Lecture Notes 1: Matrix Algebra
Part B: Determinants and Inverses

Peter J. Hammond
email: p.j.hammond@warwick.ac.uk

Autumn 2012, revised Autumn 2014

Lecture Outline

Outline

Determinants of Order 2: Definition

Consider again the pair of linear equations

$$\begin{array}{ccccc} a_{11}x_1 & + & a_{12}x_2 & = & b_1 \\ a_{21}x_1 & + & a_{12}x_2 & = & b_2 \end{array}$$

with its associated coefficient matrix

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

Let us define $D := a_{11}a_{22} - a_{21}a_{12}$.

Provided that $D \neq 0$, there is a unique solution given by

$$x_1 = \frac{1}{D}(b_1 a_{22} - b_2 a_{12}), \quad x_2 = \frac{1}{D}(b_2 a_{11} - b_1 a_{21})$$

The number D is called the determinant of the matrix $\mathbf{A}$, and denoted by either $\det(\mathbf{A})$ or more concisely, $|\mathbf{A}|$.

Determinants of Order 2: Simple Rule

Thus, for any 2×2 matrix $\mathbf{A}$, its determinant D is

$$|\mathbf{A}| = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{21}a_{12}$$

For this special case of order 2 determinants, a simple rule is:

1. multiply the diagonal elements together;
2. multiply the off-diagonal elements together;
3. subtract the product of the off-diagonal elements from the product of the diagonal elements.

Note that

$$|\mathbf{A}| = a_{11}a_{22} \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} + a_{21}a_{12} \begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix}$$

Cramer's Rule in the 2×2 Case

Using determinant notation, the solution to the equations

$$\begin{array}{ccccc} a_{11}x_1 & + & a_{12}x_2 & = & b_1 \\ a_{21}x_1 & + & a_{12}x_2 & = & b_2 \end{array}$$

can be written in the alternative form

$$x_1 = \frac{1}{D} \begin{vmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{vmatrix}, \qquad x_2 = \frac{1}{D} \begin{vmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{vmatrix}$$

This accords with Cramer's rule for the solution to $\mathbf{Ax} = \mathbf{b}$, which is the vector $\mathbf{x} = (x_i)_{i=1}^n$ each of whose components x_i is the fraction with:

1. denominator equal to the determinant D of the coefficient matrix $\mathbf{A}$ (provided, of course, that $D \neq 0$);
2. numerator equal to the determinant of the matrix $(\mathbf{A}_{-i}, \mathbf{b})$ formed from $\mathbf{A}$ by replacing its ith column with the $\mathbf{b}$ vector of right-hand side elements.

Outline

Determinants of Order 3: Definition

Determinants of order 3 can be calculated
from those of order 2 according to the formula

$$\begin{aligned} |\mathbf{A}| &= a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} \\ &= \sum\nolimits_{j=1}^{3} (-1)^{1+j} a_{1j} |\mathbf{C}_{1j}| \end{aligned}$$

where, for $j = 1, 2, 3$, the 2×2 matrix $\mathbf{C}_{1j}$ is the $(1, j)$-cofactor
obtained by removing both row 1 and column j from $\mathbf{A}$.

The result is the following sum

$$\begin{aligned} |\mathbf{A}| = a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} + a_{12}a_{23}a_{31} \\ - a_{12}a_{21}a_{33} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} \end{aligned}$$

of $3! = 6$ terms, each the product of 3 elements chosen
so that each row and each column is represented just once.

Determinants of Order 3: Cofactor Expansion

The determinant expansion

$$|\mathbf{A}| = a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} + a_{12}a_{23}a_{31} \\ - a_{12}a_{21}a_{33} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31}$$

is very symmetric, suggesting (correctly)
that the cofactor expansion along the first row (a_{11}, a_{12}, a_{13})

$$|\mathbf{A}| = \sum_{j=1}^{3} (-1)^{1+j} a_{1j} |\mathbf{C}_{1j}|$$

gives the same answer as the two cofactor expansions

$$|\mathbf{A}| = \sum_{j=1}^{3} (-1)^{r+j} a_{rj} |\mathbf{C}_{rj}| = \sum_{i=1}^{3} (-1)^{i+s} a_{is} |\mathbf{C}_{is}|$$

along, respectively:

- the rth row (a_{r1}, a_{r2}, a_{r3})
- the sth column (a_{1s}, a_{2s}, a_{3s})

Determinants of Order 3: Alternative Expressions

The same result

$$|\mathbf{A}| = a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} + a_{12}a_{23}a_{31} \\ - a_{12}a_{21}a_{33} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31}$$

can be obtained as either of the two expansions

$$\begin{aligned} |\mathbf{A}| &= \sum_{j_1=1}^{3} \sum_{j_2=1}^{3} \sum_{j_3=1}^{3} \epsilon_{j_1 j_2 j_3} a_{1 j_1} a_{2 j_2} a_{3 j_3} \\ &= \sum_{\pi \in \Pi} \operatorname{sgn}(\pi) \prod_{i=1}^{3} a_{i \pi(i)} \end{aligned}$$

Here $\epsilon_{\mathbf{j}} = \epsilon_{j_1 j_2 j_3} \in \{-1, 0, 1\}$ denotes the Levi-Civita symbol associated with the mapping $i \mapsto j_i$ from $\{1, 2, 3\}$ into itself.

Also, Π denotes the set
of all $3! = 6$ possible permutations on $\{1, 2, 3\}$,
with typical member π, whose sign is denoted by $\operatorname{sgn}(\pi)$.

Outline

The Determinant Function

When $n = 1, 2, 3$, the determinant mapping $\mathbf{A} \mapsto |\mathbf{A}| \in \mathbb{R}$ specifies the determinant $|\mathbf{A}|$ of each $n \times n$ matrix $\mathbf{A}$ as a function of its n row vectors $(\mathbf{a}_i)_{i=1}^n$.

For a general natural number $n \in \mathbb{N}$, consider any mapping

$$\mathcal{D}_n \ni \mathbf{A} \mapsto D(\mathbf{A}) = D\left((\mathbf{a}_i)_{i=1}^n\right) \in \mathbb{R}$$

defined on the domain $\mathcal{D}_n$ of $n \times n$ matrices.

