A Transfer Device for Matrix Theorems

WILLIAM P. WARDLAW
U. S. Naval Academy
Annapolis, MD 21402

Our title refers to a method for obtaining a number of results for matrices over arbitrary commutative rings by "transferring" the corresponding results for matrices over the real numbers. The technique was suggested by a proof [5] in a calculus text which showed $\det(AB) = (\det A)(\det B)$ for A and B nonsingular, and then extended the result to singular A or B by continuity. More or less, the technique described in this note is an algebraic substitute for the use of continuity which can serve as a rigorous replacement for waving the hands and stating "For commutative rings, everything goes through as for fields." The existence of the transfer device obviates the need to do undergraduate linear algebra over commutative rings and suggests that a restriction to the field $\mathbf{R}$ of real numbers (or perhaps the field $\mathbf{C}$ of complex numbers) will suffice, since many results can be "transferred" to more general settings in a graduate course.

Throughout this note, R is an arbitrary commutative ring, $R^{m\times n}$ is the collection of all $m\times n$ matrices over R, $R_n = R^{n\times n}$, and $R[t]$ is the ring of polynomials over R. Here, $R[t]$ is considered to be the ring formally generated by t and R, containing R as the constant polynomials and all of the powers t^k for positive k, even if R does *not* have an identity 1. Finally,

$$\mathbf{M}(R,t) = \bigcup_{m,n\in\mathbf{N}} R[t]^{m\times n}$$

is the partial algebra of all elements of $R = R_1$, all polynomials in $R[t] = R[t]_1$, and all matrices with entries in $R[t]$. The two operations $+$ and $\cdot$ in $\mathbf{M}(R,t)$ are ordinary matrix addition (with $A+B$ defined when A and B are the same size) and matrix (or scalar) multiplication (with $A\cdot B$ defined when A is $m\times n$ and B is $n\times p$ *or* when either A or B is 1×1). The phrase partial algebra refers to the fact that the operations are not always defined.

Recall that the determinant of a square matrix $A = (a_{ij})$ in $\mathbf{M}(R,t)$ is

$$\det A = \sum_{\sigma\in S_n} (\operatorname{sgn}\sigma)\, a_{1\sigma_1} a_{2\sigma_2}\cdots a_{n\sigma_n}.$$

The characteristic polynomial of A in R_n is $f_A = f_A(t) = \det(tI - A)$, and the (classical) adjoint or adjugate of A is $\operatorname{Adj} A = C^T$, the transpose of the cofactor matrix $C = (c_{ij})$, where $c_{ij} = (-1)^{i+j}\det A(i|j)$ and $A(i|j)$ is the matrix resulting from A by deleting the ith row and the jth column.

The transfer device and applications

TRANSFER THEOREM. *Let R and R' be commutative rings and $\theta\colon R\to R'$ be a ring homomorphism. Then θ induces a homomorphism*

$$\phi\colon \mathbf{M}(R,t)\to\mathbf{M}(R',t)$$

satisfying:

(1) $\phi(a) = \theta(a)$ *for every* $a\in R$,

(2) $\phi\left(\sum_{i=0}^{n} a_i t^i\right) = \sum_{i=0}^{n}\theta(a_i)t^i$ *for* $a_0, a_1, \dots, a_n\in R$,

(3) $A = (a_{ij})\in R[t]^{m\times n}$ *implies* $\phi(A) = (\phi(a_{ij}))$,

(4) $\phi(A+B) = \phi(A)+\phi(B)$ *when* $A+B$ *is defined*,

(5) $\phi(A\cdot B) = \phi(A)\cdot\phi(B)$ *when* $A\cdot B$ *is defined*,

(6) $\phi(\det A) = \det(\phi A)$ *when* A *is square*,

(7) $\phi(f_A) = f_{\phi(A)}$ *when* A *is square, and*

(8) $\phi(\operatorname{Adj} A) = \operatorname{Adj}(\phi A)$ *when* A *is square.*

The proof of the Transfer Theorem is not difficult: we *define* ϕ by properties (1)–(3) and then prove properties (4)–(8). It is clear that ϕ is well defined by properties (1)–(3). The proof of properties (4) and (5), which show that ϕ is a homomorphism, is straightforward but tedious, and hence is omitted. Properties (6)–(8) then follow because $\det A$, the coefficients of f_A, and the entries in $\operatorname{Adj} A$ are all polynomials in the entries of A.

