IMPA - Instituto de Matemática Pura e Aplicada, Rio de Janeiro, RJ, Brazil

Introduction to Computational Manifolds and Applications

Part 1 - Foundations

Prof. Jean Gallier

jean@cis.upenn.edu

Department of Computer and Information Science
University of Pennsylvania
Philadelphia, PA, USA

A Review of Multivariate Calculus

Normed Spaces

In order to define how close two vectors or two matrices are, and in order to define the convergence of sequences of vectors or matrices, we will use use the notion of a *norm*.

Recall that $\mathbb{R}_+ = \{x \in \mathbb{R} \mid x \geq 0\}$.

Also, recall that if $z = x + iy \in \mathbb{C}$ is a complex number, with $x, y \in \mathbb{R}$, then

$$\bar{z} = x - iy \qquad \text{and} \qquad |z| = \sqrt{x^2 + y^2}\,,$$

where $\bar{z}$ and $|z|$ are the *modulus* and the *conjugate* of the complex number z, respectively.

Normed Spaces

Definition 2.1. Let E be a vector space over a field K, where K is either the field of $\mathbb{R}$ or the field of $\mathbb{C}$. A *norm* on E is a function, $\| \ \| : E \to \mathbb{R}_+$, assigning a nonnegative real number $\|\boldsymbol{u}\|$ to any vector $\boldsymbol{u} \in E$, and satisfying the following conditions for all $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in E$:

(N1) $\|\boldsymbol{x}\| \geq 0$, and $\|\boldsymbol{x}\| = 0$ iff $\boldsymbol{x} = \boldsymbol{0}$, (positivity)

(N2) $\|\lambda \cdot \boldsymbol{x}\| = |\lambda| \cdot \|\boldsymbol{x}\|$, (scaling)

(N3) $\|\boldsymbol{x} + \boldsymbol{y}\| \leq \|\boldsymbol{x}\| + \|\boldsymbol{y}\|$. (triangle inequality)

A vector space E together with a norm $\| \ \|$ is called a *normed vector space*.

From (N3), we easily get $|\|\boldsymbol{x}\| - \|\boldsymbol{y}\|| \leq \|\boldsymbol{x} - \boldsymbol{y}\|$.

Normed Spaces

Examples:

1. Let $E = \mathbb{R}$, and $\|\boldsymbol{x}\| = |\boldsymbol{x}|$, the absolute value of $\boldsymbol{x}$.
2. Let $E = \mathbb{C}$, and $\|\boldsymbol{z}\| = |\boldsymbol{z}|$, the modulus of $\boldsymbol{z}$.
3. Let $E = \mathbb{R}^n$ (or $E = \mathbb{C}^n$). There are three standard norms.

A Review of Multivariate Calculus

Normed Spaces

For every vector $\boldsymbol{x} = (x_1, \ldots, x_n) \in E$, we have the *1-norm*, $\|\boldsymbol{x}\|_1$, defined such that

$$\|\boldsymbol{x}\|_1 = |x_1| + \cdots + |x_n|\,,$$

we have the *Euclidean norm* $\|\boldsymbol{x}\|_2$, defined such that,

$$\|\boldsymbol{x}\|_2 = \left(|x_1|^2 + \cdots + |x_n|^2\right)^{\frac{1}{2}}\,,$$

and the *sup-norm*, $\|\boldsymbol{x}\|_\infty$, defined such that,

$$\|\boldsymbol{x}\|_\infty = \max\{|x_i| \mid 1 \leq i \leq n\}\,.$$

More generally, we define the *l_p-norm* (for $p \geq 1$) by

$$\|\boldsymbol{x}\|_p = \left(|x_1|^p + \cdots + |x_n|^p\right)^{\frac{1}{p}}\,.$$

Normed Spaces

For simplicity, we will consider the vector spaces $\mathrm{M}_n(\mathbb{R})$ and $\mathrm{M}_n(\mathbb{C})$ of square $n \times n$ matrices.

Definition 2.2. A *matrix norm* $\| \ \|$ on the space of square $n \times n$ matrices in $\mathrm{M}_n(K)$, with $K = \mathbb{R}$ or $K = \mathbb{C}$, is a norm on the vector space $\mathrm{M}_n(K)$ with the additional property that

$$\|A \cdot B\| \leq \|A\| \cdot \|B\| ,$$

for all $A, B \in \mathrm{M}_n(K)$.

Since $I^2 = I$, from $\|I\| = \|I^2\| \leq \|I\|^2$, we get $\|I\| \geq 1$, for every matrix norm.

Normed Spaces

We begin with the so-called *Frobenius norm*, which is just the norm $\| \ \|_2$ on $\mathbb{C}^{n^2}$, where the $n \times n$ matrix A is viewed as the vector obtained by concatenating together the rows (or the columns) of A.

Definition 2.3. The *Frobenius norm* $\| \ \|_F$ is defined so that for every square $n \times n$ matrix $A \in \mathrm{M}_n(\mathbb{C})$,

$$\|A\|_F = \left(\sum_{i,j=1}^{n} |a_{ij}|^2 \right)^{1/2} = \sqrt{\mathrm{tr}(AA^*)} = \sqrt{\mathrm{tr}(A^*A)}.$$

A Review of Multivariate Calculus

Normed Spaces

Proposition 2.1. The Frobenius norm $\| \ \|_F$ on $\mathrm{M}_n(\mathbb{C})$ satisfies the following properties:

(1) It is a matrix norm, i.e., $\|A \cdot B\|_F \leq \|A\|_F \cdot \|B\|_F$, for all $A, B \in \mathrm{M}_n(\mathbb{C})$.

(2) It is unitarily invariant, which means that for all unitary matrices U, V, we have

$$\|A\|_F = \|U \cdot A\|_F = \|A \cdot V\|_F = \|U \cdot A \cdot V\|_F.$$

(3)

$$\sqrt{\rho(A^*A)} \leq \|A\|_F \leq \sqrt{n} \cdot \sqrt{\rho(A^*A)}, \text{ for all } A \in \mathrm{M}_n(\mathbb{C}),$$

where $\rho(A^*A)$ is the largest eigenvalue of A^*A.

