2. Convex sets

- affine and convex sets
- some important examples
- operations that preserve convexity
- generalized inequalities
- separating and supporting hyperplanes
- dual cones and generalized inequalities

Affine Set, Convex Set, Convex Conic Set

$x_1, x_2 \in C$, $\theta_1 x_1 + \theta_2 x_2 \in C$

Affine	Convex	Convex cone
$\theta_1 + \theta_2 = 1$	$\theta_1 + \theta_2 = 1$ $\theta_1, \theta_2 \geq 0$	$\theta_1, \theta_2 \geq 0$

① affine $\Rightarrow$ convex

② convex cone $\Rightarrow$ convex

③ convex $\not\Rightarrow$ affine

④ convex $\not\Rightarrow$ convex cone

Positive semidefinite cone

notation:

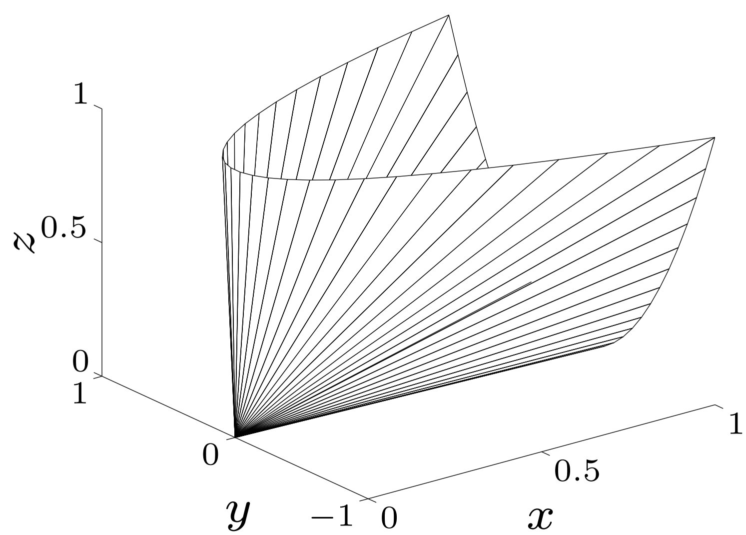
- $\mathbf{S}^n$ is set of symmetric $n \times n$ matrices
- $\mathbf{S}^n_+ = \{X \in \mathbf{S}^n \mid X \succeq 0\}$: positive semidefinite $n \times n$ matrices

$$X \in \mathbf{S}^n_+ \quad \Longleftrightarrow \quad z^T X z \geq 0 \text{ for all } z$$

  $\mathbf{S}^n_+$ is a convex cone

- $\mathbf{S}^n_{++} = \{X \in \mathbf{S}^n \mid X \succ 0\}$: positive definite $n \times n$ matrices

example: $\begin{bmatrix} x & y \\ y & z \end{bmatrix} \in \mathbf{S}^2_+$



Operations that preserve convexity

practical methods for establishing convexity of a set C

1. apply definition

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

2. show that C is obtained from simple convex sets (hyperplanes, halfspaces, norm balls, . . .) by operations that preserve convexity

   - intersection
   - affine functions
   - perspective function
   - linear-fractional functions

Intersection

$S_1, S_2,$ $S = S_1 \cap S_2$

the intersection of (any number of) convex sets is convex

$\bigcap_{z \neq 0} \{X \in S^n \mid z^T X z \geq 0\}$

$a^T x \geq b$

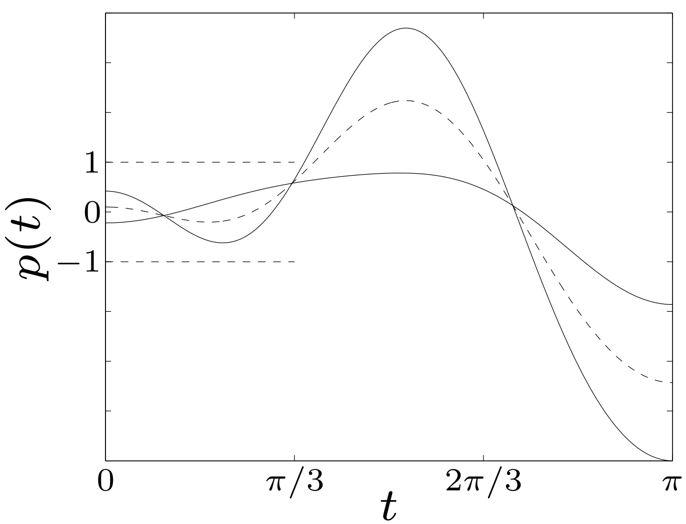
R^n

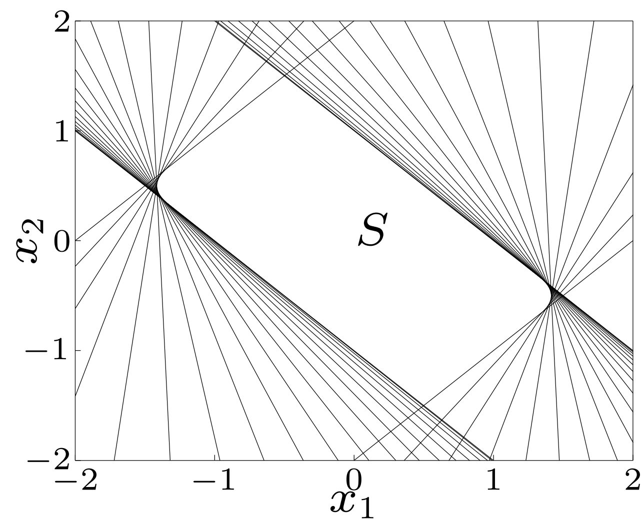
example:

$$S = \{x \in \mathbf{R}^m \mid |p(t)| \leq 1 \text{ for } |t| \leq \pi/3\}$$

where $p(t) = x_1 \cos t + x_2 \cos 2t + \cdots + x_m \cos mt$

for $m = 2$:





Affine function

suppose $f : \mathbf{R}^n \to \mathbf{R}^m$ is affine ($f(x) = Ax + b$ with $A \in \mathbf{R}^{m\times n}$, $b \in \mathbf{R}^m$)

- the image of a convex set under f is convex

$$S \subseteq \mathbf{R}^n \text{ convex} \quad \Longrightarrow \quad f(S) = \{f(x) \mid x \in S\} \text{ convex}$$

