

MATH 311

Topics in Applied Mathematics I

Lecture 31:

Differentiation in vector spaces.

The derivative

Definition. A real function f is said to be **differentiable** at a point $a \in \mathbb{R}$ if it is defined on an open interval containing a and the limit

$$\lim_{h \rightarrow 0} \frac{f(a+h) - f(a)}{h}$$

exists. The limit is denoted $f'(a)$ and called the **derivative** of f at a . An equivalent condition is

$$f(a+h) = f(a) + f'(a)h + r(h), \quad \text{where} \quad \lim_{h \rightarrow 0} r(h)/h = 0.$$

If a function f is differentiable at a point a , then it is continuous at a .

Suppose that a function f is defined and differentiable on an interval I . Then the derivative of f can be regarded as a function on I .

Problem. Find $\min_{x>0} x^x$.

The function $f(x) = x^x$ is well defined and positive on $(0, \infty)$. Hence

$$f(x) = e^{\log f(x)} = e^{\log x^x} = e^{x \log x}$$

for all $x > 0$. That is, $f(x) = g(h(x))$, where $h(x) = x \log x$ and $g(y) = e^y$. Using the Chain Rule and the Product Rule, we obtain

$$\begin{aligned} f'(x) &= e^{x \log x} (x \log x)' = x^x \left((x)' \log x + x(\log x)' \right) \\ &= x^x (\log x + 1). \end{aligned}$$

It follows that $f'(x) < 0$ for $0 < x < 1/e$ and $f'(x) > 0$ for $x > 1/e$. Hence the function f is strictly decreasing on $(0, 1/e]$ and strictly increasing on $[1/e, \infty)$. Therefore

$$\min_{x>0} f(x) = f(1/e) = (1/e)^{1/e} = e^{-1/e}.$$

Convergence in normed vector spaces

Suppose V is a vector space endowed with a norm $\|\cdot\|$. The norm gives rise to a distance function $\text{dist}(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|$.

Definition. We say that a sequence of vectors $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \dots$ converges to a vector \mathbf{u} in the normed vector space V if $\|\mathbf{v}_k - \mathbf{u}\| \rightarrow 0$ as $k \rightarrow \infty$.

In the case $V = \mathbb{R}^n$, a sequence of vectors converges with respect to a norm if and only if it converges in each coordinate. In the case $V = \mathcal{M}_{m,n}(\mathbb{R})$, a sequence of matrices converges with respect to a norm if and only if it converges in each entry.

Similarly, in the case $\dim V < \infty$ we can choose a finite basis $\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_n$. Any vector $\mathbf{v} \in V$ can be expanded into a linear combination $\mathbf{v} = x_1\mathbf{w}_1 + x_2\mathbf{w}_2 + \dots + x_n\mathbf{w}_n$. Then a sequence of vectors converges with respect to a norm if and only if each of the coordinates x_i converges.

Vector-valued functions

Suppose V is a vector space endowed with a norm $\| \cdot \|$.

Definition. We say that a function $\mathbf{v} : X \rightarrow V$ defined on a set $X \subset \mathbb{R}$ converges to a limit $\mathbf{u} \in V$ at a point $a \in \mathbb{R}$ if $\|\mathbf{v}(x) - \mathbf{u}\| \rightarrow 0$ as $x \rightarrow a$.

Further, we say that the function \mathbf{v} is continuous at a point $c \in X$ if $\mathbf{v}(c) = \lim_{x \rightarrow c} \mathbf{v}(x)$.

Finally, the function \mathbf{v} is said to be differentiable at a point $a \in \mathbb{R}$ if it is defined on an open interval containing a and the limit

$$\lim_{h \rightarrow 0} \frac{1}{h} (f(a+h) - f(a))$$

exists. The limit is denoted $\mathbf{v}'(a)$ and called the derivative of \mathbf{v} at a .

Differentiability theorems

Sum Rule If functions $\mathbf{v} : X \rightarrow V$ and $\mathbf{w} : X \rightarrow V$ are differentiable at a point $a \in \mathbb{R}$, then the sum $\mathbf{v} + \mathbf{w}$ is also differentiable at a . Moreover, $(\mathbf{v} + \mathbf{w})'(a) = \mathbf{v}'(a) + \mathbf{w}'(a)$.

Homogeneous Rule If a function $\mathbf{v} : X \rightarrow V$ is differentiable at a point $a \in \mathbb{R}$, then for any $r \in \mathbb{R}$ the scalar multiple $r\mathbf{v}$ is also differentiable at a . Moreover, $(r\mathbf{v})'(a) = r\mathbf{v}'(a)$.

Difference Rule If functions $\mathbf{v} : X \rightarrow V$ and $\mathbf{w} : X \rightarrow V$ are differentiable at a point $a \in \mathbb{R}$, then the difference $\mathbf{v} - \mathbf{w}$ is also differentiable at a . Moreover, $(\mathbf{v} - \mathbf{w})'(a) = \mathbf{v}'(a) - \mathbf{w}'(a)$.

