Nonlinear Optimization

UNCONSTRAINED MULTIVARIATE OPTIMIZATION

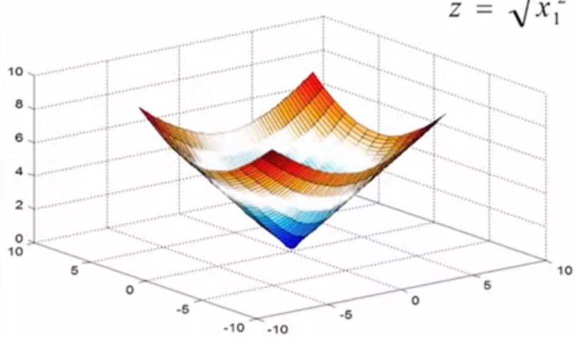
1

Multivariate optimization – Contour plots

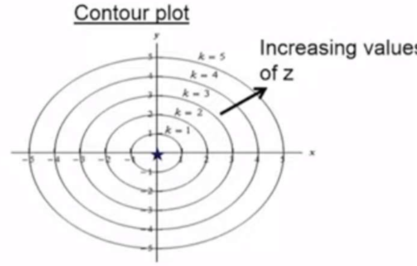
Multivariate optimization

$$z = f(x_1, x_2 .... x_n)$$

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



Contour plot



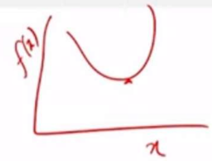
The minimum value of the function is at [0,0]

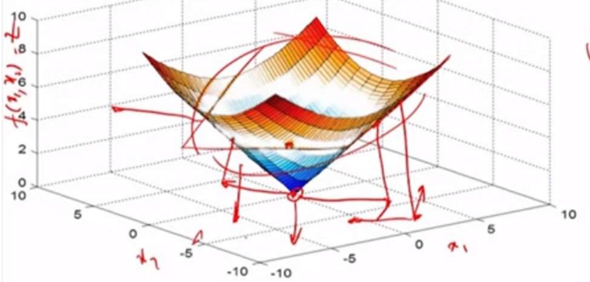
Multivariate optimization – Contour plots

Multivariate optimization

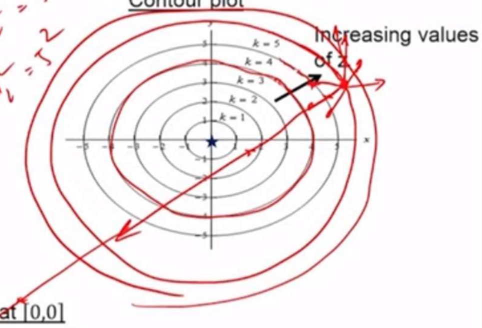
$$z = f(x_1, x_2 \dots x_n)$$

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





Contour plot



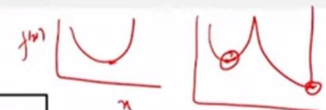
The minimum value of the function is at [0,0]

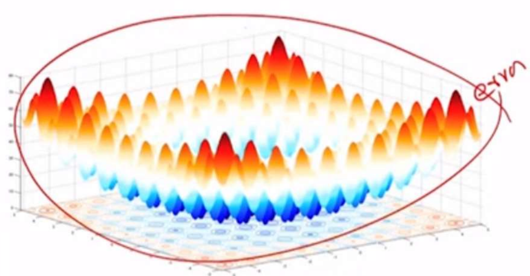
Multivariate optimization – Local and global optimum

Multivariate optimization

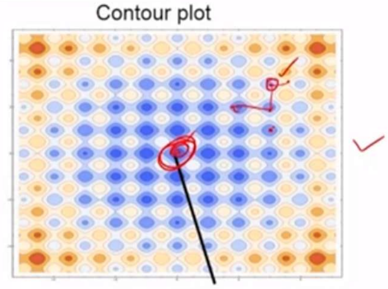
Rastrigin function

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





Contour plot



Global minimum at [0,0]

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

Multivariate optimization – Key ideas

Multivariate optimization

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

$z = f(x)$

$\frac{dz}{dx} = f'(x) = 0$

$\frac{d^2z}{dx^2} = f''(x) > 0$ for minimum

Gradient

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

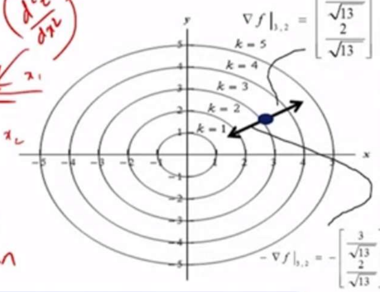
Hessian

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

$n \times n$

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

$$\nabla f|_{3,2} = \begin{bmatrix} \frac{3}{\sqrt{13}} \\ \frac{2}{\sqrt{13}} \end{bmatrix}$$



$$-\nabla f|_{3,2} = -\begin{bmatrix} \frac{3}{\sqrt{13}} \\ \frac{2}{\sqrt{13}} \end{bmatrix}$$

- 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 optimization – Conditions for local optimum

Multivariate optimization

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

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

$\delta^T H \delta = 1 \times 1$ (symmetric, $1\times n$, $n \times n$)

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

$\nabla f^T \alpha$

$-\nabla f^k$

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

$\nabla f(\overline{x^*}) = 0$

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

Multivariate optimization – Summary of conditions

<u>Multivariate optimization</u>

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

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

$\delta^T H \delta > 0$ small $\forall \delta \neq 0$ $\delta = 0$

$H = \nabla^2 f$

H as a positive definite matrix

$\lambda_1, \lambda_2 \ldots, \lambda_n$

> 0

$f(x^*)$

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

Overall Summary – Univariate and multivariate local optimum conditions

<u>Multivariate optimization</u>

$$\boxed{\begin{array}{c} \min_{x} \quad f(x) \\ x \in R \end{array}}$$

<u>Necessary condition for x^* to be the minimizer</u>

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

<u>Sufficient condition</u>

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

$$\boxed{\begin{array}{c} \min_{\overline{x}} \quad f(\overline{x}) \\ \overline{x} \in R^n \end{array}}$$

<u>Necessary condition for $\overline{x^*}$ to be the minimizer</u>

$$\nabla f(\bar{x}^*) = 0$$

<u>Sufficient condition</u>

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

Multivariate optimization – Numerical example

Multivariate optimization

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

First order condition

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

$$\Big\Downarrow \text{ solving}$$

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

Second order condition

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

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

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