Engineering Decision Theory
EN5830

Chathuranga Weeraddana

January 2020

Lecture 9: Convex Optimization Problems

Recall from Lecture 10

- standard formulation of an optimization problem:

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & g_i(x) = 0, \ i = 1, \ldots, p \\ & f_i(x) \leq 0, \ i = 1, \ldots, m \end{array} \tag{1}$$

- decision variable: x
- objective function: $f_0(x)$
- equality constraint: $g_i(x) = 0, \ i = 1, \ldots, p$
- inequality constraint: $f_i(x) \leq 0, \ i = 1, \ldots, m$

Locally Optimal Vs Globally Optimal

- we say $x^\star$ is a globally optimal point, whenever $x^\star \in \mathcal{F}$ and $f_0(x^\star) = p^\star$, where $p^\star = \inf\{f_0(x) \mid x \in \mathcal{F}\}$
- define the set

$$\mathcal{X}_{\text{opt}} = \{x \in \mathcal{F} \mid f_0(x) = p^\star\}$$

- if there exist a globally optimal point for problem (4), we say the optimal value is attained (p 128, [Boyd-2004])
- if $\mathcal{X}_{\text{opt}}$ is empty, we say the optimal value is not attained (p 128, [Boyd-2004])

Locally Optimal Vs Globally Optimal

- we say x_0 is a locally optimal point, whenever $x_0 \in \mathcal{F}$ and

  $$f_0(x_0) = \inf\{f_0(x) \mid x \in \mathcal{F}, ||x - x_0|| \leq R\} \text{ for some } R > 0$$

- i.e., $x_0 \in \mathcal{F}$ minimizes f_0 over some ball around x_0
  (p 128, [Boyd-2004])

- the word 'optimal' is used for 'globally optimal' here onwards

Convex Optimization Problems

- we say a problem in standard form is convex whenever:
  - the objective function f_0 is convex `AND`
  - the inequality constraint functions $f_i(x), i = 1, \ldots, m$ are convex `AND`
  - the equality constraint functions $g_i(x),\ i = 1, \ldots, p$ are affine

Convex Optimization Problems

- i.e., the problem

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & a_i^T x - b_i = 0, \;\; i = 1, \ldots, p \\ & f_i(x) \leq 0, \;\; i = 1, \ldots, m \end{array}$$

  is convex provided the functions $f_i, \;\; i = 0, \ldots, m$ are convex

Feasible Set of a Convex Problem

- it can easily be shown that the feasible set of a convex optimization problem is also convex
  - i.e., $\mathcal{F}$ is convex
- in other words, in the case of convex optimization problems, we are minimizing a convex function over a convex set
- geometrically interpret the idea with level curves and $\mathcal{F}$

Importance of Convexity

- a fundamental property of convex optimization problems:
  - any locally optimal point is also globally optimal
- discuss the local optimal and global optimal points (p 128, [Boyd-2004])
- optimality criterions can be specified (§ 4.2.3, [Boyd-2004])

Importance of Convexity

- by using the optimality criterions, efficient algorithms can be designed to find the optimal point
- e.g.,
  - unconstrained optimization: gradient descent algorithm (§ 9.3, [Boyd-2004]), steepest descent algorithm (§ 9.4, [Boyd-2004]), Newton's method (§ 9.5, [Boyd-2004])
  - equality constrained optimization: Newton's method with equality constraints (§ 10.2, [Boyd-2004])
  - constrained optimization: interior-point method (§ 11, [Boyd-2004])
  - simplex algorithm: a particular one to solve only linear programs (LP) [1]

[1] LPs are discussed soon.

Importance of Convexity

- the algorithms aforementioned are gracefully implemented in many software packages
- a popular one is
  - `CVX`, matlab software for convex optimization
- for a complete list, you may refer to the stanford website, `http://web.stanford.edu/~boyd/software.html`

Equivalent Convex Problems

The Idea of Equivalent Problems

- two problems are equivalent if the solution of one is readily obtained from the solution of the other, and vice-versa
- some commonly used formulations
  - eliminating equality constraints
  - introducing equality constraints
  - introducing slack variables for linear inequalities
  - epigraph form
  - minimizing over some variables

Eliminating Equality Constraints

- the problem

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \leq 0, \ i = 1, \ldots, m \\ & Ax = b \end{array}$$

  with decision variables x is equivalent to

$$\begin{array}{ll} \text{minimize} & f_0(Fz + x_0) \\ \text{subject to} & f_i(Fz + x_0) \leq 0, \ i = 1, \ldots, m \end{array}$$

  with decision variables z

- follows from $\{x \mid Ax = b\} = \{Fz + x_0\}$, see Appendix C.5 [Boyd-2004]

