IEOR 160: Nonlinear and Discrete Optimization
Optimality Conditions (Multi-variate)

Yatong Bai

February 6, 2024

This is a summarization of Professor Javad Lavaei's IEOR 160 lecture notes.

1 Regular Points and Tangent Planes

1.1 Feasible Sets Formed by Equality Constraints Only

Consider the set $\mathcal{X} := \big\{x \in \mathbb{R}^n : \ h_i(x) = 0, \ i = 1, \dots, m\big\}$.

- **Regular point:** A point $y \in \mathcal{X}$ is regular if $\nabla h_i(y), i = 1, \dots, m$ are linearly independent.
  - $\nabla h_i(y), i = 1, \dots, m$ are linearly independent if $\exists \alpha_1, \dots \alpha_m$ such that at least at least one α_i is non-zero and $\sum_{i=1}^m \alpha_i \nabla h_i(y) = 0$.
  - A zero vector is linearly dependent on any vector.
- **Tangent plane:** If a point $y \in \mathcal{X}$ is regular, then the tangent plane of $\mathcal{X}$ at y is defined as $\mathcal{T}(y) := \{\Delta y \in \mathbb{R}^n : \ \nabla h_i(y)^\top \Delta y = 0, \ i = 1, \dots, m\}$.
  - The dimension of the tangent plane is $n - m$.

1.2 Feasible Sets Formed by Equality and Inequality Constraints

Now, consider the set $\bar{\mathcal{X}} := \left\{x \in \mathbb{R}^n : \begin{array}{l} h_i(x) = 0, \ i = 1, \dots, m \\ g_j(x) \le 0, \ j = 1, \dots, l \end{array}\right\}$.

Suppose that at a point $y \in \bar{\mathcal{X}}$, some of the inequality constraints are active. I.e., there exists a set of positive integers $\mathcal{J}_{\text{active}} \subseteq \{1, 2, \dots, l\}$, such that $g_j(y) = 0$ for all $j \in \mathcal{J}_{\text{active}}$.

- **Regular point:** A point $y \in \bar{\mathcal{X}}$ is regular if all $\nabla h_i(y)$ for $i = 1, \dots, m$, as well as all $\nabla g_j(y)$ for $j \in \mathcal{J}_{\text{active}}$, are linearly independent.
- **Tangent plane:** If a point $y \in \bar{\mathcal{X}}$ is regular, then the tangent plane of $\bar{\mathcal{X}}$ at y is defined as $\mathcal{T}(y) := \left\{\Delta y \in \mathbb{R}^n : \begin{array}{l} \nabla h_i(y)^\top \Delta y = 0, \ i = 1, \dots, m \\ \nabla g_j(y)^\top \Delta y = 0, \ j \in \mathcal{J}_{\text{active}} \end{array}\right\}$.

2 Optimality Conditions

2.1 Unconstrained Optimization Problems

Consider the optimization problem $\min_x f(x)$.

- **First-order condition (FOC):** If $x_\star$ is a local minimum, then $\nabla f(x_\star) = 0$.
- **Second-order condition (SOC) necessary:** If $x_\star$ is a local minimum, then $\nabla^2 f(x_\star) \succeq 0$.
- **Second-order condition (SOC) sufficient:** If $\nabla f(x_\star) = 0$ and $\nabla^2 f(x_\star) \succ 0$, then $x_\star$ is a strict local minimum.
- If $\nabla f(x_\star) = 0$ and $\nabla^2 f(x) \succeq 0$ for all $x \in \mathbb{R}^n$, then $x_\star$ is a global min.

2.2 Equality Constrained Optimization Problems

Consider the optimization problem $\min_x f(x)$ subject to $h_i(x) = 0$ for $i = 1, \ldots, m$. Denote the dual variables associated with the constraints as $\lambda_1, \ldots, \lambda_m$. The Lagrangian function is then

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

- **FOC:** If $x_\star$ is a regular point and a local minimum, then there exist Lagrangian multipliers $\lambda_{1\star}, \ldots, \lambda_{m\star}$ such that $\nabla f(x_\star) + \sum_{i=1}^{m} \lambda_{i\star} \nabla h_i(x_\star) = 0$.
- **SOC necessary:** If $x_\star$ is a regular point and a local minimum, then $\frac{1}{2}\Delta x^\top M \Delta x \geq 0$ for all Δx on the tangent plane $\mathcal{T}(x_\star)$, where M is the Hessian of the Lagrangian, defined as

$$M \coloneqq \nabla^2 f(x_\star) + \sum_{i=1}^{m} \lambda_{i\star} \nabla^2 h_i(x_\star).$$

  - To make this condition practical and usable, we can construct a tangent plane basis matrix $E \coloneqq \begin{bmatrix} y_1 & \ldots & y_{n-m} \end{bmatrix}$, where $y_1, \ldots, y_{n-m}$ are arbitrary vectors such that
    - $y_1, \ldots, y_{n-m} \in \mathcal{T}(x_\star)$;
    - $y_1, \ldots, y_{n-m}$ are linearly independent.