Proof: $f(x), f(y) \in f(S)$ $\qquad$ $x, y \in S$

$\underline{\theta f(x) + (1-\theta) f(y)} \in f(S)$ $\qquad$ $\theta x + (1-\theta) y \in S$

$= \theta(Ax+b) + (1-\theta)(Ay+b)$

$= A(\underline{\theta x + (1-\theta)y}) + b \in f(S)$

Affine function

suppose $f : \mathbf{R}^n \to \mathbf{R}^m$ is affine ($f(x) = Ax + b$ with $A \in \mathbf{R}^{m \times n}$, $b \in \mathbf{R}^m$)

- the image of a convex set under f is convex

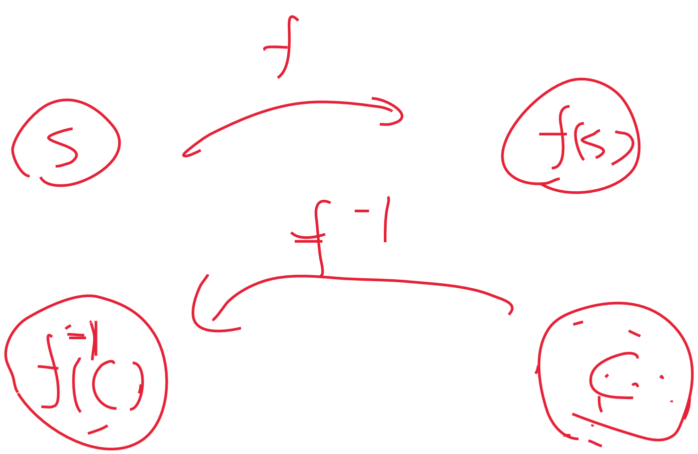
$$S \subseteq \mathbf{R}^n \text{ convex} \quad \Longrightarrow \quad f(S) = \{f(x) \mid x \in S\} \text{ convex}$$

- the inverse image $f^{-1}(C)$ of a convex set under f is convex

Quiz

$$C \subseteq \mathbf{R}^m \text{ convex} \quad \Longrightarrow \quad f^{-1}(C) = \{x \in \mathbf{R}^n \mid f(x) \in C\} \text{ convex}$$

$x, y \in f^{-1}(C)$, $\underline{\theta x + (1-\theta) y \in f^{-1}(C)}$



Affine function

suppose $f : \mathbf{R}^n \to \mathbf{R}^m$ is affine ($f(x) = Ax + b$ with $A \in \mathbf{R}^{m\times n}$, $b \in \mathbf{R}^m$)

- the image of a convex set under f is convex

$$S \subseteq \mathbf{R}^n \text{ convex} \quad \Longrightarrow \quad f(S) = \{f(x) \mid x \in S\} \text{ convex}$$

- the inverse image $f^{-1}(C)$ of a convex set under f is convex

$$C \subseteq \mathbf{R}^m \text{ convex} \quad \Longrightarrow \quad f^{-1}(C) = \{x \in \mathbf{R}^n \mid f(x) \in C\} \text{ convex}$$

$\{a + x \mid x \in S\}$ $\qquad f(x) = B - A(x)$, $\quad A(x) = \sum_{i=1}^{m} x_i A_i$

examples

$\{\alpha x \mid x \in S\}$ $\qquad \{x_1 \mid (x_1, x_2) \in S\}$ $\qquad \in S_+^n$ $\qquad \{x \in \mathbf{R}^n \mid B - A(x) \succeq 0\}$

- scaling, translation, projection
- solution set of linear matrix inequality $\{x \mid x_1A_1 + \cdots + x_mA_m \preceq B\}$ (with $A_i, B \in \mathbf{S}^p$) $\quad$ (LMI)
- hyperbolic cone $\{x \mid x^TPx \le (c^Tx)^2,\ c^Tx \ge 0\}$ (with $P \in \mathbf{S}_+^n$)

$\{(z, u) \mid \|z\|_2 \le u,\ u \ge 0\}$, $\begin{bmatrix} z \\ u \end{bmatrix}$ $\qquad f(x) = \begin{bmatrix} P^{\frac{1}{2}}x \\ c^Tx \end{bmatrix}$

Perspective and linear-fractional function

perspective function $P : \mathbf{R}^{n+1} \to \mathbf{R}^n$:

$$(x,t) = \begin{bmatrix} x_1 \\ \vdots \\ x_n \\ t \end{bmatrix} \in R^{n+1} \to \begin{bmatrix} x_1/t \\ \vdots \\ x_n/t \\ 1 \end{bmatrix} \to \begin{bmatrix} x_1/t \\ \vdots \\ x_n/t \end{bmatrix} \in R^n, \quad (t>0)$$

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \xrightarrow{P} \begin{bmatrix} x_1/x_3 \\ x_2/x_3 \end{bmatrix}, \quad x_3 > 0$$

$$P(x,t) = x/t, \qquad \mathbf{dom}\, P = \{(x,t) \mid t > 0\}$$

images and inverse images of convex sets under perspective are convex

$$f = P \circ g \qquad g(x) = \begin{bmatrix} A \\ c^T \end{bmatrix} x + \begin{bmatrix} b \\ d \end{bmatrix}, \quad A \in R^{m \times n}, \quad g: R^n \to R^{m+1}$$

$$P: R^{m+1} \to R^m$$

linear-fractional function $f : \mathbf{R}^n \to \mathbf{R}^m$:

$$f(x) = \frac{Ax+b}{c^Tx+d}, \qquad \mathbf{dom}\, f = \{x \mid c^Tx + d > 0\}$$

$$f = P \circ g$$

images and inverse images of convex sets under linear-fractional functions are convex

$$\underset{R^n}{\begin{bmatrix} \vdots \end{bmatrix}} \xrightarrow{g} \underset{R^{m+1}}{\begin{bmatrix} \vdots \end{bmatrix}} \xrightarrow{P} \underset{R^m}{\begin{bmatrix} \vdots \end{bmatrix}}$$

