Implicit differentiation and constrained local extrema in Giac/Xcas

LUKA MAROHNIĆ

Email: marohnicluka@gmail.com

July 2018

ABSTRACT. In this article, an algorithm for finding points of local extrema of a multivariate function under (several) equality constraints, which is implemented in Giac/Xcas computer algebra system, is presented. The first seven sections cover the required theoretical background. The technique of implicit differentiation is proposed for solving the problem as an alternative to the classical Lagrange method. For better classification of critical points a higher order partial derivative test is considered and a counter-example to a method proposed in 2013 is shown. An outline of the algorithm is given in the last section.

Keywords: algorithm, computer algebra, multivariate calculus, Taylor formula, constrained local extrema, Lagrange multipliers, Peano surface

TABLE OF CONTENTS

1. INTRODUCTION

Generally, the problem of finding local extrema of a multivariate function under equality constraints can be formulated as follows.

The original problem. Let $n,m\in\mathbb{N}$ and let $f:\mathbb{R}^{n+m}\to\mathbb{R}$, $\mathbf{G}:\mathbb{R}^{n+m}\to\mathbb{R}^m$ be differentiable functions. Assuming that the Jacobian matrix of $\mathbf{G}$ has maximal rank, find the exact points of local extrema of f subject to $\mathbf{G}\equiv\mathbf{0}$.

To formulate the above problem in a more convenient way, let us assume that

$$f(x_1,x_2,\ldots,x_n,y_1,y_2,\ldots,y_m)=f(\mathbf{x},\mathbf{y}),\quad \mathbf{x}=(x_1,\ldots,x_n)\in\mathbb{R}^n,\mathbf{y}=(y_1,\ldots,y_m)\in\mathbb{R}^m, \tag{1}$$

and, denoting $\mathbf{G}=(G_1,G_2,\ldots,G_m)$,

$$G_j(\mathbf{x},\mathbf{y})=0,\quad j=1,2,\ldots,m, \tag{2}$$

such that the last m columns of the Jacobian matrix

$$\mathbf{J} = \mathbf{G}'(\mathbf{x},\mathbf{y}) = \left(\begin{array}{cccc|cccc} \frac{\partial G_1}{\partial x_1} & \frac{\partial G_1}{\partial x_2} & \cdots & \frac{\partial G_1}{\partial x_n} & \frac{\partial G_1}{\partial y_1} & \frac{\partial G_1}{\partial y_2} & \cdots & \frac{\partial G_1}{\partial y_m} \\ \frac{\partial G_2}{\partial x_1} & \frac{\partial G_2}{\partial x_2} & \cdots & \frac{\partial G_2}{\partial x_n} & \frac{\partial G_2}{\partial y_1} & \frac{\partial G_2}{\partial y_2} & \cdots & \frac{\partial G_2}{\partial y_m} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial G_m}{\partial x_1} & \frac{\partial G_m}{\partial x_2} & \cdots & \frac{\partial G_m}{\partial x_n} & \frac{\partial G_m}{\partial y_1} & \frac{\partial G_m}{\partial y_2} & \cdots & \frac{\partial G_m}{\partial y_m} \end{array}\right) = \left(\begin{array}{cc} \mathbf{G}_\mathbf{x} & \mathbf{G}_\mathbf{y} \end{array}\right)$$

are linearly independent. Note that this formulation may require rearranging the variables of the original problem. Nevertheless, the critical points should always be stored with respect to the original order of variables.

2. Reducing the problem to unconstrained one

To solve the problem (1)-(2) we use the method proposed in [2] which reduces the given problem to an unconstrained one without using Lagrange multipliers.

Assuming that f and $\mathbf{G}$ are both defined and enough differentiable on an open subset $U \subset \mathbb{R}^{n+m}$, let $S = \{(\mathbf{x},\mathbf{y}) \in U : \mathbf{G}(\mathbf{x},\mathbf{y}) = 0 \wedge \det \mathbf{G}_\mathbf{y}(\mathbf{x},\mathbf{y}) \neq 0\}$ and $X_S = \{\mathbf{x} \in \mathbb{R}^n : (\mathbf{x},\mathbf{y}) \in S$ for some $\mathbf{y} \in \mathbb{R}^m\}$. Now from Theorem 6.4.1 [6] (the Implicit Function Theorem) follows the existence of an open subset $V \subset U$ containing S on which a unique continuously differentiable transformation $\mathbf{h} : \mathbb{R}^n \to \mathbb{R}^m$ is defined such that $\mathbf{h}(\mathbf{x}) = \mathbf{y}$ for every $(\mathbf{x},\mathbf{y}) \in V$. Moreover, $\mathbf{G}_\mathbf{y}$ is nonsingular on V.

For $\mathbf{x} = (x_1, x_2, \ldots, x_n) \in X = \{\mathbf{x} \in \mathbb{R}^n : (\mathbf{x},\mathbf{y}) \in V$ for some $\mathbf{y} \in \mathbb{R}^m\}$ we denote $h_j(\mathbf{x}) = y_j$, $j = 1, 2, \ldots, m$. In this context, the variables $y_1, y_2, \ldots, y_m$ depend on $x_1, x_2, \ldots, x_n$ in V which are called *free*. Now we define

$$J(\mathbf{x}) = f(\mathbf{x}, \mathbf{h}(\mathbf{x})), \quad \mathbf{x} = (x_1, x_2, \ldots, x_n) \in X. \tag{3}$$

According to Theorem 3.1 in [2], the solutions to the system

$$\nabla J(\mathbf{x}) = \mathbf{0}, \tag{4}$$

$$\mathbf{G}(\mathbf{x},\mathbf{y}) = \mathbf{0}, \tag{5}$$

which can be also written as a system of $n+m$ equations

$$\begin{aligned} \frac{\partial J}{\partial x_i}(\mathbf{x}) &= 0, \quad i = 1, 2, \ldots, n, \\ G_j(\mathbf{x},\mathbf{y}) &= 0, \quad j = 1, 2, \ldots, m, \end{aligned}$$

are the critical points in V of f subject to (2).