Throughout the remainder of this note applications of the Transfer Theorem will be demonstrated by using it to prove, first, some very well known theorems about determinants over commutative rings and, later, some less well known theorems.

THEOREM 1. *If A and B are square matrices over a commutative ring R, then*

$$\det(AB) = (\det A)(\det B).$$

Proof. Let $A = (a_{ij})$ and $B = (b_{ij})$ be $n \times n$ matrices over the ring R. Then let $\mathbf{A} = \langle A, B\rangle \equiv \langle a_{11}, \dots, a_{nn}, b_{11}, \dots, b_{nn}\rangle$ be the subring of R generated by the entries of A and B. Let $X = (x_{ij})$ and $Y = (y_{ij})$ be $n \times n$ matrices over the field $\mathbf{R}$ of real numbers with $2n^2$ independent transcendental entries x_{ij} and y_{ij} in $\mathbf{R}\setminus\mathbf{Q}$. Then let $K = \mathbf{Q}(X, Y) \equiv \mathbf{Q}(x_{11}, \dots, x_{nn}, y_{11}, \dots, y_{nn})$ be the $2n^2$-fold transcendental extension of $\mathbf{Q}$, and let $\mathbf{X} = \langle X, Y\rangle \equiv \langle x_{11}, \dots, x_{nn}, y_{11}, \dots, y_{nn}\rangle$ be the subring of K generated by the entries of X and Y. Since the transcendentals x_{ij} and y_{ij} are algebraically independent over $\mathbf{Q}$, they generate a free commutative ring (that is, the free algebra over the class of all commutative rings, with the free generating family $\{x_{ij}, y_{ij}\}$, as defined in [**1**]), which is actually just the set of all nonconstant polynomials in the polynomial ring $\mathbf{Z}[X, Y] \equiv \mathbf{Z}[x_{11}, \dots, x_{nn}, y_{11}, \dots, y_{nn}]$. Hence, there is a homomorphism $\theta\colon \mathbf{X} \to \mathbf{A}$ such that $\theta(x_{ij}) = a_{ij}$, $\theta(y_{ij}) = b_{ij}$ for each i and j. Defining ϕ as in the Transfer Theorem and using the fact that $\det(XY) = (\det X)(\det Y)$ for the matrices X and Y over $\mathbf{R}$, we obtain

$$\begin{aligned}\det(AB) &= \det(\phi X \cdot \phi Y) = \det \phi(XY) = \phi \det(XY) = \phi((\det X)(\det Y))\\ &= (\phi(\det X))(\phi(\det Y)) = (\det(\phi X))(\det(\phi Y)) = (\det A)(\det B)\end{aligned}$$

after several applications of the Transfer Theorem.

The techniques of the above proof will be repeated with minor modifications to prove the theorems which follow. To save space and to relieve tedium, many of the details given above will be omitted.

THEOREM 2 (Cayley-Hamilton). *$A \in R_n$ implies $f_A(A) = 0$.*

Proof. For any $A \in R_n$, let $\mathbf{A} = \langle A\rangle$ be the subring of R generated by the entries a_{ij} of A and let $\mathbf{X} = \langle X\rangle$ be the free subring of $K = \mathbf{Q}(X)$ generated by the n^2 transcendental entries x_{ij} of X. Let ϕ be the canonical homomorphism from $\mathbf{M}(\mathbf{X}, t)$ to $\mathbf{M}(\mathbf{A}, t)$ given by the Transfer Theorem satisfying $\phi(x_{ij}) = a_{ij}$ for all i and j. Then

$$f_A(A) = f_{\phi X}(\phi X) = (\phi f_X)(\phi X) = \phi(f_X(X)) = \phi(0) = 0$$

follows from the Transfer Theorem and the Cayley-Hamilton Theorem $f_X(X) = 0$ for matrices over $\mathbf{R}$.

