Lec5-mwf

Monday, January 19, 2009 5:29 PM

Second-order condition

Monday, January 12, 2009 2:53 PM

Assume that f is twice-differentiable, i.e., its Hessian $\nabla^2 f$ exists for each point in $\mathrm{dom}\, f$

Then f is convex if & only if $\mathrm{dom}\, f$ is convex and

$$\nabla^2 f \succeq 0 \quad \text{for all } x$$

positive semi-definite.

f is strictly convex if $\mathrm{dom}\, f$ is convex and

$$\nabla^2 f \succ 0 \quad \text{for all } x.$$

— not necessary.

This is usually the most convenient criterion.

For functions on R, need $f''(x) \geq 0$.

For functions on R^n, reduce to 1-dim

$$g(t) = f(x + tv) \quad \text{for a given vector } v.$$

$$g'(t) = \left[\nabla f(x + tv)\right]^T \cdot v$$

$$g''(t) = v^T \nabla^2 f(x + tv) \cdot v.$$

Need $g''(t) \geq 0$ for all v

$\Leftrightarrow$ $\nabla^2 f(x)$ to be positive-semidefinite

(60)

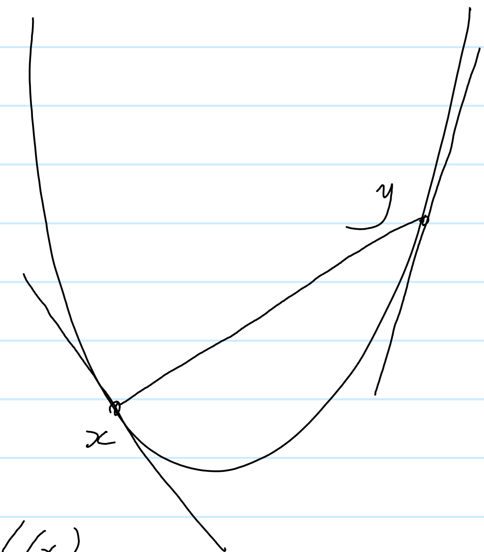
Proof of second-order condition

Monday, January 12, 2009 2:57 PM

Sketch: Focus on one-dimension case $f : R \to R$.

"$\Rightarrow$" direction: If f is convex, then $f''(x) \geq 0$.

Using first-order condition



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

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

$$\Rightarrow [f'(x) - f'(y)][x-y] \geq 0.$$

When $y > x \Rightarrow f'(y) \geq f'(x)$

$$\Rightarrow f''(x) \geq 0$$

"$\Leftarrow$" direction: If $f''(x) \geq 0$, then f is convex.