Remark 1. To avoid any generality loss in the original problem, critical points should be searched separately for each distinct partition of variables into the classes $\mathbf{x}$ (independent) and $\mathbf{y}$ (dependent) such that the the matrix $\mathbf{G}_\mathbf{y}$ is nonsingular [2]. All distinct formulations of the original problem in form (1)-(2), which form the set of *instances* of the problem, must be solved independently and the results collected into a single set. The latter, still, does not necessarily represent the whole set of critical points; there may be points in the domain of f for which $\mathbf{G} \equiv \mathbf{0}$ but some partial derivatives of J are undefined in all instances (see Example 3). In such cases one may consider to apply the classical method of LAGRANGE [3] in attempt to find the remaining critical points. Note that both methods may fail; for a simple example see [4].

3. Obtaining the critical points

Using the chain rule and the total derivative formula (see Corollary 5.4.4 in [6]), we obtain

$$\frac{\partial J}{\partial x_i} = \frac{\partial f}{\partial x_i} + \sum_{j=1}^{m} \frac{\partial f}{\partial y_j} \cdot \frac{\partial h_j}{\partial x_i}, \quad i = 1, 2, \ldots, n. \tag{6}$$

Hence, to compute $\frac{\partial J}{\partial x_i}$ for all $i = 1, 2, \ldots, n$ we need to determine the first order partial derivatives (i.e. the Jacobian) of $\mathbf{h}$. That is easy since the Theorem 6.4.1 [6] states that

$$\mathbf{h}' = -\mathbf{G}_\mathbf{y}^{-1} \cdot \mathbf{G}_\mathbf{x}. \tag{7}$$

However, the method presented here relies on the higher derivatives of J too, hence we chose a different approach which is also used in [2]. Let us define

$$\mathbf{K}(\mathbf{x})=\mathbf{G}(\mathbf{x},\mathbf{h}(\mathbf{x})),\quad \mathbf{x}\in X.$$

The function $\mathbf{K}=(K_1,K_2,\ldots,K_m)$ is constant on X_S, which implies that its (higher) partial derivatives vanish at any point of the interior of X_S. Since the partial derivatives are continuous because of the assumed differentiability of $\mathbf{G}$ (see the proof of Theorem 5.4.9 [6]), the last equation is valid for all $\mathbf{x}\in X_S$ (remember that the function $\mathbf{G}$ resp. $\mathbf{K}$ is differentiable in the open subset $V\supset S$ resp. $X\supset X_S$). Therefore in X_S (where the critical points reside) we have

$$\frac{\partial K_j}{\partial x_i}=\frac{\partial G_j}{\partial x_i}+\sum_{k=1}^{m}\frac{\partial G_j}{\partial y_k}\cdot\frac{\partial h_k}{\partial x_i}=0, \tag{8}$$

which implies

$$\sum_{k=1}^{m}\frac{\partial G_j}{\partial y_k}\cdot\frac{\partial h_k}{\partial x_i}=-\frac{\partial G_j}{\partial x_i} \tag{9}$$

for all $i=1,2,\ldots,n$ and $j=1,2,\ldots,m$. The equations (9) form a linear system of order $n\,m$ which can be solved by applying the usual methods, providing the derivatives $\frac{\partial h_j}{\partial x_i}$ of order 1 explicitly for all i,j. These values are substituted in each of the equations in (6), yielding the required derivatives $\frac{\partial J}{\partial x_i}$. Subsequently, the critical points are obtained by solving the system of equations (4)-(5).

Example 2. To illustrate the importance of Remark 1, we attempt determine all critical points of the function $f(x,y)=x^2\,y^2$ subject to $x^2+y^2=1$. It can be shown that the set of critical points of f is

$$C=\left\{(-1,0),(1,0),(0,-1),(0,1),\left(\tfrac{1}{\sqrt{2}},\tfrac{1}{\sqrt{2}}\right),\left(-\tfrac{1}{\sqrt{2}},\tfrac{1}{\sqrt{2}}\right),\left(-\tfrac{1}{\sqrt{2}},-\tfrac{1}{\sqrt{2}}\right),\left(\tfrac{1}{\sqrt{2}},-\tfrac{1}{\sqrt{2}}\right)\right\}.$$

The first four are the points of local minima while the last four are the points of local maxima (see Figure[1] 1). Also, it is easy to obtain the parametrization of the intersection of the surfaces $z=x^2\,y^2$ and $x^2+y^2=1$, which is

$$\begin{cases} x=\cos t,\\ y=\sin t,\\ z=\frac{1}{4}\sin^2 2t,\end{cases}\qquad t\in[-\pi,\pi],$$

allowing the local extrema to be easily visualized in Giac by using the `plotparam` command. To find the critical points using the the proposed method, we first formulate the problem such that x is the free and y the dependent variable. Therefore, $\mathbf{J}=(\ \mathbf{G}_x\ \ \mathbf{G}_y\ )=(\ 2\,x,\ 2\,y)$ which together with (7) implies

$$\frac{\mathrm{d}h}{\mathrm{d}x}=-\frac{x}{y}.$$

Now from (6) we have

$$\frac{\mathrm{d}J}{\mathrm{d}x}=\frac{\partial f}{\partial x}+\frac{\partial f}{\partial y}\cdot\frac{\mathrm{d}h}{\mathrm{d}x}=2\,x\,y^2-2\,x^2\,y\cdot\frac{x}{y}=2\,x\,y^2-2\,x^3=2\,x\,(y^2-x^2).$$

