Optimality Conditions for Linearly Constrained Problems

Lecture 10-1 - CMSE 382

Prof. Elizabeth Munch

Michigan State University
::
Dept of Computational Mathematics, Science & Engineering

Fri, Mar 27, 2026

This Lecture

Topics:

- Optimality conditions: motivation
- Separation and alternative theorems
- KKT conditions
- Lagrangian function

Announcements:

- Quiz Weds April 1

Section 1

Optimality Conditions

Optimality conditions

$$\begin{array}{ll} \text{minimize} & f(\mathbf{x}) \\ \text{such that} & \mathbf{x} \in C \end{array}$$

Linearly Constrained

$$\begin{array}{ll} \text{minimize} & f(\mathbf{x}) \\ \text{such that} & \mathbf{a}_i^\top \mathbf{x} \le b_i, \\ & i = 1, 2, \dots, m. \end{array}$$

KKT Conditions

Inequality Constrained

$$\begin{array}{ll} \text{minimize} & f(\mathbf{x}) \\ \text{such that} & g_i(\mathbf{x}) \le 0, \\ & i = 1, 2, \dots, m. \end{array}$$

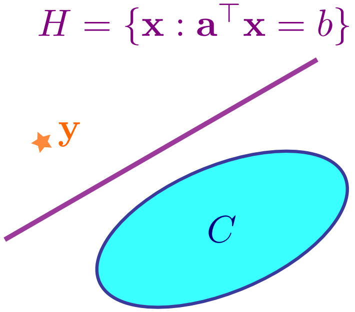
KKT Conditions

Inequality & Equality Constrained

$$\begin{array}{ll} \text{minimize} & f(\mathbf{x}) \\ \text{such that} & g_i(\mathbf{x}) \le 0, h_j(\mathbf{x}) = 0, \\ & i = 1, 2, \dots, m. \\ & j = 1, 2, \dots, p. \end{array}$$

KKT Conditions

Strict separation theorem



Theorem (Strict separation theorem)

Let $C \subseteq \mathbb{R}^n$ be a nonempty closed and convex set, and let $\mathbf{y} \notin C$. Then there exist $\mathbf{p} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$ and $\alpha \in \mathbb{R}$ such that

$$\mathbf{p}^\top \mathbf{y} > \alpha, \quad \text{and} \quad \mathbf{p}^\top \mathbf{x} \leq \alpha \text{ for all } \mathbf{x} \in C.$$

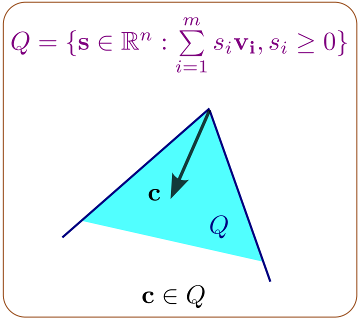
- Meaning: Given C and $\mathbf{y}$ as above, it is possible to draw a hyperplane separating $\mathbf{y}$ and C.

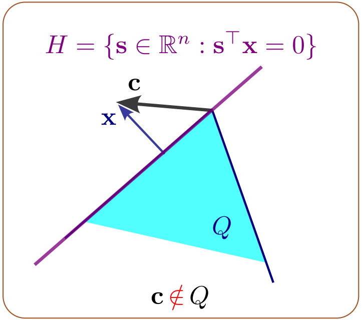
Farkas' lemma

Lemma (Farkas' lemma)

Let $\mathbf{c} \in \mathbb{R}^n$ *and* $A \in \mathbb{R}^{m\times n}$. *Then* **exactly** *one of the following systems has a solution:*

1. *There is an* $\mathbf{x} \in \mathbb{R}^n$ *such that* $A\mathbf{x} \leq \mathbf{0}$, $\mathbf{c}^\top \mathbf{x} > 0$.
2. *There is a* $\mathbf{y} \in \mathbb{R}^m$ *such that* $A^\top \mathbf{y} = \mathbf{c}$, $\mathbf{y} \geq 0$.





