Ubiquitous Doubling Algorithms, General Theory, and Applications

Ren-Cang Li
University of Texas at Arlington

with collaborators

Tsung-Ming Huang, Wen-Wei Lin

NL2A, CIRM Luminy, France

October 24, 2016

Outline

Outline

Simple and Yet Powerful Idea

Numerous approximations rely on simple iteration:

$$X_{j+1} = f(X_j) \text{ for } j = 0, 1, \ldots, \text{ gievn } X_0$$

to generate a sequence of approximations that hopefully converges to the desired target.

Often too slow, but can we make it go faster? Any ideas?

What about skipping all X_j but for $j = 2^i$?
Kind of like repeatedly squaring:

$$x, x^2, (x^2)^2 = x^4, (x^4)^2 = x^8, \ldots.$$

This is the basics of the *Doubling Algorithm* (DA).

Simple and Yet Powerful Idea

Numerous approximations rely on simple iteration:

$$X_{j+1} = f(X_j) \text{ for } j = 0, 1, \ldots, \text{ gievn } X_0$$

to generate a sequence of approximations that hopefully converges to the desired target.

Often too slow, but can we make it go faster? Any ideas?

What about skipping all X_j but for $j = 2^i$?
Kind of like repeatedly squaring:

$$x, x^2, (x^2)^2 = x^4, (x^4)^2 = x^8, \ldots .$$

This is the basics of the *Doubling Algorithm* (DA).

Simple and Yet Powerful Idea

Numerous approximations rely on simple iteration:

$$X_{j+1} = f(X_j) \text{ for } j = 0, 1, \ldots, \text{ gievn } X_0$$

to generate a sequence of approximations that hopefully converges to the desired target.

Often too slow, but can we make it go faster? Any ideas?

What about skipping all X_j but for $j = 2^i$?
Kind of like repeatedly squaring:

$$x, x^2, (x^2)^2 = x^4, (x^4)^2 = x^8, \ldots.$$

This is the basics of the *Doubling Algorithm* (DA).

Simple and Yet Powerful Idea

Numerous approximations rely on simple iteration:

$$X_{j+1} = f(X_j) \text{ for } j = 0, 1, \ldots, \text{ gievn } X_0$$

to generate a sequence of approximations that hopefully converges to the desired target.

Often too slow, but can we make it go faster? Any ideas?

What about skipping all X_j but for $j = 2^i$?
Kind of like repeatedly squaring:

$$x, x^2, (x^2)^2 = x^4, (x^4)^2 = x^8, \ldots .$$

This is the basics of the *Doubling Algorithm* (DA).

Simple and Yet Powerful Idea

Numerous approximations rely on simple iteration:

$$X_{j+1} = f(X_j) \text{ for } j = 0, 1, \ldots, \text{ gievn } X_0$$

to generate a sequence of approximations that hopefully converges to the desired target.

Often too slow, but can we make it go faster? Any ideas?

What about skipping all X_j but for $j = 2^i$?
Kind of like repeatedly squaring:

$$x, x^2, (x^2)^2 = x^4, (x^4)^2 = x^8, \ldots.$$

This is the basics of the *Doubling Algorithm* (DA).

Simple and Yet Powerful Idea (cont'd)

To accelerate $X_{j+1} = f(X_j)$ by computing X_j for $j = 2^i$ only.

Simple and Yet Powerful Idea (cont'd)

To accelerate $X_{j+1} = f(X_j)$ by computing X_j for $j = 2^i$ only.
Make sense only if there is a cost-effective path $X_{2^{i-1}} \to X_{2^i}$.

Simple and Yet Powerful Idea (cont'd)

To accelerate $X_{j+1} = f(X_j)$ by computing X_j for $j = 2^i$ only. Make sense only if there is a cost-effective path $X_{2^{i-1}} \to X_{2^i}$.

No such a path guaranteed in general, unless f is linear: $f(X) = \Phi X$, and then

$$X_{2^i} = \Phi^{2^i} X_0 = \Phi^{2^{i-1}} X_{2^{i-1}}.$$

But still need to be able to square $\Phi^{2^{i-1}}$ efficiently.

To accelerate $X_{j+1} = f(X_j)$ by computing X_j for $j = 2^i$ only.
Make sense only if there is a cost-effective path $X_{2^{i-1}} \to X_{2^i}$.

No such a path guaranteed in general, unless f is linear:
$f(X) = \Phi X$, and then

$$X_{2^i} = \Phi^{2^i} X_0 = \Phi^{2^{i-1}} X_{2^{i-1}}.$$

But still need to be able to square $\Phi^{2^{i-1}}$ efficiently.

Example (Discrete-time Algebraic Riccati Equation):
$X = A^T X(I + GX)^{-1}A + H$.

To accelerate $X_{j+1} = f(X_j)$ by computing X_j for $j = 2^i$ only. Make sense only if there is a cost-effective path $X_{2^{i-1}} \to X_{2^i}$.

No such a path guaranteed in general, unless f is linear: $f(X) = \Phi X$, and then

$$X_{2^i} = \Phi^{2^i} X_0 = \Phi^{2^{i-1}} X_{2^{i-1}}.$$

But still need to be able to square $\Phi^{2^{i-1}}$ efficiently.

Example (Discrete-time Algebraic Riccati Equation): $X = A^T X (I + GX)^{-1} A + H$. Naturally

$$X_{j+1} = f(X_j) := A^{\mathrm{T}} X_j (I + GX_j)^{-1} A + H, \quad \text{given } X_0$$

It can be turned into a linear iteration of twice the dimensions by Bernoulli substitution.

To accelerate $X_{j+1} = f(X_j)$ by computing X_j for $j = 2^i$ only. Make sense only if there is a cost-effective path $X_{2^{i-1}} \to X_{2^i}$.

No such a path guaranteed in general, unless f is linear: $f(X) = \Phi X$, and then

$$X_{2^i} = \Phi^{2^i} X_0 = \Phi^{2^{i-1}} X_{2^{i-1}}.$$

But still need to be able to square $\Phi^{2^{i-1}}$ efficiently.

Example (Discrete-time Algebraic Riccati Equation): $X = A^T X (I + GX)^{-1} A + H$. Naturally

$$X_{j+1} = f(X_j) := A^{\mathrm{T}} X_j (I + GX_j)^{-1} A + H, \quad \text{given } X_0$$

It can be turned into a linear iteration of twice the dimensions by Bernoulli substitution.

Anderson (1978) did just that.

Go Forward

In the last 12 years also saw tremendous progresses in using "doubling idea" to solve

three types nonlinear matrix equations (NMEs)

and related applications.

Go Forward

In the last 12 years also saw tremendous progresses in using "doubling idea" to solve

three types nonlinear matrix equations (NMEs)

and related applications.

Current developments: mostly equation-wise, somewhat fragmented, not totally coherent.

Go Forward

In the last 12 years also saw tremendous progresses in using “doubling idea” to solve

three types nonlinear matrix equations (NMEs)

and related applications.

Current developments: mostly equation-wise, somewhat fragmented, not totally coherent.

This talk attempts to provide a coherent theory to go forward.

References I

Brian D. O. Anderson.
Second-order convergent algorithms for the steady-state Riccati equation.
Int. J. Control, 28(2):295–306, 1978.

Morishige Kimura.
Convergence of the doubling algorithm for the discrete-time algebraic Riccati equation.
International Journal of Systems Science, 19(5):701–711, 1988.

J. Juang and W.-W. Lin.
Nonsymmetric algebraic Riccati equations and Hamiltonian-like matrices.
SIAM J. Matrix Anal. Appl., 20(1):228–243, 1998.

J. Juang.
Global existence and stability of solutions of matrix Riccati equations.
J. Math. Anal. and Appl., 258(1):1–12, 2001.

E. K.-W. Chu, H.-Y. Fan, and W.-W. Lin.
A structure-preserving doubling algorithm for continuous-time algebraic Riccati equations.
Linear Algebra Appl., 396:55 – 80, 2005.

T.-M. Hwang, E. K.-W. Chu, and W.-W. Lin.
A generalized structure-preserving doubling algorithm for generalized discrete-time algebraic Riccati equations.
Int. J. Control, 78(14):1063–1075, 2005.

W.-W. Lin and S.-F. Xu.
Convergence analysis of structure-preserving doubling algorithms for Riccati-type matrix equations.
SIAM J. Matrix Anal. Appl., 28(1):26–39, 2006.

X. Guo, W. Lin, and S. Xu.
A structure-preserving doubling algorithm for nonsymmetric algebraic Riccati equation.
Numer. Math., 103:393–412, 2006.

References II

Eric King-Wah Chu, Tsung-Min Hwang, Wen-Wei Lin, and Chin-Tien Wu.
Vibration of fast trains, palindromic eigenvalue problems and structure-preserving doubling algorithms.
J. Comput. Appl. Math., 219:237–252, 2008.

Chun-Yueh Chiang, Eric King-Wah Chu, Chun-Hua Guo, Tsung-Ming Huang, Wen-Wei Lin, and Shu-Fang Xu.
Convergence analysis of the doubling algorithm for several nonlinear matrix equations in the critical case.
SIAM J. Matrix Anal. Appl., 31(2):227–247, 2009.

D. A. Bini, B. Meini, and F. Poloni, *Transforming algebraic Riccati equations into unilateral quadratic matrix equations*, Numer. Math., 116 (2010), pp. 553–578.

C. Guo and W. Lin.
Solving a structured quadratic eigenvalue problem by a structure-preserving doubling algorithm.
SIAM J. Matrix Anal. Appl., 31(5):2784–2801, 2010.

C.-H. Guo, Y.-C. Kuo, and W.-W. Lin.
Numerical solution of nonlinear matrix equations arising from Green's function calculations in nano research.
J. Comput. Appl. Math., 236:4166–4180, 2012.

C.-H. Guo, Y.-C. Kuo, and W.-W. Lin.
On a nonlinear matrix equation arising in nano research.
SIAM J. Matrix Anal. Appl., 33(1):235–262, 2012.

Jungong Xue, Shufang Xu, and Ren-Cang Li.
Accurate solutions of M-matrix Sylvester equations.
Numer. Math., 120(4):639–670, 2012.

Jungong Xue, Shufang Xu, and Ren-Cang Li.
Accurate solutions of M-matrix algebraic Riccati equations.
Numer. Math., 120(4):671–700, 2012.

References III

Wei-Guo Wang, Wei-Chao Wang, and Ren-Cang Li.
Alternating-directional doubling algorithm for M-matrix algebraic Riccati equations.
SIAM J. Matrix Anal. Appl., 33(1):170–194, 2012.

Linzhang Lu, Fei Yuan, and Ren-Cang Li.
A new look at the doubling algorithm for a structured palindromic quadratic eigenvalue problem.
Numer. Linear Algebra Appl., 22:393–409, 2015.

G. T. Nguyen and F.Poloni.
Componentwise accurate fluid queue computations using doubling algorithms.
Numer. Math., 130:763–792, 2015.

Jungong Xue and Ren-Cang Li.
Highly accurate doubling algorithms for M-matrix algebraic Riccati equations.
Numer. Math., 2016.

Linzhang Lu, Teng Wang, Yueh-Cheng Kuo, Ren-Cang Li, and Wen-Wei Lin.
A fast algorithm for fast train palindromic quadratic eigenvalue problems.
Technical Report 2016-03, Department of Mathematics, University of Texas at Arlington, April 2016.

Outline

Continuous-time Algebraic Riccati Equation (CARE) Type

$$XDX - AX - XB + C = 0.$$

Continuous-time Algebraic Riccati Equation (CARE) Type

$$XDX - AX - XB + C = 0.$$

- CARE $A^{\top}X + XA - XGX + H = 0$ from continuous-time control system [...]

$$XDX - AX - XB + C = 0.$$

- CARE $A^\top X + XA - XGX + H = 0$ from continuous-time control system [...]
- M-matrix Algebraic Riccati equation (MARE) $XDX - AX - XB + C = 0$ from transport theory of particles [Juang (1995), C. Guo (2001), X. Guo, Lin, Xu (2006), ...]

$$XDX - AX - XB + C = 0.$$

- CARE $A^\top X + XA - XGX + H = 0$ from continuous-time control system [...]
- M-matrix Algebraic Riccati equation (MARE) $XDX - AX - XB + C = 0$ from transport theory of particles [Juang (1995), C. Guo (2001), X. Guo, Lin, Xu (2006), ...]
- MARE $XDX - AX - XB + C = 0$ from Markov-modulated fluid queue theory [Latouche & Taylor (2009), C. Guo (2001), X. Guo, Lin, Xu (2006), ...]

$$XDX - AX - XB + C = 0.$$

- CARE $A^\top X + XA - XGX + H = 0$ from continuous-time control system [...]
- M-matrix Algebraic Riccati equation (MARE) $XDX - AX - XB + C = 0$ from transport theory of particles [Juang (1995), C. Guo (2001), X. Guo, Lin, Xu (2006), ...]
- MARE $XDX - AX - XB + C = 0$ from Markov-modulated fluid queue theory [Latouche & Taylor (2009), C. Guo (2001), X. Guo, Lin, Xu (2006), ...]
- H^*ARE from the Laplace transform inversion method in Markov modulated fluid flow [Ahn & Ramaswami (2004), Liu & Xue (2012), ...]

