Solving algebraic Riccati equations

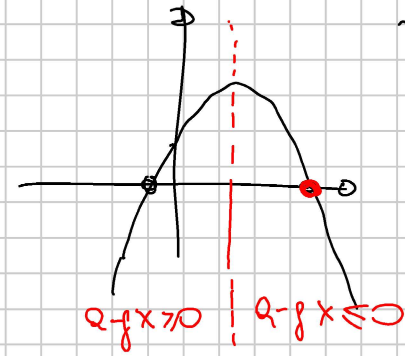
Note Title 2025-05-22

$Q \succeq 0$ $\qquad G = BR^{-1}B^* \succeq 0$

$$A^*X + XA + Q - XGX = 0$$

$\Lambda(A-GX) \subseteq LHP$ ← stabilizing solution

If $n=1$ $\qquad 2\mathrm{Re}(a)x + q - gx^2 = 0$ $\qquad q, g \geq 0$



Historically, first method: Newton's method.

$$F(X) = A^*X + XA + Q - XGX$$

$$F(X+H) = A^*(X+H) + (X+H)A - (X+H)G(X+H)$$

$$= F(X) + \underbrace{H(A-GX) + (A-GX)^*H}_{L_{F,X} \text{ Fréchet derivative}} + O(H^2)$$

Newton method:

1. Compute $F(X_k)$
2. Solve $H(A-GX_k) + (A-GX_k)^*H = -F(X_k)$ for H
3. $X_{k+1} = X_k + H$

If $\Lambda(A-GX_k) \subset LHP$, the equation is solvable.

(Ill-conditioned if $\Lambda(A-GX_k)$ contains eigenvalues close to the imaginary axis).

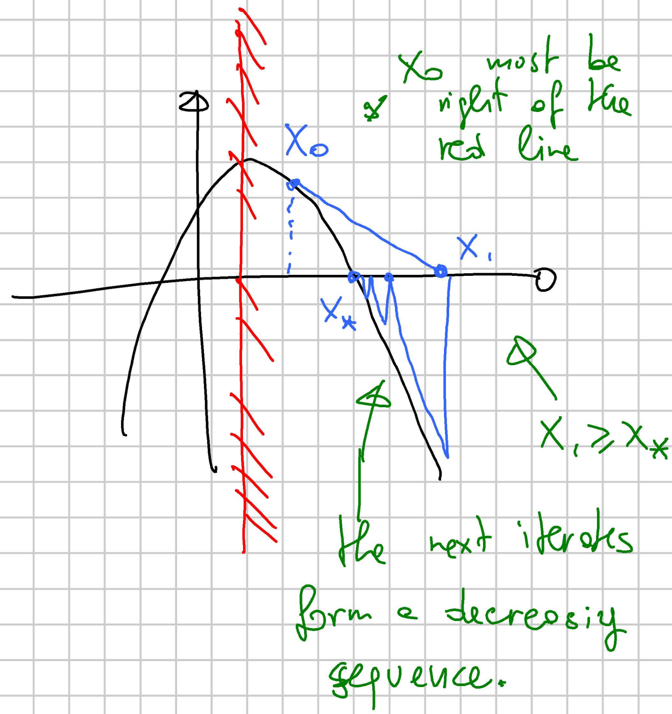
One can prove: if X_0 is s.t. $\Lambda(A-GX_0) \subset LHP$, then it holds that (if (A,G), (A^*,Q) controllable)

$$X_1 \geq X_2 \geq X_3 \geq \dots \geq X_k \geq \dots \geq X_* \geq 0$$

stabilizing solution

and the sequence converges to X_* quadratically.

⚠ $X_0 \geq X_1$ is not guaranteed!



Remark: under the usual controllability assumptions,

$$\Lambda(A - GX_*) = \{\lambda_1, \lambda_2, \dots \lambda_n\}$$

$$\Lambda(-(A - GX_*)^*) = \{-\bar{\lambda}_1, -\bar{\lambda}_2, \dots -\bar{\lambda}_n\}$$

The eigenvalues of the Fréchet derivative operator

$$H \mapsto (A - GX)^* H + H(A - GX)$$

or $$(A-GX)^T \otimes I + I \otimes (A-GX)^*$$

are $\{\lambda_i + \bar{\lambda}_j, \quad i,j = 1, \dots n\}$

These are different from 0 because they are in the LHP

$\Rightarrow$ the solution is simple and the Newton method converges quadratically.

Cost: one Lyapunov solution $\Rightarrow$ 1 Schur form at each step

$O(n^3)$ per step, but with high leading coefficient $\sim 30 n^3$ (+back-sub cost)

Often used only for solution refinement after an inaccurate method.

Another, better algorithm: Newton for the matrix sign.

X solves ARE $\iff$ $\underbrace{\begin{bmatrix} A & -G \\ -Q & -A^* \end{bmatrix}}_{H} \begin{bmatrix} I \\ X \end{bmatrix} = \begin{bmatrix} I \\ X \end{bmatrix} S$

$\begin{bmatrix} I \\ X \end{bmatrix}$ is an invariant subspace of H

$\Lambda(A-GX) \subset LHP \iff \begin{bmatrix} I \\ X \end{bmatrix}$ associated with $\Lambda(S) \subset LHP$

If we compute $\mathrm{sign}(H)$, then $\mathrm{Ker}(H+I) = \mathrm{span}\begin{bmatrix} I \\ X \end{bmatrix}$

where X is the stabilizing solution

Algorithm:

- $H_0 = H$
- $H_{k+1} = \frac{1}{2}\left(H_k + H_k^{-1}\right)$ $\quad k=0,1,2,\dots$ up to convergence
- $\begin{bmatrix} U_1 \\ U_2 \end{bmatrix}$ basis of $\mathrm{Ker}(H_\infty + I)$
- $X = U_2 U_1^{-1}$

Convergence is quadratic. Cost: one inversion $\left(2(2n)^3\right)$ per step.

