# Controllability and Stability

- Define "orth" as shorthand for "orthonormal"
- Controllability $\iff$ $span(B, AB, A^2B, \dots, A^{n-1}B) = \mathbb{R}^n$; can set eigenvals to whatever we want $(A + BF)$ with feedback gains, i.e. any unstable controllable system can be made stable
- Observability $\iff$ $\mathscr{O} = [C, CA, CA^2, \dots, CA^{n-1}]^T$ has full rank
- The general solution to $\vec{x}(t+1) = A\vec{x}(t) + B\vec{u}(t)$ is $\vec{x}(t) = A^t x(0) + \sum_{k=0}^{t-1} A^{t-1-k} B\vec{u}(k)$ for $t \geq 1$.
- Represent CT model with DT model, sampling period $\Delta$. Suppose $A$ (CT) is diagonalizable, i.e. $\exists\ V$ s.t. $V^{-1}AV = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$, then (DT) $A_d = V \begin{bmatrix} e^{\lambda_1 \Delta} & & \\ & \ddots & \\ & & e^{\lambda_n \Delta} \end{bmatrix} V^{-1}$
- Controllability matrix of original matrix: $G = [A, AB]$, so system viewed in basis of Col(A) is $\widetilde{A} = G^{-1}AG$ with $\vec{z} = G\vec{x}, \widetilde{B} = GB$, then for the canonical system $A_z = \widetilde{A}^T = \begin{bmatrix} \vec{0} & I_{n-1} \\ a_0 & a_1 \dots a_{n-1} \end{bmatrix}$ and $B_z = [0, \dots, 0, 1]^T$ (?); controllability matrix is $H = [B_z, A_zB_z]$, with $\widetilde{A} = H^{-1}A_zH$ as well, and transformation $P = \widetilde{G}G^{-1} = HG^{-1}$, giving $\vec{z} = HG^{-1}\vec{x}$ (i.e. $GH^{-1}$-basis)
- Canonical form has the same characteristic polynomial, i.e. $det(A - \lambda I) = det(A_z - \lambda I) = \lambda^n - a_{n-1}\lambda^{n-1} - \dots - a_1\lambda - a_0$ and do pattern matching ($A_z$ always take the above form)
- DT stability (BIBO) $\iff$ all eigenvalues magnitudes $|\lambda| < 1$ (max error @$\lambda = 0$); CT stability $\iff$ eigenvalues have non-positive real parts, i.e. all $Re\{\lambda\} < 0$
- System Identification with $DP = S$, least squares gives: $P = (D^TD)^{-1}D^TS$; choose our inputs randomly will ensure with high probability that $D$ has lin. indep. cols, and the work of calcing $(D^TD)^{-1}$ can be reused.
- Any observable system with an $n$-dimensional internal state can be modelled as $y[i+1] = a_{11}y[i] + \dots + a_{n1}y[i-n+1] + b_1u[i] + \dots + b_n[i-n+1]$, aka observable canonical form.
- The choice of $n$ is important when performing system identification.

## Linearization

- Defn (Linearity): (1) Scaling: $f(\alpha x) = \alpha f(x)$, e.g. must pass thru $(0,0)$; and (2) Superposition: $f(x+y) = f(x) + f(y)$
- Taylor Series approx.: If $f$ is differentiable $k$ times at $x^*$ (for any $k \geq 1$), i.e. $\frac{d^{(k)}}{dx^k}$ should be well defined. Then the Taylor expansion is: $f(x^* + \epsilon) \approx f(x^*) + \epsilon f'(x^*) + \epsilon^2 \frac{f^{(2)}(x^*)}{2!} + \cdots + \epsilon^k \frac{f^{(k)}(x^*)}{k!}$
- Linearize (set $k = 1$): $f(x^* + \epsilon) = f(x^*) + \epsilon f'(x^*)$
- Procedure to linearize $\frac{dx}{dt} = f(x) + u(t)$:
  1. Set a constant $u^*$ and find DC operating pt $(x^*, u^*)$, i.e. where $f(x^*) = -u^*$
  2. Find slope $m = \left.\frac{df(x)}{dx}\right|_{x=x^*}$
  3. Linearize $\frac{dx_l(t)}{dt} = m \cdot x_l(t) + u_l(t)$ where $x_l(t) = x(t) - x^*$ and $u_l(t) = u(t) - u^*$
- Verify the "smallness" assumption (?), i.e. $x_l(t)$ is small enough for $t \geq 0$; N.B. if $f$ is linear to begin with, then problem might arise
- Total derivative formula: $f(x^* + \delta x, y^* + \delta_y) \approx f(x^*, y^*) + \left.\frac{\delta f}{\delta x}\right|_{(x^*,y^*)} \delta x + \left.\frac{\delta f}{\delta y}\right|_{(x^*,y^*)} \delta y$

