W7

Dimensions

Def. No. of vectors in a basis for V

Denoted by dim(V)

dim({0}) = 0

cardinality of {0} = 1

dim =/= cardinality

Example

Find a basis for and determine the dimension of the solution space of the homogeneous system

$$\begin{cases} 2v + 2w - x + z = 0 \\ -v - w + 2x - 3y + z = 0 \\ x + y + z = 0 \\ v + w - 2x - z = 0 \end{cases} \xrightarrow{\text{general solution}} \begin{pmatrix} v \\ w \\ x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -s-t \\ s \\ -t \\ 0 \\ t \end{pmatrix} = s\underbrace{\begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}}_{\boldsymbol{u}_1} + t\underbrace{\begin{pmatrix} -1 \\ 0 \\ -1 \\ 0 \\ 1 \end{pmatrix}}_{\boldsymbol{u}_2}$$

linearly indep.

solution space = span$\{\boldsymbol{u}_1, \boldsymbol{u}_2\}$

basis for the solution space = $\{\boldsymbol{u}_1, \boldsymbol{u}_2\}$

dim(solution space) = 2 ← no. of parameters in the general solution

Geometrical Meaning of Dimension

1. $\dim(\mathbb{R}^n) = n$.
2. Except $\{\mathbf{0}\}$ and $\mathbb{R}^2$, all subspaces of $\mathbb{R}^2$ are lines through the origin and they are of dimension 1.
3. Except $\{\mathbf{0}\}$ and $\mathbb{R}^3$, all subspaces of $\mathbb{R}^3$ are either lines through the origin, which are of dimension 1, or planes containing the origin, which are of dimension 2.

Lines and planes that do not pass through the origin are not subspaces, and hence we cannot talk about their dimension algebraically.

Exercise 3 Q43

Let V, W be subspaces of Rn.

dim(V + W) = dim(V) + dim(W) − dim(V ∩ W)

but may not be basis

§ if V = span(S), w = span(T), then V+W = span{S ∪ T}.

§ V ∪ W is span(S) ∪ span(T) which is not a span

If V and W are subspaces of a vector space, then there exists a basis S_1 for V and a basis S_2 for W such that $S_1 \cap S_2$ is a basis for $V \cap W$. T

As long as there exists one

If V and W are subspaces of a vector space, then there exists a basis S_1 for V and a basis S_2 for W such that $S_1 \cup S_2$ is a basis for $V + W$. T

If S_1 and S_2 are basis for V and W respectively, where V and W are subspaces of a vector space, then $S_1 \cap S_2$ is a basis for $V \cap W$. (See Question 3.24.) F

If S_1 and S_2 are basis for V and W respectively, where V and W are subspaces of a vector space, then $S_1 \cup S_2$ is a basis for $V + W$. (See Question 3.20.) F

Checking for Dimensions

homogeneous system ⟶ row echelon form $\boldsymbol{R}$

number of non-pivot columns in $\boldsymbol{R}$

||

number of parameters in general solution

||

number of vectors in basis for solution space

||

the dimension of the solution space

Theorem 3.6.7

sub given basis vectors into V coord.

Let V be a vector space of dimension k and S a subset of V.

The following are equivalent:

1. S is a basis for V
2. S is linearly independent and |S| = k = dim(V)
3. S spans V and |S| = k = dim(V)

To show S is a basis for V :

S lin. indep S spans V	or	S lin. indep $\lvert S \rvert$ = dim V	or	S spans V $\lvert S \rvert$ = dim V

→ rep $\tilde{w}$ $V \subseteq$ span(S)

Let $\boldsymbol{u}_1 = (2, 0, -1)$, $\boldsymbol{u}_2 = (4, 0, 7)$ and $\boldsymbol{u}_3 = (-1, 1, 4)$.
Show that $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3\}$ is a basis for $\mathbb{R}^3$.

Solution:

$$c_1\boldsymbol{u}_1 + c_2\boldsymbol{u}_2 + c_3\boldsymbol{u}_3 = \mathbf{0}$$
$$\Rightarrow \quad c_1(2, 0, -1) + c_2(4, 0, 7) + c_3(-1, 1, 4) = (0, 0, 0)$$
$$\Rightarrow \quad c_1 = 0, \; c_2 = 0, \; c_3 = 0.$$

So $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3\}$ is linearly independent.

Since $\dim(\mathbb{R}^3) = 3$, (by Theorem 3.6.7) $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3\}$ is a basis for $\mathbb{R}^3$.

Theorem 3.6.9

Let U and V be subspaces of Rn

(i) If U is a subspace of V (i.e. U ⊆ V), then dim(U) ≤ dim(V)

(ii) If U is a subspace of V (i.e. U ⊆ V) and U ≠ V, then dim(U) < dim(V)

Note:

Let U and V be subspaces of Rn

✅ If dim(V) = n, then V = Rn

✅ If U ⊆ V and dim(U) = dim (V), then U = V

🚫 If dim(U) = dim(V), then U = V.

🚫 If dim(U) ≤ dim(V), then U ⊆ V.

There are infinitely many possible vector sets as a basis, not all are equal

Coordinate Vector

$S = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k\}$: a basis for a vector space V

$\boldsymbol{v}$: a vector in V

Write $\boldsymbol{v} = c_1\boldsymbol{u}_1 + c_2\boldsymbol{u}_2 + \cdots + c_k\boldsymbol{u}_k$

Then $(\boldsymbol{v})_S = (c_1, c_2, \ldots, c_k)$ row form of coordinate vector

$$[\boldsymbol{v}]_S = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_k \end{pmatrix}$$ column form of coordinate vector

Finding Coordinate Vectors

$S = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ basis

$\boldsymbol{v} = c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + \ldots + c_n\boldsymbol{v}_n$ $\qquad (\boldsymbol{v})_S = (c_1\ c_2 \ldots c_n)$

- Use Gaussian elimination
  - Convert vector equation into linear system
- Use orthogonal basis
  - If S is orthogonal, c_i can be found using dot product
- Use transition matrix
  - Transform from one coordinate to another

- For any $\boldsymbol{v} \in \mathbb{R}^n$, $(\boldsymbol{v})_E = \boldsymbol{v}$.
- Let $\boldsymbol{A}$ be an $m \times n$ matrix. Then
  $\boldsymbol{A}\boldsymbol{e}_i^{\mathrm{T}} =$ the i^{th} column of $\boldsymbol{A}$.

Let $S = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k\}$ be an orthogonal basis for a vector space V. Then for any $\boldsymbol{w} \in V$,

$$\boldsymbol{w} = \frac{\boldsymbol{w}\cdot\boldsymbol{u}_1}{\boldsymbol{u}_1\cdot\boldsymbol{u}_1}\boldsymbol{u}_1 + \frac{\boldsymbol{w}\cdot\boldsymbol{u}_2}{\boldsymbol{u}_2\cdot\boldsymbol{u}_2}\boldsymbol{u}_2 + \cdots + \frac{\boldsymbol{w}\cdot\boldsymbol{u}_k}{\boldsymbol{u}_k\cdot\boldsymbol{u}_k}\boldsymbol{u}_k,$$

i.e. $(\boldsymbol{w})_S = \left(\frac{\boldsymbol{w}\cdot\boldsymbol{u}_1}{\boldsymbol{u}_1\cdot\boldsymbol{u}_1}, \frac{\boldsymbol{w}\cdot\boldsymbol{u}_2}{\boldsymbol{u}_2\cdot\boldsymbol{u}_2}, \ldots, \frac{\boldsymbol{w}\cdot\boldsymbol{u}_k}{\boldsymbol{u}_k\cdot\boldsymbol{u}_k}\right)$.

Let $T = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k\}$ be an orthonormal basis for a vector space V. Then for any $\boldsymbol{w} \in V$,

$$\boldsymbol{w} = (\boldsymbol{w}\cdot\boldsymbol{v}_1)\boldsymbol{v}_1 + (\boldsymbol{w}\cdot\boldsymbol{v}_2)\boldsymbol{v}_2 + \cdots + (\boldsymbol{w}\cdot\boldsymbol{v}_k)\boldsymbol{v}_k$$

i.e. $(\boldsymbol{w})_S = (\boldsymbol{w}\cdot\boldsymbol{v}_1, \boldsymbol{w}\cdot\boldsymbol{v}_2, \ldots, \boldsymbol{w}\cdot\boldsymbol{v}_k)$.

Transition Matrices (sq matrices)

$S = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k\}$ and $T = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k\}$

two bases for a vector space V.

1. Express each $\boldsymbol{u}_i$ as linear combination of $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k\}$
2. Form the (column) coordinate vectors w.r.t. T

$$[\boldsymbol{u}_1]_T = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{k1} \end{pmatrix} \quad [\boldsymbol{u}_2]_T = \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{k2} \end{pmatrix} \quad \cdots \quad [\boldsymbol{u}_k]_T = \begin{pmatrix} a_{1k} \\ a_{2k} \\ \vdots \\ a_{kk} \end{pmatrix}$$

3. Form the matrix $\mathbf{P} = (\ [\boldsymbol{u}_1]_T\ [\boldsymbol{u}_2]_T \ldots [\boldsymbol{u}_k]_T\ )$

$$P = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kk} \end{pmatrix}$$

P is a square matrix

transition matrix from S to T

4. $\mathbf{P}\,[\boldsymbol{w}]_S = [\boldsymbol{w}]_T$ for any vector $\boldsymbol{w}$ in V.

Theorem 3.7.5 (inverse of transition matrices)

Let S and T be two bases for a vector space and let $\boldsymbol{P}$ be the transition matrix from S to T.

1. $\boldsymbol{P}$ is invertible.
2. $\boldsymbol{P}^{-1}$ is the transition matrix from T to S.

$S \xrightarrow{P_{k\times k}} T$: $[W]_T = P[W]_S$ $\qquad (T \mid S) \rightarrow \left(\begin{array}{c|c} I_k & P \\ 0 \cdots 0 & 0 \cdots 0 \end{array}\right)$

$T \longrightarrow S$: $[W]_S = P^{-1}[W]_T$ $\qquad P = ([u_1]_T \ldots [u_k]_T)$

If S and T are Orthonormal bases, then P is an orthogonal matrix. $P^{-1} = P^T$

Example

Suppose we know $\begin{pmatrix} 1 & 2 & 1 \\ 3 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix}$ is invertible.

Then we know that the linear system

$\begin{pmatrix} 1 & 2 & 1 \\ 3 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$ has only trivial solution

Write the linear system in vector equation form:

$x\begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix} + y\begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} + z\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$ has only zero coefficients $x = y = z = 0$

We conclude that $\left\{\begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}\right\}$ is linearly independent hence form a basis for $\mathbf{R}^3$

Suppose we know $\begin{pmatrix} 1 & 2 & 1 \\ 3 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix}$ is invertible.

Then we know that the determinant $\begin{vmatrix} 1 & 2 & 1 \\ 3 & 1 & 0 \\ 2 & 0 & 1 \end{vmatrix} \neq 0$

Then the transpose determinant $\begin{vmatrix} 1 & 3 & 2 \\ 2 & 1 & 0 \\ 1 & 0 & 1 \end{vmatrix} \neq 0$

So $\begin{pmatrix} 1 & 3 & 2 \\ 2 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}$ is invertible.

So the columns $\left\{\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 3 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}\right\}$ form a basis for $\mathbf{R}^3$

So the rows $\{(1\ 2\ 1), (3\ 1\ 0), (2\ 0\ 1)\}$ form a basis for $\mathbf{R}^3$

Standard basis

Example

Consider the following two bases S and T for $\mathbb{R}^3$:

$S = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3\}$, where $\boldsymbol{u}_1 = (1, 0, -1)$, $\boldsymbol{u}_2 = (0, -1, 0)$ and $\boldsymbol{u}_3 = (1, 0, 2)$, and

$T = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\}$, where $\boldsymbol{v}_1 = (1, 1, 1)$, $\boldsymbol{v}_2 = (1, 1, 0)$ and $\boldsymbol{v}_3 = (-1, 0, 0)$.

(a) Find the transition matrix from S to T.

(b) Let $\boldsymbol{w}$ be a vector in $\mathbb{R}^3$ such that $(\boldsymbol{w})_S = (2, -1, 2)$. Find $(\boldsymbol{w})_T$.

Solution

(a) First, we need to find $a_{11}, a_{21}, \ldots, a_{33}$ such that

$$\begin{aligned} \boldsymbol{u}_1 &= a_{11}\boldsymbol{v}_1 + a_{21}\boldsymbol{v}_2 + a_{31}\boldsymbol{v}_3, \\ \boldsymbol{u}_2 &= a_{12}\boldsymbol{v}_1 + a_{22}\boldsymbol{v}_2 + a_{32}\boldsymbol{v}_3, \\ \boldsymbol{u}_3 &= a_{13}\boldsymbol{v}_1 + a_{23}\boldsymbol{v}_2 + a_{33}\boldsymbol{v}_3. \end{aligned}$$

$S \rightarrow T$

By

$$\begin{array}{cccccc} v_1 & v_2 & v_3 & u_1 & u_2 & u_3 \end{array}$$

$$\underbrace{\left(\begin{array}{ccc|c|c|c} 1 & 1 & -1 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 & -1 & 0 \\ 1 & 0 & 0 & -1 & 0 & 2 \end{array}\right)}_{T \qquad S} \xrightarrow[\text{Elimination}]{\text{Gauss-Jordan}} \left(\begin{array}{ccc|c|c|c} 1 & 0 & 0 & -1 & 0 & 2 \\ 0 & 1 & 0 & 1 & -1 & -2 \\ 0 & 0 & 1 & -1 & -1 & -1 \end{array}\right),$$

$$[u_1]_T \quad [u_2]_T \quad [u_3]_T$$

we have

$$\begin{aligned} \boldsymbol{u}_1 &= -\boldsymbol{v}_1 + \boldsymbol{v}_2 - \boldsymbol{v}_3, \\ \boldsymbol{u}_2 &= \quad\ - \boldsymbol{v}_2 - \boldsymbol{v}_3, \\ \boldsymbol{u}_3 &= 2\boldsymbol{v}_1 - 2\boldsymbol{v}_2 - \boldsymbol{v}_3. \end{aligned}$$

(See Example 3.2.11 for the method of finding the linear combinations.)

So $\boldsymbol{P} = \begin{pmatrix} -1 & 0 & 2 \\ 1 & -1 & -2 \\ -1 & -1 & -1 \end{pmatrix}$ is the transition matrix from S to T.

Express S in terms of T

(b) Since $[\boldsymbol{w}]_T = \boldsymbol{P}[\boldsymbol{w}]_S = \begin{pmatrix} -1 & 0 & 2 \\ 1 & -1 & -2 \\ -1 & -1 & -1 \end{pmatrix}\begin{pmatrix} 2 \\ -1 \\ 2 \end{pmatrix} = \begin{pmatrix} 2 \\ -1 \\ -3 \end{pmatrix}$, $(\boldsymbol{w})_T = (2, -1, -3)$.

So the transition matrix from T to S is

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

Note that

$$\boldsymbol{P}^{-1}[\boldsymbol{w}]_T = \begin{pmatrix} \frac{1}{3} & \frac{2}{3} & -\frac{2}{3} \\ -1 & -1 & 0 \\ \frac{2}{3} & \frac{1}{3} & -\frac{1}{3} \end{pmatrix}\begin{pmatrix} 2 \\ -1 \\ -3 \end{pmatrix} = \begin{pmatrix} 2 \\ -1 \\ 2 \end{pmatrix} = [\boldsymbol{w}]_S.$$

W8

Row Space & Column Space

$$A = \begin{array}{c} \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{array}{c} \boldsymbol{r_1} \\ \boldsymbol{r_2} \\ \\ \boldsymbol{r_m} \end{array} \\ \begin{array}{cccc} \boldsymbol{c_1} & \boldsymbol{c_2} & & \boldsymbol{c_n} \end{array} \end{array}$$

an $m \times n$ matrix

Row(A)

The row space of A $= \text{span}\{\boldsymbol{r_1}, \boldsymbol{r_2}, \ldots, \boldsymbol{r_m}\}$ a subspace of $\mathbf{R}^n$

$\neq$ $\{\boldsymbol{r_1}, \boldsymbol{r_2}, \ldots, \boldsymbol{r_m}\}$

Col(A)

The column space of A $= \text{span}\{\boldsymbol{c_1}, \boldsymbol{c_2}, \ldots, \boldsymbol{c_n}\}$ a subspace of $\mathbf{R}^m$

$\neq$ $\{\boldsymbol{c_1}, \boldsymbol{c_2}, \ldots, \boldsymbol{c_n}\}$

Row (column) space of zero matrix $\mathbf{0}$ = zero space

$$\mathbf{0} = \begin{pmatrix} 0 & 0 & \ldots & 0 \\ 0 & 0 & \ldots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & 0 \end{pmatrix}$$

Row (column) space of $n \times n$ identity matrix $\boldsymbol{I}_n$ = $\mathbf{R}^n$

$$\boldsymbol{I}_n = \begin{pmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & 1 \end{pmatrix}$$

§ row space of A = column space of A^T
§ column space of A = row space of A^T

Theorem 4.1.7

Let A and B be row equivalent matrices (same RREF). Then,
row space of A = row space of B,
i.e. e.r.o. preserve the row space of a matrix (but e.r.o. may not preserve the column space of a matrix).

$$A = \begin{array}{c} \begin{pmatrix} 1 & 2 & 2 & 1 \\ 3 & 6 & 6 & 3 \\ 4 & 9 & 9 & 5 \\ -2 & -1 & -1 & 1 \\ 5 & 8 & 9 & 4 \\ 4 & 2 & 7 & 3 \end{pmatrix} \begin{array}{c} \boldsymbol{r_1} \\ \boldsymbol{r_2} \\ \boldsymbol{r_3} \\ \boldsymbol{r_4} \\ \boldsymbol{r_5} \\ \boldsymbol{r_6} \end{array} \\ \begin{array}{cccc} \boldsymbol{c_1} & \boldsymbol{c_2} & \boldsymbol{c_3} & \boldsymbol{c_4} \end{array} \end{array} \xrightarrow[\text{Elimination}]{\text{Gaussian}} \begin{array}{c} \text{pivot columns} \\ \begin{pmatrix} 1 & 2 & 2 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{array}{c} \mathbf{v_1} \\ \mathbf{v_2} \\ \mathbf{v_3} \\ \\ \\ \\ \end{array} \\ \text{REF} \end{array}$$

row space of $\boldsymbol{A}$ = span$\{\mathbf{r_1}\ \mathbf{r_2}\ \mathbf{r_3}\ \mathbf{r_4}\ \mathbf{r_5}\ \mathbf{r_6}\}$ $\subseteq \mathbb{R}^4$

column space of $\boldsymbol{A}$ = span$\{\mathbf{c_1}\ \mathbf{c_2}\ \mathbf{c_3}\ \mathbf{c_4}\}$ $\subseteq \mathbb{R}^6$

Basis for row space of $\boldsymbol{A}$ = $\{\mathbf{v_1}\ \mathbf{v_2}\ \mathbf{v_3}\}$

Basis for column space of $\boldsymbol{A}$ = $\{\mathbf{c_1}\ \mathbf{c_2}\ \mathbf{c_3}\}$

m x n matrix A	Subspace	Basis	Dimension
Row space	Subspace of $\mathbf{R}^n$	Non-zero rows in REF / RREF	Rank (# non-zero rows in REF)
Column space	Subspace of $\mathbf{R}^m$	Corresponding "pivot" columns in $\boldsymbol{A}$	Rank (# pivot columns in REF)
Nullspace $A\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = 0$	Subspace of $\mathbf{R}^n$ Same as solution space of $\boldsymbol{Ax} = \boldsymbol{0}$	From the spanning vectors in the general solution	Nullity (# parameters in general solution) # non-pivot columns

Let $A = (a_1\ \ a_2\ \ a_3\ \ a_4\ \ a_5)$ be a 4×5 matrix such that the columns a_1, a_2, a_3 are linearly independent while $a_4 = a_1 - 2a_2 + a_3$ and $a_5 = a_2 + a_3$.

