What Is a Householder Matrix?

Nicholas J. Higham[*]

September 15, 2020, updated November 11, 2020

A Householder matrix is an $n \times n$ orthogonal matrix of the form

$$P = I - \frac{2}{v^T v} v v^T, \qquad 0 \neq v \in \mathbb{R}^n.$$

It is easily verified that P is
- orthogonal $(P^T P = I)$,
- symmetric $(P^T = P)$,
- involutory $(P^2 = I$ that is, P is a square root of the identity matrix),

where the last property follows from the first two.

A Householder matrix is a rank-1 perturbation of the identity matrix and so all but one of its eigenvalues are 1. The eigensystem can be fully described as follows.
- P has an eigenvalue -1 with eigenvector v, since $Pv = -v$.
- P has $n-1$ eigenvalues 1 with eigenvectors any set of $n-1$ linearly independent vectors orthogonal to v, which can be taken to be mutually orthogonal: $Px = x$ for every such x.

P has trace $n-2$ and determinant -1, as can be derived directly or deduced from the facts that the trace is the sum of the eigenvalues and the determinant is the product of the eigenvalues.

For $n = 2$, a Householder matrix can be written as

$$P = \begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix}.$$

Simple examples of Householder matrices are obtained by choosing $v = e = [1, 1, \ldots, 1]^T$, for which $P = I - (2/n)ee^T$. For $n = 2, 3, 4, 5, 6$ we obtain the matrices

$$\begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}, \quad \frac{1}{3}\begin{bmatrix} 1 & -2 & -2 \\ -2 & 1 & -2 \\ -2 & -2 & 1 \end{bmatrix}, \quad \frac{1}{2}\begin{bmatrix} 1 & -1 & -1 & -1 \\ -1 & 1 & -1 & -1 \\ -1 & -1 & 1 & -1 \\ -1 & -1 & -1 & 1 \end{bmatrix},$$

$$\frac{1}{5}\begin{bmatrix} 3 & -2 & -2 & -2 & -2 \\ -2 & 3 & -2 & -2 & -2 \\ -2 & -2 & 3 & -2 & -2 \\ -2 & -2 & -2 & 3 & -2 \\ -2 & -2 & -2 & -2 & 3 \end{bmatrix}, \quad \frac{1}{3}\begin{bmatrix} 2 & -1 & -1 & -1 & -1 & -1 \\ -1 & 2 & -1 & -1 & -1 & -1 \\ -1 & -1 & 2 & -1 & -1 & -1 \\ -1 & -1 & -1 & 2 & -1 & -1 \\ -1 & -1 & -1 & -1 & 2 & -1 \\ -1 & -1 & -1 & -1 & -1 & 2 \end{bmatrix}.$$

[*]Department of Mathematics, University of Manchester, Manchester, M13 9PL, UK (nick.higham@manchester.ac.uk).

Note that the 4×4 matrix is $1/2$ times a Hadamard matrix.

Applying P to a vector x gives

$$Px = x - \left(\frac{2v^T x}{v^T v}\right) v.$$

This equation shows that P reflects x about the hyperplane $\operatorname{span}(v)^\perp$, as illustrated in the following diagram, which explains why P is sometimes called a Householder reflector. Another way of expressing this property is to write $x = \alpha v + z$, where z is orthogonal to v. Then $Px = -\alpha v + z$, so the component of x in the direction v has been reversed. If we take $v = e_i$, the ith unit vector, then $P = I - 2e_i e_i^T = \operatorname{diag}(1, 1, \ldots, -1, 1, \ldots, 1)$, which has -1 in the (i, i) position. In this case premultiplying a vector by P flips the sign of the ith component.
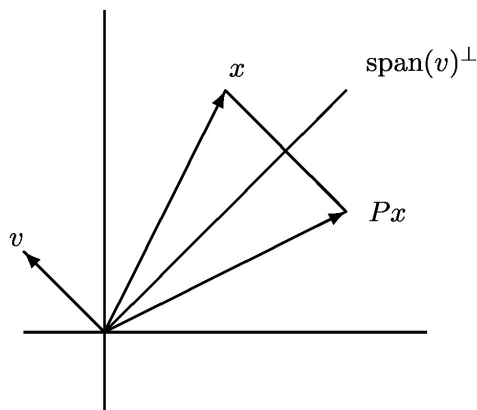
Transforming a Vector

