Note on the factorization of a square matrix into two Hermitian or symmetric matrices

SIAM REVIEW
Vol. 29, No. 3, September 1987

003

CLASSROOM NOTES

EDITED BY MURRAY S. KLAMKIN

This section contains brief notes which are essentially self-contained applications of mathematics that can be used in the classroom. New applications are preferred, but exemplary applications not well known or readily available are accepted.

Both "modern" and "classical" applications are welcome, especially modern applications to current real world problems.

Notes should be submitted to M. S. Klamkin, Department of Mathematics, University of Alberta, Edmonton, Alberta, Canada T6G 2G1.

NOTE ON THE FACTORIZATION OF A SQUARE MATRIX INTO TWO HERMITIAN OR SYMMETRIC MATRICES*

A. J. BOSCH†

Abstract. This paper presents elementary proofs of the factorization of a square matrix into two hermitian or symmetric matrices.

Key words. Jordan normal form, matrix factorization, symmetric matrices

AMS(MOS) subject classification. 15A23

1. Introduction. As we will see, every square matrix (real or complex) is a product of two symmetric matrices (real or complex, respectively). However, not every square matrix is a product of two hermitian matrices. Although these results were already published by Frobenius in 1910 (see [2]), they are still not well known to mathematicians. They are not even found in modern textbooks on matrix theory or linear algebra. Consequently, these results and their proofs (see [1], [4], [5]) are not very accessible to nonmathematicians. But they can use these results. Applications can be found in mechanics, system theory, structural analysis, etc., and we give one for a mechanical system at the end. The aim of this paper is to give elementary proofs as well as a clear summary of the conditions. The basis of all proofs is the Jordan normal form of a matrix.

2. Notation.

A is a complex or real square matrix; A^T is the transpose of A;
$A^* = \bar{A}^T$ the conjugate transpose of A; $A^{-T} = (A^T)^{-1} = (A^{-1})^T$;
Λ is a diagonal matrix of eigenvalues;
S is a real symmetric matrix: $S = \bar{S} = S^T$; C a symmetric matrix: $C^T = C$;
H denotes a hermitian matrix: $H^* = H$; U a unitary matrix: $UU^* = U^*U = I$;
$A \simeq D$ means A is similar to the matrix D: $A = BDB^{-1}$;

* Received by the editors November 12, 1985; accepted for publication (in revised form) April 23, 1986. Theorems 1 and 4 appeared in *The Factorization of a Square Matrix into Two Symmetric Matrices*, Amer. Math. Monthly, 93 (1986) pp. 462–464.

† Department of Mathematics, Eindhoven University of Technology, Eindhoven, the Netherlands.

$H > 0$ means H is positive definite: for all vectors $x \neq 0$: $x^*Hx > 0$;

$$J_k(\lambda) = \begin{bmatrix} \lambda & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda \end{bmatrix}; \qquad S_k = \begin{bmatrix} & & 0 & 1 \\ & & \cdot & \\ & \cdot & & \\ 1 & 0 & & \end{bmatrix}; \qquad C_k(\lambda) = \begin{bmatrix} 0 & & & \lambda \\ & & \cdot & 1 \\ & \cdot & \cdot & \\ \lambda & 1 & & 0 \end{bmatrix}$$

all $k \times k$-matrices;

$$A = \operatorname{diag} A_i = A_1 \oplus A_2 \oplus \cdots \oplus A_r = \begin{bmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_r \end{bmatrix};$$

the A_i are square matrices with the diagonal along the diagonal of A.

3. Preliminaries. (For the proofs of Propositions 1 and 4 see [3].)

PROPOSITION 1. *Let H be a hermitian matrix. Then there exists a unitary matrix U such that $H = U \Lambda U^*$ with Λ real.*

If $H > 0$, then all eigenvalues λ_i are positive.

PROPOSITION 2. *Let $H_1 > 0$. Then there exists an $H > 0$ such that $H_1 = H^2$.*

Proof. $H_1 = U \Lambda U^* = (U \Lambda^{1/2} U^*)(U \Lambda^{1/2} U^*) =: H^2$ with $H > 0$.

