Mathematics for AI

Lecture 32

Inequality Constraints, Primal and Dual Problems, Strong and Weak Convexity, Convex Optimization

Convex functions properties

when is $f \circ g(x)$ convex? $\quad g: \mathbb{R}^n \to \mathbb{R}$

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

$\beta = 1-\alpha$

$\cancel{f \circ g(x)} \quad f(g(\alpha x_1 + \beta x_2)) \leq \alpha f(g(x_1)) + \beta f(g(x_2))$

g convex $\quad g(\alpha x_1 + \beta x_2) \leq \alpha g(x_1) + \beta g(x_2)$

f: ~~decreasing~~ non-decreasing

$$f(g(\alpha x_1 + \beta x_2)) \leq f(\alpha g(x_1) + \beta g(x_2))$$

f: convex

$$f(\alpha g(x_1) + \beta(g(x_2)) \leq \alpha f(g(x_1)) + \beta f(g(x_2))$$

Convex functions properties

$f(x), g(x)$ convex

$h(x) = \max(f(x), g(x))$ convex

$\cancel{h(x) = \max}$

$f_i(x)$ convex for $i \in I$

$f_\lambda(x) = e^{\lambda x}$ $\quad \lambda \in \mathbb{R}^+$ convex

Maximum of convex functions

K. N. Toosi University of Technology

$f_i(x)$ convex for $i \in I$

$$\Longrightarrow \max_{i \in I} f_i(x) \quad \text{convex}$$

Minimum of concave functions

K. N. Toosi
University of Technology

$f_i(x)$ concave for $i \in I$

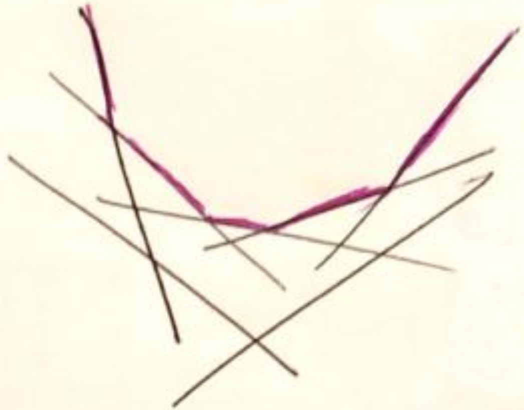
$\Rightarrow \min\limits_{i \in I} f_i(x)$ concave

Maximum and Minimum of Affine functions

$f_i(x)$ affine for $i \in I$

$\max_{i \in I} f_i(x)$ convex

$\min_{i \in I} f_i(x)$ concave



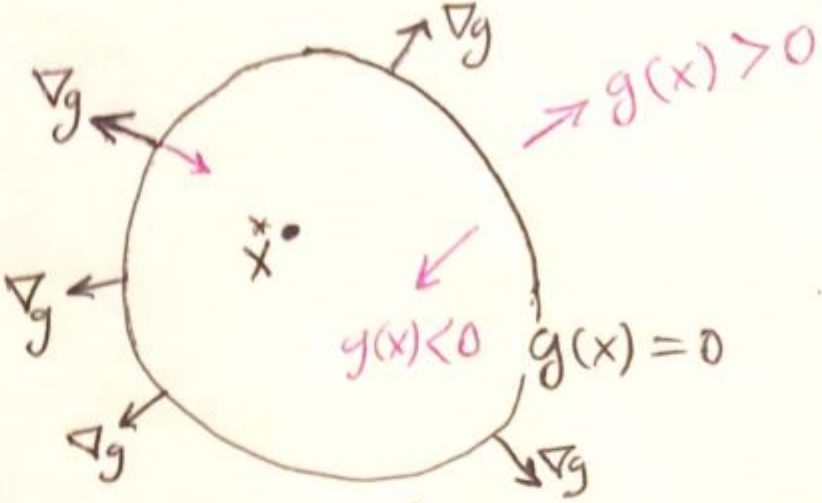
Inequality Constraints

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

$$\min f(x) \quad \text{s.t.} \quad h(x) \leqslant 0 \quad \text{inequality constraints}$$