Farkas' lemma v2. Gordan's alternative theorem

Lemma (Farkas' lemma—second formulation)

Let $\mathbf{c} \in \mathbb{R}^n$ *and* $A \in \mathbb{R}^{m\times n}$*. Then the following two claims are equivalent.*

- (A) *The implication* $A\mathbf{x} \leq \mathbf{0} \Rightarrow \mathbf{c}^\top \mathbf{x} \leq 0$ *holds true.*
- (B) *There exists* $\mathbf{y} \in \mathbb{R}^M_+$ *such that* $A^T\mathbf{y} = \mathbf{C}$*.*

Theorem (Gordon's alternative theorem)

Let $A \in \mathbb{R}^{m\times n}$*. Then exactly one of the following two systems has a solution:*

- (A) $A\mathbf{x} < 0$*.*
- (B) *There exists* $\mathbf{p} \neq \mathbf{0}$ *that satisfies* $A^\top \mathbf{p} = \mathbf{0}$*,* $\mathbf{p} \geq \mathbf{0}$

KKT for linearly constrained problems

Theorem (Necessary optimality conditions)

Consider the minimization problem

$$\begin{aligned}(P) \quad & \min_{\mathbf{x}} f(\mathbf{x}) \\ & s.t. \quad \mathbf{a}_i^T \mathbf{x} \leq b_i, \quad i = 1, 2, \ldots, m,\end{aligned}$$

where f is continuously differentiable over $\mathbb{R}^n$, $a_1, a_2, \ldots, a_m \in \mathbb{R}^n$, $b_1, b_2, \ldots, b_m \in \mathbb{R}$, and let $\mathbf{x}^*$ be a local minimum point of (P). Then there exist $\lambda_1, \lambda_2, \ldots, \lambda_m \geq 0$ such that

$$\nabla f(\mathbf{x}^*) + \sum_{i=1}^{m} \lambda_i a_i = 0, \quad \textit{and} \quad \lambda_i(\mathbf{a}_i^T \mathbf{x}^* - b_i) = 0, \quad i = 1, 2, \ldots, m.$$

- $\lambda_1, \ldots, \lambda_m$ are Lagrange multipliers. Non-negative for minimization with inequality constraints.

KKT for **convex** linearly constrained problems

Theorem (Necessary and sufficient optimality conditions)

Consider the minimization problem

$$(P) \quad \min_{\mathbf{x}} f(\mathbf{x})$$
$$\text{s.t.} \quad \mathbf{a}_i^T \mathbf{x} \leq b_i, \quad i = 1, 2, \ldots, m,$$

where f is a **convex** *continuously differentiable over $\mathbb{R}^n$, $a_1, a_2, \ldots, a_m \in \mathbb{R}^n$, $b_1, b_2, \ldots, b_m \in \mathbb{R}$, and let $\mathbf{x}^*$ be a feasible solution of (P). Then $\mathbf{x}^*$ is an optimal solution of (P)* **if and only if** *there exist $\lambda_1, \lambda_2, \ldots, \lambda_m \geq 0$ such that*

$$\nabla f(\mathbf{x}^*) + \sum_{i=1}^{m} \lambda_i a_i = 0, \quad \text{and} \quad \lambda_i(\mathbf{a}_i^T \mathbf{x}^* - b_i) = 0, \quad i = 1, 2, \ldots, m.$$

- The condition $\lambda_i(\mathbf{a}_i^T \mathbf{x}^* - b_i) = 0, \quad i = 1, 2, \ldots, m$ is called the complementary slackness condition.

The Lagrangian function

Definition (The Lagrangian function)

Consider the Nonlinear Programming Problem (NLP)

$$\text{(NLP)} \quad \min_{\mathbf{x}} f(\mathbf{x}) \quad \text{s.t.} \quad \{g_i(\mathbf{x}) \leq 0\}_{i=1}^{m}, \quad \{h_j(\mathbf{x}) = 0\}_{j=1}^{p},$$

where f, and all the g_i and h_j are continuously differentiable over $\mathbb{R}^n$.

