Lecture 1
Introduction and overview

- linear programming
- example
- course topics
- software
- integer linear programming

Linear program (LP)

$$\begin{array}{ll} \text{minimize} & \sum_{j=1}^{n} c_j x_j \\ \text{subject to} & \sum_{j=1}^{n} a_{ij} x_j \leq b_i, \quad i = 1, \ldots, m \\ & \sum_{j=1}^{n} c_{ij} x_j = d_i, \quad i = 1, \ldots, p \end{array}$$

variables: x_j
problem data: the coefficients c_j, a_{ij}, b_i, c_{ij}, d_i

- can be solved very efficiently (several 10,000 variables, constraints)
- widely available general-purpose software
- extensive, useful theory (optimality conditions, sensitivity analysis, . . .)

Example. Open-loop control problem

single-input/single-output system (with input u, output y)

$$y(t) = h_0 u(t) + h_1 u(t-1) + h_2 u(t-2) + h_3 u(t-3) + \cdots$$

output tracking problem: minimize deviation from desired output $y_{\mathrm{des}}(t)$

$$\max_{t=0,\ldots,N} |y(t) - y_{\mathrm{des}}(t)|$$

subject to input amplitude and slew rate constraints:

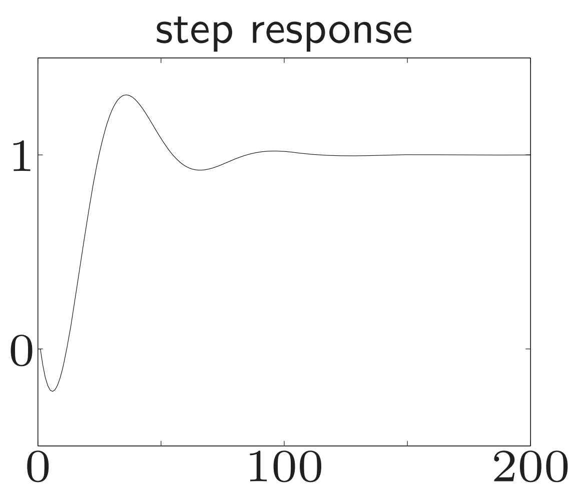
$$|u(t)| \leq U, \qquad |u(t+1) - u(t)| \leq S$$

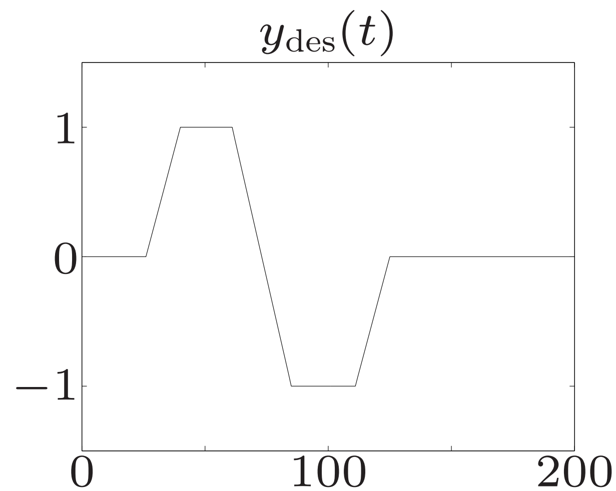
variables: $u(0), \ldots, u(M)$ (with $u(t) = 0$ for $t < 0$, $t > M$)

solution: can be formulated as an LP, hence easily solved (more later)

example

step response ($s(t) = h_t + \cdots + h_0$) and desired output:

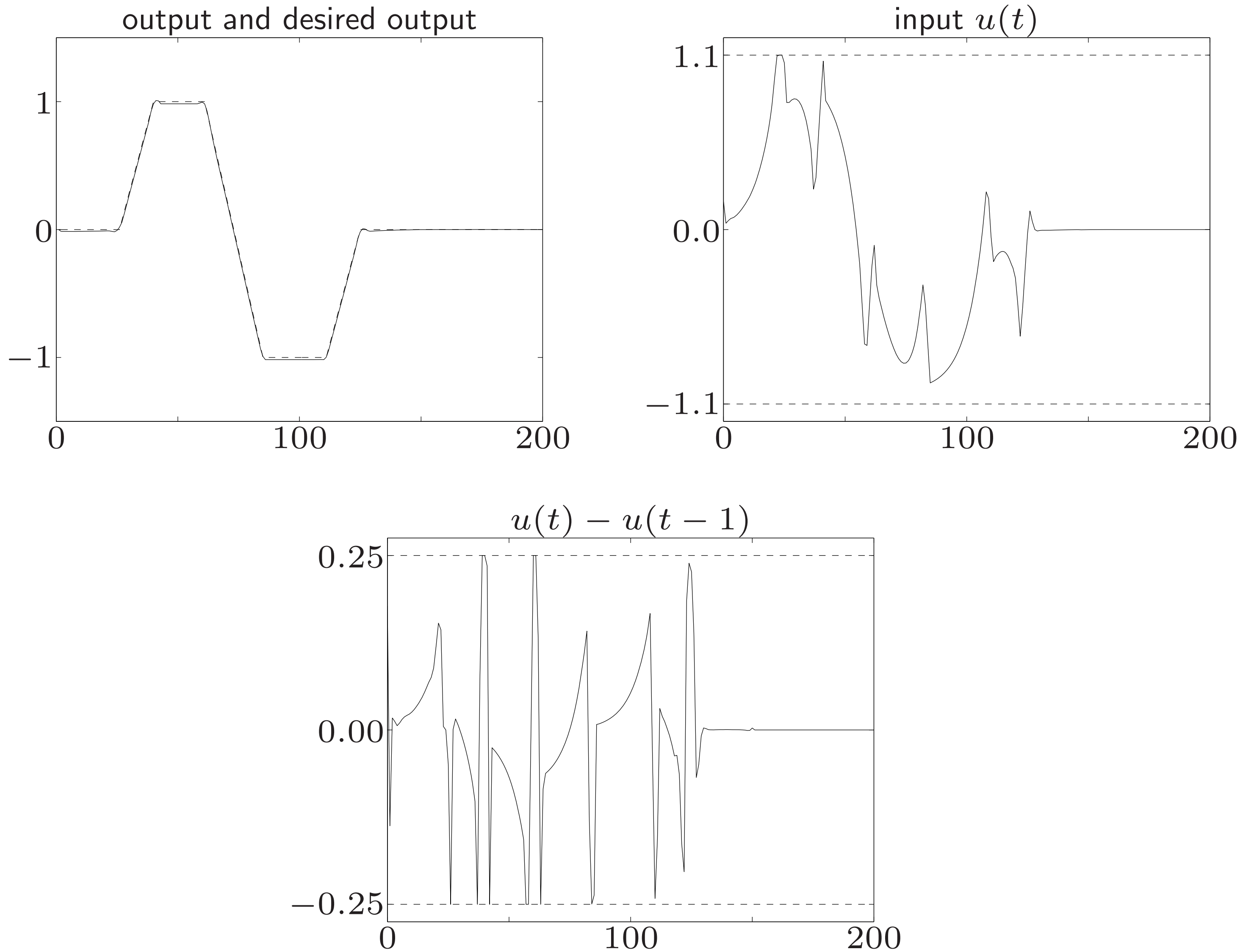




amplitude and slew rate constraint on u:

$$|u(t)| \le 1.1, \qquad |u(t) - u(t-1)| \le 0.25$$

optimal solution



Brief history

- **1930s** (Kantorovich): economic applications
- **1940s** (Dantzig): military logistics problems during WW2;
  1947: simplex algorithm
- **1950s–60s** discovery of applications in many other fields (structural optimization, control theory, filter design, . . .)
- **1979** (Khachiyan) ellipsoid algorithm: more efficient (polynomial-time) than simplex in worst case, but slower in practice
- **1984** (Karmarkar): projective (interior-point) algorithm:
  polynomial-time worst-case complexity, and efficient in practice
- **1984–today**. many variations of interior-point methods (improved complexity or efficiency in practice), software for large-scale problems

Course outline

the linear programming problem
linear inequalities, geometry of linear programming

engineering applications
signal processing, control, structural optimization . . .

duality

algorithms
the simplex algorithm, interior-point algorithms

large-scale linear programming and network optimization
techniques for LPs with special structure, network flow problems

integer linear programming
introduction, some basic techniques

Software

solvers: solve LPs described in some standard form

modeling tools: accept a problem in a simpler, more intuitive, notation and convert it to the standard form required by solvers

software for this course (see class website)

- platforms: Matlab, Octave, Python
- solvers: `linprog` (Matlab Optimization Toolbox),
- modeling tools: CVX (Matlab), YALMIP (Matlab),
- Thanks to Lieven Vandenberghe at UCLA for his slides

Integer linear program

integer linear program

$$\begin{array}{ll} \text{minimize} & \sum_{j=1}^n c_j x_j \\ \text{subject to} & \sum_{j=1}^n a_{ij} x_j \leq b_i, \quad i = 1, \ldots, m \\ & \sum_{j=1}^n c_{ij} x_j = d_i, \quad i = 1, \ldots, p \\ & x_j \in \mathbf{Z} \end{array}$$

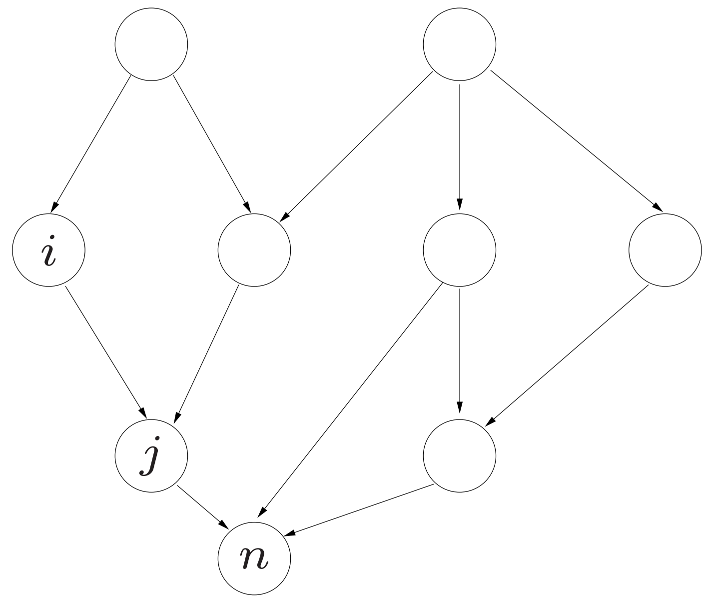
Boolean linear program

$$\begin{array}{ll} \text{minimize} & \sum_{j=1}^n c_j x_j \\ \text{subject to} & \sum_{j=1}^n a_{ij} x_j \leq b_i, \quad i = 1, \ldots, m \\ & \sum_{j=1}^n c_{ij} x_j = d_i, \quad i = 1, \ldots, p \\ & x_j \in \{0, 1\} \end{array}$$

- very general problems; can be extremely hard to solve
- can be solved as a sequence of linear programs

Example. Scheduling problem

scheduling graph $\mathcal{V}$:



- nodes represent operations (*e.g.*, jobs in a manufacturing process, arithmetic operations in an algorithm)
- $(i, j) \in \mathcal{V}$ means operation j must wait for operation i to be finished
- M identical machines/processors; each operation takes unit time

problem: determine fastest schedule

Boolean linear program formulation

variables: x_{is}, $i = 1, \ldots, n$, $s = 0, \ldots, T$:

$x_{is} = 1$ if job i starts at time s , $\qquad$ $x_{is} = 0$ otherwise

constraints:

1. $x_{is} \in \{0, 1\}$

2. job i starts exactly once:

$$\sum_{s=0}^{T} x_{is} = 1$$

3. if there is an arc (i, j) in $\mathcal{V}$, then

$$\sum_{s=0}^{T} s x_{js} - \sum_{s=0}^{T} s x_{is} \geq 1$$

4. limit on capacity (M machines) at time s:

$$\sum_{i=1}^{n} x_{is} \leq M$$

cost function (start time of job n):

$$\sum_{s=0}^{T} s x_{ns}$$

Boolean linear program

$$\begin{array}{ll} \text{minimize} & \sum_{s=0}^{T} s x_{ns} \\ \text{subject to} & \sum_{s=0}^{T} x_{is} = 1, \quad i = 1, \ldots, n \\ & \sum_{s=0}^{T} s x_{js} - \sum_{s=0}^{T} s x_{is} \geq 1, \quad (i,j) \in \mathcal{V} \\ & \sum_{i=1}^{n} x_{is} \leq M, \quad s = 0, \ldots, T \\ & x_{is} \in \{0, 1\}, \quad i = 1, \ldots, n, \quad s = 0, \ldots, T \end{array}$$

Lecture 2
Linear inequalities

- vectors
- inner products and norms
- linear equalities and hyperplanes
- linear inequalities and halfspaces
- polyhedra

Vectors

(column) vector $x \in \mathbf{R}^n$:

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

- $x_i \in \mathbf{R}$: ith *component* or *element* of x
- also written as $x = (x_1, x_2, \ldots, x_n)$

some special vectors:

- $x = 0$ (*zero vector*): $x_i = 0$, $i = 1, \ldots, n$
- $x = \mathbf{1}$: $x_i = 1$, $i = 1, \ldots, n$
- $x = e_i$ (ith *basis vector* or ith *unit vector*): $x_i = 1$, $x_k = 0$ for $k \neq i$

(n follows from context)

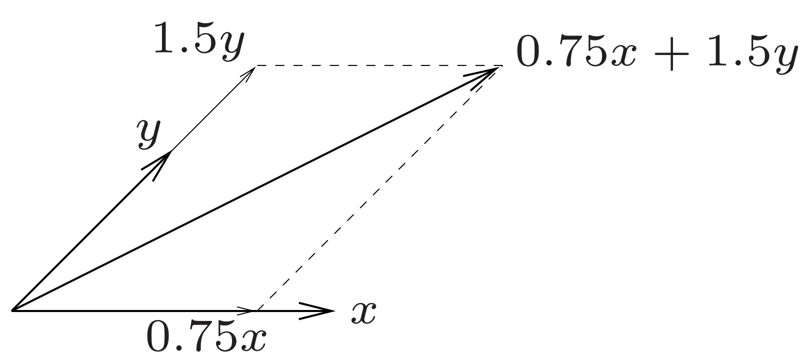
Vector operations

multiplying a vector $x \in \mathbf{R}^n$ with a scalar $\alpha \in \mathbf{R}$:

$$\alpha x = \begin{bmatrix} \alpha x_1 \\ \vdots \\ \alpha x_n \end{bmatrix}$$

adding and subtracting two vectors x, $y \in \mathbf{R}^n$:

$$x + y = \begin{bmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{bmatrix}, \qquad x - y = \begin{bmatrix} x_1 - y_1 \\ \vdots \\ x_n - y_n \end{bmatrix}$$



Inner product

$x,\ y \in \mathbf{R}^n$

$$\langle x, y\rangle := x_1y_1 + x_2y_2 + \cdots + x_ny_n = x^Ty$$

important properties

- $\langle \alpha x, y\rangle = \alpha\langle x, y\rangle$
- $\langle x+y, z\rangle = \langle x, z\rangle + \langle y, z\rangle$
- $\langle x, y\rangle = \langle y, x\rangle$
- $\langle x, x\rangle \geq 0$
- $\langle x, x\rangle = 0 \Longleftrightarrow x = 0$

linear function: $f : \mathbf{R}^n \to \mathbf{R}$ is linear, *i.e.*

$$f(\alpha x + \beta y) = \alpha f(x) + \beta f(y),$$

if and only if $f(x) = \langle a, x\rangle$ for some a

Euclidean norm

for $x \in \mathbf{R}^n$ we define the (Euclidean) norm as

$$\|x\| = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2} = \sqrt{x^T x}$$

$\|x\|$ measures *length* of vector (from origin)

important properties:

- $\|\alpha x\| = |\alpha| \|x\|$ (homogeneity)
- $\|x + y\| \leq \|x\| + \|y\|$ (triangle inequality)
- $\|x\| \geq 0$ (nonnegativity)
- $\|x\| = 0 \iff x = 0$ (definiteness)

distance between vectors: $\mathbf{dist}(x, y) = \|x - y\|$

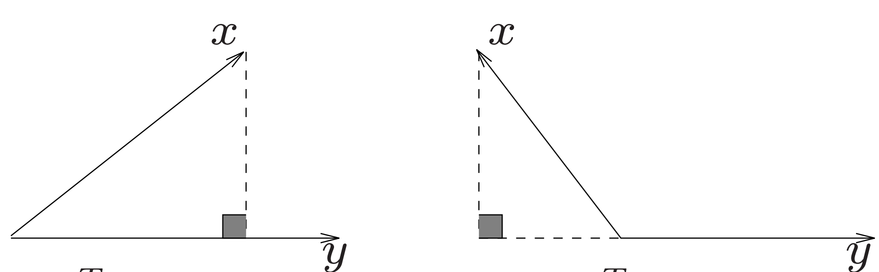
Inner products and angles

angle between vectors in $\mathbf{R}^n$:

$$\theta = \angle(x, y) = \cos^{-1} \frac{x^T y}{\|x\|\|y\|}$$

i.e., $x^T y = \|x\|\|y\| \cos\theta$

- x and y *aligned*: $\theta = 0$; $x^T y = \|x\|\|y\|$
- x and y *opposed*: $\theta = \pi$; $x^T y = -\|x\|\|y\|$
- x and y *orthogonal*: $\theta = \pi/2$ or $-\pi/2$; $x^T y = 0$ (denoted $x \perp y$)
- $x^T y > 0$ means $\angle(x, y)$ is acute; $x^T y < 0$ means $\angle(x, y)$ is obtuse

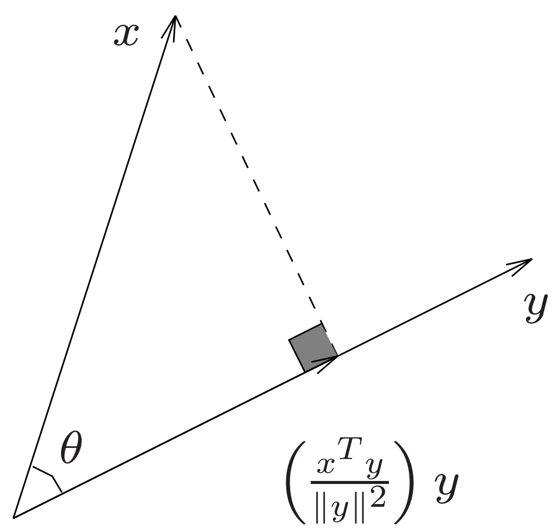


$x^T y > 0$ $\qquad$ $x^T y < 0$

Cauchy-Schwarz inequality:

$$|x^T y| \le \|x\| \|y\|$$

projection of x on y



projection is given by

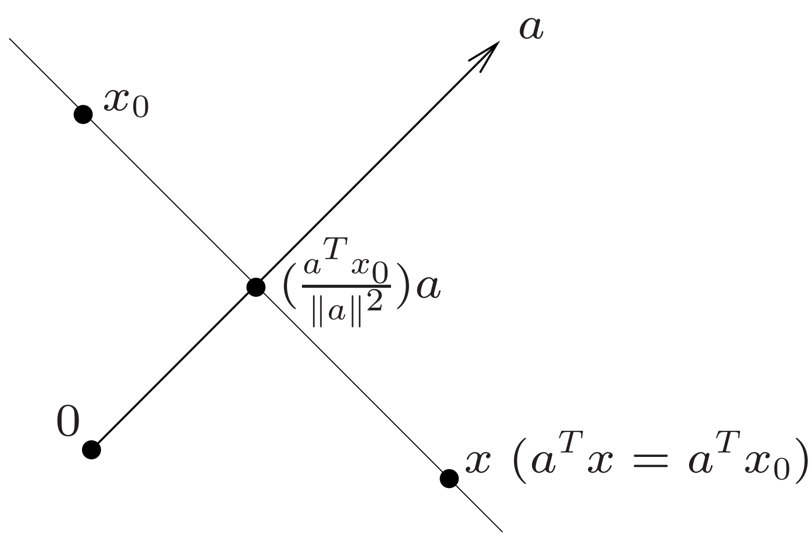
$$\left(\frac{x^T y}{\|y\|^2}\right) y$$

Hyperplanes

hyperplane in $\mathbf{R}^n$:

$$\{x \mid a^T x = b\} \quad (a \neq 0)$$

- solution set of one linear equation $a_1 x_1 + \cdots + a_n x_n = b$ with at least one $a_i \neq 0$
- set of vectors that make a constant inner product with vector $a = (a_1, \ldots, a_n)$ (the *normal* vector)



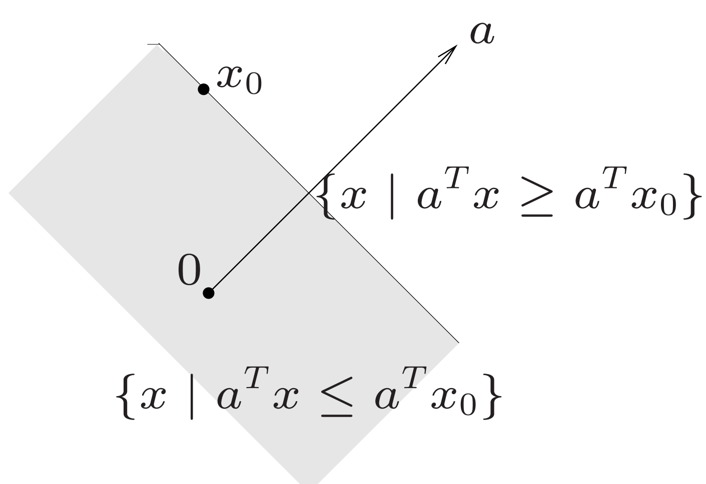
in $\mathbf{R}^2$: a line, in $\mathbf{R}^3$: a plane, . . .

