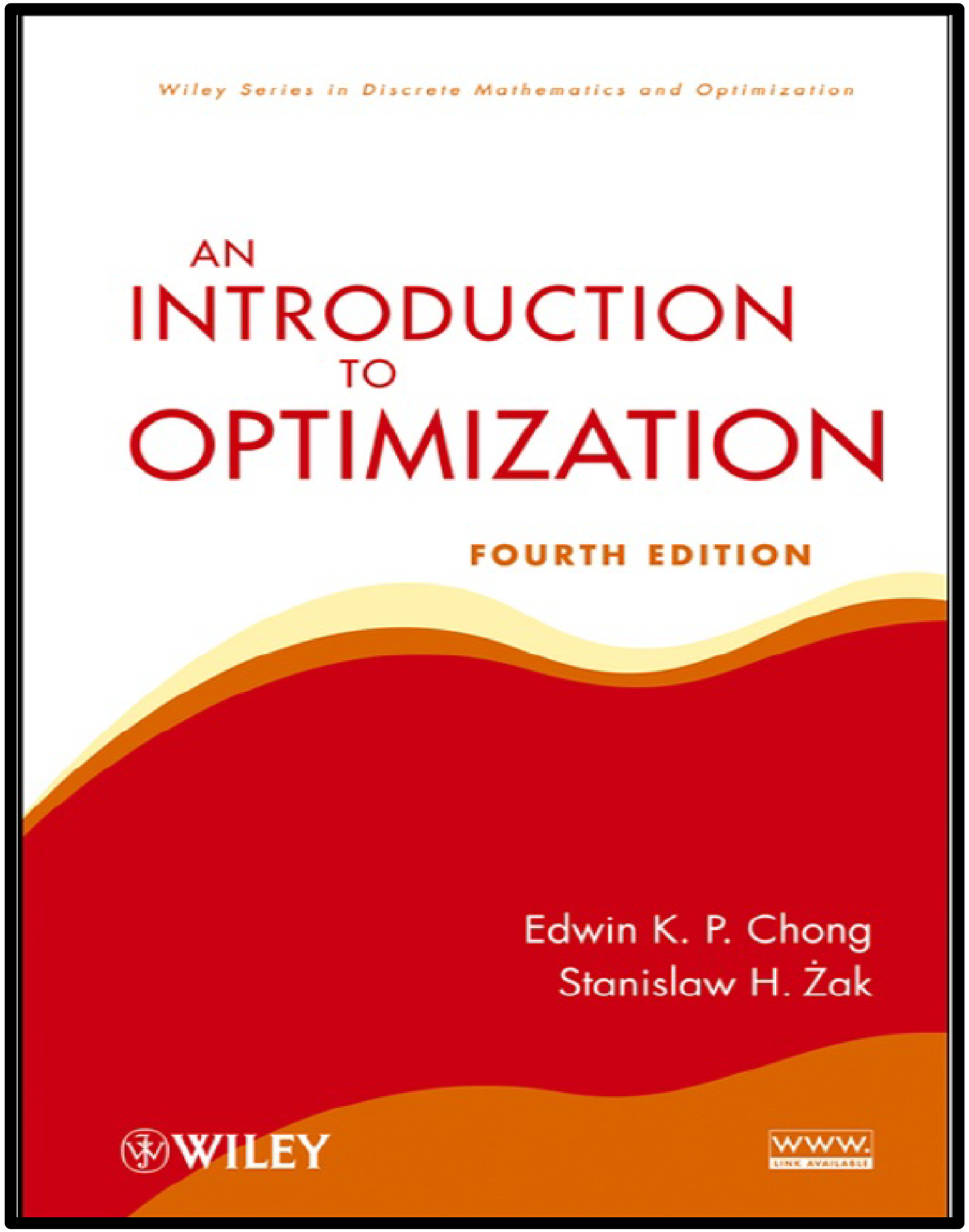
# Optimization Methods

**1.** Introduction.
**2.** Greedy algorithms for combinatorial optimization.
**3.** LS and neighborhood structures for combinatorial optimization.
**4.** Variable neighborhood search, neighborhood descent, SA, TS, EC.
**5**. Branch and bound algorithms, and subset selection algorithms.
**6.** Linear programming problem formulations and applications.
**7.** Linear programming algorithms.
**8.** Integer linear programming algorithms.
**9.** Unconstrained nonlinear optimization and gradient descent.
**10.** Newton's methods and Levenberg-Marquardt modification.
**11.** Quasi-Newton methods and conjugate direction methods.
**12.** Nonlinear optimization with equality constraints.
**13.** Nonlinear optimization with inequality constraints.
**14.** Problem formulation and concepts in multi-objective optimization.
**15.** Search for single final solution in multi-objective optimization.
**16:** Search for multiple solutions in multi-objective optimization.

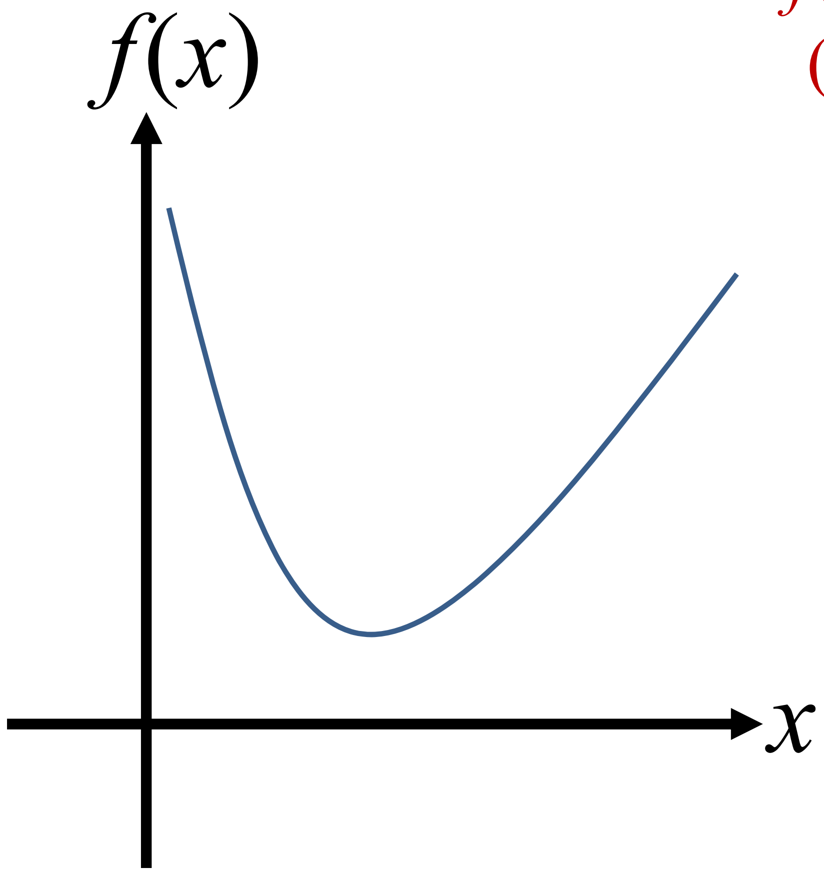
# Non-Linear Optimization is based on this book.

