

## Lecture 2: Standard LP formulation, Linear algebra : solution to simultaneous linear equations $Ax = b$ , Gaussian elimination

Lecturer: *Sundar Vishwanathan*Scribe: *Dinesh Gadge*

COMPUTER SCIENCE &amp; ENGINEERING

INDIAN INSTITUTE OF TECHNOLOGY, BOMBAY

### 1 Formulation

A linear optimization problem can be formulated as

$$\begin{aligned} \max \quad & c^T x \\ \text{subject to} \quad & Ax \leq b \end{aligned} \tag{1}$$

where,  $A$  is an  $m \times n$  matrix,  $c$  a  $n \times 1$  vector,  $b$  a  $m \times 1$  vector and  $x$  a  $n \times 1$  vector. We are given as **Input**:  $c, A, b$  and desire as **Output** :  $x$ . Among all  $x$  that satisfies  $Ax \leq b$  find one which maximises  $c^T x$ . The set  $Ax \leq b$  is a set of points which have typical properties. Our first goal is to understand them and be able to describe some of these.

### 2 Understanding the set of all solutions to $Ax = b$

Before considering the set of inequalities,  $Ax \leq b$ , we consider the set of equalities  $Ax = b$ . This is short hand notation for the following set of equalities given below.

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \cdots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + a_{m3}x_3 + \cdots + a_{mn}x_n &= b_m \end{aligned}$$

We can solve such a system of equations using Gaussian Elimination. Here is an example.

EXAMPLE 1

$$2x + 7y = 13 \tag{I}$$

$$x + 3y = 4 \tag{II}$$

Replacing II by  $-\frac{1}{2} \cdot I + II$  gives

$$\begin{aligned} -\frac{7}{2}y + 3y &= -\frac{13}{2} + 4 \\ y &= 5 \end{aligned} \tag{2}$$

### 3 Gaussian Elimination

$$\begin{aligned}
 a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \cdots + a_{1n}x_n &= b_1 \\
 a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \cdots + a_{2n}x_n &= b_2 \\
 &\vdots \\
 a_{m1}x_1 + a_{m2}x_2 + a_{m3}x_3 + \cdots + a_{mn}x_n &= b_m
 \end{aligned}$$

Assume that  $a_{11} \neq 0$ . If this is not so we exchange the first row with some other row which has a non-zero first co-ordinate. Then for each of the equations except the first, multiply the first equation by a suitable constant and subtract from the respective equations to get rid of  $x_1$  in the other equations.

We then do this with the other equations ignoring the first variable.

There are 2 operations used in Gaussian Elimination.

1. Exchange two rows.
2. Replace  $row_j$  with  $\alpha \cdot row_i + row_j$ , where  $\alpha$  is some constant.

It is possible that using these two operations the co-efficients of some  $x_i$  [for example  $x_2$ ] in all rows except one are zeroed out.

$$\begin{aligned}
 a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \cdots + a_{1n}x_n &= b_1 \\
 0 + 0 + a_{23}x_3 + \cdots + a_{2n}x_n &= b'_2 \\
 0 + 0 + a_{33}x_3 + \cdots + a_{3n}x_n &= b'_3 \\
 &\vdots \\
 0 + 0 + a_{m3}x_3 + \cdots + a_{mn}x_n &= b'_m
 \end{aligned}$$

In such a case, assume that, some  $a_{ij} \neq 0$  [for example  $a_{23} \neq 0$ ], then we repeat the process.

Finally, we get something like,

$$\begin{aligned}
 a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\
 0 + 0 + \cdots + a_{2i_1}x_{i_1} + \cdots + a_{2n}x_n &= b''_2 \\
 &\vdots \\
 0 + 0 + \cdots + 0 + a_{ki_{k-1}}x_{i_{k-1}} + \cdots + a_{kn}x_n &= b''_k \\
 &\vdots \\
 0 + 0 + \cdots + 0 + 0 + 0 + \cdots + 0 &= b''_l \\
 &\vdots \\
 0 + 0 + \cdots + 0 + 0 + 0 + \cdots + 0 &= b''_m
 \end{aligned}$$

#### Observations:

1. All zero rows occur after the non-zero rows.

2. If, from the top, the first  $t$  rows are nonzero, and the first non-zero entry of the  $i$ th row  $i = 1, \dots, t$ , is at the  $k_i$ th column, then  $k_i > k_{i-1}$ ,  $i = 2, \dots, t$ . That is, the first non-zero entries in the rows appear later and later from top to bottom.

### 3.1 Existence of a solution for $Ax = b$

$Ax = b$  does not have a solution if in a particular row  $i$ , all coefficients  $a_{ij} = 0$ , but  $b_i \neq 0$ . This is both necessary and sufficient condition for non-existence of a solution for  $Ax = b$  as we shall see later.

