QR by Householder Reflectors

Example 1. Using classical Gram-Schmidt, compute the (reduced) QR factorization of $\begin{bmatrix} 1 & 1 & 1 \\ \delta & 0 & 0 \\ 0 & \delta & 0 \\ 0 & 0 & \delta \end{bmatrix}$ where $\delta = 10^{-10}$ in double precision.

First, $\mathbf{y} = \begin{bmatrix} 1 \\ \delta \\ 0 \\ 0 \end{bmatrix}$. Since $||\mathbf{y}||_2 = \sqrt{1+\delta^2} = \sqrt{1} = 1$, $\mathbf{q}_1 = \mathbf{y}$.

$$Q = \begin{bmatrix} 1 & & \\ \delta & & \\ 0 & & \\ 0 & & \end{bmatrix} \qquad R = \begin{bmatrix} 1 & & \\ 0 & & \\ 0 & & \end{bmatrix}$$

Then setting $\mathbf{y} = \begin{bmatrix} 1 \\ 0 \\ \delta \\ 0 \end{bmatrix}$, we have $r_{12} = \begin{bmatrix} 1 & \delta & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ \delta \\ 0 \end{bmatrix} = 1$ and $\mathbf{y} = \mathbf{y} - (1) \begin{bmatrix} 1 \\ \delta \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ -\delta \\ \delta \\ 0 \end{bmatrix}$.

Then $r_{22} = ||\mathbf{y}||_2 = \sqrt{2\delta^2} = \sqrt{2}\delta$, so $\mathbf{q}_2 = \begin{bmatrix} 0 \\ -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \\ 0 \end{bmatrix}$.

$$Q = \begin{bmatrix} 1 & 0 & \\ \delta & -\frac{\sqrt{2}}{2} & \\ 0 & \frac{\sqrt{2}}{2} & \\ 0 & 0 & \end{bmatrix} \qquad R = \begin{bmatrix} 1 & 1 & \\ 0 & \sqrt{2}\delta & \\ 0 & 0 & \end{bmatrix}$$

Now for the third column, set $\mathbf{y} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \delta \end{bmatrix}$. Then $r_{13} = \begin{bmatrix} 1 & \delta & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \\ \delta \end{bmatrix} = 1$ and $\mathbf{y} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \delta \end{bmatrix} - \begin{bmatrix} 1 \\ \delta \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ -\delta \\ 0 \\ \delta \end{bmatrix}$. Then $r_{23} = \begin{bmatrix} 0 & -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \\ \delta \end{bmatrix} = 0$ so $\mathbf{y}$ is unchanged.

Then $r_{33} = ||\mathbf{y}||_2 = \sqrt{2}\delta$ and $\mathbf{q}_3 = \begin{bmatrix} 0 \\ -\frac{\sqrt{2}}{2} \\ 0 \\ \frac{\sqrt{2}}{2} \end{bmatrix}$.

$$Q = \begin{bmatrix} 1 & 0 & 0 \\ \delta & -\frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ 0 & \frac{\sqrt{2}}{2} & 0 \\ 0 & 0 & \frac{\sqrt{2}}{2} \end{bmatrix} \qquad R = \begin{bmatrix} 1 & 1 & 1 \\ 0 & \sqrt{2}\delta & 0 \\ 0 & 0 & \sqrt{2}\delta \end{bmatrix}$$

Example 2. The check that $A = QR$ works out fine here, but recall that the goal is for Q to have orthogonal columns so that Q^T is the inverse of Q. Does Q have orthogonal columns?

$\mathbf{q}_2 \cdot \mathbf{q}_3 = \begin{bmatrix} 0 & -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 0 \end{bmatrix} \begin{bmatrix} 0 \\ -\frac{\sqrt{2}}{2} \\ 0 \\ \frac{\sqrt{2}}{2} \end{bmatrix} = \frac{1}{2}$, which is nowhere near 0.

A modified version of the Gram-Schmidt Algorithm does better in this example. The algorithms produce identical answers in exact arithmetic. The only change is at the line where we update $\mathbf{y}$: use the current version of $\mathbf{y}$ instead of the original column A_j.