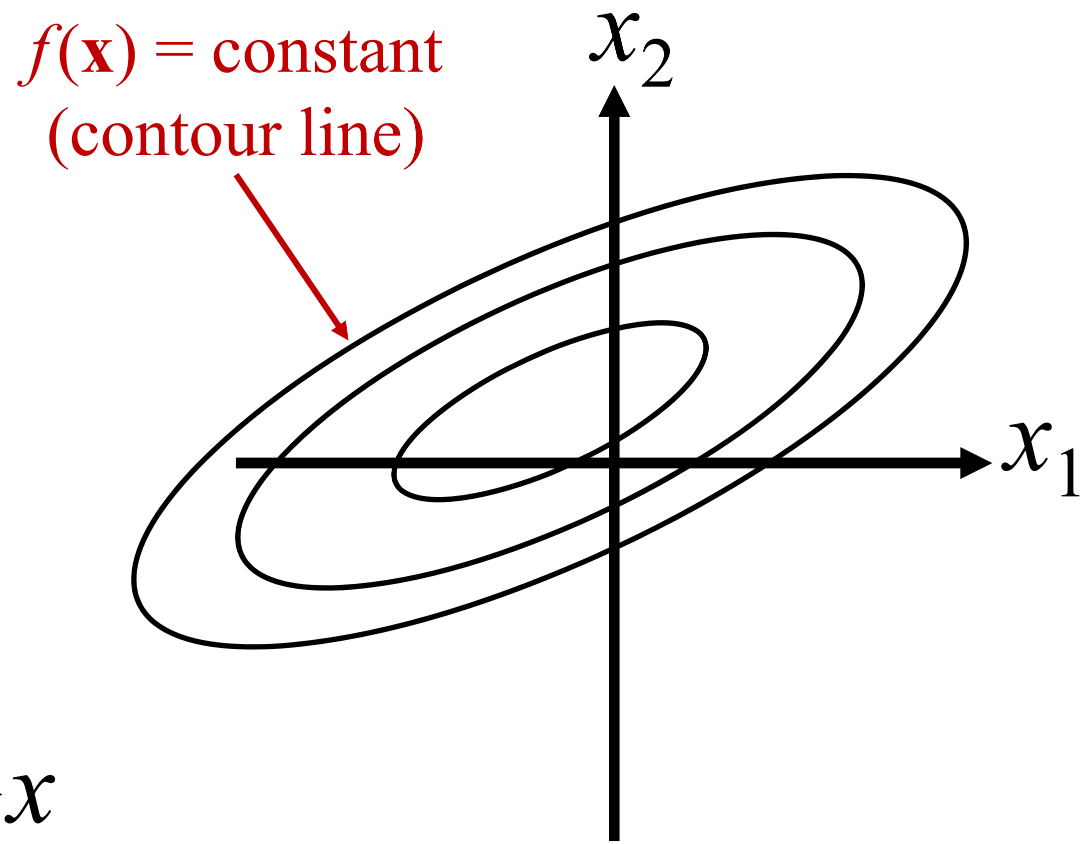


# Unconstrained nonlinear optimization Problem

Minimize $f(\mathbf{x})$, $\mathbf{x} = (x_1, x_2, \ldots, x_n)^T \in \Re^n$

$f(\mathbf{x})$: Nonlinear function

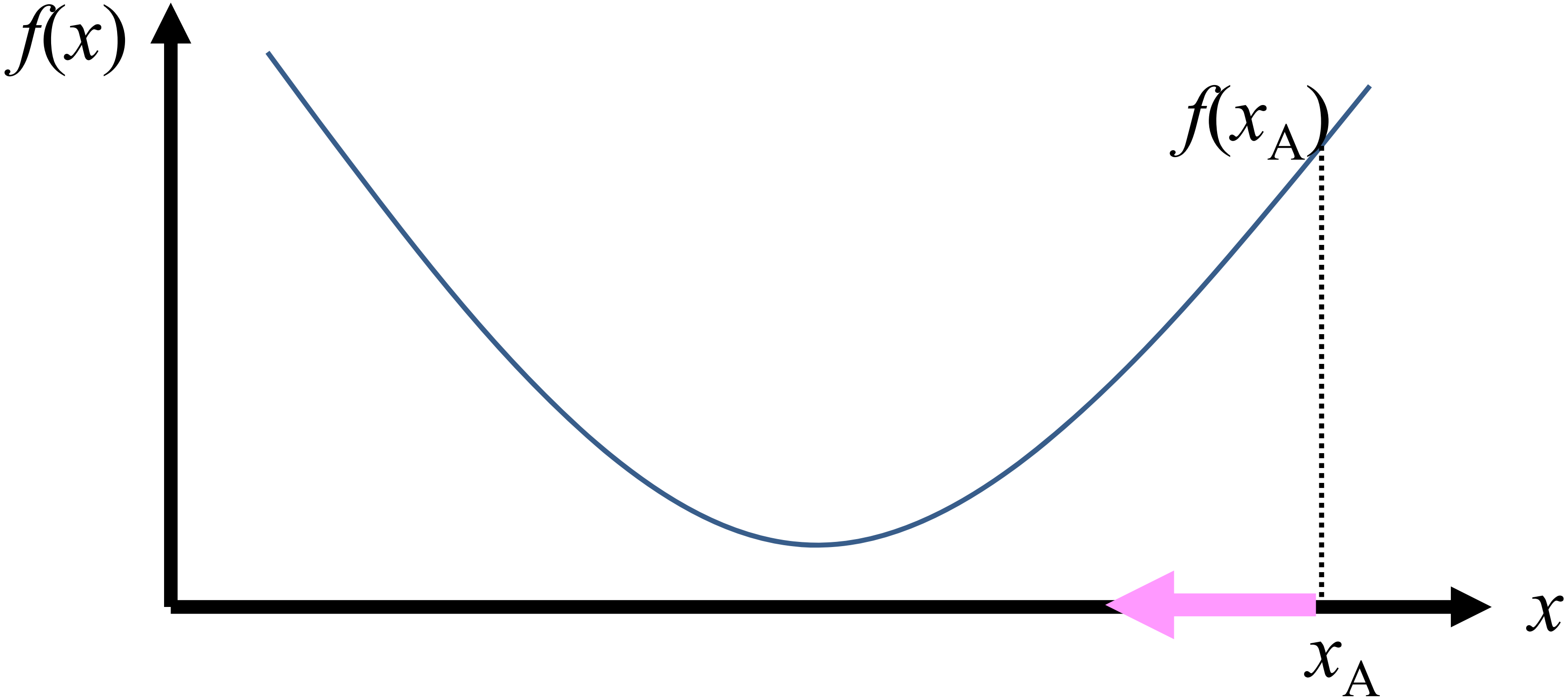


## **Gradient Descent** (Explanation)

Nonlinear minimization problem: Minimize $z = f(x)$.
Available information: derivative (numerical value) $df/dx$

