Lecture 20

§1 Quasiconvex functions

1. Sublevel set (下水平集)

$S_\alpha = \{x \in \operatorname{dom} f \mid f(x) \le \alpha\}$

注: 1° 描述的是定义域中的集合，不同的α决定不同的集合

2° 凸函数的下水平集是凸的

下水平集为凸的函数不一定凸

3° Sublevel set 与 epigraph 都建立了函数凹凸性与集合凹凸性之间的联系.

Epigraph 与凸函数能建立双向联系

Sublevel set 与凸函数不能建立双向联系

2. Quasiconvex function (拟凸函数)

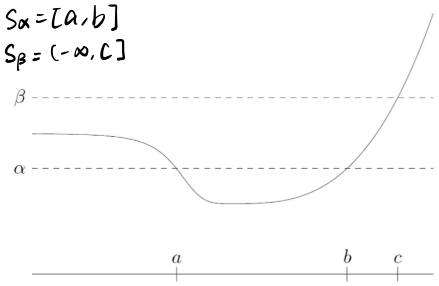
A function $f: R^n \to R$ is quasiconvex if all its sublevel sets are convex.

i.e. if $S_\alpha = \{x \mid f(x) \le \alpha\}$ is convex for each $\alpha \in R$

若f的所有下水平集都为凸集，则f为拟凸函数.

$S_\alpha = [a, b]$

$S_\beta = (-\infty, c]$



* 所有的凸函数均为拟凸函数

* 拟凸函数 can be concave ($\log x$), or discontinuous

3. Quasiconcave function (拟凹函数)

- 若 $-f$ 为 quasiconvex，则 f quasiconcave
- 所有的 superlevel sets (上水平集) $S_\alpha = \{x \mid f(x) \ge \alpha\}$ 为凸

例: $x_1 x_2$ on R_+^2, $-\log x$ (同时为凸函数)

4. Quasilinear function (拟线性函数)

- 若 f 既 quasiconvex 又 quasiconcave，则 f quasilinear
- 其所有的 sublevel 和 superlevel sets 均为 halfspaces

  所有的 level set 均为 affine

例: $\log x$ on R_+

$\operatorname{ceil}(x) = \inf\{z \in Z \mid z > x\}$

$\dfrac{c + \sum_i a_i x_i}{d + \sum_i b_i x_i}$

§2 拟凸函数的性质

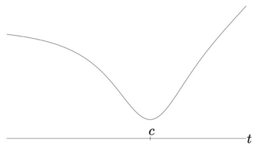
1. 性质 1

一个连续函数 $f: R \to R$ 为 quasiconvex，当且仅当（充要）以下三个情况中，至少一个成立.

- f is nondecreasing
- f is nonincreasing
- 存在一点 $c \in \mathrm{dom} f$，使得对 $t \le c$（$t \in \mathrm{dom} f$）有 f nonincreasing. 对 $t \ge c$（$t \in \mathrm{dom} f$）有 f nondecreasing.

A continuous function $f : \mathbf{R} \to \mathbf{R}$ is quasiconvex if and only if at least one of the following conditions holds:

- f is nondecreasing
- f is nonincreasing
- there is a point $c \in \mathbf{dom}\, f$ such that for $t \le c$ (and $t \in \mathbf{dom}\, f$), f is nonincreasing, and for $t \ge c$ (and $t \in \mathbf{dom}\, f$), f is nondecreasing.



证明:

- First suppose that f satisfies one of the conditions and suppose that for some number α, both x_1 and x_2 are members of the sublevel set $S_\alpha = \{x \mid f(x) \le \alpha\}$. Then $f(x_1) \le \alpha$ and $f(x_2) \le \alpha$. Therefore $f(x) \le \alpha$ for every point x between x_1 and x_2. Therefore S_α is convex.

- Suppose f is quasiconvex. If f does not satisfy any of the conditions then we can find x_1 and x_2 and x_3 such that $x_1 < x_2 < x_3$ and $f(x_2) > \max(f(x_1), f(x_3))$. Then the sublevel set S_α for $\alpha = \max(f(x_1), f(x_3))$ includes x_1 and x_3, but not x_2, this contradicts with the assumption that f is quasiconvex. Therefore, if f is quasiconvex, then f must satisfy one of the conditions.

