ME 7247: Advanced Control Systems **Supplementary notes**

Algebraic Riccati Equations

v.1.0 (11.07.2022)

In these notes, we derive conditions under which the Discrete Algebraic Riccati Equation (DARE) has a stabilizing solution. This result will also lead to a method for solving the DARE.

1 The DARE

The DARE and its associated closed-loop matrix H are given by

$$A^\mathsf{T} X A - X + Q - (A^\mathsf{T} X B + S)(B^\mathsf{T} X B + R)^{-1}(B^\mathsf{T} X A + S^\mathsf{T}) = 0 \tag{1a}$$

$$H := A - B(B^\mathsf{T} X B + R)^{-1}(B^\mathsf{T} X A + S^\mathsf{T}) \tag{1b}$$

We assume Q and R are symmetric. Moreover, $\left[\begin{smallmatrix} Q & S \\ S^\mathsf{T} & R \end{smallmatrix}\right] \succeq 0$ and $R \succ 0$. Now define:

$$E := A - BR^{-1}S^\mathsf{T}, \qquad \bar{Q} := Q - SR^{-1}S^\mathsf{T}, \qquad G := BR^{-1}B^\mathsf{T}.$$

After some algebraic manipulations, we can rewrite Eq. (1) as

$$E^\mathsf{T} X E - X + \bar{Q} - E^\mathsf{T} X B (B^\mathsf{T} X B + R)^{-1} B^\mathsf{T} X E = 0 \tag{2a}$$

$$H = E - B(B^\mathsf{T} X B + R)^{-1} B^\mathsf{T} X E \tag{2b}$$

After even more manipulations, we can rewrite Eq. (2) as

$$E^\mathsf{T}(I + XG)^{-1} X E - X + \bar{Q} = 0 \tag{3a}$$

$$H = (I + GX)^{-1} E \tag{3b}$$

Substituting (3b) into (3a), and rearranging (3b), we obtain the pair of equations

$$\begin{cases} E^\mathsf{T} X H - X + \bar{Q} = 0 \\ H + GXH - E = 0 \end{cases} \iff \begin{bmatrix} I & G \\ 0 & E^\mathsf{T} \end{bmatrix} \begin{bmatrix} I \\ X \end{bmatrix} H = \begin{bmatrix} E & 0 \\ -\bar{Q} & I \end{bmatrix} \begin{bmatrix} I \\ X \end{bmatrix} \tag{4}$$

So finding (X, H) that satisfy (3) is equivalent to finding (X, H) that satisfy (4). Now consider a similarity transform $H = P_1 J P_1^{-1}$. Setting $P_2 := X P_1$, we can rewrite (4) as

$$\begin{bmatrix} I & G \\ 0 & E^\mathsf{T} \end{bmatrix} \begin{bmatrix} P_1 \\ P_2 \end{bmatrix} J = \begin{bmatrix} E & 0 \\ -\bar{Q} & I \end{bmatrix} \begin{bmatrix} P_1 \\ P_2 \end{bmatrix}. \tag{5}$$

So finding (X, H) that satisfy (4) is equivalent to finding (P_1, P_2, J) with P_1 invertible that satisfies (5). We can construct the solution to (4) using $X = P_2 P_1^{-1}$ and $H = P_1 J P_1^{-1}$. Eq. (4) is related to the following *generalized eigenvalue problem.*

$$\lambda \underbrace{\begin{bmatrix} I & G \\ 0 & E^\mathsf{T} \end{bmatrix}}_{L} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \underbrace{\begin{bmatrix} E & 0 \\ -\bar{Q} & I \end{bmatrix}}_{M} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \tag{6}$$

Background on generalized eigenvalues. The generalized eigenvalue problem for the pair (M, L) is to find a nonzero $v \in \mathbb{C}^n$ (generalized eigenvector) and $\lambda \in \mathbb{C}$ (generalized eigenvalue) such that $\lambda L v = M v$. If L is invertible, we can left-multiply by L^{-1} and see that the generalized eigenvalues of (M, L) are just the ordinary eigenvalues of $L^{-1}M$. In this case, (6) becomes

$$\begin{bmatrix} E + GE^{-\mathsf{T}}\bar{Q} & -GE^{-\mathsf{T}} \\ -E^{-\mathsf{T}}\bar{Q} & E^{-\mathsf{T}} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \lambda \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}.$$

This matrix is *symplectic* and has nice properties. We can derive similar properties working directly from (6) (without assuming E is invertible) so we will do that instead. For any invertible matrices U and V, the generalized eigenvalues of (M, L) are the same as the generalized eigenvalues of (UMV, ULV). Similar to ordinary eigenvalues, $\lambda L - M$ is rank-deficient and all generalized eigenvalues λ satisfy the characteristic polynomial $\det(\lambda L - M) = 0$.

