18 quadratic opt 2; Hilbert spaces

Tuesday, November 3, 2020 4:12 AM

Spherical constraints

Last time we discussed quadratic optimization in the unconstrained case and when the constraints are affine or linear. Today we'll start with constraining solutions to be on the unit sphere (e.g. a direction in $\mathbb{R}^3$).

Consider the problem:

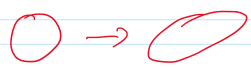
$$\text{maximize } x^T A x$$
$$\text{subject to } x^T x = 1, \quad x \in \mathbb{R}^n, \quad A \text{ a symmetric matrix}$$

Recall Rayleigh-Ritz:

$$\max_{x^T x = 1} x^T A x = \lambda_1(A) \quad \text{largest eigenvalue}$$

$$\min_{x^T x = 1} x^T A x = \lambda_n(A) \quad \text{smallest eigenvalue.}$$

What if instead of the unit sphere, we want an ellipsoid.



Consider:

$$\text{minimize } x^T A x$$
$$\text{subject to } x^T B x = 1, \quad B \text{ symmetric pos. def.}$$

Ideas? Let's diagonalize. $B = QDQ^T$, Q orthogonal, $D = \operatorname{diag}(d_1, \dots, d_n)$, $d_i > 0$.

Let $B^{\frac{1}{2}} = Q \operatorname{diag}(\sqrt{d_1}, \dots, \sqrt{d_n}) Q^T$

and $B^{-\frac{1}{2}} = Q \operatorname{diag}\left(\frac{1}{\sqrt{d_1}}, \dots, \frac{1}{\sqrt{d_n}}\right) Q^T$.

Let $x = B^{-\frac{1}{2}} y$. Then our system becomes

$$\text{minimize } y^T B^{-\frac{1}{2}} A B^{-\frac{1}{2}} y \qquad (B^{-\frac{1}{2}} A B^{-\frac{1}{2}} \text{ is symm}).$$
$$\text{subject to } y^T y = 1.$$

Thus, we can convert ellipsoid constraints into spherical ones

Combining spherical and linear constraints

Consider:

$$\text{minimize } x^T A x$$
$$\text{subj. to } x^T x = 1, \quad C^T x = 0, \quad x \in \mathbb{R}^n, \quad C \in \mathbb{R}^{n \times p}, \quad A \in \mathbb{R}^{n \times n} \text{ symmetric}$$

Again, let $C = Q^T \begin{pmatrix} R & S \\ 0 & 0 \end{pmatrix} \Pi$, Q orthogonal, R invertible upper triangular.

Let $x = Q^T \begin{pmatrix} y \\ z \end{pmatrix}$. Then $C^T x = 0$

$$\Rightarrow \Pi^T \begin{pmatrix} R^T & 0 \\ S^T & 0 \end{pmatrix} Q x = \Pi^T \begin{pmatrix} R^T & 0 \\ S^T & 0 \end{pmatrix} \begin{pmatrix} y \\ z \end{pmatrix} = 0$$

$$\Rightarrow y = 0, \quad x = Q^T \begin{pmatrix} 0 \\ z \end{pmatrix}.$$

So we get

$$\text{minimize } (y^T \; z^T) \, Q A Q^T \begin{pmatrix} y \\ z \end{pmatrix}$$
$$\text{subj. to } z^T z = 1, \quad z \in \mathbb{R}^{n-r}$$

Let $QAQ^T = (G_1 \;\; G_{12})$

So we get minimize $(y^T \ z^T) \, QAQ^T \binom{y}{z}$
subj. to $z^T z = 1$, $z \in \mathbb{R}^{n-r}$
$y = 0$, $y \in \mathbb{R}^r$.

Let $QAQ^T = \begin{pmatrix} G_1 & G_{12} \\ G_{12}^T & G_{22} \end{pmatrix}$

Can rewrite minimize $z^T G_{22} z$
subject to $z^T z = 1$, $z \in \mathbb{R}^{n-r}$. ← standard form again. just an eigenvalue problem

Chapter 12: Basics of Hilbert spaces

We are now going to move on to more general nonlinear optimization. To do so, we need to first review some Hilbert space theory.

