14 Lecture 14 – Local Extrema of Function

14.1 Taylor's Formula with Lagrangian Remainder Term

Theorem 14.1. *Let $n \in \mathbb{N} \cup \{0\}$ and $f : (a,b) \to \mathbb{R}$. We assume that there exists $f^{(n+1)}(x)$ for all $x \in (a,b)$. Then for each $x, x_0 \in (a,b)$ there exists a point ξ between x and x_0 such that*

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k + \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-x_0)^{n+1}. \tag{16}$$

The term $\frac{f^{(n+1)}(\xi)}{(n+1)!}(x-x_0)^{n+1}$ is called the **Lagrangian remainder term**.

Proof. If $x = x_0$, then formula (16) holds. We assume that $x_0 < x$ and consider a new function

$$g(z) := f(x) - \sum_{k=0}^{n} \frac{f^{(k)}(z)}{k!}(x-z)^k - \frac{L}{(n+1)!}(x-z)^{n+1}, \quad z \in [x_0, x],$$

where the number L is chosen such that $g(x_0) = 0$. We note that the function g is continuous on $[x_0, x]$ and has a derivative

$$g'(z) = -\frac{f^{(n+1)}(z)}{n!}(x-z)^n + \frac{L}{n!}(x-z)^n$$

Moreover, $g(x) = 0$. By Rolle's theorem (see Theorem 11.3), there exists $\xi \in (x_0, x)$ such that $g'(\xi) = 0$, that is,

$$g'(\xi) = -\frac{f^{(n+1)}(\xi)}{n!}(x-\xi)^n + \frac{L}{n!}(x-\xi)^n = 0.$$

Consequently, we have $L = f^{(n+1)}(\xi)$.

The case $x < x_0$ is similar. □

Remark 14.1. Formula (16) is a generalisation of the Lagrange theorem, which can be obtained taking $n = 0$.

Example 14.1. Let $f(x) = e^x$, $x \in \mathbb{R}$, and $x_0 = 0$. Then for all $n \in \mathbb{N}$ and $x \in \mathbb{R}$ there exists ξ between 0 and x such that

$$e^x = 1 + x + \frac{x^2}{2!} + \ldots + \frac{x^n}{n!} + \frac{e^\xi}{(n+1)!}x^{n+1}. \tag{17}$$

This formula follows from Theorem 14.1 and Example 13.2, since $f^{(k)}(0) = e^0 = 1$.

Remark 14.2. Formula (17) allows to obtain an approximate value of e^x, computing the value of the polynomial $1 + x + \frac{x^2}{2!} + \ldots + \frac{x^n}{n!}$. Moreover, the error is equal $\frac{e^\xi}{(n+1)!}x^{n+1}$. For instance, for $x \in [0,3]$ and $n = 12$ we have

$$\left|\frac{e^\xi}{(n+1)!}x^{n+1}\right| < \frac{e^3 3^{13}}{13!} < \frac{1}{1000}.$$

14.2 Local Extrema of Function

Let $f : (a, b) \to \mathbb{R}$ be a given function.

Definition 14.1.
- A point x_0 is called a **point of local maximum** (**local minimum**) **of** f, if there exists $\delta > 0$ such that $B(x_0, \delta) = (x_0 - \delta, x_0 + \delta) \subset (a, b)$ and $f(x) \le f(x_0)$ (resp. $f(x) \ge f(x_0)$) for all $x \in B(x_0, \delta)$.

  If x_0 is a point of local minimum or local maximum of f, then it is called a **point of local extrema of** f.

- A point x_0 is called a **point of strict local maximum** (**strict local minimum**) **of** f, if there exists $\delta > 0$ such that $B(x_0, \delta) \subset (a, b)$ and $f(x) < f(x_0)$ (resp. $f(x) > f(x_0)$) for all $x \in B(x_0, \delta) \setminus \{x_0\}$.

  If x_0 is a point of strict local minimum or strict local maximum of f, then it is called a **point of strict local extrema of** f.

