LECTURE 16

Date of Lecture: March 23, 2020

As always, $\mathbf{K} \in \{\mathbf{R}, \mathbf{C}\}$.

The symbol ☡ is for flagging a cautionary comment or a tricky argument. It occurs in the margins and is Knuth's version of Bourbaki's "dangerous bend symbol".

An n-tuple $(x_1, \dots, x_n)$ of symbols (x_i not necessarily real or complex numbers) will also be written as a column vector when convenient. Thus

$$(x_1, \dots, x_n) = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}.$$

A map $\boldsymbol{f}$ from a set S to a product set $T_1 \times \cdots \times T_n$ will often be written as an n-tuple $\boldsymbol{f} = (f_1, \dots, f_n)$, with f_i a map from S to T_i, and hence, by the above convention, as a column vector

$$\boldsymbol{f} = \begin{bmatrix} f_1 \\ \vdots \\ f_n \end{bmatrix}.$$

(See Remark 2.2.2 of Lecture 5.)

The default norm on Euclidean spaces of the form $\mathbf{R}^n$ is the Euclidean norm $\| \ \|_2$ and we will simply denote it as $\| \ \|$.

☡ Note that $(x_1, \dots, x_n) \neq [x_1 \dots x_n]$. Each side is the transpose of the other.

1. Taylor's Theorem in several variable

We recall (and prove) the approximation of an n-times differentiable function by a polynomial of degree $n-1$, which is one version of Taylor's theorem (there are various theorems, all related, which go by the same name). We then use this to prove Taylor's theorem in many variables.

1.1. The one variable case.

We need some terminology. Let $f\colon [a,b] \to \mathbf{R}$ be a function. We say f is differentiable on $[a,b]$ if:

- f is differentiable on (a,b).
- The right derivative of f at a exists and the left derivative of f at b exists.

If f is differentiable on $[a,b]$ we write $f'(a)$ for the right derivative of f at a and $f'(b)$ for the left derivative of f at b, so that in this case we have a map $f'\colon [a,b] \to \mathbf{R}$. Sometimes we use the phrase "*f' exists on $[a,b]$*" as a shorthand for "*f is differentiable on $[a,b]$*". It is clear that we can also define what it means for f to be *n-times differentiable on* $[a,b]$, and if f is so, we write $f^{(n)}$ for its n^{th} derivative (though we will prefer f'' for the second derivative and sometimes f''' for the third derivative). Further the phrase "*$f^{(n)}$ exists on $[a,b]$*" is a shorthand for "*f is n-times differentiable on $[a,b]$*".

We will be using the following version of the *Mean Value Theorem* (see [R, Theorem 5.10, p. 108]).

> *If f is a real valued continuous function on $[a, b]$ which is differentiable in (a, b), then there is a point $x \in (a, b)$ at which*
>
> $$f(b) - f(a) = f'(x)(b - a).$$

The following version (and proof) of Taylor's theorem is taken from [R, Theorem 5.15, pp.110–111].

Theorem 1.1.1. *Let n be a positive integer, $[a, b]$ a closed and bounded interval in $\mathbf{R}$, and f a real valued function on $[a, b]$ such that $f^{(n-1)}$ exists and is continuous on $[a, b]$ and $f^{(n)}$ exists on (a, b). Let s and t be two distinct points in $[a, b]$. Then there exists a number θ, strictly between s and t, such that*

$$f(t) = f(s)+f'(s)(t-s)+\frac{f''(s)}{2!}(t-s)^2+\cdots+\frac{f^{(n-1)}(s)}{(n-1)!}(t-s)^{n-1}+\frac{f^{(n)}(\theta)}{n!}(t-s)^n.$$

Proof. For $x \in \mathbf{R}$ set

$$P(x) := \sum_{k=0}^{n-1} \frac{f^{(k)}(s)}{k!}(x-s)^k.$$

We point out that P also depends on $s \in [a, b]$ and $n \in \mathbf{N}$ and so sometimes is denoted $P(s, x)$ and sometimes even more elaborately as $P(n, s, x)$. We will not have occasion to use these notations. Note that P, as a function of x, is a polynomial of degree (at most) $n - 1$.

