2-1

TURN ON RECORDING
GUNG XI FA TSAI!

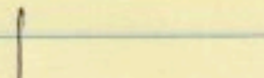
Dec 2.

BV 2.6.1 Dual Cones

let K be a cone. The set
$$K^* = \{y : x^T y \geq 0 \ \forall x \in K\}$$
is called the dual cone of K. We have
$$y \in K^* \iff -y \text{ is the normal of the supporting hyperplane for } K \text{ at } 0.$$

e.g. 1 $K = \mathbb{R}^n_+ = \{x : x \in \mathbb{R}^n, \ x \geq 0\}$

$K^* = (\mathbb{R}^n_+)^* = \mathbb{R}^n_+$, self-dual



doesn't work.

Obvious geometrically.

Pf: Let $x \in K$, so $x \geq 0$ ($x_i \geq 0$, $i = 1, \dots, n$)

Then $x^T z = \sum x_i z_i \geq 0 \ \forall z \in K$, so $x \in K^*$.

If $x \notin K$, then $x_i < 0$ for some $i = 1, \dots, n$

Let $z = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \leftarrow i$ (denote this e_i).

Then $x^T z < 0$, but $z \in K$, so $x \notin K^*$.

e.g. 2 Now, let $K = S^n_+ = \{X \in S^n : X \succeq 0 \text{ (PSD)}\}$

To define K^*, we need an inner product on S^n

We use
$$\langle X, Y \rangle = \operatorname{tr} XY = \sum_{i=1}^{n} (XY)_{ii}$$
$$= \sum_{i=1}^{n} \left(\sum_{k=1}^{n} x_{ik} y_{ik} \right)$$
$$= \sum_{i,j=1,\dots,n} x_{ik} y_{ik}$$

since $Y = Y^T$.

i^{th} row of X i^{th} col of Y

2-2 So def of K^* is $\{Y \in S^n : \operatorname{tr} XY \geq 0 \ \forall X \in K\}$

We claim $K^* = K$.

<u>Pf</u> Let $X \in K$, so $X \succeq 0$ (PSD), and

so $X = Q \Lambda Q^T$, $Q^T Q = I$, Λ diagonal, $\Lambda \geq 0$

$$= \sum_{i=1}^{n} \underbrace{\lambda_i}_{i^{\text{th}} \text{ diag entry in } \Lambda} \underbrace{q_i}_{i^{\text{th}} \text{ col of } Q} q_i^T$$

Then $\langle X, Z \rangle = \operatorname{tr}\left(\left(\sum_{i=1}^{n} \lambda_i q_i q_i^T\right) Z\right)$

$$= \operatorname{tr} \sum_{i=1}^{n} \lambda_i \underbrace{q_i^T Z q_i}_{\geq 0 \text{ for all } Z \succeq 0} \quad (\text{as } \operatorname{tr} AB = \operatorname{tr} BA)$$

so $X \in K^*$.

If $X \notin K$ then $\exists v \in \mathbb{R}^n$ with $v^T X v < 0$

(e.g. e-vector for neg. e-value)

Let $Z = v v^T$.

Then $\operatorname{tr} XZ = \operatorname{tr} X v v^T = v^T X v < 0$

although $Z \in S_+^n = K$.

So $X \notin K^*$

So $K = K^*$ for $K = S_+^n$. self-dual.

e.g. 3 Let $\|\cdot\|$ be a norm on $\mathbb{R}^n$ and

$$K = \left\{ \begin{bmatrix} x \\ t \end{bmatrix} : \|x\| \leq t \right\}$$

Define the <u>dual norm</u> by

$$\|u\|_* = \sup \{u^T x : \|x\| \leq 1\}.$$

e.g. If $\|x\| = \|x\|_p = \left(\sum |x_i|^p\right)^{1/p}$ e.g. $p=2, q=2$

then $\|x\|_* = \|x\|_q$ where $\frac{1}{p} + \frac{1}{q} = 1$. $p=1, q=\infty$

Pf: use Hölder's inequality $p \in [1, \infty]$ $[p=\infty, q=1]$.

2-3.

The dual norm $K^* = \left\{ \begin{bmatrix} x \\ t \end{bmatrix} : \|x\|_* \le t \right\}$. Pf: p.53

In particular, $K = Q^n_*$ is self-dual $(p=2)$

Actually the only self-dual cones are Q^n_+, S^n_+, the extension of S^n_+ to complex Hermitian matrices, the quaternions + octonians, + their direct sums.

What about R^n_+? It is $S^1_+ \oplus \cdots \oplus S^1_+$!

