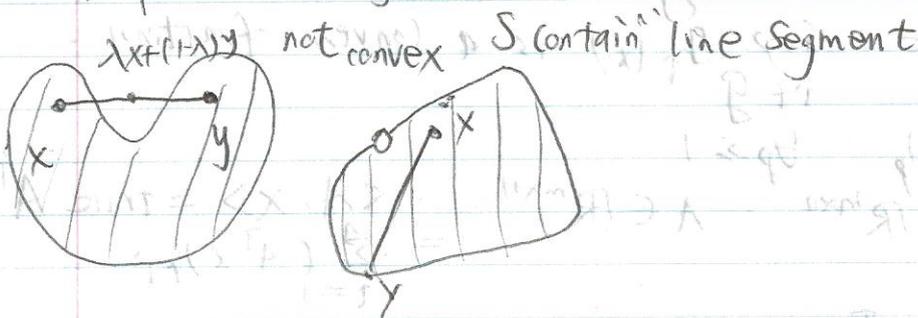
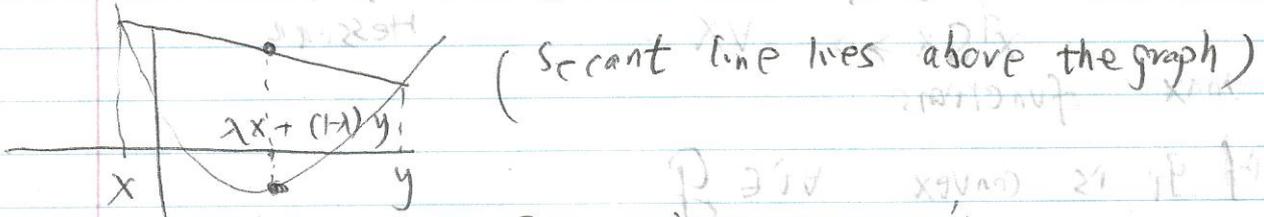


Defn: A function  $f: \mathbb{R}^n \rightarrow (-\infty, +\infty]$  is convex if:

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y) \quad \forall x, y \in \mathbb{R}^n, \forall \lambda \in [0, 1]$$

Defn: A set  $S \subseteq \mathbb{R}^n$  is a convex set if:

$$\lambda x + (1-\lambda)y \in S \quad \forall x, y \in S, \forall \lambda \in [0, 1]$$



Def: A convex optimization problem has the form

min  $f(x)$  (convex function)

subject to  $x \in S$  (convex set)

(s.t.)

usually  $S = \{x \in \mathbb{R}^n : g_i(x) \leq 0, i=1, \dots, m, h_j(x) = 0, j=1, \dots, p\}$

$x \in \Omega$

where  $g_i$  are convex  $\forall i$ ,  $h_j$  are affine  $\forall j$ .

$\Omega$  is a ("simple") convex set.

Examples: (i) linear function:  $f(x) = \langle c, x \rangle = c^T x = \sum_j c_j x_j$

(ii) affine function  $h(x) = c^T x + \alpha, \alpha \in \mathbb{R}$

(iii) exponential  $e^{ax}, a \in \mathbb{R}$

powers  $x^\alpha$  on  $\mathbb{R}_{++}$   $\alpha \geq 1$  or  $\alpha \leq 0$  ( $\mathbb{R}_+ = \{\alpha \geq 0\}$ ,  $\mathbb{R}_{++} = \{\alpha > 0\}$ )

Entropy  $x \log x$  on  $\mathbb{R}^+$ . (negative)

(iv) convex quadratic functions.

$$f(x) = \frac{1}{2} x^T Q x + c^T x + d \quad Q = Q^T \in \mathbb{R}^{n \times n}, c \in \mathbb{R}^n, d \in \mathbb{R}$$

is convex iff.  $Q \succeq 0$  positive semi-definite

(Def) Recall  $Q = Q^T$  is psd if  $\forall v, v^T Q v \geq 0$  Hessian

(v) max functions

if  $g_i$  is convex  $\forall i \in \mathcal{I}$   
then  $f(x) = \sup_{i \in \mathcal{I}} g_i(x)$  is a convex function.

(vi) norms a)  $\|x\|_p, p \geq 1$   
b)  $X \in \mathbb{R}^{m \times n}, A \in \mathbb{R}^{m \times n}$   
 $\langle A, X \rangle = \text{trace } A^T X = \sum_{i=1}^n (A^T X)_{ii}$

$$\text{vec}(A)^T \text{vec}(X)$$

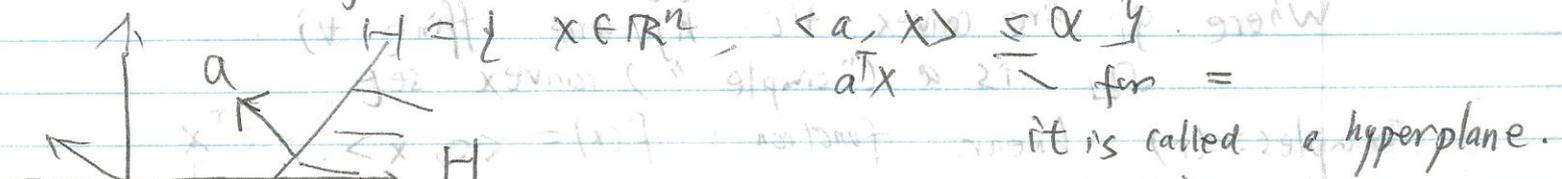
$\text{vec} \rightarrow$  stacks columns of  $A$ .

$$c) \|x\|_2 = \sigma_{\max}(X) = \sqrt{\lambda_{\max}(X^T X)}$$

Singular values  $\lambda_{\max}$  eigenvalues

Examples of convex sets

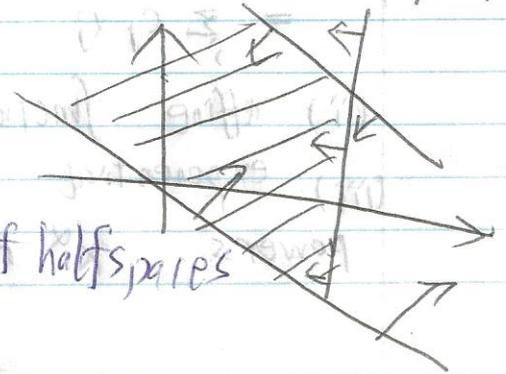
a) half-spaces  
given  $a \in \mathbb{R}^n, \alpha \in \mathbb{R}$



b) polyhedral sets

$$\{x \in \mathbb{R}^n \mid Ax \leq b\}$$

polyhedron is intersection of finite number of halfspaces and hyperplanes.

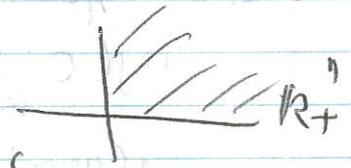


$\{x \mid (x-x_0)^T P (x-x_0) \leq 1\} \subset \mathbb{R}^n$  -  $P$  symmetric positive definite. 3.

$E = \{x \in \mathbb{R}^n : x^T Q x + p^T x + q \leq 0\}$   
 $Q$   $n \times n$  symmetric, psd.

d) Convex cones.  
 $K$  is a cone if  $\lambda K \subset K \quad \forall \lambda \geq 0$ .  
 $k \in K \quad \lambda \geq 0 \Rightarrow \lambda k \in K$  contains all rays from origin.

convex cone.  
 A cone  $K$  is a convex cone if  $K$  is a convex set.  
 (Equivalently, if  $K+K \subset K$ )



Examples of convex optimization problems.

a) Linear Programming, LP (seen already)

$\min C^T x$  linear obj.  
 st.  $Ax = b$   $x \in \mathcal{L}$  convex set.  
 $x \geq 0$   
 affine eq. constr.

(b) SDP  $x \in S^n$  space of  $n \times n$  symmetric matrices  
 $x, y \in S^n, \langle x, y \rangle = \text{trace}(x^T y) = \text{trace } xy$

$= \sum_{i,j} x_{ij} y_{ij}$

$\dim(S^n) = \frac{n(n+1)}{2}$  ( $\leq n^2$  ignore the strictly lower triangular)

linear SDP  $\min \langle C, x \rangle$   
 st.  $\langle A_i, x \rangle = b_i, i=1, \dots, m$   $A_i = A_i^T \quad b_i \in \mathbb{R} \quad \forall i$   
 $x \succeq 0$

(c) Geometric Programming

A monomial is  $m(x) = c \prod_{i=1}^n x_i^{\alpha_i}$  ( $\alpha_i \in \mathbb{R}$ )  
 $c > 0$

A posynomial  $p(x) = \sum_{j=1}^k c_j \prod_{i=1}^n x_i^{\alpha_{ij}}$  ( $\alpha_{ij} \in \mathbb{R}$ )  
 $c_j > 0$  Hilroy

## A geometric programming problem

$$\begin{aligned} \min f(x) \\ \text{st. } x \in S = \{ x \in \mathbb{R}^n \mid & \\ & \left. \begin{array}{l} g_i(x) \leq 1, \quad \forall i=1, \dots, m \\ h_j(x) = 1, \quad \forall j=1, \dots, p \end{array} \right\} \end{aligned}$$

posynomial
monomial

This is not convex in general.

Let  $y_i = \log x_i$  ( $x_i = e^{y_i}$ )

$$\log\left(C \prod_{i=1}^n x_i^{\alpha_i}\right) = \sum_{i=1}^n \alpha_i y_i + \log C$$

So  $\log(p(x))$  is convex in  $y$ .

$$\bar{f}(y) = \log f(e^y) \quad \text{and similarly for constr.}$$

Then we get a convex progr. in  $y$ .

eg. "nearest" Toeplitz PSD matrix



"Optimization Tree" at NEOS Sept 13th, 2007

— webpage → links to notes on "Convex sets and functions"

Recall

Euclidean Spaces

Def<sup>n</sup> A Euclidean space  $E$  is a finite dimensional vector space equipped with an inner product.

$$\langle \cdot, \cdot \rangle : E \times E \rightarrow \mathbb{R}$$

a)  $\langle \alpha x + \beta y, z \rangle = \alpha \langle x, z \rangle + \beta \langle y, z \rangle$

b)  $\langle x, y \rangle = \langle y, x \rangle$  ( $\langle x, y \rangle = \overline{\langle y, x \rangle}$  over complex field)

c)  $\langle x, x \rangle \geq 0$ ,  $\langle x, x \rangle = 0$  iff  $x = 0$

Examples:

(a)  $(\mathbb{R}^n, \cdot, \cdot)$

(b)  $(\mathbb{C}^n, \cdot, \cdot)$  conjugate transpose

(c)  $(\mathbb{R}^{n \times n}, \langle X, Y \rangle \mapsto \text{trace } X^T Y)$

(d) space of polynomials of  $n$  variables of degree at most  $d$

$$\left\langle \sum_{|\alpha| \leq d} C_\alpha X^\alpha, \sum_{|\alpha| \leq d} d_\alpha X^\alpha \right\rangle \mapsto X^\alpha = x_1^{\alpha_1} x_2^{\alpha_2} \dots x_n^{\alpha_n}$$

$\alpha = (\alpha_1, \dots, \alpha_n)$   
non neg. integers

$|\alpha|$  denotes  $\sum_{i=1}^n \alpha_i$

$$\sum_{|\alpha| \leq d} C_\alpha d_\alpha$$

Fact: Each euclidean space of dimension  $n$  is linearly isomorphic to  $(\mathbb{R}^n, \cdot, \cdot)$

Construction of Euclidean spaces

Def<sup>n</sup> A non empty subset  $S$  of  $E$ , is a linear subspace if

$$\lambda X + \mu Y \in S \quad \forall X, Y \in S \quad \forall \text{ scalars } \lambda, \mu$$

if  $L$  is a subspace of  $E$ , then

$$\langle \cdot, \cdot \rangle_L = \langle \cdot, \cdot \rangle \upharpoonright_{L \oplus L} \quad \text{direct sum}$$

ie. the induced inner product from  $E$

so  $L$  is a Eucl. space

Let  $x_i \in E$

and construct  $(x_1, x_2, \dots, x_p) \in (E_1 \oplus E_2 \oplus \dots \oplus E_p)$  - called

$$\langle \cdot, \cdot \rangle = \langle (x_1, \dots, x_p), (y_1, \dots, y_p) \rangle = \sum_{i=1}^p \langle x_i, y_i \rangle \quad \text{direct sum}$$

eg.  $S^n = \{ S \in \mathbb{R}^{n \times n} : S = S^T \}$   
space of symmetric matrices.

$$x, y \in S^n, \langle x, y \rangle = \text{tr } XY$$

$$E = \underbrace{\mathbb{R}^n}_{x, y} \oplus \underbrace{S^n}_{x, y}$$

$$\langle (x, X), (y, Y) \rangle = x^T y + \text{trace } XY$$

Topology of  $E$

## Euclidean norm

Def<sup>n</sup> The norm induced by  $\langle \cdot, \cdot \rangle$  is

$$\|x\| = \sqrt{\langle x, x \rangle}$$

properties (1)

$$\|x+y\|^2 = \|x\|^2 + \|y\|^2 \quad \text{if } \langle x, y \rangle = 0$$

(2) Cauchy-Schwarz

$$|\langle x, y \rangle| \leq \|x\| \|y\| \quad \text{with equality iff}$$

$x = \alpha y$  for some scalar  $\alpha$

ortho. (perpendicular)

Def<sup>n</sup> The unit ball is

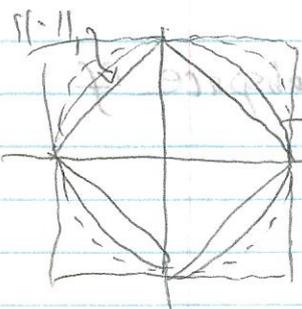
$$B = \{x \in E : \|x\| \leq 1\}$$

recall p-norm on  $\mathbb{R}^n$

$$\|x\|_p = \sqrt[p]{\sum_{i=1}^n |x_i|^p} \quad p \geq 1$$

convex functions

$$\|x\|_\infty = \lim_{p \rightarrow \infty} \|x\|_p = \max |x_i|$$



$\|x\|_1$

$\|x\|_2$

$$\|x\|_p = |x| \|x\|_p$$

Def<sup>n</sup> A set  $S$  is open if  $\forall x \in S \exists \delta > 0, \exists \epsilon > 0$  s.t.  $\{x\} + \delta B \subseteq S$

$$\{x\} + \delta B \subseteq S$$



The collection of open sets form a topology.

Def<sup>n</sup> A set  $T$  is closed if the complement  $T^c$  is open.

Def<sup>n</sup> The interior of a set  $S \subseteq E$  is the largest open set contained in  $S$ , denoted  $\text{int}(S)$

$$\text{Equivalently } \text{int}(S) = \bigcup \{ \text{open sets in } S \}$$

Def<sup>n</sup> So the closure of a set  $S \subseteq E$  is the intersection of all closed sets containing  $S$ .

(Equiv. the smallest closed set containing  $S$ )

$$\text{denoted } \underline{d(S)} \quad \underline{cl(S)}$$

$$X = \lim_{j \rightarrow \infty} x_j \quad \|X - x_j\| \rightarrow 0$$

limit point of the sequence  $\{x_j\}$  (in  $\mathbb{R}^n$ )

Note: closure  $\bar{S}$  is  $S$  union all its limit points

ie.  $\{x_j\} \subset S \quad \lim_{j \rightarrow \infty} x_j \rightarrow X \Rightarrow X \in \text{cl}(S)$

The boundary of a set  $B$ :  $\text{bd}(B) = \text{cl}(B) \setminus \text{int}(B)$

Theorem 1.1.2 in text. Bolzano-Weierstrass

Bounded sequences in  $\mathbb{E}$  have convergent subsequences,

ie.  $\{x_j\}_{j=1}^{\infty} \subset \mathbb{E}$  bounded seq.  
 $\|x_j\| \leq K, \forall j$

$\Rightarrow \exists \{x_{j_k}\} \subset \{x_j\} \exists$  hunting lions in the desert.

$$\lim_{k \rightarrow \infty} x_{j_k} = X$$

Linear manifold.

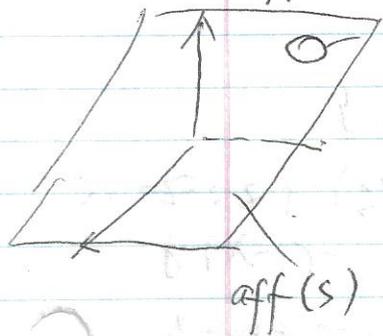
Def<sup>n</sup> A subset  $S \subseteq \mathbb{E}$  is an affine subspace

if  $\lambda x + (1-\lambda)y \in S \quad \forall x, y \in S, \lambda \in \mathbb{R}$

Affine subspace  $S$  is a translation of a (unique) linear subspace

$S = x + L, \forall x \in S$  An affine space is a vector space that's forgotten its origin.

Def<sup>n</sup> The affine hull of a set  $S \subseteq \mathbb{E}$  is the intersection of all affine subspaces containing  $S$ . The affine hull of a set of two different points is the line through them.



desk.  $\dim(S)$  is the  $\dim$  of  $\text{aff}(S)$  is the  $\dim$  of

$$L \text{ in } \text{aff}(S) = x + L$$

Linear subspace.

If  $\bar{x} \in S$ , then  $0 \in S - \bar{x}$ . The affine hull of a set of three points not on one

then  $L = \text{span}(S - \bar{x})$

$$\bar{x} + L = \text{aff}(S)$$

plane line is the plane going through them Hilroy

$$\dim(S) = \dim(S - \bar{x}) = \dim(\text{span}(S - \bar{x})) \text{ for any } \bar{x} \in S$$

Def<sup>n</sup> A set  $S$  is relatively open if  $\forall x \in S, \exists \delta > 0$  such that  $(x + \delta B) \cap \text{aff}(S) \subseteq S$ .

Def<sup>n</sup> The relative interior of  $S$  is the largest relatively open set in  $S$ .

$S = \{x_1, x_2\}$  is not relatively open

has an empty rel. int.

What about convex sets? (Exercise)

$$\text{aff}(S) = \left\{ \sum_{k=1}^K \alpha_k x_k \mid x_k \in S, \alpha_k \in \mathbb{R}, \sum_{k=1}^K \alpha_k = 1, k=1, 2, \dots \right\}$$

$$H_{\text{convex}}(S) = \left\{ \sum_{k=1}^K \alpha_k x_k \mid \alpha_k \geq 0 \right\}$$

Notice the only difference in  $\alpha_k \geq 0$  for convex hull of a set  $S$ .

Linear hull or linear span

$$\text{span}(v_1, \dots, v_r) = \left\{ \lambda_1 v_1 + \dots + \lambda_r v_r \mid \lambda_1, \lambda_2, \dots, \lambda_r \in K \right\}$$

field.

$x$  is a relative interior point of  $C$ , if  $x$  is an interior point of  $C$  relative to  $\text{aff}(C)$

Three fundamental properties of  $\text{relint}(C)$  of a convex set.

- $\text{relint}(C)$  is nonempty
- if  $x \in \text{relint}(C)$  and  $\bar{x} \in C \setminus \text{relint}(C)$ , then all points on the line segment connecting  $x$  and  $\bar{x}$ , except possibly  $\bar{x}$ , belonging to  $\text{relint}(C)$
- if  $x \in \text{relint}(C)$  and  $\bar{x} \in C$ , the line segment connecting  $\bar{x}$  and  $x$  can be prolonged beyond  $x$  without leaving  $C$ .

Convex Sets & functions — topological Sept 18th, 2007

Algebraic  
Recall: if  $S \subseteq E$  (Euclidean spaces  
( $E, \langle \cdot, \cdot \rangle$ )

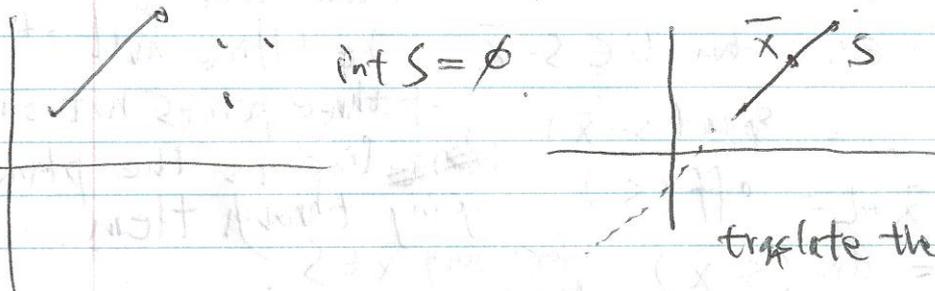
Vector space inner product

then  $\text{int } S = \{x \in S : x + \delta B \subseteq S, \text{ for some } \delta > 0\}$   
where  $B = \{x \mid \|x\| \leq 1\}$

$$\text{aff}(S) = \{z : z = \lambda x + (1-\lambda)y\}$$

for some  $x, y \in S$   
 $\lambda \in \mathbb{R}$

translate the set to origin

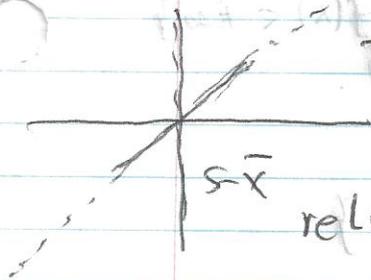


$$0 \in S - \bar{x} (= \{y \mid y = s - \bar{x} \text{ for some } s \in S\})$$

9.

we can talk about the span

Then  $\text{span}(S - \bar{x}) = \mathcal{L}$  subspace



$$\bar{x} + \mathcal{L} = \text{aff}(S)$$

$$\text{relint}(S) = \bar{x} + \text{relint}(S - \bar{x})$$

translate so  $0 \in$  new set  $\mathcal{L}$  i.e. just int when  
 $\text{aff}(S - \bar{x}) = \text{span}(S - \bar{x})$  (here  $S - \bar{x}$ ) inside  $\mathcal{L}$

all possible lin. comb

find a bases  $v_1, \dots, v_k$

$$\text{span}\{v_i\}_{i=1}^k$$



$\text{span}(x_1, x_2) \cong 2$  dimensional

translate to origin, 1 dimensional

Def<sup>n</sup> The graph of a function  $f: E \rightarrow [-\infty, +\infty]$

$$\text{is } \text{gr}(f) = \{(x, r) \in E \times \mathbb{R} : f(x) = r\}$$

Note A function is linear if its  $\text{gr}(f)$  is a linear subspace.  
 and a function is affine if its graph  $\text{gr}(f)$  is an affine subspace.

Recall (i)  $f$  is affine iff  $\exists$  a linear function  $g$  and  $b \in \mathbb{R}$ .

$$\Rightarrow f \equiv g + b$$

such that  $f(x) = g(x) + b \quad \forall x \in E$

(ii) An affine function  $f$  is linear iff  $f(0) = 0$ .

Given a linear mapping

$$A: E_1 \rightarrow E_2$$

The adjoint of  $A$  is denoted  $A^*$  and defined by

$$\langle Ax, y \rangle = \langle x, A^*y \rangle \quad \forall x \in E_1, \forall y \in E_2$$

eg.  $E_1 = \mathbb{R}^n, E_2 = \mathbb{R}^m$

$$A \in \mathbb{R}^{m \times n}$$

$$Ax = AX$$

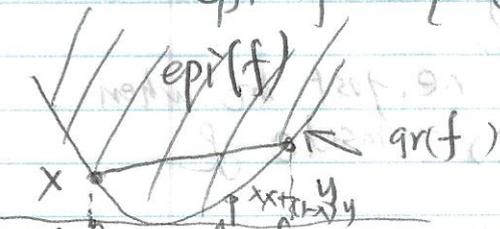
then  $A^* = A^T$

Hilroy

Def<sup>n</sup> The domain of  $f$  is  $\text{dom } f = \{x \in \mathbb{E} : f(x) < +\infty\}$

Def<sup>n</sup> The function is proper if  $\text{dom } f \neq \emptyset$ .

Def<sup>n</sup> The epigraph of a function  $f$  is  
 $\text{epi } f = \{(x, r) \in \mathbb{E} \oplus \mathbb{R} : f(x) \leq r\}$



Th<sup>m</sup> A function  $f$  is convex iff its epigraph is convex.

proof: only if (Nec) Let  $f$  be convex.

$$\forall (x, r), (y, s) \in \text{epi}(f) \quad \forall \lambda \in (0, 1)$$

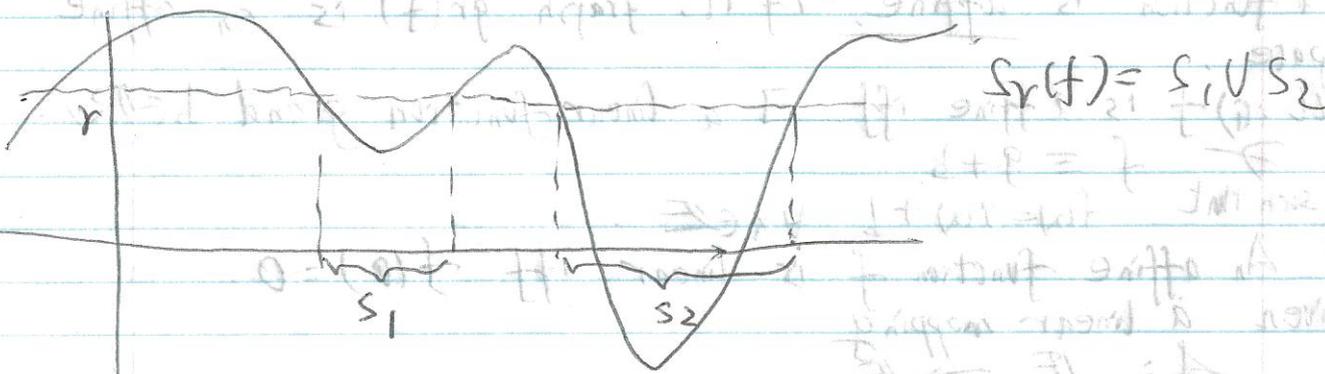
$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y) \quad \text{by convexity of } f$$

$$\leq \lambda r + (1-\lambda)s$$

$$\Rightarrow (\lambda x + (1-\lambda)y, \lambda r + (1-\lambda)s) \in \text{epi } f \quad \square$$

Def<sup>n</sup> A sublevel-set of function  $f$  at level  $r$  is

$$S_r(f) = \{x \in \mathbb{E} : f(x) \leq r\}$$



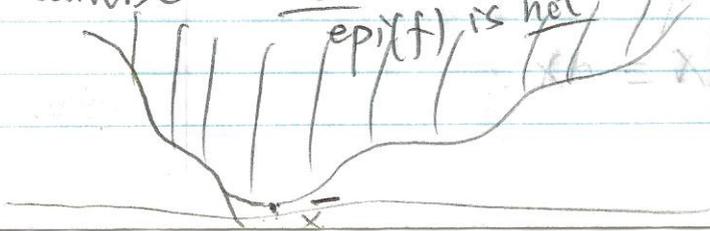
Th<sup>m</sup> If  $f$  is convex then  $S_r(f)$  is either  $\emptyset$  or a convex set

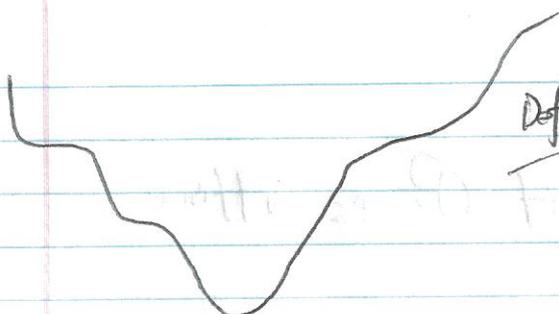
pf. Trivial.  $\square$

Converse is not true  $\text{epi}(f)$  is not convex

only stationary point if consider  $\emptyset$  take convex, remove this point

$\nabla f(\bar{x}) = 0$  is a global min. called pseudo convex.





