

1 Solutions to the Discrete Least Squares Problem (3.3)

In the previous section, we learned how to construct a Least Squares Problem given a set of points. The idea was to find a function such that the residual was minimized. We talked briefly about using QR decomposition and how it could be used to solve this problem. Now, we'll talk more about using QR decomposition.

Theorem 1.1: Full QR Decomposition

Let $A \in \mathbb{R}^{n \times m}$ such that $n \geq m$. Then, there exists an orthogonal $Q \in \mathbb{R}^{n \times m}$ and $R \in \mathbb{R}^{n \times m}$ with

$$R = \begin{bmatrix} r_{11} & r_{12} & \dots & r_{1m} \\ 0 & r_{22} & \dots & r_{2m} \\ 0 & 0 & \dots & r_{3m} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & r_{nm} \\ 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \dots & 0 \end{bmatrix} = \begin{bmatrix} \hat{R} \\ 0 \end{bmatrix},$$

with $\hat{R} \in \mathbb{R}^{m \times m}$ being an upper-triangular matrix and the 0 being the zero-matrix, such that

$$A = QR.$$

Remark: If A has full rank ($\text{rank}(A) = \min\{n, m\} = m$) and we choose $n_i > 0$, then the reduced QR is unique.

1.1 Purpose of QR

One of the main purposes of the QR decomposition is to solve the Least Squares Problem. Note that methods like LU and PLU decompositions are not useful because the matrix we're working with is a tall matrix ($n \geq m$).

Now, that being said, there are two cases to consider.

1. If A is a square matrix and invertible, then it still has a QR decomposition. Note that

$$\begin{aligned} Ax = y &\implies QRx = y \\ &\implies Q^T QRx = Q^T y \\ &\implies Rx = Q^T y. \end{aligned}$$

2. Otherwise, recall that

$$\min_{x \in \mathbb{R}^m} \|y - Ax\|_2^2$$

with $A \in \mathbb{R}^{n \times m}$ and $n \geq m$. Assume that A has a QR decomposition such that

$$A = QR,$$

we want to compute

$$\|y - Ax\|_2^2.$$

Notice that

$$\begin{aligned}
 \|\mathbf{y} - A\mathbf{x}\|_2^2 &= \|\mathbf{y} - QR\mathbf{x}\|_2^2 \\
 &= \|I\mathbf{y} - QR\mathbf{x}\|_2^2 \\
 &= \|QQ^T\mathbf{y} - QR\mathbf{x}\|_2^2 \\
 &= \|Q(Q^T\mathbf{y} - R\mathbf{x})\|_2^2 \\
 &= \|Q^T\mathbf{y} - R\mathbf{x}\|_2^2 \quad Q \text{ is orthogonal, so } \|Q\mathbf{x}\|_2^2 = \|\mathbf{x}\|_2^2 \\
 &= \left\| \begin{bmatrix} \hat{c} \\ d \end{bmatrix} - \begin{bmatrix} \hat{R} \\ 0 \end{bmatrix} \mathbf{x} \right\|_2^2 \\
 &= \left\| \begin{bmatrix} \hat{c} - \hat{R}\mathbf{x} \\ d \end{bmatrix} \right\|_2^2 \\
 &= \|\hat{c} - \hat{R}\mathbf{x}\|_2^2 + \|d\|_2^2
 \end{aligned}$$

Recall that we rewrote R so that

$$R = \begin{bmatrix} \hat{R} \\ \mathbf{0} \end{bmatrix}.$$

We can rewrite $Q^T\mathbf{y}$ so that

$$Q^T\mathbf{y} = \begin{bmatrix} \hat{c} \\ d \end{bmatrix},$$

where \hat{c} and d are vectors of length m .

With this in mind, we have

$$\min_{\mathbf{x} \in \mathbb{R}^m} \|\mathbf{y} - A\mathbf{x}\|_2^2 = \min_{\mathbf{x} \in \mathbb{R}^m} \|\hat{c} - \hat{R}\mathbf{x}\|_2^2 + \|d\|_2^2.$$

Here, $\min_{\mathbf{x} \in \mathbb{R}^m} \|\hat{c} - \hat{R}\mathbf{x}\|_2^2$ is the solution to $\hat{R}\mathbf{x} = \hat{c}$, with $\hat{c} = Q^T\mathbf{y}(1:m)$ and $\hat{R} = R(1:m, 1:m)$ and $A = QR$ is a QR decomposition. Additionally, $\hat{R}\mathbf{x} = \hat{c}$ has a unique solution \mathbf{x} , assuming $\text{rank}(A) = m$ and $\hat{c} - \hat{R}\mathbf{x} = \mathbf{0}$. The norm of the residual is $\|d\|_2$.

(Example.) Suppose we have $\min_{\mathbf{x} \in \mathbb{R}^m} \|\mathbf{y} - A\mathbf{x}\|_2^2$ with

$$A = \begin{bmatrix} -1 & -1 & 1 \\ 1 & 3 & 3 \\ -1 & -1 & 5 \\ 1 & 3 & 7 \end{bmatrix}$$

and

$$\mathbf{y} = \begin{bmatrix} 1 \\ 0 \\ -1 \\ 2 \end{bmatrix}$$

and want to find \mathbf{x} by using QR. Because we haven't talked about how to do the QR decomposition,

we'll give Q and R now:

$$\underbrace{\begin{bmatrix} -1 & -1 & 1 \\ 1 & 3 & 3 \\ -1 & -1 & 5 \\ 1 & 3 & 7 \end{bmatrix}}_{A \in \mathbb{R}^{4 \times 4}} = \frac{1}{2} \underbrace{\begin{bmatrix} -1 & 1 & -1 & 1 \\ 1 & 1 & -1 & -1 \\ -1 & 1 & 1 & -1 \\ 1 & 1 & 1 & 1 \end{bmatrix}}_{Q \in \mathbb{R}^{4 \times 4}} \underbrace{\begin{bmatrix} 2 & 4 & 2 \\ 0 & 2 & 8 \\ 0 & 0 & 4 \\ 0 & 0 & 0 \end{bmatrix}}_{R \in \mathbb{R}^{4 \times 3}}$$

$$= \begin{bmatrix} -\frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 2 & 4 & 2 \\ 0 & 2 & 8 \\ 0 & 0 & 4 \\ 0 & 0 & 0 \end{bmatrix}$$

Note that^a

$$\hat{R} = \begin{bmatrix} 2 & 4 & 2 \\ 0 & 2 & 8 \\ 0 & 0 & 4 \end{bmatrix}.$$

Additionally, note that

$$Q^T \mathbf{y} = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 2 \end{bmatrix}$$

so it follows that

$$\hat{\mathbf{c}} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}.$$

Finally, $d = [2] = 2$ (note that d is a vector of the remaining entries after $\hat{\mathbf{c}}$). Now, we can solve

$$\hat{R}\mathbf{x} = \hat{\mathbf{c}}.$$

In particular,

$$\begin{bmatrix} 2 & 4 & 2 \\ 0 & 2 & 8 \\ 0 & 0 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}.$$

This gives us

$$\mathbf{x} = \begin{bmatrix} -\frac{1}{2} \\ \frac{1}{2} \\ 0 \end{bmatrix}.$$

The norm of the residual is $\|d\|_2 = |2| = 2$.

^aRecall that $\hat{R} = R(1:m, 1:m)$. In other words, we take the first m rows and columns from R .