

## 19. TAYLOR POLYNOMIALS

If  $f: A \longrightarrow \mathbb{R}^m$  is a differentiable function, and we are given a point  $P \in A$ , one can use the derivative to write down the best linear approximation to  $f$  at  $P$ . It is natural to wonder if one can do better using quadratic, or even higher degree, polynomials. We start with the one dimensional case.

**Definition 19.1.** Let  $I \subset \mathbb{R}$  be an open interval and let  $f: I \longrightarrow \mathbb{R}$  be a  $\mathcal{C}^k$ -function. Given a point  $a \in I$ , let

$$\begin{aligned} P_{a,k}f(x) &= f(a) + f'(a)(x-a) + \frac{f''(a)}{2}(x-a)^2 + \frac{f'''(a)}{3!}(x-a)^3 + \cdots + \frac{f^{(k)}(a)}{k!}(x-a)^k \\ &= \sum_{i=0}^k \frac{f^{(i)}(a)}{i!}(x-a)^i. \end{aligned}$$

Then  $P_{a,k}f(x)$  is the  $k$ th **Taylor polynomial of  $f$** , centred at  $a$ . The **remainder** is the difference

$$R_{a,k}f(x) = f(x) - P_{a,k}f(x).$$

Note that we have chosen  $P_{a,k}f$  so that the first  $k$  derivatives of  $P_{a,k}f$  at  $a$  are precisely the same as those of  $f$ . In other words, the first  $k$  derivatives at  $a$  of the remainder are all zero. The remainder is a measure of how good the Taylor polynomial approximates  $f(x)$  and so it is very useful to estimate  $R_{a,k}(x)$ .

**Theorem 19.2** (Taylor's Theorem with remainder). Let  $I \subset \mathbb{R}$  be an open interval and let  $f: I \longrightarrow \mathbb{R}$  be a  $\mathcal{C}^{k+1}$ -function. Let  $a$  and  $b$  be two points in  $I$ .

Then there is a  $\xi$  between  $a$  and  $b$ , such that

$$R_{a,k}f(b) = \frac{f^{(k+1)}(\xi)}{(k+1)!}(b-a)^{k+1}.$$

*Proof.* If  $a = b$  then take  $\xi = a$ . The result is clear in this case. Otherwise if we put

$$M = \frac{R_{a,k}f(b)}{(b-a)^{k+1}},$$

then

$$R_{a,k}f(b) = M(b-a)^{k+1}.$$

We want to show that there is some  $\xi$  between  $a$  and  $b$  such that

$$M = \frac{f^{(k+1)}(\xi)}{(k+1)!}.$$

If we let

$$g(x) = R_{a,k}(x) - M(x - a)^{k+1},$$

then

$$g^{k+1}(x) = f^{k+1}(x) - k!M.$$

Then we are looking for  $\xi$  such that

$$g^{k+1}(\xi) = 0.$$

Now the first  $k$  derivatives of  $g$  at  $a$  are all zero,

$$g^i(a) = 0 \quad \text{for} \quad 0 \leq i \leq k.$$

By choice of  $M$ ,

$$g(b) = 0.$$

So by the mean value theorem, applied to  $g(x)$ , there is a  $\xi_1$  between  $a$  and  $b$  such that

$$g'(\xi_1) = 0.$$

Again by the mean value theorem, applied to  $g'(x)$ , there is a  $\xi_2$  between  $a$  and  $\xi_1$  such that

$$g''(\xi_2) = 0.$$

Continuing in this way, by induction we may find  $\xi_i$ ,  $1 \leq i \leq k+1$  between  $a$  and  $\xi_{i-1}$  such that

$$g^i(\xi_i) = 0.$$

Let  $\xi = \xi_{k+1}$ .

□

Let's try an easy example. Start with

$$\begin{aligned}
f(x) &= x^{1/2} \\
f'(x) &= \frac{1}{2}x^{-1/2} \\
f''(x) &= \frac{1}{2^2}x^{-3/2} \\
f'''(x) &= \frac{3}{2^3}x^{-5/2} \\
f^4(x) &= -\frac{1 \cdot 3 \cdot 5}{2^4}x^{-7/2} \\
f^5(x) &= \frac{1 \cdot 3 \cdot 5 \cdot 7}{2^5}x^{-9/2} \\
f^6(x) &= -\frac{1 \cdot 3 \cdot 5 \cdot 7 \cdot 9}{2^6}x^{-11/2} \\
f^k(x) &= (-1)^{k-1} \frac{(2k-1)!!}{2^k} x^{-(2k-1)/2} \\
f^k(9/4) &= (-1)^{k-1} \frac{(2k-1)!!}{2^k} \frac{2^{2k-1}}{3^{2k-1}} \\
&= (-1)^{k-1} \frac{(2k-1)!! 2^{k-1}}{3^{2k-1}}.
\end{aligned}$$

Let's write down the Taylor polynomial centred at  $a = 9/4$ .

$$\begin{aligned}
P_{9/4,5}f(x) &= f(9/4) + f'(9/4)(x-9/4) + f''(9/4)/2(x-9/4)^2 + f'''(9/4)/6(x-9/4)^3 \\
&\quad + f^4(9/4)/24(x-9/4)^4 + f^5(9/4)/120(x-9/4)^5.
\end{aligned}$$

So,

$$\begin{aligned}
P_{9/4,5}f(x) &= 3/2 + 1/3(x-9/4) - 1/3^3(x-9/4)^2 + 2/3^5(x-9/4)^3 \\
&\quad - \frac{1 \cdot 3 \cdot 5 \cdot 2^3}{24 \cdot 3^7}(x-9/4)^4 + \frac{1 \cdot 3 \cdot 5 \cdot 7 \cdot 2^4}{120 \cdot 3^9}(x-9/4)^5.
\end{aligned}$$

