MATHEMATICAL ANALYSIS 2
Worksheet 2.
Second and higher order derivatives. Convexity. Sylvester's criterion.

Theory outline and sample problems

The higher order partial derivatives are defined iteratively; that is, for a given function $f(x, y)$ its second order partial derivatives are defined as the partial derivatives of the first order partial derivatives $f'_x(x, y), f'_y(x, y)$, considered as new functions. For a function of two variables $f(x, y)$, there exist two partial derivatives of the first order $f'_x(x, y), f'_y(x, y)$, and four ($2 \times 2 = 4$) partial derivatives of the second order: $f''_{xx}(x, y), f''_{xy}(x, y), f''_{yx}(x, y), f''_{yy}(x, y)$. Not all of them are different, though, due to the following

Theorem 1. *(The Schwartz theorem). Let for a function $f(x, y)$ the mixed derivative f''_{xy} and f''_{yx} be well defined at the point (x_0, y_0) and f''_{xy} be continuous at this point. Then $f''_{xy}(x_0, y_0) = f''_{yx}(x_0, y_0)$.*

In other words, if the mixed derivatives are well defined and one of them is continuous, then they coincide. The continuity condition is quite mild and is typically satisfied; that is, there are typically 3 different partial derivatives of the second order for a given function $f(x, y)$.

Control question: How many different partial derivatives of the second order has a function $f(x, y, z)$? *Answer:* 6. *Explain, why!*

The partial derivatives of the 3-rd, 4-th, ... orders are defined similarly, and for them the analogue of the Schwartz theorem is true; that is, the mixed derivatives taken with the various order of variables coincide, provided they are well defined and continuous. For instance, the third order derivatives $f'''_{xxy}(x, y)$ and $f'''_{xyx}(x, y)$ coincide.

Control question: How many different partial derivatives of the third order has a function $f(x, y)$? A function $f(x, y, z)$? *Answers:* 4,10. *Explain, why!*

Sample problem **1:** Calculate all partial derivatives of the 2nd order of the function
$f(x, y) = \cos(x^2 + y^2)$
Solution: Calculate first the 1st order partial derivatives:

$$f'_x(x, y) = -\sin(x^2 + y^2)(x^2 + y^2)'_x = -2x\sin(x^2 + y^2),$$

$$f'_y(x, y) = -\sin(x^2 + y^2)(x^2 + y^2)'_y = -2y\sin(x^2 + y^2),$$

where we have used the chain rule and the table of derivatives. Then similarly calculate, using in addition the product formula,

$$f''_{xx}(x, y) = (-2x\sin(x^2+y^2))'_x = -2\sin(x^2+y^2)-2x\cos(x^2+y^2)(2x) = -2\sin(x^2+y^2)-4x^2\cos(x^2+y^2),$$

$$f''_{xy}(x, y) = (-2x\sin(x^2 + y^2))'_y = -2x\cos(x^2 + y^2)(2y) = -4xy\cos(x^2 + y^2),$$

$$f''_{yy}(x, y) = (-2y\sin(x^2 + y^2))'_y = -2\sin(x^2 + y^2) - 4y^2\cos(x^2 + y^2).$$

The function $f'_y(x, y)$ is differentiable in y; that is, $f''_{yx}(x, y)$ is well defined; in addition $f''_{xy}(x, y)$ is continuous. Then by the Schwartz theorem

$$f''_{yx}(x, y) = f''_{xy}(x, y) = -4xy\cos(x^2 + y^2).$$

Definition 1. The Hessian of the function $f(x_1, \dots, x_n)$ at the point (x_0, y_0) is the matrix

$$H_f(x_0, y_0) = \left(f''_{x_i x_j}(x_0, y_0)\right)_{i,j=1}^n.$$

For example, for the function $f(x,y)$ in the sample problem 1

$$H_f(x,y) = \begin{pmatrix} -2\sin(x^2+y^2) - 4x^2\cos(x^2+y^2) & -4xy\cos(x^2+y^2) \\ -4xy\cos(x^2+y^2) & -2\sin(x^2+y^2) - 4y^2\cos(x^2+y^2) \end{pmatrix}$$

Note that the Hessian is a symmetric matrix, provided that its entries are continuous; this is just the Schwartz theorem.

