Cyclic reduction and index reduction/shifting for a second-order probabilistic problem

Federico Poloni[1]
Joint work with Giang T. Nguyen[2]

[1]U Pisa, Computer Science Dept.
[2]U Adelaide, School of Math. Sciences

Numerical Algebra, Matrix Theory,
Differential-Algebraic Equations, and Control Theory
May 2015, Berlin

A historical finding

A historical finding

Figure: Riccati family tomb, cathedral of Treviso [photo: courtesy A. Giustiniani]

A distant cousin

One family member is still alive and very active, the Riccati equation.

1. Large-scale problems.
2. Dense Riccati equations as structured eigenproblems.
3. Nonsymmetric versions and variants in other applications.

This talk mostly about 3, but first some advertising more related to 2.

Application: linear-quadratic optimal control [Mehrmann, 91]

$$\mathcal{E}\begin{bmatrix}\dot{x}(t)\\ \dot{\mu}(t)\end{bmatrix} = \mathcal{A}\begin{bmatrix}x(t)\\ \mu(t)\end{bmatrix}, \quad x(0) = x_0, \begin{bmatrix}x(t)\\ \mu(t)\end{bmatrix} \text{ bounded.}$$

Equivalently: look for stable invariant subspace V of $\mathcal{H} = \mathcal{E}^{-1}\mathcal{A}$.

Riccati bases People first tried $V = \left[\begin{smallmatrix} I \\ X \end{smallmatrix}\right]$: X has a natural meaning.
Cons: X may have large entries.
Orthogonal bases $V = QR$ (thin), look for Q.
Cons: Structure ($X = X^T$) may be lost.

Permuted Riccati bases

Theorem [Knuth '86], [Mehrmann, P '12]

Given $V \in \mathbb{R}^{n\times m}$ with full column rank, one can factor it as

$$V = P \begin{bmatrix} I_m \\ Z \end{bmatrix} R, \quad P \text{ permutation}, \ R \in \mathbb{R}^{m\times m} \text{ invertible}, \ |Z_{ij}| \leq 1.$$

Example:

$$\begin{bmatrix} * & * & * \\ * & * & * \\ * & * & * \\ * & * & * \\ * & * & * \\ * & * & * \end{bmatrix} = \begin{bmatrix} * & * & * \\ 1 & 0 & 0 \\ * & * & * \\ 0 & 0 & 1 \\ 0 & 1 & 0 \\ * & * & * \end{bmatrix} R, \qquad |*| < 1.$$

Use it as you would use thin QR: $P \left[\begin{smallmatrix} I \\ Z \end{smallmatrix}\right]$ spans the same subspace as V and is well-conditioned.

How do we use $P \left[\begin{smallmatrix} I \\ Z \end{smallmatrix} \right] R$?

Structure preservation: if $V = \left[\begin{smallmatrix} I \\ X \end{smallmatrix} \right]$ with $X = X^*$, we can always have $Z = Z^*$ (details omitted) [Mehrmann, P '12].

We can use it to solve (generalized/extended) Riccati equations.
Great numerical properties.

Matrix pencils: we can replace $\lambda\mathcal{E} - \mathcal{A}$ with $\lambda(M\mathcal{E}) - (M\mathcal{A})$ for an invertible M (same eigenvalues and right eigenvectors).

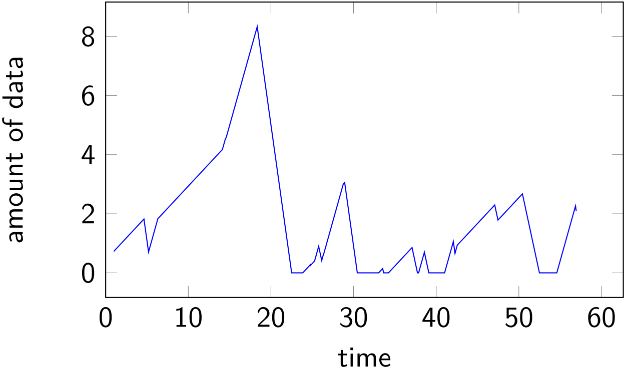
$$M(\lambda\mathcal{E} - \mathcal{A}) = \lambda \begin{bmatrix} * & 1 & * \\ * & 0 & * \\ * & 0 & * \end{bmatrix} - \begin{bmatrix} 0 & 0 & * \\ 0 & 1 & * \\ 1 & 0 & * \end{bmatrix}, \qquad |*| \leq 1.$$

