Jim Lambers
MAT 610
Summer Session 2009-10
Lecture 9 Notes

These notes correspond to Section 5.1 in the text.

The QR Factorization

Let A be an $m \times n$ matrix with full column rank. The *QR factorization* of A is a decomposition $A = QR$, where Q is an $m \times m$ orthogonal matrix and R is an $m \times n$ upper triangular matrix. There are three ways to compute this decomposition:

1. Using Householder matrices, developed by Alston S. Householder
2. Using Givens rotations, also known as Jacobi rotations, used by W. Givens and originally invented by Jacobi for use with in solving the symmetric eigenvalue problem in 1846.
3. A third, less frequently used approach is the *Gram-Schmidt orthogonalization*.

Givens Rotations

We illustrate the process in the case where A is a 2×2 matrix. In Gaussian elimination, we compute $L^{-1}A = U$ where L^{-1} is unit lower triangular and U is upper triangular. Specifically,

$$\begin{bmatrix} 1 & 0 \\ m_{21} & 1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{bmatrix} a_{11}^{(2)} & a_{12}^{(2)} \\ 0 & a_{22}^{(2)} \end{bmatrix}, \quad m_{21} = -\frac{a_{21}}{a_{11}}.$$

By contrast, the QR decomposition computes $Q^T A = R$, or

$$\begin{bmatrix} \gamma & -\sigma \\ \sigma & \gamma \end{bmatrix}^T \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{bmatrix} r_{11} & r_{12} \\ 0 & r_{22} \end{bmatrix},$$

where $\gamma^2 + \sigma^2 = 1$.

From the relationship $-\sigma a_{11} + \gamma a_{21} = 0$ we obtain

$$\begin{aligned} \gamma a_{21} &= \sigma a_{11} \\ \gamma^2 a_{21}^2 &= \sigma^2 a_{11}^2 = (1-\gamma^2) a_{11}^2 \end{aligned}$$

which yields

$$\gamma = \pm \frac{a_{11}}{\sqrt{a_{21}^2 + a_{11}^2}}.$$

It is conventional to choose the + sign. Then, we obtain

$$\sigma^2 = 1 - \gamma^2 = 1 - \frac{a_{11}^2}{a_{21}^2 + a_{11}^2} = \frac{a_{21}^2}{a_{21}^2 + a_{11}^2},$$

or

$$\sigma = \pm\frac{a_{21}}{\sqrt{a_{21}^2 + a_{11}^2}}.$$

Again, we choose the + sign. As a result, we have

$$r_{11} = a_{11}\frac{a_{11}}{\sqrt{a_{21}^2 + a_{11}^2}} + a_{21}\frac{a_{21}}{\sqrt{a_{21}^2 + a_{11}^2}} = \sqrt{a_{21}^2 + a_{11}^2}.$$

The matrix

$$Q = \begin{bmatrix} \gamma & -\sigma \\ \sigma & \gamma \end{bmatrix}^T$$

is called a *Givens rotation*. It is called a rotation because it is orthogonal, and therefore length-preserving, and also because there is an angle θ such that $\sin\theta = \sigma$ and $\cos\theta = \gamma$, and its effect is to rotate a vector clockwise through the angle θ. In particular,

$$\begin{bmatrix} \gamma & -\sigma \\ \sigma & \gamma \end{bmatrix}^T \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} \rho \\ 0 \end{bmatrix}$$

where $\rho = \sqrt{\alpha^2 + \beta^2}$, $\alpha = \rho\cos\theta$ and $\beta = \rho\sin\theta$. It is easy to verify that the product of two rotations is itself a rotation. Now, in the case where A is an $n \times n$ matrix, suppose that we have the vector

$$\begin{bmatrix} \times \\ \vdots \\ \times \\ \alpha \\ \times \\ \vdots \\ \times \\ \beta \\ \times \\ \vdots \\ \times \end{bmatrix}.$$

Then

$$\begin{bmatrix} 1 & & & & & & & & & & \\ & \ddots & & & & & & & & & \\ & & 1 & & & & & & & & \\ & & & \gamma & & & \sigma & & & & \\ & & & & 1 & & & & & & \\ & & & & & \ddots & & & & & \\ & & & & & & 1 & & & & \\ & & & -\sigma & & & & \gamma & & & \\ & & & & & & & & 1 & & \\ & & & & & & & & & \ddots & \\ & & & & & & & & & & 1 \end{bmatrix} \begin{bmatrix} \times \\ \vdots \\ \times \\ \alpha \\ \times \\ \vdots \\ \times \\ \beta \\ \times \\ \vdots \\ \times \end{bmatrix} = \begin{bmatrix} \times \\ \vdots \\ \times \\ \rho \\ \times \\ \vdots \\ \times \\ 0 \\ \times \\ \vdots \\ \times \end{bmatrix}.$$