Consider the scalar case where $M = \alpha$ and $L = \beta$. The characteristic polynomial is $\lambda\beta - \alpha$. There are three possibilities for this generalized eigenvalue problem.

(i) $\beta \neq 0$: Here, $\lambda = \alpha/\beta$ so we have a finite generalized eigenvalue.

(ii) $\beta = 0$ and $\alpha \neq 0$. There is no solution. By convention, we say $\lambda = \infty$ (eigenvalue at infinity).

(iii) $\beta = 0$ and $\alpha = 0$. This is a degenerate problem, where any $\lambda \in \mathbb{C}$ is a generalized eigenvalue.

Every square matrix has a *Schur decomposition*, $A = QUQ^*$, where Q is unitary[1] and U is upper-triangular. In such a decomposition, the diagonal entries of U are the eigenvalues of A. Every pair of square matrices (A, B) has a *generalized Schur decomposition* (also called the *QZ factorization*), $A = QSZ^*$ and $B = QTZ^*$, where Q and Z are unitary, and S and T are upper triangular. It follows that (A, B) has the same generalized eigenvalues as (S, T), so since both are upper triangular, the characteristic polynomial $\det(\lambda T - S)$ can be factored as $\Pi_i(\beta_i\lambda - \alpha_i)$ and we can consider the three cases above for each pair (α_i, β_i) separately. We conclude that the pair (M, L) has:

- r finite eigenvalues at α_i/β_i, corresponding to the cases where $\beta_i \neq 0$.
- $n - r$ eigenvalues at infinity, corresponding to the cases where $\beta_i = 0$ and $\alpha_i \neq 0$.
- If $\alpha_i = \beta_i = 0$ for any i, then the entire generalized eigenvalue problem is degenerate, the characteristic polynomial is identically zero, and every $\lambda \in \mathbb{C}$ is a generalized eigenvalue.

All solutions to the DARE. Based on the derivations above, it is clear that we can generate any solution to the DARE (1) by finding a solution (P_1, P_2, J) to (5) for which P_1 is invertible. We can do this by solving the generalized eigenvalue problem (6). For example, we could pick n generalized eigenvectors and stack them vertically to form a matrix $P \in \mathbb{R}^{2n \times n}$. As long as P_1 is invertible, we will have a solution to (5) and hence to (1).

Stabilizing solutions. We are interested in solutions of the DARE where H (and hence, J) are Schur-stable, which are the so-called *stabilizing solutions*. Our goal from now on will be to derive conditions on the matrices (A, B, Q, R, S) that ensure that a stabilizing solution exists. We will also derive other useful properties such as uniqueness and definiteness along the way.

[1] A matrix $Q \in \mathbb{C}^{n \times n}$ is *unitary* if $Q^*Q = QQ^* = I$, where Q^* denotes the conjugate transpose. Unitary matrices are analogous to orthogonal matrices, but for complex matrices. If Q is a real and unitary, then it is orthogonal.

Our first result is that the special structure of the generalized eigenvalue problem (6) leads to a special structure in the generalized eigenvalues.

Lemma 1. *If $\lambda \neq 0$ is a generalized eigenvalue of* (6)*, then so is λ^{-1}. Moreover, λ and λ^{-1} have the same algebraic multiplicity.*

Proof. Define $\Omega := \left[\begin{smallmatrix} 0 & I \\ -I & 0 \end{smallmatrix}\right]$. We can check that:

$$L\Omega L^{\mathsf{T}} = M\Omega M^{\mathsf{T}} = \begin{bmatrix} 0 & E \\ -E^{\mathsf{T}} & 0 \end{bmatrix}$$

Since $\det(\lambda L - M) = 0$, we have $\det(\lambda L^{\mathsf{T}} - M^{\mathsf{T}}) = 0$. So there exists $w \neq 0$ such that $\lambda L^{\mathsf{T}} w = M^{\mathsf{T}} w$. Multiply both sides on the left by $M\Omega$ and obtain: $\lambda M\Omega L^{\mathsf{T}} w = M\Omega M^{\mathsf{T}} w = L\Omega L^{\mathsf{T}} w$. Define $z := \Omega L^{\mathsf{T}} w$ and this becomes $\lambda M z = L z$. Note that $z \neq 0$, because otherwise we would have $L^{\mathsf{T}} w = 0$ and $M^{\mathsf{T}} w = 0$. But the former implies $w_1 = 0$ and the latter implies $w_2 = 0$, contradicting $w \neq 0$. Consequently, we have $\det(L - \lambda M) = 0$.