A Review of Multivariate Calculus

The Derivative of a Function between Normed Spaces

In most cases, we consider $E = \mathbb{R}^n$ and $F = \mathbb{R}^m$ under the Euclidean norm.

Let E and F be two normed vector spaces, let $A \subseteq E$ be some open subset of A, and let $\boldsymbol{a} \in A$ be some element of A. Even though $\boldsymbol{a}$ is a vector, we may also call it a point.

The idea behind the derivative of the function f at (a point) $\boldsymbol{a}$ is that it is a *linear approximation* of f in a small open set around $\boldsymbol{a}$. The difficulty is to make sense of the quotient

$$\frac{f(\boldsymbol{a}+\boldsymbol{h}) - f(\boldsymbol{a})}{\boldsymbol{h}}$$

where $\boldsymbol{h}$ is a vector in E. We circumvent this difficulty in two stages.

A Review of Multivariate Calculus

The Derivative of a Function between Normed Spaces

A first possibility is to consider the *directional derivative* with respect to a vector $\boldsymbol{u} \in E$, with $\boldsymbol{u} \neq \mathbf{0}$. We can consider the vector $f(\boldsymbol{a} + t \cdot \boldsymbol{u}) - f(\boldsymbol{a})$, where $t \in \mathbb{R}$. Now, the quotient

$$\frac{f(\boldsymbol{a} + t \cdot \boldsymbol{u}) - f(\boldsymbol{a})}{t}$$

makes sense.

The idea is that in E, the points

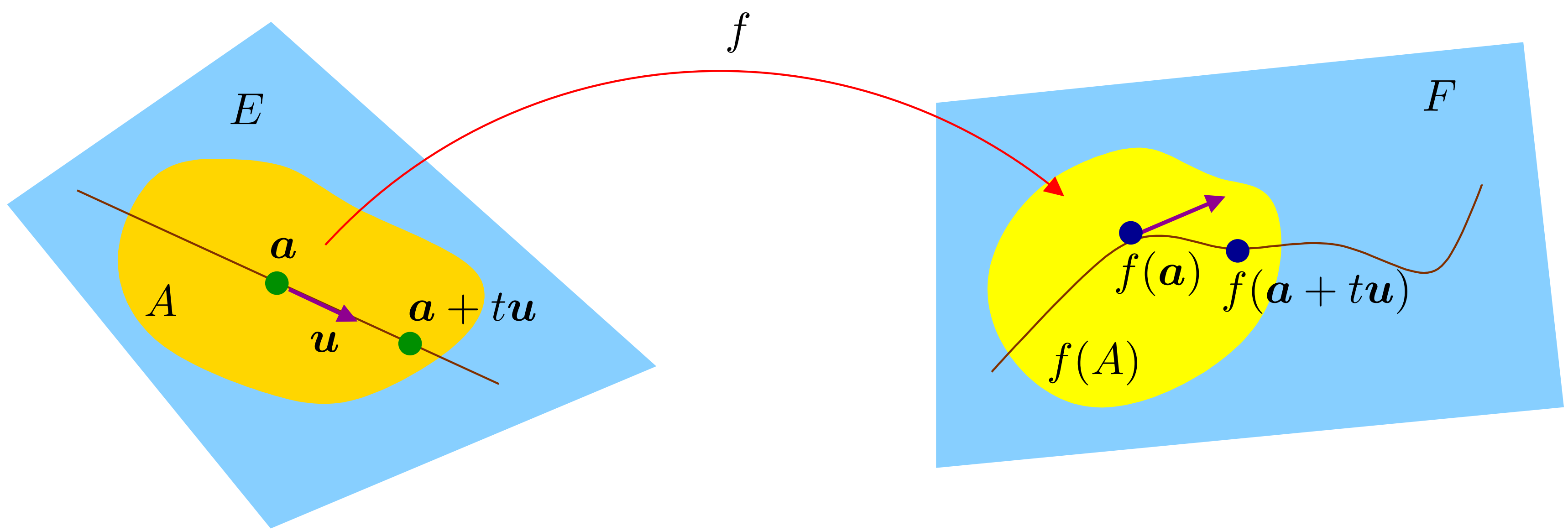
$$\boldsymbol{a} + t \cdot \boldsymbol{u},$$

for t in some small closed interval $[-\epsilon, \epsilon] \subseteq \mathbb{R}$, form a line segment containing $\boldsymbol{a}$ in the open set A and that the image of this line segment defines a small curve segment on $f(A)$.

A Review of Multivariate Calculus

The Derivative of a Function between Normed Spaces

This curve (segment) is defined by the map $t \mapsto f(\boldsymbol{a} + t \cdot \boldsymbol{u})$ from $[-\epsilon, \epsilon]$ to F, and the directional derivative $D_{\boldsymbol{u}} f(\boldsymbol{a})$ defines the direction of the tangent line at $\boldsymbol{a}$ to this curve.



A Review of Multivariate Calculus

The Derivative of a Function between Normed Spaces

Definition 2.4. Let E and F be two normed spaces, let A be a nonempty open subset of E, and let $f : A \to F$ be any function. For any $\boldsymbol{a} \in A$, for any $\boldsymbol{u}$ in E, with $\boldsymbol{u} \neq \mathbf{0}$, the *directional derivative of f at $\boldsymbol{a}$ with respect to the vector $\boldsymbol{u}$*, denoted by $D_{\boldsymbol{u}} f(\boldsymbol{a})$, is the limit

$$D_{\boldsymbol{u}} f(\boldsymbol{a}) = \lim_{t \to 0,\ t \in U} \frac{f(\boldsymbol{a} + t \cdot \boldsymbol{u}) - f(\boldsymbol{a})}{t},$$

provided that it exists, where

$$U = \{ t \in \mathbb{R} \mid \boldsymbol{a} + t \cdot \boldsymbol{u} \in A,\ t \neq 0 \}.$$

A Review of Multivariate Calculus

The Derivative of a Function between Normed Spaces

The directional derivative is sometimes called the *Gâteaux derivative*.

