Infinitesimal Calculus 3

Lecture 17, Wednesday December 28, 2022
Ari Feiglin

We now generalize Taylor polynomials to more than 1 or 2 dimensions. Suppose $f\colon \mathbb{R}^n \longrightarrow \mathbb{R}$ which is differentiable $m+1$ times ($f \in C^{m+1}$). We can define an initial point $x^0 = (x_1^0, \dots, x_n^0)$ and h be our difference vector $h = (h_1, \dots, h_n)$ then:

$$f(x^0+h) = \sum_{k=0}^{m} \frac{1}{k!}\left(\left(h_1\frac{\partial}{\partial x_1} + \cdots + h_n\frac{\partial}{\partial x_n}\right)^k f\right)(x^0) + \frac{1}{(n+1)!}\cdot\left(\left(h_1\frac{\partial}{\partial x_1} + \cdots + h_n\frac{\partial}{\partial x_n}\right)^{n+1} f\right)(x^0+\theta h)$$

for some $0 \leq \theta \leq 1$ where:

$$\left(h_1\frac{\partial}{\partial x_1} + \cdots + h_n\frac{\partial}{\partial x_n}\right)^k f = \sum_{i_1+\cdots+i_n=k} \binom{k}{i_1,\dots,i_n} \cdot \frac{\partial^k f}{\partial x_1^{i_1}\dots\partial x_n^{i_n}} h_1^{i_1}\cdots h_n^{i_n}$$

If we perform an $m=0$ order Taylor expansion:

$$f(x^0+h) = f(x^0) + \sum_{k=0}^{n} h_k f_{x_k}(x^0+\theta h) \implies f(x^0+h) - f(x^0) = \sum_{k=0}^{n} h_k f_{x_k}(x^0+\theta h)$$

And this is equal to $\nabla f(x^0+\theta h)\cdot h$, and so we get that:

$$f(y^0) - f(x^0) = \nabla f(x^0 + \theta(y^0 - x_0))\cdot(y^0 - x^0)$$

So if we let $c = \nabla f(x^0 + \theta(y^0 - x^0))$ then we get the following theorem, which is a generalization of Lagrange:

Theorem 17.1 (Lagrange's Theorem):

If $f\colon \mathbb{R}^n \longrightarrow \mathbb{R}$ is differentiable then for every $x, y \in \mathbb{R}^n$:

$$f(y) - f(x) = c\cdot(y-x)$$

where c is in $\overleftrightarrow{xy}$.

Corollary 17.2:

Suppose $D \subseteq \mathbb{R}^n$ is open and connected (and therefore path-connected), and $f\colon D \longrightarrow \mathbb{R}$ is differentiable. If ∇f is identically 0 in D, then f is constant.

Proof:

Let $x, y \in D$, since D is path connected, there is a path between x and y. Since D is open, it is polygonal connected, so we can assume that x and y are connected by a line (otherwise, we show that $f(x) = f(x_1)$ the next point in the polygonal chain connecting x and y and so on). We know that $f(x) - f(y) = \nabla f(c)\cdot(x-y)$, and c must lie on the line between x and y, so $c \in D$ and therefore $f(x) - f(y) = 0$ as required.

■

Theorem 17.3:

Suppose $f\colon \mathbb{R}^n \longrightarrow \mathbb{R}$ defined in a neighborhood of $x^0 = (x_1^0, \dots, x_n^0)$ which accepts a maximum or minimum there. If for every $1 \leq k \leq n$, $f_{x_k}(x^0)$ exists, then they are all 0.

Proof:

We define $g_k(x) = f(x_1^0, \dots, x_{k-1}^0, x, x_k^0, \dots, x_n^0)$, then g_k has a maximum or minimum at x_k^0. Furthermore, we know that $g_k'(x_k^0) = f_{x_k}(x^0)$ and since g_k has a local maximum at this point, $g'(x_k^0) = 0$ and therefore $f_{x_k}(x^0) = 0$ as well. ■

Notice then that at a local maximum or minimum, $\nabla f(x) = 0$.

Definition 17.4:

A **critical point** of f is a point x such that $\nabla f(x)$.

So local minima and maxima are critical points, but not all critical points are maxima or minima. The degree 1 Taylor expansion of $f\colon \mathbb{R}^2 \longrightarrow \mathbb{R}$ at a critical point is:

$$f(x_0 + h, y_0 + \ell) = f(x, y) + \frac{1}{2}\big(f_{xx}h^2 + 2f_{xy}h\ell + f_{yy}\ell^2\big)(x^0 + \theta\vec{h})$$

If we let $A = f_{xx}(x^0 + \theta\vec{h})$, $B = f_{xy}(x^0 + \theta\vec{h})$, and $C = f_{yy}(x^0 + \theta\vec{h})$ we have that:

$$f(x^0 + \vec{h}) - f(x^0) = \frac{1}{2}\big(Ah^2 + 2Bhk + Ck^2\big)$$

Notice then that if we focus on the polynomial:

$$Ah^2 + 2Bhk + Ck^2 = k^2\left(A\left(\frac{h}{k}\right)^2 + 2B\left(\frac{h}{k}\right) + C\right)$$

if we let $x = \frac{h}{k}$ then this is equal to $Ax^2 + 2Bx + C$ multiplied by some positive constant. Notice then that if the discriminant is negative the polynomial doesn't change its sign, that is if $B^2 - AC < 0$.

Theorem 17.5:

Suppose $f(x, y)$ is in C^2 and (x_0, y_0) is a critical point. If at (x_0, y_0)

(1) $f_{xx}f_{yy} - f_{xy} > 0$, then the discriminant is negative, and $f_{xx} > 0$ then the point is a minimum.

(2) $f_{xx}f_{yy} - f_{xy} > 0$ and $f_{xx} < 0$ then the point is a maximum.

(3) $f_{xx}f_{yy} - f_{xy} < 0$ then the point is neither a maximum nor a minimum.

(4) $f_{xx}f_{yy} - f_{xy} = 0$ then everything is possible.

Proof:

(1) Since second derivatives are continuous, the discriminant is negative in a neighborhood of (x_0, y_0) so for every point in the neighborhood, we have that

$$f(x, y) - f(x_0, y_0) = \frac{1}{2} \cdot p(x)$$

for some polynomial $p(x)$, whose sign doesn't change in this neighborhood, so $f(x_0, y_0)$ is either below or above every point in this neighborhood, and from what we know about single dimension second derivatives, since $f_{xx}(x_0, y_0) > 0$ the point is a maximum.

(2) The proof is identical to what it is above.

(3) Since the discriminant is positive, the difference in f is positive and negative in any neighborhood of (x_0, y_0) and it's therefore not a maximum nor minimum.

Definition 17.6:

The **Hessian** of a function $f\colon \mathbb{R}^n \longrightarrow \mathbb{R}$ in C^2 is a matrix $H(\vec{x}) \in \mathbb{R}^{n\times n}$ defined by $[H]_{ij} = f_{x_i x_j}(\vec{x})$.

Note that if $\vec{k} = (k_1, \dots, k_n)$ then:

$$\vec{k}^T H(\vec{x})\vec{k} = \big(f_{x_1}h_1 + \cdots + f_{x_n}h_n\big)^2(\vec{x})$$

Definition 17.7:

A matrix A is **positive** if for every vector $\vec{k}$, $\vec{k}^T A\vec{k} > 0$.