THE TEACHING OF MATHEMATICS
2017, Vol. XX, No. 2, pp. 81–88

LAGRANGE'S FORMULA FOR VECTOR-VALUED FUNCTIONS

Milosav M. Marjanović and Zoran Kadelburg

Abstract. In this paper we derive a variant of the Lagrange's formula for the vector-valued functions of severable variables, which has the form of equality. Then, we apply this formula to some subtle places in the proof of the inverse function theorem. Namely, for a continuously differentiable function f, when $f'(a)$ is invertible, the points a and $b = f(a)$ have open neighborhoods in the form of balls of fixed radii such that f, when restricted to these neighborhoods, is a bijection whose inverse is also continuously differentiable. To know the radii of these balls seems to be something hidden and tricky, but in the proof that we suggest the existence of such neighborhoods is ensured by the continuity of the involved correspondences.

MathEduc Subject Classification: I65

MSC Subject Classification: 97I60

Key words and phrases: Mean value theorem; inverse mapping theorem.

Generalizations of the Lagrange's mean value formula for the vector valued functions have the form of inequality and their accuracy of estimation of the difference $f(b) - f(a)$ may be rather poor.

For example, for a continuous function $f\colon [a,b] \to \mathbf{R}^n$, which is differentiable on (a,b), there exists $\xi \in (a,b)$ such that

$$\|f(b) - f(a)\| \leqslant (b-a)\|f'(\xi)\|,$$

what is the Lagrange's formula for this type of functions. Let us consider an example of this formula when $n = 2$. Take

$$f(x) = (\cos x, \sin x), \quad \text{i.e.} \quad \begin{cases} u_1 = \cos x \\ u_2 = \sin x, \end{cases}$$

then, for $a = 0$, $b = 2\pi$, $f(b) - f(a) = (1,0) - (1,0) = (0,0)$, while for each $x \in (a,b)$,

$$f'(x) = \begin{pmatrix} -\sin x \\ \cos x \end{pmatrix}, \quad \|f'(x)\| = 1.$$

(See [1], pp. 112–113.)

In the case of a differentiable function $f\colon U \to \mathbf{R}^m$, where U is an open and convex subset of $\mathbf{R}^n$ and K is a constant such that $(\forall i)(\forall j)\,|\partial f_i/\partial x_j| \leqslant K$, the inequality

$$\|f(b) - f(a)\| \leqslant (mn)^{1/2} K \|b - a\|$$

is also taken to be the Lagrange's formula.

The aim of this paper is the derivation of a variant of the Lagrange's formula for vector-valued functions which has the form of equality. Then, such a formula is applied to some subtle cases in the proof of the inverse function theorem. For the sake of completeness, we include some related facts, mostly taken from the books [1] and [2].

Finally, let us also add that, when preparing this paper, we have used some written notes of the lectures on analysis of the first author.

1. Operator norm

We will denote by $\mathbf{R}^n$ the Euclidean space of real n-tuples and by $\mathcal{L}(\mathbf{R}^n, \mathbf{R}^m)$ the set of all linear mappings from $\mathbf{R}^n$ into $\mathbf{R}^m$. The latter set, equipped with addition and scalar multiplication, is a linear space.

Variables, mappings and their values at a point will be denoted according to the following correspondences:

$$x \mapsto \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \qquad A \mapsto \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \qquad A(x) \mapsto A \cdot x,$$

$$y = A(x) \mapsto \begin{cases} y_1 = a_{11}x_1 + \cdots + a_{1n}x_n \\ \vdots \\ y_m = a_{m1}x_1 + \cdots + a_{mn}x_n. \end{cases}$$

The space $\mathcal{L}(\mathbf{R}^n, \mathbf{R}^m)$ is normed in the following two standard ways.

For $A \in \mathcal{L}(\mathbf{R}^n, \mathbf{R}^m)$, the number

$$\|A\| = \sup\{\, \|A(x)\| \mid \|x\| = 1\,\}$$

is called *the operator norm* of the mapping (or the operator) A. This norm is taken to be standard because of the following two inequalities.

1.1. $(\forall x)\, \|A(x)\| \leqslant \|A\| \cdot \|x\|$

(where $\|A\|$ is the best constant for this inequality to hold).

