Convex Optimization

Duality: Karush-Kuhn-Tucker (KKT) Optimality Conditions

Zubair Khalid

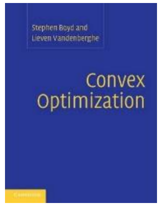
Department of Electrical Engineering
School of Science and Engineering
Lahore University of Management Sciences

https://www.zubairkhalid.org/ee563_2020.html

LUMS

Outline

- KKT Optimality Conditions
- Examples



Section 5.5.2, 5.5.3

Recap

Primal Optimization problem

$$\begin{aligned} \text{minimize} \quad & f_o(x) \\ \text{subject to} \quad & f_i(x) \le 0, \quad i = 1, 2, \ldots, m \\ & h_j(x) = 0, \quad j = 1, 2, \ldots, p \end{aligned}$$

Lagrange Dual Problem:

$$\begin{aligned} \text{maximize} \quad & g(\lambda, \mu) \\ \text{subject to} \quad & \lambda \succeq 0 \end{aligned}$$

Best Lower bound on the optimal value of the primal problem:

$$g(\lambda^*, \mu^*) = d^* \le p^*$$

Duality Gap: $p^* - d^*$

- Weak Duality, $d^* \le p^*$
- Strong Duality, $p^* = d^*$
    - Slater's condition establishes for strong duality

Optimality Conditions

Primal Optimization problem

(Does not need to be convex)

minimize $f_o(x)$

subject to $f_i(x) \leq 0, \quad i = 1, 2, \ldots, m$

$h_j(x) = 0, \quad , j = 1, 2, \ldots, p$

Lagrange Dual Problem:

maximize $g(\lambda, \mu)$

subject to $\lambda \succeq 0$

- Optimal value p^* d^*
- Optimal points x^* (λ^*, μ^*)

Q: What are the optimality conditions that are required to be satisfied by primal and dual optimal points?

A: Karush-Kuhn-Tucker Conditions or Theorem (Saddle-point theorem)

These conditions provide a uinfied framework for optimization problems.

Karush-Kuhn-Tucker Optimality Conditions

Assumptions:

- $f_1, f_2, \ldots, f_m$ and $h_1, h_2, \ldots, h_p$ are differentiable
- Optimal points: x^* and (λ^*, μ^*)
- Strong duality holds (Duality gap is zero)

KKT Conditions:

Condition 1: Primal Feasibility

$$f_i(x^*) \leq 0, \quad i = 1, 2, \ldots, m$$

$$h_j(x^*) = 0, \quad , j = 1, 2, \ldots, p$$

Condition 2: Dual Feasibility

$$\lambda^* \succeq 0 \quad \rightarrow \quad \lambda_i^* \geq 0, \quad i = 1, 2, \ldots, m$$

Karush-Kuhn-Tucker Optimality Conditions

Condition 3: Complementary Slackness

$$\lambda_i^* \, f_i(x^*) = 0, \quad i = 1, 2, \ldots, m$$

How?

For x^*, λ^*, μ^*:

since $p^* = d^*$ (strong duality assumption)

$$f_0(x^*) = g(\lambda^*, \mu^*)$$

$$= \inf_x \left(f_0(x) + \sum \lambda_i^* f_i(x) + \sum_{j=1}^{p} \mu_j^* h_j(x) \right)$$

$$\underset{=}{\leq} \; f_0(x^*) + \underbrace{\sum_{i=1}^{m} \lambda_i^* f_i(x^*) + \sum_{j=1}^{p} \mu_j^* h_j(x^*)}_{0} \; \underset{=}{\leq} \; f_0(x^*)$$

$$\sum_{i=1}^{m} \lambda_i^* f_i(x^*) = 0 \Rightarrow \lambda_i^* f_i(x^*) = 0 \Rightarrow \boxed{\begin{array}{l} \lambda_i^* > 0 \Rightarrow f_i(x^*) = 0 \\ f_i(x^*) < 0 \Rightarrow \lambda_i^* = 0 \end{array}}$$

