Search for local extrema of a single variable function

- **Short summary**
- **Creation of mathematical model**
- **Determine critical points**
- **Investigation of critical points using the corollary of Weierstrass theorem**
- **Investigation of critical points based on the definition**
- **Investigation of critical points based on the second-ordered sufficient condition of extremality**
- **Smooth extremal problem with a feasible set of a closed interval type.**

Extremal problems in which the objective function depends on a single variable are called single-variable problems and they form the simplest, but nonetheless very important subclass of extremal problems. They are found both, as independent problems when solving practical problems, and as a helper task in the process of finding the approximate value of an extremum of the functions with several variables. Thus, to study these problems is relevant. The elements of differential calculus are used to find the points of the local extremum, which naturally limits the generality of the extremal problem: the objective function must be smooth, and the feasible set must be open, semi-open, or closed.

Extremal problem is called smooth if the objective function and the functions defining the feasible set are continuously differentiable in the required order. In this chapter, wherever the opposite is not stated, we assume that the objective function is a single variable function while the feasible set is open.

To write down the smooth extremal problem on an open feasible set for a single-variable objective function, the following standard notation is used:

$$\boldsymbol{f}(\boldsymbol{x}) \rightarrow \boldsymbol{extr}, \qquad \boldsymbol{x} \in \boldsymbol{M} \subset \boldsymbol{R}, \tag{3.1}$$

Where $\boldsymbol{M}$ is an open set, i.e., it can be represented by the union of a finite or countable number of the open intervals.

Short summary

The smooth extremal problem should be investigated in the following order:

1. Create a mathematical model, give a formalization in the form (3.1).

2. Solve the equation $\boldsymbol{f'(x) = 0}$ to determine critical points of the extremal problem (3.1).

3. Investigate the critical points to choose the extremal points (based on the Weierstrass theorem, on the definition, or on sufficient conditions of the extremality).

Determining the set of critical points (which we usually denote by the symbol $\boldsymbol{K}$) is very important because often this set is finite, or the number of values of the objective function on this set is finite. In any case, if we know the set of critical points $\boldsymbol{K}$, then the problem

$$\boldsymbol{f(x) \to extr, \qquad x \in M \subset R},$$

is equivalent of the problem

$$\boldsymbol{f(x) \to extr, \qquad x \in K},$$

Which is much easier to solve in most cases.

Creation of a mathematical model

Consider the mathematical model of inventory control as an example. Many firms maintain an inventory of goods to meet future demand. Among the reasons for holding inventory is to avoid the time and cost of constant replenishment. On the other hand, to replenish only infrequently would imply large inventories that would tie up unnecessary capital and incur huge storage costs. Our goal is to determine the optimal size of the inventory.

Let us consider the easiest case. Suppose that there is a steady demand of $\boldsymbol{\lambda}$ units per year for the product. Frequent replenishment is penalized by assuming that there is a setup or ordering cost of ***$K*** each time an order is placed, irrespective of how many units are ordered. The acquisition cost of each unit is ***$c***. Excessive inventory is penalized by assuming that each unit will cost $h to store for one year. To keep things simple, we will assume that all demand must be met immediately (i.e., no back orders are allowed) and that replenishment occurs instantaneously as soon as an order is sent. Figure 3.1 graphically illustrates the change in the inventory level with respect to time.

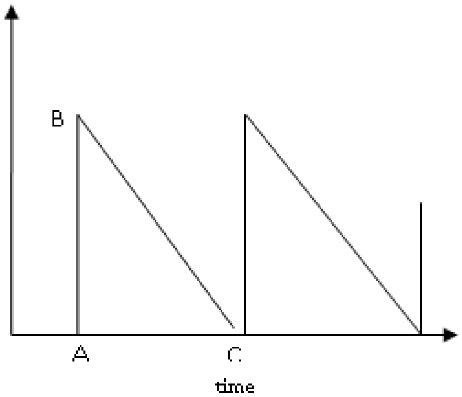


Fig. 3.1

Starting at any time A with an inventory of B, the inventory level will decrease at a rate of $\boldsymbol{\lambda}$ units per unit time until it becomes zero at time C, when a fresh order is placed. The triangle ABC represents one inventory cycle that will be repeated throughout the year. The problem is to determine the optimal order quantity B, denoted by variable $\boldsymbol{Q}$ and the common length of time C- A, denoted by $\boldsymbol{T}$, between reorders. Since T is just the length of time required to deplete Q units at rate $\boldsymbol{\lambda}$, we get $\boldsymbol{T = Q/\lambda}$

