# Sylvester Equations, Model Reduction and Error

**P. M. Van Dooren and A. Vandendorpe**

**Department of Mathematical Engineering**

**Catholic University of Louvain**

**K. A. Gallivan**

**School of Computational Science**

**Florida State University**

**MTNS 2006**

# Outline

- The Problem
- Model reduction
  - Tangential interpolation
  - Generalized Krylov subspaces
- Sylvester equations
  - Equivalent linear systems
  - Algorithms and residuals
- Sylvester residual
  - One-sided error
  - General form of error
- Numerical examples

# The Problem

- Model Reduction
  - Rational Interpolation – SISO, MIMO
  - Tangential Interpolation – MIMO
- Algorithms
  - Generalized Krylov Spaces and Projections
  - Approximations – approximate solution of linear systems
- What is the effect of approximations on solutions?
  - after the fact analysis of projector/reduced order model
  - algorithm design to control error
  - convergence analysis of series of reduced order models

# References

**Tangential interpolation, Sylvester Equations, Model Reduction:**


- D.C. Sorensen and A.C. Antoulas. *The Sylvester equation and approximate balanced truncation*. Linear Algebra and Its Applications, 351-352, pp. 671–700, 2002.
- A. Antoulas and A. Mayo – Tangential interpolation and identification, 2006
- A. Vandendorpe. Model Reduction of Linear Systems, an Interpolation Point of View. PhD thesis, Université catholique de Louvain, Louvain-la-Neuve, Belgium, 2004. Freely available at $http://www.bib.ucl.ac.be$
  - A. Vandendorpe, P. Van Dooren and K. G. SIMAX 2004.
  - A. Vandendorpe, and P. Van Dooren and K. G. JCAM 2004.
  - A. Vandendorpe, and P. Van Dooren. Proceedings of CDC 2005.

# References

**Approximate System Solves:**


- Inexact Krylov methods for linear system solving – many papers, e.g., last Householder meeting
- C. Beattie and S. Gugercin – Inexact Krylov model reduction and backward error , SIAM Annual Meeting 2006
- E.J. Grimme. Krylov projection methods for model reduction. PhD Thesis, Department of Electrical Engineering, University of Illinois at Urbana-Champaign, 1997.
  - E. Grimme, P. Van Dooren, V. DeClippel and K.G. Proceedings CDC 1997.

## MIMO Model Reduction

Approximate a given $p \times m$ rational matrix with state dimension $N$

$$T(s) := C(sE - A)^{-1}B$$

by another $p \times m$ rational matrix with state dimension $n \ll N$

$$\hat{T}(s) := \hat{C}(s\hat{E} - \hat{A})^{-1}\hat{B}$$

.

A projection formulation yields

$$\{\hat{E}, \hat{A}, \hat{B}, \hat{C}\} = \{Z^*EV, Z^*AV, Z^*B, CV\},$$

using $N \times n$ projection matrices $V,\ Z$ with $Z^*Z = V^*V = I_n$.

# Generalized Krylov Spaces

Let $x(s)$ be a $m \times 1$ polynomial vector around $\alpha \in \mathbb{C}$

$x(s) := \sum_{i=0}^{n-1} x_i (s-\alpha)^i \leftrightarrow X := \begin{bmatrix} x_0 & x_1 & \dots & x_{n-1} \end{bmatrix}$.

The $N \times n$ generalized Krylov matrix is:

$$K_n\left(A_\alpha, B_\alpha, X\right) := \begin{bmatrix} B_\alpha, A_\alpha B_\alpha & \dots & A_\alpha^{n-1} B_\alpha \end{bmatrix} \begin{bmatrix} x_0 & \dots & x_{n-1} \\ & \ddots & \vdots \\ & & x_0 \end{bmatrix}.$$

where $A_\alpha := (A - \alpha E)^{-1} E$ and $B_\alpha := (A - \alpha E)^{-1} B$.

The associated generalized Krylov subspace is:

$$\mathcal{K}_n\left(A_\alpha, B_\alpha, X\right) := Im\left\{K_k\left(A_\alpha, B_\alpha, X\right)\right\}.$$

# Tangential Interpolation

**Thm** [GVV04] If $\alpha$ is not a pole of $T(s)$ or $\hat{T}(s)$ then

$$\mathcal{K}_n\left(A_\alpha, B_\alpha, X\right) \subset Im(V) \quad \Rightarrow \quad \left(\hat{T}(s) - T(s)\right) x(s) = O(s-\alpha)^n$$

which is a right tangential interpolation condition.

One can also impose left tangential interpolation conditions

$$y^T(s)\left(\hat{T}(s) - T(s)\right) = O(s-\mu)^n$$

and two-sided interpolation conditions

$$y^T(s)\left(\hat{T}(s) - T(s)\right) x(s) = O(s-\nu)^n.$$

# Tangential Interpolation

- Can be generalized to mulitple points with higher order moments
- Simpler than MIMO rational interpolation conditions
- applies to systems with many inputs and outputs (e.g. $B = C = I_N$)
- flexibility of choices of dominant points and directions

# Projection Backward Error

- System : $\Sigma = \{E, A, B, C\}$ (known)
- Exact projections : $V_T$ and $Z_T$ (not known)
- True ROM: $\hat{\Sigma}_T = \{Z_T^* E V_T, Z_T^* A V_T, Z_T^* B, C V_T\}$ (not known)
- Computed projections : $V_C$ and $Z_C$ (known)
- Computed ROM: $\hat{\Sigma}_C = \{Z_C^* E V_C, Z_C^* A V_C, Z_C^* B, C V_C\}$ (known)
- $\hat{\Sigma}_C$ does not interpolate $\Sigma$

Find (nearby?) system that the computed projection interpolates

- $\Sigma_p = \{E_p, A_p, B_p, C_p\}$ (known)
- $\hat{\Sigma}_p = \{Z_C^* E_p V_C, Z_C^* A_p V_C, Z_C^* B_p, C_p V_C\}$ (known)
- Proximity of $\hat{\Sigma}_p$ and $\hat{\Sigma}_C$; and $\Sigma_p$ and $\Sigma$ used to assess error.

