19 Farkas Minkowski; quadratic functionals

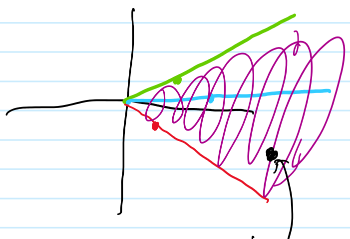
Monday, November 16, 2020 1:56 PM

Let's start by reminding ourselves of the Farkas-Minkowski Lemma, which we covered super quickly before reading week.

Let V be a real Hilbert space.

Given a finite sequence of vectors $(a_1, \dots, a_m)$, $a_i \in V$, let C be the polyhedral cone

$$C = \operatorname{cone}(a_1, \dots, a_m) = \left\{ \sum_{i=1}^{m} \lambda_i a_i \mid \lambda_i \geq 0,\ i = 1, \dots, m \right\}.$$



Every polyhedral cone of m vectors, m finite, is closed, even in an infinite-dim Hilbert space.

$b \in C$. When is b in C? Can we write a simple rule?

Thm 12.2 / 12.11 (Farkas-Minkowski in Hilbert spaces)

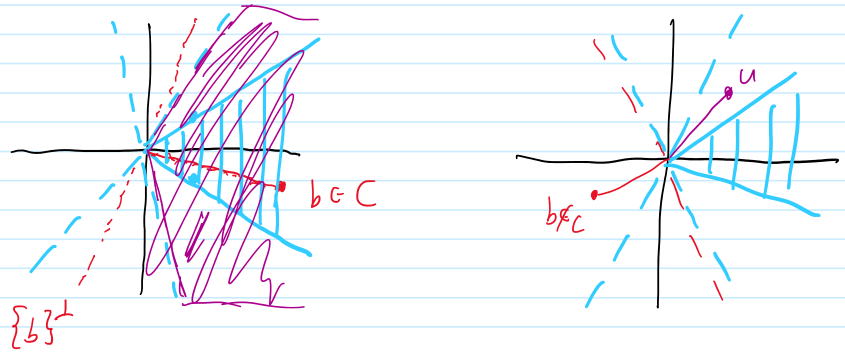
Let V be a real Hilbert space. For any finite seq $(a_1, \dots, a_m)$, $a_i \in V$, let $C = \operatorname{cone}(a_1, \dots, a_m)$. Then $\forall b \in V$, we have $b \notin C$ iff $\exists u \in V$ s.t.

$$\langle a_i, u \rangle \geq 0, \quad i = 1, \dots, m \quad \text{and} \quad \langle b, u \rangle < 0.$$

Equivalently, $b \in C$ iff $\forall u \in V$,

$$\text{if } \langle a_i, u \rangle \geq 0, \quad i = 1, \dots, m, \quad \text{then } \langle b, u \rangle \geq 0.$$

Illustration:



proof. Lemma: If C is a nonempty, closed, convex subset of a Hilbert space V, and $b \in V$ is any vector s.t. $b \notin C$, then $\exists u \in V$ and infinitely many scalars $\alpha \in \mathbb{R}$ s.t.

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

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

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

proof. Use projection lemma.

Since $b \notin C$, $\exists$ unique $c = p_C(b) \in C$ s.t.

$$\begin{cases} \|b-c\| = \inf\limits_{v \in C} \|b-v\| > 0 \\ \langle b-c, v-c\rangle \leq 0 \quad \text{for all } v \in C \end{cases}$$

$$\Rightarrow \langle v-c, c-b\rangle \geq 0 \quad \forall v \in C.$$

Note $\langle c, c-b\rangle - \langle b, c-b\rangle = \|b-c\| > 0$, since $\overset{\in C}{c} \neq \overset{\notin C}{b}$.

$$\Rightarrow \langle v, c-b\rangle \geq \langle c, c-b\rangle > \langle b, c-b\rangle.$$

Choose $u = c-b$, and any α s.t. $\langle c, c-b\rangle > \alpha > \langle b, c-b\rangle$. $\square$

Now back to proving Farkas Minkowski.

Assume $b \notin C$. Then $\exists u \in V$ and some $\alpha \in \mathbb{R}$ s.t.

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

But $0 \in C$, so $\alpha < 0$. (since $\langle 0, u\rangle = 0 > \alpha$)

Then $\forall v \in C$, $\lambda v \in C$ $\forall \lambda > 0$, so

$$\langle \lambda v, u\rangle > \alpha$$
$$\Rightarrow \langle v, u\rangle > \frac{\alpha}{\lambda} \quad \text{for every } \lambda > 0.$$

$$\Rightarrow \langle v, u\rangle \geq 0.$$

Thus, $\langle a_i, u\rangle \geq 0 \quad i = 1, \dots, m$ and $\langle b, u\rangle < \alpha < 0$.

