Computational Science:
Computational Methods in Engineering

Mathematical Preliminaries for Multivariable Optimization

Outline

- Multivariable Functions
- Scalar & Vector Fields
- Gradients & Hessians
- Revised Derivative Tests for Multiple Variables
- Conjugate Gradients

Multivariable Functions

Multivariable Functions

General form for a multivariable function:

$$f(\mathbf{x}) = f(x_1, x_2, \ldots, x_N) \qquad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix}$$

<u>Example #1</u>

$$f(x, y) = x + y + xy$$

$$f(\mathbf{x}) \qquad \mathbf{x} = \begin{bmatrix} x \\ y \end{bmatrix}$$

<u>Example #2</u>

$$f(A, \sigma, x) = A \exp\left[-\left(\frac{x}{\sigma}\right)^2\right]$$

$$f(\mathbf{x}) \qquad \mathbf{x} = \begin{bmatrix} A \\ \sigma \\ x \end{bmatrix}$$

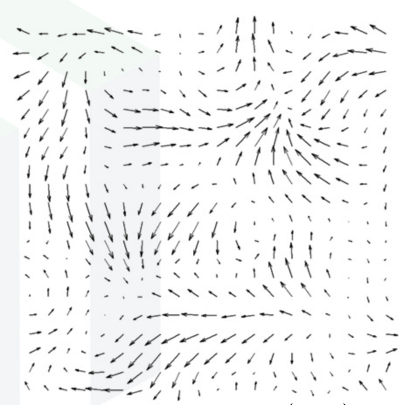
Scalar & Vector Fields

5

Scalar Field Vs. Vector Field

Scalar Field, $f(x,y)$

$\text{magnitude}(x,y,z)$



Vector Field, $\vec{v}(x,y)$

$\text{magnitude}(x,y,z)$

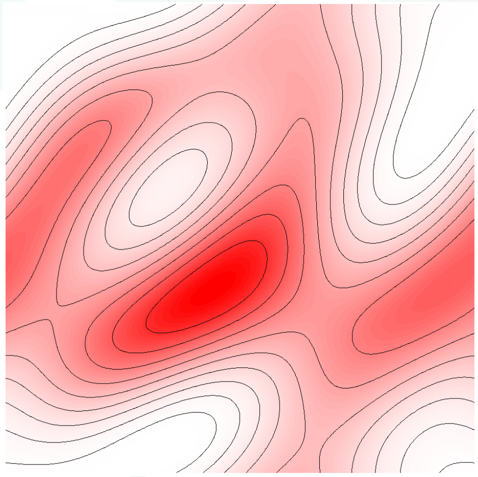
$+\text{direction}(x,y,z)$

EMPossible

6

Isocontour Lines

Isocontour lines trace the paths of equal value. Closely spaced isocontours convey that the function is varying rapidly.



EMPossible

7

Gradient of a Scalar Field (1 of 3)

Start with a scalar field…

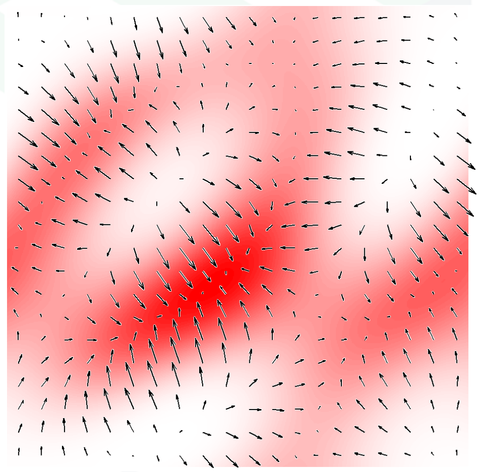
$f(x, y)$

EMPossible

8

Gradient of a Scalar Field (2 of 3)

…then plot the gradient on top of it. Color in background is the original scalar field.



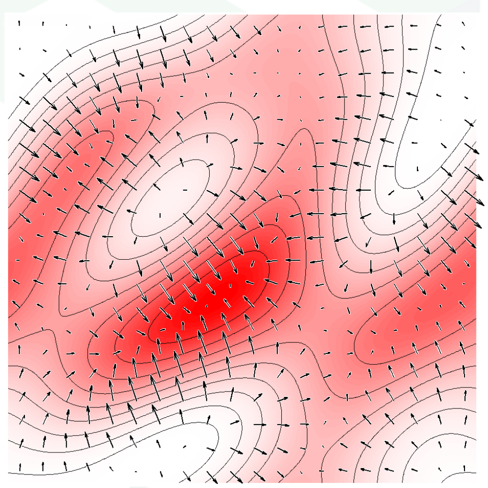
$\nabla f(x,y)$

EMPossible

9

Gradient of a Scalar Field (3 of 3)

The gradient will always be perpendicular to the isocontour lines.



