Chapter 3: Optimization for Data Science
Convex Optimization Problems

TANN Chantara

Institute of Technology of Cambodia

October 9, 2022

Table of Contents

Optimization Problems in Standard Form

$$P : \begin{cases} \min_{x\in\mathbb{R}^n} & f_0(x) \\ \text{s.t.} & f_i(x) \leq 0, \forall i = 1, .., m \\ & h_i(x) = 0, \forall i = 1, ..., p \end{cases}$$

- $x \in \mathbb{R}^n$ — opitmization or decision variable
- $f_0 : \mathbb{R}^n \to (-\infty, \infty]$ — objective or cost function
- $f_i : \mathbb{R}^n \to (-\infty, \infty]$ — inequality constraint functions
- $h_i : \mathbb{R}^n \to \mathbb{R}$ — equality constraint functions

$$\inf P = \infty \iff P \text{ is infeasible}$$

$$\inf P = -\infty \iff P \text{ is unbounded}$$

Transforming Problems into the Standard Form

- "$\geq$"-inequalities

$$\hat{f}(x) \geq 0 \iff f(x) = -\hat{f}(x) \leq 0$$

- Nonzero right-hand-sides

$$\hat{f}(x) \leq b \iff f(x) = \hat{f}(x) - b \leq 0$$

- Maximization problems

$$\sup_{x \in \mathcal{X}} \hat{f}_0(x) = -\inf_{x \in \mathcal{X}} f_0(x), \text{ where } f_0(x) = -\hat{f}_0(x)$$

Note: In principle all equality constraints can be transformed into inequality constraints

$$f(x) = 0 \iff f(x) \leq 0 \text{ and } -f(x) \leq 0$$

Convex Optimization Problem

A convex optimization problem is give as

$$P:\begin{cases}\min_{x\in\mathbb{R}^n} & f_0(x)\\ \text{s.t.} & f_i(x)\leq 0, \forall i=1,\dots,m\\ & a_i^T x = b_i, \forall i = 1,\dots,p\end{cases}$$

- Functions f_0 and f_i are all convex.
- Equality constraints are linear (affine), can be written as $Ax = b$.

Proposition: Then feasible set of a convex optimization problem is convex.

Proof: The feasible set of P is given by

$$\operatorname{dom}(f_0)\bigcap_{i=1}^{m}\{x : f_i(x)\leq 0\}\bigcap_{i=1}^{m}\{x : a_i^T = b_i\}$$

It is convex as the domains and sublevel sets of convex functions are convex, while hyperplanes are also convex. Intersections of convex sets are convex.

Local and Global Optima

Theorem: Every locally optimal solution of a convex optimization problem is also global optimally.

Proof: Consider a local optimizer x. Suppose that x is not a global minimizer, and let $y \neq x$ be a global minimizer. By convexity of the feasible set, $\theta x + (1-\theta)y$ is feasible for every $\theta \in [0,1]$. Its objective values satisfies

$$\begin{aligned} f_0(\theta x + (1-\theta)y) &\leq \theta f_0(x) + (1-\theta) f_0(y) \\ &< \theta f_0(x) + (1-\theta) f_0(x) \\ &= f_0(x) \end{aligned}$$

for all $\theta \in (0,1)$. Therefore, x cannot be a local minimizer. By contraposition, if x is a local minimizer, then it must be a global minimizer.

Optimality Criterion (Differentiable Objective)

Theorem: For P convex, x^* is optimal iff it is feasible and

$$\nabla f_0(x^*)^T(x - x^*) \geq 0 \qquad \text{for all feasible } x \qquad (1)$$

Proof: Assume (1) holds for some feasible x^* and let x be any feasible point. The 1st-order conditions for convex functions f_0 imply

$$f_0(x) \geq f_0(x^*) + \nabla f_0(x^*)^T(x - x^*) \quad \forall x \in \mathbb{R}^n$$

Using (1), we conclude that $f_0(x) \geq f_0(x^*)$. Thus, x^* is optimal.
Conversely, assume that x^* is optimal and let x be any feasible point. By convexity, the line segment $[x^*, x]$ is contained in the feasible set. By optimality of x^*, the function

$$\phi(\theta) = f_0((1-\theta)x^* + \theta x) \qquad (2)$$

Convex Optimization Problem

A convex optimization problem is give as

$$P : \begin{cases} \min_{x \in \mathbb{R}^n} & f_0(x) \\ \text{s.t.} & f_i(x) \leq 0, \forall i = 1, ..., m \\ & a_i^T x = b_i, \forall i = 1, ..., p \end{cases}$$

- Functions f_0 and f_i are all convex.
- Equality constraints are linear (affine), can be written as $Ax = b$.

