Appendix B

IN *FAST RELIABLE ALGORITHMS FOR MATRICES WITH STRUCTURE*, T. KAILATH

AND A. H. SAYED, EDS., APP. B, PP. 309–319, SIAM, PA, 1999.

ELEMENTARY TRANSFORMATIONS

Thomas Kailath

Ali H. Sayed

In this appendix we review three families of elementary (unitary and hyperbolic) transformations that can be used to annihilate selected entries in a vector and thereby reduce a matrix to triangular form. These are the Householder, Givens, and fast Givens transformations. For more details, and historical background, see [Hig96]. Special care needs to be taken when dealing with complex-valued data as compared to real-valued data, as we show in the sequel.

B.1 ELEMENTARY HOUSEHOLDER TRANSFORMATIONS

Suppose we wish to simultaneously annihilate several entries in a row vector, say to transform an $n-$dimensional vector $x = \begin{bmatrix} x_1 & x_2 & \ldots & x_{n-1} \end{bmatrix}$ to the form $\begin{bmatrix} \alpha & 0 & 0 & 0 \end{bmatrix}$, where, for general complex data, the resulting α may be complex as well.

One way to achieve this transformation is to employ a so-called *Householder reflection* Θ: it takes a row vector x and aligns it along the direction of the basis vector $e_0 = \begin{bmatrix} 1 & 0 & \ldots & 0 \end{bmatrix}$. More precisely, it performs the transformation

$$\begin{bmatrix} x_1 & x_2 & \ldots & x_{n-1} \end{bmatrix} \Theta = \alpha\, e_0 \ , \tag{B.1.1}$$

for some α to be determined. Since, as we shall promptly verify, the transformation Θ that we shall employ is unitary and also Hermitian (*i.e.,* $\Theta\Theta^* = I = \Theta^*\Theta$ and $\Theta = \Theta^*$), we can be more specific about the resulting α. In particular, it follows from (B.1.1) that the magnitude of α must be equal to $\|x\|$, *i.e.,* $|\alpha| = \|x\|$. This is because $x\Theta\Theta^* x^* = \|x\|^2 = |\alpha|^2$. Moreover, it also follows from (B.1.1) that $x\Theta x^* = \alpha x_1^*$. But since Θ will be Hermitian, we conclude that $x\Theta x^*$ is a real number and, hence, αx_1^* must be real as well.

This means that by rotating a vector x with a unitary and Hermitian transformation Θ we can achieve a post-array of the form $\begin{bmatrix} \alpha & 0 & \ldots & 0 \end{bmatrix}$, where α will in general be a complex number whose magnitude is the norm $\|x\|$ and whose phase is such that αx_1^* is real. For example, $\alpha = \pm\|x\|e^{j\phi_{x_1}}$ are possible values for α, where ϕ_{x_1} denotes the phase of x_1. [For real data, $\pm\|x\|$ are the possible values for α.]

Now, assume we define

$$\Theta \ \stackrel{\Delta}{=}\ I - 2\,\frac{g^* g}{g g^*} \quad \text{where} \quad g = x + \alpha\, e_0 \ , \tag{B.1.2}$$

and α is a complex number that is chosen as above, say $\alpha = \pm\|x\|e^{j\phi_{x_1}}$. It can be verified by direct calculation that, for any g, Θ is a unitary matrix, *i.e.,* $\Theta\Theta^* = I = \Theta^*\Theta$. It is also Hermitian.

Lemma B.1.1 (Complex Householder Transformation) *Given a row vector x with leading entry x_1, define Θ and g as in (B.1.2) where α is any complex number that satisfies the following two requirements: $|\alpha| = \|x\|$ and αx_1^* is real. Then it holds that $x\Theta = -\alpha e_0$. That is, x is rotated and aligned with e_0; the leading entry of the post-array is equal to $-\alpha$.*

(**Algebraic**) **proof**: We shall provide a geometric proof below. Here we verify our claim algebraically. Indeed, direct calculation shows that

$$\begin{aligned} gg^* &= \|x\|^2 + \alpha^* x_1 + \alpha x_1^* + |\alpha|^2 \ , \\ &= 2\|x\|^2 + 2\alpha x_1^* \ , \qquad \text{since } |\alpha|^2 = \|x\|^2 \text{ and } \alpha^* x_1 = \alpha x_1^*. \end{aligned}$$

