Week 2

Inclass

210914

Contents

- Class Review (in Korean)
- Taylor Series and the proof of the condition (in English)
- Gradient and Hessian matrix (in English)
- Solving exercises (in English and Korean)

Formulation of Optimization Problem

$$\underset{\mathbf{x}}{\text{minimize}} \quad f(\mathbf{x})$$
$$\text{subject to} \quad \mathbf{x} \in \mathcal{X}$$

- x_i : a design variable (구성하려고 하는 시스템의 속성을 수치화하여 표현하는 변수)
- $\boldsymbol{x} = (x_1,\ x_2,\ \ldots, x_n)$: a design point (수치화한 속성들을 모아서 구성한 벡터. n차원 공간상의 점에 대응)
- $\boldsymbol{f}$: The objective function (디자인포인트를 실수에 대응시키는 함수로 현재 디자인의 성능을 평가)
- $\boldsymbol{X}$: The feasible set (시스템 속성의 제한 constrints 사항을 모두 만족하는 design point의 집합)
- A design point that minimizes $\boldsymbol{f}$ is called a solution.
- Maximizing problems can be converted to minimizing problems.

No Free Lunch Theorem

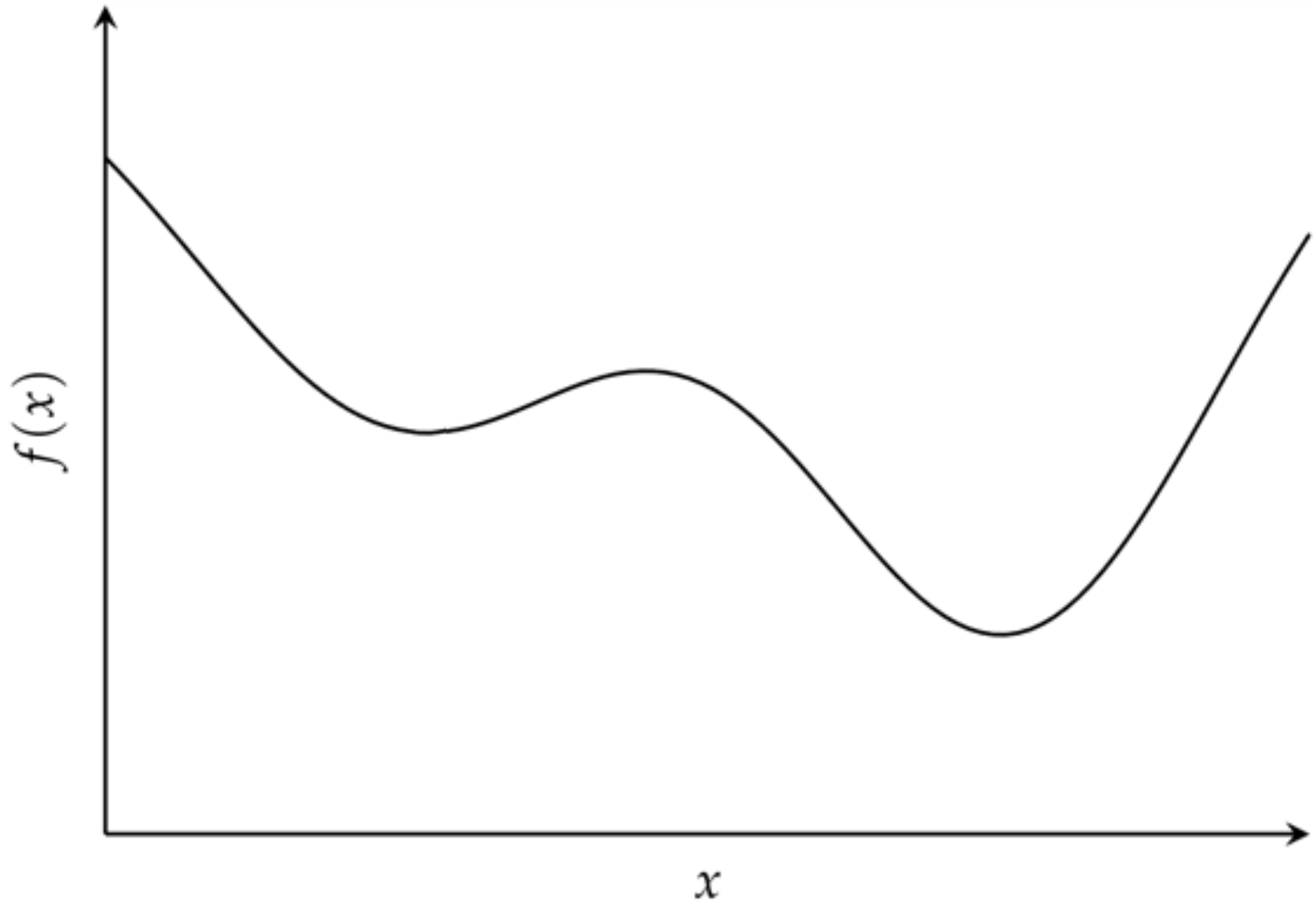
There is no magical algorithm.