(a) Determine the reduced row-echelon form of A. (Hint: The linear relations between columns will not be changed by row operations. In this question, the fifth

Let R be the reduced row-echelon form of A. Since a_1, a_2, a_3 are linearly independent, the first three columns of R are linearly independent. Thus the first three columns of R must be $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$. Together with the information given for the fourth and fifth columns, $R = \begin{pmatrix} 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & -2 & 1 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$.

solution space of $Ax=0$
= nullspace of matrix A
$\text{Null}(A) = \{u \in \mathbb{R}^n \mid Au = 0\} \subseteq \mathbb{R}^n$

Theorem 4.1.11

Let A and B be row equivalent matrices. Then the following statements hold:
(1) A given set of columns of A is linearly independent iff the set of corresponding columns of B is linearly independent.
(2) A given set of columns of A forms a basis for the column space of A iff the set of corresponding columns of B forms a basis for the column space of B.

$\forall c_i \in \mathbb{R}$ $\quad c_1a_1 + c_2a_2 + \ldots + c_na_n = 0$
$\Leftrightarrow c_1b_1 + c_2b_2 + \ldots + c_nb_n = 0$
e.r.o. preserve linear relations between columns

Row(A) = Row(B)
e.r.o. preserves row space

span the given vector set

→ Note : diff from basis of sol^n space (use general sol^n & free params)
span sol^n span

Deriving the Basis for a vector set

Row Space Method

Method 1: We place the vectors u_1, u_2, u_3, u_4, u_5, u_6 in the form of rows in a 6×4 matrix as shown.

$$A = \begin{pmatrix} 1 & 2 & 2 & 1 \\ 3 & 6 & 6 & 3 \\ 4 & 9 & 9 & 5 \\ -2 & -1 & -1 & 1 \\ 5 & 8 & 9 & 4 \\ 4 & 2 & 7 & 3 \end{pmatrix}.$$

We perform row operations to obtain a row-echelon form of A:

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

Thus by Remark 4.1.9, $\{(1,2,2,1), (0,1,1,1), (0,0,1,1)\}$ is a basis for the subspace spanned by u_1, u_2, u_3, u_4, u_5, u_6.

Column Space Method

Method 2: We place the vectors u_1, u_2, u_3, u_4, u_5, u_6 in the form of columns in a 4×6 matrix as shown.

$$B = \begin{pmatrix} 1 & 3 & 4 & -2 & 5 & 4 \\ 2 & 6 & 9 & -1 & 8 & 2 \\ 2 & 6 & 9 & -1 & 9 & 7 \\ 1 & 3 & 5 & 1 & 4 & 3 \end{pmatrix}.$$

We perform row operations to obtain a row-echelon form of B:

$$\begin{pmatrix} 1 & 3 & 4 & -2 & 5 & 4 \\ 0 & 0 & 1 & 3 & -2 & -6 \\ 0 & 0 & 0 & 0 & 1 & 5 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

We see that the first, third and fifth columns are pivot columns. Therefore by Remark 4.1.13, $\{u_1, u_3, u_5\}$ is a basis for the subspace spanned by u_1, u_2, u_3, u_4, u_5, u_6.

Extending a Set to a Basis

- Add on non-redundant vectors to the given set
- Arrange the vectors as rows of a matrix
- Look for 'missing' leading entries of the REF

Extend means use original set, not REF

CANNOT use column space method to extend basis

$S = \{(1, 4, -2, 5, 1), (2, 9, -1, 8, 2), (2, 9, -1, 9, 3)\}$

$$A = \begin{pmatrix} 1 & 4 & -2 & 5 & 1 \\ 2 & 9 & -1 & 8 & 2 \\ 2 & 9 & -1 & 9 & 3 \end{pmatrix} \xrightarrow[\text{Elimination}]{\text{Gaussian}} R = \begin{pmatrix} 1 & 4 & -2 & 5 & 1 \\ 0 & 1 & 3 & -2 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{pmatrix}$$

1. Form a matrix A using the vectors in S as rows.
2. Reduce A to a row-echelon form R.
3. Identify the non-pivot columns of R.
   Look for columns without leading entries
   the 3rd and the 5th columns

Example 4.1.14.2

form a basis for R^5
complete R to a 5x5 matrix

$$A = \begin{pmatrix} 1 & 4 & -2 & 5 & 1 \\ 2 & 9 & -1 & 8 & 2 \\ 2 & 9 & -1 & 9 & 3 \end{pmatrix} \xrightarrow[\text{Elimination}]{\text{Gaussian}} R = \begin{pmatrix} 1 & 4 & -2 & 5 & 1 \\ 0 & 1 & 3 & -2 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & x & * & * \\ 0 & 0 & 0 & 0 & y \end{pmatrix}$$

are not redundant in row space of A $\left\{ \begin{array}{l} \text{E.g. } (0\ 0\ 1\ 0\ 0) \\ \text{E.g. } (0\ 0\ 0\ 0\ 1) \end{array} \right.$

4. Form (row) vectors with leading entries at the non-pivot columns.
5. {Row vectors in A} ∪ {vectors from Step 4 } form a basis for R^n

$$\boxed{\{ (1, 4, -2, 5, 1), (2, 9, -1, 8, 2), (2, 9, -1, 9, 3), (0, 0, 1, 0, 0), (0, 0, 0, 0, 1) \}}$$

Finding a Basis from a Given Set

- Throw out redundant vectors from the given set
- Arrange the vectors as columns of a matrix
- Look for pivot columns of the REF

CANNOT use row space method to find redundant vector

Solutions of LS Ax = b

Approach 1:

- form RREF(A | b)

Approach 2:

If A is a square matrix

- A is invertible => unique solution
- A is singular => no or infinite solutions

Approach 3:

A is any matrix

- b belongs to column space of A => unique or infinite sol
- b does not belong to column space of A => no sol

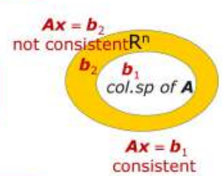
system $Ax = b$ has a solution

⇕

b can be written as a linear combination of the columns of A

⇕

b belongs to the column space of A



Theorem 4.1.16

Theorem 4.1.16 Let A be an $m \times n$ matrix. Then

$$\text{the column space of } A = \{ Au \mid u \in \mathbb{R}^n \}.$$

Hence a system of linear equations $Ax = b$ is consistent if and only if b lies in the column space of A.

span$\{c_1, c_2, \ldots, c_n\}$ = {all linear combi of the column vectors of A}

⟺ $\{u_1 \ldots u_n \mid b\}$ consistent

⟺ solve for $Ax = b$ some u s.t. $Au = b$

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

Following Remark 2.2.17, we rewrite the linear system as

$$x \begin{pmatrix} 2 \\ 1 \\ -5 \\ 1 \end{pmatrix} + y \begin{pmatrix} -1 \\ -1 \\ 1 \\ 0 \end{pmatrix} + z \begin{pmatrix} 0 \\ 3 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ 4 \\ -2 \\ 3 \end{pmatrix}.$$

We see that a solution to the system is a way of writing the vector b as a linear combination of the columns of A. For example, $x = 1$, $y = 3$ and $z = 2$ is a solution to the system. Hence $(-1, 4, -2, 3)^T$ can be written as the linear combination

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

Theorem 4.2.1

The row space and column space of a matrix have the same dimension.

Rank

Def. dimension of its row space or column space.

Denoted by rank(A)

Dimension is for vector space.

Rank is for matrix.

If R is a row-echelon form of A,

rank(A)

= no. of non-zero rows of R

= no. of leading entries in R

= no. of pivot columns in R

= largest no. of linearly independent rows in A

= largest no. of linearly independent columns in A

= no. of columns - nullity(A)

= dim(row space of A) = dim(column space of A)

Row (column) space of zero matrix $\mathbf{0}$ = zero space

rank($\mathbf{0}$) = 0

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

Row (column) space of $n \times n$ identity matrix $\boldsymbol{I}_n = \mathbf{R}^n$

rank($\boldsymbol{I}_n$) = n

$$\boldsymbol{I}_n = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

nullity(A)

= dim(nullspace of A)

= no. of non-pivot columns of A

Remark 4.2.5

(1) For an m x n matrix A, rank(A) ≤ min{m, n}.

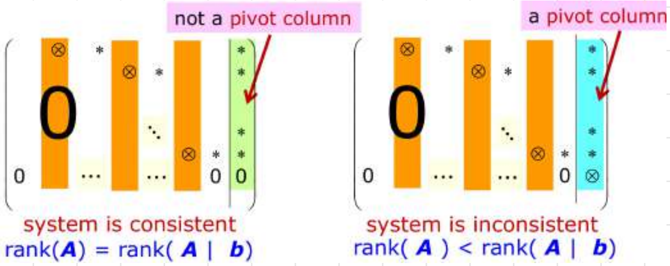
An m x n matrix A with rank(A) = min{m, n} is said to be of full rank.

(2) A square matrix A is of full rank iff det(A) ≠ 0 (invertible).

(3) rank(A) = rank(A^T) for any matrix A because row space of A = column space of A^T

Remark 4.2.6

Ax = b is consistent iff the coefficient of matrix A and the augmented matrix (A | b) have the same rank.



The nullspace of a 3×4 matrix $\boldsymbol{A}$ is given by span $\left\{ \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix} \right\}$.

Determine whether each of the following is true or false:

(i) The first two columns of $\boldsymbol{A}$ are linearly independent.

(ii) The second and fourth columns of $\boldsymbol{A}$ are identical.

Show your working below.

$\begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}$ being in the nullspace of $\boldsymbol{A}$

means the homogeneous system $\boldsymbol{A}\boldsymbol{x} = \mathbf{0}$ has a general solutions $\begin{cases} x = 0 \\ y = t \\ z = s + t \\ w = s \end{cases}$

which gives two equations $x = 0$ and $y - z + w = 0$.

Since the nullity of $\boldsymbol{A}$ is 2, by dimension theorem, the rank of $\boldsymbol{A}$ is also 2.

Hence the rref $\boldsymbol{R}$ of $\boldsymbol{A}$ is given by

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

We see that the first two columns of $\boldsymbol{R}$ are linearly independent. So the corresponding columns in $\boldsymbol{A}$ must also be linearly independent.

Hence (i) is true.

We also observe that the second and fourth columns of $\boldsymbol{R}$ are identical. So the corresponding columns in $\boldsymbol{A}$ must also be identical.

Hence (ii) is true.

Theorem 4.2.8

A: m x n matrix

B: n x p matrix

rank(AB) ≤ min{ rank(A), rank (B) }

column space of AB ⊆ column space of A

(NOT ⊆ column space of B)

row space of AB ⊆ row space of B (NOT ⊆ row space of A)

rank(AB) ≤ rank(A) = rank (B)

Exercise 4 Q23

rank(A + B) ≤ rank(A) + rank(B)

Nullspace & Nullity

A: m x n matrix

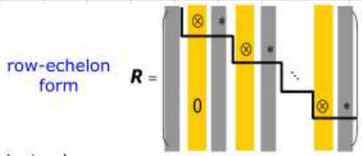
Def. Nullspace of A (subspace of Rn) = the solution space of the homogeneous system of Ax = 0

Def. Nullity of A (a no. ≤ n) is the dimension of the nullspace of A

Denoted by nullity(A) = no. of free parameters in the general solution.

Theorem 4.34 (Dimension Theorem for Matrices)

If A is a matrix with n columns, then rank(A) + nullity(A) = n.



pivot columns

(correspond to basis for column space of $\boldsymbol{A}$) rank($\boldsymbol{A}$)

non-pivot columns

(correspond to parameters in general solutions) nullity($\boldsymbol{A}$)

Size of $\boldsymbol{A}$	# column of $\boldsymbol{A}$	# column of $\boldsymbol{A}^T$	rank($\boldsymbol{A}$) rank($\boldsymbol{A}^T$)	nullity($\boldsymbol{A}$)	nullity($\boldsymbol{A}^T$)
3×4	4	3	3	1	0
7×5	5	7	2	3	5
3×2	2	3	0	2	3

Theorem 4.3.6 (Solution Set of Non-Homogenous System)

Suppose the system of linear equations $\boldsymbol{Ax} = \boldsymbol{b}$ has a (particular) solution $\boldsymbol{v}$.

The solution set of $\boldsymbol{Ax} = \boldsymbol{b}$
$= \{ \underbrace{\boldsymbol{u}}_{\text{vary}} + \underbrace{\boldsymbol{v}}_{\text{fix}} \mid \boldsymbol{u}$ is an element of the nullspace of $\boldsymbol{A} \}$

The general solution of $\boldsymbol{Ax} = \boldsymbol{b}$ can be given by

x = (the general solution of $\boldsymbol{Ax} = \boldsymbol{0}$) + $\boldsymbol{v}$

If we know the general solution of $\boldsymbol{Ax} = \boldsymbol{0}$
and one particular solution of $\boldsymbol{Ax} = \boldsymbol{b}$,
then we have the general solution for $\boldsymbol{Ax} = \boldsymbol{b}$.

Remark 4.3.7

Let Ax = b be a consistent linear system. Then

Ax = b has exactly one solution iff the nullspace of A is equal to {0} (trivial solution)

W9

Dot Product (inner product)

Dot product	$\boldsymbol{u} \cdot \boldsymbol{v}$	$u_1v_1 + u_2v_2 + \dots u_nv_n$
Norm (length)	$\lVert \boldsymbol{u} \rVert$ $\lVert \vec{u} \rVert = 1 \Rightarrow$ unit vector	$\sqrt{\boldsymbol{u} \cdot \boldsymbol{u}}$ $\sqrt{u_1^2 + u_2^2 + \dots + u_n^2}$
Distance	$\lVert \boldsymbol{u} - \boldsymbol{v} \rVert$ $d(u, v)$	$\sqrt{(\boldsymbol{u}-\boldsymbol{v}) \cdot (\boldsymbol{u}-\boldsymbol{v})}$ $\sqrt{(u_1-v_1)^2 + (u_2-v_2)^2 + \dots + (u_n-v_n)^2}$
Angle	between $\boldsymbol{u}$ and $\boldsymbol{v}$ $\vec{u} \cdot \vec{v} = \cos\theta \lVert \vec{u} \rVert \lVert \vec{v} \rVert$	$\cos^{-1}\left(\frac{\boldsymbol{u} \cdot \boldsymbol{v}}{\lVert \boldsymbol{u} \rVert \lVert \boldsymbol{v} \rVert}\right)$ $\cos^{-1}\left(\frac{u_1v_1 + u_2v_2 + \dots + u_nv_n}{\lVert \boldsymbol{u} \rVert \lVert \boldsymbol{v} \rVert}\right)$

The angle is well-defined because $-1 \le \frac{\boldsymbol{u} \cdot \boldsymbol{v}}{\lVert \boldsymbol{u} \rVert \lVert \boldsymbol{v} \rVert} \le 1$.
(See Question 5.4(a).)

Dot Product as Matrix Multiplication

$\boldsymbol{u}$ and $\boldsymbol{v}$ regarded as row matrices $\mathbf{u} \cdot \mathbf{v} = \mathbf{u}\mathbf{v}^T$

$\boldsymbol{u}$ and $\boldsymbol{v}$ regarded as column matrices $\mathbf{u} \cdot \mathbf{v} = \mathbf{u}^T\mathbf{v}$

$\lVert u \rVert = 1 \Rightarrow$ set of all unit vectors from 0
$\Rightarrow$ circle in $\mathbb{R}^2$
$\Rightarrow$ sphere in $\mathbb{R}^3$.

Theorem 5.1.5 (Properties of Dot Product)

Let c be a scalar and $\mathbf{u}, \mathbf{v}, \mathbf{w}$ vectors in $\mathbb{R}^n$.