Using the mean-value theorem

$$f(y) = f(x) + f'(x)(y-x) + \frac{1}{2} f''(x+t(y-x))(y-x)^2$$

for some $t \in [0,1]$. Since $f''(x+t(y-x)) \geq 0$

$$\Rightarrow f(y) \geq f(x) + f'(x)(y-x)$$

Results follow from first-order condition

20

Operations

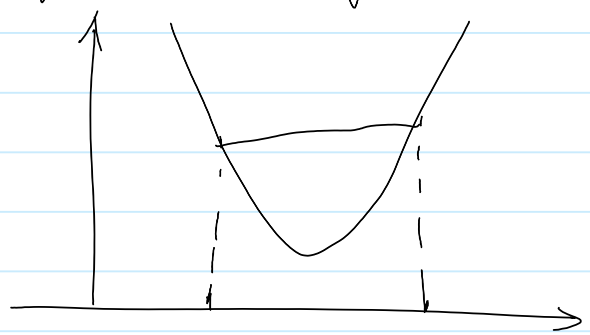
Wednesday, January 14, 2009 10:10 AM

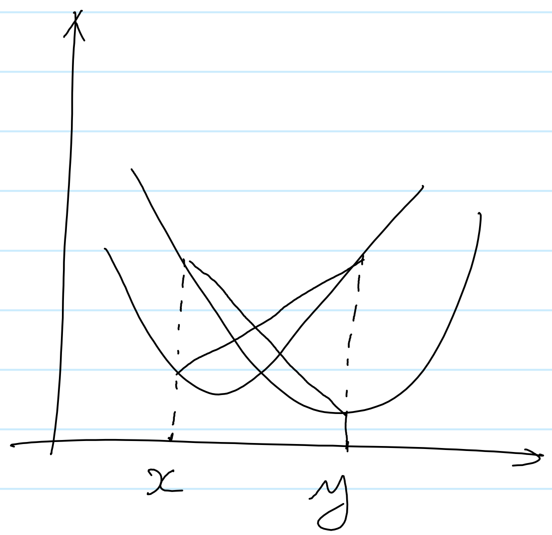
Operations that preserve convexity

① Non-negative weighted sums

$f_1, \cdots, f_m$ convex
$w_1, \cdots, w_m \geq 0$

$\Rightarrow f = w_1 f_1 + \cdots + w_m f_m$ is convex





Proof by definition.

- Can be extended to infinite sums & integrals

$f(x,y)$ convex in x for each y
$w(y) \geq 0$

Then $g(x) = \int_A w(y) f(x,y)\, dy$ is convex in x.

- Similarly, non-negative weighted sum of concave functions is concave.

Q What if some of the weights are negative?

(2) Affine mapping of the argument

Suppose $f : R^n \to R$, $A \in R^{n \times m}$, and $b \in R^n$

Define $g : R^m \to R$ by

$$g(x) \triangleq f(Ax+b)$$

with $\operatorname{dom} g = \{x \mid Ax+b \in \operatorname{dom} f\}$

Then if f is convex, so is g

if f is concave, so is g.

Proof by definition.

Examples

– $\log\left(\sum_{i=1}^{n} e^{x_i}\right)$ is convex, then

$$\log\left(e^{x_1+2x_2} + e^{2x_2+x_3} + e^{3x_1+x_3}\right)$$

is convex

– x^2 is convex, then

$$2(x_1+3x_3)^2 + (x_2+x_4)^2 \text{ is convex.}$$

Q If $f(x)$ is convex
is $f(-x)$ convex?
is $-f(x)$ convex?
(Draw a figure)

Using the above two properties, whenever we check for convexity, we can ignore the affine change of variables and non-negative weights, and focus on the simplest function form.

Examples

Monday, January 12, 2009 3:49 PM

To use convex optimization, it is very important to be able to quickly identify / convert to convex functions.

A bit exagerating: but for some problems, the key to success is to find (identify) convex functions!

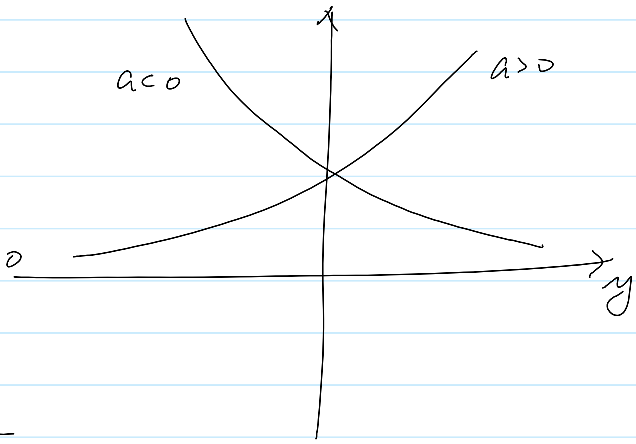
Here, we will provide some commonly-used convex functions with applications.

Keep a list of examples.

(1) exponent e^{ax}

$$\frac{df}{dx} = a e^{ax}$$

$$\frac{d^2f}{dx^2} = a^2 e^{ax} \geq 0$$



negative log $-\log x$

$$\frac{df}{dx} = -\frac{1}{x}$$

$$\frac{d^2f}{dx^2} = \frac{1}{x^2} \geq 0$$

Example: Shannon capacity

(a) $C = W \cdot \log\left(1 + \frac{P}{N}\right)$

– C is a concave function of P

- Capacity increment due to increased power diminishes as $P\uparrow$.

- "Principal of diminishing Returns"
  - often assumed for "utility" functions.

(b) $P = N\cdot\left(e^{C/W} - 1\right)$

- P is a convex function of C
- To get the same increment in capacity, the required increase in power grows quickly as $C\uparrow$.

