Congruences of a square matrix and its transpose

Roger A. Horn
Department of Mathematics, University of Utah
Salt Lake City, Utah 84103, rhorn@math.utah.edu

Vladimir V. Sergeichuk*
Institute of Mathematics, Tereshchenkivska 3
Kiev, Ukraine, sergeich@imath.kiev.ua

Abstract

It is known that any square matrix A over any field is congruent to its transpose: $A^{\mathrm{T}} = S^{\mathrm{T}}AS$ for some nonsingular S; moreover, S can be chosen such that $S^2 = I$, that is, S can be chosen to be involutory. We show that A and A^{T} are *congruent over any field $\mathbb{F}$ of characteristic not two with involution $a \mapsto \bar{a}$ (the involution can be the identity): $A^{\mathrm{T}} = \bar{S}^{\mathrm{T}}AS$ for some nonsingular S; moreover, S can be chosen such that $\bar{S}S = I$, that is, S can be chosen to be coninvolutory. The short and simple proof is based on Sergeichuk's canonical form for *congruence [*Math. USSR, Izvestiya* 31 (3) (1988) 481–501]. It follows that any matrix A over $\mathbb{F}$ can be represented as $A = EB$, in which E is coninvolutory and B is symmetric.

AMS classification: 15A21

Keywords: Congruence; Sesquilinear forms; Canonical forms

This is the authors' version of a work that was published in Linear Algebra Appl. 389 (2004) 347–353.

*The research was done while this author was visiting the University of Utah supported by NSF grant DMS-0070503.

1 Introduction

We work over a field $\mathbb{F}$ of characteristic not two with involution $a \mapsto \bar{a}$, that is, a bijection (perhaps the identity) on $\mathbb{F}$ such that

$$\overline{a+b} = \bar{a} + \bar{b}, \quad \overline{ab} = \bar{b}\bar{a}, \quad \bar{\bar{a}} = a.$$

For each matrix $A = [a_{ij}]$ over $\mathbb{F}$, we define $A^* = \overline{A}^{\mathrm{T}} = [\bar{a}_{ji}]$. If $S^*AS = B$ for some nonsingular matrix S, then A and B are said to be **congruent* (or *congruent* if the involution $a \mapsto \bar{a}$ is the identity). Except for (8), all our matrices are over $\mathbb{F}$.

In 1980, Gow used Riehm's classification of bilinear forms [4] to show that any nonsingular square matrix A over any field is congruent to its transpose: $A^{\mathrm{T}} = S^{\mathrm{T}}AS$ for some nonsingular S; moreover, Gow showed that S can be chosen such that $S^2 = I$, that is, S can be chosen to be *involutory* [3]. Independently at about the same time, Yip and Ballantine obtained the same theorem without the hypothesis of nonsingularity [8]. Apparently unaware of [3] and [8], Đoković and Ikramov (using Riehm's classification again ([4] and [5])) showed in 2002 that A and A^{T} are congruent [1].

We are interested in a broader result: Over $\mathbb{F}$, any square matrix A is *congruent to A^{T}; moreover, a matrix S that gives the *congruence can be chosen such that $\bar{S}S = I$, that is, S can be chosen to be *coninvolutory.* Since the involution on $\mathbb{F}$ can be the identity, our result includes that of [8] except for the case of a field of characteristic two.

2 A canonical form for *congruence

Our proof that A and A^{T} are *congruent over $\mathbb{F}$ is based on the classification of matrices for *congruence (up to classification of Hermitian matrices) that was obtained in [6, Theorem 3].

A matrix M is a **cosquare* if $M = A^{-*}A$ for some nonsingular A; A^{-*} denotes $(A^*)^{-1}$. If M is a *cosquare, every matrix C such that $C^{-*}C = M$ is called a **cosquare root* of M; we choose any *cosquare root and denote it by $\sqrt[*]{M}$.