1. $\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$ — commutative law
2. $(\mathbf{u} + \mathbf{v}) \cdot \mathbf{w} = \mathbf{u} \cdot \mathbf{w} + \mathbf{v} \cdot \mathbf{w}$
   $\mathbf{w} \cdot (\mathbf{u} + \mathbf{v}) = \mathbf{w} \cdot \mathbf{u} + \mathbf{w} \cdot \mathbf{v}$ — distributive law
3. $(c\mathbf{u}) \cdot \mathbf{v} = \mathbf{u} \cdot (c\mathbf{v}) = c(\mathbf{u} \cdot \mathbf{v})$ — scalar mult.
4. $||c\mathbf{u}|| = |c|\,||\mathbf{u}||$ — norm (not $c\,||\mathbf{u}||$)
5. (i) $\mathbf{u} \cdot \mathbf{u} \geq 0$ — $\mathbf{u} \cdot \mathbf{u} = ||\mathbf{u}||^2$
   (ii) $\mathbf{u} \cdot \mathbf{u} = 0$ if and only if $\mathbf{u} = \mathbf{0}$.
6. $|\mathbf{u} \cdot \mathbf{v}| \leq ||\mathbf{u}|| \cdot ||\mathbf{v}||$ — Cauchy-Schwarz Inequality

$||u+v|| \leq ||u|| + ||v||$

$u \cdot v = 0 \begin{cases} u = 0 \text{ or } v = 0 \\ \theta = \frac{\pi}{2} \end{cases}$

not always $\perp$

Additional e.g.: $Av = 0$ iff $A^TAv = 0$

- Algebraically, zero vector is orthogonal to all vectors/ subspaces, since (0 . v) = 0 for all v.
- But geometrically, we do not regard the zero vector to be perpendicular to any non-zero vector.
- If not mentioned, an orthogonal set may contain the 0 vector.
- However, orthonormal sets do not contain the 0 vector
- Orthonormal sets are linearly indep but orthogonal sets may be linearly dependent (e.g. contains the 0 vector)
- a set with only one vector is orthogonal by convention
- orthogonal sets with n **non-zero** vectors are bases for the R^n space
- there can exist a maximum of (n+1) orthogonal vectors in a R^n space with the (n+1)th vector being the 0 vector

Orthogonal & Orthonormal

=/= Linear independence

(1) Two vectors u and v in Rn are called orthogonal **u · v = 0**. (geometrically: u and v are perpendicular)

(2) A set S = {u1, u2, ..., uk} is an orthogonal set if every pair of distinct vectors in S are orthogonal:

u1 · u2 = 0, u1 · u3 = 0, ... uk-1 · uk =0

A set S of vectors in Rn is called an **orthonormal** set if S is an **orthogonal** set and every vector in S is a **unit vector**.

(3) A set S = {u1, u2, ..., uk} is an orthogonal basis of a vector space V if S is an orthogonal set and a basis for V.

(4) A vector u is orthogonal to a vector space V if u is orthogonal to every vector in V.

(5) A set S of vectors in Rn is called orthonormal if S is **orthogonal** and every vector in S is a **unit vector**.

(6) A basis S for a vector space is called an **orthogonal basis** if S is orthogonal.

(7) A basis S for a vector space is called an **orthonormal basis** if S is orthonormal.

(8) Let V be a subspace of Rn. A vector u is **orthogonal to the subspace V if u is orthogonal to all vectors in V.**

(9) Orthogonal matrix is a subset of linearly indep matrix. (E.g. (1, 0) and (1, 1) are linearly indep but not orthogonal - 45°)

(10) any set with only 1 vector is orthogonal (vacuously true)

(11) a subspace that is orthogonal to the zero space is a full rank subspace. (Whole Rn is orthogonal of zero space)

Remark 5.2.2 Given two vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ in $\mathbb{R}^n$, if they are orthogonal, then the angle between them is equal to

$$\cos^{-1}\left(\frac{\boldsymbol{u} \cdot \boldsymbol{v}}{||\boldsymbol{u}||\,||\boldsymbol{v}||}\right) = \cos^{-1}(0) = \frac{\pi}{2}.$$

Thus the concept of "orthogonal" in $\mathbb{R}^n$ is the same as the concept of "perpendicular" in $\mathbb{R}^2$ and $\mathbb{R}^3$.

$$\underset{\text{any non-zero vector}}{\mathbf{u}} \xrightarrow{\text{normalizing}} \underset{\text{unit vector}}{\frac{1}{||\mathbf{u}||}\mathbf{u}}$$ Scalar multiple of the original vector

$$\underset{\text{an orthogonal set}}{\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}} \xrightarrow{\text{normalizing}} \underset{\text{an orthonormal set}}{\left\{\frac{1}{||\mathbf{u}_1||}\mathbf{u}_1, \frac{1}{||\mathbf{u}_2||}\mathbf{u}_2, \frac{1}{||\mathbf{u}_3||}\mathbf{u}_3\right\}}$$

(a) Clearly $\text{span}(T) \subseteq \text{span}(S)$. We just need to show $\text{span}(S) \subseteq \text{span}(T)$. Since $\boldsymbol{u_1} = \boldsymbol{v_3}$, $\boldsymbol{u_2} = \frac{3}{5}\boldsymbol{v_1} + \frac{4}{5}\boldsymbol{v_2}$, $\boldsymbol{u_3} = \frac{4}{5}\boldsymbol{v_1} - \frac{3}{5}\boldsymbol{v_2}$, $\text{span}(S) \subseteq \text{span}(T)$ follows.

show $v_i \cdot v_j = \begin{cases} 0 \text{ if } i=j & \text{orthogonal} \\ 1 \text{ if } i \neq j & \text{unit vector} \end{cases}$

(b) Since S is orthonormal, $\boldsymbol{v_1} \cdot \boldsymbol{v_1} = \frac{9}{25}(\boldsymbol{u_2} \cdot \boldsymbol{u_2}) + \frac{16}{25}(\boldsymbol{u_3} \cdot \boldsymbol{u_3}) + \frac{24}{25}(\boldsymbol{u_2} \cdot \boldsymbol{u_3}) = 1$. Likewise, it can be shown that $\boldsymbol{v_2} \cdot \boldsymbol{v_2} = \boldsymbol{v_3} \cdot \boldsymbol{v_3} = 1$, $\boldsymbol{v_1} \cdot \boldsymbol{v_2} = \boldsymbol{v_1} \cdot \boldsymbol{v_3} = \boldsymbol{v_2} \cdot \boldsymbol{v_3} = 0$. Hence T is also orthonormal.

Theorem 5.24

if $S \subseteq V$ $\Leftarrow$ orthogonal basis

Let S be an orthogonal set of nonzero vectors in a vector space. Then S is linearly independent.

$\therefore$ orthonormal set is also linearly indep

$V = \text{span}\{u_1, \dots, u_k\}. = \text{col}(A)$

$u \perp V \iff u \cdot u_i = 0 \quad \forall i = 1, \dots, k$

$\iff A^T u = 0, \; A = (u_1, \dots, u_k)$

$\iff u \in \text{Null}(A)$

$A^T u = \begin{pmatrix} u_1^T u \\ \vdots \\ u_k^T u \end{pmatrix}$

Checking for Orthogonal Basis

A set S of nonzero vectors in a vector space V.

To check whether S is an orthonormal basis for V:

S is an orthogonal set $\xrightarrow{\text{Theorem 5.2.4}}$ S lin. indep

S spans V

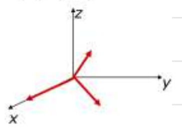
Only need to check:
(i) S is orthonormal (orthogonal) and
(ii) span(S) = V.

If we know dim V,
Only need to check:
(i) S is orthonormal (orthogonal) and
(ii) $|S|$ = dim V.

$u_1 = (2, 0, 0)$ $u_2 = (0, 1, 1)$ $u_3 = (0, 1, -1)$

$\{u_1, u_2, u_3\}$
- an orthogonal set
- has three vectors $=$ dim $\mathbf{R}^3$

$\Rightarrow$ an orthogonal basis for $\mathbf{R}^3$.



Theorem 5.2.8

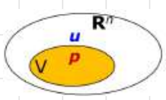
1. If $S = \{u_1, u_2, \dots, u_k\}$ is an orthogonal basis for a vector space V, then for any vector w in V, $w = c_1u_1 + c_2u_2 + \dots + c_ku_k$
$$w = \frac{w \cdot u_1}{u_1 \cdot u_1}u_1 + \frac{w \cdot u_2}{u_2 \cdot u_2}u_2 + \cdots + \frac{w \cdot u_k}{u_k \cdot u_k}u_k,$$
i.e. $(w)_S = \left(\frac{w \cdot u_1}{u_1 \cdot u_1}, \frac{w \cdot u_2}{u_2 \cdot u_2}, \dots, \frac{w \cdot u_k}{u_k \cdot u_k}\right) = \left(\frac{w \cdot u_1}{||u_1||^2}, \frac{w \cdot u_2}{||u_2||^2}, \frac{w \cdot u_3}{||u_3||^2}\right)$
2. If $T = \{v_1, v_2, \dots, v_k\}$ is an orthonormal basis for a vector space V, then for any vector w in V, $||u_i||^2 = 1$
$$w = (w \cdot v_1)v_1 + (w \cdot v_2)v_2 + \cdots + (w \cdot v_k)v_k,$$
i.e. $(w)_T = (w \cdot v_1, w \cdot v_2, \dots, w \cdot v_k)$.

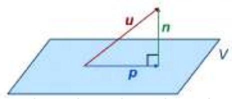
Vector Projection (Orthogonal projection)

Let V be a subspace of $\mathbb{R}^n$.

Every $u \in \mathbb{R}^n$ can be written uniquely (Exactly one) as
$$u = n + p$$
where p is a vector in V
and n is a vector orthogonal to V. $n \perp V$

The vector p is called the (orthogonal) projection of u onto V.





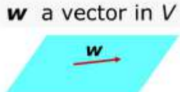
- p is the vector in V "nearest" to the given vector v
- p is the **best approximation** of v in the subspace V

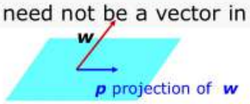
Theorem 5.2.15

Theorem 5.2.15 Let V be a subspace of $\mathbb{R}^n$ and w a vector in $\mathbb{R}^n$.

1. If $\{u_1, u_2, \dots, u_k\}$ is an orthogonal basis for V, then
$$\frac{w \cdot u_1}{u_1 \cdot u_1}u_1 + \frac{w \cdot u_2}{u_2 \cdot u_2}u_2 + \cdots + \frac{w \cdot u_k}{u_k \cdot u_k}u_k$$
is the projection of w onto V.
2. If $\{v_1, v_2, \dots, v_k\}$ is an orthonormal basis for V, then
$$(w \cdot v_1)v_1 + (w \cdot v_2)v_2 + \cdots + (w \cdot v_k)v_k$$
is the projection of w onto V.

Theorem 5.2.8	Theorem 5.2.15
w a vector in V	w need not be a vector in V





V a subspace	V a subspace
$S = \{u_1, u_2, \dots, u_k\}$ orthogonal basis	$S = \{u_1, u_2, \dots, u_k\}$ orthogonal basis

$$\frac{w \cdot u_1}{||u_1||^2}u_1 + \frac{w \cdot u_2}{||u_2||^2}u_2 + \cdots + \frac{w \cdot u_k}{||u_k||^2}u_k = \begin{cases} w & \text{if } w \in V \\ p & \text{if } w \notin V \end{cases}$$

Finding an Orthogonal Vector of a Subspace

To find a vector v that is orthogonal to a subspace $V = \text{span}\{u_1, u_2, \dots, u_k\}$ of $\mathbf{R}^n$

Let $v = (x_1, x_2, \dots, x_n)$ (unknowns)
Convert $v \cdot u_1 = 0$, $v \cdot u_2 = 0$, ... , $v \cdot u_k = 0$
into a homogeneous system.
Solve the system.

$$\left(\begin{array}{c|c} u_1 & 0 \\ \vdots & \vdots \\ u_k & 0 \end{array}\right)$$

Example

$u_1 = (1, 1, 1)$, $u_2 = (1, 0, -1)$ and $u_3 = (1, -2, 1)$.

$S = \{u_1, u_2, u_3\}$ is an orthogonal basis for $\mathbf{R}^3$.

Let $w = (1, -1, 0)$. Find $(w)_S$ coordinate vector w.r.t. basis S

$w = c_1u_1 + c_2u_2 + c_3u_3 \Rightarrow (w)_S = (c_1, c_2, c_3)$

Theorem 5.2.8 (when S is orthogonal):

$$(w)_S = \left(\frac{w \cdot u_1}{||u_1||^2}, \frac{w \cdot u_2}{||u_2||^2}, \frac{w \cdot u_3}{||u_3||^2}\right) = \left(0, \frac{1}{2}, \frac{1}{2}\right)$$

$v_1 = \left(\frac{3}{5}, \frac{4}{5}\right)$ $v_2 = \left(\frac{4}{5}, -\frac{3}{5}\right)$ $||v_i||^2 = 1$ $v_1 \cdot v_2 = 0$

$S = \{v_1, v_2\}$ is an orthonormal basis for $\mathbf{R}^2$.

Let $w = (x, y)$ be any vector in $\mathbf{R}^2$.

Express $(w)_s$ in terms of x and y

$w = (w \cdot v_1)v_1 + (w \cdot v_2)v_2$

$$\left.\begin{array}{l} w \cdot v_1 = \frac{3x+4y}{5} \\ w \cdot v_2 = \frac{4x-3y}{5} \end{array}\right\} \Rightarrow w = \frac{3x+4y}{5}v_1 + \frac{4x-3y}{5}v_2$$

$$(w)_S = \left(\frac{3x+4y}{5}, \frac{4x-3y}{5}\right)$$

1. Let V be a plane in $\mathbb{R}^3$ defined by the equation $ax + by + cz = 0$. Let $n = (a, b, c)$. For any vector $u = (x, y, z)$ in V,
$$n \cdot u = ax + by + cz = 0.$$
Thus n is orthogonal to V. In fact,
$$V = \{(x, y, z) \in \mathbb{R}^3 \mid ax + by + cz = 0\} = \{u \in \mathbb{R}^3 \mid n \cdot u = 0\}.$$
The vector n is called a *normal vector* of V.

2. Let $V = \text{span}\{u_1, u_2\}$ be a subspace of $\mathbb{R}^4$ where $u_1 = (1, 1, 1, 0)$ and $u_2 = (0, -1, -1, 1)$. Find all vectors that are orthogonal to V.

Solution Let $v = (w, x, y, z)$ be a vector in $\mathbb{R}^4$. Then
$$\begin{aligned} & v \cdot (au_1 + bu_2) = 0 \text{ for all } a, b \in \mathbb{R} \\ \Leftrightarrow\ & v \cdot u_1 = 0 \text{ and } v \cdot u_2 = 0 \\ \Leftrightarrow\ & \begin{cases} w + x + y = 0 \\ -x - y + z = 0 \end{cases} \quad \left(\begin{array}{cccc|c} 1 & 1 & 1 & 0 & 0 \\ 0 & -1 & -1 & 1 & 0 \end{array}\right) \longrightarrow \left(\begin{array}{cccc|c} 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & -1 & 0 \end{array}\right) \\ \Leftrightarrow\ & (w, x, y, z) = (-t, -s + t, s, t) \text{ for some } s, t \in \mathbb{R}. \end{aligned}$$
So a vector v is orthogonal to V if and only if
$$v = (-t, -s + t, s, t) = s(0, -1, 1, 0) + t(-1, 1, 0, 1)$$
for some $s, t \in \mathbb{R}$, i.e, $v \in \text{span}\{(0, -1, 1, 0), (-1, 1, 0, 1)\}$.

Theorem 5.2.19 (Gram-Schmidt Process)

Theorem 5.2.19 (Gram-Schmidt Process) Let $\{u_1, u_2, \ldots, u_k\}$ be a basis for a vector space V. Let

don't need consider scalar multiple

$$v_1 = u_1,$$

$$v_2 = u_2 - \frac{u_2 \cdot v_1}{v_1 \cdot v_1} v_1, \quad \text{Orthogonal to v1}$$

$$v_3 = u_3 - \frac{u_3 \cdot v_1}{v_1 \cdot v_1} v_1 - \frac{u_3 \cdot v_2}{v_2 \cdot v_2} v_2, \quad \text{Orthogonal to v1 \& v2}$$

$$\vdots$$

$$v_k = u_k - \underbrace{\frac{u_k \cdot v_1}{v_1 \cdot v_1} v_1 - \frac{u_k \cdot v_2}{v_2 \cdot v_2} v_2 - \cdots - \frac{u_k \cdot v_{k-1}}{v_{k-1} \cdot v_{k-1}} v_{k-1}}_{-\text{ projection of } u_k \text{ onto span}\{v_1, \ldots, v_{k-1}\}}. \quad \text{Orthogonal to v1, v2, ..., v(k-1)}$$

Then $\{v_1, v_2, \ldots, v_k\}$ is an orthogonal basis for V. Furthermore, let

$$w_1 = \frac{1}{||v_1||} v_1, \quad w_2 = \frac{1}{||v_2||} v_2, \quad \ldots, \quad w_k = \frac{1}{||v_k||} v_k.$$

Then $\{w_1, w_2, \ldots, w_k\}$ is an orthonormal basis for V.

if you performed the Gram-Schmidt process on a set and obtained a zero vector, then it means there are redundant vectors in the original set of vectors that you tried to GS.

can ignore scalar multiple when calculating $v_2, v_3, \ldots$

Example

Example 5.2.20 Apply the Gram-Schmidt Process to transform the basis $\{u_1, u_2, u_3\}$, where $u_1 = (1, -1, 2)$, $u_2 = (2, 1, 0)$ and $u_3 = (0, 0, 1)$, for $\mathbb{R}^3$ into an orthonormal basis.

Solution Let

$$v_1 = u_1 = (1, -1, 2),$$

$$v_2 = u_2 - \frac{u_2 \cdot v_1}{v_1 \cdot v_1} v_1 = (2, 1, 0) - \frac{1}{6}(1, -1, 2) = \left(\frac{11}{6}, \frac{7}{6}, -\frac{1}{3}\right),$$

$$v_3 = u_3 - \frac{u_3 \cdot v_1}{v_1 \cdot v_1} v_1 - \frac{u_3 \cdot v_2}{v_2 \cdot v_2} v_2 = (0, 0, 1) - \frac{2}{6}(1, -1, 2) - \frac{-1/3}{29/6}\left(\frac{11}{6}, \frac{7}{6}, -\frac{1}{3}\right) = \left(-\frac{6}{29}, \frac{12}{29}, \frac{9}{29}\right).$$

Then

$$\left\{\frac{1}{||v_1||} v_1, \frac{1}{||v_2||} v_2, \frac{1}{||v_3||} v_3\right\}$$

$$= \left\{\frac{1}{\sqrt{6}}(1, -1, 2), \frac{1}{\sqrt{29/6}}\left(\frac{11}{6}, \frac{7}{6}, -\frac{1}{3}\right), \frac{1}{\sqrt{9/29}}\left(-\frac{6}{29}, \frac{12}{29}, \frac{9}{29}\right)\right\}$$

$$= \left\{\left(\frac{1}{\sqrt{6}}, -\frac{1}{\sqrt{6}}, \frac{2}{\sqrt{6}}\right), \left(\frac{11}{\sqrt{174}}, \frac{7}{\sqrt{174}}, -\frac{2}{\sqrt{174}}\right), \left(-\frac{2}{\sqrt{29}}, \frac{4}{\sqrt{29}}, \frac{3}{\sqrt{29}}\right)\right\}$$

is an orthonormal basis for $\mathbb{R}^3$.

$W^\perp$

Pronounced as W-perp

Let W be a subspace of $\mathbf{R}^n$.

