Loc:

1

Recap:

what does this magical formula do?

A $P = A(A^TA)^{-1}A^T$

- it is suppose to produce a projection (matrix)

if I multiply P by b, I am suppose to project that vector b to the nearest point in column space

: (Suppose) if b in column space $Pb = b$, then what do I get when I apply Pb?

get b again

- Suppose that vector (b) is perpendicular to column space, imagine column space is plane and b is sticking straight up perpendicular to it.

- what is the nearest point in column space to "$\underline{b}$"

$\therefore$ $Pb = 0$

if $b \perp$ column space $Pb = 0$

those are the two extremes.

those are guys (vectors) which are in the null space of $A^{(T)}$

$\therefore$ if I am in nullspace of A^T and multiply A (big formula $\times b$) will get zero (0)

Recaps.

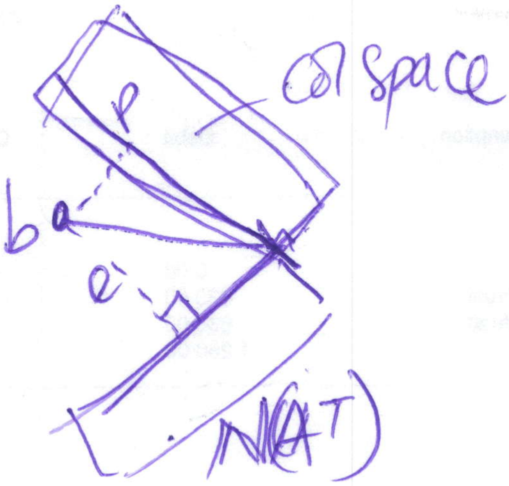
if b in column space, why $Pb = b$

b multiply times big formula (A)

$\hookrightarrow$ b are also in the column space

column space has form Ax

$\therefore$ $A(A^TA)^{-1}A^TAx = b$



∴ b typical vector, projection it to p, at same time we find other part e

$$p + e = b$$

Projection piece + error piece add up to the original b.

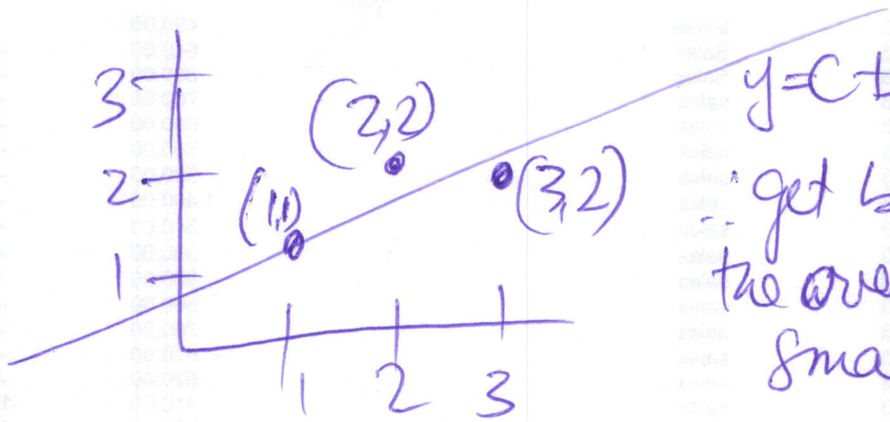
$$p \Rightarrow Pb$$

$$e \Rightarrow (I-P)b$$

also projection onto the $\perp$ space

So how do we do it?

∴ Find best Straight line:



$y = C + Dt$

∴ get best line to make the overall error as small as possible.

∴ What is that overall error? (minimise)

Now we construct our equations:

$$C + D = 1$$
$$C + 2D = 2$$
$$C + 3D = 2.$$

– they don't have a solution. But we have best solution.
⇒ what best solution?

$$Ax = b$$

3×2.

$$\underbrace{\begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} C \\ D \end{bmatrix}}_{x} = \underbrace{\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}}_{b}$$

- I do have 2 independent column,
- which is basis for Column space

, But column space does not include vector

= 1 } when calculating I will
= 2 } make error here
= 3 }

square and add up those errors

Sum of Squares.

Looking for least squares solution.

Errors are difference between:

$\|Ax - b\|^2 = \|e\|^2$

What is error vector in picture?

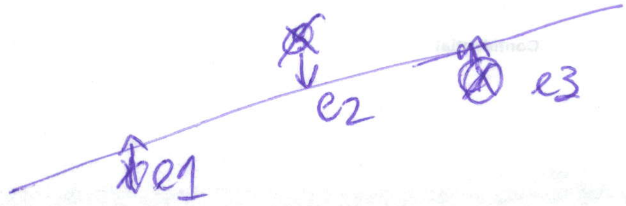
two pictures here:

1. 3 points and line

what are errors here?

vertical distance of one point to line

6



$$\therefore = e_1^2 + e_2^2 + e_3^2$$

∴ doing regression here! (linear)

∴ Using the Sum of Squares as measure of Error.