Notation: Let $D(\mathbf{A}/\mathbf{b}_r)$ denote
the new value $D(\mathbf{a}_1, \ldots, \mathbf{a}_{r-1}, \mathbf{b}_r, \mathbf{a}_{r+1}, \ldots, \mathbf{a}_n)$ of the function D
after the rth row $\mathbf{a}_r$ of the matrix $\mathbf{A}$
has been replaced by the new row vector $\mathbf{b}_r$.

Row Multilinearity

Definition

The function $\mathcal{D}_n \ni \mathbf{A} \mapsto D(\mathbf{A})$ of $\mathbf{A}$'s n rows $(\mathbf{a}_i)_{i=1}^n$
is (row) multilinear just in case,
for each row number $i \in \{1, 2, \dots, n\}$,
each pair $\mathbf{b}_i, \mathbf{c}_i \in \mathbb{R}^n$ of new versions of row i,
and each pair of scalars $\lambda, \mu \in \mathbb{R}$, one has

$$D(\mathbf{A}/\lambda\mathbf{b}_i + \mu\mathbf{c}_i) = \lambda D(\mathbf{A}/\mathbf{b}_i) + \mu D(\mathbf{A}/\mathbf{c}_i)$$

Formally, the mapping $\mathbb{R}^n \ni \mathbf{a}_i \mapsto D(\mathbf{A}/\mathbf{a}_i) \in \mathbb{R}$ should be linear,
for fixed each row $i \in \mathbb{N}_n$.

That is, D is a linear function of the ith row vector $\mathbf{a}_i$ on its own,
when all the other rows $\mathbf{a}_h$ $(h \neq i)$ are fixed.

The Three Characterizing Properties

Definition

The function $\mathcal{D}_n \ni \mathbf{A} \mapsto D(\mathbf{A})$ is alternating
just in case
for every transposition matrix $\mathbf{T}$, one has $D(\mathbf{TA}) = -D(\mathbf{A})$
— i.e., interchanging any two rows reverses its sign.

Definition

The mapping $\mathcal{D}_n \ni \mathbf{A} \mapsto D(\mathbf{A})$ is of the determinant type
just in case:

1. D is multilinear in its rows;
2. D is alternating;
3. $D(\mathbf{I}_n) = 1$ for the identity matrix $\mathbf{I}_n$.

Exercise

*Show that the mapping $\mathcal{D}_n \ni \mathbf{A} \mapsto |\mathbf{A}| \in \mathbb{R}$
is of the determinant type provided that $n \leq 3$.*

First Implication of Multilinearity in the $n \times n$ Case

Lemma

Suppose that $\mathcal{D}_n \ni \mathbf{A} \mapsto D(\mathbf{A})$ is multilinear in its rows.
For any fixed $\mathbf{B} \in \mathcal{D}_n$, the value of $D(\mathbf{AB})$ can be expressed as the linear combination

$$D(\mathbf{AB}) = \sum_{j_1=1}^{n} \sum_{j_2=1}^{n} \ldots \sum_{j_n=1}^{n} a_{1j_1} a_{2j_2} \cdots a_{nj_n} D(\mathbf{L}_{j_1 j_2 \ldots j_n} \mathbf{B})$$

of its values at all possible matrices

$$\mathbf{L_j B} = \mathbf{L}_{j_1 j_2 \ldots j_n} \mathbf{B} := (\mathbf{b}_{j_r})_{r=1}^{n}$$

whose rth row, for each $r = 1, 2, \ldots, n$, equals the j_rth row $\mathbf{b}_{j_r}$ of the matrix $\mathbf{B}$.

Characterizing 2×2 Determinants

1. In the case of 2×2 matrices, the lemma tells us that multilinearity implies

$$\begin{aligned} D(\mathbf{AB}) = a_{11}a_{21}D(\mathbf{b}_1, \mathbf{b}_1) + a_{11}a_{22}D(\mathbf{b}_1, \mathbf{b}_2) \\ + a_{12}a_{21}D(\mathbf{b}_2, \mathbf{b}_1) + a_{12}a_{22}D(\mathbf{b}_2, \mathbf{b}_2) \end{aligned}$$

   where $\mathbf{b}_1 = (b_{11}, b_{21})$ and $\mathbf{b}_2 = (b_{12}, b_{22})$ are the rows of $\mathbf{B}$.
2. If D is also alternating, then $D(\mathbf{b}_1, \mathbf{b}_1) = D(\mathbf{b}_2, \mathbf{b}_2) = 0$ and $D(\mathbf{B}) = D(\mathbf{b}_1, \mathbf{b}_2) = -D(\mathbf{b}_2, \mathbf{b}_1)$, implying that

$$\begin{aligned} D(\mathbf{AB}) &= a_{11}a_{22}D(\mathbf{b}_1, \mathbf{b}_2) + a_{12}a_{21}D(\mathbf{b}_2, \mathbf{b}_1) \\ &= (a_{11}a_{22} - a_{12}a_{21})D(\mathbf{B}) \end{aligned}$$

3. Imposing the additional restriction $D(\mathbf{B}) = 1$ when $\mathbf{B} = \mathbf{I}_2$, we obtain the ordinary determinant $D(\mathbf{A}) = a_{11}a_{22} - a_{12}a_{21}$.
4. Then, too, one derives the product rule $D(\mathbf{AB}) = D(\mathbf{A})D(\mathbf{B})$.

First Implication of Multilinearity: Proof

Each element of the product $\mathbf{C} = \mathbf{AB}$ satisfies $c_{ik} = \sum_{j=1}^{n} a_{ij} b_{jk}$.

Hence each row $\mathbf{c}_i = (c_{ik})_{k=1}^{n}$ of $\mathbf{C}$ can be expressed as the linear combination $\mathbf{c}_i = \sum_{j=1}^{n} a_{ij} \mathbf{b}_j$ of $\mathbf{B}$'s rows.

