Geometry of LP: Feasible Regions, Feasible and Improving Directions, and Optimality Test

Yinyu Ye
Department of Management Science and Engineering
Stanford University
Stanford, CA 94305, U.S.A.

Read Chapter 2.3-2.5, 4.1, Appendix B

Vectors and Matrices

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

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

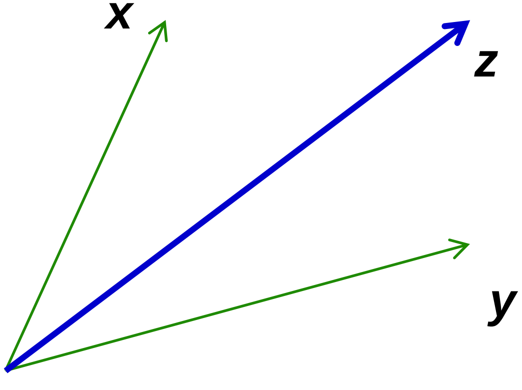
Now consider multi-product and multi-recourses:

$$Ax=b,\ \ x=A^{-1}b,\ c^Tx=c^TA^{-1}b=(c^TA^{-1})b$$

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

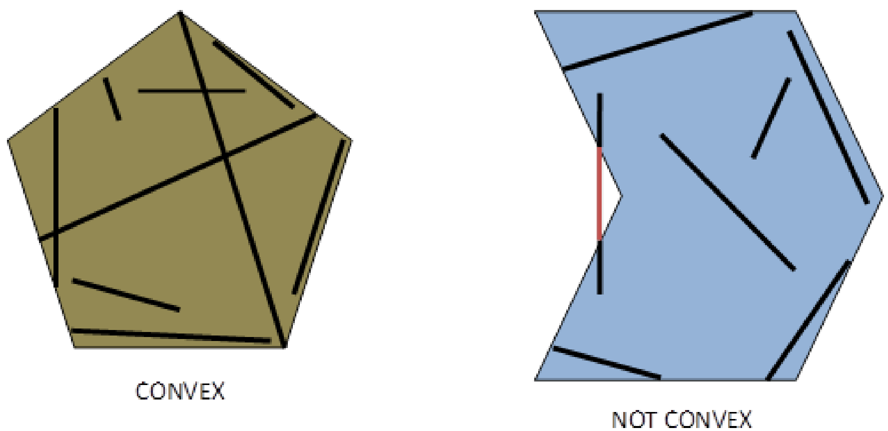
Affine, Convex and Conic Combination

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



Convex Sets

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



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

Convex and Concave Functions

- f is a **convex function** iff for $0 \le \alpha \le 1$,

  $f(\alpha \boldsymbol{x} + (1 - \alpha)\boldsymbol{y}) \le \alpha f(\boldsymbol{x}) + (1 - \alpha) f(\boldsymbol{y})$
- f is a **concave function** iff $-f$ is a **convex function**
- f is a **strictly** convex function iff for $\boldsymbol{x} \neq \boldsymbol{y}$,

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

Verification of Convex Sets and Convex Functions

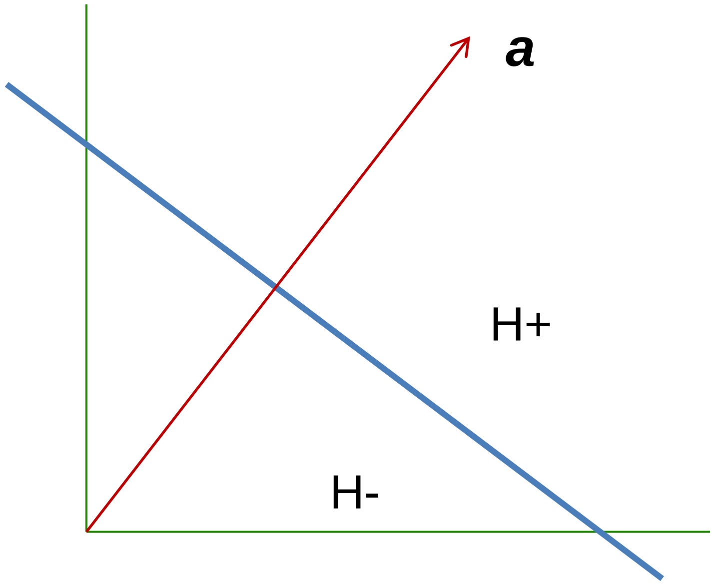
- The **epigraph** $\{(z,\boldsymbol{x}): c(\boldsymbol{x}) \le z\}$ is a convex set iff $c(.)$ is a convex function.
- The lower **level set** $\{\boldsymbol{x}: c(\boldsymbol{x}) \le 0\}$ is a convex set if $c(.)$ is a convex function.
- The upper **level set** $\{\boldsymbol{x}: c(\boldsymbol{x}) \ge 0\}$ is a convex set if $c(.)$ is a concave function.
- **Sum** of convex functions is **convex.**
- **Sum** of concave functions is **concave**
- The **composite** function : $f(\varphi(\boldsymbol{x}))$ is convex if $f(.)$ is a monotonically increasing&convex function and $\varphi(\boldsymbol{x})$ is a convex function.
  - $exp(x^2+y^2)$
- $max_i(f_i(x))$ is convex if $f_i(x)$ 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