The only remaining problem is to determine $\boldsymbol{Q}$. Note that when $\boldsymbol{Q}$ is small, $\boldsymbol{T}$ will be small, implying more frequent reorders during the year. This will result in higher reorder costs but lower inventory holding cost. On the other hand, a large inventory (large $\boldsymbol{Q}$) will result in a higher inventory cost but lower reorder costs. The basic inventory control problem is to determine the optimal $\boldsymbol{Q}$ value of that will minimize the sum of the inventory cost and reorder cost in a year. We shall now develop the necessary mathematical expression to optimize the yearly cost (cost/cycle number of cycles/year).

Number of cycles (reorders)/year $= \boldsymbol{\frac{1}{T} = \frac{\lambda}{Q}}$

Cost per cycle = reorder cost+ inventory cost $= \boldsymbol{K + cQ + \frac{Q}{2}hT = K + cQ + \frac{hQ^2}{2\lambda}}$

Note: The inventory cost per cycle is simply the cost of holding an average inventory of $\boldsymbol{Q/2}$ for a length of time $\boldsymbol{T}$. Thus, the yearly cost to be minimized is $\boldsymbol{f`(Q) = \frac{\lambda K}{Q} + \lambda c + \frac{hQ}{2}}$

For the following extremal problem:

$$\boldsymbol{f(Q) \to extr, \qquad Q \in M,}$$

$\boldsymbol{f(Q) = \frac{\lambda K}{Q} + \lambda c + \frac{hQ}{2}}$ is the objective function, and $\boldsymbol{M = (0, +\infty)}$.

Determine critical points

Theorem 1 / the first-order necessary condition of extremality /.

Suppose $\boldsymbol{M} \subset \boldsymbol{R}$ is an open set, $\boldsymbol{f}: \boldsymbol{M} \to \boldsymbol{R}$ is continuously differentiable function in some neighbourhood of the point $\hat{\boldsymbol{x}} \in \boldsymbol{M}$, and $\hat{\mathbf{x}}$ is a point of the local extremum for the extremal problem

$$\boldsymbol{f}(\boldsymbol{x}) \to \boldsymbol{extr}, \qquad \boldsymbol{x} \in \boldsymbol{M}\,, \tag{3.1}$$

Then, $\boldsymbol{f}'(\hat{\boldsymbol{x}}) = \mathbf{0}$

Proof. Consider the case of the local minimum. Suppose $\hat{\boldsymbol{x}}$ is a point of the local minimum and the function $\boldsymbol{f}(\boldsymbol{x})$ is continuously differentiable in the neighbourhood of the point $\hat{\boldsymbol{x}}$. Then Taylor's expansion of $\boldsymbol{f}(\boldsymbol{x})$ about the point $\hat{\boldsymbol{x}}$ is

$$f(\hat{\boldsymbol{x}} + \boldsymbol{\varepsilon}) = f(\hat{\boldsymbol{x}}) + \boldsymbol{\varepsilon}\frac{\boldsymbol{df}(\hat{\boldsymbol{x}})}{\boldsymbol{dx}} + \boldsymbol{o}(\boldsymbol{\varepsilon}^{\mathbf{2}}), \tag{3.2}$$

