A Collection of Benchmark Examples for
the Numerical Solution of Algebraic
Riccati Equations II:
Discrete-Time Case
A Collection of Benchmark Examples for the Numerical Solution of Algebraic Riccati Equations II: Discrete-Time Case

Peter Benner* Alan J. Laub† Volker Mehrmann*

Abstract

This is the second part of a collection of benchmark examples for the numerical solution of algebraic Riccati equations. After presenting examples for the continuous-time case in Part I, our concern in this paper is discrete-time algebraic Riccati equations. This collection may serve for testing purposes in the construction of new numerical methods, but may also be used as a reference set for the comparison of methods.

1 Introduction

We present a collection of examples for discrete-time algebraic Riccati equations (DARE) of the form

\[
0 = A^T X A - X - (A^T X B + S)(R + B^T X B)^{-1}(B^T X A + S^T) + Q
\]

(1)

where \(A, Q, X \in \mathbb{R}^{n \times n}, B, S \in \mathbb{R}^{n \times m},\) and \(R = R^T \in \mathbb{R}^{m \times m}.\) The matrix \(Q = Q^T\) may be given in factored form \(Q = C^T \tilde{Q} C\) with \(C \in \mathbb{R}^{p \times n}\) and \(\tilde{Q} = \tilde{Q}^T \in \mathbb{R}^{p \times p}.\)

As it will be described below, (1) can be solved using its relationship to the symplectic pencil defined by

\[
L - \lambda M = \begin{bmatrix} \hat{A} & 0 \\ -\tilde{Q} & I_n \end{bmatrix} - \lambda \begin{bmatrix} I_n & G \\ 0 & \hat{A}^T \end{bmatrix}
\]

(2)

where

\[
\hat{A} = A - BR^{-1}S^T,
\]

\[
G = BR^{-1}B^T,
\]

\[
\tilde{Q} = Q - SR^{-1}S^T = C^T \tilde{Q} C - SR^{-1}S^T.
\]
If $\dot{A}$ is invertible, this pencil is equivalent to the symplectic matrices

$$Z = M^{-1} L = \begin{bmatrix} \dot{A} + G\dot{A}^{-T} \dot{Q} & -G\dot{A}^{-T} \\ -\dot{A}^{-T} \dot{Q} & \dot{A}^{-T} \end{bmatrix},$$

or

$$\tilde{Z} = LM^{-1} = \begin{bmatrix} \dot{A} & -\dot{A}G\dot{A}^{-T} \\ -\dot{Q} & \dot{Q}G\dot{A}^{-T} + \dot{A}^{-T} \end{bmatrix}.$$  

The DARE (1) arises, e.g., in (a) stochastic realization problems, and (b) linear-quadratic control problems. In case (a), $R$ is the measurement noise covariance and it is not uncommon for this kind of matrix to be singular. For (b), $R$ is the control weighting matrix and in the discrete-time case, occasionally such a matrix can be singular, too. In these cases, the pencil formulation (2) is not possible. An extended symplectic pencil (ESP) can then be formed by

$$\tilde{L} - \lambda \tilde{M} = \begin{bmatrix} A & 0 & B \\ Q & -I_n & S \\ S^T & 0 & R \end{bmatrix} - \lambda \begin{bmatrix} I_n & 0 & 0 \\ 0 & -A^T & 0 \\ 0 & -B^T & 0 \end{bmatrix}.$$  

To illustrate a problem where the DARE (1) arises, we consider the discrete-time linear-quadratic control problem (case (b) from above).

Minimize

$$J(x_0, u) = \frac{1}{2} \sum_{k=0}^{\infty} \left( y_k^T \tilde{Q} y_k + 2x_k^T S u_k + u_k^T R u_k \right) dt$$

subject to the difference equation

$$x_{k+1} = Ax_k + Bu_k, \quad k = 0, 1, \ldots, \quad x_0 = \xi,$$
$$y_k = Cx_k, \quad k = 0, 1, \ldots.$$  

If, for example, $\tilde{Q} \geq 0$, $R > 0$, $(A, B)$ stabilizable$^1$, and $(A, C)$ detectable$^2$, then the solution of the optimal control problem (6)-(8) is given by the feedback law

$$u_k = -(R + B^T X B)^{-1}(A^T X B + S)^T x_k, \quad k = 0, 1, \ldots,$$

where $X$ is the unique stabilizing, positive semidefinite solution of (1) (see, e.g., [21, 28]).

One common approach to solve (1) is to compute the stable invariant subspace of the symplectic matrix $Z$ or the stable deflating subspace of the (extended) symplectic pencil given above, i.e., the invariant/deflating subspace corresponding to the generalized eigenvalues of $L - \lambda M$, $\tilde{L} - \lambda \tilde{M}$, respectively, inside the unit circle (e.g., [18, 21, 24]). If this subspace is spanned by

$$[U_1 \quad U_2]^T$$

or

$$[U_1 \quad U_2 \quad U_3]^T,$$

and $U_1$ is invertible, then $X = U_2 U_1^{-1}$ is the stabilizing solution of (1), i.e., all the eigenvalues of

$$F = A - B(R + B^T X B)^{-1}(A^T X B + S)^T$$

lie inside the unit circle.

$^1$$(A, B)$ is (d-)stabilizable, if rank $[A - \lambda M, B] = n$ for all $\lambda$ with $|\lambda| \geq 1$.

$^2$$(A, C)$ is (d-)detectable, if $(A^T, C^T)$ is (d-)stabilizable.
At this point it should be noted that it is possible to transform a continuous-time algebraic Riccati equation (CARE) into a DARE (and vice versa) via a (generalized) Cayley transformation, i.e., the Hamiltonian matrix corresponding to the CARE is transformed into a symplectic matrix/pencil. From this symplectic pencil it is possible to derive the coefficient matrices of a corresponding DARE under certain regularity assumptions; see [22]. In this way, it is possible to obtain DARE examples from the first part of our benchmark collection. We do not use this approach here, though, and restrict ourselves to data arising naturally in a discrete-time setting and/or highlighting some of the properties of discrete-time algebraic Riccati equations.

In the sequel we will use the following notation. Let $A \in \mathbb{R}^{n \times n}$. By $\sigma(A)$ we denote the set of eigenvalues or spectrum of $A$. The spectral norm of a matrix is given by
\[
\|A\| = \sqrt{\max\{|\lambda| : \lambda \in \sigma(A^T A)\}}
\]
and the given matrix condition numbers are based upon the spectral norm,
\[
\kappa(A) = \|A\| A^{-1}.
\]
The Frobenius norm of a matrix will be denoted by $|A|_F$ and is given by
\[
|A|_F = \sqrt{\sum_{i,j=1}^{n} a_{ij}^2}.
\]
All norms and condition numbers given in the sequel were computed by the MATLAB\(^3\) functions \texttt{norm} and \texttt{cond}.

The examples are grouped in three sections. The first section contains parameter-free examples of fixed dimension while the second has parameter-dependent problems of fixed size. Section 4 contains examples of scalable size.

The coefficient matrices of the presented examples are usually given in the same form as they appear in the literature. Since in most cases $S = 0$, we omit $S$ in all examples where this property holds.

