INTRODUCTION TO OPTIMIZATION THEORY

Recall the definition of an extremum:

<u>DEF</u>: Let $f: \mathbb{R}^n \supseteq D \to \mathbb{R}$ and $x^* \in D$ s.t. $\exists\ \varepsilon > 0\ \forall\ |x - x^*| < \varepsilon$: either

i) $f(x) \geq f(x^*)$ (local min.)

ii) $f(x) \leq f(x^*)$ (local max.)

then x^* is called an extremum.

<u>History</u>

① 1629 Fermat's theorem

$$\underset{x \in D}{\text{extremum}}\ f(x)$$

<u>THM</u>: Let $f: \mathbb{R}^n \supseteq D \to \mathbb{R}$ diff. at $x^* \in \mathbb{R}^n$.
x^* local extremum $\Rightarrow \nabla f(x^*) = 0$.

<u>ALG</u>: Discussion of boundary of D, non-diff. points of f, then stationary points, i.e., $f'(x^*) = 0$. So one has to work backwards as $\nabla f(x^*) = 0$ is only necessary but not suff.

② 1788 Lagrange's theorem

$$\min_{x \in D} f_0(x) \text{ subject to } \begin{cases} \forall\, i = 1 \dots m: \\ f_i(x) = 0 \end{cases}$$

<u>THM</u>: Let $f_i: \mathbb{R}^n \supseteq D \to \mathbb{R}$ be cont. diff. in neighborhood of local extremum $x^* \in D \Rightarrow \exists\ 0 \neq z^* \in \mathbb{R}^{m+1}$ s.t.

$$L(x, z) := \sum_{i=0}^{m} z_i f_i(x)$$

fulfills for $z = (z^0, \underline{z})$

$$D_{(x,z)} L(x, z)\big|_{(x,z) = (x^*, z^*)} = 0.$$

If $D_x f_i(x^*)$ $i = 1, \dots, m$ are lin. indep. $z_0^* \neq 0$.

<u>ALG</u>: Discussion of boundary of $\{x \in D \mid f_i(x) = 0,\ i = 1 \dots m\}$, non-diff. points, and stationary points of Lagrangian, i.e., for $z = (z^0, \underline{z})$

$$\nabla_{(x,z)} L(x,z) = 0 \qquad (*)$$

$$\Longleftrightarrow \begin{cases} \partial_{x_i} L(x,z) = 0 & i = 1 \ldots n \\ f_i(x) = 0 & i = 1 \ldots m \end{cases}$$

which are $n+m$ equations for $n+m+1$ unknowns as $x \in \mathbb{R}^n$ and $z \in \mathbb{R}^{m+1}$

Again one has to work backwards as $(*)$ is necessary but not sufficient.

But in case $\nabla_x f_i(x^*)$ $i = 1 \ldots m$ are lin. indep. z_0 can be chosen equal one which reduces the # of unknowns to $n+m$.

③ 1951 Karush–Kuhn–Tucker theorem

$$\min_{x \in \mathbb{R}^n} f_0(x) \quad \text{subject to} \quad \begin{cases} \forall i = 1 \ldots m \\ f_i(x) \leq 0 \\ x \in C \text{ convex} \end{cases} \qquad (P)$$

DEF: A subset of linear space is called convex if $\forall x, y \in A$ also $\{\alpha x + (1-\alpha) y \mid \alpha \in [0,1]\} \subseteq A$.

A function f is called convex if Jensen's inequality holds for any two points x, y in the domain $\forall\, 0 \leq \alpha \leq 1$

$$f(\alpha x + (1-\alpha) y) \leq \alpha f(x) + (1-\alpha) f(y)$$

THM: Let $f_i : \mathbb{R}^n \to \mathbb{R}$, $i = 0, \ldots, m$ be convex.

- If x^* is a solution to (P), then:

  $\exists\, 0 \neq z^* \in \mathbb{R}^{m+1}$ s.t.

  (i) $\inf_{x \in C} L(x, z^*) = L(x^*, z^*)$

  (ii) $z_i^* \geq 0$ for $i = 0 \ldots m$

  (iii) $z_i^* f_i(x^*) = 0$ for $i = 1 \ldots m$

- If $\exists\, z^* \in \mathbb{R}^{m+1}$ with $z_0^* \neq 0$, then:

  (i)–(iii) $\Rightarrow$ x^* solves (P)

- If $\exists\, \bar{x} \in C$ s.t. $\forall i = 1 \ldots m : f_i(\bar{x}) < 0$ } Slater cond.

  $\Rightarrow z_0^* \neq 0$

- If Slater cond. holds

(i)–(iii) $\Leftrightarrow$ $\exists\, x^* \in C,\ z^* \in \mathbb{R}^{m+1},\ z_0^* \neq 0$

s.t. $\forall\, x \in C,\ z \in \mathbb{R}^{m+1},\ z_0 \neq 0$

$$L\left(x, \frac{z^*}{z_0^*}\right) \geq L\left(x^*, \frac{z^*}{z_0^*}\right) \geq L\left(x^*, \frac{z}{z_0}\right)$$

and $\inf_{x \in C} L\left(x, \frac{z^*}{z_0^*}\right) = L\left(x^*, \frac{z^*}{z_0^*}\right) = \sup_{\substack{z \in \mathbb{R}^{m+1} \\ z_0 \neq 0}} L\left(x^*, \frac{z}{z_0}\right)$

The goal will be to understand and prove KKT based on some results of convex set theory.