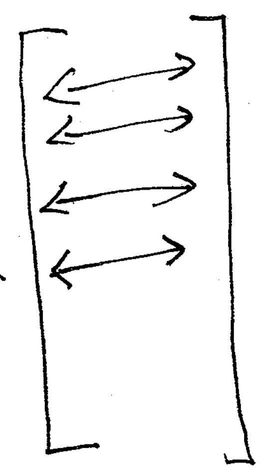
QR or Gram-Schmidt orthogonalization

LADS 23

(1)

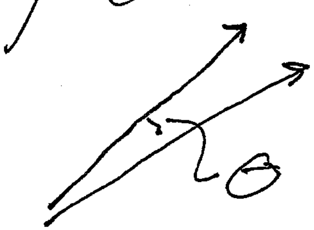
n

$B =$ m 

$m > n$

$col(B) = range(B)$ is always importants

we want a nice, stable basis for $col(B)$. As it ~~stands~~ stands, many column could be close to each other



θ small

We want an orthonormal basis for $\text{col}(B)$, but even more, the orthonormal basis $\{\vec{q}_1, \dots, \vec{q}_n\}$ should be

such that

(1) $\vec{q}_1$ is a basis for $\text{span}(\vec{b}_1)$

(2) $\{\vec{q}_1, \vec{q}_2\}$ is a basis for $\text{span}(\vec{b}_1, \vec{b}_2)$

$\vdots$

(3) $\{\vec{q}_1, \dots, \vec{q}_r\}$ is an o.n. basis for $\text{span}(\vec{b}_1, \vec{b}_2, \dots, \vec{b}_r)$.

In terms of equations

$$\vec{b}_1 = r_{11}\vec{q}_1$$
$$\vec{b}_2 = r_{12}\vec{q}_1 + r_{22}\vec{q}_2 \qquad (\#)$$
$$\vdots$$
$$\vec{b}_k = r_{1k}\vec{q}_1 + \dots + r_{kk}\vec{q}_k$$
$$\vdots$$

As a matrix

$$\underbrace{\begin{bmatrix} \vec{b}_1 & \cdots & \vec{b}_n \end{bmatrix}}_{B} = \underbrace{\begin{bmatrix} \vec{q}_1 & \cdots & \vec{q}_n \end{bmatrix}}_{\hat{Q}} \underbrace{\begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ & r_{22} & & \vdots \\ 0 & & \ddots & \\ & & & r_{nn} \end{bmatrix}}_{\hat{R}}$$

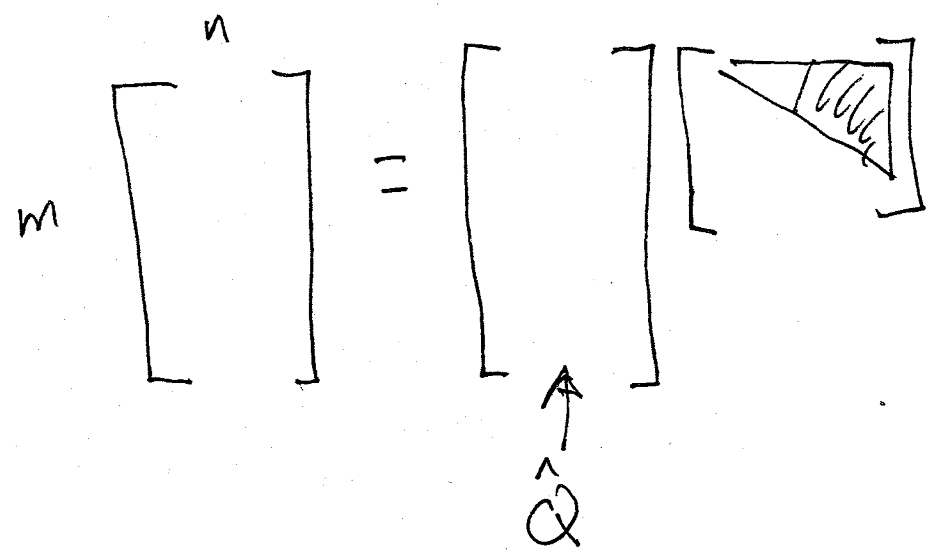
The thin or reduced QR decomposition writes

$$B = \hat{Q}\,\hat{R}$$

$\hat{R}$: $n \times n$, upper triang, $R_{ii} > 0$

$\hat{Q}$: has n orthonormal columns and is $m \times n$

when B has full rank $= n$.



