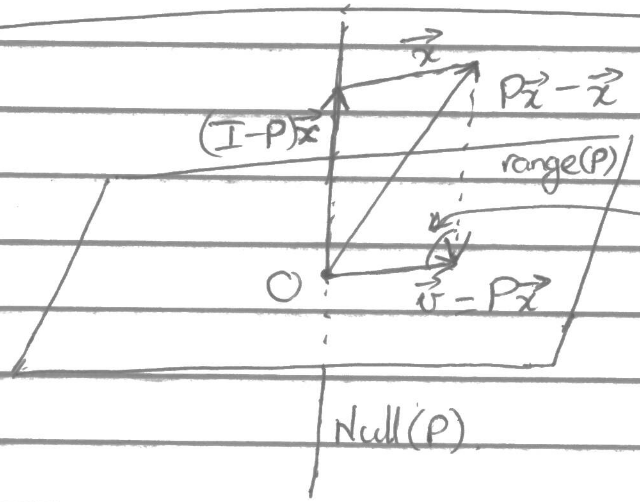
Projectors.

* Projectors:

Def: A matrix $P \in \mathbb{R}^{m \times m}$ is called a projector if $P^2 = P$



In general, this angle may not be 90°

Let $\vec{v} \in$ range(P).
Then $\exists\ \vec{x} \in \mathbb{R}^m$ such that $P\vec{x} = \vec{v}$.

$\Rightarrow \quad P\vec{v} = P(P\vec{x}) = P^2\vec{x} = P\vec{x} = \vec{v}$.

$\Rightarrow$ if $\vec{v} \in$ range(P), then applying P to $\vec{v}$ does not change $\vec{v}$.

In addition, $\forall x \in \mathbb{R}^m$, $P\vec{x} - \vec{x} \in$ Null(P).

Pf: $P(P\vec{x} - \vec{x}) = P^2\vec{x} - P\vec{x} = P\vec{x} - P\vec{x} = 0$.

Def: Let $P \in \mathbb{R}^{m \times m}$ be a projector. Then $I - P$ is also a projector and is called the complementary projector to P.

$$\begin{aligned}(I-P)^2 = (I-P)(I-P) &= I - P - PI + PP \\ &= I - 2P + P \\ &= I - P.\end{aligned}$$

$I-P$ is a projector onto $Null(P)$

$$\boxed{\begin{aligned} \underline{\text{Thm}}: \quad range(I-P) &= Null(P) \\ Null(I-P) &= range(P) \end{aligned}}$$

In other words, ~~is~~ P and $I-P$ are complementary!

$\underline{\text{Pf.}}$ Take any $\vec{v} \in Null(P)$, i.e., $P\vec{v} = 0$.

$\Rightarrow \quad (I-P)\vec{v} = \vec{v} - P\vec{v} = \vec{v}$.

$\Rightarrow \quad \vec{v} \in range(I-P)$ because $\vec{v}$ is written as a matrix-vector product $(I-P)\vec{v}$.

$\Rightarrow \quad Null(P) \subset range(I-P)$.

Now, take any $\vec{v} \in range(I-P)$.

Then, $\exists\, \vec{x} \in \mathbb{R}^m$ s.t.

$$\vec{v} = (I-P)\vec{x}.$$

$$\begin{aligned} \Rightarrow \quad P\vec{v} &= P(I-P)\vec{x} \\ &= (P-P^2)\vec{x} \\ &= 0. \end{aligned}$$

$\Rightarrow \quad \vec{v} \in Null(P)$.

$\Rightarrow \quad range(I-P) \subset Null(P)$.

(Exercise: show $Null(I-P) = range(P)$)

Thm: $\text{Null}(I-P) \cap \text{Null}(P) = \{0\}$.

i.e. $\text{range}(P) \cap \text{Null}(P) = \{0\}$.

