QR decomposition

by Marco Taboga, PhD

The QR decomposition (or QR factorization) allows to express a matrix having linearly independent columns as the product of 1) a matrix Q having orthonormal columns and 2) an upper triangular matrix R.

In order to fully understand how the QR decomposition is obtained, we should be familiar with the Gram-Schmidt process.

Overview of the decomposition

Remember that the Gram-Schmidt process is a procedure used to transform a set of linearly independent vectors into a set of orthonormal vectors (i.e., a set of vectors that have unit norm and are orthogonal to each other).

In the case of a $K\times L$ matrix A, denote its columns by $A_{\bullet 1},\ldots,A_{\bullet L}$. If these columns are linearly independent, they can be transformed into a set of orthonormal column vectors $Q_{\bullet 1},\ldots,Q_{\bullet L}$ by using the Gram-Schmidt process, which alternates normalization steps and projection steps:

1. we start with the normalization
$$Q_{\bullet 1}=\frac{1}{\|A_{\bullet 1}\|}A_{\bullet 1}$$
where $\|A_{\bullet 1}\|$ denotes the norm of $A_{\bullet 1}$;
2. we project $A_{\bullet 2}$ on $Q_{\bullet 1}$:
$$A_{\bullet 2}=\langle A_{\bullet 2},Q_{\bullet 1}\rangle Q_{\bullet 1}+\varepsilon_2$$
where $\langle A_{\bullet 2},Q_{\bullet 1}\rangle$ is the inner product between $A_{\bullet 2}$ and $Q_{\bullet 1}$ and ε_2 is the residual of the projection, orthogonal to $Q_{\bullet 1}$;
3. we normalize the residual:
$$Q_{\bullet 2}=\frac{1}{\|\varepsilon_2\|}\varepsilon_2$$
4. we project $A_{\bullet 3}$ on $Q_{\bullet 1}$ and $Q_{\bullet 2}$:
$$A_{\bullet 3}=\langle A_{\bullet 3},Q_{\bullet 1}\rangle Q_{\bullet 1}+\langle A_{\bullet 3},Q_{\bullet 2}\rangle Q_{\bullet 2}+\varepsilon_3$$
where the residual ε_3 is orthogonal to $Q_{\bullet 1}$ and $Q_{\bullet 2}$;
5. we keep on alternating normalization steps (where projection residuals are divided by their norms) and projection steps (where $A_{\bullet l}$ is projected on $Q_{\bullet 1},\ldots,Q_{\bullet l-1}$) until we have produced a set of orthonormal vectors $Q_{\bullet 1},\ldots,Q_{\bullet L}$.

Note that the residuals can be expressed in terms of normalized vectors as
$$\varepsilon_l=\|\varepsilon_l\|Q_{\bullet l}$$
for $l=1,\ldots,L$, where we have defined
$$\varepsilon_1=A_{\bullet 1}$$

Thus, the projections can be written as
$$A_{\bullet l}=\langle A_{\bullet l},Q_{\bullet 1}\rangle Q_{\bullet 1}+\ldots+\langle A_{\bullet l},Q_{\bullet l-1}\rangle Q_{\bullet l-1}+\|\varepsilon_l\|Q_{\bullet l}\quad(1)$$

The orthonormal vectors can be adjoined to form a $K\times L$ matrix
$$Q=\begin{bmatrix}Q_{\bullet 1} & \ldots & Q_{\bullet L}\end{bmatrix}$$
whose columns are orthonormal.

The coefficients of the projections can be collected in an upper triangular $L\times L$ matrix
$$R=\begin{bmatrix}\|\varepsilon_1\| & \langle A_{\bullet 2},Q_{\bullet 1}\rangle & \langle A_{\bullet 3},Q_{\bullet 1}\rangle & \ldots & \langle A_{\bullet L},Q_{\bullet 1}\rangle\\ 0 & \|\varepsilon_2\| & \langle A_{\bullet 3},Q_{\bullet 2}\rangle & \ldots & \langle A_{\bullet L},Q_{\bullet 2}\rangle\\ 0 & 0 & \|\varepsilon_3\| & & \vdots\\ \vdots & \vdots & & \ddots & \langle A_{\bullet L},Q_{\bullet L-1}\rangle\\ 0 & 0 & \ldots & 0 & \|\varepsilon_L\|\end{bmatrix}$$

By computing the matrix product between Q and R, we recover the projections in equation (1). As a matter of fact, each column of the product QR is a linear combination of the columns of Q with coefficients taken from the corresponding column of R (see the lecture on matrix products and linear combinations).

Therefore, we have that
$$A=QR$$

A formal statement

We now provide a formal statement of the QR decomposition.

Proposition Let A be a $K\times L$ matrix. If the columns of A are linearly independent, then A can be factorized as
$$A=QR$$
where Q is a $K\times L$ matrix whose columns form an orthonormal set, and R is an $L\times L$ upper triangular matrix whose diagonal entries are strictly positive.