DEFINITION 1. A Jordan matrix J is a square matrix of the form

$$J = \operatorname{diag} J_{k_i}(\lambda_i) = J_{k_1}(\lambda_1) \oplus \cdots \oplus J_{k_r}(\lambda_r);$$

the λ_i are not necessarily different.

DEFINITION 2. A special Jordan matrix is $J_0 = J_1 \oplus J_2 \oplus \bar{J}_1$ where J_2 is real and contains all real λ's of J_0; the λ's in J_1 are all different from the λ's in $\bar{J}_1$ (J_1 or J_2 can be "empty").

PROPOSITION 3. *$J \simeq J_0$ is equivalent with $J \simeq \bar{J}$.*

Proof. i) $J \simeq J_0$, hence $\bar{J} \simeq \bar{J}_0$. But $\bar{J}_0 \simeq J_0$, so $J \simeq \bar{J}$.

ii) $J \simeq \bar{J}$ means that, if λ is an eigenvalue of J, $\bar{\lambda}$ is also an eigenvalue with the same multiplicity and the same Jordan structure, so $J \simeq J_0$.

PROPOSITION 4. (Jordan normal form of a matrix). *Let A be an arbitrary square matrix, then $A \simeq J$. This means $A = BJB^{-1}$ for some invertible matrix B and some Jordan matrix J.*

COROLLARY 1. *If A is real then $A \simeq J_0$.*

Proof. Let $A \simeq J$; $\bar{A} = A$, hence $J \simeq \bar{J}$ and (Proposition 3) $J \simeq J_0$. So $A \simeq J_0$.

PROPOSITION 5. *Every Jordan matrix J is a product of two symmetric matrices.*

Proof. $J = \operatorname{diag} J_{k_i}(\lambda_i) = \operatorname{diag}(S_{k_i} C_{k_i}(\lambda_i)) = \operatorname{diag} S_{k_i} \operatorname{diag} C_{k_i}(\lambda_i) =: SC$, where the first factor S is nonsingular.

Remark 1. If we define

$$C_k(\lambda) = \begin{bmatrix} 0 & 1 & \lambda \\ & \cdot & \\ \cdot & \cdot & \\ 1 & \cdot & \\ \lambda & & 0 \end{bmatrix},$$

then $J_{k_i}(\lambda_i) = C_{k_i}(\lambda_i) S_{k_i}$ and we have $J = CS$ with the second factor nonsingular.

PROPOSITION 6. *Every Jordan matrix J_0 is a product of two hermitian matrices.*

Proof. With Proposition 5:

$$J_0 = J_1 \oplus J_2 \oplus \bar{J}_1 = S_1 C_1 \oplus S_2 C_2 \oplus S_1 \bar{C}_1 = \begin{bmatrix} 0 & & S_1 \\ & S_2 & \\ S_1 & & 0 \end{bmatrix} \begin{bmatrix} 0 & & \bar{C}_1 \\ & C_2 & \\ C_1 & & 0 \end{bmatrix} =: H_1 H_2.$$

Note that H_1 is nonsingular.

4. Theorems on factorization.

THEOREM 1. *Every square matrix A is a product of two symmetric matrices.*

Proof. If $A \simeq J$ then with Proposition 5:

$$A = BJB^{-1} = B(SC)B^{-1} = (BSB^{\mathrm{T}})(B^{-\mathrm{T}}CB^{-1}) := C_1 C_2.$$

Note that here C_1 is nonsingular.

COROLLARY 2. *Every square matrix A is similar to its transpose: $A \simeq A^{\mathrm{T}}$.*

Proof. $A = C_1 C_2$ (see Theorem 1) with C_1 nonsingular. $C_1^{-1} A C_1 = C_2 C_1 = A^{\mathrm{T}}$ and hence $A \simeq A^{\mathrm{T}}$.

THEOREM 2. *A square matrix A is a product of two hermitian matrices, of which at least one is nonsingular, iff $J \simeq J^*$.*

Proof. i) Let $A = H_1 H_2$ and suppose H_1 is nonsingular.

$$A^* = H_2 H_1 = H_1^{-1}(H_1 H_2) H_1 = H_1^{-1} A H_1.$$

Hence, $A \simeq A^*$ and $J \simeq J^*$.

ii) $J \simeq J^* \simeq \bar{J}$ hence $J \simeq J_0$ (Proposition 3) and $A \simeq J_0$. With Proposition 6: $J_0 = H_1 H_2$; $A = BJ_0 B^{-1} = B(H_1 H_2)B^{-1} = (BH_1 B^*)(B^{*-1} H_2 B^{-1})$. Both factors are hermitian and the first is nonsingular.

