Convex Optimization and Applications
4 - Convex Optimization Problems

Guillaume Sagnol
Outline

1. Definitions
2. Local vs. Global optima
3. Problem Reformulations
4. First-order Optimality Conditions
Definitions

Optimization problem (aka *nonlinear program*)

\[
\begin{align*}
\text{minimize} & \quad f_0(x) \\
\text{s.t.} & \quad f_i(x) \leq 0 \quad (\forall i \in [m]).
\end{align*}
\]  

**Definition**

- \( f_0 : \mathbb{R}^n \to \mathbb{R} \) is the *objective function* of (P)
Definitions

Optimization problem (aka *nonlinear program*)

\[
\text{minimize} \quad f_0(x) \\
\text{s.t.} \quad f_i(x) \leq 0 \quad (\forall i \in [m]).
\]

Definition

- \( f_0 : \mathbb{R}^n \rightarrow \mathbb{R} \) is the *objective function* of \((P)\)
- \( f_i(x) \leq 0 \) is a *constraint* of \((P)\)
Optimization problem (aka *nonlinear program*)

\[
\begin{align*}
\text{minimize} & \quad f_0(x) \\
\text{s.t.} & \quad f_i(x) \leq 0 \quad (\forall i \in [m]).
\end{align*}
\]

**Definition**

- \( f_0 : \mathbb{R}^n \to \mathbb{R} \) is the *objective function* of (P)
- \( f_i(x) \leq 0 \) is a *constraint* of (P)
- \( x \in \mathbb{R}^n \) is called *feasible* if \( f_i(x) \leq 0, \ \forall i \in [m] \)
Definitions

Optimization problem (aka *nonlinear program*)

\[
\begin{align*}
\text{minimize} & \quad f_0(x) \\ 
\text{s.t.} & \quad f_i(x) \leq 0 \quad (\forall i \in [m]).
\end{align*}
\]

**Definition**

- \(f_0 : \mathbb{R}^n \to \mathbb{R}\) is the **objective function** of (P)
- \(f_i(x) \leq 0\) is a **constraint** of (P)
- \(x \in \mathbb{R}^n\) is called **feasible** if \(f_i(x) \leq 0, \forall i \in [m]\)
- The **feasible set** of (P) is

\[
\mathcal{F} = \{x \in \mathbb{R}^n | f_1(x) \leq 0, \ldots, f_m(x) \leq 0\}.
\]
Optimization problem (aka *nonlinear program*)

\[
\begin{align*}
\text{minimize} \quad & f_0(x) \\
\text{s.t.} \quad & f_i(x) \leq 0 \quad (\forall i \in [m]).
\end{align*}
\]

**Definition**

- \( f_0 : \mathbb{R}^n \to \mathbb{R} \) is the *objective function* of \((P)\).
- \( f_i(x) \leq 0 \) is a *constraint* of \((P)\).
- \( x \in \mathbb{R}^n \) is called *feasible* if \( f_i(x) \leq 0, \forall i \in [m] \).
- The *feasible set* of \((P)\) is
  \[
  \mathcal{F} = \{ x \in \mathbb{R}^n \mid f_1(x) \leq 0, \ldots, f_m(x) \leq 0 \}.
  \]
- The *optimal value* of \((P)\) is
  \[
  p^* = \inf \{ f_0(x) \mid x \in \mathcal{F} \} \in \mathbb{R} \cup \{-\infty, +\infty\}.
  \]
Definitions

Optimization problem (aka *nonlinear program*)

\[
\begin{align*}
\text{minimize} & \quad f_0(x) \\
\text{s.t.} & \quad f_i(x) \leq 0 \quad (\forall i \in [m]).
\end{align*}
\]  

**Definition**

\[ p^* = +\infty \iff \mathcal{F} = \emptyset: \text{in that case, (P) is infeasible.} \]
## Definitions

**Optimization problem (aka nonlinear program)**

\[
\begin{align*}
\text{minimize} & \quad f_0(x) \\
\text{s.t.} & \quad f_i(x) \leq 0 \quad (\forall i \in [m]).
\end{align*}
\] (P)

### Definition

- **\( p^* = +\infty \iff F = \emptyset \):** in that case, (P) is **infeasible**.
- **\( p^* = -\infty \iff \exists (x_i)_{i \in \mathbb{N}} \text{ with } x_i \in F \text{ and } f_0(x_i) \to \infty \):**
  in this case, (P) is **unbounded from below**.
Definitions