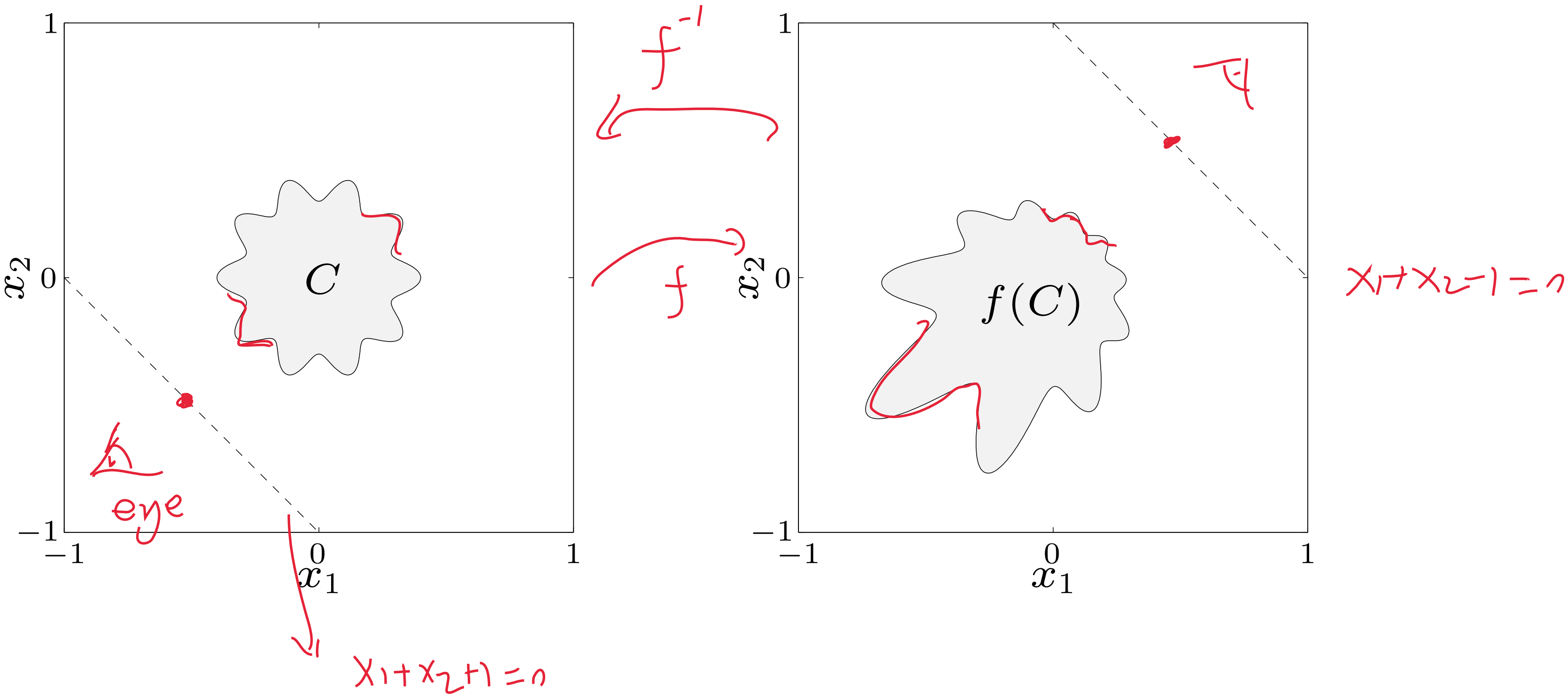
dom f: $x_1+x_2+1>0$

example of a linear-fractional function

$$f(x) = \frac{1}{x_1 + x_2 + 1} x$$

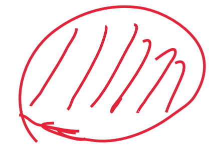
$A=1,\ b=0$

$c=\begin{bmatrix}1\\1\\0\\\vdots\\0\end{bmatrix},\ d=1$



$C = \mathrm{bd}\,C \cup \mathrm{int}\,C$ ($\mathrm{bd}\,C$, boundary)

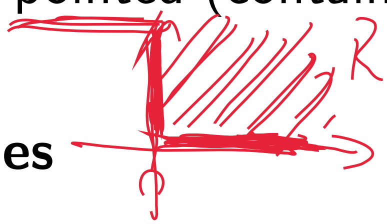
Generalized inequalities

$C =$  interior ($\mathrm{int}\,C$)

a convex cone $K \subseteq \mathbf{R}^n$ is a **proper cone** if

- K is closed (contains its boundary)
- K is solid (has nonempty interior)
- K is pointed (contains no line)

$\theta_1 x_1$ $\theta_1 x_1 + \theta_2 x_2 \in C$ $\theta_2 x_2$



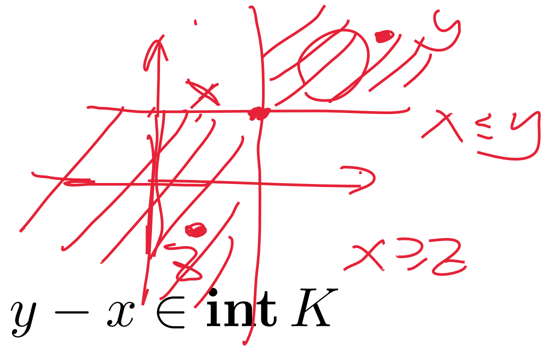
R^2_+ R_+

examples

- nonnegative orthant $K = \mathbf{R}^n_+ = \{x \in \mathbf{R}^n \mid x_i \geq 0,\ i = 1, \ldots, n\}$
- positive semidefinite cone $K = \mathbf{S}^n_+$
- nonnegative polynomials on $[0, 1]$:

$$K = \{x \in \mathbf{R}^n \mid x_1 + x_2 t + x_3 t^2 + \cdots + x_n t^{n-1} \geq 0 \text{ for } t \in [0, 1]\}$$

$(x \succeq_K 0 \iff x \in K)$ $K = R_+$



generalized inequality defined by a proper cone K:

$$x \preceq_K y \iff y - x \in K, \qquad x \prec_K y \iff y - x \in \textbf{int}\, K$$

$x \leq y \iff y - x \geq 0$
$y - x \in R_+$

examples

R^2_+

$x \leq_{R^2_+} y$
$(x \leq y)$
$x \not\leq z$
$x \not\geq z$

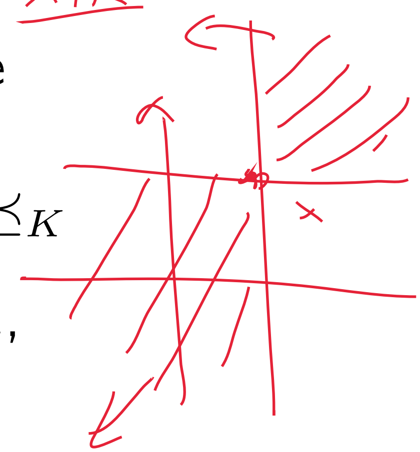
- componentwise inequality ($K = \mathbf{R}^n_+$)

$$x \preceq_{\mathbf{R}^n_+} y \iff x_i \leq y_i, \quad i = 1, \ldots, n$$

- matrix inequality ($K = \mathbf{S}^n_+$)

$$X \preceq_{\mathbf{S}^n_+} Y \iff Y - X \text{ positive semidefinite}$$

$x + K$



$X - K$

these two types are so common that we drop the subscript in $\preceq_K$

properties: many properties of $\preceq_K$ are similar to $\leq$ on $\mathbf{R}$, *e.g.*,

$$x \preceq_K y, \quad u \preceq_K v \implies x + u \preceq_K y + v$$

Minimum and minimal elements

$\preceq_K$ is not in general a *linear ordering*: we can have $x \not\preceq_K y$ and $y \not\preceq_K x$