λ_i^* is zero unless f_i is active at the optimal x^*

Karush-Kuhn-Tucker Optimality Conditions

Condition 4: Stationarity

$$L(x, \lambda^*, \mu^*) = f_o(x) + \sum_{i=1}^{m} \lambda_i^* f_i(x) + \sum_{j=1}^{p} \mu_j^* h_j(x)$$

* We know that x^* minimizes Lagrangian

* Consequently, gradiant of $L(x, \lambda^*, \mu^*)$ should be zero at x^*, that is,

$$\nabla L(x^*, \lambda^*, \mu^*) = \nabla f_o(x^*) + \sum_{i=1}^{m} \lambda_i^* \nabla f_i(x^*) + \sum_{j=1}^{p} \mu_j^* \nabla h_j(x^*) = 0$$

Karush-Kuhn-Tucker Optimality Conditions

Primal Feasibility

$$\begin{cases} f_i(x^*) \leq 0, \quad i = 1, 2, \dots, m \\ h_j(x^*) = 0, \quad , j = 1, 2, \dots, p \end{cases}$$

Dual Feasibility

$$\lambda_i^* \geq 0, \quad i = 1, 2, \dots, m$$

Complementary Slackness

$$\lambda_i^* \, f_i(x^*) = 0, \quad i = 1, 2, \dots, m$$

Stationarity

$$\nabla L(x^*, \lambda^*, \mu^*) = \nabla f_o(x^*) + \sum_{i=1}^{m} \lambda_i^* \nabla f_i(x^*) + \sum_{j=1}^{p} \mu_j^* \nabla h_j(x^*) = 0$$

Assumptions:

- Optimal points: x^* and (λ^*, μ^*)
- Strong duality holds (Duality gap is zero)

Karush-Kuhn-Tucker Optimality Conditions

- No convexity assumption
- Strong duality holds (Duality gap is zero)

☆ Optimal points: x^* and (λ^*, μ^*) satisfy KKT conditions

Non-convex Problems - Necessary Condition

Primal optimal x^* and dual optimal (λ^*, μ^*) points satisfy KKT conditions if the duality gap is zero.

Karush-Kuhn-Tucker Optimality Conditions

Convex Problems – Sufficient Condition

- Assume: Convex optimization problem

★ If $\tilde{x}$ and $(\tilde{\lambda}, \tilde{\mu})$ are any points that satisfy KKT conditions

then $\tilde{x}$ is primal **optimal** and $(\tilde{\lambda}, \tilde{\mu})$ are dual **optimal** with **zero** duality gap

Any points $\tilde{x}$ and dual optimal $(\tilde{\lambda}, \tilde{\mu})$ that satisfy KKT conditions are primal and dual optimal (solutions) with zero duality gap.

Summary:

KKT conditions are
- always sufficient
- necessary when strong duality holds

Example:

minimize $2x^2 + y^2 + 4z^2$

subject to

$$x + 2y - z = 6$$

$$2x - 2y + 3z = 12$$

Primal Feasibility Equality Constraints hold

Dual Feasibility X

No λ

Complementary Slackness X

Stationarity Gradient of Lagrangian is zero

$$L(x, y, z, \mu_1, \mu_2) = 2x^2 + y^2 + z^2 + \mu_1(x + 2y - z - 6) + \mu_2(2x - 2y + 3z - 12)$$

$$\nabla_{(x,y,z)} L = \begin{bmatrix} 4x - \mu_1 - 2\mu_2 \\ 2y - 2\mu_1 - 2\mu_2 \\ 8z + \mu_1 - 3\mu_2 \end{bmatrix} = 0$$