Pf: Take any $\vec{v} \in \text{Null}(P) \cap \text{Null}(I-P)$

$\Leftrightarrow \quad P\vec{v}$ and $(I-P)\vec{v} = 0$

$\Leftrightarrow \quad \vec{v} = 0$.

"A projector separates $\mathbb{R}^m$ into two spaces, i.e.,

$$\mathbb{R}^m = \text{range}(P) + \text{null}(P) \text{ ."}$$

In other words,

$\forall \vec{v} \in \mathbb{R}^m, \quad \exists \vec{v}_1 \in \text{range}(P)$

and $\vec{v}_2 \in \text{Null}(P)$ such that

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

and this decomposition is <u>unique</u> for a given projector P

Why? Suppose this decomposition is not unique.

Then $\exists \vec{x} \in \mathbb{R}^m, \ \vec{x} \neq 0$ such that

$$\vec{v} = \underbrace{(\vec{v}_1 + \vec{x})}_{\in \text{range}(P)} + \underbrace{(\vec{v}_2 - \vec{x})}_{\in \text{Null}(P)}$$

$\Rightarrow$ This means that

$$\vec{x} \in \text{range}(P) \quad \& \quad \vec{x} \in \text{Null}(P),$$

i.e. $\vec{x} \in \text{range}(P) \cap \text{Null}(P) = \{0\}$.

$\Rightarrow \quad \vec{x} = 0$.

E.g. $P = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \in \mathbb{R}^{2\times 2}$ $\qquad P^2 = P$

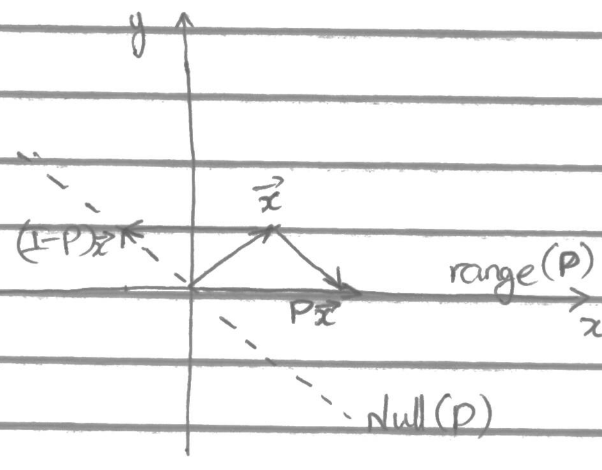
$\Rightarrow I - P = \begin{bmatrix} 0 & -1 \\ 0 & 1 \end{bmatrix}$ and $(I-P)^2 = I - P$

$\text{range}(P) = \text{span}\left\{\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right\}$.

$\text{Null}(P) = \text{span}\left\{\begin{bmatrix} -1 \\ 1 \end{bmatrix}\right\}$.

So, $\mathbb{R}^2 = \text{span}\left\{\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right\} + \text{span}\left\{\begin{bmatrix} -1 \\ 1 \end{bmatrix}\right\}$.

(But this is not an orthogonal decomposition.)



* <u>Orthogonal Projectors:</u>

<u>Def:</u> A projector $P \in \mathbb{R}^{m\times m}$ is said to be <u>orthogonal</u> if $\text{range}(P) \perp \text{null}(P)$.

Exercise: Consider $P = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ in $\mathbb{R}^2$. This is the orthogonal projector onto "x-axis". The complementary proj. is also orthogonal, i.e., orth. proj. onto "y-axis", and

$$\mathbb{R}^2 = \text{span}\left\{\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right\} \underset{\text{orthogonal.}}{\oplus} \text{span}\left\{\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right\}$$

Remark: Do not confuse an orthogonal projector P with an orthogonal matrix!

What happens if P is a projector and is an orthogonal matrix?

$$P^2 = P \qquad \text{and} \qquad P^T = P^{-1}$$

$$\Rightarrow \quad P^T P^2 = P^T P \qquad\qquad \Rightarrow \quad P^T P = I.$$

$$\boxed{P = I.}$$

Thm: A projector P is an orthogonal projector iff $P^T = P$, i.e., symmetric.