NOTE: $\hat{Q}$ has o.n. col but is not square so not an orthogonal matrix

But $\hat{Q}^T\hat{Q} = I$ ← $\hat{Q}^T$ is left inverse

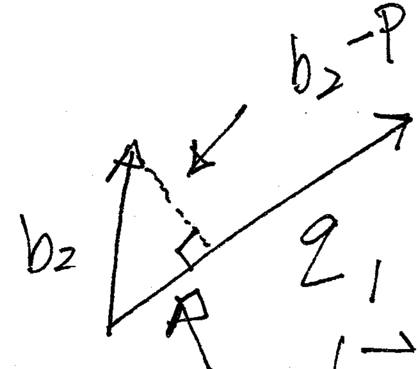
$$\underset{n \times m}{\begin{bmatrix} \longrightarrow q_1^T \longrightarrow \\ \longrightarrow \\ \longrightarrow \end{bmatrix}} \underset{m \times n}{\begin{bmatrix} | & | & | \\ | & | & | \\ | & | & | \end{bmatrix}} = n \overset{n}{\begin{bmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{bmatrix}}$$

WARNING: $\hat{Q}\hat{Q}^T \neq I$ usually

How do we compute Q and R?

Gram-Schmidt

$$\vec{q}_1 = \frac{\vec{b}_1}{\|\vec{b}_1\|_2}$$



$$\vec{p} = \underbrace{(\vec{b}_2 \cdot \vec{q}_1)}_{\text{scalar}} \vec{q}_1$$

Let $\vec{V}_2 = \vec{b}_2 - (\vec{b}_2 \cdot \vec{q}_1)\vec{q}_1$

Check $\vec{V}_2 \perp \vec{q}_1$? $\vec{V}_2 \cdot \vec{q}_1 =$

$$\vec{b}_2 \cdot \vec{q}_1 - (\vec{b}_2 \cdot \vec{q}_1)\underbrace{(\vec{q}_1 \cdot \vec{q}_1)}_{1} = 0$$

Let $\vec{q}_2 = \dfrac{\vec{V}_2}{\|\vec{V}_2\|_2}$

$\vdots$

$$\vec{v}_r = \vec{b}_r - (\vec{b}_r \cdot \vec{q}_1)\vec{q}_1 - \cdots - (\vec{b}_r \cdot \vec{q}_{r-1})\vec{q}_{r-1}$$

$$\vec{q}_r = \frac{\vec{v}_r}{\|\vec{v}_r\|}.$$

What are the R_{ij}?

$$B = \hat{Q}\hat{R}$$

$$\hat{Q}^T B = \hat{Q}^T\hat{Q}R = R$$

$$R_{ij} = \vec{q}_i \cdot \vec{b}_j$$ [note when $j < i$, this is zero by orthogonality]

8

<u>Another method:</u> Solve equation (#) for q_i will give a formula for R_{ij}

$$\vec{q}_1 = \frac{\vec{b}_1}{r_{11}}$$

$$\vec{q}_2 = \frac{\vec{b}_2 - r_{12}\vec{q}_1}{r_{22}}$$

$\vdots$

$$q_k = \frac{\vec{b}_k - \sum_{i=1}^{k-1} r_{ik}\vec{q}_i}{r_{kk}}$$

$\vdots$

$r_{ij} = \vec{q}_i \cdot \vec{b}_j$, $j > i$

Compare to Gram-Schmidt

$$|r_{kk}| = \| \vec{b}_k - \sum^{k-1} r_{ik}\vec{q}_i \|_2$$

Pseudo code for Classical Gram Schmidt

for $j = 1$ to n

$\quad \vec{v}_j = \vec{b}_j$

$\quad$ for $i = 1$ to $j-1$

$\quad\quad r_{ij} = \vec{q}_i \cdot \vec{b}_i$

$\quad\quad \vec{v}_j = \vec{v}_j - r_{ij}\vec{q}_i$

$\quad$ end

$\quad r_{jj} = \|\vec{v}_j\|_2$

$\quad \vec{q}_j = \vec{v}_j / r_{jj}$

end

Note that this computes both Q and R at the same time.

<u>NOTE</u>: Classical G.S. is unstable due to Round off error so there is an altered version that is stable.

Application to the normal equations.

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

$$\text{Cond}(A) = \frac{\sigma_1}{\sigma_n} \leftarrow 10^3$$

$$\text{Cond}(A^T A) = \frac{\sigma_1^2}{\sigma_n^2} \leftarrow 10^6.$$

So AVOID using $A^T A$ when possible

Two methods to avoid $A^T A$

1. SVD
2. QR decomp

We rewrite normal eq using QR assuming $\text{rank}(A) = n$.

$$A = \hat{Q}\hat{R}$$

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

$$\hat{R}^T \hat{Q}^T \hat{Q} \hat{R} \hat{x} = \hat{R}^T \hat{Q}^T \vec{b}$$

mult on left by $(\hat{R}^T)^{-1}$ and $\hat{Q}^T \hat{Q} = I$

$$\hat{R}\hat{x} = \hat{Q}^T \vec{b}$$

Method

(1) compute $A = \hat{Q}\hat{R}$

(2) compute $\hat{Q}^T \vec{b}$

(3) Solve $\hat{R}\hat{x} = \hat{Q}^T \vec{b}$ by back subs

upper triang.

Example next Time.