CO 355
Lecture 14

Fundamentals of Nonlinear Optimization

Vris Cheung

(Substitute class)

October 28, 2010

Outline

Extremum: unconstrained optimization

+ Given $f : S \subseteq \mathbb{R}^n \to \mathbb{R}$.
+ x^* is a (global) minimizer of f:

$$f(x^*) \leqslant f(x) \quad \text{for all } x \in S.$$

+ x^* is a *local* minimizer of f: for some nbd. N_{x^*} of x^*,

$$f(x^*) \leqslant f(x) \quad \text{for all } x \in N_{x^*} \cap S.$$

Extremum: unconstrained optimization

Necessary optimality conditions for local minimizer:

$$\begin{aligned} \text{1st order cond.:} \quad & \nabla f(x^*) = 0 \\ \text{2st order cond.:} \quad & \nabla^2 f(x^*) \succeq 0 \end{aligned}$$

Sufficient optimality condition for local minimizer:

$$\begin{aligned} & \text{1st order cond.:} \quad \nabla f(x^*) = 0 \\ \text{AND} \quad & \text{2st order cond.:} \quad \nabla^2 f(x^*) \succ 0 \end{aligned}$$

If $\nabla f(x) = 0$, x is said to be a saddle point / critical point.

Extremum: unconstrained optimization

When we try to solve

$$\min_{x \in S} f(x),$$

using only 1st order opt. cond.[1], we might end up

+ a global minimizer — great!
+ or... a local minimizer,
+ or... a saddle point.... (or maybe *nothing*...)

But if f is convex, we don't have such a problem:

Theorem

If $f : S \subseteq \mathbb{R}^n \to \mathbb{R}$ is convex (with S convex) and $\nabla f(x^*) = 0$, then x^* is a global minimizer of f.

[1] We usually don't have 2nd order information because it is too expensive to obtain in practice.

Extremum: convex case

Theorem

If $f : S \subseteq \mathbb{R}^n \to \mathbb{R}$ is convex (with S convex) and $\nabla f(x^*) = 0$, then x^* is a global minimizer of f.

Proof.

For any $x \in S$, by convexity,

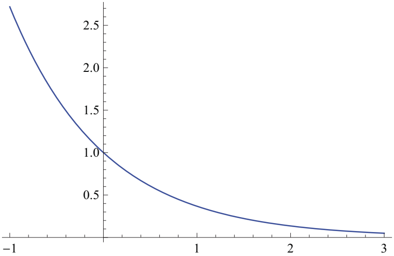
$$f(x) \geqslant f(x^*) + \nabla f(x^*)^\top (x - x^*) = f(x^*).$$

□

Existence of global minimizer

Sadly, even in convex case, it could happen that a minimizer simply does not exist: consider[2]

$$f(x) := e^{-x} \qquad (x \in \mathbb{R}).$$



$\inf_{x\in\mathbb{R}} e^{-x} = 0$ is finite, but not attained by any $x \in \mathbb{R}$.

[2] Graphics source: *Wolfram Alpha*.

Existence of global minimizer

So when would a minimizer exist?

An easy condition is *coerciveness*: $f : S \subseteq \mathbb{R}^n \to \mathbb{R}$ is coercive if [3]

$$\{x \in S : f(x) \leqslant \alpha\}$$

is bounded for all $\alpha \in \mathbb{R}$.

Lemma

Given $f : S \subseteq \mathbb{R}^n \to \mathbb{R}$, TFAE:

(1) $\{x : f(x) \leqslant \alpha\}$ is bounded for all $\alpha \in \mathbb{R}$.

(2) $f(x) \to +\infty$ as $\|x\| \to +\infty$.

[3] In course text, this set is called a *level set* of f. In literature it is also called *sublevel set* of f, to be distinguished from *level set*, which is of the form $\{x \in S : f(x) = \alpha\}$.

Existence of global minimizer

$f : S \subseteq \mathbb{R}^n \to \mathbb{R}$ is coercive if [3]

$$\{x \in S : f(x) \leqslant \alpha\}$$

is bounded for all $\alpha \in \mathbb{R}$.