$x \in S$ is **the minimum element** of S with respect to $\preceq_K$ if

$$y \in S \quad \Longrightarrow \quad x \preceq_K y$$

$S \subseteq x + K$

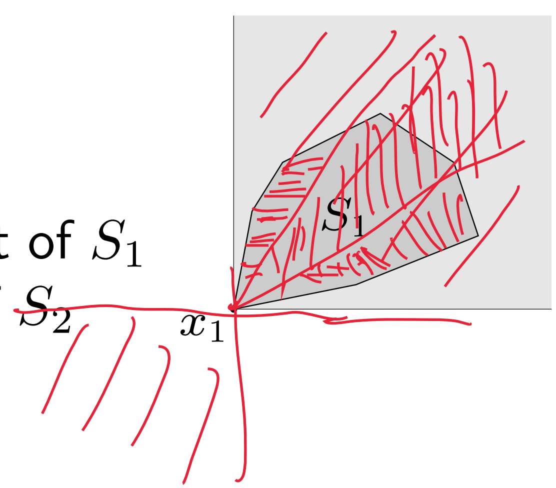
$x \in S$ is **a minimal element** of S with respect to $\preceq_K$ if

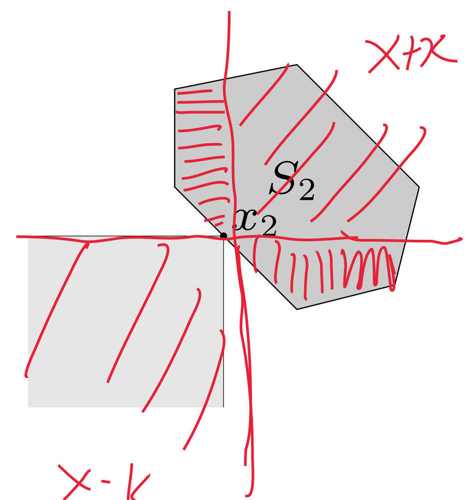
$$y \in S, \quad y \preceq_K x \quad \Longrightarrow \quad y = x$$

$S \cap (x - K) = \{x\}$

example $(K = \mathbf{R}^2_+)$

x_1 is the minimum element of S_1
x_2 is a minimal element of S_2



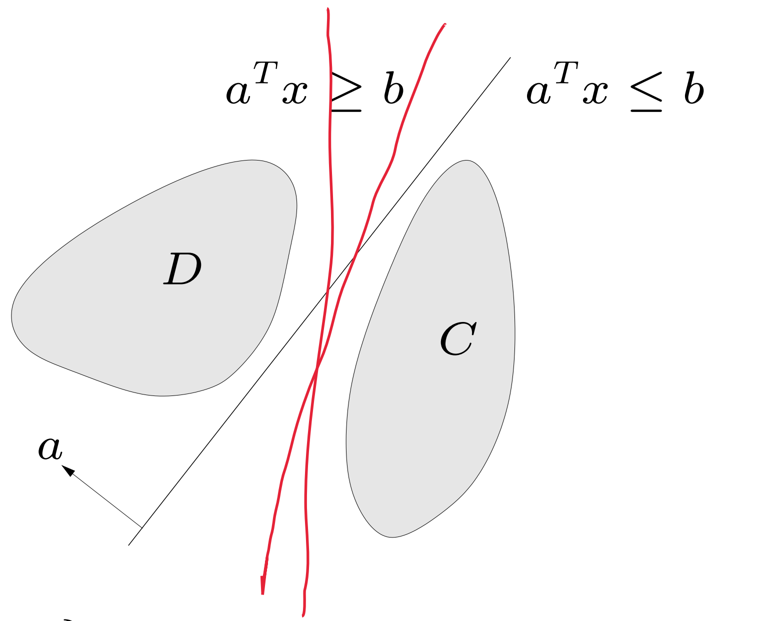


Separating hyperplane theorem

$C \cap D = \phi$

if C and D are nonempty disjoint convex sets, there exist $a \neq 0$, b s.t.

$$a^T x \leq b \text{ for } x \in C, \qquad a^T x \geq b \text{ for } x \in D$$



the hyperplane $\{x \mid a^T x = b\}$ separates C and D

strict separation requires additional assumptions ($e.g.$, C is closed, D is a singleton)

Supporting hyperplane theorem

supporting hyperplane to set C at boundary point x_0:

$$\{x \mid a^T x = a^T x_0\}$$

where $a \neq 0$ and $a^T x \leq a^T x_0$ for all $x \in C$

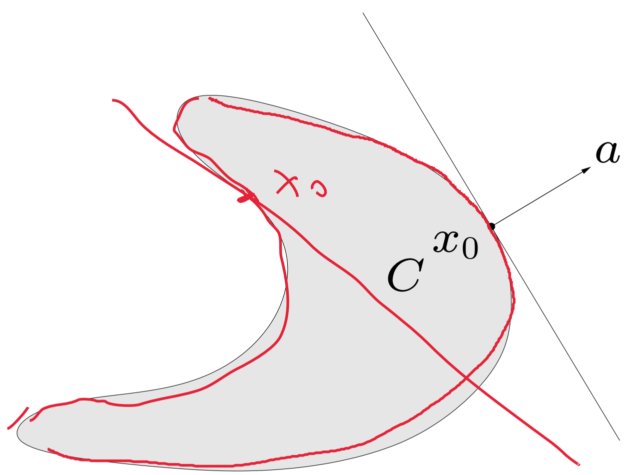
C : convex

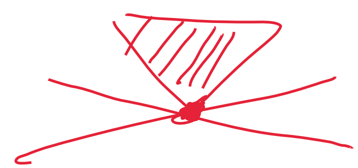
$x_0 \in \mathrm{bd}\, C$

int C

$\{x_0\} \cap \mathrm{int}\, C = \phi$

separating hyperplane





supporting hyperplane theorem: if C is convex, then there exists a supporting hyperplane at every boundary point of C

