Columbia University
in the City of New York

Math Methods for Political Science

Lecture 3: Matrix Algebra

Thibault Vatter <tv2233@columbia.edu>
Department of Statistics, Columbia University

September 13, 2017

Outline

Sums and scalar multiples

Definition 1 (Matrix equality, sum and scalar multiple)

Let A and B be two $n \times m$ matrices and $r \in \mathbb{R}$, then

- $A = B \iff A_{i,j} = B_{i,j} \ \forall i, j$.
- $Z = \mathbf{u} + \mathbf{v} \iff Z$ is a $n \times m$ matrix & $Z_{i,j} = A_{i,j} + B_{i,j} \ \forall i, j$.
- $Z = rA \iff Z$ is a $n \times m$ matrix & $Z_{i,j} = rA_{i,j} \ \forall i, j$.

Example 1 (Sums and scalar multiples)

Let $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$, $B = \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix}$, $C = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$ and $r = 2$, then

$A + B = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} + \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix} = \begin{bmatrix} 6 & 8 \\ 10 & 12 \end{bmatrix}$, $rA = 2 \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 2 & 4 \\ 6 & 8 \end{bmatrix}$,
and neither $A + C$ nor $B + C$ does exist.

Sums and scalar multiples cont'd

Definition 2 (Null and identity matrices)

- 0 is the $m \times n$ **null matrix** if $0_{i,j} = 0 \; \forall i, j$.
- I_n is the $n \times n$ **identity matrix** if $I_{i,j} = 1$ if $i = j$ and 0 if $i \neq j$.

Example 2 (Identity matrices)

Let $\mathbf{x} \in \mathbb{R}^2$, then

$$I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \Rightarrow I_2\mathbf{x} = x_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \mathbf{x}$$

Theorem 1 (Sums and scalar multiples)

Let A, B, C be matrices of the same size and $r, s \in \mathbb{R}$.

1. $A + B = B + A$
2. $(A + B) + C = A + (B + C)$
3. $A + 0 = A$
4. $r(A + B) = rB + rA$
5. $(r + s)A = rA + sA$
6. $r(sA) = (rs)A$

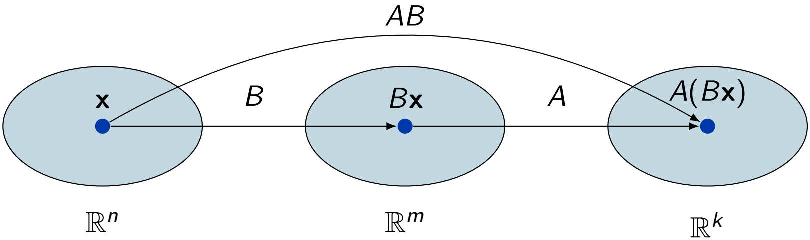
Matrix multiplication

Let A be a $k \times m$ matrix, B be a $m \times n$ matrix with columns $\mathbf{b}_1, \cdots, \mathbf{b}_n$, and $\mathbf{x} \in \mathbb{R}^n$, then

$$B\mathbf{x} = \sum_{i=1}^{n} x_i \mathbf{b}_i \in \mathbb{R}^m \text{ and } A(B\mathbf{x}) = \sum_{i=1}^{n} A(x_i \mathbf{b}_i) = \sum_{i=1}^{n} x_i A\mathbf{b}_i \in \mathbb{R}^k.$$

In other words,

$$A(B\mathbf{x}) = \sum_{i=1}^{n} x_i A\mathbf{b}_i \implies A(B\mathbf{x}) = \begin{bmatrix} A\mathbf{b}_1 & \cdots & A\mathbf{b}_n \end{bmatrix} \mathbf{x}.$$



Matrix multiplication cont'd

COLUMBIA UNIVERSITY IN THE CITY OF NEW YORK

Definition 3 (Matrix product)

Let A be a $k \times m$ matrix and B be a $m \times n$ matrix with columns $\mathbf{b}_1, \cdots, \mathbf{b}_n$, then the **product** AB is the $k \times n$ matrix whose columns are $A\mathbf{b}_1, \cdots, A\mathbf{b}_n$.