THEOREM 3. *Every square matrix A with real Jordan matrix J is a product of two hermitian matrices.*

Proof. J is real so $J = \bar{J} \simeq J^*$ (Corollary 2) and the condition of Theorem 2 is satisfied.

THEOREM 4. *Every real square matrix A is a product of two real symmetric matrices of which at least one is nonsingular.*

Proof. With Corollary 1 $A \simeq J_0$. Let $A = DJ_0 D^{-1}$ and $J_0 = SC$ as in Proposition 5 with S nonsingular.

Partition S and D in the same way as J_0 with respect to the columns:

$$D = (D_1 D_2 D_3); \quad J_0 = J_1 \oplus J_2 \oplus \bar{J}_1; \quad S = S_1 \oplus S_2 \oplus S_1;$$

$$AD = (AD_1\, AD_2\, AD_3) = DJ_0 = (D_1 J_1\, D_2 J_2\, D_3 \bar{J}_1); \quad AD_1 = D_1 J_1$$

$$\text{and hence} \quad A\bar{D}_1 = \bar{D}_1 \bar{J}_1; \quad AD_2 = D_2 J_2.$$

This means that there exists a *real* matrix B_2 of full rank such that $AB_2 = B_2 J_2$. Define the matrix $B := (D_1 B_2 \bar{D}_1)$ then $AB = BJ_0$. B is nonsingular: D_1 (so $\bar{D}_1$) and B_2 are of full rank. The columns of D_1 and $\bar{D}_1$, and the columns of $D_1(\bar{D}_1)$ and B_2 are mutually independent, being (generalized) eigenvectors associated with different eigenvalues (see Definition 2). Hence B is nonsingular.

In the proof of Theorem 1, we saw that $A = C_1 C_2 = (BSB^{\mathrm{T}})C_2$ and C_1 nonsingular. Now

$$C_1 = BSB^{\mathrm{T}} = (D_1 B_2 \bar{D}_1)(S_1 \oplus S_2 \oplus S_1)(D_1 B_2 \bar{D}_1)^{\mathrm{T}}$$
$$= (D_1 S_1 D_1^{\mathrm{T}} + \bar{D}_1 S_1 \bar{D}_1^{\mathrm{T}}) + B_2 S_2 B_2^{\mathrm{T}}$$

which is, as sum of two real matrices, also real.

$C_2 = C_1^{-1} A$ is, as a product of two real matrices, also real.

THEOREM 5. *A square matrix A is a product of two hermitian matrices of which at least one is positive definite iff $J = J^*$.*

Proof. i) Let $A = H_1H_2$ with $H_1 > 0$. With Proposition 2 $H_1 = H^2$; $A = H(HH_2H)H^{-1}$; HH_2H is hermitian, so by Proposition 1 $HH_2H = U \wedge U^*$ with $\wedge$ real; $A = H(U \wedge U^*)H^{-1} = (HU) \wedge (HU)^{-1} =: B \wedge B^{-1}$. Hence, $\wedge = J = J^*$.

ii) $J = J^*$ means $J = \wedge$ real.

$$A = B \wedge B^{-1} = (BB^*)(B^{*^{-1}} \wedge B^{-1}) =: H_1H_2 \quad \text{with} \, H_1 > 0.$$

THEOREM 6. *A real square matrix A is a product of two real symmetric matrices of which at least one is positive definite, iff $J = J^*$.*

Proof. This follows from the proof in Theorem 5:

i) Replace each H by S and U by an orthogonal matrix G.

ii) J and A are real. From $A = D \wedge D^{-1}$ it follows that there exists a *real* nonsingular matrix B such that $A = B \wedge B^{-1}$. So $H_1 = S_1$ and $H_2 = S_2$; $A = S_1S_2$ with $S_1 > 0$.

