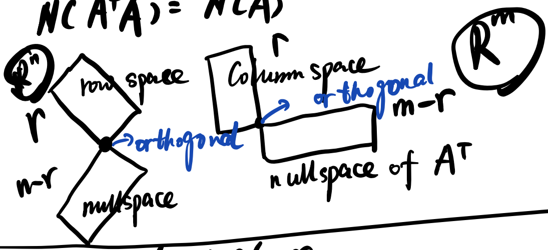
Linear Algebra lecture 14 (Lecture 13 is quiz 1. and finished)

Chapter 4.

Orthogonal vectors + subspaces

nullspace $\perp$ row space

$N(A^TA) = N(A)$



Orthogonal vectors

$x^Ty = 0$

$\|x\|^2 + \|y\|^2 = \|x+y\|^2 \Leftrightarrow \cancel{x^Tx} + \cancel{y^Ty} = (x+y)^T(x+y)$

x^Tx

$0 = \cancel{x^Tx} + \cancel{y^Ty} + 2x^Ty + \cancel{y^Tx}$

Subspace S is orthogonal to subspace T.
means: every vector in S is orthogonal to every vector in T.

"blackboard subspace" $\times$ "floor subspace"

row spce is orthogonal to nullspace.

why? $Ax=0$. $\begin{bmatrix} \text{row1 of } A \\ \text{row2 of } A \\ \cdots \\ \text{rowm of } A \end{bmatrix} \begin{bmatrix} x \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \cdots \\ 0 \end{bmatrix}$

$c_1(\text{row}_1)^T x = 0$
$+ \; c_2(\text{row}_2)^T x = 0$

$\Rightarrow (c_1(\text{row}_1)^T + c_2\,\text{row}_2^T + \ldots)\,x = 0$

nullspace and row space are $\boxed{\text{orthogonal complements}}$ in R^n.

Nullspace contains all vectors $\perp$ row space.

Coming: $Ax = b$ when There is no solution.
"solve"

$\boxed{A^TA}$
$n\times m \;\; m\times n$
$n\times n$

Symmetric
$(A^TA)^T = A^TA$

rank(AB) $\leq$ min (rank(A), rank(B)).

$A^TA\hat{x} = A^Tb$

① $N(A^TA) = N(A)$

② rank of A^TA = rank of A

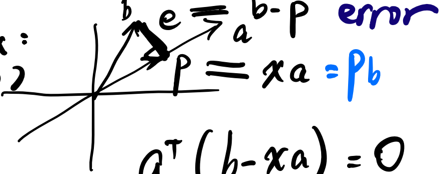
③ A^TA is invertible exactly if
A has indep. columns ($N(A) = 0$)

Linear Algebra Lecture 15

{Projections!
Least squares.

PROJECTION MATRIX

Ex: (1,2 D)



$e = b-p$ error

$p = xa = Pb$

$\underline{a^T(b-xa)=0}$

$x a^T a = a^T b$

$\underline{x = \frac{a^Tb}{a^Ta}}$ → $\underline{p = ax}$

$p = ax = \left(a\ \frac{a^Tb}{a^Ta}\right)$

projection $\underline{p} = Pb$ → $\boxed{P = \frac{aa^T}{a^Ta} \quad \text{PROJECTION MATRIX}}$

[For P:]: $\left(\frac{1}{a^Ta} \cdot \underline{aa^T}\right)$

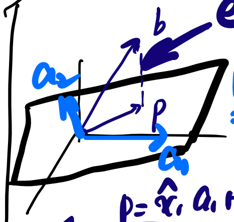
- $C(P)$ = line through a.
- $\text{rank}(P) = 1$.
- * $P^T = P$. (symmetric)
- $\underline{P^2 = P}$

• Why project?? (n-D)

Because $Ax=b$ may have no solution,

→ Solve $A\hat{x} = p$ instead ($\hat{x}$ → hat the best solution.)

p: $\boxed{\text{proj}}$ of b $\boxed{\text{onto}}$ col. space. (closest!!!)



$e = b-p$ is perpendicular to the plane.

plane of a_1, a_2
$=$ column space of $\underline{A}$: $[a_1\ a_2]$

$p = \hat{x}_1 a_1 + \hat{x}_2 a_2$

$\exists \hat{x}$: $p = A\hat{x}$

$A^TA\hat{x} = A^Tb$

⇒ $p = A\hat{x}$. Find $\hat{x}$, then we get p. (and the Proj. Matrix P), then we can can a solution which is close to the $Ax=b$ mostly.

