Explicit Polar Decomposition and a Near-Characteristic Polynomial: The 2×2 Case*

Frank Uhlig
Department of Mathematics
Oregon State University
Corvallis, Oregon 97330
and
Institut für Geometrie und Praktische Mathematik
RWTH Aachen
51 Aachen, West Germany

Submitted by Richard A. Brualdi

ABSTRACT

Explicit algebraic formulas for the polar decomposition of a nonsingular real 2×2 matrix A are given, as well as a classification of all integer 2×2 matrices that admit a rational polar decomposition. These formulas lead to a functional identity which is satisfied by all nonsingular real 2×2 matrices A as well as by exactly one type of exceptional matrix A_n for each $n>2$.

By the polar decomposition theorem, every nonsingular real square matrix A can be factored uniquely into the product of an orthogonal and a positive definite real matrix:

$$A=VP \qquad \text{with} \quad V^tV=I \text{ and } P=P^t>0,$$

where V^t, P^t denote the transposes of V, P, etc. In standard textbooks on linear algebra, the proof given for this classical result is constructive: From the eigenvalues and eigenvectors of A^tA one can construct a positive definite P with $P^2=A^tA$. Then $V:=AP^{-1}$ is orthogonal. This construction was used here initially to obtain

*Presented at the Matrix Theory Conference held at Auburn University, 19–22 March 1980.

52 Vanderbilt Ave., New York, NY 10017 0024-3795/81/040239+11$02.50

THEOREM 1 (Explicit 2×2 polar decomposition). *Every nonsingular* $A\in\mathbb{R}_{22}$ *has the polar decomposition*

$$A=VP \qquad \textit{where} \quad V^tV=I \textit{ and } P=P^t>0$$

for

$$P=\left|\det\left(A+|\det A|(A^t)^{-1}\right)\right|^{-1/2}\left(A^tA+|\det A|I\right)$$

and

$$V=\left|\det\left(A+|\det A|(A^t)^{-1}\right)\right|^{-1/2}\left(A+|\det A|(A^t)^{-1}\right).$$

Proof. Instead of giving the lengthy and rather tedious arithmetic suggested by the constructive proof for the polar decomposition theorem here, we need only verify that V and P as given above are in fact the polar factors of any nonsingular

$$A=\begin{pmatrix} a & b \\ c & d \end{pmatrix}\in\mathbb{R}_{22}.$$

Clearly P is positive definite as the sum of positive definite matrices, and

$$V^tV=\left|\det\left(A+|\det A|(A^t)^{-1}\right)\right|^{-1}\cdot$$
$$\times\left[A^tA+2|\det A|I+(\det A^2)(A^tA)^{-1}\right].$$

Now

$$A^tA=\begin{pmatrix} a^2+c^2 & ab+cd \\ ab+cd & b^2+d^2 \end{pmatrix},$$

and by Cramer's rule

$$(A^tA)^{-1}=\frac{1}{\det A^2}\begin{pmatrix} b^2+d^2 & -(ab+cd) \\ -(ab+cd) & a^2+c^2 \end{pmatrix}.$$

Thus with $\beta := |\det(A+|\det A|(A^t)^{-1})|^{1/2}$ we have

$$V^tV=\frac{1}{\beta^2}\begin{pmatrix} a^2+b^2+c^2+d^2+2|\det A| & 0 \\ 0 & a^2+b^2+c^2+d^2+2|\det A| \end{pmatrix}.$$

Since

$$(A^t)^{-1}=\frac{1}{\det A}\begin{pmatrix} d & -c \\ -b & a \end{pmatrix},$$

we have

$$|\det A|(A^t)^{-1}=\pm\begin{pmatrix} d & -c \\ -b & a \end{pmatrix},$$

the sign depending on whether $\det A$ is positive or negative. If $\det A>0$, then

$$\begin{aligned}\det\left[A+(\det A)(A^t)^{-1}\right]&=\det\begin{pmatrix} a+d & b-c \\ c-b & d+a \end{pmatrix}\\ &=(a+d)^2+(b-c)^2\\ &=a^2+b^2+c^2+d^2+2\det A\\ &=\beta^2>0.\end{aligned}$$

If $\det A<0$, then

$$\begin{aligned}\beta^2&=|\det\left(A+|\det A|(A^t)^{-1}\right)|\\ &=\left|\det\begin{pmatrix} a-d & b+c \\ c+b & d-a \end{pmatrix}\right|\\ &=|-(a-d)^2-(b+c)^2|\\ &=a^2+b^2+c^2+d^2+2|\det A|.\end{aligned}$$

Thus in either case $V^tV=I$.

