V21. Let S be the subset of $M_2(\mathbb{R})$ given by

$$S = \left\{ \begin{pmatrix} 1 & 3 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} -2 & 1 \\ 1 & -1 \end{pmatrix}, \begin{pmatrix} 1 & 10 \\ 4 & 2 \end{pmatrix} \right\}.$$

Is S linearly independent?

To show: If $c_1, c_2, c_3 \in \mathbb{R}$ and

$$c_1 \begin{pmatrix} 1 & 3 \\ 1 & 1 \end{pmatrix} + c_2 \begin{pmatrix} -2 & 1 \\ 1 & -1 \end{pmatrix} + c_3 \begin{pmatrix} 1 & 10 \\ 4 & 2 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

then $c_1 = 0$, $c_2 = 0$, $c_3 = 0$.

Assume $c_1, c_2, c_3 \in \mathbb{R}$ and

$$c_1 \begin{pmatrix} 1 & 3 \\ 1 & 1 \end{pmatrix} + c_2 \begin{pmatrix} -2 & 1 \\ 1 & -1 \end{pmatrix} + c_3 \begin{pmatrix} 1 & 10 \\ 4 & 2 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Then

$$\begin{aligned} c_1 - 2c_2 + c_3 &= 0, \\ 3c_1 + c_2 + 10c_3 &= 0, \\ c_1 + c_2 + 4c_3 &= 0, \\ c_1 - c_2 + 2c_3 &= 0, \end{aligned} \quad \text{or, equivalently,} \quad \begin{pmatrix} 1 & -2 & 1 \\ 3 & 1 & 10 \\ 1 & 1 & 4 \\ 1 & -1 & 2 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Skipping the row reduction steps (DON'T skip the row reduction steps on an exam or an assignment!), this system has solutions

$$\begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} + t \begin{pmatrix} -3 \\ 1 \\ -1 \end{pmatrix}, \qquad \text{with } t \in \mathbb{R}.$$

So $c_1 = 0$, $c_2 = 0$, $c_3 = 0$ is not the only solution.
So S is not linearly independent.

Here is a check that $c_1 = -3$, $c_2 = 1$, $c_3 = -1$ is a solution:

$$\begin{aligned} -3\begin{pmatrix} 1 & 3 \\ 1 & 1 \end{pmatrix} + \begin{pmatrix} -2 & 1 \\ 1 & -1 \end{pmatrix} - \begin{pmatrix} 1 & 10 \\ 4 & 2 \end{pmatrix} &= -3\begin{pmatrix} 1 & 3 \\ 1 & 1 \end{pmatrix} + \begin{pmatrix} -3 & -9 \\ -3 & -3 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}. \end{aligned}$$

Example V23. Is $S = \{(1,-1),(2,4)\}$ a basis of $\mathbb{R}^2$?
Let

$$A = \begin{pmatrix} 1 & 2 \\ -1 & 4 \end{pmatrix}. \qquad \text{Then} \qquad A^{-1} = \frac{1}{6}\begin{pmatrix} 4 & -2 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} \frac{2}{3} & -\frac{1}{3} \\ \frac{1}{6} & \frac{1}{6} \end{pmatrix}.$$

So

$$\begin{pmatrix} 1 & 2 \\ -1 & 4 \end{pmatrix}\begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \qquad \text{gives} \qquad \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

So S is linearly independent.

If $|a,b\rangle \in \mathbb{R}^2$ then $|a,b\rangle = c_1|1,-1\rangle + c_2|2,4\rangle$, where

$$\begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} \frac{2}{3} & -\frac{1}{3} \\ \frac{1}{6} & \frac{1}{6} \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \frac{2}{3}a - \frac{1}{3}b \\ \frac{1}{6}a + \frac{1}{6}b \end{pmatrix}.$$

So $\mathbb{R}^2 \subseteq \mathbb{R}\text{-span}(S)$. Since $S \subseteq \mathbb{R}^2$ and $\mathbb{R}^2$ is closed under addition and scalar multiplication then $\mathbb{R}\text{-span}(S) \subseteq \mathbb{R}^2$. So $\mathbb{R}\text{-span}(S) = \mathbb{R}^2$.

So S is a basis of $\mathbb{R}^2$.

Example V24. Is $S = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \right\}$ a basis of

$$\{A \in M_2(\mathbb{R}) \mid \operatorname{Tr}(A) = 0\}?$$

If $c_1, c_2, c_3 \in \mathbb{R}$ and

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

then

$$\begin{pmatrix} c_1 & c_2 \\ c_3 & -c_1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \qquad \text{and} \qquad \begin{aligned} c_1 &= 0, \\ c_2 &= 0, \\ c_3 &= 0. \end{aligned}$$

So S is linearly independent.

Then

$$
\begin{aligned}
&\{A \in M_2(\mathbb{R}) \mid \operatorname{Tr}(A) = 0\} \\
&= \left\{ \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \mid a_{11}, a_{12}, a_{21}, a_{22} \in \mathbb{R}, a_{11} + a_{22} = 0 \right\} \\
&= \left\{ \begin{pmatrix} c_1 & c_2 \\ c_3 & -c_1 \end{pmatrix} \mid c_1, c_2, c_3 \in \mathbb{R} \right\} \\
&= \left\{ c_1 \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} + c_2 \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + c_3 \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \;\middle|\; c_1, c_2, c_3 \in \mathbb{R} \right\} \\
&= \operatorname{span}(S).
\end{aligned}
$$

So S is a basis of $\{A \in M_2(\mathbb{R}) \mid \operatorname{Tr}(A) = 0\}$.

Example V25. Is

$$
S = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \right\} \qquad \text{a basis of } M_2(\mathbb{R})?
$$

Since $E_{11}, E_{12}, E_{21}, E_{22}$ is a basis of $M_2(\mathbb{R})$ then $\dim(M_2(\mathbb{R})) = 4$.
Since S contains only 3 elements then S is not a basis of $M_2(\mathbb{R})$.