Example 3 (Matrix product)

Let $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$, then

$$A\mathbf{b}_1 = 1 \begin{bmatrix} 1 \\ 3 \end{bmatrix} + 4 \begin{bmatrix} 2 \\ 4 \end{bmatrix} = \begin{bmatrix} 9 \\ 19 \end{bmatrix}, \; A\mathbf{b}_2 = 2 \begin{bmatrix} 1 \\ 3 \end{bmatrix} + 5 \begin{bmatrix} 2 \\ 4 \end{bmatrix} = \begin{bmatrix} 12 \\ 26 \end{bmatrix},$$

$$A\mathbf{b}_3 = 3 \begin{bmatrix} 1 \\ 3 \end{bmatrix} + 6 \begin{bmatrix} 2 \\ 4 \end{bmatrix} = \begin{bmatrix} 15 \\ 33 \end{bmatrix} \Longrightarrow AB = \begin{bmatrix} 9 & 12 & 15 \\ 19 & 26 & 33 \end{bmatrix}$$

Note that $(AB)_{i,j} = \sum_{l=1}^{m} A_{i,l} B_{l,j}$ for $1 \leq i \leq k$ and $1 \leq j \leq n$.

Properties of the matrix product

Theorem 2 (Properties of the matrix product)

Let A be a $m \times n$ matrix and B, C be matrices of sizes for which the indicated sums and products are defined, then

- $A(BC) = (AB)C$ (**associative** law)
- $A(B + C) = AB + AC$ (**left distributive** law)
- $(B + C)A = BA + CA$ (**right distributive** law)
- $r(AB) = (rA)B = A(rB)\ \forall r \in \mathbb{R}$
- $I_m A = A = AI_n$ (**identity for matrix multiplication**)

Warnings about the matrix product

COLUMBIA UNIVERSITY IN THE CITY OF NEW YORK

In general

- $AB \neq BA$:
  $A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \Longrightarrow AB = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \neq BA = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$.
- $AB = AC \;\not\Longrightarrow\; B = C$:
  $A = \begin{bmatrix} 3 & -6 \\ -1 & 2 \end{bmatrix}$, $B = \begin{bmatrix} -1 & 1 \\ 3 & 4 \end{bmatrix}$,
  $C = \begin{bmatrix} -3 & -5 \\ 2 & 1 \end{bmatrix} \Longrightarrow AB = AC = \begin{bmatrix} -21 & -21 \\ 7 & 7 \end{bmatrix}$.
- $AB = 0 \;\not\Longrightarrow\; A = 0$ **or** $B = 0$:
  $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$, $B = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \Longrightarrow AB = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$

Commuting matrices and matrix power

Definition 4 (Commuting matrices and matrix power)

Let A, B be two $n \times n$ matrices:

- If $AB = BA$, then we say that A **commute** with B.
- A^k denotes the product of k copies of A.

Example 4 (Commuting matrices and matrix power)

- Any $n \times n$ matrix commutes both with itself and I_n.
- $I_n^k = I_n$ for any $k \in \mathbb{N}_+$.
- $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \implies A^2 = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 7 & 15 \\ 10 & 22 \end{bmatrix}$.

Matrix transpose

COLUMBIA UNIVERSITY
IN THE CITY OF NEW YORK

Definition 5 (Matrix transpose)

Let A be a $m \times n$ matrix, then the **matrix transpose** $A^\top$ is the $n \times m$ matrix s.t. $(A^\top)_{i,j} = A_{j,i}$.

Example 5 (Matrix transpose)

- $I_n^\top = I_n$.
- $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \implies A^\top = \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix}$.

Theorem 3 (Properties of the matrix transpose)

Let A and B be matrices whose sizes are appropriate for the following sums and products:

1. $(A^\top)^\top = A$
2. $(A + B)^\top = A^\top + B^\top$
3. $\forall r \in \mathbb{R}$, *then* $(rA)^\top = rA^\top$
4. $(AB)^\top = B^\top A^\top$

Note that 4. generalizes to $(ABC)^\top = C^\top B^\top A^\top$.

Outline

Matrix inverse

COLUMBIA UNIVERSITY IN THE CITY OF NEW YORK

Definition 6 (Matrix inverse)

A $n \times n$ matrix A is called **invertible** or **non-singular** if there exists A^{-1} called its **inverse** and s.t. $A^{-1}A = AA^{-1} = I_n$.

Example 6 (Matrix inverse)

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix},\ B = \begin{bmatrix} -2 & 1 \\ 1.5 & -0.5 \end{bmatrix} \implies AB = BA = I_2$$

$$\implies B = A^{-1} \text{ and } A = B^{-1}$$

Theorem 4 (Properties of invertible matrices)

Let A and B be a $n \times n$ invertible matrices, then

1. A^{-1} *is unique.*
2. A^{-1} *is invertible and* $(A^{-1})^{-1} = A$,
3. AB *is invertible and* $B^{-1}A^{-1}$,
4. $A^\top$ *is invertible and* $(A^\top)^{-1} = (A^{-1})^\top$.