1.2. When A is an invertible mapping, then

$$(\forall x)\ \frac{\|x\|}{\|A^{-1}\|} \leqslant \|A(x)\|.$$

On the other hand, we can identify the matrix A with the vector

$$(a_{11}, \dots, a_{1n}, \dots, a_{m1}, \dots, a_{mn}) \in \mathbf{R}^{mn},$$

and, hence, we can also use the *Euclidean norm* in $\mathcal{L}(\mathbf{R}^n, \mathbf{R}^m)$:

$$\|A\|_E = \left\{\sum a_{ij}^2\right\}^{1/2}.$$

These two norms are topologically equivalent, which follows from the inequality

1.3. $\frac{1}{\sqrt{m}}\|A\|_E \leqslant \|A\| \leqslant \|A\|_E \leqslant \sqrt{m}\|A\|.$

In the case when $m = n$, the mapping $A \mapsto \det(A)$ is continuous as a polynomial function with variables a_{ij}. Further on, we know that the mapping A is invertible if and only if $\det(A) \neq 0$. As a consequence of the mentioned continuity, we deduce

1.4. *The set $\mathcal{I}(\mathbf{R}^n, \mathbf{R}^n)$ of all invertible mappings from $\mathcal{L}(\mathbf{R}^n, \mathbf{R}^n)$ is open.*

Taking into account the known formula for the matrix inverse $A^{-1} = (A_{ji}/\Delta)$, and observing coordinatewise continuity of a mapping belonging to $\mathcal{L}(\mathbf{R}^n, \mathbf{R}^n)$ ($(a_{11}, \dots, a_{nn}) \mapsto A_{ji}/\Delta$), we have

1.5. *For $A \in \mathcal{I}(\mathbf{R}^n, \mathbf{R}^n)$, the mapping*

$$A \mapsto A^{-1}$$

is continuous.

2. Differentiability

Recall that a function $f\colon U \to \mathbf{R}^m$, where U is an (open) subset of $\mathbf{R}^n$, is said to be *differentiable* at a point $x \in U$, if there exists a linear mapping $A\colon \mathbf{R}^n \to \mathbf{R}^m$ such that

$$\frac{\|f(x+h) - f(x) - A(h)\|}{\|h\|} \to 0, \quad h \to 0.$$

In this case, we write $f'(x) = A$ and this linear mapping is called the *derivative of* f at the point x.

2.1. *Let U be an (open) subset of $\mathbf{R}^n$, and let $f\colon U \to \mathbf{R}^m$ be a given function. If f is differentiable at $x \in U$ then all the partial derivatives $\frac{\partial f_i}{\partial x_j}(x)$ exist, and $f'(x)$ is represented by the matrix*

$$\begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_n} \end{pmatrix},$$

(which is called the Jacobi matrix of function f).

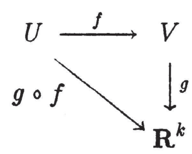
2.2. *Let U be an (open) subset of $\mathbf{R}^n$ and let V be an (open) subset of $\mathbf{R}^m$.*

(a) For differentiable functions $f_1, f_2\colon U \to \mathbf{R}^m$,

$$(f_1 + f_2)'(x) = f_1'(x) + f_2'(x)$$

holds.

(b) For differentiable functions f and g figuring in the following diagram



$$(g \circ f)'(x) = g'(f(x)) \circ f'(x)$$

holds.

A function $f\colon U \to \mathbf{R}^n$ is *continuously differentiable* if the correspondence $x \mapsto f'(x)$ is continuous, i.e. expressed in coordinates, when all the partial derivatives $\dfrac{\partial f_i}{\partial x_j}(x)$ are continuous.

Let us also state the following easy assertion.

2.3. *Let U be an (open) subset of $\mathbf{R}^n$ and let $f\colon U \to \mathbf{R}$ be a differentiable function. If f has an extremum at $x \in U$ then $(\forall j)\ \dfrac{\partial f}{\partial x_j}(x) = 0$.*

3. Lagrange's formula

Observe that for $m = 1$, according to 1.3, the operator norm is equal to the Euclidean one. In this case, i.e. for real-valued functions of several variables, the Lagrange's formula appears in the from of equality.

