Unifying Unitary and Hyperbolic Transformations

Adam Bojanczyk
School of Electrical Engineering
Cornell University
Ithaca, NY 14853-3801

Sanzheng Qiao*
Department of Computing and Software
McMaster University
Hamilton, Ontario L8S 4L7 Canada

Allan O. Steinhardt
DARPA Tactical Technology Office
3701 North Fairfax Drive
Arlington VA 22203-1714

March 2, 2000

Dedicated to Robert J. Plemmons on the occasion of his 60th birthday

Abstract

In this paper we describe unified formulas for unitary and hyperbolic reflections and rotations, and show how these unified transformations can be used to compute a Hermitian triangular decomposition $\hat{R}^{\mathrm{H}} D \hat{R}$ of a strongly nonsingular indefinite matrix $\hat{A}$ given in the form $\hat{A} = X_1^{\mathrm{H}} X_1 + \alpha X_2^{\mathrm{H}} X_2$, $\alpha = \pm 1$.

The unification is achieved by the introduction of signature matrices which determine whether the applicable transformations are unitary, hyperbolic, or their generalizations. We derive formulas for the condition numbers of the unified transformations, propose pivoting strategies for lowering the condition number of the transformations, and present a unified stability analysis for applying the transformations to a matrix.

Key Words. Cholesky factor modification, hyperbolic rotation, hyperbolic Householder transformation, error analysis.

*This work was partly supported by the Natural Sciences and Engineering Research Council of Canada under grant OGP0046301

1 Introduction

Consider an n-by-n Hermitian and positive definite matrix $\hat{A}$

$$\hat{A} = R^{\mathrm{H}} R + \alpha X^{\mathrm{H}} X \tag{1}$$

where R is upper triangular, X is a p-by-n matrix, and $\alpha = \pm 1$. The Cholesky modification problem [8] is to compute the upper triangular Cholesky factor $\hat{R}$ of $\hat{A}$: $\hat{A} = \hat{R}^H \hat{R}$, directly from R and X, without forming $\hat{A}$ explicitly. This modification of a triangular factorization is known as updating and downdating of a Cholesky factor when $\alpha = 1$ and $\alpha = -1$, respectively.

General structure of updating and downdating algorithms is virtually the same. To see this, consider matrices Y ad S defined as follows,

$$\begin{aligned} Y &= \begin{pmatrix} R \\ X \end{pmatrix} \\ S &= \begin{pmatrix} I_n & 0 \\ 0 & \alpha I_p \end{pmatrix} . \end{aligned}$$

With these definitions (1) can be rewritten in the following form:

$$\hat{A} = R^{\mathrm{H}} R + \alpha X^{\mathrm{H}} X = Y^{\mathrm{H}} S Y \ . \tag{2}$$

The building blocks in Cholesky modification problem are transformations, denoted by P_i, which operate on Y, and satisfy the relationship

$$P_i^{\mathrm{H}} S P_i = S \ .$$

Each P_i is selected so as to annihilate subdiagonal elements in the ith column of $P_{i-1} \cdots P_1 Y$. When $\alpha = 1$, P_i are unitary, when $\alpha = -1$, P_i are hyperbolic. The matrix $PY = P_n \cdots P_1 Y$ has all subdiagonal elements equal to zero and hence its upper triangular part $\hat{R}$ is the Cholesky factor of $\hat{A}$. Indeed, if we denote $\hat{Y} = PY$, then

$$\hat{Y} = \begin{pmatrix} \hat{R} \\ 0 \end{pmatrix}$$

and

$$\hat{R}^{\mathrm{H}} \hat{R} = \hat{Y}^{\mathrm{H}} S \hat{Y} = Y^{\mathrm{H}} S Y = \hat{A} \ . \tag{3}$$

The matrix S is a special case of signature matrices which have the form $\mathrm{diag}(\pm 1)$. The introduction of signature matrices leads to the following generalization of the Cholesky modification problem: Given a data matrix Y and a signature $S = \mathrm{diag}(\pm 1)$, find an upper triangular matrix $\hat{R}$ and a signature $\hat{S}$ such that

$$\hat{R}^{\mathrm{H}} \hat{S} \hat{R} = Y^{\mathrm{H}} S Y \ . \tag{4}$$

When $Y = (R^{\mathrm{H}} \ X^{\mathrm{H}})^{\mathrm{H}}$ with R upper triangular, this problem becomes the problem of modifying the triangular factor R after addition and/or deletion of data X. The signature S allows $\hat{A} = Y^{\mathrm{H}} S Y$ to be an indefinite matrix.

In this paper we introduce unified transformations which can be used to solve the general problem (4) by virtually the same algorithm irrespective of updating, downdating or modification of a symmetric triangular decomposition of an indefinite strongly nonsingular matrix $\hat{A}$. The unified transformations include unitary and hyperbolic transformations as special cases. We also present a stability analysis of the unified transformations. Based on the analysis, we propose pivoting techniques for improving numerical properties of the transformations.

This paper is organized as follows. After introducing notations in Section 2, we define the unified Householder transformations (hypernormal reflections) and the unified rotations (hypernormal rotations) in Section 3. In Section 4, we derive formulas for the condition numbers of matrices representing these transformations, and present a stability analysis of applying the transformations to a vector. We also show how pivoting can be combined with hypernormal reflections (or hypernormal rotations) to obtain decompositions (1) and (4). A numerical example is included to illustrate positive effects of pivoting.

