강의자료 추가부분

Convex Optimization

for Machine Learners

Hyunjung (Helen) Shin

Dept. of Industrial Engineering
Ajou University, Korea

shin@ajou.ac.kr
http://www.alphaminers.net
http://www.kyb.tuebingen.mpg.de/~shin

Convex Functions

Convex Optimization: Convexity Test

Hessian Matrix

The **second-order gradient** of the **twice differentiable** real function $\boldsymbol{f}: \boldsymbol{R}^n \rightarrow \boldsymbol{R}$, w.r.t. its vector argument

$$\nabla^2 f(x) = \begin{bmatrix} \frac{\partial^2 f(x)}{\partial x_1^2} & \frac{\partial^2 f(x)}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f(x)}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f(x)}{\partial x_2 \partial x_1} & \frac{\partial^2 f(x)}{\partial x_2^2} & \cdots & \frac{\partial^2 f(x)}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f(x)}{\partial x_n \partial x_1} & \frac{\partial^2 f(x)}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f(x)}{\partial x_n^2} \end{bmatrix}$$

Convex Optimization: Convexity Test

Convexity Test with Hessian

A function $\boldsymbol{f}$ is **convex** *if and only if* its $\boldsymbol{n} \times \boldsymbol{n}$ **Hessian matrix** is **positive semidefinite** for all possible values of $\boldsymbol{x} \in \boldsymbol{R}^{\boldsymbol{n}}$

Convex Optimization: Convexity Test

Convexity Test with Hessian

Quantity	Convex	Strictly Convex	Concave	Strictly Concave
$\frac{\partial^2 f(x_1,x_2)}{\partial x_1^2}\frac{\partial^2 f(x_1,x_2)}{\partial x_2^2}-\left[\frac{\partial^2 f(x_1,x_2)}{\partial x_1 \partial x_2}\right]^2$	$\geq \mathbf{0}$	$> \mathbf{0}$	$\geq \mathbf{0}$	$> \mathbf{0}$
$\frac{\partial^2 f(x_1,x_2)}{\partial x_1^2}$	$\geq \mathbf{0}$	$> \mathbf{0}$	$\leq \mathbf{0}$	$< \mathbf{0}$
$\frac{\partial^2 f(x_1,x_2)}{\partial x_2^2}$	$\geq \mathbf{0}$	$> \mathbf{0}$	$\leq \mathbf{0}$	$< \mathbf{0}$

Convex Optimization: Convexity Test

Convexity Test with Hessian: Ex)

$$f(\boldsymbol{x}) = (x_1 - x_2)^2 = x_1^2 - 2x_1x_2 + x_2^2$$

$$\frac{\partial f(x_1, x_2)}{\partial x_1} = 2x_1 - 2x_2,$$

$$\frac{\partial f(x_1, x_2)}{\partial x_2} = -2x_1 + 2x_2$$

$$\nabla^2 f(x) = \begin{bmatrix} \frac{\partial^2 f(x_1,x_2)}{\partial x_1^2} & \frac{\partial^2 f(x_1,x_2)}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f(x_1,x_2)}{\partial x_2 \partial x_1} & \frac{\partial^2 f(x_1,x_2)}{\partial x_2^2} \end{bmatrix} = \begin{bmatrix} 2 & -2 \\ -2 & 2 \end{bmatrix}$$

Convex Optimization: Convexity Test

Convexity Test with Hessian: Ex)

$$f(\boldsymbol{x}) = (x_1 - x_2)^2 = x_1^2 - 2x_1x_2 + x_2^2$$

(1) $\dfrac{\partial^2 f(x_1,x_2)}{\partial x_1^2} \dfrac{\partial^2 f(x_1,x_2)}{\partial x_2^2} - \left[\dfrac{\partial^2 f(x_1,x_2)}{\partial x_1 \partial x_2}\right]^2 = 2(2) - (-2)^2 = 0$

(2) $\dfrac{\partial^2 f(x_1,x_2)}{\partial x_1^2} = 2 > 0$

(3) $\dfrac{\partial^2 f(x_1,x_2)}{\partial x_2^2} = 2 > 0$

Since ≥ 0 holds for all three conditions $f(\boldsymbol{x})$ is **convex**.