All presented examples may be generated by the FORTRAN 77 subroutine DAREX (see Appendix A) and the MATLAB function \texttt{darex} (see Appendix B). Appendix C describes how to obtain the software.

The description of each example contains a table with some of the system properties. This information is summarized in Appendix D. For all parameters needed in the examples there exist default values that are also given in the tables. These default values are chosen in such a way that the collection of examples can be used as a testset for the comparison of methods. The tables contain information for one or more choices of the parameters. Underlined values indicate the default values.

For each example, we provide the condition number $\kappa(\hat{A})$ which shows if the symplectic matrix $Z$ in (3) can be formed safely (though it is still favorable to use the pencil approach and thereby avoiding a matrix inversion unless $\hat{A}^{-1}$ is "easy" to form). The column $|\lambda_{\text{max}}^C|$ indicates the absolute value of the closed-loop eigenvalue of largest modulus, i.e.,
\[
|\lambda_{\text{max}}^C| = \max\{|\lambda| : \lambda \in \sigma(F)\}
\]

\(^3\)MATLAB is a trademark of The MathWorks, Inc.
with $F$ as in (9). These are the (generalized) eigenvalues of the symplectic matrix (pencils) in (2), (3), and (5) inside the unit circle. Further, we give norms and condition numbers of the stabilizing solution $X$. For examples without analytical solution available, we computed approximations by the generalized Schur vector method [3, 24]. If possible, these approximations were refined by Newton’s method [3, 12, 21] to achieve the highest possible accuracy. We then chose the approximate solution with smallest residual norm and recomputed the solution using the optimal scaling strategy proposed in [11]. This computed solution was then used to determine the properties of the example.

The “right” condition number of algebraic Riccati equations is still an open problem although the problem has been attacked by several papers during recent years, see, e.g., [11, 15, 26, 30]. For simplicity, here we use the condition number proposed in [11]. This condition number measures the sensitivity of the stabilizing DARE solution with respect to first-order perturbations by means of the Fréchet derivative of the DARE at $X$. In [11] it is shown that, assuming $Q \geq 0$, $R > 0$, the so defined condition number is given by

$$K_{\text{DARE}} = \frac{\left| \begin{bmatrix} Z_1 & Z_2 & Z_3 \end{bmatrix} \right|}{\|X\|_F},$$

where

$$Z_1 = |A|_F P^{-1} \left( I_n \otimes F^T X + (F^T X \otimes I_n)T \right),$$

$$Z_2 = -|G|_F P^{-1} \left( \hat{A}^T X (I_n + GX)^{-1} \otimes \hat{A}^T X (I_n + GX)^{-1} \right),$$

$$Z_3 = |Q|_F P^{-1}.$$

Here, denoting the $j$th unit vector by $e_j$, the permutation matrix $T$ is defined by

$$T = \sum_{i,j=1}^n e_i e_j^T \otimes e_j e_i^T,$$

and $P$ is the matrix representation of the Stein (discrete Lyapunov) operator

$$\Omega(Z) = Z - F^T Z F.$$

The computation of $K_{\text{DARE}}$ therefore requires the solution of the linear equations (10)-(12). Since Kronecker products are involved, these systems get very large even for small numbers of $n$. For larger $n$, an inverse power iteration can be employed to estimate $\left| \begin{bmatrix} Z_1 & Z_2 & Z_3 \end{bmatrix} \right|_F$ (see [11]). This approach requires in each iteration step the solution of two Stein equations corresponding to $\Omega$. 

4
2 Parameter-free problems of fixed size

Example 1 [17, Example 2]

| n | m | p | parameter | $\kappa(A)$ | $|\lambda_{max}^C|$ | $|X|$ | $\kappa(X)$ | $K_{DARE}$ |
|---|---|---|-----------|-------------|----------------|--------|-------------|-----------|
| 2 | 1 | 2 |           | 114.99      | 0.50           | 21.03  | $\infty$   | 18.85     |

This is an example of stabilizable-detectable, but uncontrollable-unobservable data. We have the following system matrices:

\[ A = \begin{bmatrix} 4 & \frac{3}{2} \\ -\frac{9}{2} & -\frac{3}{2} \end{bmatrix}, \quad B = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \quad R = 1, \quad Q = \begin{bmatrix} 9 & 6 \\ 6 & 4 \end{bmatrix} \]

with stabilizing solution

\[ X = \frac{1+\sqrt{5}}{2} \begin{bmatrix} 9 & 6 \\ 6 & 4 \end{bmatrix} \]

and closed-loop spectrum \{-1/2, (3 - \sqrt{5})/3\}.

Example 2 [17, Example 3], [16, Example 6.15]

| n | m | p | parameter | $\kappa(A)$ | $|\lambda_{max}^C|$ | $|X|$ | $\kappa(X)$ | $K_{DARE}$ |
|---|---|---|-----------|-------------|----------------|--------|-------------|-----------|
| 2 | 2 | 2 |           | 1.05        | 0.69           | 5.07 x 10^{-2} | 4.97     | 4.74       |

This example illustrates a linear-quadratic control problem as defined by (6) (8). The coefficient matrices are

\[ A = \begin{bmatrix} 0.9512 & 0 \\ 0 & 0.9048 \end{bmatrix}, \quad B = \begin{bmatrix} 4.877 \\ -1.1895 \end{bmatrix}, \quad Q = \begin{bmatrix} 0.005 & 0 \\ 0 & 0.02 \end{bmatrix} \]

In [16, 17], solution matrices are given. We omit reproducing them here since they are not derived analytically.

Example 3 [31, Example II]

| n | m | p | parameter | $\kappa(A)$ | $|\lambda_{max}^C|$ | $|X|$ | $\kappa(X)$ | $K_{DARE}$ |
|---|---|---|-----------|-------------|----------------|--------|-------------|-----------|
| 2 | 1 | 1 |           | 5.83        | 0.00           | 1.00   | 1.00        | $\infty$ |

This example was used in [31] to demonstrate a compression technique for the extended pencil (5). The data are given by

\[ A = \begin{bmatrix} 2 & -1 \\ 1 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad Q = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \quad R = 0. \]

If interpreted in terms of a linear system as in (7) (8), $Q$ can be written as

\[ Q = C^T \tilde{Q} C, \quad C = \begin{bmatrix} 0 & 1 \end{bmatrix}, \quad \tilde{Q} = 1. \]

The exact stabilizing solution is $X = I_2$, and the closed-loop spectrum is \{0, 0\}. Due to the singularity of $R$, the condition number $K_{DARE}$ is not defined here (represented by a value “$\infty$” in the table). This example can be used, e.g., as a first test of any solver to deal with a singular weighting/measurement noise covariance matrix.
Example 4 [13]

| n   | m   | p   | parameter | $\kappa(A)$ | $|C_{max}|$ | $\|X\|$ | $\kappa(X)$ | $K_{DARE}$ |
|-----|-----|-----|-----------|-------------|-------------|----------|-------------|------------|
| 2   | 2   | 2   |           | $\infty$    | 0.69        | 126.99   | $2.84 \times 10^3$ | $\infty$   |

This is another example with a singular $R$ matrix. Furthermore, we have a nonzero $S$ matrix. The coefficients of (1) are given by

$$A = \begin{bmatrix} 0 & 1 \\ 0 & -1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \quad R = \begin{bmatrix} 9 & 3 \\ 3 & 1 \end{bmatrix},$$

$$Q = \frac{1}{11} \begin{bmatrix} -4 & -4 \\ -4 & 7 \end{bmatrix}, \quad S = \begin{bmatrix} 3 & 1 \\ -1 & 7 \end{bmatrix}.$$  

Again, the DARE condition number $K_{DARE}$ can not be computed due to the singular $R$.