The reason why it is natural to place the partial derivatives of second order into a matrix becomes clear when one looks at various one-dimensional *sections* (or *traces*) of the function. Recall that, for the function of two[1] variables $f(x,y)$ its section in the direction $\mathbf{v} = (a,b)$ at the point (x_0,y_0) is the function $F_{(x_0,y_0),\mathbf{v}}(t) = f(x_0+at, y_0+bt), t \in \mathbb{R}$. The derivative of this function at the point $t=0$ equals to the directional derivative of f in the direction $\mathbf{v}$ at the point (x_0,y_0), and is equal

$$\frac{d}{dt}F_{(x_0,y_0),\mathbf{v}}(t)\bigg|_{t=0} = \nabla f(x_0,y_0)\cdot\mathbf{v} = f'_x(x_0,y_0)a + f'_y(x_0,y_0)b,$$

where $\nabla f = (f'_x, f'_y)$ is the *gradient*, or the *vector derivative* of the function. For the second derivative, the similar formula is available, which involves the Hessian:

$$\frac{d^2}{dt^2}F_{(x_0,y_0),\mathbf{v}}(t)\bigg|_{t=0} = \mathbf{v}H_f(x_0,y_0)\mathbf{v}^\top = f''_{xx}(x_0,y_0)a^2 + 2f''_{xy}(x_0,y_0)ab + f''_{yy}(x_0,y_0)b^2.$$

The second derivative of a function $F(t)$ is a convenient tool for description of convexity/concavity of this function. Recall that $F(t)$ is said to be convex on an interval (t_0,t_1) if for any $s,t \in (t_0,t_1)$ and any $\alpha \in (0,1)$

$$F(\alpha s + (1-\alpha)t) \leqslant \alpha F(s) + (1-\alpha)F(t);$$

for the function to be concave the inequality should be '$\geqslant$'. Geometrically, the function is convex/concave if the the horde connecting two points on the graph (points $(t_0,F(t_0))$ and $(t_1,F(t_1))$ is located over/under the graph. The point t is called a convexity/concavity point for the function $F(t)$ if this function is convex/concave at some neighbourhood $(t-\epsilon, t+\epsilon)$ of the point t. Sufficient condition for convexity (concavity) is that $F''(t) > 0$ (resp. $F''(t) < 0$).

For a function $f(x,y)$ we say that a point (x_0,y_0) is a convexity/concavity point in a direction $\mathbf{v}$ if, for the section of f in this direction, the point $t=0$ is a convexity/concavity point.

Sample problem **2:** Find the points of convexity and concavity for the function $f(x,y) = x^3+y^3-3xy$ in the direction $\mathbf{v} = (-1,1)$.
Solution: Calculate the gradient and the Hessian of the function:

$$\nabla f(x,y) = (3x^2-3y, 3y^2-3x), \quad H_f(x,y) = \begin{pmatrix} 6x & -3 \\ -3 & 6y \end{pmatrix}.$$

The second derivative of the section function in the direction $\mathbf{v}$:

$$\mathbf{v}H_f(x_0,y_0)\mathbf{v}^\top = 6x_0(-1)^2 + 2(-1)1(-3) + 6y_0 1^2 = 6x_0 + 6y_0 + 6.$$

Then

$$\begin{cases} (x_0,y_0) \text{ is a concavity point}, & x_0+y_0 < -1 \\ (x_0,y_0) \text{ is a convexity point}, & x_0+y_0 > -1 \end{cases}$$

