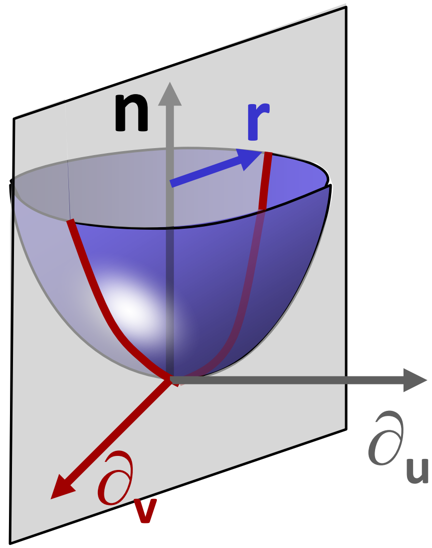
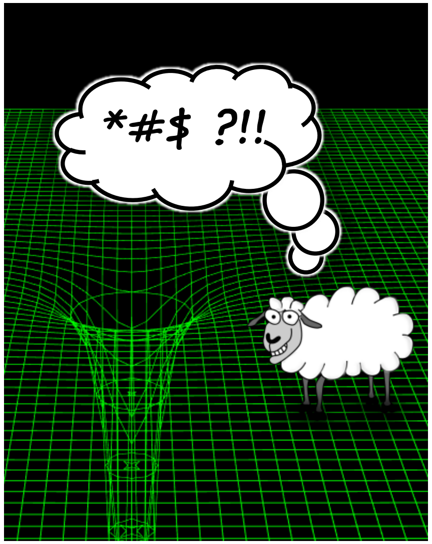
Modelling 1

SUMMER TERM 2020





LECTURE 4

Calculus

Michael Wand · Institut für Informatik · Michael.Wand@uni-mainz.de

Calculus

Overview

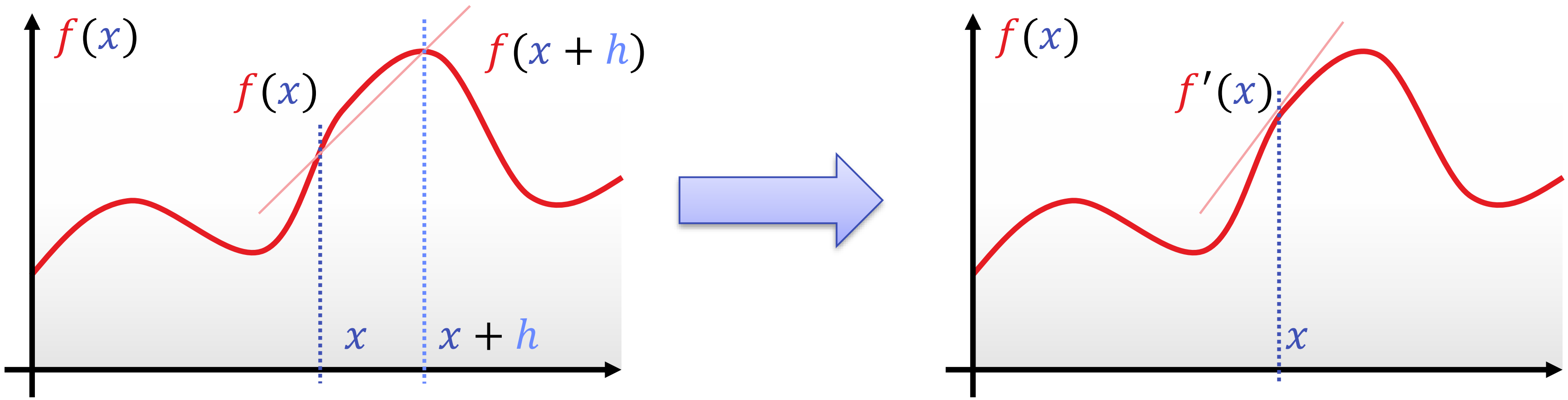
Topics

- Differentiation
- Smoothness
- Multi-dimensional derivatives
- Integration

Linear Approximation

Approx. Slope

Limit



Derivative of a Function

Derivative

$$\frac{d}{dx} f(x) := \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}$$

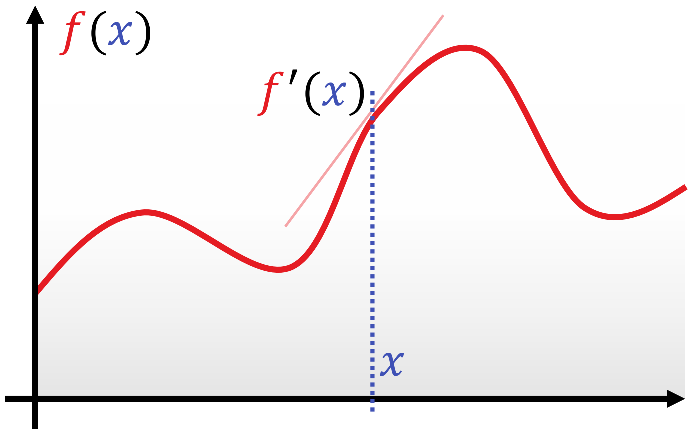
Alternative Notation:

$$\frac{d}{dx} f(x) = \underbrace{f'(x)}_{\text{variable from context}} = \underbrace{\dot{f}(x)}_{\text{time variable}}$$

$$\underbrace{\frac{d^k}{dt^k} f(x) = f^{(k)}(x)}_{\text{repeated differentiation (higher order derivatives)}}$$

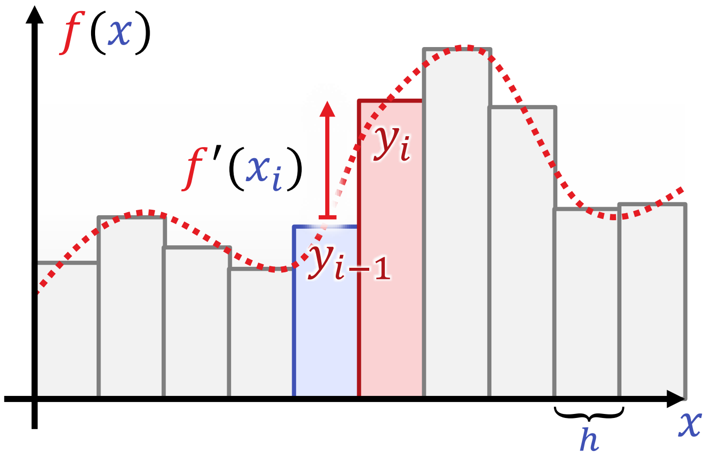
Discrete Analogy

Function *f*



tangent slope

Think of this:



neighborhood differences

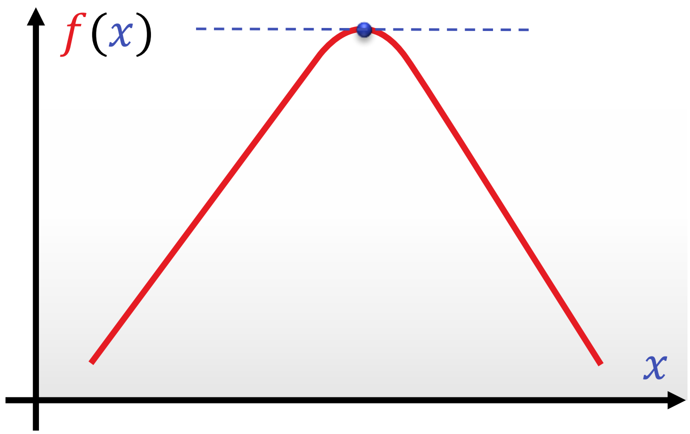
$$f : \mathbb{R} \to \mathbb{R}$$

$$f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}$$

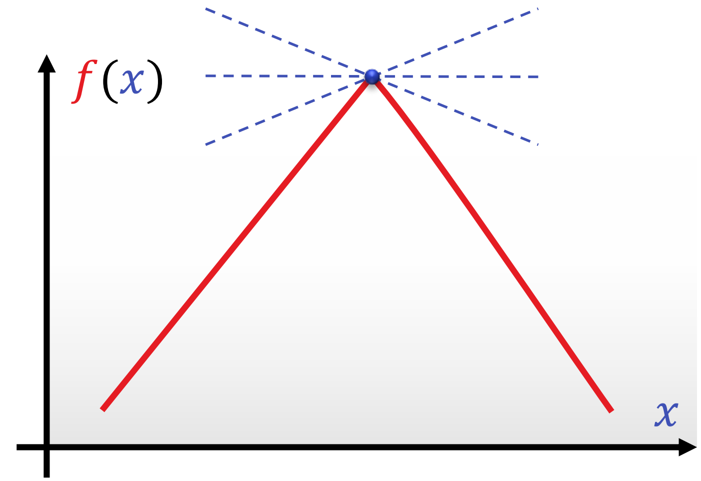
$$f = (y_1, \dots, y_n)$$

$$f'(x_i) \approx \frac{y_i - y_{i-1}}{h}$$

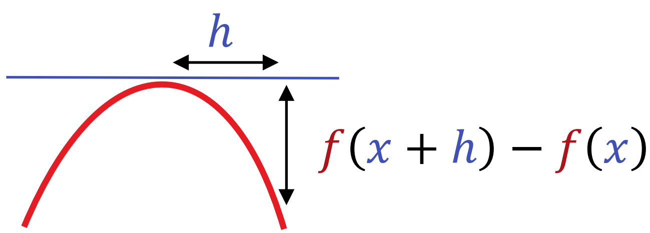
Continuous Case: Differentiability



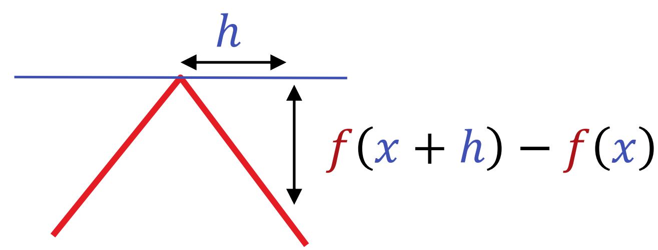
differentiable



non-differentiable



$$\frac{f(x+h)-f(x)}{h} \rightarrow \text{fixed value}$$



$$\frac{f(x+h)-f(x)}{h} \rightarrow \{+1,-1\}$$

Smoothness

Function Classes

- C^0 - “Continuous”:

$$f(x) = \lim_{y \to x} f(y)$$ (limit exists)

- “Differentiable”

$$f'(x) = \lim_{y \to x} f(y)/|y - x|$$ (limit exists)

- “Differentiable” ⇒ “Continuous” (but not vice-versa)

Smoothness

Function Classes

- C^1 - "smoothly differentiable":

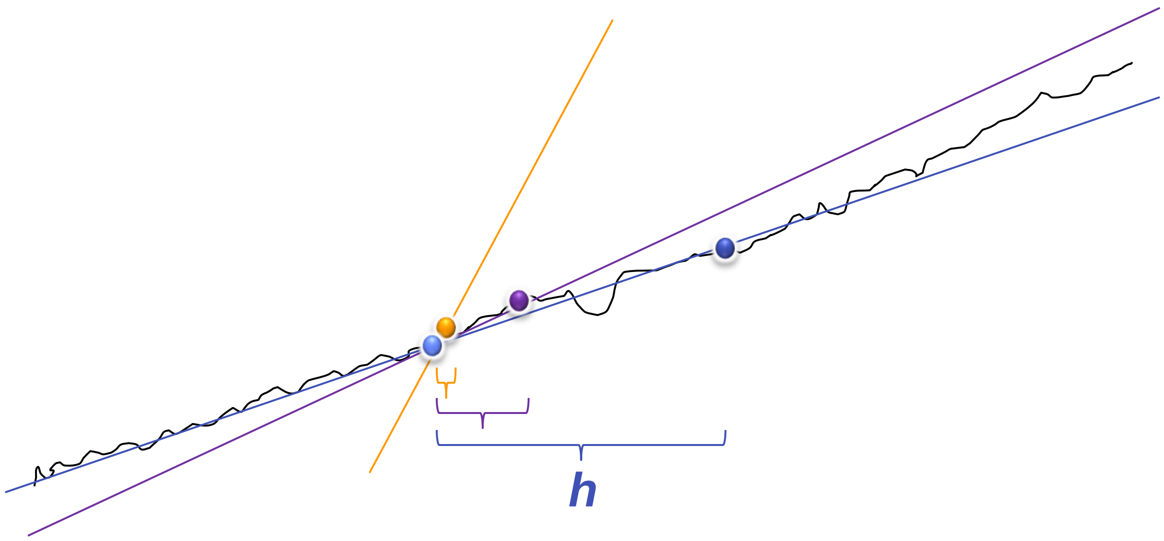
  f' exists and is continuous
- C^2 - "twice smoothly differentiable":

  f'' exists and is continuous
- C^k - "k-fold smoothly differentiable":

  $f^{(k)}$ exists and is continuous
- C^∞ - "smooth":

  f permits any order of differentiation

Differentiation is Ill-posed!