Lemma

Given $f : S \subseteq \mathbb{R}^n \to \mathbb{R}$, TFAE:
(1) $\{x : f(x) \leqslant \alpha\}$ is bounded for all $\alpha \in \mathbb{R}$.
(2) $f(x) \to +\infty$ as $\|x\| \to +\infty$.

Examples of coercive functions:

+ $f(x) := x^\top A x - c^\top x$ for PD symm. $A \in \mathbb{R}^{n \times n}$ and $c \in \mathbb{R}^n$;
+ $f(x) := \|x\|^2 \quad (= x^\top I x)$.

[3] In course text, this set is called a *level set* of f. In literature it is also called *sublevel set* of f, to be distinguished from *level set*, which is of the form $\{x \in S : f(x) = \alpha\}$.

Existence of global minimizer

$f : S \subseteq \mathbb{R}^n \to \mathbb{R}$ is coercive if

$$\{x \in S : f(x) \leqslant \alpha\}$$

is bounded for all $\alpha \in \mathbb{R}$.

Theorem (3.7)

If $S \subseteq \mathbb{R}^n$ is nonempty closed and $f : S \to \mathbb{R}$ is continuous and coercive, then f has a minimizer.

Existence of global minimizer

$f : S \subseteq \mathbb{R}^n \to \mathbb{R}$ is coercive if

$$\{x \in S : f(x) \leqslant \alpha\}$$

is bounded for all $\alpha \in \mathbb{R}$.

Theorem (3.7)

If $S \subseteq \mathbb{R}^n$ is nonempty closed and $f : S \to \mathbb{R}$ is continuous and coercive, then f has a minimizer.

Corollary (3.8)

If $S \subseteq \mathbb{R}^n$ is nonempty closed, then any $z \in \mathbb{R}^n$ has a nearest point in S, i.e. $\exists x^* \in S$ s.t.

$$\|x^* - z\| \leqslant \|x - z\|.$$

If, moreover, S is convex, then z has exactly one nearest point in S.

Outline

Constrained optimization

Given $f, g_1, \ldots, g_p, h_1, \ldots, h_q : S \subseteq \mathbb{R}^n \to \mathbb{R}$,
a nonlinear program / constr. optimization problem is of the form
(from eqn. 3.7 of course text)

$$
\text{(NLP)} \qquad \min_{x \in \mathbb{R}^n} \quad f(x)
$$

$$
\begin{aligned}
\text{s.t.} \quad & g_i(x) \leqslant 0 \quad (i = 1, \ldots, p) \;, && \text{inequality constr.} \\
& h_j(x) = 0 \quad (j = 1, \ldots, q) \;, && \text{equality constr.} \\
& x \in S.
\end{aligned}
$$

It is a convention to write min instead of inf,
even though a minimizer may not exist.

Constrained optimization

Given $f, g_1, \ldots, g_p, h_1, \ldots, h_q : S \subseteq \mathbb{R}^n \to \mathbb{R}$,
a nonlinear program / constr. optimization problem is of the form

(from eqn. 3.7 of course text)

$$\text{(NLP)} \qquad \min_{x \in \mathbb{R}^n} \quad f(x)$$

$$\begin{aligned} \text{s.t.} \quad g_i(x) &\leqslant 0 \quad (i = 1, \ldots, p) , && \text{inequality constr.} \\ h_j(x) &= 0 \quad (j = 1, \ldots, q) , && \text{equality constr.} \\ x &\in S. \end{aligned}$$

- $x \in S$ is feasible for (NLP) if
  $g_i(x) \leqslant 0 \; \forall\, i = 1, \ldots, p$ and $h_j(x) = 0 \; \forall\, j = 1, \ldots, q$.
- $x^* \in S$ is optimal for / a minimizer of (NLP) if
  x^* is feasible and $f(x^*) \leqslant f(x)$ for all feasible $x \in S$.

