1. Four Fundamental Subspaces

1.1. Column Space

- Denoted by $C(A)$

$$A = \begin{bmatrix} | & | & & | \\ u_1 & u_2 & \cdots & u_n \\ | & | & & | \end{bmatrix}$$

$$C(A) = span\{u_1,\ u_2,\ \cdots,\ u_n\} = Linear\ combination\ of\ vectors\ u_1,\ u_2,\ \cdots,\ u_n$$

Question: For what b does Ax = b have a solution ?
→ For all b $\in$ $C(A)$

Question: How to find Column space given a matrix ?
→ Let A be a matrix.
→ Find $R = RREF(A)$
→ Identify the pivot columns and pick the corresponding column vector in the original matrix A.
Example:

$$A = \begin{bmatrix} 15 & 12 & 9 \\ -5 & -4 & -3 \\ 11 & 8 & 5 \end{bmatrix}$$

$$R = RREF(A) = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix}$$

Now, as we can see only 1st and 2nd cols are pivot column so

$$C(A) = span\{(15, -5, 11), (12, -4, 8)\}$$

1.2. Null Space

$\rightarrow N(A) = \{x \mid Ax = 0,\ x \in \mathbb{R}^n\}$

→ Null space is a set of all solutions of a system of homogeneous equation $Ax = 0$.

Question: Why N(A) is a subspace ?

⟶ Because of 2 reasons

$$\begin{gathered}\rightarrow x_1,\ x_2 \in N(A)\\ Ax_1 = 0,\ Ax_2 = 0\\ A(x_1 + x_2) = 0\\ \Longrightarrow x_1 + x_2 \in N(A)\end{gathered}$$

$$\begin{gathered}\rightarrow x \in N(A),\ \alpha \in \mathbb{R}\\ A(\alpha x) = \alpha A x = 0\\ \Longrightarrow \alpha x \in N(A)\end{gathered}$$

Question: How do we interpret Null space ?

→ Find x such that $Ax = 0$

OR

→ Linear combination of the columns of A should result in $\vec{0}$.

Example:

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

$$col_1 + col_2 - col_3 = 0$$

Remark:

→ If A is invertible, then N(A) has "zero" only and C(A) is the whole space.

→ In this case, $Ax = b$ has $x_n \notin 0$ and $Ax = b$ solution are of the form $x = x_p + x_g$ where $Ax_p = b$ and $Ax_n = 0$.

Question: How to find Null space given A ?

→ A is a matrix given

→ Find $R = RREF(A)$

→ Find dependent and indpendent variable from R.

→ Assign $t_1,\ t_2,\ \cdots,\ t_n$ to independent variable and solve for dependent variable

→ Equate the R to 0 and Find basis

→ For more clarity Look these two examples.

Example-1:

$$A = \begin{bmatrix} 15 & 12 & 9 \\ -5 & -4 & -3 \\ 11 & 8 & 5 \end{bmatrix}$$

$$R = RREF(A) = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix}$$

$$x_1,\ x_2 \rightarrow dependent\ variable$$
$$x_3 \rightarrow Independent\ variable$$

$$x_3 = t$$
$$x_2 = -2t$$
$$x_1 = t$$

$$N(A) = \{(t,\ -2t,\ t)\}$$
$$N(A) = t(1, -2, 1)$$
$$N(A) = span\{(1, -2, 1)\}$$

Example-2:

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

$$R = RREF(A) = \begin{bmatrix} 1 & -1 & -2 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$x_1,\ x_4 \rightarrow dependent\ variable$$
$$x_2,\ x_3 \rightarrow independent\ variable$$

$$x_2 = t_1$$
$$x_3 = t_2$$
$$x_4 = 0$$
$$x_1 = t_1 + 2t_2$$

$$N(A) = \{(t_1 + 2t_2,\ t_1,\ t_2,\ 0)\}$$
$$N(A) = \{(t_1,\ t_1,\ 0,\ 0), (2t_2,\ 0,\ t_2,\ 0)\}$$
$$N(A) = span\{(1, 1, 0, 0), (2, 0, 1, 0)\}$$

