Lecture 25
Uncertainty models

- Data uncertainty sets
- Linear-fractional representations
- Describings bounds
- Well-posedness
- Summary

Data uncertainty sets

We consider a problem with uncertain data

WLOG, we assume the "data" of the problem is collected in an uncertain matrix $\mathbf{A}$, only known to belong to some *data uncertainty set* $\mathcal{U}$

We assume that $\mathcal{U}$ can be parametrized as follows:

$$\mathcal{U} = \{\mathbf{A}(\Delta) \mid \Delta \in \mathbf{\Delta}\}$$

where $\mathbf{A}(\cdot)$ is a known, matrix-valued function of an "uncertainty matrix" Δ, only known to belong to some set $\mathbf{\Delta}$

- the function $\mathbf{A}(\cdot)$ describes the *structure* of the uncertainty

  example: affine $\mathbf{A}(\cdot)$
- the set $\mathbf{\Delta}$ describes the *bounds* we can put on the uncertainty

  example: $\|\Delta\| \leq 1$

(we'll see later why it is convenient to consider matrix uncertainty rather than vectors)

Example

One very simple model of uncertainty is called *additive* uncertainty: (assume $\mathbf{A}$ lives in $\mathbf{R}^{m\times n}$)

$$\mathcal{U} = \left\{ \mathbf{A}(\Delta) = A + \Delta \;\middle|\; \Delta \in \mathbf{R}^{m\times n}, \;\; \|\Delta\| \le \rho \right\},$$

where A is a "nominal" matrix, and $\rho \ge 0$ a given number

(ρ measures uncertainty size)

variation: assume $\mathbf{A}$ is partitioned in four blocks, only $(1,2)$ block is subject to additive perturbation

$$\mathbf{A}(\Delta) = \begin{bmatrix} A_{11} & A_{12} + \Delta \\ A_{21} & A_{22} \end{bmatrix}$$

then

$$\mathcal{U} = \left\{ \mathbf{A}(\Delta) = A + L\Delta R \;\middle|\; \Delta \in \mathbf{R}^{m\times n}, \;\; \|\Delta\| \le \rho \right\},$$

where

$$L = \begin{bmatrix} I \\ 0 \end{bmatrix}, \quad R = \begin{bmatrix} I & 0 \end{bmatrix}.$$

above are examples with **one** (matrix) variables

Assumptions

Our uncertainty model is

$$\mathcal{U} = \{ \mathbf{A}(\Delta) \;\mid \Delta \in \mathbf{\Delta} \}$$

Assumption 1: We will assume that the structure (*i.e.*, function $\mathbf{A}(\cdot)$) is algebraic

We will introduce a "compact", matrix-based way to represent algebraic, matrix-valued functions:

linear-fractional representations

Assumption 2: we will make some assumptions about the perturbation bounds, *i.e.* the set $\mathbf{\Delta}$

(see later)

later, we use these assumptions to approximate the robust optimization problems

Linear-fractional representations (LFRs)

let $\mathbf{A}$ be a matrix-valued algebraic function of $\delta \in \mathbf{R}^l$, well-defined for $\delta = 0$.

LFR lemma: there exist matrices A, L, R, D and integers $r_1, \ldots, r_l$ such that, for every δ such that $\mathbf{A}(\delta)$ is well-defined,

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

where

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

(I_k means identity matrix of order k)

- A is the "nominal" matrix (value of $\mathbf{A}$ for $\Delta = 0$)
- L is the "left" matrix
- R is the "right" matrix
- D is the "denominator" matrix

above called a *linear-fractional representation* (LFR) of $\mathbf{A}(\cdot)$

Example: affine dependence

Assume data depends **affinely** on parameters $\delta_1, \ldots, \delta_l$, *i.e.*:

$$\mathbf{A}(\delta_1, \ldots, \delta_l) = A + \sum_{i=1}^{l} \delta_i A_i$$

Introduce matrices L_i, R_i such that

$$A_i = L_i R_i \text{ and } \textbf{Rank} L_i = \textbf{Rank} R_i = \textbf{Rank} A_i$$

then

$$\mathbf{A}(\delta_1, \ldots, \delta_l) = A + L\Delta R$$

with

$$L = \begin{bmatrix} L_1 & \ldots & L_l \end{bmatrix}, \quad R = \begin{bmatrix} R_1 \\ \vdots \\ R_l \end{bmatrix},$$

and

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

Example

Assume

$$\mathbf{A}(\delta_1, \delta_2) = 1 + 2\delta_1\delta_2$$

then, let

$$\left[\begin{array}{c|c} A & L \\ \hline R & D \end{array}\right] = \left[\begin{array}{c|cc} 1 & 2 & 0 \\ \hline 0 & 0 & 1 \\ 1 & 0 & 0 \end{array}\right]$$

