EECS 16B   Lecture 18: (Module 2 Lecture 6)

Today: - wrap up controllability
- orthonormal bases & Gram-Schmidt procedure

Reminder: MT Re-do's due tomorrow.

Recap: Controllability:

Recall soln. of discrete-time system:

$\vec{x}[i+1] = A\vec{x}[i] + Bu[i]$ (assume single input, i.e. $m=1$)

$$\circledast \quad \boxed{\vec{x}[\ell] - A^{\ell}\vec{x}[0] = \underbrace{\left[A^{\ell-1}B \mid A^{\ell-2}B \;\; .. \;\; |AB|B\right]}_{\triangleq\, C_\ell \text{ ``Controllability'' matrix}} \begin{bmatrix} u[0] \\ u[1] \\ \vdots \\ u[\ell-1] \end{bmatrix}}$$

n: dim. of $\vec{x}$ m: dim of $\vec{u}[i]$ ($m=1$ for our case)

A: $n \times n$ B: $n \times 1$

ℓ: time-index $(0, 1, 2, \ldots \ell-1, \ell)$

C_ℓ: $n \times \ell m$ ($= n \times \ell$ for our case of $m=1$)

Q) Can we find input sequence $u[0], u[1] \ldots u[\ell-1]$ that brings state $\vec{x}[0]$ @ time 0 to a target $\vec{x}_{target}$ at time ℓ?

$$\underbrace{\vec{x}[\ell]}_{\text{target state}} - A^{\ell}\underbrace{\vec{x}[0]}_{\text{init. state}} = \underbrace{\left[A^{\ell-1}B \mid A^{\ell-2}B \;\ldots\; |AB|B\right]}_{\triangleq\, C_\ell} \begin{bmatrix} u[0] \\ u[1] \\ \vdots \\ u[\ell-1] \end{bmatrix}$$

Yes, if $\vec{x}_{target} - A^{\ell}\vec{x}[0]$ lies in the column span of

$$C_\ell \triangleq \left[A^{\ell-1}B \;\ldots\; |AB|\,B\right].$$

> **Defn:** A system is called controllable if, given any target state $\vec{x}_{target}$ and initial condition $\vec{x}[0]$, we can find a time ℓ and an input sequence $\{u[0], u[1], \ldots u[\ell-1]\}$ such that $\vec{x}[\ell] = \vec{x}_{target}$.

Informally, controllable in an n-dim. state control system $\equiv$ ability to go from anywhere in $\mathbb{R}^n$ to anywhere in $\mathbb{R}^n$ by exerting the right control action

Q) How do we check this? If C_ℓ has n linearly indep. columns for some ℓ, then column space is $\mathbb{R}^n$ (where n is the state dimension) This means we can make

" $C_\ell \begin{bmatrix} u[0] \\ \vdots \\ u[\ell-1] \end{bmatrix}$ " anything we want in $\mathbb{R}^n$ by choosing the input control sequence $\{u[0], u[1], \ldots u[\ell-1]\}$.

EXAMPLE-2

(n=2)

$$\vec{x}(i+1) = \underbrace{\begin{bmatrix} 0 & 1 \\ 3 & 2 \end{bmatrix}}_{A} \vec{x}(i) + \underbrace{\begin{bmatrix} 0 \\ 1 \end{bmatrix}}_{B} u(i)$$

$$B = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \; ; \; AB = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

$C_1 = B$ (dim. = 1)

$C_2 = [AB \;\; B] = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}$ (dim = 2 = n) ✓

CONTROLLABLE

EXAMPLE-3

$$\vec{x}[i+1] = \underbrace{\begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix}}_{A} \vec{x}[i] + \underbrace{\begin{bmatrix} 1 \\ 0 \end{bmatrix}}_{B} u[i]$$

$C_1 = B$ (dim. = 1)

$C_2 = [AB \;\; B] = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$ (dim. = 1)

$C_3 = [A^2B \;\; AB \;\; B] = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$ (dim. = 1)

Dim. can't increase further: stuck at $1 < n = 2$

UNCONTROLLABLE!

$$x_1[i+1] = x_1[i] + x_2[i] + u[i]$$
$$x_2[i+1] = 2\,x_2[i]$$

Say we want to go from $\vec{x}[0] = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ to $\vec{x}_{target} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ in ℓ steps, can we do it by appropriate design of $\vec{u}[i]$?

$$\vec{x}_{target} = \begin{bmatrix} AB \mid B \end{bmatrix} \underbrace{\begin{bmatrix} u[0] \\ u[1] \end{bmatrix}}_{\text{(design choice)}}$$

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} * \\ * \end{bmatrix} ?$$

$\times$ not possible

What we observe about dim. getting stuck is due to the following result:

Theorem:

~~Lemma~~: If $A^{\ell}B$ is linearly dependent on $\{A^{\ell-1}B, \dots AB, B\}$, then $A^{\ell+1}B$ is also linearly dependent on $\{A^{\ell-1}B, \dots AB, B\}$.