For each $r = 1, 2, \ldots, n$ and arbitrary selection $\mathbf{b}_{j_1}, \ldots, \mathbf{b}_{j_{r-1}}$ of $r - 1$ rows from $\mathbf{B}$, multilinearity therefore implies that

$$
\begin{aligned}
D(\mathbf{b}_{j_1}, \ldots, \mathbf{b}_{j_{r-1}}, \mathbf{c}_r, \mathbf{c}_{r+1}, \ldots, \mathbf{c}_n) & \\
&= \sum\nolimits_{j_r=1}^{n} a_{ij_r} D(\mathbf{b}_{j_1}, \ldots, \mathbf{b}_{j_{r-1}}, \mathbf{b}_{j_r}, \mathbf{c}_{r+1}, \ldots, \mathbf{c}_n)
\end{aligned}
$$

This equation can be used to show, by induction on k, that

$$
D(\mathbf{C}) = \sum_{j_1=1}^{n} \sum_{j_2=1}^{n} \cdots \sum_{j_k=1}^{n} a_{1j_1} a_{2j_2} \cdots a_{kj_k} D(\mathbf{b}_{j_1}, \ldots, \mathbf{b}_{j_k}, \mathbf{c}_{k+1}, \ldots, \mathbf{c}_n)
$$

for $k = 1, 2, \ldots, n$, including for $k = n$ as the lemma claims. □

Additional Implications of Alternation

Lemma

Suppose $\mathcal{D}_n \ni \mathbf{A} \mapsto D(\mathbf{A})$ is both row multilinear and alternating.

Then for all possible $n \times n$ matrices $\mathbf{A}, \mathbf{B}$,
and for all possible permutation matrices $\mathbf{P}^\pi$, one has:

1. $D(\mathbf{AB}) = \sum_{\pi \in \Pi} \prod_{i=1}^n a_{i\pi(i)} D(\mathbf{P}^\pi \mathbf{B})$
2. $D(\mathbf{P}^\pi \mathbf{B}) = \text{sgn}(\pi) D(\mathbf{B})$.
3. *Under the additional assumption that $D(\mathbf{I}_n) = 1$, one has:*

   determinant formula: $D(\mathbf{A}) = \sum_{\pi \in \Pi} \text{sgn}(\pi) \prod_{i=1}^n a_{i\pi(i)}$;

   product rule: $D(\mathbf{AB}) = D(\mathbf{A}) D(\mathbf{B})$

First Additional Implication of Alternation: Proof

Because D is alternating,
one has $D(\mathbf{B}) = 0$ whenever two rows of $\mathbf{B}$ are equal.

It follows that for any matrix $(\mathbf{b}_{j_i})_{i=1}^n = \mathbf{L_j B}$
whose n rows are all rows of the matrix $\mathbf{B}$,
one has $D((\mathbf{b}_{j_i})_{i=1}^n) = 0$ unless these rows are all different.

But if all the n rows of $(\mathbf{b}_{j_i})_{i=1}^n = \mathbf{L_j B}$ are different,
there exists a permutation $\pi \in \Pi$ such that $\mathbf{L_j B} = \mathbf{P}^\pi \mathbf{B}$.

Hence, after eliminating terms that are zero, the sum

$$\begin{aligned} D(\mathbf{AB}) &= \sum_{j_1=1}^n \sum_{j_2=1}^n \cdots \sum_{j_n=1}^n a_{1j_1} a_{2j_2} \cdots a_{nj_n} D((\mathbf{b}_{j_r})_{r=1}^n) \\ &= \sum_{j_1=1}^n \sum_{j_2=1}^n \cdots \sum_{j_n=1}^n a_{1j_1} a_{2j_2} \cdots a_{nj_n} D(\mathbf{L}_{j_1 j_2 \ldots j_n} \mathbf{B}) \end{aligned}$$

as stated in part 1 of the Lemma. □

Second Additional Implication: Proof

Because D is alternating, one has $D(\mathbf{T}\mathbf{P}^\pi\mathbf{B}) = -D(\mathbf{P}^\pi\mathbf{B})$ whenever $\mathbf{T}$ is a transposition matrix.

Suppose that $\pi = \tau^1 \circ \cdots \circ \tau^q$ is one possible "factorization" of the permutation π as the composition of transpositions.

But $\text{sgn}(\tau) = -1$ for any transposition τ.

So $\text{sgn}(\pi) = (-1)^q$ by the product rule for signs of permutations.

Note that $\mathbf{P}^\pi = \mathbf{T}^1\mathbf{T}^2\cdots\mathbf{T}^q$
where $\mathbf{T}^p$ denotes the permutation matrix
corresponding to the transposition τ^p, for each $p = 1, \ldots, q$

It follows that

$$D(\mathbf{P}^\pi\mathbf{B}) = D(\mathbf{T}^1\mathbf{T}^2\cdots\mathbf{T}^q\mathbf{B}) = (-1)^q D(\mathbf{B}) = \text{sgn}(\pi)D(\mathbf{B})$$

as required. □

Third Additional Implication: Proof

In case $D(\mathbf{I}_n) = 1$, applying parts 1 and 2 of the Lemma (which we have already proved) with $\mathbf{B} = \mathbf{I_n}$ gives immediately

$$D(\mathbf{A}) = \sum_{\pi \in \Pi} \prod_{i=1}^{n} a_{i\pi(i)} D(\mathbf{P}^\pi) = \sum_{\pi \in \Pi} \operatorname{sgn}(\pi) \prod_{i=1}^{n} a_{i\pi(i)}$$

But then, applying parts 1 and 2 of the Lemma for a general matrix $\mathbf{B}$ gives

$$\begin{aligned} D(\mathbf{AB}) &= \sum_{\pi \in \Pi} \prod_{i=1}^{n} a_{i\pi(i)} D(\mathbf{P}^\pi \mathbf{B}) \\ &= \sum_{\pi \in \Pi} \operatorname{sgn}(\pi) \prod_{i=1}^{n} a_{i\pi(i)} D(\mathbf{B}) = D(\mathbf{A}) D(\mathbf{B}) \end{aligned}$$

as an implication of the first equality on this slide.

