MAX PLANCK INSTITUTE
FOR DYNAMICS OF COMPLEX
TECHNICAL SYSTEMS
MAGDEBURG

COMPUTATIONAL METHODS IN
SYSTEMS AND CONTROL THEORY

On the Solution of the Nonsymmetric T-Riccati Equation

Peter Benner Davide Palitta

METT-VIII
8th Workshop on Matrix Equations and Tensor Techniques
Magdeburg, November 6–8, 2019

Supported by:
ARC Discovery Grant No DP1801038707

Outline

Problem: Find $X \in \mathbb{R}^{n\times n}$ such that

$$0 = \mathcal{R}_T(X) := \; DX + X^TA - X^TBX + C, \qquad A, B, C, D \in \mathbb{R}^{n\times n}.$$

Problem: Find $X \in \mathbb{R}^{n \times n}$ such that

$$0 = \mathcal{R}_T(X) := DX + X^T A - X^T BX + C, \qquad A, B, C, D \in \mathbb{R}^{n \times n}.$$

T-Riccati equations arise in

- solving large-scale Dynamic Stochastic General Equilibrium (DSGE) models [Binder/Pesaran 1997, Sims 2001, Schmitt-Grohé/Uribe 2004];
- designing H_∞ controllers for descriptor systems (with additional constraints not considered here);
- special cases appear in Hamiltonian dynamics, queuing theory, etc.
- ...

The T-Riccati Equation

We consider the "nonnegative matrix" setting, i.e., we look for nonnegative solutions to the T-Riccati equation, similar to settings often considered for nonsymmetric classical algebraic Riccati equations, cf., e.g., [Guo 2001, Bini/Iannazzo/Meini (SIAM) 2012, B./Kürschner/Saak 2016].

We consider the "nonnegative matrix" setting, i.e., we look for nonnegative solutions to the T-Riccati equation, similar to settings often considered for nonsymmetric classical algebraic Riccati equations, cf., e.g., [Guo 2001, Bini/Iannazzo/Meini (SIAM) 2012, B./Kürschner/Saak 2016].

Assumptions

- B is nonnegative, $B \geq 0$.
- C is nonpositive, $C \leq 0$.
- $I \otimes D + (A^T \otimes I)\Pi$ is a nonsingular M-matrix, where $\Pi \in \mathbb{R}^{n^2 \times n^2}$ is the permutation matrix given by $\Pi := \sum_{i,j=1}^{n} e_i e_j^T \otimes e_j e_i^T$.

The T-Riccati Equation

Assumptions

- B is nonnegative, $B \geq 0$.
- C is nonpositive, $C \leq 0$.
- $I \otimes D + (A^T \otimes I)\Pi$ is a nonsingular M-matrix, where $\Pi \in \mathbb{R}^{n^2 \times n^2}$ is the permutation matrix given by $\Pi := \sum_{i,j=1}^{n} e_i e_j^T \otimes e_j e_i^T$.

Corollary

The T-Sylvester operator

$$\mathcal{S}_T : \mathbb{R}^{n \times n} \to \mathbb{R}^{n \times n}, \quad \mathcal{S}_T(X) := DX + X^T A,$$

has a nonnegative inverse, i.e., $\mathcal{S}_T^{-1}(X) \geq 0$ for $X \geq 0$.

As a consequence, the T-Sylvester equation $\mathcal{S}_T(X) + C = 0$ has a unique solution which is nonnegative.

Outline

Existence of Minimal Solution

Want: minimal nonnegative solution to T-Riccati equation

$$0 = \mathcal{R}_T(X) := \; DX + X^T A - X^T BX + C.$$

Want: minimal nonnegative solution to T-Riccati equation

$$0 = \mathcal{R}_T(X) := \; DX + X^T A - X^T BX + C.$$

Consider fixed-point iteration

$$\begin{aligned} X_0 &= 0, \\ \text{solve} \;\; DX_{k+1} + X_{k+1}^T A &= X_k^T B X_k - C \;\; \text{for } X_{k+1}, \quad k \geq 0. \end{aligned} \tag{1}$$

