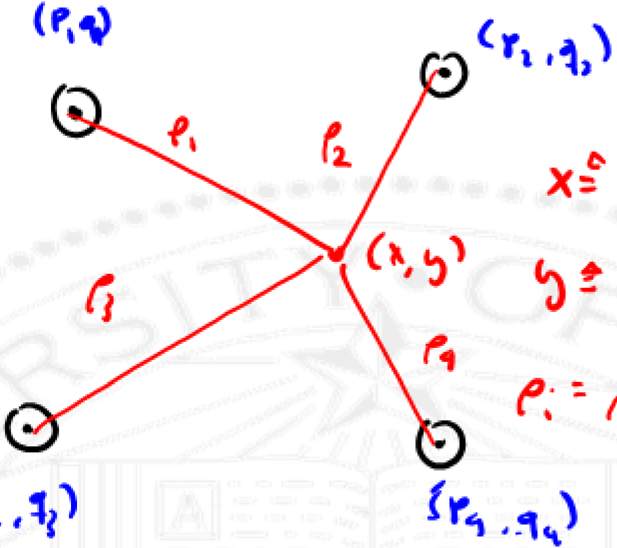
Orthonormal Vectors and QR Factorization

Gabriel Hugh Elkaim

Questions

A is full rank,

R.A.I.M.



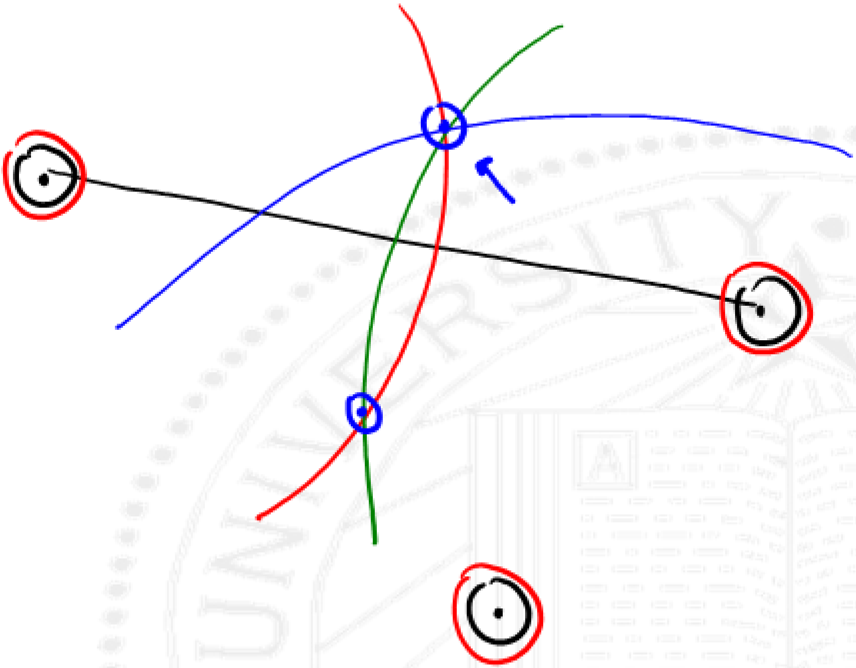
$x \doteq x_0 + \delta x$

$y \doteq y_0 + \delta y$

$\rho_i = \rho_{0i} + \delta \rho_i$

$$\rho_i = \sqrt{(p_i - x)^2 + (q_i - y)^2}$$

$$\left.\frac{\partial \rho_i}{\partial \underline{x}}\right|_{x = x_0, y_0} \qquad \begin{bmatrix} \rho_1 \\ \vdots \\ \rho_n \end{bmatrix} = \underbrace{\overset{A}{\begin{bmatrix} \vdots & \vdots \\ \vdots & \vdots \end{bmatrix}}} \begin{bmatrix} \delta x \\ \delta y \end{bmatrix}$$



$$\begin{bmatrix} x \\ y \\ z \\ t \end{bmatrix}$$

Orthonormal Vectors and QR Factorization

- Orthonormal Vectors — "orthogonal"
- Gram-Schmidt Procedure, QR Factorization
- Orthogonal Decomposition induced by a matrix

