ISEN 629: Engineering Optimization

Lecture 2

Sergiy Butenko

Industrial and Systems Engineering
Texas A& M University

Fall 2007

Local Methods in Unconstrained Optimization

- We consider an unconstrained optimization problem in the form
$$\min_{x \in \mathbb{R}^n} f(x) \tag{1}$$
where $f : \mathbb{R}^n \to \mathbb{R}$ is a smooth function.
- We will consider several methods attempting to solve this problem by generating a descending sequence $\{f(x_k) : k \geq 0\}$:
$$f(x_{k+1}) \leq f(x_k), \; k = 0, 1, \dots.$$
- Note that if $f(x)$ is bounded from below on $\mathbb{R}^n$, then the sequence $\{f(x_k) : k \geq 0\}$ converges.
- In any case, we improve the initial value of the objective.

Preliminaries

- By vector $x \in \mathbb{R}^n$ we will mean a column-vector $x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$.
- Then the corresponding row-vector is $x^T = [x_1, \dots, x_n]$.
- Unless otherwise stated, for a vector $x \in \mathbb{R}^n$, by $\|x\|$ we will mean the second norm $\|x\|_2$, that is $\|x\| = \sqrt{\sum_{i=1}^{n} x_i^2}$.
- We call $x_0 \in X$ an *interior point* of X if there exists $\varepsilon > 0$ such that $B_\varepsilon(x_0) \subset X$. If $x_0 \in X$ is not an interior point then we call it a *boundary* point of X.
- A set is open if all its points are interior.
- A set is closed if it contains all its boundary points.
- A set X is bounded if there exists $C > 0$ such that for any $x \in X : \; \|x\| \leq C$.
- A compact set is a closed bounded set.

Preliminaries

- For $x_0 \in \mathbb{R}^n$, let $B_\varepsilon(x_0) = \{x \in \mathbb{R}^n : \|x - x_0\| < \varepsilon\}$ denote the open ε-ball in x_0 and $\bar{B}_\varepsilon(x_0) = \{x \in \mathbb{R}^n : \|x - x_0\| \leq \varepsilon\}$ - the closed ε-ball in x_0.
- A vector $d \in \mathbb{R}^n$ is called a *feasible direction* for problem (1) at $x_0 \in X$ if there exists $\delta > 0$ such that $x_0 + \alpha d \in X$ for any $\alpha \leq \delta$. Note that any direction is feasible at an interior point.
- For $f : \mathbb{R}^n \to \mathbb{R}$ and $x_0, d \in \mathbb{R}^n$, the directional derivative of $f(x)$ at x_0 in the direction d is defined as
$$\frac{\partial f(x_0)}{\partial d} = \lim_{\alpha \to 0} \frac{f(x_0 + \alpha d) - f(x_0)}{\alpha}.$$
If $\|d\| = 1$, then $\frac{\partial f(x_0)}{\partial d}$ is also called the rate of increase of f at x_0 in the direction d.

Preliminaries

For two functions $f, g : \mathbb{R}^n \to \mathbb{R}$ and $x_0 \in \mathbb{R}^n$ we say that

- $f(x) = O(g(x))$ if $|f(x)|/|g(x)|$ is bounded near x_0, *i.e.*, there exist numbers $\varepsilon > 0$ and $C > 0$ such that $|f(x)| \le C|g(x)|$ for all $x \in B_\varepsilon(x_0)$.
- $f(x) = o(g(x))$ if $\lim\limits_{x \to x_0} \frac{|f(x)|}{|g(x)|} = 0$.

Preliminaries

Assume that f is differentiable. Denoting by $\phi(\alpha) = f(x_0 + \alpha d)$, we have

$$\begin{aligned}\frac{\partial f(x_0)}{\partial d} &= \lim_{\alpha \to 0} \frac{f(x_0 + \alpha d) - f(x_0)}{\alpha} \\ &= \lim_{\alpha \to 0} \frac{\phi(\alpha) - \phi(0)}{\alpha} \\ &= \phi'(0).\end{aligned}$$

On the other hand, using the chain rule,

$$\phi'(\alpha) = \nabla f(x_0 + \alpha d)^T d \quad \Rightarrow \quad \phi'(0) = \nabla f(x_0)^T d.$$

So, we obtain

$$\frac{\partial f(x_0)}{\partial d} = \nabla f(x_0)^T d.$$

First-order necessary conditions

Theorem (**FONC for a set-constrained problem)**

If $f : \mathbb{R}^n \to \mathbb{R}$ is a continuously differentiable function and x_0 is a point of local minimum for problem (1), then for any feasible direction d at x_0 the rate of increase of f at x_0 in the direction d is nonnegative:

$$\nabla f(x_0)^T d \ge 0.$$

Corollary: If x_0 is a local minimizer and an interior point of X then

$$\nabla f(x_0) = 0.$$

Theorem (**FONC for an unconstrained problem)**

If x_0 is a point of local minimum of the unconstrained problem $\min\limits_{x \in \mathbb{R}^n} f(x)$, where $f \in C^{(1)}(\mathbb{R}^n)$, then

$$\nabla f(x_0) = 0.$$

FONC and existence of a global minimum

The following example shows that FONC is not sufficient for a local minimizer. We consider

$$f(x) = x^3$$

and apply the FONC. We have

$$f'(x) = 3x^2 = 0 \quad \Leftrightarrow \quad x = 0.$$

But, obviously, $x = 0$ is not a local minimizer of $f(x)$. In fact, for any given point x_0 and any small $\varepsilon > 0$ there always exist $x_*, x^* \in B_\varepsilon(x_0)$ such that $f(x_*) < f(x_0) < f(x^*)$, so $f(x)$ does not have any local or global minimum or maximum.