If we perform a QZ factorization of (M, L), then a generalized eigenvalue λ with multiplicity r will correspond to a factor $(\lambda\beta_i - \alpha_i)^r$ in $\det(\lambda L - M)$ and a factor $(\beta_i - \lambda\alpha_i)^r$ in $\det(L - \lambda M)$. Therefore $1/\lambda$ is a generalized eigenvalue with multiplicity r as well. ∎

Theorem 1. *Suppose $\left[\begin{smallmatrix} Q & S \\ S^{\mathsf{T}} & R \end{smallmatrix}\right] \succeq 0$ and $R \succ 0$. The following statements are equivalent.*

(i) (A, B) is stabilizable and $(E, \bar{Q})$ has no unobservable modes on the unit circle.

(ii) (A, B) is stabilizable and the matrix $\left[\begin{smallmatrix} A-\lambda I & B \\ Q & S \\ S^{\mathsf{T}} & R \end{smallmatrix}\right]$ has full column rank for all $|\lambda| = 1$.

(iii) Eq. (5) *has a solution for which J is Schur-stable and P_1 is invertible.*

In this case, $X = P_2 P_1^{-1}$ and $H = P_1 J P_1^{-1}$ is the unique stabilizing solution to the DARE (1) *and $X \succeq 0$. Moreover, $X \succ 0$ if and only if $(E, \bar{Q})$ has no unobservable stable modes.*

Note. Observability means that there are no unobservable modes. Detectability means that there are no unstable unobservable modes. The condition in Theorem 1 is even weaker than detectability, since it only requires that there are no unobservable modes on the unit circle.

Proof. **(i) $\iff$ (ii).** Suppose $(E, \bar{Q})$ has an unobservable mode on the unit circle. This is equivalent to the existence of a $v \neq 0$ and $|\lambda| = 1$ satisfying $Ev = \lambda v$ and $\bar{Q} v = 0$. Substitute the definitions for E and $\bar{Q}$ and these equations become

$$\begin{bmatrix} A - BR^{-1}S^{\mathsf{T}} - \lambda I \\ Q - SR^{-1}S^{\mathsf{T}} \end{bmatrix} v = 0. \tag{7}$$

Since $R \succ 0$, we can define $w := -R^{-1}S^{\mathsf{T}} v$. Then (7) has a nonzero solution if and only if

$$\begin{bmatrix} A - \lambda I & B \\ Q & S \\ S^{\mathsf{T}} & R \end{bmatrix} \begin{bmatrix} v \\ w \end{bmatrix} = 0 \tag{8}$$

has a nonzero solution, or equivalently, the matrix in (8) has full column rank.

(i) $\implies$ (iii). First, we will show that if (A, B) is stabilizable and $(E, \bar{Q})$ is detectable, then generalized eigenvalues of (6) satisfy $|\lambda| \neq 1$.

Suppose instead that $|\lambda| = 1$ and λ satisfies (6). Then we have

$$\lambda v_1 + \lambda G v_2 = E v_1 \tag{9a}$$
$$\lambda E^\mathsf{T} v_2 = -\bar{Q} v_1 + v_2 \tag{9b}$$

Evaluate $\bar{\lambda} v_2^*$(9a) + (9b)$^* v_1$ and obtain $v_1^* \bar{Q} v_1 + v_2^* G v_2 = 0$. Since $G \succeq 0$ and $\bar{Q} \succeq 0$, we conclude that $\bar{Q} v_1 = 0$ and $G v_2 = 0$. The latter implies that $B^\mathsf{T} v_2 = 0$. Substituting these findings back into (9), we obtain $E v_1 = \lambda v_1$ and $\lambda E^\mathsf{T} v_2 = v_2$, which simplifies to $A^\mathsf{T} v_2 = \lambda^{-1} v_2$ ($\lambda \neq 0$ since $|\lambda| = 1$). Since $v \neq 0$, then either $v_1 \neq 0$ or $v_2 \neq 0$.

- If $v_2 \neq 0$, then we have $|\lambda| \geq 1$ satisfying $A^\mathsf{T} v_2 = \lambda^{-1} v_2$ and $B^\mathsf{T} v_2 \neq 0$, so from the PBH test, (A, B) is not stabilizable.
- If $v_1 \neq 0$, then we have $|\lambda| = 1$ satisfying $E v_1 = \lambda v_1$ and $\bar{Q} v_1 \neq 0$, so from the PBH test, $(E, \bar{Q})$ has an unobservable mode on the unit circle.

