1-1

Convex + Nonsmooth Optimization
Spring 2022.

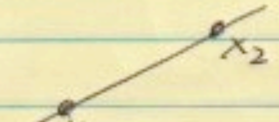
Turn on recording + show web page.

Lec 1 BV 2.1, 2.2, 2.3.1, 2.3.2, 2.5

[Read BV Ch 1 on your own]

"The great watershed in optimization isn't between linearity + nonlinearity, but between convexity + nonconvexity" — RTR, 1993.

BV 2.1 A set $C \subseteq \mathbb{R}^n$ is affine if $\forall x_1, x_2 \in C$, $+ \forall \theta \in \mathbb{R}$,
$$\theta x_1 + (1-\theta) x_2 \in C$$



For any $x_0 \in C$, $V = C - x_0 = \{x - x_0 : x \in C\}$ is a subspace (affine set containing 0)

Write $\dim C = \dim V$.

The affine hull of any set $S \subseteq \mathbb{R}^n$ is $\mathrm{aff}\, S = \{\sum_k \theta_k x_k : \{x_k\} \in S, \sum \theta_k = 1\}$

$\sum \theta_k = 1$ is needed otherwise get all linear combinations

The affine dimension of a set S is $\dim(\mathrm{aff}\, S)$

e.g. $S = \{(x_1, x_2) : x_1^2 + x_2^2 = 1\}$ a circle

$\mathrm{aff}\, S = \mathbb{R}^2$, so aff dim $S = 2$. (not 1)

$\mathrm{relint}\, S = \{x \in S : (\underbrace{B(x,r)}_{\text{ball}} \cap \mathrm{aff}\, S) \subseteq S$ for some $r > 0\}$

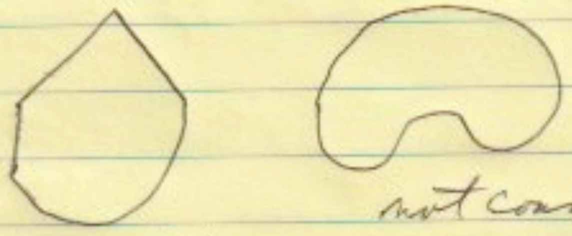
so relint of circle $= \emptyset$

e.g. $S = \{x \in \mathbb{R}^2 : x_1 \in [0,1], x_2 = 0\}$, $\mathrm{int}\, S = \emptyset$, dim relint $S = 1$.

relint $S = \{x \in \mathbb{R}^2, x_1 \in (0,1), x_2 = 0\}$

A set C is <u>convex</u> if $\forall x_1, x_2 \in C$, $\forall \theta \in [0,1]$ (not $\mathbb{R}$ now)

$$\theta x_1 + (1-\theta) x_2 \in C$$



convex

not convex

The convex hull of C is

$$\operatorname{conv} C = \left\{ \sum_k \theta_k x_k : \{x_k\} \in C, \underline{\underline{\{\theta_k\} \geq 0}}, \sum \theta_k = 1 \right\}$$

A set C is a <u>cone</u> if $\forall x \in C$ and $\theta \geq 0$,

$$\theta x \in C.$$

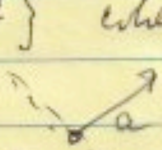
nonconvex cone nonconvex cone. convex cone.

[conic hull of C is

$$\left\{ \sum \theta_k x_k : \{x_k\} \in C, \{\theta_k\} \geq 0. \right\}]$$

A hyperplane is a set

$$\{x : a^T x = \beta\} \text{ where } a \in \mathbb{R}^n, a \neq 0, \beta \in \mathbb{R}$$



A closed halfspace is

$$\{x : a^T x \leq \beta\} \text{ where } a \in \mathbb{R}^n, a \neq 0, \beta \in \mathbb{R}$$

1-3 Some important convex cones

BV2.2 $\mathbb{R}^n_+ = \{x \in \mathbb{R}^n : \underbrace{x \geq 0}\}$ = Nonnegative Orthant

means $x_i \geq 0$, $i = 1, \ldots, n$

All "self-dual" — we'll return to this.

