APMA 4007

Lecture 13

Oct 15, 2024

Outline

1. Orthogonal complement
2. Least-square problem & sol$^{\underline{n}}$
3. Projections

In the last lecture I defined what it means for two subspaces to be <u>orthogonal</u>

$$V \perp W \quad \text{iff} \quad v^T w = 0 \text{ for all } v \in V \text{ \& } w \in W$$

The examples related to $Ax=b$ are:

$$N(A) \perp C(A^T) \qquad \text{and} \qquad N(A^T) \perp C(A)$$

goto (*)

(note that one follows from the other by replacing A w/ A^T)

A nice property of orthogonal subspaces is that their dimensions add. If $W \perp V$ then

$$\dim(W+V) = \dim W + \dim V$$

where $W + V = \{ v+w \mid v \in V \text{ \& } w \in W \}$

V & W are subspaces of the same vector space so that addition makes sense.

<u>Proof</u>

Suppose the columns of $\overset{m\times k}{A}$ are a basis for V and the columns of $\overset{m\times l}{B}$ are a basis for W with

$$A = \begin{bmatrix} | & & | \\ a_1 & \cdots & a_k \\ | & & | \end{bmatrix} \text{ \& } B = \begin{bmatrix} | & & | \\ b_1 & \cdots & b_l \\ | & & | \end{bmatrix}$$

Any element of $V + W$ looks like

$$\underbrace{Ax}_{\in V} + \underbrace{By}_{\in W} = \begin{bmatrix} A & B \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

b/c the columns of A & B are basis vectors for V & W

This shows that the columns of $[A \; B]$ <u>span</u> $V + W$.

(*) A related concept is orthogonal complement.

Def: The orthogonal complement of vector subspace V denoted $V^{\perp}$ (vee perp) is the set of all vectors $\perp$ to V.

Example

$N(A) = \{x \mid Ax = 0\} = \{x \mid x \text{ is } \perp \text{ to the } \cancel{\text{columns}} \text{ rows of } A\}$

$= [C(A^T)]^{\perp}$

Can we say that $[N(A)^{\perp}] = C(A^T)$?

In other words $[C(A^T)^{\perp}]^{\perp} = C(A^T)$?

Yes. Proof

Suppose $\overset{m\times n}{A}$ and $v \in \mathbb{R}^n$ and $v \in [N(A)]^{\perp}$ but NOT in $C(A^T)$ [this is essentially saying that $N(A)$ & $C(A^T)$ are not everything?]. In that case $\begin{bmatrix} A \\ v^T \end{bmatrix}$ ← A with extra row

has the same null space as A since $\begin{bmatrix} A \\ v^T \end{bmatrix} x = \begin{bmatrix} Ax \\ v^T x \end{bmatrix}$

and if $x \in N(A)$ then $\begin{bmatrix} Ax \\ v^T x \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ b/c $x \in N(A)$ and v is $\perp$ to $N(A)$. Since $v \notin C(A^T)$ it is not a L.C. of the rows of A. $\Rightarrow$ dim row space $\begin{bmatrix} A \\ v^T \end{bmatrix} = r+1 = \text{rank } A$ b/c we have added a L.I row. But the rule says

$\dim N\left(\begin{bmatrix} A \\ v^T \end{bmatrix}\right) = \#\text{ columns} - \text{rank} = n-(r+1)$ But $N\left(\begin{bmatrix} A \\ v \end{bmatrix}\right) = N(A)$ & $\dim N(A) = n-r$ ⨳

In order to show that the $l+k$ columns of $[A\ B]$ are a basis, it remains to show that they are L.I. Assume they are not (proof by contradiction)

$$Ax + By = \vec{0} \quad \text{for} \quad \vec{x} \neq 0 \ \text{\& } \vec{y} \neq 0.$$

Mult both sides by A^T which mean dot w/ the columns of A

$$A^T A x + \underbrace{A^T B y}_{= \vec{0}} = \vec{0}$$

$\vec{0}$ b/c the columns of A are $\perp$ to the columns of B

$A^T A x = 0$. Does it have a sol$^{\underline{n}}$ other than $x = 0$?

Dot both sides w/ x.

$$x^T A^T A x = 0$$

$$(Ax)^T Ax = 0$$

$$\|Ax\|^2 = 0$$

$$Ax = 0$$

because the only vector w/ length zero is the zero vector.

But $A\vec{x} = 0 \Leftrightarrow \vec{x} = 0$ b/c the columns of A are a basis (L.I.)

So the $k+l$ columns of $[A\ B]$ are a basis for $V+W$ and the $\dim(V+W) = k + l = \dim V + \dim W$

Note: We have shown that

1. The r pivot rows of R & the $n-r$ special solutions are a basis for $N(A) + C(A^T)$. A total of n vectors

2. The r pivot columns of A & the $m-r$ bottom rows of E ($EA = R$) are a basis of $C(A) + N(A^T)$. A total of m vectors.

