Lec21-mwf

Saturday, February 07, 2009 10:44 PM

Bring hand-out for value function.

Supporting hyperplanes

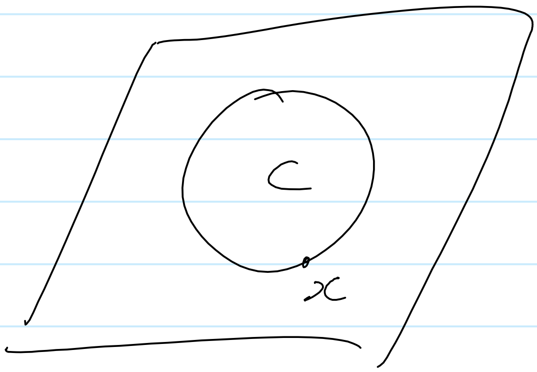
Thursday, February 05, 2009 9:52 AM

- An immediate applications of the separation theorem is regarding supporting hyperplanes.

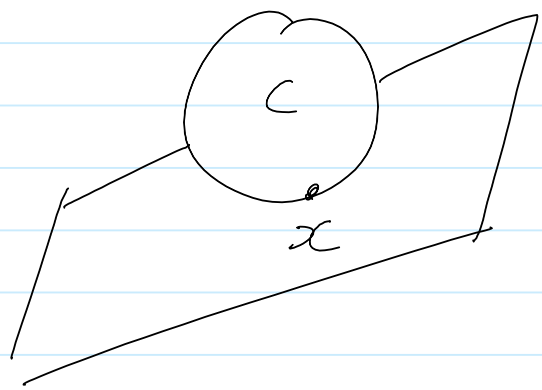
- (Simplified version of Thm 11.6 in Rockafella)

  Let C be a convex set. For every point x at the relative boundary of C, there exists a non-trivial supporting hyperplane to $cl\, C$ containing the point x.

- A non-trivial supporting hyperplane to C is one that does not contain C itself.



trivial supporting hyperplane.



non-trivial supporting hyperplane

-

Q Why is it true?

A $\mathrm{ri}\{x\} = x$ is disjoint from $\mathrm{ri}\, C$

$\Rightarrow$ There exists a hyperplane H that properly separates $\{x\}$ & C

$\Rightarrow$ Easy to check that $x \in H$, and H does not contain C.

Convex functions and affine functions

Friday, February 13, 2009 3:45 PM

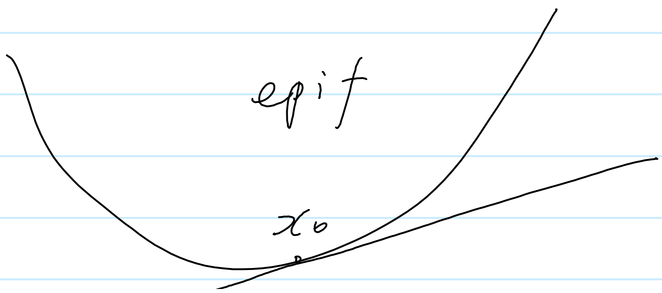
- Let f be a convex function, then epi f is a convex set
- On the other hand, any affine function defines a hyper-plane.
- We can then use the result on supporting hyperplane to study the relationship between convex functions & affine functions.
- Boyd P119. Exercise 3.28

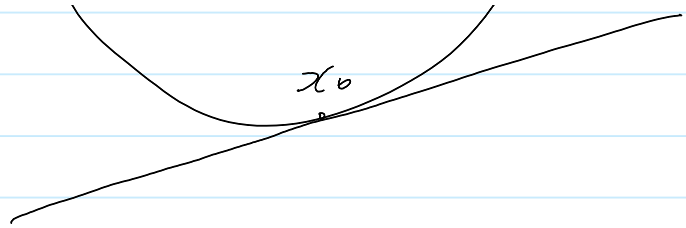
Result:

- For any point $x_0 \in \text{int dom} f$, we can find an affine function $g(x)$ such that

$$f(x) \geq g(x) \quad \forall x$$
$$f(x_0) = g(x_0)$$

- The function $g(x)$ defines a supporting hyperplane of epi f at x_0.