Existence of Minimal Solution

Want: minimal nonnegative solution to T-Riccati equation

$$0 = \mathcal{R}_T(X) := \; DX + X^T A - X^T BX + C.$$

Consider fixed-point iteration

$$\begin{aligned} X_0 &= 0, \\ \text{solve} \;\; DX_{k+1} + X_{k+1}^T A &= X_k^T BX_k - C \;\; \text{for } X_{k+1}, \quad k \geq 0. \end{aligned} \tag{1}$$

Theorem ([B./Palitta 2019])

In addition to given assumptions, suppose $\exists\, Y \geq 0$: $\mathcal{R}_T(Y) \geq 0$. *Then,*

(i) *the iterates computed by the fixed-point iteration (1) form an increasing sequence, bounded from above by* Y*:*

$$Y \geq X_{k+1} \geq X_k \quad \text{for all } k \geq 0;$$

(ii) $\{X_k\}_{k \geq 0}$ *converges (from below) to the minimal nonnegative solution* $X_{\min}$ *of the T-Riccati equation.*

Outline

- Fixed point iteration converges to minimal nonnegative solution, with linear convergence rate depending on spectral radius of S_T^{-1}.

- Fixed point iteration converges to minimal nonnegative solution, with linear convergence rate depending on spectral radius of S_T^{-1}.
- Total complexity depends on choice of T-Sylvester solver:
  - Small-scale, dense case: solver based on generalized Schur decomposition of (A, D^T) [De Terán/Dopico 2011].
  - Large-scale, sparse/low-rank case: (extended) block-Krylov subspace-type solver [Dopico/González/Kressner/Simoncini 2016].

- Fixed point iteration converges to minimal nonnegative solution, with linear convergence rate depending on spectral radius of S_T^{-1}.
- Here: Faster convergence at similar complexity per iteration step can be achieved by (inexact) Newton solver.

- Fixed point iteration converges to minimal nonnegative solution, with linear convergence rate depending on spectral radius of S_T^{-1}.
- Here: Faster convergence at similar complexity per iteration step can be achieved by (inexact) Newton solver.
- Newton's method:

$$\mathcal{R}'_T[X](X_{k+1} - X_k) = -\mathcal{R}_T(X_k),$$

where $\mathcal{R}'_T[X]$ denotes the Fréchet derivative of $\mathcal{R}_T$ at X:

$$\begin{aligned}\mathcal{R}'_T[X](Y) &= DY + Y^TA - Y^TBX - X^TBY \\ &= (D - X^TB)Y + Y^T(A - BX).\end{aligned}$$

- Fixed point iteration converges to minimal nonnegative solution, with linear convergence rate depending on spectral radius of S_T^{-1}.
- Here: Faster convergence at similar complexity per iteration step can be achieved by (inexact) Newton solver.
- Newton's method:

$$\mathcal{R}'_T[X](X_{k+1} - X_k) \quad = \quad -\mathcal{R}_T(X_k),$$

  where $\mathcal{R}'_T[X]$ denotes the Fréchet derivative of $\mathcal{R}_T$ at X:

$$\begin{aligned} \mathcal{R}'_T[X](Y) \quad &= \quad DY + Y^T A - Y^T BX - X^T BY \\ &= \quad (D - X^T B)Y + Y^T(A - BX). \end{aligned}$$

- The resulting method is analogous to the Newton-Kleinman method for standard algebraic Riccati equations.

Newton-Kleinman iteration for T-Riccati equations:

$$X_0 = 0; \quad \text{for } k \geq 0 \text{ solve T-Sylvester equation}$$

$$(D - X_k^T B)X_{k+1} + X_{k+1}^T(A - BX_k) = -X_k^T BX_k - C.$$

Newton-Kleinman iteration for T-Riccati equations:

$$X_0 = 0; \quad \text{for } k \geq 0 \text{ solve T-Sylvester equation}$$
$$(D - X_k^T B)X_{k+1} + X_{k+1}^T(A - BX_k) = -X_k^T BX_k - C.$$

Theorem ([B./Palitta 2019])

Assume $\exists\, \overline{Y}$: $\mathcal{R}_T(\overline{Y}) > 0$, *then*

(i) $I \otimes (D - X_{\min}^T B) + ((A - BX_{\min})^T \otimes I)\Pi$ *is a nonsingular M-matrix.*

(ii) *The sequence* $\{X_k\}_{k\geq 0}$ *computed by the Newton-Kleinman iteration is well-defined since* $I \otimes (D - X_k^T B) + ((A - BX_k)^T \otimes I)\Pi$ *is a nonsingular M-matrix for all* $k \geq 0$.