For a polynomial $f(x) = a_0x^n + a_1x^{n-1} + \cdots + a_n \in \mathbb{F}[x]$ we define

$$\begin{aligned} \bar{f}(x) &= \bar{a}_0x^n + \bar{a}_1x^{n-1} + \cdots + \bar{a}_n, \quad \text{and} \\ f^{\vee}(x) &= \bar{a}_n^{-1}(1 + \bar{a}_1x + \cdots + \bar{a}_nx^n) \quad \text{if } a_0 = 1 \text{ and } a_n \neq 0. \end{aligned}$$

Every square matrix is similar to a direct sum of *Frobenius blocks*

$$F_{p^t} = \begin{bmatrix} 0 & & 0 & -c_n \\ 1 & \ddots & & \vdots \\ & \ddots & 0 & -c_2 \\ 0 & & 1 & -c_1 \end{bmatrix}, \tag{1}$$

in which $p(x)^t = x^n + c_1 x^{n-1} + \cdots + c_n$ is an integer power of a polynomial $p(x)$ that is irreducible over $\mathbb{F}$.

Lemma 1 ([6, §3]). *Let $p(x)$ be irreducible over $\mathbb{F}$ and let F_{p^t} be the $n \times n$ Frobenius block* (1).

(a) *If $\mathcal{A}$ is an $n \times n$ matrix over $\mathbb{F}$ and $\mathcal{A} = F_{p^t}^* \mathcal{A} F_{p^t}$, then $\mathcal{A}$ is a* Toeplitz matrix, *that is, $\mathcal{A} = [\alpha_{i-j}]_{i,j=1}^n$ for some scalars $\alpha_{1-n}, \dots, \alpha_{-1}, \alpha_0, \alpha_1, \dots, \alpha_{n-1}$ in $\mathbb{F}$.*

(b) *If $\mathcal{A}$ is a *cosquare root of F_{p^t}, then $\mathcal{A} = \mathcal{A}^* F_{p^t} = F_{p^t}^* \mathcal{A} F_{p^t}$, and so it is a Toeplitz matrix. Moreover, it has the special form*

$$\mathcal{A} = \begin{bmatrix} a_0 & a_1 & a_2 & \cdots & a_{n-1} \\ \bar{a}_0 & a_0 & a_1 & \cdots & a_{n-2} \\ \bar{a}_1 & \bar{a}_0 & a_0 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & a_1 \\ \bar{a}_{n-2} & \bar{a}_{n-3} & \cdots & \bar{a}_0 & a_0 \end{bmatrix}. \tag{2}$$

(c) *F_{p^t} is a *cosquare if and only if*

$$\left.\begin{array}{c} p(x) \neq x, \quad p(x) = p^{\vee}(x), \ \textit{and} \\ \textit{if the involution on } \mathbb{F} \textit{ is the identity then also } p(x) \neq x + (-1)^{n+1}. \end{array}\right\} \tag{3}$$

(d) *Suppose $p(x)$ satisfies the conditions* (3) *and let m denote the integer part of $(n-1)/2$. Then one may take $\sqrt[*]{F_{p^t}}$ to have the form* (2), *in which $a_0 = \cdots = a_{m-1} = 0$,*

$$a_m = \begin{cases} 1 & \textit{if } n \textit{ is even and } p(x) \neq x - \sqrt[n-1]{1}, \\ p(-1)^t & \textit{if } n \textit{ is odd and } p(x) \neq x + 1, \\ b - \bar{b} & \textit{for any } b \in \mathbb{F} \textit{ such that } b \neq \bar{b}, \textit{ otherwise,} \end{cases}$$

and $a_{m+1},\dots,a_{n-1}$ *are determined by the identity* $\sqrt[*]{F_{p^t}} = (\sqrt[*]{F_{p^t}})^* F_{p^t}$, *i.e.,*

$$\begin{bmatrix} \bar{a}_0 & a_0 & \cdots & a_{n-2} \\ \bar{a}_1 & \bar{a}_0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & a_0 \\ \bar{a}_{n-1} & \cdots & \bar{a}_1 & \bar{a}_0 \end{bmatrix} \cdot F_{p^t} = \begin{bmatrix} a_0 & a_1 & \cdots & a_{n-1} \\ \bar{a}_0 & a_0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & a_1 \\ \bar{a}_{n-2} & \cdots & \bar{a}_0 & a_0 \end{bmatrix}. \tag{4}$$

