ELEG5481

SIGNAL PROCESSING OPTIMIZATION TECHNIQUES

4. CONVEX OPTIMIZATION PROBLEMS

Extended-Valued Extensions on Functions

- It is often convenient to extend a function domain to all of $\mathbf{R}^n$.
- Extended-valued extension of f:

$$\tilde{f}(x) = \begin{cases} f(x), & x \in \mathbf{dom} f \\ +\infty, & x \notin \mathbf{dom} f \end{cases}$$

- If f is convex then $\tilde{f}$ is also convex.

Optimization Problems in a Standard Form

$$\begin{aligned} \min\ & f_0(x) \\ \text{s.t. } & f_i(x) \leq 0, \quad i = 1, \ldots, m \\ & h_i(x) = 0, \quad i = 1, \ldots, p \end{aligned}$$

where

f_0 is the **objective function**;

$f_i, i = 1, \ldots, m$ are **inequality constraint functions**;

h_i are **equality constraint functions**.

Some Terminology

- The set

$$\mathcal{D} = \bigcap_{i=0}^{m} \textbf{dom} f_i \cap \bigcap_{i=1}^{p} \textbf{dom} h_i$$

  is the problem domain.

- The set

$$C = \{x | f_i(x) \leq 0, \; i = 1, \ldots, m, \; h_i(x) = 0, \; i = 1, \ldots, p\}$$

  is called the **feasible set**, or the **constraint set**.

- A point x is **feasible** if $x \in C$, and **infeasible** otherwise.
- The inequality constraint $f_i(x)$ is **active** at $x \in C$ if $f_i(x) = 0$.
- A point x is **strictly feasible** if

$$f_i(x) < 0, \ i = 1, \ldots, m, \quad h_i(x) = 0, \ i = 1, \ldots, p$$

  i.e., all inequality constraints are inactive.

- A problem is

  feasible if there exists $x \in C$;

  infeasible if $C = \emptyset$;

  strictly feasible if there is a strictly feasible point.

- A problem is **unconstrained** if $C = \mathbf{R}^n$.

- **Optimal value**:

$$p^\star = \inf_{x \in C} f_0(x)$$

If the problem is infeasible, we choose $p^\star = +\infty$. If $p^\star = -\infty$, the problem is **unbounded below.**

- A point $x^\star$ is **globally optimal** or simply **optimal** if $x^\star \in C$ & $f_0(x^\star) = p^\star$.

- The problem is **solvable** if an opt. point exists.

- A point x is **locally optimal** if there is an $R > 0$ such that

$$f_0(x) = \inf\{ f_0(\tilde{x}) \mid \tilde{x} \in C, \ \|\tilde{x} - x\|_2 \leq R \}$$

Related Problems

- Maximizing an objective function over a constraint set is equivalent to a minimization problem:

$$\max_{x \in C} f_0(x) = -\min_{x \in C} -f_0(x)$$

- The feasibility problem

$$\begin{aligned} &\text{find } x \\ &\text{s.t. } x \in C \end{aligned}$$

is an opt. problem where $f_0(x) = 0$. In this case $p^\star = 0$ if $C \neq \emptyset$, and $p^\star = +\infty$ if $C = \emptyset$.

Convex Optimization Problem

A standard problem

$$\begin{aligned} \min\ & f_0(x) \\ \text{s.t.}\ & f_i(x) \le 0, \quad i = 1, \ldots, m \\ & h_i(x) = 0, \quad i = 1, \ldots, p \end{aligned}$$

is **convex** if $f_0, \ldots, f_m$ are convex and $h_1, \ldots, h_p$ are affine.

A standard convex problem is often written as

$$\begin{aligned} \min\ & f_0(x) \\ \text{s.t.}\ & Ax = b,\ f_i(x) \le 0, \quad i = 1, \ldots, m \end{aligned}$$

where $A \in \mathbf{R}^{p \times n}$, $b \in \mathbf{R}^p$.

Convex Opt. Examples:

- Least squares (LS):

$$\min \ \|Ax - b\|_2^2$$

- A problem related to LS is the unconstrained quadratic program (QP):

$$\min \ x^T P x + 2q^T x + r$$

  which is convex iff $P \succeq 0$.