$x \succeq_K y \implies x - y \in K$

Dual cones and generalized inequalities

$x \in L,\ Ax \in C,\ (A \succeq 0)$

dual cone of a cone K:

$y \in K^* \iff y \succeq_{K^*} 0$ $\quad x \succeq_K 0$

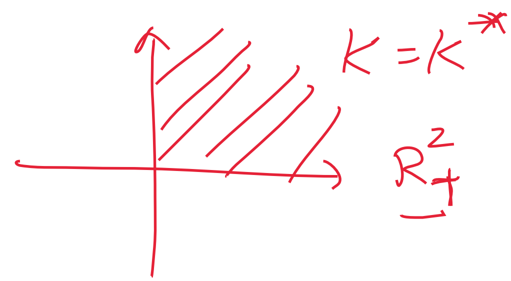
$$K^* = \{y \mid y^T x \geq 0 \text{ for all } x \in K\}$$

examples

$y^T \cdot x = \|y\| \cdot \|x\| \cos\theta$

$\theta \in [0, \frac{\pi}{2}]$

- $K = \mathbf{R}^n_+$: $K^* = \mathbf{R}^n_+$
- $K = \mathbf{S}^n_+$: $K^* = \mathbf{S}^n_+$
- $K = \{(x,t) \mid \|x\|_2 \leq t\}$: $K^* = \{(x,t) \mid \|x\|_2 \leq t\}$
- $K = \{(x,t) \mid \|x\|_1 \leq t\}$: $K^* = \{(x,t) \mid \|x\|_\infty \leq t\}$



first three examples are **self-dual** cones $(K^{**} = K$, if K is proper$)$

dual cones of proper cones are proper, hence define generalized inequalities:

$\forall y \in S$

$\lambda^T (y - x) \geq 0$

$$y \succeq_{K^*} 0 \iff y^T x \geq 0 \text{ for all } x \succeq_K 0$$

swap K^*, K

$\lambda^T (y - x) = 0$

$y \succeq_{K^*} x \iff y^T \lambda \geq x^T \lambda, \ \forall \lambda \succeq_K 0$

$y \succeq_K x \iff y^T \lambda \geq x^T \lambda, \ \forall \lambda \succeq_{K^*} 0$

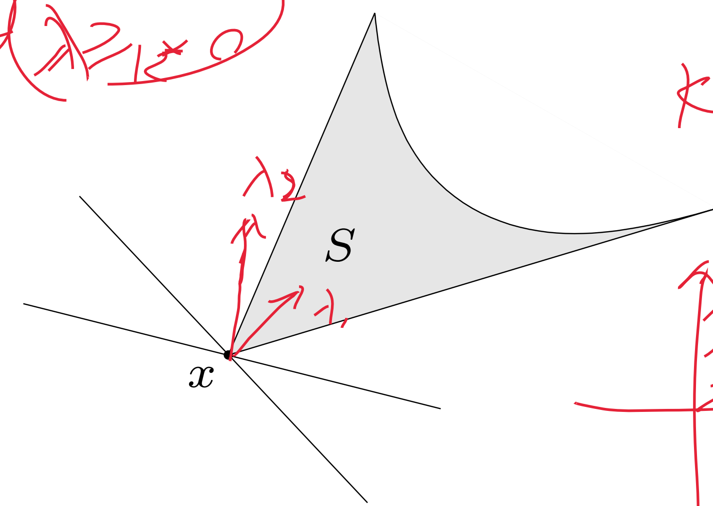
$\forall y \in S, \ x \preceq_K y$

Minimum and minimal elements via dual inequalities

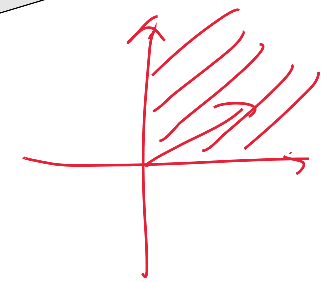
$\forall \lambda \succ_{K^*} 0$

minimum element w.r.t. $\preceq_K$

x is minimum element of S iff for all $\lambda \succ_{K^*} 0$, x is the unique minimizer of $\lambda^T z$ over S $(z \in S)$

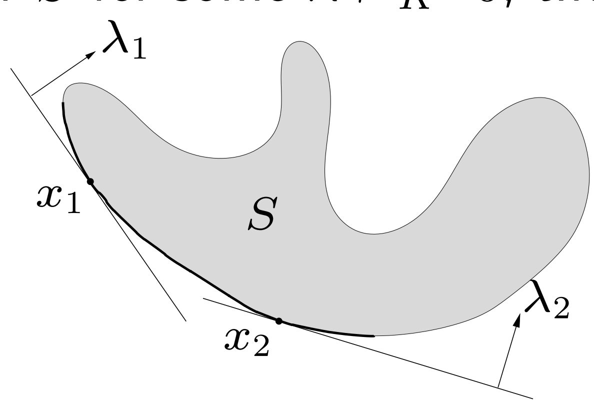


$K = \mathbb{R}^2_+$



minimal element w.r.t. $\preceq_K$

- if x minimizes $\lambda^T z$ over S for some $\lambda \succ_{K^*} 0$, then x is minimal



$K = \mathbb{R}^2_+$

- if x is a minimal element of a *convex* set S, then there exists a nonzero $\lambda \succeq_{K^*} 0$ such that x minimizes $\lambda^T z$ over S

(Pareto)

optimal production frontier

- different production methods use different amounts of resources $x \in \mathbf{R}^n$
- production set P: resource vectors x for all possible production methods
- efficient (Pareto optimal) methods correspond to resource vectors x that are minimal w.r.t. $\mathbf{R}^n_+$

