Numerical Methods

Course # CSE330

Lecture #: 9.3: QR-Decomposition

Prepared by

Abu Mohammad Khan

- In the previous lecture, we learned how to solve an over determined system using the normal matrix $A^T A$ which is a square matrix.
- The matrix A is of order $m \times n$ with $m \geq n$. But the normal matrix $A^T A$ is invertible, and hence this gives us a solution.
- But whenever the normal matrix is ill-conditioned, then it becomes difficult or sometimes impossible to solve the normal matrix. In that case, we factor the matrix A into a product of matrices, known as the QR-decomposition.
- Theorem (QR Decomposition): This theorem states that any real $m \times n$ matrix A, with $m \geq n$, can be written in the form, $A = QR$, where Q is an $m \times n$ matrix with orthonormal columns and R is an upper triangular $n \times n$ matrix.

$$A_{m\times n} = Q_{m\times n} R_{n\times n} \Rightarrow Q = \{\underbrace{q_1 | q_2 | \cdots | q_n}_{\text{Orthonormal columns}}\}$$

$$\boxed{A_{m\times n} x_{n\times 1} = b_{m\times 1}} \Rightarrow \text{Solve } x$$

BRAC UNIVERSITY
Inspiring Excellence

- To understand the QR-decomposition, we recall the Gram-Schmidt orthogonalization process which is a mathematical process to obtain an orthonormal basis or set from a set of linearly independent vectors.
- Each of these linearly independent vectors is an $m \times 1$ column vectors.
- Let's consider an $m \times n$ matrix A:

$$A = \begin{array}{cccc} u_1 & u_2 & & u_n \\ \multicolumn{4}{c}{\underbrace{\begin{pmatrix} u_{11} & u_{12} & \dots & u_{1n} \\ u_{21} & u_{22} & \cdots & u_{2n} \\ u_{31} & u_{32} & \cdots & u_{3n} \\ \vdots & \vdots & \ddots & \vdots \\ u_{m1} & u_{m2} & \cdots & u_{mn} \end{pmatrix}}_{m \times n}} \end{array} \equiv \begin{array}{c} (\,u_1 \quad u_2 \quad \cdots \quad u_n\,) \\ \downarrow \\ u_{11} \\ u_{21} \end{array}$$

- In the above, at the right-hand side, each vector u_i for $i = 1,2,\cdots n$ is column vector. In matrix form, we may write:

$$u_i = (u_{1i} \quad u_{2i} \quad \cdots \quad u_{mi})^T = m \times 1 \text{ column matrix.}$$

BRAC UNIVERSITY
Inspiring Excellence

- Therefore, we may think of the matrix A as a set of column vectors u_i.
- These column vectors, u_i, are linearly independent. Therefore, from this set of linearly independent vectors, we can construct an orthonormal set of vectors by using the Gram-Schmidt process.
- Let $\{q_1, q_2, \cdots, q_n\}$ be the set of orthonormal vectors constructed form the set ~~A~~ $= \{u_1, u_2, \cdots, u_n\}$ which is a set of n vectors in $\mathbb{R}^m$.
- The Gram-Schmidt process yields that

$u_k^T q_i = (1 \times n)(n \times 1) = 1 \times 1 =$ Scale

$p_1 = u_1$

$p_2 = u_1 - (u_k^T q_i) q_i$

$$\boxed{p_k = u_k - \sum_{i=1}^{k-1} \underbrace{\left(u_k^T\, q_i\right)}_{\text{number}} q_i \quad \text{and}} \qquad q_k = \frac{p_k}{|p_k|} = \text{Unit vector}$$

where $k = 1,2,\cdots,n$. And the norm is defined as: $|p_k| \equiv \sqrt{(p_k \cdot p_k)}$.

- It should be noted that each p_k is constructed from u_k by subtracting projections of u_k on each of the previous q_i for $i < k$.

BRAC UNIVERSITY
Inspiring Excellence

$u_1 = 3\times 1$

- Example: Let $u_1 = (3,6,0)^T$ and $u_2 = (1,2,2)^T$. Let's find the orthonormal vectors by using Gram-Schmidt process.

- Step-1: We identify that

$k=1$

$i=0$

$$p_1 = u_1 = \begin{pmatrix} 3 \\ 6 \\ 0 \end{pmatrix} \Longrightarrow |p_1| = \sqrt{p_1^T\, p_1} = \sqrt{3^2 + 6^2 + 0} = \sqrt{45}$$

$$\therefore\ q_1 \equiv \frac{p_1}{|p_1|} = \frac{1}{\sqrt{45}} \begin{pmatrix} 3 \\ 6 \\ 0 \end{pmatrix}$$

→ 1st ~~othon~~ normal vector

- Step-2: We now construct p_2 form u_2 and then normalize p_2 to obtain q_2:

$k=2 \Rightarrow i=1$

$$p_2 = u_2 - \left(u_2^T\, q_1\right) q_1$$

But, we have,

$$\left(u_2^T\, q_1\right) = \underbrace{\overset{(1\times 3)}{\begin{pmatrix} 1 & 2 & 2 \end{pmatrix}}}_{u_2^T} \underbrace{\overset{(3\times 1)}{\begin{pmatrix} 3 \\ 6 \\ 0 \end{pmatrix}}}_{q_1} \frac{1}{\sqrt{45}} = \frac{15}{\sqrt{45}} = \cancel{1\times 1}\ 1\times 1 \text{ matrix}$$

