Naval Architecture & Ocean Engineering

Computer Aided Ship Design

Part I. Optimization Method

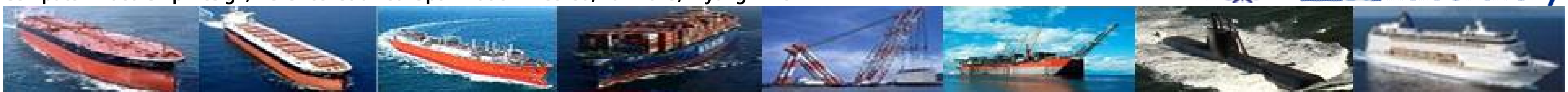
Ch. 3 Unconstrained Optimization Method

September, 2013

Prof. Myung-Il Roh

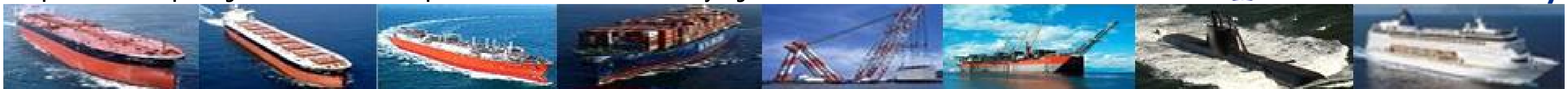
Department of Naval Architecture and Ocean Engineering,
Seoul National University of College of Engineering

SYstem Design Laboratory



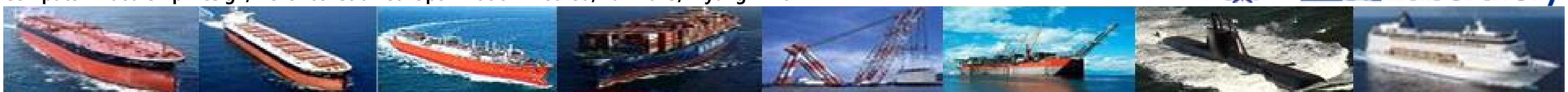
Ch. 3 Unconstrained Optimization Method

SYstem Design Laboratory



3.1 Gradient Method

1. Steepest Descent Method
2. Conjugate Gradient Method
3. Newton's Method
4. Davidon-Fletcher-Powell(DFP) Method
5. Broyden-Fletcher-Goldfarb-Shanno(BFGS) Method



3.1 Gradient Method

1. Steepest Descent Method (1/6)

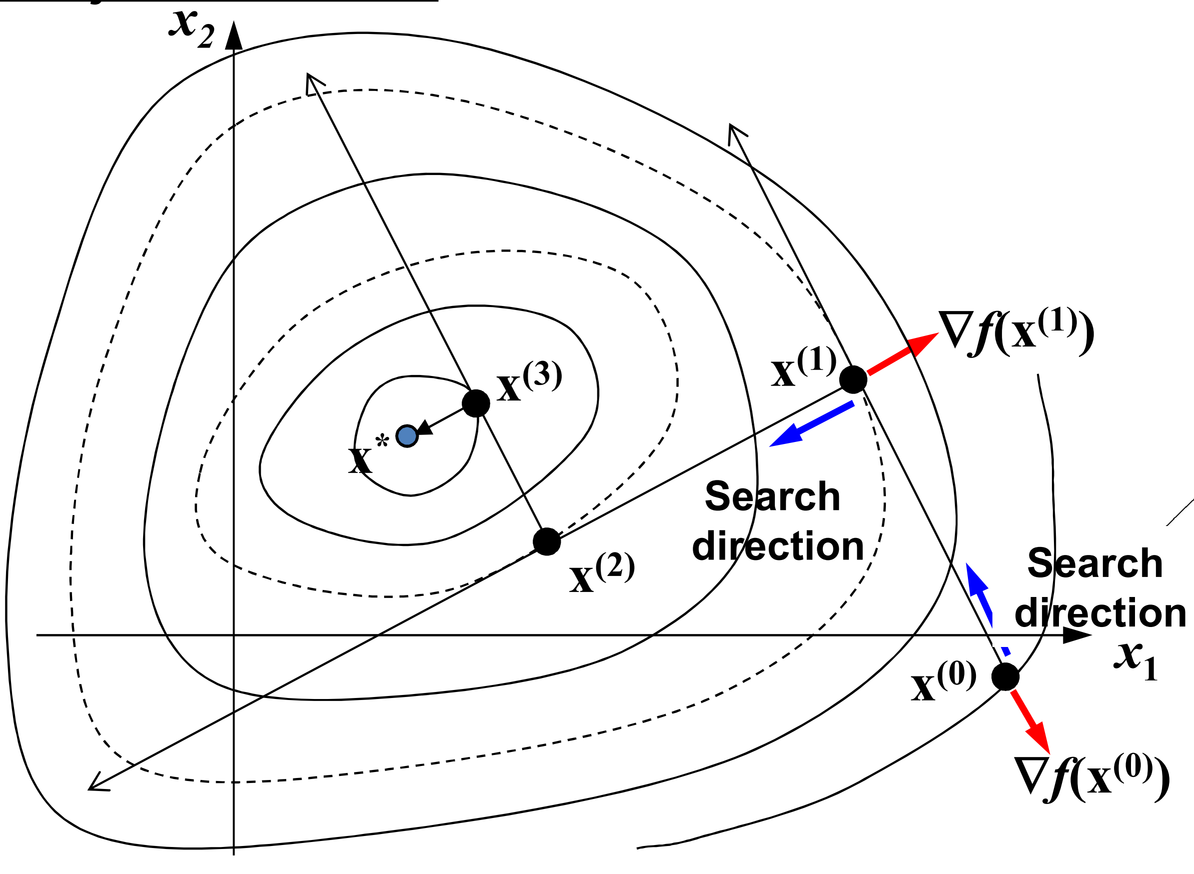
- Step 1: The search direction($\mathbf{d}$) is taken as the negative of the gradient of the objective function(f) at current iteration since the objective function decrease mostly rapidly.
- The direction of gradient vector of f , $\nabla f(\mathbf{x})$, is the direction of maximum increase of f at $\mathbf{x}$

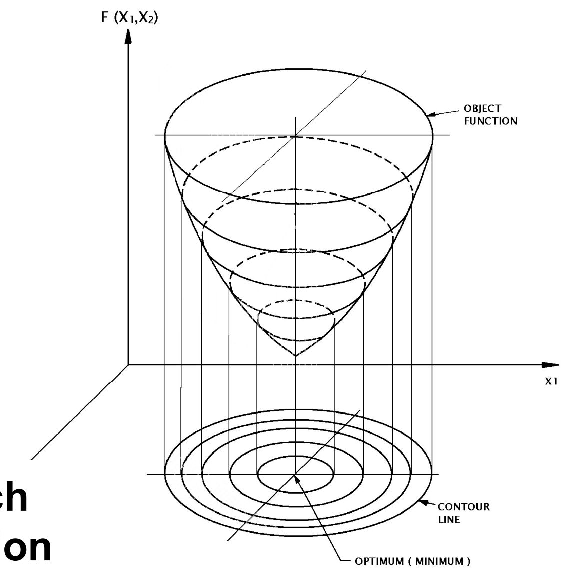
Search direction $$\mathbf{d} = -\mathbf{c} \equiv -\nabla f(\mathbf{x})$$

Ref) Appendix A.1: Directional Derivative & Gradient Vector

- Step 2: Iterate successively to find the optimum design point.

Ex) Minimize the objective function





3.1 Gradient Method
1. Steepest Descent Method (2/6): Example

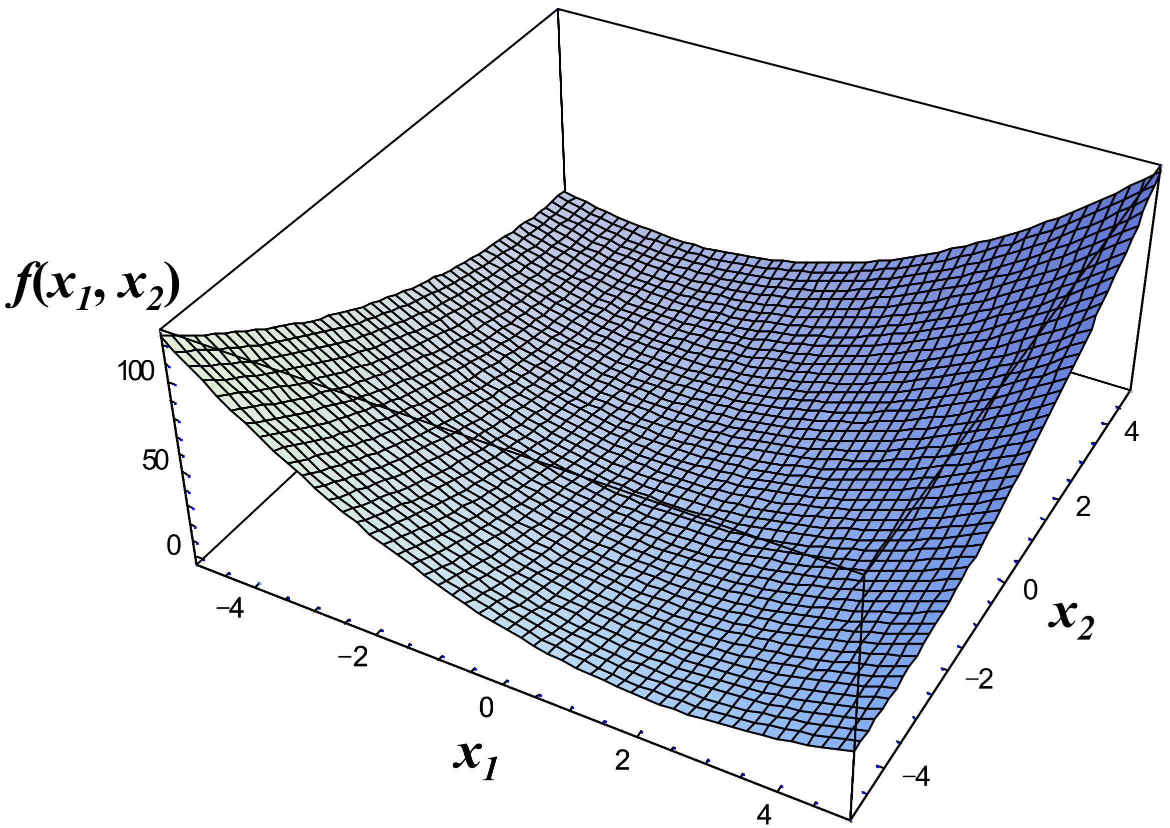
☑ **By using the steepest descent method, find the minimum design point for the following function of 2-variables.**

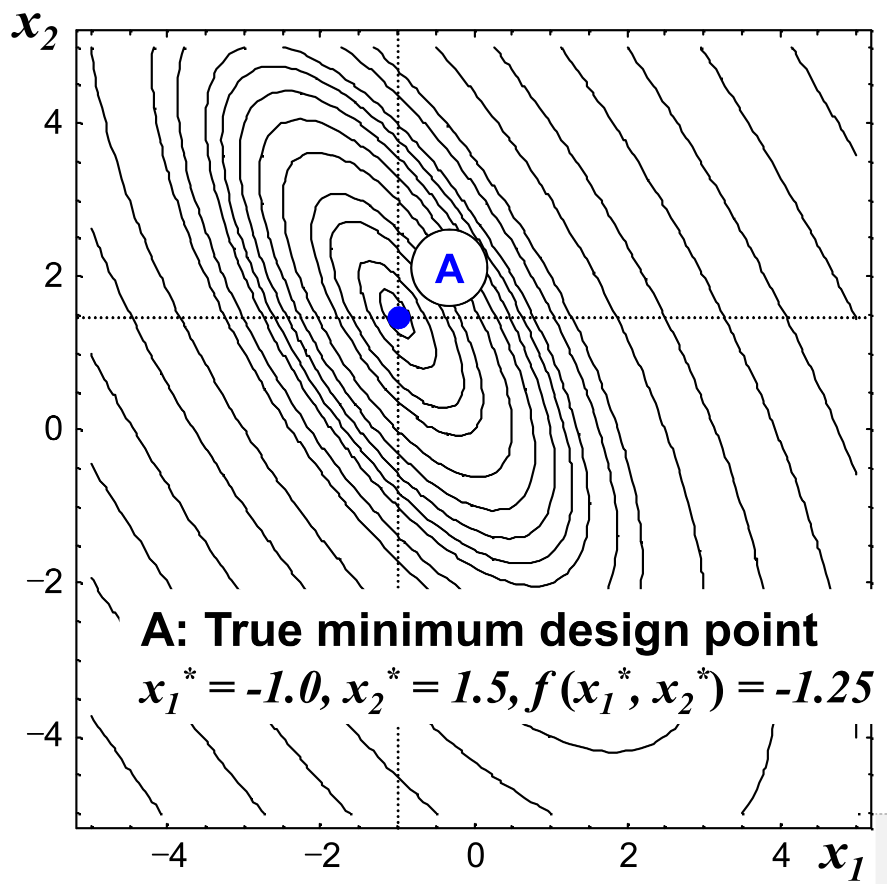
Given: Starting design point $\mathbf{x}^{(0)}$ = (0, 0), convergence tolerance ε = 0.001

Find: $\mathbf{x}^{(1)}$, $\mathbf{x}^{(2)}$

Minimize $f(x_1, x_2) = x_1 - x_2 + 2x_1^2 + 2x_1x_2 + x_2^2$

➡ **Optimization problem with two unknown variables**





3.1 Gradient Method
1. Steepest Descent Method (3/6): Example

Minimize $f(x_1, x_2) = x_1 - x_2 + 2x_1^2 + 2x_1x_2 + x_2^2$ **Starting design point** $\mathbf{x}^{(0)} = (0, 0)$

$$\nabla f(\mathbf{x}) = \nabla f(x_1, x_2) = \begin{pmatrix} 1 + 4x_1 + 2x_2 \\ -1 + 2x_1 + 2x_2 \end{pmatrix}$$

■ 1st Iteration: Find $\mathbf{x}^{(1)}$

$$\nabla f(\mathbf{x}^{(0)}) = \nabla f\begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 + 4x_1 + 2x_2 \\ -1 + 2x_1 + 2x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

$$\mathbf{x}^{(1)} = \mathbf{x}^{(0)} - \alpha^{(0)} \nabla f(\mathbf{x}^{(0)})$$

$$= \begin{pmatrix} 0 \\ 0 \end{pmatrix} - \alpha \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} -\alpha \\ \alpha \end{pmatrix}$$ **Replacing $\alpha^{(0)}$ to α for convenience**

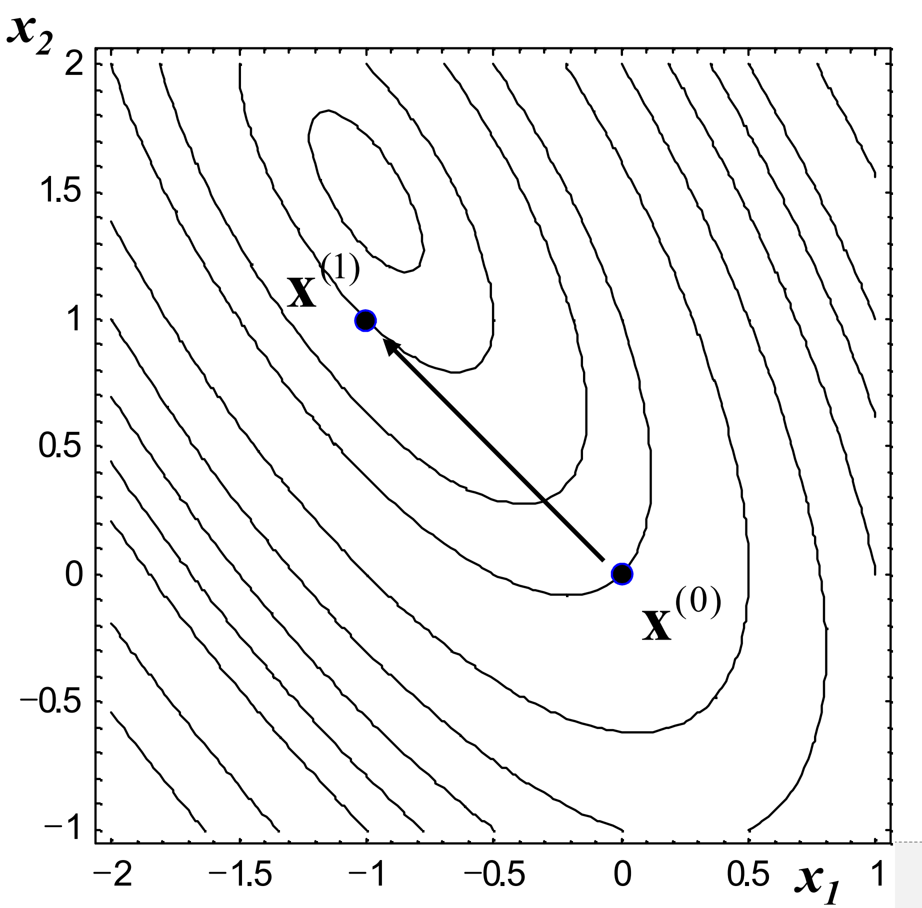
Substituting $\mathbf{x}^{(1)} = (-\alpha, \alpha)$ into the objective function

$$f(\mathbf{x}^{(1)}) = -\alpha - \alpha + 2\alpha^2 - 2\alpha^2 + \alpha^2$$
$$= \alpha^2 - 2\alpha$$

To minimize $f(\mathbf{x}^{(1)})$,

$$\boxed{\frac{df(\mathbf{x}^{(1)})}{d\alpha}} = 2\alpha - 2 = 0 \rightarrow \alpha = 1.0 \qquad \therefore \mathbf{x}^{(1)} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

How can we differentiate f with respect to α?



3.1 Gradient Method
1. Steepest Descent Method (4/6): Example

Minimize $f(x_1, x_2) = x_1 - x_2 + 2x_1^2 + 2x_1x_2 + x_2^2$ **Starting design point $\mathbf{x}^{(0)} = (0, 0)$**

■ 2nd Iteration: Find $\mathbf{x}^{(2)}$

$$\nabla f(\mathbf{x}^{(1)}) = \nabla f\begin{pmatrix} -1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 + 4x_1 + 2x_2 \\ -1 + 2x_1 + 2x_2 \end{pmatrix} = \begin{pmatrix} -1 \\ -1 \end{pmatrix}$$

$$\mathbf{x}^{(2)} = \mathbf{x}^{(1)} - \alpha^{(1)} \nabla f(\mathbf{x}^{(1)})$$

$$= \begin{pmatrix} -1 \\ 1 \end{pmatrix} - \alpha \begin{pmatrix} -1 \\ -1 \end{pmatrix} = \begin{pmatrix} -1 + \alpha \\ 1 + \alpha \end{pmatrix}$$

Replacing $\alpha^{(1)}$ to α for convenience

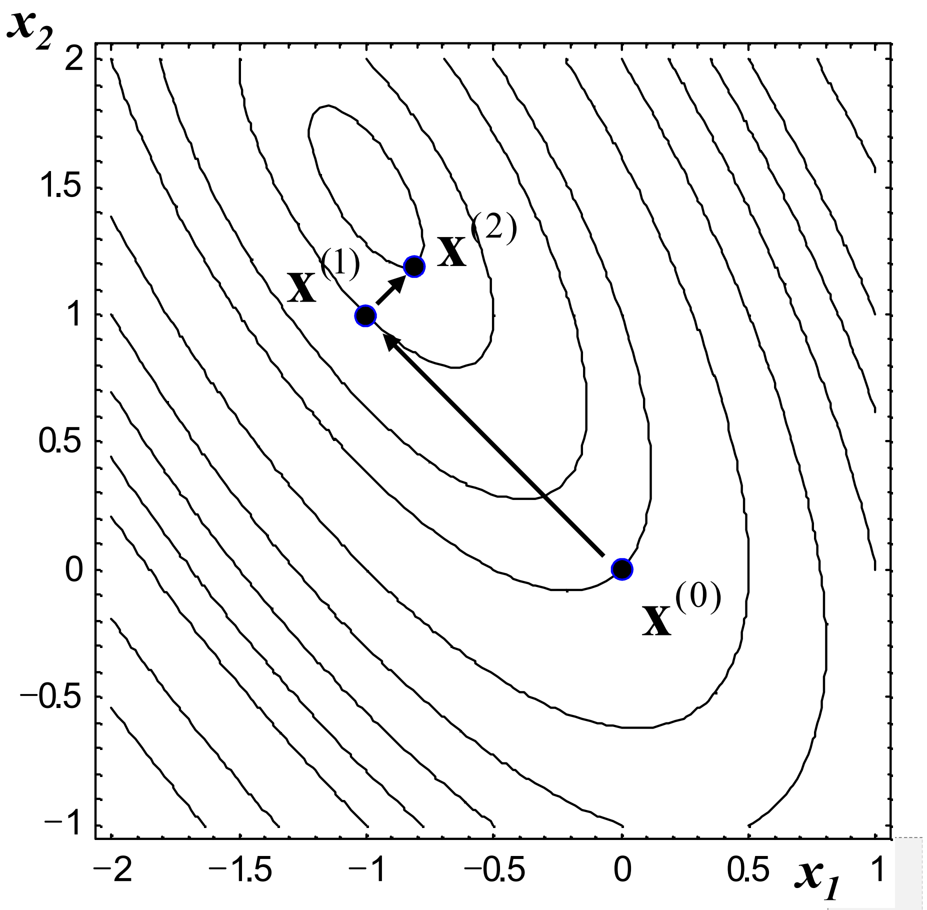
Substituting $\mathbf{x}^{(2)} = (-1 + \alpha, 1 + \alpha)$ **into the objective function**

$$f(\mathbf{x}^{(2)}) = 5\alpha^2 - 2\alpha - 1$$

To minimize $f(\mathbf{x}^{(2)})$**,**

$$\frac{df(\mathbf{x}^{(2)})}{d\alpha} = 10\alpha - 2 = 0 \quad \rightarrow \quad \alpha = 0.2$$

$$\therefore \mathbf{x}^{(2)} = \begin{pmatrix} -0.8 \\ 1.2 \end{pmatrix}$$



3.1 Gradient Method
1. Steepest Descent Method (5/6): Example

Minimize $f(x_1, x_2) = x_1 - x_2 + 2x_1^2 + 2x_1x_2 + x_2^2$ **Starting design point $\mathbf{x}^{(0)} = (0, 0)$**

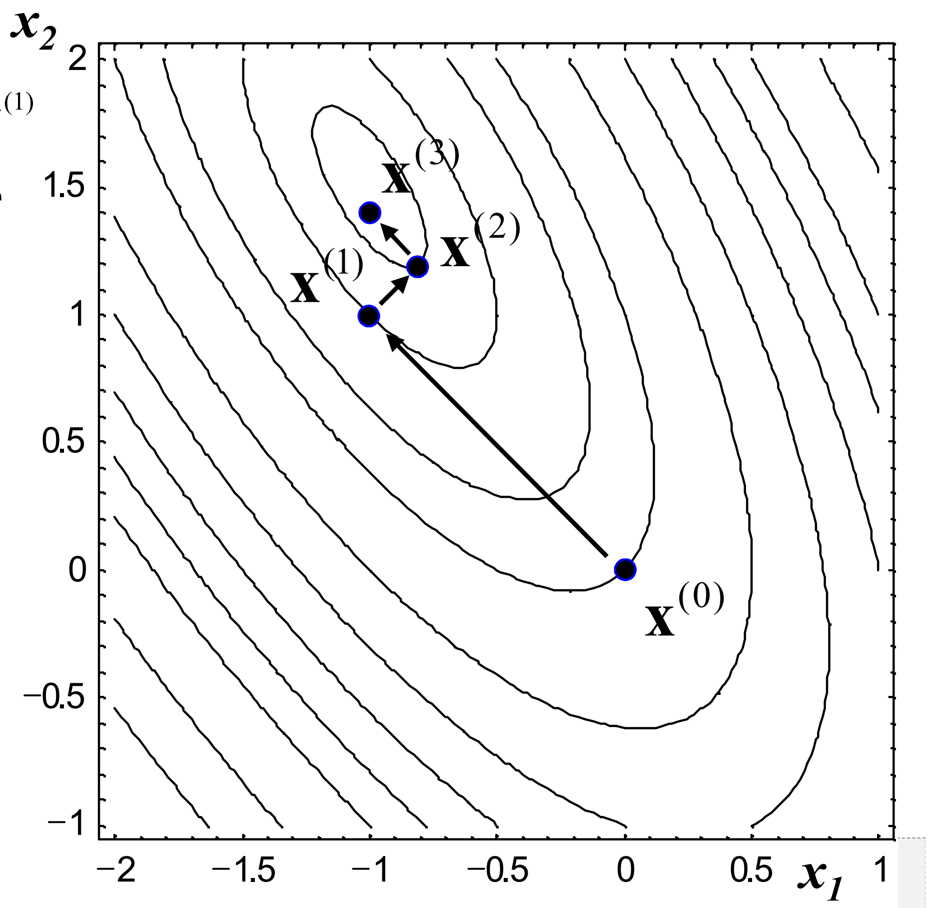
■ 3rd Iteration: Find $\mathbf{x}^{(3)}$

$$\nabla f(\mathbf{x}^{(2)}) = \nabla f\begin{pmatrix} -0.8 \\ 1.2 \end{pmatrix} = \begin{pmatrix} 1 + 4x_1 + 2x_2 \\ -1 + 2x_1 + 2x_2 \end{pmatrix} = \begin{pmatrix} 0.2 \\ -0.2 \end{pmatrix}$$

$$\mathbf{x}^{(3)} = \mathbf{x}^{(2)} - \alpha^{(2)} \nabla f(\mathbf{x}^{(2)})$$

$$= \begin{pmatrix} -0.8 \\ 1.2 \end{pmatrix} - \alpha \begin{pmatrix} 0.2 \\ -0.2 \end{pmatrix} = \begin{pmatrix} -0.8 - 0.2\alpha \\ 1.2 + 0.2\alpha \end{pmatrix}$$

Replacing $\alpha^{(1)}$ to α for convenience



Substituting $\mathbf{x}^{(3)} = (-0.8 - 0.2\alpha, 1.2 + 0.2\alpha)$ **into the objective function**

$$f(\mathbf{x}^{(3)}) = 0.04\alpha^2 - 0.08\alpha - 1.2$$

To minimize $f(\mathbf{x}^{(3)})$,

$$\frac{df(\mathbf{x}^{(3)})}{d\alpha} = 0.08\alpha - 0.08 = 0 \quad \rightarrow \quad \alpha = 1.0$$

$$\therefore \mathbf{x}^{(3)} = \begin{pmatrix} -1 \\ 1.4 \end{pmatrix}$$

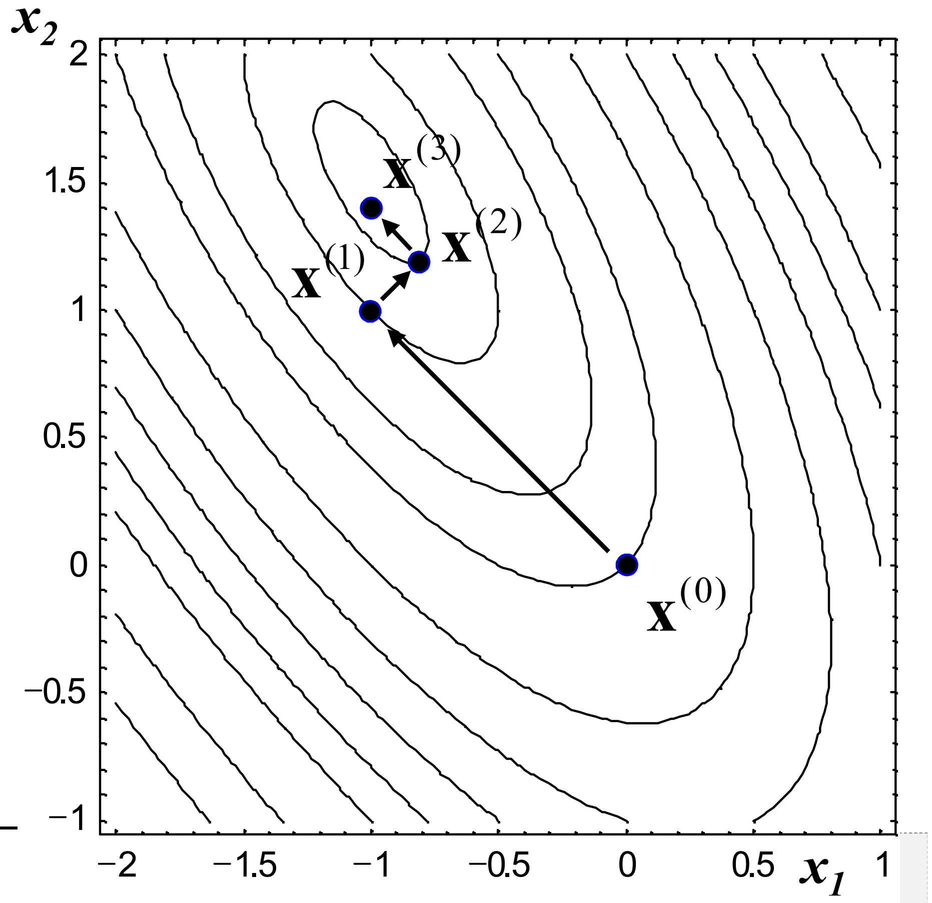
3.1 Gradient Method
1. Steepest Descent Method (6/6): Example

Minimize $f(x_1, x_2) = x_1 - x_2 + 2x_1^2 + 2x_1x_2 + x_2^2$ **Starting design point** $\mathbf{x}^{(0)} = (0, 0)$

- **4^{th} Iteration: Find the minimum design point.**

To obtain the minimum design point, we have to iterate.

If $\left|\mathrm{x}^{(k+1)} - \mathrm{x}^{(k)}\right| \le \varepsilon$**, then stop the iterative process because** $\mathbf{x}^{(k+1)}$ **can be assumed as the minimum design point.**



[Reference] Differentiation of Function of x with Respect to the Another Variable

Minimize $f(x_1, x_2) = x_1 - x_2 + 2x_1^2 + 2x_1x_2 + x_2^2$ **Starting design point** $\mathbf{x}^{(0)} = (0, 0)$

Substituting $\mathbf{x}^{(1)} = (-\alpha, \alpha)$ **into the objective function**

$$f(\mathbf{x}^{(1)}) = -\alpha - \alpha + 2\alpha^2 - 2\alpha^2 + \alpha^2$$
$$= \alpha^2 - 2\alpha$$

To minimize $f(\mathbf{x}^{(1)})$**,**

$$\frac{df(\mathbf{x}^{(1)})}{d\alpha} = 2\alpha - 2 = 0 \rightarrow \alpha = 1.0 \qquad \therefore \mathbf{x}^{(1)} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

How can we differentiate f **with respect to** α**?**

$f(x_1, x_2) = f(\mathbf{x})$ **:** f **is the function of** $\mathbf{x}$.

