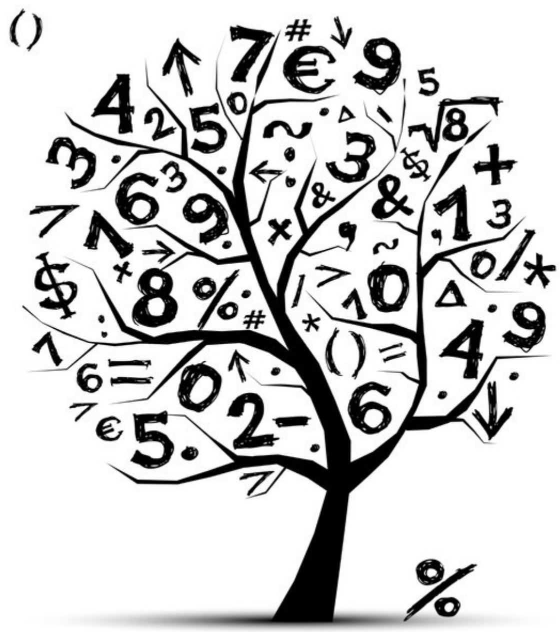
Last update: March 16, 2024

Global and local optima



- **Author**: Fabio Furini

Contents

1 Global optima and point of global optima

Definition 1: point of global minimum and global minimum

Given a scalar function $f : F \subseteq \mathbb{R}^n \to \mathbb{R}$, if there exists $\tilde{\boldsymbol{x}} \in F$ such that:

$$f(\tilde{\boldsymbol{x}}) \;\leq\; f(\boldsymbol{x}), \quad \forall \boldsymbol{x} \in F$$

*then the point $\tilde{\boldsymbol{x}}$ is called **point of global minimum** and the value $f(\tilde{\boldsymbol{x}})$ is called the **global minimum** of f in F.*

Definition 2: point of global maximum and global maximum

Given a scalar function $f : F \subseteq \mathbb{R}^n \to \mathbb{R}$, if there exists $\tilde{\boldsymbol{x}} \in F$ such that:

$$f(\tilde{\boldsymbol{x}}) \;\geq\; f(\boldsymbol{x}), \quad \forall \boldsymbol{x} \in F$$

*then the point $\tilde{\boldsymbol{x}}$ is called **point of global maximum** and the value $f(\tilde{\boldsymbol{x}})$ is called the **global maximum** of f in F.*

A maximum or a minimum if it exists it is unique, on the other hand there can exist several points of maximum or minimum (even infinite). We call **optimum** a maximum or a minimum and **point of optimum** a point of maximum or a point of minimum.

- Given a scalar function $f : F \subseteq \mathbb{R}^n \to \mathbb{R}$, we have:

$$\min_{\boldsymbol{x} \in F} f(\boldsymbol{x}) = -\max_{\boldsymbol{x} \in F} -f(\boldsymbol{x})$$

moreover we have:

$$\min_{\boldsymbol{x} \in F} \big(f(\boldsymbol{x}) + k \big) = k + \min_{\boldsymbol{x} \in F} f(\boldsymbol{x}), \quad \forall k \in \mathbb{R}$$

$$\min_{\boldsymbol{x} \in F} \big(k \cdot f(\boldsymbol{x}) \big) = k \cdot \min_{\boldsymbol{x} \in F} f(\boldsymbol{x}), \quad \forall k \in \mathbb{R}_+$$

2 Local optima and point of local optima

Definition 3: spherical neighborhood

*Given a point $\bar{\boldsymbol{x}} \in \mathbb{R}^n$ and a positive constant $\delta > 0$, the associated **spherical neighborhood** is:*

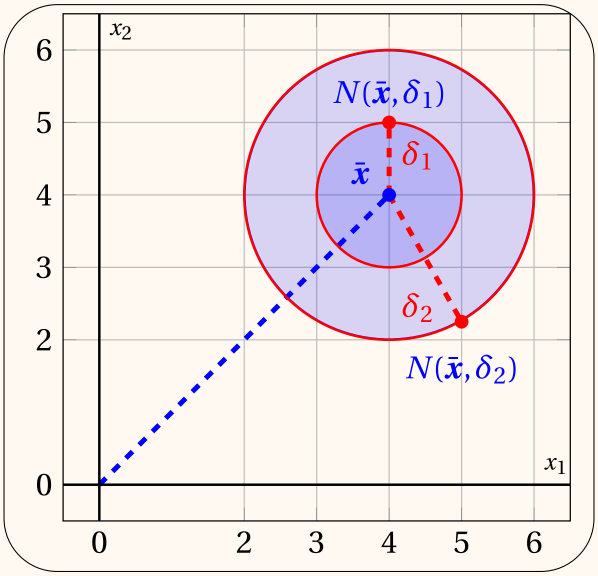
$$N(\bar{\boldsymbol{x}}, \delta) = \Big\{ \boldsymbol{x} \in \mathbb{R}^n : \; ||\bar{\boldsymbol{x}} - \boldsymbol{x}|| < \delta \Big\}$$

