11.4 多元函数的泰勒公式

一元函数的泰勒公式

带佩亚诺余项的泰勒公式

设函数$f(t)$在t_0处m阶导数存在, 则

$$f(t_0+h)=f(t_0)+f'(t_0)h+\cdots+\frac{1}{m!}f^{(m)}(t_0)h^m+R_m(h),$$

其中$R_m(h)$是f在t_0点的m阶余项. 对固定的t_0, 有

$$R_m(h)=o(|h|^m)\ (h\to 0),$$

如此表示的$R_m(h)$称之为佩亚诺余项.

带拉格朗日余项的泰勒公式

进一步, 如果$f(t)$在t_0的邻域$m+1$ 阶导数存在, 则余项可表示为

$$R_m(h)=\frac{h^{m+1}}{(m+1)!}f^{(m+1)}(t_0+\theta h),$$

其中$0<\theta<1$, 如此表示的$R_m(h)$称为拉格朗日余项.

一元函数的泰勒公式

带柯西余项的泰勒公式

如果$f(t)$在t_0的邻域$m+1$ 阶导数存在, 则

$$f(t_0+h)=f(t_0)+f'(t_0)h+\cdots+\frac{1}{m!}f^{(m)}(t_0)h^m+R_m(h),$$

余项可表示为

$$R_m(h)=\frac{h^{m+1}}{m!}f^{(m+1)}(t_0+\theta h)(1-\theta)^m,$$

其中$0<\theta<1$, 此时称为柯西余项.

带积分余项的泰勒公式

如果$m+1$阶导数在t_0的邻域内连续, 则余项可写为积分形式, 称为积分形式余项, 即

$$R_m(h)=\frac{h^{m+1}}{m!}\int_0^1 f^{(m+1)}(t_0+sh)(1-s)^m ds.$$

一元函数的泰勒公式推广到多元函数情形

一元函数的泰勒公式容易推广到多元函数情形. 设$D \subseteq \mathbb{R}^n$是一个开区域，$X_0 = (a_1, a_2, \cdots, a_n) \in D$. 取独立于$X_0$的变量$\Delta X = (\Delta x_1, \cdots, \Delta x_n)$, 使线段$X = X_0 + t\Delta X$, $0 \leqslant t \leqslant 1$全含在区域$D$内. 考虑一元函数$\varphi(t) = f(a_1 + t\Delta x_1, \cdots, a_n + t\Delta x_n)$.

如果函数$f(X)$在点X_0处m阶可微，则由本章第二节定理3知$\varphi(t)$在$t = 0$处m次可微. 不难想到可以由

$$f(X_0 + \Delta X) - f(X_0) = \varphi(1) - \varphi(0) = \sum_{k=1}^{m} \frac{\varphi^{(k)}(0)}{k!} + R_m(\Delta X)$$

来得出$f(X)$在X_0处的泰勒公式，这里$R_m(\Delta X)$是$f(X)$在X_0处泰勒公式的m阶余项.

$\varphi(t)$的各阶导数

$$\varphi'(t) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(X_0 + t\Delta X)\Delta x_i,$$

$$\varphi''(t) = \sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(X_0 + t\Delta X)\Delta x_i \Delta x_j,$$

一般地有，若$f(X)$在开区域D内m次可微，则对$t \in [0,1]$, $k = 1, \cdots, m$, 有

$$\varphi^{(k)}(t) = \sum_{i_1,\cdots,i_k=1}^{n} \frac{\partial^k f}{\partial x_{i_1} \cdots \partial x_{i_k}}(X_0 + t\Delta X)\Delta x_{i_1} \cdots \Delta x_{i_k}.$$

多元函数的泰勒公式

设函数$f(X)$在点X_0处m阶可微，则$\varphi(t)$ 在$t=0$处m次可微. 由前面的计算不难看到:

$$\varphi'(0)=\sum_{i=1}^{n}\frac{\partial f}{\partial x_i}(X_0)\Delta x_i=\mathrm{d}f(X_0)|_{\mathrm{d}x_i=\Delta x_i},$$

$$\varphi''(0)=\sum_{i,j=1}^{n}\frac{\partial^2 f}{\partial x_i\partial x_j}(X_0)\Delta x_i\Delta x_j=\mathrm{d}^2 f(X_0)\Big|_{\mathrm{d}x_i=\Delta x_i},$$

$$\cdots,$$

$$\varphi^{(m)}(0)=\sum_{i_1,\cdots,i_m=1}^{n}\frac{\partial^m f}{\partial x_{i_1}\cdots\partial x_{i_m}}(X_0)\Delta x_{i_1}\cdots\Delta x_{i_m}$$
$$=\mathrm{d}^m f(X_0)|_{\mathrm{d}x_i=\Delta x_i}.$$

