

Math Preliminary and Geometry of LP: Feasible Regions, Corners and Improving Directions

Yinyu Ye

Stanford University and CUHKSZ (Sabbatical Leave)

Currently Visiting CUHK and HK PolyU

<https://web.stanford.edu/class/msande211x/handout.shtml>

Read Chapter 2.3-2.5, 4.1, Appendix B

Vectors and Matrices

- **Column or Row Vector:** point $\mathbf{a} \in \mathbb{R}^n$, j th element: a_j
- **Transpose:** \mathbf{a}^T .
- **Matrix:** $\mathbf{A} \in \mathbb{R}^{m \times n}$, i th row: \mathbf{a}_i , j th column: \mathbf{a}_j , ij th element: a_{ij}
- **All one vector:** \mathbf{e} or $\mathbf{1}$, **All-zero matrix:** $\mathbf{0}$, and **identity matrix:** \mathbf{I}
- **Diagonal matrix:** $\mathbf{X} = \text{Diag}(x)$
- **Symmetric matrix:** $\mathbf{Q} = \mathbf{Q}^T$
- **Positive Definite (PD):** iff $\mathbf{x}^T \mathbf{Q} \mathbf{x} > 0$, for all $\mathbf{x} \neq \mathbf{0}$
- **Positive Semi-definite (PSD):** iff $\mathbf{x}^T \mathbf{Q} \mathbf{x} \geq 0$, for all \mathbf{x}

Matrix Inverse

- **Inverse of a square matrix: A^{-1} such that $A^{-1}A=I$.**

Application of inverse:

Suppose there are b unit resources, and a units of the resources can be used to produce one-unit product, and each unit product can sell for $\$c$. How much does each unit resource worth?

$$ax = b, x = a^{-1}b, cx = ca^{-1}b = (ca^{-1})b,$$

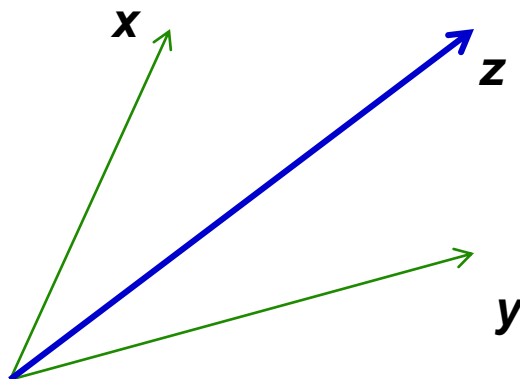
Now consider multi-product and multi-resources:

$$Ax = b, x = A^{-1}b, c^T x = c^T A^{-1}b = (c^T A^{-1})b$$

That is, the vector $c^T A^{-1}$ contains the (shadow) prices for each resources, respectively.

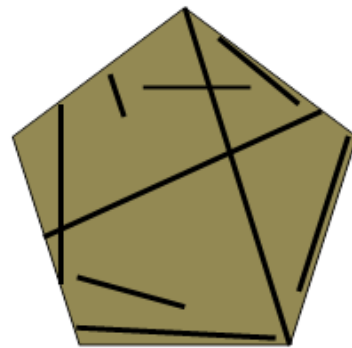
Affine, Convex and Conic Combination

- When \mathbf{x} and \mathbf{y} are two distinct points in \mathbb{R}^n and α runs over \mathbb{R} , $\{\mathbf{z} : \mathbf{z} = \alpha\mathbf{x} + (1-\alpha)\mathbf{y}\}$ is the line determined by \mathbf{x} and \mathbf{y} , called the **affine combination** of \mathbf{x} and \mathbf{y} .
- When $0 \leq \alpha \leq 1$, \mathbf{z} is called the **convex combination** of \mathbf{x} and \mathbf{y} and it is the **line segment** between \mathbf{x} and \mathbf{y}
- When $\alpha \geq 0$ and $\beta \geq 0$, $\{\mathbf{z} : \mathbf{z} = \alpha\mathbf{x} + \beta\mathbf{y}\}$ is called the **conic combination** of \mathbf{x} and \mathbf{y} and it is the **ray** between \mathbf{x} and \mathbf{y}

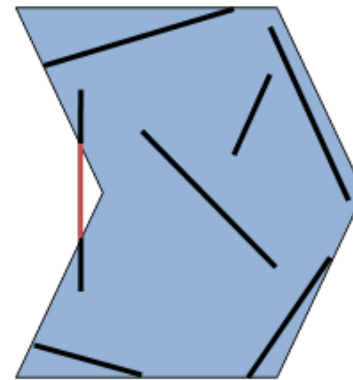


Convex Sets

- Set Ω is said to be a **convex set** iff for every $\mathbf{x}^1, \mathbf{x}^2 \in \Omega$ and every real number $\alpha \in [0, 1]$, the **convex combination** point $\alpha \mathbf{x}^1 + (1 - \alpha) \mathbf{x}^2 \in \Omega$.



CONVEX



NOT CONVEX

- The **convex hull** of a set Ω is the intersection of all convex sets containing Ω
- Intersection** of convex sets is convex
- Unit-disk** $\{(x_1, x_2): (x_1)^2 + (x_2)^2 \leq 1\}$ is a convex set.
- Ellipsoid** $\{\mathbf{x}: \mathbf{x}^T \mathbf{Q} \mathbf{x} \leq 1\}$, where \mathbf{Q} is PD, is a convex set.