BV notation: $x \succeq 0$ (\succeq)

$$Q^{n+1}_+ = \left\{ \begin{bmatrix} x \\ t \end{bmatrix} \in \mathbb{R}^{n+1} : \|x\|_2 \leq t \right\}$$

$$= \left\{ \begin{bmatrix} x \\ t \end{bmatrix} : \begin{bmatrix} x^T & t \end{bmatrix} \begin{bmatrix} I & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} x \\ t \end{bmatrix} \leq 0,\ t \geq 0 \right\}$$

Quadratic cone (2nd order cone, Lorentz cone, light cone, ice cream cone)

$$S^n_+ = \{ X \in S^n : \underbrace{X \succeq 0} \}$$

real symmetric $n \times n$ matrices $X = X^T$

X is positive semidefinite (PSD)

$v^T X v \geq 0 \ \forall v$

equivalently, eigenvalues of $X \geq 0$

$\mathbb{R}^n_{++} = \text{int}\, \mathbb{R}^n_+$

$Q^{n+1}_{++} = \text{int}\, Q^n_+$

$S^n_{++} = \text{int}\, S^n_+ = \{X : v^T X v > 0 \ \forall v \neq 0\}$ // $\{X : X \succ 0\}$

equivalently, eigenvalues > 0

Note: If $X, Y \in S^n_+$ then $\forall v \in \mathbb{R}^n$

$v^T X v \geq 0$ & $v^T Y v \geq 0$, so $\theta v^T X v + (1-\theta) v^T Y v = v^T(\theta X + (1-\theta) Y) v \geq 0 \quad \forall \theta \geq 0$

so S^n_+ is convex.

1-4

Operations that Preserve Convexity

BV 2.3.1 — If S_1, S_2 are convex, then $S_1 \cap S_2$ is convex (immed from defn) (HW)

BV 2.3.2 — Let $f: \mathbb{R}^n \to \mathbb{R}^m$ be *affine*, i.e. $f(x) = \underset{m\times n}{A}x + \underset{\in \mathbb{R}^m}{b}$

then the *image* of S under f

$$f(S) = \{f(x) : x \in S\}$$

is convex if S is convex, (HW)

and the *inverse* image of S under f,

$$f^{-1}(S) = \{x : f(x) \in S\}$$

is also convex.

e.g. *projection* of S onto some of its coordinates:

$$\left\{x_1 \in \mathbb{R}^m : \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in S \text{ for some } x_2 \in \mathbb{R}^n\right\}$$

(image of f : $f(x) = \begin{bmatrix} I & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$)

— The *Cartesian product* of two convex sets,

$$S_1 \times S_2 = \left\{\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} : x_1 \in S_1, x_2 \in S\right\}$$

is convex, + its image under f with $f(x_1, x_2) = x_1 + x_2$ is convex:

$$S_1 + S_2 = \{x_1 + x_2 : x_1 \in S_1, x_2 \in S_2\}.$$

the *sum* of S_1, S_2

1-5

- the polyhedron $\{x : Ax \le b, Cx = d\}$ is the inverse image of $\mathbb{R}_+^m \times \{0\}$ under affine f, $f(x) = \begin{bmatrix} b - Ax \\ d - Cx \end{bmatrix}$:

$$\{x : f(x) \in \mathbb{R}_+^m \times \{0\}\}$$

- solution set of Linear Matrix Inequality (LMI):

$$\{x : x_1 A_1 + \dots + x_n A_n \preceq B\}$$

$$(\text{equiv}, B - \sum x_k A_k \succeq 0)$$

is the inverse image of S_+^n under affine f. (so it is convex)

- hyperbolic cone

$$\{x : x^T P x \le (c^T x)^2,\ c^T x \ge 0\}$$

where $P \in S_+^n$, $c \in \mathbb{R}^n$ is convex since it is the inverse image of Q_+^{n+1} under affine f,

$$f(x) = \begin{bmatrix} P^{1/2} x \\ c^T x \end{bmatrix}$$

Here $P^{1/2}$ is symmetric, $(P^{1/2})(P^{1/2}) = P$

Can be defined by eigenvalues:

$$P = Q \Lambda Q^T \qquad P^{1/2} = Q \Lambda^{1/2} Q^T$$

$$Q^T Q = I \qquad \begin{bmatrix} \sqrt{\lambda_1} & & \\ & \ddots & \\ & & \sqrt{\lambda_n} \end{bmatrix}$$