Where $\boldsymbol{o}(\boldsymbol{\varepsilon}^{\mathbf{2}})$ indicates the sum of the terms in which the degree of $\boldsymbol{\varepsilon}$ is higher than 1. If $\hat{\boldsymbol{x}}$ is the point of the local minimum in $\boldsymbol{M}$, then from the definition, there must be a $\boldsymbol{\varepsilon}$ neighbourhood of $\hat{\boldsymbol{x}}$ such that for all $\boldsymbol{x}$ within a distance $\boldsymbol{\varepsilon}$:

$$\boldsymbol{f}(\boldsymbol{x}) \geq \boldsymbol{f}(\hat{\boldsymbol{x}}) \tag{3.3}$$

It follows from (3.2) and (3.3) that

$$\boldsymbol{\varepsilon}\frac{\boldsymbol{df}(\hat{\boldsymbol{x}})}{\boldsymbol{dx}} + \boldsymbol{o}(\boldsymbol{\varepsilon}^{\mathbf{2}}) \geq \mathbf{0},$$

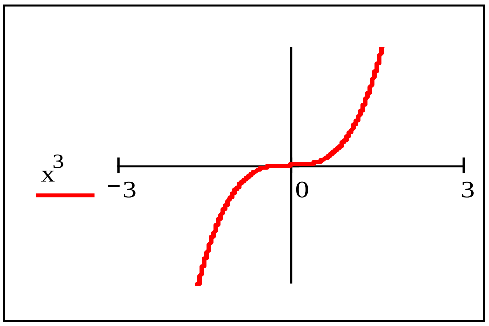
For sufficiently small $\boldsymbol{\varepsilon}$, the first term will dominate the others, and since both positive and negative $\boldsymbol{\varepsilon}$ can be chosen, it follows that inequality will hold only when

$$\frac{\boldsymbol{df}(\hat{\boldsymbol{x}})}{\boldsymbol{dx}} = \mathbf{0}$$

The proof for the local maximum is analogous. ▯

This theorem, which is known as Fermat's theorem, gives the necessary conditions for the existence of local minima and maxima, i.e., it is possible that the conditions of the theorem hold, but the function has neither a local maximum nor a local minimum. In this case, a good example is

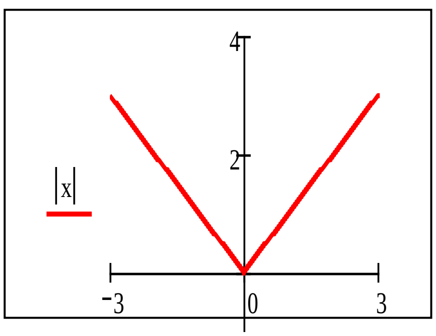
$f(x) = x^3, \mathbf{M} = \mathbf{R}\,.$



For this function the condition $f'(0) = 0$ is fulfilled, but $x = 0$ is not the point of extremum.

Remark. In this chapter we are considering smooth functions, but the function may not be differentiable and may have a local minimum or maximum.

$f(x) = |x|, \quad \mathbf{M} = \mathbf{R}\,.$



This function does not have a derivative at the point $x = 0$, although this point is the point of local minimum.

Definition. Under the conditions of the theorem discussed, the set of critical points K , for the following extremal problem:

$$f(x) \to extr, \qquad x \in M\,,$$

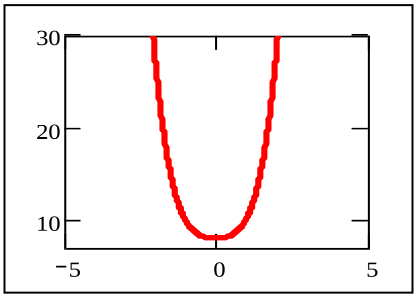
is defined as:

$$K = \{\hat{x} \in M \,|\, f'(\hat{x}) = 0\}.$$

Example. Find the points of extremum of the function $f(x) = x^4 + x^2 + 8.$

Solution. Let's write down the problem in a standard form:

$$f(x) = x^4 + x^2 + 8 \to extr, \quad x \in R. \tag{3.4}$$



To find the set of critical points for this problem, let's solve the equation: $\frac{df}{dx}=\mathbf{0}$, $\mathbf{4x^3 + 2x = 0}$, we find, that $\boldsymbol{x} = \mathbf{0}$, or $\boldsymbol{K} = \{\mathbf{0}\}$.

