Tuesday 10:30 - 11:45

(5) Orthogonalization (inner product spaces)
QR, Cauchy-Schwartz gram-schmidt

Inner Product space - Vector space $V$
w/ defined inner product $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{F}$

1) symmetry $\langle u, v \rangle = \langle v, u \rangle$
2) linearity in one argument
$\langle \alpha u + v, w \rangle = \alpha \langle u, w \rangle + \langle v, w \rangle$
$= \langle w, \alpha u + v \rangle$ (symmetry)

3) positive-definiteness: $\langle u, u \rangle \geq 0$ w/
equality $\Leftrightarrow$ $u = 0$
* can def.
most important is dot-product ($\ell_2$-inner product) $\boxed{\langle u, v \rangle = \sum_i u_i v_i = v^T u}$

Functional analysis example ($L^2$ inner product)

functions in $[0,1]$: $\langle f, g \rangle_{L^2} = \int_0^1 f(x) g(x) dx$

• used in pdes and analysis

Cauchy-Schwarz inequality: inner product between 2 vectors is bounded in size by product of natural norms

$$\boxed{|\langle u, v \rangle| \leq \|u\| \cdot \|v\|}$$

equality $\Leftrightarrow$ $v = \alpha u$
If $\vec{v} = 0$, clearly holds.

Assume $v \neq 0$

Desire $z = u - \frac{\langle u,v\rangle}{\langle v,v\rangle} v$ (projection operator we will see later)

$$\Rightarrow u = z + \frac{\langle u,v\rangle}{\langle v,v\rangle} v$$

$$\|u\|^2 = \|z\|^2 + \left|\frac{\langle u,v\rangle}{\langle v,v\rangle}\right|^2 \|v\|^2 + \cancel{2\frac{\langle u,v\rangle}{\langle v,v\rangle}\langle v,z\rangle}\ 0$$ (Pythagorean Theorem)

cross-term is 0 $z \perp v$

$$\langle z,v\rangle = \langle u,v\rangle - \frac{\langle u,v\rangle}{\langle v,v\rangle}\langle v,v\rangle \to 0 \checkmark$$

$$\Leftrightarrow \|u\|^2 = \frac{|\langle u,v\rangle|^2}{\|v\|^4}\|v\|^2 + \|z\|^2$$

$$\geq \frac{|\langle u,v\rangle|^2}{\|v\|^2}$$

$$\Leftrightarrow \boxed{|\langle u,v\rangle|^2 \leq \|u\|^2\|v\|^2}$$ (Take sqrt)

$v = \alpha u$

$$|\langle u,\alpha u\rangle| = |\alpha|\,\|u\|^2 \checkmark$$

**Triangle Inequality**) $\|x+y\| \le \|x\| + \|y\|$

Reverse triangle inequality: $\|u-v\| \ge \big| \|u\| - \|v\| \big|$

Hint: $\|u\| = \|(u-v)+v\|$

$\|u\| = \|(u-v)+v\| \le \|u-v\| + \|v\|$

$\Leftrightarrow \|u\| - \|v\| \le \|u-v\|$

$\|v\| = \|-(u-v)+u\| \le \|u-v\| + \|u\|$

$\Leftrightarrow -\|u-v\| \le \|u\| - \|v\|$ ✓

**example**) symmetric positive definite matrix can define a ~~ma~~ norm

A-norm $\|x\|_A = \sqrt{x^TAx} = \sqrt{x^TR^TRx}$
$= \sqrt{(Rx)^T(Rx)} = \|Rx\|_2$

1) $x^TAx \ge 0$ & $x^TAx = 0 \Leftrightarrow x=0$

2) $A = R^TR$ (cholesky)

$\|k\,x\|_A^2 = \cancel{\#}\; k\,x^TAx = k\,x^TR^TRx$

$= k(Rx)^TRx = \|R(kx)\| = |k|\,\|Rx\| = |k|\,\|x\|_A$

properties follow from 2 norm

3) $\|x+y\|_A = \|R(x+y)\| = \|Rx + Ry\|$

$\le \|Rx\| + \|Ry\| = \|x\|_A + \|y\|_A$ ✓

Definition) vectors $u$ and $v$ are orthogonal w/ respect to inner product if $\boxed{\langle u,v\rangle = 0}$ $\quad u \perp v$

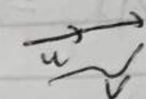
Standard definition of dot product:

$x^T y = 0 \qquad \begin{bmatrix}1\\1\end{bmatrix}^T \begin{bmatrix}1\\-1\end{bmatrix} = 0$

From vector calculus: $\boxed{\cos\theta = \dfrac{x^T y}{\|x\|\,\|y\|}}$

$\theta$ angle between vectors $\theta \in [0, \frac{\pi}{2}]$

angle is 0 when colinear





and $\frac{\pi}{2}$ when orthogonal

Orthogonal set of vectors: $\{q_j\}_{j=1}^m$ orthogonal if each vector is orthogonal to any other $\boxed{q_j^T q_k = 0 \quad j \neq k}$

Orthonormal) orthogonal and $q_j^T q_j = 1$

Orthogonal basis example

$$\left\{ \begin{bmatrix}1\\0\\0\\1\end{bmatrix}, \begin{bmatrix}0\\1\\1\\0\end{bmatrix}, \begin{bmatrix}1\\0\\0\\-1\end{bmatrix}, \begin{bmatrix}0\\1\\-1\\0\end{bmatrix} \right\}$$

Orthonormal basis:

$$\frac{1}{\sqrt{2}} \left\{ \begin{bmatrix}1\\0\\0\\1\end{bmatrix}, \begin{bmatrix}0\\1\\1\\0\end{bmatrix}, \begin{bmatrix}1\\0\\0\\-1\end{bmatrix}, \begin{bmatrix}0\\1\\-1\\0\end{bmatrix} \right\}$$

- orthogonal basis - nice properties $\{q_i\}$
- easy form for coefficients

$$\boxed{\vec{x} = \sum_{i=1}^{m} \beta_i \vec{q_i}}$$

$$\vec{q_k}^T \vec{x} = \vec{q_k}^T \sum_{i=1}^{m} \beta_i \vec{q_i}$$

$$= \sum_{i=1}^{m} \beta_i \vec{q_k}^T \vec{q_i} = \beta_k \vec{q_k}^T \vec{q_k} \quad \text{(orthogonality)}$$

$$= \beta_k \|\vec{q_k}\|_2^2 \iff \boxed{\beta_k = \frac{\vec{q_k}^T \vec{x}}{\|\vec{q_k}\|_2^2}}$$

$$\boxed{\beta_k = \vec{q_k}^T \vec{x}} \quad \text{(orthonormal)}$$

Orthogonal Matrices - $n \times n$ matrix $Q$ is orthogonal if its columns are orthonormal $\boxed{(Q^T Q = I = Q Q^T)}$

- $Q^T = Q^{-1}$ (very easy to solve)
- well-conditioned
- $\det(Q) = \pm 1$

- Non-square matrix w/ orthonormal columns
  $m \times n \quad m \neq n$
- satisfies $\boxed{Q^T Q = I \text{ but } QQ^T \neq I}$

- investigate why?

$$Q^T Q = \begin{bmatrix} q_1^T \\ | \\ q_n^T \end{bmatrix} \begin{bmatrix} q_1 & — & q_n \end{bmatrix}$$

$$= \begin{bmatrix} q_1^T q_1 & q_1^T q_2 & — & q_1^T q_n \\ | & & & \\ q_n^T q_1 & — & & q_n^T q_n \end{bmatrix} = \begin{bmatrix} 1 & 0 & — & 0 \\ | & \ddots & & \\ 0 & & & 1 \end{bmatrix} = I_n$$

- elementwise $I_{ij} = (Q^TQ)_{ij} = \sum_{k=1}^{n} q^T_{ik} q_{kj}$

$= \sum_{k=1}^{n} q_{ki} q_{kj} = \langle q_i, q_j \rangle = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases}$ ✓

Consider $QQ^T$ $n < m$ $Q$ full rank $n$

$$\begin{array}{c} \\ m \end{array}\left[\begin{array}{c} n \\ \\ \\ \end{array}\right] \begin{array}{c} n \\ \\ \end{array}\left[\begin{array}{c} \\ m \end{array}\right] = \left[\begin{array}{c} \\ \\ \\ \end{array}\right]^{m}_{m}$$

Can't be $I_m$ — rank$(n) < m$ ~~...~~
won't be invertible.