So, in order to transform A into an upper triangular matrix R, we can find a product of rotations Q such that $Q^T A = R$. It is easy to see that $O(n^2)$ rotations are required. Each rotation takes $O(n)$ operations, so the entire process of computing the QR factorization requires $O(n^3)$ operations.

It is important to note that the straightforward approach to computing the entries γ and σ of the Givens rotation,

$$\gamma = \frac{\alpha}{\sqrt{\alpha^2 + \beta^2}}, \quad \sigma = \frac{\beta}{\sqrt{\alpha + 2 + \beta^2}},$$

is not always advisable, because in floating-point arithmetic, the computation of $\sqrt{\alpha^2 + \beta^2}$ could overflow. To get around this problem, suppose that $|\beta| \geq |\alpha|$. Then, we can instead compute

$$\tau = \frac{\alpha}{\beta}, \quad \sigma = \frac{1}{\sqrt{1 + \tau^2}}, \quad \gamma = \sigma\tau,$$

which is guaranteed not to overflow since the only number that is squared is less than one in magnitude. On the other hand, if $|\alpha| \geq |\beta|$, then we compute

$$\tau = \frac{\beta}{\alpha}, \quad \gamma = \frac{1}{\sqrt{1 + \tau^2}}, \quad \sigma = \gamma\tau.$$

Now, we describe the entire algorithm for computing the QR factorization using Givens rotations. Let $[c, s] = \texttt{givens}(a, b)$ be a MATLAB-style function that computes c and s such that

$$\begin{bmatrix} c & -s \\ s & c \end{bmatrix}^T \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} r \\ 0 \end{bmatrix}, \quad r = \sqrt{a^2 + b^2}.$$

Then, let $G(i, j, c, s)^T$ be the Givens rotation matrix that rotates the ith and jth elements of a vector $\mathbf{v}$ clockwise by the angle θ such that $\cos\theta = c$ and $\sin\theta = s$, so that if $v_i = a$ and $v_j = b$, and $[c, s] = \texttt{givens}(a, b)$, then in the updated vector $\mathbf{u} = G(i, j, c, s)^T \mathbf{v}$, $u_i = r = \sqrt{a^2 + b^2}$ and $u_j = 0$. The QR factorization of an $m \times n$ matrix A is then computed as follows.

$Q = I$
$R = A$ **for** $j = 1 : n$ **do**
  for $i = m : -1 : j+1$ **do**
    $[c, s] = \texttt{givens}(r_{i-1,j}, r_{ij})$
    $R = G(i, j, c, s)^T R$
    $Q = QG(i, j, c, s)$
  end
end

Note that the matrix Q is accumulated by *column* rotations of the identity matrix, because the matrix by which A is multiplied to reduce A to upper-triangular form, a product of row rotations, is Q^T.

Example We use Givens rotations to compute the QR factorization of

$$A = \begin{bmatrix} 0.8147 & 0.0975 & 0.1576 \\ 0.9058 & 0.2785 & 0.9706 \\ 0.1270 & 0.5469 & 0.9572 \\ 0.9134 & 0.9575 & 0.4854 \\ 0.6324 & 0.9649 & 0.8003 \end{bmatrix}.$$

First, we compute a Givens rotation that, when applied to a_{41} and a_{51}, zeros a_{51}:

$$\begin{bmatrix} 0.8222 & -0.5692 \\ 0.5692 & 0.8222 \end{bmatrix}^T \begin{bmatrix} 0.9134 \\ 0.6324 \end{bmatrix} = \begin{bmatrix} 1.1109 \\ 0 \end{bmatrix}.$$

Applying this rotation to rows 4 and 5 yields

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0.8222 & -0.5692 \\ 0 & 0 & 0 & 0.5692 & 0.8222 \end{bmatrix}^T \begin{bmatrix} 0.8147 & 0.0975 & 0.1576 \\ 0.9058 & 0.2785 & 0.9706 \\ 0.1270 & 0.5469 & 0.9572 \\ 0.9134 & 0.9575 & 0.4854 \\ 0.6324 & 0.9649 & 0.8003 \end{bmatrix} = \begin{bmatrix} 0.8147 & 0.0975 & 0.1576 \\ 0.9058 & 0.2785 & 0.9706 \\ 0.1270 & 0.5469 & 0.9572 \\ 1.1109 & 1.3365 & 0.8546 \\ 0 & 0.2483 & 0.3817 \end{bmatrix}.$$