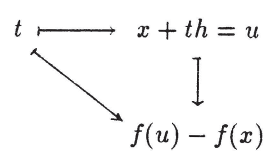
3.1. (*Lagrange's formula*). *Let U be an (open and convex) subset of $\mathbf{R}^n$, let $f\colon U \to \mathbf{R}$ be a differentiable function and let $x, x+h \in U$. Then there exists $\xi \in (0,1)$ such that*

$$f(x+h) - f(x) = [f'(x+\xi h)](h).$$

Proof. The function

$$\varphi(t) = f(x+th) - f(x), \quad (\varphi\colon (-\varepsilon, +\varepsilon) \to \mathbf{R})$$

(where x and h are constants) can be seen as a composite function



Applying Proposition 2.2, we obtain

$$\varphi'(t) = f'(u) \circ h, \quad \left(h = \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix} \text{ is a linear mapping}\right).$$

Therefore

$$f(x+h)-f(x)=\varphi(1)-\varphi(0)=\varphi'(\xi)=f'(x+\xi h)\circ h=[f'(x+\xi h)](h).$$

3.2. *Let U be an (open and convex) subset of $\mathbf{R}^n$, let $f\colon U\to\mathbf{R}$ be a differentiable function and let $x, x+h\in U$. Then*

$$(\forall j)\ \sup_{x\in U}\left|\frac{\partial f}{\partial x_j}(x)\right|\leqslant K \implies |f(x+h)-f(x)|\leqslant K\sqrt{n}\|h\|.$$

Proof. Applying 3.2, 1.1 i 1.3, we get

$$\begin{aligned}|f(x+h)-f(x)|&=|f'(x+\xi h)(h)|\leqslant\|f'(x+\xi h)\|\cdot\|h\|\\ &\leqslant\left(\sum\left[\frac{\partial f}{\partial x_j}(x+\xi h)\right]^2\right)^{1/2}\cdot\|h\|\leqslant K\cdot\sqrt{n}\cdot\|h\|.\end{aligned}$$

The following assertion is the Lagrange's formula for vector-valued functions which follows by applying 3.1 to coordinate functions.

3.3. *Let U be an (open and convex) subset of $\mathbf{R}^n$, let $f\colon U\to\mathbf{R}^m$ be a differentiable function and let $x, x+h\in U$. Then there exist $\xi_1, \dots, \xi_m$ from $(0,1)$, such that*

$$f(x+h)-f(x)=\begin{pmatrix}f_1'(x+\xi_1 h)\\ \vdots\\ f_m'(x+\xi_m h)\end{pmatrix}(h).$$

Proof.

$$\begin{aligned}f(x+h)-f(x)&=\begin{pmatrix}f_1(x+h)-f_1(x)\\ \vdots\\ f_m(x+h)-f_m(x)\end{pmatrix}=\begin{pmatrix}f_1'(x+\xi_1 h)(h)\\ \vdots\\ f_m'(x+\xi_m h)(h)\end{pmatrix}\\ &=\begin{pmatrix}f_1'(x+\xi_1 h)\\ \vdots\\ f_m'(x+\xi_m h)\end{pmatrix}(h).\end{aligned}$$

3.4. *Let U be an (open and convex) subset of $\mathbf{R}^n$, let $f\colon U\to\mathbf{R}^m$ be a differentiable function and let $x, x+h\in U$. Then*

$$(\forall i)(\forall j)\ \sup_{x\in U}\left|\frac{\partial f_i}{\partial x_j}(x)\right|\leqslant K \implies \|f(x+h)-f(x)\|\leqslant K\cdot\sqrt{mn}\cdot\|h\|.$$

Proof. Similar to the one of Proposition 3.2.

Since (in the case $m=n$ the correspondence $A=(a_{ij})\mapsto\det A$ is continuous, hence, when $\det A\neq 0$, according to 1.4, there exists $\varepsilon>0$, such that for each matrix $B=(b_{ij})$ the following implication holds

$$(\forall i)(\forall j)\, b_{ij}\in(a_{ij}-\varepsilon, a_{ij}+\varepsilon)\implies \det B\neq 0.$$

Thus, in this case the matrix A has an ε-neighborhood $\prod(a_{ij}-\varepsilon, a_{ij}+\varepsilon)$ in $\mathbf{R}^{n^2}$ such that all the matrices B contained in it are invertible.

For brevity, we shall denote the matrix

$$\begin{pmatrix} \dfrac{\partial f_1}{\partial x_1}(x+\xi_1 h) & \dots & \dfrac{\partial f_1}{\partial x_n}(x+\xi_1 h) \\ \vdots & \ddots & \vdots \\ \dfrac{\partial f_n}{\partial x_1}(x+\xi_n h) & \dots & \dfrac{\partial f_n}{\partial x_n}(x+\xi_n h) \end{pmatrix}$$

by $A(x,\xi,h)$, where $\xi=(\xi_1,\dots,\xi_n)$ and all $\xi_i\in(0,1)$. Using these denotations and the previous remarks, we can formulate and prove the following assertion.