$K^* = K = \mathbf{R}^2_+$

example $(n = 2)$

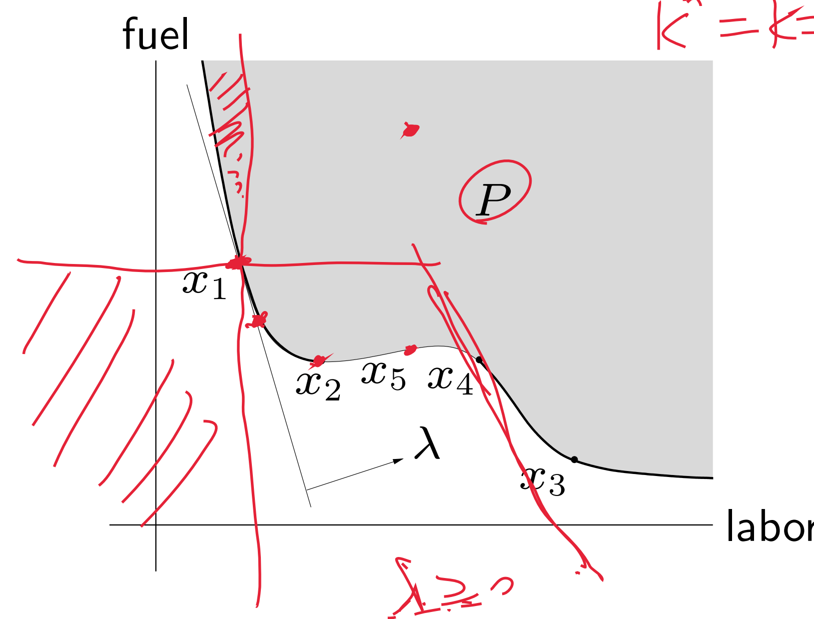
x_1, x_2, x_3 are efficient; x_4, x_5 are not

$x_1 \not\geq x_2$

$x_1 \not\leq x_2$

$x_2 \leq x_5, x_4$

$x_1 - K$

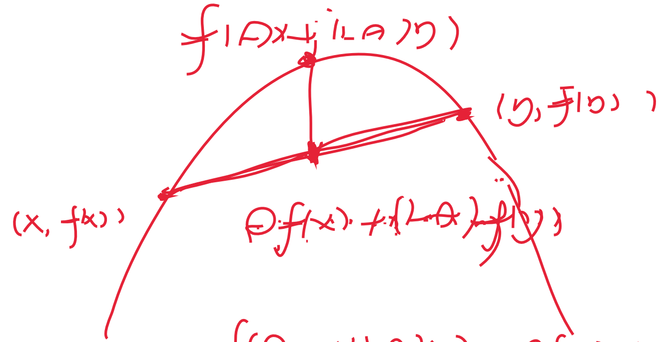


Multi-objective optimization.

3. Convex functions

- basic properties and examples
- operations that preserve convexity
- the conjugate function
- quasiconvex functions
- log-concave and log-convex functions
- convexity with respect to generalized inequalities

Definition



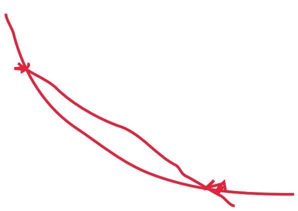
$f : \mathbf{R}^n \to \mathbf{R}$ is convex if $\mathbf{dom}\, f$ is a convex set and

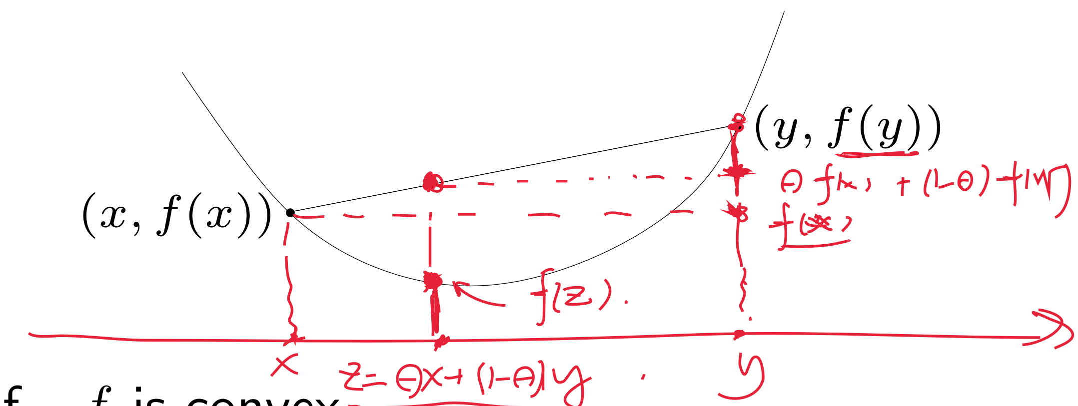
$$f(\theta x + (1-\theta)y) \leq \theta f(x) + (1-\theta) f(y)$$

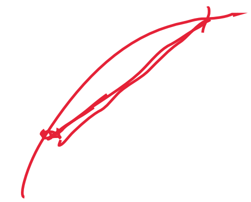
$f(\theta x+(1-\theta)y) \geq \theta f(x) + (1-\theta)f(y)$

$f(x)$: concave

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







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

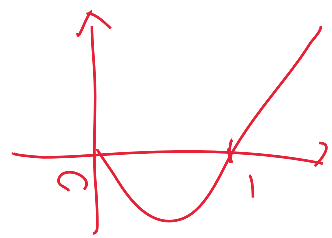
$$f(\theta x + (1-\theta)y) < \theta f(x) + (1-\theta) f(y)$$

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

Examples on R

convex:

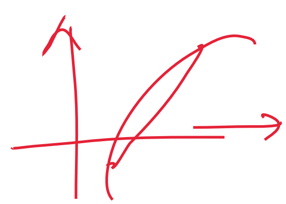
- affine: $ax + b$ on **R**, for any $a, b \in$ **R**
- exponential: e^{ax}, for any $a \in$ **R**
- powers: x^α on **R**$_{++}$, for $\alpha \geq 1$ or $\alpha \leq 0$
- powers of absolute value: $|x|^p$ on **R**, for $p \geq 1$
- ~~negative entropy~~: $x \log x$ on **R**$_{++}$



concave:

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

- logarithm: $\log x$ on **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 $f(x) = \overset{\langle a, x\rangle}{\underline{a^T x}} + b$
- 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

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

- spectral (maximum singular value) norm

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

Restriction of a convex function to a line

$f : \mathbf{R}^n \to \mathbf{R}$ is convex if and only if the function $g : \mathbf{R} \to \mathbf{R}$,

$\mathbf{R}^n$ $v \in \mathbf{R}^n$, direction (arbitrary).