Convex and Concave Functions

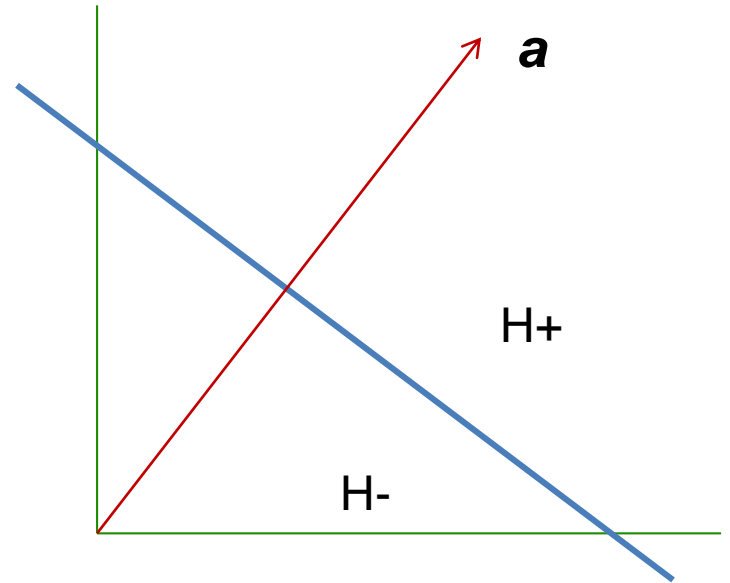
- f is a **convex function** iff for $0 \leq \alpha \leq 1$,
 $f(\alpha \mathbf{x} + (1 - \alpha) \mathbf{y}) \leq \alpha f(\mathbf{x}) + (1 - \alpha)f(\mathbf{y})$
- f is a **concave function** iff $-f$ is a **convex function**
- f is a **strictly convex function** iff for $\mathbf{x} \neq \mathbf{y}$,
 $f(0.5\mathbf{x} + 0.5\mathbf{y}) < 0.5 f(\mathbf{x}) + 0.5 f(\mathbf{y})$
- **The minimizer** of a **strictly convex function** is unique if it exists
- Gradient vector $\nabla f(\mathbf{x}) = (\partial f / \partial x_i)$: it is the **steepest ascent direction** of the function value;
- Hessian matrix $\nabla^2 f(\mathbf{x}) = (\partial^2 f / \partial x_i \partial x_j)$: the function $f(\cdot)$ is **convex** (strictly convex) iff its Hessian matrix is PSD (PD) everywhere.
- Sample **convex** functions: $\|\mathbf{x}\|$, $\|\mathbf{x}\|^2$, $\log(1 + e^{\mathbf{a}'\mathbf{x}})$
- **linear function** $\mathbf{c}^T \mathbf{x}$ is both convex and concave
- **Quadratic function** $\mathbf{x}^T \mathbf{Q} \mathbf{x}$ is convex iff \mathbf{Q} is positive semidefinite.

Verification of Convex Sets and Convex Functions

- The **epigraph** $\{(z, \mathbf{x}): c(\mathbf{x}) \leq z\}$ is a convex set iff $c(\cdot)$ is a convex function.
- The lower **level set** $\{\mathbf{x}: c(\mathbf{x}) \leq 0\}$ is a convex set if $c(\cdot)$ is a convex function.
- The upper **level set** $\{\mathbf{x}: c(\mathbf{x}) \geq 0\}$ is a convex set if $c(\cdot)$ is a concave function.
- **Sum** of convex functions is **convex**.
- **Sum** of concave functions is **concave**.
- The **composite** function $f(\varphi(\mathbf{x}))$ is convex if $f(\cdot)$ is a monotonically increasing & convex function and $\varphi(\mathbf{x})$ is a convex function.
 - $\exp(x^2+y^2)$
- $\max_i(f_i(\mathbf{x}))$ is convex if $f_i(\mathbf{x})$ is convex for all i .
- **Convex Optimization**: minimize a **convex** (or maximize a concave) function subject to a **convex** constraint set.

Hyperplane and Half-Spaces

$$\mathbf{H} = \left\{ \mathbf{x} : \mathbf{a}^T \mathbf{x} = \sum_{j=1}^n a_j x_j = b \right\}$$
$$\mathbf{H}^+ = \left\{ \mathbf{x} : \mathbf{a}^T \mathbf{x} = \sum_{j=1}^n a_j x_j \geq b \right\}$$
$$\mathbf{H}^- = \left\{ \mathbf{x} : \mathbf{a}^T \mathbf{x} = \sum_{j=1}^n a_j x_j \leq b \right\}$$



Each of them is a convex set or region, and \mathbf{a} is called the normal direction or slope vector.

They are all **convex** sets.