Solving the system of equations $2\,x\,(y^2-x^2)=0$ and $x^2+y^2=1$ using the command `solve` in Giac [5], we obtain only the last six critical points from C; the first two, $(-1,0)$ and $(1,0)$, obviously do not satisfy the first equation because $\frac{\mathrm{d}h}{\mathrm{d}x}$ is not defined when $y=0$. However, there is an alternative interpretation of the given problem, where y is the free and x the dependent variable. In that case we have $\frac{\mathrm{d}h}{\mathrm{d}x}=-\frac{y}{x}$ and $\frac{\mathrm{d}J}{\mathrm{d}y}=2\,y\,(x^2-y^2)$. Solving the system $2\,y\,(x^2-y^2)=0$ and $x^2+y^2=1$, we also get six critical points but this time the solution contains $(-1,0)$ and $(1,0)$, while $(0,-1)$ and $(0,1)$ are missing (since now $\frac{\mathrm{d}h}{\mathrm{d}x}$ is not defined when $x=0$). We conclude that it is necessary to consider both formulations to obtain the complete set of critical points.

1. All plots in this text are generated by Wolfram Alpha (http://www.wolframalpha.com/).

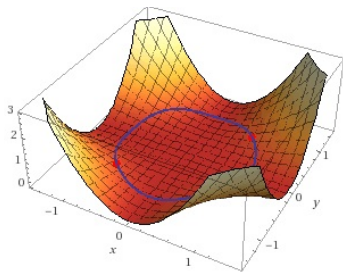

Figure 1. $f(x,y)=x^2\,y^2$ on the unit sphere S^1

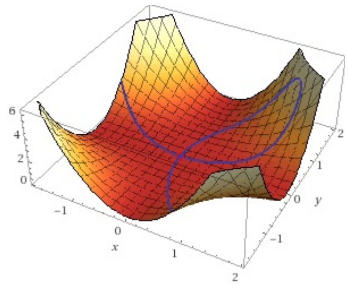

Figure 2. $f(x,y)=x^2\,y^2$ on the folium of Descartes

Example 3. We search for local extrema of the function $f(x,y)=x^2\,y^2$ on the folium of Descartes defined by $g(x,y)=x^3+y^3-3\,x\,y=0$ (see Figure 2). Let y be the dependent variable. Now we have

$$\mathbf{J}=(\ \mathbf{G}_x\ \ \mathbf{G}_y\ )=\left(\ 3\,x^2-3\,y,\ 3\,y^2-3\,x\right),$$

hence $\frac{\mathrm{d}h}{\mathrm{d}x}=-\frac{x^2-y}{y^2-x}$ and $\frac{\mathrm{d}J}{\mathrm{d}x}=\frac{\partial f}{\partial x}+\frac{\partial f}{\partial y}\cdot\frac{\mathrm{d}h}{\mathrm{d}x}=2\,x\,y^2-2\,x^2\,y\,\frac{x^2-y}{y^2-x}=\frac{2\,x^4\,y-2\,x\,y^4}{x-y^2}$. The derivative $\frac{\mathrm{d}J}{\mathrm{d}x}$ is not defined in origin $(0,0)$, but it is easy to show the latter is the point of local minimum of f subject to $g\equiv 0$. If we declare x to be the dependent variable, we obtain a similar situation; therefore, as only the local maximum at $\mathbf{c}_1=\left(\frac{3}{2},\frac{3}{2}\right)$ is detected by the above method, the critical point $(0,0)$ and its classification must be obtained by some other means. We apply the method of Lagrange [3] and define the corresponding Lagrange function $\mathcal{L}$ by

$$\mathcal{L}(x,y,\lambda)=f(x,y)-\lambda\,g(x,y),$$

where λ is the Lagrange multiplier. The system of equations $\nabla\mathcal{L}(x,y,\lambda)=\mathbf{0}$, $g(x,y)=0$ for critical points is solved by using the commands `grad` and `solve` in Giac [5]. We obtain exactly two distinct solutions, $\mathbf{c}_1$ for $\lambda=3$ and $\mathbf{c}_2=(0,0)$ for any $\lambda\in\mathbb{R}$. Note that classifying $\mathbf{c}_2$ must be done with respect to $\mathcal{L}$ (see Section 5).

4. Computing the partial derivatives of higher order

The next step is to classify the critical points. At this point, higher derivatives of J are needed. To compute these, the derivatives of $\mathbf{h}$ of the same order are required. To allow the $\frac{\partial^2}{\partial x_i\,\partial x_j}=\frac{\partial^2}{\partial x_j\,\partial x_i}$ exchange in higher order partial derivatives, we assume that f and $\mathbf{G}$ are both C^2 functions.

