Computational Foundations of Cognitive Science

Lecture 10: Algebraic Properties of Matrices; Transpose; Inner and Outer Product

Frank Keller

School of Informatics
University of Edinburgh
keller@inf.ed.ac.uk

February 23, 2010

1. Properties of Matrices
   - Addition and Scalar Multiplication
   - Matrix Multiplication
   - Zero and Identity Matrix
   - Mid-lecture Problem
2. Transpose and Trace
   - Definition
   - Properties
3. Inner and Outer Product

Reading: Anton and Busby, Ch. 3.2

Addition and Scalar Multiplication

Matrix addition and scalar multiplication obey the laws familiar from the arithmetic with real numbers.

Theorem: Properties of Addition and Scalar Multiplication

If a and b are scalars, and if the sizes of the matrices A, B, and C are such that the operations can be performed, then:

- $A + B = B + A$ (cummutative law for addition)
- $A + (B + C) = (A + B) + C$ (associative law for addition)
- $(ab)A = a(bA)$
- $(a + b)A = aA + bA$
- $(a - b)A = aA - bA$
- $a(A + B) = aA + aB$
- $a(A - B) = aA - aB$

Matrix Multiplication

However, *matrix multiplication is not cummutative,* i.e., in general $AB \neq BA$. There are three possible reasons for this:

- AB is defined, but BA is not (e.g., A is 2×3, B is 3×4);
- AB and BA are both defined, but differ in size (e.g., A is 2×3, B is 3×2);
- AB and BA are both defined and of the same size, but they are different.

Example

Assume $A = \begin{bmatrix} -1 & 0 \\ 2 & 3 \end{bmatrix}$ $\quad B = \begin{bmatrix} 1 & 2 \\ 3 & 0 \end{bmatrix}$ then

$AB = \begin{bmatrix} -1 & -2 \\ 11 & 4 \end{bmatrix}$ $\quad BA = \begin{bmatrix} 3 & 6 \\ -3 & 0 \end{bmatrix}$

Matrix Multiplication

While the cummutative law is not valid for matrix multiplication, many properties of multiplication of real numbers carry over.

Theorem: Properties of Matrix Multiplication

If a is a scalar, and if the sizes of the matrices A, B, and C are such that the operations can be performed, then:

- $A(BC) = (AB)C$ (associative law for multiplication)
- $A(B + C) = AB + AC$ (left distributive law)
- $(B + C)A = BA + CA$ (right distributive law)
- $A(B - C) = AB - AC$
- $(B - C)A = BA - CA$
- $a(BC) = (aB)C = B(aC)$

Therefore, we can write $A + B + C$ and ABC without parentheses.

Zero Matrix

A matrix whose entries are all zero is called a *zero matrix.* It is denoted as 0 or $0_{n \times m}$ if the dimensions matter.

Examples

$$\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \quad \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad [0]$$

The zero matrix 0 plays a role in matrix algebra that is similar to that of 0 in the algebra of real numbers. But again, not all properties carry over.

Zero Matrix

Theorem: Properties of $\mathit{0}$

If c is a scalar, and if the sizes of the matrices are such that the operations can be performed, then:

- $A + \mathit{0} = \mathit{0} + A = A$
- $A - \mathit{0} = A$
- $A - A = A + (-A) = \mathit{0}$
- $\mathit{0}A = \mathit{0}$
- if $cA = \mathit{0}$ then $c = 0$ or $A = \mathit{0}$

However, the *cancellation law* of real numbers does not hold for matrices: if $ab = ac$ and $a \neq 0$, then not in general $b = c$.

Zero Matrix

Example

Assume $A = \begin{bmatrix} 0 & 1 \\ 0 & 2 \end{bmatrix} \quad B = \begin{bmatrix} 1 & 1 \\ 3 & 4 \end{bmatrix} \quad C = \begin{bmatrix} 2 & 5 \\ 3 & 4 \end{bmatrix}$.

It holds that $AB = AC = \begin{bmatrix} 3 & 4 \\ 6 & 8 \end{bmatrix}$.

So even though $A \neq \mathit{0}$, we can't conclude that $B = C$.

Mid-lecture Problem

We saw that if $cA = \mathit{0}$ then $c = 0$ or $A = \mathit{0}$. Does this extend to the matrix matrix product? In other words, can we conclude that if $CA = \mathit{0}$ then $C = \mathit{0}$ or $A = \mathit{0}$?