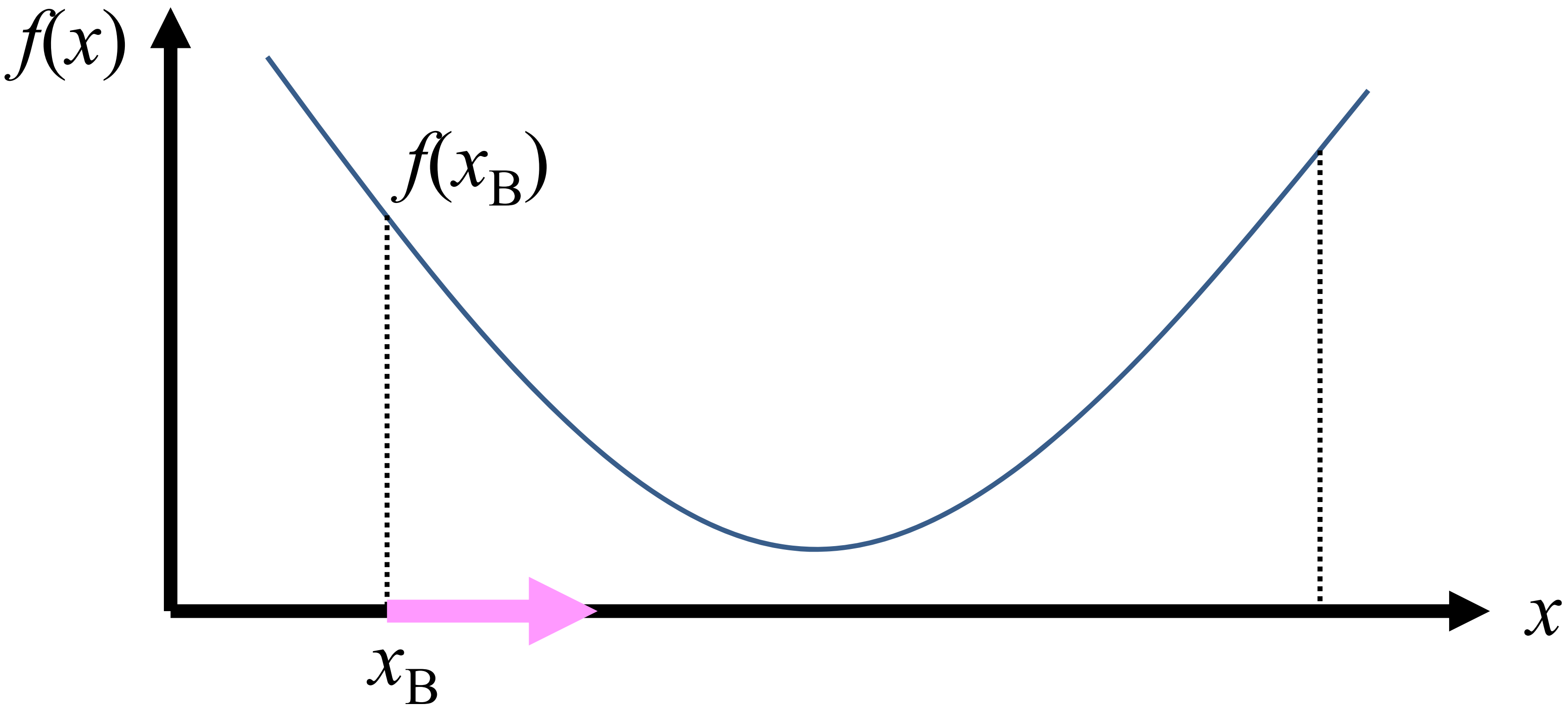
If the current solution is $x_A$, $df/dx > 0$. This means that $f(x)$ is increased by increasing $x$. Thus, we need to decrease $x$ in order to minimize $f(x)$



## **Gradient Descent** (Explanation)

Nonlinear minimization problem: Minimize $z = f(x)$.
Available information: derivative (numerical value) $df/dx$

If the current solution is $x_B$, $df/dx < 0$. This means that $f(x)$ is decreased by increasing $x$. Thus, we need to increase $x$ in order to minimize $f(x)$.



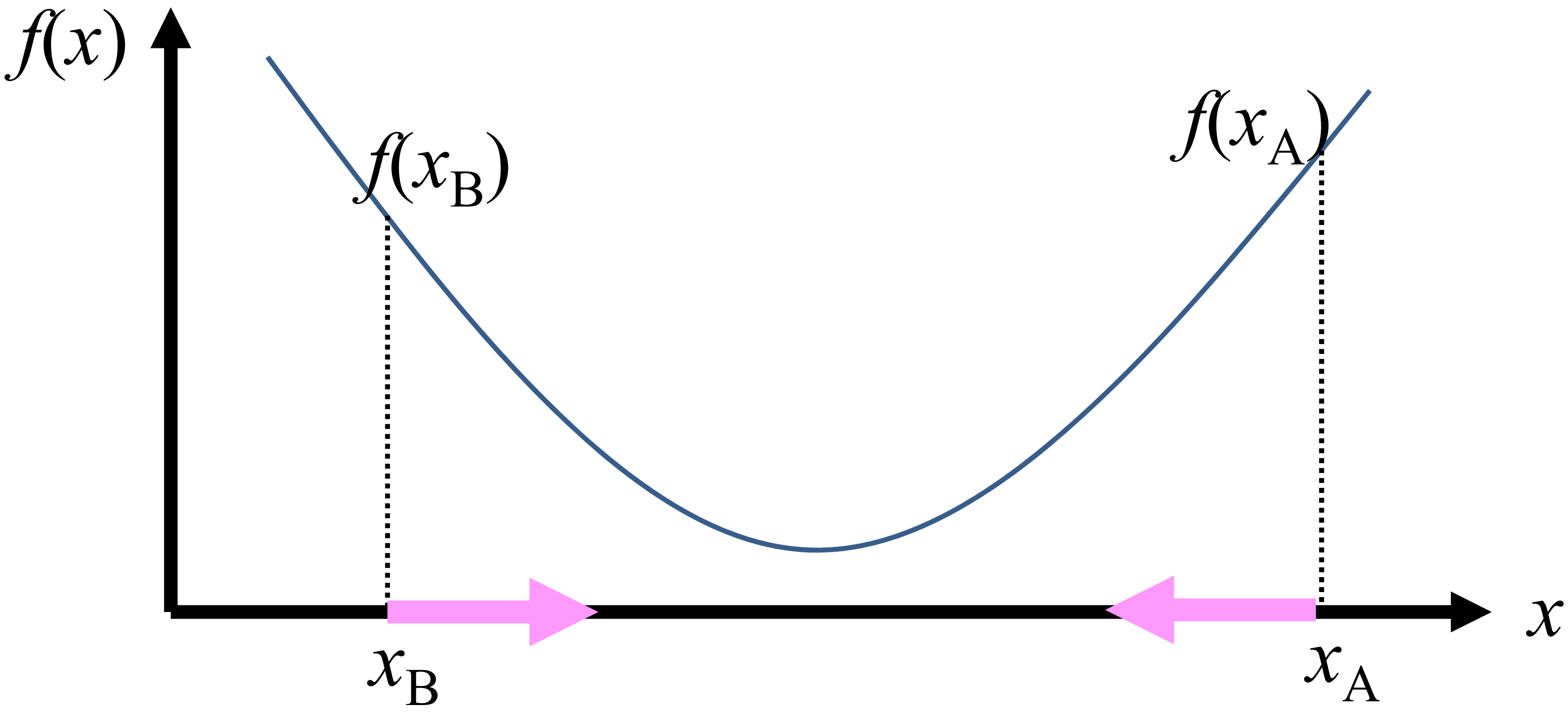
## **Gradient Descent** (Algorithm)

$$x^{(k+1)} = x^{(k)} - \alpha^{(k)} \frac{df}{dx}$$

$x^{(1)}$ : Initial value of $x$ (i.e., initial solution)

$x^{(k)}$: Value of $x$ at the $k$th iteration.

$\alpha^{(k)}$: Step size at the $k$th iteration (positive value).



