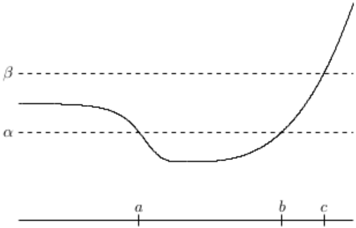
CS599: Convex and Combinatorial Optimization
Fall 2013
Lecture 8: Convex Functions Wrapup

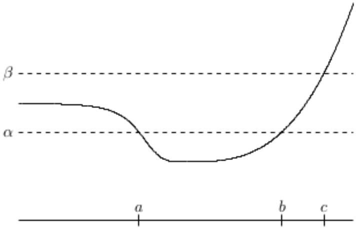
Instructor: Shaddin Dughmi

Outline



Quasiconvex Functions

A function $f : \mathbb{R}^n \to \mathbb{R}$ is quasiconvex if all its sublevel sets are convex.
i.e. if $S_\alpha = \{x | f(x) \leq \alpha\}$ is convex for each $\alpha \in \mathbb{R}$.



Quasiconvex Functions

A function $f : \mathbb{R}^n \to \mathbb{R}$ is quasiconvex if all its sublevel sets are convex.
i.e. if $S_\alpha = \{x | f(x) \leq \alpha\}$ is convex for each $\alpha \in \mathbb{R}$.

- f is quasiconcave if $-f$ is quasiconvex
  - Equivalently, all its superlevel sets are convex.
- f is quasilinear if it is both quasiconvex and quasiconcave
  - Equivalently, all its sublevel and superlevel sets are halfspaces, and all its level sets are affine

Examples

- $\log x$ is quasilinear on $\mathbb{R}_+$
- All functions $f : \mathbb{R} \to \mathbb{R}$ that are “unimodal”
- $x_1 x_2$ is quasiconcave on $\mathbb{R}^2_+$
- $\frac{a^\intercal x+b}{c^\intercal x+d}$ is quasilinear
- $||x||_0$ is quasiconcave on $\mathbb{R}^n_+$.

Alternative Definitions

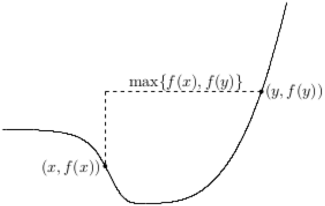
We will now look at two equivalent definitions of quasiconvex functions

Inequality Definition

f is quasiconvex if the following relaxation of Jensen's inequality holds:

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

for $0 \leq \theta \leq 1$



Alternative Definitions

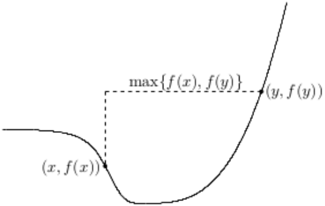
We will now look at two equivalent definitions of quasiconvex functions

Inequality Definition

f is quasiconvex if the following relaxation of Jensen's inequality holds:

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

for $0 \leq \theta \leq 1$



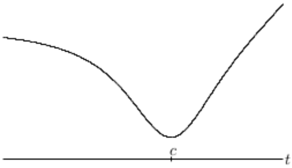
- Like Jensen's inequality, a property of f on lines in its domain

Alternative Definitions

First Order Definition

A differentiable $f : \mathbb{R}^n \to \mathbb{R}$ is quasiconvex if and only if whenever $f(y) \leq f(x)$, we have

$$\nabla f(x)^\intercal (y - x) \leq 0$$

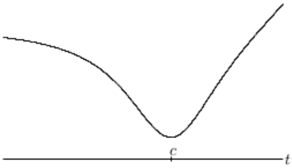


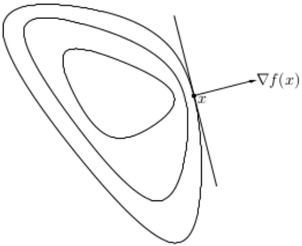
Alternative Definitions

First Order Definition

A differentiable $f : \mathbb{R}^n \to \mathbb{R}$ is quasiconvex if and only if whenever $f(y) \leq f(x)$, we have

$$\nabla f(x)^\mathsf{T}(y - x) \leq 0$$





$\nabla f(x)$ defines a supporting hyperplane for sublevel set with $\alpha = f(x)$

Operations Preserving Quasiconvexity

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 $f(x) \leq \alpha/w$,

Operations Preserving Quasiconvexity

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 $f(x) \leq \alpha/w$,

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 $f(x) \leq h^{-1}(\alpha)$

Operations Preserving Quasiconvexity

Composition with Affine Function

If $f : \mathbb{R}^n \to \mathbb{R}$ is quasiconvex, and $A \in \mathbb{R}^{n \times m}$, $b \in \mathbb{R}^n$, then

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

is a quasiconvex function from $\mathbb{R}^m$ to $\mathbb{R}$.

