QR Decompositions of Matrices

QR decomposition of matrices

QR Decomposition: Given any matrix $A \in \mathbb{R}^{n\times m}$, $n \geq m$, there exists an orthogonal matrix $Q \in \mathbb{R}^{n\times n}$ and an upper triangular matrix $R \in \mathbb{R}^{n\times m}$ such that

$$A = QR. \tag{2}$$

The decomposition (2) is called a QR decomposition of A.
If $n > m$, then $R = \begin{bmatrix} R_1 \\ 0 \end{bmatrix}$ where $R_1 \in \mathbb{R}^{m\times m}$ is upper triangular.

In particular if $n = m$, then (2) takes the form $A = QR$ where R is a square upper triangular matrix.

If $A \in \mathbb{C}^{n\times m}$, $n \geq m$, then (2) holds with $\mathbb{R}$ replaced by $\mathbb{C}$, Q being a unitary matrix.

Condensed QR decomposition

Given $A \in \mathbb{R}^{n\times m}$ with $n > m$, if $A = QR$ be a QR decomposition of A with $R = \begin{bmatrix} R_1 \\ 0 \end{bmatrix}$, then partitioning $Q = [Q_1 \;\; Q_2]$ where $Q_1 \in \mathbb{R}^{n\times m}$, gives,

$$A = QR = [Q_1 \;\; Q_2] \begin{bmatrix} R_1 \\ 0 \end{bmatrix} = Q_1 R_1 .$$

This motivates the following theorem.

Theorem Given any $n \times m$ matrix A with $n > m$, there exists an isometry $Q \in \mathbb{R}^{n\times m}$ and an upper triangular matrix R such that

$$A = QR. \qquad ((3))$$

If $\operatorname{rank} A = m$, then R is nonsingular.

The decomposition in (3) is called a *condensed QR decomposition of A*.

Unitary/Orthogonal matrices

Unitary matrices are square matrices such that $Q^* = Q^{-1}$.
A real unitary matrix is called an orthogonal matrix.

Unitary/Orthogonal matrices

Unitary matrices are square matrices such that $Q^* = Q^{-1}$.
A real unitary matrix is called an orthogonal matrix.

Evidently, a square matrix Q is unitary (orthogonal) if and only if its columns form an orthonormal basis of $\mathbb{R}^n$ ($\mathbb{C}^n$).

Unitary/Orthogonal matrices

Unitary matrices are square matrices such that $Q^* = Q^{-1}$.
A real unitary matrix is called an orthogonal matrix.

Evidently, a square matrix Q is unitary (orthogonal) if and only if its columns form an orthonormal basis of $\mathbb{R}^n$ ($\mathbb{C}^n$).

They play a very important role in matrix computations due to their nice properties. In the following let Q be an $n \times n$ unitary matrix.

Unitary/Orthogonal matrices

Unitary matrices are square matrices such that $Q^* = Q^{-1}$.
A real unitary matrix is called an orthogonal matrix.

Evidently, a square matrix Q is unitary (orthogonal) if and only if its columns form an orthonormal basis of $\mathbb{R}^n$ ($\mathbb{C}^n$).

They play a very important role in matrix computations due to their nice properties. In the following let Q be an $n \times n$ unitary matrix.

- $\langle Qx, Qy \rangle = \langle x, y \rangle$ for all $x, y \in \mathbb{C}^n$.
- $\|Qx\|_2 = \|x\|_2$.
- $\|QB\|_2 = \|B\|_2$ for any $B \in \mathbb{C}^{n \times m}$.
- $\|Q\|_2 = 1$ and $\|Q\|_F = \sqrt{n}$.
- $\kappa_2(Q) = 1$.
- $Q^* A Q$ is Hermitian if A is Hermitian.
- If A is real symmetric and Q is orthogonal, then $Q^T A Q$ is also real symmetric.
- In the presence of rounding errors, $fl(QA) = Q(A + E)$ where $\|E\|_2 / \|A\|_2$ is $O(u)$.

Unitary/Orthogonal matrices

Unitary matrices are square matrices such that $Q^* = Q^{-1}$.
A real unitary matrix is called an orthogonal matrix.

Evidently, a square matrix Q is unitary (orthogonal) if and only if its columns form an orthonormal basis of $\mathbb{R}^n$ ($\mathbb{C}^n$).