However, it is ***not*** strictly convex because the first condition only gives $= 0$ rather than > 0.

Convex Optimization: Convexity Test

1st-order condition

Differentiable f with convex domain is **convex** *iff*

$$\boldsymbol{f}(\boldsymbol{y}) \geq \boldsymbol{f}(\boldsymbol{x}) + \nabla \boldsymbol{f}(\boldsymbol{x})^{\boldsymbol{T}}(\boldsymbol{y} - \boldsymbol{x})$$

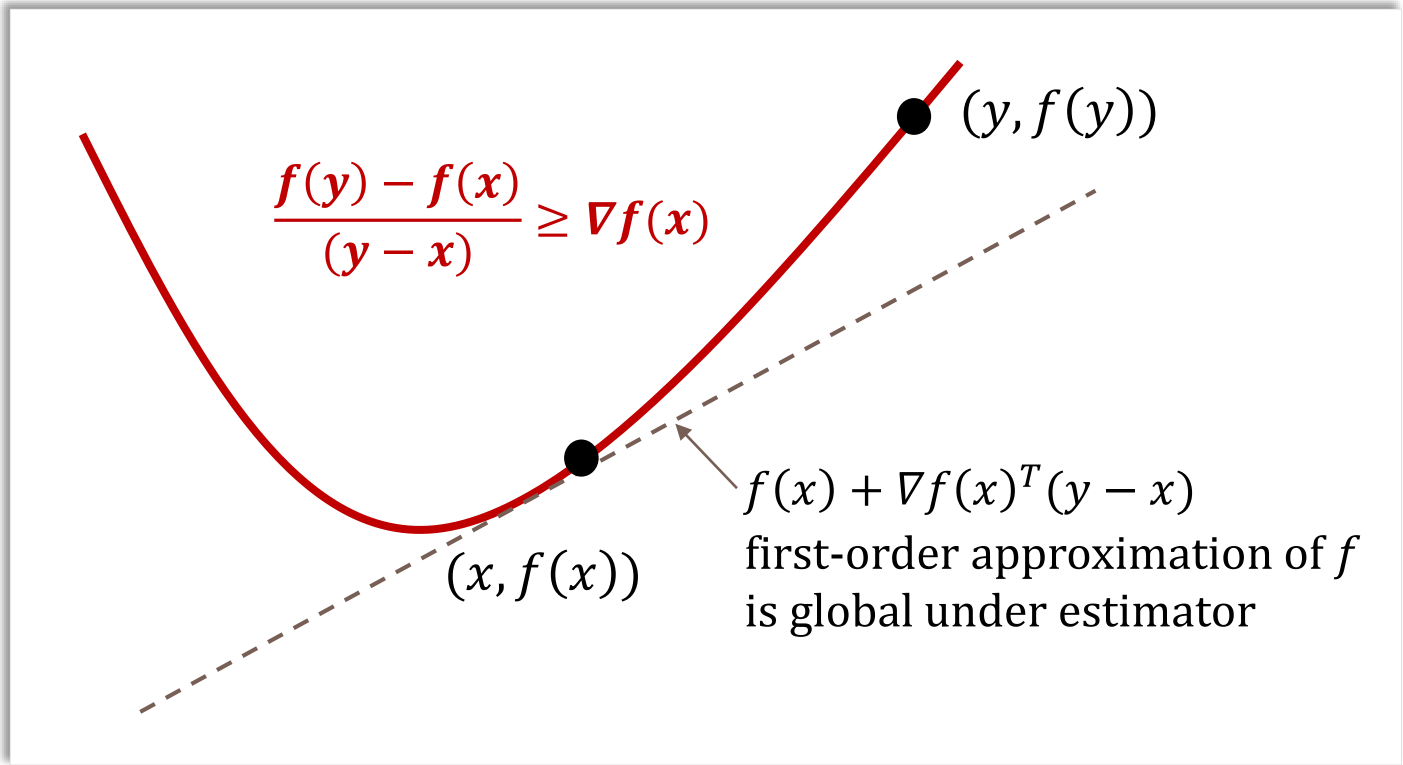
for all $\boldsymbol{x}, \boldsymbol{y} \in dom\ f$

f is **differentiable** if $dom\ f$ is open and the gradient exists at each $\boldsymbol{x} \in dom\ f$

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

Convex Optimization: Convexity Test

1st-order condition



Convex Optimization: Convexity Test

2nd-order condition

For **twice differentiable** f with convex domain f is **convex** *iff*

$$\boldsymbol{\nabla^2 f(x) \succcurlyeq 0}$$

for all $\boldsymbol{x} \in dom\ f$.