Likewise,

$$\begin{aligned} xg^*g &= x\|x\|^2 + \alpha\|x\|^2 e_0 + \alpha^* x_1 x + |\alpha|^2 x_1 e_0 \ , \\ &= x\|x\|^2 + \alpha\|x\|^2 e_0 + \alpha^* x_1 x + \alpha(\alpha^* x_1) e_0 \ , \\ &= x\|x\|^2 + \alpha\|x\|^2 e_0 + \alpha^* x_1 x + \alpha(\alpha x_1^*) e_0 \ , \\ xgg^* &= 2x\|x\|^2 + 2\alpha x_1^* x \ , \end{aligned}$$

and we obtain

$$x\Theta = \frac{xgg^* - 2xg^*g}{gg^*} = \frac{-(2\|x\|^2 + 2\alpha x_1^*)\alpha e_0}{2\|x\|^2 + 2\alpha x_1^*} \ = \ -\alpha e_0 \ .$$

◊

In other words, by defining

$$g \ \stackrel{\Delta}{=} \ x \pm e^{j\phi_{x_1}} \|x\| \ e_0 \ , \tag{B.1.3}$$

we obtain

$$x\Theta = \mp e^{j\phi_{x_1}} \|x\| e_0. \tag{B.1.4}$$

The choice of the sign in (B.1.4) depends on the choice of the sign in the expression for g. Usually, the sign in the expression for g is chosen so as to avoid a vector g of small Euclidean norm, since this norm appears in the denominator of the expression defining Θ. [In the real case, this can be guaranteed by choosing the sign in the expression for g to be the same as the sign of the leading entry of the row vector x, *viz.*, the sign of x_1.]

A Geometric Derivation

The result of the above lemma has a simple geometric interpretation: given a vector x, we would like to rotate it and align it with the vector e_0 (the careful reader will soon realize that the argument applies equally well to alignments along other vector directions). This rotation should keep the norm of x unchanged. Hence, the tip of the vector x should be rotated along a circular trajectory until it becomes aligned with e_0. The vector aligned with e_0 is equal to αe_0. Here, as indicated in Fig. B.1, we are assuming that the rotation is performed in the clockwise direction. The triangle with sides x and αe_0 and base $g = x - \alpha e_0$ is then an isosceles triangle.

We thus have $\alpha e_0 = x - g$. But we can also express this in an alternative form. If we drop a perpendicular (denoted by $g^\perp$) from the origin of x to the vector g, it will divide g into two equal parts. Moreover, the upper part is nothing but the projection of the vector x onto the vector g and is thus equal to $\langle x, g\rangle \ ||g||^{-2} \ g$. Therefore,

$$\alpha e_0 = x - 2xg^*(gg^*)^{-1}g = x \underbrace{\left[I - 2\frac{g^*g}{gg^*}\right]}_{\Theta} = x - 2 < x, g > \ ||g||^{-2} \ g.$$

A similar argument holds for the choice $g = \alpha e_0 + x$. The Householder transformation *reflects* the vector x across the line $g^\perp$ to the vector αe_0. So it is often called a Householder reflection.

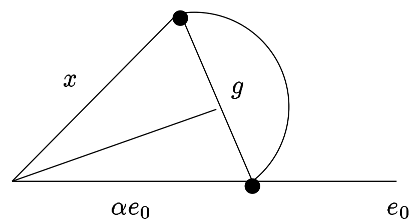

Figure B.1. Geometric interpretation of the Householder transformation.

To represent the transformation in matrix form we have two choices, depending upon whether we represent vectors as row $(1\times n)$ matrices or column $(n\times 1)$ matrices. In the first case, we have $\langle x, g\rangle = xg^*$ and $||g||^2 = gg^*$, so that

$$\alpha e_0 = x - 2(xg^*)(gg^*)^{-1}g = x\left(I - 2\frac{g^*g}{gg^*}\right) ,$$

as in (B.1.4) earlier.