2 Notations

A signature matrix S defines a class of S-unitary matrices.

Definition 1 *Let $S = \mathrm{diag}(\pm 1)$. A square matrix V is called S-unitary iff*

$$V^{\mathrm{H}} S V = S \ . \tag{5}$$

If $S = I$ then an S-unitary matrix is a unitary matrix. If S has both positive and negative ones on the diagonal, then an S-unitary matrix V is a hyperbolic matrix.

From (3) we see that S-unitary transformations are sufficient tools for finding a modified Cholesky factor of a positive definite matrix $\hat{A}$ in (1), irrespective of $\alpha = 1$ or $\alpha = -1$.

In considering S-unitary matrices it is helpful to introduce a bilinear form $[\cdot,\cdot]_S$ on C^n induced by a signature matrix S,

$$[x, y]_S = x^{\mathrm{H}} S y \ \ , \ \ x, y \in C^n \ . \tag{6}$$

This bilinear form is called indefinite scalar product [7]. The bilinear form $[.,.]_S$ defines a weighted norm $||\cdot||_S$,

$$||x||_S = \mathrm{sign}([x,x]_S) \sqrt{|[x,x]_S|} \ .$$

Note that $||v||_S$ is not a norm despite the notation, because a norm should be non-negative. From this definition, it is easy to see that if V is an $n \times n$ S-unitary matrix, then for any vector $x \in C^n$, $||Vx||_S = ||x||_S$.

A further generalization of the definition of S-unitary matrices is useful. In the situation when the difference $\hat{A} = R^{\mathrm{H}}R - X^{\mathrm{H}}X$ is an indefinite matrix, the Cholesky factor does not exist. However, if $\hat{A}$ is strongly nonsingular it is possible to find a unique triangular decomposition of the form

$$\hat{A} = \hat{R}^{\mathrm{H}}\hat{S}\hat{R} \tag{7}$$

where $\hat{S} = \mathrm{diag}(\pm 1)$ is a signature matrix and $\hat{R}$ is upper triangular.

Note that if $\hat{S} = I$ then we have a standard Cholesky modification problem. It turns out that algorithms for computing $\hat{R}$ and $\hat{S}$ in the case of indefinite $\hat{A}$ become completely analogous to those for modifying Cholesky decomposition, if the definition of S-unitary matrices is generalized as follows, see [1].

Definition 2 *Let* $S = \mathrm{diag}(\pm 1)$ *and* $\hat{S} = \mathrm{diag}(\pm 1)$ *be* $n \times n$ *diagonal matrices. A matrix* V *satisfying*

$$V^{\mathrm{H}}SV = \hat{S}$$

is called a hypernormal matrix with respect to S *and* $\hat{S}$*.*

The signature matrices S and $\hat{S}$ are allowed to be different here. However, by the Sylvester theorem, $\hat{S}$ must be a symmetric permutation of S. If $S = \hat{S}$ then hypernormal matrices are S-unitary matrices. Furthermore, if $S = I$ then they are unitary matrices.

3 Unified Transformations

In this section, we define the unified Householder transformations and unified rotations and present their algorithms.

3.1 Hypernormal Reflections

First we introduce hypernormal reflections which are extensions of Householder reflections [8] and include as special cases both unitary and hyperbolic reflections [11].

We start by recalling that an S-unitary reflections H has the form

$$H = H(S, b) = S - \frac{2bb^{\mathrm{H}}}{b^{\mathrm{H}}Sb} \tag{8}$$

for some vector b, $b^{\mathrm{H}}Sb \neq 0$. As mentioned earlier, if $S = I$ then H represents a unitary reflection.

The utility of an S-unitary reflection is to map a vector of interest v onto a vector $\hat{v}$ parallel to e_1,

$$\hat{v} = Hv = c_v \cdot e_1 \tag{9}$$

where the scalar c_v depends on v (and S). As S-unitary reflections preserve the S-norm, the relation (9) shows that the following condition must be satisfied

$$||v||_S = \mathrm{sign}(S(1,1))|c_v| \; . \tag{10}$$

As the sign on the right hand side of (10) is determined by $S(1,1)$, the conditions (9) and (10) cannot be simultaneously satisfied for an arbitrary vector v and arbitrary signature S.

The following theorem shows how S-unitary reflections (8) can be generalized so that (9) and (10) are satisfied for all v such that $v^{\mathrm{H}} S v \neq 0$.