Theorem

~~Lemma~~ implies:

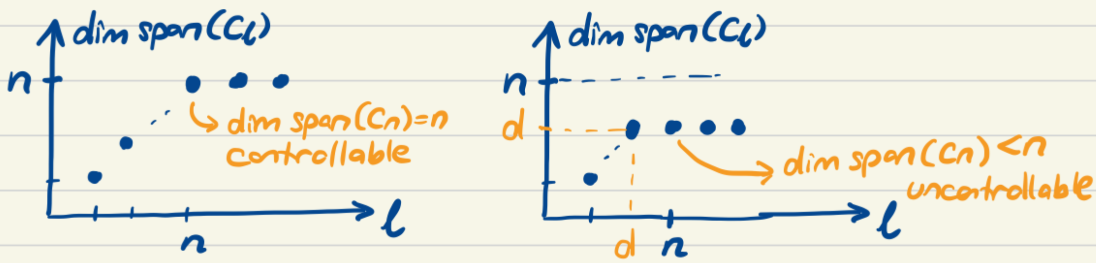
if $\dim \operatorname{span}(C_{\ell+1}) = \dim \operatorname{span}(C_{\ell}) = d$,

then $\dim \operatorname{span}(C_{\ell+2}) = \dim \operatorname{span}(C_{\ell+1}) = d$.

Once the dimension stops growing, it stops growing for good!

We know that $\dim \operatorname{span}(C_{\ell+1})$ cannot exceed n, as $C_{\ell+1}$ has n rows, and

$\dim(C_{\ell+1}) = \max[\dim(\text{row space of } C_{\ell+1}), \dim(\text{col. space of } C_{\ell+1})]$



Controllability condition: $\dim \operatorname{span}(C_n) = n$

i.e., $C_n = [A^{n-1}B \; \cdots \; AB \; B]$ has n linearly indep. columns.

Example 1:

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix} \quad B = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

Example 2:

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix} \quad B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$n=2$: AB, B linearly independent ?

Example 1 — No:

$$AB = \begin{bmatrix} 1 \\ 0 \end{bmatrix} = B$$

Uncontrollable

$$x_1[i+1] = x_1[i] + x_2[i] + u[i]$$
$$x_2[i+1] = 2x_2[i]$$

$\rightarrow x_2[l] = 2^l x_2[0]$

1. evalue of 2 remains regardless of feedback
2. we can't take x_2 component anywhere we want

Example 2 — Yes:

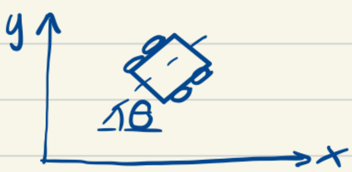
$$AB = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \text{ indep. of } B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Controllable

$$x_1[i+1] = x_1[i] + x_2[i]$$
$$x_2[i+1] = 2x_2[i] + u[i]$$

u doesn't appear in x_1 eq'n but can influence x_1 indirectly through x_2

<u>Note</u> from the above that we can control n variables with fewer than n inputs – you actually do this when you drive a car: using two inputs (steering and longitudinal forces generated by gas/brake pedals) you come not only to the desired x, y coordinates but also to the desired orientation e.g. parallel parking.



$\begin{bmatrix} x \\ y \\ \theta \end{bmatrix}$: state vector can be brought to target value.

(CCF)

Controller Canonical Form (single input, $m=1$)

A special structure of A and B in which we can arbitrarily assign evalues of

$$A_{CL} = A + BF$$

with the choice of F:

$$\underset{n}{A} = \begin{bmatrix} 0 & \boxed{\begin{matrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & 0 \cdots & 0 \\ & & \ddots & 0 \\ \cdots & & 0 & 1 \end{matrix}}^{\,I_{n-1}} \\ 0 & \\ \vdots & \\ 0 & \\ a_1 & a_2 \quad \cdots \quad a_n \end{bmatrix} \qquad B = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}$$

Example 2 had this form: $n=2$, $a_1=3$, $a_2=2$

Nice properties of this form:

$$\left(\text{Ex. 2} \quad \underset{2}{A} = \begin{bmatrix} 0 & 1 \\ 3 & 2 \end{bmatrix} \right)$$

1) Characteristic polynomial of A is transparent:

$$\det(\lambda I - A) = \lambda^n - a_n\lambda^{n-1} - a_{n-1}\lambda^{n-2} \cdots - a_2\lambda - a_1$$

Example 2: $\lambda^2 - 2\lambda - 3$

$n=2$: $A = \begin{bmatrix} 0 & 1 \\ a_1 & a_2 \end{bmatrix}$; $\lambda I - A = \begin{bmatrix} \lambda & -1 \\ -a_1 & \lambda - a_2 \end{bmatrix}$

$\det(\lambda I - A) = \lambda^2 - a_2\lambda - a_1$

For $n=3$: $\lambda^3 - a_3\lambda^2 - a_2\lambda - a_1$ is the characteristic polynomial of

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ a_1 & a_2 & a_3 \end{bmatrix}. \qquad \text{(Check.)}$$