Halfspaces

(closed) halfspace in $\mathbf{R}^n$:

$$\{x \mid a^T x \leq b\} \quad (a \neq 0)$$

- solution set of one linear inequality $a_1 x_1 + \cdots + a_n x_n \leq b$ with at least one $a_i \neq 0$
- $a = (a_1, \ldots, a_n)$ is the (outward) *normal*



- $\{x \mid a^T x < b\}$ is called an *open* halfspace

Affine sets

solution set of a set of linear equations

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_1 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{aligned}$$

intersection of m hyperplanes with normal vectors $a_i = (a_{i1}, a_{i2}, \ldots, a_{in})$ (w.l.o.g., all $a_i \neq 0$)

in matrix notation:

$$Ax = b$$

with

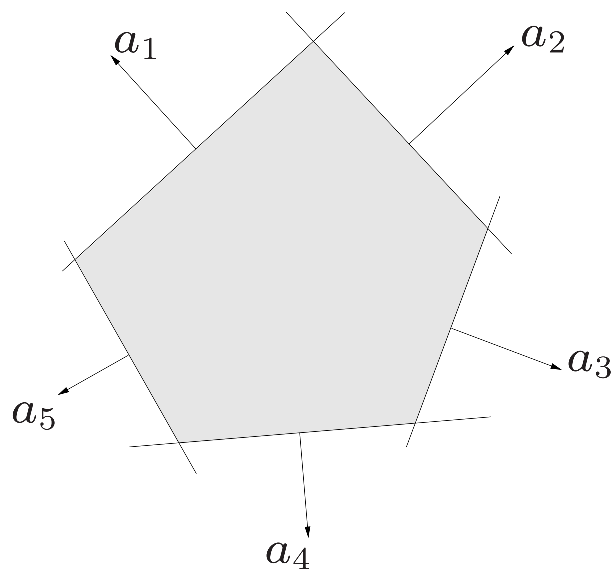
$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, \qquad b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

Polyhedra

solution set of system of linear inequalities

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &\leq b_1 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &\leq b_m \end{aligned}$$

intersection of m halfspaces, with normal vectors $a_i = (a_{i1}, a_{i2}, \ldots, a_{in})$ (w.l.o.g., all $a_i \neq 0$)



matrix notation

$$Ax \le b$$

with

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, \qquad b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

$Ax \le b$ stands for *componentwise* inequality, $i.e.$, for y, $z \in \mathbf{R}^n$,

$$y \le z \quad \Longleftrightarrow \quad y_1 \le z_1, \ldots, y_n \le z_n$$

Examples of polyhedra

- a hyperplane $\{x \mid a^T x = b\}$:

$$a^T x \leq b, \qquad a^T x \geq b$$

- solution set of system of linear equations/inequalities

$$a_i^T x \leq b_i, \quad i = 1, \ldots, m, \qquad c_i^T x = d_i, \quad i = 1, \ldots, p$$

- a slab $\{x \mid b_1 \leq a^T x \leq b_2\}$
- the probability simplex $\{x \in \mathbf{R}^n \mid \mathbf{1}^T x = 1, \;\; x_i \geq 0, \; i = 1, \ldots, n\}$
- (hyper)rectangle $\{x \in \mathbf{R}^n \mid l \leq x \leq u\}$ where $l < u$

Lecture 3
Geometry of linear programming

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

Subspaces

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

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

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

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

- $\mathcal{S} = \mathbf{R}^n$, $\mathcal{S} = \{0\}$
- $\mathcal{S} = \{\alpha v \, | \alpha \in \mathbf{R}\}$ 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 $Ax_0 = y$

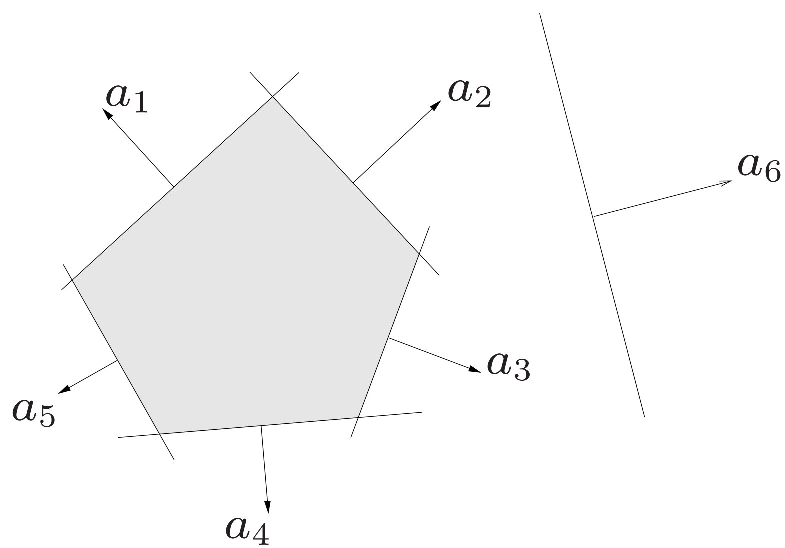
A square and invertible: unique solution for every y:

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

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:

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

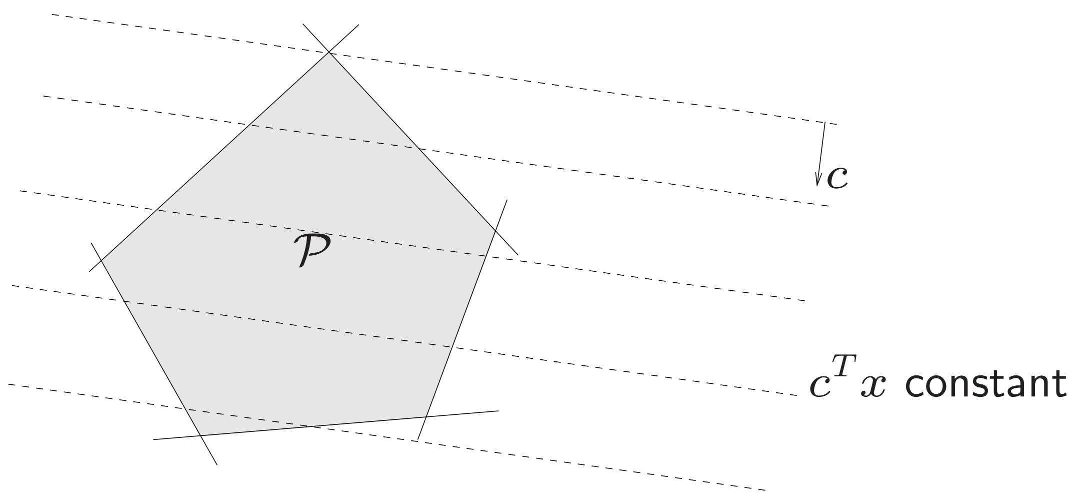
i.e., the *line segment* between any two points in $\mathcal{P}$ lies in $\mathcal{P}$

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

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

- (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 $x \in \mathcal{P}$

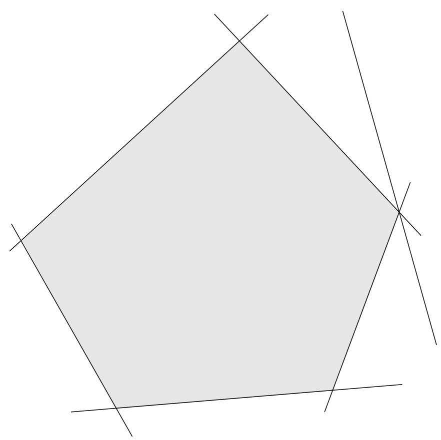
Degeneracy

set of linear inequalities $a_i^T x \leq b_i$, $i = 1, \ldots, m$

a basic feasible solution $x^\star$ with

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

is *degenerate* if #indices in I is greater than n



- a property of the *description* of the polyhedron, not its geometry
- affects the performance of some algorithms
- disappears with small perturbations of b

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

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

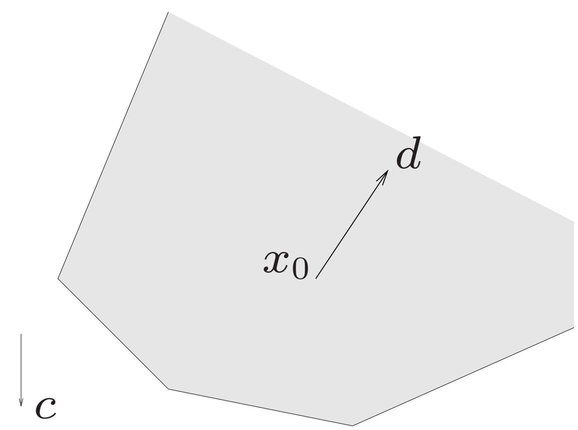
- $c = (1, 1)$: $X_{\text{opt}} = \{(0, 0)\}$, $p^\star = 0$
- $c = (1, 0)$: $X_{\text{opt}} = \{(0, x_2) \mid 0 \leq x_2 \leq 1\}$, $p^\star = 0$
- $c = (-1, -1)$: $X_{\text{opt}} = \emptyset$, $p^\star = -\infty$

Existence of optimal points

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

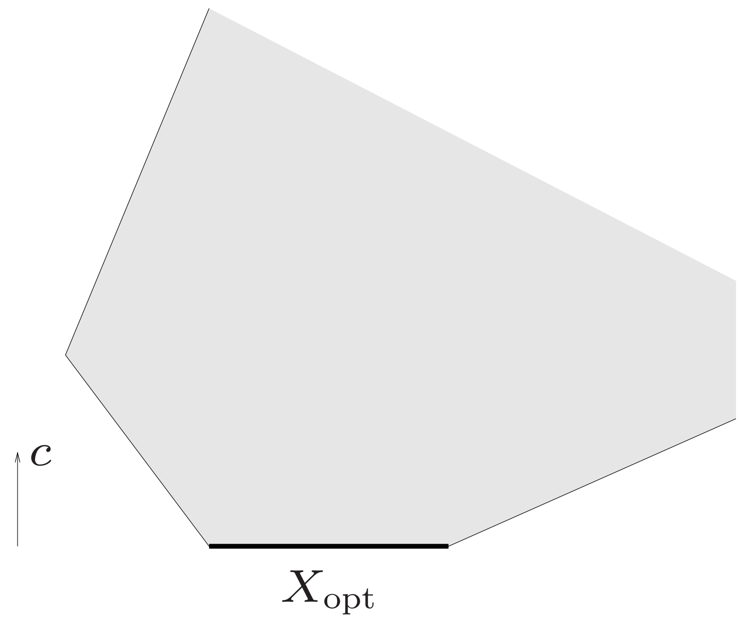
$$\{x_0 + td \mid t \geq 0\}$$

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



Lecture 4
The linear programming problem: variants and examples

- variants of the linear programming problem
- LP feasibility problem
- examples and some general applications
- linear-fractional programming

Variants of the linear programming problem

general form

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & a_i^T x \leq b_i, \quad i = 1, \ldots, m \\ & g_i^T x = h_i, \quad i = 1, \ldots, p \end{array}$$

in matrix notation:

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

where

$$A = \begin{bmatrix} a_1^T \\ a_2^T \\ \vdots \\ a_m^T \end{bmatrix} \in \mathbf{R}^{m \times n}, \qquad G = \begin{bmatrix} g_1^T \\ g_2^T \\ \vdots \\ g_p^T \end{bmatrix} \in \mathbf{R}^{p \times n}$$

inequality form LP

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & a_i^T x \leq b_i, \quad i = 1, \ldots, m \end{array}$$

in matrix notation:

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

standard form LP

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & g_i^T x = h_i, \quad i = 1, \ldots, m \\ & x \geq 0 \end{array}$$

in matrix notation:

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Gx = h \\ & x \geq 0 \end{array}$$

Reduction of general LP to inequality/standard form

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & a_i^T x \leq b_i, \quad i = 1, \ldots, m \\ & g_i^T x = h_i, \quad i = 1, \ldots, p \end{array}$$

reduction to inequality form:

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & a_i^T x \leq b_i, \quad i = 1, \ldots, m \\ & g_i^T x \geq h_i, \quad i = 1, \ldots, p \\ & g_i^T x \leq h_i, \quad i = 1, \ldots, p \end{array}$$

in matrix notation (where A has rows a_i^T, G has rows g_i^T)

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & \begin{bmatrix} A \\ -G \\ G \end{bmatrix} x \leq \begin{bmatrix} b \\ -h \\ h \end{bmatrix} \end{array}$$

reduction to standard form:

$$\begin{array}{ll} \text{minimize} & c^T x^+ - c^T x^- \\ \text{subject to} & a_i^T x^+ - a_i^T x^- + s_i = b_i, \quad i = 1, \ldots, m \\ & g_i^T x^+ - g_i^T x^- = h_i, \quad i = 1, \ldots, p \\ & x^+, x^-, s \geq 0 \end{array}$$