Once we have the matrix in this form, it is easy to get solutions to the set of equations, if one exists. Except for  $x_1, x_{i_1}, \dots, x_{i_{k-1}}$  set any values to the other variables. Now solve for the variables  $x_1, x_{i_1}, \dots, x_{i_{k-1}}$  in the reverse order.

If it has  $n$  non-zero rows then it has only one solution. Otherwise  $\{x : Ax = b\}$ , in general, has many solutions.

### 3.2 Why is this procedure correct?

The procedure being correct means that the values of the variables obtained by the procedure indeed satisfy the original set of equations. So to prove the correctness of the procedure, we have to prove that the solution set does not change on applying the operations of Gaussian Elimination. Clearly exchanging two rows does not change the solution set.

**THEOREM 1** Given a set of equations, suppose  $eq_j$  is replaced by

$$\alpha \cdot eq_i + eq_j$$

then the solutions set does not change.

**PROOF:** Let  $eq_i$  be  $a_{i1}x_1 + a_{i2}x_2 + \dots + a_{in}x_n = b_i$  and  $eq_j$  be  $a_{j1}x_1 + a_{j2}x_2 + \dots + a_{jn}x_n = b_j$ . Consider the old and the new set of equations. Note that they differ only in the  $j$ th equation. This proof consists of 2 parts.

**Part 1:** If the set of solutions satisfies original set of equations then it satisfies the new set. Let  $\langle x'_1, x'_2, \dots, x'_n \rangle$  be the solution to the original set. Then  $\langle x'_1, x'_2, \dots, x'_n \rangle$  specifically satisfies  $eq_i$  and  $eq_j$ . So,

$$\begin{aligned} a_{i1}x'_1 + a_{i2}x'_2 + \dots + a_{in}x'_n &= b_i \\ a_{j1}x'_1 + a_{j2}x'_2 + \dots + a_{jn}x'_n &= b_j \end{aligned}$$

that is,

$$\begin{aligned} a_{i1}x'_1 + a_{i2}x'_2 + \dots + a_{in}x'_n - b_i &= 0 \\ a_{j1}x'_1 + a_{j2}x'_2 + \dots + a_{jn}x'_n - b_j &= 0 \end{aligned}$$

So,

$$\begin{aligned} &\alpha(a_{i1}x'_1 + a_{i2}x'_2 + \dots + a_{in}x'_n - b_i) + (a_{j1}x'_1 + a_{j2}x'_2 + \dots + a_{jn}x'_n - b_j) \\ &= \alpha \cdot 0 + 0 \\ &= 0 \end{aligned}$$

Therefore,

$$\alpha(a_{i1}x'_1 + a_{i2}x'_2 + \cdots + a_{in}x'_n) + (a_{j1}x'_1 + a_{j2}x'_2 + \cdots + a_{jn}x'_n) = \alpha \cdot b_i + b_j$$

Hence  $\langle x'_1, x'_2, \dots, x'_n \rangle$  satisfies  $\alpha \cdot eq_i + eq_j$ . Since rest of the equations in both the sets are identical,  $\langle x'_1, x'_2, \dots, x'_n \rangle$  satisfies the new set.

**Part 2:** *If the set of solutions satisfies new set of equations then it satisfies the original set.* Let  $\langle x'_1, x'_2, \dots, x'_n \rangle$  be the solution to the new set. Then  $\langle x'_1, x'_2, \dots, x'_n \rangle$  specifically satisfies  $\alpha \cdot eq_i + eq_j$  and  $eq_i$ . So,

$$\begin{aligned} \alpha(a_{i1}x'_1 + a_{i2}x'_2 + \cdots + a_{in}x'_n) + (a_{j1}x'_1 + a_{j2}x'_2 + \cdots + a_{jn}x'_n) &= \alpha \cdot b_i + b_j \\ a_{i1}x'_1 + a_{i2}x'_2 + \cdots + a_{in}x'_n &= b_i \end{aligned}$$

that is,

$$\begin{aligned} \alpha(a_{i1}x'_1 + a_{i2}x'_2 + \cdots + a_{in}x'_n - b_i) + (a_{j1}x'_1 + a_{j2}x'_2 + \cdots + a_{jn}x'_n - b_j) &= 0 \\ a_{i1}x'_1 + a_{i2}x'_2 + \cdots + a_{in}x'_n - b_i &= 0 \end{aligned}$$

So,

$$\begin{aligned} \alpha \cdot 0 + (a_{j1}x'_1 + a_{j2}x'_2 + \cdots + a_{jn}x'_n - b_j) &= 0 \\ \Rightarrow a_{j1}x'_1 + a_{j2}x'_2 + \cdots + a_{jn}x'_n &= b_j \end{aligned}$$