$\mathbf{x}^{(1)} = (-\alpha, \alpha)$ **:** $\mathbf{x}^{(1)}$ **is the function of** α

➡ **Substituting** $\mathbf{x}^{(1)}$ **into** f**,** f **is , then, a function of** α **and can be differentiated with respect to** α **.**

In the similar way, we can consider the followings:

To minimize $f(\mathbf{x}^* + \Delta\mathbf{x})$**,**

The second-order Taylor series expansion of $f(\mathbf{x}^* + \Delta\mathbf{x})$

$$f(\mathbf{x}^* + \Delta\mathbf{x}) = f(\mathbf{x}^*) + \mathbf{c}^T\Delta\mathbf{x} + \frac{1}{2}\Delta\mathbf{x}^T\mathbf{H}(\mathbf{x}^*)\,\Delta\mathbf{x}$$
$$f(\mathbf{x}^* + \Delta\mathbf{x}) - f(\mathbf{x}^*) = \mathbf{c}^T\Delta\mathbf{x} + \frac{1}{2}\Delta\mathbf{x}^T\mathbf{H}(\mathbf{x}^*)\,\Delta\mathbf{x}$$

In the above equation, we assume that
x* is a constant and $\Delta\mathbf{x}$ **is a variable.**

$$f(\Delta\mathbf{x}) = \mathbf{c}^T\Delta\mathbf{x} + \frac{1}{2}\Delta\mathbf{x}^T\mathbf{H}(\mathbf{x}^*)\,\Delta\mathbf{x}$$

To minimize f**,**

$$\frac{df(\Delta\mathbf{x})}{d\Delta\mathbf{x}} = \mathbf{c} + \mathbf{H}(\mathbf{x}^*)\,\Delta\mathbf{x} = 0$$
$$\Rightarrow \mathbf{H}(\mathbf{x}^*)\,\Delta\mathbf{x} = -\mathbf{c}$$
$$\Rightarrow \boxed{\Delta\mathbf{x} = -\mathbf{H}(\mathbf{x}^*)^{-1}\mathbf{c}}$$

'Newton's method'

2. Conjugate Gradient Method (1/5)

☑ This method requires only a simple modification to the steepest descent method and dramatically improves the convergence rate of the optimization process.

☑ The current steepest descent direction is modified by adding a scaled direction used in the previous iteration.

- Step 1 : Estimate a starting design point as $\mathbf{x}^{(0)}$. Set the iteration counter $k = 0$. Also, specify a tolerance ε for stopping criterion. Calculate

$$\mathbf{d}^{(0)} = -\mathbf{c}^{(0)} \equiv -\nabla f(\mathbf{x}^{(0)})$$

Check stopping criterion. If $\left\|\mathbf{c}^{(0)}\right\| < \varepsilon$, then stop. Otherwise, go to Step 4.

It is noted that Step 1 of the conjugate gradient method and steepest descent method is the same.

2. Conjugate Gradient Method (2/5)

- **Step 2 : Compute the gradient of the objective function as** $\mathbf{c}^{(k)} = \nabla f(\mathbf{x}^{(k)})$.
  If $\left\|\mathbf{c}^{(k)}\right\| < \varepsilon$ **, then stop; otherwise continue.**

- **Step 3 : Calculate the new search direction as**

$$\boxed{\mathbf{d}^{(k)} = -\mathbf{c}^{(k)} + \beta_k \mathbf{d}^{(k-1)}}$$

Previous search direction

$$\beta_k = \left(\left\|\mathbf{c}^{(k)}\right\| / \left\|\mathbf{c}^{(k-1)}\right\|\right)^2$$

The current search direction is calculated by adding a scaled direction used in the previous iteration.

- **Step 4 : Compute a step size** $\alpha = \alpha_k$ **to minimize** $f(\mathbf{x}^{(k)} + \alpha \mathbf{d}^{(k)})$.

- **Step 5 : Change the design point as follows, then set** $k = k+1$ **and go to Step 2.**

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \alpha_k \mathbf{d}^{(k)}$$

3.1 Gradient Method
2. Conjugate Gradient Method (3/5) : Example

Minimize $f(x_1, x_2) = x_1 - x_2 + 2x_1^2 + 2x_1x_2 + x_2^2$ **Starting design point $\mathbf{x}^{(0)} = (0, 0)$**

$$\nabla f(\mathbf{x}) = \nabla f(x_1, x_2) = \begin{pmatrix} 1 + 4x_1 + 2x_2 \\ -1 + 2x_1 + 2x_2 \end{pmatrix}$$

☑ 1st Iteration: Find $\mathbf{x}^{(1)}$

$$\mathbf{d}^{(0)} = -\mathbf{c}^{(0)} = -\nabla f\left(\mathbf{x}^{(0)}\right) = -\nabla f \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$= -\begin{pmatrix} 1 + 4x_1 + 2x_2 \\ -1 + 2x_1 + 2x_2 \end{pmatrix} = -\begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

$$\mathbf{x}^{(1)} = \mathbf{x}^{(0)} + \alpha_0 \mathbf{d}^{(0)}$$

$$= \begin{pmatrix} 0 \\ 0 \end{pmatrix} + \alpha \begin{pmatrix} -1 \\ 1 \end{pmatrix} = \begin{pmatrix} -\alpha \\ \alpha \end{pmatrix}$$

Replacing α_0 to α for convenience

Substituting $\mathbf{x}^{(1)} = (-\alpha, \alpha)$ into the objective function

$$f(\mathbf{x}^{(1)}) = -\alpha - \alpha + 2\alpha^2 - 2\alpha^2 + \alpha^2$$

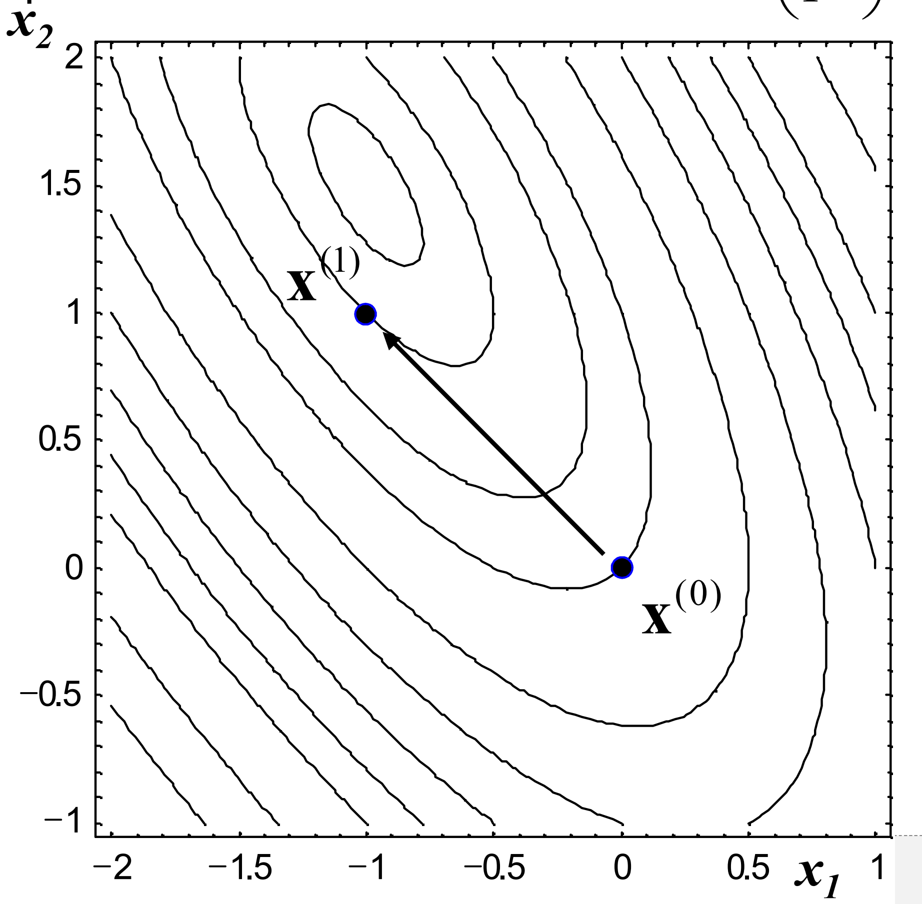
$$= \alpha^2 - 2\alpha$$

To minimize $f(\mathbf{x}^{(1)})$,

$$\frac{df(\mathbf{x}^{(1)})}{d\alpha} = 2\alpha - 2 = 0 \;\rightarrow\; \alpha = 1.0$$

$$\therefore \mathbf{x}^{(1)} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

Note: Step 1 of the conjugate gradient method and steepest descent method is the same.



3.1 Gradient Method
2. Conjugate Gradient Method (4/5): Example

Minimize $f(x_1, x_2) = x_1 - x_2 + 2x_1^2 + 2x_1x_2 + x_2^2$

■ 2nd Iteration-Find $\mathbf{x}^{(2)}$

Compute the gradient of the objective function as

$$\mathbf{c}^{(1)} = \nabla f\left(\mathbf{x}^{(1)}\right)$$

$$= \nabla f\begin{pmatrix} -1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 + 4x_1 + 2x_2 \\ -1 + 2x_1 + 2x_2 \end{pmatrix} = \begin{pmatrix} -1 \\ -1 \end{pmatrix}$$

Calculate the new search direction as

$$\mathbf{d}^{(1)} = -\mathbf{c}^{(1)} + \beta_1 \mathbf{d}^{(0)} = -\mathbf{c}^{(1)} + \frac{\left\|\nabla f\left(\mathbf{x}^{(1)}\right)\right\|^2}{\left\|\nabla f\left(\mathbf{x}^{(0)}\right)\right\|^2}\mathbf{d}^{(0)}$$

$$= -\begin{pmatrix} -1 \\ -1 \end{pmatrix} + \frac{2}{2}\begin{pmatrix} -1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \end{pmatrix}$$

$$\mathbf{x}^{(1)} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

$$\mathbf{d}^{(0)} = -\nabla f\left(\mathbf{x}^{(0)}\right) = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

$$\mathbf{d}^{(k)} = -\mathbf{c}^{(k)} + \beta_k \mathbf{d}^{(k-1)}$$

$$\beta_k = \left(\left\|\mathbf{c}^{(k)}\right\| / \left\|\mathbf{c}^{(k-1)}\right\|\right)^2$$

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \alpha_k \mathbf{d}^{(k)}$$

3.1 Gradient Method
2. Conjugate Gradient Method (5/5): Example

Minimize $f(x_1, x_2) = x_1 - x_2 + 2x_1^2 + 2x_1x_2 + x_2^2$

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \alpha_k \mathbf{d}^{(k)}$$

$$\mathbf{x}^{(1)} = \begin{pmatrix} -1 \\ 1 \end{pmatrix} \quad \mathbf{d}^{(1)} = \begin{pmatrix} 0 \\ 2 \end{pmatrix}$$

$$\nabla f(\mathbf{x}) = \nabla f(x_1, x_2) = \begin{pmatrix} 1 + 4x_1 + 2x_2 \\ -1 + 2x_1 + 2x_2 \end{pmatrix}$$

$$\mathbf{x}^{(2)} = \mathbf{x}^{(1)} + \alpha_1 \mathbf{d}^{(1)}$$

$$= \begin{pmatrix} -1 \\ 1 \end{pmatrix} + \alpha \begin{pmatrix} 0 \\ 2 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 + 2\alpha \end{pmatrix}$$ **Replacing** α_1 **to** α **for convenience**

Substituting $\mathbf{x}^{(2)} = (-1, 1 + 2\alpha)$ **into the objective function**

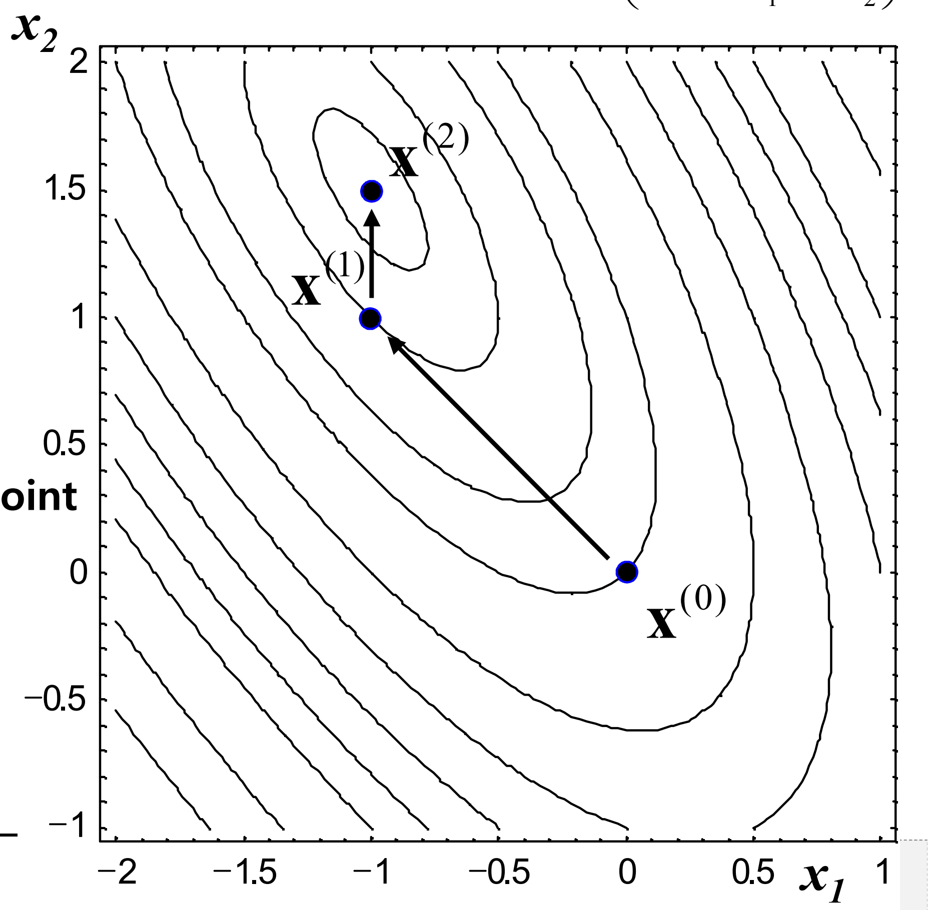
$$f(\mathbf{x}^{(2)}) = 4\alpha^2 - 2\alpha - 1$$

To minimize $f(\mathbf{x}^{(1)})$**,**

$$\frac{df(\mathbf{x}^{(2)})}{d\alpha} = 8\alpha - 2 = 0 \quad \rightarrow \quad \alpha = 0.25$$

$$\therefore \mathbf{x}^{(2)} = \begin{pmatrix} -1 \\ 1.5 \end{pmatrix}$$

→Minimum design point



Check stopping criterion.

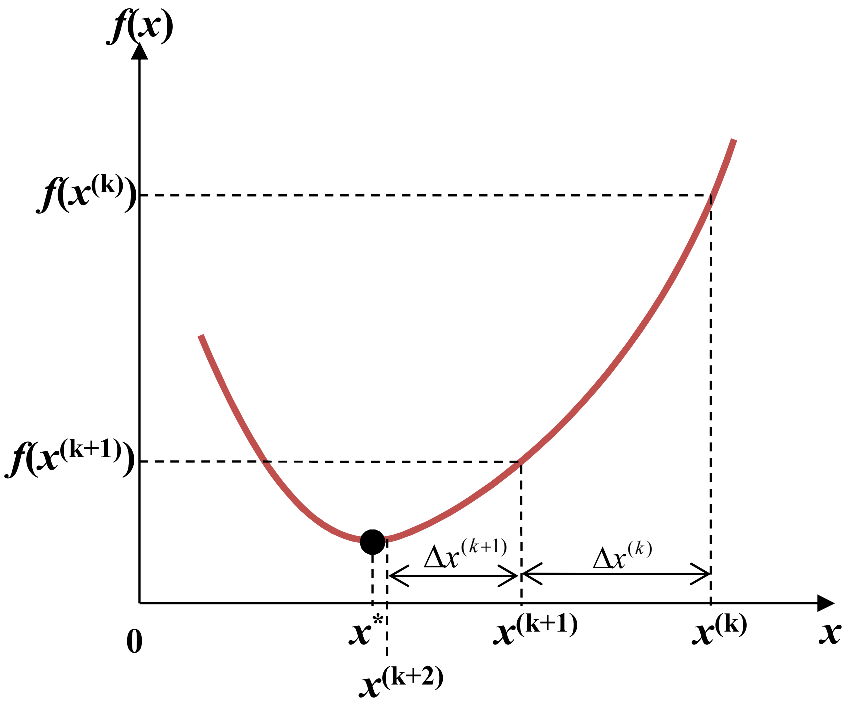
$$\mathbf{c}^{(2)} = \nabla f\left(\mathbf{x}^{(2)}\right) = \nabla f \begin{pmatrix} -1 \\ 1.5 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\left\| \mathbf{c}^{(2)} \right\| = 0 < \varepsilon$$ **→Stop!**

3.1 Gradient Method
3. Newton's Method (1)

Given: $f(x)$

Find: x^* **which minimizes** $f(x)$



Assume that $f(x)$ has minimum at $x^{(k+1)} = x^{(k)} + \Delta x^{(k)}$.

Consider the quadratic approximation of the function $f(x)$ at $x=x^{(k)}$ using the second-order Taylor expansion.

$$f(x^{(k)} + \Delta x^{(k)}) = f(x^{(k)}) + \frac{df(x^{(k)})}{dx}\Delta x^{(k)} + \frac{1}{2}\frac{d^2 f(x^{(k)})}{dx^2}\left(\Delta x^{(k)}\right)^2 + \cancel{O\left(\left(\Delta x^{(k)}\right)^3\right)}$$

In this equation, $x^{(k)}$ is a constant and $\Delta x^{(k)}$ **is a variable.** So, the following equation is a quadratic function in terms of $\Delta x^{(k)}$.

$$f(x^{(k)} + \Delta x^{(k)}) = f(x^{(k)}) + \frac{df(x^{(k)})}{dx}\Delta x^{(k)} + \frac{1}{2}\frac{d^2 f(x^{(k)})}{dx^2}\left(\Delta x^{(k)}\right)^2$$

Differentiate this equation with respect to $\Delta x^{(k)}$.

$$\frac{df(x^{(k)} + \Delta x^{(k)})}{d\Delta x^{(k)}} = \frac{df(x^{(k)})}{dx} + \frac{d^2 f(x^{(k)})}{dx^2}\Delta x^{(k)} = 0 \rightarrow$$

The necessary condition for minimization of this function

Calculate the small change $\Delta x^{(k)}$ in design.

$$\Delta x^{(k)} = \left(-\frac{df(x^{(k)})}{dx}\right) \Big/ \left(\frac{d^2 f(x^{(k)})}{dx^2}\right)$$

Is $|\Delta x^{(k)}| < \varepsilon$?

NO → $k = k+1$

YES → Set $x^* = x^{(k+1)}$ and stop the iteration.

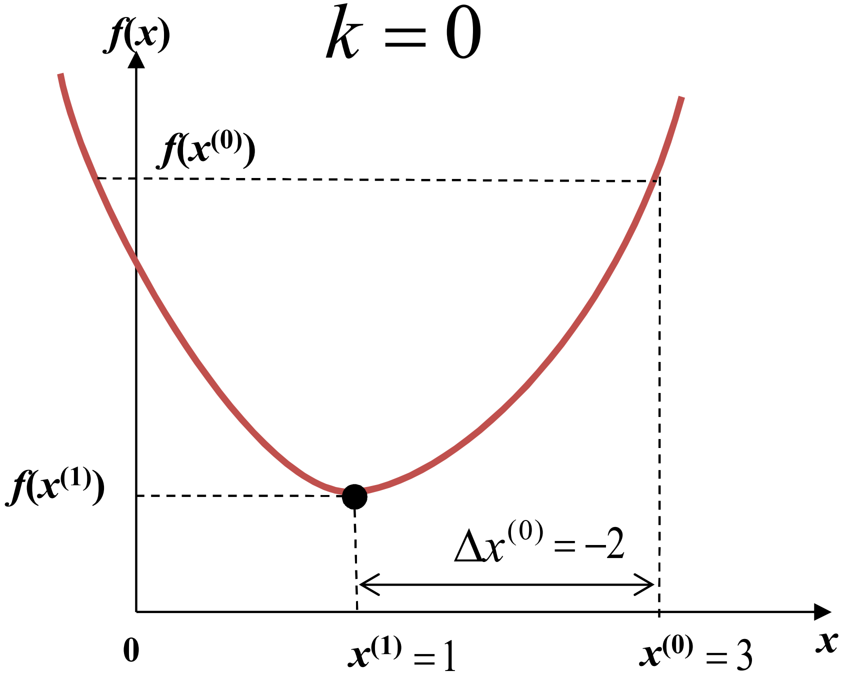
3. Newton's Method (2/9): Example

Given: $f(x) = x^2 - 2x + 2$

Starting design point $x^{(0)} = 3$

Find: x^* **which minimizes** $f(x)$

$$k = 0$$



Assume that $f(x)$ has minimum at $x^{(1)} = x^{(0)} + \Delta x^{(0)}$.

Consider the quadratic approximation of the function $f(x)$ at $x=x^{(0)}$ using the second-order Taylor expansion.

$$f(x^{(0)} + \Delta x^{(0)}) = f(x^{(0)}) + \frac{df(x^{(0)})}{dx}\Delta x^{(0)} + \frac{1}{2}\frac{d^2 f(x^{(0)})}{dx^2}\left(\Delta x^{(0)}\right)^2$$

In this equation, $x^{(0)}$ is a constant and $\Delta x^{(0)}$ is a variable. So, the following equation is a quadratic function in terms of $\Delta x^{(0)}$.

$$f(x^{(0)} + \Delta x^{(0)}) = f(x^{(0)}) + \frac{df(x^{(0)})}{dx}\Delta x^{(0)} + \frac{1}{2}\frac{d^2 f(x^{(0)})}{dx^2}\left(\Delta x^{(0)}\right)^2$$

Differentiate this equation with respect to $\Delta x^{(0)}$.

$$\frac{df(x^{(0)} + \Delta x^{(0)})}{d\Delta x^{(0)}} = \frac{df(x^{(0)})}{dx} + \frac{d^2 f(x^{(0)})}{dx^2}\Delta x^{(0)} = 0 \rightarrow$$

The necessary condition for minimization of this function

Calculate the small change $\Delta x^{(0)}$ in design.

$$\Delta x^{(0)} = \left(-\frac{df(x^{(0)})}{dx}\right) \Big/ \left(\frac{d^2 f(x^{(0)})}{dx^2}\right)$$

$$= \left(-2x + 2\right)_{x=3} / \left(2\right)_{x=3} = -2$$

Is | $\Delta x^{(0)}$ |< ε ?

NO

$$k = k + 1$$
$$= 0 + 1 = 1$$

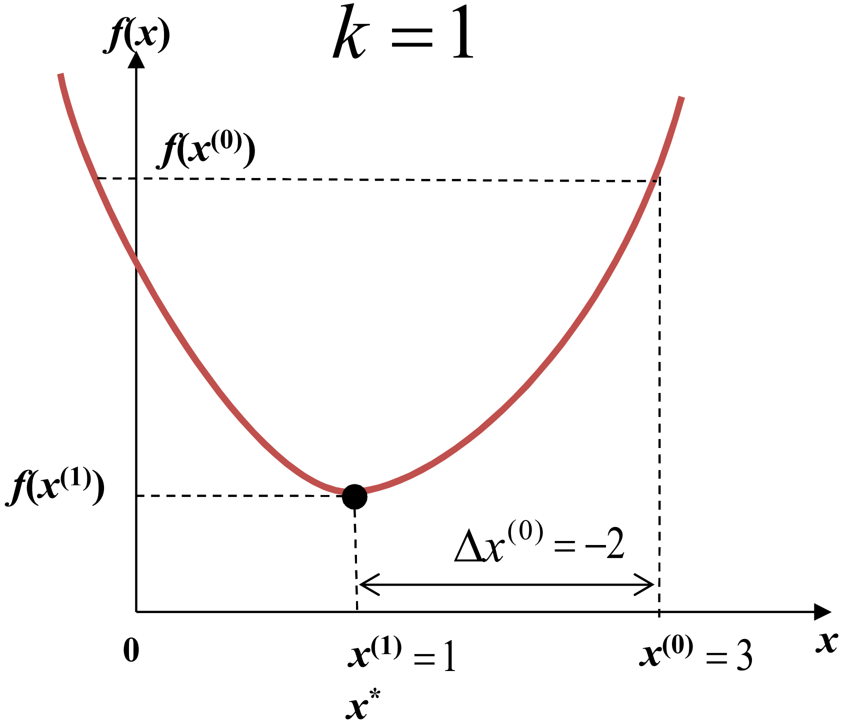
3. Newton's Method (3/9): Example

Given: $f(x) = x^2 - 2x + 2$

Starting design point $x^{(0)} = 3$

Find: x^* **which minimizes** $f(x)$

$$k = 1$$



Is it possible to find the x* which minimizes a cubic function at once?

Assume that $f(x)$ has minimum at $x^{(2)} = x^{(1)} + \Delta x^{(1)}$.

Consider the quadratic approximation of the function $f(x)$ at $x=x^{(1)}$ using the second-order Taylor expansion.

$$f(x^{(1)} + \Delta x^{(1)}) = f(x^{(1)}) + \frac{df(x^{(1)})}{dx}\Delta x^{(1)} + \frac{1}{2}\frac{d^2 f(x^{(1)})}{dx^2}\left(\Delta x^{(1)}\right)^2$$

In this equation, $x^{(1)}$ is a constant and $\Delta x^{(1)}$ is a variable. So, the following equation is a quadratic function in terms of $\Delta x^{(1)}$.

$$f(x^{(1)} + \Delta x^{(1)}) = f(x^{(1)}) + \frac{df(x^{(1)})}{dx}\Delta x^{(1)} + \frac{1}{2}\frac{d^2 f(x^{(1)})}{dx^2}\left(\Delta x^{(1)}\right)^2$$

Differentiate this equation with respect to $\Delta x^{(1)}$.

$$\frac{df(x^{(1)} + \Delta x^{(1)})}{d\Delta x^{(1)}} = \frac{df(x^{(1)})}{dx} + \frac{d^2 f(x^{(1)})}{dx^2}\Delta x^{(1)} = 0 \rightarrow$$

The necessary condition for minimization of this function

Calculate the small change $\Delta x^{(1)}$ in design.

$$\Delta x^{(1)} = \left(-\frac{df(x^{(1)})}{dx}\right) \Big/ \left(\frac{d^2 f(x^{(1)})}{dx^2}\right)$$

$$= (-2x+2)_{x=1} / (2)_{x=1} = 0$$

Is $|\Delta x^{(1)}| < \varepsilon$?

YES

Set $x^* = x^{(2)}$ and stop the iteration.

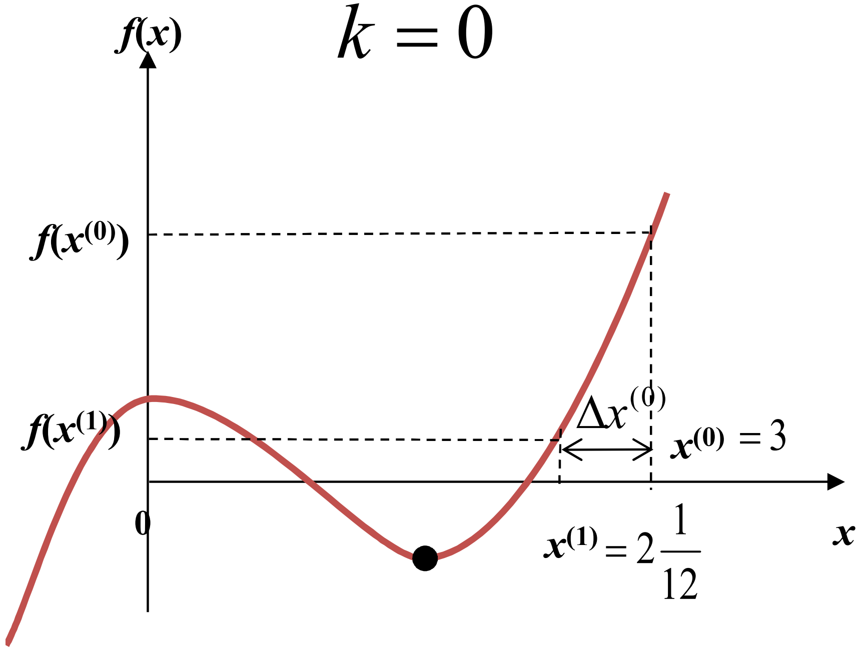
Is it possible to find the x* which minimizes a cubic function at once?

Given: $f(x) = x^3 - 3x^2 + 2x$

Starting design point $x^{(0)} = 3$

Find: x^* **which minimizes** $f(x)$

$$k = 0$$



Assume that $f(x)$ has minimum at $x^{(1)} = x^{(0)} + \Delta x^{(0)}$.

Consider the quadratic approximation of the function $f(x)$ at $x=x^{(0)}$ using the second-order Taylor expansion.

$$f(x^{(0)} + \Delta x^{(0)}) = f(x^{(0)}) + \frac{df(x^{(0)})}{dx}\Delta x^{(0)} + \frac{1}{2}\frac{d^2 f(x^{(0)})}{dx^2}\left(\Delta x^{(0)}\right)^2 + \cancel{O\left(\left(\Delta x^{(0)}\right)^3\right)}$$

In this equation, $x^{(0)}$ is a constant and $\Delta x^{(0)}$ is a variable. So, the following equation is a quadratic function in terms of $\Delta x^{(0)}$.

$$f(x^{(0)} + \Delta x^{(0)}) = f(x^{(0)}) + \frac{df(x^{(0)})}{dx}\Delta x^{(0)} + \frac{1}{2}\frac{d^2 f(x^{(0)})}{dx^2}\left(\Delta x^{(0)}\right)^2$$

Differentiate this equation with respect to $\Delta x^{(0)}$.

$$\frac{df(x^{(0)} + \Delta x^{(0)})}{d\Delta x^{(0)}} = \frac{df(x^{(0)})}{dx} + \frac{d^2 f(x^{(0)})}{dx^2}\Delta x^{(0)} = 0 \rightarrow$$

The necessary condition for minimization of this function

Calculate the small change $\Delta x^{(0)}$ in design.

$$\Delta x^{(0)} = \left(-\frac{df(x^{(0)})}{dx}\right) \Big/ \left(\frac{d^2 f(x^{(0)})}{dx^2}\right)$$

$$= \left(-3x^2 + 6x - 2\right)_{x=3} / \left(6x - 6\right)_{x=3} = -\frac{11}{12}$$

Is $|\Delta x^{(0)}| < \varepsilon$?

