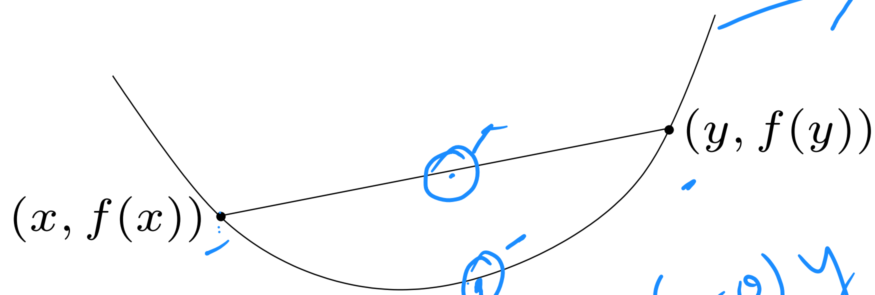
Definition

$f : \mathbf{R}^n \to \mathbf{R}$ is convex if $\mathbf{dom}\, f$ is a convex set and

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

for all $x, y \in \mathbf{dom}\, f$, $0 \le \theta \le 1$

$f(x)$



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

- f is concave if $-f$ is convex

Concave

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

$\geq$

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

for $x, y \in \mathbf{dom}\, f$, $x \ne y$, $0 < \theta < 1$

First-order condition

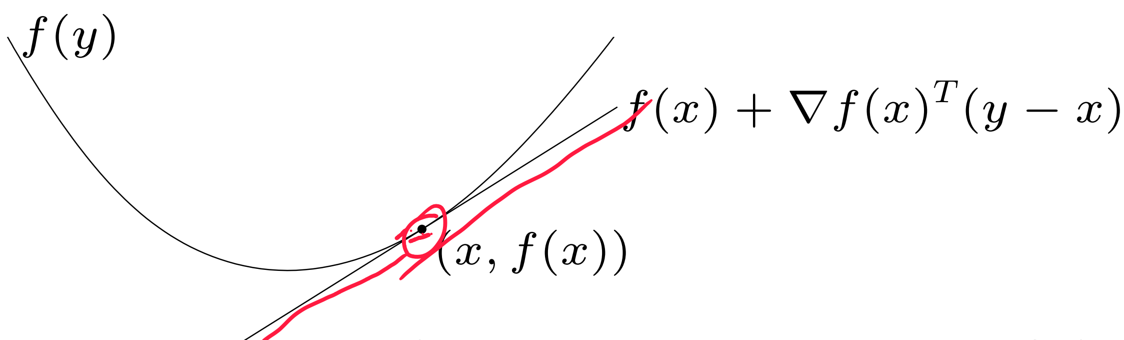
f is **differentiable** if $\mathbf{dom}\, f$ is open and the gradient

$$\nabla f(x) = \left(\frac{\partial f(x)}{\partial x_1}, \frac{\partial f(x)}{\partial x_2}, \ldots, \frac{\partial f(x)}{\partial x_n} \right)$$

exists at each $x \in \mathbf{dom}\, f$

1st-order condition: differentiable f with convex domain is convex iff

$$f(y) \geq f(x) + \nabla f(x)^T (y - x) \quad \text{for all } x, y \in \mathbf{dom}\, f$$



first-order approximation of f is global underestimator

Second-order conditions

f is **twice differentiable** if $\mathbf{dom}\, f$ is open and the Hessian $\nabla^2 f(x) \in \mathbf{S}^n$,

$$\nabla^2 f(x)_{ij} = \frac{\partial^2 f(x)}{\partial x_i \partial x_j}, \quad i, j = 1, \ldots, n,$$

$$\begin{pmatrix} \frac{\partial^2 f(x_1)}{\partial^2 x_1} & \frac{\partial^2 x}{\partial x_1 \partial x_2} \\ \checkmark & \checkmark \end{pmatrix}^{2\times 2}$$

exists at each $x \in \mathbf{dom}\, f$

2nd-order conditions: for twice differentiable f with convex domain

- f is convex if and only if

If all the principal minor ≥ 0

$$A = \begin{pmatrix} 1 & 0 \\ 1 & 2 \end{pmatrix}$$

$$\nabla^2 f(x) \succeq 0 \quad \text{for all } x \in \mathbf{dom}\, f$$

P.S.D

Det(A) = 2

Ref: Page 647

- if $\nabla^2 f(x) \succ 0$ for all $x \in \mathbf{dom}\, f$, then f is strictly convex

