Example: Solving a Least Squares Problem using Householder transformations

Problem For $A = \begin{pmatrix} 3 & -2 \\ 0 & 3 \\ 4 & 4 \end{pmatrix}$ and $b = \begin{pmatrix} 3 \\ 5 \\ 4 \end{pmatrix}$, solve $\min ||b - Ax||$.

Solution: Householder transformations

One can use Householder transformations to form a QR factorization of A and use the QR factorization to solve the least squares problem. We will present a different approach here that does not require the calculation of Q. We use an idea similar to an idea that you may have learned to solve a square linear system – the augmented equations idea. In this approach we will tack b onto the end of A forming the augmented equation $\widehat{A} = (A, b)$. If A is $m \times n$ then $\widehat{A}$ will be $m \times (n+1)$. We then change $\widehat{A}$ to simpler forms applying "legal" operations. As in solving square linear systems, our goal will be transform $\widehat{A}$ to triangular form, when the solution will be easy to find. For a least squares problem the legal operations are operations that don't change the solution to the least squares problem. Therefore the legal operations are multiplying A and b (or $\widehat{A}$) by orthogonal matrices and, in particular, we use Householder transformations.

With this approach the algorithm to solve the least square problem is:
(1) Form $\widehat{A} = (A, b)$
(2) Triangularize $\widehat{A}$ to produce the triangular matrix $\widehat{R}$. (see below)
(3) Let R be the $n \times n$ upper left corner of the $\widehat{R}$
(4) Let $c =$ the first n components of the last column of $\widehat{R}$.
(5) Solve $Rx = c$ for x. x solves least squares problem.

We illustrate this procedure for the above matrix.

(1) $\widehat{A} = \begin{pmatrix} 3 & -2 & 3 \\ 0 & 3 & 5 \\ 4 & 4 & 4 \end{pmatrix}$.

(2) We postpone the details of triangularizing $\widehat{A}$. The result is $\widehat{R} = \begin{pmatrix} -5 & -2 & -5 \\ 0 & -5 & -3 \\ 0 & 0 & -4 \end{pmatrix}$.

(3) $R = \begin{pmatrix} -5 & -2 \\ 0 & -5 \end{pmatrix}$.

(4) $c = \begin{pmatrix} -5 \\ -3 \end{pmatrix}$.

(5) Solve $Rx = \begin{pmatrix} -5 & -2 \\ 0 & -5 \end{pmatrix} x = \begin{pmatrix} -5 \\ -3 \end{pmatrix} = c$. We get $x = \begin{pmatrix} 0.76 \\ 0.60 \end{pmatrix}$.

We will accomplish step (2) by using Householder transformations. A Householder transformation can transform a vector so that all the components of the transformed vector below the k^{th} entry are zero. We apply, sequentially to $\widehat{A}$, a Householder transformation that zeros column 1 below the (1,1) diagonal element, then a second Householder transformation that zeros out column 2 below the (2,2) diagonal element, a third Householder transformation that zeros out column three below the (3,3) diagonal element, etc.. Repeating this process $n+1$ times we can convert $\widehat{A}$ to triangular form.

At step k of this procedure, where the goal is to zero column k below the diagonal, one can apply a Householder transformation to the submatrix consisting of rows k to m and columns k to $n+1$ of $\widehat{A}$. Using Matlab style notation we will represent this submatrix by $B = \widehat{A}_{k:m,k:n+1}$. Zeroing out column k of the current $\widehat{A}$ below the (k,k) diagonal entry is equivalent to zeroing out column one of B below the first entry in this column. Therefore high level pseudocode for the conversion of $\widehat{A}$ to triangular form is:

Triangularize $m \times (n+1)$ matrix $\widehat{A}$ using Householder transformations:
for k = 1 to n +1
  (1) let $z =$ the first column of the submatrix B, where B = $\widehat{A}_{k:m,k:n+1}$
  (2) Construct a Householder transformation that zeros out z below the first entry in z
  (3) Apply the Householder transformation to B, updating B and, therefore, $\widehat{A}$

We need to describe the Householder transformations in step (2). We will present the mechanics of the computations and not any of the theory. For z in step 1), let $\text{sign}(z_1)$ be $+1$ if $z_1 \geq 0$ and let $\text{sign}(z_1)$ be

-1 if $z_1 < 0$. z_1 is the first component of z. Also let e be a vector of the same dimension as z that is all zero except the first element is one. Here are details for the above algorithm:

Triangularize $m \times (n+1)$ matrix $\widehat{A}$ using Householder transformations (more detail):
for k = 1 to n +1

(1) let $z =$ the first column of the submatrix B, where B = $\widehat{A}_{k:m,k:n+1}$

(2) Construct a Householder transformation that zeros out z below the first entry in z:

(a) $v = \text{sign}(z_1)||z||_2\, e + z$ %vector normal to a Householder "mirror"

(b) $v = v/||v||_2$ % unit vector normal to a Householder "mirror"

(3) Apply the Householder transformation to B, updating B and, therefore, $\widehat{A}$:
new $B = B - 2v(v^T B)$

For the above $\widehat{A}$ matrix the calculations are:
$k = 1$

(1) $B = \widehat{A} = \begin{pmatrix} 3 & -2 & 3 \\ 0 & 3 & 5 \\ 4 & 4 & 4 \end{pmatrix}$, $z = \begin{pmatrix} 3 \\ 0 \\ 4 \end{pmatrix}$.