  LS is an unconstrained QP with $P = A^T A \succeq 0$.

  (Question: what happens when P is indefinite?)

Convex Opt. Examples (cont'd):

- Linear program (LP):

$$\begin{aligned} \min\ & c^T x \\ \text{s.t. } & x \succeq 0 \\ & Ax = b \end{aligned}$$

- Minimum norm approximation with bound constraints:

$$\begin{aligned} \min\ & \|Ax - b\| \\ \text{s.t. } & \ell_i \leq x_i \leq u_i, \quad i = 1, \ldots, m \end{aligned}$$

for any norm $\|.\|$.

Local and Global Optimality

For convex opt. problems, any locally optimal solution is globally optimal.

Let $x^\star$ be globally opt., & x be locally opt. with $f_0(x^\star) < f_0(x)$. Let $z = (1-\theta)x + \theta x^\star$:

$$f_0(z) \leq (1-\theta) f_0(x) + \theta f_0(x^\star) < f_0(x), \ \forall \theta \in [0,1] \qquad (*)$$

Since C is convex, $z \in C \ \forall \ \theta \in [0,1]$. For small enough θ, z satisfies $\|z - x\|_2 < R$ & thus

$$f_0(z) \geq f_0(x)$$

This contradicts with $(*)$.

An Optimality Criterion

Assume that f_0 is differentiable, & that the problem is convex.

A point $x \in C$ is optimal iff

$$\nabla f_0(x)^T(y - x) \geq 0, \quad \forall y \in C \qquad (*)$$

The sufficiency of this optimality criterion is straightforward.

From the 1st order condition,

$$f_0(y) \geq f_0(x) + \nabla f_0(x)^T(y - x), \ \forall y \in C$$

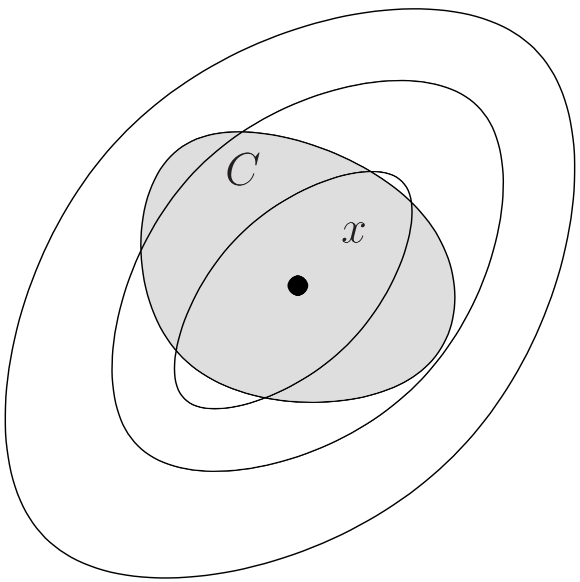
If x satisfies $(*)$ then $f_0(y) \geq f_0(x)$ $\forall y \in C$.

Interpretations of the optimality criterion:

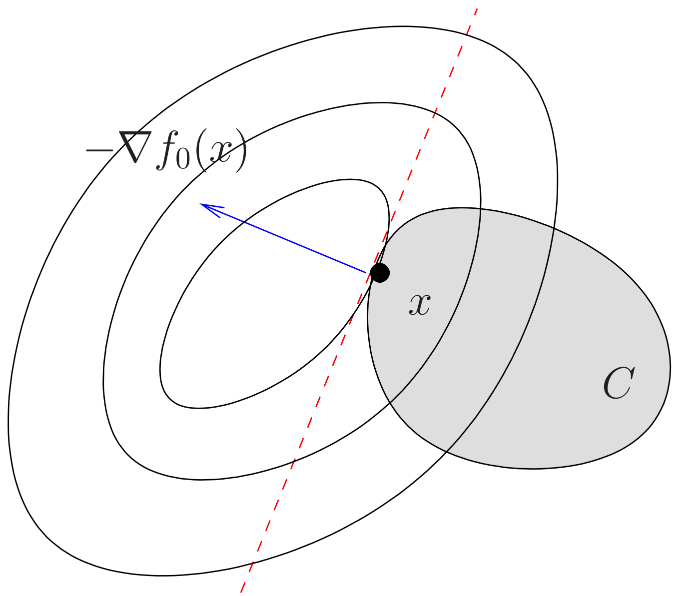
$$\nabla f_0(x)^T(y-x) \geq 0, \quad \forall y \in C$$