- Wolpert, D.H., Macready, W.G. (1997), "No Free Lunch Theorems for Optimization"
- All optimization algorithms perform equally well when their performance is averaged over all possible cases.

	Algorithm 1	Algorithm 2	Algorithm 3	Algorithm 4	...
Problem 1	74	153	12	63	...
Problem 2	35	98	34	102	...
Problem 3	163	102	23	94	...
Problem 4	223	22	723	99	...
Problem 5	62	3	54	33	...
...	...	...	...	...	...
Average	100	100	100	100	

Where is the solution?

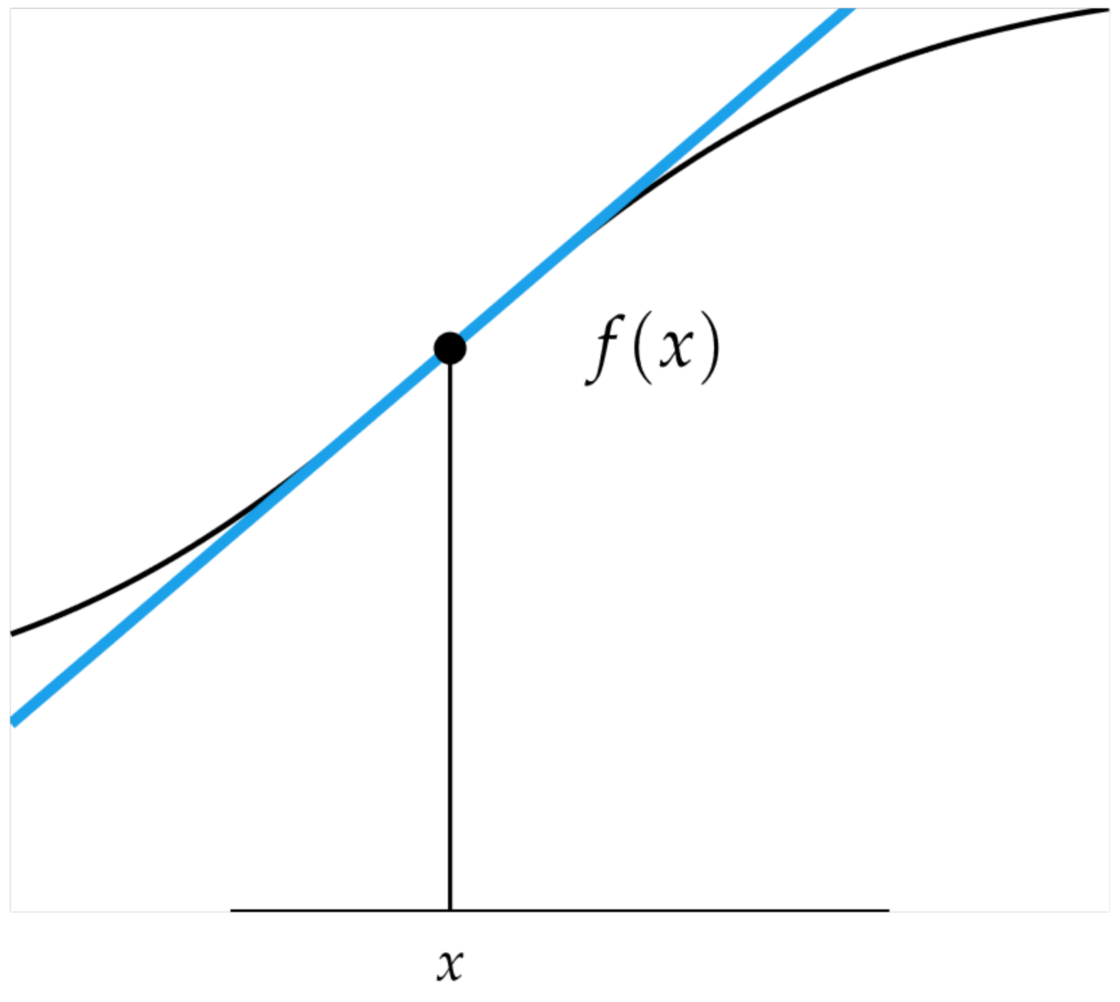
$$\begin{aligned} \underset{\mathbf{x}}{\text{minimize}} \quad & f(\mathbf{x}) \\ \text{subject to} \quad & \mathbf{x} \in \mathcal{X} \end{aligned}$$