If $\boldsymbol{\nabla^2 f(x) \succ 0}$ for all $\boldsymbol{x} \in dom\ f$, then f is **strictly convex**.

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

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

exists at each $\boldsymbol{x} \in dom\ f$.

Convex Optimization: Convexity Test

Quadratic function: $f(x) = \frac{1}{2}x^T P x + q^T + r$ (with $P \in S^n$)

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

Convex !
if $\boldsymbol{P} \geq \mathbf{0}$

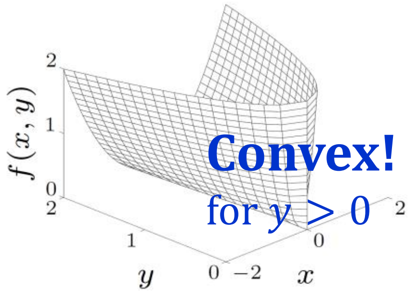
Least-Squares objective: $f(x) = \|Ax - b\|_2^2$

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

Convex!
(for any **A**)

Quadratic-over-linear: $f(x, y) = {x^2}/{y}$

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



Convex!
for $y > 0$

Convex Optimization: Convex functions

Practical methods for establishing convexity of a function

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

2. For **twice differentiable** functions, show $\boldsymbol{\nabla^2 f(x) \succcurlyeq 0}$

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

Convex Optimization: Simple Convex functions

Basic examples

- $\boldsymbol{x^p}\ for\ p \geq 1\ or\ p \leq 0; \quad -\boldsymbol{x^p}\ for\ 0 \leq p \leq 1$
- $\boldsymbol{e^x}, \quad -\mathbf{log}\,\boldsymbol{x}, \quad \boldsymbol{x\,log\,x}$
- $\boldsymbol{x^T x}, \quad \boldsymbol{x^T x}/\boldsymbol{y}\ (for\ y > 0), \quad (\boldsymbol{x^T x})^{\mathbf{1/2}}$
- $\|\ \boldsymbol{x}\ \|$ (*any norm*)
- $\mathbf{max}\,(\boldsymbol{x_1}, \cdots, \boldsymbol{x_n}), \quad \mathbf{log}(\,\boldsymbol{e^{x_1}} + \cdots + \boldsymbol{e^{x_n}})$
- $\mathbf{log}\,\boldsymbol{\Phi}(\boldsymbol{x})$ $(\Phi\ is\ Gaussian\ CDF\,)$
- $\mathbf{log}\ \ \boldsymbol{det\,X^{-1}}\ \ (\,for\ X \succ 0\,)$

Convex Optimization: Simple Convex functions

More examples

- $\boldsymbol{\lambda_{max}(X)}$ $\quad (for\ X = X^T)$
- $\boldsymbol{f(x) = x_{[1]} + \cdots + x_{[k]}}$ $\quad (sum\ of\ largest\ k\ elements\ of\ x)$
- $\boldsymbol{-\sum_{i=1}^{m} \log\left(-f_i(x)\right)}$ $\quad (\text{on } \{x | f_i(x) < 0\};\ f_i\ is\ convex)$
- $\boldsymbol{f(x) = \log \mathrm{Prob}\,(x + z \in C)}$ $\quad (C\ is\ convex, z \sim \mathcal{N}(0,\Sigma))$
- $\boldsymbol{x^T Y^{-1}}$ $\quad (x\ is\ convex\ in\ (x, Y)\ for\ Y = Y^T \succ 0)$

Convex Optimization: Convex functions

Calculus Operations

Convexity preserved under...