Structured versions available (Hamiltonian, symplectic…)

Matlab toolbox `https://bitbucket.org/fph/pgdoubling/`

A continuous queuing model

Model: data flow in/out of a buffer at different rates according to environment state (continuous-time Markov chain) [Da Silva Soares, 2005]



Same problem, different structure

Steady state: find $w(x)$ satisfying

$$\mathcal{E}\begin{bmatrix}\dot{w}_1(x)\\ \dot{w}_2(x)\end{bmatrix} = \mathcal{A}\begin{bmatrix}w_1(x)\\ w_2(x)\end{bmatrix}, \quad w_1(0) = w_{1,0}, \quad \begin{bmatrix}w_1(x)\\ w_2(x)\end{bmatrix} \text{ bounded.}$$

$\mathcal{E}$ diagonal, $\mathcal{A}$ singular M-matrix.
Equivalently: look for stable invariant subspace of $\mathcal{E}^{-1}\mathcal{A}$.

People first tried $V = \begin{bmatrix} I \\ X \end{bmatrix}$: X has a natural meaning.
This time no stability issues: $\|X\|_\infty < 1$.

Symmetry-based structure is replaced by nonnegativity-based structure

How do we solve these problems?

Idea 1 Compute $S = \exp(\mathcal{H}t)$ for a large t. (huge entries).
Stable subspace: "$\ker S$"; unstable subspace: "$\ker S^{-1}$".

Both determined without numerical trouble from a representation $S = N_*^{-1} M_*$.

Idea 2 Replace $\exp(\mathcal{H}t)$ with $(I + \frac{t}{n}\mathcal{H})^n$, computed by repeated squaring.
All we need is a method to compute $(N_2^{-1} M_2) = (N^{-1}M)^2$

$$(N, M) \mapsto (N_2, M_2), \quad \text{doubling map.}$$

The differential equation picture

$$\exp(\mathcal{H}t) \leftrightarrow \text{Propagating ODE } \dot{u} = \mathcal{H}u.$$

$$(I + \tfrac{t}{n}\mathcal{H})^n \leftrightarrow \text{Explicit Euler's method. Other methods possible.}$$

[Anderson, '78] [Chu et al, '93+] [Benner, Byers 1998] [Guo et al, '06]

How to get accuracy?

Componentwise accurate algorithms [Wang et al, '12] [Nguyen, P '14]

Error $|\tilde{X}_{ij} - X_{ij}| \leq C|X_{ij}|$ for each i, j. Tiny elements accurate.
Main idea: Avoid all subtractions!

Linear systems $Mx = b$ with diag-dominant M-matrices solved with full machine precision (regardless of condition number) given: [O'Cinneide, '93]

- off-diagonal part, and
- diagonal excess $M\mathbf{1}$.

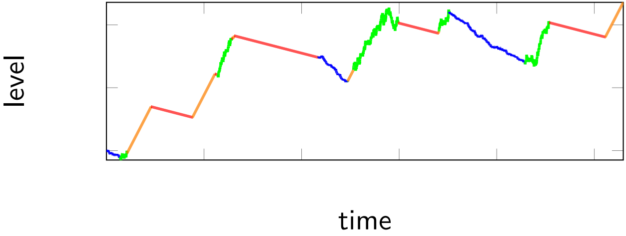
Example / Idea

$M = \begin{bmatrix} 1 & -1 \\ -1 & 1+\varepsilon \end{bmatrix}$ almost singular; can't get ε.

Instead: compute/store red entries in $\begin{bmatrix} 1 & -1 \\ -1 & 1+\varepsilon \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ \varepsilon \end{bmatrix}$. Here's ε!

Second-order problems

Now add Brownian motion in some states!