Def<sup>n</sup> all sublevel-sets are convex  
are called quasi-convex functions.

Def<sup>n</sup> The indicator function of a set  $S \subseteq \mathbb{E}$ ,  

$$f(x) = \begin{cases} 0 & \text{if } x \in S \\ +\infty & \text{if } x \notin S \end{cases}$$

Theorem: A set is convex iff its indicator function is convex.

pf: trivial  $\square$

Operations that preserve convexity.  
Suppose we have a bunch of sets

$\{S_t\}_{t \in T}$  a collection of convex sets in  $\mathbb{E}$

$\bigcap_{t \in T} S_t$  is convex.

Suppose  $m = |T|$  is finite, then  $\bigoplus_{t \in T} S_t$  is convex.  
cardinality  $m$

also  $S_1 + S_2 + \dots + S_m$  is convex.

Suppose  $S$  is convex set

$A: \mathbb{E} \rightarrow \mathbb{Y}$  is affine,

then  $A(S)$  is convex.

proof: let  $x, y \in A(S)$

$\Rightarrow \exists w, z \in S \exists x = A(w), y = A(z)$

Then  $\forall \lambda \quad \lambda x + (1-\lambda)y = \lambda A(w) + (1-\lambda)A(z)$

$= (\lambda(\mathcal{L}(w) + b) + (1-\lambda)(\mathcal{L}(z) + b))$

$= \mathcal{L}(\lambda w + (1-\lambda)z) + b$

where  $A(u) = \mathcal{L}(u) + b \leftarrow \mathbb{R} \mathbb{Y}$   
linear

Hilroy

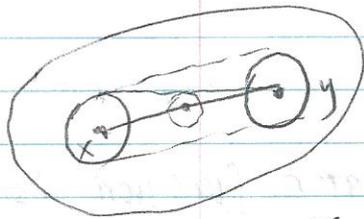
Similarly, the inverse image

$B^{-1}(S)$  is convex if  $B$  is affine.

$\text{int } S$  and  $\text{cl } S$  are convex.

proof: (for interior)

Let  $x, y \in \text{int}(S)$  (assume  $\text{int } S \neq \emptyset$ )



$$\exists \delta > 0 \rightarrow x + \delta B \subseteq S$$

$$\forall \lambda \in (0, 1) \Rightarrow (\lambda x + (1-\lambda)y) + \lambda \delta B$$

$$= \lambda x + \lambda \delta B + (1-\lambda)y$$

$$= \underbrace{\lambda(x + \delta B)}_{\subseteq S} + \underbrace{(1-\lambda)y}_{\in S} \subseteq S \quad \square$$

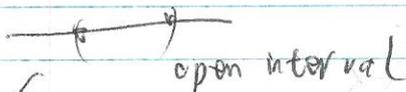
Aside: core/algebraic interior  
in finite dim.  $\text{core}(S) = \text{int}(S)$



Recall  $\text{rel int}(S) \neq \emptyset$  ( $S$  convex set)

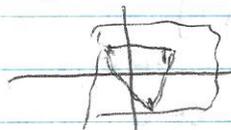
Answer  $\neq$  never empty.

proof:



1 dim

$\geq 2$  affinely ind. pts.



2 dim

$\exists$  affinely ind. pts.

proof for  $\text{cl}(S)$ ?

Let  $x, y \in \text{cl}(S)$

$$x = \lim_{i \rightarrow \infty} x_i \quad x_i \in S \quad y = \lim_{i \rightarrow \infty} y_i \quad y_i \in S$$

To show  $\lambda x + (1-\lambda)y \in \text{cl}(S)$   $\square$

Suppose  $\{f_t\}_{t \in T}$  convex functions

then  $\sup_{t \in T} \{f_t(x)\}$  is convex.

Also, finite nonnegative sum

$$\sum_{i=1}^m \alpha_i f_i(x) \text{ is convex}$$

$\alpha_i \geq 0 \forall i$

Let  $X_{(i)}$  be the  $i$ th largest component of  $X$

$$f(x) = \sum_{i=1}^K w_i X_{(i)} \text{ is convex.}$$

$w_i \geq 0$  nonneg. weights.

Support function of set  $C$ :  $f_C(x) = \sup \{ \langle x, y \rangle \mid y \in C \}$

distance to the furthest point  $f(x) = \sup_{y \in C} \|x - y\|$

$$f(x) = \lambda_{\max}(X) \quad X \in S^n \text{ symmetric}$$

$\lambda_{\max}$  largest eigenvalues

$$= \max \{ y^T X y \mid \|y\| = 1 \}$$

$$f(A) = \|A\| \text{ spectral norm (largest singular value)}$$

$$\max_{\|x\|=1} \max_{\|y\|=1} \langle Ax, y \rangle$$

Sept 20th, 2007 Convex Analysis

- Compositions of functions / sets
  - onto duality of
- examples
- i)  $g$  convex  $\Rightarrow e^{g(x)}$  is convex
  - ii)  $g$  concave  $> 0 \Rightarrow \log g(x)$  concave  
( $-g$  convex)
  - iii)  $\frac{1}{g(x)}$  convex
  - iv)  $\{X \in S^n, X \succ 0\} \Rightarrow \log \det X$  is concave (strictly)

note

- a)  $f$  convex iff.  $-f$  concave
- b)  $f$  is strictly convex on  $\Omega$  (convex) if  $f(\lambda x + (1-\lambda)y) < \lambda f(x) + (1-\lambda)f(y)$   
 $\forall \lambda \in (0,1)$   
 $\forall x \neq y, x, y \in \Omega$

Remarks:  $S$  convex set  
 $cl S = cl ri(S)$

$$\text{relint } S = \text{relint}(\text{cls } S)$$

$$\text{if } A \in \mathbb{R}^{n \times n} \quad A(\text{relint } S) = \text{relint}(AS)$$

$$A(\text{cls } S) \subseteq \text{cls}(AS)$$

with = iff  $S$  is bounded.

$$\text{relint}(S_1) \cap \text{relint}(S_2) \quad ?$$

↓  
cl

↓  
cl

Def<sup>n</sup> A convex combination is a linear combination of the form

$$X = \sum_{i=1}^m \lambda_i x_i \quad \text{where } \sum_{i=1}^m \lambda_i = 1 \quad \lambda_i \geq 0 \quad \forall i$$

$m=2$ ;

$$X \equiv [x_1, x_2]$$

$$x_1 \quad \lambda_1 x_1 + \lambda_2 x_2 \quad x_2$$

$$\lambda_1 + \lambda_2 = 1 \quad \lambda_1, \lambda_2 \geq 0$$



$m=3$

$$\lambda_1 x_1 + \lambda_2 x_2 + \lambda_3 x_3 = X$$

$$\lambda_1 + \lambda_2 + \lambda_3 = 1 \quad \lambda_1, \lambda_2, \lambda_3 \geq 0$$

equivalent

Def<sup>n</sup> (alternate) The convex hull of a set  $S \subseteq \mathbb{E}^n$  is the set of all convex combinations of its elements. Denoted  $\text{conv } S$

Proposition:  $S \subseteq \mathbb{E}^n$   $\text{conv } S$  is convex.

proof: Take  $x, y \in \text{conv } S$

$$\text{i.e. } X = \sum_{i=1}^m \lambda_i x_i \quad \sum \lambda_i = 1 \quad \lambda_i \geq 0 \quad \forall i$$

$$Y = \sum_{i=1}^p \mu_i y_i \quad \sum \mu_i = 1 \quad \mu_i \geq 0 \quad \forall i \quad y_i \in S \quad \forall i$$

Let  $\lambda \in (0, 1)$

$$Z = \lambda X + (1-\lambda) Y = \sum_{i=1}^m \lambda \lambda_i x_i + \sum_{i=1}^p (1-\lambda) \mu_i y_i$$

$$\sum_{i=1}^m \lambda \lambda_i + \sum_{i=1}^p (1-\lambda) \mu_i = \lambda + (1-\lambda) = 1 \quad \square$$

-  $S$  is convex iff  $S = \text{conv } S$

The convex hull of  $S$  is the intersection of all convex sets containing  $S$ .

proof: Since convex hull is convex,  $\Rightarrow \text{conv} S \supseteq T = \text{intersection of all convex sets containing } S$ .

Let  $C$  be a convex set,  $S \subseteq C$

$S \subseteq C \Rightarrow \text{conv} S \subseteq \text{conv} C = C$

i.e.  $\text{conv} S \subseteq T$  □

to get convex functions from general functions? use epi'f!

Def<sup>n</sup> The convex hull of a function  $f: X \rightarrow \mathbb{R} \cup \{\infty\}$  is defined on  $\text{conv} X$  by

$F(x) = \text{conv} f(x) := \inf \{r : (x, r) \in \text{conv}(\text{epi} f)\}$

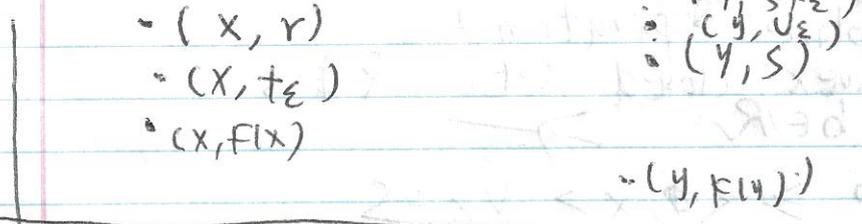
(Note: Since  $\text{conv} f(x) = \infty$  is possible,  $\text{conv} f$  is convex if  $\text{epi}(\text{conv} f)$  is convex)

Prop. The convex hull of a function  $f$  is convex.

proof: Take  $(x, r), (y, s) \in \text{epi}(\text{conv} f)$

Let  $\lambda \in (0, 1)$

Need to show  $(\lambda x + (1-\lambda)y, \lambda r + (1-\lambda)s) \in \text{epi}(\text{conv} f)$



$\forall \epsilon > 0 \exists t_\epsilon, u_\epsilon \in \mathbb{R}$   $(x, t_\epsilon) \in \text{conv}(\text{epi} f)$

$(y, u_\epsilon) \in \text{conv}(\text{epi} f)$

and  $r + \epsilon > t_\epsilon$   $s + \epsilon > u_\epsilon$

Hilroy

$$\lambda r + (1-\lambda)s + \varepsilon^{\lambda\varepsilon + (1-\lambda)\varepsilon} > \lambda t_\varepsilon + (1-\lambda)u_\varepsilon$$

Since  $\text{conv}(\text{epi } f)$  is convex.

$$\lambda(x, t_\varepsilon) + (1-\lambda)(y, u_\varepsilon) = (\lambda x + (1-\lambda)y, \lambda t_\varepsilon + (1-\lambda)u_\varepsilon)$$

$$F(\lambda x + (1-\lambda)y) \leq \inf \{ \lambda r + (1-\lambda)s + \varepsilon \mid (x, r) \in \text{conv epi}(f), (y, s) \in \text{conv epi}(f) \}$$

$$\Rightarrow F(\lambda x + (1-\lambda)y) \leq \lambda r + (1-\lambda)s \quad \square$$

The convex hull of  $f$  is the largest convex function majorized by  $f$ .

i.e. if  $h(x)$  is convex  $h(x) \leq f(x) \quad \forall x \in X$

$$\Rightarrow h(x) \leq \text{conv } f(x), \quad \forall x \in \text{conv}(X)$$

proof: Similar to above - exercise  $\square$

prop.  $\text{Conv } f(x) = \inf \left\{ \sum_{i=1}^m \lambda_i f(x^i) : \lambda_i \geq 0, \sum_{i=1}^m \lambda_i = 1, x^i \in X \right.$   
 $\left. \sum \lambda_i x^i = x \right\}$

proof: Exercise  $\square$

Theorem: Jensen's Inequality

$f$  convex on convex set  $X$ .

$$f\left(\sum_{i=1}^m \lambda_i x^i\right) \leq \sum_{i=1}^m \lambda_i f(x^i) \quad \forall x^i \in X$$

$$\lambda_i \geq 0 \quad \forall i \quad \sum \lambda_i = 1$$

$m$  is a positive integer

(integral version)

## Duality of Sets & Functions

Th<sup>m</sup> (Banc Hyperplane Separation)

If  $S \subseteq \mathbb{E}$  is convex closed set  $\bar{x} \notin S$

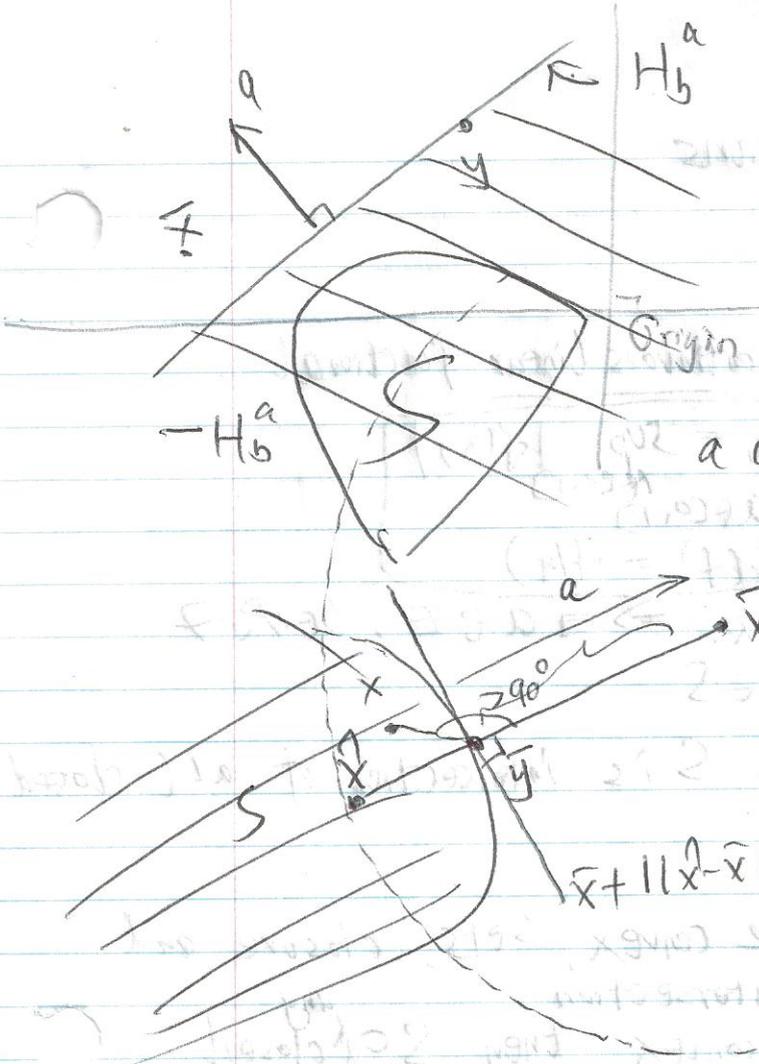
Then  $\exists a \in \mathbb{E} \quad b \in \mathbb{R}$

$$\langle a, \bar{x} \rangle > b \geq \langle a, x \rangle \quad \forall x \in S$$

Note  $H_a^b = \{x : \langle a, x \rangle = b\}$

hyperplane  $\nearrow$  normal to hyperplane

half space  $\rightarrow \{x : \langle a, x \rangle \leq b\}$



$$h = \langle a, y \rangle \quad \forall y \in H_b^a \quad (17)$$

proof: Assume  $S \neq \emptyset$  Let  $\bar{x} \in S$

$\Rightarrow T = S \cap \{ \bar{x} + \|\bar{x} - \bar{x}\| B \}$  is a compact convex non empty closed & bounded

draw ball of radius

$\|\bar{x} - \bar{x}\|$  about  $\bar{x}$

$$\min_{y \in T} f(y) = \min_{y \in T} \|y - \bar{x}\| \quad \text{continuous in } y$$

$\Rightarrow \exists \bar{y}$  which is the minimum.

$$a = \bar{x} - \bar{y}, \quad b = \langle \bar{x} - \bar{y}, \bar{y} \rangle$$

$$\text{So } \langle a, \bar{x} \rangle - b = \langle \bar{x} - \bar{y}, \bar{x} \rangle - \langle \bar{x} - \bar{y}, \bar{y} \rangle = \langle \bar{x} - \bar{y}, \bar{x} - \bar{y} \rangle > 0 \quad \text{since } \bar{x} \neq \bar{y}$$

And  $\langle a, x \rangle \leq b \quad \forall x \in S$  exercise  $\square$

Recall Convex combination

conv S = set of all convex comb.   
 convex Hull =  $\{ x : x = \sum_{i=1}^m \lambda_i x_i \mid \sum_{i=1}^m \lambda_i = 1, \lambda_i \geq 0 \forall i \}$  for some integer  $m > 0$  (conv set)

f is convex on  $\Omega = \text{dom } f$  (convex set)

$$f\left(\sum_{i=1}^m \lambda_i x_i\right) \leq \sum_{i=1}^m \lambda_i f(x_i) \quad \text{and } \otimes$$

Hilroy

# Duality of convex sets & Functions

$X$  vector space (topology)

$X^*$  dual space - set of all continuous linear functional

in  $f(x) = \langle a, x \rangle$

$$\|f\|_0 = \sup_{\|x\|_0=1} |f(x)|$$

$$a \in \mathbb{R}^n$$

$$f_a(f) = f(a)$$

## Th<sup>m</sup> Basic Separation

$\bar{x} \notin S \subseteq \mathbb{R}^n$ ,  $S$  is closed convex.  $\Rightarrow \exists a \in \mathbb{R}^n, b \in \mathbb{R}$   $\nabla x \in S$   
 $\langle a, \bar{x} \rangle > b \geq \langle a, x \rangle$

Cor.  $S$  is closed and convex iff.  $S$  is intersection of all (closed) half-spaces containing it.

proof: "Suff" ( $\Leftarrow$ ) half spaces are convex sets, closure and convexity is "closed" under intersection.

"Nec" ( $\Rightarrow$ ) suppose  $S$  is closed, convex then  $S$  is <sup>any</sup> closed half spaces containing  $S$ .

Let  $\bar{x} \notin S$  Apply th<sup>m</sup> to get the halfspace containing  $S$  but not  $\bar{x}$ .

$$\langle a, y \rangle > b > \langle a, x \rangle \quad \forall x \in S, \forall y \in T$$

strict separation.

$S, T$  convex sets,  $S \cap T = \emptyset$ .

a)  $S - T$  closed.

b)  $S$  closed,  $T$  compact

c) both  $S, T$  are obtained from quadratic ineq.

examples of strict separation:  $E = \{x: q_j(x) \leq 0, j=1, \dots, k\}$   $E_j(x) = \frac{1}{2} x^T Q_j x + b_j^T x + c_j$

of strict separation.

Suppose  $O \in S$ , closed, convex

$$\Rightarrow S = \bigcap_{(a,b) \in T} \{x: \langle a, x \rangle \leq b\} \text{ and } O \in S \Rightarrow b \geq 0$$

for some set in  $\mathbb{R} \oplus \mathbb{R}$

Zifact.  $S = \bigcap_{(a,1)} \{x : \langle a, x \rangle \leq 1\}$  (b can be fixed = 1)

proof: if  $b > 0$ , then use  $(\frac{1}{b}a, 1)$

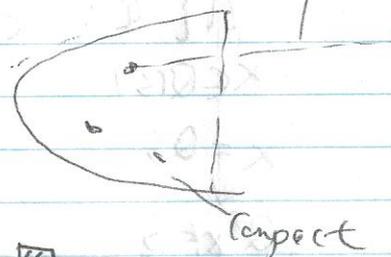
Let  $(a, 0) \in T$  and  $\bar{x} \Rightarrow \langle a, \bar{x} \rangle > 0$   
 half space going thr. origin.

unbounded

Take  $b = \frac{1}{2} \langle a, \bar{x} \rangle > 0$

$$\Rightarrow \langle a, \bar{x} \rangle > b \geq \langle a, x \rangle \quad \forall x \in S.$$

$$\Rightarrow (a, b) \notin T \quad \text{so } (\frac{1}{b}a, b) \in T \quad \square$$



compact

Def<sup>n</sup> The polar set of  $S \subseteq E$ ,  $S \neq \emptyset$ , is  $S^\circ = \{a : \langle x, a \rangle \leq 1 \quad \forall x \in S\} = \bigcap_{x \in S} \{a : \langle x, a \rangle \leq 1\}$  intersection half spaces

So  $S^\circ$  is closed convex.

Let  $\bar{x} \in S$ , so  $0 \in S - \bar{x}$

for infinite dim. order matters, has different meaning.

Def<sup>n</sup> The polar set of  $S$  wrt.  $\bar{x} \in S$  is

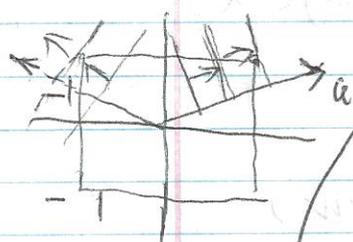
$$S^\circ(\bar{x}) := (S - \bar{x})^\circ = \{a : \langle x - \bar{x}, a \rangle \leq 1 \quad \forall x \in S\}$$

Then

$$S = (S^\circ(\bar{x}))^\circ = (S - \bar{x})^{\circ\circ} = \bigcap_{a \in S^\circ(\bar{x})} \{x : \langle a, x - \bar{x} \rangle \leq 1\}$$

if  $S$  is closed convex.

eg. A)  $B_\infty = \{x \in \mathbb{R}^2 : \max_i |x_i| \leq 1\}$



$$B_\infty^\circ = \{a \mid a^T x = a_1 x_1 + a_2 x_2 \leq 1 \quad \forall x \in B_\infty\}$$

$$\max a_1 x_1 + a_2 x_2$$

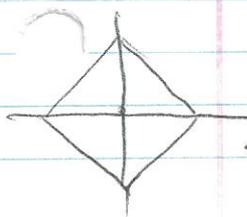
$$\text{s.t. } |x_1| + |x_2| \leq 1$$

$$\rightarrow \|a\|_1 = |a_1| + |a_2|$$

$$x_i = \text{sgn } a_i$$

$$= B_1 = \{a \mid \sum_i |a_i| \leq 1\}$$

$$\text{so } B_1^\circ = B_\infty$$



$Q = Q^T$  orth. matrix  $0 \in S$

$B_1^0 = (Q(S))^0 = Q(S^0)$  ? conjecture

$0 \in (QS)$  ✓

$\bigcap_{x \in Q(S)} \{a : \langle a, x \rangle \leq 1\} = \bigcap_{s \in S} \{a : \langle a, QS \rangle \leq 1\}$

$x = QS$   
 $\Downarrow$   
 $Q^T x = S$

$= \bigcap_{s \in S} \{a : \langle Q^T a, s \rangle \leq 1\}$

$= Q \left( \bigcap_{s \in S} \{w : \langle w, s \rangle \leq 1\} \right) = Q(S^0)$

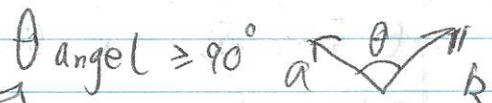
Questions a)  $B_2^0 = B_2$  ?

b)  $B_p^0 = B_q$  if  $\frac{1}{p} + \frac{1}{q} = 1$  ?  $p \geq 1$

Recall:  $K$  is a convex cone if

$K + K \in K$

$\alpha K \subset K \quad \forall \alpha > 0$



Then  $K^0 = K^- = \{a : \langle a, k \rangle \leq 0, \forall k \in K\}$  — negative polar

is a closed convex cone (C.C.C.)

a)  $S^- = (cl(S))^- = (conv(S))^- = (\text{cone } S)^-$

Suppose  $0 < \langle a, k \rangle \leq 1 \quad k \in K$  cone  $\Rightarrow k \neq 0 \quad \alpha k \in K \quad \forall \alpha \nearrow +\infty$   
 $0 < \langle a, \alpha k \rangle = \alpha \langle a, k \rangle \nearrow +\infty > 1$  for large  $\alpha$ .  
 Contradicts  $\leq 1$

b)  $S^- = cl(\text{cone } S)$

c)  $S^- = S$  iff  $S$  is a C.C.C. (on HW)

Now for duality of convex functions

"Fenchel dual" of  $f$ . "Fenchel Conjugate"

$f^*(\phi) = \sup_{x \in E} \{ \langle \phi, x \rangle - f(x) \}$

# Convex Analysis and Optimization

Sept 27 2008

21

Dimitri P.

Recall.

polar set  $S^\circ = \{y \mid \langle y, s \rangle \leq 1, \forall s \in S\}$

$= \bigcap_{s \in S} \{y \mid \langle y, s \rangle \leq 1\}$

half spaces.

$\forall \bar{x} \in S$

$$S(\bar{x})^\circ = (S - \bar{x})^\circ$$

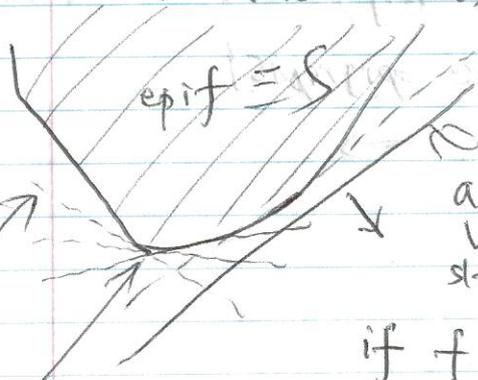
$$S(\bar{x})^{\circ\circ} = \text{cl}(\text{conv}(S - \bar{x}))$$

so  $\text{cl}(\text{conv}(S)) = \bar{x} + S(\bar{x})^{\circ\circ}$

non positive polar  
Negative

$$S^- = \{y \mid \langle y, s \rangle \leq 0, \forall s \in S\}$$

and if  $K$  is a cone, then  $K^- = K^\circ$



nonvertical halfspaces that contain epif

make closer to epif as much as possible

$$ax+rb=1$$

slope unchanged (Fermat's theorem)

if  $f$  is differentiable,  $\bar{x} \in \text{argmin}_x f(x) \Rightarrow$

$$\nabla f(\bar{x}) = 0$$

if  $f$  is NOT differentiable, we use "sub gradient".

$$\partial f(x) = \{ \phi \mid \langle \phi, y-x \rangle \leq f(y) - f(x), \forall y \}$$

rewritten as  $f(x) + \langle \phi, \Delta x \rangle \leq f(x + \Delta x)$

line

$$T = f(x) + \langle \phi, \Delta x \rangle$$

in non-diff. case (convex)

$$\bar{x} \in \text{argmin}_x f(x) \Rightarrow 0 \in \partial f(\bar{x})$$

Fenchel conjugate

Hilroy

Def<sup>n</sup>

A function  $h: E \rightarrow [-\infty, +\infty]$  is closed, if its epigraph is closed.