Algorithm 1 Modified Gram-Schmidt Algorithm

for $j = 1, \ldots, n$ **do**
  $\mathbf{y} = A_j$
  for $i = 1, \ldots, j-1$ **do**
    $r_{ij} = \mathbf{q}_i^T \mathbf{y}$ ▷ Instead of the original column, use the updated column
    $\mathbf{y} = \mathbf{y} - r_{ij}\mathbf{q}_i$
  $r_{jj} = ||\mathbf{y}||_2$
  $\mathbf{q}_j = \frac{1}{r_{jj}}\mathbf{y}$

Example 3. Recalculate the (reduced) QR factorization of $\begin{bmatrix} 1 & 1 & 1 \\ \delta & 0 & 0 \\ 0 & \delta & 0 \\ 0 & 0 & \delta \end{bmatrix}$ where $\delta = 10^{-10}$, using Modified Gram-Schmidt in double precision.

For the first two columns, there are no changes in the algorithm (if the norm of the first column isn't one, then you do have a new result in the second column). So we pick up where

$$Q = \begin{bmatrix} 1 & 0 & \\ \delta & -\frac{\sqrt{2}}{2} & \\ 0 & \frac{\sqrt{2}}{2} & \\ 0 & 0 & \end{bmatrix} \qquad R = \begin{bmatrix} 1 & 1 & \\ 0 & \sqrt{2}\delta & \\ 0 & 0 & \end{bmatrix}.$$

We also still have $r_{13} = 1$ and $\mathbf{y} = \begin{bmatrix} 0 \\ -\delta \\ 0 \\ \delta \end{bmatrix}$. Then $r_{23} = \begin{bmatrix} 0 & -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 0 \end{bmatrix} \begin{bmatrix} 0 \\ -\delta \\ 0 \\ \delta \end{bmatrix} = \frac{\sqrt{2}}{2}\delta$ and $\mathbf{y} = \begin{bmatrix} 0 \\ -\delta \\ 0 \\ \delta \end{bmatrix} - \frac{\sqrt{2}}{2}\delta \begin{bmatrix} 0 \\ -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ -\frac{1}{2}\delta \\ -\frac{1}{2}\delta \\ \delta \end{bmatrix}$. The norm of the new $\mathbf{y}$ is $\left(\sqrt{\frac{1}{4} + \frac{1}{4} + 1}\right)\delta = \frac{\sqrt{6}}{2}\delta$.

Updating $\mathbf{y}$ for Q gives the factorization

$$Q = \begin{bmatrix} 1 & 0 & 0 \\ \delta & -\frac{\sqrt{2}}{2} & -\frac{\sqrt{6}}{6} \\ 0 & \frac{\sqrt{2}}{2} & -\frac{\sqrt{6}}{6} \\ 0 & 0 & \frac{\sqrt{6}}{3} \end{bmatrix} \qquad R = \begin{bmatrix} 1 & 1 & 1 \\ 0 & \sqrt{2}\delta & \frac{\sqrt{2}}{2}\delta \\ 0 & 0 & \frac{\sqrt{6}}{2}\delta \end{bmatrix}.$$

How is orthogonality in Q now?

Well, $\mathbf{q}_2 \cdot \mathbf{q}_3 = \frac{\sqrt{2(6)}}{12} - \frac{\sqrt{2(6)}}{12} = 0$, so that's much better.

In both versions, $\mathbf{q}_1 \cdot \mathbf{q}_2 = -\frac{\sqrt{2}}{2}\delta$, which... leaves room for improvement.

Householder Reflectors

The next (final) process of finding a QR factorization bears a closer resemblance to the process of Guassian Elimination, in that we are manipulating the matrix A into upper triangular form R. However, we cannot get to R using row operations, as that will not introduce orthogonality.

The theory is to multiply A on the left by orthogonal matrices until it is transformed into R. That is, after multiplying by one orthogonal matrix Q_1, we should have:

$$Q_1 A = \begin{bmatrix} r_{11} & ? & \dots & ? \\ 0 & ? & \dots & ? \\ \vdots & & & \vdots \\ 0 & ? & \dots & ? \end{bmatrix}$$

The orthogonal matrices used to transform A are a type of orthogonal matrix called a Householder reflector, denoted by H. Householder reflectors are also symmetric, so not only is $H^T = H^{-1}$, $H^T = H$ implies $H = H^{-1}$.