They play a very important role in matrix computations due to their nice properties. In the following let Q be an $n \times n$ unitary matrix.

- $\langle Qx, Qy \rangle = \langle x, y \rangle$ for all $x, y \in \mathbb{C}^n$.
- $\|Qx\|_2 = \|x\|_2$.
- $\|QB\|_2 = \|B\|_2$ for any $B \in C^{n \times m}$.
- $\|Q\|_2 = 1$ and $\|Q\|_F = \sqrt{n}$.
- $\kappa_2(Q) = 1$. Prove these properties!
- $Q^* AQ$ is Hermitian if A is Hermitian.
- If A is real symmetric and Q is orthogonal, then $Q^T AQ$ is also real symmetric.
- In the presence of rounding errors, $fl(QA) = Q(A + E)$ where $\|E\|_2/\|A\|_2$ is $O(u)$.

Isometry

A matrix $Q \in \mathbb{C}^{n\times m}$, or $\mathbb{R}^{n\times m}$ with $n > m$, is said to be an isometry if $Q^*Q = I_m$.

Isometry

A matrix $Q \in \mathbb{C}^{n\times m}$, or $\mathbb{R}^{n\times m}$ with $n > m$, is said to be an isometry if $Q^*Q = I_m$.

Clearly a matrix Q with more rows than columns is an isometry if and only if its columns form an orthonormal set in $\mathbb{R}^n$ or $\mathbb{C}^n$.

Isometry

A matrix $Q \in \mathbb{C}^{n\times m}$, or $\mathbb{R}^{n\times m}$ with $n > m$, is said to be an isometry if $Q^*Q = I_m$.

Clearly a matrix Q with more rows than columns is an isometry if and only if its columns form an orthonormal set in $\mathbb{R}^n$ or $\mathbb{C}^n$.

Isometries have properties very similar to that of unitary matrices. Given an $n \times m$ isometry $Q = [q_1 \ \cdots q_m]$,

- $\langle Qx, Qy\rangle = \langle x, y\rangle$ for all $x, y \in \mathbb{C}^n$.
- $\|Qx\|_2 = \|x\|_2$.
- $\|QB\|_2 = \|B\|_2$ for any $B \in \mathbb{C}^{n\times m}$.
- $\|Q\|_2 = 1$ and $\|Q\|_F = \sqrt{m}$.
- $\kappa_2(Q) = 1$.
- In the presence of rounding errors, $fl(QA) = Q(A + E)$ where $\|E\|_2/\|A\|_2$ is $O(u)$.
- QQ^* is the orthogonal projection onto $\mathrm{span}\,\{q_1, \ldots, q_m\}$, that is, $QQ^*v = v$ for all $v \in \mathrm{span}\,\{q_1, \ldots, q_m\}$ and $QQ^*w = 0$ for all $w \in \{q_1, \ldots, q_m\}^{\perp}$. Prove this!

Classical Gram Schmidt Orthonormalisation

Let $\{v_1, \ldots, v_m\}$ be an ordered set of linearly independent vectors in $\mathbb{R}^n$. The Classical Gram Schmidt (CGS) process finds an ordered orthonormal set of vectors $\{q_1, \ldots, q_m\}$ in $\mathbb{R}^n$ such that

$$\mathrm{span}\{v_1, \ldots, v_k\} = \mathrm{span}\{q_1, \ldots, q_k\}, \;\; k = 1, \ldots, m.$$

Classical Gram Schmidt Orthonormalisation

Let $\{v_1, \ldots, v_m\}$ be an ordered set of linearly independent vectors in $\mathbb{R}^n$. The Classical Gram Schmidt (CGS) process finds an ordered orthonormal set of vectors $\{q_1, \ldots, q_m\}$ in $\mathbb{R}^n$ such that

$$\mathrm{span}\{v_1, \ldots, v_k\} = \mathrm{span}\{q_1, \ldots, q_k\}, \; k = 1, \ldots, m.$$

Classical Gram Schmidt (CGS):

Step 1: $q_1 := v_1 / \|v_1\|_2$.

Step 2: $q_2 := \underbrace{(v_2 - (v_2^T q_1) q_1)}_{=:\hat{q}_2} / \|v_2 - (v_2^T q_1) q_1\|_2$.

