Linear Algebra (Module 110PMA207)

Department of Pure Mathematics

Lecture Notes by Lorenz Halbeisen

Chapter 5

Determinants

In the following we will define a *determinant function* on $n \times n$ matrices as a function which assigns to each real-valued $n \times n$ matrix a real number, such that the determinant function is linear as a function of each of the rows and the columns of the matrix, and its value is 0 only in the case when the row vectors (or equivalently the column vectors) are linearly dependent. If we further require that the determinant function takes the value 1 on the $n \times n$ identity matrix, then the determinant function is even unique. As we prove the uniqueness, an explicit formula for the determinant will be obtained along with many of its useful properties. So, let us start with a formal definition of a determinant function:

Definition. A **determination function** in an arbitrary n-dimensional real vector space V is a function

$$\begin{array}{rccc} \Delta : & V^n & \longrightarrow & \mathbb{R} \\ & (x_1, \ldots, x_n) & \longmapsto & \Delta(x_1, \ldots, x_n) \end{array}$$

with the following properties:

(1) $\Delta(x_1, \ldots, x_n)$ is a linear function of each argument, *i.e.*, for all i with $1 \leq i \leq n$, for all $x, y \in V$, and for all $\lambda, \mu \in \mathbb{R}$ we have

$$\begin{aligned} &\Delta(x_1, \ldots, x_{i-1}, \lambda x + \mu y, x_{i+1}, \ldots, x_n) = \\ &\quad = \lambda\Delta(x_1, \ldots, x_{i-1}, x, x_{i+1}, \ldots, x_n) + \mu\Delta(x_1, \ldots, x_{i-1}, y, x_{i+1}, \ldots, x_n) \,. \end{aligned}$$

(2) If the vectors $x_1, \ldots, x_n$ are linearly dependent, then

$$\Delta(x_1, \ldots, x_n) = 0 \,.$$

Some properties of determinant functions:

Let $\{v_1,\dots,v_n\} \subseteq V$ be a basis of V and let $x_1,\dots,x_n$ be arbitrary vectors of V. Each vector x_i $(1 \le i \le n)$ can be written in a unique way as a linear combination of the v_i's, so, for each $1 \le i \le n$ let

$$x_i = \sum_{k_i=1}^{n} \xi_i^{k_i} v_{k_i} .$$

If Δ is a determinant function, then we have the following:

$$\begin{aligned}
\Delta(x_1,\dots,x_n) &= \Delta\big(\textstyle\sum_{k_1=1}^n \xi_1^{k_1} v_{k_1}, \sum_{k_2=1}^n \xi_2^{k_2} v_{k_2}, \dots \sum_{k_n=1}^n \xi_n^{k_n} v_{k_n}\big) \\
&= \textstyle\sum_{k_1=1}^n \xi_1^{k_1} \Delta\big(v_{k_1}, \sum_{k_2=1}^n \xi_2^{k_2} v_{k_2}, \dots \sum_{k_n=1}^n \xi_n^{k_n} v_{k_n}\big) \\
&= \textstyle\sum_{k_1=1}^n \xi_1^{k_1} \cdot \sum_{k_2=1}^n \xi_2^{k_2} \Delta\big(v_{k_1}, v_{k_2}, \sum_{k_3=1}^n \xi_3^{k_3} v_{k_3}, \dots \sum_{k_n=1}^n \xi_n^{k_n} v_{k_n}\big) \\
&= \textstyle\sum_{1\le k_1,k_2\le n} \xi_1^{k_1} \cdot \xi_2^{k_2} \Delta\big(v_{k_1}, v_{k_2}, \sum_{k_3=1}^n \xi_3^{k_3} v_{k_3}, \dots \sum_{k_n=1}^n \xi_n^{k_n} v_{k_n}\big) \\
&\qquad \dots \\
&= \textstyle\sum_{1\le k_1,\dots,k_n\le n} \xi_1^{k_1} \cdot \xi_2^{k_2} \cdot \dots \cdot \xi_n^{k_n} \Delta\big(v_{k_1}, v_{k_2}, \dots, v_{k_n}\big)
\end{aligned}$$