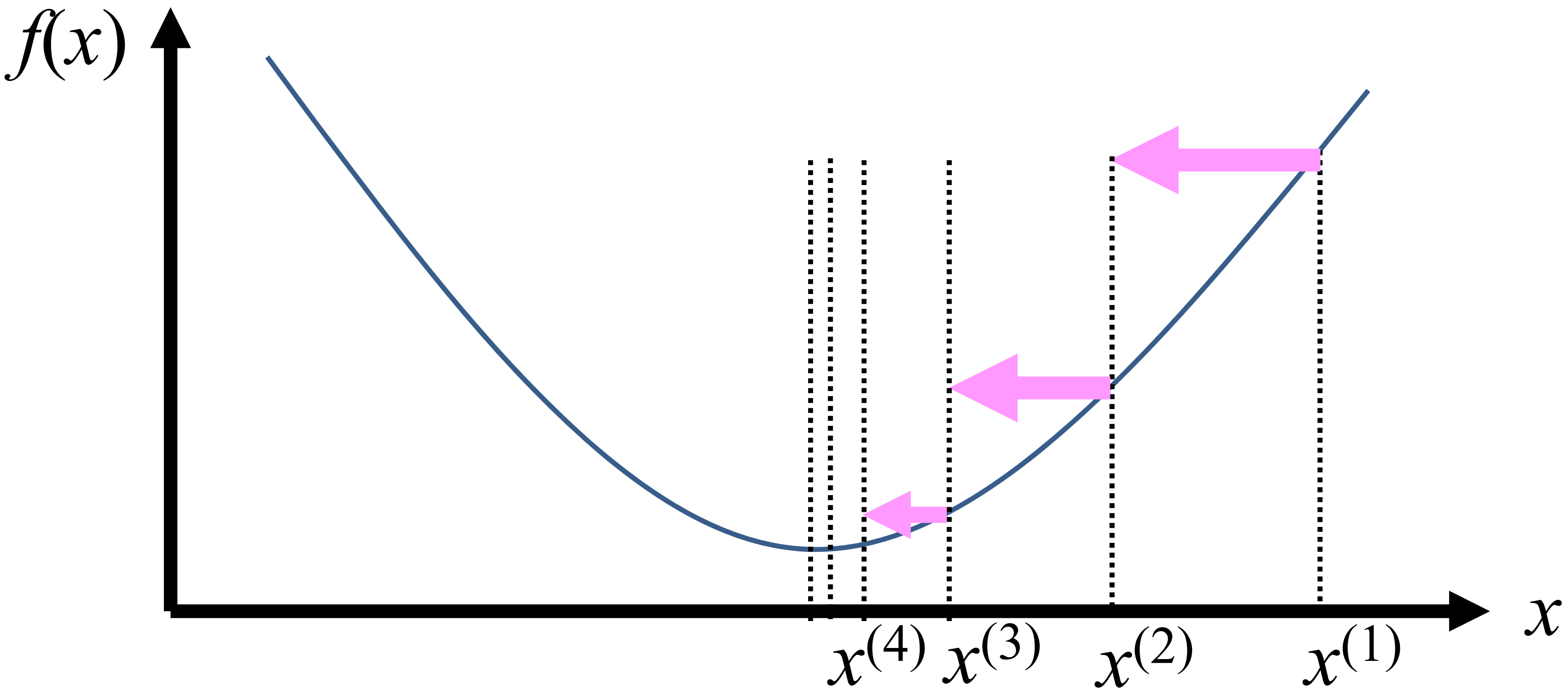
## **Gradient Descent** (Algorithm: Constant Step Size)

$$x^{(k+1)} = x^{(k)} - \alpha \frac{df}{dx}$$

$x^{(1)}$ : Initial value of $x$ (i.e., initial solution)

$x^{(k)}$: Value of $x$ at the $k$th iteration.

$\alpha$ : Step size (positive constant).



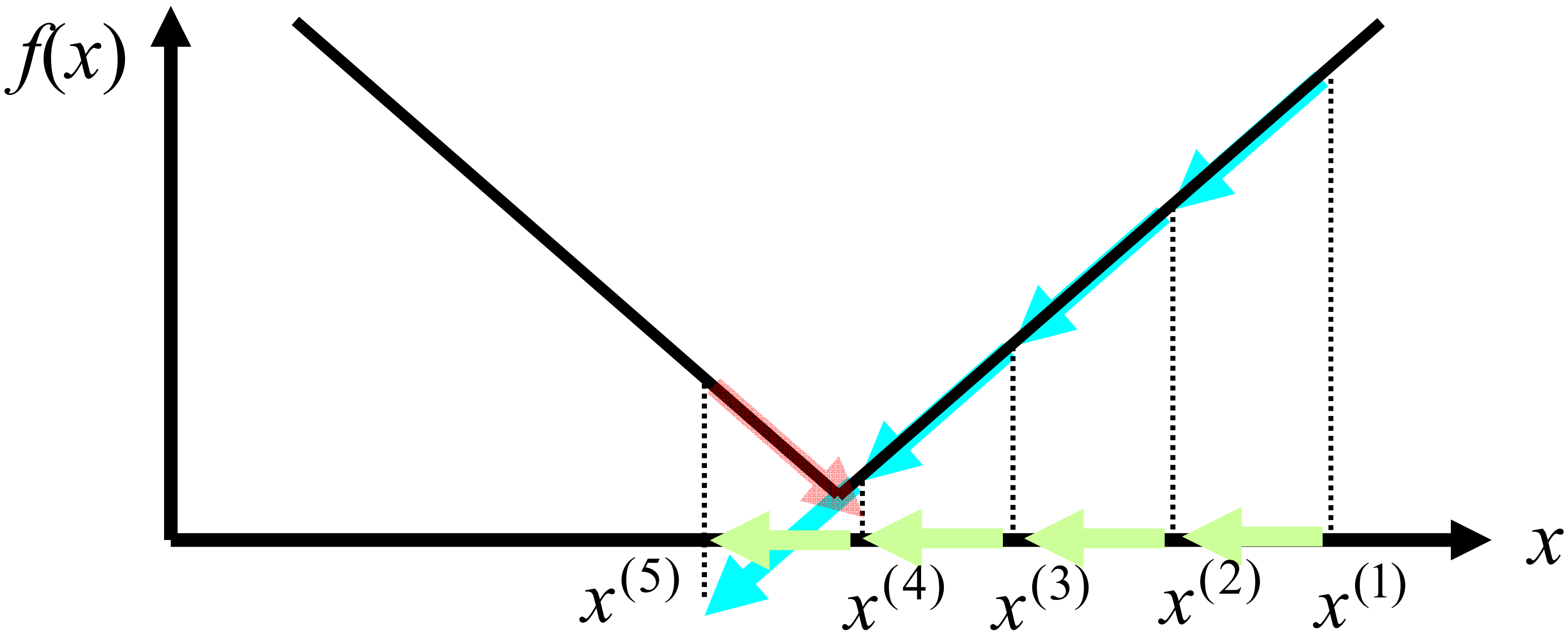
## **Gradient Descent** (Algorithm: Constant Step Size)

$$x^{(k+1)} = x^{(k)} - \alpha \frac{df}{dx}$$

$x^{(1)}$ : Initial value of $x$ (i.e., initial solution)

$x^{(k)}$: Value of $x$ at the $k$th iteration.

$\alpha$ : Step size (positive constant).