2) $A+BF$ has the same structure as A:

$$\begin{bmatrix}0&1&0&\cdots\\ &\ddots&\ddots&\\ &&&1\\ a_1&\cdots&\cdots&a_n\end{bmatrix}+\underbrace{\begin{bmatrix}0\\ \vdots\\ 0\\ 1\end{bmatrix}[f_1 \cdots f_n]}_{\begin{bmatrix}0&\cdots&0\\ \vdots&&\\ 0&\cdots&0\\ f_1&\cdots&f_n\end{bmatrix}}$$

$$=\begin{bmatrix}0&1&0&\cdots\\ &\ddots&\ddots&\\ &&&1\\ a_1+f_1&\cdots&\cdots&a_n+f_n\end{bmatrix}$$

Same structure as A, but $a_k \to a_k+f_k$, $k=1,\dots n$.

From Properties 1 and 2,

$$\det(\lambda I-A_{CL})=\lambda^n-(a_n+f_n)\lambda^{n-1}\cdots-(a_2+f_2)\lambda-(a_1+f_1).$$

Suppose we want $A_{CL}=A+BF$ to have evalues $\lambda_1,\dots\lambda_n$. Then, the characteristic polynomial of A_{CL} should be:

$$(\lambda-\lambda_1)(\lambda-\lambda_2)\cdots(\lambda-\lambda_n)$$

$$=\lambda^n-(\lambda_1+\dots\lambda_n)\lambda^{n-1}\cdots+(-1)^n\lambda_1\cdots\lambda_n$$

$$a_1+f_1=-(-1)^n\lambda_1\cdots\lambda_n=(-1)^{n+1}\lambda_1\cdots\lambda_n \Rightarrow f_1=(-1)^{n-1}\lambda_1\cdots\lambda_n-a_1$$

$$\vdots$$

$$a_n+f_n=\lambda_1+\dots+\lambda_n \Rightarrow f_n=\lambda_1+\dots\lambda_n-a_n$$

Quiz:

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 2 & 3 \end{bmatrix} \qquad B = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

$A + BF$

If we want to place eigenvalues of A_{CL} at $\lambda_1 = \lambda_2 = \lambda_3 = 0$,

(a) What is the char. poly. of A_{CL}?

(b) What are the feedback values f_1, f_2, f_3?

$$A_{CL} = A + BF = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 2 & 3 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \begin{bmatrix} f_1 & f_2 & f_3 \end{bmatrix}$$

$$= \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1+f_1 & 2+f_2 & 3+f_3 \end{bmatrix}$$

char. poly. of A_{CL}

$$= \det(\lambda I - A_{CL}) = \lambda^3 - \underbrace{(3+f_3)}_{=0}\lambda^2 - \underbrace{(2+f_2)}_{=0}\lambda - \underbrace{(1+f_1)}_{=0}$$

Want $\lambda_1 = \lambda_2 = \lambda_3 = 0$

$\det(\lambda I - A_{CL}) = \lambda^3 \Rightarrow$

$f_3 = -3$; $f_2 = -2$; $f_1 = -1$

FACT: Any controllable system

$$\vec{x}[n+1] = A\,\vec{x}[n] + B\,u[n];$$

(i.e. one having controllability matrix

$$C_n = \left[\, A^{n-1}B \;\middle|\; \cdots \;\middle|\; A^2B \;\middle|\; AB \;\middle|\; B \,\right]$$

that is rank-n & therefore invertible)

can be "transformed" to a canonical form.

$$\vec{y}[n+1] = A_y\,\vec{x}[n] + B_y\,u[n]$$

where

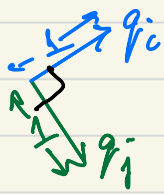
$$A_y = \begin{bmatrix} 0 & \vdots & I_{n-1} & & \\ \vdots & \vdots & & & \\ -\alpha_1 & -\alpha_2 & \cdots & \cdots & -\alpha_n \end{bmatrix} \quad \& \quad B_y = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$$

where $\vec{y}[n] = T\,\vec{x}[n]$ for a carefully chosen invertible transformation matrix T.

(HW problem on how to find T)

Orthonormal bases and Gram-Schmidt Procedure:

Column vectors $\vec{q}_1, \dots \vec{q}_k$ are called orthonormal if



$$\vec{q}_i^{\,T} \vec{q}_j = \begin{cases} 0 & \text{if } i \neq j \quad (\text{ortho}) \\ 1 & \text{if } i = j \quad (\text{normal}) \end{cases} \quad \text{--}(1)$$

A matrix $Q = [\vec{q}_1 \dots \vec{q}_k]$ with orthonormal columns

satisfies:

$$Q^TQ = \begin{bmatrix} \vec{q}_1^{\,T} \\ \vdots \\ \vec{q}_k^{\,T} \end{bmatrix} [\vec{q}_1 \dots \vec{q}_k] = \begin{bmatrix} \vec{q}_1^{\,T}\vec{q}_1 & \dots & \vec{q}_1^{\,T}\vec{q}_k \\ \vdots & & \\ \vec{q}_k^{\,T}\vec{q}_1 & \dots & \vec{q}_k^{\,T}\vec{q}_k \end{bmatrix}$$

