Singular Value Decomposition Applications

Gabriel Hugh Elkaim

Winter 2016

Singular Value Decomposition Applications

- General pseudo-inverse
- Full SVD
- Image of Unit Ball under linear transformation
- SVD in estimation/inversion
- Sensitivity of linear equations to data error
- Low rank approximation via the SVD

Full SVD (1.3)

$A \in \mathbb{R}^{m \times n}$ $\quad$ $\text{rank}(A) = r$

$$A = U_1 \Sigma_1 V_1^T = \begin{bmatrix} u_1 & \cdots & u_r \end{bmatrix} \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r \end{bmatrix} \begin{bmatrix} v_1^T \\ \vdots \\ v_r^T \end{bmatrix}$$

$U = [U_1 \,\vdots\, U_2] \in \mathbb{R}^{m \times m}$ $\quad$ $R(U) = \mathbb{R}^m$

$V = [V_1 \,\vdots\, V_2] \in \mathbb{R}^{n \times n}$ $\quad$ $R(V_1) = N(A)^{\perp}$

$$\Sigma = \left[\begin{array}{c|c} \underset{r \times r}{\Sigma_1} & \underset{r \times (n-r)}{\phi} \\ \hline \underset{(m-r) \times r}{\phi} & \underset{(m-r) \times (m-r)}{\phi} \end{array}\right]$$

Full SVD (2.3)

$$A = \begin{bmatrix} U_1 & \vdots & U_2 \end{bmatrix} \left[\begin{array}{c|c} \Sigma_1 & \phi \\ \hline \phi & 0 \end{array}\right] \begin{bmatrix} V_1^T \\ \cdots \\ V_2^T \end{bmatrix} = U \Sigma V^T$$

Σ^{-1} no longer allowed

σ_1 unambiguous

svd(A)

svd(A, 'econ')

σ_{min} ?? 0 $\underline{\text{or}}$ σ_r

Full SVD (3.3)

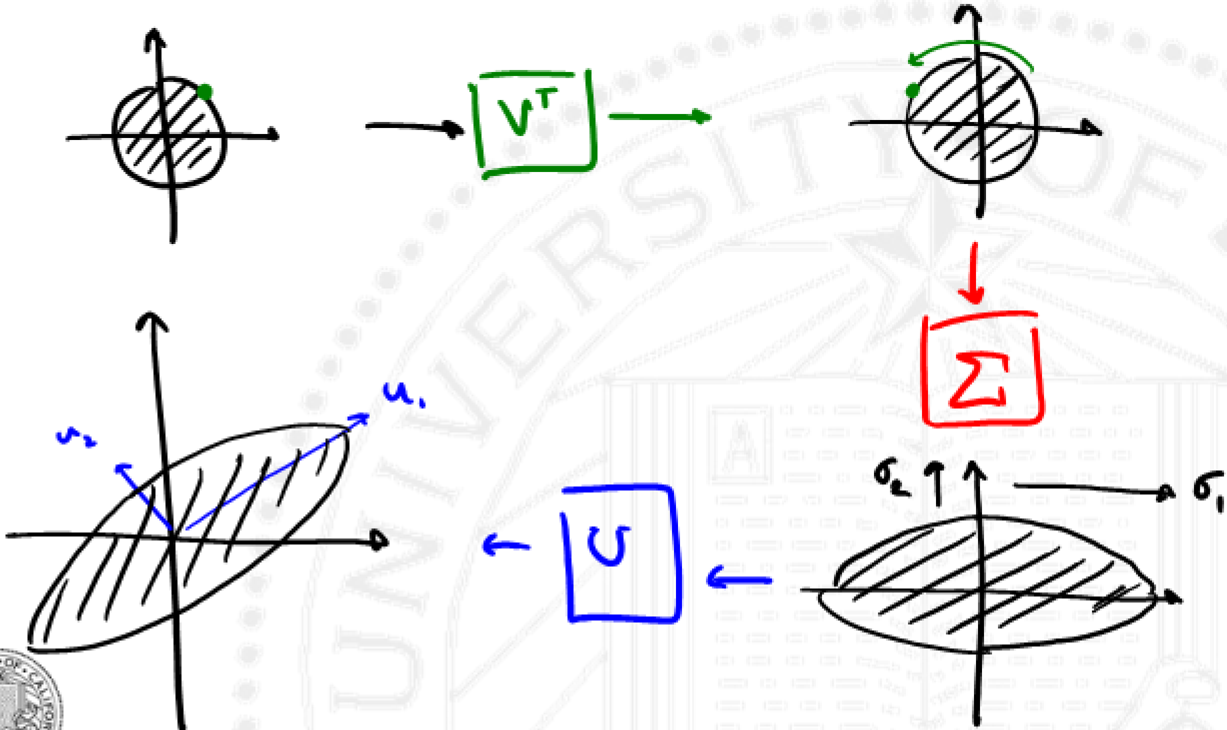
$$A = U\Sigma V^T$$

$$x \longrightarrow \boxed{V^T} \longrightarrow \boxed{\Sigma} \longrightarrow \boxed{U} \longrightarrow Ax$$