- The value $||\bar{\boldsymbol{x}} - \boldsymbol{x}||$ corresponds to the **euclidean distance** between the generic point $\boldsymbol{x} \in \mathbb{R}^n$ and the point $\bar{\boldsymbol{x}} \in \mathbb{R}^n$, the center of the neighborhood. The euclidean distance is given by the formula:

$$||\bar{\boldsymbol{x}} - \boldsymbol{x}|| = \sqrt{\sum_{j=1}^{n} (\bar{x}_j - x_j)^2}$$

Example 1: of spherical neighborhoods

Consider the point $\bar{\boldsymbol{x}} = (4,4) \in \mathbb{R}^2$, $\delta_1 = 1$ and $\delta_2 = 2$. The associated spherical neighborhoods are depicted by the light blue shaded portions of space:



Definition 4: point of local minimum and local minimum

Given a scalar function $f : F \subseteq \mathbb{R}^n \to \mathbb{R}$, if there exists $\bar{\boldsymbol{x}} \in F$ and a spherical neighborhood $N(\bar{\boldsymbol{x}}, \delta)$ with $\delta > 0$ such that:

$$f(\bar{\boldsymbol{x}}) \leq f(\boldsymbol{x}), \quad \forall \boldsymbol{x} \in N(\bar{\boldsymbol{x}}, \delta) \cap F$$

*then the point $\bar{\boldsymbol{x}}$ is called **point of local minimum** and the value $f(\bar{\boldsymbol{x}})$ is called the **local minimum** of f in $N(\bar{\boldsymbol{x}}, \delta) \cap F$.*

Definition 5: point of local maximum and local maximum

Given a scalar function $f : F \subseteq \mathbb{R}^n \to \mathbb{R}$, if there exists $\bar{\boldsymbol{x}} \in F$ and a spherical neighborhood $N(\bar{\boldsymbol{x}}, \delta)$ with $\delta > 0$ such that:

$$f(\bar{\boldsymbol{x}}) \geq f(\boldsymbol{x}), \quad \forall \boldsymbol{x} \in N(\bar{\boldsymbol{x}}, \delta) \cap F$$

*then the point $\bar{\boldsymbol{x}}$ is called **point of local maximum** and the value $f(\bar{\boldsymbol{x}})$ is called the **local maximum** of f in $N(\bar{\boldsymbol{x}}, \delta) \cap F$.*

Theorem 1

Given a convex function $f : F \subseteq \mathbb{R}^n \to \mathbb{R}$, if a point $\bar{\boldsymbol{x}} \in F$ is a point of local minimum then it is point of global minimum of f in F.

Proof. If $\bar{\boldsymbol{x}} \in F$ is a point of local minimum, there exists a spherical neighborhood $N(\bar{\boldsymbol{x}}, \delta)$, with $\delta > 0$, such that:

$$f(\bar{\boldsymbol{x}}) \;\le\; f(\hat{\boldsymbol{x}}), \quad \forall \hat{\boldsymbol{x}} \in N(\bar{\boldsymbol{x}}, \delta) \cap F$$

Since F is convex, for any $\boldsymbol{x} \in F$ there exist a point $\hat{\boldsymbol{x}} \in N(\bar{\boldsymbol{x}}, \delta) \cap F$ and $\lambda \in [0, 1]$ such that:

$$\hat{\boldsymbol{x}} = \lambda\, \boldsymbol{x} + (1 - \lambda)\, \bar{\boldsymbol{x}}$$

The point $\hat{\boldsymbol{x}}$ is a convex combination of the point of local minimum $\bar{\boldsymbol{x}}$ and the point $\boldsymbol{x} \in F$. Since f is convex, we then have:

$$f(\bar{\boldsymbol{x}}) \;\le\; f(\hat{\boldsymbol{x}}) \;\le\; \lambda\, f(\boldsymbol{x}) + (1 - \lambda)\, f(\bar{\boldsymbol{x}})$$

It follows that:

$$f(\bar{\boldsymbol{x}}) \;\le\; \lambda\, f(\boldsymbol{x}) + f(\bar{\boldsymbol{x}}) - \lambda\, f(\bar{\boldsymbol{x}}) \implies \lambda\, f(\bar{\boldsymbol{x}}) \;\le\; \lambda\, f(\boldsymbol{x}) \implies f(\bar{\boldsymbol{x}}) \;\le\; f(\boldsymbol{x})$$