Since all derivatives of K_j vanish in X_S for all $j=1,2,\ldots,m$, we proceed by obtaining second order derivatives of $\mathbf{K}$ by differentiating the equations (8). We obtain

$$\begin{aligned}\frac{\partial^2 K_p}{\partial x_i\,\partial x_j}&=\frac{\partial^2 G_p}{\partial x_i\,\partial x_j}+\sum_{q=1}^m\frac{\partial^2 G_p}{\partial x_i\,\partial y_q}\cdot\frac{\partial h_q}{\partial x_j}+\\&\quad+\sum_{k=1}^m\left(\left(\frac{\partial^2 G_p}{\partial x_j\,\partial y_k}+\sum_{q=1}^m\frac{\partial^2 G_p}{\partial y_q\,\partial y_k}\cdot\frac{\partial h_q}{\partial x_j}\right)\cdot\frac{\partial h_k}{\partial x_i}+\frac{\partial G_p}{\partial y_k}\cdot\frac{\partial^2 h_k}{\partial x_i\,\partial x_j}\right)\\&=\frac{\partial^2 G_p}{\partial x_i\,\partial x_j}+\sum_{q=1}^m\left(\frac{\partial^2 G_p}{\partial x_i\,\partial y_q}\cdot\frac{\partial h_q}{\partial x_j}+\frac{\partial^2 G_p}{\partial x_j\,\partial y_k}\cdot\frac{\partial h_k}{\partial x_i}\right)+\\&\quad+\sum_{k=1}^m\frac{\partial h_k}{\partial x_i}\sum_{q=1}^m\frac{\partial^2 G_p}{\partial y_q\,\partial y_k}\cdot\frac{\partial h_q}{\partial x_j}+\sum_{k=1}^m\frac{\partial G_p}{\partial y_k}\cdot\frac{\partial^2 h_k}{\partial x_i\,\partial x_j}=0,\end{aligned}$$

which implies

$$\sum_{k=1}^m\frac{\partial G_j}{\partial y_k}\cdot\frac{\partial^2 h_k}{\partial x_i\,\partial x_j}=-\frac{\partial^2 G_p}{\partial x_i\,\partial x_j}-\sum_{q=1}^m\left(\frac{\partial^2 G_p}{\partial x_i\,\partial y_q}\cdot\frac{\partial h_q}{\partial x_j}+\frac{\partial^2 G_p}{\partial x_j\,\partial y_q}\cdot\frac{\partial h_k}{\partial x_i}\right)-\sum_{k=1}^m\frac{\partial h_k}{\partial x_i}\sum_{q=1}^m\frac{\partial^2 G_p}{\partial y_q\,\partial y_k}\cdot\frac{\partial h_q}{\partial x_j}\tag{10}$$

for all $i,j=1,2,\ldots,n$ and $p=1,2,\ldots,m$. Substituting the first order derivatives $\mathbf{h}$ obtained by solving the system (9) in the equations (10) yields the linear system of order $n^2 m$ with the second order partial derivatives of $\mathbf{h}$ as indeterminates. (The order of the system can be reduced to $\frac{1}{2} m\,n\,(n+1)$ since $\frac{\partial^2}{\partial x_i\,\partial x_j}=\frac{\partial^2}{\partial x_j\,\partial x_i}$ holds by assumption). When these are computed, the formula

$$\frac{\partial^2 J}{\partial x_i\,\partial x_j}=\frac{\partial^2 f}{\partial x_i\,\partial x_j}+\sum_{q=1}^{m}\left(\frac{\partial^2 f}{\partial x_i\,\partial y_q}\cdot\frac{\partial h_q}{\partial x_j}+\frac{\partial^2 f}{\partial x_j\,\partial y_q}\cdot\frac{\partial h_k}{\partial x_i}\right)+\sum_{k=1}^{m}\frac{\partial h_k}{\partial x_i}\sum_{q=1}^{m}\frac{\partial^2 f}{\partial y_q\,\partial y_k}\cdot\frac{\partial h_q}{\partial x_j}+\sum_{k=1}^{m}\frac{\partial f}{\partial y_k}\cdot\frac{\partial^2 h_k}{\partial x_i\,\partial x_j},$$

obtained by differentiating the i-th equation in (6) with respect to x_j, yields the second order partial derivatives of J for each $i=1,2,\ldots,n$ and $j=1,2,\ldots,m$.

We could now proceed inductively by computing the third order derivatives, fourth order derivatives and so on, until reaching the higher derivative we require. See [2] for several practical examples demonstrating this method.

5. The second partial derivative test

To classify the critical points of J we use the second order partial derivative test. According to Theorem 3.2 [2], we look at the set E of distinct eigenvalues $e_1,e_2,\ldots,e_k$ of the $n\times n$ Hessian matrix

$$\operatorname{Hess}(J):=\left[\frac{\partial^2 J}{\partial x_i\,\partial x_j}\right]_{1\leqslant i,j\leqslant n}$$

at the critical point $\mathbf{c}=(\mathbf{x}_0,\mathbf{y}_0)\in\mathbb{R}^{n+m}$. There are four possibilities:

1. if $e>0$ for all $e\in E$ (i.e. if the Hessian is positive definite), then $\mathbf{c}$ is a local minimum,
2. if $e<0$ for all $e\in E$ (i.e. if the Hessian is negative definite), then $\mathbf{c}$ is a local maximum,
3. if $0\notin E$ and there exist $e_i,e_j\in E$ such that $e_i\,e_j<0$, then $\mathbf{c}$ is a saddle point,
4. if $0\in E$, the test is inconclusive.

