多元函数（及向量函数）的泰勒展开

Dezeming Family

2021 年 6 月 9 日

DezemingFamily 系列书和小册子因为是电子书，所以可以很方便地进行修改和重新发布。如果您获得了 DezemingFamily 的系列书，可以从我们的网站 [https://dezeming.top/] 找到最新版。对书的内容建议和出现的错误欢迎在网站留言。

目录

一 单变量函数泰勒级数

我们知道，泰勒级数是函数的一种近似工具。学习多元函数泰勒级数的计算和表示应该先弄懂一元函数的泰勒级数原理。

假设我们要在 $x = 0$ 的附近来近似一个函数 $f(x)$，我们拟使用多项式来拟合，比如二次多项式：

$$f(x)_{(x\ near\ 0)} \approx a_0 + a_1 x + a_2 x^2 \tag{一.1}$$

然后代入 $f(0)$ 到右边，得到 $a_0 = f(0)$。我们还希望近似的多项式的一阶导数和原函数的一阶导数相同，因此得到：

$$f'(0) = a_1 \tag{一.2}$$

我们再贪心一点，希望近似的多项式的二阶导数和原函数二阶导数相同，因此得到：

$$f''(0) = 1 \cdot 2 \cdot a_2 \tag{一.3}$$

如果我们想用高阶多项式来逼近，则我们可以得到：

$$f^{(n)}(0) = n! a_n \tag{一.4}$$

所以我们可以理解为，a_n 是来控制逼近多项式去逼近原函数的 n 阶导数的系数。

当我们要逼近其他点，比如逼近：

$$f(x)_{(x\ near\ x_0)} \tag{一.5}$$

我们当然也可以使用上面的逼近方法：

$$f(x)_{(x\ near\ x_0)} \approx a_0 + a_1 x + a_2 x^2 + ... + a_n x^n \tag{一.6}$$

$$f'(x_0) = a_1 + 2a_2 x_0 + 3a_3 x_0^2 + ... + n a_n x_0^{n-1} \tag{一.7}$$

$$\tag{一.8}$$

我们发现这样计算起来真的非常麻烦，假如我们要使用 n 阶多项式去逼近，则需要联立所有阶导数的方程组去解得 a_i，非常不方便。幸运的是，我们将多项式的每一项 x 换为 $x - x_0$，这个问题就解决了：

$$f(x)_{(x\ near\ x_0)} \approx a_0 + a_1(x - x_0) + a_2(x - x_0)^2 + ... + a_n(x - x_0)^n \tag{一.9}$$

$$f'(x_0) = a_1 \tag{一.10}$$

$$f''(x_0) = 1 \cdot 2 \cdot a_2 \tag{一.11}$$

其实这就相当于将原函数进行平移，使 x_0 处平移到原点位置，然后再用泰勒公式去逼近。

二　二元函数泰勒公式

我们现在有一个二元函数 $f(x_1, x_2)$，假设我们要近似点 (a_1, a_2) 的函数，其实原理也很简单，无非就是求个偏导罢了。我们知道，要想得到二元函数在某个点的变化趋势，需要求其分别对两个变量的偏导数。

我们先从直觉上思考一下

二阶近似就可以写为（设常数为 c_i）：

$$f(x_1, x_2)_{((x_1,x_2) near\ (a_1,a_2))} \tag{二.1}$$

$$\approx c_1 + c_{2,1}(x_1 - a_1) + c_{2,2}(x_2 - a_2) \tag{二.2}$$

$$+ c_{3,1}(x_1 - a_1)^2 + c_{3,2}(x_1 - a_1)(x_2 - a_2) + c_{3,3}(x_2 - a_2)^2 \tag{二.3}$$

以及泰勒展开式：

$$f(x_1, x_2)_{((x_1,x_2) near\ (a_1,a_2))} \tag{二.4}$$

$$\approx f(a_1, a_2) \tag{二.5}$$

$$+ (x_1 - a_1)\frac{\partial f}{\partial x_1}(a_1, a_2) + (x_2 - a_2)\frac{\partial f}{\partial x_2}(a_1, a_2) \tag{二.6}$$

$$+ \frac{1}{2!}(x_1 - a_1)^2 \frac{\partial^2 f}{\partial x_1^2}(a_1, a_2) + (x_1 - a_1)(x_2 - a_2)\frac{\partial^2 f}{\partial x_1 \partial x_2}(a_1, a_2) \tag{二.7}$$

$$+ \frac{1}{2!}(x_2 - a_2)^2 \frac{\partial^2 f}{\partial x_2^2}(a_1, a_2) \tag{二.8}$$

$$+ o^n \tag{二.9}$$

系数是通过下面来计算得到的（注意 $c_{3,2}$ 可以用两种混合偏导的方法得到）：

$$c_1 = f(a_1, a_2) \tag{二.10}$$

$$\frac{\partial f}{\partial x_1}(a_1, a_2) = c_{2,1} \tag{二.11}$$

$$\frac{\partial^2 f}{\partial x_1 \partial x_2}(a_1, a_2) = c_{3,2} \tag{二.12}$$

$$\frac{\partial^2 f}{\partial x_2 \partial x_1}(a_1, a_2) = c_{3,2} \tag{二.13}$$

