MAT 3007 Optimization: Tutorial 10

Guxin DU

July 8, 2025

Recap: Optimality Conditions for Unconstrained Problems

Theorem 1 (First-Order Necessary Condition).

If f is continuously differentiable and $\mathbf{x}^*$ is a local minimizer of $f(\cdot)$ for an unconstrained problem, then we must have $\nabla f(\mathbf{x}^*) = 0$.

Theorem 2 (Second-Order Necessary Condition).

If f is second-order continuously differentiable $\mathbf{x}^*$ is a local minimizer of $f(\cdot)$ for an unconstrained problem, then we must have

1. $\nabla f(\mathbf{x}^*) = 0$;
2. $\nabla^2 f(\mathbf{x}^*)$ *is positive semi-definite.*

Theorem 3 (Second-Order Sufficient Condition).

If f is second-order continuously differentiable. If $\mathbf{x}^*$ satisfies:

1. $\nabla f(\mathbf{x}^*) = 0$;
2. $\nabla^2 f(\mathbf{x}^*)$ *is positive definite.*

Then $\mathbf{x}^*$ is a local minimizer of f.

Recap: Optimality Conditions for Unconstrained Problems

Tips:

1. Check if a matrix is positive definite/semi-definite;
   $det(A) = \prod_i \lambda_i; tr(A) = \sum_i \lambda_i$
2. When to check sufficiency/necessity
   - Find candidates for optimal solutions or prove a point x is not a local optimum
   - Prove a point x is a local optimum
3. Understand the insights behind those conditions.

Exercise 1

Find all local minimizer, local maximizer and saddle points of f.

$$f(x) = x_1^4 + 2(x_1 - x_2)x_1^2 + 4x_2^2$$

Recap: Optimality Conditions for Linear Constrained Problems

The linear constrained optimization problem is

$$\begin{aligned} \min_{\mathbf{x}} \quad & f(\mathbf{x}) \\ \text{s.t.} \quad & A\mathbf{x} \geq (=)\ b. \end{aligned}$$

Theorem 4 (Equality Constraint).

If $\mathbf{x}^*$ is the minimizer, then there must exist $\mathbf{y}$ such that $A^\top \mathbf{y} = \nabla f(\mathbf{x}^*)$.

Theorem 5 (Inequality Constraint).

If $\mathbf{x}^*$ is the minimizer, then there must exist $\mathbf{y} \geq 0$ such that

$$\begin{aligned} & A^\top \mathbf{y} = \nabla f(\mathbf{x}^*), \\ & y_i(a_i^\top \mathbf{x}^* - b_i) = 0, \forall i. \end{aligned}$$

where $a_i^\top$ is the ith row of A.

Tips:

1. Check the sign in the constraint ($=$ or $\geq$);
2. Carefully check the matrix transpose;
3. Understand the derivation of the two theorems:
   - Find the feasible direction set S_F and descent direction set S_D;
   - Write the condition $S_F \bigcap S_D = \emptyset$;
   - Use dual feasibility to construct the alternative systems;

Exercise 2

Find the distance from the origin $(0,0)^\top$ to the polyhedron $S = \{(x_1, x_2)^\top | x_1 + x_2 \geq 4, 2x_1 + x_2 \geq 5\}$.
Note that the problem is equivalent to solve the following problem:

$$\begin{array}{ll} \min & x_1^2 + x_2^2 \\ \text{s.t.} & x_1 + x_2 \geq 4 \\ & 2x_1 + x_2 \geq 5 \end{array}$$

Exercise 3

Consider the following problem:

$$
\begin{aligned}
\min_{x_1, x_2} \quad & (x_1 - 3)^2 + (x_2 - 2)^2 \\
\text{s.t.} \quad & 2x_1 + x_2 - 6 \leq 0 \\
& x_1 + 2x_2 - 6 \leq 0.
\end{aligned} \tag{1}
$$

- Prove that $x_1 = 11/5$ and $x_2 = 8/5$ is a local minimizer.