It remains to show that $A=VP$. Let us take the case $\det A<0$ for example: Then

$$V=\frac{1}{\beta}\begin{pmatrix} a-d & b+c \\ c+b & d-a \end{pmatrix}$$

and

$$P=\frac{1}{\beta}\begin{pmatrix} a^2+c^2-(ad-bc) & ab+cd \\ ab+cd & b^2+d^2-(ad-bc) \end{pmatrix}$$

and

$$VP=\frac{1}{\beta^2}\begin{pmatrix} a-d & b+c \\ c+b & d-a \end{pmatrix}\begin{pmatrix} a^2+c^2-(ad-bc) & ab+cd \\ ab+cd & b^2+d^2-(ad-bc) \end{pmatrix}$$

$$=\frac{1}{\beta^2}\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix},$$

where

$$a_{11}=a\left[a^2+c^2-(ad-bc)\right]-da^2-dc^2+ad^2-bcd+ab^2+bcd+abc+c^2d$$

$$=a\left[a^2+c^2-(ad-bc)-ad+d^2+b^2+bc\right]$$

$$=a\left[a^2+b^2+c^2+d^2-2(ad-bc)\right]=a\beta^2.$$

Similarly $a_{12}=b\beta^2$, $a_{21}=c\beta^2$, and $a_{22}=d\beta^2$, and thus

$$A=\begin{pmatrix} a & b \\ c & d \end{pmatrix}=VP.$$

The case $\det A>0$ can be verified analogously. ■

One can apply Theorem 1 to classify all integer 2×2 matrices whose polar decomposition can be achieved over the rationals:

THEOREM 2 (Rational polar decomposition for nonsingular integer 2×2 matrices). *A nonsingular matrix*

$$A=\begin{pmatrix} a & b \\ c & d \end{pmatrix}\in\mathbb{Z}_{22}$$

has a polar decomposition over $\mathbb{Q}$ *iff*

> *in case* $\det A<0$: $a-d$ *and* $b+c$ *form a pythagorean pair of integers, and in case* $\det A>0$: $a+d$ *and* $b-c$ *do.*

Proof. In case $\det A>0$, we have $0<\det[A+(\det A)(A^t)^{-1}]=(a+d)^2+(b-c)^2$ from the previous proof. Thus in this case $P\in\mathbb{Q}_{22}$ and $V\in\mathbb{Q}_{22}$ iff $\beta^2=(a+d)^2+(b-c)^2$ is an integer square. The proof in case $\det A<0$ is similar. ■

Since $A=VP$ for all nonsingular real 2×2 matrices, we have

THEOREM 3 (Matrix identity for all nonsingular real 2×2 matrices). *Every nonsingular* $A\in\mathbb{R}_{22}$ *satisfies the identity*

$$\left|\det\left[A+|\det A|(A^t)^{-1}\right]\right|A=AA^tA+2|\det A|A+(\det A^2)(A^t)^{-1}. \quad (*_2)$$

Equation $(*_2)$ is a rather peculiar matrix identity. We will study it in the remainder of this paper.

If we multiply through by A^t from the right and set $X:=AA^t$, then $(*_2)$ becomes

$$\left|\det\left[A+|\det A|(A^t)^{-1}\right]\right|X=X^2+2|\det A|X+(\det X)I,$$

or

$$X^2-\left\{\left|\det\left[A+|\det A|(A^t)^{-1}\right]\right|-2|\det A|\right\}X+(\det X)I=0.$$

This last equation must hold for all positive definite 2×2 real matrices X by Theorem 3. In fact, this identity coincides with the characteristic polynomial of such X for $n=2$:

LEMMA 1. *For all nonsingular* $A\in\mathbb{R}_{22}$ *one has*

$$\operatorname{tr}AA^t=\left|\det\left[A+|\det A|(A^t)^{-1}\right]\right|-2|\det A|.$$

Proof. Without loss of generality assume

$$A=\begin{pmatrix} a & b\\ c & d\end{pmatrix}\in\mathbb{R}_{22} \qquad \text{with} \quad ad-bc>0.$$

Then $\operatorname{tr} AA^t = a^2 + b^2 + c^2 + d^2$, while $|\det[A + |\det A|(A^t)^{-1}]| - 2|\det A| = (a+d)^2 + (b-c)^2 - 2(ad-bc)$ from the proof of Theorem 2. ■