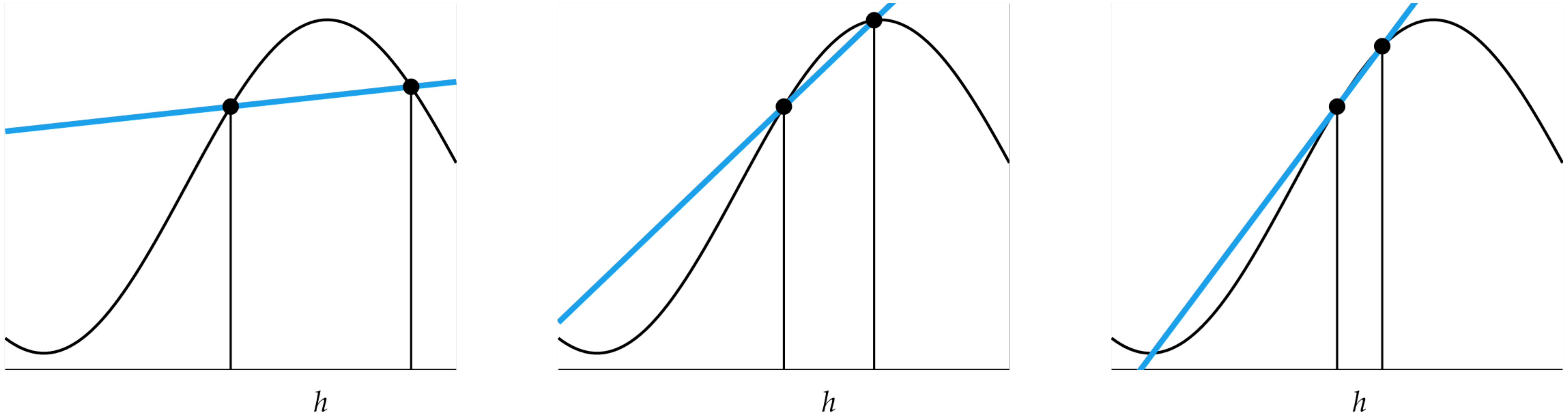


Derivatives 미분

Inspection method for functions

- The derivative $\boldsymbol{f'(x)}$ of a function $\boldsymbol{f}$ measures the behavior of $\boldsymbol{f}$ around $\boldsymbol{x}$.
- For the univariate case, $\boldsymbol{f'(x)}$ measure the slope of the tangent line(접선의 기울기) of $\boldsymbol{f}$ at $\boldsymbol{x}$.





$$f'(x) \equiv \frac{df(x)}{dx}$$

$$f'(x) \equiv \underbrace{\lim_{h \to 0} \frac{f(x+h) - f(x)}{h}}_{\text{forward difference}}$$