THEOREM 7. *A real square matrix A is a product of two symmetric positive definite matrices iff $J = J^* > 0$.*

Proof. This follows from the proof in Theorem 5 (and Theorem 6 for the real case).

i) In Theorem 5i), let $A = H_1H_2$ with H_1 and $H_2 > 0$. Then $HH_2H > 0$, so $HH_2H = U \wedge U^*$ with $\wedge > 0$.

ii) $\wedge > 0$, hence $H_2 = B^{*-1} \wedge B^{-1} > 0$.

See Table 1.

TABLE 1

Theorem	A	J	iff-condition	Factorization	
1	complex	complex	–	$A = C_1C_2$	
2	complex	complex	$J \simeq J^*$	$A = H_1H_2$	
3	complex	real	–	$A = H_1H_2$	
4	real	complex	–	$A = S_1S_2$	
5	complex	complex	$J = J^*$	$A = H_1H_2$	H_1 or $H_2 > 0$
6	real	complex	$J = J^*$	$A = S_1S_2$	S_1 or $S_2 > 0$
7	real	complex	$J = J^* > 0$	$A = S_1S_2$	S_1 and $S_2 > 0$

5. Application. Given a mechanical system governed by the equation

$$\ddot{x} + Ax = 0, \tag{1}$$

where A is a real $n \times n$ matrix and $x = x(t)$ an R^n-valued function. We try to find an invariant quadratic form (a Lyapunov function) $V(x, \dot{x})$. To this extent, we rewrite (1) as $\dot{\underset{\sim}{x}} = B\underset{\sim}{x}$, where

$$B = \begin{bmatrix} 0 & I \\ -A & 0 \end{bmatrix}, \quad \underset{\sim}{x} = \begin{bmatrix} x \\ \dot{x} \end{bmatrix},$$

and we search for V of the form $V(\underset{\sim}{x}) = \underset{\sim}{x}^{\mathrm{T}} P \underset{\sim}{x}$ (P symmetric).

$$\dot{V}(\underset{\sim}{x}) = \dot{\underset{\sim}{x}}^{\mathrm{T}} P \underset{\sim}{x} + \underset{\sim}{x}^{\mathrm{T}} P \dot{\underset{\sim}{x}} = \underset{\sim}{x}^{\mathrm{T}} B^{\mathrm{T}} P \underset{\sim}{x} + \underset{\sim}{x}^{\mathrm{T}} P B \underset{\sim}{x} = \underset{\sim}{x}^{\mathrm{T}} (B^{\mathrm{T}} P + PB) \underset{\sim}{x}.$$

The invariance of $V(\underset{\sim}{x})$ reduces to the equation $B^{\mathrm{T}}P + PB = 0$ which when P is written as

$$\begin{bmatrix} P_{11} & P_{12} \\ P_{12}^{\mathrm{T}} & P_{22} \end{bmatrix},$$

can be rewritten as

$$A^{T}P_{12}^{T}+P_{12}A=0,$$

$$P_{12}^{T}+P_{12}=0,$$

$$P_{11}-P_{22}A=0.$$

The first two equations can be satisfied by $P_{12}=0$, and the third by any symmetric P_{22}, P_{11} with P_{22} nonsingular and satisfying $A=P_{22}^{-1}P_{11}$. It follows from Theorem 4 that this factorization is always possible and moreover gives a factorization explicitly. The study of the stability of the system is now reduced to the study of the factors P_{11} and P_{22}. The origin is stable iff $V(\underset{\sim}{x})$ is definite (positive or negative) for $\underset{\sim}{x}\neq\underset{\sim}{0}$. Hence for P_{11}, $P_{22}>0$ (or both negative definite). From Theorem 7, we see that this is iff $A\simeq\Lambda>0$ (i.e., A is nondefective with positive eigenvalues).

6. Example. Consider the following mechanical system: two bars AB and BC of length 1 are connected by hinges in A and B with a torsion-stiffness k (Fig. 1). There are point-masses m in B and C. The masses of the bars are neglected. These bars move in a vertical plane under influence of a constant follower force P in C in the direction BC. We want to study small vibrations (around $\varphi_1=\varphi_2=0$) of this system, in particular the stability of the system.