(2) (a) $v = \text{sign}(z_1)||z||_2\, e + z = \text{sign}(3)\, 5 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 3 \\ 0 \\ 4 \end{pmatrix} = (+1)\, 5 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 3 \\ 0 \\ 4 \end{pmatrix} = \begin{pmatrix} 8 \\ 0 \\ 4 \end{pmatrix}$.

(b) $v = v/||v||_2 = \frac{1}{4\sqrt{5}} \begin{pmatrix} 8 \\ 0 \\ 4 \end{pmatrix} = \frac{1}{\sqrt{5}} \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}$.

(3) new $\widehat{A} =$ new $B = B - 2v(v^T B) = \begin{pmatrix} 3 & -2 & 3 \\ 0 & 3 & 5 \\ 4 & 4 & 4 \end{pmatrix} - 2\frac{1}{\sqrt{5}} \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix} \left[\frac{1}{\sqrt{5}} \begin{pmatrix} 2 & 0 & 1 \end{pmatrix} \begin{pmatrix} 3 & -2 & 3 \\ 0 & 3 & 5 \\ 4 & 4 & 4 \end{pmatrix} \right]$

$= \begin{pmatrix} 3 & -2 & 3 \\ 0 & 3 & 5 \\ 4 & 4 & 4 \end{pmatrix} - 2\frac{1}{\sqrt{5}} \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix} \left[\frac{1}{\sqrt{5}} \begin{pmatrix} 10 & 0 & 10 \end{pmatrix} \right] = \begin{pmatrix} 3 & -2 & 3 \\ 0 & 3 & 5 \\ 4 & 4 & 4 \end{pmatrix} - \begin{pmatrix} 8 & 0 & 8 \\ 0 & 0 & 0 \\ 10 & 0 & 10 \end{pmatrix}$

$= \begin{pmatrix} -5 & -2 & -5 \\ 0 & 3 & 5 \\ 0 & 4 & 0 \end{pmatrix}$.

$k = 2$

(1) $B = \begin{pmatrix} 3 & 5 \\ 4 & 0 \end{pmatrix}$, $z = \begin{pmatrix} 3 \\ 4 \end{pmatrix}$.

(2) (a) $v = \text{sign}(z_1)||z||_2\, e + z = \text{sign}(3)\, 5 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 3 \\ 4 \end{pmatrix} = (+1)\, 5 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 3 \\ 4 \end{pmatrix} = \begin{pmatrix} 8 \\ 4 \end{pmatrix}$.

(b) $v = v/||v||_2 = \frac{1}{4\sqrt{5}} \begin{pmatrix} 8 \\ 4 \end{pmatrix} = \frac{1}{\sqrt{5}} \begin{pmatrix} 2 \\ 1 \end{pmatrix}$.

(3) new $B = B - 2v(v^T B) = \begin{pmatrix} 3 & 5 \\ 4 & 0 \end{pmatrix} - 2\frac{1}{\sqrt{5}} \begin{pmatrix} 2 \\ 1 \end{pmatrix} \left[\frac{1}{\sqrt{5}} \begin{pmatrix} 2 & 1 \end{pmatrix} \begin{pmatrix} 3 & 5 \\ 4 & 0 \end{pmatrix} \right]$

$= \begin{pmatrix} 3 & 5 \\ 4 & 0 \end{pmatrix} - 2\frac{1}{\sqrt{5}} \begin{pmatrix} 2 \\ 1 \end{pmatrix} \left[\frac{1}{\sqrt{5}} \begin{pmatrix} 10 & 10 \end{pmatrix} \right] = \begin{pmatrix} 3 & 5 \\ 4 & 0 \end{pmatrix} - \begin{pmatrix} 8 & 8 \\ 4 & 4 \end{pmatrix} = \begin{pmatrix} -5 & -3 \\ 0 & -4 \end{pmatrix}$.

We now replace the lower, right 2 by 2 corner of the $\widehat{A}$ in the previous, $k = 1$, step to obtain the new $\widehat{A} = \begin{pmatrix} -5 & -2 & -5 \\ 0 & -5 & -3 \\ 0 & 0 & -4 \end{pmatrix}$.

$k = 3$

The current $\widehat{A}$ is triangular, so no work is required for $k = 3$. Set $\widehat{R} =$ the current $\widehat{A}$.

Note that the above calculatons transform the original least squares problem to a simpler problem with the same solution:

$$\min \left\| \begin{pmatrix} -5 \\ -3 \\ -4 \end{pmatrix} - \begin{pmatrix} -5 & -2 \\ 0 & -5 \\ 0 & 0 \end{pmatrix} x \right\|.$$

In this problem, by varying x, only the first two components of $\begin{pmatrix} -5 \\ -3 \\ -4 \end{pmatrix}$ can be matched. This means the least squares solution to the problem must solve $Rx = c$.