# Chapter VII
# Ellipsoidal Calculus

- linear algebra with uncertain data
- robust least-squares and variations
- worst-case simulation
- worst-case estimation
- finite-horizon control

# Lecture 27
# Linear equations with uncertain data

- uncertainty in linear algebra
- purpose of ellipsoidal calculus
- special case: interval calculus
- general case
- structured condition numbers



# Uncertainty in linear equations

Many problems can be modeled with linear equations of the forr

$$Ax = b,$$

with $A \in \mathbf{R}^{m \times n}$, $b \in \mathbf{R}^m$

In practice: $A, b$ are not well-known!

Gives rise to many possible solutions

What do we do?

Two issues:

- how do we **select** a particular candidate solution?
- evaluate how uncertainty on data results on **uncertainty on solution**

# Example

Consider a **scalar** equation

$$\mathbf{a}x = 1,$$

where $\mathbf{a}$ is only known up to an interval

$$\mathbf{a} \in \mathcal{I} = [a - \rho \; a + \rho],$$

where $a$ and $\rho \geq 0$ are given (assume $a \neq 0$, $\rho < |a|$)

**Solution set:** possible solutions lie in interval

$$\mathcal{X} = [\frac{1}{a + \rho} \; \frac{1}{a - \rho}]$$

– which solution do we choose?

– in a more complicated situation, can we estimate the solution set?



# Ellipsoidal approach

Ellipsoidal approach will seek an outer approximation of solution set, which here means an interval that contains the solution set $\mathcal{X}$

The central solution will be the center of this interval

In this case, the approximation will be exact, and the center will be given by

$$x = \frac{a}{a^2 - \rho^2}.$$

**note:** except for $\rho = 0$, differs from "nominal inverse" $\frac{1}{a}$!!

The larger the uncertainty, the larger the discrepancy between the nominal inverse and the central solution

## Purpose

Ellipsoidal calculus is a method to handle uncertainty in linear equations

- based on the notion of *ellipsoid of confidence*
- takes into account structure of uncertainty
- recovers classical models as special cases
- extends to dynamical problems



## Ellipsoidal approach and complexity

Solving $Ax = b$ ...

- $A, b$ known: classical linear algebra
- $A, b$ uncertain: leads to NP-hard problems

typical problem in "uncertain linear algebra"

- $[A\ b]$ is an *interval matrix*
- seek intervals of confidence for the solution

computing "exact" bounds is NP-hard problem

*via robust SDP:* polynomial-time algorithms for outer approximation of solution set

## Uncertainty models

We will consider linear equations

$$\mathbf{A}x = \mathbf{b}$$

where $m \times n$ matrix $\mathbf{A}$ and $m$-vector $\mathbf{b}$ are uncertain, *i.e.* all we know is

$$\begin{bmatrix} \mathbf{A} & \mathbf{b} \end{bmatrix} \in \mathcal{U}$$

where $\mathcal{U} \subseteq \mathbf{R}^{m \times (m+1)}$ is **given**

Define **solution set:**

$$\mathcal{X} = \left\{ x \;\middle|\; \mathbf{A}x = \mathbf{b} \text{ for some } \begin{bmatrix} \mathbf{A} & \mathbf{b} \end{bmatrix} \in \mathcal{U} \right\}$$

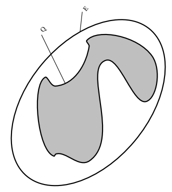
(might be empty!)

usually a quite complicated set!



## Ellipsoidal approach

With uncertain equations, we have a whole set of solutions to choose from . . .



**basic problem:** compute outer ellipsoidal approximation of the solution set

usually this is a hard problem

we'll interpret the "center" as the "central solution" of the equation

## Interpretation

It might be that the central solution does not belong to solution set!

