Computational Foundations of Cognitive Science

Lecture 9: Basic Operations on Matrices; Matrix Product

Frank Keller

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

February 23, 2010

Reading: Anton and Busby, Ch. 3.1

Matrix Notation

A *matrix* is a rectangular array of *entries*. An $m \times n$ matrix has m rows and n columns.

Examples

$$\begin{bmatrix} 1 & 2 \\ 3 & 0 \\ -1 & 4 \end{bmatrix} \quad \begin{bmatrix} 1 & 2 & -1 \\ 3 & 0 & 4 \end{bmatrix} \quad \begin{bmatrix} 2 & 1 & 0 & 3 \end{bmatrix} \quad \begin{bmatrix} 1 \\ 3 \end{bmatrix}$$

Capital letters such as A to denote matrices, lowercase letters such as a_{12} denote entries:

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$ abbreviated as $[a_{ij}]_{m \times n}$ or just $[a_{ij}]$

Matrix Notation

An $n \times n$ matrix is called a *square matrix.* The entries $a_{11}, a_{22}, \ldots, a_{nn}$ are the *main diagonal* of the matrix. $(A)_{ij}$ denotes the entries in row i and column j of matrix A.

Example

If $A = \begin{bmatrix} 3 & -3 \\ 7 & 0 \end{bmatrix}$ then $(A)_{11} = 3$, $(A)_{12} = -3$, $(A)_{21} = 7$, and $(A)_{22} = 0$.

Two matrices are equal if they have the same size and their corresponding entries are equal.

Definition: Equality

If $A = [a_{ij}]$ and $B = [b_{ij}]$ have the same size, then $A = B$ iff $(A)_{ij} = (B)_{ij}$ (or equivalently $a_{ij} = b_{ij}$), for all i and j.

Sum and Difference

For matrices of the same size, $A + B$ and $A - B$ can be obtained by adding/subtracting the corresponding entries of A and B.

Definition: Sum and Difference

If $A = [a_{ij}]$ and $B = [b_{ij}]$ have the same size, then
$(A + B)_{ij} = (A)_{ij} + (B)_{ij} = a_{ij} + b_{ij}$ and
$(A - B)_{ij} = (A)_{ij} - (B)_{ij} = a_{ij} - b_{ij}$.

Example

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

Product with Scalar

The product of a matrix A and a scalar c is obtained by multiplying each entry of A with c.

Definition: Product with a Scalar

If $A = [a_{ij}]$ and c is a scalar, then $(cA)_{ij} = c(A)_{ij} = ca_{ij}$.

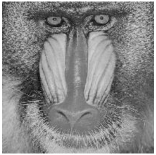
Examples

$$2\begin{bmatrix} 2 & 3 & 4 \\ 1 & 3 & 1 \end{bmatrix} = \begin{bmatrix} 4 & 6 & 8 \\ 2 & 6 & 2 \end{bmatrix} \qquad \frac{1}{3}\begin{bmatrix} 9 & -6 & 3 \\ 3 & 0 & 12 \end{bmatrix} = \begin{bmatrix} 3 & -2 & 1 \\ 1 & 0 & 4 \end{bmatrix}$$

Example: Representing Images

A greyscale image can be represented as a matrix of integers, each of which represents a shade of grey from 0 (black) to 255 (white).

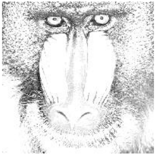
$$A = \begin{bmatrix} 106 & 147 & 145 & \cdots & 153 \\ 94 & 114 & 112 & \cdots & 98 \\ 90 & 107 & 106 & \cdots & 106 \\ \vdots & \vdots & \vdots & & \vdots \\ 117 & 112 & 148 & \cdots & 129 \end{bmatrix} =$$



Example: Representing Images

We can change the brightness of an image by multiplying its matrix with a scalar:

$2A =$



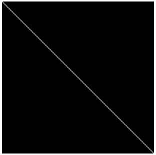
$\frac{1}{2}A =$

How do we get the inverse of an image?

Example: Representing Images

A matrix which has 1 on its diagonal and 0 everywhere else is called an *identity matrix.*

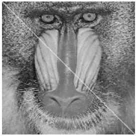
$$I = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix} \quad 255I =$$



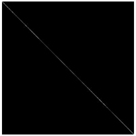
Example: Representing Images

We can add and subtract the matrices of two images:

$A + 255I =$



$255I - A =$



What happened for $255I - A$?

Row and Column Vectors

A matrix can be portioned into *row vectors* or *column vectors.* We use $\mathbf{r}_i(A)$ to denote the i-th row vector and $\mathbf{c}_j(A)$ to denote the j-th column vector of matrix A.

Examples

If $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{bmatrix}$, then $\mathbf{c}_3(A) = \begin{bmatrix} a_{13} \\ a_{23} \\ a_{33} \end{bmatrix}$ and $\mathbf{r}_1(A) = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \end{bmatrix}$ and $\mathbf{r}_2(A) = \begin{bmatrix} a_{21} & a_{22} & a_{23} & a_{24} \end{bmatrix}$

Product with Vector

The product of a matrix A and a vector $\mathbf{x}$ is the linear combination of the column vectors of A and the entries of $\mathbf{x}$.