NO

$$k = k + 1$$
$$= 0 + 1 = 1$$

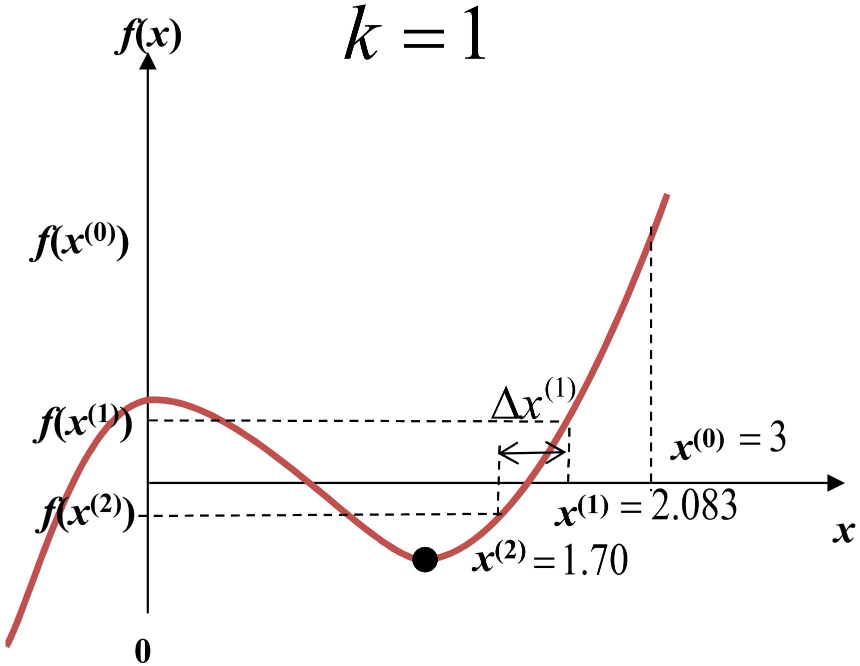
3.1 Gradient Method
3. Newton's Method (5/9): Example

Is it possible to find the x* which minimizes a cubic function at once?

Given: $f(x) = x^3 - 3x^2 + 2x$

Starting design point $x^{(0)} = 3$

Find: x^* **which minimizes** $f(x)$



Why is it not possible to find the x* which minimizes a cubic function at once?

Assume that $f(x)$ has minimum at $x^{(2)} = x^{(1)} + \Delta x^{(1)}$.

Consider the quadratic approximation of the function $f(x)$ at $x=x^{(1)}$ using the second-order Taylor expansion.

$$f(x^{(1)} + \Delta x^{(1)}) = f(x^{(1)}) + \frac{df(x^{(1)})}{dx}\Delta x^{(1)} + \frac{1}{2}\frac{d^2 f(x^{(1)})}{dx^2}(\Delta x^{(1)})^2 + \cancel{O((\Delta x^{(1)})^3)}$$

In this equation, $x^{(1)}$ is a constant and $\Delta x^{(1)}$ is a variable. So, the following equation is a quadratic function in terms of $\Delta x^{(1)}$.

$$f(x^{(1)} + \Delta x^{(1)}) = f(x^{(1)}) + \frac{df(x^{(1)})}{dx}\Delta x^{(1)} + \frac{1}{2}\frac{d^2 f(x^{(1)})}{dx^2}(\Delta x^{(1)})^2$$

Differentiate this equation with respect to $\Delta x^{(1)}$.

$$\frac{df(x^{(1)} + \Delta x^{(1)})}{d\Delta x^{(1)}} = \frac{df(x^{(1)})}{dx} + \frac{d^2 f(x^{(1)})}{dx^2}\Delta x^{(1)} = 0 \rightarrow$$

The necessary condition for minimization of this function

Calculate the small change $\Delta x^{(1)}$ in design.

$$\Delta x^{(1)} = \left(-\frac{df(x^{(1)})}{dx}\right) / \left(\frac{d^2 f(x^{(1)})}{dx^2}\right)$$

$$= (-3x^2 + 6x - 2)_{x=\frac{25}{12}} / (6x - 6)_{x=\frac{25}{12}} = -0.388$$

Is $|\Delta x^{(1)}| < \varepsilon$?

NO

$k = k + 1$

$= 1 + 1 = 2$

Since the second-order Taylor expansion is just an approximation for $f(x)$ at the point $x^{(0)}$ or $x^{(1)}$, $x^{(1)}$ or $x^{(2)}$ will probably not be the precise minimum design point of $f(x)$.

3.1 Gradient Method
3. Newton's Method (6/9): Example of Function of Two Variables

Minimize $f(\mathbf{x}) = f(x_1, x_2) = x_1 - x_2 + 2x_1^2 + 2x_1x_2 + x_2^2$, **Starting design point** $\mathbf{x}^{(0)} = (0, 0)$

$$\nabla f(\mathbf{x}) = \nabla f(x_1, x_2) = \begin{pmatrix} f_{x_1} \\ f_{x_2} \end{pmatrix} = \begin{pmatrix} 1 + 4x_1 + 2x_2 \\ -1 + 2x_1 + 2x_2 \end{pmatrix}, \quad \mathbf{H}(\mathbf{x}) = \begin{pmatrix} \dfrac{\partial^2 f}{\partial x_1^2} & \dfrac{\partial^2 f}{\partial x_1 \partial x_2} \\ \dfrac{\partial^2 f}{\partial x_1 \partial x_2} & \dfrac{\partial^2 f}{\partial x_2^2} \end{pmatrix} = \begin{pmatrix} 4 & 2 \\ 2 & 2 \end{pmatrix}$$

- **1st Iteration: Find** $\mathbf{x}^{(1)}$

Assume that $f(x)$ has minimum at $\mathbf{x}^{(1)} = \mathbf{x}^{(0)} + \Delta\mathbf{x}^{(0)}$.

Consider the quadratic approximation of the function $f(\mathbf{x})$ at $\mathbf{x}=\mathbf{x}^{(0)}$ using the second-order Taylor expansion.

How?

$$f(\mathbf{x}^{(0)} + \Delta\mathbf{x}^{(0)}) = f(\mathbf{x}^{(0)}) + \nabla f(\mathbf{x}^{(0)})^T \Delta\mathbf{x}^{(0)} + \frac{1}{2}(\Delta\mathbf{x}^{(0)})^T \mathbf{H}(\mathbf{x}^{(0)}) \Delta\mathbf{x}^{(0)}$$

In this equation, $\mathbf{x}^{(0)}$ is a constant and $\Delta\mathbf{x}^{(0)}$ is a variable. So, the following equation is a quadratic function in terms of $\Delta\mathbf{x}^{(0)}$.

$$f(\mathbf{x}^{(0)} + \Delta\mathbf{x}^{(0)}) = f(\mathbf{x}^{(0)}) + \nabla f(\mathbf{x}^{(0)})^T \Delta\mathbf{x}^{(0)} + \frac{1}{2}(\Delta\mathbf{x}^{(0)})^T \mathbf{H}(\mathbf{x}^{(0)}) \Delta\mathbf{x}^{(0)}$$

3. Newton's Method (7/9): Example of Function of Two Variables

Minimize $f(\mathbf{x}) = f(x_1, x_2) = x_1 - x_2 + 2x_1^2 + 2x_1x_2 + x_2^2$, **Starting design point** $\mathbf{x}^{(0)} = (0, 0)$

■ 1st Iteration: Find $\mathbf{x}^{(1)}$

$$f(\mathbf{x}^{(0)} + \Delta\mathbf{x}^{(0)}) = f(\mathbf{x}^{(0)}) + \nabla f(\mathbf{x}^{(0)})^T \Delta\mathbf{x}^{(0)} + \frac{1}{2}(\Delta\mathbf{x}^{(0)})^T \mathbf{H}(\mathbf{x}^{(0)})\Delta\mathbf{x}^{(0)}$$

Differentiate this equation with respect to $\Delta x^{(0)}$. **How?**

$$\nabla f(\mathbf{x}) = \nabla f(x_1, x_2) = \begin{pmatrix} f_{x_1} \\ f_{x_2} \end{pmatrix} = \begin{pmatrix} 1 + 4x_1 + 2x_2 \\ -1 + 2x_1 + 2x_2 \end{pmatrix}$$

$$\frac{\partial f(\mathbf{x}^{(0)} + \Delta\mathbf{x}^{(0)})}{\partial(\Delta\mathbf{x}^{(0)})} = \nabla f(\mathbf{x}^{(0)}) + \mathbf{H}(\mathbf{x}^{(0)})\Delta\mathbf{x}^{(0)} = 0 \longrightarrow$$

The necessary condition for minimization of function $f(x_1, x_2)$

Calculate the small change $\Delta x^{(0)}$ **in design.**

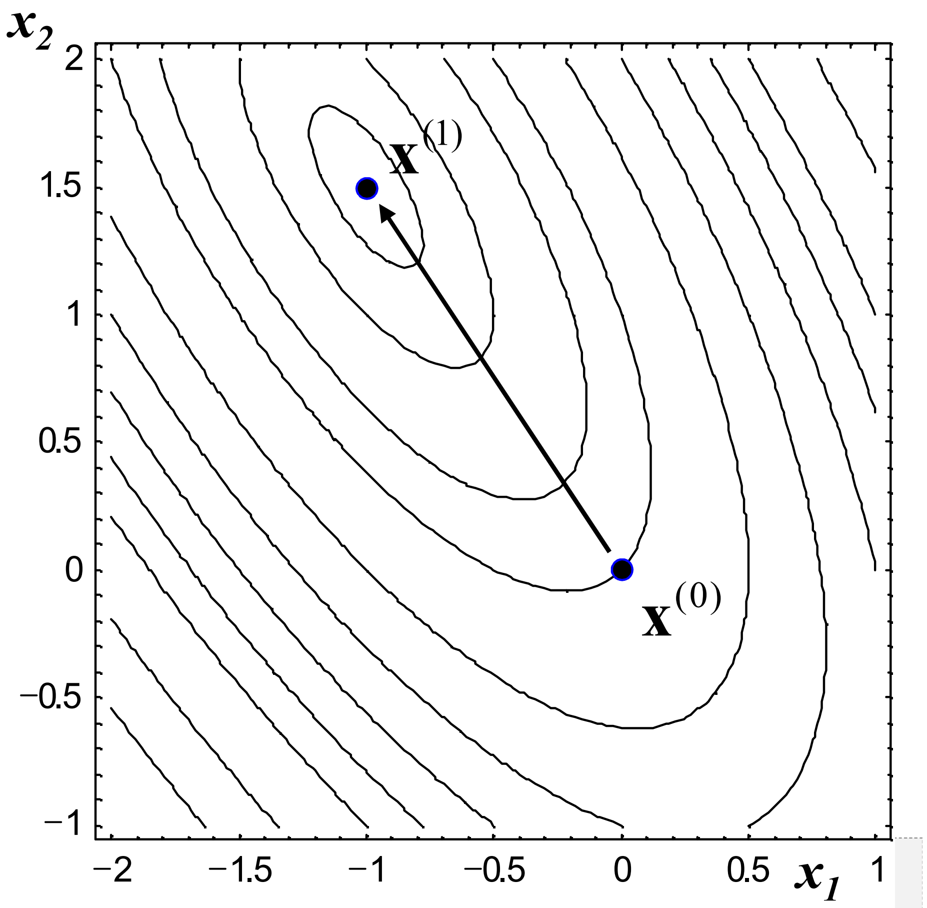
$$\mathbf{H}(\mathbf{x}^{(0)})\Delta\mathbf{x}^{(0)} = -\nabla f(\mathbf{x}^{(0)})$$

$$\Delta\mathbf{x}^{(k)} = -[\mathbf{H}(\mathbf{x}^{(k)})]^{-1}\nabla f(\mathbf{x}^{(k)})$$

$$\downarrow \quad \left(-\nabla f(\mathbf{x}^{(0)}) = \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \quad \mathbf{H}(\mathbf{x}^{(0)}) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f}{\partial x_1 \partial x_2} & \frac{\partial^2 f}{\partial x_2^2} \end{pmatrix} = \begin{pmatrix} 4 & 2 \\ 2 & 2 \end{pmatrix} \right)$$

$$\begin{pmatrix} \Delta\mathbf{x}_1^{(0)} \\ \Delta\mathbf{x}_2^{(0)} \end{pmatrix} = \begin{pmatrix} 4 & 2 \\ 2 & 2 \end{pmatrix}^{-1} \begin{pmatrix} -1 \\ 1 \end{pmatrix} \rightarrow \begin{pmatrix} \Delta\mathbf{x}_1^{(0)} \\ \Delta\mathbf{x}_2^{(0)} \end{pmatrix} = \begin{pmatrix} -1 \\ 1.5 \end{pmatrix}$$

$$\therefore x^{(1)} = x^{(0)} + \Delta\mathbf{x}^{(0)} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} + \begin{pmatrix} -1 \\ 1.5 \end{pmatrix} = \begin{pmatrix} -1 \\ 1.5 \end{pmatrix}$$



3. Newton's Method (8/9): Example of Function of Two Variables

Minimize $f(\mathbf{x}) = f(x_1, x_2) = x_1 - x_2 + 2x_1^2 + 2x_1x_2 + x_2^2$, **Starting design point** $\mathbf{x}^{(0)} = (0, 0)$

■ 2nd Iteration-Find $\mathbf{x}^{(2)}$

$x^{(1)} = \begin{pmatrix} -1 \\ 1.5 \end{pmatrix}$

In the same way as 1st Iteration,

Assume that $f(x)$ **has minimum at** $\mathbf{x}^{(2)} = \mathbf{x}^{(1)} + \Delta\mathbf{x}^{(1)}$.

Consider the quadratic approximation of the function $f(\mathbf{x})$ **at** $\mathbf{x}=\mathbf{x}^{(1)}$ **using the second-order Taylor expansion.**

$$f(\mathbf{x}^{(1)} + \Delta\mathbf{x}^{(1)}) = f(\mathbf{x}^{(1)}) + \nabla f(\mathbf{x}^{(1)})^T \Delta\mathbf{x}^{(1)} + \frac{1}{2}(\Delta\mathbf{x}^{(1)})^T \mathbf{H}(\mathbf{x}^{(1)})\Delta\mathbf{x}^{(1)}$$

In this equation, $\mathbf{x}^{(1)}$ **is a constant and** $\Delta x^{(1)}$ **is a variable. So, the following equation is a quadratic function in terms of** $\Delta x^{(1)}$**.**

$$f(\mathbf{x}^{(1)} + \Delta\mathbf{x}^{(1)}) = f(\mathbf{x}^{(1)}) + \nabla f(\mathbf{x}^{(1)})^T \Delta\mathbf{x}^{(1)} + \frac{1}{2}(\Delta\mathbf{x}^{(1)})^T \mathbf{H}(\mathbf{x}^{(1)})\Delta\mathbf{x}^{(1)}$$

Differentiate this equation with respect to $\Delta x^{(1)}$.

$$\frac{\partial f(\mathbf{x}^{(1)} + \Delta\mathbf{x}^{(1)})}{\partial(\Delta\mathbf{x}^{(1)})} = \nabla f(\mathbf{x}^{(1)}) + \mathbf{H}(\mathbf{x}^{(1)})\Delta\mathbf{x}^{(1)} = 0 \longrightarrow$$

The necessary condition for minimization of function $f(x_1, x_2)$

3. Newton's Method (9/9): Example of Function of Two Variables

Minimize $f(\mathbf{x}) = f(x_1, x_2) = x_1 - x_2 + 2x_1^2 + 2x_1x_2 + x_2^2$, **Starting design point $\mathbf{x}^{(0)}$ = (0, 0)**

■ 2nd Iteration-Find $\mathbf{x}^{(2)}$

$$x^{(1)} = \begin{pmatrix} -1 \\ 1.5 \end{pmatrix}$$

Calculate the small change $\Delta x^{(1)}$ in design.

$$\nabla f(\mathbf{x}) = \nabla f(x_1, x_2) = \begin{pmatrix} f_{x_1} \\ f_{x_2} \end{pmatrix} = \begin{pmatrix} 1 + 4x_1 + 2x_2 \\ -1 + 2x_1 + 2x_2 \end{pmatrix}$$

$$\mathbf{H}(\mathbf{x}^{(1)})\Delta\mathbf{x}^{(1)} = -\nabla f(\mathbf{x}^{(1)})$$

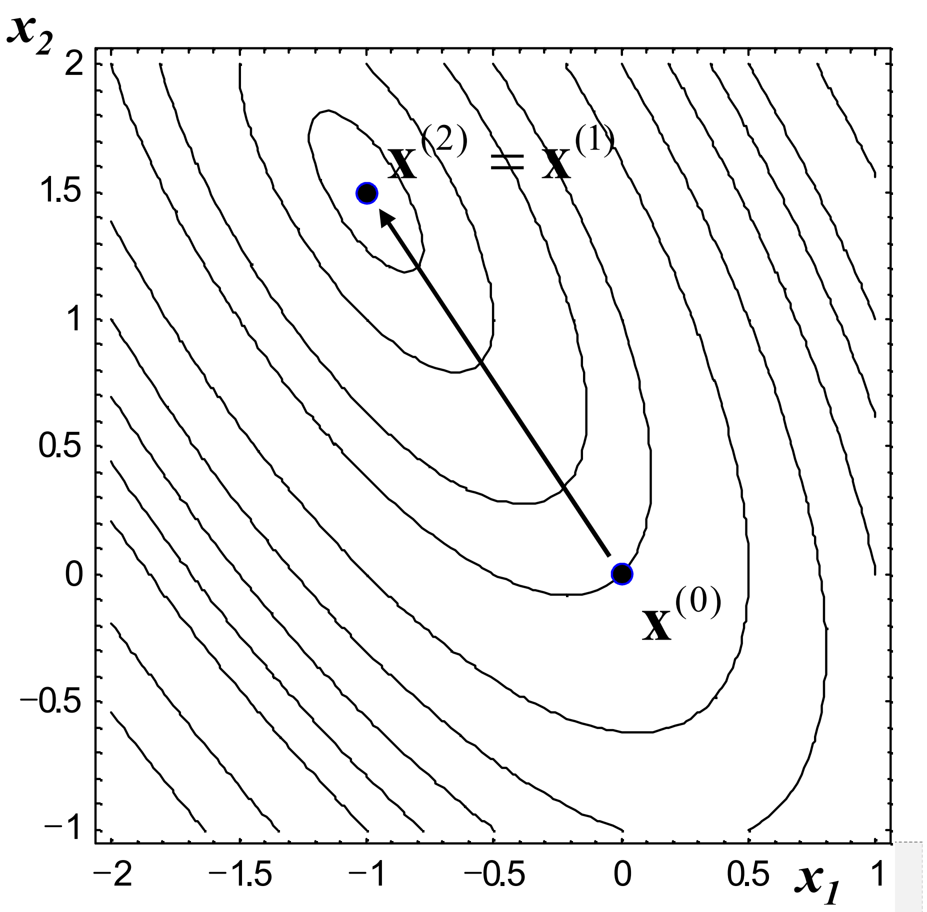
$$\Delta\mathbf{x}^{(k)} = -[\mathbf{H}(\mathbf{x}^{(k)})]^{-1}\nabla f(\mathbf{x}^{(k)})$$

$$\left(-\nabla f(\mathbf{x}^{(1)}) = \begin{pmatrix} 0 \\ 0 \end{pmatrix},\ \mathbf{H}\left(\mathbf{x}^{(1)}\right) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f}{\partial x_1 \partial x_2} & \frac{\partial^2 f}{\partial x_2^2} \end{pmatrix} = \begin{pmatrix} 4 & 2 \\ 2 & 2 \end{pmatrix} \right)$$

$$\begin{pmatrix} \Delta\mathbf{x}_1^{(1)} \\ \Delta\mathbf{x}_2^{(1)} \end{pmatrix} = \begin{pmatrix} 4 & 2 \\ 2 & 2 \end{pmatrix}^{-1} \begin{pmatrix} 0 \\ 0 \end{pmatrix} \rightarrow \begin{pmatrix} \Delta\mathbf{x}_1^{(1)} \\ \Delta\mathbf{x}_2^{(1)} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\therefore x^{(2)} = x^{(1)} + \Delta\mathbf{x}^{(1)} = \begin{pmatrix} -1 \\ 1.5 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} -1 \\ 1.5 \end{pmatrix}$$

→Optimal design point



Check stopping criterion.

$$\left|\Delta\mathbf{x}^{(1)}\right| = 0 < \varepsilon$$

→Stop!

3. Modified Newton's Method (1/2)

☑ In this method, we treat $\Delta\mathbf{x}^{(k)} = -[\mathbf{H}(\mathbf{x}^{(k)})]^{-1}\nabla f(\mathbf{x}^{(k)})$ of the Newton's method as the search direction and use any of the one-dimensional search methods to calculate the step size in the search direction.

- Step 1 : Estimate a starting design point $\mathbf{x}^{(0)}$.
  Set iteration counter $k=0$. Specify a tolerance ε for the stopping criterion.

- Step 2 : Calculate $c_i^{(k)} = \partial f(\mathbf{x}^{(k)})/\partial x_i$ for $i=1$ to n. If $\left\|\mathbf{c}^{(k)}\right\| < \varepsilon$, then stop the iterative process. Otherwise, continue.

- Step 3 : Calculate the Hessian matrix $\mathbf{H}^{(k)}$ at current design point $\mathbf{x}^{(k)}$.

$$\mathbf{H}(\mathbf{x}^{(k)}) = \left[\frac{\partial^2 f}{\partial x_i \partial x_j}\right], \quad i=1,\cdots,n; \ j=1,\cdots,n$$

3. Modified Newton's Method (2/2)

■ Step 4 : Calculate the search direction as follows:

$$\boxed{\mathbf{d}^{(k)} = \Delta\mathbf{x}^{(k)} = -\mathbf{H}^{-1}\mathbf{c}^{(k)}}$$

When $f(\mathbf{x}^* + \Delta\mathbf{x}) = f(\mathbf{x}^*) + \mathbf{c}^T\Delta\mathbf{x} + \frac{1}{2}\Delta\mathbf{x}^T\mathbf{H}(\mathbf{x}^*)\,\Delta\mathbf{x}$,

the necessary condition for minimization of this function is as follows:

$$df(\Delta\mathbf{x})/d\Delta\mathbf{x} = \mathbf{c} + \mathbf{H}(\mathbf{x}^*)\Delta\mathbf{x} = 0$$

$$\Rightarrow \mathbf{H}(\mathbf{x}^*)\,\Delta\mathbf{x} = -\mathbf{c} \Rightarrow \ \Delta\mathbf{x} = -\mathbf{H}(\mathbf{x}^*)^{-1}\mathbf{c}$$

■ Step 5 : Update the design point as $\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \alpha\mathbf{d}^{(k)}$, where α is calculated to minimize $f(\mathbf{x}^{(k)} + \alpha\mathbf{d}^{(k)})$. Any one-dimensional search method may be used to calculate α.

■ Step 6 : Set $k = k+1$ and go to Step 2.

3. Disadvantages of the Newton's Method

The Newton's method is not very useful in practice, due to following features of the method:

1. It requires the storing of the $n \times n$ matrix $\mathbf{H}(\mathbf{x}^{(k)})$.

2. It becomes very difficult and sometimes, impossible to compute the elements of the matrix $\mathbf{H}(\mathbf{x}^{(k)})$.

3. It requires the inversion of the matrix $\mathbf{H}(\mathbf{x}^{(k)})$ at each iteration.

4. It requires the evaluation of the quantity $\mathbf{H}(\mathbf{x}^{(k)})^{-1}\nabla f(\mathbf{x}^{(k)})$ at each iteration.

4. Davidon-Fletcher-Powell(DFP) Method (1/6)

☑ This method builds an approximation for the inverse of the Hessian matrix of $f(\mathbf{x})$ using only the first derivatives.

- Step 1 : Estimate a starting design point $\mathbf{x}^{(0)}$.
  Choose a symmetric positive definite nxn matrix $\mathbf{A}^{(0)}$ as an approximation for the inverse of the Hessian matrix of the objective function. In the absence of more information, $\mathbf{A}^{(0)} = \mathbf{I}$ may be chosen. Also, specify a tolerance ε for the stopping criterion. Set $k = 0$ and compute the gradient vector as $\mathbf{d}^{(0)} = -\mathbf{c}^{(0)} \equiv -\nabla f(\mathbf{x}^{(0)})$.

- Step 2 : Calculate the norm of the gradient vector as $\|\mathbf{c}^{(k)}\|$.
  If $\|\mathbf{c}^{(k)}\| < \varepsilon$, then stop the iterative process. Otherwise, continue.
  It is noted that Step 1 and 2 of this method and the steepest descent method are the same.

4. Davidon-Fletcher-Powell(DFP) Method (2/6)

■ Step 3 : Calculate the search direction as follows:

$$\boxed{\mathbf{d}^{(k)} = -\mathbf{A}^{(k)}\mathbf{c}^{(k)}}$$

Newton's method

$$\Delta\mathbf{x}^{(k)} = -[\mathbf{H}(\mathbf{x}^{(k)})]^{-1}\nabla f(\mathbf{x}^{(k)})$$

$$\therefore \mathbf{d}^{(k)} = -(\mathbf{H}^{(k)})^{-1}\mathbf{c}^{(k)}$$

Here, the matrix $\mathbf{A}$ is used as an estimate for the inverse of the Hessian matrix $\mathbf{H}^{-1}$ of the objective function.

■ Step 4 : Compute optimum step size $\alpha_k = \alpha$ to minimize $f(\mathbf{x}^{(k)} + \alpha\mathbf{d}^{(k)})$.

■ Step 5 : Update the design point as $\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \alpha_k\mathbf{d}^{(k)}$.

4. Davidon-Fletcher-Powell(DFP) Method (3/6)

$\mathbf{d}^{(k)}$: search direction
$\alpha^{(k)}$: optimum step size

- Step 6 : Update the matrix $\mathbf{A}^{(k)}$ - approximation for the inverse of the Hessian matrix of the objective function - as follows:

$$\mathbf{A}^{(k+1)} = \mathbf{A}^{(k)} + \mathbf{B}^{(k)} + \mathbf{C}^{(k)} \quad ; \quad n \times n \text{ matrix}$$

where, the correction matrices $\mathbf{B}^{(k)}$ and $\mathbf{C}^{(k)}$ are calculated as below.

$$\mathbf{B}^{(k)} = \frac{\mathbf{s}^{(k)}(\mathbf{s}^{(k)})^T}{(\mathbf{s}^{(k)} \cdot \mathbf{y}^{(k)})} \quad ; \quad n \times n \text{ matrix} \qquad \mathbf{C}^{(k)} = \frac{-\mathbf{z}^{(k)}(\mathbf{z}^{(k)})^T}{(\mathbf{y}^{(k)} \cdot \mathbf{z}^{(k)})} \quad ; \quad n \times n \text{ matrix}$$

$$\mathbf{s}^{(k)} = \alpha_k \mathbf{d}^{(k)} \quad : \quad n \times 1 \text{ matrix}$$

$$\mathbf{y}^{(k)} = \mathbf{c}^{(k+1)} - \mathbf{c}^{(k)} \quad : \quad n \times 1 \text{ matrix}$$

$$\mathbf{c}^{(k+1)} = \nabla f(\mathbf{x}^{(k+1)}) \quad : \quad n \times 1 \text{ matrix}$$

$$\mathbf{z}^{(k)} = \mathbf{A}^{(k)} \mathbf{y}^{(k)} \quad : \quad [n \times n][n \times 1] = [n \times 1] \text{ matrix}$$

- Step 7 : Set $k = k + 1$ and go to Step 2.