$$\begin{aligned}
\delta \vec{f} &\approx \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots \frac{\partial f_2}{\partial x_n} \\ \vdots & & \\ \frac{\partial f_n}{\partial x_1} & \frac{\partial f_n}{\partial x_2} & \cdots \frac{\partial f_n}{\partial x_n} \end{bmatrix} \begin{bmatrix} \delta x_1 \\ \delta x_2 \\ \vdots \\ \delta x_n \end{bmatrix} + \begin{bmatrix} \frac{\partial f_1}{\partial u_1} & \frac{\partial f_1}{\partial u_2} & \cdots \frac{\partial f_1}{\partial u_k} \\ \frac{\partial f_2}{\partial u_1} & \frac{\partial f_2}{\partial u_2} & \cdots \frac{\partial f_2}{\partial u_k} \\ \vdots & & \\ \frac{\partial f_n}{\partial u_1} & \frac{\partial f_n}{\partial u_2} & \cdots \frac{\partial f_n}{\partial u_k} \end{bmatrix} \begin{bmatrix} \delta u_1 \\ \delta u_2 \\ \vdots \\ \delta u_k \end{bmatrix} \\
&= J_{\vec{x}} \delta \vec{x} + J_{\vec{u}} \delta \vec{u}.
\end{aligned}$$

Thus, we can approximate our state transition equation as

$$\frac{\mathrm{d}\vec{x}}{\mathrm{d}t} \approx \vec{f}(\vec{x}^*, \vec{u}^*) + \delta \vec{f} = \vec{f}(\vec{x}^*, \vec{u}^*) + J_{\vec{x}} \delta \vec{x} + J_{\vec{u}} \delta \vec{u}.$$

But since we chose our DC operating point to be such that $\vec{f}(\vec{x}^*, \vec{u}^*) = \vec{0}$, we have that

$$\frac{\mathrm{d}}{\mathrm{d}t}(\delta \vec{x}) \approx J_{\vec{x}} \delta \vec{x} + J_{\vec{u}} \delta \vec{u},$$

## Orthonormalization and Gram-Schmidt

- Orthonormal matrix $S : S^T S = I$, i.e. $\vec{s}_i^T \vec{s}_j = 1$ if $i = j$ and 0 if $i \neq j$
- Gram-S: input = lin indep. vectors $\{\vec{s}_1, \ldots, \vec{s}_n\}$, output = $\{\vec{q}_1, \ldots, \vec{q}_n\}$ s.t. $Q$ is orthonormal & for $0 < k \leq n$, $span(\{\vec{s}_1, \ldots, \vec{s}_k\}) = span(\{\vec{q}_1, \ldots, \vec{q}_k\})$
- G-S Procedure
  1. $\vec{q}_1 = \frac{\vec{s}_1}{||\vec{s}_1||}$ with $||\vec{q}_1||^2 = 1$
  2. For $\vec{q}_2$, choose $\vec{z}_2 = \vec{s}_2 - \alpha\vec{q}_1 = \vec{s}_2 - proj_{\vec{q}_1}\vec{s}_2 = \vec{s}_2 - \frac{\vec{q}_1^T \vec{s}_2}{||\vec{q}_1||^2}\vec{q}_1 = \vec{s}_2 - (\vec{q}_1^T \vec{s}_2)\vec{q}_1$. Finally, set $\vec{q}_2 = \frac{\vec{z}_2}{||\vec{z}_2||}$
  3. i.e. Generalizes to: $\vec{z}_i = \vec{s}_i - \sum_{j=1}^{i-1}(\vec{s}_i^T \vec{q}_j)\vec{q}_j$; then normalizes: $\vec{q}_i = \frac{\vec{z}_i}{||\vec{z}_i||}$

  In other words: $S[i]^T(\vec{u}[i] - proj_{S[i]}\vec{u}[i]) = \vec{0}$, which implies $\vec{u}[i] - proj_{S[i]}\vec{u}[i]$ is orthogonal to Col($S$), so can be appended (after simple normalization) − producing an orthonormal basis from a set of vectors spanning a given vector space
- Let $U$ whose columns form an orth basis for Col($V$), then $proj_{Col(V)}\vec{w} = V(V^T V)^{-1} V^T \vec{w} = U(U^T U)^{-1} U^T \vec{w} = proj_{Col(U)}\vec{w}$
- $||A\vec{x}|| = \sqrt{\vec{x}^T A^T T \vec{x}}$ for general matrix $A$; important usage on symmetric matrix − for eigenvecs $U = [\vec{u}_1, \ldots, \vec{u}_n]$ of symmetric matrix $T$, then they can be orthonormal, and general vector