A this point we (w/ a ~~sooner~~ little thought) can conclude that

They have the same basis.

$$\mathbb{R}^n = C(A^T) + N(A)$$

$$\mathbb{R}^m = C(A) + N(A^T)$$

w/ the correct interpretation of addition

The 4 subspaces are everything.

[Recall that we previously showed that if we have the right number of L.I. vectors, where the right # is the dimension, they are a basis]

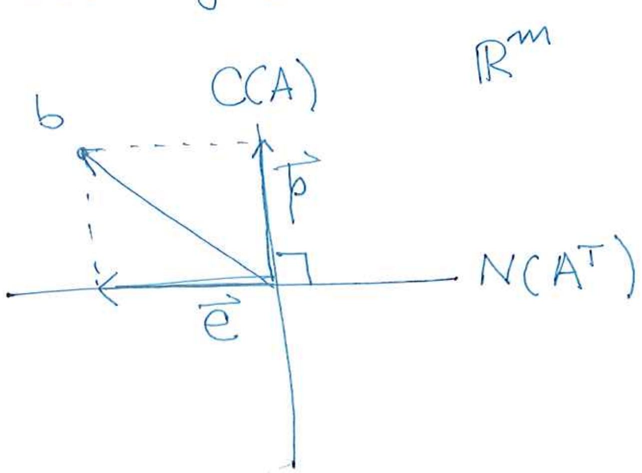
Same basis ~~bas~~ means r pivot rows of R & $n-r$ special sol$^{\underline{n}}$ are a basis for $\mathbb{R}^n$

r pivot columns of A & $m-r$ bottom rows of E are a basis for $\mathbb{R}^m$.

In particular (there is a reason this?),

for any $b \in \mathbb{R}^m$ $b = p + e$ where $p \in C(A)$ and $e \in N(A^T)$

The picture that goes with this is



Note how the picture differs from the previous one with $b \in C(A)$. In particular, if $\vec{e} \neq 0$

$b \notin C(A)$ & $Ax=b$ has no sol$^{\underline{n}}$. Another way to see this is

$$Ax = b = p + e$$

dot both sides w/ $\vec{e}$

$$\underbrace{e^T A x}_{\substack{\| \\ 0 \\ e \in N(A^T)}} = \underbrace{e^T p}_{\substack{\| \\ 0 \\ e \perp p}} + e^T e = \|e\|^2$$

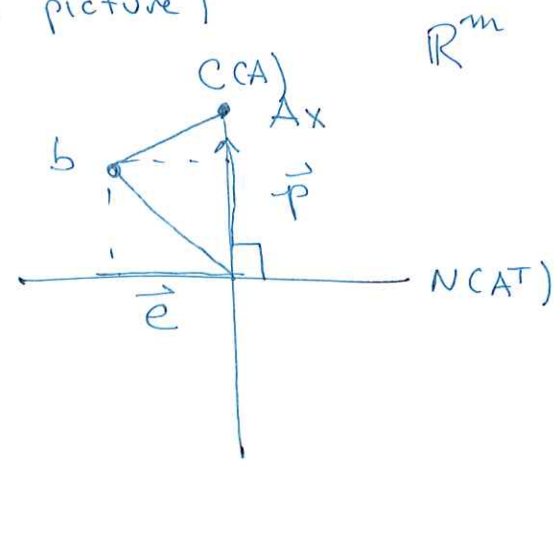
$0 = \|e\|^2$ ※ unless $\vec{e} = \vec{0}$. and $b \in C(A)$.

When $Ax=b$ has no sol$^{\underline{n}}$, we do the best that we can, which means make Ax as close to b as possible. That is,

$$\text{minimize } \|Ax-b\|^2 = \sum_{i=1}^{m} \left([Ax]_i - b_i\right)^2$$

This is called a least-squares problem.

(Add Ax to the picture)



The <u>picture</u> says (looking at the right triangle)

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

which is interesting b/c only one ~~the~~ ~~th~~ term on the rhs depends on x.

~~The~~ <u>Algebra</u> says

$$\begin{aligned}\|Ax-b\|^2 &= (Ax-b)^T(Ax-b)\\ &= (Ax-p-e)^T(Ax-p-e)\\ &= (Ax-p)^T(Ax-p) + e^Te + \overset{0}{\overset{\shortparallel}{(Ax-p)^Te}}\\ &\qquad + \underset{0}{\underset{\shortparallel}{e^T(Ax-p)}}\\ &= \|Ax-p\|^2 + \|e\|^2\end{aligned}$$

Both terms are non-negative. And we want to make their sum small. The 2nd term does not depend on x. It is unavoidable <u>error</u>. The 1st term can be made zero b/c $Ax=p$ has a sol$^{\underline{n}}$ which we call $\hat{x}$ (x hat). $\hat{x}$ is the solution of the L.S. problem. So just solve $Ax=p$? Problem: ~~plu~~ b is given <u>not</u> p. Finding p ~~solves~~ answers the ~~problem~~ question of what vector in $C(A)$ is closest to b.

p is called the <u>projection</u> of b onto $C(A)$.

