L. Vandenberghe

Lecture 3
Polyhedra

- linear algebra review
- minimal faces and extreme points

Subspace

definition: a nonempty subset S of $\mathbf{R}^n$ is a subspace if

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

- extends recursively to linear combinations of more than two vectors:

$$x_1, \ldots, x_k \in S, \quad \alpha_1, \ldots, \alpha_k \in \mathbf{R} \qquad \Longrightarrow \qquad \alpha_1 x_1 + \cdots + \alpha_k x_k \in S$$

- all subspaces contain the origin

subspaces and matrices (with $A \in \mathbf{R}^{m \times n}$)

- **range:** $\mathrm{range}(A) = \{x \in \mathbf{R}^m \mid x = Ay \text{ for some } y\}$ is a subspace of $\mathbf{R}^m$
- **nullspace:** $\mathrm{nullspace}(A) = \{x \in \mathbf{R}^n \mid Ax = 0\}$ is a subspace of $\mathbf{R}^n$

conversely, every subspace can be expressed as a range or nullspace

Linear independence

a nonempty set of vectors $\{v_1, v_2, \ldots, v_k\}$ is **linearly independent** if

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

holds only for $\alpha_1 = \alpha_2 = \cdots = \alpha_k = 0$

properties: if $\{v_1, \ldots, v_k\}$ is a linearly independent set, then

- coefficients α_k in linear combinations $x = \alpha_1 v_1 + \cdots + \alpha_k v_k$ are unique:

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

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

- none of the vectors v_i is a linear combination of the other vectors

Basis and dimension

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

- every $x \in S$ can be expressed as a linear combination of v_1, . . . , v_k
- $\{v_1, \ldots, v_k\}$ is a linearly independent set

equivalently, every $x \in S$ can be expressed in exactly one way as

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

dimension: $\dim S$ is the number of vectors in a basis of S

- key fact from linear algebra: all bases of a subspace have the same size
- a linearly independent subset of S can't have more than $\dim S$ elements
- if S is a subspace in $\mathbf{R}^n$, then $0 \leq \dim S \leq n$

Range, nullspace, and linear equations

consider a linear equation $Ax = b$ with $A \in \mathbf{R}^{m \times n}$ (not necessarily square)

range of A: determines existence of solutions

- equation is solvable for $b \in \mathrm{range}(A)$
- if $\mathrm{range}(A) = \mathbf{R}^m$, there is at least one solution for every b

nullspace of A: determines uniqueness of solutions

- if $\hat{x}$ is a solution, then the complete solution set is $\{\hat{x} + v \mid Av = 0\}$
- if $\mathrm{nullspace}(A) = \{0\}$, there is at most one solution for every b

Matrix rank

the **rank** of a matrix A is defined as

$$\operatorname{rank}(A) = \dim \operatorname{range}(A)$$

properties (assume A is $m \times n$)

- $\operatorname{rank}(A) = \operatorname{rank}(A^T)$
- $\operatorname{rank}(A) \leq \min\{m, n\}$

  if $\operatorname{rank}(A) = \min\{m, n\}$ the matrix is said to be full rank
- $\dim \operatorname{nullspace}(A) = n - \operatorname{rank}(A)$

Left-invertible matrix

definition: A is left-invertible if there exists an X with

$$XA = I$$

X is called a **left inverse** of A

equivalent properties (for an $m \times n$ matrix A)

- $\text{rank}(A) = n$
- $\text{nullspace}(A) = \{0\}$
- the columns of A form a linearly independent set
- the linear equation $Ax = b$ has at most one solution for every r.h.s. b

dimensions: if $A \in \mathbf{R}^{m \times n}$ is left-invertible, then $m \geq n$

Right-invertible matrix

definition: A is right-invertible if there exists a Y with

$$AY = I$$

Y is called a **right inverse** of A

equivalent properties (for an $m \times n$ matrix A)

- $\mathrm{rank}(A) = m$
- $\mathrm{range}(A) = \mathbf{R}^m$
- the rows of A form a linearly independent set
- the linear equation $Ax = b$ has at least one solution for every r.h.s. b

dimensions: if $A \in \mathbf{R}^{m \times n}$ is right-invertible, then $m \leq n$