Next, we compute a Givens rotation that, when applied to a_{31} and a_{41}, zeros a_{41}:

$$\begin{bmatrix} 0.1136 & -0.9935 \\ 0.9935 & 0.1136 \end{bmatrix}^T \begin{bmatrix} 0.1270 \\ 1.1109 \end{bmatrix} = \begin{bmatrix} 1.1181 \\ 0 \end{bmatrix}.$$

Applying this rotation to rows 3 and 4 yields

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0.1136 & -0.9935 & 0 \\ 0 & 0 & 0.9935 & 0.1136 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}^T \begin{bmatrix} 0.8147 & 0.0975 & 0.1576 \\ 0.9058 & 0.2785 & 0.9706 \\ 0.1270 & 0.5469 & 0.9572 \\ 1.1109 & 1.3365 & 0.8546 \\ 0 & 0.2483 & 0.3817 \end{bmatrix} = \begin{bmatrix} 0.8147 & 0.0975 & 0.1576 \\ 0.9058 & 0.2785 & 0.9706 \\ 1.1181 & 1.3899 & 0.9578 \\ 0 & -0.3916 & -0.8539 \\ 0 & 0.2483 & 0.3817 \end{bmatrix}.$$

Next, we compute a Givens rotation that, when applied to a_{21} and a_{31}, zeros a_{31}:

$$\begin{bmatrix} 0.6295 & -0.7770 \\ 0.7770 & 0.6295 \end{bmatrix}^T \begin{bmatrix} 0.9058 \\ 1.1181 \end{bmatrix} = \begin{bmatrix} 1.4390 \\ 0 \end{bmatrix}.$$

Applying this rotation to rows 2 and 3 yields

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0.6295 & -0.7770 & 0 & 0 \\ 0 & 0.7770 & 0.6295 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}^T \begin{bmatrix} 0.8147 & 0.0975 & 0.1576 \\ 0.9058 & 0.2785 & 0.9706 \\ 1.1181 & 1.3899 & 0.9578 \\ 0 & -0.3916 & -0.8539 \\ 0 & 0.2483 & 0.3817 \end{bmatrix} = \begin{bmatrix} 0.8147 & 0.0975 & 0.1576 \\ 1.4390 & 1.2553 & 1.3552 \\ 0 & 0.6585 & -0.1513 \\ 0 & -0.3916 & -0.8539 \\ 0 & 0.2483 & 0.3817 \end{bmatrix}.$$

To complete the first column, we compute a Givens rotation that, when applied to a_{11} and a_{21}, zeros a_{21}:

$$\begin{bmatrix} 0.4927 & -0.8702 \\ 0.8702 & 0.4927 \end{bmatrix}^T \begin{bmatrix} 0.8147 \\ 1.4390 \end{bmatrix} = \begin{bmatrix} 1.6536 \\ 0 \end{bmatrix}.$$

Applying this rotation to rows 1 and 2 yields

$$\begin{bmatrix} 0.4927 & -0.8702 & 0 & 0 & 0 \\ 0.8702 & 0.4927 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}^T \begin{bmatrix} 0.8147 & 0.0975 & 0.1576 \\ 1.4390 & 1.2553 & 1.3552 \\ 0 & 0.6585 & -0.1513 \\ 0 & -0.3916 & -0.8539 \\ 0 & 0.2483 & 0.3817 \end{bmatrix} = \begin{bmatrix} 1.6536 & 1.1405 & 1.2569 \\ 0 & 0.5336 & 0.5305 \\ 0 & 0.6585 & -0.1513 \\ 0 & -0.3916 & -0.8539 \\ 0 & 0.2483 & 0.3817 \end{bmatrix}.$$

Moving to the second column, we compute a Givens rotation that, when applied to a_{42} and a_{52}, zeros a_{52}:

$$\begin{bmatrix} 0.8445 & 0.5355 \\ -0.5355 & 0.8445 \end{bmatrix}^T \begin{bmatrix} -0.3916 \\ 0.2483 \end{bmatrix} = \begin{bmatrix} 0.4636 \\ 0 \end{bmatrix}.$$