Second order model

$$V\ddot{w}(x) - D\dot{w}(x) - Kw(x) = 0, \quad \text{same constraints.}$$

V diagonal with nonnegative entries (variances), D diagonal, K M-matrix.

Want: stable invariant pair (U, X) such that $VUX^2 - DUX - KU = 0$.

Solution strategy [Latouche, Nguyen '15] [Nguyen, P]

Same plan! Start from $VUX^2 - DUX - KU = 0$,

1. Discretize $\mu = 1 + h\lambda$, or $\mu = \frac{1}{1-h\lambda}$, or $\mu = \frac{1+h\lambda/2}{1-h\lambda/2}$.
   Rational transformation of matrix polynomial [Noferini '12] [M⁴, 14].
   Gives $AUY^2 - BUY + CU = 0$, $Y = f(X)$.
2. Square Given $A\mu^2 - B\mu + C$, construct $A_2\mu^2 + B_2\mu + C_2$ with squared eigenvalues. This is cyclic reduction [Bini, Meini review '09].

We need sign properties for componentwise accuracy. How to get them?

"Interesting because it's boring"

1. Step-size h needs to be small enough.
2. Only the explicit method above works, otherwise can't undo $Y = f(X)$.

Point 2 specific to 2nd-order case.

Singular V

Choice of h We need $V_{ii} - D_{ii}h - K_{ii}h^2 \geq 0$: works for small h if $V_{ii} > 0$.

What if $V_{ii} = 0$? Differentiate some equations!

Same idea as index reduction techniques [Mehrmann, Kunkel, book].
Some care needed: not all rows where V is singular, only when $D_{ii} < 0$.

$$V\lambda^2 - D\lambda - K = \begin{bmatrix} + & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \lambda^2 - \begin{bmatrix} * & 0 & 0 \\ 0 & + & 0 \\ 0 & 0 & - \end{bmatrix} \lambda + \begin{bmatrix} * & + & + \\ + & * & + \\ + & + & * \end{bmatrix}$$

$$A\mu^2 - B\mu + C = \begin{bmatrix} + & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & + \end{bmatrix} \mu^2 - \begin{bmatrix} + & 0 & - \\ 0 & + & - \\ 0 & 0 & * \end{bmatrix} \mu + \begin{bmatrix} + & + & 0 \\ + & + & 0 \\ + & + & 0 \end{bmatrix}$$

Again, interesting because it's boring. Same tools, but for entirely different reasons (sign preservation).

Experiments

Competitors:

- **New** This componentwise-accurate algorithm [Nguyen, Poloni]
- **Eig** Eigendecomposition-based [Karandikar, Kulkarni '95]
- **Sign** Sign function-based method [Agapie, Sohraby '01]
- **QZ** QZ-based method [Unknown?]

n	kind	New	Eig	Sign	QZ
12	random	**9.7e-15**	1.9e-14	8.6e-13	6.9e-14
12	rand,imbalanced	**2.1e-12**	6.0e-08	3.6e-09	1.1e-09
30	random	**4.5e-14**	2.3e-12	1.5e-09	1.4e-12
30	rand,imbalanced	**3.7e-11**	1.5e-03	7.5e+02	4.6e-04
100	random	**3.3e-13**	5.5e-01	1.0e-07	2.4e-10
100	rand,imbalanced	**2.1e-10**	1.3e-01	2.4e+01	1.5e-07

Figure: Residual $\|VUX^2 - DUX - KU\| / \|U\|$.

Conclusions

- BVPs and Riccati equations outside control theory.
- Similar problems, different structures. Different motivations, similar choices.
- Fascinating theory: differential equations, matrix polynomials, matrix functions, inverse-free methods. . .

A smaller reflection of the variety of methods, techniques and applications that I have encountered in Volker's research.

Conclusions

- BVPs and Riccati equations outside control theory.
- Similar problems, different structures. Different motivations, similar choices.
- Fascinating theory: differential equations, matrix polynomials, matrix functions, inverse-free methods. . .

A smaller reflection of the variety of methods, techniques and applications that I have encountered in Volker's research.

Many thanks!