This contradicts our assumptions in (i), so we conclude that $|\lambda| \neq 1$.

The generalized eigenvalue problem (6) cannot be degenerate, since we just showed that not every λ can be a generalized eigenvalue. Consequently, we deduce that the generalized eigenvalues must consist of exactly n eigenvalues satisfying $|\lambda| < 1$, and n eigenvalues satisfying $|\lambda| > 1$ (with some possibly at infinity). Let $P \in \mathbb{R}^{2n \times n}$ be a matrix whose columns form a basis for the n-dimensional subspace spanned by the corresponding generalized eigenvectors. Then P satisfies (5), and the corresponding J is Schur-stable. It remains to show that P_1 is invertible.

Before doing this, we first prove that $P_1^\mathsf{T} P_2 \succeq 0$. Eq. (5) yields the equations

$$P_1 J + G P_2 J = E P_1 \tag{10a}$$
$$E^\mathsf{T} P_2 J = -\bar{Q} P_1 + P_2 \tag{10b}$$

From (10b) and substituting (10a), we have:

$$\begin{aligned} P_1^\mathsf{T} P_2 &= P_1^\mathsf{T} \bar{Q} P_1 + P_1^\mathsf{T} E^\mathsf{T} P_2 J \\ &= P_1^\mathsf{T} \bar{Q} P_1 + (P_1 J + G P_2 J)^\mathsf{T} P_2 J \\ &= P_1^\mathsf{T} \bar{Q} P_1 + J^\mathsf{T} P_1^\mathsf{T} P_2 J + J^\mathsf{T} P_2^\mathsf{T} G P_2 J \end{aligned}$$

Rearrange to obtain a Lyapunov equation:

$$J^\mathsf{T} (P_1^\mathsf{T} P_2) J - (P_1^\mathsf{T} P_2) + \underbrace{P_1^\mathsf{T} \bar{Q} P_1 + J^\mathsf{T} P_2^\mathsf{T} G P_2 J}_{\succeq 0} = 0$$

Since J is Schur-stable, we conclude that $P_1^\mathsf{T} P_2$ is symmetric and positive semidefinite.

Now we prove that P_1 is invertible. Suppose instead that P_1 is singular, so there exists some $x \neq 0$ such that $P_1 x = 0$. Multiply (10a) on the right by x and obtain

$$P_1 J x + G P_2 J x = 0. \tag{11}$$

Multiply (11) on the left by $x^\mathsf{T} J^\mathsf{T} P_2^\mathsf{T}$ and obtain $x^\mathsf{T} (J^\mathsf{T} P_2^\mathsf{T} P_1 J) x + x^\mathsf{T} J^\mathsf{T} P_2^\mathsf{T} G P_2 J x = 0$. Since $P_1^\mathsf{T} P_2 \succeq 0$, and $G \succeq 0$ by assumption, both terms must be zero. So $G P_2 J x = 0$. But from (11), this implies $P_1 J x = 0$. So if $x \in \operatorname{null}(P_1)$, we conclude that $J x \in \operatorname{null}(P_1)$ ($\operatorname{null}(P_1)$ is J-invariant).

This means we can find an eigenpair (v, λ) so that $Jv = \lambda v$ and $0 \neq v \in \text{null}(P_1)$. To see why this is the case, say A is S-invariant. This means if we let U_1 be an orthonormal basis for S and U_2 be its orthogonal completion, we have $AU_1 = U_1 A_{11}$ for some matrix A_{11}. Consequently,

$$A \begin{bmatrix} U_1 & U_2 \end{bmatrix} = \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix}$$

So if we let (λ, v) be such that $A_{11} v = \lambda v$, then:

$$A(U_1 v) = A \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} v \\ 0 \end{bmatrix} = \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix} \begin{bmatrix} v \\ 0 \end{bmatrix} = U_1 A_{11} v = \lambda (U_1 v)$$

Therefore, let (λ, v) be such that $Jv = \lambda v$, $v \neq 0$, and $P_1 v = 0$. Multiplying (10b) by v, we obtain:

$$\lambda E^{\mathsf{T}} P_2 v = P_2 v \tag{12}$$

Note that $P_2 v \neq 0$ because otherwise we would have $P_1 v = 0$ and $P_2 v = 0$, which is impossible since our generalized eigenvectors must be linearly independent. So $\bar{\lambda}^{-1}$ is an eigenvalue of E.