$$\{x \mid \|x\| \le 1\}$$

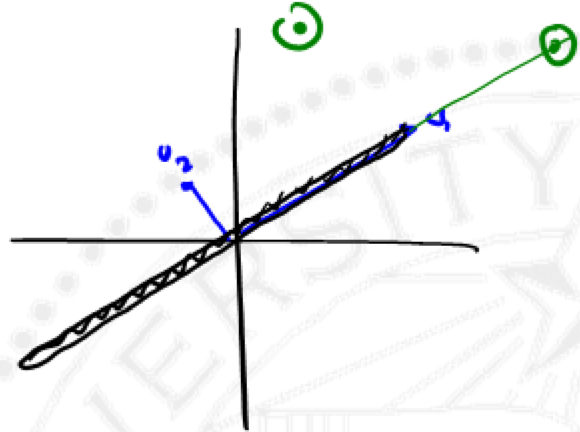
Unit Ball under Linear Transformation (1.3)

$\{ x \mid \|x\| \leq 1 \}$



Unit Ball under Linear Transformation (2.3)

$$\Sigma = \begin{bmatrix} 2 & \\ & 1/1000 \end{bmatrix}$$



$$\Sigma' = \begin{bmatrix} 10 & & \\ & 9 & \\ & & 0.01 \end{bmatrix}$$

Unit Ball under Linear Transformation (3.3)

$A = U \Sigma V^T$

$A^{+} = V \Sigma^{-1} U^T$

$$\begin{bmatrix} 1/\sigma_1 & & \\ & \ddots & \\ & & 1/\sigma_n \end{bmatrix}$$

SVD in estimation/inversion (1.3)

$y = Ax + v$

$y \in \mathbb{R}^m$ measurement

$x \in \mathbb{R}^n$ vector to be estimated

v is an unknown noise $\|v\| < \alpha$

$\hat{x} = A^{\dagger} y.$

$\hat{x} = By$ $BA = I$ unbiased estimator $\hat{x} = x$ if $v = 0$

$\tilde{x} \triangleq \hat{x} - x = B(Ax + v) - x = \cancel{BAx} + Bv - \cancel{x} = Bv$

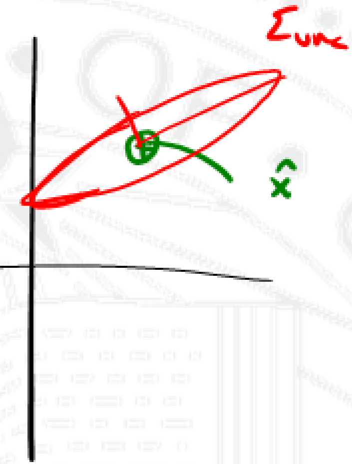
$\tilde{x} \in \mathcal{E}_{unc} = \{ Bv \mid \|v\| \leq \alpha \}$

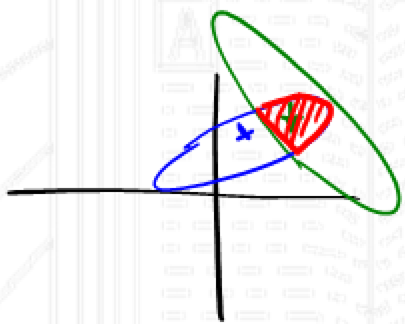
SVD in estimation/inversion (2.3)

$$\tilde{x} \in \mathcal{E}_{unc} = \{ Bv \mid \|v\| \le \alpha \}$$

$$x = \hat{x} - \tilde{x} = \hat{x} + \mathcal{E}_{unc}$$

True x lies within the $\mathcal{E}_{unc}$ centered at $\hat{x}$.





SVD in estimation/inversion (3.3)

Semi-major axes of $\mathcal{E}_{unc}$ $\alpha \cdot \sigma_i u_i$ singular values and right singular vectors of B.

