Review of Basic Linear Algebra (continued)

Let us revisit matrix A $= \begin{pmatrix} 2 & 3 & 0 \\ 0 & 3 & -3 \\ -2 & 3 & 3 \end{pmatrix}$ from Class #1. Its columns are the components of the vectors $u_i = Te_i$ ($i =$ 1,2,3) images of the basis vectors (the e_i's) by the linear map T corresponding to A. I will outline the Gram-Schmidt procedure on the three u_i vectors.

First, we normalize u_1 into $v_1 = \frac{1}{\|u_1\|}u_1 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ \frac{-1}{\sqrt{2}} \end{pmatrix}$. Next, we look for a way of writing $u_2 = w_2 + \lambda v_1$ with w_2 orthogonal to v_1 and λv_1 collinear to v_1 (then we will normalize w_2). So we calculate λ so that $w_2.v_1 = 0$. This yields

$u_2.v_1 = w_2.v_1 + \lambda v_1.v_1 = 0 + \lambda = \frac{1}{\sqrt{2}} - \frac{1}{\sqrt{2}} = 0$. So, u_2 is actually orthogonal to v_1 already (this is a coincidence) and we can write $w_2 = u_2 - (u_2.v_1)v_1 = u_2$.

Now, we normalize w_2 into $v_2 = \begin{pmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{pmatrix}$. For our last step of the Gram-Schmidt process, we seek to write $u_3 = w_3 + \lambda_1 v_1 + \lambda_2 v_2$ with w_3 orthogonal to both v_1 and v_2 (and we will finally normalize w_3). So, we write similarly (knowing that $v_1.v_2 = 0$)

$$u_3.v_1 = w_3.v_1 + \lambda_1 v_1.v_1 + \lambda_2 v_2.v_1 = 0 + \lambda_1 + 0 = \frac{-3}{\sqrt{2}}$$

$$u_3.v_2 = w_3.v_2 + \lambda_1 v_1.v_2 + \lambda_2 v_2.v_2 = 0 + 0 + \lambda_2 = \frac{-3}{\sqrt{3}} + \frac{3}{\sqrt{3}} = 0$$

Thus $w_3 = u_3 - \frac{-3}{\sqrt{2}}v_1 = \begin{pmatrix} 0 + \frac{3}{2} \\ -3 \\ 3 - \frac{3}{2} \end{pmatrix} = \begin{pmatrix} \frac{3}{2} \\ -3 \\ \frac{3}{2} \end{pmatrix}$. Finally, we normalize w_3 into $v_3 = \begin{pmatrix} \frac{1}{\sqrt{6}} \\ \frac{-2}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \end{pmatrix}$. The three vectors v_1, v_2, and v_3 are therefore normal and orthogonal to each other. Clearly, this process could easily be continued if we had been working on a $n \times n$ matrix of n column vectors of n components each. The general step would read $u_i = w_i + \sum_{j=1}^{i-1}(u_i.v_j)v_j$ followed by $v_j = \frac{1}{\|w_j\|}w_j$ (provided that $w_j \neq 0$ which follows from the linear independence assumption on the u_i's).

We can now consider the matrix $B = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} \\ 0 & \frac{1}{\sqrt{3}} & \frac{-2}{\sqrt{6}} \\ \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} \end{pmatrix}$ whose columns consist of vectors that are orthonormal (to each other). Such a matrix is called *orthogonal* although it would be more appropriately called *orthonormal*. It has the interesting property that its inverse in nothing but its transpose: $B^T = B^{-1}$ (why?). This will be particularly useful when changing bases and describing the resulting change in the matrices that represents linear maps.

Determinants

Determinants can be introduced in several equivalent ways (what serves as a definition in one way then becomes a theorem in another way). One appealing and rather standard way is to discuss the geometric intuition: a 2×2 determinant represents (up to a sign) the area of a parallelogram defined by the two row vectors while a 3×3 determinant is the volume (up to a sign) of the parallelepiped built on the three row vectors.[1] We will define the determinant (detA) of a $n \times n$ matrix A as a scalar in such a way that the function det satisfies the three following *axioms*:

(I) detA is *linear* in the rows of A;

(II) if two rows of A are identical then the determinant of A is zero;

(III) the determinant of the identity is 1.

