LECTURE 14

LECTURE OUTLINE

• Conical approximations
• Cone of feasible directions
• Tangent and normal cones
• Conditions for optimality

---

• A basic necessary condition:
  - If $x^*$ minimizes a function $f(x)$ over $x \in X$, then for every $y \in \mathbb{R}^n$, $\alpha^* = 0$ minimizes $g(\alpha) \equiv f(x + \alpha y)$ over the line subset

$$\{\alpha \mid x + \alpha y \in X\}.$$  

• Special cases of this condition ($f$: differentiable):
  - $X = \mathbb{R}^n$: $\nabla f(x^*) = 0$.
  - $X$ is convex: $\nabla f(x^*)'(x - x^*) \geq 0$, $\forall x \in X$.

• We will aim for more general conditions.
CONE OF FEASIBLE DIRECTIONS

- Consider a subset $X$ of $\mathbb{R}^n$ and a vector $x \in X$.
- A vector $y \in \mathbb{R}^n$ is a feasible direction of $X$ at $x$ if there exists an $\alpha > 0$ such that $x + \alpha y \in X$ for all $\alpha \in [0, \alpha]$.
- The set of all feasible directions of $X$ at $x$ is denoted by $F_X(x)$.
- $F_X(x)$ is a cone containing the origin. It need not be closed or convex.
- If $X$ is convex, $F_X(x)$ consists of the vectors of the form $\alpha(\bar{x} - x)$ with $\alpha > 0$ and $\bar{x} \in X$.
- Easy optimality condition: If $x^*$ minimizes a differentiable function $f(x)$ over $x \in X$, then
  \[ \nabla f(x^*)'y \geq 0, \quad \forall y \in F_X(x^*). \]
- Difficulty: The condition may be vacuous because there may be no feasible directions (other than 0).
TANGENT CONE

• Consider a subset $X$ of $\mathbb{R}^n$ and a vector $x \in X$.

• A vector $y \in \mathbb{R}^n$ is said to be a tangent of $X$ at $x$ if either $y = 0$ or there exists a sequence $\{x_k\} \subset X$ such that $x_k \neq x$ for all $k$ and

\[
x_k \to x, \quad \frac{x_k - x}{\|x_k - x\|} \to \frac{y}{\|y\|}.
\]

• The set of all tangents of $X$ at $x$ is called the tangent cone of $X$ at $x$, and is denoted by $T_X(x)$.

• $y$ is a tangent of $X$ at $x$ iff there exists $\{x_k\} \subset X$ with $x_k \to x$, and a positive scalar sequence $\{\alpha_k\}$ such that $\alpha_k \to 0$ and $\left(\frac{x_k - x}{\alpha_k}\right) \to y$. 
• In (a), $X$ is convex: The tangent cone $T_X(x)$ is equal to the closure of the cone of feas. directions $F_X(x)$.

• In (b), $X$ is nonconvex: $T_X(x)$ is closed but not convex, while $F_X(x)$ consists of just the zero vector.

• In general, $F_X(x) \subset T_X(x)$.

• For $X$: polyhedral, $F_X(x) = T_X(x)$. 
RELATION OF CONES

• Let $X$ be a subset of $\mathbb{R}^n$ and let $x$ be a vector in $X$. The following hold.

  (a) $T_X(x)$ is a closed cone.

  (b) $\text{cl}(F_X(x)) \subset T_X(x)$.

  (c) If $X$ is convex, then $F_X(x)$ and $T_X(x)$ are convex, and we have

  $$\text{cl}(F_X(x)) = T_X(x).$$

Proof: (a) Let $\{y_k\}$ be a sequence in $T_X(x)$ that converges to some $y \in \mathbb{R}^n$. We show that $y \in T_X(x)$ ...

(b) Every feasible direction is a tangent, so $F_X(x) \subset T_X(x)$. Since by part (a), $T_X(x)$ is closed, the result follows.

(c) Since $X$ is convex, the set $F_X(x)$ consists of the vectors of the form $\alpha(\bar{x} - x)$ with $\alpha > 0$ and $\bar{x} \in X$. Verify definition of convexity ...
NORMAL CONE

• Consider subset $X$ of $\mathbb{R}^n$ and a vector $x \in X$.
• A vector $z \in \mathbb{R}^n$ is said to be a *normal* of $X$ at $x$ if there exist sequences $\{x_k\} \subset X$ and $\{z_k\}$ with

\[
x_k \to x, \quad z_k \to z, \quad z_k \in T_X(x_k)^*, \quad \forall \ k.
\]

• The set of all normals of $X$ at $x$ is called the *normal cone* of $X$ at $x$ and is denoted by $N_X(x)$.
• Example:

![Diagram of a normal cone](image)

$N_X(x)$ is “usually equal” to the polar $T_X(x)^*$, but may differ at points of “discontinuity” of $T_X(x)$. 
RELATION OF NORMAL AND POLAR CONES

• We have $T_X(x)^* \subset N_X(x)$.

