Computing the QR factorization

Let A be an $m \times n$ matrix with linearly independent columns. Applying the Gram-Schmidt orthogonalization process to the columns of A produces an $m \times n$ matrix Q whose columns are orthonormal. In fact, keeping track of all column operations on A yields a factorization $A = QR$, where R is an $n \times n$ upper triangular matrix with positive entries on the diagonal.

Example 1a: $A = \begin{bmatrix} -1 & 3 \\ 1 & 5 \end{bmatrix}$. Let us carry out the Gram-Schmidt process with the columns a_1, a_2.

$$\begin{aligned}
v_1 &= a_1 = \begin{bmatrix} -1 \\ 1 \end{bmatrix} \\
u_1 &= \frac{v_1}{\|v_1\|} = \frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 1 \end{bmatrix} \\
v_2 &= a_2 - \operatorname{proj}_{u_1} a_2 = a_2 - \langle u_1, a_2 \rangle u_1 \\
&= \begin{bmatrix} 3 \\ 5 \end{bmatrix} - \left(\frac{1}{\sqrt{2}}(-3+5) \right) \frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 5 \end{bmatrix} - \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 4 \end{bmatrix} \\
u_2 &= \frac{v_2}{\|v_2\|} = \frac{1}{4\sqrt{2}} \begin{bmatrix} 4 \\ 4 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}
\end{aligned}$$

Therefore the matrix Q is

$$Q = \begin{bmatrix} u_1 & u_2 \end{bmatrix} = \begin{bmatrix} -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix}.$$

That is how the Gram-Schmidt process produces the matrix Q. Here are two methods for finding R.

Method 1. Keeping track of column operations

At each step of Gram-Schmidt, the operations on the vectors correspond to column operations on A, which correspond to multiplying by elementary matrices on the right. Let us write all those matrices.

$$\begin{aligned}
A &= \begin{bmatrix} -1 & 3 \\ 1 & 5 \end{bmatrix} \\
A \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 \\ 0 & 1 \end{bmatrix} &= \begin{bmatrix} -\frac{1}{\sqrt{2}} & 3 \\ \frac{1}{\sqrt{2}} & 5 \end{bmatrix} \\
A \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -\sqrt{2} \\ 0 & 1 \end{bmatrix} &= \begin{bmatrix} -\frac{1}{\sqrt{2}} & 4 \\ \frac{1}{\sqrt{2}} & 4 \end{bmatrix} \\
A \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -\sqrt{2} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{4\sqrt{2}} \end{bmatrix} &= \begin{bmatrix} -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} = Q.
\end{aligned}$$

From this we obtain

$$\begin{aligned} A &= Q \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{4\sqrt{2}} \end{bmatrix}^{-1} \begin{bmatrix} 1 & -\sqrt{2} \\ 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 \\ 0 & 1 \end{bmatrix}^{-1} \\ &= Q \begin{bmatrix} 1 & 0 \\ 0 & 4\sqrt{2} \end{bmatrix} \begin{bmatrix} 1 & \sqrt{2} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \sqrt{2} & 0 \\ 0 & 1 \end{bmatrix} \\ &= Q \begin{bmatrix} \sqrt{2} & \sqrt{2} \\ 0 & 4\sqrt{2} \end{bmatrix} = QR. \end{aligned} \tag{1}$$

We have found

$$R = \begin{bmatrix} \sqrt{2} & \sqrt{2} \\ 0 & 4\sqrt{2} \end{bmatrix} = \sqrt{2} \begin{bmatrix} 1 & 1 \\ 0 & 4 \end{bmatrix}.$$

Let us check the factorization:

$$QR = \frac{1}{\sqrt{2}} \begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix} \sqrt{2} \begin{bmatrix} 1 & 1 \\ 0 & 4 \end{bmatrix} = \begin{bmatrix} -1 & 3 \\ 1 & 5 \end{bmatrix} = A.$$

Note: The numbers in the product of elementary matrices in (1) "pile up" nicely in the matrix R specifically because of the order in which the column operations are performed. Those elementary matrices **do not commute** at all, so order is important.

Think of the matrix R as undoing all the operations in the Gram-Schmidt algorithm.

- The (1,1) entry $\sqrt{2}$ is undoing $u_1 = \frac{1}{\sqrt{2}} a_1$.
- The (1,2) entry $\sqrt{2}$ is undoing $v_2 = a_2 - \sqrt{2} u_1$.
- The (2,2) entry $4\sqrt{2}$ is undoing $u_2 = \frac{1}{4\sqrt{2}} v_2$.

