Introduction to Optimization

CO 255

Prof. Ricardo Fukasawa

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

P	refac	e	1
0	Info		3
1	Intr	roduction	4
2	Line	ear Optimization	6
	2.1	Determining Feasibility	7
	2.2	Fourier-Motzkin Elimination	8
	2.3	Certifying Optimality	13
	2.4	Possible Outcomes	17
	2.5	Duals of generic LPs	17
		2.5.1 Cheat Sheet	18

0

Info

Ricardo: MC 5036. OH: M $1{:}30$ - $3\mathrm{pm}$

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%

Introduction

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

$$\max_{s.t.} f(x)$$
subject to (OPT)

- \bullet 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

$$\min f(x)$$

$$s.t. \ x \in S$$

Note

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

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.



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

supremum

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

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

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

2

Linear Optimization (Programming) (LP)

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

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

$$\max_{x} c^T x$$

$$s.t. \ Ax \le b$$
(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

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

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:

$$\max 2x_1 + 0.5x_2$$
s.t.
$$x_1 \leq 2$$

$$x_1 + x_2 \leq 2$$

$$x \geq 0$$

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

$$\max \sum_{j=1}^{n} c_{j} x_{j}$$

$$s.t. \sum_{j=1}^{n} a_{ij} x_{j} \leq b_{i}, \quad \forall i = 1, \dots, m$$

$$x > 0$$

which is an LP.

2.1 Determining Feasibility

Given a polyhedron

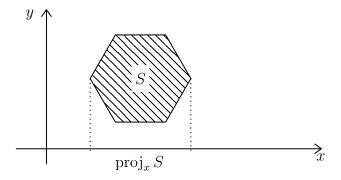
$$P = \{ x \in \mathbb{R}^n : Ax < b \}$$

either find $\overline{x} \in P$ or show $P = \emptyset$.

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 $P = \emptyset$ by looking at $\operatorname{proj}_{x_1,\dots,x_{n-1}}$ (P)

2.2 Fourier-Motzkin Elimination

Call a_{ij} entries of A. Let

$$M := \{1, 2, \dots, m\}$$

$$M^+ := \{i \in M : a_{in} > 0\}$$

$$M^- := \{i \in M : a_{in} < 0\}$$

$$M^0 := \{i \in M : a_{in} = 0\}$$

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

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

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

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

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

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

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

Define (4):

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

Theorem 2.1

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

Proof:

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

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

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

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

and

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

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

Note

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

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

Fourier-MotzKin

- \bullet $A^n = A, b^n = b$
- 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

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

then

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

and $P_{i-1} = \emptyset \iff P_i = \emptyset$.

• Keep applying projection until i = 1.

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

Let

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

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

Notice that

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

$$P_2 = \begin{cases} x_1 & +x_2 & \le 1 \\ x \in \mathbb{R}^2 : & -x_1 & \le 0 \\ & -x_2 & \le -2 \\ & -3x_1 & -3x_2 & \le -6 \end{cases}$$

draw the graph, clearly empty
$$M^{+} \colon \frac{1}{2}x_{1} + x_{2} \leq \frac{1}{2}$$

$$M^{-} \colon -x_{2} \leq -2 \qquad -x_{1} - x_{2} \leq -2$$

$$M^{0} \colon -x_{1} \leq 0$$

$$P_{1} = \begin{cases} -x_{1} & \leq 0 \\ x_{1} \in \mathbb{R} : \frac{1}{2}x_{1} & \leq -\frac{3}{2} \\ -\frac{1}{2}x_{1} & \leq -\frac{3}{2} \end{cases}$$

$$M^{+} \colon x_{1} \leq -3$$

$$M^{-} \colon -x_{1} \leq 0 \text{ and } -x_{1} \leq -3$$

$$P_{0}^{2} = \begin{cases} x \in \mathbb{R}^{2} : & 0 \leq -3 \\ 0 \leq -6 \end{cases} = \emptyset$$

$$P_0^2 = \left\{ x \in \mathbb{R}^2 : \quad \begin{array}{l} 0 \le -3 \\ 0 \le -6 \end{array} \right\} = \emptyset$$

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

Remark:

Inequality in P_i^n :

- All inequalities are obtained by a nonnegative combination of inequality in
 - 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 2.2: Farkas' Lemma

$$u^{T}A = 0$$

$$P = \{x \in \mathbb{R}^{n} : Ax \le b\} = \emptyset \iff \exists u \in \mathbb{R}^{m} : u^{T}b < 0$$

$$u \ge 0$$

Proof:

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

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

which is impossible.