Def. 12.1 A (complex) Hermitian space $\langle E, \varphi \rangle$ which is a complete normed vector space under the norm $\|u\| = \sqrt{\varphi(u,u)}$ is called a Hilbert space. A real Euclidean space $\langle E, \varphi \rangle$ which is complete under the norm $\|u\| = \sqrt{\varphi(u,u)}$ is called a real Hilbert space.

Recall: (1) metric spaces have a distance $d(x,y) \geq 0$, $d(x,y) = 0 \iff x = y$, $d(x,y) = d(y,x)$, $d(x,z) \leq d(x,y) + d(y,z)$ (e.g. $d(x,y) = \begin{cases} 0 & \text{if } x = y \\ 1 & \text{otherwise} \end{cases}$)

(2) Banach spaces are complete and have a norm $\|x\| \geq 0$, $\|x\| = 0 \iff x = 0$, $\|x+y\| \leq \|x\| + \|y\|$, $\alpha\|x\| = \|\alpha x\|$

(3) Hilbert spaces have an inner product φ, and are complete under the norm induced by φ.

Ex. 12.1 The space ℓ^2 of all countably infinite sequences $x = (x_i)_{i \in \mathbb{N}}$ of complex numbers s.t. $\sum_{i=0}^{\infty} |x_i|^2 < \infty$ is a Hilbert space under $\varphi((x_i)_{i \in \mathbb{N}}, (y_i)_{i \in \mathbb{N}}) = \sum_{i=0}^{\infty} x_i \overline{y_i}$.

Ex. 12.2 The set $C^\infty[a,b]$ of smooth functions $f: [a,b] \to \mathbb{C}$ is a Hermitian space under the Hermitian form $\langle f, g \rangle = \int_a^b f(x) \overline{g(x)} \, dx$.
Is it a Hilbert space? No, because not complete, though we can construct its completion $L^2([a,b])$ of Lebesgue integrable functions on $[a,b]$.

Thm 12.1 Given a Hermitian space $(E, \langle\,,\rangle)$ (resp. Euclidean space), there is a Hilbert space $(E_h, \langle\,,\rangle_h)$ and a linear map $\varphi: E \to E_h$ s.t.

$$\langle u, v \rangle = \langle \varphi(u), \varphi(v) \rangle_h$$

$\forall u, v \in E$, and $\varphi(E)$ is dense in E_h. Furthermore, E_h is unique up to isomorphism.

proof sketch: Use polarity equation $\langle u, v \rangle = \frac{1}{2}(\|u+v\|^2 - \|u\|^2 - \|v\|^2)$ to work in the associated Banach space instead, and make use of continuity and limits. $\blacksquare$

Remark: $\forall$ finite-dim Hermitian (resp. Euclidean) space E, with $\dim(E) = n$, we have an orthonormal basis so E is isomorphic to $\mathbb{C}^n$ (resp. $\mathbb{R}^n$), and the inner product is given by $\langle (x_1, \dots, x_n), (y_1, \dots, y_n) \rangle = \sum_{i=1}^n x_i \overline{y_i}$.

the inner product is given by $\langle (x_1,\dots,x_n), (y_1,\dots,y_n)\rangle = \sum_{i=1}^{n} x_i \overline{y_i}$.

Furthermore, every subspace $W \subseteq E$ has an orthogonal complement $W^\perp$, and the inner prod induces a natural duality between E and E^*, the space of linear forms on E. (actually between $\bar{E}$ and E^*, but $E = \bar{E}$)

Can we say something similar about infinite dim Hermitian / Euclidean spaces?
What goes wrong? E may not be guaranteed to have an orthonormal basis.

Define a Hilbert basis as an orthogonal family whose linear span is dense in E, i.e. every $v \in E$ is the limit of a sequence of finite combinations of vectors from the Hilbert basis.

e.g. If $(u_k)_{k\in K}$ is a Hilbert basis, $v \text{ "="} \sum_{k\in K} c_k u_k$, where $c_k = \dfrac{\langle v, u_k\rangle}{\|u_k\|}$ are the Fourier coefficients.