Proposition: Then feasible set of a convex optimization problem is convex.

Proof: The feasible set of P is given by

$$\mathrm{dom}(f_0) \bigcap_{i=1}^{m} \{x : f_i(x) \leq 0\} \bigcap_{i=1}^{m} \{x : a_i^T = b_i\}$$

It is convex as the domains and sublevel sets of convex functions are convex, while hyperplanes are also convex. Intersections of convex sets are convex.

Local and Global Optima

Theorem: Every locally optimal solution of a convex optimization problem is also global optimally.

Proof: Consider a local optimizer x. Suppose that x is not a global minimizer, and let $y \neq x$ be a global minimizer. By convexity of the feasible set, $\theta x + (1-\theta)y$ is feasible for every $\theta \in [0, 1]$. Its objective values satisfies

$$\begin{aligned} f_0(\theta x + (1-\theta)y) &\leq \theta f_0(x) + (1-\theta)f_0(y) \\ &< \theta f_0(x) + (1-\theta)f_0(x) \\ &= f_0(x) \end{aligned}$$

for all $\theta \in (0, 1)$. Therefore, x cannot be a local minimizer. By contraposition, if x is a local minimizer, then it must be a global minimizer.

Optimality Criterion (Differentiable Objective)

Theorem: For P convex, x^* is optimal iff it is feasible and

$$\nabla f_0(x^*)^T(x - x^*) \geq 0 \qquad \text{for all feasible } x \qquad (1)$$

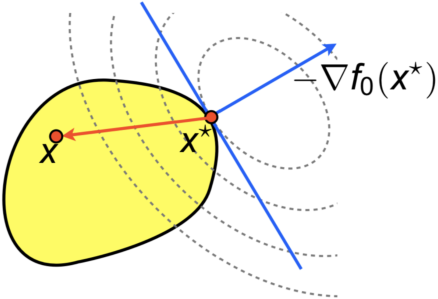
Proof: Assume (1) holds for some feasible x^* and let x be any feasible point. The 1st-order conditions for convex functions f_0 imply

$$f_0(x) \geq f_0(x^*) + \nabla f_0(x^*)^T(x - x^*) \quad \forall x \in \mathbb{R}^n$$

Using (1), we conclude that $f_0(x) \geq f_0(x^*)$. Thus, x^* is optimal.
Conversely, assume that x^* is optimal and let x be any feasible point. By convexity, the line segment $[x^*, x]$ is contained in the feasible set. By optimality of x^*, the function $\varphi(\theta) = f_0((1-\theta)x^* + \theta x) \qquad (2)$
must satisfy $\varphi(\theta) \geq \varphi(0) \forall \theta \in [0, 1]$. This only possible if

$$\varphi'(0) = \nabla f_0(x^*)^T(x - x^*) \geq 0$$

Optimality Criterion (Differentiable Objective) Cont'd



x^* is optimal if $-\nabla f_0(x^*)$ has an angle of more than 90° with $(x - x^*)$ for every feasible x

Special Case: Unconstrained problem:

$$x^* \text{optimal} \iff x^* \in \operatorname{dom}(f_0), \nabla f_0(x^*) = 0$$

Equivalent Optimization Problems

Two optimization problem P and P' are equivalent $(P \iff P')$ if the solution of P' is obtained from P via elementary transformations and vice versa.