If we represent vectors by columns, then $\langle x, g\rangle = g^*x$ and $||g||^2 = g^*g$, and we shall have

$$\alpha e_0 = x - 2(g^*x)(g^*g)^{-1}g = \left(I - 2\frac{gg^*}{g^*g}\right) x .$$

Triangularizing a Matrix

A sequence of Householder transformations of this type can be used to triangularize a given $m \times n$ matrix, say A. For this we first find a transformation Θ_0 to rotate the first row to lie along e_0, so that we have $A\Theta_0$ of the form (where α denotes entries whose exact values are not of current interest):

$$A\Theta_0 = \left[\begin{array}{c|cccc} \sigma_1 & 0 & 0 & 0 & 0 \\ \hline \alpha & & & & \\ \alpha & & & A_1 & \\ \alpha & & & & \end{array}\right].$$

Now apply a transformation of the type $(1 \oplus \Theta_1)$ where Θ_1 rotates the first row of A_1 so that it lies along e_0 in an $(n-1)$-dimensional space, and so on. This so-called Householder reduction of matrices has been found to be an efficient and stable tool for displaying rank information via matrix triangularization and it is widely used in numerical analysis (see, *e.g.*, [GV96], [Hig96]).

B.2 ELEMENTARY CIRCULAR OR GIVENS ROTATIONS

An elementary 2×2 unitary rotation Θ (also known as Givens or circular rotation) takes a 1×2 row vector $x = \begin{bmatrix} a & b \end{bmatrix}$ and rotates it to lie along the basis vector $e_0 = \begin{bmatrix} 1 & 0 \end{bmatrix}$. More precisely, it performs the transformation

$$\begin{bmatrix} a & b \end{bmatrix} \Theta = \begin{bmatrix} \alpha & 0 \end{bmatrix} , \tag{B.2.1}$$

where, for general complex data, α may be complex as well. Furthermore, its magnitude needs to be consistent with the fact that the pre-array, $\begin{bmatrix} a & b \end{bmatrix}$, and the post-array, $\begin{bmatrix} \alpha & 0 \end{bmatrix}$, must have equal Euclidean norms since

$$\begin{bmatrix} a & b \end{bmatrix} \underbrace{\Theta\Theta^*}_{I} \begin{bmatrix} a^* \\ b^* \end{bmatrix} = \begin{bmatrix} \alpha & 0 \end{bmatrix} \begin{bmatrix} \alpha^* \\ 0 \end{bmatrix} .$$

In other words, α must satisfy $|\alpha|^2 = |a|^2 + |b|^2$, and its magnitude should therefore be $|\alpha| = \sqrt{|a|^2 + |b|^2}$.

An expression for Θ that achieves the transformation (B.2.1) is given by

$$\Theta = \frac{1}{\sqrt{1 + |\rho|^2}} \begin{bmatrix} 1 & -\rho \\ \rho^* & 1 \end{bmatrix} \quad \text{where} \quad \rho = \frac{b}{a} , \quad a \neq 0. \tag{B.2.2}$$

Indeed, if we write a in polar form, say $a = |a|e^{j\phi_a}$, then it can be verified by direct calculation that we can write

$$\begin{bmatrix} a & b \end{bmatrix} \Theta = \begin{bmatrix} \pm e^{j\phi_a}\sqrt{|a|^2 + |b|^2} & 0 \end{bmatrix} .$$

That is, the above Θ leads to a post-array with a complex value α that has the same magnitude as the Euclidean norm of the pre-array, but with a phase factor that is determined by the phase of a up to a sign change. [Note in particular that αa^* is real.]

For real data $\{a, b\}$, and hence $\rho = \rho^*$, the same argument will show that we get a post-array of the form

$$\begin{bmatrix} a & b \end{bmatrix} \Theta = \begin{bmatrix} \pm\sqrt{|a|^2 + |b|^2} & 0 \end{bmatrix} .$$

In any case, the main issue is that once a and b are given, real or complex, a unitary rotation Θ can be defined that reduces the pre-array $\begin{bmatrix} a & b \end{bmatrix}$ to the form $\begin{bmatrix} \alpha & 0 \end{bmatrix}$, for some α.

