Lecture 13

13.1 Taylor's Theorem

Recall, Taylor's theorem for one variable.

Definition 13.1.1 ▶ Taylor's Polynomial

Let, $f : \mathcal{O}_1 \to \mathbb{R}$ be C^k ($k \in \mathbb{N}$). Then for all h such that $a + h \in \mathcal{O}_1$,

$$p_{a,k}(a+h) = \sum_{n=0}^{k} \frac{f^{(n)}(a)}{n!} h^n$$

is called the **Taylor's Polynomial** of f around a.

Question. "Is $f(x) \approx p_{a,k}(x)$, for x close to a"?
We have,

$$p_{a,k}(x) = \sum_{n=0}^{k} \frac{f^{(n)}(a)}{n!} (x-a)^n$$

Take, $f(x) - p_{a,k}(x) = r_{a,k}(x)$

Theorem 13.1.1 (Taylor's Theorem)

Let $f : \mathcal{O}_1 \to \mathbb{R}$ be C^{k+1}. Then, $f(x) = p_{a,k}(x) + r_{a,k}(x)$ where,

$$r_{a,k} = \frac{f^{k+1}(c)}{(k+1)!}(x-a)^{k+1}$$

for some c in between a and $x \in \mathcal{O}_1$.

We introduce the following notation for the sake of clarity in the multivariate Taylor expansion. Let $\alpha = (\alpha_1, \alpha_2, \cdots, \alpha_n) \in \mathbb{Z}^n_{\geq 0}$, and define

- $|\alpha| = \sum_{i=1}^{n} \alpha_i$
- $\alpha! = \alpha_1!\alpha_2! \cdots \alpha_n!$ (product of coordinate-wise factorials)
- $\partial^\alpha = \dfrac{\partial^{|\alpha|}}{\partial {x_1}^{\alpha_1} \cdots \partial {x_n}^{\alpha_n}}$ (α^{th} derivative)
- $\boldsymbol{\nabla} \cdot h = \displaystyle\sum_{i=1}^{n} h_i \frac{\partial}{\partial x_i}$

The last definition when iterated gives,

$$(\boldsymbol{\nabla} \cdot h)^m = \sum_{|\alpha|=m} \frac{m!}{\alpha!} h^\alpha \partial^\alpha$$

Theorem 13.1.2 (Taylor's Theorem in Multivariate Case)

Let, $f : \mathcal{O}_n \to \mathbb{R}$ be a C^{k+1} function and assume $\mathcal{O}_n$ is convex. If $h, a+h \in \mathcal{O}_n$, then

$$f(a+h) = \sum_{|\alpha| \le k} \frac{1}{\alpha!} (\partial^\alpha f)(a) h^\alpha + r_{a,k}(h)$$

where,

$$r_{a,k}(h) = \sum_{|\alpha| = k+1} \frac{1}{\alpha!} (\partial^\alpha f)(a + ch) h^\alpha \qquad \text{for some } c \in (0,1)$$

Proof. Define, $\eta : [0,1] \to \mathbb{R}$ as $\eta(t) = f(a+th)$

$$t \longmapsto a + th \longmapsto f(a+th), \qquad \eta : t \mapsto f(a+th)$$

which implies η is a C^{k+1} function around 0.

$$\therefore \eta'(t) = \boldsymbol{\nabla} f(a+th) \cdot h = (\boldsymbol{\nabla} \cdot h) f(a+th)$$

Claim

$$\eta^{(m)}(t) = (\boldsymbol{\nabla} \cdot h)^m f(a+th) \ \forall \ m \in \{0, 1, \dots, k+1\}$$

Proof. The first derivative of η,

$$\eta'(t) = \boldsymbol{\nabla} f(a+th) \cdot h = \sum_{i=1}^{n} f_{x_i}(a+th) h_i$$

which we use to compute the second derivative,

$$\begin{aligned}
\eta''(t) &= \frac{\mathrm{d}}{\mathrm{d}t} \left(\sum_{i=1}^{n} f_{x_i}(a+th) h_i \right) \\
&= \sum_{i=1}^{n} \frac{\mathrm{d}}{\mathrm{d}t} \left(f_{x_i}(a+th) h_i \right) \\
&= \sum_{i=1}^{n} h_i \sum_{j=1}^{n} f_{x_i x_j}(a+th) h_j && \text{(Chain rule of partials)} \\
&= \sum_{i,j=1}^{n} h_i h_j f_{x_i x_j}(a+th) \\
&= (\boldsymbol{\nabla} \cdot h)^2 f(a+th)
\end{aligned}$$