# ROM Backward Error

- $\Sigma = \{E, A, B, C\}$ (known)
- $\Sigma_p = \{E_p, A_p, B_p, C_p\}$ (known)
- $\hat{\Sigma}_C = \{Z_C^* E V_C, Z_C^* A V_C, Z_C^* B, C V_C\}$ (known)
- $\hat{\Sigma}_p = \{Z_C^* E_p V_C, Z_C^* A_p V_C, Z_C^* B_p, C_p V_C\}$ (known)
- $\hat{\Sigma}_C$ does not interpolate $\Sigma$
- $\hat{\Sigma}_p$ interpolates $\Sigma_p$

Require additionally that

$$\hat{\Sigma}_p = \hat{\Sigma}_C$$

- Beattie and Gugercin (2006) used this form and give an algorithm for Rational Interpolation for distinct points.
- not always possible – requires contraints on algorithm
- Petrov-Galerkin condition

# Sylvester Equations

$$AK - EKJ_\alpha - BX = 0$$

Partition $K = \begin{bmatrix} k_1 & \dots & k_n \end{bmatrix}$

$$\begin{aligned}
(A - \alpha E)k_1 = Bx_0 \quad &\rightarrow \quad k_1 = B_\alpha x_0 \\
(A - \alpha E)k_2 - Ek_1 = Bx_1 \quad &\rightarrow \quad k_2 = (A - \alpha E)^{-1} E k_1 + B_\alpha x_1 \\
&\rightarrow \quad k_2 = (A - \alpha E)^{-1} E B_\alpha x_0 + B_\alpha x_1 \\
&\vdots \\
(A - \alpha E)k_n - Ek_{n-1} = Bx_{n-1} \quad &\rightarrow \quad k_n = \sum_{j=1}^{n} A_\alpha^{j-1} B_\alpha x_{n-j}.
\end{aligned}$$

- Sylvester equation in terms of $(A, E, B)$ **not** $(A_\alpha, B_\alpha)$
- system parameters and interpolation parameters separated

# Sylvester Equations Normalization

$$
\begin{aligned}
AK - EK - J_\alpha - BX &= 0 \\
A(KT^{-1})(TS) - E(KT^{-1})(TJ_\alpha S) - BXS &= 0 \\
A\bar{K}G - E\bar{K}F - B\bar{X} &= 0
\end{aligned}
$$

- Generalized eigenvalues $det(\lambda G - F)$ are the interpolation points.
- $S = T^{-1}$ is a convenient normalization.
- $K = QR = \bar{K}T$ and $S = R^{-1}$ yields

$$AQ - EQ\bar{J}_\alpha - B\bar{X} = 0$$

- For triangular $T$, $Q$ is essentially unique but many $G$, $F$, $\bar{X}$ triples yield a Sylvester equation it solves.
- Triangular $G$, $F$ relate to recurrence-based algorithms, e.g., inverse iteration, Arnoldi etc.

# Inner Residual and Sylvester Residual

**Algorithm 1** (dangerous form)

1. Compute $k_1$ by solving with tolerance $\delta$

$$(A - \alpha E)k_1 = Bx_0 + m_1, \quad \|m_1\| \leq \delta$$

2. For $i = 2$ to $n$, compute iteratively $k_i$ by solving with tolerance $\delta$

$$(A - \alpha E)k_i = Bx_{i-1} + k_{i-1} + m_i, \quad \|m_i\| \leq \delta$$

3. Factor $K = QR$ to get $\bar{K} = Q$ orthonormal (i.e. $T = R$ and $S = R^{-1}$)

## Inner Residual and Sylvester Residual

This yields

$$\begin{aligned} AK - KJ_\alpha - BX &= M \\ A\bar{K} - \bar{K}\bar{J}_\alpha + B\bar{X} &= MR^{-1} \\ \|M\| &\leq \delta \\ \|MR^{-1}\| &\leq ? \end{aligned}$$

- For this choice of $K = QR$, $\|MR^{-1}\| >> \|M\|$.
- Which do we want to control/monitor? Why?

# Inner Residual and Sylvester Residual

**Algorithm 2** (less dangerous form – merge with Gram-Schmidt)

1. Compute $\bar{k}_1$ by solving with tolerance $\delta$

$$(A-\alpha E)\bar{k}_1 = B\bar{x}_0 + \bar{m}_1, \quad \|\bar{m}_1\| \le \delta, \quad \|\bar{k}_1\| = 1$$

2. For $i = 2$ to $n$, compute iteratively $\bar{k}_i$ by solving with tolerance $\delta$

$$(A-\alpha E)\bar{k}_i = B\bar{x}_{i-1} - \sum_{j=1}^{i-1} \bar{k}_j \bar{J}_{j,i} + \bar{m}_i, \quad \|\bar{m}_i\| \le \delta$$

   where

$$\bar{k}_i \perp \bar{k}_j,\ j = 1, \ldots, i-1, \quad \|\bar{k}_i\| = 1$$

3. Update $\bar{X}$ and $\bar{J}$ with $\bar{J}_{j,i},\ j = 1, \ldots, i-1$

This yields $A\bar{K} - \bar{K}\bar{J} + B\bar{X} = \bar{M}$ with $\|\bar{M}\| \approx \delta$ .