which is impossible. $\Longrightarrow) \ \mbox{If} \ P=\varnothing. \ \mbox{Apply Fourier-Motzkin until we get}$

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

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

If we look at corresponding constraint in \mathbb{P}_0^n is

$$0^T x \leq b_i^0$$

which can be obtained by a vector u such that $u^TA=0, u^Tb=b_j^0, u\geq 0.$

Farkas' Lemma (alternate statement)

Exactly one of the following has a solution:

a)
$$Ax \leq b$$

$$u^T A = 0$$

b)
$$u^T b < 0$$

$$u \ge 0$$

Farkas' Lemma (Different Form)

Exactly one of the following has a solution:

$$Ax = b$$
a) $x > 0$

b)
$$u^T A \ge 0$$



(Sketch)

$$P = \left\{ x : Ax = b \\ x \ge 0 \right\} = \left\{ x : \underbrace{\begin{pmatrix} A \\ -A \\ -I \end{pmatrix}}_{A'} x \le \underbrace{\begin{pmatrix} b \\ -b \\ -0 \end{pmatrix}}_{b'} \right\}$$

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

$$u_1^T A - u_2^T A - v = 0$$
$$u_1^T b - u_2^T b < 0$$
$$u_1, u_2, v > 0$$

$$u_1, u_2, v \ge 0$$
 Let $u = (u_2 - u_2)$
$$u^T A - v = 0 \implies u^T A \ge 0, \quad u^T b < 0$$

Consider a linear programming (LP):

$$\max_{s.t.} c^T x$$

$$s.t. Ax < b$$
 (LP)

Theorem 2.3: Fundamental Theorem of Linear Programming

- (LP) has exactly one of 3 outcomes:
- a) Infeasible
- b) Unbounded
- c) There exists an optimal solution.



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

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

Else, define

$$\max z$$

$$s.t. \ z - c^T x \le 0 \qquad (LP')$$

$$Ax \le b$$

(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

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

Until we are left with a polyhedron

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

Now max z s.t $A'z \le b'$ 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?)

2.3 Certifying Optimality

$$\max_{s.t} c^T x \\ s.t \quad Ax \le b$$
 (LP)

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

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

Example:

$$\max 2x_1 + x_2$$

$$x_1 + 2x_2 \le 2$$
s.t. $x_1 + x_2 \le 2$

$$x_1 - x_2 \le 0.5$$

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

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

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

In general:

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

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

$$y_1 + y_2 + y_3 = 2$$
$$2y_1 + y_2 - y_3 = 1$$

This leads to the following linear program:

min
$$2y_1 + 2y_2 + 0.5y_3$$

 $y_1 + y_2 + y_3 = 2$
 $s.t.$ $2y_1 + y_2 - y_3 = 1$
 $y_1, y_2, y_3 \ge 0$

This is called the dual LP.

In general:

$$\max_{s.t.} c^T x$$

$$s.t. Ax \le b$$

$$(P)$$

Dual of (P)

Remark:

We call (P) primal LP.

Theorem 2.4: Weak Duality

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

Proof:

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

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

Corrollary 2.5

Several results:

- If (P) is unbounded then (D) is infeasible.
- If (D) is unbounded then (P) is infeasible.

Note

- (P) and (D) can both be infeasible.
 - If \overline{x} is feasible for (P) \overline{y} feasible for (D) $c^T\overline{x} = b^T\overline{y}$, then \overline{x} optimal for (P), \overline{y} optimal for (D).

Theorem 2.6: Strong Duality

 x^* is optimal for (P) $\iff \exists y^*$ feasible for (D) such that $c^T x^* = b^T y^*$.

Proof:

$$(\iff) \checkmark$$
 $(\implies) \text{ Is (D) infeasible?}$

$$\text{Suppose } \left\{ y \in \mathbb{R}^n : \frac{A^T y = c}{y \ge 0} \right\} = \varnothing$$

(Alternate version of Farkas' Lemma) $\exists u: u^T A \geq 0 \iff \exists d: Ad \leq 0$ $c^T d > 0$

Take look at $x' = x^* + d$, then

$$Ax' = Ax^* + Ad \le b$$
$$c^T x' = c^T x^* + c^T d > c^T x^*$$

Contradiction. Thus (D) has an optimal solution y^* .

Now let
$$\gamma = b^T y^*$$
, and let $\theta := \left\{ x \in \mathbb{R}^n : \begin{matrix} Ax \leq b \\ -c^T x \leq -\gamma \end{matrix} \right\}$.