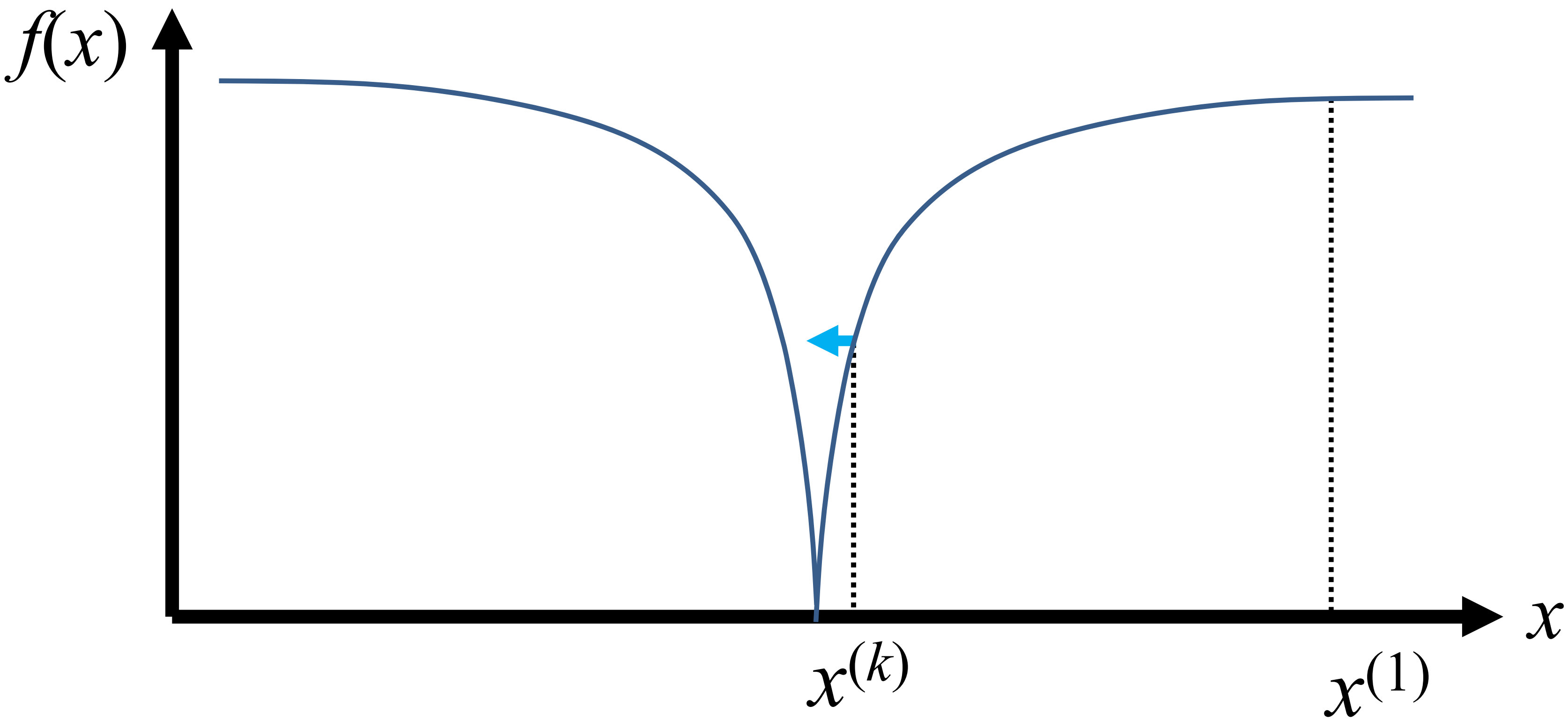



$x^{(5)} = x^{(7)} = x^{(9)} = \ldots$ $\qquad$ $x^{(4)} = x^{(6)} = x^{(8)} = \ldots$

## **Gradient Descent** (Algorithm: Constant Step Size)

$$x^{(k+1)} = x^{(k)} - \alpha \frac{df}{dx}$$

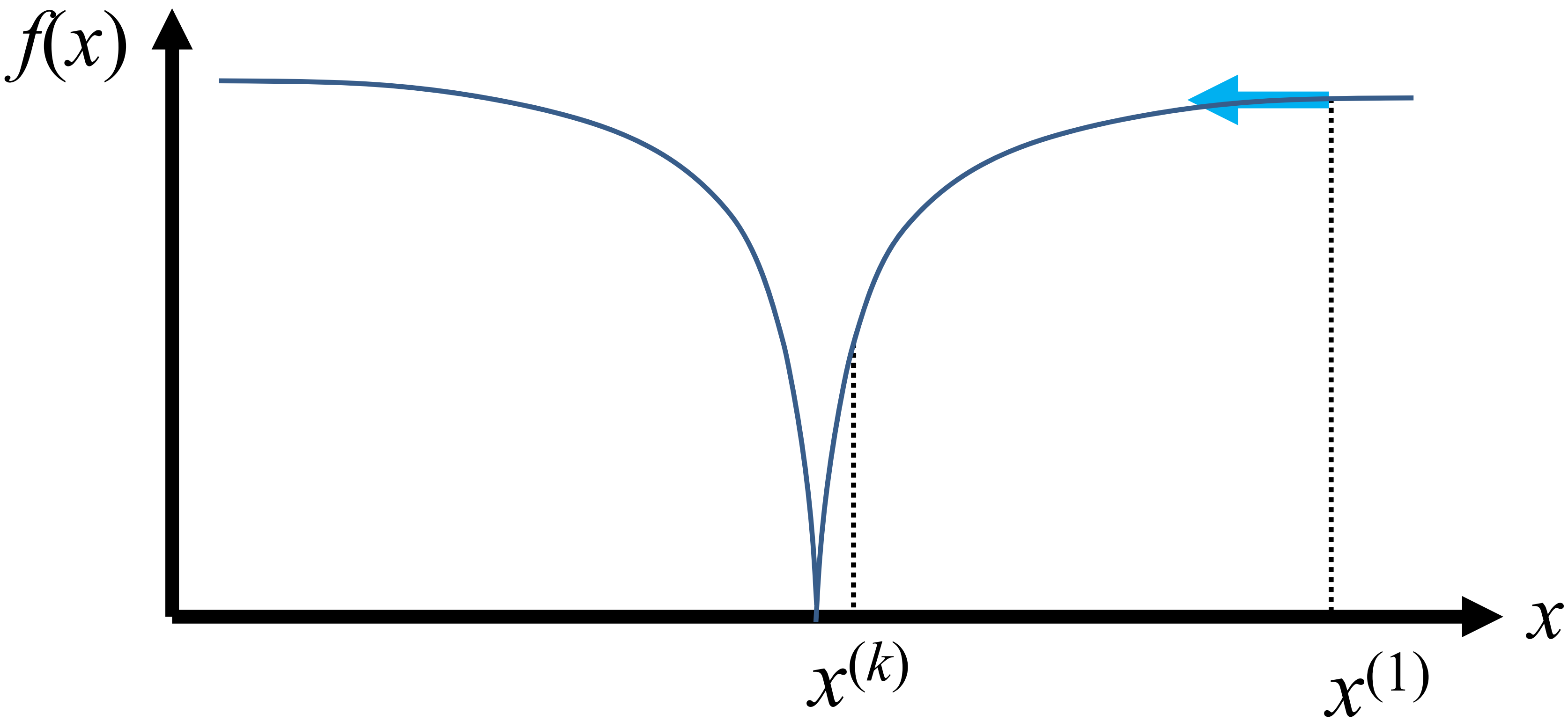
Explain a sequence of moves of a solution starting from $x^{(1)}$ when $\alpha$ is very small (which is appropriate for $x^{(k)}$): ______________________.



**Gradient Descent** (Algorithm: Constant Step Size)

$$x^{(k+1)} = x^{(k)} - \alpha \frac{df}{dx}$$

Explain a sequence of moves of a solution starting from $x^{(1)}$ when $\alpha$ is large (which is appropriate for $x^{(1)}$): ______________________.

