Lecture 2
Matrix Operations

- transpose, sum & difference, scalar multiplication
- matrix multiplication, matrix-vector product
- matrix inverse

Matrix transpose

transpose of $m \times n$ matrix A, denoted A^T or A', is $n \times m$ matrix with

$$\left(A^T\right)_{ij} = A_{ji}$$

rows and columns of A are transposed in A^T

example: $\begin{bmatrix} 0 & 4 \\ 7 & 0 \\ 3 & 1 \end{bmatrix}^T = \begin{bmatrix} 0 & 7 & 3 \\ 4 & 0 & 1 \end{bmatrix}.$

- transpose converts row vectors to column vectors, vice versa
- $\left(A^T\right)^T = A$

Matrix addition & subtraction

if A and B are both $m \times n$, we form $A + B$ by adding corresponding entries

example: $\begin{bmatrix} 0 & 4 \\ 7 & 0 \\ 3 & 1 \end{bmatrix} + \begin{bmatrix} 1 & 2 \\ 2 & 3 \\ 0 & 4 \end{bmatrix} = \begin{bmatrix} 1 & 6 \\ 9 & 3 \\ 3 & 5 \end{bmatrix}$

can add row or column vectors same way (but never to each other!)

matrix subtraction is similar: $\begin{bmatrix} 1 & 6 \\ 9 & 3 \end{bmatrix} - I = \begin{bmatrix} 0 & 6 \\ 9 & 2 \end{bmatrix}$

(here we had to figure out that I must be 2×2)

Properties of matrix addition

- commutative: $A + B = B + A$
- associative: $(A+B)+C = A+(B+C)$, so we can write as $A+B+C$
- $A + 0 = 0 + A = A$; $A - A = 0$
- $(A + B)^T = A^T + B^T$

Scalar multiplication

we can multiply a number (a.k.a. *scalar*) by a matrix by multiplying every entry of the matrix by the scalar

this is denoted by juxtaposition or $\cdot$, with the scalar on the left:

$$(-2)\begin{bmatrix} 1 & 6 \\ 9 & 3 \\ 6 & 0 \end{bmatrix} = \begin{bmatrix} -2 & -12 \\ -18 & -6 \\ -12 & 0 \end{bmatrix}$$

(sometimes you see scalar multiplication with the scalar on the right)

- $(\alpha + \beta)A = \alpha A + \beta A$; $(\alpha\beta)A = (\alpha)(\beta A)$
- $\alpha(A + B) = \alpha A + \alpha B$
- $0 \cdot A = 0$; $1 \cdot A = A$

Matrix multiplication

if A is $m \times p$ and B is $p \times n$ we can form $C = AB$, which is $m \times n$

$$C_{ij} = \sum_{k=1}^{p} a_{ik} b_{kj} = a_{i1} b_{1j} + \cdots + a_{ip} b_{pj}, \quad i = 1, \ldots, m, \quad j = 1, \ldots, n$$

to form AB, $\#$cols of A must equal $\#$rows of B; called **compatible**

- to find i, j entry of the product $C = AB$, you need the ith row of A and the jth column of B
- form product of corresponding entries, $e.g.$, third component of ith row of A and third component of jth column of B
- add up all the products

Examples

example 1: $\begin{bmatrix} 1 & 6 \\ 9 & 3 \end{bmatrix} \begin{bmatrix} 0 & -1 \\ -1 & 2 \end{bmatrix} = \begin{bmatrix} -6 & 11 \\ -3 & -3 \end{bmatrix}$

for example, to get $1,1$ entry of product:

$$C_{11} = A_{11}B_{11} + A_{12}B_{21} = (1)(0) + (6)(-1) = -6$$

example 2: $\begin{bmatrix} 0 & -1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 6 \\ 9 & 3 \end{bmatrix} = \begin{bmatrix} -9 & -3 \\ 17 & 0 \end{bmatrix}$

these examples illustrate that matrix multiplication is not (in general) commutative: we don't (always) have $AB = BA$

