Math 171A: Linear Programming

Lecture 4

Properties of the Objective Function

Philip E. Gill

http://ccom.ucsd.edu/~peg/math171a

Monday, January 11th, 2021

Visualizing the objective function: method 1

Define a new variable $z = \ell(x) \ (= c^T x)$.

For any given x, now consider $(x_1, x_2, \dots, x_n, z)$ as a point in $\mathbb{R}^{n+1}$.

The set $\mathcal{S}$ of all such points is called *the surface defined by ℓ*.

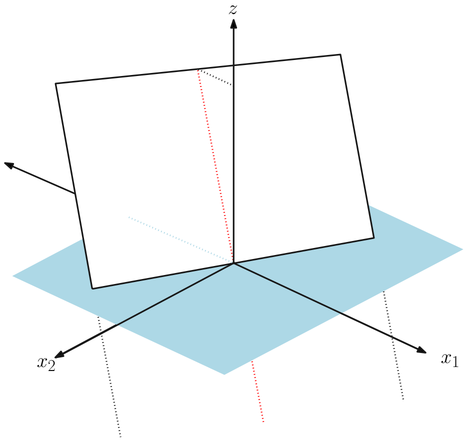
Points on the surface $\mathcal{S}$ satisfy $c^T x - z = 0$, i.e.,

$$\begin{pmatrix} c_1 & c_2 & \cdots & c_n & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \\ z \end{pmatrix} = 0.$$

Visualizing the objective function: method 1

The surface is a *hyperplane in* $\mathbb{R}^{n+1}$ with normal vector

$$\begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \\ -1 \end{pmatrix}$$



Visualizing the objective function: method 2

So far, we have:

- Defined the values of x_1, x_2, ..., x_n.
- Computed $z = c^T x$.
- Plotted $(x_1, x_2, \ldots, x_n, z)$ in $\mathbb{R}^{n+1}$.

Now we reverse the procedure:

- We *fix* z at z_0 (say).
- Plot the points x such that $c^T x = z_0$ in $\mathbb{R}^n$.

Definition (Level curve)

The set $\mathcal{L} = \{x : c^T x = z_0\}$ is called a *level curve* of $\ell(x)$.

Definition (Level curve)

The set $\mathcal{L} = \{x : c^T x = z_0\}$ is called a *level curve* of $\ell(x)$.

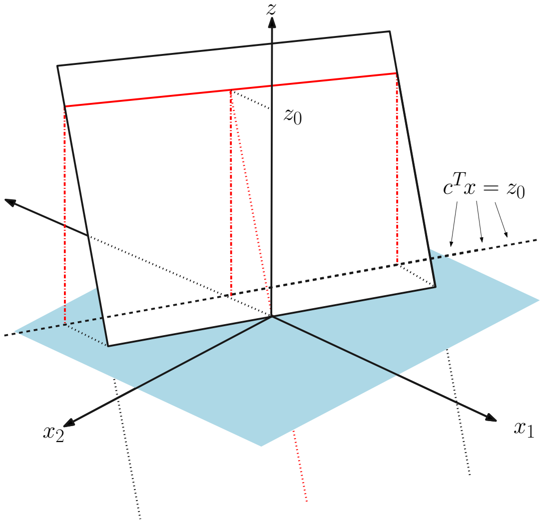
- A level curve is another hyperplane, this time in $\mathbb{R}^n$.

  The normal vector is c, the objective vector.

- By varying z_0 we get an *infinite family* of level curves.

- The level curves are parallel because they all have the same normal vector.

What is the connection between these two visualizations of ℓ?


z
z_0
$c^T x = z_0$
x_2
x_1

Every point $\bar{x} \in \mathbb{R}^n$ lies on *some* level curve.

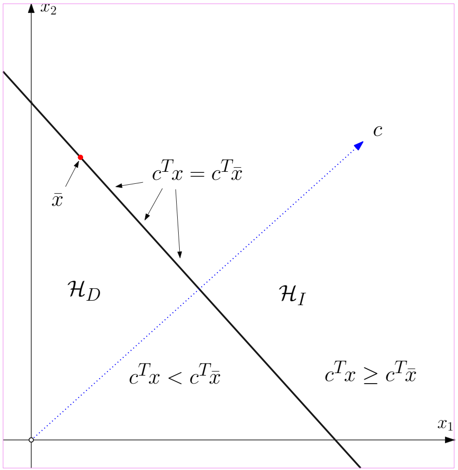
In particular, $\bar{x}$ must lie on the hyperplane $\{x : c^T x = c^T \bar{x}\}$.