Discrete-time Algebraic Riccati Equation (DARE) Type

$$BX(I + GX)^{-1}A + H - X = 0.$$

Discrete-time Algebraic Riccati Equation (DARE) Type

$$BX(I+GX)^{-1}A+H-X=0.$$

- DARE $A^\top X(I+GX)^{-1}A-X+H=0$ from Discrete-time control system [...]

Yet Another Type of Nonlinear Matrix Equation

$$X + BX^{-1}A = Q.$$

Yet Another Type of Nonlinear Matrix Equation

$$X + BX^{-1}A = Q.$$

- Solve Palindromic Quadratic Eigenvalue Problem (PQEP) $P(\lambda)z \equiv (\lambda^2 A^\mathsf{T} + \lambda Q + A)z = 0$ from high speed trains [Hilliges, Mehl, & V. Mehrmann (2004), ...]
- Compute Green function from quantum transport in nano research [S. Datta (2000), Guo & Lin (2010), ...]
- Solve QEP $P(\lambda)\boldsymbol{x} \equiv \left(\lambda^2 B + \lambda Q + A\right)\boldsymbol{x} = 0$ from Retarded Time-Delay System (TDS) [Jarlebring (2008)]
- $X + A^\mathsf{T} X^{-1} A = Q$ from surface acoustic wave simulation in telecommunication [Campbell (1998)]

Commonality

Let X be a solution to one of the equations. Then $\begin{array}{cc} & m \\ \begin{array}{c} m \\ n \end{array} & \begin{bmatrix} I \\ X \end{bmatrix} \end{array}$ is basis matrix of an eigenspace for $\mathscr{A} - \lambda\mathscr{B}$:

$$\mathscr{A}\begin{bmatrix} I \\ X \end{bmatrix} \equiv \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}\begin{bmatrix} I \\ X \end{bmatrix} = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}\begin{bmatrix} I \\ X \end{bmatrix} M \equiv \mathscr{B}\begin{bmatrix} I \\ X \end{bmatrix} M,$$

where M is $m \times m$. As a consequence, $\mathrm{eig}(M) \subset \mathrm{eig}(\mathscr{A}, \mathscr{B})$.

Let X be a solution to one of the equations. Then $\begin{array}{cc} & m \\ \begin{array}{c} m \\ n \end{array} & \begin{bmatrix} I \\ X \end{bmatrix} \end{array}$ is basis matrix of an eigenspace for $\mathscr{A} - \lambda \mathscr{B}$:

$$\mathscr{A} \begin{bmatrix} I \\ X \end{bmatrix} \equiv \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} I \\ X \end{bmatrix} = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} \begin{bmatrix} I \\ X \end{bmatrix} M \equiv \mathscr{B} \begin{bmatrix} I \\ X \end{bmatrix} M,$$

where M is $m \times m$. As a consequence, $\operatorname{eig}(M) \subset \operatorname{eig}(\mathscr{A}, \mathscr{B})$.

Blockwise we have

$$A_{11} + A_{12}X = (B_{11} + B_{12}X)M, \quad A_{21} + A_{22}X = (B_{21} + B_{22}X)M.$$

Assume $(B_{11} + B_{12}X)$ is invertible to get

$$A_{21} + A_{22}X = (B_{21} + B_{22}X) \underbrace{(B_{11} + B_{12}X)^{-1}(A_{11} + A_{12}X)}_{M}.$$

Commonality (cont'd)

$$\mathscr{A}\begin{bmatrix} I \\ X \end{bmatrix} \equiv \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}\begin{bmatrix} I \\ X \end{bmatrix} = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}\begin{bmatrix} I \\ X \end{bmatrix} M \equiv \mathscr{B}\begin{bmatrix} I \\ X \end{bmatrix} M,$$

$$A_{21} + A_{22}X = (B_{21} + B_{22}X)(B_{11} + B_{12}X)^{-1}(A_{11} + A_{12}X).$$

Commonality (cont'd)

$$\mathscr{A}\begin{bmatrix} I \\ X \end{bmatrix} \equiv \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}\begin{bmatrix} I \\ X \end{bmatrix} = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}\begin{bmatrix} I \\ X \end{bmatrix} M \equiv \mathscr{B}\begin{bmatrix} I \\ X \end{bmatrix} M,$$

$$A_{21} + A_{22}X = (B_{21} + B_{22}X)(B_{11} + B_{12}X)^{-1}(A_{11} + A_{12}X).$$

Case: $\mathscr{B} = I_{m+n}$:

$$A_{21} + A_{22}X = X(A_{11} + A_{12}X)$$

gives

$$-XA_{12}X + A_{22}X - XA_{11} + A_{21} = 0.$$

This is "CARE".

Commonality (cont'd)

$$\mathscr{A}\begin{bmatrix} I \\ X \end{bmatrix} \equiv \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}\begin{bmatrix} I \\ X \end{bmatrix} = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}\begin{bmatrix} I \\ X \end{bmatrix} M \equiv \mathscr{B}\begin{bmatrix} I \\ X \end{bmatrix} M,$$

$$A_{21} + A_{22}X = (B_{21} + B_{22}X)(B_{11} + B_{12}X)^{-1}(A_{11} + A_{12}X).$$

$$\mathscr{A}\begin{bmatrix} I \\ X \end{bmatrix} \equiv \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}\begin{bmatrix} I \\ X \end{bmatrix} = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}\begin{bmatrix} I \\ X \end{bmatrix} M \equiv \mathscr{B}\begin{bmatrix} I \\ X \end{bmatrix} M,$$

$$A_{21} + A_{22}X = (B_{21} + B_{22}X)(B_{11} + B_{12}X)^{-1}(A_{11} + A_{12}X).$$

Case: $A_{12} = 0$, $A_{22} = I$, $B_{11} = I$, and $B_{21} = 0$, i.e.,

$$\mathscr{A} = \begin{bmatrix} A_{11} & 0 \\ A_{21} & I \end{bmatrix}, \quad \mathscr{B} = \begin{bmatrix} I & B_{12} \\ 0 & B_{22} \end{bmatrix}.$$

We have $A_{21} + X = B_{22}X(I + B_{12}X)^{-1}A_{11}$, or equivalently

$$B_{22}X(I + B_{12}X)^{-1}A_{11} - A_{21} - X = 0.$$

This is “DARE”.

Commonality (cont'd)

$$\mathscr{A}\begin{bmatrix} I \\ X \end{bmatrix} \equiv \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}\begin{bmatrix} I \\ X \end{bmatrix} = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}\begin{bmatrix} I \\ X \end{bmatrix} M \equiv \mathscr{B}\begin{bmatrix} I \\ X \end{bmatrix} M,$$

$$A_{21} + A_{22}X = (B_{21} + B_{22}X)(B_{11} + B_{12}X)^{-1}(A_{11} + A_{12}X).$$

$$\mathscr{A}\begin{bmatrix} I \\ X \end{bmatrix} \equiv \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}\begin{bmatrix} I \\ X \end{bmatrix} = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}\begin{bmatrix} I \\ X \end{bmatrix} M \equiv \mathscr{B}\begin{bmatrix} I \\ X \end{bmatrix} M,$$

$$A_{21} + A_{22}X = (B_{21} + B_{22}X)(B_{11} + B_{12}X)^{-1}(A_{11} + A_{12}X).$$

Case: $m = n$, $A_{12} = 0$, $A_{22} = I$, $B_{11} = 0$, $B_{12} = I$, and $B_{22} = 0$, i.e.,

$$\mathscr{A} = \begin{bmatrix} A_{11} & 0 \\ A_{21} & I \end{bmatrix}, \quad \mathscr{B} = \begin{bmatrix} 0 & I \\ B_{21} & 0 \end{bmatrix}.$$

We have $A_{21} + X = B_{21}X^{-1}A_{11}$, or equivalently

$$X - B_{21}X^{-1}A_{11} = -A_{21}.$$

This is the NME.

Outline

Key ingredients

Three ingredients to DA:

- doubling transformation theorem
- initial setups (two standard forms)
- doubling iterative kernel

Outline

Doubling Transformation Theorem

Let $\mathscr{A} - \lambda\mathscr{B} \in \mathbb{C}^{\ell\times\ell}$ be a regular matrix pencil: $\det(\mathscr{A} - \lambda\mathscr{B}) \not\equiv 0$.

Doubling Transformation Theorem

Let $\mathscr{A} - \lambda\mathscr{B} \in \mathbb{C}^{\ell\times\ell}$ be a regular matrix pencil: det$(\mathscr{A} - \lambda\mathscr{B}) \not\equiv 0$.

Let $\mathscr{A}_\perp, \mathscr{B}_\perp \in \mathbb{C}^{\ell\times\ell}$ such that

$$\operatorname{rank}([\mathscr{A}_\perp, \mathscr{B}_\perp]) = \ell, \quad [\mathscr{A}_\perp, \mathscr{B}_\perp]\begin{bmatrix} \mathscr{B} \\ -\mathscr{A} \end{bmatrix} = 0.$$

Doubling Transformation Theorem

Let $\mathscr{A} - \lambda\mathscr{B} \in \mathbb{C}^{\ell\times\ell}$ be a regular matrix pencil: $\det(\mathscr{A} - \lambda\mathscr{B}) \neq 0$.
Let $\mathscr{A}_\perp, \mathscr{B}_\perp \in \mathbb{C}^{\ell\times\ell}$ such that

$$\operatorname{rank}([\mathscr{A}_\perp, \mathscr{B}_\perp]) = \ell, \quad [\mathscr{A}_\perp, \mathscr{B}_\perp]\begin{bmatrix} \mathscr{B} \\ -\mathscr{A} \end{bmatrix} = 0.$$

Define $\widetilde{\mathscr{A}} = \mathscr{A}_\perp\mathscr{A}$ and $\widetilde{\mathscr{B}} = \mathscr{B}_\perp\mathscr{B}$. Then $\widetilde{\mathscr{A}} - \lambda\widetilde{\mathscr{B}}$ is regular.

Doubling Transformation Theorem

Let $\mathscr{A} - \lambda\mathscr{B} \in \mathbb{C}^{\ell\times\ell}$ be a regular matrix pencil: $\det(\mathscr{A} - \lambda\mathscr{B}) \not\equiv 0$.
Let $\mathscr{A}_\perp$, $\mathscr{B}_\perp \in \mathbb{C}^{\ell\times\ell}$ such that

$$\operatorname{rank}([\mathscr{A}_\perp, \mathscr{B}_\perp]) = \ell, \quad [\mathscr{A}_\perp, \mathscr{B}_\perp]\begin{bmatrix} \mathscr{B} \\ -\mathscr{A} \end{bmatrix} = 0.$$

Define $\widetilde{\mathscr{A}} = \mathscr{A}_\perp\mathscr{A}$ and $\widetilde{\mathscr{B}} = \mathscr{B}_\perp\mathscr{B}$. Then $\widetilde{\mathscr{A}} - \lambda\widetilde{\mathscr{B}}$ is regular.
Let $\mathcal{Z} = \mathcal{R}(Z)$ be an eigenspace, i.e., $\mathscr{A}Z = \mathscr{B}ZM$. Then

$$\widetilde{\mathscr{A}}Z = \widetilde{\mathscr{B}}ZM^2.$$

Doubling Transformation Theorem

Let $\mathscr{A} - \lambda \mathscr{B} \in \mathbb{C}^{\ell \times \ell}$ be a regular matrix pencil: $\det(\mathscr{A} - \lambda \mathscr{B}) \not\equiv 0$.
Let $\mathscr{A}_\perp$, $\mathscr{B}_\perp \in \mathbb{C}^{\ell \times \ell}$ such that

$$\mathrm{rank}([\mathscr{A}_\perp, \mathscr{B}_\perp]) = \ell, \quad [\mathscr{A}_\perp, \mathscr{B}_\perp] \begin{bmatrix} \mathscr{B} \\ -\mathscr{A} \end{bmatrix} = 0.$$

Define $\widetilde{\mathscr{A}} = \mathscr{A}_\perp \mathscr{A}$ and $\widetilde{\mathscr{B}} = \mathscr{B}_\perp \mathscr{B}$. Then $\widetilde{\mathscr{A}} - \lambda \widetilde{\mathscr{B}}$ is regular.
Let $\mathcal{Z} = \mathcal{R}(Z)$ be an eigenspace, i.e., $\mathscr{A} Z = \mathscr{B} Z M$. Then

$$\widetilde{\mathscr{A}} Z = \widetilde{\mathscr{B}} Z M^2.$$

- keep "doubling" to quickly arrive at some null space, provided $\mathrm{eig}(M) \in \mathbb{D}_-$ (open unit disk);

Doubling Transformation Theorem

Let $\mathscr{A} - \lambda\mathscr{B} \in \mathbb{C}^{\ell\times\ell}$ be a regular matrix pencil: $\det(\mathscr{A} - \lambda\mathscr{B}) \neq 0$.
Let $\mathscr{A}_\perp, \mathscr{B}_\perp \in \mathbb{C}^{\ell\times\ell}$ such that

$$\operatorname{rank}([\mathscr{A}_\perp, \mathscr{B}_\perp]) = \ell, \quad [\mathscr{A}_\perp, \mathscr{B}_\perp]\begin{bmatrix} \mathscr{B} \\ -\mathscr{A} \end{bmatrix} = 0.$$

Define $\widetilde{\mathscr{A}} = \mathscr{A}_\perp\mathscr{A}$ and $\widetilde{\mathscr{B}} = \mathscr{B}_\perp\mathscr{B}$. Then $\widetilde{\mathscr{A}} - \lambda\widetilde{\mathscr{B}}$ is regular.
Let $\mathcal{Z} = \mathcal{R}(Z)$ be an eigenspace, i.e., $\mathscr{A}Z = \mathscr{B}ZM$. Then

$$\widetilde{\mathscr{A}}Z = \widetilde{\mathscr{B}}ZM^2.$$

- keep "doubling" to quickly arrive at some null space, provided $\operatorname{eig}(M) \in \mathbb{D}_-$ (open unit disk);
- maintain special forms for $\mathscr{A}$- and $\mathscr{B}$-matrices to find Z in special form.