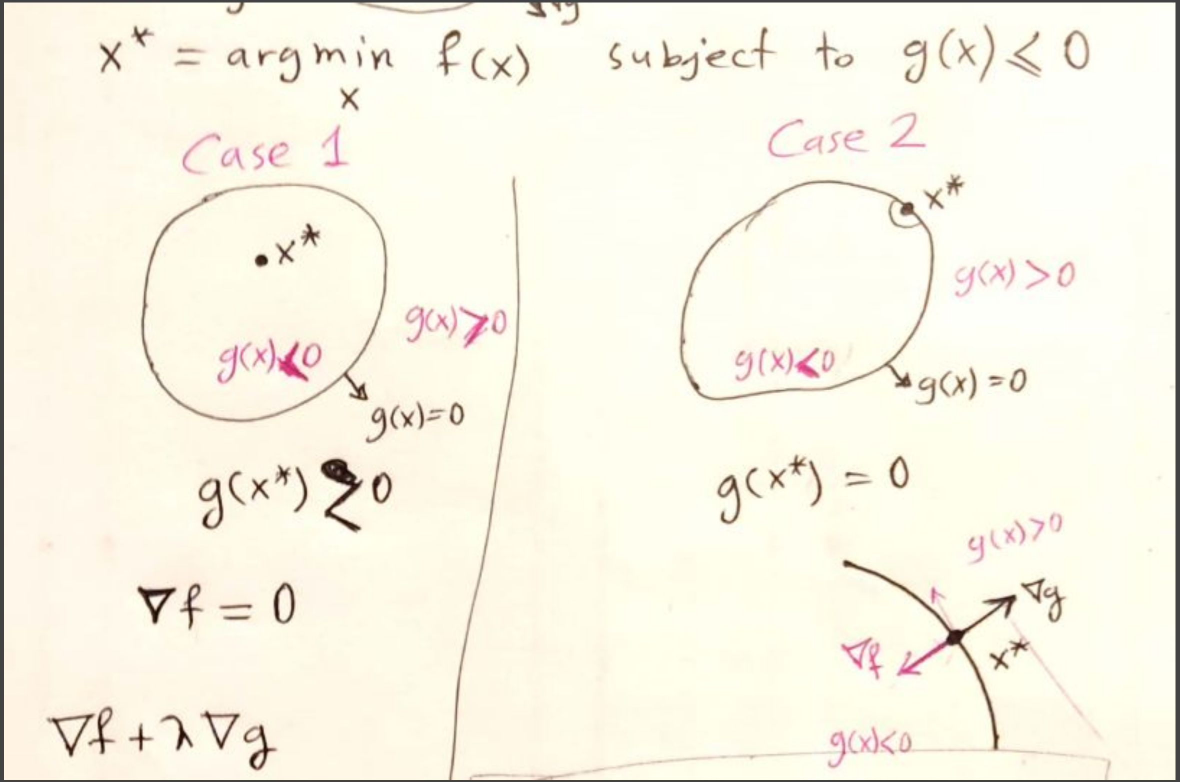
Optimality Conditions

$$\min\ f(x) \quad \text{subject to} \quad g(x) \leqslant 0$$



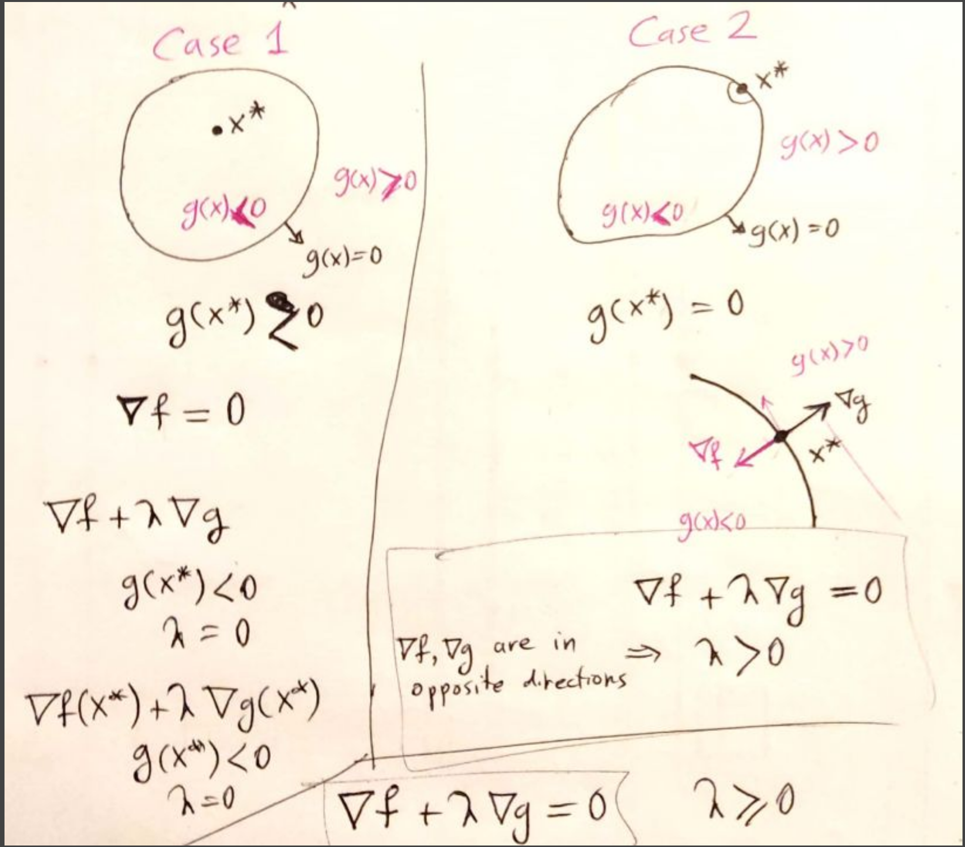
$$x^* = \underset{x}{\operatorname{argmin}}\ f(x) \quad \text{subject to} \quad g(x) \leqslant 0$$