Note that, if the method of Lagrange was used to find some critical points, as suggested for some special cases in Remark 1, we have to compute the Hessian of the Lagrange function $\mathcal{L}$ in order to apply the second derivative test to these, which is a $(n+2\,m)\times(n+2\,m)$ matrix $H_{\mathcal{L}}(\boldsymbol{\lambda},\mathbf{x})$ called the *bordered*[2] Hessian:

$$H_{\mathcal{L}}(\boldsymbol{\lambda},\mathbf{x})=\begin{pmatrix}\mathbf{0} & -\mathbf{J}\\ -\mathbf{J}^T & H_{\mathcal{L}}(\mathbf{x})\end{pmatrix},$$

where $\boldsymbol{\lambda}=(\lambda_1,\lambda_2,\ldots,\lambda_m)$ are the Lagrange multipliers, $\mathbf{x}\in\mathbb{R}^{n+m}$, $\mathbf{J}=\mathbf{G}'$ and $H_{\mathcal{L}}(\mathbf{x})$ is the $(n+m)\times(n+m)$ Hessian of $\mathcal{L}$ with multipliers treated as constants.

To perform the second partial derivative test in this case, one needs to examine the signs s_{2m+1}, $s_{2m+2},\ldots,s_{n+2m}$ of the last n principal minors of $H_{\mathcal{L}}(\boldsymbol{\lambda}_0,\mathbf{x}_0)$ for every critical point $(\boldsymbol{\lambda}_0,\mathbf{x}_0)$ of $\mathcal{L}$ (the k-th principal minor is composed of the truncated first k rows and columns)[3]:

1. if $s_k=(-1)^{k-m}$ for all $0<k-2\,m\leqslant n$, then $\mathbf{x}_0$ is a local maximum,
2. if $s_k=(-1)^m$ for all $0<k-2\,m\leqslant n$, then $\mathbf{x}_0$ is a local minimum,
3. if $s_k\neq 0$ for all $0<k-2\,m\leqslant n$ but the sequence of signs do not fit any of the above two patterns, then $\mathbf{x}_0$ is a saddle point,
4. if $s_k=0$ for some $0<k-2\,m\leqslant n$, the test is inconclusive.

6. Critical point classification by using higher order derivatives

If the second partial derivative test is inconclusive for some critical point $(\mathbf{x}_0,\mathbf{y}_0)$, we can use higher order derivatives in attempt to classify it, unless Lagrange multipliers are used (in the latter case the bordered Hessian is the only tool for classification). Firstly we define a certain family of multivariate polynomials which exhibits a set of useful properties.

2. https://en.wikipedia.org/wiki/Hessian_matrix#Bordered_Hessian
3. http://web.sgh.waw.pl/~mantosi/MO/materialy2.pdf

Definition 4. *A polynomial* $\mathbf{p}:\mathbb{R}^n\to\mathbb{R}$ *is* homogeneous *if* $\mathbf{p}(t\,\mathbf{x})=t^d\,\mathbf{p}(\mathbf{x})$ *for every* $\mathbf{x}\in\mathbb{R}^n$ *and* $t\in\mathbb{R}$, *where* $d=\deg\mathbf{p}$.

Obviously, $\mathbf{p}(\mathbf{x})=0$ if and only if $\mathbf{x}=\mathbf{0}$. Next we introduce the following types of homogeneous polynomials, which will be useful later.

Definition 5. *Homogeneous polynomial* $\mathbf{p}$ *is* positive semidefinite *resp.* negative semidefinite *if* $\mathbf{p}(\mathbf{x})\geqslant 0$ *resp.* $\mathbf{p}(\mathbf{x})\leqslant 0$ *for all* $\mathbf{x}\in\mathbb{R}^n$. *If* $\mathbf{p}(\mathbf{x})>0$ *resp.* $\mathbf{p}(\mathbf{x})<0$ *for all* $\mathbf{x}\in\mathbb{R}^n\setminus\{\mathbf{0}\}$, *then* $\mathbf{p}$ *is* positive definite *resp.* negative definite.

A practical method for determining whether a homogeneous polynomial $\mathbf{p}$ is positive resp. negative (semi)definite [2] is given in Section 7.

The technique of implicit differentiation of J presented in Section 4 allows us, assuming that f and $\mathbf{G}$ are $k+1$ times differentiable for $k\geqslant 2$, to apply the Taylor fromula (see Theorem 5.4.8 [6]) for J at the point $\mathbf{x}_0$. We obtain

$$J(\mathbf{x})-J(\mathbf{x}_0)=\sum_{r=1}^{k}\frac{1}{r!}\left(D^{(r)}_{\mathbf{x}_0,\mathbf{y}_0}J\right)(\mathbf{x}-\mathbf{x}_0)+\frac{1}{(k+1)!}\left(d^{(k+1)}_{\boldsymbol{\xi}}J\right)(\mathbf{x}-\mathbf{x}_0), \tag{11}$$

where the point $\boldsymbol{\xi}\notin\{\mathbf{x},\mathbf{x}_0\}$ lies somewhere on the line segment connecting $\mathbf{x}$ and $\mathbf{x}_0$ and $D^{(r)}_{\mathbf{x}_0,\mathbf{y}_0}J$ is defined by

$$\left(D^{(r)}_{\mathbf{x}_0,\mathbf{y}_0}J\right)(\mathbf{x})=\sum_{\mathbf{r}}\frac{r!}{r_1!\,r_2!\cdots r_n!}\cdot\frac{\partial^r J}{\partial x_1^{r_1}\,\partial x_2^{r_2}\cdots\partial x_n^{r_n}}(\mathbf{x}_0,\mathbf{y}_0)\cdot x_1^{r_1}\,x_2^{r_2}\cdots x_n^{r_n}, \tag{12}$$

where $\sum_{\mathbf{r}}$ indicates the summation over all n-tuples $\mathbf{r}=(r_1,r_2,\ldots,r_n)$ of nonnegative integers such that $r_1+r_2+\cdots+r_n=r$. From the last equation it follows that $D^{(r)}_{\mathbf{x}_0,\mathbf{y}_0}J$ is a homogeneous polynomial.

