6.252 - Lecture 2

Unconstrained Optimization - Optimality Conditions

February 5, 2015

Massachusetts Institute of Technology
Lecture Outline

- Unconstrained Optimization
- Local Minima
- Necessary Conditions for Local Minima
- Sufficient Conditions for Local Minima
- The Role of Convexity
- Quadratic Unconstrained Problems
- Existence of Optimal Solutions
Mathematical Background - Recitation

- Vectors and matrices in $\mathbb{R}^n$
- Transpose, inner product, norm
- Eigenvalues of symmetric matrices
- Positive definite and semidefinite matrices
- Convergent sequences and subsequences
- Open, closed, and compact sets
- Continuity of functions
- 1st and 2nd order differentiability of functions
- Taylor series expansions
- Mean value theorems
Local and Global Minima

Unconstrained local and global minima in one dimension
Necessary Conditions for a Local Minimum

- **1st order condition:** Zero slope at a local minimum $x^*$:
  \[ \nabla f(x^*) = 0 \]

- **2nd order condition:** Nonnegative curvature at a local minimum $x^*$:
  \[ \nabla^2 f(x^*) \succeq 0 \]  \text{(Hessian is Positive Semidefinite)}

- There may exist points that satisfy the 1st and 2nd order conditions but are not local minima

First and second order necessary optimality conditions for functions of one variable

\[ f(x) = |x|^3 \text{ (convex)} \]
\[ f(x) = x^3 \]
\[ f(x) = - |x|^3 \]

\( x^* = 0 \) \( x^* = 0 \) \( x^* = 0 \)
Proofs of Necessary Conditions

• **1st order condition** \( \nabla f(x^*) = 0 \): Fix \( d \in \mathbb{R}^n \). Then (since \( x^* \) is a local min), from 1st order Taylor

\[
\lim_{\alpha \downarrow 0} \frac{f(x^* + \alpha d) - f(x^*)}{\alpha} \geq 0,
\]

Replace \( d \) with \(-d\), to obtain

\[
d' \nabla f(x^*) = 0, \quad \forall \ d \in \mathbb{R}^n
\]

• **2nd order condition** \( \nabla^2 f(x^*) \geq 0 \): From 2nd order Taylor

\[
f(x^* + \alpha d) - f(x^*) = \alpha \nabla f(x^*)'d + \frac{\alpha^2}{2} d' \nabla^2 f(x^*)d + o(\alpha^2)
\]

Since \( \nabla f(x^*) = 0 \) and \( x^* \) is local min, there is sufficiently small \( \epsilon > 0 \) such that for all \( \alpha \in (0, \epsilon) \),

\[
0 \leq \frac{f(x^* + \alpha d) - f(x^*)}{\alpha^2} = \frac{1}{2} d' \nabla^2 f(x^*)d + \frac{o(\alpha^2)}{\alpha^2}
\]

Take the limit as \( \alpha \to 0 \).
Sufficient Conditions for a Local Minimum

• **1st order condition:** Zero slope:

\[ \nabla f(x^*) = 0 \]

• **2nd order condition:** Positive curvature:

\[ \nabla^2 f(x^*) > 0 \quad \text{(Hessian is Positive Definite)} \]

• **Proof:** Let \( \lambda > 0 \) be the smallest eigenvalue of \( \nabla^2 f(x^*) \). Using a second order Taylor expansion, we have for all \( d \)

\[
f(x^* + d) - f(x^*) = \nabla f(x^*)' d + \frac{1}{2} d' \nabla^2 f(x^*) d + o(\|d\|^2)
\]

\[
\geq \frac{\lambda}{2} \|d\|^2 + o(\|d\|^2)
\]

\[
= \left( \frac{\lambda}{2} + \frac{o(\|d\|^2)}{\|d\|^2} \right) \|d\|^2.
\]

For \( \|d\| \) small enough, \( o(\|d\|^2)/\|d\|^2 \) is negligible relative to \( \lambda/2 \).
Convexity

Convex Sets

Nonconvex Sets

Convex and nonconvex sets
A convex function. The linear interpolation $\alpha f(x) + (1 - \alpha)f(y)$ overestimates the function value $f(\alpha x + (1 - \alpha)y)$. 
Minima and Convexity

- Local minima are also global under convexity

Illustration of why local minima of convex functions are also global. Suppose that $f$ is convex and that $x^*$ is a local minimum of $f$. Let $\bar{x}$ be such that $f(\bar{x}) < f(x^*)$.

By convexity, for all $\alpha \in (0, 1)$,

$$f(\alpha x^* + (1 - \alpha)\bar{x}) \leq \alpha f(x^*) + (1 - \alpha) f(\bar{x}) < f(x^*).$$

Thus, $f$ takes values strictly lower than $f(x^*)$ on the line segment connecting $x^*$ with $\bar{x}$, and $x^*$ cannot be a local minimum which is not global.
Other Properties of Convex Functions

• $f$ is convex if and only if the linear approximation at a point $x$ based on the gradient, underestimates $f$:

$$f(z) \geq f(x) + \nabla f(x)'(z - x), \quad \forall \ z \in \mathbb{R}^n$$

- Implication:

$$\nabla f(x^*) = 0 \implies x^* \text{ is a global minimum}$$

• $f$ is convex if and only if $\nabla^2 f(x)$ is positive semidefinite for all $x$
Quadratic Unconstrained Problems

\[
\min_{x \in \mathbb{R}^n} f(x) = \frac{1}{2} x' Q x - b' x,
\]

where \( Q \) is \( n \times n \) symmetric, and \( b \in \mathbb{R}^n \).

- Necessary conditions:
  \[
  \nabla f(x^*) = Q x^* - b = 0,
  \]
  \[
  \nabla^2 f(x^*) = Q \succeq 0 : \text{positive semidefinite}.
  \]

- \( Q \succeq 0 \implies f : \text{convex, nec. conditions are also sufficient, and local minima are also global} \)

- Conclusions:
  - \( Q : \text{not } \succeq 0 \implies f \text{ has no local minima} \)
  - If \( Q \succ 0 \) (hence invertible), \( x^* = Q^{-1} b \) is unique global minimum.
  - If \( Q \succeq 0 \) but not invertible, either no solution or \( \infty \) number of solutions.
Illustration of the isocost surfaces of the quadratic cost function

\[ f(x, y) = \frac{1}{2}(\alpha x^2 + \beta y^2) - x \]

for various values of \(\alpha\) and \(\beta\).
Existence of Optimal Solutions

Consider the problem

$$\min_{x \in X} f(x)$$

- The set of optimal solutions is

$$X^* = \cap_{k=1}^{\infty} \{ x \in X \mid f(x) \leq \gamma_k \}$$

where \( \{\gamma_k\} \) is a scalar sequence such that \( \gamma_k \downarrow f^* \) with

$$f^* = \inf_{x \in X} f(x)$$

- \( X^* \) is nonempty and compact if all the sets \( \{ x \in X \mid f(x) \leq \gamma_k \} \) are compact. So:
  - A global minimum exists if \( f \) is continuous and \( X \) is compact
    (Weierstrass theorem)
  - A global minimum exists if \( X \) is closed, and \( f \) is continuous and coercive, that is, \( f(x) \to \infty \) when \( \|x\| \to \infty \)
Homework

• Lecture 1-2 readings:
  – Section 1.1.1, 1.1.2

• Homework set 1:
  – Problems 1.1.1; 1.1.2 parts (a)-(d); 1.1.3; 1.1.5.

• Homework is due Thursday, 2/12, at the beginning of the lecture.