In the trivial case $a = 0$, we simply choose Θ to be the permutation matrix,

$$\Theta = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

We finally note that, in the special case of real data, a general unitary rotation as in (B.2.2) can be expressed in the alternative form:

$$\Theta = \begin{bmatrix} c & -s \\ s & c \end{bmatrix}$$

where the so-called cosine and sine parameters, c and s, respectively, are defined by

$$c = \frac{1}{\sqrt{1+|\rho|^2}} \quad , \quad s = \frac{\rho}{\sqrt{1+|\rho|^2}} \; .$$

This justifies the name *circular rotation* for Θ, since the effect of Θ is to rotate a vector x along a circle of radius $\|x\|$, by an angle θ that is determined by the inverse of the above cosine and/or sine parameters, $\theta = \tan^{-1}\rho$, in order to align it with the basis vector $\begin{bmatrix} 1 & 0 \end{bmatrix}$. The trivial case $a = 0$ corresponds to a 90 degrees rotation in an appropriate clockwise (if $b \geq 0$) or anti-clockwise (if $b < 0$) direction.

Triangularizing a Matrix

Matrix triangularization can also be effected by a product of Givens transformations, each of which introduces a zero in a particular location. For example, suppose that $x = [\ldots, x_i, \ldots, x_j, \ldots]$, and we wish to null out x_j using x_i. All entries are possibly complex. Then let $\rho = x_j/x_i$, and define

$$\Theta = \begin{bmatrix} 1 & & & & & & & \\ & 1 & & & & & & \\ & & \frac{1}{\sqrt{1+|\rho|^2}} & & & \frac{-\rho}{\sqrt{1+|\rho|^2}} & & \\ & & & 1 & & & & \\ & & & & 1 & & & \\ & & \frac{\rho^*}{\sqrt{1+|\rho|^2}} & & & \frac{1}{\sqrt{1+|\rho|^2}} & & \\ & & & & & & 1 & \\ & & & & & & & 1 \end{bmatrix},$$

Except for the ρ terms, all off-diagonal entries are 0. Then we can see that

$$x\Theta = [\ldots, \alpha, \ldots, 0, \ldots] .$$

where the entries indicated by $\ldots$ are arbitrary and unchanged by Θ, while the resulting α will be of the general form

$$\alpha \;=\; \pm\, e^{j\phi_{x_i}} \sqrt{x_i^2 + x_j^2} \; .$$

To systematically triangularize a $p \times p$ matrix A, apply a sequence of $p-1$ such transformations to zero all entries in the first row of A expect for the first element. Proceed to the second row of the thus transformed A matrix and apply a sequence of $p-2$ transformations to zero all entries after the second one. The first row has a zero in every column affected by this sequence of transformations, so it will be undisturbed. Continuing in this fashion for all rows except the last one, we will transform A to lower triangular form.

In general, the Givens method of triangularization requires more computations than the Householder method. It requires about 30% more multiplications, and it requires one scalar square root per zero produced as opposed to one per column for the Householder method. However, the Givens method is more flexible in preserving zeros already present in the A matrix and can require fewer computations than the Householder method when A is nearly triangular to begin with (see [Gen73]) Moreover, there is a fast version, presented next, that uses 50% fewer multiplications.

B.3 HYPERBOLIC TRANSFORMATIONS

In many cases, it is necessary to use hyperbolic transformations rather than unitary transformations. We therefore exhibit here the necessary modifications.

Elementary Hyperbolic Rotations

An elementary 2×2 hyperbolic rotation Θ takes a row vector $x = \begin{bmatrix} a & b \end{bmatrix}$ and rotates it to lie either along the basis vector $e_0 = \begin{bmatrix} 1 & 0 \end{bmatrix}$ (if $|a| > |b|$) or along the basis vector $e_1 = \begin{bmatrix} 0 & 1 \end{bmatrix}$ (if $|a| < |b|$). More precisely, it performs either of the transformations:

$$\begin{bmatrix} a & b \end{bmatrix} \Theta = \begin{bmatrix} \alpha & 0 \end{bmatrix} \quad \text{if } |a| > |b| \,, \tag{B.3.1}$$

$$\begin{bmatrix} a & b \end{bmatrix} \Theta = \begin{bmatrix} 0 & \alpha \end{bmatrix} \quad \text{if } |a| < |b| \,, \tag{B.3.2}$$

