10/25/2021

Last time
→ geometry of the normal equations
→ matlab code examples.

Start with some matlab code for least squares... "least_squares_ex2.m"

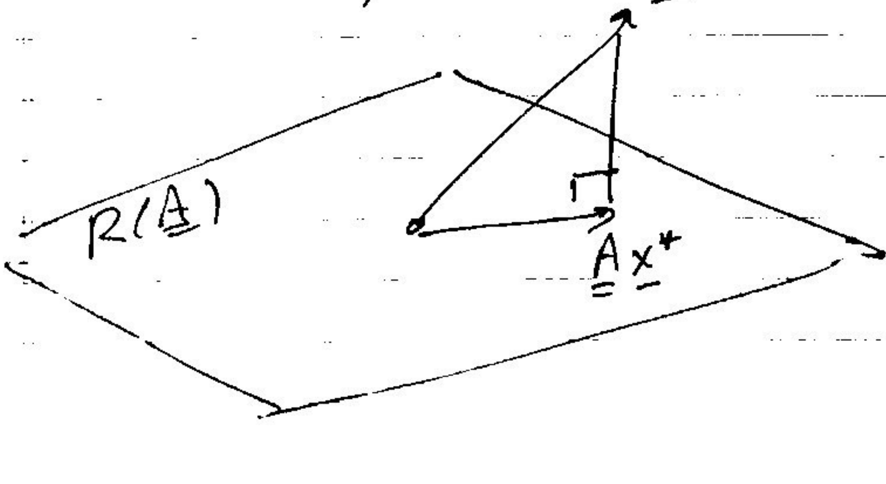
Projections

Recall from our least squares solution.

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

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

Geometrically :

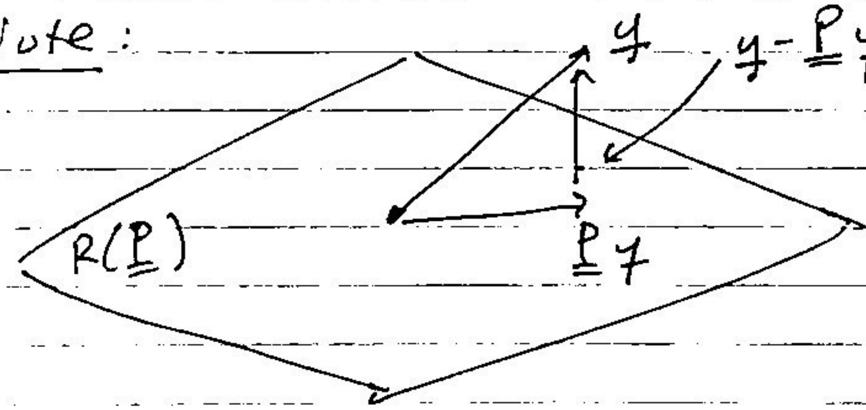


→

i.e. $\underline{\underline{A}}\,\underline{x}^+$ is the <u>orthogonal projection</u> of $\underline{b}$ onto $R(\underline{\underline{A}})$.

<u>Definition</u>: A square matrix $\underline{\underline{P}}$ that satisfies $\underline{\underline{P}}^2 = \underline{\underline{P}}\,\underline{\underline{P}} = \underline{\underline{P}}$ is called a <u>projection</u>. A projection $\underline{\underline{P}}$ that is symmetric, i.e. $\underline{\underline{P}} = \underline{\underline{P}}^T$, is an <u>orthogonal projection</u>.

<u>Note</u>:



If $\underline{\underline{P}}$ is an orthogonal projection, consider:

$$(\underline{\underline{P}}\,\underline{y})^T(\underline{y} - \underline{\underline{P}}\,\underline{y}) = \underline{y}^T\underline{\underline{P}}^T(\underline{y} - \underline{\underline{P}}\,\underline{y})$$

$$= y^T \underline{\underline{P}}^T y - y^T \underline{\underline{P}}^T \underline{\underline{P}} y$$

$$= y^T \underline{\underline{P}} y - y^T \underline{\underline{P}}^2 y \qquad (\underline{\underline{P}} = \underline{\underline{P}}^T)$$

$$= y^T \underline{\underline{P}} y - y^T \underline{\underline{P}} y \qquad (\underline{\underline{P}}^2 = \underline{\underline{P}})$$

$$= 0,$$

i.e. $\underline{\underline{P}} y$ is perpendicular to $y - \underline{\underline{P}} y$.