3.5. *Let U be an (open) subset of $\mathbf{R}^n$ and let a function $f\colon U\to\mathbf{R}^n$ be continuously differentiable. Then*

$$\det f'(x_0)\neq 0 \implies$$
$$\big[(\exists\eta>0)(\forall x)(\forall h)(\forall\xi)\, x\in B(x_0,\eta) \text{ and } x+h\in B(x_0,\eta) \implies \det A(x,\xi,h)\neq 0\big].$$

Proof. The matrix $f'(x_0)$ has an ε-neighborhood, containing matrices with non-zero determinant. The functions $\dfrac{\partial f_i}{\partial x_j}$ are continuous, hence, there exists $\eta>0$, such that

$$\|x_0-y\|<\eta \implies (\forall i)(\forall j)\left|\frac{\partial f_i}{\partial x_j}(x_0)-\frac{\partial f_i}{\partial x_j}(y)\right|<\varepsilon.$$

For $x, x+h\in B(x_0,\eta)$, we have $x+\xi_i h\in B(x_0,\eta)$, for all $\xi_i\in(0,1)$, i.e. $\|x_0-(x+\xi_i h)\|<\eta$. Then,

$$(\forall i)(\forall j)\left|\frac{\partial f_i}{\partial x_j}(x_0)-\frac{\partial f_i}{\partial x_j}(x+\xi_i h)\right|<\varepsilon,$$

and hence $\det A(x,\xi,h)\neq 0$.

4. Inverse function theorem

According to this theorem, a continuously differentiable function f is invertible over some neighborhood of a point x_0, provided that $\det f'(x_0)\neq 0$. This means that the point $f(x_0)$ has also a neighborhood such that, if f is considered as a function between these two neighborhoods, it is injective and surjective, and the inverse function f^{-1} (considered as a mapping between these neighborhoods) is also continuously differentiable. We will first present two assertions which can be extracted from any of proofs of the inverse function theorem, and which have meaning on their own.

4.1. *Let U be an (open) subset of $\mathbf{R}^n$, let a function $f\colon U\to\mathbf{R}^n$ be continuously differentiable and let $x\in U$. If $f'(x)$ is invertible then there exists a neighborhood U_x of the point x, such that the restriction $f|U_x$ is an injection.*

Proof. Take U_x to be the ball $B(x,\eta)$ from Proposition 3.5. For $x' \in B(x,\eta)$, $y = x' + h \in B(x,\eta)$, according to 3.3 and 1.2, we have

$$\|f(y) - f(x')\| = \|f(x'+h) - f(x')\| = \|A(x',\xi,h)(h)\| \geqslant \frac{\|h\|}{\|(A(x',\xi,h))^{-1}\|} > 0,$$

implying that $f(x') \neq f(y)$.

4.2. *Let U be an (open) subset of $\mathbf{R}^n$, let a function $f\colon U \to \mathbf{R}^n$ be continuously differentiable and $(\forall x \in U)\, f'(x)$ is invertible. Then $f[U]$ is an open set.*

Proof. Let a point $b \in f[U]$ be arbitrary. There exists $a \in U$, such that $f(a) = b$. Let $B(a,\eta)$ be the ball over which $\det f' \neq 0$ and f is injective, (4.1). For $\delta \in (0,\eta)$, since f is injective, $f(a) \notin f[S(a,\delta)]$, where $S(a,\delta) = \{\, x \mid \|a - x\| = \delta \,\}$. Let

$$d = \operatorname{dist}\{b, f[S(a,\delta)]\} > 0.$$

We will prove that $B(b,d/2) \subset f[U]$. For arbitrary $y_0 \in B(b,d/2)$, the function

$$\varphi(x) = \|y_0 - f(x)\|^2 = \sum (y_i^0 - f_i(x))^2$$

is continuously differentiable over U. It has the minimum at some point $x_0 \in B(a,\delta)$, because it attains the smallest value on the compact set $\overline{B(a,\delta)}$, and $(\forall x \in S(a,\delta))\, \varphi(x) \geqslant (d/2)^2$ holds, while $\varphi(a) < (d/2)^2$.