Orthonormal Set of Vectors

Set of vectors $\{u_1 \dots u_k\} \in \mathbb{R}^n$

- Normalized such that $\|u_i\| = 1$, $i = 1..k$ length of 1

  (u_i – unit vectors, direction vectors)

- Orthogonal if $u_i \perp u_j$ for $i \neq j$

  $\langle u_i, u_j \rangle = 0$ $u_i^T u_j = 0$ $\forall$ $i \neq j$

GRAM MATRIX

Orthonormal if BOTH

$U = [u_1 \dots u_k] \in \mathbb{R}^{n \times k}$

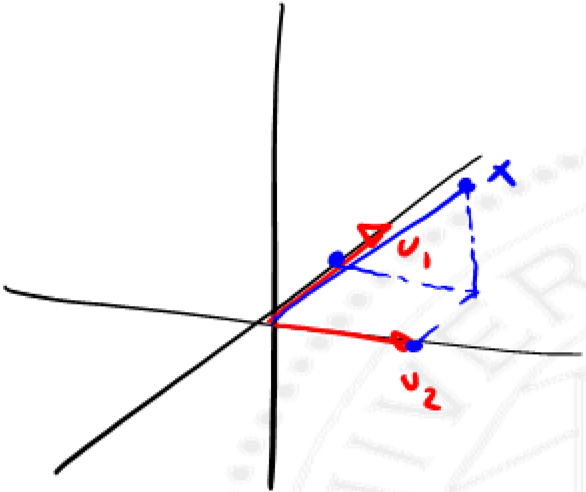
$\overline{U^T U} = I_k$

$$U = \begin{array}{cc} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \\ \quad u_1 \;\; u_2 \end{array}$$

$$U^T U = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I_{2\times 2}$$

$$U U^T = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} = \left[\begin{array}{cc|c} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \hline 0 & 0 & 0 \end{array}\right] \neq I_{3\times 3}$$

U is called ORTHOGONAL if $\{u_1 \dots u_k\}$ are orthonormal if $k < n$ then $UU^T \neq I$.



$\sum_{i}^{k} u_i u_i^T$ — won't be I.

$$\overset{\mathbb{R}^2}{y} = \underset{\mathbb{R}^{2\times 3}}{U}\, x \in \mathbb{R}^3$$

Geometric Properties

Orthonormal vectors are independent

$$\alpha_1 u_1 + \alpha_2 u_2 + \cdots + \alpha_k u_k = 0 \longleftrightarrow \alpha_i = 0.$$

$U = [u_1 \cdots u_k]$ orthonormal basis for $\mathbb{R}^n$

$$\text{span}\{u_1 \cdots u_k\} = R(U).$$

$U \in \mathbb{R}^{n \times k}$ if $\underbrace{w}_{\mathbb{R}^n} = U \underbrace{z}_{\mathbb{R}^k} \longrightarrow \|w\| = \|z\|$

Multiplication by U preserves Length within $\mathbb{R}^k$

$$\|w\|^2 = w^T w = (Uz)^T (Uz) = z^T \underbrace{U^T U}_{I} z = z^T z = \|z\|^2$$

(BE CAREFUL)

Mapping $w = Uz$ is ISOMETRIC

inner product (angle) are preserved $\in \mathbb{R}^k$.

