维基百科
自由的百科全书

偏导数

维基百科，自由的百科全书

系列条目
微积分学
$$f(x)=\frac{\mathrm{d}}{\mathrm{d}x}\int_a^x f(t)\,\mathrm{d}t$$
函数 · 极限论 · 微分学 · 积分 微积分基本定理 · 微积分发现权之争
基础概念（含极限论和级数论） [展开]
一元微分 [展开]
一元积分 [展开]
多元微积分 [展开]
微分方程 [展开]
相关数学家 [展开]
历史名作 [展开]
分支学科 [展开]

在数学中，**偏导数**（英语：partial derivative）的定义是：一个多变量的函数（或称多元函数），对其中一个变量（导数）微分，而保持其他变量恒定[注 1]。

偏导数的作用与价值在向量分析和微分几何以及机器学习领域中受到广泛认可。

函数f关于变量x的偏导数写为f_x'或$\frac{\partial f}{\partial x}$。**偏导数符号**$\partial$是全导数符号$d$的变体，由阿德里安-马里·勒让德引入，并在雅可比的重新引入后得到普遍接受。

简介

假设f是一个多元函数。例如：

$$z=f(x,y)=x^2+xy+y^2$$

因为曲面上的每一点都有无穷多条切线，描述这种函数的导数相当困难。偏导数就是选择其中一条切线，并求出它的斜率。通常，最感兴趣的是垂直于y轴（平行于xOz平面）的切线，以及垂直于x轴（平行于yOz平面）的切线。

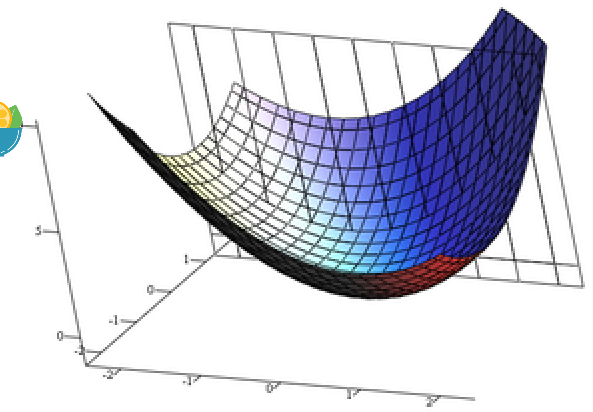
一种求出这些切线的好办法是把其他变量视为常数。例如，欲求出以上的函数在点$(1, 1)$的与xOz平面平行的切线。右图中显示了函数的图像以及这个平面。左图中显示了函数在平面$y = 1$上是什么样的。我们把变量y视为常数，通过对方程求导，我们可以发现f在点(x, y)的导数，记为：

$$\frac{\partial f}{\partial x}=2x+y$$

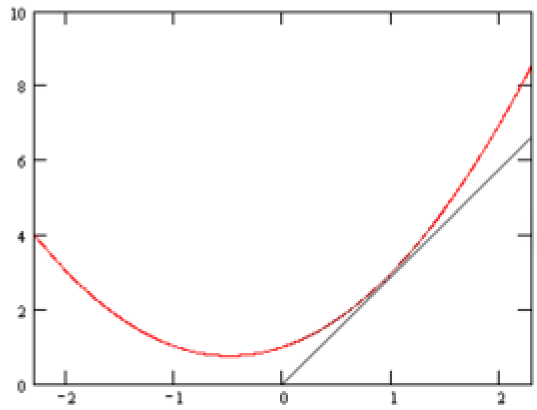
于是在点(1, 1)的xOz平面平行的切线的斜率是3。

$$\frac{\partial f}{\partial x} = 3$$

在点(1, 1)，或称“f在(1, 1)的关于x的偏导数是3”。



$f = x^2 + xy + y^2$的图像。我们希望求出函数在点(1, 1)的对x的偏导数；对应的切线与xOz平面平行。



这是上图中$y = 1$时的图像片段。

定义

函数f可以解释为y为自变量而x为常数的函数：

$$f(x,y) = f_x(y) = x^2 + xy + y^2。$$

也就是说，每一个x的值定义了一个函数，记为f_x，它是一个一元函数。也就是说：

$$f_x(y) = x^2 + xy + y^2。$$

一旦选择了一个x的值，例如a，那么$f(x,y)$便定义了一个函数f_a，把y映射到$a^2 + ay + y^2$：

$$f_a(y) = a^2 + ay + y^2。$$

在这个表达式中，a是**常数**，而不是**变量**，因此f_a是只有一个变量的函数，这个变量是y。这样，便可以使用一元函数的导数的定义：

$$f_a'(y) = a + 2y$$

以上的步骤适用于任何a的选择。把这些导数合并起来，便得到了一个函数，它描述了f在y方向上的变化：

$$\frac{\partial f}{\partial y}(x,y) = x + 2y$$

这就是f关于y的偏导数，在这里，∂是一个弯曲的d，称为**偏导数符号**。为了把它与字母d区分，∂有时读作“der”、“del”、“dah”或“偏”，而不是“dee”。