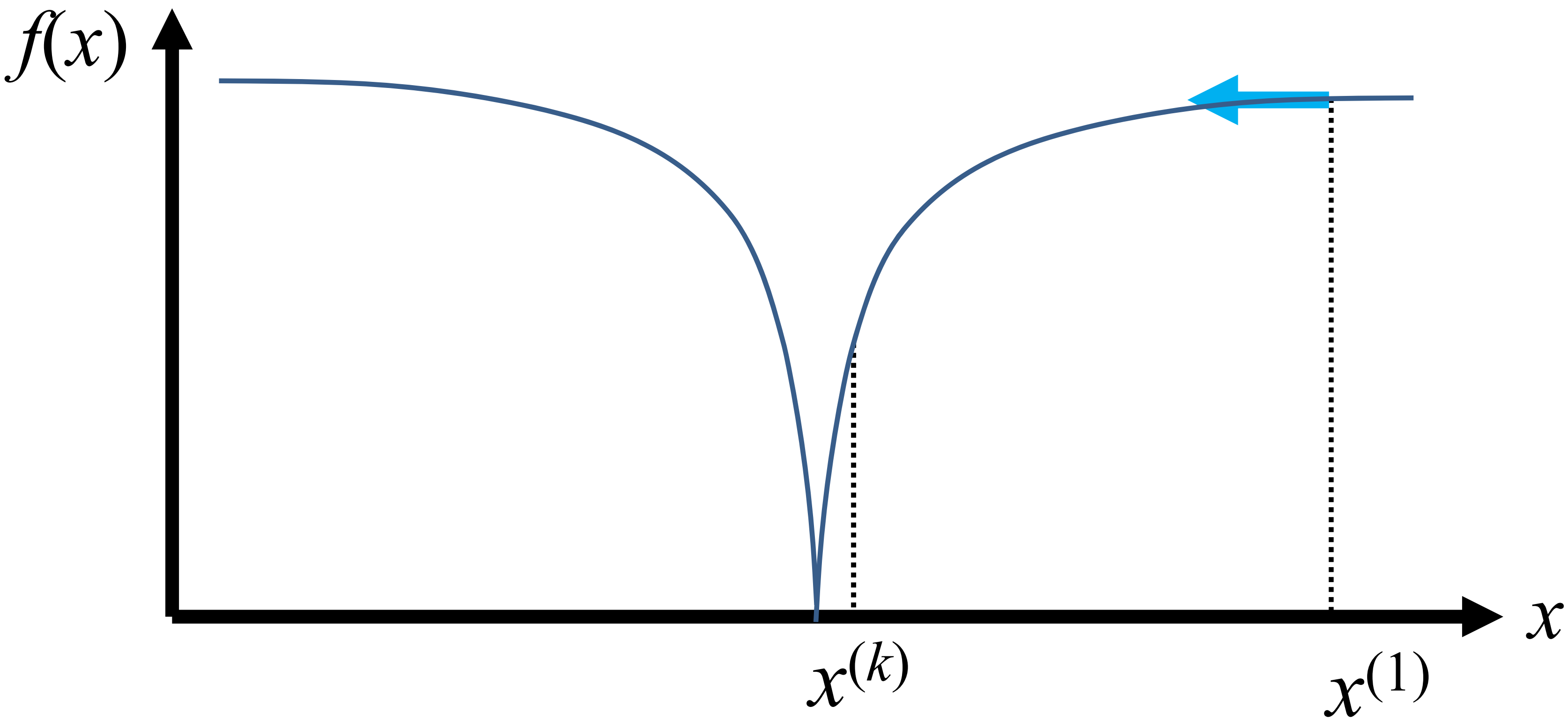


## **Gradient Descent** (Algorithm: Constant Step Size)

$$x^{(k+1)} = x^{(k)} - \alpha \frac{df}{dx} \qquad\qquad x^{(k+1)} = x^{(k)} - \alpha^{(k)} \frac{df}{dx}$$

**Discussion 1 (your idea):** How to specify the constant $\alpha$ ?

**Discussion 2 (your idea):** How to specify $\alpha^{(k)}$ for each move ?

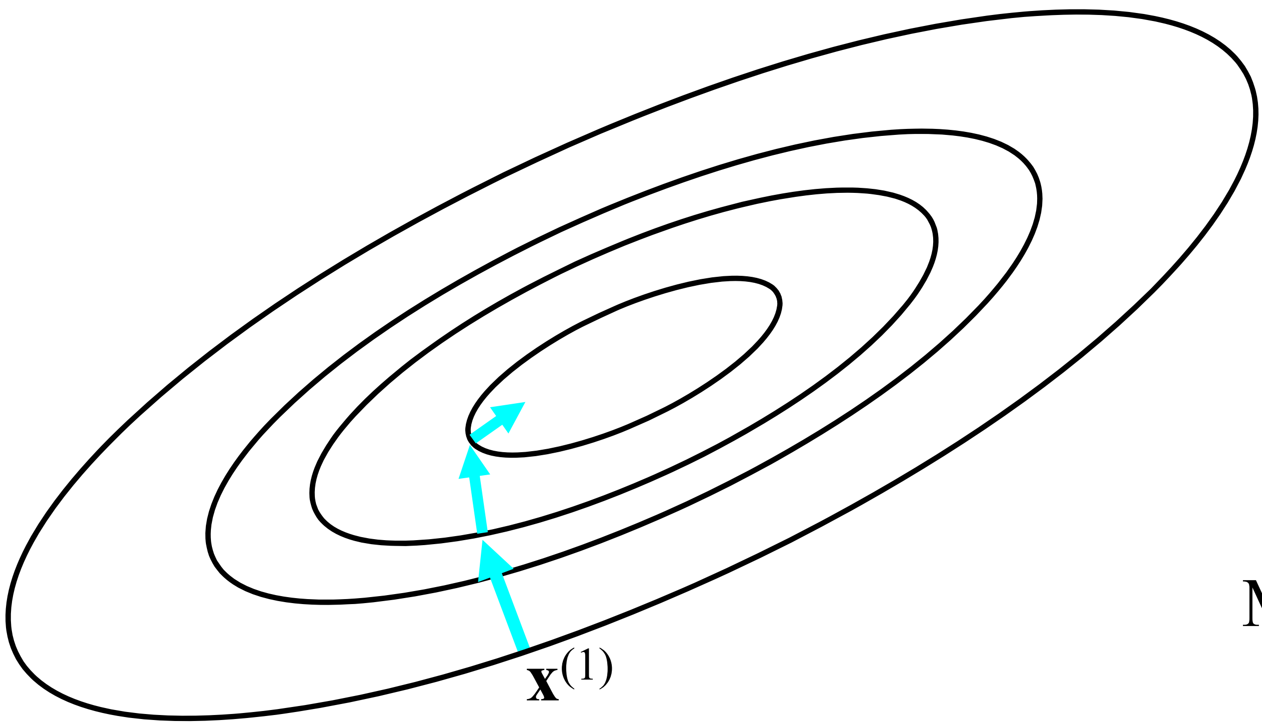


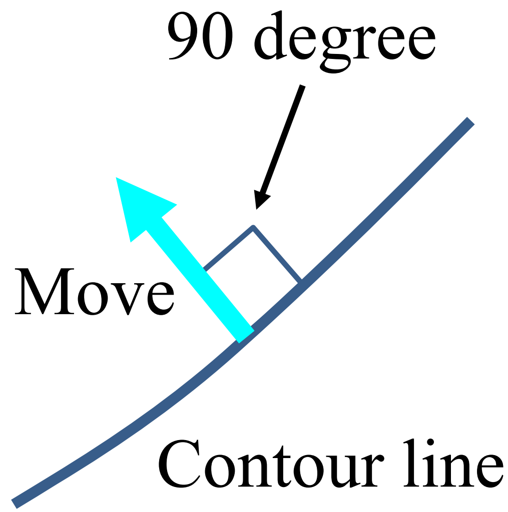
## **Gradient Descent** (Algorithm) (2D)

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} - \alpha^{(k)} \nabla f(\mathbf{x}^{(k)})$$

$$\nabla f(\mathbf{x}) = \begin{pmatrix} \dfrac{\partial f}{\partial x_1}(\mathbf{x}) \\ \dfrac{\partial f}{\partial x_2}(\mathbf{x}) \end{pmatrix}$$

