Differentiability of Bilinear Maps

S Kumaresan
kumaresa@gmail.com

9 March 2021

Definition 1. Let V_i, $i = 1,2$ and W be vector spaces over a field $\mathbb{F}$. A map $f\colon V_1 \times V_2 \to W$ is *bilinear* if f is linear in each of its variables when the other variable is fixed: $v_1 \mapsto f(v_1, v_2)$ from V_1 to W is linear for any fixed $v_2 \in V_2$ and $v_2 \mapsto f(v_1, v_2)$ from V_2 to W is linear for any fixed $v_1 \in V_1$.

Ex. 2. How do you define a k-linear map $f\colon V_1 \times \cdots \times V_k \to W$?

Example 3. Two standard and familiar examples are the inner product on V and the multiplication map on $\mathbb{R}$. To be precise, let $(V, \langle\, ,\rangle)$ be a real inner product space. Let $V_1 = V_2 = V$ and $f(v_1, v_2) := \langle v_1, v_2\rangle$. Then f is bilinear.

Let $V_1 = V_2 = \mathbb{R}$ and let $f\colon \mathbb{R}\times\mathbb{R} \to \mathbb{R}$ be the multiplication map $f(x, y) = xy$, the product of two real numbers. Then f is bilienar.

Example 4. Let $M(2,\mathbb{R})$. We may consider it as a a direct sum as follows: Let $A = \begin{pmatrix} a & b \\ c & d\end{pmatrix} \in M(2,\mathbb{R})$. Let $C_1 := \begin{pmatrix} a \\ c\end{pmatrix}$ and $C_2 := \begin{pmatrix} b \\ d\end{pmatrix}$ be the columns of A. Then $A = (C_1, C_2) \in \mathbb{R}^2 \times \mathbb{R}^2$. The determinant map $f(a) := \det A$ is bilinear.

Is f a bilinear map of the rows? Make this question precise and answer it!

Ex. 5. Can you think of the determinant function on $M(n,\mathbb{R})$ as a k-linear function for a suitable k?

Recall that a linear map between vector spaces is completely determined once we know its values on the elements of a basis. Let us see whether we have an analogue for the bilinear maps. To keep the notation simple, let $V_1 = V_2 = V$. Let $f\colon V \times V \to W$ be bilinear. Let $\{v_1, \dots, v_n\}$ be an ordered basis of V. Let $x := \sum_i x_i v_i$ and $y := \sum_i y_i v_i$ be vectors in V. We then have

$$f(x, y) = f\left(\sum_i x_i v_i, \sum_j y_j v_j\right) = \sum_{i,j} x_i y_j f(v_i, v_j). \qquad \text{(Why?)}$$

Hence it follows that if we know $f(v_i, v_j)$, $1 \le i, j \le n$, we “know" $f(x, y)$ for any $(x, y) \in V \times V$.

Let us now assume that $V := V_i = \mathbb{R}^m$, $i = 1,2$ and $W = \mathbb{R}^n$. Let $v_i = e_i$, $1 \le i \le m$, be the standard basis on $\mathbb{R}^m$. Let V and W be equipped with one of the three norms $\| \ \|$. Then it follows that

$$\| f(x,y) \| \le \sum_{i,j} |x_i| \left|y_j\right| \| f(e_i,e_j) \| \le C \| x \| \| y \|, \quad \text{where } C := \sum_{i,j} \| f(e_i,e_j) \|. \tag{1}$$

It is easy deduce the continuity of f from this estimate. (Exercise: Prove this.)

Ex. 6. Find an analogue of (1) for k-linear maps.

We keep the notation preceding (1). We now claim that f is differentiable. Let $(u,v) \in \mathbb{R}^m \times \mathbb{R}^m$. We show that f is differentiable at (u,v). Let $(h,k) \in \mathbb{R}^m \times \mathbb{R}^m$. We then observe that

$$\begin{aligned} f(u+h,v+k) - f(u,v) &= f(u+h,v) + f(u+h,k) - f(u,v) \\ &= f(u,v) + f(h,v) + f(u,k) + f(h,k) - f(u,v) \\ &= f(h,v) + f(u,k) + f(h,k). \end{aligned}$$

We let $A(h,k) := f(u,k) + f(h,v)$. It is easy to verify that A is linear. The 'error term' $E(h,k)$ is $f(h,k)$. In view of (1), we see that $\| E(h,k) \| \le C \| \| h \| \| k \| \le C \| (h,k) \| \|^2$. Thus f is differentiable and the derivative is the linear map

$$Df(u,v)\colon (h,k) \mapsto f(u,k) + f(h,v). \tag{2}$$

Ex. 7. Let $f\colon \overbrace{\mathbb{R}^m \times \cdots \times \mathbb{R}^m}^{k-\text{times}} \to \mathbb{R}^n$ be k-linear. Prove that f is differentiable and compute its derivative.