By property (2) of determinant functions, all terms in which $k_i = k_j$ (for any distinct i and j with $1 \le i,j \le j$) are equal to 0, since the vector v_{k_i} appears twice in the sequence $(v_{k_1},\dots,v_{k_n})$ which implies that the vectors $v_{k_1},\dots,v_{k_n}$ are linearly dependent. So we just have to take care of all permutations of $\{1,\dots,n\}$. Let S_n be the set of all permutations of $\{1,\dots,n\}$, or in other words, S_n is the set of all bijections $\pi : \{1,\dots,n\} \to \{1,\dots,n\}$. This leads to the following:

$$\Delta(x_1,\dots,x_n) = \sum_{\pi\in S_n} \xi_1^{\pi(1)} \cdot \xi_2^{\pi(2)} \cdot \dots \cdot \xi_n^{\pi(n)} \Delta\big(v_{\pi(1)},\dots,v_{\pi(n)}\big) .$$

Take any i,j such that $1 \le i < j \le n$ and put $x := x_i$ and $y := x_j$. Consider for the moment the determinant function Δ just as a function of the ith and jth argument, *i.e.*,

$$\Delta(x_1,\dots,x_n) =: \Delta_{i,j}(x,y) .$$

We get

$$\begin{aligned}
0 &= \Delta_{i,j}(x+y,x+y) = \Delta_{i,j}(x,x) + \Delta_{i,j}(x,y) + \Delta_{i,j}(y,x) + \Delta_{i,j}(y,y) = \\
&= \Delta_{i,j}(x,y) + \Delta_{i,j}(y,x)
\end{aligned}$$

which implies that

$$\Delta_{i,j}(x,y) = -\Delta_{i,j}(y,x) .$$

Thus, if we swap any two vectors in $(x_1,\dots,x_n)$, we change the sign of Δ.

A permutation of $\{1, \ldots, n\}$ which just swaps two numbers is called a transposition. One can show that each permutation $\pi \in S_n$ can be written as a product of transpositions. Even though this representation is not unique, for a given $\pi \in S_n$, we need either always an even number of transpositions or always an odd number of transpositions to represent π. In the former case we call π *even* and define $\operatorname{sgn}(\pi) = 1$, in the latter case we call π *odd* and define $\operatorname{sgn}(\pi) = -1$.

Therefore we have:

$$\Delta(x_1, \ldots, x_n) = \Delta(v_1, \ldots, v_n) \cdot \sum_{\pi \in S_n} \operatorname{sgn}(\pi) \cdot \xi_1^{\pi(1)} \cdot \ldots \cdot \xi_n^{\pi(n)}$$

If $\Delta(x_1, \ldots, x_n) \neq 0$ for some fixed n-tuple $(x_1, \ldots, x_n)$, then also $\Delta(v_1, \ldots, v_n) \neq 0$ for any basis $\{v_1, \ldots, v_n\}$ of V. In particular, if $\Delta(x_1, \ldots, x_n) \neq 0$ for some n-tuple $(x_1, \ldots, x_n)$, then $\Delta(v_1, \ldots, v_n) \neq 0$ for any linearly independent vectors $v_1, \ldots, v_n$.

We conclude that any determinant function Δ is determined by the value $\Delta(v_1, \ldots, v_n)$. So, if we stipulate

$$\Delta_1(e_1, \ldots, e_n) = 1$$

then the determinant function is unique and for $x_i = (\xi_i^1, \xi_i^2, \ldots, \xi_i^n)$, where $1 \leq i \leq n$, we get

$$\Delta_1(x_1, \ldots, x_n) = \sum_{\pi \in S_n} \operatorname{sgn}(\pi) \cdot \xi_1^{\pi(1)} \cdot \ldots \cdot \xi_n^{\pi(n)}.$$