This completes the proof of all three parts. □

Formal Definition and Cofactor Expansion

Definition

The determinant $|\mathbf{A}|$ of any $n \times n$ matrix $\mathbf{A}$ is defined so that $\mathcal{D}_n \ni \mathbf{A} \mapsto |\mathbf{A}|$ is the unique (row) multilinear and alternating mapping that satisfies $|\mathbf{I}_n| = 1$.

Definition

For any $n \times n$ determinant $|\mathbf{A}|$, its rs-cofactor $|\mathbf{C}_{rs}|$ is the $(n-1) \times (n-1)$ determinant of the matrix $\mathbf{C}_{rs}$ obtained by omitting row r and column s from $\mathbf{A}$.
The cofactor expansion of $|\mathbf{A}|$ along any row r or column s is

$$|\mathbf{A}| = \sum\nolimits_{j=1}^{n} (-1)^{r+j} a_{rj} |\mathbf{C}_{rj}| = \sum\nolimits_{i=1}^{n} (-1)^{i+s} a_{is} |\mathbf{C}_{is}|$$

Exercise

Prove that these cofactor expansions are valid, using the formula

$$|\mathbf{A}| = \sum\nolimits_{\pi \in \Pi} \prod\nolimits_{i=1}^{n} \operatorname{sgn}(\pi) a_{i\pi(i)}$$

Outline

Eight Basic Rules (Rules A–H of EMEA, Section 16.4)

Let $|\mathbf{A}|$ denote the determinant of any $n \times n$ matrix $\mathbf{A}$.

1. $|\mathbf{A}| = 0$ if all the elements in a row (or column) of $\mathbf{A}$ are 0.
2. $|\mathbf{A}^\top| = |\mathbf{A}|$, where $\mathbf{A}^\top$ is the transpose of $\mathbf{A}$.
3. If all the elements in a single row (or column) of $\mathbf{A}$ are multiplied by a scalar α, so is its determinant.
4. If two rows (or two columns) of $\mathbf{A}$ are interchanged, the determinant changes sign, but not its absolute value.
5. If two of the rows (or columns) of $\mathbf{A}$ are proportional, then $|\mathbf{A}| = 0$.
6. The value of the determinant of $\mathbf{A}$ is unchanged if any multiple of one row (or one column) is added to a different row (or column) of $\mathbf{A}$.
7. The determinant of the product $|\mathbf{AB}|$ of two $n \times n$ matrices equals the product $|\mathbf{A}| \cdot |\mathbf{B}|$ of their determinants.
8. If α is any scalar, then $|\alpha \mathbf{A}| = \alpha^n |\mathbf{A}|$.

The Transpose Rule 2: Verification

The transpose rule 2 is key: for any statement
about how $|\mathbf{A}|$ depends on the rows of $\mathbf{A}$,
there is an equivalent statement
about how $|\mathbf{A}|$ depends on the columns of $\mathbf{A}$.

Exercise

Verify Rule 2 directly for 2×2 and then for 3×3 matrices.

Proof of Rule 2 The expansion formula implies that

$$|\mathbf{A}| = \sum_{\pi \in \Pi} \operatorname{sgn}(\pi) \prod_{i=1}^{n} a_{i\pi(i)} = \sum_{\pi \in \Pi} \operatorname{sgn}(\pi) \prod_{j=1}^{n} a_{\pi^{-1}(j)j}$$

But the product rule for signs of permutations implies
that $\operatorname{sgn}(\pi) \operatorname{sgn}(\pi^{-1}) = \operatorname{sgn}(\iota) = 1$, with $\operatorname{sgn}(\pi) = \pm 1$.
Hence $\operatorname{sgn}(\pi^{-1}) = 1/\operatorname{sgn}(\pi) = \operatorname{sgn}(\pi)$.
So, because $\pi \leftrightarrow \pi^{-1}$ is a bijection,

$$|\mathbf{A}| = \sum_{\pi^{-1} \in \Pi} \operatorname{sgn}(\pi^{-1}) \prod_{j=1}^{n} a^{\top}_{j\pi^{-1}(j)} = |\mathbf{A}^{\top}|$$

after using the expansion formula with π replaced by π^{-1}. $\square$

Verification of Rule 6

Exercise

Verify Rule 6 directly for 2×2 *and then for* 3×3 *matrices.*

Proof of Rule 6 Recall the notation $\mathbf{E}_{r+\alpha q}$ for the matrix resulting from adding the multiple of α times row q of $\mathbf{I}$ to its rth row.

Recall too that $\mathbf{E}_{r+\alpha q}\mathbf{A}$ is the matrix that results from applying the same row operation to the matrix $\mathbf{A}$.

Finally, recall the formula $|\mathbf{A}| = \sum_{j=1}^{n} a_{rj}|\mathbf{C}_{rj}|$ for the cofactor expansion of $|\mathbf{A}|$ along the rth row.

The corresponding cofactor expansion of $\mathbf{E}_{r+\alpha q}\mathbf{A}$ is then

$$|\mathbf{E}_{r+\alpha q}\mathbf{A}| = \sum_{j=1}^{n}(a_{rj} + \alpha a_{qj})|\mathbf{C}_{rj}| = |\mathbf{A}| + \alpha|\mathbf{B}|$$

where $\mathbf{B}$ is derived from $\mathbf{A}$ by replacing row r with row q.

Unless $q = r$, the matrix $\mathbf{B}$ will have its qth row repeated, implying that $|\mathbf{B}| = 0$ because the determinant is alternating.

So $q \neq r$ implies $|\mathbf{E}_{r+\alpha q}\mathbf{A}| = |\mathbf{A}|$ for all α, which is Rule 6. □

Verification of the Other Rules

Apart from Rules 2 and 6,
note that we have already proved the product Rule 7,
whereas the interchange Rule 4 just restates alternation.

Now that we have proved Rule 2,
note that Rules 1 and 3 follow from multilinearity,
applied in the special case when one row of the matrix
is multiplied by a scalar.

Also, the proportionality Rule 5 follows
from combining Rule 4 with multilinearity.

