Convex Optimization and Applications

4 - Convex Optimization Problems

Guillaume Sagnol

Technische Universität Berlin

Outline

Definitions

Optimization problem (aka *nonlinear program*)

$$\begin{aligned} \underset{\boldsymbol{x}\in\mathbb{R}^n}{\textbf{minimize}} \quad & f_0(x) \\ s.t. \quad & f_i(x) \leq 0 \qquad (\forall i \in [m]). \end{aligned} \tag{P}$$

Definition

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

Definitions

Optimization problem (aka *nonlinear program*)

$$\begin{aligned} \underset{x\in\mathbb{R}^n}{\textbf{minimize}} \quad & f_0(x) \\ s.t. \quad & f_i(x) \leq 0 \qquad (\forall i \in [m]). \end{aligned} \tag{P}$$

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)

Definitions

Optimization problem (aka *nonlinear program*)

$$\begin{aligned} \underset{\boldsymbol{x}\in\mathbb{R}^n}{\textbf{minimize}} \quad & f_0(\boldsymbol{x}) \\ s.t. \quad & f_i(\boldsymbol{x}) \leq 0 \qquad (\forall i \in [m]). \end{aligned} \tag{P}$$

Definition

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

Definitions

Optimization problem (aka *nonlinear program*)

$$\begin{aligned} \underset{\boldsymbol{x}\in\mathbb{R}^n}{\textbf{minimize}} \quad & f_0(\boldsymbol{x}) \\ s.t. \quad & f_i(\boldsymbol{x}) \leq 0 \qquad (\forall i \in [m]). \end{aligned} \tag{P}$$

Definition

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

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

Definitions

Optimization problem (aka *nonlinear program*)

$$\begin{aligned} \underset{x\in\mathbb{R}^n}{\textbf{minimize}} \quad & f_0(x) & & \quad \text{(P)} \\ s.t. \quad & f_i(x) \leq 0 & (\forall i \in [m]). \end{aligned}$$

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 $\quad f_i(x) \leq 0, \ \forall i \in [m]$
- The *feasible set* of (P) is

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

- The *optimal value* of (P) is

$$p^* = \inf\{f_0(x) |\ x \in \mathcal{F}\} \in \mathbb{R} \cup \{-\infty, +\infty\}.$$

Definitions

Optimization problem (aka *nonlinear program*)

$$\begin{aligned} \underset{\boldsymbol{x}\in\mathbb{R}^n}{\textbf{minimize}} \quad & f_0(\boldsymbol{x}) \\ s.t. \quad & f_i(\boldsymbol{x}) \leq 0 \qquad (\forall i \in [m]). \end{aligned} \tag{P}$$

Definition

- $p^* = +\infty \iff \mathcal{F} = \emptyset$: in that case, (P) is *infeasible*.

Definitions

Optimization problem (aka *nonlinear program*)

$$\begin{aligned} \underset{\boldsymbol{x}\in\mathbb{R}^n}{\textbf{minimize}} \quad & f_0(\boldsymbol{x}) \\ s.t. \quad & f_i(\boldsymbol{x}) \leq 0 \qquad (\forall i \in [m]). \end{aligned} \tag{P}$$

Definition

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

Definitions

Optimization problem (aka *nonlinear program*)

$$\begin{aligned} \underset{x\in\mathbb{R}^n}{\textbf{minimize}} \quad & f_0(x) \\ s.t. \quad & f_i(x) \leq 0 \qquad (\forall i \in [m]). \end{aligned} \tag{P}$$

Definition

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

Definitions

Optimization problem (aka *nonlinear program*)

$$\begin{aligned} \underset{\boldsymbol{x}\in\mathbb{R}^n}{\textbf{minimize}} \quad & f_0(\boldsymbol{x}) \qquad \text{(P)} \\ s.t. \quad & f_i(\boldsymbol{x}) \leq 0 \qquad (\forall i \in [m]). \end{aligned}$$

Definition

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

Definitions

Optimization problem (aka *nonlinear program*)

$$\begin{aligned} \mathop{\textbf{minimize}}_{\boldsymbol{x}\in\mathbb{R}^n} \quad & f_0(\boldsymbol{x}) \\ s.t. \quad & f_i(\boldsymbol{x}) \leq 0 \qquad (\forall i \in [m]). \end{aligned} \tag{P}$$

Definition

- $\boldsymbol{x} \in \mathcal{F}$ is called ϵ-*suboptimal* if $f_0(\boldsymbol{x}) \leq p^* + \epsilon$.
- The set of all ϵ-suboptimal solutions is called the ϵ-*suboptimal set*.

Definitions

Optimization problem (aka *nonlinear program*)

$$\begin{aligned} \underset{\boldsymbol{x}\in\mathbb{R}^n}{\textbf{minimize}} \quad & f_0(\boldsymbol{x}) \\ s.t. \quad & f_i(\boldsymbol{x}) \leq 0 \qquad (\forall i \in [m]). \end{aligned} \tag{P}$$

Definition

- $\boldsymbol{x} \in \mathcal{F}$ is called *ϵ-suboptimal* if $f_0(\boldsymbol{x}) \leq p^* + \epsilon$.
- The set of all ϵ-suboptimal solutions is called the *ϵ-suboptimal set*.
- **Remark:** The constraints are understood in the sense of extended functions, i.e., $\boldsymbol{x} \notin \textbf{dom}\, f_i \implies f_i(\boldsymbol{x}) = \infty$. Hence,

$$\mathcal{F} \subseteq \bigcap_{i=0}^{m} \textbf{dom}\, f_i.$$

Local optimum

Definition (Local optimum).