$$\mathrm{H} = \{x : a^T x = \sum_{j=1}^{n} a_j x_j = b\}$$

$$\mathrm{H}^+ = \{x : a^T x = \sum_{j=1}^{n} a_j x_j \geq b\}$$

$$\mathrm{H}^- = \{x : a^T x = \sum_{j=1}^{n} a_j x_j \leq b\}$$



Each of them is a convex set or region, and ***a*** is called the normal direction or slope vector.

They are all **convex** sets.

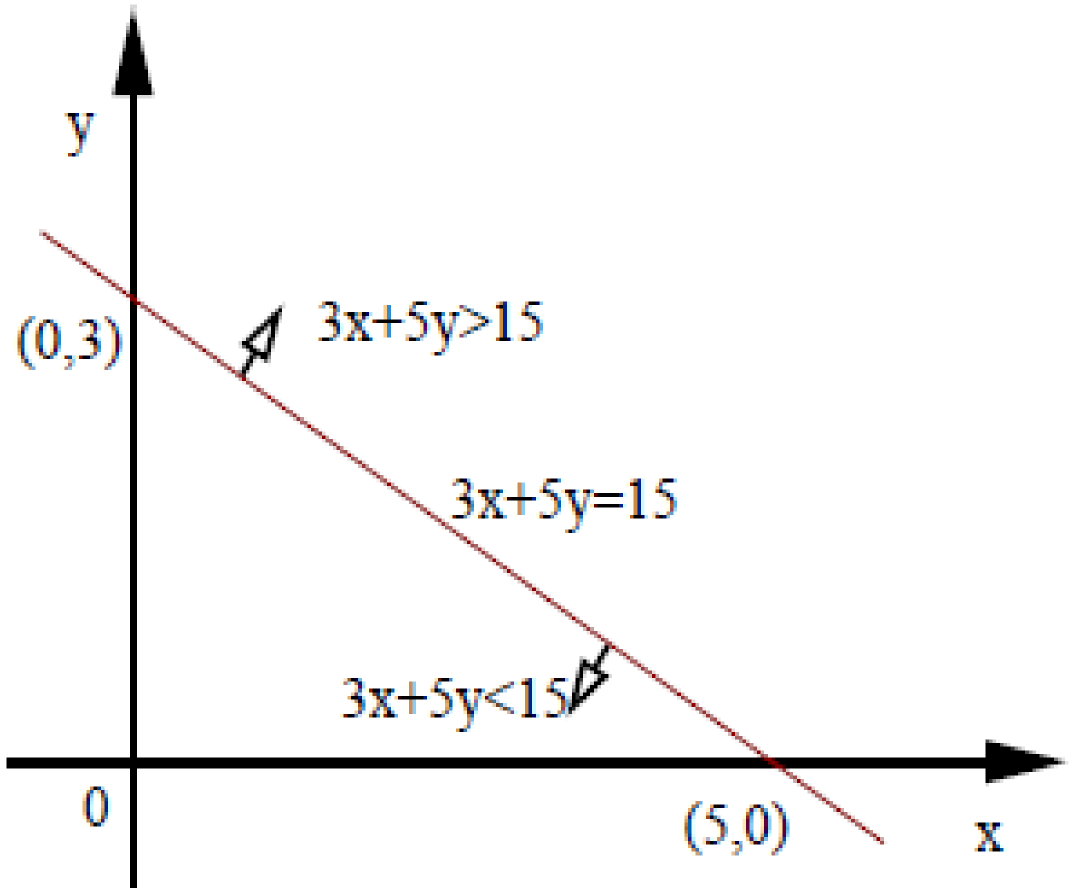


Figure 1: Plane and Half-Spaces