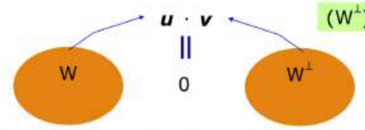
Define $W^\perp = \{ \boldsymbol{u} \in \mathbf{R}^n \mid \boldsymbol{u}$ is orthogonal to W$\}$

Exercise 5 Q7 $W^\perp$ is also a subspace of $\mathbf{R}^n$

Every vector in $W^\perp$ is orthogonal to every vector in W.

$\boldsymbol{u} \cdot \boldsymbol{v}$ = 0

$(W^\perp)^\perp = W$



(b) Let $\{w_1, \ldots, w_k\}$ be a basis for W.

$$u \in W^\perp \Leftrightarrow \begin{cases} w_1 \cdot u = 0 \\ \vdots \\ w_k \cdot u = 0 \end{cases} \Leftrightarrow \begin{pmatrix} w_1 \\ \vdots \\ w_k \end{pmatrix} u^T = 0$$

So $W^\perp$ is a solution set of a homogeneous system of linear equations. By Theorem 3.3.6, $W^\perp$ is a subspace of $\mathbb{R}^n$.

Alternative proof: Since $w \cdot 0 = 0$ for all $w \in W$, $0 \in W^\perp$. So $W^\perp$ is nonempty. Let u and v be any vectors in $W^\perp$, i.e. $w \cdot u = 0$ and $w \cdot v = 0$ for all $w \in W$, and let $a, b \in \mathbb{R}$. Then for all $w \in W$, $w \cdot (au + bv) = a(w \cdot u) + b(w \cdot v) = 0 + 0 = 0$. Hence $u + v \in W^\perp$. By Remark 3.3.8, $W^\perp$ is a subspace of $\mathbb{R}^n$.

$W^\perp = \text{null}(A^T)$

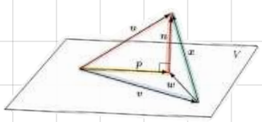
$A: n \times k$

$A^T = k \times n$

$\dim(W) = k = \text{rank}(A) = \text{rank}(A^T)$

$\dim(W^\perp) = \text{nullity}(A^T) = n - \text{rank}(A^T)$
$= n - k$

W10



Theorem 5.3.2

Theorem 5.3.2 Let V be a subspace in $\mathbb{R}^n$. If $\boldsymbol{u}$ is a vector in $\mathbb{R}^n$ and $\boldsymbol{p}$ is the projection of $\boldsymbol{u}$ onto V, then

$$\underbrace{d(\boldsymbol{u}, \boldsymbol{p})}_{||\vec{n}||} \le \underbrace{d(\boldsymbol{u}, \boldsymbol{v})}_{||\vec{x}||} \quad \text{for all } \boldsymbol{v} \in V,$$

i.e. $\boldsymbol{p}$ is the *best approximation* of $\boldsymbol{u}$ in V.

Let $V = \text{span}\{(1, 0, 1), (1, 1, 1)\}$ which is a plane in $\mathbb{R}^3$ containing the origin.

Find the (shortest) distance from $\boldsymbol{u} = (1, 2, 3)$ to V.

For plane,

$$d_\perp = \frac{|a \cdot n - q \cdot n|}{||n||}$$

position vector on plane $= d$ in $ax + by + cz = d$

$n = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$

Solution: The shortest distance from $\boldsymbol{u}$ to V is $d(\boldsymbol{u}, \boldsymbol{p})$ where $\boldsymbol{p}$ is the projection of $\boldsymbol{u}$ onto V (by Theorem 5.3.2).

First, applying the Gram-Schmidt Process (Theorem 5.2.19), the vectors

$$(1, 0, 1) \text{ and } (1, 1, 1) - \frac{(1, 1, 1) \cdot (1, 0, 1)}{(1, 0, 1) \cdot (1, 0, 1)}(1, 0, 1) = (0, 1, 0)$$

form an orthogonal basis for V.

Thus (by Theorem 5.2.15)

$$\boldsymbol{p} = \frac{(1, 2, 3) \cdot (1, 0, 1)}{(1, 0, 1) \cdot (1, 0, 1)}(1, 0, 1) + \frac{(1, 2, 3) \cdot (0, 1, 0)}{(0, 1, 0) \cdot (0, 1, 0)}(0, 1, 0)$$

$$= (2, 2, 2)$$

and the distance from $\boldsymbol{u}$ to V is

$$d(\boldsymbol{u}, \boldsymbol{p}) = ||\boldsymbol{u} - \boldsymbol{p}|| = ||(1, 2, 3) - (2, 2, 2)||$$
$$= ||(-1, 0, 1)|| = \sqrt{2}.$$

Least Square Solutions

When a linear system $\mathbf{Ax} = \mathbf{b}$ is inconsistent

A least squares solution $\mathbf{x}_0$ of $\mathbf{Ax} = \mathbf{b}$:

- $\mathbf{x}_0$ is the best approximation to a solution of $\mathbf{Ax} = \mathbf{b}$

$\mathbf{Ax}_0 \approx \mathbf{b}$ (closest) $\quad ||\mathbf{b} - \mathbf{Ax}_0||$ (smallest)

i.e. $||\boldsymbol{b} - \boldsymbol{Ax}_0|| \le ||\boldsymbol{b} - \boldsymbol{Av}||$ for all $\boldsymbol{v}$ in $\mathbf{R}^n$

projection of b onto col(A) i.e. $(Ax_0 - b) \perp \text{col}(A)$

To find the least squares solution $\mathbf{x}_0$ of $\mathbf{Ax} = \mathbf{b}$:

solve the new system $\mathbf{A}^T\mathbf{Ax} = \mathbf{A}^T\mathbf{b}$ $\Leftrightarrow$ $A^T(Ax_0 - b) = 0$

$\mathbf{Ax} = \mathbf{b}$

consistent	inconsistent
$\mathbf{A}^T\mathbf{Ax} = \mathbf{A}^T\mathbf{b}$ ← Always consistent →	$\mathbf{A}^T\mathbf{Ax} = \mathbf{A}^T\mathbf{b}$
gives the same solution set as $\mathbf{Ax} = \mathbf{b}$	gives the least squares solutions of $\mathbf{Ax} = \mathbf{b}$
Ex 5 Q27	There may be more than one LSS

Projection is unique but least sq solution is not unique.

if $A^TAx = A^Tb$ has infinitely many sol^n

Theorem 5.3.8

Theorem 5.3.8 Let $\boldsymbol{Ax} = \boldsymbol{b}$ be a linear system, where $\boldsymbol{A}$ is an $m \times n$ matrix, and let $\boldsymbol{p}$ be the projection of $\boldsymbol{b}$ onto the column space of $\boldsymbol{A}$. Then

$$||\boldsymbol{b} - \boldsymbol{p}|| \le ||\boldsymbol{b} - \boldsymbol{Av}|| \quad \text{for all } \boldsymbol{v} \in \mathbb{R}^n,$$ d(b,p)≤d(b,w) for all w∈V,

i.e. $\boldsymbol{u}$ is a least squares solution to $\boldsymbol{Ax} = \boldsymbol{b}$ if and only if $\boldsymbol{Au} = \boldsymbol{p}$.

p=Au is the best approximation of b onto V.

Example

Example 5.3.9 Let $\boldsymbol{A} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 1 \end{pmatrix}$, $\boldsymbol{b} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ and

$$V = \text{the column space of } \boldsymbol{A} = \text{span}\left\{ \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\}.$$

By Example 5.3.3, the projection of $\boldsymbol{b}$ onto V is $\boldsymbol{p} = \begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix}$. By Theorem 5.3.8, $\begin{pmatrix} x \\ y \end{pmatrix}$ is a least squares solution to $\boldsymbol{Ax} = \boldsymbol{b}$ if and only if

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix} \quad \Leftrightarrow \quad \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \end{pmatrix}.$$

Theorem 5.3.10

Theorem 5.3.10 (A method to find the least squares solution) Let $\boldsymbol{Ax} = \boldsymbol{b}$ be a linear system. Then $\boldsymbol{u}$ is a least squares solution to $\boldsymbol{Ax} = \boldsymbol{b}$ if and only if $\boldsymbol{u}$ is a solution to $\boldsymbol{A}^T\boldsymbol{Ax} = \boldsymbol{A}^T\boldsymbol{b}$. $\Leftrightarrow$ $A^T(Ax-b)=0$

Example

In this example, we demonstrate how to find the projection using a least squares solution:

Let $V = \text{span}\{(1,-1,1,-1), (1,2,0,1), (2,1,1,0)\}$. Find the projection of $(1,1,1,1)$ onto V.

Solution Let $\boldsymbol{A} = \begin{pmatrix} 1 & 1 & 2 \\ -1 & 2 & 1 \\ 1 & 0 & 1 \\ -1 & 1 & 0 \end{pmatrix}$ and $\boldsymbol{b} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$. We first obtain a least squares solution of $\boldsymbol{Ax} = \boldsymbol{b}$. The equation $\boldsymbol{A}^T\boldsymbol{Ax} = \boldsymbol{A}^T\boldsymbol{b}$ is

$$\begin{pmatrix} 4 & -2 & 2 \\ -2 & 6 & 4 \\ 2 & 4 & 6 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 4 \\ 4 \end{pmatrix}.$$

Solving this linear system, we have $\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -t + \frac{2}{5} \\ -t + \frac{4}{5} \\ t \end{pmatrix}$ where t is an arbitrary parameter. Take any one of the least squares solutions, say $\boldsymbol{u} = \begin{pmatrix} \frac{2}{5} \\ \frac{4}{5} \\ 0 \end{pmatrix}$, and compute $\boldsymbol{Au}$:

$$\begin{pmatrix} 1 & 1 & 2 \\ -1 & 2 & 1 \\ 1 & 0 & 1 \\ -1 & 1 & 0 \end{pmatrix} \begin{pmatrix} \frac{2}{5} \\ \frac{4}{5} \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{6}{5} \\ \frac{6}{5} \\ \frac{2}{5} \\ \frac{2}{5} \end{pmatrix}.$$

By Theorem 5.3.8, $(\frac{6}{5}, \frac{6}{5}, \frac{2}{5}, \frac{2}{5})$ is the projection of $(1,1,1,1)$ onto V.

Let $S = \{v_1, \dots, v_k\}$ & $A = (v_1 \ \dots \ v_k)$

(1) S is an orthogonal set $\Leftrightarrow$ A^TA is a diagonal matrix

(2) S is an orthonormal set $\Leftrightarrow$ $A^TA = I_k$

Finding Least Squares Solutions

If a linear system $\mathbf{Ax} = \mathbf{b}$ is consistent

- A least squares solution $\mathbf{x}_0$ of the system is an actual solution of $\mathbf{Ax} = \mathbf{b}$ itself

If a linear system $\mathbf{Ax} = \mathbf{b}$ is inconsistent

- A least squares solution $\mathbf{x}_0$ of the system is given by the actual solution of $\mathbf{A}^T\mathbf{Ax} = \mathbf{A}^T\mathbf{b}$

A least squares solution $\mathbf{x}_0$ of $\mathbf{Ax} = \mathbf{b}$:

- is the best approximation to a solution of the system $\underset{\text{closest}}{\mathbf{Ax}_0 \approx \mathbf{b}}$ $\quad \underset{\text{smallest}}{||\mathbf{Ax}_0 - \mathbf{b}||}$
- $\mathbf{Ax}_0$ = projection of $\mathbf{b}$ onto column space of $\mathbf{A}$
- $\mathbf{x}_0 \neq$ projection of $\mathbf{b}$

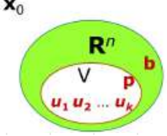
Finding Projection Using Least Squares Methods

Let $\mathbf{x}_0$ be a least squares solution of $\mathbf{Ax} = \mathbf{b}$

Then $\mathbf{Ax}_0$ = projection of $\mathbf{b}$ onto column space of $\mathbf{A}$

To find projection $\mathbf{p}$ of $\mathbf{b}$ onto a subspace V:

- Find any basis $\{\boldsymbol{u}_1, \dots, \boldsymbol{u}_k\}$ for V = column space of $\mathbf{A}$
- Form matrix $\mathbf{A}$ using $\boldsymbol{u}_1, \dots, \boldsymbol{u}_k$ as column vectors
- Solve the system $\mathbf{A}^T\mathbf{Ax} = \mathbf{A}^T\mathbf{b}$ to get $\mathbf{x}_0$
- $\mathbf{Ax}_0$ = projection $\mathbf{p}$



Project $\mathbf{v}$ onto subspace $\mathbf{W}$

- If you have an orthonormal basis $\{\boldsymbol{u}_1, \dots, \boldsymbol{u}_r\}$ for W
  - Use the formula

Projection = $(\boldsymbol{v} \cdot \boldsymbol{u}_1)\boldsymbol{u}_1 + (\boldsymbol{v} \cdot \boldsymbol{u}_2)\boldsymbol{u}_2 + \dots + (\boldsymbol{v} \cdot \boldsymbol{u}_r)\boldsymbol{u}_r$

orthogonal basis :

projection = $\frac{v \cdot u_1}{||u_1||^2}u_1 + \dots + \frac{v \cdot u_r}{||u_r||^2}u_r$

If S is a basis for V, then

projection of w onto $V = A(A^TA)^{-1}A^Tw$

if asked to extend orthogonal basis given v

a new vector would be v - projectn

$v \in Null(A^TA) = Null(A)$

Orthogonal Matrices (sq, invertible matrices)

A square matrix A is called *orthogonal* if $A^{-1} = A^{\mathrm{T}}$.

Remark 5.4.4 By Theorem 2.4.12, a square matrix A is orthogonal if and only if $AA^{\mathrm{T}} = I$ (or $A^{\mathrm{T}}A = I$).

- All orthogonal matrices are invertible(linearly independent).
- Their transposes are also orthogonal matrices
- The transition matrix between two orthonormal bases is an orthogonal matrix
- Note: There is no such things as orthonormal matrix.

The rows of an orthogonal matrix form an orthonormal basis, and so their norms are all equal to 1.

$A = (c_1\ c_2 \dots c_k)$ is an $n \times k$ matrix such that the columns $\{c_1, c_2, \dots, c_k\}$ of A form an orthonormal set.
Can we conclude that
(I) $A^{\mathrm{T}}A$ = identity matrix and (II) AA^{T} = identity matrix?
True False

$$A^{\mathrm{T}}A = \begin{pmatrix} c_1^{\mathrm{T}} \\ \vdots \\ c_k^{\mathrm{T}} \end{pmatrix}(c_1 \cdots c_k) = \begin{pmatrix} c_1^{\mathrm{T}}c_1 & \cdots & c_1^{\mathrm{T}}c_k \\ \vdots & & \vdots \\ c_k^{\mathrm{T}}c_1 & \cdots & c_k^{\mathrm{T}}c_k \end{pmatrix}$$

$$AA^{\mathrm{T}} = \begin{pmatrix} r_1 \\ \vdots \\ r_n \end{pmatrix}(r_1^{\mathrm{T}} \cdots r_n^{\mathrm{T}}) = \begin{pmatrix} r_1r_1^{\mathrm{T}} & \cdots & r_1r_n^{\mathrm{T}} \\ \vdots & & \vdots \\ r_nr_1^{\mathrm{T}} & \cdots & r_nr_n^{\mathrm{T}} \end{pmatrix}$$

If the matrix is non-square, the columns as the orthonormal set does not imply the rows will also be orthonormal.

e.g. $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{pmatrix}$ $A^TA = I$

$AA^T = \begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

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

Theorem 5.4.6

Let A be a square matrix of order n.
The following statements are equivalent:

1. A is an orthogonal matrix.
2. The rows of A form an orthonormal basis for Rn.
3. The columns of A form an orthonormal basis for Rn.

Theorem 5.4.7

Let S and T be two orthonormal bases for a vector space and let P be the transition matrix from S to T.

1. P is orthogonal.
2. P^{T} is the transition matrix from T to S. $= P^{-1}$

Transition matrix between orthonormal bases

P is an orthogonal matrix

Let $S = \{u_1, u_2, \dots, u_n\}$ and $T = \{v_1, v_2, \dots, v_n\}$ be two bases for $\mathbf{R}^n$

Transition matrix from S to T: $P = ([u_1]_T\ [u_2]_T \cdots [u_n]_T)$

S is an orthonormal basis for $\mathbf{R}^n$, T is the standard basis for $\mathbf{R}^n$ → $P = (u_1\ u_2 \cdots u_n)$

S is the standard basis for $\mathbf{R}^n$, T is an orthonormal basis for $\mathbf{R}^n$ → $P = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}$

S is an orthonormal basis for $\mathbf{R}^n$, T is an orthonormal basis for $\mathbf{R}^n$ → $P = \begin{pmatrix} u_1 \cdot v_1 & u_2 \cdot v_1 & \cdots & u_k \cdot v_1 \\ u_1 \cdot v_2 & u_2 \cdot v_2 & \cdots & u_k \cdot v_2 \\ \vdots & \vdots & & \vdots \\ u_1 \cdot v_k & u_2 \cdot v_k & \cdots & u_k \cdot v_k \end{pmatrix}$

orthonormal

$[u_i]_T \cdot [u_j]_T = u_i \cdot u_j = \begin{cases} 1 & \text{if } i=j \\ 0 & \text{if } i \neq j \end{cases}$

$P[w]_S = [w]_T$

$S = \{u_1, u_2, \dots, u_n\}$ and $T = \{v_1, v_2, \dots, v_n\}$ two orthonormal bases for $\mathbf{R}^n$

$A = (u_1\ u_2 \dots u_n)$ and $B = (v_1\ v_2 \dots v_n)$ orthogonal matrices

P the transition matrix from S to T

True or false: $PA = B$

Correct relation: $BP = A$ $\Leftrightarrow P = B^{-1}A$

orthogonal matrices

Hint: $(B \mid A) \to (I \mid P)$

Pre-multiply by B^{-1}

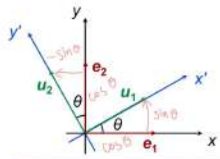
$(B^{-1}B \mid B^{-1}A)$

$\boxed{5.32}$ A: orthogonal matrix of order n
$u, v \in \mathbb{R}^n$
(a) $\|u\| = \|Au\|$
(b) $d(u,v) = d(Au, Av)$
(c) $\measuredangle$ b/w u & v = $\measuredangle$ b/w Au & Av

Let $E = \{ \boldsymbol{e}_1, \boldsymbol{e}_2 \}$ be the standard bases for $\mathbb{R}^2$ where
$\boldsymbol{e}_1 = (1, 0)$ is in the same direction as the x-axis,
$\boldsymbol{e}_2 = (0, 1)$ is in the same direction as the y-axis.