Invertible matrix

definition: A is invertible (nonsingular) if it is left- and right-invertible

- A is necessarily square
- the linear equation $Ax = b$ has exactly one solution for every r.h.s. b

inverse: if left and right inverses exist, they must be equal and unique

$$XA = I, \quad AY = I \qquad \Longrightarrow \qquad X = X(AY) = (XA)Y = Y$$

we use the notation A^{-1} for the left/right inverse of an invertible matrix

Affine set

definition: a subset S of $\mathbf{R}^n$ is affine if

$$x, y \in S, \quad \alpha + \beta = 1 \qquad \Longrightarrow \qquad \alpha x + \beta y \in S$$

- the line through any two distinct points x, y in S is in S
- extends recursively to affine combinations of more than two vectors

$$x_1, \ldots, x_k \in S, \quad \alpha_1 + \cdots + \alpha_k = 1 \qquad \Longrightarrow \qquad \alpha_1 x_1 + \cdots + \alpha_k x_k \in S$$

parallel subspace: a nonempty set S is affine if and only if the set

$$L = S - \hat{x},$$

with $\hat{x} \in S$, is a subspace

- the parallel subspace L is independent of the choice of $\hat{x} \in S$
- we define the dimension of S to be $\dim L$

Matrices and affine sets

linear equations: the solution set of a system of linear equations

$$S = \{x \mid Ax = b\}$$

is an affine set; moreover, all affine sets can be represented this way

range parametrization: a set defined as

$$S = \{x \mid x = Ay + c \text{ for some } y\}$$

is affine; all nonempty affine sets can be represented this way

Affine hull

definition

- the affine hull of a set C is the smallest affine set that contains C
- equivalently, the set of all affine combinations of points in C:

$$\{\alpha_1 v_1 + \cdots + \alpha_k v_k \mid k \geq 1,\ v_1, \ldots, v_k \in C,\ \alpha_1 + \cdots + \alpha_k = 1\}$$

notation: $\operatorname{aff} C$

example: the affine hull of $C = \{(x, y, z) \in \mathbf{R}^3 \mid x^2 + y^2 = 1, z = 1\}$ is

$$\operatorname{aff} C = \{(x, y, z) \in \mathbf{R}^3 \mid z = 1\}$$

Affine independence

a set of vectors $\{v_1, v_2, \ldots, v_k\}$ in $\mathbf{R}^n$ is **affinely independent** if

$$\operatorname{rank}(\begin{bmatrix} v_1 & v_2 & \cdots & v_k \\ 1 & 1 & \cdots & 1 \end{bmatrix}) = k$$

- the set $\{v_2 - v_1, v_3 - v_1, \ldots, v_k - v_1\}$ is linearly independent
- the affine hull of $\{v_1, v_2, \ldots, v_k\}$ has dimension $k - 1$
- this implies $k \leq n + 1$

example

$$\{\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}\}$$

Outline

- linear algebra review
- **minimal faces and extreme points**

Polyhedron

a **polyhedron** is the solution set of a finite number of linear inequalities

- definition can include linear equalities ($Cx = d \Leftrightarrow Cx \leq d, -Cx \leq -d$)
- note '*finite*': the solution of the infinite set of linear inequalities

$$a^T x \leq 1 \quad \text{for all } a \text{ with } \|a\| = 1$$

  is the unit ball $\{x \mid \|x\| \leq 1\}$ and not a polyhedron

notation: in the remainder of the lecture we consider a polyhedron

$$P = \{x \mid Ax \leq b,\, Cx = d\}$$

- we assume P is not empty
- A is $m \times n$ with rows a_i^T

Lineality space

the **lineality space** of P is

$$L = \text{nullspace}(\left[\begin{array}{c} A \\ C \end{array}\right])$$

if $x \in P$, then $x + v \in P$ for all $v \in L$:

$$A(x+v) = Ax \leq b, \quad C(x+v) = Cx = d \qquad \forall v \in L$$

pointed polyhedron

- a polyhedron with lineality space $\{0\}$ is called pointed
- a polyhedron is pointed if it does not contain an entire line

Examples

not pointed

- a halfspace $\{x \mid a^T x \leq b\}$ $(n \geq 2)$: lineality space is $\{x \mid a^T x = 0\}$
- a 'slab' $\{x \mid -1 \leq a^T x \leq 1\}$ $(n \geq 2)$: lineality space is $\{x \mid a^T x = 0\}$
- $\{(x, y, z) \mid |x| \leq 1, |y| \leq 1\}$ has lineality space $\{(0, 0, z) \mid z \in \mathbf{R}\}$

examples of pointed polyhedra

- probability simplex $\{x \in \mathbf{R}^n \mid \mathbf{1}^T x = 1, x \geq 0\}$
- $\{(x, y, z) \mid |x| \leq z, |y| \leq z\}$