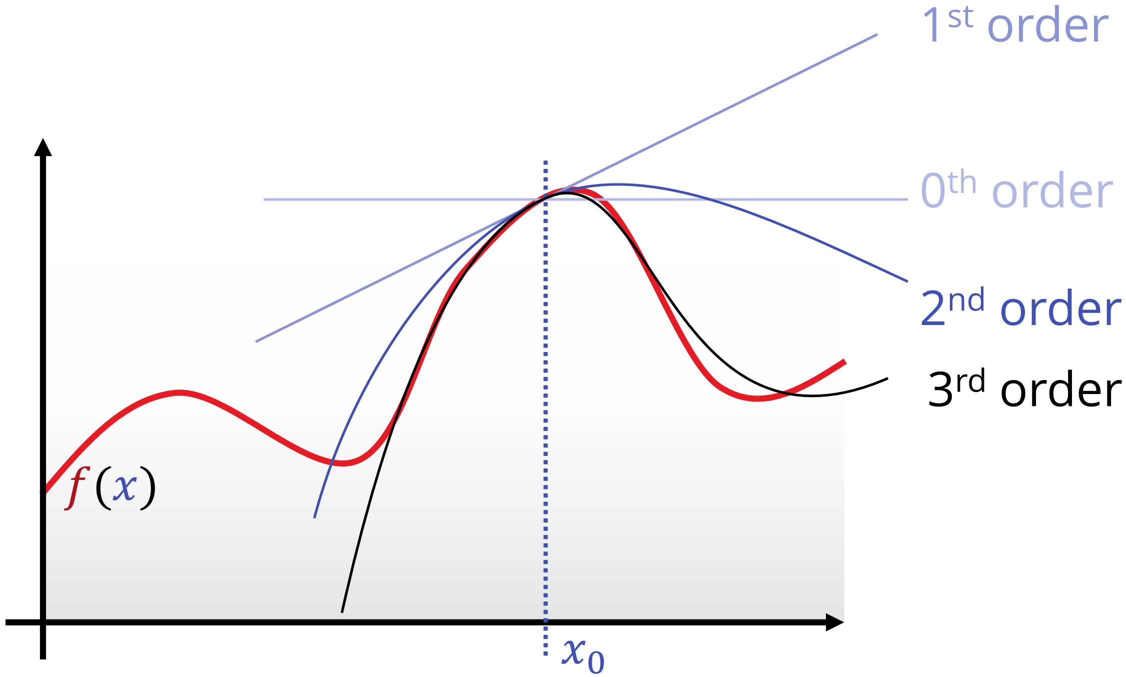


Regularization

- Numerical differentiation needs regularization
  - Higher order is more problematic
- Finite differences (larger *h*)
- Averaging (polynomial fitting) over finite domain

Taylor Approximation

Taylor Approximation to f



(schematic, not actually computed)

Taylor's Series

Local approximation:

$$f(x) \approx f(x_0)$$
$$+ \left[\frac{d}{dx} f(x_0)\right] (x - x_0)$$
$$+ \frac{1}{2}\left[\frac{d^2}{dx^2} f(x_0)\right] (x - x_0)^2$$
$$+ \cdots$$
$$+ \frac{1}{k!}\left[\frac{d^k}{dx^k} f(x_0)\right] (x - x_0)^k$$
$$+ \mathcal{O}\left((x - x_0)^{k+1}\right)$$

When is it useful?

- Good local approximation for C^k functions" (local convergence)
- Converges globally for all holomorphically extensible functions C^ω

Rule of Thumb

Derivatives and Polynomials

- Polynomial: $f(x) = c_0 + c_1 x + c_2 x^2 + c_3 x^3 \ldots$
  - 0th-order derivative: $f(0) = c_0$
  - 1st-order derivative: $f'(0) = c_1$
  - 2nd-order derivative: $f''(0) = 2c_2$
  - 3rd-order derivative: $f'''(0) = 6c_3$
  - ...

Rule of Thumb:

- Derivatives correspond to polynomial coefficients
- Estimate derivatives $\leftrightarrow$ polynomial fitting
- Same in multi-variate case

Multivariate Case

Scalar function

$$f: \mathbb{R}^n \to \mathbb{R}$$

Taylor series

$$\begin{aligned} f(\mathbf{x}) \approx\ & f(\mathbf{x}_0) \\ & + \nabla f(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0) \\ & + \frac{1}{2} (\mathbf{x} - \mathbf{x}_0)^{\mathrm{T}} \cdot \mathbf{H}_f(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0) \\ & + \cdots \end{aligned}$$

Multivariate Derivatives

Partial Derivative

Multivariate Notation

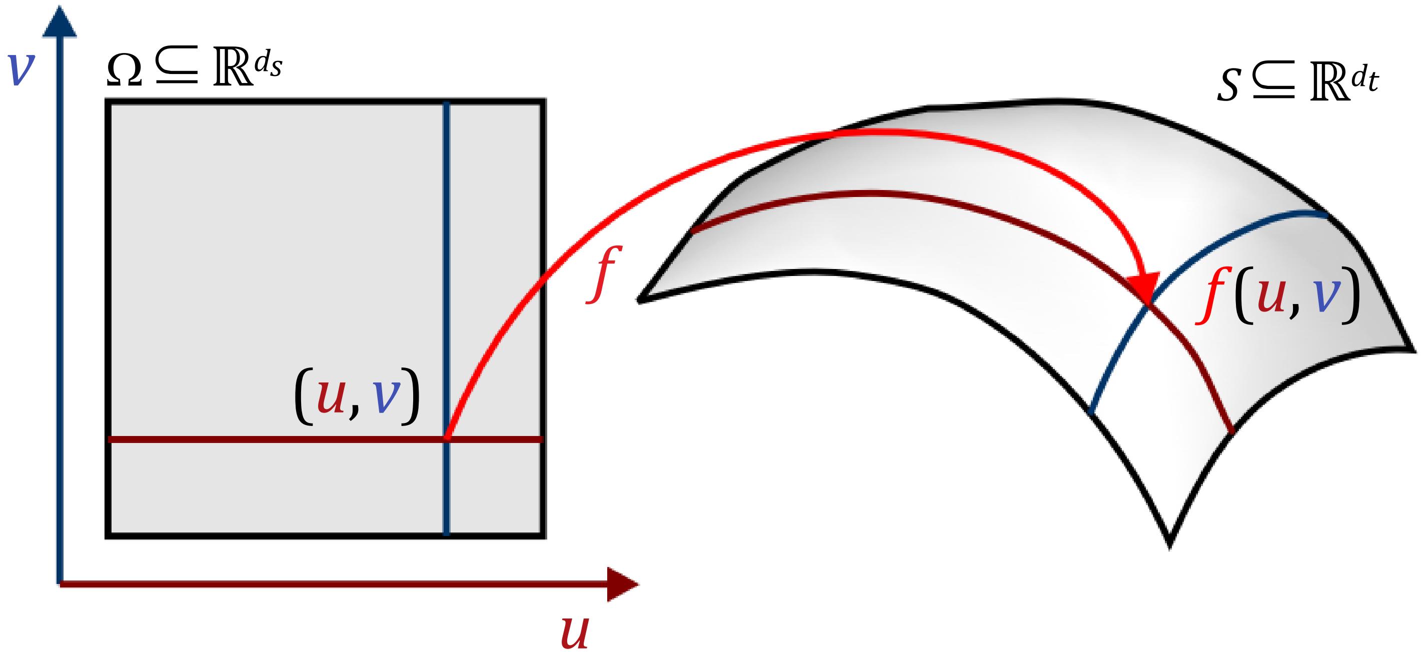
use curly-d

$$\frac{\partial}{\partial x_k} f(x_1, \dots, x_k, \dots, x_n) := \lim_{h \to 0} \frac{f(x_1, \dots, x_k + h, \dots, x_n) - f(x_1, \dots, x_k, \dots, x_n)}{h}$$