In the following paragraphs we explain how critical points are investigated. We will also finish solving of this example.

Investigation of critical points using the corollary of Weierstrass theorem

We can use the corollary of the Weierstrass theorem only if the feasible set is closed. If $\boldsymbol{M} = \boldsymbol{R}$, it is simultaneously open and closed set.

Using the results of the Weierstrass theorem if we establish that for the following extremal problem:

$$\boldsymbol{f}(\boldsymbol{x}) \to \boldsymbol{extr}, \quad \boldsymbol{x} \in \boldsymbol{R}, \qquad (3.5)$$

$\mathbf{glextr}(\mathbf{3.5}) \neq \emptyset$, then we get extremal problem with smaller feasible set and with the same extremum points as the problem (3.5):

$$\boldsymbol{f}(\boldsymbol{x}) \to \boldsymbol{extr}, \quad \boldsymbol{x} \in \boldsymbol{K} \qquad (3.6)$$

Because

$$(\hat{\boldsymbol{x}} \in \boldsymbol{K} \quad and \ \hat{\boldsymbol{x}} \in \ \mathbf{glextr}(\mathbf{3.5}) \Rightarrow \hat{\boldsymbol{x}} \in \ \mathbf{glextr}(\mathbf{3.6}).$$

The case of global maximum is analogous.

Example. Let's continue to solve the problem (3.4):

$$\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{x}^4 + \boldsymbol{x}^2 + \mathbf{8} \to \boldsymbol{extr}, \ \ \boldsymbol{x} \in \boldsymbol{R}.$$

As we have seen $\boldsymbol{K} = \{\mathbf{0}\}$. Since $\lim\limits_{|x|} \boldsymbol{f}(\boldsymbol{x}) = +\infty$, therefore $\mathbf{glextr}(\mathbf{3.4}) \neq \emptyset$

Because K involves only a single element $\boldsymbol{x} = \mathbf{0}$, it is therefore a global minimum for the extremal problem

$$\boldsymbol{f}(\boldsymbol{x}) \to \boldsymbol{extr}, \qquad \boldsymbol{x} \in \boldsymbol{K},$$

i.e. $\mathbf{0} \in \ \mathbf{glextr}(\mathbf{3.4})$.

Investigation of critical points based on the definition

Consider again the extremal problem (3.5). Suppose the conditions of the ***theorem 1*** are fulfilled and $\widehat{\boldsymbol{x}} \in \boldsymbol{K}$. Denote each $\boldsymbol{x} \in \boldsymbol{R}$ in the following form: $\boldsymbol{x} = \widehat{\boldsymbol{x}} + \boldsymbol{h}$ and consider the difference:

$$\boldsymbol{\Delta f} = \boldsymbol{f}(\widehat{\boldsymbol{x}} + \boldsymbol{h}) - \boldsymbol{f}(\widehat{\boldsymbol{x}})$$

It is essential to determine the dependence of $\boldsymbol{\Delta f}$ on the $\boldsymbol{h}$'s sign because:

1) if $\boldsymbol{\Delta f} \geq \mathbf{0}$ for all $\boldsymbol{h}$, then $\widehat{\boldsymbol{x}} \in \ \mathbf{glmin}(\mathbf{3.5})$;

2) if $\boldsymbol{\Delta f} \leq \mathbf{0}$ for all $\boldsymbol{h}$, then $\widehat{\boldsymbol{x}} \in \ \mathbf{glmax}(\mathbf{3.5})$;

3) if $\boldsymbol{\Delta f} \geq \mathbf{0}$ for sufficiently small values of $\boldsymbol{h}$, then$\widehat{\boldsymbol{x}} \in \ \mathbf{locmin}(\mathbf{3.5})$;

4) If $\boldsymbol{\Delta f} \leq \mathbf{0}$ for sufficiently small values of $\boldsymbol{h}$ then $\widehat{\boldsymbol{x}} \in \ \mathbf{locmax}(\mathbf{3.5})$.

