Calculus II

Essential Terminology for Multivariate Vector Functions

Uwe Naumann

Informatik 12: Software and Tools for Computational Engineering
RWTH Aachen University
Contents

Objective and Learning Outcomes

Continuity

Differentiability

Chain Rule

DAG

Directional Derivative

Taylor Series

Summary and Next Steps
Outline

Objective and Learning Outcomes

Continuity

Differentiability

Chain Rule

DAG

Directional Derivative

Taylor Series

Summary and Next Steps
Objective

- Introduction to essential vector calculus

Learning Outcomes

- You will understand
  - continuity
  - differentiability
  - gradient, Jacobian, Hessian
  - chain rule of differential calculus
  - partial, total, directional derivative
  - directional derivative DAG $\times$ vector product
  - Taylor series.
Outline

Objective and Learning Outcomes

Continuity

Differentiability

Chain Rule

DAG

Directional Derivative

Taylor Series

Summary and Next Steps
Let $\mathbb{R}^n$ be the domain of the multivariate scalar function $f : \mathbb{R}^n \rightarrow \mathbb{R}$. The function $f$ is **continuous** at a point $\tilde{x} \in \mathbb{R}^n$ if

$$\lim_{x \to \tilde{x}} f(x) = f(\tilde{x})$$

A multivariate vector function

$$F = \begin{pmatrix} f_1 \\ \vdots \\ f_m \end{pmatrix} : \mathbb{R}^n \rightarrow \mathbb{R}^m$$

is continuous if and only if all its **component functions** $f_i, i = 1, \ldots, m$ are continuous.
Outline

Objective and Learning Outcomes

Continuity

Differentiability

Chain Rule

DAG

Directional Derivative

Taylor Series

Summary and Next Steps
Calculus II
Differentiability

Let \( \mathbb{R}^n \) be the domain of the multivariate scalar function \( f : \mathbb{R}^n \to \mathbb{R} \). The function \( f \) is differentiable at point \( \tilde{x} \in \mathbb{R}^n \) if there is a vector \( f' \in \mathbb{R}^n \) such that

\[
f(\tilde{x} + \Delta x) = f(\tilde{x}) + f' \cdot \Delta x + r
\]

with asymptotically vanishing remainder \( r = r(\tilde{x}, \Delta x) \in \mathbb{R} \), such that

\[
\lim_{\Delta x \to 0} \frac{r}{\|\Delta x\|_2} = 0,
\]

where \( \|v\|_2 \equiv \sqrt{v^T \cdot v} = \sqrt{\sum_{i=0}^{n-1} v_i^2} \) denotes the Euclidean norm of the vector \( v \in \mathbb{R}^n \).

\[
f' = f'(x) = \frac{df}{dx}(x) : \mathbb{R}^n \to \mathbb{R}^n
\]

is called the gradient of \( f \).
Let $f : \mathbb{R}^n \to \mathbb{R}$ be differentiable at $\tilde{x} \in \mathbb{R}^n$. Then

$$f'(\tilde{x}) = \begin{pmatrix}
\frac{dy}{dx_0} \\
\vdots \\
\frac{dy}{dx_{n-1}}
\end{pmatrix}$$

where $y = f(x)$ and

$$\frac{dy}{dx_i} = \frac{dy}{dx_i}(\tilde{x}) = \lim_{\Delta x \to \pm 0} \frac{f(\tilde{x} + \Delta x \cdot e^i) - f(\tilde{x})}{\Delta x} < \infty$$

with the $i$-th Cartesian basis vector in $\mathbb{R}^n$ denoted by $e^i$. 

Calculus II, info@stce.rwth-aachen.de
Let

$$y = f(x) = e^{\sin(\|x\|_2^2)} = e^{\sin(x^T \cdot x)} = e^{\sin(\sum_{i=0}^{n-1} x_i^2)}$$

Differentiation wrt. $x$ yields the gradient

$$f'(x) = \left(2 \cdot x_j \cdot \cos\left(\sum_{i=0}^{n-1} x_i^2\right) \cdot e^{\sin(\sum_{i=0}^{n-1} x_i^2)}\right)_{j=0, \ldots, n-1}$$
Let $\mathbb{R}^n$ be the domain of the multivariate vector function $F : \mathbb{R}^n \to \mathbb{R}^m$. The function $F$ is differentiable at point $\tilde{x} \in \mathbb{R}^n$ if there is a matrix $F' \in \mathbb{R}^{m \times n}$ such that

$$F(\tilde{x} + \Delta x) = F(\tilde{x}) + F' \cdot \Delta x + r$$

with asymptotically vanishing remainder $r = r(\tilde{x}, \Delta x) \in \mathbb{R}^m$, such that