EMPossible

10

Gradients & Hessians

The Multidimensional Gradient

The standard 3D gradient in Cartesian coordiantes is

$$\nabla f(x,y,z) = \frac{\partial f(x,y,z)}{\partial x}\hat{a}_x + \frac{\partial f(x,y,z)}{\partial y}\hat{a}_y + \frac{\partial f(x,y,z)}{\partial z}\hat{a}_z$$

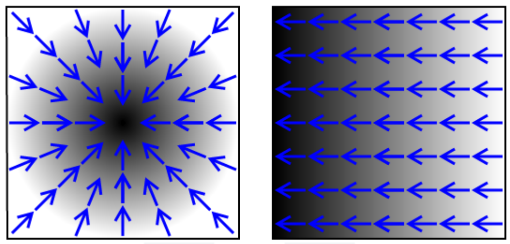
When more dimensions are involved, it is written as

$$\nabla f(\mathbf{x}) = \begin{bmatrix} \dfrac{\partial f(\mathbf{x})}{\partial x_1} \\ \dfrac{\partial f(\mathbf{x})}{\partial x_2} \\ \vdots \\ \dfrac{\partial f(\mathbf{x})}{\partial x_N} \end{bmatrix} \qquad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix}$$

EMPossible

Properties of the Gradient

1. It only makes sense to calculate the gradient of a scalar field*.
2. ∇f points in the direction of the maximum rate of change in f.
3. ∇f at any point is perpendicular to the constant f surface that passes through that point.
4. The gradient points toward big positive numbers in the scalar field.



* It is actually possible to calculate the gradient of a vector field. It is a tensor quantity and is called the Jacobian. It is commonly used in coordinate transformations, but is outside the scope of this course.

EMPossible 13

Numerical Calculation of the Gradient

There may not always be a closed form expression for the function so it may not be possible to calculate an analytical expression for the gradient.

When this is the case, the gradient can still be calculated numerically.

$$\nabla f(\mathbf{x}) = \begin{bmatrix} \dfrac{\partial f(\mathbf{x})}{\partial x_1} \\ \dfrac{\partial f(\mathbf{x})}{\partial x_2} \\ \vdots \\ \dfrac{\partial f(\mathbf{x})}{\partial x_N} \end{bmatrix} \approx \begin{bmatrix} \dfrac{f(x_1+\delta_1)-f(x_1-\delta_1)}{2\delta_1} \\ \dfrac{f(x_2+\delta_2)-f(x_2-\delta_2)}{2\delta_2} \\ \vdots \\ \dfrac{f(x_N+\delta_N)-f(x_N-\delta_N)}{2\delta_N} \end{bmatrix}$$

This can be a very intensive computation!

EMPossible 14

Derivative Tests in Multiple Dimensions

Suppose there is a 2D function $f(x, y)$.

$$\frac{\partial f}{\partial x} = 0 \qquad \frac{\partial^2 f}{\partial x^2} > 0$$

$$\frac{\partial f}{\partial y} = 0 \qquad \frac{\partial^2 f}{\partial y^2} > 0$$

Does this indicate a minimum? No!

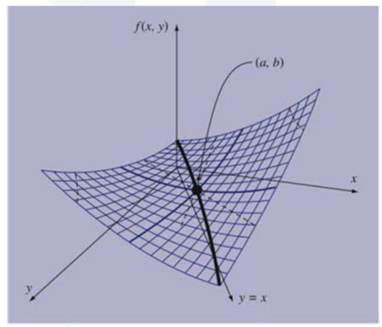


FIGURE 14.8
A saddle point ($x = a$ and $y = b$). Notice that when the curve is viewed along the x and y directions, the function appears to go through a minimum (positive second derivative), whereas when viewed along an axis $x = y$, it is concave downward (negative second derivative).

Figure borrowed from:
Steven Chapra, *Numerical Methods for Engineers*, 7th Ed., McGraw Hill.

EMPossible 15

The Hessian

The Hessian describes "curvature" of multiple variable functions.

The Hessian will be used to determine whether a points is really at a maximum or minimum.