Outline

Road Map

$$\mathscr{A}\begin{bmatrix} I \\ X \end{bmatrix} \equiv \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}\begin{bmatrix} I \\ X \end{bmatrix} = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}\begin{bmatrix} I \\ X \end{bmatrix} M \equiv \mathscr{B}\begin{bmatrix} I \\ X \end{bmatrix} M.$$

But eig(M) may not be "favorable", i.e., lying in $\mathbb{D}$ or even $\mathbb{D}_{0-}$ (closed unit disk).

Road Map

$$\mathscr{A}\begin{bmatrix} I \\ X \end{bmatrix} \equiv \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}\begin{bmatrix} I \\ X \end{bmatrix} = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}\begin{bmatrix} I \\ X \end{bmatrix} M \equiv \mathscr{B}\begin{bmatrix} I \\ X \end{bmatrix} M.$$

But eig(M) may not be "favorable", i.e., lying in $\mathbb{D}$ or even $\mathbb{D}_{0-}$ (closed unit disk).

1) Transform $\mathscr{A} - \lambda\mathscr{B} \to \mathscr{A}' - \lambda\mathscr{B}'$ such that

$$\mathscr{A}'\begin{bmatrix} I \\ X \end{bmatrix} = \mathscr{B}'\begin{bmatrix} I \\ X \end{bmatrix}\mathscr{M}, \ \text{eig}(\mathscr{M}) \subset \mathbb{D}_{-} \ \text{or even} \ \text{eig}(\mathscr{M}) \subset \mathbb{D}_{0-}.$$

Road Map

$$\mathscr{A}\begin{bmatrix} I \\ X \end{bmatrix} \equiv \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}\begin{bmatrix} I \\ X \end{bmatrix} = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}\begin{bmatrix} I \\ X \end{bmatrix} M \equiv \mathscr{B}\begin{bmatrix} I \\ X \end{bmatrix} M.$$

But eig(M) may not be "favorable", i.e., lying in $\mathbb{D}$ or even $\mathbb{D}_{0-}$ (closed unit disk).

1) Transform $\mathscr{A} - \lambda\mathscr{B} \to \mathscr{A}' - \lambda\mathscr{B}'$ such that

$$\mathscr{A}'\begin{bmatrix} I \\ X \end{bmatrix} = \mathscr{B}'\begin{bmatrix} I \\ X \end{bmatrix}\mathscr{M}, \ \text{eig}(\mathscr{M}) \subset \mathbb{D}_{-} \ \text{or even} \ \text{eig}(\mathscr{M}) \subset \mathbb{D}_{0-}.$$

One such a transformation is the *Möbius transformation*:

$$\begin{bmatrix} \mathscr{A} \\ \mathscr{B} \end{bmatrix} \to \begin{bmatrix} \mathscr{A}' \\ \mathscr{B}' \end{bmatrix} := \begin{bmatrix} \beta I & -\gamma I \\ \alpha I & \delta I \end{bmatrix}\begin{bmatrix} \mathscr{A} \\ \mathscr{B} \end{bmatrix}.$$

Road Map

$$\mathscr{A}\begin{bmatrix} I \\ X \end{bmatrix} \equiv \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}\begin{bmatrix} I \\ X \end{bmatrix} = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}\begin{bmatrix} I \\ X \end{bmatrix} M \equiv \mathscr{B}\begin{bmatrix} I \\ X \end{bmatrix} M.$$

But eig(M) may not be "favorable", i.e., lying in $\mathbb{D}$ or even $\mathbb{D}_{0-}$ (closed unit disk).

1) Transform $\mathscr{A} - \lambda\mathscr{B} \to \mathscr{A}' - \lambda\mathscr{B}'$ such that

$$\mathscr{A}'\begin{bmatrix} I \\ X \end{bmatrix} = \mathscr{B}'\begin{bmatrix} I \\ X \end{bmatrix}\mathscr{M}, \;\; \text{eig}(\mathscr{M}) \subset \mathbb{D}_{-} \;\text{ or even }\; \text{eig}(\mathscr{M}) \subset \mathbb{D}_{0-}.$$

One such a transformation is the *Möbius transformation*:

$$\begin{bmatrix} \mathscr{A} \\ \mathscr{B} \end{bmatrix} \to \begin{bmatrix} \mathscr{A}' \\ \mathscr{B}' \end{bmatrix} := \begin{bmatrix} \beta I & -\gamma I \\ \alpha I & \delta I \end{bmatrix}\begin{bmatrix} \mathscr{A} \\ \mathscr{B} \end{bmatrix}.$$

2) Perform $\mathscr{A}_0 - \lambda\mathscr{B}_0 = P(\mathscr{A}' - \lambda\mathscr{B}')$ for preferable $\mathscr{A}_0 - \lambda\mathscr{B}_0$.

Road Map

$$\mathscr{A}\begin{bmatrix} I \\ X \end{bmatrix} \equiv \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}\begin{bmatrix} I \\ X \end{bmatrix} = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}\begin{bmatrix} I \\ X \end{bmatrix} M \equiv \mathscr{B}\begin{bmatrix} I \\ X \end{bmatrix} M.$$

But eig(M) may not be "favorable", i.e., lying in $\mathbb{D}$ or even $\mathbb{D}_{0-}$ (closed unit disk).

1) Transform $\mathscr{A} - \lambda\mathscr{B} \to \mathscr{A}' - \lambda\mathscr{B}'$ such that

$$\mathscr{A}'\begin{bmatrix} I \\ X \end{bmatrix} = \mathscr{B}'\begin{bmatrix} I \\ X \end{bmatrix}\mathscr{M}, \ \text{eig}(\mathscr{M}) \subset \mathbb{D}_{-} \ \text{or even} \ \text{eig}(\mathscr{M}) \subset \mathbb{D}_{0-}.$$

One such a transformation is the *Möbius transformation*:

$$\begin{bmatrix} \mathscr{A} \\ \mathscr{B} \end{bmatrix} \to \begin{bmatrix} \mathscr{A}' \\ \mathscr{B}' \end{bmatrix} := \begin{bmatrix} \beta I & -\gamma I \\ \alpha I & \delta I \end{bmatrix}\begin{bmatrix} \mathscr{A} \\ \mathscr{B} \end{bmatrix}.$$

2) Perform $\mathscr{A}_0 - \lambda\mathscr{B}_0 = P(\mathscr{A}' - \lambda\mathscr{B}')$ for preferable $\mathscr{A}_0 - \lambda\mathscr{B}_0$.

3) Perform doubling iteration kernel.

First standard form (SF1)

$$\text{(SF1)} \qquad \mathscr{A}_0 = \begin{array}{c c} & \begin{array}{cc} m & n \end{array} \\ \begin{array}{c} m \\ n \end{array} & \begin{bmatrix} E_0 & 0 \\ -X_0 & I \end{bmatrix} \end{array}, \quad \mathscr{B}_0 = \begin{array}{c c} & \begin{array}{cc} m & n \end{array} \\ \begin{array}{c} m \\ n \end{array} & \begin{bmatrix} I & -Y_0 \\ 0 & F_0 \end{bmatrix} \end{array}.$$

$$
\text{(SF1)} \qquad \mathscr{A}_0 = \begin{array}{c c} & \begin{array}{cc} m & n \end{array} \\ \begin{array}{c} m \\ n \end{array} & \begin{bmatrix} E_0 & 0 \\ -X_0 & I \end{bmatrix} \end{array}, \quad \mathscr{B}_0 = \begin{array}{c c} & \begin{array}{cc} m & n \end{array} \\ \begin{array}{c} m \\ n \end{array} & \begin{bmatrix} I & -Y_0 \\ 0 & F_0 \end{bmatrix} \end{array}.
$$

Essentially Gaussian elimination (assuming involved inversions exist):

$$
\begin{aligned}
[\mathscr{A}' \mid \mathscr{B}'] = \left[\begin{array}{cc|cc} X & X & \color{red}{X} & X \\ X & X & X & X \end{array}\right] &\rightarrow \left[\begin{array}{cc|cc} X & X & I & X \\ X & X & \color{red}{X} & X \end{array}\right] \\
\rightarrow \left[\begin{array}{cc|cc} X & X & I & X \\ X & \color{red}{X} & 0 & X \end{array}\right] &\rightarrow \left[\begin{array}{cc|cc} X & \color{red}{X} & I & X \\ X & I & 0 & X \end{array}\right] \\
\rightarrow \left[\begin{array}{cc|cc} X & 0 & I & X \\ X & I & 0 & X \end{array}\right] &
\end{aligned}
$$

in SF1

Second standard form (SF2)

$$\text{(SF2)} \qquad \mathscr{A}_0 = \begin{array}{c c} & \begin{array}{cc} n & n \end{array} \\ \begin{array}{c} n \\ n \end{array} & \begin{bmatrix} E_0 & 0 \\ -X_0 & I \end{bmatrix} \end{array}, \quad \mathscr{B}_0 = \begin{array}{c c} & \begin{array}{cc} n & n \end{array} \\ \begin{array}{c} n \\ n \end{array} & \begin{bmatrix} -Y_0 & I \\ F_0 & 0 \end{bmatrix} \end{array}.$$

Second standard form (SF2)

$$\text{(SF2)} \qquad \mathscr{A}_0 = \begin{array}{c} \\ n \\ n \end{array}\!\!\begin{array}{c} \begin{array}{cc} n & n \end{array} \\ \begin{bmatrix} E_0 & 0 \\ -X_0 & I \end{bmatrix} \end{array}, \quad \mathscr{B}_0 = \begin{array}{c} \\ n \\ n \end{array}\!\!\begin{array}{c} \begin{array}{cc} n & n \end{array} \\ \begin{bmatrix} -Y_0 & I \\ F_0 & 0 \end{bmatrix} \end{array}.$$

Essentially Gaussian elimination (assuming involved inversions exist):

$$[\mathscr{A}' \mid \mathscr{B}'] = \left[\begin{array}{cc|cc} X & X & X & X \\ X & X & X & X \end{array}\right] \to \left[\begin{array}{cc|cc} X & X & X & X \\ X & I & X & X \end{array}\right]$$
$$\to \left[\begin{array}{cc|cc} X & 0 & X & X \\ X & I & X & X \end{array}\right] \to \left[\begin{array}{cc|cc} X & 0 & X & I \\ X & I & X & X \end{array}\right]$$
$$\to \left[\begin{array}{cc|cc} X & 0 & X & I \\ X & I & X & 0 \end{array}\right]$$

in SF2

Outline

To keep in mind ...

To generate a sequence $\{\mathscr{A}_i - \lambda \mathscr{B}_i\}$ such that

- maintain respective standard form, and
- at the same time satisfy

$$\mathscr{A}_i \begin{bmatrix} I \\ X \end{bmatrix} = \mathscr{B}_i \begin{bmatrix} I \\ X \end{bmatrix} \mathscr{M}^{2^i} \quad \text{for } i = 0, 1, \ldots$$

When the spectral radius $\rho(\mathscr{M}) < 1$, $\{\mathscr{M}^{2^i}\}$ goes to 0 quadratically.

To keep in mind ...

To generate a sequence $\{\mathscr{A}_i - \lambda\mathscr{B}_i\}$ such that

- maintain respective standard form, and
- at the same time satisfy

$$\mathscr{A}_i \begin{bmatrix} I \\ X \end{bmatrix} = \mathscr{B}_i \begin{bmatrix} I \\ X \end{bmatrix} \mathscr{M}^{2^i} \quad \text{for } i = 0, 1, \ldots$$

When the spectral radius $\rho(\mathscr{M}) < 1$, $\{\mathscr{M}^{2^i}\}$ goes to 0 quadratically.

Guidance – doubling transformation theorem:

$$\operatorname{rank}([\mathscr{A}_\perp, \mathscr{B}_\perp]) = \ell, \quad [\mathscr{A}_\perp, \mathscr{B}_\perp] \begin{bmatrix} \mathscr{B} \\ -\mathscr{A} \end{bmatrix} = 0,$$

$$\widetilde{\mathscr{A}} := \mathscr{A}_\perp \mathscr{A}, \quad \widetilde{\mathscr{B}} := \mathscr{B}_\perp \mathscr{B},$$

$$\mathscr{A} Z = \mathscr{B} Z M \quad \Rightarrow \quad \widetilde{\mathscr{A}} Z = \widetilde{\mathscr{B}} Z M^2.$$

DA for SF1

To construct $\mathscr{A}_i - \lambda \mathscr{B}_i$ with

$$\mathscr{A}_i = \begin{array}{c c} & \begin{array}{cc} m & n \end{array} \\ \begin{array}{c} m \\ n \end{array} & \begin{bmatrix} E_i & 0 \\ -X_i & I \end{bmatrix} \end{array}, \quad \mathscr{B}_i = \begin{array}{c c} & \begin{array}{cc} m & n \end{array} \\ \begin{array}{c} m \\ n \end{array} & \begin{bmatrix} I & -Y_i \\ 0 & F_i \end{bmatrix} \end{array}.$$

Enough to know how to go from step i to step $i+1$.