So we'll have a series of Householder reflectors $H_1, \dots, H_n$ where

$$H_n \dots H_2 H_1 A = R,$$

and then because each H is orthogonal and symmetric,

$$A = H_1 H_2 \dots H_n R.$$

Example 4. Verify that the product $H_1 \dots H_n$ is an orthogonal matrix.

Using the theorem/definition that Q is orthogonal if $Q^{-1} = Q^T$, we check the product

$$(H_1 \dots H_n)^T (H_1 \dots H_n) = H_n \dots H_2 H_1 H_1 H_2 \dots H_n = H_n \dots H_2 H_2 \dots H_n = I.$$

Therefore, since $H_1 \dots H_n$ is square, $H_1 \dots H_n$ is orthogonal.

We still need more information to define a Householder reflector. We begin with $\mathbf{x}$ equal to the first column of A; we desire H that will reflect $\mathbf{x}$ to the vector where the first element is nonzero, and all remaining entries are zero (basically, the x-axis of $\mathbb{R}^n$). So we require

$$H\mathbf{x} = \mathbf{u} \text{ and } ||\mathbf{x}||_2 = ||\mathbf{u}||_2.$$

Question 1. So what is the nonzero element of $\mathbf{u}$ when $\mathbf{x} = A_1$?

It must be that $r_{11} = ||\mathbf{x}||_2 = ||A_1||_2$, or else the length would not be preserved.

More generally, the goal at a particular step of the process is to take a vector $\mathbf{x}$ with m entries and reflect it over an $m-1$ dimensional plane to another vector $\mathbf{u}$ of the same length. (Draw triangle with $\mathbf{x}$ and $\mathbf{u}$)

Question 2. Verify that $\mathbf{u} - \mathbf{x}$ is orthogonal to $\mathbf{u} + \mathbf{x}$.

$(\mathbf{u}-\mathbf{x}) \cdot (\mathbf{u}+\mathbf{x}) = \mathbf{u} \cdot (\mathbf{u}+\mathbf{x}) - \mathbf{x} \cdot (\mathbf{u}+\mathbf{x}) = \mathbf{u}\cdot\mathbf{u} + \mathbf{x}\cdot\mathbf{u} - \mathbf{x}\cdot\mathbf{u} - \mathbf{x}\cdot\mathbf{x} = ||\mathbf{u}||_2^2 - ||\mathbf{x}||_2^2 = 0$

For convenience, let $\mathbf{v}$ represent $\mathbf{u}-\mathbf{x}$. We want to reflect $\mathbf{x}$ over $\mathbf{u}+\mathbf{x}$. To do so, we remove from $\mathbf{x}$, twice the projection of $\mathbf{x}$ onto $\mathbf{v}$.

The projection of $\mathbf{x}$ onto $\mathbf{v}$ is traditionally given in linear algebra as

$$\frac{\mathbf{x}\cdot\mathbf{v}}{\mathbf{v}\cdot\mathbf{v}}\mathbf{v}.$$

We rewrite the projection formula in a matrix form (working towards our orthogonal matrix H). Since $\frac{\mathbf{x}\cdot\mathbf{v}}{\mathbf{v}\cdot\mathbf{v}}$ is a scalar, we can multiply by it second. We can also reverse the order of the vectors in the dot product, by the properties of the dot product.

$$\mathbf{v}\frac{\mathbf{v}\cdot\mathbf{x}}{\mathbf{v}\cdot\mathbf{v}} = \frac{1}{\mathbf{v}\cdot\mathbf{v}}\mathbf{v}(\mathbf{v}^T\mathbf{x}) = \frac{1}{\mathbf{v}\cdot\mathbf{v}}(\mathbf{v}\mathbf{v}^T)\mathbf{x} = \frac{\mathbf{v}\mathbf{v}^T}{\mathbf{v}^T\mathbf{v}}\mathbf{x}$$

Let P represent $\frac{\mathbf{v}\mathbf{v}^T}{\mathbf{v}^T\mathbf{v}}$ (which is now a matrix, not a scalar).

Definition 3. A *projection matrix* is a matrix that satisfies $P^2 = P$.

Question 4. Is P a projection matrix?