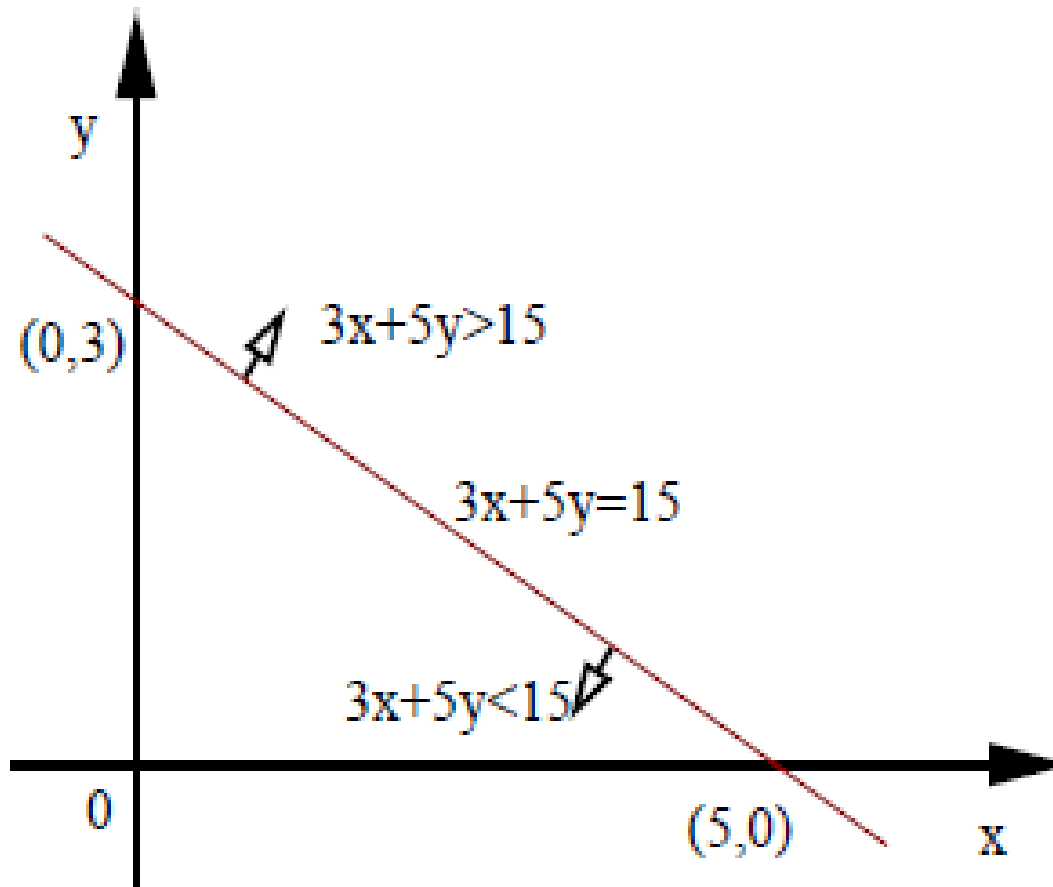


Figure 1: Plane and Half-Spaces

LP Feasible Region in the Inequality Form

\mathbf{x} simultaneously satisfy

$$\mathbf{a}_1^T \mathbf{x} = \sum_{j=1}^n a_{1j} x_j \leq b_1$$

$$\mathbf{a}_2^T \mathbf{x} = \sum_{j=1}^n a_{2j} x_j \leq b_2$$

...

$$\mathbf{a}_m^T \mathbf{x} = \sum_{j=1}^n a_{mj} x_j \leq b_m$$

This is the intersection of the m Half-spaces, and it is a **convex** (polyhedron) set

