Lec8-mwf

Monday, January 19, 2009 10:00 PM

Vector function composition

Saturday, January 17, 2009 2:41 PM

If $g: R^n \to R$, $h: R \to R$, take a restriction to a line

$$f(\lambda) = h(g(x+\lambda v))$$

g convex/concave in x

$\Leftrightarrow g(x+\lambda v)$ convex/concave in λ

$\Rightarrow$ Exactly the same rule apply!

$$f(x) = h(g(x)) \quad : \quad g = [g_1 \cdots g_K]^T$$

Assume without loss of generality that $x \in R$ (e.g. restricted the function to a line).

$h: R^K \to R \qquad h(y_1, \cdots, y_K)$

$g_i: R \to R \qquad i = 1, 2, \cdots, K$

$$g = [g_1, g_2, \cdots, g_K]^T$$

$$f(x) = h(g_1(x), g_2(x), \cdots g_K(x))$$

Then

$$f'(x) = \nabla h(g(x))^T \cdot g'(x) \qquad \in R$$

$$f''(x) = g'(x)^T \nabla^2 h(g(x)) \cdot g'(x) + \nabla h(g(x))^T g''(x) \qquad \in R$$

where

$$\nabla h = \begin{bmatrix} \frac{\partial h}{\partial y_1} \\ \vdots \\ \frac{\partial h}{\partial y_K} \end{bmatrix} \qquad g' = \begin{bmatrix} g_1' \\ \vdots \\ g_K' \end{bmatrix} \qquad g'' = \begin{bmatrix} g_1'' \\ \vdots \\ g_K'' \end{bmatrix}$$

where
$$\nabla h = \begin{bmatrix} \frac{\partial h}{\partial y_1} \\ \vdots \\ \frac{\partial h}{\partial y_k} \end{bmatrix} \qquad g' = \begin{bmatrix} g_1' \\ \vdots \\ g_k' \end{bmatrix} \qquad g'' = \begin{bmatrix} g_1'' \\ \vdots \\ g_k'' \end{bmatrix}$$

$$\nabla^2 h = \begin{bmatrix} \frac{\partial^2 h}{\partial y_i \partial y_j} \end{bmatrix}_{k \times k}$$
positive or negative semi-definite

Assume g_i is convex, $g_i''(x) \geq 0$

If $\nabla^2 h(\cdot) \geq 0$ then

$$g'(x)^T \nabla^2 h(g(x)) \, g'(x) \geq 0$$

If in addition $\nabla h(g(x)) \geq 0$ for each element, then

$$\nabla h(g(x))^T g''(x) \geq 0$$

We then obtain the following rule

- The convexity of f must follow that of h
- If h & g are of the same type, need h increasing in every argument
- If h & g are of different types, need h decreasing in every argument.

Finally, if $\text{dom } h \neq \mathbb{R}^k$, we will need to replace h' by extend-value function $\tilde{h}$.

Example:

ⓐ $h(z) = z_{[1]} + \cdots + z_{[r]}$, the sum of the r largest component. skip

- $h(z)$ is convex

$g_1(x), \cdots, g_k(x)$ are convex.

$\Rightarrow h(g_1(x), \cdots, g_k(x))$ is convex because $h(z)$ is non-decreasing in each argument.

ⓑ $h(z) = \log\left(\sum_{i=1}^{k} e^{z_i}\right)$

$g_1(x), \cdots g_k(x)$ are convex

$\Rightarrow h(g_1, \cdots, g_k) = \log\left(\sum_{i=1}^{k} e^{g_i(x)}\right)$ is convex.

ⓒ $f(x,u,v) = \sqrt{uv - x^T x}$ $\quad x \in R^n, \; u, v \in R$

on $u>0, v>0, uv > x^T x$

See Boyd Ex 3.22(b), p 117

$$f(x,u,v) = \sqrt{u\left(v - \frac{x^T x}{u}\right)}$$
$$= \sqrt{g_1 \cdot g_2}$$

$g_1(x,u,v) = u, \quad g_2(x,u,v) = v - \frac{x^T x}{u}$

- $\sqrt{g_1 g_2}$ is concave in g_1, g_2

- $g_2(x,u,v)$ is concave in (x,u,v)
- $g_1(u)=u$ is concave
- $\sqrt{g_1 g_2}$ extended to R^2 is non-decreasing in g_1 & g_2.