Differentiability theorems

Product Rule #1 If functions $f : X \rightarrow \mathbb{R}$ and $\mathbf{v} : X \rightarrow V$ are differentiable at a point $a \in \mathbb{R}$, then the scalar multiple $f\mathbf{v}$ is also differentiable at a . Moreover,
$$(f\mathbf{v})'(a) = f'(a)\mathbf{v}(a) + f(a)\mathbf{v}'(a).$$

Product Rule #2 Assume that the norm on V is induced by an inner product $\langle \cdot, \cdot \rangle$. If functions $\mathbf{v} : X \rightarrow V$ and $\mathbf{w} : X \rightarrow V$ are differentiable at a point $a \in \mathbb{R}$, then the inner product $\langle \mathbf{v}, \mathbf{w} \rangle$ is also differentiable at a . Moreover,
$$(\langle \mathbf{v}, \mathbf{w} \rangle)'(a) = \langle \mathbf{v}'(a), \mathbf{w}(a) \rangle + \langle \mathbf{v}(a), \mathbf{w}'(a) \rangle.$$

Chain Rule If a function $f : X \rightarrow \mathbb{R}$ is differentiable at a point $a \in \mathbb{R}$ and a function $\mathbf{v} : Y \rightarrow V$ is differentiable at $f(a)$, then the composition $\mathbf{v} \circ f$ is differentiable at a . Moreover,
$$(\mathbf{v} \circ f)'(a) = f'(a)\mathbf{v}'(f(a)).$$

Partial derivative

Consider a function $f : X \rightarrow V$ that is defined in a domain $X \subset \mathbb{R}^n$ and takes values in a normed vector space V . The function f depends on n real variables: $f = f(x_1, x_2, \dots, x_n)$.

Let us select a point $\mathbf{a} = (a_1, a_2, \dots, a_n) \in X$ and a variable x_j . Now we go to the point \mathbf{a} and fix all variables except x_j . That is, we introduce a function of one variable

$$\phi(x) = f(a_1, \dots, a_{j-1}, x, a_{j+1}, \dots, a_n).$$

If the function ϕ is differentiable at a_j , then the derivative $\phi'(a_j)$ is called the **partial derivative** of f at the point \mathbf{a} with respect to the variable x_j .

Notation: $\frac{\partial f}{\partial x_j}(\mathbf{a})$, $\frac{\partial}{\partial x_j} f(\mathbf{a})$, $(D_{x_j} f)(\mathbf{a})$.

Directional derivative

Consider a function $f : X \rightarrow V$ that is defined on a subset $X \subset W$ of a vector space W and takes values in a normed vector space V . For every point $\mathbf{a} \in X$ and vector $\mathbf{v} \in W$ we introduce a function of real variable $\phi(t) = f(\mathbf{a} + t\mathbf{v})$. If the function ϕ is differentiable at 0, then the derivative $\phi'(0)$ is called the **directional derivative** of f at the point \mathbf{a} along the vector \mathbf{v} . Notation: $(D_{\mathbf{v}}f)(\mathbf{a})$.

The partial derivative is a particular case of the directional derivative, when $W = \mathbb{R}^n$ and \mathbf{v} is from the standard basis.

Homogeneity $(D_{r\mathbf{v}}f)(\mathbf{a}) = r(D_{\mathbf{v}}f)(\mathbf{a})$ for all $r \in \mathbb{R}$ whenever $(D_{\mathbf{v}}f)(\mathbf{a})$ exists.

Linearity Suppose W is a normed vector space, $(D_{\mathbf{v}}f)(\mathbf{a})$ exists for all \mathbf{v} and depends continuously(?) on \mathbf{a} . Then $\mathbf{v} \mapsto (D_{\mathbf{v}}f)(\mathbf{a})$ is a linear transformation.

Limit of a function and continuity

Let V and W be normed vector spaces. Suppose $f : E \rightarrow V$ is a function defined on a set $E \subset W$.

Definition. We say that the function f **converges to a limit** $L \in V$ at a point $\mathbf{w}_0 \in W$ if for every $\varepsilon > 0$ there exists $\delta = \delta(\varepsilon) > 0$ such that for all $\mathbf{w} \in E$,

$$0 < \|\mathbf{w} - \mathbf{w}_0\| < \delta \quad \text{implies} \quad \|f(\mathbf{w}) - L\| < \varepsilon.$$

An equivalent condition is that for any sequence $\mathbf{w}_1, \mathbf{w}_2, \dots$ of vectors from E , $\lim_{n \rightarrow \infty} \mathbf{w}_n = \mathbf{w}_0$ implies $\lim_{n \rightarrow \infty} f(\mathbf{w}_n) = L$.