• $\boxed{\text{Key}}$ equation: $\underset{(e)}{b - A\hat{x}}$ is perp. to plane. use it ~

⇓

$\underbrace{{}^{1}\ a_1^T(b-A\hat{x}) = 0. \quad {}^{2}\ a_2^T(b-A\hat{x}) = 0.}$

put these equations together as a matrix equation:

$$\begin{bmatrix} a_1^T \\ a_2^T \end{bmatrix}(b - A\hat{x}) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

⇒ $A^T(\overset{e}{b-A\hat{x}}) = 0$. ⇒ e in $N(A^T)$

$e \perp C(A)$ YES!

then, $\boxed{A^TA\hat{x} = A^Tb}$

not as the 1-D, they are matrices instead of numbers

So, $\hat{x} = (A^TA)^{-1}A^Tb \xrightarrow{1D} \frac{aa^T}{a^Ta}$

and, $p = A\hat{x} = \underline{A(A^TA)^{-1}A^Tb}$ ✗ $AA^{-1}(A^T)^{-1}A^T = I$

A is not square ... it doesn't have an inverse.

and, matrix $P = A(A^TA)^{-1}A^T$ → symmetric

$\boxed{\begin{matrix} P^T = P \\ P^2 = P \end{matrix}}$

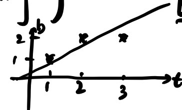
$A(A^TA^{-1})\cancel{A^TA(A^TA)^{-1}}A^T$

if A is square. Column space of A is R^n — the whose space. (different from subspace) so b is definitely in it and $P = I$ for sure.

The analysis before is based on the given "A". we use the connection (all ⊥) to get p and its $\hat{x}$. ~~But in the question of least square, we don't know the A, by ourselves~~ the A is what we need to compute.

Least squares

Fitting by a line



(1,1). (2,2). (3,2)

$\underline{b = C + Dt}$ find the best line

$$\begin{cases} C+D=1 \\ C+2D=2 \\ C+3D=2 \end{cases} \quad \underset{A}{\begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix}} \underset{x}{\begin{bmatrix} C \\ D \end{bmatrix}} = \underset{b}{\begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}}$$

→ $\underline{A^TA\hat{x} = A^Tb}$!

→ .... So there must be more equation to be listed by us. let $\sum e_i^2$ be min.!

Linear Algebra Lecture 16

Projections.

Least squares and best straight line.

If b in column space $Pb = b$.

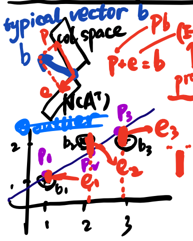
If $b \perp$ column space $Pb = 0$.

b is in $N(A^T)$! $p = Pb = A(A^TA)^{-1}\underline{A^Tb} = 0$.

$b = Ax$ form exsits. $Pb = A(A^TA)^{-1}A^TAx = Ax = b$

The average vector has a component P in the column space and a component perpendicular to it, and what the prediction does is to ~~kill~~ kill the later part and preserve the former part.

[be got no component in the column space]



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

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

pic.1 : 3 points & best line & errors (in "blackboard plane")

pic.2 : ⇒ Showing the vectors

$\llcorner$ in this pic., C and D never show up directly, and they are the combination of the columns of A that gives P.

Minimize $\|A\hat{x} - b\|^2 = \|e\|^2$ "linear regression" system.

$= e_1^2 + e_2^2 + e_3^2$ "using this sum of squares as the measure of errors"

fitting things by a few parameters

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

$\boxed{A^TA\hat{x} = A^Tb}$ normal eqns

[venture to call it the most important equation in statistics and in estimation]

$P = A\hat{x}$

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} = \begin{bmatrix} 3 & 6 \\ 6 & 14 \end{bmatrix} \quad \begin{bmatrix} 5 \\ 11 \end{bmatrix}$$

take on b, A^Tb

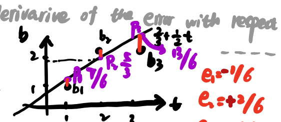
$3C + 6D = 5$

$6C + 14D = 11$

$\Rightarrow \boxed{\begin{matrix} D = \frac{1}{2} \\ C = \frac{2}{3} \end{matrix}}$ Best line $\frac{2}{3} + \frac{1}{2}t$

Here we want to find these equations from calculus:

$(C+D-1)^2 + (C+2D-2)^2 + (C+3D-2)^2$ take partial derivatives with respect to ... is the derivative of the err with respect to C being zero ---- D ----



$e_1 = -1/6$

$e_2 = +2/6$

$e_3 = -1/6$

$(e + p = b)$

$b = p + e$

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

$p^Te = 0$

$v^Te = 0$ ($\forall v$ in column space)

(this is what makes sure the existence of P, on the other words, making sure we can do all projections at the same times.

