QR Decomposition with Gram-Schmidt

The QR decomposition (also called the QR factorization) of a matrix is a decomposition of the matrix into an orthogonal matrix and a triangular matrix. A QR decomposition of a real square matrix A is a decomposition of A as

$$A = QR,$$

where Q is an orthogonal matrix (i.e. $Q^T Q = I$) and R is an upper triangular matrix. If A is nonsingular, then this factorization is unique.

There are several methods for actually computing the QR decomposition. One of such method is the Gram-Schmidt process.

1 Gram-Schmidt process

Consider the GramSchmidt procedure, with the vectors to be considered in the process as columns of the matrix A. That is,

$$A = \left[\mathbf{a}_1 \mid \mathbf{a}_2 \mid \cdots \mid \mathbf{a}_n \right].$$

Then,

$$\begin{aligned}
\mathbf{u}_1 &= \mathbf{a}_1, \qquad \mathbf{e}_1 = \frac{\mathbf{u}_1}{||\mathbf{u}_1||}, \\
\mathbf{u}_2 &= \mathbf{a}_2 - (\mathbf{a}_2 \cdot \mathbf{e}_1)\mathbf{e}_1, \qquad \mathbf{e}_2 = \frac{\mathbf{u}_2}{||\mathbf{u}_2||}. \\
\mathbf{u}_{k+1} &= \mathbf{a}_{k+1} - (\mathbf{a}_{k+1} \cdot \mathbf{e}_1)\mathbf{e}_1 - \cdots - (\mathbf{a}_{k+1} \cdot \mathbf{e}_k)\mathbf{e}_k, \qquad \mathbf{e}_{k+1} = \frac{\mathbf{u}_{k+1}}{||\mathbf{u}_{k+1}||}.
\end{aligned}$$

Note that $|| \cdot ||$ is the L_2 norm.

1.1 QR Factorization

The resulting QR factorization is

$$A = \left[\mathbf{a}_1 \mid \mathbf{a}_2 \mid \cdots \mid \mathbf{a}_n \right] = \left[\mathbf{e}_1 \mid \mathbf{e}_2 \mid \cdots \mid \mathbf{e}_n \right] \begin{bmatrix} \mathbf{a}_1 \cdot \mathbf{e}_1 & \mathbf{a}_2 \cdot \mathbf{e}_1 & \cdots & \mathbf{a}_n \cdot \mathbf{e}_1 \\ 0 & \mathbf{a}_2 \cdot \mathbf{e}_2 & \cdots & \mathbf{a}_n \cdot \mathbf{e}_2 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbf{a}_n \cdot \mathbf{e}_n \end{bmatrix} = QR.$$

Note that once we find $\mathbf{e}_1, \ldots, \mathbf{e}_n$, it is not hard to write the QR factorization.

2 Example

Consider the matrix

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix},$$

with the vectors $\mathbf{a}_1 = (1,1,0)^T,\ \mathbf{a}_2 = (1,0,1)^T,\ \mathbf{a}_3 = (0,1,1)^T$.
Note that all the vectors considered above and below are column vectors. From now on, I will drop ^{T} notation for simplicity, but we have to remember that all the vectors are column vectors.
Performing the Gram-Schmidt procedure, we obtain:

$$\begin{aligned}
\mathbf{u}_1 &= \mathbf{a}_1 = (1,1,0), \\
\mathbf{e}_1 &= \frac{\mathbf{u}_1}{||\mathbf{u}_1||} = \frac{1}{\sqrt{2}}(1,1,0) = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\right), \\
\mathbf{u}_2 &= \mathbf{a}_2 - (\mathbf{a}_2 \cdot \mathbf{e}_1)\mathbf{e}_1 = (1,0,1) - \frac{1}{\sqrt{2}}\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\right) = \left(\frac{1}{2}, -\frac{1}{2}, 1\right), \\
\mathbf{e}_2 &= \frac{\mathbf{u}_2}{||\mathbf{u}_2||} = \frac{1}{\sqrt{3/2}}\left(\frac{1}{2}, -\frac{1}{2}, 1\right) = \left(\frac{1}{\sqrt{6}}, -\frac{1}{\sqrt{6}}, \frac{2}{\sqrt{6}}\right), \\
\mathbf{u}_3 &= \mathbf{a}_3 - (\mathbf{a}_3 \cdot \mathbf{e}_1)\mathbf{e}_1 - (\mathbf{a}_3 \cdot \mathbf{e}_2)\mathbf{e}_2 \\
&= (0,1,1) - \frac{1}{\sqrt{2}}\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\right) - \frac{1}{\sqrt{6}}\left(\frac{1}{\sqrt{6}}, -\frac{1}{\sqrt{6}}, \frac{2}{\sqrt{6}}\right) = \left(-\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right), \\
\mathbf{e}_3 &= \frac{\mathbf{u}_3}{||\mathbf{u}_3||} = \left(-\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right).
\end{aligned}$$

Thus,

$$\begin{aligned}
Q &= \left[\ \mathbf{e}_1 \mid \mathbf{e}_2 \mid \cdots \mid \mathbf{e}_n\ \right] = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ 0 & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{3}} \end{bmatrix}, \\
R &= \begin{bmatrix} \mathbf{a}_1 \cdot \mathbf{e}_1 & \mathbf{a}_2 \cdot \mathbf{e}_1 & \mathbf{a}_3 \cdot \mathbf{e}_1 \\ 0 & \mathbf{a}_2 \cdot \mathbf{e}_2 & \mathbf{a}_3 \cdot \mathbf{e}_2 \\ 0 & 0 & \mathbf{a}_3 \cdot \mathbf{e}_3 \end{bmatrix} = \begin{bmatrix} \frac{2}{\sqrt{2}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 0 & \frac{3}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ 0 & 0 & \frac{2}{\sqrt{3}} \end{bmatrix}.
\end{aligned}$$

Method 1: QR **method for eigenvalues - simultaneous iteration**

Instead of repeating the iterations for each of the eigen vectors we start with a complete set of orthogonal vectors and successively multiply with $\mathbb{A}$ and orthogonalize them again.

Start:Orthonormal matrix $\mathbb{Q}_0$
For $j = 1 \rightarrow M$
$\mathbb{Z} = \mathbb{A}\mathbb{Q}_{j-1}$ (Mult)
$\mathbb{Z} \rightarrow \mathbb{Q}_j \mathbb{R}_j$ (QR decomp)
$\mathbb{A}_j = \mathbb{Q}_j^T \mathbb{A} \mathbb{Q}_j$ (Similarity)
End For

We get $\mathbb{D} \approx \mathbb{A}_M$ and $\mathbb{V} \approx \mathbb{Q}_M$ for large M

Method 2: QR **method for eigenvalues**

There is another algorithm that uses smaller number of operations and is very similar to the simultaneous iteration.

$\mathbb{A}_0 = \mathbb{A}$ (Set value)
$\mathbb{V}_0 = \mathbb{I}$ (Set value)
For $j = 1 \rightarrow M$
$\mathbb{A}_{j-1} \rightarrow \mathbb{Q}_j \mathbb{R}_j$ (QR decomp)
$\mathbb{A}_j = \mathbb{R}_j \mathbb{Q}_j$ (Mult)
$\mathbb{V}_j = \mathbb{V}_{j-1} \mathbb{Q}_j$ (Mult)
End For

We get $\mathbb{D} \approx \mathbb{A}_M$ and $\mathbb{V} \approx \mathbb{V}_M$ for large M

Note that :

$$\begin{aligned} \mathbb{A}_j &= \mathbb{R}_j \mathbb{Q}_j \\ &= \mathbb{Q}_j^T \mathbb{Q}_j \mathbb{R}_j \mathbb{Q}_j \\ &= \mathbb{Q}_j^T \mathbb{A}_{j-1} \mathbb{Q}_j \end{aligned}$$