THE GRAM-SCHMIDT ALGORITHM AND QR [6.4 + 6.5]

Orthogonality – The Gram-Schmidt algorithm

1. Two vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ are orthogonal if $\boldsymbol{u}.\boldsymbol{v} = \boldsymbol{0}$.
2. They are orthonormal if in addition $\|\boldsymbol{u}\| = \|\boldsymbol{v}\| = \boldsymbol{1}$
3. In this case the matrix $\boldsymbol{Q} = [\boldsymbol{u}, \boldsymbol{v}]$ is such

$$\boldsymbol{Q^T Q = I}$$

- We say that the system $\{\boldsymbol{u}, \boldsymbol{v}\}$ is orthonormal ..
- .. and that the matrix $\boldsymbol{Q}$ has orthonormal columns
- .. or that it is orthogonal [Text reserves this term to $\boldsymbol{n} \times \boldsymbol{n}$ case]

Example: An orthonormal system $\{u, v\}$

$$u = \frac{1}{2}\begin{pmatrix} 1 \\ -1 \\ 1 \\ 1 \end{pmatrix} \quad v = \frac{1}{2}\begin{pmatrix} 1 \\ 1 \\ -1 \\ 1 \end{pmatrix}$$

Generalization: (to n vectors)

➤ A system of vectors $\{v_1, \ldots, v_n\}$ is orthogonal if $v_i.v_j = 0$ for $i \neq j$; and orthonormal if in addition $\|v_i\| = 1$ for $i = 1, \cdots, n$

- A matrix is orthogonal if its columns are orthonormal
- Then: $V = [v_1, \ldots, v_n]$ has orthonormal columns

[Note: The term 'orthonormal matrix' is not used. 'orthogonal' is often used for square matrices only (textbook)]

Question: We are given the set $\{u_1, u_2, \cdots, u_n\}$ which is not orthogonal. How do we get a set of vectors $\{q_1, q_2, \cdots, q_n\}$ that is orthonormal and spans the same subspace as $\{u_1, u_2, \cdots, u_n\}$?

Rationale: Orthonormal systems are easier to use.

Answer: Gram-Schmidt process - to be described next.

✍ See section 6.4 of text – example 1 with 2 vectors.

The Gram-Schmidt algorithm

Problem: Given a set $\{u_1, u_2\}$ how can we generate another set $\{q_1, q_2\}$ from linear combinations of u_1, u_2 so that $\{q_1, q_2\}$ is orthonormal?

Step 1 Define first vector: $q_1 = u_1/\|u_1\|$ ('Normalization')

Step 2: Orthogonalize u_2 against q_1: $\hat{q} = u_2 - (u_2.q_1)\ q_1$

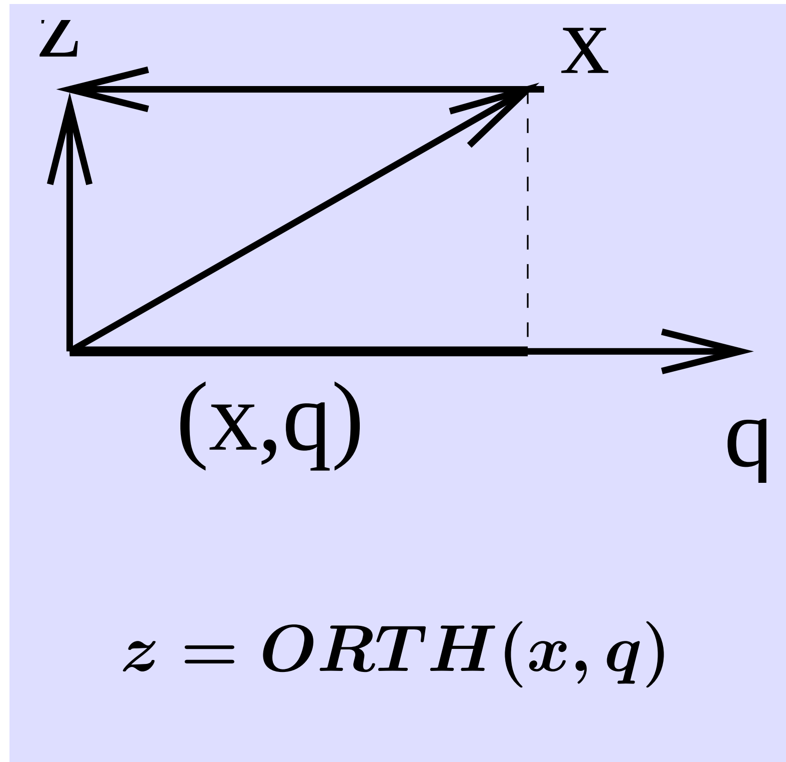
Step 3 Normalize to get second vector: $q_2 = \hat{q}/\|\hat{q}\|$

