derivative

Canonical name	Derivative
Date of creation	2013-03-22 12:41:41
Last modified on	2013-03-22 12:41:41
Owner	rmilson (146)
Last modified by	rmilson (146)
Numerical id	35
Author	rmilson (146)
Entry type	Definition
Classification	msc 26B05
Classification	msc 46G05
Classification	msc 26A24
Related topic	TableOfDerivatives
Related topic	DerivativeOfInverseFunction
Related topic	PartialDerivative
Related topic	Gradient
Related topic	RelatedRates
Related topic	CalculatingLipschitzRatios
Defines	directional derivative
Defines	Fréchet derivative

Qualitatively the *derivative* is a of the change of a function in a small around a specified point.

Motivation

The idea behind the derivative comes from the straight line. What characterizes a straight line is the fact that it has constant "slope". In other words,

Figure 1: The straight line $y = mx + b$

for a line given by the equation $y = mx + b$, as in Fig. **??**, the ratio of Δy over Δx is always constant and has the value $\dfrac{\Delta y}{\Delta x} = m$.

Figure 2: The parabola $y = x^2$ and its tangent at (x_0, y_0)

For other curves we cannot define a "slope", like for the straight line, since such a quantity would not be constant. However, for sufficiently smooth curves, each point on a curve has a tangent line. For example consider the curve $y = x^2$, as in Fig. **??**. At the point (x_0, y_0) on the curve, we can draw a tangent of slope m given by the equation $y - y_0 = m(x - x_0)$.

Suppose we have a curve of the form $y = f(x)$, and at the point $(x_0, f(x_0))$ we have a tangent given by $y - y_0 = m(x - x_0)$. Note that for values of x sufficiently close to x_0 we can make the approximation $f(x) \approx m(x-x_0)+y_0$. So the slope m of the tangent describes how much $f(x)$ changes in the vicinity of x_0. It is the slope of the tangent that will be associated with the derivative of the function $f(x)$.

Formal definition

More formally for any real function $f\colon \mathbb{R} \to \mathbb{R}$, we define the *derivative* of f at the point x as the following limit (if it exists)

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

This definition turns out to be with the motivation introduced above.

The derivatives for some elementary functions are (cf. derivative notation)

1. $\frac{d}{dx}c = 0, \qquad \text{where } c \text{ is constant;}$
2. $\frac{d}{dx}x^n = nx^{n-1};$
3. $\frac{d}{dx}\sin x = \cos x;$
4. $\frac{d}{dx}\cos x = -\sin x;$
5. $\frac{d}{dx}e^x = e^x;$
6. $\frac{d}{dx}\ln x = \frac{1}{x}.$

While derivatives of more complicated expressions can be calculated algorithmically using the following rules

Linearity $\frac{d}{dx}\left(af(x) + bg(x)\right) = af'(x) + bg'(x);$

Product rule $\frac{d}{dx}\left(f(x)g(x)\right) = f'(x)g(x) + f(x)g'(x);$

Chain rule $\frac{d}{dx}g(f(x)) = g'(f(x))f'(x);$

Quotient Rule $\frac{d}{dx}\frac{f(x)}{g(x)} = \frac{f'(x)g(x) - f(x)g'(x)}{g(x)^2}.$

Note that the quotient rule, although given as much importance as the other rules in elementary calculus, can be derived by succesively applying the product rule and the chain rule to $\frac{f(x)}{g(x)} = f(x)\frac{1}{g(x)}$. Also the quotient rule does not generalize as well as the other ones.

Since the derivative $f'(x)$ of $f(x)$ is also a function x, higher derivatives can be obtained by applying the same procedure to $f'(x)$ and so on.

Generalization

Banach Spaces

Unfortunately the notion of the "slope of the tangent" does not directly generalize to more abstract situations. What we can do is keep in mind the facts that the tangent is a linear function and that it approximates the function near the point of tangency, as well as the formal definition above.

Very general conditions under which we can define a derivative in a manner much similar to the above areas follows. Let $f\colon \mathsf{V} \to \mathsf{W}$, where V and W are Banach spaces. Let $\mathbf{h} \neq 0$ be an element of V. We define the *directional derivative* $(D_{\mathbf{h}} f)(\mathbf{x})$ at $\mathbf{x}$ as the following limit (when it exists):

$$(D_{\mathbf{h}} f)(\mathbf{x}) := \lim_{\epsilon \to 0} \frac{f(\mathbf{x} + \epsilon \mathbf{h}) - f(\mathbf{x})}{\epsilon},$$

where ϵ is a scalar. Note that $f(x + \epsilon \mathbf{h}) \approx f(\mathbf{x}) + \epsilon (D_{\mathbf{h}} f)(\mathbf{x})$, which is with our original motivation. In certain contexts, this directional derivative is also called the *Gâteaux derivative*.

Finally we define the *derivative* at $\mathbf{x}$ as the bounded linear map $(Df)(\mathbf{x})\colon \mathsf{V} \to \mathsf{W}$ such that for any non-zero $\mathbf{h} \in \mathsf{V}$

$$\lim_{\|\mathbf{h}\| \to 0} \frac{(f(\mathbf{x} + \mathbf{h}) - f(\mathbf{x})) - (Df)(\mathbf{x}) \cdot \mathbf{h}}{\|\mathbf{h}\|} = 0.$$

Once again we have $f(\mathbf{x} + \mathbf{h}) \approx f(\mathbf{x}) + (Df)(\mathbf{x}) \cdot \mathbf{h}$. In fact, if the derivative $(Df)(\mathbf{x})$ exists, the directional derivatives can be obtained as $(D_{\mathbf{h}} f)(\mathbf{x}) = (Df)(\mathbf{x}) \cdot \mathbf{h}$.[1] However, the existence of $(D_{\mathbf{h}} f)$ for each non-zero $\mathbf{h} \in \mathsf{V}$ does not guarantee the existence of $(Df)(\mathbf{x})$. This derivative is also called the *Fréchet derivative*. In the more familiar case $f\colon \mathbb{R}^n \to \mathbb{R}^m$, the derivative Df is simply the Jacobian of f.

