Chapter 11

Wednesday, July 27, 2011

11.0.1 Determinants of Square Matrices

The ***determinant*** of a matrix is a number that one can associate to a square matrix that gives a lot of information about its column vectors and its invertibility. We will only give the definition of determinant for (1×1), (2×2), and (3×3)- matrices. Given a (1×1)-matrix $A = [a_{11}]$, the determinant is given by its single entry:

$$\det(A) = a_{11}.$$

Given a (2×2)-matrix

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix},$$

its determinant is given by

$$\det(A) = ad - bc.$$

Given a (3×3)-matrix

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix},$$

its determinant is given by

$$\det(A) = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32}$$

$$-a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33}.$$

The definition of the determinant of an $(n \times n)$-matrix becomes more computationally complicated as n increases. Certainly, for $n = 1, 2, 3$, the complexity of each determinant grows quickly. One of the most important properties of determinants is that they behave well with respect to matrix multiplication.

Theorem. Let A and B be two $(n \times n)$-matrices. Then

$$\det(AB) = \det(A) \cdot \det(B).$$

For certain kinds of matrices, the determinant is of a particularly easy form. Consider an $(n \times n)$-matrix A which is ***upper triangular***; these are precisely the matrices with 0's below the diagonal and arbitrary elements everywhere else. Another way of stating this is that $A = (a_{ij})$ with $a_{ij} = 0$ if $i < j$.

Theorem. Let A be an upper triangular $(n \times n)$-matrix. Then,

$$\det(A) = a_{11}a_{22}\cdots a_{nn};$$

that is, the determinant of A is given by the product of its diagonal terms.

Given a square matrix A, its transpose A^T is also square and we may thus discuss its determinant. In fact, the determinant of A and A^T are related in the best possible way:

Theorem. Let A be a square matrix, then $\det(A) = \det(A^T)$.

Combining the above two theorems tells us that the determinant of a ***lower triangular*** matrix is also the product of its diagonal elements.

11.0.2 Determinants and Invertibility

One of the most important uses of determinants is that they give a quick method for deciding if a square matrix is invertible. An $(n \times n)$-matrix A is called ***invertible*** if there exists another $(n \times n)$-matrix A^{-1} such that

$$A \cdot A^{-1} = I_n,$$

where I_n is the (square) identity matrix with 1's along the diagonal and 0's everywhere else.

Note that the identity matrix I_n is an upper triangular matrix (since there are only 0 entries below the diagonal). Thus, by our theorem above, the determinant of I_n will be the product of its diagonal entries and thus $\det(I_n) = 1 \cdot 1 \cdots 1 = 1$. Using this and the above theorem, we see that if A is an invertible, then $A \cdot A^{-1} = I_n$ and thus

$$1 = \det(I_n) = \det(A \cdot A^{-1}) = \det(A)\det(A^{-1}).$$

For this to be true, we must have that $\det(A) \neq 0$ and

$$\det(A^{-1}) = \frac{1}{\det(A)}.$$

In fact, the converse is true and we have the following very important characterization of invertibility.

Theorem. An $(n \times n)$-matrix A is invertible if and only if $\det(A) \neq 0$.

Another useful aspect of the determinant is that it detects when a set of n n-vectors is a basis. By the theorem about the number of elements in a basis, there must be n vectors; given these vectors, we can form an $(n \times n)$-matrix A with the n vectors as its columns. We have the following theorem.

Theorem. A collection of n n-vectors is a basis for $\mathbb{R}^n$ if and only if the matrix A with these vectors as its n columns is invertible. Thus, the vectors are a basis if and only if $\det(A) \neq 0$.

11.1 Eigenvalues, Eigenvectors, and Eigenspaces

Recalling that an $(n \times m)$-matrix A gives a map from $\mathbb{R}^m$ to $\mathbb{R}^n$, one becomes interested in the possibility that geometric information about the transformation A may be encoded algebraically in A. In particular, if we restrict ourselves to a *square* $(n \times n)$-matrix A, the map is now from $\mathbb{R}^n$ to itself. Of interest is when the matrix A stretches a particular $\mathbf{x} \in \mathbb{R}^n$ in some direction. If this stretch did occur, then we have algebraically that

$$A\mathbf{x} = \lambda\mathbf{x}$$

for this particular $\mathbf{x}$ and for some real number λ. When this stretching does occur, we say that λ is an ***eigenvalue*** for A and that $\mathbf{x}$ is an ***eigenvector*** with eigenvalue λ. We must also add the stipulation that $\mathbf{0}$ is never an eigenvector since, if allowed, would correspond to infinitely many eigenvalues.