Now we can state the higher derivative test. Assume that $D^{(r)}_{\mathbf{x}_0,\mathbf{y}_0}J\equiv 0$ for $r=1,2,\ldots,k-1$ and $\mathbf{p}_k:=D^{(k)}_{\mathbf{x}_0,\mathbf{y}_0}J\not\equiv 0$. According to Theorem 5.4.10 [6], the critical point $(\mathbf{x}_0,\mathbf{y}_0)$ is a

1. saddle point if $\mathbf{p}_k$ is not semidefinite (in particular when k is odd),
2. local minimum if $\mathbf{p}_k$ is positive definite,
3. local maximum if $\mathbf{p}_k$ is negative definite.

The test is inconclusive when $\mathbf{p}_k$ is semidefinite but not positive nor negative definite. However, if $\mathbf{p}_k$ is positive semidefinite resp. negative semidefinite there is still a possibility that $(\mathbf{x}_0,\mathbf{y}_0)$ is a point of local minimum resp. local maximum of f. This information is reported by our algorithm.

7. On the difficulty of classifying critical points correctly

Theorem 6.5 [2] was proposed as an alternative, more detailed method of classifying critical points using homogeneous polynomials, supposedly handling the case in which the first non-vanishing polynomial $\mathbf{p}_k$ is semidefinite. However, it turns to be incorrect as we have found a counterexample: the proposed chain of reasoning leads to a false conclusion when classifying the critical point $(0,0)$ of the *Peano surface* (see Figure 3 and also Fig. 1–2 in [1]) defined by

$$f(x,y)=(2\,x^2-y)\,(y-x^2). \tag{13}$$

This example was devised by Giuseppe Peano to demonstrate the falsity of some of the hitherto accepted criteria for maxima and minima of functions of several variables [1]. The Maclaurin expansion of f is trivial:

$$T_4(x,y)=-y^2+3\,x^2\,y-2\,x^4.$$

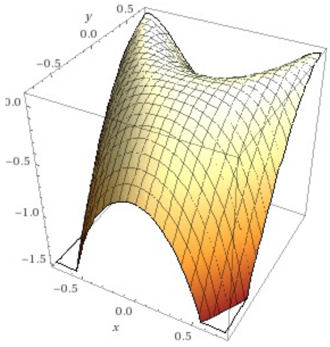

Figure 3. Peano surface

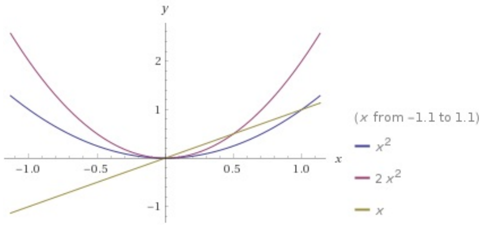

Figure 4. the set of zeros of the function f defined in (13)

According to Theorem 6.5 [2], we find the first non-vanishing homogeneous polynomial. That happens to be $\mathbf{p}_2(x,y)=-y^2$. Computing its image on the unit sphere S^1, we conclude that it attains the maximal value 0 in the points $(-1,0)$ and $(1,0)$, which leads to the case c_5) of Theorem 6.5. The next homogeneous polynomial which does not vanish in $\mathbf{p}_2^{-1}(0)\cap S^1=\{(-1,0),(1,0)\}$ is $\mathbf{p}_4(x,y)=-2\,x^4$, which has negative values on $\mathbf{p}_2^{-1}(0)\cap S^1$. This leads to case c_{53}) which classifies the point $(0,0)$ as being a point of strict local maximum. However, this conclusion coincides with the one often being falsely "deduced" in the past. The truth is that $(0,0)$ is a saddle point, since f is a polynomial, $(0,0)$ is its only critical point, $f(0,0)=0$, $f\big(\frac{3}{4},\frac{3}{4}\big)=\frac{9}{128}$ and $f\big(\frac{3}{4},-\frac{3}{4}\big)=-\frac{315}{128}$. This means that, unfortunately, the proof of Theorem 6.5 [2] is flawed.

The fact that was overlooked in the proof is that $\mathbf{x}_0$ may be a saddle point for the function J even if the examination of its directional derivatives, with respect to each possible direction, suggests that $\mathbf{x}_0$ is a point of local extremum. Simply, there is not enough data to assert that $J(\mathbf{x})-J(\mathbf{x}_0)$ will not change sign when $\mathbf{x}$ passes through $\mathbf{x}_0$ when moving along a continuous (not necessarily linear) curve in $\mathbb{R}^n$ which contains $\mathbf{x}_0$. That is precisely the reason why the false classification of the critical point of Peano curve occurred.

Observe that the function f defined in (13) has zero values on two parabolas $y=x^2$ and $y=2\,x^2$ shown in Figure 4. The narrow area enclosed by the parabolas, but not including them, is the domain on which f is strictly positive. On the interior of the complement of that set f is strictly negative. It is obvious that the restriction of f to an arbitrary line passing through origin attains strict local maximum in $(0,0)$, as demonstrated in Figure 4 for the line $x-y=0$. However, no more data can be squeezed out the Taylor expansion of f; that is unfortunate because, as verified by this example, we are far from being certain that $(0,0)$ is a point of local maximum of f on $\mathbb{R}^2$. The rigorous, valid check is to test the critical point $(0,0)$ with the restriction of f to each smooth curve passing through origin; hence there is little hope that Theorem 6.5 [2] could be easily fixed.