Finally, Rule 8, concerning the effect of multiplying
all elements of a matrix by the same scalar, is easily checked
because the expansion of $|\mathbf{A}|$ is the sum of many terms,
each of which involves the product of exactly n elements of $\mathbf{A}$.

Outline

Expansion by Alien Cofactors

Expanding along either row r or column s gives

$$|\mathbf{A}| = \sum_{j=1}^{n} a_{rj}|\mathbf{C}_{rj}| = \sum_{i=1}^{n} a_{is}|\mathbf{C}_{is}|$$

when one uses matching cofactors.

Expanding by alien cofactors, however,
from either the wrong row $i \neq r$
or the wrong column $j \neq s$, gives

$$0 = \sum_{j=1}^{n} a_{rj}|\mathbf{C}_{ij}| = \sum_{i=1}^{n} a_{is}|\mathbf{C}_{ij}|$$

This is because the answer will be the determinant
of an alternative matrix in which:

- either row i has been duplicated and put in row r;
- or column j has been duplicated and put in column s.

The Adjugate Matrix

Definition

The adjugate (or "(classical) adjoint") $\mathbf{adj}\,\mathbf{A}$ of an order n square matrix $\mathbf{A}$ has elements given by $(\mathbf{adj}\,\mathbf{A})_{ij} = |\mathbf{C}_{ji}|$.

It is therefore the transpose of the cofactor matrix $\mathbf{C}^+$ whose elements are the respective cofactors of $\mathbf{A}$.

Main Property of the Adjugate Matrix

Theorem

$(\mathbf{adj}\,\mathbf{A})\mathbf{A} = \mathbf{A}(\mathbf{adj}\,\mathbf{A}) = |\mathbf{A}|\mathbf{I}_n$ *for every* $n \times n$ *square matrix* $\mathbf{A}$.

Proof.

The (i, j) elements of the two product matrices are

$$[(\mathbf{adj}\,\mathbf{A})\mathbf{A}]_{ij} = \sum_{k=1}^{n} |\mathbf{C}_{ki}| a_{kj} \text{ and } [\mathbf{A}(\mathbf{adj}\,\mathbf{A})]_{ij} = \sum_{k=1}^{n} a_{ik} |\mathbf{C}_{jk}|$$

These are expansions by:

- alien cofactors in case $i \neq j$, implying that they equal 0;
- matching cofactors in case $i = j$, implying that they equal $|\mathbf{A}|$.

Hence $[(\mathbf{adj}\,\mathbf{A})\mathbf{A}]_{ij} = [\mathbf{A}(\mathbf{adj}\,\mathbf{A})]_{ij} = |\mathbf{A}|(\mathbf{I}_n)_{ij}$ for each pair (i, j).

□

Outline

Minor Determinants: Definition

Definition

Given any $m \times n$ matrix $\mathbf{A}$, a minor determinant of order k is the determinant $|\mathbf{A}_{i_1 i_2 \ldots i_k, j_1 j_2 \ldots j_k}|$ of a $k \times k$ submatrix (a_{ij}), with $1 \leq i_1 < i_2 < \ldots < i_k \leq m$ and $1 \leq j_1 < j_2 < \ldots < j_k \leq n$, that is formed by selecting all the elements that lie both:

- in one of the chosen rows i_r $(r = 1, 2, \ldots, k)$;
- in one of the chosen columns j_s $(s = 1, 2, \ldots, k)$.

Example

1. In case $\mathbf{A}$ is an $n \times n$ matrix:
   - the whole determinant $|\mathbf{A}|$ is the only minor of order n;
   - each of the n^2 cofactors $\mathbf{C}_{ij}$ is a minor of order $n - 1$;
2. In case $\mathbf{A}$ is an $m \times n$ matrix:
   - each element of the mn elements of the matrix is a minor of order 1;
   - there are $\binom{m}{k} \cdot \binom{n}{k}$ minors of order k.

Principal and Leading Principal Minors

Exercise

Verify that the set of elements that make up the minor $|\mathbf{A}_{i_1 i_2 \ldots i_k, j_1 j_2 \ldots j_k}|$ *of order* k *is completely determined by its* k *diagonal elements* a_{i_h, j_h} $(h = 1, 2, \ldots, k)$.

Definition

If $\mathbf{A}$ is an $n \times n$ matrix,
the minor $|\mathbf{A}_{i_1 i_2 \ldots i_k, j_1 j_2 \ldots j_k}|$ of order k is:

- a principal minor if all its diagonal elements are diagonal elements of $\mathbf{A}$;
- a leading principal minor if its diagonal elements are a_{hh} $(h = 1, 2, \ldots, k)$.

Exercise

Explain why an $n \times n$ *determinant has* $2^n - 1$ *principal minors.*

Outline

Definition of Inverse Matrix

Exercise

Suppose that $\mathbf{A}$ *is any "invertible"* $n \times n$ *matrix for which there exist* $n \times n$ *matrices* $\mathbf{B}$ *and* $\mathbf{C}$ *such that* $\mathbf{AB} = \mathbf{CA} = \mathbf{I}$*.*

1. *By writing* $\mathbf{CAB}$ *in two different ways, prove that* $\mathbf{B} = \mathbf{C}$*.*
2. *Use this result to show that the equal matrices* $\mathbf{B} = \mathbf{C}$*, if they exist, must be unique.*

Definition

The $n \times n$ matrix $\mathbf{X}$ is the unique inverse of the invertible $n \times n$ matrix $\mathbf{A}$ provided that $\mathbf{AX} = \mathbf{XA} = \mathbf{I}_n$.

In this case we write $\mathbf{X} = \mathbf{A}^{-1}$, so $\mathbf{A}^{-1}$ denotes the unique inverse.

Big question: does the inverse exist?

Existence Conditions

Theorem

An $n \times n$ matrix $\mathbf{A}$ has an inverse if and only if $|\mathbf{A}| \neq 0$,
which holds if and only if
at least one of the equations $\mathbf{AX} = \mathbf{I}_n$ and $\mathbf{XA} = \mathbf{I}_n$ has a solution.

Proof.