Definition: Product with a Vector

If A is an $m \times n$ matrix with the column vectors $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$, and $\mathbf{x}$ is an $n \times 1$ column vector then

$$A\mathbf{x} = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n$$

Example

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

Mid-lecture Problem

Assume you have a system of four linear equations, each of which has three variables.

How can you represent this system using a matrix A and a vector $\mathbf{x}$? What are the dimensions of A and $\mathbf{x}$? What does the product $A\mathbf{x}$ correspond to?

Example:

$$\begin{aligned} 4a + 2b - c &= 3 \\ -2a + b - 3c &= 0 \\ a - 5b &= 4 \\ -b + 6c &= -3 \end{aligned}$$

Product with Vector

Alternative notation for $A\mathbf{x}$ without using column vectors:

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} x_1a_{11} + x_2a_{12} + \cdots + x_na_{1n} \\ x_1a_{21} + x_2a_{22} + \cdots + x_na_{2n} \\ \vdots \\ x_1a_{m1} + x_2a_{m2} + \cdots + x_na_{mn} \end{bmatrix}$$

Theorem: Linearity Properties

If A is an $m \times n$ matrix, then the following holds for all column vectors $\mathbf{u}$ and $\mathbf{v}$ in R^n and every scalar c:

1. $A(c\mathbf{u}) = c(A\mathbf{u})$
2. $A(\mathbf{u} + \mathbf{v}) = A\mathbf{u} + A\mathbf{v}$

Matrix Product

Definition: Matrix Product

If A is an $m \times s$ matrix and B is an $s \times n$ matrix and if the column vectors of B are $\mathbf{b_1}, \mathbf{b_2}, \ldots, \mathbf{b_n}$, then the product AB is the $m \times n$ matrix $AB = \begin{bmatrix} A\mathbf{b}_1 & A\mathbf{b}_2 & \cdots & A\mathbf{b}_n \end{bmatrix}$.

Note that the number of columns of A has to be the same as the number or row of B, otherwise the product is undefined.

Example

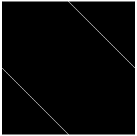
$$AB = \begin{bmatrix} 1 & 2 & 4 \\ 2 & 6 & 0 \end{bmatrix} \begin{bmatrix} 4 & 1 & 4 & 3 \\ 0 & -1 & 3 & 1 \\ 2 & 7 & 5 & 2 \end{bmatrix} = \begin{bmatrix} 12 & 27 & 30 & 13 \\ 8 & -4 & 26 & 12 \end{bmatrix}$$

Note that BA is undefined, as B is 3×4 and A is 2×3.

Example: Representing Images

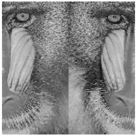
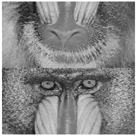
Let's assume a variant of the identity matrix with two diagonals containing ones, as in:

$$B = \begin{bmatrix} 0&0&0&0&0&1&0&0&0&0 \\ 0&0&0&0&0&0&1&0&0&0 \\ 0&0&0&0&0&0&0&1&0&0 \\ 0&0&0&0&0&0&0&0&1&0 \\ 0&0&0&0&0&0&0&0&0&1 \\ 1&0&0&0&0&0&0&0&0&0 \\ 0&1&0&0&0&0&0&0&0&0 \\ 0&0&1&0&0&0&0&0&0&0 \\ 0&0&0&1&0&0&0&0&0&0 \\ 0&0&0&0&1&0&0&0&0&0 \end{bmatrix} \quad 255B =$$



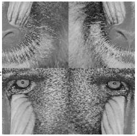
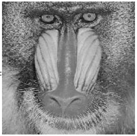
Example: Representing Images

Examples for matrix multiplication:

$AB =$  $BA =$ 

Example: Representing Images

Examples for matrix multiplication:

$BAB =$  $ABB =$ 

Row-Column Rule

Sometimes we want to compute a specific entry in a matrix product, without computing the entire column.

The entry in row i and column j of AB is the i-th row vector of A times the j-th column vector of B:

Theorem: Row-Column Rule or Dot Product Rule

$$(AB)_{ij} = \mathbf{r}_i(A)\mathbf{c}_j(B) = \mathbf{r}_i(A) \cdot \mathbf{c}_j(B) = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{is}b_{sj}$$

In the same way, the j-th column of AB is A times the j-th column of B. The i-th row of AB is the i-th row of A times B.

Theorem: Column Rule and Row Rule

$$\mathbf{c}_j(AB) = A\mathbf{c}_j(B) \qquad \mathbf{r}_i(AB) = \mathbf{r}_i(A)B$$

Summary

- Matrix addition: $(A + B)_{ij} = (A)_{ij} + (B)_{ij} = a_{ij} + b_{ij}$;
- matrix subtraction: $(A - B)_{ij} = (A)_{ij} - (B)_{ij} = a_{ij} - b_{ij}$;
- product with scalar: $(cA)_{ij} = c(A)_{ij} = ca_{ij}$;
- i-th row vector of A: $\mathbf{r}_i(A)$; j-th column vector: $\mathbf{c}_j(A)$;
- product with vector: $A\mathbf{x} = x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n$;
- matrix product: $AB = \begin{bmatrix} A\mathbf{b}_1 & A\mathbf{b}_2 & \cdots & A\mathbf{b}_n \end{bmatrix}$;
- row-column rule: $(AB)_{ij} = \mathbf{r}_i(A)\mathbf{c}_j(B)$.