Tutorial 12: KKT condition

MAT3007 Optimization

Kangran ZHAO

School of Data Science,
Chinese University of Hong Kong, Shen Zhen

July 15, 2025

Table of Contents

1. KKT condition
   - Review
   - Exercise

KKT Condition

We consider a general constrained optimization problem

$$\begin{aligned} \min \quad & f(x) \\ s.t. \quad & g_i(x) \leq 0 \quad \forall i \in 1, ..., m \\ & h_j(x) = 0 \quad \forall i \in 1, ..., p \end{aligned}$$

KKT Condition

1. Main Condition

$$\nabla f(x) + \sum_{i=1}^{m} \lambda_i \nabla g_i(x) + \sum_{j=1}^{p} \mu_j \nabla h_j(x) = 0.$$

2. Dual Feasibility

$$\lambda_i \geq 0 \quad \forall i = 1, ..., m.$$

3. Complementarity

$$\lambda_i \cdot g_i(x) = 0 \quad \forall i = 1, ..., m.$$

4. Primal Feasibility

$$g_i(x) \leq 0, h_j(x) = 0 \quad \forall i, \quad \forall j.$$

KKT Condition: Remark

1. Is KKT condition sufficient or necessary for being a minimizer?

KKT Condition: Remark

1. Is KKT condition sufficient or necessary for being a minimizer?
   - Neither. However, if a point satisfy the constraint qualification (CQ), then KKT condition is a necessary condition for it to be a minimizer.

KKT Condition: Remark

1. Is KKT condition sufficient or necessary for being a minimizer?
   - Neither. However, if a point satisfy the constraint qualification (CQ), then KKT condition is a necessary condition for it to be a minimizer.
2. For convex problem, is KKT condition necessary or sufficient for a global minimizer?

KKT Condition: Remark

1. Is KKT condition sufficient or necessary for being a minimizer?
   - Neither. However, if a point satisfy the constraint qualification (CQ), then KKT condition is a necessary condition for it to be a minimizer.
2. For convex problem, is KKT condition necessary or sufficient for a global minimizer?
   - Sufficient, not necessary.

KKT Condition: Remark

Proof (KKT's sufficiency in convex optimization)

Define $d := x - x^*$

$$\begin{aligned} f(x) - f(x^*) &\geq \nabla f(x^*)^T d \\ &= -(\lambda^*)^T \nabla g(x^*)^T d - (\mu^*)^T \nabla h(x^*)^T d \\ &= -(\lambda^*)^T \nabla g(x^*)^T d \geq 0 \end{aligned}$$

The second equality is due to affine-linearity of h:

$$\nabla h(x^*)^T d = h(x) - h(x^*) = 0.$$

The last inequality is by the complementarity condition:

$$\lambda_i^* \nabla g_i(x^*)^T d \leq \lambda_i^* (g_i(x) - g_i(x^*)) = \lambda_i^* g_i(x) \leq 0$$

Exercise 1

Consider the problem of projecting a point $v \in \mathbb{R}^n$ onto an ellipsoid

$$\begin{aligned} \min \quad & \|x - v\|^2 \\ s.t. \quad & \sum_{i=1}^{n} b_i x_i^2 \leq 1, \end{aligned}$$

where $b_i > 0$ for $i = 1, ..., n$. v is outside the ellipsoid, i.e. $\sum_{i=1}^{n} b_i v_i^2 > 1$.

1. Derive the KKT condition for the problem.
2. Describe a way to solve the problem by the KKT condition.

Exercise 1: Solution

1. Define $B := diag\{b_1, b_2, ..., b_n\}$. The Lagragian function is

$$\mathcal{L}(x, \lambda) = \|x - v\|^2 + \lambda(x^T Bx - 1)$$

- main condition:$\nabla_x \mathcal{L}(x, \lambda) = 2(x - v) + 2\lambda Bx = 0.$
- primal feasiblity: $x^T Bx \leq 1.$
- dual feasibility: $\lambda \geq 0.$
- complementarity condition: $\lambda(x^T Bx - 1) = 0.$

Exercise 1: Solution

2. We discuss two situations:
   - $\lambda = 0$. By main condition we have $x = v$. By primal feasibility, $v^T Bv \leq 1$, which contradicts with $\sum_{i=1}^n b_i v_i^2 > 1$.
   - $\lambda > 0$. By main condition we have $x = (I + \lambda B)^{-1} v$. ($I + \lambda B$ is invertible due to its positive diagonal), we can substitute it into $x^T Bx = 1$ to get λ, then use it to calculate x.

Exercise 2: SVM

Consider the Support Vector Machine (SVM) problem with linear model

$$\begin{aligned} \min_{w,b,\epsilon} \quad & \frac{1}{2} w^T w \\ s.t. \quad & y_i(w^T x_i + b) \geq 1, \quad \forall i \end{aligned}$$

Write down the KKT condition for SVM.

Exercise 2: Solution

Define the dual variables $\alpha, \beta \in \mathbb{R}^n$. The Lagrangian function is

$$L(w, b, \epsilon, \alpha, \beta) = \frac{1}{2} w^T w + \sum_{i=1}^{n} \alpha_i (1 - y_i(w^T x_i + b)).$$

Main condition

$$w - \sum_{i=1}^{n} \alpha_i y_i x_i = 0, \quad \sum_{i=1}^{n} \alpha_i y_i = 0, \quad i = 1..., n.$$

Primal feasibility

$$y_i(w^T x_i + b) \geq 1, \quad i = 1, ..., n.$$

Dual feasibility

$$\alpha_i \geq 0, \quad i = 1, ..., n.$$

Complementarity

$$\alpha_i (1 - y_i(w^T x_i + b)) = 0, \quad i = 1, ..., n$$

Exercise 3: SVM with Slack Variables

Consider the Support Vector Machine (SVM) problem with linear model

$$
\begin{aligned}
\min_{w,b,\epsilon} \quad & \frac{1}{2}w^T w + C\sum_{i=1}^{n} \epsilon_i \\
s.t. \quad & y_i(w^T x_i + b) \geq 1 - \epsilon_i, \quad \forall i \\
& \epsilon_i \geq 0, \quad \forall i.
\end{aligned}
$$

Write down the KKT condition for SVM.

Exercise 3: Solution

Define the dual variables $\alpha, \beta \in \mathbb{R}^n$. The Lagrangian function is

$$L(w, b, \epsilon, \alpha, \beta) = \frac{1}{2} w^T w + C \sum_{i=1}^{n} \epsilon_i + \sum_{i=1}^{n} \alpha_i (1 - \epsilon_i - y_i (w^T x_i + b)) + \sum_{i=1}^{n} \beta_i (-\epsilon_i).$$

Main condition

$$w - \sum_{i=1}^{n} \alpha_i y_i x_i = 0, \quad \sum_{i=1}^{n} \alpha_i y_i = 0, \quad C - \alpha_i - \beta_i = 0, \quad i = 1..., n.$$

Primal feasibility

$$y_i (w^T x_i + b) \geq 1 - \epsilon_i, \quad \epsilon_i \geq 0, \quad i = 1, ..., n.$$

Dual feasibility

$$\alpha_i \geq 0, \quad \beta_i \geq 0, \quad i = 1, ..., n.$$

Complementarity

$$\alpha_i (1 - \epsilon_i - y_i (w^T x_i + b)) = 0, \quad i = 1, ..., n$$

$$\beta_i (-\epsilon_i) = 0, \quad i = 1, ..., n.$$

Thanks for coming!