Example 5 [14], [22, Example 2].

| n   | m   | p   | parameter | $\kappa(A)$ | $|C_{max}|$ | $\|X\|$ | $\kappa(X)$ | $K_{DARE}$ |
|-----|-----|-----|-----------|-------------|-------------|----------|-------------|------------|
| 2   | 1   | 2   |           | $\infty$    | 0.38        | 5.19     | 114.13      | 1.88       |

This example shows one of the major differences between the properties of continuous-time algebraic Riccati equations and their discrete counterparts. Consider the DARE defined by

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad Q = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}, \quad R = 1.$$  

The spectrum of the pencil $L - \lambda M$ in (2) is $\{0, \infty, -(3 \pm \sqrt{5})/2\}$. The DARE has exactly two solutions,

$$X_1 = \begin{bmatrix} 1 \\ 2 + \sqrt{5} \end{bmatrix}, \quad X_2 = \begin{bmatrix} 1 \\ 2 - \sqrt{5} \end{bmatrix},$$

but neither of them is negative semidefinite. On the other hand, $(A, B)$ is controllable. In the case of a continuous-time system, this property would assure the existence of a negative semidefinite solution. The stabilizing solution in the control-theoretic sense is the positive definite solution $X_1$.

Example 6 [1]

| n   | m   | p   | parameter | $\kappa(A)$ | $|C_{max}|$ | $\|X\|$ | $\kappa(X)$ | $K_{DARE}$ |
|-----|-----|-----|-----------|-------------|-------------|----------|-------------|------------|
| 4   | 2   | 4   |           | 1.01        | 0.94        | 35.36    | 3.34        | 30.58      |

The data of this example represent a simple control problem for a satellite. The system is given by equations describing the small angle altitude variations about the roll and yaw axes of a satellite in circular orbit. These equations originally form a second-order differential equation. A first-order realization of this model and sampling every 0.1 seconds yields the system matrices

$$A = \begin{bmatrix} 0.998 & 0.067 & 0 & 0 \\ -0.067 & 0.998 & 0 & 0 \\ 0 & 0 & 0.998 & 0.153 \\ 0 & 0 & -0.153 & 0.998 \end{bmatrix}, \quad B = \begin{bmatrix} 0.0033 & 0.02 \\ 0.1 & -0.0007 \\ 0.04 & 0.0073 \\ -0.0028 & 0.1 \end{bmatrix}.$$  

The weighting matrices used in the performance index $J(x, u)$ in (6) are given by

$$Q = \hat{Q} = \begin{bmatrix} 1.87 & 0 & 0 & -0.244 \\ 0 & 0.744 & 0.205 & 0 \\ 0 & 0.205 & 0.589 & 0 \\ -0.244 & 0 & 0 & 1.048 \end{bmatrix}, \quad R = I_2.$$
Example 7 [19]

| \( n \) | \( m \) | \( p \) | parameter | \( \kappa(A) \) | \( |\lambda_{\text{max}}^C| \) | \( \|X\| \) | \( \kappa(X) \) | \( K_{\text{DARE}} \) |
|---|---|---|---|---|---|---|---|---|
| 4 | 2 | 4 | parameter | 19.86 | 0.99 | 2.06 | 183.33 | 790.37 |

This is a simple example of a control system having slow and fast modes.

\[
A = 10^{-3} \times \begin{bmatrix} 984.75 & -79.903 & 0.9054 & -1.0765 \\ 41.588 & 998.99 & -35.855 & 12.684 \\ -54.662 & 44.916 & -329.91 & 193.18 \\ 266.24 & -100.45 & -924.55 & -263.25 \end{bmatrix},
\]

\[
B = 10^{-4} \times \begin{bmatrix} 37.112 & 7.361 \\ -870.51 & 0.93411 \\ -11984.0 & -4.1378 \\ -31927.0 & 9.2535 \end{bmatrix}, \quad R = I_2, \quad Q = 0.01 I_4.
\]

One complex conjugate pair of the computed closed-loop eigenvalues is located on a circle with radius \( \approx 0.99 \) around the origin, i.e., is relatively close to the unit circle. Requiring that this distance should not cause any problems for any DARE solver seems to be reasonable.

Example 8 [20, Example 4.3]

| \( n \) | \( m \) | \( p \) | parameter | \( |\lambda_{\text{max}}^C| \) | \( \|X\| \) | \( \kappa(X) \) | \( K_{\text{DARE}} \) |
|---|---|---|---|---|---|---|---|
| 4 | 4 | 4 | parameter | 37.860 | \( \approx 1 - 1.8 \times 10^{-5} \) | 65.77 | 6.18 \times 10^{11} | 5.12 \times 10^4 |

Here, the coefficient matrices of (1) are constructed as follows. Given

\[
A_0 = \begin{bmatrix} 0.4 & 0 & 0 & 0 \\ 1 & 0.6 & 0 & 0 \\ 0 & 1 & 0.8 & 0 \\ 0 & 0 & 0 & -0.999982 \end{bmatrix}, \quad Q_0 = \begin{bmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.
\]

\[
V = \begin{bmatrix} 1 & -1 & -1 & -1 \\ 0 & 1 & -1 & -1 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad V^{-1} = \begin{bmatrix} 1 & 1 & 2 & 4 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix},
\]

we obtain

\[
A = VA_0 V^{-1} = \begin{bmatrix} -0.6 & -2.2 & -3.6 & -5.400018 \\ 1 & 0.6 & 0.8 & 3.399982 \\ 0 & 1 & 1.8 & 3.799982 \\ 0 & 0 & 0 & -0.999982 \end{bmatrix}, \quad B = VI_4 = V,
\]

\[
Q = V^{-T} Q_0 V^{-1} = \begin{bmatrix} 2 & 1 & 3 & 6 \\ 1 & 2 & 2 & 5 \\ 3 & 2 & 6 & 11 \\ 6 & 5 & 11 & 22 \end{bmatrix}, \quad R = I_4.
\]

A factorization in the control-theoretic sense, \( Q = C^T \bar{Q} C \), is therefore given by \( C := V^{-1} \) and \( \bar{Q} := Q_0 \). All the generalized eigenvalues of \( L - \lambda M \) are real. The distance of the largest closed-loop eigenvalue to the unit circle is \( \approx 1.8 \times 10^{-5} \). The problem is designed so that \( \kappa(L + M) \approx 4 \times 10^{11} \). Due to this large condition number and the eigenvalues close to the unit circle, problems with the convergence of the iteration process are to be expected when the DARE is solved via a method based on the sign function iteration (e.g., the methods in [4, 9]). Note that \( K_{\text{DARE}} \) signals only a very mild ill-conditioning of the DARE.
Example 9 [8, Section 2.7.4]