➤ Result: $\{q_1, q_2\}$ is an orthonormal set of vectors which spans the same space as $\{u_1, u_2\}$.

➤ The operations in step 2 can be written as

$$\hat{q} := ORTH(u_2, q_1)$$

ORTH (u_2, q_1): "orthogonalize u_2 against q_1"



➤ $ORTH(x, q)$ denotes the operation of orthogonalizing a vector x against a unit vector q.

$$ORTH(x, q) = x - (x.q)q$$

Example: $\boldsymbol{u_1 = \begin{pmatrix} 1 \\ -1 \\ 1 \\ 1 \end{pmatrix}}$ $\boldsymbol{u_2 = \begin{pmatrix} 2 \\ 0 \\ 0 \\ 2 \end{pmatrix}}$

Step 1: $\boldsymbol{q_1 = \frac{1}{2}\begin{pmatrix} 1 \\ -1 \\ 1 \\ 1 \end{pmatrix}}$

Step 2: First compute $\boldsymbol{u_2 . q_1 = ... = 2}$. Then:

$$\boldsymbol{\hat{q} = \begin{pmatrix} 2 \\ 0 \\ 0 \\ 2 \end{pmatrix} - 2 \times \frac{1}{2}\begin{pmatrix} 1 \\ -1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ -1 \\ 1 \end{pmatrix}}$$

Step 3: Normalize $\boldsymbol{q_2 = \frac{1}{2}\begin{pmatrix} 1 \\ 1 \\ -1 \\ 1 \end{pmatrix}}$

Generalization: 3 vectors

- How to generalize to 3 or more vectors?
- For 3 vectors : $[u_1, u_2, u_3]$
  - First 2 steps are the same $\rightarrow q_1, q_2$
  - Then orthogonalize u_3 against q_1 and q_2:

$$\hat{q} = u_3 - (u_3.q_1)q_1 - (u_3.q_2)q_2$$

  - Finally, normalize: $$q_3 = \hat{q}/\|\hat{q}\|$$

General problem: Given $U = [u_1, \ldots, u_n]$, compute $Q = [q_1, \ldots, q_n]$ which is orthonormal and s.t. $\text{Col}(Q) = \text{Col}(U)$.

ALGORITHM : 1. *Classical Gram-Schmidt*

1. *For* $j = 1 : n$ *Do:*
2. $\quad \hat{q} = u_j$
3. $\quad$ *For* $i = 1 : j - 1$
4. $\quad\quad \hat{q} := \hat{q} - (u_j.q_i)q_i$ $\qquad$ / *set* $r_{ij} = (u_j.q_i)$
5. $\quad$ *End*
6. $\quad q_j := \hat{q}/\|\hat{q}\|$ $\qquad$ / *set* $r_{jj} = \|\hat{q}\|$
7. *End*

➤ All n steps can be completed iff $u_1, u_2, \ldots, u_n$ are linearly independent.