Ex: Define $\underline{\underline{P}} = \dfrac{\underline{v}\,\underline{v}^T}{\underline{v}^T\underline{v}} = \dfrac{\underline{v}\,\underline{v}^T}{\|\underline{v}\|_2^2}$.

Check: $\underline{\underline{P}}\,\underline{\underline{P}} = \dfrac{\underline{v}\,\underline{v}^T\underline{v}\,\underline{v}^T}{(\underline{v}^T\underline{v})^2} = \dfrac{\underline{v}(\underline{v}^T\underline{v})\underline{v}^T}{(\underline{v}^T\underline{v})^2}$

$$= \frac{\underline{v}\,\underline{v}^T}{\underline{v}^T\underline{v}} = \underline{\underline{P}} \checkmark$$

$$\underline{\underline{P}}^T = \left(\frac{\underline{v}\,\underline{v}^T}{\underline{v}^T\underline{v}}\right)^T = \left(\frac{\underline{v}^{T^T}\underline{v}^T}{\underline{v}^T\underline{v}}\right) = \frac{\underline{v}\,\underline{v}^T}{\underline{v}^T\underline{v}} = \underline{\underline{P}},$$

so $\underline{\underline{P}} = \dfrac{\underline{v}\,\underline{v}^T}{\underline{v}^T\underline{v}}$ is an orthogonal projector. onto what ???

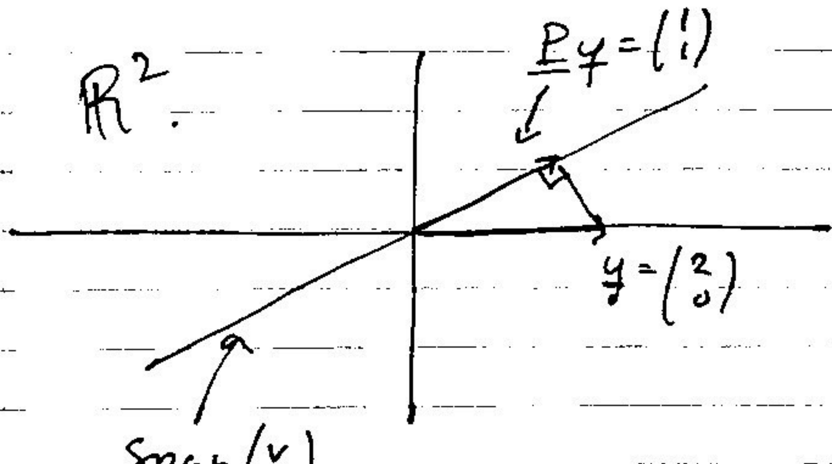
the span$(\underline{v})$.

Take $\underline{v} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

Then $\underline{\underline{P}} = \dfrac{\underline{v}\,\underline{v}^T}{\underline{v}^T\underline{v}} = \dfrac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$.

Let $\underline{y} = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$. Then $\underline{\underline{P}}\underline{y} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}\begin{pmatrix} 2 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

Picture: $\mathbb{R}^2$.



Ex: Define $\tilde{\underline{\underline{P}}} = \underline{\underline{I}} - \underline{\underline{P}}$, and suppose $\underline{\underline{P}}$ is a projector. Then $\tilde{\underline{\underline{P}}}$ is also a projector. If $\underline{\underline{P}}$ is an orthogonal projector, the $\tilde{\underline{\underline{P}}}$ is also an orthogonal projector. For example:

$$\tilde{\underline{\underline{P}}}^2 = (\underline{\underline{I}} - \underline{\underline{P}})^2 = \underline{\underline{I}}^2 - 2\underline{\underline{P}} + \underline{\underline{P}}^2$$

$$= \underline{\underline{I}} - 2\underline{\underline{P}} + \underline{\underline{P}} = \underline{\underline{I}} - \underline{\underline{P}} = \tilde{\underline{\underline{P}}}.$$

What does $\underline{\underline{\tilde{P}}}$ do geometrically?

