Available online at www.sciencedirect.com

SCIENCE DIRECT

Journal of Algebra 257 (2002) 97–105

JOURNAL OF Algebra

www.academicpress.com

A square matrix is congruent to its transpose

Dragomir Ž. Đoković [a,*,1] and Khakim D. Ikramov [b,2]

[a] *Department of Pure Mathematics, University of Waterloo, Waterloo, ON, Canada N2L 3G1*
[b] *Faculty of Computational Mathematics and Cybernetics, Moscow State University, Moscow 119899, Russia*

Received 12 September 2001

Communicated by Efim Zelmanov

Abstract

For any matrix X let X' denote its transpose. We show that if A is an n by n matrix over a field K, then A and A' are congruent over K, i.e., $P'AP = A'$ for some $P \in \mathrm{GL}_n(K)$.

Keywords: Congruence of matrices; Bilinear forms; Sesquilinear forms; Kronecker modules

1. Introduction

For any matrix X let X' denote its transpose. Let us start by recalling the following known fact (see, e.g., [2, Theorem 4, p. 205] or [3, Theorem 11]): if A is a complex n by n matrix, then A and A' are congruent, i.e., $P'AP = A'$ for some invertible complex matrix P. Our main objective is to prove that the same assertion is valid for matrices over an arbitrary field K. In spite of its elementary character, the proof of this result is quite involved.

* Corresponding author.
E-mail addresses: djokovic@uwaterloo.ca (D.Ž. Đoković), ikramov@cs.msu.su (K.D. Ikramov).
1 The author was supported in part by the NSERC Grant A-5285.
2 The author was supported in part by the NSERC Grant OGP 000811.

0021-8693/02/$ – see front matter
PII: S0021-8693(02)00126-6

Theorem 1.1. *If A is an n by n matrix over a field K, then $P'AP = A'$ for some $P \in \mathrm{GL}_n(K)$.*

The following result is an immediate consequence.

Corollary 1.2. *For $A \in \mathrm{GL}_n(K)$, the matrices A and A^{-1} are congruent over K.*

Proof. Indeed, if $P \in \mathrm{GL}_n(K)$ is such that $P'AP = A'$, then $Q'AQ = A^{-1}$ for $Q = PA^{-1}$. □

The proof of Theorem 1.1 is based on the paper of Riehm [3] and an addendum to it by Gabriel [1]. This paper solves the equivalence problem for bilinear forms on finite-dimensional vector spaces over a field. For technical reasons we prefer to use the subsequent paper of Riehm and Shrader-Frechette [4], which gives a solution of the equivalence problem for sesquilinear forms on finitely generated modules over semisimple (Artinian) rings. We need only apply this general theory to bilinear forms over K. Let us reformulate the above theorem in the language of bilinear forms.

If $f : V \times V \to K$ is a bilinear form on a finite-dimensional K-vector space V, we shall say that (V, f) is a *bilinear space.* The definition of equivalence of two bilinear forms is the usual one.

Definition 1.3. Two bilinear forms $f : V \times V \to K$ and $g : W \times W \to K$ are equivalent if there exists a vector space isomorphism $\varphi : V \to W$ such that $g(\varphi(x), \varphi(y)) = f(x, y)$, $\forall x, y \in V$. In that case, assuming that V and W are finite-dimensional, we also say that the bilinear spaces (V, f) and (W, g) are isometric and that φ is an isometry.

Let us define the transpose of a bilinear form.

Definition 1.4. The transpose of a bilinear form $f : V \times V \to K$ is the bilinear form $g : V \times V \to K$ such that $g(x, y) = f(y, x)$ for all $x, y \in V$. We shall denote the transpose of f by f'.

Let f and g be as in Definition 1.3 and assume that $\dim(V) = \dim(W) = n < \infty$. We fix a basis $\{v_1, v_2, \ldots, v_n\}$ of V. Then the n by n matrix $A = (a_{ij})$ where $a_{ij} = f(v_i, v_j)$ is *the matrix of* f with respect to this basis. The matrix of f', with respect to the same basis, is A'. Similarly, let $B = (b_{ij})$ be the matrix of g with respect to a fixed basis $\{w_1, w_2, \ldots, w_n\}$ of W. To say that f and g are equivalent is the same as to say that $P'AP = B$ for some $P \in \mathrm{GL}_n(K)$.