This hyperplane splits $\mathbb{R}^n$ into two half-spaces

$$\mathcal{H}_I = \{x : c^T x \geq c^T \bar{x}\} \quad (\text{"I" for increasing})$$

$$\mathcal{H}_D = \{x : c^T x < c^T \bar{x}\} \quad (\text{"D" for decreasing})$$

This implies that the vector c must point *into* the half-space $\mathcal{H}_I$.



$\Rightarrow$ level curves with *larger* objective value lie in $\mathcal{H}_I$

$\Rightarrow$ level curves with *smaller* objective value lie in $\mathcal{H}_D$

Consider the one-dimensional variation of $\ell(x)$ as we move along the linear path $x(\alpha) = \bar{x} + \alpha p$, for $\alpha \geq 0$.

At any point on the path, the objective is

$$\begin{aligned} \ell(\bar{x} + \alpha p) &= c^T(\bar{x} + \alpha p) \\ &= c^T\bar{x} + \alpha c^T p \\ &= \ell(\bar{x}) + \alpha c^T p. \end{aligned}$$

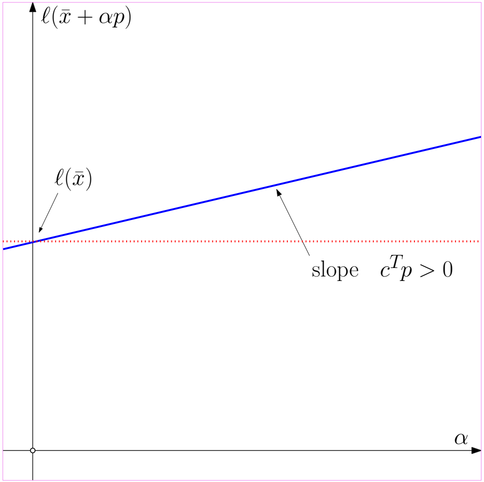
As expected, $\ell(\bar{x} + \alpha p)$ is a linear function of α, with rate of change

$$\frac{d}{d\alpha}\ell(\bar{x} + \alpha p)\bigg|_{\alpha=0} = c^T p,$$

$\Rightarrow$ $c^T p$ gives the (constant) rate of change of $\ell(x)$.

