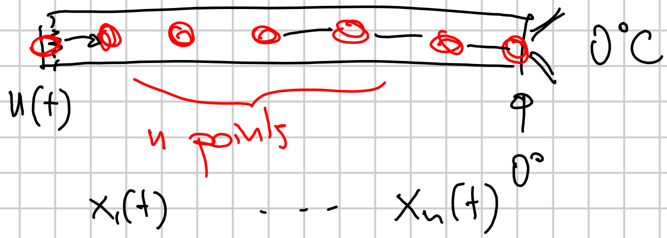
Matrix equations in control theory

Heating a corridor:



$$\dot{x}_k(t) = \alpha\big(x_{k+1}(t) - x_k(t)\big) + \alpha\big(x_{k-1}(t) - x_k(t)\big)$$

$$\frac{d}{dt}\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}(t) = \alpha \begin{bmatrix} -2 & 1 & & 0 \\ 1 & \ddots & \ddots & \\ & \ddots & \ddots & 1 \\ 0 & & 1 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}(t) + \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} u(t)$$

$$\frac{d}{dt} x(t) = Ax(t) + bu(t)$$

linear control system

Goal: reach a comfortable $x(t)$ by setting $u(t)$ appropriately

$A \in \mathbb{C}^{n\times n}$, $x : [0,\infty) \to \mathbb{C}^n$, $u(t) : [0,\infty) \to \mathbb{C}$

$b \in \mathbb{C}^n$

Stein equation (discrete-time Lyapunov equation)

(S) $\quad X - A^* X A = Q \qquad A \in \mathbb{C}^{n\times n}, \quad Q \in \mathbb{C}^{n\times n}$

$X \in \mathbb{C}^{n\times n} \qquad Q \succcurlyeq 0$ (pos. semidefinite)

Discrete-time dynamical system

$$\begin{cases} x_{k+1} = Ax_k \\ x_0 \text{ given} \end{cases}$$

$x_k = A^k x_0$ if $\rho(A) < 1$,

then $\lim_{k\to\infty} x_k = 0$ for all x_0

Lemma: Let $Q \succ 0$ be given, and suppose the equation (S) has solution $X \succ 0$, then $\rho(A) < 1$

Proof: $Av = v\lambda$

$$v^* Q v = v^* X v - v^* A^* X A v$$
$$= v^* X v - \bar{\lambda} v^* X v \lambda$$
$$= (1 - \bar{\lambda}\lambda)\, v^* X v$$

$$1 - \bar{\lambda}\lambda = \frac{v^* Q v}{v^* X v} > 0 \Rightarrow |\lambda| < 1 \text{ for all } \lambda. \quad \square$$

1. Choose $Q \succ 0$
2. Solve equation
3. If $X \succ 0$, then $\rho(A) < 1$

Lemma: suppose $\boxed{\rho(A) < 1,}$ then the solution of (S) is

$$X = \sum_{k=0}^{\infty} (A^*)^k Q A^k = Q + A^* Q A + (A^*)^2 Q A^2 + \dots$$

Proof:

$$X = Q + A^* \left(\sum_{k=0}^{\infty} (A^*)^k Q A^k \right) A = Q + A^* X A$$

$$\| A^k \| \sim \rho(A)^k$$

From the formula for X, we see $X \succ 0$ if $Q \succ 0$.

Lyapunov's argument: consider the "energy function"

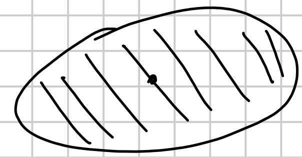
$V(x) = x^* X x.$

$$\begin{cases} x_{k+1} = A x_k \\ x_0 \text{ given} \end{cases}$$

$$V(x_k) - V(x_{k+1}) = x_k^* X x_k - x_{k+1}^* X x_{k+1}$$
$$= x_k^* X x_k - x_k^* A^* X A x_k$$

$$= x_k^* Q x_k \geq 0$$

$x_k^* X x_k \leq \ell$



$V(x_k)$ decreasing $\Rightarrow$ x_k bounded

A slightly stronger argument:

$$x_k^* Q x_k \geq \lambda_{min}(Q) \|x_k\|^2$$

If $\|x_k\| \geq \varepsilon$ for all k, contradiction

Numerical methods for $\quad X - A^* X A = Q$

#1: vectorization.

$$\begin{bmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{bmatrix} \rightarrow \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8 \\ 9 \end{bmatrix}$$

Given X, $\operatorname{vec} X$ is the vector obtained by stacking its columns.

$$\mathbb{C}^{n \times n} \rightarrow \mathbb{C}^{n^2}$$

The map $\quad X \rightarrow BXA \qquad A, B, X \in \mathbb{C}^{n \times n}$

is linear.