(iii) $X_k \leq X_{k+1} \leq X_{\min}$ *for any* $k \geq 0$.

(iv) $\lim_{k\to\infty} X_k = X_{\min}$.

Newton-Kleinman iteration for T-Riccati equations:

$$X_0 = 0; \quad \text{for } k \geq 0 \text{ solve T-Sylvester equation}$$
$$(D - X_k^T B)X_{k+1} + X_{k+1}^T (A - BX_k) = -X_k^T BX_k - C.$$

Theorem ([B./Palitta 2019])

Assume $\exists\, \overline{Y}$*:* $\mathcal{R}_T(\overline{Y}) > 0$*, then*

(i) $I \otimes (D - X_{\min}^T B) + ((A - BX_{\min})^T \otimes I)\Pi$ *is a nonsingular M-matrix.*

(ii) *The sequence* $\{X_k\}_{k\geq 0}$ *computed by the Newton-Kleinman iteration is well-defined since* $I \otimes (D - X_k^T B) + ((A - BX_k)^T \otimes I)\Pi$ *is a nonsingular M-matrix for all* $k \geq 0$*.*

(iii) $X_k \leq X_{k+1} \leq X_{\min}$ *for any* $k \geq 0$*.*

(iv) $\lim_{k\to\infty} X_k = X_{\min}$.

Conjecture: $\mathcal{R}_T(\overline{Y}) > 0$ can be replaced by $\mathcal{R}_T(\overline{Y}) \geq 0$.

Newton-Kleinman iteration for T-Riccati equations:

$$X_0 = 0; \quad \text{for } k \geq 0 \text{ solve T-Sylvester equation}$$
$$(D - X_k^T B)X_{k+1} + X_{k+1}^T (A - BX_k) = -X_k^T BX_k - C.$$

Implementation details:

- Need to solve a T-Sylvester equation in each step — we use generalized Schur approach from [De Terán/Dopico 2011].
- Local convergence is quadratic in all experiments, as expected.
- Initial convergence can be accelerated by exact line search as suggested in [B./Byers 1998] for continuous-time symmetric algebraic Riccati equations:
  - use Newton direction $S_k := X_{k+1} - X_k$ as a descent direction;
  - minimize $\|\mathcal{R}_T(X_k + \lambda_k S_k)\|_F$ — optimizer $\lambda_k^{\mathrm{opt}} \in [0, 2]$ can be computed analytically;
  - set $X_{k+1} := X_k + \lambda_k^{\mathrm{opt}} S_k$.
- Stop when $\|\mathcal{R}_T(X_{k+1})\|_F \leq \varepsilon \|C\|_F$ for user-specified tolerance $\varepsilon > 0$.

- In a large-scale setting, dense solver for the T-Sylvester equation in the Newton steps becomes infeasible.

- In a large-scale setting, dense solver for the T-Sylvester equation in the Newton steps becomes infeasible.
- Here, we consider large-scale T-Riccati equations

$$0 \;=\; DX + X^TA - X^TBX + C.$$

with A, D sparse and B, C of low rank, i.e.,
  - $B = B_1B_2^T, \quad B_1, B_2 \in \mathbb{R}^{n\times p}$ with $p \ll n$,
  - $C = C_1C_2^T, \quad C_1, C_2 \in \mathbb{R}^{n\times q}$ with $q \ll n$,

so that $B, -C$ are nonnegative.

- In a large-scale setting, dense solver for the T-Sylvester equation in the Newton steps becomes infeasible.
- Here, we consider large-scale T-Riccati equations

$$0 = DX + X^T A - X^T BX + C.$$

with A, D sparse and B, C of low rank, i.e.,

  - $B = B_1 B_2^T, \quad B_1, B_2 \in \mathbb{R}^{n\times p}$ with $p \ll n$,
  - $C = C_1 C_2^T, \quad C_1, C_2 \in \mathbb{R}^{n\times q}$ with $q \ll n$,

so that $B, -C$ are nonnegative.

