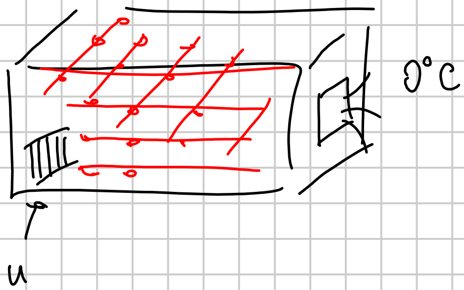
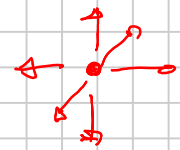
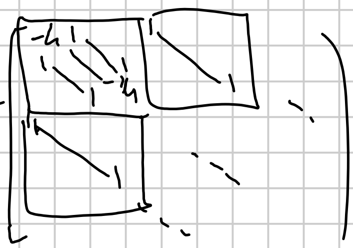
Large and sparse control systems

Ex:
3D version of heating control problem:





$$A = \begin{bmatrix} -2 & 1 & & \\ 1 & \ddots & \ddots & \\ & \ddots & \ddots & 1 \\ & & 1 & -2 \end{bmatrix}$$ in the 1D version,

$A=$  2D

In general, A large and sparse.

$v \to Av$
$v \to (A-\alpha I)^{-1} v$ the only maps that we will need.

$A \in \mathbb{C}^{n\times n}$ matrix n very large
$B \in \mathbb{C}^{n\times m}$ tall thin $m << n$

How do find stabilizability, optimal controls, etc.

In this lecture: just one simple problem,
solving the Lyapunov equation

$$\boxed{AX + XA^* + bb^* = 0}$$

$A \in \mathbb{C}^{n\times n}$, $\Lambda(A) \subset LHP$
$b \in \mathbb{C}^{n\times 1}$

Once we know how to do this,
if B has dimension $m>1$

$$B=[b_1\ b_2\ \cdots\ b_m]$$

then we can compute the solution X to $AX+XA^*+BB^*$
as $X=X_1+X_2+\cdots+X_m$

X_i solves $AX_i+X_iA^*+b_ib_i^*=0$

since $BB^*=b_1b_1^*+b_2b_2^*+\cdots+b_mb_m^*$, and the equation is linear.

Also, to solve Riccati equations we can use Newton's method once we have an algorithm for Lyapunov eqs.

Problem: even if A sparse, $Q=bb^*$ low-rank,
the solution to $AX+XA^*+bb^*=0$
is not sparse; how to store it?
(and also if (A,b) controllable $X>0$ has full rank)

Solution: in many practical cases, X has eigenvalues that decay quickly (exponentially), so we can approximate it with a low-rank matrix.

Lyapunov equation $\boxed{AX+XA^*+bb^*=0}$ (L)

We wish to convert it to a discrete-time problem equiv to a Stein equation

$$\boxed{X-\hat{A}X\hat{A}^*=\hat{b}\hat{b}^*} \quad (S)$$

Theorem: Let $\tau>0$. X solves (L) if and only if

X solves the Stein equation (S) with

$$\hat{A} = (A-\tau I)^{-1}(A+\tau I), \quad \hat{b} = \sqrt{2\tau}\,(A-\tau I)^{-1} b$$

<u>Proof</u>: First note that $\Lambda(A-\tau I) \subset LHP$ if $\tau > 0$, so it is invertible

$$(S) \Leftrightarrow X - (A-\tau I)^{-1}(A+\tau I)\, X (A+\tau I)^{*}(A-\tau I)^{-*} = 2\tau\, (A-\tau I)^{-1} b b^{*} (A-\tau I)^{-*}$$

Multiply by $(A-\tau I)$ and $(A-\tau I)^{*}$ on the left and right

$$\Leftrightarrow (A-\tau I)X(A-\tau I)^{*} - (A+\tau I)X(A+\tau I)^{*} = 2\tau\, b b^{*}$$

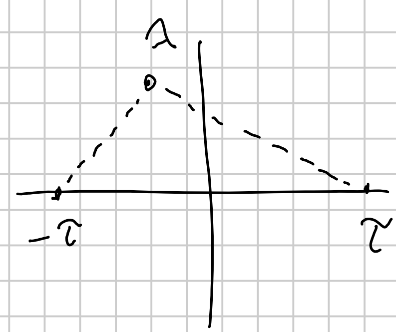
$$\Leftrightarrow -\tau X A^{*} - A X \tau - \tau X A^{*} - A X \tau = 2\tau\, b b^{*}$$

$$\Leftrightarrow 2\tau\left(AX + XA^{*} + bb^{*}\right) = 0 \Leftrightarrow (L)$$

<u>Lemma</u> if $\Lambda(A) \subset LHP$, then $\Lambda(\hat{A}) \subset$ Unit disc (with $\hat{A} = (A-\tau I)^{-1}(A+\tau I)$)

<u>Proof</u>: If $\{\lambda_1, \lambda_2, \dots \lambda_n\} = \Lambda(A)$, then

$$\Lambda(\hat{A}) = \left\{ \frac{\lambda_i + \tau}{\lambda_i - \tau}, \; i = 1, 2, \dots n \right\}$$



If $\lambda \in LHP$, $\mathrm{dist}(\lambda, -\tau) < \mathrm{dist}(\lambda, \tau)$

hence $\left|\frac{\lambda+\tau}{\lambda-\tau}\right| < 1$.

