Stein (discrete Lyapunov) equation

$\to \quad X - A^* X A = Q$

$X - A X A^* = G, \quad G = bb^*$ $\qquad x_{k+1} = Ax_k + bu_k$

$A \in \mathbb{C}^{n\times n} \quad x_k, b \in \mathbb{C}^n$

$$X = \sum_{k=0}^{\infty} (A^*)^k Q A^k \qquad \rho(A) < 1$$

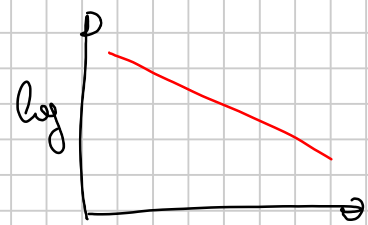
Smith method: $X_{k+1} = Q + A^* X_k A \qquad X_0 = 0$

$$\underline{X \;= Q - A^* X A}$$

$$X_{k+1} - X = A^*(X_k - X)A$$

$$X_k - X = (A^*)^k (X_0 - X) A^k$$

$$\|X_k - X\| \sim \rho(A)^{2k}$$



<u>A dense</u>: "doubling variant"

<u>Idea</u>: construct a recurrence for $\boxed{Q_k = X_{2^k}}$

$$X_k = Q + A^*QA + (A^*)^2 Q A^2 + \dots + (A^*)^{k-1} Q A^{k-1}$$

$$\mathrm{vec}(X_k) = \left(I + M + M^2 + \dots + M^{k-1}\right) \mathrm{vec}\, Q \qquad M = A^T \otimes A^*$$

$$I + M + M^2 + \dots + M^{2^{k+1}-1} = \left(I + M + M^2 + \dots + M^{2^k-1}\right) + M^{2^k}\left(I + M + M^2 + \dots + M^{2^k-1}\right)$$

$$\mathrm{vec}\, Q_{k+1} = \mathrm{vec}\, Q_k + M^{2^k} \mathrm{vec}\, Q_k$$

$$Q_{k+1} = Q_k + (A^*)^{2^k} Q_k A^{2^k}$$

$$\begin{cases} A_0 = A, \quad Q_0 = Q \\ A_{k+1} = A_k^2 \\ Q_{k+1} = Q_k + A_k^* Q_k A_k \end{cases}$$

"Squared Smith"

$\boxed{Q_k = X_{2^k} \text{ at each step}}$

Cost: 3 matmul

Convergence: $\|X - Q_k\| \sim \rho(A)^{2 \cdot 2^k}$

$$X_0 \preccurlyeq X_1 \preccurlyeq X_2 \preccurlyeq X_3 \preccurlyeq \dots \preccurlyeq X$$

Sparse A

$$X - AXA^* = G = \boxed{bb^*}$$

problem: X dense, and (in general) not low-rank

"solution": in many problems, X is approximately low rank!

We want to store the solution as $X \approx ZZ^*$, Z tall thin

Smith:

$$X_k = bb^* + Abb^*A^* + A^2bb^*(A^*)^2 + \dots + A^{k-1}b\,b^*(A^*)^{k-1}$$

$$X_k = Z_k Z_k^* \qquad Z_k = \left[b \ Ab \ A^2b \ \cdots \ A^{k-1}b \right] \in \mathbb{C}^{n \times k}$$

$$v_0 = b \qquad v_{k+1} = Av_k$$

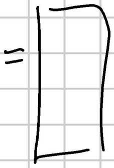
$$Z_k = [v_0 \ \dots \ v_{k-1}] \qquad Z_1 = [b] \qquad Z_{k+1} = [Z_k \ v_k]$$

$$\|Z_k Z_k^* - X\| \sim \rho(A)^{2k}$$

$$\sigma_k(X) \sim \rho(A)^{2k}$$

$$\sigma_{k+1} = \min_{\operatorname{rk} Y = k} \|X - Y\|$$

(More on this algorithm with Lyapunov equations)

Compression of Z_k: $\quad Z_k =$  $\quad Z_k = U_k S_k V_k^*$

If any of $\mathrm{diag}(S_k)$ are small, truncate U, S, V.

Remark: for the dense case you do not want to solve

$$(I - M)\,\mathrm{vec}\, X = \mathrm{vec}\, Q$$

with a direct $O(n^6)$ method.