Determinants of Matrices. Let $\varphi : V \to V$ be a linear mapping from the n-dimensional vector space V into itself, let $\{v_1, \ldots, v_n\}$ be a basis of V, and let A be the matrix which corresponds to φ (with respect to the basis $v_1, \ldots, v_n$). Then

$$\Delta_\varphi(v_1, \ldots, v_n) := \frac{\Delta_1\big(\varphi(v_1), \ldots, \varphi(v_n)\big)}{\Delta_1(v_1, \ldots, v_n)}$$

is the determinant of the mapping φ and we define

$$\det(A) := \Delta_\varphi(v_1, \ldots, v_n).$$

In particular, for the standard basis $e_1, \ldots, e_n$ and for

$$A = \begin{pmatrix} \xi_1^1 & \xi_2^1 & \ldots & \xi_n^1 \\ \xi_1^2 & \xi_2^2 & \ldots & \xi_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ \xi_1^n & \xi_2^n & \ldots & \xi_n^n \end{pmatrix}$$

we get

$$\det(A) = \sum_{\pi \in S_n} \operatorname{sgn}(\pi) \cdot \xi_1^{\pi(1)} \cdot \ldots \cdot \xi_n^{\pi(n)}.$$

Computing Determinants of Matrices. In general, if we compute the determinant of an $n \times n$ matrix, we have to add and subtract $n!$ products of n numbers. For $n = 3$, one can do it by hand, but for say $n = 5$, it is very likely that we make a mistake somewhere. So, let us check if there are better methods than just using the formula given above.

First notice that for an $n \times n$ matrix $A = (a_{i,j})_{1 \le i,j \le n}$ we have

$$\det(A) = \sum_{\pi \in S_n} \operatorname{sgn}(\pi) \cdot a_{\pi(1),1} \cdot \ldots \cdot a_{\pi(n),n} = \sum_{\pi \in S_n} \operatorname{sgn}(\pi) \cdot a_{1,\pi(1)} \cdot \ldots \cdot a_{n,\pi(n)}$$

which shows that $\det(A) = \det(A^t)$. Further notice that by linearity of the determinant, no elementary row or column operation on A changes the value of $\det(A)$, and finally notice that the determinant of a triangular matrix of the form

$$\begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} & \cdots & a_{1,n} \\ 0 & a_{2,2} & a_{2,3} & \cdots & a_{2,n} \\ 0 & 0 & a_{3,3} & \cdots & a_{3,n} \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & a_{n,n} \end{pmatrix}$$

is equal to $a_{1,1} \cdot a_{2,2} \cdot \ldots \cdot a_{n,n}$. So, in order to compute the determinant of an $n \times n$ matrix A, by elementary row and column operations we can bring first the matrix A in triangular form and then just compute the product of the elements on the diagonal.

However, this procedure does not work if some of the entries of A are variables. In this case there is no other way as to compute the determinant of A by the formula given above (but at least there are some techniques which make the computation easier).

Defining Properies of Determinants. A function "det" which assigns to each $n \times n$ matrix A a real number $\det(A)$ is a determinant function if it is:

(1) Multi-linear, *i.e.*, linear as a function of *each* column and of *each* row.

(2) If two columns (or equivalently two rows) of A are equal, then $\det(A) = 0$.

(3) $\det(I_n) = 1$, where I_n denotes the $n \times n$ identity matrix.

As a consequence we get that $\det(A) = 0$ if and only if the column vectors (or equivalently the row vectors) are linearly dependent, and that if A' is a matrix obtained from A by interchanging two columns (or two rows) of A, then $\det(A') = -\det(A)$.

Chapter 6

Eigenvalues and Eigenvectors of Matrices

In the sequel let V be an n-dimensional vector space and let A be an $n \times n$ matrix.

DEFINITION. A real number λ is called an **Eigenvalue** of A if there is a *non-zero* vector $x \in V$ such that $A.x = \lambda x$, and a *non-zero* vector $x \in V$ is called an **Eigenvector** of A if there is a real number λ such that $A.x = \lambda x$.

Notice that $\lambda x = (\lambda \cdot I_n).x$, and therefore, $A.x = \lambda x$ is equivalent to $A.x = (\lambda \cdot I_n).x$, and since A is linear, this is equivalent to $(A - \lambda \cdot I_n).x = 0$, which implies that the matrix $(A - \lambda \cdot I_n)$ is not regular. So, for $n \times n$ matrices A, non-zero vectors $x \in V$ and real numbers λ, the following are equivalent:

$$A.x = \lambda x$$

$$(A - \lambda \cdot I_n).x = 0$$

$$\det(A - \lambda \cdot I_n) = 0$$

Thus, for a given matrix A, in order to find the Eigenvalues of A, we have to write $\det(A-\lambda \cdot I_n)$ as a polynomial in λ and compute its roots. This gives us the Eigenvalues of A and with these values we can compute the corresponding Eigenvectors.

