Convex Optimal Control and Sequential Convex Programming

Moritz Diehl

March 27, 2015

Overview

- Convex Optimization
- Convex Optimal Control
- Sequential Convex Programming

Recall:

- convex optimization $=$
  convex objective $f(x)$ and convex feasible set
- rules to get convex functions (if f_i are convex):
  - positive sum: $\sum_i w_i f_i(x)$ $(w_i \geq 0)$
  - affine input transformation: $f(Ax + b)$
  - maximum: $\max_i f_i(x)$
  - partial minimization: $\min_y f(x, y)$ (convex in x, y together)
  - examples: affine functions $a^\top x + b$, all norms $\|x\|$
- rules to get convex sets
  - hyperplanes (linear equalities): $a^\top x + b = 0$
  - sublevel sets of convex functions: $f(x) \leq 0$
  - intersection
  - projection: $\{x | \exists y : (x, y) \in \Omega\}$ (Ω convex)
  - Linear Matrix Inequalities (LMI): $\sum_{i=1}^n x_i A_i \succeq B$
    (B, A_i symmetric)

Examples: which function is convex?

- $f(x) = a^\top x + b$
- $f(x) = \min_{\|a - \bar{a}\|_2 \leq \gamma} a^\top x + b$

Examples: which function is convex?

- $f(x) = a^\top x + b$
- $f(x) = \min_{\|a-\bar{a}\|_2 \leq \gamma} a^\top x + b$ concave

Examples: which function is convex?

- $f(x) = a^\top x + b$
- $f(x) = \min_{\|a-\bar{a}\|_2 \leq \gamma} a^\top x + b$ concave
- $f(x) = \max_{\|a-\bar{a}\|_2 \leq \gamma} a^\top x + b$

Examples: which function is convex?

- $f(x) = a^\top x + b$
- $f(x) = \min_{\|a-\bar{a}\|_2 \leq \gamma} a^\top x + b$ concave
- $f(x) = \max_{\|a-\bar{a}\|_2 \leq \gamma} a^\top x + b \quad = \bar{a}^\top x + \gamma \|x\|_2 + b$ (robust LP = SOCP)

Examples: which function is convex?

- $f(x) = a^\top x + b$
- $f(x) = \min_{\|a-\bar{a}\|_2 \leq \gamma} a^\top x + b$ concave
- $f(x) = \max_{\|a-\bar{a}\|_2 \leq \gamma} a^\top x + b \quad = \bar{a}^\top x + \gamma \|x\|_2 + b$ (robust LP = SOCP)
- $f(x) = \min_{u,v} u^2 + 2Mv \text{ s.t. } -u-v \leq x \leq u+v, \; v \geq 0$

Examples: which function is convex?

- $f(x) = a^\top x + b$
- $f(x) = \min_{\|a-\bar{a}\|_2 \leq \gamma} a^\top x + b$ concave
- $f(x) = \max_{\|a-\bar{a}\|_2 \leq \gamma} a^\top x + b \quad = \bar{a}^\top x + \gamma\|x\|_2 + b$ (robust LP = SOCP)
- $f(x) = \min_{u,v} u^2 + 2Mv \text{ s.t. } -u-v \leq x \leq u+v,\ v \geq 0$ yes, because convex in x, u, v (Huber penalty)

Examples: which function is convex?

- $f(x) = a^\top x + b$
- $f(x) = \min_{\|a-\bar{a}\|_2 \leq \gamma} a^\top x + b$ concave
- $f(x) = \max_{\|a-\bar{a}\|_2 \leq \gamma} a^\top x + b \quad = \bar{a}^\top x + \gamma \|x\|_2 + b$ (robust LP = SOCP)
- $f(x) = \min_{u,v} u^2 + 2Mv \text{ s.t. } -u - v \leq x \leq u + v,\ v \geq 0$ yes, because convex in x, u, v (Huber penalty)
- $f(x) = \max_{\|y\| \leq \gamma} \begin{bmatrix} x \\ y \end{bmatrix}^\top \begin{bmatrix} Q & S \\ S^\top & R \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$

Examples: which function is convex?

