11/05/2019 Lecture 18

- Homework: due Today
- Quiz 3: on Thursday; II.3, III.1.1, III.1.2; At end of class!!
- Final exam date: Sat, Dec 7, 8:30 AM

<u>Last time</u>: Projecting onto lines ~~[illegible]~~. $L = \text{span}\{x\}$.

The <u>projection</u> of y onto the line $L = \text{span}\{x\}$ is the vector in L closest to y. Denote this vector by $P_x y$.

↳ $P_x y = sx$ for some s, and we showed that s is the minimizer of $f(s) = \|y - sx\|_2^2$ ~~[illegible]~~.

After minimizing f, we got $s = \frac{\langle x, y\rangle}{\|x\|^2}$, thus,

$$\boxed{P_x y = \frac{\langle x, y\rangle}{\|x\|^2} x.}$$

$$\frac{\langle x, y\rangle}{\|x\|^2} x = \frac{x^T y}{\|x\|^2} x = \frac{x x^T y}{\|x\|^2}$$

Then, we showed that

$$\Rightarrow \boxed{P_x y = \underbrace{\frac{1}{\|x\|^2} x x^T}_{n \times n} \underbrace{y}_{n \times 1}}$$

$P_x := \frac{1}{\|x\|^2} x x^T$ is the projection matrix/operator.

<u>Example</u>: Let $x = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}$. Then,

$$P_x = \frac{1}{\|x\|^2} x x^T = \frac{1}{5} \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} \begin{bmatrix} 1 & 2 & 0 \end{bmatrix} = \begin{bmatrix} \frac{1}{5} & \frac{2}{5} & 0 \\ \frac{2}{5} & \frac{4}{5} & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

If $y = \begin{bmatrix} 2 \\ 8 \\ 3 \end{bmatrix}$, then

$$P_x y = \begin{bmatrix} \frac{1}{5} & \frac{2}{5} & 0 \\ \frac{2}{5} & \frac{4}{5} & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 2 \\ 8 \\ 3 \end{bmatrix} = \begin{bmatrix} \frac{18}{5} \\ \frac{36}{5} \\ 0 \end{bmatrix}.$$

Another way to rewrite $P_x y$ is:

$$P_x y = \frac{\langle x, y\rangle}{\|x\|^2} x = \left\langle \frac{x}{\|x\|}, y \right\rangle \frac{x}{\|x\|}$$

What is $\frac{x}{\|x\|}$? It is the unit vector in the direction of x.

Call it $\hat{x}$.

Then, $\boxed{P_x y = \langle \hat{x}, y\rangle \hat{x}.}$

Some important properties of projection operators $P = P_x$:

① $P(Py) = Py$, i.e. $P^2 = P$

Proof: $P = \frac{1}{\|x\|^2}(xx^T) \Rightarrow P^2 = \left(\frac{1}{\|x\|^2} xx^T\right)\left(\frac{1}{\|x\|^2} xx^T\right)$

$$= \frac{1}{\|x\|^4} x \underbrace{x^T x}_{\|x\|^2} x^T = \frac{1}{\|x\|^2} xx^T = P. \checkmark$$

②. $P^T = P$

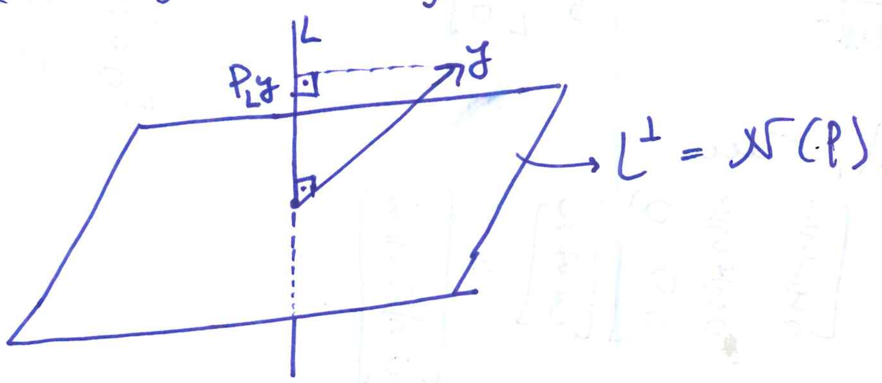
Proof: $P^T = \left(\frac{1}{\|x\|^2} xx^T\right)^T = \frac{1}{\|x\|^2}(x^T)^T x^T = \frac{1}{\|x\|^2} xx^T = P. \checkmark$

Consequences: (a) $\langle z, Py\rangle = \langle Pz, y\rangle$ since $P^T = P$

(b). $\langle Pz, Py\rangle = \langle P^2 z, y\rangle = \langle Pz, y\rangle$ since $P^T = P$ and $P^2 = P$.