Constrained optimization

Given $f, g_1, \ldots, g_p, h_1, \ldots, h_q : S \subseteq \mathbb{R}^n \to \mathbb{R}$,
a nonlinear program / constr. optimization problem is of the form
(from eqn. 3.7 of course text)

$$
\text{(NLP)} \qquad \min_{x \in \mathbb{R}^n} \quad f(x)
$$

$$
\begin{aligned}
\text{s.t.} \quad & g_i(x) \leqslant 0 \quad (i = 1, \ldots, p) \ , && \text{inequality constr.} \\
& h_j(x) = 0 \quad (j = 1, \ldots, q) \ , && \text{equality constr.} \\
& x \in S.
\end{aligned}
$$

The set of optimal solutions of (NLP) is denoted by

$$
\arg\min_x \{ f(x) : g_i(x) \leqslant 0 \ (i = 1, \ldots, p), h_j(x) = 0 \ (j = 1, \ldots, q) \} .
$$

Constrained optimization

$$
\text{(NLP)} \qquad \min_{x \in \mathbb{R}^n} \quad f(x)
$$
$$
\begin{aligned}
\text{s.t.} \quad & g_i(x) \leqslant 0 \quad (i = 1, \ldots, p) \ , \\
& h_j(x) = 0 \quad (j = 1, \ldots, q) \ , \\
& x \in S.
\end{aligned}
$$

Special classes of NLP:

+ linear program: all functions are affine, and $S = \mathbb{R}^n$.
+ convex program: $f, g_1, \ldots, g_p$ are convex, and $h_1, \ldots, h_q$ are affine.

Issues in constrained optimization

$$
\text{(NLP)} \qquad \min_{x \in \mathbb{R}^n} \quad f(x)
$$
$$
\begin{aligned}
\text{s.t.} \quad & g_i(x) \leqslant 0 \quad (i = 1, \ldots, p) \;, \\
& h_j(x) = 0 \quad (j = 1, \ldots, q) \;, \\
& x \in S.
\end{aligned}
$$

Four possibilities:

(1) (NLP) is infeasible.

(2) (NLP) is feasible...

(a) but unbounded (in objective value).

(b) and bounded (in objective value)...

(i) but does not attain optimality.

(ii) *and* attains optimality.

Issues in constrained optimization

$$
\text{(NLP)} \qquad \min_{x \in \mathbb{R}^n} \; f(x) \quad \text{s.t.} \quad g_i(x) \leqslant 0 \;\; (i = 1, \dots, p), \quad h_j(x) = 0 \;\; (j = 1, \dots, q), \quad x \in S.
$$

Four possibilities:

(1) (NLP) is infeasible.

(2) (NLP) is feasible...

(a) but unbounded (in objective value).

(b) and bounded (in objective value)...

(i) but does not attain optimality. $\Longleftarrow$ This is not possible in LP!

(ii) *and* attains optimality. (An LP is infeas., unbdd, or attains opt.)

Issues in constrained optimization

- + feasibility
- + boundedness of the opt. problem
- + existence of optimal solutions
- + uniqueness of optimal solutions
- + computability of an (approximate) optimal solution:
  — *how do we find an optimal solution?*
- + certificate of optimality
  — *how can we prove that what we find is optimal?*

Issues in constrained optimization

+ feasibility
+ boundedness of the opt. problem
+ existence of optimal solutions
+ uniqueness of optimal solutions
+ computability of an (approximate) optimal solution:
  — *how do we find an optimal solution?*
  Necessary opt. cond.
+ certificate of optimality
  — *how can we prove that what we find is optimal?*
  Sufficient opt. cond.

Outline

Sufficient optimality conditions

$$\text{(NLP)} \qquad \min_{x \in S} \quad f(x)$$
$$\text{s.t.} \quad g_i(x) \leqslant 0 \quad (i = 1, \ldots, p) \ ,$$
$$h_j(x) = 0 \quad (j = 1, \ldots, q) \ .$$

The Lagrangian of (NLP) is defined by the function

$$L(x, \lambda, \mu) := f(x) + \sum_{i=1}^{p} \lambda_i g_i(x) + \sum_{j=1}^{q} \mu_j h_j(x),$$

for $x \in \mathbb{R}^n$, $\lambda \in \mathbb{R}^p$ and $\mu \in \mathbb{R}^q$.