Λ diagonal

BV 2.5.1 <u>Separating Hyperplane Thm.</u>

Suppose C, D are <u>nonempty</u>, <u>disjoint</u>, <u>convex</u> sets $\in \mathbb{R}^n$

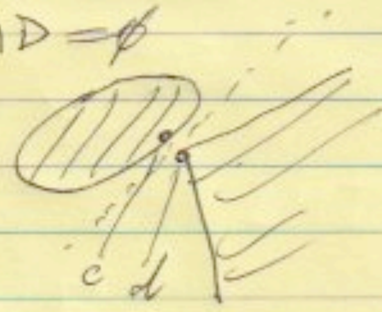
$\leftarrow C \cap D = \emptyset$

Then $\exists\, a \in \mathbb{R}^n,\ \beta \in \mathbb{R}$ s.t.

$a^T x \leq \beta \ \forall\, x \in C$

& $a^T x \geq \beta \ \forall\, x \in D$

(not nec strict)



<u>Pf</u> assuming

$$\text{dist}(C, D) = \inf\{\|u - v\|_2 : u \in C,\ v \in D\}$$

is <u>POSITIVE</u>,

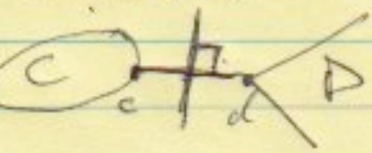
and that $\exists\, c, d$ with

$$\|c - d\|_2 = \text{dist}(C, D).$$

Define

$$a = d - c$$

$$\beta = \frac{\|d\|_2^2 - \|c\|_2^2}{2}$$



and affine fcn

$$f(x) = a^T x - \beta = (d - c)^T\left(x - \tfrac{1}{2}(d + c)\right)$$

Now show f is nonnegative on D.

Suppose not, then $\exists\, u \in D$ with

$$f(u) = (d - c)^T\left(u - \tfrac{1}{2}(d + c)\right) < 0 \quad (*)$$

$$= (d - c)^T\left(u - d + \tfrac{1}{2}(d - c)\right)$$

$$= (d - c)^T(u - d) + \tfrac{1}{2}\|d - c\|_2^2$$

Then $(*)$ implies $(d - c)^T(u - d) < 0$.

cont'd.

1-7

Let $g(t) = \|(d-c) + t(u-d)\|_2^2$

$$= [(d-c) + t(u-d)]^T [(d-c) + t(u-d)]$$

$$= \|d-c\|_2^2 + 2t(u-d)^T(d-c) + O(t^2)$$

so $\left.\frac{d}{dt} g(t)\right|_{t=0} = 2(u-d)^T(d-c) < 0$

so for suff. small $t > 0$ $(t < 1)$ we have

$$g(t) < \|d-c\|_2^2$$

$$g(t)^{1/2} < \|d-c\|_2$$

so $d + t(u-d)$ is closer to c than d is.

Contradiction.

A similar argument shows that f is nonpositive on C. Done.

[In fact BV say that, under the assumption given, can show that f is positive on D and negative on C – strict separation, but they don't give the proof.]

In general, strict separation does not hold, even when both sets are closed. See HW.

However we do have ...

<u>Strict</u> separation of a point + a closed convex set.
$x_0 \notin C$ $\qquad$ C

Since $x_0 \notin C$, and C is closed, $\exists \varepsilon > 0$ s.t.
$B(x_0, \varepsilon) \cap C = \phi$.