# Appropriate Sylvester Equation for Arnoldi

Compute SISO single-point RI via Arnoldi on $(A_\alpha, B_\alpha) = ((A - \alpha E)^{-1}E, (A - \alpha E)^{-1}b)$ yields

$$AV_mG - EV_mF - be_1^T = 0$$

$$\begin{aligned} G &= H_\alpha J_0 + \beta e_1 e_1^T \\ F &= \alpha G + J_0 \end{aligned}$$

For these $G$ and $F$, $\det(\lambda G - F) = 0$ for $\lambda = \alpha$ with $alg(\alpha) = m$ and $geo(\alpha) = 1$.

# Sylvester Residual and Backward Error

$$\{\hat{E}, \hat{A}, \hat{B}, \hat{C}\} = \{Z_n^* E V_n, Z_n^* A V_n, Z_n^* B, C V_n\}.$$

Consider a one-sided approach, $Z_n$ is not subject to a Sylvester equation and $V_n$ is an orthonormal basis that satisfies

$$A V_n G - E V_n F - B X = R_s.$$

$\Delta_A$ and $\Delta_E$ that yield a $0$ residual solve the Sylvester equation satisfy

$$\Delta_A V_n G - \Delta_E V_n F = -R_s.$$

$V_n^* V_n = I_n \rightarrow \Delta_A = R_A V_n^*$ and $\Delta_E = R_E V_n^*$

$$R_A G - R_E F = -R_s$$

# Sylvester Residual and Backward Error

Forms of error dictate residual of interest

$$\Delta_E = 0, \quad R_E = 0, \quad R_A = -R_s G^{-1}$$

- Algorithm 2 has correct residual control, since the normalization has made $G = I$, i.e.,

$$A\bar{K} - \bar{K}\bar{J} + B\bar{X} = \bar{M}$$

  with $\|\bar{M}\| \approx \delta$ .

- $\hat{A}_p = Z_n^*(A - R_S G^{-1} V_n^*)V_n = \hat{A} - Z_n^* R_S G^{-1}$
- backward ROM error if $Z_n^* R_S = 0$
- consistent with Beattie and Gugercin particular case.

## Sylvester Residual and Backward Error

- $\Delta_E = 0, \;\; R_E = 0, \;\; R_A = RG^{-1}$
- $\Delta_E = \Delta_A, \;\; \bar{R} = R_E, \;\; \bar{R} = R_A, \;\; \bar{R} = R(G - F)^{-1}$
- Weighted Error. $R_E = \gamma_E \bar{R}, \;\; R_A = \gamma_A \bar{R}, \;\; \bar{R} = R(\gamma_A G - \gamma_E F)^{-1}$

# General Two-sided Backward Error

Given orthonormal $V_n$ and $Z_n$ defining the reduced order model

$$\{\hat{E}, \hat{A}, \hat{B}, \hat{C}\} = \{Z_n^* E V_n, Z_n^* A V_n, Z_n^* B, C V_n\}$$

find perturbation $\{\Delta_E, \Delta_A, \Delta_B, \Delta_C\}$ such that

$$\begin{aligned} \Delta_A V_n G_v - \Delta_E V_n F_v - \Delta_B X &= -R_v \\ G_z^* Z_n^* \Delta_A - F_z^* Z_n^* \Delta_E - Y^* \Delta_C &= -R_z \end{aligned}$$

Work in a coordinate system defined by unitary $Q_v$ and $Q_z$ where such that

$$Q_v V_n = \begin{bmatrix} I_n \\ 0 \end{bmatrix} \text{ and } Q_z Z_n = \begin{bmatrix} I_n \\ 0 \end{bmatrix}.$$

# General Two-sided Backward Error

In new coordinate system, the reduced order model is simply truncation:

$$\begin{aligned}
\hat{A} &= A_{11} \\
\hat{E} &= E_{11} \\
\hat{B} &= B_1 \\
\hat{C} &= C_1
\end{aligned}$$

# General Two-sided Backward Error

Defining

$$Q_z \Delta_A Q_v^* = \begin{bmatrix} \mathcal{A}_{11} & \mathcal{A}_{12} \\ \mathcal{A}_{21} & \mathcal{A}_{22} \end{bmatrix}, \quad Q_z \Delta_E Q_v^* = \begin{bmatrix} \mathcal{E}_{11} & \mathcal{E}_{12} \\ \mathcal{E}_{21} & \mathcal{E}_{22} \end{bmatrix}, \quad Q_z \Delta_B = \begin{bmatrix} \mathcal{B}_1 \\ \mathcal{B}_2 \end{bmatrix},$$

$$\Delta_C Q_v^* = \begin{bmatrix} \mathcal{C}_1 & \mathcal{C}_2 \end{bmatrix}, \quad -R_z Q_v^* = \begin{bmatrix} R_3 & R_4 \end{bmatrix} \text{ and } -Q_z R_v = \begin{bmatrix} R_1 \\ R_2 \end{bmatrix},$$