Alternative notations:

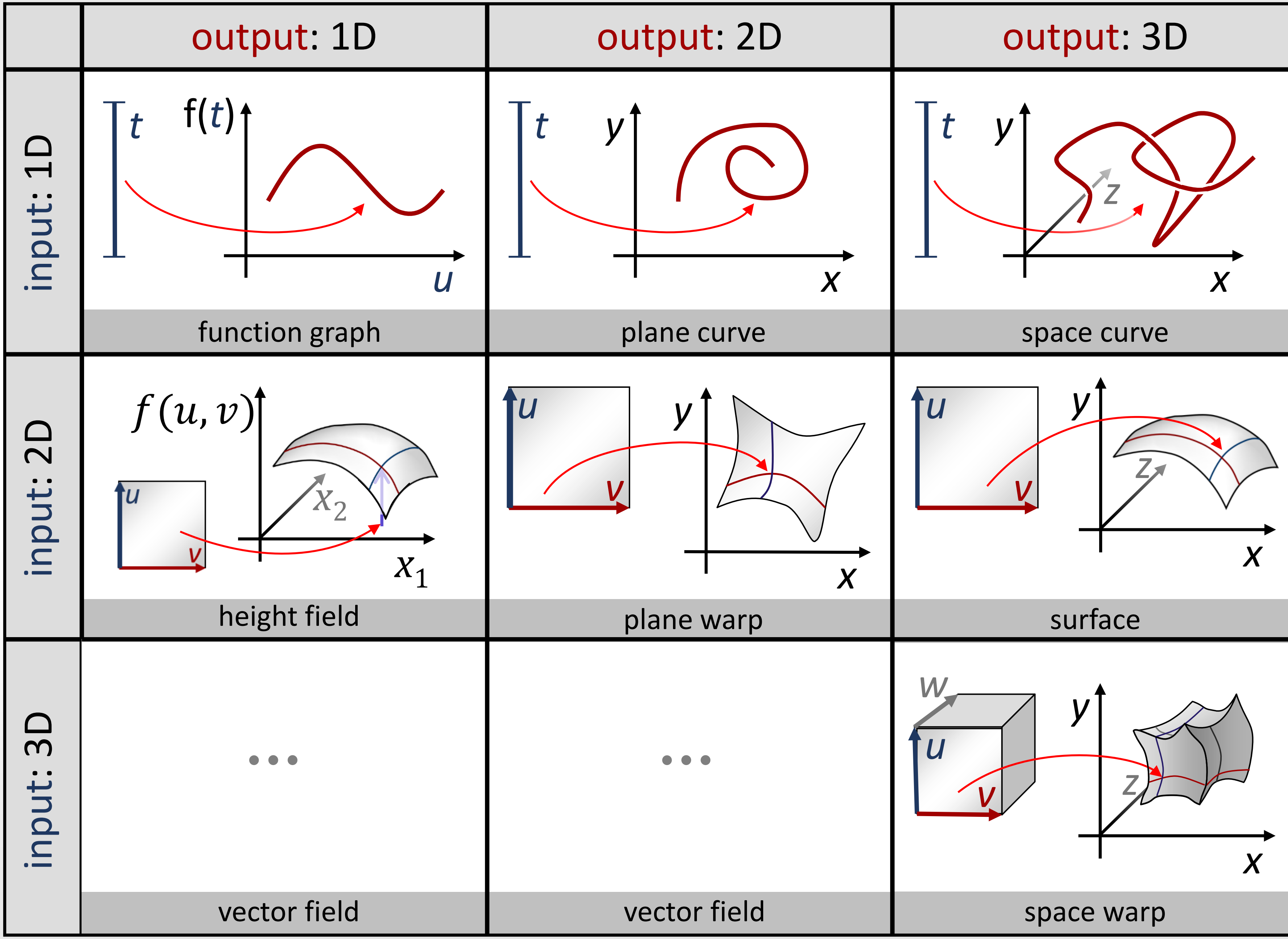
$$\frac{\partial}{\partial x_k} f(\mathbf{x}) = \partial_k f(\mathbf{x}) = f_{x_k}(\mathbf{x})$$

Parametric Models



Parametric Models

- f maps from parameter domain Ω to target space
- Evaluation of f: one point on the model



output: 1D
output: 2D
output: 3D
input: 1D
input: 2D
input: 3D
t
f(t)
u
function graph
t
y
x
plane curve
t
y
z
x
space curve
f (u, v)
u
v
x_2
x_1
height field
u
v
y
x
plane warp
u
v
y
z
x
surface
• • •
vector field
• • •
vector field
w
u
v
y
z
x
space warp

Visualization

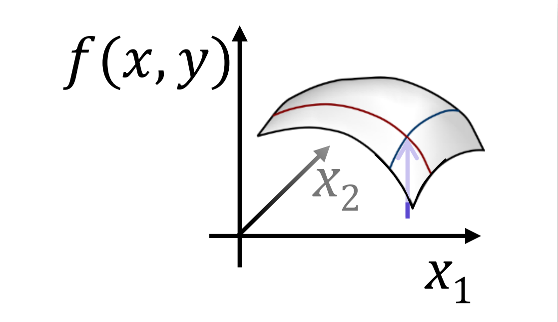
Derivatives for:

- Functions $f: \mathbb{R}^n \to \mathbb{R}$ ("scalar field")
- Functions $f: \mathbb{R} \to \mathbb{R}^n$ ("curves")
- Functions $f: \mathbb{R}^n \to \mathbb{R}^m$ (general case)

Visualization

Derivatives for:

- Functions $f: \mathbb{R}^n \to \mathbb{R}$ ("scalar field")
- Functions $f: \mathbb{R} \to \mathbb{R}^n$ ("curves")
- Functions $f: \mathbb{R}^n \to \mathbb{R}^m$ (general case)



height field

Gradient

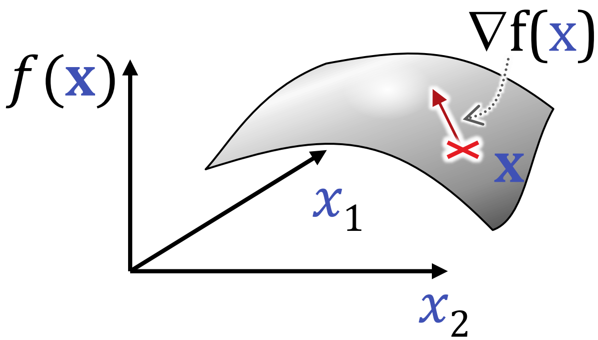
Scalar field

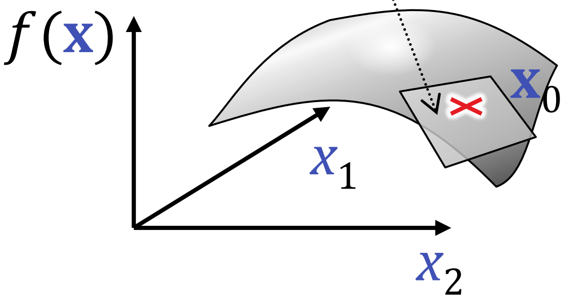
$$f\colon \mathbb{R}^n \to \mathbb{R}$$

Gradient

$$\nabla f(\mathbf{x}) = \begin{pmatrix} \dfrac{\partial}{\partial x_1} \\ \vdots \\ \dfrac{\partial}{\partial x_n} \end{pmatrix} f(\mathbf{x}) = \begin{pmatrix} \dfrac{\partial}{\partial x_1} f(\mathbf{x}) \\ \vdots \\ \dfrac{\partial}{\partial x_n} f(\mathbf{x}) \end{pmatrix}$$

Gradient





Direction of steepest ascent

- Local linear approximation:

$$f(\mathbf{x}) \approx f(\mathbf{x}_0) + \nabla f(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0)$$

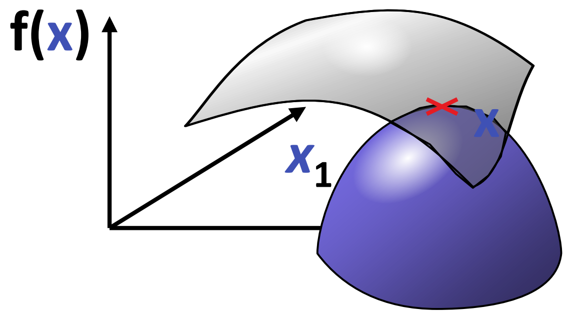
Hessian Matrix

Second derivative

$$\mathbf{H}_f(\mathbf{x}) \coloneqq \begin{pmatrix} \frac{\partial^2}{\partial x_1^2} & \frac{\partial}{\partial x_2}\frac{\partial}{\partial x_1} & \cdots & \frac{\partial}{\partial x_n}\frac{\partial}{\partial x_1} \\ \frac{\partial}{\partial x_1}\frac{\partial}{\partial x_2} & \frac{\partial^2}{\partial x_2^2} & & \frac{\partial}{\partial x_n}\frac{\partial}{\partial x_2} \\ \vdots & & \ddots & \vdots \\ \frac{\partial}{\partial x_1}\frac{\partial}{\partial x_n} & \frac{\partial}{\partial x_2}\frac{\partial}{\partial x_n} & \cdots & \frac{\partial^2}{\partial x_n^2} \end{pmatrix} f(\mathbf{x})$$

- "Hessian" matrix (symmetric for $f \in C^2$)
- Even higher order: symmetric tensors

Taylor Approximation



2nd order approximation
(schematic)

Second order Taylor approximation:

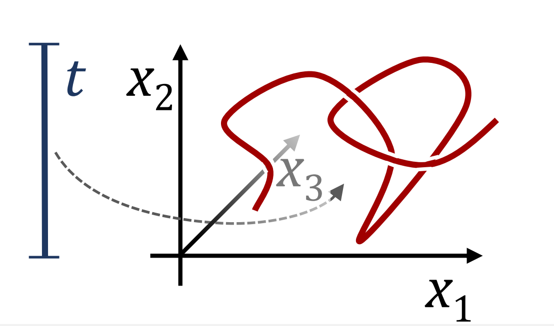
- Fit a paraboloid to a general function

$$f(\mathbf{x}) \approx f(\mathbf{x}_0) + \nabla f(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0)$$
$$+ \frac{1}{2} (\mathbf{x} - \mathbf{x}_0)^{\mathrm{T}} \cdot \mathbf{H}_f(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0)$$
$$+ \mathcal{O}(\|\Delta \mathbf{x}\|^3)$$

Special Cases

Derivatives for:

- Functions $f: \mathbb{R}^n \to \mathbb{R}$ (“scalar field”)
- Functions $f: \mathbb{R} \to \mathbb{R}^n$ (“curves”)
- Functions $f: \mathbb{R}^n \to \mathbb{R}^m$ (general case)



curves

Derivatives of Curves

Derivatives of vector valued functions:

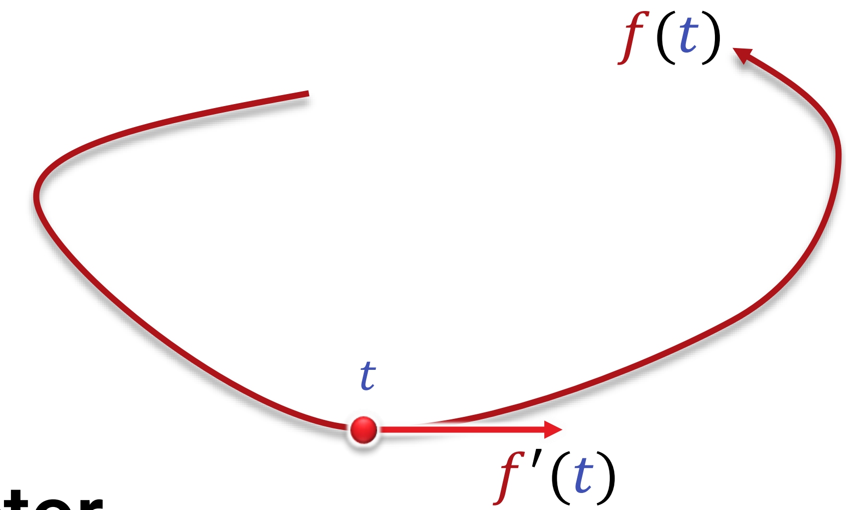
- Function $f: \mathbb{R} \to \mathbb{R}^n$ ("curve")

$$f(t) = \begin{pmatrix} f_1(t) \\ \vdots \\ f_n(t) \end{pmatrix}$$

- Derivatives for all output dimension

$$\frac{d}{dt} f(t) := \begin{pmatrix} \frac{d}{dt} f_1(t) \\ \vdots \\ \frac{d}{dt} f_n(t) \end{pmatrix} = f'(t) = \dot{f}(t)$$

