Subject § 5.3

Given $A_{m\times n}$ and $b \in \mathbb{R}^m$. Consider solving

$$Ax = b$$

1) Consistent. Already studied

2) Inconsistent. no solution. It is natural to find $x \in \mathbb{R}^n$ such that Ax is closest to b.

For each $x \in \mathbb{R}^n$, the residual to $Ax = b$ is

$$r(x) = b - Ax.$$

The norm $\|r(x)\| = \|b - Ax\|$. Consider the problem: find $x \in \mathbb{R}^n$ such that $\|r(x)\| = \|b - Ax\|$ is minimized. Such a solution x is called a least square solution to $Ax = b$.

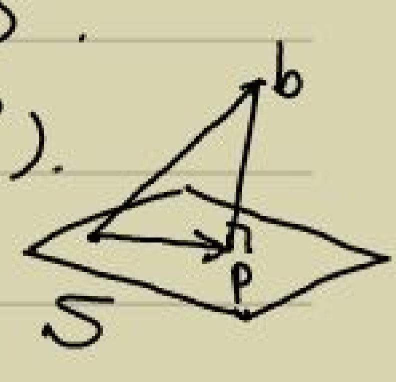
* $\min \|r(x)\| \Longleftrightarrow \min \|r(x)\|^2$.

<u>THM</u>. Let S be a subspace of $\mathbb{R}^m$. For each $b \in \mathbb{R}^m$, there is a unique p in S such that

$$\|b - p\| < \|b - y\| \quad \forall y \in S,\ y \neq p. \qquad (1)$$

Furthermore, a vector p in S satisfying (1) if and only if $b - p \perp S$ or $b - p \in S^\perp$.

(p is called the projection of b onto S).



Proof: We have $\mathbb{R}^m = S \oplus S^\perp$.

$b \in \mathbb{R}^m$. $b = \underbrace{p + p^\perp}_{\text{unique}}$ for some $p \in S$, $p^\perp \in S^\perp$.

$\Rightarrow b - p = p^\perp \perp S$.

To verify $\|b-p\| < \|b-y\| \quad \forall y \in S,\ y \neq p$.

we have

$\|b-y\|^2 = \|b-p+p-y\|^2$ $\quad$ $b-p \perp S$, $p-y \in S$ by Pythegorean's Law

$= \|b-p\|^2 + \|p-y\|^2 > \|b-p\|^2$. if $y \neq p$.

Thus any $y \neq p$ cannot be a solution. $\Rightarrow$ uniqueness.

$\divideontimes$

consider $\quad Ax = b$.

$S = \mathcal{R}(A) = CS(A)$ a subspace of $\mathbb{R}^m$.

$x^* \in \mathbb{R}^n$ is a least square solution if and only if $p = Ax^*$ is closest to b. $\Leftrightarrow$ $p = Ax^*$ is the projection of b onto $S = \mathcal{R}(A)$. or if and only if

$b - p = b - Ax^* \perp S = \mathcal{R}(A) \Rightarrow$

$b - Ax^* \in S^\perp = \mathcal{R}(A)^\perp = \mathcal{N}(A^T)$. i.e.,

$A^T(b - Ax^*) = \theta \quad$ or $\quad A^TAx^* = A^Tb$ --- normal equation.

from which we can solve for a least square solution x^*.

Steps to find a least square solution to $Ax=b$.

1) set $L=A^TA_{n\times n}$ always symmetric;
2) compute $d=A^Tb$
3). solve the normal equation $Lx=d$ for x^*.

it is always consistent.

we may have x_1^*, x_2^*. But $Ax_1^*=Ax_2^*$ unique in $\mathcal{R}(A)$. Both x_1^* and x_2^* are least square solutions.

If $A_{m\times n}$ has rank $=n$, then $(A^TA)_{n\times n}$ has rank $=n$ thus invertible. The unique least square solution

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

The projection of b onto $\mathcal{R}(A)$ is

$$p=Ax^*=A(A^TA)^{-1}A^Tb.$$

$P=A(A^TA)^{-1}A^T$ --- projection operator matrix.