$$\|\hat{x} - x\| \le \alpha \|B\|$$

$$BA = I$$

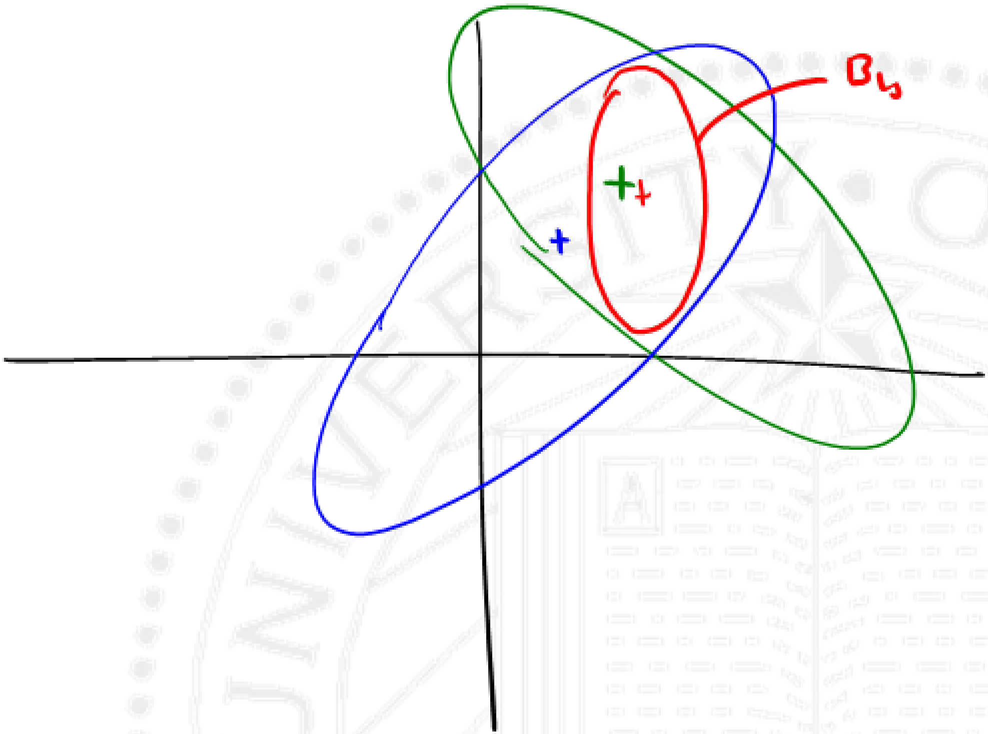
$$B_{ls} = A^{\dagger}$$

$$B_{ls} B_{ls}^T \le B B^T$$

$$\sigma_i(B_{ls}) \le \sigma_i(B) \quad i = 1 \ldots n$$

$$\|B_{ls}\| \le \|B\| \qquad \mathcal{E}_{ls} \le \mathcal{E}_{unc_B}$$

Example (1.3)

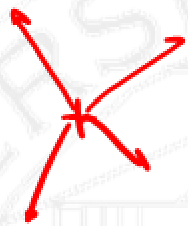


Example (2.3)

$y_i = -k_i^T x + v_i$

$A_1 = \begin{bmatrix} k_1^T \\ k_2^T \end{bmatrix}$

$A_2 = \begin{bmatrix} k_1^T \\ \vdots \\ k_4^T \end{bmatrix}$

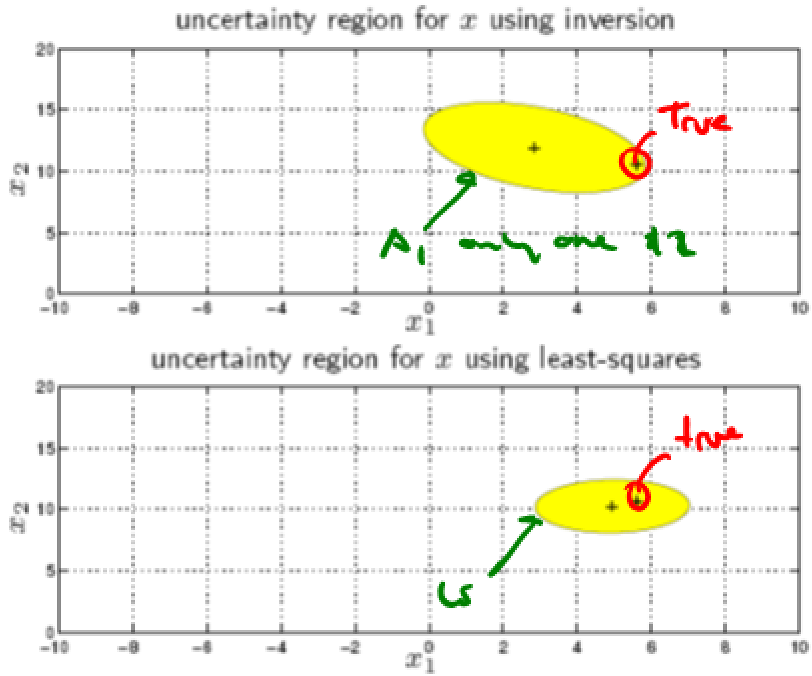


$\hat{x} = A_1^{-1} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$

$\hat{x} = A_2^{\dagger} \begin{bmatrix} y_1 \\ \vdots \\ y_4 \end{bmatrix}$

$\alpha = 1$

Example (3.3)


uncertainty region for x using inversion
true
A_1 only one t2
x_1
x_2
uncertainty region for x using least-squares
true
LS