THEOREM 3. *If A is a square matrix over the commutative ring R, then*

$$A(\operatorname{Adj} A) = (\det A) I = (\operatorname{Adj} A) A.$$

Proof. Let A, X, and ϕ be as in the proof of Theorem 2. Then

$$X(\operatorname{Adj} X) = (\det X) I = (\operatorname{Adj} X) X$$

holds for the matrix X over $\mathbf{R}$, and so the images under ϕ of the above three expressions must also be equal, i.e.,

$$A(\operatorname{Adj} A) = (\det A) I = (\operatorname{Adj} A) A.$$

So far, we have used the Transfer Theorem only to "transfer" a theorem that is well known for matrices over the real numbers to obtain the corresponding theorem for matrices over an arbitrary commutative ring R. The next three theorems are interesting not only because they are less well known than the preceding three, but also because even their proof for arbitrary matrices over the field $\mathbf{R}$ of real numbers makes use of the transfer theorem. The first two of these will be proved by obtaining the result for invertible matrices over $\mathbf{R}$ and then applying the transfer theorem to obtain the corresponding results for any matrices over an arbitrary commutative ring R. In particular, this establishes the results for singular matrices over the real numbers. (Note that this approach could also have been taken for Theorem 1; indeed, such a proof would be the algebraic equivalent of the continuity proof **[5]** that motivated this paper.)

THEOREM 4. *Let A and B be $n \times n$ matrices over a commutative ring R. Then*

$$\operatorname{Adj}(AB) = (\operatorname{Adj} B)(\operatorname{Adj} A).$$

Proof. Case 1. Assume A and B are invertible over $\mathbf{R}$. Then

$$\begin{aligned}\operatorname{Adj}(AB) &= \det(AB)\cdot(AB)^{-1} = (\det A)(\det B)\,B^{-1}A^{-1}\\ &= (\det B)\,B^{-1}\cdot(\det A)\,A^{-1} = (\operatorname{Adj} B)(\operatorname{Adj} A).\end{aligned}$$

Case 2. $A, B \in R_n$. Let X, Y, and ϕ be chosen as in the proof of Theorem 1. Since the elements x_{ij} of X are algebraically independent, $\det X$ can be considered as a polynomial in the x_{ij} with rational coefficients, and X is singular if and only if this polynomial is identically 0. However, the substitution $x_{ij} = \delta_{ij}$ (where δ_{ij} is the Kronecker delta defined by $\delta_{ii} = 1$ and $\delta_{ij} = 0$ if $i \neq j$) gives $X = I$ and $\det X = 1$, so the polynomial $\det X$ cannot be identically 0. Thus X is an invertible matrix over $\mathbf{R}$, and so is Y. Hence $\operatorname{Adj}(XY) = (\operatorname{Adj} Y)(\operatorname{Adj} X)$ by Case 1, and the Transfer Theorem gives

$$\operatorname{Adj}(AB) = \phi(\operatorname{Adj}(XY)) = \phi((\operatorname{Adj} Y)(\operatorname{Adj} X)) = (\operatorname{Adj} B)(\operatorname{Adj} A).$$

THEOREM 5. *If $A \in R_n$, then $\operatorname{Adj} A = p_A(A)$ is a polynomial p_A evaluated at A, where*

$$p_A(t) = (-1)^{n+1}[f_A(t) - f_A(0)]/t.$$

Proof. Case 1. Assume A is invertible over $\mathbf{R}$. Then

$$Ap_A(A) = (-1)^{n+1}[f_A(A) - f_A(0)\cdot I] = (-1)^n f_A(0)\cdot I = (\det A)\cdot I = A(\operatorname{Adj} A)$$

implies that

$$p_A(A) = \operatorname{Adj} A$$

upon left multiplication by A^{-1}.

Case 2. Let A, X, and ϕ be as in the proof of Theorem 2. Then X is invertible over $\mathbf{R}$, so $p_X(X) = \operatorname{Adj} X$. Recalling the definition of p_A, it is clear from the Transfer Theorem that the image under ϕ is $p_A(A) = \operatorname{Adj} A$.

The last two theorems were first proved for invertible matrices over the field $\mathbf{R}$ of real numbers and then "transferred" to matrices over arbitrary commutative rings. To carry out the transfers we needed to know that the "transcendental matrices" X and Y are invertible. To prove our last theorem, we will need a more subtle property of the matrices X and Y, namely, that their product XY has distinct eigenvalues in the field $\mathbf{C}$ of complex numbers.