Provided $|\mathbf{A}| \neq 0$, the identity $(\textbf{adj}\,\mathbf{A})\mathbf{A} = \mathbf{A}(\textbf{adj}\,\mathbf{A}) = |\mathbf{A}|\mathbf{I}_n$
shows that the matrix $\mathbf{X} := (1/|\mathbf{A}|)\,\textbf{adj}\,\mathbf{A}$ is well defined
and satisfies $\mathbf{XA} = \mathbf{AX} = \mathbf{I}_n$, so $\mathbf{X}$ is the inverse $\mathbf{A}^{-1}$.

Conversely, if either $\mathbf{XA} = \mathbf{I}_n$ or $\mathbf{AX} = \mathbf{I}_n$ has a solution,
then the product rule for determinants implies
that $1 = |\mathbf{I}_n| = |\mathbf{AX}| = |\mathbf{XA}| = |\mathbf{A}||\mathbf{X}|$, and so $|\mathbf{A}| \neq 0$.
The rest follows from the paragraph above. □

Singularity

So $\mathbf{A}^{-1}$ exists if and only if $|\mathbf{A}| \neq 0$.

Definition

1. In case $|\mathbf{A}| = 0$,
   the matrix $\mathbf{A}$ is said to be singular;
2. In case $|\mathbf{A}| \neq 0$,
   the matrix $\mathbf{A}$ is said to be non-singular or invertible.

Example and Application to Simultaneous Equations

Exercise

Verify that

$$\mathbf{A} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \Longrightarrow \mathbf{A}^{-1} = \mathbf{C} := \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{pmatrix}$$

by using direct multiplication to show that $\mathbf{AC} = \mathbf{CA} = \mathbf{I}_2$.

Example

Suppose that a system of n simultaneous equations in n unknowns is expressed in matrix notation as $\mathbf{Ax} = \mathbf{b}$.

Of course, $\mathbf{A}$ must be an $n \times n$ matrix.

Suppose $\mathbf{A}$ has an inverse $\mathbf{A}^{-1}$.

Premultiplying both sides of the equation $\mathbf{Ax} = \mathbf{b}$ by this inverse gives $\mathbf{A}^{-1}\mathbf{Ax} = \mathbf{A}^{-1}\mathbf{b}$, which simplifies to $\mathbf{Ix} = \mathbf{A}^{-1}\mathbf{b}$.

Hence the unique solution of the equation is $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$.

Cramer's Rule: Statement

Notation

Given any $m \times n$ matrix $\mathbf{A}$,
let $[\mathbf{A}_{-j}, \mathbf{b}]$ denote the new $m \times n$ matrix
in which column j has been replaced by the column vector $\mathbf{b}$.

Evidently $[\mathbf{A}_{-j}, \mathbf{a}_j] = \mathbf{A}$.

Theorem

Provided that the $n \times n$ matrix $\mathbf{A}$ is invertible,
the simultaneous equation system $\mathbf{A}\mathbf{x} = \mathbf{b}$
has a unique solution $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$ whose ith component
is given by the ratio of determinants $x_i = |[\mathbf{A}_{-i}, \mathbf{b}]|/|\mathbf{A}|$.

This result is known as Cramer's rule.

Cramer's Rule: Proof

Proof.

Given the equation system $\mathbf{AX} = \mathbf{b}$,
each cofactor $|C_{ji}|$ of the coefficient matrix $\mathbf{A}$
is also the (j, i) cofactor of the matrix $|[\mathbf{A}_{-i}, \mathbf{b}]|$.

Expanding the determinant $|[\mathbf{A}_{-i}, \mathbf{b}]|$ by cofactors along column i
therefore gives $\sum_{j=1}^n b_j |C_{ji}| = \sum_{j=1}^n (\mathbf{adj}\,\mathbf{A})_{ij} b_j$,
by definition of the adjugate matrix.

Hence the unique solution to the equation system has components

$$x_i = (\mathbf{A}^{-1}\mathbf{b})_i = \frac{1}{|\mathbf{A}|} \sum_{j=1}^n (\mathbf{adj}\,\mathbf{A})_{ij} b_j = \frac{1}{|\mathbf{A}|} |[\mathbf{A}_{-i}, \mathbf{b}]|$$

for $i = 1, 2, \ldots, n$.

□

Rule for Inverting Products

Theorem

Suppose that $\mathbf{A}$ *and* $\mathbf{B}$ *are two invertible* $n \times n$ *matrices.*

Then the inverse of the matrix product $\mathbf{AB}$ *exists, and is the reverse product* $\mathbf{B}^{-1}\mathbf{A}^{-1}$ *of the inverses.*

Proof.

Using the associative law for matrix multiplication repeatedly gives:

$$(\mathbf{B}^{-1}\mathbf{A}^{-1})(\mathbf{AB}) = \mathbf{B}^{-1}(\mathbf{A}^{-1}\mathbf{A})\mathbf{B} = \mathbf{B}^{-1}(\mathbf{I})\mathbf{B} = \mathbf{B}^{-1}(\mathbf{IB}) = \mathbf{B}^{-1}\mathbf{B} = \mathbf{I}$$

and

$$(\mathbf{AB})(\mathbf{B}^{-1}\mathbf{A}^{-1}) = \mathbf{A}(\mathbf{BB}^{-1})\mathbf{A}^{-1} = \mathbf{A}(\mathbf{I})\mathbf{A}^{-1} = (\mathbf{AI})\mathbf{A}^{-1} = \mathbf{AA}^{-1} = \mathbf{I}.$$

These equations confirm that $\mathbf{X} := \mathbf{B}^{-1}\mathbf{A}^{-1}$ is the unique matrix satisfying the double equality $(\mathbf{AB})\mathbf{X} = \mathbf{X}(\mathbf{AB}) = \mathbf{I}$. □

Rule for Inverting Chain Products

Exercise

Prove that, if $\mathbf{A}$, $\mathbf{B}$ *and* $\mathbf{C}$ *are three invertible* $n \times n$ *matrices, then* $(\mathbf{ABC})^{-1} = \mathbf{C}^{-1}\mathbf{B}^{-1}\mathbf{A}^{-1}$.