Next we want to investigate the nonsingular matrices $A \in \mathbb{R}_{nn}$ that satisfy $(*_2)$ for $n \neq 2$. In fact, we are able to classify all such $n \times n$ matrices:

THEOREM 4. *A nonsingular matrix $A \in \mathbb{R}_{nn}$ satisfies $(*_2)$ iff either A is 2×2 or $(1/\sqrt{\alpha_n})A$ is orthogonal for $n > 2$, where α_n is the unique positive root of the equation $X^{n/2-1}(1 + X^{n/2-1})^{n-2} - 1 = 0$.*

Proof. Assume that $A \in \mathbb{R}_{nn}$ nonsingular satisfies $(*_2)$. Then $X := AA^t$ satisfies the quadratic polynomial

$$X^2 + \left\{2|\det A| - \left|\det\left[A + |\det A|(A^t)^{-1}\right]\right|\right\}X + (\det X)I = 0, \qquad \text{(NCP)}$$

which we will call a *near-characteristic polynomial* of X for reasons that will become clear later. Thus X has at most two distinct eigenvalues. And either

Case (a): X is similar to

$$\left(\begin{array}{c|c} \lambda I_p & 0 \\ \hline 0 & \mu I_q \end{array}\right)$$

with $p, q \geqslant 1$, $p + q = n$, $\lambda \neq \mu$, $\lambda, \mu > 0$ as eigenvalues of $AA^t = X$, or
Case (b): X is a multiple of the identity, $X = \alpha I$, $\alpha \in \mathbb{R}$.

We will show first that case (a) cannot occur unless $n = 2$. In case (a) X satisfies both its minimum polynomial

$$(X - \lambda I)(X - \mu I) = 0$$

and its "near-characteristic polynomial" (NCP). Thus

$$X^2 - (\lambda + \mu)X + \lambda\mu I = X^2 + \left\{2|\det A| - \left|\det\left[A + |\det A|(A^t)^{-1}\right]\right|\right\}X$$
$$+ (\det X)I = 0,$$

where $\det X = \lambda^p \mu^q = (\det A)^2$. By a uniqueness argument we conclude that

$$\lambda\mu = \det X = \lambda^p \mu^q \qquad (1)$$

and

$$\lambda+\mu=\left|\det\left[A+|\det A|(A^t)^{-1}\right]\right|-2|\det A|. \tag{2}$$

Multiplying (2) by $|\det A^t|$, we get

$$|\det A|(\lambda+\mu)=|\det(X+|\det A|I)|-2\det X. \tag{3}$$

In further investigating (3) we may without loss of generality assume that X is in fact equal to $\begin{pmatrix} \lambda I_p & 0 \\ 0 & \mu I_q \end{pmatrix}$, since a similarity transform of X does not affect the determinants involving X in (3). Then (3) can be rewritten as

$$(\lambda+\mu)|\det A|=(\lambda+|\det A|)^p(\mu+|\det A|)^q-2\lambda^p\mu^q. \tag{4}$$

Set $w:=|\det A|>0$. Then $w=\sqrt{\lambda\mu}$ from (1) and (4) is equivalent to

$$(\lambda+w)^p(\mu+w)^q-(\lambda+\mu)w-2\lambda\mu=0. \tag{5}$$

Expanding the pth and qth powers respectively, one gets

$$\left[w^p+pw^{p-1}\lambda+\cdots+\binom{p}{2}w^2\lambda^{p-2}+pw\lambda^{p-1}+\lambda^p\right]$$
$$\times\left[w^q+qw^{q-1}\mu+\cdots+\binom{q}{2}w^2\mu^{q-2}+qw\mu^{q-1}+\mu^q\right]$$
$$-(\lambda+\mu)w-2\lambda\mu=0,$$

or equivalently

$$w^{p+q}+w^{p+q-1}(q\mu+\lambda p)+\cdots$$
$$+w^2\left[\binom{q}{2}\lambda^p\mu^{q-2}+pq\lambda^{p-1}\mu^{q-1}+\binom{p}{2}\lambda^{p-2}\mu^q\right]$$
$$+w\left[p\lambda^{p-1}\mu^q+q\lambda^p\mu^{q-1}-(\lambda+\mu)\right]+\lambda^p\mu^q-2\lambda\mu=0.$$