Pf. ($\Leftarrow$) Take any $\vec{v}_1 \in \text{range}(P)$

any $\vec{v}_2 \in \text{Null}(P)$

Then $\exists \vec{x} \in \mathbb{R}^m$ such that $\vec{v}_1 = P\vec{x}$.

and
$$v_1^T v_2 = (P\vec{x})^T v_2 = x^T P^T v_2$$
$$= x^T P v_2 \quad \text{as } P^T = P.$$
$$= 0.$$

$\Rightarrow$ $\text{range}(P) \perp \text{Null}(P)$.

($\Rightarrow$) ~~Suppos~~ Since range(P) $\perp$ Null(P),
$\exists$ orthonormal basis of $\mathbb{R}^m$ such that
$\{q_1, \dots, q_m\}$ s.t.

$$\text{range}(P) = \text{span}\{q_1, \dots, q_n\}.$$
$$\text{Null}(P) = \text{span}\{q_{n+1}, \dots, q_m\}$$

Then, $Pq_j = \begin{cases} q_j & \text{for } 1 \le j \le n \\ 0 & \text{for } n+1 \le j \le m. \end{cases}$

Then

$$PQ \overset{\text{why?}}{=} Q \underbrace{\left[\begin{array}{c|c} I_{n\times n} & O_{n\times(m-n)} \\ \hline O_{(m-n)\times n} & O_{(m-n)\times(m-n)} \end{array}\right]}_{\Lambda}$$

$$\Rightarrow \quad PQ = Q\Lambda$$
$$\Rightarrow \quad P\underbrace{QQ^T}_{I} = Q\Lambda Q^T.$$
I as columns of Q are orthonormal.
$$\Rightarrow \quad P = Q\Lambda Q^T$$
$$\Rightarrow \quad P^T = (Q\Lambda Q^T)^T = (Q^T)^T\Lambda^T Q^T = Q\Lambda Q^T = P.$$

Projection with Bases.

$*$ Projection with an orthonormal basis:

In general, consider an orthogonal projector $P \in \mathbb{R}^{m\times m}$ where $\dim(\text{range}(P)) = n \leq m$.

Let's define the matrix

$$\hat{Q} := [q_1 \dots q_n] \in \mathbb{R}^{m\times n}.$$

where $\{q_1, \dots, q_n\}$ forms an orthonormal basis (O.N.B) of range(P).

Then, $\quad P = \hat{Q}\hat{Q}^T$

Recall that for any $\vec{v} \in \mathbb{R}^m$, $\exists\ \vec{r} \in \mathbb{R}^m$ (residual) such that

$$\vec{v} = \vec{r} + \sum_{i=1}^{n} (\vec{q}_i \vec{q}_i^T)\vec{v}.$$

The mapping $v \mapsto \sum_{i=1}^{n} (q_i q_i^T)\vec{v} = y$ is an orthogonal projection onto $\text{range}(\hat{Q}) = \text{range}(P)$.

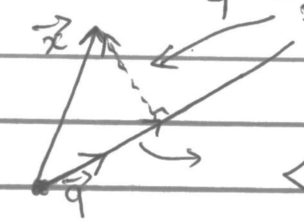
- Special case: The rank-one orthogonal projection with a unit vector $q \in \mathbb{R}^m$.

$$P_q := \underbrace{qq^T}_{\text{rank-1 matrix}} \in \mathbb{R}^{m\times m}.$$

Its complementary projection is

$$P_q^{\perp} := I - P_q.$$

$x - qq^Tx = (I - qq^T)x$



$\langle q, x\rangle q = qq^T x.$

For a general vector $a \in \mathbb{R}^m$ with $a \neq 0$, $\|a\| \neq 1$, the orthogonal projection onto span$\{a\}$ becomes

$$P_a = \frac{aa^T}{a^Ta} \quad \text{and} \quad P_a^{\perp} = I - \frac{aa^T}{a^Ta}.$$

(Pf: Let $q = \frac{a}{\|a\|_2}$.)