yields

$$\begin{bmatrix} \mathcal{A}_{11} & \mathcal{A}_{12} \\ \mathcal{A}_{21} & \mathcal{A}_{22} \end{bmatrix} \begin{bmatrix} I_n \\ 0 \end{bmatrix} G_v - \begin{bmatrix} \mathcal{E}_{11} & \mathcal{E}_{12} \\ \mathcal{E}_{21} & \mathcal{E}_{22} \end{bmatrix} \begin{bmatrix} I_n \\ 0 \end{bmatrix} F_v - \begin{bmatrix} \mathcal{B}_1 \\ \mathcal{B}_2 \end{bmatrix} X = \begin{bmatrix} R_1 \\ R_2 \end{bmatrix}$$

$$\begin{bmatrix} \mathcal{A}_{11}^* & \mathcal{A}_{21}^* \\ \mathcal{A}_{12}^* & \mathcal{A}_{22}^* \end{bmatrix} \begin{bmatrix} I_n \\ 0 \end{bmatrix} G_z - \begin{bmatrix} \mathcal{E}_{11}^* & \mathcal{E}_{21}^* \\ \mathcal{E}_{12}^* & \mathcal{E}_{22}^* \end{bmatrix} \begin{bmatrix} I_n \\ 0 \end{bmatrix} F_z - \begin{bmatrix} \mathcal{C}_1^* \\ \mathcal{C}_2^* \end{bmatrix} Y = \begin{bmatrix} R_3^* \\ R_4^* \end{bmatrix}$$

# General Two-sided Backward Error

We have

$$\begin{bmatrix} \mathcal{A}_{11} \\ \mathcal{A}_{21} \end{bmatrix} G_v - \begin{bmatrix} \mathcal{E}_{11} \\ \mathcal{E}_{21} \end{bmatrix} F_v - \begin{bmatrix} \mathcal{B}_1 \\ \mathcal{B}_2 \end{bmatrix} X = \begin{bmatrix} R_1 \\ R_2 \end{bmatrix}$$

$$G_z^* \begin{bmatrix} \mathcal{A}_{11} & \mathcal{A}_{12} \end{bmatrix} - F_z^* \begin{bmatrix} \mathcal{E}_{11} & \mathcal{E}_{12} \end{bmatrix} - Y^* \begin{bmatrix} \mathcal{C}_1 & \mathcal{C}_2 \end{bmatrix} = \begin{bmatrix} R_3 & R_4 \end{bmatrix}.$$

and we can set $\mathcal{A}_{22} = \mathcal{E}_{22} = 0$.

$R_2$ and $R_4$ equations independent. $R_1$ and $R_3$ equations coupled.

$$\begin{aligned} \mathcal{A}_{21} G_v - \mathcal{E}_{21} F_v - \mathcal{B}_2 X &= R_2 \\ G_z^* \mathcal{A}_{12} - F_z^* \mathcal{E}_{12} - Y^* \mathcal{C}_2 &= R_4 \\ \mathcal{A}_{11} G_v - \mathcal{E}_{11} F_v - \mathcal{B}_1 X &= R_1 \\ G_z^* \mathcal{A}_{11} - F_z^* \mathcal{E}_{11} - Y^* \mathcal{C}_1 &= R_3 \end{aligned}$$

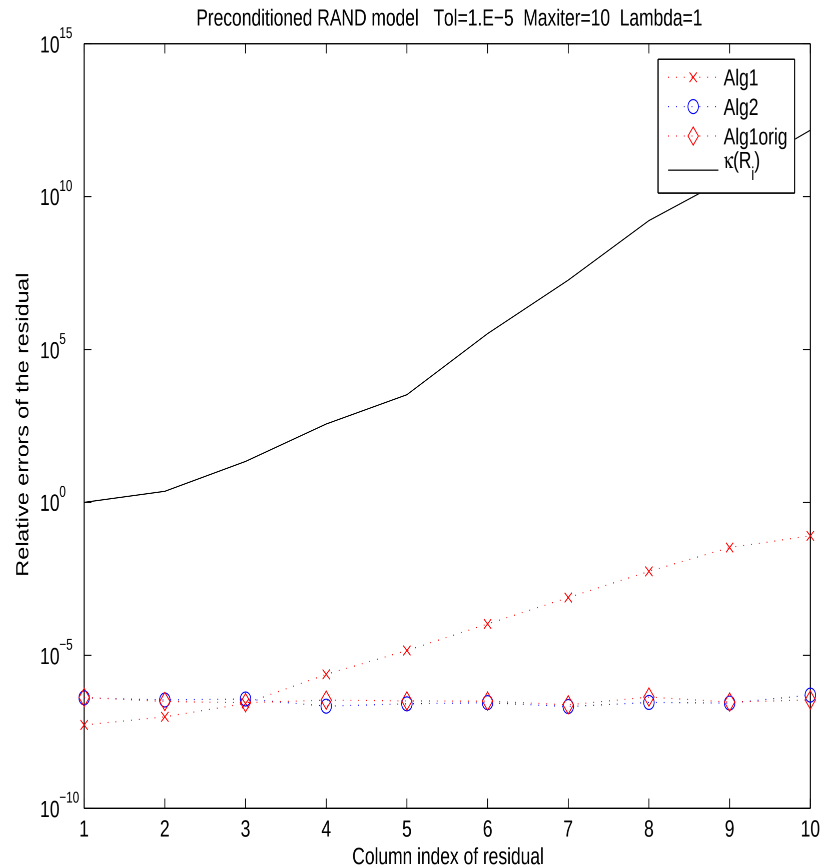
# General Two-sided Backward Error

$$\begin{bmatrix} G_v^* & -F_v^* & -X^* \end{bmatrix} \begin{bmatrix} \mathcal{A}_{21}^* \\ \mathcal{E}_{21}^* \\ \mathcal{B}_2^* \end{bmatrix} = R_2^* \,, \quad \begin{bmatrix} G_z^* & -F_z^* & -Y^* \end{bmatrix} \begin{bmatrix} \mathcal{A}_{12} \\ \mathcal{E}_{12} \\ \mathcal{C}_2 \end{bmatrix} = R_4,$$