- $f(x) = a^\top x + b$
- $f(x) = \min_{\|a-\bar{a}\|_2 \le \gamma} a^\top x + b$ concave
- $f(x) = \max_{\|a-\bar{a}\|_2 \le \gamma} a^\top x + b \quad = \bar{a}^\top x + \gamma \|x\|_2 + b$ (robust LP = SOCP)
- $f(x) = \min_{u,v} u^2 + 2Mv \text{ s.t. } -u - v \le x \le u + v,\ v \ge 0$ yes, because convex in x, u, v (Huber penalty)
- $f(x) = \max_{\|y\| \le \gamma} \begin{bmatrix} x \\ y \end{bmatrix}^\top \begin{bmatrix} Q & S \\ S^\top & R \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$ yes if $Q \succeq 0$

Examples: which function is convex?

- $f(x) = a^\top x + b$
- $f(x) = \min_{\|a-\bar{a}\|_2 \leq \gamma} a^\top x + b$ concave
- $f(x) = \max_{\|a-\bar{a}\|_2 \leq \gamma} a^\top x + b \quad = \bar{a}^\top x + \gamma \|x\|_2 + b$ (robust LP = SOCP)
- $f(x) = \min_{u,v} u^2 + 2Mv \text{ s.t. } -u-v \leq x \leq u+v, \; v \geq 0$ yes, because convex in x, u, v (Huber penalty)
- $f(x) = \max_{\|y\| \leq \gamma} \begin{bmatrix} x \\ y \end{bmatrix}^\top \begin{bmatrix} Q & S \\ S^\top & R \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$ yes if $Q \succeq 0$
- $f(x) = \min_{\|y\| \leq \gamma} \begin{bmatrix} x \\ y \end{bmatrix}^\top \begin{bmatrix} Q & S \\ S^\top & R \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$

Examples: which function is convex?

- $f(x) = a^\top x + b$
- $f(x) = \min_{\|a-\bar{a}\|_2 \le \gamma} a^\top x + b$ concave
- $f(x) = \max_{\|a-\bar{a}\|_2 \le \gamma} a^\top x + b \quad = \bar{a}^\top x + \gamma \|x\|_2 + b$ (robust LP = SOCP)
- $f(x) = \min_{u,v} u^2 + 2Mv \text{ s.t. } -u - v \le x \le u + v, \; v \ge 0$ yes, because convex in x, u, v (Huber penalty)
- $f(x) = \max_{\|y\| \le \gamma} \begin{bmatrix} x \\ y \end{bmatrix}^\top \begin{bmatrix} Q & S \\ S^\top & R \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$ yes if $Q \succeq 0$
- $f(x) = \min_{\|y\| \le \gamma} \begin{bmatrix} x \\ y \end{bmatrix}^\top \begin{bmatrix} Q & S \\ S^\top & R \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$ yes if $\begin{bmatrix} Q & S \\ S^\top & R \end{bmatrix} \succeq 0$

Examples: which function is convex?

- $f(x) = a^\top x + b$
- $f(x) = \min_{\|a-\bar{a}\|_2 \le \gamma} a^\top x + b$ concave
- $f(x) = \max_{\|a-\bar{a}\|_2 \le \gamma} a^\top x + b \quad = \bar{a}^\top x + \gamma \|x\|_2 + b$
  (robust LP = SOCP)
- $f(x) = \min_{u,v} u^2 + 2Mv \text{ s.t. } -u - v \le x \le u + v,\ v \ge 0$ yes, because convex in x, u, v (Huber penalty)
- $f(x) = \max_{\|y\| \le \gamma} \begin{bmatrix} x \\ y \end{bmatrix}^\top \begin{bmatrix} Q & S \\ S^\top & R \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$ yes if $Q \succeq 0$
- $f(x) = \min_{\|y\| \le \gamma} \begin{bmatrix} x \\ y \end{bmatrix}^\top \begin{bmatrix} Q & S \\ S^\top & R \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$ yes if $\begin{bmatrix} Q & S \\ S^\top & R \end{bmatrix} \succeq 0$
  Note: $f(x) = x^\top (Q - SR^{-1}S^\top) x$ if $\gamma = \infty$ and $R \succ 0$
  (Schur complement)