- * 5 variables
- * 5 equations

$$x + 2y - z = 6$$

$$2x - 2y + 3z = 12$$

Example:

minimize $2x_1 + x_2$

$x \in R^2$

subject to $x_1^2 + x_2^2 \leq 9$ $\quad$ $x_1^2 + x_2^2 - 9 \leq 0$

$x_1 \leq x_2$ $\quad$ $x_1 - x_2 \leq 0$

Lagrangian

$$\mathcal{L}(x_1, x_2, \lambda_1, \lambda_2) = 2x_1 + x_2 + \lambda_1(x_1^2 + x_2^2 - 9) + \lambda_2(x_1 - x_2)$$

Primal Feasibility

$$x_1^2 + x_2^2 \leq 9$$

$$x_1 - x_2 \leq 0$$

Dual Feasibility

$$\lambda_1 \geq 0$$

$$\lambda_2 \geq 0$$

Complementary Slackness

$$\lambda_1(x_1^2 + x_2^2 - 9) = 0$$

$$\lambda_2(x_1 - x_2) = 0$$

Stationarity

$$2 + 2\lambda_1 x_1 + \lambda_2 = 0$$

$$1 + 2x_2\lambda_1 - \lambda_2 = 0$$

Example:

Primal Feasibility

$x_1^2 + x_2^2 \le 9$

$x_1 - x_2 \le 0$

Dual Feasibility

$\lambda_1 \ge 0$

$\lambda_2 \ge 0$

Complementary Slackness

$\lambda_1 (x_1^2 + x_2^2 - 9) = 0$

$\lambda_2 (x_1 - x_2) = 0$

Stationarity

$2 + 2\lambda_1 x_1 + \lambda_2 = 0$

$1 + 2x_2\lambda_1 - \lambda_2 = 0$

1 $\lambda_1 = 0, \quad \lambda_2 = 0$

2 $\lambda_1 = 0, \quad x_1 = x_2$

3 $\lambda_2 = 0, \quad x_1^2 + x_2^2 - 9 = 0$

4 $x_1 = x_2, \quad x_1^2 + x_2^2 - 9 = 0$

Case 1 : (stationarity eq.)

$2 = 0, \quad 1 = 0$ (Not possible)

Case 2:

$\lambda_2 = -2, \quad \lambda_2 = -1$ (Not possible)

Case 3 :

$2 + 2\lambda_1 x_1 = 0 \Rightarrow x_1 = -\frac{1}{\lambda_1}$

$1 + 2x_2\lambda_1 = 0 \quad x_2 = -\frac{1}{2\lambda_1}$

$x_1^2 + x_2^2 = 9$

$\Rightarrow \frac{1}{\lambda_1^2} + \frac{1}{4\lambda_1^2} = 9$

$\Rightarrow \lambda_1^2 = \frac{5}{36} \Rightarrow \boxed{\lambda_1 = \frac{\sqrt{5}}{6}}$

$\Rightarrow \boxed{x_1 = -\frac{6}{\sqrt{5}}}$

$\boxed{x_2 = -\frac{3}{\sqrt{5}}}$

* optimal points

Example:

minimize $\quad \frac{1}{2}x^TPx + q^Tx + r \qquad P \in S_+^n$

$x \in R^n$

subject to $\quad Ax = b \qquad A \in R^{p \times n}$

Lagrangian

$$\mathcal{L}(x, \mu) = \frac{1}{2}x^TPx + q^Tx + r + (Ax - b)^T\mu$$

Primal Feasibility

$$Ax = b$$

Stationarity

$$\nabla_x \mathcal{L}(x, \mu) = Px + q + A^T\mu = 0$$

$$\begin{bmatrix} 0 & A \\ A^T & P \end{bmatrix} \begin{bmatrix} \mu \\ x \end{bmatrix} = \begin{bmatrix} b \\ -q \end{bmatrix}$$

Feedback: Questions or Comments?

Email: zubair.khalid@lums.edu.pk

Slides available at: https://www.zubairkhalid.org/ee563_2020.html
(Let me know should you need latex source)