In the special case where $E = \mathbb{R}$ and $F = \mathbb{R}$, and we let $\boldsymbol{u} = \mathbf{1}$ (i.e., the real number 1, viewed as a vector), it is immediately verified that $D_{\mathbf{1}}f(\boldsymbol{a}) = f'(\boldsymbol{a})$, for every $\boldsymbol{a} \in A$.

When $E = \mathbb{R}$ and F is any normed vector space, the derivative $D_{\mathbf{1}}f(\boldsymbol{a})$, also denoted by $f'(\boldsymbol{a})$, provides a suitable generalization of the notion of derivative.

A Review of Multivariate Calculus

The Derivative of a Function between Normed Spaces

However, when E has dimension greater than or equal to 2, directional derivatives present a serious problem, which is that their definition is not sufficiently uniform.

As a consequence, a function can have all directional derivatives at $\boldsymbol{a}$, and yet not be continuous at $\boldsymbol{a}$. Two functions may have all directional derivatives in some open sets, and yet their composition may not. Thus, we introduce a more uniform notion of derivative.

Indeed, there is no reason to believe that the directional derivatives at a given point $\boldsymbol{a}$ with respect to all nonzero vectors $\boldsymbol{u}$ share something in common.

A Review of Multivariate Calculus

The Derivative of a Function between Normed Spaces

Let E and F be two normed vector spaces, let A be a nonempty open subset of E, and let $f : A \to F$ be any function. For any $\boldsymbol{a} \in A$, we say that f is *differentiable at $\boldsymbol{a} \in A$* if there is a *linear continuous* map, $L : E \to F$, and a function, $\epsilon : E \to \mathbb{R}$, such that

$$f(\boldsymbol{a} + \boldsymbol{h}) = f(\boldsymbol{a}) + L(\boldsymbol{h}) + \epsilon(\boldsymbol{h}) \cdot \|\boldsymbol{h}\|$$

for every $(\boldsymbol{a} + \boldsymbol{h}) \in A$, where

$$\lim_{\boldsymbol{h} \to \boldsymbol{0},\, \boldsymbol{h} \in U} \epsilon(\boldsymbol{h}) = \boldsymbol{0}\,,$$

with

$$U = \{\boldsymbol{h} \in E \mid \boldsymbol{a} + \boldsymbol{h} \in A,\, \boldsymbol{h} \neq \boldsymbol{0}\}\,.$$

The linear map L is denoted by $Df(\boldsymbol{a})$, or $Df_{\boldsymbol{a}}$, or $df(\boldsymbol{a})$, or $df_{\boldsymbol{a}}$, or $f'(\boldsymbol{a})$, and it is called the *Fréchet derivative*, or *derivative*, or *total derivative*, or *total differential*, or *differential*, of f at $\boldsymbol{a}$.

A Review of Multivariate Calculus

The Derivative of a Function between Normed Spaces

Note that for every $\boldsymbol{h} \in U$, since $\boldsymbol{h} \neq \mathbf{0}$, the value of $\epsilon(\boldsymbol{h})$ is uniquely determined as

$$\epsilon(\boldsymbol{h}) = \frac{f(\boldsymbol{a}+\boldsymbol{h}) - f(\boldsymbol{a}) - L(\boldsymbol{h})}{\|\boldsymbol{h}\|},$$

and the value $\epsilon(\mathbf{0})$ plays absolutely no role in this definition. As a consequence, it does no harm to assume that $\epsilon(\mathbf{0})$ is equal to zero, and we will assume this from now on.

Note that *the continuous linear map L is unique* (if it exists), and this fact is implied as a corollary of the proposition given below, which shows that our new definition of derivative is consistent with the definition of the directional derivative we saw before.

A Review of Multivariate Calculus

The Derivative of a Function between Normed Spaces

For instance, if $E = M_n(\mathbb{R})$ and $F = M_n(\mathbb{R})$ and $f : E \to F$ is the function defined as $f(A) = A^T A - I$, for every $A \in E$, then $Df(A)(H) = A^T H + H^T A$, for all $H \in E$ (check it!).

A Review of Multivariate Calculus

The Derivative of a Function between Normed Spaces

Proposition 2.2. Let E and F be two normed spaces, let A be a nonempty open subset of E, and let $f : A \to F$ be any function. For any $\boldsymbol{a} \in A$, if $Df(\boldsymbol{a})$ is defined, then f is continuous at $\boldsymbol{a}$ and f has a directional derivative $D_{\boldsymbol{u}}f(\boldsymbol{a})$ for every $\boldsymbol{u} \neq \boldsymbol{0}$ in E. Furthermore,

$$D_{\boldsymbol{u}}f(\boldsymbol{a}) = Df(\boldsymbol{a})(\boldsymbol{u}) .$$

The uniqueness of L follows from Proposition 2.2. Also, when E is of finite dimension, it is easily shown that every linear map is continuous and this assumption is then redundant.

A Review of Multivariate Calculus

The Derivative of a Function between Normed Spaces

If $Df(\boldsymbol{a})$ exists for every $\boldsymbol{a} \in A$, we get a map

$$Df : A \to \mathcal{L}(E;F)\,,$$

called the *derivative of f on A*, and also denoted by df. Here, $\mathcal{L}(E;F)$ denotes the vector space of continuous linear maps from E to F.

The Derivative of a Function between Normed Spaces

When E is of finite dimension n, for any basis $(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n)$ of E, we can define the directional derivatives with respect to the vectors in the basis $(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n)$ (actually, we can also do it for an infinite basis). This way, we obtain the definition of partial derivatives below:

For any two normed spaces E and F, if E is of finite dimension n, for every basis $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n$ for E, for every $\boldsymbol{a} \in E$, for every function $f : E \to F$, the directional derivatives $D_{\boldsymbol{u}_j} f(\boldsymbol{a})$ (if they exist) are called the *partial derivatives of f with respect to the basis $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n$*. The partial derivative $D_{\boldsymbol{u}_j} f(\boldsymbol{a})$ is also denoted by $\partial_j f(\boldsymbol{a})$ or $\frac{\partial f}{\partial x_j}(\boldsymbol{a})$.

