Lecture 3
Geometry of linear programming

- subspaces and affine sets, independent vectors
- matrices, range and nullspace, rank, inverse
- polyhedron in inequality form
- extreme points
- the optimal set of a linear program

Subspaces

$\mathcal{S} \subseteq \mathbf{R}^n$ ($\mathcal{S} \neq \emptyset$) is called a *subspace* if

$$x, y \in \mathcal{S}, \quad \alpha, \beta \in \mathbf{R} \quad \Longrightarrow \quad \alpha x + \beta y \in \mathcal{S}$$

$\alpha x + \beta y$ is called a *linear combination* of x and y

examples (in $\mathbf{R}^n$)

- $\mathcal{S} = \mathbf{R}^n$, $\mathcal{S} = \{0\}$
- $\mathcal{S} = \{\alpha v \mid \alpha \in \mathbf{R}\}$ where $v \in \mathbf{R}^n$ ($i.e.$, a line through the origin)
- $\mathcal{S} = \mathrm{span}(v_1, v_2, \ldots, v_k) = \{\alpha_1 v_1 + \cdots + \alpha_k v_k \mid \alpha_i \in \mathbf{R}\}$, where $v_i \in \mathbf{R}^n$
- set of vectors orthogonal to given vectors $v_1, \ldots, v_k$:

$$\mathcal{S} = \{x \in \mathbf{R}^n \mid v_1^T x = 0, \ldots, v_k^T x = 0\}$$

Independent vectors

vectors $v_1, v_2, \ldots, v_k$ are *independent* if and only if

$$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_k v_k = 0 \quad \Longrightarrow \quad \alpha_1 = \alpha_2 = \cdots = 0$$

some equivalent conditions:

- coefficients of $\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_k v_k$ are uniquely determined, $i.e.$,

  $$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_k v_k = \beta_1 v_1 + \beta_2 v_2 + \cdots + \beta_k v_k$$

  implies $\alpha_1 = \beta_1, \alpha_2 = \beta_2, \ldots, \alpha_k = \beta_k$

- no vector v_i can be expressed as a linear combination of the other vectors $v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_k$

Basis and dimension

$\{v_1, v_2, \ldots, v_k\}$ is a *basis* for a subspace $\mathcal{S}$ if

- $v_1, v_2, \ldots, v_k$ span $\mathcal{S}$, $i.e.$, $\mathcal{S} = \mathrm{span}(v_1, v_2, \ldots, v_k)$
- $v_1, v_2, \ldots, v_k$ are independent

equivalently: every $v \in \mathcal{S}$ can be uniquely expressed as

$$v = \alpha_1 v_1 + \cdots + \alpha_k v_k$$

fact: for a given subspace $\mathcal{S}$, the number of vectors in any basis is the same, and is called the *dimension* of $\mathcal{S}$, denoted $\dim \mathcal{S}$

Affine sets

$\mathcal{V} \subseteq \mathbf{R}^n$ ($\mathcal{V} \neq \emptyset$) is called an *affine set* if

$$x, y \in \mathcal{V},\ \alpha + \beta = 1 \quad \Longrightarrow \quad \alpha x + \beta y \in \mathcal{V}$$

$\alpha x + \beta y$ is called an *affine combination* of x and y

examples (in $\mathbf{R}^n$)

- subspaces
- $\mathcal{V} = b + \mathcal{S} = \{x + b \mid x \in \mathcal{S}\}$ where $\mathcal{S}$ is a subspace
- $\mathcal{V} = \{\alpha_1 v_1 + \cdots + \alpha_k v_k \mid \alpha_i \in \mathbf{R}, \sum_i \alpha_i = 1\}$
- $\mathcal{V} = \{x \mid v_1^T x = b_1, \ldots, v_k^T x = b_k\}$ (if $\mathcal{V} \neq \emptyset$)

every affine set $\mathcal{V}$ can be written as $\mathcal{V} = x_0 + \mathcal{S}$ where $x_0 \in \mathbf{R}^n$, $\mathcal{S}$ a subspace ($e.g.$, can take any $x_0 \in \mathcal{V}$, $\mathcal{S} = \mathcal{V} - x_0$)