We will see in this section that there exists a (unique) function "det" that satisfies the above three axioms and we will review techniques to calculate it. To begin the investigation, we recall the elementary row operations of Class #1 and investigate the consequences of our three axioms on the determinant of the matrix A as well as that of the elementary matrices representing the row operations.

(1) Adding a row (or a multiple of it) to another does not change anything to the determinant of A by Axioms I and II combined. In particular, the corresponding elementary matrix E must have determinant 1 since it is just the identity matrix modified by the addition of a row (or multiple of it). It follows that, for such an E:

detA = det(EA) = (detE)(detA)

---

[1]One may use column instead since the determinant of the transpose of A is the same as the determinant of A. I make this choice so that the discussion connects easily with row operations.

(2) Exchanging two rows is easily seen to only change the sign of the determinant (see Exercise 1 in the homework). So, if P is the corresponding permutation matrix, we must have:

$$\det(\mathrm{PA}) = -\det\mathrm{A} = (\det\mathrm{P})(\det\mathrm{A})$$

(3) Multiplying a row (say the ith) of A by a scalar λ_i is in fact the same as multiplying A by the diagonal matrix $\Lambda_i = \begin{pmatrix} 1 & .. & 0 & .. & 0 \\ .. & .. & .. & .. & .. \\ 0 & .. & \lambda_i & .. & 0 \\ .. & .. & .. & .. & .. \\ 0 & .. & .. & .. & 1 \end{pmatrix}$. By Axioms I and III, $\det\Lambda_i = \lambda_i$ and we have:[2]

$$\det(\Lambda_i \mathrm{A}) = \lambda_i \det\mathrm{A} = (\det\Lambda_i)(\det\mathrm{A})$$

So, a picture now emerges: elementary row operations either leave the determinant unchanged, or change its sign, or multiply it by a known constant. Moreover, the determinant of the product of A by such an elementary matrix is the corresponding product of determinants. By induction, we can already write for any product of elementary matrices $\mathrm{E}_k ... \mathrm{E}_1$ representing k successive row operations:

$$\det(\mathrm{E}_k ... \mathrm{E}_1 \mathrm{A}) = (\det\mathrm{E}_k)...(\det\mathrm{E}_1)(\det\mathrm{A}) \qquad (1)$$

If we let $\mathrm{B} = \mathrm{E}_k ... \mathrm{E}_1$, we may therefore claim (in this case) that

$$\det(\mathrm{BA}) = (\det\mathrm{B})(\det\mathrm{A}) \qquad (2)$$

Now, we are left with only two cases: either A is row equivalent to the identity or it is singular. Let us begin by the latter: at one point, in the row operations, we must encounter a zero diagonal element. Indeed, the whole piece of column below this diagonal element must also be zero or we would perform an appropriate row exchange. So, leave this zero alone and move on to the next diagonal element and continue row operations, including the removal (whenever possible) of the upper triangular elements. This must eventually leave at least one identically zero row (see homework problem 2)! But a matrix with one identically zero row must have zero determinant (why?). So, a singular matrix has zero determinant. Moreover, with this understanding, equation (2) still holds. Indeed, it is also the case that

$$\det(\mathrm{BA}) = (\det\mathrm{B})(\det\mathrm{A}) = (\det\mathrm{A})(\det\mathrm{B}) = \det(\mathrm{AB}) \qquad (3)$$

[2] Note that $\lambda_i = 0$ implies $\det\Lambda_i = 0$.

Now assume that A is row equivalent to the identity. Then, the row operations can be carried out until $E_k ... E_1 A = I$ and (1), (2), and (3) still hold. Moreover, according to the LDU decomposition (with P a permutation matrix)

$$\det PA = (\det P)(\det A) = (\det L)(\det D)(\det U) = \det D \tag{4}$$

Since the determinant of a diagonal matrix is the product of diagonal elements by Axiom I, we have found detA up to a sign (detP). Finally, (3) clearly hold in all cases. So, from a computational viewpoint, if a matrix is given explicitly (by actual numbers rather than variables), finding its determinant can easily be achieved by Gaussian elimination. But in some important cases, we will need determinants involving unknowns, and that will be more easily calculated using a formula that we will now establish. All of it can be done by induction on the number n of rows and columns. We assume we have solved the problem up to $(n-1)$ and now turn to n.

