Unconstrained optimization

Sridharakumar Narasimhan
Professor,
Department of Chemical Engineering
IIT Madras

Outline

- Unconstrained optimization
- Univariate
- Multivariate

Unconstrained optimization

- Class of optimization problems
  - Objective function
  - No constraints on decision variables
- Classification
  - Univariate – single decision variable
  - Multivariate – More than one decision variables

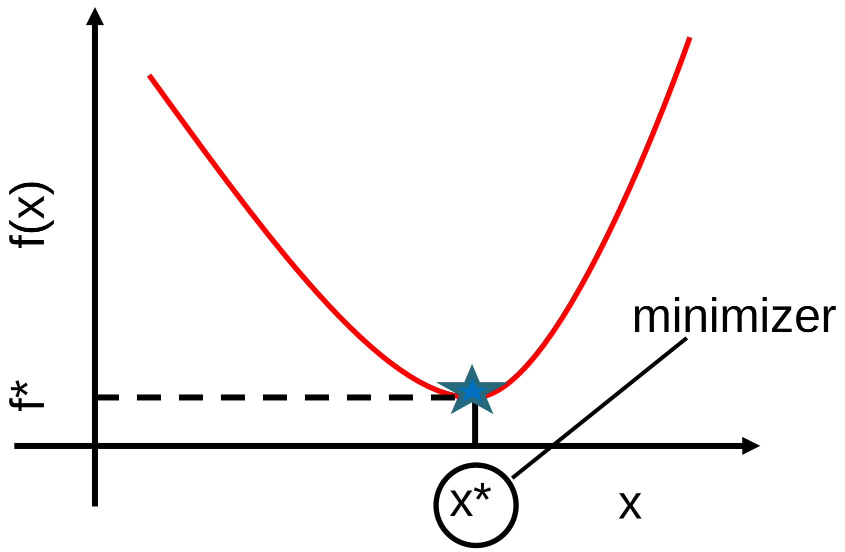
Univariate unconstrained optimization

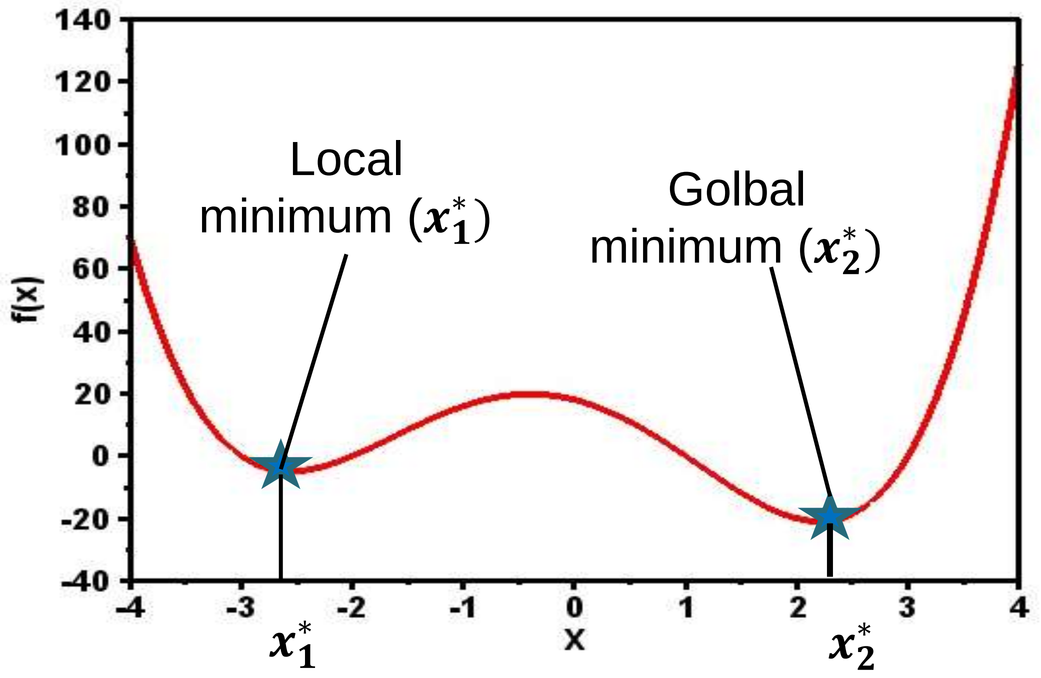
$$\min_{x} \; f(x)$$

$$x \in R$$

Decision variable

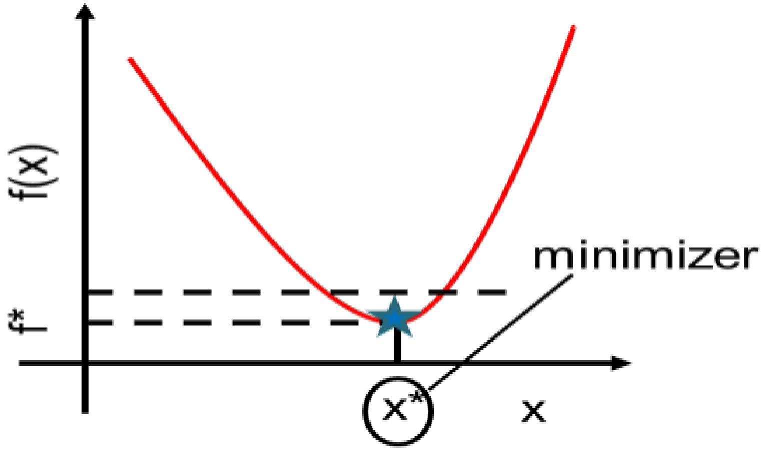
Objective function





Univariate unconstrained optimization

$$\boxed{\begin{array}{c}\min\limits_{x} \; f(x) \\ x \in R\end{array}}$$



Approximate $f(x)$ as a quadratic function using taylor series at a point x^k

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

When $x^k = x^*$,

$$f(x) = f(x^*) + \underbrace{\cancel{\frac{1}{1!}f'(x^*)(x - x^*)}}_{0} + \frac{1}{2!}f''(x^*)(x - x^*)^2$$

$$\underbrace{f(x) - f(x^*)}_{\text{Positive}} = \frac{1}{2!}\underbrace{f''(x^*)}_{\text{Has to be positive}}\underbrace{(x - x^*)^2}_{\text{Always positive}}$$

Univariate unconstrained optimization

$$\boxed{\begin{array}{c}\min\limits_{x} \; f(x) \\ x \in R\end{array}}$$

Necessary and sufficient conditions for x^* to be the minimizer of the function $f(x)$

