Methods to solve Riccati equations

Note Title 2023-05-24

Continuous-time algebraic Riccati equation (CARE, ARE)

$$A^*X + XA + Q - XGX = 0 \qquad \text{all } n\times n$$

$$G = BR^{-1}B^*$$

$$\underset{n\times m}{[\,]} \; \underset{m\times m}{\square} \; \underset{m\times n}{\square}$$

$Q \succeq 0$, $G \succeq 0$

(A,G), (A^*,Q) stabilizable

X s.t. $\Lambda(A-GX) \subset LHP$ (stabilizing solution)

1) Newton's method

$$F(X) = A^*X + XA + Q - XGX \qquad \mathbb{C}^{n\times n} \to \mathbb{C}^{n\times n}$$

$$\Updownarrow$$

$$\begin{bmatrix} A & -G \\ -Q & -A^* \end{bmatrix}\begin{bmatrix} I \\ X \end{bmatrix} = \begin{bmatrix} I \\ X \end{bmatrix}(A-GX)$$

$$F(X+H) = F(X) + L_{F,X}[H] + o(\|H\|)$$

$$L_{F,X}[H] = A^*H + HA - HGX - XGH$$

$$= (A-GX)^*H + H(A-GX)$$

(We can restrict to symmetric matrices, because the stabilizing solution is so)

$$\begin{cases} X_0 \text{ given (Hermitian)} \\ X_{k+1} = X_k - L_{F,X_k}^{-1} F(X_k) \end{cases}$$

i.e., we get $H = X_k - X_{k+1}$ by solving $L_{F,X_k}[H] = F(X_k)$

$$\boxed{(A-GX_k)^*H + H(A-GX_k) = A^*X_k + X_kA + Q - X_kGX_k}$$

Lyapunov equation. In particular, H is Hermitian

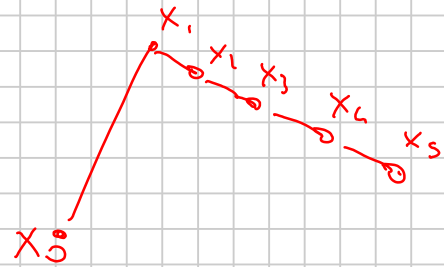
$\Rightarrow X_{k+1}$ is Hermitian if X_k is so.

Thm: if X_0 is stabilizing, then the sequence satisfies
$\Lambda(A-GX_0) \subset LHP$

$$X_1 \succcurlyeq X_2 \succcurlyeq X_3 \succcurlyeq \dots \succcurlyeq X \quad \text{(stabilizing solution)}$$

and $\lim_{k\to\infty} X_k = X$. Convergence is quadratic.

⚠ There is no $X_0 \succcurlyeq X_1$.



At each step, one needs to solve a Lyapunov equation
$\to$ compute the Schur factorization of $A - GX_k$.

Sign function method: $\begin{bmatrix} I \\ X \end{bmatrix}$ is the stable inv. subspace of H, so we can get it from $\text{sign}(H)$

1) Run the iteration
$$H_0 = H \qquad H_{k+1} = \frac{1}{2}\left(H_k + H_k^{-1}\right)$$
which converges quadratically to $\text{sign}(H)$

$$H = V \begin{bmatrix} J_{LHP} & 0 \\ 0 & J_{RHP} \end{bmatrix} V^{-1} \qquad \text{sign}(H) = V \begin{bmatrix} -I & 0 \\ 0 & I \end{bmatrix} V^{-1}$$

$$\text{sign}(H) + I = V \begin{bmatrix} 0 & 0 \\ 0 & 2I \end{bmatrix} V^{-1}$$

2) $\boxed{\text{Compute } \begin{array}{c} \\ n \\ n \end{array}\begin{array}{c} n \\ \begin{bmatrix} U_1 \\ U_2 \end{bmatrix} \end{array} = \text{Ker}\left[\text{sign}(H) + I\right]}$

3) Get $X = U_2 U_1^{-1}$

$$\begin{bmatrix} U_1 \\ U_2 \end{bmatrix} U^{-1} = \begin{bmatrix} I \\ U_2 U_1^{-1} \end{bmatrix} = \begin{bmatrix} I \\ X \end{bmatrix}$$

Cost per it.: $2n^3$ vs. $30n^3$ for Schur fact.

Structure preservation:

$\underbrace{H = \begin{bmatrix} A & -G \\ -Q & -A^* \end{bmatrix} \text{ with } G = G^*,\ Q = Q^*}_{\text{"Hamiltonian matrices"}}$ $\Rightarrow -H^* = JHJ^{-1} = J^{-1}HJ$

$\Rightarrow$ eigenvalue symmetry: $\lambda, -\bar{\lambda}$

$J = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ $\quad J^{-1} = -J$

Def: H is called a <u>Hamiltonian matrix</u> if $H_{22} = -H_{11}^*$

$$H_{12} = H_{12}^*$$
$$H_{21} = H_{21}^*$$

<u>Lemma:</u>

- If H is Hamiltonian, so is H^{-1} (if it exists)
- If H is Hamiltonian, so is αH for $\alpha \in \mathbb{R}$
- If H_1, H_2 are Hamiltonian, so is $H_1 + H_2$

Hence in the sign function method if H_0 is Hamiltonian then all iterates are Hamiltonian. (and hence $\operatorname{sign}(H)$)