This proves Farkas-Minkowski in the forward direction.

Backward direction is easy ?

Assume $\exists u \in V$ and $\langle u, a_i\rangle \geq 0$ and $\langle b, u\rangle < 0$.

Suppose $b \in C$. Then $b = \sum_i \lambda_i a_i$, $\lambda_i \geq 0$.

$$\Rightarrow \sum_i \langle \lambda_i a_i, u\rangle < 0$$

$$\Rightarrow \sum_i \lambda_i \langle u, a_i\rangle < 0.$$

But this is a contradiction, as $\lambda_i \geq 0$ and $\langle u, a_i\rangle \geq 0$.

But this is a contradiction, as $\lambda_i \geq 0$ and $\langle u, a_i \rangle \geq 0$.
Thus, $b \notin C$. $\square$

General results of optimization theory

We're now going to cover some of the techniques from the last couple of weeks in greater generality.

Def. 13.1 A real-valued function $J : V \to \mathbb{R}$ defined on a normed vector space V is *coercive* iff $\forall$ sequence $(v_k)_{k \geq 1}$, $v_k \in V$, if $\lim_{k \to \infty} \|v_k\| = \infty$, then $\lim_{k \to \infty} J(v_k) = +\infty$.

Ex. $f(x) = x^2 + 2x$ is coercive, but $f(x) = ax + b$ is not (consider $-1, -2, -3, \ldots$)

Prop. 13.1 Let $U \subseteq \mathbb{R}^n$ be nonempty and closed, $J : \mathbb{R}^n \to \mathbb{R}$ continuous and coercive (if U unbounded). Then $\exists u \in U$ s.t. $J(u) = \inf_{v \in U} J(v)$.

proof. Pick $u_0 \in U$. Then $\exists r > 0$ s.t. for all $\|v\| > r$, $J(u_0) < J(v)$.
Let $U_0 = U \cap \{ v \in \mathbb{R}^n \mid \|v\| \leq r \}$, which is compact.
Then $\exists u \in U_0$ s.t. $J(u) = \inf_{v \in U_0} J(v)$, since a continuous function on a compact set has a min.
But then $J(u) \leq J(u_0) < J(v)$, $\Rightarrow J(u) = \inf_{v \in U_0} J(v)$. $\square$

Proof was easy because U_0 is compact. We can generalize to Hilbert spaces.

Def. 13.2 Let V be a Hilbert space. A sequence $(u_k)_{k \geq 1}$ of vectors $u_k \in V$ *converges weakly* if $\exists u \in V$ s.t. $\lim_{k \to \infty} \langle v, u_k \rangle = \langle v, u \rangle$ for every $v \in V$.

Contrast weak convergence to normal convergence in the norm.

Def. A.4 A Hilbert space is *separable* iff it has a countable Hilbert basis.

Def. 13.3 Let V be a Hilbert space. If $f : V \to \mathbb{R}$ is differentiable at $u \in V$, then the derivative $df_u : V \to \mathbb{R}$ is a continuous linear form, so by the Riesz rep. theorem, $\exists$ unique vector $\nabla f_u \in V$ s.t. $df_u(v) = \langle u, \nabla f_u \rangle$ $\forall v \in V$ called the *gradient* of f at u.

Similarly, we can define a generalization of the Hessian $\nabla^2 f_u : V \to V$ by
$D^2 f(\ldots) = \langle \nabla^2 f(\ldots) \ldots \rangle \ \forall \ldots \in V$

Similarly, we can define a generalization of the Hessian $\nabla^2 f_u : V \to V$ by
$$D^2 f_u(v, w) = \langle \nabla^2 f_u(v), w \rangle \quad \forall v, w \in V.$$

Thm 13.1/2 Let U be a nonempty, convex, closed subset of a separable Hilbert space V, and let $J : V \to \mathbb{R}$ be a convex, differentiable function which is coercive if U is unbounded. Then $\exists\, u \in V$ s.t.
$$u \in U \quad \text{and} \quad J(u) = \inf_{v \in U} J(v).$$

proof sketch Very involved proof, but makes use of weak convergence, Riesz representation theorem, etc.

Many of our earlier results can be similarly generalized to Hilbert spaces.

Def. 13.4 Let V be a real Hilbert space. A function $J : V \to \mathbb{R}$ is called a *quadratic functional* if it is of the form
$$J(v) = \frac{1}{2} a(v, v) - h(v),$$
where $a : V \times V \to \mathbb{R}$ is a symmetric and continuous bilinear form, $h : V \to \mathbb{R}$ is a continuous linear form.

In $\mathbb{R}^n$, $a(u, v) = \langle Au, v \rangle$, $h(v) = \langle b, v \rangle$.
We get Hilbert space analogues of our quadratic optimization results.