一般地，函数$f(x_1,...,x_n)$在点$(a_1,...,a_n)$关于x_i的偏导数定义为：

$$\frac{\partial f}{\partial x_i}(a_1,\ldots,a_n) = \lim_{h\to 0}\frac{f(a_1,\ldots,a_i+h,\ldots,a_n) - f(a_1,\ldots,a_n)}{h}$$

在以上的差商中，除了x_i以外的所有变量都是固定的。这个固定值的选择决定了一个一元函数$f_{a_1,\ldots,a_{i-1},a_{i+1},\ldots,a_n}(x_i) = f(a_1,\ldots,a_{i-1},x_i,a_{i+1},\ldots,a_n)$，根据定义，

$$\frac{df_{a_1,\ldots,a_{i-1},a_{i+1},\ldots,a_n}}{dx_i}(a_1,\ldots,a_n) = \frac{\partial f}{\partial x_i}(a_1,\ldots,a_n)$$

这个表达式说明了偏导数的计算可以化为一元导数的计算。

多变量函数的一个重要的例子，是欧几里德空间$\mathbf{R}^n$（例如$\mathbf{R}^2$或$\mathbf{R}^3$）上的标量值函数$f(x_1,...x_n)$。在这种情况下，f关于每一个变量x_j具有偏导数$\partial f/\partial x_j$。在点a，这些偏导数定义了一个向量：

$$\nabla f(a)=\left(\frac{\partial f}{\partial x_1}(a),\ldots,\frac{\partial f}{\partial x_n}(a)\right)$$

这个向量称为f在点a的**梯度**。如果f在定义域中的每一个点都是可微的，那么梯度便是一个向量值函数∇f，它把点a映射到向量$\nabla f(a)$。这样，梯度便决定了一个向量场。

一个常见的符号滥用是在欧几里得空间$\mathbf{R}^3$中用单位向量 $\hat{\mathbf{i}},\hat{\mathbf{j}},\hat{\mathbf{k}}$来定义Nabla算子 ($\nabla$) 如下:

$$\nabla=\left[\frac{\partial}{\partial x}\right]\hat{\mathbf{i}}+\left[\frac{\partial}{\partial y}\right]\hat{\mathbf{j}}+\left[\frac{\partial}{\partial z}\right]\hat{\mathbf{k}}$$

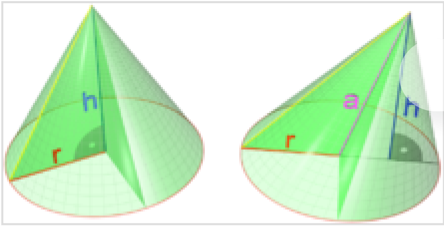
或者，更一般地，对于n维欧几里得空间$\mathbf{R}^n$ 的坐标$(x_1, x_2, x_3,\ldots,x_n)$和单位向量$(\hat{\mathbf{e}}_1,\hat{\mathbf{e}}_2,\hat{\mathbf{e}}_3,\ldots,\hat{\mathbf{e}}_n)$:

$$\nabla=\sum_{j=1}^{n}\left[\frac{\partial}{\partial x_j}\right]\hat{\mathbf{e}}_j=\left[\frac{\partial}{\partial x_1}\right]\hat{\mathbf{e}}_1+\left[\frac{\partial}{\partial x_2}\right]\hat{\mathbf{e}}_2+\left[\frac{\partial}{\partial x_3}\right]\hat{\mathbf{e}}_3+\cdots+\left[\frac{\partial}{\partial x_n}\right]\hat{\mathbf{e}}_n$$

例子

考虑一个圆锥的体积V；它与高度h和半径r有以下的关系：

$$V(r,h)=\frac{\pi r^2 h}{3}。$$



圆锥的体积与它的高度和半径有关

V关于r的偏导数为：

$$\frac{\partial V}{\partial r}=\frac{2\pi rh}{3}，$$ 它描述了高度固定而半径变化时，圆锥的体积的变化率。

V关于h的偏导数为：

$$\frac{\partial V}{\partial h}=\frac{\pi r^2}{3}，$$ 它描述了半径固定而高度变化时，圆锥的体积的变化率。

现在考虑V关于r和h的全导数。它们分别是：

$$\frac{\mathrm{d}V}{\mathrm{d}r}=\overbrace{\frac{2\pi rh}{3}}^{\frac{\partial V}{\partial r}}+\overbrace{\frac{\pi r^2}{3}}^{\frac{\partial V}{\partial h}}\frac{\partial h}{\partial r}$$

