Existence of antiderivatives in multivariable calculus

Rayan Charrier

Abstract

I prove that a differentiable vector field is a gradient if and only if its Jacobian is symmetric, a result which is extended to higher-order tensors. The assumption of C^1-differentiability is relaxed by differentiating under the integral sign without domination (allowed by the particular formulation of the problem).

Introduction

The Heinstock-Kurzweil integration theory provides necessary and sufficient conditions to be a derivative (see [1, p.243]). If $f\colon [a;b] \to \mathbb{R}$ is HK-integrable then its primitive integral $F\colon x \mapsto \int_a^x f$ is differentiable almost everywhere with $F' = f$ [1, p.80]. Conversely, if $F\colon [a;b] \to \mathbb{R}$ is differentiable with at most countably exceptions then its derivative F' is HK-integrable with $F(x) = F(a) + \int_a^x F'$ for all $x \in [a;b]$ [1, p.60]. Cantor's function shows that almost-everywhere differentiability does not suffice [1, p.67]. However, in multivariable calculus the question is not only about "smoothness", but also about "shape": some symmetry is required (and sufficient when continuous). This "shape" condition being already known, the main contribution of this work is to drop the assumption of continuity for the differential in the converse implication.

Results

Theorem 1. *Let $\mathrm{O} \subset \mathbb{R}^d$ open and $V\colon \mathrm{O} \to \mathbb{R}^d$ differentiable.*
Then, V is the gradient of a functional on O if and only if $J_V(x)$ is symmetric for all $x \in \mathrm{O}$.

As a reminder, a matrix function is a Jacobian if and only if all its rows are (transposed) gradients.

Corollary 2. *A matrix function is a Hessian if and only if its rows are (transposed) gradients and the matrix images are symmetric, i.e it is a symmetric Jacobian.*

Theorem 1 extends to the next orders.

Theorem 3. *Let $\mathrm{O} \subset \mathbb{R}^d$ open, and $T\colon \mathrm{O} \to \mathbb{R}^{d_1 \times \cdots \times d_{k-1} \times d}$ differentiable.*
Then, T is the derivative of a $(k-1)^{th}$-order tensor on O if and only if the matrices $\mathrm{d}T(x)_{i_1,\ldots,i_{k-1}}$ are symmetric for all $(i_1,\ldots,i_{k-1}) \in \prod_{1 \leqslant j \leqslant k-1} [\![1;d_j]\!]$ and $x \in \mathrm{O}$.

Theorem 4. *Let $\mathrm{O} \subset \mathbb{R}^d$ open and $T\colon \mathrm{O} \to \mathbb{R}^{d \times \cdots \times d}$ differentiable.*
Then, T is the k^{th}-derivative of a functional on O if and only if the tensors $T(x)$ and $\mathrm{d}T(x)$ are symmetric for all $x \in \mathrm{O}$.

Remark 5. Vector fields, which are one-sided, are already symmetric, hence a single symmetry condition in Theorem 1.

Remark 6. In all the above results, differentiability can be considered on the complementary of a countable set. However, it cannot be relaxed by considering finite differences.

The most basic case is that of matrices (2^{nd}-order tensors).

Corollary 7. *Let $\mathrm{O} \subset \mathbb{R}^d$ open and $M\colon \mathrm{O} \to \mathbb{R}^{d \times d}$ differentiable.*
*Then, M is the Hessian of a functional on O if and only if $M(x)$ **and** $\mathrm{d}M(x)$ are symmetric for all $x \in \mathrm{O}$.*

Remark 8. Nor the symmetry of the images of M nor that of its derivative's alone are sufficient conditions.

Being a Hessian is not preserved by basic function operations.

Remark 9. The square of a Hessian is not necessarily a Hessian.

Proof of Clairaut's theorem (a.k.a Schwarz')

The direct implication of Theorem 1 is known as Clairaut's theorem.

Theorem. *Let* $\mathrm{O} \subset \mathbb{R}^2$ *open,* $f \colon \mathrm{O} \to \mathbb{R}$ *differentiable, and* $(u, v) \in \mathrm{O}$ *such that* f *is twice differentiable at* (u, v)*. Then,*

$$\partial_1 \partial_2 f\big|_{(u,v)} = \partial_2 \partial_1 f\big|_{(u,v)}.$$

Proof. We proceed as in [9], [3, p.175]. We prove that

$$\frac{1}{t}\left[\frac{f(u+t, v+t) - f(u, v+t)}{t} - \frac{f(u+t, v) - f(u, v)}{t}\right] \underset{t \to 0}{\longrightarrow} \partial_1 \partial_2 f\big|_{(u,v)}. \tag{1}$$

Noting that the sum of the numerators can be re-formulated as

$$\big[f(u+t, v+t) - f(u+t, v)\big] - \big[f(u, v+t) - f(u, v)\big],$$

this suffices to prove the result by switching the order of the arguments.

