

Syracuse University

SURFACE

Mathematics - Faculty Scholarship

Mathematics

8-31-2004

The Adjoint of an Even Size Matrix Factors

Ragnar-Olaf Buchweitz

University of Toronto

Graham J. Leuschke

Syracuse University

Follow this and additional works at: <https://surface.syr.edu/mat>



Part of the [Mathematics Commons](#)

Recommended Citation

Buchweitz, Ragnar-Olaf and Leuschke, Graham J., "The Adjoint of an Even Size Matrix Factors" (2004).
Mathematics - Faculty Scholarship. 32.

<https://surface.syr.edu/mat/32>

This Article is brought to you for free and open access by the Mathematics at SURFACE. It has been accepted for inclusion in Mathematics - Faculty Scholarship by an authorized administrator of SURFACE. For more information, please contact surface@syr.edu.

THE ADJOINT OF AN EVEN SIZE MATRIX FACTORS

RAGNAR-OLAF BUCHWEITZ AND GRAHAM LEUSCHKE

ABSTRACT. We show that the adjoint matrix of a generic square matrix of even size can be factored nontrivially. This answers a question of G. Bergman. This note should be considered a preliminary report on work in progress.

1. DETERMINANTS AND DERIVATIONS

1.1. Let K be a commutative ring, $X = (x_{ij})$ the generic $(n \times n)$ -matrix, whose entries thus form a family of n^2 indeterminates, and set $S = K[x_{ij}]$, the polynomial ring over K in these variables.

1.2. The determinant $\det(X)$ of the generic matrix X is a nonzerodivisor in S , and the classical adjoint matrix $\text{adj}(X)$ of X is uniquely determined through either of the following two matrix equations

$$(*) \quad \text{adj}(X)X = \det(X) \text{id}_n \quad \text{and} \quad X \text{adj}(X) = \det(X) \text{id}_n,$$

where id_n represents the $n \times n$ identity matrix.

1.3. We will use the following notation for minors of the generic matrix X : Let $[i_1 i_2 \dots i_k \mid j_1 j_2 \dots j_k]$ denote the (unsigned) determinant of the $(k \times k)$ -submatrix of X that consists of the rows indexed $1 \leq i_1 < \dots < i_k \leq n$, and of the columns indexed $1 \leq j_1 < \dots < j_k \leq n$.

The symbol $[i_1 i_2 \dots i_k \mid \widehat{j_1 j_2 \dots j_k}]$ will denote the complementary minor, thus, the determinant of the $(n - k) \times (n - k)$ -submatrix of X obtained by removing the rows indexed i_ν and the columns indexed j_ν . For consistency, the empty determinant, for $k = n$, has value 1.

We extend the symbols $[? \mid ?]$ and $[? \mid \widehat{?}]$ to not necessarily strictly increasing index sets by requiring them to be alternating in both the left and right arguments. In particular, each symbol vanishes if there is repetition of indices either before or after the vertical bar.

1.4. If U is any $(n \times n)$ -matrix over some K -algebra R , then there exists a unique K -algebra homomorphism $ev_U : S \rightarrow R$, $x_{ij} \mapsto u_{ij}$ that transforms the entries of X to those of U . Let $[\dots](U) = ev_U([\dots])$ represent the corresponding minor of the matrix U , and write $I_t(U) \subseteq R$ for the ideal generated by all the $(t \times t)$ -minors of U . The transpose of a matrix U will be denoted U^T .

Example 1.5. The (i, j) -th entry of the adjoint matrix can be written as

$$\text{adj}(X)_{ij} = (-1)^{i+j} [j \mid \widehat{i}] = (-1)^{i+j} [1 \dots \widehat{j} \dots n \mid 1 \dots \widehat{i} \dots n].$$

Date: March 16, 2011.

The first author was partly supported by NSERC grant 3-642-114-80.

1.6. Recall that a map $D : R \rightarrow R$, on a not necessarily commutative ring R , is a *derivation* if $D(ab) = D(a)b + aD(b)$ for any elements $a, b \in R$.

For example, the *partial derivation* $\partial_{i_j} = \frac{\partial}{\partial x_{ij}}$ with respect to the variable x_{ij} defines a derivation on S that is furthermore K -linear. These partial derivations form indeed a basis of the free S -module $\text{Der}_K(S)$ of all K -linear derivations on S ,

$$\text{Der}_K(S) \cong \bigoplus_{1 \leq i, j \leq n} S \partial_{i_j}.$$

Now we state the facts on derivations and minors that we will use.