Using the same argument we did with x, we conclude that since $P_1 v = 0$, we must have $G P_2 v = 0$, so $B^{\mathsf{T}} P_2 v = 0$. We also have from (12) that $E^{\mathsf{T}} P_2 v = A^{\mathsf{T}} P_2 v = \lambda^{-1} P_2 v$. So $A^{\mathsf{T}} P_2 v = \lambda^{-1} P_2 v$ and $B^{\mathsf{T}} P_2 v = 0$ and $P_2 v \neq 0$, with $|\lambda^{-1}| > 1$ (since $|\lambda| < 1$ because J is Schur-stable). This contradicts the stabilizability of (A, B). So we conclude that P_1 must be invertible. Using $X = P_2 P_1^{-1}$ and $H = P_1 J P_1^{-1}$, we have constructed a solution to the DARE (1).

The stabilizing solution must be unique, because all solution arise from picking an n-dimensional invariant subspace in (5), and there are exactly n stable generalized eigenvalues, so our choice is fixed. The solution $X = P_2 P_1^{-1}$ is invariant under different choices of coordinates for the basis, because any invertible transformation $P \mapsto PT$ leads to $X \mapsto (P_2 T)(P_1 T)^{-1} = P_2 P_1^{-1} = X$.

Now we prove definiteness. Write $X = P_2 P_1^{-1} = P_1^{-\mathsf{T}} (P_1^{\mathsf{T}} P_2) P_1^{-1}$. Since $P_1^{\mathsf{T}} P_2 \succeq 0$, it follows that $X \succeq 0$. Now we investigate conditions under which $X \succ 0$. Rewrite the DARE (3) as

$$H^{\mathsf{T}} X H - X + \left(\bar{Q} + E^{\mathsf{T}} (I + XG)^{-1} X G X (I + GX)^{-1} E \right) = 0$$

This is a Lyapunov equation with H stable and the constant term in brackets is positive semidefinite. So from the main Lyapunov theorem, $X \succ 0$ if and only if $(H, \bar{Q} + E^{\mathsf{T}} (I + XG)^{-1} X G X (I + GX)^{-1} E)$ is observable. By the PBH test, this fails if and only if there exists some (λ, v) such that

$$Hv = \lambda v \quad \text{and} \quad \left(\bar{Q} + E^{\mathsf{T}} (I + XG)^{-1} X G X (I + GX)^{-1} E \right) v = 0$$

Since H is stable, this forces $|\lambda| \leq 1$. Multiply the right equation on the left by v^* and obtain

$$\bar{Q} v = 0 \quad \text{and} \quad G X (I + GX)^{-1} E v = 0 \tag{13}$$

Now rearrange the other equation using (3b) and apply (13).

$$Hv = (I + GX)^{-1} E v = \left(I - G X (I + GX)^{-1} \right) E v = E v$$

Therefore $X \succ 0$ if and only if $(E, \bar{Q})$ has no stable unobservable modes.

(iii) $\Longrightarrow$ (i). Suppose (6) has a solution for which J is Schur-stable and P_1 is invertible. Then we can use $X = P_2 P_1^{-1}$ and $H = P_1 J P_1^{-1}$ to construct a stabilizing solution to the DARE (1). From (1b), $H = A + BK$, where $K := -(B^{\mathsf{T}} X B + R)^{-1} (B^{\mathsf{T}} X A + S^{\mathsf{T}})$. Therefore (A, B) is stabilizable. Suppose by contradiction that $(E, \bar{Q})$ has an unobservable mode on the unit circle. Then there is some $w \neq 0$ and $|\lambda| = 1$ with $Ew = \lambda w$ and $\bar{Q} w = 0$. By letting $v_1 = w$ and $v_2 = 0$, we satisfy (6), so λ is an eigenvalue of J, which contradicts Schur-stability of J. ■

Theorem 1 says that the stabilizing solution is unique, and that it is positive semidefinite. However, it does not say the converse. Namely, a positive semidefinite solution need not be stabilizing. However, if we slightly strengthen the assumptions of Theorem 1, we can make this claim.