A Review of Multivariate Calculus

The Derivative of a Function between Normed Spaces

The notation $\frac{\partial f}{\partial x_j}(\boldsymbol{a})$ for a partial derivative, although customary and going back to Leibniz, is a "logical obscenity." Indeed, the variable x_j really has nothing to do with the formal definition. This is one of these situations where tradition is just too hard to overthrow!

We now consider a number of standard results about derivatives.

A Review of Multivariate Calculus

The Derivative of a Function between Normed Spaces

Proposition 2.3. Given two normed spaces E and F, if $f : E \longrightarrow F$ is a constant function, then $Df(\boldsymbol{a}) = \mathbf{0}$, for every $\boldsymbol{a} \in E$. Furthermore, if $f : E \longrightarrow F$ is a continuous affine map, then $Df(\boldsymbol{a}) = \boldsymbol{g}$, for every $\boldsymbol{a} \in E$, where $\boldsymbol{g}$ is the linear map associated with map f.

Proposition 2.4. Given a normed space E and a normed vector space F, for any two functions $f, g : E \longrightarrow F$, for every $\boldsymbol{a} \in E$, if $Df(\boldsymbol{a})$ and $Dg(\boldsymbol{a})$ exist, then $D(f+g)(\boldsymbol{a})$ and $D(\lambda f)(\boldsymbol{a})$ exist, and

$$D(f+g)(\boldsymbol{a}) = Df(a) + Dg(\boldsymbol{a}),$$

$$D(\lambda f)(\boldsymbol{a}) = \lambda \cdot Df(\boldsymbol{a}).$$

A Review of Multivariate Calculus

The Derivative of a Function between Normed Spaces

Proposition 2.5. Given three normed vector spaces E_1, E_2, and F, for any continuous bilinear map $f : E_1 \times E_2 \to F$, for every $(\boldsymbol{a}, \boldsymbol{b}) \in E_1 \times E_2$, $Df(\boldsymbol{a}, \boldsymbol{b})$ exists, and for every $\boldsymbol{u} \in E_1$ and $\boldsymbol{v} \in E_2$,

$$Df(\boldsymbol{a}, \boldsymbol{b})(\boldsymbol{u}, \boldsymbol{v}) = f(\boldsymbol{u}, \boldsymbol{b}) + f(\boldsymbol{a}, \boldsymbol{v}) .$$

A Review of Multivariate Calculus

The Derivative of a Function between Normed Spaces

We now state the very useful chain rule.

Theorem 2.6. Given three normed spaces E, F, and G, let A be an open set in E, and let B an open set in F. For any functions $f : A \to F$ and $g : B \to G$, such that $f(A) \subseteq B$, for any $\mathbf{a} \in A$, if $Df(\mathbf{a})$ exists and $Dg(f(\mathbf{a}))$ exists, then $D(g \circ f)(\mathbf{a})$ exists, and

$$D(g \circ f)(\mathbf{a}) = Dg(f(\mathbf{a})) \circ Df(\mathbf{a}).$$

Theorem 2.6 has many interesting consequences. We mention one corollary.

Proposition 2.7. Given two normed spaces E and F, let A be some open subset in E, let B be some open subset in F, let $f : A \to B$ be a bijection from A to B, and assume that Df exists on A and that Df^{-1} exists on B. Then, for every point $\mathbf{a} \in A$, we have that

$$Df^{-1}(f(\mathbf{a})) = (Df(\mathbf{a}))^{-1}.$$

A Review of Multivariate Calculus

The Derivative of a Function between Normed Spaces

Proposition 2.7 has the remarkable consequence that the two vector spaces E and F have the same dimension. In other words, the existence of a bijection f between an open set A of E and an open set B of F, such that f is differentiable on A and f^{-1} is differentiable on B, implies that the two vector spaces E and F have the same dimension.

If both E and F are of finite dimension, for any basis $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n$ of E and any basis $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ of F, every function $f : E \to F$ is determined by m functions $f_i : E \to \mathbb{R}$, where

$$f(\boldsymbol{x}) = f_1(\boldsymbol{x})\boldsymbol{v}_1 + \cdots + f_m(\boldsymbol{x})\boldsymbol{v}_m\,,$$

for every $\boldsymbol{x} \in E$.

A Review of Multivariate Calculus

The Derivative of a Function between Normed Spaces

Then, we get

$$Df(\boldsymbol{a})(\boldsymbol{u}_j) = Df_1(\boldsymbol{a})(\boldsymbol{u}_j)\boldsymbol{v}_1 + \cdots + Df_i(\boldsymbol{a})(\boldsymbol{u}_j)\boldsymbol{v}_i + \cdots + Df_m(\boldsymbol{a})(\boldsymbol{u}_j)\boldsymbol{v}_m\,,$$

that is,

$$Df(\boldsymbol{a})(\boldsymbol{u}_j) = \partial_j f_1(\boldsymbol{a})\boldsymbol{v}_1 + \cdots + \partial_j f_i(\boldsymbol{a})\boldsymbol{v}_i + \cdots + \partial_j f_m(\boldsymbol{a})\boldsymbol{v}_m\,.$$