Symbolic Differentiation

Producing a new formula which is the derivative of a given formula

- *Derivatives of powers*:

$$\frac{d}{dx}x^a = ax^{a-1}.$$

- *Exponential and logarithmic functions*:

$$\frac{d}{dx}e^x = e^x.$$

$$\frac{d}{dx}a^x = a^x \ln(a), \qquad a > 0$$

$$\frac{d}{dx}\ln(x) = \frac{1}{x}, \qquad x > 0.$$

$$\frac{d}{dx}\log_a(x) = \frac{1}{x\ln(a)}, \qquad x, a > 0$$

- *Trigonometric functions*:

$$\frac{d}{dx}\sin(x) = \cos(x).$$

$$\frac{d}{dx}\cos(x) = -\sin(x).$$

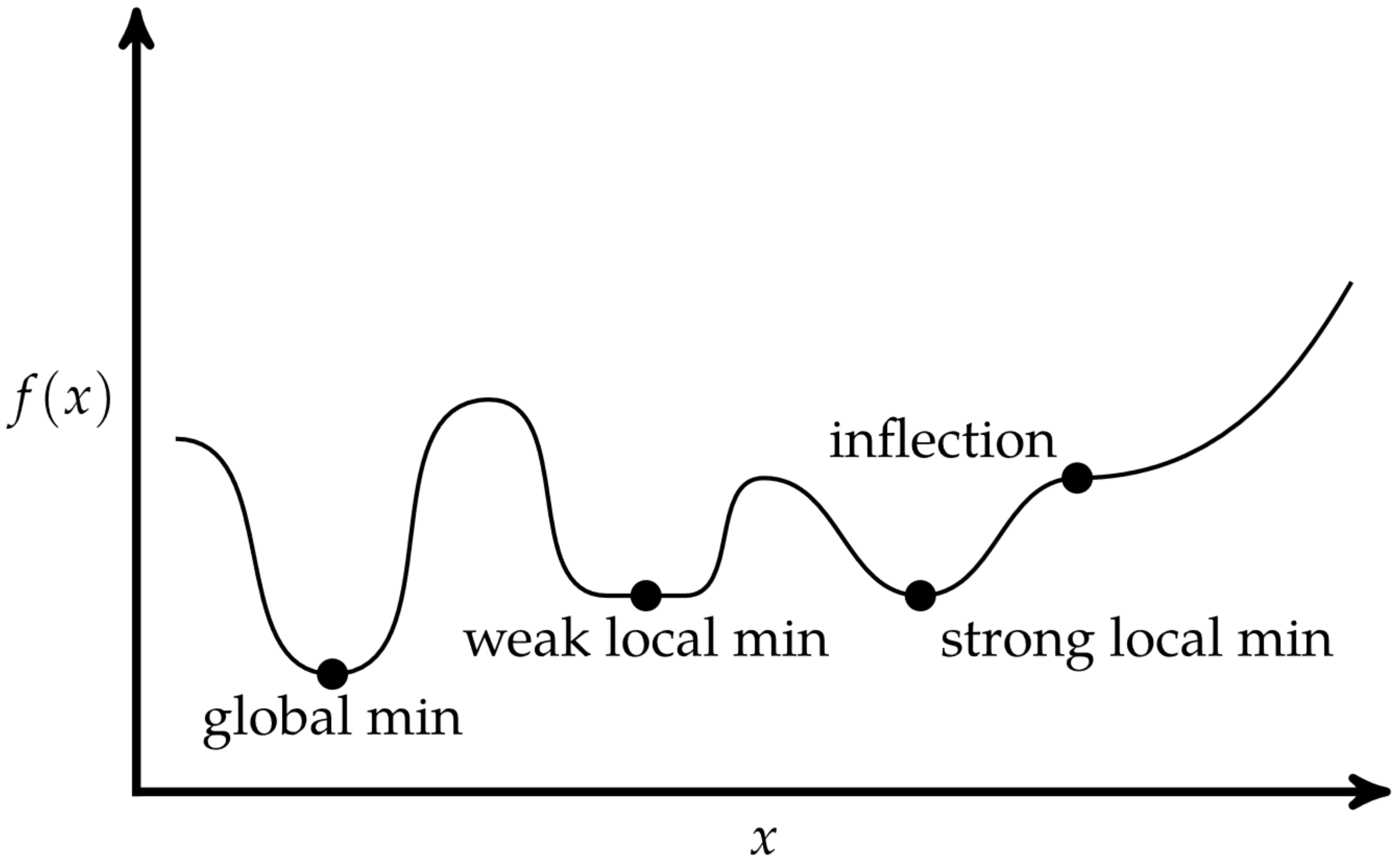
$$\frac{d}{dx}\tan(x) = \sec^2(x) = \frac{1}{\cos^2(x)} = 1 + \tan^2(x).$$

```julia
julia> using SymEngine
julia> @vars x; # define x as a symbolic variable
julia> f = x^2 + x/2 - sin(x)/x;
julia> diff(f, x)
1/2 + 2*x + sin(x)/x^2 - cos(x)/x
```