Optimization problem (aka \textit{nonlinear program})

$$\begin{align*}
\text{minimize} & \quad f_0(x) \\
\text{s.t.} & \quad f_i(x) \leq 0 \quad (\forall i \in [m]).
\end{align*}$$

\textbf{Definition}

\begin{itemize}
\item $p^* = +\infty \iff \mathcal{F} = \emptyset$: in that case, (P) is \textit{infeasible}.
\item $p^* = -\infty \iff \exists (x_i)_{i \in \mathbb{N}}$ with $x_i \in \mathcal{F}$ and $f_0(x_i) \to \infty$; in this case, (P) is \textit{unbounded from below}.
\item When $f_0$ is constant, we just have to find any $x \in \mathcal{F}$; in this case (P) is a \textit{feasibility problem}.
\end{itemize}
Definitions

Optimization problem (aka *nonlinear program*)

\[
\begin{align*}
\text{minimize} & \quad f_0(x) \\
\text{s.t.} & \quad f_i(x) \leq 0 \quad (\forall i \in [m]).
\end{align*}
\]  

**Definition**

- \( p^* = +\infty \iff \mathcal{F} = \emptyset \): in that case, (P) is infeasible.
- \( p^* = -\infty \iff \exists (x_i)_{i \in \mathbb{N}} \text{ with } x_i \in \mathcal{F} \text{ and } f_0(x_i) \to \infty \); in this case, (P) is unbounded from below.
- When \( f_0 \) is constant, we just have to find any \( x \in \mathcal{F} \); In this case (P) is a feasibility problem.
- If \( x^* \in \mathcal{F} \) satisfies \( f_0(x^*) = p^* \), we say that \( x^* \) solves (P), or that \( x^* \) is a (global) optimal solution to (P).
Optimization problem (aka *nonlinear program*)

\[
\begin{align*}
\text{minimize} & \quad f_0(x) \\
\text{s.t.} & \quad f_i(x) \leq 0 \quad (\forall i \in [m]).
\end{align*}
\]

**Definition**

- \(x \in F\) is called \(\epsilon\)-suboptimal if \(f_0(x) \leq p^* + \epsilon\).
- The set of all \(\epsilon\)-suboptimal solutions is called the \(\epsilon\)-suboptimal set.
Definitions

Optimization problem (aka *nonlinear program*)

\[
\begin{align*}
\text{minimize} \quad & f_0(x) \\
\text{s.t.} \quad & f_i(x) \leq 0 \quad (\forall i \in [m]).
\end{align*}
\]

Definition

- \( x \in \mathcal{F} \) is called \( \epsilon \)-suboptimal if \( f_0(x) \leq p^* + \epsilon \).
- The set of all \( \epsilon \)-suboptimal solutions is called the \( \epsilon \)-suboptimal set.
- **Remark:** The constraints are understood in the sense of extended functions, i.e., \( x \notin \text{dom} f_i \implies f_i(x) = \infty \). Hence,
  \[ \mathcal{F} \subseteq \bigcap_{i=0}^{m} \text{dom} f_i. \]
The vector $x$ is called a \textit{local optimum} for Problem (P) if it solves the problem

\[\begin{align*}
\text{minimize} & \quad f_0(z) \\
\text{s.t.} & \quad f_i(z) \leq 0 \quad (\forall i \in [m]); \\
& \quad \|z - x\| \leq R
\end{align*}\]

for some $R > 0$. In other words, there is a neighbourhood of $x$ in $\mathcal{F}$ in which $f_0$ is minimized at $x$. 
Proposition (Differential characterization of local optima).

Assume the objective function $f_0$ is twice differentiable, and let $x \in \text{int} \ F$. Then, the following holds:

- If $\nabla f_0(x) = 0$ and $\nabla^2 f_0(x) \succ 0$, then $x$ is a local optimum.
- Conversely, if $x$ is a local optimum, then $\nabla f_0(x) = 0$ and $\nabla^2 f_0(x) \succeq 0$.

Remarks:

- Above proposition is valid for interior points only!
Proposition
(Differential characterization of local optima).

Assume the objective function $f_0$ is twice differentiable, and let $x \in \text{int } \mathcal{F}$. Then, the following holds:

- If $\nabla f_0(x) = 0$ and $\nabla^2 f_0(x) \succ 0$, then $x$ is a local optimum.
- Conversely, if $x$ is a local optimum, then $\nabla f_0(x) = 0$ and $\nabla^2 f_0(x) \succeq 0$.

