Convex set

line segment between x_1 and x_2: all points

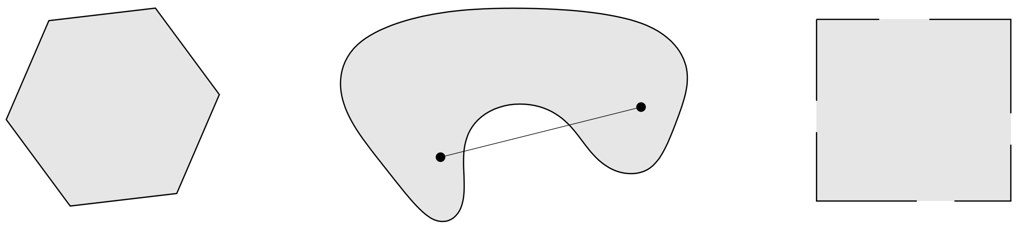
$$x = \theta x_1 + (1 - \theta) x_2$$

with $0 \leq \theta \leq 1$

convex set: contains line segment between any two points in the set

$$x_1, x_2 \in C, \quad 0 \leq \theta \leq 1 \quad \Longrightarrow \quad \theta x_1 + (1 - \theta) x_2 \in C$$

examples (one convex, two nonconvex sets)



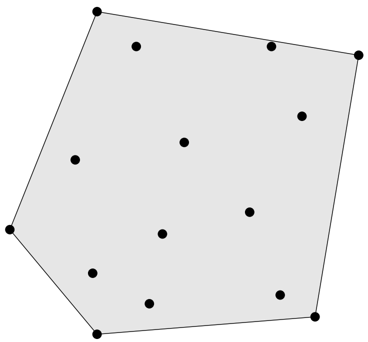
Convex combination and convex hull

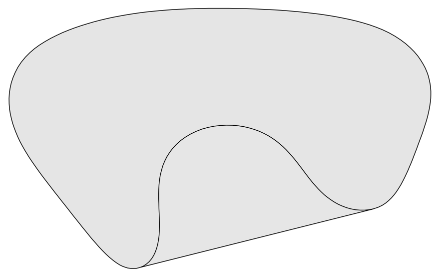
convex combination of x_1,. . . , x_k: any point x of the form

$$x = \theta_1 x_1 + \theta_2 x_2 + \cdots + \theta_k x_k$$

with $\theta_1 + \cdots + \theta_k = 1$, $\theta_i \geq 0$

convex hull $\mathbf{conv}\, S$: set of all convex combinations of points in S



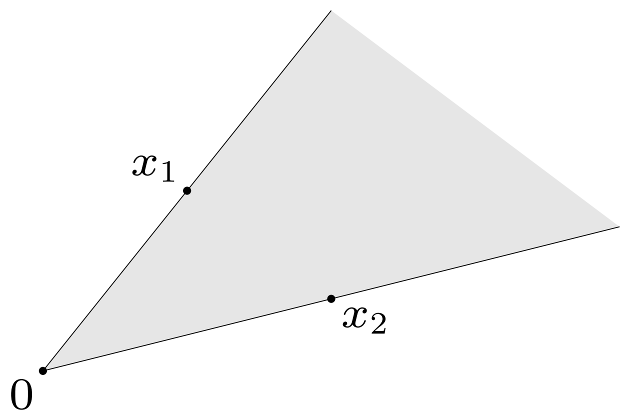


Convex cone

conic (nonnegative) combination of x_1 and x_2: any point of the form

$$x = \theta_1 x_1 + \theta_2 x_2$$

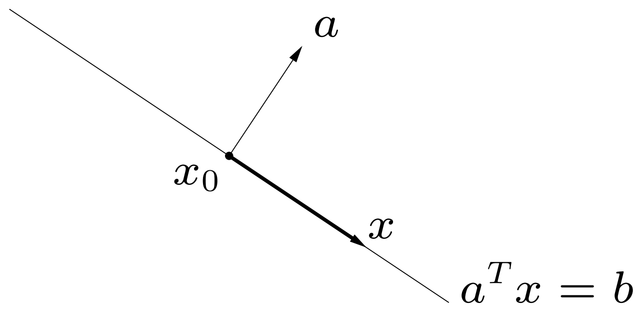
with $\theta_1 \geq 0$, $\theta_2 \geq 0$