- The solution of large-scale T-Sylvester equation can then be approximated in low-rank format by (extended) block-Krylov subspace-type solver [Dopico/González/Kressner/Simoncini 2016].

Inexact Newton-Kleinman(-Krylov) method:

- T-Sylvester equation is solved using (extended) block-Krylov method only up to a residual

$$L_{k+1} := C_1^T C_2 + X_k^T B_1 B_2^T X_k - (D - X_k^T B_1 B_2^T)\widetilde{X}_{k+1} - \widetilde{X}_{k+1}^T (A - B_1 B_2^T X_k).$$

Note: L_{k+1} can be computed in low-rank format!

Inexact Newton-Kleinman(-Krylov) method:

- T-Sylvester equation is solved using (extended) block-Krylov method only up to a residual

  $$L_{k+1} := C_1^T C_2 + X_k^T B_1 B_2^T X_k - (D - X_k^T B_1 B_2^T)\widetilde{X}_{k+1} - \widetilde{X}_{k+1}^T (A - B_1 B_2^T X_k).$$

  Note: L_{k+1} can be computed in low-rank format!

- Accuracy parameter $0 < \eta_k < 1$ for T-Sylvester equation is chosen to achieve at least superlinear convergence:

  $$\|L_{k+1}\|_F \leq \eta_k \|\mathcal{R}_T(X_k)\|_F.$$

  Typical choice, also used here: $\eta_k = 1/(1 + k^3)$.

Inexact Newton-Kleinman(-Krylov) method:

- T-Sylvester equation is solved using (extended) block-Krylov method only up to a residual

$$L_{k+1} := C_1^T C_2 + X_k^T B_1 B_2^T X_k - (D - X_k^T B_1 B_2^T)\widetilde{X}_{k+1} - \widetilde{X}_{k+1}^T (A - B_1 B_2^T X_k).$$

  Note: L_{k+1} can be computed in low-rank format!

- Accuracy parameter $0 < \eta_k < 1$ for T-Sylvester equation is chosen to achieve at least superlinear convergence:

$$\|L_{k+1}\|_F \leq \eta_k \|\mathcal{R}_T(X_k)\|_F.$$

  Typical choice, also used here: $\eta_k = 1/(1 + k^3)$.

- To ensure convergence, enforce sufficient decrease condition (Armijo rule) by line search

$$\|\mathcal{R}_T(X_k + \lambda_k S_k)\|_F \leq (1 - \lambda_k \alpha)\|\mathcal{R}_T(X_k)\|_F, \quad \alpha > 0.$$

Theorem ([B./Palitta 2019])

Under the assumptions for convergence of the exact Newton-Kleinman iteration, suppose that furthermore, for all $k \geq 0$, $\exists\, \widetilde{X}_{k+1}$ *satisfying*

$$(D - X_k^T B_1 B_2^T)\widetilde{X}_{k+1} + \widetilde{X}_{k+1}^T (A - B_1 B_2^T X_k) = -X_k^T B_1 B_2^T X_k - C_1^T C_2 + L_{k+1}$$

where $\|L_{k+1}\|_F \leq \eta_k \|\mathcal{R}_T(X_k)\|_F$. *Then:*

(i) *If the step size parameters* λ_k *are bounded away from zero, i.e.,* $\lambda_k \geq \lambda_{\min} > 0$ *for all* k, *then* $\|\mathcal{R}_T(X_k)\|_F \to 0$.

(ii) *If, in addition to (i), the matrices* L_{k+1} *are nonnegative for all* $k \geq 0$, *then the sequence* $\{X_k\}_{k\geq 0}$ *generated by the inexact Newton-Kleinman method with* $X_0 = 0$ *is well-defined and* $X_k \leq X_{k+1} \leq X_{\min}$. *Moreover,*

$$\lim_{k\to\infty} X_k = X_{\min}.$$

Outline

- We use the same coefficient matrices as in "Numerical test 7.1" from [Dopico et al. 2016]: $D, A \in \mathbb{R}^{n\times n}$ represent the finite difference discretization on the unit square of the 2-dimensional differential operators

$$\begin{aligned} \mathcal{L}_D(u) &= -u_{xx} - u_{yy} + y(1-x)u_x + 10^4 u, \\ \mathcal{L}_A(u) &= -u_{xx} - u_{yy} = -\Delta u \end{aligned}$$