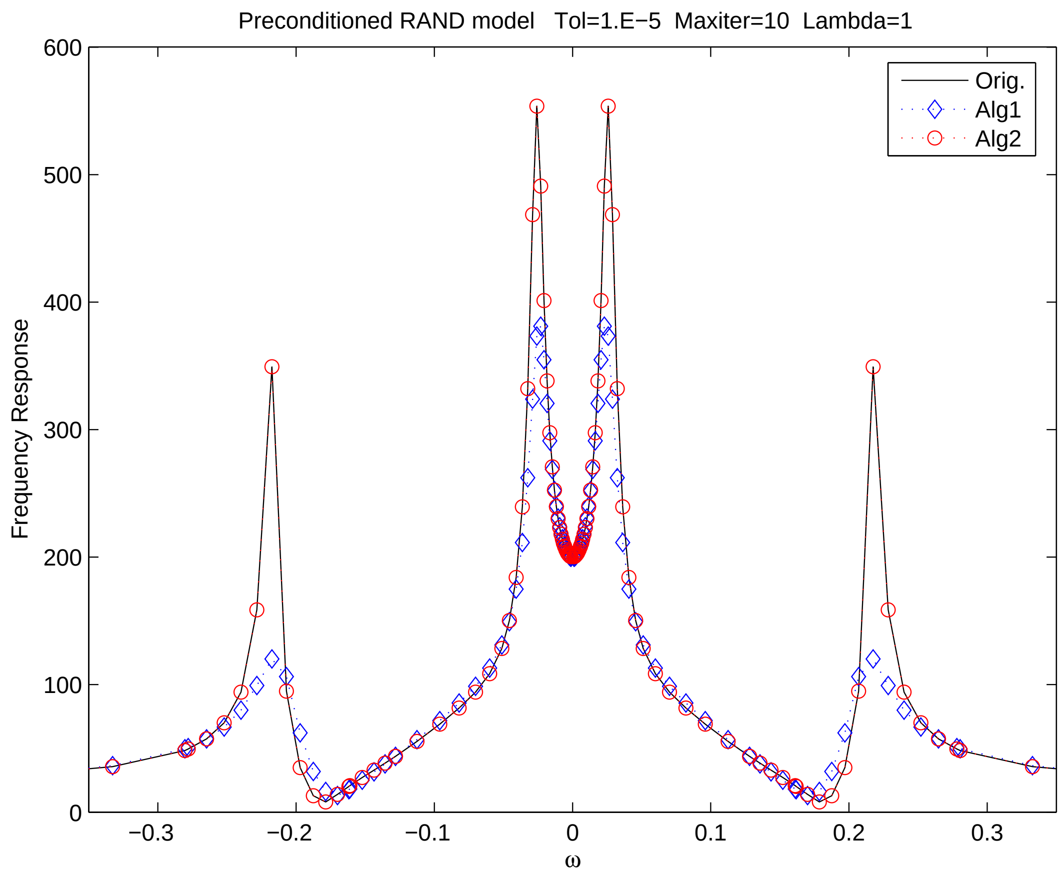
and

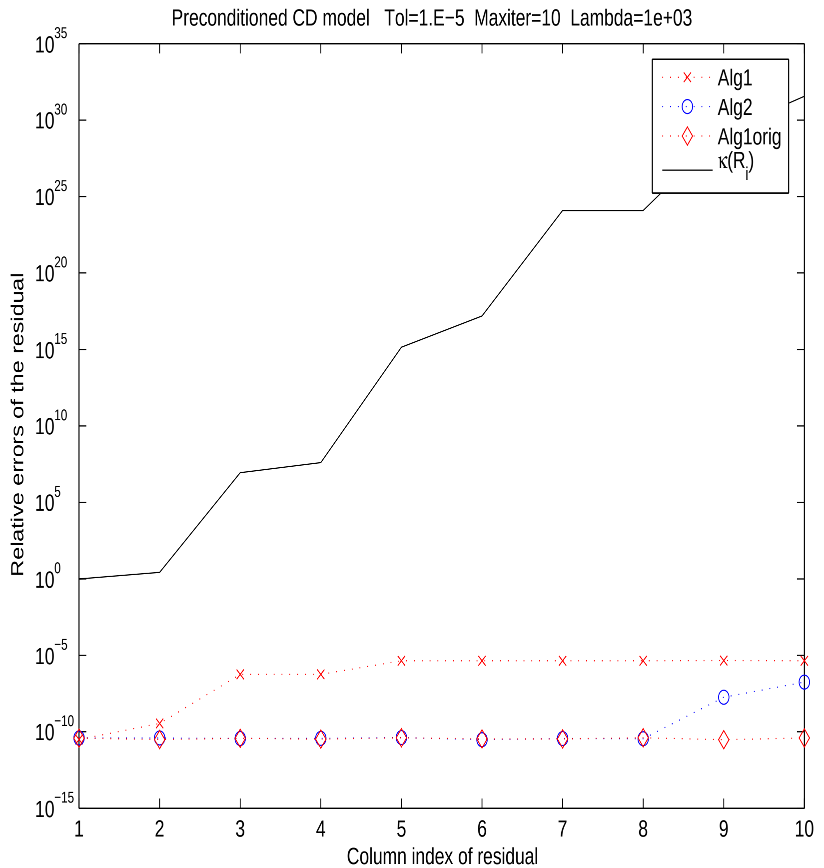
$$\begin{bmatrix} I \otimes G_v^* & -I \otimes F_v^* & -I \otimes X^* & 0 \\ G_z^* \otimes I & -F_v^* \otimes I & 0 & -Y^* \otimes I \end{bmatrix} \begin{bmatrix} vec(\mathcal{A}_{11}) \\ vec(\mathcal{E}_{11}) \\ vec(\mathcal{B}_1) \\ vec(\mathcal{C}_1) \end{bmatrix} = \begin{bmatrix} vec(R_1) \\ vec(R_3) \end{bmatrix}.$$