Again, order of operations is very important!

Let us rewrite the equalities using symbols, in order to obtain the general formula.

$$\begin{aligned} &A = \begin{bmatrix} a_1 & a_2 \end{bmatrix} = \begin{bmatrix} v_1 & a_2 \end{bmatrix} \\ &A \begin{bmatrix} \frac{1}{\|v_1\|} & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} \frac{v_1}{\|v_1\|} & a_2 \end{bmatrix} = \begin{bmatrix} u_1 & a_2 \end{bmatrix} \\ &A \begin{bmatrix} \frac{1}{\|v_1\|} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -\langle u_1, a_2 \rangle \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} u_1 & a_2 - \langle u_1, a_2 \rangle u_1 \end{bmatrix} = \begin{bmatrix} u_1 & v_2 \end{bmatrix} \\ &A \begin{bmatrix} \frac{1}{\|v_1\|} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -\langle u_1, a_2 \rangle \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{\|v_2\|} \end{bmatrix} = \begin{bmatrix} u_1 & \frac{v_2}{\|v_2\|} \end{bmatrix} = \begin{bmatrix} u_1 & u_2 \end{bmatrix} = Q. \end{aligned} \tag{2}$$

From this we obtain

$$\begin{aligned} A &= Q \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{\|v_2\|} \end{bmatrix}^{-1} \begin{bmatrix} 1 & -\langle u_1, a_2 \rangle \\ 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} \frac{1}{\|v_1\|} & 0 \\ 0 & 1 \end{bmatrix}^{-1} \\ &= Q \begin{bmatrix} 1 & 0 \\ 0 & \|v_2\| \end{bmatrix} \begin{bmatrix} 1 & \langle u_1, a_2 \rangle \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \|v_1\| & 0 \\ 0 & 1 \end{bmatrix} \\ &= Q \begin{bmatrix} \|v_1\| & \langle u_1, a_2 \rangle \\ 0 & \|v_2\| \end{bmatrix} = QR. \end{aligned}$$

This formula for R generalizes to any value of n. For example, if n were 3, R would be given by

$$R = \begin{bmatrix} \|v_1\| & \langle u_1, a_2 \rangle & \langle u_1, a_3 \rangle \\ 0 & \|v_2\| & \langle u_2, a_3 \rangle \\ 0 & 0 & \|v_3\| \end{bmatrix}. \tag{3}$$

Knowing this, there is no need to rewrite the whole computation (2) every time. We can just build the matrix R as we go along the Gram-Schmidt process, using (3).

Method 2. $R = Q^T A$

The fact that Q has orthonormal columns can be restated as $Q^T Q = I$. In particular, Q has a left inverse, namely Q^T. From this we can find R:

$$A = QR \Rightarrow Q^T A = Q^T QR = R.$$

In other words, the formula $R = Q^T A$ holds, no matter what m and n are. It doesn't matter if the matrix is square or not.

Example: In example 1a, we had $A = \begin{bmatrix} -1 & 3 \\ 1 & 5 \end{bmatrix}$ and $Q = \frac{1}{\sqrt{2}} \begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix}$.

$$R = Q^T A = \frac{1}{\sqrt{2}} \begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} -1 & 3 \\ 1 & 5 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 2 & 2 \\ 0 & 8 \end{bmatrix} = \sqrt{2} \begin{bmatrix} 1 & 1 \\ 0 & 4 \end{bmatrix}.$$

Finally, let us make sure the two methods agree. We have:

$$\langle u_k, a_k \rangle = \langle \frac{v_k}{\|v_k\|}, a_k \rangle = \frac{1}{\|v_k\|} \langle v_k, a_k \rangle = \frac{1}{\|v_k\|} \langle v_k, v_k \rangle = \|v_k\|$$

and hence formula (3) can be rewritten as

$$R = \begin{bmatrix} \|v_1\| & \langle u_1, a_2 \rangle & \langle u_1, a_3 \rangle \\ 0 & \|v_2\| & \langle u_2, a_3 \rangle \\ 0 & 0 & \|v_3\| \end{bmatrix} = \begin{bmatrix} \langle u_1, a_1 \rangle & \langle u_1, a_2 \rangle & \langle u_1, a_3 \rangle \\ 0 & \langle u_2, a_2 \rangle & \langle u_2, a_3 \rangle \\ 0 & 0 & \langle u_3, a_3 \rangle \end{bmatrix} = Q^T A.$$