Advanced Algorithms

Floriano Zini

Free University of Bozen-Bolzano
Faculty of Computer Science

Academic Year 2013-2014

Lecture 6 – Linear Programming (cont.)

Example 2 – Pension fund

- A **pension fund** has **30 million dollar** to invest
- The **money** is to be **divided** among **Treasury notes**, **bonds**, and **stocks**
- The **rules** for administration of the fund require that
  - **at least 3 million be invested in each type of investment**
  - **at least half the money be invested in Treasury notes and bonds**
  - **the amount invested in bonds not exceed twice the amount invested in Treasury notes**
- The annual **yields** for the various **investments** are:
  - **7% for Treasury notes**
  - **8% for bonds**
  - **9% for stocks**
- **Question: How should the money be allocated among the various investments to produce the largest return?**

Example 2 – Pension fund (cont.)

First step – Identification of the problem:

- Individuate (or choose) the objective
  - ***maximize the total return***
- Individuate the available resources
  - ***30 million dollars***
- Individuate available data
  - ***annual yields for the various investments***

Example 2 – Pension fund (cont.)

Second step – Formulation of the linear programming problem:

- Define the variables
  - x_t: amount in Treasury notes
  - x_b: amount in bonds
  - x_s: amount in stocks
- Define the mathematical relations among variables and data
  - Constraints
    - at least 3 million be invested in each type of investment $\boxed{x_t, x_b, x_s \geq 3}$
    - at least half the money be invested in Treasury notes and bonds $\boxed{x_t + x_b \geq 15}$
    - the amount invested in bonds not exceed twice the amount invested in Treasury notes $\boxed{x_b \leq 2 \cdot x_t}$
    - the total amount of investments corresponds to the budget $\boxed{x_t + x_b + x_s = 30}$

Example 2 – Pension fund (cont.)

Second step – Formulation of the linear programming problem:

- Define the mathematical relations among variables and data
  - Objective function:
    - **Maximizing the total yearly return**

$$\boxed{max\ 0.07x_t + 0.08x_b + 0.09x_s}$$

Example 2 – Pension fund (cont.)

To summarize:

Mathematical model:

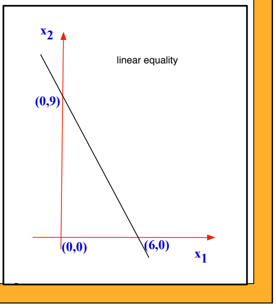
x_t:Treasury notes, x_b: bonds, x_s: stocks.

Maximizing the total return

$$max\ 0.07x_t + 0.08x_b + 0.09x_s$$
$$x_t + x_b \geq 15$$
$$x_b \leq 2 \cdot x_t$$
$$x_t + x_b + x_s = 30$$
$$x_t, x_b, x_s \geq 3$$

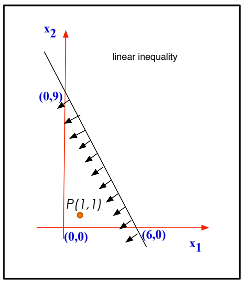
Feasible Region

- What is the Feasible Region of a LP problem?
  - **It is the set of all points that satisfy all the constraints**
- Only **points in the feasible region** can be **used to assign values to variables**
- **A linear equation defines a hyper plan in the k-D space**
  - E.g., consider the equation $3x_1 + 2x_2 = 18$ It is represented in the 2-D space by a line intersecting the axes in $(x_1=0, x_2=18/2=9)$ and $(x_1=18/3=6, x_2=0)$



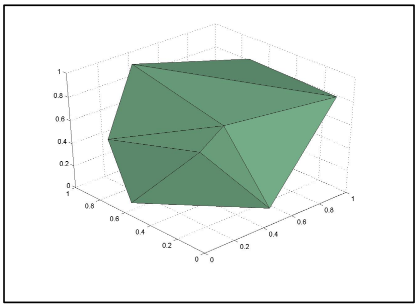
Feasible Region (cont.)