* <u>Projection with an arbitrary basis.</u>

Let $\{a_1, \dots, a_n\} \subset \mathbb{R}^m$ be a set of linearly independent vectors. ($\Rightarrow n \leq m$).

Let $A = [a_1 \dots a_n]$.

What is the orthogonal projection onto

$$\text{span}\{a_1, \dots, a_n\} = \text{range}(A)?$$

<u>FTOLA.</u>
$\text{Null}(A) \perp \text{im}(A^T)$
$\text{Null}(A^T) \perp \text{im}(A)$

Take any $v \in \mathbb{R}^m$. Then

$\exists\, v_1 \in \text{range}(A)$ and $\exists\, v_2 \in \text{range}(A)^{\perp}$ such that

$$v = v_1 + v_2 \quad (\text{where } v_1 \perp v_2).$$

$$\Rightarrow \underset{\text{s.t.}}{\exists x \in \mathbb{R}^n} \; v = Ax + v_2.$$

$$\Rightarrow v_2 = v - Ax.$$

and $a_j \perp v_2$ for $1 \leq j \leq n$.

$$\Rightarrow a_j^T (v - Ax) = 0, \quad 1 \leq j \leq n$$

$$A^T (v - Ax) = 0$$

$$A^Tv - A^TAx = 0$$

$$A^TAx = A^Tv.$$

Since A is of full rank, $(A^TA)^{-1}$ exists,

$$\Rightarrow \quad x = (A^T A)^{-1} A^T v.$$

Hence, the orthogonal projection onto range(A):

$$\boxed{P_A = A(A^T A)^{-1} A^T.}$$

E.x. $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}$.

1) What's the orthogonal projector P_A onto the Column space of A?

2) Let $v = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$. Compute its ~~co~~ orthogonal projection onto range(A)?

Sol.

1) $P_A = A(A^T A)^{-1} A^T$.

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

$$\Rightarrow (A^T A)^{-1} = \frac{1}{2} \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} 1/2 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\Rightarrow P_A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1/2 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 1/2 & 0 & 1/2 \\ 0 & 1 & 0 \\ 1/2 & 0 & 1/2 \end{bmatrix}.$$

2) $$P_A v = \begin{bmatrix} 1/2 & 0 & 1/2 \\ 0 & 1 & 0 \\ 1/2 & 0 & 1/2 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \\ 2 \end{bmatrix}.$$

* Mean Values: (the average of data points).

Def: Given a data set $D = \{x_1, x_2, \dots, x_n\} \subset \mathbb{R}^m$. The mean of the data set ~~as~~ is defined as

$$\mathbb{E}[D] = \frac{1}{n}\sum_{i=1}^{n} x_i$$

Examples:

1) Quiz grades: $D = \{8, 7, 6, 10, 10\}$.

$$\mathbb{E}[D] = \frac{1}{5}(8+7+6+10+10) = \frac{41}{5} \approx 8.2$$

2) Consider a data set D which includes students' grades:

$$D = \left\{ \underset{\text{student 1}}{\begin{bmatrix} 8 \\ 9 \\ 10 \end{bmatrix}}, \underset{\text{student 2}}{\begin{bmatrix} 10 \\ 10 \\ 10 \end{bmatrix}}, \underset{\text{student 3}}{\begin{bmatrix} 9 \\ 9 \\ 9 \end{bmatrix}} \right\} \begin{array}{l} \leftarrow \text{midterm 1} \\ \leftarrow \text{midterm 2} \\ \leftarrow \text{midterm 3} \end{array}$$

Then,

$$\mathbb{E}[D] = \frac{1}{3}\left(\begin{bmatrix} 8 \\ 9 \\ 10 \end{bmatrix} + \begin{bmatrix} 10 \\ 10 \\ 10 \end{bmatrix} + \begin{bmatrix} 9 \\ 9 \\ 9 \end{bmatrix}\right) = \begin{bmatrix} 9 \\ 28/3 \\ 29/3 \end{bmatrix}.$$

~~* Variance of one-dimensional data sets~~

* Variances and covariances:

Var: measures how far a set of (random) numbers are spread out from their mean value.