Lemma 1.7. *If R is a commutative ring, $D : R \rightarrow R$ a derivation, and U an $(n \times n)$ -matrix over R , then*

$$D(\det U) = \sum_{i=1}^n \begin{vmatrix} u_{11} & \cdots & u_{1n} \\ \vdots & & \vdots \\ D(u_{i1}) & \cdots & D(u_{in}) \\ \vdots & & \vdots \\ u_{n1} & \cdots & u_{nn} \end{vmatrix} = \sum_{i=1}^n \begin{vmatrix} u_{11} & \cdots & D(u_{1j}) & \cdots & u_{1n} \\ \vdots & & \vdots & & \vdots \\ u_{n1} & \cdots & D(u_{nj}) & \cdots & u_{nn} \end{vmatrix}$$

where $|V|$ denotes the determinant of the matrix V .

Proof. This follows immediately from the Leibnitz rule for derivations applied to the complete expansion of the determinant. \square

Lemma 1.8. *Let X be again the generic matrix and S the associated polynomial ring over K .*

- (1) *For any pair of indices $1 \leq i, j \leq n$,*

$$\partial_{i_j}(\det X) = \text{adj}(X)_{ji}$$

equivalently,

$$\text{adj}(X)^T = (\partial_{i_j}(\det X))_{ij}.$$

- (2) *For any pair of indices $1 \leq i, j \leq n$,*

$$\sum_{\nu=1}^n x_{i\nu} \partial_{j\nu}(\det X) = \delta_{ij} \det(X) = \sum_{\nu=1}^n x_{\nu i} \partial_{\nu j}(\det X),$$

where δ_{ij} is the Kronecker symbol.

- (3) *For any indices $1 \leq i_1, i_2, \dots, i_k \leq n$ and $1 \leq j_1, j_2, \dots, j_k \leq n$,*

$$\partial_{i_1 j_1} \cdots \partial_{i_k j_k}(\det X) = (-1)^{i_1 + \cdots + i_k + j_1 + \cdots + j_k} [i_1 \dots i_k \widehat{j_1, \dots, j_k}],$$

in particular, these terms vanish whenever there is a repetition among the i 's or the j 's.

Proof. Claim (1) follows from 1.7 with $D = \partial_{i_j}$ and $U = X$. In view of (1), claim (2) is simply a reformulation of the equation (*) above. To see (3), apply first 1.7 or (1) to the generic matrix using the derivation $\partial_{i_k j_k}$, and then use induction on $k \geq 1$. \square

2. THE FACTORIZATIONS

We now use the “differential calculus” from the previous section to establish two factorization results about products of the transpose of the adjoint matrix with *alternating* matrices on one or both sides. Recall that an $(n \times n)$ -matrix $A = (a_{kl})$ is alternating if $A^T = -A$ and the diagonal elements vanish, $a_{kk} = 0$ for each $k = 1, \dots, n$. The latter condition is of course a consequence of the first as soon as 2 is a nonzerodivisor in K .

Theorem 2.1. *Let U, A be $(n \times n)$ -matrices over a commutative ring R , with A alternating. The $(n \times n)$ -matrix $B = (b_{rs})$ with entries from $I_1(A) \cdot I_{n-2}(U) \subseteq R$, given by*

$$b_{rs} = \sum_{k < l} a_{kl} (-1)^{k+l+r+s} [kl \widehat{\mid} rs](U)$$

is then alternating as well and satisfies the matrix equation

$$(**) \quad A \cdot \text{adj}(U)^T = UB.$$

If $\det U$ is a nonzerodivisor in R , then B is the unique solution to this equation.

Proof. As $[kl \widehat{\mid} sr] = -[kl \widehat{\mid} rs]$ and $[kl \widehat{\mid} rr] = 0$, the matrix B is alternating. To verify that B satisfies (**), it suffices to establish the generic case, where $R = S$ and $U = X$. Let E_{ij} denote the elementary $(n \times n)$ -matrix with 1 at position (i, j) as its only nonzero entry. Recall that $E_{rs}E_{uv} = \delta_{su}E_{rv}$ for any indices $1 \leq r, s, u, v \leq n$. As $\partial_{kr}\partial_{ls}(\det X) = (-1)^{k+l+r+s}[kl \widehat{\mid} rs]$ by Lemma 1.8(3), the right hand side of (**) expands now first as

$$\begin{aligned} XB &= \left(\sum_{i, \nu} x_{i\nu} E_{i\nu} \right) \left(\sum_{\mu, j} \sum_{k < l} a_{kl} \partial_{k\mu} \partial_{lj}(\det X) E_{\mu j} \right) \\ &= \sum_{k < l} a_{kl} \sum_{i, j} \left(\sum_{\nu} x_{i\nu} \partial_{k\nu} \partial_{lj}(\det X) \right) E_{ij}. \end{aligned}$$