- underdetermined and properties of systems $(G_v, F_v, X)$ and $(G_z, F_z, Y)$ give ranks
- can find minimal solution, i.e., “nearest” system
- backward ROM error requires $Z_n^* R_v = 0$ and $V_n^* R_z = 0$.

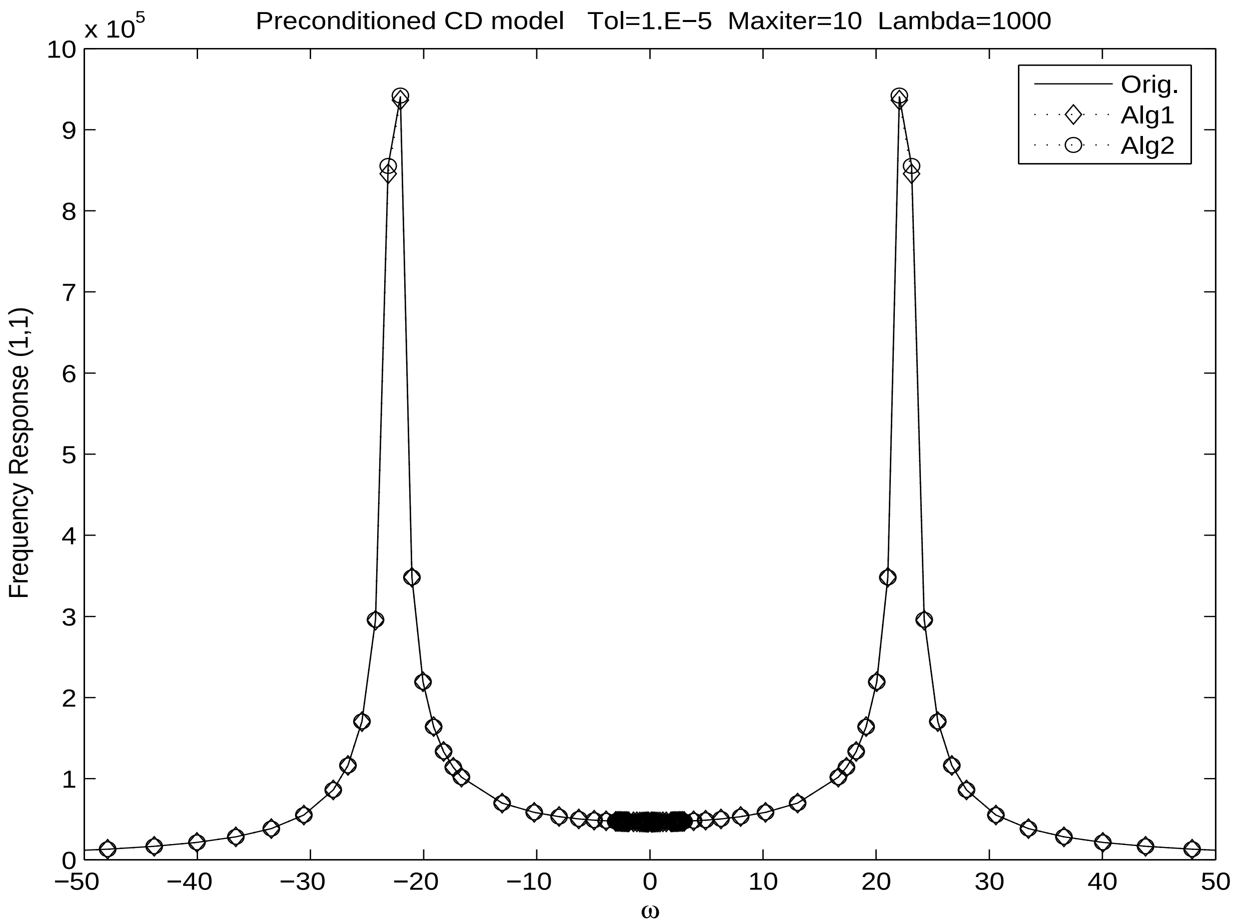



| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Alg1 | 24 | 23 | 23 | 23 | 24 | 25 | 25 | 25 | 25 | 25 |
| Alg2 | 12 | 12 | 13 | 15 | 17 | 19 | 20 | 22 | 23 | 24 |

