SE/EC/ME 724 Advanced Optimization Theory and Methods

Yannis Paschalidis
yannisp@bu.edu, http://sites.bu.edu/paschalidis/

Department of Electrical and Computer Engineering,
Division of Systems Engineering,
Department of Biomedical Engineering
and Center for Information and Systems Engineering,
Boston University

Lecture 3: Outline

1. Derivatives, differentiable convex functions, subgradients.
2. Unconstrained optimization: basic definitions.
3. Existence of optimal solutions.
4. Unconstrained optimization: Optimality conditions.

Derivatives

- Gradient:
  - $f : \mathbb{R}^n \to \mathbb{R} \Rightarrow \nabla f(\mathbf{x}) = (\frac{\vartheta f(\mathbf{x})}{\vartheta x_1}, \ldots, \frac{\vartheta f(\mathbf{x})}{\vartheta x_n})$.
  - $f = (f_1, \ldots, f_m) : \mathbb{R}^n \to \mathbb{R}^m \Rightarrow \nabla f(\mathbf{x}) = [\nabla f_1(\mathbf{x}) \cdots \nabla f_m(\mathbf{x})]$.
- Hessian:
  - $f : \mathbb{R}^n \to \mathbb{R} \Rightarrow \nabla^2 f(\mathbf{x}) = \nabla(\nabla f(\mathbf{x})) = \left(\frac{\vartheta^2 f(\mathbf{x})}{\vartheta x_i \vartheta x_j}\right)$.
- Taylor expansion:

$$f(\mathbf{x}+\mathbf{y}) = f(\mathbf{x}) + \mathbf{y}'\nabla f(\mathbf{x}) + \frac{1}{2}\mathbf{y}'\nabla^2 f(\mathbf{x})\mathbf{y} + o(||\mathbf{y}||^2).$$

Subgradients

Definition

Let $f : \mathbb{R}^n \to \mathbb{R}$ be convex. A $\mathbf{d} \in \mathbb{R}^n$ is a **subgradient** of f at $\mathbf{x} \in \mathbb{R}^n$ if $f(\mathbf{z}) \geq f(\mathbf{x}) + (\mathbf{z}-\mathbf{x})'\mathbf{d}$, $\forall \mathbf{z} \in \mathbb{R}^n$. If f is concave, $\mathbf{d} \in \mathbb{R}^n$ is a **subgradient** of f at $\mathbf{x} \in \mathbb{R}^n$ if $-\mathbf{d}$ is a subgradient of $-f$ at $\mathbf{x}$.

Definition

The **subdifferential** of f at $\mathbf{x} \in \mathbb{R}^n$, denoted by $\vartheta f(\mathbf{x})$, is the set of all subgradients at $\mathbf{x}$.

Proposition

Let $f : \mathbb{R}^n \to \mathbb{R}$ be convex. For all $\mathbf{x} \in \mathbb{R}^n$ $\vartheta f(\mathbf{x})$ is nonempty, convex, compact, and it holds

$$\dot{f}(\mathbf{x};\mathbf{y}) = \lim_{a \downarrow 0} \frac{f(\mathbf{x}+a\mathbf{y}) - f(\mathbf{x})}{a} = \max_{\mathbf{d} \in \vartheta f(\mathbf{x})} \mathbf{y}'\mathbf{d}, \quad \forall \mathbf{y} \in \mathbb{R}^n.$$

Unconstrained optimization

- Unconstrained optimization problem:

$$\begin{aligned} \min \quad & f(\mathbf{x}) \\ \text{s.t.} \quad & \mathbf{x} \in \mathbb{R}^n \end{aligned}$$

- $\mathbf{x}^*$ is a **local minimum** if $\exists \epsilon > 0$ s.t. $f(\mathbf{x}^*) \leq f(\mathbf{x})$ $\forall \mathbf{x}$ with $||\mathbf{x} - \mathbf{x}^*|| < \epsilon$.
- $\mathbf{x}^*$ is a **global minimum** if $f(\mathbf{x}^*) \leq f(\mathbf{x})$ $\forall \mathbf{x} \in \mathbb{R}^n$.

Existence of minima

Theorem

(**Weierstrass**) Let $\mathscr{A} \subset \mathbb{R}^n$ where $\mathscr{A}$ is closed and non-empty. Let $f : \mathscr{A} \to \mathbb{R}^n$ be lower-semicontinuous for all $\mathbf{x} \in \mathscr{A}$.

a. If $\mathscr{A}$ is compact then $\exists \mathbf{x} \in \mathscr{A}$ s.t. $f(\mathbf{x}) = \inf_{\mathbf{z} \in \mathscr{A}} f(\mathbf{z})$.
b. If f is coercive, then $\exists \mathbf{x} \in \mathscr{A}$ s.t. $f(\mathbf{x}) = \inf_{\mathbf{z} \in \mathscr{A}} f(\mathbf{z})$.

Necessary Conditions

Proposition

Let $\mathbf{x}^*$ be an unconstrained local min and $f : \mathbb{R}^n \to \mathbb{R}$ continuously differentiable in an open set $\mathscr{S}$ containing $\mathbf{x}^*$. Then

$$\nabla f(\mathbf{x}^*) = \mathbf{0}. \qquad (\textbf{1st order})$$

If $f : \mathbb{R}^n \to \mathbb{R}$ is twice continuously differentiable within $\mathscr{S}$ then

$$\nabla^2 f(\mathbf{x}^*) \succeq 0. \qquad (\textbf{2nd order})$$

Sufficient Conditions

Proposition

Let $f : \mathbb{R}^n \to \mathbb{R}$ be twice continuously differentiable in an open set $\mathscr{S} \subset \mathbb{R}^n$. Let also $\mathbf{x}^* \in \mathscr{S}$ s.t.

$$\nabla f(\mathbf{x}^*) = \mathbf{0},$$

$$\nabla^2 f(\mathbf{x}^*) \succ 0.$$

Then $\mathbf{x}^*$ is a **strict** unconstrained local min of f, that is, $\exists \gamma, \epsilon > 0$ s.t.

$$f(\mathbf{x}) \geq f(\mathbf{x}^*) + \frac{\gamma}{2}||\mathbf{x} - \mathbf{x}^*||^2, \qquad \forall \mathbf{x} \text{ with } ||\mathbf{x} - \mathbf{x}^*|| < \epsilon.$$