The innermost sum can be simplified using first that partial derivatives commute, then applying the product rule, and finally invoking Lemma 1.8(2) together with the fact that $\partial_{lj}(x_{i\nu}) = \delta_{il}\delta_{j\nu}$. In detail, these steps yield the following equalities:

$$\begin{aligned} \sum_{\nu} x_{i\nu} \partial_{k\nu} \partial_{lj}(\det X) &= \sum_{\nu} x_{i\nu} \partial_{lj} \partial_{k\nu}(\det X) \\ &= \sum_{\nu} \partial_{lj}(x_{i\nu} \partial_{k\nu}(\det X)) - \sum_{\nu} \partial_{lj}(x_{i\nu}) \partial_{k\nu}(\det X) \\ &= \partial_{lj} \left(\sum_{\nu} x_{i\nu} \partial_{k\nu}(\det X) \right) - \delta_{il} \sum_{\nu} \delta_{j\nu} \partial_{k\nu}(\det X) \\ &= \delta_{ik} \partial_{lj}(\det X) - \delta_{il} \partial_{kj}(\det X) \end{aligned}$$

In light of this simplification, we may expand XB further as follows:

$$\begin{aligned}
XB &= \sum_{k < l} a_{kl} \sum_{i,j} \left(\sum_{\nu} x_{i\nu} \partial_{k\nu} \partial_{l_j}(\det X) \right) E_{ij} \\
&= \sum_{k < l} a_{kl} \sum_{i,j} \left(\delta_{ik} \partial_{l_j}(\det X) - \delta_{il} \partial_{k_j}(\det X) \right) E_{ij} \\
&= \sum_{k < l} a_{kl} \sum_j \left(\partial_{l_j}(\det X) E_{k_j} - \partial_{k_j}(\det X) E_{l_j} \right) \\
&= \sum_{k < l} a_{kl} \left(E_{kl} \sum_j \partial_{l_j}(\det X) E_{l_j} - E_{lk} \sum_j \partial_{k_j}(\det X) E_{k_j} \right) \\
&= \sum_{k < l} a_{kl} \left(E_{kl} \sum_{i,j} \partial_{ij}(\det X) E_{ij} - E_{lk} \sum_{i,j} \partial_{ij}(\det X) E_{ij} \right) \\
&= \sum_{k < l} a_{kl} (E_{kl} - E_{lk}) \sum_{i,j} \partial_{ij}(\det X) E_{ij} \\
&= A \cdot \text{adj}(X)^T
\end{aligned}$$

with the last equality using that A is alternating, thus $A = \sum_{k < l} a_{kl} (E_{kl} - E_{lk})$, and that $\text{adj}(X)^T = \sum_{i,j} \partial_{ij}(\det X) E_{ij}$, in view of 1.8(1).

The final assertion about uniqueness follows from (**) by multiplying from the left with $\text{adj}(U)$ and using equation (*) to obtain

$$\text{adj}(U) \cdot A \cdot \text{adj}(U)^T = \det(U)B.$$

□

Remark 2.2. One may formulate 2.1 equally well for multiplication of the transpose of the adjoint matrix from the right by an alternating matrix. Namely, assume A, B are $(n \times n)$ -matrices over S satisfying $A \text{adj}(X)^T = XB$. Let $\varphi : S \rightarrow S$ be the K -algebra automorphism uniquely determined through $\varphi(x_{ij}) = x_{ji}$. Clearly, φ is involutive and exchanges X and its transpose, $\varphi(X) = X^T$. Moreover, $\varphi(\text{adj}(X)) = \text{adj}(X)^T$, in view of equation (*). Now

$$\begin{aligned}
A \text{adj}(X)^T &= XB && \text{if, and only if,} \\
\varphi(A) \varphi(\text{adj}(X)^T) &= \varphi(X) \varphi(B) && \text{if, and only if,} \\
\varphi(A) \text{adj}(X) &= X^T \varphi(B) && \text{if, and only if,} \\
\text{adj}(X)^T \varphi(A)^T &= \varphi(B)^T X.
\end{aligned}$$

In case A, B are alternating, then so are $\varphi(A), \varphi(B)$ and the last equation is equivalent to

$$\text{adj}(X)^T \varphi(A) = \varphi(B)X.$$

We now investigate what happens when multiplying simultaneously from both left and right.

