

# AMS 553.361 - Introduction to Optimization

Lectures by Donniell Fishkind  
Notes by Kaushik Srinivasan

Johns Hopkins University  
Fall 2018

<b>Lecture 1 (2018-08-30)</b>	<b>1</b>	<b>Lecture 4 (2018-09-07)</b>	<b>7</b>
<b>Lecture 2 (2018-08-31)</b>	<b>3</b>	<b>Lecture 5 (2018-09-12)</b>	<b>8</b>
<b>Lecture 3 (2018-09-05)</b>	<b>4</b>		

## 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.

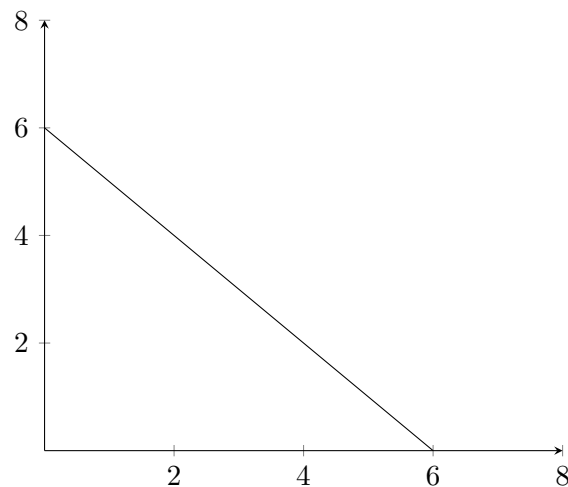
Please email any corrections or suggestions to [ksriniv4@jhu.edu](mailto:ksriniv4@jhu.edu).

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

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



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

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

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

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

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

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

**Example 5.** minimize  $3 + (x - 1)^2$  s.t.  $1 \leq x \leq 3 \rightarrow$  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.  $x \geq 10 \rightarrow$  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 open if every point in  $S$  is an interior point. example: open ball.
- Set  $S \subseteq \mathbb{R}^n$  is closed 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  $x^* \in S$ :

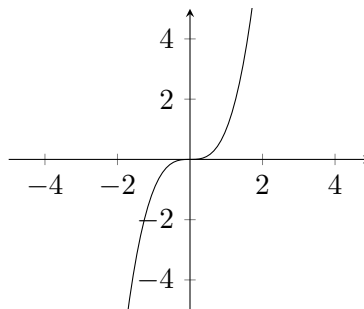
- $x^*$  is a global maximizer if  $\forall y \in S, f(x^*) \leq f(y)$
- $x^*$  is a strict global minimizer if  $\forall y \in S, \text{ s.t. } y \neq x^*, f(x^*) < f(y)$
- $x^*$  is a local minimizer if  $\exists$  any neighborhood  $N$  of  $x$  s.t.  $\forall y \in N \cap S, f(x^*) \leq f(y)$
- $x^*$  is a strict 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**  $f'(x^*) = 0 \not\Rightarrow x^*$  local min/max



**BUT**

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

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

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

$\nabla f(x^*) = \vec{0} \ \& \ [?] \Rightarrow x^*$  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
- $x_2$  = 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  $\begin{bmatrix} 6.2 \\ 2 \\ 1.6 \\ 3.2 \end{bmatrix}$

$$\begin{aligned} &\text{minimize} && 6.2x_1 + 2x_2 + 1.6x_3 + 3.2x_4 \\ &\begin{cases} 1.2x_1 + 2.6x_2 + 0x_3 + 9.2x_4 \geq 6.2 \\ 3.9x_1 + x_2 + 0.8x_3 + 2x_4 \geq 11.9 \\ 3.9x_1 + 1x_2 + 0.8x_3 + 2x_4 \geq 11.9 \\ 6x_1 + 0x_2 + 4x_3 + 3.1x_4 \geq 11.9 \end{cases} && \text{and } x_1, x_2, x_3, x_4 \geq 0 \end{aligned}$$