Det(A_{11}) = 2

Det(A_{22}) = 1

Operations that preserve convexity

practical methods for establishing convexity of a function

1. verify definition (often simplified by restricting to a line)

2. for twice differentiable functions, show $\nabla^2 f(x) \succeq 0$

3. show that f is obtained from simple convex functions by operations that preserve convexity

   - nonnegative weighted sum
   - composition with affine function
   - pointwise maximum and supremum
   - composition
   - minimization
   - perspective

conv $\to$ C

$\underline{f_1 + f_2} \to$ convex.

H.W.

Determine if

$$f(x_1, x_2) = \frac{1}{x_1 x_2} \text{ on } R^2_+ \text{ is convex or concave.}$$

Derive the Hessian matrix.

$$\nabla^2 f(x_1, x_2) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f}{\partial x_1 \partial x_2} & \frac{\partial^2 f}{\partial x_2^2} \end{pmatrix} = \frac{1}{x_1 x_2} \begin{pmatrix} \boxed{2/x_1^2} & 1/x_1 x_2 \\ \frac{1}{x_1 x_2} & \boxed{2/x_2^2} \end{pmatrix}$$

Show that the positive definite matrix.

$$\text{Det}\left(\nabla^2 f(x_1, x_2)\right) = \frac{3}{x_1^4 x_2^4} > 0$$

The $f(x_1, x_2)$ is a convex function.

Ex. 2

X is a random variable.

X takes values $\{a_1, a_2, \ldots, a_n\}$

$P(X = a_i) = p_i, \; i = 1(1)n.$

Determine if $E(X)$ is a convex function of $p_1, \ldots, p_n$.

1. $E(X) = \underbrace{a_1 p_1 + \cdots + a_n p_n}_{\text{convex or not in } p}$.

2. $\mathrm{Prob}(X > a)$
$= \sum_{j=i}^{n} p_j, \; i = \min\{i \mid a_i \geq a\}$

3. $\mathrm{Prob}(c \leq X \leq d)$ is convex in p or not?

Positive weighted sum & composition with affine function

nonnegative multiple: αf is convex if f is convex, $\alpha \geq 0$

H.W

sum: $f_1 + f_2$ convex if f_1, f_2 convex (extends to infinite sums, integrals)

composition with affine function: $f(Ax + b)$ is convex if f is convex

examples

- log barrier for linear inequalities

$$f(x) = -\sum_{i=1}^{m} \log(b_i - a_i^T x), \qquad \mathbf{dom}\, f = \{x \mid a_i^T x < b_i, i = 1, \ldots, m\}$$

- (any) norm of affine function: $f(x) = \|Ax + b\|$

Pointwise maximum

if $f_1, \ldots, f_m$ are convex, then $f(x) = \max\{f_1(x), \ldots, f_m(x)\}$ is convex

f_1 convex
f_2 convex. $f = \max\{f_1(x), f_2(x)\}$

examples

- piecewise-linear function: $f(x) = \max_{i=1,\ldots,m}(a_i^T x + b_i)$ is convex
- sum of r largest components of $x \in \mathbf{R}^n$:

$$f(x) = x_{[1]} + x_{[2]} + \cdots + x_{[r]}$$

  is convex ($x_{[i]}$ is ith largest component of x)

  proof:

$$f(x) = \max\{x_{i_1} + x_{i_2} + \cdots + x_{i_r} \mid 1 \le i_1 < i_2 < \cdots < i_r \le n\}$$

Ref: Section 3.2.5 of Boyd's book.