First order necessary conditon: $f'(x^*) = 0$

Second order sufficiency condition: $f''(x^*) > 0$

Univariate unconstrained optimization : Example

$$\min_{x} \; f(x)$$

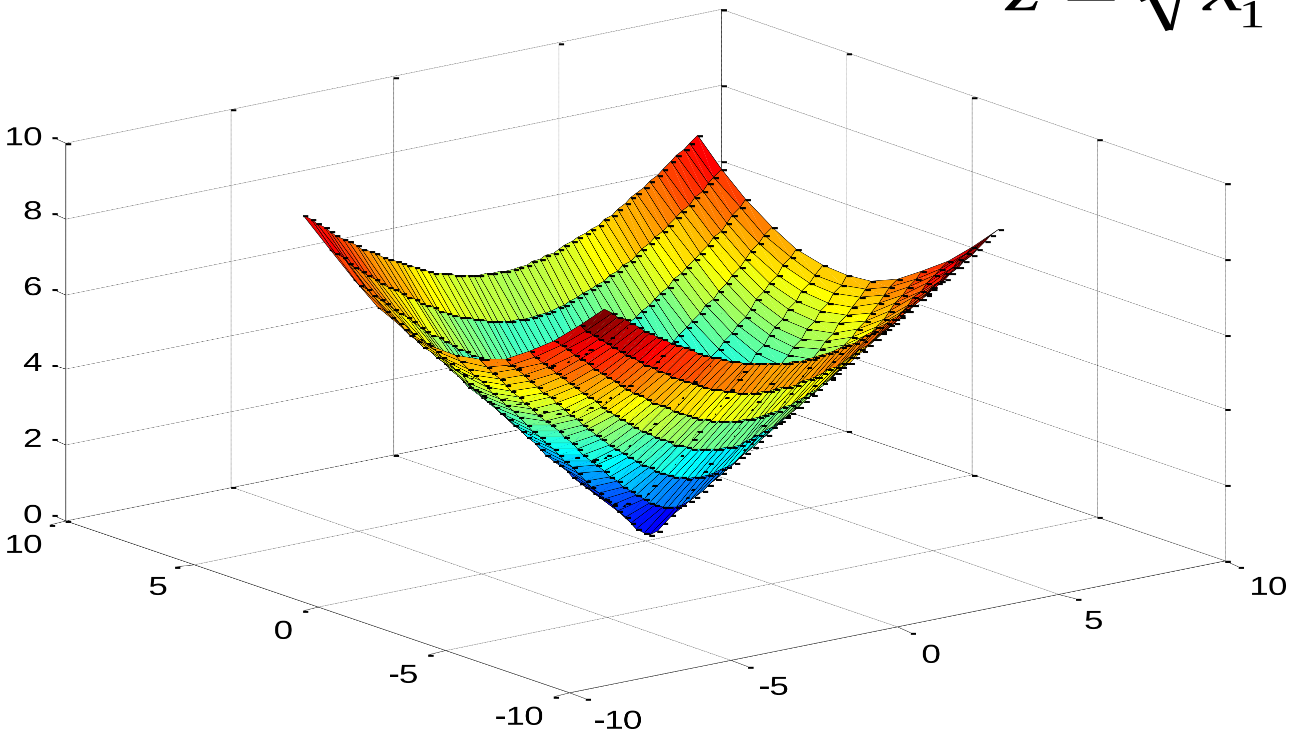
$$f(x) = 3x^4 - 4x^3 - 12x^2 + 3$$

First order condition	Second order condition
$f'(x) = 12x^3 - 12x^2 - 24x = 0$ $= 12x(x^2 - x - 2x) = 0$ $= 12x(x+1)(x-2) = 0$ $x = 0, x = -1, x = 2$	$f''(x) = 36x^2 - 24x - 24$ $f''(x)\vert_{x=0} = -24$ $f''(x)\vert_{x=-1} = 36 > 0$ $f''(x)\vert_{x=2} = 72 > 0$
$f(-1) = -2$	$f(2) = -29$
$x^* = -1$, is a local minimizer of $f(x)$	$x^* = 2$, is a global minimizer of $f(x)$