以及

$$\frac{\mathrm{d}V}{\mathrm{d}h}=\overbrace{\frac{\pi r^2}{3}}^{\frac{\partial V}{\partial h}}+\overbrace{\frac{2\pi rh}{3}}^{\frac{\partial V}{\partial r}}\frac{\partial r}{\partial h}$$

现在假设，由于某些原因，高度和半径的比k需要是固定的：

$$k=\frac{h}{r}=\frac{\partial h}{\partial r}$$

这便给出了关于r的全导数：

$$\frac{\mathrm{d}\,V}{\mathrm{d}\,r} = \frac{2\pi rh}{3} + k\frac{\pi r^2}{3}$$

可以化简为：

$$\frac{\mathrm{d}\,V}{\mathrm{d}\,r} = k\pi r^2$$

类似地，关于h的全导数是：

$$\frac{\mathrm{d}\,V}{\mathrm{d}\,h} = \pi r^2$$

含有未知函数的偏导数的方程，称为偏导数方程，它在物理学、工程学，以及其它应用科学中经常会见到。

与关于r和h二者相关的全导数是由雅可比矩阵给出的，它的形式为梯度向量$\nabla V = (\frac{\partial V}{\partial r}, \frac{\partial V}{\partial h}) = (\frac{2}{3}\pi rh, \frac{1}{3}\pi r^2)$。

记法

在以下的例子中，设f为x、y和z的函数。

f的一阶偏导数为：

$$\frac{\partial f}{\partial x} = f_x = \partial_x f$$

二阶偏导数为：

$$\frac{\partial^2 f}{\partial x^2} = f_{xx} = \partial_{xx} f$$

二阶混合偏导数为：

$$\frac{\partial^2 f}{\partial y\,\partial x} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) = f_{xy} = \partial_{yx} f$$

高阶偏导数为：

$$\frac{\partial^{i+j+k} f}{\partial x^i\,\partial y^j\,\partial z^k} = f^{(i,j,k)}$$

当处理多变量函数时，有些变量可能互相有关，这样就需要明确指定哪些变量是固定的。在诸如统计力学的领域中，f关于x的偏导数，把y和z视为常数，通常记为：

$$\left(\frac{\partial f}{\partial x}\right)_{y,z}$$

正式定义和性质

像导数一样，偏导数也是定义为一个极限。设U为$\mathbf{R}^n$的一个开子集，$f: U \to \mathbf{R}$是一个函数。我们定义f在点$\boldsymbol{a} = (a_1, \ldots, a_n) \in U$关于第$i$个变量$x_i$的偏导数为：

$$\frac{\partial}{\partial x_i} f(\mathbf{a}) = \lim_{h \to 0} \frac{f(a_1, \ldots, a_{i-1}, a_i + h, a_{i+1}, \ldots, a_n) - f(a_1, \ldots, a_n)}{h}$$

即使在某个给定的点a，所有的偏导数$\partial f/\partial x_i(a)$都存在，函数仍然不一定在该点连续。然而，如果所有的偏导数在a的一个邻域内存在并连续，那么f在该邻域内完全可微分，且全导数是连续的。在这种情况下，我们称f是一个C^1函数。

偏导数$\frac{\partial f}{\partial x}$可以视为定义在$U$内的另外一个函数，并可以再次求偏导数。如果所有的混合二阶偏导数在某个点（或集合）连续，我们便称f为在该点（或集合）的一个C^2函数；在这种情况下，根据克莱罗定理，偏导数可以互相交换：

$$\frac{\partial^2 f}{\partial x_i \, \partial x_j} = \frac{\partial^2 f}{\partial x_j \, \partial x_i}。$$

参考文献

- George B. Thomas & Ross L. Finney. Calculus and Analytic Geometry. Addison-Wesley Publishing Company, Inc. 1994: 833–840. ISBN 0-201-52929-7.

注释

1. 相对于全导数，在其中所有变量都允许变化

参见

- 达朗贝尔算子
- 复合函数求导法则
- 旋度
- 方向导数
- 散度
- 外导数
- 梯度
- 雅可比矩阵
- 拉普拉斯算子
- 二阶导数的对称性
- 三乘积法则，又称为循环链式法则。

-