Lecture 28
Linear equations with structured data

- uncertainty: structure vs. bounds
- example: least squares
- exploiting structure
- example: total least squares
- structured total least squares
- ellipsoidal approach

Uncertainty: structure vs. bounds

Last lecture: we considered a linear equation

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

is "uncertain", meaning we know a **set of confidence** $\mathcal{U}$ for matrix $[\mathbf{A}\ \mathbf{b}]$: in particular,

- We know bounds on $\mathbf{A}$, $\mathbf{b}$
- We know the structure of the uncertainty
  (*e.g.*, some elements are uncertain, others are not, etc)

What if we don't know the bounds?

That is, we have a *partial information* on set $\mathcal{U}$

Example: least-squares

Suppose we measure the data A, b in a linear equation

$$Ax = b,$$

where $A \in \mathbf{R}^{m \times n}$, $b \in \mathbf{R}^m$, and $m >> n$ (overdetermined system)

In general, the above equations have **no** solutions (above equation is "inconsistent")

Why? The above equation is actually an **idealized model** of the reality ... the correct equation is different!

Let us **assume** that the **true** equation is actually linear, and of the form

$$Ax = b + \Delta b \tag{1}$$

for some (yet unknown) Δb

Least squares solution

The **least-squares** method seeks the smallest norm of errors $\|\Delta b\|_2$ that makes equation (1) consistent:

$$\min \ \|\Delta b\|_2 \text{ subject to } Ax = b + \Delta b$$

Underlying assumptions:

- true model is linear
- all errors in b (no error in A)
- errors are "uniformly distributed" (all directions around b are permitted)

Here, the "distance to inconsistency" is measured with $\|\Delta b\|_2$

Exploiting structure

In practice, we have some *sructure information* about the way A, b can be perturbed

let us assume that the true model is of the form

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

where $[\mathbf{A}\ \mathbf{b}]$ belongs to a given set $\mathcal{S}$

we are given a "nominal" matrix $[A\ b]$ that belongs to $\mathcal{S}$; the "nominal equation"

$$Ax = b$$

has no solutions

We seek a model $[\mathbf{A}\ \mathbf{b}]$ in $\mathcal{S}$ that is

- "closest" to $[A\ b]$;
- consistent, *i.e.* $\mathbf{A}x = \mathbf{b}$ has a solution

We need a definition of what "closest" means

Example: affine perturbations

Assume $\mathcal{S}$ is an affine space: since it contains the "nominal matrix" $[A\ b]$ we can write

$$\mathcal{S} = \left\{= [\mathbf{A}\ \mathbf{b}] = [A\ b] + \sum_{i=1}^{l} \mathbf{c}_i [A_i\ b_i] \;\middle|\; \mathbf{c} \in \mathbf{R}^l \right\},$$

where $[A_i\ b_i]$'s are given matrices, and $\mathbf{c}$ is a free parameter

We'll measure distance between $[A\ b]$ and a matrix in $\mathcal{S}$ by the Euclidean norm of parameter $\mathbf{c}$

Total least-squares

Total Least-Squares (TLS) is a variation on ordinary LS

In some cases, assumption that A has no errors is not valid

Total least-squares:

$$\min \ \|[\Delta A \ \Delta b]\|_2 \text{ subject to } (A + \Delta A)x = b + \Delta b$$

Interpretation:

- true model is linear
- errors in **both** A, b (on b only in LS)
- errors are "uniformly distributed" (all directions around data matrix $[A \ b]$ are equally permitted)

SDP solution to total least-squares

Ordinary least-squares is obviously an SDP ...

How about total least-squares?

Ellipsoidal approach:

- Assume $\|[\Delta A \ \Delta b]\|_2 \le \rho$, find ellipsoid of confidence for solutions of
$$(A + \Delta A)x = b + \Delta b$$
- Minimize ρ until ellipsoid of confidence is reduced to a singleton

Since the ellipsoidal approach yields exact results via SDP (unstructured uncertainty), the above approach solves the problem

Direct solution

we can recover the solution directly ...