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

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

$$\begin{aligned}
& \text{where } x_1, x_2, \dots, x_n \geq 0 \\
& \quad \Updownarrow \\
& \min c^T x \quad \text{s.t.} \quad A\vec{x} \geq \vec{b}, x \geq \vec{0} \\
& A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m, C \in \mathbb{R}^n, x \in \mathbb{R}^n
\end{aligned}$$

**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
$\alpha \rightarrow u$	$x_1$	31.7
$\alpha \rightarrow v$	$x_2$	28.6
$\beta \rightarrow u$	$x_3$	17.6
$\beta \rightarrow v$	$x_4$	37.4
$\gamma \rightarrow u$	$x_5$	22.8
$\gamma \rightarrow v$	$x_6$	29.7

$$\min \quad 31.7x_1 + 28.6x_2 + 17.6x_3 + 37.4x_4 + 22.8x_5 + 29.7x_6$$

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

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

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

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

$$c^T = [31.7 \quad 28.6 \quad 17.6 \quad 37.4 \quad 22.8 \quad 29.7]$$

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

$$\begin{aligned}
& Ax (\geq) b \rightarrow (\leq) \\
& Ax \geq b \leftrightarrow \boxed{-A} x \leq \boxed{-b}
\end{aligned}$$

### 3. Standard form $\rightarrow$ 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 $\rightarrow$ Standard form

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



## Lecture 4 (2018-09-07)

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

$$\left. \begin{array}{l} Ax \geq b \\ x \geq \vec{0} \end{array} \right\} \rightarrow \begin{bmatrix} A \\ I \end{bmatrix} x \geq \begin{bmatrix} \vec{b} \\ \vec{0} \end{bmatrix}$$

Simplex algorithm only works with non-negative elements.

### 6. unconstrained sign → non-negativity by substitution

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

$$\left. \begin{array}{ll} \min & 5x_1 + 6x_2 \\ \text{s.t.} & 2x_1 - 3x_2 \geq 9 \\ & x_1 + x_2 \geq -8 \\ & x_1 \geq 0 \quad x_2 \text{unconst} \end{array} \right\} \rightarrow \left\{ \begin{array}{l} X_2 = \underbrace{X'_2}_{\geq 0} - \underbrace{X''_2}_{\geq 0} \end{array} \right\} \rightarrow \left\{ \begin{array}{ll} \min & 5x_1 + 6x'_2 - 6x''_2 \\ \text{s.t.} & 2x_1 - 3x'_2 + 3x''_2 \geq 9 \\ & x_1 + x'_2 - x''_2 \geq -8 \\ & x_1, x'_2, x''_2 \geq 0 \end{array} \right.$$

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

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

Unchanged by row operations

- 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  $[A \mid b] \rightarrow [A' \mid b']$  then  $\exists$  invertible matrix  $C$

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

**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

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

## Lecture 5 (2018-09-12)

### Half spaces, Hyperplanes, Polyhedral sets

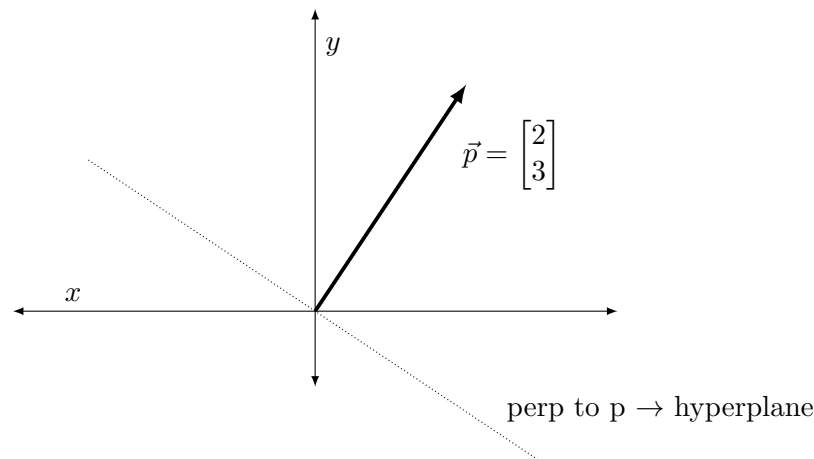
**Recall:** Suppose  $p, x \in \mathbb{R}^n$ . The inner product of  $p, x$

$$p^T x = \|p\| \|x\| \cos \theta$$

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

**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  $\{x \in \mathbb{R}^n, p^T x \geq \alpha\}$  for non-zero  $p \in \mathbb{R}^n$ , and  $\alpha \in \mathbb{R}$