LP Feasible Region in the Inequality Form

x simultaneously satisfy

$$a_1^T x = \sum_{j=1}^{n} a_{1j} x_j \le b_1$$

$$a_2^T x = \sum_{j=1}^{n} a_{2j} x_j \le b_2$$

$$\dots$$

$$a_m^T x = \sum_{j=1}^{n} a_{mj} x_j \le b_m$$

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

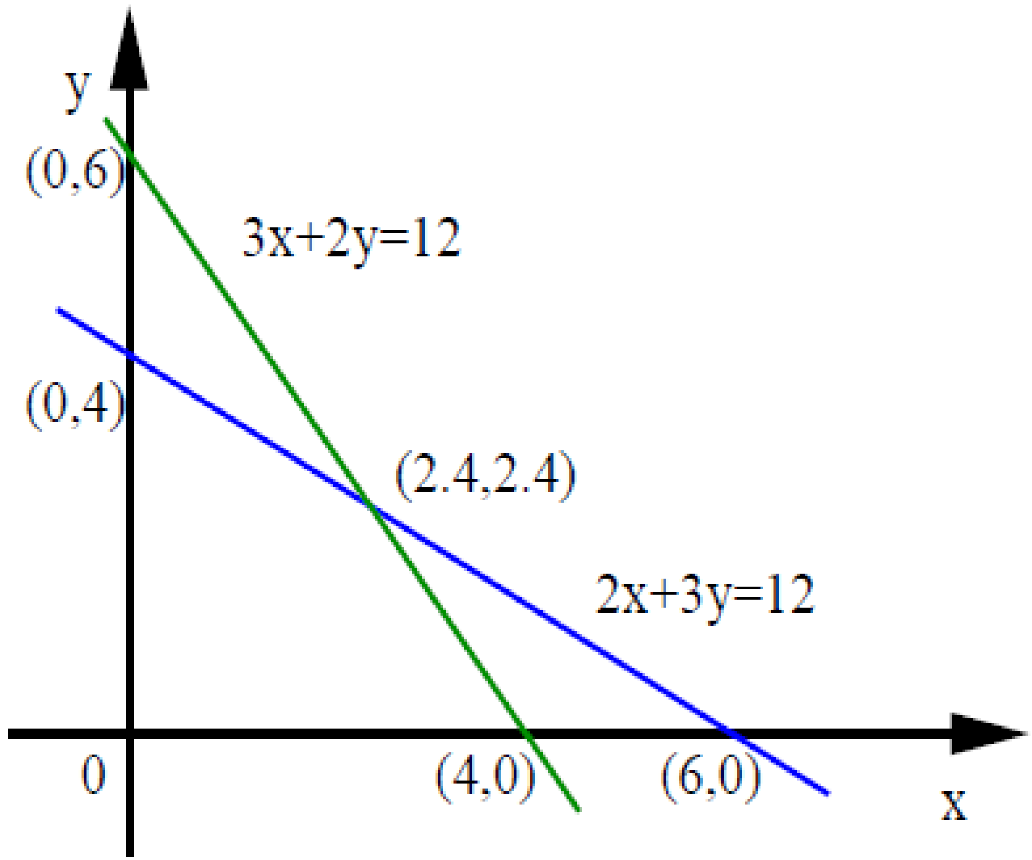


Figure 2: System of Linear Equations

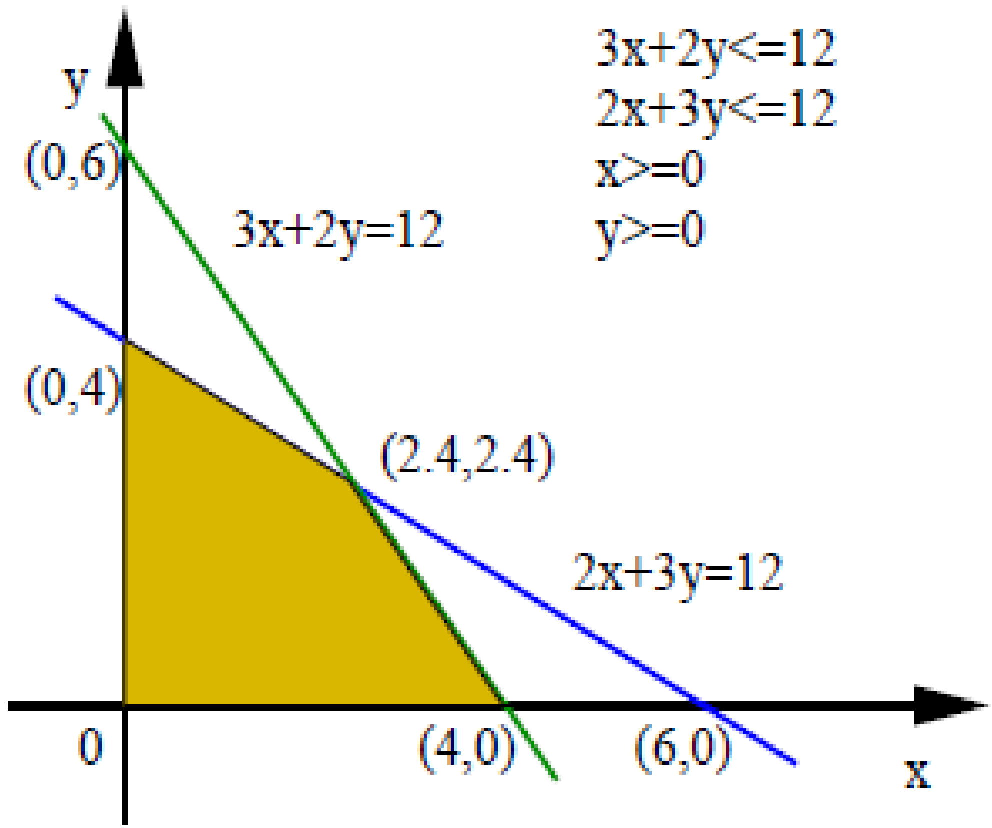
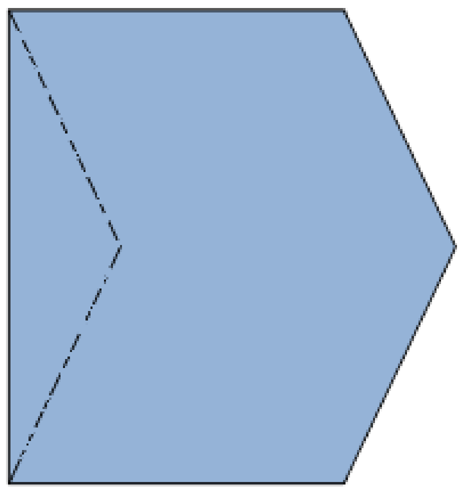