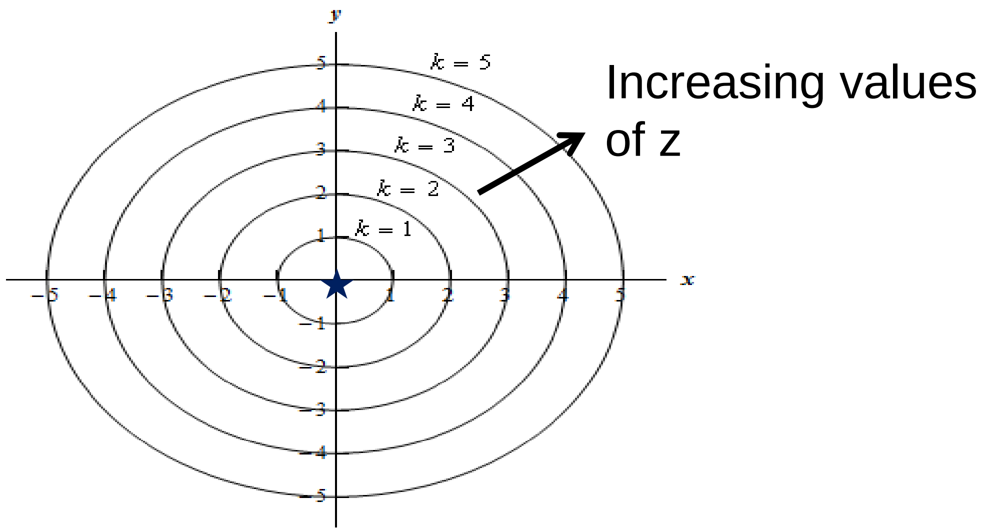
Multivariate unconstrained optimization

$$z = f(x_1, x_2 \ldots x_n)$$

$$z = \sqrt{x_1^2 + x_2^2}$$



Contour plot

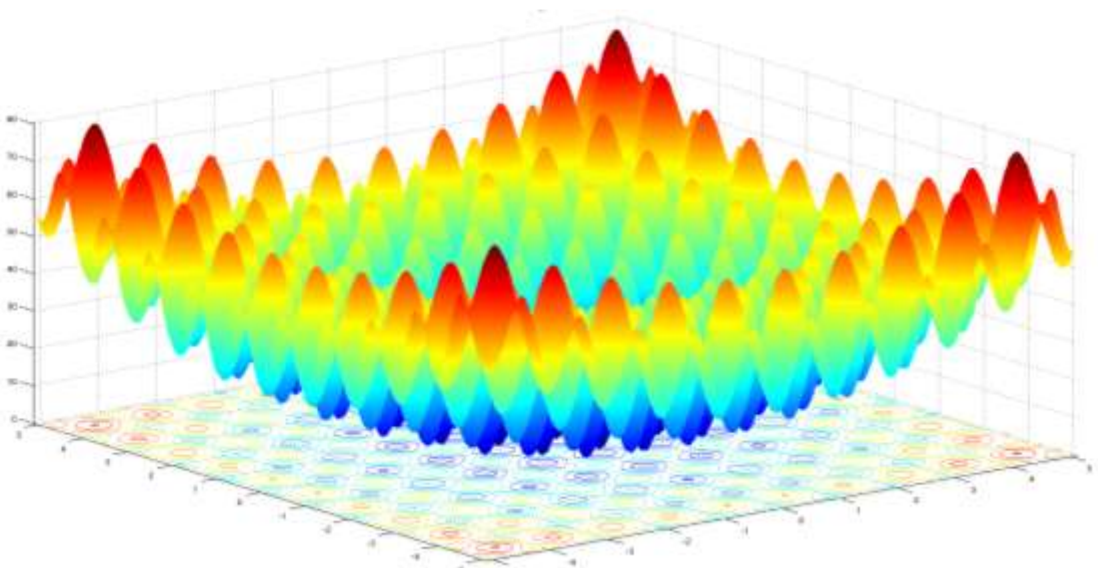


The minimum value of the function is at [0,0]

Multivariate unconstrained optimization

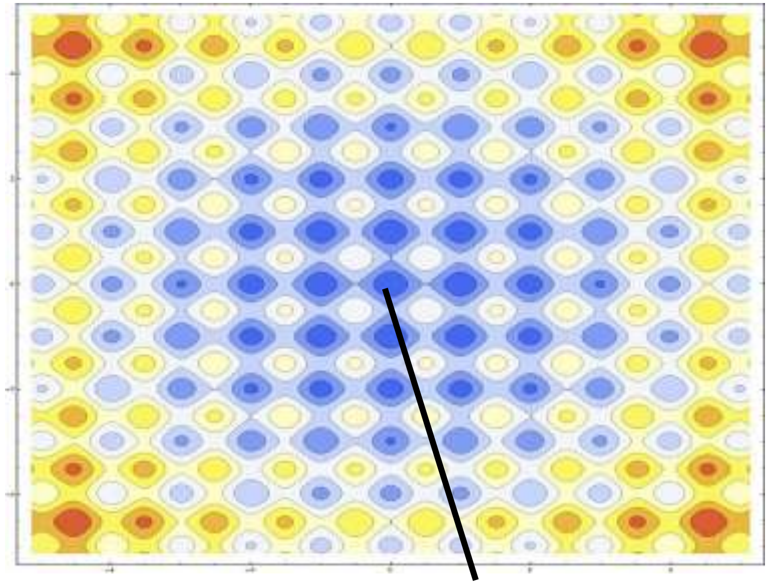
Rastrigin function

$$f(x_1, x_2) = 20 + \sum_{i=1}^{2} [x_i^2 - 10\cos(2\pi x_i)]$$



