Matrix Algebra and Linear Transformations

September 5, 2007

Recall that a matrix is a rectangular array of numbers. More precisely, an $n \times m$ *matrix* is an array with n rows and m columns. (The number of rows is always specified first, and the number of columns second; the choice of the letters n and m is not important or standard.) Thus an $n \times m$ matrix is determined by nm numbers, called its *entries*. One denotes by A_{ij} the entry in the ith row and jth column, and by M_{nm} the set of all $n \times m$ matrices. The notation $A \in M_{nm}$ means that A belongs to this set, so that it is an $n \times m$ matrix. Recall that a *row vector* is a matrix with just one row, and a *column vector* is a matrix with just one column. The work *vector* will usually mean a column vector. For example, we will often write a row vector of length n as $\mathbf{v} = (v_1, v_2, \ldots v_n)$.

Here are the definitions of the basic matrix operations.

- Matrix addition: If $A, B \in M_{nm}$, then $A + B \in M_{nm}$ is obtained from A and B by adding the corresponding entries:

  $$(A + B)_{ij} : A_{ij} + B_{ij} \text{ for all } i, j.$$

- Multipication of a matrix by a number: If $A \in M_{nm}$ and $r \in \mathbf{R}$, then rA is obtained by multiplying all the entries of A by r:

  $$(rA)_{ij} := rA_{ij} \text{ for all } i, j.$$

- The *dot product* of two row vectors (or two column vectors) of the same length n is obtained by adding together the products of corresponding entries:

  $$\mathbf{v} \cdot \mathbf{w} := v_1 w_1 + v_2 w_2 + \cdots v_n w_n.$$

  Here the answer is a *number*, not a matrix.

- The *product* of an $n \times m$ matrix A by a column vector $\mathbf{v}$ of length m is the column vector of length m whose ith entry is computed from the ith row of A and $\mathbf{v}$ in a way quite a bit like the dot product:

$$(A\mathbf{v})_i := a_{i1}v_1 + a_{i2}v_2 + \cdots a_{im}v_m \text{ for } i = 1, \cdots n.$$

The way these operations work can by indicated symbolically as follows:

- Addition: $M_{nm} \times M_{nm} \xrightarrow{+} M_{nm}$
- Scalar multiplication: $\mathbf{R} \times M_{nm} \longrightarrow M_{nm}$
- Dot product: $M_{n1} \times M_{n1} \xrightarrow{\cdot} \mathbf{R}$
- Matrix multiplication: $M_{nm} \times M_{m1} \longrightarrow M_{n1}$

Later we will extend the definition of matrix multiplication: $M_{nm} \times M_{mp} \longrightarrow M_{np}$ It is sometimes useful to change row vectors to column vectors. If $\mathbf{v}$ is a row vector of length n. then its *transpose* $\mathbf{v}^T$ is the column vector with the same entries. More generally, if A is an $n \times m$ matrix, then the transpose A^T of A is the $m \times n$ matrix defined by changing all rows to columns:

$$(A^T)_{ij} := A_{ji}.$$

Remark: Note that numbers are not the same thing as 1×1 matrices. In particular we cannot define the product of a 1×1 matrix and a 2×2 matrix. There is also a pedantic looking distinction that the book is not careful about but which can be important. The product of a row vector of length n and a column vector of length n is a 1×1 matrix. The dot product of two column vectors of length n is a number. To relate these precisely, we could say the following: If $\mathbf{v}$ is a row vector of length n and $\mathbf{w}$ is a column vector of length n, then the matrix product $\mathbf{vw}$ is the 1×1 matrix whose only entry is the dot product $\mathbf{v}^T \cdot \mathbf{w}$.

More generally, one can combine scalar multiplication and addition into one step. This is usually expressed in the case of vectors of vectors as follows.

Definition 1 *Let $\mathbf{v}_1, \mathbf{v}_2, \ldots \mathbf{v}_m$ be a list of vectors. Then a vector $\mathbf{v}$ is said to be a linear combination of $\mathbf{v}_1, \mathbf{v}_2, \ldots \mathbf{v}_m$ if there exists a list $c_1, \ldots c_n$ of numbers such that $\mathbf{v} = c_1\mathbf{v}_1 + \cdots c_m\mathbf{v}_m$.*

It is extremely useful to know that matrix operations satisfy the following familiar looking properties: They are all easy to check.

- Associativity: $(A+B)+C = A+(B+C)$.
- Commutativity: $A+B = B+A$.
- Neutral element: $0+A = A+0 = A$, where A is the zero matrix (all entries are zero).
- Associativity: $(ab)A = a(bA)$.
- Unit element: $1A = A$.
- Distributivity: $c(A+B) = cA+cB$.
- Distributivity: $(a+b)A = aA+bA$.
- Distributivity: $A(\mathbf{v}+\mathbf{w}) = A\mathbf{v}+A\mathbf{w}$.
- Commutativity: $A(c\mathbf{v}) = c(A\mathbf{v})$.

An important consequence of the last two is the following form of the *principal of superposition*, which is often useful when solving linear equations:

Principle of superposition: If $A\mathbf{v}_1 = \mathbf{w}_1$ and $A\mathbf{v}_2 = \mathbf{w}_2$, then

$$A(a_1\mathbf{v}_1 + a_2\mathbf{v}_2) = a_1\mathbf{w}_1 + a_2\mathbf{w}_2.$$