Theorem 2.3. *Let U, A, B denote the same matrices as introduced in 2.1. If A' is another alternating $(n \times n)$ -matrix, then the $(n \times n)$ -matrix $C = (c_{wm})$ with entries*

from $I_1(A) \cdot I_{n-3}(U) \cdot I_1(A') \subseteq R$ given by

$$c_{wm} = \sum_{k < l, u < v} (-1)^{k+l+m+u+v+w} a_{kl} [klm \widehat{\mid} uvw](U) a'_{uv}$$

satisfies

$$(***) \quad BA' = r \operatorname{id}_n + CU,$$

where

$$r = - \sum_{k < l, u < v} (-1)^{k+l+u+v} a_{kl} [kl \widehat{\mid} uv](U) a'_{uv} \in R.$$

Proof. It suffices again to verify the result for the generic matrix $U = X$, in which case we can employ once more the description of minors as given in 1.8(3). The straightforward calculation proceeds then as follows:

$$\begin{aligned} (BA' - r \operatorname{id}_n)_{ij} &= \sum_m \sum_{k < l} (-1)^{k+l+i+m} a_{kl} [kl \widehat{\mid} im] a'_{mj} \\ &\quad + \delta_{ij} \sum_{k < l} \sum_{u < v} (-1)^{k+l+u+v} a_{kl} [kl \widehat{\mid} uv] a'_{uv} \\ &= \sum_{k < l} a_{kl} \left(\sum_m \partial_{ki} \partial_{lm} (\det X) a'_{mj} + \sum_{u < v} \partial_{ku} \partial_{lv} (\det X) a'_{uv} \delta_{ij} \right) \\ &= \sum_{k < l, m} a_{kl} \left(\partial_{ki} \partial_{lm} (\det X) a'_{mj} + \sum_{u < v} \partial_{ku} \partial_{lv} (\partial_{mi} (\det X) x_{mj}) a'_{uv} \right) \end{aligned}$$

where we have used 1.8(2) in the last step. Using the product rule twice together with $\partial_{rs}(x_{mn}) = \delta_{rm} \delta_{sn}$, we find next

$$\begin{aligned} \partial_{ku} \partial_{lv} (\partial_{mi} (\det X) x_{mj}) &= \partial_{ku} \partial_{mi} (\det X) \delta_{lm} \delta_{vj} + \partial_{lv} \partial_{mi} (\det X) \delta_{km} \delta_{uj} \\ &\quad + \partial_{ku} \partial_{lv} \partial_{mi} (\det X) x_{mj} \end{aligned}$$

Substituting and evaluating the Kronecker symbols yields

$$\begin{aligned} (BA' - r \operatorname{id}_n)_{ij} &= \sum_{k < l, m} a_{kl} \left(\partial_{ki} \partial_{lm} (\det X) a'_{mj} + \sum_{u < v} \partial_{ku} \partial_{lv} (\partial_{mi} (\det X) x_{mj}) a'_{uv} \right) \\ &= \sum_{k < l} a_{kl} \left(\sum_m \partial_{ki} \partial_{lm} (\det X) a'_{mj} + \sum_{u < j} \partial_{ku} \partial_{li} (\det X) a'_{uj} \right. \\ &\quad \left. + \sum_{j < v} \partial_{ki} \partial_{lv} (\det X) a'_{jv} + \sum_{u < v, m} \partial_{ku} \partial_{lv} \partial_{mi} (\det X) a'_{uv} x_{mj} \right) \end{aligned}$$

The terms involving only second order derivatives of the determinant cancel. To see this, rename summation indices, use that $\partial_{km} \partial_{li} (\det X) = -\partial_{ki} \partial_{lm} (\det X)$ and

that A' is alternating, whence its entries satisfy $a'_{mm} = 0, a'_{jm} = -a'_{mj}$. In detail,

$$\begin{aligned}
(BA' - r \operatorname{id}_n)_{ij} &= \sum_{k < l} a_{kl} \left(\sum_m \partial_{ki} \partial_{lm} (\det X) a'_{mj} - \sum_{m < j} \partial_{km} \partial_{li} (\det X) a'_{mj} \right. \\
&\quad \left. - \sum_{j < m} \partial_{ki} \partial_{lm} (\det X) a'_{mj} + \sum_{u < v} \sum_m \partial_{ku} \partial_{lv} \partial_{mi} (\det X) a'_{uv} x_{mj} \right) \\
&= \sum_m \left(\sum_{k < l, u < v} a_{kl} \partial_{ku} \partial_{lv} \partial_{mi} (\det X) a'_{uv} \right) x_{mj} \\
&= \sum_m \left(\sum_{k < l, u < v} (-1)^{k+l+m+u+v+i} a_{kl} [klm \widehat{uv} i] a'_{uv} \right) x_{mj} \\
&= (CX)_{ij}
\end{aligned}$$

where we have evaluated the third order derivatives of the determinant according to 1.8(3). \square