Figure 3: System of Linear Inequalities

Corner or Extreme Points

Convex Hull:



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

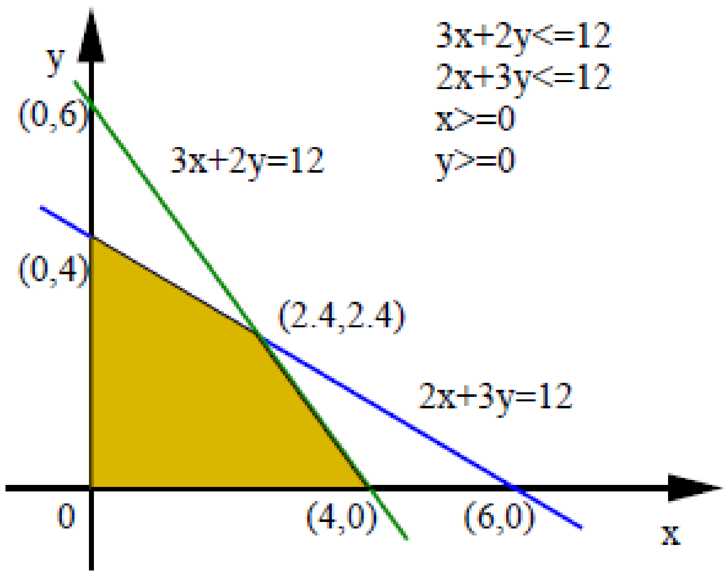


Figure 3: System of Linear Inequalities

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 $\boldsymbol{d}$

It is always associated with a given point $\boldsymbol{x}$

Together a point and a direction vector define a ray:

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

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

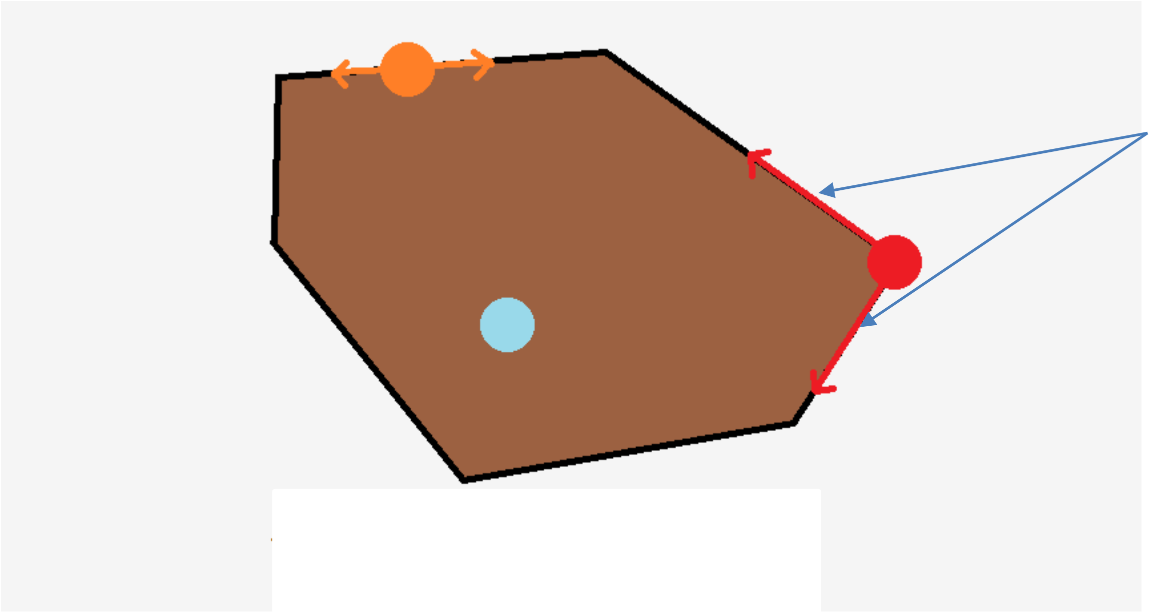
Feasible Direction: A direction, $\boldsymbol{d}$, is said to be "feasible" (relative to a given feasible point $\boldsymbol{x}$) if $\boldsymbol{x} + \epsilon \boldsymbol{d}$ is feasible for some $\epsilon > 0$ and small enough.

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:



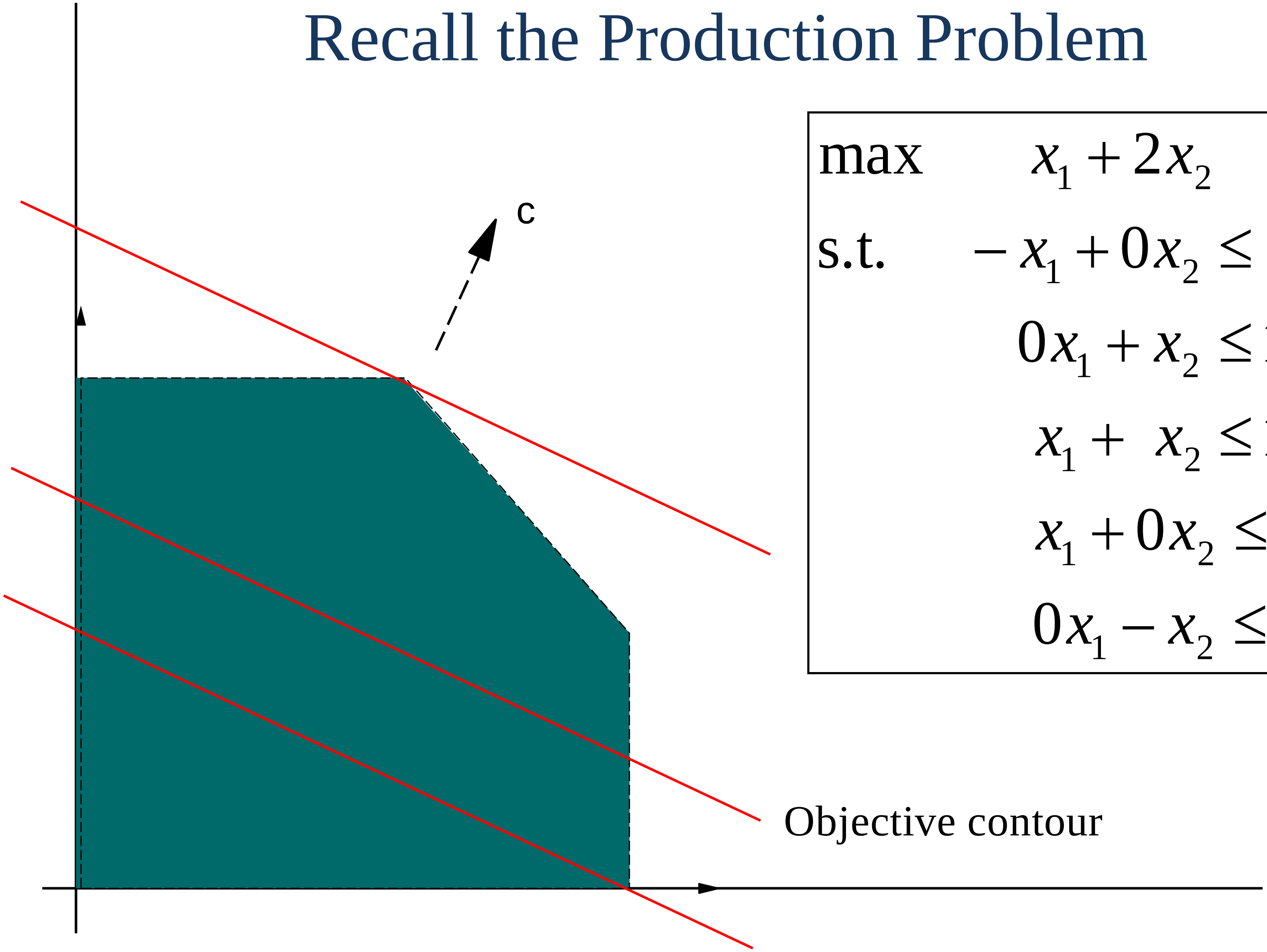
Two **extreme** directions

Interior Point is a point ***x*** where every direction is feasible

LP Problem in Inequality Form

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

Recall the Production Problem



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

Basic Theorems 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

Sketch Proof of Local Optimality Implies Global*

$$\begin{array}{lll} \text{(P)} & \text{minimize} & f(\mathbf{x}) \\ & \text{subject to} & \mathbf{x} \in \Omega \subset R^n, \end{array}$$