THEOREM 6. *Let R be a commutative ring, let A be an $m \times n$ matrix over R, and let B be an $n \times m$ matrix over R, where $m \leqslant n$. Then*

$$f_{BA}(t) = t^{n-m} f_{AB}(t).$$

Proof. Case 1. Let A and B be $m \times n$ and $n \times m$ matrices, respectively, over the field $\mathbf{R}$ of real numbers such that AB has distinct nonzero eigenvalues $\lambda_1, \ldots, \lambda_m$ in the field $\mathbf{C}$ of complex numbers. If λ is a nonzero eigenvalue of AB with eigenvector $\mathbf{v}$, then $AB\mathbf{v} = \lambda\mathbf{v}$ implies $BA(B\mathbf{v}) = \lambda(B\mathbf{v})$ and λ is an eigenvalue of BA, also. Thus, $\lambda_1, \ldots, \lambda_m$ are distinct nonzero eigenvalues of BA. Hence the $n \times n$ matrix BA has rank m and nullity $n - m$, so 0 is an $(n-m)$-fold eigenvalue of BA. Therefore,

$$f_{BA}(t) = t^{n-m}(t - \lambda_1) \cdots (t - \lambda_m) = t^{n-m} f_{AB}(t).$$

Case 2. Let A and B be $m \times n$ and $n \times m$ matrices, respectively, over the commutative ring R, and let $X = (x_{ij})$ and $Y = (y_{ij})$ be $m \times n$ and $n \times m$ matrices, respectively, with $2mn$ algebraically independent entries x_{ij}, y_{ij} in $\mathbf{R} \setminus \mathbf{Q}$. The $m \times m$ matrix XY is invertible over the reals, since its determinant is nonzero. This can be seen by considering $\det(XY)$ as a polynomial in x_{ij}, y_{ij}. The substitutions $x_{ij} = \delta_{ij}$ and $y_{ij} = \delta_{ij}$ give $\det(XY) = \det I = 1$, showing that $\det(XY)$ cannot be 0.

Moreover, XY has m distinct eigenvalues in $\mathbf{C}$. This can be seen as follows: The discriminant $D(f_{XY})$ of f_{XY} is a polynomial in the coefficients of f_{XY}, which are in turn polynomials in the entries x_{ij} and y_{ij} of X and Y. (See [**2**] or [**6**] for a description of the discriminant of a polynomial and its properties.) The matrix XY has a repeated eigenvalue if and only if the discriminant $D(f_{XY})$ is zero. However, considering $D(f_{XY})$ as a polynomial in x_{ij}, y_{ij}, the substitutions $x_{ij} = \delta_{ij}$ and $y_{ij} = i \cdot \delta_{ij}$ give $XY = \operatorname{diag}(1, 2, \ldots, m)$ with m distinct eigenvalues, showing that the polynomial $D(f_{XY})$ cannot be identically 0.

Defining the transfer homomorphism ϕ more or less as in the proof of Theorem 1, we use Case 1 to obtain

$$f_{YX}(t) = t^{n-m} f_{XY}(t),$$

and hence, upon taking images under ϕ,

$$f_{BA}(t) = t^{n-m} f_{AB}(t).$$

As an algebraist, I was unhappy when years ago I first encountered the "continuity" proof that $\det(AB) = (\det A)(\det B)$ given in [**5**], especially because the extension to singular A or B was so easy to carry out algebraically. However, it did motivate me to look for an algebraic equivalent of the continuity argument. My solution to this problem was improved by my exposure to the notion of a "generic element" in [**4**] while taking a graduate seminar in Lie algebras. Prior to obtaining these proofs, I had not seen Theorems 4–6 in the literature, but it was later pointed out to me that Theorem 6 for matrices over a field can be found in [**3**]. All of these results are so easy and natural that they probably appear somewhere in the literature. However, my interest was more in the technique than in the specific results. Perhaps some teachers and students of linear algebra may find some pleasure and utility in these ideas, just as I have.

References

[1] George Grätzer, Universal Algebra, Van Nostrand, Princeton, 1968, p. 162.
[2] Emil Grosswald, Topics from the Theory of Numbers, Macmillan, New York, 1966, p. 59.
[3] Nathan Jacobson, Lectures in Abstract Algebra, vol. II, Van Nostrand, New York, 1953, p. 106, Th. 16.
[4] ______, Lie Algebras, Interscience, New York, 1962, pp. 60–61.
[5] Murray H. Protter and Charles B. Morrey, Jr., Modern Mathematical Analysis, Addison-Wesley, Palo Alto, 1964, pp. 327–328.
[6] B. L. van der Waerden, Algebra, vol. 1, Ungar, New York, 1970, pp. 101–107.