Then use mathematical induction to extend this result in order to find the inverse of the product $\mathbf{A}_1\mathbf{A}_2 \cdots \mathbf{A}_k$ *of any finite chain of invertible* $n \times n$ *matrices.*

Matrices for Elementary Row Operations

Example

Consider the following two
out of the three possible kinds of elementary row operation:

1. of multiplying the rth row by $\alpha \in \mathbb{R}$,
   represented by the matrix $\mathbf{S}_r(\alpha)$;
2. of multiplying the qth row by $\alpha \in \mathbb{R}$,
   then adding the result to row r,
   represented by the matrix $\mathbf{E}_{r+\alpha q}$.

Exercise

Find the determinants and, when they exist, the inverses of the matrices $\mathbf{S}_r(\alpha)$ *and* $\mathbf{E}_{r+\alpha q}$.

Outline

Inverting Orthogonal Matrices

An n-dimensional square matrix $\mathbf{Q}$ is said to be orthogonal just in case its columns form an orthonormal set — i.e., they must be pairwise orthogonal unit vectors.

Theorem

A square matrix $\mathbf{Q}$ is orthogonal if and only if it satisfies $\mathbf{Q}^\top \mathbf{Q} = \mathbf{I}$.

Proof.

The elements of the matrix product $\mathbf{Q}^\top \mathbf{Q}$ satisfy

$$(\mathbf{Q}^\top \mathbf{Q})_{ij} = \sum_{k=1}^{n} q_{ik}^\top q_{kj} = \sum_{k=1}^{n} q_{ki} q_{kj} = \mathbf{q}_i \cdot \mathbf{q}_j$$

where $\mathbf{q}_i$ (resp. $\mathbf{q}_j$) denotes the ith (resp. jth) column vector of $\mathbf{Q}$.

But the columns of $\mathbf{Q}$ are orthonormal iff $\mathbf{q}_i \cdot \mathbf{q}_j = \delta_{ij}$ for all $i, j = 1, 2, \ldots, n$, and so iff $\mathbf{Q}^\top \mathbf{Q} = \mathbf{I}$. □

Exercises on Orthogonal Matrices

Exercise
Show that if the matrix $\mathbf{Q}$ *is orthogonal,*
then so is $\mathbf{Q}^\top$.

Use this result to show that a matrix is orthogonal
if and only if its row vectors also form an orthonormal set.

Exercise
Show that any permutation matrix is orthogonal.

Rotations in $\mathbb{R}^2$

Example

In $\mathbb{R}^2$, consider the anti-clockwise rotation through an angle θ of the unit circle $S^1 = \{(x_1, x_2) \in \mathbb{R}^2 \mid x_1^2 + x_2^2 = 1\}$.

It maps:

1. the first unit vector $(1, 0)$ of the canonical basis to the column vector $(\cos\theta, \sin\theta)^\top$;
2. the second unit vector $(0, 1)$ of the canonical basis to the column vector $(-\sin\theta, \cos\theta)^\top$.

So the rotation can be represented by the rotation matrix

$$\mathbf{R}_\theta := \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

with these vectors as its columns.

Rotations in $\mathbb{R}^2$ Are Orthogonal Matrices

Because $\sin(-\theta) = -\sin(\theta)$ and $\cos(-\theta) = -\cos(\theta)$, the transpose of $\mathbf{R}_\theta$ satisfies $\mathbf{R}_\theta^\top = \mathbf{R}_{-\theta}$, and so is the clockwise rotation through an angle θ of the unit circle S^1.

Since clockwise and anti-clockwise rotations are inverse operations, it is no surprise that $\mathbf{R}_\theta^\top \mathbf{R}_\theta = \mathbf{I}$.

We verify this algebraically by using matrix multiplication

$$\mathbf{R}_\theta^\top \mathbf{R}_\theta = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \mathbf{I}$$

because $\cos^2\theta + \sin^2\theta = 1$, thus verifying orthogonality.

Similarly

$$\mathbf{R}_\theta \mathbf{R}_\theta^\top = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \mathbf{I}$$

Outline

Partitioned Matrices: Definition

A partitioned matrix is a rectangular array of different matrices.

Example

Consider the $(m + \ell) \times (n + k)$ matrix

$$\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix}$$

where the four submatrices $\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}$
are of dimension $m \times n$, $m \times k$, $\ell \times n$ and $\ell \times k$ respectively.

For any scalar $\alpha \in \mathbb{R}$,
the scalar multiple of a partitioned matrix is

$$\alpha \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix} = \begin{pmatrix} \alpha\mathbf{A} & \alpha\mathbf{B} \\ \alpha\mathbf{C} & \alpha\mathbf{D} \end{pmatrix}$$

Partitioned Matrices: Addition

Suppose the two partitioned matrices

$$\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} \mathbf{E} & \mathbf{F} \\ \mathbf{G} & \mathbf{H} \end{pmatrix}$$

have the property that the following four pairs
of corresponding matrices have equal dimensions:
(i) $\mathbf{A}$ and $\mathbf{E}$; (ii) $\mathbf{B}$ and $\mathbf{F}$; (iii) $\mathbf{C}$ and $\mathbf{G}$; (iv) $\mathbf{D}$ and $\mathbf{H}$.

Then the sum of the two matrices is

$$\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix} + \begin{pmatrix} \mathbf{E} & \mathbf{F} \\ \mathbf{G} & \mathbf{H} \end{pmatrix} = \begin{pmatrix} \mathbf{A} + \mathbf{E} & \mathbf{B} + \mathbf{F} \\ \mathbf{C} + \mathbf{G} & \mathbf{D} + \mathbf{H} \end{pmatrix}$$

Partitioned Matrices: Multiplication

Provided that the two partitioned matrices

$$\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} \mathbf{E} & \mathbf{F} \\ \mathbf{G} & \mathbf{H} \end{pmatrix}$$

along with their sub-matrices are all compatible for multiplication, the product is defined as

$$\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix} \begin{pmatrix} \mathbf{E} & \mathbf{F} \\ \mathbf{G} & \mathbf{H} \end{pmatrix} = \begin{pmatrix} \mathbf{AE} + \mathbf{BG} & \mathbf{AF} + \mathbf{BH} \\ \mathbf{CE} + \mathbf{DG} & \mathbf{CF} + \mathbf{DH} \end{pmatrix}$$

This adheres to the usual rule for multiplying rows by columns.