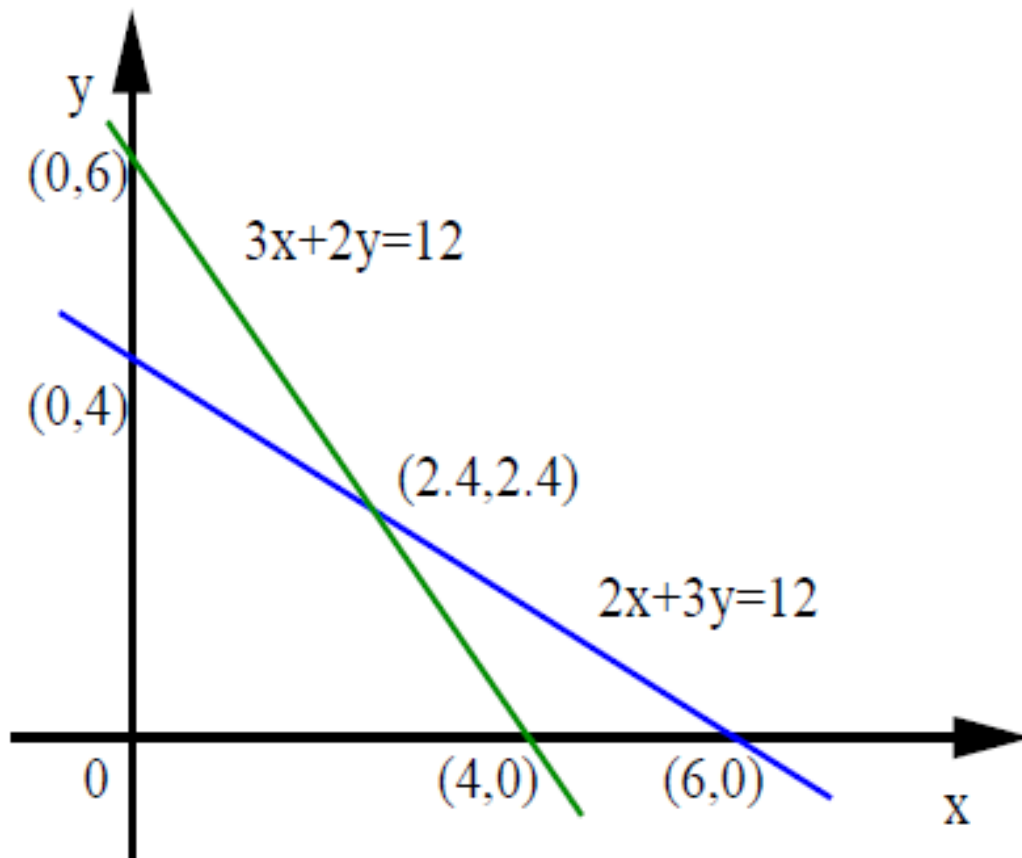


Figure 2: System of Linear Equations

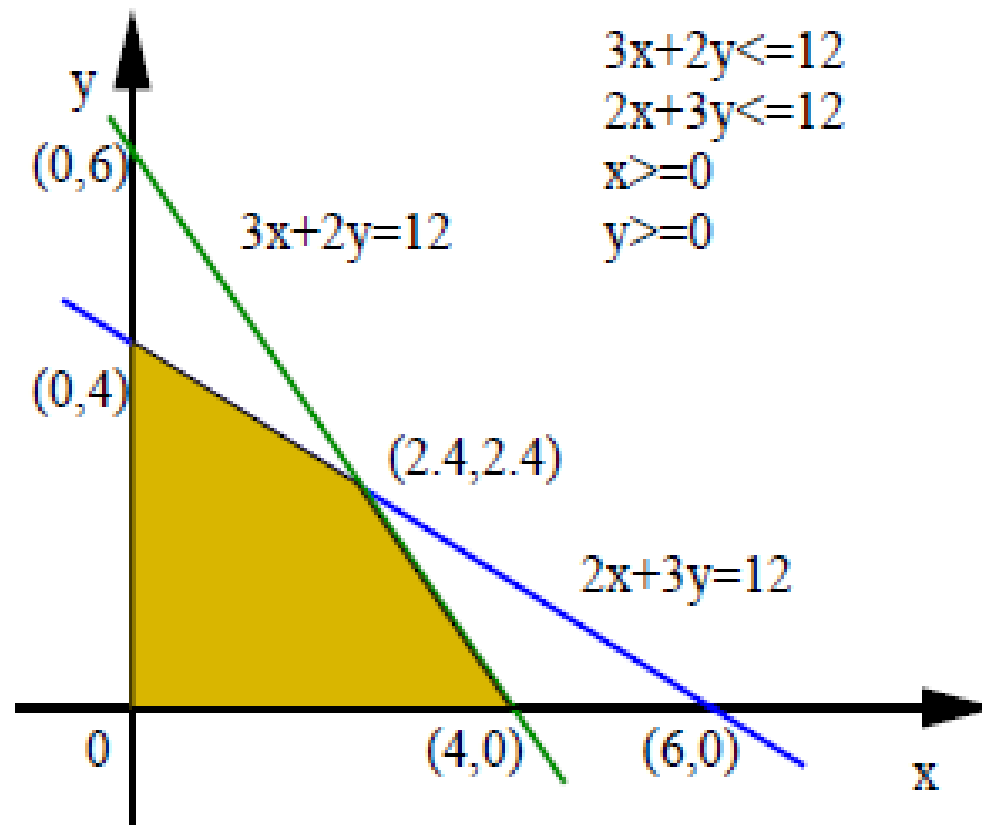
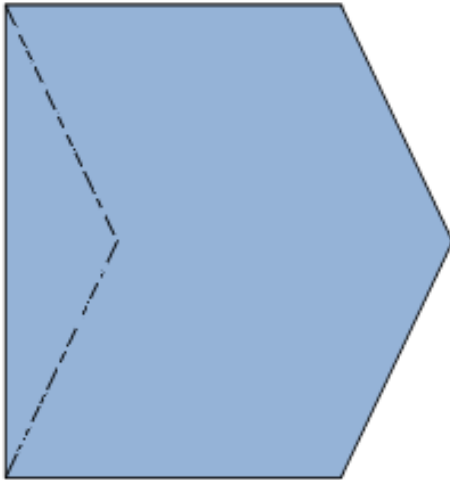


Figure 3: System of Linear Inequalities

Corner or Extreme Points

Convex Hull:



The convex hull of a region, R , is the smallest convex region containing it.