Example 14.2. For the function $f(x) = x^2$, $x \in \mathbb{R}$, the point $x_0 = 0$ is a point of strict local minimum of f and f takes the smallest value at this point.

Example 14.3. For the function $f(x) = x$, $x \in [0, 1]$, the points $x_* = 0$ and $x^* = 1$ are points at which the function takes the smallest and the largest values, respectively. But they are not points of local extrema.

Theorem 14.2. *If x_0 is a point of local extrema of f and f has a derivative at x_0, then $f'(x_0) = 0$.*

Proof. Let x_0 be a point of local maximum. Then by Definition 14.1, there exists $\delta > 0$ such that $B(x_0, \delta) \subset (a, b)$ and $f(x) \le f(x_0)$ for all $x \in B(x_0, \delta)$. In particular, $f(x_0) = \max_{x \in B(x_0,\delta)} f(x)$. Applying the Fermat theorem (see Theorem 11.2) to the function f defined on $(x_0 - \delta, x_0 + \delta)$, we obtain $f'(x_0) = 0$. □

Remark 14.3. Theorem 14.2 gives only a necessary condition of local extrema. If $f'(x_0) = 0$ at some point $x_0 \in (a, b)$, then it does not imply that x_0 is a point of local extrema. For instance, for the function $f(x) = x^3$, $x \in \mathbb{R}$, the point $x_0 = 0$ is not a point of a local extrema while $f'(0) = 0$.

Remark 14.4. A point at which derivative does not exist can also be a point of local extrema. For example, for the function $f(x) = |x|$, $x \in \mathbb{R}$, the point $x_0 = 0$ is a point of local minimum but the derivative at $x_0 = 0$ does not exist (see Example 10.2).

Definition 14.2. A point $x_0 \in (a, b)$ is said to be a **critical point** or **stationary point of** f, if $f'(x_0) = 0$.

Remark 14.5. Point of local extrema of f belong to the set of all critical points of f and points where the derivative of f does not exist.

Theorem 14.3. *Let x_0 be a critical point of f and the function f be differentiable on some neighbourhood of the point x_0.*

a) If for some $\delta > 0$ $f'(x) > 0$ for all $x \in (x_0 - \delta, x_0)$ and $f'(x) < 0$ for all $x \in (x_0, x_0 + \delta)$, then x_0 is a point of strict local maximum of f.

b) If for some $\delta > 0$ $f'(x) < 0$ for all $x \in (x_0 - \delta, x_0)$ and $f'(x) > 0$ for all $x \in (x_0, x_0 + \delta)$, then x_0 is a point of strict local minimum of f.

Proof. We will only prove a). Since $f'(x) > 0$ for all $x \in (x_0 - \delta, x_0)$, the function f strictly increases on $(x_0 - \delta, x_0]$, by Remark 12.1. Hence, $f(x) < f(x_0)$ for all $x \in (x_0 - \delta, x_0)$. Similarly, $f(x_0) > f(x)$ for all $x \in (x_0, x_0 + \delta)$, since the function f strictly decreases on $[x_0, x_0 + \delta)$ due to $f'(x) < 0$, $x \in (x_0, x_0 + \delta)$. Thus, x_0 is a point of strict local maximum. □

Example 14.4. For the function $f(x) = x^3 - 3x$, $x \in \mathbb{R}$, the points 1 and -1 are critical points of f, since the derivative $f(x) = 3x^2 - 3$, $x \in \mathbb{R}$, equals zero at those points. The point -1 is a point of strict local maximum because the derivative changes its sign from "+" to "−", passing through -1. The point 1 is a point of strict local minimum because the derivative changes its sign from "−" to "+", passing through 1.

Exercise 14.1. Find points of local extrema of the following functions:
a) $f(x) = x^2 e^x$, $x \in \mathbb{R}$; b) $f(x) = x + \frac{1}{x}$, $x > 0$; c) $f(x) = x^x$, $x > 0$; d) $f(x) = |x| e^{-x^2}$, $x \in \mathbb{R}$.