unsigned

$$\langle Uz, U\tilde{z} \rangle = \langle z, \tilde{z} \rangle$$

Orthonormal Basis for R^n

$k = n \quad \{u_1, \dots u_n\}$ orthonormal basis for $\mathbb{R}^n$

$U = [u_1 \dots u_n]$ orthogonal $\mathbb{R}^{n \times n}$

$U^T U = I_{n \times n}$

$U U^T = I_{n \times n}$

$$\boxed{U^{-1} = U^T}$$

$p q^T$

diad, outer product

$$\sum_{i=1}^{n} u_i u_i^T = I$$

$u_i \quad u_i^T$

n rank 1 matrices

$\rightarrow n \times n \ [\quad]$

$$I = UU^T = \sum_{i=1}^{n} u_i u_i^T$$

$$I = \sum_{i=1}^{n} |u_i\rangle\langle u_i|$$

$$y = \sum_{i=1}^{n} |u_i\rangle\langle u_i|x\rangle \leftarrow y = Ux.$$

Expansion in Orthonormal Basis

$\underline{k=n}$

U orthogonal $\quad x = UU^Tx \qquad [u_1 \cdots u_n]\begin{bmatrix} u_1^Tx \\ \vdots \\ u_n^Tx \end{bmatrix}$

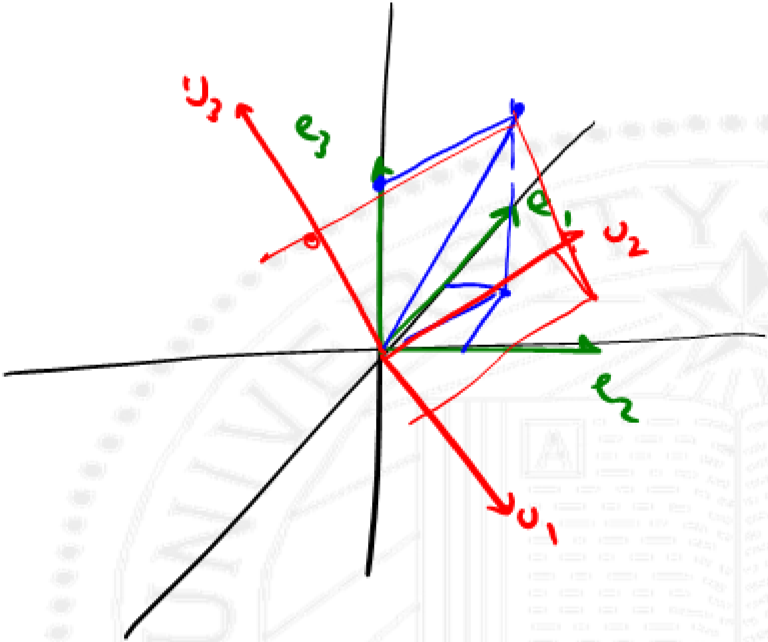
$n \times 1$

$$x = \sum_{i=1}^{n} (u_i^Tx)u_i$$

u_i^Tx component of x in the direction of u_i

$a = U^Tx$ expands x into its u_i components

$x = Ua$ reconstructs x from its u_i components



$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

$U^T x$

$\downarrow$

coordinate $\begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}$

Geometric Interpretation

$\mathbb{R}^3 \longrightarrow$ Rotation Matrix — SO(3).

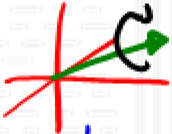
U orthogonal $\in \mathbb{R}^{n\times n}$ $\quad w = Uz$.

$\|w\| = \|z\| \Rightarrow w^T w = (Uz)^T Uz = z^T U^T U z = z^T z$.

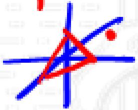
$\angle(w,\tilde{w}) = \angle(z,\tilde{z})$

