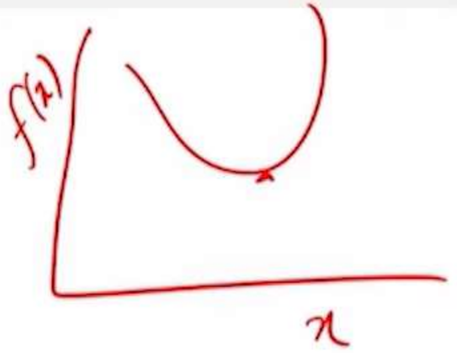
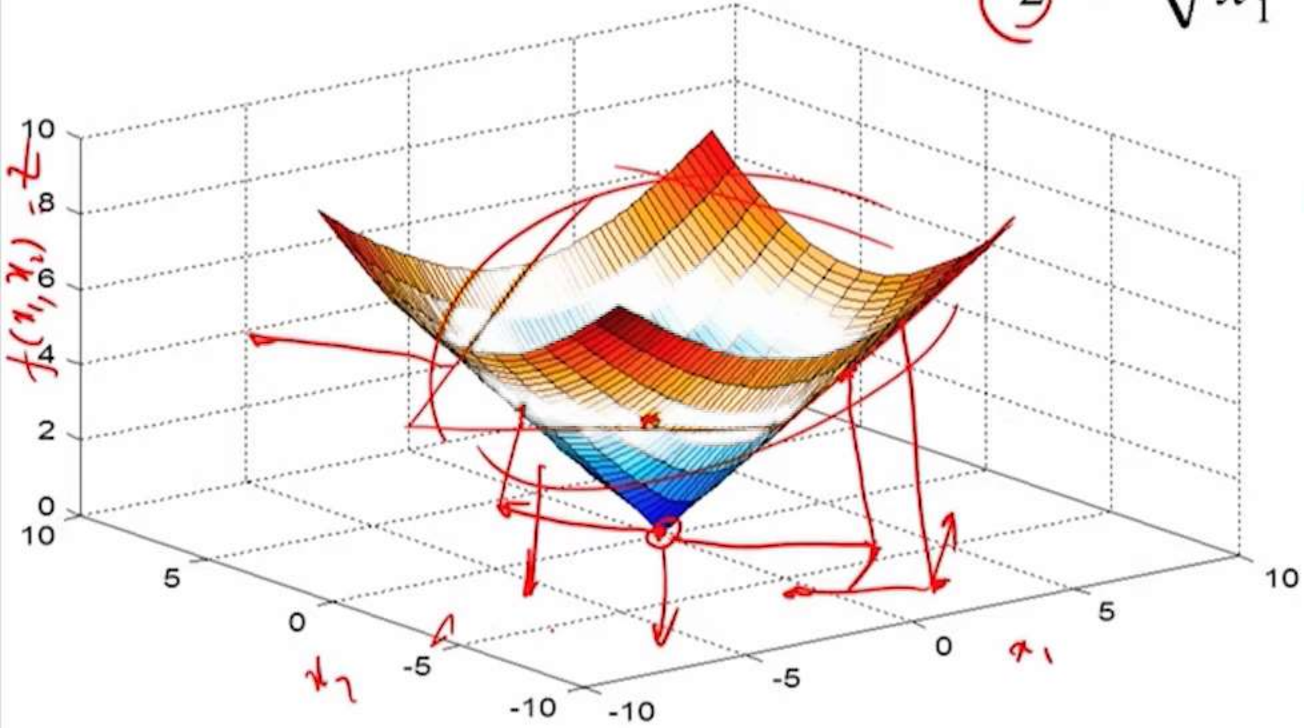
Multivariate optimization – Contour plots