Proof of Optimality Property (1.3)

$A \in \mathbb{R}^{m \times n}$ $m \geq n$ full rank

$A = U \Sigma V^T$
Σ: $n \times n$

$B_{ls} = A^{\dagger} = V \Sigma^{-1} U^T$ $B_{ls} \cdot A = V \Sigma^{-1} U^T U \Sigma V^T = I$

$Z \triangleq B - B_{ls}$ $\therefore$ $B = B_{ls} + Z$

$ZA = ZU\Sigma V^T = 0$

$(ZU\Sigma V^T) V \Sigma^{-1}$

$ZU = 0$

$Z = \phi.$

Proof of Optimality Property (2.3)

$$BB^T = (B_{ls} + Z)(B_{ls} + Z)^T$$

$$= B_{ls}B_{ls}^T + B_{ls}Z^T + ZB_{ls}^T + ZZ^T$$

$$= B_{ls}B_{ls}^T + ZZ^T$$

$$ZB_{ls}^T = (ZU)\Sigma^{-1}V^T = 0$$

$$\underline{B_{ls}B_{ls}^T \le B_{ls}B_{ls}^T + ZZ^T}$$

Sensitivity of Linear Equations to Data Error (1.3)

$y = Ax \quad A \in \mathbb{R}^{n \times n} \quad A^{-1}$ exists

$$x = A^{-1}y \longrightarrow (x + \delta x) = A^{-1}(y + \delta y)$$

$$\delta x = A^{-1}\delta y$$

$$A = U\Sigma V^T$$

$$A^{-1} = V\Sigma^{-1}U^T = \underbrace{VR}_{\tilde{V}}\,\tilde{\Sigma}\,\underbrace{RU^T}_{\tilde{U}^T}$$

$$R = \begin{bmatrix} & & 1 \\ & \iddots & \\ 1 & & \end{bmatrix}$$

$$\tilde{\Sigma} = \begin{bmatrix} 1/\sigma_n & & \\ & \ddots & \\ & & 1/\sigma_1 \end{bmatrix}$$

Sensitivity of Linear Equations to Data Error (2.3)

$$\|\delta x\| = \|A^{-1}\delta y\| \le \|A^{-1}\|\|\delta y\|$$

if $\|A^{-1}\|$ is large

- small errors in y can lead to big errors in x
- can't solve for x given y with small errors
- A may be full rank, but in practice is singular

Sensitivity of Linear Equations to Data Error (3.3)

$y = Ax \qquad \|y\| \leq \|A\| \|x\|$

$$\frac{\|\delta x\|}{\|x\|} \leq \|A\| \|A^{-1}\| \frac{\|\delta y\|}{\|y\|}$$

perturbation in y as %

perturbation in x as %

$$\kappa(A) = \|A\| \cdot \|A^{-1}\| = \sigma_{max}(A) / \sigma_{min}(A) \geq 1$$

$\mathrm{cond}(A) \qquad \mathrm{rcond}(A) = \frac{1}{\kappa(A)} \qquad 1 \to 0$

Relative error in solution $\leq$ κ · relative error in data

Low rank approximations (1.3)

bits in solution $\simeq$ # bits in data $-\log_2 \kappa$

κ small $\to 1$ $\Rightarrow$ "well conditioned"

κ large $\to \infty$ $\Rightarrow$ "ill conditioned"

$$\kappa(A) \triangleq \frac{\sigma_{max}(A)}{\sigma_{min}(A)}$$

Low rank approximations (2.3)

$A \in \mathbb{R}^{m \times n}$ rank$(A) = 1$

$A = U \Sigma V^T = \sum_{i=1}^{r} \sigma_i \underbrace{u_i v_i^T}_{\text{rank 1 dyads}}$

$\hat{A} \rightarrow \text{rank}(\hat{A}) \leq p < r$ such that $\hat{A} \approx A$

$\min_{\hat{A}} \|A - \hat{A}\|$

$\|A\|_{FR} = \|A(:)\|$

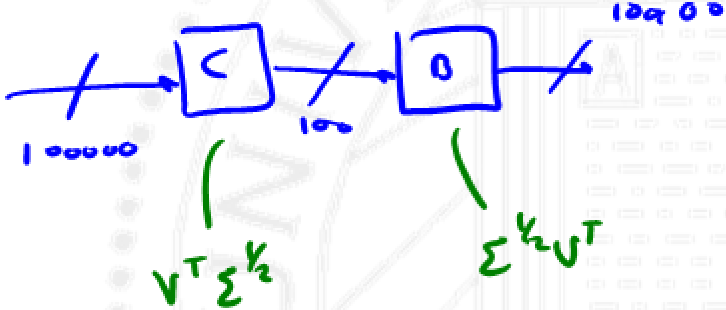
$\hat{A} = \sum_{i=1}^{p} \sigma_i u_i v_i^T$

