Introduction to Optimization

CO 255

Prof. Ricardo Fukasawa

LaTeXed by Sibelius Peng

Preface

Disclaimer Much of the information on this set of notes is transcribed directly/indirectly from the lectures of CO 255 during Winter 2020 as well as other related resources. I do not make any warranties about the completeness, reliability and accuracy of this set of notes. Use at your own risk.

For any questions, send me an email via `https://notes.sibeliusp.com/contact/`.

You can find my notes for other courses on `https://notes.sibeliusp.com/`.

Sibelius Peng

Contents

0

Info

Ricardo: MC 5036. OH: M 1:30 - 3pm
TA: Adam Brown: MC 5462. OH: F 10-11am

Books (not required)

- Intro to Linear Opt. Bertsimas
- Int Programming. Conforti

Grading

- assns: 20% (≈ 5)
- mid: 30% (Feb 11 in class)
- final: 50%

1 Introduction

Given a set S, and a function $f : S \to \mathbb{R}$. An optimization problem is:

$$\begin{aligned} &\max f(x) \\ &\underbrace{s.t.}_{\text{subject to}} \quad x \in S \end{aligned} \qquad \text{(OPT)}$$

- S **feasible region**
- A point $\overline{x} \in S$ is a **feasible solution**
- $f(x)$ is **objective function**

(OPT) means: "Find a feasible solution x^* such that $f(x) \leq f(x^*), \forall x \in S$"

- Such x^* is an **optimal solution**
- $f(x^*)$ is **optimal value**

Other ways to write (OPT):

$$\max\{f(x), x \in S\}$$

$$\max_{x \in S} f(x)$$

Analogous problem

$$\begin{aligned} &\min f(x) \\ &s.t. \quad x \in S \end{aligned}$$

Note

$$\begin{matrix} \max f(x) \\ s.t. \quad x \in S \end{matrix} = -1 \begin{pmatrix} \min -f(x) \\ s.t. \quad x \in S \end{pmatrix}$$

Problem x^* may not exist

a) Problem is unbounded:

$$\forall M \in \mathbb{R}, \exists \overline{x} \in S, \ s.t. \ f(\overline{x}) > M$$

b) $S = \phi$, i.e. (OPT) is **INFEASIBLE**

c) There may not exist x^* achieving supremum.

Example:

$$\begin{array}{ll} \max & x \\ \text{s.t} & x < 1 \end{array}$$

supremum

$$\sup\{f(x) : x \in S\} = \begin{cases} +\infty & \text{if OPT unbounded} \\ -\infty & \text{if } S = \varnothing \\ \min\{x : x \geq f(x), \forall x \in S\} & \text{otherwise} \end{cases}$$

always exist and are well-defined

infimum

$$\inf\{f(x) : x \in S\} = -1 \cdot \sup\{-f(x) : x \in S\}$$

From this point on, we will abuse notation and say $\max\{f(x) : x \in S\}$ is $\sup\{f(x) : x \in S\}$.

One way to specify that I want an opt. sol. (if exists) is

$$x^* \in \operatorname{argmax}\{f(x) : x \in S\}$$

1.1 Linear Optimization (Programming)

or (LP).

$$S = \{x \in \mathbb{R}^n : Ax \leq b\}$$

where $A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m$ and $f(x) = c^T x, c \in \mathbb{R}^n$.

$$\downarrow$$

$$\begin{array}{l} \max c^T x \\ s.t. \ Ax \leq b \end{array} \qquad (LP)$$

Note

$$A = \begin{pmatrix} | & & | \\ A_1 & \cdots & A_n \\ | & & | \end{pmatrix} \qquad A = \begin{pmatrix} - & a_1^T & - \\ & \vdots & \\ - & a_m^T & - \end{pmatrix}$$

Clarifying

$$u, v \in \mathbb{R}^n, \qquad u \leq v \iff u_j \leq v_j, \forall j \in 1, \dots, n$$

Note

$u \not\leq v$ is not the same as $u > v$

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} \not\leq \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Example:

$$\begin{array}{llll} \max & 2x_1 + & 0.5x_2 & \\ s.t. & x_1 & & \leq 2 \\ & x_1 + & x_2 & \leq 2 \\ & x & & \geq 0 \end{array}$$

- Strict ineq. not allowed

halfspace, hyperplane, polyhedron

Let $h \in \mathbb{R}^n, h_0 \in \mathbb{R}$.