3.1 Gradient Method
4. Davidon-Fletcher-Powell(DFP) Method (4/6): Example

Minimize $f(\mathbf{x}) = f(x_1, x_2) = x_1 - x_2 + 2x_1^2 + 2x_1x_2 + x_2^2$, **Starting design point** $\mathbf{x}^{(0)} = (0, 0)$

$$\nabla f(\mathbf{x}) = \nabla f(x_1, x_2) = \begin{pmatrix} 1 + 4x_1 + 2x_2 \\ -1 + 2x_1 + 2x_2 \end{pmatrix}$$

■ 1st Iteration: Find $\mathbf{x}^{(1)}$

$$\mathbf{x}^{(0)} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \ \mathbf{A}^{(0)} = \mathbf{I}$$

$$\mathbf{c}^{(0)} = \nabla f(\mathbf{x}^{(0)}) = \begin{pmatrix} 1 + 4 \cdot 0 + 2 \cdot 0 \\ -1 + 2 \cdot 0 + 2 \cdot 0 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

Check stopping criterion.

$$\left\| \mathbf{c}^{(0)} \right\| = \sqrt{1^2 + (-1)^2} = \sqrt{2} > \varepsilon$$

$$\mathbf{d}^{(0)} = -\mathbf{A}^{(0)}\mathbf{c}^{(0)} = -\mathbf{I}\,\mathbf{c}^{(0)} = -\mathbf{c}^{(0)} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

$$\mathbf{x}^{(1)} = \mathbf{x}^{(0)} + \alpha_0 \mathbf{d}^{(0)}$$

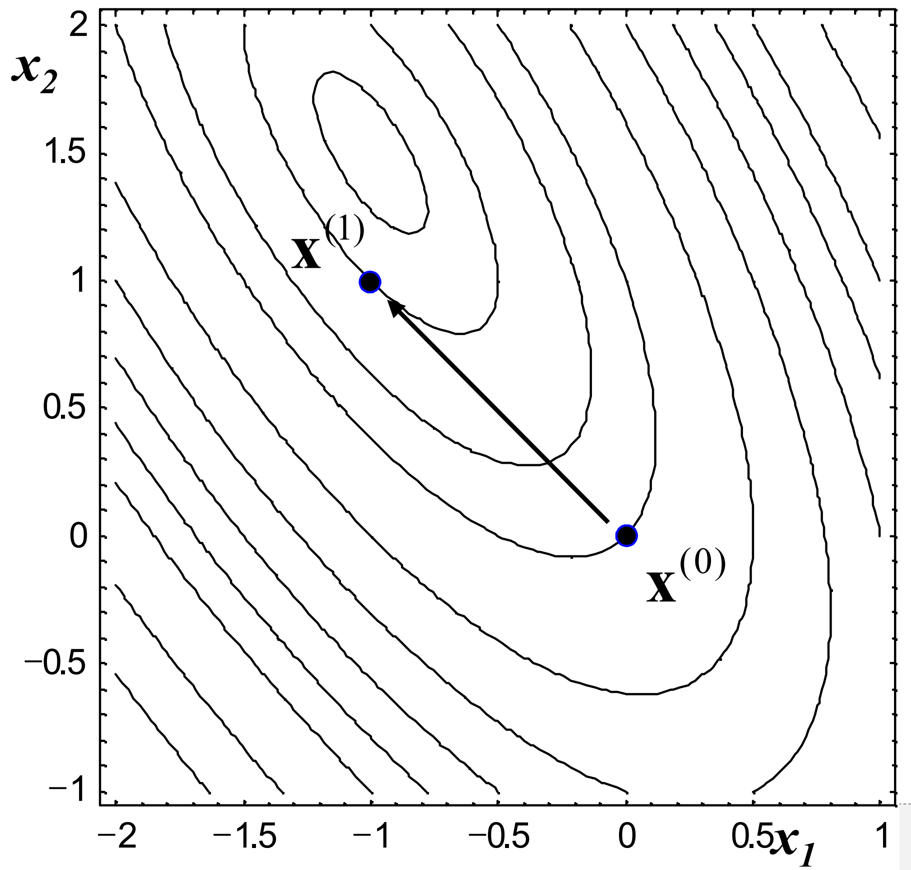
$$= \begin{pmatrix} 0 \\ 0 \end{pmatrix} + \alpha \begin{pmatrix} -1 \\ 1 \end{pmatrix} = \begin{pmatrix} -\alpha \\ \alpha \end{pmatrix}$$ **Replacing** α_0 **to** α **for convenience**

Substitute $\mathbf{x}^{(1)} = (-\alpha, \alpha)$ **into the objective function**

$$f(\mathbf{x}^{(1)}) = \alpha^2 - 2\alpha$$

To minimize $f(\mathbf{x}^{(1)})$,

$$\frac{df(\mathbf{x}^{(1)})}{d\alpha} = 2\alpha - 2 = 0 \ \rightarrow \ \alpha = 1.0 \qquad \therefore \mathbf{x}^{(1)} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$



$$\nabla f(\mathbf{x}) = \nabla f(x_1, x_2) = \begin{pmatrix} 1 + 4x_1 + 2x_2 \\ -1 + 2x_1 + 2x_2 \end{pmatrix}$$

4. Davidon-Fletcher-Powell(DFP) Method (5/6): Example

Minimize $f(\mathbf{x}) = f(x_1, x_2) = x_1 - x_2 + 2x_1^2 + 2x_1x_2 + x_2^2$, **Starting design point $\mathbf{x}^{(0)} = (0, 0)$**

■ 2nd Iteration: Find $\mathbf{x}^{(2)}$

Update the matrix $\mathbf{A}^{(1)}$ - approximation for the inverse of the Hessian matrix of the objective function - as follows:

$$\mathbf{A}^{(1)} = \mathbf{A}^{(0)} + \mathbf{B}^{(0)} + \mathbf{C}^{(0)}$$

$$\mathbf{B}^{(0)} = \frac{\mathbf{s}^{(0)}\mathbf{s}^{(0)^T}}{\mathbf{s}^{(0)} \cdot \mathbf{y}^{(0)}}$$

$$\mathbf{s}^{(0)} = \alpha \mathbf{d}^{(0)} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

$$\mathbf{c}^{(0)} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad \mathbf{c}^{(1)} = \begin{pmatrix} -1 \\ -1 \end{pmatrix}$$

$$\mathbf{y}^{(0)} = \mathbf{c}^{(1)} - \mathbf{c}^{(0)} = \begin{pmatrix} -2 \\ 0 \end{pmatrix}$$

$$\mathbf{s}^{(0)}\mathbf{s}^{(0)^T} = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$$

$$\mathbf{s}^{(0)} \cdot \mathbf{y}^{(0)} = 2$$

$$= \begin{pmatrix} 0.5 & -0.5 \\ -0.5 & 0.5 \end{pmatrix}$$

$$\mathbf{C}^{(0)} = \frac{-\mathbf{z}^{(0)}\mathbf{z}^{(0)T}}{\mathbf{y}^{(0)} \cdot \mathbf{z}^{(0)}}$$

$$\mathbf{A}^{(0)} = \mathbf{I} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$\mathbf{z}^{(0)} = \mathbf{A}^{(0)}\mathbf{y}^{(0)} = \begin{pmatrix} -2 \\ 0 \end{pmatrix}$$

$$\mathbf{y}^{(0)} \cdot \mathbf{z}^{(0)} = 4$$

$$\mathbf{z}^{(0)}\mathbf{z}^{(0)^T} = \begin{pmatrix} 4 & 0 \\ 0 & 0 \end{pmatrix}$$

$$= \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix}$$

$$\mathbf{A}^{(1)} = \mathbf{A}^{(0)} + \mathbf{B}^{(0)} + \mathbf{C}^{(0)}$$

$$= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 0.5 & -0.5 \\ -0.5 & 0.5 \end{pmatrix} + \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix}$$

$$= \begin{pmatrix} 0.5 & -0.5 \\ -0.5 & 1.5 \end{pmatrix}$$

3.1 Gradient Method
4. Davidon-Fletcher-Powell(DFP) Method (6/6): Example

$$\nabla f(\mathbf{x}) = \nabla f(x_1, x_2) = \begin{pmatrix} 1+4x_1+2x_2 \\ -1+2x_1+2x_2 \end{pmatrix}$$

Minimize $f(\mathbf{x}) = f(x_1, x_2) = x_1 - x_2 + 2x_1^2 + 2x_1x_2 + x_2^2$, **Starting design point** $\mathbf{x}^{(0)} = (0, 0)$

■ 2nd Iteration: Find $\mathbf{x}^{(2)}$

Check stopping criterion.

$$\left\| \mathbf{c}^{(1)} \right\| = \sqrt{2} > \varepsilon$$

$$\mathbf{d}^{(1)} = -\mathbf{A}^{(1)}\mathbf{c}^{(1)} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

$$\mathbf{x}^{(2)} = \mathbf{x}^{(1)} + \alpha_1 \mathbf{d}^{(1)}$$

$$= \begin{pmatrix} -1 \\ 1 \end{pmatrix} + \alpha \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ 1+\alpha \end{pmatrix}$$ **Replacing** α_1 **to** α **for convenience**

Substitute $\mathbf{x}^{(2)} = (-1, 1+\alpha)$ **into the objective function**

$$f(\mathbf{x}^{(2)}) = \alpha^2 - \alpha - 1$$

To minimize $f(\mathbf{x}^{(2)})$**,**

$$\frac{df(\mathbf{x}^{(2)})}{d\alpha} = 2\alpha - 1 = 0 \;\rightarrow\; \alpha = 0.5$$

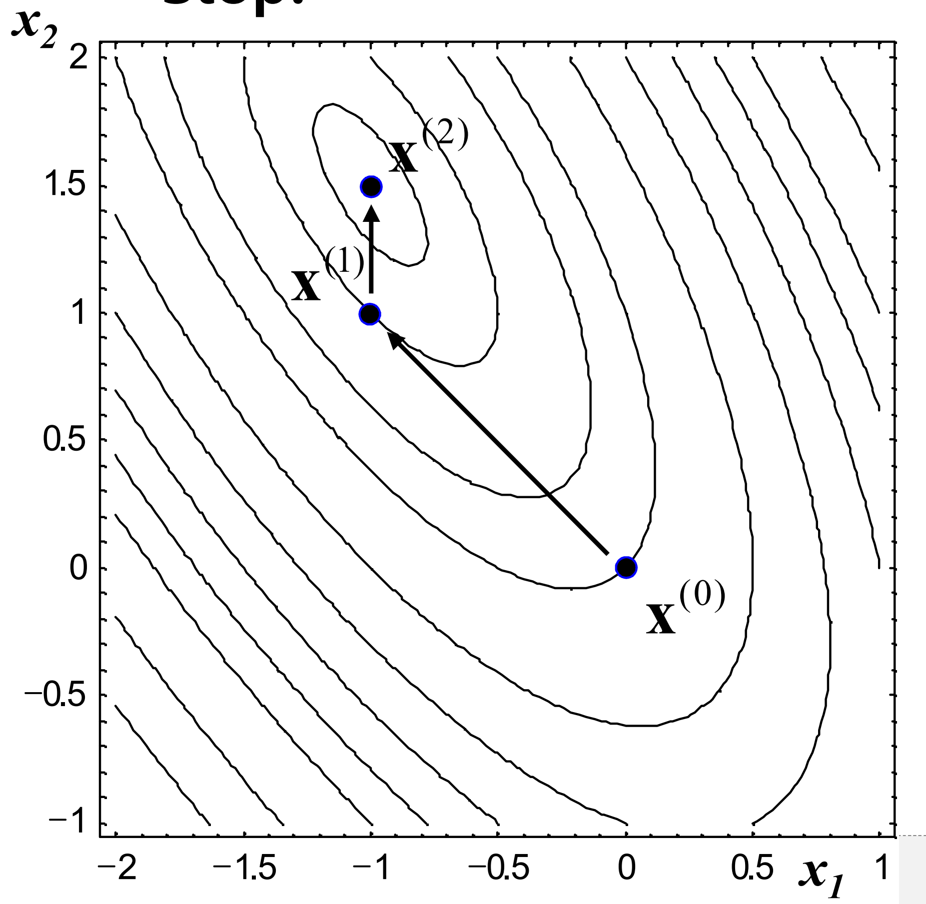
$$\therefore \mathbf{x}^{(2)} = \begin{pmatrix} -1 \\ 1.5 \end{pmatrix}$$ **→Optimal design point**

$$\mathbf{c}^{(2)} = \nabla f(\mathbf{x}^{(2)}) = \begin{pmatrix} 1+4\cdot(-1)+2\cdot 1.5 \\ -1+2\cdot(-1)+2\cdot 1.5 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Check stopping criterion.

$$\left\| \mathbf{c}^{(2)} \right\| = 0 \le \varepsilon$$

→Stop!



5. Broyden-Fletcher-Goldfarb-Shanno(BFGS) Method (1/6)

☑ This method updates the Hessian matrix rather than its inverse at every iteration.

- Step 1 : Estimate a starting design point $\mathbf{x}^{(0)}$.
  Choose a symmetric positive definite nxn matrix $\tilde{\mathbf{H}}^{(0)}$ as an approximation for the Hessian matrix of the objective function.
  In the absence of more information, let $\tilde{\mathbf{H}}^{(0)} = \mathbf{I}$. Specify a tolerance ε for the stopping criterion. Set $k = 0$, and compute the gradient vector as $\mathbf{c}^{(0)} = \nabla f(\mathbf{x}^{(0)})$.

- Step 2 : Calculate the norm of the gradient vector as $\left\|\mathbf{c}^{(k)}\right\|$.
  If $\left\|\mathbf{c}^{(k)}\right\| < \varepsilon$, then stop the iterative process. Otherwise, continue.
  It is noted that Step 1 and 2 of this method and the steepest descent method are the same.

5. Broyden-Fletcher-Goldfarb-Shanno(BFGS) Method (2/6)

- Step 3 : Solve the linear system of the following equation to obtain the search direction.

$$\boxed{\mathbf{d}^{(k)} = -(\tilde{\mathbf{H}}^{(k)})^{-1}\mathbf{c}^{(k)}}$$

Newton's method

$$\Delta\mathbf{x}^{(k)} = -[\mathbf{H}(\mathbf{x}^{(k)})]^{-1}\nabla f(\mathbf{x}^{(k)})$$

$$\therefore \mathbf{d}^{(k)} = -(\mathbf{H}^{(k)})^{-1}\mathbf{c}^{(k)}$$

This equation looks like $\mathbf{d}^{(k)} = -(\mathbf{H}^{(k)})^{-1}\mathbf{c}^{(k)}$ of the Newton's method, but $\tilde{\mathbf{H}}^{(k)}$ is an approximated Hessian matrix $\mathbf{H}^{(k)}$, comprised of the first order derivatives.

- Step 4 : Compute optimum step size $\alpha_k = \alpha$ to minimize $f(\mathbf{x}^{(k)} + \alpha\mathbf{d}^{(k)})$.

- Step 5 : Update the design point as $\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \alpha_k\mathbf{d}^{(k)}$.

5. Broyden-Fletcher-Goldfarb-Shanno(BFGS) Method (3/6)

- Step 6 : Update the matrix $\tilde{\mathbf{H}}^{(k)}$- approximation for the Hessian matrix of the objective function - as follows:

$$\tilde{\mathbf{H}}^{(k+1)} = \tilde{\mathbf{H}}^{(k)} + \mathbf{D}^{(k)} + \mathbf{E}^{(k)} \qquad : \quad n \times n \text{ matrix}$$

where, the correction matrices $D^{(k)}$ and $E^{(k)}$ are given as below.

$$\mathbf{D}^{(k)} = \frac{\mathbf{y}^{(k)}\mathbf{y}^{(k)^T}}{(\mathbf{y}^{(k)} \cdot \mathbf{s}^{(k)})}; \qquad \mathbf{E}^{(k)} = \frac{\mathbf{c}^{(k)}\mathbf{c}^{(k)^T}}{(\mathbf{c}^{(k)} \cdot \mathbf{d}^{(k)})};$$

$$\mathbf{s}^{(k)} = \alpha_k \mathbf{d}^{(k)}$$: change in design

$$\mathbf{y}^{(k)} = \mathbf{c}^{(k+1)} - \mathbf{c}^{(k)}$$: change in gradient

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

$\mathbf{d}^{(k)}$: search direction

$\alpha^{(k)}$: optimum step size

- Step 7 : Set $k = k + 1$ and go to Step 2.

3.1 Gradient Method
5. Broyden-Fletcher-Goldfarb-Shanno(BFGS) Method (4/6): Example

Minimize $f(\mathbf{x}) = f(x_1, x_2) = x_1 - x_2 + 2x_1^2 + 2x_1x_2 + x_2^2$, **Starting design point** $\mathbf{x}^{(0)} = (0, 0)$

$$\nabla f(\mathbf{x}) = \nabla f(x_1, x_2) = \begin{pmatrix} 1 + 4x_1 + 2x_2 \\ -1 + 2x_1 + 2x_2 \end{pmatrix}$$

■ 1st Iteration: Find $\mathbf{x}^{(1)}$

$$\mathbf{x}^{(0)} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad \tilde{\mathbf{H}}^{(0)} = \mathbf{I}$$

$$\mathbf{c}^{(0)} = \nabla f(\mathbf{x}^{(0)}) = \begin{pmatrix} 1 + 4 \cdot 0 + 2 \cdot 0 \\ -1 + 2 \cdot 0 + 2 \cdot 0 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

Check stopping criterion.

$$\left\| \mathbf{c}^{(0)} \right\| = \sqrt{1^2 + (-1)^2} = \sqrt{2} > \varepsilon$$

$$\mathbf{d}^{(0)} = -(\tilde{\mathbf{H}}^{(0)})^{-1} \mathbf{c}^{(0)} = -\mathbf{I}\,\mathbf{c}^{(0)} = -\mathbf{c}^{(0)} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

$$\mathbf{x}^{(1)} = \mathbf{x}^{(0)} + \alpha_0 \mathbf{d}^{(0)}$$

$$= \begin{pmatrix} 0 \\ 0 \end{pmatrix} + \alpha \begin{pmatrix} -1 \\ 1 \end{pmatrix} = \begin{pmatrix} -\alpha \\ \alpha \end{pmatrix}$$

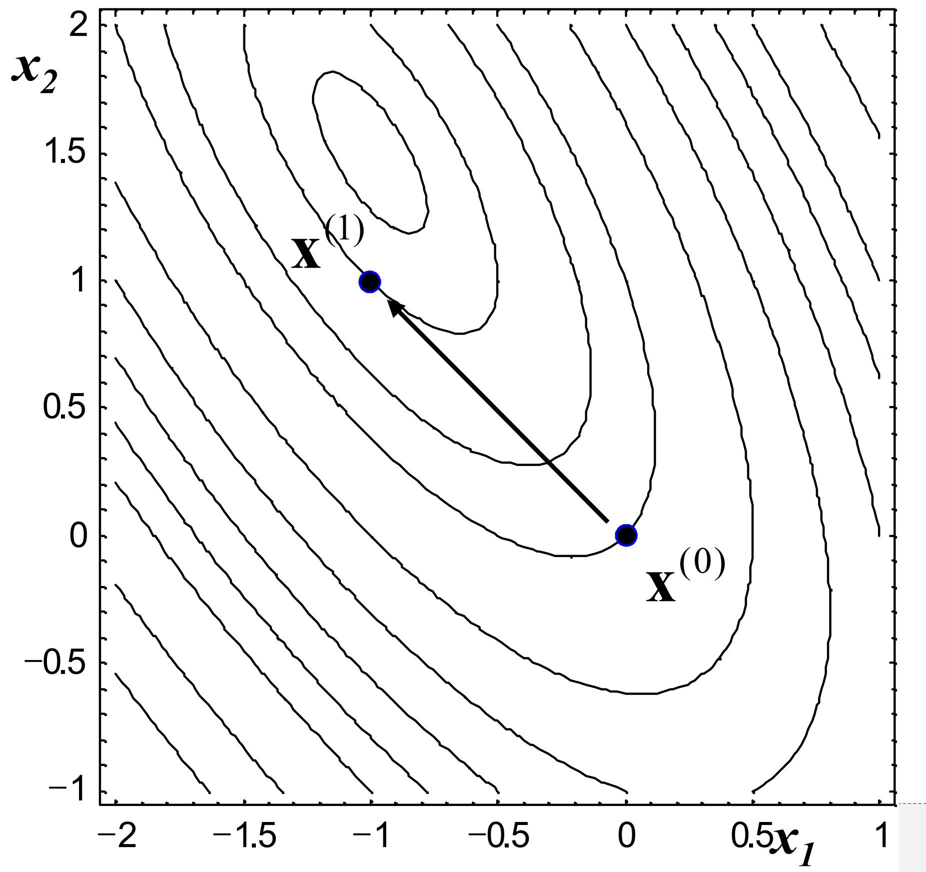
Replacing α_0 **to** α **for convenience**

Substitute $\mathbf{x}^{(1)} = (-\alpha, \alpha)$ **into the objective function**

$$f(\mathbf{x}^{(1)}) = \alpha^2 - 2\alpha$$

To minimize $f(\mathbf{x}^{(1)})$,

$$\frac{df(\mathbf{x}^{(1)})}{d\alpha} = 2\alpha - 2 = 0 \;\rightarrow\; \alpha = 1.0 \qquad \therefore \mathbf{x}^{(1)} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$



5. Broyden-Fletcher-Goldfarb-Shanno(BFGS) Method (5/6): Example

$$\nabla f(\mathbf{x}) = \nabla f(x_1, x_2) = \begin{pmatrix} 1+4x_1+2x_2 \\ -1+2x_1+2x_2 \end{pmatrix}$$

Minimize $f(\mathbf{x}) = f(x_1, x_2) = x_1 - x_2 + 2x_1^2 + 2x_1x_2 + x_2^2,$ **Starting design point $\mathbf{x}^{(0)} = (0, 0)$**

■ 2nd Iteration: Find $\mathbf{x}^{(2)}$

Update the matrix $\tilde{\mathbf{H}}^{(0)}$- approximation for the Hessian matrix of the objective function - as follows:

$$\tilde{\mathbf{H}}^{(1)} = \tilde{\mathbf{H}}^{(0)} + \mathbf{D}^{(0)} + \mathbf{E}^{(0)}$$

$$\mathbf{D}^{(0)} = \frac{\mathbf{y}^{(0)}\mathbf{y}^{(0)^T}}{\mathbf{y}^{(0)} \cdot \mathbf{s}^{(0)}}$$

$$\mathbf{s}^{(0)} = \alpha\mathbf{d}^{(0)} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

$$\mathbf{c}^{(0)} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad \mathbf{c}^{(1)} = \begin{pmatrix} -1 \\ -1 \end{pmatrix}$$

$$\mathbf{y}^{(0)} = \mathbf{c}^{(1)} - \mathbf{c}^{(0)} = \begin{pmatrix} -2 \\ 0 \end{pmatrix}$$

$$\mathbf{y}^{(0)}\mathbf{y}^{(0)^T} = \begin{pmatrix} 4 & 0 \\ 0 & 0 \end{pmatrix}$$

$$\mathbf{y}^{(0)} \cdot \mathbf{s}^{(0)} = 2$$

$$= \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}$$

$$\mathbf{E}^{(0)} = \frac{-\mathbf{c}^{(0)}\mathbf{c}^{(0)T}}{\mathbf{c}^{(0)} \cdot \mathbf{d}^{(0)}}$$

$$\mathbf{c}^{(0)} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad \mathbf{d}^{(0)} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

$$\mathbf{c}^{(0)}\mathbf{c}^{(0)^T} = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$$

$$\mathbf{c}^{(0)} \cdot \mathbf{d}^{(0)} = -2$$

$$= \begin{pmatrix} -0.5 & 0.5 \\ 0.5 & -0.5 \end{pmatrix}$$

$$\begin{aligned} \tilde{\mathbf{H}}^{(1)} &= \tilde{\mathbf{H}}^{(0)} + \mathbf{D}^{(0)} + \mathbf{E}^{(0)} \\ &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} -0.5 & 0.5 \\ 0.5 & -0.5 \end{pmatrix} \\ &= \begin{pmatrix} 2.5 & 0.5 \\ 0.5 & 0.5 \end{pmatrix} \end{aligned}$$

$\nabla f(\mathbf{x}) = \nabla f(x_1, x_2) = \begin{pmatrix} 1+4x_1+2x_2 \\ -1+2x_1+2x_2 \end{pmatrix}$

5. Broyden-Fletcher-Goldfarb-Shanno(BFGS) Method (6/6): Example

Minimize $f(\mathbf{x}) = f(x_1, x_2) = x_1 - x_2 + 2x_1^2 + 2x_1x_2 + x_2^2$, Starting design point $\mathbf{x}^{(0)} = (0, 0)$

■ 2nd Iteration: Find $\mathbf{x}^{(2)}$

Check stopping criterion.

$$\left\|\mathbf{c}^{(1)}\right\| = \sqrt{2} > \varepsilon$$

$$\tilde{\mathbf{H}}^{(1)}\mathbf{d}^{(1)} = -\mathbf{c}^{(1)} \quad \mathbf{d}^{(1)} = -(\tilde{\mathbf{H}}^{(1)})^{-1}\mathbf{c}^{(1)}$$

$$\tilde{\mathbf{H}}^{(1)} = \begin{pmatrix} 2.5 & 0.5 \\ 0.5 & 0.5 \end{pmatrix}, \quad \mathbf{d}^{(1)} = \begin{pmatrix} 0 \\ 2 \end{pmatrix} \quad \mathbf{c}^{(1)} = \begin{pmatrix} -1 \\ -1 \end{pmatrix}$$

$$\mathbf{x}^{(2)} = \mathbf{x}^{(1)} + \alpha_1 \mathbf{d}^{(1)}$$

$$= \begin{pmatrix} -1 \\ 1 \end{pmatrix} + \alpha \begin{pmatrix} 0 \\ 2 \end{pmatrix} = \begin{pmatrix} -1 \\ 1+2\alpha \end{pmatrix}$$

Replacing α_1 to α for convenience

Substitute $\mathbf{x}^{(2)} = (-1, 1+2\alpha)$ into the objective function

$$f(\mathbf{x}^{(2)}) = 4\alpha^2 - 2\alpha - 1$$

To minimize $f(\mathbf{x}^{(2)})$,

$$\frac{df(\mathbf{x}^{(2)})}{d\alpha} = 8\alpha - 2 = 0 \;\rightarrow\; \alpha = 0.25$$

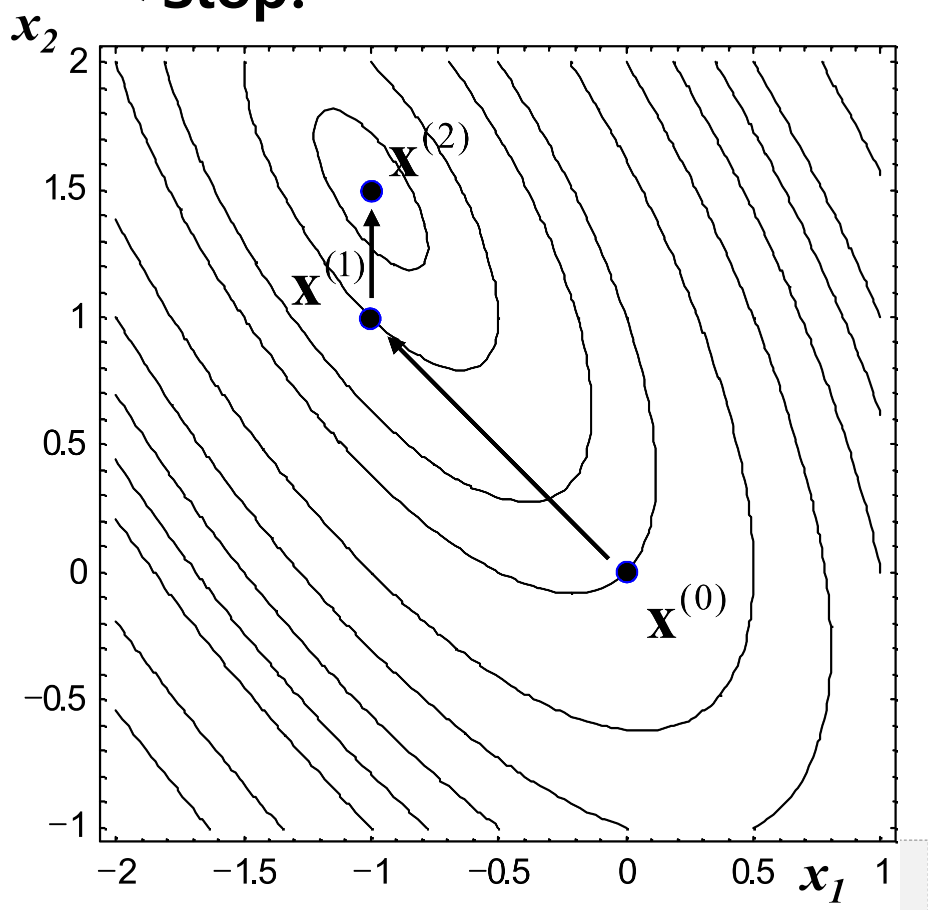
$$\therefore \mathbf{x}^{(2)} = \begin{pmatrix} -1 \\ 1.5 \end{pmatrix}$$ **→Optimal design point**

$$\mathbf{c}^{(2)} = \nabla f(\mathbf{x}^{(2)}) = \begin{pmatrix} 1+4\cdot(-1)+2\cdot 1.5 \\ -1+2\cdot(-1)+2\cdot 1.5 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

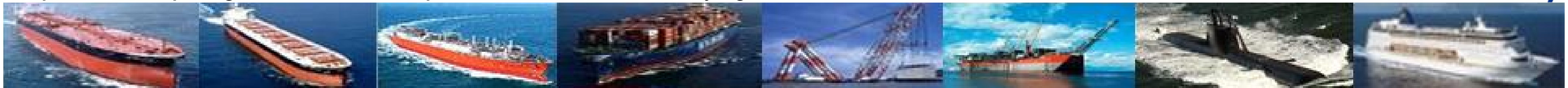
Check stopping criterion.

$$\left\|\mathbf{c}^{(2)}\right\| = 0 \le \varepsilon$$

→Stop!



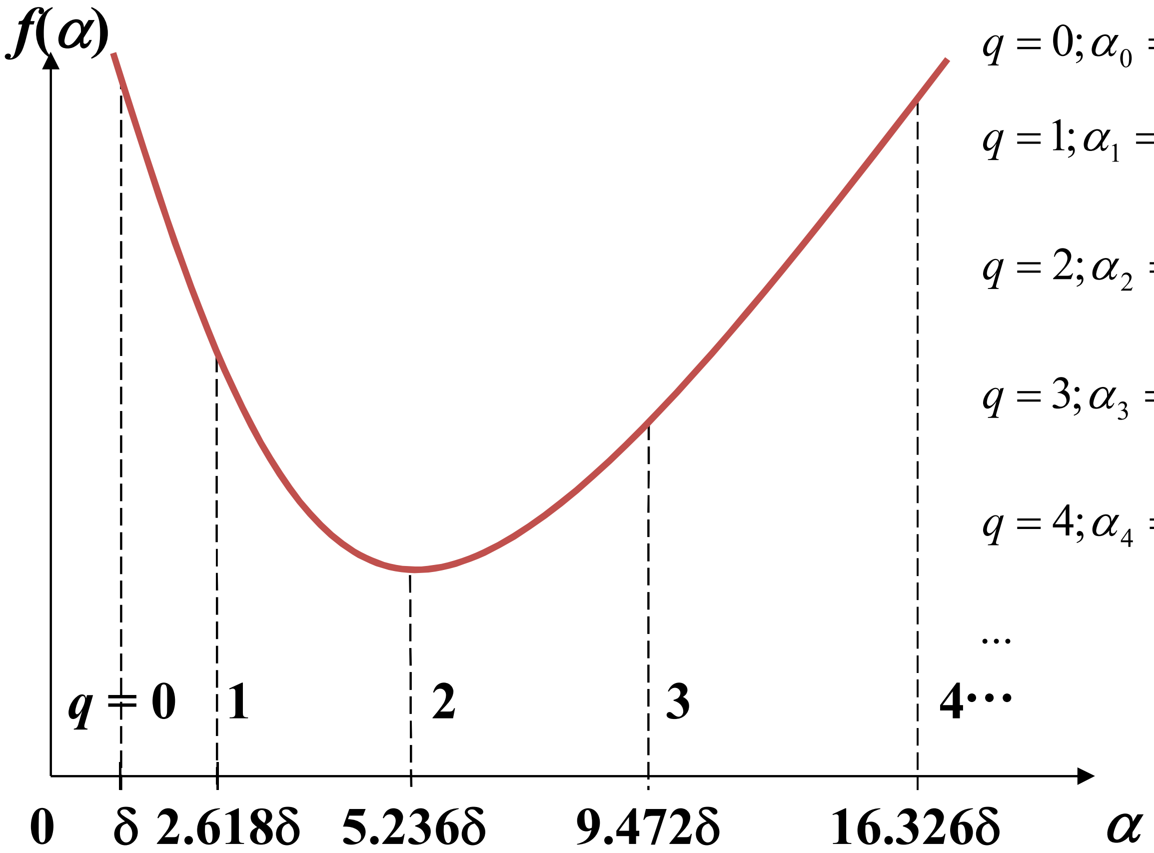
3.2 Golden Section Search Method (One Dimensional Search Method)



3.2 Golden Section Search Method
- Phase 1: Global Search (1/2)

☑ **Searching for the interval in which the minimum lies**

- **In the figure, starting at $q=0$, we evaluate $f(\alpha)$ at $\alpha=\delta$, where $\delta>0$ is a small number. If the value $f(\delta)$ is smaller than the value $f(0)$, we then take an increment of 1.618δ in the step size(i.e., the increment is 1.618 times the previous increment δ). (See Fibonacci sequence)**



$$q=0;\alpha_0=\delta$$

$$q=1;\alpha_1=\delta+1.618\delta=2.618\delta=\sum_{j=0}^{1}\delta(1.618)^j$$

$$q=2;\alpha_2=2.618\delta+1.618(1.618\delta)=5.236\delta=\sum_{j=0}^{2}\delta(1.618)^j$$

$$q=3;\alpha_3=5.236\delta+(1.618)^3\delta=9.472\delta=\sum_{j=0}^{3}\delta(1.618)^j$$

$$q=4;\alpha_4=9.472\delta+(1.618)^4\delta=16.326\delta=\sum_{j=0}^{4}\delta(1.618)^j$$

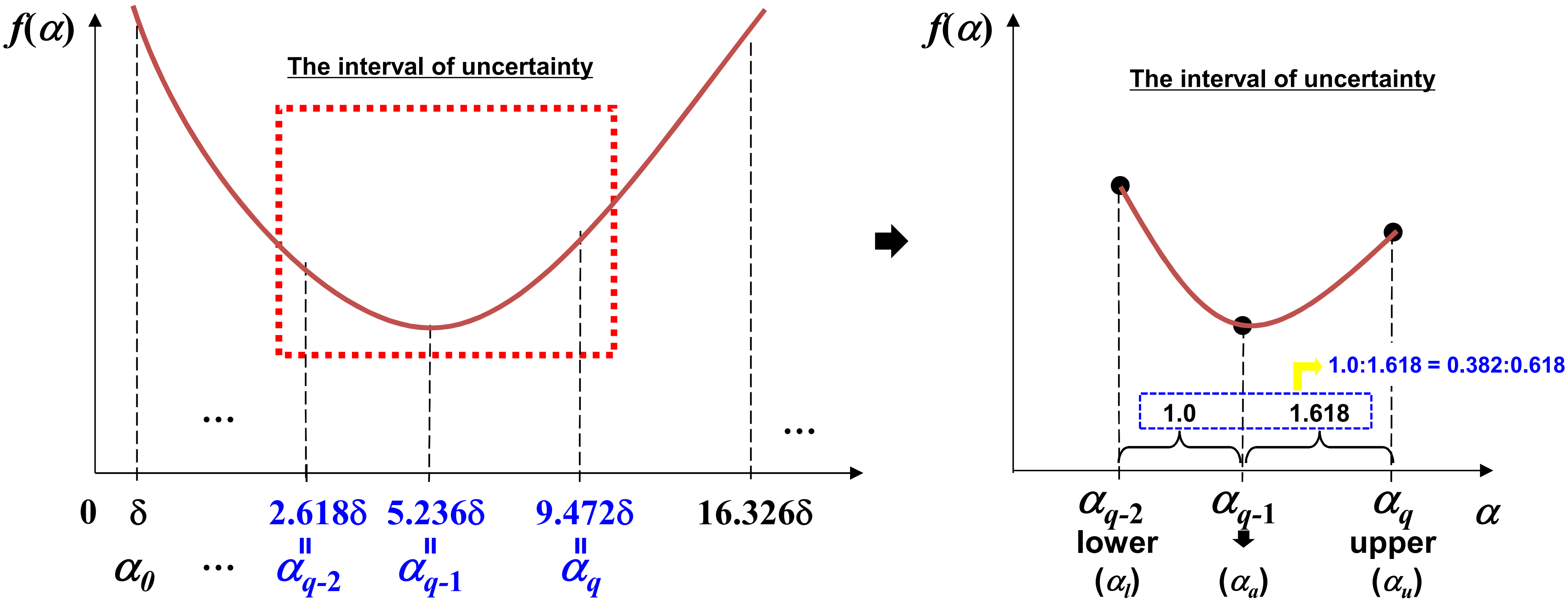
$$\cdots$$

$$\boxed{\therefore\ \alpha_q=\sum_{j=0}^{q}\delta(1.618)^j,\ \ q=0,1,2,\ldots}$$

3.2 Golden Section Search Method - Phase 1: Global Search (2/2)

- If the function at α_{q-1} is smaller than that at the previous point α_{q-2} and the next point α_q, (i.e., $f(\alpha_{q-1}) < f(\alpha_{q-2}),\ f(\alpha_{q-1}) < f(\alpha_q)$) the minimum point lies between α_q and α_{q-2}.

