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:**
  \[ f : \mathbb{R}^n \to \mathbb{R} \Rightarrow \nabla f(x) = \left( \frac{\partial f(x)}{\partial x_1}, \ldots, \frac{\partial f(x)}{\partial x_n} \right). \]
  \[ f = (f_1, \ldots, f_m) : \mathbb{R}^n \to \mathbb{R}^m \Rightarrow \nabla f(x) = [\nabla f_1(x) \cdots \nabla f_m(x)]. \]

- **Hessian:**
  \[ f : \mathbb{R}^n \to \mathbb{R} \Rightarrow \nabla^2 f(x) = \nabla(\nabla f(x)) = \left( \frac{\partial^2 f(x)}{\partial x_i \partial x_j} \right). \]

- **Taylor expansion:**
  \[ f(x + y) = f(x) + y' \nabla f(x) + \frac{1}{2} y' \nabla^2 f(x)y + o(||y||^2). \]

Subgradients

**Definition**

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

**Definition**

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

**Proposition**

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

\[ \tilde{f}(x; y) = \lim_{a \downarrow 0} \frac{f(x + ay) - f(x)}{a} = \max_{d \in \partial f(x)} y'd, \quad \forall y \in \mathbb{R}^n. \]
Unconstrained optimization

- Unconstrained optimization problem:
  \[
  \min \ f(x) \\
  \text{s.t.} \quad x \in \mathbb{R}^n
  \]

- \(x^*\) is a **local minimum** if \(\exists \epsilon > 0\) s.t. \(f(x^*) \leq f(x)\ \forall x\) with \(||x - x^*|| < \epsilon\).

- \(x^*\) is a **global minimum** if \(f(x^*) \leq f(x)\ \forall x \in \mathbb{R}^n\).

Existence of minima

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

- If \(\mathcal{A}\) is compact then \(\exists x \in \mathcal{A}\) s.t. \(f(x) = \inf_{z \in \mathcal{A}} f(z)\).
- If \(f\) is coercive, then \(\exists x \in \mathcal{A}\) s.t. \(f(x) = \inf_{z \in \mathcal{A}} f(z)\).
Necessary Conditions

**Proposition**
Let \( x^* \) be an unconstrained local min and \( f : \mathbb{R}^n \rightarrow \mathbb{R} \) continuously differentiable in an open set \( \mathcal{S} \) containing \( x^* \). Then
\[
\nabla f(x^*) = 0. \quad (1st \ order)
\]
If \( f : \mathbb{R}^n \rightarrow \mathbb{R} \) is twice continuously differentiable within \( \mathcal{S} \) then
\[
\nabla^2 f(x^*) \succeq 0. \quad (2nd \ order)
\]

Sufficient Conditions

**Proposition**
Let \( f : \mathbb{R}^n \rightarrow \mathbb{R} \) be twice continuously differentiable in an open set \( \mathcal{S} \subset \mathbb{R}^n \). Let also \( x^* \in \mathcal{S} \) s.t.
\[
\nabla f(x^*) = 0,
\]
\[
\nabla^2 f(x^*) \succ 0.
\]
Then \( x^* \) is a **strict** unconstrained local min of \( f \), that is, \( \exists \gamma, \epsilon > 0 \) s.t.
\[
f(x) \geq f(x^*) + \frac{\gamma}{2} ||x - x^*||^2, \quad \forall x \text{ with } ||x - x^*|| < \epsilon.
\]