Linear Algebra for Computer Science

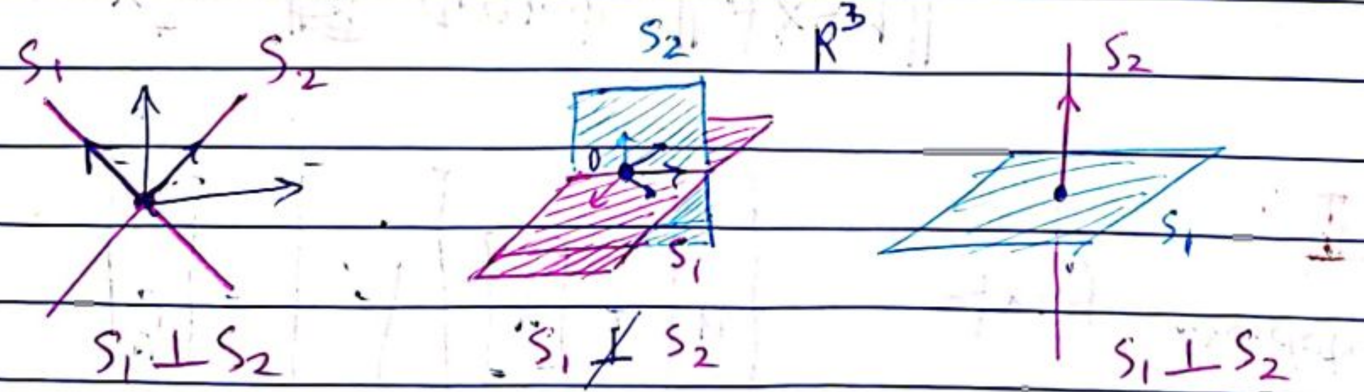
Lecture 18

Gram-Schmidt Orthogonalization, QR decompostion

Orthogonal subspaces

$S_1, S_2 \subseteq V$ are linear subspaces of V.

S_1, S_2 are orthogonal if for all $v_1 \in S_1$, $v_2 \in S_2$ $\quad v_1 \perp v_2$



Four Basic Subspaces

K. N. Toosi
University of Technology

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

$\operatorname{rank}(A) = r$

$C(A) = \{Ax \mid x \in \mathbb{R}^n\}$ column space

$\boxed{R(A) =} C(A^T) = \{A^T x \mid x \in \mathbb{R}^m\}$ row space

$(x^T A)^T$

$N(A) = \{x \in R^n \mid Ax = 0\}$ null space

$N(A^T) = \left\{ x \in \mathbb{R}^m \;\middle|\; \begin{matrix} x^T A = 0^T \\ A^T x = 0 \end{matrix} \right\}$

$x^T A = 0^T$ left null space

Four Basic Subspaces

K. N. Toosi
University of Technology

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

$\text{rank}(A) = r$

$C(A) = \{ Ax \mid x \in \mathbb{R}^n \}$ column space

$\boxed{R(A) =} C(A^T) = \{ A^T x \mid x \in \mathbb{R}^m \}$ row space

$(x^T A)^T$

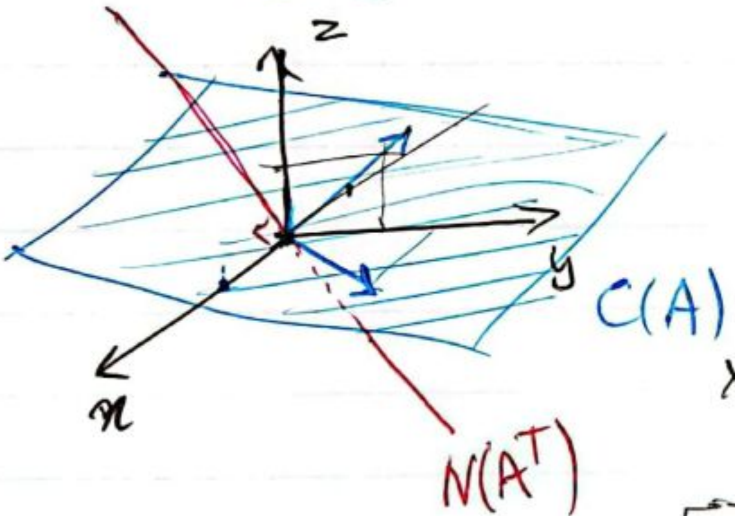
$N(A) = \{ x \in \mathbb{R}^n \mid Ax = 0 \}$ null space