(The interval in which the minimum lies is called <u>the interval of uncertainty</u>.)



- Therefore, upper and lower limits on the interval of uncertainty are

$$\underset{\text{upper}}{\alpha_u} \equiv \alpha_q = \sum_{j=0}^{q} \delta(1.618)^j, \underset{\text{lower}}{\alpha_l} \equiv \alpha_{q-2} = \sum_{j=0}^{q-2} \delta(1.618)^j, \alpha_a \equiv \alpha_{q-1} = \sum_{j=0}^{q-1} \delta(1.618)^j$$

[Reference] Fibonacci Sequence

Fibonacci sequence defined as

$$F_0 = 0; \quad F_1 = 1; \qquad F_n = F_{n-1} + F_{n-2}, \quad n = 2,\ 3,\ \cdots$$

Any number of the Fibonacci sequence for n(>1) is obtained by adding the previous two numbers, so the sequence is given as follows.

➔ 0, 1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, …

General term: $F_n = \dfrac{\varphi^n - (1-\varphi)^n}{\sqrt{5}}, \quad \varphi = \dfrac{1+\sqrt{5}}{2} \approx 1.6180339887\cdots$

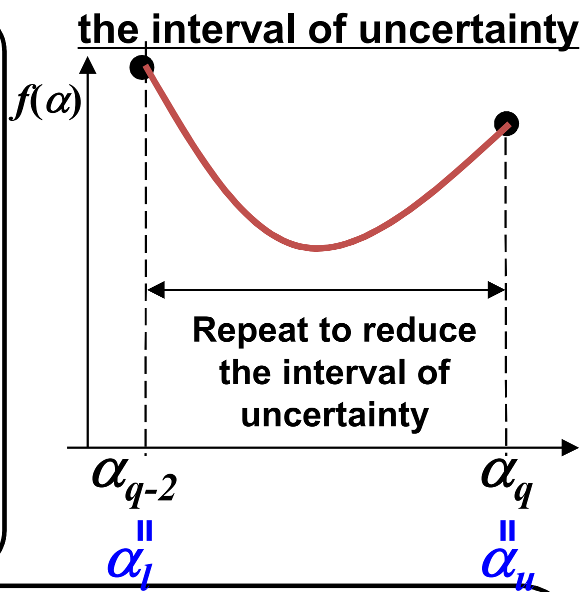
Property: $\lim\limits_{n\to\infty} \dfrac{F_n}{F_{n-1}} = \varphi, \quad 1-\varphi = -\dfrac{1}{\varphi}$

$$\left(\because \frac{1-\varphi}{\varphi} < 1 \right)$$

$$\lim_{n\to\infty} \frac{F_n}{F_{n-1}} = \lim_{n\to\infty} \frac{\varphi^n - (1-\varphi)^n}{\varphi^{n-1} - (1-\varphi)^{n-1}} = \lim_{n\to\infty} \frac{\dfrac{\varphi^n - (1-\varphi)^n}{\varphi^{n-1}}}{\dfrac{\varphi^{n-1} - (1-\varphi)^{n-1}}{\varphi^{n-1}}} = \lim_{n\to\infty} \frac{\varphi - (1-\varphi)\cancel{\left(\dfrac{1-\varphi}{\varphi}\right)^{n-1}}}{1 - \cancel{\left(\dfrac{1-\varphi}{\varphi}\right)^{n-1}}} = \varphi$$

3.2 Golden Section Search Method
- Phase 2: Local Search (1/3)

- **Reduction of the interval of uncertainty** by comparing function values at α_a and α_b
  - We consider two points symmetrically located from either end as shown in the figure – points α_a and α_b are located at a distance of $\tau\, I^{(k)}$ from either end of the interval.
  - Comparing function values at α_a and α_b, either the left $(\alpha_l\,,\ \alpha_a)$ or the right $(\alpha_b\,,\ \alpha_u)$ portion of the interval gets discarded because the minimum cannot lie there.

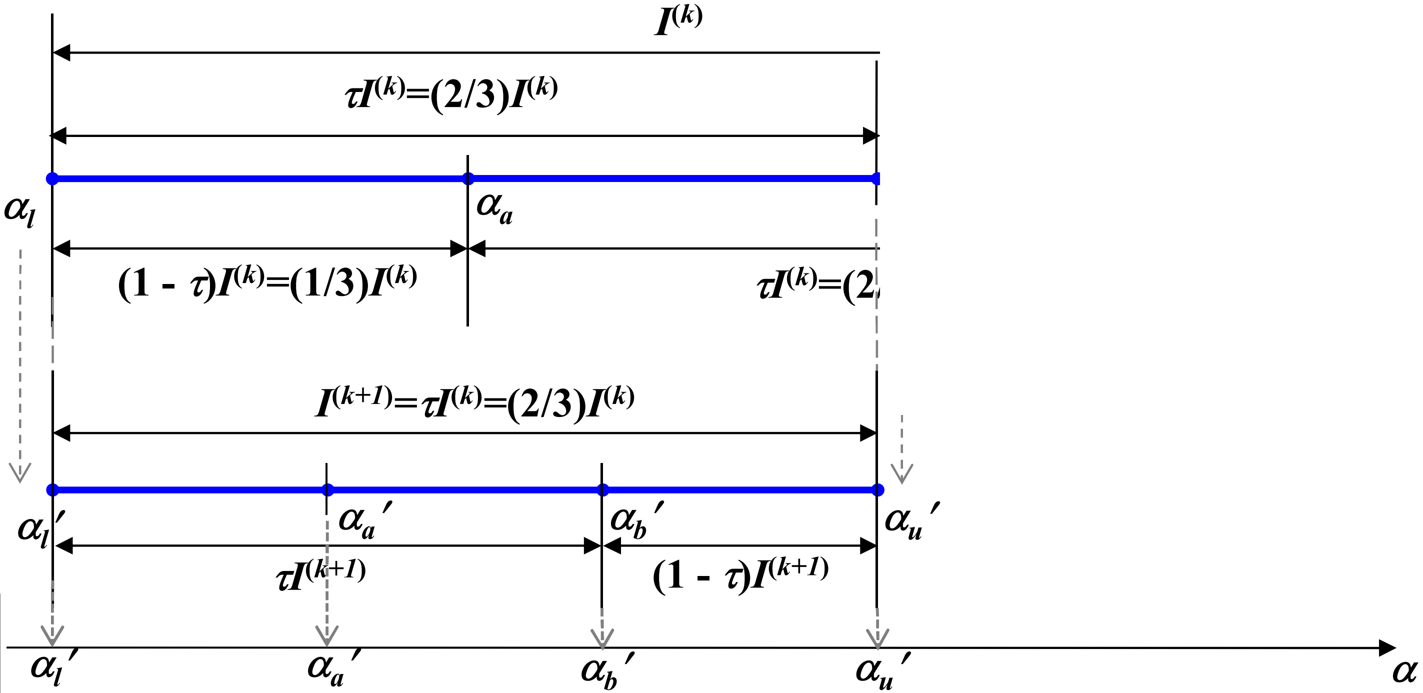


< If τ = 2/3 >

(a)

(b)

If $f(\alpha_a) < f(\alpha_b)$, then minimum point lies between α_l and α_b.



For new interval of uncertainty, we always have to compute $f(\alpha_a')$, $f(\alpha_b')$.

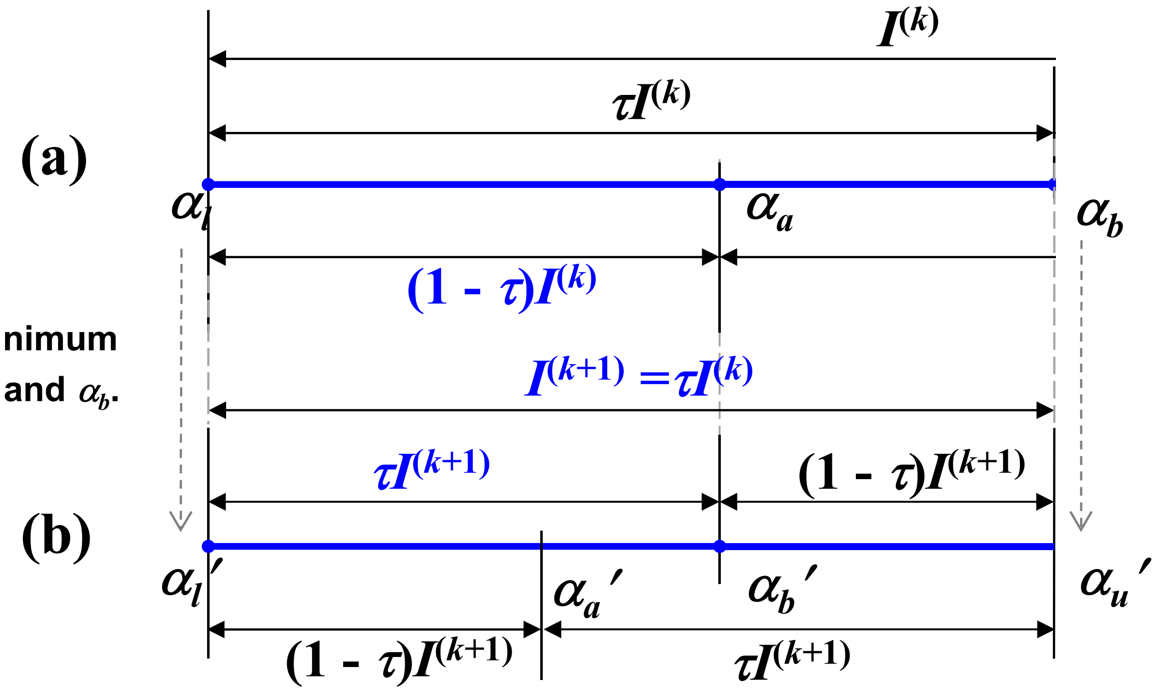
<Question>

Is there any method to use the previous function values?

3.2 Golden Section Search Method
- Phase 2: Local Search (2/3)

- **Reduction of the interval of uncertainty by comparing function values at α_a and α_b**
  - **We consider two points symmetrically located from either end as shown in the figure – points α_a and α_b are located at a distance of $\tau I^{(k)}$ from either end of the interval.**

If $f(\alpha_a) < f(\alpha_b)$, then minimum point lies between α_l and α_b.



1. $f(\alpha_a)$ will be used for the next interval of uncertainty $I^{(k+1)}$.
2. α_a can be equal to α_a' or α_b' of the next interval of uncertainty $I^{(k+1)}$.

3-1. Assume that α_a is equal to α_a'.

$$\alpha_a = \alpha_a'$$
$$(1-\tau)I^{(k)} = (1-\tau)I^{(k+1)}$$
$$(1-\tau)I^{(k)} = (1-\tau)\tau I^{(k)}$$
$$I^{(k)} = \tau I^{(k)}$$

Because τ=1, this assumption is wrong.

3-2. Assume that α_a is equal to α_b'.

$$\alpha_a = \alpha_b'$$
$$(1-\tau)I^{(k)} = \tau I^{(k+1)}$$
$$(1-\tau)I^{(k)} = \tau \cdot \tau \cdot I^{(k)}$$
$$\tau \cdot \tau I^{(k)} - (1-\tau)I^{(k)} = 0$$
$$\tau^2 + \tau - 1 = 0$$
$$\Rightarrow \tau = 0.618, -1.618 \Rightarrow 0.618$$

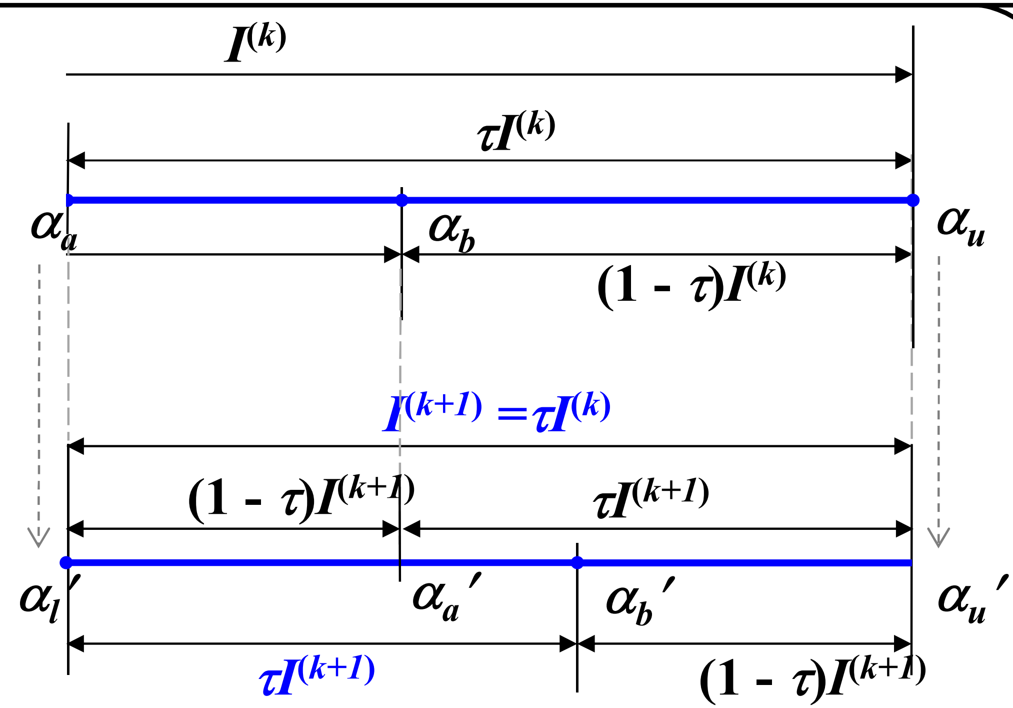
3.2 Golden Section Search Method
- Phase 2: Local Search (3/3)

- **Reduction of the interval of uncertainty by comparing function values at α_a and α_b**
  - **We consider two points symmetrically located from either end as shown in the figure – points α_a and α_b are located at a distance of $\tau I^{(k)}$ from either end of the interval.**

(a)

If $f(\alpha_a) > f(\alpha_b)$, then minimum point lies between α_a and α_u.

(b)



1. $f(\alpha_b)$ will be used for the next interval of uncertainty $I^{(k+1)}$.
2. α_b can be equal to α_a' or α_b' of the next interval of uncertainty $I^{(k+1)}$.

3-1. Assume that α_b is equal to α_b'.

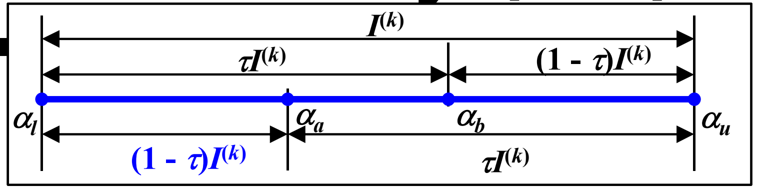
$$\alpha_b = \alpha_b'$$
$$(1-\tau)I^{(k)} = (1-\tau)I^{(k+1)}$$
$$(1-\tau)I^{(k)} = (1-\tau)\tau I^{(k)}$$
$$I^{(k)} = \tau I^{(k)}$$

Because τ=1, this assumption is wrong.

3-2. Assume that α_b is equal to α_a'.

$$\alpha_b = \alpha_a'$$
$$(1-\tau)I^{(k)} = \tau I^{(k+1)}$$
$$(1-\tau)I^{(k)} = \tau \cdot \tau \cdot I^{(k)}$$
$$\tau \cdot \tau I^{(k)} - (1-\tau)I^{(k)} = 0$$
$$\tau^2 + \tau - 1 = 0$$
$$\Rightarrow \tau = 0.618, -1.618 \Rightarrow 0.618$$

3.2 Golden Section Search Method: Summary (1/3)



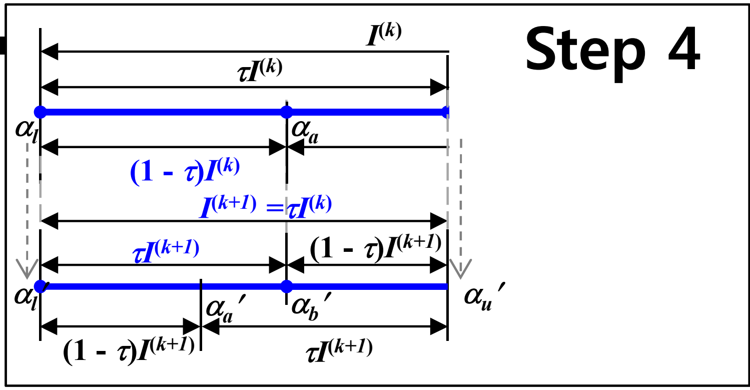
- **Step 1: For a chosen small number δ, let q be the smallest integer to satisfy $f(\alpha_{q-1}) < f(\alpha_{q-2}), f(\alpha_{q-1}) < f(\alpha_q)$ where α_q, α_{q-1}, and α_{q-2} are calculated from $\alpha_q = \sum_{j=0}^{q} \delta(1.618)^j, \ (q = 0,1,2,\ldots)$. The upper and lower bounds on α^* (the optimum value for α) are given as follows.**

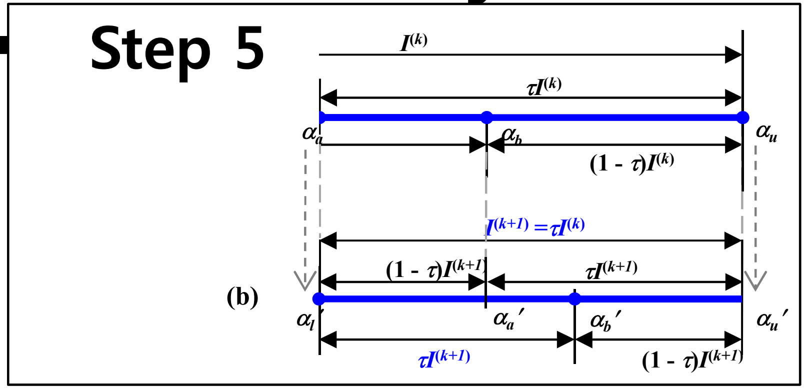
$$\alpha_u \equiv \alpha_q = \sum_{j=0}^{q} \delta(1.618)^j, \alpha_l \equiv \alpha_{q-2} = \sum_{j=0}^{q-2} \delta(1.618)^j$$

- **Step 2 : Compute $f(\alpha_a)$ and $f(\alpha_b)$ where $\alpha_a = \alpha_l + 0.382I$ and $\alpha_b = \alpha_l + 0.618I$ (interval of uncertainty $I = \alpha_u - \alpha_l$).**

- **Step 3 : Compute $f(\alpha_a)$ and $f(\alpha_b)$, and go to Step 4, Step 5 or Step 6.**

3.2 Golden Section Search Method: Summary (2/3)



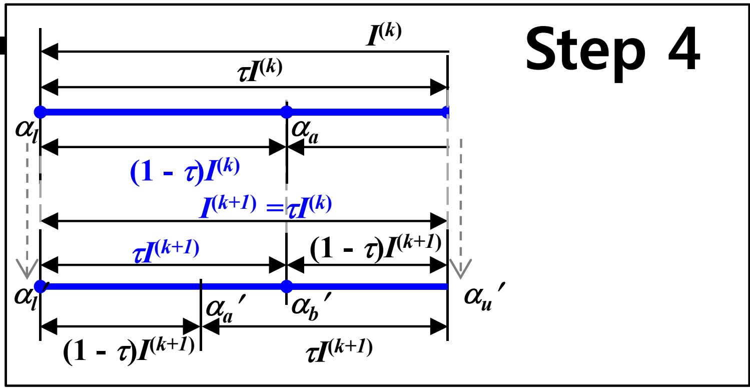


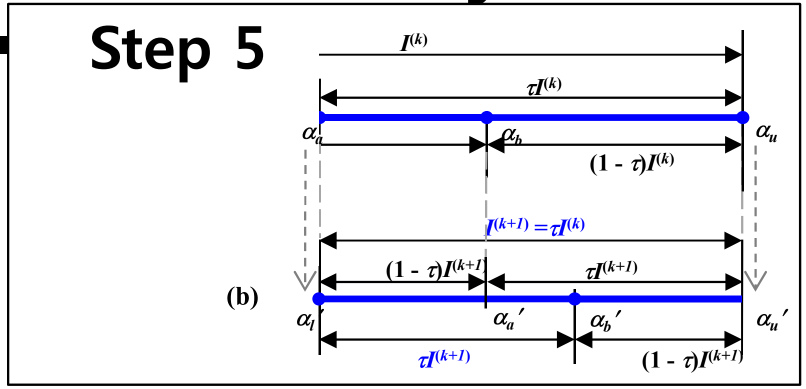
- **Step 4 : If $f(\alpha_a) < f(\alpha_b)$, then minimum point α^* lies between α_l and α_b, i.e., $\alpha_l \le \alpha^* \le \alpha_b$. The new limits for the reduced interval of uncertainty are $\alpha_l' = \alpha_l$ and $\alpha_u' = \alpha_b$. Also, $\alpha_b' = \alpha_a$. Compute $f(\alpha_a')$, where $\alpha_a' = \alpha_l' + 0.382(\alpha_u' - \alpha_l')$ and go to Step 7.**

- **Step 5 : If $f(\alpha_a) > f(\alpha_b)$, then minimum point α^* lies between α_a and α_u, i.e., $\alpha_a \le \alpha^* \le \alpha_u$. Similar to the procedure in Step 4, let $\alpha_l' = \alpha_a$ and $\alpha_u' = \alpha_u$, so that $\alpha_a' = \alpha_b$. Compute $f(\alpha_b')$, where $\alpha_b' = \alpha_l' + 0.618(\alpha_u' - \alpha_l')$ and go to Step 7.**

- **Step 6 : If $f(\alpha_a) = f(\alpha_b)$, let $\alpha_l = \alpha_a$ and $\alpha_u = \alpha_b$ and return to Step 2.**

3.2 Golden Section Search Method: Summary (3/3)

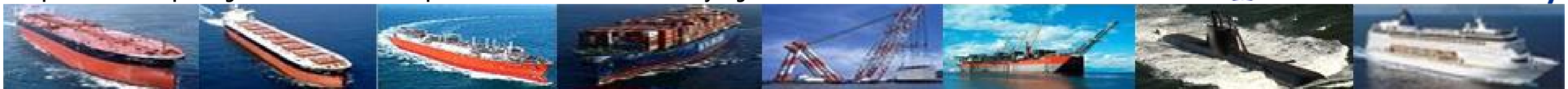




■ **Step 7 : If the new interval of uncertainty $I' = \alpha_u' - \alpha_l'$ is small enough to satisfy a stopping criterion (i.e., $I' < \varepsilon$), let $\alpha^* = (\alpha_u' - \alpha_l')/2$ and stop. Otherwise, delete the primes(') on $\alpha_l', \alpha_a', \alpha_b'$ and α_u' and return to Step 3.**

3.3 Direct Search Method

1. Hooke & Jeeves Direct Search Method
2. Nelder & Mead Simplex Method

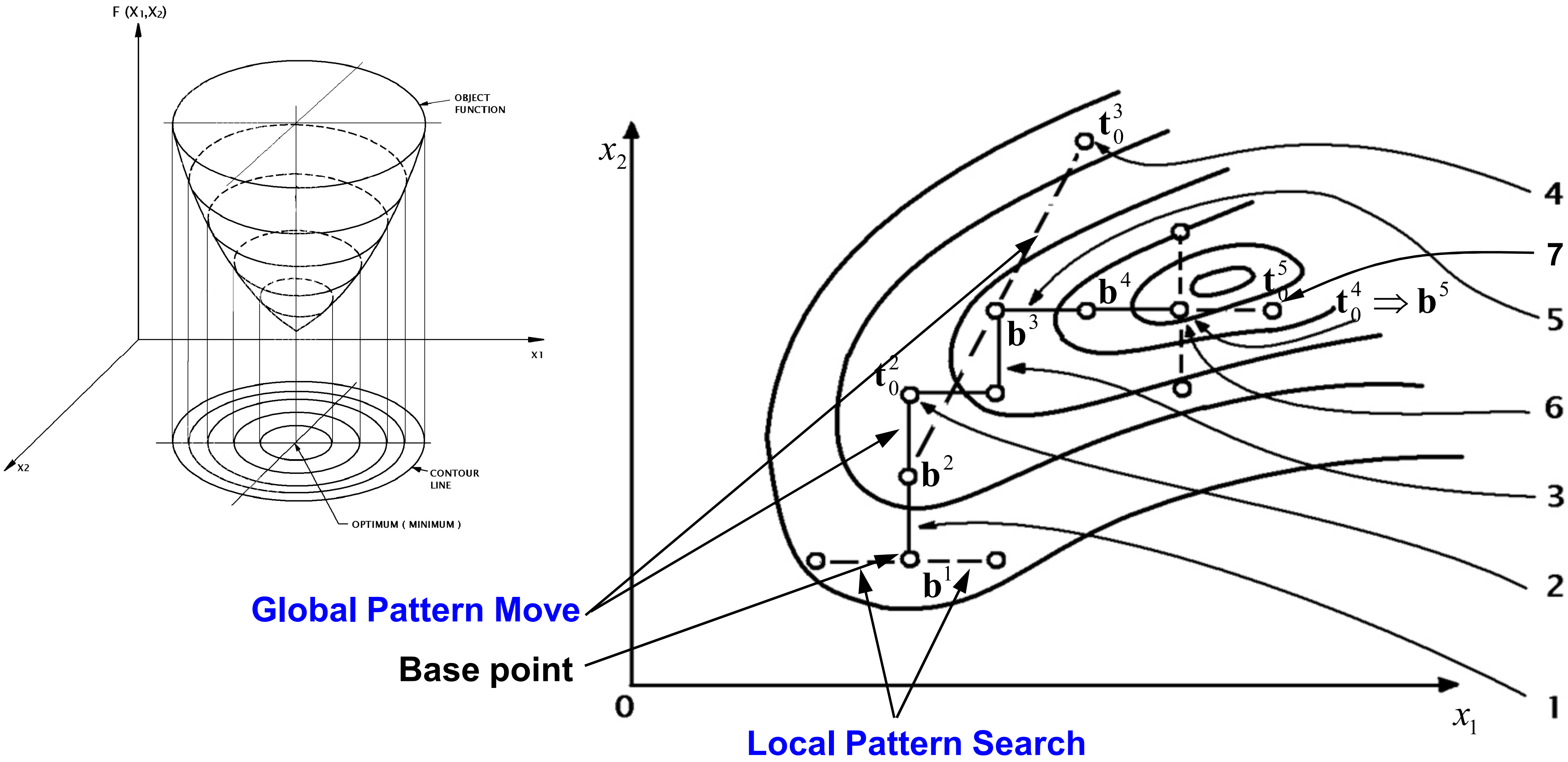


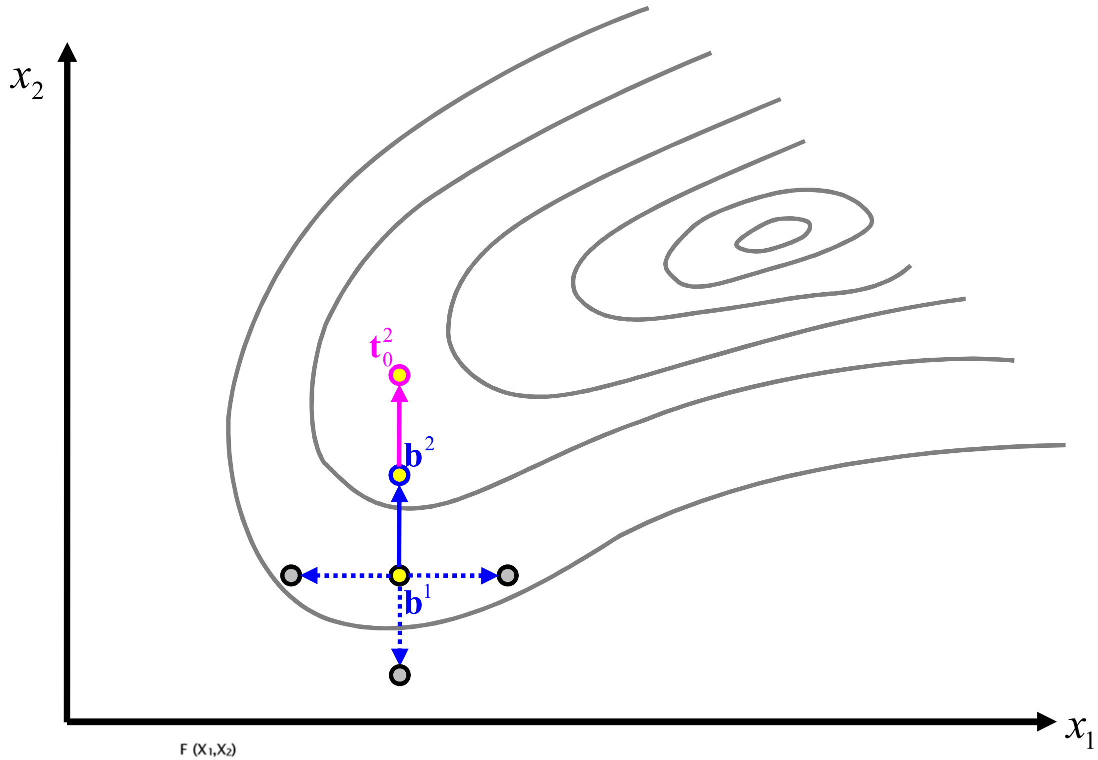
3.3 Direct Search Method

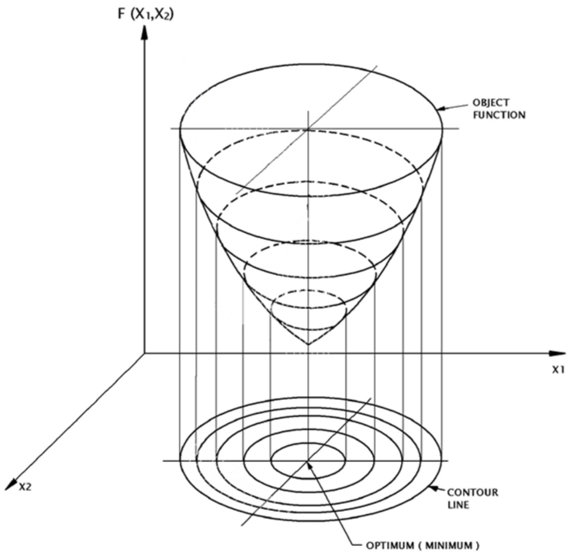
1. Hooke & Jeeves Direct Search Method (1/16)

