AMS 553.361 - Introduction to Optimization

Lectures by Donniell Fishkind
Notes by Kaushik Srinivasan

Johns Hopkins University
Fall 2018

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 Ti*k*Z 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.

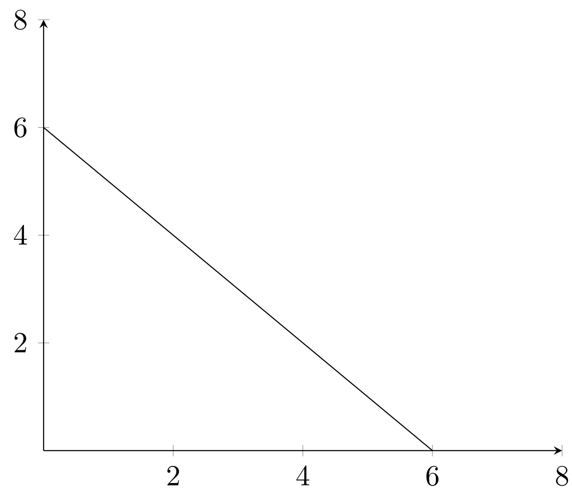
Please email any corrections or suggestions to ksriniv4@jhu.edu.

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 atleast (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 $1 + x_1^2(x_2-1)^3e^{-x_1-x_2}$

$$\begin{aligned} \text{Maximise} \quad & 1 + x_1^2(x_2-1)^3e^{-x_1-x_2} \\ & x_1 + x_2 \leq 6 \\ & x_2 > \log x_1 \\ & x_1 \geq 0 \\ & x_2 \geq 0 \end{aligned}$$



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

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

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

however optimal objective function value is1.2677

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. $-\infty \leq x \leq 7 \to$ unbounded, has no solution.

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

Example 4. Find min $\log x$ s.t. $x > 1$, $x \leq 0.5 \to$ infeasible!!!

Example 5. minimize $3 + (x-1)^2$ s.t. $1 \leq x \leq 3 \to$ 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. $x \geq 10 \to$ optimal solution is $x^* = 10$. but $f'(x) \neq 0$. 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 <u>open</u> 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.

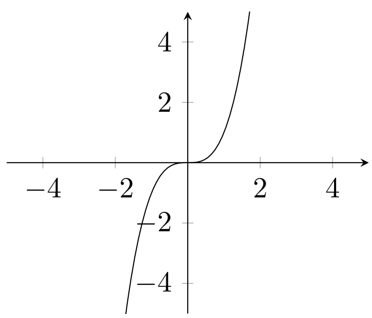
Note: $\forall S \subseteq \mathbb{R}^n, S$ is open **iff** S^c is closed.
In (P) min $f(x)$ s.t. $x \in S$, suppose $x^* \in S$:

- x^* is a <u>global maximizer</u> if $\forall y \in S$, $f(x^*) \leq f(y)$
- x^* is a <u>strict</u> 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* $\Rightarrow f'(x^*) = 0$ $\qquad$ x^* is a *local maximizer* $\Rightarrow f'(x^*) = 0$

$$\textbf{BUT } f'(x^*) = 0 \not\Longrightarrow x^* \text{ local min/max}$$



BUT

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

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

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

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_2 = units of ingredient 2
- x_3 = units of ingredient 3
- 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 $\begin{bmatrix} 6.2 \\ 2 \\ 1.6 \\ 3.2 \end{bmatrix}$

$$\text{minimize} \qquad 6.2x_1 + 2x_2 + 1.6x_3 + 3.2x_4$$

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

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 <u>canonical form</u>, an example

$$\begin{gathered} \text{minimize} \qquad c_1x_1 + c_2x_2 + \cdots + c_nx_n \\ \text{s.t.} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \geq b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \geq b_2 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \geq b_m \\ \text{where} \qquad x_1, x_2, \cdots, x_n \geq 0 \end{gathered}$$

$$\Updownarrow$$

$$\min \quad c^T x \quad \text{s.t.} \quad A\vec{x} \geq \vec{b}, x \geq \vec{0}$$
$$A \in \mathbb{R}^{mxn}, b \in \mathbb{R}^m, C \in \mathbb{R}^n, x \in \mathbb{R}^n$$