Hence  $\langle x'_1, x'_2, \dots, x'_n \rangle$  satisfies  $eq_j$ . Since rest of the equations in both the sets are identical,  $\langle x'_1, x'_2, \dots, x'_n \rangle$  satisfies the old set.  $\square$

## 4 Understanding $Ax = b$ geometrically

Another way of looking at  $Ax = b$  is through geometry. The operation of adding a constant times another equation to an equation rotates one of the hyperplanes in  $\mathbb{R}^n$ . Fig. 1 illustrates this as the case of rotation of lines (1-dimensional hyperplane) in a  $x$ - $y$  plane (i.e.  $\mathbb{R}^2$ ). Here the solid lines are from example 1 and one of the dotted lines is obtained by rotating the line corresponding to eq. II to the one corresponding to eq. 2. The other dotted line can be obtained by similarly manipulating eq. I

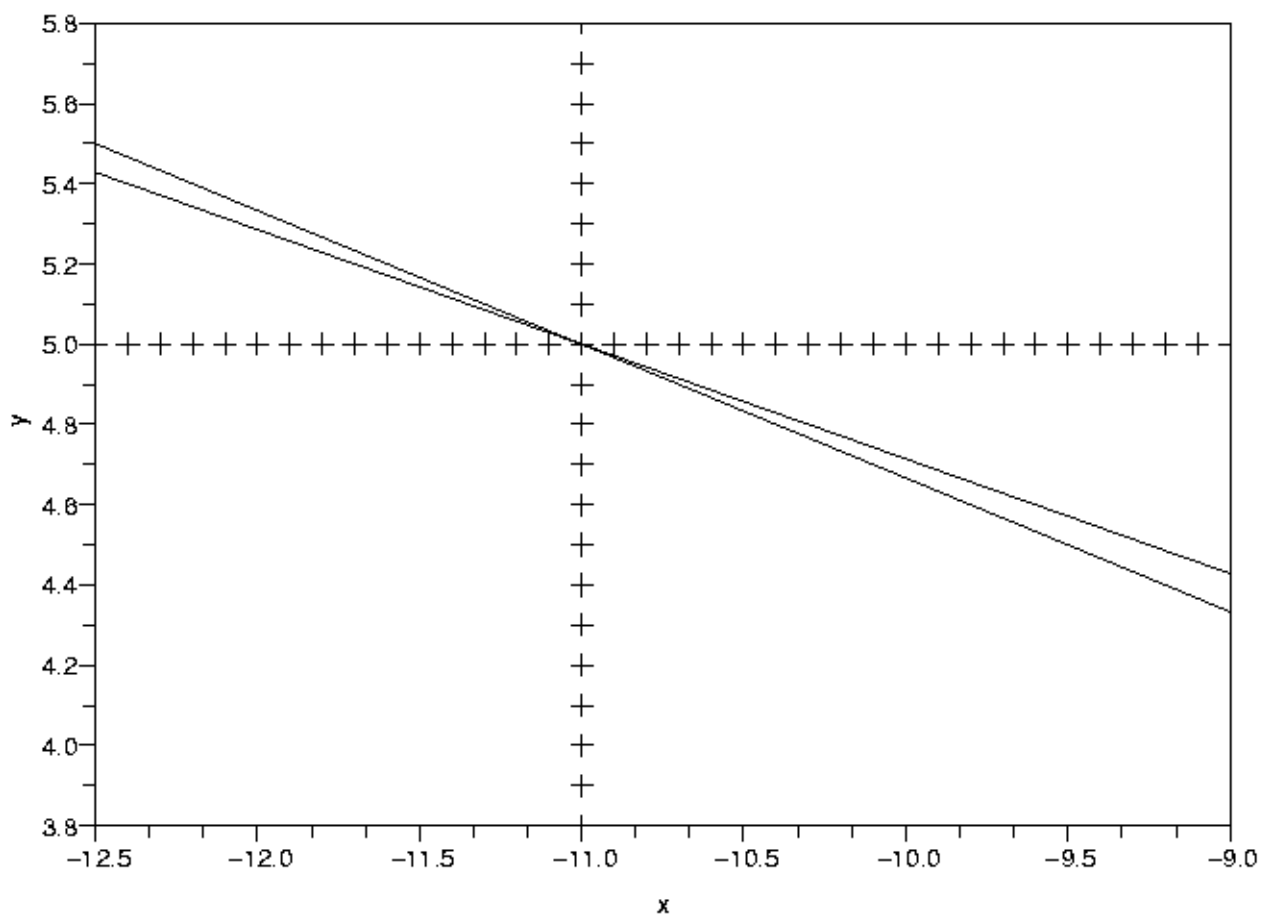


Figure 1: Geometrical way of looking at  $Ax = b$