Warm up. $n=1$. A has 1 column.

$$A = \begin{bmatrix} | \\ a \\ | \end{bmatrix} = \vec{a}$$

b is not in $C(A)$

[here means $\vec{b}$ is not a multiple of $\vec{a}$]

1x1 scalar

$$\vec{a}\, x = \vec{b}$$ has no sol^n

unknown

$b = p + e$ where $p = A\hat{x}$

unknown

$e = b - p$ & $A^T e = 0$ b/c $e \in N(A^T)$

So $A^T(b-p) = 0 \Rightarrow A^T b = A^T p$

here $A = a$ so

$$a^T b = a^T p = a^T a\, \hat{x}$$

$$\Rightarrow \hat{x} = \frac{a^T b}{a^T b} \quad \text{and} \quad p = \vec{a}\, \frac{a^T b}{a^T a}$$

which is the familar 1D projection from physics (inclined plane)

Example

Project $b = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ onto $\vec{a} = \begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix}$

$$\hat{x} = \frac{a^T b}{a^T a} = \frac{4}{14} = \frac{2}{7}$$

$$p = \frac{2}{7} \begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix} \qquad (\vec{a}\, x)$$

$$\vec{e} = \text{error} = b - p = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \frac{2}{7}\begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix} = \frac{1}{7}\begin{bmatrix} 3 \\ 9 \\ 1 \end{bmatrix}$$

check that $\vec{e} \perp \vec{a}$ $\quad \frac{1}{7}\begin{bmatrix} 3 \\ 9 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix} = \frac{1}{7}(6 - 9 + 3)$

$= 0$ ✓

Define a projection matrix $\underline{P}$ such that

$\underline{P}\, b = p$ here $p = \frac{a^T b}{a^T a}\vec{a}$ so rank-1 matrix.

$\underline{P} = \frac{a a^T}{a^T a}$, here in example $\underline{P} = \frac{1}{14}\begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix}\begin{bmatrix} 2 & -1 & 3 \end{bmatrix}$

If A has more than one column, same idea

$$b = p + e \qquad A\hat{x} = p$$

$$e = b - p$$

$$A^T e = 0 \Rightarrow A^T b = A^T p = A^T A \hat{x}$$

$$\hat{x} = (A^T A)^{-1} A^T b \quad \checkmark$$

L.S. sol$\underline{^n}$

$$p = A\hat{x} = \underbrace{A(A^TA)^{-1}A^T}\, b$$

projection

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

projection matrix

Remarks.

1. $P = A(A^TA)^{-1}A^T$ Don't all the A's cancel?

$$= A A^{-1}(A^T)^{-1}A^T = I \ ?$$

Generally, no b/c A is not square. If A^{-1} exists there is no need for L.S.

2. What happens if you project twice?

$$P^2 = A(A^TA)^{-1}(A^T A)(A^TA)^{-1}A^T = A(A^TA)^{-1}A^T = P$$

nothing happen the 2nd time $P[Pb]$
$\uparrow$ already in $C(A)$

3. What does $(I-P)$ do? Projects onto $N(A^T)$

$(I-P)b = b - p = e$ ✓ closest vector in $N(A^T)$ to b.

4. Does (A^TA) have an inverse?

- it is square $\underset{n\times m}{A^T}\ \underset{m\times n}{A}$ $\;n\times n$

- are the columns L.I.? Yes iff the columns of A are.

Proof

I will show $N(A) = N(A^TA)$.

This is useful b/c columns A L.I iff $N(A) = \vec{0}$

$\Leftrightarrow N(A^TA) = \vec{0} \Leftrightarrow A^{-1}$ exists. (columns of A^TA L.I)

$$x \in N(A) \Rightarrow Ax = 0 \Rightarrow A^TAx = 0 \Rightarrow x \in N(A^TA)$$

$$N(A) \subseteq N(A^TA)$$

$$x \in N(A^TA) \Rightarrow A^TAx = 0 \Rightarrow x^TA^TAx = 0 \Rightarrow \|Ax\|^2 = 0$$

$$\Rightarrow Ax = 0$$

$$\Rightarrow x \in N(A)$$

$$N(A^TA) \subseteq N(A)$$

$\Rightarrow N(A^TA) = N(A)$ ✓

So if the columns of A are L.I, the sol$^{\underline{n}}$ of the L.S. problem is

$$\hat{x} = A(A^TA)^{-1}A^Tb.$$