A natural approach might be to apply the mean-value theorem at each fraction of the left-hand term in (1), which yields terms in $\partial_1 f$, before using the definition of differentiation to obtain $\partial_2 \partial_1 f|_{(u,v)}$ (and $\partial_1 \partial_2$ with the re-formulation). The approximations can be bounded using the mean-value inequality but such an approach uses the continuity of the second-order differential (and in particular the fact that it is defined on a neighborhood).

Instead, consider for all t small enough the function defined on a common neighborhood of 0 by

$$g_t(s) := f(u+t, v+s) - f(u, v+s),$$

so that the left-hand term in (1) corresponds to $(g_t(t) - g_t(0))/t^2$. For any such t the mean-value theorem provides the existence of $\varepsilon_t \in (0, t)$ such that

$$\frac{g_t(t) - g_t(0)}{t} = g_t'(\varepsilon_t) = \partial_2 f|_{(u+t, v+\varepsilon_t)} - \partial_2 f|_{(u, v+\varepsilon_t)}. \tag{2}$$

Moreover, by the definition of the differential,

$$\partial_2 f|_{(u+t, v+\varepsilon_t)} = \partial_2 f|_{(u,v)} + \partial_1 \partial_2 f|_{(u,v)} \times t + \partial_2 \partial_2 f|_{(u,v)} \times \varepsilon_t + \mathrm{o}(|t| + |\varepsilon_t|).$$

Similarly,

$$\partial_2 f|_{(u, v+\varepsilon_t)} = \partial_2 f|_{(u,v)} + \partial_2 \partial_2 f|_{(u,v)} \times \varepsilon_t + \mathrm{o}(|t| + |\varepsilon_t|).$$

Plugging into (2) yields, noting that $|\varepsilon_t| \leqslant |t|$:

$$\frac{g_t(t) - g_t(0)}{t} = \partial_1 \partial_2 f|_{(u,v)} \times t + \mathrm{o}(t),$$

and hence the result when t goes to 0 in (1).

We now prove the converse implication of Theorem 1.

Proof of Theorem 1

We start with the case O star-shaped [7], i.e there exists $x_\star \in \mathrm{O}$ such that $[x_\star; x] \subset \mathrm{O}$ for all $x \in \mathrm{O}$. Without loss of generality we can assume that $x_\star = 0$. Define then for all $x \in \mathrm{O}$ (it must be that by the fundamental theorem of calculus):

$$f(x) := \int_0^1 V(tx)^\top x \, \mathrm{d}t.$$

If we can differentiate under the integral sign (for instance when J_V is continuous), then

$$\nabla f(x) = \int_0^1 t J_V(tx)^\top x \, \mathrm{d}t + \int_0^1 V(tx) \, \mathrm{d}t, \tag{3}$$

where integrating by parts provides

$$\int_0^1 V(tx) \, \mathrm{d}t = \Big[tV(tx)\Big]_0^1 - \int_0^1 t J_V(tx) x \, \mathrm{d}t. \tag{4}$$

Using the symmetry of the Jacobian, plugging into (3) yields $\nabla f(x) = V(x)$.

We now drop the assumption that O is star-shaped. Since we can consider the connected components of O separately, they are indeed open and disjoint, we assume that O is connected. Write using Lemma 12 in appendix:

$$\mathrm{O} = \underset{n\in\mathbb{N}}{\cup} \mathrm{B}\left(x_n, r_n\right).$$

We define f on $\cup_{n\leqslant N} \mathrm{B}\left(x_n, r_n\right)$ by induction on $N \in \mathbb{N}$. The first part of the proof provides the initialization. When adding the $N + 1^{th}$ ball, with some arbitrary $x_\star$ in the intersection with $\cup_{n\leqslant N} \mathrm{B}\left(x_n, r_n\right)$ (see Lemma 12), we define on the new ball:

$$f(x) := f(x_\star) + \int_0^1 V(x_\star + t(x - x_\star))^\top (x - x_\star) \mathrm{d}t.$$

This function is well-defined on the intersection where it must already satisfy this equality by the fundamental theorem of calculus and the inductive hypothesis that $V = \nabla f$.