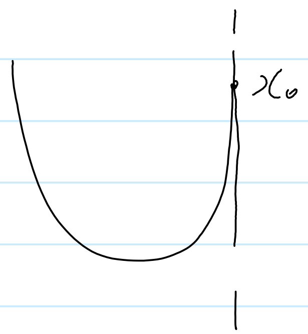


Proof Sketch:

- For any point $(x_0, f(x_0))$ on the boundary of epi f, we can find a supporting hyperplane.

Ⓠ Does such a hyperplane always correspond to an affine function?

Ⓐ Need x_0 be in the interior of dom f



No corresponding affine function

- If $x_0 \in$ int dom f, f is well-defined in a neighborhood of x_0,

- Then the situation in the above figure will not occur.

Subgradients

Sunday, February 15, 2009 7:23 AM

- the vector h is a subgradient of the function f at x_0 if

$$f(x) \geq \underbrace{f(x_0) + h'(x - x_0)}_{\text{affine } g(x):\, g(x_0) = f(x_0)} \quad \forall x$$

- The set of all subgradients of f at x_0 is called subdifferentials of f at x_0.
  - denoted by $\partial f(x_0)$

Through the earlier discussion, we can conclude that:

- A convex function f has non-empty sub-differentials $\partial f(x)$ at any $x \in \mathrm{int}(\mathrm{dom} f)$

- If f is differentiable, then

$$\partial f = \{ f'(x) \}$$

  - contains only one element.

Convex sets and hyperplanes - skip

Thursday, February 05, 2009 9:58 AM

- We know that any intersection of half spaces & hyperplanes is a convex set.
- In particular, a polyhedra is the intersection of a finite number of half-spaces & hyperplanes.

<u>The converse</u>

- A closed convex set C is the intersections of the closed half-spaces that contain it.
- Let $\{S_\alpha\}$ be the set of closed half-spaces that contain C

then $C \subset \cap S_\alpha$

- Let us now show that $C \supset \cap S_\alpha$.
- Assume in the contrary that there exists a point x such that

$$x \in \cap S_\alpha \text{ but } x \notin C$$

$\Rightarrow$ There exists a closed half-space containing C but does not contain x (strong separation)

$\Rightarrow$ $x \notin \bigcap S_\alpha$ (contradiction)

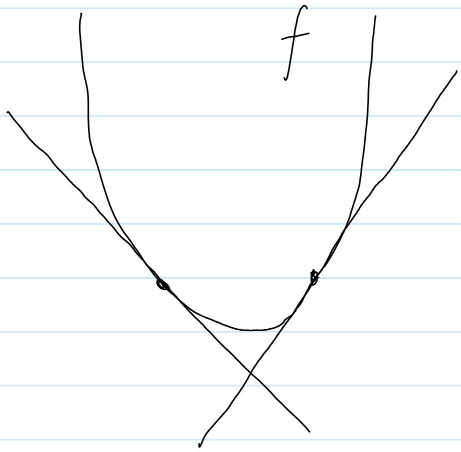
Sup of affine functions - skip

Sunday, February 15, 2009 7:28 AM

- A function f is closed if its epigraph is closed

- If f is closed and convex, then f equals to the pointwise supremum of all affine functions that are global underestimators of f.

$$f = \sup \{ g(x) \mid g \text{ affine}, \ g(z) \le f(z), \ \forall z \}$$

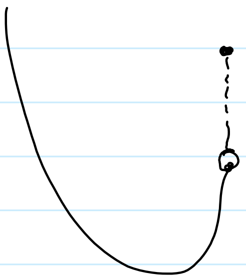


- Proof: use the fact that any closed convex set is the intersection of all half-spaces that contain it.

  - epi f is closed

(25)

- Is there a convex function that is not closed?