From (1), $\lambda^p\mu^q=\lambda\mu$; thus $\lambda^{p-1}\mu^{q-1}=1$ and $\lambda^{p-1}\mu^q=\mu$, $\lambda^p\mu^{q-1}=\lambda$, while $w^2=\lambda\mu$. Thus we get

$$w^{p+q}+w^{p+q-1}(q\mu+\lambda p)+\cdots$$
$$+w^2\left[\binom{q}{2}\lambda^p\mu^{q-2}+pq+\binom{p}{2}\lambda^{p-2}\mu^{q-1}\right]$$
$$+w\left[(p-1)\mu+(q-1)\lambda\right]=0. \quad (6)$$

Unless $p=q=1$, all terms on the left hand side of (6) are positive, contradicting the fact that their sum is zero. Hence if A satisfies $(*_2)$, then $X=AA^t$ cannot have two distinct roots $\lambda\neq\mu$ unless $n=2$.

There remains case (b): $X=AA^t=\alpha I$ for $\alpha>0$. Here we multiply the "near-characteristic polynomial" (NCP) by $|\det A^t|=|\det A|$ to obtain

$$|\det A|X^2+\left[2|\det X|-|\det(X+|\det A|I)|\right]X+|\det A^3|I=0. \quad (7)$$

Now $\det A^2=\det X=\alpha^n$ and $w=|\det A|=\alpha^{n/2}>0$. In case (b), (7) is a matrix identity for the diagonal matrices I, X, X^2, where each diagonal element satisfies the same equation

$$\alpha^{n/2}\alpha^2+\left(2\alpha^n-(\alpha+\alpha^{n/2})^n\right)\alpha+\alpha^{3n/2}=0. \quad (8)$$

Equation (8) can be rewritten thus:

$$\alpha^{n/2}\left(\alpha^2+2\alpha^{n/2}\alpha+(\alpha^{n/2})^2\right)=(\alpha+\alpha^{n/2})^n\alpha,$$

or

$$\alpha^{n/2}(\alpha+\alpha^{n/2})^2=(\alpha+\alpha^{n/2})^n\alpha.$$

And thus

$$\alpha^{n/2-1}=(\alpha+\alpha^{n/2})^{n-2}=\alpha^{n-2}(1+\alpha^{n/2-1})^{n-2},$$

and

$$(1+\alpha^{n/2-1})^{n-2}=\alpha^{-n/2+1}. \quad (9)$$

A matrix $A \in \mathbb{R}_{nn}$ with $n>2$ will satisfy $(*_2)$ iff $AA^t=\alpha_n I$ where $\alpha_n>0$ satisfies (9).

It remains to find out whether the equation (9) has any positive roots. For $n>2$ (9) cannot have a root $\alpha_n>1$, for then the left side of (9) would exceed 1 while the right hand side would be less than one. If $\alpha_n=1$, then (9) becomes $2^{n-2}=1$, or $n=2$. Hence we are left to look for solutions α_n of (9) with $0<\alpha_n<1$. Set $Z:=\alpha^{n/2-1}$, where $0<Z<1$ and $n>2$. Then finding solutions α_n of (9) is equivalent to finding positive roots for the polynomial

$$P(Z)=Z(1+Z)^{n-2}-1. \tag{10}$$

Since $P(0)=-1$ and $P(1)=2^{n-2}-1>0$ for $n>2$, P must have at least one root Z_n between 0 and 1. It can have no more than one root $Z_n>0$, for $Z=-1$ is an $(n-2)$-fold root and $Z=0$ is a simple root of $P(Z)+1$, and hence $P(Z)$ must be monotonic for $Z>0$ (as well as for $Z<-1$). Thus for each $n>2$ there is exactly one number $\alpha_n>0$ such that all $A\in\mathbb{R}_{nn}$ with $AA^t=\alpha_n I$ satisfy $(*_2)$. For $n=1$, $AA^t=(\alpha)$ for $\alpha>0$ and Eq. $(*_2)$ does not hold, as can readily be seen.

Note that the sequence $\{\alpha_n^{n/2-1}\}$ must converge to zero, but slowly enough so that $\{(1+\alpha_n^{n/2-1})^{n-2}\}$ diverges to ∞, their product being equal to one for each $n>2$. ■

COMMENT 1. For the exceptional matrices $A\in\mathbb{R}_{nn}$ that satisfy $(*_2)$ for $n>2$, $|\det[A+|\det A|(A^t)^{-1}]|-2|\det A|$ does not equal $\operatorname{tr} AA^t$.