Def<sup>n</sup> A function  $h$  is lower semi-continuous at  $\bar{x}$  if

$$\liminf h(x_j) \geq h(\bar{x}) \text{ when } x_j \rightarrow \bar{x}$$

$h$  is lower semi-continuous (LSC) at all  $\bar{x}$

A function is closed iff. all its sub-level sets are closed iff  $A$  is LSC.

prop. A function is closed and convex iff.  $\forall x \in E, f(x) = \sup \{Ax\}$ ;  $A$  is affine and majorized by  $f$ .

proof: Suff. ("E")

the epigraph of affine functions are halfspace (closed & convex) and  $\text{epi}(f) = \{ \text{intersection of these epigraphs} \}$

Nec ( $\Rightarrow$ ):  $f$  is closed & convex.  $\Rightarrow$   $\text{epi}(f)$  is closed, convex

WLOG.  $0 \in \text{epi}(f)$

$$\text{Let } T = (\text{epi}(f))^0$$

$$\text{So } \text{epi}(f) = T^0 = \text{intersection } \bigcap \{ (x, r) : \langle a, x \rangle + br \leq 1 \}$$

Since  $(x, r) \in \text{epi}(f)$  we can let  $r \nearrow +\infty$  so  $\Rightarrow b \leq 0$

Now show  $\exists (a, b) \in T, \Rightarrow b < 0$

Suppose not, let  $\bar{x} \in \text{dom } f$

$$(\bar{x}, f(\bar{x})) \in \text{epi}(f) \Rightarrow (\bar{x}, r) \in \text{epi}(f) \forall r \in \mathbb{R}$$

$\Rightarrow f(\bar{x}) = -\infty$ , contradiction

$$\Rightarrow (a, b) \in T \text{ with } b < 0$$

Pick any  $(\bar{a}, 0) \in T$  where  $\langle \bar{a}, \bar{x} \rangle > 1$

$T$  is convex  $\Rightarrow$

$$(\lambda \bar{a} + (1-\lambda)\hat{a}, (1-\lambda)b) \in T \forall \lambda \in (0, 1)$$

For any  $r \in \mathbb{R}$ .

$$\langle \lambda \bar{a} + (1-\lambda)a, \bar{x} \rangle + (1-\lambda)br$$

$$= \lambda \langle \bar{a}, \bar{x} \rangle + (1-\lambda) [\langle a, \bar{x} \rangle + br]$$

$\rightarrow$  If  $\lambda$  is close to 1 (i.e. as  $\lambda \rightarrow 1$ )

So  $\exists \bar{\lambda} \in (0, 1) \nearrow$

$$\langle \bar{\lambda} \bar{a} + (1-\bar{\lambda})a, \bar{x} \rangle + (1-\bar{\lambda})br > 1$$

$$\Rightarrow \text{epi}(f) = \bigcap_{\substack{(a,b) \in T \\ b < 0}} \{ (x,r) \mid \langle a, x \rangle + br \leq 1 \}$$

So  $(a,b) \in T \iff \langle a, x \rangle + br \leq 1 \quad \forall (x,r) \in \text{epi} f$

$$\iff \langle -\frac{a}{b}, x \rangle + \frac{1}{b} \leq r \quad \forall (x,r) \in \text{epi} f$$

Affine Function majorized by  $f(x)$  (can choose  $-f(x)$ )

$$\text{or } -\frac{1}{b} \geq \sup_{x \in E} \{ \langle -\frac{a}{b}, x \rangle - f(x) \}$$

$$\Leftrightarrow (-\frac{a}{b}, -\frac{1}{b}) \in \text{epi}(g) \text{ where } g: \mathcal{P} \rightarrow \sup_{x \in E} \{ \langle \mathcal{P}, x \rangle - f(x) \}$$

Note: denote

Def<sup>n</sup> The Fenchel conjugate of  $f$  is denoted

$$f^* \mathcal{P} \Rightarrow \sup_{x \in E} \{ \langle \mathcal{P}, x \rangle - f(x) \}$$

Note  $f \geq h \Rightarrow h^* \geq f^*$  (reverse ordering)

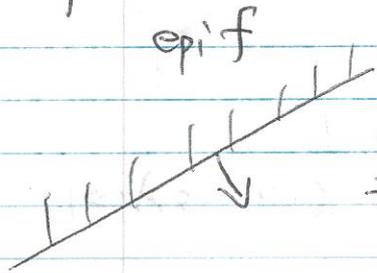
Theorem (Fenchel-Young Inequality)

$$\forall \mathcal{P}, x \in E \quad f(x) + f^*(\mathcal{P}) \geq \langle \mathcal{P}, x \rangle$$

(Equality holds if  $x \in \text{dom} f$ , and  $\mathcal{P} \in \partial f(x)$ )

Examples:

①  $f: X \rightarrow \langle a, x \rangle + \alpha$



$$f^*(\rho) = \sup_{x \in E} \{ \langle \rho, x \rangle - f(x) \}$$

$$= \sup_{x \in E} \{ \langle \rho - a, x \rangle - \alpha \}$$

$$= \begin{cases} +\infty & \text{if } \rho \neq a \\ -\alpha & \text{if } \rho = a \end{cases} = \delta_{\{a\}}(\rho) - \alpha$$

②  $f: X \rightarrow \delta_{\{a\}}(x) - \alpha$

$$f^*(\rho) = \langle a, \rho \rangle + \alpha$$

$$f^{**} = f$$

③  $f: x \in \mathbb{R}^n \rightarrow \frac{1}{2} x^T Q x$  (where  $Q = Q^T \succ 0$ )

$$f^*(\rho) = \sup_{x \in \mathbb{R}^n} \{ \rho^T x - \frac{1}{2} x^T Q x \}$$

$$0 = \nabla = \rho - Qx \Rightarrow x = Q^{-1}\rho$$

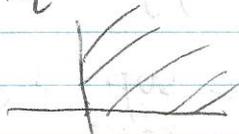
$$\rho^T Q^{-1}\rho - \frac{1}{2} \rho^T Q^{-1} Q Q^{-1} \rho = \frac{1}{2} \rho^T Q^{-1} \rho$$

So  $f^{**}(x) = \frac{1}{2} x^T Q x = f(x)$

④ log-barrier function.

$$f: x \in \mathbb{R}^n \rightarrow -\sum_{i=1}^n \log x_i$$

$$\min C^T x - \alpha (\log x_1 + \log x_2)$$



$$f^*(\rho) = \sup_{x > 0} \{ \rho^T x + \sum_{i=1}^n \log x_i \}$$

$$0 = \nabla = \rho + \begin{bmatrix} \frac{1}{x_1} \\ \vdots \\ \frac{1}{x_n} \end{bmatrix}$$

$$\Rightarrow x = - \begin{pmatrix} \frac{1}{\rho_1} \\ \vdots \\ \frac{1}{\rho_n} \end{pmatrix} > 0 \quad \dots \quad f^*(\rho) = -n + \sum_{i=1}^n \log \left( -\frac{1}{\rho_i} \right) \text{ if } \rho_i < 0 \forall i$$

pp 25-28

OMG, I lost these 4 pages.

Next  $f_S(x), f^*(x)$  support functions.  
 Sept - Oct 4 2007

$$f^*(\phi) = \sup_x \{ \langle \phi, x \rangle - f(x) \}$$

so  $f^*$  is a sup of linear in  $\phi$  (for each  $x$ )

So  $f^*$  is convex, epi closed convex

We saw  $f(x) = -\log \det(x)$   $X \in S_{++}^n$

Then  $f^*(\Phi) = -n + f(-\Phi)$   $\Phi \in S_{--}^n$

$Z_S f^{***} = f$  ?

"Comments" on convex fns

$$f: \mathbb{R}^n \rightarrow \bar{\mathbb{R}} \quad (= (-\infty, +\infty])$$

$$\partial f(x) = \{ \phi \mid \phi^T(y-x) \leq f(y) - f(x), \forall y \}$$

sub differential at  $x$

subgradient

$$x \in S_{\Omega}, \bar{x} \in \arg \min_{x \in \Omega} f(x) \iff 0 \in \partial f(\bar{x})$$

open, convex set

$\partial f(\bar{x})$  is a singleton iff  $f$  is differentiable at  $\bar{x}$ .

ie.  $\partial f(\bar{x}) = \{ \nabla f(\bar{x}) \}$

Fenchel-Young Ineq.

$$f(x) + f^*(y) \geq \langle x, y \rangle \quad \forall x, y$$

equality holds iff  $y \in \partial f(x)$  iff  $x \in \partial f^*(y)$

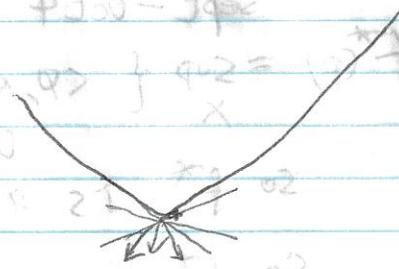
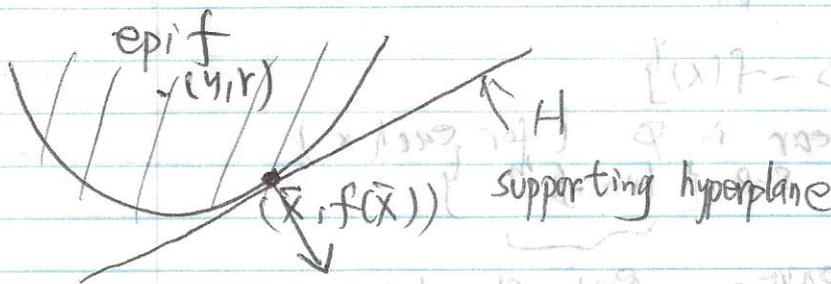
Very soon we will get to optimality conditions and duality for convex programs.

eg.  $p = \inf_{x \in \mathbb{R}^n} \{ f(x) + g(Ax) \}$

primal  $d = \sup_{z \in \mathbb{R}^m} \{ f^*(-A^T z) + g^*(z) \}$

dual

$f$  convex



Recall  $f$  is convex iff  $f(\bar{x} + \Delta x) \geq f(\bar{x}) + \nabla f(\bar{x})^T \Delta x = f(\bar{x}) + \nabla f(\bar{x})^T (y - \bar{x})$   
 $y = \bar{x} + \Delta x = f(\bar{x}) + \nabla f(\bar{x})^T y - \nabla f(\bar{x})^T \bar{x}$

$\Rightarrow \nabla f(\bar{x})^T \bar{x} - f(\bar{x}) \geq \nabla f(\bar{x})^T y - 1 \cdot r$   
 $\nabla f(\bar{x})^T \bar{x} - f(\bar{x}) \in \langle \nabla f(\bar{x})^T, -1 \rangle (y, r)$

$f^*(\nabla f(\bar{x}))$  normal to hyperplane

eg.  $f(x) = \frac{1}{2} \|x\|^2$  then  $\nabla f(x) = x$

so  $f^*(\nabla f(\bar{x})) = \bar{x} \bar{x} - f(\bar{x}) = f(\bar{x})$

$\frac{\bar{x}}{\|x\|^2} f^* = f$

Recall for  $f(x) = \frac{1}{2} x^T Q x$ ,  $Q \succ 0$

$f^*(\phi) = \frac{1}{2} \phi^T Q^{-1} \phi$

converse? Suppose  $f^* = f$

then  $2f(x) = f(x) + f^*(x) \geq \langle x, x \rangle = \|x\|^2$

i.e.  $f(x) \geq \frac{1}{2} \|x\|^2$

But also.  $f(x) = f^*(x) = \sup_y \{ \frac{1}{2} \|x\|^2 \langle x, y \rangle - f(y) \}$

$\Rightarrow f(x) = \frac{1}{2} \|x\|^2 \leq \frac{1}{2} \|x\|^2$

For  $\mathbb{E} = \mathbb{S}^n$   
 $\|x\|_2^2 = \|x\|_F^2 = \text{tr } X^2$

when is  $f^{**} = f$ ? (we know  $f^{**}$  is lsc, convex)

consider  $f = \delta_S \leftarrow$  indicator function  
 $0 \neq S \in \mathbb{E}$

$$\delta_S^*(\rho) = \sup_{x \in \mathbb{E}} \{ \langle \rho, x \rangle - \delta_S(x) \} = \sup_{x \in S} \langle \rho, x \rangle$$

$$= \sigma_S(\rho) \text{ called support function of } S.$$

$S = \{e_1, e_2, \dots, e_n\}$  unit vectors

$$f(x) = \max_i x_i = \max_i e_i^T x = \sigma_S(x)$$

always replace by  $\text{conv } S$

$\mathbb{Z}_S f^*(\rho) = \delta_{\text{conv}(S)}(\rho)$ ?

$x_{[1]}$  ordered vectors  $x_{[1]} \geq x_{[2]} \geq \dots \geq x_{[n]}$

$$f(x) = \max \sum_{i=1}^r w_i x_{[i]} \quad w_i \geq 0$$

$\mathbb{Z}_S f(x) = \sigma_S(x)$ ?

ie.  $S = \{v = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} : v_i \in \{0, 1\}, \text{ vectors } e_n^T v = r\}$

$S = \{v = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ w_1 \\ \vdots \\ w_r \end{pmatrix} : \text{with } w_1, \dots, w_r \text{ occurring exactly once in each vector}\}$

eg.  $n=3$

	0	5	5	7	0	7	
$w_1=5$	5	0	7	5	7	0	...
$w_2=7$	7	7	0	0	5	5	$\frac{r}{1}$

Def<sup>n</sup> A function is positively homogeneous if

$$f(\lambda x) = \lambda f(x) \quad \forall \lambda > 0$$

$$\forall x \in \mathbb{E}$$

Hilroy

(equiv.  $f(\lambda x) \leq \lambda f(x) \quad \forall \lambda > 0, \forall x \in \mathbb{E}$ )  
 ex ec.

suppose  $S$  is closed and convex. Then  $\delta_S$  is ccp. and Fenchel-Young applies:

$$\begin{aligned} z \in \partial \delta_S(y) &\Leftrightarrow y \in \partial \sigma_S(z) \Leftrightarrow \delta_S(y) - y^T z + \sigma_S = 0 \\ &\Leftrightarrow y \in S \quad \forall y \in S \quad z^T (y - y) \leq 0 \end{aligned}$$

Oct 9th, 2007

When is  $f^{**} = f$ ?

(related to when is  $S^{\circ\circ} = \bar{S}$ ?)

Recall.  $\delta_S = \begin{cases} 0 & \text{if } x \in S \\ +\infty & \text{if } x \notin S \end{cases}$       indicator function.       $\delta_S^*(\varphi) = \sup_{x \in \mathbb{E}} \{ \langle \varphi, x \rangle - \delta_S(x) \}$   
 $= \sup_{x \in S} \langle \varphi, x \rangle =: \sigma_S(\varphi)$

support function of  $S = \sigma_{\text{conv}(S)}(\varphi) = \sigma_{\text{cl}(\text{conv}(S))}(\varphi)$

Def<sup>n</sup>  $f$  is positively homogeneous if  $f(\lambda x) = \lambda f(x) \quad \forall \lambda > 0, \forall x \in \mathbb{E}$

$\sigma_S(\cdot)$  is positively homogeneous and convex.

called sublinear

ie. equiv.  $f(\lambda x + \mu y) \leq \lambda f(x) + \mu f(y) \quad \forall \lambda, \mu > 0, \forall x, y \in \mathbb{E}$

Note if  $f$  pos. homog., then  $f^* = \delta_{\{ \varphi \mid \langle \varphi, x \rangle \leq f(x) \quad \forall x \}}$

Def<sup>n</sup> The set supported by  $f$  is

$$S_f = \{ \varphi \mid \langle \varphi, x \rangle \leq f(x) \quad \forall x \}$$

suppose  $K$  is a cone,

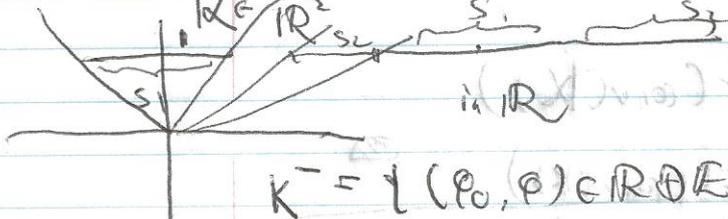
$$\begin{aligned} \delta_{S_K} &= \{ \varphi \mid \langle \varphi, x \rangle \leq \delta_K(x) \quad \forall x \} \\ &= \{ \varphi \mid \langle \varphi, x \rangle \leq 0 \quad \forall x \in K \} \\ &= K^- = K^\circ \end{aligned}$$

For a general set, we can homogenize - change to a cone.

$S$  need not to be convex

Let  $\emptyset \neq S \subseteq \mathbb{E}$

$$K = \{ (x_0, x) \in \mathbb{R} \oplus \mathbb{E} : x_0 > 0, \frac{1}{x_0} x \in S \}$$



$$K^- = \{ (\rho_0, \rho) \in \mathbb{R} \oplus \mathbb{E} : \rho_0 x_0 + \langle \rho, x \rangle \leq 0, \forall (x_0, x) \in K \}$$

and  $(1, x) \in K$  iff  $x \in S$   
 so  $(-1, \rho) \in K^-$  iff  $-x_0 + \langle \rho, x \rangle \leq 0 \quad \forall (x_0, x) \in K$   
 iff  $-1 + \langle \rho, x \rangle \leq 0 \quad \forall (1, x) \in K$   
 $\Leftrightarrow \langle \rho, x \rangle \leq 1 \quad \forall x \in S$

$$\Leftrightarrow \rho \in S^0 \quad Ax \leq b$$

$$\min x_1 + x_2 + x_3 \quad \text{st} \quad x_1 + x_2 \leq 3, -1 \leq x_1 \leq 1$$

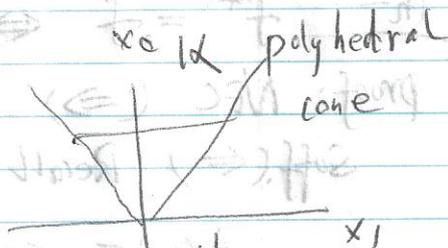
$$K = \{ (x_0, x_1, 0) : x_0 > 0, -1 \leq \frac{x_1}{x_0} \leq 1 \}$$

We can change

Summary: set constraints to nice cone constraints

$$\text{e.g. } g(x) \in S$$

$$(x_0, g) \in K$$



prop.  $\delta_S^{**} = \delta_S$  iff  $S$  is closed and convex

proof: Nec. ( $\Rightarrow$ ) follows since  $\delta_S^{**}$  is convex and LSC

Suffi. ( $\Leftarrow$ ) Suppose  $S$  is closed and convex

$$\text{We know } (\delta_S^{**})^* = \delta_{S^*} = \delta_T, \quad T = S^0_{S^*}$$

$$\text{i.e. } T = \{ x : \langle \rho, x \rangle \leq \sigma_S(\rho) \quad \forall \rho \}$$

We need only show  $S = T$

$$S \subseteq T \text{ clear since } \langle \rho, \bar{x} \rangle \leq \sup_{x \in S} \{ \langle \rho, x \rangle \} = \sigma_S(\rho), \quad \forall \rho \in \mathbb{E}^*$$

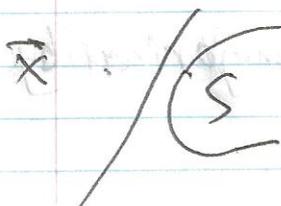
$$\Rightarrow \bar{x} \in T$$

Suppose  $\bar{x} \notin S$  a closed convex set.  $\Rightarrow \sup_{s \in S} \{ \langle \rho, s \rangle \}$

Apply Hyperplane separation theorem:

$$\exists \rho, \alpha \text{ s.t.}$$

$$\langle \rho, \bar{x} \rangle > \alpha > \sigma_S(\rho)$$



$$\Rightarrow \bar{x} \notin T \quad \square$$

prop.  $\delta_S^{**} = \delta_{cl(conv(S))}$

$\forall \emptyset \neq S \subseteq \mathbb{E}$

Pf: exercise  $\square$

cor. If  $\emptyset \neq K$  cone, then  $K^{\circ\circ} = cl(conv(K))$

Pf:  $\delta_{K^{\circ\circ}} = \delta_{K^{\circ}}^* = \delta_{K^{\circ\circ}}^{**} = \delta_{cl(conv(K))}$   $\square$

Also  $S^{\circ\circ} = cl(conv(S \cup \{0\}))$   $\forall S \neq \emptyset$

And  $\delta_S^{**} = \delta_{cl(conv(S))}^*$

Th<sup>m</sup>  $f^{**} = f \iff f$  is closed and convex.

proof: NEC ( $\implies$ ) clear.

Suff ( $\impliedby$ ) Recall  $f(x) = \sup \{ A(x) : A \text{ is affine } A \leq f \}$

$A(x) = \langle a, x \rangle + b$

$\implies A^*(x) = \delta_{\{a\}}(x) - b$   $f^{**}(x) = \sup_{\emptyset} \{ \langle \phi, x \rangle - f^*(\phi) \}$

$\implies A^{**}(x) = \langle a, x \rangle + b = A(x)$   $= \sup_{\phi} \{ \langle \phi, x \rangle - \sup_y \{ \langle \phi, y \rangle - f(y) \} \}$

If  $A$  is affine and  $A \leq f$ , then  $f \geq f^{**}$

$\geq A^{**} = A$

So  $f(x) \geq f^{**}(x) \geq \sup \{ A(x) : A \text{ affine } A \leq f \} = f(x) \quad \forall x$

Choose  $y=x$   
get a larger value

$\implies f = f^{**}$   $\square$

$A \leq f$  see text for gauge functions

$f^* \leq A^*$  web page with definitions - in progress

$A^{**} \leq f^{**}$

Def<sup>n</sup> The closure of a function  $f$  is

$cl(f)(x) := \inf_{k \rightarrow \infty} \{ \lim_{k \rightarrow \infty} f(x^k) : x^k \rightarrow x \}$

so ①  $epi(cl(f)) = cl(epi(f))$

②  $cl(f)$  is the largest closed function majorized by  $f$

Cor.  $f^{**} = cl(\text{conv}(f))$

proof:  $f \geq cl(\text{conv}(f)) \Rightarrow f^{**} \geq cl(\text{conv}(f))^{**}$   
 $= cl(\text{conv}(f))$  largest closed function, convex,  
majorized by  $f \geq f^{**}$   $\square$

More on support functions:

$\sigma_S(\varphi) = \sup_{x \in S} \langle \varphi, x \rangle$  is a proper, closed, sublinear.  
 $\forall \varphi \in S^{\circ} \subseteq E$

prop. If  $f$  is a proper, closed, sublinear, then  $f = \sigma_{S_f}$   
( $S_f =$  set supported by  $f$ )

proof:  $f$  is positively homogeneous,

$\Rightarrow f^* = \delta_{S_f} \Rightarrow f^{**} = \sigma_{S_f} \Rightarrow f = f^{**} = \sigma_{S_f}$   $\square$

So we have a bijection between {closed, convex set}  $\leftrightarrow$  {closed, sublinear functions}  
easy def.  $\rightarrow$  a little tricky  $\leftarrow$

Now suppose  $S$  is closed convex set,  $0 \in S$

$\sigma_{S^{\circ}}(x) = \sup_{\varphi \in S^{\circ}} \langle \varphi, x \rangle$   
 $S^{\circ} = \{\varphi : \langle \varphi, x \rangle \leq 1, \forall x \in S\}$

So  $\Rightarrow \sigma_{S^{\circ}}(x) \leq 1 \quad \forall x \in S$

If  $x \in \lambda S \quad \lambda \geq 0$ , then  $\langle x, \varphi \rangle = \lambda \langle \frac{1}{\lambda}x, \varphi \rangle \leq \lambda$   
 $\forall \varphi \in S^{\circ}$

So  $\sigma_{S^{\circ}}(x) \leq \gamma_S(x) := \inf \{\lambda : x \in \lambda S\}$   
called the gauge of  $S$ .

properties of  $\gamma_S$

- ①  $\gamma_S \geq 0 \quad \gamma_S(0) = 0$
- ②  $\gamma_S$  is positively homogeneous.
- ③ If  $S$  is convex, then  $\gamma_S$  is sublinear.
- ④ If  $S$  is closed & convex, then  $\gamma_S$  is closed and sublinear.

Hilroy

... ④ cont.  $\Rightarrow \mathcal{D}_S = \mathcal{D}_S^{**} = \mathcal{D}_{S^0}^* = \mathcal{D}_S^0$

⑤  $\mathcal{D}_S(x) = \inf\{\lambda > 0 : \langle \phi, x \rangle \leq \lambda \mathcal{D}_S(\phi)\}$

A gauge function is a nonnegative, <sup>∅</sup>sublinear function with  $\mathcal{D}_S(0) = 0$

Def<sup>n</sup> The polar of a function  $g$  is  $g^0(\phi) := \inf\{\lambda > 0 : \langle \phi, x \rangle \leq \lambda g(x) \forall x \in X\}$

So  $\mathcal{D}_S = \mathcal{D}_S^0$  and  $\mathcal{D}_{S^0} = \mathcal{D}_S^0$  when  $S$  is closed, convex,  $0 \in S$

Example: A norm is a gauge function.

$\|\cdot\| : E \rightarrow \mathbb{R}$   $\text{dom}(\|\cdot\|) = E$

a)  $\|x\| \geq 0$  with equality iff  $x=0$

b)  $\|\alpha x\| = |\alpha| \|x\|$

c)  $\|x+y\| \leq \|x\| + \|y\| \quad \forall x, y \in E$

def<sup>n</sup>  $\|\cdot\|^0(\phi) := \inf\{\lambda > 0 : \langle \phi, x \rangle \leq \lambda \|x\|, \forall x\}$

is a norm called the dual norm of  $\|\cdot\|$ .

Also  $\|\phi\|^0 = \sup_{x \neq 0} \frac{\langle \phi, x \rangle}{\|x\|}$

$\|x\|_p^0 = \|x\|_q$  where  $\frac{1}{p} + \frac{1}{q} = 1 \quad (p > 1)$

$\|x\|_1^0 = \|x\|_\infty$

$S_{\|\cdot\|} = \{\phi : \langle \phi, x \rangle \leq \|x\| \forall x\}$

is a closed convex set.

conversely, the gauge function of a closed convex set containing the origin is a norm.