Implication:

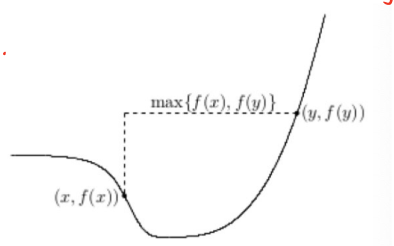
- 若 f 为 quasiconvex，则 optimal minimizers 形成一个凸集
- 若 $\vec{x}^*$ 为一个 strict local minimizer（存在 $\varepsilon > 0$，使 $\forall \vec{y} \in S \cap B(\vec{x}^*, \varepsilon)$ 且 $\vec{y} \ne \vec{x}$，有 $f(\vec{x}^*) < f(\vec{y})$），则 $\vec{x}^*$ 同时也是 global minimizer.

Example

- Show $f(a) = 2a^4 - a^3 + 4a^2 - 3a$ is quasiconvex.
- $f'(a) = (8a-3)(1 + a^2)$
- $f'(a) = 0$ has a single solution as 3/8.

2. 性质 2

f 为 quasiconvex 当且仅当（充要）$f(\theta \vec{x} + (1-\theta)\vec{y}) \le \max\{f(\vec{x}), f(\vec{y})\}$，for $0 \le \theta \le 1$ and $\forall \vec{x}, \vec{y}$.



证明：

- Suppose f is quasiconvex. Pick any **x** and **y** in the domain of f. Let α = max(f(**x**) , f(**y**)), then **x** and **y** belongs to the sublevel set S_α = {x| f(x) ≤ α}. Since f is quasiconvex, then S_α is convex, and θ**x** + (1 − θ)**y** ∈ S_α. Therefore f(θ**x** + (1 − θ)**y**) ≤ α = max {f(**x**), f(**y**)}.

- Pick any sublevel set S_α and **x**, **y** ∈ S_α. Then max {f(**x**), f(**y**)} ≤ α. Since f(θ**x** + (1 − θ)**y**) ≤ max {f(**x**), f(**y**)} for 0 ≤ θ ≤1, therefore f(θ**x** + (1 − θ)**y**) ≤ max {f(**x**), f(**y**)} ≤ α, which means θ**x** + (1 − θ)**y** ∈ S_α. Therefore S_α is convex, and f is quasiconvex.

3. 性质3 (First-order conditions)

若 $f: R^n \to R$ 可微，则 f 为 quasiconvex 当且仅当(充要) 对 $\forall f(\vec{y}) \le f(\vec{x})$，令 $\vec{y} = \vec{x} + s\vec{e}$，有

$$s\frac{df(\vec{x}+t\vec{e})}{dt}\Big|_{t=0} \le 0$$ (y处的方向导数 ≤ 0)

* 一维：$f(y) \le f(x) \Rightarrow f'(x)(y-x) \le 0$

- A differentiable f : $R^n \to R$ is quasiconvex if and only if
  - For any f(**y**) ≤ f(**x**), letting **y** = **x** + s **e**, we have

$$s\frac{d\,f(x+te)}{dt}\Big|_{t=0} \le 0$$

One dimension: $f(y) \le f(x) \Rightarrow f'(x)(y-x) \le 0$.

证明：

We show the proof for the case $f : \mathbf{R} \to \mathbf{R}$

First suppose f is a differentiable function on $\mathbf{R}$ and satisfies

$$f(y) \le f(x) \Longrightarrow f'(x)(y-x) \le 0. \qquad (3.43.A)$$

Suppose $f(x_1) \ge f(x_2)$ where $x_1 \ne x_2$. We assume $x_2 > x_1$ (the other case can be handled similarly), and show that $f(z) \le f(x_1)$ for $z \in [x_1, x_2]$. Suppose this is false, *i.e.*, there exists a $z \in [x_1, x_2]$ with $f(z) > f(x_1)$. Since f is differentiable, we can choose a z that also satisfies $f'(z) < 0$. By (3.43.A), however, $f(x_1) < f(z)$ implies $f'(z)(x_1 - z) \le 0$, which contradicts $f'(z) < 0$.

To prove sufficiency, assume f is quasiconvex. Suppose $f(x) \ge f(y)$. By the definition of quasiconvexity $f(x + t(y - x)) \le f(x)$ for $0 < t \le 1$. Dividing both sides by t, and taking the limit for $t \to 0$, we obtain

$$\lim_{t\to 0} \frac{f(x + t(y-x)) - f(x)}{t} = f'(x)(y-x) \le 0,$$

which proves (3.43.A).

