QR Factorization of a Matrix A

Let A be an $m \times n$ matrix, and suppose that A has rank n. That is, the columns of A are linearly independent. The QR factorization of A factors A into the product:

$$A = QR\,,$$

where Q is an $m \times n$ matrix whose columns form an orthonormal set of vectors in $\mathbb{R}^m$, and R is an $n \times n$ upper triangle matrix all of whose diagonal entries are positive.

The proof that such a factorization is possible depends on the Gram-Schmidt algorithm. So let's review it. Given a set $\{\vec{x}_i\}_{i=1}^k$ of linearly independent vectors we can construct a set $\{\vec{u}_i\}_{i=1}^k$ of vectors with the following properties:

1. $\{\vec{u}_i\}_{i=1}^k$ is an orthonormal set of vectors.

2. For each j the spans of $\{\vec{x}_i\}_{i=1}^j$ and $\{\vec{u}_i\}_{i=1}^j$ are the same, and there is a positive constant λ_j such that

   $$\vec{x}_j = \lambda_j \vec{u}_j + \sum_{i=1}^{j-1} c_i \vec{u}_i\,,$$

   where there is no second summand when $j = 1$, and the c_j are some scalars, which depend on j.

Now let $\vec{c}_i$ for $1 \le i \le n$ denote the columns of A. By assumption this set is linearly independent. Let $\{\vec{u}_i\}_{i=1}^n$ be the orthonormal set constructed from them via the Gram-Schmidt algorithm. The above two properties imply the following set of equations:

$$\begin{aligned}
\vec{c}_1 &= \langle \vec{c}_1, \vec{u}_1 \rangle \vec{u}_1 \text{ with } \langle \vec{c}_1, \vec{u}_1 \rangle > 0 \\
\vec{c}_2 &= \langle \vec{c}_2, \vec{u}_2 \rangle \vec{u}_2 + \langle \vec{c}_2, \vec{u}_1 \rangle \vec{u}_1 \text{ with } \langle \vec{c}_2, \vec{u}_2 \rangle > 0 \\
&\vdots \\
\vec{c}_j &= \langle \vec{c}_j, \vec{u}_j \rangle \vec{u}_j + \sum_{i=1}^{j-1} \langle \vec{c}_j, \vec{u}_i \rangle \vec{u}_i \text{ with } \langle \vec{c}_j, \vec{u}_j \rangle > 0 \\
&\quad \text{for } 1 \le j \le n.
\end{aligned}$$

Let Q be that $m \times n$ matrix whose j^{th} column is the vector $\vec{u}_j$, let R be the $n \times n$ upper triangle matrix whose j^{th} column equals the vector

$$\begin{bmatrix} \langle \vec{c}_j, \vec{u}_1 \rangle \\ \langle \vec{c}_j, \vec{u}_2 \rangle \\ \vdots \\ \langle \vec{c}_j, \vec{u}_j \rangle \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Thus,

$$\begin{aligned} A &= \begin{bmatrix} \vec{c}_1 & \vec{c}_2 & \cdots & \vec{c}_n \\ \downarrow & \downarrow & \cdots & \downarrow \end{bmatrix} \\ &= \begin{bmatrix} \vec{q}_1 & \vec{q}_2 & \cdots & \vec{q}_n \\ \downarrow & \downarrow & \cdots & \downarrow \end{bmatrix} \begin{bmatrix} \langle \vec{c}_1, \vec{u}_1 \rangle & \langle \vec{c}_2, \vec{u}_1 \rangle & \langle \vec{c}_3, \vec{u}_1 \rangle & \cdots & \langle \vec{c}_n, \vec{u}_1 \rangle \\ 0 & \langle \vec{c}_2, \vec{u}_2 \rangle & \langle \vec{c}_3, \vec{u}_2 \rangle & \cdots & \langle \vec{c}_2, \vec{u}_n \rangle \\ 0 & 0 & \langle \vec{c}_3, \vec{u}_3 \rangle & \cdots & \langle \vec{c}_3, \vec{u}_n \rangle \\ \downarrow & \downarrow & \downarrow & \cdots & \downarrow \\ 0 & 0 & 0 & \cdots & \langle \vec{c}_n, \vec{u}_n \rangle \end{bmatrix} \\ &= QR \end{aligned}$$

The following gives a useful result about matrices like Q.

Theorem Let Q be an $m \times n$ matrix whose columns form an orthonormal collection of vectors. Then

$$Q^T Q = I_n .$$

Proof The i, j entry of the matrix $Q^T Q$ is the scalar product of the i^{th} row of Q^T with the j^{th} column of Q. However, the i^{th} row of Q^T is the i^{th} column of Q. Thus, the i, j entry of the product is the inner product of two columns of Q. Since the columns of Q form an orthonormal set, the inner product is zero when $i \neq j$ and 1 when $i = j$. Thus, $Q^T Q = I_n$.