Remarks:

- Above proposition is valid for interior points only!
- cannot replace $\succ$ by $\succeq$ in the 1st statement ($f(x) = x^3$)
- cannot replace $\succeq$ by $\succ$ in the 2d statement ($f(x) = x^4$)
Convex optimization problem

Optimization problem

\[
\begin{align*}
\text{minimize} & \quad f_0(x) \\
\text{s.t.} & \quad f_i(x) \leq 0 \quad (\forall i \in [m]).
\end{align*}
\]

Problem (P) is said to be convex if \( f_0, \ldots, f_m \) are convex.
Convex optimization problem

Optimization problem

\[
\begin{align*}
\text{minimize} & \quad f_0(x) \\
\text{s.t.} & \quad f_i(x) \leq 0 \quad (\forall i \in [m]).
\end{align*}
\]

Problem (P) is said to be convex if \( f_0, \ldots, f_m \) are convex.

Remark: We also say that the maximization problem

\[
\begin{align*}
\text{maximize} & \quad f_0(x) \\
\text{s.t.} & \quad f_i(x) \geq 0 \quad (\forall i \in [m]).
\end{align*}
\]

is convex if \( f_0, \ldots, f_m \) are concave.
Convex optimization problem

Optimization problem

\[
\begin{align*}
\text{minimize} & \quad f_0(x) \\
\text{s.t.} & \quad f_i(x) \leq 0 \quad (\forall i \in [m]).
\end{align*}
\]

Problem (P) is said to be convex if \( f_0, \ldots, f_m \) are convex.

Equality constraints \( f_i(x) = 0 \) can be handled as \( f_i(x) \leq 0, f_i(x) \geq 0 \). Hence, in a convex optimization problem, equality constraints must be linear.

It is often convenient to write equality constraints separately:

\[
\begin{align*}
\text{minimize} & \quad f_0(x) \\
\text{s.t.} & \quad f_i(x) \leq 0 \quad (\forall i \in [m]); \\
& \quad Ax = b.
\end{align*}
\]
Fundamental result

Theorem

Let (P) be a convex optimization problem. Then, any local optimum $x^*$ is also a global optimum.
Equivalence of Problems

“informal” definition

We say that Problems (P) and (Q) are equivalent, and we use the (nonstandard) notation $P \sim Q$, if there is a “simple transformation” which maps an optimal solution of $P$ to an optimal solution of $Q$, and vice-versa.
Change of variables

\[
\text{minimize} \quad f_0(x) \\
\text{s.t.} \quad f_i(x) \leq 0 \quad (\forall i \in [m]). 
\]

If \( \phi : \mathbb{R}^n \to \mathbb{R}^n \) is one-to-one, and every feasible \( x \) can be written as \( x = \phi(z) \) for some \( z \), then

\[
P \sim \text{minimize} \quad f_0(\phi(z)) \\
\text{s.t.} \quad f_i(\phi(z)) \leq 0 \quad (\forall i \in [m]).
\]
Transformation of objective or constraints

\[ \begin{align*}
\text{minimize} & \quad f_0(x) \\
\text{s.t.} & \quad f_i(x) \leq 0 \quad (\forall i \in [m]).
\end{align*} \] (P)

If

- \( \psi_0 : \mathbb{R} \to \mathbb{R} \) is strictly increasing
- \( \forall i \in [m], \psi_i : \mathbb{R} \to \mathbb{R} \) satisfies \( \psi_i(u) \leq 0 \iff u \leq 0 \),

then we have:

\[ \begin{align*}
P & \quad \sim \quad \text{minimize} \quad \psi_0(f_0(x)) \\
\text{s.t.} & \quad \psi_i(f_i(x)) \leq 0 \quad (\forall i \in [m]).
\end{align*} \]
Eliminating equality constraints

\[ \min_{x \in \mathbb{R}^n} f_0(x) \quad (P_{Eq}) \]
\[ \text{s.t.} \quad f_i(x) \leq 0 \quad (\forall i \in [m]) \]
\[ Ax = b. \]

The equality constraints define an affine subspace

\[ L = \{ x \in \mathbb{R}^n : Ax = b \} = \{ Cz + d : z \in \mathbb{R}^r \} \]
for some \( C \in \mathbb{R}^{n \times r}, d \in \mathbb{R}^n \). Hence,

\[ P_{Eq} \sim \minimize_{z \in \mathbb{R}^r} f_0(Cz + d) \]
\[ \text{s.t.} \quad f_i(Cz + d) \leq 0 \quad (\forall i \in [m]). \]
Slack variables