We now prove that we can write $\nabla f = V$ with the only assumption that V is differentiable. We want to be able to differentiate under the integral sign. But even for the integrals to be defined, we need to consider a generalization of the Riemann integral: the Heinstock-Kurzweil integral (or Denjoy's or Perron's integral [1]), which has the desirable property that any derivative of a univariate function is integrable (see [1, Theorem 4.5] for instance). Note that it does not extend to multivariate functions (see [6, Exemple 4] for a differentiable function $g\colon \mathbb{R}^2 \to \mathbb{R}$ such that $\partial_2 g$ is not integrable on $[0; 1]$ in the first variable). We must thus proceed coordinate-wise.

The key is the following lemma [10, Corollary 8-ii]. *We introduce a particular case since on a line segment, differentiability with at most countable exceptions implies being ACG* (see [10, p.3], or [1, Theorem 4.7, Theorem 14.22] for a proof).*

Lemma 10. *Let* $g\colon [s_1; s_2] \times [0; 1] \subset \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ *such that*

(i) $g(\cdot, t)$ *is differentiable on* $[s_1; s_2]$ *for all* $t \in [0; 1]$,

(ii) $\int_0^1 \partial_s g(\cdot, t) \mathrm{d}t$ *is continuous on* $[s_1; s_2]$,

(iii) we can write for all $[s_1'; s_2'] \subset [s_1; s_2]$:

$$\int_0^1 \int_{s_1'}^{s_2'} \partial_s g \, \mathrm{d}s\mathrm{d}t = \int_{s_1'}^{s_2'} \int_0^1 \partial_s g \, \mathrm{d}t\mathrm{d}s.$$

Then we can differentiate under the integral sign: on $]s_1; s_2[$,

$$\partial_s \left[\int_0^1 g(\cdot, t) \, \mathrm{d}t\right] = \int_0^1 \partial_s g(\cdot, t) \, \mathrm{d}t.$$

Let $x \in \mathrm{O}$ and $i \in [\![1;d]\!]$. For all $s \in \mathbb{R}$ write $x_s := x - x_i e_i + s e_i$ (we replace the i^{th}-coordinate by s). Define then for all s in a neighborhood $]s_1; s_2[$ of x_i and $t \in]0;1[$:

$$g(s,t) := V(tx_s)^\top x_s,$$

so that for all $t \in [0;1]$ and $s \in [s_1; s_2]$ we can write

$$\partial_1 g(\cdot,t)|_s = \left[tJ_V(tx_s)^\top x_s + V(tx_s)\right]_i. \tag{5}$$

Let $[s_1'; s_2'] \subset [s_1; s_2]$.

The first integral in Lemma 10's assumption (iii) is well-defined and equal to [1, p.59]:

$$\begin{aligned}\int_0^1 \int_{s_1'}^{s_2'} \partial_1 g \; \mathrm{d}s\mathrm{d}t &= \int_0^1 \left(g(s_2',t) - g(s_1',t)\right) \mathrm{d}t \\ &= \int_0^1 \left(V(tx_{s_2'})^\top x_{s_2'} - V(tx_{s_1'})^\top x_{s_1'}\right) \mathrm{d}t. \end{aligned}\tag{6}$$

Regarding the second integral in the assumption, we must first prove that it is well-defined, and in particular the integral in t. The function $t \mapsto V(tx_s)$ in (5) is continuous hence integrable, but $t \mapsto tJ_V(tx_s)^\top x_s$ is not a priori (see [6, Exemple 4]). The trick is that for all $s \in [s_1'; s_2']$ the function $t \mapsto V(tx_s)$ has for derivative $t \mapsto J_V(tx_s)x_s$, and integrating by parts with $t \mapsto t$ for second function allows to write (see [1, Theorem 4.19]):

$$\int_0^1 tJ_V(tx_s)x_s \mathrm{d}t = [tV(tx_s)]_0^1 - \int_0^1 V(tx_s)\mathrm{d}t. \tag{7}$$

Plugging into (5) and using the symmetry of the Jacobian we deduce

$$\int_{s_1'}^{s_2'} \int_0^1 \partial_1 g \; \mathrm{d}t\mathrm{d}s = \int_{s_1'}^{s_2'} \int_0^1 \left[tJ_V(tx_s)^\top x_s + V(tx_s)\right]_i \mathrm{d}t\mathrm{d}s = \int_{s_1'}^{s_2'} V(x_s)_i \mathrm{d}s. \tag{8}$$

With (6) and (8) the integral assumption in Lemma 10 becomes

$$\int_0^1 \left(V(tx_{s_2})^\top x_{s_2} - V(tx_{s_1})^\top x_{s_1}\right) \mathrm{d}t = \int_{s_1'}^{s_2'} V(x_s)_i \mathrm{d}s.$$