Proceeding with induction on the order of the derivative, we get $\eta^{(m)}(t) = (\boldsymbol{\nabla} \cdot h)^m f(a+th)$ for all $0 \le m \le k+1$ which is our claim. □

By one-variable Taylor's Theorem,

$$\eta(1) = p_{0,k}(1) + r_{0,k}(c) \text{ for some } c \in (0,1) \tag{13.1}$$

with

$$p_{0,k}(1) = \eta(0) + \frac{\eta'(0)}{1!} + \cdots + \frac{\eta^{(k)}(0)}{k!} \tag{13.2}$$

and

$$r_{0,k}(c) = \frac{\eta^{(k+1)}(c)}{(k+1)!} \tag{13.3}$$

Substituting $\eta^{(m)}(t)$ in (13.1) we have,

$$f(a+h) = \sum_{|\alpha|\leq k} \frac{1}{\alpha!}(\partial^\alpha f)(a)h^\alpha + r_{a,k}(h)$$

□

We note that, in particular, if $f : \mathcal{O}_2 \to \mathbb{R}$ is a C^2 function then we have,

$$f(a+h) = f(a) + \nabla f(a) \cdot h + \frac{1}{2}h^t H_f(a+ch)h \tag{13.4}$$

where,

$$H_f(a) = \begin{pmatrix} f_{xx}(a) & f_{xy}(a) \\ f_{xy}(a) & f_{yy}(a) \end{pmatrix}$$

and $c \in (0,1)$.

Theorem 13.1.3 (Extremum)

Let $f : \mathcal{O}_2 \to \mathbb{R}$ be a C^2 function such that $Df(a) = 0$. We write

$$H_f(a) = \begin{pmatrix} f_{xx}(a) & f_{xy}(a) \\ f_{xy}(a) & f_{yy}(a) \end{pmatrix}$$

Then,

(i) $f(a)$ is a local maximum if $f_{xx}(a) < 0$ and $\det(H_f(a)) > 0$

(ii) $f(a)$ is a local minimum if $f_{xx}(a) > 0$ and $\det(H_f(a)) > 0$

(iii) a is a saddle point if $\det(H_f(a)) < 0$

Proof. As a is an interior point of $\mathcal{O}_2$ we can get an $r > 0$ such that $a, a+h \in B_r(a) \subseteq \mathcal{O}_2$. By (13.4),

$$f(a+h) - f(a) = \nabla f(a) \cdot h + \frac{1}{2}h^t H_f(a+ch)h$$

We will prove (ii), other statements can be proved similarly. Our assumptions tell that $H_f(a)$ is positive definite. Hence, by the Lemma 12.2.1, there exist $\epsilon > 0$ such that $H_f(x)$ is positive definite $\forall\ x \in B_\epsilon(a)$. So for every $x \in B_\epsilon(a)$ the quadratic form $h^t H_f(x)h > 0$ with $h \neq 0$ which implies $f(x) - f(a) > 0$ in $B_\epsilon(a)$ that means a is a point of local minimum. □

Example 13.1.1 (Finding Critical points of a function and their nature)

Find the critical points and discuss the nature of the function

$$f(x,y) = x^3 - 6x^2 - 8y^2$$

Solution. Setting $\nabla f(x,y) = 0$, i.e., $(f_x, f_y)(x,y) = 0$, we get the system of equations

$$3x^2 - 12x = 0 \text{ and } -16y = 0$$

whose solution set is $\{(0,0),(4,0)\}$ implying that $(0,0),(4,0)$ are critical points.
The 2nd derivatives are,

$$f_{xx}(x,y) = 6x - 12,\ \ f_{yy}(x,y) = -16 \text{ and } f_{xy}(x,y) = 0$$

Now, we compute the determinant of the hessian at these points to tell their nature. For $(0, 0)$,

$$\det(H_f(0,0)) = \begin{vmatrix} -12 & 0 \\ 0 & -16 \end{vmatrix} > 0 \text{ and } f_{xx}(0,0) = -12 < 0$$

So, f has a local maximum at $(0, 0)$. And at $(4, 0)$,

$$\det(H_f(4,0)) = \begin{vmatrix} 12 & 0 \\ 0 & -16 \end{vmatrix} < 0$$

which shows that $(4, 0)$ is a saddle point. ∎