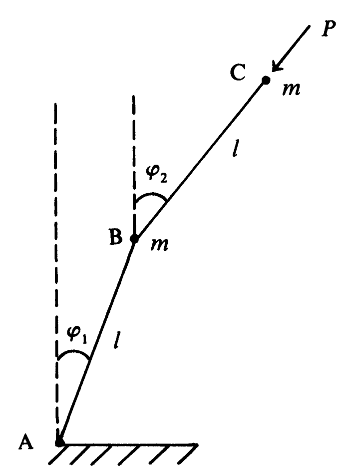

FIG. 1

The system has 2 degrees of freedom. As generalized coordinates, we take the angles φ_1 and φ_2 (see Fig. 1). We give, without derivation, the relations:

kinetic energy:	$T=\frac{1}{2}ml^2(2\dot{\varphi}_1^2+2\dot{\varphi}_1\dot{\varphi}_2\cos(\varphi_1-\varphi_2)+\dot{\varphi}_2^2)$;
elastic energy:	$W=\frac{1}{2}k(\varphi_1^2+(\varphi_2-\varphi_1)^2)$;
generalized follower force:	$Q_{\varphi_1}=-Pl\sin(\varphi_2-\varphi_1)$; $Q_{\varphi_2}=0$.

The equations of motion are:

$$\frac{d}{dt}\frac{\delta T}{\delta\dot{\varphi}_i}-\frac{\delta T}{\delta\varphi_i}+\frac{\delta W}{\delta\varphi_i}=Q_{\varphi_i},\qquad i=1,2;$$

After linearizing and replacing sin a by a, cos a by 1 for small a, we get:

$$2ml^2\ddot{\varphi}_1+ml^2\ddot{\varphi}_2+2k\varphi_1-k\varphi_2-Pl\varphi_1+Pl\varphi_2=0,$$

$$ml^2\ddot{\varphi}_1+ml^2\ddot{\varphi}_2-k\varphi_1+k\varphi_2=0.$$

In matrix notation: $M\ddot{\varphi}+K\varphi=0$ or $\ddot{\varphi}+A\varphi=0$, where $A=M^{-1}K$;

$$\varphi=\begin{bmatrix}\varphi_1\\ \varphi_2\end{bmatrix};\quad M=\begin{bmatrix}2ml^2 & ml^2\\ ml^2 & ml^2\end{bmatrix};\quad K=\begin{bmatrix}2k-Pl & -k+Pl\\ -k & k\end{bmatrix}.$$

i) Suppose $P=0$. Then M and K are both symmetric and positive definite. The system is stable which follows from §5.

ii) Suppose $P>0$. Then K is not symmetric.

$$A=\frac{1}{ml^2}\begin{bmatrix}3k-Pl & -2k+Pl\\ -4k+Pl & 3k-Pl\end{bmatrix}$$

with eigenvalues

$$\lambda_{1,2}=\frac{1}{ml^2}(3k-Pl\pm\sqrt{(3k-Pl)^2-k^2}).$$

From §5 it follows that λ must be real and positive. Hence the system is stable for $P<2k/l$; in this case $\lambda_1\neq\lambda_2$ and A is not defective. However, if $P=2k/1$ then $\lambda_1=\lambda_2=k/ml^2$ and

$$A=\frac{1}{ml^2}\begin{bmatrix}k & 0\\ -2k & k\end{bmatrix}$$

is defective. Theorem 7 tells us that a defective matrix is not a product of two positive definite matrices. Hence for $P=2k/l$ the system is not stable.

Acknowledgments. I thank Dr. Laffey, University College, Dublin, who drew my attention to Corollary 2 and Professors M. L. J. Hautus and W. J. Kuÿpers, Eindhoven University, for the application in §5.

REFERENCES

[1] D. H. Carlson, *On real eigenvalues of complex matrices*, Pacific J. Math., 15 (1965), pp. 1119–1129.

[2] G. Frobenius, *Ueber die mit einer Matrix vertauschbaren Matrizen*, Sitzungsber. Preuss. Akad. für Wiss., (1910) pp. 3–15.

[3] Ben Noble and J. W. Daniel, *Applied Linear Algebra*, Prentice-Hall, Englewood Cliffs, NJ, 1977.

[4] Chi Song Wong, *Characterization of products of symmetric matrices*, Linear Algebra Appl., 42 (1982), pp. 243–251.

[5] O. Taussky. *The role of symmetric matrices in the study of general matrices*, Linear Algebra Appl., 5 (1972), pp. 147–154.