10/22/2021

last time:

→ derivation of normal equations

$$\underline{\underline{A}}^T \underline{\underline{A}}\, \underline{x} = \underline{\underline{A}}^T \underline{b}$$

using FTLA.

Geometry of the Normal Equations

goal: $\min_{\underline{x}} \| \underline{\underline{A}}\, \underline{x} - \underline{b} \|_2^2$.

* If we find the closest element $\underline{\underline{A}}\, \underline{x}^*$ to $\underline{b}$ (in $R(\underline{\underline{A}})$), we have achieved our goal.

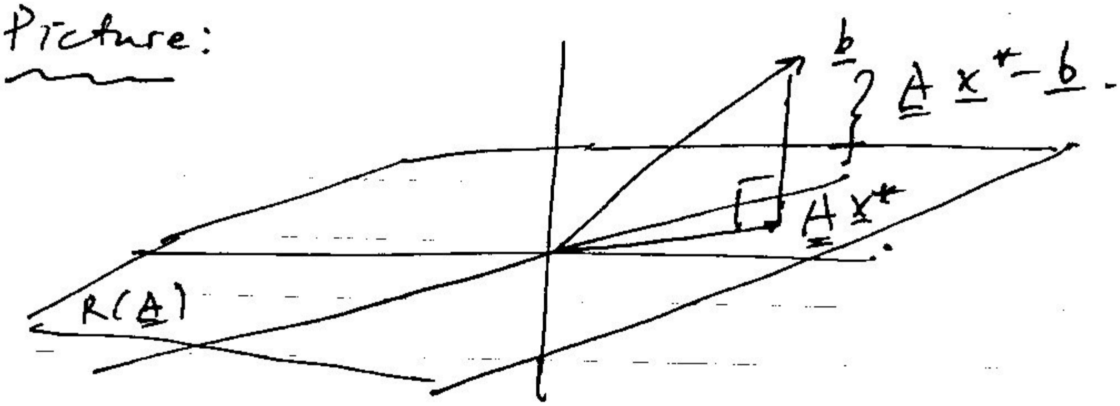
Notice: $\underline{\underline{A}}^T \underline{\underline{A}}\, \underline{x}^* = \underline{\underline{A}}^T b \Leftrightarrow \underline{\underline{A}}^T(\underline{\underline{A}}\, \underline{x}^* - \underline{b}) = \underline{0}$.

Take the inner product of above with $y \in \mathbb{R}^n$ to conclude:

$$\underbrace{y^T \underline{\underline{A}}^T}_{(\underline{\underline{A}} y)^T} (\underline{\underline{A}}\, \underline{x}^* - \underline{b}) = 0$$

i.e. we have $R(\underline{\underline{A}}) \perp \underline{\underline{A}}\,\underline{x}^* - \underline{b}$.

Picture:



Observation: we can see from the picture that $\underline{\underline{A}}\,\underline{x}^*$ is the orthogonal projection of $\underline{b}$ onto $R(\underline{\underline{A}})$.

To see this in equation form, using the normal equations:

$$\underline{\underline{A}}^T \underline{\underline{A}}\,\underline{x}^* = \underline{\underline{A}}^T \underline{b} \iff \underline{x}^* = (\underline{\underline{A}}^T\underline{\underline{A}})^{-1}\underline{\underline{A}}^T\underline{b}$$

So, $\underline{\underline{A}}\,\underline{x}^* = \underline{\underline{A}}\,(\underline{\underline{A}}^T\underline{\underline{A}})^{-1}\underline{\underline{A}}^T\underline{b}$.

Check: $\underline{\underline{P}} = \underline{\underline{A}}\,(\underline{\underline{A}}^T\underline{\underline{A}})^{-1}\underline{\underline{A}}^T$ is an orthogonal projector on $R(\underline{\underline{A}})$.

Ex: (fitting a parabola)

Data: $(t_1, y_1) = (0, 2)$

$(t_2, y_2) = (-1, 5)$

$(t_3, y_3) = (-2, 4)$

$(t_4, y_4) = (1, 1)$

goal: fit a parabola $g(t) = \gamma t^2 + \alpha t + \beta$ to the data.

$$\underline{b} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix} = \begin{pmatrix} 2 \\ 5 \\ 4 \\ 1 \end{pmatrix}.$$

$$\underline{\underline{A}} = \begin{pmatrix} t_1^2 & t_1 & 1 \\ t_2^2 & t_2 & 1 \\ t_3^2 & t_3 & 1 \\ t_4^2 & t_4 & 1 \end{pmatrix}, \quad \underline{x} = \begin{pmatrix} \gamma \\ \alpha \\ \beta \end{pmatrix}.$$

$$\Rightarrow \underline{\underline{A}} = \begin{pmatrix} 0 & 0 & 1 \\ 1 & -1 & 1 \\ 4 & -2 & 1 \\ 1 & 1 & 1 \end{pmatrix}.$$

matlab example in class...