Definition. Given a set $E \subset W$, a function $f : E \rightarrow V$, and a point $\mathbf{w}_0 \in E$, the function f is **continuous at \mathbf{w}_0** if

$$f(\mathbf{w}_0) = \lim_{\mathbf{w} \rightarrow \mathbf{w}_0} f(\mathbf{w}).$$

We say that the function f is **continuous on** a set $E_0 \subset E$ if f is continuous at every point of E_0 .

Continuity of a linear transformation

Theorem Suppose V and W are normed vector spaces and $L : W \rightarrow V$ is a linear transformation. Then the following conditions are equivalent:

- (i) L is continuous everywhere on W ,
- (ii) L is continuous at the zero vector,
- (iii) $\|L(\mathbf{w})\| \leq C\|\mathbf{w}\|$ for some $C > 0$ and all $\mathbf{w} \in W$.

Example. • If $\dim W < \infty$ then any linear transformation $L : W \rightarrow V$ is continuous. Otherwise it is not so.

The (Fréchet) differential

Suppose V and W are normed vector spaces and consider a function $F : X \rightarrow V$, where $X \subset W$.

Definition. We say that the function F is **differentiable** at a point $\mathbf{a} \in X$ if it is defined in a neighborhood of \mathbf{a} and there exists a continuous linear transformation $L : W \rightarrow V$ such that

$$F(\mathbf{a} + \mathbf{v}) = F(\mathbf{a}) + L(\mathbf{v}) + R(\mathbf{v}),$$

where $\|R(\mathbf{v})\|/\|\mathbf{v}\| \rightarrow 0$ as $\|\mathbf{v}\| \rightarrow 0$. The transformation L is called the **differential** of F at \mathbf{a} and denoted $(DF)(\mathbf{a})$.

Theorem If a function F is differentiable at a point \mathbf{a} , then the directional derivatives $(D_{\mathbf{v}}F)(\mathbf{a})$ exist for all \mathbf{v} and $(D_{\mathbf{v}}F)(\mathbf{a}) = (DF)(\mathbf{a})[\mathbf{v}]$.

Chain Rule If a function F is differentiable at a point \mathbf{a} and a function G is differentiable at the point $\mathbf{b} = F(\mathbf{a})$, then $(D(G \circ F))(\mathbf{a}) = DG(\mathbf{b}) \circ DF(\mathbf{a})$.

Examples

- Any linear transformation $L : \mathbb{R} \rightarrow \mathbb{R}$ is a scaling $L(x) = rx$ by a scalar r . If L is the differential of a function $f : X \rightarrow \mathbb{R}$ at a point $a \in \mathbb{R}$, then $r = f'(a)$.
- Any linear transformation $L : \mathbb{R}^n \rightarrow \mathbb{R}$ is the dot product with a fixed vector, $L(\mathbf{x}) = \mathbf{x} \cdot \mathbf{v}_0$. If L is the differential of a function $f : X \rightarrow \mathbb{R}$ at a point $\mathbf{a} \in \mathbb{R}^n$, then $\mathbf{v}_0 = \nabla f(\mathbf{a})$.
- Any linear transformation $L : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is a matrix transformation: $L(\mathbf{x}) = B\mathbf{x}$, where $B = (b_{ij})$ is an $m \times n$ matrix. If L is the differential of a function $\mathbf{F} : X \rightarrow \mathbb{R}^m$ at a point $\mathbf{a} \in \mathbb{R}^n$, then $b_{ij} = \frac{\partial F_i}{\partial x_j}(\mathbf{a})$.

The matrix B of partial derivatives is called the **Jacobian**

matrix of \mathbf{F} and denoted $\frac{\partial(F_1, \dots, F_m)}{\partial(x_1, \dots, x_n)}$.

Problem. Find the directional derivative of the function $f(x, y) = e^{x+y} \sin(x - y)$ at the point $(0, 0)$ along the vector $\mathbf{v} = (2, 1)$.

By definition, the directional derivative $D_{\mathbf{v}}f(0, 0)$ equals the derivative of the function $\phi(t) = f(t\mathbf{v})$ at the point $t = 0$. We have $\phi(t) = f(2t, t) = e^{3t} \sin t$. Then $\phi'(t) = 3e^{3t} \sin t + e^{3t} \cos t = e^{3t}(3 \sin t + \cos t)$. Finally, $\phi'(0) = 1$.

Alternatively, we can find the directional derivative by the formula $D_{\mathbf{v}}f(0, 0) = \nabla f(0, 0) \cdot \mathbf{v}$. We obtain

$$\frac{\partial f}{\partial x} = e^{x+y} \sin(x - y) + e^{x+y} \cos(x - y),$$

$$\frac{\partial f}{\partial y} = e^{x+y} \sin(x - y) - e^{x+y} \cos(x - y).$$

Then $\nabla f(0, 0) = (1, -1)$. Consequently, $D_{\mathbf{v}}f(0, 0) = (1, -1) \cdot (2, 1) = 1$.