implies that an optimal $x \in C$ either achieves $\nabla f_0(x) = 0$,

or has $-\nabla f_0(x)$ forming a supporting hyperplane for C.



Case 1: $\nabla f_0(x) = 0$



Case 2: $\nabla f_0(x)^T(y-x) \geq 0, \ \forall y \in C$

For an unconstrained convex problem, $x \in C$ is optimal iff

$$\nabla f_0(x) = 0.$$

Example: The unconstrained QP:

$$\min \; x^T P x + 2q^T x + r$$

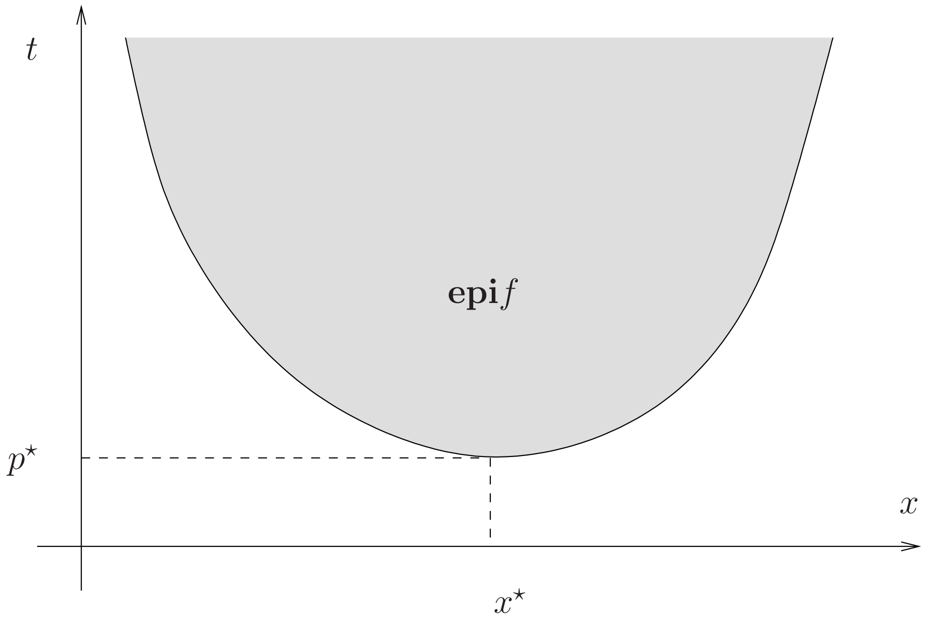
The optimality criterion of LS is

$$\nabla f_0(x) = 2Px + 2q$$

If P is invertible (or $P \succ 0$), then $x = -P^{-1}q$ is the optimal solution.

Equivalent Problem: Epigraph Form

$$
\begin{aligned}
\min\ & t \\
\text{s.t.}\ & f_0(x) - t \leq 0 \\
& f_i(x) \leq 0, \quad i = 1, \ldots, m \\
& h_i(x) = 0, \quad i = 1, \ldots, p
\end{aligned}
$$



- Now we minimize the problem with respect to x & t.
- $f_0(x) - t$ is convex in x and t, and hence the problem is still convex.