Properties of matrix multiplication

- $0A = 0$, $A0 = 0$ (here 0 can be scalar, or a compatible matrix)
- $IA = A$, $AI = A$
- $(AB)C = A(BC)$, so we can write as ABC
- $\alpha(AB) = (\alpha A)B$, where α is a scalar
- $A(B + C) = AB + AC$, $(A + B)C = AC + BC$
- $(AB)^T = B^T A^T$

Matrix-vector product

very important special case of matrix multiplication: $y = Ax$

- A is an $m \times n$ matrix
- x is an n-vector
- y is an m-vector

$$y_i = A_{i1}x_1 + \cdots + A_{in}x_n, \quad i = 1, \ldots, m$$

can think of $y = Ax$ as

- a function that transforms n-vectors into m-vectors
- a set of m linear equations relating x to y

Inner product

if v is a row n-vector and w is a column n-vector, then vw makes sense, and has size 1×1, $i.e.$, is a scalar:

$$vw = v_1 w_1 + \cdots + v_n w_n$$

if x and y are n-vectors, $x^T y$ is a scalar called *inner product* or *dot product* of x, y, and denoted $\langle x, y \rangle$ or $x \cdot y$:

$$\langle x, y \rangle = x^T y = x_1 y_1 + \cdots + x_n y_n$$

(the symbol $\cdot$ can be ambiguous — it can mean dot product, or ordinary matrix product)

Matrix powers

if matrix A is square, then product AA makes sense, and is denoted A^2

more generally, k copies of A multiplied together gives A^k:

$$A^k = \underbrace{A\,A\,\cdots\,A}_{k}$$

by convention we set $A^0 = I$

(non-integer powers like $A^{1/2}$ are tricky — that's an advanced topic)

we have $A^k A^l = A^{k+l}$

Matrix inverse

if A is square, and (square) matrix F satisfies $FA = I$, then

- F is called the *inverse* of A, and is denoted A^{-1}
- the matrix A is called *invertible* or *nonsingular*

if A doesn't have an inverse, it's called *singular* or *noninvertible*

by definition, $A^{-1}A = I$; a basic result of linear algebra is that $AA^{-1} = I$

we define negative powers of A via $A^{-k} = \left(A^{-1}\right)^k$

Examples

example 1: $\begin{bmatrix} 1 & -1 \\ 1 & 2 \end{bmatrix}^{-1} = \frac{1}{3}\begin{bmatrix} 2 & 1 \\ -1 & 1 \end{bmatrix}$ (you should check this!)

example 2: $\begin{bmatrix} 1 & -1 \\ -2 & 2 \end{bmatrix}$ does not have an inverse; let's see why:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} 1 & -1 \\ -2 & 2 \end{bmatrix} = \begin{bmatrix} a-2b & -a+2b \\ c-2d & -c+2d \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

. . . but you can't have $a - 2b = 1$ and $-a + 2b = 0$

Properties of inverse

- $\left(A^{-1}\right)^{-1} = A$, *i.e.*, inverse of inverse is original matrix (assuming A is invertible)
- $(AB)^{-1} = B^{-1}A^{-1}$ (assuming A, B are invertible)
- $\left(A^T\right)^{-1} = \left(A^{-1}\right)^T$ (assuming A is invertible)
- $I^{-1} = I$
- $(\alpha A)^{-1} = (1/\alpha)A^{-1}$ (assuming A invertible, $\alpha \neq 0$)
- if $y = Ax$, where $x \in \mathbf{R}^n$ and A is invertible, then $x = A^{-1}y$:

$$A^{-1}y = A^{-1}Ax = Ix = x$$

Inverse of 2×2 matrix

it's useful to know the general formula for the inverse of a 2×2 matrix:

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

provided $ad - bc \neq 0$ (if $ad - bc = 0$, the matrix is singular)

there are similar, but much more complicated, formulas for the inverse of larger square matrices, but the formulas are rarely used