Calculus on $\mathbb{R}^n$
Homework 10 **2023**

1 Local Extremum Points for Twice Differentiable Functions

Algorithm for determining the local extremum points of a C^2 function

Let $\emptyset \neq A \subseteq \mathbb{R}^n$ an open set, and let $f : A \to \mathbb{R}$ be a C^2 function on A.

Step 1 Determine the **first order partial derivatives** of f

Step 2 Determine the critical/stationary points of f, i.e all the points $a \in A$ s.t.

$$\frac{\partial f}{\partial x_j}(a) = 0, \quad \forall j \in \{1, ..., n\}.$$

More precisely, determine the set

$$\mathcal{C} = \{c \in A : \nabla f(c) = 0_n\}$$

If f **has no critical points**, (i.e $\mathcal{C} = \emptyset$) then f **has no local extremum points.**
If f **has some critical points,** (i.e. $\mathcal{C} \neq \emptyset$) **GO TO Step 3**

Step 3 Determine all the **second-order partial derivatives** of f at a random point in $a \in A$. Construct the Hessian matrix attached to the function f at the point a, denoted by $H_f(a)$, and defined by:

$$H_f(a) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(a) & \frac{\partial^2 f}{\partial x_2 \partial x_1}(a) & \dots & \frac{\partial^2 f}{\partial x_n \partial x_1}(a) \\ \frac{\partial^2 f}{\partial x_1 \partial x_2}(a) & \frac{\partial^2 f}{\partial x_2^2}(a) & \dots & \frac{\partial^2 f}{\partial x_n \partial x_2}(a) \\ \dots & \dots & \dots & \dots \\ \frac{\partial^2 f}{\partial x_1 \partial x_n}(a) & \frac{\partial^2 f}{\partial x_2 \partial x_n}(a) & \dots & \frac{\partial^2 f}{\partial x_n^2}(a) \end{pmatrix}$$

Step 4 For **every** critical point $c \in \mathcal{C}$, analyze the Hessian matrix, hence, analyze $H_f(c)$

- If $H_f(c)$ is **POSITIVE DEFINITE**, then c is a **local minimum point** of f.
- If $H_f(c)$ is **NEGATIVE DEFINITE**, then c is a **local maximum point** of f.

- If $H_f(c)$ is **INDEFINITE**, then c is **not a local extremum point** of f, case in which it is called a **a saddle point**.

- Recall to repeat **Step 4** for each critical point $c \in \mathcal{C}$

Now, the natural question arising is concerning the nature of the Hessian matrix. We use **Sylvester's Theorem** applied to the quadratic matrix $H_f(c)$. In order to formulate this theorem, we make the following notations for the determinants for the minors (on the first diagonal) of this matrix. Thus, for

$$H_f(a) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(a) & \frac{\partial^2 f}{\partial x_2 \partial x_1}(a) & \dots & \frac{\partial^2 f}{\partial x_n \partial x_1}(a) \\ \\ \frac{\partial^2 f}{\partial x_1 \partial x_2}(a) & \frac{\partial^2 f}{\partial x_2^2}(a) & \dots & \frac{\partial^2 f}{\partial x_n \partial x_2}(a) \\ \\ \dots & \dots & \dots & \dots \\ \frac{\partial^2 f}{\partial x_1 \partial x_n}(a) & \frac{\partial^2 f}{\partial x_2 \partial x_n}(a) & \dots & \frac{\partial^2 f}{\partial x_n^2}(a) \end{pmatrix}$$

$$\Delta_1 = \left| \frac{\partial^2 f}{\partial x_1^2}(a) \right|, \quad \Delta_2 = \begin{vmatrix} \frac{\partial^2 f}{\partial x_1^2}(a) & \frac{\partial^2 f}{\partial x_2 \partial x_1}(a) \\ \frac{\partial^2 f}{\partial x_1 \partial x_2}(a) & \frac{\partial^2 f}{\partial x_2^2}(a) \end{vmatrix} \dots$$