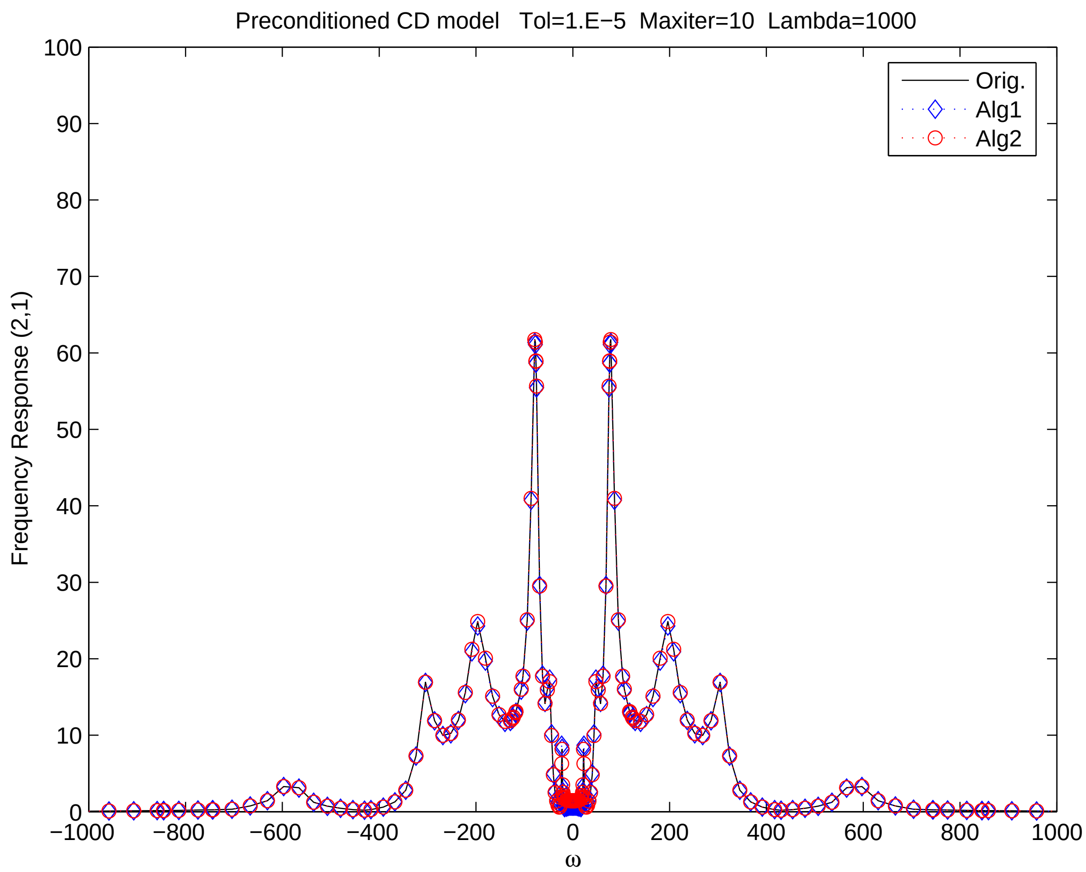

Preconditioned RAND model   Tol=1.E−5  Maxiter=10  Lambda=1
Orig.
Alg1
Alg2
Frequency Response
600
500
400
300
200
100
0
−0.3
−0.2
−0.1
0
0.1
0.2
0.3
ω



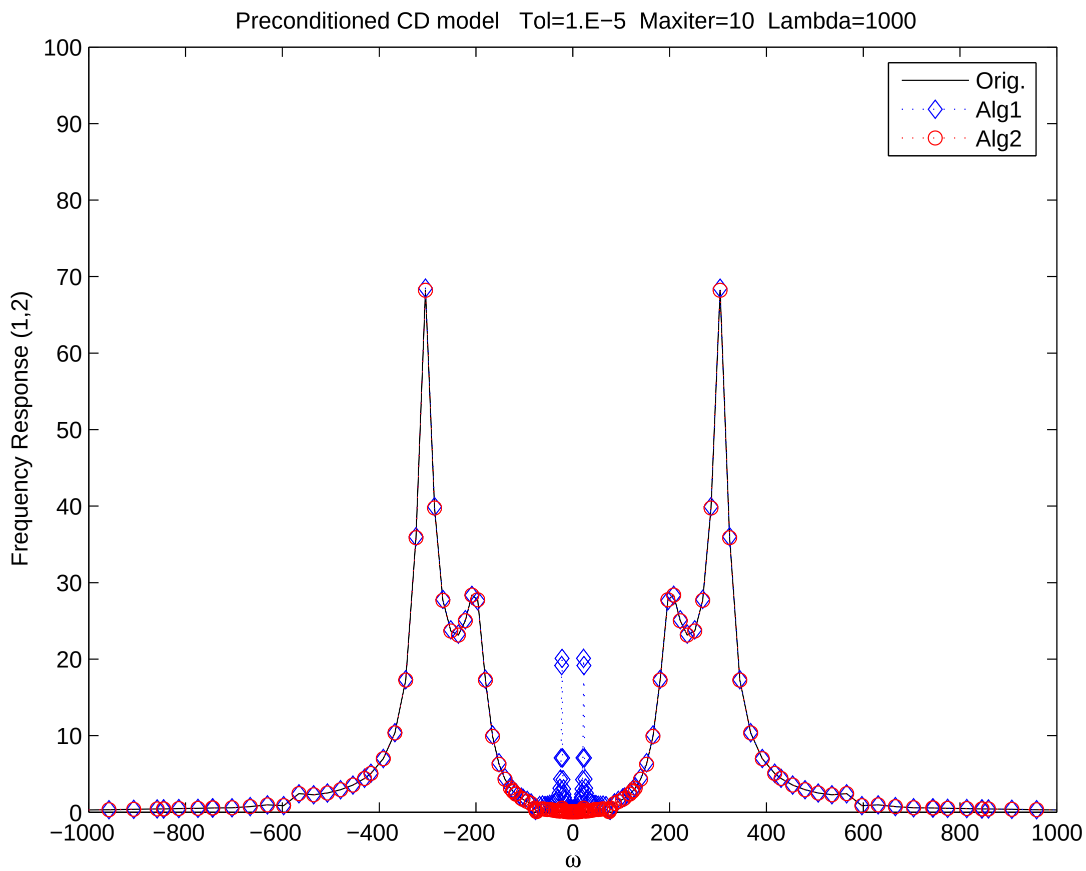

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Alg1 | 28 | 24 | 29 | 28 | 26 | 29 | 29 | 26 | 28 | 29 |
| Alg2 | 17 | 20 | 43 | 45 | 73 | 76 | 99 | 94 | 100 | 100 |