$\dim(\mathcal{V} - x_0)$ is called the dimension of $\mathcal{V}$

Matrices

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \in \mathbf{R}^{m \times n}$$

some special matrices:

- $A = 0$ (zero matrix): $a_{ij} = 0$
- $A = I$ (identity matrix): $m = n$ and $A_{ii} = 1$ for $i = 1, \ldots, n$, $A_{ij} = 0$ for $i \neq j$
- $A = \mathbf{diag}(x)$ where $x \in \mathbf{R}^n$ (diagonal matrix): $m = n$ and

$$A = \begin{bmatrix} x_1 & 0 & \cdots & 0 \\ 0 & x_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & x_n \end{bmatrix}$$

Matrix operations

- addition, subtraction, scalar multiplication
- transpose:

$$A^T = \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \vdots & \vdots & & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{bmatrix} \in \mathbf{R}^{n \times m}$$

- multiplication: $A \in \mathbf{R}^{m \times n}$, $B \in \mathbf{R}^{n \times q}$, $AB \in \mathbf{R}^{m \times q}$:

$$AB = \begin{bmatrix} \sum_{i=1}^n a_{1i}b_{i1} & \sum_{i=1}^n a_{1i}b_{i2} & \cdots & \sum_{i=1}^n a_{1i}b_{iq} \\ \sum_{i=1}^n a_{2i}b_{i1} & \sum_{i=1}^n a_{2i}b_{i2} & \cdots & \sum_{i=1}^n a_{2i}b_{iq} \\ \vdots & \vdots & & \vdots \\ \sum_{i=1}^n a_{mi}b_{i1} & \sum_{i=1}^n a_{mi}b_{i2} & \cdots & \sum_{i=1}^n a_{mi}b_{iq} \end{bmatrix}$$

Rows and columns

rows of $A \in \mathbf{R}^{m \times n}$:

$$A = \begin{bmatrix} a_1^T \\ a_2^T \\ \vdots \\ a_m^T \end{bmatrix}$$

with $a_i = (a_{i1}, a_{i2}, \ldots, a_{in}) \in \mathbf{R}^n$

columns of $B \in \mathbf{R}^{n \times q}$:

$$B = \begin{bmatrix} b_1 & b_2 & \cdots & b_q \end{bmatrix}$$

with $b_i = (b_{1i}, b_{2i}, \ldots, b_{ni}) \in \mathbf{R}^n$

for example, can write AB as

$$AB = \begin{bmatrix} a_1^T b_1 & a_1^T b_2 & \cdots & a_1^T b_q \\ a_2^T b_1 & a_2^T b_2 & \cdots & a_2^T b_q \\ \vdots & \vdots & & \vdots \\ a_m^T b_1 & a_m^T b_2 & \cdots & a_m^T b_q \end{bmatrix}$$

Range of a matrix

the *range* of $A \in \mathbf{R}^{m \times n}$ is defined as

$$\mathcal{R}(A) = \{Ax \mid x \in \mathbf{R}^n\} \subseteq \mathbf{R}^m$$

- a subspace
- set of vectors that can be 'hit' by mapping $y = Ax$
- the span of the columns of $A = [a_1 \ \cdots \ a_n]$

$$\mathcal{R}(A) = \{a_1 x_1 + \cdots + a_n x_n \mid x \in \mathbf{R}^n\}$$

- the set of vectors y s.t. $Ax = y$ has a solution

$\mathcal{R}(A) = \mathbf{R}^m \iff$

- $Ax = y$ can be solved in x for any y
- the columns of A span $\mathbf{R}^m$
- $\dim \mathcal{R}(A) = m$

Interpretations

$v \in \mathcal{R}(A)$, $w \notin \mathcal{R}(A)$

- $y = Ax$ represents output resulting from input x
  - v is a possible result or output
  - w cannot be a result or output

  $\mathcal{R}(A)$ characterizes the *achievable outputs*

- $y = Ax$ represents measurement of x
  - $y = v$ is a *possible* or *consistent* sensor signal
  - $y = w$ is *impossible* or *inconsistent*; sensors have failed or model is wrong