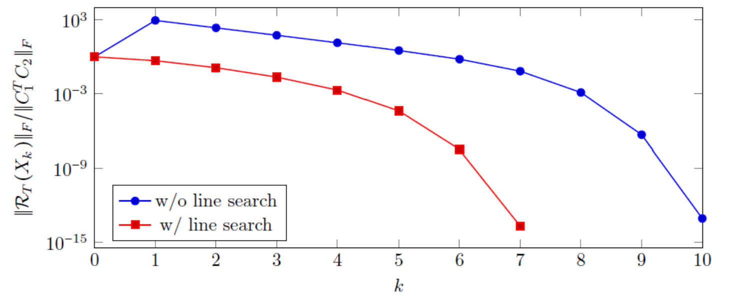
  with homogeneous Dirichlet boundary conditions.
- For small-scale tests, $B, C \in \mathbb{R}^{n\times n}$ are full random matrices.
- For large-scale tests, we consider low-rank matrices
  - $B = B_1 B_2^T$, $B_1, B_2 \in \mathbb{R}^{n\times p}$,
  - $C = C_1 C_2^T$, $C_1, C_2 \in \mathbb{R}^{n\times q}$,

  such that B_i, C_i have unit norm and random entries for $i = 1, 2$.

Results for small-scale / "exact" Newton-Kleinman

	n	Its	Rel. Res	Time (secs)
w/o line search	324	8	9.0e-15	11.28
w/ line search		5	1.1e-14	7.54
w/o line search	784	10	7.5e-14	99.94
w/ line search		7	2.4e-14	73.73

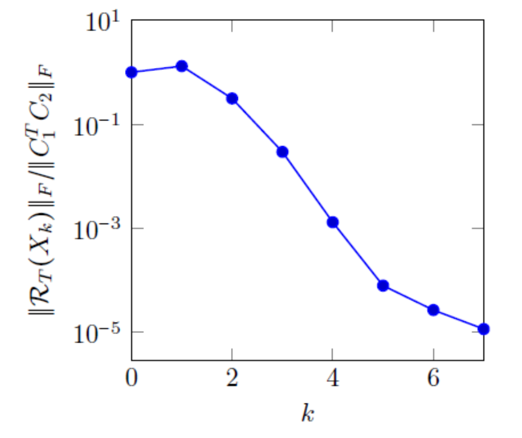
Results for small-scale ($n = 784$) / "exact" Newton-Kleinman



Results for large-scale / "inexact" Newton-Kleinman

n	p	q	Its (inner)	Mem.	Rank(X)	Rel. Res.	Time (s)
10,000	1	1	13 (7.46)	192	24	6.7e-7	15.65
	1	5	8 (8.5)	672	93	6.3e-7	52.15
	5	10	6 (6.3)	1560	213	4.5e-7	110.12
22,500	1	1	14 (9.86)	256	30	5.9e-7	69.19
	1	5	no convergence				
	5	10	no convergence				
32,400	1	1	10 (10.6)	384	26	6.8e-7	127.61
	1	5	no convergence				
	5	10	no convergence				

Results for large-scale / "inexact" Newton-Kleinman



Similar to Example 6.1 in [Guo 2001], define

$$
\begin{aligned}
R &= \texttt{rand}(2n, 2n) \in \mathbb{R}^{2n\times 2n},\\
W &= \mathrm{diag}(R\mathbf{1}) - R, \text{ where } \mathbf{1} = (1,\dots,1)^T \in \mathbb{R}^{2n}.
\end{aligned}
$$

$A, D \in \mathbb{R}^{n\times n}$ are chosen according to the partition

$$
W = \begin{bmatrix} D & M \\ N & A \end{bmatrix}, \quad \text{while } B = -N/\|N\|_F.
$$

Then define $X_{\texttt{exact}}$ with uniformly distributed random entries and unit norm, and compute

$$
C := DX_{\texttt{exact}} + X_{\texttt{exact}}^T A - X_{\texttt{exact}}^T B X_{\texttt{exact}}.
$$

Results for small-scale / "exact" Newton-Kleinman

	n	Its	Rel. Res.	Rel. Err.	Time (s)
w/o line search	500	3	1.0e-14	1.1e-10	10.80
w/ line search		3	1.1e-14	7.6e-11	10.84
w/o line search	1,000	4	1.5e-14	2.2e-9	78.59
w/ line search		4	1.5e-14	4.2e-9	78.99

Here, line search brings no advantage.