Applying this rotation to rows 4 and 5 yields

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0.8445 & 0.5355 \\ 0 & 0 & 0 & -0.5355 & 0.8445 \end{bmatrix}^T \begin{bmatrix} 1.6536 & 1.1405 & 1.2569 \\ 0 & 0.5336 & 0.5305 \\ 0 & 0.6585 & -0.1513 \\ 0 & -0.3916 & -0.8539 \\ 0 & 0.2483 & 0.3817 \end{bmatrix} = \begin{bmatrix} 1.6536 & 1.1405 & 1.2569 \\ 0 & 0.5336 & 0.5305 \\ 0 & 0.6585 & -0.1513 \\ 0 & -0.4636 & -0.9256 \\ 0 & 0 & -0.1349 \end{bmatrix}.$$

Next, we compute a Givens rotation that, when applied to a_{32} and a_{42}, zeros a_{42}:

$$\begin{bmatrix} 0.8177 & 0.5757 \\ -0.5757 & 0.8177 \end{bmatrix}^T \begin{bmatrix} 0.6585 \\ -0.4636 \end{bmatrix} = \begin{bmatrix} 0.8054 \\ 0 \end{bmatrix}.$$

Applying this rotation to rows 3 and 4 yields

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0.8177 & 0.5757 & 0 \\ 0 & 0 & -0.5757 & 0.8177 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}^T \begin{bmatrix} 1.6536 & 1.1405 & 1.2569 \\ 0 & 0.5336 & 0.5305 \\ 0 & 0.6585 & -0.1513 \\ 0 & -0.4636 & -0.9256 \\ 0 & 0 & -0.1349 \end{bmatrix} = \begin{bmatrix} 1.6536 & 1.1405 & 1.2569 \\ 0 & 0.5336 & 0.5305 \\ 0 & 0.8054 & 0.4091 \\ 0 & 0 & -0.8439 \\ 0 & 0 & -0.1349 \end{bmatrix}.$$

Next, we compute a Givens rotation that, when applied to a_{22} and a_{32}, zeros a_{32}:

$$\begin{bmatrix} 0.5523 & -0.8336 \\ 0.8336 & 0.5523 \end{bmatrix}^T \begin{bmatrix} 0.5336 \\ 0.8054 \end{bmatrix} = \begin{bmatrix} 0.9661 \\ 0 \end{bmatrix}.$$

Applying this rotation to rows 3 and 4 yields

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0.5523 & -0.8336 & 0 & 0 \\ 0 & 0.8336 & 0.5523 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}^T \begin{bmatrix} 1.6536 & 1.1405 & 1.2569 \\ 0 & 0.5336 & 0.5305 \\ 0 & 0.8054 & 0.4091 \\ 0 & 0 & -0.8439 \\ 0 & 0 & -0.1349 \end{bmatrix} = \begin{bmatrix} 1.6536 & 1.1405 & 1.2569 \\ 0 & 0.9661 & 0.6341 \\ 0 & 0 & -0.2163 \\ 0 & 0 & -0.8439 \\ 0 & 0 & -0.1349 \end{bmatrix}.$$

Moving to the third column, we compute a Givens rotation that, when applied to a_{43} and a_{53}, zeros a_{53}:

$$\begin{bmatrix} 0.9875 & -0.1579 \\ 0.1579 & 0.9875 \end{bmatrix}^T \begin{bmatrix} -0.8439 \\ -0.1349 \end{bmatrix} = \begin{bmatrix} 0.8546 \\ 0 \end{bmatrix}.$$

Applying this rotation to rows 4 and 5 yields

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0.9875 & -0.1579 \\ 0 & 0 & 0 & 0.1579 & 0.9875 \end{bmatrix}^T \begin{bmatrix} 1.6536 & 1.1405 & 1.2569 \\ 0 & 0.9661 & 0.6341 \\ 0 & 0 & -0.2163 \\ 0 & 0 & -0.8439 \\ 0 & 0 & -0.1349 \end{bmatrix} = \begin{bmatrix} 1.6536 & 1.1405 & 1.2569 \\ 0 & 0.9661 & 0.6341 \\ 0 & 0 & -0.2163 \\ 0 & 0 & -0.8546 \\ 0 & 0 & 0 \end{bmatrix}.$$