Theorem 1 *Let v and S be a vector and a signature matrix, respectively, such that $v^{\mathrm{H}} S v \neq 0$. Let J be a permutation for which*

$$\mathrm{sign}(e_1^{\mathrm{T}} JSJe_1) = \mathrm{sign}(v^{\mathrm{H}} S v) \; , \tag{11}$$

and set

$$\tilde{v} = Jv \quad \textit{and} \quad \tilde{S} = JSJ \; . \tag{12}$$

Let us define a vector b by

$$b = \tilde{S}\tilde{v} + \theta \, \mathrm{abs}(||v||_S) e_1 \tag{13}$$

$$\theta = \begin{cases} \mathrm{sign}(e_1^{\mathrm{T}} \tilde{S} e_1) \frac{\tilde{v}_1}{|\tilde{v}_1|} & \textit{if } \tilde{v}_1 \neq 0 \\ \mathrm{sign}(e_1^{\mathrm{T}} \tilde{S} e_1) & \textit{otherwise.} \end{cases} \tag{14}$$

Then the transformation

$$H_J = JS - 2b(Jb)^{\mathrm{H}} / (b^{\mathrm{H}} \tilde{S} b) \tag{15}$$

satisfies

$$\hat{v} = H_J v = -\theta \, \mathrm{abs}(||v||_S) e_1 \tag{16}$$

$$||\hat{v}||_{\tilde{S}} = ||v||_S \tag{17}$$

$$(H_J)^{\mathrm{H}} \tilde{S} (H_J) = S \; . \tag{18}$$

Proof:

In order to show (16) note that, using (12) and (13),

$$b^{\mathrm{H}} \tilde{S} b = \mathrm{sign}(v^{\mathrm{H}} S v) ||v||_S^2 + \theta \, \tilde{v}_1^* \, \mathrm{abs}(||v||_S) + \theta^* \tilde{v}_1 \, \mathrm{abs}(||v||_S) + |\theta|^2 \, ||v||_S^2 \, \mathrm{sign}(e_1^{\mathrm{T}} \tilde{S} e_1). \tag{19}$$

With the choice of (14), (19) becomes

$$b^{\mathrm{H}} \tilde{S} b = 2 \, \mathrm{sign}(e_1^{\mathrm{T}} \tilde{S} e_1) (||v||_S^2 + \mathrm{abs}(||v||_S) |\tilde{v}_1|) \; .$$

We also have

$$b^{\mathrm{H}}\tilde{v} = \mathrm{sign}(e_1^{\mathrm{T}}\tilde{S}e_1)(||v||_S^2 + \mathrm{abs}(||v||_S)|\tilde{v}_1|)\ .$$

Now it is easy to check that (16) holds.

The relationship (17) follows immediately from (11), (12) and the fact that $H_J J$ is $\tilde{S}$-unitary.

The relation (18) states that the transformation H_J is hypernormal with respect to S and $\tilde{S}$. This leads to to the following definition.

Definition 3 (Unified Householder Transformation) *Let S be a signature matrix and v a vector with nonzero S-norm. Let $\tilde{S} = JSJ$, and b be as in Theorem 1. Then the transformation H_J defined by (15) is called a unified Householder transformation for the vector v and the signatures S and $\tilde{S}$.*

The relation (16) states that any vector v with nonzero S-norm can be mapped by a unified reflection onto the first coordinate e_1. The vector e_1 can be replaced by any direction d for which $d^{\mathrm{H}}Sd \neq 0$.

Note that if H_J is hypernormal with respect to S and $\hat{S}$ then $P = H_J J$ is $\hat{S}$-unitary. Instead of dealing with H_J, it is convenient to consider P. The following algorithm shows how P (and hence H_J) can be computed.

Algorithm 1 (Unified Householder Transformation) *Given a signature matrix $S = \mathrm{diag}(\sigma_i)$, $\sigma_i = \pm 1$, and a vector x with nonzero S-norm, this algorithm computes the vector u, scalars α and δ, a signature matrix $\hat{S}$ and a permutation matrix J so the unified Householder transformation*

$$P = \hat{S} - \delta^{-1}uu^{\mathrm{H}} \tag{20}$$

has the following properties,

$$(PJ)x = -\alpha e_1 \quad \textit{and} \quad (PJ)^{\mathrm{H}}\hat{S}(PJ) = S,$$

where $\hat{S} = JSJ$.

> $m = \max\{|x_i|\}$; $x = x/m$;
> $J = I$;
> **if** $\mathrm{sign}(x^{\mathrm{H}}Sx) \neq \sigma_1$
>   find j so that $\sigma_j = \mathrm{sign}(x^{\mathrm{H}}Sx)$;
> $J =$ permutation $(1, j)$;
> **end**
> $y = Jx$; $\hat{S} = JSJ$;
> **if** $y_1 \neq 0$
> $\theta = \mathrm{sign}(\hat{\sigma}_1)y_1/|y_1|$;

else

$\theta = \mathrm{sign}(\hat{\sigma}_1)$;

end

$\alpha = \theta\sqrt{\left|y^{\mathrm{H}}\hat{S}y\right|}$;

$u = \hat{S}y + \alpha e_1$;

$\delta = \hat{\sigma}_1 \alpha^* u_1$;

$\alpha = m\alpha$.