The vector $\boldsymbol{x}$ is called a *local optimum* for Problem (P) if it solves the problem

$$\begin{aligned} \underset{\boldsymbol{z}\in\mathbb{R}^n}{\textbf{minimize}} \quad & f_0(\boldsymbol{z}) \\ s.t. \quad & f_i(\boldsymbol{z}) \leq 0 \qquad (\forall i \in [m]); \\ & \|\boldsymbol{z}-\boldsymbol{x}\| \leq R \end{aligned} \tag{P_R}$$

for some $R > 0$. In other words, there is a neighbourhood of $\boldsymbol{x}$ in $\mathcal{F}$ in which f_0 is minimized at $\boldsymbol{x}$.

Local optimum

Proposition (Differential characterization of local optima).

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

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

Remarks:

- Above proposition is valid **for interior points only** !

Local optimum

Proposition (Differential characterization of local optima).

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

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

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

$$\underset{\boldsymbol{x}\in\mathbb{R}^n}{\textbf{minimize}} \quad f_0(\boldsymbol{x}) \qquad \text{(P)}$$
$$s.t. \quad f_i(\boldsymbol{x}) \leq 0 \qquad (\forall i \in [m]).$$

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

Convex optimization problem

Optimization problem

$$\begin{aligned} \underset{\boldsymbol{x}\in\mathbb{R}^n}{\textbf{minimize}} \quad & f_0(\boldsymbol{x}) \\ s.t. \quad & f_i(\boldsymbol{x}) \leq 0 \qquad (\forall i \in [m]). \end{aligned} \tag{P}$$

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

Remark: We also say that the maximization problem

$$\begin{aligned} \underset{\boldsymbol{x}\in\mathbb{R}^n}{\textbf{maximize}} \quad & f_0(\boldsymbol{x}) \\ s.t. \quad & f_i(\boldsymbol{x}) \geq 0 \qquad (\forall i \in [m]). \end{aligned}$$

is convex if $f_0, \ldots, f_m$ are concave.

Convex optimization problem

Optimization problem

$$\underset{\boldsymbol{x}\in\mathbb{R}^n}{\textbf{minimize}} \quad f_0(\boldsymbol{x}) \qquad \text{(P)}$$
$$s.t. \quad f_i(\boldsymbol{x}) \leq 0 \qquad (\forall i \in [m]).$$

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

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

It is often convenient to write equality constraints separately:

$$\underset{\boldsymbol{x}\in\mathbb{R}^n}{\textbf{minimize}} \quad f_0(\boldsymbol{x}) \qquad (\mathrm{P}_{\mathrm{Eq}})$$
$$s.t. \quad f_i(\boldsymbol{x}) \leq 0 \qquad (\forall i \in [m]);$$
$$A\boldsymbol{x} = \boldsymbol{b}.$$

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 $P \dot{\sim} Q$, 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

$$\begin{aligned}\underset{\boldsymbol{x}\in\mathbb{R}^n}{\textbf{minimize}} \quad & f_0(x) \\ s.t. \quad & f_i(x) \leq 0 \qquad (\forall i \in [m]).\end{aligned} \tag{P}$$

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

$$\text{P} \quad \dot{\sim} \quad \begin{aligned}\underset{\boldsymbol{z}\in\mathbb{R}^n}{\textbf{minimize}} \quad & f_0(\phi(z)) \\ s.t. \quad & f_i(\phi(z)) \leq 0 \qquad (\forall i \in [m]).\end{aligned}$$

Transformation of objective or constraints

$$\begin{aligned} \underset{\boldsymbol{x}\in\mathbb{R}^n}{\textbf{minimize}} \quad & f_0(\boldsymbol{x}) && \text{(P)} \\ s.t. \quad & f_i(\boldsymbol{x}) \leq 0 \qquad (\forall i \in [m]). \end{aligned}$$

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:

$$\text{P} \quad \dot{\sim} \quad \begin{aligned} \underset{\boldsymbol{x}\in\mathbb{R}^n}{\textbf{minimize}} \quad & \psi_0(f_0(\boldsymbol{x})) \\ s.t. \quad & \psi_i(f_i(\boldsymbol{x})) \leq 0 \qquad (\forall i \in [m]). \end{aligned}$$

Eliminating equality constraints

$$\begin{aligned}
\underset{\boldsymbol{x}\in\mathbb{R}^n}{\textbf{minimize}} \quad & f_0(\boldsymbol{x}) && \text{(P}_{\mathrm{Eq}}\text{)}\\
s.t. \quad & f_i(\boldsymbol{x}) \leq 0 \qquad (\forall i \in [m])\\
& A\boldsymbol{x} = \boldsymbol{b}.
\end{aligned}$$