$x^T A = 0^T$ left null space

$N(A^T) = \left\{ x \in \mathbb{R}^m \;\middle|\; \begin{array}{l} x^T A = 0^T \\ A^T x = 0 \end{array} \right\}$

Four Basic Subspaces - Example

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



$$N(A^T) = \{ x \in \mathbb{R}^3 \mid A^T x = 0 \}$$

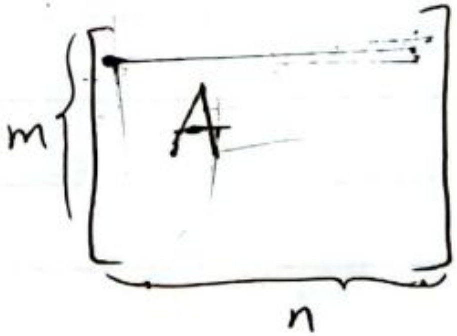
$$\begin{bmatrix} 1 & 2 & 0 \\ -1 & 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 0$$

$$x \in N(A^T) \left\{ \begin{array}{l} x \perp (1,2,0)^T \\ x \perp (-1,1,1)^T \end{array} \right\} \Rightarrow x \perp C(A)$$

Four Basic Subspaces

$N(A^T)$

$$C(A) \perp N(A^T)$$
$$R(A) \perp N(A)$$



$\dim(R(A)) = r$

$\dim(C(A)) = r$

$\dim(N(A)) = n-r$

$\dim(N(A^T)) = m-r$

$\mathrm{rank}(A) = r$

$R(A), N(A) \subseteq \mathbb{R}^n$ $\qquad \dim(R(A)) + \dim(N(A)) = n$

$\mathbb{R}^n = R(A) \oplus N(A)$

$C(A), N(A^T) \subseteq \mathbb{R}^m$ $\qquad \dim(C(A)) + \dim(N(A^T)) = m$

Four Basic Subspaces

$N(A^T)$

$$\boxed{\begin{aligned} C(A) &\perp N(A^T) \\ R(A) &\perp N(A) \end{aligned}}$$

$$m\left\{\underbrace{\begin{bmatrix} & & \\ & A & \\ & & \end{bmatrix}}_{n}\right.$$

$\dim(R(A)) = r$

$\dim(C(A)) = r$

$\dim(N(A)) = n-r$

$\dim(N(A^T)) = m-r$

$\operatorname{rank}(A) = r$

$R(A),\ N(A) \subseteq \mathbb{R}^n \qquad \dim(R(A)) + \dim(N(A)) = n$

$\mathbb{R}^n = R(A) \oplus N(A)$

$C(A),\ N(A^T) \subseteq \mathbb{R}^m \qquad \dim(C(A)) + \dim(N(A^T)) = m$

Four Basic Subspaces

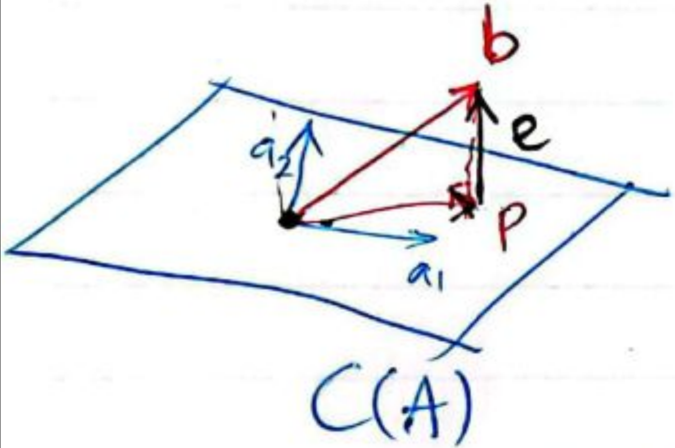
$c \in C(A^T) = R(A) \qquad c = A^T x \qquad c^T = \underline{x^T A}$