Optimality Conditions



Optimality Conditions

K. N. Toosi
University of Technology



Lagrangian and Lagrange Multipliers

K. N. Toosi
University of Technology

$$\min_{x} f(x) \quad \text{s.t.} \quad g_i(x) \leq 0 \quad i = 1, 2, \ldots, p$$

$$h_i, g_i, f : \mathbb{R}^n \longrightarrow \mathbb{R} \qquad h_i(x) = 0 \quad i = 1, 2, \ldots, q$$

$$L(x, \vec{\lambda}, \vec{\mu}) = f(x) + \sum_{i=1}^{p} \lambda_i g_i(x) + \sum_{i=1}^{q} \mu_i h_i(x)$$

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

$\lambda \in \mathbb{R}^p$

$\vec{\mu} \in \mathbb{R}^q$

Lagrangian in vector form

K. N. Toosi
University of Technology

$$\min_{x} f(x) \quad \text{s.t.} \quad g(x) \leqslant \vec{0} \ , \ h(x) = \vec{0}$$

$$L(x, \lambda, \mu) = f(x) + \lambda^T g(x) + \mu^T h(x)$$

Lagrangian

K. N. Toosi
University of Technology

$$\min_{x} f(x) \quad \text{s.t.} \quad g(x) \leqslant \vec{0}, \quad h(x) = \vec{0}$$

$$L(x, \lambda, \mu) = f(x) + \lambda^T g(x) + \mu^T h(x)$$

Assume that x is feasible $(g(x) \leqslant \vec{0},\ h(x) = \vec{0})$

and $\lambda \geqslant \vec{0}$ $(\lambda_i \geqslant 0$ for all $i = 1 \cdots p)$ then

$$L(x, \lambda, \mu) \leqslant f(x)$$

since $h_i(x) = 0 \Rightarrow \sum \mu_i h_i(x) = 0$

$\lambda_i \geqslant 0 \quad g_i(x) \leqslant 0 \Rightarrow \sum \lambda_i g_i(x) \leqslant 0$

Lagrangian dual function

$$\boxed{\lambda_i \geqslant 0 \quad g_i(x) \leqslant 0 \Rightarrow \sum \lambda_i g_i(x) \leqslant 0}$$

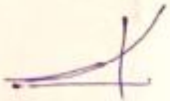
دوگان

$$\underbrace{f^*(\lambda, \mu)}_{\text{dual variables}} = \min_x L(x, \lambda, \mu)$$

(Lagrangian) dual function of f

If $\lambda \geqslant 0$
x feasible
$g(x) \leqslant 0,\ h(x) = 0$

$$\Rightarrow f^*(\lambda, \mu) \leqslant L(x, \lambda, \mu) \leqslant f(x)$$



$\geqslant f^* = \min f(x)$ s.t. $g(x) \leqslant \vec{0}$, $h(x) = \vec{0}$

Dual function as a lower bound

$$x^* = \operatorname{argmin} f(x) \quad \text{s.t.} \quad g(x) \leqslant \vec{0}, \; h(x) = \vec{0}$$

$$\Rightarrow \boxed{f^*(\lambda, \mu) \leqslant f(\underline{x^*}) \qquad \text{if } \lambda \geqslant 0}$$