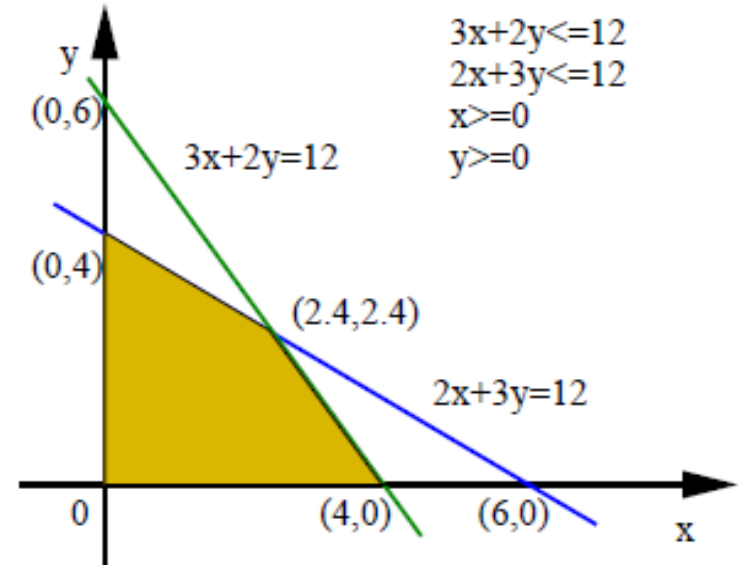


Figure 3: System of Linear Inequalities

Extreme Points: A point in the set that is not on the line segment (convex combination) of other two different points in the convex hull of the set. For LP in inequality form, an extreme point is the intersection of n hyperplanes associated with the inequality constraints that is also feasible – called **Basic Feasible Solution**.

Feasible Direction I

Direction Vector: A direction is notated by a vector d

It is always associated with a given point x

Together a point and a direction vector define a ray:

$$x + \epsilon d, \text{ for all } \epsilon > 0$$

where d and αd are considered the **same** direction for all $\alpha > 0$

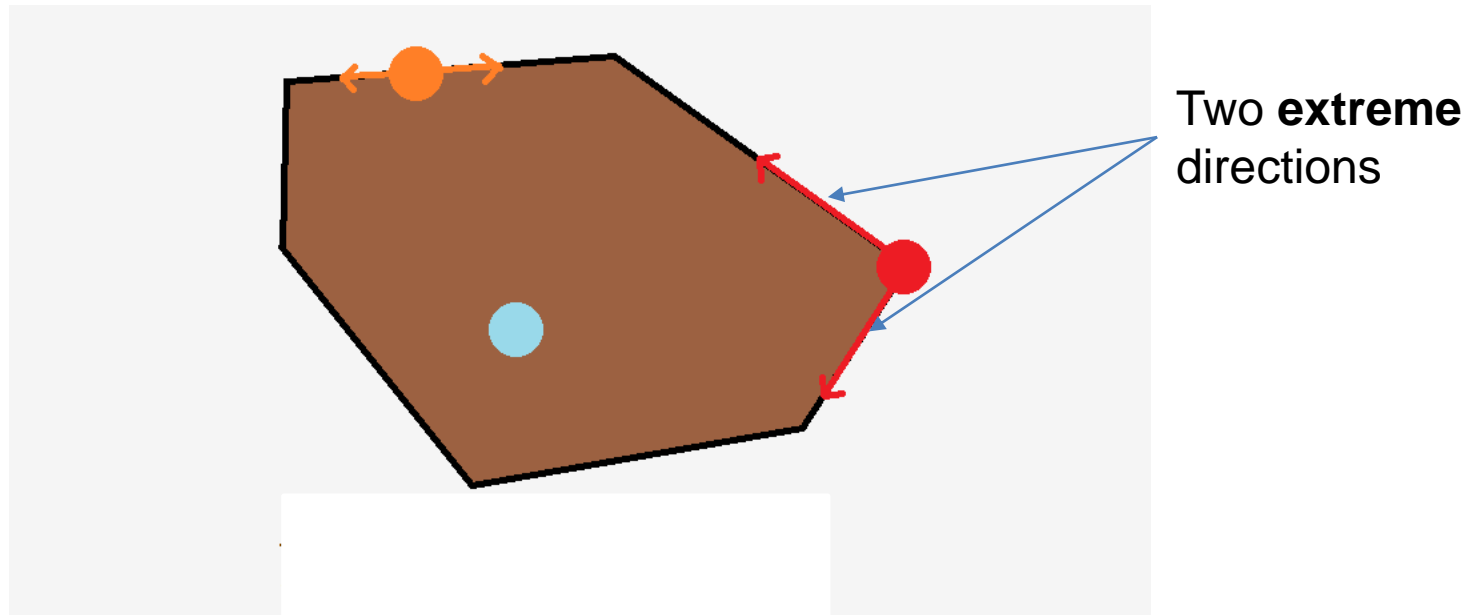
Feasible Direction: A direction, d , is said to be “feasible” (relative to a given feasible point x) if $x + \epsilon d$ is feasible for some $\epsilon > 0$ and small enough.

Extreme Feasible direction: direction to its nearby extreme points.

For LP, all feasible directions at a feasible point form a **convex (cone) set: conic combination** of feasible (extreme) directions from the point.