It only involves V and no derivatives of V, and can be written equivalently with Riemann integrals since V is continuous [1, p.14]. This integral equation is satisfied when V is C^1-differentiable by the classical theorems of interchange. Otherwise, the Stone-Weierstrass theorem allows to uniformly approximate V by C^1-differentiable functions, that will necessarily satisfy this equation, and so will V by uniform convergence under the integral sign. Assumption (iii) thus holds. Using (7) and the symmetry of the Jacobian also proves that assumption (ii) holds. We can then use Lemma 10 to establish (3) coordinate-wise:

$$\begin{aligned}\partial_i f(x) = \partial_i \left[\int_0^1 V(t\cdot)^\top \cdot \mathrm{d}t\right](x) &= \partial_s \left[\int_0^1 g(\cdot,t)\mathrm{d}t\right](x_i) \\ &= \left[\int_0^1 \partial_s g(\cdot,t)\mathrm{d}t\right](x_i) = \int_0^1 \left[tJ_V(tx)x + V(tx)\right]_i \mathrm{d}t.\end{aligned}$$

The integration by parts (4) has been proved in (7), hence $\nabla f(x) = V(x)$.

Proof of Corollary 2

To begin with, a Hessian is a Jacobian, symmetric by Clairaut's theorem.

Conversely, a symmetric Jacobian is the derivative of a vector field that satisfy the conditions of Theorem 1, i.e is a gradient.

Proof of Theorem 3

The theorem informs whether there exists or not some differentiable tensor function $\mathit{\Gamma}\colon \mathrm{O} \subset \mathbb{R}^d \to \mathbb{R}^{d_1 \times \cdots \times d_{k-1}}$ such that $\mathtt{d}\mathit{\Gamma} = T$ on O. This can be brought down to $d_1 \times \cdots \times d_{k-1}$ independent conditions on vector fields $V_{i_1,\ldots,i_{k-1}}$ with $i_j \in [\![1; d_j]\!]$ for all $j \in [\![1; k-1]\!]$ (see Appendix). It then suffices to apply Theorem 1 to each of them.

Proof of Theorem 4

The direct implication follows inductively from Clairaut's theorem and is well known. We prove the converse by induction [2].

To begin with, the symmetry of $\mathtt{d}T(x)$ proves that T is the derivative of a $(k-1)^{th}$-order tensor-valued function $\mathit{\Gamma}$ on O by Theorem 3. Moreover, the fundamental theorem of calculus allows to write for all $[x; y] \subset \mathrm{O}$:

$$\mathit{\Gamma}(y) - \mathit{\Gamma}(x) = \int_0^1 T(x + t(y-x)) \times (y-x)\mathtt{d}t.$$

Since $T(x+t(y-x))$ is symmetric for all $t \in [0;1]$, and thus $T(x+t(y-x)) \times (y-x)$, if $\mathit{\Gamma}(x)$ is symmetric then $\mathit{\Gamma}(y)$ is too. Since $\mathit{\Gamma}$ is defined within an additive constant, we can thus choose $\mathit{\Gamma}$ symmetric by setting for each connected component of O a symmetric image *(one may then verify that the set of points with a symmetric image is both open and closed in this connected component)*. This allows to proceed by induction until $\mathit{\Gamma}$ is a vector field, when Theorem 1 concludes the proof.

Proof of Remark 6 (countable set)

Lemma 10 is actually stated in [10, Corollary 8-i] with differentiability everywhere except on a countable set, and the HK-integration theorems we use are extended for that in [1, Theorem 4.7, Theorem 12.2.b].

Counterexample for Remark 6 (finite differences)

Consider the vector field defined on $\mathbb{R}^2$ by $V(u,v) := (v^2, 2uv)$, which is the gradient of $f\colon (u,v) \mapsto uv^2$. Comparing the finite differences for V_1 and V_2 instead of $\partial_2 V_1$ and $\partial_1 V_2$ amounts to considering for all $h, k \in \mathbb{R}^*$,

$$\begin{aligned} & \frac{V_1(u,v+k) - V_1(u,v)}{k} && \text{and} \quad \frac{V_2(u+h,v) - V_2(u,v)}{h} \\ = & \frac{(v+k)^2 - v^2}{k} && \quad = \frac{2(u+h)v - 2uv}{h} \\ = & \ 2v + k && \quad = 2v. \end{aligned}$$

Equalities as necessary conditions thus require to take the limit when $k \to 0$ (i.e differentiating).