  $\vec{x} = \sum_{i=1}^{n}(\vec{u}_i^T \vec{x})\vec{u}_i$ with $||\vec{x}|| = \sqrt{\sum_{i=1}^{n}(\vec{u}_i^T \vec{x})^2}$, and so

  $\vec{y} = T\vec{x} = \sum_{i=1}^{n}(\vec{u}_i^T \vec{x})\lambda_i \vec{u}_i$ for corresponding eigenval $\lambda's$ with $||\vec{y}|| = \sqrt{\sum_{i=1}^{n}(\vec{u}_i^T \vec{x}\lambda_i)^2}$
- For orthonormal matrices $U, V$ then $||UA||_F = ||AV||_F = ||A||_F = \sqrt{\sum_{i=1}^{N}\sigma_i^2}$ where $\sigma's$ are singular values of $A$ (using SVD decomp)
- For symmetric $A$ and $V$ with orth cols, then $V^T AV$ is symmetric

## SVD/PCA

- Schur decomposition: $A = QTQ^{-1}$ where $Q^TQ = I$ and $T$ is upper-triangular; if A is symmetric, then $T$ is diagonal (Spectral Theorem)
- A $\mathbb{R}$ symmetric matrix will always have $\mathbb{R}$ eigenvals
- SVD Procedure (tall matrix $A \in \mathbb{R}^{m\times n}, m > n$; pick $A^T$ if $m < n$)
  1. Compute the symmetric matrix $A^TA \in \mathbb{R}^{n\times n}$ ($AA^T$ if $m < n$)
  2. Find eigenvals & normalized eigenvecs $(\lambda_1, \vec{v}_1), \dots, (\lambda_n, \vec{v}_n)$ of $A^TA$. By the spectral theorem for real symmetric matrices, these eigenvectors are orth. ($\vec{u}_i's$ if $m < n$)
  3. $\sigma_i = \sqrt{\lambda_i}$ where $\lambda_i's$ are sorted in desc order, and they're non-negative.
  4. Compute $U$ where $\vec{u}_i = \frac{A\vec{v}_i}{\sigma_i}$ s.t. they are normalized since $\sigma_i = ||A\vec{v}_i||$ and orth since $(A\vec{v}_i)^T(A\vec{V}_j) = 0$ if $i \neq j$ ($\vec{v}_i = \frac{A^T\vec{u}_i}{\sigma_i}$ if $m < n$)
  5. N.B. The eigenvec(s) corresponding to $\lambda = 0$ for $A^TA/AA^T$ (?) gives Null($A$); The columns of $U$ gives range($A$)
- SVD applies to arbitrary rectangle matrix: for $A \in \mathbb{R}^{m\times n}$, then $A = U\Sigma V^T$ where $U \in \mathbb{R}^{m\times m}, V \in \mathbb{R}^{n\times n}$ are orthogonal, and $\Sigma \in \mathbb{R}^{m\times n} = \begin{bmatrix} \sigma_1 & 0 & 0 & ... \\ 0 & \sigma_2 & 0 & ... \\ & \dots & & \\ 0 & \dots & 0 & \sigma_n \\ 0 & \dots & \dots & 0 \\ & \dots & & \\ 0 & \dots & \dots & 0 \end{bmatrix}$ is "diagonal"
- SVD: For $A \in \mathbb{R}^{m\times n}$ of rank $r$, then can be decomposed to sum of $r$ rank-1's: $A = \sum_{i=1}^{r} \sigma_i \vec{u}_i \vec{v}_i^T$
- WLOG, let $m \geq n$, and (SVD) $A = U\Sigma V^T = \begin{bmatrix} \vec{u}_1 & \dots & \vec{u}_m \end{bmatrix} \cdot \Sigma \cdot \begin{bmatrix} \vec{v}_1^T \\ \dots \\ \vec{v}_n^T \end{bmatrix} = \sum_{k=1}^{k=n} \sigma_k \vec{u}_k \vec{v}_k^T$, i.e. any $m \times n$ matrix $A$ ($m \geq n$) can be expressed as the sum of $n$ rank-1 matrices (sim for $m \leq n$).
- Minimum Energy Control: only have nonzero entries in directions orthogonal to Null($\mathscr{C}_t$), i.e. only first $n$ entries, $v_1, \dots, v_n \neq 0$ gives min $||\vec{u}||$
- Existence of such $v_i's$ is shown by constructing an orthonormal basis Consider $Q = \mathscr{C}_t^T\mathscr{C}_t$, and eigenvals: $\lambda_i = ||\mathscr{C}_t\vec{v}_i||^2$, so can order: $\lambda_1 \geq \dots \geq \lambda_n > \lambda_{n+1} = \dots = \lambda_t = 0$