Sums, nonnegative scaling: $if\ f\ is\ convex,$ $then\ g(x) = f(Ax + b)\ is\ convex$

Pointwise sup: $if\ f_a\ is\ convex\ for\ each\ \alpha \in A,$ $then\ g(x) = sup_{a \epsilon A} f_a(x)\ is\ convex$

Minimization: $if\ f(x, y)\ is\ convex,$ $then\ g(x) = \inf_y f(x, y)\ is\ convex$

Composition rules: $if\ h\ is\ convex\ \&\ increasing, f\ is\ convex,$ $then\ g(x) = h\big(f(x)\big)\ is\ convex$

Perspective transformation: $if\ f\ is\ convex,$ $then\ g(x, t) = tf(x/t)\ is\ convex\ for\ t > 0$

. . . and many, many others

Convex Sets

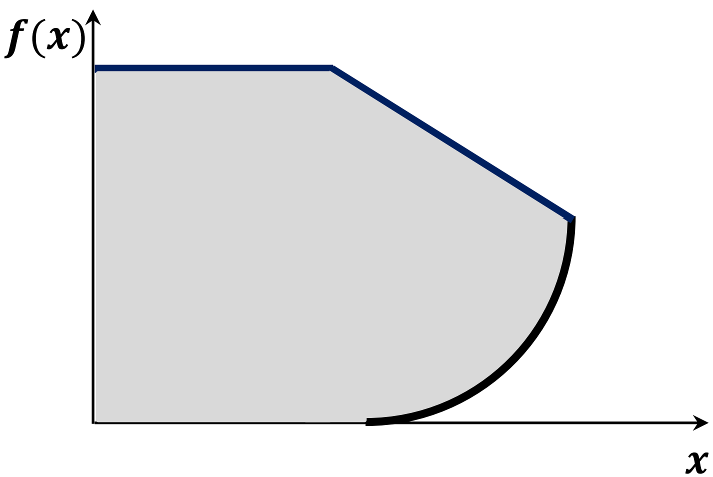
Convex Optimization: Convex Set

Convex Set

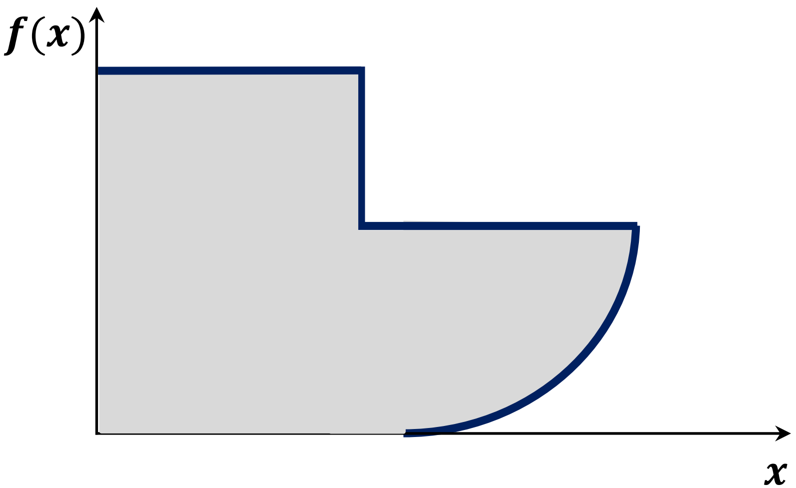
A **convex set** is a collection of points that, for each **pair of points** in the collection, the **entire line segment** joining these two points is **also in the collections.**

A set $\boldsymbol{C}$ is **convex** *iff* for all $\boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{C}$ and $0 \leq \alpha \leq 1$,

$$\alpha \boldsymbol{x} + (1 - \alpha)\boldsymbol{y} \in \boldsymbol{C}$$



A convex set



A set that is not convex

Convex Optimization: Convex Set

Practical methods for establishing convexity of a set $\boldsymbol{C}$

1. Apply **definition**

$$\alpha \boldsymbol{x} + (1-\alpha)\boldsymbol{y} \in \boldsymbol{C}, \qquad \boldsymbol{x}, \boldsymbol{y} \in \boldsymbol{C} \text{ and } 0 \le \alpha \le 1$$

2. Show that $\boldsymbol{C}$ is obtained from **simple convex sets**
   :**hyperplanes**, **half-spaces**, **norm balls**, . . .