Suppose $F_{p(x)^t}$ is a *cosquare. Since $p(x) = p^\vee(x)$, the field

$$\mathbb{F}[\kappa] = \mathbb{F}[x]/p(x)\mathbb{F}[x], \qquad \kappa := x + p(x)\mathbb{F}[x], \tag{5}$$

possesses the involution

$$f(\kappa) \mapsto f(\kappa)^\circ := \overline{f}(\kappa^{-1}). \tag{6}$$

It was proved in [6, Lemma 7] that if $f(\kappa) \in \mathbb{F}[\kappa]$ and $f(\kappa) = f(\kappa)^\circ$, then $f(\kappa)$ is uniquely representable as $f(\kappa) = \varphi(\kappa)$, in which

$$\varphi(x) = \bar{b}_r x^{-r} + \bar{b}_{r-1} x^{-r+1} + \cdots + b_0 + \cdots + b_{r-1} x^{r-1} + b_r x^r, \tag{7}$$

r is the integer part of $(\deg p(x))/2$, $b_0, b_1, \dots, b_r \in \mathbb{F}$, $b_0 = \bar{b}_0$, and if $\deg p(x)$ is even then

$$b_r = \begin{cases} 0 & \text{if the involution } b \mapsto \bar{b} \text{ is the identity,} \\ \bar{b}_r & \text{if } b \mapsto \bar{b} \text{ is not the identity and } p(0) \neq 1, \\ -\bar{b}_r & \text{if } b \mapsto \bar{b} \text{ is not the identity and } p(0) = 1. \end{cases}$$

Theorem 2 ([6, Theorem 3])**.** *Let* $\mathbb{F}$ *be a field of characteristic not two with involution* (*the involution can be the identity*)*. Every square matrix* A *over* $\mathbb{F}$ *is* **congruent to a direct sum of matrices of the three types:*

(i) *a singular Jordan block* $J_n(0)$*;*

(ii) $\sqrt[*]{F_{p^t}}\varphi(F_{p^t})$*, in which* F_{p^t} *is the* n*-by-*n *Frobenius block* (1)*,* $p(x)$ *satisfies* (3)*, and* $\varphi(x)$ *is a nonzero function of the form* (7)*;*

(iii) $\begin{bmatrix} 0 & I_n \\ F_{p^t} & 0 \end{bmatrix}$*, in which* $p(x) \neq x$ *and* $p(x)$ *does not satisfy* (3)*.*

Any matrix of type (ii) *is a* **cosquare root of* F_{p^t} *and hence is a Toeplitz matrix. The summands are determined by* A *to the following extent:*

Type (i) *uniquely.*

Type (ii) *up to replacement of the whole group of summands*

$$\sqrt[*]{F_{p^t}}\varphi_1(F_{p^t}) \oplus \cdots \oplus \sqrt[*]{F_{p^t}}\varphi_s(F_{p^t})$$

with the same $p(x)^t$ *by*

$$\sqrt[*]{F_{p^t}}\psi_1(F_{p^t}) \oplus \cdots \oplus \sqrt[*]{F_{p^t}}\psi_s(F_{p^t})$$

in which each $\psi_i(x)$ *is a nonzero function of the form* (7) *and the Hermitian matrices*

$$\operatorname{diag}(\varphi_1(\kappa), \dots, \varphi_s(\kappa)) \qquad \textit{and} \qquad \operatorname{diag}(\psi_1(\kappa), \dots, \psi_s(\kappa)) \tag{8}$$

over the field $\mathbb{F}[\kappa]$ *defined in* (5) *with the involution* (6) *are* **congruent.*

Type (iii) *up to replacement of* F_{p^t} *by* $F_{(p^\vee)^t}$.

