ĐẠI HỌC BÁCH KHOA

25 YEARS ANNIVERSARY SOICT

HA NOI UNIVERSITY OF SCIENCE AND TECHNOLOGY

SCHOOL OF INFORMATION AND COMMUNICATION TECHNOLOGY

HA NOI UNIVERSITY OF SCIENCE AND TECHNOLOGY
SCHOOL OF INFORMATION AND COMMUNICATION TECHNOLOGY

FUNDAMENTALS OF OPTIMIZATION

Constrained convex optimization

CONTENT

- Lagrange dual function
- Lagrange dual problem
- KKT condition

Lagrange dual function

- Optimization problem in the standard form

$$
\begin{aligned}
(P)\quad & \text{minimize } f(x) \\
\text{s.t.}\quad & g_i(x) \le 0,\ i = 1, 2, \ldots, m \\
& h_i(x) = 0,\ i = 1, 2, \ldots, p \\
& x \in X \subseteq R^n
\end{aligned}
$$

with $x \in R^n$, and assume $D = (\cap_{i=1}^{m} \textbf{dom } g_i) \cap (\cap_{i=1}^{m} \textbf{dom } h_i)$ is not empty.

- Denote f^* the optimal value of $f(x)$
- If f, g_i (i = 1,2,…,m) are convex functions, h_i (i = 1,…,p) are linear → **convex program**

Lagrange dual function

- Define Lagrangian function L: $R^n \times R^m \times R^p \rightarrow R$

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

- Lagrange dual function (or dual function)

$$q(\lambda, \mu) = inf_{x \in D} L(x, \lambda, \mu)$$

- Lagrange dual problem

(D) maximize $q(\lambda, \mu)$

$\lambda, \mu \geq 0$

Lagrange dual function

- **Weak duality theorem** if x^* is an optimal solution to the primal problem and (λ^*,μ^*) is an optimal solution to the dual problem, then $f(x^*) \geq q(\lambda^*,\mu^*)$
- **Corollary** If there exist x^* and (λ^*,μ^*) such that $f(x^*) = q(\lambda^*,\mu^*)$, then x^* and (λ^*,μ^*) are respectively optimal solutions to the primal and dual problems

KKT Conditions

- **Theorem** (Fritz John necessary conditions) Let x* be a feasible solution of (*P*). If *x** is a local minimum of (*P*), then there exists (u,λ,μ) such that:
  - $u\nabla f(x^*) + \sum_{i=1}^{m} \lambda_i \nabla g_i(x^*) + \sum_{i=1}^{p} \mu_i \nabla h_i(x^*) = 0$
  - $u,\lambda \geq 0$, $(u,\lambda,\mu) \neq 0$
  - $\lambda_i g_i(x^*) = 0$, $i = 1,\ldots, m$

KKT Conditions

- **Theorem** (Karush-Kuhn-Tucker (KKT) necessary conditions) Let x* be a feasible solution of (P) and $I = \{i: g_i(x^*) = 0\}$. Further, suppose that $\nabla h_i(x^*)$ for $i = 1,\dots, p$ and $g_i(x^*)$ for $i \in I$ are linearly independent. If x^* is a local minimum of (P), then there exists (λ,μ) such that:
  - $\nabla f(x^*) + \sum_{i=1}^{m} \lambda_i \nabla g_i(x^*) + \sum_{i=1}^{p} \mu_i \nabla h_i(x^*) = 0$
  - $\lambda \geq 0$,
  - $\lambda_i g_i(x^*) = 0,\ i = 1,\dots, m$

KKT Conditions

- **Theorem** (KKT sufficient conditions) Let x^* be a feasible solution of (P) which is convex program (f, g_i are convex functions, h_i are linear functions). If there exists (λ,μ) such that:
  - $\nabla f(x^*) + \sum_{i=1}^{m} \lambda_i \nabla g_i(x^*) + \sum_{i=1}^{p} \mu_i \nabla h_i(x^*) = 0$
  - $\lambda \geq 0$,
  - $\lambda_i g_i(x^*) = 0,\ i = 1, \ldots, m$

  then x^* is a global optimal solution of (P)

Example

minimize $f(x,y) = 2x - y$

s.t. $g_1(x,y) = x^2 + y^2 - 2 \leq 0$

$g_2(x,y) = x - y - 1 \leq 0$

- $\nabla f(x,y) = [2 \quad 1]^T$, $\nabla g_1(x,y) = [2x \quad 2y]^T$, $\nabla g_2(x,y) = [1 \quad -1]^T$
- f and g_2 are linear, so they are convex
- $\nabla^2 g_1(x,y) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$ which is positive definite, so g_1 is convex
- **Exercise** Apply KKT condition, solve the problem ?

Lagrange relaxation for Integer Programming

$$
\begin{aligned}
Z_{IP} = \ \text{minimize } f(x) = \quad & c^{\mathrm{T}}x \\
\text{s.t.} \quad & Ax \geq b \\
& Dx \geq d \\
& x \text{ integer}
\end{aligned}
$$

- Let $X = \{x \text{ integer} \mid Dx \geq d\}$
- Assume optimizing over X can be solved easily, but adding constraint $Ax \geq b$ makes the problem too difficult

Lagrange relaxation for Integer Programming

$$
\begin{aligned}
Z(\lambda) = \min_x \quad & c^T x + \lambda^T (b - Ax) \\
\text{s.t.} \quad & Dx \geq d \\
& x \text{ integer} \qquad\qquad \text{c}
\end{aligned}
$$

- For a fixed λ, $Z(\lambda)$ is assumed to be computed easily
- Important: compute the best lower bound

$$Z_D = \max_{\lambda \geq 0} Z(\lambda)$$

Lagrange relaxation for Integer Programming

- **Exercise** Given a constant value $\lambda = \lambda^{(k)}$, suppose $x^{(k)}$ is an optimal solution to the problem

$$
\begin{aligned}
Z(\lambda) = \min_x \quad & c^T x + \lambda^T (b - Ax) \\
\text{s.t.} \quad & Dx \geq d \\
& x \text{ integer} \qquad \text{c}
\end{aligned}
$$

Explain why $s^{(k)} = b - Ax^{(k)}$ is a subgradient of function Z at $\lambda^{(k)}$?

Lagrange relaxation for Integer Programming

- Subgradient method for computing Z_D

```
Choose starting point λ^(0) (e.g., λ^(0) = 0); k = 0
while (STOP condition not reach) {
    x^(k) is the solution of Z(λ^(0))
    Compute subgradient s^(k) = b - Ax^(k) of function Z at λ^(k)
    if s^(k) = 0 then BREAK
    λ^(k+1) = max{0, λ^(k) + α^(k)s^(k)}  /* α^(k) denote the step size */
    k = k + 1
}
```

VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG
SCHOOL OF INFORMATION AND COMMUNICATION TECHNOLOGY

Thank you for your attentions!