于是可以把多元函数的泰勒公式写成下面形式

$$f(X_0+\Delta X)=f(X_0)+\sum_{k=1}^{m}\frac{1}{k!}\mathrm{d}^k f(X_0)\Big|_{\mathrm{d}x_i=\Delta x_i}+R_m(\Delta X).$$

带佩亚诺余项的泰勒公式与带拉格朗日余项的泰勒公式

定理 1 (局部泰勒公式)

设$D \subseteq \mathbb{R}^n$是一个开区域，$X_0 = (a_1, a_2, \cdots, a_n) \in D$，函数$f(X)$在点$X_0$处$m$次可微. 则有

$$f(X_0+\Delta X) = f(X_0) + \sum_{k=1}^{m} \frac{1}{k!} \mathrm{d}^k f(X_0)\Big|_{\mathrm{d}x_i=\Delta x_i} + o(|\Delta X|^m), |\Delta X| \to 0.$$

定理 2 (整体泰勒公式)

设$X_0 = (a_1, a_2, \cdots, a_n) \in \mathbb{R}^n$，函数$f(X)$在点$X_0$的$r$邻域$B_r(X_0) = \{X \in \mathbb{R}^n \mid |X - X_0| < r\}$内$m+1$次可微. 则对于任何$X \in B_r(X_0)$, 存在点$C = X_0 + \theta(X - X_0)$, $0 < \theta < 1$, 使得

$$\begin{aligned} & f(X_0+\Delta X) \\ = \ & f(X_0) + \sum_{k=1}^{m} \frac{1}{k!} \mathrm{d}^k f(X_0)\Big|_{\mathrm{d}x_i=\Delta x_i} + \frac{1}{(m+1)!} \mathrm{d}^{m+1} f(C)\Big|_{\mathrm{d}x_i=\Delta x_i}. \end{aligned}$$

局部泰勒公式的一个有问题的证明

请指出下面证明的问题所在

$$
\begin{aligned}
& f(a_1+t\Delta x_1,\cdots,a_n+t\Delta x_n)-f(a_1,\cdots,a_n) \\
=& \sum_{k=1}^{m}\frac{t^k}{k!}\sum_{i_1,\cdots,i_k=1}^{n}\frac{\partial^k f}{\partial x_{i_1}\cdots\partial x_{i_k}}(X_0)\Delta x_{i_1}\cdots\Delta x_{i_k}+R_m(t\Delta X).
\end{aligned}
$$

当ΔX固定时，$R_m(t\Delta X)=o(|t|^m), t\to 0.$ 令$h_i=t\Delta x_i$，则$\sqrt{h_1^2+\cdots+h_n^2}=|t|\sqrt{\Delta x_1^2+\cdots\Delta x_n^2}$，从而如果$\sqrt{\Delta x_1^2+\cdots\Delta x_n^2}\neq 0$，则有$o(|t|^m)=o\left(\left(\sqrt{h_1^2+\cdots+h_n^2}\right)^m\right)$. 由此可得具有佩亚诺型余项的泰勒公式

$$
\begin{aligned}
& f(a_1+h_1,\cdots,a_n+h_n)-f(a_1,\cdots,a_n) \\
=& \sum_{k=1}^{m}\frac{1}{k!}\sum_{i_1,\cdots,i_k=1}^{n}\frac{\partial^k f}{\partial x_{i_1}\cdots\partial x_{i_k}}(X_0)h_{i_1}\cdots h_{i_k} \\
& +o\left(\left(\sqrt{h_1^2+\cdots+h_n^2}\right)^m\right).
\end{aligned}
$$

局部泰勒公式的证明

对m用数学归纳法. 对$m=1$，由函数的微分定义，知定理的结论成立. 现设$m>1$, 并且定理的结论对$m-1$时已成立.

为方便起见, 记$\Delta X=U=(u_1,u_2,\cdots,u_n)$, 并记

$$\begin{aligned} P_m(U) &= f(X_0)+\sum_{k=1}^{m}\frac{1}{k!}\mathrm{d}^k f(X_0)\Big|_{\mathrm{d}x_i=u_i} \\ &= f(X_0)+\sum_{k=1}^{m}\frac{1}{k!}\sum_{i_1,\cdots,i_k=1}^{n}\frac{\partial^k f}{\partial x_{i_1}\cdots\partial x_{i_k}}(X_0)u_{i_1}\cdots u_{i_k}. \end{aligned}$$

从而$R_m(U)=f(X_0+U)-P_m(U)$.