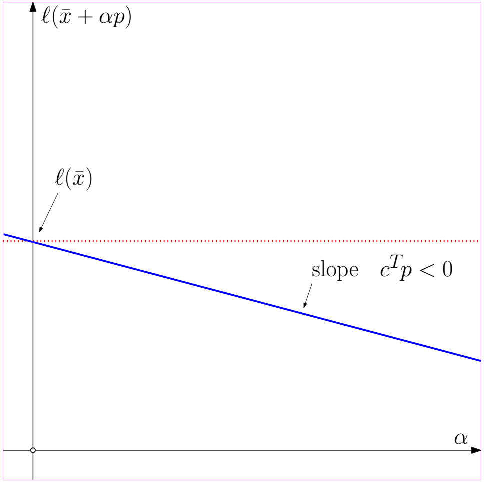
One-dimensional variation of ℓ

Case 1: $c^T p > 0$



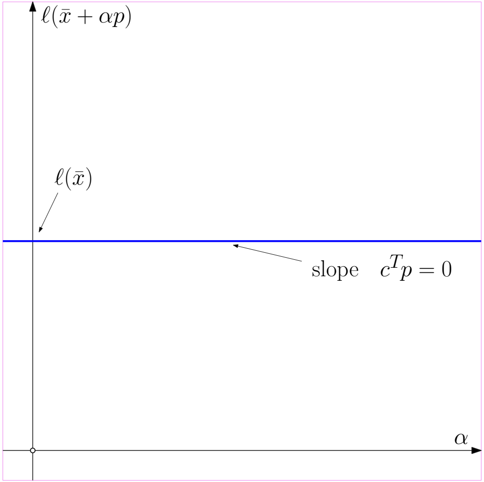
One-dimensional variation of ℓ

Case 2: $c^T p < 0$



One-dimensional variation of ℓ

Case 3: $c^T p = 0$



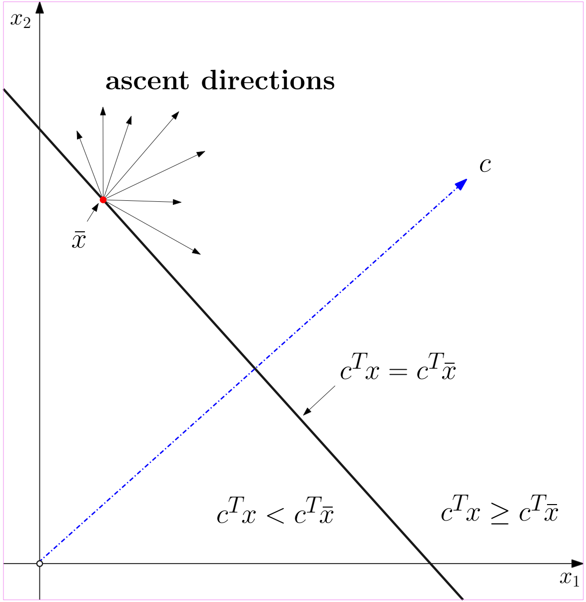
Definition (Ascent direction)

The vector p is an *ascent direction* for $\ell(x)$ if $c^T p > 0$.

At the point $\bar{x}$, an *ascent direction* points into the half-space

$$\mathcal{H}_I = \{x : c^T x > c^T \bar{x}\},$$

i.e., the set of points with *bigger* objective value.


x_2
ascent directions
c
$\bar{x}$
$c^T x = c^T \bar{x}$
$c^T x < c^T \bar{x}$
$c^T x \geq c^T \bar{x}$
x_1

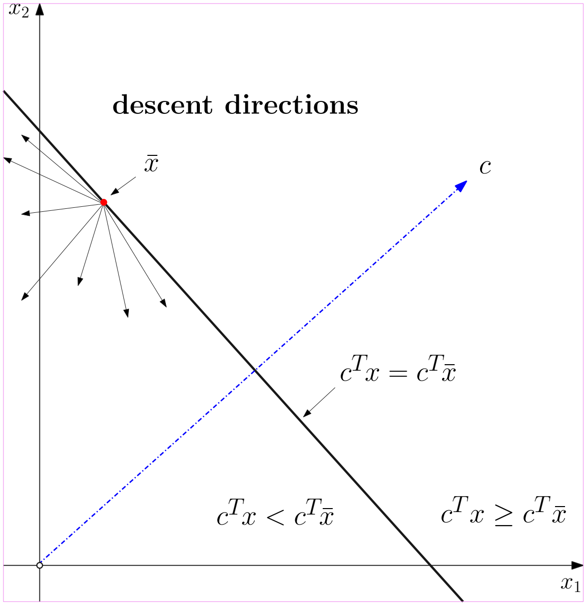
Definition (Descent direction)

The vector p is a *descent direction* for $\ell(x)$ if $c^T p < 0$.

At the point $\bar{x}$, a *descent direction* points into the half-space

$$\mathcal{H}_D = \{x : c^T x < c^T \bar{x}\},$$

i.e., the set of points with *smaller* objective value.


x_2
descent directions
$\bar{x}$
c
$c^T x = c^T \bar{x}$
$c^T x < c^T \bar{x}$
$c^T x \geq c^T \bar{x}$
x_1

Result

If p is a *descent direction* for $\ell(x) = c^T x$, then

$$\ell(\bar{x} + \alpha p) < \ell(\bar{x}) \quad \text{for } \textit{all } \alpha > 0,$$

i.e., $\ell(x)$ is *unbounded below* along any descent direction.

Similarly, $\ell(x)$ is *unbounded above* along any ascent direction.

Summary: basic properties of an LP

- An LP is either *infeasible*, *unbounded* or has an *optimal solution*.
- An *optimal solution* always lies on the boundary of the feasible region.
- Every point on the boundary of the feasible region satisfies a *linear system of equations* that is either square, underdetermined or overdetermined.