Low rank approximations (3.3)

$$\|A - \hat{A}\| = \left\| \sum_{i=p+1}^{r} \sigma_i u_i v_i^T \right\| = \sigma_{p+1}$$

$u_i v_i^T$ — in order of importance

$$A = BC$$



Distance to Singularity (1.3)

σ_i :

$$\sigma_i = \min \{ \|A - B\| \mid \operatorname{rank}(B) \le (i-1) \}$$

$\{15, 10, 3, 2, 1\times10^{-5}\}$

$y = Ax + v \qquad \|v\| \le 0.01$

small σ_{min} means that A is close to singular matrix.

Distance to Singularity (2.3)

$y = \underbrace{Ax}_{\text{signal}} + \overbrace{v}^{\text{noise}}$

$A \in \mathbb{R}^{100 \times 30}$

$\Sigma = \{10, 7, 3, \tfrac{1}{2}, \boxed{\tfrac{1}{100}, \tfrac{1}{1000} \ldots}\}$

$\|x\|$ is on the order of 1

$\|v\|$ is on the order of 0.1

$\sigma_i u_i v_i^T = \phi$ for $i = 5 \ldots 30$

$$\boxed{y = \sum_{i=1}^{4} \sigma_i u_i v_i^T x + v}$$

Better model

Distance to Singularity (3.3)

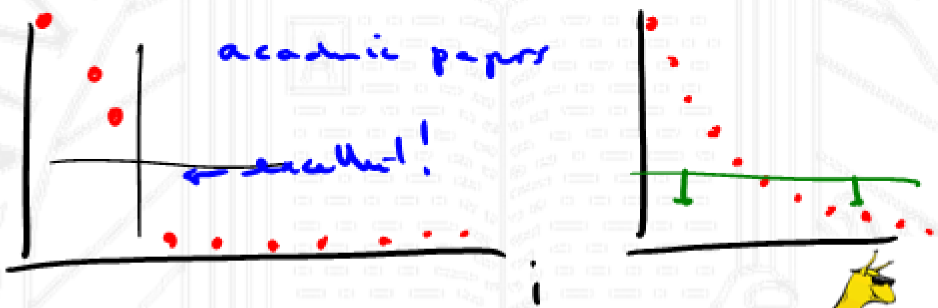
$\{a_1 \ldots a_{100}\} \in \mathbb{R}^{10}$

stock price data in time (100 days)

$[a_1 \ldots a_{100}] = A \in \mathbb{R}^{10 \times 100}$ QR → 10.

SVD: A → 3 significant singular values.

log(σ_i)



Questions?

//

Final

- Thursday — Friday　　8
- Friday — Saturday
- Tuesday/Wed (3)

24 hours open everything except published solution or people.

Controllability and State Transfer

Gabriel Hugh Elkaim

Winter 2016

Controllability and State Transfer

- State transfer
- Reachable set
- Controllability Matrix
- Minimum Norm inputs
- Infinite-horizon minimum norm transfer

State Transfer (1.3)

$\dot{x} = Ax + Bu$ $\quad$ $x_{k+1} = Ax_k + Bu_k$ $\quad$ $[t_i, t_f]$

$u : [t_i, t_f] \to \mathbb{R}^m$ steers or transfers $x(t_i)$ to $x(t_f)$

– where can $x(t_f)$ be given $x(t_i)$ @ $t = t_i$?

– how quickly can I get to x_{target} ?

– how do I determine u that gets me there?

– how do I find a small or "efficient" u that takes me from $x(t_i)$ to $x_{des}(t_f)$.

State Transfer (2.3)

$t_i = \phi \qquad x(t_i) = x_0 = \phi \leftarrow$ relaxed state

$x(t)$ is REACHABLE in t seconds

$R_t \subseteq \mathbb{R}^n$ set of reachable points in t seconds (t-epochs)

$\dot{x} = Ax + Bu$

$$R_t = \left\{ \int_0^t e^{A(t-\tau)} Bu(\tau)\, d\tau \;\middle|\; u[0,t] \to \mathbb{R}^m \right\}$$

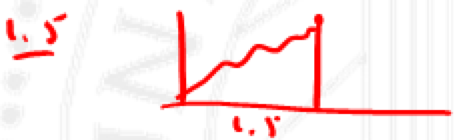
State Transfer (3.3)

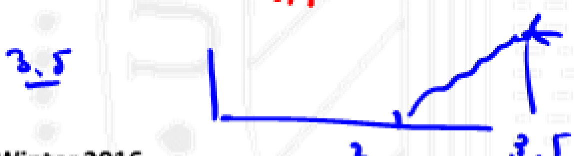
$x_{k+1} = Ax_k + Bu_k$

$$R_t = \left\{ \sum_{\tau=0}^{t-1} A^{t-1-\tau} B u(\tau) \;\middle|\; u_k \in \mathbb{R}^m,\ k=0,\dots,t-1 \right\}$$

R_t subspace of $\mathbb{R}^n$

$R_t \subseteq R_s$ if $t \le s$ (can always reach more points given more time).