A Review of Multivariate Calculus

The Derivative of a Function between Normed Spaces

Since the j-th column of the $m \times n$-matrix representing $Df(\boldsymbol{a})$ with respect to the bases $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n$ and $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m$ is equal to the components of the vector $Df(\boldsymbol{a})(\boldsymbol{u}_j)$ over the basis $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m$, the linear map $Df(\boldsymbol{a})$ is determined by the $m \times n$-matrix $J(f)(\boldsymbol{a})$:

$$J(f)(\boldsymbol{a}) = \left(\frac{\partial f_i}{\partial x_j}(\boldsymbol{a})\right) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(\boldsymbol{a}) & \frac{\partial f_1}{\partial x_2}(\boldsymbol{a}) & \cdots & \frac{\partial f_1}{\partial x_n}(\boldsymbol{a}) \\ \frac{\partial f_2}{\partial x_1}(\boldsymbol{a}) & \frac{\partial f_2}{\partial x_2}(\boldsymbol{a}) & \cdots & \frac{\partial f_2}{\partial x_n}(\boldsymbol{a}) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1}(\boldsymbol{a}) & \frac{\partial f_m}{\partial x_2}(\boldsymbol{a}) & \cdots & \frac{\partial f_m}{\partial x_n}(\boldsymbol{a}) \end{pmatrix}.$$

A Review of Multivariate Calculus

The Derivative of a Function between Normed Spaces

This matrix is called the *Jacobian matrix of Df at $\boldsymbol{a}$*. When $m = n$, the determinant, $\det\big(J(f)(\boldsymbol{a})\big)$, of $J(f)(\boldsymbol{a})$ is called the *Jacobian of $Df(\boldsymbol{a})$*. This determinant only depends on $Df(\boldsymbol{a})$, and not on specific bases. Partial derivatives give a means for computing it.

When $E = \mathbb{R}^n$ and $F = \mathbb{R}^m$, for any function $f : \mathbb{R}^n \to \mathbb{R}^m$, it is easy to compute the partial derivatives $\frac{\partial f_i}{\partial x_j}(\boldsymbol{a})$. We simply treat the function $f_i : \mathbb{R}^n \to \mathbb{R}$ as a function of its j-th argument, leaving the others fixed, and compute the derivative as the usual derivative.

The Derivative of a Function between Normed Spaces

In the case where $E = \mathbb{R}$, for any function $f : \mathbb{R} \rightarrow F$, the Jacobian matrix of $Df(\boldsymbol{a})$ is a column vector. In fact, this column vector is just $D_{\mathbf{1}}f(\boldsymbol{a})$. Then, for every $\lambda \in \mathbb{R}$,

$$Df(\boldsymbol{a})(\lambda) = \lambda \cdot D_{\mathbf{1}}f(\boldsymbol{a})\,.$$

This case is sufficiently important to warrant a definition.

Definition 2.7. Given a function $f : \mathbb{R} \rightarrow F$, where F is a normed space, the vector

$$Df(\boldsymbol{a})(\mathbf{1}) = D_{\mathbf{1}}f(\boldsymbol{a})$$

is called the *vector derivative or velocity vector (in the real case) at $\boldsymbol{a}$*. We usually identify $Df(\boldsymbol{a})$ with its Jacobian matrix $D_{\mathbf{1}}f(\boldsymbol{a})$, which is the column vector corresponding to $D_{\mathbf{1}}f(\boldsymbol{a})$.

A Review of Multivariate Calculus

The Derivative of a Function between Normed Spaces

By abuse of notation, we also let $Df(\boldsymbol{a})$ denote the vector $Df(\boldsymbol{a})(\mathbf{1}) = D_{\mathbf{1}}f(\boldsymbol{a})$.

When $E = \mathbb{R}$, the physical interpretation is that f defines a (parametric) curve that is the trajectory of some particle moving in $\mathbb{R}^m$ as a function of time, and the vector $D_{\mathbf{1}}f(\boldsymbol{a})$ is the *velocity* of the moving particle $f(t)$ at $t = \boldsymbol{a}$.

A Review of Multivariate Calculus

The Derivative of a Function between Normed Spaces

When $A = (0,1)$ and $F = \mathbb{R}^3$, a function

$$f : (0,1) \longrightarrow \mathbb{R}^3$$

defines a (parametric) curve in $\mathbb{R}^3$. If $f = (f_1, f_2, f_3)$, its Jacobian matrix at $\boldsymbol{a} \in \mathbb{R}$ is

$$J(f)(\boldsymbol{a}) = \begin{pmatrix} \frac{\partial f_1}{\partial t}(\boldsymbol{a}) \\ \frac{\partial f_2}{\partial t}(\boldsymbol{a}) \\ \frac{\partial f_3}{\partial t}(\boldsymbol{a}) \end{pmatrix} .$$

A Review of Multivariate Calculus

The Derivative of a Function between Normed Spaces

When $E = \mathbb{R}^2$ and $F = \mathbb{R}^3$, a function

$$\varphi : \mathbb{R}^2 \to \mathbb{R}^3$$

defines a parametric surface. Letting $\varphi = (f, g, h)$, its Jacobian matrix at $\boldsymbol{a} \in \mathbb{R}^2$ is

$$J(\varphi)(\boldsymbol{a}) = \begin{pmatrix} \frac{\partial f}{\partial u}(\boldsymbol{a}) & \frac{\partial f}{\partial v}(\boldsymbol{a}) \\ \frac{\partial g}{\partial u}(\boldsymbol{a}) & \frac{\partial g}{\partial v}(\boldsymbol{a}) \\ \frac{\partial h}{\partial u}(\boldsymbol{a}) & \frac{\partial h}{\partial v}(\boldsymbol{a}) \end{pmatrix} .$$

A Review of Multivariate Calculus

The Derivative of a Function between Normed Spaces

When $E = \mathbb{R}^3$ and $F = \mathbb{R}$, for a function

$$f : \mathbb{R}^3 \to \mathbb{R}\,,$$

the Jacobian matrix at $\boldsymbol{a} \in \mathbb{R}^3$ is

$$J(f)(\boldsymbol{a}) = \begin{pmatrix} \frac{\partial f}{\partial x}(\boldsymbol{a}) & \frac{\partial f}{\partial y}(\boldsymbol{a}) & \frac{\partial f}{\partial z}(\boldsymbol{a}) \end{pmatrix}.$$