- **A linear inequality selects a half k-D space (i.e., the space region on one side of the hyper plan)**
    - E.g., consider the inequality $3x_1 + 2x_2 \leq 18$
      It is represented in the 2-D space by one of the half-spaces delimited by the previous line
        - The points belonging to this half space are those under and on the line. E.g., the point $P(x_1=1, x_2=1)$ satisfies the inequality



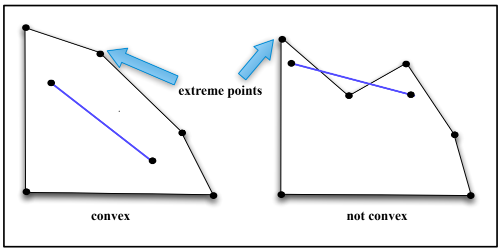
Feasible Region (cont.)

- In general, **if we have *k* variables, each linear equation defines a k-D plane, and each inequality a half-space on one side of the plane**
- Thus, **all the feasible solutions**, that are all the points in the *k* space that satisfy all the constraints, **are in the intersection of all the half-spaces identified by the linear inequalities**



Feasible Region (cont.)

- Property
  - **the feasible region is always convex**
- Convex set
  - **if two points *a* and *b* are in the set, then so it is *½*(a + b)***
- Extreme point
  - **a point in the set that can't be written as ½*(a + b) where a and b are two distinct points in the set**

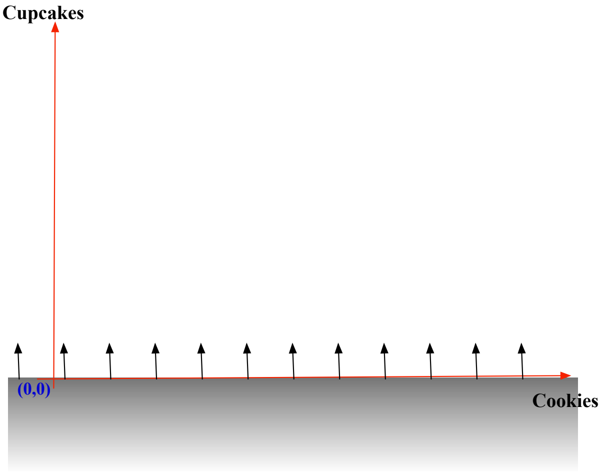


Feasible Region - Confectionery's problem

Feasible Region - Confectionery’s problem (cont.)

$x_2 \geq 0$

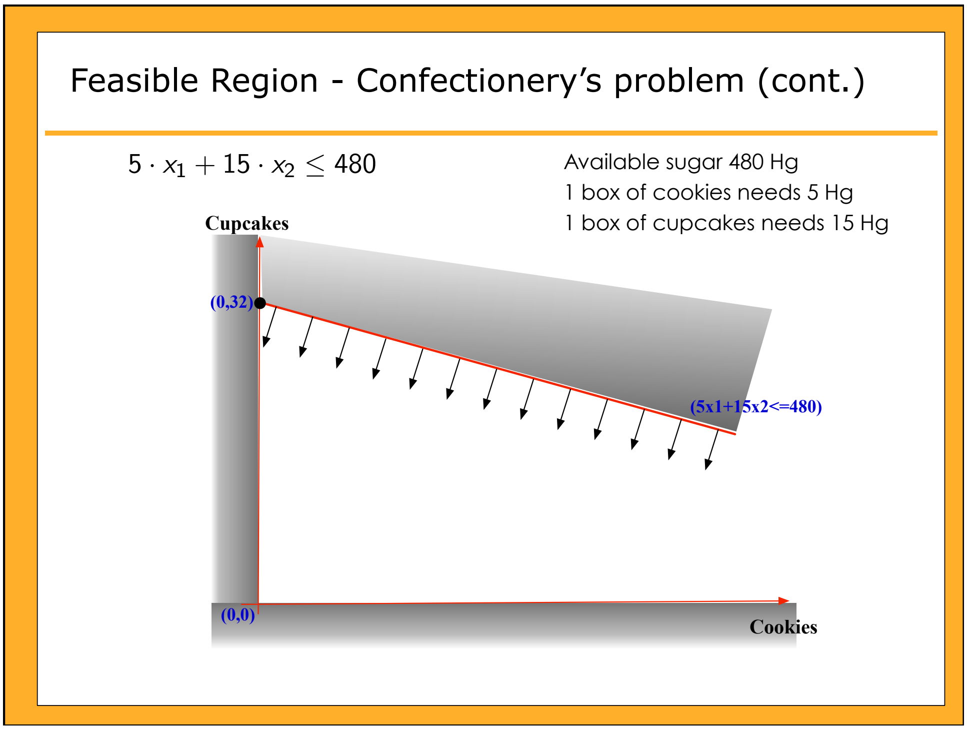
The number of boxes of cupcakes is positive