Two Methods for Toy Linear Programs

The graphical method

1. Graph the feasible region.
2. Find a corner point of $\mathbb{F}$ and a level curve $\{x : c^T x = z_0\}$ such that
   (a) z_0 is as small as possible; and
   (b) the level curve passes through the corner point.
3. Select two hyperplanes that past through the corner point and solve for x.

Example 1:

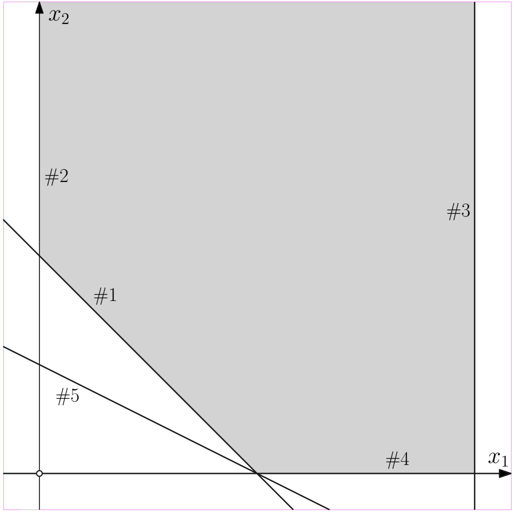
$$\begin{aligned} \underset{x_1,x_2}{\text{minimize}} \quad & 2x_1 + x_2 \\ \text{subject to} \quad & x_1 + x_2 \geq 1 \\ & x_1 \geq 0 \\ & -x_1 \geq -2 \\ & x_2 \geq 0 \\ & x_1 + 2x_2 \geq 1. \end{aligned}$$

$$\begin{aligned} \underset{x\in\mathbb{R}^2}{\text{minimize}} \quad & c^T x \\ \text{subject to} \quad & Ax \geq b, \end{aligned}$$

$$\text{with} \quad c = \begin{pmatrix} 2 \\ 1 \end{pmatrix},$$

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \\ -1 & 0 \\ 0 & 1 \\ 1 & 2 \end{pmatrix}, \quad b = \begin{pmatrix} 1 \\ 0 \\ -2 \\ 0 \\ 1 \end{pmatrix}.$$

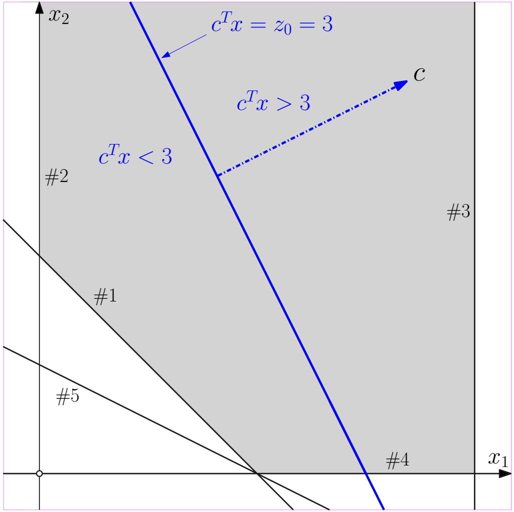
First, we graph the 5 constraints.



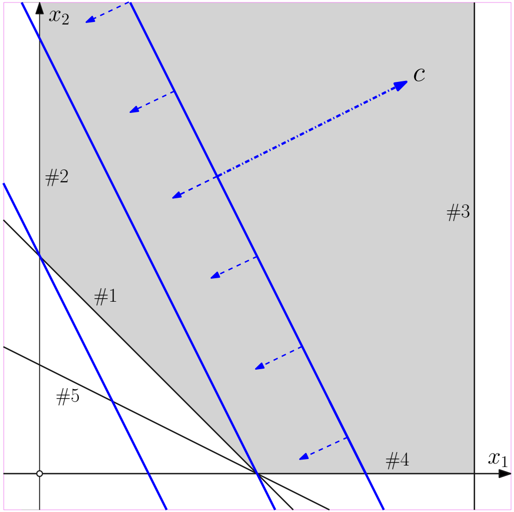
Next we graph just one level curve.