1. Base Point
2. Global Pattern Move
3. Local Pattern Search

☑ **This method is a sequential technique, each step of which consists of two kinds of move, the 'Local Pattern Search' at a base point and 'Global Pattern Move' to the optimal design point.**





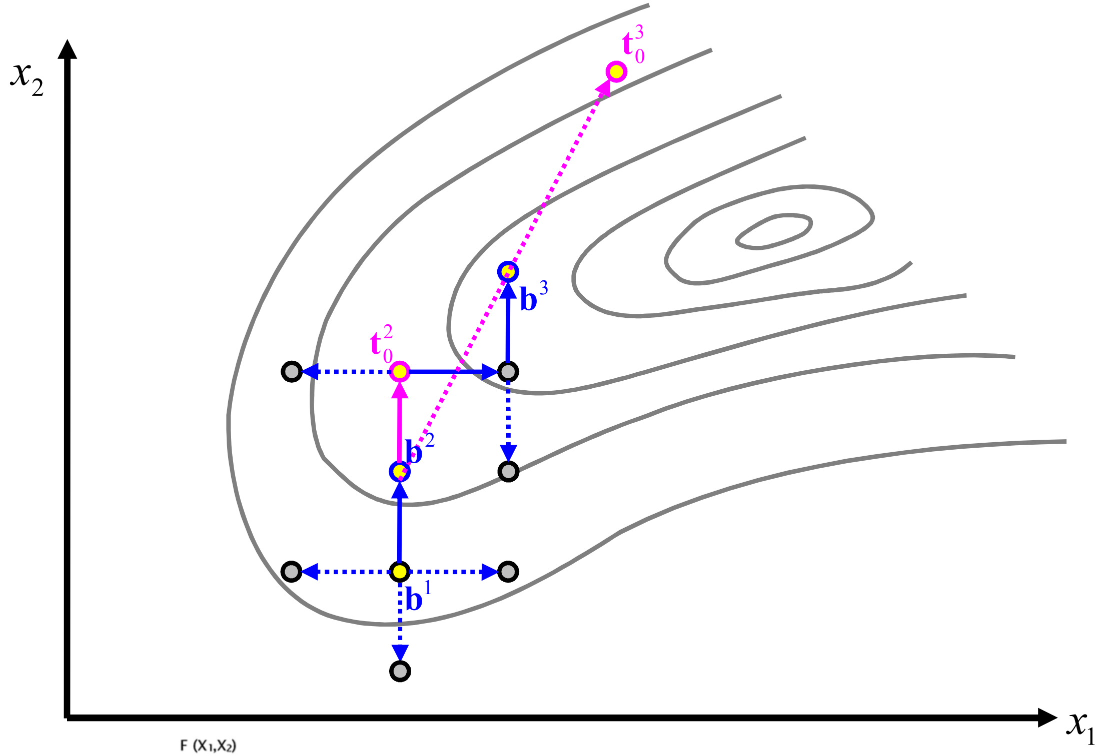


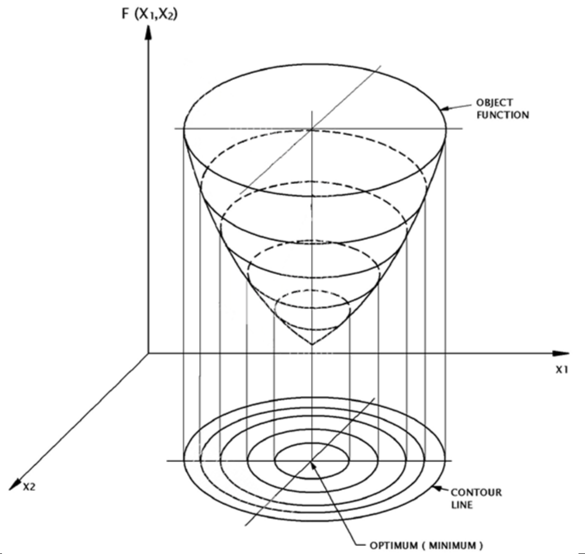
1. 'Local Pattern Search' at the base point b^1

- Search in x_1 direction.
  - No improvement of the value of the objective function in x_1 direction → No movement in x_1 direction
- Search in x_2 direction.
  - Improvement of the value of the objective function in x_2 direction → Movement in the positive x_2 direction
- Move and define the base point b^2.

2. 'Global Pattern Move' at the base point b^2

- Find a temporary base point t_0^2 by symmetrical displacement of b^1 to b^2.
- Because the value of the objective function at t_0^2 is better than that at b^2, perform the 'Local Pattern Search' at t_0^2.





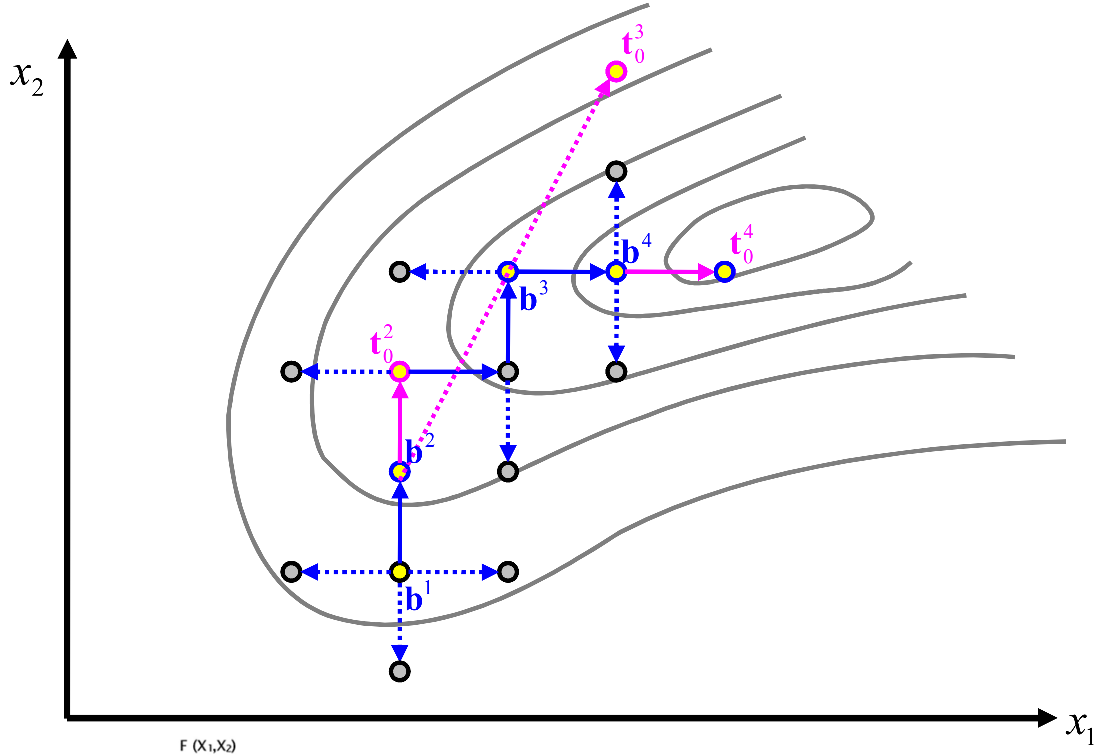
3. 'Local Pattern Search' at the temporary base point t_0^2

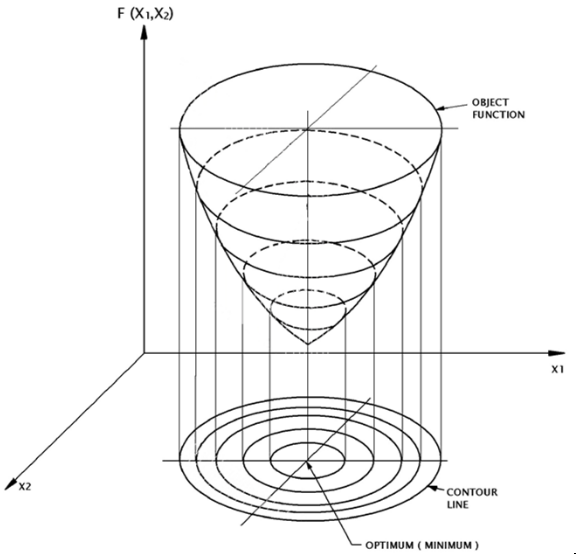
- **Search in x_1 direction.**
  - Improvement of the value of the objective function in x_1 direction → Movement in the positive x_1 direction
- **Search in x2 direction.**
  - Improvement of the value of the objective function in x_2 direction → Movement in the positive x_2 direction
- **Move and define the base point b^3.**

4. 'Global Pattern Move' at the base point b^3

- **Find a temporary base point t_0^3 by symmetrical displacement of b^2 to b^3.**
- **Because the value of the objective function at t_0^3 is not better than that at b^3, perform the 'Local Pattern Search' at b^3.**

3.3 Direct Search Method
1. Hooke & Jeeves Method (4/16)



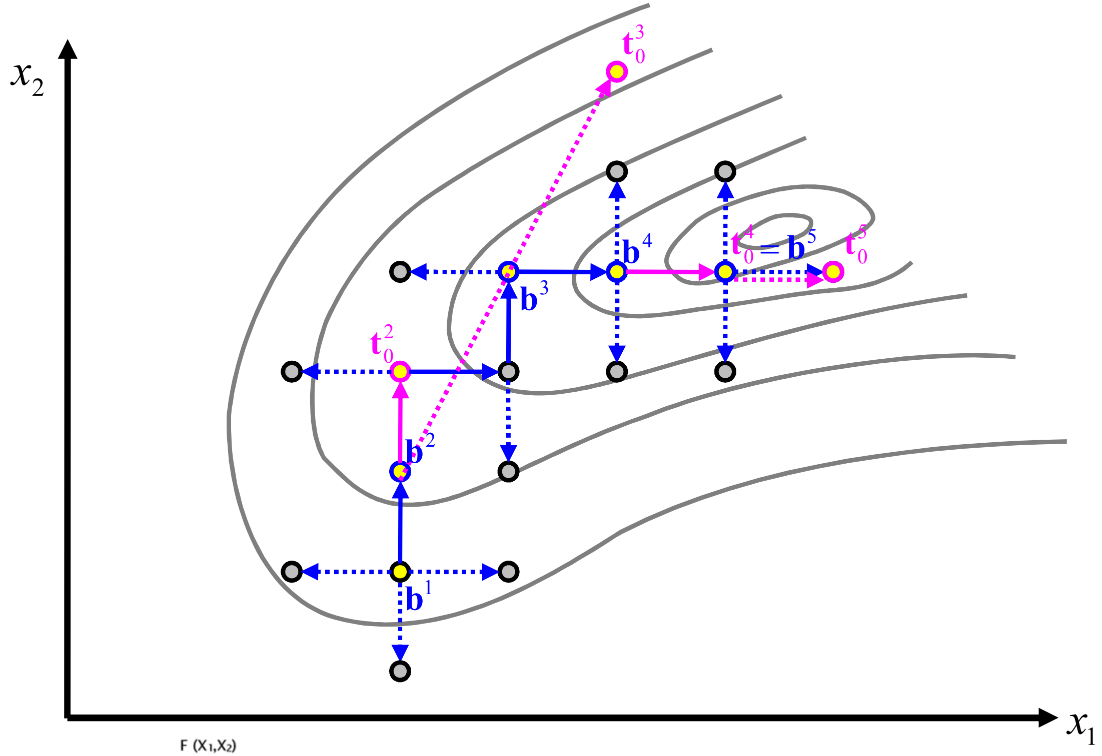


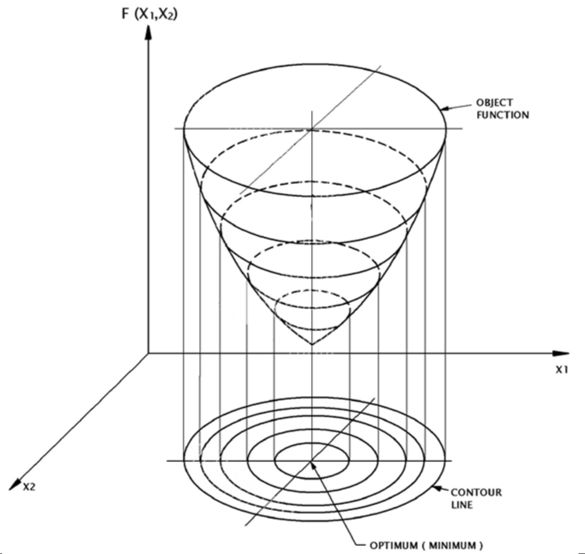
5. 'Local Pattern Search' at the base point b^3

- **Search in x_1 direction.**
  - **Improvement of the value of the objective function in x_1 direction → Movement in the positive x_1 direction**
- **Search in x_2 direction.**
  - **No improvement of the value of the objective function in x_2 direction → No movement in x_2 direction**
- **Move and define the base point b^4.**

6. 'Global Pattern Move' at the base point b^4

- **Find a temporary base point t_0^4 by symmetrical displacement of b^3 to b^4.**
- **Because the value of the objective function at t_0^4 is better than that at b^4, perform the 'Local Pattern Search' at t_0^4 .**





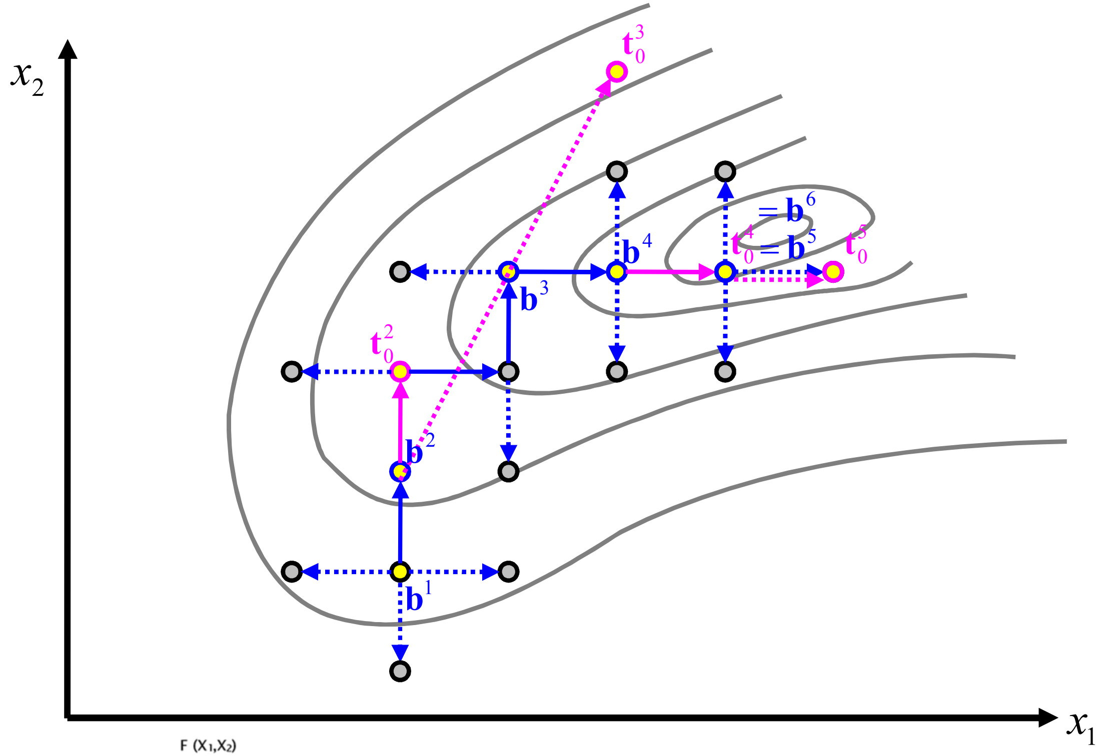
7. 'Local Pattern Search' at the temporary base point t_0^4

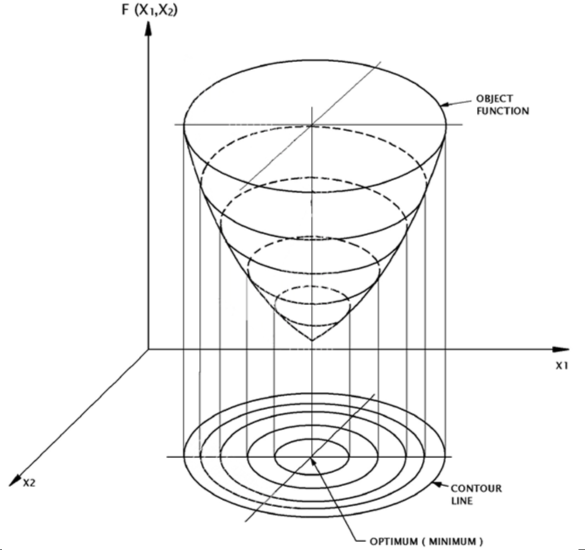
- Search in x_1 direction.
  - No improvement of the value of the objective function in x_1 direction → No movement in x_1 direction
- Search in x_2 direction.
  - No improvement of the value of the objective function in x_2 direction → No movement in x_2 direction
- Because there is no improvement of the value of the objective function in x_1 and x_2 direction, the current base point is defined as the base point b^5.

8. 'Global Pattern Move' at the base point b^5

- Find a temporary base point t_0^5 by symmetrical displacement of b^4 to b^5.
- Because the value of the objective function at t_0^5 is not better than at b^5, perform the 'Local Pattern Search' at b^5.

1. Base Point
2. Global Pattern Move
3. Local Pattern Search





9. 'Local Pattern Search' at the base point b^5

- Search in x_1 direction.
  - No improvement of the value of the objective function in x_1 direction → No movement in x_1 direction

- Search in x_2 direction.
  - No improvement of the value of the objective function in x_2 direction → No movement in x_2 in x_2 direction

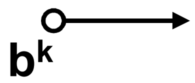
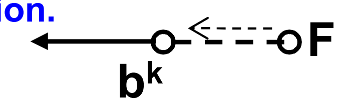
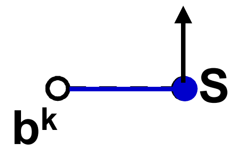
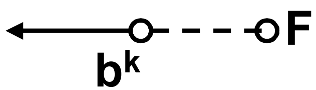
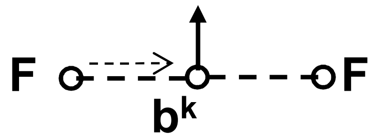
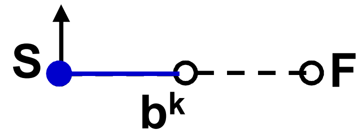
- Because there is no improvement of the value of the objective function in x1 and x2 direction, the current base point defined as base point b^6.

- Because b^5 = b^6, reduce the step size by half and perform the 'Local Pattern Search' at b^6.

3.3 Direct Search Method
1. Hooke & Jeeves Method (7/16): Rule of the 'Local Pattern Search' (1)

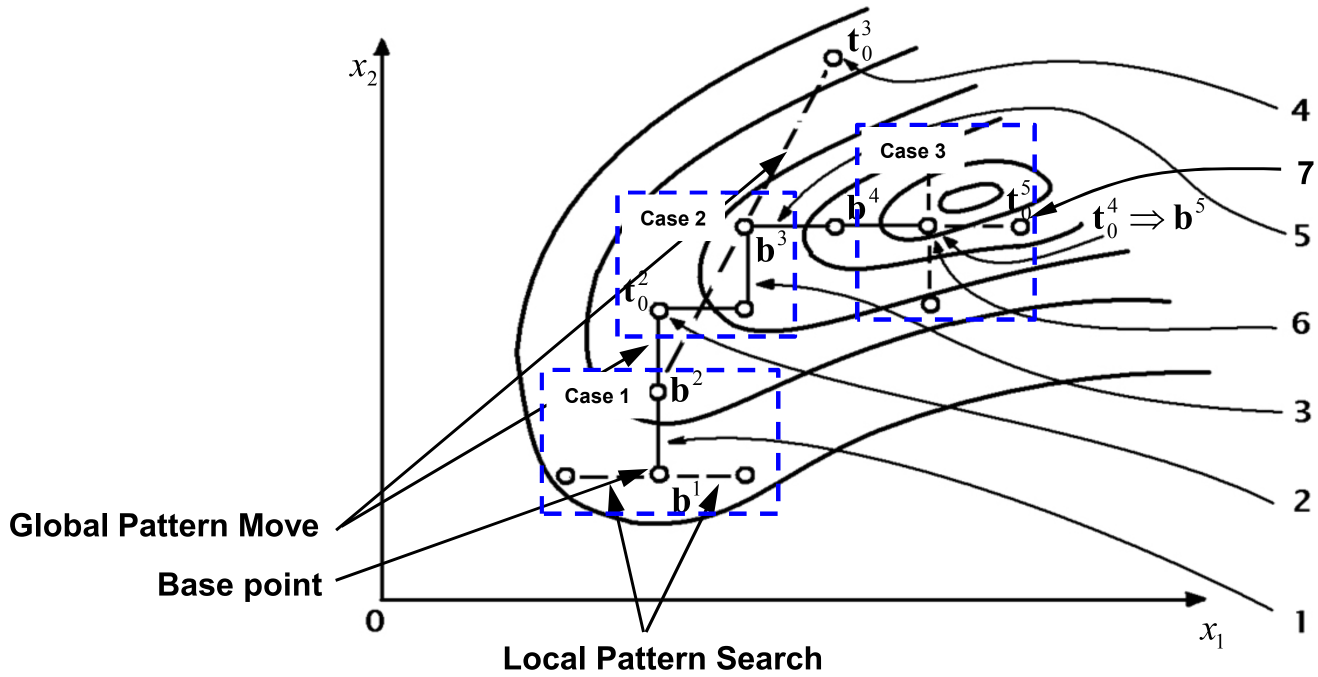
Rule of the 'Local Pattern Search'

(F: Fail, S: Success)

<table>
<tr><td colspan="3">Rule ① Search in the positive x_i direction.</td></tr>
<tr><td rowspan="2">- Move the exploratory point in the positive x_i direction and evaluate the value of the objective function at that point.

</td><td>- If the value of the objective function is increased (Fail)</td><td>- Come back to the previous point and search in the negative x_i direction.

</td></tr>
<tr><td>- If the value of the objective function is decreased (Success)</td><td>- Search in the x_{i+1} direction at the current point.

</td></tr>
<tr><td colspan="3">Rule ② Search in the negative x_i direction.</td></tr>
<tr><td rowspan="2">- If the search in the positive x_i direction is failed, move the exploratory point in the negative x_i direction and evaluate the value of the objective function at that point.

</td><td>- If the value of the objective function is increased (Fail)</td><td>- Come back to the previous point and search in x_{i+1} direction.

</td></tr>
<tr><td>- If the value of the objective function is decreased (Success)</td><td>- Search in the x_{i+1} direction at the current point.

</td></tr>
</table>

- This process of the 'Local Pattern Search' is continued for $i = 1,\dots, n$.
- After searching in x_n direction, the current point is defined as new base point b^{k+1}.

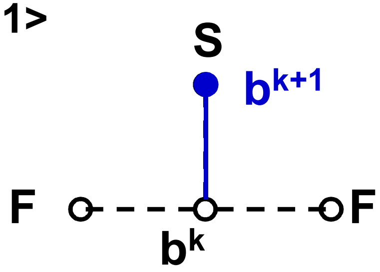
1. Hooke & Jeeves Method (8/16): Rule of the 'Local Pattern Search' (2)



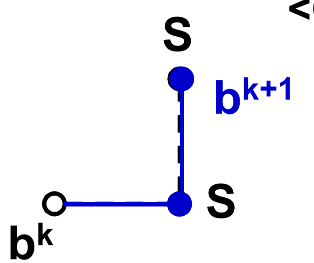
* Super script 'k' means the number of step.

- **Rule of the Local Pattern Search**(F: Fail, S: Success)

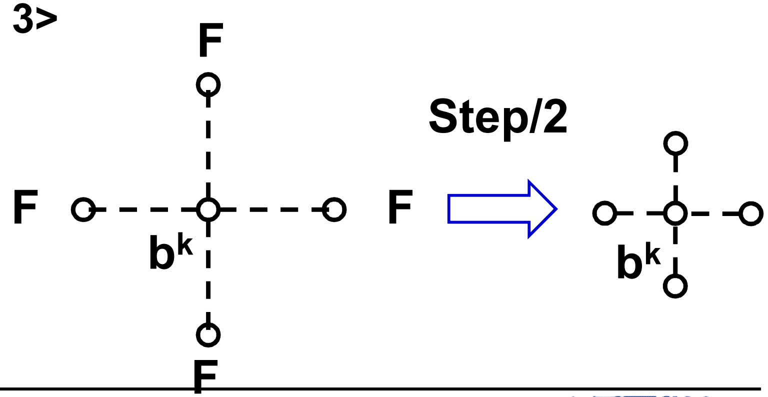
<Case 1>



<Case 2>



<Case 3>



3.3 Direct Search Method
1. Hooke & Jeeves Method (9/16): Algorithm Summary (1)

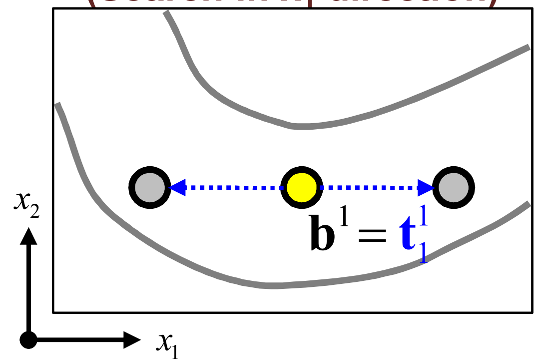
1) Local Pattern Search (Problem with n design variables)

1. Compute the value of the objective function at the starting base point b^1.

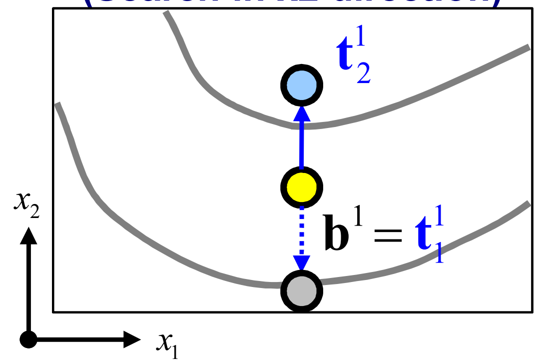
2. Compute the value of the objective function at $b^1 \pm \delta_1$, where δ_1 is input step size and a vector with n elements($\delta_1 = [\delta_1, 0, 0, \ldots, 0]^T$). If the value of the objective function is decreased, $b^1 \pm \delta_1$ is adopted as t_1^1 and the search is continued.

3. Compute the value of the objective function at $t_1^1 \pm \delta_2$, where δ_2 is also input step size and a vector with n elements($\delta_2 = [0, \delta_2, 0, \ldots, 0]^T$). If the value of the function is decreased, $t_1^1 \pm \delta_2$ is adopted as t_2^1.

Example of the 'Local Pattern Search' in the problem with two design variables(x_1, x_2)
(Search in x_1 direction)



Example of the 'Local Pattern Search' in the problem with two design variables(x_1, x_2)
(Search in x2 direction)



3.3 Direct Search Method
1. Hooke & Jeeves Method (10/16): Algorithm Summary (2)

1) Local Pattern Search (Problem with n design variables)

4. After the 'Local Pattern Search' for all design variables, new base point is defined. (new base point $b^2 = t_n^1$)

5. Perform the 'Global Pattern Move' from the previous base point along the line from the previous to current base point.