Feasible Region - Confectionery’s problem (cont.)

$x_1 \geq 0$

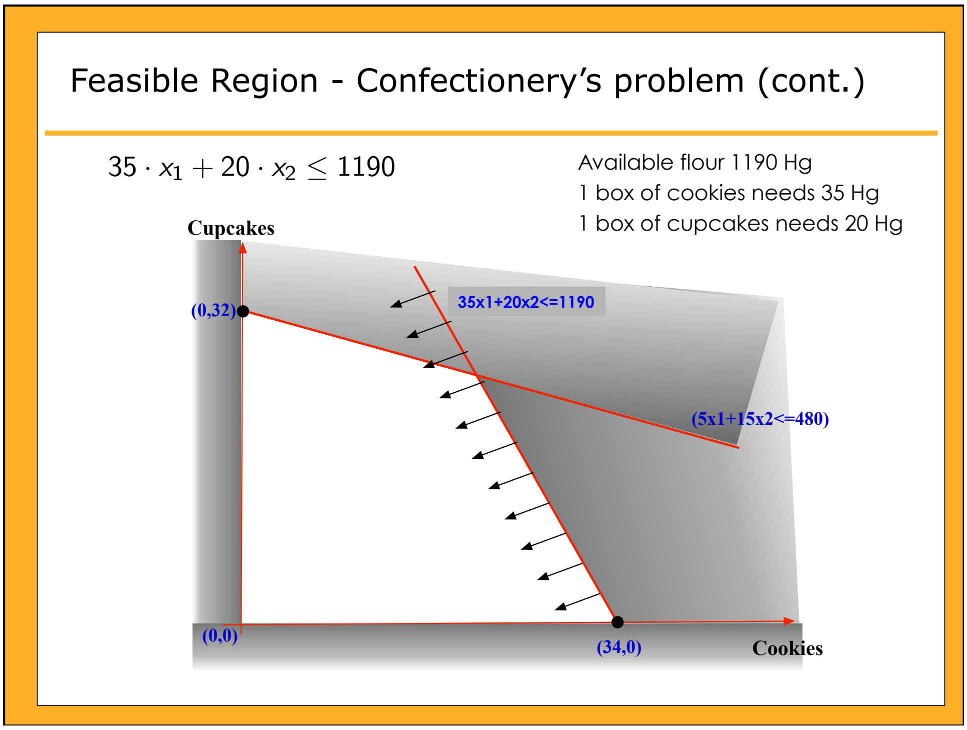
The number of boxes of cookies is positive



Feasible Region - Confectionery's problem (cont.)