Proof. We have $\operatorname{tr} AA^t=\operatorname{tr} X=\operatorname{tr}\alpha_n I=n\alpha_n$, while

$$\Big\{\Big|\det\Big[A+|\det A|(A^t)^{-1}\Big]\Big|-2|\det A|\Big\}(\det A^t)$$
$$=|\det(X+|\det A|I)|-2|\det X|$$
$$=\big(\alpha_n+\alpha_n^{n/2}\big)^n-2\alpha_n^n.$$

The question is whether an $\alpha_n>0$ that satisfies (9) might also satisfy

$$n\alpha_n\alpha_n^{n/2}=\big(\alpha_n+\alpha_n^{n/2}\big)^n-2\alpha_n^n. \tag{11}$$

Equation (11) is equivalent to $n\alpha_n^{-n/2+1}=(1+\alpha_n^{n/2-1})^n-2$, and by combining this with (9) we get

$$n\alpha_n^{-n/2+1}=\alpha_n^{-n/2+1}\big(1+\alpha_n^{n/2-1}\big)^2-2,$$

or

$$n\alpha_n^{-n/2+1} = \alpha_n^{-n/2+1}\left(1+2\alpha_n^{n/2-1}+\alpha_n^{n-2}\right)-2$$
$$= \alpha_n^{-n/2+1}+2+\alpha_n^{n/2-1}-2.$$

Thus

$$0=\alpha_n^{-n/2+1}\left(1+\alpha_n^{n-2}-n\right). \quad (12)$$

Since $\alpha_n>0$ for all $n>2$, (12) implies $n=1+\alpha_n^{n-2}$. But $0<\alpha_n<1$, so that $1+\alpha_n^{n-2}<2<n$. Hence no $A\in\mathbb{R}_{nn}$ with $AA^t=\alpha_n I$, α_n as in Theorem 4, satisfies (11) if $n>2$. ∎

The "scaled orthogonal matrices" A_n from Theorem 4 that satisfy $(*_2)$ are such that $X=A_nA_n^t$ does not satisfy the 2×2 characteristic polynomial $X^2-(\operatorname{tr}X)X+(\det X)I=0$, but instead each $X=A_nA_n^t$ satisfies the "near-characteristic polynomial"

$$X^2+\left\{2|\det A|-\left|\det\left[A+|\det A|(A^t)^{-1}\right]\right|\right\}X+(\det X)I=0. \quad \text{(NCP)}$$

These A_n are truly exceptional

Comment 2. The function $|\det[A+|\det A|(A^t)^{-1}]|-2|\det A|$ acts like the trace function of AA^t for all nonsingular real 2×2 matrices A, as shown in Lemma 1. It also acts like the trace function for various $n\times n$ matrices, though not for the exceptional matrices A_n of Theorem 4.

Example. If $A=\beta I$ for $\beta>0$, $n>2$, then

$$\left|\det\left[A+|\det A|(A^t)^{-1}\right]\right|-2|\det A|=(\beta+\beta^{n-1})^n-2\beta^n$$

and $\operatorname{tr}AA^t=n\beta^2$. Hence for equality, we have to solve

$$\beta^{n-2}(1+\beta^{n-2})^n-2\beta^{n-2}-n=0. \quad (13)$$

With $Z:=\beta^{n-2}>0$ we need to find a positive solution of

$$q(Z)=Z(1+Z)^n-2Z-n=0. \quad (14)$$

Now $q(0)=-n$ and $q^{(1)}=2^n-2-n>0$ for $n>2$. Thus for each $n>2$ there is at least one $\beta_n>0$ such that for $A=\beta_n I$, the function

$$\left|\det\left[A+|\det A|(A^t)^{-1}\right]\right|-2|\det A|$$

acts like $\operatorname{tr} AA^t$.

The complete set of $A\in\mathbb{R}_{nn}$ for which

$$\left|\det\left[A+|\det A|(A^t)^{-1}\right]\right|-2|\det A|=\operatorname{tr} AA^t$$

is not known.

COMMENT 3. The significance of the "near-characteristic polynomial" or of $(*_2)$ is not fully understood at the moment, and neither is the role of the exceptional "scaled orthogonal matrices" A_n for $n>2$. It would be of great interest to obtain explicit algebraic formulas for the polar decomposition of 3×3 matrices and the corresponding "near-characteristic polynomial" and $(*_3)$ (or $(*_n)$ in general). Amazingly, there seems to be no literature on the subject thus far, except—of course—for various algorithms, like the singular value decomposition or the Cholesky decomposition, which can be used for an approximative polar decomposition of A.

I am grateful to Tom Laffey for his help with the "exceptional" A_n.

Received 27 May 1980