Multivariate optimization

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



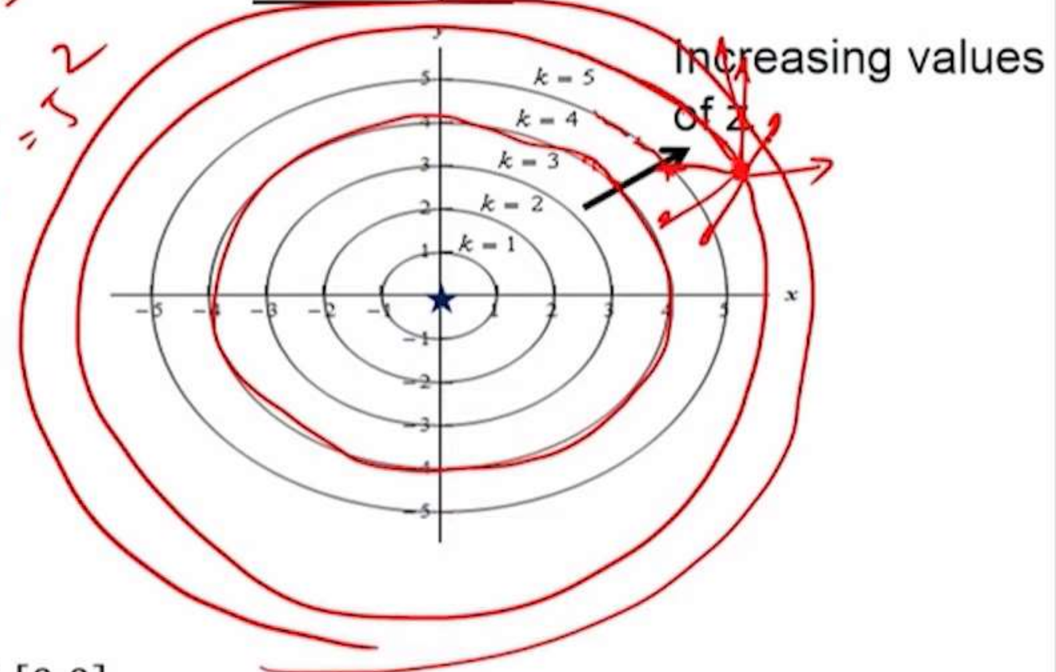
$$z = \sqrt{x_1^2 + x_2^2}$$



$\sqrt{x_1^2 + x_2^2} = 5$

$x_1^2 + x_2^2 = 5^2$

Contour plot



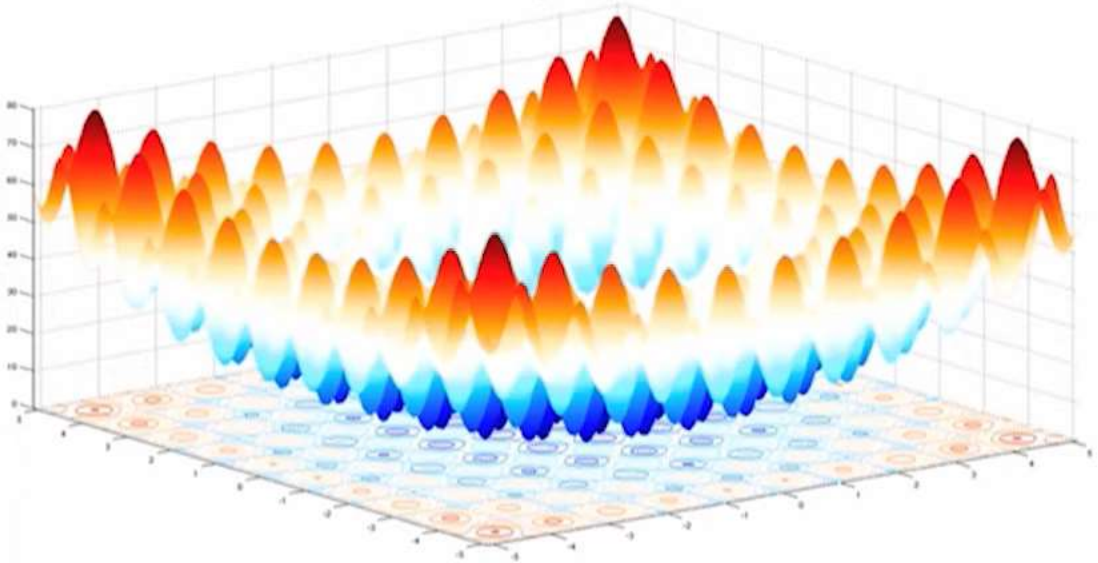
The minimum value of the function is at [0,0]

