For more comprehensive treatments of the least squares problem, see Björck (NMLS) and Lawson and Hansen (SLS). Other useful global references include Stewart (MABD), Higham (ASNA), Watkins (FMC), Trefethen and Bau (NLA), Demmel (ANLA), and Ipsen (NMA).

5.1 Householder and Givens Transformations

Recall that $Q \in \mathbb{R}^{m\times m}$ is *orthogonal* if

$$Q^TQ = QQ^T = I_m.$$

Orthogonal matrices have an important role to play in least squares and eigenvalue computations. In this section we introduce Householder reflections and Givens rotations, the key players in this game.

5.1.1 A 2-by-2 Preview

It is instructive to examine the geometry associated with rotations and reflections at the $m = 2$ level. A 2-by-2 orthogonal matrix Q is a *rotation* if it has the form

$$Q \;=\; \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix}.$$

If $y = Q^Tx$, then y is obtained by rotating x counterclockwise through an angle θ.

A 2-by-2 orthogonal matrix Q is a *reflection* if it has the form

$$Q \;=\; \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ \sin(\theta) & -\cos(\theta) \end{bmatrix}.$$

If $y = Q^Tx = Qx$, then y is obtained by reflecting the vector x across the line defined by

$$S = \mathsf{span}\left\{ \begin{bmatrix} \cos(\theta/2) \\ \sin(\theta/2) \end{bmatrix} \right\}.$$

Reflections and rotations are computationally attractive because they are easily constructed and because they can be used to introduce zeros in a vector by properly choosing the rotation angle or the reflection plane.

5.1.2 Householder Reflections

Let $v \in \mathbb{R}^m$ be nonzero. An m-by-m matrix P of the form

$$P = I - \beta vv^T, \qquad \beta = \frac{2}{v^Tv} \tag{5.1.1}$$

is a *Householder reflection.* (Synonyms are Householder matrix and Householder transformation.) The vector v is the *Householder vector.* If a vector x is multiplied by P, then it is reflected in the hyperplane $\mathsf{span}\{v\}^\perp$. It is easy to verify that Householder matrices are symmetric and orthogonal.

Householder reflections are similar to Gauss transformations introduced in §3.2.1 in that they are rank-1 modifications of the identity and can be used to zero selected components of a vector. In particular, suppose we are given $0 \neq x \in \mathbb{R}^m$ and want

$$Px = \left(I - \frac{2vv^T}{v^Tv}\right)x = x - \frac{2v^Tx}{v^Tv}v$$

to be a multiple of $e_1 = I_m(:,1)$. From this we conclude that $v \in \text{span}\{x, e_1\}$. Setting

$$v = x + \alpha e_1$$

gives

$$v^Tx = x^Tx + \alpha x_1$$

and

$$v^Tv = x^Tx + 2\alpha x_1 + \alpha^2.$$

Thus,

$$\begin{aligned} Px &= \left(1 - 2\frac{x^Tx + \alpha x_1}{x^Tx + 2\alpha x_1 + \alpha^2}\right)x - 2\alpha\frac{v^Tx}{v^Tv}e_1 \\ &= \left(\frac{\alpha^2 - \| x \|_2^2}{x^Tx + 2\alpha x_1 + \alpha^2}\right)x - 2\alpha\frac{v^Tx}{v^Tv}e_1. \end{aligned}$$

In order for the coefficient of x to be zero, we set $\alpha = \pm\| x \|_2$ for then

$$v = x \pm \| x \|_2 e_1 \;\Rightarrow\; Px = \left(I - 2\frac{vv^T}{v^Tv}\right)x = \mp\| x \|_2 e_1. \tag{5.1.2}$$

It is this simple determination of v that makes the Householder reflections so useful.

5.1.3 Computing the Householder Vector

There are a number of important practical details associated with the determination of a Householder matrix, i.e., the determination of a Householder vector. One concerns the choice of sign in the definition of v in (5.1.2). Setting

$$v_1 = x_1 - \| x \|_2$$

leads to the nice property that Px is a positive multiple of e_1. But this recipe is dangerous if x is close to a positive multiple of e_1 because severe cancellation would occur. However, the formula

$$v_1 = x_1 - \| x \|_2 = \frac{x_1^2 - \| x \|_2^2}{x_1 + \| x \|_2} = \frac{-(x_2^2 + \cdots + x_n^2)}{x_1 + \| x \|_2}$$

suggested by Parlett (1971) does not suffer from this defect in the $x_1 > 0$ case.

In practice, it is handy to normalize the Householder vector so that $v(1) = 1$. This permits the storage of $v(2{:}m)$ where the zeros have been introduced in x, i.e., $x(2{:}m)$. We refer to $v(2{:}m)$ as the *essential part* of the Householder vector. Recalling that $\beta = 2/v^Tv$ and letting $\mathsf{length}(x)$ specify vector dimension, we may encapsulate the overall process as follows:

Algorithm 5.1.1 (Householder Vector) Given $x \in \mathbb{R}^m$, this function computes $v \in \mathbb{R}^m$ with $v(1) = 1$ and $\beta \in \mathbb{R}$ such that $P = I_m - \beta vv^T$ is orthogonal and $Px = \| x \|_2 e_1$.

function $[v, \beta] = \mathsf{house}(x)$

$m = \mathsf{length}(x),\ \sigma = x(2{:}m)^T x(2{:}m),\ v = \begin{bmatrix} 1 \\ x(2{:}m) \end{bmatrix}$

if $\sigma = 0$ and $x(1) >= 0$

$\beta = 0$

elseif $\sigma = 0$ & $x(1) < 0$

$\beta = -2$

else

$\mu = \sqrt{x(1)^2 + \sigma}$

if $x(1) <= 0$

$v(1) = x(1) - \mu$

else

$v(1) = -\sigma/(x(1) + \mu)$

end

$\beta = 2v(1)^2/(\sigma + v(1)^2)$

$v = v/v(1)$

end

Here, $\mathsf{length}(\cdot)$ returns the dimension of a vector. This algorithm involves about $3m$ flops. The computed Householder matrix that is orthogonal to machine precision, a concept discussed below.

5.1.4 Applying Householder Matrices

It is critical to exploit structure when applying $P = I - \beta vv^T$ to a matrix A. Premultiplication involves a matrix-vector product and a rank-1 update:

$$PA \;=\; \left(I - \beta vv^T\right) A \;=\; A - (\beta v)(v^T A).$$

The same is true for post-multiplication,

$$AP \;=\; A\left(I - \beta vv^T\right) \;=\; A - (Av)(\beta v)^T.$$

In either case, the update requires $4mn$ flops if $A \in \mathbb{R}^{m\times n}$. Failure to recognize this and to treat P as a general matrix increases work by an order of magnitude. *Householder updates never entail the explicit formation of the Householder matrix.*

In a typical situation, house is applied to a subcolumn or subrow of a matrix and $(I - \beta vv^T)$ is applied to a submatrix. For example, if $A \in \mathbb{R}^{m\times n}$, $1 \le j < n$, and $A(j{:}m, 1{:}j-1)$ is zero, then the sequence

$$\begin{aligned} [v, \beta] \;&=\; \mathsf{house}(A(j{:}m, j)) \\ A(j{:}m, j{:}n) \;&=\; A(j{:}m, j{:}n) - (\beta v)(v^T A(j{:}m, j{:}n)) \\ A(j+1{:}m, j) \;&=\; v(2{:}m-j+1) \end{aligned}$$