If $\theta = \emptyset$, by Farkas'

$$\exists \left(\overline{\overline{y}} \right) : \begin{cases} \left(\overline{\overline{y}} \right)^T \begin{pmatrix} A \\ -c^T \end{pmatrix} = 0 \\ \left(\overline{\overline{y}} \right)^T \begin{pmatrix} b \\ -\gamma \end{pmatrix} < 0 & \iff \begin{matrix} A^T \overline{y} = c\overline{\lambda} \\ \overline{y} < \gamma \overline{\lambda} \\ \overline{y} \ge 0 \\ \overline{\lambda} \ge 0 \end{cases}$$

Case 1: $\overline{\lambda} > 0$.

Let $y' = \frac{\overline{y}}{\overline{\lambda}}$. Then we have

$$A^T y' = A^T \frac{\overline{y}}{\overline{\lambda}} = c$$
 and $b^T y' = b^T \frac{\overline{y}}{\overline{\lambda}} < \gamma$ and $y' = \frac{\overline{y}}{\overline{\lambda}} \ge 0$

Contradicts optimality of y^* .

$$A^T y = 0$$

Case 2: $\overline{\lambda} = 0$. Then $b^T y < 0$

$$\overline{y} \ge 0$$

Now we can do the same thing previously. Let $y' = y^* + \overline{y}$, then

$$A^T y' = A^T y^* + A^T \overline{y} = c$$

and

$$y' = y^* + \overline{y} \ge 0$$
$$b^T y' = b^T y^* + b^T \overline{y} < b^T y^*$$

Contradicts optimality of y^* .

Thus $\theta \neq \emptyset$.

Let $\overline{x} \in \theta$,

$$c^T x^* \underbrace{\leq}_{\text{weak duality}} b^T y^* = \gamma \underbrace{\leq}_{\overline{x} \in \theta} c^T \overline{x} \leq c^T x^*$$

where the last inequality is because \overline{x} feasible for (P), x^* optimal for (P).

2.4 Possible Outcomes

See here.

2.5 Duals of generic LPs

$$\max (2,3,-4)x
\begin{pmatrix}
1 & 0 & 7 \\
0 & -2 & 1 \\
1 & 0 & 1 \\
-1 & 0 & -1 \\
0 & 1 & 0 \\
-1 & 0 & 0 \\
0 & 1 & 0
\end{pmatrix} x \le \begin{pmatrix}
5 \\
-3 \\
8 \\
-8 \\
6 \\
0 \\
0
\end{pmatrix}$$

and dual

min
$$(5, -3, 8, -8, 6, 0, 0)y$$

s.t $\begin{pmatrix} 1 & 0 & 1 & -1 & 0 & -1 & 0 \\ 0 & -2 & 0 & 0 & 1 & 0 & 1 \\ 7 & 1 & 1 & -1 & 0 & 0 & 0 \end{pmatrix} y = \begin{pmatrix} 2 \\ 3 \\ -4 \end{pmatrix}$ and $y \ge 0$ (D_1)

min
$$(5, -3, 8, -8, 6)y$$

s.t $\begin{pmatrix} 1 & 0 & 1 & -1 & 0 \\ 0 & -2 & 0 & 0 & 1 \\ 7 & 1 & 1 & -1 & 0 \end{pmatrix} y \stackrel{\geq}{\leq} \begin{pmatrix} 2 \\ 3 \\ -4 \end{pmatrix}$ and $y \geq 0$ (D_2)

Claim (y_1^*, \ldots, y_5^*) is optimal for $(D_2) \iff (y_1^*, \ldots, y_5^*, y_6^*, y_7^*)$ optimal for (D_1) with

$$y_6^* = y_1^* + y_3^* - y_4^* - 2$$

$$y_7^* = 3 - (-2y_2^* + y_5^*)$$

min
$$(5,3,8,6)y$$

s.t $\begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 2 & 0 & 1 \\ 7 & -1 & 1 & 0 \end{pmatrix} y \stackrel{\geq}{\leq} \begin{pmatrix} 2 \\ 3 \\ -4 \end{pmatrix}$ and $y_1 \geq 0, y_2 \leq 0$ $y_4 \geq 0$ (D_3)

Claim Opt value of (D_2) and (D_3) are same.

In general

2.5.1 Cheat Sheet

Here or

Primal (m	ax)	Dual (min)	
Constraint	\\ \\	≥ 0 ≤ 0 free	Variable
Variable	≥ ≤ free		Constraint

Remark:

This is not symmetric... The way you can remember it is by thinking natural variables in real life, like you cannot have negative number of cars and so on...

Index

Н	S
halfspace, hyperplane, polyhedron 7	supremum 5
infimum 5	