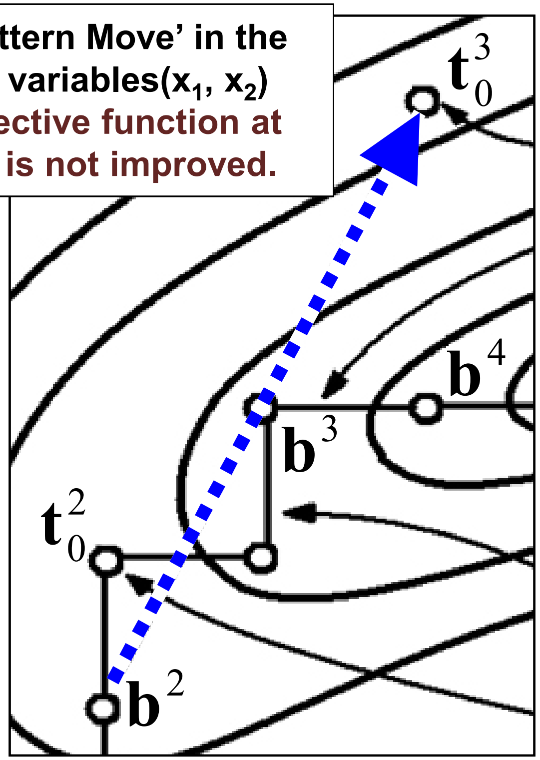
1. Hooke & Jeeves Method (11/16): Algorithm Summary (3)

2) Global Pattern Move

1. **Define the temporary base point located the same distance between the previous and current base point(obtained from 'Local Pattern Search') from the current base point ('Global Pattern Move'), and calculate the value of the objective function at this point. The temporary base point is calculated by 'Global Pattern Move' as follows.**

$$\mathbf{t}_0^{k+1} = \mathbf{b}^k + 2(\mathbf{b}^{k+1} - \mathbf{b}^k) = 2\mathbf{b}^{k+1} - \mathbf{b}^k$$

Example of the 'Global Pattern Move' in the problem with two design variables(x_1, x_2) when the value of the objective function at the temporary base point is not improved.



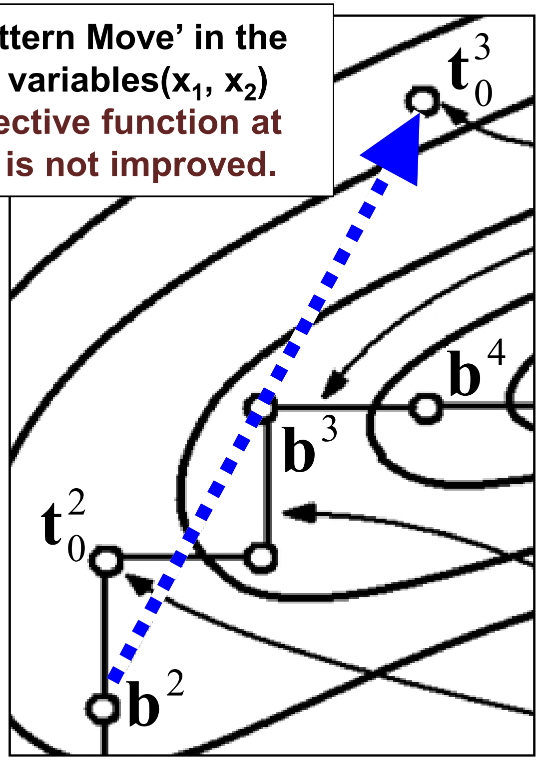
2. **If the result of the temporary base point is a better point than the previous base point, perform the 'Local Pattern Search' at the temporary base point. Otherwise, come back to the previous base point and perform the 'Local Pattern Search'.**

3.3 Direct Search Method
1. Hooke & Jeeves Method (12/16): Algorithm Summary (4)

3) Closing Condition (Stopping Criterion)

1. **When even this 'Local Pattern Search' fails($b^{k+1} = b^k$, there is no improvement), reduce the step sizes δ_i by half, $\delta_i/2$, and resume the 'Local Pattern Search'.**

Example of the 'Global Pattern Move' in the problem with two design variables(x_1, x_2) when the value of the objective function at the temporary base point is not improved.



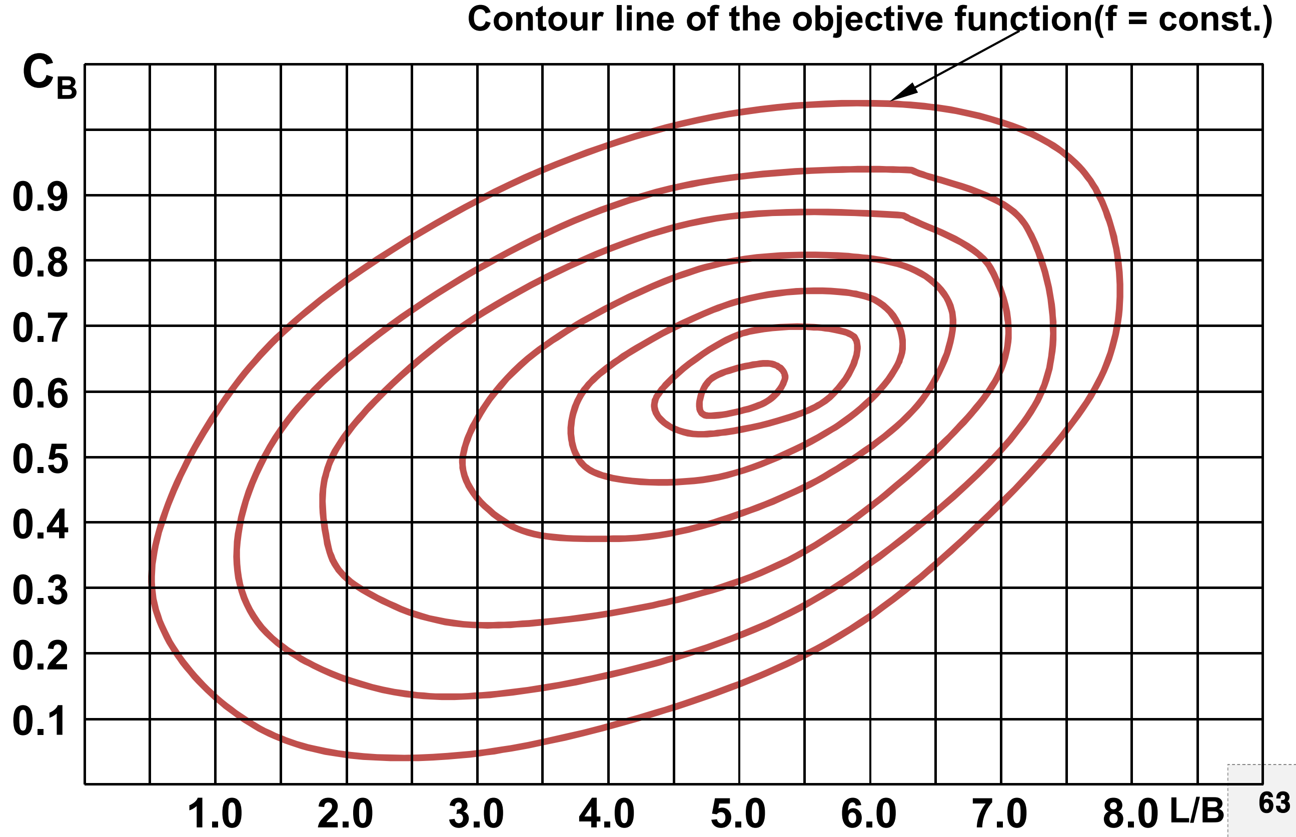
2. **If the step size δ_i is smaller than ε_i, stop the iteration and current base point is the optimal design point.**

3.3 Direct Search Method
1. Hooke & Jeeves Method (13/16): Example

☑ If the contour line of the objective function of shipbuilding cost with two design variables, L/B and C_B, is given as shown in the Figure, find the optimal value of the L/B and C_B to minimize the shipbuilding cost by using the 'Hooke & Jeeves Direct Search Method' and plot the procedures in the graph.

- Hooke & Jeeves Direct Search Method
  - Starting design point: L/B = 7.0, C_B = 0.2
  - Step size at the starting design point: Δ(L/B) = 0.5, $\Delta(C_B)$ = 0.1

Optimization problem with two unknown variables



3.3 Direct Search Method
1. Hooke & Jeeves Method (14/16): Example

$$x_1 = L/B,\ x_2 = C_B$$

- Iteration 1 : Local Pattern Search 1

$\mathbf{b}^0 = (7,\ 0.2),\ \Delta x_1 = 0.5,\ \Delta x_2 = 0.1,$

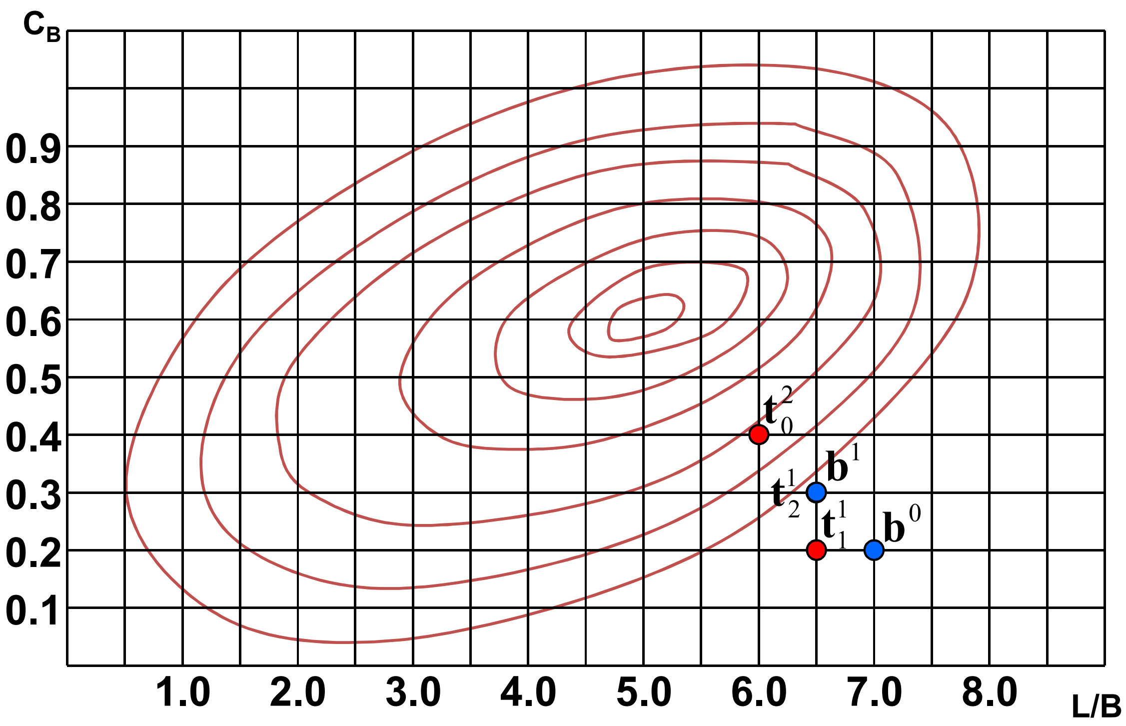
$\mathbf{t}_0^1 = \mathbf{b}^0$

Search from $\mathbf{t}_0^1$ in $-x_1$ direction $\rightarrow \mathbf{t}_1^1 = (6.5,\ 0.2)$

Search from $\mathbf{t}_1^1$ in $+x_2$ direction $\rightarrow \mathbf{t}_2^1 = (6.5,\ 0.3)$

Because the value of the objective function at $\mathbf{t}_2^1$ is improved, this point is adopted as a new base point.

$$\mathbf{b}^1 = \mathbf{t}_2^1$$



- Iteration 2 : Global Pattern Move 1

Define the temporary base point by using $\mathbf{b}^0$ and $\mathbf{b}^1$

$\rightarrow \mathbf{t}_0^2 = (6,\ 0.4)$

Because the value of the objective function at $\mathbf{t}_0^2$ is improved, perform the 'Local Pattern Search' at this point.

3.3 Direct Search Method
1. Hooke & Jeeves Method (15/16): Example

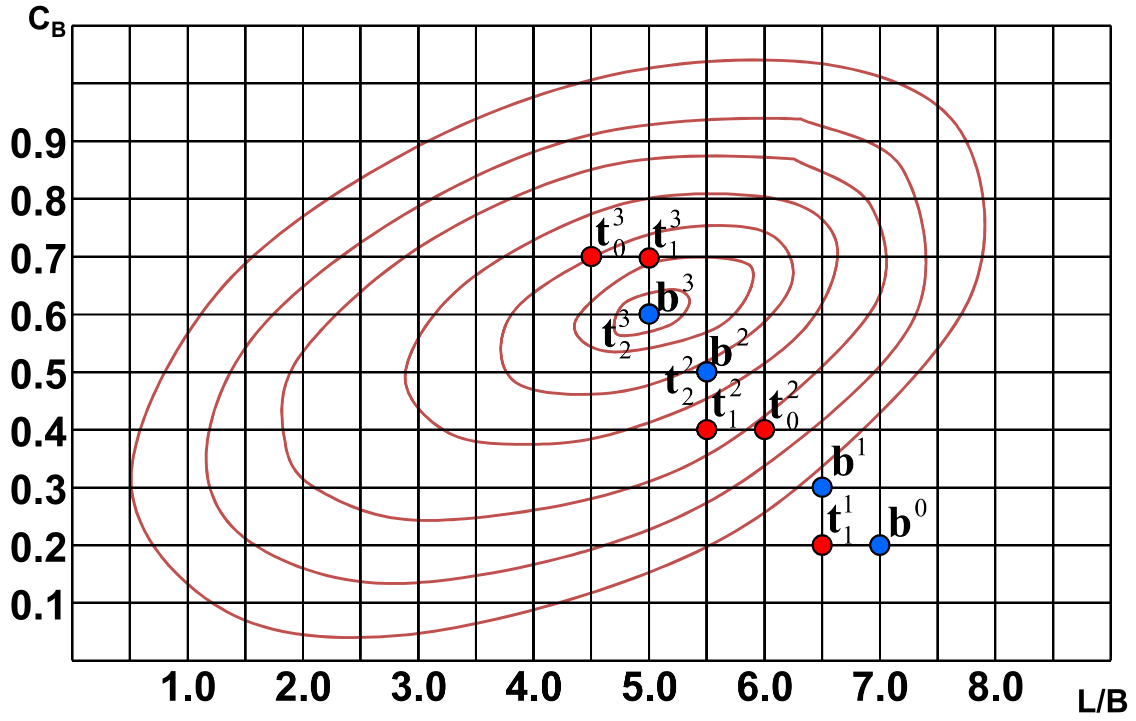
- Iteration 3 : Local Pattern Search 2

Search from $\mathbf{t}_0^2$ in $-x_1$ direction $\rightarrow \mathbf{t}_1^2 = (5.5,\ \ 0.4)$

Search from $\mathbf{t}_1^2$ in $+x_2$ direction $\rightarrow \mathbf{t}_2^2 = (5.5,\ \ 0.5)$

Because the value of the objective function at $\mathbf{t}_2^2$ is improved, this point is adopted as a new base point.

$$\mathbf{b}^2 = \mathbf{t}_2^2$$



- Iteration 4 : Global Pattern Move 2

Define the temporary base point by using $\mathbf{b}^1$ and $\mathbf{b}^2$

$\rightarrow \mathbf{t}_0^3 = (4.5,\ \ 0.7)$

- Iteration 5 : Local Pattern Search 3

Search from $\mathbf{t}_0^3$ in $+x_1$ direction $\rightarrow \mathbf{t}_1^0 = (5,\ \ 0.7)$

Search from $\mathbf{t}_1^3$ in $-x_2$ direction $\rightarrow \mathbf{t}_2^3 = (5,\ \ 0.6)$

Because the value of the objective function at $\mathbf{t}_2^3$ is improved, this point is adopted as a new base point.

$$\mathbf{b}^3 = \mathbf{t}_2^3$$

1. Hooke & Jeeves Method (16/16): Example

•Iteration 6 : Global Pattern Move 3

Define the temporary base point by using $\mathbf{b}^2$ and $\mathbf{b}^3$

$$\rightarrow \mathbf{t}_0^4 = (4.5,\ \ 0.7)$$

Because the value of the objective function at $\mathbf{t}_0^4$ is not improved,

$$\mathbf{t}_0^4 = \mathbf{b}^3$$

•Iteration 7 : Local Pattern Search 4

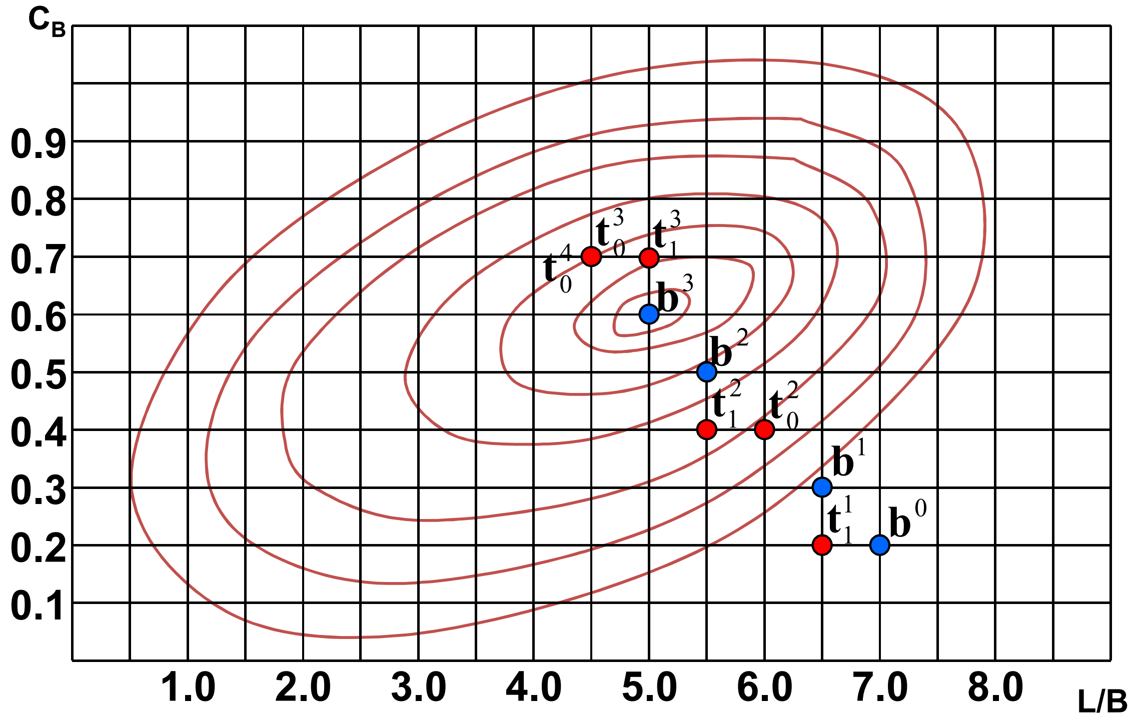
Because there is no improvement of the value of the objective function from the temporary base design point $\mathbf{t}_0^4$ in x_1 direction and x_2 direction,

$$\mathbf{t}_2^4 = \mathbf{t}_1^4 = \mathbf{t}_0^4$$

•Iteration 8 : Global Pattern Move 4

$\mathbf{b}^4 = \mathbf{b}^3 \rightarrow \ \Delta x_1 = 0.25,\ \ \Delta x_2 = 0.05,$

$\mathbf{t}_0^5 = \mathbf{b}^4$

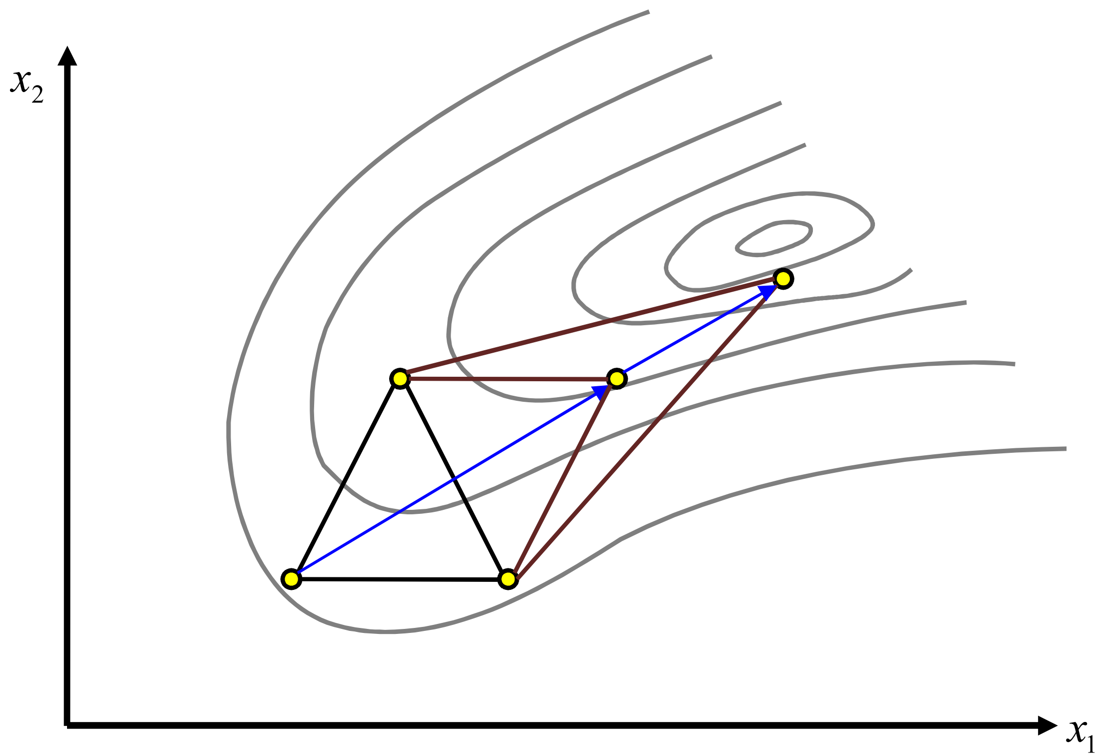


•Iteration 9 : Stopping the iteration of search

Because there is no improvement of the value of the objective function from base design point $(x_1,\ x_2) = (L/B,\ C_B) = (5.0,\ 0.6)$ in x_1 direction and x_2 direction by performing the 'Local Pattern Search' and 'Global Pattern Move', the optimal design point is $L/B = 5.0,\ \ C_B = 0.6$.

2. Nelder & Mead Simplex Method (1/14)

☑ This method is used to find optimal design point by successively reflecting, expanding, contracting, and reducing the simplex with (n+1) corners in the function of n design variables.



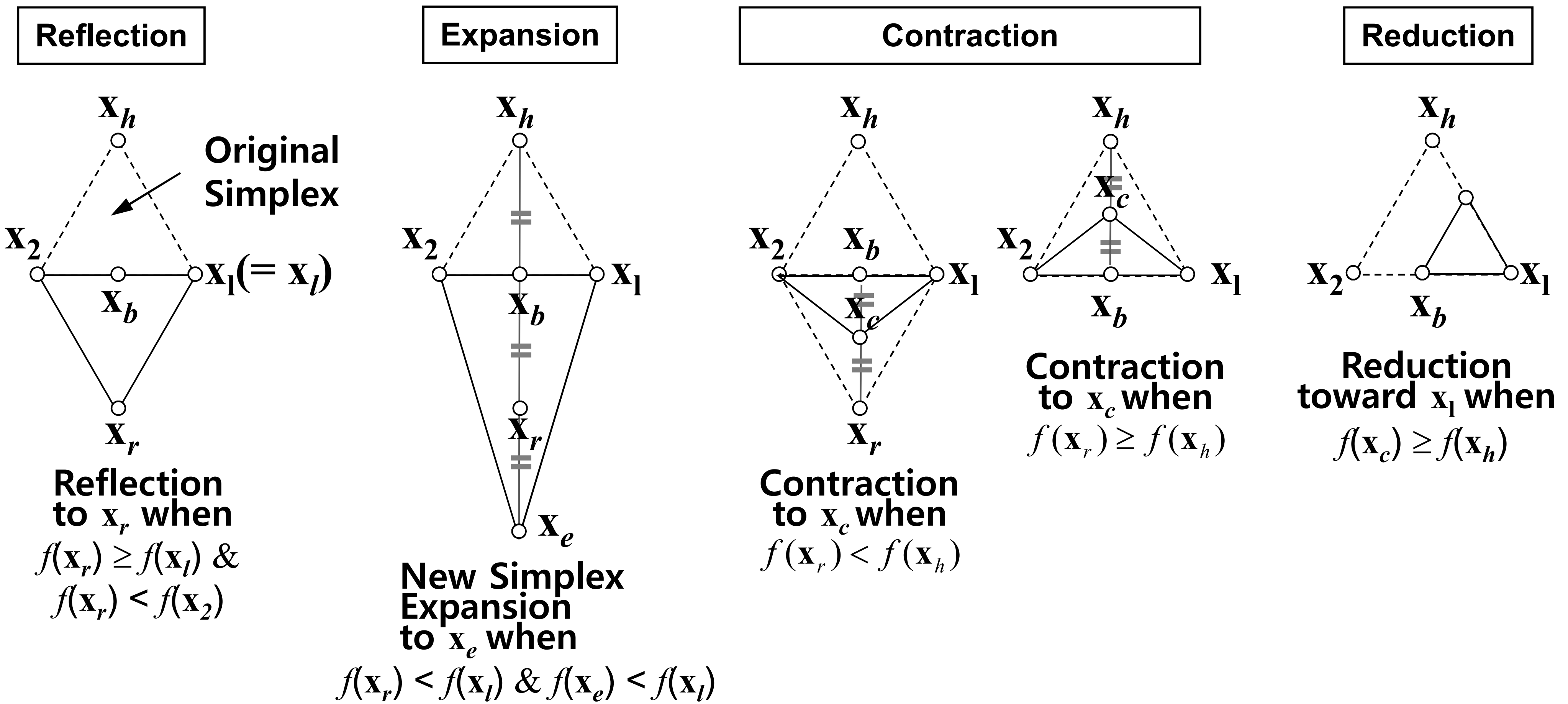
1. This method uses n+1 points in the function of n design variables.
   Ex) If the number of the design variables is two, this method use three points, i.e., triangle.

2. The simplex is reflected in the direction where the value of the objective function is improved.

3. If the value of the objective function is improved, the simplex is expanded. Otherwise, the simplex is reduced.

2. Nelder & Mead Simplex Method (2/14)

☑ The following figure shows various operations (Reflection, Expansion, Contraction, Reduction) for 2-dimensional case.



$\mathbf{x}_h$: Simplex point having the largest objective function value
$\mathbf{x}_l$: Simplex point having the smallest objective function value
$\mathbf{x}_b$: Center point between $\mathbf{x}_1$ and $\mathbf{x}_2$

2. Nelder & Mead Simplex Method (3/14)

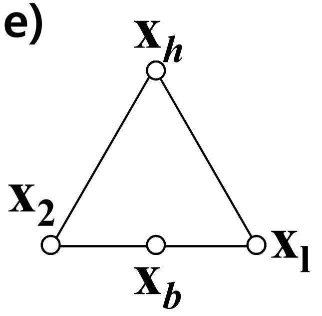
☑ **Step 1 :** Calculate the value of the objective function f at the $n+1$ corners of the simplex.

☑ **Step 2 :** Establish the corners which yield the highest, $\mathbf{x}_h$, and lowest, $\mathbf{x}_l$, of $f(\mathbf{x})$ in the current simplex.

☑ **Step 3 :** Calculate the value of the objective function f at the centroid($\mathbf{x}_b$) of all $\mathbf{x}_i$ except $\mathbf{x}_h$, i.e.,

$$\mathbf{x}_b = \frac{1}{n}\sum_{i=1}^{n+1}\mathbf{x}_i(\text{with } \mathbf{x}_h \text{ excluded})$$

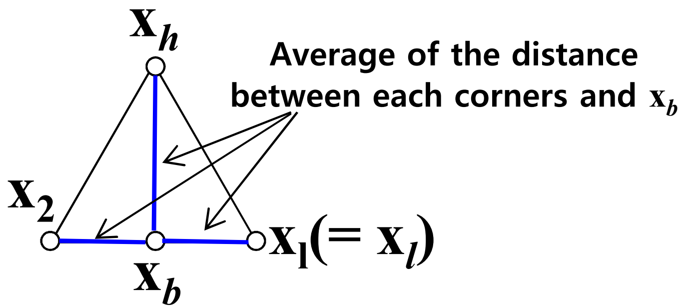
Example)



$$\mathbf{x}_b = \frac{\mathbf{x}_1 + \mathbf{x}_2}{2}$$

☑ **Step 4 : Test stopping criterion:**

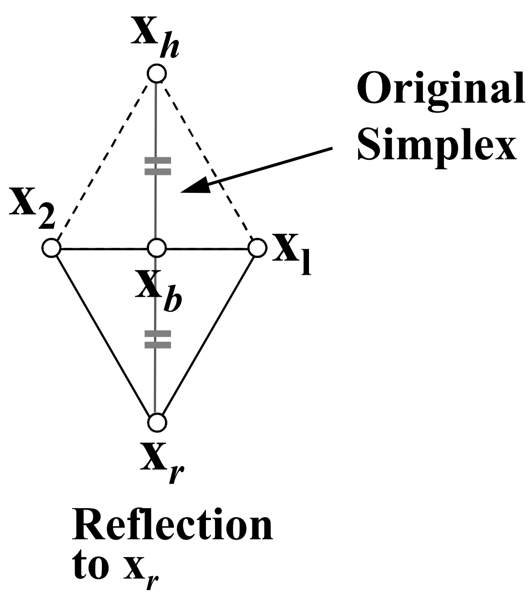
$$\left\{\frac{1}{n+1}\sum_{i=1}^{n+1}[f(\mathbf{x}_i)-f(\mathbf{x}_b)]^2\right\}^{1/2} \le \varepsilon$$



- If the stopping criterion is satisfied, stop and return $f(\mathbf{x}_l)$ as minimum. Otherwise, continue.

☑ **Step 5 : Reflection**

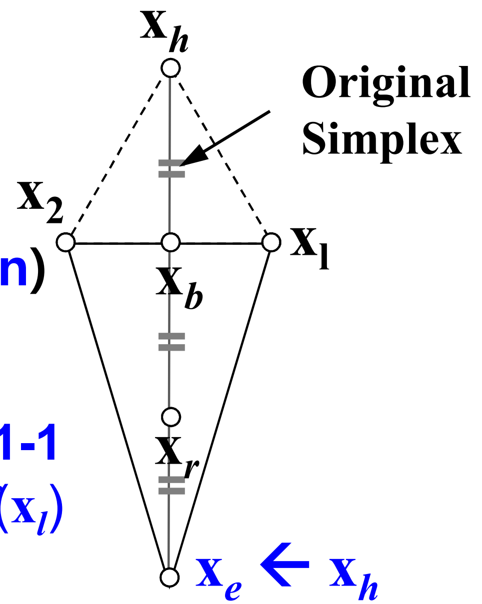
- Reflect $\mathbf{x}_h$ through $\mathbf{x}_b$ to give $\mathbf{x}_r = 2\mathbf{x}_b - \mathbf{x}_h$.
  Calculate the value of the objective function f at $\mathbf{x}_r$ and change the simplex as following conditions.



