Theory & Practice of Least Squares

recall $x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^n$.

the distance between vectors x & y is.

$$\|x - y\| = \sqrt{(x-y)^T(x-y)}$$

our problem: The least-squares problem in $\mathbb{R}^n$

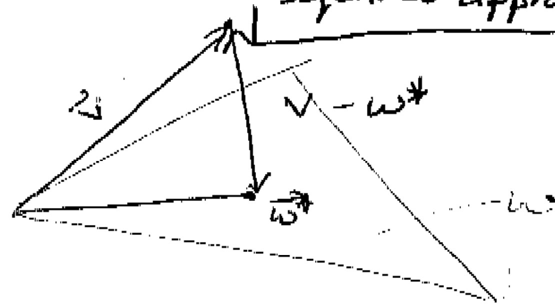
Let W be a p-dim. subspace of $\mathbb{R}^n$. Given a vector v in $\mathbb{R}^n$, find a vector w^* in W s.t.

$$\|v - w^*\| \leq \|v - w\| \quad \forall w \in W$$

The vector w^* is called the least-squares approximation to v.

Thm: Let W be a p-dim subspace of $\mathbb{R}^n$ & let $v \in \mathbb{R}^n$. Suppose $\exists w^* \in W$ s.t. $(v - w^*)^T w = 0$ $\forall w \in W$. Then w^* is the best least-squares approx to v.

□ proof. Let w be any vector in W & consider the following calculation for the distance from v to w.



$$\|v - w\|^2 = \|(v - w^*) + (w^* - w)\|^2$$

$$= (v - w^*)^T(v - w) + 2(v - w^*)^T(w^* - w) + (w^* - w)^T(w^* - w)$$

$$= \|v - w^*\|^2 + \underbrace{\|w^* - w\|^2}_{\geq 0}$$

$$\therefore \quad \|v - w\|^2 \geq \|v - w^*\| \quad ■$$

the previous proof does not guarantee the existence or uniqueness of w^*.

the previous thm forced us to check all vectors $w \in W$ to see if $(v-w^*)^T w = 0$.

Thm: Let W be a p-dim. subspace of $\mathbb{R}^n$ & let $\{u_1, u_2, \dots, u_p\}$ be a basis for W. Let $v \in \mathbb{R}^n$. Then $(v-w^*)^T w = 0$ ~~iff~~ $\forall w \in W$ iff $(v-w^*)^T u_i = 0$ for $1 \le i \le p$.

□ proof.

($\Rightarrow$): obvious since you can take $w = u_i$, $1 \le i \le p$.

($\Leftarrow$): Suppose $(v-w^*)^T u_i = 0$ for $1 \le i \le p$ & let $w \in W$ be given.

$\Rightarrow \; w = a_1 u_1 + \dots a_p u_p$

$\Rightarrow \; (v-w^*)^T w = (v-w^*)^T (a_1 u_1 + \dots + a_p u_p)$

$= a_1 (v-w^*)^T u_1 + \dots + a_p (v-w^*)^T u_p$

$= 0.$ ∎

we still don't have existence or uniqueness, but @ least we now have a finite number of conditions to check.

we will show existence thru a constructive proof.

$\boxed{\begin{matrix}3.9\\3/4\end{matrix}}$

Thm: Let W be a p-dim. subspace of $\mathbb{R}^n$ & let $v \in \mathbb{R}^n$. Then there is one & only one best least squares approx. in W to v.

(one: exists; only one: unique)

$\square$ proof.

Assume $\{u_1, \dots, u_p\}$ is an orthogonal basis to W (why can we assume this?)

Let $w^* \in W$ be s.t. $w^* = a_1 u_1 + \dots + a_p u_p$.

$\Rightarrow (v - w^*)^T u_i = 0 \iff (v - (a_1 u_1 + \dots + a_p u_p))^T u_i = 0$ for $1 \le i \le p$.

$\iff v^T u_i - a_i u_i^T u_i = 0$ for $1 \le i \le p$

$\iff a_i = \dfrac{v^T u_i}{u_i^T u_i}$ for $1 \le i \le p$.

Each a_i is well defined, so we have shown that not only does w^* exist, but we know how to construct it.

NTS: w^* is unique.

Suppose w is another best least-squares approx.

$$\|v - w\|^2 = \|(v - w^*) + (w^* - w)\|^2$$

$$= \|v - w^*\|^2 + \|\underbrace{w^* - w}_{\ge 0}\|^2 \text{ since } (v - w^*) \perp (w^* - w)$$

$\Rightarrow$ $w^* - w = \theta$.

Hence w^* exists & is unique $\blacksquare$

$\boxed{\begin{matrix}3.9\\4/4\end{matrix}}$

ex): Let $A = \begin{bmatrix} 4 & 2 & 1 \\ -2 & 0 & -1 \\ 3 & 1 & 1 \end{bmatrix}$ and define $W = R(A)$.

If $v = \begin{bmatrix} 3 \\ 0 \\ 3 \end{bmatrix}$, find w^* in W that is the best approx to v.

(A) A basis for $R(A)$: $\overset{w_1}{\begin{bmatrix} 1 \\ 0 \\ 1/2 \end{bmatrix}}$ & $\overset{w_2}{\begin{bmatrix} 0 \\ 1 \\ -1/2 \end{bmatrix}}$

(B) Find an orthogonal basis using Gram-Schmidt. $\{u_1, u_2\}$

(a) $u_1 = w_1 = \begin{bmatrix} 1 \\ 0 \\ 1/2 \end{bmatrix}$

(b) find a s.t. $u_2 = w_2 + a u_1 \perp u_1$

$$\Rightarrow u_1^T u_2 = 0 = u_1^T (w_2 + a u_1)$$

$$\Rightarrow a = -\frac{u_1^T w_2}{u_1^T u_1}$$

$$= -\frac{-1/4}{5/4}$$

$$= \frac{1}{5}$$

$$\Rightarrow u_2 = \begin{bmatrix} 0 \\ 1 \\ -1/2 \end{bmatrix} + \frac{1}{5} \begin{bmatrix} 1 \\ 0 \\ 1/2 \end{bmatrix} = \begin{bmatrix} 1/5 \\ 1 \\ -2/5 \end{bmatrix}$$

(C) Now $w^* = a_1 u_1 + a_2 u_2$

$$a_1 = \frac{v^T u_1}{u_1^T u_1} = \frac{9/2}{5/4} = \frac{18}{5}$$

$$a_2 = \frac{v^T u_2}{u_2^T u_2} = \frac{1}{30/25} = \frac{5}{6}$$

Best linear approx.

$$\Rightarrow w^* = \frac{18}{5} \begin{bmatrix} 1 \\ 0 \\ 1/2 \end{bmatrix} + \frac{5}{6} \begin{bmatrix} 1/5 \\ 1 \\ -2/5 \end{bmatrix} = \begin{bmatrix} 18/5 + 1/6 \\ 5/6 \\ 18/10 - 2/6 \end{bmatrix} = \begin{bmatrix} 113/6 \\ 5/6 \\ 22/15 \end{bmatrix}$$