$$\lim_{\Delta x \to 0} \frac{\|r\|_2}{\|\Delta x\|_2} = 0 .$$

The matrix

$$F' = F'(x) = \frac{dF}{dx}(x) : \mathbb{R}^n \to \mathbb{R}^{m \times n}$$

is called the Jacobian of $F$. 
Let $F : \mathbb{R}^n \to \mathbb{R}^n$ be differentiable at $\tilde{x} \in \mathbb{R}^n$. Then

$$F'(\tilde{x}) = \begin{pmatrix} \frac{dy_0}{dx_0} & \cdots & \frac{dy_0}{dx_{n-1}} \\ \vdots & \ddots & \vdots \\ \frac{dy_{n-1}}{dx_0} & \cdots & \frac{dy_{n-1}}{dx_{n-1}} \end{pmatrix}$$

where $y_j = F_j(x)$ and

$$\frac{dy_j}{dx_i} = \frac{dy_j}{dx_i}(\tilde{x}) = \lim_{\Delta x \to \pm 0} \frac{F_j(\tilde{x} + \Delta x \cdot e^i) - F_j(\tilde{x})}{\Delta x} < \infty.$$
Let \( f : \mathbb{R}^n \to \mathbb{R} \) be differentiable at \( \tilde{x} \in \mathbb{R}^n \). It is twice differentiable at \( \tilde{x} \) if \( f' : \mathbb{R}^n \to \mathbb{R}^n \) is differentiable at \( \tilde{x} \).

The matrix

\[
\begin{pmatrix}
\frac{d^2 y}{dx_0^2} & \cdots & \frac{d^2 y}{dx_0 dx_{n-1}} \\
\vdots & \ddots & \vdots \\
\frac{d^2 y}{dx_{n-1} dx_0} & \cdots & \frac{d^2 y}{dx_{n-1}^2}
\end{pmatrix}
\]

is called the Hessian matrix of \( f \) at point \( \tilde{x} \).

If \( f' \) is continuous [at some point, within some subdomain], then \( f \) is called continuously differentiable [at this point, within this subdomain].

If \( f \) is twice continuously differentiable at \( \tilde{x} \), then its Hessian is symmetric, i.e.,

\[
\begin{pmatrix}
\frac{d^2 y}{dx_0^2} & \cdots & \frac{d^2 y}{dx_0 dx_{n-1}} \\
\vdots & \ddots & \vdots \\
\frac{d^2 y}{dx_{n-1} dx_0} & \cdots & \frac{d^2 y}{dx_{n-1}^2}
\end{pmatrix}
= \begin{pmatrix}
\frac{d^2 y}{dx_{n-1} dx_0} & \cdots & \frac{d^2 y}{dx_{n-1}^2} \\
\vdots & \ddots & \vdots \\
\frac{d^2 y}{dx_0^2} & \cdots & \frac{d^2 y}{dx_0 dx_{n-1}}
\end{pmatrix}
\]

Hessians of multivariate vector functions \( F : \mathbb{R}^n \to \mathbb{R}^m \) are 3-tensors. So are third derivatives of \( f \), and so forth.
Outline

Objective and Learning Outcomes

Continuity

Differentiability

Chain Rule

DAG

Directional Derivative

Taylor Series

Summary and Next Steps
Let \( y = F(x) : \mathbb{R}^n \to \mathbb{R}^m \) be such that

\[
y = F(x) = F_2(F_1(x), x) = F_2(z, x)
\]

with (continuously) differentiable \( F_1 : \mathbb{R}^n \to \mathbb{R}^p \) and \( F_2 : \mathbb{R}^p \times \mathbb{R}^n \to \mathbb{R}^m \).

Then \( F \) is continuously differentiable over \( \mathbb{R}^n \) and

\[
\frac{dF}{dx}(\tilde{x}) = \frac{dF_2}{dx}(\tilde{z}, \tilde{x}) = \frac{dF_2}{dz}(\tilde{z}, \tilde{x}) \cdot \frac{dF_1}{dx}(\tilde{x}) + \frac{\partial F_2}{\partial x}(\tilde{z}, \tilde{x})
\]

for all \( \tilde{x} \in \mathbb{R}^n \) and \( \tilde{z} = F_1(\tilde{x}) \).