Theorem 14.4. *Let a function $f : (a, b) \to \mathbb{R}$ and a point $x_0 \in (a, b)$ satisfy the following properties:*

1) there exists $\delta > 0$ such that f is differentiable on $(x_0 - \delta, x_0 + \delta)$;

2) $f'(x_0) = 0$;

3) there exists $f''(x_0)$ and $f''(x_0) \neq 0$.

If $f''(x_0) < 0$, then x_0 is a point of strict local maximum. If $f''(x_0) > 0$, then x_0 is a point of strict local minimum.

Proof. We write for the function f and the point x_0 the Taylor formula (see Theorem13.4). So,

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2!}(x - x_0)^2 + o((x - x_0)^2), \quad x \to x_0.$$

Hence, for $x \neq x_0$ we have

$$f(x) - f(x_0) = (x - x_0)^2 \left(\frac{f''(x_0)}{2!} + \frac{o((x - x_0)^2)}{(x - x_0)^2} \right)$$

and, hence, $f(x) - f(x_0)$ has the same sign as $f''(x_0)$ on some neighbourhood of x_0, since $\frac{o((x-x_0)^2)}{(x-x_0)^2} \to 0$, $x \to x_0$. □

Example 14.5. For the function $f(x) = x^2 - x$, $x \in \mathbb{R}$, the point $\frac{1}{2}$ is a point of strict local minimum, since $f'\left(\frac{1}{2}\right) = 0$ and $f''\left(\frac{1}{2}\right) = 2 < 0$.

Theorem 14.5. *Let $f : (a, b) \to \mathbb{R}$, a point x_0 belong to (a, b) and $m \in \mathbb{N}$, $m \geq 2$. We also assume that the following conditions hold:*

1) there exists $\delta > 0$ such that $f^{(m-1)}(x)$ exists for all $x \in (x_0 - \delta, x_0 + \delta)$;

2) $f'(x_0) = f''(x_0) = \ldots = f^{(m-1)}(x_0) = 0$;

3) there exists $f^{(m)}(x_0)$ and $f^{(m)}(x_0) \neq 0$.

If m is even and $f^{(m)}(x_0) < 0$, then x_0 is a point of local maximum.
If m is even and $f^{(m)}(x_0) > 0$, then x_0 is a point of local minimum.
If m is odd, then x_0 is not a point of local extrema.

Proof. The proof of Theorem 14.5 is similar to the proof of Theorem 14.4. □

Exercise 14.2. Prove Theorem 14.5.

Exercise 14.3. Find points of local extrema of the following functions:
a) $f(x) = x^4(1-x)^3$, $x \in \mathbb{R}$; b) $f(x) = \frac{x^2}{2} - \frac{1}{4} + \frac{9}{4(2x^2+1)}$, $x \in \mathbb{R}$; c) $f(x) = \begin{cases} e^{-\frac{1}{x^2}}, & x \neq 0, \\ 0, & x = 0, \end{cases}$, $x \in \mathbb{R}$.

14.3 Convex and Concave Functions

Let $-\infty \le a < b \le +\infty$.

Definition 14.3.
- A function $f : (a,b) \to \mathbb{R}$ is said to be a **convex function on** (a,b), if for each $x_1, x_2 \in (a,b)$ and $\alpha \in (0,1)$
$$f(\alpha x_1 + (1-\alpha)x_2) \le \alpha f(x_1) + (1-\alpha) f(x_2).$$
- A function $f : (a,b) \to \mathbb{R}$ is said to be a **concave function on** (a,b), if for each $x_1, x_2 \in (a,b)$ and $\alpha \in (0,1)$
$$f(\alpha x_1 + (1-\alpha)x_2) \ge \alpha f(x_1) + (1-\alpha) f(x_2).$$