• When $N_X(x) = T_X(x)^*$, we say that $X$ is regular at $x$.

• If $X$ is convex, then for all $x \in X$, we have $z \in T_X(x)^*$ if and only if $z'(\bar{x} - x) \leq 0$, $\forall \bar{x} \in X$.

Furthermore, $X$ is regular at all $x \in X$. In particular, we have

$$T_X(x)^* = N_X(x), \quad T_X(x) = N_X(x)^*.$$  

• Note that convexity of $T_X(x)$ does not imply regularity if $X$ at $x$.

• Important fact in nonsmooth analysis: If $X$ is closed and regular at $x$, then

$$T_X(x) = N_X(x)^*.$$  

In particular, $T_X(x)$ is convex.
OPTIMALITY CONDITIONS I

• Let $f : \mathbb{R}^n \mapsto \mathbb{R}$ be a smooth function. If $x^*$ is a local minimum of $f$ over a set $X \subset \mathbb{R}^n$, then

$$\nabla f(x^*)'y \geq 0, \quad \forall \ y \in T_X(x^*).$$

Proof: Let $y \in T_X(x^*)$ with $y \neq 0$. Then, there exist $\{\xi_k\} \subset \mathbb{R}$ and $\{x_k\} \subset X$ such that $x_k \neq x^*$ for all $k$, $\xi_k \to 0$, $x_k \to x^*$, and

$$(x_k - x^*)/\|x_k - x^*\| = y/\|y\| + \xi_k.$$

By the Mean Value Theorem, we have for all $k$

$$f(x_k) = f(x^*) + \nabla f(\tilde{x}_k)'(x_k - x^*),$$

where $\tilde{x}_k$ is a vector that lies on the line segment joining $x_k$ and $x^*$. Combining these equations,

$$f(x_k) = f(x^*) + (\|x_k - x^*\|/\|y\|)\nabla f(\tilde{x}_k)'y_k,$$

where $y_k = y + \|y\|\xi_k$. If $\nabla f(x^*)'y < 0$, since $\tilde{x}_k \to x^*$ and $y_k \to y$, for sufficiently large $k$, $\nabla f(\tilde{x}_k)'y_k < 0$ and $f(x_k) < f(x^*)$. This contradicts the local optimality of $x^*$. 
OPTIMALITY CONDITIONS II

Let $f : \mathbb{R}^n \mapsto \mathbb{R}$ be a convex function. A vector $x^*$ minimizes $f$ over a convex set $X$ if and only if there exists a subgradient $d \in \partial f(x^*)$ such that

$$d'(x - x^*) \geq 0, \quad \forall \ x \in X.$$ 

Proof: If for some $d \in \partial f(x^*)$ and all $x \in X$, we have $d'(x - x^*) \geq 0$, then, from the definition of a subgradient we have $f(x) - f(x^*) \geq d'(x - x^*)$ for all $x \in X$. Hence $f(x) - f(x^*) \geq 0$ for all $x \in X$.

Conversely, suppose that $x^*$ minimizes $f$ over $X$. Then, $x^*$ minimizes $f$ over the closure of $X$, and we have

$$f'(x^*; x-x^*) = \sup_{d \in \partial f(x^*)} d'(x-x^*) \geq 0, \quad \forall \ x \in \text{cl}(X).$$

Therefore,

$$\inf_{x \in \text{cl}(X) \cap \{z \mid \|z-x^*\| \leq 1\}} \sup_{d \in \partial f(x^*)} d'(x - x^*) = 0.$$ 

Apply the saddle point theorem to conclude that ”infsup=supinf” and that the supremum is attained by some $d \in \partial f(x^*)$. 
OPTIMALITY CONDITIONS III

• Let \( x^* \) be a local minimum of a function \( f : \mathbb{R}^n \rightarrow \mathbb{R} \) over a subset \( X \) of \( \mathbb{R}^n \). Assume that the tangent cone \( T_X(x^*) \) is convex, and that \( f \) has the form

\[
f(x) = f_1(x) + f_2(x),
\]

where \( f_1 \) is convex and \( f_2 \) is smooth. Then

\[
-\nabla f_2(x^*) \in \partial f_1(x^*) + T_X(x^*)^*.
\]

• The convexity assumption on \( T_X(x^*) \) (which is implied by regularity) is essential in general.

• Example: Consider the subset of \( \mathbb{R}^2 \)

\[
X = \{(x_1, x_2) \mid x_1 x_2 = 0\}.
\]

Then \( T_X(0)^* = \{0\} \). Take \( f \) to be any convex non-differentiable function for which \( x^* = 0 \) is a global minimum over \( X \), but \( x^* = 0 \) is not an unconstrained global minimum. Such a function violates the necessary condition.