$$\Delta_n = \begin{vmatrix} \frac{\partial^2 f}{\partial x_1^2}(a) & \frac{\partial^2 f}{\partial x_2 \partial x_1}(a) & \dots & \frac{\partial^2 f}{\partial x_n \partial x_1}(a) \\ \\ \frac{\partial^2 f}{\partial x_1 \partial x_2}(a) & \frac{\partial^2 f}{\partial x_2^2}(a) & \dots & \frac{\partial^2 f}{\partial x_n \partial x_2}(a) \\ \\ \dots & \dots & \dots & \dots \\ \frac{\partial^2 f}{\partial x_1 \partial x_n}(a) & \frac{\partial^2 f}{\partial x_2 \partial x_n}(a) & \dots & \frac{\partial^2 f}{\partial x_n^2}(a) \end{vmatrix}$$

Sylvester's Theorem:

A quadratic matrix of the form $H_f(a)$ is

- **POSITIVE DEFINITE if $\Delta_k > 0$, for all $k \in \{1, ..., n\}$**
- **NEGATIVE DEFINITE if $(-1)^K \Delta_k > 0$, for all $k \in \{1, ..., n\}$**
- **INDEFINITE if we are not in the two cases above, and still $\Delta_k > 0$ for all $k \in \{1, ..., n\}$**

This theorem can be used in most of the cases, but it cannot be applies when some of the determinants are 0.

For the case when there exists $k \in \{1, ..., n\}$ such that $\Delta_k = 0$, we have to consider the second order differential function attached to f at the point a, which is

$$d^2 f(a) : \mathbb{R}^n \to \mathbb{R}$$

given by

$$d^2 f(a)(h) = h^T \cdot H_f(a) \cdot h, \quad \forall h \in \mathbb{R}^n,$$

which, if you do the matrix multiplication turns into

$$d^2 f(a)(h) = \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(a) \cdot h_i \cdot h_j, \quad \forall h = (h_1, ..., h_n) \in \mathbb{R}^n .$$

If we may emphasize two particular points $u, v \in \mathbb{R}^n$ such that

$$d^2 f(a)(u) < 0 < d^2 f(a)(v)$$

this means that the $H_f(a)$ is indefinite, therefore, at **Step 4** of the algorithm, the conclusion is that

$$a \text{ is a saddle point,}$$

hence it is not a local extremum point.

$H_f(a)$ is said to be semidefinite if

$$d^2 f(a)(h) \leq 0$$

or ≥ 0 for all $h \in \mathbb{R}^n$ (it is not stict). In this case the nature of the point a has to be analyzed by going back to the original form of the function (see some of the following examples)

Moreover, consider the following briefly solved examples

Example 1:Determine all local extrema, their type (minima or maxima) and the corresponding extreme values of the following functions:

a) $f\colon \mathbb{R}^3 \to \mathbb{R}$, $f(x, y, z) = 2x^2 - xy + 2xz - y + y^3 + z^2$; b) $f\colon \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = x^2 + y^2(1 - x)^3$.

Solution

a) $f\colon \mathbb{R}^3 \to \mathbb{R}$, $f(x, y, z) = 2x^2 - xy + 2xz - y + y^3 + z^2$.

For all $(x, y, z) \in \mathbb{R}^3$ it holds $\frac{\partial f}{\partial x}(x, y, z) = 4x - y + 2z$, $\frac{\partial f}{\partial y}(x, y, z) = -x - 1 + 3y^2$ and $\frac{\partial f}{\partial z}(x, y, z) = 2x + 2z$. The stationary points of f are the solutions of the system

$$\begin{cases} 4x - y + 2z = 0 \\ -x - 1 + 3y^2 = 0 \\ 2x + 2z = 0. \end{cases}$$

From the last equation it follows that $x = -z$. By replacing this in the first equaiton we get $y = 2x$, equality which, together with the second equation, provides us with $12x^2 - x - 1 = 0$. Thus $x \in \{-\frac{1}{4}, \frac{1}{3}\}$. It follows that $(-\frac{1}{4}, -\frac{1}{2}, \frac{1}{4})$ and $(\frac{1}{3}, \frac{2}{3}, -\frac{1}{3})$ are the stationary points of the function.