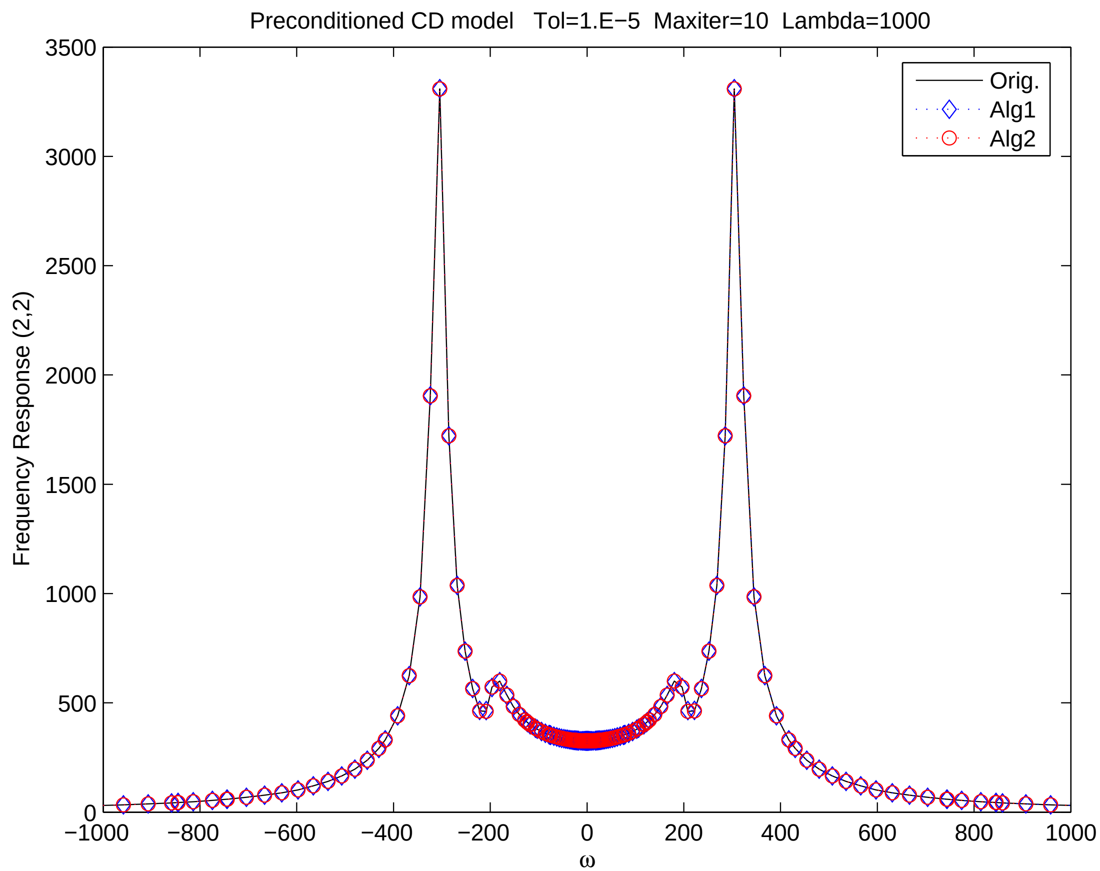
Preconditioned CD model Tol=1.E−5 Maxiter=10 Lambda=1000
x 10^5
Frequency Response (1,1)
ω
Orig.
Alg1
Alg2



Preconditioned CD model   Tol=1.E−5  Maxiter=10  Lambda=1000
Frequency Response (2,1)
ω
Orig.
Alg1
Alg2
100
90
80
70
60
50
40
30
20
10
0
−1000
−800
−600
−400
−200
0
200
400
600
800
1000



Preconditioned CD model   Tol=1.E−5  Maxiter=10  Lambda=1000
Frequency Response (1,2)
ω
Orig.
Alg1
Alg2
100
90
80
70
60
50
40
30
20
10
0
−1000
−800
−600
−400
−200
0
200
400
600
800
1000



Preconditioned CD model   Tol=1.E−5  Maxiter=10  Lambda=1000
Orig.
Alg1
Alg2
Frequency Response (2,2)
ω
3500
3000
2500
2000
1500
1000
500
0
−1000
−800
−600
−400
−200
0
200
400
600
800
1000

# Conclusion

- sufficient conditions for backward error depend on $(G_v, F_v, X)$ and $(G_z, F_z, Y)$
- can add constraints to backward error form, e.g., earlier one-sided forms are consistent with these equations
- more constraints might allow control in algorithm
- can find “nearest” system based on minimal norm solutions
- useful in analyzing algorithms (?)
- useful in defining algorithms (?)
- structure in Sylvester residual (?)
- insight into sensitivitiy (?)