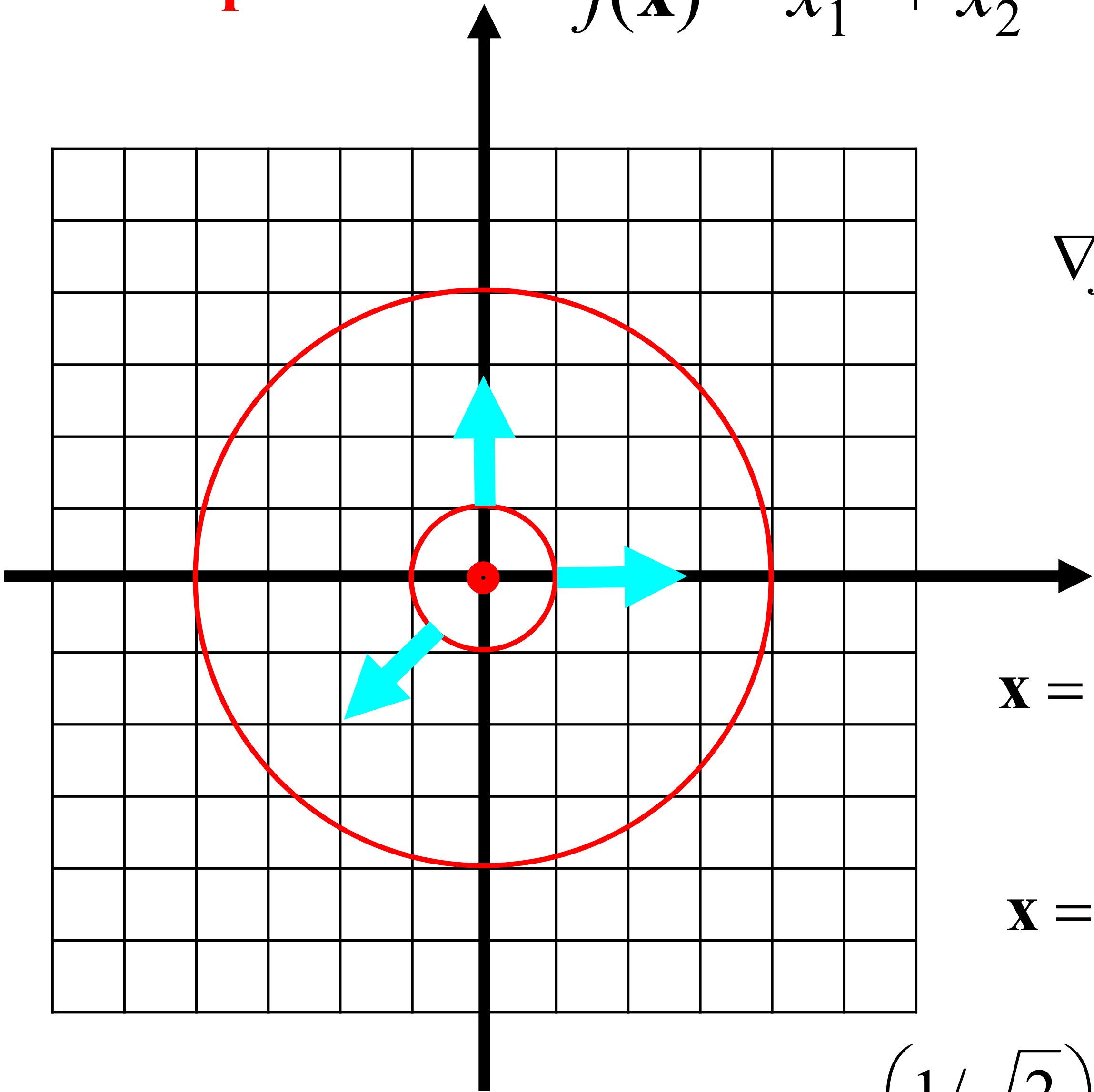
$\alpha^{(k)}$: Step size at the $k$th iteration (positive value).



**Example**

$$f(\mathbf{x}) = x_1{}^2 + x_2{}^2$$



$$\nabla f(\mathbf{x}) = \begin{pmatrix} \dfrac{\partial f}{\partial x_1}(\mathbf{x}) \\ \dfrac{\partial f}{\partial x_2}(\mathbf{x}) \end{pmatrix} = \begin{pmatrix} 2x_1 \\ 2x_2 \end{pmatrix}$$

$$\mathbf{x} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \Rightarrow \quad \nabla f(\mathbf{x}) = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$$

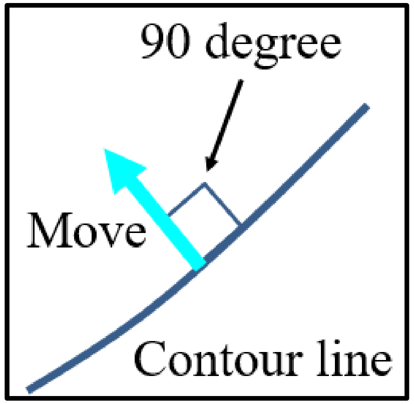
$$\mathbf{x} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad \Rightarrow \quad \nabla f(\mathbf{x}) = \begin{pmatrix} 0 \\ 2 \end{pmatrix}$$

$$\mathbf{x} = \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix} \quad \Rightarrow \quad \nabla f(\mathbf{x}) = \begin{pmatrix} 2/\sqrt{2} \\ 2/\sqrt{2} \end{pmatrix}$$

## **Gradient Descent** (Algorithm: Constant Step Size)

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} - \alpha \nabla f(\mathbf{x}^{(k)})$$

$\alpha$ : Step size (positive constant).



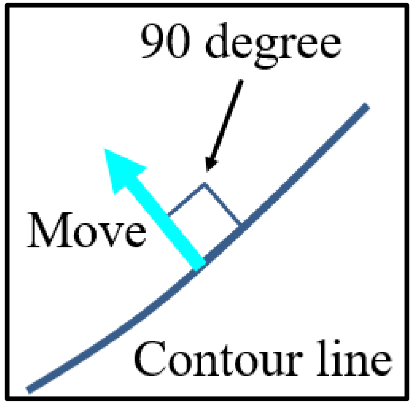

## Lab Session Task 1

Choose an initial solution in [-5, 5]x[-5, 5], specify the step size, and show a sequence of moves from the initial solution using the gradient decent algorithm with the constant step size for the following problem (see below). By iterating these steps using a different initial solution and a different step size, show several sequences of moves. Then, choose the most appropriate step size for this problem. That is, your task is to clearly show the search behavior and to find the best specification of the constant step size.