As we will see later, not every matrix has real Eigenvalues, and so, not every matrix has real Eigenvectors. Further we will see that Eigenvectors which correspond to different Eigenvalues are always linearly independent. But first let us consider an example:

Let us compute the Eigenvalues and corresponding Eigenvectors of the following 3×3 matrix:

$$A = \begin{pmatrix} -1 & 0 & 2 \\ 0 & -7 & 8 \\ 4 & 0 & -3 \end{pmatrix}$$

The determinant of the matrix $(A - \lambda \cdot I_3)$ is the following polynomial in λ, which is called the **characteristic polynomial** of A:

$$\mathrm{char}_A(\lambda) = -\lambda^3 - 11\lambda^2 - 23\lambda + 35$$

The roots of this polynomial give us the Eigenvalues of A. It is not hard to see that

$$\lambda_1 = 1$$

is a root of the polynomial $\mathrm{char}_A(\lambda)$, and therefore, $(1 - \lambda)$ divides $\mathrm{char}_A(\lambda)$. Now,

$$\mathrm{char}_A(\lambda) : (1 - \lambda) = \lambda^2 + 12\lambda + 35$$

which gives us the other two roots of $\mathrm{char}_A(\lambda)$:

$$\lambda_2 = -7\,, \quad \lambda_3 = -5$$

To each of these three different Eigenvalues of A we find an Eigenvector by solving the equation $A.x_i = \lambda_i x_i$ for $i = 1, 2, 3$, and get for example

$$x_1 = (1, 1, 1)\,, \quad x_2 = (0, 1, 0)\,, \quad x_3 = (-1, 8, 2)\,.$$

Notice that not every matrix has real Eigenvalues, and thus, not every matrix has real Eigenvectors. For example the characteristic polynomial of the rotation matrix

$$\begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix}$$

is $\lambda^2 - 2\cos(\alpha)\lambda + 1$, with roots

$$\lambda_{1,2} = \frac{2\cos(\alpha) \pm \sqrt{4\cos^2(\alpha) - 4}}{2}\,.$$

Now, these roots are real numbers if and only if α is a multiple of π, and we get $\lambda_1 = \lambda_2 = \cos(\alpha)$ (which is either 1 or -1). If α is not a multiple of π, then the roots of the characteristic polynomial are complex and in this case, the matrix has no real Eigenvalue and consequently no real Eigenvector.

Let us now show that Eigenvectors which correspond to distinct Eigenvalues are linearly independent. For this assume that $\lambda_1, \ldots, \lambda_k$ are k pairwise distinct Eigenvalues of some $n \times n$ matrix A (where $n \geq k$) with corresponding Eigenvectors $x_1, \ldots, x_k$. So, for all i with $1 \leq i \leq k$ we have $A.x_i = \lambda_i x_i$.

The proof is by induction on k: For $k = 1$ the statement is obvious. So, let us assume that the statement is true for some k_0 and let $k = k_0 + 1$. Assume towards a contradiction that the x_i's are linearly dependent. Thus, there are real numbers $\xi_1, \ldots, \xi_k$ such that

$$\sum_{i=1}^{k} \xi_i x_i = 0\,.$$

Notice that by induction hypothesis, all ξ_i's are non-zero. Now $A.0 = 0$, which implies that

$$A.\Big(\sum_{i=1}^{k} \xi_i x_i\Big) = 0\,.$$

Since A is linear and the x_i's are Eigenvectors of A we get

$$A.\Big(\sum_{i=1}^{k} \xi_i x_i\Big) = \sum_{i=1}^{k} \lambda_i \xi_i x_i$$

which implies that

$$\sum_{i=1}^{k} \lambda_i \xi_i x_i = 0\,.$$

On the other hand, since $\sum_{i=1}^{k} \xi_i x_i = 0$, we have

$$\lambda_1 \cdot \sum_{i=1}^{k} \xi_i x_i = 0$$

and therefore,

$$\sum_{i=1}^{k} \lambda_i \xi_i x_i - \lambda_1 \cdot \sum_{i=1}^{k} \xi_i x_i = \sum_{i=2}^{k} (\lambda_i - \lambda_1) \xi_i x_i = 0\,.$$

In particular, since all the λ_i's are distinct and all ξ_i's are non-zero, we can write the zero-vector as an non-trivial linear combination of the vectors $x_2, \ldots, x_k$, which contradicts our induction hypothesis and completes the proof.