$n \in N(A)$

$$c^T n = (x^T A)\, n = x^T (A n) = x^T \vec{0} = 0$$

$$\Rightarrow \quad R(A) \perp N(A)$$

the residual vector



$$A$$

$$p = Ax$$

$$x = (A^TA)^{-1}A^Tb$$

$$p = A(A^TA)^{-1}A^Tb = Pp$$

$$Ax = b$$

$$Ax + e = b$$

$$Ax = b - \underline{\underline{e}}$$

$$\boxed{e = b - Ax}$$

residual projection

$$e = b - Ax = b - p = b - A(A^TA)^{-1}A^Tb$$

$$= \underbrace{\left(I - A(A^TA)^{-1}A^T\right)}_{P_N \ ?} b = (I - P)b$$

$$P_N^T = P_N$$

$$P_N P_N = (I-P)(I-P) = I - P - P + PP = I - P - P + P = I - P = P_N$$

P_N is a projection into some space

$$P_N P = (I-P)P = P - PP = \cancel{0}\, P - P = 0 \qquad P_N(Px) = 0$$

$$x \in C(A) \Rightarrow \quad x = Ay \Rightarrow P_N x = \left(I - A(A^TA)^{-1}A^T\right)Ay = Ay - Ay = 0$$

residual projection

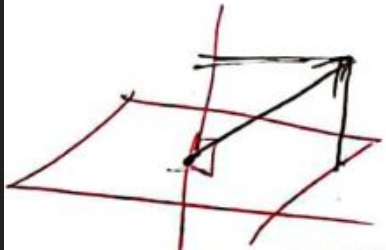
K. N. Toosi University of Technology

III

$\underline{\underline{P_A}}$: Projection into C(A)

$P_N = \underline{I - P_A}$ projection into $N(A^T)$

left null space of A



$$x = P_A x + P_N x \qquad (P_A + P_N = I)$$

P_S : P_S projection into linear subspace S

P_N : is projection into the orthogonal complement of S

Orthogonal Complement

$S^{\perp} = \{x \mid x \perp y \text{ for all } y \in S\}$

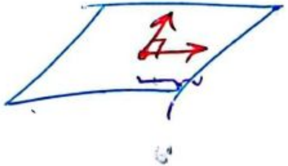
Projection matrix with Orthonormal Columns

K. N. Toosi
University of Technology

Assume A has orthonormal columns

$$A = \begin{bmatrix} a_1 & a_2 & \cdots & a_n \end{bmatrix} \qquad \begin{cases} a_i^T a_i = 1 \\ a_i^T a_j = 0 \quad i \neq j \end{cases} \qquad A^T A = I$$

$$P = A(A^T A)^{-1} A^T = A A^T$$

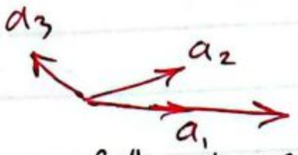


$$Ax = b - e \qquad x = (A^T A)^{-1} A^T b = A^T b$$

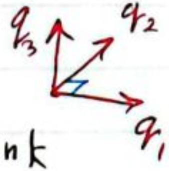
$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} a_1^T \\ a_2^T \\ \vdots \\ a_n \end{bmatrix} b \qquad x_i = \langle a_i, b \rangle$$

Orthogonalization

Let $A = \begin{bmatrix} a_1 & a_2 \cdots & a_n \end{bmatrix}$



has full column rank



$\|q_i\| = 1$

$q_i^T q_j = 0$
$i \neq j$

find Q such that ~~da~~that $\begin{cases} C(Q) = C(A) \\ Q^T Q = I \end{cases}$ Q has orthonormal columns

$$A \xrightarrow{\text{orthogonalization}} Q$$

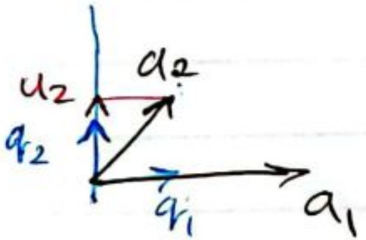
find an orthonormal basis $q_1, \cdots, q_n$

s.t.