Left for homework!

$$\mathbf{u} = \mathbf{x} - 2P\mathbf{x} = (I - 2P)\mathbf{x}$$

So we define $H = I - 2P = I - 2\dfrac{\mathbf{v}\mathbf{v}^T}{\mathbf{v}^T\mathbf{v}}$, where $\mathbf{v}$ is the difference between $\mathbf{x}$ and the desired vector $\mathbf{u}$.

Example 5. Find a Householder reflector that transforms $\mathbf{x} = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$ into an equal length vector on the x-axis.

First, the goal is to transform into the vector $\mathbf{u} = \begin{bmatrix} ||\mathbf{x}||_2 \\ 0 \end{bmatrix}$, and here, $||\mathbf{x}||_2 = 5$. So $\mathbf{u} = \begin{bmatrix} 5 \\ 0 \end{bmatrix}$.

Now we subtract to form $\mathbf{v} = \begin{bmatrix} 2 \\ -4 \end{bmatrix}$.

Next, $P = \dfrac{\mathbf{v}\mathbf{v}^T}{\mathbf{v}^T\mathbf{v}} = \left(\dfrac{1}{20}\right) \begin{bmatrix} 4 & -8 \\ -8 & 16 \end{bmatrix}$. Then

$$H = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \frac{1}{10} \begin{bmatrix} 4 & -8 \\ -8 & 16 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \frac{1}{5} \begin{bmatrix} 2 & -4 \\ -4 & 8 \end{bmatrix} = \begin{bmatrix} \frac{3}{5} & \frac{4}{5} \\ \frac{4}{5} & -\frac{3}{5} \end{bmatrix}.$$

We can check our answer:

$H\mathbf{x} = \begin{bmatrix} \frac{3}{5} & \frac{4}{5} \\ \frac{4}{5} & -\frac{3}{5} \end{bmatrix} \begin{bmatrix} 3 \\ 4 \end{bmatrix} = \begin{bmatrix} \frac{9+16}{5} \\ \frac{12-12}{5} \end{bmatrix} = \begin{bmatrix} 5 \\ 0 \end{bmatrix}$ as desired.

To complete a process for the QR factorization, we only need to know how to handle the later columns. On the first step, H_1 will be exactly as described above, and we form $H_1 A$ which looks like:

$$\begin{bmatrix} r_{11} & r_{12} & \dots & r_{1n} \\ 0 & ? & \dots & ? \\ \vdots & & & \vdots \\ 0 & ? & \dots & ? \end{bmatrix}$$

We repeat this process with just the vector below those that have been completed. So on the second step, we use the $n-1$ elements from the diagonal down to generate a Householder reflector. However, doing so generates $\hat{H}_2$ of size $(n-1) \times (n-1)$. We want to apply $\hat{H}_2$ only to the elements below the completed rows, so to leave the previous rows unchanged we set H_2 as the lower part of a matrix with identity on the diagonal. That is, in general, the H_k Householder reflector is the block matrix

$$\begin{bmatrix} I_k & 0 \\ 0 & \hat{H}_k \end{bmatrix}.$$

Example 6. Compute the full QR factorization of $\begin{bmatrix} 2 & 0 \\ 1 & -1 \\ 2 & -7 \end{bmatrix}$.

First, $\mathbf{x} = \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix}$ with norm 3, so $\mathbf{u} = \begin{bmatrix} 3 \\ 0 \\ 0 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} 1 \\ -1 \\ -2 \end{bmatrix}$. Then

$$\begin{aligned} H_1 &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \left(\frac{2}{1+1+4}\right) \begin{bmatrix} 1 \\ -1 \\ -2 \end{bmatrix} \begin{bmatrix} 1 & -1 & -2 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \left(\frac{1}{3}\right) \begin{bmatrix} 1 & -1 & -2 \\ -1 & 1 & 2 \\ -2 & 2 & 4 \end{bmatrix} \\ &= \begin{bmatrix} \frac{2}{3} & \frac{1}{3} & \frac{2}{3} \\ \frac{1}{3} & \frac{2}{3} & -\frac{2}{3} \\ \frac{2}{3} & -\frac{2}{3} & -\frac{1}{3} \end{bmatrix} \end{aligned}$$

Then $H_1 A = \begin{bmatrix} \frac{2}{3} & \frac{1}{3} & \frac{2}{3} \\ \frac{1}{3} & \frac{2}{3} & -\frac{2}{3} \\ \frac{2}{3} & -\frac{2}{3} & -\frac{1}{3} \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 1 & -1 \\ 2 & -7 \end{bmatrix} = \begin{bmatrix} 3 & -5 \\ 0 & 4 \\ 0 & 3 \end{bmatrix}$.