Operations Preserving Quasiconvexity

Composition with Affine Function

If $f : \mathbb{R}^n \to \mathbb{R}$ is quasiconvex, and $A \in \mathbb{R}^{n \times m}$, $b \in \mathbb{R}^n$, then

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

is a quasiconvex function from $\mathbb{R}^m$ to $\mathbb{R}$.

Proof

The α sublevel of $f(Ax + b) \leq \alpha$ is the inverse image of the α-sublevel of f under the affine map $x \to Ax + b$.

Operations Preserving Quasiconvexity

Composition with Affine Function

If $f : \mathbb{R}^n \to \mathbb{R}$ is quasiconvex, and $A \in \mathbb{R}^{n\times m}$, $b \in \mathbb{R}^n$, then

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

is a quasiconvex function from $\mathbb{R}^m$ to $\mathbb{R}$.

Proof

The α sublevel of $f(Ax + b) \leq \alpha$ is the inverse image of the α-sublevel of f under the affine map $x \to Ax + b$.

Note: extends to linear fractional maps $x \to \frac{Ax+b}{c^T x+d}$.

Operations Preserving Quasiconvexity

Maximum

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

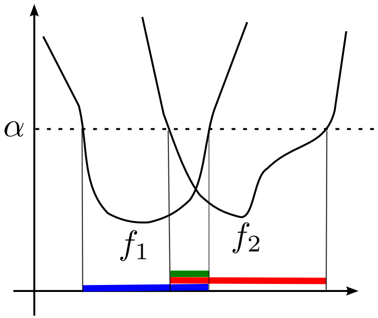
Generalizes to the maximum of any number of functions, $\max_{i=1}^{k} f_i(x)$, and also to the supremum of an infinite set of functions $\sup_y f_y(x)$.

Operations Preserving Quasiconvexity

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 $\sup_y f_y(x)$.



Operations Preserving Quasiconvexity

Minimization

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

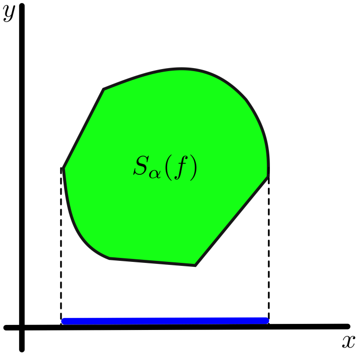
Operations Preserving Quasiconvexity

Minimization

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

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

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



Operations NOT preserving quasiconvexity

Sum

$f_1 + f_2$ is NOT necessarily quasiconvex when f_1 and f_2 are quasiconvex.

Operations NOT preserving quasiconvexity

Sum

$f_1 + f_2$ is NOT necessarily quasiconvex when f_1 and f_2 are quasiconvex.

Composition Rules

The composition rules for convex functions do NOT carry over in general.

Outline

We now briefly look at a class of quasiconcave functions which pops up in "multiplication" and "volume" maximization problems.

Log-concave Functions

A function $f : \mathbb{R}^n \to \mathbb{R}_+$ is log-concave if $\log f(x)$ is a concave function. Equivalently:

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

for $x, y \in \mathbb{R}^n$ and $\theta \in [0, 1]$.

We now briefly look at a class of quasiconcave functions which pops up in “multiplication” and “volume” maximization problems.

Log-concave Functions

A function $f : \mathbb{R}^n \to \mathbb{R}_+$ is log-concave if $\log f(x)$ is a concave function. Equivalently:

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

for $x, y \in \mathbb{R}^n$ and $\theta \in [0, 1]$.

- i.e. concave if plotted on log-scale paper

We now briefly look at a class of quasiconcave functions which pops up in “multiplication” and “volume” maximization problems.

Log-concave Functions

A function $f : \mathbb{R}^n \to \mathbb{R}_+$ is log-concave if $\log f(x)$ is a concave function. Equivalently:

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

for $x, y \in \mathbb{R}^n$ and $\theta \in [0, 1]$.

- i.e. concave if plotted on log-scale paper
- Concave functions are log-concave, and both are quasiconcave.

We now briefly look at a class of quasiconcave functions which pops up in "multiplication" and "volume" maximization problems.

Log-concave Functions

A function $f : \mathbb{R}^n \to \mathbb{R}_+$ is log-concave if $\log f(x)$ is a concave function. Equivalently:

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

for $x, y \in \mathbb{R}^n$ and $\theta \in [0, 1]$.

