Counterexamples in Gramian based model reduction

Carl Christian Kjelgaard Mikkelsen

HPC2N-NLA
Umeå
January 18th
2010
Overview of talk

• The main problem
• Model reduction
• Counterexamples
• Open questions
The main problem

**Theorem** If $P \geq 0$ and $Q \geq 0$, then

$$\exists V, W : PV = W\Sigma, \quad QW = V\Sigma, \quad V^T W = I$$

where

$$\Sigma = \text{diag}\{\sigma_1, \sigma_2, \ldots, \sigma_n\}$$

and

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n \geq 0.$$  

**Problem** Compute the first $k$ columns of $V$ and $W$. 
An equivalent problem

If

\[ PV = W\Sigma, \quad QW = V\Sigma \]

then

\[ (QP)V = Q(PV) = Q(W\Sigma) = V\Sigma^2, \]
\[ (PQ)W = P(QW) = P(V\Sigma) = W\Sigma^2. \]
Motivation: Model reduction of LTI systems

A Linear Time Invariant (LTI) system

\[ \Sigma = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \]

is a dynamical system

\[ \dot{x} = Ax + Bu \]

\[ y = Cx + Du \]

**Problem** Determine the output \( y \) as a function of the input \( u \).
Motivation: Model reduction of LTI systems

Let $P$ and $Q$ be given by

\[
AP + PA^T + BB^T = 0, \\
A^T Q + QA + C^T C = 0.
\]

Compute the rank $k$ dominant eigenspace for $PQ$ and $QP$, i.e.

\[
PV_k = W_k \Sigma_k, \quad QW_k = V_k \Sigma_k, \quad V_k^T W_k = I_k
\]

and form

\[
\Sigma_k = \begin{pmatrix} V_k^T AW_k & V_k^T B \\ CW_k & D \end{pmatrix}.
\]

Then

\[
\| \Sigma - \Sigma_k \|_{\mathcal{H}_\infty} \leq 2(\sigma_{k+1} + \sigma_{k+2} + \cdots + \sigma_n).
\]
Counterexamples

There exists a system

$$\Sigma = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

such that

$$P > 0, \quad \|P\|_2 \geq 1$$
$$Q > 0, \quad \|Q\|_2 \geq 1$$

but

$$\forall x \in \mathbb{R}^n : \text{fl}(P\text{fl}(Qx)) = 0.$$
Counterexamples, . . .

Specifically,

\[
A = - \begin{bmatrix}
  d & & \\
  -1 & \ddots & \\
  & \ddots & -1 \\
  & & -1 & d
\end{bmatrix}, \quad B = (\sqrt{2d}, 0, \ldots, 0)^T
\]

and

\[C = (0, \ldots, 0, \sqrt{2d})\]

will do the trick where \(d > 1\).
Counterexamples, . . .

For $n = 4$ we have

$$P = \begin{bmatrix}
1 & \nu \\
\nu & \nu^2 \\
\nu^2 & \nu^3 
\end{bmatrix} \begin{bmatrix}
1 & 1 & 1 & 1 \\
1 & 2 & 3 & 4 \\
1 & 3 & 6 & 10 \\
1 & 4 & 10 & 20 
\end{bmatrix} \begin{bmatrix}
1 \\
\nu \\
\nu^2 \\
\nu^3
\end{bmatrix}$$

where $\nu = \frac{1}{2d} < \frac{1}{2}$.

In general,

$$p_{ij} = \binom{i + j - 2}{j - 1} \nu^{i-j-2}.$$
Counterexamples, ...

The sparsity pattern for \( \text{fl}(P) \) for \( n = 500 \), and \( d = 2, 4, 8, 16 \).

The matrix \( Q \) satisfies

\[
Q = JPJ
\]

where \( J \) is the anti-diagonal matrix.
This is a general problem!

**Observation:**

Given a negative definite matrix $A$ and a vector $B$, such that

$$K(A, B) = \mathbb{R}^n$$

then there exists $C$, such that

$$K(A^T, C) = \mathbb{R}^n$$

and

$$\text{fl}(P)\text{fl}(Q) = 0.$$  

**Proof:** Incomplete.

Hinges on the convergence theory for a particular solver (KPIK).
KPIK, briefly

Approximate $P$ with

$$P_j = V_j Y_j V_j^T$$

where

$$\text{Ran} V_j = \text{Ran} \left[ A^{-j} B \ A^{-j+1} B \ \ldots \ \ B \ AB \ \ldots \ A^{j-1} B \right]$$

and $Y_j$ solves

$$H_j Y_j + Y_j H_j^T + B_j B_j^T = 0$$

where

$$H_j = V_j^T AV_j, \quad B_j = V_j^T B.$$
If KPIK runs to completion, then

$$AV = VH$$

where both $H$ and $H^{-1}$ are upper block Hessenberg!

The choice of

$$A = H, \quad B = e_1, \quad C = e_n$$

ensures

$$P_k \perp Q_k$$

for all realistic values of $k$.

A proof of the observed (rapid) convergence would show

$$\text{fl}(P)\text{fl}(Q) = 0.$$
Counterexamples, . . .

**Theorem** Given $P \geq 0$ and $Q \geq 0$, such that $PQ$ has distinct eigenvalues, then there exists a stable system

$$\Sigma = \begin{bmatrix} A & B \\ C & D \end{bmatrix}$$

where $B$ and $C^T$ are vectors, such that

$$AP + PA^T + BB^T = 0,$$

$$A^TQ + QA + C^CC = 0.$$ 

**Proof** Simple and constructive, but I find it hard to control the conditioning of $A$. 


Counterexamples, . . .

In particular, systems for which

\[ P \approx \text{diag}\{1, 10^{-2}I_{100}, 10^{-9}I_N, 10^{14}I_{100}, 1\} \]
\[ Q \approx \text{diag}\{1, 10^{14}I_{100}, 10^{-9}I_N, 10^{-2}I_{100}, 1\} \]
\[ PQ \approx \text{diag}\{1, 10^{12}I_{100}, 10^{-18}I_N, 10^{12}I_{100}, 1\} \]

do exists! The problem is

\[ P_{202} \approx \text{diag}\{1, 0_{100}, 0_N, 10^{14}I_{100}, 1\} \]
\[ Q_{202} \approx \text{diag}\{1, 10^{14}I_{100}, 0_N, 0_{100}, 1\} \]

such that

\[ P_{202}Q_{202} \approx \text{diag}\{1, I_{100}, 0_N, 0_{100}, 1\} \]

while

\[ (PQ)_{202} \approx \text{diag}\{1, 10^{12}I_{100}, 0_N, 10^{12}I_{100}, 1\} \]

Information is lost!
Summary

• Grammian based model reduction hinges on approximating

\[ P \quad \text{and} \quad Q \]

with low rank matrices.

• This is unavoidable, but risky as we can lose all relevant information about

\[ PQ \quad \text{and} \quad QP \]

• Even a partial loss is hard to detect.
Open questions

• Convergence of KPIK; requires good understanding of complex analysis.

• Approximation of matrix exponential in KPIKs extended Krylov subspaces.

• Construction of a real problems which will yield specific Gramians.

• When can (approximate) balanced truncation be trusted?
Questions?
Thank you for your attention.