But what if one datapoint is way off. would this line be best fit? No

there will be a giant error

& squared it will be Big

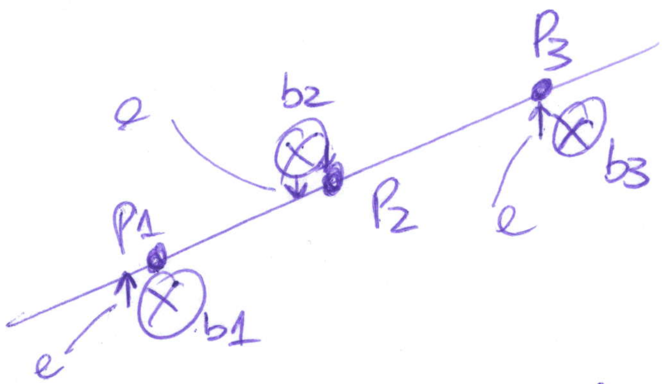
∴ Called ~~as the~~ outlier →

(Suppose we do not have this point)

[illegible]

2

e is distance



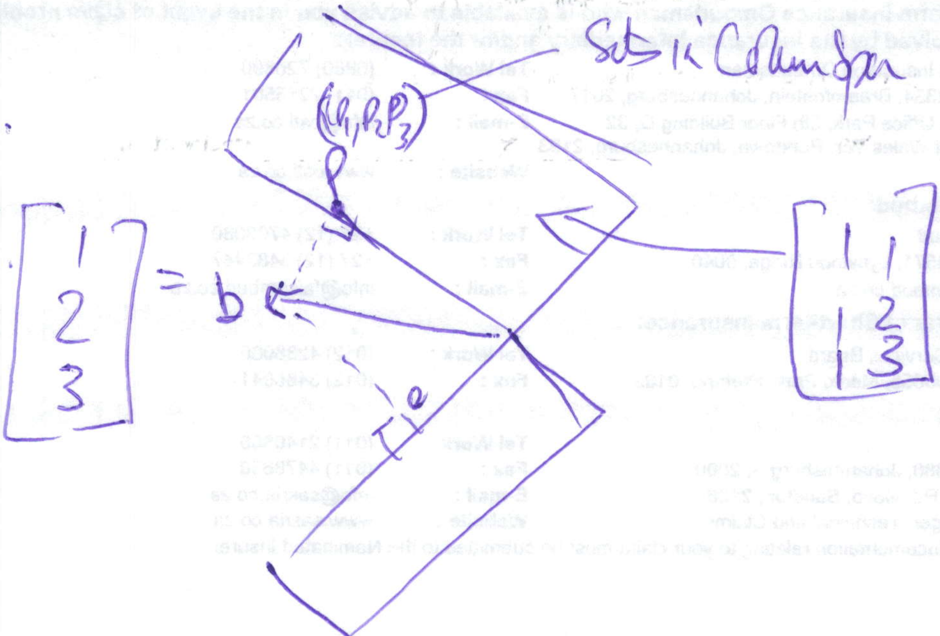
– in picture, what are the points really on the line?

∴ what are P_1, P_2, P_3?

⇒ (and we also have b_1, b_2, b_3)

∴ Instead of b_1, b_2, b_3, we can put P_1, P_2 and P_3 in Equation. How would it affect the three equations. If can solve it.

∴ P_1, P_2 and P_3 are in the column space. They are a combination of columns of Matrix A



Find $\hat{x} = \begin{bmatrix}\hat{C}\\ \hat{D}\end{bmatrix}$, and find P?

$\hat{}$ to denote they are estimates

⇒ best line, not perfect line

(8)

$\therefore \hat{x}: \quad A^T A \hat{x} = A^T b$ } most important equation in statistics 28/1/19

or in Estimation

⇒ where you have noise, etc

⇒ wherever you fitting things

$A^T A$

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} = \begin{bmatrix} 3 & 6 \\ 6 & 14 \end{bmatrix}$$

⇒ expect to be:

⇒ symmetric

⇒ invertible.

$A^T b$

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{bmatrix} \left[\begin{array}{cc|c} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 1 & 3 & 2 \end{array}\right] = \left[\begin{array}{cc|c} 3 & 6 & 5 \\ 6 & 14 & 11 \end{array}\right]$$

⇒ Equations:

$$3C + 6D = 5$$

$$6C + 14D = 11$$

Solved for C & D

~~Key~~

Aside: lets try to find the Equations from the minimise:

$$\text{Minimise } \|Ax-b\|^2 = \|e\|^2$$

$$= e_1^2 + e_2^2 + e_3^2$$

take square of something then take derivative, then get something linear.

$$= (C+D-1)^2 + (C+2D-2)^2 + (C+3D-2)^2$$

(using Calculus), partial derivates →

(and we'll get the equation.