More examples

- $f(X) = \textbf{trace}(X^{-1})$

More examples

- $f(X) = \textbf{trace}(X^{-1}) = \min_Y \textbf{trace}(Y) \text{ s.t. } Y - X^{-1} \succeq 0$
  $= \min_Y \textbf{trace}(Y) \text{ s.t. } \begin{bmatrix} Y & I \\ I & X \end{bmatrix} \succeq 0.$ (A-criterion)
- $f(X) = \lambda_{\max}(X^{-1})$

More examples

- $f(X) = \textbf{trace}(X^{-1}) = \min_Y \textbf{trace}(Y) \ \text{s.t.}\ Y - X^{-1} \succeq 0$
  $= \min_Y \textbf{trace}(Y) \ \text{s.t.}\ \begin{bmatrix} Y & I \\ I & X \end{bmatrix} \succeq 0$. (A-criterion)
- $f(X) = \lambda_{\max}(X^{-1}) = \min_\lambda \lambda \ \text{s.t.}\ \lambda I - X^{-1} \succeq 0$
  $= \min_\lambda \lambda \ \text{s.t.}\ \begin{bmatrix} \lambda I & I \\ I & X \end{bmatrix} \succeq 0$. (E-criterion)

Examples: which set or inequality is convex?

- $\|x\| \leq 1$
- $\|x\| \geq 1$

Examples: which set or inequality is convex?

- $\|x\| \leq 1$
- $\|x\| \geq 1$ not

Examples: which set or inequality is convex?

- $\|x\| \leq 1$
- $\|x\| \geq 1$ not
- $\|x\|_2^2 + \|y\|_2^2 \leq 1$
- $x^\top y \leq 1$

Examples: which set or inequality is convex?

- $\|x\| \leq 1$
- $\|x\| \geq 1$ not
- $\|x\|_2^2 + \|y\|_2^2 \leq 1$
- $x^\top y \leq 1$ not

Examples: which set or inequality is convex?

- $\|x\| \leq 1$
- $\|x\| \geq 1$ not
- $\|x\|_2^2 + \|y\|_2^2 \leq 1$
- $x^\top y \leq 1$ not
- $\|x\|_2^2 + \|y\|_2^2 \leq 1 - 2x^\top y$

Examples: which set or inequality is convex?

- $\|x\| \leq 1$
- $\|x\| \geq 1$ not
- $\|x\|_2^2 + \|y\|_2^2 \leq 1$
- $x^\top y \leq 1$ not
- $\|x\|_2^2 + \|y\|_2^2 \leq 1 - 2x^\top y \quad \Leftrightarrow \|x + y\|_2^2 \leq 1$ yes

Examples: which set or inequality is convex?

- $\|x\| \leq 1$
- $\|x\| \geq 1$ not
- $\|x\|_2^2 + \|y\|_2^2 \leq 1$
- $x^\top y \leq 1$ not
- $\|x\|_2^2 + \|y\|_2^2 \leq 1 - 2x^\top y \quad \Leftrightarrow \|x+y\|_2^2 \leq 1$ yes
- $\{w, M | (a_i - w)^\top M^{-1}(a_i - w) \leq 1, i = 1, \ldots, m\}$

Examples: which set or inequality is convex?

- $\|x\| \leq 1$
- $\|x\| \geq 1$ not
- $\|x\|_2^2 + \|y\|_2^2 \leq 1$
- $x^\top y \leq 1$ not
- $\|x\|_2^2 + \|y\|_2^2 \leq 1 - 2x^\top y \quad \Leftrightarrow \|x+y\|_2^2 \leq 1$ yes
- $\{w, M | (a_i - w)^\top M^{-1}(a_i - w) \leq 1, i = 1, \ldots, m\}$
  $= \left\{ w, M \middle| \begin{bmatrix} 1 & (a_i - w)^\top \\ a_i - w & M \end{bmatrix} \succeq 0, i = 1, \ldots, m \right\}$
  (bounding ellipsoid = LMI)

Interesting Example

$$
\begin{aligned}
f(x_{\text{init}}) = \quad & \min_{x,u} && \sum_{k=0}^{N-1} \|x_k\|_2^2 + \|u_k\|_2^2 \\
& \text{subject to} && x_{k+1} = Ax_k + Bu_k, \quad && k = 0, \ldots, N-1, \\
& && -1 \le u_k \le 1, && k = 0, \ldots, N-1, \\
& && x_0 = x_{\text{init}}, \\
& && x_N = 0.
\end{aligned}
$$

Question: is f convex ?