Example

Assume $C = \begin{bmatrix} 0 & 1 \\ 0 & 2 \end{bmatrix}$.

Can you come up with an $A \neq \mathit{0}$ so that $CA = \mathit{0}$?

Identity Matrix

A square matrix with ones on the main diagonal and zeros everywhere else is an *identity matrix.* It is denoted as I or I_n to indicate the size

Examples

$$[1] \quad \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

The identity matrix I plays a role in matrix algebra similar to that of 1 in the algebra of real numbers, where $a \cdot 1 = 1 \cdot a = a$.

Identity Matrix

Multiplying an matrix with I will leave that matrix unchanged.

Theorem: Identity

If A is an $n \times m$ matrix, then $AI_m = A$ and $I_n A = A$.

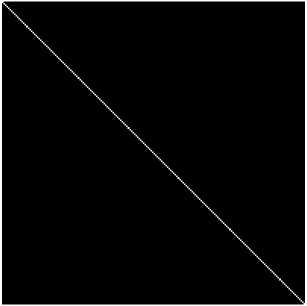
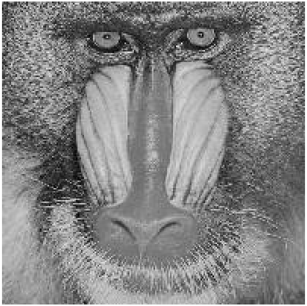
Examples

$$AI_3 = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} = A$$

$$I_2 A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} = A$$

Example: Representing Images

Assume we use matrices to represent greyscale images. If we multiply an image with I then it remains unchanged:

$255I =$  $AI =$ 

Recall that 0 is black, and 255 is white.

Definition of the Transpose

Definition: Transpose

If A is an $m \times n$ matrix, then the transpose of A, denoted by A^T, is defined to be the $n \times m$ matrix that is obtained by making the rows of A into columns: $(A)_{ij} = (A^T)_{ji}$.

Examples

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{bmatrix} \quad B = \begin{bmatrix} 2 & 3 \\ 1 & 4 \\ 5 & 6 \end{bmatrix} \quad C = \begin{bmatrix} 1 & 3 & -5 \end{bmatrix} \quad D = [4]$$

$$A^T = \begin{bmatrix} a_{11} & a_{21} & a_{31} \\ a_{12} & a_{22} & a_{32} \\ a_{13} & a_{23} & a_{33} \\ a_{14} & a_{24} & a_{34} \end{bmatrix} \quad B^T = \begin{bmatrix} 2 & 1 & 5 \\ 3 & 4 & 6 \end{bmatrix} \quad C^T = \begin{bmatrix} 1 \\ 3 \\ -5 \end{bmatrix} \quad D^T = [4]$$

Definition of the Trace

If A is a square matrix, we can obtain A^T by interchanging the entries that are symmetrically positions about the main diagonal:

$$A = \begin{bmatrix} -1 & 2 & 4 \\ 3 & 7 & 0 \\ 5 & 8 & -6 \end{bmatrix} \quad A^T = \begin{bmatrix} -1 & 3 & 5 \\ 2 & 7 & 8 \\ 4 & 0 & -6 \end{bmatrix}$$

Definition: Trace

If A is a square matrix, then the trace of A, denoted by $\mathrm{tr}(A)$, is defined to be the sum of the entries on the main diagonal of A.

$$\mathrm{tr}(A) = \mathrm{tr}(A^T) = -1 + 7 + (-6) = 0$$

Properties of the Transpose

Theorem: Properties of the Transpose

If the sizes of the matrices are such that the stated operations can be performed, then:

- $(A^T)^T = A$
- $(A + B)^T = A^T + B^T$
- $(A - B)^T = A^T - B^T$
- $(kA)^T = kA^T$
- $(AB)^T = B^T A^T$

Properties of the Trace

Theorem: Transpose and Dot Product

$A\mathbf{u} \cdot \mathbf{v} = \mathbf{u} \cdot A^T\mathbf{v}$
$\mathbf{u} \cdot A\mathbf{v} = A^T\mathbf{u} \cdot \mathbf{v}$