Let's continue discussing extremal problem:

$$\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{x}^{\mathbf{4}} + \boldsymbol{x}^{\mathbf{2}} + \mathbf{8} \to \boldsymbol{extr}, \ \ \boldsymbol{x} \in \boldsymbol{R}$$

$\boldsymbol{K} = \{\mathbf{0}\}$ and we have a unique critical point $\widehat{\boldsymbol{x}} = \mathbf{0}$. So, $\boldsymbol{x} = \widehat{\boldsymbol{x}} + \boldsymbol{h}$ and $\boldsymbol{x} = \mathbf{0} + \boldsymbol{h}$.

Let's consider the difference:

$$\boldsymbol{f}(\boldsymbol{h}) - \boldsymbol{f}(\mathbf{0}) = \boldsymbol{h}^{\mathbf{4}} + \boldsymbol{h}^{\mathbf{2}} + \mathbf{8} - \mathbf{8} = \boldsymbol{h}^{\mathbf{4}} + \boldsymbol{h}^{\mathbf{2}} \geq \mathbf{0}.$$

The difference is non-negative for any $\boldsymbol{h}$,which means, that $\widehat{\boldsymbol{x}} = \mathbf{0}$ is the point of global minimum.

Investigation of critical points based on the second-ordered sufficient condition of extremality

Theorem 2. Suppose the first derivative of a function $\boldsymbol{f}: \boldsymbol{M} \rightarrow \boldsymbol{R}$ at a point $\widehat{\boldsymbol{x}} \boldsymbol{\epsilon} \boldsymbol{R}$ is equal to zero, while the second derivative is different from zero. Then,

A. If $\frac{d^2 f(\widehat{x})}{dx^2} > \mathbf{0}$, then $\widehat{\boldsymbol{x}}$ is the local minimum.

B. If $\frac{d^2 f(\widehat{x})}{dx^2} < \mathbf{0}$, then $\widehat{\boldsymbol{x}}$ is the local maximum.

Proof. Let us represent the function $\boldsymbol{f}(\boldsymbol{x})$ by Taylor series in the neighbourhood of the point $\widehat{\boldsymbol{x}}$. Since the first-order derivatives are equal to zero, we get:

$$\boldsymbol{f}(\widehat{\boldsymbol{x}}+\boldsymbol{\varepsilon})-\boldsymbol{f}(\widehat{\boldsymbol{x}})=\frac{\varepsilon^2}{2} \frac{d^2 f(\widehat{x})}{dx^2}+\boldsymbol{o}(\boldsymbol{\varepsilon}^3), \quad (3.7)$$

The term $\boldsymbol{\varepsilon}^2$ is always positive and for all sufficiently small $\boldsymbol{\varepsilon}$ the sign of equation (3.7), will be dominated by the first term. Hence, if $\frac{d^2 f(\widehat{x})}{dx^2} > \mathbf{0}$, the first term of (3.7) is positive, $\boldsymbol{f}(\widehat{\boldsymbol{x}}+\boldsymbol{\varepsilon})-\boldsymbol{f}(\widehat{\boldsymbol{x}})>\mathbf{0}$ and $\widehat{\boldsymbol{x}}$ corresponds to a local minimum. If $\frac{d^2 f(\widehat{x})}{dx^2} < \mathbf{0}$, the first term of (3.7) is negative, $\boldsymbol{f}(\widehat{\boldsymbol{x}}+\boldsymbol{\varepsilon})-\boldsymbol{f}(\widehat{\boldsymbol{x}})<\mathbf{0}$ and $\widehat{\boldsymbol{x}}$ corresponds to a local maximum. □