Householder matrices are powerful tools for introducing zeros into vectors. Suppose we are given vectors x and y and wish to find a Householder matrix P such that $Px = y$. Since P is orthogonal, we require that $\|x\|_2 = \|y\|_2$, and we exclude the trivial case $x = y$. Now

$$Px = y \quad \Longleftrightarrow \quad x - 2\left(\frac{v^T x}{v^T v}\right) v = y,$$

and this last equation has the form $\alpha v = x - y$ for some α. But P is independent of the scaling of v, so we can set $\alpha = 1$. Now with $v = x - y$ we have

$$v^T v = x^T x + y^T y - 2x^T y$$

and, since $x^T x = y^T y$,

$$v^T x = x^T x - y^T x = \frac{1}{2} v^T v.$$

Therefore

$$Px = x - v = y,$$

as required. Most often we choose y to be zero in all but its first component.

2

Square Roots

What can we say about square roots of a Householder matrix, that is, matrices X such that $X^2 = P$?

We note first that the eigenvalues of X are the square roots of those of P and so $n-1$ of them will be ± 1 and one will be $\pm$i. This means that X cannot be real, as the nonreal eigenvalues of a real matrix must appear in complex conjugate pairs.

Write $P = I - 2vv^T$, where v is normalized so that $v^T v = 1$. It is natural to look for a square root of the form $X = I - \theta vv^T$. Setting $X^2 = P$ leads to the quadratic equation $\theta^2 - 2\theta + 2 = 0$, and hence $\theta = 1 \pm$ i. As expected, these two square roots are complex even though P is real. As an example, $\theta = 1 -$ i gives the following square root of the matrix above corresponding to $v = e/n^{1/2}$ with $n = 3$:

$$X = \frac{1}{3} \begin{bmatrix} 2+\mathrm{i} & -1+\mathrm{i} & -1+\mathrm{i} \\ -1+\mathrm{i} & 2+\mathrm{i} & -1+\mathrm{i} \\ -1+\mathrm{i} & -1+\mathrm{i} & 2+\mathrm{i} \end{bmatrix}.$$

A good way to understand all the square roots is to diagonalize P, which can be done by a similarity transformation with a Householder matrix! Normalizing $v^T v = 1$ again, let $w = v - e_1$ and $H = I - 2ww^T/(w^T w)$. Then from the construction above we know that $Hv = e_1$. Hence

$$H^T P H = HPH = I - 2Hvv^T H = I - 2e_1 e_1^T = \mathrm{diag}(-1, 1, 1, \ldots, 1) =: D.$$

Then $P = HDH^T$ and so $X = H\sqrt{D}H^T$ gives 2^n square roots on taking all possible combinations of signs on the diagonal for $\sqrt{D}$. Because P has repeated eigenvalues these are not the only square roots. The infinitely many others are obtained by taking non-diagonal square roots of D, which are of the form $\mathrm{diag}(\pm i, Y)$, where Y is any non-diagonal square root of the $(n-1) \times (n-1)$ identity matrix, which in particular could be a Householder matrix!

Block Householder Matrix

It is possible to define an $n \times n$ block Householder matrix in terms of a given $Z \in \mathbb{R}^{n \times p}$, where $n \geq p$, as
$$P = I - 2Z(Z^T Z)^+ Z^T.$$

Here, "+" denotes the Moore–Penrose pseudoinverse. For $p = 1$, P clearly reduces to a standard Householder matrix. It can be shown that $(Z^T Z)^+ Z^T = Z^+$ (this is most easily proved using the SVD), and so

$$P = I - 2ZZ^+ = I - 2P_Z,$$

where $P_Z = ZZ^+$ is the orthogonal projector onto the range of Z (that is, range(P_Z) = range(Z), $P_Z^2 = P_Z$, and $P_Z = P_Z^T$). Hence, like a standard Householder matrix, P is symmetric, orthogonal, and involutory. Furthermore, premultiplication of a matrix by P has the effect of reversing the component in the range of Z.