In a canonical form for similarity, one may choose as the direct summands (canonical blocks) any matrices that are similar to the Frobenius blocks (1). Over some fields, this freedom of choice can make it possible to achieve a pleasantly simple and convenient canonical form for *congruence. For example, if $\mathbb{F} = \mathbb{C}$ is the field of complex numbers, then the irreducible polynomials are all of the form $p(x) = x - \lambda$, and $F_{(x-\lambda)^n}$ is similar to the n-by-n Jordan block $J_n(\lambda)$ with eigenvalue λ. The conditions (3) tell us that when the involution on $\mathbb{C}$ is complex conjugation, then $F_{(x-\lambda)^n}$ is a *cosquare if and only if $|\lambda| = 1$; for the identity involution on $\mathbb{C}$, $F_{(x-\lambda)^n}$ is a cosquare if and only if $\lambda = (-1)^{n+1}$.

Define the n-by-n matrices

$$\Gamma_n = \begin{bmatrix} 0 & & & & & \iddots \\ & & & & 1 & \iddots \\ & & & -1 & -1 & \\ & & 1 & 1 & & \\ & -1 & -1 & & & \\ 1 & 1 & & & & 0 \end{bmatrix}, \quad \Delta_n = \begin{bmatrix} 0 & & & 1 \\ & & \iddots & i \\ & 1 & \iddots & \\ 1 & i & & 0 \end{bmatrix}.$$

Then $\Gamma_n^{-*}\Gamma_n$ is similar to $J_n((-1)^{n+1})$ and $\Delta_n^{-*}\Delta_n$ is similar to $J_n(1)$. Thus, for complex matrices we have the following canonical forms for congruence and for *congruence with respect to complex conjugation:

(i) Every square complex matrix is congruent to a direct sum, determined uniquely up to permutation of summands, of matrices of the form

$$J_n(0), \quad \Gamma_n, \quad \begin{bmatrix} 0 & I_n \\ J_n(\lambda) & 0 \end{bmatrix},$$

in which $\lambda \neq 0$, $\lambda \neq (-1)^{n+1}$, and λ is determined up to replacement by λ^{-1}.

(ii) Every square complex matrix is *congruent to a direct sum, determined uniquely up to permutation of summands, of matrices of the form

$$J_n(0), \quad \lambda\Gamma_n, \quad \begin{bmatrix} 0 & I_n \\ J_n(\mu) & 0 \end{bmatrix},$$

in which $|\lambda| = 1$ and $|\mu| > 1$. Alternatively, one may use the symmetric matrix Δ_n instead of Γ_n.

3 *Congruence of A and A^{T}

The problem of showing that A and A^{T} are congruent has been said to be difficult. In [1], the authors write, "In spite of its elementary character, the proof of this result is quite involved." In [2] we read that "The proofs...are rather complicated." However, the difficulty has been in the methods, not in the results. The canonical forms in Theorem 2 permit us to give a short and simple proof of a broader result.

Theorem 3. *Over any field* $\mathbb{F}$ *of characteristic not two with involution* $a \mapsto \bar{a}$ (*the involution can be the identity*), *every square matrix* A *is *congruent to its transpose. Moreover, there is a coninvolutory matrix* S *over* $\mathbb{F}$ *such that* $A^{\mathrm{T}} = S^*AS$.