(<u>Remark</u>: note that solving $\dot{x} = Ax$ with the trapezoidal rule would have returned up to scaling $x_{k+1} = \hat{A} x_k$)

How to solve $X - \hat{A} X \hat{A}^{*} = \hat{b}\hat{b}^{*}$?

<u>Idea</u>: fixed point iteration

$$\begin{cases} X_0 = 0 \\ X_{k+1} = \hat{A} X_k \hat{A}^* + \hat{b}\hat{b}^* \end{cases}$$

$X_1 = \hat{b}\hat{b}^*$ $\quad X_2 = \hat{A}\hat{b}\hat{b}^*\hat{A}^* + \hat{b}\hat{b}^*$ $\quad X_3 = \hat{A}^2\hat{b}\hat{b}^*\hat{A}^* + \hat{A}\hat{b}\hat{b}^*\hat{A}^* + \hat{b}\hat{b}^*$

$$X_k = \sum_{i=0}^{k-1} \hat{A}^i \hat{b}\hat{b}^* (\hat{A}^*)^i$$

Since $\Lambda(\hat{A}) \subset$ unit disc, $\hat{A}^i \to 0$

Lemma: if X is the exact solution of the Stein equation, then $X_k - X = \hat{A}^k (X_0 - X) \hat{A}^k$

Proof: induction: for $k=0$ it is trivial, then

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

$$= \hat{A}\, \hat{A}^k (X_0 - X) \hat{A}^{k*}\, \hat{A}^*$$

ind. hypothesis

Hence $\|X_k - X\| \sim \rho(\hat{A})^{2k} \|X_0 - X\|$, linear convergence

Low rank formulation:

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

$$= \underbrace{\begin{bmatrix} \hat{b} & \hat{A}\hat{b} & \hat{A}^2\hat{b} & \cdots & \hat{A}^{k-1}\hat{b} \end{bmatrix}}_{Z} \cdot \underbrace{\begin{bmatrix} \hat{b} & \hat{A}b & \hat{A}^2 b & \cdots & \hat{A}^{k-1} b \end{bmatrix}^*}_{Z^*}$$

Algorithm: (ADI iteration with single shift)

Start from $v = \hat{b}$, $Z = [\hat{b}]$

for $k = 1, 2, 3, \dots$

$v \leftarrow \hat{A} v = (A - \tau I)^{-1}(A + \bar{\tau} I) v = (A - \tau I)^{-1}(A - \tau I + 2\tau I) v$

$Z \leftarrow [Z \;\; v]$ $\qquad\qquad = v + 2\tau(A-\tau I)^{-1}v$

end

We can return Z s.t. $X \approx ZZ^*$

From the convergence speed bound

$$X_k - X = \hat{A}^k (X_0 - X)\hat{A}^{*k}$$

we get

$$\|X_k - X\| \leq \|\hat{A}^k\|^2 \cdot \|X\|$$

$$\frac{\|X_k - X\|}{\|X\|} \leq \underbrace{\|\hat{A}^k\|^2}_{\rho(A)^k}$$

If I stop after k iterations,

$$\frac{\|X_k - X\|}{\|X\|} \leq \underbrace{\|\hat{A}^k\|^2}_{\rho(\hat{A})^k}$$

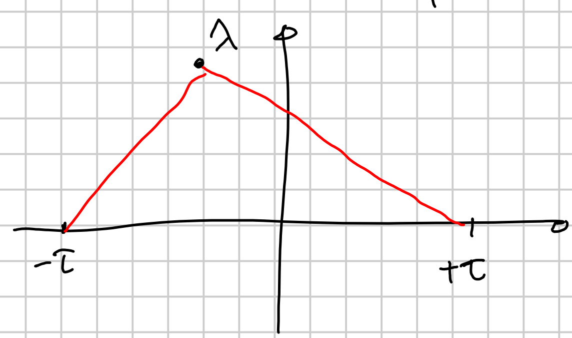
In particular, X is at distance at most $\|\hat{A}^k\|^2$ (relative)

from a matrix $X_k = Z_k Z_k^*$ of rank k.