1.3. Row Space

→ Column space of A^T ↮ span of rows of A
→ Non zero rows of row space

Note:
→ Col rank = dim(C(A))
→ row rank = dim(R(A))
→ col rank = row rank

Question: How to find Row space of A ?
→ span of non-zero rows of RREF

Example:

$$A = \begin{bmatrix} 5 & 10 & 15 & 20 \\ 1 & 2 & 3 & 4 \\ -24 & -48 & -72 & -96 \\ -2 & -4 & -6 & -8 \end{bmatrix} R = RREF(A) = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$x_1 \rightarrow dependent$$
$$x_2,\ x_3,\ x_4 \rightarrow independent$$

$$Row\ space = \{(1, 2, 3, 4)\}$$

1.4. Left Null Space

→ Null space of A^T

→ set of all y such that $A^T y = 0$.

$$N(A^T) = \{y | A^T y = 0\} = \{y | y^T A = 0\}$$

→ For a $m \times n$ matrix A,

$$\underbrace{[y_1, \cdots, y_m]\,[A]}_{\text{Linear combination of rows leading to zero}} = [0, \cdots, 0]$$

Remark:

→ A is a $m \times n$ matrix

→ dim(C(A)) + dim(N(A)) = # columns of A = n

$$r + (n - r) = n$$

→ A^T is a $n \times m$ matrix

→ dim(C(A^T)) + dim(N(A^T)) = # rows = m

$$r + dim(N(A^T)) = m$$

$$dim(N(A^T)) = m - r$$

1.5. Important Notes

→ In general, if $A \in M_{m\times n}(\mathbb{R})$,

→ The row space is a subspace of $\mathbb{R}^n$.

→ The column space is a subspace of $\mathbb{R}^m$.

→ The RREF of A allows us to find:

→ row space (non-zero rows of REF)

→ column space (pivot column in the original matrix)

→ Null space (we know, how)

→ Nullity

→ The Dimension of the null space is called nullity

→ # Independent variables

→ Rank

→ # Non-zero rows in RREF

→ # Dependent variable

→ # pivots

→ Rank-Nullity Theorem

→ If $A \in M_{m\times n}(\mathbb{R})$, then we have:

$$rank(A) + nullity(A) = n$$

Question: Find all Fundamental subspace for the following matrices.

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

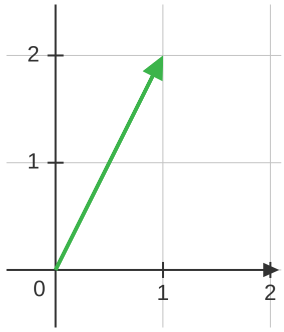
Answer:

$$R = \begin{bmatrix} & & & \\ & & & \\ & & & \end{bmatrix}$$

2. Least Squares

2.1. Norm of a vector

→ Length of a vector



$$||x||^2 = x_1^2 + x_2^2 \quad \textit{'Squared Length'}$$

$$Here, \left|\left|\begin{bmatrix}1\\2\end{bmatrix}\right|\right|^2 = 1^2 + 2^2 = 5$$

In general,

$$for\ x = \begin{bmatrix}x_1\\x_2\\\vdots\\x_n\end{bmatrix} \in \mathbb{R}^n$$

$$||x||^2 = x_1^2 + x_2^2 + \cdots + x_n^2$$

2.2. Orthogonal vectors

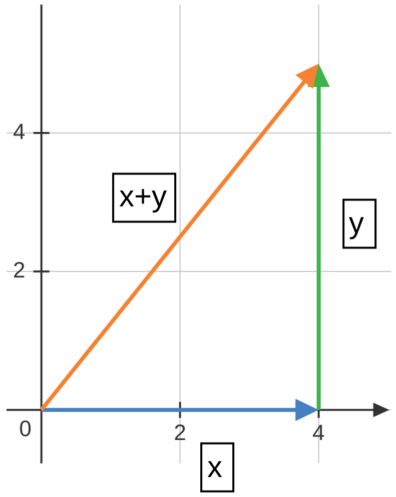
→ Dot Product or Inner Product: $x^T y$

$$x \perp y\ if\ x^T y = 0$$

$$x = \begin{bmatrix}x_1\\x_2\\\vdots\\x_n\end{bmatrix}, y = \begin{bmatrix}y_1\\y_2\\\vdots\\y_n\end{bmatrix}, then$$

$$x^T y = \sum x_i y_i$$

From Pythagoras theorem:



$$||x||^2 + ||y||^2 = ||x+y||^2$$
$$x^Tx + y^Ty = (x+y)^T(x+y)$$
$$x^Tx + y^Ty = x^Tx + x^Ty + y^Tx + y^Ty$$

$$2x^Ty = 0 \implies x^Ty = 0 \implies x \text{ is orthogonal to } y$$

Remark:

→ 0 is orthogonal to every x

→ If V = $\{v_1, \cdots, v_k\}$ is mutually orthogonal "non-trivial" set of vectors, then V is a linearly independent set.

→ Why ?

→ suppose

$$c_1v_1 + c_2v_2 + \cdots + c_kv_k = 0$$
$$\implies v_1^T(c_1v_1 + \cdots + c_kv_k) = 0$$
$$\implies c_1\, v_1^Tv_1 = 0 \text{ but } ||v_1|| \neq 0 \implies c_1 = 0$$

$$\text{Similarly } c_i = 0 \; \forall \, i$$

2.3. Orthonormal Vectors

→ $\{u, v\}$ are orthonormal if $v^T u \ = \ 0$ and $||u|| \ = \ ||v|| \ = \ 1$.

2.4. Orthogonal subspaces

U, V are orthogonal subspaces if

$$x^T y \ = 0 \ \forall \ x \ \in U, \ y \ \in V$$

Note:
→ $\{0\}$ is $\perp$ to every subspace

Example:

$$U \ = \ \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \right\} V \ = \ \left\{ \begin{bmatrix} 0 \\ 0 \\ 2 \\ 1 \end{bmatrix} \right\}$$

$$U \ \perp V$$

$$W \ = \ \left\{ \begin{bmatrix} 0 \\ 0 \\ -1 \\ 2 \end{bmatrix} \right\}$$

$$Then, \ W \ \perp U, \ W \ \perp \ V$$

2.5. Orthogonality wrt to Four Fundamentals subspace

Claim-1: $R(A) \perp N(A)$
Proof:

$$A = \begin{bmatrix} - & row_1 & - \\ - & row_2 & - \\ & \cdots & \\ - & row_m & - \end{bmatrix}_{m \times n} \quad x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \vec{0} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

$$x \in N(A)$$
$$Ax = \vec{0}$$

The above equation implies:

$$row_1 \perp x$$
$$row_2 \perp x$$
$$\vdots$$
$$row_m \perp x$$

It is pretty obvious, as we are getting zero when taking the dot product of any row to vector x.

Thus, any linear combination

$$\underbrace{(c_1 row_1 + \cdots + c_m row_m)}_{R(A)} \perp x$$

$$So,\ R(A) \perp N(A)$$

$$\leftrightarrow C\left(A^T\right) \perp N(A)$$

Claim-2: $C(A) \perp N\left(A^T\right)$
Proof: follows from claim 1

Example:

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

rank(A) = dim(C(A) = 1

$$x = \begin{bmatrix} -2 \\ 1 \end{bmatrix} \in N(A)$$

Row space: $R(A) = line\ through \begin{bmatrix} 1 & 2 \end{bmatrix}$
Null space: $N(A) = line\ through \begin{bmatrix} -2 \\ 1 \end{bmatrix}$

$$[1,2]\begin{bmatrix} -2 \\ 1 \end{bmatrix} = 0\ verifies\ R(A) \perp N(A)$$

Column space: $C(A) = \textit{line through} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$

Left Null space: $N\left(A^T\right) = y_1 + 2y_2 + 3y_3 = 0 \leftarrow \textit{is a plane}$