➤ Define a matrix R as follows

$$r_{ij} = \begin{cases} u_j.q_i & \text{if} \quad i < j \text{ (see line 4)} \\ \|\hat{q}\| & \text{if} \quad i = j \text{ (see line 6)} \\ 0 & \text{if} \quad i > j \text{ (lower part)} \end{cases}$$

➤ We have from the algorithm: (For $j = 1, 2, \cdots, n$)

$$u_j = r_{1j} q_1 + r_{2j} q_2 + \ldots + r_{jj} q_j$$

➤ If $U = [u_1, u_2, \ldots, u_n]$, $Q = [q_1, q_2, \ldots, q_n]$, and if R is the $n \times n$ upper triangular matrix defined above:

$$R = \{r_{ij}\}_{i,j=1,\ldots,n}$$

then the above relation can be written as

$$U = QR$$

➤ This is called the QR factorization of U.

- Q has orthonormal columns. It satisfies:

$$Q^TQ = I$$

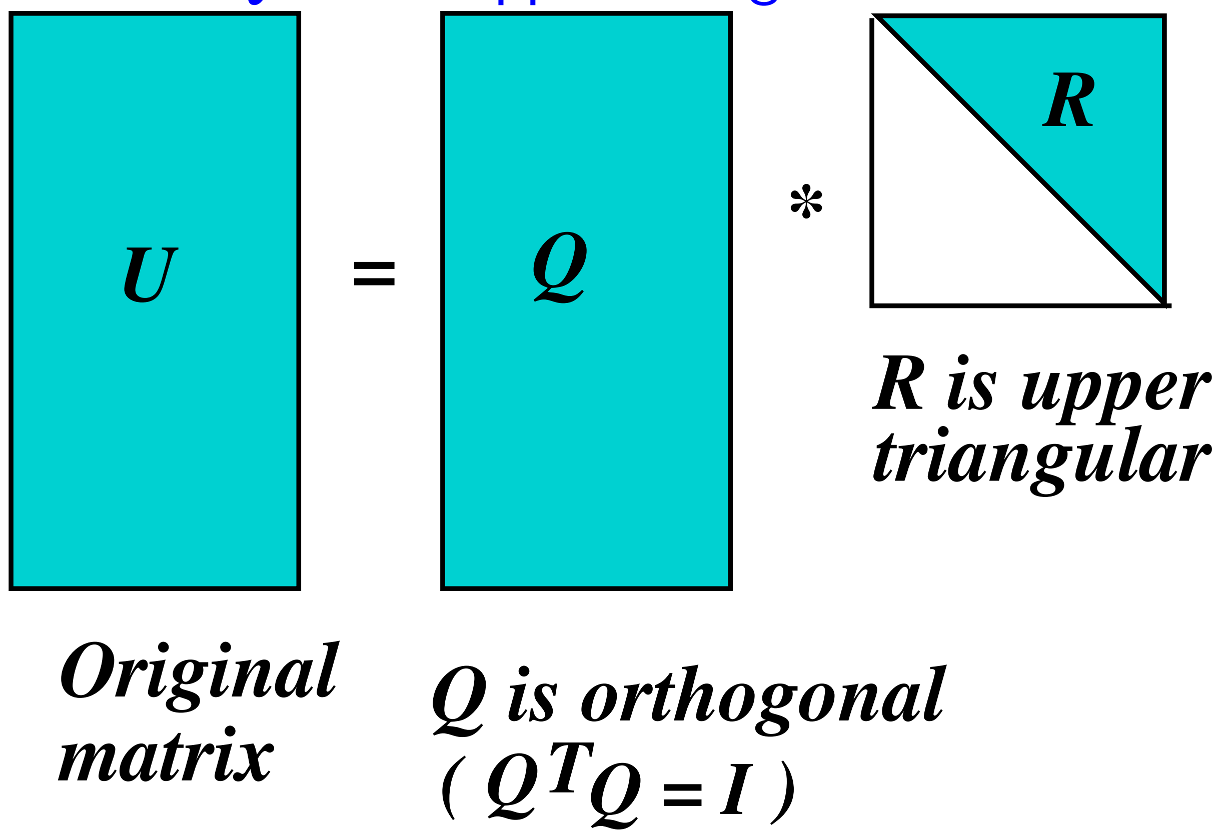
- It is said to be orthogonal
- R is upper triangular

✍ What is the inverse of an orthogonal $n \times n$ matrix?