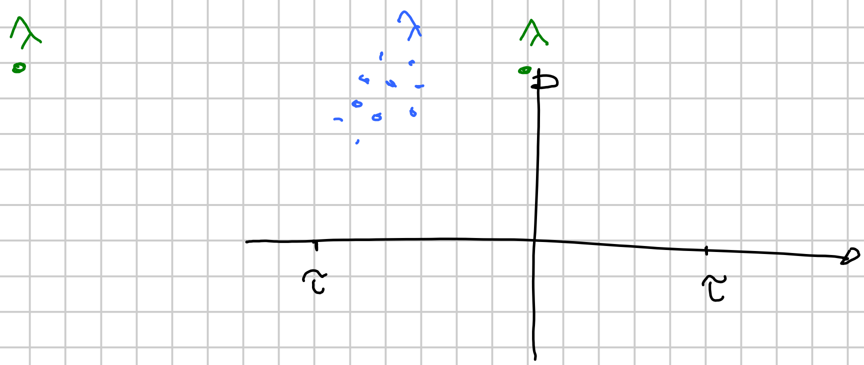
$\Rightarrow$ if $\rho(\hat{A})$ is small, X is well-approximated by a low-rank matrix.

What choice of τ makes $\rho(\hat{A})$ as small as possible?

$$\Lambda(\hat{A}) = \left\{ \frac{\lambda+\tau}{\lambda-\tau} : \lambda \in \Lambda(A) \right\}$$



Observation: τ should be "around" the eigenvalues of A



If $\mathrm{Re}(\lambda) \ll \tau$ $\qquad \dfrac{\lambda-\tau}{\lambda+\tau} \approx 1$

If $|\mathrm{Re}(\lambda)| \ll \tau$ $\qquad \dfrac{\lambda-\tau}{\lambda+\tau} \approx -1$

In particular, if A has both very small and large $|\mathrm{Re}(\lambda)|$ then $\rho(\hat{A})$ will be very close to 1, and the algorithm will be very slow

Idea to avoid this: change τ at every step!

$$\hat{A}_k = (A-\tau_k I)^{-1}(A+\tau_k I) \qquad \hat{b}_k = \sqrt{2\tau}\,(A-\tau_k I)^{-1} b$$

$$X_{k+1} = \hat{A}_k X_k \hat{A}^* + \hat{b}_k \hat{b}_k^*$$ a different fixed-pt eqn at each step

One can still rearrange it in a low-rank formulation.

Error bound (proved analogously):

$$X_k - X = \underbrace{\hat{A}_k \hat{A}_{k-1} \cdots \hat{A}_1}(X_0 - X)\hat{A}_1^* \hat{A}_2^* \hat{A}_3^* \cdots \hat{A}_k^*$$

$$\rho(\hat{A}_k \hat{A}_{k-1} \cdots A_1) = \max_{\lambda \in \Lambda(A)} \frac{|(\lambda+\tau_1)(\lambda+\tau_2)\cdots(\lambda+\tau_k)|}{|(\lambda-\tau_1)(\lambda-\tau_2)\cdots(\lambda-\tau_k)|}$$

If $\Lambda(A)$ is composed of only k distinct real negative eigenvalues, then after K steps I can reach $\rho(\hat{A}_k \cdots \hat{A}_1) = 0$ by choosing $\tau_i = -\lambda_i$ (numerator becomes zero).

This is an interesting approximation problem: find $\tau_1, \tau_2, \ldots \tau_k$ s.t. $p(x) = (x-\tau_1)(x-\tau_2)\cdots(x-\tau_k)$ is a polynomial of degree k with

$|p(\lambda_i)|$ small for all i, $|p(-\lambda_i)|$ large for all i

---

With multiple shifts, $\hat{A}_k = (A-\tau_k I)^{-1}(A+\tau_k I)$

$$Z = \left[\hat{b}_k \quad \hat{A}_k \hat{b}_{k-1} \quad \hat{A}_k \hat{A}_{k-1} \hat{b}_{k-2} \quad \cdots \quad \hat{A}_k \hat{A}_{k-1} \cdots \hat{A}_2 \hat{b}_1\right]$$

Im Z lives in a certain rational Krylov subspace: all the columns are of the form

$r(A)b$, where $r(x) = \dfrac{p(x)}{q(x)}$ and

$$q(x) = (x-\tau_k)(x-\tau_{k-1})\cdots(x-\tau_2)$$

This suggests other methods based on Arnoldi: first compute a basis V_k s.t. $AV_{k+1} = V_k H_{k,k+1}$

then look for $X \approx V_k Y V_k^*$

$AX + XA^* + bb^* = 0 \iff$

$$\iff AV_k Y V_k^* + V_k Y V_k^* A^* + bb^* = 0$$

$$V_k^* A V_k Y + Y V_k^* A V_k + V_k^* b b^* V_k = 0$$

$$A_k Y + Y A_k + \beta e_1 e_1^* \beta^* = 0$$ a smaller-size Lyapunov equation.