Reachability (1.3)

The reachable set of point $\mathcal{R}$ for some t: $\mathcal{R} = \bigcup_{t \geq 0} \mathcal{R}_t$

Reachability for Discrete-Time LDS (1.2)

$x_{k+1} = Ax_k + Bu_k \quad x_k \in \mathbb{R}^n$

$x(0) = 0$

$x_1 = Bu_0$

$x_2 = ABu_0 + Bu_1$

$x_3 = A^2Bu_0 + ABu_1 + Bu_2$

$\vdots$

$$x_k = \mathcal{C}_k \begin{bmatrix} u_{k-1} \\ \vdots \\ u_0 \end{bmatrix}$$

$$\mathcal{C}_k = \begin{bmatrix} B & AB & A^2B & \cdots & A^{k-1}B \end{bmatrix} \leftarrow \text{controllability matrix}$$

Krylov matrix

$$R_k = \text{range}(\mathcal{C}_k)$$

Reachability for Discrete-Time LDS (2.2)

C-H $\rightarrow$ A^k is a linear combination $A^0 \ldots A^{k-1}$

$k \geq n$ $\quad \text{Range}(\mathcal{C}_k) = \text{Range}(\mathcal{C}_n)$

$$R_k = \begin{cases} \text{Range}(\mathcal{C}_k) & k < n \\ \text{Range}(\mathcal{C}_n) & k \geq n \end{cases}$$

$\mathcal{C} = [B \;\; AB \;\; A^2B \;\; \ldots \;\; A^{n-1}B]$ $\quad$ range($\mathcal{C}$) anything I can hit with $\underline{x}$.

if ($\mathcal{C}$ is full rank) $\leftarrow$ reach each of $\mathbb{R}^n$ in n steps.

Controllable System (1.2)

Controllable: if all state are reachable

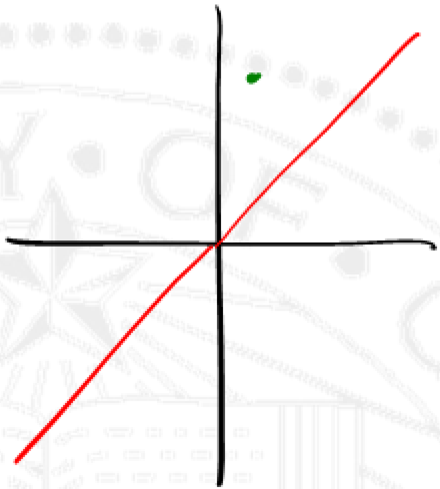
$$\mathcal{R} = \mathbb{R}^n \quad \text{Rank}(\mathcal{C}) = n$$

Controllable System (2.2)

$$x_{k+1} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} x_k + \begin{bmatrix} 1 \\ 1 \end{bmatrix} u_k$$

$$\mathcal{C} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \quad \text{rank}(\mathcal{C}) \neq 2$$

$$R = \text{range}(\mathcal{C}) = \{x \mid x_1 = x_2\}$$



General State Transfer (1.2)

$t_f > t_i$

$$x(t_f) = A^{t_f - t_i}\, x(t_i) + \mathcal{C}_{t_f - t_i} \begin{bmatrix} u(t_f - 1) \\ \vdots \\ u(t_i) \end{bmatrix}$$

$x(t_i) \longrightarrow x_{des}$ @ $t = t_f$

- general state transfer — reachability problem
- if controllable I can get there in n steps
- drive $x \rightarrow 0$ (regulator)
- if state starts in R will never leave R

General State Transfer (2.2)

$$z \in \text{range}(\mathcal{C}) \longrightarrow Az \in \text{Range}(\mathcal{C})$$

$$z = \sum_{i=1}^{n-1} A^i B u_i$$

$$Az = \sum_{i=0}^{n-1} A^{i+1} B u(i) = \sum_{i=1}^{n-1} A^i B u(i-1) + A^n B u(n-1)$$

$$= \sum_{i=0}^{n-1} A^i B\left(u(i-1) - \alpha_i u(n-1)\right) - \alpha_0 u(n-1) B \in \text{Range}(\mathcal{C})$$

$$\mathcal{X}(s) = s^n + \alpha_{n-1} s^{n-1} + \ldots + \alpha_0 = \det(sI - A)$$

$$A^n = -\alpha_0 I - \ldots - \alpha_0 X^{n-1}$$

Least-Norm Input for Reachability (1.4)