Corollary 1. *Suppose* $\begin{bmatrix} Q & S \\ S^\mathsf{T} & R \end{bmatrix} \succeq 0$ *and* $R \succ 0$*. Further suppose* (A, B) *is stabilizable and* $(E, \bar{Q})$ *is detectable. Then the DARE* (1) *has a unique positive semidefinite solution. Moreover, that solution is stabilizing. This solution is positive definite if and only if* $(E, \bar{Q})$ *is observable.*

Proof. From Theorem 1, we know there is a unique stabilizing solution and that this solution is positive semidefinite. It remains to show that any positive semidefinite solution must be stabilizing. As in the proof of Theorem 1, rewrite the DARE (3) as

$$H^\mathsf{T} X H - X + \left(\bar{Q} + E^\mathsf{T}(I + XG)^{-1} XGX(I + GX)^{-1} E\right) = 0. \tag{14}$$

This is a Lyapunov equation and the constant term in brackets is positive semidefinite. Suppose (λ, v) is an unobservable mode of $(H, \bar{Q} + E^\mathsf{T}(I + XG)^{-1}XGX(I + GX)^{-1}E)$. Since the second term is positive semidefinite, this means that

$$Hv = \lambda v, \qquad \bar{Q}v = 0, \qquad \text{and} \quad GX(I + GX)^{-1}Ev = 0.$$

Manipulating the definition for H, we have:

$$Hv = (I + GX)^{-1}Ev = \left(I - GX(I + GX)^{-1}\right) Ev = Ev.$$

Therefore, $Ev = \lambda v$ and $\bar{Q}v = 0$. In other words, (λ, v) is an unobservable mode of $(E, \bar{Q})$. Properties like detectability and observability involve the *absence* of unobservable modes in certain regions of the complex plane, so since $(E, \bar{Q})$ is detectable, we have that $(H, \bar{Q} + E^\mathsf{T}(I + XG)^{-1}XGX(I + GX)^{-1}E)$ is detectable. Applying the main Lyapunov theorem to (14), we conclude that H must be stable, so X is a stabilizing solution to the DARE (1).

From Theorem 1, $X \succ 0$ if and only if $(E, \bar{Q})$ has no unobservable stable modes. Since $(E, \bar{Q})$ is already detectable, it has no unobservable unstable modes. Therefore, $X \succ 0$ if and only if $(E, \bar{Q})$ has no unobservable modes, i.e., $(E, \bar{Q})$ is observable. ■

Note on numerical stability. Solving the generalized eigenvalue problem (6) requires computing E and $\bar{Q}$, which requires inverting R. Typically $R \succ 0$, but in some cases, we may have an application where $R \succeq 0$ so R will not be invertible. Even if R is invertible, it may not always be desirable to invert it. Eq. (6) can be transformed using the fact that $\lambda L - M$ is a Schur complement in the (A, B, Q, S, R) coordinates. Solving (6) is equivalent to solving

$$\lambda \begin{bmatrix} I & 0 & 0 \\ 0 & A^\mathsf{T} & 0 \\ 0 & B^\mathsf{T} & 0 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} A & 0 & B \\ Q & I & S \\ S^\mathsf{T} & 0 & R \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}. \tag{15}$$

For example, the Matlab DARE solver `idare` solves the generalized eigenvalue problem (15) rather than computing E and $\bar{Q}$ and solving (6).

2 Convergence of the Riccati difference equation

Now we prove that the Riccati difference equation converges to the stabilizing solution.

Theorem 2. *Suppose* $\begin{bmatrix} Q & S \\ S^\mathsf{T} & R \end{bmatrix} \succeq 0$, $R \succ 0$, (A, B) *stabilizable and* $(E, \bar{Q})$ *has no unobservable modes on the unit circle, then the discrete Riccati difference equation (RDE)*

$$X_{t+1} = A^\mathsf{T} X_t A + Q - (A^\mathsf{T} X_t B + S)(B^\mathsf{T} X_t B + R)^{-1}(B^\mathsf{T} X_t A + S^\mathsf{T}) \tag{16}$$

converges to the stabilizing solution of the DARE for any initial condition $X_0 \succ 0$.

Proof. **Step 1.** X_t is bounded below. We have

$$\begin{bmatrix} A^\mathsf{T} X_t A + Q & A^\mathsf{T} X_t B + S \\ B^\mathsf{T} X_t A + S^\mathsf{T} & B^\mathsf{T} X_t B + R \end{bmatrix} = \begin{bmatrix} A & B \end{bmatrix}^\mathsf{T} X_0 \begin{bmatrix} A & B \end{bmatrix} + \begin{bmatrix} Q & S \\ S^\mathsf{T} & R \end{bmatrix} \succeq 0$$

since we assumed $X_0 \succeq 0$ and $\begin{bmatrix} Q & S \\ S^\mathsf{T} & R \end{bmatrix} \succeq 0$. Since $R \succ 0$, this is equivalent to the Schur complement being positive semidefinite, so $X_1 \succeq 0$. Iterating, we obtain $X_t \succeq 0$ for all t.