Example: Piecewise linear objective

$$\min_{x \in C} \max_{i=1,\ldots,m} \{a_i^T x + b_i\} \Longleftrightarrow \begin{array}{ll} \displaystyle\min_{x\in C,\ t\in\mathbf{R}} & t \\ \text{s.t.} & \displaystyle\max_{i=1,\ldots,m} \{a_i^T x + b_i\} \leq t, \end{array}$$

$$\Longleftrightarrow \begin{array}{ll} \displaystyle\min_{x\in C,\ t\in\mathbf{R}} & t \\ \text{s.t.} & a_i^T x + b_i \leq t,\ i = 1, \ldots, m \end{array}$$

Equiv. Problem: Equality Constraint Elimination

Let $A^\dagger$ be the pseudo-inverse of A, $\mathcal{N}(A)$ be the nullspace of A, & $\mathcal{R}(A)$ be the range space of A.

$$\begin{aligned} Ax = b \Longleftrightarrow x = A^\dagger b + \nu, \quad \nu \in \mathcal{N}(A) \\ \Longleftrightarrow x = A^\dagger b + Fz, \quad z \in \mathbb{R}^d \end{aligned}$$

for some $F \in \mathbb{R}^{n\times d}$ such that $\mathcal{R}(F) = \mathcal{N}(A)$, $d = \dim(\mathcal{N}(A))$.

A standard convex problem can be rewritten as

$$\begin{aligned} \min_z \ & f_0(A^\dagger b + Fz) \\ \text{s.t.} \ & f_i(A^\dagger b + Fz) \le 0, \quad i = 1, \ldots, m \end{aligned}$$

Example:

$$\begin{aligned} &\min\ x^T R x \\ &\text{s.t. } Ax = b \end{aligned}$$

can be turned to

$$\min_z\ (x_0 - Fz)^T R (x_0 - Fz).$$

where $x_0 = A^\dagger b$.

Suppose that $F^T R F \succ 0$. The optimal solution is

$$z^\star = (F^T R F)^{-1} F^T R x_0.$$

Equivalent Problem: Function Transformations

Suppose $\psi_0 : \mathbf{R} \to \mathbf{R}$ is monotone increasing, & $\psi_i : \mathbf{R} \to \mathbf{R}$, $i = 1, \ldots, m$ satisfy $\psi_i(u) \leq 0$ iff $u \leq 0$.

A standard problem is equivalent to

$$\begin{aligned} \min\ & \psi_0(f_0(x)) \\ \text{s.t.}\ & \psi_i(f_i(x)) \leq 0, \quad i = 1, \ldots, m \\ & h_i(x) = 0, \quad i = 1, \ldots, p \end{aligned}$$

Examples:

- The inequality constraint

$$\log x \leq 1$$

has a nonconvex inequality function. But

$$\log x \leq 1 \Longleftrightarrow x \leq e^1$$

- Consider

$$1 - x_1 x_2 \leq 0, \quad x_1 \geq 0, \quad x_2 \geq 0$$

$1 - x_1 x_2$ is non-convex. But we can replace it by

$$-\log x_1 - \log x_2 \leq 0, \quad x_1 \geq 0, \quad x_2 \geq 0$$

Examples (Cont'd):

- The least 2-norm problem

$$\min \|Ax - b\|_2 \tag{$*$}$$

  is equiv. to the LS problem $\min \|Ax - b\|_2^2$.

  In fact we want to avoid ($*$), which exhibits less desirable differentiation properties.

- Suppose that $P \succeq 0$. The problem

$$\max_{x \in C} \frac{1}{x^T P x}$$

  is equivalent to

$$\min_{x \in C} x^T P x$$

Example: maximum-likelihood (ML) estimation of mean

Likelihood function for m Gaussian samples with mean μ and covariance Σ:

$$L(\mu, \Sigma) = \frac{1}{(2\pi)^{\frac{nm}{2}} (\det \Sigma)^{\frac{m}{2}}} \exp \left\{ -\frac{1}{2} \sum_{i=1}^{m} (x_i - \mu)^T \Sigma^{-1} (x_i - \mu) \right\}$$

Given Σ, the ML estimation of μ is

$$\max_{\mu} L(\mu, \Sigma)$$

We often put logarithm on $L(\mu, \Sigma)$ to obtain an equiv. but convex problem

$$\max_{\mu} \log L(\mu, \Sigma) \propto \max_{\mu} - \textstyle\sum_{i=1}^{m} (x_i - \mu)^T \Sigma^{-1} (x_i - \mu)$$

Equivalent Problem: Change of Variables

Suppose $\phi : \mathbf{R}^n \to \mathbf{R}^n$ is one-to-one with image covering the problem domain.

