These are brief notes for the lecture on Monday September 6, 2010: they are not complete, but they are a guide to what I want to say today. They are not guaranteed to be correct.

2.1. Matrix Operations

Let A be an $m \times n$ matrix, that is, m rows and n columns. We'll refer to the entries of A by their row and column indices. The entry in the i^{th} row and j^{th} column is denoted by a_{ij}, and is called the (i, j)-entry of A.

$$\begin{pmatrix} a_{11} & \dots & a_{1j} & \dots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{i1} & \dots & a_{ij} & \dots & a_{in} \\ \vdots & & \vdots & & \vdots \\ a_{m1} & \dots & a_{mj} & \dots & a_{mn} \end{pmatrix}$$

The columns of A are vectors in $\mathbb{R}^m$, and are denoted in the book by $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ and in my notes by $\underline{a}_1, \underline{a}_2, \dots, \underline{a}_n$. In order to focus attention on the columns we write

$$A = [\underline{a}_1 \quad \underline{a}_2 \quad \dots \quad \underline{a}_n]$$

Just as we can add two vectors if they have the same number of rows (and since they only have one column, the same number of columns!) we can define the sum of two matrices precisely when they have the same number of rows and the same number of columns. To do so, add corresponding elements: the (i, j)-entry of $A + B$ is $a_{ij} + b_{ij}$. Focusing on the column vectors of A and B,

$$[\underline{a}_1 \quad \underline{a}_2 \quad \dots \quad \underline{a}_n] + [\underline{b}_1 \quad \underline{b}_2 \quad \dots \quad \underline{b}_n] = [\underline{a}_1 + \underline{b}_1 \quad \underline{a}_2 + \underline{b}_2 \quad \dots \quad \underline{a}_n + \underline{b}_n]$$

Example:

$$\begin{pmatrix} 1 & 2 \\ 3 & -1 \\ 2 & 4 \end{pmatrix} + \begin{pmatrix} -1 & 3 \\ 2 & -1 \\ -3 & 4 \end{pmatrix} = \begin{pmatrix} & & & \\ & & & \\ & & & \end{pmatrix}$$

We denote by 0 the matrix all of whose elements are zero.

THEOREM 1. *Let A, B, C be matrices of the same size, and let r, s be scalars. Then*

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

Note: observe how similar this is to the corresponding theorem for vector addition.

Matrix Multiplication In calculus we meet composition of functions, such as $f(g(x))$: for example if $f(y) = y^2$ and $g(x) = \sin(x)$, then $f(g(x)) = (\sin(x))^2$. The main functions we've met so far for vectors are linear transformations from $\mathbb{R}^n$ to $\mathbb{R}^m$. We saw last week that a linear transformation $T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ corresponds to multiplying a vector $\underline{x} \in \mathbb{R}^n$ by an $m \times n$ matrix A.

If we are going to compose functions, then certain things have to match up: if we are going to compute $f(g(x))$ then $g(x)$ has to be in the domain of f. Likewise, we can consider compositions of linear transformations:

$$\mathbb{R}^p \xrightarrow{U} \mathbb{R}^n \xrightarrow{T} \mathbb{R}^m$$

Here $U : \mathbb{R}^p \longrightarrow \mathbb{R}^n$, and $T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$. So, since $U(\underline{x})$ is in the domain of T, we can compute $T(U(\underline{x}))$.

We can of course write down matrices B ($n \times p$ corresponding to the transformation U) and A ($m \times n$ corresponding to the transformation T). Then $U(\underline{x}$ is equal to $B\underline{x}$, and $T(\underline{v}$ is equal to $A\underline{v}$. Replacing $\underline{v}$ by $B\underline{x}$, we obtain

$$T(U(\underline{x})) = A(B\underline{x})$$

Now,

$$B\underline{x} = x_1\underline{b}_1 + x_2\underline{b}_2 + \cdots + x_p\underline{b}_p$$

and so

$$\begin{aligned} A(B\underline{x}) &= A(x_1\underline{b}_1 + x_2\underline{b}_2 + \cdots + x_p\underline{b}_p \\ &= Ax_1\underline{b}_1 + Ax_2\underline{b}_2 + \cdots + Ax_p\underline{b}_p \\ &= x_1A\underline{b}_1 + x_2A\underline{b}_2 + \cdots + x_pA\underline{b}_p \\ &= [A\underline{b}_1 \quad A\underline{b}_2 \quad \ldots \quad A\underline{b}_p]\ \underline{x} \end{aligned}$$

That is, the composition of the transformations U followed by T corresponds to multiplication by a matrix with column form

$$[A\underline{b}_1 \quad A\underline{b}_2 \quad \ldots \quad A\underline{b}_p]$$

DEFINITION. *If A is an $m \times n$ matrix, and if B is a $n \times p$ matrix with columns $\underline{b}_1, \underline{b}_2 \ldots, \underline{b}_p$, then the matrix product AB is the $m \times p$ matrix whose columns are $A\underline{b}_1, A\underline{b}_2, \ldots, A\underline{b}_p$. That is*

$$AB = [A\underline{b}_1 \quad A\underline{b}_2 \quad \ldots \quad A\underline{b}_p]$$

Row-Column rule. If A is $m \times n$ and if B is $n \times p$ the (i, j)-entry of AB is given by

$$(AB)_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$$

Note: $\text{Row}_i(AB) = \text{Row}_i(A).B$.

Define the $m \times m$ identity matrix to be the $m \times m$ matrix with 1's down the diagonal and zeros elsewhere.

THEOREM 2. *With A, B and C appropriately sized matrices and r a scalar*

(1) $(AB)C = A(BC)$
(2) $A(B + C) = AB + AC$
(3) $(B + C)A = BA + CA$
(4) $r(AB) = (rA)B = A(rB)$
(5) $I_m A = A = A I_n$

Proof:

Powers of a square matrix

Transpose of a matrix