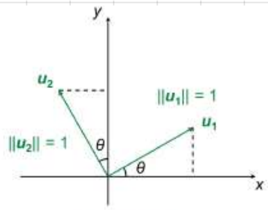
Consider a new $x'y'$-coordinate system obtained by rotating the original xy-coordinates anti-clockwise about the origin through an angle θ.

Let $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$ be the unit vectors such that
$\boldsymbol{u}_1$ is in the direction of the x'-axis,
$\boldsymbol{u}_2$ is in the direction of the y'-axis.



$S = \{ \boldsymbol{u}_1, \boldsymbol{u}_2 \}$ is an orthonormal basis for $\mathbb{R}^2$.

$\boldsymbol{u}_1 = (\cos(\theta), \sin(\theta))$
$= \cos(\theta)\, \boldsymbol{e}_1 + \sin(\theta)\, \boldsymbol{e}_2,$
$\boldsymbol{u}_2 = (-\sin(\theta), \cos(\theta))$
$= -\sin(\theta)\, \boldsymbol{e}_1 + \cos(\theta)\, \boldsymbol{e}_2.$



The transition matrix from S to E is

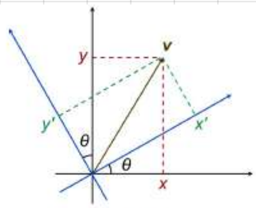
$$\boldsymbol{P} = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}.$$

Thus (by Theorem 5.4.7) the transition matrix from E to S is

$$\boldsymbol{P}^{\mathrm{T}} = \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix}.$$

Let $\boldsymbol{v} = (x, y) \in \mathbb{R}^2$ and let $(\boldsymbol{v})_S = (x', y')$.

In here, (x', y') is the coordinates of $\boldsymbol{v}$ using the new $x'y'$-coordinate system.



Since the transition matrix from E to S is $\boldsymbol{P}^{\mathrm{T}}$,

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = [\boldsymbol{v}]_S = \boldsymbol{P}^{\mathrm{T}}[\boldsymbol{v}]_E = \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}.$$

So $x' = x\cos(\theta) + y\sin(\theta),$
$y' = -x\sin(\theta) + y\cos(\theta).$

Let $S = \{ \boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3 \}$, where

$$\boldsymbol{u}_1 = \tfrac{1}{\sqrt{3}}(1, 1, 1),\ \boldsymbol{u}_2 = \tfrac{1}{\sqrt{2}}(1, 0, -1),\ \boldsymbol{u}_3 = \tfrac{1}{\sqrt{6}}(1, -2, 1),$$

and $T = \{ \boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3 \}$, where

$$\boldsymbol{v}_1 = (0, 0, 1),\ \boldsymbol{v}_2 = \tfrac{1}{\sqrt{2}}(1, -1, 0),\ \boldsymbol{v}_3 = \tfrac{1}{\sqrt{2}}(1, 1, 0).$$

Both S and T are orthonormal based for $\mathbb{R}^3$.

$$\boldsymbol{u}_1 = (\boldsymbol{u}_1 \cdot \boldsymbol{v}_1)\boldsymbol{v}_1 + (\boldsymbol{u}_1 \cdot \boldsymbol{v}_2)\boldsymbol{v}_2 + (\boldsymbol{u}_1 \cdot \boldsymbol{v}_3)\boldsymbol{v}_3 = \tfrac{1}{\sqrt{3}}\boldsymbol{v}_1 + \tfrac{2}{\sqrt{6}}\boldsymbol{v}_3,$$
$$\boldsymbol{u}_2 = (\boldsymbol{u}_2 \cdot \boldsymbol{v}_1)\boldsymbol{v}_1 + (\boldsymbol{u}_2 \cdot \boldsymbol{v}_2)\boldsymbol{v}_2 + (\boldsymbol{u}_2 \cdot \boldsymbol{v}_3)\boldsymbol{v}_3 = \tfrac{-1}{\sqrt{2}}\boldsymbol{v}_1 + \tfrac{1}{2}\boldsymbol{v}_2 + \tfrac{1}{2}\boldsymbol{v}_3,$$
$$\boldsymbol{u}_3 = (\boldsymbol{u}_3 \cdot \boldsymbol{v}_1)\boldsymbol{v}_1 + (\boldsymbol{u}_3 \cdot \boldsymbol{v}_2)\boldsymbol{v}_2 + (\boldsymbol{u}_3 \cdot \boldsymbol{v}_3)\boldsymbol{v}_3 = \tfrac{1}{\sqrt{6}}\boldsymbol{v}_1 + \tfrac{3}{\sqrt{12}}\boldsymbol{v}_2 + \tfrac{-1}{\sqrt{12}}\boldsymbol{v}_3.$$

The transition matrix form S to T is

$$\boldsymbol{P} = \begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ 0 & \frac{1}{2} & \frac{3}{\sqrt{12}} \\ \frac{2}{\sqrt{6}} & \frac{1}{2} & \frac{-1}{\sqrt{12}} \end{bmatrix}.$$

The transition matrix form T to S is

$$\boldsymbol{P}^{\mathrm{T}} = \begin{bmatrix} \frac{1}{\sqrt{3}} & 0 & \frac{2}{\sqrt{6}} \\ \frac{-1}{\sqrt{2}} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{\sqrt{6}} & \frac{3}{\sqrt{12}} & \frac{-1}{\sqrt{12}} \end{bmatrix}.$$

$S \xrightarrow{P_{k \times k}} E$: $[W]_E = P[W]_S$ $\quad (E \mid S) \rightarrow \left(\begin{array}{c|c} I_k & P \\ 0 \cdots 0 & 0 \cdots 0 \end{array}\right)$

$E \longrightarrow S$: $[W]_S = P^{-1}[W]_E$ $\quad P = ([u_1]_E \cdots [u_k]_E)$
$= P^{T}[W]_E$

Eigenvalue & Eigenvector

☆ must use non-zero det in proof

Definition 6.1.3 Let $\boldsymbol{A}$ be a square matrix of order n. A nonzero column vector $\boldsymbol{u}$ in $\mathbb{R}^n$ is called an *eigenvector* of $\boldsymbol{A}$ if

$$\boldsymbol{A}\boldsymbol{u} = \lambda \boldsymbol{u}$$

for some scalar λ. The scalar λ is called an *eigenvalue* of $\boldsymbol{A}$ and $\boldsymbol{u}$ is said to be an eigenvector of $\boldsymbol{A}$ *associated* with the eigenvalue λ.

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

Characteristic polynomial:

$$\det(\lambda \boldsymbol{I} - \boldsymbol{A}) = \begin{vmatrix} \lambda & 0 & -1 & 0 \\ 0 & \lambda & 0 & -1 \\ -1 & 0 & \lambda & 0 \\ 0 & -1 & 0 & \lambda \end{vmatrix}$$
$$= \lambda \times \begin{vmatrix} \lambda & 0 & -1 \\ 0 & \lambda & 0 \\ -1 & 0 & \lambda \end{vmatrix} - 1 \times \begin{vmatrix} 0 & \lambda & -1 \\ -1 & 0 & 0 \\ 0 & -1 & \lambda \end{vmatrix}$$
$$= \lambda(\lambda^3 - \lambda) - (1)(\lambda^2 - 1)$$
$$= \lambda^2(\lambda^2 - 1) - (1)(\lambda^2 - 1)$$
$$= (\lambda^2 - 1)(\lambda^2 - 1)$$
$$= (\lambda - 1)(\lambda + 1)(\lambda - 1)(\lambda + 1) = (\lambda - 1)^2(\lambda + 1)^2$$

Multiplicities of the eigenvalues

So the eigenvalues of $\boldsymbol{A}$ are 1 (repeated) and -1 (repeated).

- eigenvalues can be non-integers, and even complex numbers
- Gaussian elimination does not preserve eigenvalues:
if λ is an eigenvalue of A, λ may not be an eigenvalue of rref(A)
- n x n matrix has n eigenvalues, counting multiplicities (repeated roots)

$Au \in \operatorname{span}\{u\}$
$A^2u = A\lambda u = \lambda Au = \lambda(\lambda u) = \lambda^2 u$

A behaves like (λI) but not necessarily $= \lambda I$.

Example

$$\mathbf{B} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \quad \mathbf{x} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \quad \mathbf{y} = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \quad \mathbf{z} = \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}$$

$$\mathbf{Bx} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 3 \\ 3 \end{pmatrix} = 3\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = 3\mathbf{x}$$

$\mathbf{x}$ is an eigenvector associated with eigenvalue 3

$$\mathbf{B}(2\mathbf{x}) = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}\begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix} = \begin{pmatrix} 6 \\ 6 \\ 6 \end{pmatrix} = 3\begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix} = 3(2\mathbf{x})$$

$2\mathbf{x}$ is an eigenvector associated with eigenvalue 3

$$\mathbf{B}(k\mathbf{x}) = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}\begin{pmatrix} k \\ k \\ k \end{pmatrix} = 3\begin{pmatrix} k \\ k \\ k \end{pmatrix} = 3(k\mathbf{x})$$

$k\mathbf{x}$ is an eigenvector associated with eigenvalue 3

- If x is an eigenvector of B, then 2x is also an eigenvector of B with the same eigenvalue.
- if x is an eigenvector of 2B with twice the eigenvalue

$Bx = \lambda x$

$B(2x) = \lambda(2x) \qquad (2B)x = 2\lambda x$

$E_\lambda = \{ v \in \mathbb{R}^n \mid Av = \lambda v \} = \text{Null}(\lambda I - A)$

$(\lambda I - A)v = 0$

Remark 6.1.5

Remark 6.1.5 Let $\boldsymbol{A}$ be a square matrix of order n. Then

λ is an eigenvalue of $\boldsymbol{A}$

$\Leftrightarrow$ $\boldsymbol{Au} = \lambda\boldsymbol{u}$ for some nonzero column vector $\boldsymbol{u}$

$\Leftrightarrow$ $\lambda\boldsymbol{u} - \boldsymbol{Au} = \boldsymbol{0}$ for some nonzero column vector $\boldsymbol{u}$

$\Leftrightarrow$ $(\lambda\boldsymbol{I} - \boldsymbol{A})\boldsymbol{u} = \boldsymbol{0}$ for some nonzero column vector $\boldsymbol{u}$

singular $\Leftrightarrow$ the linear system $(\lambda\boldsymbol{I} - \boldsymbol{A})\boldsymbol{x} = \boldsymbol{0}$ has non-trivial solution

$\Leftrightarrow$ $\det(\lambda\boldsymbol{I} - \boldsymbol{A}) = 0$ (by Theorem 3.6.11). ← Solve this eqn onto find the eigenvalues of A

If expanded, $\det(\lambda\boldsymbol{I} - \boldsymbol{A})$ is a polynomial in λ of degree n. $\Leftrightarrow$ λ is a root of the characteristic eqn

coeff of $\lambda^n = 1$; constant $= \det(A)$

$$\det(\lambda\mathbf{I} - \mathbf{A}) = \begin{vmatrix} \lambda - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & \lambda - a_{22} & \cdots & -a_{2n} \\ \vdots & & \ddots & \vdots \\ -a_{n1} & -a_{n1} & \cdots & \lambda - a_{nn} \end{vmatrix}$$

eigenspace = { eigenvectors } ∪ {0}

non-zero

Characteristic Equation & Characteristic Polynomial

Definition 6.1.6 Let $\boldsymbol{A}$ be a square matrix of order n. The equation

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

is called the *characteristic equation* of $\boldsymbol{A}$ and the polynomial

$$\det(\lambda\boldsymbol{I} - \boldsymbol{A})$$

is called the *characteristic polynomial* of $\boldsymbol{A}$.

- det(xI - A) is the characteristic polynomial.
- det(A - xI) may differ by a negative sign.
- Nevertheless, as far as finding eigenvalue is concern, both det will give the same answer.

Theorem 6.1.8 (main theorem)***

Let A be an n x n matrix. The following statements are equivalent.

(1) A is invertible.

(2) Ax=0 has only the trivial solution.

(3) RREF(A) = I

(4) A can be expressed as a product of elementary matrices.

(5) det(A) =/= 0

(6) The rows of A form a basis for Rn.

(7) The columns of A form a basis for Rn.

} row(col) space of $A = \mathbb{R}^n$

(8) rank(A) = n, nullity(A) = 0

(9) 0 is not an eigenvalue of A

Examples

1. Let $\boldsymbol{A}$ be the 2×2 matrix in Example 6.1.1. The characteristic polynomial of $\boldsymbol{A}$ is

$$\begin{aligned} \det(\lambda\boldsymbol{I} - \boldsymbol{A}) &= \det\left(\lambda\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} 0.96 & 0.01 \\ 0.04 & 0.99 \end{pmatrix}\right) \\ &= \begin{vmatrix} \lambda - 0.96 & -0.01 \\ -0.04 & \lambda - 0.99 \end{vmatrix} \\ &= (\lambda - 0.96)(\lambda - 0.99) - (-0.01)(-0.04) \\ &= \lambda^2 - 1.95\lambda + 0.95 \\ &= (\lambda - 1)(\lambda - 0.95). \end{aligned}$$

Hence $\det(\lambda\boldsymbol{I} - \boldsymbol{A}) = 0$ if and only if $\lambda = 1$ or 0.95. The eigenvalues of $\boldsymbol{A}$ are 1 and 0.95.

2. Let $\boldsymbol{B}$ be the 3×3 matrix in Example 6.1.4.2. The characteristic polynomial of $\boldsymbol{B}$ is

$$\begin{aligned} \det(\lambda\boldsymbol{I} - \boldsymbol{B}) &= \begin{vmatrix} \lambda - 1 & -1 & -1 \\ -1 & \lambda - 1 & -1 \\ -1 & -1 & \lambda - 1 \end{vmatrix} \\ &= \lambda^3 - 3\lambda^2 = (\lambda - 3)(\lambda - 0)^2. \end{aligned}$$

Hence $\det(\lambda\boldsymbol{I} - \boldsymbol{B}) = 0$ if and only if $\lambda = 3$ or 0. The eigenvalues of $\boldsymbol{B}$ are 3 and 0.

Theorem 6.1.9

If A is a triangular matrix (either upper or lower), in particular, diagonal matrix the eigenvalues of A are the diagonal entries of A.

Example

$$\begin{pmatrix} -1 & 3.5 & 14 \\ 0 & 5 & -26 \\ 0 & 0 & 2 \end{pmatrix}$$

The eigenvalues are −1, 5 & 2.

$$\begin{pmatrix} -2 & 0 & 0 \\ 99 & 0 & 0 \\ 10 & -4.5 & 10 \end{pmatrix}$$

The eigenvalues are −2, 0 & 10.

Finding Eigenvalues

- If an eigenvector **u** is given, multiply it by the matrix: $\boldsymbol{Au} = \lambda\boldsymbol{u}$
- Solve characteristic equation $\det(\lambda\boldsymbol{I} - \boldsymbol{A}) = 0$
- If the matrix is triangular, take the diagonal entries
- If λ is an eigenvalue of $\boldsymbol{A}$, then
  - λ is an eigenvalue of $\boldsymbol{A}^T$
  - λ^n is an eigenvalue of $\boldsymbol{A}^n$
  - λ^{-1} is an eigenvalue of $\boldsymbol{A}^{-1}$ (when $\boldsymbol{A}$ is invertible)

Exercise 6 Q3

$k\lambda$ is an eigenvalue of kA

If you are given the diagonalization of $A = PDP^{-1}$, then eigenvalues are given by the diagonals of D

Eigenspace

$\boldsymbol{A}$: n x n matrix and λ is (one of the) eigenvalue of $\boldsymbol{A}$

Solution space of $(\lambda\boldsymbol{I} - \boldsymbol{A})\boldsymbol{x} = 0$ ⟷ eigenspace E_λ of $\boldsymbol{A}$ associated with eigenvalue λ

trivial solution

Non-zero nontrivial solutions ⟷ eigenvector of $\boldsymbol{A}$ associated with eigenvalue λ

- the sum of the dimensions of all the eigen space of a n*n matrix does not necessarily equal to n.
- it is equal only for matrices that diagonalizable

Diagonalizable Matrix (Square matrix)

Def. A square matrix A is called diagonalizable if there exists an **invertible matrix P** such that **$P^{-1}AP$ is a diagonal matrix.**

- matrix P diagonalizes A

$$P^{-1}AP = \begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix} \quad \text{diagonal}$$

diagonalizable

All diagonal matrices are Diagonalizable including zero matrix (P can be any inv matrix)

- Show that a matrix is non-diagonalisable: proof by contradiction

The matrix $\boldsymbol{M} = \begin{pmatrix} 2 & 0 \\ 1 & 2 \end{pmatrix}$ in Example 6.1.12.4 is not diagonalizable, i.e. there is no invertible matrix that can diagonalize $\boldsymbol{M}$.

Proof Assume the contrary, i.e. there exists an invertible matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ such that

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1}\begin{pmatrix} 2 & 0 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix}.$$

Then $\begin{pmatrix} 2 & 0 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix}$ which implies

$= \begin{pmatrix} 2a & 2b \\ a+2c & b+2d \end{pmatrix}$

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

$$\begin{cases} 2a = \lambda a \\ 2b = \mu b \\ a + 2c = \lambda c \\ b + 2d = \mu d. \end{cases}$$

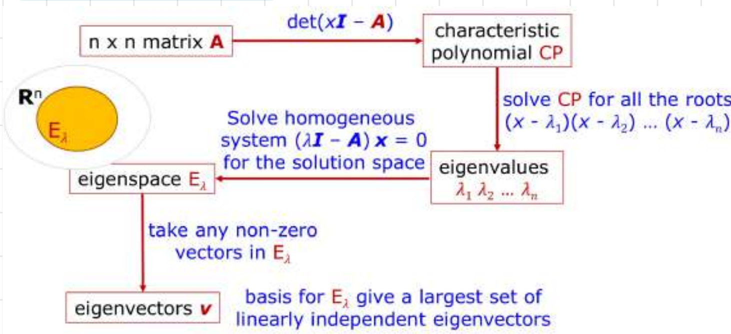
Solving the equations, we obtain $a = 0$ and $b = 0$. However, $\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ c & d \end{pmatrix}$ is not invertible, a contradiction.