Proof of the Separation Theorem - skip

Thursday, February 05, 2009 10:20 AM

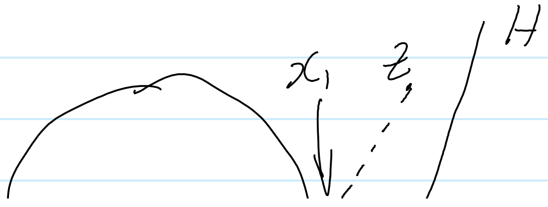
- The easiest case. (Boyd P46)
- More complete cases are shown in handout. (separation.pdf)

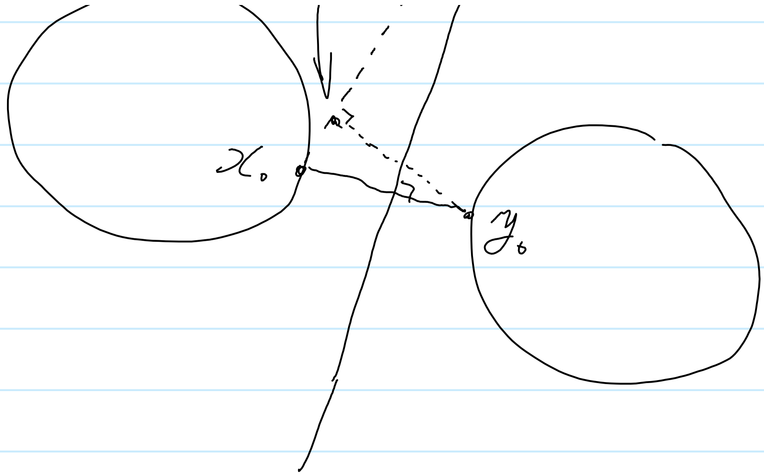
- Let C & D be two closed and bounded convex sets. Assume that C & D are disjoint from each other. Then there exists a hyperplane that separates C & D <u>strongly</u>

<u>Sketch of proof:</u>

- The function $\|x-y\|$, $x \in C$, $y \in D$ is continuous.
- From real analysis, any continuous function must attain the minimum on a closed and bounded set.
- Hence, there exist $x_0 \in C$, $y_0 \in D$ such that

$$\|x_0 - y_0\| \le \|x-y\|, \quad x \in C, y \in D$$





- $\|x_0 - y_0\|$ must be greater than 0 since $C \cap D = \emptyset$.

- Choose H to be the hyperplane that bisects the line segment $x_0 y_0$. Let

$$\varepsilon = \frac{\|x_0 - y_0\|}{4} > 0$$

- We can show that there is no point in C that is within a distance ε to H.

  - If there is such a point $z \in C$, then $\angle z x_0 y_0 < \pi/2$

  - There must then exist a point $x_1 \in C$ (by convexity), that is closer to y_0 than x_0 (a contradiction).

- Other cases see handout. (separation.pdf)

35

Back to duality

Sunday, February 15, 2009 7:52 AM

- use handout (valuefunction.pdf).
- Primal problem

$$\begin{array}{ll} \min & f_0(x) \\ \text{sub to} & f_i(x) \le 0 \\ & h_i(x) = 0 \end{array}$$

- Lagrangian

$$L(x, \lambda, \nu) = f_0(x) + \sum_i \lambda_i f_i(x) + \sum_i \nu_i h_i(x)$$

$$g(\lambda, \nu) = \min_x L(x, \lambda, \nu)$$

- Dual problem

$$\begin{array}{ll} \max & g(\lambda, \nu) \\ \text{sub to} & \lambda \ge 0 \end{array}$$

- Let p^*, d^* denote the optimal value of the primal/dual problems, respectively.

Weak duality:

- $d^* \le p^*$

Strong duality

- There exists an $x \in \operatorname{ri} D$ such that

$$f_i(x) < 0 \quad \forall i$$
$$Ax = b \quad \leftarrow \text{equality constraints } h_i(\cdot)$$

Main Theorem: If the slater condition holds, and in addition rank A is equal to the number of equality constraints, then the duality gap is zero

---

Proof (Sketch).

- In light of weak duality, we only need to show that there exists $\lambda^* \geq 0$, ν^* such that $g(\lambda^*, \nu^*) = p^*$.

- To find these λ^*, ν^*, we use the Separation Theorem

- Define the value function

lower boundary of A

c plays the same role as u.

$$v(\vec{c}, \vec{b}) = \min f_0(x)$$
$$\text{sub to } f_i(x) \leq c^i, \; i=1,\cdots,M$$
$$h_i(x) = b^i, \; i=1,\cdots,N$$

Claim 1:

- $v(\vec{c}, \vec{b})$ is a convex function of $\vec{c}, \vec{b}$

- exercise

Claim 2:

- If Slater condition holds, and rank A = # of equality constraints, then $v(\vec{c}, \vec{b})$ is well-defined (i.e, not equal to $+\infty$) at a neighborhood of the origin.