$\vec{q}_1 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix}$

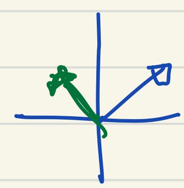
$\vec{q}_2 = \frac{1}{2}\begin{bmatrix} 1 \\ -1 \end{bmatrix}$

$Q = \begin{bmatrix} \vec{q}_1 & \vec{q}_2 \end{bmatrix} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$; $Q^T = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$

$= I_{k \times k}$ by def'n (1).

$$\boxed{Q^TQ = I}$$

If Q is square $Q^TQ = I$, means:



$$Q^T = Q^{-1}.$$

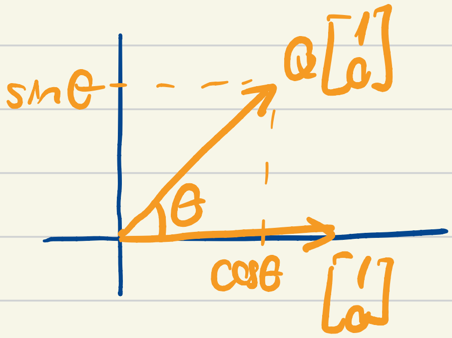
(Q is called orthogonal.)

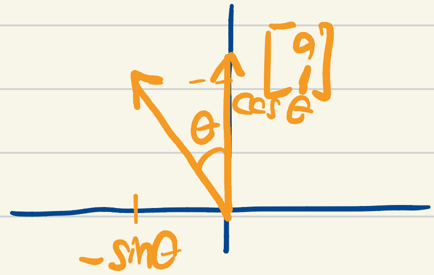
Example: $Q = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$ (rotation matrix)

$$Q = \big[\, \underbrace{\begin{matrix}\cos\theta \\ \sin\theta\end{matrix}}_{\vec{q}_1} \quad \underbrace{\begin{matrix}-\sin\theta \\ \cos\theta\end{matrix}}_{\vec{q}_2} \,\big]$$

$\vec{q}_1^{\,T}\vec{q}_2 = -\cos\theta\sin\theta + \sin\theta\cos\theta = 0$

$\vec{q}_i^{\,T}\vec{q}_i = \cos^2\theta + \sin^2\theta = 1$





so called because $Q\vec{x}$ rotates $\vec{x}$ by angle θ in the plane without changing its length

Useful features of matrices with orthonormal columns:

1) $\|Q\vec{x}\| = \|\vec{x}\|$ (preserves length: what we observed in example above is true for other Q with orthonormal columns)

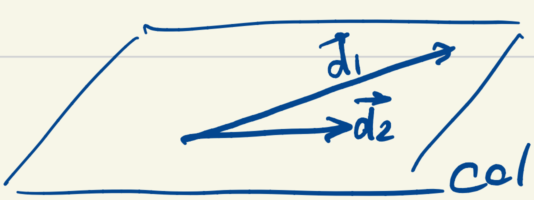
$$\underbrace{\|Q\vec{x}\|}_{\sqrt{(Q\vec{x})^T(Q\vec{x})}} = \underbrace{\|\vec{x}\|}_{\sqrt{\vec{x}^T\vec{x}}}$$

$$\sqrt{(Q\vec{x})^T(Q\vec{x})} = \sqrt{\vec{x}^T \underbrace{Q^TQ}_{=I} \vec{x}} = \sqrt{\vec{x}^T\vec{x}}$$

2) Q also preserves dot product: $(Q\vec{x})^T(Q\vec{y}) = \vec{x}^T \underbrace{Q^TQ}_{=I} \vec{y} = \vec{x}^T\vec{y}$

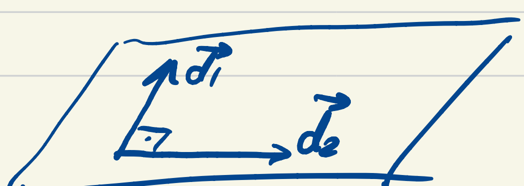
3) Easy visualization of column space:

column space of $D = [\vec{d}_1\ \vec{d}_2]$



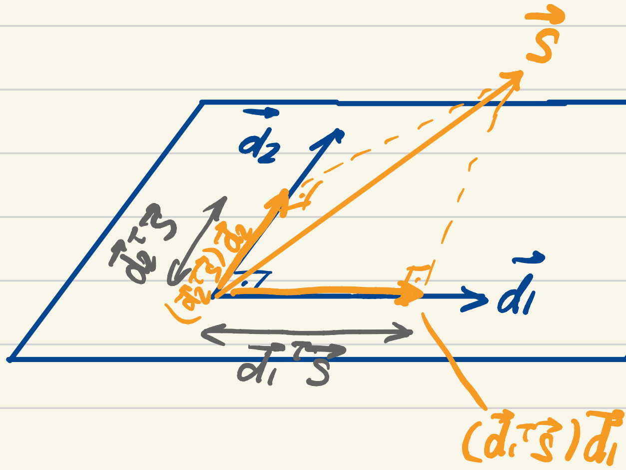
col. space of D

if $\vec{d}_1, \vec{d}_2$ were orthonormal



prettier!