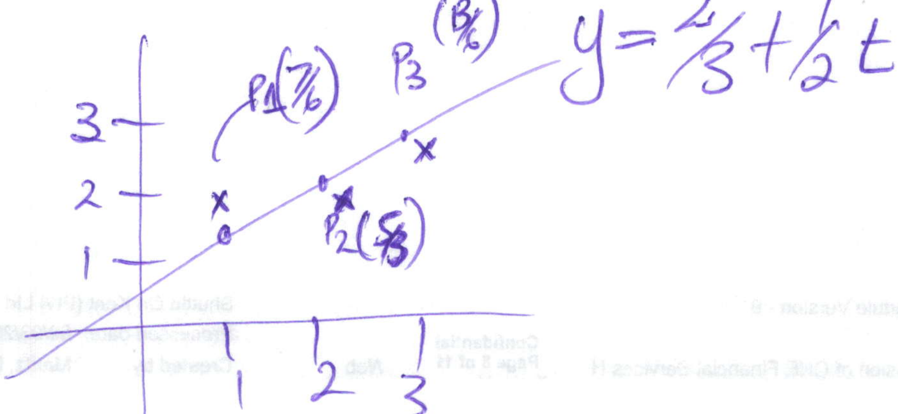
Back

$$3C + 6D = 5$$
$$6C + 14D = 11$$

Solve using Back Substitution:

$$D = 1/2, \quad C = 2/3.$$

Best line $2/3 + 1/2\, t$



$e_1 = -1/6$

$e_2 = +2/6$

$e_3 = -1/6$

$\therefore b = p + e$

$$\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 7/6 \\ 10/6 \\ 13/6 \end{bmatrix} + \begin{bmatrix} -1/6 \\ 2/6 \\ -1/6 \end{bmatrix}$$

- what else about these two vectors.

$\therefore$ they are $\perp$

~~, not only are they $p \perp$~~

(projection)

, not only $\perp$ to column space, ~~but~~

(which other vector is it $\perp$ to?)

[~~p~~ $\perp$ to other stuff]

that's in column space.

$\therefore$ another vector in column space?

$\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ and e is suppose to be

$\perp$ to it

$\therefore P = A\hat{x}$

If A has independent columns, then A^TA is invertible. [to prove x must be 0]

Suppose $A^TAx = 0$

∴ What do I want to prove now?

∴ what conclusion do I want to reach

~~x must be~~ 0.

⇒ to show x must be zero.

Two ways to do! (Prove)

$A^TAx = 0$

∴ take the dot product of both sides with x

multiply both sides by x^T

∴ $x^TA^TAx = 0$

(what can we conclude from that equation?

$\Theta = (Ax)^T(Ax) \Rightarrow [x=0]$ (below)

vector has to be zero

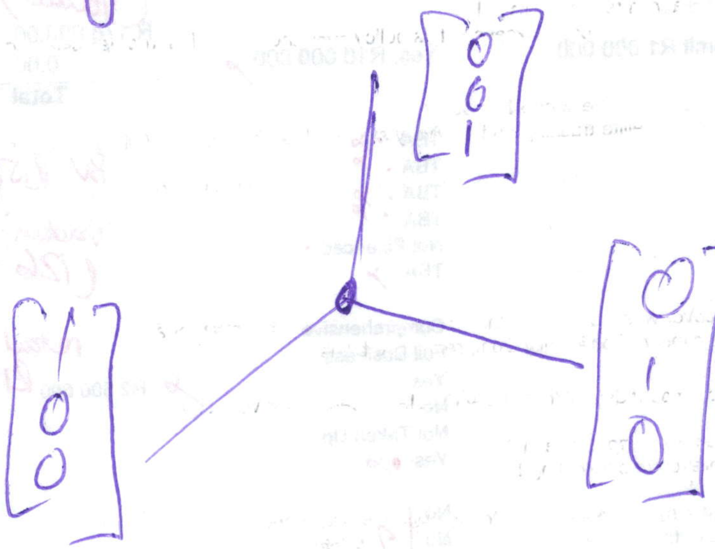
$Ax = 0$ ← and, then what? (B)

Now we can use hypothes [A has independent cols]

(B) $x = 0$

Columns are definitely independent if they are $\perp$ unit vectors, 'orthonormal' vectors

like vectors:



these vectors are unit vectors

— they $\perp$

— and independent.