Theorem 1.5. *Let V be a finite-dimensional K-vector space, $f : V \times V \to K$ a bilinear form on V, and f' its transposed form. Then f and f' are equivalent.*

The orthogonality of subspaces of a bilinear space is defined as follows.

Definition 1.6. Let (V, f) be a bilinear space. We say that the subspaces U and W of V are orthogonal to each other if $f(U, W) = 0$ and $f(W, U) = 0$.

Answering the question of a referee, we point out that Theorem 1.1 does not generalize to "hermitian conjugacy," i.e., if $\sigma \in \mathrm{Aut}(K)$ with $\sigma^2 = 1$ and we set $A^* = (A')^\sigma$, then A and A^* in general are not $*$-congruent. (Two n by n matrices A and B over K are said to be $*$-congruent if $B = P^*AP$ for some $P \in \mathrm{GL}_n(K)$.) A simple counter example is provided by $(K, \sigma) =$ (the complex field, the complex conjugation) and the 1 by 1 matrix $A = [i]$, where i is the imaginary unit.

2. Kronecker modules

In this section we recall some facts about the Kronecker modules which are special cases of the general Kronecker modules discussed in [4]. The reader should consult this reference and [1] for more details.

We define a *Kronecker module* as a four-tuple (X, u, v, Y) where X and Y are finite-dimensional K-vector spaces and $u, v : X \to Y$ are linear maps. To a bilinear space (Z, h) we assign the Kronecker module $K(Z, h) = K(Z) = (Z, h_l, h_r, Z^*)$, where Z^* is the dual space of Z and $h_l, h_r : Z \to Z^*$ are defined by

$$h_l(x)(y) = h(x, y), \qquad h_r(x)(y) = h(y, x), \quad \forall x, y \in Z.$$

Every Kronecker module is a direct sum of indecomposable ones which are unique up to ordering and isomorphism. There are five types of indecomposable Kronecker modules (X, u, v, Y):

I. Both u and v are isomorphisms and $u^{-1}v$ is indecomposable (i.e., it has only one elementary divisor).

II. The spaces X and Y have the same dimension and the pencil $\lambda u + \mu v$, with respect to suitable bases of X and Y, has the matrix

$$\begin{bmatrix} \lambda & & & & \\ \mu & \lambda & & & \\ & \mu & \ddots & & \\ & & \ddots & \lambda & \\ & & & \mu & \lambda \end{bmatrix}.$$

II*. Similar to II with the matrix

$$\begin{bmatrix} \mu & \lambda & & & \\ & \mu & \lambda & & \\ & & \mu & \ddots & \\ & & & \ddots & \lambda \\ & & & & \mu \end{bmatrix}.$$

III. In this case $\dim(Y) = \dim(X) + 1$ and the pencil $\lambda u + \mu v$ has the matrix

$$\begin{bmatrix} \lambda & & & \\ \mu & \lambda & & \\ & \mu & \ddots & \\ & & \ddots & \lambda \\ & & & \mu \end{bmatrix}.$$

III*. In this case $\dim(X) = \dim(Y) + 1$ and the matrix is

$$\begin{bmatrix} \mu & \lambda & & & \\ & \mu & \ddots & & \\ & & \ddots & \lambda & \\ & & & \mu & \lambda \end{bmatrix}.$$

The following theorem is an immediate consequence of the theory of Kronecker modules (also known as the theory of matrix pencils).

Theorem 2.1. *If (V, f) is a bilinear space, then the Kronecker modules $K(V, f)$ and $K(V, f')$ are isomorphic.*

In terms of matrices, this can be restated as follows.

Theorem 2.2. *If A is an n by n matrix over a field K, then the matrix pencils $\lambda A + \mu A'$ and $\lambda A' + \mu A$ are equivalent.*

If (Z, h) is a bilinear space and $Z = Z_1 + Z_2$ is a direct decomposition of Z such that $h(Z_1, Z_2) = 0$ and $h(Z_2, Z_1) = 0$, then we say that this space is the orthogonal direct sum of the bilinear spaces (Z_1, h_1) and (Z_2, h_2), where h_1 and h_2 are the corresponding restrictions of h. The following theorem is a very special case of the general result stated in [4, Section 9].