$\text{span}(q_1 - q_n) = \text{span}(a_1 - a_n)$

Orthogonalization

$$\operatorname{span}(q_1 - q_n) = \operatorname{span}(a_1 - a_n)$$



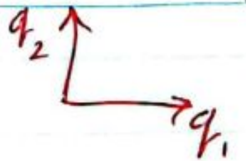
$$\rightarrow q_1 = \frac{a_1}{\|a_1\|}$$

$$u_1 = a_1 \qquad q_1 = \frac{u_1}{\|u_1\|} = \frac{a_1}{\|a_1\|}$$

$$u_2 = a_2 - \frac{a_1 a_1^T}{a_1^T a_1} a_2 = a_2 - \frac{a_1 a_1^T}{\|a_1\|^2} a_2 = a_2 - \left(\frac{a_1}{\|a_1\|}\right)\left(\frac{a_1}{\|a_1\|}\right)^T a_2$$

$$\boxed{u_2 = a_2 - q_1 \, q_1^T a_2} = a_2 - q_1 (q_1^T a_2)$$

$$q_2 = \frac{u_2}{\|u_2\|}$$



$$\Rightarrow \underbrace{\|u_2\| q_2}_{u_2} + (q^T a_2) q_1 = a_2$$

Orthogonalization

$$a_1 = \|u_1\|\, q_1 = \alpha\, q_1$$

$$a_2 = \|u_2\|\, q_2 + (q_1^T a_2)\, q_1 = \beta\, q_1 + \gamma\, q_2$$

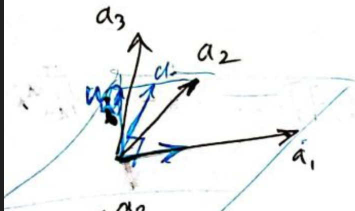
$$\begin{bmatrix} a_1 & a_2 \end{bmatrix} = \begin{bmatrix} q_1 & q_2 \end{bmatrix} \begin{bmatrix} \alpha & \beta \\ 0 & \gamma \end{bmatrix}$$

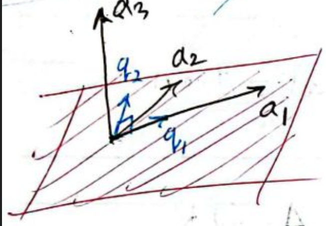
$$A = Q\, R$$

Orthogonalization

V

$$A = \begin{bmatrix} a_1 & a_2 & a_3 \end{bmatrix}$$





$$u_1 = a_1 \qquad q_1 = \frac{u_1}{\|a_1\|} = \frac{a_1}{\|a_1\|}$$

$$u_2 = a_2 - q_1 q_1^T a_2 \qquad q_2 = \frac{u_2}{\|u_2\|}$$

$$u_3 = a_3 - P_{\mathrm{span}(a_1, a_2) = \mathrm{span}(q_1, q_2)} \, a_3$$

$$u_3 = a_3 - P_{[a_1\, a_2]} \, a_3 = a_3 - P_{[q_1\, q_2]} \, a$$

A

$$u_3 = a_3 - \begin{bmatrix} q_1 & q_2 \end{bmatrix} \begin{bmatrix} q_1^T \\ q_2^T \end{bmatrix} a_3$$

$$u_3 = a_3 - q_1 \boxed{q_1^T a_3} - q_2 \boxed{q_2^T a_3}$$

$$q_3 = \frac{u_3}{\|u_3\|} \qquad -\gamma q_3 = u_3 = a_3 - q_1 \alpha - q_2 \beta$$

$$a_3 = \alpha q_1 + \beta q_2 + \gamma q_3$$

Orthogonalization

K. N. Toosi
University of Technology

$$u_3 = a_3 - q_1 \boxed{q_1^T a_3} - q_2 \boxed{q_2^T a_3}$$

$$q_3 = \frac{u_3}{\|u_3\|} \qquad \gamma q_3 = u_3 = a_3 - q_1 \alpha - q_2 \beta$$