$\{x \in \mathbb{R}^n : h^T \leq h_0\}$ is a **halfspace**.

$\{x \in \mathbb{R}^n : h^T = h_0\}$ is a **hyperplane**.

$Ax \leq b$ is a **polyhedron** (i.e. intersection of finitely many halfspaces).

Example:

n products, m resources. Producing $j \in \{1, \dots, n\}$ given c_j profit/unit and consumes a_{ij} units of resource i, $\forall i \in \{1, \dots, m\}$. There are b_i units available $\forall i \in \{1, \dots, m\}$.

$$\begin{aligned} \max & \sum_{j=1}^{n} c_j x_j \\ s.t. & \sum_{j=1}^{n} a_{ij} x_j \leq b_i, \qquad \forall i = 1, \dots, m \\ & x \geq 0 \end{aligned}$$

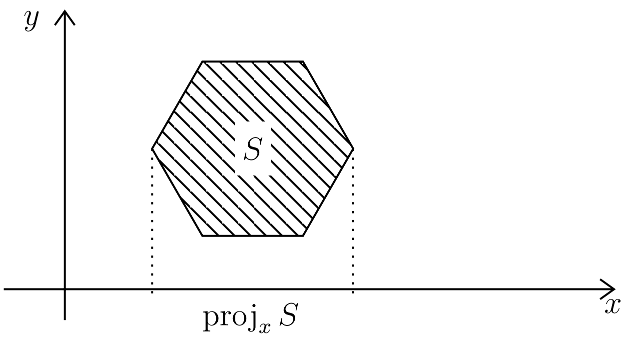
which is an LP.

1.1.1 Determining Feasibility

Given a polyhedron
$$P = \{x \in \mathbb{R}^n : Ax \leq b\}$$
either find $\overline{x} \in P$ or show $P = \varnothing$.

Idea In 1-d, easy. $\rightarrow$ Reduce problem in dimension n to one in dimension $n-1$.

Notation Let $S = \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^p : Ax + Gy \leq b\}$, then
$$\operatorname{proj}_x S := \{x \in \mathbb{R}^n : \exists y \text{ so that } (x, y) \in S\}$$
is the (orthogonal) *projection* if S onto x.



We will find if $P = \varnothing$ by looking at $\operatorname{proj}_{x_1,\ldots,x_{n-1}}$ (P)

1.1.2 Fourier-MotzKin Elimination

Call a_{ij} entries of A. Let

$$\begin{aligned}
M &:= \{1, 2, \ldots, m\} \\
M^+ &:= \{i \in M : a_{in} > 0\} \\
M^- &:= \{i \in M : a_{in} < 0\} \\
M^0 &:= \{i \in M : a_{in} = 0\}
\end{aligned}$$

For $i \in M^+$ (1):

$$a_i^T \leq b_i \iff \sum_{j=1}^{n} a_{ij} x_j \leq b_i \iff \sum_{j=1}^{n-1} \frac{a_{ij}}{a_{in}} x_j + x_n \leq \frac{b_i}{a_{in}}, \qquad \forall i \in M^+$$

For $i \in M^-$ (2):

$$a_i^T x \leq b_i \iff \sum_{j=1}^{n-1} \frac{a_{ij}}{a_{in}} x_j - x_n \leq \frac{b_i}{-a_{in}}, \qquad \forall i \in M^-$$