Eliminating Equality Constraints

- the problem

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \leq 0,\ i = 1, \ldots, m \\ & Ax = b \end{array}$$

  with decision variables x is equivalent to

$$\begin{array}{ll} \text{minimize} & f_0(Fz + x_0) \\ \text{subject to} & f_i(Fz + x_0) \leq 0,\ i = 1, \ldots, m \end{array}$$

  with decision variables z

- follows from $\{x \mid Ax = b\} = \{Fz + x_0\}$, see Appendix C.5 [Boyd-2004]

Introducing Equality Constraints

- the problem

$$\begin{array}{ll} \text{minimize} & f_0(A_0x + b_0) \\ \text{subject to} & f_i(A_ix + b_i) \leq 0,\ i = 1, \ldots, m \end{array}$$

with decision variables x is equivalent to

$$\begin{array}{ll} \text{minimize} & f_0(y_0) \\ \text{subject to} & f_i(y_i) \leq 0,\ i = 1, \ldots, m \\ & y_i = A_ix + b_i,\ i = 0, 1, \ldots, m \end{array}$$

with decision variables $(x, y_0, y_1, \ldots, y_m)$

Introducing Slack Variables

- the problem

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & a_i^{\mathsf{T}} x \leq b_i, \ i = 1, \ldots, m \end{array}$$

with decision variables x is equivalent to

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & a_i^{\mathsf{T}} x + s_i = b_i, \ i = 1, \ldots, m \\ & s_i \geq 0, \ i = 1, \ldots, m \end{array}$$

with decision variables $(x, s_1, \ldots, s_m)$

Epigraph Form

- the problem

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \leq 0, \ \ i = 1, \ldots, m \\ & Ax = b \end{array}$$

with decision variables x is equivalent to

$$\begin{array}{ll} \text{minimize} & t \\ \text{subject to} & f_0(x) \leq t \\ & f_i(x) \leq 0, \ \ i = 1, \ldots, m \\ & Ax = b \end{array}$$

with decision variables (x, t)

Minimizing Over Some Variables

- the problem

$$\begin{array}{ll} \text{minimize} & f_0(x,y) \\ \text{subject to} & f_i(x) \leq 0,\ i=1,\ldots,m \\ & Ax = b \end{array}$$

with decision variables (x,y) is equivalent to

$$\begin{array}{ll} \text{minimize} & \tilde{f}_0(x) \\ \text{subject to} & f_i(x) \leq 0,\ i=1,\ldots,m \\ & Ax = b \end{array}$$

with decision variables (x,t)

- where $\tilde{f}_0(x) = \inf_y f_0(x,y)$

A Catalogue of Convex Problems

Linear Programming

- when the objective and constraint functions are all affine, the problem is called a linear program (LP)

- a general linear program has the form

$$\begin{array}{ll} \text{minimize} & c^T x + d \\ \text{subject to} & Gx \preceq h \\ & Ax = b, \end{array} \tag{2}$$

  where $G \in \mathbb{R}^{m \times n}$ and $A \in \mathbb{R}^{p \times n}$

- geometrically interpret, what is an LP (Figure 4.4, p. 147, [Boyd-2004])

Linear Programming

- diet problem, to find cheapest healthy diet
- piecewise-linear minimization
- many others

Quadratic Optimization Problems

- when the objective the objective function is (convex) quadratic and the constraint functions are affine, the problem is called a quadratic program (QP)

- a general QP has the form

$$\begin{array}{ll} \text{minimize} & (1/2)x^T Px + q^T x + r \\ \text{subject to} & Gx \preceq h \\ & Ax = b, \end{array} \tag{3}$$

  where P is symmetric positive semidefinite, $G \in \mathbb{R}^{m \times n}$, and $A \in \mathbb{R}^{p \times n}$

- geometrically interpret, what is a QP (Figure 4.5, p. 153, [Boyd-2004])

Quadratic Programming

- least-squares and constrained least-squares
- linear program with random cost

QCQP

- quadratically constrained quadratic program (QP)
- a general QP has the form

$$\begin{array}{ll} \text{minimize} & (1/2)x^T P_0 x + q_0^T x + r_0 \\ \text{subject to} & (1/2)x^T P_i x + q_i^T x + r_i \leq 0, \ i = 1, \ldots, m \\ & Ax = b, \end{array}$$

  where P_is are symmetric positive semidefinite

- feasible set ?
- geometrically interpret the idea

SOCP

- second order cone program (SOCP)
- the problem has the form

$$\begin{array}{ll} \text{minimize} & f^T x \\ \text{subject to} & ||A_i x + b_i||_2 \leq c_i^{\mathsf{T}} x + d_i, \ i = 1, \ldots, m \\ & Fx = g, \end{array}$$

- inequalities are called second-order cone (SOC) constraints (suppose $A_i \in \mathbb{R}^{n_i \times m}$):

$$(A_i x + b_i, c_i^{\mathsf{T}} x + d_i) \in \text{the second-order cone in } \mathbb{R}^{n_i+1}$$