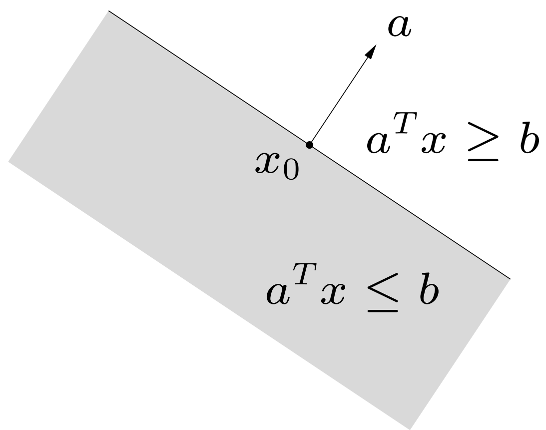
convex cone: set that contains all conic combinations of points in the set

Hyperplanes and halfspaces

hyperplane: set of the form $\{x \mid a^T x = b\}$ $(a \neq 0)$



halfspace: set of the form $\{x \mid a^T x \leq b\}$ $(a \neq 0)$



- a is the normal vector
- hyperplanes are affine and convex; halfspaces are convex

Euclidean balls and ellipsoids

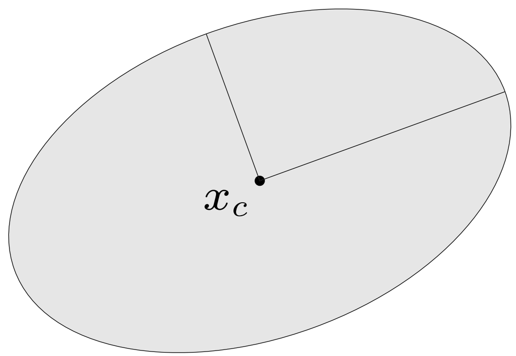
(Euclidean) ball with center x_c and radius r:

$$B(x_c, r) = \{x \mid \|x - x_c\|_2 \leq r\} = \{x_c + ru \mid \|u\|_2 \leq 1\}$$

ellipsoid: set of the form

$$\{x \mid (x - x_c)^T P^{-1} (x - x_c) \leq 1\}$$

with $P \in \mathbf{S}^n_{++}$ (*i.e.*, P symmetric positive definite)



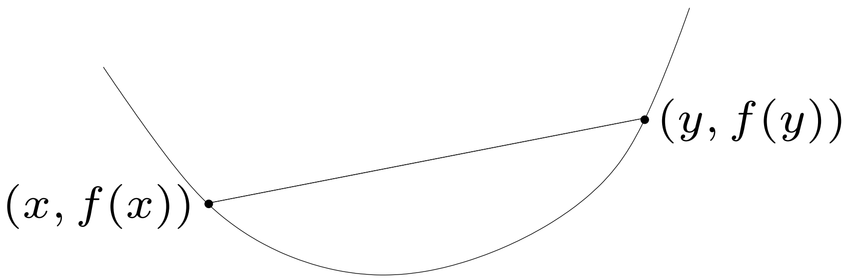
other representation: $\{x_c + Au \mid \|u\|_2 \leq 1\}$ with A square and nonsingular

Convex Functions
Definition

f : $\mathbf{R}^n \to \mathbf{R}$ is convex if

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

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



- f is concave if $-f$ is convex
- f is strictly convex if

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

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

Examples on R

convex:

- affine: $ax + b$ on **R**, for any $a, b \in$ **R**
- exponential: e^{ax}, for any $a \in$ **R**
- powers: x^α on $\mathbf{R}_{++}$, for $\alpha \geq 1$ or $\alpha \leq 0$
- powers of absolute value: $|x|^p$ on **R**, for $p \geq 1$
- negative entropy: $x \log x$ on $\mathbf{R}_{++}$

concave:

- affine: $ax + b$ on **R**, for any $a, b \in$ **R**
- powers: x^α on $\mathbf{R}_{++}$, for $0 \leq \alpha \leq 1$
- logarithm: $\log x$ on $\mathbf{R}_{++}$

Examples on $\mathbf{R}^n$ and $\mathbf{R}^{m\times n}$

affine functions are convex and concave; all norms are convex

examples on $\mathbf{R}^n$

- affine function $f(x) = a^T x + b$
- norms: $\|x\|_p = (\sum_{i=1}^n |x_i|^p)^{1/p}$ for $p \geq 1$; $\|x\|_\infty = \max_k |x_k|$

examples on $\mathbf{R}^{m\times n}$ ($m \times n$ matrices)

- affine function

$$f(X) = \mathbf{tr}(A^T X) + b = \sum_{i=1}^{m}\sum_{j=1}^{n} A_{ij}X_{ij} + b$$