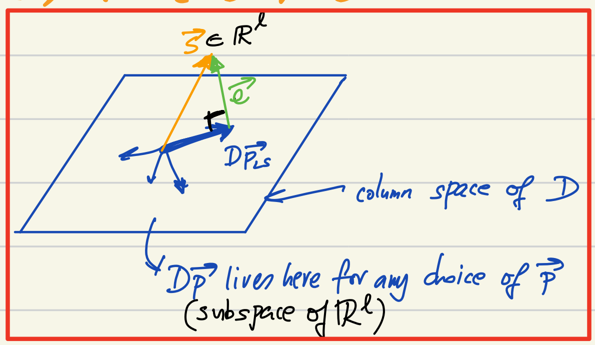
and projection onto column space is trivial:



Projection of $\vec{s}$ onto column space of D:

$$= (\vec{d}_1^T\vec{s})\vec{d}_1 + (\vec{d}_2^T\vec{s})\vec{d}_2$$

Recall Least-Squares:



$$\vec{s} = D\vec{p} + \vec{e}$$

$$\hat{p} = (D^TD)^{-1}D^T\vec{s}$$

What if (by some miracle) D had orthonormal columns?

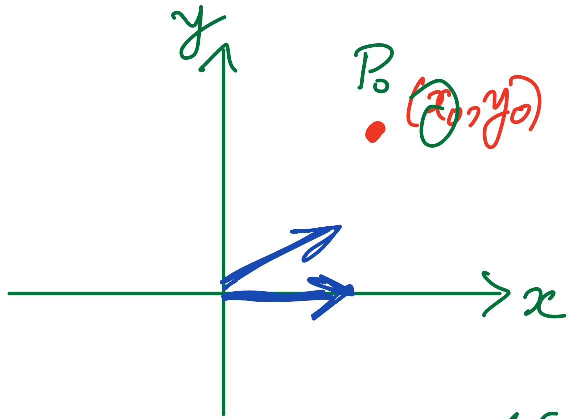
$$D^TD = I \Rightarrow \hat{p} = D^T\vec{s} \quad \text{(no matrix inversion!)}$$

$$\left(D^T = \begin{bmatrix} \leftarrow q_1^T \rightarrow \\ \vdots \\ \leftarrow q_n^T \rightarrow \end{bmatrix};\ D = \begin{bmatrix} | & | & & | \\ \vec{q}_1 & \vec{q}_2 & \cdots & \vec{q}_n \\ | & | & & | \end{bmatrix};\ \boxed{D^TD = I_n} \right)$$

$$\langle q_i, q_j \rangle = \begin{cases} 1 & i=j \\ 0 & i \neq j \end{cases}$$

Gram-Schmidt:

Even if columns of D are not orthonormal, we can construct an orthonormal basis for the column space

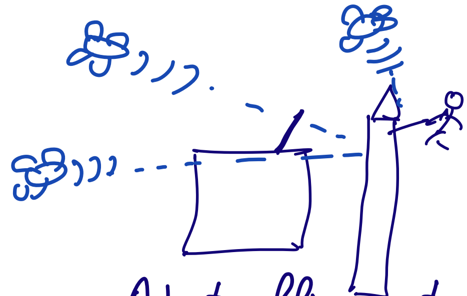


$$\mathrm{Proj}_{x\text{-axis}}\, P_0 = \langle (x_0, y_0), (1, 0) \rangle$$
$$= x_0$$

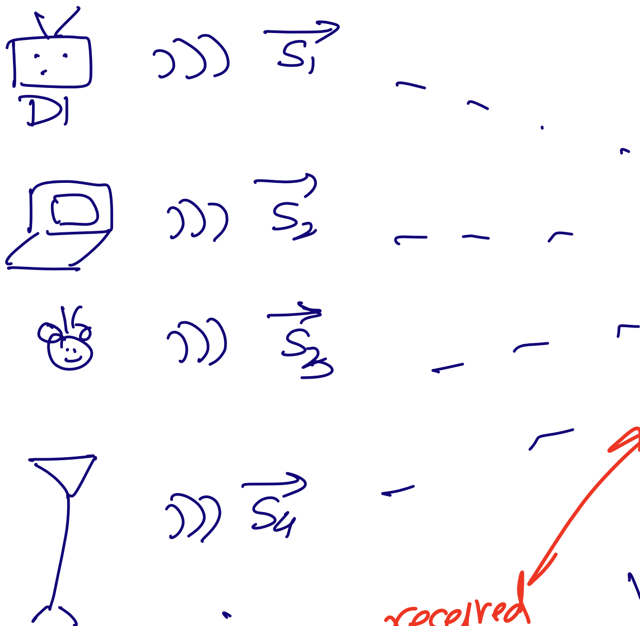
Example

"Internet of Things"

"drone" network



Air traffic controller for drones



$\vec{r}$

"signatures" of drones

$$\vec{r} = \alpha_1 \vec{s}_1 + \alpha_2 \vec{s}_2 + \cdots$$

received signal

I want to know: Who is transmitting?