http://en.wikipedia.org/wiki/Rastrigin_function

Contour plot

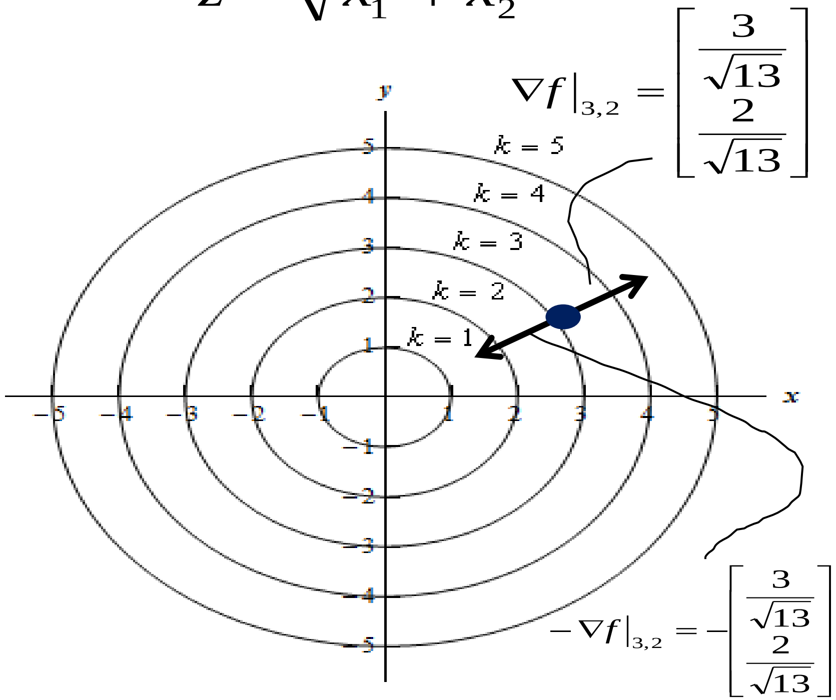


Global minimum at $[0,0]$

Multivariate unconstrained optimization

$$z = f(x_1, x_2 \ldots x_n)$$

$$z = \sqrt{x_1^2 + x_2^2}$$



<u>Gradient</u>

$$\nabla f = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \\ \cdots \\ \cdots \\ \frac{\partial f}{\partial x_n} \end{bmatrix}$$