Example 2. *Transportation problem*: There are 3 electricity generation plants α, β, γ and two cities u, v. α, β, γ 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
$\beta \to u$	x_3	17.6
$\beta \to v$	x_4	37.4
$\gamma \to u$	x_5	22.8
$\gamma \to v$	x_6	29.7

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

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

$$\left.\begin{array}{ll} u & x_1 + x_3 + x_5 \geq 86.2 \\ v & x_2 + x_4 + x_6 \geq 110.1 \end{array}\right\} \text{Need for electricity for } u, v$$

Converting one form of (LP) to another

1. **(LP)** in standard form $\qquad$ minimize $c^T x$ $\qquad$ s.t. $Ax = b, \quad x \geq \vec{0}$
   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 $\to$ min $\quad$ or $\quad$ min $\to$ max $\quad \Longrightarrow \quad$ $\max 6x_1 + 3x_2 - 4x_3 \leftrightarrow \min \; -6x_1 - 3x_2 + 4x_3$
$$Ax \geq b \to \leq$$
$$Ax \geq b \leftrightarrow \boxed{-A} x \leq \boxed{-b}$$

3. **Standard form $\to$ canonical form**
$$\left.\begin{array}{l} \min \; c^T x \\ \text{s.t. } Ax = b \\ x \geq 0 \end{array}\right\} \leftrightarrow \left\{\begin{array}{l} \min \; c^T x \\ \text{s.t. } Ax \geq b \\ Ax \leq b \\ x \geq \vec{0} \end{array}\right\} \leftrightarrow \left\{\begin{array}{l} \min \; c^T x \\ \text{s.t. } Ax \geq b \\ -Ax \geq -b \\ x \geq \vec{0} \end{array}\right\} \leftrightarrow \left\{\begin{array}{l} \min \; c^T x \text{ s.t.} \\ \begin{bmatrix} A \\ -A \end{bmatrix} x \geq \begin{bmatrix} b \\ -b \end{bmatrix} \\ x \geq \vec{0} \end{array}\right.$$

4. **Canoconical form → Standard form**

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13} &\geq b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23} &\geq b_2 \\ x_1, x_2, x_3 &\geq 0 \end{aligned} \iff \begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13} + \underbrace{x_4 \qquad\quad \\ a_{21}x_1 + a_{22}x_2 + a_{23} \qquad\quad + x_5}_{\text{slack variables}} &= b_1 \\ &= b_2 \\ x_1, x_2, x_3, x_4, x_5 &\geq 0 \end{aligned}$$

Lecture 4 (2018-09-07)

5. **non-negativity → include in matrix**

$$\left.\begin{matrix} Ax \geq b \\ x \geq \vec{0} \end{matrix}\right\} \longrightarrow \begin{bmatrix} A \\ I \end{bmatrix} x \geq \begin{bmatrix} \vec{b} \\ \vec{0} \end{bmatrix}$$

   Simplex algorithm only works with non-negative elements.

6. **unconstrained sign → non-negativity** by <u>substitution</u>

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

$$\left.\begin{matrix} min & 5x_1 + 6x_2 & \\ s.t. & 2x_1 - 3x_2 \geq 9 & \\ & x_1 + x_2 \geq -8 & \\ & x_1 \geq 0 & x_2 \text{unconst} \end{matrix}\right\} \to \left\{ X_2 = \underbrace{X_2'}_{\geq 0} - \underbrace{X_2''}_{\geq 0} \right\} \to \begin{cases} min & 5x_1 + 6x_2' - 6x_2'' \\ s.t. & 2x_1 - 3x_2' + 3x_2'' \geq 9 \\ & x_1 + x_2' - x_2'' \geq -8 \\ & x_1, x_2', x_2'' \geq 0 \end{cases}$$

7. **Have constraints** $Ax = b$

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

   Unchanged by row opeartions

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

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

   If row reduce $\left[\begin{array}{c:c} A & b \end{array}\right] \to \left[\begin{array}{c:c} A' & b' \end{array}\right]$ then $\exists$ invertible matrix C

$$C \left[\begin{array}{c:c} A & b \end{array}\right] \to \left[\begin{array}{c:c} A' & b' \end{array}\right]$$

   Converse: if D is invertible, then $\left[\begin{array}{c:c} A' & b' \end{array}\right]$ can be reduced to $D \left[\begin{array}{c:c} A & b \end{array}\right]$

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

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