11/07/2019

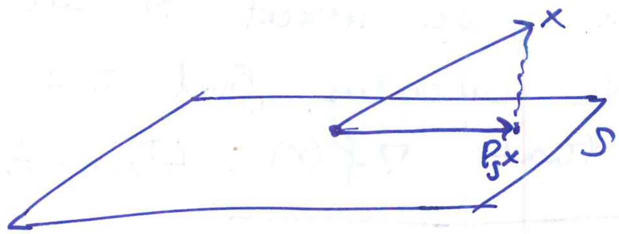
Lecture 19

<u>Orthogonal projections:</u> Let $S \subseteq \mathbb{R}^n$ subspace ~~[illegible]~~

How do we project orthogonally onto S?

I.e., ~~[illegible]~~ for $x \in \mathbb{R}^n$, we want:

$$\min_{y \in S} \|y - x\| \quad , \text{ and we call } y^* = P_S x \text{ the minimize}$$



<u>Facts:</u> P_S is a linear operator $P_S : \mathbb{R}^n \to \mathbb{R}^n$
$x \mapsto P_S x$

For every $x \in \mathbb{R}^n$, $P_S x$ is the orth. proj. of x onto S.

$P = P_S$ satisfies the following properties:

(a) $P^2 = P$

(b) $P^T = P$

We call any $n \times n$ matrix P that satisfies (a) and (b) an orthogonal proj. matrix. For such a matrix:

(1) $Q = I - P$ is also an orth. proj. matrix

(2) $PQ = QP = 0$

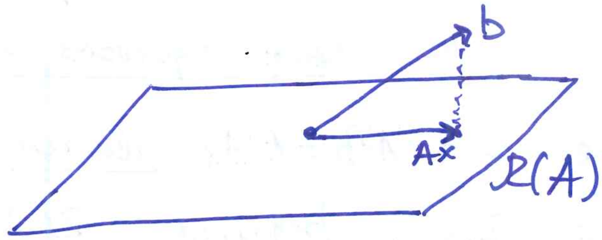
(3). P projects orthogonally onto $S = \mathcal{R}(P)$, i.e. $\forall x$, Px is the closest to x in $S = \mathcal{R}(P)$.

(4). Q projects orthogonally onto $S^\perp = \mathcal{R}(P)^\perp = \mathcal{N}(P)$.

Given a subspace S, how do we project onto it?

<u>linear regression:</u> $b = \overset{\text{known}}{A} \underset{\text{to find}}{x} + z$

We want to project b onto $\mathcal{R}(A)$!!!



We want to find x so that Ax is as close as possible to b, i.e. $\min_{x \in \mathbb{R}^m} \|b - Ax\|$

$P = P_{\mathcal{R}(A)}$

This will happen when $Ax = Pb$ is the projection of b onto $\mathcal{R}(A)$. Two ways to find x:

①. If $f(x) = \|b - Ax\|^2$, we want to minimize f, so we use multivariable calculus, find ∇f and set it to 0. It turns out that $\nabla f(x) = A^TAx - A^Tb \Rightarrow \boxed{A^TAx = A^Tb.}$

②.

In this case, $Qb = (I-P)b = b - Ax$ is the projection onto $\mathcal{R}(A)^\perp$

$$\Rightarrow Qb \perp \mathcal{R}(A)$$

$$\Rightarrow (b-Ax) \perp \mathcal{R}(A)$$

$$\Leftrightarrow (b-Ax) \in \mathcal{R}(A)^\perp$$

$$\Leftrightarrow b - Ax \in \mathcal{N}(A^T)$$

$$\Leftrightarrow A^T(b-Ax) = 0$$

$$\Leftrightarrow \boxed{A^Tb = A^TAx.}$$ The least squares equations. (LSE)

__Remarks:__

1. Any solution to the LSE minimizes $\|b-Ax\|$ over $x \in \mathbb{R}^n$.
2. If A^TA is invertible, then $\exists$ a unique solution

$$x^* = (A^TA)^{-1}A^Tb$$

the least squares solution.

3. The LSE $A^Tb = A^TAx$ __always__ has a solution.

__Proof:__ This is because $\mathcal{R}(A^T) = \mathcal{R}(A^TA)$.

or equivalently $\mathcal{R}(A^T)^\perp = \mathcal{R}(A^TA)^\perp \Leftrightarrow \mathcal{N}(A) = \mathcal{N}(A^TA)$.

We showed this in the graphs & networks chapter.

6. Projection onto a line $L = \mathrm{span}\{a\}$:

$$A = \begin{bmatrix} a \end{bmatrix}_{m\times 1}$$

Then $A^TA = \langle a, a\rangle = \|a\|^2$ invertible if $x \neq \vec{0}$.

Then, $x = (A^TA)^{-1}A^Tb = \frac{1}{\|a\|^2} a^Tb = \frac{\langle a, b\rangle}{\|a\|^2}$, $\boxed{A(A^TA)^{-1}A^T = \frac{aa^T}{\|a\|^2}}$.

which is exactly the projection onto span$\{a\}$!!!

4. A^TA is invertible. $\iff N(A^TA) = \{\vec{0}\} \iff N(A) = \{\vec{0}\}$

$\iff$ $\boxed{\text{Columns of } A \text{ are linearly independent!!!}}$

5. If A^TA is invertible, then $x = (A^TA)^{-1}A^Tb$, so $Ax = A(A^TA)^{-1}A^Tb$

$\boxed{P = A(A^TA)^{-1}A^T}$ is the projection matrix onto $R(A)$!

7. If A^TA is invertible, then $(A^TA)^{-1}A^T$ is a left inverse of A:

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

Notation: $A^+ = (A^TA)^{-1}A^T$ the Moore-Penrose pseudoinverse of A.

In MATLAB: pinv(A) gives A^+.

8. If A^TA is not invertible, then how can we find the projection Pb onto $R(A)$?

A^TA not invertible $\iff$ columns of A are l.d.

- Find a basis for $R(A)$, e.g. by Gaussian elimination.
- Create a new matrix $\tilde{A}$ whose columns are the basis vectors
- Then $R(A) = R(\tilde{A})$, and $P = P_{R(A)} = P_{R(\tilde{A})} = \tilde{A}(\tilde{A}^T\tilde{A})^{-1}\tilde{A}^T$.

Example: If $A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$, set $\tilde{A} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$.

Then, $\tilde{A}^T\tilde{A} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$

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

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

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

the projection matrix onto $\text{span}\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix} \right\}$.