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

$\mathbf{R}$

is convex (in t) for any $x \in \mathbf{dom}\, f$, $v \in \mathbf{R}^n$

can check convexity of f by checking convexity of functions of one variable

example. $f : \mathbf{S}^n \to \mathbf{R}$ with $f(X) = \log\det X$, $\mathbf{dom}\, f = \mathbf{S}^n_{++}$

$$\begin{aligned} g(t) = \log\det(X + tV) &= \log\det X + \log\det(I + tX^{-1/2}VX^{-1/2}) \\ &= \log\det X + \sum_{i=1}^{n} \log(1 + t\lambda_i) \end{aligned}$$

$\det(AB) = \det(A)\cdot\det(B)$

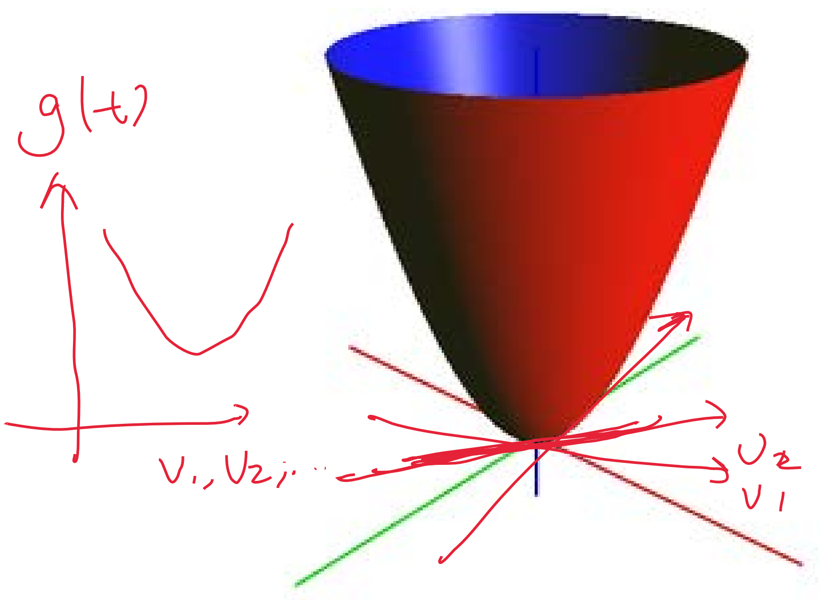
$X^{\frac{1}{2}}(I + tX^{-\frac{1}{2}}VX^{-\frac{1}{2}})X^{\frac{1}{2}}$

where λ_i are the eigenvalues of $X^{-1/2}VX^{-1/2}$

g is concave in t (for any choice of $X \succ 0$, V); hence f is concave

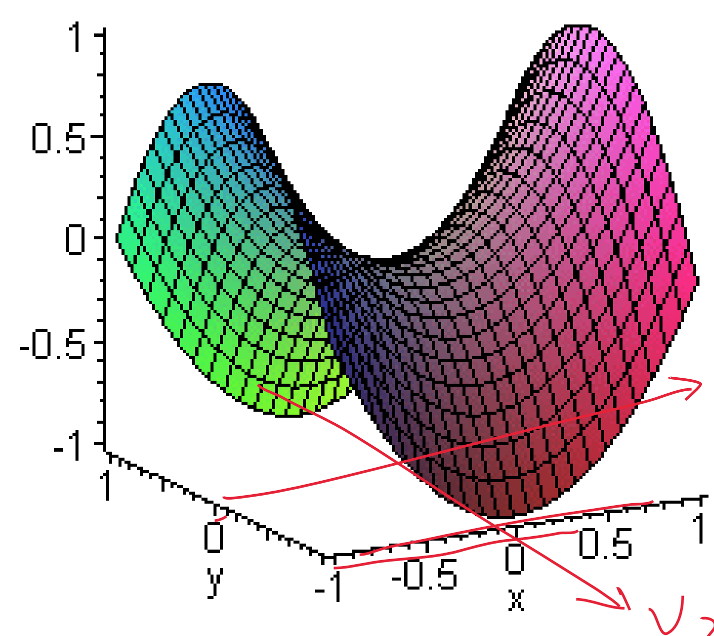
Restriction of a convex function to a line

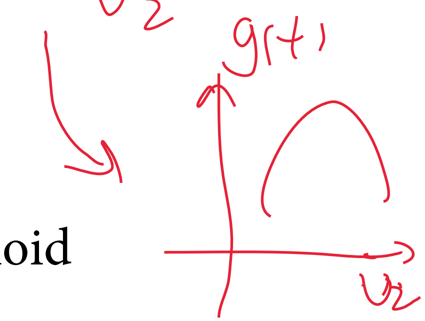
Restricting a function to a line
Draw a line in the domain of the function, and evaluate the function only along that line.



$$z = \frac{x^2}{a^2} + \frac{y^2}{b^2}.$$

Paraboloid





$$z = \frac{x^2}{a^2} - \frac{y^2}{b^2}$$

Hyperbolic paraboloid

https://math.stackexchange.com/questions/292693/what-does-it-mean-to-restricting-a-function-to-a-line-in-convex-optimization

Extended-value extension

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

extended-value extension $\tilde f$ of f is

$\mathrm{dom}\, f = [a, b]$

$$\tilde f(x) = f(x), \quad x \in \mathbf{dom}\, f, \qquad \tilde f(x) = \infty, \quad x \notin \mathbf{dom}\, f$$

often simplifies notation; for example, the condition

$$0 \le \theta \le 1 \quad \Longrightarrow \quad \tilde f(\theta x + (1-\theta)y) \le \theta \tilde f(x) + (1-\theta)\tilde f(y)$$

(as an inequality in $\mathbf{R} \cup \{\infty\}$), means the same as the two conditions

- $\mathbf{dom}\, f$ is convex
- for $x, y \in \mathbf{dom}\, f$,

$$0 \le \theta \le 1 \quad \Longrightarrow \quad f(\theta x + (1-\theta)y) \le \theta f(x) + (1-\theta) f(y)$$

First-order condition

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

$$\nabla f(x) = \left(\frac{\partial f(x)}{\partial x_1}, \frac{\partial f(x)}{\partial x_2}, \ldots, \frac{\partial f(x)}{\partial x_n} \right) \quad \in \mathbb{R}^n$$