The Hessian is defined as

$$\left[H(f)\right] = \begin{bmatrix} \dfrac{\partial^2 f}{\partial x_1^2} & \dfrac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \dfrac{\partial^2 f}{\partial x_1 \partial x_N} \\ \dfrac{\partial^2 f}{\partial x_2 \partial x_1} & \dfrac{\partial^2 f}{\partial x_2^2} & \cdots & \dfrac{\partial^2 f}{\partial x_2 \partial x_N} \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial^2 f}{\partial x_N \partial x_1} & \dfrac{\partial^2 f}{\partial x_N \partial x_2} & \cdots & \dfrac{\partial^2 f}{\partial x_N^2} \end{bmatrix}$$

In two dimensions, the Hessian is

$$[H] = \begin{bmatrix} \dfrac{\partial^2 f}{\partial x^2} & \dfrac{\partial^2 f}{\partial x \partial y} \\ \dfrac{\partial^2 f}{\partial y \partial x} & \dfrac{\partial^2 f}{\partial y^2} \end{bmatrix}$$

$$\det[H] = \frac{\partial^2 f}{\partial x^2}\frac{\partial^2 f}{\partial y^2} - \left(\frac{\partial^2 f}{\partial x \partial y}\right)^2 \quad \text{determinant}$$

EMPossible 16

Revised Derivative Tests for Multiple Variables

17

Derivative Tests Revised for Multiple Dimensions

If $\det[H] > 0$ and $\frac{\partial^2 f}{\partial x^2} > 0$, then $f(x, y)$ has a local minumum.

If $\det[H] > 0$ and $\frac{\partial^2 f}{\partial x^2} < 0$, then $f(x, y)$ has a local maximum.

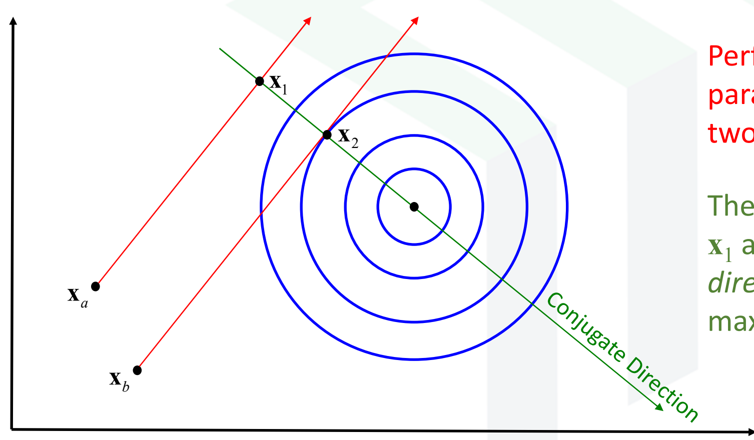
If $\det[H] < 0$ then $f(x, y)$ has a saddle point.

EMPossible

18

Conjugate Gradients

Conjugate Direction



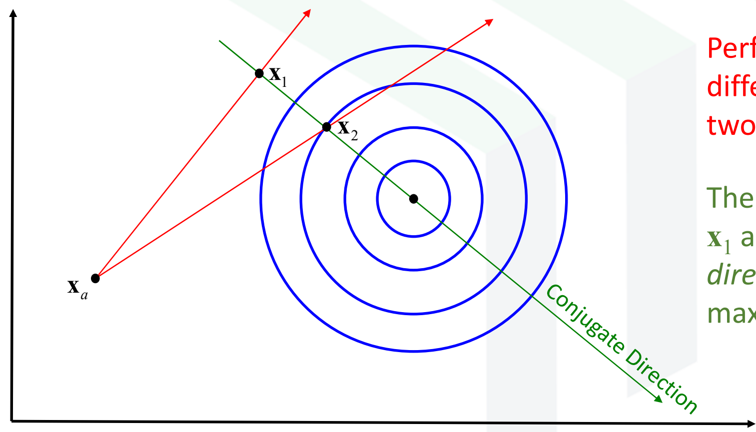
Start at two different points, $\mathbf{x}_a$ and $\mathbf{x}_b$.

Perform 1D optimizations along parallel directions to arrive at the two extremas $\mathbf{x}_1$ and $\mathbf{x}_2$.

The direction of the line connecting $\mathbf{x}_1$ and $\mathbf{x}_2$ is called a *conjugate direction* and is directed toward the maximum.

EMPossible

Alternative Conjugate Direction



Start at common point $\mathbf{x}_a$.

Perform 1D optimizations along two different directions to arrive at the two extremas $\mathbf{x}_1$ and $\mathbf{x}_2$.

The direction of the line connecting $\mathbf{x}_1$ and $\mathbf{x}_2$ is also a *conjugate direction* and is directed toward the maximum.

EMPossible

21