Notes for "Convex Optimization"

Chapter 3: Convex Function

Summary

In this chapter, the author first introduces the definition of convex function and its first-, second-order conditions. Then, the concepts of sublevel sets and epigraph are involved, which are the useful tool to judge the function convexity. An important inequlity "Jensen's inequality" is derived.
Secondly, the author introduces some kinds of operations which can preserve the functions' convexity, including Nonnegative weighted sums, Composition with an affine mapping, Pointwise maximum and supremum, Composition, and Perspective of a function.
Thirdly, the author extends the convex funtion into some other functions, such as the conjugate function, Quasiconvex function, and log-concave and log-convex functions. The basic definitions, determination condition, and properties are descibed.
Finally, the author introduces the convexity of generalized inequalities, which extends the scalar function ($f: R^n \rightarrow R$) into vector function ($f: R^n \rightarrow R^p$).

Convex Function Definition

A function: $f: R^n \rightarrow R$ is convex if $\mathbf{dom} f$ is convex set and if for all $x, y \in \mathbf{dom} f$, and θ with $0 \leq \theta \leq 1$, we have

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

A function is convex if and only if it is convex when restricted to any line that intersects its domain. We can construct a line:

$$g(t) = \{x + tv\}$$

The intersection between g and $\mathbf{dom} f$ is $\{x + tv \in \mathbf{dom} f\}$. Then $f(x)$ is convex if and only if

$$g(t) = f(x + tv), x + tv \in \mathbf{dom} f$$

is convex.

Extended-value extensions:

$$\tilde{f}(x) = \begin{cases} f(x) & x \in \mathbf{f} \\ \infty & x \notin \mathbf{f} \end{cases}$$

First order condition

Suppose f is differentiable, f is convex if and only if $\mathbf{dom} f$ is convex and

$$f(y) \geq f(x) + \nabla f(x)^T (y - x)$$

holds for all $x, y \in \mathbf{dom} f$.

Second order condition

Assume that f is twice differentiable. Then f is convex if and only if $mathbf f f$ is convex and its Hessian is positive semidefinate: for all $x \in \mathbf{dom} f$,

$$\nabla^2 f(x) \succeq 0$$

Note: The separate requirement that $\mathbf{dom} f$ be convex cannot be dropped from the first- and second-order characterizations of convexity.

Proof for First- and Second-order conditions

Details can be found in:
http://yangzhou301.xyli.me/2016/03/14/826442654/

Sublevel sets

The α-sublevel set of a function $f : R^n \rightarrow$ is defined as

$$C_a = \{x \in \mathbf{dom} f | f(x) \leq \alpha\}$$

Sublevel set of a convex function is convex.

Epigraph

Epi meas above so epigraph means above the graph

$$\mathbf{epi} f = \{(x, f(x)) | x \in \mathbf{dom} f\}$$

A function is convex if and only if its epigraph is convex set.

Jensen's inequality and extensions

If f is convex, $x_1, ..., x_k \in \mathbf{dom} f$, and $\theta_1, ..., \theta_k \leq 0$ with $\theta_1 + ... + \theta_k = 1$, then

$$f(\theta_1 x_1 + ... + \theta_k x_k) \leq \theta_1 f(x_1) + ... + \theta_k f(x_k)$$

It can extends to **integrals and probability**.

Operations that preserve convexity

1. Nonnegative weighted sums

$$f = w_1 f_1 + ... + w_k f_k, w_i \geq 0$$

It can extends to **integrals and probability**.

2. Composition with an affine mapping

Suppose $f : R^n \rightarrow R, A \in R^{n \times m}$ and $b \in R^n$, Define $g : R^n \rightarrow R$ by

$$g(x) = f(Ax + b)$$

with $\mathbf{dom} g = \{x | Ax + b \in \mathbf{dom} f\}$. Then, if f is convex, so is g.