So by the SHT, $\exists a \neq 0, \beta$ s.t.
$a^T x \leq \beta$ for $x \in C$ & $a^T x \geq \beta$ for $x \in B(x_0, \varepsilon)$

Since $B(x_0, \varepsilon) = \{x_0 + u : \|u\|_2 \leq \varepsilon\}$,
we have $\underbrace{a^T(x_0+u)}_{a^T x_0 + a^T u} \geq \beta \quad \forall \|u\|_2 \leq \varepsilon$

min of this over u: $u = \frac{-a}{\|a\|_2} \cdot \varepsilon$

Then $a^T x_0 + a^T u = a^T x_0 - \varepsilon \|a\|_2 \geq \beta$

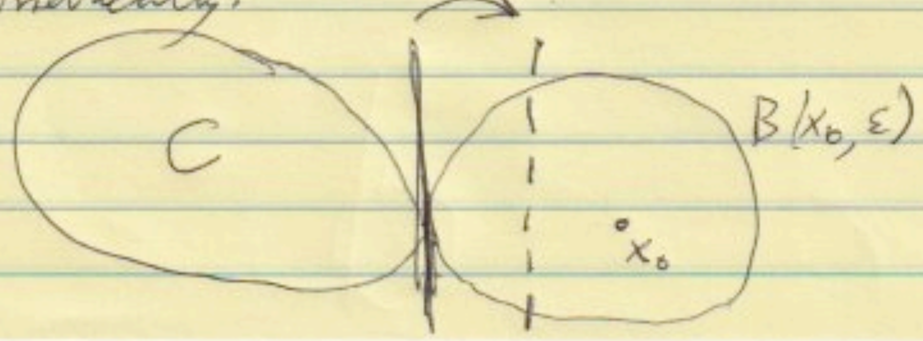
So define new hyperplane with same a
but now the constant $\frac{\beta + \varepsilon\|a\|_2}{2}$ $\hat{\beta}$

$a^T x \leq \beta < \hat{\beta}$ for $x \in C$

and $a^T x \geq \beta + \varepsilon \|a\|_2 > \hat{\beta}$ for $x = x_0$.

so we have STRICT separation.

Geometrically:



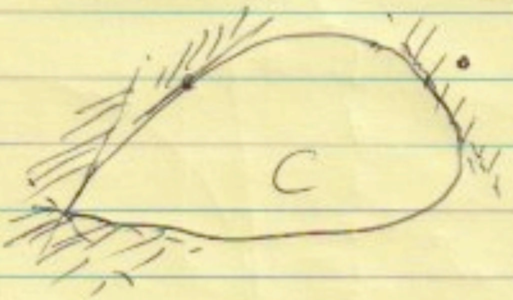
1-9

Consequence: Let C be closed + convex, + let $S = \bigcap \{\text{all } \underline{\text{half spaces}} \text{ containing } C\}$

Clearly $x \in C \Rightarrow x \in S$

Suppose $\exists\, x \in S,\ x \notin C$.

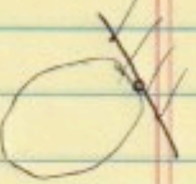
By strict SHT, $\exists$ hyperplane strictly separating C and $\{x\}$, i.e. a halfspace containing C but not x.



But then $x \notin S$ by defn of $\underline{S}$. (since S is defined by intersection of all halfspaces containing C)

So, a closed convex set equals the intersection of all halfspaces that contain it.

1-10



Supporting hyperplanes. not nec convex

Suppose $C \subseteq \mathbb{R}^n$, and
$x_0 \in \operatorname{bd} C \quad (= \operatorname{cl} C \setminus \operatorname{int} C)$

If $a \neq 0$ satisfies $a^T x \leq a^T x_0 \ \forall x \in C$,
then H $\{x : a^T x = a^T x_0\}$ is called a **supporting hyperplane** to C at x_0.

Equivalently, $\{x_0\}$ and C are separated by H.

Thm. For any convex set C, & any $x_0 \in \operatorname{bd} C$, $\exists$ supporting hyperplane to C at x_0

Pf (i) If $\operatorname{int} C \neq \emptyset$, apply SHT to $\{x_0\}$ and $\operatorname{int} C$. (since they **don't** intersect).

(not relint)

(ii) If $\operatorname{int} C = \emptyset$, then C must lie in an affine set with dim $< n$, so any hyperplane containing it contains C and x_0 & is therefore a trivial supp hyperplane, with $a^T x = \beta = a^T x_0$ for $x \in C$.