$$y^T A = 0$$

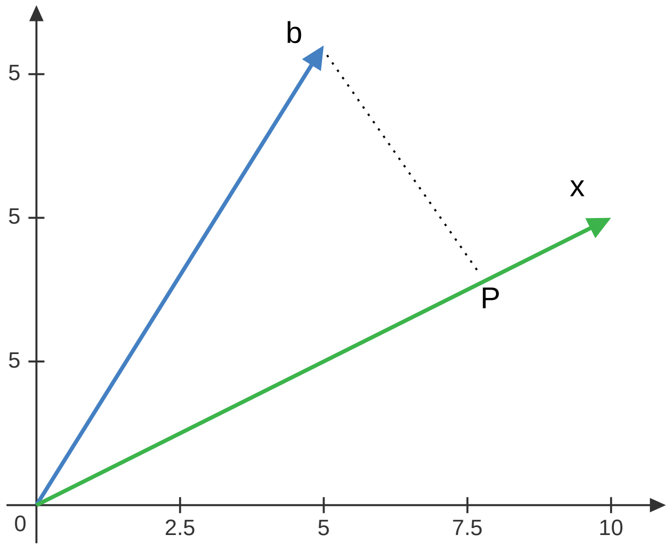
Dimension check:

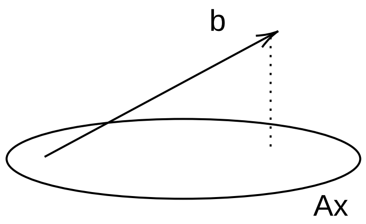
$\rightarrow dim(C(A)) + dim(N(A)) = 2 = number\ of\ columns$

$\rightarrow dim\left(C\left(A^T\right)\right) + dim\left(N\left(A^T\right)\right) = 3 = number\ of\ rows$

2.6. Projections

→ Want to project b onto the line through x, Or more generally onto the column space of a matrix A.





Question: What is need of Projection ?

→ Suppose we are given $(x_1, b_1), \quad \cdots, \ (x_n, b_n)$

$$\begin{array}{rl} & 2x = b_1 \\ e.g. & 3x = b_2 \\ & 4x = b_3 \end{array} \quad Or \quad \begin{array}{rcl} x + 2y & = & 4 \\ x + 3y & = & 5 \\ 2x + 4y & = & 6 \end{array}$$

These are inconsistent $\Longrightarrow$ No solution that satisfies this system of equations

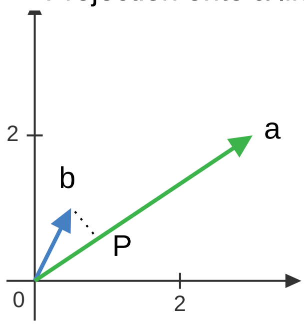
Matrix View: $Ax = b$
Inconsistent if

$$b \notin C(A)$$

In such situation, if makes sense to project b onto C(A)).

Question: Given a basis for a subspace S (e.g. spanned by columns of A), if there an easy way to calculate the projection P of b onto S ?

→ Projection onto a line



$$P = \hat{x}a$$
$$e = b - p = b - \hat{x}a$$
$$e \perp a$$

$$(b - \hat{x}a) \perp a$$

$$a\,(b-\hat{x}a) \;=\; 0 \; leading\ to \;\; \hat{x} \;=\; \frac{a^T b}{a^T a} \;\Longrightarrow\; P \;=\; \hat{x}a \;=\; \left(\frac{a^T b}{a^T a}\right)a$$

Cauchy-Schwarz inequality:

$$||e||^2 \;=\; ||b-p|| \;\geqslant\; 0$$

$$||b-\left(\frac{a^T b}{a^T a}\right)a||^2 \;=\; b^T b \;-2\frac{\left(a^T b\right)^2}{a^T a} \;+\; \left(\frac{a^T b}{a^T a}\right)^2 a^T a$$

$$= \;\frac{\left(b^T b\right)\left(a^T a\right) - \left(a^T b\right)^2}{\left(a^T a\right)} \;\geqslant\; 0$$

$$= \;\left(b^T b\right)\left(a^T a\right) \;\geqslant\; \left(a^T b\right)^2$$

$$= \;|a^T b| \;\leqslant\; ||a||\;||b|| \;\rightarrow\; cauchy\ schwarz\ inequality$$

Projection Matrix :

$$P \;=\; \left(\frac{a^T b}{a^T a}\right)a \;=\; \left(\frac{aa^T}{a^T a}\right)b$$

$$Let\ \mathbb{P} \;=\; \frac{aa^T}{a^T a}\,.\;\; Then$$

$$Projection\ of\ b\ onto\ a\ is$$

$$\mathbb{P}b$$

To project any vector b, just left multiply by the projectio matrix $\mathbb{P}$.