There must be a matrix such that

$$\boxed{\operatorname{vec}(BXA) = M \operatorname{vec}(X)} \quad \leftarrow$$

$$M = A^T \otimes B = \begin{bmatrix} a_{11} B & a_{21} B & B & \cdots & a_{n1} B \\ a_{12} B & & & & \\ \vdots & & & & \\ a_{1n} B & \cdots & \cdots & \cdots & a_{nn} B \end{bmatrix} \quad \text{block matrix}$$

$\uparrow$

⚠ Truly a T, not a $*$ (or an H, $\dagger$).

$$(A^T \otimes B)(C^T \otimes D) = A^T C^T \otimes BD$$

With these tools, we can turn (S) into a vector problem:

$$X - A^* X A = Q \iff \mathrm{vec}(X) - \mathrm{vec}(A^* X A) = \mathrm{vec}(Q)$$

$$\iff \mathrm{vec}(X) - (A^T \otimes A^*)\,\mathrm{vec}(X) = \mathrm{vec}(Q)$$

$$= \Big(I_{n^2} - \underbrace{\boxed{A^T \otimes A^*}}_{M}\Big)\,\mathrm{vec}\,X = \mathrm{vec}\,Q$$

$n^2 \times n^2$ linear system.

$O(n^6)$

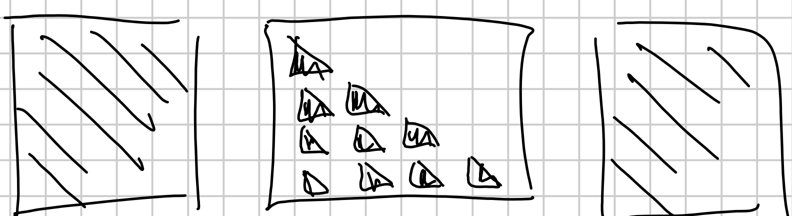
Let $M = A^T \otimes A^*$, let $A = U T U^*$ Schur factorization

U unitary, T upper triangular

$$T = \begin{bmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix} \qquad \mathrm{diag}(T) = \Lambda(A)$$

$$M = (UTU^*)^T \otimes (UTU^*)^* = \bar{U} T^T U^T \otimes U T^* U^*$$

$$= (\bar{U} \otimes U) \cdot (T^T \otimes T^*) \cdot (U^T \otimes U^*)$$



orthogonal

orthogonal*

This is a Schur-like factorization!

$$\Lambda(M) = \Lambda(\text{lower tr. matrix}) = (\lambda_i \bar{\lambda}_j \quad i,j = 1, \dots n)$$

where $(\lambda_1, \dots \lambda_n) = \mathrm{eig}(A)$

Stein equation $\Leftrightarrow$ $(I-M)\,vec(X) = vec(Q)$

$\Lambda(M)$ given by expression above

If $\rho(A) < 1$, then $\rho(M) < 1$

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

$$vec(X) = (I + M + M^2 + \dots)\, vec(Q) = vec(Q) + M\, vec(Q) + M^2 vec(Q) + \dots$$

$$X = Q + A^* X A + (A^*)^2 X A^2 + \dots$$

$$\begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix} \otimes \begin{bmatrix} \mu & 1 \\ 0 & \mu \end{bmatrix} = \begin{bmatrix} \lambda\mu & \lambda & \mu & 1 \\ & \lambda\mu & 0 & \mu \\ & & \lambda\mu & \lambda \\ 0 & & & \lambda\mu \end{bmatrix} - \lambda\mu I = \text{rank } 2$$

Controllability of dynamical systems

$$x_{k+1} = A x_k + b u_k \quad \Leftrightarrow \quad \dot{x}(t) = Ax(t) + bu(t)$$

We can alter the behavior of this system by adding a multiple of $b \in \mathbb{C}^n$ at each step

$$x_1 = A x_0 + b u_0$$
$$x_2 = A^2 x_0 + A b u_0 + b u_1$$
$$x_3 = A^3 x_0 + A^2 b u_0 + A b u_1 + b u_2$$
$$\vdots$$
$$x_k = A^k x_0 + A^{k-1} b u_0 + A^{k-2} b u_1 + \dots + b u_{k-1}$$

If I choose u appropriately, I can reach every possible x_k in the space

$$A^k x_0 + \underbrace{\mathrm{span}(b, Ab, A^2b, \dots A^{k-1}b)}_{}$$

$= K_k(A,b)$ Krylov space

$= \{p(A)b : p(z) \text{ polynomial of degree } < k\}$

The vectors x_F then I can reach as final states starting from $x_0 = 0$ are those in

$$K(A,B) = \mathrm{span}([b, Ab, A^2b, \dots, A^k b, \dots])$$

remark: if $A \in \mathbb{C}^{n\times n}$ matrix, then A^n is a linear combination of $I, A, A^2, \dots, A^{n-1}$ (Cayley-Hamilton theorem)

Def: a pair (A,b) (or a dynamical system $x_{k+1} = Ax_k + bu_k$) is called controllable if $K(A,B) = \mathbb{C}^n$

$\Rightarrow$ we can choose (u_k) to reach every possible vector x_F.