$5 \cdot x_1 + 15 \cdot x_2 \leq 480$

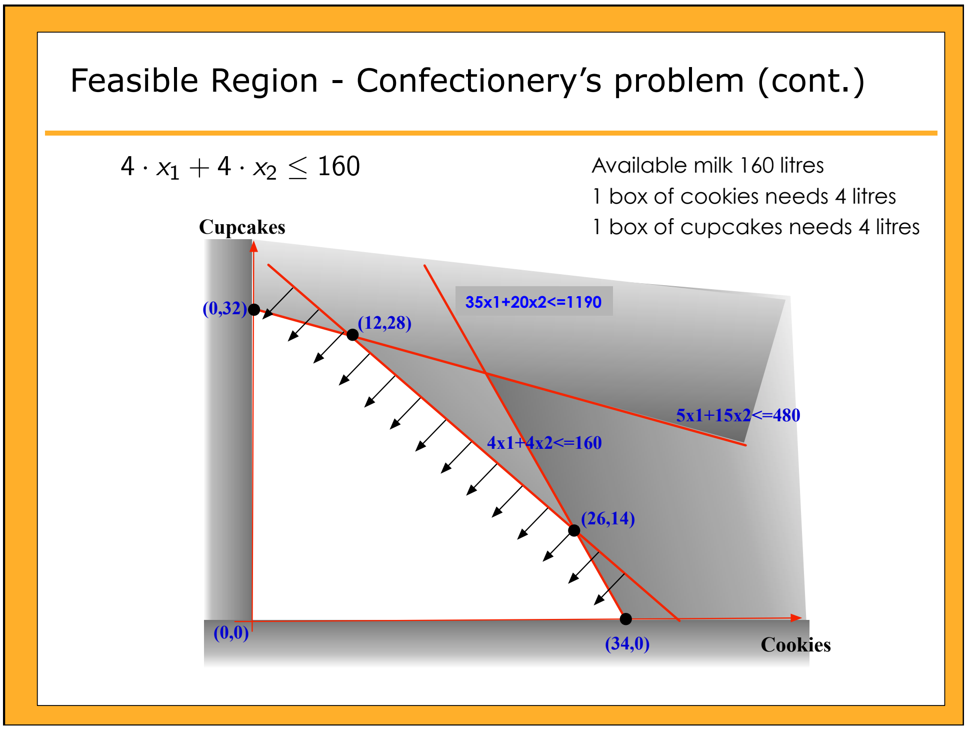
Available sugar 480 Hg
1 box of cookies needs 5 Hg
1 box of cupcakes needs 15 Hg



Feasible Region - Confectionery's problem (cont.)

$35 \cdot x_1 + 20 \cdot x_2 \leq 1190$

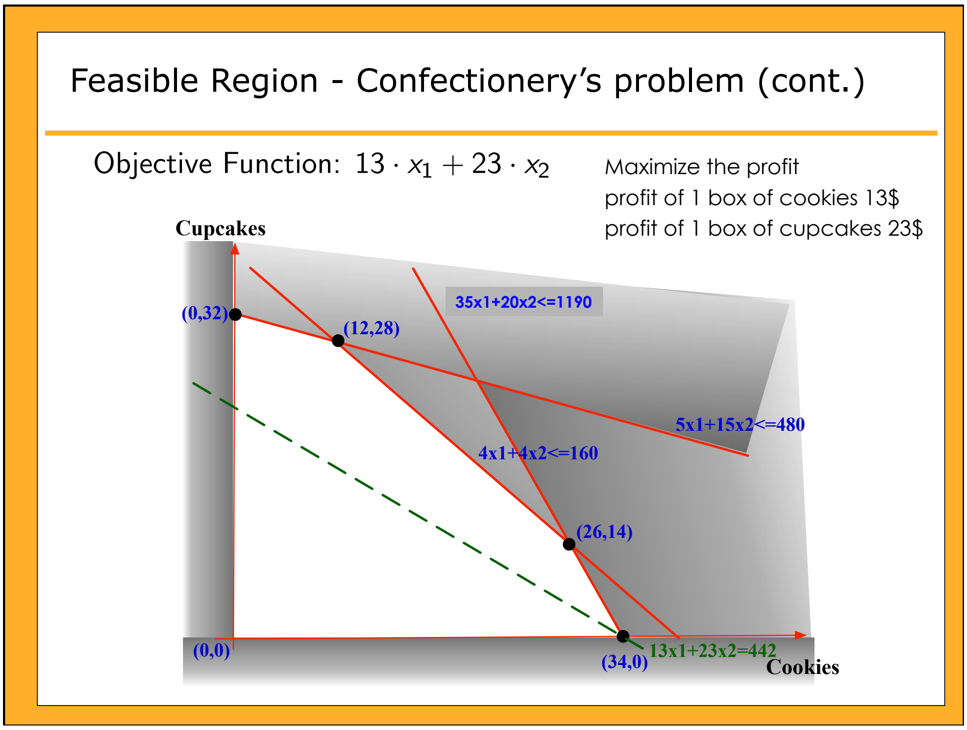
Available flour 1190 Hg
1 box of cookies needs 35 Hg
1 box of cupcakes needs 20 Hg