Epigraphical reformulation: The standard form convex optimization problem is equivalent to

$$P' : \begin{cases} \min_{x\in\mathbb{R}^n, t\in\mathbb{R}} & t \\ \text{s.t.} & f_0(x) - t \leq 0 \\ & f_i(x) \leq 0, \forall i = 1, ..., m \\ & Ax = b \end{cases}$$

$\implies$ One can always assume a linear objective function

Equivalent Optimization Cont'd

Partial minimization: The optimization problem

$$P : \begin{cases} \min_{x_1, x_2} & f_0(x_1, x_2) \\ \text{s.t.} & f_i(x_1) \leq 0, \forall i = 1, \ldots, m \\ & g_i(x_2) \leq 0, \forall i = 1, \ldots, q \end{cases}$$

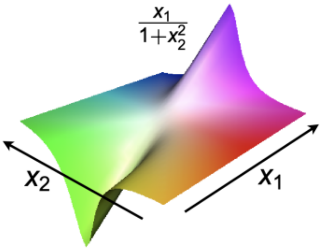
is equivalent to

$$P' : \begin{cases} \min_{x_1} & \bar{f}_0(x_1) \\ \text{s.t.} & f_i(x_1) \leq 0, \forall i = 1, \ldots, m \end{cases}$$

$$\bar{f}_0(x_1) = \inf_{x_2} \{f_0(x_1, x_2) : g_i(x_2) \leq 0, \forall i = 1, \ldots, q\}$$

Equivalent Optimization Cont'd

Example: The following problems are equivalent



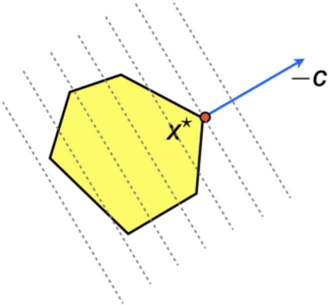
$$P: \begin{cases} \min_{x_1,x_2} & x_1^2 + x_2^2 \\ \text{s.t.} & x_1/(1+x_2^2) \le 0 \\ & (x_1+x_2)^2 = 0 \end{cases}$$

$$P': \begin{cases} \min_{x_1,x_2} & x_1^2 + x_2^2 \\ \text{s.t.} & x_1 \le 0 \\ & x_1+x_2 = 0 \end{cases}$$

- P is nonconvex (involves nonconvex inequality constriant and nonlinear equality constraint)
- P′ is convex

Linear program (LP)

$$\begin{cases} \min_{x \in \mathbb{R}^n} & c^T x + d \\ \text{s.t.} & Ax = b \\ & Cx \leq g \end{cases}$$

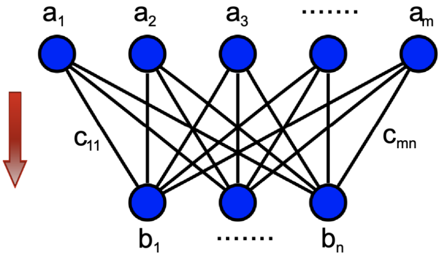


Affine objective and constraints, polyhedral feasible set

Problems with n variables and encoded in L input bits can be solved in $O(n^{3.5}L)$ arithmetic operations via interior point methods.

Transportation problem

$$\begin{cases} \min_{x \in \mathbb{R}^{m \times n}} & \sum_{i=1}^{m} \sum_{j=1}^{n} c_{ij} x_{ij} & \text{transportation costs} \\ \text{s.t.} & \sum_{j=1}^{n} x_{ij} = a_i, \forall i & \text{total shipment from source i} \\ & \sum_{i=1}^{m} x_{ij} = b_j, \forall j & \text{total shipment to destination j} \\ & x_{ij} \geq 0, \forall i, j & \text{nonnegative shipment only} \end{cases}$$



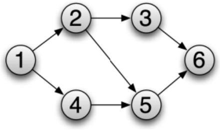
Problem feasible if supply = total demand

$$\sum_{i=1}^{m} a_i = \sum_{j=1}^{n} b_j$$

Project Scheduling

Project: directed acyclic graph $G = (V, E)$

- tasks $V = \{1, 2, \ldots, n\}$
- precendences $E \subset V \times V$
- task durations $\delta_v, v \in V$



Early start policy:
Find earliest start times of project n via an LP

$$\begin{cases} \min_{S \in \mathbb{R}^n_+} & S_n \\ \text{s.t.} & S_v \geq S_u + \delta_u, \forall (u, v) \in E \end{cases}$$

Quadratic program (QP)

$$\begin{cases} \min_{x\in\mathbb{R}^n} & \frac{1}{2}x^TPx + q^Tx + r \\ \text{s.t.} & Ax = b \\ & Cx \leq g \end{cases}$$

Convex quadratic objective ($P \in \mathbb{S}^n_+$) and affine constraints, i.e., a polyhedral feasible set.

