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.

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?

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.

How is orthogonality in Q now?

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.

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 $\mathbf{x} = A_1$?

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 $\mathbf{u} + \mathbf{x}$.

So we want to project $\mathbf{x}$ over $\mathbf{x} - \mathbf{u}$. For convenience, let $\mathbf{v}$ represent $\mathbf{x} - \mathbf{u}$. We want to 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 wish to rewrite this in a matrix form (working towards our orthogonal matrix H).

Let P represent $\dfrac{\mathbf{v}\mathbf{v}^T}{\mathbf{v}^T\mathbf{v}}$ (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!

So as described, we want to remove 2 times the projection onto $\mathbf{v}$ from $\mathbf{x}$ to get $\mathbf{u}$. That is,

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

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.

We can check our answer:

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_1A 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, this 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 $\begin{bmatrix} 2 & 0 \\ 1 & -1 \\ 2 & -7 \end{bmatrix}$.

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, this 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?