4. 性质4 (Second-order conditions)

- 给出 f 可微
- 若 f quasiconvex，则对 $\forall \vec{x}$ 与 $\vec{e}$，
  若 $\frac{df(\vec{x}+t\vec{e})}{dt}\Big|_{t=0} = 0$，则 $\frac{d^2f(\vec{x}+t\vec{e})}{dt^2}\Big|_{t=0} \ge 0$
- 若 f 满足以下条件，则 f quasiconvex
  对 $\forall \vec{x}$ 与 $\vec{e}$，若 $\frac{df(\vec{x}+t\vec{e})}{dt}\Big|_{t=0} = 0$，则始终有 $\frac{d^2f(\vec{x}+t\vec{e})}{dt^2}\Big|_{t=0} > 0$

- Given f differentiable

- If f is quasiconvex, then for any **x** and **e,**
  - If $\frac{d\,f(x+te)}{dt}|_{t=0} = 0$, we must have $\frac{d^2 f(x+te)}{dt\,dt}|_{t=0} \geq 0$

- If f satisfies the following property, f is quasiconvex.
  - For any **x** and **e,** if $\frac{d\,f(x+te)}{dt}|_{t=0} = 0$, we always have $\frac{d^2 f(x+te)}{dt\,dt}|_{t=0} > 0$

证明:

We show the proof for the case $f : \mathbf{R} \to \mathbf{R}$

We first show that if $f : \mathbf{R} \to \mathbf{R}$ is quasiconvex on an interval (a, b), then it must satisfy (3.21), *i.e.*, if $f'(c) = 0$ with $c \in (a, b)$, then we must have $f''(c) \geq 0$. If $f'(c) = 0$ with $c \in (a, b)$, $f''(c) < 0$, then for small positive ϵ we have $f(c-\epsilon) < f(c)$ and $f(c+\epsilon) < f(c)$. It follows that the sublevel set $\{x \mid f(x) \leq f(c) - \epsilon\}$ is disconnected for small positive ϵ, and therefore not convex, which contradicts our assumption that f is quasiconvex.

Now we show that if the condition (3.22) holds, then f is quasiconvex. Assume that (3.22) holds, *i.e.*, for each $c \in (a, b)$ with $f'(c) = 0$, we have $f''(c) > 0$. This means that whenever the function f' crosses the value 0, it is strictly increasing. Therefore it can cross the value 0 at most once. If f' does not cross the value 0 at all, then f is either nonincreasing or nondecreasing on (a, b), and therefore quasiconvex. Otherwise it must cross the value 0 exactly once, say at $c \in (a, b)$. Since $f''(c) > 0$, it follows that $f'(t) \leq 0$ for $a < t \leq c$, and $f'(t) \geq 0$ for $c \leq t < b$. This shows that f is quasiconvex.

Convex Optimization — Stephen Boyd and Lieven Vandenberghe

Example

- Show $f(a) = 2a^4 - a^3 + 4a^2 - 3a$ is quasiconvex.

- $f'(a)$ = (8a-3) $(1 + a^2)$

- $f'(a)$ =0 has a single solution as 3/8.

- $f''(3/8)$ = 73/8>0

§3 Operations that perserve quasiconvex

1. Scaling (缩放)

若f拟凸, $w>0$, 则 wf 也拟凸.

f与wf有相同的下水平集: 当且仅当 $wf(x) \leq \alpha$ 且 $f(x) \leq \alpha/w$

Scaling

If f is quasiconvex and $w > 0$, then wf is also quasiconvex.

f and wf have the same sublevel sets: $wf(x) \leq \alpha$ iff $f(x) \leq \alpha/w$,

2、与非减函数复合

若 $f: R^n \to R$ 为 quasiconvex, h 为 non-decreasing, 则 $h \circ f$ 为 quasiconvex.

f与 $h \circ f$ 有相同的下水平集: 当且仅当 $h(f(x)) \leq \alpha$ 且 $f(x) \leq h^{-1}(\alpha)$

Composition with Nondecreasing Function

If $f : \mathbb{R}^n \to \mathbb{R}$ is quasiconvex $h : \mathbb{R} \to \mathbb{R}$ is non-decreasing, then $h \circ f$ is quasiconvex.