Proof

Note that R is invertible because a triangular matrix is invertible if its diagonal entries are strictly positive.

It is time to make an example.

Example Define the 3×2 matrix
$$A=\begin{bmatrix}1 & \sqrt{5}\\ 2 & 0\\ 0 & -\sqrt{5}\end{bmatrix}$$

The norm of the first column is
$$\|A_{\bullet 1}\|=\sqrt{1^2+2^2+0^2}=\sqrt{5}$$
so that the first normalized vector is
$$Q_{\bullet 1}=\frac{1}{\|A_{\bullet 1}\|}A_{\bullet 1}=\begin{bmatrix}\frac{1}{\sqrt{5}}\\ \frac{2}{\sqrt{5}}\\ 0\end{bmatrix}$$

The inner product between $A_{\bullet 2}$ and $Q_{\bullet 1}$ is
$$\begin{aligned}&\langle A_{\bullet 2},Q_{\bullet 1}\rangle\\ &=Q_{\bullet 1}^{\intercal}A_{\bullet 2}\\ &=\begin{bmatrix}\frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} & 0\end{bmatrix}\begin{bmatrix}\sqrt{5}\\ 0\\ -\sqrt{5}\end{bmatrix}\\ &=\frac{1}{\sqrt{5}}\cdot\sqrt{5}+\frac{2}{\sqrt{5}}\cdot 0+0\cdot\left(-\sqrt{5}\right)\\ &=1\end{aligned}$$

The projection of the second column on $Q_{\bullet 1}$ is
$$\begin{aligned}&\langle A_{\bullet 2},Q_{\bullet 1}\rangle Q_{\bullet 1}\\ &=1\cdot\begin{bmatrix}\frac{1}{\sqrt{5}}\\ \frac{2}{\sqrt{5}}\\ 0\end{bmatrix}=\begin{bmatrix}\frac{1}{\sqrt{5}}\\ \frac{2}{\sqrt{5}}\\ 0\end{bmatrix}\end{aligned}$$
and the residual of the projection is
$$\begin{aligned}\varepsilon_2&=A_{\bullet 2}-\langle A_{\bullet 2},Q_{\bullet 1}\rangle Q_{\bullet 1}\\ &=\begin{bmatrix}\sqrt{5}\\ 0\\ -\sqrt{5}\end{bmatrix}-\begin{bmatrix}\frac{1}{\sqrt{5}}\\ \frac{2}{\sqrt{5}}\\ 0\end{bmatrix}\\ &=\begin{bmatrix}\frac{4}{\sqrt{5}}\\ -\frac{2}{\sqrt{5}}\\ -\frac{5}{\sqrt{5}}\end{bmatrix}\end{aligned}$$

The norm of the residual is
$$\begin{aligned}\|\varepsilon_2\|&=\sqrt{\left(\frac{4}{\sqrt{5}}\right)^2+\left(-\frac{2}{\sqrt{5}}\right)^2+\left(-\frac{5}{\sqrt{5}}\right)^2}\\ &=\sqrt{\frac{16}{5}+\frac{4}{5}+\frac{25}{5}}=\sqrt{\frac{45}{5}}=\sqrt{9}=3\end{aligned}$$

Thus,
$$Q_{\bullet 2}=\frac{1}{\|\varepsilon_2\|}\varepsilon_2=\frac{1}{3}\begin{bmatrix}\frac{4}{\sqrt{5}}\\ -\frac{2}{\sqrt{5}}\\ -\frac{5}{\sqrt{5}}\end{bmatrix}=\begin{bmatrix}\frac{4}{3\sqrt{5}}\\ -\frac{2}{3\sqrt{5}}\\ -\frac{5}{3\sqrt{5}}\end{bmatrix}$$

Let us verify that $Q_{\bullet 1}$ and $Q_{\bullet 2}$ are orthogonal:
$$\begin{aligned}&\langle Q_{\bullet 1},Q_{\bullet 2}\rangle\\ &=Q_{\bullet 2}^{\intercal}Q_{\bullet 1}\\ &=\begin{bmatrix}\frac{4}{3\sqrt{5}} & -\frac{2}{3\sqrt{5}} & -\frac{5}{3\sqrt{5}}\end{bmatrix}\begin{bmatrix}\frac{1}{\sqrt{5}}\\ \frac{2}{\sqrt{5}}\\ 0\end{bmatrix}\\ &=\frac{4}{3\sqrt{5}}\cdot\frac{1}{\sqrt{5}}-\frac{2}{3\sqrt{5}}\cdot\frac{2}{\sqrt{5}}-\frac{5}{3\sqrt{5}}\cdot 0\\ &=\frac{4}{15}-\frac{4}{15}=0\end{aligned}$$

We now have performed all the calculations that lead to the QR factorization
$$A=QR$$