3. Pointwise maximum and supremum

If $f_1, ..., f_m$ are convex, then their pointwise maximum

$$f(x) = \max\{f_1(x), ..., f_m(x)\}$$

4. Composition

$h : R^k \rightarrow R$ and $g : R^n \rightarrow R^k$, guarantee convexity of their composition $f = h(g(x)) : R^n \rightarrow R$, defined by

$$f(x) = h(g(x)), \quad \mathbf{dom} f = \{x \in \mathbf{dom} g | g(x) \in \mathbf{dom} h\}$$

it $\mathbf{dom}\, g = \mathbf{R}^n$ and $\mathbf{dom}\, h = \mathbf{R}$:

f is convex if h is convex, $\tilde{h}$ is nondecreasing, and g is convex,
f is convex if h is convex, $\tilde{h}$ is nonincreasing, and g is concave,
f is concave if h is concave, $\tilde{h}$ is nondecreasing, and g is concave,
f is concave if h is concave, $\tilde{h}$ is nonincreasing, and g is convex.

Note the requirement that monotonicity hold for the extended-value extension $\widetilde{h}$ and not just the function h.

Vector composition:

$$f(x) = h(g(x)) = h(g_1(x), ..., g_k(x))$$
$$f''(x) = g'(x)^T \nabla^2 h(g(x)) g'(x) + \nabla h(g(x))^T g''(x)$$

f is convex if h is convex, h is nondecreasing in each argument, and g_i are convex,
f is convex if h is convex, h is nonincreasing in each argument, and g_i are concave,
f is concave if h is concave, h is nondecreasing in each argument, and g_i are concave.

5. Perspective of a function

If $f : R^n \to R$, then the perspective of f is the function $g : R^{n+1} \to R$ defined by

$$g(x,t) = tf(x/t)$$

with domain

$$\mathbf{dom} g = \{(x,t) | x/t \in \mathbf{dom} f, t > 0\}$$

If f is convex function, then so is tis perspective function.

The conjugate function

Definition:

Let $f : R^n \to R$. The function $f^* : R^n \to R$, defined as

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

is called the conjugate of the function f.

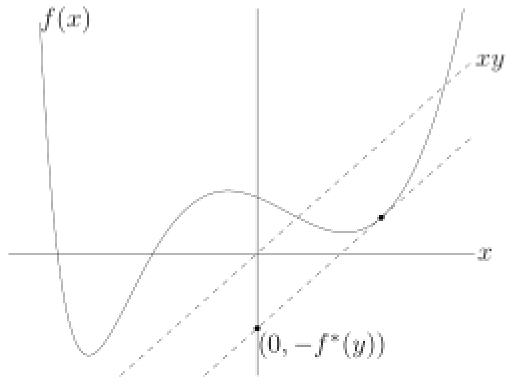

Figure 3.8 A function $f : \mathbf{R} \to \mathbf{R}$, and a value $y \in \mathbf{R}$. The conjugate function $f^*(y)$ is the maximum gap between the linear function yx and $f(x)$, as shown by the dashed line in the figure. If f is differentiable, this occurs at a point x where $f'(x) = y$.

f^* is a convex function whether or not f is convex.

Some examples are listed to explain how to conduct the conjugate function.

Example 3.21 We derive the conjugates of some convex functions on $\mathbf{R}$.

- *Affine function.* $f(x) = ax + b$. As a function of x, $yx - ax - b$ is bounded if and only if $y = a$, in which case it is constant. Therefore the domain of the conjugate function f^* is the singleton $\{a\}$, and $f^*(a) = -b$.
- *Negative logarithm.* $f(x) = -\log x$, with $\mathbf{dom}\, f = \mathbf{R}_{++}$. The function $xy + \log x$ is unbounded above if $y \geq 0$ and reaches its maximum at $x = -1/y$ otherwise. Therefore, $\mathbf{dom}\, f^* = \{y \mid y < 0\} = -\mathbf{R}_{++}$ and $f^*(y) = -\log(-y) - 1$ for $y < 0$.
- *Exponential.* $f(x) = e^x$. $xy - e^x$ is unbounded if $y < 0$. For $y > 0$, $xy - e^x$ reaches its maximum at $x = \log y$, so we have $f^*(y) = y \log y - y$. For $y = 0$,

Basic properties:

1. Fenchel's inequality

$$f(x) + f^*(y) \geq x^T y$$

2. Conjugate of the conjugate
   if f is convex and f is closed, then $f^{**} = f$
3. Scaling and composition with affine transformation
   For $a > 0$ and $b \in R$, the conjugate of $g(x) = af(x) + b$ is $g^*(y) = af^*(y/a) - b$
   Suppose $A \in R^{n \times n}$ is onosingular and $b \in R^n$, the conjugate of g(x)=f(Ax+b) is

$$g^*(y) = f^*(A^{-T} y) - b^T A^{-T} y$$

   with $\mathbf{dom} g^* = A^T \mathbf{dom} f^*$

4. Sums of independent functions

$$f^*(w,z) = f_1^*(w) + f_2^*(z)$$

Quasiconvex function

1. Definition

A function $f : R^n \to R$ is called quasiconvex if its domain and all its sublevel sets

$$S_a = \{x \in \mathbf{dom} f | f(x) \le \alpha\}$$

for $\alpha \in R$ are convex.