(45)

Summary

Monday, January 19, 2009 10:00 PM

- Convex functions

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

- First order condition

$$f(y) \ge f(x) + f'(x)(y-x)$$

Second order condition

$$f''(x) \ge 0$$

- Operations
    - Non-negative weighted sum
    - Affine change of variable
    - Pointwise maximum
    - Pointwise minimum (more demanding)
    - Composition
    - Perspective

Jensen's inequality - skip

Monday, January 12, 2009 4:46 PM

If a function f is convex, then

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

for any $0 \le \theta \le 1$.

Easily extended to finite convex combinations

$$f(\theta_1 x_1 + \cdots + \theta_k x_k)$$
$$\le \theta_1 f(x_1) + \cdots + \theta_k f(x_k)$$

for all $\theta_1 + \cdots + \theta_k = 1$, $0 \le \theta_1, \ldots, \theta_k \le 1$

Can also be extended to infinite convex combinations

Let $\theta_1, \cdots, \theta_k, \cdots \ge 0$ satisfies

$$\sum_{k=1}^{+\infty} \theta_k = 1$$

Let $x_1, \cdots x_k, \cdots \in \operatorname{dom} f$, which is convex and the limit $\sum_{k=1}^{+\infty} \theta_k x_k$ exists

Then ① $\sum_{k=1}^{+\infty} \theta_k x_k \in \operatorname{dom} f$

② $f\left(\sum_{k=1}^{+\infty} \theta_k x_k\right) \le \sum_{k=1}^{+\infty} \theta_k f(x_k)$.

The proof of part 1 will need the Separation Theorem (to be discussed later)

The proof of part 2 uses the first-order condition (assuming f is differentiable), or use subgradients (to be discussed later).

We can also extend to integrals

- Suppose $p(\cdot)$ is a density on $S \subseteq \operatorname{dom} f$

  $p(x) \geq 0$ and $\int_S p(x)\, dx = 1$

- If f is convex, then

  $$f\left(\int_S p(x) \cdot x \, dx\right) \leq \int_S f(x)\, p(x)\, dx$$

Same for expectations

$$f(E X) \leq E f(X).$$

30

Convex optimization problems

Saturday, January 17, 2009 3:22 PM

- A convex optimization problem minimizes a convex function over a convex set

$$\min \; f_0(x) \quad \leftarrow \text{convex func.}$$
$$\text{sub to} \; x \in C \quad \leftarrow \text{convex set.}$$

- However, for the most part we will use the following "standard form":

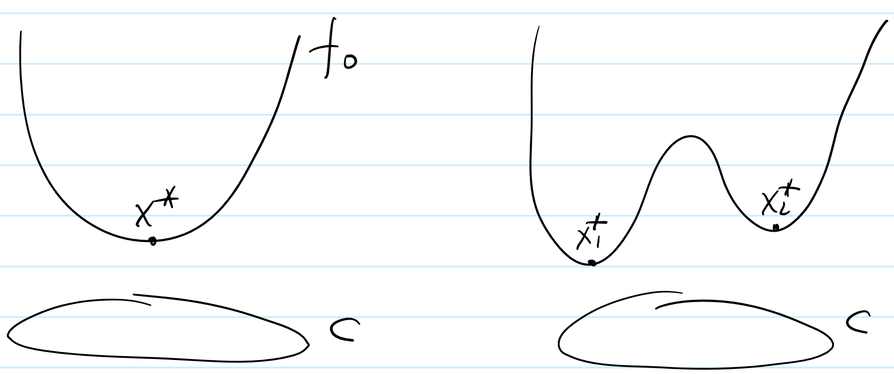
Standard form:

$$\begin{aligned} \min \quad & f_0(x) \\ \text{sub to} \quad & f_i(x) \le 0, \quad i = 1, \cdots, m \\ & h_i(x) = 0, \quad i = 1, \cdots, p \end{aligned}$$

where f_0, f_i are convex functions
h_i are linear functions

- x: variable to optimize
  $f_0(x)$: objective
  $f_i(x) \le 0$: inequality constraints
  $h_i(x) = 0$: equality constraints