  So, we can write $Q = V\Lambda V^{-1}$ where $\Lambda = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ & & \ddots & \\ 0 & \dots & 0 & \lambda_t \end{bmatrix}$ and thus, let $\sigma_i = \sqrt{\lambda_i} = ||\mathscr{C}_t\vec{v}_i||$, the singular values.
- Create another orth basis since $\mathscr{C}_t\vec{v}_i$ are mutually orth: $\vec{w}_i = \frac{\mathscr{C}_t\vec{v}_i}{\sigma_i}$ (Check: ?) to get $\mathscr{C}_t = W\Sigma V^T$ (SVD!)

  $\mathscr{C}_t\vec{v}_i = \sigma_i\vec{w}_i$ for $i \leq n$; and $\mathscr{C}_t\vec{v}_i = \vec{0}$ for $i > n$
- Min Energy Control: for $\mathscr{C}_t\vec{u} = \vec{x}^*$, obtain control: $\vec{u} = \sum_{i=1}^{n} \frac{(\vec{x}^* \cdot \vec{w}_i)\vec{w}_i}{\sigma_i}\vec{v}_i$

- Frobenius norm: $||A||_F = \sqrt{\sum\limits_{i=1}^{m} \sum\limits_{j=1}^{n} a_{ij}^2}$

- $\widetilde{A} = \begin{bmatrix} A_{11} - \mu_1 & A_{12} - \mu_2 & \dots & A_{1m} - \mu_m \\ \dots & & & \\ A_{n1} - \mu_1 & A_{n2} - \mu_2 & \dots & A_{nm} - \mu_m \end{bmatrix}$, and the covariance matrix $S = \frac{1}{n} \widetilde{A}^T \widetilde{A}$ is symmetric, and (PCA) can be written as $S = P \Lambda P^T$ where $P$ is orth (its columns are principal components) and $\Lambda \geq 0$.

  Define $S_i = \sqrt{S_{ii}}$, so correlation matrix: $R_{ij} = \frac{S_{ij}}{S_i S_j} = \cos\theta$

- SVD to PCA

- PCA to SVD

## Extra Sanity Checks

- $x(0)e^{\lambda t} + \frac{u}{\lambda}(e^{\lambda t} - 1)$ solves $\frac{d}{dt}x(t) = \lambda x(t) + u$
- Proving some $c$ is real $\iff c = \overline{c}$
- For $\vec{x} = [a + bj, c + dj]$, then $||\vec{x}||^2 = a^2 + b^2 + c^2 + d^2$
- $||\vec{x}||^2 = \overline{\vec{x}}^T \vec{x} = 0 \iff \vec{x} = \vec{0}$
- $\vec{u}, \vec{v}$ are orth. $\iff \vec{u}^T \vec{v} = 0$ and $||\vec{u}|| = 1$
- Finding Null($A$) $\implies$ span of eigenvec(s) for eigenval $= 0$ (for $A$ or $A^T A / AA^T$ (?))
- For SVD, the diagonal entries of $\Sigma$ MUST be non-negative.
- (Disc) For symmetric matrix $A$, any two eigenvectors corresponding to distinct eigenvals of $A$ are orthogonal.
- Correlation matrix: if the data is exactly on a horizontal line, then the correlation becomes undefined
- Ex) Suppose that the matrix $A \in \mathbb{R}^{N \times M}$ has lin indep columns. The vector in $\mathbb{R}^N$ is not in the subspace spanned by the columns of $A$. Then $proj_{Col(A)}\vec{y} = A\vec{\hat{y}} = A(A^T A)^{-1} A^T \vec{y}$ where $\vec{\hat{y}}$ minimizes $||\vec{y} - A\vec{\hat{y}}||$
- For $\vec{x}_u$ in basis $U$ and its corresponding $\vec{x}_v$ in basis $V$, then $\vec{x} = U\vec{x}_u = V\vec{x}_v$

  For orth basis $U$, decompose generic $\vec{x} = U^{-1}\vec{x}U = U^T\vec{x}U$
- For invertible $T$, $det(A - \lambda I) = det(TT^{-1}(A - \lambda I)) = det(T(A - \lambda I)T^{-1})$