Critical Points

Candidates for the solution

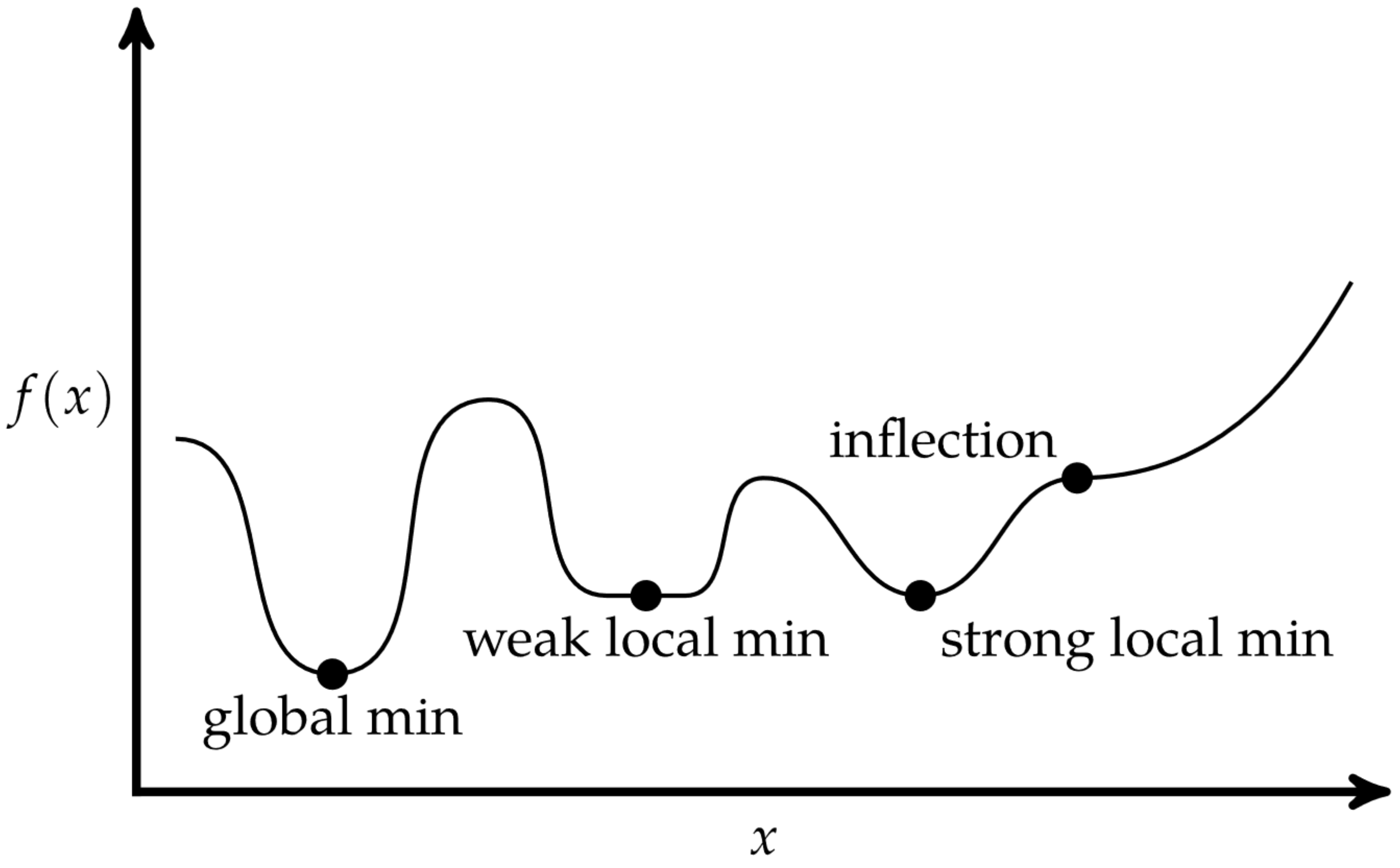
- $\boldsymbol{x}^*$ is a local minimum if there exists $\delta > 0$ such that $\boldsymbol{f}\left(\boldsymbol{x}^*\right) \leq \boldsymbol{f}(\boldsymbol{x})$ for all $\boldsymbol{x}$ with $\left|\boldsymbol{x} - \boldsymbol{x}^*\right| < \boldsymbol{\delta}$.
- The derivative is zero at all local and global minima of a differentiable function.