- If A has indep. columns, then A^TA is invertible. ∴ To PROVE

Suppose $A^TAx = 0$. x must be 0

(A has indep cols) $\Rightarrow x = 0$

$\boxed{\text{IDEA}}$ $x^TA^TAx = 0 = (Ax)^TAx \Rightarrow Ax = 0$ (the length squared)

of course: $Ax = 0 \rightarrow A^TAx = 0$

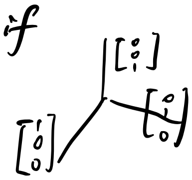
$N(A^TA) = N(A)$

∴ A^TA and A share the same nullspace,

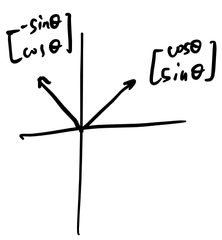
$\text{rank}(A^TA) = \text{rank}(A)$

- Columns definitely independent if they are perp. unit vectors.

orthonormal vectors



$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{I}$$



Linear Algebra Lecture 17

Orthogonal basis $q_1, \dots, q_n$	Orthonormal Vectors
Orthogonal matrix (called) Q (when square)	$q_i^T q_j = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j \end{cases}$
Gram-Schmidt $A \to Q$	

$$Q = \begin{bmatrix} q_1 & \cdots & q_n \end{bmatrix} \quad Q^TQ = \begin{bmatrix} q_1^T \\ \vdots \\ q_n^T \end{bmatrix} \begin{bmatrix} q_1 & \cdots & q_n \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ & 1 & & \\ 0 & & \ddots & \\ & & & 1 \end{bmatrix}$$

$$\underline{Q^T \quad Q \quad = \quad I} \checkmark$$

(and we can get $|Q|^2 = 1$)

If Q is square, then $Q^TQ = I$ tells us $Q^T = Q^{-1}$ (and if "square" and "col. independent" → columnspace is the whole space!)

Examples

permutation $Q = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} = I.$

$Q = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}, \ \theta = \frac{\pi}{4}$

$Q = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$

$Q = \frac{1}{3} \begin{bmatrix} 1 & -2 & 2 \\ 2 & -1 & -2 \\ 2 & 2 & 1 \end{bmatrix}$

$Q = \frac{1}{2} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}$

Adhemar (form)

independent ⇐ orthogonal basis headed off all at 90 degrees

there's no combination that gives zero.

Q has orthonormal columns.

Project onto its column space

$\to P = Q(Q^TQ)^{-1}Q^T = \underline{QQ^T}$ ($= I$ if Q is square)

symmetric

$(QQ^T)(QQ^T) = QQ^T \checkmark$

the normal equation: $A^TA\hat{x} = A^Tb$. Now A is Q.

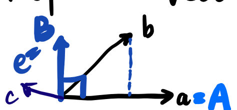
$\to Q^TQ\hat{x} = Q^Tb$

$\Rightarrow \underline{\hat{x} = Q^Tb}$

$\boxed{\hat{x}_i = q_i^T b} \bigstar$

Gram-Schmidt

independent Vectors a, b → orthogonal $A \perp B$ C → orthonormal $q_1 = \frac{A}{\|A\|}$ $q_2 = \frac{B}{\|B\|}$ $q_3 = \frac{C}{\|C\|}$



$B = b - \frac{A^Tb}{A^TA}A$

$A^TB = A^Tb - \frac{A^Tb}{A^TA}A^TA = 0 \checkmark$

(when we think about this question, the space is arbitrary — 2-D and 3-D are all OK!)

(So what we should figure out is just the connection between a and b.)

$C = c - \frac{A^Tc}{A^TA}A - \frac{B^Tc}{B^TB}B$ [subtract off its components in the A and B directions]

orthogonal (components based on the "orthogonal")

$\to C \perp A.$

$C \perp B.$

Ex: $a = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ $b = \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}$ $B = \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} - \frac{3}{3}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}$. $q_1 = \frac{A}{\sqrt{3}} = \begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix}$ $q_2 = \frac{B}{\sqrt{2}} = \begin{bmatrix} 0 \\ -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$ $Q = \begin{bmatrix} q_1 & q_2 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{3} & 0 \\ 1/\sqrt{3} & -1/\sqrt{2} \\ 1/\sqrt{3} & 1/\sqrt{2} \end{bmatrix}$

$\| \ A$

elimination: $A = LU$

We are always working here with this in the same space. (column space of A) $A = \begin{bmatrix} 1 & 0 \\ 1 & -1 \\ 1 & 1 \end{bmatrix}$ share the same column space

and we're just like getting 90° angles in there.

$A = QR$ matrix form.

upper triangular and square.

$\begin{bmatrix} a_1 & a_2 \end{bmatrix} = \begin{bmatrix} q_1 & q_2 \end{bmatrix} \begin{bmatrix} a_1^Tq_1 & a_2^Tq_1 \\ a_1^Tq_2 & a_2^Tq_2 \end{bmatrix}$

inner product

[under the diagonal, q's subscript > a's subscript]

a_i's projection along q_i ⇐ their sum must be a_i. because {a_i and $q_1, \dots, q_n$} are dependent!] independent

The whole point of Gram-Schmidt was that we constructed these later q's to be perpendicular to all the earlier vectors

[make new q to be totally new direction !!