DA for SF1

To construct $\mathscr{A}_i - \lambda \mathscr{B}_i$ with

$$\mathscr{A}_i = \begin{array}{c c} & \begin{array}{cc} m & n \end{array} \\ \begin{array}{c} m \\ n \end{array} & \begin{bmatrix} E_i & 0 \\ -X_i & I \end{bmatrix} \end{array}, \quad \mathscr{B}_i = \begin{array}{c c} & \begin{array}{cc} m & n \end{array} \\ \begin{array}{c} m \\ n \end{array} & \begin{bmatrix} I & -Y_i \\ 0 & F_i \end{bmatrix} \end{array}.$$

Enough to know how to go from step i to step $i+1$.

To find $\mathscr{A}_{i\perp}, \mathscr{B}_{i\perp} \in \mathbb{C}^{(m+n)\times(m+n)}$ such that

$$\text{rank}([\mathscr{A}_{i\perp}, \mathscr{B}_{i\perp}]) = m+n, \quad [\mathscr{A}_{i\perp}, \mathscr{B}_{i\perp}] \begin{bmatrix} \mathscr{B}_i \\ -\mathscr{A}_i \end{bmatrix} = 0.$$

DA for SF1

To construct $\mathscr{A}_i - \lambda \mathscr{B}_i$ with

$$\mathscr{A}_i = \begin{array}{c c} & \begin{array}{cc} m & n \end{array} \\ \begin{array}{c} m \\ n \end{array} & \begin{bmatrix} E_i & 0 \\ -X_i & I \end{bmatrix} \end{array}, \quad \mathscr{B}_i = \begin{array}{c c} & \begin{array}{cc} m & n \end{array} \\ \begin{array}{c} m \\ n \end{array} & \begin{bmatrix} I & -Y_i \\ 0 & F_i \end{bmatrix} \end{array}.$$

Enough to know how to go from step i to step $i+1$.

To find $\mathscr{A}_{i\perp}, \mathscr{B}_{i\perp} \in \mathbb{C}^{(m+n)\times(m+n)}$ such that

$$\operatorname{rank}([\mathscr{A}_{i\perp}, \mathscr{B}_{i\perp}]) = m+n, \quad [\mathscr{A}_{i\perp}, \mathscr{B}_{i\perp}] \begin{bmatrix} \mathscr{B}_i \\ -\mathscr{A}_i \end{bmatrix} = 0.$$

Can again perform Gaussian-like block eliminations

$$\begin{bmatrix} I_m & -Y_i \\ 0 & F_i \\ -E_i & 0 \\ X_i & -I_n \end{bmatrix} \xrightarrow{L_1} \begin{bmatrix} I_m & -Y_i \\ 0 & F_i \\ 0 & -E_i Y_i \\ 0 & -I_n + X_i Y_i \end{bmatrix} \xrightarrow{L_2} \begin{bmatrix} I_m & -Y_i \\ 0 & F_i \\ 0 & -E_i Y_i \\ 0 & -I_n \end{bmatrix} \xrightarrow{L_3} \begin{bmatrix} I_m & -Y_i \\ 0 & 0 \\ 0 & 0 \\ 0 & -I_n \end{bmatrix} \xrightarrow{L_4} \begin{bmatrix} I_m & -Y_i \\ 0 & -I_n \\ 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

DA for SF1 (cont'd)

Post-process the last $m+n$ rows of $L_4L_3L_2L_1$ to give

$$\mathscr{A}_{i\perp} = \begin{bmatrix} E_i(I_m - Y_iX_i)^{-1} & 0 \\ -F_i(I_n - X_iY_i)^{-1}X_i & I_n \end{bmatrix}, \quad \mathscr{B}_{i\perp} = \begin{bmatrix} I_m & -E_i(I_m - Y_iX_i)^{-1}Y_i \\ 0 & -F_i(I_n - X_iY_i)^{-1} \end{bmatrix}.$$

DA for SF1 (cont'd)

Post-process the last $m+n$ rows of $L_4L_3L_2L_1$ to give

$$\mathscr{A}_{i\perp} = \begin{bmatrix} E_i(I_m - Y_iX_i)^{-1} & 0 \\ -F_i(I_n - X_iY_i)^{-1}X_i & I_n \end{bmatrix}, \quad \mathscr{B}_{i\perp} = \begin{bmatrix} I_m & -E_i(I_m - Y_iX_i)^{-1}Y_i \\ 0 & -F_i(I_n - X_iY_i)^{-1} \end{bmatrix}.$$

Finally

$$\mathscr{A}_{i+1} = \mathscr{A}_{i\perp}\mathscr{A}_i, \quad \mathscr{B}_{i+1} = \mathscr{B}_{i\perp}\mathscr{B}_i$$

to give

$$E_{i+1} = E_i(I_m - Y_iX_i)^{-1}E_i, \tag{1a}$$

$$F_{i+1} = F_i(I_n - X_iY_i)^{-1}F_i, \tag{1b}$$

$$X_{i+1} = X_i + F_i(I_n - X_iY_i)^{-1}X_iE_i, \tag{1c}$$

$$Y_{i+1} = Y_i + E_i(I_m - Y_iX_i)^{-1}Y_iF_i. \tag{1d}$$

DA for SF1 (cont'd)

Post-process the last $m+n$ rows of $L_4L_3L_2L_1$ to give

$$\mathscr{A}_{i\perp} = \begin{bmatrix} E_i(I_m - Y_iX_i)^{-1} & 0 \\ -F_i(I_n - X_iY_i)^{-1}X_i & I_n \end{bmatrix}, \quad \mathscr{B}_{i\perp} = \begin{bmatrix} I_m & -E_i(I_m - Y_iX_i)^{-1}Y_i \\ 0 & -F_i(I_n - X_iY_i)^{-1} \end{bmatrix}.$$

Finally

$$\mathscr{A}_{i+1} = \mathscr{A}_{i\perp}\mathscr{A}_i, \quad \mathscr{B}_{i+1} = \mathscr{B}_{i\perp}\mathscr{B}_i$$

to give

$$E_{i+1} = E_i(I_m - Y_iX_i)^{-1}E_i, \tag{1a}$$

$$F_{i+1} = F_i(I_n - X_iY_i)^{-1}F_i, \tag{1b}$$

$$X_{i+1} = X_i + F_i(I_n - X_iY_i)^{-1}X_iE_i, \tag{1c}$$

$$Y_{i+1} = Y_i + E_i(I_m - Y_iX_i)^{-1}Y_iF_i. \tag{1d}$$

At convergence, X_i goes to some solution X.

DA for SF2

To construct $\mathscr{A}_i - \lambda\mathscr{B}_i$ with

$$\mathscr{A}_i = \begin{array}{c c} & \begin{array}{cc} n & n \end{array} \\ \begin{array}{c} n \\ n \end{array} & \begin{bmatrix} E_i & 0 \\ -X_i & I \end{bmatrix} \end{array}, \quad \mathscr{B}_i = \begin{array}{c c} & \begin{array}{cc} n & n \end{array} \\ \begin{array}{c} n \\ n \end{array} & \begin{bmatrix} -Y_i & I \\ F_i & 0 \end{bmatrix} \end{array}.$$

Enough to know how to go from step i to step $i+1$.

To construct $\mathscr{A}_i - \lambda \mathscr{B}_i$ with

$$\mathscr{A}_i = \begin{array}{c c} & \begin{array}{cc} n & n \end{array} \\ \begin{array}{c} n \\ n \end{array} & \begin{bmatrix} E_i & 0 \\ -X_i & I \end{bmatrix} \end{array}, \quad \mathscr{B}_i = \begin{array}{c c} & \begin{array}{cc} n & n \end{array} \\ \begin{array}{c} n \\ n \end{array} & \begin{bmatrix} -Y_i & I \\ F_i & 0 \end{bmatrix} \end{array}.$$

Enough to know how to go from step i to step $i+1$.

To find $\mathscr{A}_{i\perp}, \mathscr{B}_{i\perp} \in \mathbb{C}^{(m+n)\times(m+n)}$ such that

$$\operatorname{rank}([\mathscr{A}_{i\perp}, \mathscr{B}_{i\perp}]) = m+n, \quad [\mathscr{A}_{i\perp}, \mathscr{B}_{i\perp}] \begin{bmatrix} \mathscr{B}_i \\ -\mathscr{A}_i \end{bmatrix} = 0.$$

DA for SF2

To construct $\mathscr{A}_i - \lambda \mathscr{B}_i$ with

$$\mathscr{A}_i = \begin{array}{c c} & \begin{array}{cc} n & n \end{array} \\ \begin{array}{c} n \\ n \end{array} & \begin{bmatrix} E_i & 0 \\ -X_i & I \end{bmatrix} \end{array}, \quad \mathscr{B}_i = \begin{array}{c c} & \begin{array}{cc} n & n \end{array} \\ \begin{array}{c} n \\ n \end{array} & \begin{bmatrix} -Y_i & I \\ F_i & 0 \end{bmatrix} \end{array}.$$

Enough to know how to go from step i to step $i+1$.

To find $\mathscr{A}_{i\perp}, \mathscr{B}_{i\perp} \in \mathbb{C}^{(m+n)\times(m+n)}$ such that

$$\operatorname{rank}([\mathscr{A}_{i\perp}, \mathscr{B}_{i\perp}]) = m+n, \quad [\mathscr{A}_{i\perp}, \mathscr{B}_{i\perp}] \begin{bmatrix} \mathscr{B}_i \\ -\mathscr{A}_i \end{bmatrix} = 0.$$

Can again perform Gaussian-like block eliminations

$$\begin{bmatrix} -Y_i & I_n \\ F_i & 0 \\ E_i & 0 \\ -X_i & I_n \end{bmatrix} \xrightarrow{L_1} \begin{bmatrix} X_i - Y_i & 0 \\ F_i & 0 \\ E_i & 0 \\ -X_i & I_n \end{bmatrix} \xrightarrow{L_2} \begin{bmatrix} I_n & 0 \\ F_i & 0 \\ E_i & 0 \\ -X_i & I_n \end{bmatrix} \xrightarrow{L_3} \begin{bmatrix} I_n & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & I_n \end{bmatrix} \xrightarrow{L_4} \begin{bmatrix} I_n & 0 \\ 0 & I_n \\ 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

DA for SF2 (cont'd)

Post-process the last $2n$ rows of $L_4L_3L_2L_1$ to give

$$\mathscr{A}_{i\perp} = \begin{bmatrix} E_i(X_i - Y_i)^{-1} & 0 \\ -F_i(X_i - Y_i)^{-1} & I \end{bmatrix}, \quad \mathscr{B}_{i\perp} = \begin{bmatrix} I & E_i(X_i - Y_i)^{-1} \\ 0 & -F_i(X_i - Y_i)^{-1} \end{bmatrix}.$$

Post-process the last $2n$ rows of $L_4L_3L_2L_1$ to give

$$\mathscr{A}_{i\perp} = \begin{bmatrix} E_i(X_i - Y_i)^{-1} & 0 \\ -F_i(X_i - Y_i)^{-1} & I \end{bmatrix}, \quad \mathscr{B}_{i\perp} = \begin{bmatrix} I & E_i(X_i - Y_i)^{-1} \\ 0 & -F_i(X_i - Y_i)^{-1} \end{bmatrix}.$$

Finally

$$\mathscr{A}_{i+1} = \mathscr{A}_{i\perp}\mathscr{A}_i, \quad \mathscr{B}_{i+1} = \mathscr{B}_{i\perp}\mathscr{B}_i$$

to give

$$E_{i+1} = E_i(X_i - Y_i)^{-1}E_i, \tag{2a}$$
$$F_{i+1} = F_i(Y_i - X_i)^{-1}F_i, \tag{2b}$$
$$X_{i+1} = X_i + F_i(X_i - Y_i)^{-1}E_i, \tag{2c}$$
$$Y_{i+1} = Y_i + E_i(Y_i - X_i)^{-1}F_i. \tag{2d}$$

DA for SF2 (cont'd)

Post-process the last $2n$ rows of $L_4L_3L_2L_1$ to give

$$\mathscr{A}_{i\perp} = \begin{bmatrix} E_i(X_i - Y_i)^{-1} & 0 \\ -F_i(X_i - Y_i)^{-1} & I \end{bmatrix}, \quad \mathscr{B}_{i\perp} = \begin{bmatrix} I & E_i(X_i - Y_i)^{-1} \\ 0 & -F_i(X_i - Y_i)^{-1} \end{bmatrix}.$$

Finally

$$\mathscr{A}_{i+1} = \mathscr{A}_{i\perp}\mathscr{A}_i, \quad \mathscr{B}_{i+1} = \mathscr{B}_{i\perp}\mathscr{B}_i$$

to give

$$E_{i+1} = E_i(X_i - Y_i)^{-1}E_i, \tag{2a}$$
$$F_{i+1} = F_i(Y_i - X_i)^{-1}F_i, \tag{2b}$$
$$X_{i+1} = X_i + F_i(X_i - Y_i)^{-1}E_i, \tag{2c}$$
$$Y_{i+1} = Y_i + E_i(Y_i - X_i)^{-1}F_i. \tag{2d}$$

At convergence, X_i goes to some solution X.