$\vec{s}_1, \vec{s}_2, \vec{s}_3, \vec{s}_4$

$$\underbrace{\begin{bmatrix} \vec{s}_1 & \vec{s}_2 & \vec{s}_3 & \vec{s}_4 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \end{bmatrix}}_{\vec{x}} \approx \underbrace{\begin{bmatrix} \vec{r} \end{bmatrix}}_{\vec{b}}$$

Goal:

Estimate $\{\alpha_1, \alpha_2, \alpha_3, \alpha_4\}$ using Least squares!

$$\vec{\hat{x}}_{LS} = (A^T A)^{-1} A^T \vec{b}$$

What if we learn about each device/drone one-by-one ("on-line" fashion)

Initially: $\vec{s}_1, \vec{r}$

Q) How well does $\vec{s}_1$ explain received $\vec{r}$?

$$LS: \quad \underbrace{\begin{bmatrix} | \\ \vec{s}_1 \\ | \end{bmatrix}}_{A} \underbrace{[\alpha_1]}_{x} = \underbrace{\begin{bmatrix} | \\ \vec{r} \\ | \end{bmatrix}}_{\vec{b}}$$

$$\boxed{\alpha_1 = (\vec{s}_1^{\,T} \vec{s}_1)^{-1} \vec{s}^{\,T} \vec{r}}$$

LS est. of α_1

Expanded set: $\{\vec{s}_1, \vec{s}_2, \vec{r}\}$

$$LS: \quad \underbrace{\begin{bmatrix} | & | \\ \vec{s}_1 & \vec{s}_2 \\ | & | \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix}}_{\vec{x}} = \underbrace{\begin{bmatrix} | \\ \vec{r} \\ | \end{bmatrix}}_{\vec{b}}$$

New set: $\{\vec{s}_1, \vec{s}_2, \vec{s}_3, \vec{r}\}$

$$LS: \quad \begin{bmatrix} | & | & | \\ \vec{s}_1 & \vec{s}_2 & \vec{s}_3 \\ | & | & | \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} = \begin{bmatrix} | \\ \vec{r} \\ | \end{bmatrix}$$

Observe: A keeps getting "fatter"

$$\underbrace{\begin{bmatrix}\vec{s}_1\end{bmatrix}}_{A_1} \Rightarrow \underbrace{\begin{bmatrix}\vec{s}_1 & \vec{s}_2\end{bmatrix}}_{A_2} \Rightarrow \underbrace{\begin{bmatrix}\vec{s}_1 & \vec{s}_2 & \vec{s}_3\end{bmatrix}}_{A_3} \Rightarrow \cdots$$

$A^T A \longrightarrow$ more & more work "from scratch"

$(A^T A)^{-1} \longrightarrow$ inverses of growing matrices

. OUCH!

Recall properties of orthonormality:

Least-squares is easy if columns of matrix A are orthonormal.

$$\underbrace{\begin{bmatrix} | & | & & | \\ \vec{q}_1 & \vec{q}_2 & \cdots & \vec{q}_n \\ | & | & & | \end{bmatrix}}_{A}$$

$\|\vec{q}_i\| = 1 = \langle \vec{q}_i, \vec{q}_i \rangle$ Normalized

$\langle \vec{q}_i, \vec{q}_j \rangle = 0 \quad \forall\, i, j,\ i \neq j$

Orthogonal.

$$A\vec{x} = \vec{b}$$

$$A = \begin{bmatrix} | & | \\ \vec{q}_1 & \vec{q}_2 \\ | & | \end{bmatrix}$$

$\vec{q}_1, \vec{q}_2$ Orthonormal.

LS

$$\hat{\vec{x}} = (A^T A)^{-1} A^T \vec{b}$$

$$= \Bigg(\underbrace{\begin{bmatrix} - \vec{q}_1^T - \\ - \vec{q}_2^T - \end{bmatrix} \begin{bmatrix} | & | \\ \vec{q}_1 & \vec{q}_2 \\ | & | \end{bmatrix}}_{\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}}\Bigg)^{-1} \begin{bmatrix} - \vec{q}_1^T - \\ - \vec{q}_2^T - \end{bmatrix} \vec{b}$$

$$\hat{\vec{x}}_{LS} = \begin{bmatrix} - \vec{q}_1^T - \\ - \vec{q}_2^T - \end{bmatrix} \vec{b} = \begin{bmatrix} \langle \vec{q}_1, \vec{b} \rangle \\ \langle \vec{q}_2, b \rangle \end{bmatrix}$$

$$A = \begin{bmatrix} \vec{q}_1 & \vec{q}_2 & \vec{q}_3 \end{bmatrix} \qquad A\vec{x} = \vec{b}$$

Orthonormal.