Consider $\underline{\underline{P}}\,\underline{y} \in R(\underline{\underline{P}})$ and

$$\underline{\underline{\tilde{P}}}\,\underline{z} = (\underline{\underline{I}} - \underline{\underline{P}})\underline{z} \in R(\underline{\underline{I}} - \underline{\underline{P}}).$$

look at $(\underline{\underline{P}}\,\underline{y})^T(\underline{\underline{\tilde{P}}}\,\underline{z}) = (\underline{\underline{P}}\,\underline{y})^T(\underline{\underline{I}} - \underline{\underline{P}})\underline{z}$

$$= \underline{y}^T \underline{\underline{P}}^T \underline{z} - \underline{y}^T \underline{\underline{P}}^T \underline{\underline{P}}\,\underline{z}$$

$$= \underline{y}^T \underline{\underline{P}}\,\underline{z} - \underline{y}^T \underline{\underline{P}}\,\underline{z} = 0,$$

i.e. $R(\underline{\underline{P}})$ and $R(\underline{\underline{I}} - \underline{\underline{P}})$ are orthogonal subspaces.

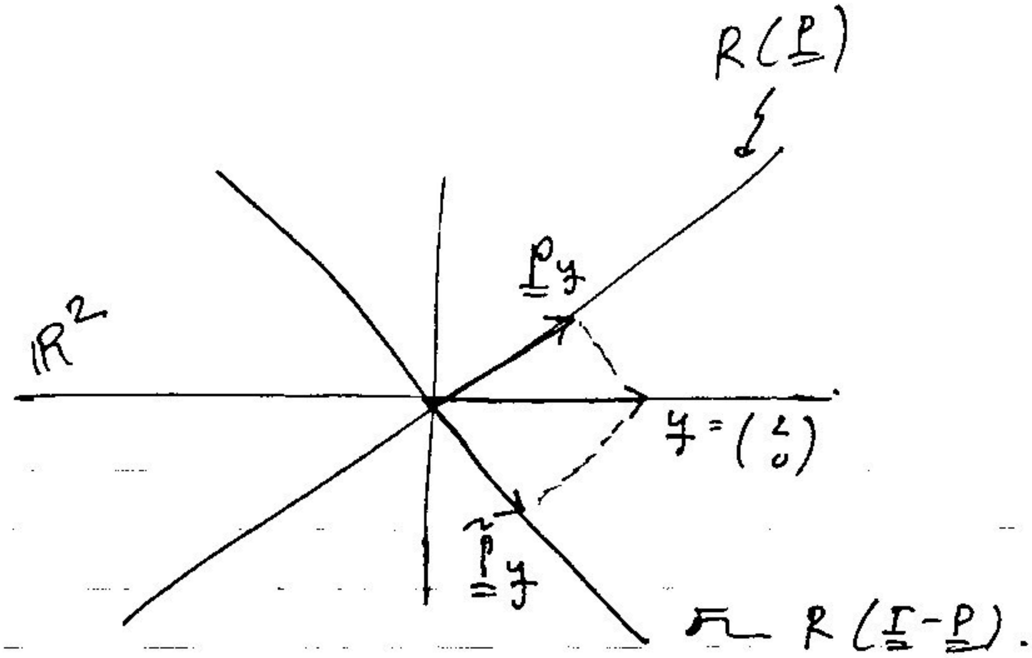
We say that $\underline{\underline{I}} - \underline{\underline{P}}$ projects vectors orthogonally onto the <u>orthogonal complement</u> to $R(\underline{\underline{P}})$.

Again, If $\underline{\underline{P}} = \dfrac{\underline{v}\,\underline{v}^T}{\underline{v}^T\underline{v}} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$ with

$\underline{v} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, then

$$\underline{\underline{\tilde{P}}} = \underline{\underline{I}} - \underline{\underline{P}} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$

Picture:



look at $\underline{\underline{\tilde{P}}}\, y = \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} 2 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$.

In those previous examples, we were considering a single vector $\underline{v}$ and defining $\underline{\underline{P}} = \frac{\underline{v}\,\underline{v}^T}{\underline{v}^T\underline{v}}$. Note that

$$\underline{\underline{P}} = \frac{\underline{v}\,\underline{v}^T}{\underline{v}^T\underline{v}} = \underline{v}\,(\underline{v}^T\underline{v})^{-1}\underline{v}^T,$$ which

is starting to look like: $\underline{\underline{A}}\,(\underline{\underline{A}}^T\underline{\underline{A}})^{-1}\underline{\underline{A}}^T$.

!!!