Counterexample for Remark 8 (symmetry of M)

Symmetry does not even guarantee to be a Jacobian. Consider for example the matrix function

$$M\colon (u,v) \mapsto \begin{pmatrix} u & 0 \\ 0 & u \end{pmatrix}.$$

The vector field of the last row, defined on $\mathbb{R}^2$ by $V(u,v) := (0,u)$, has for Jacobian $\left(\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right)$ and therefore is not a gradient by Theorem 1.

One can also compute the vector field $W\colon (u,v) \mapsto \int_0^1 M(0 + t(u,v)) \times (u,v)\mathrm{d}t$ and check that its Jacobian J_W is not equal to M. If M was the Jacobian of a vector field, then by the fundamental theorem of calculus it would be equal to W within an additive constant, and thus have for Jacobian both J_W and M.

Counterexample for Remark 8 (symmetry of $\mathtt{d}M$)

The condition of symmetry for the images of M is necessary by Clairaut's theorem. It is not met by the constant matrix function equal to $\left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$, even if it has a null, and thus symmetric, derivative.

Counterexample for Remark 9

Consider the function $f(u,v) = \frac{u^2 v}{2}$, such that

$$\nabla^2 f(u,v) = \begin{pmatrix} v & u \\ u & 0 \end{pmatrix} \quad \text{and} \quad M(u,v) := \left[\nabla^2 f(u,v)\right]^2 = \begin{pmatrix} u^2 + v^2 & uv \\ uv & u^2 \end{pmatrix}.$$

The matrix function M has symmetric images but its tensor derivative has not:

$$\partial_1 M_{12}|_{(u,v)} = v \quad \text{and} \quad \partial_2 M_{11}|_{(u,v)} = 2v.$$

Without using tensors, only vectors and matrices as in Corollary 2, this also proves by Theorem 1 that the matrix function M is not even a Jacobian.

Discussion

Helmholtz's decomposition

Helmholtz' decomposition, when it does exist, splits a vector field into the sum of a conservative and of a solenoidal vector field. It is of significant importance in electrodynamics for instance (in particular for the integral characterizations of these properties). The fundamental theorem of vector calculus provides sufficient conditions for the existence of such a decomposition [4, 5, 8].

As a reminder, a *conservative* vector field is the gradient of a functional, which is characterized in Theorem 1 by a symmetric Jacobian, and a *solenoidal* vector field is a divergence-free one, which is characterized by a Jacobian of null trace.

In order to guarantee the symmetry condition in Theorem 1, one could be tempted to consider the symmetric-antisymmetric decomposition of Jacobians to obtain Helmholtz' decompositions. It is indeed well-known that any square matrix M can be written $M =: S + A$ with S symmetric and A antisymmetric, and in particular of null trace, by considering

$$S := \frac{M + M^\top}{2} \quad \text{and} \quad A := \frac{M - M^\top}{2}.$$

However, as pointed out in Remark 8, the symmetry of the images of a matrix function alone does not suffice to be a Hessian, and this approach does not necessarily work for non-linear vector fields. Consider the counterexample defined on $\mathbb{R}^2$ by

$$V(u,v) := \begin{bmatrix} u^2/2 + v^2/2 \\ -uv \end{bmatrix}.$$

One can verify that V admits a Helmholtz's decomposition $V = \nabla f + W$ with

$$f\colon (u,v) \mapsto \frac{u^3}{6} + \frac{u^2 v}{2} - \frac{uv^2}{2} - \frac{v^3}{6} \quad \text{and} \quad W\colon (u,v) \mapsto (-uv + v^2, \frac{v^2 - u^2}{2}).$$

However,

$$J_V(u,v) = \begin{pmatrix} u & v \\ -v & -u \end{pmatrix} \quad \text{and} \quad S(u,v) = \frac{J_V(u,v) + J_V(u,v)^\top}{2} = \begin{pmatrix} u & 0 \\ 0 & -u \end{pmatrix},$$

which is not the Jacobian of a conservative vector field (see previous section).

As a remark, the differentiable vector fields that are both conservative and solenoidal are the gradients of harmonic functions. Helmholtz decompositions are thus unique within addition-substraction of such vector fields.

Regularity

Theorem 1 shows that symmetry is a kind of regularity property, in particular to go back in differentiation orders (the vector field must be sufficiently structured for that).

Note also on the other hand that if a conservative vector field has the property to be solenoidal, then this makes it infinitely differentiable (as the gradient of a harmonic function).

Appendices

Differential

Let $(\mathrm{E}, \|\cdot\|_E)$ and $(\mathrm{F}, \|\cdot\|_F)$ be two Banach spaces, $\mathrm{O} \subset \mathrm{E}$ open and $f \colon \mathrm{O} \to \mathrm{F}$.