2. Nelder & Mead Simplex Method (5/14)

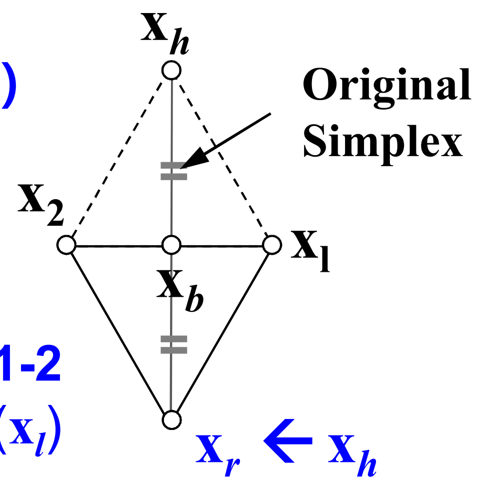
☑ **Step 6 : Expansion**

- **Step 6-1 : If $f(\mathbf{x}_r) < f(\mathbf{x}_l)$, reflect $\mathbf{x}_b$ through $\mathbf{x}_r$ to give $\mathbf{x}_e = 2\mathbf{x}_r - \mathbf{x}_b$.**
  And then, calculate $f(\mathbf{x}_e)$ and compare $f(\mathbf{x}_e)$ and $f(\mathbf{x}_l)$.



- **Step 6-1-1 : If $f(\mathbf{x}_e) < f(\mathbf{x}_l)$, replace $\mathbf{x}_h$ by $\mathbf{x}_e$(expansion) and return to Step 2.**

➡ **Step 6-1-1**
$f(\mathbf{x}_e) < f(\mathbf{x}_l)$



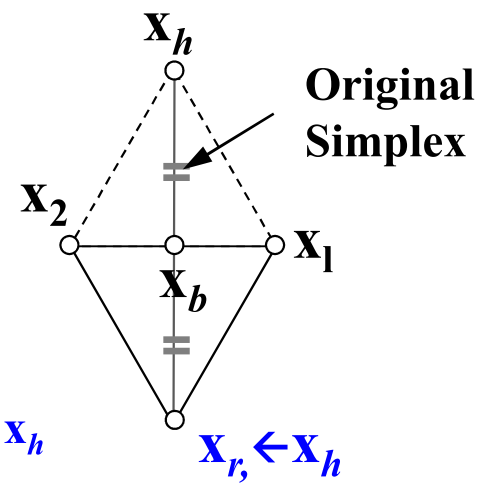
- **Step 6-1-2 : If $f(\mathbf{x}_e) \geq f(\mathbf{x}_l)$, replace $\mathbf{x}_h$ by $\mathbf{x}_r$(reflection) and return to Step 2.**

➡ **Step 6-1-2**
$f(\mathbf{x}_e) \geq f(\mathbf{x}_l)$

2. Nelder & Mead Simplex Method (6/14)

☑ **Step 6 : Expansion**

- **Step 6-2 : If $f(\mathbf{x}_r) \geq f(\mathbf{x}_l)$,**
  - Step 6-2-1 : test $f(\mathbf{x}_r) < f(\mathbf{x}_i)$ for all $\mathbf{x}_i$ except $\mathbf{x}_h$.
    If true, replace $\mathbf{x}_h$ by $\mathbf{x}_r$(reflection)
    and return to Step 2.



➡ Step 6-2-1
For all $\mathbf{x}_i$ except $\mathbf{x}_h$
$f(\mathbf{x}_r) < f(\mathbf{x}_i)$

  - Step 6-2-2 : If false, continue.

2. Nelder & Mead Simplex Method (7/14)

☑ **Step 7 : Contraction**

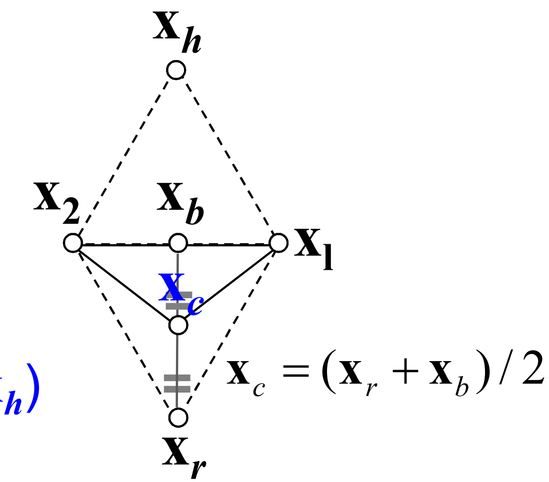
- **Step 7-1 : If $f(\mathbf{x}_r) < f(\mathbf{x}_h)$,**

  calculate the value of the objective function f

  at $\mathbf{x}_c = (\mathbf{x}_r + \mathbf{x}_b)/2.$

➡ **Step 7-1**

$f(\mathbf{x}_r) < f(\mathbf{x}_h)$



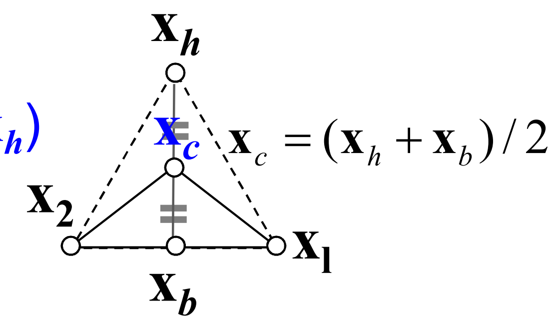
- **Step 7-2 : If $f(\mathbf{x}_r) \ge f(\mathbf{x}_h)$,**

  calculate the value of the objective function f

  at $\mathbf{x}_c = (\mathbf{x}_h + \mathbf{x}_b)/2.$

➡ **Step 7-2**

$f(\mathbf{x}_r) \ge f(\mathbf{x}_h)$

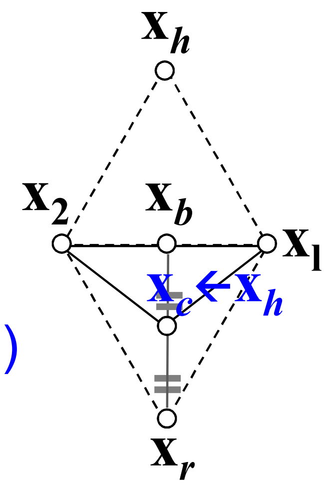


2. Nelder & Mead Simplex Method (8/14)

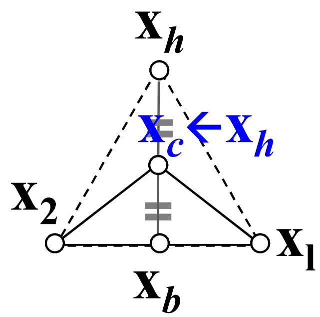
☑ **Step 8 : Reduction**

- **Step 8-1 : If $f(\mathbf{x}_c) < f(\mathbf{x}_h)$,**
  replace $\mathbf{x}_h$ by $\mathbf{x}_c$ (contraction)
  and return to Step 2.

➡ Step 8-1
$f(\mathbf{x}_c) < f(\mathbf{x}_h)$

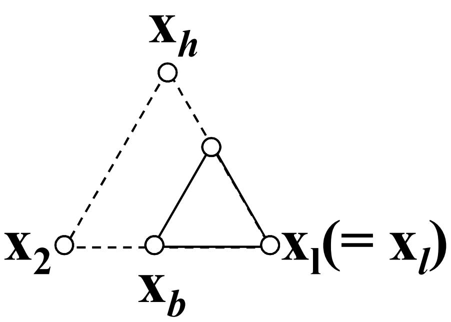


or



- **Step 8-2 : If $f(\mathbf{x}_c) \geq f(\mathbf{x}_h)$,**
  reduce the simplex toward $\mathbf{x}_l$ using $\mathbf{x}_i = (\mathbf{x}_i + \mathbf{x}_l)/2$
  (reduction) and return to Step 2.

➡ Step 8-2
$f(\mathbf{x}_c) \geq f(\mathbf{x}_h)$



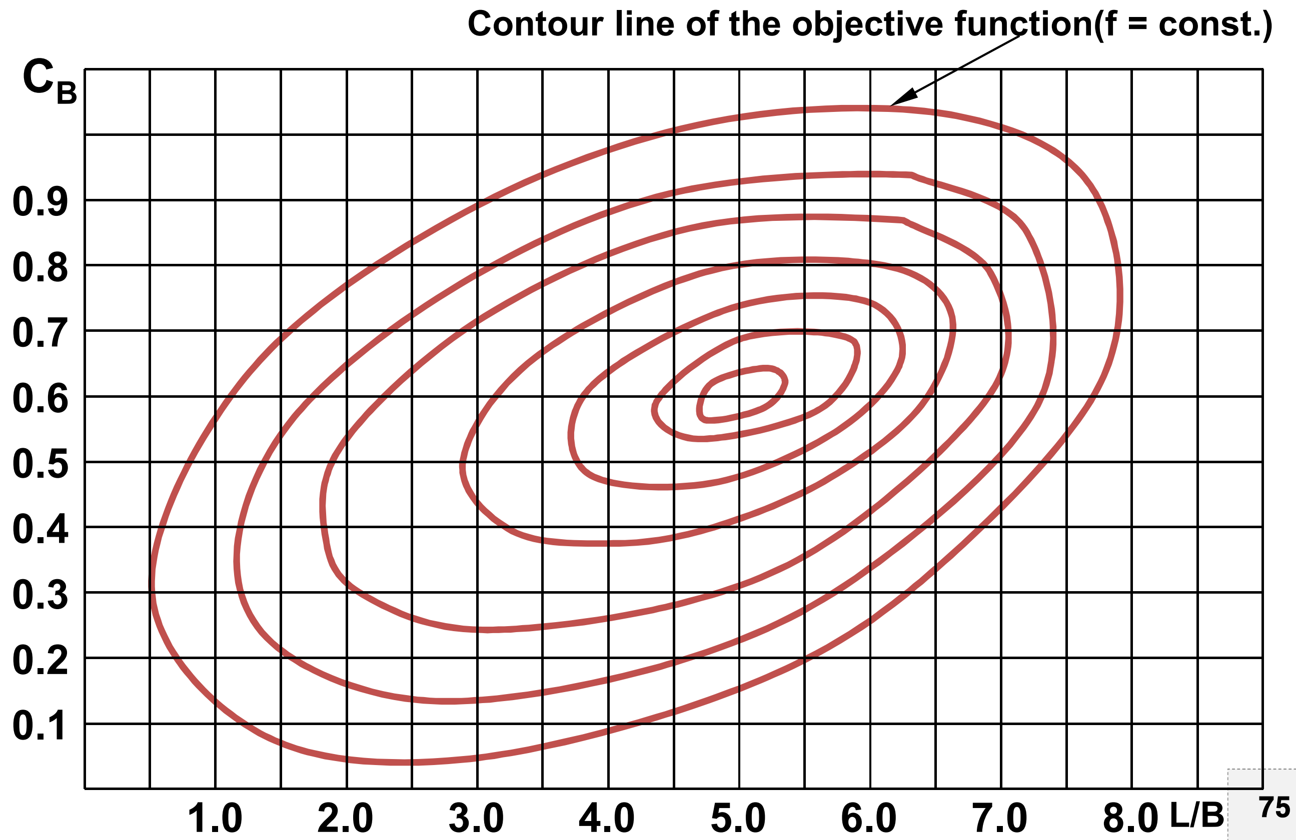
Reduction toward $\mathbf{x}_1$

3.3 Direct Search Method
2. Nelder & Mead Simplex Method (9/14): Example

☑ If the contour line of the objective function of shipbuilding cost with two design variables, L/B and C_B, is given as shown in Fig, find the value of the L/B and C_B to minimize the shipbuilding cost by using the 'Nelder & Mead Simplex Method' and plot the procedures in the graph.

- Nelder & Mead Simplex Method
  - Starting corners of the simplex: (L/B, CB) = (1, 0.1), (1.5, 0.1), (1.5, 0.2)
  - Stopping criterion: 0.01

Optimization problem with two unknown variables



2. Nelder & Mead Simplex Method (10/14): Example

$$x_1 = L/B,\ x_2 = C_B$$

Triangle 1 : $x_1,\ x_2,\ x_3$

Iteration 1) Because x_2 is x_h, reflect x_2 through the center between x_1 and x_3. $\rightarrow x_r$

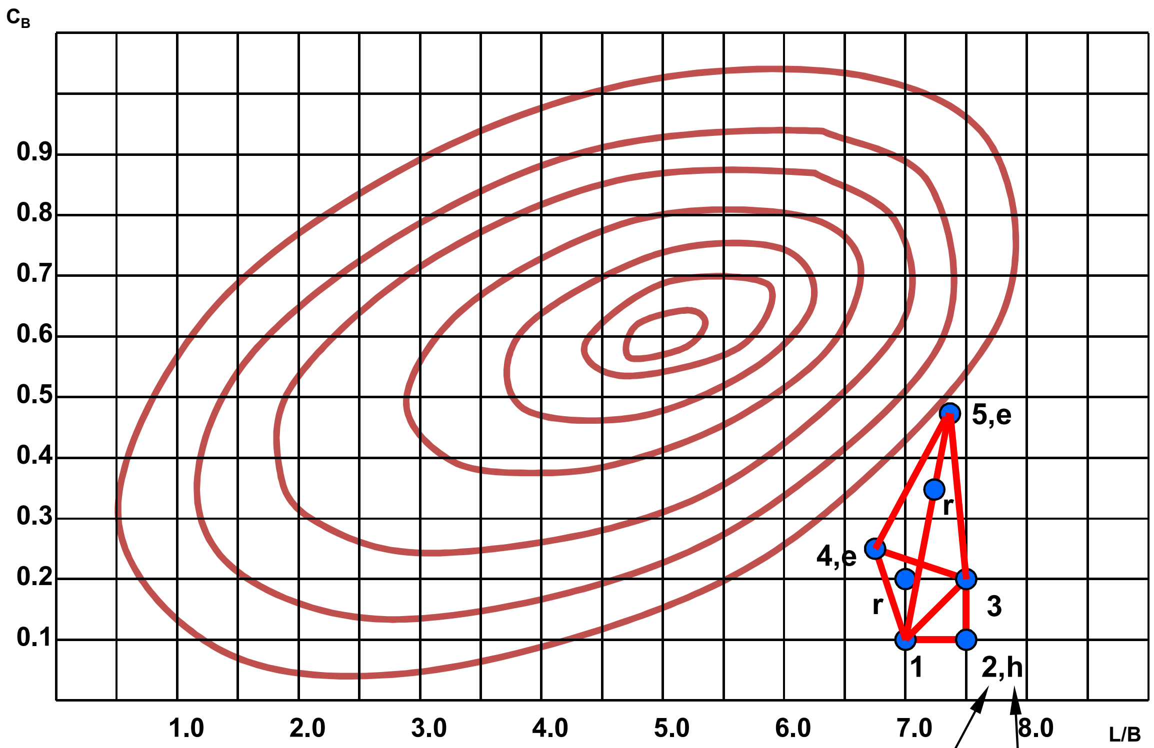
Because $f(x_r) < f(x_1)$ and $f(x_3)$, perform the expansion $\rightarrow x_{4,e}$

$\rightarrow$ Triangle 2 : $x_1,\ x_3,\ x_4$

Iteration 2) Because x_1 is x_h, reflect x_1 through the center between x_3 and x_4. $\rightarrow x_r$

Because $f(x_r) < f(x_3)$ and $f(x_4)$, perform the expansion $\rightarrow x_{5,e}$

$\rightarrow$ Triangle 3 : $x_3,\ x_4,\ x_5$



Number means the index 'i' of x_i.

Alphabet means the kind of x_i .

h: maximum point of the corner in the simplex(triangle)

r: reflection

e: expansion

c: contraction

2. Nelder & Mead Simplex Method (11/14): Example

$$x_1 = L/B,\ x_2 = C_B$$

Iteration 3) Because x_3 is x_h, reflect x_3 through the center between x_4 and x_5. $\rightarrow x_r$

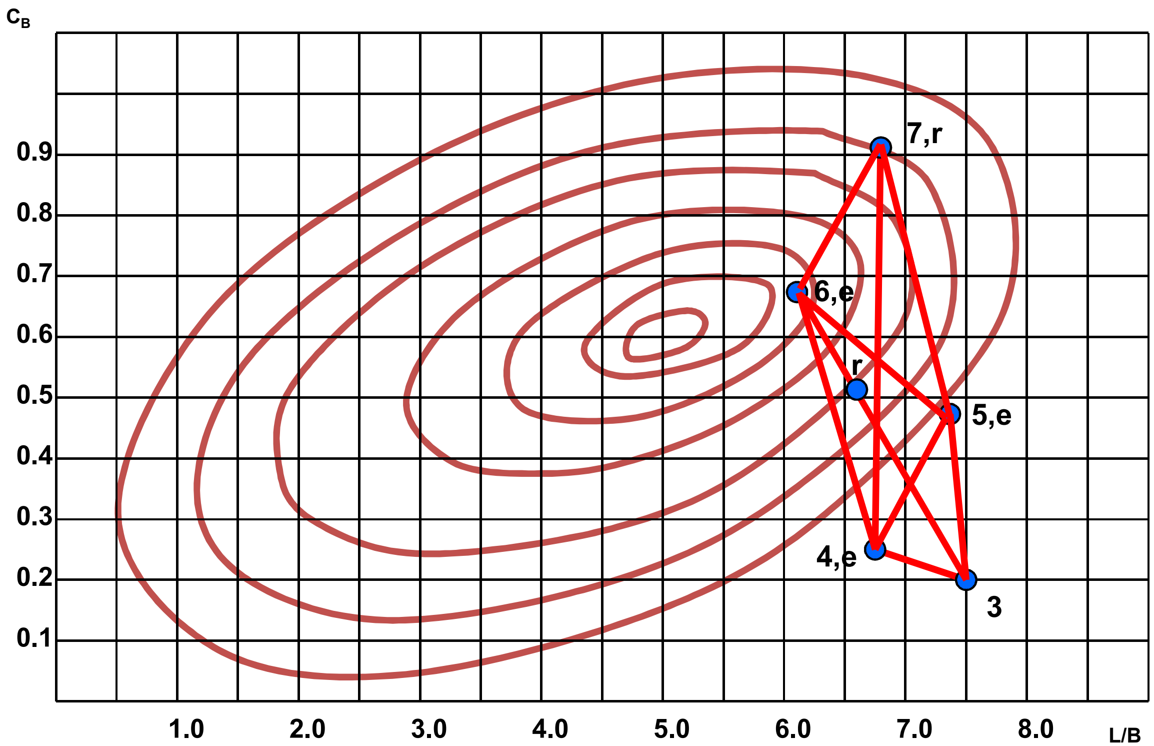
Because $f(x_r) < f(x_4)$ and $f(x_5)$, perform the expansion $\rightarrow x_{6,e}$

$\rightarrow$ Triangle 4 : $x_4,\ x_5,\ x_6$

Iteration 4) Because x_4 is x_h, reflect x_4 through the center between x_5 and x_6. $\rightarrow x_{7,r}$

Because $f(x_{7,r}) > f(x_6)$, go to the next iteration.

$\rightarrow$ Triangle 5 : $x_5,\ x_6,\ x_7$



2. Nelder & Mead Simplex Method (12/14): Example

Iteration 5) Because x_5 is x_h, reflect x_5 through the center between x_6 and x_7. $\rightarrow x_r$

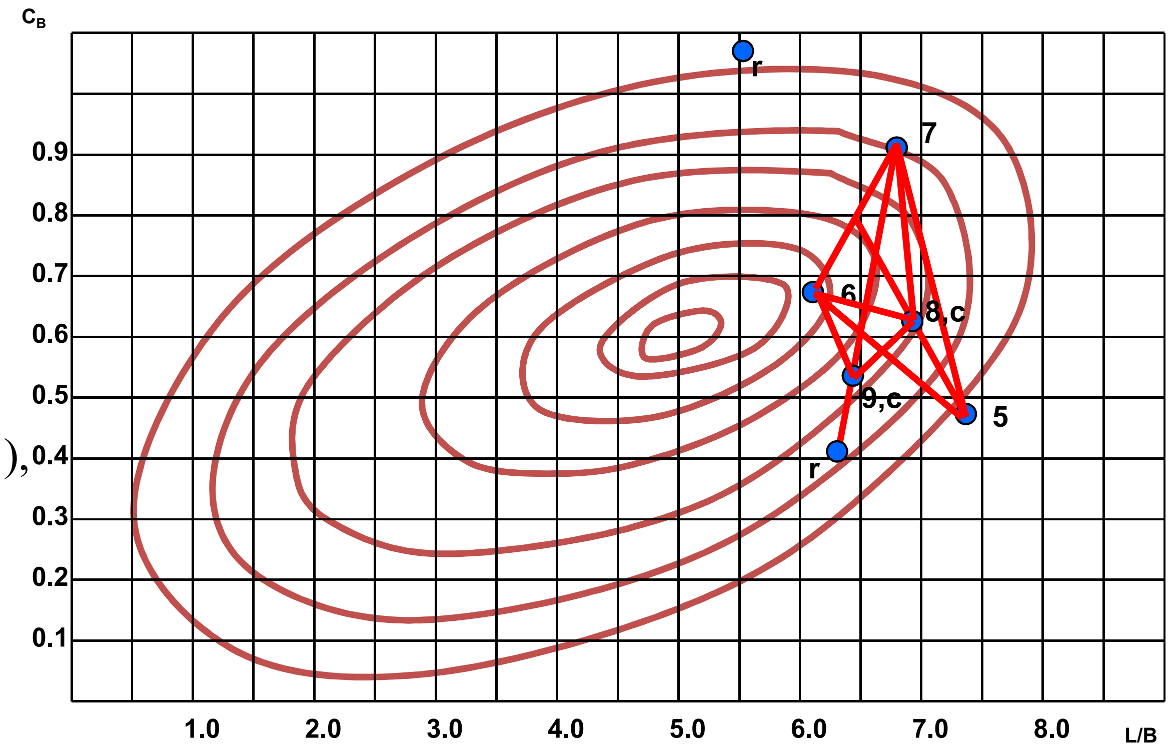
Because $f(x_r) > f(x_5), f(x_6)$ and $f(x_7)$, perform the constraction. $\rightarrow x_{8,c}$

$\rightarrow$ Triangle 6 : $x_6,\ x_7,\ x_8$

Iteration 6) Because x_7 is x_h, reflect x_7 through the center between x_6 and x_8. $\rightarrow x_r$

Because $f(x_r) > f(x_6), f(x_8)$ and $f(x_r) < f(x_7)$, contract the simplex toward $\mathrm{x}_r \rightarrow x_{9,c}$

$\rightarrow$ Triangle 7 : $x_6,\ x_8,\ x_9$



3.3 Direct Search Method
2. Nelder & Mead Simplex Method (13/14): Example

Iteration 7) Because x_8 is x_h, reflect x_8 through the center between x_6 and x_9. $\rightarrow x_r$

Because $f(x_r) < f(x_6), f(x_9)$,

preforme the expansion $\rightarrow\ x_{10,c}$

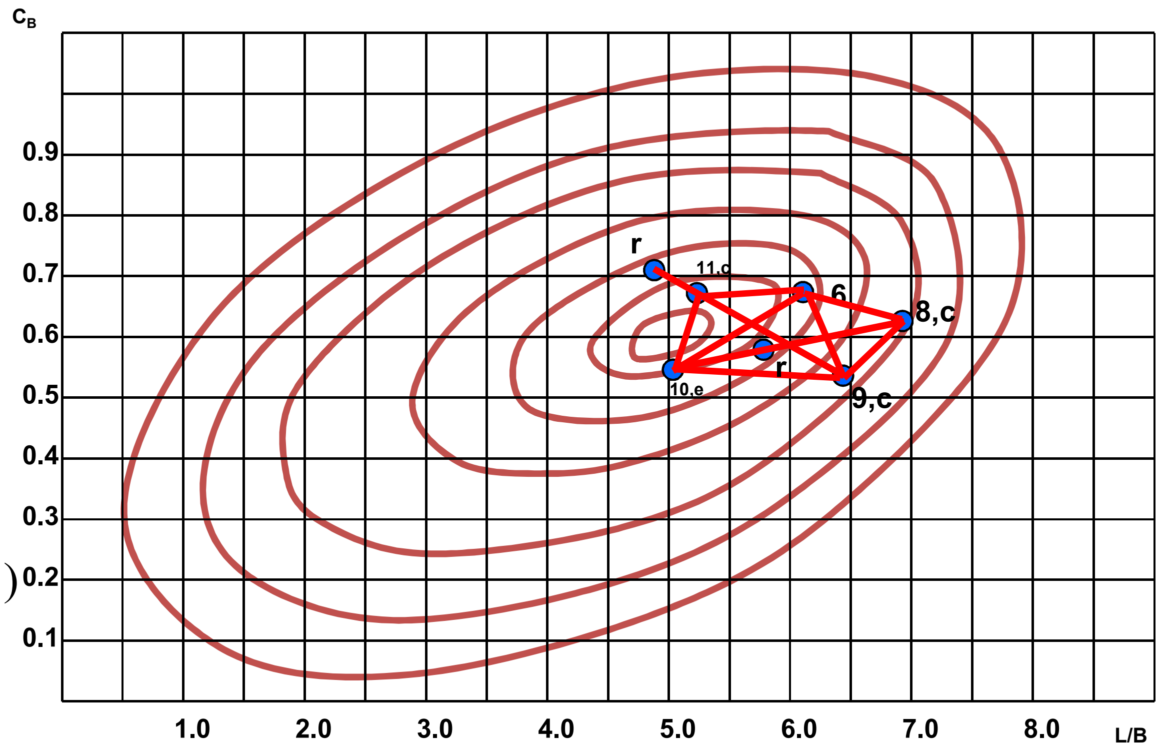
$\rightarrow$ Triangle 8 : $x_6,\ x_9,\ x_{10}$

Iteration 8) Because $x_{9,c}$ is x_h, reflect $x_{9,c}$ through the center between x_6 and x_{10}. $\rightarrow x_r$

Because $f(x_r) > f(x_6), f(x_{10})$ and $f(x_r) < f(x_9)$

contract the simplex toward $x_r \rightarrow\ x_{11,c}$

$\rightarrow$ Triangle 9 : $x_6,\ x_{10},\ x_{11}$

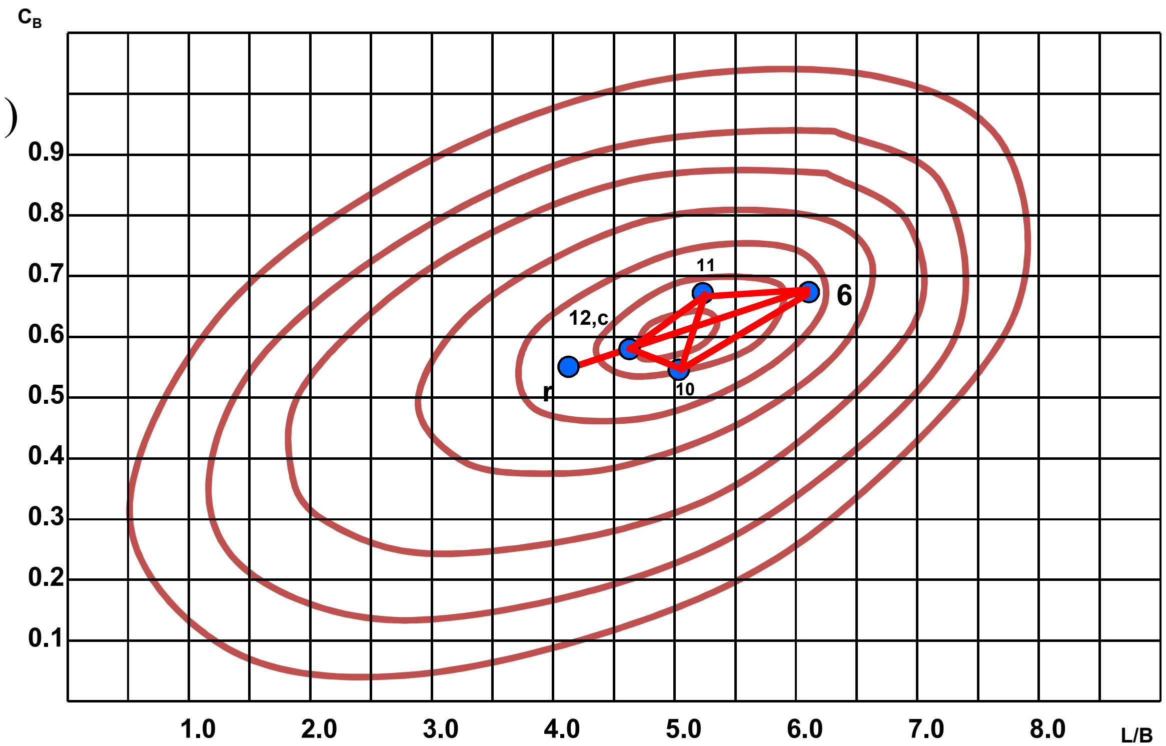


3.3 Direct Search Method
2. Nelder & Mead Simplex Method (14/14): Example

Iteration 9) Because x_6 is x_h, reflect x_6 through the center between x_{10} and x_{11}. $\rightarrow x_r$

Because $f(x_r) > f(x_{10}), f(x_{11})$ and $f(x_r) < f(x_6)$ contract the simplex toward $x_r \rightarrow x_{12,c}$

$\rightarrow$ Triangle 10 : $x_{10}, \; x_{11}, \; x_{12}$



$x_1(7, \; 0.1)$ $\quad x_2(7.5, \; 0.1)$

$x_3(7.5, \; 0.2)$ $\quad x_4(6.75, \; 0.25)$

$x_5(7.375, \; 0.475)$ $\quad x_6(6.1875, \; 0.6875)$

$x_7(6.8125, \; 0.9125)$ $\quad x_8(6.9375, \; 0.6375)$

$x_9(6.4375, \; 0.5375)$ $\quad x_{10}(5.0625, \; 0.5625)$

$x_{11}(5.21875, \; 0.66875)$ $\quad x_{12}(4.6171875, \; 0.5796875)$

Performing 10 times iterations, we can recognize that the simplex(triangle) has the tendency to approach the result obtained by the 'Hooke & Jeeves direct search method'.