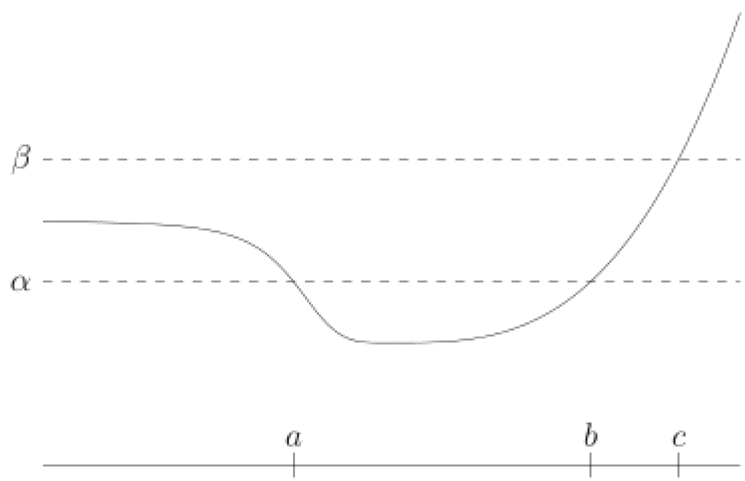

Figure 3.9 A quasiconvex function on $\mathbf{R}$. For each α, the α-sublevel set S_α is convex, *i.e.*, an interval. The sublevel set S_α is the interval $[a, b]$. The sublevel set S_β is the interval $(-\infty, c]$.

2. Basic properties

A function f is quasiconvex if and only if $\mathbf{dom} f$ is convex and for any $x, y \in \mathbf{dom} f$ and $0 \le \theta \le 1$,

$$f(\theta x + (1-\theta)y) \le \max\{f(x), f(y)\}$$

3. Differentiable quasiconvex functions

First-order conditions:
Suppose $f : R^n \to R$ is differentiable. Then f is quasiconvex if and only if $\mathbf{dom} f$ is convex and for all $x, y \in \mathbf{dom} f$

$$f(y) \le f(x) \Rightarrow \nabla f(x)^T (y - x) \le 0$$

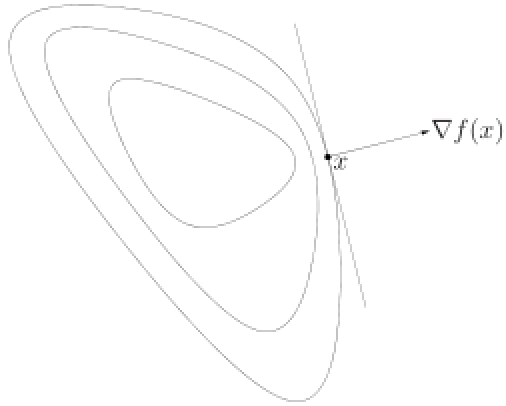

Figure 3.12 Three level curves of a quasiconvex function f are shown. The vector $\nabla f(x)$ defines a supporting hyperplane to the sublevel set $\{z \mid f(z) \leq f(x)\}$ at x.

Proof: https://blog.csdn.net/wang136958280/article/details/86549034

Second-order conditions:

Suppose f is twice differentiable. If f is quasiconvex, then for all $x \in \mathbf{dom} f$ and all $y \in R^n$, we have

$$y^T \nabla f(x) = 0 \Rightarrow y^T \nabla^2 f(x) y \geq 0$$

Operations that preserve quasiconvexity

1. Nonnegative weighted maximum

$$f = \max\{w_1 f_1, ..., w_m f_m\}$$

$$f(x) = \sup_{y \in C}(w(y) g(x, y))$$

2. Composition

If $g : R^n \rightarrow R$ is quasiconvex and $h : R \rightarrow R$ is nondecreasing, then $f = h(g(x))$ is quasiconvex.