- preserves 2-norm $\|Qx\|_2^2 = \|x\|_2^2$
$x^T \underbrace{Q^TQ}_{I} x = x^T x$ ✓

## Gram-Schmidt

- how to construct orthonormal basis for a subspace — orthogonalization

- Uses idea of vector projection of $\vec{u}$ onto $\vec{v}$

$$\text{Proj}_v(\vec{u}) = \frac{\langle \vec{u}, \vec{v} \rangle}{\langle \vec{v}, \vec{v} \rangle} \vec{v} = (\|u\| \cos\theta) \hat{v}$$

$$\cos\theta = \frac{\langle u, v \rangle}{\|\vec{u}\| \|\vec{v}\|}, \quad \hat{v} = \frac{v}{\|\vec{v}\|} \text{ (unit vector)}$$

- results in vector with direction colinear w/ $\vec{v}$ but magnitude dependent on inner product.

Recall in Cauchy-Schwarz we used
$\vec{z} = \vec{u} - \frac{\langle \vec{u}, \vec{v} \rangle}{\langle \vec{v}, \vec{v} \rangle} \vec{v} = \vec{u} - \mathrm{proj}_v(\vec{u})$

Gram-Schmidt: Have basis
$\{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_n\}$
Want orthogonal basis $\{\vec{v}_1, \ldots, \vec{v}_n\}$

$\vec{v}_1 = \vec{u}_1$
$\vec{v}_2 = \vec{u}_2 - \frac{\langle \vec{u}_2, \vec{v}_1 \rangle}{\langle \vec{v}_1, \vec{v}_1 \rangle} \vec{v}_1 \quad (\vec{v}_2 \perp \vec{v}_1)$

$\vec{v}_3 = \vec{u}_3 - \underbrace{\frac{\langle \vec{u}_3, \vec{v}_1 \rangle}{\langle \vec{v}_1, \vec{v}_1 \rangle} \vec{v}_1}_{\mathrm{proj}_{v_1}(\vec{u}_3)} - \underbrace{\frac{\langle \vec{u}_3, \vec{v}_2 \rangle}{\langle \vec{v}_2, \vec{v}_2 \rangle} \vec{v}_2}_{\mathrm{proj}_{v_2}(\vec{u}_3)}$

$\vdots \quad \boxed{(\vec{v}_3 \perp \mathrm{span}\{\vec{v}_1, \vec{v}_2\})}$

$$\boxed{\vec{v}_n = \vec{u}_n - \sum_{i=1}^{n-1} \frac{\langle \vec{u}_n, \vec{v}_i \rangle}{\langle \vec{v}_i, \vec{v}_i \rangle} \vec{v}_i}$$

$v_i \perp v_{i-1} = 0$
$i = 1, \ldots, n$

check: $\langle \vec{v}_1, \vec{v}_2 \rangle = \langle \vec{v}_1, \vec{u}_2 \rangle - \frac{\langle \vec{u}_2, \vec{v}_1 \rangle}{\cancel{\langle \vec{v}_1, \vec{v}_1 \rangle}} \cancel{\langle \vec{v}_1, \vec{v}_1 \rangle}$
$= 0$ ✓

check $\langle \vec{v}_n, \vec{v}_j \rangle = 0$ for $j < n$
$= \langle \vec{u}_n, \vec{v}_j \rangle - \sum_{i=1}^{n-1} \frac{\langle \vec{u}_n, \vec{v}_i \rangle}{\langle \vec{v}_i, \vec{v}_i \rangle} \langle \vec{v}_j, \vec{v}_i \rangle$ $\quad$ "0 unless $j = i$"
$= \langle \vec{u}_n, \vec{v}_j \rangle - \frac{\langle \vec{u}_n, \vec{v}_j \rangle}{\cancel{\langle \vec{v}_j, \vec{v}_j \rangle}} \cancel{\langle \vec{v}_j, \vec{v}_j \rangle} = \boxed{0}$ ✓

- Divide each vector by length for orthonormal basis

- $\boxed{\text{QR decomposition}}$ another matrix factorization
- $\boxed{A = QR}$, $Q$ orthogonal, $R$ upper triangular
- help us solve $Ax = b$

- In Gauss elim, we computed LU by taking linear combo of rows

$\boxed{LA = U}$ premultiply due to row operations

- In Gram-Schmidt (GS) process we perform column manipulations, not row

Can we write GS as matrix-matrix op (post multiplication by S)