Proof. Let $A_{\mathrm{can}} = S^*AS$ be a canonical form of A for *congruence, that is, a direct sum of matrices of the three types described in Theorem 2. If A_{can} is *congruent to $A_{\mathrm{can}}^{\mathrm{T}}$, then A is *congruent to A^{T} since $R^*A_{\mathrm{can}}R = A_{\mathrm{can}}^{\mathrm{T}}$ implies

$$(SR\bar{S}^{-1})^*A(SR\bar{S}^{-1}) = A^{\mathrm{T}}.$$

Hence it suffices to prove that all matrices of the three types described in Theorem 2 are *congruent to their transposes, and that R can be chosen to be coninvolutory.

Matrices of types (i) and (ii) are always *congruent to their transposes since they are Toeplitz matrices, and for any Toeplitz matrix B we have

$$\begin{bmatrix} 0 & & 1 \\ & \iddots & \\ 1 & & 0 \end{bmatrix} B \begin{bmatrix} 0 & & 1 \\ & \iddots & \\ 1 & & 0 \end{bmatrix} = B^{\mathrm{T}}.$$

Notice that the congruence is achieved via a real involutory matrix.

For each matrix of type (iii) we have

$$\begin{bmatrix} 0 & S^{-1} \\ S^* & 0 \end{bmatrix} \cdot \begin{bmatrix} 0 & I \\ F & 0 \end{bmatrix} \cdot \begin{bmatrix} 0 & S \\ S^{-*} & 0 \end{bmatrix} = \begin{bmatrix} 0 & F^{\mathrm{T}} \\ I & 0 \end{bmatrix} = \begin{bmatrix} 0 & I \\ F & 0 \end{bmatrix}^{\mathrm{T}},$$

in which S is any nonsingular matrix such that $S^{-1}FS = F^{\mathrm{T}}$. However, S always can be chosen to be symmetric [7], and if we do so then

$$\mathcal{S} = \begin{bmatrix} 0 & S \\ S^{-*} & 0 \end{bmatrix} = \begin{bmatrix} 0 & S \\ \bar{S}^{-1} & 0 \end{bmatrix}$$

and $\bar{\mathcal{S}}\mathcal{S} = I$. □

It was proved in [3, p. 329] that any nonsingular matrix over a field can be represented as $A = EB$, in which E is involutory and B is symmetric. As an immediate consequence of Theorem 3, we have the following factorization theorem.

Corollary 4. *Over any field $\mathbb{F}$ of characteristic not two with involution $a \mapsto \bar{a}$ (the involution can be the identity), any square matrix A can be represented as $A = EB$, in which E is coninvolutory and B is symmetric.*

Proof. Theorem 3 ensures that $A = EA^TE^*$ for some coninvolutory matrix E. Therefore, $\bar{E}A = \bar{E}EA^TE^* = A^T\bar{E}^T = (\bar{E}A)^T \equiv B$ is symmetric and $A = EB$. □

References

[1] D. Ž. Đocović and K. D. Ikramov, A square matrix is congruent to its transpose, *J. Algebra* 257 (2002) 97–105.

[2] D. Ž. Đocović, F. Szechtman, and K. Zhao, An algorithm that carries a square matrix into its transpose by an involutory congruence transformation, *Electron. J. Linear Algebra* 10 (2003) 320-340.

[3] R. Gow, The equivalence of an invertible matrix to its transpose, *Linear and Multilinear Algebra* 8 (1980) 329-336.

[4] C. Riehm, The equivalence of bilinear forms, *J. Algebra* 31 (1974) 45-66.

[5] C. Riehm and M. Shrader-Frechette, The equivalence of sesquilinear forms, *J. Algebra* 42 (1976) 495-530.

[6] V. V. Sergeichuk, Classification problems for system of forms and linear mappings, *Math. USSR, Izvestiya* 31 (3) (1988) 481–501.

[7] O. Taussky and H. Zassenhaus, On the similarity transformation between a matrix and its transpose, *Pacific J. Math.* 9 (1959) 893-896.

[8] E. L. Yip and C. S. Ballantine, Congruence and conjunctivity of matrices to their adjoints, *Linear Algebra Appl.* 41 (1981) 33-72.