Lecture 9

Vector-valued functions of several variables

Let $n, m \in \mathbb{N}$, $m \geq 2$. For $j \in \{1, \ldots, m\}$, consider the projection mapping $pr_j : \mathbb{R}^m \to \mathbb{R}$, $pr_j(y) = y_j$, $\forall y = (y_1, \ldots, y_m) \in \mathbb{R}^m$.

Definition 1. Let $A \subseteq \mathbb{R}^n$ nonempty. A function $f : A \to \mathbb{R}^m$ is called a *vector-valued function of n variables*. The *components of f* are the real-valued functions $f_1, \ldots, f_m : A \to \mathbb{R}$ defined by $f_j = pr_j \circ f$, $\forall j \in \{1, \ldots, m\}$.
Notation: $f = (f_1, \ldots, f_m)$.

Example 1. Let $A \subseteq \mathbb{R}^n$ nonempty and open, and $f : A \to \mathbb{R}$ partially differentiable.
Gradient of f: $\nabla f : A \to \mathbb{R}^n$, $\nabla f(c) = \left(\frac{\partial f}{\partial x_1}(c), \ldots, \frac{\partial f}{\partial x_n}(c)\right)$, $\forall c \in A$.

In general, properties of vector-valued functions can be studied by considering their components one at a time.

The chain rule

Theorem 1 (Chain rule)**.** *Let $I \subseteq \mathbb{R}$ be an interval, $c \in I$, $A \subseteq \mathbb{R}^n$, and $g = (g_1, \ldots, g_n) : I \to A$ such that $\forall i \in \{1, \ldots, n\}$, g_i is differentiable at c and $g(c) \in \operatorname{int} A$. If $f = f(x_1, \ldots, x_n) : A \to \mathbb{R}$ is C^1 near $g(c)$ (that is, there exists $r > 0$ such that $B(g(c), r) \subseteq A$ and $f|_{B(g(c),r)} \in C^1(B(g(c), r))$), then $f \circ g : I \to \mathbb{R}$ is differentiable at c and*

$$(f \circ g)'(c) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(g(c)) g_i'(c).$$

Example 2. $g : \mathbb{R} \to \mathbb{R}^2$, $g(t) = (t^2, 3t)$, $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = x^2 y - x$.

Remark 1. There also exists a chain rule for the more general case when g is a vector-valued function of several variables.

Local extrema and partial derivatives

Definition 2. Let $A \subseteq \mathbb{R}^n$ nonempty and $f : A \to \mathbb{R}$. We say that $c \in A$ is a

- *local maximum point (local minimum point)* for f if there exists $V \in \mathcal{V}(c)$ such that for every $x \in V \cap A$,
$$f(c) \geq f(x) \quad (f(c) \leq f(x)). \tag{1}$$
- *local extremum point* for f if it is either a local maximum point or a local minimum point for f.
- *global maximum point (global minimum point)* for f if (1) holds for every $x \in A$.
- *global extremum point* for f if it is either a global maximum point or a global minimum point for f.

If c is a local maximum (local minimum / local extremum) point for f, then we also say that f *attains a local maximum (local minimum / local extremum) at* c.
If c is a global maximum (global minimum / global extremum) point for f, then we also say that f *attains a global maximum (global minimum / global extremum) at* c.
We say that f *attains its maximum (minimum)* if it has at least one global maximum point (global minimum point).

Definition 3. A set $A \subseteq \mathbb{R}^n$ is called

- *bounded* if there exists $r > 0$ such that $A \subseteq B(0_n, r)$.
- *closed* if $\mathbb{R}^n \setminus A$ is open.

Theorem 2 (Maximum-Minimum Theorem, Weierstrass)**.** *Let $A \subseteq \mathbb{R}^n$ be nonempty, closed and bounded. If $f : A \to \mathbb{R}$ is continuous, then f attains both its maximum and its minimum.*

Theorem 3 (Fermat)**.** *Let $A \subseteq \mathbb{R}^n$ be nonempty and open, and $f : A \to \mathbb{R}$. If $c \in A$, f is partially differentiable at c, and f attains a local extremum at c, then $\nabla f(c) = 0_n$.*

Definition 4. Let $A \subseteq \mathbb{R}^n$ be nonempty and open, and $f : A \to \mathbb{R}$. A point $c \in A$ at which f is partially differentiable is called a *stationary point* (or *critical point*) for f if $\nabla f(c) = 0_n$.

Remark 2. Local extremum points of a function which is defined on an open set and which is partially differentiable are found among its stationary points, but not all stationary points are local extremum points. Stationary points which are not local extremum points are sometimes called *saddle points*.

Example 3. $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = x^2 - y^2$, $c = 0_2$.

Definition 5. Let $C = (c_{ij})_{1 \leq i,j \leq n}$ be a symmetric $n \times n$ matrix of real numbers. The function $\Phi_C : \mathbb{R}^n \to \mathbb{R}$ defined by

$$\Phi_C(h) = \sum_{i=1}^{n} \sum_{j=1}^{n} c_{ij} h_i h_j, \quad \forall h = (h_1, \ldots, h_n) \in \mathbb{R}^n$$

is called the *quadratic form associated to C*.
We say that Φ_C (or, equivalently, C) is

- *positive definite (negative definite)*: if for every $h \in \mathbb{R}^n \setminus \{0_n\}$, $\Phi_C(h) > 0$ ($\Phi_C(h) < 0$).
- *positive semidefinite (negative semidefinite)*: if for every $h \in \mathbb{R}^n$, $\Phi_C(h) \geq 0$ ($\Phi_C(h) \leq 0$).
- *indefinite*: if there exist $a, b \in \mathbb{R}^n$ such that $\Phi_C(a) < 0 < \Phi_C(b)$.

Example 4. (i) $n = 2$,

$$C = \begin{pmatrix} c_{11} & c_{12} \\ c_{12} & c_{22} \end{pmatrix}.$$

(ii) $n = 3$,

$$C = \begin{pmatrix} c_{11} & c_{12} & c_{13} \\ c_{12} & c_{22} & c_{23} \\ c_{13} & c_{23} & c_{33} \end{pmatrix}.$$

Theorem 4 (Sylvester)**.** *Let $C = (c_{ij})_{1 \le i,j \le n}$ be a symmetric $n \times n$ matrix of real numbers. For every $k \in \{1, \dots, n\}$, let*

$$\Delta_k = \det(c_{ij})_{1 \le i,j \le k} = \begin{vmatrix} c_{11} & \dots & c_{1k} \\ \vdots & & \vdots \\ c_{k1} & \dots & c_{kk} \end{vmatrix}.$$

Then

(i) C is positive definite $\iff \Delta_k > 0$, $\forall k \in \{1, \dots, n\}$.

(ii) C is negative definite $\iff (-1)^k \Delta_k > 0$, $\forall k \in \{1, \dots, n\}$.

Example 5. Let

$$C = \begin{pmatrix} -2 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 2 \end{pmatrix}.$$

Theorem 5. *Let $A \subseteq \mathbb{R}^n$ be nonempty and open, $c \in A$ and $f \in C^2(A)$. Then*

(i) if c is a local minimum (local maximum) point of f, then $\nabla f(c) = 0_n$ and $H_f(c)$ positive semidefinite (negative semidefinite).

(ii) if $\nabla f(c) = 0_n$ and $H_f(c)$ is positive definite (negative definite), then c is a local minimum (local maximum) point of f.

Remark 3. If $\nabla f(c) = 0_n$ and $H_f(c)$ is indefinite, then c is not a local extremum point of f.