Notation: \( \frac{\partial F_2}{\partial x} \) partial derivative; \( \frac{dF_2}{dx} \) total derivative
Chain Rule
Directed Acyclic Graph

A composite function \( y = F(x) \) such as

\[
\begin{align*}
  z &= F_1(x) \\
  y &= F_2(z, x)
\end{align*}
\]

induces a directed acyclic graph (DAG) \( G = (V, E) \) with vertices in \( V \) representing variables (e.g, \( x, z \) and \( y \)) and with local (partial) derivatives associated with the edges in \( E \).

\[
F'(x) \equiv \frac{dy}{dx} = \sum_{\text{path} \in \text{DAG}} \prod_{(i,j) \in \text{path}} \frac{\partial v_j}{\partial v_i} = \frac{\partial y}{\partial x} + \frac{\partial y}{\partial z} \cdot \frac{\partial z}{\partial x} = \frac{\partial y}{\partial x} + \frac{dy}{dz} \cdot \frac{dz}{dx}
\]
Chain Rule

Directional Derivative

The directional derivative (Jacobian $\times$ vector product)

$$y^{(1)} = \frac{dF}{dx}(\tilde{x}) \cdot x^{(1)}$$

of $y = F(x)$ at $\tilde{x}$ can be represented as the derivative of $y = y(x(\dot{c}))$ with respect to (wrt.) an auxiliary variable $\dot{c} \in \mathbb{R}$ at $\tilde{x}$ such that

$$\frac{dx}{d\dot{c}} = x^{(1)}.$$

The chain rule yields

$$y^{(1)} \equiv \frac{dF}{d\dot{c}} = \frac{dF}{dx}(\tilde{x}) \cdot \frac{dx}{d\dot{c}} = \frac{dF}{dx}(\tilde{x}) \cdot x^{(1)}.$$

Directional derivatives are marked with the superscript $(1)$. 
Outline

Objective and Learning Outcomes

Continuity

Differentiability

Chain Rule

DAG

Directional Derivative

Taylor Series

Summary and Next Steps
In scientific computing the multivariate vector functions

\[ F : \mathbb{R}^n \rightarrow \mathbb{R}^m : y = F(x) \]

of interest are implemented as differentiable computer programs.

Such programs decompose into sequences of \( q = p + m \) differentiable elemental functions \( \varphi_j \) evaluated as a single assignment code\(^1\)

\[ v_j = \varphi_j(v_{k})_{k \prec j} \quad \text{for } j = n, \ldots, n + q - 1 \]

and where \( v_i = x_i \) for \( i = 0, \ldots, n - 1 \), \( y_k = v_{n+p+k} \) for \( k = 0, \ldots, m - 1 \) and \( k \prec j \) if \( v_k \) is an argument of \( \varphi_j \).

A DAG \( G = (V, E) \) is induced. Partial derivatives of the elemental functions wrt. their arguments are associated with the corresponding edges.

\(^1\)Variables are written once.
Single Assignment Code ⇒ DAG

Example

\[ y = f(x) = e^{\sin(\|x\|_2^2)} = e^{\sin(x^T \cdot x)} = e^{\sin(\sum_{i=0}^{n-1} x_i^2)}, \quad n = 2 \]

\[ \begin{align*}
    v_0 &= x_0 \\
    v_1 &= x_1 \\
    v_2 &= v_0^2; & \quad & \frac{dv_2}{dv_0} = 2 \cdot v_0 \\
    v_3 &= v_1^2; & \quad & \frac{dv_3}{dv_1} = 2 \cdot v_1 \\
    v_4 &= v_2 + v_3; & \quad & \frac{dv_4}{dv_2} = \frac{dv_4}{dv_3} = 1 \\
    v_5 &= \sin(v_4); & \quad & \frac{dv_5}{dv_4} = \cos(v_4) \\
    v_6 &= e^{v_5}; & \quad & \frac{dv_6}{dv_5} = v_6 \\
    y &= v_6
\end{align*} \]
Outline

Objective and Learning Outcomes

Continuity

Differentiability

Chain Rule

DAG

Directional Derivative

Taylor Series

Summary and Next Steps
The DAG $G = G(\tilde{x})$ of $F$ induces a **linear mapping** (generalized Jacobian $\times$ Vector Product)

$$G : \mathbb{R}^n \rightarrow \mathbb{R}^m : \quad y^{(1)} = G \cdot x^{(1)}$$

defined by the chain rule applied to $F(x(\dot{c}))$ at $x = \tilde{x}$ and for

$$\frac{dx}{dc} \equiv x^{(1)} \in \mathbb{R}^n .$$

This **DAG $\times$ vector** product is evaluated as

$$v_i^{(1)} = \sum_{j<i} \frac{d\varphi_i(v_k)_{k<i}}{dv_j} \cdot v_j^{(1)} \quad \text{for } i = n, \ldots, n + q - 1$$

and where $v_i^{(1)} = x_i^{(1)}$ for $i = 0, \ldots, n - 1$ and $y_k^{(1)} = v_{p+k}^{(1)}$ for $k = 0, \ldots, m - 1$. 
DAG $\times$ Vector Product