  $\mathcal{R}(A)$ characterizes the *possible results*

Nullspace of a matrix

the *nullspace* of $A \in \mathbf{R}^{m \times n}$ is defined as

$$\mathcal{N}(A) = \{ \; x \in \mathbf{R}^n \mid Ax = 0 \; \}$$

- a subspace
- the set of vectors mapped to zero by $y = Ax$
- the set of vectors orthogonal to all rows of A:

  $$\mathcal{N}(A) = \left\{ \; x \in \mathbf{R}^n \mid a_1^T x = \cdots = a_m^T x = 0 \right\}$$

  where $A = [a_1 \;\; \cdots \;\; a_m]^T$

zero nullspace: $\mathcal{N}(A) = \{0\} \iff$

- x can always be uniquely determined from $y = Ax$
  ($i.e.$, the linear transformation $y = Ax$ doesn't 'lose' information)
- columns of A are independent

Interpretations

suppose $z \in \mathcal{N}(A)$

- $y = Ax$ represents output resulting from input x
  - z is input with no result
  - x and $x + z$ have same result

  $\mathcal{N}(A)$ characterizes *freedom of input choice* for given result

- $y = Ax$ represents measurement of x
  - z is undetectable — get zero sensor readings
  - x and $x + z$ are indistinguishable: $Ax = A(x + z)$

  $\mathcal{N}(A)$ characterizes *ambiguity* in x from $y = Ax$

Inverse

$A \in \mathbf{R}^{n \times n}$ is *invertible* or *nonsingular* if $\det A \neq 0$

equivalent conditions:

- columns of A are a basis for $\mathbf{R}^n$
- rows of A are a basis for $\mathbf{R}^n$
- $\mathcal{N}(A) = \{0\}$
- $\mathcal{R}(A) = \mathbf{R}^n$
- $y = Ax$ has a unique solution x for every $y \in \mathbf{R}^n$
- A has an inverse $A^{-1} \in \mathbf{R}^{n \times n}$, with $AA^{-1} = A^{-1}A = I$

Rank of a matrix

we define the *rank* of $A \in \mathbf{R}^{m \times n}$ as

$$\mathbf{rank}(A) = \dim \mathcal{R}(A)$$

(nontrivial) facts:

- $\mathbf{rank}(A) = \mathbf{rank}(A^T)$
- $\mathbf{rank}(A)$ is maximum number of independent columns (or rows) of A, hence
$$\mathbf{rank}(A) \leq \min\{m, n\}$$
- $\mathbf{rank}(A) + \dim \mathcal{N}(A) = n$

Full rank matrices

for $A \in \mathbf{R}^{m \times n}$ we have $\mathbf{rank}(A) \leq \min\{m, n\}$

we say A is *full rank* if $\mathbf{rank}(A) = \min\{m, n\}$

- for *square* matrices, full rank means nonsingular
- for *skinny* matrices ($m > n$), full rank means columns are independent
- for *fat* matrices ($m < n$), full rank means rows are independent

Sets of linear equations

$$Ax = y$$

given $A \in \mathbf{R}^{m \times n}$, $y \in \mathbf{R}^m$

- solvable if and only if $y \in \mathcal{R}(A)$
- unique solution if $y \in \mathcal{R}(A)$ and $\mathbf{rank}(A) = n$
- general solution set:
$$\{x_0 + v \mid v \in \mathcal{N}(A)\}$$
  where $Ax_0 = y$

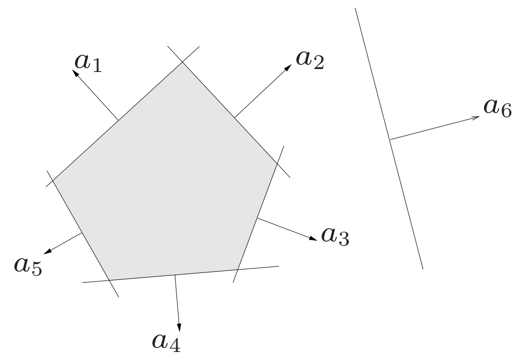
A square and invertible: unique solution for every y:

$$x = A^{-1}y$$

Polyhedron (inequality form)