Finally, we compute a Givens rotation that, when applied to a_{33} and a_{43}, zeros a_{43}:

$$\begin{bmatrix} 0.2453 & -0.9694 \\ 0.9694 & 0.2453 \end{bmatrix}^T \begin{bmatrix} -0.2163 \\ -0.8546 \end{bmatrix} = \begin{bmatrix} 0.8816 \\ 0 \end{bmatrix}.$$

Applying this rotation to rows 3 and 4 yields

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0.2453 & -0.9694 & 0 \\ 0 & 0 & 0.9694 & 0.2453 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}^T \begin{bmatrix} 1.6536 & 1.1405 & 1.2569 \\ 0 & 0.9661 & 0.6341 \\ 0 & 0 & -0.2163 \\ 0 & 0 & -0.8546 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1.6536 & 1.1405 & 1.2569 \\ 0 & 0.9661 & 0.6341 \\ 0 & 0 & -0.8816 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = R.$$

Applying the transpose of each Givens rotation, in order, to the *columns* of the identity matrix yields the matrix

$$Q = \begin{bmatrix} 0.4927 & -0.4806 & 0.1780 & -0.7033 & 0 \\ 0.5478 & -0.3583 & -0.5777 & 0.4825 & 0.0706 \\ 0.0768 & 0.4754 & -0.6343 & -0.4317 & -0.4235 \\ 0.5523 & 0.3391 & 0.4808 & 0.2769 & -0.5216 \\ 0.3824 & 0.5473 & 0.0311 & -0.0983 & 0.7373 \end{bmatrix}$$

such that $Q^TA = R$ is upper triangular. □

Householder Reflections

It is natural to ask whether we can introduce more zeros with each orthogonal rotation. To that end, we examine *Householder reflections*. Consider a matrix of the form $P = I - \tau\mathbf{u}\mathbf{u}^T$, where $\mathbf{u} \neq \mathbf{0}$ and τ is a nonzero constant. It is clear that P is a symmetric rank-1 change of I. Can we choose τ so that P is also orthogonal? From the desired relation $P^TP = I$ we obtain

$$\begin{aligned} P^TP &= (I - \tau\mathbf{u}\mathbf{u}^T)^T(I - \tau\mathbf{u}\mathbf{u}^T) \\ &= I - 2\tau\mathbf{u}\mathbf{u}^T + \tau^2\mathbf{u}\mathbf{u}^T\mathbf{u}\mathbf{u}^T \\ &= I - 2\tau\mathbf{u}\mathbf{u}^T + \tau^2(\mathbf{u}^T\mathbf{u})\mathbf{u}\mathbf{u}^T \\ &= I + (\tau^2\mathbf{u}^T\mathbf{u} - 2\tau)\mathbf{u}\mathbf{u}^T \\ &= I + \tau(\tau\mathbf{u}^T\mathbf{u} - 2)\mathbf{u}\mathbf{u}^T. \end{aligned}$$

It follows that if $\tau = 2/\mathbf{u}^T\mathbf{u}$, then $P^TP = I$ for any nonzero $\mathbf{u}$. Without loss of generality, we can stipulate that $\mathbf{u}^T\mathbf{u} = 1$, and therefore P takes the form $P = I - 2\mathbf{v}\mathbf{v}^T$, where $\mathbf{v}^T\mathbf{v} = 1$.

Why is the matrix P called a reflection? This is because for any nonzero vector $\mathbf{x}$, $P\mathbf{x}$ is the reflection of $\mathbf{x}$ across the hyperplane that is normal to $\mathbf{v}$. To see this, we consider the 2×2 case and set $\mathbf{v} = \begin{bmatrix} 1 & 0 \end{bmatrix}^T$ and $\mathbf{x} = \begin{bmatrix} 1 & 2 \end{bmatrix}^T$. Then

$$\begin{aligned} P &= I - 2\mathbf{v}\mathbf{v}^T \\ &= I - 2\begin{bmatrix} 1 \\ 0 \end{bmatrix}\begin{bmatrix} 1 & 0 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - 2\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \\ &= \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \end{aligned}$$

Therefore

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

Now, let $\mathbf{x}$ be any vector. We wish to construct P so that $P\mathbf{x} = \alpha\mathbf{e}_1$ for some α, where $\mathbf{e}_1 = \begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix}^T$. From the relations

$$\|P\mathbf{x}\|_2 = \|\mathbf{x}\|_2, \quad \|\alpha\mathbf{e}_1\|_2 = |\alpha|\|\mathbf{e}_1\|_2 = |\alpha|,$$

we obtain $\alpha = \pm\|\mathbf{x}\|_2$. To determine P, we begin with the equation

$$P\mathbf{x} = (I - 2\mathbf{v}\mathbf{v}^T)\mathbf{x} = \mathbf{x} - 2\mathbf{v}\mathbf{v}^T\mathbf{x} = \alpha\mathbf{e}_1.$$