explanation

Rotation about some axis



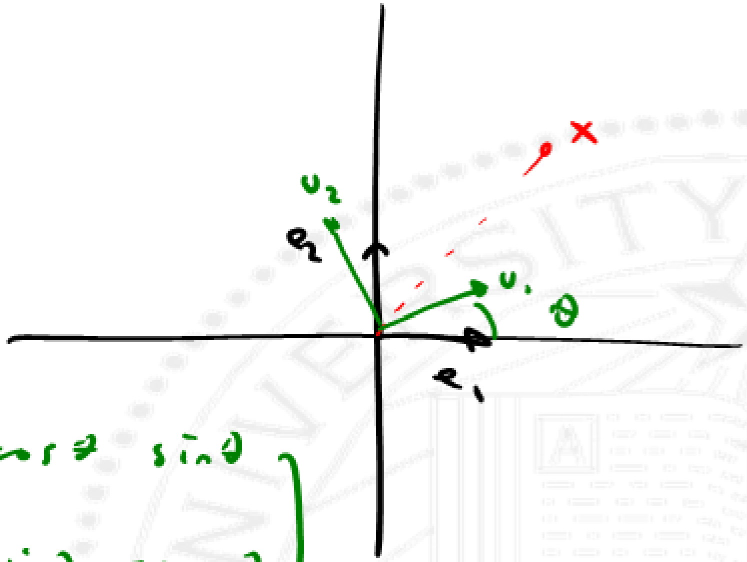
Reflection across a plane



$y = R_\theta x$

$\mathbb{R}^3$

$$\left[\begin{array}{cc|c} c & s & 0 \\ s & -c & 0 \\ \hline 0 & 0 & 1 \end{array}\right]$$



$$R_\theta = \begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix}$$

Examples: Rotation/Reflection

Quaternions

$$q = \begin{bmatrix} q_0 \\ \cdots \\ q_1 \\ q_2 \\ q_3 \end{bmatrix} \begin{array}{l} \leftarrow \sin\frac{\theta}{2} \\ \\ \left.\begin{array}{l} \\ \\ \\ \end{array}\right\} \hat{e}\cos\left(\frac{\theta}{2}\right) \end{array}$$

$$\dot{x} = Ax$$

$$q^T q = \underbrace{(\sin^2 + \cos^2)}_{1}\underbrace{e^T e}_{1} = 1$$

$$\dot{q} = -\frac{1}{2}\Omega q$$

$$\downarrow$$

$$\begin{bmatrix} \omega^T \\ \omega & (\omega\times) \end{bmatrix}$$

Gram-Schmidt Procedure (1.3)

Given a set of independent vectors $\{a_1 \ldots a_k\} \in \mathbb{R}^n$

G-S finds $\{q_1 \ldots q_k\}$ such that:

span $\{a_1 \ldots a_r\}$ = span $\{q_1 \ldots q_r\}$ $r \le k$.

$\{q_1 \ldots q_r\}$ orthonormal basis for $\{a_1 \ldots a_r\}$

$A \in \mathbb{R}^{n \times k} \longrightarrow R(A) = R(Q)$ ~200 years old

Gram-Schmidt Procedure (2.3)

$\tilde{q}_1 := a_1$

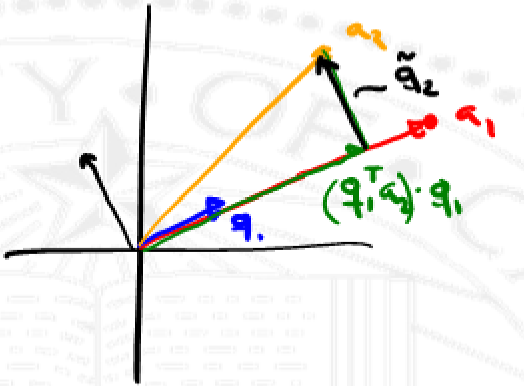
$q_1 = \tilde{q}_1 / \|\tilde{q}_1\|$ ← normalize