Def. A.2 $\quad v \text{ "="} \sum_{k\in K} u_k$ if $\forall \varepsilon > 0$, $\exists$ finite $I \subseteq K$ s.t. $\left\| x - \sum_{j\in J} u_j \right\| < \varepsilon$ for every finite J s.t. $I \subseteq J \subseteq K$.

Def A.3 Given any non-empty index set K, the space $\ell^2(K)$ is the set of all sequences $(z_k)_{k\in K}$, where $z_k \in \mathbb{C}$, s.t. $(|z_k|^2)_{k\in K}$ is summable (i.e. $\sum_{k\in K} |z_k|^2 < \infty$)

Remark: (1) When $|K| = n < \infty$, $\ell^2(K)$ is isomorphic to $\mathbb{C}^n$.
(2) When $K = \mathbb{N}$, $\ell^2(\mathbb{N}) = \ell^2$ above, a Hilbert space.

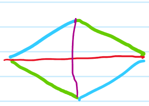
Thm (Riesz-Fischer) For every Hilbert space E, there is some nonempty set K s.t. E is isomorphic to the Hilbert space $\ell^2(K)$. More specifically, for any Hilbert basis $(u_k)_{k\in K}$ of E, the maps $f: \ell^2(K) \to E$ and $g: E \to \ell^2(K)$ defined s.t. $f((\lambda_k)_{k\in K}) = \sum_{k\in K} \lambda_k u_k$ and $g(u) = \left(\dfrac{\langle u, u_k\rangle}{\|u_k\|^2}\right)_{k\in K} = (c_k)_{k\in K}$, are bijective linear isometries s.t. $g\circ f = id$ and $f \circ g = id$.

i.e. All Hilbert spaces are isomorphic to $\ell^2(K)$, for some K.

We want to show that closed subspaces of a Hilbert space have an orthogonal complement, and that we have a notion of duality, though we'll need to define the dual E' of continuous linear maps.

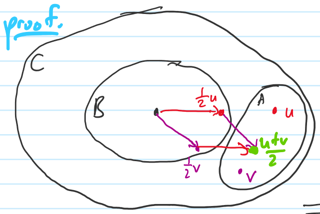
Prop. 12.1/2 If E is a Hermitian space, for any two vectors $u, v \in E$,
$$\|u+v\|^2 + \|u-v\|^2 = 2(\|u\|^2 + \|v\|^2) \quad \text{(parallelogram law)}$$

proof. Left as exercise



Prop 12.2/3 If E is a Hermitian space, for $d, \delta \in \mathbb{R}$ s.t. $0 \le \delta < d$, let $B = \{u \in E \mid \|u\| < d\}$ and $C = \{u \in E \mid \|u\| \le d + \delta\}$.
For any convex set A s.t. $A \subseteq C - B$, we have
$$\|v-u\| \le \sqrt{12 d\delta}, \quad \forall u, v \in A.$$

proof.



Since A is convex, $\frac{u+v}{2} \in A$.
$\Rightarrow \left\|\frac{u+v}{2}\right\| \ge d$

By parallelogram law, $\left\|\frac{u+v}{2}\right\|^2 + \left\|\frac{u-v}{2}\right\|^2 = \frac{1}{2}\left(\|u\|^2 + \|v\|^2\right)$

$\Rightarrow \frac{1}{4}\|u-v\|^2 = \frac{1}{2}\left[\|u\|^2 + \|v\|^2\right] - \left\|\frac{u+v}{2}\right\|^2 \le (d+\delta)^2 - d^2 = 2d\delta + \delta^2 \le 3d\delta$

$\Rightarrow \|v-u\| \le \sqrt{12 d\delta}$. $\square$

Def. 12.2 If $a \in E$, $X \subseteq E$, then the *distance* $d(a, X) = \inf_{b \in X} d(a,b)$, where d is the metric on E.