As an example, here is the block Householder matrix corresponding to $Z = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \end{bmatrix}^T$:

$$\frac{1}{5} \begin{bmatrix} -2 & -4 & -1 & 2 \\ -4 & 2 & -2 & -1 \\ -1 & -2 & 2 & -4 \\ 2 & -1 & -4 & -2 \end{bmatrix}.$$

One can show (using the SVD again) that the eigenvalues of P are -1 repeated r times and 1 repeated $n - r$ times, where $r = \text{rank}(Z)$. Hence $\text{trace}(P) = n - 2r$ and $\det(P) = (-1)^r$.

Schreiber and Parlett (1988) note the representation for $n = 2k$,

$$P = \pm\text{diag}(Q_1, Q_2) \begin{bmatrix} \cos(2\Theta) & \sin(2\Theta) \\ \sin(2\Theta) & -\cos(2\Theta) \end{bmatrix} \text{diag}(Q_1, Q_2)^T,$$

where Q_1 and Q_2 are orthogonal and Θ is symmetric positive definite. This formula neatly generalizes the formula for a standard Householder matrix for $n = 2$ given above, and a similar formula holds for odd n.

Schreiber and Parlett also show how given $E \in \mathbb{R}^{n \times p}$ ($n > p$) one can construct a block Householder matrix H such that

$$HE = \begin{bmatrix} F \\ 0 \end{bmatrix}, \qquad F \in \mathbb{R}^{p \times p}.$$

The polar decomposition plays a key role in the theory and algorithms for such H.

Rectangular Householder Matrix

We can define a rectangular Householder matrix as follows. Let $m > n$, $u \in \mathbb{R}^n$, $v \in \mathbb{R}^{m-n}$, and

$$P = \begin{bmatrix} I_n \\ 0 \end{bmatrix} + \alpha \begin{bmatrix} u \\ v \end{bmatrix} u^T = \begin{bmatrix} I_n + \alpha u u^T \\ \alpha v u^T \end{bmatrix} \in \mathbb{R}^n.$$

Then $P^T P = I$, that is, P has orthonormal columns, if

$$\alpha = \frac{-2}{u^T u + v^T v}.$$

Of course, P is just the first n columns of the Householder matrix built from the vector $[u^T \ v^T]^T$.

Historical Note

The earliest appearance of Householder matrices is in the book by Turnbull and Aitken (1932). These authors show that if $\|x\|_2 = \|y\|_2$ ($x \neq -y$) then a unitary matrix of the form $R = \alpha z z^* - I$ (in their notation) can be constructed so that $Rx = y$. They use this result to prove the existence of the Schur decomposition. The first systematic use of Householder matrices for computational purposes was by Householder (1958) who used them to construct the QR factorization.

References

This is a minimal set of references, which contain further useful references within.

- Massimiliano Fasi and Nicholas J. Higham, Generating Extreme-Scale Matrices with Specified Singular Values or Condition Numbers, MIMS EPrint 2020.8, Manchester Institute for Mathematical Sciences, The University of Manchester, UK, March 2020. (For the use of rectangular Householder matrices.)
- Nicholas J. Higham, Accuracy and Stability of Numerical Algorithms, second edition, Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2002. (Chapter 19.)
- Nicholas J. Higham, Functions of Matrices: Theory and Computation, Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2008. (Sections 1.5 and 1.6 for the theory of matrix square roots.)
- Robert S. Schreiber and Beresford N. Parlett, Block Reflectors: Theory and Computation, SIAM J. Numer. Anal. 25(1), 189–205, 1988.

Related Blog Posts

- What Is a Generalized Inverse? (2020)
- What Is a Hadamard Matrix? (2020)
- What Is an Orthogonal Matrix? (2020)
- What Is the Polar Decomposition? (2020)

This article is part of the "What Is" series, available from `https://nhigham.com/category/what-is` and in PDF form from the GitHub repository `https://github.com/higham/what-is`.