Step 2. X_t is bounded above. Using the same approach as in Section 1, rewrite the RDE as

$$X_{t+1} = E^\mathsf{T}(I + X_t G)^{-1} X_t E + \bar{Q}. \tag{17}$$

Define $\Delta_t := X_t - X$, where $X \succeq 0$ is the stabilizing solution to the DARE. Subtracting (3a) from (17) and substituting $H := (I + GX)^{-1}E$ from (3b), we obtain an expression for the error Δ_{t+1}:

$$\begin{aligned} \Delta_{t+1} &= E^\mathsf{T}\Big((I + X_t G)^{-1} X_t - (I + XG)^{-1} X\Big)E \\ &= H^\mathsf{T}\Big((I + XG)(I + X_t G)^{-1} X_t (I + GX) - (I + XG)X\Big)H \\ &= H^\mathsf{T}\Big((I + XG)(I + X_t G)^{-1}\big(X_t(I + GX) - (I + X_t G)X\big)\Big)H \\ &= H^\mathsf{T}(I + XG)(I + X_t G)^{-1}\Delta_t H \\ &= H^\mathsf{T}\Big(\Delta_t - \Delta_t G^{1/2}\big(I + G^{1/2} X_t G^{1/2}\big)^{-1} G^{1/2}\Delta_t\Big)H \\ &\preceq H^\mathsf{T}\Delta_t H \end{aligned} \tag{18}$$

Therefore, we have $\Delta_t \preceq (H^\mathsf{T})^t \Delta_0 H^t \to 0$ since H is Schur-stable. Therefore, Δ_t, and hence X_t, is uniformly bounded: $0 \preceq X_t \preceq \bar{X}$ for some $\bar{X}$.

Step 3. Analogously to H, define $H_t := A + BK_t$, where $K_t := -(B^\mathsf{T} X_t B + R)^{-1}(B^\mathsf{T} X_t A + S^\mathsf{T})$. The DARE (1) can be rewritten as

$$X_{t+1} = H_t^\mathsf{T} X_t H_t + \begin{bmatrix} I \\ K_t \end{bmatrix}^\mathsf{T} \begin{bmatrix} Q & S \\ S^\mathsf{T} & R \end{bmatrix} \begin{bmatrix} I \\ K_t \end{bmatrix} \succeq H_t^\mathsf{T} X_t H_t$$

Define the transition matrix $\Psi_t := H_0 H_1 \cdots H_{t-1}$. Iterating the above, we have $X_t \succeq \Psi_t^\mathsf{T} X_0 \Psi_t$. Since X_t is bounded above, we conclude that Ψ_t must also be bounded.

Step 4. The error recursion (18) can take the simple form $\Delta_{t+1} = H^\mathsf{T}\Delta_t H_t$. Recursing, we obtain $\Delta_t = (H^\mathsf{T})^t \Delta_0 \Psi_t$. Since Ψ_t is bounded and H is Schur-stable, we conclude that $\Delta_t \to 0$. ■

3 Riccati inequalities

Discrete Algebraic Riccati Equations (DARE) satisfy certain monotonicity conditions, which means that we can often solve the associated Discrete Algebraic Riccati Inequality (DARI) instead.

Theorem 3. *Suppose the assumptions of Theorem 1 hold. Define the Riccati operator*

$$f(X) := A^\mathsf{T} X A - X + Q - (A^\mathsf{T} X B + S)(B^\mathsf{T} X B + R)^{-1}(B^\mathsf{T} X A + S^\mathsf{T}).$$

Let $X_0 \succeq 0$ be the stabilizing solution to the DARE $f(X_0) = 0$.

- *If $X \succeq 0$ and $f(X) \succ 0$, then $X_0 \succ X$. So X_0 is the* maximal solution *to this DARI.*
- *If $X \succeq 0$ and $f(X) \prec 0$, then $X_0 \prec X$. So X_0 is the* minimal solution *to this DARI.*

Proof. Define $K := -(B^\mathsf{T} X B + R)^{-1}(B^\mathsf{T} X A + S^\mathsf{T})$. We can rewrite the Riccati operator as

$$(A + BK)^\mathsf{T} X (A + BK) - X + \begin{bmatrix} I \\ K \end{bmatrix}^\mathsf{T} \begin{bmatrix} Q & S \\ S^\mathsf{T} & R \end{bmatrix} \begin{bmatrix} I \\ K \end{bmatrix} = f(X) \tag{19}$$