**Example 1.**

$$\begin{array}{lll} (2D) & 2x_1 + 5x_2 = 6 & \text{hyperplane} \\ & 2x_1 + 5x_2 \geq 6 & \text{half space} \\ (3D) & 3x_1 + 2x_2 - 7x_3 = 8 & \text{hyperplane} \\ & 3x_1 + 2x_2 - 7x_3 \geq 8 & \text{half space} \end{array}$$

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

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

$$\begin{array}{l} \vec{p}_1 \rightarrow \\ \vec{p}_2 \rightarrow \\ \vdots \\ \vec{p}_m \rightarrow \end{array} \begin{bmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \geq \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

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

**Example 3.**

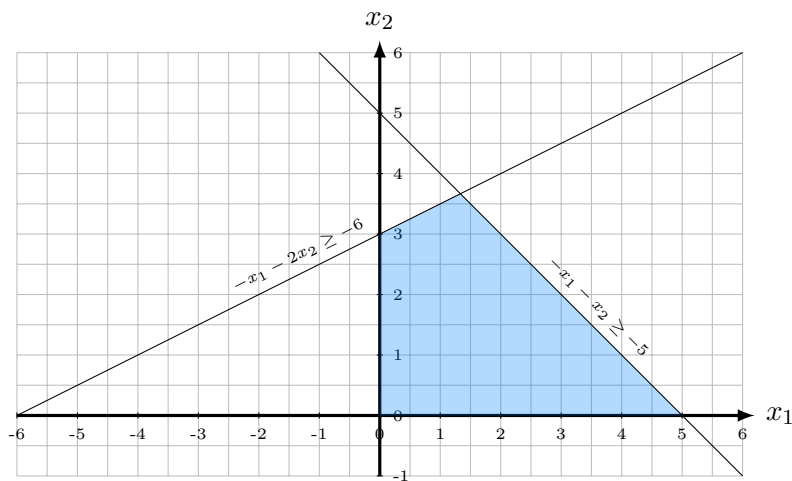
$$\begin{bmatrix} 1 & -2 \\ -1 & -1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \geq \begin{bmatrix} -6 \\ -5 \\ 0 \\ 0 \end{bmatrix}$$

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

3.  $x_1 \geq 0$

2.  $-x_1 - x_2 \geq 5$

4.  $x_2 \geq 0$



**Solving (LP) geometrically**

$$\begin{aligned} \min \quad & \begin{bmatrix} 1 \\ -3 \end{bmatrix}^T \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \\ \text{s.t.} \quad & \begin{bmatrix} 1 & -2 \\ -1 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \geq \begin{bmatrix} -6 \\ -5 \end{bmatrix} \\ & x_1 \geq 0, x_2 \geq 0 \end{aligned}$$

**polyhedral sets:**

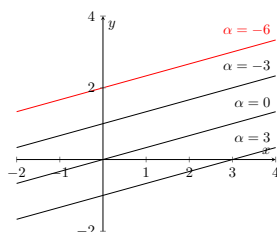
$$x : Ax \leq b \equiv [(-A)x \geq (-b)]$$

$$x : Ax = b \equiv \begin{bmatrix} Ax \geq b \\ Ax \leq b \end{bmatrix} \quad \text{i.e.} \quad \begin{bmatrix} A \\ -A \end{bmatrix} x \geq \begin{bmatrix} b \\ -b \end{bmatrix}$$

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

**Definition.**  $\alpha$ -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,  $\alpha = -6$