BV 3.1 Convex Functions

TWO DIFFERENT NOTATIONS

BV: $f : \mathbb{R}^n \to \mathbb{R}$ with f defined only on its domain $\operatorname{dom} f$. e.g. $\operatorname{dom} \log = \mathbb{R}_{++}$

OR $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ defined everywhere

e.g. $\log x = \begin{cases} \log x & \text{if } x > 0 \\ +\infty & \text{if } x \le 0. \end{cases}$

We now say $\operatorname{dom} f = \{x : f(x) < \infty\}$

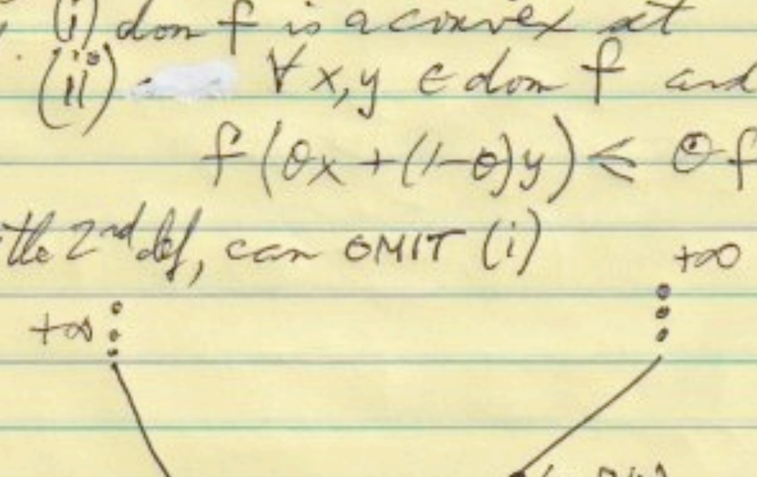
We say f is convex using the BV def'n if

(i) $\operatorname{dom} f$ is a convex set

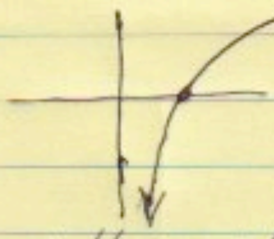
and (ii) $\forall x, y \in \operatorname{dom} f$ and $\theta \in [0,1]$

$$f(\theta x + (1-\theta) y) \le \theta f(x) + (1-\theta) f(y)$$

Using the 2nd def, can OMIT (i)



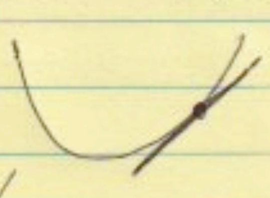
e.g. $-\log$ is convex (using either def)



Thm

Suppose f is differentiable, so dom f is open and $\nabla f \doteq \begin{bmatrix} \partial f/\partial x_1 \\ \vdots \\ \partial f/\partial x_n \end{bmatrix}$ is defined everywhere on dom f

Then f is convex iff (i) dom f is convex and
(ii) $f(y) \geq f(x) + \nabla f(x)^T (y-x) \quad \forall x, y \in \text{dom } f$



function lies above the tangent line

Pf: BV

Thm. Suppose f is twice continuously differentiable; so dom f is open ← not in BV, seems to be needed so $\nabla^2 f(x)^T = \nabla^2 f(x)$
and its Hessian $\nabla^2 f(x) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \vdots & & \\ \frac{\partial^2 f}{\partial x_1 \partial x_n} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{bmatrix}$ is continuous (& therefore symmetric) on dom f.
Then f is convex iff (i) dom f is convex and (ii) $\nabla^2 f(x) \succeq 0$ (PSD) $\forall x \in \text{dom } f$.

Examples on $\mathbb{R}$

all convex:

- e^{ax} for any $a \in \mathbb{R}$
- x^a on $\mathbb{R}_{++}$ when $a \geq 1$ or $a \leq 0$
- $|x|^p$ on $\mathbb{R}$ for $p \geq 1$
- $-\log x$ on $\mathbb{R}_{++}$
- $x \log x$ (negative entropy) on $\mathbb{R}_{++}$ (or $\mathbb{R}_+$, taking value 0 at 0)

2-5

- $\|\cdot\|$ on $\mathbb{R}^n$ — Pf $\|\theta x+(1-\theta)y\| \le \|\theta x\| + \|(1-\theta)y\|$ triangle ineq.
  $= \theta\|x\|+(1-\theta)\|y\|$.
- $\max(x_1, \dots, x_n)$ on $\mathbb{R}^n$
- log sum exp
  $\log(e^{x_1}+\dots+e^{x_n})$
- geometric mean $\left(\prod_{i=1}^{n} x_i\right)^{1/n}$ on $\mathbb{R}_{++}$
- $-\log\det$: $f(X) = -\log\det X$ on S_{++}^n.