Definition 14.4.
- A function $f : (a,b) \to \mathbb{R}$ is said to be a **strictly convex function on** (a,b), if for each $x_1, x_2 \in (a,b)$, $x_1 \neq x_2$, and $\alpha \in (0,1)$
$$f(\alpha x_1 + (1-\alpha)x_2) < \alpha f(x_1) + (1-\alpha) f(x_2).$$
- A function $f : (a,b) \to \mathbb{R}$ is said to be a **strictly concave function on** (a,b), if for each $x_1, x_2 \in (a,b)$, $x_1 \neq x_2$, and $\alpha \in (0,1)$
$$f(\alpha x_1 + (1-\alpha)x_2) > \alpha f(x_1) + (1-\alpha) f(x_2).$$

Example 14.6. Let $M, L \in \mathbb{R}$. The function $f(x) = Mx + L$, $x \in \mathbb{R}$, is both convex and concave on $\mathbb{R}$. Indeed, for each $x_1, x_2 \in \mathbb{R}$ and $\alpha \in (0,1)$ we have
$$f(\alpha x_1 + (1-\alpha)x_2) = M(\alpha x_1 + (1-\alpha)x_2) + L = \alpha(Mx_1 + L) + (1-\alpha)(Mx_2 + L) = \alpha f(x_1) + (1-\alpha) f(x_2).$$

Example 14.7. The function $f(x) = |x|$, $x \in \mathbb{R}$, is convex on $\mathbb{R}$. Indeed, for each $x_1, x_2 \in (a,b)$ and $\alpha \in (0,1)$
$$f(\alpha x_1 + (1-\alpha)x_2) = |\alpha x_1 + (1-\alpha)x_2| \le \alpha|x_1| + (1-\alpha)|x_2| = \alpha f(x_1) + (1-\alpha) f(x_2),$$
by the triangular inequality (see Theorem 2.5).

Example 14.8. The function $f(x) = x^2$, $x \in \mathbb{R}$, is strictly convex on $\mathbb{R}$. To prove this, we fix $x_1, x_2 \in \mathbb{R}$, $x_1 \neq x_2$, $\alpha \in (0,1)$ and use the inequality $2x_1x_2 < x_1^2 + x_2^2$ which trivially follows from $(x_1 - x_2)^2 > 0$. Thus,
$$\begin{aligned} f(\alpha x_1 + (1-\alpha)x_2) &= (\alpha x_1 + (1-\alpha)x_2)^2 = \alpha^2 x_1^2 + 2\alpha(1-\alpha)x_1x_2 + (1-\alpha)^2 x_2^2 \\ &< \alpha^2 x_1^2 + \alpha(1-\alpha)(x_1^2 + x_2^2) + (1-\alpha)^2 x_2^2 = \alpha x_1^2 + (1-\alpha)x_2^2 = \alpha f(x_1) + (1-\alpha) f(x_2). \end{aligned}$$

Theorem 14.6. *Let a function $f : (a, b) \to \mathbb{R}$ has the derivative $f'(x)$ for all $x \in (a, b)$.*

(i) The function f is convex (strictly convex) on (a, b), if f' increases (strictly increases) on (a, b).

(ii) The function f is concave (strictly concave) on (a, b), if f' decreases (strictly decreases) on (a, b).

Combining theorems 14.6, 12.1 and 12.2 we obtain the following statement.

Theorem 14.7. *Let a function $f : (a, b) \to \mathbb{R}$ have the second derivative $f''(x)$ for all $x \in (a, b)$.*

(i) The function f is convex (concave) on (a, b) iff $f''(x) \geq 0$ (resp. $f''(x) \leq 0$) for all $x \in (a, b)$.

(ii) The function f is strictly convex (strictly concave) on (a, b) iff $f''(x) \geq 0$ (resp. $f''(x) \leq 0$) for all $x \in (a, b)$ and there is no interval $(\alpha, \beta) \subset (a, b)$ such that $f''(x) = 0$ for all $x \in (\alpha, \beta)$.

Exercise 14.4. Identify intervals on which the following functions are convex or concave:
a) $f(x) = e^x$, $x \in \mathbb{R}$; b) $f(x) = \ln x$, $x > 0$; c) $f(x) = \sin x$, $x \in \mathbb{R}$; d) $f(x) = \arctan x$, $x \in \mathbb{R}$;
e) $f(x) = x^\alpha$, $x > 0$, $\alpha \in \mathbb{R}$.