Unified Householder transformations can be used to find the decomposition (7). Namely, we can construct a sequence of unified Householder transformations H_{J_1}, H_{J_2}, $\cdots$, H_{J_k} such that

$$H_{J_k} \cdots H_{J_2} H_{J_1} Y = \begin{bmatrix} \hat{R} \\ 0 \end{bmatrix} \tag{21}$$

where H_{J_i} are hypernormal with respect to S_i and S_{i-1}, $S_i = J_i S_{i-1} J_i$, and J_i are permutations determined by the relation (11). Recalling (2), from (21) we obtain S_k and $\hat{R}$ such that

$$\hat{A}^{\mathrm{H}} S \hat{A} = \hat{R}^{\mathrm{H}} S_k \hat{R}.$$

At this point a problem that should be addressed is: what happens when $||v||_S = 0$? The answer is that both procedures per se fail (see [5] for some implications of this problem). What we rely upon in recovering from a situation of $||v||_S = 0$ is that the unified Householder transformation is applied to whole matrices, not merely to isolated column vectors. If in step i the working column v_i and the signature matrix S_i are such that

$$v_i^{\mathrm{H}} S_i v_i = 0$$

then a suitable permutation of the remaining columns of $H_{J_{i-1}} \cdots H_{J_1} Y$ has to be chosen so for the new working column v_i

$$v_i^{\mathrm{H}} S v_i \neq 0 \ .$$

This is possible when $\hat{A}^{\mathrm{H}} S \hat{A}$ is strongly nonsingular [8]. Column and row pivoting can be introduced so as to decrease a potential magnification of rounding errors, see Section 4. On completion we get the triangular factorization of the form

$$K^{\mathrm{T}} \hat{A} K = \hat{R}^{\mathrm{H}} S_k \hat{R} \ , \tag{22}$$

where K represents permutation of columns of $\hat{A}$.

3.2 Hypernormal Rotation

In analogy to the unified Householder reflection we propose the following unified definition of a rotation. This definition includes both unitary and hyperbolic rotations as special cases.

Definition 4 (Unified Rotation) *Given a signature matrix $S = \mathrm{diag}(\sigma_1, \sigma_2)$, a unified rotation has the form:*

$$Q = \begin{pmatrix} c^* & (\sigma_2/\sigma_1)s^* \\ -s & c \end{pmatrix} \tag{23}$$

where the pair (c, s) satisfies

$$\sigma_1|c|^2 + \sigma_2|s|^2 = \hat{\sigma}_1, \quad \hat{\sigma}_1 = \pm 1.$$

Note that

$$Q^{\mathrm{H}}SQ = \begin{pmatrix} \sigma_1|c|^2 + \sigma_2|s|^2 & 0 \\ 0 & \sigma_2|c|^2 + \sigma_1|s|^2 \end{pmatrix} = \begin{pmatrix} \hat{\sigma}_1 & 0 \\ 0 & \sigma_1\sigma_2\hat{\sigma}_1 \end{pmatrix}.$$

Denoting $\hat{\sigma}_2 = \sigma_1\sigma_2\hat{\sigma}_1$, we have

$$Q^{\mathrm{H}}SQ = \hat{S},$$

where $\hat{S} = \mathrm{diag}(\hat{\sigma}_1, \hat{\sigma}_2)$, that is Q is a hypernormal matrix with respect to S and $\hat{S}$.

It can be verified that when $\sigma_1\sigma_2 = 1$, Q is a unitary rotation, when $\sigma_1 = -\sigma_2 = 1$ and $\sigma_1|x_1|^2 + \sigma_2|x_2|^2 > 0$, Q is a hyperbolic rotation, and when $\sigma_1 = \sigma_2 = 1$ and $\sigma_1|x_1|^2 + \sigma_2|x_2|^2 < 0$, Q is a purely hypernormal transformation.

Similar to the unified Householder reflections, unified rotations can be used to zero selected elements of a vector. The following algorithm summarizes how this can be done.

Algorithm 2 (Unified Rotation) *For a given vector $x = (x_1, x_2)^{\mathrm{T}}$ and a signature matrix $S = \mathrm{diag}(\sigma_1, \sigma_2)$ this algorithm computes a unified rotation*

$$Q = \begin{pmatrix} c^* & (\sigma_2/\sigma_1)s^* \\ -s & c \end{pmatrix}$$

and a signature matrix $\hat{S} = \mathrm{diag}(\hat{\sigma}_1, \hat{\sigma}_2)$ so that

$$Q\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \rho \\ 0 \end{pmatrix} \quad \textit{and} \quad Q^{\mathrm{H}}SQ = \mathrm{diag}(\hat{\sigma}_1, \hat{\sigma}_2).$$

if $x_2 = 0$
    $c = 1$; $s = 0$; $\hat{\sigma}_1 = \sigma_1$; $\hat{\sigma}_2 = \sigma_2$; return
end
if $\sigma_1|x_1| = -\sigma_2|x_2|$
    error("Indefinite norm of x is zero")
end
if $|x_1| > |x_2|$
    $\tau = x_2/x_1$; $\delta = 1 + (\sigma_2/\sigma_1)|\tau|^2$;
    $c = \mathrm{sign}(x_1)/\sqrt{|\delta|}$; $s = \tau c$;
else
    $\tau = x_1/x_2$; $\delta = \sigma_2/\sigma_1 + |\tau|^2$;
    $s = \mathrm{sign}(x_2)/\sqrt{|\delta|}$; $c = \tau s$;
end
if $\delta > 0$
    $\hat{\sigma}_1 = \sigma_1$; $\hat{\sigma}_2 = \sigma_2$;
else
    $\hat{\sigma}_1 = -\sigma_1$; $\hat{\sigma}_2 = -\sigma_2$;
end.