$\boxed{AS = Q}$

$\boxed{\text{Nonsingular } A}$ - square + invertible

- Then $AS = Q$ w/ $A$ and $Q$ nonsingular

$\Rightarrow$ $S$ is nonsingular

$\boxed{S = A^{-1}Q}$ $\boxed{S^{-1} = Q^T A}$

$\boxed{A = QS^{-1}}$ Let $\boxed{R = S^{-1}}$

$= \boxed{QR}$

Find expression for $R$ explicitly:

$Q^T A = Q^T Q R = R$

$$\begin{bmatrix} q_1^T \\ | \\ q_n^T \end{bmatrix} [\vec{a_1} \text{ --- } \vec{a_n}] \Rightarrow \boxed{r_{ij} = \vec{q_i}^T \vec{a_j}}$$

* <u>Key</u>: R is upper triangular

From GS: $\{\vec{q_i}\}$ is orthonormal basis for columns of A

$\vec{a_1}$ is multiple of $\vec{q_1}$
$\vec{a_2}$ is linear combo $\vec{q_2}, \vec{q_1}$
⋮
$\vec{a_j}$ is linear combo $\vec{q_1}, \dots, \vec{q_j}$

$$\boxed{\vec{a_j} = \sum_{k=1}^{j} \alpha_i \vec{q_k}}$$

$$r_{ij} = \vec{q_i}^T \vec{a_j} = \sum_{k=1}^{j} \alpha_i \vec{q_i}^T \vec{q_k}$$

- use <u>orthogonality</u> for $i > j \geq k$

$\boxed{r_{ij} = 0}$ distinct $k \neq i$ $\boxed{q_i^T q_k = 0}$

upper tri by def
if $i \leq j$ then $= \alpha_i q_i^T q_i = \boxed{\alpha_i}$ ✓

- <u>What are the diagonal elements?</u>

$r_{ii} = \vec{q_i}^T \vec{a_i}$

$r_{11} = \vec{q_1}^T \vec{a_1}$, $\vec{q_1} = \frac{\vec{a_1}}{\|\vec{a_1}\|} \Rightarrow \frac{\vec{a_1}^T \vec{a_1}}{\|\vec{a_1}\|} = \frac{\|\vec{a_1}\|^2}{\|\vec{a_1}\|}$

$\Rightarrow \boxed{\|\vec{a_1}\|} = \|\vec{w_1}\|$

Let $\{\vec{w_i}\}$ be orthogonal basis $\left(\vec{q_i} = \frac{w_i}{\|w_i\|}\right.$ ⟵ orthonormal$\left.\right)$

From GS: $\boxed{\vec{w_2} = \vec{a_2} - (\vec{a_2}^T \vec{q_1}) \vec{q_1}}$

* orthogonal to any scalar multiple

$\Rightarrow \vec{a_2} = \vec{w_2} + (\vec{a_2}^T \vec{q_1}) \vec{q_1}$

$r_{22} = \vec{q_2}^T \vec{a_2} = \vec{q_2}^T (\vec{w_2} + (\vec{a_2}^T \vec{q_1}) \vec{q_1})$

$= \vec{q_2}^T \vec{w_2}$, since $\vec{q_2}^T \vec{q_1} = 0$ (orthogonal)

$= \dfrac{\vec{w_2}^T \vec{w_2}}{\|\vec{w_2}\|} = \boxed{\|\vec{w_2}\|}$

$\boxed{r_{ii} = \|\vec{w_i}\| \quad \forall i}$ norm constants

- GS in matrix form $A = QR$

$Q = [\vec{q_1}, \ldots, \vec{q_n}]$ $\boxed{R_{ij} = \begin{cases} \vec{q_i}^T \vec{a_j} & i \leq j \\ 0 & i > j \end{cases}}$

orthogonal
$i \neq j$ no overlapping indices

- GS is a triangular process for orthogonalizing columns of A (triangular orthogonalization)

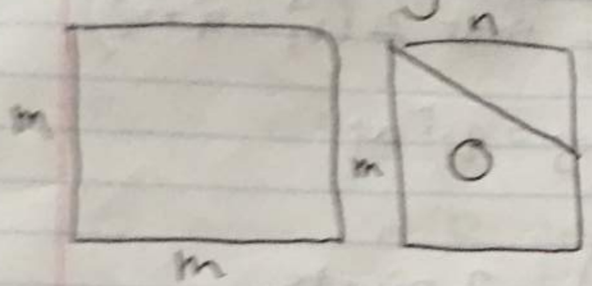
Other methods in 302: Householder
Givens

