

Sign-like methods for CAREs

Matrix sign iteration

$$X_{k+1} = \frac{1}{2}(X_k + X_k^{-1}), \quad X_0 = \mathcal{H}.$$

It is not difficult to see that X_k is Hamiltonian at each step (i.e., $JX_k = -X_k^*J$). Just show that

- ▶ If M is Hamiltonian, then M^{-1} is Hamiltonian, too.
- ▶ If M_1, M_2 are Hamiltonian, then $M_1 + M_2$ is Hamiltonian, too.

(Guiding idea: Hamiltonian matrices are ‘like antisymmetric ones’: properties that you expect for antisymmetric matrices will often hold for Hamiltonian, too.)

Lemma: Se H_1, H_2 Hermitiane (i.e. $H_i^* J = -J H_i$)
allora

• $H_1 + H_2$ è Hermitiana

(basta sommare (*) per $i=1,2$)

• H_1^{-1} è Hermitiana:

$$H_1^* J = -J H_1 \rightarrow J H_1^{-1} = -H_1^* J$$

$$H^* J = -JH$$

$$J^* = -J$$

$$J = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix}$$

$$(JH)^* = -(H^* J) = H^* J^*$$

$$\begin{bmatrix} \textcircled{A} & \textcircled{-Q} \\ \textcircled{-G} & \textcircled{-A^T} \end{bmatrix}$$

$$JH = \begin{bmatrix} -G & -A^T \\ -A & Q \end{bmatrix}$$

$$X_{k+1} = \frac{1}{2}(X_k + X_k^{-1})$$

$$Z_k := J X_k$$

$$\boxed{Z_{k+1}} = J X_{k+1} = J \cdot \frac{1}{2}(X_k + X_k^{-1}) = \frac{1}{2}(J X_k + J X_k^{-1}) =$$

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

(per calcolare Z_k^{-1} posso usare un metodo per matrici simmetriche, ad es. LDL^T , che costa la metà di un metodo per matrici generiche)

Structure-preserving sign iteration

In machine arithmetic, the X_k won't be exactly Hamiltonian — unless we modify our algorithm to ensure that they are.

Recall: X is Hamiltonian iff $Z = JX$ is symmetric.

Rewrite the iteration in terms of $Z_k := JX_k$:

$$Z_{k+1} = \frac{1}{2}(Z_k + \underbrace{JZ_k^{-1}J}), \quad Z_0 = J\mathcal{H}.$$

Will preserve symmetry exactly (assuming the method we use for inversion does).

We can incorporate scaling.

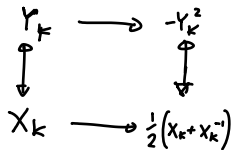
It is in some sense 'working on even pencils': given an even pencil $\lambda J - Z_k$, construct $\lambda J - Z_{k+1}$ (will see more of this idea in the following).

$J\lambda - z_k$ ha gli stessi autovalori

$$\text{di } J^{-1}z_k = X_k$$

Problema del metodo: serve invertire X_k o z_k a ogni passo. Se z_0 , o X_0 , è quasi singolare, allora posso fare errori al primo passo e non recupero più.

Towards doubling



Recall: in the ^{sign} doubling iteration, if we set $Y_k = (I - X_k)^{-1}(I + X_k)$, then $Y_{k+1} = -Y_k^2$.

In an ideal world without rounding errors, we could compute Y_0, Y_1, Y_2, \dots , and then get the stable invariant subspace as $\ker Y_\infty$.

We can do something similar, if we work in a suitable format.

X_0 has n eigenvalues in LHP, n in RHP

Y_0 has n eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$ with $|\lambda_i| < 1$, and n with $|\lambda_i| > 1$

$Y_k = (Y_0)^{2^k}$ has eigenvalues $\lambda_i^{2^k}$ with n tending to 0 and n tending to ∞

Vogliamo avere ad ogni passo

$$Y_k = \begin{bmatrix} I & G_k \\ 0 & F_k \end{bmatrix}^{-1} \begin{bmatrix} E_k & 0 \\ H_k & I \end{bmatrix}$$

per opportune matrici E_k, F_k, G_k, H_k

$$Y_0 = (I - \mathcal{H})^{-1} (I + \mathcal{H})$$

Trucco: cerco una matrice $M \in \mathbb{C}^{2n \times 2n}$ invertibile tale che

$$M \begin{bmatrix} I - \mathcal{H} & \vdots & I + \mathcal{H} \\ 2n & & 2n \end{bmatrix} = \begin{bmatrix} I & G_0 & \vdots & E_0 & 0 \\ 0 & F_0 & \vdots & H_0 & I \\ 2n & & & 2n & \end{bmatrix}$$