Example:

$$a \;=\; \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

$$\mathbb{P} \;=\; \left(\frac{aa^T}{a^T a}\right) \;=\; \frac{1}{a^T a}\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \;=\; \frac{1}{3}\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \;=\; \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix}$$

Observe that

→ $\mathbb{P}$ is symmetric

→ $\mathbb{P}^2 = \mathbb{P}$ i.e. $\mathbb{P}^2 b = \mathbb{P}b$
→ $C(\mathbb{P}) = \textit{line through a}$
→ $N(\mathbb{P}) = \textit{plane orthogonal to a}$
→ rank r($\mathbb{P}$) $= 1$

Note:
→ $\mathbb{P}^2 = \mathbb{P}$ implies that if you again left multiply a with Projection matrix then you will get a.
→ The idea is $\mathbb{P}b$ is already on the line through a. So, another round of projection won't change it.

Example-2:

$$a = \begin{bmatrix} 2 \\ 2 \\ 2 \end{bmatrix}$$

$$\mathbb{P} = \left(\frac{aa^T}{a^Ta}\right) = \frac{1}{a^Ta}\begin{bmatrix} 4 & 4 & 4 \\ 4 & 4 & 4 \\ 4 & 4 & 4 \end{bmatrix} = \frac{1}{12}\begin{bmatrix} 4 & 4 & 4 \\ 4 & 4 & 4 \\ 4 & 4 & 4 \end{bmatrix} = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix}$$

2.7. Least Squares and Projections onto a subspace

Consider this system of equations:

$$2x = b_1$$
$$3x = b_2$$
$$4x = b_3$$

when this system is solvable: if b is on the line through $\begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix}$

→ Suppose we have a vector "b" that leads to an "inconsistent" system.
→ We could pick a subset of equations & solve it exactly
  → The Problem with approach: large error in some inputs & no error in others.

REASONABLE ALTERNATIVE: minimise the avg error.

$$E^2 = (2x - b_1)^2 + (3x - b_2)^2 + (4x - b_3)^2$$

→ We want to minimize the sum of squares E^2.

$$\frac{dE^2}{dx} = 0$$

$$2(2x - b1)(2) + 2(3x - b_2)(3) + 2(4x - b_3)(4) = 0$$

$$2[2(2x - b_1) + 3(3x - b_2) + 4(4x - b_3)] = 0$$

$$\textit{leading to}$$

$$\hat{x} = \frac{2b_1 + 3b_2 + 4b_3}{2^2 + 3^2 + 4^2} = \frac{a^T b}{a^T a} \; with \; a = \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix}$$

Bottom line: Taking derivative & finding the minima turns out to be the same as performing a projection.

Projection onto a subspace

$$Ax = b, \; A \; is \; m \times n, \; m > n$$

want: Projection of b onto column space C(A).

$$S = span(\; columns \; of \; A \;)$$

Projection of b onto S is $P = A\hat{x}$

$$Orthogonal \; vector \; e = b - p = b - A\hat{x}$$

Q: How to find $\hat{x}$?
→ Observe e $\perp$ every vector in C(A) .
→ Recall that $C(A) \perp N\left(A^T\right)$, *i.e.* $N\left(A^T\right)$ is the orthogonal complement of C(A). i.e. every vector in C(A) is orthogonal to every vector in $N\left(A^T\right)$ & any given vector is in either $C(A) \; or \; N\left(A^T\right)$.

→ Where does e belong ?

$$e \in N\left(A^T\right) \implies A^T e = 0 \implies A^T(b - A\hat{x}) = 0$$

leading to

$$A^T A\hat{x} = A^T b$$

equation to solve to obtain the projection of b onto $C(A)$.

Note: Even if $Ax = b$ is not solvable, $A^T A\hat{x} = A^T b$ has a solution.

