Mathematics for AI

Lecture 30

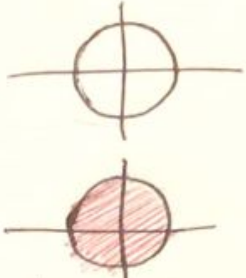
Multiple Equality Constraints
Convex Sets

Remember: Constrained Optimization

min $f(x)$ subject to $h(x)=0$

equality constraints $x^Tx=1$

inequality constraints $x^Tx\leq 1$



$g(x)\geq 1$

$1-g(x)\leq 0$

min $f(x)$ $\quad \boxed{h_1(x)=0},\ h_2(x)=0,\ \dots,\ h_p(x)=0$

$f: \mathbb{R}^n \to \mathbb{R}$ $\quad g_1(x)\leq 0,\ g_2(x)\leq 0,\ \dots,\ g_q(x)\leq 0$

min $f(x)$ s.t. $h(x)=\vec{0}$

$g(x)\leq 0$

$$h(x)=\begin{bmatrix} h_1(x) \\ h_2(x) \\ \vdots \\ h_p(x) \end{bmatrix}$$

$h: \mathbb{R}^n \to \mathbb{R}^p$

Remember: Equality Constraint

K. N. Toosi
University of Technology

$\min f(x)$ sub. to $h(x) = 0$ $\quad h: \mathbb{R} \to \mathbb{R}$

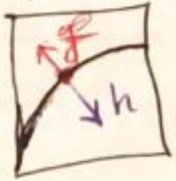
$$L(x, \lambda) = f(x) + \lambda h(x)$$

$$\frac{\partial L}{\partial \lambda} = 0 \Rightarrow \boxed{h(x) = 0}$$

$$\frac{\partial L}{\partial x} = \nabla_x L = 0 \Rightarrow \boxed{\nabla f(x) + \lambda \nabla h(x) = 0}$$