Triangle Inequality for p Norm

p norm: For $p \geq 1$, the p-norm of a vector $\mathbf{x} = (x_1, x_2, \ldots, x_n)^\top$ is defined as

$$\|\mathbf{x}\|_p = \left(\sum_{i=1}^{n} |x_i|^p \right)^{1/p}.$$

Particularly, for $p = \infty$, the infinity norm is defined as

$$\|\mathbf{x}\|_\infty = \max_{1 \leq i \leq n} |x_i|.$$

Triangle Inequality: For any vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, and any $p \geq 1$ (∞ included), the triangle inequality states that

$$\|\mathbf{x} + \mathbf{y}\|_p \leq \|\mathbf{x}\|_p + \|\mathbf{y}\|_p.$$

The inequality above is also known as Minkowski's inequality.

Proof of Triangle Inequality for p Norm

For $p = 1$, the triangle inequality follows directly from $|x_i + y_i| \leq |x_i| + |y_i|$ for each $1 \leq i \leq n$.
For $p = \infty$, we have

$$\max_{1 \leq i \leq n} |x_i + y_i| \leq \max_{1 \leq i \leq n} (|x_i| + |y_i|) \leq \max_{1 \leq i \leq n} |x_i| + \max_{1 \leq i \leq n} |y_i|,$$

For $1 < p < \infty$, we first assume that $x \neq 0$ and $y \neq 0$. Otherwise the inequality would be trivial. Then we consider the function $\phi(t) = t^p$ defined on $t \in (0, \infty)$. $\phi(t)$ is convex as $\phi''(t) = p(p-1)t^{p-2} > 0$ for $t > 0$. By the convexity of ϕ, we have for any $s, t > 0$ and $\lambda \in [0, 1]$,

$$(\lambda s + (1 - \lambda)t)^p \leq \lambda s^p + (1 - \lambda)t^p.$$

Proof of Triangle Inequality for p Norm (continued)

For any $1 \le i \le n$, we let $s = \frac{|x_i|}{\|\mathbf{x}\|_p}$, $t = \frac{|y_i|}{\|\mathbf{y}\|_p}$, $\lambda = \frac{\|\mathbf{x}\|_p}{\|\mathbf{x}\|_p + \|\mathbf{y}\|_p}$ and $1 - \lambda = \frac{\|\mathbf{y}\|_p}{\|\mathbf{x}\|_p + \|\mathbf{y}\|_p}$, by the above inequality, we have

$$\left(\frac{|x_i| + |y_i|}{\|\mathbf{x}\|_p + \|\mathbf{y}\|_p}\right)^p \le \frac{\|\mathbf{x}\|_p}{\|\mathbf{x}\|_p + \|\mathbf{y}\|_p}\left(\frac{|x_i|}{\|\mathbf{x}\|_p}\right)^p + \frac{\|\mathbf{y}\|_p}{\|\mathbf{x}\|_p + \|\mathbf{y}\|_p}\left(\frac{|y_i|}{\|\mathbf{y}\|_p}\right)^p.$$

Sum the above inequality over all i from 1 to n, we obtain

$$\frac{\sum_{i=1}^n (|x_i| + |y_i|)^p}{(\|\mathbf{x}\|_p + \|\mathbf{y}\|_p)^p} \le \frac{\|\mathbf{x}\|_p}{\|\mathbf{x}\|_p + \|\mathbf{y}\|_p}\frac{\|\mathbf{x}\|_p^p}{\|\mathbf{x}\|_p^p} + \frac{\|\mathbf{y}\|_p}{\|\mathbf{x}\|_p + \|\mathbf{y}\|_p}\frac{\|\mathbf{y}\|_p^p}{\|\mathbf{y}\|_p^p} = 1.$$

Therefore, we have

$$\|\mathbf{x} + \mathbf{y}\|_p \le \|\mathbf{x}\|_p + \|\mathbf{y}\|_p.$$

Coercive

Definition 6.

A continuous function $f : \mathbb{R}^n \to \mathbb{R}$ is said to be **coercive** if

$$\lim_{\|x\|\to\infty} f(x) = +\infty.$$

i.e. , $\forall B > 0$, $\exists r > 0$ such that if $\|x\| \geq r$ then $f(x) > B$.

Theorem 7.

Let f be a continuous and coercive function. Then for all $\alpha > 0$, the level set

$$L_{\leq\alpha} := \{x : f(x) \leq \alpha\}$$

is compact and f has at least one global minimizer.