$\det \neq 0$

Finding Eigenvectors

- Solve the homogeneous system $(\lambda\mathbf{I} - \mathbf{A})\mathbf{x} = \mathbf{0}$
- Look for vectors **u** such that **Au** = k**u**
- If you are given the diagonalization of $\mathbf{A} = \mathbf{PDP}^{-1}$, then eigenvectors are given by the $\boxed{\text{columns}}$ of **P**.
- If you are given the eigenspace E_λ, any $\boxed{\text{nonzero}}$ vector in it is an eigenvector.
- If **u** is an eigenvector w.r.t. λ, then $\boxed{k\mathbf{u}}$ is also an eigenvector w.r.t. λ, for any $k \neq 0$.
- If **u**, **v** are eigenvectors w.r.t. λ, then $\boxed{s\mathbf{u} + t\mathbf{v}}$ is also an eigenvector w.r.t. λ, for any s, t not both 0.

Finding Eigenspace



Example

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

Solve $(\lambda\boldsymbol{I} - \boldsymbol{A})\,\boldsymbol{x} = \boldsymbol{0}$ for eigenvalues $\lambda = 1$ and -1

$$\begin{pmatrix} \lambda & 0 & -1 & 0 \\ 0 & \lambda & 0 & -1 \\ -1 & 0 & \lambda & 0 \\ 0 & -1 & 0 & \lambda \end{pmatrix}\begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

$\lambda = 1$

$$\left(\begin{array}{cccc|c} 1 & 0 & -1 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 \\ -1 & 0 & 1 & 0 & 0 \\ 0 & -1 & 0 & 1 & 0 \end{array}\right) \xrightarrow{G.E.} \left(\begin{array}{cccc|c} 1 & 0 & -1 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right) \Rightarrow \begin{array}{c} w = t,\ z = s, \\ y = t,\ x = s \end{array}$$

Gen. soln: $\begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} = \begin{pmatrix} s \\ t \\ s \\ t \end{pmatrix} = s\begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} + t\begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}$

Eigenspace for $\lambda = 1$:

$E_1 = \text{span}\left\{\begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}\right\}$

Any non-trivial linear combination is an eigenvector associated to $\lambda = 1$

Eigenspace for $\lambda = -1$: $E_{-1} = \text{span}\left\{\begin{pmatrix} -1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ 0 \\ 1 \end{pmatrix}\right\}$

$\boldsymbol{AB} = \boldsymbol{A}(\boldsymbol{b}_1\ \ \boldsymbol{b}_2\ \cdots\ \boldsymbol{b}_n) = (\boldsymbol{Ab}_1\ \ \boldsymbol{Ab}_2\ \cdots\ \boldsymbol{Ab}_n)$

$$\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}\begin{array}{c}\begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 3 & 5 \end{pmatrix} \\ \boldsymbol{b}_1\ \boldsymbol{b}_2\ \boldsymbol{b}_3\end{array} = \begin{array}{c}\begin{pmatrix} 2 & 3 & 4 \\ 2 & 5 & 8 \\ 2 & 4 & 6 \end{pmatrix} \\ \boldsymbol{Ab}_1\,\boldsymbol{Ab}_2\,\boldsymbol{Ab}_3\end{array}$$

$\boldsymbol{BD} = (\boldsymbol{b}_1\ \ \boldsymbol{b}_2\ \cdots\ \boldsymbol{b}_n)\,\boldsymbol{D} = (d_1\boldsymbol{b}_1\ \ d_2\boldsymbol{b}_2\ \cdots\ d_n\boldsymbol{b}_n)$

diagonal matrix with diagonal entries d_i

$$\begin{array}{c}\begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 3 & 5 \end{pmatrix} \\ \boldsymbol{b}_1\ \boldsymbol{b}_2\ \boldsymbol{b}_3\end{array}\begin{pmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 4 \end{pmatrix} = \begin{array}{c}\begin{pmatrix} 2 & 3 & 4 \\ 2 & 6 & 12 \\ 2 & 9 & 20 \end{pmatrix} \\ 2\boldsymbol{b}_1\ 3\boldsymbol{b}_2\ 4\boldsymbol{b}_3\end{array}$$

Theorem 6.2.3

Let A be a square matrix of order **n**. Then

A is diagonalizable iff A has **n linearly independent eigenvectors** (basis)

(May be associated to the same eigenvalues)

Algorithm 6.24 (Diagonalization)

Algorithm 6.2.4 Given a square matrix $\boldsymbol{A}$ of order n, we want to determine whether $\boldsymbol{A}$ is diagonalizable. Also, if $\boldsymbol{A}$ is diagonalizable, find an invertible matrix $\boldsymbol{P}$ such that $\boldsymbol{P}^{-1}\boldsymbol{AP}$ is a diagonal matrix.

Step 1: Find all distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_k$. (By Remark 6.1.5, eigenvalues can be obtained by solving the characteristic equation $\det(\lambda\boldsymbol{I} - \boldsymbol{A}) = 0$.)

Step 2: For each eigenvalue λ_i, find a basis S_{λ_i} for the eigenspace E_{λ_i}. $(\lambda_i I - A)x = 0$

$\dim = \text{non-pivot}$

Step 3: Let $S = S_{\lambda_1} \cup S_{\lambda_2} \cup \cdots \cup S_{\lambda_k}$. All bases are linearly indep

(a) If $|S| < n$, then $\boldsymbol{A}$ is not diagonalizable.

(b) If $|S| = n$, say $S = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_n\}$, then $\boldsymbol{P} = (\boldsymbol{u}_1 \ \ \boldsymbol{u}_2 \ \ \cdots \ \ \boldsymbol{u}_n)$ is an invertible matrix that diagonalizes $\boldsymbol{A}$.

Remark 6.2.5

(1) In Step 1, the matrix A may have eigenvalues that are complex numbers, i.e. the characteristic equation det(xI- A) = 0 has complex solutions. We can still use the algorithm to diagonalize the matrix. (Not in scope)

(2)

Suppose the characteristic polynomial of the matrix $\boldsymbol{A}$ can be factorized as

$$\det(\lambda\boldsymbol{I} - \boldsymbol{A}) = (\lambda - \lambda_1)^{r_1}(\lambda - \lambda_2)^{r_2} \cdots (\lambda - \lambda_k)^{r_k}$$

where $\lambda_1, \lambda_2, \ldots, \lambda_k$ are distinct eigenvalues of $\boldsymbol{A}$. Then for each eigenvalue λ_i,

$$1 \leq \dim(E_{\lambda_i}) \leq r_i.$$ Multiplicity of the eigenvalue

Furthermore, $\boldsymbol{A}$ is diagonalizable if and only if in Step 2, for each eigenvalue λ_i, $\dim(E_{\lambda_i}) = r_i$, i.e. $|S_{\lambda_i}| = r_i$. Sum of multiplicities = n

(3) In Step 3, the set S is always linearly independent.

Example

$$\boldsymbol{B} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

Step 1: By solving characteristic polynomial, the eigenvalues are 3 and 0.

Step 2: For $\lambda = 3$, solve $(3\boldsymbol{I} - \boldsymbol{B})\boldsymbol{x} = 0$

For $\lambda = 0$, solve $(0\boldsymbol{I} - \boldsymbol{B})\boldsymbol{x} = 0$

$S_3 = \left\{\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}\right\}$ a basis for E_3 $\quad S_0 = \left\{\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}\right\}$ a basis for E_0

$\dim(E_3) = 1$
$\Rightarrow$ multiplicity of eigenval $3 = 1$

$\dim(E_0) = 2$
$\Rightarrow$ multiplicity of eigenval $0 = 2$

Step 3: $|S| = |S_3| + |S_0| = 1 + 2 =$ order of $\boldsymbol{B}$

So $\boldsymbol{B}$ is diagonalizable

Step 3:

$$\boldsymbol{P} = \begin{pmatrix} 1 & -1 & -1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \quad \text{Then} \quad \boldsymbol{P}^{-1}\boldsymbol{BP} = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$\boldsymbol{u}_1 \quad \boldsymbol{u}_2 \quad \boldsymbol{u}_3$

you do not need to multiply this out!!!

∴ diagonal matrix has one 3 & two 0

$\boldsymbol{P}$ is not unique $\quad \boldsymbol{P} = \begin{pmatrix} 2 & -7 & 1 \\ 2 & 7 & 0 \\ 2 & 0 & -1 \end{pmatrix}$

$2\boldsymbol{u}_1 \quad 7\boldsymbol{u}_2 \quad -\boldsymbol{u}_3$

$$\boldsymbol{Q} = \begin{pmatrix} -1 & -1 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} \quad \text{Then} \quad \boldsymbol{Q}^{-1}\boldsymbol{BQ} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

Theorem 6.2.7

P is a sq matrix

Let A be a square matrix of order n.

If A has **n distinct eigenvalues**, then A is diagonalizable.

Remark 6.2.9

The converse of Theorem 6.2.7 is not true, i.e. a diagonalizable matrix of order n may not need to have n distinct eigenvalue

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

$\boldsymbol{A}$ has 4 distinct eigenvalues 1, 2, 3, 4.

So $\boldsymbol{A}$ is diagonalizable.

$$\boldsymbol{B} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}$$ diagonal matrices are diagonalizable

$\boldsymbol{B}$ has only 2 distinct eigenvalues 1, 2.

And $\boldsymbol{B}$ is also diagonalizable.

Checking Diagonalizable Matrix

- When **A** is a diagonal matrix $\quad P = I \ ; \ A = IDI^{-1}$
- When **A** is a symmetric matrix
- When **A** has n distinct eigenvalues $\rightarrow$ diagonalizable

Sufficient conditions

- When **A** has n linearly independent eigenvectors
- When dim E_λ = multiplicity of λ for every eigenvalue λ of **A**

Equivalent conditions

$\sum_{i=1}^{k} r_i = n$

To show that a matrix is not diagonalizable:

Find one eigenvalue such that dim E_{λ_i} < multiplicity of λ_i

or $\det(\lambda I - A)$ does not split into linear factors

Char. Poly $= (x - \lambda_1)^{r_1}(x - \lambda_2)^{r_2} \cdots (x - \lambda_k)^{r_k}$

· n linearly indep eigenvectors $\Leftrightarrow \exists$ basis $\{u_1, \ldots, u_n\} \in \mathbb{R}^n$ of eigenvectors

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

Step 1: The eigenvalues are 1 and 2. (We can get the eigenvalues of $\boldsymbol{A}$ without solving the equation $\det(\lambda\boldsymbol{I} - \boldsymbol{A}) = 0$. Why?)

Step 2: For $\lambda = 1$, the linear system $(\lambda\boldsymbol{I} - \boldsymbol{A})\boldsymbol{x} = \boldsymbol{0}$ is

$$\begin{pmatrix} 0 & 0 & 0 \\ -1 & -1 & 0 \\ 3 & -5 & -1 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Solving the system, we have $\begin{pmatrix} x \\ y \\ z \end{pmatrix} = t\begin{pmatrix} 1 \\ -1 \\ 8 \end{pmatrix}$ where t is an arbitrary parameter.

So $\left\{\begin{pmatrix} 1 \\ -1 \\ 8 \end{pmatrix}\right\}$ is a basis for E_1.

For $\lambda = 2$, the linear system $(\lambda\boldsymbol{I} - \boldsymbol{A})\boldsymbol{x} = \boldsymbol{0}$ is

$$\begin{pmatrix} 1 & 0 & 0 \\ -1 & 0 & 0 \\ 3 & -5 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Solving the system, we have $\begin{pmatrix} x \\ y \\ z \end{pmatrix} = t\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ where t is an arbitrary parameter.

So $\left\{\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}\right\}$ is a basis for E_2.

Step 3: Since we only have two linearly independent eigenvectors, $\boldsymbol{A}$ is not diagonalizable.

Finding powers of a matrix

Example 6.2.11.1 invertible

$$A = \begin{pmatrix} -4 & 0 & -6 \\ 2 & 1 & 2 \\ 3 & 0 & 5 \end{pmatrix}$$

Use Algorithm 6.2.4 to find the eigenvalues and eigenvectors

We have

$$P = \begin{pmatrix} -2 & 0 & -1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \qquad P^{-1}AP = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

obtain this diagonal matrix from eigenvalues, not matrix multiplication!

$$A^m = P\begin{pmatrix} (-1)^m & 0 & 0 \\ 0 & 1^m & 0 \\ 0 & 0 & 2^m \end{pmatrix}P^{-1} \qquad A^{-1} = P\begin{pmatrix} (-1)^{-1} & 0 & 0 \\ 0 & 1^{-1} & 0 \\ 0 & 0 & 2^{-1} \end{pmatrix}P^{-1}$$

Discussion 6.2.10 As we have seen in Example 6.1.1, one of the applications of diagonalization is to compute powers of square matrices: Let A be a square matrix of order n and P an invertible matrix such that

$$P^{-1}AP = \begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix}.$$

premultiply by P, post-multiply by P^{-1}

$(P^{-1}AP)^m = P^{-1}A^mP$

1. For all positive integer m, $A^m = P\begin{pmatrix} \lambda_1^m & & & 0 \\ & \lambda_2^m & & \\ & & \ddots & \\ 0 & & & \lambda_n^m \end{pmatrix}P^{-1}$.

for $m<0$, true only if A inv

2. Suppose A is invertible. By Theorem 6.1.8, we know that $\lambda_i \neq 0$ for all i. Then

$$A^{-1} = P\begin{pmatrix} \lambda_1^{-1} & & & 0 \\ & \lambda_2^{-1} & & \\ & & \ddots & \\ 0 & & & \lambda_n^{-1} \end{pmatrix}P^{-1}$$

and hence for all positive integer m, $A^{-m} = P\begin{pmatrix} \lambda_1^{-m} & & & 0 \\ & \lambda_2^{-m} & & \\ & & \ddots & \\ 0 & & & \lambda_n^{-m} \end{pmatrix}P^{-1}$.

Solving Linear Recurrence Relation

$a_0 = 0,\ a_1 = 1,\ a_n = a_{n-1} + a_{n-2}$ for $n \geq 2$

Solving linear recurrence relation

$a_0 = u \qquad a_1 = v \qquad a_n = pa_{n-1} + qa_{n-2}$ for $n \geq 2$

Form the recurrence matrix A $\qquad A = \begin{pmatrix} 0 & 1 \\ q & p \end{pmatrix} \qquad A = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$

Find the eigenvalues of A $\qquad \lambda_1 = \frac{1+\sqrt{5}}{2} \qquad \lambda_2 = \frac{1-\sqrt{5}}{2}$

If A is diagonalizable, find the matrix P that diagonalizes A $\qquad P = \begin{pmatrix} 1 & 1 \\ \frac{1+\sqrt{5}}{2} & \frac{1-\sqrt{5}}{2} \end{pmatrix}$

Set up $x_n = A^n x_0$ and diagonalize A^n $\qquad \begin{pmatrix} a_n \\ a_{n+1} \end{pmatrix} = P\begin{pmatrix} \lambda_1^n & 0 \\ 0 & \lambda_2^n \end{pmatrix}P^{-1}\begin{pmatrix} a_0 \\ a_1 \end{pmatrix}$

Multiply out the RHS and equate the first component $\qquad a_n = s(\lambda_1)^n + t(\lambda_2)^n$

↳ free params

Example

Example 6.1.1

$$\underbrace{\begin{pmatrix} a_n \\ b_n \end{pmatrix}}_{x_n} = \underbrace{\begin{pmatrix} 0.96 & 0.01 \\ 0.04 & 0.99 \end{pmatrix}}_{A}\underbrace{\begin{pmatrix} a_{n-1} \\ b_{n-1} \end{pmatrix}}_{x_{n-1}} \Rightarrow \begin{pmatrix} a_n \\ b_n \end{pmatrix} = \underbrace{\begin{pmatrix} 0.96 & 0.01 \\ 0.04 & 0.99 \end{pmatrix}^n}_{A^n}\underbrace{\begin{pmatrix} a_0 \\ b_0 \end{pmatrix}}_{x_0}$$

$x_n = Ax_{n-1} = A^2x_{n-2} = A^3x_{n-3} = \cdots = A^n x_0$ current population

long term effect ⟶ a_n and b_n for large n

⟶ x_n for large n

⟶ A^n for large n

$$A^{(\text{big } n)} = \begin{pmatrix} 1 & 1 \\ 4 & -1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 0.95^{(\text{big } n)} \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 4 & -1 \end{pmatrix}^{-1} \approx \begin{pmatrix} 0.2 & 0.2 \\ 0.8 & 0.8 \end{pmatrix}$$

$\to 0$

$$\begin{pmatrix} a_{(\text{big } n)} \\ b_{(\text{big } n)} \end{pmatrix} \approx \begin{pmatrix} 0.2 & 0.2 \\ 0.8 & 0.8 \end{pmatrix}\begin{pmatrix} a_0 \\ b_0 \end{pmatrix} = \begin{pmatrix} 0.2(a_0+b_0) \\ 0.8(a_0+b_0) \end{pmatrix}$$

2. Let $(a_0, a_1, a_2, \ldots)$ be a sequence of numbers such that $a_0 = 0$, $a_1 = 1$ and $a_n = a_{n-1} + a_{n-2}$ for $n \geq 2$. These numbers are known as the *Fibonacci numbers*. In the following, we demonstrate a method to find the value of a_n by using the eigenvalue technique.

In order to formulate the problem in terms of a matrix equation, for $n = 1, 2, 3, \ldots$, we write

$$\begin{cases} a_n = a_n & n=1 \\ a_{n+1} = a_{n-1} + a_n & n\geq 2 \end{cases} \Leftrightarrow \begin{pmatrix} a_n \\ a_{n+1} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} a_{n-1} \\ a_n \end{pmatrix}.$$

Let $x_n = \begin{pmatrix} a_n \\ a_{n+1} \end{pmatrix}$ and $A = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$. Then

$$x_n = Ax_{n-1} = A^2x_{n-2} = \cdots = A^n x_0.$$

Following Algorithm 6.2.4, we find an invertible matrix $P = \begin{pmatrix} 1 & 1 \\ \frac{1+\sqrt{5}}{2} & \frac{1-\sqrt{5}}{2} \end{pmatrix}$ such that $P^{-1}AP = \begin{pmatrix} \frac{1+\sqrt{5}}{2} & 0 \\ 0 & \frac{1-\sqrt{5}}{2} \end{pmatrix}$. Then

$$\begin{pmatrix} a_n \\ a_{n+1} \end{pmatrix} = x_n = A^n x_0$$
$$= \begin{pmatrix} 1 & 1 \\ \frac{1+\sqrt{5}}{2} & \frac{1-\sqrt{5}}{2} \end{pmatrix}\begin{pmatrix} \left(\frac{1+\sqrt{5}}{2}\right)^n & 0 \\ 0 & \left(\frac{1-\sqrt{5}}{2}\right)^n \end{pmatrix}\begin{pmatrix} 1 & 1 \\ \frac{1+\sqrt{5}}{2} & \frac{1-\sqrt{5}}{2} \end{pmatrix}^{-1}\begin{pmatrix} 0 \\ 1 \end{pmatrix}$$
$$= \begin{pmatrix} \frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^n - \frac{1}{\sqrt{5}}\left(\frac{1-\sqrt{5}}{2}\right)^n \\ \frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^{n+1} - \frac{1}{\sqrt{5}}\left(\frac{1-\sqrt{5}}{2}\right)^{n+1} \end{pmatrix}.$$

Thus $a_n = \frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^n - \frac{1}{\sqrt{5}}\left(\frac{1-\sqrt{5}}{2}\right)^n$.