Example: Least-squares problem

$$\min_{x\in\mathbb{R}^n} ||Ax - b||_2$$

If A has full column rank, we have

$$0 = \nabla||Ax - b||_2^2 = 2A^TAx - 2A^Tb \iff x = (A^TA)^{-1}A^Tb$$

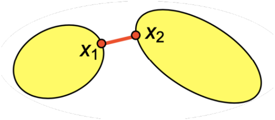
Quadratically constrained QP (QCQP)

$$\begin{cases} \min_{x\in\mathbb{R}^n} & \frac{1}{2}x^T P_0 x + q_0^T x + r_0 \\ \text{s.t.} & \frac{1}{2}x^T P_i x + q_i^T x + r_i \leq 0, i = 1, \ldots, m \end{cases}$$

Convex quadratic objective and constrains ($P_i \in \mathbb{S}^n_+$), i.e., the easible set is an intersection of ellipsoids (for $P_i \succ 0$) and halfspaces ($P_0 = 0$). Every LP is a QCQP with $P_i = 0$.

Example: Distance between ellipsoids

$$\begin{cases} \min_{x_1,x_2\in\mathbb{R}^n} & ||x_1 - x_2||_2^2 \\ \text{s.t.} & (x_1 - \mu_1)^T \sum_1^{-1}(x_1 - \mu_1) \leq 1 \\ & (x_1 - \mu_1)^T \sum_2^{-1}(x_2 - \mu_1) \leq 1 \end{cases}$$



Second-Order Cone Programming (SOCP)

$$\begin{cases} \min_{x\in\mathbb{R}^n} & f^T x \\ \text{s.t.} & ||A_i x - b_i||_2 \le c_i^T x + d_i, i = 1, ..., m \\ & Fx = g \end{cases}$$

$(f \in \mathbb{R}^n, A_i \in \mathbb{R}^{n_i \times n}, b_i \in \mathbb{R}^{n_i}, d_i \in \mathbb{R}, F \in \mathbb{R}^{p\times p}, g \in \mathbb{R}^p)$

- The inequalities are called second-order cone constraints are the $(n_i + 1)$-dimensional vector $(A_i x + b_i, c_i^T x + d_i)$ belongs to the second-order cone in $\mathbb{R}^{n_i+1}$.
- Every LP is an SOCP with $A_i = 0, b_i = 0$. Every SOCP with $c_i = 0$ is a QCQP, and every QCQP is an SOCP.
- ε-optimal solution can be found in $O((\sum_{i=1}^m n_i)^{3.5} \log(\varepsilon^{-1}))$ arithmetic operations via interior point methods.

"Hidden" SOCP Constraints

Quadratic constraints:

$$||x||_2^2 \leq t \iff \left\| \begin{pmatrix} 2x \\ t-1 \end{pmatrix} \right\|_2 \leq t+1$$

Hyperbolic constraints:

$$||x||_2^2 \leq st, s \geq 0, t \geq 0 \iff \left\| \begin{pmatrix} 2x \\ s-t \end{pmatrix} \right\|_2 \leq s+t, s \geq 0, t \geq 0$$

Problem with Generalized Inequalities

Standard form convex optimization problems with generalized inequalities

$$P : \begin{cases} \min_{x \in \mathbb{R}^n} & f_0(x) \\ \text{s.t.} & f_i(x) \preceq_{K_i} 0, \forall i = 1, \dots, m \\ & Ax = b \end{cases}$$

Assumption:

- $f_0 : \mathbb{R}^n \to \mathbb{R}$ convex, $f_i : \mathbb{R}^n \to \mathbb{R}^{k_i}$ K_i-convex.
- $K_i \subset \mathbb{R}^{k_i}$ is a proper convex cone.