The interpretation of the central solution is as follows:

- Assume that the outer ellipsoidal approximation is shperical
- Then, the central solution is a point that is the closest possible to any solution

Using non-spherical ellipsoids allows to reflect non-spherical shapes of solution set



## LFR model

LFR model: $\mathcal{U}$ is described by

$$\mathcal{U} = \{[\, A \;\; b \,] + L\Delta(I - D\Delta)^{-1} [\, R_A \;\; R_b \,] \;\mid \Delta \in \mathbf{\Delta}\},$$

where $A, L, R, D$ are given matrices of appropriate size, and $\mathbf{\Delta}$ satisfies property $(P)$

that is:

**property** $(P)$**:** for every vector pair $p, q$, the condition

$$p = \Delta q \text{ for some } \Delta \in \mathbf{\Delta}$$

can be characterized as a *linear matrix inequality* on the symmetric (PSD) rank-one matrix

$$X = \begin{bmatrix} q \\ p \end{bmatrix} \begin{bmatrix} q \\ p \end{bmatrix}^T \succeq 0$$

We denote this inequality by

$$\Phi_{\mathbf{\Delta}}(X) \succeq 0$$

where $\Phi_{\mathbf{\Delta}}$ is a linear map

# Examples

**Classical least-squares model:** assume equation is of the form

$$Ax = b + \Delta b, \quad \|\Delta b\|_2 \leq \rho$$

where $\rho \geq 0$, $A, b$ are given

interpretation: we assume that all errors are in the measurement vector $b$

LFR model:

$$\mathcal{U} = \left\{ \begin{bmatrix} A & b \end{bmatrix} + \Delta \begin{bmatrix} 0 & 1 \end{bmatrix} \;\middle|\; \Delta \in \mathbf{\Delta} \right\},$$

where

$$\mathbf{\Delta} = \{\Delta b \mid \|\Delta b\|_2 \leq 1\} \subseteq \mathbf{R}^m$$

satisfies property $(P)$ with

$$\mathbf{\Phi}_{\mathbf{\Delta}} = \rho^2 q^T q - p^2$$

(since $p = \Delta q$ with $\Delta \in \mathbf{R}^m$, $\|\Delta\|_2 \leq \rho$ is equivalent to $p^2 \leq \rho^2 q^T q$)



# Solution set

In the classical least-squares model, the solution set is an ellipsoid

For example: if $A$ is square, invertible, then

$$\mathcal{X} = A^{-1}b + \rho A^{-1} \{u \mid \|u\|_2 \leq 1\}$$

(So, $\mathcal{X}$ is the affine transformation of the unit ball)

A similar result holds with $A$ singular

# Total least squares model

assume additive uncertainty in **both A**, **b**:

$$[\mathbf{A}\ \mathbf{b}] \in \mathcal{U} = \{[A + \Delta A\ b + \Delta b]\ \mid\ \|[\Delta A\ \Delta b]\| \leq \rho\},$$

where $[A\ b] \in \mathbf{R}^{n\times(m+1)}$ and $\rho \geq 0$ are given

Solution set is

$$\mathcal{X} = \left\{x \ \middle|\ \|Ax - b\| \leq \rho\sqrt{\|x\|^2 + 1}\right\},$$

which is

- unbounded if $\rho \geq \sigma_{\min}([A\ b])$
- an ellipsoid if $\rho < \sigma_{\min}([A\ b])$, with center

$$\hat{x} = (A^T A - \rho^2 I)^{-1} A^T b$$



# Parametrization of ellipsoids

model ellipsoids as

$$\mathcal{E}(P, \hat{x}) = \left\{x \ \middle|\ \begin{bmatrix} P & x - \hat{x} \\ (x - \hat{x})^T & 1 \end{bmatrix} \succeq 0\right\},$$

where

- $\hat{x}$ is the center
- $P = P^T \succeq 0$ determines the "shape" (flat ellipsoids allowed)
- **size** of $\mathcal{E}$ related to

$$\phi(P) = \mathbf{Tr}\, P$$

(interpretation: sum of squares of semi-axis lengths)

other measures of size are possible (including volume)

## Flat vs. non-flat ellipsoids

When $P \succ 0$, ellipsoid takes a more classical form:

$$\mathcal{E}(P, \hat{x}) = \left\{ x \;\middle|\; (x - \hat{x})^T P^{-1} (x - \hat{x}) \leq 1 \right\}$$

Our parametrization allows for $P$ to be singular

**Interpretation:** non-zero nullspace of $P$ means ellipsoid is flat (*i.e.*, lives in a linear space)

nullspace of $P$ defines this linear space



## Special case: interval calculus

basic problem of *interval calculus:* given

$$\mathbf{f} \;: \mathbf{R}^l \longrightarrow \mathbf{R}^m,$$

and a set $\mathcal{D} \subseteq \mathbf{R}^l$, find a **upper and lower bounds** for the elements of $\mathbf{f}(\delta)$ when $\delta$ varies over $\mathcal{D}$