Face

definition: for $J \subseteq \{1, 2, \ldots, m\}$, define

$$F_J = \{x \in P \mid a_i^T x = b_i \text{ for } i \in J\}$$

if F_J is nonempty, it is called a **face** of P

properties

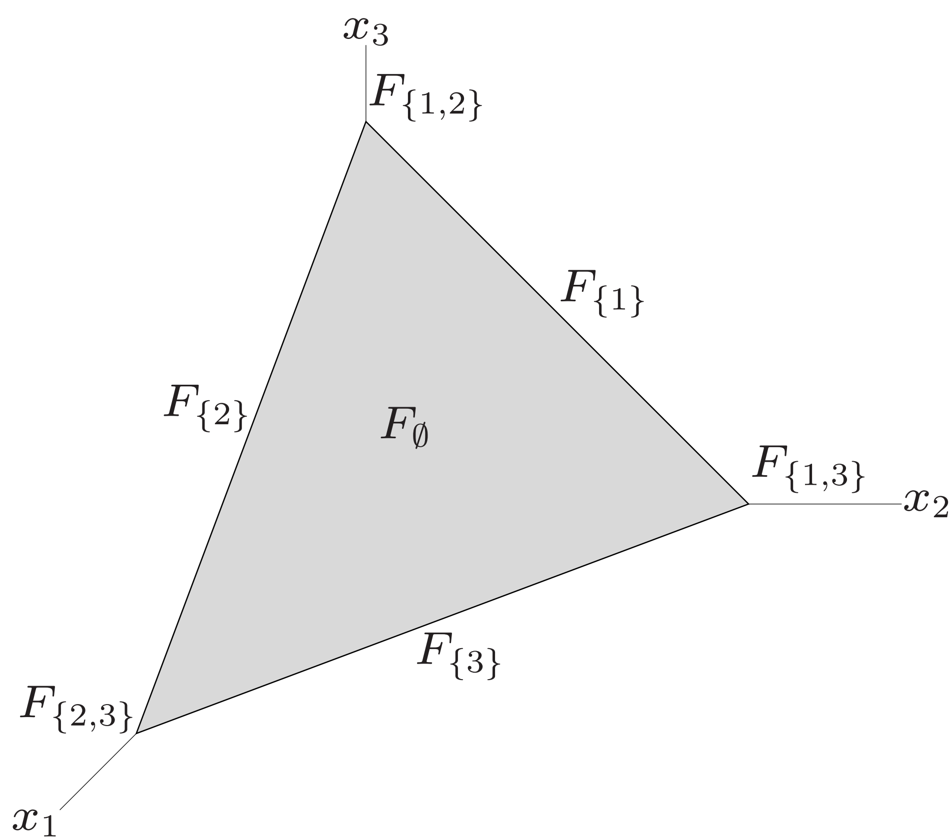
- F_J is a nonempty polyhedron, defined by the inequalities and equalities

$$a_i^T x \le b_i \text{ for } i \notin J, \qquad a_i^T x = b_i \text{ for } i \in J, \qquad Cx = d$$

- faces of F_J are also faces of P
- all faces have the same lineality space as P
- the number of faces is finite and at least one (P itself is a face: $P = F_\emptyset$)

Example

$$\begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \leq \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \qquad x_1 + x_2 + x_3 = 1$$



Example

$$\begin{bmatrix} 1 & -1 & 1 \\ 1 & -1 & -1 \\ -1 & 1 & 1 \\ -1 & 1 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \leq \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

- solution set is a (non-pointed) polyhedron

$$P = \{x \in \mathbf{R}^3 \mid |x_1 - x_2| + |x_3| \leq 1\}$$

- the lineality space is the line $L = \{(t, t, 0) \mid t \in \mathbf{R}\}$

faces of P

- three-dimensional face: $F_\emptyset = P$
- two-dimensional faces:

$$\begin{aligned}
F_{\{1\}} &= \{x \mid x_1 - x_2 + x_3 = 1,\ x_1 \geq x_2,\ x_3 \geq 0\} \\
F_{\{2\}} &= \{x \mid x_1 - x_2 - x_3 = 1,\ x_1 \geq x_2,\ x_3 \leq 0\} \\
F_{\{3\}} &= \{x \mid -x_1 + x_2 + x_3 = 1,\ x_1 \leq x_2,\ x_3 \geq 0\} \\
F_{\{4\}} &= \{x \mid -x_1 + x_2 - x_3 = 1,\ x_1 \leq x_2,\ x_3 \leq 0\}
\end{aligned}$$