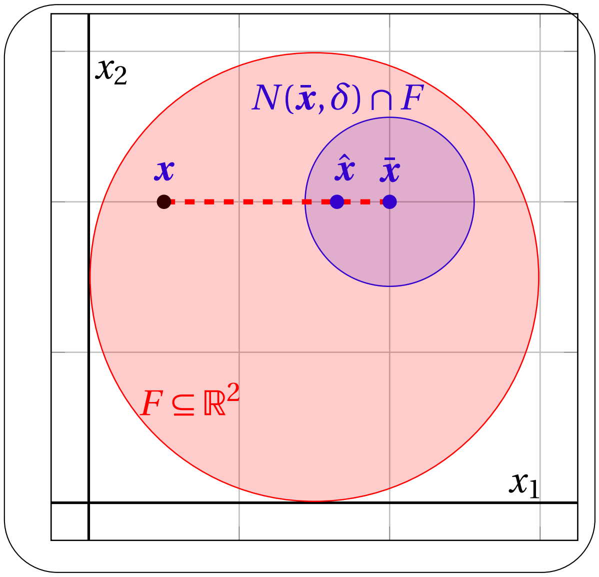
The last implication is true since $0 \le \lambda \le 1$. We have then shown that:

$$f(\bar{\boldsymbol{x}}) \le f(\boldsymbol{x}), \quad \forall \boldsymbol{x} \in F$$

and, accordingly, $\bar{\boldsymbol{x}}$ is also a point of global minimum.

□

In $\mathbb{R}^2$, the idea of the proof is given by the picture which shows the reasoning taking as an example a point $\boldsymbol{x}$ in F and a point $\hat{\boldsymbol{x}}$ in $N(\bar{\boldsymbol{x}}, \delta) \cap F$.



- Given a convex function $f: F \subseteq \mathbb{R}^n \to \mathbb{R}$, we have shown that:

$$\text{point of local minimum} \quad \Longrightarrow \quad \text{point of global minimum}$$

so "point of local minimum" is a sufficient condition to "point of global minimum" and "point of global minimum" is a necessary condition to "point of local minimum". The implication clearly works also in the other direction:

$$\text{point of global minimum} \quad \Longrightarrow \quad \text{point of local minimum}$$

so "point of global minimum" is a sufficient condition to "point of local minimum" and "point of local minimum" is a necessary condition to "point of global minimum".

So we finally have:

$$\text{point of global minimum} \quad \Longleftrightarrow \quad \text{point of local minimum}$$

so "point of local minimum" is a necessary and sufficient condition to "point of global minimum " and vice versa.

Corollary 1

Given a convex function $f: F \subseteq \mathbb{R}^n \to \mathbb{R}$, a point $\bar{\boldsymbol{x}} \in F$ is a point of global minimum if and only if it is point of local minimum of f in F.

- Clearly the previous corollary extends also to the local minimum and global minimum.

Corollary 2

Given a convex function $f: F \subseteq \mathbb{R}^n \to \mathbb{R}$, the value $f(\bar{\boldsymbol{x}})$ with $\bar{\boldsymbol{x}} \in F$ is a global minimum if and only if it is local minimum of f in F.

- For **concave functions** we have the following results:

Theorem 2

Given a concave function $f: F \subseteq \mathbb{R}^n \to \mathbb{R}$, if a point $\bar{\boldsymbol{x}} \in F$ is a point of local maximum then it is point of global maximum of f in F.

Proof. It is sufficient to follow the same reasoning of the previous proof and replacing "$\leq$" with "$\geq$' and the term convex with concave for the function f. $\square$

Corollary 3

Given a concave function $f : F \subseteq \mathbb{R}^n \to \mathbb{R}$, a point $\bar{\boldsymbol{x}} \in F$ is a point of global maximum if and only if it is point of local maximum of f in F.

Corollary 4

Given a concave function $f : F \subseteq \mathbb{R}^n \to \mathbb{R}$, the value $f(\bar{\boldsymbol{x}})$ with $\bar{\boldsymbol{x}} \in F$ is a global maximum if and only if it is local maximum of f in F.

3 Gradients and contours lines

Definition 6: gradient

*Given a scalar function $f : F \subseteq \mathbb{R}^n \to \mathbb{R}$, the **gradient** $\nabla f(\boldsymbol{x})$ of the function f is the vector whose components along the axes are the partial derivatives of the function f with respect to each variable:*

$$\nabla f(\boldsymbol{x}) = \begin{bmatrix} \frac{\partial f(\boldsymbol{x})}{\partial x_1} & \frac{\partial f(\boldsymbol{x})}{\partial x_2} & \dots & \frac{\partial f(\boldsymbol{x})}{\partial x_n} \end{bmatrix}$$

The direction of the gradient $\nabla f(\boldsymbol{x})$ (a vector in $\mathbb{R}^n$) is that in which the derivative of the function f has its **maximum value**, i.e., where the function f increases the most. Accordingly $-\nabla f(\boldsymbol{x})$ is the direction in which the derivative of the function f has its **minimum value**, i.e., where the function f decreases the most.

Definition 7: contour lines

*Given a scalar function $f : F \subseteq \mathbb{R}^n \to \mathbb{R}$, a **contour line** of the function f is the set of points $\boldsymbol{x} \in F$ such that $f(\boldsymbol{x}) = k$ for some constant $k \in \mathbb{R}$.*