A Review of Multivariate Calculus

The Derivative of a Function between Normed Spaces

More generally, when

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

the Jacobian matrix at $\boldsymbol{a} \in \mathbb{R}^n$ is the row vector

$$J(f)(\boldsymbol{a}) = \begin{pmatrix} \frac{\partial f}{\partial x_1}(\boldsymbol{a}) & \cdots & \frac{\partial f}{\partial x_n}(\boldsymbol{a}) \end{pmatrix}.$$

Its transpose is a column vector called the *gradient* of f at $\boldsymbol{a}$, denoted by $\operatorname{grad} f(\boldsymbol{a})$ or $\nabla f(\boldsymbol{a})$. Then, given any $\boldsymbol{v} \in \mathbb{R}^n$, note that $Df(\boldsymbol{a})(\boldsymbol{v})$ is the scalar product of $\operatorname{grad} f(\boldsymbol{a})$ and $\boldsymbol{v}$:

$$Df(\boldsymbol{a})(\boldsymbol{v}) = \frac{\partial f}{\partial x_1}(\boldsymbol{a}) \cdot v_1 + \cdots + \frac{\partial f}{\partial x_n}(\boldsymbol{a}) \cdot v_n = \langle \operatorname{grad} f(\boldsymbol{a}), \boldsymbol{v} \rangle\,.$$

In matrix form, we write

$$Df(\boldsymbol{a})(\boldsymbol{v}) = \left(\operatorname{grad} f(\boldsymbol{a})\right)^{\mathrm{T}} \cdot \boldsymbol{v}\,.$$

A Review of Multivariate Calculus

The Derivative of a Function between Normed Spaces

Let E, F, and G be finite dimensional vector spaces, $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_p$ be a basis for E, $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ be a basis for F, $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_m$ be a basis for G, A be an open subset of E, and B be an open subset of F. Then, for any functions $f : A \to F$ and $g : B \to G$, such that $f(A) \subseteq B$, for any $\boldsymbol{a} \in A$, letting $\boldsymbol{b} = f(\boldsymbol{a})$ and $h = g \circ f$, if $Df(\boldsymbol{a})$ exists and $Dg(\boldsymbol{b})$ exists, by Theorem 2.6, the Jacobian matrix $J(h)(\boldsymbol{a}) = J(g \circ f)(\boldsymbol{a})$ w.r.t. the bases

$$\boldsymbol{u}_1, \ldots, \boldsymbol{u}_p \qquad \text{and} \qquad \boldsymbol{w}_1, \ldots, \boldsymbol{w}_m$$

is the product, $J(g)(\boldsymbol{b}) \cdot J(f)(\boldsymbol{a})$, of the Jacobian matrices, $J(g)(\boldsymbol{b})$ and $J(f)(\boldsymbol{a})$, w.r.t the bases

$$\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n \qquad \text{and} \qquad \boldsymbol{w}_1, \ldots, \boldsymbol{w}_m\,,$$

respectively:

A Review of Multivariate Calculus

The Derivative of a Function between Normed Spaces

$$J(h)(\boldsymbol{a}) = \begin{pmatrix} \frac{\partial g_1}{\partial y_1}(\boldsymbol{b}) & \frac{\partial g_1}{\partial y_2}(\boldsymbol{b}) & \cdots & \frac{\partial g_1}{\partial y_n}(\boldsymbol{b}) \\ \frac{\partial g_2}{\partial y_1}(\boldsymbol{b}) & \frac{\partial g_2}{\partial y_2}(\boldsymbol{b}) & \cdots & \frac{\partial g_2}{\partial y_n}(\boldsymbol{b}) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial g_m}{\partial y_1}(\boldsymbol{b}) & \frac{\partial g_n}{\partial y_2}(\boldsymbol{b}) & \cdots & \frac{\partial g_m}{\partial y_n}(\boldsymbol{b}) \end{pmatrix} \cdot \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(\boldsymbol{a}) & \frac{\partial f_1}{\partial x_2}(\boldsymbol{a}) & \cdots & \frac{\partial f_1}{\partial x_p}(\boldsymbol{a}) \\ \frac{\partial f_2}{\partial x_1}(\boldsymbol{a}) & \frac{\partial f_2}{\partial x_2}(\boldsymbol{a}) & \cdots & \frac{\partial f_2}{\partial x_p}(\boldsymbol{a}) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1}(\boldsymbol{a}) & \frac{\partial f_n}{\partial x_2}(\boldsymbol{a}) & \cdots & \frac{\partial f_n}{\partial x_p}(\boldsymbol{a}) \end{pmatrix}.$$

Thus, we have the familiar formula

$$\frac{\partial h_i}{\partial x_j}(\boldsymbol{a}) = \sum_{k=1}^{n} \frac{\partial g_i}{\partial y_k}(\boldsymbol{b}) \cdot \frac{\partial f_k}{\partial x_j}(\boldsymbol{a}).$$

A Review of Multivariate Calculus

The Derivative of a Function between Normed Spaces

Given two normed spaces E and F of finite dimension, given an open subset A of E, if a function $f : A \to F$ is differentiable at $\boldsymbol{a} \in A$, then its Jacobian matrix is well defined.

The converse is false. There are functions such that all the partial derivatives exist at some $\boldsymbol{a} \in A$, but yet the function is not differentiable at $\boldsymbol{a}$, and not even continuous at $\boldsymbol{a}$.

However, there are sufficient conditions on the partial derivatives for $Df(\boldsymbol{a})$ to exist, namely, continuity of the partial derivatives. If f is differentiable on A, then f defines a function,

$$Df : A \to \mathcal{L}(E;F).$$

It turns out that the continuity of the partial derivatives on A is a necessary and sufficient condition for Df to exist and to be continuous on A, as stated in the following theorem:

The Derivative of a Function between Normed Spaces

Theorem 2.8. Given two normed spaces E and F, where E is of finite dimension n and where

$$(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n)$$

is a basis of E, given any open subset A of E and any function $f : A \to F$, the derivative

$$Df : A \to \mathcal{L}(E; F)$$

is defined and continuous on A iff every partial derivative $\frac{\partial f}{\partial x_j}$ is defined and continuous on A, for all j, $1 \leq j \leq n$. As a corollary, if F is of finite dimension m, and $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m$ is a basis of F, the derivative $Df : A \to \mathcal{L}(E; F)$ is defined and continuous on A iff every partial derivative $\frac{\partial f_i}{\partial x_j}$ is defined and continuous on A, for all i, j, $1 \leq i \leq m, 1 \leq j \leq n$.