The function f is said to be *differentiable at some* $x \in \mathrm{O}$ if there exists a continuous linear function $g \colon \mathrm{E} \to \mathrm{F}$ such that

$$f(x+h) \underset{h \to 0}{=} f(x) + g(h) + \mathrm{o}(h).$$

(Note that linear functions in finite dimensions are necessarily continuous.) If this holds then we define the *differential of f at x* as $\mathtt{d}f_x := g$, so that

$$f(x+h) \underset{h \to 0}{=} f(x) + \mathtt{d}f_x \cdot h + \mathrm{o}(h).$$

We say that f is *differentiable on* O if it is at all $x \in \mathrm{O}$, and we then define the *differential of f on* O as the function

$$\begin{aligned} \mathtt{d}f \colon \mathrm{E} &\longrightarrow \mathcal{L}(\mathrm{E}, \mathrm{F}) \\ x &\longmapsto \mathtt{d}f_x, \end{aligned}$$

where $\mathcal{L}(\mathrm{E}, \mathrm{F})$ denotes the set of continuous linear functions from E to F.

The differential of the differential is then a function from E to $\mathcal{L}(\mathrm{E}, \mathcal{L}(\mathrm{E}, \mathrm{F}))$, and its differential (the third-order one) from E to $\mathcal{L}(\mathrm{E}, \mathcal{L}(\mathrm{E}, \mathcal{L}(\mathrm{E}, \mathrm{F})))$. We assimilate the codomain of the n^{th}-order differential with $\mathcal{L}^n(\mathrm{E} \times \cdots \times \mathrm{E}, \mathrm{F})$ the set of continuous multilinear functions from E^n to F.

In light of the next appendix' last remark, in finite dimension we can in turn assimilate each function $\mathtt{d}^n f_x$ with a tensor, which contains the partial derivatives of order n with respect to the corresponding coordinates.

If the vector field is $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$, and E is Hilbert (i.e endowed with an inner product $<,>$), we can also use Riesz' representation theorem.

For example, Riesz' representation theorem provides that the linear functions $g \colon \mathbb{R}^p \to \mathbb{R}$ are uniquely determined by a (the) vector $v \in \mathbb{R}^p$ such that $g(h) = v^\top h$ for all $h \in \mathbb{R}^p$. When applied to the differential of a functional f at some x, this allows to define the *gradient of f at x* as the $v =: \nabla f(x)$ such that $\mathtt{d}f_x \cdot h = \nabla f(x)^\top h$ for all $h \in \mathbb{R}^p$, i.e

$$f(x+h) = f(x) + \nabla f(x)^\top h + \mathrm{o}(h).$$

The differential can be assimilated with the gradient function $x \mapsto \nabla f(x)$, whose i^{th}-coordinate is the derivative along e_i (the partial derivative $\partial_i f$ with respect to the i^{th}-coordinate).

In the case of a multivariate and vector-valued function $V \colon \mathbb{R}^p \to \mathbb{R}^q$ this yields

$$\begin{bmatrix} V_1(x+h) \\ \vdots \\ V_q(x+h) \end{bmatrix} = \begin{bmatrix} V_1(x) \\ \vdots \\ V_q(x) \end{bmatrix} + \begin{pmatrix} \nabla V_1(x)^\top \\ \vdots \\ \nabla V_q(x)^\top \end{pmatrix} \times h + \mathrm{o}(h),$$

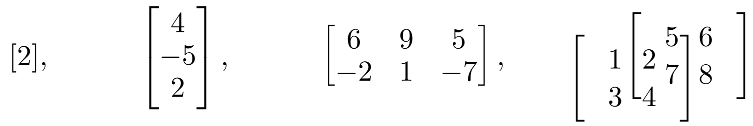
which writes

$$V(x+h) = V(x) + J_V(x) h + \mathrm{o}(h),$$

and we can assimilate the differential with the Jacobian matrix.

Tensors

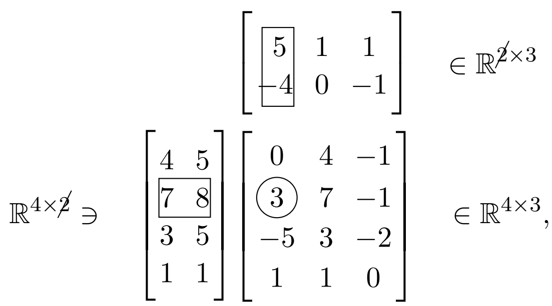
Tensors extend the definition of scalars, vectors and matrices to higher orders:



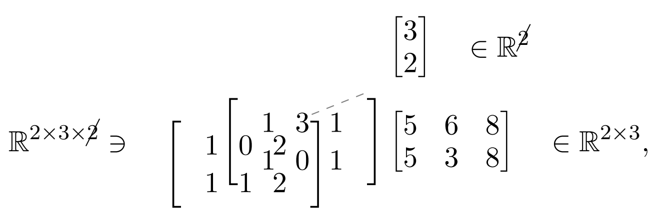
so that a scalar is a 0^{th}-order tensor, the above vector is a 1^{st}-order tensor of dimension 3, the above matrix is a 2^{nd}-order tensor of dimensions $(2, 3)$, and the last object is a 3^{rd}-order tensor of dimensions $(2, 2, 2)$.

A symmetric tensor is a tensor invariant by permuting the dimensions' ordering (as for matrices when swapping i and j).

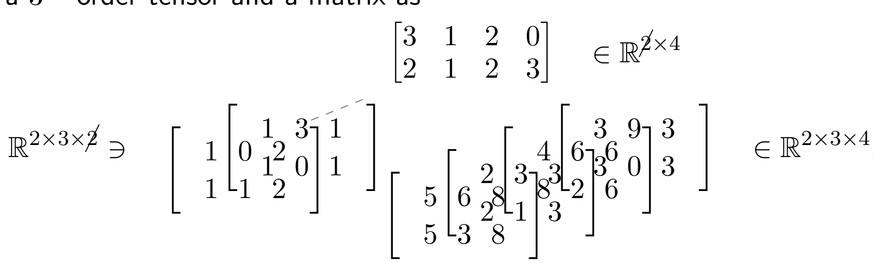
The canonical matrix product can be generalized to tensors. For example in



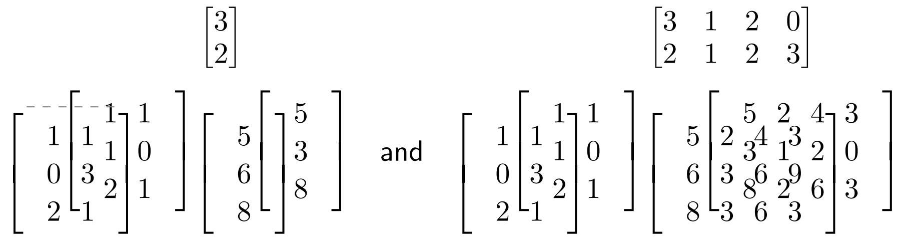
each new coefficient is a scalar product (as in the multiplication of a vector by a matrix and, trivially, in the scalar product of two vectors). The *inner product* of a 3^{rd}-order tensor and a vector is then defined as



and that of a 3^{rd}-order tensor and a matrix as



One may argue that with those visualizations, the inner product correctly displays the initial and final dimensions, but do not visually conserve orders and dimensions as with matrices (for example, in the product of a 3^{rd}-order tensor with a matrix, the dimension 4 starts as a width and ends up as a depth). We can change perspectives with a permutation of the dimensions to represent those products as



On the other hand, such a visualization is not clear anymore regarding how dimensions evolve: the represented 3^{rd}-order tensor has dimensions $(3,2,2)$, the vector (2), and the resulting tensor $(3,1,2)$, whereas for the second product, dimensions $(3,2,2)$ and $(2,4)$ produce a tensor of dimensions $(3,4,2)$.

The inner product is defined for all $n, m \in \mathbb{N}^*$ and $p_1, \dots, p_{n-1}, d, q_2, \dots, q_m \in \mathbb{N}^*$ by:

$$\begin{aligned} \mathbb{F}^{p_1 \times \dots \times p_{n-1} \times d} \times \mathbb{F}^{d \times q_2 \times \dots \times q_m} &\longrightarrow \mathbb{F}^{p_1 \times \dots \times p_{n-1} \times q_2 \times \dots \times q_m} \\ (R, S) &\longmapsto T := R \times S \quad \text{such that} \quad T_{IJ} = R_I^\top S_J \\ &\qquad \text{for all } (I, J) \in \prod_{i=1}^{n-1} [\![1; p_i]\!] \times \prod_{j=2}^{m} [\![1; q_j]\!]. \end{aligned}$$

The order of the resulting tensor is the sum of the arguments' orders minus 2 since the common dimension (the last one of the first tensor and the first one of the second tensor) disappears by the computation of the scalar products.