Define $M \in \mathbf{R}$ by the formula

$$M = \frac{f(t) - P(t)}{(t-s)^n}.$$

Since $s \neq t$, M is a well-defined real number. Moreover, we have

$$(*) \qquad f(t) = P(t) + M(t-s)^n.$$

We have a map $g\colon [a, b] \to \mathbf{R}$ given by

$$(**) \qquad g(x) = f(x) - P(x) - M(x-s)^n \qquad (x \in [a, b]).$$

It is immediate from the definition of M that

$$(\dagger) \qquad g(t) = 0.$$

Now for $j, k \in \{0, \dots, n-1\}$ we have the following relation.

$$\frac{\mathrm{d}^k}{\mathrm{d}x^k}\Big(\frac{f^{(j)}(s)(x-s)^j}{j!}\Big)\Big|_{x=s} = \begin{cases} 0, & j \neq k, \\ f^{(k)}(s), & j = k. \end{cases}$$

It follows that $P^{(k)}(s) = f^{(k)}(s)$ for $k = 0, 1, \dots, n-1$. In particular, from $(**)$, we have

$$(\ddagger) \qquad g^{(k)}(s) = 0 \qquad (k = 0, 1, \dots, n-1).$$

Setting $k = 0$ in $(\ddagger)$ we see that $g(s) = 0$, and this together with $(\dagger)$ and the mean value theorem tells us that there is a θ_1 strictly between s and t such that $g'(\theta_1) = 0$. Using $(\ddagger)$ and the mean value theorem again (now applied to g') we see that there exists θ_2 strictly between s and θ_1 such that $g''(\theta_2) = 0$. Continuing in this mannner, we can find θ_n strictly between s and θ_{n-1} such that $g^{(n)}(\theta_n) = 0$. Let $\theta = \theta_n$. Then θ lies strictly between s and t and

$$0 = g^{(n)}(\theta) = f^{(n)}(\theta) - P^{(n)}(\theta) - n!M = f^{(n)}(\theta) - n!M$$

since P is a polynomial of degree at most $n-1$. This means $M = \frac{f^{(n)}(\theta)}{n!}$, and hence by $(*)$ we are done. ☐

Corollary 1.1.2. *If $f^{(n)}$ is bounded on (a,b), then*

$$\lim_{h\to 0} \frac{1}{h^{n-1}} \left\{ f(s+h) - \sum_{k=0}^{n-1} \frac{f^{(k)}(s)}{k!} h^k \right\} = 0.$$

Proof. Let $|f^{(n)}(x)| \leq C$ for $x \in (a,b)$ where C is a constant, and set $t = s+h$. The expression within the curly braces, according to the theorem, is $h^n f^{(n)}(\theta)/n!$ for some θ strictly between s and $s+h$. Thus

$$\left| f(s+h) - \sum_{k=0}^{n-1} \frac{f^{(k)}(s)}{k!} h^k \right| \leq \frac{C}{n!} |h|^n.$$

The result follows. ☐

Perhaps the most useful version, if one's focus is on maxima and minima problems, is the following:

Corollary 1.1.3. *Suppose f is $\mathscr{C}^n$ on (a,b) and $s \in (a,b)$. Then*

$$\lim_{h\to 0} \frac{1}{h^{n-1}} \left\{ f(s+h) - \sum_{k=0}^{n-1} \frac{f^{(k)}(s)}{k!} h^k \right\} = 0.$$

Proof. We can find $\alpha, \beta \in (a,b)$ such that $s \in (\alpha,\beta) \subset [\alpha,\beta] \subset (a,b)$. Now, $f^{(n)}$ is continuous on (a,b) and hence is bounded on $[\alpha,\beta]$, whence on (α,β). We can now apply the previous corollary to $f|_{[\alpha,\beta]}$. ☐

Remark 1.1.4. Suppose f is $\mathscr{C}^3$ on (a,b) and $f'(s) = 0$ at some $s \in (a,b)$. According to Corollary 1.1.3, if $\eta(h) = f(s+h) - f(s) - \frac{1}{2}f''(s)h^2$, then $\eta(h)/h^2 \to 0$ as $h \to 0$. Now

$$\frac{f(s+h) - f(s)}{h^2} = \frac{f''(s)}{2} + \frac{\eta(h)}{h^2}.$$

If $f''(s)$ is positive then there exists a $\delta > 0$ such that $|\eta(h)/h^2| < \frac{1}{2}f''(s)$ for all $0 < |h| < \delta$. This means $\frac{1}{2}f''(s) + \eta(h)/h^2 > 0$ for $0 < |h| < \delta$. In particular $f(s+h) - f(s) > 0$ for such $0 < |h| < \delta$. Thus

$$f(s) < f(s+h) \qquad (0 < |h| < \delta),$$

i.e. f has a *local minimum* at s. This is the famous second derivative test for local minima that you are all familiar with. One can similarly deduce the second derivative test for local maxima. We can also deduce that by considering the function $-f$ and applying the second derivative test for local minima.