Feasible Direction II

Feasible direction d is location-dependent of the point:

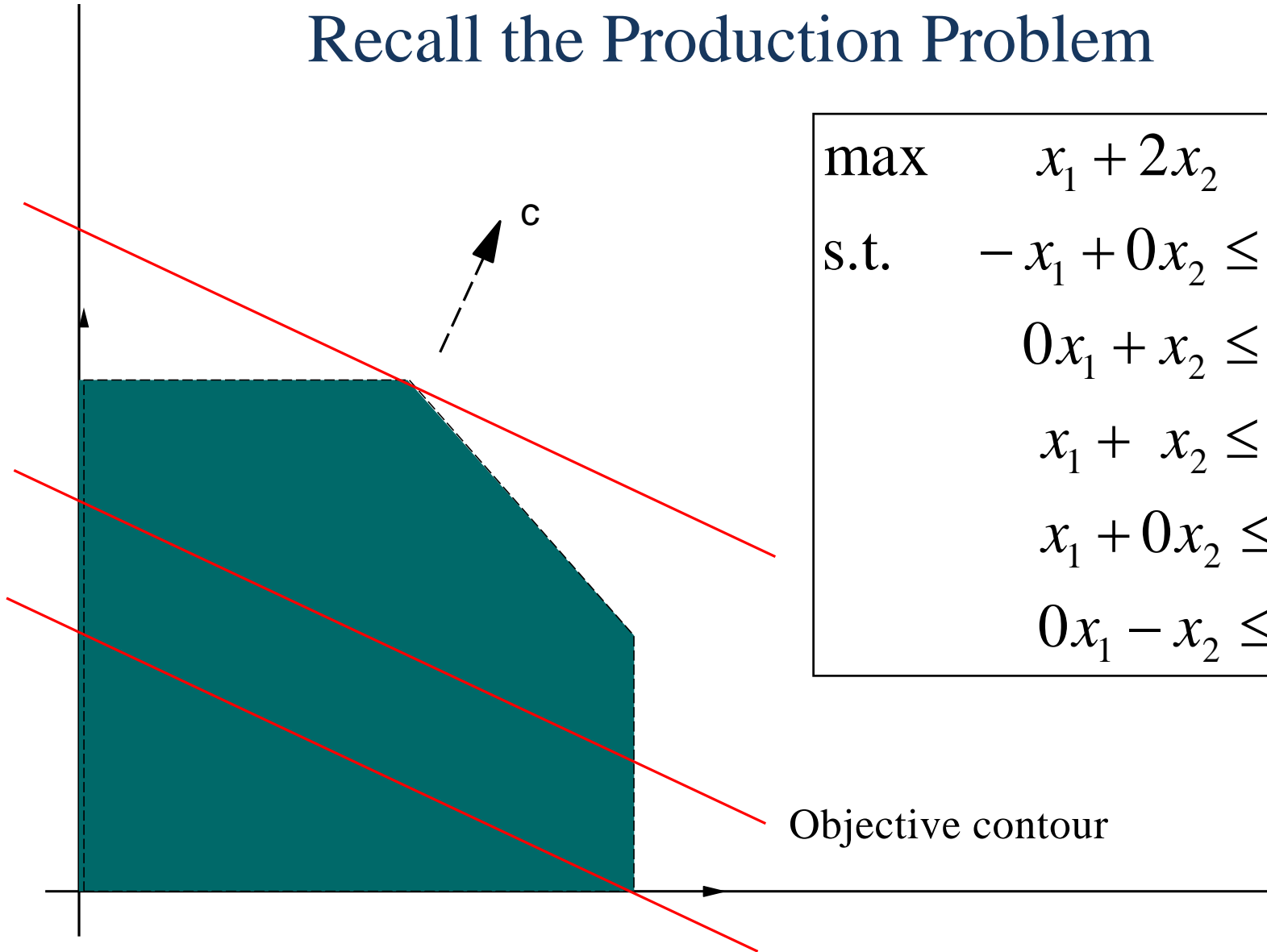


Interior Point is a point x where every direction is feasible

LP Problem in Inequality Form

$$\begin{array}{ll} \max & c^T x = \sum_{j=1}^n c_j x_j \\ \text{s.t.} & a_1^T x = \sum_{j=1}^n a_{1j} x_j \leq b_1 \\ & a_2^T x = \sum_{j=1}^n a_{2j} x_j \leq b_2 \\ & \dots \\ & a_m^T x = \sum_{j=1}^n a_{mj} x_j \leq b_m \end{array}$$

Recall the Production Problem



$$\begin{aligned} \max \quad & x_1 + 2x_2 \\ \text{s.t.} \quad & -x_1 + 0x_2 \leq 0 \\ & 0x_1 + x_2 \leq 1 \\ & x_1 + x_2 \leq 1.5 \\ & x_1 + 0x_2 \leq 1 \\ & 0x_1 - x_2 \leq 0 \end{aligned}$$