Sufficient optimality conditions

From Calculus class:

Theorem

Consider the equality constrained problem

$$\min_{x\in S} f(x) \quad s.t. \quad h_j(x) = 0 \ \ (j = 1,\dots,q) \ . \tag{1}$$

If, for some $\bar{\mu} \in \mathbb{R}^q$, $\bar{x} \in S$ *is feasible for* (1) *and solves*

$$\min_{x\in S} L(x,\bar{\mu}) := f(x) + \sum_{j=1}^{q} \bar{\mu}_j h_j(x),$$

then $\bar{x}$ *is optimal for* (1).

Sufficient optimality conditions

Theorem (full version FYI)

Consider the constrained optimization problem

$$\min_{x \in S} \quad f(x) \quad s.t. \quad g_i(x) \leqslant 0 \quad (i = 1, \ldots, p) \ , \qquad h_j(x) = 0 \quad (j = 1, \ldots, q) \ . \tag{2}$$

Suppose $\bar{x} \in S$. If
(a) $\bar{x}$ is feasible for (2),
(b) there exist non-negative $\bar{\lambda} \in \mathbb{R}^p$ and $\bar{\mu} \in \mathbb{R}^q$ s.t.
(i) $\bar{x} \in S$ solves

$$\min_{x \in S} L(x, \bar{\lambda}, \bar{\mu}) := f(x) + \sum_{i=1}^{p} \bar{\lambda}_i g_i(x) + \sum_{j=1}^{q} \bar{\mu}_j h_j(x),$$

(ii) complementary slackness condition holds:

$$\text{for each } i = 1, \ldots, p, \quad \bar{\lambda}_i g_i(\bar{x}) = 0,$$

then $\bar{x}$ is optimal for (2).

Sufficient optimality conditions

Theorem

Consider the inequality constrained problem

$$\min_{x \in S} \quad f(x) \quad s.t. \quad g_i(x) \leqslant 0 \quad (i = 1, \ldots, p) \;,. \tag{3}$$

Suppose $\bar{x} \in S$. If
(a) $\bar{x}$ is feasible for (3),
(b) there exists non-negative $\bar{\lambda} \in \mathbb{R}^p$ s.t.
(i) $\bar{x} \in S$ solves

$$\min_{x \in S} L(x, \bar{\lambda}) := f(x) + \sum_{i=1}^{p} \bar{\lambda}_i g_i(x),$$

(ii) complementary slackness condition holds:

$$\text{for each } i = 1, \ldots, p, \quad \bar{\lambda}_i g_i(\bar{x}) = 0,$$

then $\bar{x}$ is optimal for (3).

Sufficient optimality conditions: convex case

Recall: if a function ϕ is convex, then $\nabla\phi(\bar{x}) = 0$ implies $\bar{x}$ is a minimizer.

Theorem (3.11, convex programs)

Consider the equality constrained problem

$$\min_{x \in S} \quad f(x) \quad s.t. \quad g_i(x) \leqslant 0 \quad (i = 1, \ldots, p) \ , \tag{4}$$

where $f, g_1, \ldots, g_p$ are convex functions. Suppose $\bar{x} \in S$. If
(a) $\bar{x}$ is feasible for (4),
(b) there exists non-negative $\bar{\lambda} \in \mathbb{R}^p$ s.t.
(i) $\bar{x}$ satisfies

$$\nabla_x L(\bar{x}, \bar{\lambda}) = \nabla f(\bar{x}) + \sum_{i=1}^{p} \nabla \bar{\lambda}_i g_i(\bar{x}) = 0,$$

(ii) complementary slackness condition holds:

$$\text{for each } i = 1, \ldots, p, \quad \bar{\lambda}_i g_i(\bar{x}) = 0,$$

then $\bar{x}$ is optimal for (4).