Example

\[ y = f(x) = e^{\sin(\|x\|_2^2)} = e^{\sin(x^T \cdot x)} = e^{\sin(\sum_{i=0}^{n-1} x_i^2)}, \quad n = 2 \]

\[ \begin{align*}
    v_0 &= x_0 & v_0^{(1)} &= x_0^{(1)} \\
    v_1 &= x_1 & v_1^{(1)} &= x_1^{(1)} \\
    v_2 &= v_0^2 & v_2^{(1)} &= 2 \cdot v_0 \cdot v_0^{(1)} \\
    v_3 &= v_1^2 & v_3^{(1)} &= 2 \cdot v_1 \cdot v_1^{(1)} \\
    v_4 &= v_2 + v_3 & v_4^{(1)} &= v_2^{(1)} + v_3^{(1)} \\
    v_5 &= \sin(v_4) & v_5^{(1)} &= \cos(v_4) \cdot v_4^{(1)} \\
    v_6 &= e^{v_5} & v_6^{(1)} &= v_6 \cdot v_5^{(1)} \\
    y &= v_6 & y^{(1)} &= v_6^{(1)}
\end{align*} \]
Outline

Objective and Learning Outcomes

Continuity

Differentiability

Chain Rule

DAG

Directional Derivative

Taylor Series

Summary and Next Steps
Taylor Series

$n$D Case

- $f : \mathbb{R}^n \rightarrow \mathbb{R}$

\[
f(x + \Delta x) = f(x) + f'(x)^T \cdot \Delta x + \frac{1}{2} \cdot \Delta x^T \cdot f''(x) \cdot \Delta x + O(\|\Delta x\|_2^3)
\]

- $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$

\[
F(x + \Delta x) = F(x) + F'(x) \cdot \Delta x + O(\|\Delta x\|_2^2)
\]

Higher-order terms are omitted to avoid tensor notation.
A function $F : \mathbb{R}^n \to \mathbb{R}^m$ is linear if

- $F(a + b) = F(a) + F(b)$
- $F(\alpha \cdot a) = \alpha \cdot F(a)$

for all $a, b \in \mathbb{R}^n$ and $\alpha \in \mathbb{R}$.

Example: $F(x) = M \cdot x$ with $M \in \mathbb{R}^{m \times n}$ is linear.

$$F(a + b) = M \cdot (a + b) = M \cdot a + M \cdot b = F(a) + F(b)$$
$$F(\alpha \cdot a) = M \cdot \alpha \cdot a = \alpha \cdot M \cdot a = \alpha \cdot F(a)$$

Functions $F(x) = M \cdot x + v$ with $v \in \mathbb{R}^m$ are called affine.

Affine functions define linear systems $m = n$ as well as linear least-squares problems $m \neq n$. 
Convexity and Concavity

\( f : \mathbb{R}^n \rightarrow \mathbb{R} \)

A function \( f : \mathbb{R}^n \rightarrow \mathbb{R} \) is convex if and only if its Hessian \( f'' \) is positive semi-definite for all \( x \in \mathbb{R}^n \), i.e., \( \forall 0 \neq v \in \mathbb{R}^n \)

\[ v^T \cdot f''(x) \cdot v \geq 0. \]

One can show that \( f \) is strictly convex over \( \mathbb{R}^n \) if \( f'' \) is positive definite for all \( x \in \mathbb{R}^n \), i.e,

\[ v^T \cdot f''(x) \cdot v > 0. \]

The other direction does not hold in general.

Similarly, concavity is defined in terms of negative (semi-)definiteness of the Hessian.

The concepts can be generalized for multivariate vector functions \( F : \mathbb{R}^n \rightarrow \mathbb{R}^m \).
Outline

Objective and Learning Outcomes

Continuity

Differentiability

Chain Rule

DAG

Directional Derivative

Taylor Series

Summary and Next Steps
Calculus II

Summary and Next Steps

Summary

▶ continuity
▶ differentiability
▶ gradient, Jacobian, Hessian
▶ chain rule of differential calculus
▶ partial, total, directional derivative
▶ directional derivative as \( \text{DAG} \times \text{vector product} \)
▶ Taylor series

Next Steps

▶ practice \( \text{DAG} \times \text{vector product} \)
▶ Continue the course to find out more ...