Standard Symplectic Form

Goal: write $Y_0 = (I - \mathcal{H})^{-1}(I + \mathcal{H})$ as

$$Y_0 = \begin{bmatrix} I & -G_0 \\ 0 & F_0 \end{bmatrix}^{-1} \begin{bmatrix} E_0 & 0 \\ -H_0 & I \end{bmatrix}.$$

Can be rewritten as: find M such that

$$M \begin{bmatrix} (I - \mathcal{H}) & (I + \mathcal{H}) \end{bmatrix} = \begin{bmatrix} I & G_0 & E_0 & 0 \\ 0 & F_0 & H_0 & I \end{bmatrix}.$$

Solution: M is the inverse of block columns (1, 4).

Structural properties: if \mathcal{H} is Hamiltonian, Y_0 is symplectic. If Y_0 is symplectic, $E_0 = F_0^*$, $G_0 = G_0^*$, $H_0 = H_0^*$.

Moreover, if $G \succeq 0$, $H \succeq 0$, then $G_0 \succeq 0$, $H_0 \preceq 0$.

(All these manipulations can be reformulated as left-multiplication of the pencil $(I - \mathcal{H}, I + \mathcal{H})$.)

Doubling algorithm

Plan Given $Y_k = \begin{bmatrix} I & G_k \\ 0 & E_k^* \end{bmatrix}^{-1} \begin{bmatrix} E_k & 0 \\ H_k & I \end{bmatrix}$, compute

$$Y_{k+1} = -Y_k^2 = \begin{bmatrix} I & G_{k+1} \\ 0 & E_{k+1}^* \end{bmatrix}^{-1} \begin{bmatrix} E_{k+1} & 0 \\ H_{k+1} & I \end{bmatrix}.$$

Similar to the 'inverse-free sign method' described earlier.

If $Y_k = \mathcal{M}_k^{-1} \mathcal{N}_k$, then $-Y_k^2 = -\mathcal{M}_k^{-1} \mathcal{N}_k \mathcal{M}_k^{-1} \mathcal{N}_k = \mathcal{M}_k^{-1} \widehat{\mathcal{M}}_k^{-1} \widehat{\mathcal{N}}_k \mathcal{N}_k = (\widehat{\mathcal{M}}_k \mathcal{M}_k)^{-1} (\widehat{\mathcal{N}}_k \mathcal{N}_k)$, where $\widehat{\mathcal{M}}_k, \widehat{\mathcal{N}}_k$ satisfy $\widehat{\mathcal{M}}_k^{-1} \widehat{\mathcal{N}}_k = -\mathcal{N}_k \mathcal{M}_k^{-1}$, i.e.,

$$\begin{bmatrix} \widehat{\mathcal{M}}_k & \widehat{\mathcal{N}}_k \end{bmatrix} \begin{bmatrix} \mathcal{N}_k \\ \mathcal{M}_k \end{bmatrix} = 0.$$

Doubling: inversion trick

$$\begin{bmatrix} I & \widehat{G}_k & \widehat{E}_k & 0 \\ 0 & \widehat{F}_k & \widehat{H}_k & I \end{bmatrix} \begin{bmatrix} E_k & 0 \\ H_k & I \\ I & G_k \\ 0 & E_k^* \end{bmatrix} = 0$$

holds if

$$\begin{aligned} \begin{bmatrix} \widehat{G}_k & \widehat{E}_k \\ \widehat{F}_k & \widehat{H}_k \end{bmatrix} &= - \begin{bmatrix} E_k & 0 \\ 0 & E_k^* \end{bmatrix} \begin{bmatrix} H_k & I \\ I & G_k \end{bmatrix}^{-1} \\ &= \begin{bmatrix} E_k & 0 \\ 0 & E_k^* \end{bmatrix} \begin{bmatrix} G_k(I - H_k G_k)^{-1} & -(I - G_k H_k)^{-1} \\ -(I - H_k G_k)^{-1} & H_k(I - G_k H_k)^{-1} \end{bmatrix}. \end{aligned}$$

Doubling: the formulas

Putting everything together,

$$\begin{aligned}\begin{bmatrix} E_{k+1} & 0 \\ H_{k+1} & I \end{bmatrix} &= \begin{bmatrix} -E_k(I - G_k H_k)^{-1} & 0 \\ E_k^* H_k (I - G_k H_k)^{-1} & I \end{bmatrix} \begin{bmatrix} E_k & 0 \\ H_k & I \end{bmatrix} \\ &= \begin{bmatrix} -E_k(I - G_k H_k)^{-1} E_k & 0 \\ H_k + E_k^* H_k (I - G_k H_k)^{-1} E_k & I \end{bmatrix}\end{aligned}$$

and an analogous computation gives E_{k+1}^* , G_{k+1} :