$\mathrm{Rank}(\mathcal{C}) = n \qquad x(0) = 0 \text{ to } x_k = x_{des}$

$$x_{des} = \mathcal{C} \begin{bmatrix} u_{k-1} \\ \vdots \\ u_0 \end{bmatrix}$$

Fat, Full rank

$$\sum_{n=0}^{k-1} \|u_n\|^2 \leftarrow \text{energy.}$$

$$\begin{bmatrix} u_{ln}(k-1) \\ \vdots \\ u_{ln}(0) \end{bmatrix} = \underbrace{\mathcal{C}_k^T (\mathcal{C}_k \mathcal{C}_k^T)^{-1}}_{\text{least norm.}} x_{des}$$

Least-Norm Input for Reachability (2.4)

$$u_{ln}(\tau) = B^T (A^T)^{(t-1-\tau)} \left[\sum_{s=0}^{t-1} A^s B B^T (A^T)^s \right]^{-1} x_{des}$$

$\tau = 0 \ldots t-1$ "matched filtering"

Least-Norm Input for Reachability (3.4)

$\mathcal{E}_{min}$ min value of $\sum_{\tau=0}^{t-1} \|u(\tau)\|^2$ required to reach x_{des} in t epochs

$$\mathcal{E}_{min} = \sum_{\tau=0}^{t-1} \|u_{ln}(\tau)\|^2 = \left(\mathcal{C}_t^T(\mathcal{C}_t\mathcal{C}_t^T)^{-1}x_{des}\right)^T \mathcal{C}_t^T(\mathcal{C}_t\mathcal{C}_t^T)^{-1}x_{des}$$

$$= x_{des}^T \cancel{(\mathcal{C}_t\mathcal{C}_t^T)^{-1}} \cancel{(\mathcal{C}_t\mathcal{C}_t^T)} (\mathcal{C}_t\mathcal{C}_t^T)^{-1} x_{des}$$

$$= \boxed{x_{des}^T (\mathcal{C}_t\mathcal{C}_t^T)^{-1} x_{des}}$$

$$= x_{des}^T \left(\sum_{\tau=0}^{t-1} A^\tau B B^T (A^T)^\tau \right)^{-1} x_{des}$$

Least-Norm Input for Reachability (4.4)

$\Sigma_{min}(x_{des}, t)$ measure of how hard it is to reach x_{des} from $x(0)=0$ in t epoch's

$\Sigma_{min}(x_{des}, t)$ practical measure of controllability/reachability

→ ellipsoid $\{z \mid \Sigma_{min}(z, t) \le 1\}$ reachable area with (unit of energy)

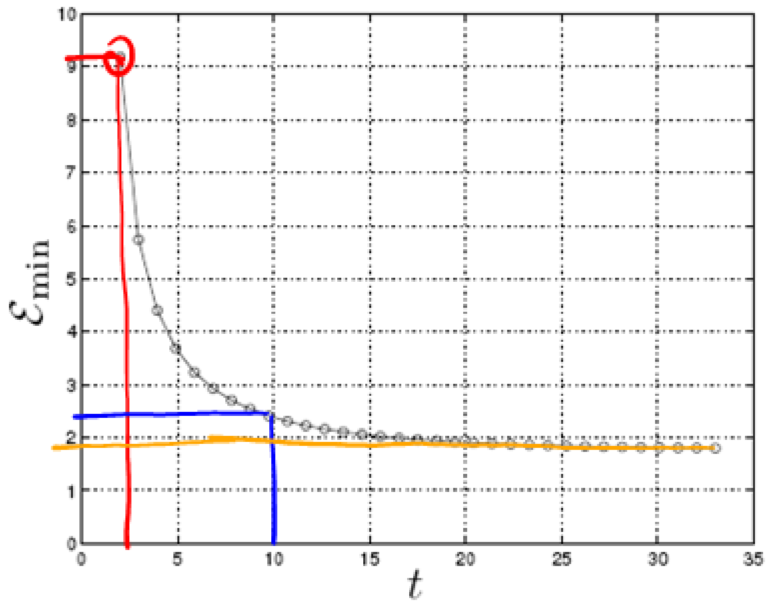
$\Sigma_{min} \simeq f(t) \quad t \ge s$

