Minimization Problems

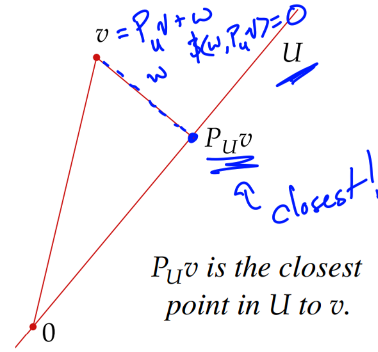
6.61 *minimizing distance to a subspace*

Suppose U is a finite-dimensional subspace of V, $v \in V$, and $u \in U$. Then

$$\|v - P_U v\| \le \|v - u\|.$$

i.e. $\operatorname{dist}(v, P_U v) \le \operatorname{dist}(v, u)$ for all $u \in U$

Furthermore, the inequality above is an equality if and only if $u = P_U v$.



$P_U v$ is the closest point in U to v.

Proof $0 \le \|P_U v - u\|^2$, so we have

$$\|v - P_U v\|^2 - \|v - P_U v\|^2 \le \|P_U v - u\|^2$$

$$\|v - P_U v\|^2 \le \|\underbrace{P_U v - u}_{\text{this is in } U}\|^2 + \|\underbrace{v - P_U v}_{\text{this is in } U^\perp}\|^2$$

$$= \|(\cancel{P_U v} - u) + (v - \cancel{P_U v})\|^2 \quad \text{(via pythagorean)}$$

$$= \|v - u\|^2$$

So, we are done!

Idea

We can use this to solve minimation problems! That is, if we want to minimize some quantity described by a distance (as induced by an inner product), then we just use a projection to get the minimizing value.

$\boxed{Ex}$ Suppose we want to find a poly. that approximates a non-poly function on a closed interval as well as possible given a constraint on its degree. For instance, suppose we want a poly of degree at most 5 to approx sin as well as possible on $[-\pi, \pi]$, in the following sense: $p(x)$, the poly, should minimize

$$\int_{-\pi}^{\pi} |\sin(x) - p(x)|^2 dx$$

Let $C([-\pi,\pi])$ denote the vector space of continuous real-valued functions on $[-\pi,\pi]$ & define the inner product $\langle f,g \rangle = \int_{-\pi}^{\pi} f(x)g(x)dx$ (this is an inner product, which you can check). Let U be the subspace of $C[-\pi,\pi]$ that is $\mathcal{P}_5(\mathbb{R})$. Then, our task becomes finding $p(x) \in U$ s.t. $\|p(x) - \sin(x)\|$ is as small as possible, since $\|p(x)-\sin(x)\| = \langle p(x)-\sin(x), p(x)-\sin(x)\rangle$

$$= \int_{-\pi}^{\pi} |p(x)-\sin(x)|^2 dx.$$

So, we just use the previous result & find our answer is $P_U(\sin(x)) = p(x)$. To actually do this computation, we would GS the basis $1, x, x^2, x^3, x^4, x^5$ of U, then compute $P_U \sin(x)$ as $P_U \sin(x) = \langle \sin(x), e_1 \rangle e_1 + \dots + \langle \sin(x), e_6 \rangle e_6$ where $e_1, \dots, e_6$ is the ONB we obtained via GS. This gives

$p(x) \approx 0.987862x - 0.155271x^3 + 0.00564312x^5$, which is a far better approx. than the Taylor poly $x - \frac{x^3}{3!} + \frac{x^5}{5!}$

Pseudoinverse

$\boxed{\text{Idea}}$ Suppose $T \in \mathcal{L}(V,W)$, where V is an inner product space, and suppose we want to solve $Tx = b$ for some $b \in W$. If T is invertible, $x = T^{-1}b$ is the unique solution. If not, then there are either many solutions or none. If there are many, then we could ask for the "best one" in the sense of finding $\|x\|$ as small as possible or, if W is also an inner product space, finding x s.t. $\|Tx - b\|$ is as small as possible if there are no solutions.