For all $(x,y,z)\in\mathbb{R}^3$ it holds

$$\frac{\partial^2 f}{\partial x^2}(x,y,z)=4,\quad \frac{\partial^2 f}{\partial y^2}(x,y,z)=6y,\quad \frac{\partial^2 f}{\partial z^2}(x,y,z)=2,$$

$$\frac{\partial^2 f}{\partial x\partial y}(x,y,z)=-1,\quad \frac{\partial^2 f}{\partial x\partial z}(x,y,z)=2,\quad \frac{\partial^2 f}{\partial y\partial z}(x,y,z)=0,$$

thus

$$H(f)\left(-\frac{1}{4},-\frac{1}{2},\frac{1}{4}\right)=\begin{pmatrix}4&-1&2\\-1&-3&0\\2&0&2\end{pmatrix},\quad H(f)\left(\frac{1}{3},\frac{2}{3},-\frac{1}{3}\right)=\begin{pmatrix}4&-1&2\\-1&4&0\\2&0&2\end{pmatrix}.$$

The matrix $H(f)\left(\frac{1}{3},\frac{2}{3},-\frac{1}{3}\right)$ is positive definite, thus $(\frac{1}{3},\frac{2}{3},-\frac{1}{3})$ is a local minimum point. Moreover

$$f(\frac{1}{3},\frac{2}{3},-\frac{1}{3})=\frac{-13}{27}.$$

Since Sylvester's criterion cannot be applied to the matrix $H(f)\left(-\frac{1}{4},-\frac{1}{2},\frac{1}{4}\right)$, we will study its nature by other means. The quadratic from associated to this matrix is

$$\Phi(h_1,h_2,h_3)=4(h_1)^2-2h_1h_2+4h_1h_3-3(h_2)^2+2(h_3)^2.$$

Since $\Phi(0,1,0)=-3<0$ and $\Phi(0,0,1)=2>0$, it follos that this quadratic form (and its corresponding matrix) is indefinite. Therefore, $(-\frac{1}{4},-\frac{1}{2},\frac{1}{4})$ is not a local extremum, it is a saddle point.

b) $f\colon\ \mathbb{R}^2\to\mathbb{R}$, $f(x,y)=x^2+y^2(1-x)^3$.

For all $(x,y)\in\mathbb{R}^2$ it holds $\frac{\partial f}{\partial x}(x,y)=2x-3y^2(1-x)^2$ and $\frac{\partial f}{\partial y}(x,y)=2y(1-x)^3$. The stationary points of f are the solutions of the system

$$\begin{cases}2x-3y^2(1-x)^2=0\\ 2y(1-x)^3=0.\end{cases}$$

From the last equation we get $y=0$ or $x=1$. If $y=0$, then from the first equation of the system we get $x=0$. We notice that $x=1$ does not obey the first equation of the system, thus it is excluded. In conclusion $(0,0)$ is the only stationary point of the function.

For all $(x,y)\in\mathbb{R}^2$ it holds

$$\frac{\partial^2 f}{\partial x^2}(x,y)=2+6y^2(1-x),\quad \frac{\partial^2 f}{\partial x\partial y}(x,y)=-6y(1-x)^2,\quad \frac{\partial^2 f}{\partial y^2}(x,y)=2(1-x)^3,$$

thus

$$H(f)(0,0)=\begin{pmatrix}2&0\\0&2\end{pmatrix}.$$

Since $2>0$ and $4>0$, it follows that $H(f)(0,0)$ is positive definite, thus $(0,0)$ is a local minimum point. Moreover

$$f(0,0)=0.$$

Exerciţii:

Determine all local extrema, their type (minima or maxima) and the corresponding extreme values of the following functions:

a) $f\colon \mathbb{R}^3 \to \mathbb{R}, \quad f(x,y,z) = x^3 - 3x + y^2 + z^2;$

b) $f\colon \mathbb{R}^2 \to \mathbb{R} \quad f(x,y) = x^4 + y^4 - 4(x-y)^2$**;**

c) $f\colon \mathbb{R}^3 \to \mathbb{R} \quad f(x,y,z) = z^2(1+xy) + xy$**;**

d) $f\colon \mathbb{R}^2 \to \mathbb{R} \quad f(x,y) = x^3 + 3xy^2 - 15x - 12y.$