我们知道，初等函数的二阶偏导连续，则与求偏导的先后次序无关，所以二阶混合偏导相等，因此二阶项可以拆成：

$$\frac{1}{2!}(x_1-a_1)^2\frac{\partial^2 f}{\partial x_1^2}(a_1,a_2)+\frac{1}{2!}(x_1-a_1)(x_2-a_2)\frac{\partial^2 f}{\partial x_1\partial x_2}(a_1,a_2) \tag{二.14}$$

$$+\frac{1}{2!}(x_2-a_2)(x_1-a_1)\frac{\partial^2 f}{\partial x_2\partial x_1}(a_1,a_2)+\frac{1}{2!}(x_2-a_2)^2\frac{\partial^2 f}{\partial x_2^2}(a_1,a_2) \tag{二.15}$$

这就是一般常见的二元函数泰勒公式。但是毕竟把 $c_{3,2}$ 拆成两部分并不是很容易想到，我们用辅助函数的方法来证明一下。

辅助函数推导

我们引入一个辅助函数 Φ 来解决该问题。定义点 (a_1,a_2) 在周边邻域的近似，泰勒定理需要研究的是在 (a_1,a_2) 周围邻域上的偏移的函数值近似 (a_1+u,a_2+v)。我们设 $x_1=a_1+tu$，$x_2=a_2+tv$。（我在最后一章给出了构造辅助函数的原理性说明。）

$$\Phi(t)=f(a_1+tu,a_2+tv)\quad(0\le t\le 1) \tag{二.16}$$

把 Φ 在 $t=0$ 展开（其实就是在 (a_1,a_2) 点展开），去近似 $t=1$ 上的值：

$$\Phi(t)=\Phi(0)+\frac{\Phi'(0)}{1!}t+\frac{\Phi''(0)}{2!}t^2+...+\frac{\Phi^n(0)}{n!}t^n+\frac{\Phi^{n+1}(\theta t)}{(n+1)!}t^{n+1}\quad(0<\theta<1) \tag{二.17}$$

$$\Phi(1)=\Phi(0)+\frac{\Phi'(0)}{1!}+\frac{\Phi''(0)}{2!}+...+\frac{\Phi^n(0)}{n!}+\frac{\Phi^{n+1}(\theta)}{(n+1)!} \tag{二.18}$$

根据链式法则：

$$\frac{df}{dt}=\frac{\partial f}{\partial x_1}\frac{dx_1}{dt}+\frac{\partial f}{\partial x_2}\frac{dx_2}{dt} \tag{二.19}$$

$$\phi'(t)=\frac{df(a_1+tu,a_2+tv)}{dt}=u\frac{\partial f}{\partial x_1}+v\frac{\partial f}{\partial x_2} \tag{二.20}$$

$$\phi'(0)=u\frac{\partial f}{\partial x_1}(a_1,a_2)+v\frac{\partial f}{\partial x_2}(a_1,a_2) \tag{二.21}$$

$$\phi''(t)=\frac{d\left(u\frac{\partial f}{\partial x_1}+v\frac{\partial f}{\partial x_2}\right)}{dt} \tag{二.22}$$

$$=u^2\frac{\partial^2 f}{\partial x_1^2}+uv\frac{\partial^2 f}{\partial x_1\partial x_2}+uv\frac{\partial^2 f}{\partial x_2\partial x_1}+v^2\frac{\partial^2 f}{\partial x_2^2} \tag{二.23}$$

$$\phi''(0)=... \tag{二.24}$$

我们令 $t=1$，$u=x_1-a_1$，$v=x_2-a_2$，就能得到前面的二元函数泰勒公式形式：

$$\Phi(1)=\Phi(t)=f(x_1,x_2)_{((x_1,x_2)near\ (a_1,a_2))} \tag{二.25}$$

$$=f(a_1,a_2) \tag{二.26}$$

$$+\left((x_1-a_1)\frac{\partial f}{\partial x_1}(a_1,a_2)+(x_2-a_2)\frac{\partial f}{\partial x_2}(a_1,a_2)\right) \tag{二.27}$$