$A = [a_1 \ \cdots \ a_m]^T \in \mathbf{R}^{m \times n}$, $b \in \mathbf{R}^m$

$$\mathcal{P} = \{x \mid Ax \leq b\} = \{x \mid a_i^T x \leq b_i, \ \ i = 1, \ldots, m\}$$



$\mathcal{P}$ is convex:

$$x, y \in \mathcal{P}, \ \ 0 \leq \lambda \leq 1 \quad \Longrightarrow \quad \lambda x + (1-\lambda) y \in \mathcal{P}$$

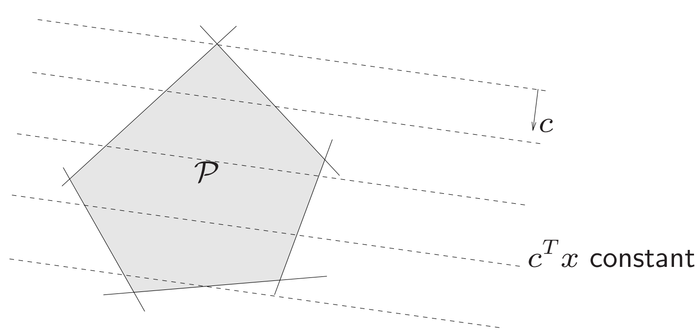
i.e., the *line segment* between any two points in $\mathcal{P}$ lies in $\mathcal{P}$

Extreme points and vertices

$x \in \mathcal{P}$ is an **extreme point** if it cannot be written as

$$x = \lambda y + (1-\lambda) z$$

with $0 \leq \lambda \leq 1$, $y, z \in \mathcal{P}$, $y \neq x$, $z \neq x$



$x \in \mathcal{P}$ is a **vertex** if there is a c such that $c^T x < c^T y$ for all $y \in \mathcal{P}$, $y \neq x$

fact: x is an extreme point $\Longleftrightarrow$ x is a vertex (proof later)

Basic feasible solution

define I as the set of indices of the *active* or *binding* constraints (at $x^\star$):

$$a_i^T x^\star = b_i, \quad i \in I, \qquad a_i^T x^\star < b_i, \quad i \notin I$$

define $\bar{A}$ as

$$\bar{A} = \begin{bmatrix} a_{i_1}^T \\ a_{i_2}^T \\ \vdots \\ a_{i_k}^T \end{bmatrix}, \qquad I = \{i_1, \ldots, i_k\}$$

$x^\star$ is called a *basic feasible solution* if

$$\mathbf{rank}\,\overline{A} = n$$

fact: $x^\star$ is a vertex (extreme point) $\Longleftrightarrow$ $x^\star$ is a basic feasible solution (proof later)

Example

$$\begin{bmatrix} -1 & 0 \\ 2 & 1 \\ 0 & -1 \\ 1 & 2 \end{bmatrix} x \le \begin{bmatrix} 0 \\ 3 \\ 0 \\ 3 \end{bmatrix}$$

- (1,1) is an extreme point
- (1,1) is a vertex: unique minimum of $c^T x$ with $c = (-1, -1)$
- (1,1) is a basic feasible solution: $I = \{2, 4\}$ and $\mathbf{rank}\,\overline{A} = 2$, where

$$\overline{A} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$

Equivalence of the three definitions

vertex $\Longrightarrow$ extreme point

let $x^\star$ be a vertex of $\mathcal{P}$, *i.e.*, there is a $c \neq 0$ such that

$$c^T x^\star < c^T x \quad \text{for all } x \in \mathcal{P},\ x \neq x^\star$$

let $y, z \in \mathcal{P}$, $y \neq x^\star$, $z \neq x^\star$:

$$c^T x^\star < c^T y, \qquad c^T x^\star < c^T z$$

so, if $0 \leq \lambda \leq 1$, then

$$c^T x^\star < c^T(\lambda y + (1-\lambda) z)$$

hence $x^\star \neq \lambda y + (1-\lambda) z$

extreme point $\Longrightarrow$ basic feasible solution

suppose $x^\star \in \mathcal{P}$ is an extreme point with

$$a_i^T x^\star = b_i, \quad i \in I, \qquad a_i^T x^\star < b_i, \quad i \notin I$$