Multivariate optimization – Local and global optimum

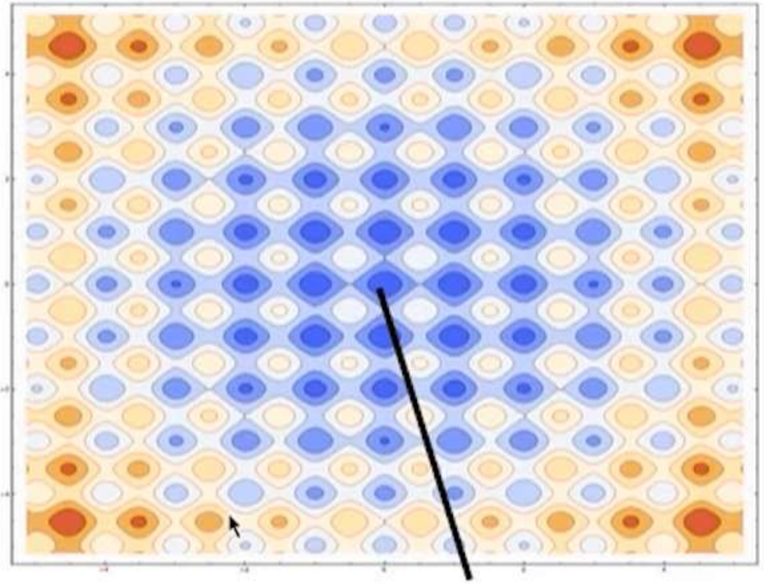
Multivariate optimization

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

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

$z = f(x)$

$\frac{d z}{d x} = f'(x) = 0$

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

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

Gradient

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

Hessian

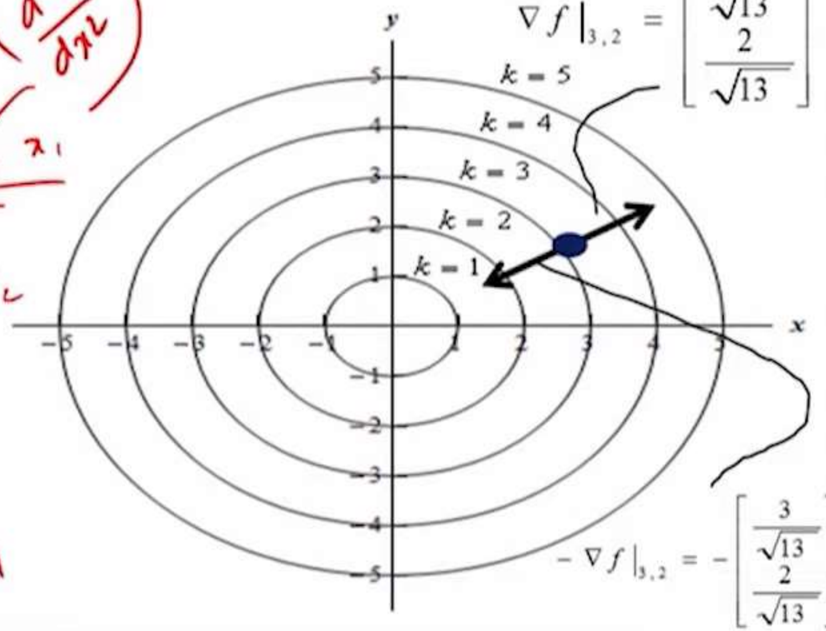
$$\nabla^2 f = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial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

x_1

x_2

$n \times n$

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



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

- 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 $\overline{x^k}$

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

$\delta^T H \delta = 1\times1$; $1\times n$, $n\times n$ symmetric, $n\times 1$

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

$\nabla f^T \alpha$, $-\nabla f^T \alpha$

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

$\neq 0$

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

Multivariate optimization – Summary of conditions

Multivariate optimization

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

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

$f(x^*)$

Condition for Hessian to be positive definite

symmetric

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

$\delta = 0$

$H = \nabla^2 f$

H as a positive definite matrix

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

> 0

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

Multivariate optimization

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

$$\min_{\bar{x}} \; f(\bar{x})$$
$$\bar{x} \in R^n$$

Necessary condition for x^* to be the minimizer

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

Sufficient condition

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

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

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

Sufficient condition

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

Multivariate optimization – Numerical example

Multivariate optimization

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

First order condition

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

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

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

Second order condition

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

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

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