$h \circ f$ and f have the same sublevel sets: $h(f(x)) \leq \alpha$ iff $f(x) \leq h^{-1}(\alpha)$

3. linear transformation

若 f 为 quasiconvex，则 $g(\vec{x}) = f(a\vec{x}+b)$ 为 quasiconvex.

证明：

- Want to show $g(x)$ is quasiconvex

  That is equivalent to showing, $\forall \alpha$, $S_\alpha^{(g)}$ is convex

- Pick any $x_1, x_2 \in S_\alpha^{(g)}$.

  Want to show $\theta x_1 + (1-\theta)x_2 \in S_\alpha^{(g)}$

  That is equavalent to showing, $\forall \alpha$, $g(\theta x_1 + (1-\theta)x_2) \le \alpha$

- $g(\theta x_1 + (1-\theta)x_2) = f(a[\theta x_1 + (1-\theta)x_2] + b)$

  $= f(\theta(ax_1+b) + (1-\theta)(ax_2+b))$

  Since $f(ax_1+b) = g(x_1) \le \alpha$; $f(ax_2+b) = g(x_2) \le \alpha$,

  This means ax_1+b, $ax_2+b \in S_\alpha^{(f)}$

  Since f is quasiconvex, $\theta(ax_1+b) + (1-\theta)(ax_2+b) \in S_\alpha^{(f)}$

4. Maximum

若 f_1, f_2 为 quasiconvex，则 $g(x) = \max\{f_1(x), f_2(x)\}$ 也为 quasiconvex.

推广：

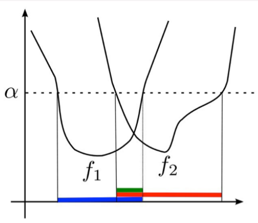
maximum of any number of functions : $\max_{i=1}^{k} f_i(x)$

supremum of an infinite set of functions : $\sup_y f_y(x)$

Maximum

If f_1, f_2 are quasiconvex, then $g(x) = \max\{f_1(x), f_2(x)\}$ is also quasiconvex.

Generalizes to the maximum of any number of functions, $\max_{i=1}^{k} f_i(x)$, and also to the supremum of an infinite set of functions $\sup_y f_y(x)$.



5. Minimization

若 $f(x,y)$ 为 quasiconvex 且 C 为非空凸集，则 $g(x) = \inf_{y\in C} f(x,y)$ 为 quasiconvex

Minimization

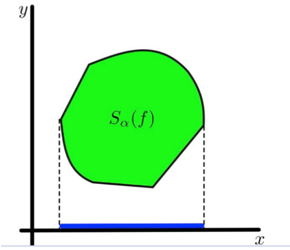
If $f(x,y)$ is quasiconvex and C is convex and nonempty, then $g(x) = \inf_{y\in C} f(x,y)$ is quasiconvex.

注：若 $C \in R$，可理解为将 $f(x,y)$ 投影到一垂直于定义域平面的平面上，所构成图形的下边缘构成的函数为拟凸.

证明：

Proof (for $C = \mathbb{R}^k$)

$S_\alpha(g)$ is the projection of $S_\alpha(f)$ onto hyperplane $y = 0$.



- Pick x_1, x_2, α such that $g(x_1) \le \alpha$, $g(x_2) \le \alpha$

  Want to show $g(\theta x_1 + (1-\theta) x_2) \le \alpha$

- $g(\theta x_1 + (1-\theta) x_2) = \min\limits_{y \in C} f(\theta x_1 + (1-\theta) x_2, y)$

  $\le f(\theta x_1 + (1-\theta) x_2, \theta y_1 + (1-\theta) y_2)$

  where y_1, y_2 satisfy

  $g(x_1) = f(x_1, y_1) \le \alpha$

  $g(x_2) = f(x_2, y_2) \le \alpha$

- Then $(\theta x_1 + (1-\theta) x_2, \theta y_1 + (1-\theta) y_2) = \theta (x_1, y_1) + (1-\theta)(x_2, y_2)$

  $g(\theta x_1 + (1-\theta) x_2) \le f(\theta x_1 + (1-\theta) x_2, \theta y_1 + (1-\theta) y_2)$

  $= f(\theta (x_1, y_1) + (1-\theta)(x_2, y_2))$

  $\le \alpha$

6、*Sum

f_1, f_2 为 quasiconvex，则 $f_1 + f_2$ 不一定 quasiconvex

例：$x^3 + x^2$