Theorem 2.3. *Every bilinear space (Z, h) can be decomposed into an orthogonal direct sum*

$$Z = Z_{\mathrm{I}} + Z_{\mathrm{II}} + Z_{\mathrm{III}}$$

such that all indecomposable direct summands of $K(Z_{\mathrm{I}})$ *are of type* I*, those of* $K(Z_{\mathrm{II}})$ *are of type* II *or* II**, and those of* $K(Z_{\mathrm{III}})$ *are of type* III *or* III**. If* h_{I} *is the restriction of* h *to* $Z_{\mathrm{I}} \times Z_{\mathrm{I}}$*, etc., the bilinear spaces* $(Z_{\mathrm{I}}, h_{\mathrm{I}})$*,* $(Z_{\mathrm{II}}, h_{\mathrm{II}})$*,* $(Z_{\mathrm{III}}, h_{\mathrm{III}})$ *are uniquely determined by* (Z, h) *up to isometry.*

Moreover, if $Z = Z_{\mathrm{II}}$ *or* $Z = Z_{\mathrm{III}}$*, then* $K(Z)$ *determines* (Z, h) *up to isometry.*

3. Reduction to the nondegenerate case

We now begin the proof of Theorem 1.5. We warn the reader that the notation and the definitions of the invariants of bilinear spaces in [3] and [4] do not agree. The second paper is more general and we shall exclusively use the definitions given there.

By Theorem 2.3, there is an orthogonal direct decomposition $V = V_{\mathrm{I}} + V_{\mathrm{II}} + V_{\mathrm{III}}$ where the summands V_{I}, V_{II}, and V_{III} have the properties stated there.

Let f_{I} be the restriction of f to $V_{\mathrm{I}} \times V_{\mathrm{I}}$, etc., and f'_{I} the restriction of f' to $V_{\mathrm{I}} \times V_{\mathrm{I}}$, etc. Clearly f'_{I} is the transpose of f_{I}, etc. We claim that the bilinear spaces $(V_{\mathrm{II}}, f_{\mathrm{II}})$ and $(V_{\mathrm{II}}, f'_{\mathrm{II}})$ are isometric, and so are $(V_{\mathrm{III}}, f_{\mathrm{III}})$ and $(V_{\mathrm{III}}, f'_{\mathrm{III}})$. The Kronecker module $K(V_{\mathrm{II}}, f_{\mathrm{II}})$ is a direct sum of indecomposable summands of type II or II*. By Theorem 2.2, $K(V_{\mathrm{II}}, f_{\mathrm{II}}) \cong K(V_{\mathrm{II}}, f'_{\mathrm{II}})$ and the last assertion of Theorem 2.3 implies that $(V_{\mathrm{II}}, f_{\mathrm{II}})$ and $(V_{\mathrm{II}}, f'_{\mathrm{II}})$ are isometric. The same argument shows that also $(V_{\mathrm{III}}, f_{\mathrm{III}})$ and $(V_{\mathrm{III}}, f'_{\mathrm{III}})$ are isometric, and so our claim is true.

It remains to show that the bilinear spaces $(V_{\mathrm{I}}, f_{\mathrm{I}})$ and $(V_{\mathrm{I}}, f'_{\mathrm{I}})$ are isometric. As f_{I} is nondegenerate, the proof of our theorem has been reduced to the nondegenerate case, i.e., the case where A is a nonsingular matrix.

4. Reduction to the primary case

We assume in this section that f is nondegenerate. We shall use a number of results of [4] without explicit reference and the reader should consult this paper for the claims made but not proved here.

We recall from [4] that the *asymmetry* of f is the invertible linear operator $\alpha : V \to V$ such that $f(x, y) = f(\alpha(y), x)$, $\forall x, y \in V$. Its matrix, with respect to our fixed basis of V, is $(A')^{-1}A$. The asymmetry of f' is $\alpha' = \alpha^{-1}$ and its matrix is $A^{-1}A'$. As any matrix is similar to its transpose, the asymmetries α and α' are similar operators.