- one-dimensional faces:

$$\begin{aligned}
F_{\{1,2\}} &= \{x \mid x_1 - x_2 = 1,\ x_3 = 0\} \\
F_{\{1,3\}} &= \{x \mid x_1 = x_2,\ x_3 = 1\} \\
F_{\{2,4\}} &= \{x \mid x_1 = x_2,\ x_3 = -1\} \\
F_{\{3,4\}} &= \{x \mid x_1 - x_2 = -1,\ x_3 = 0\}
\end{aligned}$$

- F_J is empty for all other J

Minimal face

a face of P is a **minimal face** if it does not contain another face of P

examples

- polyhedron on page 3–18: the faces $F_{\{1,2\}}$, $F_{\{1,3\}}$, $F_{\{2,3\}}$
- polyhedron on page 3–19: the faces $F_{\{1,2\}}$, $F_{\{1,3\}}$, $F_{\{2,4\}}$, $F_{\{3,4\}}$

property

- a face is minimal if and only if it is an affine set (see next page)
- all minimal faces are translates of the lineality space of P

  (since all faces have the same lineality space) *proof:* let F_J be the face defined by

$$a_i^T x \le b_i \text{ for } i \notin J, \qquad a_i^T x = b_i \text{ for } i \in J, \qquad Cx = d$$

partition the inequalities $a_i^T x \le b_i$ ($i \notin J$) in three groups:

1. $i \in J_1$ if $a_i^T x = b_i$ for all x in F_J
2. $i \in J_2$ if $a_i^T x < b_i$ for all x in F_J
3. $i \in J_3$ if there exist points $\hat{x}, \tilde{x} \in F_j$ with $a_i^T \hat{x} < b_i$ and $a_i^T \tilde{x} = b_i$

- inequalities in J_2 are redundant (can be omitted without changing F_J)
- if J_3 is not empty and $j \in J_3$, then $F_{J \cup \{j\}}$ is a proper face of F_J:
  - $F_{J \cup \{j\}}$ is not empty because it contains $\tilde{x}$
  - $F_{J \cup \{j\}}$ is not equal to F_J because it does not contain $\hat{x}$

therefore, if F_J is a minimal face then $J_3 = \emptyset$ and F_J is the solution set of

$$a_i^T x = b_i \text{ for } i \in J_1 \cup J, \qquad Cx = d$$

Extreme points

extreme point (vertex): a minimal face of a pointed polyhedron

rank test: given $\hat{x} \in P$, is $\hat{x}$ an extreme point?

- let $J(\hat{x}) = \{i_1, \ldots, i_k\}$ be the indices of the active constraints at $\hat{x}$:

$$a_i^T \hat{x} = b_i \text{ for } i \in J(\hat{x}), \qquad a_i^T \hat{x} < b_i \text{ for } i \notin J(\hat{x})$$

- $\hat{x}$ is an extreme point if

$$\operatorname{rank}\left(\begin{bmatrix} A_{J(\hat{x})} \\ C \end{bmatrix}\right) = n \quad \text{where } A_{J(\hat{x})} = \begin{bmatrix} a_{i_1}^T \\ \vdots \\ a_{i_k}^T \end{bmatrix}$$

$A_{J(\hat{x})}$ is the submatrix of A with rows indexed by $J(\hat{x})$

proof: the face $F_{J(\hat{x})}$ is defined as the set of points x that satisfy

$$a_i^T x = b_i \text{ for } i \in J(\hat{x}), \qquad a_i^T x \leq b_i \text{ for } i \notin J(\hat{x}), \qquad Cx = d \tag{1}$$

$x = \hat{x}$ satisfies (1) by definition of $J(\hat{x})$

- if the rank condition is satisfied, $x = \hat{x}$ is the only point that satisfies (1)

  therefore $F_{J(\hat{x})}$ is a minimal face ($\dim F_{J(\hat{x})} = 0$)