$S, \langle \phi, x \rangle \leq \|x\| \forall x \iff \frac{\langle \phi, x \rangle}{\|x\|} \leq 1 \forall x \neq 0$

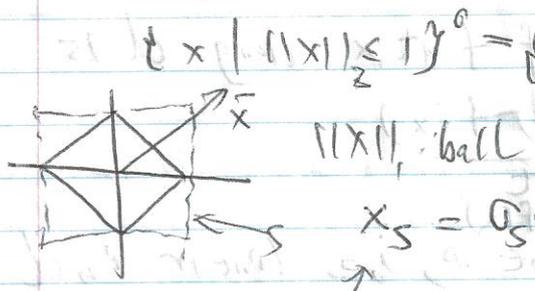
$\iff \|\phi\|^0 \leq 1$

i.e. unit ball of dual norm.

$\Rightarrow S_{\|\cdot\|} = \{\phi : \|\phi\|^0 \leq 1\}$

$S_{\|\cdot\|}^0 = \{\phi : \|\phi\| \leq 1\} = (S_{\|\cdot\|})^0$

$\|\cdot\|_2$  in  $\mathbb{R}^2$



primal geometry      dual geometry

Unconstrained Optimization  $0$  is hidden

$$\inf_{x \in E} \{f(x)\} = -\sup_{x \in E} \{ -f(x) \}$$

$$= -f^*(0)$$

Suppose  $x \in \arg \min_{x \in E} f(x)$  Then

$$f(x) = -f^*(0) \Rightarrow \langle x, 0 \rangle + f(x) = \langle x, 0 \rangle$$

where  $\langle x, 0 \rangle + f(x)$  is equivalent to  $\langle x, 0 \rangle$  (Fenchel-Young)

$$\text{then } \langle x, 0 \rangle - f(x) = \sup_{y \in E} \{ \langle y, 0 \rangle - f(y) \}$$

$$\langle x, 0 \rangle - f(x) \geq \langle y, 0 \rangle - f(y) \quad \forall y \in E$$

$$\Rightarrow f(y) - f(x) \geq \langle 0, y - x \rangle \quad \forall y \in E$$

called subgradient of  $f$  at  $x$ .

Def<sup>n</sup> The subdifferential of  $f$  at  $x$  is

$$\partial f(x) = \{ \text{subgradient at } x \}$$

Th<sup>m</sup> So  $x \in \arg \min_{x \in E} f(x)$  iff  $0 \in \partial f(x)$

Comment:  $f$  is differentiable at  $x \in \text{dom}(f)$ , iff  $\partial f(x)$  is a singleton.

Fix  $d \in E$ .

$$0 \in \partial f(x) \Rightarrow f(x+td) - f(x) \geq t \langle 0, d \rangle \quad \forall t > 0$$

Def<sup>n</sup>  $\Rightarrow \frac{f(x+td) - f(x)}{t} \geq \langle \rho, d \rangle \quad \forall t > 0$

The directional derivative of  $f$  at  $x$  along  $d$  is

$$f'(x; d) = \lim_{t \downarrow 0} \frac{f(x+td) - f(x)}{t}$$

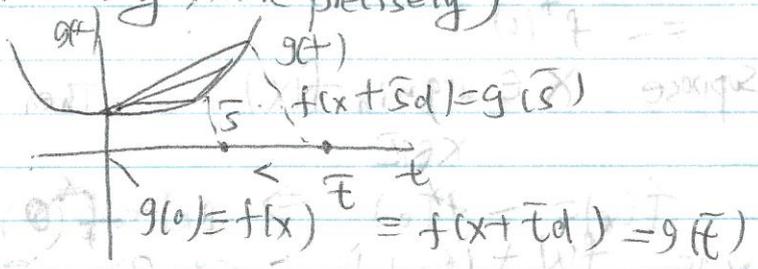
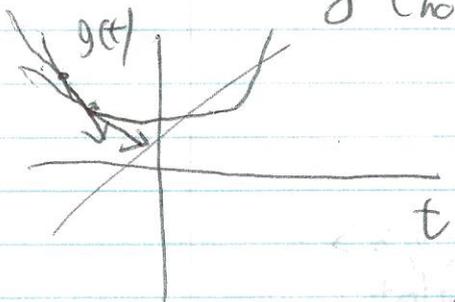
if it exists.  $\leftarrow$  gradient  $\nabla f(x)$

if  $f'(x; d) = \langle \rho, d \rangle$  for some  $\rho$ , i.e. linear in  $d$ .

then  $f$  is "differentiable" (called Gateaux differentiable)

Lemma. If  $f$  is convex,  $x \in \text{dom}(f)$ ,  $\forall d \in \mathbb{E}$

then function  $h(t) := \frac{f(x+td) - f(x)}{t}$  is monotonically increasing (non-decreasing, more precisely)

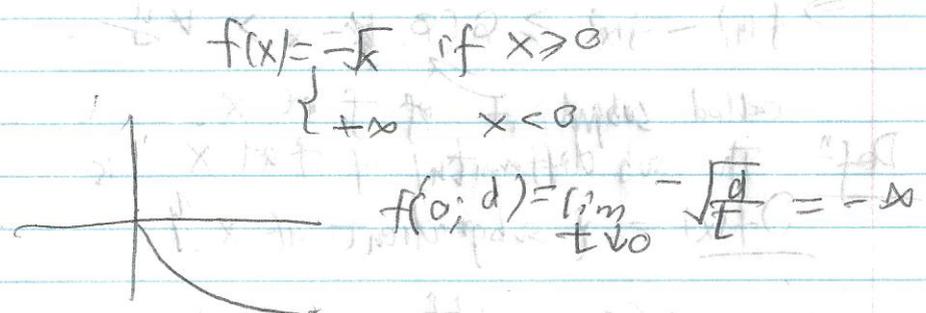
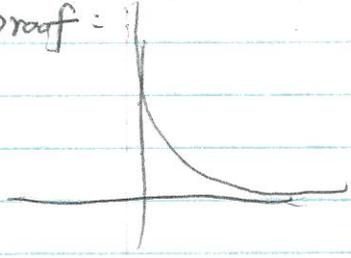


proof: See def<sup>n</sup> of convexity to get slope of the chords are monotonic in  $t$ .

Th<sup>m</sup> If  $f$  is convex, then  $\forall x \in \text{dom}(f)$ ,  $\forall d \in \mathbb{E}$

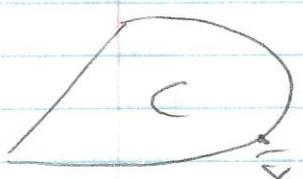
$$f'(x; d) = \lim_{t \downarrow 0} \frac{f(x+td) - f(x)}{t} \text{ exists in } [-\infty, +\infty]$$

Proof:



$$f(x) = \begin{cases} -x & \text{if } x \geq 0 \\ +\infty & \text{if } x < 0 \end{cases}$$

$$f'(0; d) = \lim_{t \downarrow 0} \frac{-\sqrt{d}}{t} = -\infty$$



$\bar{x} \in \text{argmin}_{x \in C} f(x)$  iff optimality condition.

Oct 6th 2007

Aside, Recall

Suppose  $f$  is convex and differentiable,  $(f: \mathbb{E} \rightarrow (-\infty, +\infty])$  proper.  
 Then TEAE:

(i)  $f$  is convex on  $S$  (ii)  $f(x) - f(y) \geq \langle \nabla f(y), x - y \rangle \quad \forall x, y \in S$

(iii)  $\langle \nabla f(x) - \nabla f(y), x - y \rangle \geq 0 \quad \forall x, y \in S$

(non decreasing slope)

(iv)  $\nabla^2 f(x) \succeq 0 \quad \forall x \in S$  if  $f$  is twice differentiable.  
 $\partial f(x), f(x; \cdot) \rightarrow$  extend to nonsmooth functions.

$\phi \in \partial f(x)$  if  $f(y) - f(x) \geq \langle \phi, y - x \rangle$

called subgradient subdifferential



For  $y = x + td$ , then

$\phi \in \partial f(x) \Rightarrow \frac{f(x+td) - f(x)}{t} \geq \langle \phi, d \rangle \quad \forall t > 0, \forall d \in \mathbb{E}$

So:  $f'(x; d) = \lim_{t \downarrow 0} \frac{f(x+td) - f(x)}{t} \geq \langle \phi, d \rangle \quad \forall d \in \mathbb{E}$   
 directional derivative support function

$\Rightarrow \partial f(x) \subset S_{f'(x; \cdot)}$

We also showed

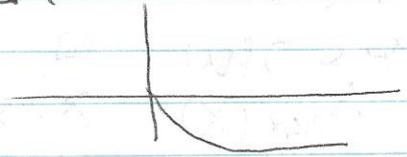
a)  $g(t) = \frac{f(x+td) - f(x)}{t}$  for  $f$  convex and  $x \in \text{dom}(f)$

any  $d \in \mathbb{E}$ , we get  $g$  is nondecr. in  $t$  ( $t \neq 0$ )

b)  $f'(x; d)$  exists in  $[-\infty, +\infty]$ .

Example:

①  $f(x) = \begin{cases} -\sqrt{x} & (x \geq 0) \\ +\infty & (x < 0) \end{cases}$



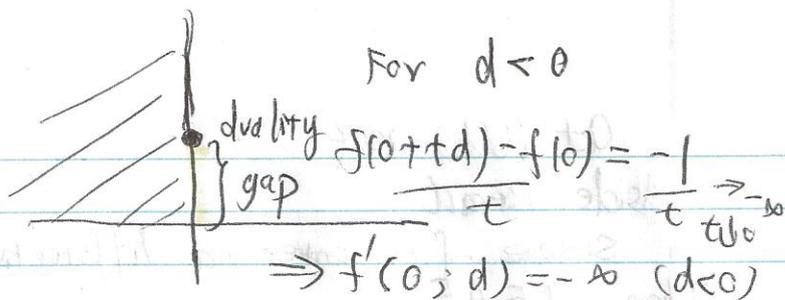
For  $d > 0$  at  $x=0$ :

$\frac{f(0+td) - f(0)}{t} = \frac{-\sqrt{td} - 0}{t} \rightarrow -\infty$  as  $t \downarrow 0$

$\exists \phi \ni \langle \phi, d \rangle \leq f(d) - f(0) \quad \forall d$   
 $\phi d \in \mathbb{R}^+ - \sqrt{d} \quad \phi \leq -\frac{1}{\sqrt{d}}$

ANS: NO  $\phi$  exists.

$$\textcircled{2} f(x) = \begin{cases} 0 & \text{if } x < 0 \\ 1 & \text{if } x = 0 \\ +\infty & \text{if } x > 0 \end{cases}$$



More properties of  $f'(x; \cdot)$  (directional derivative)

We saw:  $\textcircled{1} \varphi \in \partial f(x)$ , iff  $f'(x; d) \geq \langle \varphi, d \rangle \quad \forall d$   
 $x \in \text{dom}(f)$

$\textcircled{2} f'(x; \cdot)$  is positively homogenous, if it exists.

proof:  $\forall \lambda > 0 \quad f'(x; \lambda d) = \lim_{t \downarrow 0} \frac{\lambda f(x + \lambda t d) - f(x)}{\lambda t} = \lambda f'(x; d)$

Since  $t > 0$  iff  $\lambda t > 0$

$\textcircled{3}$  If  $f$  is convex, then  $f'(x; \cdot)$  is convex (hence sublinear)

proof: exercise (see epigraph)

$\textcircled{4}$  If  $f$  is convex,  $x \in \text{dom}(f)$ , then  $N_S(\bar{x}) = (S - \bar{x})^\circ$  Defn of normal cone.

$$\partial f(x) = S_{f'(x; \cdot)} (= \{ \varphi : \langle \varphi, d \rangle \leq f'(x; d) \quad \forall d \}) \text{ important}$$

proof: " $\subseteq$ " done already.

" $\supseteq$ ": since  $f$  is convex, then  $g(t) = \frac{f(x+td) - f(x)}{t}$  is

nondecr.  $\forall t > 0 \Rightarrow f'(x; d) \leq \frac{f(x+td) - f(x)}{t} \quad (t=1 \text{ i.e.})$

$\Rightarrow$  For any  $\varphi \in S_{f'(x; \cdot)}$  we get

$$\langle \varphi, d \rangle \leq f'(x; d) \quad \forall d \in \mathbb{E}$$

$$\Rightarrow \forall y \in \mathbb{E} \quad \langle \varphi, y - x \rangle \leq f'(x; y - x) \leq f(y) - f(x)$$

$$\Rightarrow \varphi \in \partial f(x) \quad \square$$

example:  $\lambda_{\max}(X) \quad X \in S^n$  largest eigenvalue.

Aside:  $Xv = \lambda v \quad v^T v = 1$   $\lambda$  singleton implies differentiable.  
 $\uparrow$   
 eigenpair

$$\frac{dX}{dt} = \dot{X} \quad \dot{X} v + X \dot{v} = \lambda v + \dot{\lambda} v$$

$$2v^T \dot{v} = 0 \quad v^T \dot{X} v + v^T X \dot{v} = \lambda v^T v + \dot{\lambda} v^T v$$

$$v^T \dot{X} v = \dot{\lambda} \quad \frac{d\lambda}{dt} = v^T \frac{\partial X}{\partial t} v = \frac{(Xv)^T \dot{v}}{v^T v} = \frac{\lambda v^T \dot{v}}{1} = \dot{\lambda}$$

eg.  $X = \begin{bmatrix} t & x_{12} & x_{13} \\ x_{21} & x_{22} & \dots \\ \dots & \dots & \dots \end{bmatrix} \Rightarrow \dot{X} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$

$\lambda = v^T \dot{X} v = v_1^2$   
found at current  $\bar{x}$

$X = \begin{bmatrix} t & t^2 & 3 \\ t^2 & 4s & 5 \\ 3+5 & 7 & 1 \\ 5 & 6 & 2 \end{bmatrix}$   
 $\uparrow$   
(s,t)

$\dot{x}_t = \frac{\partial X}{\partial t} = \begin{bmatrix} 1 & 2t & 0 & 0 \\ 2t & 4 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$  at  $t=1, s=2$

Rayleigh principle (Quotient)

Aside 2: Lagrangian

$\lambda_{\max}(X) = \max_{\|v\|=1} v^T X v$   $\nabla_v (v^T X v + \lambda (1 - v^T v)) = 0$

Variational formula

$2Xv - 2\lambda v = 0$

$\Rightarrow Xv = \lambda v$

Let  $V = \{v \in \mathbb{R}^n : \|Xv\| = \lambda_{\max}(X) \|v\|\}$

$v^T X v = \lambda v^T v = \lambda$

called eigen space for  $\lambda_{\max}$

So  $v^T v = 1 \implies v^T X v = \lambda_{\max}(X)$  iff  $(v \in V \implies v^T v = 1)$

so  $\lambda_{\max}(X+tD) - \lambda_{\max}(X)$  (direction is  $D \in S^n$ )

$= \max_{\|v\|=1} v^T (X+tD) v - \lambda_{\max}(X)$

$\geq v^T (X+tD) v - v^T X v$   $v \in V, \|v\|=1$   
 $= v^T t D v$   $\underbrace{v^T X v}_{= \lambda_{\max}(X)}$

so  $f'(x; D) \geq \max_{\|v\|=1} v^T D v$   $\uparrow$   $M y = y^T M^T D M y$

$v = \{ M y = y \in \mathbb{R}^k \}$  wlog.  $M$  orthon. cols  $k = \dim V$

show  $\leq$  as well

Exercise, similar proof

so  $f'(x; D) = \max_{\|v\|=1} v^T D v = \max_{\|v\|=1} \text{trace } D v v^T$

$\max_{\|v\|=1} \text{trace } D v v^T$

$\leq \langle D, W \rangle$   
 $W = v v^T$

$= \sigma_S(D)$  where  $S = \{W = v v^T : \|v\|=1, v \in V\}$

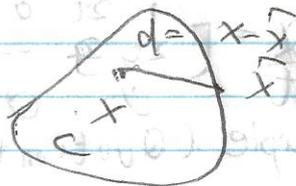
$$\|V\| = 1 \iff v^T v = 1 \iff \text{tr } VV^T = 1$$

and we can use  $\text{conv } S \subseteq S^1$  (compact convex set)  
 so we get

$$f'(x; \cdot) = \nabla_T; \quad T = S_{f'(x; \cdot)}$$

$$df(x) = S_{f'(x; \cdot)} = T$$

$$\text{So } \boxed{f'(x; \cdot) \text{ is } = df(x)}$$



Defn Oct 18th 2007

$C$  convex set,  $\bar{x}$  local min of  $f$  on  $C$  if  $\exists \epsilon > 0$   $\forall x \in C, \|x - \bar{x}\| < \epsilon \implies f(\bar{x}) \leq f(x)$

if  $\bar{x}$  local min of  $f$  on  $C$ , then

$$f(\bar{x} + t(x - \bar{x})) \geq f(\bar{x}) \quad \forall x \in C, \text{ for small } t > 0$$

$$\implies f'(\bar{x}; d) \geq 0 \quad \forall d = x - \bar{x}, x \in C$$

in diff<sup>ble</sup> case  $f'(\bar{x}; d) = \langle \nabla f(\bar{x}), d \rangle$

$$\text{So } \implies \nabla f(\bar{x}) \in (C - \bar{x})^\circ$$

$$\text{(or } -\nabla f(\bar{x}) \in N_C(\bar{x}))$$

i.e. a necessary condition (geometric nec. condition)

Sufficient condition?

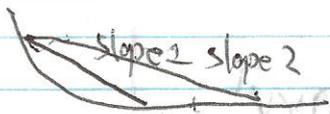
Lemma: Suppose  $f$  is convex on the convex set  $C$ ,  $x, \bar{x} \in C$  and  $f(x) < f(\bar{x})$

$$\text{Then } f'(\bar{x}; d = x - \bar{x}) < 0$$

proof: Let  $0 \leq t \leq 1$  and

$$y = \bar{x} + t(x - \bar{x})$$

$$f(y) = f(\bar{x} + t(x - \bar{x})) = f(t x + (1-t)\bar{x}) \leq t f(x) + (1-t)f(\bar{x}) < f(\bar{x}) \quad \square$$



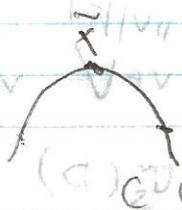
$$\text{slope } 2 \geq \text{slope } 1$$

exercise - complete proof.

Therefore,

$$f'(\bar{x}, x - \bar{x}) \geq 0 \quad \forall x \in C$$

$\implies \bar{x}$  global min of  $f$  (convex) on  $C$



So a sufficient condition if convexity holds.

In a diffble case,  $\bar{x} \in \text{argmin}_x f(x)$  iff  $\nabla f(\bar{x}) \in (C - \bar{x})^\circ$   
 (f convex, C convex)  $x \in C$   
 iff  $-\nabla f(\bar{x}) \in N_C(\bar{x})$  } ⊗

what about non differentiable?

Recall  $\nabla f(\bar{x})$  exists  $\iff$  Gateau differentiable.

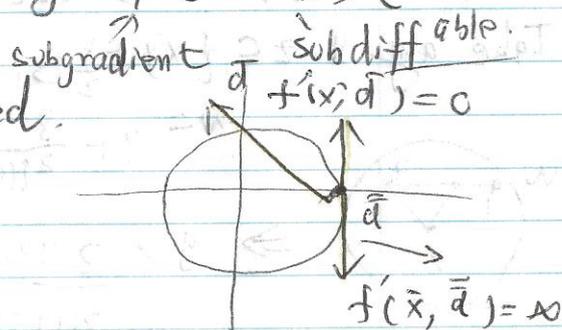
iff  $\partial f(\bar{x}) = \{ \nabla f(\bar{x}) \}$  (Singleton)

can be replaced  $\nabla f(\bar{x})$  in ⊗ by any  $\phi \in \partial f(\bar{x})$ ?

Note  $f'(x; \cdot)$  may not be closed.

even if f is closed, convex.

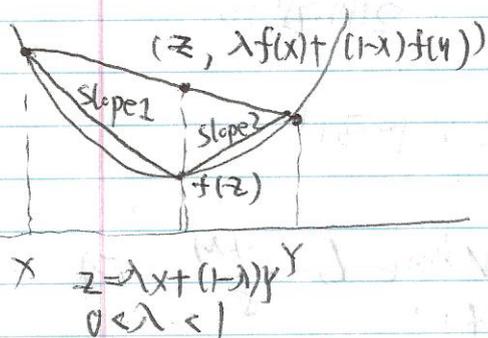
Ex.  $f = \delta_B$  in  $\mathbb{R}^n$ ,  $n \geq 2$   
 $x = e_1 = (1, 0, \dots, 0)^T$



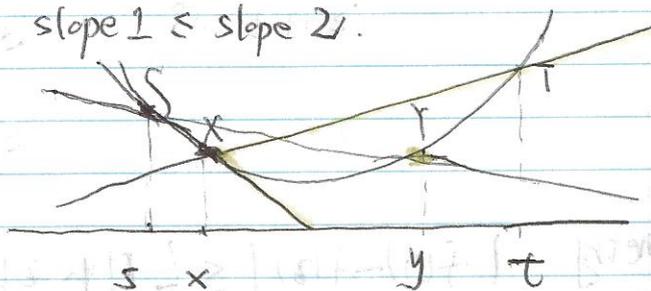
So  $f'(x; d) = \delta_{\{d: d_1 < 0\}}$  is not closed.

{Recall closed  $\iff$  l.s.c.}  $\liminf_{x^k \rightarrow \bar{x}} f(x^k) \geq f(\bar{x})$

Note, f convex  $f: \mathbb{R} \rightarrow \mathbb{R}$



slope 1  $\leq$  slope 2.



$S = (s, f(s))$ ,  $\begin{cases} x \text{ lies below } x \text{ or on line } SX \\ Y \text{ lies above or on line } SX \\ Y \text{ lies below or on line } XT \end{cases}$

Now let  $y \rightarrow x \implies f \rightarrow x$  i.e.  $f(y) \rightarrow f(x)$



So, prop.  $f$  convex on  $S$ , an open convex set, Then  $f$  is continuous on  $S$ .

Def<sup>n</sup> A function  $f$  is locally Lipschitz continuous at point  $x \in E$  if  $\exists \delta > 0, \exists L$  s.t.  $|f(x) - f(y)| \leq L \|x - y\| \forall y \in \{x\} + \delta B$  (for some  $L$ )

Locally Lipschitz  $\Rightarrow$  cont.  $\Rightarrow$  L.S.C.

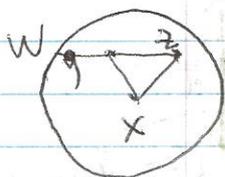
Th<sup>m</sup> If  $f$  is convex and  $x \in \text{int}(\text{dom } f)$ , then  $f$  is locally Lip. cont. at  $x$

proof: Step 1:  $f$  is locally bounded at  $x$  since we know  $f$  is locally continuous, ie.  $\exists \delta > 0, \exists M$  s.t.  $f(y) \leq M, \forall y \in \{x\} + \delta B$ .

Step 2: Show it is loc. Lip.

Take any  $y, z \in \{x\} + \frac{\delta}{2} B, y \neq z$

$$w = y + \frac{\delta}{2\|y-z\|}(y-z)$$



$$\Rightarrow y = \frac{2\|y-z\|}{2\|y-z\| + \delta} w + \frac{\delta}{2\|y-z\| + \delta} z$$

$$\Rightarrow f(y) \leq \frac{2\|y-z\|}{2\|y-z\| + \delta} f(w) + \frac{\delta}{2\|y-z\| + \delta} f(z)$$

$M$  (inside the  $\delta$  ball)

But  $\|w-x\| \leq \|w-y\| + \|y-x\| \leq \delta$

$$\Rightarrow f(y) - f(z) \leq \frac{2\|y-z\|}{2\|y-z\| + \delta} (M - f(z)) \leq \frac{4M}{2\|y-z\| + \delta} \|y-z\|$$

$$\leq \frac{4M}{\delta} \|y-z\|$$

By symmetry,  $|f(y) - f(z)| \leq L \|y-z\|$ , where  $L = \frac{4M}{\delta}$   $\square$

Th<sup>m</sup> If  $f$  is convex and  $x \in \text{int}(\text{dom } f)$

Then  $f'(x; \cdot)$  is closed.

proof: outline;  $f'(x; \cdot)$  is sublinear.

The only thing need to do is to show  $\text{dom}(f'(x; \cdot)) = E$

Summary: (i)  $f'(x; \cdot)$  is finite everywhere if  $x \in \text{int}(\text{dom}(f))$   
 (ii)  $\partial f(x) = \{\varphi : \langle \varphi, d \rangle \leq f'(x; d) \quad \forall d\}$   
 $(\equiv \bigcap_{d \in S} f'(x; d))$   
 So for  $x \in \text{int}(\text{dom}(f))$ ,  $f'(x; \cdot) = (f'(x; \cdot))^*$   
 $= \nabla \partial f(x) = \sup \{ \langle \varphi, \cdot \rangle \}$   
 $\varphi \in \partial f(x)$   
 closed set, convex.  
 From definition.

And  $\exists \delta > 0, L > 0 \rightarrow f(y) - f(x) \leq L \|y - x\| \quad \forall y \in \{x\} + \delta B$   
 $\Rightarrow f'(x; d) \leq L \quad \forall d$   
 $\Rightarrow \partial f(x) \subseteq LB$

so  $\partial f(x)$  is bounded, convex, closed set (compact)

$f'(x; d) = \max_{\varphi \in \partial f(x)} \langle \varphi, d \rangle \quad \forall d$

Summary

- ①  $\inf_{x \in E} f(x) = -f^*(0) \xrightarrow{\text{Definition}} \sup_{x \in E} \{ \langle 0, x \rangle - f(x) \} = \inf_{x \in E} f(x)$
- ②  $\bar{x} \in \arg \min_{x \in E} f(x) \Leftrightarrow 0 \in \partial f(\bar{x})$   
 $\rightarrow f(\bar{x}) \leq f(x) \quad \forall x \in E$
- ③  $\partial f(x) = \bigcap_{d \in S} f'(x; d)$   
 $\langle 0, x - \bar{x} \rangle \leq f(x) - f(\bar{x}) \quad \forall x \in E$
- ④  $\partial f(x)$  is closed, convex, and it is bounded if  $x \in \text{int}(\text{dom}(f))$  (otherwise,  $\partial f(x)$  is unbounded or empty)
- ⑤  $\nabla_n$  diff  $\rightarrow \partial f(x) = \{ \langle \nabla f(x), d \rangle \}$   
 $\partial f(x)$  is singleton.
- ④  $x \in \text{int}(\text{dom}(f))$  then  
 $f'(\bar{x}; d) = \max_{\varphi \in \partial f(\bar{x})} \langle \varphi, d \rangle \quad \forall d$
- ⑤  $\partial f(\bar{x})$  is closed and bounded when  $x \in \text{int}(\text{dom}(f))$   
 otherwise it is unbounded or empty
- ⑥  $f$  is loc. lip. at  $x \in \text{int}(\text{dom}(f))$

⑦  $\partial f(x)$  is a singleton iff  $f$  is Gateaux Differentiable at  $x$   
 i.e.  $f'(x; d) = \langle a, d \rangle \forall d$   
 i.e.  $\partial f(x) = \{ \nabla f(x) \}$

Set Constrained Minimization  
 $\emptyset \neq S \subseteq \mathbb{E}$ , convex

$f$  is proper and convex

$$\inf_{x \in S} \{ f(x) \} = \inf_{x \in \mathbb{E}} \{ f(x) + \delta_S(x) \}$$

$\bar{x} \in \text{argmin}_{x \in S} f(x)$  iff  $0 \in \partial f(\bar{x}) + \partial \delta_S(\bar{x})$

Question: iff?  $0 \in \partial f(\bar{x}) + \partial \delta_S(\bar{x})$ ?