Define

$$\tilde{f}_i(z) = f_i(\phi(z)), i = 0, \ldots, m, \quad \tilde{h}_i(z) = h_i(\phi(z)), i = 1, \ldots, p$$

A standard problem is equivalent to

$$\begin{aligned} \min \; & \tilde{f}_0(z) \\ \text{s.t.} \; & \tilde{f}_i(z) \le 0, \quad i = 1, \ldots, m \\ & \tilde{h}_i(z) = 0, \quad i = 1, \ldots, p \end{aligned}$$

Example: ML estimation of covariance

Consider the log likelihood in the last example, which can be expressed as

$$\log L(\mu, \Sigma) \propto -\tfrac{m}{2} \log\det\Sigma - \tfrac{m}{2}\mathbf{tr}(\hat{C}_\mu \Sigma^{-1})$$

where $\hat{C}_\mu = \frac{1}{m}\sum_{i=1}^{m}(x_i - \mu)(x_i - \mu)^T$.

Given μ, the ML estimation of Σ is

$$\max_{\Sigma \succ 0} \; -\log\det\Sigma - \mathbf{tr}(\hat{C}_\mu \Sigma^{-1})$$

which can be shown to have a nonconcave objective function.

Let $\Psi = \Sigma^{-1} \succ 0$. The ML problem is equiv. to

$$\max_{\Psi \succ 0} \; \log\det\Psi - \mathbf{tr}(\hat{C}_\mu \Psi)$$

which is a convex problem.

More about ML mean and covariance estimation Joint ML estimation of μ and Σ:

$$\begin{aligned} &\max_{\mu,\Psi=\Sigma^{-1}\succ 0} \underbrace{\log\det\Psi - \mathbf{tr}(\hat{C}_\mu\Psi)}_{\triangleq f_0(\mu,\Psi)} \\ &= \max_{\Psi\succ 0} \max_{\mu} f_0(\mu,\Psi) \end{aligned}$$

Consider the inner maximization given Ψ:

$$\nabla_\mu f_0 = -\tfrac{1}{m}\textstyle\sum_{i=1}^m \Psi x_i - \Psi\mu$$

$$\Longleftrightarrow \arg\max_\mu f_0(\mu,\Psi) = \tfrac{1}{m}\textstyle\sum_{i=1}^m x_i \triangleq \hat{\mu}$$

A neat result is that $\hat{\mu}$ does not depend on Ψ.

Putting the optimal μ back to the ML problem

$$\max_{\Psi \succ 0} \; \log\det\Psi - \mathbf{tr}(\hat{C}_{\hat{\mu}}\Psi)$$

Using

$$\nabla_{\Psi} f_0 = \Psi^{-1} - \hat{C}_{\mu},$$

we arrive at

$$\hat{\Sigma} = \hat{\Psi}^{-1} = \hat{C}_{\hat{\mu}},$$

assuming that $\hat{C}_{\hat{\mu}} \succ 0$.

Remark: the sampled mean and covariance are indeed the ML estimates, under the Gaussian assumption.

Question: $f_0(\mu, \Psi)$ is concave in either μ or $\Psi \succ 0$, but is it concave in μ and $\Psi \succ 0$?

Standard Form with Generalized Inequalities

A convex problem with generalized inequalities:

$$\begin{aligned} \min\ & f_0(x) \\ \text{s.t. } & f_i(x) \preceq_{K_i} 0, \quad i = 1, \ldots, L \\ & Ax = b \end{aligned}$$

where

- $\preceq_{K_i}$ are generalized inequalities on $\mathbf{R}^m$,
- $f_i : \mathbf{R}^n \rightarrow \mathbf{R}^m$ are K_i-convex.