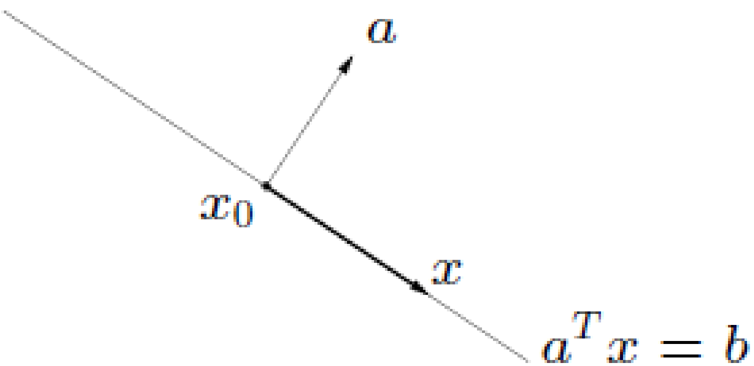
3. **Operations** that **preserve convexity**

   Intersection
   Affine functions
   Perspective function
   Linear-fractional functions

Convex Optimization: Simple Convex Set

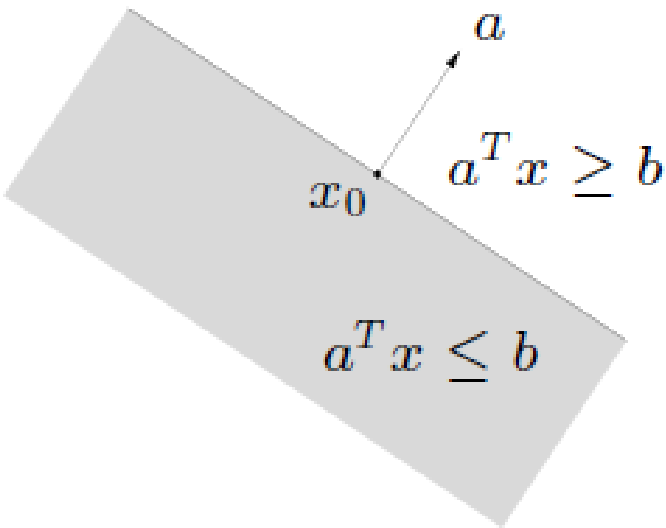
Hyperplane:
set of the form $\{x | a^T x = b\} (a \neq 0)$

