# AMS 553.361 - Introduction to Optimization

# Lectures by Donniell Fishkind Notes by Kaushik Srinivasan

### Johns Hopkins University Fall 2018

Lecture 1 (2018-08-30)	1	Lecture 4 (2018-09-07)	
Lecture 2 (2018-08-31)	3	Lecture 5 (2018-09-12)	8
Lecture 3 (2018-09-05)	4	Lecture 6 (2018-09-14)	10

### Introduction

Math 110.304 is one of the semi-important courses that is required/recommended for the engineering-based majors at Johns Hopkins University.

These notes are being live-TeXed, through I edot for Typos and add diagrams requiring the TikZ package separately. I am using Texpad on Mac OS X.

I would like to thank Zev Chonoles from The University of Chicago and Max Wang from Harvard University for providing me with the inspiration to start live-TeXing my notes. They also provided me the starting template for this, which can be found on their personal websites.

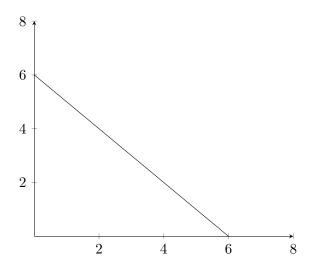
Please email any corrections or suggestions to ksriniv4@jhu.edu.

# Lecture 1 (2018-08-30)

### Introduction

**Example 1.** Up to 6 units of two nutrients can be added to solution and we require the number of units of nutrient 2 has to be at least (natural) logarithm of # units of nutrient one. Goal: choose  $x_1$  = number units of nutrient 1, choose  $x_2$  = number units of nutrient 2. To maximize expected height of plant  $1 + x_1^2(x_2 - 1)^3 e^{-x_1 - x_2}$ 

Maximise 
$$1 + x_1^2(x_2 - 1)^3 e^{-x_1 - x_2}$$
  
 $x_1 + x_2 \le 6$   
 $x_2 > \log x_1$   
 $x_1 \ge 0$   
 $x_2 \ge 0$ 



generic optimization problem: say  $\subseteq \mathbb{R}$ 

$$f: \underbrace{\mathbb{S}}_{\text{feasible region}} \to \underbrace{\mathbb{R}}_{\text{objective function}}$$

$$x^* \begin{bmatrix} 2 \\ 4 \end{bmatrix}$$
 optimal

however optimal objective function value is 1.2677

**Definition 1.** We say  $x^*$  is an optimal solution if

- $x^* \in \mathbb{S}$
- For any  $y \in \mathbb{S}$

**Example 2.** Find min  $\log x$  s.t.  $-\infty \le x \le 7 \to \text{unbounded}$ , has no solution.

**Example 3.** Find min  $\log x$  s.t.  $1 < x \le 7 \to \text{bounded}$ , but is also no solution. (as we can go 1.000001)

**Example 4.** Find min  $\log x$  s.t. x > 1,  $x \le 0.5 \rightarrow$  infeasible!!!

**Example 5.** minimize  $3 + (x - 1)^2$  s.t.  $1 \le x \le 3 \to$  feasible as optimal solution  $x^* = 2$  is an interior part of feasible region.

$$f(x) = 3 + (x - 1)^2$$
$$f'(x) = 0$$
$$f''(x) > 0$$

**Example 6.** Minimize  $3 + (x - 2)^2$  s.t.  $x \ge 10 \to \text{optimal solution is } x^* = 10$ . but  $f'(x) \ne 0$ . But is not an interior point - it is a boundary point of feasible region.

# Lecture 2 (2018-08-31)

**Definition 1.**  $\forall x \in \mathbb{R}^n$  Euclidian length of x is  $||x|| = \left(\sum_{i=1}^n x_i^2\right)^{\frac{1}{2}}$ 

 $\forall x, y \in \mathbb{R}^n$  Euclidian distance from x to y is ||x - y||

**Definition 2.**  $\forall S \subseteq \mathbb{R}^n$  point  $x \in S$  is an interior point of S if J a neighborhood of x which is a subset of S point  $S \in \mathbb{R}^n$  is a boundary point of S if every neighborhood of x contains a point in S and a point not in S

- Set  $S \subseteq \mathbb{R}^n$  is open if every point in S is an interior point. example: open ball.
- Set  $S \subseteq \mathbb{R}^n$  is <u>closed</u> if S contains all boundary points of S.