Outline

Primary and Dual

In a DA, X_i, if it does, converges to a solution of "Primary equation". What about Y_i? "Dual" (to the primary equation) is the key for answering that.

In a DA, X_i, if it does, converges to a solution of "Primary equation". What about Y_i? "Dual" (to the primary equation) is the key for answering that.

Framework:

$$\text{(primary)} \quad \mathscr{A}_0 \begin{bmatrix} I \\ X \end{bmatrix} = \mathscr{B}_0 \begin{bmatrix} I \\ X \end{bmatrix} \mathscr{M},$$

Primary and Dual

In a DA, X_i, if it does, converges to a solution of "Primary equation". What about Y_i? "Dual" (to the primary equation) is the key for answering that.

Framework:

$$\begin{aligned} &\text{(primary)} \quad \mathscr{A}_0 \begin{bmatrix} I \\ X \end{bmatrix} = \mathscr{B}_0 \begin{bmatrix} I \\ X \end{bmatrix} \mathscr{M}, \\ &\text{(dual)} \qquad \hat{\mathscr{A}}_0 \begin{bmatrix} I \\ Y \end{bmatrix} = \hat{\mathscr{B}}_0 \begin{bmatrix} I \\ Y \end{bmatrix} \mathscr{N}. \end{aligned}$$

Primary and Dual

In a DA, X_i, if it does, converges to a solution of "Primary equation". What about Y_i? "Dual" (to the primary equation) is the key for answering that.

Framework:

$$\text{(primary)} \quad \mathscr{A}_0 \begin{bmatrix} I \\ X \end{bmatrix} = \mathscr{B}_0 \begin{bmatrix} I \\ X \end{bmatrix} \mathscr{M},$$

$$\text{(dual)} \quad \hat{\mathscr{A}}_0 \begin{bmatrix} I \\ Y \end{bmatrix} = \hat{\mathscr{B}}_0 \begin{bmatrix} I \\ Y \end{bmatrix} \mathscr{N}.$$

Let $\Pi_{m,n} = \begin{bmatrix} 0 & I_n \\ I_m & 0 \end{bmatrix}$.

Case: SF1

$$\text{(SF1)} \quad \mathscr{A}_0 = \begin{array}{c c} & \begin{array}{cc} m & n \end{array} \\ \begin{array}{c} m \\ n \end{array} & \begin{bmatrix} E_0 & 0 \\ -X_0 & I \end{bmatrix} \end{array}, \quad \mathscr{B}_0 = \begin{array}{c c} & \begin{array}{cc} m & n \end{array} \\ \begin{array}{c} m \\ n \end{array} & \begin{bmatrix} I & -Y_0 \\ 0 & F_0 \end{bmatrix} \end{array}.$$

Case: SF1

$$
\text{(SF1)} \quad \mathscr{A}_0 = \begin{array}{c c} & \begin{array}{cc} m & n \end{array} \\ \begin{array}{c} m \\ n \end{array} & \begin{bmatrix} E_0 & 0 \\ -X_0 & I \end{bmatrix} \end{array}, \quad \mathscr{B}_0 = \begin{array}{c c} & \begin{array}{cc} m & n \end{array} \\ \begin{array}{c} m \\ n \end{array} & \begin{bmatrix} I & -Y_0 \\ 0 & F_0 \end{bmatrix} \end{array}.
$$

Define

$$
\hat{\mathscr{A}}_0 := \Pi_{m,n}^T \mathscr{B}_0 \Pi_{m,n}, \;\; \hat{\mathscr{B}}_0 := \Pi_{m,n}^T \mathscr{A}_0 \Pi_{m,n}
$$

to give

$$
\hat{\mathscr{A}}_0 = \begin{array}{c c} & \begin{array}{cc} n & m \end{array} \\ \begin{array}{c} n \\ m \end{array} & \begin{bmatrix} F_0 & 0 \\ -Y_0 & I \end{bmatrix} \end{array}, \quad \hat{\mathscr{B}}_0 = \begin{array}{c c} & \begin{array}{cc} n & m \end{array} \\ \begin{array}{c} n \\ m \end{array} & \begin{bmatrix} I & -X_0 \\ 0 & E_0 \end{bmatrix} \end{array}.
$$

Case: SF1

$$
\text{(SF1)}\quad \mathscr{A}_0 = \begin{array}{c c} & \begin{array}{cc} m & n \end{array} \\ \begin{array}{c} m \\ n \end{array} & \left[\begin{array}{cc} E_0 & 0 \\ -X_0 & I \end{array}\right] \end{array}, \quad \mathscr{B}_0 = \begin{array}{c c} & \begin{array}{cc} m & n \end{array} \\ \begin{array}{c} m \\ n \end{array} & \left[\begin{array}{cc} I & -Y_0 \\ 0 & F_0 \end{array}\right] \end{array}.
$$

Define

$$
\hat{\mathscr{A}}_0 := \Pi_{m,n}^T \mathscr{B}_0 \Pi_{m,n}, \quad \hat{\mathscr{B}}_0 := \Pi_{m,n}^T \mathscr{A}_0 \Pi_{m,n}
$$

to give

$$
\hat{\mathscr{A}}_0 = \begin{array}{c c} & \begin{array}{cc} n & m \end{array} \\ \begin{array}{c} n \\ m \end{array} & \left[\begin{array}{cc} F_0 & 0 \\ -Y_0 & I \end{array}\right] \end{array}, \quad \hat{\mathscr{B}}_0 = \begin{array}{c c} & \begin{array}{cc} n & m \end{array} \\ \begin{array}{c} n \\ m \end{array} & \left[\begin{array}{cc} I & -X_0 \\ 0 & E_0 \end{array}\right] \end{array}.
$$

reciprocal relationship in eigenvalues

If (λ, z) is an eigenpair of $\mathscr{A}_0 - \lambda\mathscr{B}_0$, then $(1/\lambda, \Pi_{m,n}^T z)$ is an eigenpair of $\hat{\mathscr{A}}_0 - \lambda\hat{\mathscr{B}}_0$ and vice versa.

Case: SF1 (cont'd)

$$\text{(primary)} \quad \mathscr{A}_0 \begin{bmatrix} I \\ X \end{bmatrix} = \mathscr{B}_0 \begin{bmatrix} I \\ X \end{bmatrix} \mathscr{M}; \qquad \text{(dual)} \quad \hat{\mathscr{A}}_0 \begin{bmatrix} I \\ Y \end{bmatrix} = \hat{\mathscr{B}}_0 \begin{bmatrix} I \\ Y \end{bmatrix} \mathscr{N}.$$

Case: SF1 (cont'd)

$$\text{(primary)} \quad \mathscr{A}_0 \begin{bmatrix} I \\ X \end{bmatrix} = \mathscr{B}_0 \begin{bmatrix} I \\ X \end{bmatrix} \mathscr{M}; \qquad \text{(dual)} \quad \hat{\mathscr{A}}_0 \begin{bmatrix} I \\ Y \end{bmatrix} = \hat{\mathscr{B}}_0 \begin{bmatrix} I \\ Y \end{bmatrix} \mathscr{N}.$$

Plug in $\mathscr{A}_0$ and $\mathscr{B}_0$ to get

$$\begin{aligned} E_0 &= (I - Y_0 X)\mathscr{M}, \quad & X - X_0 &= F_0 X \mathscr{M}, \\ F_0 &= (I - X_0 Y)\mathscr{N}, \quad & Y - Y_0 &= E_0 Y \mathscr{N}. \end{aligned}$$

$$\text{(primary)} \quad \mathscr{A}_0 \begin{bmatrix} I \\ X \end{bmatrix} = \mathscr{B}_0 \begin{bmatrix} I \\ X \end{bmatrix} \mathscr{M}; \qquad \text{(dual)} \quad \hat{\mathscr{A}}_0 \begin{bmatrix} I \\ Y \end{bmatrix} = \hat{\mathscr{B}}_0 \begin{bmatrix} I \\ Y \end{bmatrix} \mathscr{N}.$$

Plug in $\mathscr{A}_0$ and $\mathscr{B}_0$ to get

$$\begin{aligned} E_0 &= (I - Y_0 X)\mathscr{M}, \quad & X - X_0 &= F_0 X \mathscr{M}, \\ F_0 &= (I - X_0 Y)\mathscr{N}, \quad & Y - Y_0 &= E_0 Y \mathscr{N}. \end{aligned}$$

Eliminate $\mathscr{M}$ and $\mathscr{N}$ to get

$$\begin{aligned} \text{primary:} \quad X &= X_0 + F_0 X (I - Y_0 X)^{-1} E_0, \\ \text{dual:} \quad Y &= Y_0 + E_0 Y (I - X_0 Y)^{-1} F_0. \end{aligned}$$

Case: SF2

$$\text{(SF2)} \quad \mathscr{A}_0 = \begin{array}{c c} & \begin{array}{cc} n & n \end{array} \\ \begin{array}{c} n \\ n \end{array} & \begin{bmatrix} E_0 & 0 \\ -X_0 & I \end{bmatrix} \end{array}, \quad \mathscr{B}_0 = \begin{array}{c c} & \begin{array}{cc} n & n \end{array} \\ \begin{array}{c} n \\ n \end{array} & \begin{bmatrix} -Y_0 & I \\ F_0 & 0 \end{bmatrix} \end{array}.$$

Case: SF2

$$\text{(SF2)} \quad \mathscr{A}_0 = \begin{array}{cc} & \begin{array}{cc} n & n \end{array} \\ \begin{array}{c} n \\ n \end{array} & \left[\begin{array}{cc} E_0 & 0 \\ -X_0 & I \end{array} \right] \end{array}, \quad \mathscr{B}_0 = \begin{array}{cc} & \begin{array}{cc} n & n \end{array} \\ \begin{array}{c} n \\ n \end{array} & \left[\begin{array}{cc} -Y_0 & I \\ F_0 & 0 \end{array} \right] \end{array}.$$

Define

$$\hat{\mathscr{A}}_0 := \Pi_{n,n}^T \mathscr{B}_0, \ \hat{\mathscr{B}}_0 := \Pi_{n,n}^T \mathscr{A}_0$$

to give

$$\hat{\mathscr{A}}_0 = \begin{array}{cc} & \begin{array}{cc} n & n \end{array} \\ \begin{array}{c} n \\ n \end{array} & \left[\begin{array}{cc} F_0 & 0 \\ -Y_0 & I \end{array} \right] \end{array}, \quad \hat{\mathscr{B}}_0 = \begin{array}{cc} & \begin{array}{cc} n & n \end{array} \\ \begin{array}{c} n \\ n \end{array} & \left[\begin{array}{cc} -X_0 & I \\ E_0 & 0 \end{array} \right] \end{array}.$$

Case: SF2

$$\text{(SF2)}\quad \mathscr{A}_0 = \begin{array}{c c} & \begin{array}{cc} n & n \end{array} \\ \begin{array}{c} n \\ n \end{array} & \left[\begin{array}{cc} E_0 & 0 \\ -X_0 & I \end{array}\right] \end{array}, \quad \mathscr{B}_0 = \begin{array}{c c} & \begin{array}{cc} n & n \end{array} \\ \begin{array}{c} n \\ n \end{array} & \left[\begin{array}{cc} -Y_0 & I \\ F_0 & 0 \end{array}\right] \end{array}.$$

Define

$$\hat{\mathscr{A}}_0 := \Pi_{n,n}^T \mathscr{B}_0, \;\; \hat{\mathscr{B}}_0 := \Pi_{n,n}^T \mathscr{A}_0$$

to give

$$\hat{\mathscr{A}}_0 = \begin{array}{c c} & \begin{array}{cc} n & n \end{array} \\ \begin{array}{c} n \\ n \end{array} & \left[\begin{array}{cc} F_0 & 0 \\ -Y_0 & I \end{array}\right] \end{array}, \quad \hat{\mathscr{B}}_0 = \begin{array}{c c} & \begin{array}{cc} n & n \end{array} \\ \begin{array}{c} n \\ n \end{array} & \left[\begin{array}{cc} -X_0 & I \\ E_0 & 0 \end{array}\right] \end{array}.$$

reciprocal relationship in eigenvalues

If (λ, z) is an eigenpair of $\mathscr{A}_0 - \lambda\mathscr{B}_0$, then $(1/\lambda, z)$ is an eigenpair of $\hat{\mathscr{A}}_0 - \lambda\hat{\mathscr{B}}_0$ and vice versa.

$$\text{(primary)} \quad \mathscr{A}_0 \begin{bmatrix} I \\ X \end{bmatrix} = \mathscr{B}_0 \begin{bmatrix} I \\ X \end{bmatrix} \mathscr{M}; \qquad \text{(dual)} \quad \hat{\mathscr{A}}_0 \begin{bmatrix} I \\ Y \end{bmatrix} = \hat{\mathscr{B}}_0 \begin{bmatrix} I \\ Y \end{bmatrix} \mathscr{N}.$$