Show that if A is not a scalar matrix that only 1 eigenvalue, then A is not diagonalizable.

Suppose A has only 1 eigenvalue λ. If A is diagonalizable, then

$A = P\begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}P^{-1} = P(\lambda I)P^{-1} = \lambda(PP^{-1}) = \lambda I$

$\Rightarrow$ A is a scalar matrix.

A diagonalizable $\Rightarrow$ A scalar

$\Leftrightarrow$ A not scalar $\Rightarrow$ A not diagonalizable

if wan show counterEg for non-diagonalizable, eg. $\begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}$

find a non-scalar matrix w 1 eigenval $\Rightarrow$ use Δ matrix

Orthogonal Diagonalizable Matrix

A square matrix A is called orthogonally diagonalizable if there exists an orthogonal matrix P s.t. P^TAP is a diagonal matrix. (diagonal entries = λ)

- matrix P orthogonally diagonalizes A.

general diagonalizable matrix:

$P^{-1}AP$ is a diagonal matrix

↓ invertible

Thm: $\lambda_1 \neq \lambda_2$ distinct eigenvalues of orthogonally diagonalizable matrix A, then $E_{\lambda_1} \perp E_{\lambda_2}$.

i.e. for any $v_1 \in E_{\lambda_1}$, $v_2 \in E_{\lambda_2}$, $v_1 \cdot v_2 = 0$

↓

E_{λ_1} & E_{λ_2} are linearly indep for diagonalizable matrix A.

Theorem 6.34

A square matrix is orthogonally diagonalizable iff it is symmetric.

Algorithm 6.3.5. (Orthogonally Diagonalizable Matrix)

orthogonally diagonalizable matrices $\subset$ diagonalizable matrices

Algorithm 6.3.5 Given a symmetric matrix $\boldsymbol{A}$ of order n, we want to find an orthogonal matrix $\boldsymbol{P}$ such that $\boldsymbol{P}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{P}$ is a diagonal matrix.

Step 1: Find all distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_k$.

Step 2: For each eigenvalue λ_i,

Step 2a: find a basis S_{λ_i} for the eigenspace E_{λ_i} and then $E_{\lambda_i} = \text{null}(\lambda I - A)$

Step 2b: use the Gram-Schmidt Process (Theorem 5.2.19) to transform S_{λ_i} to an orthonormal basis T_{λ_i}.

Step 3: Let $T = T_{\lambda_1} \cup T_{\lambda_2} \cup \cdots \cup T_{\lambda_k}$, say $T = \{\boldsymbol{v_1}, \boldsymbol{v_2}, \ldots, \boldsymbol{v_n}\}$.

Then $\boldsymbol{P} = \begin{pmatrix} \boldsymbol{v_1} & \boldsymbol{v_2} & \cdots & \boldsymbol{v_n} \end{pmatrix}$ is an orthogonal matrix that diagonalizes $\boldsymbol{A}$.

Remark 6.3.6

Remark 6.3.6 The following are some remarks on Algorithm 6.3.5:

1. In Step 1, the eigenvalues of a symmetric matrix are always real numbers.
2. Since $\boldsymbol{A}$ is diagonalizable, by Remark 6.2.5.2, we have the following result:

   Suppose the characteristic polynomial of the matrix $\boldsymbol{A}$ can be factorized as

   $$\det(\lambda \boldsymbol{I} - \boldsymbol{A}) = (\lambda - \lambda_1)^{r_1}(\lambda - \lambda_2)^{r_2} \cdots (\lambda - \lambda_k)^{r_k}$$

   where $\lambda_1, \lambda_2, \ldots, \lambda_k$ are distinct eigenvalues of $\boldsymbol{A}$. Then in Step 2, for each eigenvalue λ_i, $\dim(E_{\lambda_i}) = r_i$, i.e. $|S_{\lambda_i}| = |T_{\lambda_i}| = r_i$.
3. In Step 3, the set T is always orthonormal. (See Question 6.26.)
4. Since T is always orthonormal, by Theorem 5.4.6, the square matrix $\boldsymbol{P}$ in Step 3 is always orthogonal.

$r_1 + r_2 + \cdots + r_k$ = degree of polynomial = order of A

$\dim(E_{\lambda_1}) + \dim(E_{\lambda_2}) + \cdots + \dim(E_{\lambda_k})$ = no. of linearly indep eigenvectors

Example

$$\mathbf{B} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

Step 1: The eigenvalues are 3 and 0.

Step 2a: Bases for E_3 and E_0: $\left\{\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}\right\}$ and $\left\{\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}\right\}$

Step 2b: Orthonormal bases: $\left\{\begin{pmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{pmatrix}\right\}$ and $\left\{\begin{pmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{pmatrix}, \begin{pmatrix} -\frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{6}} \\ \frac{2}{\sqrt{6}} \end{pmatrix}\right\}$

just normalize

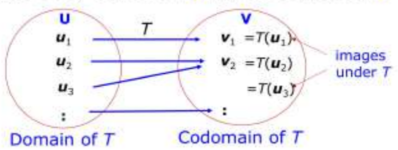
Gram-Schmidt

Step 3: $\mathbf{P} = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & \frac{2}{\sqrt{6}} \end{pmatrix}$ and $\mathbf{P}^T\mathbf{B}\mathbf{P} = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$

Mapping

A mapping from **U** to **V** $T: U \to V$

assigns every element of **U** with an element of **V**



Linear Transformation

Definition 7.1.1 A *linear transformation* is a mapping $T : \mathbb{R}^n \to \mathbb{R}^m$ of the form

$$T\left(\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}\right) = \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{pmatrix} \quad \text{for} \quad \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$$

where $a_{11}, a_{12}, \ldots, a_{mn}$ are real numbers. In particular, if $n = m$, T is also called a *linear operator* on $\mathbb{R}^n$. We can rewrite the formula of T as

$T: \mathbb{R}^n \to \mathbb{R}^n$

Domain of T = Codomain of T

standard matrix for T is a square matrix.

$$T\left(\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}\right) = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

linear transformatn

$T: \mathbb{R}^n \to \mathbb{R}^m$

codomain $\mathbb{R}^m$

domain $\mathbb{R}^n$

$T(u) = Au \quad \forall u \in \mathbb{R}^n$

input standard matrix

The matrix $(a_{ij})_{m \times n}$ above is called the *standard matrix* for T.

Formal definition of Linear Transformation

A linear transformation $T : \mathbf{R}^n \to \mathbf{R}^m$

is a mapping from $\mathbf{R}^n$ to $\mathbf{R}^m$

that satisfies the following condition:

For all vectors $\boldsymbol{u}, \boldsymbol{v}$ in $\mathbf{R}^n$ and scalars a, b

$T(a\boldsymbol{u} + b\boldsymbol{v}) = aT(\boldsymbol{u}) + bT(\boldsymbol{v})$ To check for Linear Tfm

Linearity conditions of T

1. $T(\boldsymbol{u} + \boldsymbol{v}) = T(\boldsymbol{u}) + T(\boldsymbol{v})$ T preserves addition
2. $T(c\boldsymbol{u}) = cT(\boldsymbol{u})$ T preserves scalar multiplication

Identity Transformation

Let $I : \mathbb{R}^n \to \mathbb{R}^n$ be the *identity transformation* defined by

$$I(\boldsymbol{u}) = \boldsymbol{u} \quad \text{for } \boldsymbol{u} \in \mathbb{R}^n.$$

It is a linear operator on $\mathbb{R}^n$ and the standard matrix is the identity matrix $\boldsymbol{I_n}$.

Zero Transformation

Let $O : \mathbb{R}^n \to \mathbb{R}^m$ be the *zero transformation* defined by

$$O(\boldsymbol{u}) = \boldsymbol{0} \quad \text{for } \boldsymbol{u} \in \mathbb{R}^n.$$

It is a linear transformation and the standard matrix is the zero matrix $\boldsymbol{0}_{m\times n}$.

O is a linear operator if domain = codomain

Theorem 7.1.4

Theorem 7.1.4 Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation.

1. $T(\boldsymbol{0}) = \boldsymbol{0}$. T preserves zero vector
2. If $\boldsymbol{u_1}, \boldsymbol{u_2}, \ldots, \boldsymbol{u_k} \in \mathbb{R}^n$ and $c_1, c_2, \ldots, c_k \in \mathbb{R}$, then T preserves linear combi

$$T(c_1\boldsymbol{u_1} + c_2\boldsymbol{u_2} + \cdots + c_k\boldsymbol{u_k}) = c_1T(\boldsymbol{u_1}) + c_2T(\boldsymbol{u_2}) + \cdots + c_kT(\boldsymbol{u_k}).$$

1

Discussion 7.1.6 Let $\{\boldsymbol{u_1}, \boldsymbol{u_2}, \ldots, \boldsymbol{u_n}\}$ be a basis for $\mathbb{R}^n$. Any vector $\boldsymbol{v} \in \mathbb{R}^n$ can be written as

Linearity condition

$$\boldsymbol{v} = c_1\boldsymbol{u_1} + c_2\boldsymbol{u_2} + \cdots + c_n\boldsymbol{u_n}$$

for some $c_1, c_2, \ldots, c_n \in \mathbb{R}$. Suppose $T : \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation. By Theorem 7.1.4.2,

Images of the basis vectors

Image of a general vector

$$T(\boldsymbol{v}) = c_1\boxed{T(\boldsymbol{u_1})} + c_2\boxed{T(\boldsymbol{u_2})} + \cdots + c_n\boxed{T(\boldsymbol{u_n})}.$$

It follows that the image $T(\boldsymbol{v})$ of $\boldsymbol{v}$ is completely determined by the images $T(\boldsymbol{u_1}), T(\boldsymbol{u_2}), \ldots, T(\boldsymbol{u_n})$ of the basis vectors $\boldsymbol{u_1}, \boldsymbol{u_2}, \ldots, \boldsymbol{u_n}$.

Example

Example 7.1.7 Let $T : \mathbb{R}^3 \to \mathbb{R}^2$ be the linear transformation such that

$$T\left(\begin{pmatrix}1\\1\\1\end{pmatrix}\right) = \begin{pmatrix}1\\3\end{pmatrix}, \quad T\left(\begin{pmatrix}0\\1\\1\end{pmatrix}\right) = \begin{pmatrix}-1\\2\end{pmatrix}, \quad T\left(\begin{pmatrix}2\\0\\-1\end{pmatrix}\right) = \begin{pmatrix}4\\-1\end{pmatrix}.$$

1. Find the image of the vector $(-1, 4, 6)^T$ under T.
2. Find the formula of T.

Solution

1. We first observe that the vectors

$$\left\{\begin{pmatrix}1\\1\\1\end{pmatrix}, \begin{pmatrix}0\\1\\1\end{pmatrix}, \begin{pmatrix}2\\0\\-1\end{pmatrix}\right\}$$

form a basis for $\mathbb{R}^3$ (check it). Thus the given information determines T completely.

To find the image of $(-1, 4, 6)^T$, we first write it as a linear combination of vectors from our basis. That is, we find constants c_1, c_2 and c_3 such that

$$\begin{pmatrix}-1\\4\\6\end{pmatrix} = c_1\begin{pmatrix}1\\1\\1\end{pmatrix} + c_2\begin{pmatrix}0\\1\\1\end{pmatrix} + c_3\begin{pmatrix}2\\0\\-1\end{pmatrix}.$$

Solving the equation, we get $c_1 = 3$, $c_2 = 1$ and $c_3 = -2$. Thus the required image is

$$T\left(\begin{pmatrix}-1\\4\\6\end{pmatrix}\right) = T\left(3\begin{pmatrix}1\\1\\1\end{pmatrix} + \begin{pmatrix}0\\1\\1\end{pmatrix} - 2\begin{pmatrix}2\\0\\-1\end{pmatrix}\right)$$

$$= 3T\left(\begin{pmatrix}1\\1\\1\end{pmatrix}\right) + T\left(\begin{pmatrix}0\\1\\1\end{pmatrix}\right) - 2T\left(\begin{pmatrix}2\\0\\-1\end{pmatrix}\right)$$

$$= 3\begin{pmatrix}1\\3\end{pmatrix} + \begin{pmatrix}-1\\2\end{pmatrix} - 2\begin{pmatrix}4\\-1\end{pmatrix} = \begin{pmatrix}-6\\13\end{pmatrix}.$$

Non-examples

$T(0) = 0$

$T(\alpha u) = \alpha T(u)$

$T(u+v) = T(u) + T(v)$

Let $T_1 : \mathbb{R}^2 \to \mathbb{R}^2$ be defined by

$$T_1\left(\begin{pmatrix}x\\y\end{pmatrix}\right) = \begin{pmatrix}x+1\\y+3\end{pmatrix} \quad \text{for } \begin{pmatrix}x\\y\end{pmatrix} \in \mathbb{R}^2.$$

It is a not a linear transformation. Zero vector not preserved

Note that $T_1\left(\begin{pmatrix}0\\0\end{pmatrix}\right) = \begin{pmatrix}1\\3\end{pmatrix}$ which violates Theorem 7.1.4.1.

Let $T_2 : \mathbb{R}^3 \to \mathbb{R}^2$ be defined by

$$T_2\left(\begin{pmatrix}x\\y\\z\end{pmatrix}\right) = \begin{pmatrix}x^2\\yz\end{pmatrix} \quad \text{for } \begin{pmatrix}x\\y\\z\end{pmatrix} \in \mathbb{R}^3.$$

It is a not a linear transformation.

Note that T_5 violates Theorem 7.1.4.2. For example,

$$T_2\left(\begin{pmatrix}1\\0\\0\end{pmatrix} + \begin{pmatrix}1\\0\\0\end{pmatrix}\right) = T_2\left(\begin{pmatrix}2\\0\\0\end{pmatrix}\right) = \begin{pmatrix}4\\0\end{pmatrix}$$

and Linear combi not preserved

$$T_2\left(\begin{pmatrix}1\\0\\0\end{pmatrix}\right) + T_2\left(\begin{pmatrix}1\\0\\0\end{pmatrix}\right) = \begin{pmatrix}1\\0\end{pmatrix} + \begin{pmatrix}1\\0\end{pmatrix} = \begin{pmatrix}2\\0\end{pmatrix}.$$

2

$T : \mathbf{R}^n \to \mathbf{R}^m$: any linear transformation

$\{\mathbf{e_1}, \mathbf{e_2}, \ldots, \mathbf{e_n}\}$: the standard basis for $\mathbf{R}^n$

$$\boldsymbol{A} = \begin{pmatrix}a_{11} & a_{12} & \cdots & a_{1n}\\ a_{21} & a_{22} & \cdots & a_{2n}\\ \vdots & \vdots & & \vdots\\ a_{m1} & a_{m2} & \cdots & a_{mn}\end{pmatrix}$$ the standard matrix of T

$$T(\mathbf{e}_1) = \mathbf{A}\mathbf{e}_1 = \begin{pmatrix}a_{11} & a_{12} & \cdots & a_{1n}\\ a_{21} & a_{22} & \cdots & a_{2n}\\ \vdots & \vdots & & \vdots\\ a_{m1} & a_{m2} & \cdots & a_{mn}\end{pmatrix}\begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix} = \begin{pmatrix}a_{11}\\a_{21}\\\vdots\\a_{m1}\end{pmatrix}$$

The image $T(\boldsymbol{e_j})$ = the jth column of $\boldsymbol{A}$

$$\boldsymbol{A} = (T(\mathbf{e_1}) \ \ T(\mathbf{e_2}) \ \cdots \ T(\mathbf{e_n}))$$

Remark 7.1.3

Prove Remark 7.1.3:

Show that a mapping $T : \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation if and only if

$$T(c\boldsymbol{u} + d\boldsymbol{v}) = cT(\boldsymbol{u}) + dT(\boldsymbol{v}) \quad \text{for all } \boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^n \text{ and } c, d \in \mathbb{R}.$$

Prove Linear Transformation

2. Is the mapping $T\begin{pmatrix}x\\y\\z\end{pmatrix} = \left[\begin{pmatrix}x\\y\\z\end{pmatrix} \cdot \begin{pmatrix}1\\2\\3\end{pmatrix}\right]\begin{pmatrix}1\\2\\3\end{pmatrix}$ a linear transformation?

$$T\begin{pmatrix}x\\y\\z\end{pmatrix} = (x+2y+3z)\begin{pmatrix}1\\2\\3\end{pmatrix} = \begin{pmatrix}x+2y+3z\\2x+4y+6z\\3x+6y+9z\end{pmatrix} = x\begin{pmatrix}1\\2\\3\end{pmatrix} + y\begin{pmatrix}2\\4\\6\end{pmatrix} + z\begin{pmatrix}3\\6\\9\end{pmatrix}$$

$$= \begin{pmatrix}1&2&3\\2&4&6\\3&6&9\end{pmatrix}\begin{pmatrix}x\\y\\z\end{pmatrix}.$$ T is a linear transformation.

$$T(x) = (x \cdot v)v = v(v \cdot x) = v(v^T x) = (vv^T)x = Ax$$

Finding the formula for T

Method 1: GJE

Let $(x, y, z)^T$ be any vector in $\mathbb{R}^3$. Following the procedure of Part (a), by solving the equation

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = c_1 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + c_2 \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} + c_3 \begin{pmatrix} 2 \\ 0 \\ -1 \end{pmatrix},$$

where x, y, z are regarded as constants, we obtain $c_1 = x - 2y + 2z$, $c_2 = -x + 3y - 2z$ and $c_3 = y - z$, i.e.