For $i \in M^0$ (3):

$$a_i^T x \le b_i \iff \sum_{j=1}^{n-1} a_{ij} x_j \le b_i, \qquad \forall i \in M^0$$

$$P = \{x \in \mathbb{R}^n : (1)(2)(3)\}$$

Define (4):

$$\sum_{j=1}^{n-1} \left(\frac{a_{ij}}{a_{in}} - \frac{a_{kj}}{a_{kn}} \right) x_j \le \frac{b_i}{a_{in}} - \frac{b_i}{a_{kn}}, \qquad \forall i \in M^+, \forall k \in M^-$$

Theorem 1.1

$$(\overline{x}_1, \ldots, \overline{x}_{n-1}) \text{ satisfies (3), (4)} \iff \exists \overline{x}_n : (\overline{x}_1, \ldots, \overline{x}_n) \in P$$

Proof:

$\Longleftarrow$ If $(\overline{x}_1, \ldots, \overline{x}_n)$ satisfies (1), (2), (3) then $(\overline{x}_1, \ldots, \overline{x}_{n-1})$ satisfies (3) and adding (1), (2) $\implies (\overline{x}_1, \ldots, \overline{x}_{n-1})$ satisfies (4)

$\Longrightarrow$ If $(\overline{x}_1, \ldots, \overline{x}_{n-1})$ satisfies (4)

$$\sum_{j=1}^{n-1} \frac{a_{ij}}{a_{in}} \overline{x}_j - \frac{b_i}{a_{in}} \le \sum_{j=1}^{n-1} \frac{a_{kj}}{a_{kn}} \overline{x}_j - \frac{b_k}{a_{kn}}, \quad \forall i \in M^+, k \in M^-$$

Let

$$\overline{x}_n := \max_{i \in M^+} \left\{ \sum_{j=1}^{n-1} \frac{a_{ij}}{a_{in}} \overline{x}_j - \frac{b_i}{a_{in}} \right\}$$

$$\implies \sum_{j=1}^{n-1} \frac{a_{ij}}{a_{in}} \overline{x}_j - \frac{b_i}{a_{in}} \le -\overline{x}_n, \quad \forall i \in M^+$$

and

$$-\overline{x}_n \le \sum_{j=1}^{n-1} \frac{a_{kj}}{a_{kn}} \overline{x}_j - \frac{b_k}{a_{kn}}, \qquad \forall k \in M^-$$

$$\implies (\overline{x}_1, \ldots, \overline{x}_n) \in P$$

□

Note

Proof assumes M^+, M^- are nonempty. But statement holds regardless.

(if M^+ or $M^- = \varnothing$ then (4) yields no constraints)

Fourier-MotzKin

- $A^n = A, b^n = b$
- given A^i, b^i obtain A^{i-1}, b^{i-1} (A^{i-1} has one less column than A^i column than A^i) by applying the steps described

  $$P_i := \{x \in \mathbb{R}^i : A^i x \leq b^i\}$$

  then

  $$P_{i-1} = \operatorname{proj}_{x_1,\ldots,x_{i-1}} P_i$$

  and $P_{i-1} = \varnothing \iff P_i = \varnothing$.
- Keep applying projection until $i = 1$.

$$P_0 = \varnothing \iff P_n = P = \varnothing$$

Let

$$P_i^n = P_i \times \mathbb{R}^{n-i} = \{x \in \mathbb{R}^n (A^i, 0)x \leq b^i\}$$

not hard to see $P_i^n = \varnothing \iff P_i = \varnothing$

Notice that

$$P_0 = \varnothing \iff P_0^n = \varnothing, P_0^n = \{0 \leq b^0\}$$

Example:

$$P_2 = \left\{ x \in \mathbb{R}^2 : \begin{array}{rrl} x_1 & +x_2 & \leq 1 \\ -x_1 & & \leq 0 \\ & -x_2 & \leq -2 \\ -3x_1 & -3x_2 & \leq -6 \end{array} \right\}$$

draw the graph, clearly empty

M^+: $\frac{1}{2}x_1 + x_2 \leq \frac{1}{2}$

M^-: $-x_2 \leq -2 \qquad -x_1 - x_2 \leq -2$

M^0: $-x_1 \leq 0$

$$P_1 = \left\{ x_1 \in \mathbb{R} : \begin{array}{rl} -x_1 & \leq 0 \\ \frac{1}{2}x_1 & \leq -\frac{3}{2} \\ -\frac{1}{2}x_1 & \leq -\frac{3}{2} \end{array} \right\}$$