When applied between a tensor and a vector, the inner product reduces the order of the tensor by 1. For all $T \in \mathbb{F}^{d_1 \times \dots \times d_k}$, the multiplication with k vectors (to obtain a scalar) is then defined as the multilinear functional

$$\begin{aligned} T \colon \mathbb{F}^{d_1} \times \cdots \mathbb{F}^{d_k} &\longrightarrow \mathbb{F} \\ (v_1, \dots, v_k) &\longmapsto T(v_1, \dots, v_k) = (T \times v_k) \times \cdots \times v_1 \\ &\qquad = \sum_{\substack{1 \leqslant i_1 \leqslant d_i \\ 1 \leqslant i_k \leqslant d_k}} T_{i_1, \dots, i_k} \times [v_1]_{i_1} \times \cdots \times [v_k]_{i_k}. \end{aligned}$$

It appears in the sum that computing the product in any order does not affect the result (starting with T and an arbitrary v_i with d_i for common dimension).

Conversely, in finite dimensions all multilinear functions can be expressed via a tensor (see the above sum).

Open sets

Lemma 11. *An open set of $\mathbb{R}^d$ can be written as a countable union of open balls.*

Proof. Let $x \in \mathrm{O}$ and $r \in \mathbb{Q} \cap (0; +\infty)$ such that $\mathrm{B}(x, r) \subset \mathrm{O}$. By the density of the rational numbers in the real numbers there exists $q \in \mathbb{Q}^d \cap \mathrm{B}(x, r/2)$, so that $x \in \mathrm{B}(q, r/2) \subset \mathrm{B}(x, r) \subset \mathrm{O}$.

Lemma 12. *A connected open set of $\mathbb{R}^d$ can be written as a countable union of open balls that successively intersect the union of the previous ones.*

Proof. By Lemma 11 we can write $\mathrm{O} =: \cup_{n\in\mathbb{N}} \mathrm{B}(x_n, r_n)$. Starting with $\varphi(0) := 0$, we define by induction, while it is possible:

$$\varphi(n+1) := \mathtt{min}\{k \in \mathbb{N} \smallsetminus \varphi([\![0; n]\!]) \ : \ \mathrm{B}(x_k, r_k) \cap \bigcup_{i\in\varphi([\![0;n]\!])} \mathrm{B}(x_i, r_i) \neq \emptyset\}.$$

Moreover, since

$$\mathrm{O} = \bigcup_{n\in\mathrm{Im}\,\varphi} \mathrm{B}(x_n, r_n) \uplus \bigcup_{n\in\mathbb{N}\smallsetminus\mathrm{Im}\,\varphi} \mathrm{B}(x_n, r_n)$$

is a disjoint union of open sets, it must be that $\mathrm{Im}\,\varphi = \mathbb{N}$. Therefore,

$$\mathrm{O} = \bigcup_{n\in\mathbb{N}} \mathrm{B}\left(x_{\varphi(n)}, r_{\varphi(n)}\right),$$

where by construction each ball intersects the union of the previous balls (those with lower indices).

References

[1] R. G. Bartle, *A modern theory of integration. Graduate studies in Mathematics, 32*, American Mathematical Society, Providence, RI, 2001.

[2] Deane (https://math.stackexchange.com/users/10584/deane), *When is a matrix-valued function the Hessian of some function?* Mathematics Stack Exchange. https://math.stackexchange.com/q/4827549 (version: 2023-12-14).

[3] J. Dieudonne, *Foundations of modern analysis. Pure and Applied Mathematics, Vol. X*, Academic Press, New York-London, 1960.

[4] E. Glötzl and O. Richters, *Helmholtz decomposition and potential functions for n-dimensional analytic vector fields*, J. Math. Anal. Appl., 525 (2023).

[5] D. J. Griffiths, *Introduction to electrodynamics. Fourth edition*, Pearson, 2013.

[6] J. Jarník, J. Kurzweil, and S. Schwabik, *On Mawhin's approach to multiple nonabsolutely convergent integral*, Časopis Pěst. Mat., 108 (1983), pp. 356–380.

[7] Kajelad (https://math.stackexchange.com/users/354840/kajelad), *Can every symmetric Jacobian matrix be a Hessian matrix?* Mathematics Stack Exchange. https://math.stackexchange.com/q/4425665 (version: 2022-04-11).

[8] A. Kustepeli, *On the Helmholtz theorem and its generalization for multi-layers*, Electromagnetics, 36 (2016), pp. 135–148.

[9] J. Morrow, *Mixed Partials - Young's Theorem. Available: https://sites.math.washington.edu/~morrow/334_15/mixed%20partials%20young.pdf*, tech. rep., University of Washington.

[10] E. Talvila, *Necessary and sufficient conditions for differentiating under the integral sign*, Amer. Math. Monthly, 108 (2001), pp. 544–548.