Objective contour

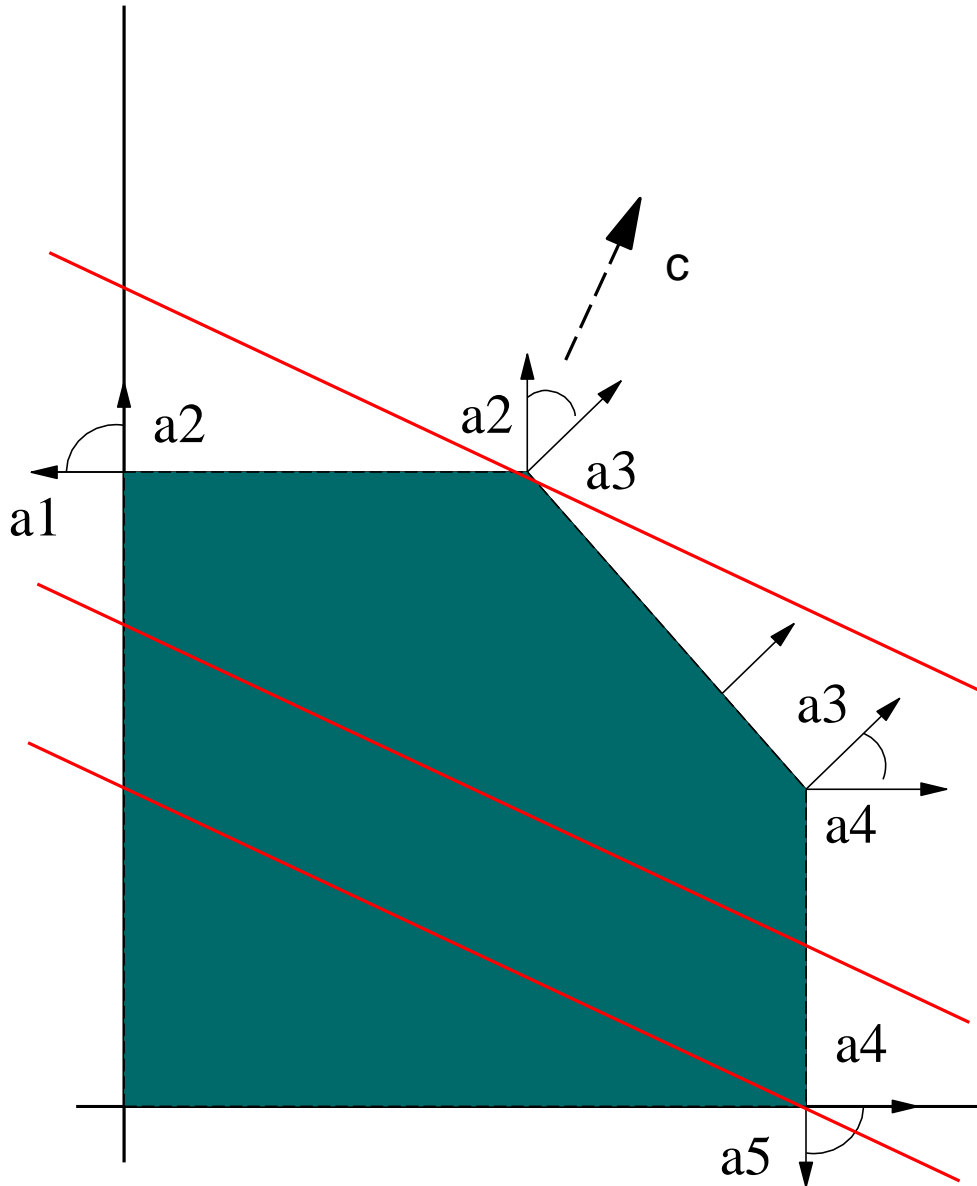
Facts of Linear Programming

All LP problems fall into one of three cases:

- Problem is **infeasible**: Feasible region is empty.
- Problem is **unbounded**: Feasible region is unbounded towards the optimizing direction.
- Problem is **feasible and bounded**; and in this case:
 - there exists an **optimal solution or optimizer**.
 - There may be a **unique** optimizer or **multiple** optimizers.
 - All optimizers form a convex set, and they are on a **face** of the feasible region.
 - There is always at least one **corner (extreme)** optimizer if the feasible region has a corner point.

LP is a convex optimization problem; moreover, **local optimality implies global optimality**