$\tilde{q}_2 := a_2 - (q_1^T a_2) \cdot q_1$

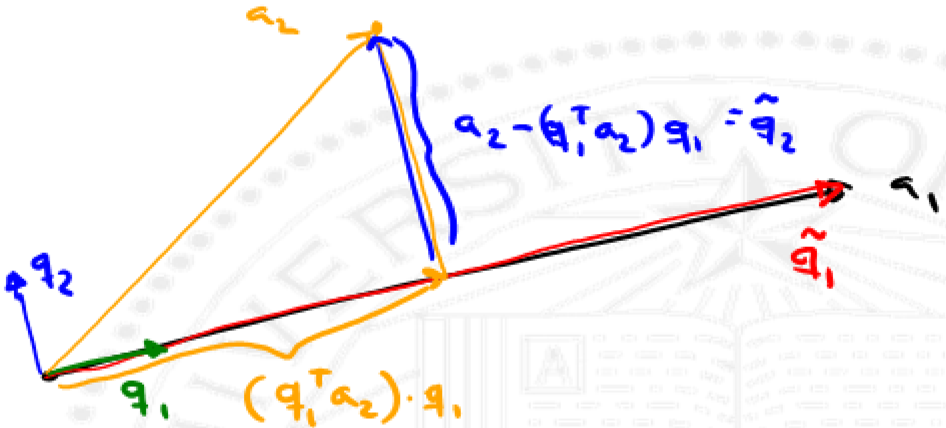
$q_2 = \tilde{q}_2 / \|\tilde{q}_2\|$

$\tilde{q}_3 := a_3 - (q_1^T a_3) q_1 - (q_2^T a_3) q_2$

$q_3 = \tilde{q}_3 / \|\tilde{q}_3\|$



Gram-Schmidt Procedure (3.3)



$$a_i = \underbrace{(q_1^T a_i)}_{r_{1i}} q_1 + \underbrace{(q_2^T a_i)}_{r_{2i}} q_2 + \dots + \underbrace{(q_{i-1}^T a_i)}_{r_{i-1,i}} q_{i-1} + \underbrace{(q_i^T a_i)}_{r_{ii}} q_i$$

QR decomposition

$A \in \mathbb{R}^{n \times k}$ $A = QR$ $Q \in \mathbb{R}^{n \times k}$ $R \in \mathbb{R}^{k \times k}$

$$\underbrace{\begin{bmatrix} a_1 \cdots a_k \end{bmatrix}}_{A} = \underbrace{\begin{bmatrix} q_1 \cdots q_k \end{bmatrix}}_{Q} \underbrace{\begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1k} \\ 0 & r_{22} & \cdots & r_{2k} \\ & & \ddots & \\ & & & r_{kk} \end{bmatrix}}_{R}$$

upper triangular
+ on diagonal
R^{-1} exists

orthonormal basis
for span(A)
$Q^T Q = I_k$
$\mathcal{R}(Q) = \mathcal{R}(A)$

General Gram-Schmidt Procedure

$\{a_1 \ldots a_k\}$ are dependent $\quad \tilde{q}_j = 0$, for some j.

a_j is linearly <u>dependent</u> on $\{a_1 \ldots a_{j-1}\}$.

$r = 0$

for $i = 1 \ldots k$ {

$\quad \tilde{a} = a_i - \sum_{j=1}^{i} q_j (q_j^T a_i)$

$\quad$ if ($\tilde{a}$ > epsilon)

$\quad\quad r$++;

$\quad\quad q_i = \tilde{a} / \|\tilde{a}\|$

$[q_1 \cdots q_r]$ basis for $\mathcal{R}(A)$

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

General Gram-Schmidt Procedure

$\cancel{A = QR}$

$A = Q\left[\tilde{R} \,\vdots\, S\right]P \qquad Q^TQ = I_{r\times r}$

$\tilde{R}$ - upper triangular & $\tilde{R}^{-1}$ exists

P - permutation matrix

$AP = Q\tilde{R}$

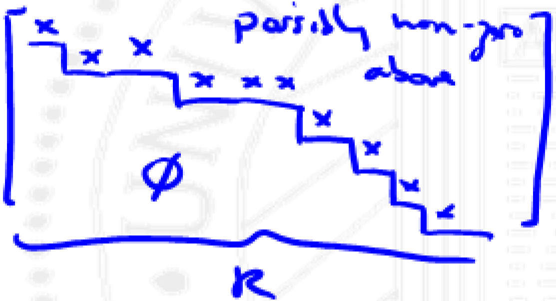
Rank Revealing QR factorization

General Gram-Schmidt Procedure

$[q_1 \cdots q_r]$ orthonormal basis for $\mathcal{R}(A)$

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

$A = QR \qquad Q^T Q = I_{r \times r} \qquad R \in \mathbb{R}^{r \times k} \qquad \text{rank}(R) = r.$