M^+: $x_1 \leq -3$

M^-: $-x_1 \leq 0$ and $-x_1 \leq -3$

$$P_0^2 = \left\{ x \in \mathbb{R}^2 : \begin{array}{c} 0 \leq -3 \\ 0 \leq -6 \end{array} \right\} = \varnothing$$

Here $b^0 = \binom{-3}{-6}$

Remark:

Inequality in P_i^n:

- All inequalities are obtained by a nonnegative combination of inequality in P_{i+1}^n
  $\implies$ all nonnegative combination of inequalities in P.
- If all A, b are rational then so are all A^i, b^i
- If $b = 0, b_i = 0, \forall i$

Theorem 1.2: Farkas' Lemma

$$P = \{x \in \mathbb{R}^n : Ax \leq b\} = \varnothing \iff \exists u \in \mathbb{R}^m : \begin{array}{l} u^T A = 0 \\ u^T b < 0 \\ u \geq 0 \end{array}$$

Proof:

$(\impliedby)$ Suppose $\overline{x}$ satisfies $A\overline{x} \leq b$.

$$0 = u^T A\overline{x} \leq u^T b < 0$$

which is impossible.

$(\implies)$ If $P = \varnothing$. Apply Fourier-Motzkin until we get

$$P_0^n = \varnothing = \{x \in \mathbb{R}^n : 0x \leq b^0\}$$

i.e. there exists j for which $b_j^0 < 0$.

If we look at corresponding constraint in P_0^n is

$$0^T x \leq b_j^0$$

which can be obtained by a vector u such that $u^T A = 0, u^T b = b_j^0, u \geq 0$.

□

Farkas' Lemma (alternate statement)

Exactly one of the following has a solution:

a) $Ax \leq b$

b) $\begin{aligned} u^T A &= 0 \\ u^T b &< 0 \\ u &\geq 0 \end{aligned}$

Farkas' Lemma (Different Form)

Exactly one of the following has a solution:

a) $\begin{aligned} Ax &= b \\ x &\geq 0 \end{aligned}$

b) $\begin{aligned} u^T A &\geq 0 \\ u^T b &< 0 \end{aligned}$

Proof:

(Sketch)

$$P = \left\{ x : \begin{aligned} Ax &= b \\ x &\geq 0 \end{aligned} \right\} = \left\{ x : \underbrace{\begin{pmatrix} A \\ -A \\ -I \end{pmatrix}}_{A'} x \leq \underbrace{\begin{pmatrix} b \\ -b \\ -0 \end{pmatrix}}_{b'} \right\}$$

Apply original Farkas' Lemma to get $P = \varnothing \iff \exists u_1 \in \mathbb{R}^m, u_2 \in \mathbb{R}^m, v \in \mathbb{R}^n$:

$$\begin{aligned} u_1^T A - u_2^T A - v &= 0 \\ u_1^T b - u_2^T b &< 0 \\ u_1, u_2, v &\geq 0 \end{aligned}$$

Let $u = (u_2 - u_2)$

$$u^T A - v = 0 \implies u^T A \geq 0, \quad u^T b < 0$$

$\square$

Consider a linear programming (LP):

$$\begin{aligned} \max \;& c^T x \\ s.t. \;& Ax \leq b \end{aligned} \qquad (LP)$$

Theorem 1.3: Fundamental Theorem of Linear Programming

(LP) has exactly one of 3 outcomes:

a) Infeasible

b) Unbounded

c) There exists an optimal solution.

Proof:

Let's assume a), b) don't hold.

If $n = 1$, then (LP) has an optimal solution. (Why?)

Else, define

$$\begin{aligned} &\max z \\ &s.t. \quad z - c^T x \leq 0 \qquad (LP') \\ &\qquad Ax \leq b \end{aligned}$$

(LP') is also not in case a) or b). (Why?)