More general;  $f, g$  convex, proper. (Not need have actually)

and if  $\varphi \in \partial f(x) + \partial g(x)$

$$\varphi = \varphi_1 + \varphi_2$$

$$\text{So } \langle \varphi_1, y-x \rangle \leq f(y) - f(x) \quad \forall y$$

$$\langle \varphi_2, y-x \rangle \leq g(y) - g(x) \quad \forall y$$

$$\Rightarrow \langle \varphi_1 + \varphi_2, y-x \rangle \leq (f+g)(y-x) \quad \forall y$$

$$\Rightarrow \partial f(x) + \partial g(x) \subset \partial (f+g)(x)$$

What about  $\supset$ ?

First Suppose  $0 \in \partial (f+g)(x)$

$$\Rightarrow \langle 0, y-x \rangle \leq (f+g)(y) - (f+g)(x) \quad \forall y \in \mathbb{E}$$

$$\Rightarrow g(y) - g(x) \leq f(y) - f(x) \quad \forall y \in \mathbb{E}$$

$$\textcircled{xxx} \quad -(g(x+d) - g(x)) \leq f(x+d) - f(x) \quad \forall d \in \mathbb{E}$$

Theorem (a Sandwich Theorem)

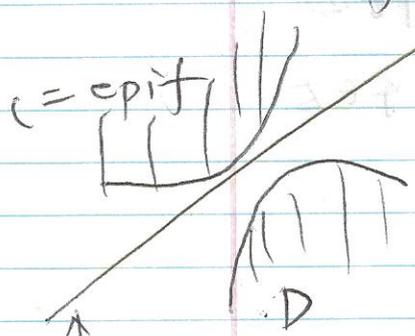
$f, g$  proper, convex,

$$\text{int}(\text{dom}(f) \cap \text{dom}(g)) \neq \emptyset$$

$$f \geq -g$$

Then ~~the~~  $\exists \varphi \in E, P \in \mathbb{R}, \rightarrow$

$$f(x) \geq \underbrace{\langle \varphi, x \rangle + P}_{\text{affine}} \geq -g(x) \quad \forall x \in E$$



proof: Let  $C = \{(x, r) \mid r \geq f(x)\}$  epi f.

$$D = \{(x, r) \mid r \leq -g(x)\}$$

$C, D$  are convex.

$\Rightarrow C - D$  is convex.

hyperplane Supp. at  $(0, 0)$

Note  $(0, 0) \notin \text{int}(C - D)$

Suppose not, i.e.  $(0, 0) \in \text{int}(C - D)$

Then  $(0, \delta) \in (C - D)$  for  $\delta < \delta < 0$

$\parallel$   
 $r - s$  where  $r \geq f(x), s \leq -g(x)$  but  $r \geq s$  Contrad.

So  $(0, 0) \notin \text{int}(C - D) \neq \emptyset$

$\Rightarrow \exists (a, b) \neq 0, \rightarrow$

$$0 = \langle (0, 0), (a, b) \rangle \geq \langle (z, t), (a, b) \rangle$$

i.e.  $\langle a, x \rangle + br \geq \langle a, y \rangle + bs$   $\forall (z, t) \in C - D$

First note that  $b < 0$

if  $b > 0$  we can choose  $s \uparrow +\infty$  (contrad.)

so  $b < 0$

Can  $b = 0$ ?

Let  $z \in \text{int}(\text{dom}(f)) \cap \text{dom}(g)$

$\Rightarrow \exists \delta > 0 \rightarrow \{z\} + \delta B \subseteq \text{dom}(f)$

$$\langle a, z \rangle + 0 \geq \langle a, z + \frac{\delta a}{\|a\|^2} \rangle + 0$$

$$\Rightarrow 0 \geq \delta \|a\| = \delta \quad \text{contrad.}$$

so  $b < 0$ , let  $\phi = -a/b$

$$\Rightarrow \langle \phi, x \rangle + g(x) \geq \langle \phi, y \rangle - f(y) \quad \forall x, y \in \mathbb{R}$$

$$\Rightarrow \langle \phi, y \rangle + f(y) \geq \langle -\phi, x \rangle - g(x) \quad \forall x, y \in \mathbb{R}$$

$$\text{Let } p = \inf_{y \in \mathbb{R}} \{ -\langle \phi, y \rangle + f(y) \}$$

$$\Rightarrow f(y) \geq \langle \phi, y \rangle + p$$

$$\text{at the same time } -g(x) \leq \langle \phi, x \rangle + p \quad \square$$

ref. Halpern

student

in AMM (late 80's)

Pourciau

→

proofs of Lagr. mult. th<sup>m</sup> use hyperpl. separ.

Recall ~~(x)~~

$$-(g(x+d) - g(x)) \leq f(x+d) - f(x) \quad \forall d$$

$$\text{if } 0 \in \partial(f+g)(x)$$

$$\text{Let } \bar{f}(d) = f(x+d) - f(x)$$

$$\bar{g}(d) = g(x+d) - g(x)$$

$$\text{then } -\bar{g} \leq \bar{f}$$

$$\text{with } \text{int}(\text{dom } \bar{f}) \cap \text{dom } \bar{g} = \text{int}(\text{dom } f) \cap \text{dom } g - \{x\} \neq \emptyset$$

$$\Rightarrow \exists (\rho, \rho) \ni \bar{f}(d) \geq \langle \rho, d \rangle + \rho \geq -\bar{g}(d)$$

by sandwich theorem.

$$\uparrow_{=0} \Rightarrow \rho = 0$$

need to show

$$\partial f(x) + \partial g(x) \supset \partial(f+g)(x)$$

$$\Rightarrow \phi \in \partial f(x), \quad -\phi \in \partial g(x).$$

$$\Rightarrow 0 \in \partial f(x) + \partial g(x)$$

$$\text{so } 0 \in \partial(f+g)(x) \Rightarrow 0 \in \partial f(x) + \partial g(x)$$

$$\text{Now } \phi \in \partial(f+g)(x) \Rightarrow 0 \in \partial(f+g)(x) - \{\phi\}$$

$\Rightarrow \phi \in \partial(f_\phi + g)(x)$  where  $f_\phi(x) = f(x) - \langle \phi, x \rangle$

Big Theorem  $\Rightarrow \phi \in \partial f(x) + \partial g(x)$  i.e. if  $\text{int}(\text{dom}(f)) \cap \text{dom}(g) \neq \emptyset$

$\partial(f+g)(x) = \partial f(x) + \partial g(x)$

$x \in \arg \min_{x \in S} f(x)$  iff  $\phi \in \partial f(x) + \partial \delta_S(x)$  (Need  $\text{int}(\text{dom}(f)) \cap S \neq \emptyset$ )  
 iff  $(-\partial f(x)) \cap \partial \delta_S(x) \neq \emptyset$

Note  $\phi \in \partial \delta_S(x)$

$\Leftrightarrow \langle \phi, y-x \rangle \leq \delta_S(y) - \delta_S(x) \quad \forall y$

( $x \in S$ !) iff  $\langle \phi, y-x \rangle \leq 0$  polar

$\Leftrightarrow \phi \in N_S(x) = (S-x)^\circ = -(S-x)^\circ$

$N_S(x)$

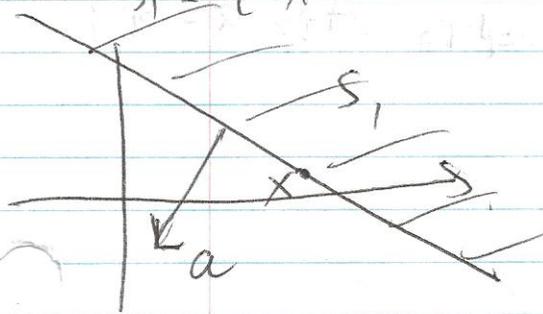
①  $x \in \text{int}(S) \Rightarrow N_S(x) = \{0\}$

②  $S = S_1 \cap S_2$  then  $N_S(x) = \partial \delta_S(x) = \partial(\delta_{S_1} + \delta_{S_2})(x)$   
 $\begin{matrix} \nearrow \\ \text{convex} \\ \searrow \end{matrix}$   
 $\text{int}(S_1) \cap \text{int}(S_2) \neq \emptyset$   
 $= \partial \delta_{S_1}(x) + \partial \delta_{S_2}(x)$   
 $= N_{S_1}(x) + N_{S_2}(x)$

special cases

$S_1 = \{x \mid \langle a, x \rangle \leq b\}$  half space  
 $x$  on the boundary.

$N_{S_1}(x) =$  cone generated by  $\{a\}$ ,  
 closed convex cone.



$S_2 = \{x \mid \max_{1 \leq i \leq m} A_i x \leq b_i\}$

$A = \begin{bmatrix} a_1^T \\ a_2^T \\ \vdots \\ a_m^T \end{bmatrix}$   $x \in S_2$   
 $N_{S_2}(x)$  We use active set of constraints. Hilroy

will use  $S = S^{++}$  iff  $S$  c.c.c.?

Oct 25th, 2007 Convex analysis and Optimization  
(Half of the term has passed, wow)

Motivation

characterize  $\bar{x} \in \arg \min_{x \in S} f(x)$  iff  $\bar{x} \in \arg \min_{x \in \mathbb{R}^n} f(x) + \delta_S(x)$

iff  $0 \in \partial(f + \delta_S)(\bar{x})$  iff  $0 \in \partial f(\bar{x}) + \partial \delta_S(\bar{x})$   $N_S(\bar{x})$

Assume  $\text{int dom}(f) \cap S \neq \emptyset$  (regularity assumption)

iff  $(-\partial f(\bar{x})) \cap N_S(\bar{x}) \neq \emptyset$

(In diff'ble case,  $-\nabla f(\bar{x}) \in N_S(\bar{x})$ )

e.g.  $S = \{x : \langle a, x \rangle \leq b\}$  half space

$N_S(\bar{x}) = \{\rho : \langle \rho, y - \bar{x} \rangle \leq 0 \forall y \in S\}$

$y - \bar{x} = d \leftarrow$  feasible direction  
 $N_S(\bar{x}) = \text{cone}\{a\}$

$$e(x) = \frac{1}{2} x^T Q x + c^T x = \alpha \quad Q \succ 0$$

$$L_\alpha = \{x : e(x) = \alpha\}$$

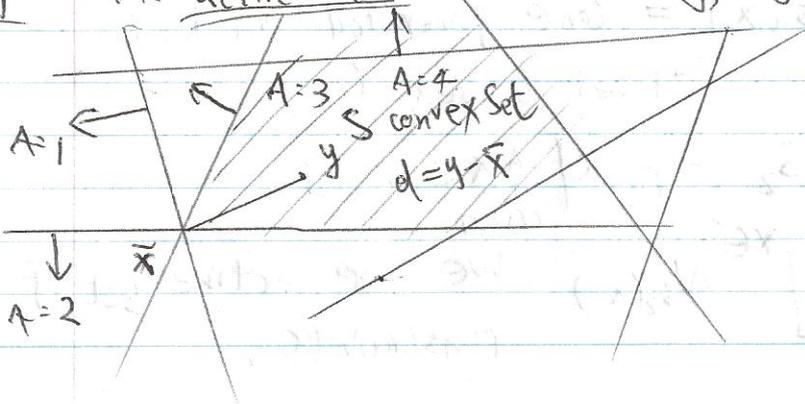
$$\frac{1}{2} \|Hx - \beta\|^2$$

example 2;  $S$  polyhedral set,  $S = \{x \mid Ax \leq b\} \subset \mathbb{R}^n$

$\bar{x}$  feasible, i.e.  $A\bar{x} \leq b$   $A_{i:} \bar{x} \leq b_i, i=1, 2, \dots, m$

$N_S(\bar{x}) = ?$

Def<sup>n</sup> The active set at  $\bar{x}$  is  $\mathcal{A} = \mathcal{A}(\bar{x}) := \{i : A_{i:} \bar{x} = b_i\}$



$$N_S(\bar{x}) = \{ \phi : \langle \phi, y - \bar{x} \rangle \leq 0 \quad \forall y \in S \}$$

Def:  $S$  convex set,  $\bar{x} \in S$ .  $d$  is a feasible direction at  $\bar{x}$  if  $\bar{x} + \alpha d \in S \quad \forall 0 < \alpha < \bar{\alpha}$  for some  $\bar{\alpha} > 0$

clear:  $d$  is a feasible direction iff:  $\alpha d = y - \bar{x}$  for some  $\alpha > 0$  and some  $y \in S$  at  $\bar{x}$

$$N_S(\bar{x}) = \{ \phi : \langle \phi, d \rangle \leq 0 \quad \forall \text{ feasible dir. } d \text{ at } \bar{x} \}$$

So normal cone  $N_S(\bar{x})$  depends only on  $\mathcal{L}(\bar{x})$

Farkas Lemma. (Several lines of Proof if we have the knowledge)

$$B = [b_1, \dots, b_n]_{m \times n} \quad b_i \in \mathbb{R}^m$$

(I).  $B\lambda = b$ , for some  $\lambda \geq 0$  iff.

$$\text{II} \quad B^T d \geq 0 \Rightarrow b^T d \geq 0$$

or  $(\leq 0) \quad (\leq 0)$

proof:  $K = \text{cone} \{ b_i \}_{i=1}^n$  is a finitely generated cone so a ccc.

Then we know  $K$  ccc. iff  $K = K^{++}$ .

so  $b = B\lambda, \lambda \geq 0$  iff  $b \in K$  iff  $b \in (K^+)^+$

$$Bd \geq 0 \Rightarrow b^T d \geq 0$$

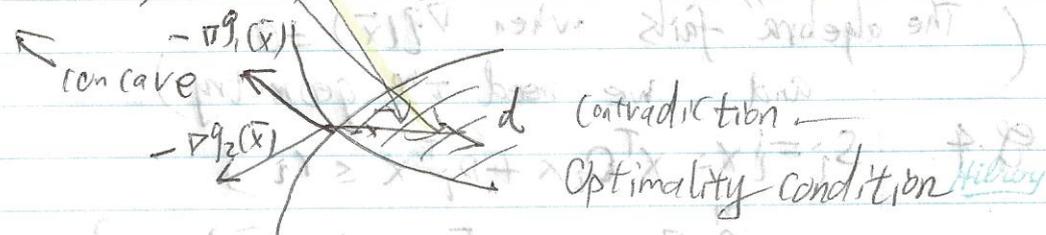
iff Farkas' Lemma is often called a theorem of the alternative.

II holds iff  $B^T d \geq 0$  and III holds iff  $B^T d < 0$  is inconsistent. Exactly one of I, III holds

II holds iff  $0 = \min_{B^T d \geq 0} b^T d$  Taylor's theory estimation (1st order)

eg. if  $b = \nabla f(\bar{x}), b_i = \nabla g_i(\bar{x}) \quad i \in J(\bar{x})$

$\min f(x)$  st.  $g_k(x) \geq 0 \quad k \in I, \dots, m$   
 convex



So  $Bx=b, \lambda \geq 0$  is a test for opt.

So  $\nabla f(x) + \sum \lambda_i \nabla g_i(x) = 0$  is a test for opt

i.e. Karush-Kuhn-Tucker conditions

Aside 1.

$$g(x) = \begin{cases} x^2 & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases}$$

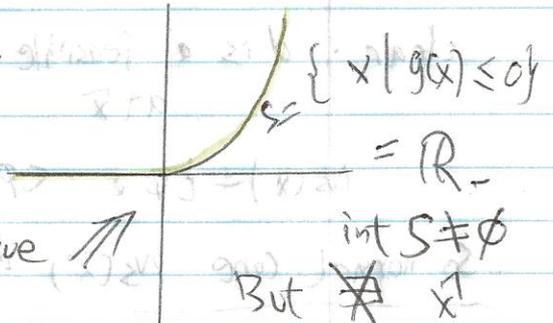
Ex 3.  $S = \{x : \frac{1}{2} x^T Q x + p^T x \leq r\}$   $Q \succeq 0$

Let  $q(x) = \frac{1}{2} x^T Q x + p^T x$

if  $q(\bar{x}) < r$ , then  $\bar{x} \in \text{int}(S)$

and  $N_S(\bar{x}) = \{0\}$

not always true



Aside 2 (backside)

$$S = \{x : Ax \leq b\}$$

$$N_S(\bar{x}) = \text{cone} \left\{ \begin{matrix} (A_{i_1}^T) \\ \vdots \\ (A_{i_m}^T) \end{matrix} \right\}$$

$A_{i_j}^T \bar{x} = b_{i_j}$   
 $i \in \{1, 2, \dots, m\}$

cone generated by active normals.

alternate view:  $S_i = \{x : A_{i_j} x \leq b_{i_j}\}$

$$S = \bigcap_{i=1}^m S_i$$

Then  $N_S(\bar{x}) = (S - \bar{x})^\circ = \sum_{i \in \mathcal{I}(\bar{x})} (S_i - \bar{x})^\circ$

active set  
= cone  $\{\nabla q(\bar{x})\}$

For quadratic

$$= \text{cone} \{Q\bar{x} + p\}$$

$$\boxed{\text{if } Q\bar{x} + p \neq 0}$$

eg.  $q(x) = x_2^2$   $x \in \mathbb{R}^2$   $\bar{x} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$   $q(\bar{x}) = 0$

$$\nabla q(x) = 2x_2 = \begin{pmatrix} 0 \\ 2x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$S = \{x : x_2^2 \leq 0\} = \left\{ \begin{pmatrix} x_1 \\ 0 \end{pmatrix} \mid \forall x_1 \in \mathbb{R} \right\}$$

$$N_S(\bar{x}) = \left\{ \begin{pmatrix} 0 \\ x_2 \end{pmatrix} \mid x_2 \in \mathbb{R} \right\}$$

(The "algebra" fails when  $\nabla q(\bar{x}) = 0$  and we need the geometry)

eg. 4.  $S_i = \{x : x^T Q_i x + p_i^T x \leq r_i\}$

$$S = \bigcap_{i=1}^m S_i, \Rightarrow \bar{x} \quad N_S(\bar{x}) = ?$$

Assume  $\nabla q_i(\bar{x}) \neq 0 \forall i \in \mathcal{I}(\bar{x})$  and  $\exists \hat{x} \in \text{int} S_i \forall i$  (so  $\hat{x} \in \text{int} S$ )

~~$\{\nabla q_i(\bar{x})\}$~~   $\in \mathcal{I}(\bar{x})$   
 So  $N_S(\bar{x}) = \sum_{i=1}^m N_{S_i}(\bar{x}) = \dots$

Oct 30, 2007 Tuesday

(Using  $N_S(\bar{x})$  to charact. opt.)

Suppose  $S = \{x \in E : g(x) \leq 0\}$   $g$  convex, proper.  
 $\bar{x} \in \text{int dom}(g)$

First, suppose  $\bar{x} \in \text{int} S$   
 (eg. if  $g(\bar{x}) < 0$ )

$\Rightarrow N_S(\bar{x}) = (S - \bar{x})^\circ = \{0\}$

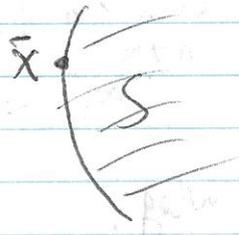
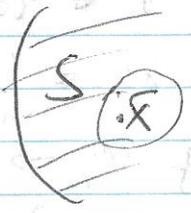
Next, suppose  $g(\bar{x}) = 0$

then

$\phi \in \partial g(\bar{x}) \Rightarrow \langle \phi, y - \bar{x} \rangle \leq g(y) - g(\bar{x})$   
 $\forall y$

$\Rightarrow \langle \phi, y - \bar{x} \rangle \leq 0 \quad \forall y \in S$

$\Rightarrow \phi \in N_S(\bar{x})$  by def<sup>n</sup>



Conclusion: cone generated by  $\text{cone } \partial g(\bar{x}) \subset N_S(\bar{x})$

eg.  $g(x) = x^2, x \in \mathbb{R}$

$\bar{x} = 0$

so  $S = \{0\}$  and  $N_S(\bar{x}) = \mathbb{R}$

But  $\partial g(\bar{x}) = \{xg(\bar{x})\} = \{2\bar{x}\} = \{0\}$

so  $\text{cone } \{\partial g(\bar{x})\} = \{0\} \subsetneq N_S(\bar{x})$

Algebraic (analytic) convex, compact description vs geometric description of

But suppose  $0 \notin \partial g(\bar{x})$

Then  $\exists y \neq \bar{x} : g(y) < g(\bar{x}) = 0$

$\{i.e. \exists y \neq \bar{x} : g(y) < 0\}$

Hilroy

$y$  is called a Slater point

Slater's constraint qualification

Lemma 1.  $g, \bar{x}$  as above, suppose  $0 \notin \partial g(\bar{x})$ , then  $\text{cone } \partial g(\bar{x})$  is closed.

eg.  cones = open half space union  $\{0\}$   
Not closed.

proof: Since  $\partial g(\bar{x})$  is convex, we get  
 $\text{cone}(\partial g(\bar{x})) = \mathbb{R}_+(\partial g(\bar{x})) = \bigcup_{\lambda \geq 0} (\lambda \partial g(\bar{x}))$

Let  $\gamma \in \text{cone}(\partial g(\bar{x}))$  i.e.

$$\gamma = \lim_{n \rightarrow \infty} \lambda_n \phi_n, \lambda_n \geq 0, \phi_n \in \partial g(\bar{x})$$

wlog.  $\phi_n \rightarrow \bar{\phi} \in \partial g(\bar{x})$

$$\Rightarrow \lim_{n \rightarrow \infty} \lambda_n \|\phi_n\| = \|\gamma\|$$

$$\Rightarrow \lambda_n \leq \frac{2.2 \|\gamma\|}{\|\bar{\phi}\|} \text{ for sufficiently large } n.$$

So wlog.  $\lambda_n \rightarrow \bar{\lambda} \geq 0$

so  $\bar{\gamma} = \bar{\lambda} \bar{\phi} \in \text{cone}(\partial g(\bar{x})) \Rightarrow$  closed  $\square$

To show  $N_S(\bar{x}) \subset \text{cone } \partial g(\bar{x})$

Let  $\bar{\phi} \in N_S(\bar{x})$ , but assume  $\bar{\phi} \notin \text{cone } \partial g(\bar{x})$   
 a. c.c.c.

Now derive a contradict.