We now apply this to the determinant map on $M(2,\mathbb{R})$. Let A be as in Example 4. Let $M(2,\mathbb{R}) \ni H := \begin{pmatrix} x & y \\ z & w \end{pmatrix} = (H_1,H_2) \in \mathbb{R}^2 \times \mathbb{R}^2$ where H_i is the i-th column of H. By (2), we see that

$$\begin{aligned} Df(A)(H) &= \det(C_1,H_2) + \det(H_1,C_2) \\ &= \det\begin{pmatrix} a & y \\ c & w \end{pmatrix} + \det\begin{pmatrix} x & b \\ z & d \end{pmatrix} \\ &= (aw - cy) + (dx - bz). \end{aligned} \tag{3}$$

Let us recall that we have an inner product on $M(2,\mathbb{R})$ which as obtained by identifying $M(2,\mathbb{R})$ with $\mathbb{R}^4$:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto (a,b,c,d) \quad \text{so that } \left\langle \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \begin{pmatrix} x & y \\ z & w \end{pmatrix} \right\rangle = ax + by + cz + dw.$$

We also know that it can be intrinsically defined as follows:

$$\left\langle \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \begin{pmatrix} x & y \\ z & w \end{pmatrix} \right\rangle = \operatorname{Tr}\left(\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x & y \\ z & w \end{pmatrix}^T \right).$$

If f is the determinant map on $M(2,\mathbb{R})$, we know it is differentiable and we have computed the derivative in (3). We now ask the question: What is its gradient? We would like to exhibit the gradient as an element of $M(2,\mathbb{R})$ as a matrix, not as a vector in $\mathbb{R}^4$! Thus, we need to find a matrix B such that $Df(A)(H) = \langle H, B\rangle = \operatorname{Tr}(HB^T)$. A little inspection of (3) and reflection leads us to $B^T = \begin{pmatrix} d & -c \\ -b & a \end{pmatrix}$. Do you recognize how B is related to A? If not, you may need to review the algebra of matrices! B is the so-called adjunct matrix of A.

Ex. 8. You know what this exercise is about. State the exercise and solve it!

Now let us have some fun. The following material is not true to the title!

Let us denote an arbitrary element $A \in M(2,\mathbb{R})$ by $A = \begin{pmatrix} x & y \\ z & w \end{pmatrix}$. Thus $f(A) := \det(A) = xw - yz$. This is a polynomial function and hence is differentiable. So, to find the gradient it suffices to find the partial derivatives! They are

$$\frac{\partial f}{\partial x} = w, \quad \frac{\partial f}{\partial y} = -z, \quad \frac{\partial f}{\partial z} = -y \quad \text{and} \quad \frac{\partial f}{\partial w} = x.$$

So, if we wish to visualize the gradient of f as an element of $M(2,\mathbb{R})$, it should look like

$$\begin{pmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \\ \frac{\partial f}{\partial z} & \frac{\partial f}{\partial w} \end{pmatrix} = \begin{pmatrix} w & -z \\ -y & x \end{pmatrix}.$$

Viola, we got it so easily!

Ex. 9. What should be the next question? Formulate it and try to answer it!

Hint for Ex. 9: The question is: Is there a similar easy method to compute the gradient of the determinant function on $M(n,\mathbb{R})$?

The answer is: Yes, if you remember the facts on determinants! Recall the Laplace expansion of determinants in terms of cofactors. (We are not going to define cofactors!) Given $A := (a_{ij}) \in M(n,\mathbb{R})$, let C_{ij} denote the (i,j)-th cofactor of a_{ij}. Then Laplace expansion says

$$a_{i1}C_{j1} + \cdots + a_{in}C_{jn} = \delta_{ij}\det(A) = a_{1i}C_{1j} + \cdots + a_{ni}C_{nj},. \tag{4}$$

where δ is Kronecker delta. Can you now go ahead and complete the answer? Try on your own. If you are lazy, follow the hint below.

Hint: Take $j = i$ in (4). Observe that C_{ij} is independent of the (i,j)-th variable a_{ij}. Compute $\frac{\partial \det}{\partial a_{ik}}$. Does your answer agree with what we found when $n = 2$? ◇