Feasible Region - Confectionery's problem (cont.)

$$4 \cdot x_1 + 4 \cdot x_2 \leq 160$$

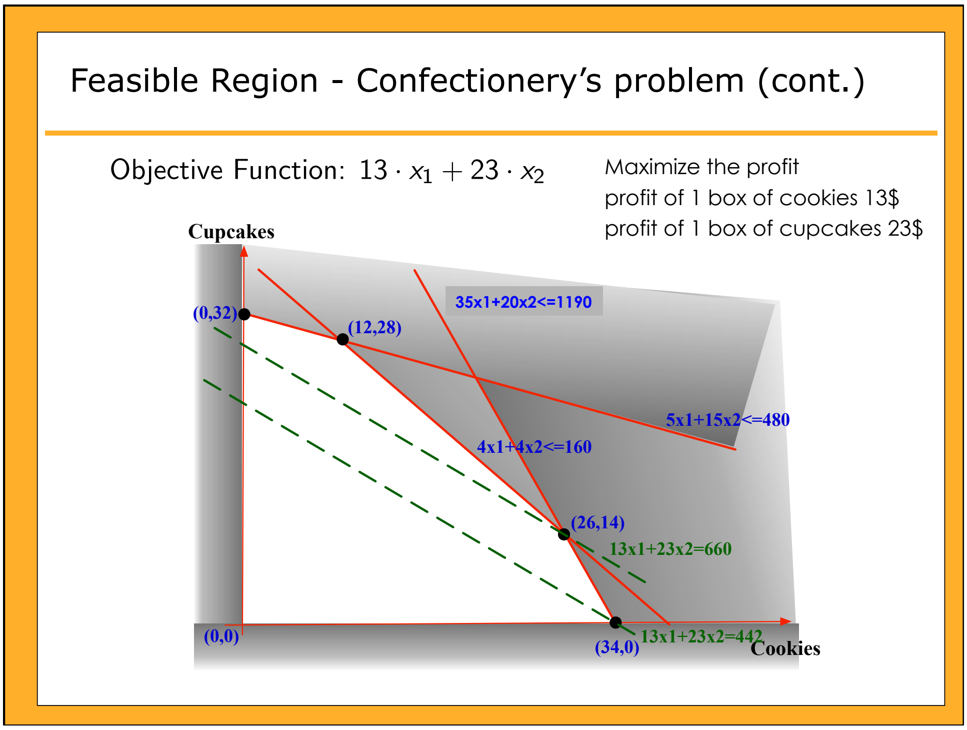
Available milk 160 litres
1 box of cookies needs 4 litres
1 box of cupcakes needs 4 litres



Feasible Region - Confectionery's problem (cont.)