a is the normal vector



Hyperplanes are **affine** and **convex**

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



Halfspaces are **convex**

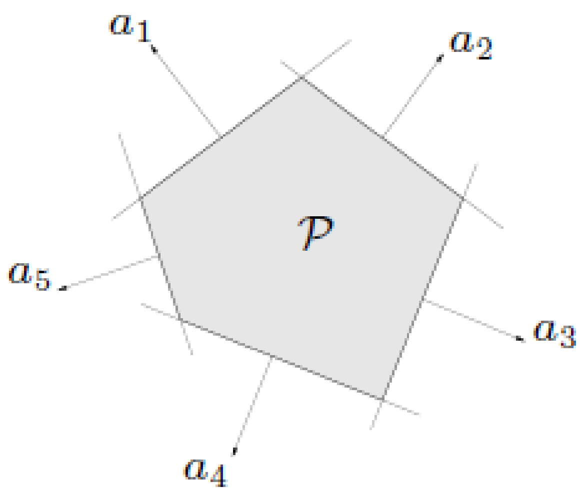
Convex Optimization: Simple Convex Set

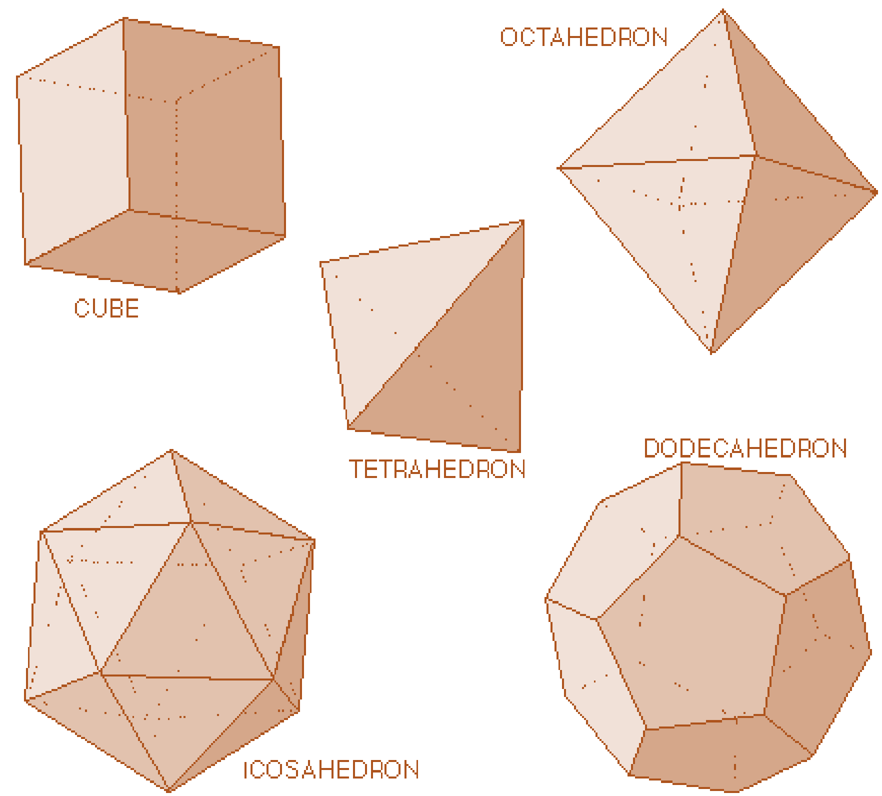
Polyhedra

Solution set of finitely many **linear inequalities** and **equalities**

$$Ax \preccurlyeq b, \qquad Cx = d$$

($A \in \mathbb{R}^{m \times n}, C \in \mathbb{R}^{p \times n}$,
$\preccurlyeq$is componentwise inequality)





Polyhedron is **intersection** of finite number of **halfspaces** and **hyperplanes**

Convex Optimization: Simple Convex Set

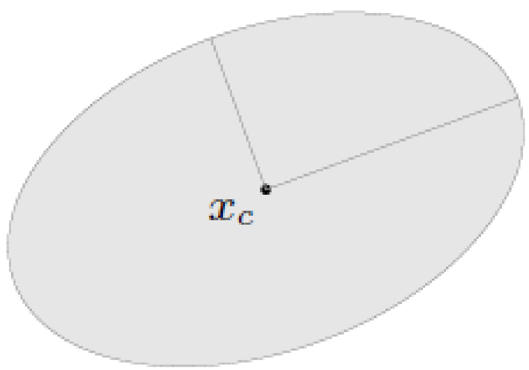
(Euclidean) ball with center x_c and radius r:

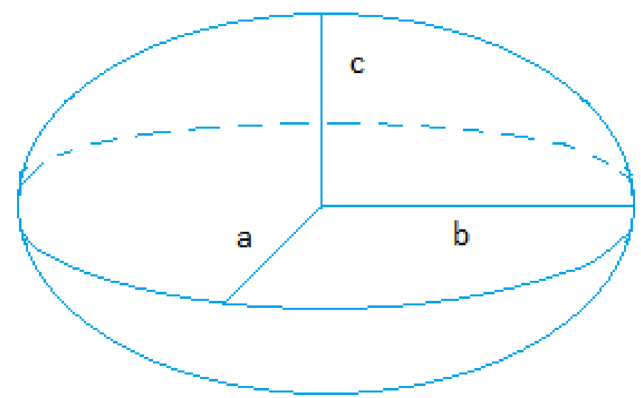
$$B(x_c, r) = \{x \mid \|x - x_c\|_2 \le r\} = \{x_c + ru \mid \|u\|_2 \le 1\}$$

