LECTURE APPENDIX

Kyle Kloster

October 12, 2014

How unique is QR?

Full rank, $m = n$

In class we looked at the special case of full rank, $n \times n$ matrices, and showed that the QR decomposition is unique up to a factor of a diagonal matrix with entries ± 1. Here we'll see that the other full rank cases follow the $m = n$ case somewhat closely. Any full rank QR decomposition involves a square, upper-triangular partition $\boldsymbol{R}$ within the larger (possibly rectangular) $m \times n$ matrix. The gist of these uniqueness theorems is that $\boldsymbol{R}$ is unique, up to multiplication by a diagonal matrix of ± 1s; the extent to which the orthogonal matrix is unique depends on its dimensions.

Theorem $(m = n)$ If $\boldsymbol{A} = \boldsymbol{Q}_1 \boldsymbol{R}_1 = \boldsymbol{Q}_2 \boldsymbol{R}_2$ are two QR decompositions of full rank, square $\boldsymbol{A}$, then

$$\begin{aligned} \boldsymbol{Q}_2 &= \boldsymbol{Q}_1 \boldsymbol{S} \\ \boldsymbol{R}_2 &= \boldsymbol{S} \boldsymbol{R}_1 \end{aligned}$$

for some square diagonal $\boldsymbol{S}$ with entries ± 1. If we require the diagonal entries of $\boldsymbol{R}$ to be positive, then the decomposition is unique.

Theorem $(m < n)$ If $\boldsymbol{A} = \boldsymbol{Q}_1 \begin{bmatrix} \boldsymbol{R}_1 & \boldsymbol{N}_1 \end{bmatrix} = \boldsymbol{Q}_2 \begin{bmatrix} \boldsymbol{R}_2 & \boldsymbol{N}_2 \end{bmatrix}$ are two QR decompositions of a full rank, $m \times n$ matrix $\boldsymbol{A}$ with $m < n$, then

$$\boldsymbol{Q}_2 = \boldsymbol{Q}_1 \boldsymbol{S}, \qquad \boldsymbol{R}_2 = \boldsymbol{S} \boldsymbol{R}_1, \quad \text{and} \quad \boldsymbol{N}_2 = \boldsymbol{S} \boldsymbol{N}_1$$

for square diagonal $\boldsymbol{S}$ with entries ± 1. If we require the diagonal entries of $\boldsymbol{R}$ to be positive, then the decomposition is unique.

Theorem $(m > n)$ If $\boldsymbol{A} = \begin{bmatrix} \boldsymbol{Q}_1 & \boldsymbol{U}_1 \end{bmatrix} \begin{bmatrix} \boldsymbol{R}_1 \\ 0 \end{bmatrix} = \begin{bmatrix} \boldsymbol{Q}_2 & \boldsymbol{U}_2 \end{bmatrix} \begin{bmatrix} \boldsymbol{R}_2 \\ 0 \end{bmatrix}$ are two QR decompositions of a full rank, $m \times n$ matrix $\boldsymbol{A}$ with $m > n$, then

$$\boldsymbol{Q}_2 = \boldsymbol{Q}_1 \boldsymbol{S}, \qquad \boldsymbol{R}_2 = \boldsymbol{S} \boldsymbol{R}_1, \quad \text{and} \quad \boldsymbol{U}_2 = \boldsymbol{U}_1 \boldsymbol{T}$$

for square diagonal $\boldsymbol{S}$ with entries ± 1, and square orthogonal $\boldsymbol{T}$. If we require the diagonal entries of $\boldsymbol{R}$ to be positive, then $\boldsymbol{Q}$ and $\boldsymbol{R}$ are unique.