Conditions for Local Minima

Candidates for the solution

- If $\boldsymbol{x}^*$ is a local minimum, then $\boldsymbol{f}'\left(\boldsymbol{x}^*\right) = \mathbf{0}$ and $\boldsymbol{f}''\left(\boldsymbol{x}^*\right) \geq \mathbf{0}$.
- If $\boldsymbol{f}'\left(\boldsymbol{x}^*\right) = \mathbf{0}$ and $\boldsymbol{f}''\left(\boldsymbol{x}^*\right) > \mathbf{0}$, then $\boldsymbol{x}^*$ is a strong local minimum.

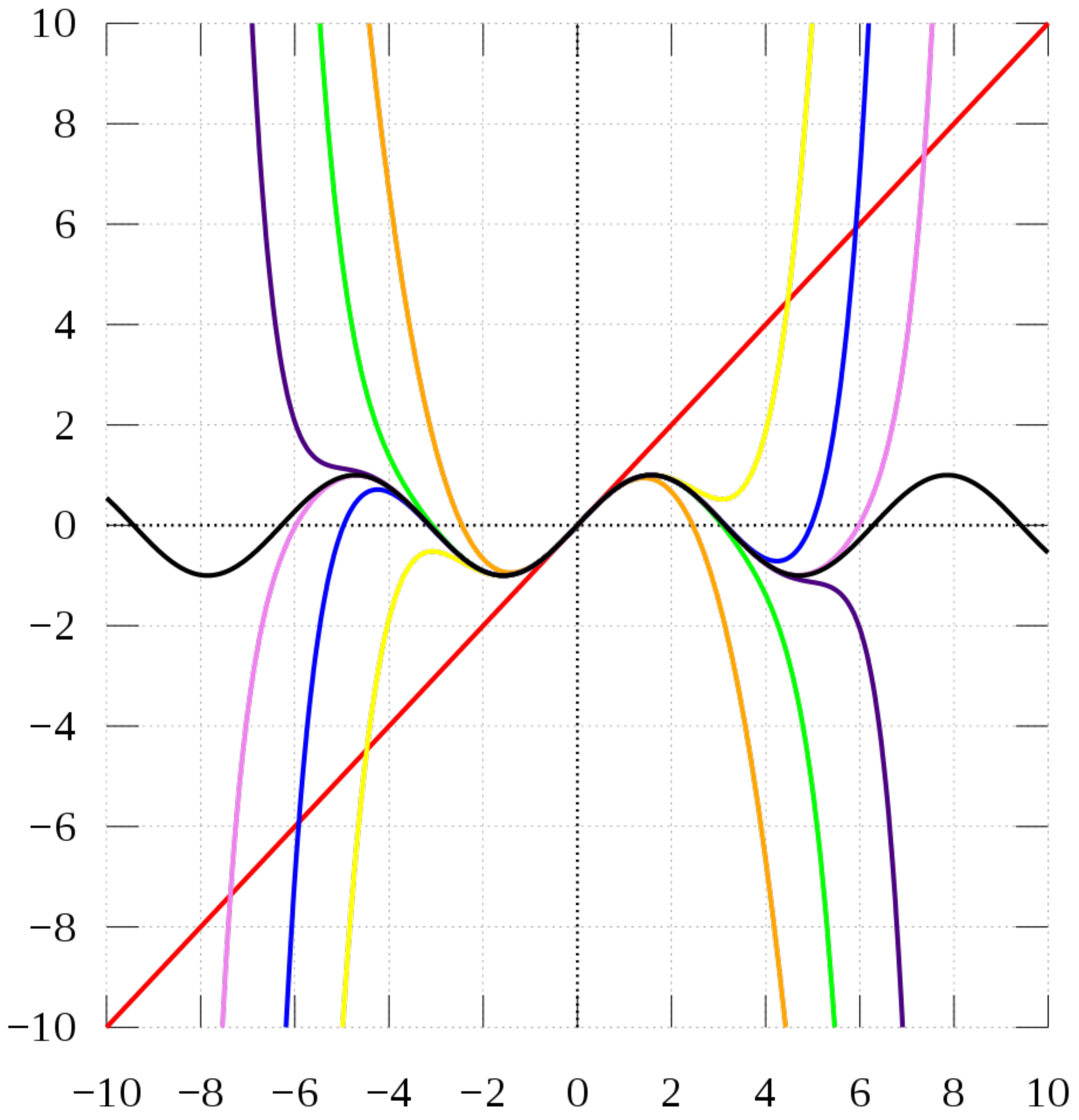


Taylor series

- an infinite sum of terms that are expressed in terms of the function's derivatives at a single point

- The Taylor series of a function $f(x)$ that is infinitely differentiable at a:

$$f(x) = f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \frac{f'''(a)}{3!}(x-a)^3 + \ldots$$



Conditions for Local Minima - proof

Candidates for the solution

$$f(x) = f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \frac{f'''(a)}{3!}(x-a)^3 \ldots$$

- If $\boldsymbol{x}^*$ is a local minimum, then $\boldsymbol{f}'\left(\boldsymbol{x}^*\right) = \mathbf{0}$.

$$f\left(\boldsymbol{x}^* + h\right) = f(x^*) + \frac{f'\left(x^*\right)}{1!}(h) + \frac{f''\left(x^*\right)}{2!}\left(h\right)^2 + \ldots$$

$$f\left(\boldsymbol{x}^* - h\right) = f\left(x^*\right) + \frac{f'\left(x^*\right)}{1!}(-h) + \frac{f''\left(x^*\right)}{2!}\left(-h\right)^2 + \ldots$$

- If $\boldsymbol{x}^*$ is a local minimum, then $\boldsymbol{f}''\left(\boldsymbol{x}^*\right) \geq \mathbf{0}$.