*Note:*  $\forall S \subseteq \mathbb{R}^n$ , S is open **iff**  $S^c$  is closed.

In (P) min f(x) s.t.  $x \in S$ , suppose  $x^* \in S$ :

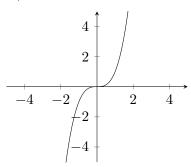
- $x^*$  is a global maximizer if  $\forall y \in S, f(x^*) \leq f(y)$
- $x^*$  is a strict global minimizer if  $\forall y \in S$ , s.t.  $y \neq x^*$ ,  $f(x^*) < f(y)$
- $x^*$  is a <u>local minimizer</u> if  $\exists$  any neighborhood N of x s.t.  $\forall y \in N \cap S, f(x^*) \leq f(y)$
- $x^*$  is a <u>strict</u> local minimizer if  $\exists$  any neighborhood N of x s.t.  $\forall y \in (S \cap N) \underbrace{\setminus x^*}_{\text{besides}}$ ,  $f(x^*) < f(y)$

*Note:* If  $S \subseteq \mathbb{R}^1$ ,  $x^*$  is interior of S, f is suitably differentiable at  $x^*$ .

If  $x^*$  is a local minimizer  $\Rightarrow f'(x^*) = 0$ 

 $x^*$  is a local maximizer  $\Rightarrow f'(x^*) = 0$ 

**BUT** 
$$f'(x^*) = 0 \implies x^* \text{ local min/max}$$



BUT

$$f'(x^*) = 0\&f''(x^*) > 0 \Longrightarrow x^*$$
 strict local max  $f'(x^*) = 0\&f''(x^*) < 0 \Longrightarrow x^*$  strict local min

If  $S \subseteq \mathbb{R}^n$ ,  $x^*$  interior part of S, f strictly differentiable,  $x^*$  local min/max  $\Longrightarrow \nabla f(x^*) = \vec{0}$ 

$$\nabla f(x^*) = \vec{0}$$
 & [?]  $\Rightarrow x^*$  strict local min/max

# Lecture 3 (2018-09-05)

# Linear Programming

Example 1. Diet Problem: You will pick levels of four ingredients for chicken feed.

•  $x_1$  = units of ingredient 1

•  $x_3$  = units of ingredient 3

•  $x_2$  = units of ingredient 2

•  $x_4$  = units of ingredient 4

All are real numbers, so fractions allowed. Given minimum levels of 3 nutrients  $\begin{bmatrix} 6.2\\11.9\\10 \end{bmatrix} = \begin{bmatrix} \text{nutrient 1}\\ \text{nutrient 2}\\ \text{nutrient 3} \end{bmatrix}$ 

Given how many units of nutrient per unit of ingredient

nutrients \ ingredients	1	2	3	4
1 (protein)	1.2	2.6	0	9.2
2 (carbs)	3.9	1	.8	2
3 (cholesterol)	6	0	4	3.1

**Problem**: Find yand of ingredients that meet nutritional requirements cheaply as possible.

Minimum Cost of the ingredients  $\begin{bmatrix} 6.2 \\ 2 \\ 1.6 \\ 3.2 \end{bmatrix}$ 

minimize 
$$6.2x_1 + 2x_2 + 1.6x_3 + 3.2x_4$$

$$\begin{cases}
1.2x_1 + 2.6x_2 + 0x_3 + 9.2x_4 \ge 6.2 \\
3.9x_1 + x_2 + 0.8x_3 + 2x_4 \ge 11.9 \\
3.9x_1 + 1x_2 + 0.8x_3 + 2x_4 \ge 11.9 \\
6x_1 + 0x_2 + 4x_3 + 3.1x_4 \ge 11.9
\end{cases}$$
 and  $x_1, x_2, x_3, x_4 \ge 0$ 

What makes a ingredient list special?