The boundary case $x_0+y_0 = -1$ is not covered by the sufficient condition, and should be studied separately. Note that $\mathbf{v}$ is the direction vector for the line $\ell = \{x+y=1\}$, thus if $(x_0,y_0) \in \ell$ then for any $t \in \mathbb{R}$ the point $(x_0,y_0) + t\mathbf{v} \in \ell$. Therefore for such (x_0,y_0) one has $\frac{d^2}{dt^2}F_{(x_0,y_0),\mathbf{v}}(t) \equiv 0$, which means that the section $F_{(x_0,y_0),\mathbf{v}}(t)$ is a linear function of t. The linear function is both convex and concave, hence the final answer is

$$\begin{cases} (x_0,y_0) \text{ is a concavity point}, & x_0+y_0 < -1 \\ (x_0,y_0) \text{ is a convexity point}, & x_0+y_0 > -1 \\ (x_0,y_0) \text{ is both a concavity and a convexity point}, & x_0+y_0 = -1 \end{cases}$$

[1] For simplicity of notation, only; the same relations hold true for arbitrary number of variables

Definition 2. A set $D \subset \mathbb{R}^2$ is convex if, for any two points $X, Y \in D$, the segment $[X, Y]$ connecting these points belongs to D. Function $f(x, y), (x, y) \in D$ is called convex if D is convex and for any $(x_1, y_1), (x_2, y_2) \in D$ and any $\alpha \in (0, 1)$

$$f(\alpha x_1 + (1-\alpha)x_2, \alpha y_1 + (1-\alpha)y_2) \leqslant \alpha f(x_1, y_1) + (1-\alpha) f(x_2, y_2);$$

for the function to be concave the inequality should be '$\geqslant$'.

Geometric meaning of this definition is the same as for the above one-dimensional one: for a function to be convex/concave, any horde connecting two points on the graph has to be located over/under the graph; note however that now the graph is a surface, not a curve. The requirement that the domain D is convex, from this point of view, is just the requirement that for each point on the horde there should be a point on the graph to compare with.

The point (x_0, y_0) is called a convexity/concavity point for the function $f(x, y)$ if this function is convex/concave when restricted to some (small) ball $B_\epsilon(x_0, y_0)$ centered at this point, $B_\epsilon(x_0, y_0) = \{(x, y) : \sqrt{(x-x_0)^2 + (y-y_0)^2} < \epsilon\}$. For (x_0, y_0) to be a convexity/concavity point it is equivalent that (x_0, y_0) is a convexity/concavity point at each direction $\mathbf{v}$. This leads to the useful sufficient condition for convexity in the terms of the Hessian.

Definition 3. A symmetric $n \times n$-matrix A is called

- positive definite, if $\mathbf{v} A \mathbf{v}^\top > 0$ for any $\mathbf{v} \neq 0$;
- negative definite, if $\mathbf{v} A \mathbf{v}^\top < 0$ for any $\mathbf{v} \neq 0$.

Theorem 2. *For (x_0, y_0) to be a convexity/concavity point for a function $f(x, y)$ it is sufficient that the Hessian of the function in this point is positive/negative definite.*

There exists a convenient criterion for a matrix to be positive/negaitive definite.

Theorem 3. *(The Sylvester criterion) For a symmetric matrix A to be positive defined it is necessary and sufficient that for each sub-matrix A_k of the size $k \times k, k = 1, \ldots, n$, which has the same upper left corner with the original matrix, the determinant is positive.*
For the matrix to be negative defined, the determinants of A_k should have the signs $-1, +1, -1, \ldots$ for $k = 1, 2, 3, \ldots$.

Sample problem **3:** Study if the following matrix is positive/negative definite: $\begin{pmatrix} 3 & 3 \\ 3 & 4 \end{pmatrix}$.
Solution: We have

$$\det A_1 = \det(3) = 3 > 0, \quad \det A_2 = \det \begin{pmatrix} 3 & 3 \\ 3 & 4 \end{pmatrix} = 3 > 0,$$

hence the matrix is positive definite.