✍ Show that when $U \in \mathbb{R}^{m\times n}$ the total cost of Gram-Schmidt is $\approx 2mn^2$.

Another decomposition:

A matrix U, with linearly independent columns, is the product of an orthogonal matrix Q and a upper triangular matrix R.



✍ Orthonormalize the system of vectors:

$$U = [u_1, u_2, u_3] = \begin{pmatrix} 1 & -4 & 3 \\ -1 & 2 & -1 \\ 1 & 0 & 1 \\ 1 & -2 & -1 \end{pmatrix}$$

For this example:

✍ 1) what is Q? what is R?

✍ 2) Verify (matlab) that $U = QR$

✍ 3) Compute Q^TQ. [Result should be the identity matrix]

Solution: [values for $\boldsymbol{R}$ are in red]

Step 1: $q_1 = \frac{u_1}{\|u_1\|} = \frac{1}{2}\begin{bmatrix} 1 \\ -1 \\ 1 \\ 1 \end{bmatrix}$ $r_{11} = \|u_1\| = 2$

Step 2: $\hat{q}_2 = u_2 - (u_2.q_1)q_1 \quad \rightarrow$

$$\hat{q}_2 = \begin{bmatrix} -4 \\ 2 \\ 0 \\ -2 \end{bmatrix} - \frac{-8}{2} \times \frac{1}{2}\begin{bmatrix} 1 \\ -1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -2 \\ 0 \\ 2 \\ 2 \end{bmatrix} \qquad r_{12} = \frac{-8}{2} = -4$$

$$\rightarrow \quad q_2 = \frac{\hat{q}_2}{\|\hat{q}_2\|} = \frac{1}{\sqrt{8}}\begin{bmatrix} -2 \\ 0 \\ 2 \\ 0 \end{bmatrix} = \frac{1}{\sqrt{2}}\begin{bmatrix} -1 \\ 0 \\ 1 \\ 0 \end{bmatrix} \qquad r_{22} = \sqrt{8}$$

Step 3: $\hat{q}_3 = u_3 - (u_3.q_1)q_1 - (u_3.q_2)q_2 \rightarrow$

$$\hat{q}_3 = \begin{bmatrix} 3 \\ -1 \\ 1 \\ -1 \end{bmatrix} - \frac{4}{2} \times \frac{1}{2} \begin{bmatrix} 1 \\ -1 \\ 1 \\ 1 \end{bmatrix} - \frac{-2}{\sqrt{2}} \times \frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 1 \\ -2 \end{bmatrix}$$

$$q_3 = \frac{1}{\sqrt{6}} \begin{bmatrix} 1 \\ 0 \\ 1 \\ -2 \end{bmatrix} \quad r_{13} = 2; \quad r_{23} = -\sqrt{2}; \quad r_{33} = \sqrt{6}$$

$$Q = \begin{bmatrix} 1/2 & -1/\sqrt{2} & 1/\sqrt{6} \\ -1/2 & 0 & 0 \\ 1/2 & 1/\sqrt{2} & 1/\sqrt{6} \\ 1/2 & 0 & -2/\sqrt{6} \end{bmatrix} \qquad R = \begin{bmatrix} 2 & -4 & 2 \\ 0 & \sqrt{8} & -\sqrt{2} \\ 0 & 0 & \sqrt{6} \end{bmatrix}$$

Solving LS systems via QR factorization

➤ In practice: not a good idea to solve the system $A^TAx = A^Tb$. Use the QR factorization instead. How?

➤ Answer in the form of an exercise

Problem: $Ax \approx b$ in least-squares sense

A is an $m \times n$ (full-rank) matrix. Consider the QR factorization of A

$$A = QR$$

✍ Approach 1: Write the normal equations – then 'simplify'

✍ Approach 2: Write the condition $b - Ax \perp \text{Col}(A)$ and recall that A and Q have the same column space.

✍ Total cost?