把原点记为$O=(0,0,\cdots,0)$, 显然$R_m(O)=0$. 由条件, $R_m(U)$在原点O的邻域内有直到$m-1$阶的一切偏导数, 在原点O处m阶可微, 并且直到m阶所有偏导数全为零.

局部泰勒公式的证明（续）

我们写

$$
\begin{aligned}
& R_m(U) \\
= \quad & R_m(U) - R_m(O) \\
= \quad & \sum_{i=1}^{n} \left[R_m(u_1, \cdots, u_i, 0, \cdots, 0) - R_m(u_1, \cdots, u_{i-1}, 0, \cdots, 0)\right].
\end{aligned}
$$

由一元函数的拉格朗日中值定理, 存在$\theta_i \in (0, 1)$, 对于任意$i = 1, \cdots, n$, 有

$$
\begin{aligned}
& R_m(u_1, \cdots, u_i, 0, \cdots, 0) - R_m(u_1, \cdots, u_{i-1}, 0, \cdots, 0) \\
= \quad & \frac{\partial R_m}{\partial u_i}(u_1, u_2, \cdots, \theta_i u_i, 0, \cdots, 0) u_i.
\end{aligned}
$$

局部泰勒公式的证明（再续）

于是

$$R_m(U)=\sum_{i=1}^{n}\frac{\partial R_m}{\partial u_i}(u_1,u_2,\cdots,\theta_i u_i,0,\cdots,0)u_i.$$

而$\frac{\partial R_m}{\partial u_i}(u_1,\cdots,u_n)$在原点处$m-1$次可微，从而由归纳假设，并注意到所有偏导数满足$\frac{\partial^k R_m}{\partial u_{i_1}\cdots\partial u_{i_k}}(O)=0$，从而有

$$\begin{aligned}&\frac{\partial R_m}{\partial u_i}(u_1,u_2,\cdots,\theta_i u_i,0,\cdots,0)\\ =\;& o(|(u_1,u_2,\cdots,\theta_i u_i,0,\cdots,0)|^{m-1})\\ =\;& o(|U|^{m-1}).\end{aligned}$$

因此$R_m(U)=\sum_{i=1}^{n}o(|U|^{m-1})u_i=o(|U|^m)$.

整体泰勒公式的证明

由前面的计算，对一元函数$\varphi(t)$用泰勒公式，从而有

$$
\begin{aligned}
& f(a_1+t\Delta x_1,\cdots,a_n+t\Delta x_n)-f(a_1,\cdots,a_n) \\
&= \sum_{k=1}^{m}\frac{t^k}{k!}\sum_{i_1,\cdots,i_k=1}^{n}\frac{\partial^k f}{\partial x_{i_1}\cdots\partial x_{i_k}}(X_0)\Delta x_{i_1}\cdots\Delta x_{i_k}+R_m(t\Delta X).
\end{aligned}
$$

若取$t=1$, 并且$f(X)$在D内$m+1$次可微时,

$$
\begin{aligned}
& f(a_1+\Delta x_1,\cdots,a_n+\Delta x_n)-f(a_1,\cdots,a_n) \\
&= \sum_{k=1}^{m}\frac{1}{k!}\sum_{i_1,\cdots,i_k=1}^{n}\frac{\partial^k f}{\partial x_{i_1}\cdots\partial x_{i_k}}(X_0)\Delta x_{i_1}\cdots\Delta x_{i_k}+R_m(\Delta X),
\end{aligned}
$$

对于$\varphi(t)$取拉格朗日余项, 可知存在$0<\theta<1$，使

$$
\begin{aligned}
& R_m(\Delta X) \\
= \ & \frac{1}{(m+1)!}\sum_{i_1,\cdots,i_{m+1}=1}^{n}\frac{\partial^{m+1} f}{\partial x_{i_1}\cdots\partial x_{i_{m+1}}}(X_0+\theta\Delta X)\Delta x_{i_1}\cdots\Delta x_{i_{m+1}}.
\end{aligned}
$$

带积分余项的泰勒公式

设$X_0 \in \mathbb{R}^n$，函数$f(X)$在点X_0的r邻域$B_r(X_0) = \{X \in \mathbb{R}^n | \ |X - X_0| < r\}$内$m+1$次连续可微，则对于任何$X \in B_r(X_0)$, 记$\Delta X = X - X_0$, 令$\varphi(t) = f(X_0 + t\Delta X)$, 有

$$\begin{aligned} & f(X_0 + \Delta X) \\ = \ & f(X_0) + \sum_{k=1}^{m} \frac{1}{k!} \mathrm{d}^k f(X_0) \bigg|_{\mathrm{d}x_i = \Delta x_i} + \frac{1}{m!} \int_0^1 (1-s)^m \varphi^{(m+1)}(s) \mathrm{d}s, \end{aligned}$$