- variables x^+, x^-, s
- recover x as $x = x^+ - x^-$
- $s \in \mathbf{R}^m$ is called a *slack* variable

in matrix notation:

$$\begin{array}{ll} \text{minimize} & \widetilde{c}^T \widetilde{x} \\ \text{subject to} & \widetilde{G}\widetilde{x} = \widetilde{h} \\ & \widetilde{x} \geq 0 \end{array}$$

where

$$\widetilde{x} = \begin{bmatrix} x^+ \\ x^- \\ s \end{bmatrix}, \qquad \widetilde{c} = \begin{bmatrix} c \\ -c \\ 0 \end{bmatrix}, \qquad \widetilde{G} = \begin{bmatrix} A & -A & I \\ G & -G & 0 \end{bmatrix}, \qquad \widetilde{h} = \begin{bmatrix} b \\ h \end{bmatrix}$$

LP feasibility problem

feasibility problem: find x that satisfies $a_i^T x \le b_i$, $i = 1, \ldots, m$

solution via LP (with variables t, x)

$$\begin{array}{ll} \text{minimize} & t \\ \text{subject to} & a_i^T x \le b_i + t, \quad i = 1, \ldots, m \end{array}$$

- variables t, x
- if minimizer $x^\star$, $t^\star$ satisfies $t^\star \le 0$, then $x^\star$ satisfies the inequalities

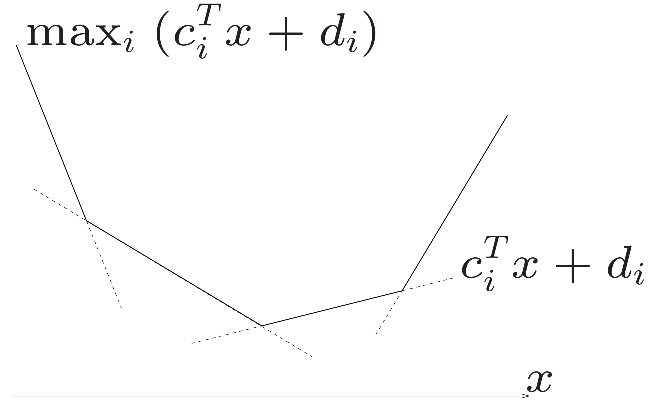
LP in matrix notation:

$$\begin{array}{ll} \text{minimize} & \widetilde{c}^T \widetilde{x} \\ \text{subject to} & \widetilde{A}\widetilde{x} \le \widetilde{b} \end{array}$$

$$\widetilde{x} = \begin{bmatrix} x \\ t \end{bmatrix}, \qquad \widetilde{c} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \qquad \widetilde{A} = \begin{bmatrix} A & -\mathbf{1} \end{bmatrix}, \qquad \widetilde{b} = b$$

Piecewise-linear minimization

piecewise-linear minimization: minimize $\max_{i=1,\ldots,m}(c_i^T x + d_i)$



equivalent LP (with variables $x \in \mathbf{R}^n$, $t \in \mathbf{R}$):

$$\begin{array}{ll} \text{minimize} & t \\ \text{subject to} & c_i^T x + d_i \leq t, \quad i = 1, \ldots, m \end{array}$$

in matrix notation:

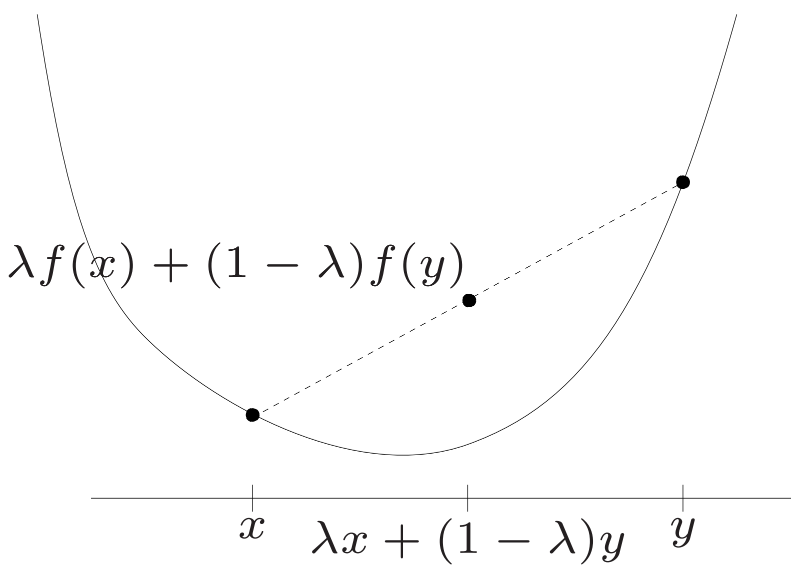
$$\begin{array}{ll} \text{minimize} & \widetilde{c}^T \widetilde{x} \\ \text{subject to} & \widetilde{A}\widetilde{x} \leq \widetilde{b} \end{array}$$

$$\widetilde{x} = \left[\begin{array}{c} x \\ t \end{array}\right], \qquad \widetilde{c} = \left[\begin{array}{c} 0 \\ 1 \end{array}\right], \qquad \widetilde{A} = \left[\begin{array}{cc} C & -\mathbf{1} \end{array}\right], \qquad \widetilde{b} = \left[\begin{array}{c} -d \end{array}\right]$$

Convex functions

$f : \mathbf{R}^n \to \mathbf{R}$ is convex if for $0 \le \lambda \le 1$

$$f(\lambda x + (1-\lambda)y) \le \lambda f(x) + (1-\lambda)f(y)$$

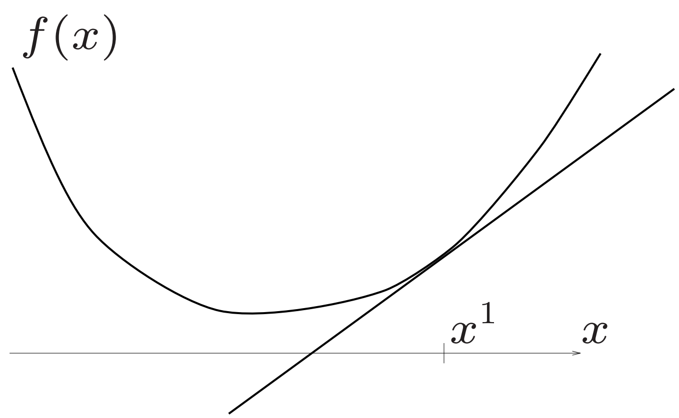


Piecewise-linear approximation

assume $f : \mathbf{R}^n \to \mathbf{R}$ differentiable and convex

- 1st-order approximation at x^1 is a *global lower bound* on f:

$$f(x) \geq f(x^1) + \nabla f(x^1)^T(x - x^1)$$



- evaluating f, ∇f at several x^i yields a *piecewise-linear* lower bound:

$$f(x) \geq \max_{i=1,\ldots,K} \left(f(x^i) + \nabla f(x^i)^T(x - x^i) \right)$$

Convex optimization problem

$$\text{minimize} \quad f_0(x)$$

(f_i convex and differentiable)

LP approximation (choose points x^j, $j = 1, \ldots, K$):

$$\begin{array}{ll} \text{minimize} & t \\ \text{subject to} & f_0(x^j) + \nabla f_0(x^j)^T(x - x^j) \leq t, \quad j = 1, \ldots, K \end{array}$$

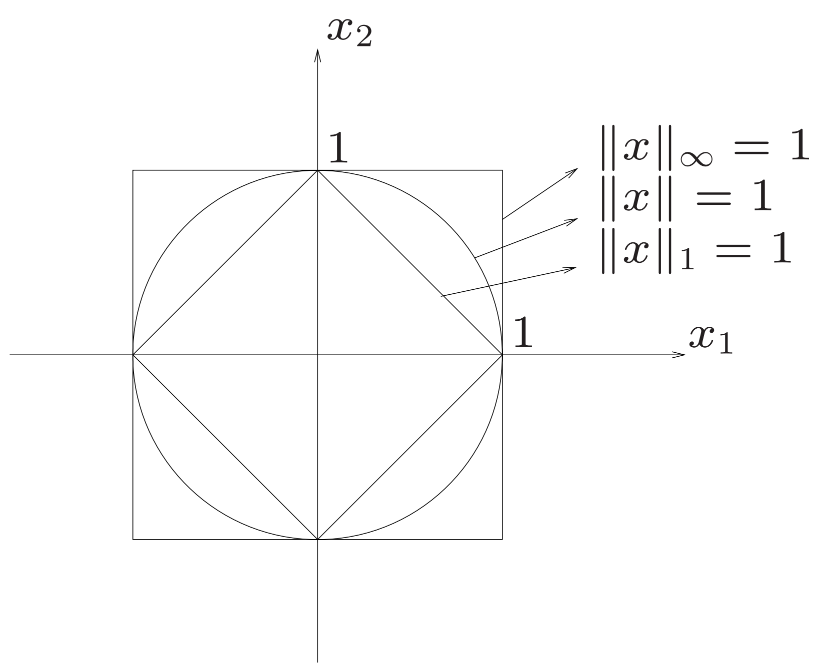
(variables x, t)

- yields lower bound on optimal value
- can be extended to nondifferentiable convex functions
- more sophisticated variation: cutting-plane algorithm (solves convex optimization problem via sequence of LP approximations)

Norms

norms on $\mathbf{R}^n$:

- Euclidean norm $\|x\|$ (or $\|x\|_2$) $= \sqrt{x_1^2 + \cdots + x_n^2}$
- ℓ_1-norm: $\|x\|_1 = |x_1| + \cdots + |x_n|$
- ℓ_∞- (or Chebyshev-) norm: $\|x\|_\infty = \max_i |x_i|$



Norm approximation problems

$$\text{minimize} \quad \|Ax - b\|_p$$

- $x \in \mathbf{R}^n$ is variable; $A \in \mathbf{R}^{m \times n}$ and $b \in \mathbf{R}^m$ are problem data
- $p = 1, 2, \infty$
- $r = Ax - b$ is called *residual*
- $r_i = a_i^T x - b_i$ is ith residual (a_i^T is ith row of A)
- usually overdetermined, *i.e.*, $b \notin \mathcal{R}(A)$ (*e.g.*, $m > n$, A full rank)

interpretations:

- approximate or fit b with linear combination of columns of A
- b is corrupted measurement of Ax; find 'least inconsistent' value of x for given measurements

examples:

- $\|r\| = \sqrt{r^T r}$: least-squares or ℓ_2-approximation (a.k.a. regression)

- $\|r\| = \max_i |r_i|$: Chebyshev, ℓ_∞, or minimax approximation

- $\|r\| = \sum_i |r_i|$: absolute-sum or ℓ_1-approximation

solution:

- ℓ_2: closed form expression

$$x_{\text{opt}} = (A^T A)^{-1} A^T b$$

  (assume $\mathbf{rank}(A) = n$)
- ℓ_1, ℓ_∞: no closed form expression, but readily solved via LP

ℓ_1-approximation via LP

ℓ_1-approximation problem

$$\text{minimize} \quad \|Ax - b\|_1$$

write as

$$\begin{array}{ll} \text{minimize} & \sum_{i=1}^m y_i \\ \text{subject to} & -y \leq Ax - b \leq y \end{array}$$

an LP with variables y, x:

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

with

$$\widetilde{x} = \begin{bmatrix} x \\ y \end{bmatrix}, \qquad \widetilde{c} = \begin{bmatrix} 0 \\ \mathbf{1} \end{bmatrix}, \qquad \widetilde{A} = \begin{bmatrix} A & -I \\ -A & -I \end{bmatrix}, \qquad \widetilde{b} = \begin{bmatrix} b \\ -b \end{bmatrix}$$

ℓ_∞-approximation via LP

ℓ_∞-approximation problem

$$\text{minimize} \quad \|Ax - b\|_\infty$$

write as

$$\begin{array}{ll} \text{minimize} & t \\ \text{subject to} & -t\mathbf{1} \leq Ax - b \leq t\mathbf{1} \end{array}$$

an LP with variables t, x:

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

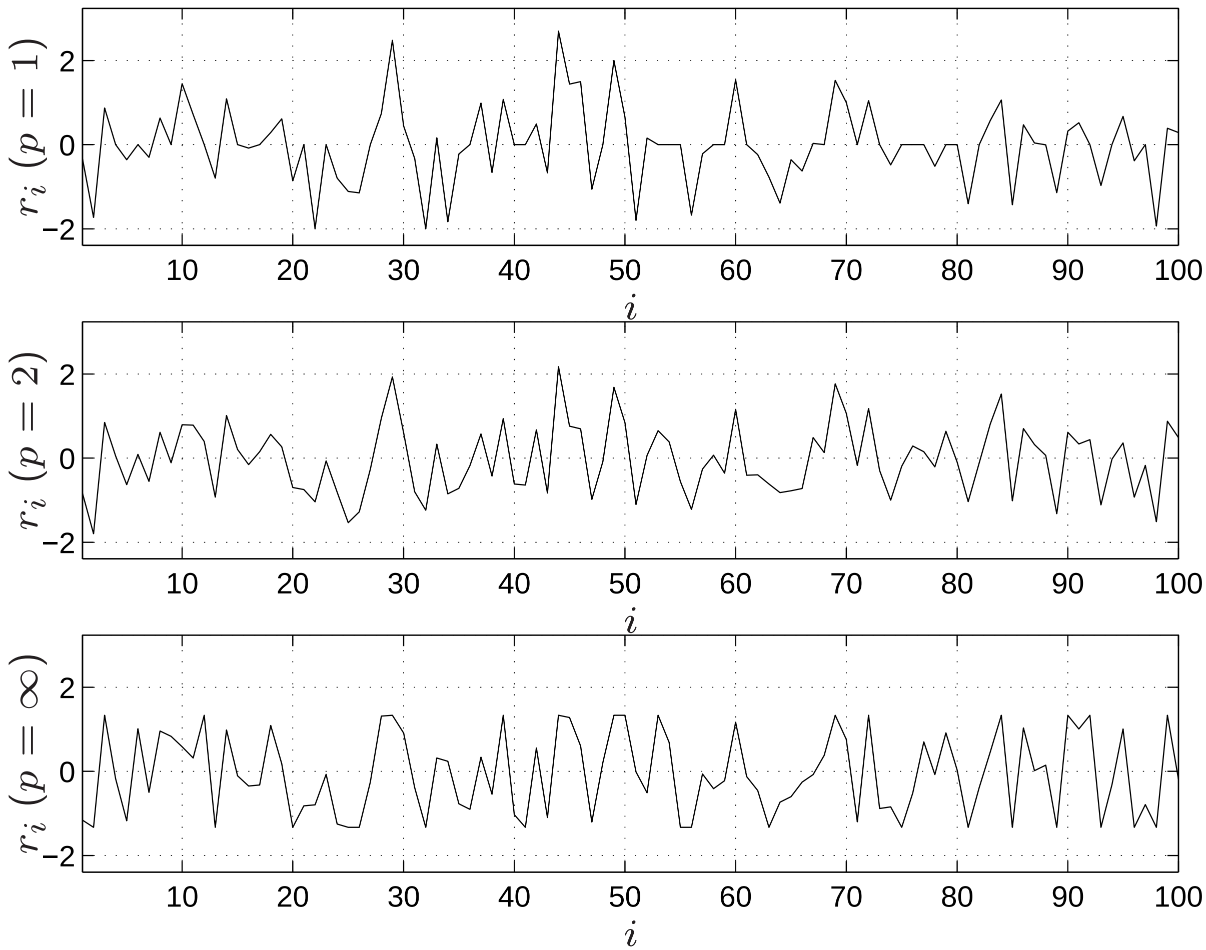
with

$$\widetilde{x} = \begin{bmatrix} x \\ t \end{bmatrix}, \qquad \widetilde{c} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \qquad \widetilde{A} = \begin{bmatrix} A & -\mathbf{1} \\ -A & -\mathbf{1} \end{bmatrix}, \qquad \widetilde{b} = \begin{bmatrix} b \\ -b \end{bmatrix}$$

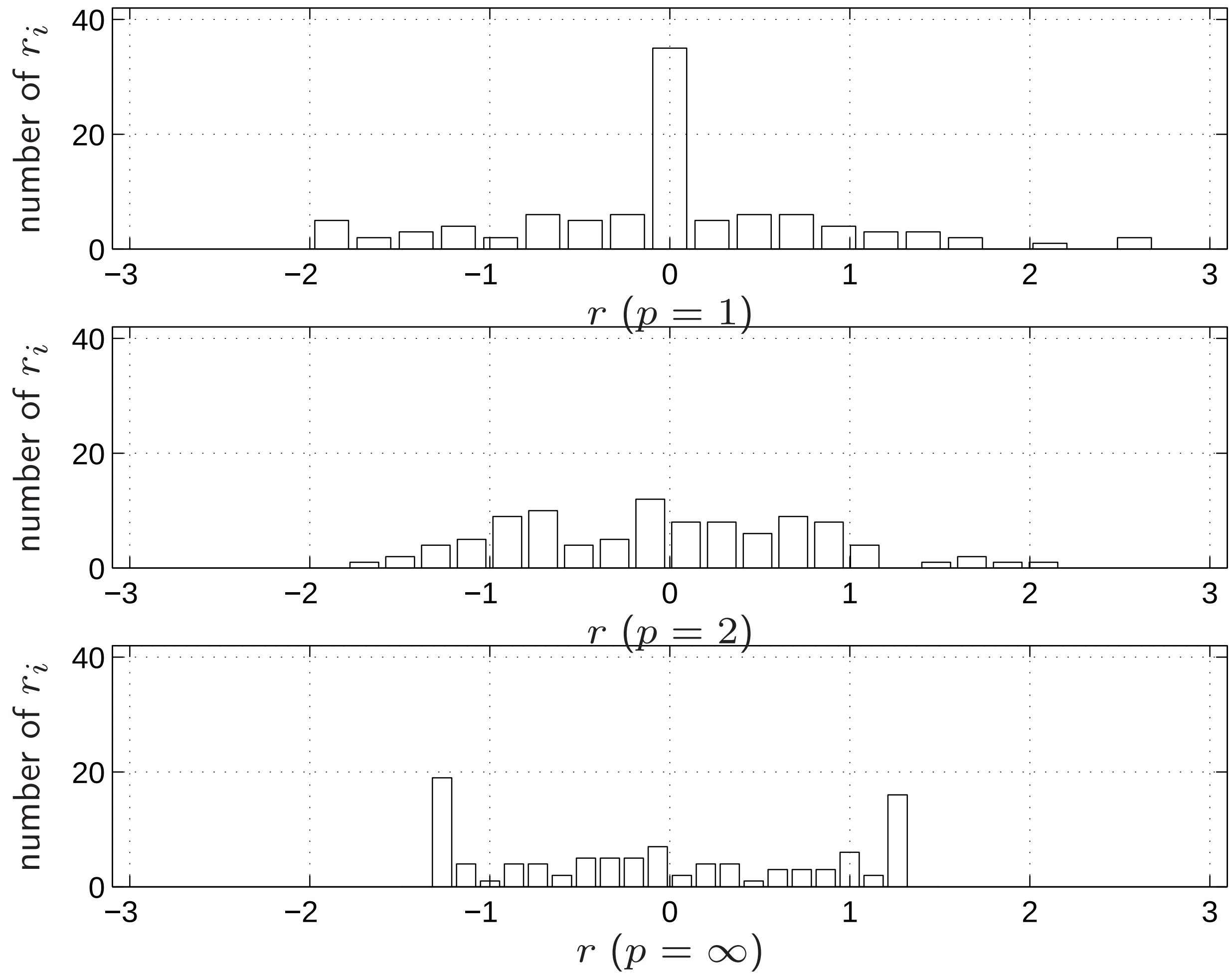
Example

minimize $\|Ax - b\|_p$ for $p = 1,\ 2,\ \infty$ ($A \in \mathbf{R}^{100\times 30}$)

resulting residuals:



histogram of residuals:



- $p = \infty$ gives 'thinnest' distribution; $p = 1$ gives widest distribution
- $p = 1$ most very small (or even zero) r_i

Interpretation: maximum likelihood estimation

m linear measurements $y_1, \ldots, y_m$ of $x \in \mathbf{R}^n$:

$$y_i = a_i^T x + v_i, \quad i = 1, \ldots, m$$

- v_i: measurement noise, IID with density p
- y is a random variable with density $p_x(y) = \prod_{i=1}^m p(y_i - a_i^T x)$

log-likelihood function is defined as

$$\log p_x(y) = \sum_{i=1}^m \log p(y_i - a_i^T x)$$

maximum likelihood (ML) estimate of x is

$$\hat{x} = \underset{x}{\operatorname{argmax}} \sum_{i=1}^m \log p(y_i - a_i^T x)$$

examples

- v_i Gaussian: $p(z) = 1/(\sqrt{2\pi}\sigma)e^{-z^2/2\sigma^2}$

  ML estimate is ℓ_2-estimate $\hat{x} = \mathrm{argmin}_x \|Ax - y\|_2$

- v_i double-sided exponential: $p(z) = (1/2a)e^{-|z|/a}$

  ML estimate is ℓ_1-estimate $\hat{x} = \mathrm{argmin}_x \|Ax - y\|_1$

- v_i is one-sided exponential: $p(z) = \begin{cases} (1/a)e^{-z/a} & z \geq 0 \\ 0 & z < 0 \end{cases}$

  ML estimate is found by solving LP

$$\begin{array}{ll} \text{minimize} & \mathbf{1}^T(y - Ax) \\ \text{subject to} & y - Ax \geq 0 \end{array}$$

- v_i are uniform on $[-a, a]$: $p(z) = \begin{cases} 1/(2a) & -a \leq z \leq a \\ 0 & \text{otherwise} \end{cases}$

  ML estimate is any x satisfying $\|Ax - y\|_\infty \leq a$

Linear-fractional programming

$$\begin{array}{ll} \text{minimize} & \dfrac{c^T x + d}{f^T x + g} \\ \text{subject to} & Ax \le b \\ & f^T x + g \ge 0 \end{array}$$

(asume $a/0 = +\infty$ if $a > 0$, $a/0 = -\infty$ if $a \le 0$)

- nonlinear objective function
- like LP, can be solved very efficiently

equivalent form with linear objective (vars. x, γ):

$$\begin{array}{ll} \text{minimize} & \gamma \\ \text{subject to} & c^T x + d \le \gamma(f^T x + g) \\ & f^T x + g \ge 0 \\ & Ax \le b \end{array}$$

Bisection algorithm for linear-fractional programming

given: interval $[l, u]$ that contains optimal γ
repeat: solve feasibility problem for $\gamma = (u + l)/2$

$$\begin{array}{l} c^T x + d \leq \gamma(f^T x + g) \\ f^T x + g \geq 0 \\ Ax \leq b \end{array}$$

if feasible $u := \gamma$; if infeasible $l := \gamma$
until $u - l \leq \epsilon$

- each iteration is an LP feasibility problem
- accuracy doubles at each iteration
- number of iterations to reach accuracy ϵ starting with initial interval of width $u - l = \epsilon_0$:

$$k = \lceil \log_2(\epsilon_0/\epsilon) \rceil$$

Generalized linear-fractional programming

$$\begin{array}{ll} \text{minimize} & \max_{i=1,\ldots,K} \dfrac{c_i^T x + d_i}{f_i^T x + g_i} \\ \text{subject to} & Ax \leq b \\ & f_i^T x + g_i \geq 0, \quad i = 1, \ldots, K \end{array}$$

equivalent formulation:

$$\begin{array}{ll} \text{minimize} & \gamma \\ \text{subject to} & Ax \leq b \\ & c_i^T x + d_i \leq \gamma (f_i^T x + g_i), \quad i = 1, \ldots, K \\ & f_i^T x + g_i \geq 0, \quad i = 1, \ldots, K \end{array}$$