According to 2.3,

$$(\forall j)\, \frac{\partial \varphi}{\partial x_j}(x_0) = -\sum_i 2(y_i^0 - f_i(x_0)) \frac{\partial f_i}{\partial x_j}(x_0) = 0 \quad (j = 1,\dots,n).$$

This system is homogenous in $y_i^0 - f_i(x_0)$, so its determinant (after reducing by 2) $\det f'(x_0) \neq 0$. Therefore,

$$(\forall i)\, y_i^0 = f_i(x_0),$$

i.e. $f(x_0) = y_0$. Thus, $y_0 \in f[U]$.

4.3. (*Inverse function theorem*) *Let U be an (open) subset of $\mathbf{R}^n$ and let a function $f\colon U \to \mathbf{R}^n$ be continuously differentiable. If the mapping $f'(a)$ is invertible for some $a \in U$, and $b = f(a)$, then:*

(a) there exist open sets V and W with $a \in V$, $b \in W$, such that the restriction $f\colon V \to W$ is a bijection.

(b) the function $f^{-1}\colon W \to V$ is continuously differentiable and $(f^{-1})'(y) = (f'(x))^{-1}$, $(y = f(x))$.

Proof. (a) Since $\det f'(a) \neq 0$, there exists a neighborhood V of the point a over which $\det f' \neq 0$ and, according to 4.1, over which f is an injection. According to 4.2, the set $f[V] = W$ is open. The function $f^{-1}\colon W \to V$ is continuous, since for an (open) subset A of V we have

$$(f^{-1})^{-1}[A] = f[A],$$

which is, again according to 4.2, an open set.

(b) Let $y \in W$, $f(x) = y$, $A_x = f'(x)$. We will prove that $(f^{-1})'(y) = A_x^{-1}$. Note that by the relation $y + k = f(x + h)$ we have that $h \neq 0$ implies $k \neq 0$, since f is injective; also, $h \to 0$ implies that $k \to 0$, since f and f^{-1} are continuous functions. Now we have

$$\frac{\|f^{-1}(y+k) - f^{-1}(y) - A_x^{-1}(k)\|}{\|k\|} = \frac{\|h - A_x^{-1}(k)\|}{\|k\|} = \frac{\|A_x^{-1}(A_x(h) - k)\|}{\|k\|}$$
$$\leqslant \|A_x^{-1}\| \frac{\|A_x(h) - k\|}{\|k\|} = \|A_x^{-1}\| \frac{\|A_x(h) - k\|}{\|h\|} \cdot \frac{\|h\|}{\|k\|}.$$

The expression $\|A_x^{-1}\|$ does not depend on k, while $\dfrac{\|k - A_x(h)\|}{\|h\|}$ tends to zero as k, i.e. h tends to zero (since f is differentiable). By 3.3,

$$\frac{\|h\|}{\|k\|} = \frac{\|h\|}{\|f(x+h) - f(x)\|} = \frac{\|h\|}{\|A(x,\xi,h)(h)\|},$$

and by 3.5, $A(x, \xi, h)$ is invertible over certain neighborhood of the point x. Hence,

$$\frac{\|h\|}{\|k\|} \leqslant \frac{\|h\|}{\|h\|/\|(A(x,\xi,h))^{-1}\|} \leqslant \|(A(x,\xi,h))^{-1}\|.$$

Since $h \to 0$ implies $A(x, \xi, h) \to A_x$, and since the mapping $A \mapsto A^{-1}$ is continuous, we have that $\|A(x, \xi, h)^{-1}\| \to \|A_x^{-1}\|$. Therefore, $(f^{-1})'(y) = A_x^{-1}$, and so f^{-1} is a differentiable function.

Since all the correspondences

$$y \overset{f^{-1}}{\longmapsto} x \mapsto A_x \mapsto A_x^{-1} = (f^{-1})'(y)$$

are continuous, f^{-1} is also a continuously differentiable function.

We conclude with the following

Open question. Can the Lagrange's formula in the form of equality be deduced for mappings in arbitrary Banach spaces?

REFERENCES

[1] Rudin, W., *Principles of Mathematical Analysis*, McGraw-Hill, Inc., 1976.
[2] Spivak, M., *Calculus on Manifolds*, W.A.Benjamin, Inc, 1965.

M.M.M.: Serbian Academy of Sciences and Arts, Kneza Mihaila 35, 11000 Beograd, Serbia
E-mail: milomar@beotel.rs

Z.K.: Faculty of Mathematics, University of Belgrade, Studentski trg 16, 11000 Beograd, Serbia
E-mail: kadelbur@matf.bg.ac.rs