6.67 *restriction of a linear map to obtain a one-to-one and onto map*

Suppose V is finite-dimensional and $T \in \mathcal{L}(V,W)$. Then $T|_{(\operatorname{null} T)^\perp}$ is a one-to-one map of $(\operatorname{null} T)^\perp$ onto range T. ie. $T|_{(\operatorname{null} T)^\perp} \in \mathcal{L}((\operatorname{null} T)^\perp, \operatorname{range} T)$ is invertible

$\boxed{\text{Proof}}$ (sketch) If $v \in (\operatorname{null} T)^\perp$ & $T|_{(\operatorname{null} T)^\perp} v = 0$, then $v \in \operatorname{null} T \cap (\operatorname{null} T)^\perp = \{0\}$ & $v = 0$. So, $T|_{(\operatorname{null} T)^\perp}$ has trivial null space & is injective.
For surjectivity, if $\hat{w} = Tv \in \operatorname{range} T$, then write $v = u + w$ where $u \in \operatorname{null} T$ & $w \in (\operatorname{null} T)^\perp$. So, $T|_{(\operatorname{null} T)^\perp} w = Tw = Tw + 0 = Tw + Tu = Tv = \hat{w}$ & $\hat{w} \in \operatorname{range}(T|_{(\operatorname{null} T)^\perp})$. Hence, $\operatorname{range} T \subseteq \operatorname{range} T|_{(\operatorname{null} T)^\perp}$ & the other direction is immediate, so they are equal.

6.68 definition: *pseudoinverse*, $T^\dagger$

Suppose that V is finite-dimensional and $T \in \mathcal{L}(V,W)$. The *pseudoinverse* $T^\dagger \in \mathcal{L}(W,V)$ of T is the linear map from W to V defined by

$$\underbrace{T^\dagger}_{\text{this is the pseudoinverse of } T} w = (T|_{(\operatorname{null} T)^\perp})^{-1} P_{\operatorname{range} T}\, w$$

for each $w \in W$.

6.69 *algebraic properties of the pseudoinverse*

Suppose V is finite-dimensional and $T \in \mathcal{L}(V, W)$.

(a) If T is invertible, then $T^\dagger = T^{-1}$.

(b) $TT^\dagger = P_{\operatorname{range} T}$ = the orthogonal projection of W onto range T.

(c) $T^\dagger T = P_{(\operatorname{null} T)^\perp}$ = the orthogonal projection of V onto $(\operatorname{null} T)^\perp$.

Projections are identities when restricted (i.e. if $P_u \in \mathcal{L}(V)$ w/ $u \subseteq V$, then $P_u|_u = I$)

$\boxed{\text{Proof}}$ (a) If T is invertible, then $\operatorname{null} T = \{0\}$ & $(\operatorname{null} T)^\perp = V$, as well as range $T = W$.

So, $T^\dagger = \left(T|_{(\operatorname{null} T)^\perp}\right)^{-1} P_{\operatorname{range} T} = \left(T|_V\right)^{-1} P_W = (T)^{-1} I = T^{-1}$

(b) $TT^\dagger = T\left(T|_{(\operatorname{null} T)^\perp}\right)^{-1} P_{\operatorname{range} T}$, so if $w \in \operatorname{range} T$, we get $TT^\dagger w = T\left(T|_{(\operatorname{null} T)^\perp}\right)^{-1} P_{\operatorname{range} T} w = T\left(T|_{(\operatorname{null} T)^\perp}\right)^{-1} w = w = P_{\operatorname{range} T} w$

Likewise, if $w \in (\operatorname{range} T)^\perp$, then $TT^\dagger w = T\left(T|_{(\operatorname{null} T)^\perp}\right)^{-1} P_{\operatorname{range} T} w = 0 = P_{\operatorname{range} T} w$

So, $TT^\dagger = P_{\operatorname{range} T}$ since they agree on both range T & $(\operatorname{range} T)^\perp$.

(c) (basically the same process as part (b))

6.70 *pseudoinverse provides best approximate solution or best solution*

Suppose V is finite-dimensional, $T \in \mathcal{L}(V, W)$, and $b \in W$.

(a) If $x \in V$, then

$$\|T(T^\dagger b) - b\| \leq \|Tx - b\|,$$

with equality if and only if $x \in T^\dagger b + \operatorname{null} T$.

i.e. minimize $\|Tx - b\|$ when there is no solution

(b) If $x \in T^\dagger b + \operatorname{null} T$, then

$$\|T^\dagger b\| \leq \|x\|,$$

with equality if and only if $x = T^\dagger b$.

i.e. minimize $\|x\|$ when there are many solutions