4 Analysis

In this section we derive formulas for condition numbers of unified transformations and show that their condition numbers can be arbitrarily large. Then, we propose pivoting strategies for avoiding large condition numbers when a sequence of transformations is applied to a matrix. Finally, we present an error analysis of the direct application of a unified transformation to a vector.

4.1 Unified Householder Transformations

Consider P defined in (20) (or PJ). Suppose that the diagonal of the signature matrix S has p positive elements and m negative elements. Partitioning $S = M + N$ where M is the positive part of S and N the negative part, we have

$$P = M + N - \xi u u^{\mathrm{H}} \quad \text{where} \quad \xi = \frac{2}{u^{\mathrm{H}} S u}$$

and

$$\mathrm{rank}(M) = p, \quad \mathrm{rank}(N) = m, \quad \text{where } p + m = n.$$

Thus $\mathrm{rank}(N - \xi u u^{\mathrm{H}}) \le m + 1$ and P has at least $n - (m+1) = p - 1$ eigenvalues equal to $+1$. Similarly, P has at least $m - 1$ eigenvalues equal to -1 since $\mathrm{rank}(M - \xi u u^{\mathrm{H}}) \le p + 1$. It follows that P has at least $(m-1) + (p-1) = n - 2$ unit singular values.

Now, we derive formulas for the other two singular values of P. We can verify that

$$(P + SPS)(u + Su) = -\xi(u^{\mathrm{H}}u)(u + Su).$$

Thus $-\xi(u^{\mathrm{H}}u)$ is an eigenvalue of $P + SPS$. Since P is S-unitary and Hermitian, we have $SPS = P^{-1}$ and hence $-\xi(u^{\mathrm{H}}u)$ is an eigenvalue of $P + P^{-1}$. Since eigenvalues of P and P^{-1} are reciprocals of each other, a non-unit eigenvalue λ of P satisfies

$$\lambda + \lambda^{-1} = -\xi u^{\mathrm{H}} u \ .$$

This in turn means that a non-unit singular value σ of P satisfies

$$|\xi u^{\mathrm{H}} u| = \sigma + \sigma^{-1} \ .$$

Solving for σ in the above equation, we obtain

$$\sigma_{\min}^{-1} = \sigma_{\max} = \frac{u^{\mathrm{H}}u}{|u^{\mathrm{H}}Su|} + \sqrt{\left(\frac{u^{\mathrm{H}}u}{u^{\mathrm{H}}Su}\right)^2 - 1}.$$

When $S = \pm I$, then $\sigma_{\min} = \sigma_{\max} = 1$, $J = I$, and PJ is a unitary Householder matrix. In the case when $S \neq \pm I$ and $J = I$, P is the hyperbolic Householder defined in [9, 10], the above result on the condition number coincides with [3].

For a general signature matrix S, the ratio $(u^{\mathrm{H}}u)/|u^{\mathrm{H}}Su|$ can be arbitrarily large. Thus $\mathrm{cond}(PJ)$ can also be arbitrarily large. However, as shown in Algorithm 1, when a Householder transformation is constructed to introduce zeros into a given vector x, we have some freedom of choosing the permutation J. We propose the following pivoting scheme, which is analogous to the one described in [14], in which J is chosen to minimize the condition number of the Householder transformation.

Algorithm 3 (Householder with row pivoting) *Given a complex vector x and a signature matrix $S = \mathrm{diag}(\sigma_i)$, this algorithm computes the vector u, scalars α and δ, and permutation $(1, j)$ in the unified Householder transformation PJ where*

$$P = \hat{S} - \delta^{-1}uu^{\mathrm{H}}.$$

Let matrix J represent the permutation $(1, j)$, then $\hat{S} = JSJ$,

$$(PJ)x = -\alpha e_1 \quad \textit{and} \quad (PJ)^{\mathrm{H}}\hat{S}(PJ) = S.$$

This algorithm incorporates pivoting to minimize the condition number of P.

$m = \max\{|x_i|\}$; $x = x/m$;
$J = I$;
if $S \neq \pm I$
  find j so that $\sigma_j = \mathrm{sign}(x^{\mathrm{H}}Sx)$ and $|x_j|$ is maximized;

$J =$ permutation $(1, j)$;
end
$y = Jx$; $\hat{S} = JSJ$;
$\theta = \text{sign}(\hat{\sigma}_1) y_1 / |y_1|$;
$\alpha = \theta \sqrt{\left| y^{\mathrm{H}} \hat{S} y \right|}$;
$u = \hat{S} y + \alpha e_1$;
$\delta = \hat{\sigma}_1 \alpha^* u_1$;
$\alpha = m\alpha$.