If we plug in  $x = 2$ , so that  $x - 9/4 = -1/4$  we get an approximation to  $f(2) = \sqrt{2}$ .

$$P_{9/4,3}(2) = 3/2 + 1/3(-1/4) - 1/3^3(1/4)^2 - 2/3^5(1/4)^3 = \frac{10997}{7776} \approx 1.41422 \dots$$

On the other hand,

$$|R_3(2, 9/4)| = \frac{1 \cdot 3}{4!}(\xi)^{-7/2}(1/4)^4 < \frac{1 \cdot 3}{4!}(1/2) = 1/16.$$

In fact

$$|R_3(2, 9/4)| = \frac{10997}{7776} - \sqrt{2} \approx 4 \times 10^{-6}.$$

**Definition 19.3.** Let  $A \subset \mathbb{R}^n$  be an open subset which is convex (if  $\vec{a}$  and  $\vec{b}$  belong to  $A$ , then so does every point on the line segment between them). Suppose that  $f: A \rightarrow \mathbb{R}$  is  $\mathcal{C}^k$ .

Given  $\vec{a} \in A$ , the  $k$ th **Taylor polynomial** of  $f$  centred at  $\vec{a}$  is

$$P_{\vec{a},k}f(\vec{x}) = f(\vec{a}) + \sum_{1 \leq i \leq n} \frac{\partial f}{\partial x_i}(\vec{a})(x_i - a_i) + \frac{1}{2} \sum_{1 \leq i,j \leq n} \frac{\partial^2 f}{\partial x_i \partial x_j}(\vec{a})(x_i - a_i)(x_j - a_j) + \dots$$

$$+ \frac{1}{k!} \sum_{1 \leq i_1, i_2, \dots, i_k \leq n} \frac{\partial^k f}{\partial x_{i_1} \partial x_{i_2} \dots \partial x_{i_k}}(\vec{a})(x_{i_1} - a_{i_1})(x_{i_2} - a_{i_2}) \dots (x_{i_k} - a_{i_k}).$$

The **remainder** is the difference

$$R_{\vec{a},k}f(\vec{x}) = f(\vec{x}) - P_{\vec{a},k}f(\vec{x}).$$

**Theorem 19.4.** Let  $A \subset \mathbb{R}^n$  be an open subset which is convex. Suppose that  $f: A \rightarrow \mathbb{R}$  is  $\mathcal{C}^{k+1}$ , and let  $\vec{a}$  and  $\vec{b}$  belong to  $A$ .

Then there is a vector  $\vec{\xi}$  on the line segment between  $\vec{a}$  and  $\vec{b}$  such that

$$R_{\vec{a},k}(\vec{b}) = \frac{1}{(k+1)!} \sum_{1 \leq l_1, l_2, \dots, l_{k+1} \leq n} \frac{\partial^{k+1} f}{\partial x_{l_1} \partial x_{l_2} \dots \partial x_{l_{k+1}}}(\vec{\xi})(b_{l_1} - a_{l_1})(b_{l_2} - a_{l_2}) \dots (b_{l_{k+1}} - a_{l_{k+1}}).$$

*Proof.* As  $A$  is open and convex, we may find  $\epsilon > 0$  so that the parametrised line

$$\vec{r}: (-\epsilon, 1 + \epsilon) \rightarrow \mathbb{R}^n \quad \text{given by} \quad \vec{r}(t) = \vec{a} + t(\vec{b} - \vec{a}),$$

is contained in  $A$ . Let

$$g: (-\epsilon, 1 + \epsilon) \rightarrow \mathbb{R},$$

be the composition of  $\vec{r}(t)$  and  $f(\vec{x})$ .

**Claim 19.5.**

$$P_{0,k}g(t) = P_{\vec{a},k}f(\vec{r}(t)).$$

*Proof of (19.5).* This is just the chain rule;

$$g'(t) = \sum_{1 \leq i \leq n} \frac{\partial f}{\partial x_i}(\vec{r}(t))(b_i - a_i)$$

$$g''(t) = \sum_{1 \leq i,j \leq n} \frac{\partial^2 f}{\partial x_i \partial x_j}(\vec{r}(t))(b_i - a_i)(b_j - a_j)$$

and so on. □

So the result follows by the one variable result. □

We can write out the first few terms of the Taylor series of  $f$  and get something interesting. Let  $\vec{h} = \vec{x} - \vec{a}$ . Then

$$P_{\vec{a},2}f(x) = f(\vec{a}) + \sum_{1 \leq i \leq n} \frac{\partial f}{\partial x_i}(\vec{a})h_i + 1/2 \sum_{1 \leq i < j \leq n} \frac{\partial^2 f}{\partial x_i \partial x_j}(\vec{a})h_i h_j.$$

The middle term is the same as multiplying the row vector formed by the gradient of  $f$ ,

$$\nabla f(\vec{a}) = \left( \frac{\partial f}{\partial x_1}(\vec{a}), \frac{\partial f}{\partial x_2}(\vec{a}), \dots, \frac{\partial f}{\partial x_n}(\vec{a}) \right),$$

and the column vector given by  $\vec{h}$ . The last term is the same as multiplying the matrix with entries

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(\vec{a}),$$

on the left by  $\vec{h}$  and on the right by the column vector given by  $\vec{h}$  and dividing by 2.

The matrix

$$Hf(\vec{a}) = \left( \frac{\partial^2 f}{\partial x_i \partial x_j}(\vec{a}) \right),$$

is called the **Hessian** of  $f(\vec{x})$ .