where, for general complex data, α may be complex as well. Furthermore, its magnitude needs to be consistent with the fact that the pre-array, $\begin{bmatrix} a & b \end{bmatrix}$, and the post-array, $\begin{bmatrix} \alpha & 0 \end{bmatrix}$, must have equal Euclidean $J-$norms, *e.g.,* when $|a| > |b|$ we get

$$\begin{bmatrix} a & b \end{bmatrix} \underbrace{\Theta J \Theta^*}_{J} \begin{bmatrix} a^* \\ b^* \end{bmatrix} = \begin{bmatrix} \alpha & 0 \end{bmatrix} J \begin{bmatrix} \alpha^* \\ 0 \end{bmatrix} \,,$$

where $J = (1 \oplus -1)$. By the $J-$norm of a row vector x we mean the indefinite quantity xJx^*, which can be positive, negative, or even zero. Hence, for $|a| > |b|$, α must satisfy $|\alpha|^2 = |a|^2 - |b|^2$, and its magnitude should therefore be $|\alpha| = \sqrt{|a|^2 - |b|^2}$. When $|b| > |a|$ we should get $|\alpha| = \sqrt{|b|^2 - |a|^2}$.

An expression for a $J-$unitary hyperbolic rotation Θ that achieves (B.3.1) or (B.3.2) is given by

$$\Theta = \frac{1}{\sqrt{1-|\rho|^2}} \begin{bmatrix} 1 & -\rho \\ -\rho^* & 1 \end{bmatrix} \quad \text{where} \quad \rho = \frac{b}{a} \;,\;\; a \neq 0 \;,\;\; \text{if } |a| > |b| \qquad \text{(B.3.3)}$$

$$\Theta = \frac{1}{\sqrt{1-|\rho|^2}} \begin{bmatrix} 1 & -\rho \\ -\rho^* & 1 \end{bmatrix} \quad \text{where} \quad \rho^* = \frac{a}{b} \;,\;\; b \neq 0 \;,\;\; \text{if } |a| < |b| \qquad \text{(B.3.4)}$$

It can be verified by direct calculation that these transformations lead to post-arrays of the form

$$\begin{bmatrix} a & b \end{bmatrix} \Theta \;=\; \begin{bmatrix} \pm e^{j\phi_a} \sqrt{|a|^2-|b|^2} & 0 \end{bmatrix} \quad \text{when } |a| > |b| \;,$$

$$\begin{bmatrix} a & b \end{bmatrix} \Theta \;=\; \begin{bmatrix} 0 & \pm e^{j\phi_b} \sqrt{|b|^2-|a|^2} \end{bmatrix} \quad \text{when } |b| > |a| \;,$$

For real data, a general hyperbolic rotation as in (B.3.3) or (B.3.4) can be expressed in the alternative form:

$$\Theta = \begin{bmatrix} ch & -sh \\ -sh & ch \end{bmatrix}$$

where the so-called hyperbolic cosine and sine parameters, ch and sh, respectively, are defined by

$$ch = \frac{1}{\sqrt{1-|\rho|^2}} \;\;,\;\; sh = \frac{\rho}{\sqrt{1-|\rho|^2}} \;.$$

This justifies the name *hyperbolic rotation* for Θ, since the effect of Θ is to rotate a vector x along the *hyperbola* of equation

$$x^2 - y^2 = |a|^2 - |b|^2,$$

by an angle θ that is determined by the inverse of the above hyperbolic cosine and/or sine parameters, $\theta = \tanh^{-1}\rho$, in order to align it with the appropriate basis vector. Note also that the special case $|a| = |b|$ corresponds to a row vector $x = \begin{bmatrix} a & b \end{bmatrix}$ with zero hyperbolic norm since $|a|^2 - |b|^2 = 0$. It is then easy to see that there does not exist a hyperbolic rotation that will rotate x to lie along the direction of one basis vector or the other.

There exist alternative implementations of the hyperbolic rotation Θ that exhibit better numerical properties. Here we briefly mention two modifications.