Case: SF2 (cont'd)

$$\text{(primary)} \quad \mathscr{A}_0 \begin{bmatrix} I \\ X \end{bmatrix} = \mathscr{B}_0 \begin{bmatrix} I \\ X \end{bmatrix} \mathscr{M}; \qquad \text{(dual)} \quad \hat{\mathscr{A}}_0 \begin{bmatrix} I \\ Y \end{bmatrix} = \hat{\mathscr{B}}_0 \begin{bmatrix} I \\ Y \end{bmatrix} \mathscr{N}.$$

Plug in $\mathscr{A}_0$ and $\mathscr{B}_0$ to get

$$E_0 = (X - Y_0)\mathscr{M}, \quad X - X_0 = F_0\mathscr{M},$$
$$F_0 = (Y - X_0)\mathscr{N}, \quad Y - Y_0 = E_0\mathscr{N}.$$

$$\text{(primary)} \quad \mathscr{A}_0 \begin{bmatrix} I \\ X \end{bmatrix} = \mathscr{B}_0 \begin{bmatrix} I \\ X \end{bmatrix} \mathscr{M}; \qquad \text{(dual)} \quad \hat{\mathscr{A}}_0 \begin{bmatrix} I \\ Y \end{bmatrix} = \hat{\mathscr{B}}_0 \begin{bmatrix} I \\ Y \end{bmatrix} \mathscr{N}.$$

Plug in $\mathscr{A}_0$ and $\mathscr{B}_0$ to get

$$\begin{aligned} E_0 &= (X - Y_0)\mathscr{M}, \quad & X - X_0 &= F_0\mathscr{M}, \\ F_0 &= (Y - X_0)\mathscr{N}, \quad & Y - Y_0 &= E_0\mathscr{N}. \end{aligned}$$

Eliminate $\mathscr{M}$ and $\mathscr{N}$ to get

$$\begin{aligned} \text{primary:} \quad X &= X_0 + F_0(X - Y_0)^{-1}E_0, \\ \text{dual:} \quad Y &= Y_0 + E_0(Y - X_0)^{-1}F_0. \end{aligned}$$

Case: SF2 (cont'd)

$$\text{(primary)} \quad \mathscr{A}_0 \begin{bmatrix} I \\ X \end{bmatrix} = \mathscr{B}_0 \begin{bmatrix} I \\ X \end{bmatrix} \mathscr{M}; \qquad \text{(dual)} \quad \hat{\mathscr{A}}_0 \begin{bmatrix} I \\ Y \end{bmatrix} = \hat{\mathscr{B}}_0 \begin{bmatrix} I \\ Y \end{bmatrix} \mathscr{N}.$$

Plug in $\mathscr{A}_0$ and $\mathscr{B}_0$ to get

$$E_0 = (X - Y_0)\mathscr{M}, \quad X - X_0 = F_0\mathscr{M},$$
$$F_0 = (Y - X_0)\mathscr{N}, \quad Y - Y_0 = E_0\mathscr{N}.$$

Eliminate $\mathscr{M}$ and $\mathscr{N}$ to get

$$\text{primary:} \quad X = X_0 + F_0(X - Y_0)^{-1}E_0,$$
$$\text{dual:} \quad Y = Y_0 + E_0(Y - X_0)^{-1}F_0.$$

self-dual-ness

"primary" is "dual".

DA on Primary and Dual

What happen if DA is applied to $\mathscr{A}_0 - \lambda \mathscr{B}_0$, generating

$$X_i,\ Y_i,\ E_i,\ F_i$$

and to $\hat{\mathscr{A}}_0 - \lambda \hat{\mathscr{B}}_0$, generating

$$\hat{X}_i,\ \hat{Y}_i,\ \hat{E}_i,\ \hat{F}_i?$$

DA on Primary and Dual

What happen if DA is applied to $\mathscr{A}_0 - \lambda \mathscr{B}_0$, generating

$$X_i,\ Y_i,\ E_i,\ F_i$$

and to $\hat{\mathscr{A}}_0 - \lambda \hat{\mathscr{B}}_0$, generating

$$\hat{X}_i,\ \hat{Y}_i,\ \hat{E}_i,\ \hat{F}_i?$$

DA on primary and dual

For all $i \geq 0$,

$$\hat{X}_i = Y_i, \quad \hat{Y}_i = X_i,$$
$$\hat{E}_i = F_i, \quad \hat{F}_i = E_i.$$

Outline

General

When a DA is applied to both $\mathscr{A}_0 - \lambda\mathscr{B}_0$ and $\hat{\mathscr{A}}_0 - \lambda\hat{\mathscr{B}}_0$, two sequences $\{\mathscr{A}_i - \lambda\mathscr{B}_i\}_{i=0}^{\infty}$ and $\{\hat{\mathscr{A}}_i - \lambda\hat{\mathscr{B}}_i\}_{i=0}^{\infty}$ are produced, assuming no breakdown occurs.

General

When a DA is applied to both $\mathscr{A}_0 - \lambda\mathscr{B}_0$ and $\hat{\mathscr{A}}_0 - \lambda\hat{\mathscr{B}}_0$, two sequences $\{\mathscr{A}_i - \lambda\mathscr{B}_i\}_{i=0}^{\infty}$ and $\{\hat{\mathscr{A}}_i - \lambda\hat{\mathscr{B}}_i\}_{i=0}^{\infty}$ are produced, assuming no breakdown occurs.

The sequences are essentially “identical”. Moreover,

$$\mathscr{A}_i \begin{bmatrix} I \\ X \end{bmatrix} = \mathscr{B}_i \begin{bmatrix} I \\ X \end{bmatrix} \mathscr{M}^{2^i}, \quad \hat{\mathscr{A}}_i \begin{bmatrix} I \\ Y \end{bmatrix} = \hat{\mathscr{B}}_i \begin{bmatrix} I \\ Y \end{bmatrix} \mathscr{N}^{2^i} \quad \text{for } i = 0, 1, \ldots.$$

General

When a DA is applied to both $\mathscr{A}_0 - \lambda\mathscr{B}_0$ and $\hat{\mathscr{A}}_0 - \lambda\hat{\mathscr{B}}_0$, two sequences $\{\mathscr{A}_i - \lambda\mathscr{B}_i\}_{i=0}^{\infty}$ and $\{\hat{\mathscr{A}}_i - \lambda\hat{\mathscr{B}}_i\}_{i=0}^{\infty}$ are produced, assuming no breakdown occurs.

The sequences are essentially "identical". Moreover,

$$\mathscr{A}_i \begin{bmatrix} I \\ X \end{bmatrix} = \mathscr{B}_i \begin{bmatrix} I \\ X \end{bmatrix} \mathscr{M}^{2^i}, \quad \hat{\mathscr{A}}_i \begin{bmatrix} I \\ Y \end{bmatrix} = \hat{\mathscr{B}}_i \begin{bmatrix} I \\ Y \end{bmatrix} \mathscr{N}^{2^i} \quad \text{for } i = 0, 1, \ldots.$$

Certain conditions on $\rho(\mathscr{M})$ and $\rho(\mathscr{N})$ should be imposed.

General

When a DA is applied to both $\mathscr{A}_0 - \lambda\mathscr{B}_0$ and $\hat{\mathscr{A}}_0 - \lambda\hat{\mathscr{B}}_0$, two sequences $\{\mathscr{A}_i - \lambda\mathscr{B}_i\}_{i=0}^{\infty}$ and $\{\hat{\mathscr{A}}_i - \lambda\hat{\mathscr{B}}_i\}_{i=0}^{\infty}$ are produced, assuming no breakdown occurs.

The sequences are essentially "identical". Moreover,

$$\mathscr{A}_i \begin{bmatrix} I \\ X \end{bmatrix} = \mathscr{B}_i \begin{bmatrix} I \\ X \end{bmatrix} \mathscr{M}^{2^i}, \quad \hat{\mathscr{A}}_i \begin{bmatrix} I \\ Y \end{bmatrix} = \hat{\mathscr{B}}_i \begin{bmatrix} I \\ Y \end{bmatrix} \mathscr{N}^{2^i} \quad \text{for } i = 0, 1, \ldots.$$

Certain conditions on $\rho(\mathscr{M})$ and $\rho(\mathscr{N})$ should be imposed.

DA computes special solutions $X = \Phi$ and $Y = \Psi$ to Primary and Dual.

Case: SF1

$$\mathscr{A}_i = \begin{array}{c c} & \begin{array}{cc} m & n \end{array} \\ \begin{array}{c} m \\ n \end{array} & \begin{bmatrix} E_i & 0 \\ -X_i & I \end{bmatrix} \end{array}, \quad \mathscr{B}_i = \begin{array}{c c} & \begin{array}{cc} m & n \end{array} \\ \begin{array}{c} m \\ n \end{array} & \begin{bmatrix} I & -Y_i \\ 0 & F_i \end{bmatrix} \end{array},$$

$$\hat{\mathscr{A}}_i = \begin{array}{c c} & \begin{array}{cc} n & m \end{array} \\ \begin{array}{c} n \\ m \end{array} & \begin{bmatrix} F_i & 0 \\ -Y_i & I \end{bmatrix} \end{array}, \quad \hat{\mathscr{B}}_i = \begin{array}{c c} & \begin{array}{cc} n & m \end{array} \\ \begin{array}{c} n \\ m \end{array} & \begin{bmatrix} I & -X_i \\ 0 & E_i \end{bmatrix} \end{array},$$

Case: SF1

$$\mathscr{A}_i = \begin{array}{c c} & \begin{array}{cc} m & n \end{array} \\ \begin{array}{c} m \\ n \end{array} & \begin{bmatrix} E_i & 0 \\ -X_i & I \end{bmatrix} \end{array}, \quad \mathscr{B}_i = \begin{array}{c c} & \begin{array}{cc} m & n \end{array} \\ \begin{array}{c} m \\ n \end{array} & \begin{bmatrix} I & -Y_i \\ 0 & F_i \end{bmatrix} \end{array},$$

$$\hat{\mathscr{A}}_i = \begin{array}{c c} & \begin{array}{cc} n & m \end{array} \\ \begin{array}{c} n \\ m \end{array} & \begin{bmatrix} F_i & 0 \\ -Y_i & I \end{bmatrix} \end{array}, \quad \hat{\mathscr{B}}_i = \begin{array}{c c} & \begin{array}{cc} n & m \end{array} \\ \begin{array}{c} n \\ m \end{array} & \begin{bmatrix} I & -X_i \\ 0 & E_i \end{bmatrix} \end{array},$$

$$E_i = (I - Y_i \Phi)\mathscr{M}^{2^i}, \quad \Phi - X_i = F_i \Phi \mathscr{M}^{2^i},$$

$$F_i = (I - X_i \Psi)\mathscr{N}^{2^i}, \quad \Psi - Y_i = E_i \Psi \mathscr{N}^{2^i}.$$

Case: SF1

$$\mathscr{A}_i = \begin{array}{c c} & \begin{array}{cc} m & n \end{array} \\ \begin{array}{c} m \\ n \end{array} & \left[\begin{array}{cc} E_i & 0 \\ -X_i & I \end{array}\right] \end{array}, \quad \mathscr{B}_i = \begin{array}{c c} & \begin{array}{cc} m & n \end{array} \\ \begin{array}{c} m \\ n \end{array} & \left[\begin{array}{cc} I & -Y_i \\ 0 & F_i \end{array}\right] \end{array},$$

$$\hat{\mathscr{A}}_i = \begin{array}{c c} & \begin{array}{cc} n & m \end{array} \\ \begin{array}{c} n \\ m \end{array} & \left[\begin{array}{cc} F_i & 0 \\ -Y_i & I \end{array}\right] \end{array}, \quad \hat{\mathscr{B}}_i = \begin{array}{c c} & \begin{array}{cc} n & m \end{array} \\ \begin{array}{c} n \\ m \end{array} & \left[\begin{array}{cc} I & -X_i \\ 0 & E_i \end{array}\right] \end{array},$$

$$E_i = (I - Y_i\Phi)\mathscr{M}^{2^i}, \quad \Phi - X_i = F_i\Phi\mathscr{M}^{2^i},$$
$$F_i = (I - X_i\Psi)\mathscr{N}^{2^i}, \quad \Psi - Y_i = E_i\Psi\mathscr{N}^{2^i}.$$

Convergence Theorem

Suppose that there are solutions $X = \Phi$ and $Y = \Psi$ to Primary and Dual such that $\rho(\mathscr{M}) \cdot \rho(\mathscr{N}) < 1$, and suppose that DA executes without any breakdown. Then X_i and Y_i converge to Φ and Ψ quadratically, and moreover,

$$\limsup_{i\to\infty} \|X_i - \Phi\|^{1/2^i} \le \rho(\mathscr{M}) \cdot \rho(\mathscr{N}), \quad \limsup_{i\to\infty} \|Y_i - \Psi\|^{1/2^i} \le \rho(\mathscr{M}) \cdot \rho(\mathscr{N}).$$