One important feature of eigenvectors is that they form very important subspaces of $\mathbb{R}^n$ called *eigenspaces*. This space is defined as

$$E_\lambda(A) = \{\mathbf{x} \in \mathbb{R}^n \mid A\mathbf{x} = \lambda\mathbf{x} \;\text{ or }\; \mathbf{x} = \mathbf{0}\}.$$

This set $E_\lambda(A)$ is called the ***eigenspace*** corresponding to the eigenvalue λ. Since $\mathbf{0}$ is not an eigenvector, we must purposefully include it in our definition of the eigenspace. We wish to show that this subset $E_\lambda(A)$ forms a subspace by showing closure under scalar multiplication and vector addition. If $\mathbf{x}$ is an eigenvector with eigenvalue λ, then $\alpha\mathbf{x}$ is also an eigenvector with eigenvalue λ since

$$A(\alpha\mathbf{x}) = \alpha A\mathbf{x} = \alpha\lambda\mathbf{x} = \lambda(\alpha\mathbf{x}).$$

Similarly, if both $\mathbf{x}$ and $\mathbf{y}$ are eigenvectors with eigenvalue λ, then the sum $\mathbf{x} + \mathbf{y}$ is also an eigenvector with eigenvalue λ since

$$A(\mathbf{x} + \mathbf{y}) = A\mathbf{x} + A\mathbf{y} = \lambda\mathbf{x} + \lambda\mathbf{y} = \lambda(\mathbf{x} + \mathbf{y}).$$

11.1.1 Computing Eigenvalues and Eigenvectors

To find the eigenvalues and eigenvectors of the square matrix A, we must solve the matrix equation $A\mathbf{x} = \lambda\mathbf{x}$. By subtraction, this equation is equivalent to $A\mathbf{x} - \lambda\mathbf{x} = \mathbf{0}$. Since the identity matrix I has no effect on vectors, we may insert it to obtain $A\mathbf{x} - \lambda I\mathbf{x} = \mathbf{0}$. Factoring out the common $\mathbf{x}$, we see that for $\mathbf{x}$ to be an eigenvector of A with eigenvalue λ, the following equation must hold:

$$(A - \lambda I)\mathbf{x} = \mathbf{0}.$$

So, the matrix $A - \lambda I$ must send the non-zero vector $\mathbf{x}$ to the zero vector $\mathbf{0}$. Such a matrix $A - \lambda I$ is not invertible since it sends a non-zero vector to the zero vector, so its determinant must be zero. Thus, we must solve the polynomial equation

$$\det(A - \lambda I) = 0.$$

This determinant $\det(A - \lambda I)$ is a polynomial in the variable λ and is known as the ***characteristic polynomial of*** A.

As an example, consider the (2×2)-matrix

$$A = \begin{bmatrix} 2 & -4 \\ -1 & -1 \end{bmatrix}.$$

Its characteristic polynomial is found by computing the determinant of $A - \lambda I$:

$$\det(A - \lambda I) = \det\left(\begin{bmatrix} 2-\lambda & -4 \\ -1 & -1-\lambda \end{bmatrix}\right) = (2-\lambda)(-1-\lambda) - (-4)(-1) =$$

$$\lambda^2 - \lambda - 6 = (\lambda - 3)(\lambda + 2).$$

The roots of this characteristic polynomial are precisely the eigenvalues. Thus, A has eigenvalues -2 and 3.

The next step is to compute the eigenspace of the individual eigenvalues. If we begin with the eigenvalue $\lambda = -2$, then an eigenvector will satisfy $A\mathbf{x} = -2\mathbf{x}$. Of course, this is equivalent to $\mathbf{x}$ satisfying $(A + 2I)\mathbf{x} = \mathbf{0}$. For this equation to hold, we must have the following:

$$(A + 2I)\mathbf{x} = \begin{bmatrix} 4 & -4 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

This gives us the following system of equations:

$$\begin{array}{r} 4x_1 - 4x_2 = 0 \\ -x_1 + x_2 = 0 \end{array}.$$

Solving, we see that $x_1 = x_2$. Thus, any vector of the following form will be an eigenvector for $\lambda = -2$:

$$\begin{bmatrix} x_1 \\ x_1 \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

Thus, the eigenspace $E_{-2}(A)$ for the eigenvalue -2 is all scalar multiples of the vector $[1, 1]$. Geometrically, this is simply the line in $\mathbb{R}^2$ going through the origin and the point $(1, 1)$. Since this is an eigenspace, the matrix A transforms $\mathbb{R}^2$ by stretching this line by flipping it (since -2 is negative) and then scaling by 2.

We compute the eigenspace for $\lambda = 3$ in a similar way. The matrix equation $(A - 3I)\mathbf{x} = \mathbf{0}$ gives us equation

$$(A - 3I)\mathbf{x} = \begin{bmatrix} -1 & -4 \\ -1 & -4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Once again, this leads to the following system of equations:

$$\begin{array}{r} -x_1 - 4x_2 = 0 \\ -x_1 - 4x_2 = 0 \end{array},$$

which is, of course, redundant. This equation thus tells us that $x_1 = -4x_2$. Thus, any vector of the following form will be an eigenvector for $\lambda = 3$:

$$\begin{bmatrix} -4x_2 \\ x_2 \end{bmatrix} = x_2 \begin{bmatrix} -4 \\ 1 \end{bmatrix}.$$

Thus, the eigenspace $E_3(A)$ corresponding to the eigenvalue $\lambda = 3$ consists of all scalar multiples of the vector $[-4, 1]$. Thus, A transforms $\mathbb{R}^2$ by stretching the line spanned by $[-4, 1]$ by stretching it by a factor of 3.