Nevertheless, the approach proposed in [2] contains some useful ideas. For example, examining the image of a nontrivial homogeneous polynomial $\mathbf{p}:\mathbb{R}^n\to\mathbb{R}$ is significantly simplified by using the fact, stated in Lemma 6.3 [2], that the image of the restriction of $\mathbf{p}$ to the n-sphere S^{n-1} is a segment $[a,b]\in\mathbb{R}$ such that $\mathbf{p}$ is

1. positive semidefinite resp. positive definite if and only if $0=a<b$ resp. $0<a<b$,

2. negative semidefinite resp. negative definite if and only if $a<b=0$ resp. $a<b<0$.

These results are used in our algorithm when applying the higher derivative test. The proof is straightforward and relies on the fact that, by Definition 4, for $\mathbf{x}\in\mathbb{R}^n\setminus\{\mathbf{0}\}$ we have

$$\mathbf{p}(\mathbf{x})=\mathbf{p}\bigg(\frac{\|\mathbf{x}\|}{\|\mathbf{x}\|}\,\mathbf{x}\bigg)=\|\mathbf{x}\|^d\,\mathbf{p}\bigg(\frac{1}{\|\mathbf{x}\|}\,\mathbf{x}\bigg),$$

where $d=\deg\mathbf{p}$. Note that $\frac{1}{\|\mathbf{x}\|}\,\mathbf{x}\in S^{n-1}$, hence the image of $\mathbf{p}$ is completely determined by the image of the restriction $\mathbf{p}|_{S^{n-1}}$ (except for $\mathbf{x}=\mathbf{0}$, but recall that $\mathbf{p}(\mathbf{0})=0$ always holds).

8. The algorithm

In this section we outline the algorithm, implemented in Giac as the command `extrema` [5], which attempts to solve the problem 1 using implicit differentiation or the method of Lagrange (the former is the default, while the latter may be specified by passing the option `lagrange` to the command `extrema`).

The branch of the algorithm which uses Lagrange multipliers is rather straightforward: after computing critical points, the bordered Hessian is used as the only method of classification. The pseudocode of this branch is given as Algorithm 1.

Algorithm 1
Input: the original constrained local extrema problem 1
Output: the set C of classified critical points

begin
  $\boldsymbol{\lambda} \leftarrow (\lambda_1, \lambda_2, \ldots, \lambda_m)$, the symbols representing Lagrange multipliers
  $\mathcal{L}(\boldsymbol{\lambda}, \mathbf{x}) \leftarrow f(\mathbf{x}) - \boldsymbol{\lambda} \cdot \mathbf{G}(\mathbf{x})$, the Lagrangian function ($\cdot$ denotes the scalar product)
  solve the system $\nabla \mathcal{L} = \mathbf{0}$, $\mathbf{G} \equiv \mathbf{0}$ for critical points and add them to the set C
  label each critical point with unclassified
  $H_{\mathcal{L}} \leftarrow$ the bordered Hessian of $\mathcal{L}$ w.r.t. $\boldsymbol{\lambda}, \mathbf{x}$
  for each critical point $\mathbf{x}_0 \in C$ **do**
    for $k = 1, 2, \ldots, n$ **do**
      $M \leftarrow$ the $(2\,m + k)$-th principal minor of $H_{\mathcal{L}}(\boldsymbol{\lambda}_0, \mathbf{x}_0)$
      $s \leftarrow \operatorname{sgn}(\det M) \in \{-1, 0, 1\}$
      if $s = 0$ **then**
        label $\mathbf{x}_0$ with unclassified
        break
      end
      if $\mathbf{x}_0$ is labeled with saddle **then**
        continue
      end
      if $s = (-1)^m$ and $\mathbf{x}_0$ is not labeled with strict_max **then**
        label $\mathbf{x}_0$ with strict_min
      else if $s = (-1)^{m+k}$ and $\mathbf{x}_0$ is not labeled with strict_min **then**
        label $\mathbf{x}_0$ with strict_max
      else label $\mathbf{x}_0$ with saddle
      end
    od
  od
end

Here we analyze the branch of the algorithm which uses implicit differentiation. The procedure, given as Algorithm 2, is applied to every distinct formulation in form (1)-(2) of the original problem 1 (see Remark 1). At each instance, the critical points are added to the common set C and labeled with unclassified. The set C is implemented as a static map with points as keys and labels as values. Unless $L = 1$, an attempt is made to classify critical points: each of them is either a strict local minimum (strict_min), a strict local maximum (strict_max), a saddle point (saddle), a possible (non-strict) local minimum (min?) or a possible (non-strict) local maximum (max?). Also, it may remain unclassified.

The partial derivatives of J are found in succession: the first order derivatives are needed for computing the critical points, but also for obtaining the second order derivatives when applying the Hessian test. If this attempt to classify a critical point fails, the third order derivatives are computed and so on, until the classification succeeds or the limit L is reached. Hence the partial derivatives are always stored once computed.