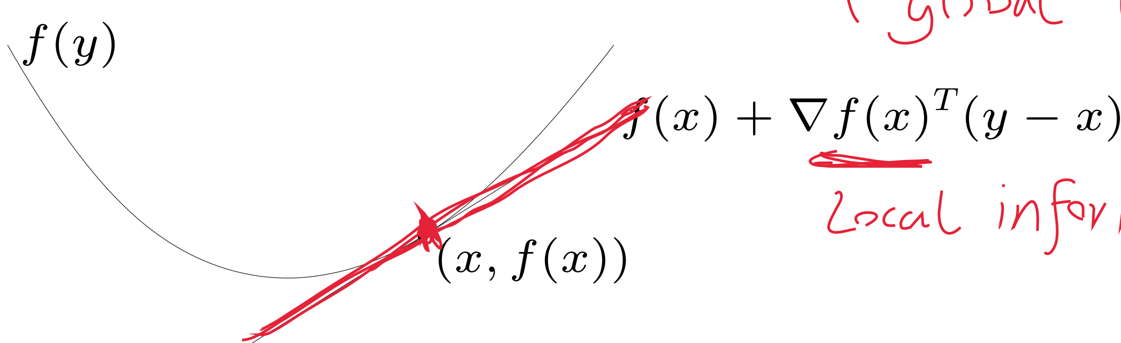
exists at each $x \in \mathbf{dom}\, f$

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

first-order Taylor approximation.

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

↑ global underestimator



Local information.

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

($\nabla^2 f(x) \in S^n_+$)

→ Hessen matrix is PSD

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

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

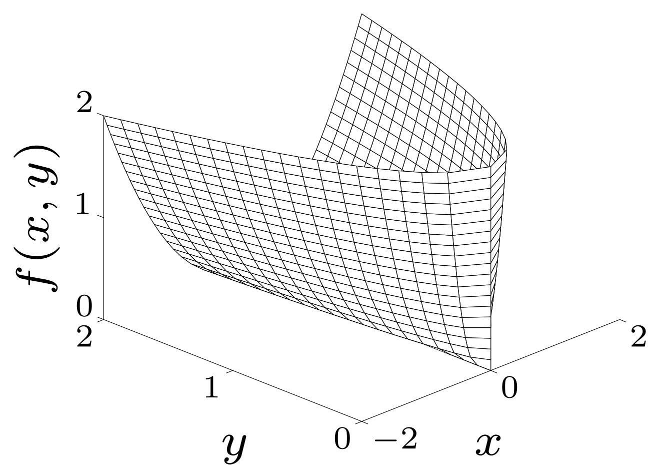
$$\nabla f(x) = 2A^T(Ax - b), \qquad \nabla^2 f(x) = 2A^T A$$

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 $y > 0$



log-sum-exp: $f(x) = \log \sum_{k=1}^{n} \exp x_k$ is convex

$$\nabla^2 f(x) = \frac{1}{\mathbf{1}^T z} \mathbf{diag}(z) - \frac{1}{(\mathbf{1}^T z)^2} z z^T \qquad (z_k = \exp x_k)$$

to show $\nabla^2 f(x) \succeq 0$, we must verify that $v^T \nabla^2 f(x) v \geq 0$ for all v:

$$v^T \nabla^2 f(x) v = \frac{(\sum_k z_k v_k^2)(\sum_k z_k) - (\sum_k v_k z_k)^2}{(\sum_k z_k)^2} \geq 0$$

since $(\sum_k v_k z_k)^2 \leq (\sum_k z_k v_k^2)(\sum_k z_k)$ (from Cauchy-Schwarz inequality)

geometric mean: $f(x) = (\prod_{k=1}^{n} x_k)^{1/n}$ on $\mathbf{R}_{++}^n$ is concave

(similar proof as for log-sum-exp)

Epigraph and sublevel set

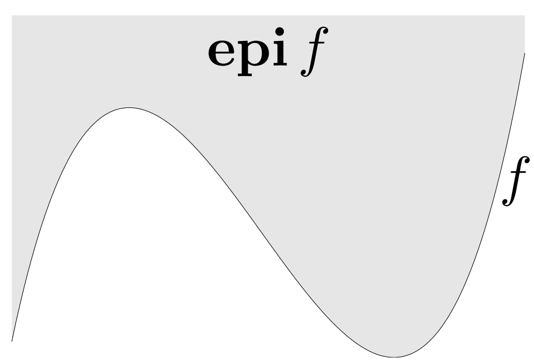
α-**sublevel set** of $f : \mathbf{R}^n \to \mathbf{R}$:

$$C_\alpha = \{x \in \mathbf{dom}\, f \mid f(x) \leq \alpha\}$$

sublevel sets of convex functions are convex (converse is false)

epigraph of $f : \mathbf{R}^n \to \mathbf{R}$:

$$\mathbf{epi}\, f = \{(x,t) \in \mathbf{R}^{n+1} \mid x \in \mathbf{dom}\, f,\ f(x) \leq t\}$$



f is convex if and only if $\mathbf{epi}\, f$ is a convex set

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 and supremum
   - composition
   - minimization
   - perspective

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 $f_1, \ldots, f_m$ 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:

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

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 $X \in \mathbf{S}^n$,

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

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)

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

- note: monotonicity must hold for extended-value extension $\tilde{h}$

examples

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

Vector composition

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

$$f(x) = h(g(x)) = h(g_1(x), g_2(x), \ldots, g_k(x))$$

f is convex if $\begin{array}{l} g_i \text{ convex, } h \text{ convex, } \tilde{h} \text{ nondecreasing in each argument} \\ g_i \text{ concave, } h \text{ convex, } \tilde{h} \text{ nonincreasing in each argument} \end{array}$

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

$$f''(x) = g'(x)^T \nabla^2 h(g(x)) g'(x) + \nabla h(g(x))^T g''(x)$$

examples

- $\sum_{i=1}^m \log g_i(x)$ is concave if g_i are concave and positive
- $\log \sum_{i=1}^m \exp g_i(x)$ is convex if g_i are convex

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

Perspective

the **perspective** of a function $f : \mathbf{R}^n \to \mathbf{R}$ is the function $g : \mathbf{R}^n \times \mathbf{R} \to \mathbf{R}$,

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

g is convex if f is convex

examples

- $f(x) = x^T x$ is convex; hence $g(x,t) = x^T x/t$ is convex for $t > 0$
- negative logarithm $f(x) = -\log x$ is convex; hence relative entropy $g(x,t) = t\log t - t\log x$ is convex on $\mathbf{R}^2_{++}$
- if f is convex, then

$$g(x) = (c^T x + d) f\left((Ax+b)/(c^T x + d)\right)$$

is convex on $\{x \mid c^T x + d > 0,\ (Ax+b)/(c^T x + d) \in \mathbf{dom}\, f\}$