Ellipsoid: set of the form

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

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



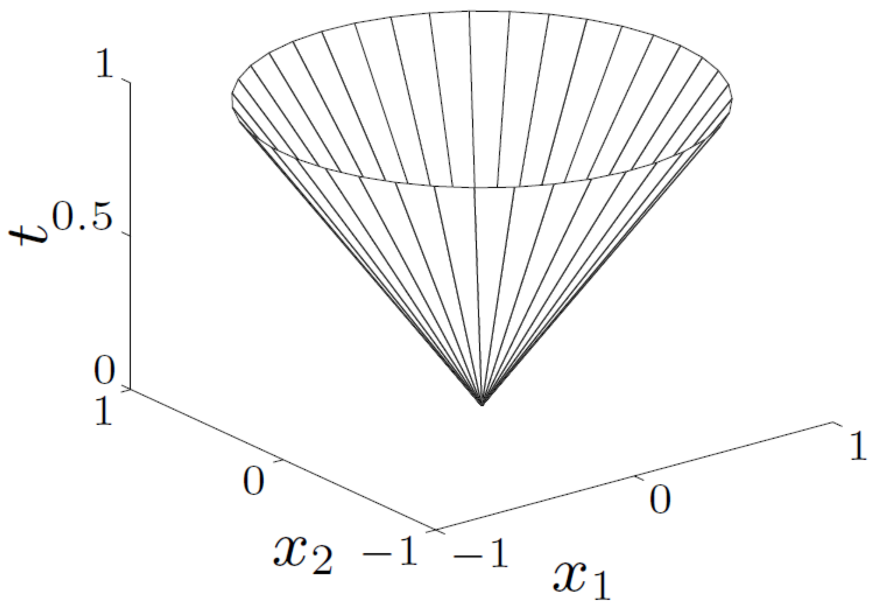


Other representation: $\{x_c + Au \mid \|u\|_2 \le 1\}$ with A square and nonsingular

Convex Optimization: Simple Convex Set

Norm cone

$\{(x, t) \mid \|x\| \leq t\}$



Norm cones are **convex**

Euclidean norm cone is called **second-order cone**

Convex Optimization: Simple Convex Set

PSD Cone

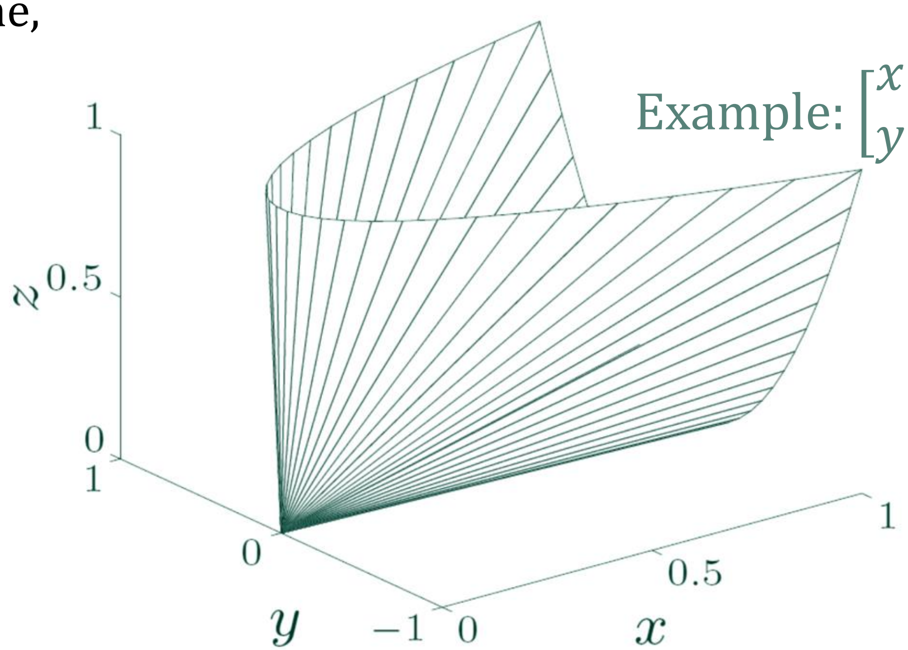
$S_+^n = \{X \in S^n | X \geq 0\}$
positive semidefinite $n \times n$ matrices