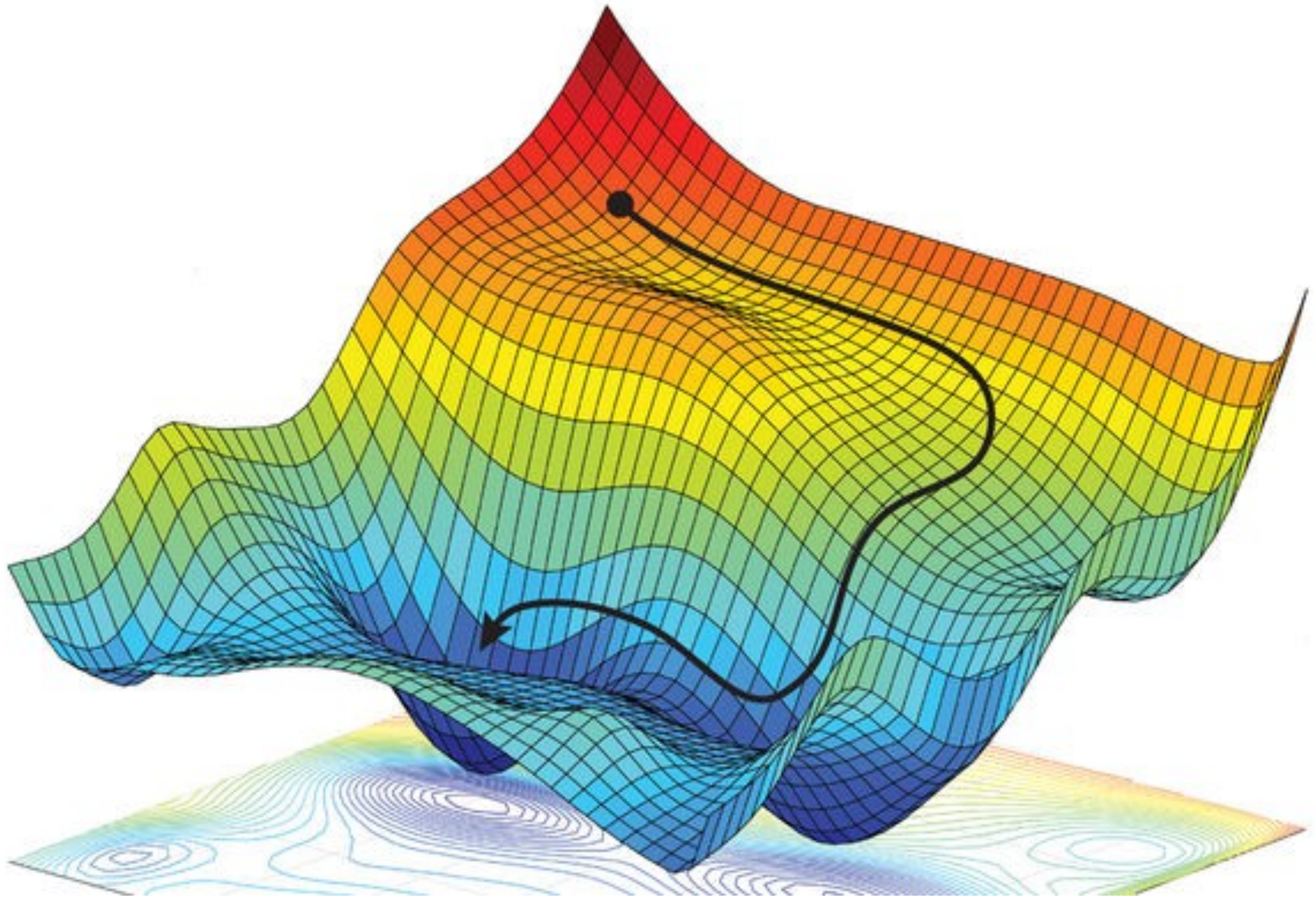
$$f\left(\boldsymbol{x}^* + h\right) = f\left(x^*\right) + \frac{f'\left(x^*\right)}{1!}(h) + \frac{f''\left(x^*\right)}{2!}\left(h\right)^2 + \frac{f'''\left(x^*\right)}{3!}\left(h\right)^3 \ldots$$

$$f\left(\boldsymbol{x}^* - h\right) = f\left(x^*\right) + \frac{f'\left(x^*\right)}{1!}(-h) + \frac{f''\left(x^*\right)}{2!}\left(h\right)^2 + \frac{f'''\left(x^*\right)}{3!}\left(-h\right)^3 \ldots$$

Gradient and Hessian matrix

- Partial derivative of a function of several variables is its derivative with respect to one of those variables, with the others held constant

$$\nabla f = \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{pmatrix} \quad \text{and} \quad \nabla^2 f = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{pmatrix}$$



Excercise

- **Exercise 1.1.** Give an example of a function with a local minimum that is not a global minimum.

Excercise

- **Exercise 1.2.** What is the minimum of the function $f(x) = x^3 - x$?

Excercise

- **Exercise 1.3.** Does the first-order condition $f\,'(x) = 0$ hold when x is the optimal solution of a constrained problem?

Excercise

- **Exercise 1.4.** How many minima does $f(x, y) = x^3 + y$, subject to $x > y \geq 1$, have?

Excercise

- **Exercise 1.5.** How many inflection points does $x^3 - 10$ have?