- more general than LP, QP, and QCQP

SDP

- semidefinite program (SDP)
- the problem has the form

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & x_1 F_1 + x_2 F_2 + \cdots + x_n F_n + G \preceq 0 \\ & Ax = b, \end{array}$$

  with F_i, G symmetric

- inequality constraint is called linear matrix inequality (LMI)
- discuss LMI's, see example 2.10 od [Boyd-2004]
- more general than LP, QP, QCQP, and SOCP

What if Many LMIs

- we can put them all in one LMI
- e.g.,

$$x_1\bar{F}_1 + \cdots + x_n\bar{F}_n + \bar{G} \preceq 0,\ x_1\hat{F}_1 + \cdots + x_n\hat{F}_n + \hat{G} \preceq 0,$$

is equivalent to

$$x_1 \begin{bmatrix} \bar{F}_1 & 0 \\ 0 & \hat{F}_1 \end{bmatrix} + \cdots + x_n \begin{bmatrix} \bar{F}_2 & 0 \\ 0 & \hat{F}_2 \end{bmatrix} + \begin{bmatrix} \bar{G} & 0 \\ 0 & \hat{G} \end{bmatrix} \preceq 0$$

LP as SDP



- the problem

$$\begin{array}{ll} \text{minimize} & c^{\mathsf{T}} x \\ \text{subject to} & Ax \preceq b \end{array}$$

  is equivalent to

$$\begin{array}{ll} \text{minimize} & c^{\mathsf{T}} x \\ \text{subject to} & \mathtt{diag}(Ax - b) \preceq 0 \end{array}$$

SOCP as SDP

- the problem

$$\begin{array}{ll} \text{minimize} & f^{\mathsf{T}}x \\ \text{subject to} & ||A_i x + b_i||_2 \leq c_i^{\mathsf{T}} x + d_i, \ i = 1, \ldots, m \end{array}$$

  is equivalent to

$$\begin{array}{ll} \text{minimize} & f^{\mathsf{T}}x \\ \text{subject to} & \begin{bmatrix} (c_i^{\mathsf{T}} x + d_i) I & (A_i x + b_i) \\ (A_i x + b_i)^{\mathsf{T}} & c_i^{\mathsf{T}} x + d_i \end{bmatrix} \succeq 0 \ i = 1, \ldots, m \end{array}$$

- see appendix A.5.5 of [Boyd-2004] for Schur complement and section 4.4.7 of [Calafiore-2014]

GP

- geometric program (GP)
- monomial function: $f(x) = cx_1^{a_1}x_2^{a_2}\cdots x_n^{a_n}$, $\mathbf{dom}\ \mathbb{R}^n_{++}$
- posynomial function is a sum of monomials:

$$f(x) = \sum_{k=1}^{K} c_k x_1^{a_{1k}} x_2^{a_{2k}} \cdots x_n^{a_{nk}}, \quad \mathbf{dom}\ \mathbb{R}^n_{++}$$

- GP is of the form

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \leq 1 \ i = 1, \ldots, m \\ & h_i(x) = 1, \ i = 1, \ldots, p \end{array}$$

with f_is are posynomials and h_is are monomials

Convex Formulation of GP

- is GP convex in its original form?

Convex Formulation of GP

- is GP convex in its original form? no
- but considered as a convex problem because GP has an equivalent convex reformulation
- to get the convex reformulation
  - change of variables: $y_i = \log x_i$
  - take logarithm of cost and constraints

Convex Formulation of GP

- monomial function $f(x) = cx_1^{a_1} x_2^{a_2} \cdots x_n^{a_n}$, $\mathbf{dom}\ \mathbb{R}^n_{++}$ transforms to

$$\log f(e^{y_1}, \ldots, e^{y_n}) = a^{\mathsf{T}} x + \log c$$

- posynomial function

$$f(x) = \sum_{k=1}^{K} c_k x_1^{a_{1k}} x_2^{a_{2k}} \cdots x_n^{a_{nk}}, \quad \mathbf{dom}\ \mathbb{R}^n_{++}$$

transforms to

$$\log f(e^{y_1}, \ldots, e^{y_n}) = \log \left(\sum_{k=1}^{K} e^{a_k^{\mathsf{T}} y + \log c_k} \right)$$

Convex Formulation of GP

- thus, the convex formulation is simply

$$\begin{array}{ll} \text{minimize} & \log\left(\sum_{k=1}^{K} e^{a_0^{\mathsf{T}} y + \log c_0}\right) \\ \text{subject to} & \log\left(\sum_{k=1}^{K} e^{a_k^{\mathsf{T}} y + \log c_k}\right) \leq 1, \;\; i = 1, \ldots, m \\ & Gy + d = 0, \;\; i = 1, \ldots, p \end{array}$$