  - If $\vec{c}$ is changed from zero a little bit, we can still find x such that
  $$f_i(x) < c^i, \; i=1,2,\cdots,M$$

- If $\vec{b}$ is changed from zero a little bit, we can still find x such that $h_i(x) = b^i$, $i = 1, \cdots, N$

- Similarly if both $\vec{c}$ & $\vec{b}$ change from zero a little bit.

- Further, we can show that $v(\vec{c}, \vec{b}) > -\infty$ (see handout).

Claim 3 :

- $v(\vec{0}, \vec{0}) < +\infty$

- since there exists some feasible point.

- By claim 2, $v(\vec{c}, \vec{b})$ must also take real values in a neighborhood of $(\vec{0}, \vec{0})$.

- Hence, from the separation theorem, there is a subgradient $(-\vec{\lambda}, -\vec{\nu})$ at $(\vec{0}, \vec{0})$.

Claim 4 :

- $\vec{\lambda}, \vec{\nu}$ is the Lagrange multiplier such that

$$g(\vec{\lambda}, \vec{\nu}) = p^* = f_0(x^*) \quad \text{and} \quad \vec{\lambda} \geq 0.$$

Proof of Claim 4:

① To show $\vec{\lambda} \geq 0$, note that by definition of subgradients

$$v(\vec{c}, \vec{b}) \geq v(0, 0) + (-\vec{\lambda}) \cdot \vec{c} + (-\vec{\nu}) \cdot \vec{b} \quad \forall \vec{c}, \vec{b}.$$

- For any $\vec{c} \geq 0$, $\vec{b} = 0$, $v(\vec{c}, \vec{b}) \leq v(\vec{0}, \vec{0})$

$$\Rightarrow \quad -\vec{\lambda}\vec{c} \le 0 \quad \forall \vec{c} \ge 0$$
$$\Rightarrow \quad \vec{\lambda} \ge 0$$

② To show $g(\vec{\lambda}, \vec{\nu}) = p^* = f_0(x^*)$

Note that for any $\vec{x}$, let

$$\vec{c}(x) = [f_1(x), f_2(x), \cdots, f_M(x)]$$
$$\vec{b}(x) = [h_1(x), h_2(x), \cdots, h_N(x)]$$

- Then by the definition of subgradient $(-\vec{\lambda}, -\vec{\nu})$,

$$\nu(\vec{c}, \vec{b}) \ge \nu(0, 0) + (-\vec{\lambda}) \cdot \vec{c} + (-\vec{\nu}) \cdot b$$

- But $f_0(x) \ge \nu(\vec{c}(x), \vec{b}(x))$ since x is a feasible point of the problem $\nu(\vec{c}(x), \vec{b}(x))$

Also $\nu(0, 0) = f_0(x^*)$

$$\Rightarrow \quad f_0(x) \ge f_0(x^*) - \vec{\lambda} \cdot \vec{c}(x) - \vec{\nu}\vec{b}(x)$$

This is true for all x. Rearrange:

$$\Rightarrow \quad f_0(x) + \vec{\lambda} \cdot \vec{c}(x) + \vec{\nu} \cdot \vec{b}(x) \ge f_0(x^*) \quad \forall x$$

$$\Rightarrow \quad L(x, \vec{\lambda}, \vec{\nu}) \ge f_0(x^*) \quad \forall x$$

$$\Rightarrow \quad g(\vec{\lambda}, \vec{\nu}) = \min_x L(x, \vec{\lambda}, \vec{\nu}) \ge f_0(x^*)$$

However, we know that

$$g(\vec{\lambda}, \vec{\nu}) \le g(\lambda^*, \nu^*) \le f_0(x^*)$$

$$\Rightarrow \quad g(\vec{\lambda}, \vec{\nu}) = g(\lambda^*, \nu^*) = f_0(x^*)$$

This then proves strong duality. An immediate consequence is the complementary slackness condition.

JJ

Graphical

Sunday, February 20, 2011 9:58 PM