<u>Proof:</u> $-H^* = JHJ^{-1} \rightarrow -H^{-*} = JH^{-1}J^{-1}$

$-H_1^* = JH_1J^{-1},\ -H_2^* = JH_2J^{-1} \rightarrow -(H_1+H_2)^* = J(H_1+H_2)J^{-1}$

<u>Geometrical angle</u>: Consider the (skew-symmetric, indefinite) bilinear form associated to J

$$\langle u, v \rangle_J = u^* J v = [u_1^* \ u_2^*] \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = u_1^* v_2 - u_2^* v_1$$

$u, v \in \mathbb{C}^{2n}$

H Hamiltonian $\Leftrightarrow$ H is skew-symmetric w.r.t. $\langle u, v \rangle_J$,

i.e. $\langle u, Hv \rangle_J = \langle -Hu, v \rangle_J \quad \forall u, v$

$$\underbrace{u^* JHv = -u^* H^* J v}_{} \quad \forall u, v \Leftrightarrow JH = -H^* J$$

Note that H Hamiltonian $\Leftrightarrow$ $JH = -H^*J = H^*J^* = (JH)^*$

i.e., JH is Hermitian

$$JH = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix} \begin{bmatrix} A & -G \\ -Q & -A^* \end{bmatrix} = \begin{bmatrix} -Q & -A^* \\ -A & G \end{bmatrix} \quad \text{is Hermitian}$$

In particular, we can rewrite the iteration

$$H_{k+1} = \frac{1}{2}\left(H_k + H_k^{-1}\right)$$

$H_k = J^{-1} Z_k$

in terms of the Hermitian matrices $Z_k = J H_k$

$H_k^{-1} = Z_k^{-1} J$

$$Z_{k+1} = JH_{k+1} = J\left(\frac{1}{2}(H_k + H_k^{-1})\right) = \frac{1}{2}\left(JH_k + J\underbrace{H_k^{-1}}_{= Z_k^{-1}J}\right)$$

$$= \frac{1}{2}\left(Z_k + J Z_k^{-1} J\right)$$

This is cheaper, because Z_k^{-1} can be computed with a symmetric method (LDL), for half the cost (n^3 vs $2n^3$) (and Z_k^{-1} is guaranteed to be <u>exactly</u> symmetric).

<u>Schur method</u>:

- → compute $H = QUQ^*$ Schur dec.
- → reorder $\hat{Q}$, $\hat{U} = \begin{bmatrix} \hat{U}_{11} & \hat{U}_{12} \\ 0 & \hat{U}_{22} \end{bmatrix}$ s.t. $\Lambda(\hat{U}_{11}) \subset LHP$
- → The inv. subspace is spanned by the first n columns of $\hat{Q}$.

We can prove that the schur decomposition is backward stable, i.e., the computed $\tilde{\hat{Q}}$, $\tilde{\hat{U}}$ satisfy exactly

$$\tilde{\hat{Q}} \tilde{\hat{U}} \tilde{\hat{Q}}^* = H + \delta_H \qquad \|\delta_H\| / \|H\| = O(u)$$

However, nothing ensures that the computed eigenvalues on $\mathrm{diag}(U)$, $\mathrm{diag}(\hat{U})$ will come in symmetric pairs.

In very ill-conditioned problems, they might not even be n in the LHP, n in the RHP.

(On the other hand, the sign function method computes an *exactly* Hamiltonian H_k at each step).

What transformations preserve the Hamiltonian structure?

Hamiltonian $\iff$ skew-symmetric wrt $\langle u, v\rangle_J$

We must look for transformations that are orthogonal w.r.t. $\langle u, v\rangle_J$

S <u>symplectic matrices</u> s.t.

$$\langle u, v\rangle_J = \langle Su, Sv\rangle_J \quad \forall\, u, v \in \mathbb{C}^{2n}$$
$$\Updownarrow$$
$$u^* J v = u^* S^* J S v \quad \forall\, u, v$$
$$\Updownarrow$$
$$J = S^* J S$$

<u>Def</u>: Matrices $S \in \mathbb{C}^{2n \times 2n}$ s.t. $S^* J S = J$ are called <u>symplectic matrices</u>

Ideally, we would like to compute a "structured Schur form" by $\underbrace{\text{orthogonal}}_{\text{stability}}$ <u>and</u> $\underbrace{\text{symplectic}}_{\text{preserve structure}}$ transformations

Unfortunately, symplectic alone is not enough: symplectic matrices can be ill-conditioned, e.g. for all M invertible

$S = \begin{bmatrix} M & 0 \\ 0 & -M^{-*} \end{bmatrix}$ is symplectic

Observation (Laub): let $U = \begin{bmatrix} U_1 \\ U_2 \end{bmatrix}$ our stable invariant subspace, computed e.g. via Schur form, and assume that $U = \begin{bmatrix} U_1 \\ U_2 \end{bmatrix}$ is an orthogonal basis. Then,

1. $V = \begin{bmatrix} U_1 & -U_2 \\ U_2 & U_1 \end{bmatrix}$ is orthogonal and symplectic

2. $V^* H V = \begin{bmatrix} T_{11} & T_{12} \\ 0 & T_{22} \end{bmatrix}$, with $T_{22} = -T_{11}^*$
$T_{12} = T_{12}^*$

Q: is there a better algorithm that uses "elementary symplectic Givens transformations" at each step to compute this Hamiltonian Schur-like form while preserving structure exactly?

A: Very likely no (also for reasons related to stability when H has eigenvalues on the imaginary axis)

Similar feats have been achieved, e.g. a Schur-like form for H^2