Structured doubling algorithm

$$\begin{aligned}E_{k+1} &= -E_k(I - G_k H_k)^{-1} E_k, \\ G_{k+1} &= G_k + E_k G_k (I - H_k G_k)^{-1} E_k^*, \\ H_{k+1} &= H_k + E_k^* H_k (I - G_k H_k)^{-1} E_k.\end{aligned}$$

SDA: details

Note that (even when the middle term does not converge)

$$G_k(I - H_k G_k)^{-1} = G_k + G_k H_k G_k + G_k H_k G_k H_k G_k + \dots = (I - G_k H_k)^{-1} G_k,$$

and this matrix is symmetric. If $G_k = B_k B_k^*$, then it can also be rewritten as $B_k(I - B_k^* H_k B_k)^{-1} B_k^*$ (inverting a symmetric matrix).

Monotonicity If $H_k \preceq 0$ then $G_k(I - H_k G_k)^{-1} \succeq 0$. Hence, $0 \preceq G_0 \preceq G_1 \preceq \dots$, and $0 \succeq H_0 \succeq H_1 \succeq H_2 \succeq \dots$

Cost As much as a $2n \times 2n$ inversion $M^{-1}N$, if you put everything together. Unlike the sign algorithm, we have a bound $\sigma_{\min}(I - H_k G_k) \geq 1$.

SDA: the dual equation

To analyze convergence, we need to introduce another matrix. Let Y be the matrix such that

$$\mathcal{H} \begin{bmatrix} -Y \\ I \end{bmatrix} = \begin{bmatrix} A & -G \\ -Q & -A^* \end{bmatrix} \begin{bmatrix} -Y \\ I \end{bmatrix} = \begin{bmatrix} -Y \\ I \end{bmatrix} \hat{\mathcal{R}}$$

is the **anti-stable** invariant subspace of \mathcal{H} , i.e., $\Lambda(\hat{\mathcal{R}}) \subset RHP$.

Why does it exist? Because $\begin{bmatrix} I \\ Y \end{bmatrix}$ spans the stable subspace of $\mathcal{H}^* = -J\mathcal{H}J$, and we can repeat our arguments on it (in the version with (A, B) and (A^T, C^T) controllable)

SDA: convergence (intuitively)

Intuitive view E_k approximately squared at each time. Hence

$$\mathcal{H}_k = \begin{bmatrix} I & G_k \\ 0 & E_k^* \end{bmatrix}^{-1} \begin{bmatrix} E_k & 0 \\ H_k & I \end{bmatrix}$$

has n eigenvalues $\rightarrow 0$ and n that $\rightarrow \infty$. $\ker \mathcal{H}_k \approx \begin{bmatrix} I \\ -H_k \end{bmatrix}$, so
 $-H_k \rightarrow X$.

Dually, “ $\ker \mathcal{H}_k^{-1}$ ” (a thing that shouldn't exist...) $\approx \begin{bmatrix} -G_k \\ I \end{bmatrix}$, so
 $G_k \rightarrow Y$.

SDA convergence (formally)

More formally

$$\mathcal{H}_0 \begin{bmatrix} I \\ X \end{bmatrix} = (I - \mathcal{H})^{-1}(I + \mathcal{H}) \begin{bmatrix} I \\ X \end{bmatrix} = \begin{bmatrix} I \\ X \end{bmatrix} (I - \mathcal{R})^{-1}(I + \mathcal{R}).$$

where $\mathcal{S} = (I - \mathcal{R})^{-1}(I + \mathcal{R})$ has all eigenvalues in the unit circle.

$$\begin{bmatrix} I & G_k \\ 0 & E_k^* \end{bmatrix} \begin{bmatrix} I \\ X \end{bmatrix} = \begin{bmatrix} I \\ X \end{bmatrix} \begin{bmatrix} E_k & 0 \\ H_k & I \end{bmatrix} \begin{bmatrix} I \\ X \end{bmatrix} \mathcal{S}^{2^k}.$$

which implies

$$E_k = (I + G_k X) \mathcal{S}^{2^k},$$
$$H_k + X = E_k^* X \mathcal{S}^{2^k} = (\mathcal{S}^{2^k})^* (I + X G_k) \mathcal{S}^{2^k} \succeq 0.$$

The same computation on the dual equation gives $G_k \preceq Y$, so G_k is bounded and $E_k \rightarrow 0, H_k + X \rightarrow 0$ (quadratically as \mathcal{S}^{2^k}).