- Therefore, we obtain,

$$p_2 = \underbrace{\begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}}_{u_2} - \frac{\overbrace{15}^{(u_2^T q_1)}}{(\sqrt{45})^2} \overbrace{\begin{pmatrix} 3 \\ 6 \\ 0 \end{pmatrix}}^{q_1} = \begin{pmatrix} 0 \\ 0 \\ 2 \end{pmatrix}$$

- Clearly: $|p_2| = 2 \implies \boxed{q_2 = \frac{p_2}{|p_2|} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}}.$

$\underbrace{\{u_1, u_2\}} \rightarrow \underbrace{\{q_1, q_2\}}$ same number of elements

- Finally, the set $\{q_1, q_2\}$ is an orthonormal set constructed from the linearly independent set of two vectors $\{u_1, u_2\}$.

- The above can be checked by computing: $q_1 \cdot q_2 = 0$ and $|q_1| = |q_2| = 1$. In other words, these two vectors satisfies the following relation:

$$\boxed{q_i \cdot q_j = \delta_{ij} = \begin{cases} 1; & \text{if } i = j \\ 0; & \text{if } i \neq j \end{cases}}$$

unit norm
orthogonal
orthonormal

- The above process will repeatedly be applied if we have more vectors.
- The above process is used to construct the QR-decomposition of the matrix A.
- Here, the $m \times n$ matrix is written as a set of n linearly independent vectors u_i for $i = 1,2,\cdots,\cancel{m}$. n
- Each u_i is an $m \times 1$ column vector. Hence: $A = (\, u_1 \,|u_2\,|\cdots|\, u_n \,)$ which are now vectors in $\mathbb{R}^m$.
- Applying the Gram-Schmidt process will produce a set of orthogonal vectors $q_i \in \mathbb{R}^m$.
- Since, $q_k = \frac{p_k}{|p_k|} \Longrightarrow p_k = |p_k| q_k$. The Gram-Schmidt process gives,

$$
\begin{aligned}
p_1 &= |p_1| q_1 = u_1 \\
p_2 &= |p_2| q_2 = u_2 - (\, u_2^T \, q_1 \,)\, q_1 \\
&\;\vdots \\
p_k &= |p_k| q_k = u_k - \sum_{i=1}^{k-1} (\, u_k^T \, q_i \,)\, q_i
\end{aligned}
$$

- Rearranging the last term, we find:

$$u_k = |p_k| q_k + \sum_{i=1}^{k-1} (\, u_k^T \, q_i \,) \, q_i$$

- Taking the inner product with q_k of the above equation and using the orthonormality of the q_i's, we find;

$$u_k \cdot q_k = |p_k| \underbrace{q_k \cdot q_k}_{=1} + \sum_{i=1}^{k-1} (\, u_k^T \, q_i \,) \underbrace{q_i \cdot q_k}_{=0} = |p_k|$$

- Therefore, we finally, write:

$$u_k = (u_k \cdot q_k) q_k + \sum_{i=1}^{k-1} (\, u_k^T \, q_i \,) \, q_i$$

$$\therefore \boxed{u_k = \sum_{i=1}^{k} (\, u_k^T \, q_i \,) \, q_i \; ; \;\; k = 1,2,\cdots, m.}$$

- Clearly, we see that:

$$u_1 = (u_1^T q_1) q_1 = (q_1)(u_1^T q_1)$$

$$u_2 = (u_2^T q_1) q_1 + (u_2^T q_2) q_2 = \boxed{\underbrace{(q_1 \quad q_2)}_{2\times 1} \begin{pmatrix} u_2^T q_1 \\ u_2^T q_2 \end{pmatrix}}$$

2×1

- Therefore, up to 2nd term, we obtain,

$$(u_1 \quad u_2) = \boxed{(q_1 \quad q_2) \begin{pmatrix} u_1^T q_1 & u_2^T q_1 \\ 0 & u_2^T q_2 \end{pmatrix}}$$

- The 3rd term is,

$$u_3 = (u_3^T q_1) q_1 + (u_3^T q_2) q_2 + (u_3^T q_3) q_3$$

$$= \boxed{(q_1 \quad q_2 \quad q_3) \begin{pmatrix} u_3^T q_1 \\ u_3^T q_2 \\ u_3^T q_3 \end{pmatrix}}$$

- Hence, up to 3rd term, we write,

$$\begin{pmatrix} u_1 & u_2 & u_3 \end{pmatrix} = \begin{pmatrix} q_1 & q_2 & q_3 \end{pmatrix} \begin{pmatrix} u_1^T q_1 & u_2^T q_1 & u_3^T q_1 \\ 0 & u_2^T q_2 & u_3^T q_2 \\ 0 & 0 & u_3^T q_3 \end{pmatrix}$$

- This way, we can write up to n-th term:

$$\underbrace{\left(u_1 | u_2 | \cdots | u_n \right)}_{A_{m\times n}} = \underbrace{\left(q_1 | q_2 | \cdots | q_n \right)}_{Q_{m\times n}} \underbrace{\begin{pmatrix} u_1^T q_1 & u_2^T q_1 & \cdots & u_n^T q_1 \\ 0 & u_2^T q_2 & \cdots & u_n^T q_2 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & u_n^T q_n \end{pmatrix}}_{R_{n\times n}}$$

- In compact form, we write,

$$\boxed{A = QR}$$

This is known as the QR-decomposition of the matrix A.