Let $p \in K[X]$ be a monic irreducible polynomial and assume that $p \neq X$. For such p we define the monic irreducible polynomial $p^* \in K[X]$ by $p^*(X) =$

$p(0)^{-1}X^d p(X^{-1})$, where d is the degree of p. Let us decompose V into primary components with respect to α

$$V = \bigoplus_p V_p,$$

where the sum is over the monic irreducible polynomials $p \in K[X]$, $p \neq X$. The subspaces V_p and V_q are orthogonal if $q \neq p^*$. If $p^* \neq p$ then the assertion of our theorem is true for the restriction of f to $V_p + V_{p^*}$ by [4, Theorem 16, Corollary]. It remains to deal with the case $p^* = p$.

Hence the proof of Theorem 1.4 has been reduced to the primary case: The minimal polynomial of α is a power of p, where $p = p^*$ is an irreducible polynomial in $K[X]$.

5. Reduction to the homogeneous primary case

In this section we consider the primary case as just described above. By [4, Proposition 25], there exists an orthogonal direct decomposition

$$V = \bigoplus_{s \geqslant 1} V_s$$

such that $V_s \subseteq \ker(p(\alpha)^s)$ and the induced map

$$V_s/p(\alpha)(V_s) \to \ker\bigl(p(\alpha)^s\bigr)/\bigl(\ker\bigl(p(\alpha)^{s-1}\bigr) + p(\alpha)\ker\bigl(p(\alpha)^{s+1}\bigr)\bigr)$$

is an isomorphism for each s. Hence, without any loss of generality, we may assume that $V = V_s$ for some s. In other words, we may assume that α has only one elementary divisor with arbitrary multiplicity. We refer to this case as the *homogeneous primary case*.

6. The homogeneous primary case

The minimal polynomial of α is a power of p, say p^s, where p is as in the previous section, and all elementary divisors of α are equal to p^s. Let r be the number of these elementary divisors. In order to prove that f and f' are equivalent, it suffices to check that they have the same invariants attached to them by [4, Theorems 27 and 31]. That is exactly what we are going to show.

Assume first that $p = X - 1$. Set $\pi = 1 - \alpha^{-1}$ and $\pi' = 1 - (\alpha')^{-1} = 1 - \alpha$. Define $\widetilde{V} = V/\pi(V)$ and note that $\pi'(V) = \pi(V)$. If s is even, then the bilinear invariant attached to f (see [4, p. 517]) is a nondegenerate skew-symmetric form on the r-dimensional K-vector space $\widetilde{V}$. (Hence if s is even then r must be even.) Therefore this invariant is unique up to equivalence.

We now assume that s is odd, in which case the bilinear invariant is the nondegenerate symmetric form $\tilde{f}$ on $\widetilde{V}$ defined by

$$\tilde{f}(\tilde{x}, \tilde{y}) = f\left(\pi^{s-1}(x), y\right), \quad \forall x, y \in V,$$

where $\tilde{x}$ denotes the canonical image of x in $\widetilde{V}$. The analogous invariant $\tilde{f}'$ of the bilinear form f' is defined similarly (using π' instead of π). Since $\pi^{s-1}(V)$ is the eigenspace of α for the eigenvalue 1 and $\pi' = -\alpha\pi$, we obtain that

$$\begin{aligned} \tilde{f}'(\tilde{x}, \tilde{y}) &= f'\left((\pi')^{s-1}(x), y\right) = f'\left((\alpha\pi)^{s-1}x, y\right) = f\left(y, (\alpha\pi)^{s-1}x\right) \\ &= f\left(\alpha(\alpha\pi)^{s-1}x, y\right) = f\left(\alpha^s\pi^{s-1}x, y\right) = f\left(\pi^{s-1}x, y\right) = \tilde{f}(\tilde{x}, \tilde{y}) \end{aligned}$$

for all $x, y \in V$. Hence $\tilde{f} = \tilde{f}'$.

If the characteristic of K is 2 then there are additional invariants: The quadratic forms F_i, $i \geqslant 0$. It is immediate from the definition of these forms (see [4, Section 8]) that these invariants are the same for f and f'.

In the case $p = X + 1$ (we may assume that the characteristic of K is not 2) the proof is similar.

It remains to consider the case where p has degree $d > 1$. As $p = p^*$, it follows that d is even (see [3]) and $p(0)^2 = 1$. We set $\pi = \alpha^{-d}p(\alpha)$, $\pi' = \alpha^d p(\alpha^{-1})$ and $\widetilde{V} = V/\pi(V) = V/\pi'(V)$. The algebra $K[\alpha]$, which is isomorphic to the quotient ring $K[X]/(p^s)$, has an involution J such that $\alpha^J = \alpha^{-1}$. The corresponding involution, also denoted by J, of $K[X]/(p^s)$ sends the element $\zeta = X + (p^s)$ to its inverse.