Sample problem **4:** Write the Hessian of the function and specify the domains where the Hessian is positively/negatively defined, $f(x, y) = \cos(x^2 + y^2)$.
Solution: We have already calculated the Hessian,

$$H_f(x, y) = \begin{pmatrix} -2\sin(x^2+y^2) - 4x^2\cos(x^2+y^2) & -4xy\cos(x^2+y^2) \\ -4xy\cos(x^2+y^2) & -2\sin(x^2+y^2) - 4y^2\cos(x^2+y^2) \end{pmatrix}.$$

For the Hessian to be either positive or negative defined it is necessary that the second minor is positive, i.e.

$$\begin{aligned} \det H_f(x, y) &= (2\sin(x^2+y^2) + 4x^2\cos(x^2+y^2))(2\sin(x^2+y^2) + 4y^2\cos(x^2+y^2)) - 16x^2y^2\cos(x^2+y^2)^2 \\ &= 4\sin(x^2+y^2)^2 + 8(x^2+y^2)\sin(x^2+y^2)\cos(x^2+y^2) > 0 \end{aligned}$$

If this condition is satisfied, then the Hessian is positive definite if $-2\sin(x^2+y^2)-4x^2\cos(x^2+y^2)>0$, and negative definite if the sign is $<$. Thus, the answer is that $H_f(x,y)$ is positive defined if

$$\begin{cases} 4\sin(x^2+y^2)^2+8(x^2+y^2)\sin(x^2+y^2)\cos(x^2+y^2)>0, \\ -2\sin(x^2+y^2)-4x^2\cos(x^2+y^2)>0, \end{cases}$$

and is negative defined if

$$\begin{cases} 4\sin(x^2+y^2)^2+8(x^2+y^2)\sin(x^2+y^2)\cos(x^2+y^2)>0, \\ -2\sin(x^2+y^2)-4x^2\cos(x^2+y^2)<0. \end{cases}$$

Concavity/convexity naturally applies for the study of local extrema of the functions of several variables.

Definition 4. A point (x_0,y_0) is a local maximum of a function $f(x,y)$ if there exists a (small) ball $B_\epsilon(x_0,y_0)$ centered at this point, such that

$$f(x,y)\leqslant f(x_0,y_0), \quad (x,y)\in B_\epsilon(x_0,y_0).$$

A point (x_0,y_0) is a local minimum, if

$$f(x,y)\geqslant f(x_0,y_0), \quad (x,y)\in B_\epsilon(x_0,y_0).$$

Local extremum point is a point of either local maximum or local minimum.

Theorem 4. *I* (Necessary condition of a local extremum) *If (x_0,y_0) is the interior point of the domain of a differentiable function $f(x,y)$ and (x_0,y_0) is a local extremum, then*

$$\nabla f(x_0,y_0)=\vec{0}. \tag{*}$$

Any point (x_0,y_0) satisfying (*) is called a critical point of the function $f(x,y)$.

II (Sufficient condition of a local extremum) *If (x_0,y_0) is a critical point and a point of convexity/concavity for $f(x,y)$, then it is a local minimum/maximum.*

The following classification of the critical points on a plane is standard: if $\det H_f(x_0,y_0)\neq 0$, the eigenvalues λ_1,λ_2 of this matrix is non-zero, and the following three possibilities are available, only:

- $\lambda_1,\lambda_2>0$, $H_f(x_0,y_0)$ is positive definite, (x_0,y_0) is a convexity point and a local minimum;
- $\lambda_1,\lambda_2<0$, $H_f(x_0,y_0)$ is negative definite, (x_0,y_0) is a concavity point and a local maximum;
- λ_1,λ_2 have different signs, f is convex/concave in different directions (namely, eigenvectors for $H_f(x_0,y_0)$), in this case (x_0,y_0) is called a *saddle point*