$\text{dom} f_1$, $\text{dom} f = \text{dom} f_1 \cap \text{dom} f_2$

$\text{dom} f_2$,

$\underline{\underline{f(x)}} = \max\{f_1(x), f_2(x)\}$

$x \in \text{dom} f_1 \cap \text{dom} f_2$.

<u>Proof</u>:

$0 \leq \theta \leq 1$, $x, y \in \text{dom} f$.

$$f(\underline{\theta x + (1-\theta) y}) = \max\{\underline{f_1(\theta x + (1-\theta) y)},$$

$$\underline{f_2(\theta x + (1-\theta) y)}\}$$

$$\leq \max[\{\theta f_1(x) + (1-\theta) f(y)\}, \{\theta f_2(x) + (1-\theta) f_2(y)\}]$$

$$\leq \theta \{\overline{f_1(x)}, \overline{f_2(x)}\} + (1-\theta)\{f_1(y), f_2(y)\}$$

$$= \underline{\theta f(x)} + (1-\theta) f(y).$$

Pointwise supremum

if $f(x, y)$ is convex in x for each $y \in \mathcal{A}$, then

$$g(x) = \sup_{y \in \mathcal{A}} f(x, y)$$

is convex

examples

- support function of a set C: $S_C(x) = \sup_{y \in C} y^T x$ is convex
- distance to farthest point in a set C:

$$f(x) = \sup_{y \in C} \|x - y\|$$

- maximum eigenvalue of symmetric matrix: for $X \in \mathbf{S}^n$,

$$\lambda_{\max}(X) = \sup_{\|y\|_2 = 1} y^T X y$$

Minimization

if $f(x, y)$ is convex in (x, y) and C is a convex set, then

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

is convex

examples

- $f(x, y) = x^T A x + 2x^T B y + y^T C y$ with

$$\begin{bmatrix} A & B \\ B^T & C \end{bmatrix} \succeq 0, \qquad C \succ 0$$

  minimizing over y gives $g(x) = \inf_y f(x, y) = x^T (A - BC^{-1} B^T) x$

  g is convex, hence Schur complement $A - BC^{-1} B^T \succeq 0$

- distance to a set: $\mathbf{dist}(x, S) = \inf_{y \in S} \|x - y\|$ is convex if S is convex

Log-concave and log-convex functions

a positive function f is log-concave if $\log f$ is concave:

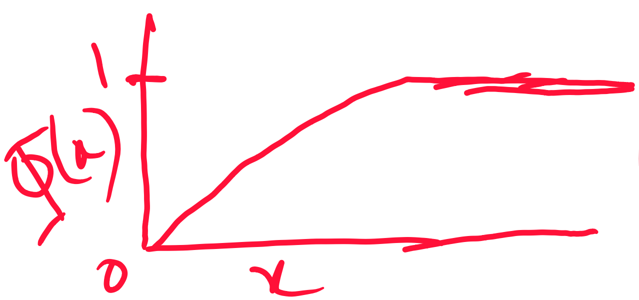
$$\boxed{f(\theta x + (1-\theta)y) \geq f(x)^\theta f(y)^{1-\theta}} \quad \text{for } 0 \leq \theta \leq 1$$

f is log-convex if $\log f$ is convex

- powers: x^a on $\mathbf{R}_{++}$ is log-convex for $a \leq 0$, log-concave for $a \geq 0$
- many common probability densities are log-concave, *e.g.*, normal:

$$f(x) = \frac{1}{\sqrt{(2\pi)^n \det \Sigma}} e^{-\frac{1}{2}(x-\bar{x})^T \Sigma^{-1}(x-\bar{x})}$$

- cumulative Gaussian distribution function Φ is log-concave



$$\left(\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-u^2/2}\, du\right)$$

x^2; e^x

1. S. T. $\Phi(x)$ is log concave:

Home work 2:

S.T. $f(x) = e^{ax}$ is both
$a \in \mathbb{R}, x > 0$.

log convex and log concave.

A is Positive definite

if $\underline{x}' \underline{A} \underline{x} > \underline{0} \ \forall \underline{x}$.