Objective Function: $13 \cdot x_1 + 23 \cdot x_2$

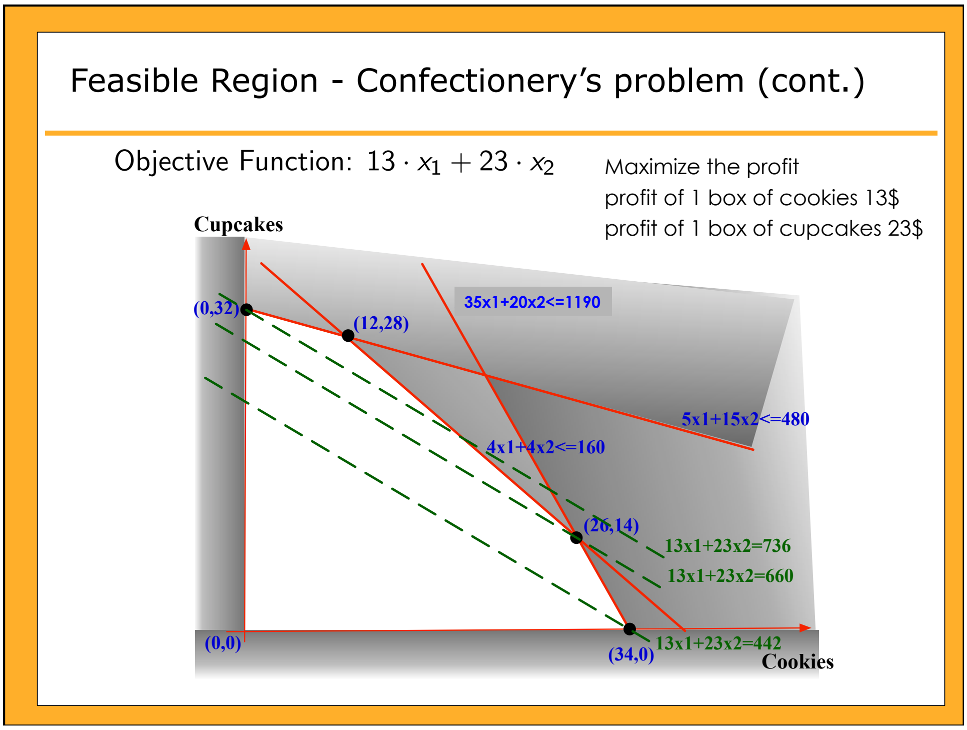
Maximize the profit
profit of 1 box of cookies 13$
profit of 1 box of cupcakes 23$



Feasible Region - Confectionery's problem (cont.)

Objective Function: $13 \cdot x_1 + 23 \cdot x_2$

Maximize the profit
profit of 1 box of cookies 13$
profit of 1 box of cupcakes 23$



Feasible Region - Confectionery's problem (cont.)

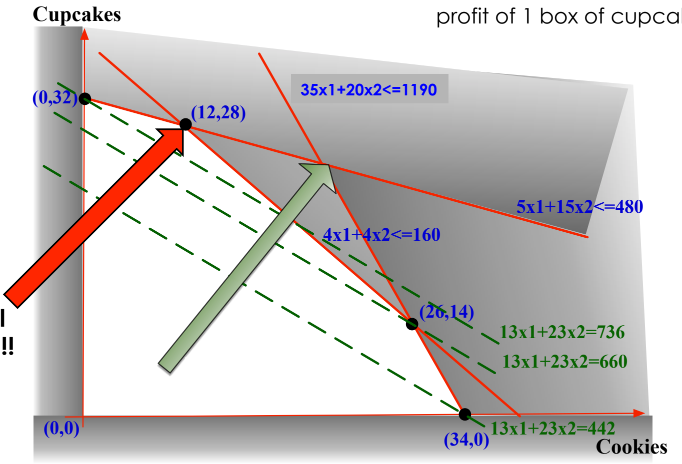
Objective Function: $13 \cdot x_1 + 23 \cdot x_2$

Maximize the profit
profit of 1 box of cookies 13$
profit of 1 box of cupcakes 23$

Feasible Region - Confectionery's problem (cont.)

Objective Function: $13 \cdot x_1 + 23 \cdot x_2$

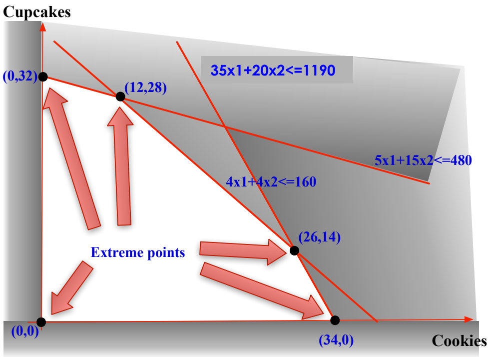
Maximize the profit
profit of 1 box of cookies 13$
profit of 1 box of cupcakes 23$



Optimal solution!!