Also if (x^*, z^*) is an optimal solution to (LP'), then x^* is an optimal solution to (LP). (Why?)

Apply Fourier-Motzkin to

$$\left\{ (x, z) : \begin{aligned} z - c^T x &\leq 0 \\ Ax &\leq b \end{aligned} \right\}$$

Until we are left with a polyhedron

$$\{z \in \mathbb{R} : A'z \leq b'\}$$

Now $\max z \qquad s.t \quad A'z \leq b'$ is not cases a) or b). (Why?)

$\rightarrow$ can get an optimal solution z^* to such problem. Apply Fourier-Motzkin back to get (x^*, z^*) optimal solution to (LP'). (Why?) $\square$

1.2 Certifying Optimality

$$\begin{aligned} &\max \ c^T x \\ &s.t \quad Ax \leq b \end{aligned} \qquad (LP)$$

and let $\overline{x} \in P = \{x : Ax \leq b\}$

Question Can we certify that $\overline{x}$ is optimal?

Example:

$$\begin{aligned} &\max\ 2x_1 + x_2 \\ &\quad x_1 + 2x_2 \leq 2 \\ s.t.\ &\quad x_1 + x_2 \leq 2 \\ &\quad x_1 - x_2 \leq 0.5 \end{aligned}$$

Consider $\overline{x} = (0,1)^T$ is clearly NOT optimal.

$x^* = (1, 0.5)^T$ and $c^T x^* = 2.5$. Any feasible solution satisfies

$$\begin{array}{rll} x_1 + 2x_2 & \leq 2 & \times 1/3 \\ x_1 + x_2 & \leq 2 & \times 1 \\ +\ \ x_1 - x_2 & \leq 0.5 & \times 2/3 \\ \hline 2x_1 + x_2 & \leq 3 & \end{array}$$

Instead do $1 \times 1st$ constraint $+ 1 \times 3rd$ constraint $\implies 2x_1 + x_2 \leq 2.5$

In general:

$$\begin{array}{rll} x_1 + 2x_2 & \leq 2 & \times y_1 \\ x_1 + x_2 & \leq 2 & \times y_2 \\ +\ \ x_1 - x_2 & \leq 0.5 & \times y_3 \\ \hline \end{array}$$
$$(y_1 + y_2 + y_3)x_1 + (2y_1 + y_2 - y_3)x_2 \leq 2y_1 + 2y_2 + 0.5y_3$$

As long as $y_1, y_2, y_3 \geq 0$ and

$$\begin{aligned} y_1 + y_2 + y_3 &= 2 \\ 2y_1 + y_2 - y_3 &= 1 \end{aligned}$$

This leads to the following linear program:

$$\begin{aligned} &\min\ 2y_1 + 2y_2 + 0.5y_3 \\ &\quad y_1 + y_2 + y_3 = 2 \\ s.t.\ &\quad 2y_1 + y_2 - y_3 = 1 \\ &\quad y_1, y_2, y_3 \geq 0 \end{aligned}$$

This is called the dual LP.

In general:

$$\begin{aligned} &\max\ c^T x \\ &s.t.\ \ Ax \leq b \end{aligned} \qquad (LP)$$

Dual of (LP)

$$\begin{aligned} &\min\ b^T y \\ s.t.\ &\quad y^T A = c^T \\ &\quad y \geq 0 \end{aligned} \qquad (D)$$

Theorem 1.4: Weak Duality

Let $\overline{x}$ feasible for (LP), $\overline{y}$ feasible for (D). Then $c^T x \leq b^T y$.

Proof:

$$c^T\overline{x} = \overline{y}^T(A\overline{x}) \leq \overline{y}^T b$$

where we used $A\overline{x} \leq b$ and $\overline{y} \geq 0$. $\square$

Theorem 1.5: Strong Duality

x^* is optimal for (LP) $\iff$ $\exists y^*$ feasible for (D) such that

$$c^T x^* = b^T y^*$$

Index