ضریب لاگرانژ

$x \in \mathbb{R}^2$



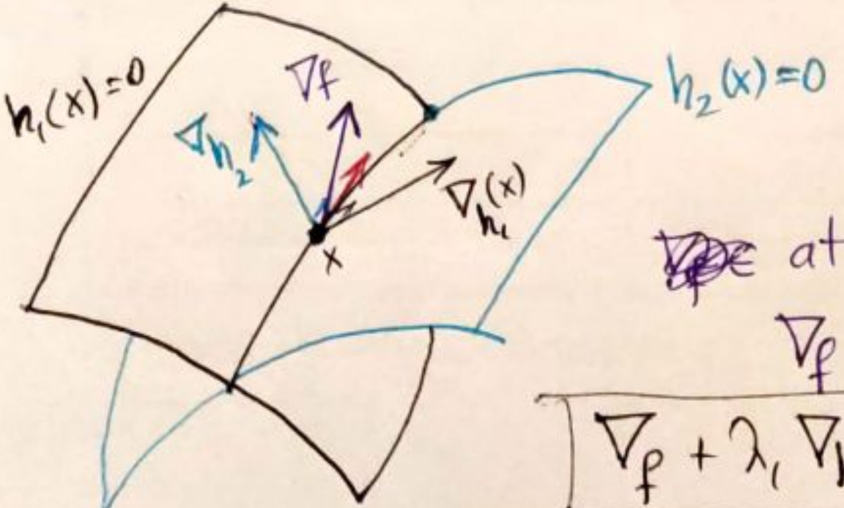
$x \in \mathbb{R}^3$

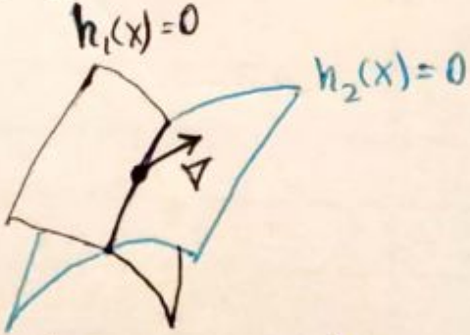
Multiple Equality Constraints

K. N. Toosi
University of Technology

$$\min f(x) \quad \text{s.t.} \quad h_1(x)=0\ ,\ h_2(x)=0$$

Assume that $x \in \mathbb{R}^3$





at stationary point

$$\nabla f \in \text{span}(\nabla h_1, \nabla h_2)$$

$$\boxed{\nabla f + \lambda_1 \nabla h_1 + \lambda_2 \nabla h_2 = 0}$$

Multiple Equality Constraints

$$\min_{x} f(x) \quad \text{s.t.} \quad h_1(x) = 0$$
$$h_2(x) = 0$$

$f: \mathbb{R}^n \to \mathbb{R}$

$$L(x, \lambda_1, \lambda_2) = f(x) + \lambda_1 h_1(x) + \lambda_2 h_2(x)$$

$$\frac{\partial L}{\partial \lambda_1} = 0 \Rightarrow h_1(x) = 0$$

$$\frac{\partial L}{\partial \lambda_2} = 0 \Rightarrow h_2(x) = 0$$

$$\nabla_x L = \frac{\partial L}{\partial x} = 0 \Rightarrow \nabla f(x) + \lambda_1 \nabla h_1(x) + \lambda_2 \nabla h_2(x) = 0$$

$n+2$ equations, $n+2$ unknowns $(x, \lambda_1, \lambda_2)$

Multiple Equality Constraints - Vector Form

K. N. Toosi
University of Technology

$$h(x) = \begin{bmatrix} h_1(x) \\ h_2(x) \end{bmatrix} \qquad h: \mathbb{R}^n \to \mathbb{R}^2 \Rightarrow \lambda_1 h_1(x) + \lambda_2 h_2(x) =$$

$$\underbrace{\begin{bmatrix} \lambda_1 & \lambda_2 \end{bmatrix}}_{\vec{\lambda}^T} \underbrace{\begin{bmatrix} h_1(x) \\ h_2(x) \end{bmatrix}}_{h} = \vec{\lambda}^T h(x)$$

$$\min \; f(x) \quad \text{s.t.} \quad h(x) = \vec{0}$$

$$\cdot\, L(x, \lambda) = f(x) + \lambda^T h(x) \qquad f: \mathbb{R}^n \to \mathbb{R} \quad x \in \mathbb{R}^n$$

$$L: \mathbb{R}^n \times \mathbb{R}^p \to \mathbb{R} \qquad h: \mathbb{R}^p \to \mathbb{R} \quad \lambda \in \mathbb{R}^p$$

λ: the vector of Lagrange multipliers (dual variable) $\quad \lambda = \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_p \end{bmatrix}$

$$\left.\begin{aligned} \nabla_x L &= \frac{\partial L}{\partial x} = \nabla_x f + \nabla_x \lambda^T h = 0 \\ \nabla_\lambda L &= \frac{\partial L}{\partial \lambda} = h(x) = 0 \end{aligned}\right\} \quad \begin{aligned} &n+p \text{ equations} \\ &n+p \text{ unknowns} \end{aligned}$$

Remember: Quadratic Function

K. N. Toosi
University of Technology

Quadratic function $f : \mathbb{R}^n \to \mathbb{R}$ MA

$f(x) = \frac{1}{2} x^T A x + b^T x + c$ $A \in \mathbb{R}^{n \times n}$, symmetric

$b \in \mathbb{R}^n$, $c \in \mathbb{R}$

stationary point $\nabla f = 0 \Rightarrow \nabla f = Ax + b = 0$

$Ax = -b$

$x = A^{-1} b$

$\to$ x minimum if

A is PD

Optimize Quadratic subject to linear constraints

$$\min \; f(x) = \tfrac{1}{2} x^T H x + b^T x + c \quad \text{s.t.} \quad Ax = v$$

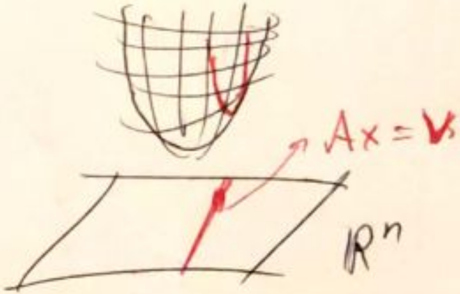
$$H \in \mathbb{R}^{n\times n},\; b \in \mathbb{R}^n,\; c \in \mathbb{R} \qquad A \in \mathbb{R}^{p\times n} \qquad v \in \mathbb{R}^p$$

$$\begin{bmatrix} a_1^T x \\ a_2^T x \\ \vdots \\ a_p^T x \end{bmatrix} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_p \end{bmatrix}$$