- Similar view to LU - premult by Q to get triangular R

- Like LU, QR can be used to effectively compute solutions to
$A\vec{x} = \vec{b} \Rightarrow QR\vec{x} = \vec{b}$

1) $\boxed{R\vec{x} = Q^T\vec{b}}$ solve for $\vec{x}$ using back substitution $\boxed{\vec{x} = R^{-1}\vec{y} \quad \vec{y} = Q^T\vec{b}}$

**Theorem** Every matrix $A \in \mathbb{R}^{m \times n}$ $m \geq n$ has full QR factorization $A = QR$, $Q$ orthogonal $\in \mathbb{R}^{m \times m}$ and upper triangular $R \in \mathbb{R}^{m \times n}$



- Note typically LU is preferred among direct methods for solving systems because it takes $\frac{1}{2}$ as many operations

- QR is still 1 of most useful factorizations due to orthogonality property & invariance

- $\|A\vec{x} - \vec{b}\|_2 = \|QR\vec{x} - \vec{b}\|_2 = \|Q^T QR\vec{x} - Q^T\vec{b}\|_2$

$= \boxed{\|R\vec{x} - Q^T\vec{b}\|_2}$ (solve least squares later)

**Singular example** $A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix}$

**Gram-Schmidt** $\vec{q}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$

$$\vec{q}_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} - \frac{\langle \vec{q}_1, \vec{a}_2 \rangle}{\langle \vec{q}_1, \vec{q}_1 \rangle} \vec{a}_2 = \boxed{\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}} - 0$$

$$\vec{q}_3 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} - \frac{\langle \vec{q}_1, \vec{a}_3 \rangle}{\langle \vec{q}_1, \vec{q}_1 \rangle} \vec{q}_1 - \frac{\langle \vec{q}_2, \vec{a}_3 \rangle}{\langle \vec{q}_2, \vec{q}_2 \rangle} \vec{q}_2$$

$$= \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} - 1 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} - 1 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \boxed{\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}}$$

- Note $\vec{a}_3 = 0$ because 3rd column of A is dependent on first 2 ($\vec{v}_3 = \vec{v}_1 + \vec{v}_2$)

$$\boxed{Q = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}} \qquad R_{ij} = q_i^T a_j$$

$$\boxed{R = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}}$$

2 pivots
0 on diagonal
$\det(R) = 0$ not invertible

Check $A = QR$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix} \checkmark$$

- Problem is that Q is not invertible (can't have $\vec{0}$ in a basis)

We can define skinny QR only taking nonzero columns from Q

$$\boxed{A = Q_{skinny} R_{skinny} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}}$$

$3 \times 2$ $\qquad 2 \times 3$

- ignore 0 irrelevant row of R & column of Q

<u>careful</u>: $Q^T Q = I$ but $QQ^T \neq I$
$= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ rk$=2$

- we could also leave 3rd row of R & make Q orthogonal

make orthogonal basis

$$A = Q_F R = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

Full QR - Q orthogonal/square can invert

- 0 row in R indicates A singular & 3rd column of $Q_F$ not in column space

$$Q^T Q = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix} = I_2 \checkmark$$

$$\text{But } Q Q^T = \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

3×3 but only 2 pivots - can't be $I_3$

Extra exercises | Show product of 2 lower triangular matrices is lower triangular

$$e_i^T A B e_j = \boxed{a_i^T b_j}$$

$$= \sum_{k=1}^{n} a_{ik} b_{kj}$$

Suppose A lower triangular

$\Rightarrow a_{ik} = 0 \quad k > i$

$$(B)_{ij} = \sum_{k=1}^{i} a_{ik} b_{kj} \qquad \text{But } \boxed{b_{kj} = 0 \text{ if } j > k}$$

$= \boxed{0} \quad j > k \qquad$ Now $\boxed{1 \le k \le i}$

So if $j > i \ge k \quad (AB)_{ij} = 0 \checkmark$

• Show $\|A\| \geq \|A^{-1}\|^{-1}$

$$1 = \|I\| = \|A^{-1}A\| \leq \|A^{-1}\|\,\|A\|$$

$$\Leftrightarrow \boxed{\|A\| \geq \frac{1}{\|A^{-1}\|} = \|A^{-1}\|^{-1}}$$