Case: SF2

$$\mathscr{A}_i = \begin{array}{c c} & \begin{array}{cc} n & n \end{array} \\ \begin{array}{c} n \\ n \end{array} & \begin{bmatrix} E_i & 0 \\ -X_i & I \end{bmatrix} \end{array}, \quad \mathscr{B}_i = \begin{array}{c c} & \begin{array}{cc} n & n \end{array} \\ \begin{array}{c} n \\ n \end{array} & \begin{bmatrix} -Y_i & I \\ F_i & 0 \end{bmatrix} \end{array},$$

$$\hat{\mathscr{A}}_i = \begin{array}{c c} & \begin{array}{cc} n & n \end{array} \\ \begin{array}{c} n \\ n \end{array} & \begin{bmatrix} F_i & 0 \\ -Y_i & I \end{bmatrix} \end{array}, \quad \hat{\mathscr{B}}_i = \begin{array}{c c} & \begin{array}{cc} n & n \end{array} \\ \begin{array}{c} n \\ n \end{array} & \begin{bmatrix} -X_i & I \\ E_i & 0 \end{bmatrix} \end{array},$$

Case: SF2

$$\mathscr{A}_i = \begin{array}{c c} & \begin{array}{cc} n & n \end{array} \\ \begin{array}{c} n \\ n \end{array} & \left[\begin{array}{cc} E_i & 0 \\ -X_i & I \end{array}\right] \end{array}, \quad \mathscr{B}_i = \begin{array}{c c} & \begin{array}{cc} n & n \end{array} \\ \begin{array}{c} n \\ n \end{array} & \left[\begin{array}{cc} -Y_i & I \\ F_i & 0 \end{array}\right] \end{array},$$

$$\hat{\mathscr{A}}_i = \begin{array}{c c} & \begin{array}{cc} n & n \end{array} \\ \begin{array}{c} n \\ n \end{array} & \left[\begin{array}{cc} F_i & 0 \\ -Y_i & I \end{array}\right] \end{array}, \quad \hat{\mathscr{B}}_i = \begin{array}{c c} & \begin{array}{cc} n & n \end{array} \\ \begin{array}{c} n \\ n \end{array} & \left[\begin{array}{cc} -X_i & I \\ E_i & 0 \end{array}\right] \end{array},$$

$$E_i = (\Phi - Y_i)\mathscr{M}^{2^i}, \quad \Phi - X_i = F_i \mathscr{M}^{2^i},$$

$$F_i = (\Psi - X_i)\mathscr{N}^{2^i}, \quad \Psi - Y_i = E_i \mathscr{N}^{2^i}.$$

Case: SF2

$$\mathscr{A}_i = \begin{array}{c c} & \begin{array}{cc} n & n \end{array} \\ \begin{array}{c} n \\ n \end{array} & \begin{bmatrix} E_i & 0 \\ -X_i & I \end{bmatrix} \end{array}, \quad \mathscr{B}_i = \begin{array}{c c} & \begin{array}{cc} n & n \end{array} \\ \begin{array}{c} n \\ n \end{array} & \begin{bmatrix} -Y_i & I \\ F_i & 0 \end{bmatrix} \end{array},$$

$$\hat{\mathscr{A}}_i = \begin{array}{c c} & \begin{array}{cc} n & n \end{array} \\ \begin{array}{c} n \\ n \end{array} & \begin{bmatrix} F_i & 0 \\ -Y_i & I \end{bmatrix} \end{array}, \quad \hat{\mathscr{B}}_i = \begin{array}{c c} & \begin{array}{cc} n & n \end{array} \\ \begin{array}{c} n \\ n \end{array} & \begin{bmatrix} -X_i & I \\ E_i & 0 \end{bmatrix} \end{array},$$

$$E_i = (\Phi - Y_i)\mathscr{M}^{2^i}, \quad \Phi - X_i = F_i\mathscr{M}^{2^i},$$
$$F_i = (\Psi - X_i)\mathscr{N}^{2^i}, \quad \Psi - Y_i = E_i\mathscr{N}^{2^i}.$$

Convergence Theorem

Suppose that there are solutions $X = \Phi$ and $Y = \Psi$ to Primary and Dual such that $\rho(\mathscr{M}) \cdot \rho(\mathscr{N}) < 1$, and suppose that DA executes without any breakdown. Then X_i and Y_i converge to Φ and Ψ quadratically, and moreover,

$$\limsup_{i\to\infty} \|X_i - \Phi\|^{1/2^i} \le \rho(\mathscr{M}) \cdot \rho(\mathscr{N}), \quad \limsup_{i\to\infty} \|Y_i - \Psi\|^{1/2^i} \le \rho(\mathscr{M}) \cdot \rho(\mathscr{N}).$$

Outline

critical case

What if $\rho(\mathscr{M}) \cdot \rho(\mathscr{N}) = 1$, the critical case?

What if $\rho(\mathscr{M}) \cdot \rho(\mathscr{N}) = 1$, the critical case?

Complicated but still possible to have linear convergence at the respectable rate $1/2$.

critical case

What if $\rho(\mathscr{M}) \cdot \rho(\mathscr{N}) = 1$, the critical case?

Complicated but still possible to have linear convergence at the respectable rate $1/2$.

More to come ...

Outline

CARE: $-XGX + A^{\mathrm{H}}X + XA + H = 0$

$$\mathscr{A} - \lambda\mathscr{B} := \mathscr{H} - \lambda I_{2n} \equiv \begin{bmatrix} A & -G \\ -H & -A^{\mathrm{H}} \end{bmatrix} - \lambda I_{2n},$$

$$\begin{bmatrix} \mathscr{A}' \\ \mathscr{B}' \end{bmatrix} = \begin{bmatrix} I_{2n} & -\gamma I_{2n} \\ I_{2n} & \gamma I_{2n} \end{bmatrix} \begin{bmatrix} \mathscr{H} \\ I_{2n} \end{bmatrix} \qquad (\gamma > 0),$$

$P(\mathscr{A}' - \lambda\mathscr{B}') = \mathscr{A}_0 - \lambda\mathscr{B}_0$ in SF1 and then apply DA.

(a) All $I - X_iY_i$ and $I - Y_iX_i$ are nonsingular;

(b) $0 \preceq X_0 \preceq X_i \preceq X_{i+1} \preceq \Phi$, $0 \succeq Y_0 \succeq Y_i \succeq Y_{i+1} \succeq \Psi$ and

$$\limsup_{i\to\infty} \|\Phi - X_i\|^{1/2^i} \le \rho(\mathscr{M})^2, \quad \limsup_{i\to\infty} \|\Psi - Y_i\|^{1/2^i} \le \rho(\mathscr{N})^2,$$

where

$\rho(\mathscr{M}) = \rho((I - Y_0\Phi)^{-1}E_0) = \rho(\mathscr{N}) = \rho((I - X_0\Psi)^{-1}E_0^{\mathrm{H}})$;

(c) $\mathrm{eig}(A - G\Phi) \subset \mathbb{C}_-$, $\mathrm{eig}(-H\Psi - A^{\mathrm{H}}) \subset \mathbb{C}_+$, and

$$\mathrm{eig}(\mathscr{H}) = \mathrm{eig}(A - G\Phi) \cup \mathrm{eig}(-H\Psi - A^{\mathrm{H}}).$$

Assume

$$\boxed{W = \begin{bmatrix} B & -D \\ -C & A \end{bmatrix} \text{ is a nonsingular or an irreducible singular } M\text{-matrix.}} \tag{3}$$

MARE with (3) *has a unique minimal nonnegative solution* Φ, *i.e.,*

$$0 \le \Phi \le X \quad \textit{for any other nonnegative solution } X\,.$$

Known in applied probability, stochastic fluid models, but rigorous matrix proof by C. Guo (2000):

Assume

$$\boxed{W = \begin{bmatrix} B & -D \\ -C & A \end{bmatrix} \text{ is a nonsingular or an irreducible singular } M\text{-matrix.}} \tag{3}$$

MARE with (3) *has a unique minimal nonnegative solution* Φ, *i.e.,*

$$0 \le \Phi \le X \quad \textit{for any other nonnegative solution } X\,.$$

Known in applied probability, stochastic fluid models, but rigorous matrix proof by C. Guo (2000):

$$\mathscr{A} - \lambda\mathscr{B} := \mathscr{H} - \lambda I_{m+n} \equiv \begin{bmatrix} B & -D \\ C & -A^{\mathrm{H}} \end{bmatrix} - \lambda I_{m+n},$$

$$\begin{bmatrix} \mathscr{A}' \\ \mathscr{B}' \end{bmatrix} = \begin{bmatrix} I_{m+n} & -\beta I_{m+n} \\ I_{m+n} & \alpha I_{m+n} \end{bmatrix} \begin{bmatrix} \mathscr{H} \\ I_{m+n} \end{bmatrix} \qquad (\alpha \geq \max_i A_{ii},\ \beta \geq \max_i B_{ii}),$$

$P(\mathscr{A}' - \lambda\mathscr{B}') = \mathscr{A}_0 - \lambda\mathscr{B}_0$ in SF1 and then apply DA.

(a) All $I - X_i Y_i$ and $I - Y_i X_i$ are nonsingular M-matrices for all $i \geq 0$;

(b) $0 \leq X_0 \leq X_i \leq X_{i+1} \leq \Phi$, $0 \leq Y_0 \leq Y_i \leq Y_{i+1} \leq \Psi$ and

$$\limsup_{i\to\infty} \|\Phi - X_i\|^{1/2^i} \leq \rho(\mathscr{M})^2, \quad \limsup_{i\to\infty} \|\Psi - Y_i\|^{1/2^i} \leq \rho(\mathscr{N})^2,$$

where
$\rho(\mathscr{M}) = \rho((I - Y_0\Phi)^{-1}E_0) = \rho(\mathscr{N}) = \rho((I - X_0\Psi)^{-1}E_0^{\mathrm{H}})$;

(c) $\mathrm{eig}(B - D\Phi) \subset \mathbb{C}_-$, $\mathrm{eig}(-A - \Phi D) \subset \mathbb{C}_-$, and

$$\mathrm{eig}(\mathscr{H}) = \mathrm{eig}(B - D\Phi) \cup \mathrm{eig}(-A - \Phi D).$$

Moreover ...

Fast Train Eigenvalue Problem

$$P(\lambda)z \equiv (\lambda^2 A^\mathsf{T} + \lambda Q + A)z = 0.$$

Fast Train Eigenvalue Problem

$$P(\lambda)z \equiv (\lambda^2 A^\mathsf{T} + \lambda Q + A)z = 0.$$

Framework of solvent approach:

1. Compute the stabilizing solution Φ of the matrix equation $X + A^\mathsf{T} X^{-1} A = Q$. Then

$$P(\lambda) = \lambda^2 A^\mathsf{T} + \lambda Q + A = (\lambda A^\mathsf{T} + X)X^{-1}(\lambda X + A).$$

2. Solve the (linear) eigenvalue problems for matrix pencils $\lambda A^\mathsf{T} + X$ and $\lambda X + A$.

Fast Train Eigenvalue Problem

$$
Q = \begin{array}{cc}
 & \begin{array}{ccccc} k & k & k & \cdots & k \end{array} \\
\begin{array}{c} k \\ k \\ k \\ \vdots \\ k \end{array} &
\begin{bmatrix}
H_0 & H_1^T & & & \\
H_1 & H_0 & H_1^T & & \\
 & H_1 & \ddots & \ddots & \\
 & & \ddots & \ddots & H_1^T \\
 & & & H_1 & H_0
\end{bmatrix}
\end{array},
\quad
A = \begin{array}{cc}
 & \begin{array}{cccc} k & \cdots & k & k \end{array} \\
\begin{array}{c} k \\ k \\ \vdots \\ k \end{array} &
\begin{bmatrix}
0 & \dots & 0 & H_1 \\
0 & \dots & 0 & 0 \\
\vdots & & \vdots & \vdots \\
0 & \dots & 0 & 0
\end{bmatrix}
\end{array}.
$$

Fast Train Eigenvalue Problem

$$Q = \begin{array}{c} \\ k \\ k \\ k \\ \vdots \\ k \end{array} \begin{array}{c} \begin{array}{ccccc} k & k & k & \cdots & k \end{array} \\ \begin{bmatrix} H_0 & H_1^T & & & \\ H_1 & H_0 & H_1^T & & \\ & H_1 & \ddots & \ddots & \\ & & \ddots & \ddots & H_1^T \\ & & & H_1 & H_0 \end{bmatrix} \end{array}, \quad A = \begin{array}{c} \\ k \\ k \\ \vdots \\ k \end{array} \begin{array}{c} \begin{array}{cccc} k & \cdots & k & k \end{array} \\ \begin{bmatrix} 0 & \dots & 0 & H_1 \\ 0 & \dots & 0 & 0 \\ \vdots & & \vdots & \vdots \\ 0 & \dots & 0 & 0 \end{bmatrix} \end{array}.$$

Numerous notorious numerical difficulties:

- most eigenvalues are 0 and ∞, $2(m-1)k$ in all;
- problem size n can range from 10^3 to 10^5
- most seriously, badly scaled finite eigenvalues: $10^{-50} \sim 10^{50}$ or to an even greater extreme
- all finite nonzero eigenvalues and eigenvectors are to be computed

Fast Train Eigenvalue Problem

$$
Q = \begin{array}{c} \\ k \\ k \\ k \\ \vdots \\ k \end{array}
\begin{array}{c}
\begin{array}{ccccc} k & k & k & \cdots & k \end{array} \\
\begin{bmatrix}
H_0 & H_1^T & & & \\
H_1 & H_0 & H_1^T & & \\
 & H_1 & \ddots & \ddots & \\
 & & \ddots & \ddots & H_1^T \\
 & & & H_1 & H_0
\end{bmatrix}
\end{array},
\quad
A = \begin{array}{c} \\ k \\ k \\ \vdots \\ k \end{array}
\begin{array}{c}
\begin{array}{cccc} k & \cdots & k & k \end{array} \\
\begin{bmatrix}
0 & \dots & 0 & H_1 \\
0 & \dots & 0 & 0 \\
\vdots & & \vdots & \vdots \\
0 & \dots & 0 & 0
\end{bmatrix}
\end{array}.
$$

Numerous notorious numerical difficulties:

- most eigenvalues are 0 and ∞, $2(m-1)k$ in all;
- problem size n can range from 10^3 to 10^5
- most seriously, badly scaled finite eigenvalues: $10^{-50} \sim 10^{50}$ or to an even greater extreme
- all finite nonzero eigenvalues and eigenvectors are to be computed

Cleverly implemented solvent approach based on DA gets it done!