For given $\rho \geq 0$, the set of solutions to

$$(A + \Delta A)x = b + \Delta b$$

is the quadratic set:

$$\left\{ x \;\middle|\; \begin{bmatrix} x \\ 1 \end{bmatrix}^T \left(\begin{bmatrix} A^T \\ b^T \end{bmatrix} \begin{bmatrix} A & b \end{bmatrix} - \rho^2 I_{m+1} \right) \begin{bmatrix} x \\ 1 \end{bmatrix} \leq 0 \right\}.$$

- if $\rho < \sigma_{\min}([A\ b])$, then the ellipsoid of confidence is empty.
- if $\sigma_{\min}([A\ b]) \leq \rho < \sigma_{\min}(A)$, then the center of the ellipsoid is

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

TLS: summary

Except in the degenerate case when

$$\sigma_{\min}([A\ b]) = \sigma_{\min}(A),$$

the TLS solution is given by

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

where

$$\rho = \sigma_{\min}([A\ b])$$

The ellipsoidal approach yields the same answer

Structured total least squares

The structured total least-squares problem is as follows

- given a "nominal model" $[A\ b]$
- given a structure set $\mathcal{S}$, subset of $\mathbf{R}^{m\times(n+1)}$;
- given a distance function defined on $\mathcal{S}$, $\mathbf{dist}_{\mathcal{S}}$;

solve

$$\min_{[\mathbf{A}\ \mathbf{b}]\in\mathcal{S}} \mathbf{dist}_{\mathcal{S}}\left([\mathbf{A}\ \mathbf{b}], [A\ b]\right) \text{ subject to } \mathbf{A}x = \mathbf{b}$$

Examples

- LS:

$$\mathcal{S} = \{[\mathbf{A}\ \mathbf{b}] = [A\ b] + [0\ \Delta b]\},$$
$$\mathbf{dist}\left([\mathbf{A}_1\ \mathbf{b}_1], [\mathbf{A}_2\ \mathbf{b}_2]\right) = \|\mathbf{b}_1 - \mathbf{b}_2\|_2$$

- TLS:

$$\mathcal{S} = \{[\mathbf{A}\ \mathbf{b}] = [A\ b] + [\Delta A\ \Delta b]\}$$
$$\mathbf{dist}\left([\mathbf{A}_1\ \mathbf{b}_1], [\mathbf{A}_2\ \mathbf{b}_2]\right) = \|[\mathbf{A}_1\ \mathbf{b}_1] - [\mathbf{A}_2\ \mathbf{b}_2\|_2$$

- Affine uncertainty:

$$\mathcal{S} = \left\{= [\mathbf{A}\ \mathbf{b}] = [A\ b] + \sum_{i=1}^{l} \mathbf{c}_i [A_i\ b_i]\ \middle|\ \mathbf{c} \in \mathbf{R}^l\right\},$$
$$\mathbf{dist}\left([\mathbf{A}_1\ \mathbf{b}_1], [\mathbf{A}_2\ \mathbf{b}_2]\right) = \|\mathbf{c}_1 - \mathbf{c}_2\|_2$$

Ellipsoidal approach

Similar approach

- For given $\rho \geq 0$, let

  $$\mathcal{S}_\rho := \mathcal{S} \cup \{[\mathbf{A}\ \mathbf{b}] \mid \mathbf{dist}_{\mathcal{S}}([\mathbf{A}\ \mathbf{b}], [A\ b]) \leq \rho\}$$

  Find ellipsoid of confidence for solutions of the equations

  $$\mathbf{A}x = \mathbf{b} \quad [\mathbf{A}\ \mathbf{b}] \in \mathcal{S}_\rho$$

- Minimize ρ until ellipsoid of confidence is reduced to a singleton

gives an upper bound on optimal objective

returned solution not necessarily consistent!