Alternative Route to above equation:

$$A = \begin{bmatrix} | & | & & | \\ a_1 & a_2 & \cdots & a_n \\ | & | & & | \end{bmatrix} \quad \begin{matrix} a_1^T e = 0 \\ \vdots \\ a_n^T e = 0 \end{matrix} \quad \leftrightarrow \quad \begin{matrix} a_1^T(b - A\hat{x}) = 0 \\ \vdots \\ a_n^T(b - A\hat{x}) = 0 \end{matrix}$$

$$\begin{bmatrix} a_1^T \\ \vdots \\ a_n^T \end{bmatrix} \left[b - A\hat{x} \right] = 0$$

$$A^T(b - A\hat{x}) = 0$$

$$\textit{leading to}$$

$$A^T A\hat{x} = A^T b$$

Bottom line: $A^T A\hat{x} = A^T b$ leads to that $\hat{x}$ that minimises $||Ax - b||^2$

This is connection of projections to least squares.

Remarks:

- Suppose columns of A are linearly independent (l.i.) Then,

$$A^T A \text{ is invertible}$$

→ $A^T A$ is square and symmetric

→ Solving $A^T A\hat{x} = A^T b$ when $A^T A$ is invertible

→ $\hat{x} = \left(A^T A\right)^{-1} A^T b$

→ Projection $\mathbb{P} = A\hat{x} = A\left(A^T A\right)^{-1} A^T b$

- $b \in C(A)$ *i.e.*, $b = Ax$

$$P = A\left(A^T A\right)^{-1} A^T b = A\left(A^T A\right)^{-1} A^T A x = AIx = Ax = b$$

- $b \in N\left(A^T\right)$

$$P = A\left(A^T A\right)^{-1} A^T b = 0 \; since \; A^T b = 0$$

- A is square & invertible → $C(A) = \mathbb{R}^n$

$$P = A\left(A^TA\right)^{-1}A^Tb = AA^{-1}\left(A^T\right)^{-1}A^Tb = b$$

- A is rank one i.e., $A = \begin{bmatrix} | \\ a \\ | \end{bmatrix}$ Then,

$$\hat{x} = \frac{a^Tb}{a^Ta}$$

coincides with what we derived earlier for projection onto a line.

Projection matrix: $P = A\left(A^TA\right)^{-1}A^T$

- Symmetric

$$\mathbb{P}^T = \mathbb{P}$$

$$\left(A\left(A^TA\right)^{-1}A^T\right)^T = \left(A^T\right)^T\left(\left(A^TA\right)^{-1}\right)^T(A)^T$$

$$= A\left(A^TA\right)^{-1}A^T = \mathbb{P}$$

The middle term comes as it as because A^TA is a symmetric matrix so it is going to be itself.

- $\mathbb{P}^2 = \mathbb{P}$

$$A\left(A^TA\right)^{-1}A^TA\left(A^TA\right)^{-1}A^T$$

$$= A\left(A^TA\right)A^T = \mathbb{P}$$

So, Projection matrix is symmetric & satifies $\mathbb{P}^2 = \mathbb{P}$

The converse is also true.

$$If\ \mathbb{P}^2 = \mathbb{P}\ \&\ P^T = \mathbb{P}$$

$$then\ \mathbb{P}\ is\ a\ projection\ matrix.$$

$\mathbb{P}b$ = projection of b onto the column space of $\mathbb{P}$.

Examples: Least Squares

Simple case: One dimension

Dataset: $(x_1,\ b_1),\ \ \cdots,\ (x_m,\ b_m)$

$$b_i\ =\ \theta x_i + \theta' \tag{1}$$

Here, $\theta\ is\ scaling\ factor$ and θ' is constant offset and the above equation is a Linear fit

We want to find such a fit for all $i\ =\ 1,\ \ \cdots,\ m$ i.e. for all data points

System of equations 1 is equivalent to

$$\begin{bmatrix} x_1 & 1 \\ \vdots & \vdots \\ x_m & 1 \end{bmatrix} \begin{bmatrix} \theta' \\ \theta'' \end{bmatrix} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}$$

This is like

$$A\theta\ =\ b,\ where\ \theta\ =\ \begin{bmatrix} \theta' \\ \theta'' \end{bmatrix}$$

$A\theta\ =\ b$ may be inconsistent.