The Derivative of a Function between Normed Spaces

Definition 2.8. Given two normed vector spaces E and F, and an open subset A of E, we say that a function $f : A \to F$ is a *C^0-function on A* if f is continuous on A. We say that $f : A \to F$ is a *C^1-function on A* if the derivative Df exists and is continuous on A.

Let E and F be two normed vector spaces, let $A \subseteq E$ be an open subset of E and let $f : E \to F$ be a function such that $Df(\boldsymbol{a})$ exists for all $\boldsymbol{a} \in A$. If $Df(\boldsymbol{a})$ is *injective* for all $\boldsymbol{a} \in A$, we say that f is an *immersion* (on A) and if $Df(\boldsymbol{a})$ is *surjective* for all $\boldsymbol{a} \in A$, we say that f is a *submersion* (on A).

When E and F are finite dimensional with $\dim(E) = n$ and $\dim(F) = m$, if $m \geq n$, then f is an immersion iff the Jacobian matrix, $J(f)(\boldsymbol{a})$, has full rank (n) for all $\boldsymbol{a} \in E$ and if $n \geq m$, then f is a submersion iff the Jacobian matrix, $J(f)(\boldsymbol{a})$, has full rank (m), for all $\boldsymbol{a} \in E$.

A Review of Multivariate Calculus

The Derivative of a Function between Normed Spaces

A very important theorem is the inverse function theorem. In order for this theorem to hold for infinite dimensional spaces, it is necessary to assume that our normed spaces are complete.

Given a normed vector space, E, we say that a sequence, $(\boldsymbol{u}_n)_n$, with $\boldsymbol{u}_n \in E$, is a *Cauchy sequence* iff for every $\epsilon > 0$, there is some $N > 0$ so that for all $m, n \geq N$, we have

$$\|\boldsymbol{u}_n - \boldsymbol{u}_m\| < \epsilon\,.$$

A normed vector space, E, is said to be *complete* iff every Cauchy sequence converges.

A Review of Multivariate Calculus

The Derivative of a Function between Normed Spaces

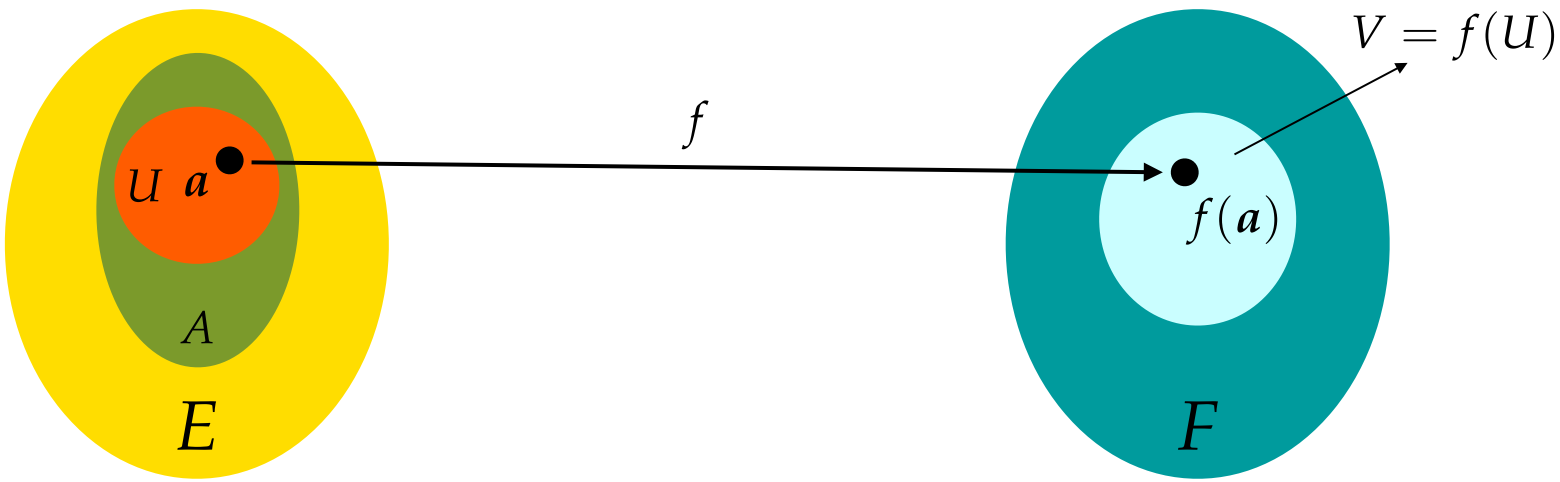
A complete normed vector space is also called a *Banach space*, after Stefan Banach (1892-1945).

Fortunately, $\mathbb{R}$, and every finite dimensional (real or complex) normed vector space is complete.

A real (resp. complex) vector space, E, is a real (resp. complex) *Hilbert space* if it is complete as a normed space with the norm $\boldsymbol{u} = \sqrt{\langle \boldsymbol{u}, \boldsymbol{u} \rangle}$ induced by its Euclidean (resp. Hermitian) inner product.