suppose $x^\star$ is not a basic feasible solution; then there exists a $d \neq 0$ with

$$a_i^T d = 0, \quad i \in I$$

and for small enough $\epsilon > 0$,

$$y = x^\star + \epsilon d \in \mathcal{P}, \quad z = x^\star - \epsilon d \in \mathcal{P}$$

we have

$$x^\star = 0.5y + 0.5z,$$

which contradicts the assumption that $x^\star$ is an extreme point

basic feasible solution $\Longrightarrow$ vertex

suppose $x^\star \in \mathcal{P}$ is a basic feasible solution and

$$a_i^T x^\star = b_i \quad i \in I, \qquad a_i^T x^\star < b_i \quad i \notin I$$

define $c = -\sum_{i \in I} a_i$; then

$$c^T x^\star = -\sum_{i \in I} b_i$$

and for all $x \in \mathcal{P}$,

$$c^T x \geq -\sum_{i \in I} b_i$$

with equality only if $a_i^T x = b_i$, $i \in I$

however the only solution to $a_i^T x = b_i$, $i \in I$, is $x^\star$; hence $c^T x^\star < c^T x$ for all $x \in \mathcal{P}$

Unbounded directions

$\mathcal{P}$ contains a **half-line** if there exists $d \neq 0$, x_0 such that

$$x_0 + td \in \mathcal{P} \text{ for all } t \geq 0$$

equivalent condition for $\mathcal{P} = \{x \mid Ax \leq b\}$:

$$Ax_0 \leq b, \quad Ad \leq 0$$

fact: $\mathcal{P}$ unbounded $\Longleftrightarrow$ $\mathcal{P}$ contains a half-line

$\mathcal{P}$ contains a **line** if there exists $d \neq 0$, x_0 such that

$$x_0 + td \in \mathcal{P} \text{ for all } t$$

equivalent condition for $\mathcal{P} = \{x \mid Ax \leq b\}$:

$$Ax_0 \leq b, \quad Ad = 0$$

fact: $\mathcal{P}$ has no extreme points $\Longleftrightarrow$ $\mathcal{P}$ contains a line

Optimal set of an LP

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax \leq b \end{array}$$

- optimal value: $p^\star = \min\{c^T x \mid Ax \leq b\}$ ($p^\star = \pm\infty$ is possible)
- optimal point: $x^\star$ with $Ax^\star \leq b$ and $c^T x^\star = p^\star$
- optimal set: $X_{\text{opt}} = \{x \mid Ax \leq b,\ c^T x = p^\star\}$

example

$$\begin{array}{ll} \text{minimize} & c_1 x_1 + c_2 x_2 \\ \text{subject to} & -2x_1 + x_2 \leq 1 \\ & x_1 \geq 0, \quad x_2 \geq 0 \end{array}$$

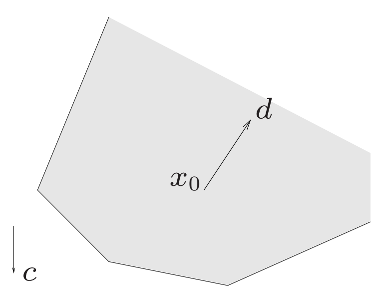
- $c = (1,1)$: $X_{\text{opt}} = \{(0,0)\}$, $p^\star = 0$
- $c = (1,0)$: $X_{\text{opt}} = \{(0,x_2) \mid 0 \leq x_2 \leq 1\}$, $p^\star = 0$
- $c = (-1,-1)$: $X_{\text{opt}} = \emptyset$, $p^\star = -\infty$

Existence of optimal points

- $p^\star = -\infty$ if and only if there exists a feasible half-line

$$\{x_0 + td \mid t \geq 0\}$$

  with $c^T d < 0$



- $p^\star = +\infty$ if and only if $\mathcal{P} = \emptyset$
- $p^\star$ is finite if and only if $X_{\text{opt}} \neq \emptyset$

property: if $\mathcal{P}$ has at least one extreme point and $p^\star$ is finite, then there exists an extreme point that is optimal