其中

$$\varphi^{(m+1)}(s) = \sum_{i_1, \cdots, i_{m+1}=1}^{n} \frac{\partial^{m+1} f}{\partial x_{i_1} \cdots \partial x_{i_{m+1}}} (X_0 + s\Delta X) \Delta x_{i_1} \cdots \Delta x_{i_{m+1}}.$$

经常使用的两个特殊情形下的泰勒公式

柯西余项也可以类似地导出, 在此不再赘述. $m = 1, 2$是特别重要的, 我们写下它来以备后面使用.

$$\begin{aligned} & f(X) - f(X_0) \\ = \ & \langle \nabla f(X_0), \Delta X \rangle + \frac{1}{2} \Delta X \cdot H_f(X_0) \cdot \Delta X^T + o(|\Delta X|^2), \ |\Delta X| \to 0. \end{aligned}$$

$$\begin{aligned} & f(X) - f(X_0) \\ = \ & \langle \nabla f(X_0), \Delta X \rangle + \frac{1}{2} \Delta X \cdot H_f(X_0 + \theta \Delta X) \cdot \Delta X^T, \ 0 < \theta < 1. \end{aligned}$$

显然这里的行向量$\Delta X = (\Delta x_1, \cdots, \Delta x_n) = X - X_0$看作是$1 \times n$矩阵.

主观题 10分

设置

写出函数$f(x,y)=x^y$在点$(1,1)$处的二阶泰勒展开式.

正常使用主观题需2.0以上版本雨课堂

作答

例题

例 1 (凸函数的充要条件)

(i) 在凸区域D内定义的函数$f \in C^1(D)$是凸函数的充要条件为

$$f(X) - f(X_0) \geqslant \langle \nabla f(X_0), X - X_0 \rangle, \quad \forall X, X_0 \in D.$$

(ii) 在凸区域D内定义的函数$f \in C^2(D)$是凸函数的充要条件是$H_f(X) \geqslant 0$(即$H_f(X)$半正定).

例1的证明(i)

(i) 必要性见上节的公式(7), 下面证明充分性. 设X_1, $X_2 \in D$, $\lambda \in (0,1)$, 记$X_0 = \lambda X_1 + (1-\lambda)X_2$, 则有

$$f(X_1) - f(X_0) \geqslant \langle \nabla f(X_0), X_1 - X_0 \rangle = (1-\lambda)\langle \nabla f(X_0), X_1 - X_2 \rangle,$$

$$f(X_2) - f(X_0) \geqslant \langle \nabla f(X_0), X_2 - X_0 \rangle = \lambda \langle \nabla f(X_0), X_2 - X_1 \rangle,$$

第一个不等式两边乘λ与第二个不等式两边乘$1-\lambda$相加, 可得

$$\lambda f(X_1) + (1-\lambda) f(X_2) - f(\lambda X_1 + (1-\lambda)X_2) \geqslant 0.$$

例1的证明(ii)

(ii) 由(i)的结论, $f(X)$在D内是凸函数的充要条件为

$$f(X_0+\Delta X)-f(X_0)\geqslant\langle\nabla f(X_0),\Delta X\rangle$$

对任意$X_0\in D$, $X_0+\Delta X\in D$成立, 由公式(2), 有

$$\frac{1}{2}\Delta X\cdot H_f(X_0)\cdot\Delta X^T+o(|\Delta X|^2)\geqslant 0$$

成立. 从而必有二次型

$$V\cdot H_f(X_0)\cdot V^T\geqslant 0,\ \forall\, V\in\mathbb{R}^n$$

成立, 即$H_f(X_0)\geqslant 0$. 事实上, 若存在$V_0\neq 0$使$V_0\cdot H_f(X_0)\cdot V_0^T=\lambda<0$, 则令$\Delta X=tV_0$, 有$\Delta X\cdot H_f(X_0)\cdot\Delta X^T+o(|\Delta X|^2)=\lambda t^2+o(t^2)<0$在$|t|$很小时成立. 矛盾. 反之, 若$\Delta X\cdot H_f(X)\cdot\Delta X^T\geqslant 0$ 在$X\in D$成立, 则由(3), 有(4)成立, 从而$f(X)$在D内是凸函数. $\square$