For simplicity of exposition, we write the matrix $A = \begin{pmatrix} u_1 \\ .. \\ u_i \\ .. \\ u_n \end{pmatrix}$, where the u_i's are row vectors: $u_i = (a_{i1},..,a_{ij},..,a_{in})$. This vector can also be written

$u_i = a_{i1}e_1 + .. + a_{ij}e_j + .. + a_{in}e_n$. And by linearity of the determinant (with $i = 1$):

$$\det A = a_{11}\det\begin{pmatrix} e_1 \\ .. \\ u_i \\ .. \\ u_n \end{pmatrix} + .. + a_{1j}\det\begin{pmatrix} e_j \\ .. \\ u_i \\ .. \\ u_n \end{pmatrix} + .. + a_{1n}\det\begin{pmatrix} e_n \\ .. \\ u_i \\ .. \\ u_n \end{pmatrix} \tag{5}$$

Now, in the jth term above, we may permute the first j first rows until e_j lies in the jth row. In doing so, we multiply that determinant by $(-1)^{j+1}$. We then reach a jth term reading

$$(-1)^{j+1}a_{1j}\det\begin{pmatrix} u_2 \\ .. \\ e_j \\ .. \\ u_n \end{pmatrix} = (-1)^{j+1}a_{1j}\det\begin{pmatrix} B_{11} & 0 & B_{12} \\ 0 & 1 & 0 \\ B_{21} & 0 & B_{22} \end{pmatrix} = (-1)^{j+1}a_{1j}\det M_j$$

where $M_j = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}$ is the "minor" corresponding to the removal from A of the first row and the jth column. If we know how to calculate the determinant of this $(n-1)\times(n-1)$ matrix, we are done. But this can be done by the same reduction to $(n-2)$ dimensional minors. This yields the usual method of calculating determinants that is best suited for the use we will make of it in the study of eigenvalues and the characteristic polynomial.

In fact, a slight modification of the above reduction also shows that there exists a single determinant function on $n \times n$ matrices. One begins with (5) and, instead of permuting rows, one further writes for each j

$$\det\begin{pmatrix} e_j \\ u_2 \\ .. \\ .. \\ u_n \end{pmatrix} = \mathrm{a}_{21}\det\begin{pmatrix} e_j \\ e_1 \\ .. \\ .. \\ u_n \end{pmatrix} + .. + \mathrm{a}_{2k}\det\begin{pmatrix} e_j \\ e_k \\ .. \\ .. \\ u_n \end{pmatrix} + .. + \mathrm{a}_{2n}\det\begin{pmatrix} e_j \\ e_n \\ .. \\ .. \\ u_n \end{pmatrix}$$

and continues on until only e_i's appear in the determinants which are then either $0, 1$, or -1, depending on the order and repetition of the e_i's. This yields a single formula

$$\det\mathrm{A} = \sum_{\pi \in \mathrm{S}(n)} (\mathrm{sign}\pi)\mathrm{a}_{1\pi(1)}...\mathrm{a}_{n\pi(n)}$$

where $\mathrm{S}(n)$ is the set of all permutation of the first n positive integers (there are $n!$ such permutations).

Homework

1. Given the linearity axiom for determinants, prove that Axiom II is equivalent to Axiom II': when two rows are exchanged, the determinant changes sign. Hint: denote by $\det(\mathrm{u}_1,..,\mathrm{u}_i,..,\mathrm{u}_j,..,\mathrm{u}_n)$ the determinant of the row vectors $\mathrm{u}_1,..,\mathrm{u}_i,..,\mathrm{u}_j,..,\mathrm{u}_n$ and evaluate $\det(\mathrm{u}_1,..,\mathrm{u}_i - \mathrm{u}_j,..,\mathrm{u}_i - \mathrm{u}_j,..,\mathrm{u}_n)$.

2. Using as definition of a singular matrix that it is *not* nonsingular, argue that (i) A is singular if and only if it is row equivalent to a matrix that has at least one identically zero row (Hint: use the theorem of Class #2); (ii) A is singular if and only if there exists a non-zero vector x such that $\mathrm{Ax} = 0$; and (iii) for any matrix B, if A is singular, then so are AB and BA (we assume everywhere here that A and B are $n \times n$).

3. Consider elementary matrices of type E (row addition), P (row exchange) and Λ_i (multiplication of row i by λ_i). What are the effects on A of the products AE, AP, and $\mathrm{A}\Lambda_i$?