

泰勒公式

前言

我喜欢对数学公式有一个形象化理解，但是当空间维度超过三维时，数学又抽象的没法理解了，在此推荐一个UP主[3Blue1Brown](#)他讲了不少三维空间下的数学形象化理解。

这个[泰勒公式的形象化介绍](#)讲泰勒公式挺棒的。

正文

首先要明确泰勒公式是对一个存在复杂表达式函数在某一点展开的等效表示（带余项）或近似表示（不带余项），将其转化成简单的形式，即局部近似。

本部分不考虑余项。

一元函数

假设 $f(x)$ 为真实存在的一个复杂函数，为 n 次多项式。在 $x = 0$ 处，我们使用一个 n 次多项式 $g(x)$ 来等效。

（假设 $f(x)$ 最高项为 n 次，这样我们才能使用如下的全等式），则有

$$g(x) = f(x) , \tag{1}$$

$$g(x) = a_0 + a_1x^1 + a_2x^2 + \dots + a_nx^n , \tag{2}$$

此时 $g(x)$ 中的 $\{a_0, a_1, \dots, a_n\}$ 等参数未知。因为两函数在 $x = 0$ 处相等，则有

$$g(0) = a_0 = f(0) , \quad (3)$$

此时 a_0 已知，但其他参数依然未知需要求解。对公式（1）、（2）求导则有

$$g'(x) = f'(x) , \quad (4)$$

$$g'(x) = a_1 + 2a_2x^1 + \dots + na_nx^{n-1} 。 \quad (5)$$

因为是在 $x = 0$ 处展开，所以有

$$g'(0) = a_1 = f'(0) , \quad (6)$$

此时 $\{a_0, a_1\}$ 已知，但其他参数依然未知需要求解。对公式（4）、（5）求导则有(需要注意的是到了高阶的求导，因此就有了阶乘项)

$$g''(x) = f''(x) , \quad (7)$$

$$g''(x) = 2!a_2 + \dots + n * (n - 1) * a_nx^{n-2} 。 \quad (8)$$

因为是在 $x = 0$ 处展开，所以有

$$g''(0) = 2!a_2 = f''(0) , \quad (9)$$

$$a_2 = \frac{f''(0)}{2!} 。 \quad (10)$$

其他阶基本同理，当进行 n 阶求导后可得

$$a_n = \frac{f^n(0)}{n!}。 \quad (11)$$

自此， $\{a_0, a_1, \dots, a_n\}$ 已全部求解，则 $f(x)$ 在 $x = 0$ 处的泰勒展开式为

$$g(x) = f(0) + f'(0)x + \frac{f''(0)}{2!}x^2 + \dots + \frac{f^n(0)}{n!}x^n。 \quad (12)$$

当在 x_0 处展开时，公式（2）变成

$$g(x) = a_0 + a_1(x - x_0)^1 + a_2(x - x_0)^2 + \dots + a_n(x - x_0)^n。 \quad (13)$$

基于公式（13）， $f(x)$ 的最终展开式为

$$g(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2!}(x - x_0)^2 + \dots + \frac{f^n(x_0)}{n!}(x - x_0)^n。 \quad (14)$$

高维函数

二元函数

如果我们按照一元函数的近似思路，如公式（2）所示，在该部分则变成

$$g(x, y) = a_{01} + a_{11}x + a_{12}y + a_{21}x^2 + a_{22}xy + a_{23}y^2 + \dots。 \quad (15)$$

按照近似点的值相等、导数相等、二阶导相等、...、 n 阶导相等的思路上述对应的 $g(x, y)$ 是可以求解出来的，但是一元函数通过 n 阶方程近似是很容易理解的，但公式（15）为什么是这种形式就不好说清楚，更不要说形象化理解了。这里将使用该[陕师大阳光青椒-陈森](#)的思路进行二元函数泰勒公式的推导。

因为泰勒展开式是局部近似，因此令 $f(x, y)$ 在 (x_0, y_0) 展开式为

$$g(x_0 + h, y_0 + k) = f(x_0 + h, y_0 + k) \text{。} \quad (16)$$

重新定义一个函数为

$$F(t) = f(x_0 + ht, y_0 + kt) \text{ ,} \quad (17)$$

基于该函数我们可得

$$g(x_0 + h, y_0 + k) = F(1) \text{ ,} \quad (18)$$

因求解 $g(x)$ 就变成了求解 $F(1)$ 。由公式 (17) 的定义可知 $F(t)$ 为一元函数, 则 $F(t)$ 在 $(0, 0)$ 处的泰勒展开为

$$F(t) = F(0) + F'(0)t + \frac{F''(0)}{2!}t^2 + \cdots \text{。} \quad (19)$$

对上述未知参数依次求解: 首先易知

$$F(0) = f(x_0, y_0) \text{。} \quad (20)$$

其次, 对 $F(t)$ 求一阶导为

$$F'(t) = \frac{\partial f(x_0 + ht, y_0 + kt)}{\partial t} \text{ ,} \quad (21)$$