Rearranging, we obtain

$$\frac{1}{2}(\mathbf{x} - \alpha\mathbf{e}_1) = (\mathbf{v}^T\mathbf{x})\mathbf{v}.$$

It follows that the vector $\mathbf{v}$, which is a unit vector, must be a scalar multiple of $\mathbf{x} - \alpha\mathbf{e}_1$. Therefore, $\mathbf{v}$ is defined by the equations

$$\begin{aligned}
v_1 &= \frac{x_1 - \alpha}{\|\mathbf{x} - \alpha\mathbf{e}_1\|_2} \\
&= \frac{x_1 - \alpha}{\sqrt{\|\mathbf{x}\|_2^2 - 2\alpha x_1 + \alpha^2}} \\
&= \frac{x_1 - \alpha}{\sqrt{2\alpha^2 - 2\alpha x_1}} \\
&= -\frac{\alpha - x_1}{\sqrt{2\alpha(\alpha - x_1)}} \\
&= -\mathrm{sgn}(\alpha)\sqrt{\frac{\alpha - x_1}{2\alpha}}, \\
v_2 &= \frac{x_2}{\sqrt{2\alpha(\alpha - x_1)}} \\
&= -\frac{x_2}{2\alpha v_1}, \\
&\vdots \\
v_n &= -\frac{x_n}{2\alpha v_1}.
\end{aligned}$$

To avoid catastrophic cancellation, it is best to choose the sign of α so that it has the opposite sign of x_1. It can be seen that the computation of $\mathbf{v}$ requires about $3n$ operations.

Note that the matrix P is never formed explicitly. For any vector $\mathbf{b}$, the product $P\mathbf{b}$ can be computed as follows:

$$P\mathbf{b} = (I - 2\mathbf{v}\mathbf{v}^T)\mathbf{b} = \mathbf{b} - 2(\mathbf{v}^T\mathbf{b})\mathbf{v}.$$

This process requires only $4n$ operations. It is easy to see that we can represent P simply by storing only $\mathbf{v}$.

Now, suppose that that $\mathbf{x} = \mathbf{a}_1$ is the first column of a matrix A. Then we construct a Householder reflection $H_1 = I - 2\mathbf{v}_1\mathbf{v}_1^T$ such that $H\mathbf{x} = \alpha\mathbf{e}_1$, and we have

$$A^{(2)} = H_1 A = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ 0 & & & \\ \vdots & \mathbf{a}^{(2)}_{2:m,2} & \cdots & \mathbf{a}^{(2)}_{2:m,n} \\ 0 & & & \end{bmatrix}.$$

where we denote the constant α by r_{11}, as it is the $(1,1)$ element of the updated matrix $A^{(2)}$. Now, we can construct $\tilde{H}_2$ such that

$$\tilde{H}_2\mathbf{a}^{(2)}_{2:m,2} = \begin{bmatrix} r_{22} \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad A^{(3)} = \begin{bmatrix} 1 & 0 \\ 0 & \tilde{H}_2 \end{bmatrix} A^{(2)} = \begin{bmatrix} r_{11} & r_{12} & r_{13} & \cdots & r_{1n} \\ 0 & r_{22} & r_{23} & \cdots & r_{2n} \\ 0 & 0 & & & \\ \vdots & \vdots & \mathbf{a}^{(3)}_{3:m,3} & \cdots & \mathbf{a}^{(3)}_{3:m,n} \\ 0 & 0 & & & \end{bmatrix}.$$

Note that the first column of $A^{(2)}$ is unchanged by $\tilde{H}_2$, because $\tilde{H}_2$ only operates on rows 2 through m, which, in the first column, have zero entries. Continuing this process, we obtain

$$H_n \cdots H_1 A = A^{(n+1)} = R,$$

where, for $j = 1, 2, \ldots, n$,

$$H_j = \begin{bmatrix} I_{j-1} & 0 \\ 0 & \tilde{H}_j \end{bmatrix}$$

and R is an upper triangular matrix. We have thus factored $A = QR$, where $Q = H_1 H_2 \cdots H_n$ is an orthogonal matrix.