The implementation of the subroutine for finding the partial derivatives of the implicit functions $h_1, h_2, \ldots, h_m$, which are needed for computing the partial derivatives of J, is realized using the simplest data structures—integer arrays—alongside the linear system solver already available in

Algorithm 2
Input: the constrained local extrema problem (1)-(2), limit $L \in \mathbb{N}$
Output: the set C of (classified) critical points

```
begin
    compute the first order partial derivatives of J
    solve the system ∇J = 0, G ≡ 0 for critical points and add them to the set C
    label each newly found critical point with unclassified
    if L = 1 then return
    end
    compute the second order partial derivatives of J
    H_J ← the Hessian of J
    for each critical point c = (x_0, y_0) found in this instance with label unclassified do
        {e_1, e_2, ..., e_k} ← the set of distinct eigenvalues of H_J(c)
        e ← 0
        for i = 1, 2, ..., k do
            if e_i = 0 then
                label c with unclassified
                break
            else if e = 0 then
                e ← e_i
                if e > 0 then label c with strict_min
                else label c with strict_max
                end
            else if e_i e < 0 then label c with saddle
            end
        od
        if c is labeled unclassified then
            for k = 2, 3, ..., L do
                compute the k-th order partial derivatives of J unless already available
                p_k ← the k-th hom. polynomial in the Taylor expansion of J w.r.t. c
                if p_k ≢ 0 then
                    if k is odd then label c with saddle
                    else
                        [a, b] ← p_k(S^{n-1}), the image of p_k|_{S^{n-1}} (a = b is possible)
                        if a < 0 < b then label c with saddle
                        else if a > 0 then label c with strict_min
                        else if b < 0 then label c with strict_max
                        else if a = 0 then label c with min?
                        else if b = 0 then label c with max?
                        end
                    end
                    break
                end
            od
        end
    od
end
```

Giac [5]. It is possible because each term in the equations (9), (10) and the analogous equations of higher order is a product of an integer A, a partial derivative $\frac{\partial^r G_j}{\partial x_1^{r_1}\,\partial x_2^{r_2}\cdots\partial x_n^{r_n}}$ for some $j \in \{1,2,\ldots,m\}$ and $r = r_1 + r_2 + \cdots + r_n$, and possibly a number of nonnegative powers of partial derivatives of h_1, $h_2, \ldots, h_j$. Each of these factors, except A, can be represented as an array of nonnegative integers. These arrays together with the number A are gathered in a larger structure representing the entire term. Differentiating these terms symbolically was relatively easy to implement since only three rules of differentiation are used: the total derivative formula (6), the chain rule and the product rule.

The homogeneous polynomial $\mathbf{p}_k$ for $k=2,3,\ldots,L$ is computed using the formula (12), which is easy once all partial derivatives of J of order k are available. To obtain a and b such that $\mathbf{p}_k(S^{n-1})=[a,b]$ (which is always possible as $\mathbf{p}_k$ is continuous, so it maps the connected compact subset $S^{n-1}\subset\mathbb{R}^n$ to a connected compact subset of $\mathbb{R}$, hence $[a,b]$, note that $a=b$ is allowed), we use the same subroutine utilized by `minimize` and `maximize` commands in Giac [5], which computes the exact global extrema of a multivariate continuous function on a compact domain by applying Karush-Kuhn-Tucker (KKT) conditions. Because of the polynomial input, solving the system of conditions is straightforward.

The final solution is recorded in the set C, which (in the end) contains all found critical points paired with their class labels. The local minima and local maxima, kept in two separate lists, are returned as the solution to the original problem. The rest of data collected in the classification process is dumped to the log stream to provide useful information about saddle points and possible local extrema. Note that setting L to 1 will force the algorithm to skip the classification process (as the order of partial derivatives that are allowed to be computed is limited by L). In that case, only the set of classification points is returned.

It is recommended to set the limit L of the order of partial derivatives to 10 or lower, as the complexity of the algorithm for computing the partial derivatives of order $k\geqslant 2$ is greater than $O(n^{2k}m^2)$, which is the necessary (but not sufficient) number of operations performed when solving a linear system of order $n^k m$.

Conclusion

The algorithm described in this text first reduces the original problem to an unconstrained one. Then, two attempts are made: the first one to find the exact locations of the critical points and the second one to classify them. Note that both may fail. For example, some functions have infinitely many critical points. On the other hand, in some cases computing higher partial derivatives—as seen in PEANO's example discussed in Section 7—may not be enough to determine if a critical point is indeed a local extremum. Also, since our algorithm performs exact computations, classifying a critical point can only use the data available at that precise point. Inspecting the function behavior in a neighborhood of that point can surely help to determine the classification in special cases, but generally cannot be used: it is not clear how to determine the size of the neighborhood, and it is also difficult to assert that no critical points, undetected by the solver, reside in it; there is also the problem of checking whether f is well-defined in each of its points. This alternative might, however, be useful to consider when f and $\mathbf{G}$ are polynomials, as the issues mentioned above do not seem to pose a problem in that case.

Bibliography

[1] Arnold Emch. A Model for the Peano Surface. *The American Mathematical Monthly*, 29:388–391, 1922. `https://www.jstor.org/stable/2299024`.

[2] Salvador Gigena. Constrained local extrema without Lagrange multipliers and the higher derivative test. `https://arxiv.org/abs/1303.3184`, 2013.

[3] Lagrange multipliers. Encyclopedia of Mathematics. `http://www.encyclopediaofmath.org/index.php?title=Lagrange_multipliers`.

[4] Jeffrey Nunemacher. Lagrange Multipliers Can Fail to Determine Extrema. *The College Mathematics Journal*, 34:60–62, 2003. `https://www.maa.org/sites/default/files/nunemacher01010325718.pdf`.

[5] Bernard Parisse and Renée De Graeve. Giac/Xcas, version 1.4.9. `http://www-fourier.ujf-grenoble.fr/~parisse/giac.html`, 2018.

[6] W. F. Trench. *Introduction to Real Analysis*. Prentice Hall/Pearson Education, 2003. `ramanujan.math.trinity.edu/wtrench/texts/TRENCH_REAL_ANALYSIS.PDF`.