<table>
<thead>
<tr>
<th>n</th>
<th>m</th>
<th>p</th>
<th>parameter</th>
<th>( \kappa(A) )</th>
<th>( A_{max} )</th>
<th>( | X | )</th>
<th>( \kappa(X) )</th>
<th>( K_{DARE} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>2</td>
<td>5</td>
<td></td>
<td>23.52</td>
<td>0.98</td>
<td>73.90</td>
<td>73.73</td>
<td>100.81</td>
</tr>
</tbody>
</table>

The fifth-order linearized state-space model of a chemical plant presented in [10, 29] is discretized by sampling every 0.5 seconds, yielding a discrete-time linear-quadratic control problem of the form (6) (7) defined by

\[
A = 10^{-4} \times \begin{bmatrix}
9540.70 & 196.43 & 35.97 & 6.73 & 1.90 \\
4084.90 & 4131.70 & 1608.40 & 446.79 & 119.71 \\
1221.70 & 2632.60 & 3614.90 & 1593.00 & 1238.30 \\
411.18 & 1285.80 & 2729.90 & 2144.20 & 4097.60 \\
13.05 & 58.08 & 187.50 & 361.62 & 9428.00
\end{bmatrix}, \quad B = 10^{-4} \times \begin{bmatrix}
4.34 & -1.22 \\
266.06 & -104.53 \\
375.30 & -551.00 \\
360.76 & -660.00 \\
46.17 & -91.48
\end{bmatrix}.
\]

The weighting matrices in the cost functional (6) are chosen as identities, i.e., we have \( Q = \bar{Q} = I_5 \) and \( R = I_2 \).

If we modify the optimal control problem (6) (8) by allowing the observation to depend upon the control, we obtain the following problem:

Minimize

\[
J(x_0, u) = \frac{1}{2} \sum_{k=0}^{\infty} \left( y_k^T \bar{Q} y_k + u_k^T \bar{R} u_k \right) dt
\]

subject to the difference equation

\[
x_{k+1} = Ax_k + Bu_k, \quad k = 0, 1, \ldots, \quad x_0 = \xi,
\]

\[
y_k = Cx_k + Du_k, \quad k = 0, 1, \ldots,
\]

then we can rewrite the cost functional (13) as

\[
J(x_0, u) = \frac{1}{2} \sum_{k=0}^{\infty} \left( x^T Q x + x^T S u + u^T S^T x + u^T Ru \right) dt
\]

where \( Q = C^T \bar{Q} C, \quad R = \bar{R} + D^T \bar{Q} D, \quad \text{and} \quad S = C^T \bar{Q} D. \)

The data of the following example come from such a problem.

Example 10 [7]

<table>
<thead>
<tr>
<th>n</th>
<th>m</th>
<th>p</th>
<th>parameter</th>
<th>( \kappa(A) )</th>
<th>( A_{max} )</th>
<th>( | X | )</th>
<th>( \kappa(X) )</th>
<th>( K_{DARE} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>2</td>
<td>2</td>
<td></td>
<td>( \infty )</td>
<td>6.70</td>
<td>2.53</td>
<td>37.38</td>
<td>3.94</td>
</tr>
</tbody>
</table>

The matrices of the linear system \( (A, B, C, D) \) are given by

\[
A = \begin{bmatrix}
0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 \\
\end{bmatrix}, \quad B = \begin{bmatrix}
0 & 0 \\
0 & 0 \\
1 & 0 \\
0 & 0 \\
0 & 0 \\
0 & 1 \\
\end{bmatrix},
\]

\[
C = \begin{bmatrix}
1 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & -1 & 0
\end{bmatrix}, \quad D = \begin{bmatrix}
1 & 0 \\
1 & 0
\end{bmatrix}.
\]
With $\tilde{Q} = \tilde{R} = I_2$, we obtain the following coefficient matrices for the DARE: $A$, $B$, $C$, $\tilde{Q}$ are defined above, and

$$Q = C^T \tilde{Q} C = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad R = \begin{bmatrix} 3 & 0 \ 0 & 1 \end{bmatrix}, \quad S = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}. $$

The system properties are good-natured. The system can easily be transformed to a standard system as in (2). Therefore, this example is helpful for first verifications of any DARE solver based on the extended formulation given in (5) since the results can be compared to those obtained by any other solver based on the formulation by the symplectic pencil (2).

**Example 11** [25]

| $n$ | $m$ | $p$ | $\nu$ | $\kappa(A)$ | $|\lambda_{max}^C|$ | $|X|$ | $\kappa(X)$ | $K_{DARE}$ |
|-----|-----|-----|------|-------------|-----------------|-----|-------------|------------|
| 9   | 3   | 2   | 1.58 $\times 10^6$ | 0.96 | 607.66 | $\infty$ | 74.23 |

This is the data for a 9th-order discrete state-space model of a tubular ammonia reactor. It should be noted that the underlying model includes a disturbance term which is neglected in this context.

The continuous state-space model of this problem was presented as Example 5 in the first part of the benchmark collection [5]. Sampling every 30 seconds yields the following system matrices for the discrete model:

$$A = 10^{-2} \begin{bmatrix} 87.01 & 13.50 & 1.159 & 0.05014 & -3.722 & 0.03484 & 0.4242 & 0.7249 \\ 7.655 & 89.74 & 1.272 & 0.05504 & -4.016 & 0.03743 & 0.4530 & 0.7499 \\ -12.72 & 35.75 & 81.70 & 0.1455 & -10.28 & 0.0987 & 1.185 & 1.872 \\ -36.35 & 63.39 & 7.491 & 79.66 & -27.35 & 0.2653 & 3.172 & 4.882 \\ -96.00 & 164.59 & -12.89 & -0.5597 & 7.142 & 0.7108 & 8.452 & 12.59 \\ -66.44 & 11.296 & -8.889 & -0.3854 & 8.447 & 1.36 & 14.43 & 10.16 \\ -41.02 & 69.30 & -5.471 & -0.2371 & 6.649 & 1.249 & 0.01063 & 9.997 & 6.967 \\ -17.99 & 30.17 & -2.393 & -0.1035 & 6.059 & 2.216 & 21.39 & 3.554 \\ -34.51 & 58.04 & -4.596 & -0.1989 & 10.56 & 1.986 & 20.91 & 21.52 \end{bmatrix}. $$


In the discrete model, only the first and fifth state variables are used as outputs, i.e.,

$$C = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}$$

and the weighting matrices are chosen as $\tilde{Q} = 50I_2$ and $\tilde{R} = I_3$. 