<u>Hessian</u>

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

- Gradient of a function at a point is orthogonal to the contours
- Gradient points in the direction of greatest increase of the function
- Negative gradient points in the direction of the greatest decrease of the function
- Hessian is a symmetric matrix

Multivariate unconstrained optimization

Approximate $f(\bar{x})$ as a quadratic using taylor series at a point $\overline{x^k}$

$$f(\bar{x}) = f(\overline{x^k}) + [\nabla f(\overline{x^k})]^T(\bar{x} - \overline{x^k}) + \frac{1}{2}(\bar{x} - \overline{x^k})^T \nabla^2 f(x^k)(\bar{x} - \overline{x^k})$$

At $\overline{x^k} = \overline{x^*}$ (minimizer of $f(\bar{x})$)

$$f(\bar{x}) = f(\overline{x^*}) + \cancel{[\nabla f(\overline{x^*})]^T(\bar{x} - \overline{x^*})}^{0} + \frac{1}{2}(\bar{x} - \overline{x^*})^T \nabla^2 f(x^*)(\bar{x} - \overline{x^*})$$

$$\underbrace{f(\bar{x}) - f(\overline{x^*})}_{\text{Positive}} = \frac{1}{2}\underbrace{(\bar{x} - \overline{x^*})^T \nabla^2 f(\overline{x^*})(\bar{x} - \overline{x^*})}_{\text{Has to be Positive}}$$

Multivariate unconstrained optimization

$$(\bar{x} - \overline{x^*})^T \nabla^2 f(\overline{x^*})(\bar{x} - \overline{x^*}) > 0$$

$$(\bar{v})^T \nabla^2 f(\overline{x^*})(\bar{v}) > 0$$

Condition for Hessian to be positive definite

Hessian matrix is said to be positive definite at a point if all the eigen values of the Hessian matrix are positive

Multivariate unconstrained optimization

$$\boxed{\begin{array}{c}\min\limits_{x} \; f(x) \\ x \in R\end{array}}$$

$$\boxed{\begin{array}{c}\min\limits_{\bar{x}} \; f(\bar{x}) \\ \bar{x} \in R^n\end{array}}$$

Necessary condition for x^* to be the minimizer

$$f'(x^*) = 0$$

Sufficient condition

$$f''(x^*) > 0$$

Necessary condition for $\overline{x^*}$ to be the minimizer

$$\nabla f(x^*) = 0$$

Sufficient condition

$\nabla^2 f(\overline{x^*})$ has to be positive definite

Multivariate unconstrained optimization: Example

$$\min_{x_1, x_2} \; x_1 + 2x_2 + 4x_1^2 - x_1x_2 + 2x_2^2$$

First order condition

$$\nabla f = \begin{bmatrix} \dfrac{\partial f}{\partial x_1} \\ \dfrac{\partial f}{\partial x_2} \end{bmatrix} = \begin{bmatrix} 1 + 8x_1 - x_2 \\ 2 - x_1 + 4x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

solving

$$\begin{bmatrix} x_1^* \\ x_2^* \end{bmatrix} = \begin{bmatrix} -0.19 \\ -0.54 \end{bmatrix}$$

Second order condition

$$\nabla^2 f = \begin{bmatrix} \dfrac{\partial^2 f}{\partial {x_1}^2} & \dfrac{\partial^2 f}{\partial x_1 x_2} \\ \dfrac{\partial^2 f}{\partial x_2 x_1} & \dfrac{\partial^2 f}{\partial {x_2}^2} \end{bmatrix} = \begin{bmatrix} 8 & -1 \\ -1 & 4 \end{bmatrix}$$

$$\begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix} = \begin{bmatrix} 3.76 \\ 8.23 \end{bmatrix}$$

$$\begin{bmatrix} x_1^* \\ x_2^* \end{bmatrix} = \begin{bmatrix} -0.19 \\ -0.54 \end{bmatrix}$$