(In previous problem, $\mathbf{f}$ is given in implicit form)

assumption: we can describe $\mathbf{f}$ in LFR

$$\mathbf{f}(\delta) = f + L\Delta \left(I - D\Delta\right)^{-1} r,$$

where

$$\Delta = \mathbf{diag}\left(\delta_1 I_{r_1}, \ldots, \delta_l I_{r_l}\right).$$

We assume the corresponding uncertainty set

$$\boldsymbol{\Delta} = \left\{ \Delta = \mathbf{diag}\left(\delta_1 I_{r_1}, \ldots, \delta_l I_{r_l}\right) \;\middle|\; \delta \in \mathcal{D} \right\}$$

satisfies property $(P)$

## Minimizing size of ellipsoid of confidence

The set

$$\mathcal{E} = \{x \mid (x - \hat{x})(x - \hat{x})^T \preceq P\}$$

is an ellipsoid of confidence for the set

$$\{\mathbf{f}(\delta) \mid \delta \in \mathcal{D}\}$$

if and only if

$$P \succeq (\mathbf{f}(\delta) - \hat{x})(\mathbf{f}(\delta) - \hat{x})^T \text{ for every } \delta \in \mathcal{D}$$

or, using Schur complements:

$$\begin{bmatrix} P & (\mathbf{f}(\delta) - \hat{x}) \\ (\mathbf{f}(\delta) - \hat{x})^T & 1 \end{bmatrix} \succeq 0 \text{ for every } \delta \in \mathcal{D} \qquad (*)$$

the problem of minimizing size:

$$\min \ \mathbf{Tr}\, P \text{ subject to } (*)$$

is a robust semidefinite programming problem



## Intervals of confidence

Allowing flat ellipsoids we can estimate the best interval of confidence for a particular component of solution $x$

Assume we want intervals of confidence for $x_1$

$$\min \ e_1^T P e_1 \text{ subject to } (*)$$

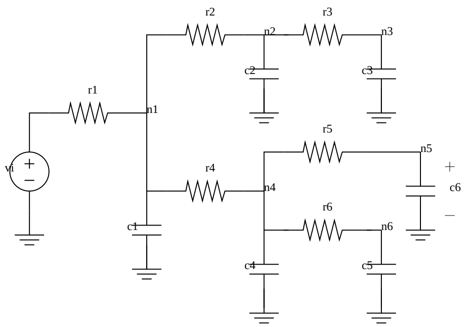
where $e_1$ is first unit vector

Can do the same for other components to get a “box” of confidence

# Example: worst-case timing of RC circuits

recall from earlier lecture: **Elmore delay for RC tree**

## RC tree

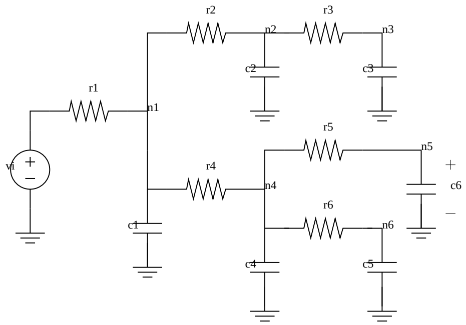



- one input voltage source
- resistors form a tree with root at voltage source
- all capacitors are grounded



**Elmore delay** to node $k$:

$$\sum_i C_i \,(\textstyle\sum R\text{'s upstream from node } k \text{ and node } i)$$




**Example**: Elmore delay to node $3$

$$\begin{aligned} T_3^{\mathrm{elm}} &= C_3(R_1+R_2+R_3)+C_2(R_1+R_2)+C_1R_1 \\ &\quad + C_4R_1+C_5R_1+C_6R_1 \end{aligned}$$

In general case, Elmore delay to node $i$ can be written as

$$T_i^{\mathrm{elm}} = C^T Q_i R,$$

where $C$, $R$ are vectors of capacitances & resistances, and $Q$ is a square matrix (that specifies tree topology, and node being considered)