$$+\frac{1}{2!}\left((x_1-a_1)^2\frac{\partial^2 f}{\partial x_1^2}+(x_1-a_1)(x_2-a_2)\frac{\partial^2 f}{\partial x_1\partial x_2}\right. \tag{二.28}$$

$$\left.+(x_1-a_1)(x_2-a_2)\frac{\partial^2 f}{\partial x_2\partial x_1}+(x_2-a_2)^2\frac{\partial^2 f}{\partial x_2^2}\right)+... \tag{二.29}$$

我们还可以写为矩阵向量形式，设:

$$\boldsymbol{x}=\begin{bmatrix}x_1\\x_2\end{bmatrix}\qquad \boldsymbol{a}=\begin{bmatrix}a_1\\a_2\end{bmatrix} \tag{二.30}$$

则矩阵向量形式为:

$$f(\boldsymbol{x})\approx f(\boldsymbol{a})+\nabla f(\boldsymbol{a})\cdot(\boldsymbol{x}-\boldsymbol{a})+\frac{1}{2}(\boldsymbol{x}-\boldsymbol{a})^T H(\boldsymbol{a})(\boldsymbol{x}-\boldsymbol{a}) \tag{二.31}$$

其中 H 为海森（Hessian）矩阵：

$$H = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} \end{bmatrix} \tag{二.32}$$

三　泰勒公式拟合邻域点函数值

我们已经有了 $f(x_1, x_2)_{((x_1,x_2) near\ (a_1,a_2))}$ 的表示方法，如果我们要求某点 $(a_1 + \Delta a_1, a_2 + \Delta a_2)$ 处的函数值，则我们需要计算：

$$f(a_1 + \Delta a_1, a_2 + \Delta a_2) \tag{三.1}$$
$$= f(a_1, a_2) \tag{三.2}$$
$$+ \left(\Delta a_1 \frac{\partial f}{\partial x_1}(a_1, a_2) + \Delta a_2 \frac{\partial f}{\partial x_2}(a_1, a_2)\right) \tag{三.3}$$
$$+ \frac{1}{2!}\left((\Delta a_1)^2 \frac{\partial f}{\partial x_1^2} + \Delta a_1 \Delta a_2 \frac{\partial f}{\partial x_1 \partial x_2}\right. \tag{三.4}$$
$$\left. + \Delta a_1 \Delta a_2 \frac{\partial f}{\partial x_2 \partial x_1} + (\Delta a_2)^2 \frac{\partial f}{\partial x_2^2}\right) + ... \tag{三.5}$$

其实很多书上的公式都是表示为上面的形式。

四　多元函数泰勒公式

其实掌握了二元函数泰勒公式以后，多元变量就很好扩展了，只是需要一定的方法来方便表示。

我们令：

$$\left(\Delta a_1 \frac{\partial f}{\partial x_1}(a_1, a_2) + \Delta a_2 \frac{\partial f}{\partial x_2}(a_1, a_2)\right)^2 \tag{四.1}$$
$$= \left((\Delta a_1)^2 \frac{\partial f}{\partial x_1^2} + \Delta a_1 \Delta a_2 \frac{\partial f}{\partial x_1 \partial x_2}\right. \tag{四.2}$$
$$\left. + \Delta a_1 \Delta a_2 \frac{\partial f}{\partial x_2 \partial x_1} + (\Delta a_2)^2 \frac{\partial f}{\partial x_2^2}\right) \tag{四.3}$$

因此多元函数泰勒展开就很容易写为：

$$f(x_1, x_2, ..., x_n) = f(a_1 + \Delta a_1, a_2 + \Delta a_2, ..., a_n + \Delta a_n) \tag{四.4}$$
$$= f(a_1, a_2, ..., a_n) \tag{四.5}$$
$$+ \frac{1}{1!}\left(\Delta a_1 \frac{\partial f}{\partial x_1} + \Delta a_2 \frac{\partial f}{\partial x_2} + ... + \Delta a_n \frac{\partial f}{\partial x_n}\right)(a_1, a_2, ..., a_n) \tag{四.6}$$
$$+ \frac{1}{2!}\left(\Delta a_1 \frac{\partial f}{\partial x_1} + \Delta a_2 \frac{\partial f}{\partial x_2} + ... + \Delta a_n \frac{\partial f}{\partial x_n}\right)^2(a_1, a_2, ..., a_n) \tag{四.7}$$
$$+ ... \tag{四.8}$$
$$+ \frac{1}{n!}\left(\Delta a_1 \frac{\partial f}{\partial x_1} + \Delta a_2 \frac{\partial f}{\partial x_2} + ... + \Delta a_n \frac{\partial f}{\partial x_n}\right)^n(a_1, a_2, ..., a_n) \tag{四.9}$$