Numerical Example Consider the following modification of triangular decomposition problem:

$$\hat{A} = R^{\mathrm{T}} R - x x^{\mathrm{T}} + y y^{\mathrm{T}} - z z^{\mathrm{T}} = \hat{R}^{\mathrm{T}} \text{diag}(1, -1) \hat{R}$$

where

$$R = \begin{pmatrix} d & 1 \\ 0 & 1 \end{pmatrix}, \quad x^{\mathrm{T}} = (d + \epsilon \quad d - \epsilon), \quad y^{\mathrm{T}} = (1 \quad -9), \quad z^{\mathrm{T}} = (1 - d \quad 1),$$

and $\hat{R}$ is upper triangular and ϵ is the machine precision. Note that the matrix $\hat{A}$ is indefinite. In matrix form, we want to find two unified Householder transformations $P_1 J_1$ and $P_2 J_2$ such that

$$(P_2 J_2)(P_1 J_1) \begin{pmatrix} R \\ x^{\mathrm{T}} \\ y^{\mathrm{T}} \\ z^{\mathrm{T}} \end{pmatrix} = P_2 J_2 \begin{pmatrix} \hat{r}_{11} & \hat{r}_{12} \\ 0 & 1 \\ 0 & x_2 \\ 0 & y_2 \\ 0 & z_2 \end{pmatrix} = \begin{pmatrix} \hat{r}_{11} & \hat{r}_{12} \\ 0 & \hat{r}_{22} \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \equiv \begin{pmatrix} \hat{R} \\ 0 \end{pmatrix}.$$

For simplicity, we denote $\hat{R}$ as the computed result. We use

$$E = \| \hat{R}^{\mathrm{T}} \text{diag}(1, -1) \hat{R} - (R^{\mathrm{T}} R - x x^{\mathrm{T}} + y y^{\mathrm{T}} - z z^{\mathrm{T}}) \|_2$$

to measure the error.

When $d = 1.0$, no pivoting is necessary and both Algorithm 3 and Algorithm 1 give identical results. Table 1 shows the errors and condition numbers of $P_1 J_1$ and $P_2 J_2$ in both cases of row pivoting (Algorithm 3) and no pivoting (Algorithm 1) for $d = 10^{-2}$, 10^{-4}, and 10^{-6}.

Although as shown in Table 1 row pivoting can improve the accuracy, further improvement in the accuracy of the decomposition can be achieved by combining row pivoting with column pivoting (22). In column pivoting, in step i, one chooses a column v_j with the largest ratio $\text{abs}(||v_j||_{S_i})/||v_j||$ as the pivot column. Table 2 shows the residual errors in the triangular decomposition of $\hat{A}$ when column pivoting is used in addition to row pivoting.

d	Row Pivoting			No Pivoting		
	cond(P_2J_2)	cond(P_1J_1)	E	cond(P_2J_2)	cond(P_1J_1)	E
1.0e−2	2.0e2	1.1	7.1e−13	8.8e3	6.4e3	2.3e−11
1.0e−4	2.0e4	1.0	7.0e−11	9.9e7	9.5e7	2.8e−7
1.0e−6	2.0e6	1.0	1.5e−8	1.0e12	1.0e12	5.5e−3

Table 1: Errors and condition numbers of the unified Householder transformations with row pivoting and without pivoting.

d	Row and Column Pivoting		
	cond(P_2J_2)	cond(P_1J_1)	E
1.0e−2	1.3	1.2	6.2e−14
1.0e−4	1.3	1.2	6.0e−14
1.0e−6	1.3	1.2	9.0e−14

Table 2: Errors and condition numbers of the unified Householder transformations with row and column pivoting.

This simultaneous row and column pivoting is akin to Gaussian elimination with complete pivoting.

Now, we present a stability analysis of the direct application of a unified Householder transformation. Our analysis includes both the unitary and hyperbolic Householder transformations as special cases. Since the permutation J in the unified Householder transformation PJ does not cause rounding errors, we only consider the direct application of the S-unitary P to a vector z:

$$\hat{z} \equiv \mathrm{fl}(Pz). \tag{24}$$

Suppose that G is a permutation matrix such that

$$\breve{S} \equiv GSG = \mathrm{diag}(I_p, -I_m).$$

If we partition $S = M + N$ where M is the positive part of S and N the negative part, then

$$GMG = \mathrm{diag}(I_p, 0) \quad \text{and} \quad GNG = \mathrm{diag}(0, -I_m). \tag{25}$$

Thus $\breve{P} \equiv GPG$ is $\breve{S}$-unitary or hyperbolic. Rewriting (24) as $\hat{z} = \mathrm{fl}(Pz) = \mathrm{fl}(G\breve{P}Gz)$, we obtain

$$G\hat{z} = \mathrm{fl}(\breve{P}(Gz)).$$

It is shown in [14] that the direct application of a hyperbolic Householder is relationally stable. Using the partition

$$G\hat{z} = \begin{pmatrix} \hat{x} \\ \hat{y} \end{pmatrix} \quad \text{and} \quad Gz = \begin{pmatrix} x \\ y \end{pmatrix}, \tag{26}$$

the relational stability of the direct application of the hyperbolic transformation $\breve{P}$ means that there is a unitary Householder matrix H such that

$$H \begin{pmatrix} \hat{x} \\ y \end{pmatrix} = \begin{pmatrix} x \\ \hat{y} \end{pmatrix} + f$$

where $\|f\|$ is small compared to $\|x\|$, $\|y\|$, $\|\hat{x}\|$, and $\|\hat{y}\|$ [14]. Using the partition (26), it follows from the above equation that

$$\begin{aligned} & GHG \left[G \begin{pmatrix} I_p & 0 \\ 0 & 0 \end{pmatrix} G\hat{z} - G \begin{pmatrix} 0 & 0 \\ 0 & -I_m \end{pmatrix} Gz \right] \\ = \;& G \begin{pmatrix} I_p & 0 \\ 0 & 0 \end{pmatrix} Gz - G \begin{pmatrix} 0 & 0 \\ 0 & -I_m \end{pmatrix} G\hat{z} + f. \end{aligned}$$