Geometric Meaning



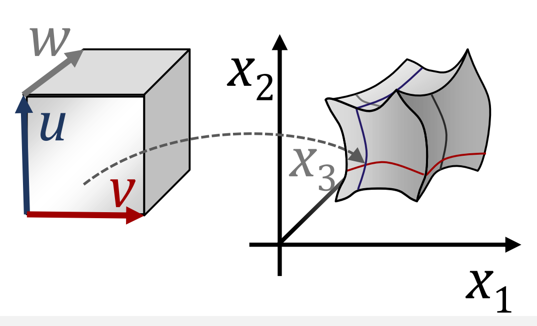
Tangent Vector

- f' is the tangent vector
  - Higher order derivatives: also vectors
- Physical particle
  - First derivative $\dot{f}$ ≙ velocity.
  - Second derivative $\ddot{f}$ ≙ acceleration

Special Cases

Derivatives for:

- Functions $f: \mathbb{R}^n \to \mathbb{R}$ ("scalar field")
- Functions $f: \mathbb{R} \to \mathbb{R}^n$ ("curves")
- Functions $f: \mathbb{R}^n \to \mathbb{R}^m$ (general case)



"space warp" $\mathbb{R}^3 \to \mathbb{R}^3$

Jacobian Matrix

Jacobian Matrix:

$$\nabla f(\mathbf{x}) = \mathrm{J}_f(\mathbf{x}) = \nabla f(x_1, \dots, x_n)$$

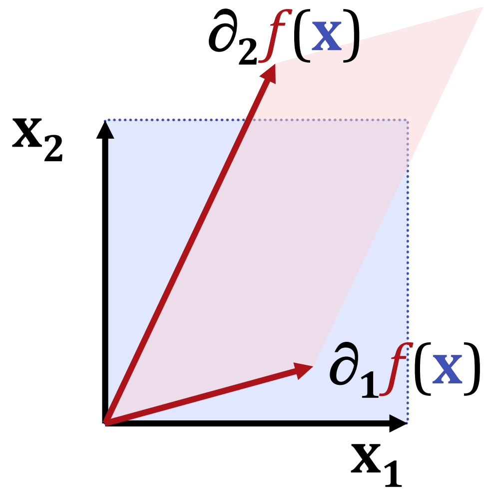
$$= \begin{pmatrix} \nabla f_1(x_1, \dots, x_n) \\ \vdots \\ \nabla f_m(x_1, \dots, x_n) \end{pmatrix} = \begin{pmatrix} \partial_{x_1} f_1(\mathbf{x}) & \cdots & \partial_{x_n} f_1(\mathbf{x}) \\ \vdots & \ddots & \vdots \\ \partial_{x_1} f_m(\mathbf{x}) & \cdots & \partial_{x_n} f_m(\mathbf{x}) \end{pmatrix}$$

First-order Taylor approximation:

$$f(\mathbf{x}) \approx f(\mathbf{x_0}) + \mathbf{J}_f(\mathbf{x_0}) \cdot (\mathbf{x} - \mathbf{x_0})$$

matrix / vector
product

Intuition



Jacobian Matrix / ∇f:

- Think of basis vectors of input space
- Mapped to parallelepiped in output space

Tensor Formulation

General case: tensors

- Input $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$
- Developed at $\mathbf{y} = (y_1, \dots, y_n) \in \mathbb{R}^n$
- Output $f(\mathbf{x}) = \left(f^1(\mathbf{x}), \dots, f^m(\mathbf{x})\right) \in \mathbb{R}^m$

using y instead of x_0 to avoid index confusion

$$f^j(\mathbf{x}) \approx \sum_{k=0}^{ord} \frac{1}{k!} \sum_{\substack{(i_1,\dots,i_k) \\ \in\{1,\dots n\}^k}} \underbrace{\left[\frac{\partial}{\partial_{i_1} \cdots \partial_{i_k}} f^j\right](\mathbf{x}_0)}_{\text{Tensors } (d_k)^j_{i_1 \cdots i_k}} \left(x_{i_1} - y_{i_1}\right) \cdots \left(x_{i_k} - y_{i_k}\right)$$

Are partial derivatives canonical?

Partial Derivatives – Coordinate Systems

use curly-d

$$\frac{\partial}{\partial x_k} f(x_1, \ldots, x_k, \ldots, x_n) := \lim_{h \to 0} \frac{f(x_1, \ldots, x_k + h, \ldots, x_n) - f(x_1, \ldots, x_k, \ldots, x_n)}{h}$$

Problem:

- What happens, if the coordinate system changes?
- Partial derivatives go into different directions then.
- Do we get the same result?

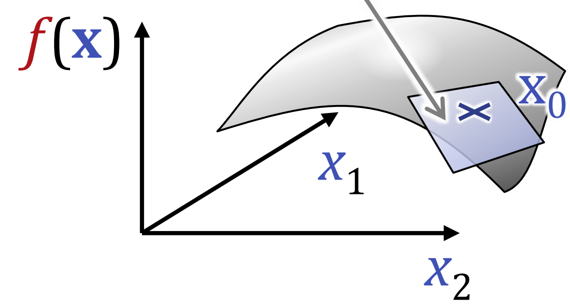
Coordinate Systems

Problem:

- What happens, if the coordinate system changes?
- Partial derivatives go into different directions then.
- Do we get the same result?

Total Derivative

$$f(\mathbf{x_0}) + \mathbf{J}_f(\mathbf{x_0}) \cdot (\mathbf{x} - \mathbf{x_0})$$



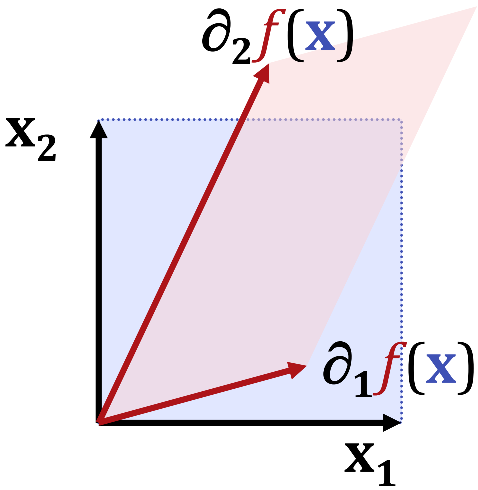
First order Taylor approx.:

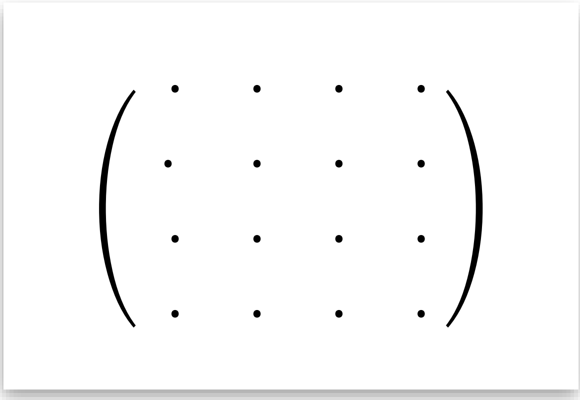
- $f(\mathbf{x}) = f(\mathbf{x_0}) + \mathbf{J}_f(\mathbf{x_0}) \cdot (\mathbf{x} - \mathbf{x_0}) + \mathbf{r}_{\mathbf{x}_0}(\mathbf{x})$
- Converges for C^1 functions $f: \mathbb{R}^n \to \mathbb{R}^m$:

$$\lim_{\mathbf{x} \to \mathbf{x}_0} \frac{\mathbf{r}_{\mathbf{x}_0}(\mathbf{x})}{\|\mathbf{x} - \mathbf{x}_0\|} = 0$$