PD Cone

$S_{++}^n = \{X \in S^n | X > 0\}$:
positive definite $n \times n$ matrices

$$X \in S_+^n \quad \Leftrightarrow \quad z^T X z \geq 0 \text{ for all } z$$

where S_+^n is convex cone,
S^n is set of symmetric
$n \times n$ matrices

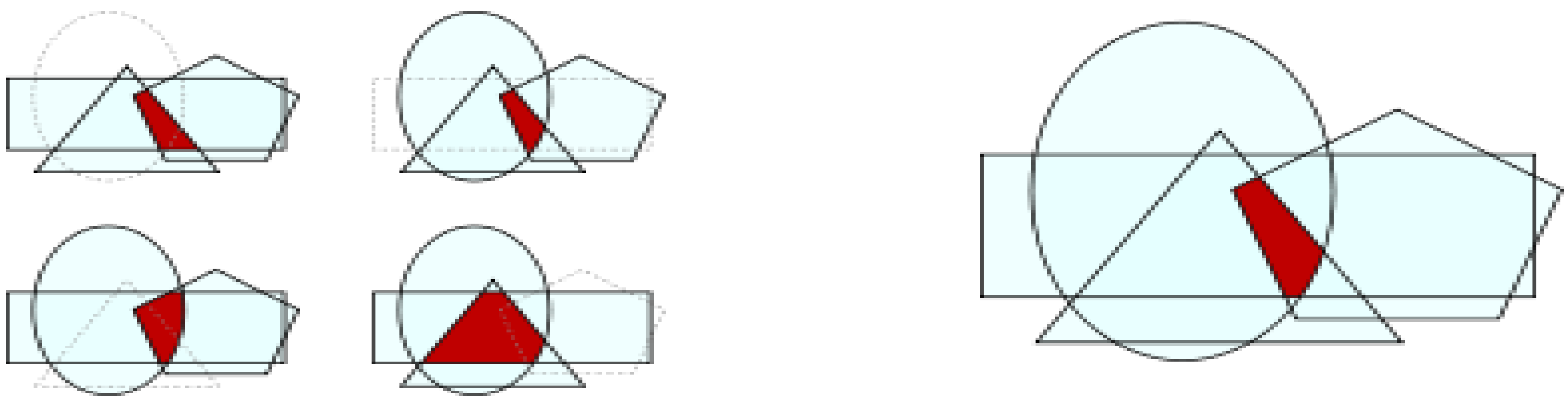


Example: $\begin{bmatrix} x & y \\ y & z \end{bmatrix} \in S_+^2$

Convex Optimization: Convex Set

Intersection

The **intersection** of **(any number of) convex sets** is **convex**



Convex Optimization: Convex Set

Affine Function Suppose $f\colon \mathbb{R}^n \to \mathbb{R}^m$ is affine

$$f(\boldsymbol{x}) = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{b} \qquad \boldsymbol{A} \in \mathbb{R}^{m \times n}, \boldsymbol{b} \in \mathbb{R}^m$$

The **image of a convex set** under f is **convex**

$S \subseteq \mathbb{R}^n$ convex $\longrightarrow$ $f(S) = \{f(x) \mid x \in S\}$

The **inverse image** $f^{-1}(\boldsymbol{C})$ **of a convex set** under f is **convex**

$C \subseteq \mathbb{R}^m$ convex $\longrightarrow$ $f^{-1}(C) = \{x \in \mathbb{R}^n \mid f(x) \in C\}$

Scaling, Translation, Projection

Solution set of **linear matrix inequality**,
$\{x \mid x_1 A_1 + \cdots + x_m A_m \leq B\}$ with $A_i, B \in S^p$

Hyperbolic cone $\{x \mid x^T P x \leq (c^T x)^2, c^T x \geq 0\}$ with $P \in S_+^n$

Convex Optimization: Convex Set

Perspective Function $P : \mathbb{R}^{n+1} \rightarrow \mathbb{R}^{n}$,

$$P(x, t) = {}^{x}/_{t}, \qquad dom\ P = \{(x, t) | t > 0\}$$

The **(inverse) images of convex sets** under **perspective** are **convex**

Linear-fractional Function $f : \mathbb{R}^{n} \rightarrow \mathbb{R}^{m}$

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

The **(inverse) images of convex sets** under **linear-fractional functions** are **convex**