Finally, using (25), we have the following unified equation for the stability:

$$GHG\,(M\hat{z} - Nz) = Mz - N\hat{z} + f \tag{27}$$

where $\|f\|$ is small compared to $\|z\|$ and $\|\hat{z}\|$. Note that Mz is the vector obtained by setting the elements of z corresponding to -1 in S to zero and Nz can be obtained similarly. When $G = I$ ($S = \mathrm{diag}(I_p, -I_m)$), the above equation (27) shows the relational stability of the hyperbolic Householder transformation presented in [14]. When $S = I = M$ (or $S = -I = N$), the above equation (27) becomes

$$H\hat{z} = z + f \quad \text{or} \quad Hz = \hat{z} + f,$$

which shows the backward stability of the unitary Householder transformation.

4.2 Unified Rotations

Now we turn to the unified rotation defined in (23). Similar to the previous section, we first derive the condition number of a unified rotation. Then we propose a pivoting strategy. Finally, we discuss the stability.

It can be verified that the two eigenvalues of $Q^{\mathrm{H}}Q$, where Q is defined in (23), are:

$$\lambda = |c|^2 + |s|^2 \pm |\sigma_2/\sigma_1 - 1|\,|c|\,|s|.$$

It then follows that the condition number of Q is

$$\mathrm{cond}(Q) = \sqrt{\frac{|c|^2 + |s|^2 + |\sigma_2/\sigma_1 - 1|\,|c|\,|s|}{|c|^2 + |s|^2 - |\sigma_2/\sigma_1 - 1|\,|c|\,|s|}}\,.$$

In particular, when $\sigma_2 = \sigma_1$, then Q is unitary and $\mathrm{cond}(Q) = 1$. When $\sigma_2 = -\sigma_1 = -1$,

$$\mathrm{cond}(Q) = \frac{|c| + |s|}{||c| - |s||} \tag{28}$$

which can be arbitrarily large.

To avoid rotations with large condition numbers during a sequence of updating and downdating, we can apply the following pivoting strategy. If $\sigma_2 = -\sigma_1$ when a rotation is computed to eliminate an element, we can choose, if possible, the element to be eliminated so that the condition number (28) is minimized. In particular, if the rotation is computed by Algorithm 2, we choose x_2 so that $(|x_1| + |x_2|)/\,|\,|x_1| - |x_2|\,|)$ is minimized.

Finally, we discuss the stability of the unified rotation. It is shown [13] that the direct application of hyperbolic rotation is neither backward stable nor relationally stable. However, Chambers' algorithm [4] and the LINPACK algorithm [6] are relationally stable [2, 12], but not backward stable. It is also shown [13] that the relational stability extends to a sequence of updating and downdating. In the following, we present a unified algorithm which performs a direct application when the unified rotation Q is unitary and LINPACK method when Q is hyperbolic. Thus it is backward stable when Q is unitary and relationally stable when Q is hyperbolic.

The unified rotation Q satisfies

$$Q^{\mathrm{H}} \begin{pmatrix} \hat{\sigma}_1 & 0 \\ 0 & \hat{\sigma}_2 \end{pmatrix} Q = \begin{pmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{pmatrix} \tag{29}$$

where $\sigma_1|c|^2 + \sigma_2|s|^2 = \hat{\sigma}_1$ ($\hat{\sigma}_1 = \pm 1$) and $\hat{\sigma}_2 = (\sigma_2/\sigma_1)\hat{\sigma}_1$. Consider the direct application of Q to a 2-by-1 vector:

$$\begin{pmatrix} \breve{x} \\ \breve{y} \end{pmatrix} = Q \begin{pmatrix} x \\ y \end{pmatrix}. \tag{30}$$

It follows from (29) that

$$\begin{pmatrix} \breve{x} \\ \breve{y} \end{pmatrix}^{\mathrm{H}} \begin{pmatrix} \hat{\sigma}_1 & 0 \\ 0 & \hat{\sigma}_2 \end{pmatrix} \begin{pmatrix} \breve{x} \\ \breve{y} \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix}^{\mathrm{H}} \begin{pmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

and

$$|\breve{x}|^2 + \frac{\sigma_1}{\sigma_2}|\breve{y}|^2 = \frac{\sigma_1}{\hat{\sigma}_1}\left(|x|^2 + \frac{\sigma_1}{\sigma_2}|y|^2\right). \tag{31}$$

When $\sigma_1 = \sigma_2$, thus $\hat{\sigma}_1 = \hat{\sigma}_2$, Q is unitary. When $\sigma_1 = -\sigma_2$ and $\sigma_1 = \hat{\sigma}_1$, (31) becomes $|\breve{x}|^2 + |y|^2 = |x|^2 + |\breve{y}|^2$ and there exists a unitary $\breve{Q}$ such that

$$\breve{Q} \begin{pmatrix} \breve{x} \\ y \end{pmatrix} = \begin{pmatrix} x \\ \breve{y} \end{pmatrix}. \tag{32}$$