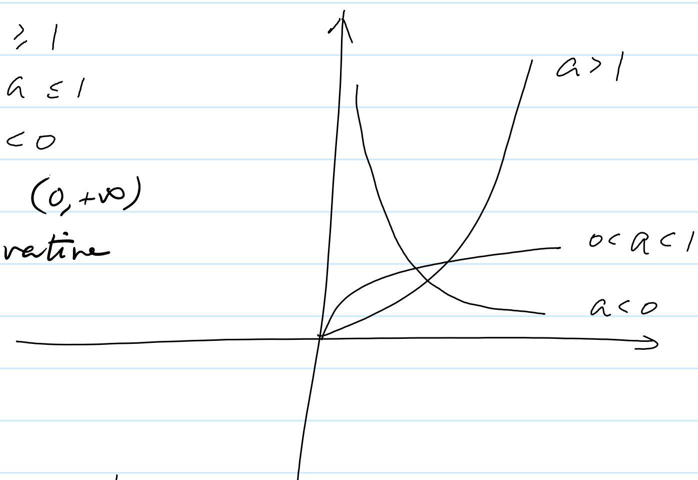
(2) powers $\quad x^a \quad a \geq 1$

$\qquad\qquad -x^a \quad 0 \leq a \leq 1$

$\qquad\qquad x^a \quad a < 0$

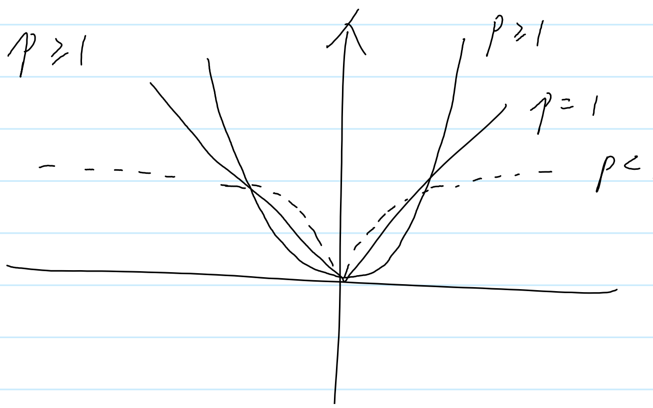
- restricted domain to $(0, +\infty)$

check 2nd-order derivative



power of absolute values

$|x|^p \quad p \geq 1$



norms

$$\|x\|_p = \left(\sum_{i=1}^{n} |x_i|^p \right)^{1/p} \qquad p \geq 1$$

— through triangular inequality (homework)

Parabola

$$(x - x_c)^T P (x - x_c)$$

↑ positive semi-definite.

Examples:

(a) Mean square errors, regression.

Given (x_i, y_i) $i = 1, 2, \cdots, N$
What is the parameter a, b such that
$y = ax + b$ has the smallest mean square error, given by

$$\sum_{i=1}^{N} \left[y_i - (ax_i + b) \right]^2$$

(Q) Can we use other types of norm?

(A) Yes. e.g.

$$\sum_{i=1}^{N} \left| y_i - (ax_i + b) \right|^p$$

still convex when $p \geq 1$

(b) Detection & Likelihood functions.

Skip

Take x as a parameter, Given x,
the distribution of y is given by $P(y|x)$.

Problem: estimate the value of parameter x based on observing one sample y from the distribution

e.g. x is the transmitted signal
y is the received signal corrupted by Gaussian noise, $y = x + z$

$$P(z) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{z^2}{2\sigma^2}}$$

Now given y, what is the estimate of x?

Let $l(x) = \log P(y|x)$ — log-likelihood function

Maximum Likelihood detection:

Estimate x as

$$\hat{x} = \arg\max_x P(y|x) = \arg\max_x l(x)$$

i.e, to choose as our estimate a value of the parameter that maximizes the likelihood function of the observed value of y.

For Gaussian noise

$$P(y|x) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(y-x)^2}{2\sigma^2}}$$

$$\Rightarrow \quad l(x) = \log\frac{1}{\sqrt{2\pi}\,\sigma} - \frac{1}{2\sigma^2}(y-x)^2$$

which is a concave function of x.

Other forms of noise also possible

– Cost of energy generation

$$C(p) = a \cdot p + b p^2$$

- Utility function of an electricity consumer

$$U(p) = U_0 - b(p - p_0)^2$$

most-desirable consumption