

# 3D Computer Vision - WT 2005

## Multivariate Calculus Reminders

Martin Groher

CAMP - Chair for Computer Aided Medical Procedures

November 9, 2005



- 1 Multivariate Real-Valued Functions
  - Partial Derivative
  - Gradient
  - Hessian
- 2 Multivariate Vector-Valued Functions
  - Jacobian
  - Application: Linear Approximation
- 3 Summary and References
  - Summary
  - References

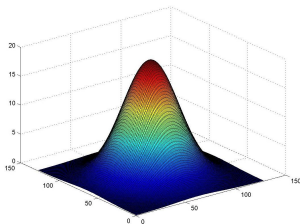
# Why?

- Introduction/reminder of methods and concepts required not only in CV but in almost all applications
- Makes life **much** more easier
- Direct use of the concepts and techniques in:
  - Parameter Estimation
  - Algorithm Analysis
  - Error Propagation
  - Reconstruction
  - Optimization Methods
  - ...

# How?

- “Mathematics is a full meal with many courses - not a buffet”  
(Walter Gander)
- To keep that picture - We will now have some fast food ;-)

# Assumptions



- In the first part (Partial Derivative, Gradient, Hessian) we will consider real-valued functions.
- We will assume that a function  $f : \mathbf{R}^n \rightarrow \mathbf{R}$  is sufficiently nice so that it fulfills all the necessary properties required for anything we want to do with it.

## Definition of the Partial Derivative

- We apply the derivative only to  $x_i$  (the  $i$ -th variable of  $f$ )
- All other variables  $x_j$  with  $i \neq j$  are assumed constant
- The formula is evaluated for an arbitrary  $a \in \mathbf{R}^n$ :

$$\frac{\partial f}{\partial x_i}(a) =: \partial_i f(a)$$

## Properties of the Partial Derivative

- Partial derivatives give the derivatives along the main coordinate system axes
- Higher partial derivatives are also possible:

$$\partial_{xy}f(a) = \partial_y\partial_x f(a) = \frac{\partial}{\partial y} \left( \frac{\partial f}{\partial x}(a) \right)$$

- Schwarz theorem: if  $f$  has an *open* domain and is two times continuously differentiable

$$\frac{\partial}{\partial y} \left( \frac{\partial f}{\partial x}(a) \right) = \frac{\partial}{\partial x} \left( \frac{\partial f}{\partial y}(a) \right)$$

## Definition of the Gradient

Taking all partial derivatives  $\partial_i f(x)$  of a function and putting them together in a vector gives a thing called a gradient:

$$\mathit{grad} f(x) = \begin{bmatrix} \partial_1 f(x) \\ \partial_2 f(x) \\ \vdots \\ \partial_n f(x) \end{bmatrix}$$

Like the partial derivative, the gradient is a function that can be evaluated for all  $x \in \mathbf{R}^n$ .

# Properties of the Gradient

- The gradient  $\text{grad } f(x)$  shows for every point  $x$  in the direction of the steepest slope of the function  $f$  at the point  $x$
- The magnitude of the gradient  $\text{grad } f(x)$  gives the “steepness” of the function  $f$  at the point  $x$
- The gradient can be interpreted as the first derivative of a function  $f : \mathbf{R}^n \rightarrow \mathbf{R}$

# Directional Derivative

- The directional derivative is the generalization of the concept of partial derivatives
- Partial derivatives give the derivatives along the main coordinate system axes
- We want to have a derivative along an arbitrary vector  $v$

## Directional Derivative II

- Given a vector  $v$  with  $\|v\|_2 = 1$
- Then the derivative along  $v$  is

$$\partial_v f(x) = \text{grad } f(x) \cdot v = \sum_{i=1}^n f_{x_i}(x) v_i$$

- The directional derivative can be interpreted as the projection of the gradient  $\text{grad } f(x)$  onto the vector  $v$