- i.e. concave if plotted on log-scale paper
- Concave functions are log-concave, and both are quasiconcave.
- Taking the logarithm of a non-concave (yet quasiconcave) function can "concavify" it

We now briefly look at a class of quasiconcave functions which pops up in "multiplication" and "volume" maximization problems.

Log-concave Functions

A function $f : \mathbb{R}^n \to \mathbb{R}_+$ is log-concave if $\log f(x)$ is a concave function. Equivalently:

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

for $x, y \in \mathbb{R}^n$ and $\theta \in [0, 1]$.

- i.e. concave if plotted on log-scale paper
- Concave functions are log-concave, and both are quasiconcave.
- Taking the logarithm of a non-concave (yet quasiconcave) function can "concavify" it
- Most common form of "concavification" and "convexification" of objective functions, which to a large extent is an art.

Examples

- All concave functions
- x^a for $a \geq 0$
- e^x
- $\prod_i x_i$
- Determinant of a PSD matrix
- The pdf of many common distributions such as Gaussian and exponential
  - Intuitively, those distributions which decay faster than exponential (i.e. $e^{-\lambda x)}$)

Operations Preserving Log-Concavity

Scaling

If f is logconcave and $w \in \mathbb{R}$ then wf is also logconcave.

Operations Preserving Log-Concavity

Scaling

If f is logconcave and $w \in \mathbb{R}$ then wf is also logconcave.

Composition with Affine

If f is logconcave then so is $f(Ax + b)$.

Operations Preserving Log-Concavity

Scaling

If f is logconcave and $w \in \mathbb{R}$ then wf is also logconcave.

Composition with Affine

If f is logconcave then so is $f(Ax + b)$.

Multiplication

If f_1, f_2 are log-concave, then so is $f_1 f_2$

Operations Preserving Log-Concavity

Scaling

If f is logconcave and $w \in \mathbb{R}$ then wf is also logconcave.

Composition with Affine

If f is logconcave then so is $f(Ax + b)$.

Multiplication

If f_1, f_2 are log-concave, then so is $f_1 f_2$

Log-concavity NOT preserved by sums.

Operations Preserving Log-Concavity

Theorem (Prekopa & Liendler)

If $f(x, y)$ is log-concave, then $g(x) = \int_y f(x, y)$ is also log-concave.

Operations Preserving Log-Concavity

Theorem (Prekopa & Liendler)

If $f(x,y)$ is log-concave, then $g(x) = \int_y f(x,y)$ is also log-concave.

Example (Yield)

- Design parameters $x \in \mathbb{R}^n$ of a product

Operations Preserving Log-Concavity

Theorem (Prekopa & Liendler)

If $f(x, y)$ is log-concave, then $g(x) = \int_y f(x, y)$ is also log-concave.

Example (Yield)

- Design parameters $x \in \mathbb{R}^n$ of a product
- Noise $\delta \in \mathbb{R}^n$ drawn from a log-concave distribution with pdf f
- Resulting product has configuration $x + \delta$.

Operations Preserving Log-Concavity

Theorem (Prekopa & Liendler)

If $f(x, y)$ is log-concave, then $g(x) = \int_y f(x, y)$ is also log-concave.

Example (Yield)

- Design parameters $x \in \mathbb{R}^n$ of a product
- Noise $\delta \in \mathbb{R}^n$ drawn from a log-concave distribution with pdf f
- Resulting product has configuration $x + \delta$.
- Set $S \subseteq \mathbb{R}^n$ of "desired" configurations

Operations Preserving Log-Concavity

Theorem (Prekopa & Liendler)

If $f(x,y)$ is log-concave, then $g(x) = \int_y f(x,y)$ is also log-concave.

Example (Yield)

- Design parameters $x \in \mathbb{R}^n$ of a product
- Noise $\delta \in \mathbb{R}^n$ drawn from a log-concave distribution with pdf f
- Resulting product has configuration $x + \delta$.
- Set $S \subseteq \mathbb{R}^n$ of "desired" configurations
- Probability of desirable configuration is $\int_{y \in S} f(y - x)$

Operations Preserving Log-Concavity

Theorem (Prekopa & Liendler)

If $f(x, y)$ is log-concave, then $g(x) = \int_y f(x, y)$ is also log-concave.

Example (Yield)

- Design parameters $x \in \mathbb{R}^n$ of a product
- Noise $\delta \in \mathbb{R}^n$ drawn from a log-concave distribution with pdf f
- Resulting product has configuration $x + \delta$.
- Set $S \subseteq \mathbb{R}^n$ of "desired" configurations
- Probability of desirable configuration is $\int_{y \in S} f(y - x)$
- By above theorem, choosing x to optimize this probability is convex optimization problem

Next Week

Convex Optimization Problems!