Note that for each $j = 1, 2, \ldots, n$, $\tilde{H}_j$ is also a Householder reflection, based on a vector whose first $j - 1$ components are equal to zero. Therefore, application of H_j to a matrix does not affect the first j rows or columns. We also note that

$$A^T A = R^T Q^T Q R = R^T R,$$

and thus R is the Cholesky factor of $A^T A$.

Example We apply Householder reflections to compute the QR factorization of the matrix from the previous example,

$$A^{(1)} = A = \begin{bmatrix} 0.8147 & 0.0975 & 0.1576 \\ 0.9058 & 0.2785 & 0.9706 \\ 0.1270 & 0.5469 & 0.9572 \\ 0.9134 & 0.9575 & 0.4854 \\ 0.6324 & 0.9649 & 0.8003 \end{bmatrix}.$$

First, we work with the first column of A,

$$\mathbf{x}_1 = \mathbf{a}^{(1)}_{1:5,1} = \begin{bmatrix} 0.8147 \\ 0.9058 \\ 0.1270 \\ 0.9134 \\ 0.6324 \end{bmatrix}, \quad \|\mathbf{x}_1\|_2 = 1.6536.$$

The corresponding Householder vector is

$$\tilde{\mathbf{v}}_1 = \mathbf{x}_1 + \|\mathbf{x}_1\|_2 \mathbf{e}_1 = \begin{bmatrix} 0.8147 \\ 0.9058 \\ 0.1270 \\ 0.9134 \\ 0.6324 \end{bmatrix} + 1.6536 \begin{bmatrix} 1.0000 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 2.4684 \\ 0.9058 \\ 0.1270 \\ 0.9134 \\ 0.6324 \end{bmatrix}.$$

From this vector, we build the Householder reflection

$$c = \frac{2}{\tilde{\mathbf{v}}_1^T \tilde{\mathbf{v}}_1} = 0.2450, \quad \tilde{H}_1 = I - c\tilde{\mathbf{v}}_1 \tilde{\mathbf{v}}_1^T.$$

Applying this reflection to $A^{(1)}$ yields

$$\tilde{H}_1 A^{(1)}_{1:5,1:3} = \begin{bmatrix} -1.6536 & -1.1405 & -1.2569 \\ 0 & -0.1758 & 0.4515 \\ 0 & 0.4832 & 0.8844 \\ 0 & 0.4994 & -0.0381 \\ 0 & 0.6477 & 0.4379 \end{bmatrix}, \quad A^{(2)} = \begin{bmatrix} -1.6536 & -1.1405 & -1.2569 \\ 0 & -0.1758 & 0.4515 \\ 0 & 0.4832 & 0.8844 \\ 0 & 0.4994 & -0.0381 \\ 0 & 0.6477 & 0.4379 \end{bmatrix}.$$

Next, we take the "interesting" portion of the second column of the updated matrix $A^{(2)}$, from rows 2 to 5:

$$\mathbf{x}_2 = \mathbf{a}^{(2)}_{2:5,2} = \begin{bmatrix} -0.1758 \\ 0.4832 \\ 0.4994 \\ 0.6477 \end{bmatrix}, \quad \|\mathbf{x}_2\|_2 = 0.9661.$$

The corresponding Householder vector is

$$\tilde{\mathbf{v}}_2 = \mathbf{x}_2 - \|\mathbf{x}_2\|_2 \mathbf{e}_1 = \begin{bmatrix} -0.1758 \\ 0.4832 \\ 0.4994 \\ 0.6477 \end{bmatrix} - 0.9661 \begin{bmatrix} 1.0000 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} -1.1419 \\ 0.4832 \\ 0.4994 \\ 0.6477 \end{bmatrix}.$$

From this vector, we build the Householder reflection

$$c = \frac{2}{\tilde{\mathbf{v}}_2^T \tilde{\mathbf{v}}_2} = 0.9065, \quad \tilde{H}_2 = I - c\tilde{\mathbf{v}}_2 \tilde{\mathbf{v}}_2^T.$$

Applying this reflection to $A^{(2)}$ yields

$$\tilde{H}_2 A^{(2)}_{2:5,2:3} = \begin{bmatrix} 0.9661 & 0.6341 \\ 0 & 0.8071 \\ 0 & -0.1179 \\ 0 & 0.3343 \end{bmatrix}, \quad A^{(3)} = \begin{bmatrix} -1.6536 & -1.1405 & -1.2569 \\ 0 & 0.9661 & 0.6341 \\ 0 & 0 & 0.8071 \\ 0 & 0 & -0.1179 \\ 0 & 0 & 0.3343 \end{bmatrix}.$$