We can replace linear inequalities by linear equalities by introducing *slack variables*. For example,

\[
\begin{align*}
\text{minimize} \quad & f_0(x) \\
\text{s.t.} \quad & f_i(x) \leq 0 \quad (\forall i \in [m]) \\
& Ax \leq b,
\end{align*}
\]

where \( A \in \mathbb{R}^{p \times n} \), is equivalent to

\[
\begin{align*}
\text{minimize} \quad & f_0(x) \\
\text{s.t.} \quad & f_i(x) \leq 0 \quad (\forall i \in [m]) \\
& Ax + s = b \\
& s \geq 0.
\end{align*}
\]
Epigraph formulation

\[
\begin{align*}
\text{minimize} & \quad f_0(x) \\
\text{s.t.} & \quad f_i(x) \leq 0 \quad (\forall i \in [m]).
\end{align*}
\]

We can always assume that the objective function is a linear form: \( f_0(x) = c^T x \), with \( \|c\| = 1 \), by passing to the epigraph formulation:

\[
P \sim \minimize_{x \in \mathbb{R}^n, t \in \mathbb{R}} t \\
\text{s.t.} & \quad f_i(x) \leq 0 \quad (\forall i \in [m]) \\
& \quad f_0(x) \leq t.
\]
Partial minimization

It is possible to reduce a problem by solving it (partially) for some blocks of variables. For example, the following two problems are equivalent:

\[
\begin{align*}
\text{minimize} & \quad f_0(x_1, x_2) \\
\text{s.t.} & \quad f_i(x_1) \leq 0 \quad (\forall i \in [m_1]) \\
& \quad g_j(x_1, x_2) \leq 0 \quad (\forall j \in [m_2])
\end{align*}
\]

\[
\sim \begin{align*}
\text{minimize} & \quad \tilde{f}_0(x_1) \\
\text{s.t.} & \quad f_i(x_1) \leq 0 \quad (\forall i \in [m_1]),
\end{align*}
\]

where we have defined

\[
\tilde{f}_0(x_1) := \inf \left\{ f_0(x_1, x_2) \mid x_2 \in \mathbb{R}^{n_2}, g_j(x_1, x_2) \leq 0, \forall j \in [m_2] \right\}.
\]
First-order optimality condition

This result gives a simple optimality condition, which depends only on \( \nabla f_0 \) and the feasibility set

\[
\mathcal{F} = \{ x \in \mathbb{R}^n | f_1(x) \leq 0, \ldots, f_m(x) \leq 0 \}.
\]

**Theorem**

Let \( f_0 \) be differentiable. Then, a vector \( x^* \) is optimal for the convex problem \((P)\) if and only if \( x^* \) is feasible (i.e., \( x^* \in \mathcal{F} \)), and

\[
\forall y \in \mathcal{F}, \quad \nabla f_0(x^*)^T(y - x^*) \geq 0.
\]

Geometrically, this means that either \( \nabla f_0(x^*) = 0 \), or \( \nabla f_0(x^*) \) defines a supporting hyperplane to \( \mathcal{F} \) at \( x^* \).
Equality-constrained optimization

We can use the previous theorem to characterize optimal solutions of equality-constrained convex programs:

**Proposition**

Consider the optimization problem

\[
\begin{align*}
\text{minimize} & \quad f_0(x) \\
\text{s.t.} & \quad Ax = b,
\end{align*}
\]

where \( f_0 \) is convex and differentiable. Then, \( x^* \) is optimal iff \( \nabla f_0(x^*) \in \text{Im} \ A^T \).
Optimization over $\mathbb{R}^n_+$

We can also obtain a characterization of optimal solutions of convex programs over the nonnegative orthant:

**Proposition**

Consider the optimization problem

$$
\begin{align*}
\text{minimize} \quad & f_0(x) \\
\text{s.t.} \quad & x \geq 0,
\end{align*}
$$

where $f_0$ is convex and differentiable. Then, $x^*$ is optimal iff

- $x^* \geq 0$
- $\nabla f_0(x^*) \geq 0$
- $\forall i \in [n], \ (x_i = 0 \text{ or } \frac{\partial f_0}{\partial x_i}(x^*) = 0)$