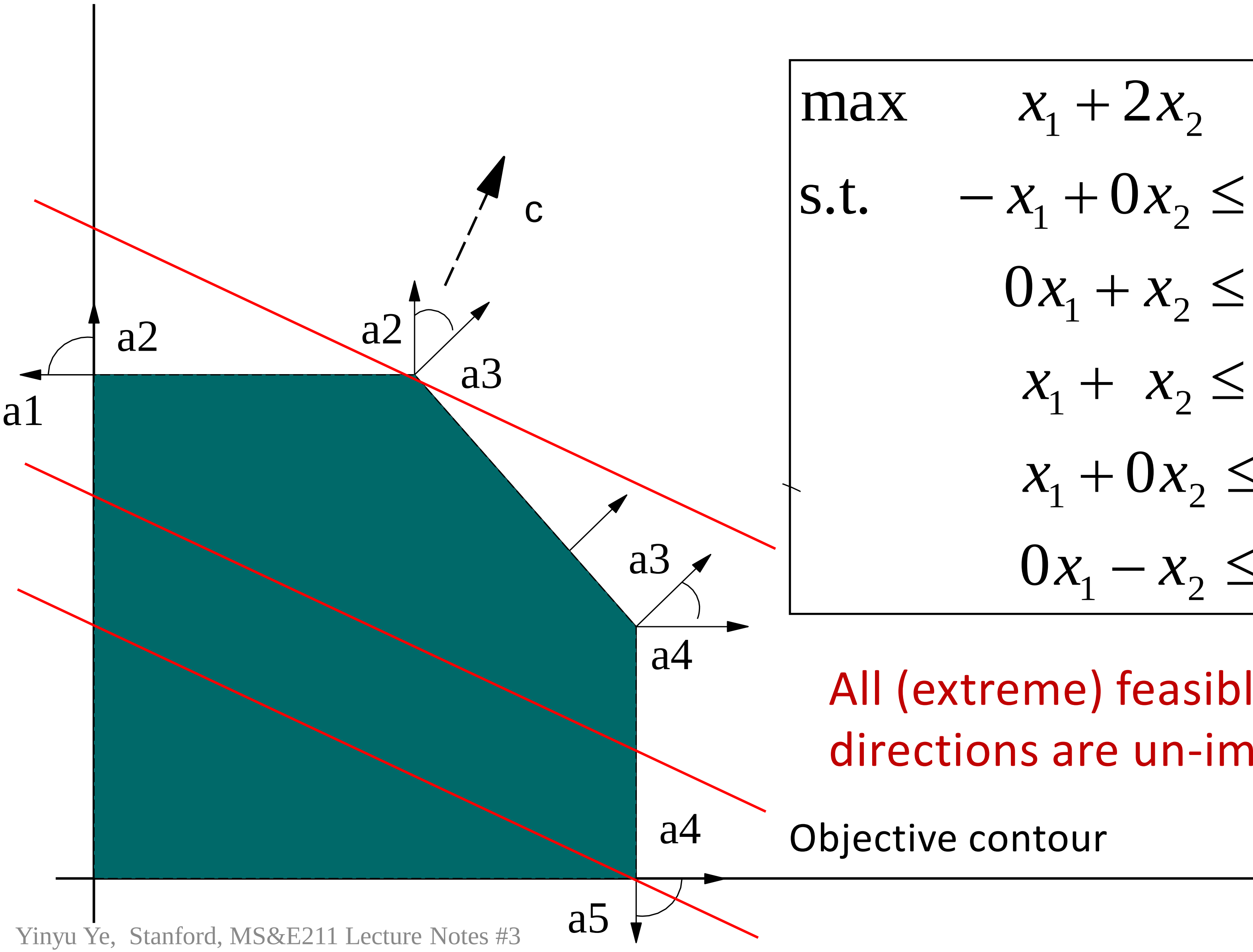
Again Convex Optimization problem if $f(\mathbf{x})$ is a convex function subject to a convex feasible region Ω.

Proof by contradiction. Suppose $\mathbf{x}'$ is a local minimizer but not a global minimizer $\mathbf{x}^*$, that is, $\mathbf{x}' \in \Omega$ and $\mathbf{x}^* \in \Omega$ but $f(\mathbf{x}^*) < f(\mathbf{x}')$. Now the convex combination point $\alpha\mathbf{x}' + (1-\alpha)\mathbf{x}^*$ must be feasible (why?), and

$$f(\alpha\mathbf{x}' + (1-\alpha)\mathbf{x}^*) \le \alpha f(\mathbf{x}') + (1-\alpha) f(\mathbf{x}^*) < f(\mathbf{x}')$$

for any $0 \le \alpha < 1$. This contradicts the local optimality as α can be arbitrarily close to 1 so that $\alpha\mathbf{x}' + (1-\alpha)\mathbf{x}^*$ can be arbitrarily close to $\mathbf{x}'$.