The *diameter* $\delta(X) = \sup\{d(a,b) \mid a, b \in X\}$. $\quad (\delta(X) \in \mathbb{R} \cup \{\infty\})$

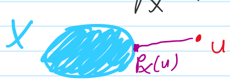
Prop. 12.3/4 Let E be a metric space

(1) For any subset $X \subseteq E$, $\delta(X) = \delta(\bar{X})$

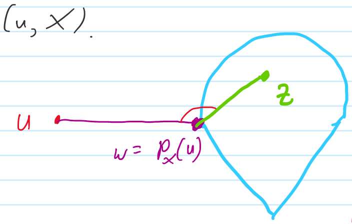
(2) If E is a complete metric space, $\forall$ sequence (F_n) of closed nonempty subsets of E s.t. $F_{n+1} \subseteq F_n$, if $\lim_{n\to\infty} \delta(F_n) = 0$, then $\bigcap_{n=1}^{\infty} F_n$ is a single point.

Prop 12.4/5 (*Projection lemma*) Let E be a Hilbert space.

(1) For any nonempty, convex, and closed subset $X \subseteq E$, $\forall u \in E$, $\exists$ a unique vector $p_X(u) \in U$ s.t. $\|u - p_X(u)\| = \inf_{v \in X} \|u - v\| = d(u, X)$.



(2) The vector $p_X(u)$ is the unique $w \in E$ s.t.
$$w \in X \text{ and } \operatorname{Re}\langle u-w, z-w\rangle \le 0 \quad \forall z \in X$$



i.e. the "angle" is at least 90°.
i.e. X and u lie on diff sides of the "tangent space"

(3) If X is a nonempty closed subspace of E, then $p_X(u)$ is the unique vector $w \in E$ s.t.
$$w \in X \text{ and } \langle u - w, z\rangle = 0 \quad \forall x \in X.$$

proof. (1) Let $d = d(u, X)$. Define a sequence X_n by
$$X_n = \{v \in X \mid \|u - v\| \le d + \tfrac{1}{n}\}. \quad \text{(clearly nonempty subsets of } X)$$

Then $X_{n+1} \subseteq X_n$, $X_n \neq \emptyset$, $X_n \subseteq X$.

Also, $\sup\{\|z - v\| \mid v, z \in X_n\} \le \sqrt{12\, d/n}$, $\quad$ (letting $B = \{v \in E \mid \|u - v\| < d\}$, $C = \{v \in E \mid \|u-v\| \le d + \frac{1}{n}\}$, $A = X_n$)

$\Rightarrow \bigcap_{n \ge 1} X_n$ contains at most one pt.

Now consider a seq $(w_n)_{n\ge1}$ by picking an arbitrary $w_n \in X_n$ for every $n \ge 1$.

Now consider a seq $(w_n)_{n\geq 1}$ by picking an arbitrary $w_n \in X_n$ for every $n \geq 1$.

Given any $\varepsilon > 0$, let $N > \frac{12d}{\varepsilon^2}$, $\|w_m - w_n\| \leq \sqrt{\frac{12d}{N}} < \varepsilon$, so (w_n) is Cauchy.

Let $w = \lim_{n\to\infty} w_n \in X$ as X is closed.

Note that $\|u-w\| \leq \underbrace{\|u-w_n\|}_{\leq d + \frac{1}{n}} + \|w_n - w\|$.

So given any $\varepsilon > 0$, can choose n sufficiently large that $\frac{1}{n} < \frac{\varepsilon}{2}$ and $\|w_n - w\| \leq \frac{\varepsilon}{2}$,

$\Rightarrow \|u-w\| \leq d + \varepsilon$.

$\Rightarrow \|u-w\| \leq d$

$\Rightarrow \|u-w\| = d$. (since $d = d(u,X)$, and $w \in X$)

Of course, any $z \in X$ s.t. $\|u-z\| = d$ belongs to every X_n, so $z = w$, proving uniqueness $\boxtimes$

(2) Let $z \in X$. Then $v = (1-\lambda)p_X(u) + \lambda z \in X$ $\forall\, 0 \leq \lambda \leq 1$, since X is convex.