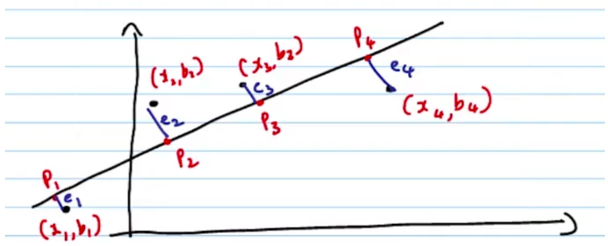
Least squares approach:

$$minimize \qquad E^2\ =\ ||b - A\theta||^2$$

$$=\ (b_1 - \theta' x_1 - \theta'')^2 +\ \cdots\ + (b_m - \theta' x_m - \theta'')^2$$

$$(\widehat{\theta}',\ \widehat{\theta}'')\ =\ argmin\ ||b - A\theta||^2$$

We want to find the line that minimises the sum of the square distance.



If the point $b_1,\ b_2,\ b_3,\ b_4$ lie on a line Then,

$$P_1\ =\ b_1,\ P_2\ =\ b_2,\ P_3\ =\ b_3,\ P_4\ =\ b_4$$

$$\&\ E^2\ =\ 0$$

$$\&\ Ax\ =\ b\ can\ be\ solved.$$

If not then minimise the square error

$$min \quad E^2 \; = \; ||A\theta - b||^2$$

Example:

$$A\theta \; = \; b$$

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

Is $A\theta \; = \; b$ consistent ? or does b belong to C(A) ?
→ We can check this by using Gaussian elimination

$$[A|b] \; = \; \begin{bmatrix} -1 & 1 & 1 \\ 1 & 1 & 1 \\ 2 & 1 & 3 \end{bmatrix}$$

$$[R|c] \; = \; \begin{bmatrix} 1 & -1 & -1 \\ 0 & 1 & 1 \\ 0 & 0 & 2 \end{bmatrix}$$

Since last entry of last column has non-zero value this system of equations cannot be solved. (Inconsistent or b does not belong to C(A))

First we have to cheeck if given a system is it even consistent ?
→ If it is already consistent, there is no point going and solving an alternative equation, the original set of equations is already solvable.

So, Now we will do Least Squares:

$$A^T A\widehat{\theta} \; = A^T b, \quad where\; \widehat{\theta} \; = \begin{bmatrix} \widehat{\theta}' \\ \widehat{\theta}'' \end{bmatrix}$$

$$A^T A \; = \; \begin{bmatrix} 6 & 2 \\ 2 & 3 \end{bmatrix} and\; A^T b \; = \; \begin{bmatrix} 6 \\ 5 \end{bmatrix}$$

$$\begin{bmatrix} 6 & 2 \\ 2 & 3 \end{bmatrix}\begin{bmatrix} \widehat{\theta}' \\ \widehat{\theta}'' \end{bmatrix} \; = \; \begin{bmatrix} 6 \\ 5 \end{bmatrix}$$

Solving this system of equations we have:

$$\widehat{\theta}' \; = \; \frac{4}{7} \quad and \quad \widehat{\theta}'' \; = \; \frac{9}{7}$$

$$\widehat{\theta} \; = \; \begin{bmatrix} 4/7 \\ 9/7 \end{bmatrix}$$

The Best Line (in the least square sense through the given data is:

$$\frac{4}{7}x + \frac{9}{7}$$

What are the Projections ?

$$P_1 = \frac{4}{7}(-1) + \frac{9}{7} = \frac{5}{7}$$

$$P_2 = \frac{4}{7}(1) + \frac{9}{7} = \frac{13}{7}$$

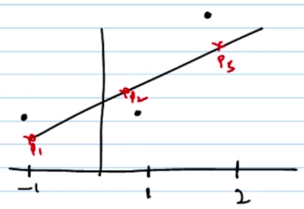
$$P_3 = \frac{4}{7}(2) + \frac{9}{7} = \frac{17}{7}$$

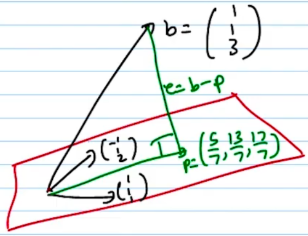
The Original data is not on a line, so $E^2 > 0$

$$E^2 = ||b - A\widehat{\theta}||^2 = ||e||^2$$

$$e = \left[1 - \left(\frac{-4}{7} + \frac{9}{7}\right)\right], \left[1 - \left(\frac{4}{7} + \frac{9}{7}\right)\right], \left[3 - \left(\frac{8}{7} + \frac{9}{7}\right)\right]$$

$$e = \frac{2}{7}, \frac{-6}{7}, \frac{4}{7}$$





$e = \left(\frac{+2}{7}, \frac{-6}{7}, \frac{4}{7}\right)$

$e \perp (-1, 1, 2)$ and $e \perp (1, 1, 1)$

IMPORTANT POINTS OF THIS WEEK

Least Squares Method

Where are we going to use this Least squares method → to find the line of best fit for a set of data.