Combining the results from 2.1 and 2.3 yields the following.

Corollary 2.4. *Let U, A, A' be $(n \times n)$ -matrices over a commutative ring R , with A, A' alternating. One then has an equality of matrices*

$$A \operatorname{adj}(U)^T A' = rU + UCU,$$

where r and C are as specified in 2.3. \square

Remark 2.5. The element $r \in I_1(A) \cdot I_{n-2}(U) \cdot I_1(A') \subseteq R$ is a “half trace” of BA' , as

$$\begin{aligned}
\operatorname{tr}(BA') &= \sum_{k < l} \sum_{i, j} a_{kl} (-1)^{k+l+i+j} [kl \widehat{ij}] a'_{ji} \\
&= 2 \sum_{k < l} \sum_{i < j} a_{kl} (-1)^{k+l+i+j} [kl \widehat{ij}] a'_{ji} \\
&= 2r
\end{aligned}$$

invoking once again that A' is alternating. Equivalently, $\operatorname{tr}(CU) = (2 - n)r$.

Remark 2.6. If $n = 2$, all expressions of the form $[klm \widehat{uv} i]$ vanish, and 2.1 together with 2.3 specialize to the easily established identity

$$\begin{pmatrix} 0 & a \\ -a & 0 \end{pmatrix} \begin{pmatrix} x_{22} & -x_{21} \\ -x_{12} & x_{11} \end{pmatrix} \begin{pmatrix} 0 & b \\ -b & 0 \end{pmatrix} = -ab \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix}.$$

If the size $n = 2m$ is even, then there are *invertible* alternating matrices of that size over any commutative ring. For example, the alternating “hyperbolic matrix”

$$\begin{pmatrix} 0 & \operatorname{id}_m \\ -\operatorname{id}_m & 0 \end{pmatrix}$$

has determinant equal to 1 over any ring.

Corollary 2.7. *If n is even, then the adjoint of the generic matrix admits nontrivial factorizations*

$$\operatorname{adj}(X) = YZ = Y'Z'$$

into products of $(n \times n)$ -matrices over S with $\det(Y) = \det(Z') = \det(X)$.

More precisely, any pair of alternating $(n \times n)$ -matrices A, A' of determinant equal to 1 over S gives rise to such factorizations. With r and C the data associated to A, A' as in 2.3, one may take

$$\begin{aligned} Y &= (A')^{-1}X^T \quad \text{and} \quad Z = (r \operatorname{id}_n + C^T X^T)A^{-1}, \\ Y' &= (A')^{-1}(r \operatorname{id}_n + X^T C^T) \quad \text{and} \quad Z' = X^T A^{-1}. \end{aligned}$$

Proof. Transposing the equation in 2.4 for $U = X$ yields first

$$(A')^T \operatorname{adj}(X)A^T = rX^T + X^T C^T X^T.$$

As A, A' are invertible and alternating, this equality is equivalent to

$$\operatorname{adj}(X) = (A')^{-1}(rX^T + X^T C^T X^T)A^{-1}.$$

□

Remark 2.8. Bergman [1] shows that, over a field K of characteristic zero, in any factorization $\operatorname{adj}(X) = YZ$ of the generic adjoint matrix into noninvertible factors, either $\det(Y) = \det(X)$ or $\det(Z) = \det(X)$, up to units of S .

REFERENCES

- [1] G.M. Bergman, *Can one factor the classical adjoint of a generic matrix?*, preprint (arXiv: math.AC/0306126).

DEPT. OF MATH., UNIVERSITY OF TORONTO, TORONTO, ONT. M5S 3G3, CANADA
E-mail address: ragnar@math.utoronto.ca

DEPT. OF MATH., UNIVERSITY OF TORONTO, TORONTO, ONT. M5S 3G3, CANADA
E-mail address: gleuschk@math.toronto.edu