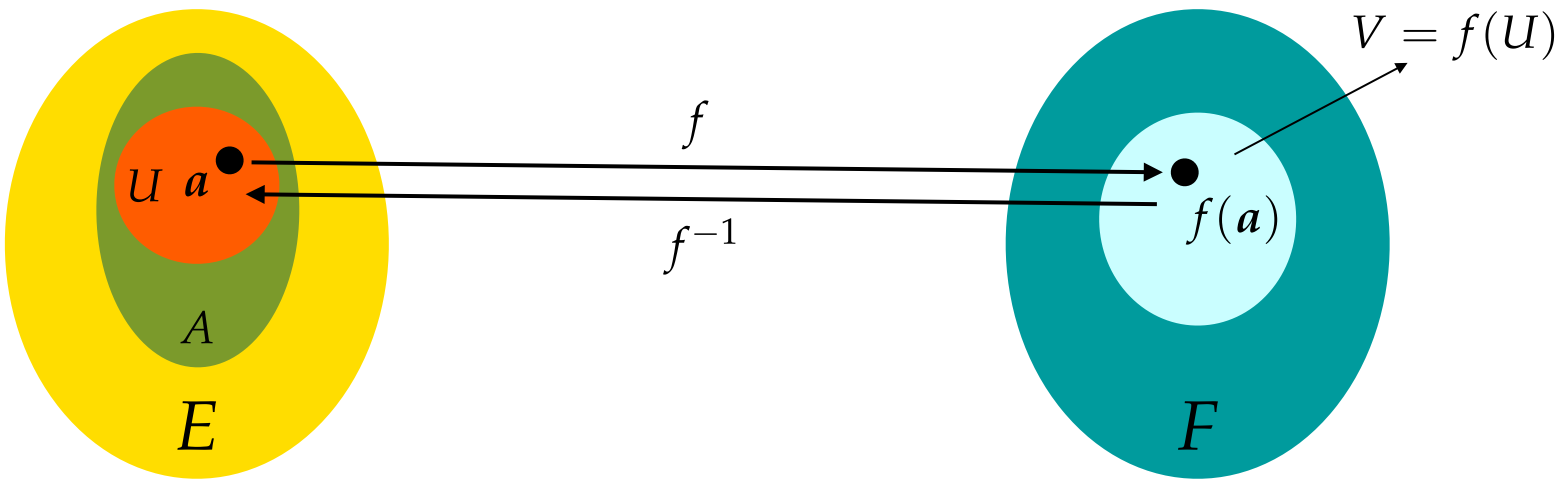
The Derivative of a Function between Normed Spaces

Definition 2.9. Given two topological spaces E and F and an open subset A of E, we say that a function $f : A \to F$ is a *local homeomorphism from A to F* if for every $\boldsymbol{a} \in A$, there is an open set $U \subseteq A$ containing $\boldsymbol{a}$ and an open set V containing $f(\boldsymbol{a})$ such that f is a homeomorphism from U to $V = f(U)$. If B is an open subset of F, we say that $f : A \to F$ is a *(global) homeomorphism from A to B* if f is a homeomorphism from A to $B = f(A)$.



The Derivative of a Function between Normed Spaces

Definition 2.10. If E and F are normed spaces, we say that $f : A \longrightarrow F$ is a *local diffeomorphism from A to F* if for every $\boldsymbol{a} \in A$, there is an open set $U \subseteq A$ containing $\boldsymbol{a}$ and an open set V containing $f(\boldsymbol{a})$ such that f is a bijection from U to V, f is a C^1-function on U, and f^{-1} is a C^1-function on $V = f(U)$. We say that $f : A \longrightarrow F$ is a *(global) diffeomorphism from A to B* if f is a homeomorphism from A to $B = f(A)$, f is a C^1-function on A, and f^{-1} is a C^1-function on B.



A Review of Multivariate Calculus

The Derivative of a Function between Normed Spaces

Note that a local diffeomorphism is a local homeomorphism. Also, as a consequence of Theorem 2.6, if f is a diffeomorphism on A, then $Df(\boldsymbol{a})$ is a bijection for every $\boldsymbol{a} \in A$.

Recall...

Theorem 2.6. Given three normed spaces E, F, and G, let A be an open set in E, and let B an open set in F. For any functions $f : A \longrightarrow F$ and $g : B \longrightarrow G$, such that $f(A) \subseteq B$, for any $\boldsymbol{a} \in A$, if $Df(\boldsymbol{a})$ exists and $Dg(f(\boldsymbol{a}))$ exists, then $D(g \circ f)(\boldsymbol{a})$ exists, and

$$D(g \circ f)(\boldsymbol{a}) = Dg(f(\boldsymbol{a})) \circ Df(\boldsymbol{a}) \,.$$

The Derivative of a Function between Normed Spaces

Theorem 2.9. (**Inverse Function Theorem**) Let E and F be complete normed spaces, let A be an open subset of E, and let $f : A \to F$ be a C^1-function on A. The following properties hold:

(1) For every $\boldsymbol{a} \in A$, if $Df(\boldsymbol{a})$ is invertible, then there exist some open subset $U \subseteq A$ containing $\boldsymbol{a}$, and some open subset V of F containing $f(\boldsymbol{a})$, such that f is a diffeomorphism from U to $V = f(U)$. Furthermore,

$$Df^{-1}(f(\boldsymbol{a})) = (Df(\boldsymbol{a}))^{-1}.$$

For every neighborhood N of $\boldsymbol{a}$, the image $f(N)$ of N is a neighborhood of $f(\boldsymbol{a})$, and for every open ball $U \subseteq A$ of center $\boldsymbol{a}$, the image $f(U)$ of U contains some open ball of center $f(\boldsymbol{a})$.

(2) If $Df(\boldsymbol{a})$ is invertible for every $\boldsymbol{a} \in A$, then $B = f(A)$ is an open subset of F, and f is a local diffeomorphism from A to B. Furthermore, if f is injective, then f is a diffeomorphism from A to B.

A Review of Multivariate Calculus

The Derivative of a Function between Normed Spaces

Part (1) of Theorem 2.9 is often referred to as the "(local) inverse function theorem." It plays an important role in the study of manifolds and (ordinary) differential equations.

If E and F are both of finite dimension, the case where $Df(\boldsymbol{a})$ is just injective or just surjective is also important for defining manifolds, which is done using implicit definitions.