Applications of G-S

- directly yields an orthonormal basis for $\mathcal{R}(A)$
- yields a factorization for $A = BC$ $\quad B \in \mathbb{R}^{n \times r}$ $\quad C \in \mathbb{R}^{r \times k}$

$y = Ax$

$$x \xrightarrow{n} \boxed{C} \xrightarrow{r} \boxed{B} \xrightarrow{k} y$$

given a new vector, $b \in \mathbb{R}^n$, check if it is in the span of $\{a_1 \ldots a_k\}$

$b \in \text{span}\{a_1 \ldots a_k\}$, is b in $\mathcal{R}(A)$? $\quad GS \rightarrow [a_1 \ldots a_k \; b]$

GS works incrementally on $[a_1 \cdots a_p]$ $p \le k$.

$$[a_1 \cdots a_p] = [q_1 \cdots q_p] R_p$$

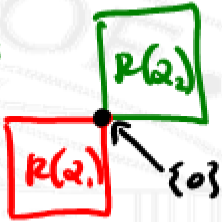
s — rank of $[a_1 \cdots a_p]$

R_p is the leading $s \times p$ sub-block of R.

"Full" *QR* Factorization (1.2)

extend any orthogonal basis to cover all of $\mathbb{R}^n$.

$\mathcal{R}(Q_1) \perp \mathcal{R}(Q_2)$ complementary subspaces



$$\langle Q_1 u, Q_2 v \rangle = \phi$$

$$\mathcal{R}(Q_1) \overset{\perp}{+} \mathcal{R}(Q_2) = \mathbb{R}^n$$

$$\mathcal{R}(Q_2) = \mathcal{R}(Q_1)^\perp \qquad \mathcal{R}(Q_1) = \mathcal{R}(Q_2)^\perp$$

"Full" *QR* Factorization (2.2)

$$A = Q_1 R_1 = \begin{bmatrix} Q_1 & \vdots & Q_2 \end{bmatrix} \begin{bmatrix} R_1 \\ \cdots \\ 0 \end{bmatrix}$$