3 Parameter-dependent problems of fixed size

Example 12

| $n$ | $m$ | $p$ | parameter | $\kappa(A)$ | $|\lambda_{\text{max}}^C|$ | $|X|$ | $\kappa(X)$ | $K_{\text{DARE}}$ |
|-----|-----|-----|-----------|--------------|----------------|------|------------|-------------|
| 2   | 1   | 2   | $\varepsilon = 100$ | $\infty$ | 0.00 | $1.00 \times 10^2$ | 1.00 | $1.00 \times 10^4$ | 2.65 |
|     |     |     | $\varepsilon = 10^6$ | $\infty$ | 0.00 | $1.00 \times 10^{12}$ | 1.00 | $1.00 \times 10^{12}$ | 2.65 |

Here, the matrix $A$ has a parameter and the coefficient matrices of the DARE (1) are:

$$A = \begin{bmatrix} 0 & \varepsilon \\ 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad R = 1, \quad Q = I_2.$$

The stabilizing solution is given by

$$X = \begin{bmatrix} 1 & 0 \\ 0 & 1 + \varepsilon^2 \end{bmatrix}$$

and the closed-loop spectrum is $\{0, 0\}$. For $\varepsilon = 100$, this is Example 2 from [11]. As $\varepsilon \rightarrow \infty$, this becomes an example of a DARE which is badly scaled in the sense of [27] due to the fact that $|A|_F \gg ||G||_F, ||Q||_F$. Obviously, the norm (and condition) of the stabilizing solution $X$ grow like $\varepsilon^2$ whereas the DARE condition number $K_{\text{DARE}}$ remains constant.

Example 13 [27]

| $n$ | $m$ | $p$ | parameter | $\kappa(A)$ | $|\lambda_{\text{max}}^C|$ | $|X|$ | $\kappa(X)$ | $K_{\text{DARE}}$ |
|-----|-----|-----|-----------|--------------|----------------|------|------------|-------------|
| 3   | 3   | 3   | $\varepsilon = 1$ | $\infty$ | 0.38 | 9.11 | 9.11 | 2.51 |
|     |     |     | $\varepsilon = 10^6$ | $\infty$ | 0.38 | $9.11 \times 10^6$ | 9.11 | 2.51 |

This example is constructed as follows. Let

$$A_0 = \text{diag}(0, 1, 3), \quad V = I - \frac{2}{3}v^T, \quad v^T = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}.$$  

Then

$$A = VA_0V, \quad G = \frac{1}{\varepsilon}I_3, \quad Q = \varepsilon I_3.$$  

A factorization $Q = C^T \tilde{Q}C$ can be obtained by setting $C := V$ and $\tilde{Q} := Q$; a factorization $G = BR^{-1}B^T$ is given by $B = I_3$ and $R = \varepsilon$. This is used in both the FORTRAN 77 and MATLAB implementations if a factored form is required.

As solution we get

$$X = V \text{diag}(x_1, x_2, x_3) V$$

where

$$x_1 = \varepsilon, \quad x_2 = \frac{(1 + \sqrt{5})}{2}, \quad x_3 = \frac{(9 + \sqrt{225})}{2}.$$  

The closed-loop spectrum is given by $\lambda_1 = 0, \lambda_2 = \frac{(3 - \sqrt{5})}{2}$, and $\lambda_3 = \frac{(11 - \sqrt{225})}{6}$.

For growing $\varepsilon$, the corresponding symplectic pencil (2) becomes more and more badly scaled which leads to a significant loss of accuracy in all DARE solvers based on eigenvalue methods. This demonstrates the need to use an appropriate scaling as proposed in [11].
The following system describes a very simple process control of a paper machine. The continuous-time model with a time delay is sampled at intervals of length $D$ which yields a singular transition matrix $A$. The time delay is equal to the length of three sampling intervals. The other parameters defining the system are a first-order time constant $\tau$ and the steady-state gain $K$. The linear system \((7) (8)\) is then given by

$$A = \begin{bmatrix}
1 - D/\tau & 0 & 0 & 0 \\
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0
\end{bmatrix}, \quad B = \begin{bmatrix}
KD/\tau \\
0 \\
0 \\
0
\end{bmatrix}, \quad C = \begin{bmatrix}
0 & 0 & 0 & 1
\end{bmatrix}.$$  

The weighting matrices used in this example are $R = r$ and $\tilde{Q} = 1$. Defining

$$\alpha = 1 - \frac{D}{\tau}, \quad \beta = \frac{KD}{\tau},$$

it can be shown that the solutions of the DARE \((1)\) are given by

$$X = \text{diag}(x_i, 1, 1, 1)$$

where $x_i$, $i = 1, 2$, solve the scalar quadratic equation

$$(\alpha^2 - 1)x + 1 - \frac{\alpha^2 \beta^2}{r + \beta^2 x} x^2 = 0,$$

whence

$$x_i = \frac{1}{2\beta^2} \left( \tau (\alpha^2 - 1) + \beta^2 \pm \sqrt{(\tau (\alpha^2 - 1) + \beta^2)^2 + 4\beta^2 r} \right). \quad (17)$$

The stabilizing positive semidefinite solution of \((1)\) is thus defined by the unique positive solution $x_1$ of \((17)\) and the closed-loop eigenvalues are

$$\lambda_1 = \frac{\alpha r}{r + \beta^2 x_1} = \frac{(\tau - D)\tau r}{\tau^2 r + (DK)^2 x_1}, \quad \lambda_2 = \lambda_3 = \lambda_4 = 0.$$  

Due to the variety of parameters in this example, it is possible to investigate DAREs with critical properties in many aspects. Since these properties merely rely on $\alpha$, $\beta$, and $r$, these effects can be produced by keeping $K$ and $D$ constant and varying $\tau$ (and $r$). Since $|\lambda_{\text{max}}| = \lambda_1$, for $\tau \gg D, K$ the largest closed-loop eigenvalue approaches the unit disk. For $\tau \ll D, K$ the norm and condition of $X$ become large and the DARE becomes ill conditioned with respect to $K_{\text{DARE}}$. 

| $n$ | $m$ | $p$ | parameter | $|A|$ | $|X|$ | $\kappa(X)$ | $K_{\text{DARE}}$
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>1</td>
<td>1</td>
<td>$\tau = 2.0, D = 1.0, K = 2.0, r = 0.25$</td>
<td>$\infty$</td>
<td>$9.62 \times 10^{-2}$</td>
<td>1.05</td>
<td>1.05</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$\tau = 10^8, D = 1.0, K = 1.0, r = 0.25$</td>
<td>$\infty$</td>
<td>$\approx 1 - \sqrt{7} \times 10^{-8}$</td>
<td>$5.69 \times 10^7$</td>
<td>$5.69 \times 10^7$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>$\tau = 10^{-6}, D = 1.0, K = 1.0, r = 0.25$</td>
<td>$\infty$</td>
<td>$2.0 \times 10^{-7}$</td>
<td>1.25</td>
<td>1.25</td>
</tr>
</tbody>
</table>
4 Examples of scalable size

Example 15 [24, Example 3]

| n | m | p | parameter | $\kappa(A)$ | $|\lambda_{\text{max}}|$ | $|X|$ | $\kappa(X)$ | $K_{\text{DARE}}$ |
|---|---|---|-----------|------------|----------------|------|--------------|-------------|
| n | 1 | n | $n = 10$, $r = 1$ | $\infty$ | 0.00 | 10.00 | 10.00 | 11.01 |
| n | 1 | n | $n = 100$, $r = 1$ | $\infty$ | 0.00 | 100.0 | 100.0 | 279.75 |

Consider the DARE defined by

$$ A = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & 0 \\ 0 & \cdots & 0 & 1 & 0 \end{bmatrix} \in \mathbb{R}^{n \times n}, \quad B = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}, \quad R = r, \quad Q = I_n. $$

The stabilizing solution has a very simple form, namely,

$$ X = \text{diag}(1, 2, \ldots, n). $$

The closed-loop eigenvalues are all zero, that is, the spectrum of the symplectic pencil $L - \lambda M$ in (2) is given by the generalized eigenvalues $\lambda_1 = \ldots = \lambda_n = 0$ and $\lambda_{n+1} = \ldots = \lambda_{2n} = \infty$.