Mixed Downdating

Assume we apply a hyperbolic rotation Θ to a row vector $\begin{bmatrix} x & y \end{bmatrix}$, say

$$\begin{bmatrix} x_1 & y_1 \end{bmatrix} = \begin{bmatrix} x & y \end{bmatrix} \frac{1}{\sqrt{1-|\rho|^2}} \begin{bmatrix} 1 & -\rho \\ -\rho^* & 1 \end{bmatrix} . \qquad \text{(B.3.5)}$$

Then, more explicitly,

$$x_1 = \frac{1}{\sqrt{1-|\rho|^2}}\left[x - \rho^* y\right] , \tag{B.3.6}$$

$$y_1 = \frac{1}{\sqrt{1-|\rho|^2}}\left[-\rho x + y\right] . \tag{B.3.7}$$

Solving for x in terms of x_1 from the first equation and substituting into the second equation we obtain

$$y_1 = -\rho x_1 + \sqrt{1-|\rho|^2}\, y . \tag{B.3.8}$$

An implementation that is based on (B.3.6) and (B.3.8) is said to be in mixed downdating form. It has better numerical stability properties than a direct implementation of Θ as in (B.3.5) – see [BBDH87]. In the above mixed form, we first evaluate x_1 and then use it to compute y_1. We can obtain a similar procedure that first evaluates y_1 and then uses it to compute x_1. For this purpose, we solve for y in terms of y_1 from (B.3.7) and substitute into (B.3.6) to obtain

$$x_1 = -\rho^* y_1 + \sqrt{1-|\rho|^2}\, x . \tag{B.3.9}$$

Eqs. (B.3.7) and (B.3.9) represent the second mixed form.

The OD Method

The OD (Orthogonal-Diagonal) procedure is based on using the SVD of the hyperbolic rotation Θ. Assume ρ is real and write $\rho = b/a$, where $|a| > |b|$. Then it is straightforward to verify that any hyperbolic rotation of this form admits the following eigen-decomposition:

$$\Theta = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} \sqrt{\frac{a+b}{a-b}} & 0 \\ 0 & \sqrt{\frac{a-b}{a+b}} \end{bmatrix}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}\frac{1}{\sqrt{2}} \triangleq QDQ^T, \tag{B.3.10}$$

where the matrix

$$Q = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$$

is orthogonal ($QQ^T = I$), and T denotes transposition.

Due to the special form of the factors (Q, D), a real hyperbolic rotation, with $|\rho| < 1$ can then be applied to a row vector $\begin{bmatrix} x & y \end{bmatrix}$ to yield $\begin{bmatrix} x_1 & y_1 \end{bmatrix}$ as follows (note that the first and last steps involve simple additions and subtractions):

$$\begin{bmatrix} x' & y' \end{bmatrix} \leftarrow \begin{bmatrix} x & y \end{bmatrix}\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$$

$$\begin{bmatrix} x'' & y'' \end{bmatrix} \leftarrow \begin{bmatrix} x' & y' \end{bmatrix} \begin{bmatrix} \frac{1}{2}\sqrt{\frac{a+b}{a-b}} & 0 \\ 0 & \frac{1}{2}\sqrt{\frac{a-b}{a+b}} \end{bmatrix}$$

$$\begin{bmatrix} x_1 & y_1 \end{bmatrix} \leftarrow \begin{bmatrix} x'' & y'' \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$$

This procedure is numerically stable, as shown in [CS96] (and also Ch. 2). An alternative so-called H-procedure is also described in the same reference. It is costlier than the OD method, but is more accurate and can be shown to be "forward" stable, which is a desirable property for finite precision implementations.

When ρ is a complex number, the unitary matrix Q becomes complex. If we now write $\rho = b/a$, then $|\rho| = |b|/|a|$ and $|a| > |b|$, and we obtain

$$Q = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -\frac{|a|}{a}\frac{b}{|b|} & \frac{|a|}{a}\frac{b}{|b|} \end{bmatrix}, \quad D = \begin{bmatrix} \frac{\sqrt{|a|+|b|}}{\sqrt{|a|-|b|}} & \\ & \frac{\sqrt{|a|-|b|}}{\sqrt{|a|+|b|}} \end{bmatrix}.$$

Hyperbolic Householder Transformations

One can also use hyperbolic or $J-$unitary Householder reflections to simultaneously annihilate several entries in a row, *e.g.*, to transform $\begin{bmatrix} x & x & x & x \end{bmatrix}$ directly to the form $\begin{bmatrix} x' & 0 & 0 & 0 \end{bmatrix}$.