- efficiently solved via bisection on γ
- each iteration is an LP feasibility problem

Von Neumann economic growth problem

simple model of an economy: m goods, n economic sectors

- $x_i(t)$: 'activity' of sector i in current period t
- $a_i^T x(t)$: amount of good i consumed in period t
- $b_i^T x(t)$: amount of good i produced in period t

choose $x(t)$ to maximize *growth rate* $\min_i x_i(t+1)/x_i(t)$:

$$\begin{array}{ll} \text{maximize} & \gamma \\ \text{subject to} & Ax(t+1) \leq Bx(t), \quad x(t+1) \geq \gamma x(t), \quad x(t) \geq \mathbf{1} \end{array}$$

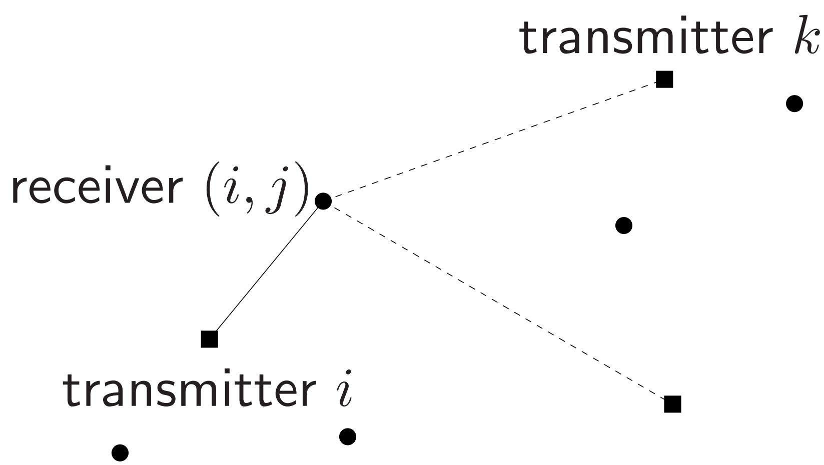
or equivalently (since $a_{ij} \geq 0$):

$$\begin{array}{ll} \text{maximize} & \gamma \\ \text{subject to} & \gamma Ax(t) \leq Bx(t), \quad x(t) \geq \mathbf{1} \end{array}$$

(linear-fractional problem with variables $x(0)$, γ)

Optimal transmitter power allocation

- m transmitters, mn receivers all at same frequency
- transmitter i wants to transmit to n receivers labeled (i, j), $j = 1, \ldots, n$



- A_{ijk} is path gain from transmitter k to receiver (i, j)
- N_{ij} is (self) noise power of receiver (i, j)
- variables: transmitter powers p_k, $k = 1, \ldots, m$

at receiver (i, j):

- signal power: $S_{ij} = A_{iji}p_i$
- noise plus interference power: $I_{ij} = \sum_{k \neq i} A_{ijk}p_k + N_{ij}$
- signal to interference/noise ratio (SINR): S_{ij}/I_{ij}

problem: choose p_i to maximize smallest SINR:

$$\begin{array}{ll} \text{maximize} & \displaystyle\min_{i,j} \frac{A_{iji}p_i}{\sum_{k \neq i} A_{ijk}p_k + N_{ij}} \\ \text{subject to} & 0 \leq p_i \leq p_{\max} \end{array}$$

- a (generalized) linear-fractional program
- special case with analytical solution: $m = 1$, no upper bound on p_i (see exercises)

Lecture 5
Applications in control

- optimal input design
- robust optimal input design
- pole placement (with low-authority control)

Linear dynamical system

$$y(t) = h_0 u(t) + h_1 u(t-1) + h_2 u(t-2) + \cdots$$

- single input/single output: input $u(t) \in \mathbf{R}$, output $y(t) \in \mathbf{R}$
- h_i are called *impulse response* coefficients
- *finite impulse response* (FIR) system of order k: $h_i = 0$ for $i > k$

if $u(t) = 0$ for $t < 0$,

$$\begin{bmatrix} y(0) \\ y(1) \\ y(2) \\ \vdots \\ y(N) \end{bmatrix} = \begin{bmatrix} h_0 & 0 & 0 & \cdots & 0 \\ h_1 & h_0 & 0 & \cdots & 0 \\ h_2 & h_1 & h_0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ h_N & h_{N-1} & h_{N-2} & \cdots & h_0 \end{bmatrix} \begin{bmatrix} u(0) \\ u(1) \\ u(2) \\ \vdots \\ u(N) \end{bmatrix}$$

a linear mapping from input to output sequence

Output tracking problem

choose inputs $u(t)$, $t = 0, \ldots, M$ $(M < N)$ that

- minimize *peak deviation* between $y(t)$ and a desired output $y_{\rm des}(t)$, $t = 0, \ldots, N$,

$$\max_{t=0,\ldots,N} |y(t) - y_{\rm des}(t)|$$

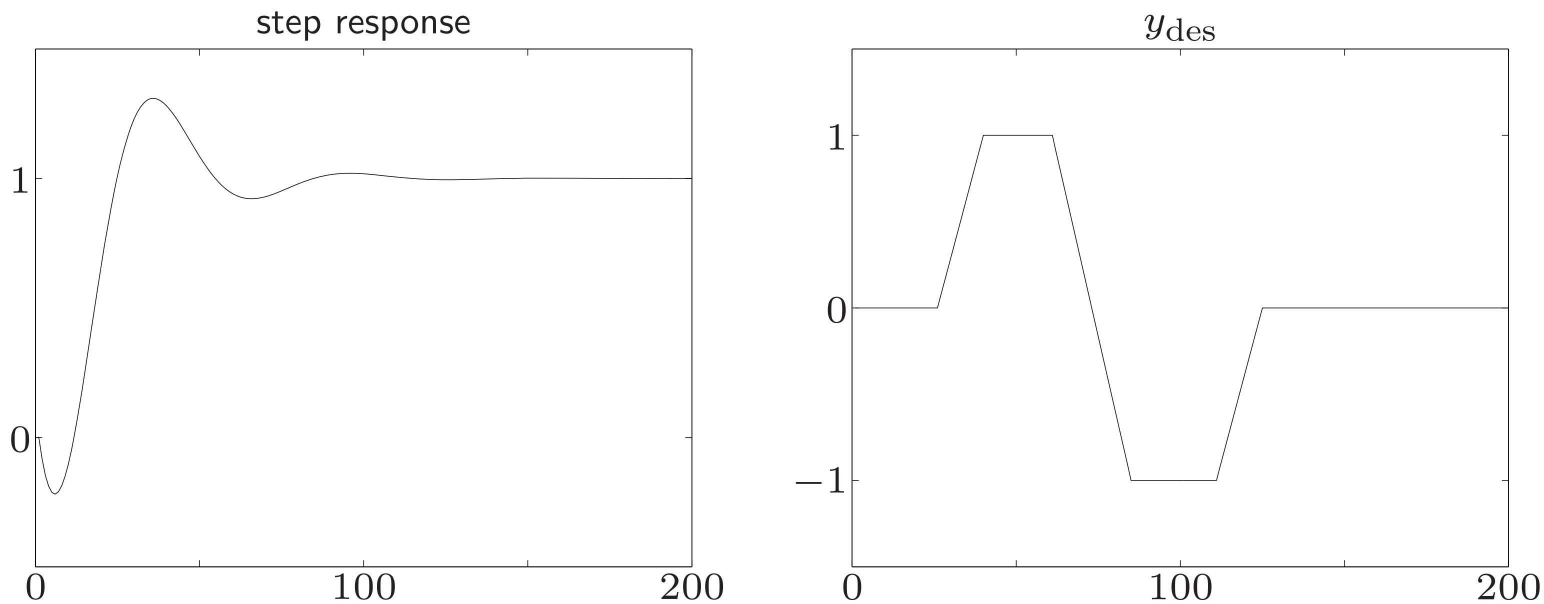
- satisfy amplitude and slew rate constraints:

$$|u(t)| \leq U, \quad |u(t+1) - u(t)| \leq S$$

as a linear program (variables: w, $u(0)$, . . . , $u(N)$):

$$\begin{array}{ll} \mbox{minimize.} & w \\ \mbox{subject to} & -w \leq \sum_{i=0}^{t} h_i u(t-i) - y_{\rm des}(t) \leq w, \quad t = 0, \ldots, N \\ & u(t) = 0, \quad t = M+1, \ldots, N \\ & -U \leq u(t) \leq U, \quad t = 0, \ldots, M \\ & -S \leq u(t+1) - u(t) \leq S, \quad t = 0, \ldots, M+1 \end{array}$$

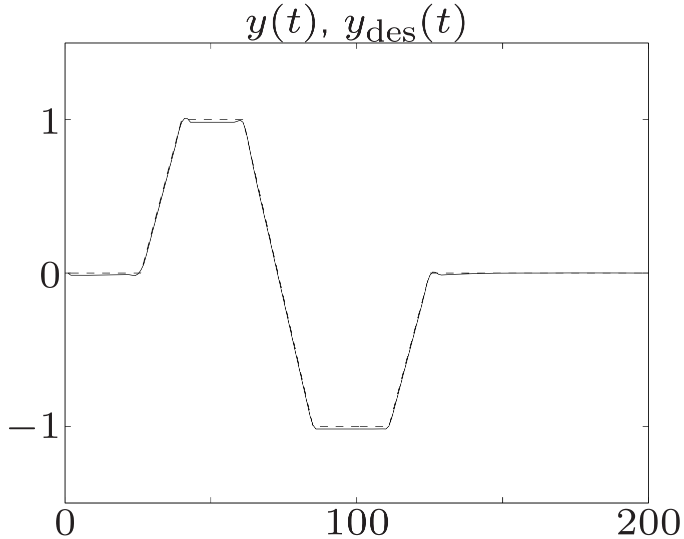
example. single input/output, $N = 200$



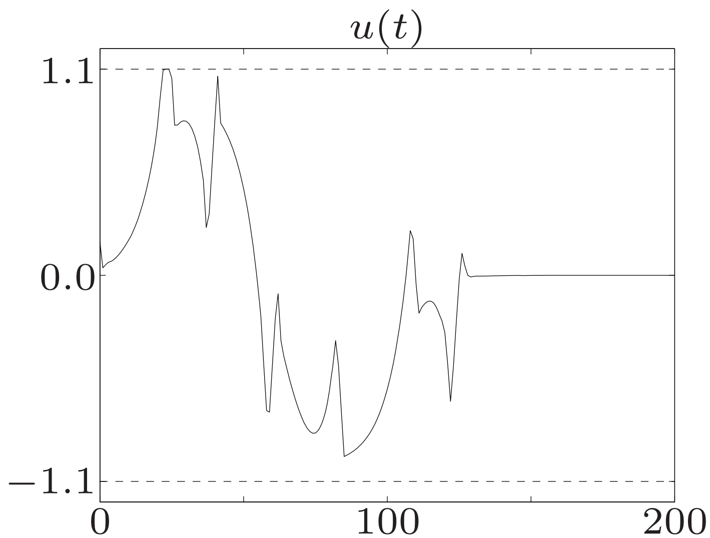
constraints on u:

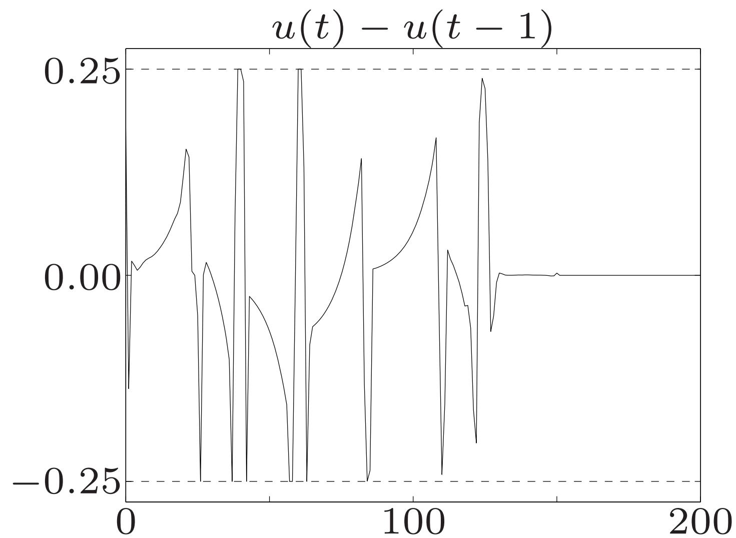
- input horizon $M = 150$
- amplitude constraint $|u(t)| \leq 1.1$
- slew rate constraint $|u(t) - u(t-1)| \leq 0.25$

output and desired output:



optimal input sequence u:





Robust output tracking (1)

- impulse response is not exactly known; it can take two values:

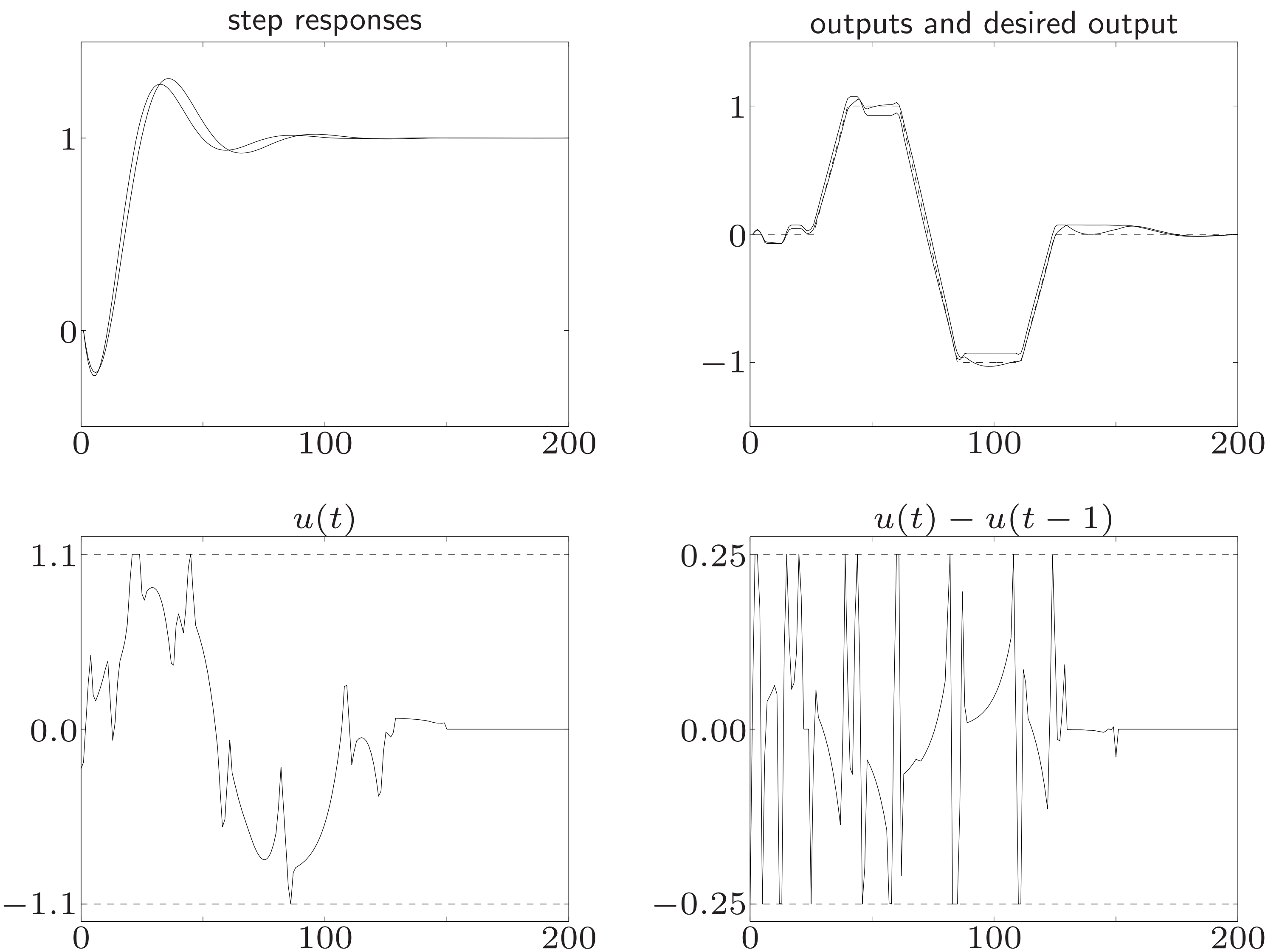
$$(h_0^{(1)}, h_1^{(1)}, \ldots, h_k^{(1)}), \qquad (h_0^{(2)}, h_1^{(2)}, \ldots, h_k^{(2)})$$

- design an input sequence that minimizes the worst-case peak tracking error

$$\begin{array}{ll}
\mbox{minimize} & w \\
\mbox{subject to} & -w \leq \sum_{i=0}^{t} h_i^{(1)} u(t-i) - y_{\mathrm{des}}(t) \leq w, \quad t = 0, \ldots, N \\
& -w \leq \sum_{i=0}^{t} h_i^{(2)} u(t-i) - y_{\mathrm{des}}(t) \leq w, \quad t = 0, \ldots, N \\
& u(t) = 0, \quad t = M+1, \ldots, N \\
& -U \leq u(t) \leq U, \quad t = 0, \ldots, M \\
& -S \leq u(t+1) - u(t) \leq S, \quad t = 0, \ldots, M+1
\end{array}$$

an LP in the variables w, $u(0)$, . . . , $u(N)$

example



step responses
1
0
0
100
200
outputs and desired output
1
0
−1
0
100
200
u(t)
1.1
0.0
−1.1
0
100
200
u(t) − u(t − 1)
0.25
0.00
−0.25
0
100
200

Robust output tracking (2)

$$\begin{bmatrix} h_0(s) \\ h_1(s) \\ \vdots \\ h_k(s) \end{bmatrix} = \begin{bmatrix} \bar{h}_0 \\ \bar{h}_1 \\ \vdots \\ \bar{h}_k \end{bmatrix} + s_1 \begin{bmatrix} v_0^{(1)} \\ v_1^{(1)} \\ \vdots \\ v_k^{(1)} \end{bmatrix} + \cdots + s_K \begin{bmatrix} v_0^{(K)} \\ v_1^{(K)} \\ \vdots \\ v_k^{(K)} \end{bmatrix}$$

$\bar{h}_i$ and $v_i^{(j)}$ are given; $s_i \in [-1,+1]$ is unknown

robust output tracking problem (variables w, $u(t)$):

$$\begin{array}{ll} \text{min.} & w \\ \text{s.t.} & -w \leq \sum_{i=0}^{t} h_i(s)u(t-i) - y_{\text{des}}(t) \leq w, \quad t = 0,\ldots,N, \quad \forall s \in [-1,1]^K \\ & u(t) = 0, \quad t = M+1,\ldots,N \\ & -U \leq u(t) \leq U, \quad t = 0,\ldots,M \\ & -S \leq u(t+1) - u(t) \leq S, \quad t = 0,\ldots,M+1 \end{array}$$

straightforward (and very inefficient) solution: enumerate all 2^K extreme values of s

simplification: we can express the 2^{K+1} linear inequalities

$$-w \leq \sum_{i=0}^{t} h_i(s)u(t-i) - y_{\mathrm{des}}(t) \leq w \text{ for all } s \in \{-1,1\}^K$$

as two nonlinear inequalities

$$\sum_{i=0}^{t} \bar{h}_i u(t-i) + \sum_{j=1}^{K} \left| \sum_{i=0}^{t} v_i^{(j)} u(t-i) \right| \leq y_{\mathrm{des}}(t) + w$$

$$\sum_{i=0}^{t} \bar{h}_i u(t-i) - \sum_{j=1}^{K} \left| \sum_{i=0}^{t} v_i^{(j)} u(t-i) \right| \geq y_{\mathrm{des}}(t) - w$$

proof:

$$\begin{aligned}
&\max_{s\in\{-1,1\}^K} \sum_{i=0}^{t} h_i(s)u(t-i) \\
&= \sum_{i=0}^{t} \bar{h}_i u(t-i) + \sum_{j=1}^{K} \max_{s_j\in\{-1,+1\}} s_j \sum_{i=0}^{t} v_i^{(j)} u(t-i) \\
&= \sum_{i=0}^{t} \bar{h}_i u(t-i) + \sum_{j=1}^{K} \left| \sum_{i=0}^{t} v_i^{(j)} u(t-i) \right|
\end{aligned}$$

and similarly for the lower bound

robust output tracking problem reduces to:

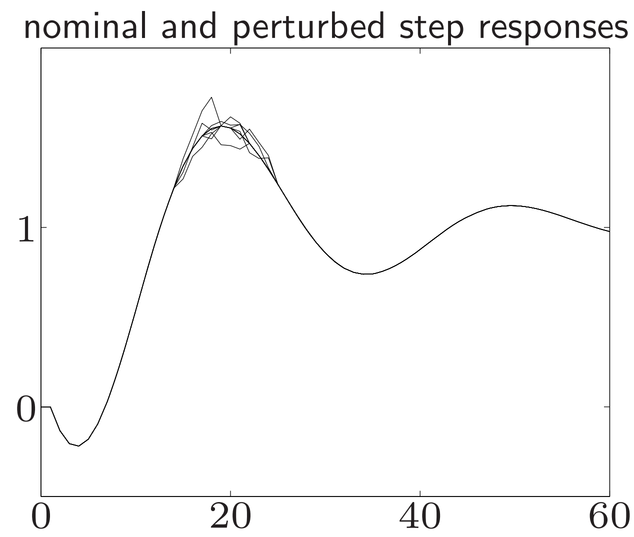
$$
\begin{array}{ll}
\text{min.} & w \\
\text{s.t.} & \sum_{i=0}^{t} \bar{h}_i u(t-i) + \sum_{j=1}^{K} \left| \sum_{i=0}^{t} v_i^{(j)} u(t-i) \right| \leq y_{\rm des}(t) + w, \quad t = 0, \ldots, N \\
& \sum_{i=0}^{t} \bar{h}_i u(t-i) - \sum_{j=1}^{K} \left| \sum_{i=0}^{t} v_i^{(j)} u(t-i) \right| \geq y_{\rm des}(t) - w, \quad t = 0, \ldots, N \\
& u(t) = 0, \quad t = M+1, \ldots, N \\
& -U \leq u(t) \leq U, \quad t = 0, \ldots, M \\
& -S \leq u(t+1) - u(t) \leq S, \quad t = 0, \ldots, M+1
\end{array}
$$

(variables $u(t)$, w)

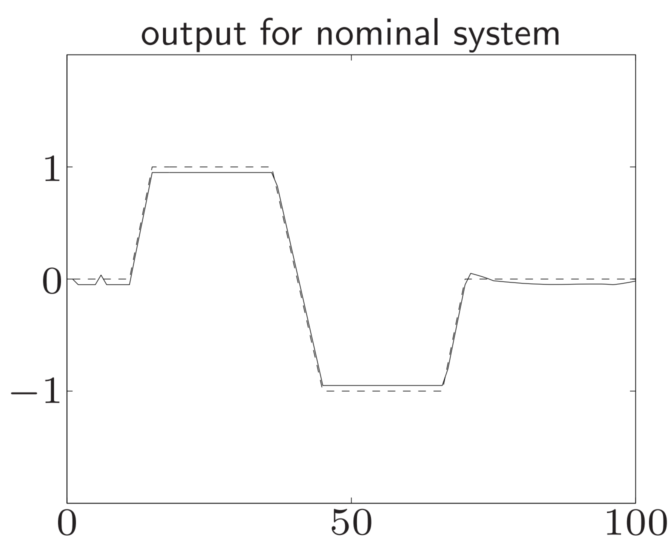
to express as an LP:

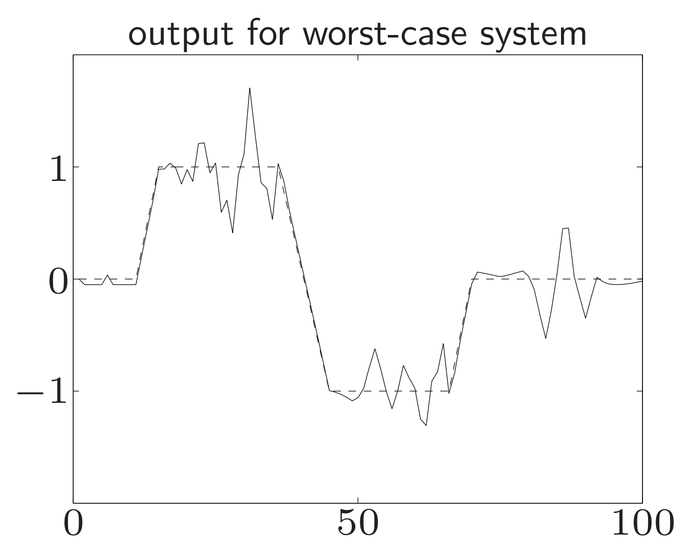
- for $t = 0, \ldots, N$, $j = 1, \ldots, K$, introduce new variables $p^{(j)}(t)$ and constraints
$$
-p^{(j)}(t) \leq \sum_{i=0}^{t} v_i^{(j)} u(t-i) \leq p^{(j)}(t)
$$
- replace $|\sum_i v_i^{(j)} u(t-i)|$ by $p^{(j)}(t)$

example $(K = 6)$



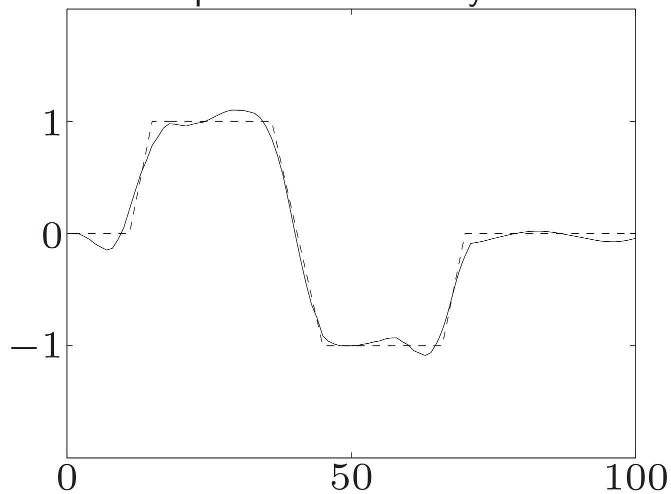
design for nominal system



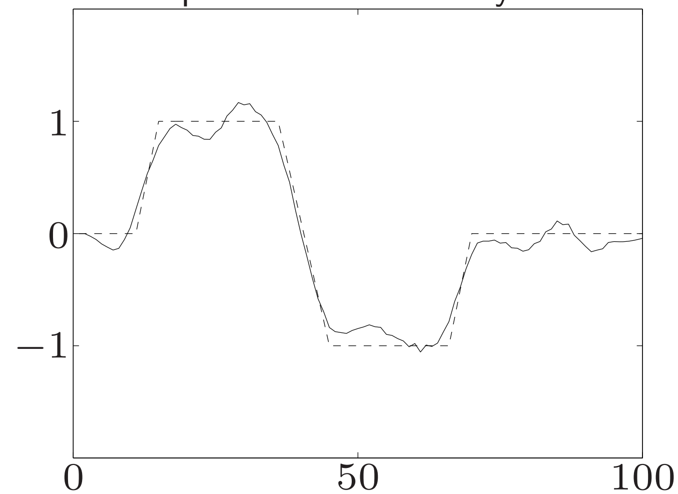


robust design

output for nominal system



output for worst-case system



State space description

input-output description:

$$y(t) = H_0 u(t) + H_1 u(t-1) + H_2 u(t-2) + \cdots$$

if $u(t) = 0$, $t < 0$:

$$\begin{bmatrix} y(0) \\ y(1) \\ y(2) \\ \vdots \\ y(N) \end{bmatrix} = \begin{bmatrix} H_0 & 0 & 0 & \cdots & 0 \\ H_1 & H_0 & 0 & \cdots & 0 \\ H_2 & H_1 & H_0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ H_N & H_{N-1} & H_{N-2} & \cdots & H_0 \end{bmatrix} \begin{bmatrix} u(0) \\ u(1) \\ u(2) \\ \vdots \\ u(N) \end{bmatrix}$$

block Toeplitz structure (constant along diagonals)

state space model:

$$x(t+1) = Ax(t) + Bu(t), \qquad y(t) = Cx(t) + Du(t)$$

with $H_0 = D$, $H_i = CA^{i-1}B$ $(i > 0)$

$x(t) \in \mathbf{R}^n$ is *state sequence* alternative description:

$$
\begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ y(0) \\ y(1) \\ \vdots \\ y(N) \end{bmatrix}
=
\begin{bmatrix}
A & -I & 0 & \cdots & 0 & B & 0 & \cdots & 0 \\
0 & A & -I & \cdots & 0 & 0 & B & \cdots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & \cdots & -I & 0 & 0 & \cdots & B \\
C & 0 & 0 & \cdots & 0 & D & 0 & \cdots & 0 \\
0 & C & 0 & \cdots & 0 & 0 & D & \cdots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & \cdots & C & 0 & 0 & \cdots & D
\end{bmatrix}
\begin{bmatrix} x(0) \\ x(1) \\ x(2) \\ \vdots \\ x(N) \\ u(0) \\ u(1) \\ \vdots \\ u(N) \end{bmatrix}
$$

- we don't eliminate the intermediate variables $x(t)$
- matrix is larger, but very sparse (interesting when using general-purpose LP solvers)

Pole placement

linear system

$$\dot{z}(t) = A(x)z(t), \qquad z(0) = z_0$$

where $A(x) = A_0 + x_1 A_1 + \cdots + x_p A_p \in \mathbf{R}^{n \times n}$

- solutions have the form

$$z_i(t) = \sum_k \beta_{ik} e^{\sigma_k t} \cos(\omega_k t - \phi_{ik})$$

  where $\lambda_k = \sigma_k \pm j\omega_k$ are the eigenvalues of $A(x)$

- $x \in \mathbf{R}^p$ is the design parameter
- goal: place eigenvalues of $A(x)$ in a desired region by choosing x

Low-authority control

eigenvalues of $A(x)$ are very complicated (nonlinear, nondifferentiable) functions of x

first-order perturbation: if $\lambda_i(A_0)$ is *simple*, then

$$\lambda_i(A(x)) = \lambda_i(A_0) + \sum_{k=1}^{p} \frac{w_i^* A_k v_i}{w_i^* v_i} x_k + o(\|x\|)$$

where w_i, v_i are the left and right eigenvectors:

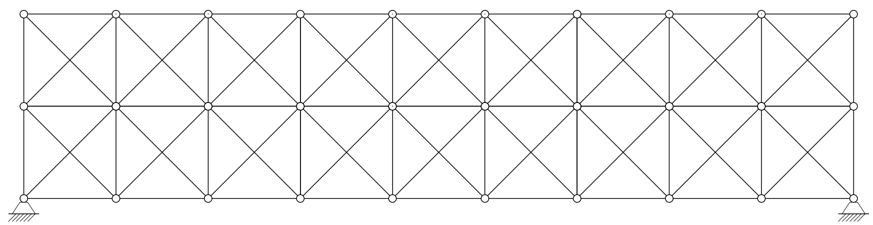
$$w_i^* A_0 = \lambda_i(A_0) w_i^*, \quad A_0 v_i = \lambda_i(A_0) v_i$$

'low-authority' control:

- use linear first-order approximations for λ_i
- can place λ_i in a polyhedral region by imposing linear inequalities on x
- we expect this to work only for small shifts in eigenvalues

Example

truss with 30 nodes, 83 bars



$$M\ddot{d}(t) + D\dot{d}(t) + Kd(t) = 0$$

- $d(t)$: vector of horizontal and vertical node displacements
- $M = M^T > 0$ (mass matrix): masses at the nodes
- $D = D^T > 0$ (damping matrix); $K = K^T > 0$ (stiffness matrix)

to increase damping, we attach dampers to the bars:

$$D(x) = D_0 + x_1 D_1 + \cdots + x_p D_p$$

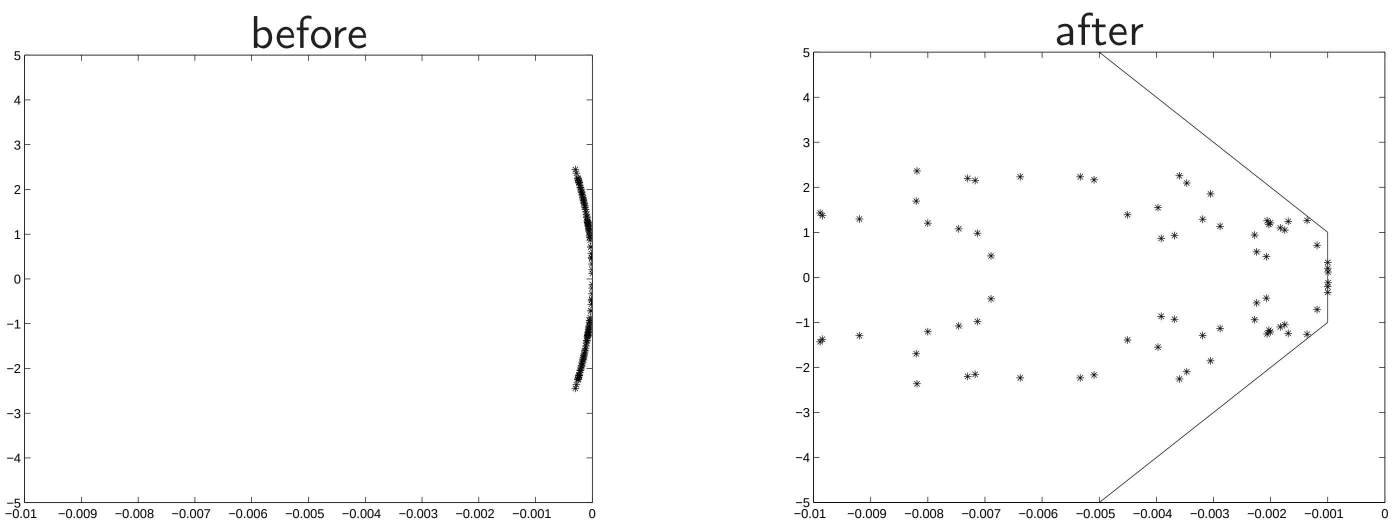
$x_i > 0$: amount of external damping at bar i

eigenvalue placement problem

$$\begin{array}{ll} \text{minimize} & \sum_{i=1}^{p} x_i \\ \text{subject to} & \lambda_i(M, D(x), K) \in \mathcal{C}, \quad i = 1, \ldots, n \\ & x \geq 0 \end{array}$$

an LP if $\mathcal{C}$ is polyhedral and we use the 1st order approximation for λ_i

eigenvalues



location of dampers