· 1-dim: Given $D = \{x_1, \dots, x_n\}$ where $x_i \in \mathbb{R}$,
The variance of D is defined as

$$\cancel{\mathbb{E}[D]} \quad Var[D] = \frac{1}{n} \sum_{i=1}^{n} (x_i - \mathbb{E}[D])^2$$

E.g. Consider

$D_1 = \{1, 2, 4, 5\} \Rightarrow \mathbb{E}[D_1] = 3.$

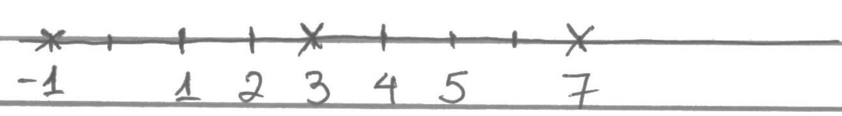
$$Var[D_1] = \frac{(1-3)^2 + (2-3)^2 + (4-3)^2 + (5-3)^2}{4} = \frac{10}{4}.$$

$D_2 = \{-1, 3, 7\} \Rightarrow \mathbb{E}[D_2] = 3.$

$$Var[D_2] = \frac{(-1-3)^2 + (3-3)^2 + (7-3)^2}{3} = \frac{32}{3}.$$

$Var[D_2] > Var[D_1].$

$\Rightarrow$ the spread of the data is higher in D_2 than D_1.



- Positive - definite matrix:

Def: A symmetric $m \times m$ real matrix M is said to be positive definite if

$$x^T M x > 0 \qquad \forall x \in \mathbb{R}^m \setminus \{0\}.$$

E.g. 1) The identity ~~matrices~~ ~~is~~ matrices are p.d.

Consider $I_{2\times 2} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$. Then

$$\forall z \neq 0, \; z^T I z = [z_1 \;\; z_2] \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}$$

$$= [z_1 \;\; z_2] \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}$$

$$= z_1^2 + z_2^2 > 0.$$

2) ~~Let~~ $M = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}$ is p.d. (show!)

Fact: For any real invertible matrix A, $A^T A$ is a positive definite matrix (show!)

Def: Given a data set $D = \{x_1, x_2, \dots, x_n\} \subset \mathbb{R}^m$, The variance of D is defined as

$$Var[D] = \frac{1}{n} \sum_{i=1}^{n} \underbrace{(x_i - \mu)(x_i - \mu)^T}_{\text{covariance matrix}} \in \mathbb{R}^{m\times m}$$

where $\mu = E[D]$.

E.g. $D = \left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 5 \\ 4 \end{bmatrix} \right\}.$

$$\cancel{Var =}\ \mathbb{E}[D] = \frac{1}{2}\left(\begin{bmatrix} 1 \\ 2 \end{bmatrix} + \begin{bmatrix} 5 \\ 4 \end{bmatrix} \right) = \begin{bmatrix} 3 \\ 3 \end{bmatrix} = \mu.$$

$$Var[D] = \frac{1}{2}\left(\begin{bmatrix} -2 \\ -1 \end{bmatrix} \begin{bmatrix} -2 & -1 \end{bmatrix} + \begin{bmatrix} 2 \\ 1 \end{bmatrix} \begin{bmatrix} 2 & 1 \end{bmatrix} \right).$$

$$= \frac{1}{2}\left(\begin{bmatrix} 4 & 2 \\ 2 & 1 \end{bmatrix} + \begin{bmatrix} 4 & 2 \\ 2 & 1 \end{bmatrix} \right).$$

$$= 4 \begin{bmatrix} 4 & 2 \\ 2 & 1 \end{bmatrix}.$$