With some algebraic manipulations, we can prove the identity

$$(A + B\hat{K})^\mathsf{T}(\hat{X} - X)(A + B\hat{K}) - (\hat{X} - X) + (\hat{K} - K)^\mathsf{T}(B^\mathsf{T} X B + R)(\hat{K} - K) = f(\hat{X}) - f(X). \tag{20}$$

Suppose $f(X) \succ 0$ and $X \succeq 0$. Let $\hat{X} \mapsto X_0$ in (20) and obtain

$$\begin{aligned}(A + BK_0)^\mathsf{T}(X_0 - X)(A + BK_0) - (X_0 - X) &= -(K_0 - K)^\mathsf{T}(B^\mathsf{T} X B + R)(K_0 - K) - f(X) \\ &\prec 0.\end{aligned}$$

Since X_0 is stabilizing, $A + BK_0$ is Schur-stable, so we conclude that $X_0 \succ X$. Note that in order to make this work, we only used the fact that $B^\mathsf{T} X B + R \succ 0$. We made the stronger assumption $X \succeq 0$ for simplicity.

Now suppose $f(X) \prec 0$ and $X \succeq 0$. Examining (19), we deduce that

$$(A + BK)^\mathsf{T} X (A + BK) - X = f(X) - \begin{bmatrix} I \\ K \end{bmatrix}^\mathsf{T} \begin{bmatrix} Q & S \\ S^\mathsf{T} & R \end{bmatrix} \begin{bmatrix} I \\ K \end{bmatrix} \prec 0.$$

Therefore, $(A + BK)$ must be Schur-stable. Let $(\hat{X}, X) \mapsto (X, X_0)$ in (20) and obtain

$$\begin{aligned}(A + BK)^\mathsf{T}(X - X_0)(A + BK) - (X - X_0) &= -(K - K_0)^\mathsf{T}(B^\mathsf{T} X_0 B + R)(K - K_0) + f(X) \\ &\prec 0.\end{aligned}$$

Since $A + BK$ is Schur-stable, we conclude that $X_0 \prec X$. ■

We can leverage the results from Theorem 3 to derive equivalent linear matrix inequalities that can be used to find the solution to the DARE. This is an alternative approach to solving it directly using the stable generalized eigenvalue approach described in Section 1. The first result uses the maximal property of the DARI $f(X) \succ 0$.

Corollary 2. *Suppose the assumptions of Theorem 1 hold. The following are equivalent.*

(i) X_0 *is the stabilizing solution to the DARE*

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

(ii) For any $W \succ 0$, X_0 *is the solution to the optimization problem*

$$\begin{aligned} \underset{X}{\text{maximize}} \quad & \operatorname{trace}(WX) \\ \text{subject to} \quad & A^\mathsf{T} X A - X + Q - (A^\mathsf{T} X B + S)(B^\mathsf{T} X B + R)^{-1}(B^\mathsf{T} X A + S^\mathsf{T}) \succeq 0 \\ & X \succeq 0. \end{aligned}$$

(iii) For any $W \succ 0$, X_0 *is the solution to the optimization problem*

$$\begin{aligned} \underset{X}{\text{maximize}} \quad & \operatorname{trace}(WX) \\ \text{subject to} \quad & \begin{bmatrix} A^\mathsf{T} X A - X + Q & A^\mathsf{T} X B + S \\ B^\mathsf{T} X A + S^\mathsf{T} & B^\mathsf{T} X B + R \end{bmatrix} \succeq 0 \\ & X \succeq 0. \end{aligned}$$

Proof. From the maximality property of Theorem 3, if $f(X) \succeq 0$, we have $X_0 \succeq X$. Therefore, $W^{1/2} X_0 W^{1/2} \succeq W^{1/2} X W^{1/2}$, and so $\operatorname{trace}(WX_0) \geq \operatorname{trace}(WX)$. We can then apply Schur complement rules to transform the DARI into an LMI and prove the third item. ∎

The third item in Corollary 2 is the critical one, because it is a *linear matrix inequality* (LMI). It is a convex optimization problem that can be readily solved by modern solvers.

Although we could have written the second item of Corollary 2 using the minimality property instead, by changing the maximization to a minimization and replacing $f(X) \succeq 0$ by $f(X) \preceq 0$, we would not be able to derive the third item this way. This is because the Schur complement identity we used no longer works (the inequality goes the wrong way now).

There are other approaches that leverage the minimality property, however. One example is to jointly solve for (X_0, K_0). This turns out to be convexifiable once you make a change of variables and also leads to an LMI. In any case, this document is already too long so I'll stop here!