## Application: Linear Approximation of Real-Valued-Functions

Given the function  $f : \mathbf{R}^n \rightarrow \mathbf{R}$  and the point  $x_0$  we can approximate  $f$  linearly in the neighbourhood of  $x_0$  by using the gradient:

$$f(x) \approx f(x_0) + \text{grad } f(x_0) \cdot (x - x_0)$$

# Hessian

- For a real-valued function  $f : \mathbf{R}^n \rightarrow \mathbf{R}$  we can define a  $n \times n$  matrix containing second partial derivatives.
- It shall be called the **Hessian** of the function  $f$  and denoted

$$Hess_f(x) = \begin{bmatrix} \partial_{x_1x_1} f(x) & \partial_{x_1x_2} f(x) & \dots & \partial_{x_1x_n} f(x) \\ \partial_{x_2x_1} f(x) & \partial_{x_2x_2} f(x) & \dots & \partial_{x_2x_n} f(x) \\ \vdots & \vdots & \ddots & \vdots \\ \partial_{x_nx_1} f(x) & \partial_{x_nx_2} f(x) & \dots & \partial_{x_nx_n} f(x) \end{bmatrix}$$

# Properties of the Hessian

- The Hessian is a symmetric matrix
- The Hessian can be interpreted as the second derivative for real-valued functions

## Application: Approximation by Taylor Expansion

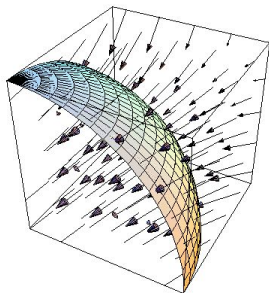
- Given a real-valued function  $f : \mathbf{R}^n \rightarrow \mathbf{R}$
- We can approximate the function in the neighbourhood of  $x_0$  using the Taylor expansion
- Linear Approximation

$$f(x) \approx f(x_0) + \text{grad } f(x_0) \cdot (x - x_0)$$

- Quadratic Approximation

$$f(x) \approx f(x_0) + \text{grad } f(x_0) \cdot (x - x_0) + \frac{1}{2} (x - x_0)^T \text{Hess}_f(x_0) (x - x_0)$$

# New Assumptions



- For the things to come (Jacobian) we will take a look at vector-valued functions.
- i.e. the function  $f$  will be  $f : \mathbf{R}^n \rightarrow \mathbf{R}^m$

image:

[www.math.armstrong.edu/mmacalc/animations/](http://www.math.armstrong.edu/mmacalc/animations/)

# Jacobian

The Jacobian  $J_f(x)$  of the function  $f : \mathbf{R}^n \rightarrow \mathbf{R}^m$  is the generalization of the gradient for vector-valued functions.

$$J_f(x) = \begin{bmatrix} \text{grad } f_1(x) \\ \text{grad } f_2(x) \\ \vdots \\ \text{grad } f_m(x) \end{bmatrix} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(x) & \frac{\partial f_1}{\partial x_2}(x) & \dots & \frac{\partial f_1}{\partial x_n}(x) \\ \frac{\partial f_2}{\partial x_1}(x) & \frac{\partial f_2}{\partial x_2}(x) & \dots & \frac{\partial f_2}{\partial x_n}(x) \\ \vdots & \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1}(x) & \frac{\partial f_m}{\partial x_2}(x) & \dots & \frac{\partial f_m}{\partial x_n}(x) \end{bmatrix}$$

# Properties of the Jacobian

- The Jacobian is itself a function that must be evaluated
- The Jacobian  $J_f(x)$  for  $f : \mathbf{R}^n \rightarrow \mathbf{R}^m$  is a  $m \times n$  matrix.

# Properties of the Jacobian

- The Jacobian  $J_f(x)$  of the function  $f : \mathbf{R}^n \rightarrow \mathbf{R}^m$  can be interpreted as the first derivative of the vector-valued function  $f$
- Thus  $J_f(x)$  can be used for the linear approximation of that function

## Application: Linear Approximation

- The Jacobian  $J(x)$  of the function  $f : \mathbf{R}^n \rightarrow \mathbf{R}^m$  can be interpreted as the linear approximation of that function.
- This can be used for local approximation around some value  $x_0$ .
- Hence in the neighbourhood of  $x_0$  we can set:

$$f(x) \approx \tilde{f}(x) = f(x_0) + J(x_0) \cdot (x - x_0)$$

# Summary

- Operations on real-valued functions:
  - Partial Derivative
  - Gradient
  - Hessian
- Operations on vector-valued functions:
  - Jacobian
- Most often used for Linear or Quadratic Local Approximation (Taylor Expansion)
- Used in CV for:
  - Estimation
  - Optimization (which itself is also used e.g. in Estimation)
  - Algorithm Analysis and Error Propagation
  - Image Processing

# References

- Vachenauer. *Höhere Mathematik 1*