Lemma: All matrices H_k are Hamiltonian, i.e. $J H_k = H_k^* J^{-1} = -H_k^* J$ $\qquad J = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix}$

or $H_k = \left[\begin{array}{c|c} H_{11} & H_{12} \\ \hline H_{21} & H_{22} \end{array}\right]$ $\qquad H_{12} = H_{12}^*, \quad H_{21} = H_{21}^*$

$H_{22} = -H_{11}^*$

<u>Proof</u>: H Hamiltonian $\Leftrightarrow JH = -H^* J$

H_1, H_2 Hamiltonian $\Rightarrow H_1 + H_2$ Hamiltonian

$J H_1 = -H_1^* J$
$J H_2 = -H_2^* J$
$J(H_1 + H_2) = -(H_1 + H_2)^* J$

H Hamiltonian $\Rightarrow H^{-1}$ Hamiltonian
invertible

$JH = -H^* J$
multiply by H^{-*} and H^{-1}
$H^{-*} J = -J H^{-1}$

<u>Remark</u>: if one considers the antisymmetric scalar product

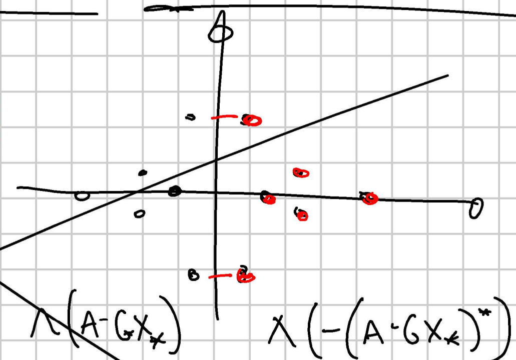
$\langle u, v \rangle = u^* J v$

H Hamiltonian $\Leftrightarrow \langle u, Hv \rangle = -\langle Hu, v \rangle$.

$\| J_{oc}^{-1} \| = \mathrm{sep}(A - GX, -(A-GX)^*)$

$K(J_{oc}) = \| J_{oc} \| \cdot \| J_{oc}^{-1} \| = \dfrac{\| J_{oc} \|}{\mathrm{sep}}$

$2\|A - GX\|$



$H_k \qquad Z_k = J H_k$ $\qquad J = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$

H Hamiltonian $\Leftrightarrow Z = JH$ symmetric (Hermitian)

$JH = -H^* J \quad \Leftrightarrow \quad (JH) = (JH)^* = H^* J^* = -H^* J$

We rewrite the iteration in terms of $Z_k = J H_k$

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

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

<u>Structure-preserving sign iteration:</u>

$$\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} A & -G \\ -Q & -A^* \end{bmatrix}$$

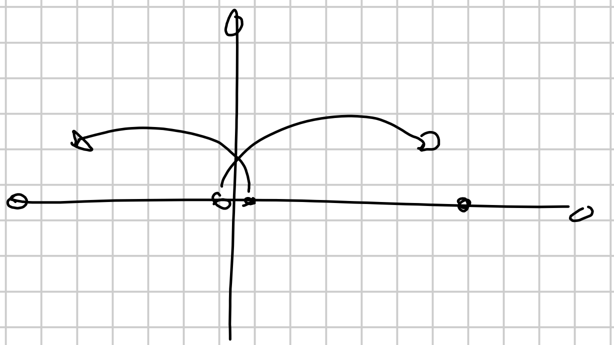
$$Z_0 = \begin{bmatrix} -Q & -A^* \\ -A & G \end{bmatrix}$$

$$Z_{k+1} = \frac{1}{2}\left(Z_k + J Z_k^{-1} J\right) \qquad k = 0, 1, 2, \dots$$

$$\begin{bmatrix} U_1 \\ U_2 \end{bmatrix} = \ker(H_\infty + I) = \ker(J^{-1} Z_\infty + I) = \ker(Z_\infty + J)$$

$$X = U_2 U_1^{-1}$$

Z_k symmetric $\Rightarrow$ inv(Z_k) has half the cost: $(2n)^3$ instead of $2(2n)^3$



Third method: reordered Schur form

1. Compute $[U, T] = \text{schur}(H)$
2. Reorder s.t. T has stable eigenvalues first
3. $X = U_2 U_1^{-1}$, where $U = \begin{bmatrix} U_1 & * \\ U_2 & * \end{bmatrix}$

The Schur factorization is backward stable:

the computed $\tilde{U}, \tilde{T}$ satisfy $\dfrac{\|\tilde{U}\tilde{T}\tilde{U}^H - H\|}{\|H\|} = O(u)$

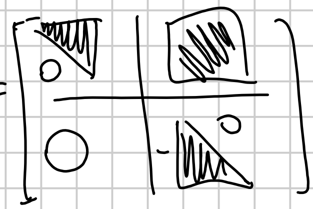
(backward stability).

However, $\tilde{U}\tilde{T}\tilde{U}^H$ is <u>not</u> guaranteed to be Hamiltonian

<u>Q</u>: can one make a structure-preserving version of the Schur factorization?

In principle, yes: if $\begin{bmatrix} U_1 \\ U_2 \end{bmatrix}$ is the exact invariant subspace, then $U = \begin{bmatrix} U_1 & -U_2 \\ U_2 & U_1 \end{bmatrix}$ is orthogonal ($U^*U = I$)

and preserves J: $U^*JU = J$

$U^*HU =$ 

... but there is no algorithm to compute it based on Givens-like transformations ("curse of Van Loan").