Def. 13.5 A bilinear form $a : V \times V \to \mathbb{R}$ s.t. $\exists\, \alpha > 0$ s.t. $a(v, v) \geq \alpha \|v\|^2 \quad \forall v \in V$ is said to be *coercive*.

Thm 13.4 (Lions - Stampacchia). Given a Hilbert space V, let
$$J(v) = \frac{1}{2} a(v, v) - h(v)$$
be a quadratic functional.
If a is coercive, then $\forall\, U \subseteq V$ nonempty, closed, and convex, then $\exists$ unique $u \in U$ s.t.
$$a(u, v - u) \geq h(v - u) \quad \forall v \in U.$$
If a is symmetric, then $u \in U$ is the unique element of U s.t.
$$J(u) = \inf_{v \in U} J(v).$$

proof. By Prop. 12.8/9, $\exists$ unique continuous linear map $A : V \to V$ s.t.
$$a(u, v) = \langle Au, v \rangle \quad \forall u, v \in V, \text{ with } \|A\| = \|a\| = C.$$
By Riesz representation theorem, $\exists$ unique $b \in V$ s.t.
$$h(v) = \langle b, v \rangle \quad \text{for all } v \in V$$

By Riesz representation theorem, $\exists$ unique $b \in V$ s.t.

$$h(v) = \langle b, v \rangle \quad \text{for all } v \in V.$$

Thus $\quad J(v) = \frac{1}{2}\langle Av, v \rangle - \langle b, v \rangle \quad \forall\, v \in V.$

Note $\|Av\| \le \|A\|\,\|v\| = C\|v\|.$

Let $\rho > 0$ be a constant to be determined later.

Now $\forall v \in U$, $\quad a(u, v-u) \ge h(v-u)$

$\iff \quad \langle Au, v-u \rangle \ge \langle b, v-u \rangle$

$\iff \quad \rho\langle Au, v-u \rangle \ge \rho\langle b, v-u \rangle$

$\iff \quad \langle \underbrace{\rho b - \rho Au + u} - u, v-u \rangle \le 0.$

By the projection lemma, this is equivalent to finding $u \in U$ s.t.

$$u = p_U(\rho b - \rho Au + u)$$

Consider the function $F : V \to V$ s.t. $F(v) = p_U(\rho b - \rho Av + v)$.

Then u is a fixed pt of F.

Recall that projections do not increase distance, so

$$\|F(v_1) - F(v_2)\| \le \|v_1 - v_2 - \rho(Av_1 - Av_2)\|$$

Then $\quad \|F(v_1) - F(v_2)\|^2 \le \|v_1 - v_2\|^2 - 2\rho\langle Av_1 - Av_2, v_1 - v_2\rangle + \rho^2\|Av_1 - Av_2\|^2$

coercive $\qquad\qquad \|A\| = C$

$$= \|v_1 - v_2\|^2\left[1 - 2\rho\alpha + \rho^2 C^2\right]$$

Pick $\rho > 0$ s.t. $\rho < \frac{2\alpha}{C^2}$. Then $k^2 = 1 - 2\rho\alpha + \rho^2C^2 < 1$, $k > 0$, so

$$\|F(v_1) - F(v_2)\| \le k\|v_1 - v_2\|, \; k < 1, \text{ so } F \text{ is a contraction.}$$

By the Banach fixed pt theorem, F has a unique fixed pt $u \in U$.

If a is symmetric, then a is an inner product on V, so

$$J(v) = \tfrac{1}{2}a(v,v) - a(c,v) \quad \text{for some } c \text{ by the Riesz rep. thm.}$$

$$= \tfrac{1}{2}a(v-c, v-c) - \tfrac{1}{2}a(c,c)$$

So minimizing $J(v)$ over U is equivalent to minimizing $a(v-c, v-c)$,

which by the proj. lemma is equiv. to finding the proj. $p_U(c)$ w.r.t. inner prod. a.

Since U is closed and convex, let $u' = p_U(c)$ w.r.t. a which is unique.

which by the proj. lemma is equiv. to finding the proj. $p_U(c)$ w.r.t. inner prod. a.
Since U is closed and convex, let $u' = p_U(c)$ w.r.t. a, which is unique.

But $a(u', v-u') - h(v-u')$

$= a(u', v-u') - a(c, v-u')$

$= a(u'-c, v-u') = -a(c-u', v-u') \geq 0$ by the proj. lemma,

$\Rightarrow u = u'$.

Thm 13.5 (Lax-Milgram) If $U = V$, i.e. we are optimizing over all of V.

Then $\exists$ unique $u \in V$ s.t.

$$a(u, v) = h(v) \quad \forall v \in V.$$

If a is symmetric, then $u \in V$ is the unique element of V s.t.

$$J(u) = \inf_{v \in V} J(v).$$