Proofs

Proof: $(m < n)$ Let $\boldsymbol{Q}_1 \begin{bmatrix} \boldsymbol{R}_1 & \boldsymbol{N}_1 \end{bmatrix} = \boldsymbol{Q}_2 \begin{bmatrix} \boldsymbol{R}_2 & \boldsymbol{N}_2 \end{bmatrix}$ with $\boldsymbol{Q}_i$ being $m \times m$ and orthogonal, $\boldsymbol{R}_i$ being $m \times m$ and upper triangular, and $\boldsymbol{N}_i$ being an arbitrary $m \times (n - m)$ matrix. Then multiplying through yields $\boldsymbol{Q}_1 \boldsymbol{R}_1 = \boldsymbol{Q}_2 \boldsymbol{R}_2$, two QR decompositions of a full rank, $m \times m$ matrix. Using the theorem above, we get that $\boldsymbol{Q}_2 = \boldsymbol{Q}_1 \boldsymbol{S}$ and $\boldsymbol{R}_2 = \boldsymbol{S} \boldsymbol{R}_1$ for a diagonal matrix $\boldsymbol{S}$ with entries ± 1. Looking at the right-most partition of the original product yields $\boldsymbol{Q}_1 \boldsymbol{N}_1 = \boldsymbol{Q}_2 \boldsymbol{N}_2$. But we've shown $\boldsymbol{Q}_2 = \boldsymbol{Q}_1 \boldsymbol{S}$, so now we have $\boldsymbol{Q}_1 \boldsymbol{N}_1 = \boldsymbol{Q}_1 \boldsymbol{S} \boldsymbol{N}_2$. Left-multiplying by $\boldsymbol{Q}_1^T$ and then by $\boldsymbol{S}$ then proves $\boldsymbol{N}_2 = \boldsymbol{S} \boldsymbol{N}_1$, completing the theorem.

Proof: $(m > n)$ Let $\boldsymbol{A}$ be full rank and $m \times n$ with $m > n$. Suppose it has decompositions

$$\boldsymbol{A} = \tilde{\boldsymbol{Q}}_1 \tilde{\boldsymbol{R}}_1 = \tilde{\boldsymbol{Q}}_2 \tilde{\boldsymbol{R}}_2$$

for $m \times m$ orthogonal matrices $\tilde{\boldsymbol{Q}}_i$, $m \times n$ and upper-triangular matrices $\tilde{\boldsymbol{R}}_i$. (We know we can do this because the QR decomposition always exists).

Since $m > n$, we can write $\tilde{\boldsymbol{Q}}_i = \begin{bmatrix} \boldsymbol{Q}_i & \boldsymbol{U}_i \end{bmatrix}$ and $\tilde{\boldsymbol{R}}_i = \begin{bmatrix} \boldsymbol{R}_i \\ 0 \end{bmatrix}$ where $\boldsymbol{Q}_i$ is $m \times n$ and $\boldsymbol{U}_i$ is $m \times (m-n)$. Then

$$\boldsymbol{A} = \tilde{\boldsymbol{Q}}_i \tilde{\boldsymbol{R}}_i = \begin{bmatrix} \boldsymbol{Q}_i & \boldsymbol{U}_i \end{bmatrix} \begin{bmatrix} \boldsymbol{R}_i \\ 0 \end{bmatrix} = \boldsymbol{Q}_i \boldsymbol{R}_i$$

where $\boldsymbol{R}_i$ is square, upper-triangular, invertible (because $\boldsymbol{A}$ is full rank), and the columns of $\boldsymbol{Q}_i$ are orthonormal so $\boldsymbol{Q}_i$ satisfies $\boldsymbol{Q}_i^T \boldsymbol{Q}_i = \boldsymbol{I}$.

Then we have

$$\boldsymbol{Q}_1 \boldsymbol{R}_1 = \boldsymbol{Q}_2 \boldsymbol{R}_2, \tag{1}$$

and left-multiplying by $\boldsymbol{Q}_2^T$ and right-multiplying by $\boldsymbol{R}_1^{-1}$ yields

$$\boldsymbol{Q}_2^T \boldsymbol{Q}_1 = \boldsymbol{R}_2 \boldsymbol{R}_1^{-1}. \tag{2}$$

Note that the right-hand side of Eqn (2) is upper-triangular (since $\boldsymbol{R}_i$ is). On the other hand, left-multiplying Eqn (1) by $\boldsymbol{Q}_1^T$ and right-multiplying by $\boldsymbol{R}_2^{-1}$ gives $\boldsymbol{Q}_1^T \boldsymbol{Q}_2 = \boldsymbol{R}_1 \boldsymbol{R}_2^{-1}$, and taking the transpose yields a *lower-triangular* expression for $\boldsymbol{Q}_2^T \boldsymbol{Q}_1$. Therefore $\boldsymbol{Q}_1^T \boldsymbol{Q}_2 = \boldsymbol{R}_1 \boldsymbol{R}_2^{-1}$ is both lower- and upper-triangular, and so it is diagonal. Call it $\boldsymbol{D}$. Then right-multiplying Eqn (1) by $\boldsymbol{R}_2^{-1}$ yields