This example can be used to test any DARE solver for growing dimension of the problem. The DARE condition number $K_{\text{DARE}}$ increases only slowly and for any order of the DARE, $|X| = \kappa(X) = n$.

Note further that the choice of $r$ does not influence the stabilizing solution $X$ but for $r < 1$, the condition number $K_{\text{DARE}}$ behaves like $1/r$. 

12
A The FORTRAN 77 subroutine DAREX

This is the prolog of a FORTRAN 77 subroutine for generating all presented examples. The subroutine was documented according to standards for SLICOT\(^4\) [23]. Besides calls to LAPACK and BLAS [2], DAREX calls the subroutines SP2SY and SY2SP which are used to convert symmetric matrices from general storage mode to packed storage mode and vice versa. These subroutines are provided together with \textit{darex.f}. If you have no access to LAPACK and BLAS, please contact the authors.

For some of the examples, DAREX reads the data from data files delivered together with \textit{darex.f}. These are Examples 6–9 and 11. The corresponding data files (in ASCII format) are \textit{DAREX6.DAT}, \textit{DAREX7.DAT}, \textit{DAREX8.DAT}, \textit{DAREX9.DAT}, and \textit{DAREX11.DAT}.

Note that the references given in the prolog of DAREX refer to those given at the end of the prolog and do not correspond to the references of this paper.

\begin{verbatim}
SUBROUTINE DAREX(NO, N, M, P, NPAR, DPAM, A, LDA, B, LDB, C, 1
LDC, Q, LDQ, R, LDR, S, LDS, X, LDX, NOTE, 2
STORE, WITHC, WITHG, WITHS, RWORK, IERR)
C
C PURPOSE
C
C To generate the benchmark examples for the numerical solution of the
C discrete-time algebraic Riccati equation (DARE)
C
C \begin{equation}
C 0 = A X A^T X - (A X B + S) (R + B X B) (B X A + S)^T + Q.
C \end{equation}
C
C as presented in [1]. Here, A, Q, X are real N-by-N matrices, B, S are
C N-by-M, and R is M-by-M. The matrices Q and R are symmetric and Q
C may be given in factored form
C
C \begin{equation}
C (I) \quad Q = C Q_0 C^T.
C \end{equation}
C
C Here, C is P-by-N and Q_0 is P-by-P. If R is nonsingular, the DARE
C can be rewritten equivalently as
C
C \begin{equation}
C 0 = X - A X (I_n + G X) A - Q
C \end{equation}
C
C where I_n is the N-by-N identity matrix and
C
C \begin{equation}
C (II) \quad G = B R B^T.
C \end{equation}
C
\end{verbatim}

\(^{4}\text{Subroutine Library in Control and Systems Theory}\)
ARGUMENT LIST

ARGUMENTS IN

NO - INTEGER.
The number of the benchmark example to generate according to [1].

N - INTEGER.
This integer determines the actual state dimension, i.e., the order of the matrix A as follows:
N is fixed for the examples of Sections 2 and 3 of [1], i.e., currently Examples 1-14.
NOTE that N is overwritten for Examples 1-14 and for the other example(s) if N is set by default.

M, P - INTEGER.
M is the number of columns in the matrix B and the order of the matrix \( R \) (in control problems, the number of inputs of the system).
P is the number of rows in the matrix C from (I) (in control problems, the number of outputs of the system).
Currently, M and P are fixed or determined by N for all examples and thus are not referenced on input.
NOTE that M and P are overwritten and M .LE. N and P .LE. N for all examples.

NPAR - INTEGER.
Number of input parameters supplied by the user.
Examples 1-11 (Section 2 of [1]) have no parameters.
Examples 12-13 (Section 3 of [1]) each have one DOUBLE PRECISION parameter which may be supplied in DPARAM(1).
Example 14 has 4 DOUBLE PRECISION parameters which may be supplied in DPARAM(1) - DPARAM(4).
Example 15 has one INTEGER parameter which determines the size of the problem. This parameter may be supplied in the input argument N. Besides, this example has one DOUBLE PRECISION parameter which may be supplied in DPARAM(1).
If the input value of NPAR is less than the number of parameters of the Example NO (according to [1]), the missing parameters are set by default.

DPARAM - DOUBLE PRECISION array of DIMENSION at least ndp.
Double precision parameter vector where ndp is the number of DOUBLE PRECISION parameters of Example NO (according to [1]). For all examples, ndp <= 4. For explanation of the parameters see [1].
DPARAM(1) defines the parameters 'epsilon' for the examples in Section 3 (NO = 12,13), the parameter 'tau' for NO = 14, and the parameter 'r' for NO = 15. For Example 14, DPARAM(2) - DPARAM(4) define in consecutive order 'D', 'K', and 'r'. If NPAR is smaller than the number of used parameters in Example NO (as described in [1]), default values are used and returned in corresponding components of DPARAM. NOTE that those entries of DPARAM are overwritten which are used to generate the example but were not supplied by the user.

LDA - INTEGER.
The leading dimension of array A as declared in the calling program.
LDA .GE. N where N is the order of the matrix A, i.e., the output value of the integer N.

LDB - INTEGER.
The leading dimension of array B as declared in the calling program.
LDB .GE. N (output value of N).

LDC - INTEGER.
The leading dimension of array C as declared in the calling program.
LDC .GE. P where P is either defined by default or depends upon N. (For all examples, P .LE. N, where N is the output value of the argument N.)

LDQ - INTEGER.
If full storage mode is used for Q, i.e., STORE = 'F' or 'f', then Q is stored like a 2-dimensional array with leading dimension LDQ. If packed symmetric storage mode is used, then LDQ is not referenced. That is, if STORE = 'F' or STORE = 'f', then LDQ .GE. N if WITHC = .FALSE.
LDQ .GE. P if WITHC = .TRUE.

LDR - INTEGER.
If full storage mode is used for the array R, i.e., STORE = 'F' or 'f', then R is stored like a 2-dimensional array with leading dimension LDR. If packed symmetric storage mode is used, then LDR is not referenced. That is, if STORE = 'F' or STORE = 'f', then LDR .GE. M if WITHG = .FALSE.
LDR .GE. N if WITHG = .TRUE.
LDS - INTEGER.
The leading dimension of array S as declared in the
calling program.
LDS .GE. N if S is to be returned (see MODE PARAMETER
WITHS). Otherwise, LDS is not referenced.

LDX - INTEGER.
The leading dimension of array X as declared in the
calling program.
LDX .GE. N if an exact solution is available (Examples
1,3,5,12-15). Otherwise, X is not referenced.