Feasible Region - Confectionery's problem (cont.)

- Regardless of objective function coefficients, an **optimal solution occurs at an extreme point**



- This is **true for all LP problems**

Feasible Region – pension fund

- Let's change a bit the **mathematical model** and reduce to a 2-D space

x_t:Treasury notes, x_b: bonds, x_s: stocks.

x_t:Treasury notes, x_b: bonds, $30-(x_t+x_b)$: stocks.

$$\begin{array}{c} max\ 0.07x_t + 0.08x_b + 0.09x_s \\ x_t + x_b \geq 15 \\ x_b \leq 2 \cdot x_t \\ x_t + x_b + x_s = 30 \\ x_t, x_b, x_s \geq 3 \end{array} \Rightarrow \begin{array}{c} max\ 0.07x_t + 0.08x_b + 0.09(30 - \\ (x_t + x_b)) = 2.7 - .02x_t - 0.01x_b \\ x_t + x_b \geq 15 \\ x_b \leq 2 \cdot x_t \\ \text{X} \\ x_t, x_b \geq 3 \\ 30 - (x_t + x_b) \geq 3 \end{array}$$

Feasible Region – pension fund (cont.)

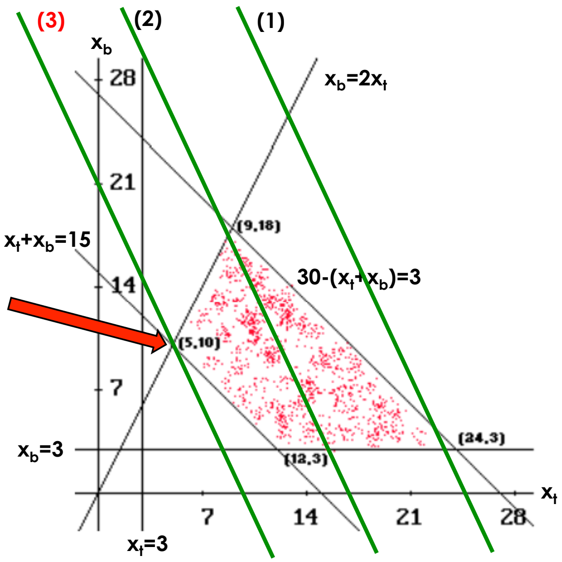
(1): $2.7-0.02*x_t-0.01*x_b=2.2$

(2): $2.7-0.02*x_t-0.01x_b=2.35$

(3): $2.7-0.02*x_t-0.01*x_b=2.5$ OPTIMUM!!

Optimal Solution:
$x_t=5$, $x_b=10$, $x_s=30-(5+10)=15$

Value = $2.7-0.02*x_t-0.01*x_b=$
$= 2.7-0.1-0.1=2.5$



Optimal solution

- Extreme point property
  - **if there exists an optimal solution to a linear program P, then there exists one that is an extreme point**
- Good news
  - **only need to consider finitely many possible solutions**
- Bad news
  - **the number of extreme points is an exponential function of the number of equalities/inequalities and the number of variables**
- Greedy property
  - **extreme point is optimal iff no neighboring extreme point is better**
- Generic algorithm for solving a LP problem
  1. **Start at some extreme point** (the origin is a good candidate)
  2. **Move to the neighboring point where the objective function increase (or decrease)**
  3. **Repeat until optimal**

Matrix-vector notation

- A linear function $c_1*x_1 + c_2*x_2$ can be written as the dot product of two vectors

$$c = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} \text{ and } x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

denoted by $c^T x$ where $c^T=(c_1\ c_2)$ is the transpose of c

- Similarly, linear constraints can be compiled into matrix-vector form

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \\ a_{3,1} & a_{3,2} \end{pmatrix} \text{ and } x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \text{ and } b = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$$

denoted by $Ax \leq b$, where each row of A corresponds to a constraint

Matrix-vector notation – Confectionery's problem

x_1: cookies, x_2: cupcakes

$\max\ 13x_1 + 23x_2$ → $\max\ c^T x$

$5x_1 + 15x_2 \leq 480$
$4x_1 + 4x_2 \leq 160$
$35x_1 + 20x_2 \leq 1190$
→ $Ax \leq b$