Observations:

- P has a convex feasible set.
- All local optima of P are global optima

Conic Form Problems

Special case where objective and constraints are affine in x

$$P : \begin{cases} \min_{x\in\mathbb{R}^n} & c^T x \\ \text{s.t.} & Fx + g \preceq_K 0 \\ & Ax = b \end{cases}$$

This conic form problem reduces to an LP if $K = \mathbb{R}^n_+$

Note that also the SOCP

$$\begin{cases} \min_{x\in\mathbb{R}^n} & f^T x \\ \text{s.t.} & ||A_i x - b_i||_2 \leq c_i^T x + d_i, i = 1, \ldots, m \\ & Fx = g \end{cases}$$

is equivalent to

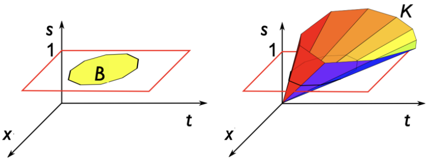
$$\begin{cases} \min_{x\in\mathbb{R}^n} & f^T x \\ \text{s.t.} & (A_i x \mid b_i, c^T x \mid d_i) \preceq_{K_i} 0, i = 1, \ldots, m \\ & Fx = g, \text{ where } K_i = \{(y,t) \in \mathbb{R}^{n_i+1} : ||y||_2 \leq t\} \end{cases}$$

Every Convex Problem is Conic

Theorem: Every convex optimization problem can be reformulated as a conic form problem.

Proof idea: Consider an arbitrary convex problem $\inf_{x \in C} f_0(x)$, where C is a convex feasible set. This problem is equivalent to

$$
\begin{aligned}
&\inf\{t : f_0(x) \le t, x \in C\} \\
\iff &\inf\{t : (t, x) \in B\} && B = \{(t, x) : f_0(x) \le t, x \in C\} \\
\iff &\inf\{t : s = 1, (s, t, x) \succeq_K 0\} && K = \{(s, t, x) : (t, x)/s \in B, s \ge 0\}
\end{aligned}
$$



Note: K is a convex cone (why?)

Semidefinite Program (SDP)

$$\begin{cases} \min_{x\in\mathbb{R}^n} & c^T x \\ \text{s.t.} & F_1x_1 + \cdots + F_nx_n \preceq G \\ & Ax = b \end{cases}$$

$(c \in \mathbb{R}^n, F_1, ..., F_n, G \in \mathbb{S}^k, A \in \mathbb{R}^{m\times n}, b \in \mathbb{R}^m)$
The semidefinite constraint is called a linear matrix inequality (LMI). Note that several LMIs can be combined to a single LMI:

$$\tilde{F}_1x_1 + \cdots +_n \preceq \text{ and } \hat{F}_1x_1 + \cdots + \hat{F}_n \preceq \hat{G}$$

$$\Longleftrightarrow$$

$$\begin{pmatrix} \tilde{F}_1 & 0 \\ 0 & \hat{F}_1 \end{pmatrix} x_1 + \cdots + \begin{pmatrix} {}_n & 0 \\ 0 & \hat{F}_n \end{pmatrix} x_n \preceq \begin{pmatrix} & 0 \\ 0 & \hat{G} \end{pmatrix}$$

An ε-optimal solution can be found in $O(n^2k^{2.5}\log(\varepsilon^{-1}))$ arithmetic operations via interior point methods.

Every SOCP is an SDP

The SOCP

$$\begin{cases} \min_{x\in\mathbb{R}^n} & f^T x \\ \text{s.t.} & \|A_i x - b_i\|_2 \le c_i^T x + d_i, i = 1, \dots, m \\ & Fx = g \end{cases}$$

is equivalent to an SDP by Schur's lemma

$$\|A_i x - b_i\|_2 \le c_i^T x + d_i \iff \begin{pmatrix} c_i^T x + d_i & (A_i x - b_i)^T \\ A_i x - b_i & (c_i^T x + d_i) I \end{pmatrix} \succeq 0$$