ARGUMENTS OUT

N - INTEGER.
The order of the matrix A.

M - INTEGER.
The number of columns of matrix B and the order of the
matrix R.

P - INTEGER.
The number of rows of the matrix C from (I).

DPARAM - DOUBLE PRECISION array of DIMENSION at least 7.
Double precision parameter vector. For explanation of the
parameters see [1].
DPARAM(1) defines the parameters 'epsilon' for the
examples in Section 3 (NO = 12,13), the parameter 'tau'
if NO = 14, and the parameter 'r' if NO = 15.
For Example 14, DPARAM(2) = DPARAM(4) define in
consecutive order 'D', 'K', and 'r'.

A - DOUBLE PRECISION array of DIMENSION (LDA,N).
The leading N by N part of this array contains the
coefficient matrix A of the DARE.

B - DOUBLE PRECISION array of DIMENSION (LDB,M).
If WITHG = .FALSE., then array B contains the coefficient
matrix B of the DARE.
Otherwise, B is used as workspace.

C - DOUBLE PRECISION array of DIMENSION (LDC,N).
If WITHC = .TRUE., then array C contains the matrix C of
the factored form (I) of Q.
Otherwise, C is used as workspace.
Q - DOUBLE PRECISION array of DIMENSION at least qdim.
   If STORE = 'F' or 'f', then qdim = LDQ*nq.
   If STORE = 'U', 'u', 'L' or 'l', then qdim = nq*(nq+1)/2.
   If WITHC = .FALSE., then nq = N and the array Q
   contains the coefficient matrix Q of the DARE.
   If WITHC = .TRUE., then nq = P and the array Q contains
   the matrix Q0 from (I).
   The symmetric matrix contained in array Q is stored
   according to MODE PARAMETER STORE.

R - DOUBLE PRECISION array of DIMENSION at least rdim.
   If STORE = 'F' or 'f' then rdim = LDR*nr.
   If STORE = 'U', 'u', 'L' or 'l' then rdim = nr*(nr+1)/2.
   If WITHG = .FALSE., then nr = M and the array R
   contains the coefficient matrix R of the DARE.
   If WITHG = .TRUE., then nr = N and the array R contains
   the matrix G from (II).
   The symmetric matrix contained in array R is stored
   according to MODE PARAMETER STORE.

X - DOUBLE PRECISION array of DIMENSION (LDX,xdim).
   If an exact solution is available (NO = 1,3,5,12-15),
   then xdim = N and the leading N-by-N part of this array
   contains the solution matrix X. Otherwise, X is not
   referenced.

NOTE - CHARACTER*70.
   String containing short information about the chosen
   example.

WORK SPACE

RWORK - DOUBLE PRECISION array of DIMENSION at least N*N.

MODE PARAMETERS

STORE - CHARACTER.
   Specifies the storage mode for arrays Q and R.
   STORE = 'F' or 'f': Full symmetric matrices are stored in
   Q and R, i.e., the leading N-by-N
   (M-by-M, P-by-P) parts of these
   arrays each contain a symmetric
   matrix.
   STORE = 'L' or 'l': Matrices contained in arrays Q and R
   are stored in lower packed mode, that
   is, the lower triangle of a k-by-k
(k=N,M,P) symmetric matrix is stored by columns, i.e., the matrix entry \( Q(i,j) \) is stored in the array entry \( Q(i+(2k-j)*(j-1)/2) \) for \( j \leq i \).

STORE = 'U' or 'u': Matrices contained in arrays Q and R are stored in upper packed mode, that is, the upper triangle of a k-by-k (k=N,M,P) symmetric matrix is stored by columns, i.e., the matrix entry \( G(i,j) \) is stored in the array entry \( G(i+j*(j-1)/2) \) for \( i \leq j \).

Otherwise, CAREX returns with an error.

WITHC - LOGICAL.
Indicates whether the matrices C, Q0 as in (I) are to be returned as follows.
WITHC = .TRUE., C is returned in array C and Q0 is returned in array Q.
WITHC = .FALSE., the coefficient matrix Q of the DARE is returned in array Q, whereas C and Q0 are not returned.

WITHG - LOGICAL.
Indicates whether the matrix G in (II) or the matrices B and R are returned as follows.
WITHG = .TRUE., the matrix G from (II) is returned in array R, whereas the matrices B and R are not returned.
WITHG = .FALSE., the coefficient matrices B and R of the DARE are returned in arrays B and R.

WITHS - LOGICAL.
Indicates whether the coefficient matrix S of the DARE is returned as follows.
WITHS = .TRUE., the coefficient matrix S of the DARE is returned in array S.
WITHS = .FALSE., the coefficient matrix S of the DARE is not returned.

ERROR INDICATOR

IERR - INTEGER.
Unless the routine detects an error (see next section), IERR contains 0 on exit.

WARNINGS AND ERRORS DETECTED BY THE ROUTINE
IERR = 1: (NO .LT. 1) or (NO .GT. NEX).
IERR = 2: (N .LT. 1) or (N .GT. LDA) or (N .GT. LDB) or
or (P .GT. LDC) or (WITHS and N .GT. LDS) or
N .GT. LDX and solution is available) or
((STORE = 'F' or STORE = 'f') and
((WITHC .EQ. .FALSE. and N .GT. LDQ) or
((WITH .EQ. .TRUE. and P .GT. LDQ)) or
((WITHG .EQ. .FALSE. and M .GT. LDR) or
= .TRUE. and N .GT. LDR))
IERR = 3: MODE PARAMETER STORE had an illegal value on input.
IERR = 4: Data file could not be opened or had wrong format.
IERR = 5: Division by zero.
IERR = 6: G can not be computed as in (II) due to a singular R
matrix. This error can only occur if
(WITHG .EQ. .TRUE.).

REFERENCES

A Collection of Benchmark Examples for the Numerical Solution
of Algebraic Riccati Equations II: Discrete-Time Case.
Technical Report SPC 95.23, Fak. f. Mathematik,
TU Chemnitz-Zwickau (Germany), December 1995.


CONTRIBUTOR

Peter Benner and Volker Mehrmann (TU Chemnitz-Zwickau)
Alan J. Laub (University of California, Santa Barbara)

KEYWORDS

discrete-time, algebraic Riccati equation, Hamiltonian matrix

REVISIONS

B The MATLAB function darex

The prolog of the MATLAB function darex is listed below. For all listed examples, it is possible to return the matrices $A$, $B$, $R$, $Q$, and the factors $C$, $Q = Q_0$. $G = BR^{-1}B^T$ can also be returned if $R$ is nonsingular. Otherwise, the output argument $G$ will contain an empty matrix. If the solution is not available, the output argument $X$ contains an empty matrix, too. Otherwise, $X$ is returned as well as the DARE condition number $K_{DARE}$ computed by the MATLAB function darecond.