By def. $\|u-v\| \geq \|u - p_X(u)\|$, so

$$\|u-v\|^2 = \|u - p_X(u) - \lambda(z - p_X(u))\|^2$$

$$= \|u - p_X(u)\|^2 + \lambda^2\|z - p_X(u)\|^2 - 2\lambda \operatorname{Re}\langle u - p_X(u), z - p_X(u)\rangle.$$

$$\Rightarrow \operatorname{Re}\langle u - p_X(u), z - p_X(u)\rangle = \frac{1}{2\lambda}\underbrace{\left(\|u-p_X(u)\|^2 - \|u-v\|^2\right)}_{\leq 0} + \underbrace{\frac{\lambda}{2}\|z - p_X(u)\|^2}_{\text{Can choose } \lambda \text{ small}}$$

If $\|u-p_X(u)\|^2 - \|u-v\|^2 < 0$, then choose $\lambda > 0$ small so $\operatorname{Re}\langle u-p_X(u), z-p_X(u)\rangle < 0$.

Else, take limit as $\lambda \to 0$, so $\|u-p_X(u)\|^2 - \|u-v\|^2 = 0$

$\Rightarrow \operatorname{Re}\langle u - p_X(u), z - p_X(u)\rangle \leq 0$.

Conversely, if $w \in X$ satisfies $\operatorname{Re}\langle u-w, z-w\rangle \leq 0$ $\forall z \in X$, then

$$\|u-z\|^2 = \|u-w\|^2 + \|z-w\|^2 - 2\operatorname{Re}\langle u-w, z-w\rangle \geq \|u-w\|^2$$

$\Rightarrow \|u-w\| = d(u,X) = d \Rightarrow w = p_X(u)$. $\boxtimes$

(3) If X is a subspace of E and $w \in X$, then $\{z - w \mid z \in X\} = X$, so

$w \in X$ and $\operatorname{Re}\langle u-w, z\rangle \leq 0$ $\forall z \in X$.

But $\operatorname{Re}\langle u-w, z\rangle \leq 0$ and $\operatorname{Re}\langle u-w, -z\rangle \leq 0$ (since if $z \in X$, $-z \in X$)

$\Rightarrow \operatorname{Re}\langle u-w, z\rangle = 0$. $\forall z \in X$.

Further, $\operatorname{Re}\langle u-w, iz\rangle \leq 0 \Rightarrow -\operatorname{Im}\langle u-w, z\rangle = 0$

$\Rightarrow \langle u-w, z\rangle = 0$ $\forall z \in X$. $\boxtimes$

Further, $\text{Re}\langle u-w, iz\rangle \geq 0 \Rightarrow -\text{Im}\langle u-w, z\rangle = 0$

$$\Rightarrow \langle u-w, z\rangle = 0 \quad \forall z \in X. \qquad \boxed{/}$$

Def. 12.3 The vector $p_X(u)$ is called the *projection of u onto X*, and the map $p_X : E \to X$ is called the *projection of E onto X*.

Prop. 12.5/6 Let E be a Hilbert space. For any nonempty convex closed subset $X \subseteq E$, the map $p_X : E \to X$ is continuous. In fact, p_X satisfies the Lipschitz condition

$$\|p_X(v) - p_X(u)\| \leq \|v-u\| \quad \forall u, v \in E.$$

proof. $\forall u, v \in E$, let $x = p_X(u) - u$, $y = p_X(v) - p_X(u)$, $z = v - p_X(v)$.

Then $v - u = x + y + z$.

$$\Rightarrow \text{Re}\langle x, y\rangle = \text{Re}\langle p_X(u) - u, p_X(v) - p_X(u)\rangle = -\text{Re}\langle u - p_X(u), p_X(v) - p_X(u)\rangle \geq 0$$

$\text{Re}\langle z, y\rangle \geq 0.$

$$\begin{aligned}\Rightarrow \|v-u\|^2 &= \|x+y+z\|^2 \\ &= \|x+z\|^2 + \|y\|^2 + 2\,\text{Re}\langle x, y\rangle + 2\,\text{Re}\langle z, y\rangle \\ &\geq \|y\|^2 = \|p_X(v) - p_X(u)\|^2\end{aligned}$$

$$\Rightarrow \|p_X(v) - p_X(u)\| \leq \|v-u\|$$

$$\Rightarrow p_X \text{ is continuous} \qquad \boxed{/}$$

Prop. 12.6/7 Let E be a Hilbert space.

(1) For any closed subspace $V \subseteq E$, $E = V \oplus V^\perp$, and the map $p_V : E \to V$ is linear and continuous.