Step k: Assuming that $q_1, \ldots, q_{k-1}$ are calculated as above,

$$q_k = \underbrace{(v_k - \sum_{i=1}^{k-1} (v_k^T q_i) q_i)}_{=:\hat{q}_k} / \|v_k - \Sigma_{i=1}^{k-1} (v_k^T q_i) q_i\|_2.$$

Classical Gram Schmidt Orthonormalisation

Let $\{v_1, \ldots, v_m\}$ be an ordered set of linearly independent vectors in $\mathbb{R}^n$. The Classical Gram Schmidt (CGS) process finds an ordered orthonormal set of vectors $\{q_1, \ldots, q_m\}$ in $\mathbb{R}^n$ such that

$$\operatorname{span}\{v_1, \ldots, v_k\} = \operatorname{span}\{q_1, \ldots, q_k\}, \;\; k = 1, \ldots, m.$$

Classical Gram Schmidt (CGS):

Step 1: $q_1 := v_1 / \|v_1\|_2$.

Step 2: $q_2 := \underbrace{(v_2 - (v_2^T q_1)q_1)}_{=:\hat{q}_2} / \|v_2 - (v_2^T q_1)q_1\|_2$.

Step k: Assuming that $q_1, \ldots, q_{k-1}$ are calculated as above,

$$q_k = \underbrace{(v_k - \sum_{i=1}^{k-1}(v_k^T q_i)q_i)}_{=:\hat{q}_k} / \|v_k - \Sigma_{i=1}^{k-1}(v_k^T q_i)q_i\|_2.$$

Exercise: Show that CGS applied to the basis $\{e_1 + e_2, e_2, e_2 + e_3\}$ in $\mathbb{R}^3$ produces the ordered orthonormal basis

$$\{(e_1 + e_2)/\sqrt{2}, (e_2 - e_1)/\sqrt{2}, e_3\}.$$

Equivalence of CGS and condensed QR decomposition

CGS $\equiv$ condensed QR

Supose $\{v_1, \ldots, v_m\}$ is an ordered linearly independent subset of $\mathbb{R}^n$ and $\{q_1, \ldots, q_m\}$ is the output of CGS on $\{v_1, \ldots, v_m\}$.

CGS $\equiv$ condensed QR

Supose $\{v_1, \dots, v_m\}$ is an ordered linearly independent subset of $\mathbb{R}^n$ and $\{q_1, \dots, q_m\}$ is the output of CGS on $\{v_1, \dots, v_m\}$. Then,

$$\underbrace{[v_1 \cdots v_m]}_{=:V} = \underbrace{[q_1 \cdots q_m]}_{=:Q} \underbrace{\begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1m} \\ & r_{22} & \cdots & r_{2m} \\ & & \ddots & \vdots \\ & & & r_{mm} \end{bmatrix}}_{=:R}$$

where $r_{ij} = v_j^T q_i$ for $j > i$, $r_{jj} = \|\hat{q}_j\|_2$ and $r_{ij} = 0$ otherwise. Clearly, $V = QR$ is a condensed QR decomposition of V.

CGS $\equiv$ condensed QR

Supose $\{v_1, \ldots, v_m\}$ is an ordered linearly independent subset of $\mathbb{R}^n$ and $\{q_1, \ldots, q_m\}$ is the output of CGS on $\{v_1, \ldots, v_m\}$. Then,

$$\underbrace{[v_1 \cdots v_m]}_{=:V} = \underbrace{[q_1 \cdots q_m]}_{=:Q} \underbrace{\begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1m} \\ & r_{22} & \cdots & r_{2m} \\ & & \ddots & \vdots \\ & & & r_{mm} \end{bmatrix}}_{=:R}$$

where $r_{ij} = v_j^T q_i$ for $j > i$, $r_{jj} = \|\hat{q}_j\|_2$ and $r_{ij} = 0$ otherwise. Clearly, $V = QR$ is a condensed QR decomposition of V.

Exercise: Conversely if $V = QR$ be a condensed QR decomposition of $V = [v_1 \cdots v_m] \in \mathbb{R}^{n\times m}$ where $R = [r_{ij}]_{m\times m}$ with $r_{ii} > 0$ for all $i = 1, \ldots, m$, then the columns $q_1, \ldots, q_m$ of Q are equal to those obtained via CGS on the columns of V with

$$r_{ij} = \begin{cases} v_j^T q_i, & i < j, \\ \|\hat{q}_j\|_2, & i = j \\ 0, & \text{otherwise.} \end{cases}$$