Example. Let's consider an example,

$$\boldsymbol{f}(\boldsymbol{x})=\boldsymbol{x}^4+\boldsymbol{x}^2+\mathbf{8} \rightarrow \boldsymbol{extr}, \quad \boldsymbol{x} \in \boldsymbol{R}.$$

we have found that $\boldsymbol{x}=\mathbf{0}$ is a critical point. Write down sufficient conditions.

$$\boldsymbol{f}'(\boldsymbol{x})=\mathbf{4}\boldsymbol{x}^3+\mathbf{2}\boldsymbol{x} \text{ and } \boldsymbol{f}'(\mathbf{0})=\mathbf{0}$$

$$\boldsymbol{f}''(\boldsymbol{x})=\mathbf{12}\boldsymbol{x}^2+\mathbf{2} \text{ and } \boldsymbol{f}''(\mathbf{0})=\mathbf{2}$$

$\boldsymbol{f}''(\mathbf{0})=\mathbf{2}$ is positive, e.g., $\boldsymbol{x}=\mathbf{0}$ is the local minimum.

Theorem 3. Suppose the first derivative of a function $\boldsymbol{f}: \boldsymbol{M} \rightarrow \boldsymbol{R}$ at a point $\widehat{\boldsymbol{x}} \boldsymbol{\epsilon} \boldsymbol{R}$ is equal to zero and the first nonzero higher order derivative is denoted by $\boldsymbol{n}$. Then,

- If n is odd, then $\widehat{\boldsymbol{x}}$ is a point of inflection.
- If n is even, then $\widehat{\boldsymbol{x}}$ is a local extremum.

A. If $\frac{d^n f(\widehat{x})}{dx^n} > \mathbf{0}$, then $\widehat{\boldsymbol{x}}$ is the local minimum.

B. If $\frac{d^n f(\widehat{x})}{dx^n} < \mathbf{0}$, then $\widehat{\boldsymbol{x}}$ is the local maximum.

Example. $\boldsymbol{f(x) = x^3, \mathrm{M} = \mathrm{R}}$

For this extremal problem, $\boldsymbol{K = \{0\}}$.

$$\frac{\boldsymbol{df(0)}}{\boldsymbol{dx}} = \boldsymbol{0} \qquad \frac{\boldsymbol{d^2f(\hat{0})}}{\boldsymbol{dx^2}} = \boldsymbol{0} \quad \frac{\boldsymbol{d^3f(\hat{0})}}{\boldsymbol{dx^3}} = \boldsymbol{6.}$$

Thus, the first non-vanishing derivative is 3, which is an odd number and $\boldsymbol{x = 0}$ **is** an inflection point.

Smooth extremal problem with a feasible set of a closed interval type.

Consider the extremal problem:

$$\boldsymbol{f(x) \to extr}, \qquad \boldsymbol{x \in [a, b]} \qquad (3.8)$$

This problem is investigated according to the following scheme:

1. Determine critical points:
$$\boldsymbol{K = \{x \in (a, b) | f'(x) = 0\} \cup \{a, b\}};$$

2. Investigate critical points:

If, $\boldsymbol{\hat{x} \epsilon K} \cap \boldsymbol{(a, b)}$, then the investigation is to be carried out by the methods discussed in the previous section, and for the ends of the closed interval a and b , we have:

$$\boldsymbol{f'(a) > 0 \Rightarrow\ a \in locmin(3.8)}$$

$$\boldsymbol{f'(a) < 0 \Rightarrow\ a \in locmax(3.8)}$$

$$\boldsymbol{f'(b) > 0 \Rightarrow\ b \in locmax(3.8)}$$

$$\boldsymbol{f'(b) < 0 \Rightarrow\ b \in locmin(3.8)}$$

3. Find the points, where largest and smallest value of $\boldsymbol{f(x)}$ out of $\boldsymbol{f(a), f(b), f(x_1)}$, $\boldsymbol{f(x_2), \dots, f(x_n)}$ is attained. ($\boldsymbol{x_1, x_2, \dots, x_n}$ are critical points of the function $\boldsymbol{f(x)}$). These points become the global maximum and global minimum respectively.