(2) For any $u \in E$, $p_V(u)$ is the unique vector $w \in E$ s.t. $w \in V$ and $\langle v - w, z\rangle = 0 \quad \forall z \in V$.

proof. (1) By Prop. 12.4(3), $\langle u - p_V(u), v\rangle = 0 \quad \forall v \in V$.

Thus, $u - p_V(u) \in V^\perp \quad \forall u \in E$.

But $u = p_V(u) + u - p_V(u) \quad \forall u \in E$, so $E = V + V^\perp$.

And $V \cap V^\perp = \{0\}$ because $\langle\,,\,\rangle$ is pos. def. so $E = V \oplus V^\perp$.

We already proved continuity, so only need to prove linearity of $p_V : E \to V$.

$$p_V(\lambda u + \mu v) - (\lambda p_V(u) + \mu p_V(u)) = p_V(\lambda u + \mu v) - (\lambda u + \mu v) + \lambda(u - p_V(u)) + \mu(v - p_V(u))$$

$$\underbrace{p_V(\lambda u+\mu v)}_{\in V} - (\underbrace{\lambda p_V(u)}_{\in V} + \underbrace{\mu p_V(u)}_{\in V}) = \underbrace{p_V(\lambda u+\mu v) - (\lambda u+\mu v)}_{\in V^T} + \underbrace{\lambda(u - p_V(u))}_{\in V^T} + \underbrace{\mu(v - p_V(u))}_{\in V^T}$$

$\Rightarrow \quad = 0.$ since $V \cap V^\perp = \{0\}$.

(2) Repeat of Prop 12.4 (3) $\boxed{\diagup}$

Remark: If $p_V : E \to V$ is linear, then V is a subspace of E. Thus, if V is a closed convex subset of E, p_V is linear iff V is a subspace.

Ex. Let's go back to least squares solution of $Ax=b$ again.

Let $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$. Find $\min_{x\in\mathbb{R}^n} \|Ax-b\|_2$.

Equiv to finding $x \in \mathbb{R}^n$ s.t. $\|Ax-b\| = \inf_{y\in\mathbb{R}^n} \|Ay-b\|$

$\iff$ is there some $z \in \mathrm{Im}(A)$ s.t. $\|z-b\| = d(b, \mathrm{Im}(A))$.

$\mathrm{Im}(A)$ is a closed subspace of $\mathbb{R}^m$, (all subspaces of finite $\mathbb{R}^m$ are closed)

so $\exists$ unique $z \in \mathrm{Im}(A)$ s.t. $\|z-b\| = \inf_{y\in\mathbb{R}^n} \|Ay-b\|$.

$\Rightarrow \exists x \in \mathbb{R}^n$ (not necessarily unique) s.t. $Ax = z$.

By Prop. 12.6, $\langle z-b, w\rangle = 0 \quad \forall w \in \mathrm{Im}(A)$

$\iff \langle Ax-b, Ay\rangle = 0 \quad \forall y \in \mathbb{R}^n$

$\iff \langle A^T(Ax-b), y\rangle = 0 \quad \forall y \in \mathbb{R}^n$

$\iff A^T(Ax-b) = 0$

$\iff A^TAx = A^Tb.$ (normal eq., can be solved with pseudoinverse)

Cor 12.1 For any continuous nonnull linear map $h : E \to \mathbb{C}$,

$$H = \mathrm{Ker}(h) = \{u \in E \mid h(u) = 0\} = h^{-1}(0)$$

is a closed hyperplane H, and thus, $H^\perp$ is a subspace of dim 1 s.t. $E = H \oplus H^\perp$.

proof. Let $y_0 \in H^\perp$, $y_0 \neq 0$. WLOG, say $h(y_0) = 1$.

