EECS 16A Designing Information Devices and Systems I
Spring 2015 Note 4

Lecture notes by Mihir Patil (01/30/2015)

Inner Products

An inner product is a procedure in which the components of vectors are used to output a single value, usually by means of a sum of products. The specific inner product that will be used in the class is known as the **dot product**, but an inner product is a more general term for an operation between two vectors that outputs a single value.

$$x^T y = \begin{bmatrix} x_1 & x_2 & \dots & x_n \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ . \\ . \\ . \\ y_3 \end{bmatrix} = x_1y_1 + x_2y_2 + \dots + x_ny_n = \sum_{i=1}^{n} x_i y_i$$

$x^T y$ is commonly denoted as $x \cdot y$, which is the dot product. Here's an example of matrix multiplication of two vectors.

$$\begin{bmatrix} -1 \\ 3.5 \\ 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} = -1 + 0 + 0 = 1$$

Orthogonal Vectors

Two vectors x, y are said to be **orthogonal** if $x^T y = 0$. In the 2-D and 3-D coordinate spaces, only perpendicular and orthogonal mean the same thing; in higher-dimension spaces, it is harder to visualize vectors being perpendicular, so the term orthogonal comes in handy to abstract away the need to visualize.

Here's an example, in $\mathbb{R}^3$. Let's say $x = \begin{bmatrix} 1 \\ 1 \\ 3 \end{bmatrix}$ and $y = \begin{bmatrix} 4 \\ -1 \\ -1 \end{bmatrix}$

Then, $x^T y = (1)(4) + (1)(-1) + (3)(1) = 0$. Thus x and y are orthogonal.

Special Vector Operations

In mathematics and computer science, vectors are widely used, but there are a few that are used quite a bit.

(a) **Unit Vector**
The unit vector is a vector that has a magnitude of 1. Here is an example operation with a unit vector.

$$\begin{bmatrix} 0 & 0 & \dots & 1 & \dots & 0 & 0 \end{bmatrix} \begin{bmatrix} x_0 \\ x_1 \\ \cdot \\ \cdot \\ x_i \\ \cdot \\ \cdot \\ x_n \end{bmatrix} = x_i$$

(b) **Sum of Components**

$$\begin{bmatrix} 1 & 1 & \dots & 1 \end{bmatrix} \begin{bmatrix} x_0 \\ x_1 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{bmatrix} = x_1 + x_2 + \dots + x_n$$

(c) **Average**

$$\begin{bmatrix} \frac{1}{n} & \frac{1}{n} & \dots & \frac{1}{n} \end{bmatrix} \begin{bmatrix} x_0 \\ x_1 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{bmatrix} = \frac{x_1 + x_2 + \dots + x_n}{n}$$

(d) **Sum of Squares**

$$\begin{bmatrix} x_0 & x_1 & \dots & x_n \end{bmatrix} \begin{bmatrix} x_0 \\ x_1 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{bmatrix} = x_1^2 + x_2^2 + \dots + x_n^2$$

(e) **Selective Sum**
Here, the values of the horizontal vector has a set of 0's and 1's, which correspond to which values of the variable matrix (the vertical one) are to be used, and which ones are to be thrown away for the sum.

$$\begin{bmatrix} 0 & 0 & 1 & 0 & 1 & ... & 0 & 1 \end{bmatrix} \begin{bmatrix} x_0 \\ x_1 \\ . \\ . \\ . \\ x_n \end{bmatrix}$$

Introduction to Norms

The **Euclidean Norm** of a vector is given by:

$$\|x\|_2 = \sqrt{x_1^2 + x_2^2 + ... + x_n^2} = \sqrt{x^T x}$$

This corresponds to the usual notion of distance in $\mathbb{R}^2$ or $\mathbb{R}^3$. It is interesting to note that the set of points with equal Euclidean norm is a circle in $\mathbb{R}^2$, or a sphere in $\mathbb{R}^3$.

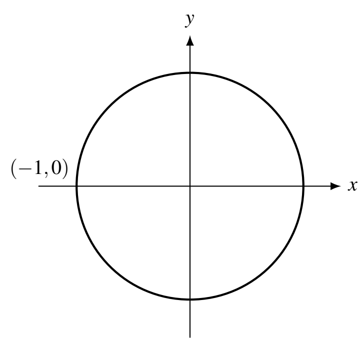

Figure 1: The unit circle.

You may have noticed that the subscript 2 in the definition of the norm given above. The subscript differentiates the Euclidean norm (or 2-norm) from other useful norms. Here are some examples of other norms.

$$\|x\|_1 = \sum_{i=1}^{n} |x_i|$$

$$\|x\|_\infty = max_{1 \le i \le n} |x_i|$$

Properties of Norms

$\|\alpha x\| = \|\alpha\| \, \|x\|$
$\|x+y\| \leq \|x\| + \|y\|$ known as the "triangle inequality"
$\|x\| \geq 0$
$\|x\| = 0$ only if $x = 0$

Introduction to Matrices

A **matrix** is a rectangular array of numbers, written as:

Each A_{ij} (where i is the row index and j is the column index) is a **component**, or **element** of the matrix A.