Outline

MARE: $XDX - AX - XB + C = 0$

$$W = \begin{array}{cc} & \begin{array}{cc} m & n \end{array} \\ \begin{array}{c} m \\ n \end{array} & \begin{bmatrix} B & -D \\ -C & A \end{bmatrix} \end{array},$$

and W is a nonsingular or an irreducible singular M-matrix.

MARE: $XDX - AX - XB + C = 0$

$$W = \begin{array}{c c} & \begin{array}{cc} m & n \end{array} \\ \begin{array}{c} m \\ n \end{array} & \begin{bmatrix} B & -D \\ -C & A \end{bmatrix} \end{array},$$

and W is a nonsingular or an irreducible singular M-matrix.

$$B = \begin{bmatrix} 3 & -1 & & \\ & 3 & \ddots & \\ & & \ddots & -1 \\ -1 & & & 3 \end{bmatrix} \in \mathbb{R}^{n \times n}, \; C = 2I_n, \; A = \xi B, \; D = \xi C,$$

where $\xi > 0$ is a parameter.

$$W = \begin{array}{c c} & \begin{array}{cc} m & n \end{array} \\ \begin{array}{c} m \\ n \end{array} & \begin{bmatrix} B & -D \\ -C & A \end{bmatrix} \end{array},$$

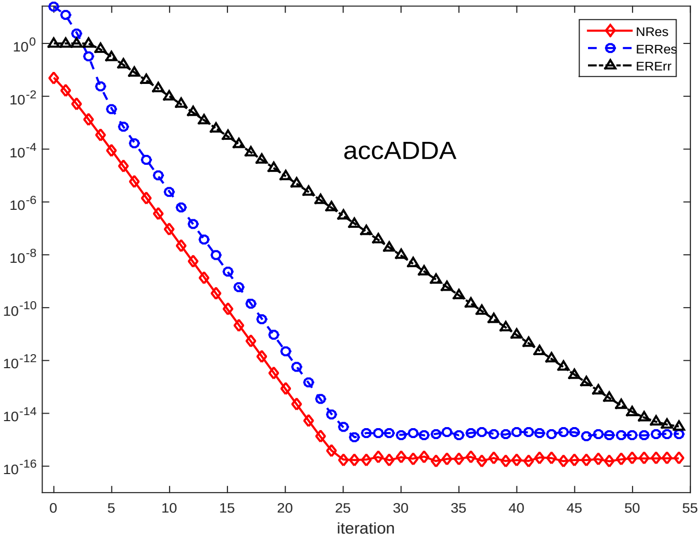
and W is a nonsingular or an irreducible singular M-matrix.

$$B = \begin{bmatrix} 3 & -1 & & \\ & 3 & \ddots & \\ & & \ddots & -1 \\ -1 & & & 3 \end{bmatrix} \in \mathbb{R}^{n\times n},\ C = 2I_n,\ A = \xi B,\ D = \xi C,$$

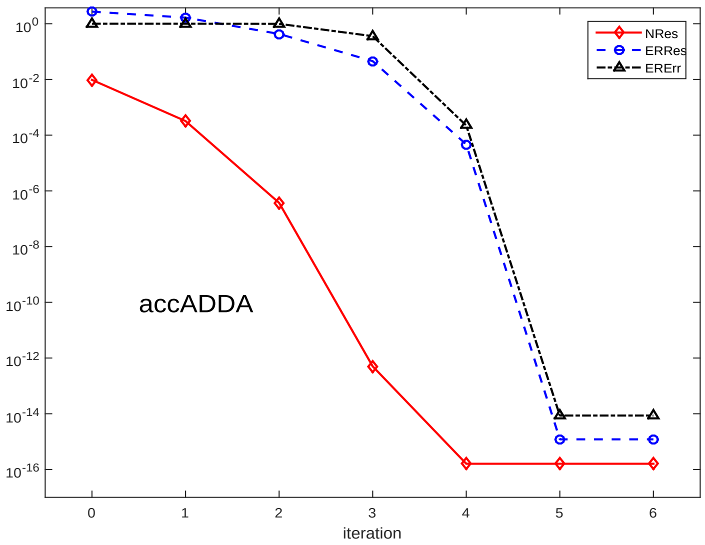
where $\xi > 0$ is a parameter.

Quadratic convergence for $\xi \neq 1$ and linear convergence otherwise.

MARE $\xi = 1$ (cont'd)



MARE $\xi = 10^4$ (cont'd)



Fast Train Eigenvalue Problem

$$P(\lambda)z \equiv (\lambda^2 A^{\mathsf{T}} + \lambda Q + A)z = 0, \quad \text{with}$$

$$Q = \begin{array}{c c} & \begin{array}{ccccc} k & k & k & \cdots & k \end{array} \\ \begin{array}{c} k \\ k \\ k \\ \vdots \\ k \end{array} & \begin{bmatrix} H_0 & H_1^{\mathsf{T}} & & & \\ H_1 & H_0 & H_1^{\mathsf{T}} & & \\ & H_1 & \ddots & \ddots & \\ & & \ddots & \ddots & H_1^{\mathsf{T}} \\ & & & H_1 & H_0 \end{bmatrix} \end{array}, \quad A = \begin{array}{c c} & \begin{array}{cccc} k & \cdots & k & k \end{array} \\ \begin{array}{c} k \\ k \\ \vdots \\ k \end{array} & \begin{bmatrix} 0 & \dots & 0 & H_1 \\ 0 & \dots & 0 & 0 \\ \vdots & & \vdots & \vdots \\ 0 & \dots & 0 & 0 \end{bmatrix} \end{array}.$$

Fast Train Eigenvalue Problem

$$P(\lambda)z \equiv (\lambda^2 A^{\mathsf{T}} + \lambda Q + A)z = 0, \quad \text{with}$$

$$Q = \begin{array}{c} \\ k \\ k \\ k \\ \vdots \\ k \end{array} \begin{array}{c} \begin{array}{ccccc} k & k & k & \cdots & k \end{array} \\ \begin{bmatrix} H_0 & H_1^{\mathsf{T}} & & & \\ H_1 & H_0 & H_1^{\mathsf{T}} & & \\ & H_1 & \ddots & \ddots & \\ & & \ddots & \ddots & H_1^{\mathsf{T}} \\ & & & H_1 & H_0 \end{bmatrix} \end{array}, \quad A = \begin{array}{c} \\ k \\ k \\ \vdots \\ k \end{array} \begin{array}{c} \begin{array}{cccc} k & \cdots & k & k \end{array} \\ \begin{bmatrix} 0 & \dots & 0 & H_1 \\ 0 & \dots & 0 & 0 \\ \vdots & & \vdots & \vdots \\ 0 & \dots & 0 & 0 \end{bmatrix} \end{array}.$$

Can prove that it is equivalent to
$\widehat{P}(\hat{\lambda})y := (\hat{\lambda}^2 H_1^{\mathsf{T}} + \hat{\lambda} H_0 + H_1)y = 0$: $\lambda = \hat{\lambda}^m$.

Fast Train Eigenvalue Problem

$$P(\lambda)z \equiv (\lambda^2 A^\mathsf{T} + \lambda Q + A)z = 0, \quad \text{with}$$

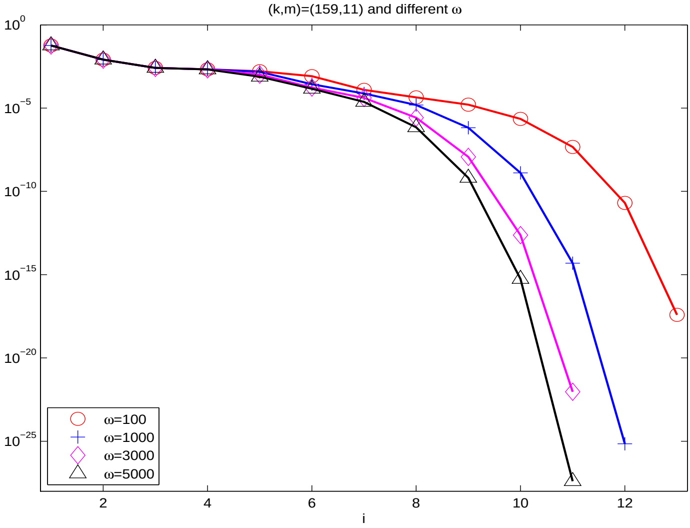
$$Q = \begin{array}{c c} & \begin{array}{ccccc} k & k & k & \cdots & k \end{array} \\ \begin{array}{c} k \\ k \\ k \\ \vdots \\ k \end{array} & \begin{bmatrix} H_0 & H_1^\mathsf{T} & & & \\ H_1 & H_0 & H_1^\mathsf{T} & & \\ & H_1 & \ddots & \ddots & \\ & & \ddots & \ddots & H_1^\mathsf{T} \\ & & & H_1 & H_0 \end{bmatrix} \end{array}, \quad A = \begin{array}{c c} & \begin{array}{cccc} k & \cdots & k & k \end{array} \\ \begin{array}{c} k \\ k \\ \vdots \\ k \end{array} & \begin{bmatrix} 0 & \dots & 0 & H_1 \\ 0 & \dots & 0 & 0 \\ \vdots & & \vdots & \vdots \\ 0 & \dots & 0 & 0 \end{bmatrix} \end{array}.$$

Can prove that it is equivalent to
$\widehat{P}(\hat{\lambda})y := (\hat{\lambda}^2 H_1^\mathsf{T} + \hat{\lambda} H_0 + H_1)y = 0$: $\lambda = \hat{\lambda}^m$.

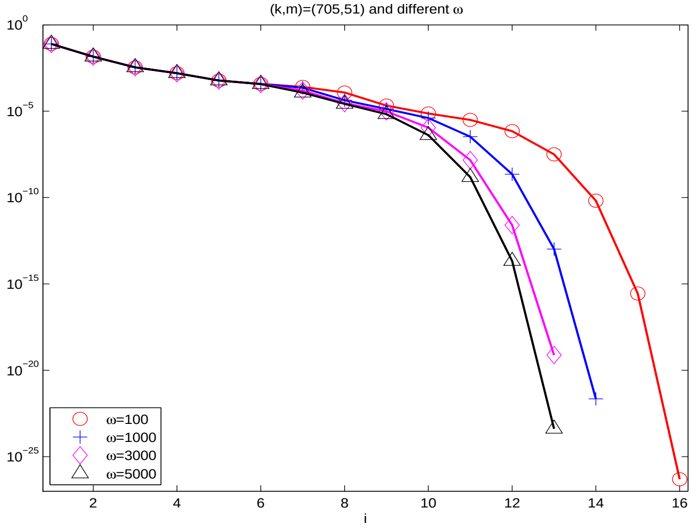
$\widehat{P}(\lambda) = (\lambda H_1^\mathsf{T} + \widehat{\Phi})\widehat{\Phi}^{-1}(\lambda\widehat{\Phi} + H_1)$ for a solution $\widehat{\Phi}$ of

$$\widehat{X} + H_1^\mathsf{T} \widehat{X}^{-1} H_1 = H_0$$

Fast Train Eigenvalue Problem (cont'd)



Fast Train Eigenvalue Problem (cont'd)



Outline

Summary

- Three types of nonlinear matrix equations (NMEs) from various applications

Summary

- Three types of nonlinear matrix equations (NMEs) from various applications
- Eigen-connections of NMEs to matrix pencils $\mathscr{A} - \lambda\mathscr{B}$

Summary

- Three types of nonlinear matrix equations (NMEs) from various applications
- Eigen-connections of NMEs to matrix pencils $\mathscr{A} - \lambda\mathscr{B}$
- A coherent general theory of doubling algorithms, actively researched in last decade or so

Summary

- Three types of nonlinear matrix equations (NMEs) from various applications
- Eigen-connections of NMEs to matrix pencils $\mathscr{A} - \lambda\mathscr{B}$
- A coherent general theory of doubling algorithms, actively researched in last decade or so
- Overwhelming favorable numerical evidences