How ?

→ By minimising the sum of the offsets or residuals of points from the plotted line.
→ Represents general trend of the data
→ Used for regression analysis

1. The column space of the projection matrix of a vector v is → a line passing through v.
2. The Null space of the projection matrix of a vector v is → a plane orthogonal to v.

Questions

1. Let $S = \{(1, 2, 4, 0),\ (-2, 3, -1, 0),\ (0, 2, 6, -1)\}$. Which pair(s) of vectors in this given set are orthogonal ?

→ Lets assume:

$$\begin{aligned} a &= (1, 2, 4, 0) \\ b &= (-2, 3, -1, 0) \\ c &= (0, 2, 6, -1) \end{aligned}$$

$$\begin{aligned} a.b &= 0 \\ a.c &\neq 0 \\ b.c &= 0 \end{aligned}$$

So, only (a,b) and (b,c) are orthogonal pairs.

2. (a) Find the projection matrix for $a = [2, -1, 2, 3]^T$

(b) Obtain the projection of $b = [1, 3, -2, 5]^T$onto a and compute the error.

Answer: (a)

$$\mathbb{P} = \frac{aa^T}{a^Ta}$$

$$aa^T = \begin{bmatrix} 4 & -2 & 4 & 6 \\ -2 & 1 & -2 & -3 \\ 4 & -2 & 4 & 6 \\ 6 & -3 & 6 & 9 \end{bmatrix} and\ a^Ta = 18$$

$$So,\ \mathbb{P} = \begin{bmatrix} 2/9 & -1/9 & 2/9 & 1/3 \\ -1/9 & 1/18 & -1/9 & -1/6 \\ 2/9 & -1/9 & 2/9 & 1/3 \\ 1/3 & -1/6 & 1/3 & 1/2 \end{bmatrix}$$

Answer (b) :

$$P = \mathbb{P}.b = \begin{bmatrix} 10/9 \\ -5/9 \\ 10/9 \\ 5/9 \end{bmatrix}$$

Once we have obatined the projection we can compute the error.

$$e = b - p$$

$$= \begin{bmatrix} 1 \\ 3 \\ -2 \\ 5 \end{bmatrix} - \begin{bmatrix} 10/9 \\ -5/9 \\ 10/9 \\ 5/9 \end{bmatrix} = 1/9 \begin{bmatrix} -1 \\ 32 \\ 28 \\ 30 \end{bmatrix}$$

3. Build a model that studies the relationship btw x and y given in the table using least squares method.

x	1	2	3	4	5
y	2.6	3.4	7.1	10.2	13.5

Solution: Least Square Method: $A^T A\hat{x} = A^T b$ where $\hat{\theta} = \begin{bmatrix} \hat{\theta}' \\ \hat{\theta}'' \end{bmatrix}$

The Matrix A is a matrix containing x_i and last column is of ones.

$$A = \begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 3 & 1 \\ 4 & 1 \\ 5 & 1 \end{bmatrix} \; b = \begin{bmatrix} 2.6 \\ 3.4 \\ 7.1 \\ 10.2 \\ 13.5 \end{bmatrix}$$

$$A^T A = \begin{bmatrix} 55 & 15 \\ 15 & 5 \end{bmatrix} A^T b = \begin{bmatrix} 139 \\ 36.8 \end{bmatrix}$$

$$\begin{bmatrix} 55 & 15 \\ 15 & 5 \end{bmatrix} \begin{bmatrix} \hat{\theta}' \\ \hat{\theta}'' \end{bmatrix} = \begin{bmatrix} 139 \\ 36.8 \end{bmatrix}$$

Solving the above system of linear equations

$$\hat{\theta}' = 2.86$$

$$\hat{\theta}'' = -1.22$$

So, the best fit line is:

$$y = 2.86x - 1.22$$