Assume H is positive definite

$$h(x) = Ax - v$$

$$L(x,\lambda) = f(x) + \underbrace{\lambda^T}_{1\times p}\underbrace{(\underbrace{A}_{p\times n}\underbrace{x}_{n\times 1} - \underbrace{v}_{p\times 1})}$$

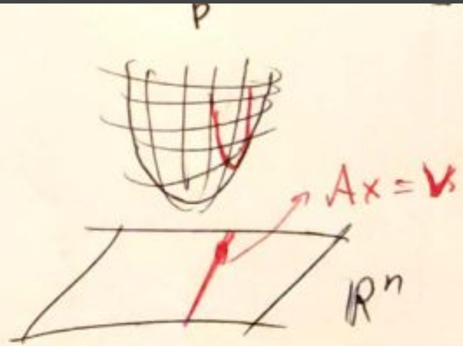


$$\frac{\partial L}{\partial x} = \nabla f + \frac{\partial}{\partial x} \lambda^T (Ax - v)^{p\times 1} = 0$$

Optimize Quadratic subject to linear constraints

K. N. Toosi University of Technology

$$L(x,\lambda) = f(x) + \underbrace{\lambda^T}_{1\times p}\underbrace{(\underbrace{A}_{p\times n}\underbrace{x}_{n\times 1}-\underbrace{v}_{p\times 1})}_{p\times 1}$$



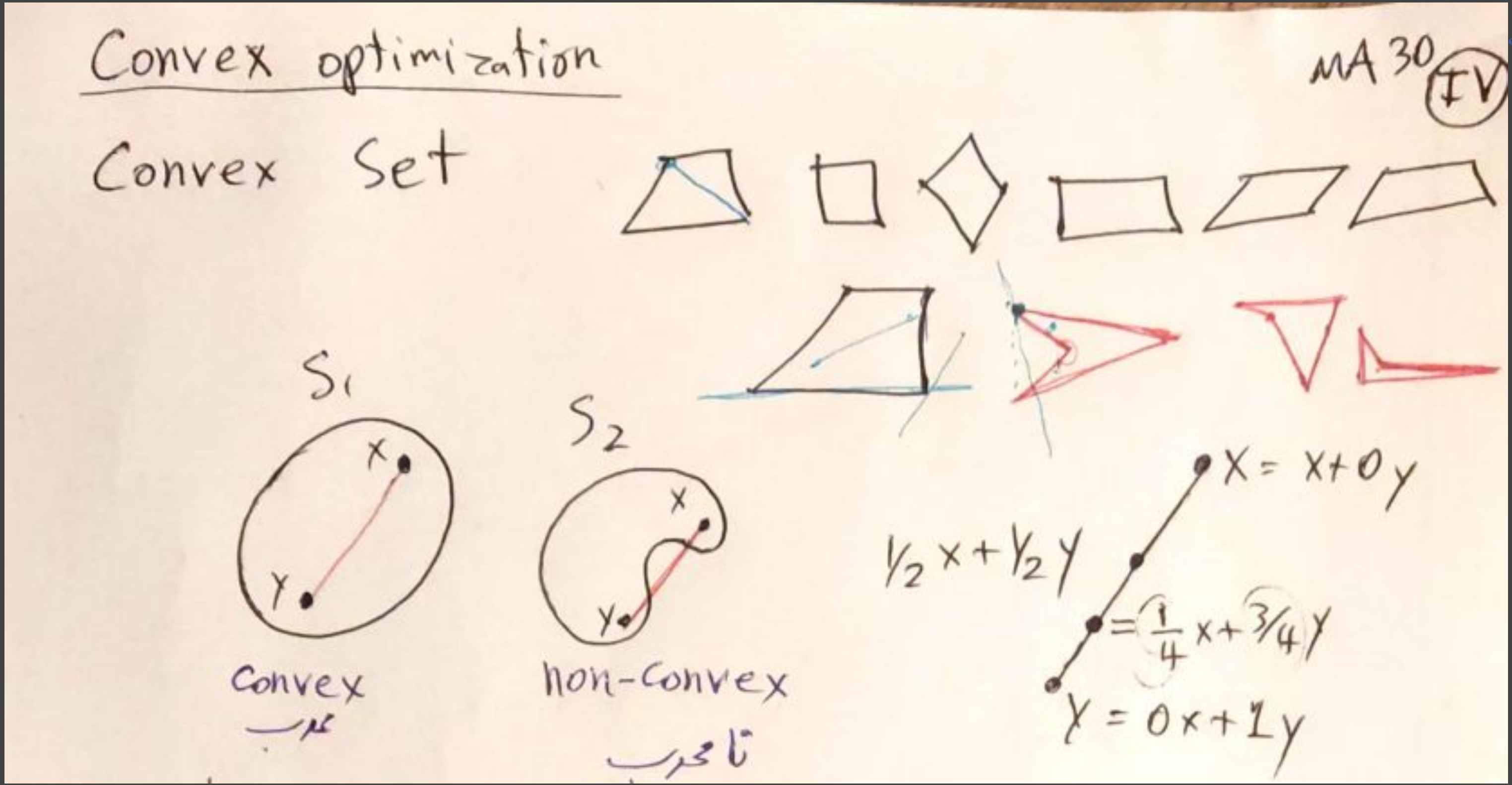
$$\frac{\partial L}{\partial x} = \nabla f + \frac{\partial}{\partial x}\lambda^T(Ax-v) = 0$$