- if the rank condition does not hold, then there exists a $v \neq 0$ with

$$a_i^T v = 0 \text{ for } i \in J(\hat{x}), \qquad Cv = 0$$

  this implies that $x = \hat{x} \pm tv$ satisfies (1) for small positive and negative t

  therefore the face $F_{J(\hat{x})}$ is not minimal ($\dim F_{J(\hat{x})} > 0$)

Example

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

- $\hat{x} = (1,1)$ is in P:

$$\begin{bmatrix} -1 & 0 \\ 2 & 1 \\ 0 & -1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 3 \\ -1 \\ 3 \end{bmatrix} \leq \begin{bmatrix} 0 \\ 3 \\ 0 \\ 3 \end{bmatrix}$$

- the active constraints at $\hat{x}$ are $J(\hat{x}) = \{2, 4\}$
- the matrix $A_{J(\hat{x})} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$ had rank 2, so $\hat{x}$ is an extreme point

Example

the polyhedron on page 3–18 has three extreme points

- $\hat{x} = (1, 0, 0)$:

$$J(\hat{x}) = \{2, 3\}, \qquad \operatorname{rank}(\begin{bmatrix} A_{J(\hat{x})} \\ C \end{bmatrix}) = \operatorname{rank}(\begin{bmatrix} 0 & -1 & 0 \\ 0 & 0 & -1 \\ 1 & 1 & 1 \end{bmatrix}) = 3$$

- $\hat{x} = (0, 1, 0)$:

$$J(\hat{x}) = \{1, 3\}, \qquad \operatorname{rank}(\begin{bmatrix} A_{J(\hat{x})} \\ C \end{bmatrix}) = \operatorname{rank}(\begin{bmatrix} -1 & 0 & 0 \\ 0 & 0 & -1 \\ 1 & 1 & 1 \end{bmatrix}) = 3$$

- $\hat{x} = (0, 0, 1)$:

$$J(\hat{x}) = \{1, 2\}, \qquad \operatorname{rank}(\begin{bmatrix} A_{J(\hat{x})} \\ C \end{bmatrix}) = \operatorname{rank}(\begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 1 & 1 & 1 \end{bmatrix}) = 3$$

Exercise: polyhedron in standard form

consider a nonempty polyhedron P defined by

$$x \geq 0, \qquad Cx = d$$

note that P is pointed (regardless of values of C, d)

- show that $\hat{x}$ is an extreme point if $\hat{x} \in P$ and

$$\mathrm{rank}(\begin{bmatrix} c_{i_1} & c_{i_2} & \cdots & c_{i_k} \end{bmatrix}) = k$$

  where c_j is column j of C and $\{i_1, i_2, \ldots, i_k\} = \{i \mid \hat{x}_i > 0\}$

- show that an extreme point $\hat{x}$ has at most $\mathrm{rank}(C)$ nonzero elements *solution:* without loss of generality, assume $\{i_1, \ldots, i_k\} = \{1, \ldots, k\}$

- apply rank test to

$$\begin{bmatrix} -I \\ C \end{bmatrix} = \begin{bmatrix} -I_k & 0 \\ 0 & -I_{n-k} \\ D & E \end{bmatrix},$$

  with $D = \begin{bmatrix} c_1 & \cdots & c_k \end{bmatrix}$ and $E = \begin{bmatrix} c_{k+1} & \cdots & c_n \end{bmatrix}$

- inequalities $k+1$, . . . , n are active at $\hat{x}$

- $\hat{x}$ is an extreme point if the submatrix of active constraints has rank n:

$$\text{rank}(\begin{bmatrix} 0 & -I_{n-k} \\ D & E \end{bmatrix}) = n - k + \text{rank}(D) = n$$

  i.e., $\text{rank}(D) = k$

Exercise: Birkhoff's theorem

doubly stochastic matrix: an $n \times n$ matrix X is doubly stochastic if

$$X_{ij} \geq 0, \quad i,j = 1,\ldots,n, \qquad X\mathbf{1} = 1, \qquad X^T\mathbf{1} = 1$$

- a nonnegative matrix with column and row sums equal to one
- set of doubly stochastic matrices form is a pointed polyhedron in $\mathbf{R}^{n \times n}$

question: show that the extreme points are the permutation matrices

(a permutation matrix is a doubly stochastic matrix with elements 0 or 1)