Transposes and Some Special Matrices

The rule for transposing a partitioned matrix is

$$\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix}^\top = \begin{pmatrix} \mathbf{A}^\top & \mathbf{C}^\top \\ \mathbf{B}^\top & \mathbf{D}^\top \end{pmatrix}$$

So the original matrix is symmetric
iff $\mathbf{A} = \mathbf{A}^\top$, $\mathbf{D} = \mathbf{D}^\top$, $\mathbf{B} = \mathbf{C}^\top$, and $\mathbf{C} = \mathbf{B}^\top$.

It is diagonal iff $\mathbf{A}, \mathbf{D}$ are both diagonal,
while $\mathbf{B} = \mathbf{0}$ and $\mathbf{C} = \mathbf{0}$.

The identity matrix is diagonal with $\mathbf{A} = \mathbf{I}$, $\mathbf{D} = \mathbf{I}$,
possibly identity matrices of different dimensions.

Partitioned Matrices: Inverses, I

For an $(m+n) \times (m+n)$ partitioned matrix to have an inverse, the equation

$$\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix} \begin{pmatrix} \mathbf{E} & \mathbf{F} \\ \mathbf{G} & \mathbf{H} \end{pmatrix} = \begin{pmatrix} \mathbf{AE} + \mathbf{BG} & \mathbf{AF} + \mathbf{BH} \\ \mathbf{CE} + \mathbf{DG} & \mathbf{CF} + \mathbf{DH} \end{pmatrix} = \begin{pmatrix} \mathbf{I}_m & \mathbf{0}_{n \times m} \\ \mathbf{0}_{m \times n} & \mathbf{I}_n \end{pmatrix}$$

should have a solution for the matrices $\mathbf{E}, \mathbf{F}, \mathbf{G}, \mathbf{H}$, given $\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}$.

Assuming that $\mathbf{A}$ has an inverse, we can:

1. construct new first m equations by premultiplying the old ones by $\mathbf{A}^{-1}$;
2. construct new second n equations by:
   - premultiplying the new first m equations by $\mathbf{C}$;
   - then subtracting this product from the old second n equations.

The result is

$$\begin{pmatrix} \mathbf{I}_m & \mathbf{A}^{-1}\mathbf{B} \\ \mathbf{0} & \mathbf{D} - \mathbf{CA}^{-1}\mathbf{B} \end{pmatrix} \begin{pmatrix} \mathbf{E} & \mathbf{F} \\ \mathbf{G} & \mathbf{H} \end{pmatrix} = \begin{pmatrix} \mathbf{A}^{-1} & \mathbf{0}_{n \times m} \\ -\mathbf{CA}^{-1} & \mathbf{I}_n \end{pmatrix}$$

Partitioned Matrices: Inverses, II

To take the next step, assume the matrix $\mathbf{X} := \mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B}$ also has an inverse $\mathbf{X}^{-1} = (\mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B})^{-1}$.

Given $\begin{pmatrix} \mathbf{I}_m & \mathbf{A}^{-1}\mathbf{B} \\ \mathbf{0} & \mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B} \end{pmatrix} \begin{pmatrix} \mathbf{E} & \mathbf{F} \\ \mathbf{G} & \mathbf{H} \end{pmatrix} = \begin{pmatrix} \mathbf{A}^{-1} & \mathbf{0}_{n\times m} \\ -\mathbf{C}\mathbf{A}^{-1} & \mathbf{I}_n \end{pmatrix}$,
we can then premultiply the second n equations by $\mathbf{X}^{-1}$, then subtract $\mathbf{A}^{-1}\mathbf{B}$ times the new second n equations from the old first m equations to obtain

$$\begin{pmatrix} \mathbf{I}_m & \mathbf{0}_{n\times m} \\ \mathbf{0}_{m\times n} & \mathbf{I}_n \end{pmatrix} \begin{pmatrix} \mathbf{E} & \mathbf{F} \\ \mathbf{G} & \mathbf{H} \end{pmatrix} = \begin{pmatrix} \mathbf{E} & \mathbf{F} \\ \mathbf{G} & \mathbf{H} \end{pmatrix} = \mathbf{Z}$$

where $\mathbf{Z} := \begin{pmatrix} \mathbf{A}^{-1} + \mathbf{A}^{-1}\mathbf{B}\mathbf{X}^{-1}\mathbf{C}\mathbf{A}^{-1} & -\mathbf{A}^{-1}\mathbf{B}\mathbf{X}^{-1} \\ -\mathbf{X}^{-1}\mathbf{C}\mathbf{A}^{-1} & \mathbf{X}^{-1} \end{pmatrix}$

Exercise

Use direct multiplication twice in order to verify that

$$\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix} \mathbf{Z} = \mathbf{Z} \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix} = \begin{pmatrix} \mathbf{I}_m & \mathbf{0}_{n\times m} \\ \mathbf{0}_{m\times n} & \mathbf{I}_n \end{pmatrix}$$

Partitioned Matrices: Extension

Exercise

Suppose that the two partitioned matrices

$$\mathbf{A} = (\mathbf{A}_{ij})^{k \times \ell} \quad \textit{and} \quad \mathbf{B} = (\mathbf{B}_{ij})^{k \times \ell}$$

are both $k \times \ell$ arrays of respective $m_i \times n_j$ matrices $\mathbf{A}_{ij}, \mathbf{B}_{ij}$.

1. *Under what conditions can the product* $\mathbf{AB}$ *be defined as a $k \times \ell$ array of matrices?*
2. *Under what conditions can the product* $\mathbf{BA}$ *be defined as a $k \times \ell$ array of matrices?*
3. *When either* $\mathbf{AB}$ *or* $\mathbf{BA}$ *can be so defined, give a formula for its product, using summation notation.*
4. *Express* $\mathbf{A}^\top$ *as a partitioned matrix.*
5. *Under what conditions is the matrix* $\mathbf{A}$ *symmetric?*