Let J be an $n \times n$ signature matrix such as $J = (I_p, -I_q)$ with $p + q = n$. We are now interested in a $J-$unitary Householder transformation Θ that takes a $1 \times n$ row vector x and aligns it either along the basis vector $e_0 = \begin{bmatrix} 1 & 0 \end{bmatrix}$ (if $xJx^* > 0$) or along the basis vector $e_{n-1} = \begin{bmatrix} 0 & 1 \end{bmatrix}$ (if $xJx^* < 0$). [One can also require Θ to align x along the direction of some other basis vector depending on the sign of xJx^* and on the order of the sequence of $\pm 1's$ in J. We shall, without loss of generality, focus here on the special directions e_0 and e_{n-1} and assume that J is of the form $J = (I_p \oplus -I_q)$.]

Hence, we require Θ to perform either of the transformations

$$x\Theta = \pm\alpha\, e_0 \quad \text{if} \quad xJx^* > 0 \tag{B.3.11}$$

$$x\Theta = \pm\alpha\, e_{n-1} \quad \text{if} \quad xJx^* < 0\,, \tag{B.3.12}$$

where, for general complex data, the resulting α may be complex as well.

When $xJx^* > 0$, we define

$$\Theta = I - 2\,\frac{Jg^*g}{gJg^*} \quad \text{where} \quad g = x + \alpha\, e_0\,, \tag{B.3.13}$$

and α is a complex number that satisfies $|\alpha|^2 = xJx^*$ and αx_1^* is real. It can be verified by direct calculation that Θ is $J-$unitary, *i.e.*, $\Theta J \Theta^* = J = \Theta^* J \Theta$. When $xJx^* < 0$, we use the same expression for Θ but with

$$g = x + \alpha\, e_{n-1}\,, \tag{B.3.14}$$

and α a complex number that satisfies $|\alpha|^2 = -xJx^*$ and αx^*_{n-1} is real.

Lemma B.3.1 (Complex Hyperbolic Householder Transformation) *Given a row vector x with leading entry x_1 and $xJx^* > 0$, define Θ and g as in (B.3.13) where α is any complex number that satisfies the following two requirements (see below): $|\alpha| = \sqrt{xJx^*}$ and αx_1^* is real. Then it holds that $x\Theta = -\alpha e_0$. That is, x is rotated and aligned with e_0; the leading entry of the post-array is equal to $-\alpha$.*

For a vector x that satisfies instead $xJx^* < 0$, and with trailing entry x_{n-1}, we choose g as in (B.3.14) where α is any complex number that satisfies: $|\alpha| = \sqrt{|xJx^*|}$ and αx^*_{n-1} is real. Then it holds that $x\Theta = -\alpha e_{n-1}$.

(**Algebraic**) **proof**: We prove the first statement only since the second one follows from a similar argument. Direct calculation shows that

$$\begin{aligned} gJg^* &= 2xJx^* + 2\alpha x_1^* , \\ xJg^*g &= x(xJx^*) + \alpha(xJx^*)e_0 + \alpha^* x_1 x + \alpha(\alpha x_1^*)e_0 , \\ xgJg^* &= 2x(xJx^*) + 2\alpha x_1^* x . \end{aligned}$$

Therefore,

$$x\Theta = \frac{xgJg^* - 2xJg^*g}{gJg^*} = -\alpha e_0 .$$

$\diamondsuit$

Specific choices for α, and hence g, are:

$$g = \begin{cases} x \pm e^{j\phi_{x_1}} \sqrt{xJx^*}\, e_0 & \text{when } xJx^* > 0 \\ x \pm e^{j\phi_{x_{n-1}}} \sqrt{|xJx^*|}\, e_{n-1} & \text{when } xJx^* < 0 \end{cases}$$

and they lead to

$$x\Theta = \begin{cases} \mp e^{j\phi_{x_1}} \sqrt{xJx^*}\, e_0 & \text{when } xJx^* > 0 \\ \mp e^{j\phi_{x_{n-1}}} \sqrt{|xJx^*|}\, e_{n-1} & \text{when } xJx^* < 0 \end{cases}$$