当然也可以写为累加和的形式：

$$f(x_1, x_2, ..., x_n) = f(a_1 + \Delta a_1, a_2 + \Delta a_2, ..., a_n + \Delta a_n) \tag{四.10}$$
$$= f(a_1, a_2, ..., a_n) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x_i - a_i) + \frac{1}{2!}\sum_{i=1}^{n}\sum_{j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(x_i - a_i)(x_j - a_j) \tag{四.11}$$
$$+ \frac{1}{3!}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n} \frac{\partial^3 f}{\partial x_i \partial x_j \partial x_k}(x_i - a_i)(x_j - a_j)(x_k - a_k) + ... \tag{四.12}$$

以及描述为向量的形式：

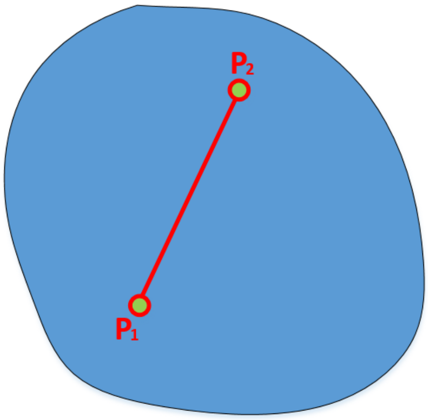
$$f(\boldsymbol{x}) = \sum_{l\geq 0} \frac{(\boldsymbol{x}-\boldsymbol{a})^l}{l!}(\partial^l f)(\boldsymbol{a}) \tag{四.13}$$

最终其实可以描述为矩阵向量的形式，来近似到二阶，很多时候二阶近似就已经完全够用了，海森矩阵描述如下：

$$H = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ . & . & . & . \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{bmatrix} \tag{四.14}$$

二元函数的辅助函数

函数 $f(x_1, x_2)$ 在开区域 R 中有二阶连续偏导，其中 $P_1(a_1, a_2)$ 是该区域的一个点，我们在开区域中任选另一个点 $P_2(a_1+u, a_2+v)$，并且我们设 u 和 v 足够小，来保证从 P_1 沿直线运动到 P_2 的路径仍然在开区域中：



则描述从 P_1 到 P_2 的运动轨迹的参数方程为 $(a_1 + t\cdot u, a_2 + t\cdot v)$，

因此定义参数方程 $F(t) = f(a_1 + t\cdot u, a_2 + t\cdot v)$，我们知道当 u 和 v 变化时，P_2 可以表示开区域附近邻域上的任意一个点。因为 $f(a_1 + t\cdot u, a_2 + t\cdot v)$ 中，$x_1 = a_1 + t\cdot u$，$x_2 = a_2 + t\cdot v$，因此其实对 t 求导就可以应用链式法则。

同理，因为 t 在 $[0, 1]$ 之间是连续的，所以我们可以对 F 在 $t = 0$ 进行泰勒展开（其实就是在 (a_1, a_2) 点展开），令 $t = 1$ 就能得到：

$$F(1) = F(0) + \frac{F'(0)}{1!} + \frac{F''(0)}{2!} + ... \tag{四.15}$$

换句话说，$f(a_1+u, a_2+v)$ 就等于 $F(1)$，也就是说可以用上面的公式来进行近似。因此，只要当 u 和 v 任意取值时，我们就能得到 R 上 $P_1(a_1, a_2)$ 点附近函数值所有的近似值了！

现在，我们重新令变量 $x_1 = u + a_1$，$x_2 = v + a_2$，这样代入到上式，我们就能得到 $f(x_1, x_2)$ 在 (a_1, a_2) 处的泰勒展开。

参考文献

[1] https://blog.csdn.net/weixin_34297863/article/details/114015833

[2] https://zhuanlan.zhihu.com/p/32274749

[3] https://www.bilibili.com/video/BV1Gx411Y7cz

[4] https://blog.csdn.net/red_stone1/article/details/70260070

[5] https://wenku.baidu.com/view/26fc62c689eb172ded63b7dd.html

[6] https://mathinsight.org/

[7] Finney, Giordano. 托马斯微积分. 第 10 版 [M]. 高等教育出版社, 2003.