The algebra $K[X]/(p)$ is a finite field extension of K. It also has an involution, J, which sends the element $\xi = X + (p)$ to its inverse. The space V is naturally a module over the algebra $K[X]/(p^s)$ with ζ acting as α. Similarly, $\widetilde{V}$ is naturally a module over the algebra $K[\xi] = K[X]/(p)$ in two different ways: First we let ξ act as $\tilde{\alpha}$ (the linear transformation induced by α), and second we let ξ act as $\tilde{\alpha}' = (\tilde{\alpha})^{-1}$. We shall distinguish these two actions by writing $\xi \circ \tilde{x} = \tilde{\alpha}(\tilde{x})$ for the former and $\xi * \tilde{x} = (\tilde{\alpha})^{-1}(\tilde{x})$ for the latter.

Recall that a J-*sesquilinear* form h on a $K[\xi]$-vector space M is a K-bilinear map $h : M \times M \to K[\xi]$ such that $h(ax, by) = a^J bh(x, y)$ for all $a, b \in K[\xi]$ and $x, y \in M$. Let $\mu \in K[\xi]$ satisfy $\mu\mu^J = 1$. A J-sesquilinear form h on M is called μ-*hermitian* if $h(y, x) = \mu h(x, y)^J$ for all vectors $x, y \in M$. A J-sesquilinear form h on M is called *hermitian* if it is μ-Hermitian for $\mu = 1$.

From now on we set $\mu = p(0)^{s-1}\xi^{(s-1)d+1}$. As in [4, p. 512], let ν be 1 if the characteristic of K is 0, and otherwise let ν be the greatest power of the characteristic such that p is a polynomial in $X_1 = X^\nu$. Then $\{1, \xi, \ldots, \xi^{\nu-1}\}$ is a basis of $K[\xi]$ as a vector space over its subfiled $K[\xi_1]$, where $\xi_1 = \xi^\nu$. Define the $K[\xi_1]$-linear functional $\tau_1 : K[\xi] \to K[\xi_1]$ by

$$\tau_1\left(\sum_{i=0}^{\nu-1} a_i\xi^i\right) = a_0, \quad a_i \in K[\xi_1].$$

We now define the K-linear map $\tau : K[\xi] \to K$ by $\tau = \mathrm{Tr} \circ \tau_1$ where Tr is the trace map $K[\xi_1] \to K$ of the separable field extension $K[\xi_1]$ of K.

Apart from the asymmetry α, the bilinear form f has only one invariant (see [4]): The unique nondegenerate μ-Hermitian form $\tilde{f}$ on the $K[\xi]$-vector space $(\widetilde{V}, \circ)$ such that

$$\tau \tilde{f}(\tilde{x}, \tilde{y}) = f\big(\pi^{s-1}(x), y\big), \quad \forall x, y \in V.$$

Similarly, the analogous invariant of the bilinear form f' is the unique nondegenerate μ-Hermitian form $\tilde{f}'$ on the $K[\xi]$-vector space $(\widetilde{V}, *)$ such that

$$\tau \tilde{f}'(\tilde{x}, \tilde{y}) = f'\big((\pi')^{s-1}(x), y\big), \quad \forall x, y \in V.$$

In order to complete the proof of the theorem, it suffices to show that the μ-Hermitian forms $\tilde{f}$ and $\tilde{f}'$ are equivalent, i.e., that there exists an isomorphism $\varphi : (\widetilde{V}, \circ) \to (\widetilde{V}, *)$ of $K[\xi]$-vector spaces such that $\tilde{f}'(\varphi(\tilde{x}), \varphi(\tilde{y})) = \tilde{f}(\tilde{x}, \tilde{y})$ for all $x, y \in V$.