The matrix with orthonormal columns is
$$Q=\begin{bmatrix}Q_{\bullet 1} & Q_{\bullet 2}\end{bmatrix}=\begin{bmatrix}\frac{1}{\sqrt{5}} & \frac{4}{3\sqrt{5}}\\ \frac{2}{\sqrt{5}} & -\frac{2}{3\sqrt{5}}\\ 0 & -\frac{5}{3\sqrt{5}}\end{bmatrix}$$
and the upper triangular matrix is
$$R=\begin{bmatrix}\|A_{\bullet 1}\| & \langle A_{\bullet 2},Q_{\bullet 1}\rangle\\ 0 & \|\varepsilon_2\|\end{bmatrix}=\begin{bmatrix}\sqrt{5} & 1\\ 0 & 3\end{bmatrix}$$

Let us check that indeed the product of Q and R equals A:
$$\begin{aligned}QR&=\begin{bmatrix}\frac{1}{\sqrt{5}} & \frac{4}{3\sqrt{5}}\\ \frac{2}{\sqrt{5}} & -\frac{2}{3\sqrt{5}}\\ 0 & -\frac{5}{3\sqrt{5}}\end{bmatrix}\begin{bmatrix}\sqrt{5} & 1\\ 0 & 3\end{bmatrix}\\ &=\begin{bmatrix}\frac{1}{\sqrt{5}}\cdot\sqrt{5}+\frac{4}{3\sqrt{5}}\cdot 0 & \frac{1}{\sqrt{5}}\cdot 1+\frac{4}{3\sqrt{5}}\cdot 3\\ \frac{2}{\sqrt{5}}\cdot\sqrt{5}-\frac{2}{3\sqrt{5}}\cdot 0 & \frac{2}{\sqrt{5}}\cdot 1-\frac{2}{3\sqrt{5}}\cdot 3\\ 0\cdot\sqrt{5}-\frac{5}{3\sqrt{5}}\cdot 0 & 0\cdot 1-\frac{5}{3\sqrt{5}}\cdot 3\end{bmatrix}\\ &=\begin{bmatrix}1 & \sqrt{5}\\ 2 & 0\\ 0 & -\sqrt{5}\end{bmatrix}\end{aligned}$$

Uniqueness of the decomposition

The QR decomposition is unique.

Proposition Under the assumptions of the previous proposition, the QR decomposition is unique, that is, the matrices Q and R satisfying the stated properties are unique.

Proof

Pre-multiplication by the Q factor

An important fact that we have discussed in the previous proof but we have not separately stated until now is that the Q matrix in the decomposition is such that
$$Q^*Q=I$$
where Q^* is the conjugate transpose of Q. As a consequence,
$$Q^*A=Q^*QR=R$$

If Q has only real entries, then the conjugate transpose coincides with the transpose and the two equations above become
$$Q^{\intercal}Q=I$$
and
$$Q^{\intercal}A=R$$

Square matrices

When the matrix A being decomposed is a square $K\times K$ matrix, then
$$A=QR$$
where Q and R are both square $K\times K$ matrices.

But a square matrix Q having orthonormal columns is a unitary matrix.

Therefore, the QR decomposition of a square matrix having linearly independent columns is the product of a unitary matrix and an upper triangular matrix with strictly positive entries.

Application to linear regression

The QR method is often used to estimate linear regressions.

In a linear regression we have an $N\times 1$ vector y of outputs and an $N\times K$ matrix of inputs whose columns are assumed to be linearly independent. We need to find the $K\times 1$ coefficient vector β that minimizes the mean squared errors made by using the fitted values
$$\widehat{y}=X\beta$$
to predict the actual values y.

The well-known solution to this problem is the so-called ordinary least squares (OLS) estimator
$$\beta=(X^{\intercal}X)^{-1}X^{\intercal}y$$

We can simplify the formula for the OLS estimator, avoid to invert a matrix and thus reduce the computational burden (and the possible numerical instabilities) by computing the QR decomposition of X:
$$X=QR$$
where Q is $N\times K$ and R is $K\times K$.

Then, the OLS estimator becomes
$$\begin{aligned}\beta&=(X^{\intercal}X)^{-1}X^{\intercal}y\\ &=(R^{\intercal}Q^{\intercal}QR)^{-1}R^{\intercal}Q^{\intercal}y\\ &=(R^{\intercal}R)^{-1}R^{\intercal}Q^{\intercal}y\\ &=R^{-1}(R^{\intercal})^{-1}R^{\intercal}Q^{\intercal}y\\ &=R^{-1}Q^{\intercal}y\end{aligned}$$
or
$$R\beta=Q^{\intercal}y$$

The latter way of writing the solution is more convenient: since R is upper triangular, we do not need to invert it, but we can use the back-substitution algorithm to find the solution β.

Solved exercises

Below you can find some exercises with explained solutions.

Exercise 1

Compute the QR decomposition of
$$A=\begin{bmatrix}1-i & 1+2i\\ 1+i & 1-2i\end{bmatrix}$$

Solution

How to cite

Please cite as:

Taboga, Marco (2017). "QR decomposition", Lectures on matrix algebra. https://www.statlect.com/matrix-algebra/QR-decomposition.