When $\sigma_1 = -\sigma_2$ and $\sigma_1 = -\hat{\sigma}_1$, (31) becomes $|\breve{x}|^2 + |x|^2 = |y|^2 + |\breve{y}|^2$ and there exists a unitary $\tilde{Q}$ so that

$$\tilde{Q} \begin{pmatrix} \breve{x} \\ x \end{pmatrix} = \begin{pmatrix} y \\ \breve{y} \end{pmatrix}. \tag{33}$$

Suppose the (c,s)-pair in the unified rotation Q has been computed, we first use (23) and (30) to compute $\breve{x}$:

$$\breve{x} = c^* x + \frac{\sigma_2}{\sigma_1} s^* y. \tag{34}$$

Also, from (23) and (30), we have

$$\breve{y} = -sx + cy. \tag{35}$$

If $\sigma_1 = -\sigma_2$ and $\sigma_1 = \hat{\sigma}_1$, we compute $\breve{y}$ from $\breve{x}$ and y using (32). Solving for x in (34), we have

$$x = \frac{1}{c^*}\left(\breve{x} - \frac{\sigma_1}{\sigma_2} s^* y\right)$$

Substituting x in (35) with the above equation, we get

$$\breve{y} = \frac{1}{c^*}(-s\breve{x} + y).$$

Note that in this case $|c|^2 - |s|^2 = 1$. Similarly, if $\sigma_1 = -\sigma_2$ and $\sigma_1 = -\hat{\sigma}_1$, we compute $\breve{y}$ from $\breve{x}$ and x using (33). ¿From (34) and (35), we obtain

$$\breve{y} = -\frac{1}{s^*}(c\breve{x} + x),$$

noting that in this case $|c|^2 - |s|^2 = -1$.

In summary, we have the following algorithm.

Algorithm 4 (Application of a unified rotation) *Given the (c,s)-pair in a unified rotation Q with its associated σ_i and $\hat{\sigma}_i$ for $i = 1,2$ and a vector $(x,y)^{\mathrm{T}}$, this algorithm computes the vector $(\breve{x},\breve{y})^{\mathrm{T}} = Q(x,y)^{\mathrm{T}}$.*

$\breve{x} = c^* x + (\sigma_2/\sigma_1) s^* y$;
if $\sigma_1 = \sigma_2$
    $\breve{y} = -sx + cy$;
elseif $\sigma_1 = \hat{\sigma}_1$
    $\breve{y} = (-s\breve{x} + y)/c^*$;
else
    $\breve{y} = -(c\breve{x} + x)/s^*$;
end

Conclusion This paper presents unified hypernormal transformations which are generalizations of unitary and hyperbolic transformations. The unified transformations can be applied to the problem of the triangularization of a strongly nonsingular indefinite matrix. We give algorithms for computing the unified transformations, derive their condition numbers, present unified error analysis, and propose pivoting schemes.

References

[1] A.W. Bojanczyk and A. Steinhardt. Hyperbolic Transformations in Signal Processing and Control *Transactions of the Ninth Army Conference on Applied Mathematics and Control*, 479–487, 1992.

[2] A. Bojanczyk, R. P. Brent, P. Van Dooren, and F. de Hoog. A note on downdating the Cholesky factorization. *SIAM J. Scientific Statist. Computing*. 8:210–221, 1987.

[3] A. Bojanczyk and A. O. Steinhardt. Stability analysis of a Householder-based algorithm for downdating the Cholesky factorization. *SIAM Journal on Scientific and Statistical Computing*, 12:1255–1265, 1991.

[4] J. M. Chambers. Regression updating. *J. Am. Statist. Assoc.*, 66:744–748, 1971.

[5] G. Cybenko and M. Berry. Hyperbolic Householder algorithms for factoring structured matrices. *SIAM J. Matrix Anal. Appl.*, 11(4):499–520, 1990.

[6] J. J. Dongarra, J. R. Bunch, C. B. Moler, and G. W. Stewart. *LINPACK User's Guide*. SIAM, Philadelphia, PA, 1979.

[7] I. Gohberg, P. Lancaster, and L. Rodman, *Matrices and Indefinite Scalar Products*. Birkhauser, 1984.

[8] G. H. Golub and C. F. Van Loan. *Matrix Computations*. 3rd edition, The Johns Hopkins University Press, Baltimore, Maryland, 1996.

[9] C. M. Rader and A. O. Steinhardt. Hyperbolic Householder transformations. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 34:1589–1602, 1986.

[10] C. M. Rader and A. O. Steinhardt. Hyperbolic Householder transformations. *SIAM Journal on Matrix Analysis and Applications*, 9:269–290, 1988.

[11] A. Steinhardt. Householder Transformations in Signal Processing. *IEEE ASSP Magazine*, July, 1988.

[12] G. W. Stewart. The effects of rounding error on an algorithm for downdating a Cholesky factorization. *J. Inst. Maths. Applics.*, 23:203-213, 1979.

[13] G. W. Stewart. On the stability of sequential updates and downdates, *IEEE Transactions on Signal Processing*, 43(11):2642–2648, 1995.

[14] M. Stewart and G. W. Stewart. On hyperbolic triangularization: stability and pivoting, *SIAM J. Matrix Anal. Appl.*, 19(4):847–860, 1998.