- Compute two sparse nonnegative matrices $F, G \in \mathbb{R}^{n\times n}$ with random entries.
- Define

$$\begin{aligned} D &:= F + (\rho(F) + 1)I, \\ A &:= G + (\rho(G) + 20)I. \end{aligned}$$

- Construct low-rank matrices
  - $B = B_1 B_2^T,\ B_1, B_2 \in \mathbb{R}^{n\times p}$,
  - $C = C_1 C_2^T,\ C_1, C_2 \in \mathbb{R}^{n\times q}$

  such that $B_i,\ C_i$ have unit norm and random entries for $i = 1, 2$.

Results for large-scale / "inexact" Newton-Kleinman

n	p	q	Its (inner)	Mem.	Rank(X)	Rel. Res.	Time (s)
10,000	1	1	4 (1.5)	32	10	6.1e-7	0.16
	1	5	5 (2.2)	192	36	5.5e-9	1.11
	5	10	5 (2)	360	76	4.9e-9	3.27
50,000	1	1	4 (1.5)	32	6	9.7e-8	0.79
	1	5	5 (2)	192	36	5.5e-9	5.44
	5	10	5 (2)	360	76	3.8e-9	14.88
100,000	1	1	4 (1.5)	32	6	5.8e-8	1.48
	1	5	5 (2.2)	192	36	5.5e-9	11.33
	5	10	5 (2)	360	76	3.9e-9	24.49

Numerical results confirm linear complexity of inexact Newton-Kleinman-Krylov solver!

Summary & Outlook

- We have established sufficient conditions for the existence and uniqueness of a minimal nonnegative solution of T-Riccati equations.
- The minimum nonnegative solution can be computed by a Newton(-Kleinman) method.
- Line search can accelerate the convergence.
- In the large scale setting, low-rank approximate solutions can be computed by inexact Newton-Kleinman, where line search guarantees convergence as long as T-Sylvester equations can be solved in the Newton steps.

Summary & Outlook

- We have established sufficient conditions for the existence and uniqueness of a minimal nonnegative solution of T-Riccati equations.
- The minimum nonnegative solution can be computed by a Newton(-Kleinman) method.
- Line search can accelerate the convergence.
- In the large scale setting, low-rank approximate solutions can be computed by inexact Newton-Kleinman, where line search guarantees convergence as long as T-Sylvester equations can be solved in the Newton steps.
- Convergence analysis of (extended) Krylov-type solvers for T-Sylvester equations?
- Alternative solvers for T-Sylvester equations, e.g. of ADI-type, possible?
- Projection techniques working directly on the T-Riccati equations?
- Other (than nonnegative matrix) settings for T-Riccati equations?

References

P. Benner and R. Byers.
An exact line search method for solving generalized continuous-time algebraic Riccati equations.
IEEE Transactions on Automatic Control, 43:101–107, 1998.

P. Benner, P. Kürschner, and J. Saak.
Low-rank Newton-ADI methods for large nonsymmetric algebraic Riccati equations.
Journal of the Franklin Institute, 353(5):1147–1167, 2016.

P. Benner and D. Palitta.
On the solution of the nonsymmetric T-Riccati equation.
arXiv:1911.xxxxx [math.NA], November 2019.

D. A. Bini, B. Iannazzo, and B. Meini.
Numerical solution of algebraic Riccati equations.
Fundamentals of Algorithms, SIAM, Philadelphia, PA, 2012.

F. De Terán and F. M. Dopico.
Consistency and efficient solution of the Sylvester equation for $\star$-congruence.
Electronic Journal of Linear Algebra, 22:849–863, 2011.

F. M. Dopico, J. González, D. Kressner, and V. Simoncini.
Projection methods for large-scale T-Sylvester equations.
Mathematics of Computation, 85:2427–2455, 2016.

C.-H. Guo.
Nonsymmetric algebraic Riccati equations and ´Wiener-Hopf factorization for M-matrices.
SIAM Journal of Matrix Analysis and Applications, 23:225–242, 2001.