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

So the general formula of T is

$$T\left(\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right) = (x - 2y + 2z) \begin{pmatrix} 1 \\ 3 \end{pmatrix} + (-x + 3y - 2z) \begin{pmatrix} -1 \\ 2 \end{pmatrix} + (y - z) \begin{pmatrix} 4 \\ -1 \end{pmatrix}$$
$$= \begin{pmatrix} 2x - y \\ x - y + 3z \end{pmatrix}.$$

$$\begin{pmatrix} u_1 & u_2 & u_3 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

$$\begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = \begin{pmatrix} u_1 & u_2 & u_3 \end{pmatrix}^{-1} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

Method 2: Find T(e1), T(e2), T(e3)

$$T\left(\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}\right) = \begin{pmatrix} 1 \\ 3 \end{pmatrix} \quad T\left(\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}\right) = \begin{pmatrix} -1 \\ 2 \end{pmatrix} \quad T\left(\begin{pmatrix} 2 \\ 0 \\ -1 \end{pmatrix}\right) = \begin{pmatrix} 4 \\ -1 \end{pmatrix}$$

$\mathbf{A} = (\ T(\mathbf{e}_1)\ T(\mathbf{e}_2)\ T(\mathbf{e}_3)\ )$

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

Find $T(\mathbf{e}_1)$, $T(\mathbf{e}_2)$, $T(\mathbf{e}_3)$

Find $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ in terms of $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ $\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$ $\begin{pmatrix} 2 \\ 0 \\ -1 \end{pmatrix}$

$$\left(\begin{array}{ccc|c|c|c} 1 & 0 & 2 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & -1 & 0 & 0 & 1 \end{array}\right) \xrightarrow{\text{Gauss-Jordan elimination}} \left(\begin{array}{ccc|c|c|c} 1 & 0 & 0 & 1 & -2 & 2 \\ 0 & 1 & 0 & -1 & 3 & -2 \\ 0 & 0 & 1 & 0 & 1 & -1 \end{array}\right)$$

$$\mathbf{e}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} - \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \qquad \mathbf{e}_2 = -2\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + 3\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} + \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix} \qquad \mathbf{e}_3 = 2\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} - 2\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} - \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}$$

$$T(\mathbf{e}_1) = T\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} - T\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \end{pmatrix} - \begin{pmatrix} -1 \\ 2 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

$$T(\mathbf{e}_2) = -2T\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + 3T\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} + T\begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix} = -2\begin{pmatrix} 1 \\ 3 \end{pmatrix} + 3\begin{pmatrix} -1 \\ 2 \end{pmatrix} + \begin{pmatrix} 4 \\ -1 \end{pmatrix} = \begin{pmatrix} -1 \\ -1 \end{pmatrix}$$

$$T(\mathbf{e}_3) = 2T\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} - 2T\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} - T\begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix} = 2\begin{pmatrix} 1 \\ 3 \end{pmatrix} - 2\begin{pmatrix} -1 \\ 2 \end{pmatrix} - \begin{pmatrix} 4 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 3 \end{pmatrix}$$

Method 3: Stacking matrices

$$T\left(\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}\right) = \begin{pmatrix} 1 \\ 3 \end{pmatrix} \quad T\left(\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}\right) = \begin{pmatrix} -1 \\ 2 \end{pmatrix} \quad T\left(\begin{pmatrix} 2 \\ 0 \\ -1 \end{pmatrix}\right) = \begin{pmatrix} 4 \\ -1 \end{pmatrix}$$

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

$T(u_1), T(u_2), \ldots, T(u_k)$

works only if $\{u_1, \ldots, u_k\}$ is a basis.

$$\left(\begin{array}{ccc|c|c|c} 1 & 0 & 2 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & -1 & 0 & 0 & 1 \end{array}\right) \xrightarrow{\text{Gauss-Jordan elimination}} \left(\begin{array}{ccc|c|c|c} 1 & 0 & 0 & 1 & -2 & 2 \\ 0 & 1 & 0 & -1 & 3 & -2 \\ 0 & 0 & 1 & 0 & 1 & -1 \end{array}\right)$$

$$\mathbf{A}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \end{pmatrix}, \mathbf{A}\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ 2 \end{pmatrix}, \mathbf{A}\begin{pmatrix} 2 \\ 0 \\ -1 \end{pmatrix} = \begin{pmatrix} 4 \\ -1 \end{pmatrix}$$

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

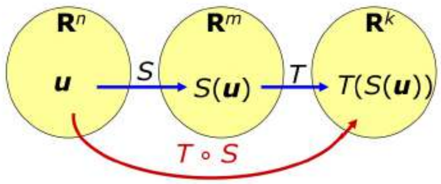
$$\mathbf{A} = \begin{pmatrix} 1 & -1 & 4 \\ 3 & 2 & -1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 2 \\ 1 & 1 & 0 \\ 1 & 1 & -1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & -1 & 4 \\ 3 & 2 & -1 \end{pmatrix}\begin{pmatrix} 1 & -2 & 2 \\ -1 & 3 & -2 \\ 0 & 1 & -1 \end{pmatrix}$$

Composition of Linear Combinations

Let $S : \mathbf{R}^n \to \mathbf{R}^m$ and $T : \mathbf{R}^m \to \mathbf{R}^k$ be linear transformations.

The composition of T with S, denoted by $T \circ S$ First S, then T is a mapping from $\mathbf{R}^n$ to $\mathbf{R}^k$ such that

$(T \circ S)(\boldsymbol{u}) = T(S(\boldsymbol{u}))$ for all $\boldsymbol{u}$ in $\mathbf{R}^n$.



Theorem 7.1.1 1

If $S : \mathbf{R}^n \to \mathbf{R}^m$ and $T : \mathbf{R}^m \to \mathbf{R}^k$ are linear transformations

S, T have standard matrices $\boldsymbol{A}$, $\boldsymbol{B}$ respectively

then $T \circ S : \mathbf{R}^n \to \mathbf{R}^k$ is again a linear transformation.

$T \circ S$ has standard matrix $\boldsymbol{BA}$

$$S\left(\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right) = \begin{pmatrix} x+y \\ z \end{pmatrix} \qquad T\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = \begin{pmatrix} y \\ y \\ x \end{pmatrix} \qquad (T \circ S)\left(\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right) = \begin{pmatrix} z \\ z \\ x+y \end{pmatrix}$$

$$\boldsymbol{A} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad \boldsymbol{B} = \begin{pmatrix} 0 & 1 \\ 0 & 1 \\ 1 & 0 \end{pmatrix}$$

$$\boldsymbol{BA} = \begin{pmatrix} 0 & 1 \\ 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$ standard matrix of $T \circ S$

$$(T \circ S)\left(\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right) = \mathbf{BA}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} z \\ z \\ x+y \end{pmatrix}$$

Example

Example 7.1.12 Let $S : \mathbb{R}^3 \to \mathbb{R}^2$ be the linear transformation defined by

$$S\left(\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right) = \begin{pmatrix} x+y \\ z \end{pmatrix} \quad \text{for} \quad \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3$$

and $T : \mathbb{R}^2 \to \mathbb{R}^3$ the linear transformation defined by

$$T\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = \begin{pmatrix} y \\ y \\ x \end{pmatrix} \quad \text{for} \quad \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2.$$

Then $T \circ S : \mathbb{R}^3 \to \mathbb{R}^3$ is given by

$$(T \circ S)\left(\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right) = T\left(S\left(\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right)\right) = T\left(\begin{pmatrix} x+y \\ z \end{pmatrix}\right) = \begin{pmatrix} z \\ z \\ x+y \end{pmatrix} \quad \text{for} \quad \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3.$$

The standard matrices for S, T and $T \circ S$ are

$$\begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 \\ 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$

respectively. Note that

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

$T(S(u)) = T(Au) = BAu$

B A

Range

Column space of A

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

- denoted by $R(T)$
- the set of images of T.

$R(T) = \{\text{images of } T\}$

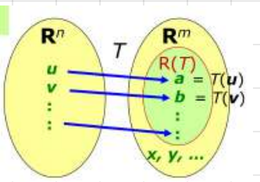
$R(T) = \{\, T(\boldsymbol{u}) \mid \boldsymbol{u} \in \mathbf{R}^n \,\}$

explicit set notation

$R(T)$ is a subset of $\mathbf{R}^m$

$R(T)$ may not be equal to $\mathbf{R}^m$

range of T ⊆ codomain of T



$T : \mathbf{R}^2 \to \mathbf{R}^3$: the linear transformation defined by

$$T\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = \begin{pmatrix} x+y \\ y \\ x \end{pmatrix} = x\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + y\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \quad \text{for all} \quad \begin{pmatrix} x \\ y \end{pmatrix} \quad \text{in} \quad \mathbf{R}^2.$$

What is $R(T)$?

standard matrix $\boldsymbol{A} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \end{pmatrix}$

$$R(T) = \left\{ \begin{pmatrix} x+y \\ y \\ x \end{pmatrix} \;\middle|\; x, y \in \mathbf{R} \right\} = \operatorname{span}\left\{ \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \right\}$$ a plane in $\mathbb{R}^3$

explicit set notation

linear span form
column space of $\boldsymbol{A}$

Theorem 7.24

$T : \mathbf{R}^n \to \mathbf{R}^m$: a linear transformation

$\boldsymbol{A}$ the standard matrix for T

Then $R(T)$ = span{ columns of $\boldsymbol{A}$ }
= the column space of $\boldsymbol{A}$

Discussion 7.2.3

$T : \mathbf{R}^n \to \mathbf{R}^m$ a linear transformation

$R(T)$ = span{ columns of $\boldsymbol{A}$ }
= span $\{ T(\boldsymbol{e}_1), T(\boldsymbol{e}_2), \ldots, T(\boldsymbol{e}_n) \}$

If $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_n\}$ is any basis for $\mathbf{R}^n$

then $\boxed{R(T) = \operatorname{span}\{ T(\boldsymbol{u}_1), T(\boldsymbol{u}_2), \ldots, T(\boldsymbol{u}_n) \}}$

Rank of a Linear Transformation

Let T be a linear transformation.

The dimension of R(T) = dimension of column space of $\boldsymbol{A}$

called the rank of T denoted by rank(T)

$\boldsymbol{A}$ the standard matrix for T rank(T) = rank($\boldsymbol{A}$)

Finding the range R(T) and its basis

T: $\mathbf{R}^n \to \mathbf{R}^m$

I. if formula of T is given

➢ R(T) ={formula in $x_1, x_2, \dots, x_n \mid x_1, x_2, \dots, x_n \in \mathbf{R}$ }

II. if standard matrix $\mathbf{A}$ is given

➢ R(T) = span{ columns of $\mathbf{A}$ }

or part I above

Find basis for column space of $\mathbf{A}$

III. if image of a basis $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n\}$ for $\mathbf{R}^n$ is given

➢ R(T) = span{T($\mathbf{u}_1$), T($\mathbf{u}_2$), ..., T($\mathbf{u}_n$)}

Throw out the redundant vectors in the span
(use column space method if necessary)

Kernel

Nullspace of A

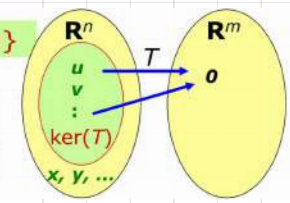
Let $T: \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation.

- denoted by ker(T)
- the set of vectors in Rn whose image is the zero vector in Rm.
- ker(T) = {vectors that map to 0 under T }

ker(T) = { $\boldsymbol{u} \in \mathbf{R}^n \mid T(\boldsymbol{u}) = \mathbf{0}$ }
implicit set notation

ker(T) is a subset of $\mathbf{R}^n$

ker(T) may not be equal to $\mathbf{R}^n$



Theorem 7.2.9

Theorem 7.2.9 Let $T: \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation and $\boldsymbol{A}$ the standard matrix for T. Then

$$\mathrm{Ker}(T) = \text{the nullspace of } \boldsymbol{A}$$

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

Dimension of Kernel of a LT

ker(T) = the nullspace of standard matrix A

dim(ker(T)) = nullity(T) = nullity(A)

Theorem 7.2.12

Let $T: \mathbf{R}^n \to \mathbf{R}^m$ be any linear transformation.

rank(T) + nullity(T) = n

By Thm 4.3.4. rank($\boldsymbol{A}$) + nullity($\boldsymbol{A}$) = n (number of columns)

The standard matrix A of T is of size m x n

Example

Example 7.2.6 Let $T: \mathbb{R}^4 \to \mathbb{R}^4$ be a linear transformation defined by

$$T\left(\begin{pmatrix} w \\ x \\ y \\ z \end{pmatrix}\right) = \begin{pmatrix} x + 2y + z \\ x + 3y \\ x + 4y - z \\ y - z \end{pmatrix} \quad \text{for} \quad \begin{pmatrix} w \\ x \\ y \\ z \end{pmatrix} \in \mathbb{R}^4.$$

Find a basis for the range of T and determine the rank of T.

Solution The range of T is equal to the column space of the matrix

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

By Gaussian Elimination, we reduce $\boldsymbol{A}$ to

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

So $\{(1,1,1,0)^T, (2,3,4,1)^T\}$ is a basis for R(T) and

$$\mathrm{rank}(T) = \dim(\mathrm{R}(T)) = 2.$$

Let $T_1: \mathbb{R}^3 \to \mathbb{R}^4$ be a linear transformation defined by

$$T_1\left(\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right) = \begin{pmatrix} 2x - y \\ x - y + 3z \\ -5x + y \\ x - z \end{pmatrix} \quad \text{for} \quad \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3.$$

The kernel of T_1 is the set of vectors $\boldsymbol{u} \in \mathbb{R}^3$ such that $T_1(\boldsymbol{u}) = \mathbf{0}$. That is, we need to solve

$$T_1\left(\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right) = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

This is equivalent to

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

By solving the linear system, we get only the trivial solution $x = 0$, $y = 0$, $z = 0$. Thus

$$\mathrm{Ker}(T_1) = \{(0,0,0)^T\}.$$ nullity(T) = 0

Let $T_2: \mathbb{R}^3 \to \mathbb{R}^3$ be a linear transformation defined by

$$T_2\left(\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right) = \begin{pmatrix} z - y \\ 0 \\ x \end{pmatrix} \quad \text{for} \quad \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3.$$

By equating $\begin{pmatrix} z - y \\ 0 \\ x \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$, we get $x = 0$ and $y = z$. So

nullity(T) = 1

$$\mathrm{Ker}(T_2) = \left\{ \begin{pmatrix} 0 \\ y \\ y \end{pmatrix} \middle| \, y \in \mathbb{R} \right\} = \mathrm{span}\{(0,1,1)^T\}.$$

Finding Kernel of a Linear Transformation

- set formula = 0 and solve this homogeneous system the general solution gives ker(T)

OR

- use nullspace of A

Using Range and Kernel in Proof

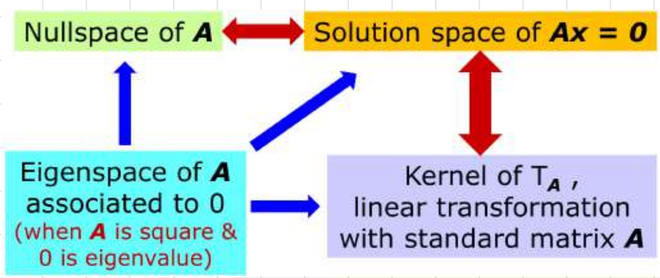
Let $T : \mathbf{R}^n \to \mathbf{R}^m$ be a linear transformation.

$\mathrm{Ker}(T) = \{ \mathbf{v} \in \mathbf{R}^n \mid T(\mathbf{v}) = \mathbf{0} \}$

if you want to show:
In a proof, if you start with: $\mathbf{v} \in \ker(T)$,
try to show:
you should follow by: $T(\mathbf{v}) = \mathbf{0}$.

$R(T) = \{ T(\mathbf{v}) \mid \mathbf{v} \in \mathbf{R}^n \}$

if you want to show:
In a proof, if you start with: $\mathbf{v} \in R(T)$,
try to show:
you should follow by: $\mathbf{v} = T(\mathbf{u})$ for some $\mathbf{u} \in \mathbf{R}^n$.



Linear transformation vs Subspaces

$T: \mathbf{R}^n \to \mathbf{R}^m$ linear transformation	U is a subspace of $\mathbf{R}^n$
Linearity conditions	Closure Properties
(i) T $\boxed{\text{preserves}}$ addition Let $\mathbf{u}, \mathbf{v} \in \mathbf{R}^n$. Then $T(\mathbf{u}+\mathbf{v}) = T(\mathbf{u})+T(\mathbf{v})$.	(a) U is $\boxed{\text{closed}}$ under addition Let $\mathbf{u}, \mathbf{v} \in U$. Then $\mathbf{u}+\mathbf{v} \in U$.
(ii) T $\boxed{\text{preserves}}$ scalar mult Let $\mathbf{u} \in \mathbf{R}^n$, $c \in \mathbf{R}$. Then $T(c\mathbf{u}) = cT(\mathbf{u})$.	(b) U is $\boxed{\text{closed}}$ under scalar mult Let $\mathbf{u} \in U$, $c \in \mathbf{R}$. Then $c\mathbf{u} \in U$.
(iii) T $\boxed{\text{preserves}}$ zero vector $T(\mathbf{0}) = \mathbf{0}$	(c) U $\boxed{\text{contains}}$ the zero vector $\mathbf{0} \in U$
If one of (i), (ii), (iii) is violated, T is not a linear transformation	If one of (a), (b), (c) is violated, U is not a subspace of $\mathbf{R}^n$

Linear transformation vs standard matrix

Linear Transformation	Standard matrix
$T: \mathbf{R}^n \to \mathbf{R}^m$ $S: \mathbf{R}^m \to \mathbf{R}^k$	$\mathbf{A}$ is an $m \times n$ matrix $\mathbf{B}$ is an $k \times m$ matrix
$T(\mathbf{u})$	$\mathbf{Au}$
$T(\mathbf{e}_1), T(\mathbf{e}_2), \ldots, T(\mathbf{e}_n)$	columns of $\mathbf{A}$
$R(T)$	column space of $\mathbf{A}$
$\mathrm{Ker}(T)$	nullspace of $\mathbf{A}$
$S \circ T$	$\mathbf{BA}$

Row equivalence preserves

✅ row space
✅ linear dependency of rows
⛔ linearity relations of rows

⛔ column space
✅ linear dependency of columns
✅ linearity relations of columns

✅ rank & nullity
✅ nullspace
✅ Solutions
✅ Invertibiltiy

⛔ Transpose
⛔ Determinant
⛔ Eigenvalues/Eigenvectors/ charac poly