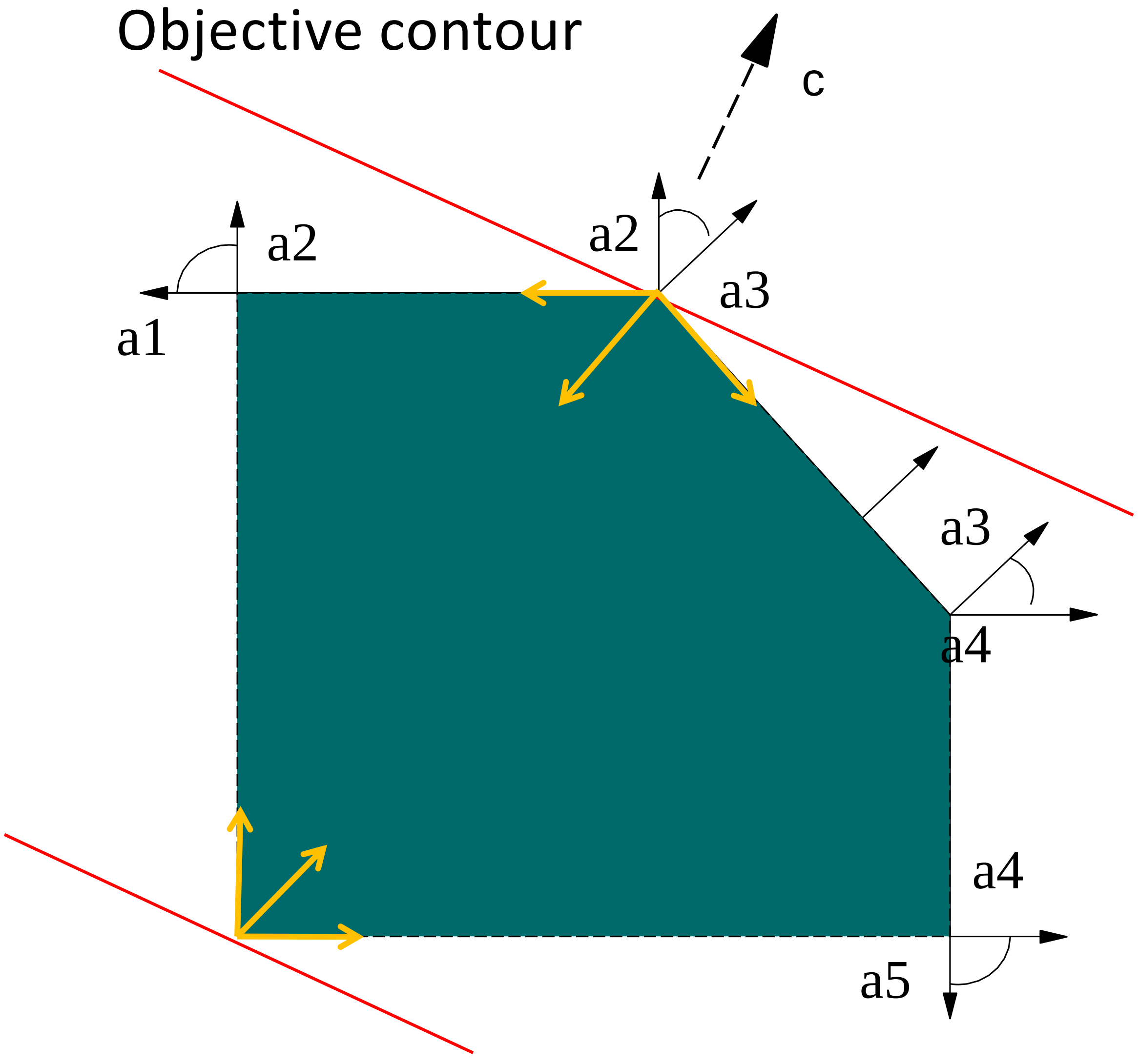
Optimality Certification of the Production Problem



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

All (extreme) feasible directions are un-improving

Feasible Directions at the Optimal Corner



At the optimal corner, ***c*** must a **conic combination** of $\boldsymbol{a_2}$ and $\boldsymbol{a_3}$, 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

$$\boldsymbol{c} = \alpha_2 \boldsymbol{a_2} + \alpha_3 \boldsymbol{a_3},$$

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, $\boldsymbol{c}^T\boldsymbol{d} \le 0$ in this case (where $\boldsymbol{c}$ is the steepest ascent direction of the objective function).
- Recall at the optimal corner, objective direction $\boldsymbol{c}$ is a conic combination of the normal directions, $\boldsymbol{a}_{j1}$ and $\boldsymbol{a}_{j2}$, at the corner point, that is, there are **multipliers** $\alpha_{j1} \ge 0$ and $\alpha_{j2} \ge 0$, such as

$$\boldsymbol{c} = \alpha_{j1}\boldsymbol{a}_{j1} + \alpha_{j2}\boldsymbol{a}_{j2},$$ 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

$$\boldsymbol{c} = \alpha_1\boldsymbol{a}_{i1} + \ldots + \alpha_n\boldsymbol{a}_{in}$$

with nonnegative multipliers α's where vectors $\boldsymbol{a}_{i1}, \ldots, \boldsymbol{a}_{in}$ are the normal directions of hyperplanes associated with the corner point.

- This is the essential idea led to the **Simplex** method by Dantzig: if $\boldsymbol{c}$ 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 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.