In the example above, A has 3 rows and 4 columns (a 3×4 matrix). If the components of A are real numbers, we say that $A \in \mathbb{R}^{m \times n}$.

Two matrices A and B are said to be equal, $A = B$, if they are of the same size and $A_{ij} = B_{ij}$ for all i, j.

Square matrices are matrices that have the same number of rows as columns ($n \times n$ matrices).

Tall matrices are matrices that have more rows than columns. ($m > n$). **Wide matrices** are matrices that have more columns than rows. ($m < n$).

Matrices have **column vectors** and **row vectors**. Column vectors are the vectors that are oriented vertically along the columns of a matrix. Likewise, row vectors are the transpose of the vectors oriented horizontally along the rows. An example is given below:

$$A = \begin{bmatrix} -1 & 3 \\ 3.5 & 2 \\ 0 & -0.1 \end{bmatrix}$$

This matrix has two column vectors:

$$a_1 = \begin{bmatrix} -1 \\ 3.5 \\ 0 \end{bmatrix} \text{ and } a_2 = \begin{bmatrix} 3 \\ 2 \\ -0.1 \end{bmatrix}$$

and three row vectors:

$$b_1 = \begin{bmatrix} -1 \\ 3 \end{bmatrix}, \; b_2 = \begin{bmatrix} 3.5 \\ 2 \end{bmatrix} \text{ and } b_3 = \begin{bmatrix} 0 \\ -0.1 \end{bmatrix}$$

Thus, A can be written as:

$$A = \begin{bmatrix} a_1 & a_2 \end{bmatrix} \text{ or } A = \begin{bmatrix} b_1^T \\ b_2^T \\ b_3^T \end{bmatrix}$$

Uses of Matrices

One use of matrices is for images. A black and white image with $m \times n$ pixels can be represented as an $m \times n$ matrix, where each component is the greyscale value corresponding to that pixel in the image. Often, however, we use vectors instead of matrices to represent images, because operations with vectors are often computationally easier than they are on matrices.

Another use of a matrix is to represent a network. Digraphs are often represented as matrices. For example, the matrix:

$$\begin{bmatrix} 0 & 1 & 2 \\ 8 & 4 & 3 \\ 0 & 0 & 0 \end{bmatrix}$$

can be used to represent connections between nodes, and their corresponding weights.

Special Matrices

A **zero matrix** is a matrix with all the components equal to zero, usually just represented as 0, where its size is implied from context.

An **identity matrix** is a square matrix whose diagonal elements are 1 and whose off-diagonal elements are all 0:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Note that the column vectors (and the transpose of the row vectors) of an $n \times n$ identity matrix are the unit vectors in $\mathbb{R}^n$.

The **transpose** of an $m \times n$ matrix A, denoted A^T is the $n \times m$ matrix given by $(A^T)_{ij} = A_{ij}$.

A square matrix is said to be **symmetric** if $A = A^T$, which means that $A_{ji} = A_{ij}$ for all i, j.

Matrix Addition

Two matrices of the same size can be added together by adding the corresponding components.

$$\begin{bmatrix} -1 & 3 \\ 3.5 & 2 \\ 0 & -0.1 \end{bmatrix} + \begin{bmatrix} 2 & -1 \\ -1 & -2 \\ 3 & 0.1 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 2.5 & 0 \\ 3 & 0 \end{bmatrix}$$

Matrix addition has the following properties:

(a) **Commutativity** $A+B=B+A$

(b) **Associativity** $(A+B)+C=A+(B+C)$

(c) **Additive Identity (Zero)** $A+0=A$

(d) **Additive Inverse** $A+(-A)=0$

Scalar Matrix Multiplication

As with vectors, multiply each component of the matrix by the scalar:

$$(-3)\begin{bmatrix} -1 & 3 \\ 3.5 & 2 \\ 0 & -0.1 \end{bmatrix} = \begin{bmatrix} 3 & -9 \\ -10.5 & -6 \\ 0 & 0.3 \end{bmatrix}$$

Additionally, we know that $-A=(-1)A$, and $(0)A=0$.

(a) **Associativity** $(\alpha\beta)A=(\alpha)(\beta A)$

(b) **Distributive** $(\alpha+\beta)=\alpha A+\beta B$

(c) **Multiplicative Identity (One)** $(1)A=A$

Matrix-Vector Multiplication

A matrix and a vector, $A \in \mathbb{R}^{m\times n}, x \in \mathbb{R}^n$, can be multiplied together as follows:

$$Ax = \begin{bmatrix} A_{11}x_1 & A_{12}x_2 & \dots & A_{1n}x_n \\ A_{21}x_1 & A_{22}x_2 & \dots & A_{2n}x_n \\ . & . & . & . \\ . & . & . & . \\ . & . & . & . \\ A_{m1}x_1 & A_{m2}x_2 & \dots & A_{mn}x_n \end{bmatrix}$$

here's an example:

$$\begin{bmatrix} -1 & 3 \\ 3.5 & 2 \\ 0 & -0.1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \\ -0.1 \end{bmatrix}$$