- 1. proportionality
- 2. additivity
- 3. divisibility
  - fractions are allowed referring to  $x_1, x_2, x_3$

(Linear Programming) in canonical form, an example

minimize 
$$c_1x_1 + c_2x_2 + \dots + c_nx_n$$
  
s.t. $a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \ge b_1$   
 $a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n \ge b_2$   
 $\vdots$   
 $a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n \ge b_m$ 

where 
$$x_1, x_2, \cdots, x_n \ge 0$$

$$\updownarrow$$

$$\min \quad c^T x \quad \text{s.t.} \quad A\vec{x} \ge \vec{b}, x \ge \vec{0}$$

$$A \in \mathbb{R}^{mxn}, b \in \mathbb{R}^m, C \in \mathbb{R}^n, x \in \mathbb{R}^n$$

**Example 2.** Transportation problem: There are 3 electricity generation plants  $\alpha, \beta, \gamma$  and two cities u, v.  $\alpha, \beta, \gamma$  respectively produce 65.2, 98.6, 32.5 units of electricity. u, v respectively use 86.2, 110.1 units of electricity

Question: How to supply transportation as cheaply as possible?

generator to city	cost variable	cost
$\alpha \to u$	$x_1$	31.7
$\alpha \to v$	$x_2$	28.6
$\beta \to u$	$\begin{vmatrix} x_3 \end{vmatrix}$	17.6
$\beta \to v$	$x_4$	37.4
$\gamma \to u$	$x_5$	22.8
$\gamma \to v$	$x_6$	29.7

min 
$$31.7x_1 + 28.6x_2 + 17.6x_3 + 374x_4 + 22.8x_5 + 29.7x_6$$

$$\left. \begin{array}{ll} x_1+x_2 \leq 65.2 & \alpha \\ x_3+x_4 \leq 98.6 & \beta \\ x_5+x_6 \leq 32.5 & \gamma \end{array} \right\} \text{ electricity produced at } \alpha,\beta,\gamma \\ u & x_1+x_3+x_5 \geq 86.2 \\ v & x_2+x_4+x_6 \geq 110.1 \end{array} \right\} \text{ Need for electricity for } u,v$$

#### Converting one form of (LP) to another

1. **(LP)** in standard form minimize  $c^T x$  s.t. Ax = b,  $x \ge \vec{0}$  Here

$$A = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \end{bmatrix} \qquad b = \begin{bmatrix} 65.2 \\ 98.6 \\ 32.5 \\ 86.2 \\ 110.1 \end{bmatrix}$$

$$c^{T} = \begin{bmatrix} 31.7 & 28.6 & 17.6 & 37.4 & 22.8 & 29.7 \end{bmatrix}$$

2.  $\max \rightarrow \min$  or  $\min \rightarrow \max \implies \max 6x_1 + 3x_2 - 4x_3 \leftrightarrow \min -6x_1 - 3x_2 + 4x_3$ 

$$Ax \ge b \to \le$$

$$Ax \ge b \leftrightarrow Ax \le -b$$

3. Standard form  $\rightarrow$  canonical form

$$\begin{vmatrix}
\min c^T x \\
\text{s.t. } Ax = b \\
x \ge 0
\end{vmatrix}
\leftrightarrow
\begin{cases}
\min c^T x \\
\text{s.t. } Ax \ge b \\
Ax \le b \\
x \ge \vec{0}
\end{cases}
\leftrightarrow
\begin{cases}
\min c^T x \\
\text{s.t. } Ax \ge b \\
-Ax \ge -b \\
x \ge \vec{0}
\end{cases}
\leftrightarrow
\begin{cases}
\min c^T x \\
Ax \le b \\
-Ax \ge -b \\
x \ge \vec{0}
\end{cases}$$

4. Canoconical form  $\rightarrow$  Standard form

# Lecture 4 (2018-09-07)

5. non-negativity  $\rightarrow$  include in matrix

$$\left. \begin{array}{l} Ax \ge b \\ x \ge \vec{0} \end{array} \right\} \longrightarrow \left[ \begin{matrix} A \\ I \end{matrix} \right] x \ge \left[ \begin{matrix} \vec{b} \\ \vec{0} \end{matrix} \right]$$

Simplex algorithm only works with non-negative elements.