③. $\mathcal{R}(P) = L = \text{span}\{x\}$.

④. $\mathcal{N}(P) = (\mathcal{R}(P^T))^\perp = (\mathcal{R}(P))^\perp = (\text{span}\{x\})^\perp = L^\perp$.



How do we project onto $L^\perp$? Let $P_{L^\perp} y = y - P_L y = (I - P_L) y$

$$\Rightarrow \quad y = P_L y + P_{L^\perp} y \qquad (*)$$

$\cancel{\Rightarrow P_{L^\perp} y = y - P_L y = (I - P_L) y.}$

Conclusion: $P_{L^\perp} = I - P_L$ is the projection operator onto $L^\perp$.

Example continued: $P_{L^\perp} = I - P_L = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \begin{bmatrix} \frac{1}{5} & \frac{2}{5} & 0 \\ \frac{2}{5} & \frac{4}{5} & 0 \\ 0 & 0 & 0 \end{bmatrix}$

$$= \begin{bmatrix} \frac{4}{5} & -\frac{2}{5} & 0 \\ -\frac{2}{5} & \frac{1}{5} & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Thus, $P_{L^\perp} \begin{bmatrix} 2 \\ 8 \\ 3 \end{bmatrix} = \begin{bmatrix} \frac{4}{5} & -\frac{2}{5} & 0 \\ -\frac{2}{5} & \frac{1}{5} & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} -\frac{8}{5} \\ \frac{4}{5} \\ 3 \end{bmatrix}.$

~~Example~~ $(*)$ For any vector y, show that $P_L y \perp P_{L^\perp} y$:

$$\langle P_L y, P_{L^\perp} y \rangle = \langle P_L y, (I - P_L) y \rangle$$

$$= \langle (I - P_L)^T P_L y, y \rangle$$

But $(I - P_L)^T = I^T - P_L^T = I - P_L$

$$\Rightarrow (I - P_L)^T P_L = (I - P_L) P_L = P_L - P_L^2 = P_L - P_L = 0$$

$$\Rightarrow \langle P_L y, P_{L^\perp} y \rangle = \langle 0, y \rangle = 0.$$

Projecting onto general linear subspaces:

An $n \times n$ matrix P is an <u>orthogonal projection matrix</u> if

(a) $P^2 = P$, and

(b) $P^T = P$.

For such a matrix define $Q = I - P$. Then:

(1). Q is also an orthogonal projection matrix

(2). $P + Q = I$ and $PQ = QP = 0$.

(3). P projects orthogonally onto $\mathcal{R}(P)$

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

<u>Proofs</u>: see typed notes.

<u>One immediate consequence</u>: <u>Pythagorean Theorem</u>:
Let $x \in \mathbb{R}^n$ and let P be an orthogonal projection matrix with $Q = I - P$. Then, $\|x\|^2 = \|Px\|^2 + \|Qx\|^2$.

<u>Proof</u>: $\|x\|^2 = \langle x, x \rangle$

and $x = Px + Qx$

$$\Rightarrow \|x\|^2 = \langle Px + Qx, Px + Qx \rangle \overset{\text{expand}}{=} \langle Px, Px \rangle + \overbrace{\langle Px, Qx \rangle}^{0}$$

$$+ \underbrace{\langle Qx, Px \rangle}_{0} + \langle Qx, Qx \rangle = \langle Px, Px \rangle + \langle Qx, Qx \rangle$$

$$= \|Px\|^2 + \|Qx\|^2.$$

$$\langle Px, Qx \rangle = \langle Q^T Px, x \rangle \overset{(b)+(1)}{=} \langle QPx, x \rangle \overset{(2)}{=} \langle 0x, x \rangle = 0.$$

Next: linear regression and least squares!!!