Geometric Derivation

The geometric derivation presented earlier for Householder transformations still applies provided we use "J-inner products", *i.e.*, provided we interpret

$$\begin{aligned} \langle x, g\rangle_J &= xJg^* , \quad \text{when } \{x, g\} \text{ are rows} , \\ \langle x, g\rangle_J &= g^*Jx , \quad \text{when } \{x, g\} \text{ are columns} . \end{aligned}$$

Then using rows, we can write, for example, when $\|x\|_J = \sqrt{xJx^*} > 0$,

$$\mp\alpha e_0 = x - 2\langle x, g\rangle_J \, \|g\|_J^{-2}\, g = x\left(I - 2\frac{Jg^*g}{gJg^*}\right) ,$$

where $g \triangleq x \pm \alpha\, e_0$. Tab. B.1 collects the expressions for the several rotations that we have considered in the earlier discussion.

Table B.1. Unitary and hyperbolic rotations.

Rotation	Expression	Effect
Circular or Givens	$\Theta = \frac{1}{\sqrt{1+\|\rho\|^2}} \begin{bmatrix} 1 & -\rho \\ \rho^* & 1 \end{bmatrix}$ $\rho = \frac{b}{a}\ ,\ a \neq 0$	$\begin{bmatrix} a & b \end{bmatrix} \Theta =$ $\begin{bmatrix} \pm e^{j\phi_a}\sqrt{\|a\|^2+\|b\|^2} & 0 \end{bmatrix}$
Permutation	$\Theta = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$	$\begin{bmatrix} 0 & b \end{bmatrix} \Theta = \begin{bmatrix} b & 0 \end{bmatrix}$
Hyperbolic I	$\Theta = \frac{1}{\sqrt{1-\|\rho\|^2}} \begin{bmatrix} 1 & -\rho \\ -\rho^* & 1 \end{bmatrix}$ $\rho = \frac{b}{a}\ ,\ a \neq 0\ ,\ \|a\| > \|b\|$	$\begin{bmatrix} a & b \end{bmatrix} \Theta =$ $\begin{bmatrix} \pm e^{j\phi_a}\sqrt{\|a\|^2-\|b\|^2} & 0 \end{bmatrix}$
Hyperbolic II	$\Theta = \frac{1}{\sqrt{1-\|\rho\|^2}} \begin{bmatrix} 1 & -\rho \\ -\rho^* & 1 \end{bmatrix}$ $\rho^* = \frac{a}{b}\ ,\ b \neq 0\ ,\ \|a\| < \|b\|$	$\begin{bmatrix} a & b \end{bmatrix} \Theta =$ $\begin{bmatrix} 0 & \pm e^{j\phi_b}\sqrt{\|b\|^2-\|a\|^2} \end{bmatrix}$
Unitary Householder	$\Theta = I_n - 2\,\frac{g^* g}{g g^*}$ $g = x \pm e^{j\phi_{x_1}} \\|x\\|\ e_0$	$\begin{bmatrix} x_1 & \dots & x_{n-1} \end{bmatrix} \Theta =$ $\mp e^{j\phi_{x_1}} \\|x\\| e_0$
Hyperbolic Householder I	$\Theta = I - 2\,\frac{J g^* g}{g J g^*}$ $g = x \pm e^{j\phi_{x_1}} \sqrt{\|xJx^*\|}\ e_0$ $xJx^* > 0$	$\begin{bmatrix} x_1 & \dots & x_{n-1} \end{bmatrix} \Theta =$ $\mp e^{j\phi_{x_1}} \sqrt{\|xJx^*\|} e_0$ $e_0 = \begin{bmatrix} 1 & 0 & \dots & 0 \end{bmatrix}$
Hyperbolic Householder II	$\Theta = I - 2\,\frac{J g^* g}{g J g^*}$ $g = x \pm e^{j\phi_{x_{n-1}}} \sqrt{\|xJx^*\|}\ e_{n-1}$ $xJx^* < 0$	$\begin{bmatrix} x_1 & \dots & x_{n-1} \end{bmatrix} \Theta =$ $\mp e^{j\phi_{x_{n-1}}} \sqrt{\|xJx^*\|} e_{n-1}$ $e_{n-1} = \begin{bmatrix} 0 & \dots & 0 & 1 \end{bmatrix}$