$$\Rightarrow Hx + b + \frac{\partial}{\partial x}\left((\underset{p\times 1}{\underset{\downarrow}{A^T\lambda}})^T x - \lambda^T v\right) = 0$$

$$\left.\begin{array}{l} \Rightarrow Hx + b + A^T\lambda = 0 \\ \text{constran} \quad Ax = v \end{array}\right\} \begin{bmatrix} H & A^T \\ A & 0 \end{bmatrix}\begin{bmatrix} x \\ \lambda \end{bmatrix} = \begin{bmatrix} -b \\ v \end{bmatrix}$$

$$\Rightarrow x,\lambda = \checkmark$$

Convex Sets


Convex optimization
MA 30
IV
Convex Set
S1
S2
x
y
convex
محدب
non-convex
نامحدب
X = x+0y
1/2 x + 1/2 y
= 1/4 x + 3/4 y
Y = 0x+1y

Convex Sets

A set X is convex if for ~~all~~ any

$x, y \in S$ $\alpha x + (1-\alpha) y \in S$ for all $\alpha \in [0,1]$.

$$\underbrace{\alpha x + \beta y \quad \alpha, \beta > 0, \quad \alpha + \beta = 1}_{\text{a convex combination of } x, y}$$

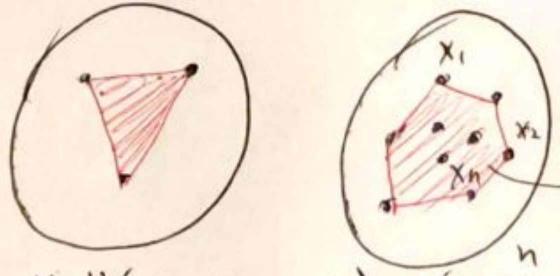
Convex Hull

for a convex set S

$x_1, x_2, \ldots x_n \in S$ any convex combination of $x_1, x_2, \ldots x_n$ is also in S

$\alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_n x_n \in S$ for all $\alpha_1, \alpha_2, \ldots, \alpha_n \geq 0$