    The SOC necessary condition can be written as $E^\top M E \succeq 0$.
- **SOC sufficient:** If $x_\star$ is a regular point satisfying the FOC, then it is a local minimum if $\frac{1}{2}\Delta x^\top M_{x_\star} \Delta x > 0$ for all $\Delta x \in \mathcal{T}(x_\star)$ satisfying $\Delta x \neq 0$.
  - The SOC sufficient condition can also be written as $E^\top M E \succ 0$.

2.3 Equality and Inequality Constrained Optimization Problems

Consider the optimization problem $\begin{array}{ll} \min_x & f(x) \\ \text{s.t.} & h_i(x) = 0,\ i = 1, \ldots, m, \\ & g_j(x) \leq 0,\ j = 1, \ldots, l. \end{array}$

Denote the dual variables associated with the equality constraints as $\lambda_1, \ldots, \lambda_m$. Similarly, denote the dual variables associated with the inequality constraints as $\mu_1, \ldots, \mu_l$. The Lagrangian of this problem is then

$$L(x, \lambda, \mu) \coloneqq f(x) + \sum_{i=1}^{m} \lambda_i h_i(x) + \sum_{j=1}^{l} \mu_j g_j(x).$$

- **FOC:** If $x_\star$ is a regular point and a local minimum, then there exist Lagrangian multipliers $\lambda_{1\star}, \ldots, \lambda_{m\star}$ and $\mu_{1\star}, \ldots, \mu_{l\star}$ that makes the following holds true:
  1. **Primal Feasibility**: $h_i(x_\star) = 0$ for all $i = 1, \ldots, m$ and $g_j(x_\star) \leq 0$ for all $j = 1, \ldots, l$;
  2. **Dual Feasibility**: $\mu_{j\star} \geq 0$ for all $j = 1, \ldots l$;
  3. **Lagrangian Stationarity**: $\nabla f(x_\star) + \sum_{i=1}^{m} \lambda_{i\star} \nabla h_i(x_\star) + \sum_{j=1}^{l} \mu_{j\star} \nabla g_j(x_\star) = 0$;
  4. **Complementary Slackness**: $\mu_{j\star} \cdot g_j(x_\star) = 0$ for all $j = 1, \ldots l$.

  These conditions are called the Karush–Kuhn–Tucker (**KKT**) conditions.

- **SOC necessary:** If $x_\star$ is a regular point and a local minimum, then $E^\top M E \succeq 0$, where

  $$M \coloneqq \nabla^2 f(x_\star) + \sum_{i=1}^{m} \lambda_{i\star} \nabla^2 h_i(x_\star) + \sum_{j=1}^{l} \mu_{j\star} \nabla^2 g_j(x_\star)$$

  is the Hessian of the Lagrangian w.r.t. x, and E is the tangent plane basis matrix defined the same as in Section 2.2.

- **SOC sufficient:** If $x_\star$ is a regular point satisfying the FOC and is non-degenerate, then it is a local minimum if $E^\top M E \succ 0$.
  - Non-degeneracy: $x_\star$ is degenerate if there exists some j such that $\mu_{j\star} = g_j(x_\star) = 0$.

- For convex optimization problems, the KKT conditions can be used to find the global optima.

3 Sensitivity

If we compare the optimization problem

$$\min_x f(x) \;\text{ s. t. }\; h_i(x) = 0,\; g_j(x) \leq 0 \tag{1}$$

with the problem that has perturbed constraints

$$\min_x f(x) \;\text{ s. t. }\; h_i(x) = \epsilon_i,\; g_j(x) \leq \bar{\epsilon}_j \tag{2}$$

where each ϵ_i and $\bar{\epsilon_j}$ is a small number, then the optimal objective value of (2) f_ϵ can be approximated by

$$f_\epsilon \approx f_\star - \sum_i \lambda_i \epsilon_i - \sum_j \mu_j \bar{\epsilon}_j,$$

where $f_\star$ is the optimal objective value of (1), and λ_i and μ_j are the Lagrangian multipliers of (1).