Optimality Certification using Local Information



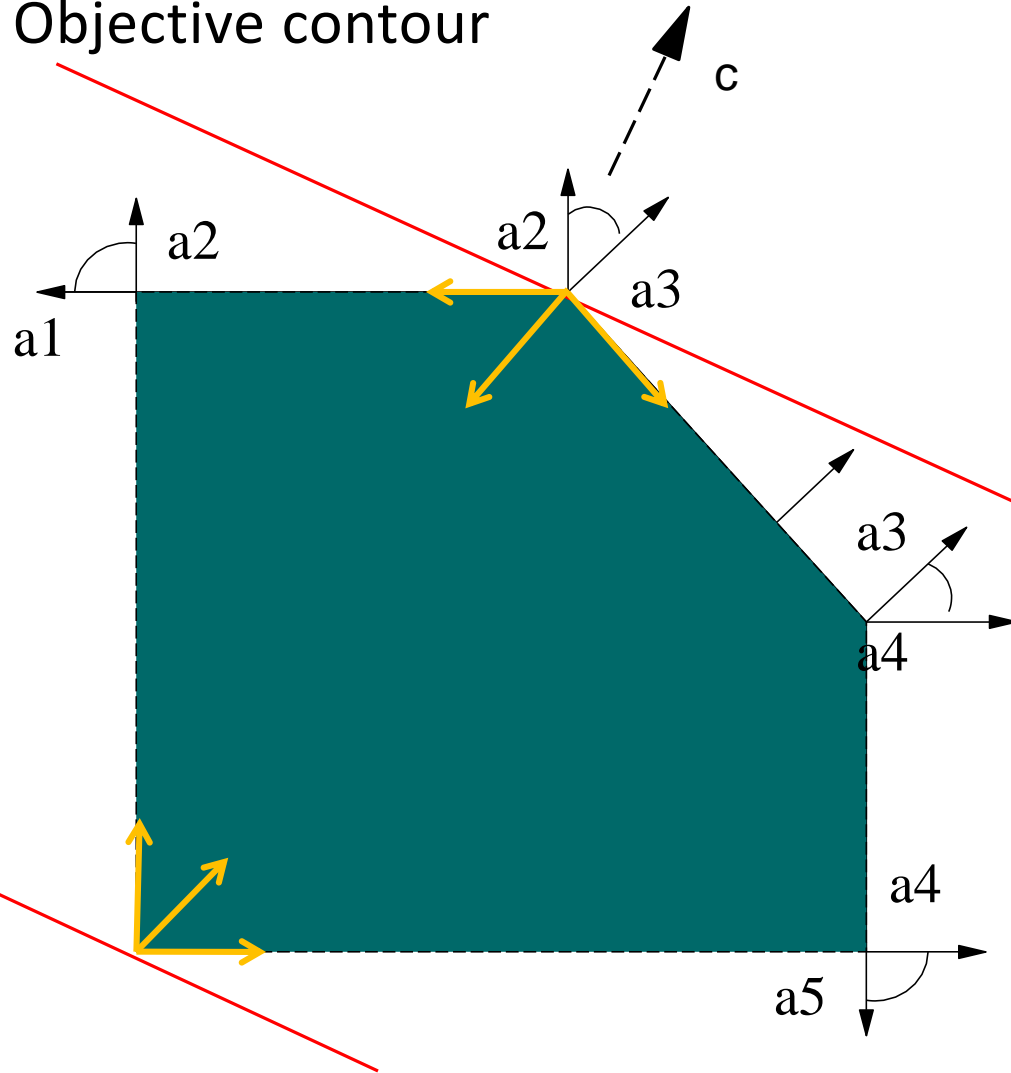
$$\begin{array}{ll} \max & x_1 + 2x_2 \\ \text{s.t.} & -x_1 + 0x_2 \leq 0 \\ & 0x_1 + x_2 \leq 1 \\ & x_1 + x_2 \leq 1.5 \\ & x_1 + 0x_2 \leq 1 \\ & 0x_1 - x_2 \leq 0 \end{array}$$

All (extreme) feasible directions are un-improving

Objective contour

Feasible Directions at the Optimal Corner

Objective contour



At the optimal corner, c must be a **conic combination** of a_2 and a_3 , the two normal direction vectors of the intersection constraints. Or it has an obtuse angle with any (extreme) feasible directions.

Recall conic comb means there are **multipliers** $\alpha_2 \geq 0$ and $\alpha_3 \geq 0$, such as

$$c = \alpha_2 a_2 + \alpha_3 a_3,$$

where multipliers α 's can be seen as the decomposed unit resultant force in c direction.

How to Certify a Corner Point being an Optimizer

- Every feasible direction at the point is an un-**improving** (non-ascent in this case) direction, that is, $\mathbf{c}^T \mathbf{d} \leq 0$ in this case (where \mathbf{c} is the steepest ascent direction of the objective function).
- Recall at the optimal corner, objective direction \mathbf{c} is a conic combination of the normal directions, \mathbf{a}_{j_1} and \mathbf{a}_{j_2} , at the corner point, that is, there are **multipliers** $\alpha_{j_1} \geq 0$ and $\alpha_{j_2} \geq 0$, such as
$$\mathbf{c} = \alpha_{j_1} \mathbf{a}_{j_1} + \alpha_{j_2} \mathbf{a}_{j_2},$$
 in the 2-dimensional case.
- **Theorem:** For LP of inequality form in n-dimensional case, a feasible corner is maximal if and only if its objective vector
$$\mathbf{c} = \alpha_1 \mathbf{a}_{i_1} + \dots + \alpha_n \mathbf{a}_{i_n}$$
with nonnegative multipliers α 's where vectors $\mathbf{a}_{i_1}, \dots, \mathbf{a}_{i_n}$ are the normal directions of hyperplanes associated with the corner point.
- This is the essential idea led to the **Simplex** method by Dantzig: if \mathbf{c} has an acute angle with a norm vector, then go along the extreme feasible direction, while stay feasible, till hit the next corner point...

Simplex Method

George B. Dantzig's **Simplex Method** for linear programming stands as one of the most significant algorithmic achievements of the 20th century. It is now over 60 years old and still going strong.

The basic idea of the simplex method is to confine the search to **corner points** of the feasible region (of which there are only **finitely** many) in a most intelligent way.

The key for the simplex method is to make computers **see** corner points; and the key for interior-point methods is to **stay** in the interior of the feasible region.