Theorem 14.8 (Jensen's inequality). *Let $f : (a, b) \to \mathbb{R}$ be a convex function. Then for each $n \geq 2$, $x_1, \ldots, x_n \in (a, b)$ and $\alpha_1, \ldots, \alpha_n \in [0, 1]$, $\alpha_1 + \ldots + \alpha_n = 1$, the inequality*

$$f(\alpha_1 x_1 + \ldots + \alpha_n x_n) \leq \alpha_1 f(x_1) + \ldots + \alpha_n f(x_n) \tag{18}$$

holds.

Proof. We are going to use the mathematical induction to prove the theorem. For $n = 2$ inequality (18) is true due to the convexity of f.

Next, we assume that inequality (18) holds for some $n \geq 2$ and each $x_1, \ldots, x_n \in (a, b)$ and each $\alpha_1, \ldots, \alpha_n \in [0, 1]$, $\alpha_1 + \ldots + \alpha_n = 1$, and prove (18) for $n + 1$ and $x_1, \ldots, x_{n+1} \in (a, b)$, $\alpha_1, \ldots, \alpha_{n+1} \in [0, 1]$, $\alpha_1 + \ldots + \alpha_{n+1} = 1$. We remark that there exists k such that $\alpha_k < 1$. So, let $\alpha_{n+1} < 1$. Then, by Definition 14.3 and the induction assumption,

$$f\left(\sum_{k=1}^{n+1} \alpha_k x_k\right) = f\left(\alpha_{n+1} x_{n+1} + \sum_{k=1}^{n} \alpha_k x_k\right) \leq \alpha_{n+1} f(x_{n+1}) + (1 - \alpha_{n+1}) f\left(\sum_{k=1}^{n} \frac{\alpha_k}{1 - \alpha_{n+1}} x_k\right)$$
$$\leq \alpha_{n+1} f(x_{n+1}) + (1 - \alpha_{n+1}) \sum_{k=1}^{n} \frac{\alpha_k}{1 - \alpha_{n+1}} f(x_k) = \sum_{k=1}^{n+1} \alpha_k f(x_k).$$

□

Example 14.9. The function $f(x) = -\ln x$, $x > 0$, is convex on $(0, +\infty)$, since $f''(x) = \frac{1}{x^2} > 0$, $x > 0$ (see Theorem 14.7 (i)). Applying (18) to f, for each $n \geq 2$, $x_1, \ldots, x_n \in (0, +\infty)$ and $\alpha_1, \ldots, \alpha_n \in [0, 1]$, $\alpha_1 + \ldots + \alpha_n = 1$, we have

$$\ln\left(\sum_{k=1}^{n} \alpha_k x_k\right) \geq \sum_{k=1}^{n} \alpha_k \ln x_k.$$

This implies

$$\prod_{k=1}^{n} x_k^{\alpha_k} \leq \sum_{k=1}^{n} \alpha_k x_k \tag{19}$$

for all $n \geq 2$, $x_1, \ldots, x_n \in (0, +\infty)$ and $\alpha_1, \ldots, \alpha_n \in [0, 1]$, $\alpha_1 + \ldots + \alpha_n = 1$. In particular, taking $\alpha_1 = \ldots = \alpha_n = \frac{1}{n}$, we get

$$\sqrt[n]{\prod_{k=1}^{n} x_k} \leq \frac{1}{n} \sum_{k=1}^{n} x_k$$

for all $n \geq 2$, $x_1, \ldots, x_n \in (0, +\infty)$, which is the **inequality of arithmetic and geometric means**.

Exercise 14.5 (Young's inequality)**.** Let $p > 1$, $q > 1$ and $\frac{1}{p} + \frac{1}{q} = 1$. Prove that $xy \leq \frac{x^p}{p} + \frac{y^q}{q}$ for all $x, y \in (0, +\infty)$.

(*Hint:* Use inequality (19))