令 $p = x_0 + ht, q = y_0 + kt$,则按照复合函数求导的规则上式则变为

$$F'(t) = \frac{\partial f(p, q)}{\partial p} \frac{\partial p}{\partial t} + \frac{\partial f(p, q)}{\partial q} \frac{\partial q}{\partial t} = \frac{\partial f(p, q)}{\partial p} h + \frac{\partial f(p, q)}{\partial q} k \text{。} \quad (22)$$

在一阶导的基础上求解二阶导为

$$F''(t) = \frac{\partial^2 f(p, q)}{\partial p^2} h^2 + \frac{\partial^2 f(p, q)}{\partial p \partial q} hk + \frac{\partial^2 f(p, q)}{\partial q \partial p} kh + \frac{\partial^2 f(p, q)}{\partial q^2} k^2。 \quad (23)$$

高阶导数求解思路基本一样。基于公式 (22) 和 (23) 可知

$$F'(0) = \frac{\partial f(x_0, y_0)}{\partial p} h + \frac{\partial f(x_0, y_0)}{\partial q} k, \quad (24)$$

$$F''(0) = \frac{\partial^2 f(x_0, y_0)}{\partial p^2} h^2 + \frac{\partial^2 f(x_0, y_0)}{\partial p \partial q} hk + \frac{\partial^2 f(x_0, y_0)}{\partial q \partial p} kh + \frac{\partial^2 f(x_0, y_0)}{\partial q^2} k^2。 \quad (25)$$

所以,

$$g(p, q) = F(1) = f(x_0, y_0) + \frac{\partial f(x_0, y_0)}{\partial p} h + \frac{\partial f(x_0, y_0)}{\partial q} k + \frac{\partial^2 f(x_0, y_0)}{\partial p^2} h^2 + \quad (26-1)$$

$$\frac{\partial^2 f(x_0, y_0)}{\partial p \partial q} hk + \frac{\partial^2 f(x_0, y_0)}{\partial q \partial p} kh + \frac{\partial^2 f(x_0, y_0)}{\partial q^2} k^2 + \dots \quad (26-2)$$

上式即为二元泰勒展开。使用 $\{x, y\}$ 替换 $\{p, q\}$ 则有 $\{h = x - x_0, k = y - y_0\}$, 代入上式可得常见的 $\{x, y\}$ 形式

$$g(x, y) = F(1) = f(x_0, y_0) + \frac{\partial f(x_0, y_0)}{\partial x} (x - x_0) + \frac{\partial f(x_0, y_0)}{\partial y} (y - y_0) + \frac{\partial^2 f(x_0, y_0)}{\partial x^2} (x - x_0)^2 \quad (27-1)$$

$$+ \frac{\partial^2 f(x_0, y_0)}{\partial x \partial y} (x - x_0)(y - y_0) + \frac{\partial^2 f(x_0, y_0)}{\partial y \partial x} (y - y_0)(x - x_0) + \frac{\partial^2 f(x_0, y_0)}{\partial y^2} (y - y_0)^2 + \dots \quad (27-2)$$

当使用矩阵的形式进行描述时，则有

$$g(x, y) = f(x_0, y_0) + \begin{bmatrix} \frac{\partial f(x_0, y_0)}{\partial x} & 0 \\ 0 & \frac{\partial f(x_0, y_0)}{\partial y} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \frac{1}{2!} \begin{bmatrix} x \\ y \end{bmatrix}^T \begin{bmatrix} \frac{\partial^2 f(x_0, y_0)}{\partial x^2} & \frac{\partial^2 f(x_0, y_0)}{\partial x \partial y} \\ \frac{\partial^2 f(x_0, y_0)}{\partial y \partial x} & \frac{\partial^2 f(x_0, y_0)}{\partial y^2} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \dots \quad (28)$$

其中在控制中一阶导的矩阵形式成为雅克比矩阵 J ，二级导的矩阵形式为海森矩阵 H ，至于三元及以上函数的推导基本同上，不再赘述。

PS

1.为什么是在一个已知的点展开?

从上述完整的推导过程可以看出（至少按照这个思路来说），需要一个点来提供确定的信息进行未知参数的求解。曾看到过这样一个说法：泰勒级数的局部性质比较好，但是整体确很难收敛，这源于它是以多项式函数为基的，而傅里叶级数却表现了很好的整体性，也是因为它以三角函数为基的。

2.关于余项

上面的内容是没有考虑余项的，因此严格意义上数学推导是不严谨。

3.其他

这个vscode里mk真的是拉胯啊，长公式没法换行，latex一些语法只能预览有效，输出成PDF乱码（pdf里公式28是有问题，不知道该怎么解决）。markdown源码和pdf请参见[tusha-github](https://github.com/tusha)。