("*totally differentiable*")

Intuition





Coordinates do not matter for C^k $(k \geq 1)$

- Differentiation: approximate with linear map
- Linear map is fixed by mapping the basis vectors

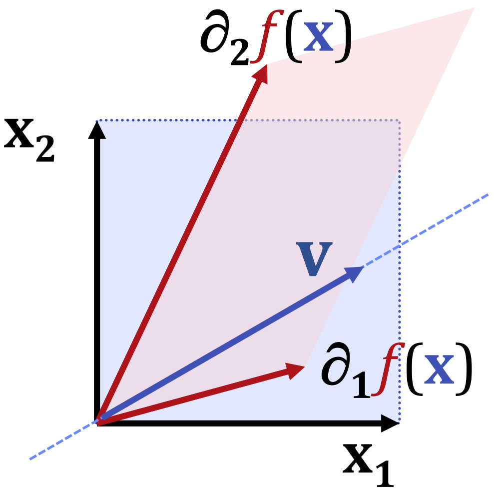
Directional Derivative

The directional derivative is defined as:

- Given $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$ and $\mathbf{v} \in \mathbb{R}^n, ||\mathbf{v}|| = 1$
- Directional derivative

$$\nabla_{\mathbf{v}} f(\mathbf{x}) = \frac{\partial f}{\partial \mathbf{v}}(\mathbf{x}) := \frac{d}{dt} f(\mathbf{x} + t\mathbf{v})$$

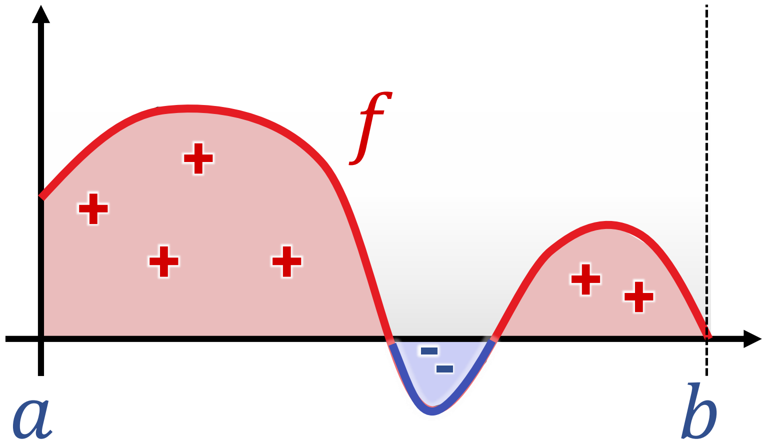
$$\nabla_{\mathbf{v}} f(\mathbf{x}) = \nabla f(\mathbf{x}) \frac{\mathbf{v}}{||\mathbf{v}||}$$



- Compute from Jacobian matrix
  (total differentiability required)

Integration

Integral



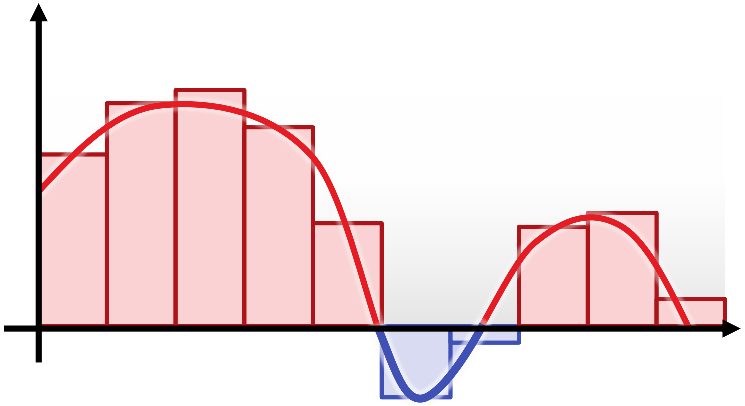
Integral of a function

- Function $f: \mathbb{R} \to \mathbb{R}$
- Integral

$$\int_a^b f(x)dx$$

measures signed area under curve

Integral



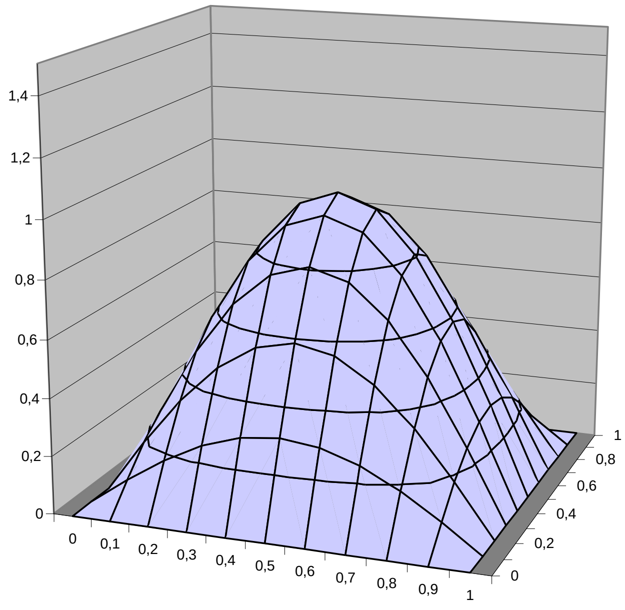
Numerical Approximation

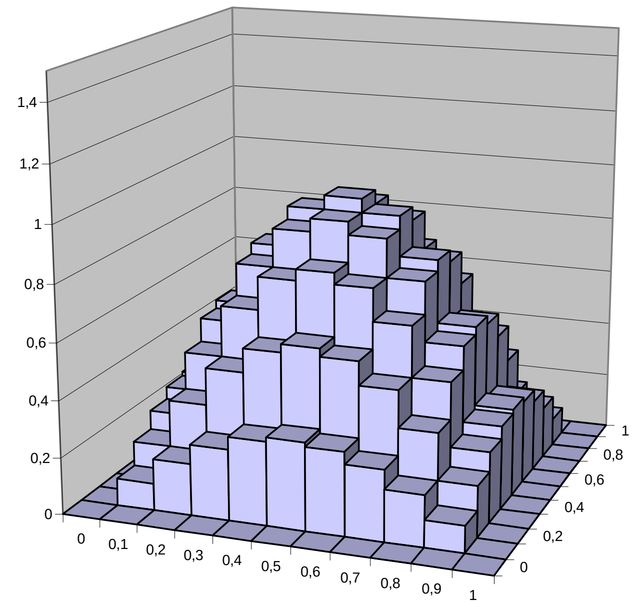
- Sum up a series of approximate shapes
- (Riemannian) Definition: limit for baseline → zero
- Intuition: Sum of numbers in array