The associated **Lagrangian function** is of the form

$$L(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) = f(\mathbf{x}) + \sum_{i=1}^{m} \lambda_i g_i(\mathbf{x}) + \sum_{j=1}^{p} \mu_j h_j(\mathbf{x}).$$

The **necessary KKT condition (stationarity condition)** is

$$\nabla_{\mathbf{x}} L(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) = \nabla f(\mathbf{x}) + \sum_{i=1}^{m} \lambda_i \nabla g_i(\mathbf{x}) + \sum_{j=1}^{p} \mu_j \nabla h_j(\mathbf{x}) = \mathbf{0}.$$

The Lagrangian function for linearly constrained optimization

Recall the minimization problem **with linear constraints**

$$(\mathrm{Q}) \quad \min_{\mathbf{x}} f(\mathbf{x}) \quad \text{s.t.} \quad A\mathbf{x} \leq \mathbf{b}, \quad C\mathbf{x} = \mathbf{d}.$$

The associated **Lagrangian function** is of the form

$$L(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) = f(\mathbf{x}) + \boldsymbol{\lambda}^\top (A\mathbf{x} - \mathbf{b}) + \boldsymbol{\mu}^\top (C\mathbf{x} - \mathbf{d}).$$

The **necessary KKT condition** $\nabla f(\mathbf{x}^*) + \sum_{i=1}^m \lambda_i \mathbf{a}_i + \sum_{j=1}^p \mu_j \mathbf{c}_j = \mathbf{0}$ can be written in terms of the Lagrangian as

$$\nabla_{\mathbf{x}} L(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) = \nabla f(\mathbf{x}) + A^\top \boldsymbol{\lambda} + C^T \boldsymbol{\mu} = \mathbf{0}.$$

Steps for finding the stationary points for a linearly constrained problem

- Write the problem in the standard form

$$\min_{\mathbf{x}} f(\mathbf{x}) \quad \text{s.t.} \quad A\mathbf{x} \leq \mathbf{b}, \quad C\mathbf{x} = \mathbf{d}.$$

- Write down the Lagrangian function

$$L(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) = f(\mathbf{x}) + \boldsymbol{\lambda}^\top (A\mathbf{x} - \mathbf{b}) + \boldsymbol{\mu}^\top (C\mathbf{x} - \mathbf{d}).$$

- Write down the KKT conditions

$$\nabla_{\mathbf{x}} L(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) = \nabla f(\mathbf{x}) + A^\top \boldsymbol{\lambda} + C^T \boldsymbol{\mu} = \mathbf{0}, \text{ and } \lambda_i(\mathbf{a}_i^\top \mathbf{x}^* - b_i) = 0.$$

- Write down the feasibility constraints

$$(A\mathbf{x} - \mathbf{b}) \quad \leq 0 \quad \text{and} \quad (C\mathbf{x} - \mathbf{d}) = 0$$

  - If inequality constraints are present, include $\boldsymbol{\lambda} \geq \mathbf{0}$ as a constraint.
- Solve the stationarity and feasibility constraints for the stationary points of the problem.
- If the problem is convex, then stationarity implies optimality.

Groups - Round 5

Group 1
Michal, Kyle, Daniel, Purvi

Group 2
Joseph, Jack, Scott, Breena

Group 3
Saitej, Dori, Noah, Tianjian

Group 4
Dev, Shanze, Lowell, Andrew

Group 5
Lora, Aidan, Arjun, Monirul Amin

Group 6
Anthony, Abigail, Atticus, Yousif

Group 7
Luis, Vinod, Morgan, Dominic

Group 8
Jay, Jonid, Alice, Aaron

Group 9
Arya, Jake, K M Tausif, Lauryn

Group 10
Maye, Ha, Zheng, Sai

Group 11
Jamie, Karen, Brandon, Quang Minh

Group 12
Long, Sanskaar, Braedon, Igor