Note that the references given in the prolog of darex refer to those given at the end of the prolog and do not correspond to the references of this paper.

```matlab
function [A,B,Q,R,S,X,parout,G,C,Q0]=darex(index,parin)

% DAREX
% Test examples for the discrete-time algebraic Riccati equation (DARE)
% 
% (I) 0 = DR(X) = A'XA - X - (A'XB + S) (R + B'XB) (B'XA + S') + Q
% 
% Here, $A$, $Q$, and $X$ are $n$-by-$n$ matrices, $B$ and $S$ are $n$-by-$m$, and $R$ is $m$-by-$m$. $Q$ and $R$ are symmetric and $X$ is the required solution matrix.
% One common approach to solve DAREs is to compute a deflating subspace of the symplectic pencil
% 
% \[-1\]
% \[
% \begin{pmatrix}
% A & B & R & S & 0 \\
% 0 & I & G \\
% \end{pmatrix}
% \]
% 
% (II) $L - s M := \begin{pmatrix} -1 & \phantom{0} \\ S & R & S' - Q & I \end{pmatrix} - s \begin{pmatrix} 0 & \phantom{0} \\ (A - B'R'S')' \end{pmatrix}$
% 
% \[-1\]
% 
% where $G = BRB'$ is a symmetric $n$-by-$n$ matrix. $Q$ may also be given in factored form, $Q = C'Q_0C$, where $C$ is a $p$-by-$n$ and $Q_0$ is a $p$-by-$p$ matrix.
% NOTE that for DAREs, $R$ being a singular matrix is not uncommon. In this case, the symplectic pencil cannot be formed as in (II), but a solution of the DARE can be computed via a deflating subspace of the extended pencil
% 
% \[-1\]
% \[
% \begin{pmatrix}
% A & 0 & B \\
% 0 & I & 0 & 0 \\
% \end{pmatrix}
% \]
% 
% (III) $LL - s MM := \begin{pmatrix} Q & -I & S \\
% S' & 0 & R \end{pmatrix} - s \begin{pmatrix} 0 & -A' & 0 \\
% 0 & -B' & 0 \end{pmatrix}$
% 
% For examples with singular $R$-matrix, $G$ can not be computed and is thus not returned.
% 
% Input:
% - index: number of example to generate, indices refer to example
numbers in [1].
% - parin: input parameters (optional, defaults values given in [1])
% For Example number
%   + 1-11: not referenced ([1], Section 2).
%   + 12-13: parin(1) = real-valued scalar.
%   + 14 : parin(1:4) = [tau, D, K, r], real-valued scalars.
%   + 15 : parin(1) = n = problem size.
%         parin(2) = r = real-valued scalar.
% Output:
%   - A, B, Q, R, S: coefficient matrices of DARE as in (I).
%   - X : exact solution of DARE (if available), usually the
%         stabilizing solution.
%         If an exact solution is not available, the empty matrix
%         is returned.
%   - parout : vector with system properties,
%     parout(1:3) = [n, m, p].
%     parout(4) = 2-norm condition number of A.
%     The following parameters are only returned if an
%     solution of the DARE is available:
%     parout(5) = radius of smallest circle enclosing the
%                 closed-loop spectrum.
%     parout(6) = 2-norm of X.
%     parout(7) = 2-norm condition number of X.
%     parout(8) = condition number of DARE as defined in [2].
%   - G, C, Q0 : optional output matrices as defined above. NOTE that
%     G can only be computed if R is nonsingular. Otherwise,
%     G contains on output the empty matrix.
% References:
%     Examples for the Numerical Solution of Algebraic Riccati
%     Equations II: Discrete-Time Case', Tech. Report SPC 95-23,
%     Fak. f. Mathematik, TU Chemnitz-Zwickau (Germany), December 1995.
%     Algebraic Riccati Equation to Enhance Stability of the Schur
%     Solution Method', IEEE Transactions on Automatic Control, vol. 37,
%     no. 4, pp. 513-518, 1992.

Peter Benner, Volker Mehrmann (TU Chemnitz-Zwickau, Germany),
Alan J. Laub (University of California at Santa Barbara)
12-14-1995
C How to obtain the software

The codes corresponding to this paper may be obtained via anonymous ftp at TU Chemnitz-Zwickau. Proceed as follows.

> ftp ftp.tu-chemnitz.de
> Name: anonymous
> Password: your complete e-mail address
> cd pub/Local/mathematik/Benner

Observe the capital “L” in Local!

Now get the compressed FORTRAN 77 subroutines darex.f, sp2sy.f, sy2sp.f, data files, a sample Makefile, and a sample program example.f together with an introductory README file by

> get darex.f.tar.Z

or the compressed MATLAB function files darex.m, darecond.m and an introductory README file by

> get darex.m.tar.Z

After exiting ftp, extracting the MATLAB codes and data files is achieved by the following commands:

> uncompress darex.m.tar.Z
> tar xf darex.m.tar

Analogously, the FORTRAN 77 codes and corresponding data files are obtained by

> uncompress darex.f.tar.Z
> tar xf darex.f.tar

In both cases, the command tar xf creates a directory containing all required files. For darex.m.tar.Z, this directory is called darex.m and for darex.f.tar.Z, it will be darex.f. If any problems occur in obtaining or running the codes, please contact one of the authors.
Table I

<table>
<thead>
<tr>
<th>X</th>
<th>( p )</th>
<th>( \alpha )</th>
<th>( \beta )</th>
<th>( \gamma )</th>
<th>( \delta )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>2</td>
<td>2.000</td>
<td>2.000</td>
<td>2.000</td>
<td>2.000</td>
<td>2.000</td>
</tr>
<tr>
<td>3</td>
<td>3.000</td>
<td>3.000</td>
<td>3.000</td>
<td>3.000</td>
<td>3.000</td>
</tr>
<tr>
<td>4</td>
<td>4.000</td>
<td>4.000</td>
<td>4.000</td>
<td>4.000</td>
<td>4.000</td>
</tr>
</tbody>
</table>

Table I summarizes the properties of all the presented examples. A value \( \infty \) for a condition

\( D \) is available for this example. The column \( \alpha \) indicates whether an analytical stabilizing solution

\( \Rightarrow \) defined for the example. The column \( \beta \) indicates whether a numerical stabilizing solution

\( \Rightarrow \) and condition number computed by MATLAB \( \sqrt{\|D\| \|D^T\|} \) is not

\( \Rightarrow \) infinite that the corresponding matrix is not invertible with respect to the numerical

\( \Rightarrow \) number of all the presented examples. A value \( \infty \) for a condition

\( \Rightarrow \)
References


Other titles in the SPC series:

95_9 H. Matthes. A DD preconditioner for the clamped plate problem. February 1995
95_11 M. Bollhöfer. Algebraic Domain Decomposition. March 1995
95_15 B. Heinrich, B. Lang, B. Weber. Parallel computation of Fourier-finite-element approximations and some experiments. May 1995
95_16 W. Rath. Canonical forms for linear descriptor systems with variable coefficients. May 1995
95_17 C. He, A. J. Laub, V. Mehrmann. Placing plenty of poles is pretty preposterous. May 1995
95_18 J. J. Hench, C. He, V. Kučera, V. Mehrmann. Dampening controllers via a Riccati equation approach. May 1995


Some papers can be accessed via anonymous ftp from server ftp.tu-chemnitz.de, directory pub/Local/mathematik/SPC. (Note the capital L in Local!)

The complete list of current and former preprints is available via http://www.tu-chemnitz.de/~pester/sfb/spc95pr.html.