Matrix inverse and linear systems

Theorem 5 (Matrix inverse and linear systems)

Let A be a $n \times n$ invertible matrix, then $\forall \mathbf{b} \in \mathbb{R}^n$, $A\mathbf{x} = \mathbf{b}$ has a the unique solution $\mathbf{x} = A^{-1}\mathbf{b}$.

Example 7 (Matrix inverse and linear systems)

Let $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$, then $A^{-1} = \begin{bmatrix} -2 & 1 \\ 1.5 & -0.5 \end{bmatrix}$. If $\mathbf{b} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$, then the solution to $A\mathbf{x} = \mathbf{b}$ is $\mathbf{x} = A^{-1}\mathbf{b} = \begin{bmatrix} -2 & 1 \\ 1.5 & -0.5 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 1/2 \end{bmatrix}$.

Matrix inverse and row equivalence

COLUMBIA UNIVERSITY IN THE CITY OF NEW YORK

Theorem 6 (Invertible matrices and row equivalence)

Let A be a $n \times n$ matrix, then A is invertible $\iff$ A is r.e. to I_n. Furthermore, any sequence of elementary row operations reducing A to I_n also transforms I_n into A^{-1}.

Example 8 (Matrix inverse by row reduction)

If A is invertible, then $\begin{bmatrix} A & I_n \end{bmatrix}$ is r.e. to $\begin{bmatrix} I_n & A^{-1} \end{bmatrix}$.

$$
\begin{aligned}
A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} &\implies \begin{bmatrix} A & I_n \end{bmatrix} = \begin{bmatrix} 1 & 2 & 1 & 0 \\ 3 & 4 & 0 & 1 \end{bmatrix} \\
&\implies \begin{bmatrix} 1 & 2 & 1 & 0 \\ 0 & -2 & -3 & 1 \end{bmatrix} \implies \begin{bmatrix} 1 & 0 & -2 & 1 \\ 0 & -2 & -3 & 1 \end{bmatrix} \\
&\implies \begin{bmatrix} 1 & 0 & -2 & 1 \\ 0 & 1 & 3/2 & 1/2 \end{bmatrix} \implies A^{-1} = \begin{bmatrix} -2 & 1 \\ 3/2 & 1/2 \end{bmatrix}
\end{aligned}
$$

Characterization of the matrix inverse

COLUMBIA UNIVERSITY IN THE CITY OF NEW YORK

Theorem 7 (Characterization of the matrix inverse)

Let A be a $n \times n$ matrix with columns $\mathbf{a}_1, \cdots, \mathbf{a}_n$, then the following statements are equivalent:

1. *A is invertible.*
2. *A is r.e. to I_n.*
3. *A has n pivot positions.*
4. *The only solution of $A\mathbf{x} = \mathbf{0}$ is $\mathbf{x} = \mathbf{0}$.*
5. *$A\mathbf{x} = \mathbf{b}$ has at least one solution for each $\mathbf{b} \in \mathbb{R}^n$.*
6. *$Span\{\mathbf{a}_1, \cdots, \mathbf{a}_n\} = \mathbb{R}^n$.*
7. *There exists C $n \times n$ s.t. $CA = I_n$.*
8. *There exists B $n \times n$ s.t. $AB = I_n$.*
9. *$A^\top$ is invertible.*

Outline

The determinant

Let A be a $n \times n$ matrix.

Definition 7 (Crossing out)

$A_{-i,-j}$ is the matrix obtained by **crossing out** row i and column j.

Definition 8 (Determinant)

The **determinant** of A is $\det A = \sum_{j=1}^{n} (-1)^{1+j} A_{1,j} \det A_{-1,-j}$.

Example 9 (Crossing out and determinant)

$$A = \begin{bmatrix} 1 & 5 & 0 \\ 2 & 4 & -1 \\ 0 & -2 & 0 \end{bmatrix} \implies A_{-2,-3} = \begin{bmatrix} 1 & 5 \\ 0 & -2 \end{bmatrix}$$

$$\begin{aligned} \det A &= 1 \det \begin{bmatrix} 4 & -1 \\ -2 & 0 \end{bmatrix} - 5 \det \begin{bmatrix} 2 & -1 \\ 0 & 0 \end{bmatrix} + 0 \det \begin{bmatrix} 2 & 4 \\ 0 & -2 \end{bmatrix} \\ &= 1(0-2) - 5(0-0) + 0(-4-0) = -2 \end{aligned}$$

Cofactor

COLUMBIA UNIVERSITY IN THE CITY OF NEW YORK

Let A be a $n \times n$ matrix.

Definition 9 (Cofactor)

The **(i,j)-cofactor** is $C_{i,j} = (-1)^{i+j} \det A_{-i,-j}$.