6. unconstrained sign  $\rightarrow$  non-negativity by substitution

$$\underbrace{Z_{\text{uncons}}} := \underbrace{Z_1}_{\geq 0} - \underbrace{Z_2}_{\geq 0}$$

$$min \quad 5x_1 + 6x_2$$

$$s.t. \quad 2x_1 - 3x_2 \geq 9$$

$$x_1 + x_2 \geq -8$$

$$x_1 \geq 0 \quad x_2 \text{unconst}$$

$$\Rightarrow \left\{ X_2 = \underbrace{X_2'}_{\geq 0} - \underbrace{X_2''}_{\geq 0} \right\} \rightarrow \begin{cases} min \quad 5x_1 + 6x_2' - 6x_2'' \\ s.t. \quad 2x_1 - 3x_2' + 3x_2'' \geq 9 \\ x_1 + x_2' - x_2'' \geq -8 \\ x_1, x_2', x_2'' \geq 0 \end{cases}$$

7. Have constraints Ax = b

illustration 
$$x_1, x_2, x_3$$
 
$$\begin{cases} 2x_1 - 2x_2 + 6x_3 = 8\\ 3x_1 + x_2 - x_3 = 2 \end{cases}$$

Unchanged by row opeartions

- swap rows
- multiply by non-zero constant
- add one row to another

$$\left[\begin{array}{cc|cc} 1 & -1 & 3 & 4 \\ 3 & 1 & -1 & 2 \end{array}\right] \xrightarrow{\text{rref}} \left[\begin{array}{cc|cc} 1 & 0 & \frac{1}{2} & -\frac{3}{2} \\ 0 & 1 & -\frac{5}{2} & -\frac{5}{2} \end{array}\right]$$

If row reduce  $[A' \mid b] \rightarrow [A' \mid b']$  then  $\exists$  invertible matrix C

$$C \left[ A \mid b \right] \rightarrow \left[ A' \mid b' \right]$$

**Converse:** if D is invertible, then  $[A' \mid b']$  can be reduced to  $D[A \mid b]$ 

8. Changing the order of columns in Ax = b, it doesn't matter as long as you change the order of the variables conformally

$$\min \begin{bmatrix} 17 \\ 83 \\ -11 \end{bmatrix}^T \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \\
s.t. \begin{bmatrix} 2 & -2 & 6 \\ 3 & 1 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 8 \\ 2 \end{bmatrix} \right\} \longrightarrow \begin{cases} \min \begin{bmatrix} 83 \\ -11 \\ 17 \end{bmatrix}^T \begin{bmatrix} x_2 \\ x_3 \\ x_1 \end{bmatrix} \\
s.t. \begin{bmatrix} -2 & 6 & 3 \\ 1 & -1 & 3 \end{bmatrix} \begin{bmatrix} x_2 \\ x_3 \\ x_1 \end{bmatrix} = \begin{bmatrix} 8 \\ 2 \end{bmatrix}$$

### Lecture 5 (2018-09-12)

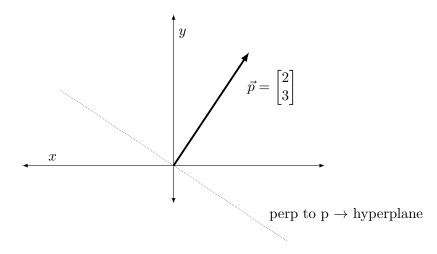
### Half spaces, Hyperplanes, Polyhedral sets

**Recall:** Suppose  $p, x \in \mathbb{R}^n$ . The inner product of p, x

$$p^{T}x = ||p|| \, ||x|| \cos \theta$$
thus  $p^{T}x \begin{cases} > 0 & \text{if } \theta \text{ acute} \\ = 0 & \text{if } p \perp x \text{ (perpendicular)} \\ < 0 & \text{if } \theta \text{ obtuse} \end{cases}$ 

**Definition 1.** Suppose  $p \in \mathbb{R}^n$  non-zero.  $\{\vec{x} \in \mathbb{R}^n : p^T x = 0\}$  is a hyperplane through origin with normal vector p.  $\{\vec{x} \in \mathbb{R}^n : p^T x \leq 0\}$  is (associated) and (closed) half space.