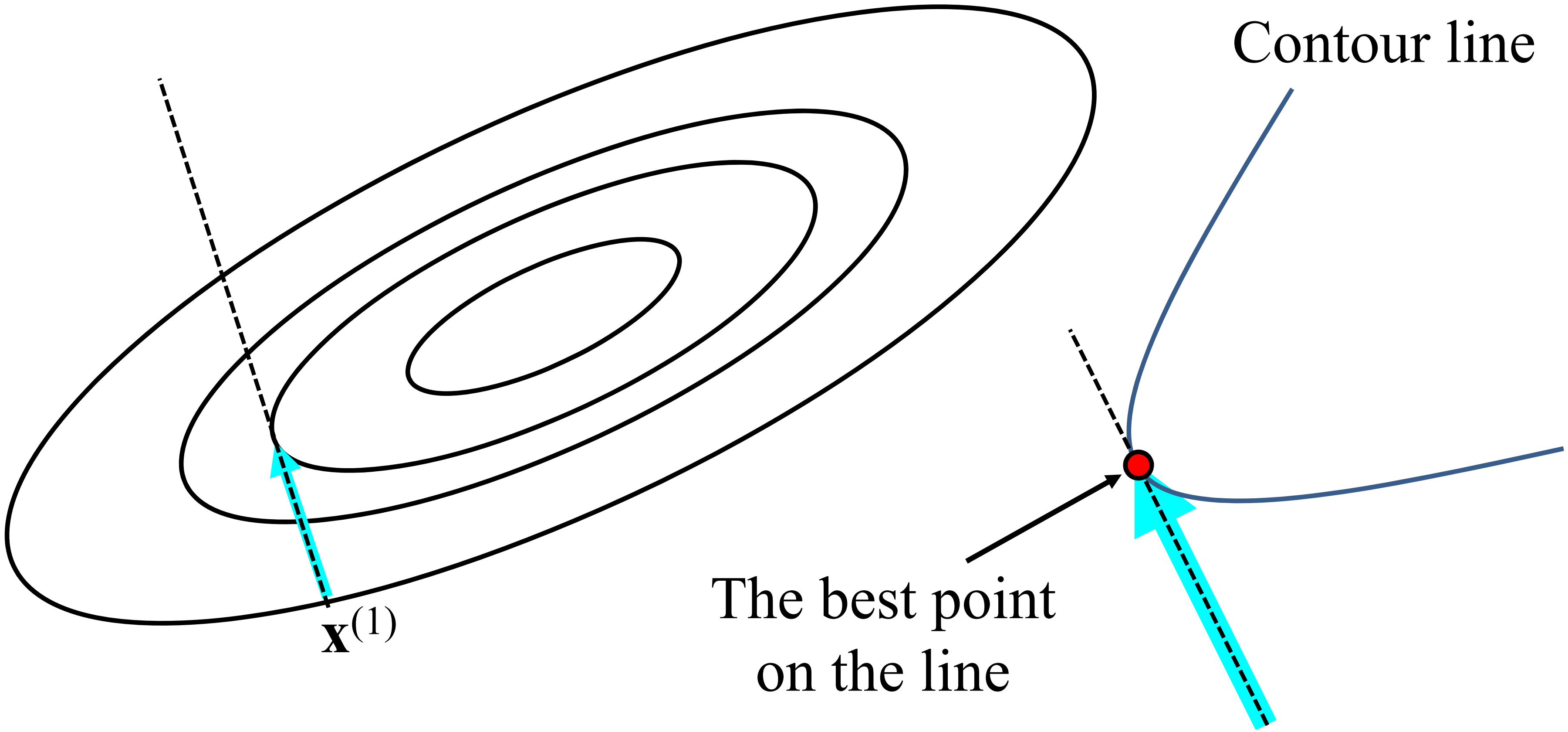
Minimize $f(\mathbf{x}) = (x_1 + 1)^2 / 9 + (x_2 + 1)^2$

## Steepest Descent

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} - \alpha^{(k)} \nabla f(\mathbf{x}^{(k)})$$



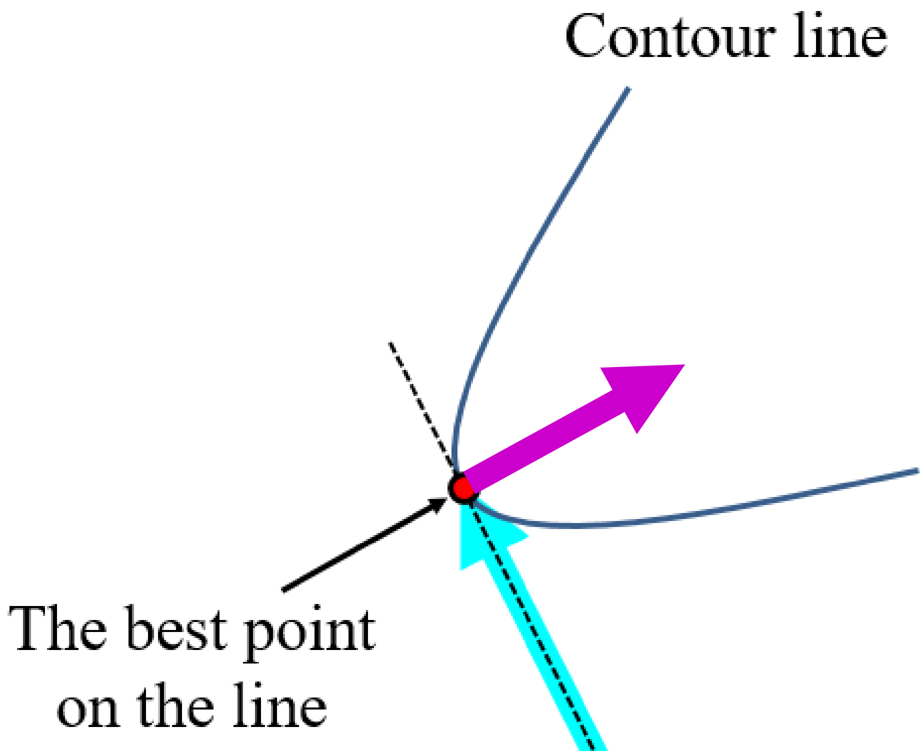

In each iteration, the size of the move is specified so that the objective function is minimized on the move direction line.

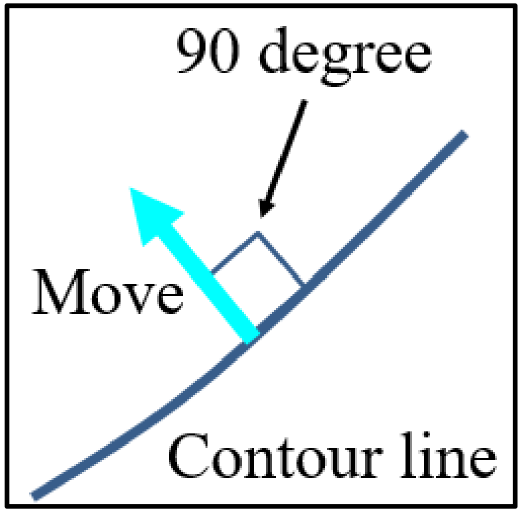


## Steepest Descent

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} - \alpha^{(k)} \nabla f(\mathbf{x}^{(k)})$$

**Feature:** the ($k$+1)-th move is orthogonal to the $k$-th move

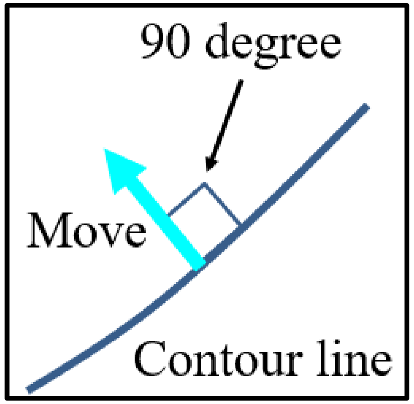


## **Gradient Descent** (Algorithm: Constant Step Size)

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} - \alpha \nabla f(\mathbf{x}^{(k)})$$

$\alpha$ : Step size (positive constant).




## Lab Session Task 1

Choose an initial solution in [-5, 5]x[-5, 5], specify the step size, and show a sequence of moves from the initial solution using the gradient decent algorithm with the constant step size for the following problem (see below). By iterating these steps using a different initial solution and a different step size, show several sequences of moves. Then, choose the most appropriate step size for this problem. That is, your task is to clearly show the search behavior and to find the best specification of the constant step size.

Minimize $f(\mathbf{x}) = (x_1 + 1)^2 / 9 + (x_2 + 1)^2$