Then $\forall x \in E$, $h(x - h(x)y_0) = h(x) - h(x)h(y_0) = 0$

$\Rightarrow x - h(x)y_0 \in H.$

$\Rightarrow \langle x - h(x)y_0, y_0\rangle = 0$

$\Rightarrow h(x)\|y_0\| = \langle x, y_0\rangle \quad \forall x \in H$

$\Rightarrow h(x) = \langle x, \frac{y_0}{\|y_0\|}\rangle \quad \forall x \in H.$ (so h is completely defined by $\frac{y_0}{\|y_0\|}$)

Then for any $u \in E$, $\langle u - (u - \langle u, \frac{y_0}{\|y_0\|}\rangle y_0), z\rangle \quad \forall z \in H$

$= \langle u, \frac{y_0}{\ldots}\rangle \cdot \langle y_0, z\rangle = 0.$

Then for any $u \in E$, $\langle u - (u - \langle u, \frac{y_0}{\|y_0\|} \rangle y_0), z \rangle \quad \forall z \in H$

$$= \langle u, \frac{y_0}{\|y_0\|} \rangle \cdot \langle y_0, z \rangle = 0.$$

And $h(u - h(u) y_0) = 0$, so $u - h(u) y_0 \in H$.

So $P_H(u) = u - h(u) y_0$, but this proves uniqueness of y_0, so $\dim(H^\perp) = 1$. ∎

This leads us to define a slightly different notion of duality.

Def. 12.4 Given a Hilbert space E, we define the dual space E' of E as the vector space of all continuous linear forms $h: E \to \mathbb{C}$. Maps in E' are also called bounded linear operators, bounded linear functionals, or simply operators or functionals.

Def. $\forall u, v \in E$, $\varphi_u^l : E \to \mathbb{C}$ is given by $\varphi_u^l(v) = \overline{\langle u, v \rangle}$

$\varphi_v^r : E \to \mathbb{C}$ is given by $\varphi_v^r(u) = \langle u, v \rangle$.

Note: $\varphi_u^l = \varphi_u^r \equiv \varphi_u \in E'$.

Prop. 12.7/8 (Riesz representation theorem) Let E be a Hilbert space. Then the map $\flat: E \to E'$ defined s.t. $\flat(v) = \varphi_v$, is semilinear, continuous and bijective. Furthermore, for any continuous linear map $\psi \in E'$, if $u \in E$ is the unique vector s.t. $\psi(v) = \langle v, u \rangle \quad \forall v \in E$,

then $\|\psi\| = \|u\|$, where $\|\psi\| = \sup \left\{ \frac{|\psi(v)|}{\|v\|}, v \in E, v \neq 0 \right\}$

proof. Same as proof of Theorem 13.6 (Vol 1), except we need a different argument for surjectivity of $\flat: E \to E'$, as E may not be finite dimensional.

For any nonnull linear operator $h \in E'$, the hyperplane $H = \operatorname{Ker}(h) = h^{-1}(0)$ is a closed subspace of E, and $H^\perp$ is a subspace of dim 1 s.t. $E = H \oplus H^\perp$.

Then picking any $w \in H^\perp$, $H = \operatorname{Ker} \varphi_w$, where $\varphi_w(u) = \langle u, w \rangle$.

Thus, since φ_w and h define the same hyperplane, $\exists \lambda \in \mathbb{C}$ s.t. $h = \lambda \varphi_w$.

$\Rightarrow h = \varphi_{\bar{\lambda} w}$, which proves surjectivity of $\flat: E \to E'$.

By Cauchy-Schwarz, $|\psi(v)| = |\langle v, u \rangle| \leq \|v\| \|u\|$,

so by def. $\|\psi\| \leq \|u\|$.

But $\psi = 0$ iff $u = 0$, so assume $u \neq 0$. Then

$$\|u\|^2 = \langle u, u \rangle = \psi(u) \leq \|\psi\| \|u\|$$

$\Rightarrow \|u\| \leq \|\psi\|$

$\Rightarrow \|\psi\| = \|u\|$. ∎

$$\Rightarrow \|u\| \le \|\Psi\|$$
$$\Rightarrow \|\Psi\| = \|u\|.$$

Define: For $\varphi : E \times E \to \mathbb{C}$ a sesquilinear map, $\displaystyle \|\varphi\| = \sup_{\substack{\|x\|\le 1\\ \|y\|\le 1}} \{ |\varphi(x,y)| \}$.