Note that there are indeed examples where the $A^i b$ are very linearly dependent, e.g.

$$\boxed{A = \begin{bmatrix} * & * \\ 0 & 1 \end{bmatrix} \quad b = \begin{bmatrix} 1 \\ 0 \end{bmatrix}}$$

$x_{k+1} = Ax_k + b u_k$

second entry constant!

$$A^k = \begin{bmatrix} * & * \\ 0 & 1 \end{bmatrix} \qquad A^k b = \begin{bmatrix} * \\ 0 \end{bmatrix}$$

$$K(A,b) = \mathrm{span}\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} * \\ 0 \end{bmatrix}, \begin{bmatrix} * \\ 0 \end{bmatrix}, \dots\right)$$

$$= \mathrm{span}\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right)$$

Stein equations and controllability:

Consider the equation

$$x_{k+1} = A x_k \;\to\; \boxed{X - AXA^* = G \qquad G = bb^*} \quad (CS)$$

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

(CS) has solution

$$X = G + AGA^* + A^2 G (A^*)^2 + \dots$$
$$= bb^* + (Ab)(Ab)^* + (A^2 b)(A^2 b)^* + \dots$$
$$= [b \;\; Ab \;\; A^2 b \;\; \cdots] \cdot [b \;\; Ab \;\; A^2 b \;\; \cdots]^*$$

$X \succ 0 \iff [b \;\; Ab \;\; A^2 b \;\; \cdots]$ has full row rank

$\iff \operatorname{span}(b, Ab, A^2 b, \dots) = \mathbb{C}^n$

$\iff (A, b)$ controllable.

To check if I can reach every possible state x_F by altering (u_k), it is sufficient to

1. solve $X - AXA^* = bb^*$
2. check if $X \succ 0$ or not.

$$\operatorname{Im}[b \;\; Ab \;\; A^2 b \;\; \cdots] = \operatorname{Im}[b \;\; Ab \;\; A^2 b \;\; \cdots \;\; A^{n-1} b]$$

Suppose you want to reach x_F (given) starting from $x_0 = 0$. Among all possible sequences that get to x_F at time k, i.e. $x_k = x_F$, we want to find the one with smallest

$$|u_0|^2 + |u_1|^2 + |u_2|^2 + \dots + |u_{k-1}|^2$$

$$x_k = \cancel{A^k x_0} + A^{k-1} b u_0 + \cdots + A b u_{k-2} + b u_{k-1}$$

$$\begin{bmatrix} A^{k-1}b & A^{k-2}b & \cdots & Ab & b \end{bmatrix} \begin{bmatrix} u_0 \\ u_1 \\ \vdots \\ u_{k-1} \end{bmatrix} = x_k$$

smallest-norm solution :
$$\begin{bmatrix} u_0 \\ \vdots \\ u_{k-1} \end{bmatrix} = \begin{bmatrix} A^{k-1}b & \cdots & Ab & b \end{bmatrix}^+ x_F$$

$$\|u\|^2 = x_F^* \left(\begin{bmatrix} A^{k-1}b & \cdots & Ab & b \end{bmatrix}^+ \right)^* \begin{bmatrix} A^{k-1}b & \cdots & Ab & b \end{bmatrix}^+ x_F$$

$$x_F^* \left(bb^* + Abb^*A^* + A^2bb^*(A^*)^2 + \cdots + A^{k-1}bb^*(A^*)^{k-1} \right)^{-1} x_F$$

If $k \to \infty$, we get

$x_F^* X^{-1} x_F$ = minimum amount of energy needed to reach x_F from $x_0 = 0$

$$F = \begin{bmatrix} b & Ab & \cdots & A^{k-1}b \end{bmatrix}$$

$$F^* = QR$$

a

$$(F^+)(F^+)^* = R^{-1} \cancel{Q^* Q} R^{-*}$$
$$= R^{-1} R^{-*}$$

$$R^{-1} Q^* = A^* \cdot a$$