Choosing the objective value $z_0 = 3$ defines the level curve:

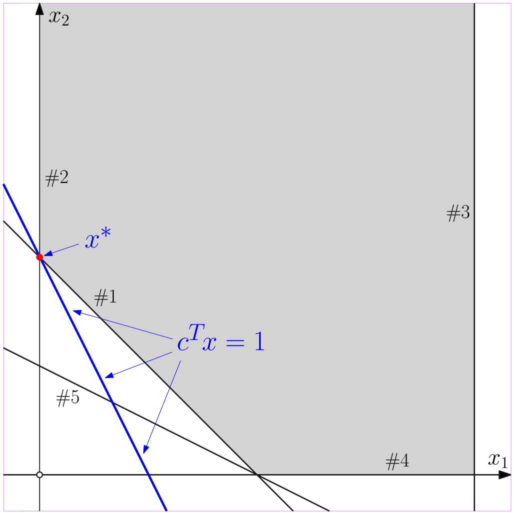
$$2x_1 + x_2 = 3.$$



Next we choose the corner point corresponding to the best level curve.



Next we choose the corner point corresponding to the best level curve.



The optimal corner point lies at the intersection of the hyperplanes

$$\mathcal{H}_1 = \{x : x_1 + x_2 = 1\}$$
$$\mathcal{H}_2 = \{x : x_1 = 0\}.$$

This implies that x^* satisfies the equations

$$\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \text{i.e.,} \quad x^* = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

The optimal objective value is then

$$c^T x^* = \begin{pmatrix} 2 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = 1.$$

Example 2:

$$\begin{aligned} \underset{x_1,x_2}{\text{minimize}} \quad & x_1 + x_2 \\ \text{subject to} \quad & x_1 + x_2 \geq 1 \\ & x_1 \geq 0 \\ & -x_1 \geq -2 \\ & x_2 \geq 0 \\ & x_1 + 2x_2 \geq 1. \end{aligned}$$

Only the objective has changed from Example 1.

$$\underset{x \in \mathbb{R}^2}{\text{minimize}} \quad c^T x$$
$$\text{subject to} \quad Ax \geq b,$$

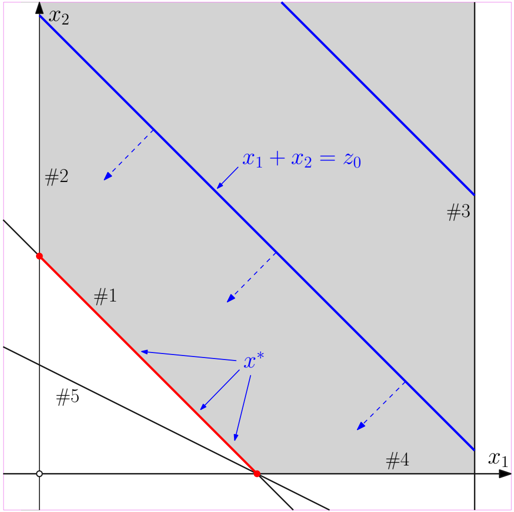
$$\text{with} \qquad c = \begin{pmatrix} 1 \\ 1 \end{pmatrix},$$

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \\ -1 & 0 \\ 0 & 1 \\ 1 & 2 \end{pmatrix}, \quad b = \begin{pmatrix} 1 \\ 0 \\ -2 \\ 0 \\ 1 \end{pmatrix}.$$

In this case, c is parallel to a_1 and there are *infinitely many* solutions *(all with the same optimal objective value)*.

Nevertheless, there are two optimal corner points, at

$$x^* = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad \text{and} \quad x^* = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$



Example 3:

$$\begin{array}{lrcrcr} \underset{x_1,x_2}{\text{minimize}} & -2x_1 & - & x_2 & & \\ \text{subject to} & x_1 & + & x_2 & \geq & 1 \\ & x_1 & & & \geq & 0 \\ & -x_1 & & & \geq & -2 \\ & & & x_2 & \geq & 0 \\ & x_1 & + & 2x_2 & \geq & 1. \end{array}$$

Again, only the objective has been changed.

$$\begin{array}{ll} \underset{x\in\mathbb{R}^2}{\text{minimize}} & c^T x \\ \text{subject to} & Ax \geq b, \end{array}$$