### Example



Everything on either side of the hyperplane is the half spaces.

**Definition 2.** In general hyperplanes in  $\mathbb{R}^n$  are sets  $\{x \in \mathbb{R}^n, p^T x = \alpha\}$  for non-zero  $p \in \mathbb{R}^n$ , and  $\alpha \in \mathbb{R}$ 

**Definition 3.** half spaces in  $\mathbb{R}^n$  are sets  $\{x \in \mathbb{R}^n, p^T x \ge \alpha\}$  for non-zero  $p \in \mathbb{R}^n$ , and  $\alpha \in \mathbb{R}$  **Example 1.** 

(2D) 
$$2x_1 + 5x_2 = 6$$
 hyperplane  $2x_1 + 5x_2 \ge 6$  half space (3D)  $3x_1 + 2x_2 - 7x_3 = 8$  hyperplane  $3x_1 + 2x_2 - 7x_3 \ge 8$  half space

**Definition 4.** A polyhedron (polyhedral space) is the intersection of finely may half spaces.

**Example 2.**  $x \cdot Ax \geq b$  for  $A \in \mathbb{R}^{mxn}$ ,  $b \in \mathbb{B}^m$ 

 $\vec{x}$  satisfies  $Ax \geq b$  iff  $\vec{x}$  satisfies  $p_1^T x \geq b_1$ , and  $\vec{x}$  satisfies  $p_2^T x \geq b_2, \ldots, p_m^T x \geq b_m$ 

### Example 3.

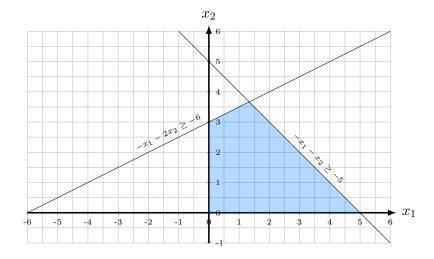
$$\begin{bmatrix} 1 & -2 \\ -1 & -1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \ge \begin{bmatrix} -6 \\ -5 \\ 0 \\ 0 \end{bmatrix}$$

1. 
$$x_1 - 2x_2 \ge -6$$

3. 
$$x_1 \ge 0$$

2. 
$$-x_1 - x_2 \ge 5$$

4. 
$$x_2 \ge 0$$



### Solving (LP) geometrically

min 
$$\begin{bmatrix} 1 \\ -3 \end{bmatrix}^T \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$
s.t. 
$$\begin{bmatrix} 1 & -2 \\ -1 & -1 \end{bmatrix} \ge \begin{bmatrix} -6 \\ -5 \end{bmatrix}$$

$$x_1 \ge 0, x_2 \ge 0$$

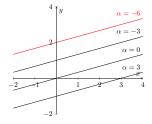
$$x: Ax \le b \equiv [(-A)x \ge (-b)]$$

$$x: Ax = b \equiv \begin{bmatrix} Ax \ge b \\ Ax \le b \end{bmatrix} \text{ i.e. } \begin{bmatrix} A \\ -A \end{bmatrix} x \ge \begin{bmatrix} b \\ -b \end{bmatrix}$$

$$x: \underbrace{Ax = b}_{x \ge 0} \equiv \begin{bmatrix} A \\ -A \\ I \end{bmatrix} x \ge \begin{bmatrix} b \\ -b \\ \vec{0} \end{bmatrix}$$