(there are many choices possible)

gives idea of proof: we know to form an LFR of a product of two variables ...

can extend this to several variables, and products/sums/inverses, etc, of LFRs?

(see later)

LFRs of functions with matrix variables

In some cases, the uncertain variables are not **scalars** $\delta_1, \ldots, \delta_l$, but more generally **matrices** $\Delta_1, \ldots, \Delta_l$

A similar result holds: we can always represent $\mathbf{A}(\Delta_1, \ldots, \Delta_l)$ in the LFR

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

where

$$\Delta = \textbf{diag}\left(\underbrace{\Delta_1, \ldots, \Delta_1}_{r_1 \text{ times}}, \ldots, \underbrace{\Delta_l, \ldots, \Delta_l}_{r_l \text{ times}}\right)$$

(previous scalar case is obvious special case)

A feedback view

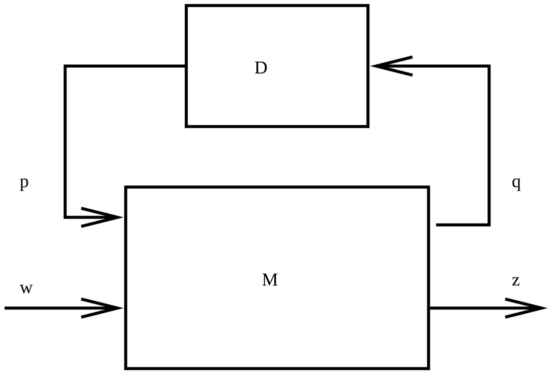
Consider input-output relationship:

$$z = \mathbf{A}(\delta)w = \left(A + L\Delta(I - D\Delta)^{-1}R\right)w$$

if LFR is well-posed, then it can be represented in LFR form:

$$\begin{bmatrix} z \\ q \end{bmatrix} = \begin{bmatrix} A & L \\ R & D \end{bmatrix} \begin{bmatrix} w \\ p \end{bmatrix}, \quad p = \Delta q$$

feedback picture:



morale: feedback systems are "universal"; LFR puts all known (linear) part on one side, and all uncertainty on the other

Proof of LFR lemma

constructing an LFR can be done via LFR algebra:

If we know the LFR of two matrix-valued functions, we can build the LFR of the sum, product, etc

example (product of two LFRs): if

$$\mathbf{A}_i(\Delta_i) = A_i + L_i\Delta_i(I - D_i\Delta_i)^{-1}R_i, \quad i = 1, 2$$

then the product $\mathbf{A}_1(\Delta_1)\mathbf{A}_2(\Delta_2)$ has LFR

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

where

$$\left[\begin{array}{c|c} A & L \\ \hline R & D \end{array}\right] = \left[\begin{array}{c|c} A_1A_2 & \begin{bmatrix} L_1 & A_1L_2 \end{bmatrix} \\ \hline \begin{bmatrix} R_1A_2 \\ R_2 \end{bmatrix} & \begin{bmatrix} D_1 & R_1L_2 \\ 0 & D_2 \end{bmatrix} \end{array}\right],$$

$$\Delta = \textbf{diag}\,(\Delta_1, \Delta_2)$$

A special case

in the case of **one** variable, the LFR is well-known in control theory as the *state-space representation* of transfer functions

Fact: (Kalman's "state-space" representation theory) every algebraic function of one variable s that is well-defined in $s = \infty$ can be written

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

where A, L, R, D are given matrices, and $\Delta = 1/s \cdot I$

Uniqueness, invariance

for a given algebraic functions, there are **many** different choices of LFRs

Given one choice, and an invertible matrix T, we have (whenever well-posed)