Lyapunov equations: (continuous-time equivalent of Smith)

$$Q \succeq 0$$

$$A^* X + XA + Q = 0 \qquad \dot{x}(t) = A x(t) \qquad x(t) = \exp(tA)\, x_0$$

$$V(x) = x^* X x$$

$$\frac{d}{dt} V(x(t)) = \frac{d}{dt} x(t)^* X x(t) = \dot{x}(t)^* X x(t) + x(t)^* X \dot{x}(t)$$

$$= x(t)^* (A^* X + XA)\, x(t) = -x(t)^* Q\, x(t) \leq 0$$

table:

cont. time	discr. time
$\dot{x}(t) = A x(t)$	$x_{k+1} = A x_k$
stable if $\Lambda(A) \subset LHP$ (left half-plane)	stable if $\Lambda(A) \subset$ disc
$x(t) = \exp(At)\, x_0$	$x_k = A^k x_0$
Lyap $\quad A^* X + XA + Q = 0$	Stein $\quad X - A^* X A = Q$
$X = \int_0^\infty \exp(A^* t)\, Q \exp(At)\, dt$	$X = \sum_{k=0}^{\infty} (A^*)^k Q A^k$

(if A stable) | (if A stable)

Control: $\dot{x}(t) = Ax(t) + bu(t)$

$$x(t) = \exp(At)x_0 + \int_0^t \exp(A(t-\tau))\, bu(\tau)\,d\tau$$

since $\exp(A(t-\tau))b \in \operatorname{Im}[b\ Ab\ A^2b \cdots]$

controllable if
$\operatorname{rank}[b\ Ab\ A^2b\ \cdots] = n$

Control: $x_{k+1} = Ax_k + bu_k$

$$x_k = A^k x_0 + A^{k-1}bu_0 + A^{k-2}bu_1 + \dots + bu_{k-1}$$

controllable if
$\operatorname{rank}[b\ Ab\ \cdots\ A^{k-1}b] = n$

def: A system is controllable if for given $x_F \in \mathbb{C}^n$ and t_F, I can choose $u(t)$ such that

$x(t_F) = x_F$

min $\|u\|$ s.t. $x(t_F) = x_F$ for some t_F (including $t_F = \infty$) $= x_F^* X^{-1} x_F$

where X solves $AX + XA^* + bb^* = 0$

"Controllability Gramian"

min $\|u\|$ s.t. $x_k = x_F$ for some k $= x_F^* X^{-1} x_F$

where X solves $X - AXA^* = bb^*$

$t_F = 10$, $x_F^* X_{t_F}^{-1} x_F$, where

$$X_{t_F} = \int_0^{t_F} \exp(At)\, bb^* \exp(A^*t)\,dt$$

$$[b\ Ab \cdots A^{k-1}b]\begin{bmatrix} b^* \\ b^*A^* \\ \vdots \\ b^*A^{*k-1} \end{bmatrix}$$

$A^*XI + IXA$

$$\operatorname{vec} X = \underbrace{(I \otimes A^* + A^T \otimes I)}^{-1} \operatorname{vec} Q$$

$$\operatorname{vec} X = (I - A^T \otimes A^*)^{-1} \operatorname{vec} Q$$

$A = UTU^*$:

$$(U \otimes \bar{U})\,(I \otimes T^* + T^T \otimes I)\,(U^T \otimes U^*)$$

orth | | orth*

Singular if there are two eigenvalues $\lambda_i, \lambda_j \in \Lambda(A)$ s.t.

$$\lambda_i = -\overline{\lambda_j}$$

Singular if there are two eigenvalues $\lambda_i, \lambda_j \in \Lambda(A)$ s.t.

$$\lambda_i = \frac{1}{\overline{\lambda_j}}$$

$$I \otimes T^* + T^T \otimes I$$

$$= \begin{bmatrix} \blacktriangle & & & \\ & \blacktriangle & & \\ & & \ddots & \\ & & & \blacktriangle \end{bmatrix} + \begin{bmatrix} \ddots & & & 0 \\ \ddots & \ddots & & \\ \vdots & \ddots & \ddots & \\ \ddots & \cdots & & \ddots \end{bmatrix}$$

$$\begin{bmatrix} \lambda_1 & & & & & & & & \\ & \ddots & & & & & & & \\ & & \overline{\lambda_n} & & & & & & \\ & & & \overline{\lambda_1} & & & & & \\ & & & & \ddots & & & & \\ & & & & & \overline{\lambda_n} & & & \\ & & & & & & \ddots & & \\ & & & & & & & \overline{\lambda_1} & & \\ & & & & & & & & \ddots & \\ & & & & & & & & & \overline{\lambda_n} \end{bmatrix} + \begin{bmatrix} \lambda_1 & & & & & & & \\ & \ddots & & & & & & \\ & & \lambda_1 & & & & & \\ & & & \lambda_2 & & & & \\ & & & & \ddots & & & \\ & & & & & \lambda_2 & & \\ & & & & & & \ddots & \\ & & & & & & & \lambda_n \\ & & & & & & & & \ddots \\ & & & & & & & & & \lambda_n \end{bmatrix}$$