## Uncertain data in Elmore delay estimation

In practice, we only know the $R_i$'s and $C_i$'s up to intervals:

$$\mathbf{R}_i \in [\underline{R}_i\ \overline{R}_i], \quad \mathbf{C}_i \in [\underline{C}_i\ \overline{C}_i]$$

where $\underline{R}_i \le \overline{R}_i$, $\underline{C}_i \le \overline{C}_i$ are given

**Problem:** estimate upper and lower bounds on Elmore delay to node $i$, for example

$$\max\ \mathbf{C}^T Q_i \mathbf{R} \text{ subject to } \mathbf{C} \in [\underline{C}\ \overline{C}], \quad \mathbf{R} \in [\underline{R}\ \overline{R}]$$

Define the interval data vector:

$$\mathbf{a} = \begin{bmatrix} \mathbf{R} \\ \mathbf{C} \end{bmatrix}$$

and a "weighting matrix" that defines topology to node $i$:

$$W_i = \begin{bmatrix} 0 & Q_i \\ Q_i^T & 0 \end{bmatrix}$$

so that Elmore delay to node $i$ is

$$T_i^{\mathrm{elm}} = \mathbf{a}^T W_i \mathbf{a}$$

where $\mathbf{a}$ is a vector that is only known to belong to a given set $\mathcal{U}$ (a polytope)



## LFR model

Since problem data $\mathbf{a}$ (*i.e.*, capacitances & resistances) is interval data, we can assume the following LFR model structure:

$$\mathbf{a} = a + L\Delta r,$$

where $a$ is the nominal data, $L$ is a matrix, $r$ a vector, and

$$\Delta = \textbf{diag}\left(\frac{\delta R_1}{R_1}, \frac{\delta R_2}{R_2}, \ldots, \frac{\delta C_1}{C_1}, \frac{\delta C_2}{C_2}, \ldots\right)$$

describes relative uncertainty in capacitances & resistances

The uncertainty set (*i.e.*, the set of allowable $\Delta$'s) is

$$\boldsymbol{\Delta} = \left\{\Delta \in \mathbf{R}^{l\times l} \text{ diagonal}, \ \|\Delta\| \le \rho\right\}$$

where $1 > \rho \ge 0$ is the uncertainty level (assumed uniform overs $R_i$'s and $C_i$'s)

($l$ denotes total number of capacitances & resistances)

Above $\boldsymbol{\Delta}$ satisfies property $(P)$ with

$$\Phi_{\boldsymbol{\Delta}}\left(\begin{bmatrix} q \\ p \end{bmatrix}\begin{bmatrix} q \\ p \end{bmatrix}^T\right) = \textbf{diag}\,(q_i^2 - p_i^2)_{i=1}^{l}$$

## Upper bound on Elmore delay

We seek an upper bound on

$$\max \ \mathbf{a}^T W \mathbf{a} \ \ \mathbf{a} \in \mathcal{U}$$

where $W$ is given

apply quadratic matrix inequality robustness lemma (from earlier lecture)

$\lambda$ is an upper bound on the maximum Elmore delay if there exist a matrix $S$ of appropriate size such that

$$\begin{bmatrix} a^T \\ L^T \end{bmatrix} W \begin{bmatrix} a & L \end{bmatrix} + \begin{bmatrix} r^T S r - \lambda & 0 \\ 0 & -S \end{bmatrix} \preceq 0,$$
$$S \succeq 0$$

minimizing $\lambda$ subject to the above constraint is an SDP that yields an upper bound on the original problem

similar results for lower estimate



## General case

**Problem:** ellipsoidal outer approximation for solution set

$$\mathcal{X} = \{x \ | \ \mathbf{A}(\Delta)x = \mathbf{b}(\Delta) \text{ for some } \Delta \in \mathbf{\Delta}\}$$

where

$$\begin{bmatrix} \mathbf{A}(\Delta) & \mathbf{b}(\Delta) \end{bmatrix} = \begin{bmatrix} A & b \end{bmatrix} + L\Delta(I - D\Delta)^{-1} \begin{bmatrix} R_A & R_b \end{bmatrix}$$

**Assumptions:**

- Uncertainty set $\mathbf{\Delta}$ satisfies property $(P)$
- LFR is well-posed over $\mathbf{\Delta}$