Dual function as a lower bound

$$y^* = \min_x f(x) \quad \text{s.t.} \quad g(x) \leq 0 \;,\; h(x) = 0$$

primal problem

$$L(x, \lambda, \mu) = f(x) + \lambda^T g(x) + \mu^T h(x)$$

$$f^*(\lambda, \mu) = \min_x L(x, \lambda, \mu)$$

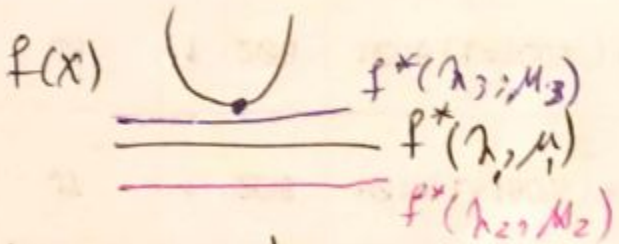
$$f^*(\lambda, \mu) \leq y^* \qquad \text{for all } \mu \in \mathbb{R}^q,\ \lambda \in \mathbb{R}^p,\ \lambda \geq 0$$

dual function

The best (largest) lower bound

K. N. Toosi
University of Technology

$$f^*(\lambda,\mu) \leqslant y^*$$



what is the best lower bound?

$$z^* = \max_{\lambda,\mu} f^*(\lambda,\mu) \quad \text{s. t.} \quad \lambda \geqslant 0$$

dual problem

$$z^* \leqslant y^*$$

Primal and Dual Problems

K. N. Toosi
University of Technology

Primal Problem

$$y^* = \min_{x} f(x)$$

s. t.

$$g(x) \preccurlyeq \vec{0}$$

$$h(x) = \vec{0}$$

Dual Problem

$$z^* = \max_{\lambda, \mu} f^*(\lambda, \mu)$$

s. t.

$$\lambda \succcurlyeq \vec{0}$$

Weak Duality and Duality Gap

Primal Problem	Dual Problem
$y^* = \min_{x} f(x)$	$z^* = \max_{\lambda,\mu} f^*(\lambda,\mu)$
s.t.	s.t.
$g(x) \leq \vec{0}$	$\lambda \geq \vec{0}$
$h(x) = \vec{0}$	

$y^* \geq z^*$ weak duality

$y^* - z^*$ duality gap

Strong Duality (zero duality gap)

$y^* \geq z^*$ weak duality

$y^* - z^*$ duality gap

For some problem we have $y^* = z^*$. In those cases we say that strong duality holds.

Dual function is always concave

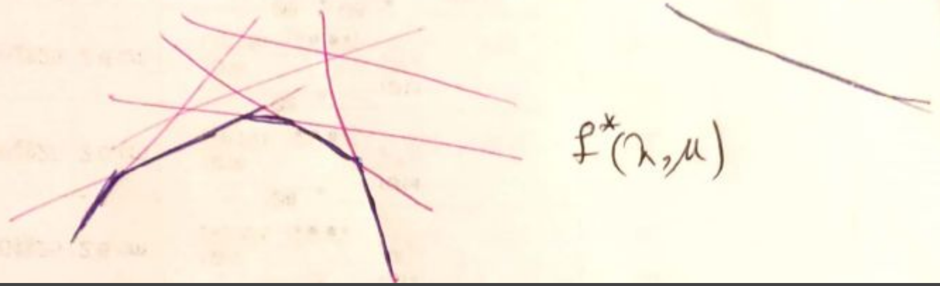
MA32 V

$$f^*(\lambda,\mu) = \min_x L(x,\lambda,\mu)$$

$$= \min_x \underbrace{f(x) + \lambda^T g(x) + \mu^T h(x)}_{\text{Affine function on } \begin{bmatrix}\lambda \\ \mu\end{bmatrix}}$$

$$= \min_x f(x) + \begin{bmatrix} g(x)^T & h(x)^T \end{bmatrix} \begin{bmatrix}\lambda \\ \mu\end{bmatrix}$$

$f^*(\lambda,\mu)$ = min over a set of affine functions



Convex Optimization (Abstract form)

K. N. Toosi
University of Technology

~~Convex~~ Convex Optimization

$\min f(x)$ subject to $x \in C$

C is a convex set.

abstract form

Convex Optimization (Standard form)

K. N. Toosi
University of Technology

Convex Optimization (standard form)

$$\min f(x) \quad \text{subject to} \quad g_i(x) \leq 0 \qquad i = 1, 2, \ldots, p$$

$$Ax = b \qquad A \in \cancel{1, 2, \ldots} \ \mathbb{R}^{q \times n}$$

$$b \in \mathbb{R}^q$$

f convex

g_i convex

for most convex optimization problems strong duality holds. (slater's conditions, etc.)

$$z^* = y^*$$