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

where

$$\left[\begin{array}{c|c} \tilde{A} & \tilde{L} \\ \hline \tilde{R} & \tilde{D} \end{array}\right] = \left[\begin{array}{c|c} A & LT \\ \hline T^{-1}R & T^{-1}DT \end{array}\right]$$

LFR of symmetric matrices

Consider a function $\mathbf{F}(\delta)$ taking values in the space of symmetric matrices

If $\mathbf{F}(\cdot)$ is algebraic in δ, then we can represent the upper-triangular matrix extracted from $\mathbf{F}$ using an LFR

So, $\mathbf{F}(\cdot)$ admits LFR

$$\begin{aligned}\mathbf{F}(\delta) \;=\; & F + L\Delta(I - D\Delta)^{-1}R + \\ & \left(L\Delta(I - D\Delta)^{-1}R\right)^T\end{aligned}$$

for appropriate F, L, R, D and

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

LFR of symmetric affine matrix function

Assume now $\mathbf{F}$ also depends on another variable x

Assume $\mathbf{F}(x, \delta)$ is **affine** in x

For every y, we can write

$$\begin{aligned}\mathbf{F}(x, \delta) \;=\; & F(x) + L(x)\Delta(I - D\Delta)^{-1}R + \\ & \left(L(x)\Delta(I - D\Delta)^{-1}R\right)^T\end{aligned}$$

for appropriate **affine** matrix-valued functions $F(x)$, $L(x)$ and matrices R, D, and

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

Describing bounds

Uncertainty model is

$$\mathcal{U} = \{\mathbf{A}(\Delta) \mid \Delta \in \mathbf{\Delta}\}$$

LFRs are about describing *structure* of the algebraic, matrix-valued map $\mathbf{A}(\cdot)$

We now need to describe which sets $\mathbf{\Delta}$ we consider

We've seen that LFR lemma gives rise to sets $\mathbf{\Delta}$ comprising matrices of the form

$$\Delta = \textbf{diag}\left(\underbrace{\Delta_1, \ldots, \Delta_1}_{r_1 \text{ times}}, \ldots, \underbrace{\Delta_l, \ldots, \Delta_l}_{r_l \text{ times}}\right)$$

Example of bound: $\|\Delta\| \leq 1$

What kind of bounds would we like to handle?

- componentwise bounds on scalars: $|\delta_i| \leq 1$, $i = 1, \ldots, l$
- Euclidean bounds on vectors: $\|\delta\|_2 \leq 1$
- Norm bounds on matrices: $\|\Delta_i\| \leq 1$
- sign-bounds: $\delta_i > 0$, $i = 1, \ldots, l$
- other (??)

...and a mix of the above

Assumption on uncertainty set $\mathbf{\Delta}$

main assumption on $\mathbf{\Delta}$ 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

In this characterization, the uncertainty Δ is not explicit anymore

Example: largest singular value bounds

unstructured case: if

$$\mathbf{\Delta} = \{\Delta \mid \|\Delta\| \leq 1\},$$

then $p = \Delta q$ for some $\Delta \in \mathbf{\Delta}$ is equivalent to a **scalar** inequality

$$p^T p \leq q^T q,$$

(indeed an LMI in X)

Maximum-norm bounds

This is a “structured” version of previous case

assume

$$\mathbf{\Delta} = \left\{ \mathbf{diag}\left(\underbrace{\Delta_1, \ldots, \Delta_1}_{r_1 \text{ times}}, \ldots, \underbrace{\Delta_l, \ldots, \Delta_l}_{r_l \text{ times}} \right) \;\middle|\; \|\Delta_i\| \leq 1 \right\},$$

then $p = \Delta q$ for some $\Delta \in \mathbf{\Delta}$ is equivalent to

$$\mathbf{diag}\,(q_i q_i^T - p_i p_i^T)_{i=1}^l \succeq 0$$

(here, q_i, p_i stands for the i-th block-vector of length r_i in q, p)

(again an LMI in X)

Example: Euclidean-norm bounds

assume

$$\mathbf{\Delta} = \left\{ \mathbf{diag}\,(\delta_i I_{r_i})_{i=1}^l \;\middle|\; \delta \in \mathbf{R}^l,\ \|\delta\|_2 \leq 1 \right\},$$

then $p = \Delta q$ for some $\Delta \in \mathbf{\Delta}$ is equivalent to

$$\mathbf{diag}\,(q_i^T q_i)_{i=1}^l - pp^T \succeq 0$$

(here, q_i stands for a block-vector of q of length r_i)

proof: hinges on the fact that

$$p = M\delta \text{ for some } \delta \in \mathbf{R}^l,\quad \|\delta\|_2 \leq 1$$

iff

$$pp^T \preceq MM^T$$

(try to prove this!)