$$\Sigma_{min}(t) \le \Sigma_{min}(s)$$

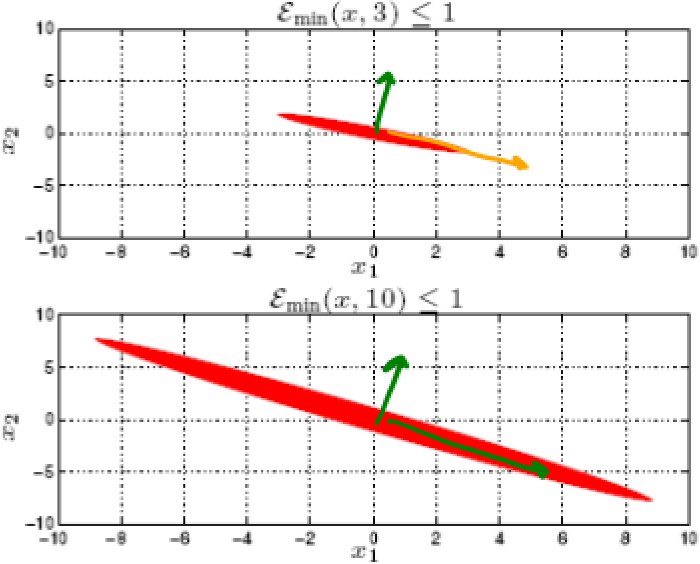
Least-Norm Input Example (1.2)

$$x_{k+1} = \begin{bmatrix} 1.75 & 0.8 \\ -0.95 & 0 \end{bmatrix} x_k + \begin{bmatrix} 1 \\ 0 \end{bmatrix} u_k$$

$$\mathcal{E}_{min}(z, t) \qquad z = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$



Least-Norm Input Example (1.2)



Minimum Energy of Infinite Horizon (1.3)

$$P = \lim_{t\to\infty} \left(\mathcal{C}_t \mathcal{C}^T\right)^{-1} = \lim_{t\to\infty} \left(\sum_{\tau=0}^{t-1} A^\tau B B^T (A^T)^\tau\right)^{-1}$$

min energy to get to x_{des} for any amount of time.

$$\min\left\{ \sum_{\tau=0}^{t-1} \|u(\tau)\|^2 \;\middle|\; x(0)=0 \quad x(t)=x_{des}\right\} = x_{des}^T P x_{des}$$

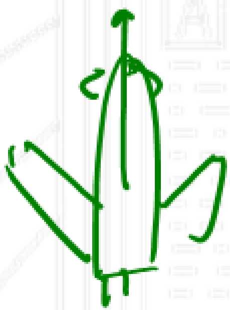
If A is stable $P > 0$ can't get anywhere for free.

Minimum Energy of Infinite Horizon (2.3)

if A is not stable P can have non-zero nullspace.

$Pz = 0$ $z \neq 0$ ⇒ can get to z using u's with very small control energy.

- Basis is highly maneuverable military aircraft



Continuous-time Reachability (1.3)

$\dot{x} = Ax + Bu \qquad x(t) \in \mathbb{R}^n$

$$R_t = \left\{ \int_0^t e^{A(t-\tau)} B u(\tau)\, d\tau \;\middle|\; u:[0,t] \to \mathbb{R}^m \right\}$$

for $t > 0 \quad R_t = R = \text{range}(\mathcal{C})$

$$\mathcal{C} \triangleq [B \;\; AB \;\; A^2B \;\; \dots \;\; A^{n-1}B]$$

Same Reachability as the discrete case

any reachable point can be reached in infinitesimal time given large enough u.

Continuous-time Reachability (2.3)

for any u, $x(0)=0$ $\quad x(t) \in \mathcal{R}_{\text{range}}(\mathcal{C})$

$$e^{At} = I + At + \frac{A^2t^2}{2!} + \dots \quad \text{use C-H: } A^n, A^{n+1} \dots$$

$$e^{At} = \alpha_0(t)I + \alpha_1(t)A + \dots \alpha_{n-1}(t)A^{n-1}$$

$$x(t) = \int_0^t e^{A\tau} B u(t-\tau)\, d\tau$$

$$= \int_0^t \left(\sum_{i=0}^{n-1} \alpha_i(\tau) A^i\right) B u(t-\tau)\, d\tau$$

$\mathcal{C}$

$$\boxed{\sum_{i=0}^{n-1} A^i B} \int_0^t \underbrace{\alpha_i(\tau) u(t-\tau)\, d\tau}_{z}$$

Continuous-time Reachability (3.3)

$= \mathcal{C} z$ where $z_i = \int_0^t \alpha_i(\tau)\, u(t-\tau)\, d\tau$

$x(t)$ is always in Range of $\mathcal{C}$.

~~Impulsive Inputs (1.3)~~

Observability matrix

$$\mathcal{O}_k = \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{k-1} \end{bmatrix}$$

$$\tilde{x}(0) = \mathcal{E}_{ME} = \left\{ \left(\mathcal{O}_k \mathcal{O}_k^T \right)^{-1} \right.$$

~~Impulsive Inputs (2.3)~~

Parting thoughts