Every LP is an SDP

We have already seen that every LP is an SOCP and every SOCP is an SDP. Thus, every LP is an SDP.
Direct argument: The LP

$$\begin{cases} \min_{x\in\mathbb{R}^n} & c^T x + d \\ \text{s.t.} & Ax = b \\ & Cx \leq g \end{cases}$$

is equivalent to the SDP

$$\begin{cases} \min_{x\in\mathbb{R}^n} & c^T x + d \\ \text{s.t.} & \begin{pmatrix} \operatorname{diag}(b-Ax) & 0 & 0 \\ 0 & \operatorname{diag}(Ax-b) & 0 \\ 0 & 0 & \operatorname{diag}(g-Cx) \end{pmatrix} \succeq 0 \end{cases}$$

Eigenvalue Minimization

Minimize the largest eigenvalue of a matrix, i.e., solve

$$\min_{x\in\mathbb{R}^n} \lambda_{\max}(A(x)) \qquad (3)$$

where $A(x) = A_0 + A_1x_1 + \cdots + A_nx_n, A_j \in \mathbb{S}^k, \forall i = 0, \ldots, n$

Proposition: Problem (3) is equivalent to the SDP

$$\min_{x\in\mathbb{R},t\geq 0} \{t : A(x) \preceq tI\}$$

Proof: Consider $X = R \wedge R^T \in \mathbb{S}^k$, with $\wedge = \mathrm{diag}(\lambda_1(X), \ldots, \lambda_k(X))$ and R orthogonal $(R^T = R^{-1})$. Then,

$$\begin{aligned}
\lambda_{\max}(X) \leq t &\iff \lambda_i(X) \leq t, \forall i = 1, \ldots, k \\
&\iff \wedge \preceq tI \\
&\iff X = R \wedge R^T \preceq tRIR^T = tI
\end{aligned}$$

"Hidden" SDP Constraints

Upper bounds on inverse matrices:

$$X \succeq Y^{-1} \quad \Longleftrightarrow \quad \begin{pmatrix} X & I \\ I & Y \end{pmatrix} \succeq 0$$

Upper bounds on squared matrices:

$$X \succeq YY^T \quad \Longleftrightarrow \quad \begin{pmatrix} X & Y \\ Y^T & I \end{pmatrix} \succeq 0$$

Quadratic matrix inequalities:

$$(AXB)^T AXB + CXD + (CXD)^T \succeq Y$$

$$\Longleftrightarrow \begin{pmatrix} Y - CXD_(CXD)^T & (AXB)^T \\ AXB & I \end{pmatrix} \succeq 0$$

To prove this inequalities: use Schur's Lemma.

Popular Solvers

- **CPLEX** (https://www.ibm.com/analytics/cplex-optimizer): Caters for LP, QP, SOCP and MILP (mixed-integer LP); free for academic use.
- **Gurobi** (www.gurobi.com): caters for LP, QP, SOCP and MISOCP (mixed-integer SOCP); can be deployed on the cloud; free for academic use.
- **MOSEK** (http://www.mosek.com): caters for LP, QP, SOCP, SDP and MISDP (mixed-integer SDP); ideal for large-scale sparse problems; free for academic use.
- **SDPT3** (http://www.math.cmu.edu/ reha/sdpt3.html): caters for SDP; free, runs on Matlab.
- **IPOPT** (https://github.com/coin-or/Ipopt): caters for general NLPs (nonlinear programs); free.
- **YALMIP** (https://yalmip.github.io): modeling language for optimization problems implemented as a free toolbox for Matlab. A list of solvers that can be used with Yalmip is available from

Summary

- **Convex Optimization Problems:** convex objective and inequality constraints, linear equality constraints; feasible set is always convex; every local minimizer is a global minimizer.
- **Optimality Criterion** x^* is optimal iff the gradient of the objective makes an acute angle with every feasible direction.
- **Problem Classes:** definitions of LP, QP, QCQP, SOCP, SDP and conic form problems; (strict!) subset relations between problem classes.
- **Modeling Tricks:** use Schur's lemma to express hyperbolic constraints, bounds on inverse matrices and quadratic matrix inequalities as explicit SOCP constraints or LMIs.