$\alpha_1 + \alpha_2 + \cdots + \alpha_n = 1$



convex hull of $x_1, x_2, \ldots x_n$

$$\text{Convex-Hull}(x_1, x_2, \ldots, x_n) = \left\{ \sum_{i=1}^{n} \alpha_i x_i \;\middle|\; \alpha_1, \alpha_2, \ldots, \alpha_n \geq 0, \sum_{i=1}^{n} \alpha_i = 1 \right\}$$

intersection and union of convex sets

S_1, S_2 convex

MA30 #V

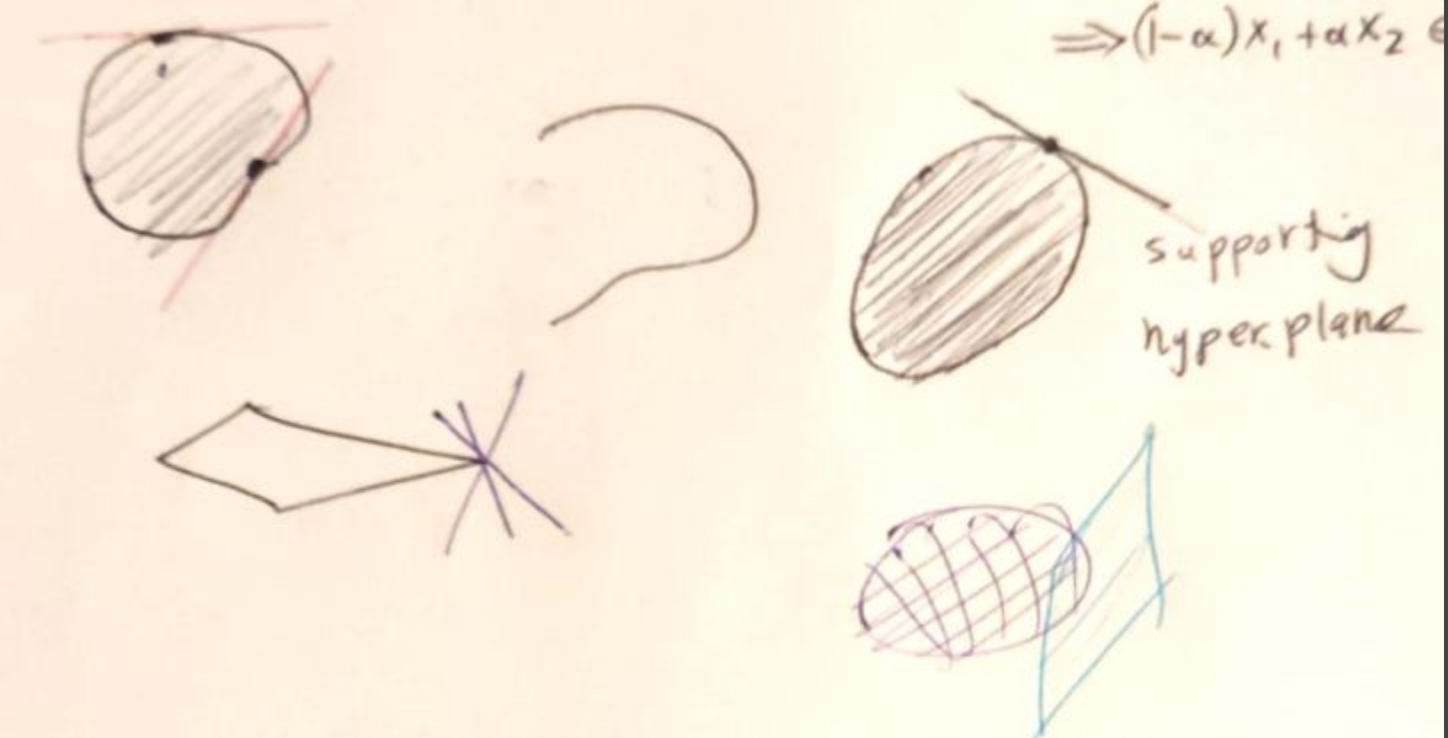
$\boxed{S_1 \cap S_2 \text{ convex}}$

$S_1 \cup S_2$ might not be convex

$$x_1, x_2 \in S_1 \cap S_2 \Rightarrow \left.\begin{array}{l} x_1, x_2 \in S_1 \Rightarrow (1-\alpha)x_1 + \alpha x_2 \in S_1 \\ x_1, S_2 \in S_2 \Rightarrow (1-\alpha)x_1 + \alpha x_2 \in S_2 \end{array}\right\} \Rightarrow$$

$$\Rightarrow (1-\alpha)x_1 + \alpha x_2 \in S$$

Supporting Hyperplane



Convex Functions