Sample problem **5:** Find and classify all the critical points of $f(x,y)=4+x^3+y^3-3xy$.
Solution: To find the critical points, we have to solve (*). Calculate the gradient:

$$\nabla f(x,y)=(3x^2-3y,3y^2-3x).$$

Then (*) is equivalent to the system of equations

$$\begin{cases} 3x^2-3y=0 \\ 3y^2-3x=0 \end{cases} \Longleftrightarrow \begin{cases} y=x^2 \\ x=y^2 \end{cases} \Longleftrightarrow \begin{cases} y=x^2 \\ x=x^4 \end{cases} \Longleftrightarrow \left[\begin{array}{l} (x,y)=(0,0) \\ (x,y)=(1,1) \end{array} \right.$$

We have

$$H_f(x,y)=\begin{pmatrix} 6x & -3 \\ -3 & 6y \end{pmatrix},$$

and $H_f(1,1)$ is positive definite, while $H_f(0,0)$ is neither positive nor negative definite and $\det H_f(0,0) = -9 \neq 0$. Thus $(1,1)$ is a local minimum, $(0,0)$ is a saddle point, and there is no other critical points.

Problems to solve

1. Calculate all partial derivatives of the 2nd order of the functions

(a) $f(x,y) = \sin(x^3 - y^2)$, (b) $f(x,y) = ye^{x^2+y^2}$, (c) $f(x,y) = \operatorname{tg} x + \dfrac{y^3}{x}$,

(d) $f(x,y) = \ln 1 + xy$, (e) $f(x,y,z) = \dfrac{x}{\sqrt{x^2+z^2+z^2}}$ (f) $f(x,y,z) = \ln(1 + x^2 + y^3 + z^4)$.

2. Write the Hessian of the function and specify the domains where the Hessian is positive/negative definite.

(a) $f(x,y) = \sin(x^2 + y^2)$, (b) $f(x,y) = xye^{x+y}$, (c) $f(x,y) = x + \dfrac{y}{x^2}$,

(d) $f(x,y) = \ln 1 + xy$, (e) $f(x,y) = \dfrac{1}{\sqrt{x^2+y^2}}$ (f) $f(x,y) = \ln(1 + x^2 + y^2)$.

3. Study the following matrices for being positive/negative definite.

(a) $\begin{pmatrix} -3 & 3 \\ 3 & -4 \end{pmatrix}$, (b) $\begin{pmatrix} -1 & 2 \\ 2 & 4 \end{pmatrix}$, (c) $\begin{pmatrix} -1 & 2 & 4 \\ 2 & 2 & 2 \\ 4 & 2 & 1 \end{pmatrix}$,

(d) $\begin{pmatrix} 3 & 4 & 1 \\ 4 & 5 & 2 \\ 1 & 2 & 17 \end{pmatrix}$, (e) $\begin{pmatrix} 5 & 2 & 3 \\ 2 & 1 & 2 \\ 3 & 2 & 6 \end{pmatrix}$, (f) $\begin{pmatrix} 2 & 2 & -3 & -1 \\ 2 & 5 & 1 & 2 \\ -3 & 1 & -2 & 3 \\ -1 & 2 & 3 & 4 \end{pmatrix}$.

4. Find the points of convexity/concavity for the function $f(x,y) = x^4 + y^4 - 4xy$ in the direction $\mathbf{v} = (1,1)$.

5. Find and classify all the critical points of $f(x,y) = 2x^2 + y^2 - 3xy$.

6. Find and classify all the critical points of $f(x,y) = 2x^3 - 3x^2y - 12x^2 - 3y^2$.

7. Find and classify all the critical points of $f(x,y) = 2xy - y^2 + x^3 + x^2$.

8. Calculate all partial derivatives of the 3rd order of the functions

(a) $f(x,y) = \cos(x^2 + y^2)$ (b) $f(x,y) = e^{xy}$ (c) $f(x,y) = \sqrt{x+y}$.