Example: sign bounds

assume

$$\boldsymbol{\Delta} = \left\{ \mathbf{diag}\,(\delta_i I_{r_i})_{i=1}^{l} \;\middle|\; \delta \in \mathbf{R}^l,\ \delta_i \geq 0,\quad i = 1, \ldots, l \right\},$$

then $p = \Delta q$ for some $\Delta \in \boldsymbol{\Delta}$ is equivalent to

$$\mathbf{diag}\,(q_i p_i^T + p_i q_i^T)_{i=1}^{l} \succeq 0$$

(again, q_i stands for a block-vector of q of length r_i)

proof: hinges on the fact that

$$p = \delta q \text{ for some } \delta \in \mathbf{R},\quad \delta \geq 0$$

if and only if

$$pq^T + qp^T \succeq 0$$

Cases covered by framework

- data is a algebraic function of bounded parametric uncertainty
- bounds on parameters can be
  - Euclidean
  - maximum-norm
  - sector bounded (*e.g.*, sign)
  - mix of above
- matrix-norm bounds

special case: "unstructured perturbation", means Δ is a matrix ball in $\mathbf{R}^{n_p \times n_q}$ (for standard operator norm)

in this case, our results will be exact

Well-posedness

"linear-fractional representation" (LFR) is *well-posed* over a set $\mathbf{\Delta} \in \mathbf{R}^{N \times N}$ $(N = r_1 + \ldots + r_l)$ if

$$\det(I - D\Delta) \neq 0 \text{ for every } \Delta \in \mathbf{\Delta}$$

For now we implicitely assume well-posedness (more later)

fact: if map $\mathbf{A}$ is polynomial, then we can always construct D to be strictly upper-triangular

proof: see later

then, as expected, LFR is always well-posed

proof: since Δ is a diagonal matrix, $D\Delta$ is strictly upper-triangular, and

$$\det(I - D\Delta) = 1 \text{ for every } \Delta \in \mathbf{\Delta}$$

Summary: what is an uncertainty model?

In this course, we assume the data uncertainty model

$$\mathcal{U} = \{\mathbf{A}(\Delta) \mid \Delta \in \mathbf{\Delta}\}$$

is such that:

- the matrix-valued function $\mathbf{A}(\cdot)$ is algebraic, and admits an LFR

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

  that is valid, and well-posed everywhere on $\mathbf{\Delta}$
- The uncertainty set satisfies the property (P)

A rank-constrained view

If we look at the input-output map

$$z = \mathbf{A}(\Delta)w, \quad \Delta\ in \boldsymbol{\Delta}$$

then the above assumption show we can describe this "uncertain system" as follows:

$$\begin{bmatrix} z \\ q \end{bmatrix} = \begin{bmatrix} A & L \\ R & D \end{bmatrix} \begin{bmatrix} w \\ p \end{bmatrix}, \quad \Phi_{\boldsymbol{\Delta}}(X) \succeq 0, \quad X = \begin{bmatrix} q \\ p \end{bmatrix} \begin{bmatrix} q \\ p \end{bmatrix}^T \succeq 0$$

the above can be interpreted as a set of linear matrix equalities and inequalities on the rank-one, symmetric, PSD matrix:

$$\tilde{X} = \begin{bmatrix} z \\ w \\ q \\ p \\ 1 \end{bmatrix} \begin{bmatrix} z \\ w \\ q \\ p \\ 1 \end{bmatrix}^T$$

Abstract point of view

Following previous reasoning, a data uncertainty set can be modeled as

$$\tilde{A}\tilde{x} = \tilde{b}, \quad \tilde{x} \succeq 0, \quad \textbf{Rank}\tilde{x} = 1$$

where $\tilde{x}$ represents a matrix, and $\tilde{A}$ a linear map

above constraints are $\mathbf{not}$ convex, due to rank constraints

The robust optimization approximations we will study hinge on relaxing (*i.e.*, dropping) the rank constraint