General Convex Optimal Control Problem

$$
\begin{aligned}
V(x_{\text{init}}) = \min_{x,u} \quad & \sum_{k=0}^{N-1} L_k(x_k, u_k) \; + \; E(x_N) \\
\text{subject to} \quad & x_{k+1} = A_k x_k + B_k u_k + c_k, \quad && k = 0, \dots, N-1, \\
& (x_k, u_k) \in \Omega_k, && k = 0, \dots, N-1, \\
& x_0 = x_{\text{init}}, \\
& x_N \in \Omega_N.
\end{aligned}
$$

Note: Cost-to-go of a convex optimal control problem is convex

Overview

- Convex Optimization
- Convex Optimal Control
- **Sequential Convex Programming**

NLP with Convex Structure

$$\begin{aligned} \underset{x}{\text{minimize}} \quad & \phi(F(x)) \\ \text{subject to} \quad & G(x) = 0, \\ & H(x) \in \Omega. \end{aligned}$$

with ϕ and Ω convex and F, G, H smooth functions.
SCP idea: linearize F, G, H to obtain convex problem.
Notation: $F_{\mathrm{L}}(x; \bar{x}) = F(\bar{x}) + \frac{\partial F}{\partial x}(\bar{x})(x - \bar{x})$

SCP subproblem

At solution guess x_k solve

$$\begin{aligned} \underset{x}{\text{minimize}} \quad & \phi(F_{\mathrm{L}}(x; x_k)) + \frac{1}{2}(x - x_k)^\top M_k (x - x_k) \\ \text{subject to} \quad & G_{\mathrm{L}}(x; x_k) = 0, \\ & H_{\mathrm{L}}(x; x_k) \in \Omega. \end{aligned}$$

with $M_k \succeq 0$. Take solution $x^*(x_k)$ as new solution guess x_{k+1}.

Examples for SCP

- Gauss-Newton $\min_x \|F_\mathrm{L}(x; x_k)\|_2^2$
- Generalized Gauss-Newton [Bock 1983]: $\min_x \|F_\mathrm{L}(x; x_k)\|_2^2 \;\; \mathrm{s.t.} \; G_\mathrm{L}(x; x_k) = 0$
- Sequential Linear Programming (SLP)
- Sequential Semi-Definite Programming [Fares, Noll, Apkarian, SICON, 2002]
- Method of Moving Asymptotes (slightly different)

SCP for mathematical analysis

Note: can remove F, H with slacks and new equality

$$g_{\text{new}}(x, y, z) = \begin{bmatrix} F(x) - y \\ H(x) - z \\ G(x) \end{bmatrix} = 0$$

The SCP algorithm will remain identical.
Can also move ϕ into Ω with epigraph trick $(t \geq \phi(x))$. SCP algorithm still remains the same.

$$\begin{array}{ll} \underset{x}{\text{minimize}} & c^\top x \\ \text{subject to} & g(x) = 0, \\ & x \in \Omega. \end{array}$$

Define partial Lagrangian $L(x, \lambda) = c^\top x + \lambda^\top g(x)$.

SCP convergence theory

$$
\begin{aligned}
\underset{x}{\text{minimize}} \quad & c^\top x + \frac{1}{2}(x - x_k)^\top M_k (x - x_k) \\
\text{subject to} \quad & g_{\mathrm{L}}(x; x_k) = 0, \\
& x \in \Omega.
\end{aligned}
$$

Note: "SCP" would converge quadratically if nonconvex subproblems would be allowed with $M_k = \nabla_x^2 L(x_k, \lambda_k)$.
But: can prove that with convex subproblems and bounded regularization ($R \succeq M_k \succeq 0$) superlinear convergence is impossible [Diehl, Jarre, Vogelbusch, SIOPT, 2006].
Problem: negative curvature in partial Hessian $\nabla_x^2 L$.
Making M_k bigger than true Hessian slows down the algorithm.
General convergence theory can be extended to inexact Jacobians (if gradient correction is used) and requires sufficiently good KKT matrix approximation [Quoc Tran Dinh et al, SIOPT, 2012]

Is regularization needed ?