Under these general conditions the following properties of the derivative remain

1. $D\mathbf{h} = 0, \qquad$ where $\mathbf{h}$ is a constant;
2. $D(A \cdot \mathbf{x}) = A, \qquad$ where A is linear.

[1] The notation $A \cdot \mathbf{h}$ is used when $\mathbf{h}$ is a vector and A a linear operator. This notation can be considered advantageous to the usual notation $A(\mathbf{h})$, since the latter is rather bulky and the former incorporates the intuitive distributive properties of linear operators also associated with usual multiplication.

Linearity $D(af(\mathbf{x}) + bg(\mathbf{x})) \cdot \mathbf{h} = a(Df)(\mathbf{x}) \cdot \mathbf{h} + b(Dg)(\mathbf{x}) \cdot \mathbf{h}$;

"Product" rule $D(B(f(\mathbf{x}), g(\mathbf{x}))) \cdot \mathbf{h} = B((Df)(\mathbf{x}) \cdot \mathbf{h}, g(\mathbf{x})) + B(f(\mathbf{x}), (Dg)(\mathbf{x}) \cdot \mathbf{h})$, where B is bilinear;

Chain rule $D(g(f(\mathbf{x})) \cdot \mathbf{h} = (Dg)(f(\mathbf{x})) \cdot ((Df)(\mathbf{x}) \cdot \mathbf{h})$.

Note that the derivative of f can be seen as a function $Df \colon \mathsf{V} \to L(\mathsf{V}, \mathsf{W})$ given by $Df \colon \mathbf{x} \mapsto (Df)(\mathbf{x})$, where $L(\mathsf{V}, \mathsf{W})$ is the space of bounded linear maps from V to W. Since $L(\mathsf{V}, \mathsf{W})$ can be considered a Banach space itself with the norm taken as the operator norm, higher derivatives can be obtained by applying the same procedure to Df and so on.

0.1 Partial derivatives

A straightforward extension of the derivatives defined above is that of partial derivatives for functions of several independent variables. Partial derivatives have numerous applications, as for example in physics and engineering; wave equations are among such important examples of the use of partial derivatives in physics and engineering.

Manifolds

Let V be a Banach space (for finite dimensional manifolds $\mathsf{V} = \mathbb{R}^n$). A manifold modeled on V is a topological space that is locally homeomorphic to V and is endowed with enough structure to define derivatives. Since the notion of a manifold was constructed specifically to generalize the notion of a derivative, this seems like the end of the road for this entry. The following discussion is rather technical, a more intuitive explanation of the same concept can be found in the entry on related rates.

Consider manifolds V and W modeled on Banach spaces V and W, respectively. Say we have $y = f(x)$ for some $x \in V$ and $y \in W$, then, by definition of a manifold, we can find charts $(X, \mathbf{x})$ and $(Y, \mathbf{y})$, where X and Y are neighborhoods of x and y, respectively. These charts provide us with canonical isomorphisms between the Banach spaces V and W, and the respective tangent spaces T_xV and T_yW:

$$\mathrm{d}\mathbf{x}_x \colon T_xV \to \mathsf{V}, \quad \mathrm{d}\mathbf{y}_y \colon T_yW \to \mathsf{W}.$$

Now consider a map $f\colon V \to W$ between the manifolds. By composing it with the chart maps we construct the map

$$g_{(X,\mathbf{x})}^{(Y,\mathbf{y})} = \mathbf{y} \circ f \circ \mathbf{x}^{-1} \colon \mathsf{V} \to \mathsf{W},$$

defined on an appropriately domain. Since we now have a map between Banach spaces, we can define its derivative at $\mathbf{x}(x)$ in the sense defined above, namely $Dg_{(X,\mathbf{x})}^{(Y,\mathbf{y})}(\mathbf{x}(x))$. If this derivative exists for every choice of admissible charts $(X, \mathbf{x})$ and $(Y, \mathbf{y})$, we can say that the derivative of $Df(x)$ of f at x is defined and given by

$$Df(x) = \mathrm{d}\mathbf{y}_y^{-1} \circ Dg_{(X,\mathbf{x})}^{(Y,\mathbf{y})}(\mathbf{x}(x)) \circ \mathrm{d}\mathbf{x}_x$$

(it can be shown that this is well defined and independent of the choice of charts).

Note that the derivative is now a map between the tangent spaces of the two manifolds $Df(x)\colon T_xV \to T_yW$. Because of this a common notation for the derivative of f at x is T_xf. Another alternative notation for the derivative is $f_{*,x}$ because of its connection to the category-theoretical pushforward.

Distributions

Derivatives can also be generalized in less "smooth" contexts. For example the derivative is one of `http://planetmath.org/OperationsOnDistributions`operation that can be defined for distributions.

Standard connection of $\mathbb{R}^n$

Let Ω be an open set in $\mathbb{R}^n$. There is an operator on vectors fields in Ω which measure how a pair of them, $X, Y : \Omega \to \mathbb{R}^n$ vary, one with respect to the other:

$$D_X Y = (JY)X$$

Here JY is the Jacobian of Y, so when we multiply, we can see that the components of $D_X Y$ are the directional variations of the components of Y in the direction X.

Additional Topic

- Non-Newtonian calculus