The equality constraints define an affine subspace $L = \{\boldsymbol{x} \in \mathbb{R}^n : A\boldsymbol{x} = \boldsymbol{b}\} = \{C\boldsymbol{z} + \boldsymbol{d} : \boldsymbol{z} \in \mathbb{R}^r\}$ for some $C \in \mathbb{R}^{n\times r}, \boldsymbol{d} \in \mathbb{R}^n$. Hence,

$$\begin{aligned}
\mathrm{P}_{\mathrm{Eq}} \quad \dot{\sim} \quad \underset{\boldsymbol{z}\in\mathbb{R}^r}{\textbf{minimize}} \quad & f_0(C\boldsymbol{z} + \boldsymbol{d})\\
s.t. \quad & f_i(C\boldsymbol{z} + \boldsymbol{d}) \leq 0 \qquad (\forall i \in [m]).
\end{aligned}$$

Slack variables

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

$$\begin{aligned} \underset{\boldsymbol{x}\in\mathbb{R}^n}{\textbf{minimize}} \quad & f_0(\boldsymbol{x}) \\ s.t. \quad & f_i(\boldsymbol{x}) \leq 0 \qquad (\forall i \in [m]) \\ & A\boldsymbol{x} \leq \boldsymbol{b}, \end{aligned}$$

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

$$\begin{aligned} \underset{\boldsymbol{x}\in\mathbb{R}^n, \boldsymbol{s}\in\mathbb{R}^p}{\textbf{minimize}} \quad & f_0(\boldsymbol{x}) \\ s.t. \quad & f_i(\boldsymbol{x}) \leq 0 \qquad (\forall i \in [m]) \\ & A\boldsymbol{x} + \boldsymbol{s} = \boldsymbol{b} \\ & \boldsymbol{s} \geq \boldsymbol{0}. \end{aligned}$$

Epigraph formulation

$$\begin{aligned} \underset{\boldsymbol{x}\in\mathbb{R}^n}{\textbf{minimize}} \quad & f_0(\boldsymbol{x}) \\ s.t. \quad & f_i(\boldsymbol{x}) \leq 0 \qquad (\forall i \in [m]). \end{aligned} \tag{P}$$

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

$$\begin{aligned} \text{P} \quad \dot{\sim} \quad \underset{\boldsymbol{x}\in\mathbb{R}^n, t\in\mathbb{R}}{\textbf{minimize}} \quad & t \\ s.t. \quad & f_i(\boldsymbol{x}) \leq 0 \qquad (\forall i \in [m]) \\ & f_0(\boldsymbol{x}) \leq t. \end{aligned}$$

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:

$$\begin{aligned} \underset{\boldsymbol{x}_1 \in \mathbb{R}^{n_1}, \boldsymbol{x}_2 \in \mathbb{R}^{n_2}}{\textbf{minimize}} \quad & f_0(\boldsymbol{x}_1, \boldsymbol{x}_2) \\ s.t. \quad & f_i(\boldsymbol{x}_1) \leq 0 \quad (\forall i \in [m_1]) \\ & g_j(\boldsymbol{x}_1, \boldsymbol{x}_2) \leq 0 \quad (\forall j \in [m_2]) \end{aligned}$$

$$\dot{\sim} \quad \begin{aligned} \underset{\boldsymbol{x}_1 \in \mathbb{R}^{n_1}}{\textbf{minimize}} \quad & \tilde{f}_0(\boldsymbol{x}_1) \\ s.t. \quad & f_i(\boldsymbol{x}_1) \leq 0 \quad (\forall i \in [m_1]), \end{aligned}$$

where we have defined

$$\tilde{f}_0(\boldsymbol{x}_1) := \inf \left\{ f_0(\boldsymbol{x}_1, \boldsymbol{x}_2) \mid \boldsymbol{x}_2 \in \mathbb{R}^{n_2}, g_j(\boldsymbol{x}_1, \boldsymbol{x}_2) \leq 0, \forall j \in [m_2] \right\}.$$

First-order optimality condition

This result gives a simple optimality condition, which depends only on ∇f_0 and the feasibility set

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

Theorem

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

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

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

Equality-constrained optimization

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

Proposition

Consider the optimization problem

$$\begin{aligned} \underset{\boldsymbol{x}\in\mathbb{R}^n}{\textbf{minimize}} \quad & f_0(\boldsymbol{x}) \\ s.t. \quad & A\boldsymbol{x} = \boldsymbol{b}, \end{aligned}$$

where f_0 is convex and differentiable. Then, $\boldsymbol{x}^*$ is optimal iff $\nabla f_0(\boldsymbol{x}^*) \in \textbf{Im}\, A^T$.

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

$$\begin{aligned} \underset{\boldsymbol{x}\in\mathbb{R}^n}{\textbf{minimize}} \quad & f_0(\boldsymbol{x}) \\ s.t. \quad & \boldsymbol{x} \geq \mathbf{0}, \end{aligned}$$

where f_0 is convex and differentiable. Then, $\boldsymbol{x}^*$ is optimal iff

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