Problem: \( \min_{x \in X} f(x) \), where:

(a) \( X \subset \mathbb{R}^n \) is nonempty, convex, and closed.
(b) \( f \) is continuously differentiable over \( X \).

- Local and global minima. If \( f \) is convex local minima are also global.
Proposition (Optimality Condition)

(a) If \( x^* \) is a local minimum of \( f \) over \( X \), then

\[
\nabla f(x^*)' (x - x^*) \geq 0, \quad \forall \ x \in X.
\]

(b) If \( f \) is convex over \( X \), then this condition is also sufficient for \( x^* \) to minimize \( f \) over \( X \).

At a local minimum \( x^* \), the gradient \( \nabla f(x^*) \) makes an angle less than or equal to 90 degrees with all feasible variations \( x - x^* \), \( x \in X \).

Illustration of failure of the optimality condition when \( X \) is not convex. Here \( x^* \) is a local min but we have \( \nabla f(x^*)' (x - x^*) < 0 \) for the feasible vector \( x \) shown.
Proof: (a) Suppose that \( \nabla f(x^*)' (x - x^*) < 0 \) for some \( x \in X \). By the Mean Value Theorem, for every \( \epsilon > 0 \) there exists an \( s \in [0, 1] \) such that

\[
f(x^* + \epsilon(x - x^*)) = f(x^*) + \epsilon \nabla f(x^* + s\epsilon(x - x^*))' (x - x^*).
\]

Since \( \nabla f \) is continuous, for suff. small \( \epsilon > 0 \),

\[
\nabla f(x^* + s\epsilon(x - x^*))' (x - x^*) < 0
\]

so that \( f(x^* + \epsilon(x - x^*)) < f(x^*) \). The vector \( x^* + \epsilon(x - x^*) \) is feasible for all \( \epsilon \in [0, 1] \) because \( X \) is convex, so the local optimality of \( x^* \) is contradicted.

(b) Using the convexity of \( f \)

\[
f(x) \geq f(x^*) + \nabla f(x^*)' (x - x^*)
\]

for every \( x \in X \). If the condition \( \nabla f(x^*)' (x - x^*) \geq 0 \) holds for all \( x \in X \), we obtain \( f(x) \geq f(x^*) \), so \( x^* \) minimizes \( f \) over \( X \). Q.E.D.
OPTIMIZATION SUBJECT TO BOUNDS

- Let $X = \{ x \mid x \geq 0 \}$. Then the necessary condition for $x^* = (x_1^*, \ldots, x_n^*)$ to be a local min is

$$ \sum_{i=1}^{n} \frac{\partial f(x^*)}{\partial x_i} (x_i - x_i^*) \geq 0, \quad \forall x_i \geq 0, \, i = 1, \ldots, n. $$

- Fix $i$. Let $x_j = x_j^*$ for $j \neq i$ and $x_i = x_i^* + 1$:

$$ \frac{\partial f(x^*)}{\partial x_i} \geq 0, \quad \forall i. $$

- If $x_i^* > 0$, let also $x_j = x_j^*$ for $j \neq i$ and $x_i = \frac{1}{2} x_i^*$. Then $\frac{\partial f(x^*)}{\partial x_i} \leq 0$, so

$$ \frac{\partial f(x^*)}{\partial x_i} = 0, \quad \text{if} \, x_i^* > 0. $$
OPTIMIZATION OVER A SIMPLEX

\[ X = \left\{ x \mid x \geq 0, \sum_{i=1}^{n} x_i = r \right\} \]

where \( r > 0 \) is a given scalar.

• Necessary condition for \( x^* = (x^*_1, \ldots, x^*_n) \) to be a local min:

\[
\sum_{i=1}^{n} \frac{\partial f(x^*)}{\partial x_i} (x_i - x^*_i) \geq 0, \quad \forall x_i \geq 0 \text{ with } \sum_{i=1}^{n} x_i = r.
\]

• Fix \( i \) with \( x^*_i > 0 \) and let \( j \) be any other index. Use \( x \) with \( x_i = 0, x_j = x^*_j + x^*_i, \) and \( x_m = x^*_m \) for all \( m \neq i, j \):

\[
\left( \frac{\partial f(x^*)}{\partial x_j} - \frac{\partial f(x^*)}{\partial x_i} \right) x^*_i \geq 0,
\]

\[ x^*_i > 0 \quad \Rightarrow \quad \frac{\partial f(x^*)}{\partial x_i} \leq \frac{\partial f(x^*)}{\partial x_j}, \quad \forall j. \]
OPTIMAL ROUTING

- Given a data net, and a set $W$ of OD pairs $w = (i, j)$. Each OD pair $w$ has input traffic $r_w$.

- Optimal routing problem:

$$
\text{minimize } D(x) = \sum_{(i,j)} D_{ij} \left( \sum_{\text{all paths } p \text{ containing } (i,j)} x_p \right)
$$

subject to

$$
\sum_{p \in P_w} x_p = r_w, \quad \forall \ w \in W,
$$

$$
x_p \geq 0, \quad \forall \ p \in P_w, \ w \in W
$$

- Optimality condition

$$
x_p^* > 0 \quad \Rightarrow \quad \frac{\partial D(x^*)}{\partial x_p} \leq \frac{\partial D(x^*)}{\partial x_{p'}}, \quad \forall \ p' \in P_w.
$$


TRAFFIC ASSIGNMENT

- Transportation network with OD pairs $w$. Each $w$ has paths $p \in P_w$ and traffic $r_w$. Let $x_p$ be the flow of path $p$ and let $T_{ij}\left(\sum_p: \text{crossing } (i,j) x_p\right)$ be the travel time of link $(i, j)$.

- User-optimization principle: Traffic equilibrium is established when each user of the network chooses, among all available paths, a path requiring minimum travel time, i.e., for all $w \in W$ and paths $p \in P_w$,

\[ x_p^* > 0 \implies t_p(x^*) \leq t_{p'}(x^*), \quad \forall p' \in P_w, \forall w \in W \]

where $t_p(x)$, is the travel time of path $p$

\[ t_p(x) = \sum_{\text{all arcs } (i,j) \text{ on path } p} T_{ij}(F_{ij}), \quad \forall p \in P_w, \forall w \in W. \]

Identical with the optimality condition of the routing problem if we identify the arc travel time $T_{ij}(F_{ij})$ with the cost derivative $D'_{ij}(F_{ij})$. 
PROJECTION OVER A CONVEX SET

• Let $z \in \mathbb{R}^n$ and a closed convex set $X$ be given. Problem:

  minimize $f(x) = \|z - x\|^2$

  subject to $x \in X$.

Proposition (Projection Theorem) Problem has a unique solution $[z]^+$ (the projection of $z$).

Necessary and sufficient condition for $x^*$ to be the projection. The angle between $z - x^*$ and $x - x^*$ should be greater or equal to 90 degrees for all $x \in X$, or $(z - x^*)'(x - x^*) \leq 0$

• If $X$ is a subspace, $z - x^* \perp X$.

• The mapping $f : \mathbb{R}^n \leftrightarrow X$ defined by $f(x) = [x]^+$ is continuous and nonexpansive, that is,

  $$\|[x]^+ - [y]^+\| \leq \|x - y\|, \quad \forall x, y \in \mathbb{R}^n.$$