- spectral (maximum singular value) norm

$$f(X) = \|X\|_2 = \sigma_{\max}(X) = (\lambda_{\max}(X^T X))^{1/2}$$

First-order condition

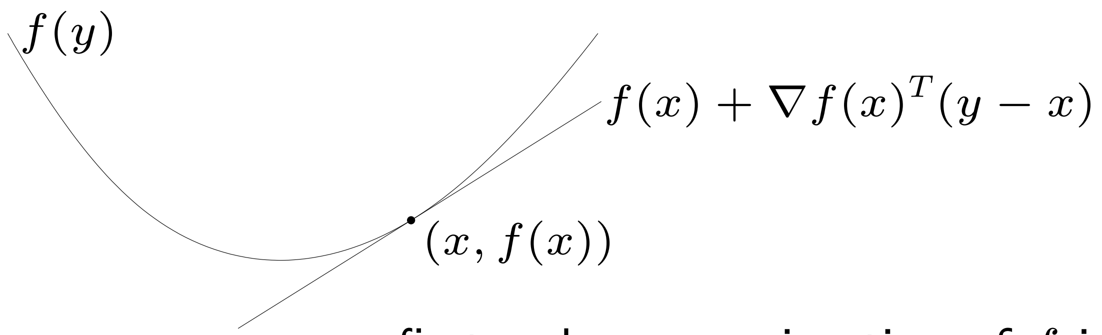
f is **differentiable** if $\mathbf{dom}\, f$ is open and the gradient

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

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 $\nabla^2 f(x) \in \mathbf{S}^n$,

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

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

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

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

Examples

quadratic function: $f(x) = (1/2)x^T Px + q^T x + r$ (with $P \in \mathbf{S}^n$)

$$\nabla f(x) = Px + q, \qquad \nabla^2 f(x) = P$$

convex if $P \succeq 0$

least-squares objective: $f(x) = \|Ax - b\|_2^2$

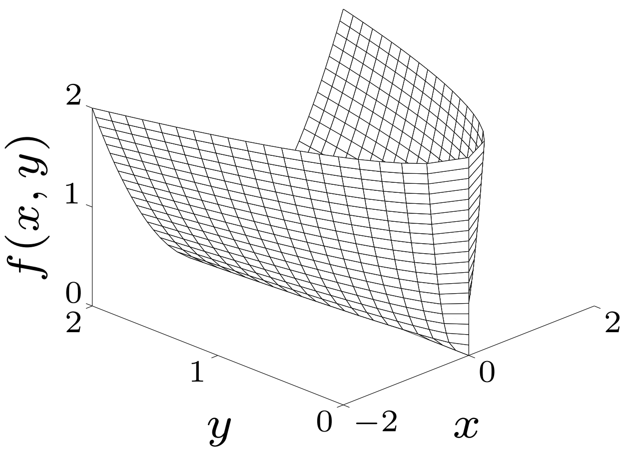
$$\nabla f(x) = 2A^T(Ax - b), \qquad \nabla^2 f(x) = 2A^T A$$

convex (for any A)

quadratic-over-linear: $f(x, y) = x^2/y$

$$\nabla^2 f(x, y) = \frac{2}{y^3} \begin{bmatrix} y \\ -x \end{bmatrix} \begin{bmatrix} y \\ -x \end{bmatrix}^T \succeq 0$$

convex for $y > 0$



Jensen's inequality

basic inequality: if f is convex, then for $0 \leq \theta \leq 1$,

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

extension: if f is convex, then

$$f(\mathbf{E}\, z) \leq \mathbf{E}\, f(z)$$

for any random variable z

basic inequality is special case with discrete distribution

$$\mathbf{prob}(z = x) = \theta, \qquad \mathbf{prob}(z = y) = 1 - \theta$$

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
   - composition

Positive weighted sum & composition with affine function

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

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

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:

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

Composition with scalar functions

composition of $g : \mathbf{R}^n \to \mathbf{R}$ and $h : \mathbf{R} \to \mathbf{R}$:

$$f(x) = h(g(x))$$

f is convex if $\begin{array}{l} g \text{ convex, } h \text{ convex, } \tilde{h} \text{ nondecreasing} \\ g \text{ concave, } h \text{ convex, } \tilde{h} \text{ nonincreasing} \end{array}$

- proof (for $n = 1$, differentiable g, h)

$$f''(x) = h''(g(x))g'(x)^2 + h'(g(x))g''(x)$$

examples

- $\exp g(x)$ is convex if g is convex
- $1/g(x)$ is convex if g is concave and positive