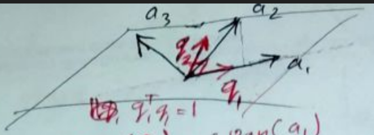
$$a_3 = \alpha q_1 + \beta q_2 + \gamma q_3$$

$$\begin{bmatrix} a_1 & a_2 & a_3 \end{bmatrix} = \begin{bmatrix} q_1 & q_2 & q_3 \end{bmatrix} \begin{bmatrix} e & f & h \\ & g & i \\ & & j \end{bmatrix}$$

$$A = \underbrace{Q}_{\text{orthonormal}} \; \underbrace{R}_{\text{upper triangular}}$$

Q-R decomposition

Orthogonalization



$q_1^T q_1 = 1$

$u_1 = a_1$

$q_1 = \frac{u_1}{\|u_1\|}$

$\text{span}(q_1) = \text{span}(a_1)$

$u_2 = a_2 - q_1 q_1^T a_2$

$q_2 = \frac{u_2}{\|u_2\|}$

$\begin{bmatrix} q_1^T \\ q_2^T \end{bmatrix} \begin{bmatrix} q_1 & q_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

$\text{span}(q_1, q_2) = \text{span}(a_1, a_2)$

$u_3 = a_3 - q_1 q_1^T a_3 - q_2 q_2^T a_3$

$q_3 = \frac{u_3}{\|u_3\|}$

$\begin{bmatrix} q_1 & q_2 & q_3 \end{bmatrix}^T \begin{bmatrix} q_1 & q_2 & q_3 \end{bmatrix} = I$

$\text{span}(q_1, q_2, q_3) = \text{span}(a_1, a_2, a_3)$

$\vdots$

$u_i = a_i - \sum_{j=1}^{i-1} q_j q_j^T a_i$

$q_i = \frac{u_i}{\|u_i\|}$

$a_1, a_2, \dots, a_n \Longrightarrow \underbrace{q_1, q_2, \dots, q_n}_{\text{orthonormal}}$

$\text{span}(q_1, q_2, \dots, q_n) = \text{span}(a_1, a_2, \dots, a_n)$

QR decomposition

K. N. Toosi
University of Technology

LA 19 (II)

$$\underbrace{\begin{bmatrix} a_1 & a_2 & \cdots & a_n \end{bmatrix}}_{m\times n} = \underbrace{\begin{bmatrix} q_1 & q_2 & \cdots & q_n \end{bmatrix}}_{m\times n} \underbrace{\begin{bmatrix} r_{11} & r_{12} & r_{13} & r_{1n} \\ 0 & r_{22} & r_{23} & r_{2n} \\ 0 & 0 & r_{33} & \vdots \\ 0 & 0 & 0 & r_{nn} \end{bmatrix}}_{n\times n}$$

upper-triangular

$$[A] = QR$$

$R \to$ upper triangular

$Q^T Q = I$ (Q has orthonormal columns)

$m = n$

$$\underset{n\times n}{A} = \underset{n\times n}{Q}\ \underset{n\times n}{R}$$

$R \to$ upper-triangular

$Q \to$ orthogonal

$$Q^T Q = Q Q^T = I$$

Q-R decomposition

(RQ, QL, LQ)

Orthogonalization

K. N. Toosi
University of Technology

$a_1, a_2, \cdots, a_n$ idependent

$$q_1 = \frac{a_1}{\|a_1\|}$$

for $i = 2 \cdots n$

$$u_i = a_i - q_1 q_1^T a_i - q_2 q_2^T a_i - \cdots - q_{i-1} q_{i-1}^T a_i$$

$$q_i = \frac{u_i}{\|u_i\|}$$

Gram-Schmidt orthogonalization

Orthogonalization

A : full column rank $\Rightarrow A = QR$

Q: orthonormal columns

R: upper triangular

$A \in \mathbb{R}^{n \times n}$ square & non-singular

$A = QR$

Q: orthogonal (square)

R: upper triangular

(RQ-decompision)