upper triangular

non-unique

$[Q_1 | Q_2]$ square, orthogonal, spans $\mathbb{R}^n$

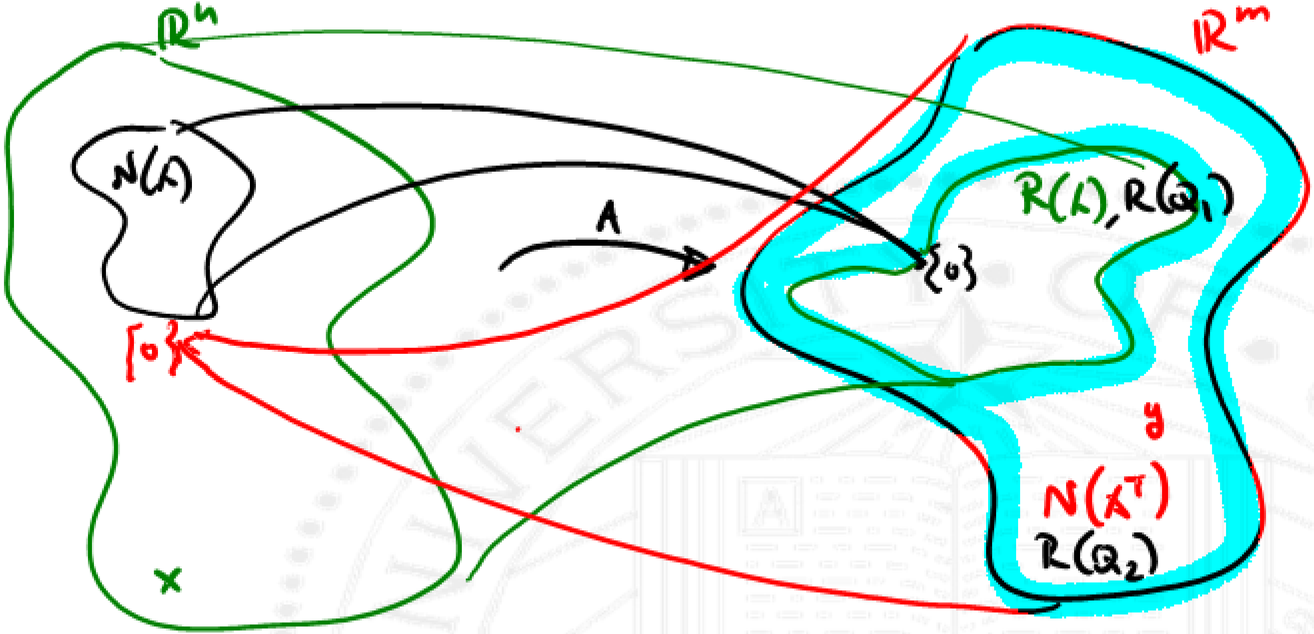
$R(Q_1) = R(A)$, columns of $Q_2 \in \mathbb{R}^{n\times(n-r)}$ $\perp Q_1$

any matrix $[A \mid \tilde{A}]$ that is full rank $(\tilde{A} = I)$

$[A \mid I] \longrightarrow$ full GS $\longrightarrow Q_1$

Q_1 orthonormal basis built from GS of A.

Q_2 orthonormal basis built from GS of I.



$$y = Ax$$

Orthogonal Decomposition Induced by A

$$A = \begin{bmatrix} Q_1 & \vdots & Q_2 \end{bmatrix} \begin{bmatrix} R_1 \\ \cdots \\ 0 \end{bmatrix} \qquad A^T = \begin{bmatrix} R_1^T & \vdots & 0 \end{bmatrix} \begin{bmatrix} Q_1^T \\ Q_2^T \end{bmatrix}$$

$z \in N(A^T) \longrightarrow Q_1^T z = 0. \quad z \in R(Q_2)$

columns of Q_2 are an orthonormal basis for $N(A^T)$

$$0 = A^T z = \begin{bmatrix} R_1^T \mid 0 \end{bmatrix} \begin{bmatrix} Q_1^T \\ Q_2^T \end{bmatrix} z = \begin{bmatrix} R_1^T \mid 0 \end{bmatrix} \begin{bmatrix} Q_1^T z \\ \cdots \\ Q_2^T z \end{bmatrix} = R_1^T Q_1^T z.$$

$$(R_1^T)^{-1} 0 = (R_1^T)^{-1} A^T z = (R_1^T)^{-1} R_1^T Q_1^T z = Q_1^T z.$$

Orthogonal Decomposition Induced by *A*

$R(Q_2) = N(A^T)$ columns of Q_2 are a basis for $R(A^T)$

$R(A) \perp N(A^T)$ perpendicular subspace.

$R(Q_1) \overset{\perp}{+} R(Q_2) \in \mathbb{R}^n$

$R(A)^\perp = N(A^T)$ $\quad$ $N(A)^\perp = R(A)$

$$\boxed{\mathcal{R}(A) \overset{\perp}{+} \mathcal{N}(A^T) = \mathbb{R}^n}$$

orthogonal decomposition induced by A.

$y = Ax$

$$\boxed{\mathcal{R}(Q_1) + \mathcal{R}(Q_2) = \mathbb{R}^n}$$

Questions?