Bartels-Stewart algorithm (1970s)

$O(n^3)$ direct algorithm to solve $A^*X + XA + Q = 0$:

1. Compute $A = UTU^*$, Schur factorization
2. Change of variables in the equation:

$$\cancel{U^* U} T^* U^* X U + U^* X U T \cancel{U^* U} + U^* Q U = 0$$

$$T^* \hat{X} + \hat{X} T + \hat{Q} = 0 \qquad \hat{X} = U^* X U$$

$$\hat{Q} = U^* Q U$$

3. Solve the triangular version entry by entry:

$$\underbrace{\begin{bmatrix} \times & & & \\ \times & \times & & \\ \times & \times & \times & \\ \times & \times & \times & \times \end{bmatrix}}_{T^*}\underbrace{\begin{bmatrix} \times & \times & \times & \times \\ \times & \times & \times & \times \\ \times & \times & \times & \times \\ \times & \times & \times & \times \end{bmatrix}}_{\hat X}+\underbrace{\begin{bmatrix} \times & \times & \times & \times \\ \times & \times & \times & \times \\ \times & \times & \times & \times \\ \times & \times & \times & \times \end{bmatrix}}_{\hat X}\underbrace{\begin{bmatrix} \times & \times & \times & \times \\ & \times & \times & \times \\ & & \times & \times \\ & & & \times \end{bmatrix}}_{T}+\underbrace{\begin{bmatrix} \times & \times & \times & \times \\ \times & \times & \times & \times \\ \times & \times & \times & \times \\ \times & \times & \times & \times \end{bmatrix}}_{\hat Q}=0$$

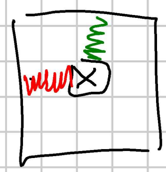
(1,1) entry:

$$\bar t_{11}x_{11}+x_{11}t_{11}+q_{11}=0 \to x_{11}=\frac{-q_{11}}{\bar t_{11}+t_{11}}$$

(2,1) $$\bar t_{21}x_{11}+\bar t_{22}x_{21}+x_{21}t_{11}+q_{21}=0 \to x_{21}=\frac{-q_{21}-\bar t_{21}x_{11}}{\bar t_{22}+t_{11}}$$

(i,j) $$\sum_{k=1}^{i}\bar t_{ik}x_{kj}+\sum_{k=1}^{j}x_{ik}t_{kj}+q_{ij}=0$$

$$x_{ij}=\frac{-q_{ij}-\boxed{\sum_{k=1}^{i-1}\bar t_{ik}x_{kj}}-\boxed{\sum_{k=1}^{j-1}x_{ik}t_{kj}}}{\bar t_{ii}+t_{jj}}.$$



We can compute $\hat X$ entry by entry

4. $\boxed{X=U\hat X U^*}$

Idea: this is inverting $I\otimes A^*+A^T\otimes I$ using the Schur factorization:

$$\underbrace{(I\otimes A^*+A^T\otimes I)^{-1}\mathrm{vec}Q=(\bar U\otimes U)\underbrace{(I\otimes T^*+T^T\otimes I)^{-1}\underbrace{(U^T\otimes U^*)\mathrm{vec}Q}_{\mathrm{vec}(\hat Q)}}_{\text{forward substitution to compute }\mathrm{vec}(\hat X)}}_{\mathrm{vec}(X)=\mathrm{vec}(U\hat X U^*)}$$

$O(n^3)$ beats $O(n^6)$

Same idea for Stein equations:

$$\hat{X} - T^* \hat{X} T = \hat{Q}$$

$$\square - \triangle \, \square \, \nabla = \square$$

entries can be computed one by one, because $I - T^T \otimes T^*$ is lower triangular.

The same idea works for $AX + XB = C$ (Sylvester equations)

$$\square\square + \square\square = \square$$

and for $AXB + CXD = E$ (generalized Sylvester equations)

with a generalization of the Schur factorization ($qz(A,C)$, $qz(B,D)$)

The Bartels-Stewart idea does ~~not~~ work for 3-term equations:

$$AX + XB + DXE = C$$

(no cubic algorithms known for those

$$\begin{cases} A_1^* X + X A_1 \preceq 0 \\ A_2^* X + X A_2 \preceq 0 \end{cases} \qquad V(x) = x^* X x$$