- A key feature is that any local optimum must be a global optimum



- What about concave functions?
  - min a concave func won't work

  - max a concave func f_0 is the same as min the convex func $-f_0$
  - $f_1(x) \geq 0$ for a concave func is the same as $-f_1(x) \leq 0$ for the convex func. $-f_1(x)$.
- Why $h_i(x)$ must be linear?
- Some properties of convex opt. problems hold for the general form, but some required the stricter form.

- Domain of the problem:

$$D = \bigcap_{i=0}^{m} \mathrm{dom} f_i \bigcap_{i=1}^{p} \mathrm{dom}\, h_i$$

- A point x is feasible if $x \in D$ and x satisfies the constraints $f_i(x) \leq 0$ and $h_i(x) = 0$.

- Optimal value

$$p^* = \inf \{ f_0(x) \mid f_i(x) \leq 0,\ i=1,\cdots,m,\ h_i(x) = 0,\ i=1,\cdots,p \}$$

we allow
- $p^* = +\infty$ (if no point x is feasible)
- $p^* = -\infty$ (if there exists a sequence of

feasible x_k, such that $f(x_k) \to -\infty$).

— Optimal point:

If the optimal value p^* is attained at feasible point x^*

$\Rightarrow$ x^* is an optimal point.

— We have taken the inequality constraints as "$\leq$". Although some results also apply for "$<$".

The "$\leq$" part often implies that the feasible set is closed. Hence, the optimal value is always attainable if $f_0(x)$ is continuous.

(10)

Continuity of convex functions

Saturday, January 17, 2009 3:15 PM

Skip. Related to directional derivative

The simple property that

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

is in fact very strong and have important implications both geometrically & algebraically.

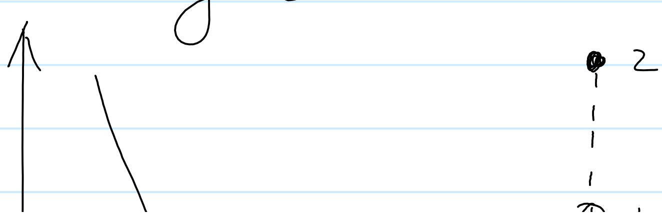
For example:

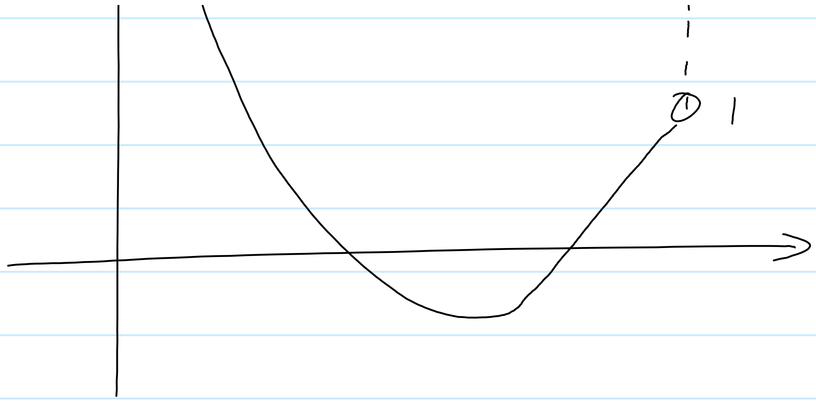
Continuity:

Let f be a convex function, then f is continuous on the interior of its domain

Discontinuity cannot occur in the interior of the domain.

Discontinuity can occur at the boundary.





(65)

More on this when we discuss separation Theorems.