$$\boldsymbol{Q}_2 \boldsymbol{R}_2 \boldsymbol{R}_2^{-1} = \boldsymbol{Q}_2 = \boldsymbol{Q}_1 \boldsymbol{R}_1 \boldsymbol{R}_2^{-1} = \boldsymbol{Q}_1 \boldsymbol{D}$$

and so $\boldsymbol{Q}_2 = \boldsymbol{Q}_1 \boldsymbol{D}$. Multiplying this by its transpose and using orthogonality of $\boldsymbol{Q}_i$ we get $\boldsymbol{I} = \boldsymbol{Q}_2^T \boldsymbol{Q}_2 = (\boldsymbol{Q}_1 \boldsymbol{D})^T (\boldsymbol{Q}_1 \boldsymbol{D}) = \boldsymbol{D}^T \boldsymbol{Q}_1^T \boldsymbol{Q}_1 \boldsymbol{D} = \boldsymbol{D}^T \boldsymbol{D} = \boldsymbol{D}^2$. This proves $\boldsymbol{D}^2 = \boldsymbol{I}$, so $\boldsymbol{D} = \boldsymbol{S}$, a diagonal matrix with entries ± 1. So $\boldsymbol{Q}_2 = \boldsymbol{Q}_1 \boldsymbol{S}$. Left multiplying Eqn (1) by $\boldsymbol{Q}_2^T = \boldsymbol{S} \boldsymbol{Q}_1^T$ then yields

$$\boldsymbol{S} \boldsymbol{Q}_1^T \boldsymbol{Q}_1 \boldsymbol{R}_1 = \boldsymbol{S} \boldsymbol{R}_1 = \boldsymbol{Q}_2^T \boldsymbol{Q}_2 \boldsymbol{R}_2 = \boldsymbol{R}_2,$$

proving that $\boldsymbol{R}_2 = \boldsymbol{S} \boldsymbol{R}_1$.

Handling $\boldsymbol{U}_i$ Finally, we consider $\boldsymbol{U}_i$. To make $\tilde{\boldsymbol{Q}}_i = \begin{bmatrix} \boldsymbol{Q}_i & \boldsymbol{U}_i \end{bmatrix}$ orthonormal, $\boldsymbol{U}_i$ can be *any* set of columns that are orthonormal to $\boldsymbol{Q}_i$. Since there is such a vast choice for $\boldsymbol{U}_i$, we then want to know if there is a relationship between $\boldsymbol{U}_1$ and $\boldsymbol{U}_2$.

Since $\boldsymbol{Q}_2 = \boldsymbol{Q}_1 \boldsymbol{S}$, those two sets of columns (i.e. $\boldsymbol{Q}_1$ and $\boldsymbol{Q}_2$) span the same subspace of $\mathbb{R}^m$. Because the matrices $\tilde{\boldsymbol{Q}}_i$ are full rank, their range must be all of $\mathbb{R}^m$, and so we must have $\mathbb{R}^m = \mathrm{col}(\boldsymbol{Q}_i) \oplus \mathrm{col}(\boldsymbol{U}_i)$. But $\mathrm{col}(\boldsymbol{Q}_1) = \mathrm{col}(\boldsymbol{Q}_2)$, so we must have that $\mathrm{col}(\boldsymbol{U}_1) = \mathrm{col}(\boldsymbol{U}_2)$. This means there exists an invertible matrix $\boldsymbol{T}$ such that $\boldsymbol{U}_2 = \boldsymbol{U}_1 \boldsymbol{T}$ because the columns of $\boldsymbol{U}_i$ are bases for the same subspace of $\mathbb{R}^m$.

Using the orthogonality of $\boldsymbol{U}_i$, the fact that $\boldsymbol{U}_i$ are $m \times (m-n)$ (hence tall and narrow), and the fact that $\boldsymbol{U}_2 = \boldsymbol{U}_1 \boldsymbol{T}$, we have that $\boldsymbol{I} = \boldsymbol{U}_2^T \boldsymbol{U}_2 = (\boldsymbol{U}_1 \boldsymbol{T})^T (\boldsymbol{U}_1 \boldsymbol{T}) = \boldsymbol{T}^T \boldsymbol{U}_1^T \boldsymbol{U}_1 \boldsymbol{T} = \boldsymbol{T}^T \boldsymbol{I} \boldsymbol{T} = \boldsymbol{T}^T \boldsymbol{T}$, proving that $\boldsymbol{T}$ is in fact orthogonal.