Examples

- Second order cone program:

$$\begin{aligned} \min\ & c_0^T x \\ \text{s.t. } & \|A_i x + b_i\|_2 \leq c_i^T x + d_i, \quad i = 1, \ldots, m \\ & Fx = g \end{aligned}$$

- Semidefinite program:

$$\begin{aligned} \min\ & \mathbf{tr}(CX) \\ \text{s.t. } & X \succeq 0 \\ & \mathbf{tr}(A_i X) = b_i, \quad i = 1, \ldots, p \end{aligned}$$

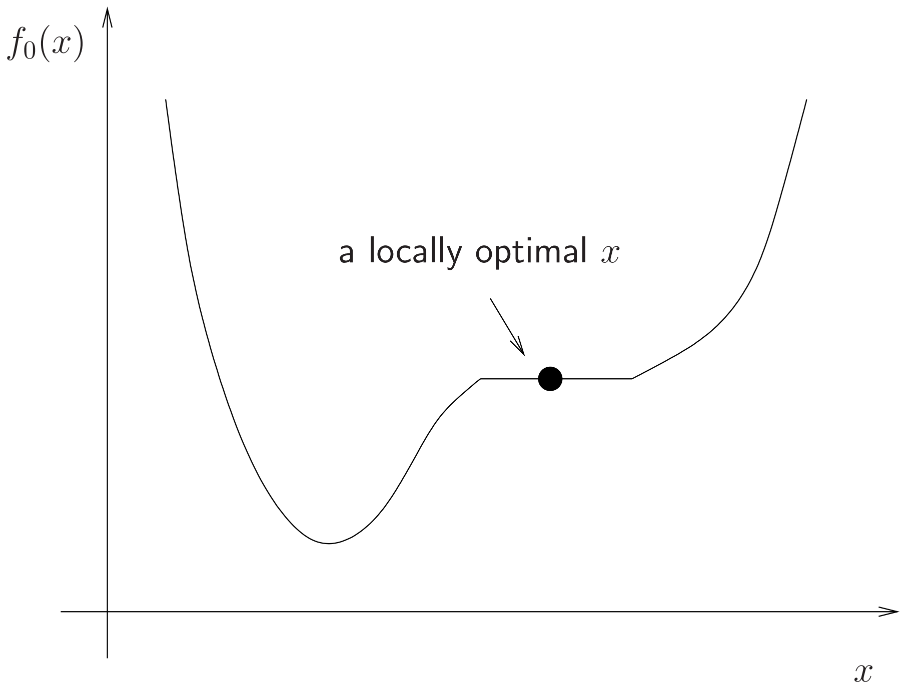
Quasiconvex Optimization

A quasiconvex opt. problem has a standard form

$$\begin{aligned} \min\ & f_0(x) \\ \text{s.t.}\ & f_i(x) \leq 0, \quad i = 1, \ldots, m \\ & Ax = b \end{aligned}$$

where $f_0(x)$ is quasiconvex, and $f_i(x)$ convex for $i > 1$.

- A quasiconvex problem can have locally optimal points



- A point $x \in C$ is optimal if

$$\nabla f_0(x)^T (y - x) > 0, \ \forall y \in C$$

But the converse is not always true.

Let $\phi_t : \mathbf{R}^n \to \mathbf{R}$, $t \in \mathbf{R}$, be a family of convex functions that satisfy

$$f_0(x) \leq t \Longleftrightarrow \phi_t(x) \leq 0$$

and also, for each x, $\phi_s(x) \leq \phi_t(x)$ whenever $s \leq t$.

Fixing a t, the following feasibility problem is convex

$$\begin{aligned} &\text{find } x \\ &\text{s.t. } \phi_t(x) \leq 0,\ f_i(x) \leq 0, \quad i = 1, \ldots, m \\ &\qquad Ax = b \end{aligned}$$

Idea: decrease t until the convex feasibility problem is infeasible.

Assume that $p^\star$ is known to lie within $[\ell, u]$.

Bisection method for quasiconvex optimization

given bounds ℓ, u, and tolerance ϵ.

repeat

1. $t := (\ell + u)/2$.
2. Solve the convex feasibility problem.
3. **if** the problem is feasible, $u := t$; **otherwise** $\ell := t$.

until $u - \ell \leq \epsilon$.