$Pb=$ projection of b onto $\mathcal{R}(A)$.

Ex: Find the least square solution to $\begin{cases} x_1+x_2=3 \\ -2x_1+3x_2=1 \\ 2x_1-x_2=2 \end{cases}$.

Sol: we have $A = \begin{bmatrix} 1 & 1 \\ -2 & 3 \\ 2 & -1 \end{bmatrix}$, $b = \begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix}$. $A^T = \begin{bmatrix} 1 & -2 & 2 \\ 1 & 3 & -1 \end{bmatrix}$.

$$L = A^T A = \begin{bmatrix} 1 & -2 & 2 \\ 1 & 3 & -1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -2 & 3 \\ 2 & -1 \end{bmatrix} = \begin{bmatrix} 9 & -7 \\ -7 & 11 \end{bmatrix} \text{ symmetric}$$

$$d = A^T b = \begin{bmatrix} 1 & -2 & 2 \\ 1 & 3 & -1 \end{bmatrix} \begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 5 \\ 4 \end{bmatrix}.$$

Solve the normal equation $Lx = d$ or

$$\begin{bmatrix} 9 & -7 \\ -7 & 11 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 5 \\ 4 \end{bmatrix}.$$

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \frac{1}{99-49} \begin{bmatrix} 11 & 7 \\ 7 & 9 \end{bmatrix} \begin{bmatrix} 5 \\ 4 \end{bmatrix} = \frac{1}{50} \begin{bmatrix} 83 \\ 71 \end{bmatrix}.$$

<u>Curve Fitting</u>:

If experimental (sample) data have been collected

x	x_1	$\cdots$	x_m
y	y_1	$\cdots$	y_m

where x = time, position or sequential index

wish to find a polynomial $p(x) = c_0 + c_1 x + \cdots + c_n x^n$ of degree n, where $m \gg n$, such that

$$p(x_i) = y_i, \quad i = 1, \cdots, m \gg n \text{ — overdetermined or inconsistent.}$$

$$\Downarrow$$

$$\begin{array}{c} \begin{matrix} x^0 & x^1 & x^2 & \cdots & x^n \end{matrix} \\ \begin{bmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^n \\ 1 & x_2 & x_2^2 & \cdots & x_2^n \\ \cdots \\ 1 & x_m & x_m^2 & \cdots & x_m^n \end{bmatrix} \end{array} \begin{bmatrix} c_0 \\ c_1 \\ \vdots \\ c_n \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}, \text{ denoted by } AC = Y.$$

Since $m \gg n$, most time, it is inconsistent. So find the best least square polynomial by solving $A^T A C = A^T Y$ ... normal equation.

for $c_0, c_1, \cdots, c_n$, and then $p(x) = c_0 + c_1 x + \cdots + c_n x^n$.

Ex: Given

x	0	1	2	3
y	3	2	4	4

Find the best quadratic least square fitting.

sol: $n=2$.

$$\begin{array}{ccc} x^0 & x^1 & x^2 \end{array}$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \\ 4 \\ 4 \end{bmatrix}.$$

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \end{bmatrix},\ A^T = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \\ 0 & 1 & 4 & 9 \end{bmatrix},\ y = \begin{bmatrix} 3 \\ 2 \\ 4 \\ 4 \end{bmatrix} \quad A^T y = \begin{bmatrix} 13 \\ 22 \\ 54 \end{bmatrix}.$$

$$A^T A = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \\ 0 & 1 & 4 & 9 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \end{bmatrix} = \begin{bmatrix} 4 & 6 & 14 \\ 6 & 14 & 36 \\ 14 & 36 & 98 \end{bmatrix}.$$

solve the normal equation

$$\begin{bmatrix} 4 & 6 & 14 \\ 6 & 14 & 36 \\ 14 & 36 & 98 \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 13 \\ 22 \\ 54 \end{bmatrix}$$

$\Rightarrow (c_0, c_1, c_2) = (2.75, -0.25, 0.25)$.

$\Rightarrow p(x) = 2.75 - 0.25x + 0.25x^2$. ※.