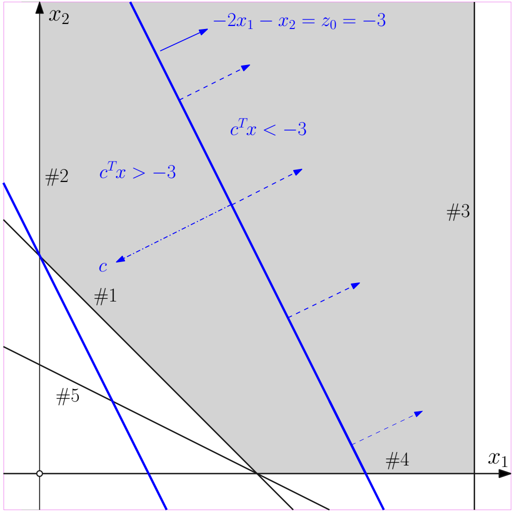
$$\text{with} \quad c = \begin{pmatrix} -2 \\ -1 \end{pmatrix},$$

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \\ -1 & 0 \\ 0 & 1 \\ 1 & 2 \end{pmatrix}, \quad b = \begin{pmatrix} 1 \\ 0 \\ -2 \\ 0 \\ 1 \end{pmatrix}.$$

In this case, we can reduce $z_0 = c^T x$ to minus infinity and never encounter a corner point.

The objective function is *unbounded below* in the feasible region.

We say that the "problem is unbounded".



The "Brute-Force" Method

How do we solve the problem when $n > 2$?

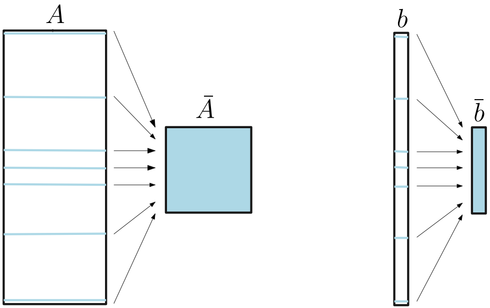
Assume that $m \geq n$.

A bounded solution always lies at a corner point (well, almost always).

$\Rightarrow$ evaluate $\ell(x)$ at *every* corner point.

A corner point must lie on n hyperplanes.

$\Rightarrow$ find a corner point by *gathering* n hyperplanes from A and b.



If $\bar{A}$ is nonsingular and the solution of $\bar{A}\bar{x} = \bar{b}$ is feasible, then $\bar{x}$ is a corner point.

1. Define $\ell^* = +\infty$
2. Gather a subset of n hyperplanes from the m rows of A.
   (a) if $\bar{A}$ is singular, continue at Step 2.
   (b) Solve $\bar{A}\bar{x} = \bar{b}$
   (c) If $\bar{x}$ is infeasible, continue at Step 2.
   (d) $\ell^* = \min\{\ell^*, c^T\bar{x}\}$
   (e) If all subsets of hyperplanes have been examined, stop.
   Otherwise, continue at Step 2.

How efficient is this method?

Recall that there are at most $\binom{m}{n} = \dfrac{m!}{n!(m-n)!}$ corner points

Case 1: $m = 5,\ n = 2$:

$$\binom{5}{2} = \frac{5!}{2!\,3!} = 10 \quad \text{corner points (at most)}$$

Case 2: $m = 20,\ n = 10$:

$$\binom{20}{10} = \frac{20!}{10!\,10!} = 184756 \quad \text{corner points (at most)}$$

Assuming n^3 operations to solve for x

$\Rightarrow\ \approx 1.8$ billion calculations!

Summary: basic properties of an LP

- An LP is either *infeasible*, *unbounded* or has an *optimal solution*.
- An *optimal solution* always lies on the boundary of the feasible region.
- Every point on the boundary of the feasible region satisfies a *linear system of equations* that is either square, underdetermined or overdetermined.

Aim of the class

(I) Given a boundary point $\bar{x}$, can we determine if $\bar{x}$ is optimal without the need to evaluate $c^T x$ at every other corner point?

(II) If $\bar{x}$ is not a solution, can we determine a direction p along which $c^T x$ decreases?

Review of Linear Equations

Review of linear equations

We review properties of systems of linear equations

$$Ax = b,$$

where A is an $m \times n$ matrix and b is an m-vector.

We say $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$.

We make no assumptions on the *shape of A*.

We cannot say that $x = A^{-1}b$, in general.