**Definition.**  $\underline{\alpha}$ -level set is  $\{x: f(x) = \alpha\}$ 

Here 
$$\alpha$$
-level set is  $x: x_1 - 3x_2 = \alpha$  rewrite:  $x_2 = \underbrace{\frac{1}{3}}_{\text{slope}} x - \underbrace{\frac{\alpha}{3}}_{\text{y-int}}$ 



The solution is the level set with the least  $\alpha$ . In this case,  $\alpha=-6$ 

# Lecture 6 (2018-09-14)

<u>Hyperplane</u>:-  $x: p^T x = \alpha$  <u>Halfspace</u>:-  $x: p^T x \ge \alpha$ Polyhedral set - intersection of halfspaces. The example  $Ax \ge b$ 

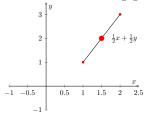
Example 1.

min 
$$c^T x$$
 min  $c^T x$   
s.t.  $Ax = b$  s.t.  $Ax \ge b$   
 $x > 0$   $x > 0$ 

Level set  $c^T x = \alpha$ 

**Definition 1.** Let's say  $x, y \in \mathbb{R}^n$ . A <u>complex combination</u> of xy is  $\lambda x + (1 - \lambda)y \in \mathbb{R}^n$  where  $\lambda \in [0, 1]$ 

**Example 2.** 
$$x = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad y = \begin{bmatrix} 2 \\ 3 \end{bmatrix}, \quad \text{let } \lambda = \frac{1}{2}$$



**Note.** In general, a convex combination of x, y are points on the line segment from x to y as  $\lambda x + (1 - \lambda)y = y + \lambda(x - y)$ 

**Definition 2.**  $S \subseteq \mathbb{R}^n$  is <u>convex</u> if  $\forall x, y \in S, \forall x \in [0, 1]$   $\lambda x + (1 - \lambda)y \in S$ 

**Example 3.** Every polyhedron  $S := \{x \in \mathbb{R}^n : Ax \geq b\}$  is convex since if  $y, z \in S$  [i.e.  $Ay \geq b$ ,  $Ax \geq b$ ] and  $\lambda \in [0, 1]$ , then

$$A(\lambda y + (1 - \lambda)z) = \lambda Ay + (1 - \lambda)Az \ge \lambda b + (1 - \lambda)b = b$$
$$\lambda y + (1 - \lambda)z \in \mathcal{S}$$

**Definition 3.** Let S be a convex set. X is an extreme point of S if  $X = \lambda y + (1 - \lambda)z$  for,  $y, z \in S$ ,  $\lambda \in (0,1) \Rightarrow x = y = z$ 

Example 4. Consider

min 
$$10x_1 + 10x_2 + 0x_3 - 3x_4 - 5x_5 - 3x_6$$
s.t. 
$$-x_1 + 2x_2 + 3x_3 + 6x_4 + 9x_5 + 8x_6 = 26$$

$$-2x_1 + 3x_2 + x_3 + x_4 + 6x_5 + 8x_6 = 17$$

$$x_1 + x_2 - x_3 + x_4 + x_5 + 3x_6 = 1$$

$$x_1, x_2, x_3, \dots, x_6 \ge 0$$

What are feasible x? What if **JUST FOR NOW** we set  $x_4, x_5, x_6 = 0$ 

$$\underbrace{\begin{bmatrix} -1 & 2 & 3 \\ -2 & 3 & 1 \\ 1 & 1 & -1 \end{bmatrix}}_{B \to \text{``Basic''}} \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}}_{x_B \to \text{basic variable}} = \underbrace{\begin{bmatrix} 26 \\ 17 \\ 1 \end{bmatrix}}_{b}$$

Lucky # 1  $\rightarrow$  B is invertible

$$B^{-1} = \frac{1}{13} \begin{bmatrix} 4 & -5 & -1 \\ 1 & 2 & 5 \\ 5 & -3 & -1 \end{bmatrix}$$
 so  $x_B = B^{-1}b = \begin{bmatrix} 2 \\ 5 \\ 6 \end{bmatrix}$   
Hence the basic feasible solutions are 
$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{bmatrix} = \begin{bmatrix} 2 \\ 5 \\ 6 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$