Now apply hyperplane separation:

$\exists \bar{d} \neq 0$  and  $\beta \in \mathbb{R}, \exists$

$$\langle \bar{d}, \bar{\phi} \rangle > \beta \geq \langle \bar{d}, \phi \rangle \quad \forall \phi \in \text{cone } \partial g(\bar{x})$$

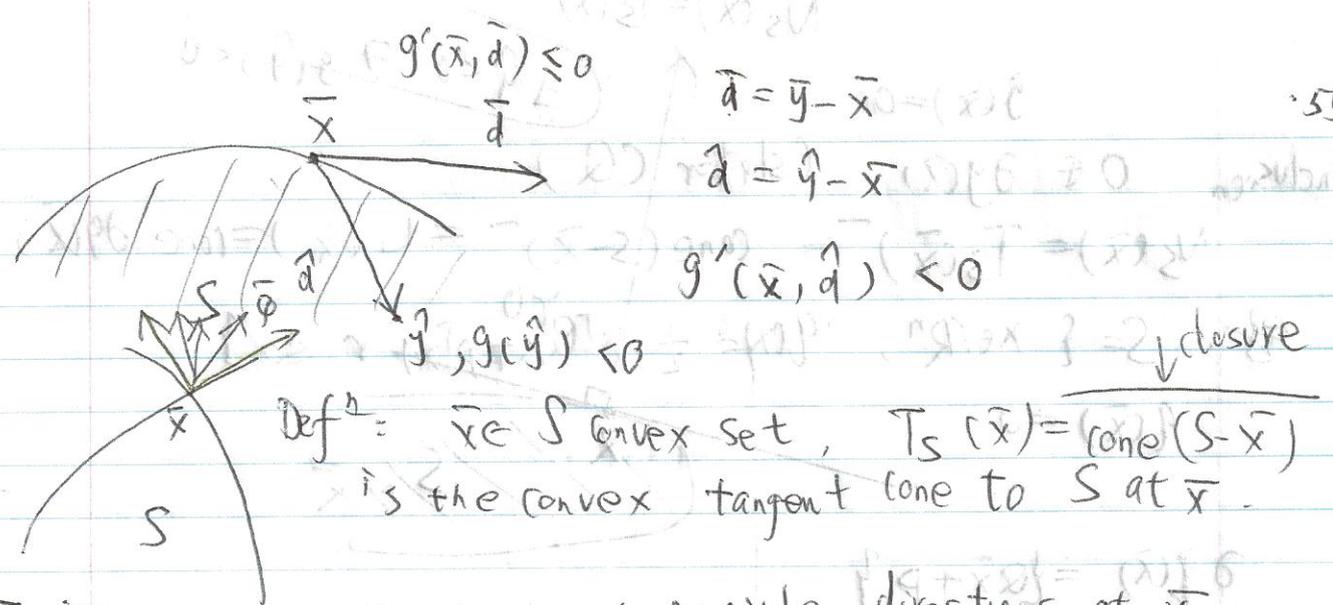
since we have a cone,  $\Rightarrow \beta \leq 0$  (wlog.  $\beta = 0$ )

$$\Rightarrow \bar{d} \in \partial g(\bar{x})$$

$$\Rightarrow g'(\bar{x}, \bar{d}) = \sigma_{\partial g(\bar{x})}(\bar{d}) \leq 0$$

$$\Rightarrow g'(\bar{x}, \bar{d}) = \lim_{t \downarrow 0} \frac{1}{t} (g(\bar{x} + t\bar{d}) - g(\bar{x}))$$

$$\leq 0$$



$T_S(\bar{x}) =$  closure of set of all feasible directions at  $\bar{x}$

$$= \{ d : \bar{x} + \alpha d \in S, \forall 0 < \alpha < \bar{\alpha} \text{ for some } \bar{\alpha} > 0 \}$$

if  $\text{int} S \neq \emptyset$ , then  $T_S(\bar{x}) = T_{\text{int}(S)}(\bar{x})$

Lemma 2.  $\bar{x} \in S$  convex,  
Then  $T_S(\bar{x}) = N_S(\bar{x})$

proof: By definition  $S \ni$

Def<sup>n</sup>  $g, S, \bar{x}$  as above

Lemma 3 Then  $T_S(\bar{x}) = \{ d : g'(\bar{x}; d) \leq 0 \}$

$L(\bar{x})$  notation

? linearizing cone? Called linearizing cone. relates cone  $\partial g(\bar{x})$ ?

$$g'(\bar{x}; \bar{d}) = \max_{\varphi \in \partial g(\bar{x})} \langle \varphi, \bar{d} \rangle \quad (\text{ref pg. 45})$$

$g'(\bar{x}; \bar{d}) \leq 0$  iff  $\langle \varphi, \bar{d} \rangle \leq 0 \forall \varphi \in \partial g(\bar{x})$

iff  $\langle \bar{x}, \bar{d} \rangle \leq 0 \forall \bar{x} \in \text{cone}(\partial g(\bar{x}))$

$$L(\bar{x}) = (\partial g(\bar{x}))^\vee$$

proof of lemma 3

$d \in T_S(\bar{x}) \Rightarrow d \in L(\bar{x})$  simple from def<sup>n</sup>  $S$

$\Leftarrow$  difficult

(use the picture, i.e. points from interior, i.e.  $\hat{y}, g(\hat{y}) < 0$ )  $\square$

Ailroy

$$N_S(\bar{x}) = T_S(\bar{x})^\circ$$

$$g(\bar{x}) = 0, \bar{x} = \bar{p} = \bar{r}$$

$$(\exists \eta < 0 \exists g(\eta) < 0)$$

conclusion

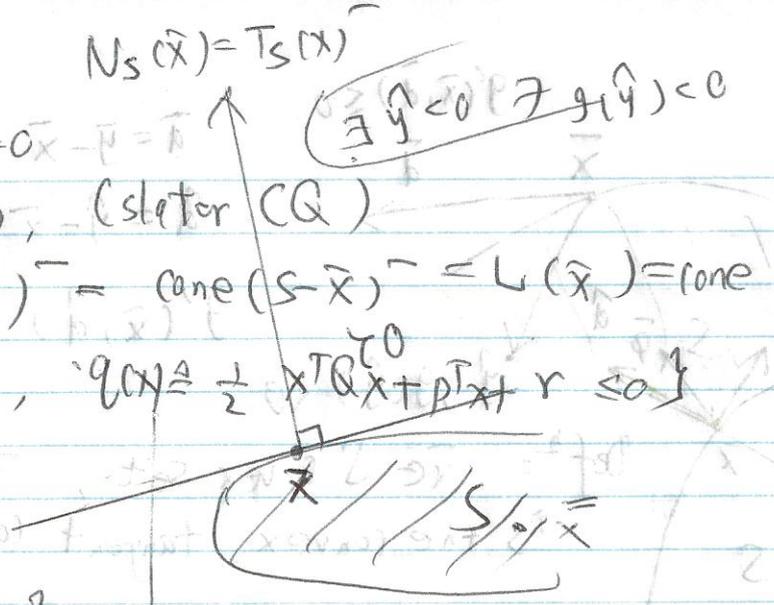
$0 \notin \partial g(\bar{x})$ , (stator CQ)

$$N_S(\bar{x}) = T_S(\bar{x})^\circ = \text{cone}(S - \bar{x})^\circ = L(\bar{x}) = \text{cone } \partial g(\bar{x})$$

eg.  $S = \{x \in \mathbb{R}^n, q(x) \triangleq \frac{1}{2} x^T Q x + p^T x + r \leq 0\}$

$$q(\bar{x}) = 0$$

$$\partial q(\bar{x}) = \{Q\bar{x} + p\}$$



$T_S(\bar{x})$  is a half space, a geometric description of  $S$  local at  $\bar{x}$

over all convex functions in the world

which ones have a min at  $\bar{x}$  over  $S$ ?

i.e. those  $f \ni$

$$(-\partial f(\bar{x})) \cap N_S(\bar{x}) \neq \emptyset$$

what if  $\nabla q(\bar{x}) = Q\bar{x} + p = 0$ ?

$$\Rightarrow \exists \eta < 0 \exists g(\eta) < 0$$

What about  $S_i = \{x : q_i(x) \leq 0\} \quad i=1, \dots, m$

$$\text{and } S = \bigcap_{i=1}^m S_i$$

$$\emptyset \neq \text{int } S$$

$$N_S(x) = ?$$

### Nov 1st, 2007 Convex analysis & Optimization

Suppose  $C$  convex set  $\bar{x} \in C$

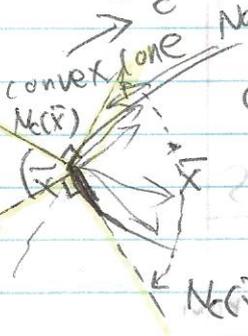
$D_C(\bar{x}) = \text{cone}(C - \bar{x})$  cone of feasible direction at  $\bar{x}$  in  $C$   
NOT NEC. closed.

$$T_C(\bar{x}) = \text{cl } D_C(\bar{x}) = \overline{D_C(\bar{x})}$$

(convex) tangent cone at  $\bar{x}$  in  $C$

"Local primal approximation of  $C$  at  $\bar{x}$ "

$$N_C(\bar{x})^\circ = T_C(\bar{x})$$



$$N_C(\bar{x}) = \{ \varphi = \langle \varphi, d \rangle \leq 0 \quad \forall d = y - \bar{x}, y \in C \}$$

normal cone to  $C$  at  $\bar{x}$ .

so  $= D_C^-(\bar{x}) = T_C^-(\bar{x})$

"A local dual approx. of  $C$  at  $\bar{x}$ ".

Prop.  $f$  convex proper,  $\bar{x} \in \operatorname{argmin}_{x \in S} f(x)$ ,  $S$  convex set (\*)

iff.  $\{ d \in D_C(\bar{x}) \Rightarrow f'(\bar{x}, d) < 0 \}$  (\*\*)

$\bar{x}$   $\swarrow$   $d$  proof: <sup>NEC</sup>  $\Rightarrow$  (for "any"  $f$ )

Suppose (\*) holds, but  $\exists d \in D_S(\bar{x}) \Rightarrow$

$f'(\bar{x}; d) < 0$ , then

$\exists \bar{t} > 0 \Rightarrow \bar{x} + td \in S \quad \forall 0 < t < \bar{t}$

and  $0 > f'(\bar{x}; d) = \lim_{t \downarrow 0} \frac{1}{t} [f(\bar{x} + td) - f(\bar{x})]$

$\Rightarrow 0 > \frac{1}{t} (f(\bar{x} + td) - f(\bar{x})) \quad \forall 0 < t < \bar{t} < \bar{t}$  for some  $\bar{t} > 0$

contrad. (We don't need convexity in NEC)

Suff ( $\Leftarrow$ ) (need convexity)

Assume (\*\*) But suppose (\*) fails

Then  $\exists x^1 \in S \Rightarrow f(x^1) < f(\bar{x})$

$\Rightarrow f'(\bar{x}; x^1 - \bar{x}) < 0$ , since  $f$  is a convex function.

contrad. □

(\*\*) iff.  $f'(\bar{x}; d) \geq 0 \quad \forall d \in D_C(\bar{x})$   
 primal characterisation of optimality.

we know  $f'(\bar{x}; d) = \underbrace{\sigma_{\partial f(\bar{x})}}_{\text{sublinear}}(d) \quad (\bar{x} \in \operatorname{int} \operatorname{dom} f)$   
 $= \max_{\varphi \in \partial f(\bar{x})} \langle \varphi, d \rangle$   
dual space  
primal space

$\bar{x}$  opt. ( $*$  holds) iff.  $(**)$   $f'(\bar{x}; d) \geq 0 \quad \forall d \in T_S(\bar{x})$

epi  $f_1$  / epi  $f_2$  iff  $0 \leq \inf_{d \in T_S(\bar{x})} \max_{\varphi \in \partial f(\bar{x})} \langle \varphi, d \rangle$  (add  $\|d\|=1$  here?)

$T_S(\bar{x})$  (compact convex)  $\leftarrow$  convex function. loc. Lip.

We will see (from duality theory) that iff.

$$0 \leq \max_{\varphi \in \partial f(\bar{x})} \left( \inf_{\substack{d \in T_S(\bar{x}) \\ \|d\|=1}} \langle \varphi, d \rangle \right) \stackrel{\min}{=} \begin{cases} \geq 0 & \text{if } \varphi \in (S - \bar{x})^\top \\ < 0 & \text{otherwise} \end{cases}$$

iff.  $\partial f(\bar{x}) \cap (S - \bar{x})^\top \neq \emptyset$  dual picture.

iff.  $(-\partial f(\bar{x})) \cap N_S(\bar{x}) \neq \emptyset$

For  $S = \{x \mid g(x) \leq 0\}$   $g$  convex, proper  
 $\exists \bar{x}, g(\bar{x}) < 0$ , (Slater  $\mathbb{Q}$ )

then  $g(\bar{x}) = 0$   
 $T_S(\bar{x}) = L(\bar{x}) (\equiv g'(\bar{x}; d) \leq 0)$

$$T_S(\bar{x})^\circ = N_S(\bar{x}) = \text{cone } \partial g(\bar{x})$$

$\bar{x} \in \text{dom}(f) \cap \text{int dom}(g)$

$$\text{so } \partial(f+g)(\bar{x}) = \partial f(\bar{x}) + \partial g(\bar{x})$$

So  $\bar{x} \in \arg \min_{x \in S} f(x)$  iff.  $\exists \varphi_f \in \partial f(\bar{x}) \varphi_g \in \partial g(\bar{x})$   
 $-\varphi_f = \lambda \varphi_g$  for some  $\lambda \geq 0$

iff  $\varphi_f + \lambda \varphi_g = 0$ , for some  $\lambda \geq 0 \quad \varphi_f \in \partial f(\bar{x}) \quad \varphi_g \in \partial g(\bar{x})$   
 KKT

So  $\bar{x}$  opt. for minimizing  $\min f(x)$   
 $g(x) \leq 0$

iff.  $\varphi_f + \lambda \varphi_g = 0$  for some  $\lambda \geq 0$   $\varphi_f \in \partial f(\bar{x})$   $\varphi_g \in \partial g(\bar{x})$

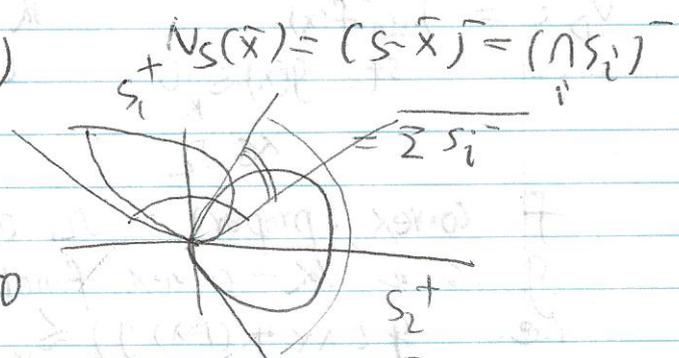
(dual feasibility)  $g(\bar{x}) \leq 0$   $\lambda g(\bar{x}) = 0$  (complementary slackness)  
 (primal feasibility)

for  $S = \{x \mid g_i(x) \leq 0 \quad i=1, \dots, m\}$  assume  $\exists \bar{x} \in \text{Slater CQ}$   
 $\varphi_i(\bar{x}) < 0 \quad \forall i$   
 Slater point

$S_i = \{x \mid g_i(x) \leq 0\}$   $\text{int } S_i \neq \emptyset$   
 $\text{int} \bigcap_i S_i = \bigcap_i \text{int } S_i \neq \emptyset$

$\Rightarrow N(S) = \sum_i N(S_i)$  (exer)

$N(S_i) = \begin{cases} \{0\} & \text{if } g_i(\bar{x}) < 0 \\ \text{cone } \partial g_i(\bar{x}) & \text{if } g_i(\bar{x}) = 0 \end{cases}$



$\bar{x} \in \arg \min_{x \in S} f(x)$  iff.  $-\varphi_f \in N_S(\bar{x})$  for some  $\varphi_f \in \partial f(\bar{x})$   
 $= \sum_i N_{S_i}(\bar{x})$

$= \sum_i \lambda_i \varphi_i$  for some  $\lambda_i \geq 0$   
 $\varphi_i \in \partial g_i(\bar{x})$  if  $g_i(\bar{x}) = 0$  Active set at  $\bar{x}$

$S = \bigcap_i S_i$   $S = \{\varphi : \langle \varphi, d \rangle \leq 0 \quad \forall d \in S\}$

$\sum S_i^- = \{\varphi = \sum_{i=1}^m \varphi_i : \langle \varphi_i, d \rangle \leq 0 \quad \forall d \in S_i\}$

Let  $d \in S = \bigcap_i S_i$   $\varphi \in \sum_i S_i^-$   
 $\Rightarrow \langle \varphi, d \rangle = \sum_{i=1}^m \langle \varphi_i, d \rangle \leq 0$  since  $d \in \bigcap_i S_i$  dual feasibility

$\bar{x} \in \arg \min_{x \in S} f(x)$  iff  $\varphi_f + \sum_{i=1}^m \lambda_i \varphi_i = 0$  for some  $\lambda_i \geq 0$   
 $\lambda_i \geq 0 \quad \varphi_i \in \partial g_i(\bar{x})$

Can be written as

primal feas.  $g_i(\bar{x}) \leq 0 \quad \forall i$   
 Complementary slackness  $\lambda_i g_i(\bar{x}) = 0 \quad \forall i$

$$\sum \lambda_i e_i = 0$$

$\swarrow$   $\text{ring}(z)$      $\searrow$   $\text{ring}(\lambda)$

will now use chain rule  $\rightarrow E$

$$\partial(f \circ A)(x) \supseteq A^{\text{adj}} \partial f(Ax) \quad \forall Ax \in \text{dom} f$$

equality holds if  $\text{int} \text{dom} f \cap A(Y) \neq \emptyset$

$f: E \rightarrow \mathbb{R}$

Next Fenchel Duality Nov. 6 2007

But first Lagrangian Duality

$$v_p := \min_{x \in \Omega} f(x) \quad \text{st. } g(x) \leq_K 0$$

$K$  is CCC.  $U \leq_K V$  means  $V - U \in K$

$f$  convex, proper.  $\Omega$  convex set (e.g.  $\Omega \subset (\text{epi} f) \cap \text{epi}(g)$ )  
 $g$  is a  $K$ -convex function (proper)  
 i.e.  $g(\lambda x + (1-\lambda)y) \leq_K \lambda g(x) + (1-\lambda)g(y)$

Exercise

- a) level sets for  $g$  are convex sets
- b)  $g$  is  $K$ -convex iff.  $g_\varphi(\cdot) := \langle \varphi, g(\cdot) \rangle$  is convex  
 iff.  $g_\varphi$  is convex  $\forall \varphi \in \mathcal{U}, \forall \varphi \in K^*$

for any set  $\mathcal{U} \Rightarrow K^* = \text{cone}(\mathcal{U})$  a "generating set"

Lagrangian  $L(x, \lambda) = f(x) + \langle \lambda, g(x) \rangle$

"payoff function" from player  $X$  to player  $\Omega$

game:  $X$  plays  $x \in \Omega$   
 $\Omega$  plays  $\lambda \in K^*$

$L(x, \lambda)$  from  $X$  to  $\Omega$

Best (worst-case) strategy for  $X$ :

$$\min_{x \in \Omega} \max_{\lambda \in K^*} L(x, \lambda) \quad (= f(x) + \langle \lambda, g(x) \rangle)$$

$X$  has a "hidden constraint"

(Since  $-k = -k^{++}$ )  $(g(x) \leq k^0)$

Now the max is attained at  $\lambda = 0$

We get  $V_p = \min_{x \in \Omega} \max_{\lambda \in K^+} L(x, \lambda)$

(p) back ie. player X best strategy is the primal problem (P) best (worst case strategy) for  $\Omega$ :

$V_p \geq V_D = \max_{\lambda \in K^+} (\min_{x \in \Omega} L(x, \lambda))$

weak duality th<sup>n</sup>  $\phi(\lambda)$  dual functional

(D)  $\max_{\lambda \in K^+} \phi(\lambda)$  dual problem

special case.  $f(x) = \langle c, x \rangle$   
 $g(x) = b - Ax, A: E \rightarrow Y$  lin. transf.  $b \in Y$   
 $\Omega$  ccc.

ie.  $x \succ_{\Omega} 0$  if  $x \in \Omega$   
 partial order

(D)  $\max_{\lambda \in K^+} \min_{x \in \Omega} \langle c, x \rangle + \langle \lambda, b - Ax \rangle$

$\lambda \in K^+ \quad x \in \Omega$   
 $\langle \lambda, b \rangle + \langle c - A^{adj} \lambda, x \rangle$   
 player  $\lambda$  has a hidden constraint  $\lambda \geq 0 \quad \forall x \in \Omega$

ie.  $(c - A^{adj} \lambda) \geq 0$  and then the min is attained at

$x = 0$

we get  $V_D = \max_{\lambda \in K^+} \langle \lambda, b \rangle$   
 st.  $A^{adj} \lambda \leq c, \lambda \geq 0$

$$V_p = \min_{\text{st.}} \langle c, x \rangle$$

$$Ax \geq_k b$$

$$x \geq_{\Omega} 0$$

So in LP.  $\min c_1^T x_1 + c_2^T x_2 + c_3^T x_3$

$$\begin{cases} A_{11}x_1 + A_{12}x_2 + A_{13}x_3 = b_1 & \text{--- (1)} \\ A_{21}x_1 + A_{22}x_2 + A_{23}x_3 \leq b_2 & \text{--- (2)} \\ A_{31}x_1 + A_{32}x_2 + A_{33}x_3 \geq b_3 & \text{--- (3)} \end{cases}$$

$x_1 \geq 0, x_2 \leq 0, x_3 \text{ free}$

Dual:

$$\max b_1^T \lambda_1 + b_2^T \lambda_2 + b_3^T \lambda_3$$

$$\text{st. } \begin{cases} A_{11}^T \lambda_1 + A_{21}^T \lambda_2 + A_{31}^T \lambda_3 \leq c_1 & \lambda_1 \text{ free} \\ A_{12}^T \lambda_1 + A_{22}^T \lambda_2 + A_{32}^T \lambda_3 \geq c_2 & \lambda_2 \leq 0 \\ A_{13}^T \lambda_1 + A_{23}^T \lambda_2 + A_{33}^T \lambda_3 = c_3 & \lambda_3 \geq 0 \end{cases}$$

In the LP case, (unless  $V_p + \infty > V_D = -\infty$  both infeasible) we get  $V_p = V_D$ , a zero duality gap.

(true for any polyhedral cones,  $K, \Omega$ )

ex. (p)  $V_p = \min_{x \in \mathbb{R}} 2x$  st.  $\frac{1}{2}x^2 \leq 0$

feasible set  $S = \{0\}$

$-\nabla f(x) = -2 \in N_S(0) = \mathbb{R}$  so 0 is opt.

KKT conditions,

p.f.  $\frac{1}{2}x^2 \leq 0$ , df.  $0 = \nabla L(x, \lambda) = \nabla (2x + \frac{\lambda}{2}x^2) = 2 + \lambda x = 2$   
 $\lambda \geq 0$  infeas.

CS  $\lambda (\frac{1}{2}x^2) = 0$

So **KKT FAILS !!**

(D)  $V_D = \max_{\lambda \geq 0} \min_x L(x, \lambda) = 2x + \frac{\lambda}{2}x^2$

$0 = \nabla_x L_0 = 2 + \lambda x$   
 $x_\lambda = -2/\lambda$

$L(x_\lambda, \lambda) = 2(-2/\lambda) + \frac{\lambda}{2}(4/\lambda^2) = -4/\lambda + \frac{2}{\lambda} = -2/\lambda$

$V_D = \max_{\lambda > 0} -2/\lambda = 0 = V_p$

ie. Strong duality fails,

zero duality gap but no opt  $\lambda$  !

i.e. strong duality  $V_p = V_D$  zero duality gap and  $V_D$  is attained. ( $v_p = p(\lambda^*)$ )

Note  $V_D = \max_{\lambda \geq 0} \min_x 2x x_0 + \frac{\lambda}{2} \cdot x^2 + t(1-x_0^2) \equiv Q(x, x_0) + t$   
 $\begin{matrix} \nearrow \\ \text{homogenize the Lagrangian, Note } V_D \text{ is not changed} \end{matrix}$   
 (always true for quadratic) since homog quad.  
 and  $1-x_0^2=0$  to Lagr. (we will see that  $V_D$  is still unchanged)

New "hidden constraint" for player  $\lambda$   $\rightarrow$  Hessian is  $\geq 0$   
 $\nabla^2 L(x, x_0, \lambda, t) \geq 0$   
 if  $H y = \lambda y$   $\lambda < 0$  then  $(\xi y)^T H (\xi y) = \xi^2 y^T H y \downarrow$   
 as  $\xi \rightarrow \infty$   $t$  now is in the Hessian  $< 0$   $-\infty$

$\Rightarrow$  min is 0 at  $x=0, x_0=0$

New dual is

$$\max t$$

$$\begin{cases} \begin{bmatrix} -2t & 0 \\ 0 & -\lambda \end{bmatrix} \succeq 0 \\ \lambda \geq 0 \\ t \geq 0 \end{cases} \Leftrightarrow \begin{pmatrix} 2t & 0 \\ 0 & -\lambda \end{pmatrix} \preceq \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

bex  
 $A: X \rightarrow Y$   
 lin. transf.  $C \in X(X^T)$   
 $\lambda \in Y (= Y^*)$

exercise dual of the SDP? Nov 8th, 2007

Recall Cone Duality (linear case)

(p)  $\min \langle c, x \rangle$   
 st.  $Ax \preceq_K b$   
 $x \succeq_{\Omega} 0$   
 $K, \Omega$  CCC.

(d)  $\max \langle b, \lambda \rangle$   
 st.  $A^* \lambda \succeq_{\Omega^*} c$   
 $\lambda \succeq_{K^*} 0$

eg.  $K = \mathbb{R}^p \oplus \mathbb{R}_+^n \oplus S_+^h$  similarly for  $\Omega$

eg. SDP - Semidefinite Programming

(P)  $\min \langle C, X \rangle = \text{trace } CX$

$\forall x$

$$C = \begin{bmatrix} \alpha & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \alpha > 0$$

$K = \{0\}$

$\Omega = S_+^n, n=3$

st.  $AX = \begin{pmatrix} x_{22} \\ x_{11} + 2x_{23} \end{pmatrix} = b$

$\mathbb{R}^2 \approx \mathbb{R}^b$

$b = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \in Y = \mathbb{R}^2$

$S_+^n = \{ X = X^T, n \times n, \text{ symmetric, real, } \lambda_i(X) \geq 0 \forall i \}$   
 $= \{ X = X^T; y^T X y \geq 0 \forall y \}$

$X \succeq 0 \quad A: S^3 \rightarrow \mathbb{R}^2$

(D)  $\max \langle b, \lambda \rangle$

$(S_+^n)^T = S_+^n$  (self-polar cone)

proof:  $X \succeq 0$  iff  $X = S^2$  for some  $S \succeq 0$

st.  $A^* \lambda \succeq C$

$A^{adj}$

$\lambda \succeq_{K^*} 0$

$\uparrow \rightarrow \{0\}^T = \mathbb{R}^2$

free.

$AX = \begin{pmatrix} x_{22} \\ x_{11} + 2x_{23} \\ x_{23} + x_{32} \end{pmatrix}$

What is  $A^* \lambda$ ?

$\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \approx \begin{bmatrix} 0 & 0 & 0 \\ c & 1 & 0 \\ c & 0 & 0 \end{bmatrix} = A_1, \quad A_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$

so  $AX = \begin{pmatrix} \langle A_1, X \rangle \\ \langle A_2, X \rangle \end{pmatrix}$

$A^* \equiv A^{adj}$  satisfies

$\lambda \in \mathbb{R}^2$

$\langle AX, \lambda \rangle = \langle X, A^* \lambda \rangle$

$\lambda_1 \langle A_1, X \rangle + \lambda_2 \langle A_2, X \rangle = \text{tr } X (A^* \lambda)$

$$\lambda_1 \text{trace}(A_1 X) + \lambda_2 \text{trace}(A_2 X) = \text{tr} X (A^* \lambda)$$

$$\text{trace}(\lambda_1 A_1 + \lambda_2 A_2) X = \text{tr} X (A^* \lambda)$$

$$A^* \lambda = \lambda_1 A_1 + \lambda_2 A_2$$

so if  $A: S^n \rightarrow \mathbb{R}^m$ , then  $\exists A_1, A_2, \dots, A_m \in S^n \rightarrow$

$$AX = (\text{trace}(A_i X)) \in \mathbb{R}^m \text{ and } A^* \lambda = \sum \lambda_i A_i \in S^n$$

so (D) here is

$$(D) \max_{\lambda} b^T \lambda = \lambda_2$$

$$A^* \lambda = \lambda_1 \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ c & c & c \end{bmatrix} + \lambda_2 \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \end{bmatrix} \preceq C = \begin{bmatrix} \alpha & 0 & 0 \\ 0 & \alpha & 0 \\ c & c & c \end{bmatrix}$$

$\lambda$  free.

and  $\det \begin{bmatrix} \alpha & 0 \\ 1 & \alpha \end{bmatrix} \geq 0 \Rightarrow \lambda_2 = 0$

(C) feasible set =  $\{ \lambda \in \mathbb{R}^2 : \lambda_1 \leq 0, \lambda_2 = 0 \}$

$\Rightarrow b^T \lambda = 0$  on feas. set  $\Rightarrow V_D = 0 \leq V_P$  (weak duality)

feasible set for (D) =  $\{ X \succeq 0 : \begin{matrix} X_{22} = 0 \\ X_{11} + 2X_{23} = 1 \end{matrix} \}$   $\Rightarrow X_{33} = 0 \Rightarrow X_{11} = 1$

$$\begin{bmatrix} X_{11} & X_{12} & X_{13} \\ X_{21} & X_{22} & X_{23} \\ X_{31} & X_{32} & X_{33} \end{bmatrix} \succeq 0 \Rightarrow X_{23} = 0$$

i.e.  $V_P > V_D$  (duality gap)  
both finite

(For polyhedral cones,  $V_P = V_D$  nec)

onto Fenchel duality  
ex.  $\min f_1(x) - f_2(x)$

st.  $x \in X_1 \cap X_2$   
both  $f_1, f_2$  are finite valued.