## Making the problem quadratic

Using LFR, we can turn the problem into a quadratic one

First, represent solution set using quadratic (matrix) inequalities: $x \in \mathcal{X}$ iff there exist $p, q$ such that

$$(*) \quad \begin{cases} Ax = b + Lp, \quad q = Rx + Dp \\ \Phi_{\mathbf{\Delta}}\left(\begin{bmatrix} q \\ p \end{bmatrix}\begin{bmatrix} q \\ p \end{bmatrix}^T\right) \succeq 0 \end{cases}$$

**note:** used well-posedness assumption!

Assume we look for a non-flat ellipsoid of confidence

Our problem is thus:

$$\min \ \mathbf{Tr}\, P \text{ subject to}$$
$$(x-\hat{x})^T P^{-1}(x-\hat{x}) \leq 1 \text{ for every } x, p \text{ such that } (*)$$

This problem involves quadratic (matrix) inequalities only ...



## Lagrange relaxation

$x$ belongs to ellipsoid of confidence if there exist a matrix $S \succeq 0$ and a scalar $\tau \geq 0$ such that:

for every $x, p$:

$$(x-\hat{x})^T P^{-1}(x-\hat{x}) + \tau\|Ax - b - Lp\|^2$$
$$+\mathbf{Tr}\, S\Phi_{\mathbf{\Delta}}\left(\begin{bmatrix} Rx + Dp \\ p \end{bmatrix}\begin{bmatrix} Rx + Dp \\ p \end{bmatrix}^T\right) \leq 1$$

Above is readily transformed into an LMI in $\hat{x}, P$ and $S, \tau$

**note:** we've used $\|Ax - b - Lp\|^2 \leq 0$ to handle linear constraint (WLOG)

# Example: robust polynomial interpolation

seek a set of confidence for the coefficients of a fourth-order polynomial

$$p(t) = x_1 + tx_2 + \ldots + t^{n-1}x_n$$

such that the polynomial interpolates points $(\mathbf{a}_i, \mathbf{b}_i)$, $i = 1, \ldots, 4$ that are known only within intervals.

- leads to an uncertain system of the form

  $$\mathbf{A}x = \mathbf{b}$$

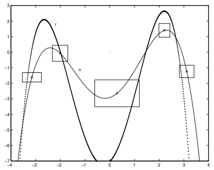
  where $\mathbf{b}$ is an interval vector and $\mathbf{A}$ is a Vandermonde matrix (constructed from interval vector $\mathbf{a}$)

- amenable to ellipsoidal approach
- yields ellipsoid of confidence for the solution



# Minimum sensitivity interpolant

the center of the confidence ellipsoid is at the smallest "distance" from any other interpolant



## Structured condition numbers

let

$$\mathbf{A}(\Delta) = A + L\Delta(I - D\Delta)^{-1}R$$

be a structured matrix, well-posed over $\Delta \in \mathbf{\Delta}_\rho$, where

$$\mathbf{\Delta}_\rho := \{\Delta \in \mathbf{\Delta} \mid \|\Delta\| \le \rho\}$$

(here, $\mathbf{\Delta}$ is a linear set)

*structured inversion error* at $X$:

$$\lambda(\mathbf{A}, \mathbf{\Delta}, \rho, X) = \frac{1}{\rho} \max \left\{ \|\mathbf{A}(\Delta)^{-1} - X\| \ : \ \Delta \in \mathbf{\Delta} \right\}.$$

(structured, absolute) condition number obtains with $\rho \to 0$, $X = A^{-1}$

results:

- upper bounds can be computed, or minimized over $X$, via SDP
- when $\mathbf{\Delta}$ is unit-ball of $\mathbf{R}^{n_p \times n_q}$, obtain exact value of structured condition number: $\|A^{-1}L\| \cdot \|RA^{-1}\|$ (thus, recover absolute condition number $\|A^{-1}\|^2$ with additive, "classical" perturbation structure)



## What is the condition number of a Vandermonde matrix?

- from classical perturbation theory (additive, unstructured perturbation): usually large
- from specialized algorithms (*e.g.*, displacement rank): small
- from structured condition number analysis: explain the existence of well-conditioned algorithms
  - INV: inversion error at $X = A^{-1}$
  - OPT: optimized (over $X$) inversion error
  - USR, SR: unstructured, structured invertibility radius