Next, $\mathbf{x} = \begin{bmatrix} 4 \\ 3 \end{bmatrix}$ with norm 5, so $\mathbf{u} = \begin{bmatrix} 5 \\ 0 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} 1 \\ -3 \end{bmatrix}$.

Then,

$$\begin{aligned} \hat{H}_2 &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \frac{2}{10} \begin{bmatrix} 1 & -3 \\ -3 & 9 \end{bmatrix} \\ &= \begin{bmatrix} \frac{4}{5} & \frac{3}{5} \\ \frac{3}{5} & -\frac{4}{5} \end{bmatrix}. \end{aligned}$$

This makes $H_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{4}{5} & \frac{3}{5} \\ 0 & \frac{3}{5} & -\frac{4}{5} \end{bmatrix}$. Then R equals $H_2 H_1 A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{4}{5} & \frac{3}{5} \\ 0 & \frac{3}{5} & -\frac{4}{5} \end{bmatrix} \begin{bmatrix} 3 & -5 \\ 0 & 4 \\ 0 & 3 \end{bmatrix} = \begin{bmatrix} 3 & -5 \\ 0 & 5 \\ 0 & 0 \end{bmatrix}$.

Then Q of the QR factorization is $H_1 H_2 = \begin{bmatrix} \frac{2}{3} & \frac{1}{3} & \frac{2}{3} \\ \frac{1}{3} & \frac{2}{3} & -\frac{2}{3} \\ \frac{2}{3} & -\frac{2}{3} & -\frac{1}{3} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{4}{5} & \frac{3}{5} \\ 0 & \frac{3}{5} & -\frac{4}{5} \end{bmatrix} = \frac{1}{15} \begin{bmatrix} 10 & 10 & -5 \\ 5 & 2 & 14 \\ 10 & -11 & -2 \end{bmatrix}$.

We should be able to check: $\frac{1}{15} \begin{bmatrix} 10 & 10 & -5 \\ 5 & 2 & 14 \\ 10 & -11 & -2 \end{bmatrix} \begin{bmatrix} 3 & -5 \\ 0 & 5 \\ 0 & 0 \end{bmatrix} = \frac{1}{15} \begin{bmatrix} 30 & 0 \\ 15 & -15 \\ 30 & -105 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 1 & -1 \\ 2 & -7 \end{bmatrix}$.

A note on full vs. reduced: Householder Reflectors are naturally producing a full factorization. If you wanted to give a reduced factorization, take the answer and erase any below diagonal elements in R and the corresponding columns in Q. For Example 6, we have

$$R = \begin{bmatrix} 3 & -5 \\ 0 & 5 \\ 0 & 0 \end{bmatrix} \text{ which becomes } R = \begin{bmatrix} 3 & -5 \\ 0 & 5 \end{bmatrix}$$

and

$$Q = \frac{1}{15}\begin{bmatrix} 10 & 10 & -5 \\ 5 & 2 & 14 \\ 10 & -11 & -2 \end{bmatrix} \text{ which becomes } Q = \frac{1}{15}\begin{bmatrix} 10 & 10 \\ 5 & 2 \\ 10 & -11 \end{bmatrix}.$$

I leave it to you to check again that $QR = A$.

Stability and Operation Counts

Computing a QR factorization by Householder reflectors takes about $2mn^2 - \frac{2}{3}n^3$ operations, whereas using Gram-Schmidt is about $2mn^2$. It also typically requires less memory.

As for numerical stability, the Householder reflectors method is used in practice and is known to deliver better orthogonality in Q.

Question 5. In practice, routines will use $\pm||\mathbf{x}||_2$ where the sign is chosen to be the opposite of the first element of $\mathbf{x}$. Why might this be better, numerically, than always choosing the positive norm?

It avoids subtracting nearly equal numbers, which we discussed previously may result in loss of significance.