Example: linear regression.

$$\text{price of house in Vancouver} = \boxed{\text{size}} \cdot x_1 + \boxed{\text{year built}} \cdot x_2 + \boxed{\text{distance to beach}} \cdot x_3 + \dots$$

Observe data: for m houses:

$$b_1 = a_{11} x_1 + a_{12} x_2 + a_{13} x_3$$

$$b_2 = a_{21} x_1 + a_{22} x_2 + a_{23} x_3$$

$$\vdots$$

$$b_m = a_{m1} x_1 + a_{m2} x_2 + a_{m3} x_3$$

$$b = A x \underbrace{+ z}_{\text{in real data there is always noise!}}$$

Can we find x?

known

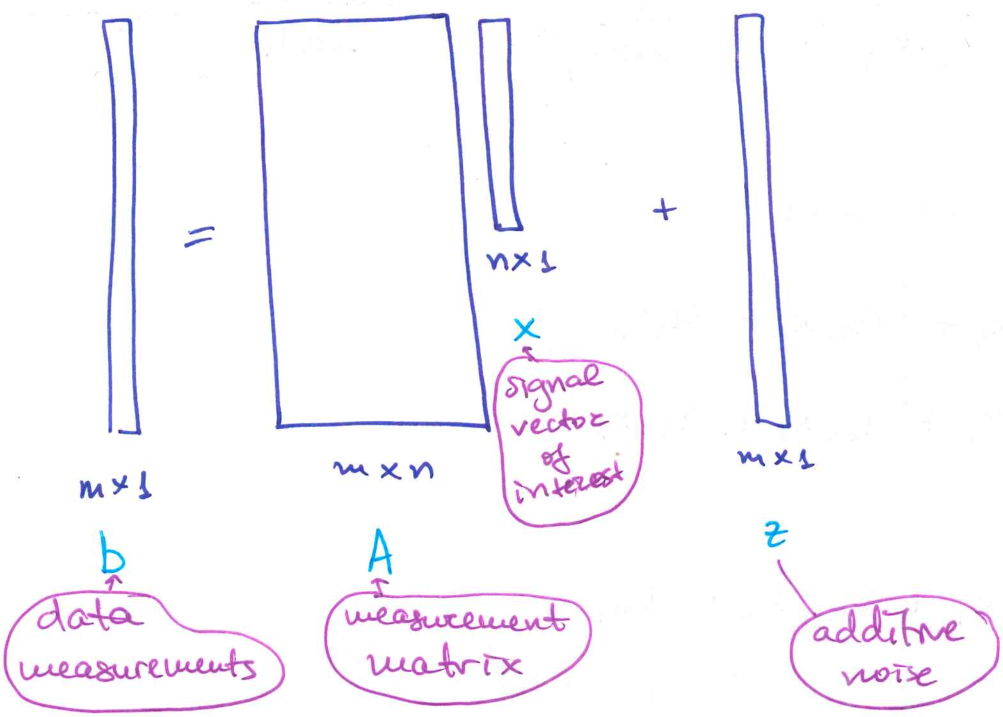
Note: We want to find a linear function

$$f(a) = a_1 x_1 + a_2 x_2 + a_3 x_3$$

that fits our data!

"Linear regression".

III.1.3. Least squares and projections onto $\mathcal{R}(A)$



$m > n$

We know: A, b

Want to learn: x.

$\mathcal{R}(A)$ is n-dimensional, $\mathcal{R}(A) \subseteq \mathbb{R}^m$
$b \in \mathbb{R}^m$

But because of the noise z, $b \notin \mathcal{R}(A)$.

So, if we try to solve $Ax = b$, there will be NO solution!

Thus, we want the "best" solution, i.e. we want to minimize the error $\|b - Ax\|$, for $x \in \mathbb{R}^n$.

Thus, we want

$$x^* = \underbrace{\arg\min_{x \in \mathbb{R}^n}}_{\text{pick the minimizer } x^* = x.} \|b - Ax\|$$

We are projecting b onto $\mathcal{R}(A)$!!! The projection vector is