Finally, we take the interesting portion of the third column of $A^{(3)}$, from rows 3 to 5:

$$\mathbf{x}_3 = \mathbf{a}^{(3)}_{3:5,3} = \begin{bmatrix} 0.8071 \\ -0.1179 \\ 0.3343 \end{bmatrix}, \quad \|\mathbf{x}_3\|_2 = 0.8816.$$

The corresponding Householder vector is

$$\tilde{\mathbf{v}}_3 = \mathbf{x}_3 + \|\mathbf{x}_3\|_2 \mathbf{e}_1 = \begin{bmatrix} 0.8071 \\ -0.1179 \\ 0.3343 \end{bmatrix} + 0.8816 \begin{bmatrix} 1.0000 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1.6887 \\ -0.1179 \\ 0.3343 \end{bmatrix}.$$

From this vector, we build the Householder reflection

$$c = \frac{2}{\tilde{\mathbf{v}}_3^T \tilde{\mathbf{v}}_3} = 0.6717, \quad \tilde{H}_3 = I - c\tilde{\mathbf{v}}_3 \tilde{\mathbf{v}}_3^T.$$

Applying this reflection to $A^{(3)}$ yields

$$\tilde{H}_3 A^{(3)}_{3:5,3:3} = \begin{bmatrix} -0.8816 \\ 0 \\ 0 \end{bmatrix}, \quad R = A^{(4)} = \begin{bmatrix} -1.6536 & -1.1405 & -1.2569 \\ 0 & 0.9661 & 0.6341 \\ 0 & 0 & -0.8816 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Applying these same Householder reflections, in order, on the *right* of the identity matrix, yields the orthogonal matrix

$$Q = H_1 H_2 H_3 = \begin{bmatrix} -0.4927 & -0.4806 & 0.1780 & -0.6015 & -0.3644 \\ -0.5478 & -0.3583 & -0.5777 & 0.3760 & 0.3104 \\ -0.0768 & 0.4754 & -0.6343 & -0.1497 & -0.5859 \\ -0.5523 & 0.3391 & 0.4808 & 0.5071 & -0.3026 \\ -0.3824 & 0.5473 & 0.0311 & -0.4661 & 0.5796 \end{bmatrix}$$

such that

$$R = Q^T A = H_3 H_2 H_1 A = \begin{bmatrix} -1.6536 & -1.1405 & -1.2569 \\ 0 & 0.9661 & 0.6341 \\ 0 & 0 & -0.8816 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

is upper triangular, where

$$H_1 = \tilde{H}_1, \quad H_2 = \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \tilde{H}_2 \end{bmatrix}, \quad H_3 = \begin{bmatrix} 1 & 0 & \mathbf{0} \\ 0 & 1 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \tilde{H}_3 \end{bmatrix},$$

are the same Householder transformations before, defined in such a way that they can be applied to the *entire* matrix A. Note that for $j = 1, 2, 3,$,

$$H_j = I - 2\mathbf{v}_j \mathbf{v}_j^T, \quad \mathbf{v}_j = \begin{bmatrix} \mathbf{0} \\ \tilde{\mathbf{v}}_j \end{bmatrix}, \quad \|\mathbf{v}_j\|_2 = \|\tilde{\mathbf{v}}_j\|_2 = 1,$$

where the first $j - 1$ components of $\mathbf{v}_j$ are equal to zero.

Also, note that the first $n = 3$ columns of Q are the same as those of the matrix Q that was computed in the previous example, except for possible negation. The fourth and fifth columns are not the same, but they do span the same subspace, the orthogonal complement of the range of A. □

Givens Rotations vs. Householder Reflections

We showed how to construct Givens rotations in order to rotate two elements of a column vector so that one element would be zero, and that approximately $n^2/2$ such rotations could be used to transform A into an upper triangular matrix R. Because each rotation only modifies two rows of A, it is possible to interchange the order of rotations that affect different rows, and thus apply sets of rotations in parallel. This is the main reason why Givens rotations can be preferable to Householder reflections. Other reasons are that they are easy to use when the QR factorization needs to be updated as a result of adding a row to A or deleting a column of A. They are also more efficient when A is sparse.