Prop. 12.8/9 Given a Hilbert space E, for every continuous sesquilinear map $\varphi : E \times E \to \mathbb{C}$, there is a unique continuous linear map $f_\varphi : E \to E$ s.t. $\varphi(u,v) = \langle u, f_\varphi(v)\rangle$ for all $u, v \in E$.

We also have $\|f_\varphi\| = \|\varphi\|$. If φ is Hermitian, then f_φ is self-adjoint, that is $\langle u, f_\varphi(v)\rangle = \langle f_\varphi(u), v\rangle$ for all $u, v \in E$.

(Idea: $\varphi(x,y) = x^* A y$, $f : E \times E \to \mathbb{C}$.
$f_\varphi : E \to E$ by $\varphi(x,y) = x^T(Ay)$. i.e. $f_\varphi(y) = Ay$)

Prop. 12.9/10 Given a Hilbert space E, for every continuous linear map $f : E \to E$, there is a unique continuous linear map $f^* : E \to E$ s.t.
$$\langle f(u), v\rangle = \langle u, f^*(v)\rangle \quad \text{for all } u, v \in E,$$
and we have $\|f^*\| = \|f\|$. The map f^* is called the *adjoint* of f.

Thm 12.2/11 (Farkas–Minkowski Lemma in Hilbert spaces)
Let $(V, \langle\,,\rangle)$ be a real Hilbert space. For any finite sequence of vectors $(a_1, \dots, a_m)$ with $a_i \in V$, if C is the polyhedral cone $C = \text{cone}(a_1, \dots, a_m)$, for any vector $b \in V$, we have $b \notin C$ iff $\exists$ a vector $u \in V$ s.t.
$$\langle a_i, u\rangle \ge 0 \quad i = 1, \dots, m \quad \text{and} \quad \langle b, u\rangle < 0$$

Equivalently, $b \in C$ iff $\forall\, u \in V$,
$$\text{if } \langle a_i, u\rangle \ge 0 \quad i = 1, \dots, m, \text{ then } \langle b, u\rangle \ge 0.$$

Recall: $\text{cone}((S_n)) = \left\{ \sum_{i=1}^{k} \lambda_i u_i,\ u_i \in S,\ \lambda_i \in \mathbb{R},\ \lambda_i \ge 0 \right\}$

proof. Claim: If C is a closed, nonempty, subset of a Hilbert space V, and $b \in V$ is any vector s.t. $b \notin C$, then $\exists$ some $u \in V$ and infinitely many scalars $\alpha \in \mathbb{R}$ s.t.
$$\langle v, u\rangle > \alpha \quad \forall v \in C$$
$$\langle b, u\rangle < \alpha.$$

proof. By the projection lemma, $\exists\, c = p_C(b) \in C$ s.t.
$$\|b - c\| = \inf_{v \in C} \|b - v\| > 0$$

$$\|b-c\| = \inf_{v\in C} \|b-v\| > 0$$

$$\langle b-c, v-c\rangle \le 0 \quad \text{for all } v \in C.$$

Equivalently,

$$\|b-c\| = \inf_{v\in C} \|b-v\| > 0$$

$$\langle v-c, c-b\rangle \ge 0 \quad \text{for all } v \in C$$

$$\Rightarrow (v, c-b) \ge \langle c, c-b\rangle > \langle b, c-b\rangle.$$

Now we simply pick $u = c-b$ and any α s.t.

$$\langle c, c-b\rangle > \alpha > \langle b, c-b\rangle \qquad \square$$

Now assume $b \notin C$. Then $\exists\, \alpha \in \mathbb{R}$ s.t.

$$\langle v, u\rangle > \alpha \quad \forall v \in C$$

$$\langle b, u\rangle < \alpha.$$

But C is a polyhedral cone containing 0, so $\alpha < 0$. (otherwise, $\langle 0, u\rangle = 0$)

Then $\forall v \in C$, $\langle v, u\rangle > \frac{\alpha}{\lambda}$ for every $\lambda > 0$ (since if $v\in C$, $\lambda v \in C$)

Thus, $\langle v, u\rangle \ge 0$.

Since $a_i \in C$ for $i = 1, \dots, m$, $\langle a_i, u\rangle \ge 0$.