Use Lagr. duality

primal  $\rightarrow$

$$\min_{y \in X_1, z \in X_2} \max_{\lambda} L(y, z, \lambda)$$

$$\min f_1(y) - f_2(z)$$

$$\text{st. } y \in X_1, z \in X_2 \quad [z=y] \quad \text{FCR}$$

both  $f_1, f_2$  are finite valued.

dual functional  $\Phi(\lambda)$  in dual program  $\max_{\lambda} \min_{y \in X_1, z \in X_2} L(y, z, \lambda)$

$$\rightarrow = \min_{y \in X_1, z \in X_2} \{ f_1(y) - f_2(z) + \lambda^T (z - y) \}$$

separable in  $y, z$

$$= \min_{y \in X_1, z \in X_2} \{ z^T \lambda - f_2(z) \} + \min_{y \in X_1} \{ f_1(y) - y^T \lambda \}$$

$$= g_2(\lambda) - g_1(\lambda)$$

$$g_1(\lambda) = \sup_{x \in X_1} \{ x^T \lambda - f_1(x) \}$$

$$g_2(\lambda) = \inf_{x \in X_2} \{ x^T \lambda - f_2(x) \}$$

So dual problem is

$$\max g_2(\lambda) - g_1(\lambda)$$

$$\text{st. } \lambda \in \Omega_1 \cap \Omega_2$$

$$\text{where } \Omega_1 = \text{dom}(g_1)$$

$$\Omega_2 = \text{dom}(-g_2)$$

can use  $X_1 = \text{dom } f_1$   
 $X_2 = \text{dom}(-f_2)$   
 so if we allow extended functions,  $X_1, X_2$  can be discarded.

$\mathbb{R} \rightarrow (-\infty, +\infty]$  allow  $+\infty$  - extended

So now use "extended value" function

$$(b) \quad V_p = \inf \{ f(x) + g(x) \} \quad (f, g, \text{convex})$$

$$\text{Then } \forall \rho, x: f(x) \geq \langle -\rho, x \rangle - f^*(-\rho)$$

(def)

$$g(x) \geq \langle \rho, x \rangle - g^*(\rho)$$

Weak Duality

$$\circ V_p \geq \sup_{\rho} \{ -f^*(-\rho) - g^*(\rho) \}$$

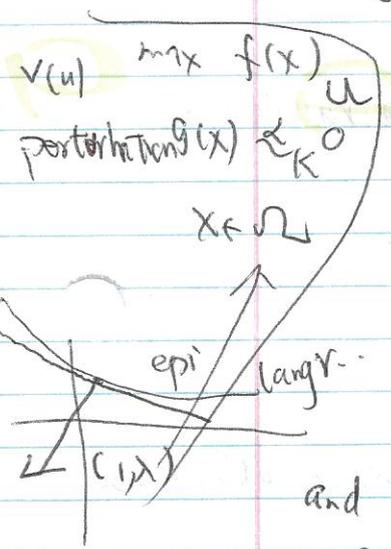
→ The Fenchel Dual.

$V_p - V_D$  is the duality gap.

We now show that  $\text{int}(\text{dom}(f)) \cap \text{dom}(g) \neq \emptyset \Rightarrow$  Strong Duality

Let  $h(u) = \inf_x \{ f(x) + g(x-u) \}$   
 =  $\inf_x \{ f(x) + \tilde{g}(u-x) \}$  where  $\tilde{g}(y) = g(-y)$   
 $h(u) = (f \circ \tilde{g})(u)$  called Infimal convolution.

*Annotations:*  
 -  $h(u)$  is labeled "perturbation".  
 -  $\tilde{g}(y) = g(-y)$  is underlined.  
 - To the right:  $V_p = V_D$  and  $V_D$  is attained if finite.



Lemma Let  $f_i$  be proper, convex on  $E$ , and  $f(x) = \inf_x \{ \sum_{i=1}^m f_i(x_i) : x_i \in E, \sum_{i=1}^m x_i = x \}$

Then  $f$  is convex.

Proof: Let  $F_i = \text{epi } f_i$ .

$\Rightarrow F = \sum_{i=1}^m F_i$  is a convex set.

and  $F = \text{epi } f$  since

$(x, u) \in F$  iff  $\exists x_i \in E, \mu_i \in \mathbb{R} \rightarrow$

$\mu_i \geq f_i(x_i) \quad u = \sum \mu_i \quad x = \sum x_i$  trivial by def.

Nov 3th 2007 Birthday ☺

Theorem (Fenchel Duality)

If  $f$  and  $g$  are proper and convex then

$$\inf_x \{ f(x) + g(x) \} \geq \sup_p \{ -f^*(-p) - g^*(p) \}$$

If  $\text{int}(\text{dom}(f)) \cap \text{dom}(g) \neq \emptyset$ , then equality holds and the sup is attained when finite.

Proof: Recall  $h(u) = \inf_x \{ f(x) + \tilde{g}(u-x) \}$ ,  $\tilde{g}(x) = g(-x)$   
 $h(u) = (f \circ \tilde{g})(u)$  infimal convolution.

Hilroy

Then  $h$  is convex.

To show equality holds and attainment. pick  $\bar{x} \in \text{int}(\text{dom}(f)) \cap \text{dom}(g)$

Now show  $h$  is proper.

$\exists \delta > 0$ , st.  $\{\bar{x}\} + \delta B \subseteq \text{dom}(f)$   
unit ball

$$\forall u \in \delta B \quad h(u) \leq f(\bar{x}+u) + g(\bar{x}) < \infty$$

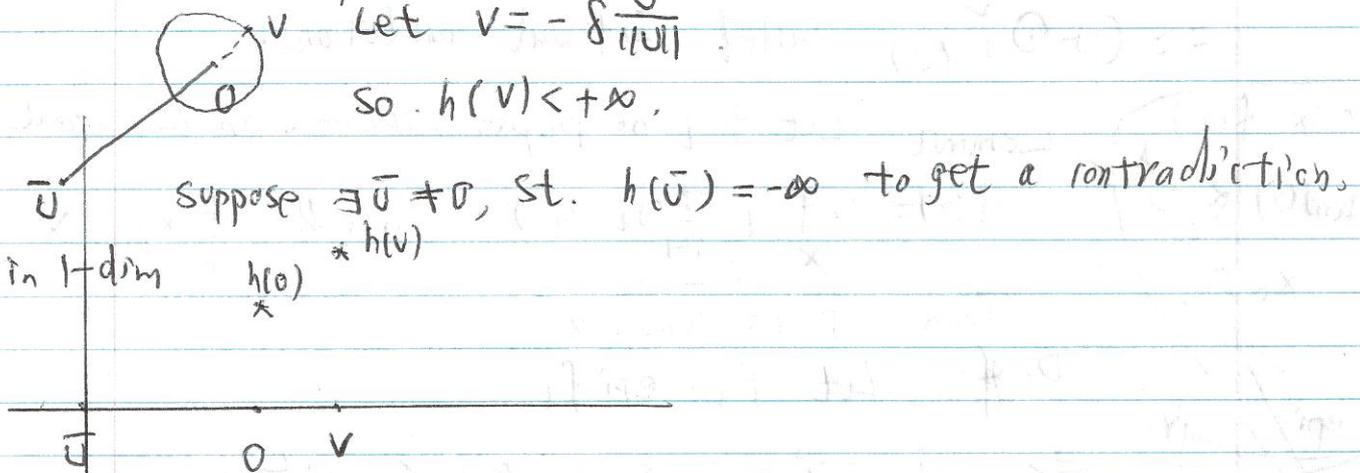
Thus  $0 \in \text{int} \text{dom}(h)$

In addition, wlog.  $h(0) > -\infty$ .

(If  $h(0) = -\infty$ , then  $-\infty \equiv \inf \equiv \sup$ )

Let  $v = -\frac{u}{\|u\|}$

so  $h(v) < +\infty$ .



Let  $S \geq \lambda(v) \quad \forall r \in \mathbb{R}, (\bar{u}, r) \in \text{epi}(h)$  since  $h(\bar{u}) = -\infty$   
 $(v, S) \in \text{epi}(h)$

$$\text{so } \left(0, \frac{\delta r + \|\bar{u}\| S}{\delta + \|\bar{u}\|}\right) = \left(\frac{\delta}{\delta + \|\bar{u}\|} \bar{u} + \frac{\|\bar{u}\|}{\delta + \|\bar{u}\|} v, \frac{\delta r + \|\bar{u}\| S}{\delta + \|\bar{u}\|}\right) \in \text{epi}(h)$$

Let  $r \downarrow -\infty$ , which contradicts  $h(0) > -\infty$ .

so  $h$  is proper,  $0 \in \text{int}(\text{dom}(h)) \Rightarrow \partial h(0) \neq \emptyset$ .

Pick  $-\rho \in \partial h(0)$

$$\textcircled{a} \quad -h(0) = h^*(-\rho)$$

$$= (f \oplus g)^*(-\rho) \quad (\text{on assign 4})$$

$$= f^*(-\rho) + g^*(-\rho)$$

$$\langle -\rho, y - 0 \rangle \leq h(y) + h(0)$$

$$\langle -\rho, y \rangle - h(y) \leq h(0)$$

use  $y = 0$  to get  $-\rho \in \partial h(0)$

Now  $f^*(-\phi) = \sup_x \{ \langle -\phi, x \rangle - g(-x) \}$   
 $= \sup_x \{ \langle \phi, x \rangle - g(x) \} = g^*(\phi)$

So  $-h(0) = -\inf_x \{ f(x) + g(x) \} = f^*(-\phi) + g^*(\phi) = \square$   
 (attained with zero duality gap)

eg. Take  $f = \delta_S$ ,  $S = \text{convex set}$

Then  $\text{int}(\text{dom } f) \cap \text{dom } g = \text{int}(\text{dom } f) \cap S \subset \mathbb{R}^n$

Fenchel duality  $\Rightarrow$   
 $x \in \text{argmin}_{x \in S} f(x)$

$\Leftrightarrow \exists \phi$ , st.  $f(x) = -f^*(-\phi) - \delta_S^*(\phi)$   
 and  $x \in S$

$\Leftrightarrow \exists \phi$ , st.  $f(x) + f^*(-\phi) - \langle -\phi, x \rangle$   
 $x \in S$  and  $= -\sup_{y \in S} \{ \langle \phi, y \rangle \} - \langle -\phi, x \rangle$

$\Leftrightarrow x \in S$  and  $\exists \phi$  st.  $f(x) + f^*(-\phi) - \langle -\phi, x \rangle$   
 $= + \inf_{y \in S} \{ \langle -\phi, y-x \rangle \}$

$\Leftrightarrow x \in S$  and  $\exists \phi$  st.  $f(x) + f^*(-\phi) = \langle -\phi, x \rangle$   
 and  $\inf_{y \in S} \{ \langle -\phi, y-x \rangle \} = 0$  use Fenchel-Young

$\Leftrightarrow \exists -\phi \in \partial f(x) \Rightarrow \langle \phi, y-x \rangle \leq 0 \quad \forall y \in S$

$\Leftrightarrow (-\partial f(x)) \cap N_S(x) \neq \emptyset$

LP duality

$V_p(\epsilon) = \inf \{ \langle C, x \rangle : Ax \leq b, Dx = e \}$   $\leftarrow$  expect free dual variable

$\frac{\partial V_p(\epsilon)}{\partial \epsilon_i} \downarrow$  As  $\epsilon_i \uparrow$  feasible set  $\uparrow \Rightarrow$  expect  $\phi \leq 0$

$= \inf_{x, y} \{ \langle C, x \rangle + \delta_{\{Ax+y=b\}} + \delta_{\{Dx=e\}} + \delta_{\{y \geq 0\}} \}$   
 $= \inf_{x, y} \{ \langle C, x \rangle + \delta_{\{Ax+y=b\}} + \delta_{\{Dx=e\}} + \delta_{\{y \geq 0\}} \}$   
 (with annotations:  $A$  is  $m \times n$ ,  $D$  is  $p \times n$ ,  $y$  is stack variable,  $\delta_{\{y \geq 0\}}$  is  $\delta_m$  over  $\mathbb{R}_+$ )

$$f(x, y) = \langle c, x \rangle + \delta_{\mathbb{R}_+^m}(y)$$

$$\langle (\varphi, \psi), (x, y) \rangle$$

$$f^*(\varphi, \psi) = \sup_{x, y} \{ \langle \varphi, x \rangle + \langle \psi, y \rangle \}$$

(eliminate x, y)

$$= \left( \delta_{\langle c, \cdot \rangle}(\varphi) + \delta_{\mathbb{R}_+^m}(\psi) \right)$$

$$g^*(\varphi, \psi) = \sup_{x, y} \{ \langle \varphi, x \rangle + \langle \psi, y \rangle : Ax + y = b, Dx = e \}$$

(hidden constraint)

$$= \sup_x \{ \langle \varphi, x \rangle + \langle \psi, b - Ax \rangle : Dx = e \}$$

$$= \sup_x \{ \langle \varphi - A^{\text{adj}} \psi, x \rangle : Dx = e \} + \langle \psi, b \rangle$$

using adjoint

$$= \delta_{\mathcal{R}(\Delta^{\text{adj}})}(\varphi - A^{\text{adj}} \psi) + \langle \varphi - A^{\text{adj}} \psi, x^* \rangle + \langle \psi, b \rangle$$

Simplify sup over  $\varphi, \psi$

----- exercise -----

$\mathcal{R}(\Delta^{\text{adj}})^\perp = \mathcal{N}(\Delta)$   
Range space Null space

$$\begin{aligned} 0 = \langle \Delta x, \varphi \rangle &\text{ iff } \Delta x = 0 \\ &\text{ iff } \langle x, \Delta^{\text{adj}} \psi \rangle = 0 \quad \forall \psi \\ &\text{ iff } x \in \mathcal{R}(\Delta^{\text{adj}})^\perp \end{aligned}$$

$$Dx = e \text{ iff } x = x' + n \quad \text{some } n \in \mathcal{N}(D)$$

Nov 15th 2007 Cloudy 1:00pm

Recall  $f^*(-\varphi) + f(\varphi) \geq \langle -\varphi, x \rangle$

$$g^*(\varphi) + g(x) \geq \langle \varphi, x \rangle$$

$$\inf_x f(x) + g(x) \geq \sup_{\varphi} \{ -f^*(-\varphi) - g^*(\varphi) \}$$

with equality if  $\text{int dom}(f) \cap \text{dom}(g) \neq \emptyset$

We applied it to LP  $\leftarrow$  Constraint Qualification

convex conic opt.

$$\inf \{ \langle c, x \rangle : Ax \preceq_K b, Dx = e \}$$

$\leftarrow$  convex cone

$$\text{slack variable } y = b - Ax$$

if  $V_p$  (or  $V_h$ ) is finite, then  $V_p = V_h$  and both are attained

$$= \inf \{ \langle c, x \rangle + \delta_{\{b\}}(Ax + y) + \delta_{\{e\}}(Dx) + \delta_K(y) \}$$

$f(x, y) \leftarrow \underbrace{\langle c, x \rangle + \langle y, y \rangle}_{g(x, y)}$

So  $f^*(\varphi, \psi) = \sup_{x, y} \{ \langle \varphi, x \rangle + \langle \psi, y \rangle - \langle c, x \rangle - \delta_K(y) \}$

$$= \delta_{\{c\}}(\varphi) + \delta_{K^*}(\psi) \quad \text{with } \varphi = e$$

$$g^*(\varphi, \psi) = \delta_{R(D^*)}(\varphi - A^*\psi) + \langle \varphi - A^*\psi, x \rangle + \langle \psi, b \rangle$$

$A^* \equiv A^{adj}$       Note  $g_v(x) = \langle v, Ax \rangle$

$$g_v^*(\varphi) = \sup_x \{ \langle \varphi, x \rangle - g_v(x) \}$$

$$= \sup_x \{ \langle \varphi - A^*v, x \rangle \}$$

$$= \delta_{R(A^*)}(\varphi - A^*v)$$

So the dual is

$$\sup_{\varphi, \psi} \{ -f^*(\varphi, -\psi) - g^*(\varphi, \psi) \}$$

after  $\psi \leftarrow -\psi$      $c = A^*\varphi + D^*\eta$

$$\Rightarrow \underbrace{-c - A^*\varphi = -D^*\eta}_{\text{for some } \eta}$$

$$= \sup_{\varphi, \psi} \{ -\delta_{\{c\}}(\varphi) - \delta_{K^*}(-\psi) - \delta_{R(D^*)}(\varphi - A^*\psi) \}$$

$\Rightarrow \varphi = -c \quad -\psi \in K^*$

$$\langle \varphi - A^*\psi, x \rangle - \langle \psi, b \rangle$$

$A: X \rightarrow Y$   
 $D: X \rightarrow Z$

$$-\langle -c - A^*\psi, x \rangle \quad \text{after } \psi \leftarrow -\psi$$

$X, Y, Z$  finite dim.

$$\langle D^*\eta, x \rangle = \langle \eta, e \rangle + \langle \psi, b \rangle$$

$\eta \in K^*$

$$= \sup_{\varphi, \eta} \{ \langle \eta, e \rangle + \langle \psi, b \rangle : A^*\varphi + D^*\eta = c \quad \varphi \in K^* \}$$

$\eta$  free

Linear, always finite

$$f(x, y) = \langle c, x \rangle + \delta_K(y) = +\infty \quad \text{iff } y \in K \neq \emptyset$$

$$\Rightarrow \text{dom}(f) = K \subset Y$$

$$\text{int}(\text{dom}(f)) = ? \quad K \neq \emptyset$$

$\text{int} K = K$

$$\text{dom}(g) = ? \quad \{(x, y) : Ax + y = b, Dx = e\}$$

$\neq \emptyset$  if lin. syst. are consist.

("generalized") Slater CQ:

$$\exists \hat{x}, \hat{y}, \hat{z} \quad Ax + \hat{y} = b \quad Dx = e \quad \hat{y} \in \text{int} K$$

$$A\hat{x} \prec_K b \quad D\hat{x} = e$$

Note The Mangasarian - Fromwittz CQ here needs  $D$  is "onto"

When Slater CQ holds, we get strong duality, i.e. opt. values are equal and the dual is obtained if finite.

So here M-F means Slater and  $D$  onto. (if Slater's holds)  $x^*$  is opt. (for primal) iff.

equiv.  $\mathcal{N}(D^*) = \{0\}$

$$\boxed{\exists \eta, \gamma \quad \gamma \in K^* \quad A^* \gamma + D^* \eta = c} \quad \boxed{Ax^* \preceq_K b \quad Dx^* = e} \quad (\text{pf.})$$

(b.f)

and  $\langle c, x^* \rangle = \langle e, \eta \rangle + \langle b, \gamma \rangle \dots (*)$

From  $(*)$  :  $0 = \langle c, x^* \rangle - \langle e, \eta \rangle - \langle b, \gamma \rangle$

$$= \langle A^* \gamma + D^* \eta, x^* \rangle - \langle e, \eta \rangle - \langle b, \gamma \rangle$$

$$= \langle \gamma, Ax^* \rangle + \langle \eta, Dx^* \rangle - \langle e, \eta \rangle - \langle b, \gamma \rangle$$

$$= \langle -\gamma, b - Ax^* \rangle \geq 0$$

$\underbrace{\gamma}_{K^*} \quad \underbrace{b - Ax^*}_{\in K} \geq 0$

$$\boxed{\langle \gamma, b - Ax^* \rangle = 0} \quad \text{Complementary Slackness}$$

Remark

If  $K = S_+^n$  so  $y = b - Ax^* \in S_+^n \quad \forall x \in S_+^n$

Lemma  $U, V \succeq 0 \Rightarrow \langle U, V \rangle = 0$  iff  $UV = 0$

proof:  $\text{tr} UV = \text{tr} U^{1/2} U^{1/2} V = \text{tr} U^{1/2} V U^{1/2}$  etc. ---

In  $\mathbb{R}_+^n \quad U, V \succeq 0 \Rightarrow \langle U, V \rangle = 0$  iff  $U_i V_i = 0 \quad \forall i$

$$\sum U_i V_i = 0$$

eg. **Qnstr. Opt.**

primal  $V_p = \inf \{ f(x) : g_i(x) \leq 0 \quad i=1, \dots, m \}$

convex function

$$g_i(x) + y_i = 0 \quad y_i \geq 0$$

$$V_p = \inf \left\{ f(x) + \sum_{i=1}^m \delta_{\{0\}}(g_i(x) + y_i) + \delta_K(y) \right\}$$

$K = \mathbb{R}_+^m$

Let  $\hat{f}(x, y) = \delta_K(y)$

$\hat{g}(x, y) = f(x) + \sum_{i=1}^m \delta_{\{0\}}(g_i(x) + y_i)$

$\hat{f}^*(\rho, \psi) = \underbrace{\delta_{\{0\}}(\rho)}_{\text{since no } x} + \delta_{K^-}(\psi)$

$$K^- = -\mathbb{R}_+^m = \mathbb{R}_+^m$$

$\hat{g}^*(\rho, \psi) = \sup_{x, y} \{ \langle \rho, x \rangle + \langle \psi, y \rangle - f(x) : g_i(x) + y_i = 0 \quad \forall i \}$

$= \sup_x \{ \langle \rho, x \rangle - f(x) - \sum_{i=1}^m \psi_i g_i(x) \quad x \in \bigcap_{i=1}^m \text{dom} g_i \}$

$V_p \geq V_D = \sup_{\rho, \psi} \{ -\delta_{\{0\}}(-\rho) - \delta_{K^-}(-\psi) \} - \sup_x \{ \langle \rho, x \rangle - f(x) - \sum \psi_i g_i(x) \}$

$\rho = 0 \quad \psi \in K^+$

$= \sup_{\psi \in K^+} \inf_x \{ f(x) + \sum \psi_i g_i(x) : x \in \bigcap \text{dom} g_i \}$

Lagrangian / payoff function

$\text{dom } \hat{f} = \mathbb{R}^n \oplus \mathbb{R}^m \quad \text{dom } \hat{g} = \{ (x, y) : x \in \text{dom} f \cap (\bigcap \text{dom} g_i) \}$

Hilroy



$\exists y \notin H$ , Apply hyperplane separation theorem,

$\exists \phi \Rightarrow \langle \phi, y \rangle < \langle \phi, h \rangle \forall h \in H$ .

use  $-h \in H \Rightarrow \phi \in H^\perp$

$\Rightarrow 0 \neq \phi \in (S^+)^{\perp} = S^{+ \perp} = (S^{++}) \cap (S^{+-}) = S \cap (-S)$

Bit of Summary / Review.

To solve minimize  $f(x)$   
 $x \in C$

We can redefine  $f$  so  $f(x) = +\infty$  if  $x \notin C$

parameterized problem, e.g.

$h(u) = \inf_{x \in X} F(x, u)$   $F$  convex in  $(x, u) \Rightarrow h$  is convex, (use epigraph argument)

Saddle-point problems / game theory.

$\mathcal{L} = X \times \Lambda \rightarrow [-\infty, +\infty]$

$f(x) := \sup_{\lambda \in \Lambda} \mathcal{L}(x, \lambda)$  }  $\text{min}_{x \in X} f(x)$

$\phi(x) = \inf_{\lambda \in \Lambda} \mathcal{L}(x, \lambda)$  }  $\text{max}_{\lambda \in \Lambda} \phi(x)$

$f(x) \geq \inf_{x \in X} f(x) \geq \sup_{\lambda \in \Lambda} \phi(x) \geq \phi(x)$

$\geq \alpha(x, \lambda) \geq$

called a saddle-point if it exists.

eg. (of Fenchel Duality)

$f(x) := \max_{i=1, \dots, m} f_i(x)$  proper convex functions.

$f^*(\phi) = \sup_x \{ \langle \phi, x \rangle - \max_{i=1, \dots, m} f_i(x) \}$

$= \sup_{x, t} \{ \langle \phi, x \rangle - t : f_i(x) - t \leq 0 \}$

$$= - \inf_{x, t} \{ t - \langle \varphi, x \rangle : f_i(x) - t \leq 0 \quad \forall i \}$$

Assume  $\bigcap_{i=1}^m \text{dom } f_i \neq \emptyset \Rightarrow$  Slater CQ holds

$$\Rightarrow f^*(\varphi) = - \sup_{\psi \geq 0} \inf_{x, t} \{ t - \langle \varphi, x \rangle + \sum_{i=1}^m \psi_i (f_i(x) - t) \}$$

strict feasibility

$$= - \inf_{x, t} \{ t - \langle \varphi, x \rangle + \sum_{i=1}^m \psi_i (f_i(x) - t) \}$$

for some  $\psi_i \geq 0$

$$= - \inf_x \{ - \langle \varphi, x \rangle + \sum \psi_i f_i(x) \} - \inf_t \{ t (1 - \sum \psi_i) \}$$

(dual attained)

$$\Rightarrow \exists \psi \geq 0 \quad \sum_{i=1}^m \psi_i = 1$$

$$f^*(\varphi) = \sup_x \{ \langle \varphi, x \rangle - \sum_{i=1}^m \psi_i f_i(x) \}$$

$$= (\sum_{i=1}^m \psi_i f_i)^*(\varphi)$$

$\psi$  dependent on  $\varphi$ .