Theorem 8 (The determinant)

$\forall j$, we have $\det A = \sum_{i=1}^n A_{i,j} C_{i,j} = \sum_{i=1}^n A_{j,i} C_{j,i}$

Example 10 (Cofactor and the determinant)

$$A = \begin{bmatrix} 1 & 5 & 0 \\ 2 & 4 & -1 \\ 0 & -2 & 0 \end{bmatrix} \implies C_{3,2} = (-1)^{3+2} \det \begin{bmatrix} 1 & 0 \\ 2 & -1 \end{bmatrix} = 1$$

$$\det A = \sum_{i=1}^n A_{3,i} C_{3,i} = 0C_{3,1} - 2C_{3,2} + 0C_{3,3} = -2$$

Elementary row operations

COLUMBIA UNIVERSITY IN THE CITY OF NEW YORK

Let A be a $n \times n$ matrix.

Theorem 9 (Determinant and elementary row operations)

Let B be obtained from A by a single elementary row operation:

- *replacement* $\implies \det B = \det A$,
- *interchange* $\implies \det B = -\det A$,
- *scaling a row by k* $\implies \det B = k \det A$.

Example 11 (Determinant and elementary row operations)

$$\det \begin{bmatrix} 1 & -4 & 2 \\ -2 & 8 & -9 \\ -1 & 7 & 0 \end{bmatrix} = \det \begin{bmatrix} 1 & -4 & 2 \\ 0 & 0 & -5 \\ -1 & 7 & 0 \end{bmatrix} = \det \begin{bmatrix} 1 & -4 & 2 \\ 0 & 0 & -5 \\ 0 & 3 & 2 \end{bmatrix}$$

$$= -\det \begin{bmatrix} 1 & -4 & 2 \\ 0 & 3 & 2 \\ 0 & 0 & -5 \end{bmatrix} = 15$$

Invertible matrices

Let A be a $n \times n$ matrix.

Theorem 10 (Determinant of invertible matrices)

A is invertible $\iff \det A \neq 0$.

Example 12 (Determinant of an invertible matrix)

$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ with $a, b, c, d \in \mathbb{R}$ s.t. $ad - bc \neq 0 \implies A$ invertible.

Theorem 11 (Properties of the determinant)

- $\det A^\top = \det A$
- *If B is a $n \times n$ matrix, then* $\det AB = \det A \det B$.
- *If* $\det A \neq 0$, *then* $\det A^{-1} = 1/ \det A$.

Cramer's rule

Let A be a $n \times n$ matrix with columns $\mathbf{a}_1, \cdots, \mathbf{a}_n$, $\mathbf{b} \in \mathbb{R}^n$ and

$$A_i(\mathbf{b}) = \begin{bmatrix} \mathbf{a}_1 & \cdots & \mathbf{a}_{i-1} & \mathbf{b} & \mathbf{a}_{i+1} & \cdots & \mathbf{a}_n \end{bmatrix}$$

Theorem 12 (Cramer's rule)

If A is invertible, then the unique solution of $A\mathbf{x} = \mathbf{b}$ is s.t.

$$x_i = \det A_i(\mathbf{b}) / \det A.$$

Example 13 (Cramer's rule)

$$\begin{array}{rrcr} 3x_1 & -2x_2 & = & 6 \\ -5x_1 & +4x_2 & = & 8 \end{array} \implies A = \begin{bmatrix} 3 & -2 \\ -5 & 4 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 6 \\ 8 \end{bmatrix} \text{ and } \det A = 2$$

$$\implies A_1(\mathbf{b}) = \begin{bmatrix} 6 & -2 \\ 8 & 4 \end{bmatrix} \text{ and } A_2(\mathbf{b}) = \begin{bmatrix} 3 & 6 \\ 5 & 8 \end{bmatrix}$$

$$\implies \mathbf{x} = \frac{1}{\det A} \begin{bmatrix} \det A_1(\mathbf{b}) \\ \det A_2(\mathbf{b}) \end{bmatrix} = \begin{bmatrix} \frac{24+16}{2} \\ \frac{24+30}{2} \end{bmatrix} = \begin{bmatrix} 20 \\ 27 \end{bmatrix}$$

Adjugate matrix and the inverse

Let A be a $n \times n$ matrix.

Definition 10 (Adjugate matrix)

The **adjugate matrix** $\operatorname{adj} A$ is the $n \times n$ matrix s.t. $\operatorname{adj} A_{i,j} = C_{j,i}$.

Theorem 13 (Adjugate matrix and the inverse)

If A is invertible, then $A^{-1} = \frac{\operatorname{adj} A}{\det A}$.

Example 14 (Adjugate matrix and the inverse)

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \implies \operatorname{adj} A = \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} \text{ and } \det A = ad - bc$$

$$\implies A^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$