Existence of a global minimum

- An important question is, when does a minimum exist? This question is, in general, very difficult to answer.
- Even a bounded function may not have any minimizers. For example, the function $f(x) = e^x$ is bounded ($f(x) > 0$ for any x), but it does not have any local or global minimizers.
- However, there are some cases where the existence of a global maximum is guaranteed. One of such cases is given by the Weierstrass theorem:

Theorem (Weierstrass)

For a continuous function $f : \mathbb{R}^n \to \mathbb{R}$ and a compact set $X \subset \mathbb{R}^n$ there always exist $x_*, x^* \in X$ such that

$$f(x_*) = \min_{x \in X} f(x) \text{ and } f(x^*) = \max_{x \in X} f(x).$$

Existence of a global minimum

Another special case, where a global minimum is guaranteed to exist is represented by *coercive functions*. A function $f : \mathbb{R}^n \to \mathbb{R}$ is called coercive, if

$$\lim_{\|x\| \to \infty} f(x) = +\infty.$$

Theorem

Any coercive continuous function $f : \mathbb{R}^n \to \mathbb{R}$ has at least one point of global minimum in $\mathbb{R}^n$.

Convex sets and functions

- Convex problems represent one of the most important categories of optimization problems. Their very special properties put them among the few "tractable" optimization problem classes.
- A set X is said to be convex if for any $x, y \in X$ and any $\alpha \in (0, 1): \ \alpha x + (1 - \alpha)y \in X$. In other words, all points located on the line segment connecting x and y are in X.
- Given a convex set X, a function $f : X \to \mathbb{R}$ is convex if for any $x_1, x_2 \in X$ and any $\alpha \in (0, 1)$ we have $f(\alpha x_1 + (1 - \alpha)x_2) \leq \alpha f(x_1) + (1 - \alpha)f(x_2)$.
- If the last inequality is strict, *i.e.*, $f(\alpha x_1 + (1 - \alpha)x_2) < \alpha f(x_1) + (1 - \alpha)f(x_2)$ for any $x_1, x_2 \in X, \alpha \in (0, 1)$, then $f(x)$ is called strictly convex.

Convex problems

A problem $\min_{x \in X} f(x)$ is called a convex minimization problem (or simply a convex problem) if $f(x)$ is a convex function and X is a convex set.

Theorem

Any local minimum of a convex problem is global.

First order characterization of a convex function

Theorem

If $f(x)$ is differentiable on a convex set X then it is convex if and only if for any $x, y \in X$:

$$f(y) \geq f(x) + \nabla f(x)^T (y - x).$$

FONC are sufficient for convex problems

Consider a convex problem $\min_{x \in X} f(x)$. For any $x, y \in X$ a differentiable convex function satisfies

$$f(y) \geq f(x) + \nabla f(x)^T (y - x),$$

so, if for $x = x^*$: $\nabla f(x^*) = 0$, we obtain

$$f(y) \geq f(x^*),$$

for any $y \in X$. Thus, x^* is a point of global minimum for this problem. This implies that the FONC in the case of an unconstrained convex problem become sufficient conditions. In other words, for an unconstrained problem

$$\min_{x \in \mathbb{R}^n} f(x),$$

where $f(x)$ is differentiable and convex, x^* is a global minimizer if and only if $\nabla f(x^*) = 0$.

Quadratic forms and functions

- A quadratic form $q(x)$ is given by

$$q(x) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j,$$

  where a_{ij}, $i, j = 1, \dots, n$ are the coefficients of the quadratic form, $x = [x_1, \dots, x_n]^T$ is the vector of variables.
- Let $A = [a_{i,j}]_{n \times n}$ be the matrix of coefficients of $q(x)$. Then $q(x) = x^T A x = x^T Q x$, where $Q = (A + A^T)/2$ is a symmetric matrix. Therefore any quadratic form can be associated with the symmetric matrix of its coefficients.
- A function $f : \mathbb{R}^n \to \mathbb{R}$ in the form $f(x) = \frac{1}{2} x^T Q x + c^T x$, where Q is a real $n \times n$ symmetric matrix and $c \in \mathbb{R}^n$ is the vector of linear coefficients, is called a quadratic function.

Positive definite forms and matrices

- A symmetric matrix Q is called positive definite (semidefinite), and associated with Q quadratic form $q(x) = x^T Q x$ is positive definite (semidefinite), if

$$q(x) = x^T Q x > 0 \ (\geq 0)$$

  for all nonzero $x \in \mathbb{R}^n$.
- It is known that a symmetric matrix A is positive definite (semidefinite) if and only if all the eigenvalues of A are positive (nonnegative).
- Another equivalent characterization of a positive definite matrix is positivity of all its leading principal minors [A k-th leading principal minor of a square matrix A is the determinant of a matrix that consists of the intersection of first k rows and first k columns of A].

Second order characterization of a convex function

Theorem

A twice continuously differentiable function $f(x)$ is convex on an open convex set X if and only if $\nabla^2 f(x)$ is positive semidefinite for any $x \in X$.