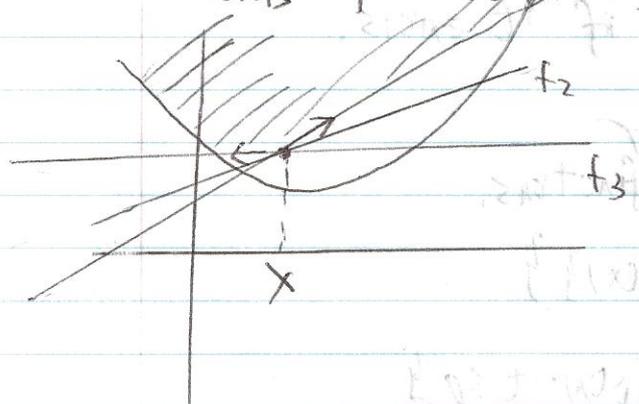
Using inf-conv. .... get

$$\underline{\text{Th}}^m \quad (\max f_i)^*(\varphi) \leq \inf_{\varphi', \dots, \varphi^m} \left\{ \sum_{i=1}^m \psi_i f_i^*(\varphi^i) : \sum \psi_i \varphi^i = \varphi, \sum \psi_i = 1, \psi_i \geq 0 \quad \forall i \right\}$$

Th<sup>m</sup> (subdiff)

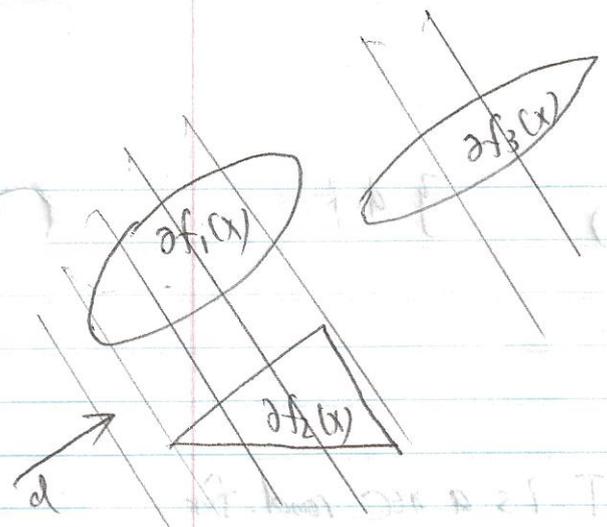
$$\partial f(x) \subseteq \text{conv} \left( \bigcup_{i \in I} \partial f_i(x) \right) \text{ where}$$

$$= \text{holds if } \text{int}(\text{dom } f) \neq \emptyset \quad I = \{ i : f_i(x) = f(x) \}$$

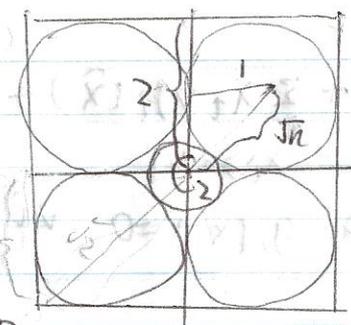


$$f'(x; d) = \sigma_{\partial f(x)}(d)$$

$$= \sigma(d) = \max_{i \in I} f'_i(x; d)$$



Question =



radius  $\frac{r_{n-1}}{2}$

$B_2$  in  $\mathbb{R}^3$

$E_3, B_3$

Q:  $\lim_{n \rightarrow \infty} \frac{|C_n|}{|B_n|} = +\infty?$

**KKT OPT. Cond.**

$\inf_x \{ f(x) = g_i(x) \leq 0, i=1, \dots, m \}$   
 $(x \in \Omega)$

$\bar{x}$  is an opt. sol<sup>n</sup> iff.

d.f.  $0 = \nabla f(\bar{x}) + \sum_{i=1}^m \lambda_i \nabla g_i(\bar{x}) \quad \lambda_i \geq 0 \quad \forall i$

p.f.  $g_i(\bar{x}) \leq 0 \quad \forall i$

c.s.  $\lambda_i g_i(\bar{x}) = 0 \quad \forall i$

Nov 22th, Snow Thursday 2007 1:00pm

**General NLP**

$\min f(x)$   
 st.  $g_i(x) \leq 0$  (or  $g_i(x) \leq_k 0$ )

generalized Slater CQ.  $i=1, \dots, m$

Convex case  $\exists \bar{x} \quad g_i(\bar{x}) < 0$

(Assume diff.)  $h_j(x) = 0, j=1, \dots, p$  At  $\bar{x}$  a feasible point, Monjarian - From Witz

CQ (MFCC)

$\exists d \neq 0 \quad g_i'(\bar{x}; d) = \nabla g_i(\bar{x})^T d < 0 \quad \forall i \in I = \{i : g_i(\bar{x}) = 0\}$

$h_j'(\bar{x}; d) = \nabla h_j(\bar{x})^T d = 0 \quad \forall j$

and  $\{ \nabla h_j(\bar{x}) \}_{j=1}^p$  is lin. indep.

Th<sup>m</sup> (KKT) If  $\bar{x}$  is a "local" opt. for NLP and MFCC hold

at  $\bar{x}$ , then  $g_i(\bar{x}) \leq 0 \quad \forall i$   
 $h_j(\bar{x}) = 0 \quad \forall j$  } P.f.

$$0 = \nabla f(\bar{x}) + \sum_i \lambda_i \nabla g_i(\bar{x}) + \sum_j \mu_j \nabla h_j(\bar{x}) \quad \} \text{ d.f.}$$

C.S.  $\lambda_i g_i(\bar{x}) = 0 \quad \forall i$

In Convex Case, if GSCQ holds, then KKT is a nec. cond. for opt at  $\bar{x}$ .

Note KKT is always a suff condition for opt. at  $\bar{x}$

Now look at the Parameterizations / Lagrangian to motivate conjugate duality.

We saw advantages of using extended valued functions

where  $f$  is redefined to  $+\infty$  for  $x$  not feasible.

Parameterized Problem

$$F(x) = \inf_{x \in X} F(x, u) \quad \text{where } F(x, 0) \equiv f(x)$$

We use "Lagrangian / payoff"

$$K : X \times Y \rightarrow [-\infty, +\infty]$$

$$f(x) = \sup_{y \in Y} K(x, y)$$

$$f(x) \geq K(x, y) \geq g(y)$$

$$g(y) = \inf_{x \in X} K(x, y)$$

$$\otimes \quad \inf_{x \in X} \sup_{y \in Y} K(x, y) = \inf_{x \in X} f(x)$$

(P/Q)  $\min_{x \in X} f(x)$

$$\geq \sup_{y \in Y} g(y) = \sup_{y \in Y} \inf_{x \in X} K(x, y)$$

(D/Q)  $\max_{y \in Y} g(y)$

if  $=$  holds, then we have a saddle value.

$$\otimes \otimes \quad K(x, \bar{y}) \geq K(\bar{x}, \bar{y}) \geq K(\bar{x}, y) \quad \forall x \in X \quad \forall y \in Y$$

called a saddle point.

eg.  $K(x, y) = x^2 - y^2$

$\Rightarrow (\bar{x}, \bar{y}) = (0, 0)$  is a saddle point.

Th<sup>m</sup> the pair  $(\bar{x}, \bar{y})$  is a saddle-point iff  $\bar{x}$  solves ① and  $\bar{y}$  solves ② and saddle value exists.

pf --- Follow from  $\text{inf}_{x \in X} f(x) = \text{sup}_{y \in Y} g(y)$

Note For each  $y \in Y$ :  $\text{inf}_{x \in X} f(x) \geq \text{inf}_{x \in X} K(x, y)$

eg LP diet problem provides a lower bound.

$c_j$  - # unit cost for food  $j = 1, \dots, n$

$b_i$  - # unit min nutrient requirement  $i = 1, \dots, m$

$A_{ij}$  - # units of nutrient  $i$  in Food  $j$

$x_j$  - # units of food  $j$  in diet

$y_i$  - : unit cost of nutrient  $i$  (shadow price)

$\min \sum c_j x_j$  (=  $c^T x$ ) cost of diet

st.  $Ax \geq b$  (meet min. nutrient requir...)

$x \geq 0$

lagrangian / payoff

$K(x, y) = c^T x + y^T (b - Ax)$   $x$  to  $y$

$X = \mathbb{R}_+^n$   $Y = \mathbb{R}_+^m$

② becomes the dual LP  $\max b^T y$

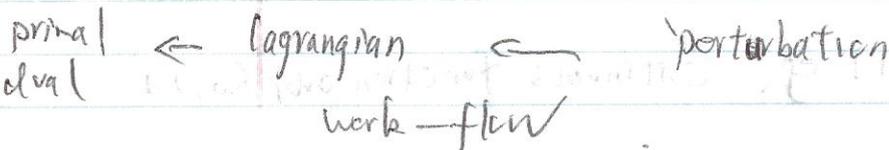
st.  $A^T y \leq c$

$y \geq 0$

How do we construct  $K$  for general problems?

We get that conjugation provides a 1-1 map from ①

$K(x, y) \leftrightarrow F(x, 0)$  where  $F: X \times Y \rightarrow [-\infty, +\infty]$



eg.  $f_i, \text{convex}, i=0, \dots, m$

$\min_{x \in X} f(x)$

convex set

$$f(x) = \begin{cases} f_0(x) & \text{if } x \in C, f_i(x) \leq 0 \quad \forall i=1, \dots, m \\ +\infty & \text{otherwise} \end{cases}$$

(circled 0)  $\rightarrow U_i \text{ for } F$

embed in family  $F(x, u)$

replace  $f_i, i=0, \dots, m$  by  $h_i(A_i x + q_i) + l_i(x)$

$\uparrow$  linear maps       $\uparrow$  Affine

$u = (u_1, \dots, u_m, z_0, \dots, z_m)$  For  $F(x, u)$  add  $(-z_i)$

and  $\leq u_i$  again.

again.  $F$  maintains convexity.  $F(x, 0) \equiv f(x)$   
 or with fixed  $r > 0$

add  $f_0(x) + (r \|u\|_2^2)$

this modification leads to a "different" dual.

eg.  $\min_{x \in C} f_0(x)$  st  $h(x, s) \leq 0 \quad \forall s \in S$

convex in  $x$  for each  $s$

could be infinite # constr.

$$f(x) = \begin{cases} f_0(x) & \text{if } x \in C, h(x, s) \leq 0 \quad \forall s \in S \\ +\infty & \text{otherwise} \end{cases}$$

(circled 0)  $\uparrow$   $U(s)$  function.  
 add perturbation

For  $F(x, u)$

eg. Chebyshev Approx.

$h_i = [0, 1] \rightarrow \mathbb{R} \quad i=0, \dots, m$

$\min_{x \in X} f(x) = \|h_0 - \sum_{i=1}^m x_i h_i\|_\infty$

$= \max_{0 \leq t \leq 1} |h_0(t) - \sum_{i=1}^m x_i h_i(t)|$

$+ \in \mathbb{R}^m$

it is a min max problem

parametrize add  $(+u) \quad u \equiv 0(t)$

$u \in C [0, 1]$  eg. continuous function over  $[0, 1]$

$0 \leq x \leq 1$   $\Rightarrow$   $A^T x \leq b$   $\Leftrightarrow$   $x \geq 0$   $\Rightarrow$   $0 \leq x \leq 1$

Paired duality

$X \leftrightarrow X^*$  dual space  
 Linear Space — bilinear form  
 $\langle x, v \rangle$   
 $\langle \cdot, v \rangle : X \rightarrow \mathbb{R}$  for  $\forall v \in X^*$   
 linear functional

eg.  $x(t) \in L_2[0,1]$   $y(t) \in [0,1]$

$$\langle x, y \rangle = \int_{S=[0,1]} x(t) y(t) dt$$

Nov 27th, SNOWING Tuesday 2007

Prob. 1.2. Solution

$T, S_i, i=1,2$  polyhedral cones,  $S_1$  pointed,  $A_i \in m \times n_i$   $A_1 \neq 0$

Exactly one is true:

- I  $\sum_{i=1}^3 A_i x^i \in T$   $0 \neq x^1 \in S_1, x^2 \in S_2$
- II  $y \in -T^+, A_1^T y \in \text{int}(S_1^+), A_2^T y \in S_2^+, A_3^T y = 0$

proof: part 1: show both I, II cannot be feasible.

i.e. take  $y$ , and  $x^i$  and  $0 \geq y^T \sum_{i=1}^3 A_i x^i, y^T A_3 = 0$

$$0 \geq x_1^T A_1^T y + x_2^T A_2^T y$$

part 2: Suppose I is inconsistent.

Apply Farkas' lemma to show II consistent.

$\Rightarrow$  I:  $Ax = b, x \geq 0$  } exactly one holds  
 II:  $A^T y \geq 0, b^T y < 0$  }  
 I:  $A^T y = b, y \geq 0$   
 II:  $Ax \geq 0, b^T x < 0$

First,  $S_1$  pointed, so  $\text{int} S_1^+ \neq \emptyset$  (from Assign. Prob 1)

so let  $\phi \in \text{int} S_1^+$  and we get

$0 \neq x_1 \in S_1$  iff  $x^1 \in S_1$  and  $\langle \phi, x^1 \rangle > 0$

$T = \{ x \in \mathbb{R}^m \mid B^T x \geq 0 \}$   $T^+ = \{ B \lambda : \lambda \geq 0 \}$

$S_1 = \{ x \in \mathbb{R}^{n_1} \mid B_1^T x \geq 0 \}$   $S_1^+ = \{ B_1 \lambda : \lambda \geq 0 \}$

Therefore I is equiv. to

$$\begin{bmatrix} B_1^T & 0 & 0 \\ 0 & B_2^T & 0 \\ 0 & 0 & B_3^T \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \geq 0 \quad (-\phi^T, x^1) < 0$$

Hilroy

$$\begin{array}{l} \text{II } A^T \lambda = b \quad \lambda \geq 0 \\ \text{I } Ax \geq 0 \quad b^T x < 0 \end{array}$$

from Farkas, if this (I) is inconsistent, then the following is consistent.

$$\begin{bmatrix} B_1 & 0 & A_1^T B \\ 0 & B_2 & A_2^T B \\ 0 & 0 & A_3^T B \end{bmatrix} \lambda = \begin{bmatrix} -\rho \\ 0 \\ 0 \end{bmatrix} \quad \lambda = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{pmatrix} \geq 0$$

Set:  $y = B \lambda$

get  $A_1^T y = -\rho - B_1 \lambda_1 \in \text{int } S_1^T$

$A_2^T y = -B_2 \lambda_2 \in -S_2^T$  so set  $y \leftarrow -y$

$A_3^T y = 0$

opt. probs. e.g.

$f(x) = \begin{cases} f_0(x) & \text{if } x \text{ feasible} \\ +\infty & \text{otherwise} \end{cases}$

$\min_{x \in X} c, g \text{ feasible } f_i(x) \leq 0 \text{ and } x \in e$

parametrize  $F(x, u) : X \times U \rightarrow [-\infty, +\infty]$

with  $F(x, 0) = f(x)$

Use conjugation to get a saddle function  $k(x, y)$

conjugation is with a "paired space".

$X, X^*$

$\langle \cdot, \cdot \rangle$  bilinear form.

$\langle \cdot, \cdot \rangle$  contin. lin functional  $\forall v \in X^*$

$\langle x, \cdot \rangle$   $\forall x \in X$

$x \in \ell_2 \quad x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \\ \vdots \end{pmatrix} \quad \sum x_i^2 < \infty$   
 $\|x\| = \sqrt{\sum x_i^2}$

$x(t) \in \ell_2[a, b] \quad \int_a^b \|x(t)\|^2 dt < \infty$

$X^* \equiv X \quad \langle x, y \rangle = \int_a^b x(t)y(t) dt$

Recall  $f^*(v) = \sup_{x \in X} \{ \langle x, v \rangle - f(x) \}$   
 conjugate in "convex sense"

$$g^*(v) = \inf_{x \in X} \{ \langle x, v \rangle - g(x) \}$$

conjugate in "concave sense"

$f(x) = f_0(x)$  if  $x \in C$   
 $+\infty$  otherwise

Given opt. prob.  
 $\min_{x \in X} f(x)$

parametrize:  $F(x, u)$  with  $F(x, 0) \equiv f(x)$

Let  $h_x(u) \equiv -F(x, u)$

set  $K(x, y) = h_x^*(y)$  in CONCAVE sense

So  $K(x, y) = \inf_{u \in U} \{ F(x, u) + \langle u, y \rangle \}$

called Lagrange/saddle func

For each  $x$ ,  $K(x, y)$  is conjugate in concave sense of  $-F(x, y)$

we can take 2<sup>nd</sup> conjugate of  $(-F)(x, \cdot)$

$$F(x, u) = \sup_{y \in Y} \{ K(x, y) - \langle u, y \rangle \}$$

if  $F(x, \cdot)$  is closed convex.

$$K(x, y) = \begin{cases} f_0(x) + \sum_{i=1}^m y_i f_i(x) & \text{if } x \in C, y \geq 0 \\ -\infty & \text{if } x \in C, y \text{ not } \geq 0 \\ +\infty & \text{otherwise} \end{cases}$$

Th<sup>m</sup> 6.  $K(x, y)$  is closed concave in  $y$  for each  $x$ ,  
 $F(x, u)$  convex in  $u$

$x^*$   $f(x) = \sup_{y \in Y} K(x, y)$

(yields primal problem  $\min_{x \in X} f(x)$ )

conversely, we can start with  $f, F$  and get  $K$ .

So set  $g(y) = \inf_{x \in X} K(x, y)$

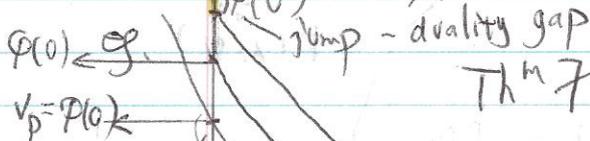
and get a dual problem

$$V_D = \max_{y \in Y} g(y) \quad \text{for primal} \quad V_P = \min_{x \in X} f(x)$$

Define Optimal value function.

$$\phi(u) = \inf_{x \in X} F(x, u)$$

Recall  $F$  convex in  $u \Rightarrow \phi$  convex in  $u$



Thm 7

$g$  is closed concave

$$g = (-\phi)^*$$

(concave sense)

Thus,

$$V_D = \sup_{y \in Y} g(y) = (cl(\text{co } \phi))(0) \equiv (-\phi)^{**}(0)$$

if  $F(x, u)$  convex in  $(x, u)$  we have

$$-g^* = cl \phi$$

(except when  $\phi$  is nowhere finite)

ie.  $V_D = \lim_{u \rightarrow 0} \inf_{u \in U} \phi(u)$

Proof:

$$g(y) = \inf_{x \in X} \inf_{u \in U} \{ F(x, u) + \langle u, y \rangle \}$$

↓  
ϕ(u) when swapping

$$\equiv (-\phi)^*(y)$$

So  $V_P = V_D$ , saddle value, equiv to

$$\phi(0) = (cl(\text{co } \phi))(0)$$

↑ convex hull

This works well when  $\phi$  is CONVEX!

true for convex program.

when else? Hot Research Area.

eg. generalized trust-region subproblem

$$\begin{aligned} \min & f_0(x) \\ \text{st.} & f_1(x) \leq 0 \end{aligned}$$

quadratic

$$\frac{1}{2}x^T Q_1 x + b_1^T x + c_1$$

Assume Slater CQ.

but No convexity ( $Q_1$  can be indefinite)

Nov 29th, Thursday Last Class for convex analysis & Opt.

Problem:  $f: E \rightarrow \mathbb{R}$  proper, convex

$$A: Y \rightarrow E$$

lin. transformation

Show  $\partial(f \circ A)(y) \supseteq A^{adj} \partial f(Ay) \quad \forall Ay \in \text{dom}(f)$

= if  $\text{int dom}(f) \cap A(y) \neq \emptyset$

suppose that  $Ay \in \text{dom}(f)$  and  $\varphi \in \partial f(Ay)$

First show  $A\varphi \in \partial(f \circ A)(y)$

We have

$$\langle \varphi, w - Ay \rangle \leq f(w) - f(Ay) \quad \forall w$$

$$\Rightarrow \langle \varphi, A\bar{y} - Ay \rangle \leq f(A\bar{y}) - f(Ay) \quad \forall A\bar{y}$$

Then  $\langle A^{adj} \varphi, \bar{y} - y \rangle \leq (f \circ A)(\bar{y}) - (f \circ A)(y)$

Now suppose

$$\bar{w} = A\bar{y} \in \text{int}(\text{dom}(f)) \cap A(y) \quad (\neq \emptyset)$$

Define  $F(y) := (f \circ A)(y)$

suppose, on the contrary that

$$\exists \varphi \in \partial F(\bar{y}) \text{ but } \varphi \notin A^{adj} \partial f(\bar{w})$$

$f$  cont. at  $\bar{w}$ , so  $\partial f(\bar{w})$  is convex, compact.

$\Rightarrow A^{adj} \partial f(\bar{w})$  is also

$\Rightarrow \exists d \neq 0$  by compact./closure

$$\alpha = \langle d, \varphi \rangle > \langle d, \psi \rangle \quad \forall \psi \in A^{adj} \partial f(\bar{w})$$

$$\Rightarrow \alpha > \max_{\varphi \in A^{adj} \partial f(\bar{w})} \langle d, \varphi \rangle = \max_{\varphi \in \partial f(\bar{w})} \langle d, A^{adj} \varphi \rangle$$

$$= f'(A\bar{y}, Ad) = (f \circ A)'(\bar{y}; d)$$

$$= F'(\bar{y}; d)$$

But  $F'(\bar{y}; d) = \max_{\delta \in \partial F(\bar{y})} \langle d, \delta \rangle$

$\geq \langle d, \phi \rangle = d$  contrad.

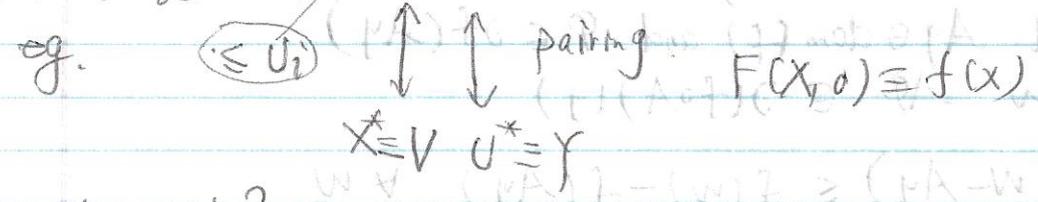
eg. NLP

(P)  $\inf f_0(x)$  st.  $f_i(x) \leq 0 \quad x \in C, i=1, \dots, m$

(P\*)  $f(x) := \begin{cases} f_0(x) & \text{if } x \in C, f_i(x) \leq 0, i=1, \dots, m \\ +\infty & \text{elsewhere} \end{cases}$

parametrize  $x \in X$

$$F: X \times U \rightarrow [-\infty, +\infty]$$



Construct dual?

use  $K(x, y) \equiv (-F)^*(x, y) = L(x, y)$  std. Lagrangian  
conjugate in concave sense

eg. get  $K(x, y) = \begin{cases} f_0(x) + \sum_{i=1}^m y_i f_i(x) & \text{if } x \in C, y \geq 0 \\ -\infty & \text{otherwise } (x \notin C) \end{cases}$

Now define dual functional

$$g(y) := \inf_{x \in X} K(x, y) \quad (D) \quad \max_{y \in Y} g(y)$$

(Note  $f(x) = \sup_{y \in Y} K(x, y)$ )

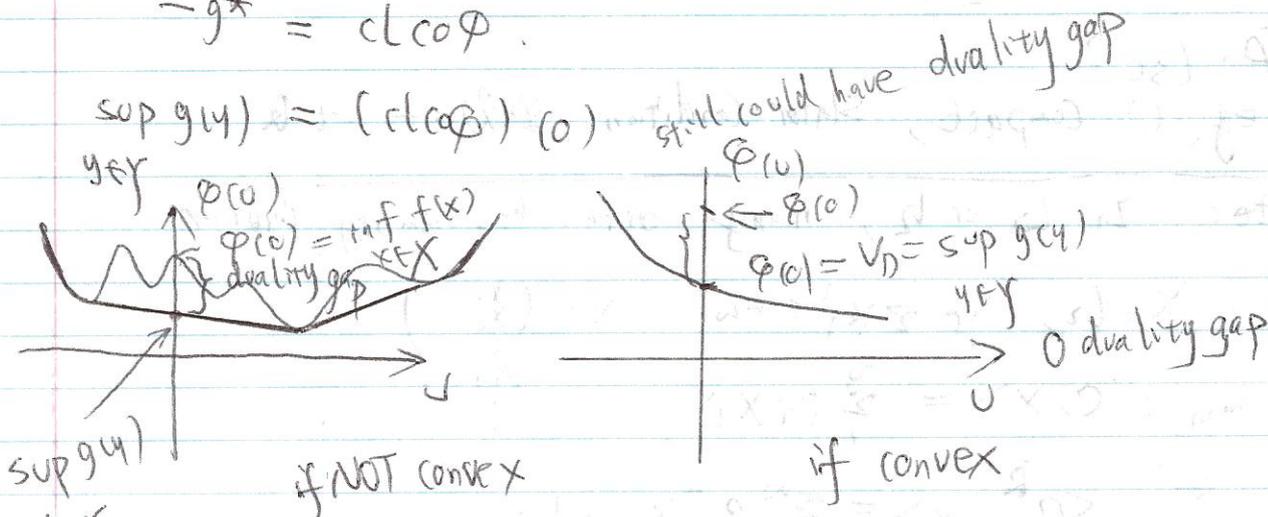
recover  $f(x)$  in this way

eg. Apply to LP  
 i.e.  $f(x) = \langle q, x \rangle + \sum_{i=1}^m c_i x_i$

as expected, we get the dual LP.

Also,  $\phi(u) = \inf_{x \in X} f(x, u)$  (value  $f^*$ )

get  $g \equiv (-\phi)^*$  (concave case)  
 $-g^* = \text{clco}\phi$



eg. NLP (generalized)

$Q$  convex cone, (closed)

$\min f_0(x)$  s.t.  $x \in C, \Phi(x) \in Q$   
 $\uparrow$  convex set  $\uparrow$   $Q$

$F(x, u) = \begin{cases} f_0(x) & \text{if } x \in C, \Phi(x) \in Q \\ +\infty & \text{elsewhere} \end{cases}$

$K(x, y) = \inf_{u \in U} \{ F(x, u) + \langle u, y \rangle \}$

Case 1:  $x \notin C$  then  $K(x, y) = +\infty$

Case 2:  $x \in C$

part a)  $y \in \mathbb{R}^+$ , then  $\inf_{u \in U} \{ f_0(x) + \langle u, y \rangle : \Phi(x) \in Q \}$   
 a "Lagrangian"

Then  $K(x, y) = f_0(x) + \langle y, \Phi(x) \rangle$

part b)  $y \notin Q^+$  then  $\exists \bar{U} \in \langle \bar{U}, y \rangle < 0$

$$U = \Phi(x) + t(\bar{U}) \quad t \nearrow +\infty$$

$$K(x, y) = -\infty$$

When do we get a 0 duality gap.

where  $g(y) = \inf_{x \in X} K(x, y)$   
 $= -\infty$

$\Phi = \text{lsc?}$

eg.  $C$  compact, Slater condition  $\Phi(x) \in -\text{int } Q$

Note: In  $L_2$  or  $l_2$ , many, orth. has empty interior.

eg.  $X \in l_2$ ,  $\forall \epsilon \exists x_1^2 \in +\infty$   $x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \\ \vdots \end{pmatrix}$

$$\min \langle c, x \rangle = \sum_1^{\infty} c_i x_i$$

st.  $\langle a, x \rangle = \sum_1^{\infty} a_i x_i = b_i$   
 $x_i \geq 0 \quad \forall i = 1, 2, \dots$

and  $\exists \bar{x} > 0$  ie,  $x_i^2 > 0 \quad \forall i$ .

But, given any  $\epsilon > 0$ , let  $\bar{x}$  satisfy.

$$|x_{i_0}^1| < \frac{\epsilon}{2}$$

set  $\bar{x}_1 = x_{i_0}^1$ , let  $x_{i_0}^2 = -\frac{\epsilon}{4}$

then  $\|x - \bar{x}\| = \sqrt{\sum_1^{\infty} (x_i^2 - \bar{x}_i)^2} = \sqrt{x_{i_0}^2 - \epsilon} < \epsilon$

$\Rightarrow X \notin \text{int } \mathbb{R}_+ \downarrow$

$$\text{int } \mathbb{R}_+^{\infty} = \emptyset.$$