$$A^T A = I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\hat{\vec{x}} = \underbrace{(A^TA)^{-1}}_{I} A^T \vec{b}$$

LS

$$= \begin{bmatrix} - & \vec{q}_1^{\,T} & - \\ - & \vec{q}_2^{\,T} & - \\ - & \vec{q}_3^{\,T} & - \end{bmatrix} \vec{b}$$

$$= \begin{bmatrix} \langle \vec{q}_1, \vec{b} \rangle \\ \langle \vec{q}_2, \vec{b} \rangle \\ \langle \vec{q}_3, \vec{b} \rangle \end{bmatrix}$$

$$\boxed{\begin{aligned} &\langle \vec{x}, \vec{y} \rangle \\ &= \vec{x}^T \vec{y} \\ &= \vec{y}^T \vec{x} \end{aligned}}$$

We can reuse our previous work!!!

Goal: To find $\vec{x}$ as a linear combination of $\vec{S_1}, \vec{S_2}, \vec{S_3} \dots$, where we only find out the $\vec{S_i}$ vectors one at a time.

$\vec{S_1}, \vec{S_2}, \vec{S_3} \longrightarrow$ Convert them to $\vec{q_1}, \vec{q_2}, \vec{q_3}$

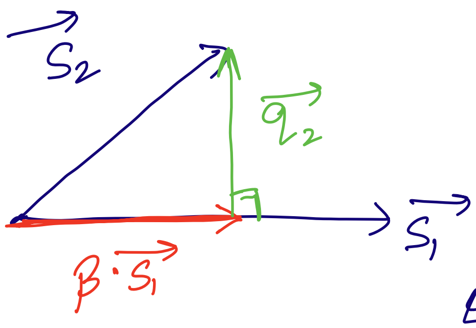
such that $\vec{q_1}, \vec{q_2}, \vec{q_3}$ etc. are all orthonormal.

Focus now: How to orthogonalize vectors?

$\{\vec{s_1}, \vec{s_2}\}$.

$\vec{s_2}$ in terms of $\vec{s_1}$

$$\vec{s_2} = \underbrace{\beta \cdot \vec{s_1}}_{\text{aligned with } \vec{s_1}} + \underbrace{\vec{q_2}}_{\text{"new" dimension: orthogonal to } \vec{s_1}}$$



$$\vec{s_2} = \underline{\underline{\beta}} \cdot \vec{s_1} + \vec{q_2}$$

$$\vec{q_2} = \vec{s_2} - \beta \vec{s_1}$$

Review of projections

$$\langle \vec{s_1}, \vec{q_2} \rangle = 0$$

$$\Rightarrow \langle \vec{s_1}, \vec{s_2} - \beta \vec{s_1} \rangle = 0$$

$$\Rightarrow \langle \vec{s_1}, \vec{s_2} \rangle - \beta \cdot \|\vec{s_1}\|^2 = 0$$

$$\Rightarrow \quad \beta = \frac{\langle \vec{s_1}, \vec{s_2} \rangle}{\|\vec{s_1}\|^2}$$

$$\mathrm{Proj}_{\vec{s_1}} \vec{s_2} = \overbrace{\frac{\langle \vec{s_2}, \vec{s_1} \rangle}{\|\vec{s_1}\|^2}}^{\beta} \vec{s_1}$$

Gram-Schmidt Algorithm / Procedure

Given $\{\vec{s}_1, \vec{s}_2, \dots \vec{s}_n\}$

Convert this to a set

$$\{\vec{q}_1, \vec{q}_2, \dots \vec{q}_n\}.$$

such that: $\langle \vec{q}_i, \vec{q}_i \rangle = \|\vec{q}_i\|^2 = 1$

$$\langle \vec{q}_i, \vec{q}_j \rangle = 0$$

and:

$$\text{span}\{\vec{s}_1\} = \text{span}\{\vec{q}_1\}$$

$$\text{span}\{\vec{s}_1, \vec{s}_2\} = \text{span}\{\vec{q}_1, \vec{q}_2\}.$$

$$\vdots$$

$$\text{span}\{\vec{s}_1, \vec{s}_2, \dots \vec{s}_n\} = \text{span}\{\vec{q}_1, \vec{q}_2, \dots \vec{q}_n\}.$$

Consider: linearly independent set.

$\{\vec{s}_1, \vec{s}_2, \dots \vec{s}_n\}$ lin indep.

Gram-Schmidt - Alg.

$\{\vec{s_1}, \vec{s_2}, \vec{s_3}\}$.

① $\{\vec{s_1}\} \rightarrow \vec{q_1}$ find.

$$\frac{\vec{s_1}}{\|\vec{s_1}\|} = \vec{q_1} \rightarrow \text{unit norm.}$$

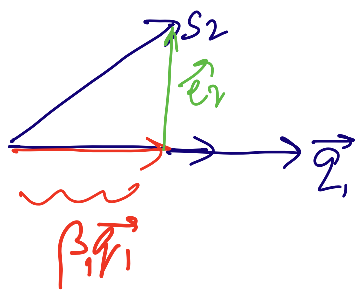
$$\text{span}\{\vec{q_1}\} = \text{span}\{\vec{s_1}\} \checkmark$$

② $\{\vec{s_1}, \vec{s_2}\}$

$$\boxed{\vec{q_1} = \frac{\vec{s_1}}{\|\vec{s_1}\|}}$$

What is new in $\vec{s_2}$, that is not captured by $\vec{q_1}$.