If f is quasiconvex, then $g(x) = f(Ax + b)$ is quasiconvex.

3.Minimization

If $f(x, y)$ is quasiconvex jointly in x and y and C is a convex setm then the function

$$g(x) = \inf_{y \in C} f(x, y)$$

is quasiconvex.

log-concave and log-convex functions

1. Definition

A function $f : R^n \to R$ is logarithmically concave or log-concave if $f(x) > 0$ for all $x \in \mathbf{dom} f$ and $\log f$ is concave.
f is log-convex if and only if $1/f$ is log-concave.
A function $f : R^n \to R$ with convex domain and $f(x) > 0$ for all $x \in \mathbf{dom} f$, is log-concave if and only if for all $x, y \in \mathbf{dom} f$ and $0 \le \theta \le 1$, we have

$$f(\theta x + (1-\theta)y) \ge f(x)^\theta f(y)^{1-\theta}$$

2. Properties

Twice differentiable log-convex/concave functions
f is log-convex if and only if for all $x \in \mathbf{dom} f$

$$f(x)\nabla^2 f(x) \succeq \nabla f(x) \nabla f(x)^T$$

and log-concave if and only if for all $x \in \mathbf{dom} f$

$$f(x)\nabla^2 f(x) \preceq \nabla f(x) \nabla f(x)^T$$

Log-convexity and log-concavity are closed under multiplication and positive scaling.
Integration of log-concave functions
If $f : R^n \times R^m \to R$ is log-concave, then

$$g(x) = \int f(x, y) dy$$

is a log-concave function of x(on R^n).
If f and g are log-concave on R^n, then so is the convolution

$$(f * g)(x) = \int f(x-y) g(y) dy$$

Convexity with respect to generalized inequalities

1. Monotonicity

Suppose $K \subseteq R^n$ is a proper cone with associated generalized inequality $\preceq_K$. A function $f : R^n \to R$ is called K-nondecreasing if

$$x \preceq_K y \Rightarrow f(x) \leq f(y)$$

and K-increasing if

$$x \preceq_K y, x \neq y \Rightarrow f(x) < f(y)$$

A dfifferentiable function f with convex domain, is K-nondecreasing if and only if

$$\nabla f(x) \succeq_{K^*} 0$$

For all $x \in \mathbf{dom} f$, f is K-increasing if

$$\nabla f(x) \succ_{K^*} 0$$

Convexity with respect to a generalized inequality

Suppose $K \subseteq R^m$ is a proper cone with associated generalized inequality $\preceq_K$. We say $f : R^n \to R^m$ is K-convex if for all x, y and $0 \leq \theta \leq 1$,

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

.

The function is strictly K-convex if

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

for all $x \neq y$ and $0 \leq \theta \leq 1$

Dual characterization

A function f is K-convex if and only if for every $w \succeq_{K^*} 0$, the function $w^T f$ is convex.

Differentiable K-convex function

A differentiable function f is K-convex if and only if its domain is convex, and for all $x, y \in \mathbf{dom} f$

$$f(y) \succeq_K f(x) + Df(x)(y-x)$$

Here $Df(x) \in R^{m \times n}$ is the derivative or Jacobian matrix of f at x.

Composition theorem

If $g : R^n \to R^p$ is convex, $h : R^p \to R$ is convex, and $\widetilde{h}$ is K-nondecreasing, then $h(g(x))$ is convex.

Example:

$g : R^{m \times n} \to S^n$ defined by

$$g(X) = X^T AX + B^T + X^T B + C$$

where $A \in S^m, B \in R^{m\times n}$ and $C \in S^n$ is convex when $A \succeq 0$.
The function $h : S^n \to R$ defined by $h(Y) = -\log\det(-Y)$ is convex and increasing on $\mathbf{h} = -S^n_{++}$.
Then

$$f(X) = -\log\det(-(X^T AX + B^T + X^T B + C))$$

is convex on

$$\mathbf{dom} f = \{X \in R^{m\times n} | X^T AX + B^T + X^T B + C \prec 0\}$$

This generalizes the fact that

$$-log(-(ax^2 + bx + c))$$

is convex on

$$\{x \in R | ax^2 + bx + c \in 0\}$$

provided $a \geq 0$