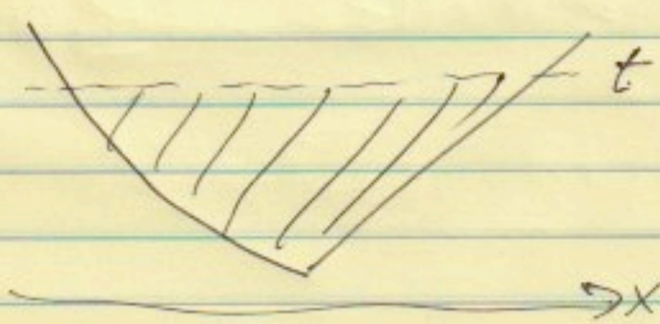
  Nice fact: NOTE.
  $$\nabla f(X) = X^{-1}.$$

Epigraph

$$\text{graph } f = \left\{ \begin{bmatrix} x \\ f(x) \end{bmatrix} : x \in \text{dom } f \right\} \subseteq \mathbb{R}^{n+1}$$

set above the graph. $$\text{epi(graph) } f = \left\{ \begin{bmatrix} x \\ t \end{bmatrix} : f(x) \le t \right\} \subseteq \mathbb{R}^{n+1}$$

f is a convex fun $\Leftrightarrow$ epi f is a convex set



Operations that Preserve Convexity of Functions

— $f = \sum_{i=1}^{m} w_i f_i$ is convex if f_i convex, $w_i \geq 0$.

— $g(x) = \int_A w(y) f(x,y)\,dy$ is convex if $f(x,y)$ is convex $\forall y \in A$ and $w(y) \geq 0$ $\forall y \in A$.

— $g(x) = f(Ax+b)$ is convex if f is convex

BV 3.2.3 — $f(x) = \max(f_1(x), \ldots, f_m(x))$ is convex if $f_1, \ldots, f_m$ are

e.g. sum of r largest components
For $x \in \mathbb{R}^n$, let $x_{[i]}$ mean i^{th} largest of $x_1, \ldots, x_n$
Then $f(x) = \sum_{i=1}^{r} x_{[i]}$ is convex on $\mathbb{R}^n$,
because it can be written as
$$\max\{x_{i_1} + \ldots + x_{i_r} : 1 \leq i_1 \leq i_2 \leq \ldots \leq i_r \leq n\}$$
the pointwise max of $\binom{n}{r}$ linear functions
which are convex.

— $g(x) = \sup_{y \in A} f(x,y)$ is convex if $f(x,y)$ is convex $\forall y \in A$.

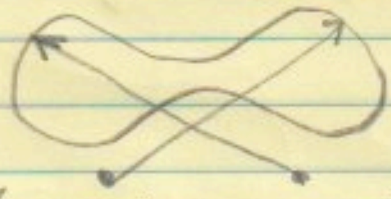
e.g. support function for a set.
Let $C \subseteq \mathbb{R}^n$, $C \neq \emptyset$.
The support function is $S_C(x) = \sup\{\underbrace{x^T y}_{\text{linear, } \therefore \text{ convex}} : y \in C\}$

e.g. distance to farthest point in a set (in any norm)

$$f(x) = \sup_{y \in C} \|x-y\| \quad \text{convex on } \mathbb{R}^n$$

$C \subseteq \mathbb{R}^n$

Don't need to assume C is convex



Not distance to nearest point.

Pf $\|x-y\|$ is convex in x for any y.

e.g. max eigenvalue: $f(X) = \lambda_{max}(X) \quad X \in S^n$

$$= \sup\{v^T X v : \|v\|_2 = 1\}$$

$$= \sup \langle X, vv^T\rangle : \|v\|_2 = 1\}.$$

max singular value: $f(X) = \|X\|_2$

$$= \sup\{u^T X v : \|u\|_2 = \|v\|_2 = 1\}$$

$$= \sup\{\langle X, uv^T\rangle : \quad " \quad \}$$

Alternatively:

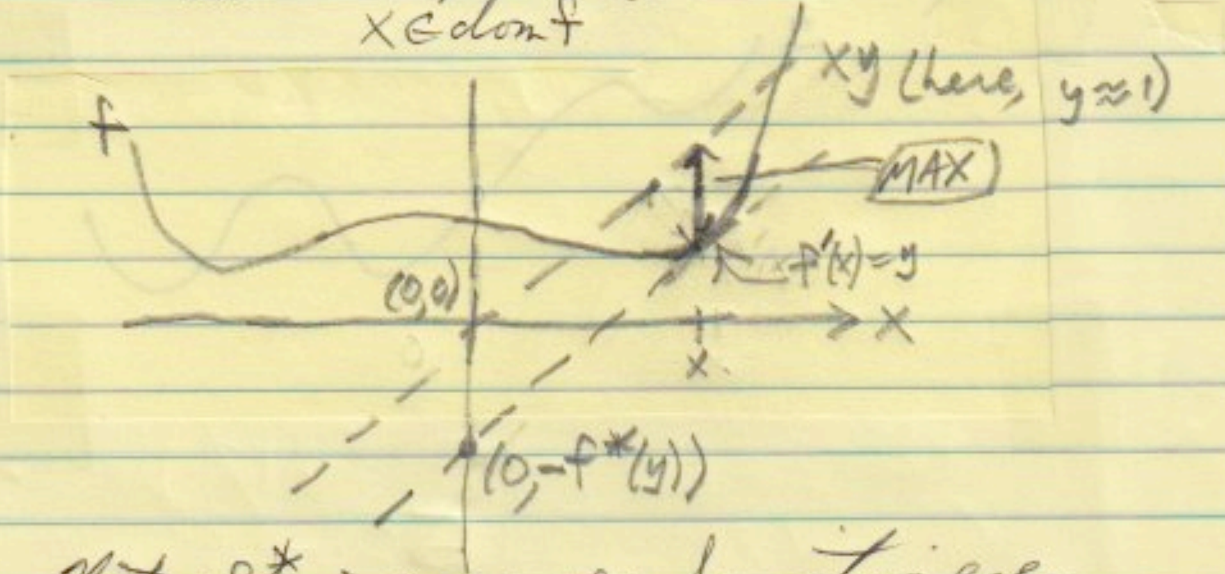
$$f(X) = \|X\|_2 = \max\{\|Xu\| : \|u\| = 1\}$$

max of convex functions

Conjugate Function

Let $f: \mathbb{R}^n \to \mathbb{R}$, not nec. convex. The conjugate of f is

$$f^*(y) = \sup_{x \in \mathrm{dom} f} \left(y^T x - f(x)\right)$$



Note: f^* is a convex function even if f is not, as it is max of affine functions of y.

e.g. $f(x) = ax + b$

$$f^*(y) = \begin{cases} -b & \text{if } y = a \\ +\infty & \text{if } y \neq b \end{cases}$$

e.g. $f(x) = \begin{cases} -\log x & x > 0 \\ \infty & x \leq 0 \end{cases}$

$$f^*(y) = \sup_{x>0} \; yx + \log x = \begin{cases} +\infty & \text{if } y \geq 0 \\ -1 - \log(-y) & \text{if } y < 0 \end{cases}$$

Let $g(x) = yx + \log x$ for $y < 0$

$\rightarrow -\infty$ as $x \to \infty$

$\rightarrow -\infty$ as $x \to 0$

$g'(x) = y + \frac{1}{x} = 0 \Rightarrow x = -\frac{1}{y}$

2-9 $f: S^n \to \mathbb{R}$

e.g. $f(X) = \begin{cases} -\log\det X & \text{if } X > 0 \\ +\infty & \text{otherwise} \end{cases}$

$$f^*(Y) = \sup_{X \succ 0} \{ \operatorname{tr}(YX) + \log\det X \}$$

If $Y \nprec 0$ (not neg. def.) then let λ be an evalue of Y with e-vector v, with $\lambda \geq 0$

Let $X = I + tvv^T$

then $\operatorname{tr} YX + \log\det X = \operatorname{tr} Y + t\lambda + \log\det(\underbrace{I + tvv^T}_{\text{e-values are } 1, \dots, 1, 1+t})$

$= \underbrace{\operatorname{tr} Y}_{\text{const}} + \underbrace{t\lambda}_{\text{either } \to\infty \text{ or is } 0.} + \underbrace{\log(1+t)}_{\to\infty \text{ as } t\to\infty}$

So $f^*(Y) = \begin{cases} \infty & \text{if } Y \nprec 0 \\ -n - \log\det(-Y) & \text{if } Y < 0 \end{cases}$

For 2nd part, take gradient.

$$\nabla \{ \operatorname{tr} YX + \log\det X \} = Y + X^{-1} = 0$$

so max is at $X = -Y^{-1}$ so get

$$-n + \log\det(-Y^{-1}) = -n + \log\left(\frac{1}{\det(-Y)}\right)$$

$$= -n - \log\det(-Y)$$

Demo CVX using lsqsol.m, simpleLP.m