$x_1, x_2 \geq 0$ → $x_1, x_2 \geq 0$

$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ $c = \begin{pmatrix} 13 \\ 23 \end{pmatrix}$

$A = \begin{pmatrix} 5 & 15 \\ 4 & 4 \\ 35 & 20 \end{pmatrix}$ $b = \begin{pmatrix} 480 \\ 160 \\ 1190 \end{pmatrix}$

$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ $c = \begin{pmatrix} 13 \\ 23 \end{pmatrix}$

Is $c^T x = x c^T$?

NO!

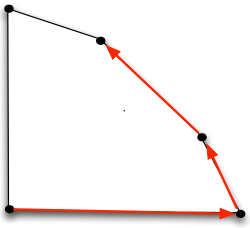
$$c^T x = \begin{pmatrix} 13 \\ 23 \end{pmatrix}^T * \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = (13\ 23) * \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 13x_1 + 23x_2$$

$$xc^T = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} * \begin{pmatrix} 13 \\ 23 \end{pmatrix}^T = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} * (13\ 23) = \begin{pmatrix} 13x_1 & 23x_1 \\ 13x_2 & 23x_2 \end{pmatrix}$$

Solving a LP problem – Simplex Algorithms

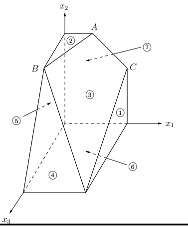
- Developed in 1947 by **George Dantzig**
- One of the most successful algorithms of all time
- Generic algorithm

```
let v be any vertex of the feasible region
while there is a neighbor v' of v with better objective value
    set v = v'
```



Simplex method (cont.)

- Each **vertex** is the unique **point at which some subset of hyper planes meet**
- If we have ***m* constraints** and ***n* variables**, **each vertex is defined by *n* inequalities satisfied with equality**
  - E.g., vertex A is the sole point where constraints (2), (3), and (7) are satisfied with equality
- Two vertices are **neighbors** if they have ***n−1* defining inequalities in common**
  - E.g., vertices A and C share the two defining inequalities (3) and (7)



$$\max \quad x_1 + 6x_2 + 13x_3$$

$$\begin{aligned} x_1 &\le 200 & ① \\ x_2 &\le 300 & ② \\ x_1 + x_2 + x_3 &\le 400 & ③ \\ x_2 + 3x_3 &\le 600 & ④ \\ x_1 &\ge 0 & ⑤ \\ x_2 &\ge 0 & ⑥ \\ x_3 &\ge 0 & ⑦ \end{aligned}$$

Simplex method (cont.)

- At each iteration the simplex algorithm must
  1. check whether the current vertex is optimal
  2. determine where to move next
- The above tasks 1 and 2 are easy if the vertex is the origin
- Why is the origin so convenient?

$$\boxed{\begin{array}{c} \max\ \mathbf{c}^T\mathbf{x} \\ \mathbf{A}\mathbf{x} \leq \mathbf{b} \\ \mathbf{x} \geq 0 \end{array}}$$

  - Suppose we have a generic LP problem and that the origin is in the feasible region
  - The origin is a vertex since it is the unique point at which the ***n* inequalities {$x_1 \geq 0, \ldots, x_n \geq 0$} are tight**
  - The **origin is optimal iff all $c_i \leq 0$**
    - If all $c_i \leq 0$ then we can't hope a better objective value because **x** ≥ 0
    - If there is at least a $c_i > 0$ then we can increase the objective function by increasing the variable x_i
  - So if the origin is not optimal we can **move by increasing some x_i for which $c_i > 0$ until we hit some other constraint**
    - i.e., until **some constraints** other than **x** ≥ 0 **becomes tight**
  - At that point, we again have exactly *n* tight inequalities, so we are at a new vertex
  - We then **transform the coordinate systems** to move the new vertex to the origin!