Recall that $p(0)^2 = 1$. It is easy to check that $\tau(a^J) = \tau(a)$ for all $a \in K[\xi]$. Since $\pi' = p(0)\alpha^d \pi$ and $f(x, y) = f(\alpha(y), x)$ for arbitrary $x, y \in V$, we obtain that

$$\begin{aligned}
\tau \tilde{f}'(\tilde{x}, \tilde{y}) &= f'\big((\pi')^{s-1}(x), y\big) = f'\big(\big(p(0)\alpha^d \pi\big)^{s-1}(x), y\big) \\
&= f\big(y, \big(p(0)\alpha^d \pi\big)^{s-1}(x)\big) = f\big(p(0)^{s-1}\alpha^{1+d(s-1)}\pi^{s-1}(x), y\big) \\
&= \tau \tilde{f}\big(p(0)^{s-1}\xi^{1+d(s-1)} \circ \tilde{x}, \tilde{y}\big) = \tau\big(p(0)^{s-1}\xi^{d(1-s)-1}\tilde{f}(\tilde{x}, \tilde{y})\big) \\
&= \tau\big(p(0)^{2(s-1)}\tilde{f}(\tilde{y}, \tilde{x})^J\big) = \tau \tilde{f}(\tilde{y}, \tilde{x}).
\end{aligned}$$

As both $(\tilde{x}, \tilde{y}) \to \tilde{f}'(\tilde{x}, \tilde{y})$ and $(\tilde{x}, \tilde{y}) \to \tilde{f}(\tilde{y}, \tilde{x})$ are μ-Hermitian forms on the $K[\xi]$-vector space $(\widetilde{V}, *)$, the above equality and [4, Theorem 22] imply that

$$\tilde{f}'(\tilde{x}, \tilde{y}) = \tilde{f}(\tilde{y}, \tilde{x}), \quad \forall x, y \in X. \tag{6.1}$$

One should keep in mind that this identity is possible only because $\tilde{f}$ and $\tilde{f}'$ are μ-Hermitian forms for two different $K[\xi]$-vector space structures on the K-vector space $\widetilde{V}$. The identity map from $(\widetilde{V}, \circ)$ to $(\widetilde{V}, *)$ is a J-linear isomorphism (not $K[\xi]$-linear).

We remark that every basis of $(\widetilde{V}, \circ)$ is also a basis of $(\widetilde{V}, *)$. By [5, Theorem 6.3, p. 259], we can choose vectors $x_1, \ldots, x_r \in V$ such that $\{\tilde{x}_1, \ldots, \tilde{x}_r\}$ is an orthogonal basis of $(\widetilde{V}, \circ)$ with respect to the form $\tilde{f}$. By (6.1), this basis is also an orthogonal basis of $(\widetilde{V}, *)$ with respect to the form $\tilde{f}'$. Moreover, (6.1) entails that the μ-Hermitian forms $\tilde{f}$ and $\tilde{f}'$ have the same matrix with respect to the above basis. Hence these two forms are equivalent and the proof of Theorem 1.4 is completed.

7. An application to real orthogonal groups

Let $\mathrm{O}(p,q)$, $p+q=n$, be the subgroup of $\mathrm{GL}_n(\mathbf{R})$ consisting of all matrices A such that $A'J_{p,q}A=J_{p,q}$, where $J_{p,q}=\mathrm{diag}(1,\dots,1,-1,\dots,-1)$ with the first p (respectively last q) diagonal entries equal $+1$ (respectively -1). Consider the action of $\mathrm{O}(p,q)$ on the space $\mathcal{K}_n$ of all n by n skew-symmetric matrices given by $X\to AXA'$, $X\in\mathcal{K}_n$, $A\in\mathrm{O}(p,q)$. Then the following result is valid.

Proposition 7.1. *For any $X\in\mathcal{K}_n$, the matrices X and $-X$ belong to the same orbit of* $\mathrm{O}(p,q)$.

Proof. Apply Theorem 1.1 to the matrix $J_{p,q}+X$ whose transpose is $J_{p,q}-X$. □

References

[1] P. Gabriel, Appendix: Degenerate bilinear forms, J. Algebra 31 (1974) 67–72.
[2] A.I. Mal'tsev, Foundations of Linear Algebra, Freeman, San Francisco, 1963.
[3] C. Riehm, The equivalence of bilinear forms, J. Algebra 31 (1974) 45–66.
[4] C. Riehm, M.A. Shrader-Frechette, The equivalence of sesquilinear forms, J. Algebra 42 (1976) 495–530.
[5] W. Scharlau, Quadratic and Hermitian Forms, Springer, New York, 1985.