Remove from $\vec{s_2}$, the projection of $\vec{s_2}$ onto $\vec{q_1}$



$$\vec{e_2} = \vec{s_2} - \underbrace{\frac{\langle \vec{s_2}, \vec{q_1} \rangle}{\|\vec{q_1}\|^2}}_{\beta_1} \cdot \vec{q_1}$$

$$\vec{e_2} = \vec{s_2} - \langle \vec{s_2}, \vec{q_1} \rangle \cdot \vec{q_1}$$

$$\vec{q_2} = \frac{\vec{e_2}}{\|\vec{e_2}\|} \rightarrow \underline{\underline{\text{unit norm}}}$$

<u>Check</u>: $\langle \vec{q_2}, \vec{q_1} \rangle = \left\langle \frac{\vec{e_2}}{\|\vec{e_2}\|}, \vec{q_1} \right\rangle$

$$= \left\langle \frac{\vec{s_2} - \langle \vec{s_2}, \vec{q_1} \rangle \cdot \vec{q_1}}{\|\vec{e_2}\|}, \vec{q_1} \right\rangle$$

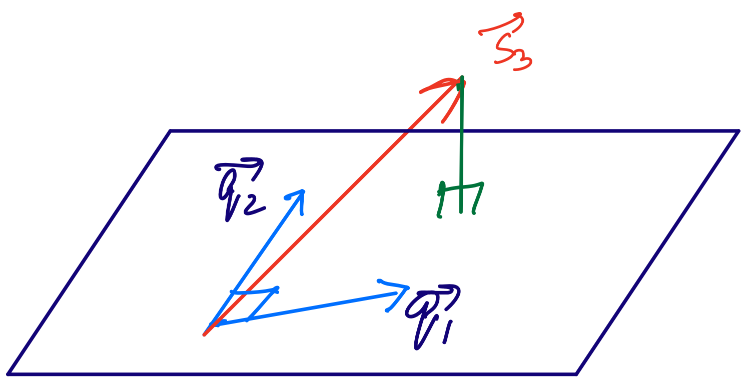
$$= \frac{1}{\|\vec{e_2}\|} \left[\langle \vec{s_2}, \vec{q_1} \rangle - \langle \vec{s_2}, \vec{q_1} \rangle \underbrace{\langle \vec{q_1}, \vec{q_1} \rangle}_{1} \right]$$

$$= 0$$

Furthermore: $\operatorname{span}\{\vec{s_1}, \vec{s_2}\} = \operatorname{span}\{\vec{q_1}, \vec{q_2}\}$.

---

3 $\{\vec{s_1}, \vec{s_2}, \vec{s_3}\}$. $\boxed{\vec{q_1}, \vec{q_2}}$



Project $\vec{s_3}$ onto $\operatorname{span}\{\vec{q_1}, \vec{q_2}\}$

$$\text{proj} = \begin{bmatrix} \langle \vec{s_3}, \vec{q_1} \rangle \\ \langle \vec{s_3}, \vec{q_2} \rangle \end{bmatrix} \begin{bmatrix} \vec{q_1} & \vec{q_2} \end{bmatrix}$$

$$= \underbrace{\langle \vec{s_3}, \vec{q_1} \rangle \vec{q_1} + \langle \vec{s_3}, \vec{q_2} \rangle \vec{q_2}}_{\text{true by projection formula \& the fact that } \vec{q_1}, \vec{q_2} \text{ are orthonormal}}$$

$$\vec{e_3} = \vec{s_3} - \left(\langle \vec{s_3}, \vec{q_1} \rangle \vec{q_1} + \langle \vec{s_3}, \vec{q_2} \rangle \vec{q_2} \right)$$

$$\vec{q_3} = \frac{\vec{e_3}}{\|\vec{e_3}\|} \quad \checkmark$$

<u>Check:</u> $\text{span}\{\vec{q_1}, \vec{q_2}, \vec{q_3}\} = \text{span}\{\vec{s_1}, \vec{s_2}, \vec{s_3}\}$

<u>Check:</u> $\langle \vec{q_3}, \vec{q_1} \rangle = \langle \vec{q_3}, \vec{q_2} \rangle = 0.$

$$\underbrace{\overbrace{\begin{bmatrix} | & | & | \\ \vec{s_1} & \vec{s_2} & \vec{s_3} \\ | & | & | \end{bmatrix}}^{A}}_{\text{Columnspace}} \longleftrightarrow \underbrace{\begin{bmatrix} | & | & | \\ \vec{q_1} & \vec{q_2} & \vec{q_3} \\ | & | & | \end{bmatrix}}_{\text{Basis for columnspace of } A.}$$