Multi-Dimensional Integral

Integration in higher dimensions

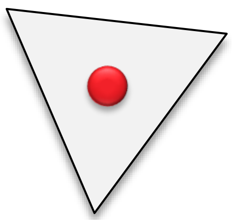
- Functions $f: \mathbb{R}^n \to \mathbb{R}$
- Tessellate domain and sum up cuboids



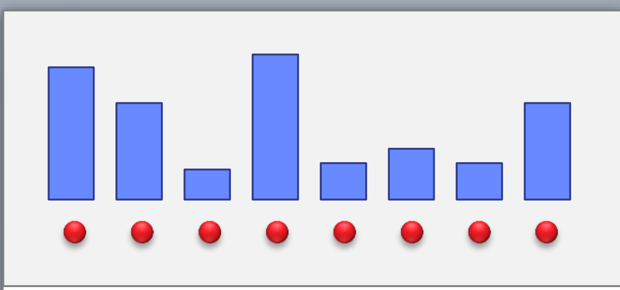


Surface Integrals

Line / Surface / Volume / Hypervolume Elements

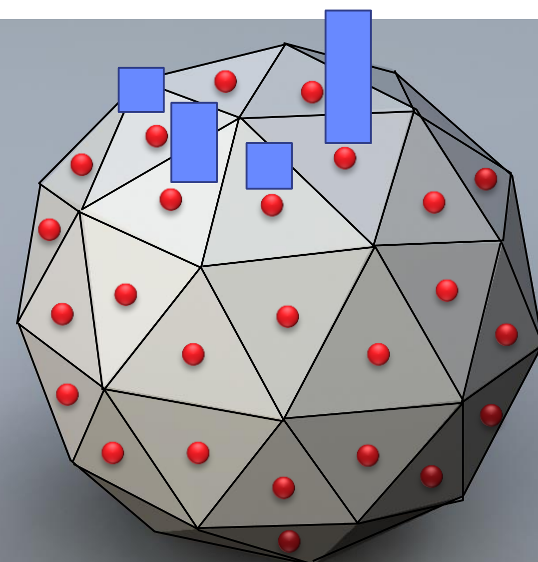


$$\int_{\mathcal{S}} f(\mathbf{x})\,d\mathbf{x} = \lim_{\text{smaller}\,\triangledown} \sum_{i=1}^{n} f(\mathbf{x}_i)\cdot|\triangledown_i|$$



function f on surface $\mathcal{S}$

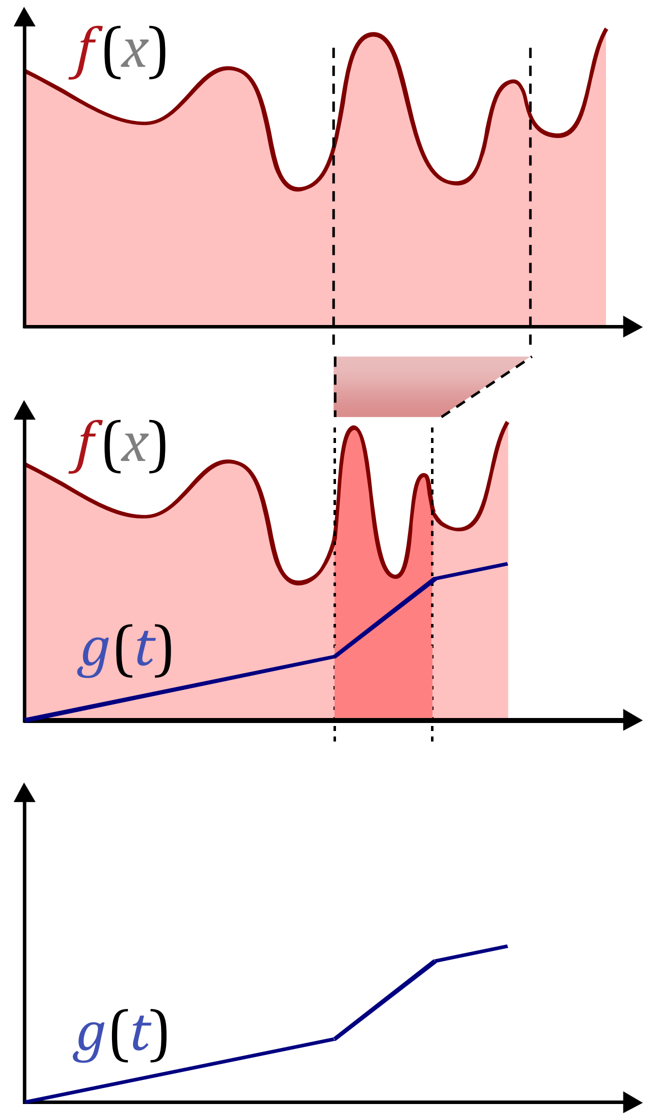
$f : \mathcal{S} \to \mathbb{R}$



Integral Transformations

Integration by substitution:

$$\int_a^b f(x)dx = \int_{g^{-1}(a)}^{g^{-1}(b)} f\big(g(t)\big) \cdot g'(t)dt$$



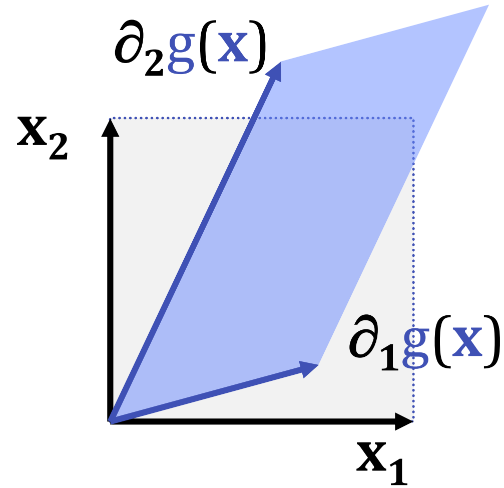
Need to compensate

- Speed of movement affects measured area
  - **Faster:** shrinks measured area
  - **Slower:** inflates

Multi-Dimensional Substitution

Transformation of Integrals:

$$\int_{\Omega} f(\mathbf{x})d\mathbf{x} = \int_{g^{-1}(\Omega)} f\big(g(\mathbf{y})\big) \cdot |\det[\nabla g(\mathbf{y})]|d\mathbf{y}$$



- $g \in C^1$, invertible
- Jacobian approximates local behavior of $g(\cdot)$
- Determinant: local area/volume change
- In particular: $\big|\det\big(\nabla g(\mathbf{y})\big)\big| = 1$ means $g(\cdot)$ is *area/volume conserving*.