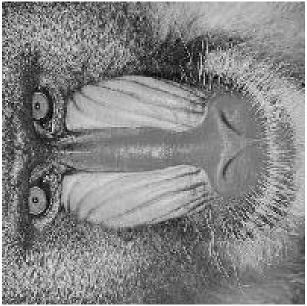
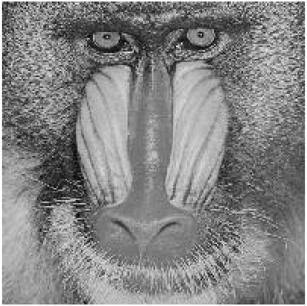
Theorem: Properties of the Trace

If A and B are square matrices of the same size, then:

- $\text{tr}(A^T) = \text{tr}(A)$
- $\text{tr}(cA) = c\ \text{tr}(A)$
- $\text{tr}(A + B) = \text{tr}(A) + \text{tr}(B)$
- $\text{tr}(A - B) = \text{tr}(A) - \text{tr}(B)$
- $\text{tr}(AB) = \text{tr}(BA)$

Example: Representing Images

We can transpose a matrix representing an image:

$A^T =$  $(A^T)^T =$ 

Inner and Outer Product

Definition: Inner and Outer Product

If $\mathbf{u}$ and $\mathbf{v}$ are column vectors with the same size, then $\mathbf{u}^T\mathbf{v}$ is the inner product of $\mathbf{u}$ and $\mathbf{v}$; if $\mathbf{u}$ and $\mathbf{v}$ are column vectors of any size, then $\mathbf{u}\mathbf{v}^T$ is the outer product of $\mathbf{u}$ and $\mathbf{v}$.

Theorem: Properties of Inner and Outer Product

$$\mathbf{u}^T\mathbf{v} = \mathrm{tr}(\mathbf{u}\mathbf{v}^T)$$
$$\mathbf{u}^T\mathbf{v} = \mathbf{u}\cdot\mathbf{v} = \mathbf{v}\cdot\mathbf{u} = \mathbf{v}^T\mathbf{u}$$
$$\mathrm{tr}(\mathbf{u}\mathbf{v}^T) = \mathrm{tr}(\mathbf{v}\mathbf{u}^T) = \mathbf{u}\cdot\mathbf{v}$$

Inner and Outer Product

Examples

$$\mathbf{u} = \begin{bmatrix} -1 \\ 3 \end{bmatrix} \quad \mathbf{v} = \begin{bmatrix} 2 \\ 5 \end{bmatrix} \quad \mathbf{u}^T\mathbf{v} = \begin{bmatrix} -1 & 3 \end{bmatrix} \begin{bmatrix} 2 \\ 5 \end{bmatrix} = -1 \cdot 2 + 3 \cdot 5 = [13] = 13$$

$$\mathbf{u}\mathbf{v}^T = \begin{bmatrix} -1 \\ 3 \end{bmatrix} \begin{bmatrix} 2 & 5 \end{bmatrix} = \begin{bmatrix} -1 \cdot 2 & -1 \cdot 5 \\ 3 \cdot 2 & 3 \cdot 5 \end{bmatrix} = \begin{bmatrix} -2 & -5 \\ 6 & 15 \end{bmatrix}$$

$$\mathbf{u}^T\mathbf{v} = \begin{bmatrix} u_1 & u_2 & \cdots & u_n \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = [u_1 v_1 + u_2 v_2 + \cdots + u_n v_n] = \mathbf{u} \cdot \mathbf{v}$$

$$\mathbf{u}\mathbf{v}^T = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} \begin{bmatrix} v_1 & v_2 & \cdots & v_n \end{bmatrix} = \begin{bmatrix} u_1 v_1 & u_1 v_2 & \cdots & u_1 v_n \\ u_2 v_1 & u_2 v_2 & \cdots & u_2 v_n \\ \vdots & \vdots & & \vdots \\ u_n v_1 & u_n v_2 & \cdots & u_n v_n \end{bmatrix}$$

Summary

- Matrix addition and scalar multiplication are cummutative, associative, and distributive;
- matrix multiplication is associative and distributive, but not cummutative: $AB \neq BA$;
- the zero matrix *0* consists of only zeros, the identity matrix I consists of ones on the diagonal and zeros everywhere else;
- transpose A^T: $(A)_{ij} = (A^T)_{ji}$;
- trace $\mathrm{tr}(A)$: sum of the entries on the main diagonal;
- the trace and the transpose are distributive;
- inner product: $\mathbf{u}^T\mathbf{v}$;
- outer product: $\mathbf{u}\mathbf{v}^T$.