Regard example $\min_{x,y} y \;\text{ s.t. }\; y = x^2$ and linearize equality to $y = \bar{x}^2 + 2\bar{x}(x - \bar{x})$. SCP subproblem is unbounded for $\bar{x} \neq 0$. For small regularization, SCP still diverges.

Note: at solution, we have $\nabla^2_{x,y} L(x, y, \lambda_*) = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}$.

Observation: algorithm converges if we set $M_k \succ \frac{1}{2}\nabla^2_{x,y} L(x, y, \lambda_*)$

Conjecture: local convergence of SCP at solution (x_*, λ_*) can be guaranteed if $M_k \succeq \nabla^2 L(x_*, \lambda_*)$ [to be proven]

Nicest Case for SCP: Concave Inequalities

Regard concave H (i.e. tangents lie above graph).
SCP subproblem:

$$\begin{aligned} \underset{x}{\text{minimize}} \quad & \phi(x) \\ \text{subject to} \quad & H_{\mathrm{L}}(x; x_k) \leq 0, \qquad x \in \Omega \end{aligned}$$

Lemma: If x_0 is feasible, all iterates remain feasible.
Proof: If $H(x_k) \leq 0$ then $H(x_{k+1}) \leq H_{\mathrm{L}}(x_{k+1}; x_k) \leq 0$.
No globalisation necessary. Easiest SCP method.
Note: partial Hessian $\sum_i \mu_i \nabla^2 H_i$ *negative* semidefinite.
Possibly slow but robust convergence.

Variants of Nicest Case: Convex Concave Decompositions
Regard $H(x) = H_1(x) - H_2(x)$ with H_1, H_2 convex.
Form SCP subproblem:

$$\begin{aligned} \underset{x}{\text{minimize}} \quad & \phi(x) \\ \text{subject to} \quad & H_1(x) - H_{2,\mathrm{L}}(x; x_k) \leq 0, \qquad x \in \Omega \end{aligned}$$

Again, all iterates remain feasible, same nice method as before.
Also known as DC decomposition (Difference of Convex functions).
Can generalize to nonconvex matrix inequalities [Quoc Tran Dinh et al, IEEE TAC 2012]

Parametric Pathfollowing with Real-Time SCP(RTSCP)

Regard parameter dependent problems:

$$\begin{aligned} \underset{x}{\text{minimize}} \quad & \phi(F(x,p)) \\ \text{subject to} \quad & G(x,p) = 0, \\ & H(x,p) \in \Omega. \end{aligned}$$

Idea 1: Use slacks to make dependence *affine* on parameter.

$$\begin{aligned} \underset{x}{\text{minimize}} \quad & \phi(F(x) + A_F p) \\ \text{subject to} \quad & G(x) + A_G p = 0, \\ & H(x) + A_H p \in \Omega. \end{aligned}$$

Idea 2: linearize at x_k but use new value p_{k+1} in SCP subproblem. Get generalized tangential predictor for true solution [Diehl 2001, Quoc Tran Dinh et al., SIOPT, 2012] (tangential if $x_k = x^*(p_k)$ and $M_k = \nabla_x^2 L(x_k, \lambda_k)$).

Optimal Control Problem with Convex Structure

$$
\begin{aligned}
\underset{x,u}{\text{minimize}} \quad & \sum_{k=0}^{N-1} L_k(x_k, u_k) \;+\; E(x_N) \\
\text{subject to} \quad & x_{k+1} = f_k(x_k, u_k), && k = 0, \ldots, N-1, \\
& (x_k, u_k) \in \Omega_k, && k = 0, \ldots, N-1, \\
& x_0 = x_{\text{init}}, \\
& x_N \in \Omega_N.
\end{aligned}
$$

SCP idea: linearize nonlinear equalities (system dynamics).
RTSCP for Nonlinear Model Predictive Control: Solve for latest state estimate x_{init} (note: x_{init} enters problem affinely).