Chapter 7

Inner Products and Orthogonality

As usual, let $V = \mathbb{R}^n$ for some $n \geq 1$.

DEFINITION. An **inner product** on V is function which assigns to each ordered pair of vectors x, y in V a real number $\langle x, y \rangle$ in such a way that

(1) $\langle x, y \rangle = \langle y, x \rangle$,

(2) $\langle \alpha x + \beta x', y \rangle = \alpha \langle x, y \rangle + \beta \langle x', y \rangle$,

(3) If $x \neq 0$, then $\langle x, x \rangle > 0$.

Notice that by (2), $\langle x, x \rangle = 0$ if and only if $x = 0$.

Property (3) leads to the notion of **length** or **norm** of a vector x defined by

$$\|x\| := \sqrt{\langle x, x \rangle}\,.$$

Unlike determinants, inner products are by no means unique. For example we can define an inner product $\langle \cdot\,, \cdot \rangle$ on $\mathbb{R}^2$ as follows: For $x = (x_1, x_2)$ and $y = (y_1, y_2)$ let

$$\langle x, y \rangle = x_1 y_1 - x_2 y_1 - x_1 y_2 + 4 x_2 y_2\,.$$

However, there is a **standard inner product** on $\mathbb{R}^n$, called the **dot product**, defined as follows: For $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$ in $\mathbb{R}^n$ let

$$x \bullet y := x_1 y_1 + \ldots + x_n y_n\,.$$

Two vectors x and y are called **orthogonal** if $x \bullet y = 0$. Notice that the zero-vector is orthogonal to every vector. A set X of vectors is **orthonormal** if whenever both x and y are in X, then either $x \bullet y = 0$ (if $x \neq y$) or $x \bullet y = 1$ (if $x = y$).

Now, an $n \times n$ matrix is called **orthogonal** if the set of its column vectors is orthonormal. Orthogonal matrices can be characterized as follows:

(1) A is orthogonal iff the column vectors of A are orthonormal.

(2) A is orthogonal iff $A^t.A = I_n$.

(3) A is orthogonal iff $A.A^t = I_n$.

(4) A is orthogonal iff the row vectors of A are orthonormal.

(5) A is orthogonal iff for all $x \in \mathbb{R}^n$, $\|A.x\| = \|x\|$.

(6) A is orthogonal iff for all $x, y \in \mathbb{R}^n$, $A.x \bullet A.y = x \bullet y$.

Notice that (2) implies $A^t = A^{-1}$ and that for any $n \times n$ matrices B and C and for any vectors $x, y \in R^n$ we have

$$B.x \bullet C.y = x \bullet (B^t.C).y\,.$$

As a consequence we get the following:

(a) The transpose of an orthogonal matrix is orthogonal.

(b) The inverse of an orthogonal matrix is orthogonal.

(c) A product of orthogonal matrices is orthogonal.

(d) If A is orthogonal, then $\det(A) = 1$ or $\det(A) = -1$.

For (d) notice that for any $n \times n$ matrices B and C we have

$$\det(B.C) = \det(B) \cdot \det(C)\,.$$

Example: Let us consider $\mathbb{R}^2$. Since rotations about the origin and reflections about lines through the origin preserve length and orthogonality, the standard matrices of these linear mappings must be orthogonal. In fact, the matrices corresponding to these mappings are the *only* orthogonal 2×2 matrices. In other words, every orthogonal 2×2 matrix is expressible in the form

$$R_\alpha = \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix} \quad \text{or} \quad H_{\alpha/2} = \begin{pmatrix} \cos(\alpha) & \sin(\alpha) \\ \sin(\alpha) & -\cos(\alpha) \end{pmatrix}$$

That is, every orthogonal 2×2 matrix A is either a rotation through some angle α about the origin, in which case $\det(A) = 1$, or a reflection about the straight line $t \cdot \big(\cos(\alpha/2), \sin(\alpha/2)\big)$, in which case $\det(A) = -1$.