1.2. **Combinatorial identity.** Let $P(k; s_1, \dots, s_n)$ denote the number of arrangements of k objects, the objects being of n types, with s_1 of type 1, s_2 of type 2, ..., s_n of type n (so that $s_1 + \cdots + s_n = k$). Then it is well known and easy to prove that

$$P(k; s_1, \dots, s_n) = \frac{k!}{s_1! s_2! \dots s_n!}. \tag{1.2.1}$$

1.3. **Taylor's in several variables.** Let U be an open subset of $\mathbf{R}^n$ and $f\colon U \to \mathbf{R}$ a $\mathscr{C}^m$ function. Fix a point $\boldsymbol{a}$ in U. There exists an open ball B centred at $\mathbf{0}$ such that $\overline{B} + \boldsymbol{a} \subset U$, where $\overline{B}$ is, as usual the closure of the open ball B. For $\boldsymbol{h}$ in B (or even $\overline{B}$), $\boldsymbol{a} + t\boldsymbol{h} \in U$ for $t \in [0, 1]$. This means the following. If $\phi\colon \mathbf{R} \times \mathbf{R}^n \to \mathbf{R}^n$ is the map given by the formula

$$\phi(t, \boldsymbol{h}) = \boldsymbol{a} + t\boldsymbol{h},$$

then $\phi^{-1}(U)$ contains the compact set

$$K := [0, 1] \times \overline{B}$$

in $\mathbf{R}^{n+1} = \mathbf{R} \times \mathbf{R}^n$. It is not difficult to see that there is an open rectangle $(a, b) \times V$ in $\phi^{-1}(U)$ such that $K \subset (a, b) \times V$.

Fix $\boldsymbol{h} \in B$. Define $g\colon (a, b) \to \mathbf{R}$ by the formula

$$g(t) = f(\boldsymbol{a} + t\boldsymbol{h}).$$

Since g is the composite of a $\mathscr{C}^\infty$ function and a $\mathscr{C}^m$ function, it is a $\mathscr{C}^m$ function.

For $k \in \{1, \dots, m\}$, a repeated application of the chain rule gives

$$g^{(k)}(t) = \sum (D_{i_1 \dots i_k} f)(\boldsymbol{a} + t\boldsymbol{h}) h_{i_1} \dots h_{i_k},$$

where the sum is taken over $(i_1, \dots, i_k) \in \{1, \dots, n\}^k$. Since $k \in \{1, \dots, m\}$ and f is $\mathscr{C}^m$, it is $\mathscr{C}^k$. By Theorem 2.2.1 of Lecture 11, if $(j_1, \dots, j_k)$ is a permutation of $(i_1, \dots, i_k)$, then

$$D_{j_1 \dots j_k} = D_{i_1 \dots i_k}.$$

If $(i_1, \dots, i_k)$ has 1 occurring as a component s_1 times, 2 occurring s_2 times, ..., n occurring s_n times, it is clear that the number of $(j_1, \dots, j_k)$ which are permutations of $(i_1, \dots, i_k)$ is $P(k; s_1, \dots, s_n)$. Thus, using (1.2.1), we can rewrite the displayed formula for $g^{(k)}(t)$ as

$$g^{(k)}(t) = \sum_{s_1 + \cdots + s_n = k} \frac{k!}{s_1! \dots s_n!} (D_1^{s_1} \dots D_n^{s_n} f)(\boldsymbol{a} + t\boldsymbol{h}) h_1^{s_1} \dots h_n^{s_n} \tag{1.3.1}$$

According to the Taylor's theorem for one variable, i.e Theorem 1.1.1, we have

$$g(1) = \sum_{k=0}^{m-1} \frac{g^{(k)}(0)}{k!} + \frac{g^{(m)}(\theta)}{m!},$$

for some $\theta \in (0, 1)$. This, together with (1.3.1) gives

$$f(\boldsymbol{a} + \boldsymbol{h}) = \sum_{k=0}^{m-1} \sum_{s_1 + \cdots + s_n = k} \frac{(D_1^{s_1} \dots D_n^{s_n} f)(\boldsymbol{a})}{s_1! \dots s_n!} h_1^{s_1} \dots h_n^{s_n} + r(\boldsymbol{h}) \tag{1.3.2}$$

where

$$r(\boldsymbol{h}) = \sum_{s_1 + \cdots + s_n = m} \frac{(D_1^{s_1} \dots D_n^{s_n} f)(\boldsymbol{a} + \theta\boldsymbol{h})}{s_1! \dots s_n!} h_1^{s_1} \dots h_n^{s_n}. \tag{1.3.3}$$

Since $D_1^{s_1} \dots D_n^{s_n} f$ is continuous on U for $s_1, \dots, s_n$ such that $\sum_{i=1}^n s_i = m$ and since $\overline{B} + \boldsymbol{a}$ is compact, there is a constant C such that

$$\tfrac{1}{s_1! \dots s_n!} |(D_1^{s_1} \dots D_n^{s_n} f)(\boldsymbol{a} + \boldsymbol{x})| \leq C \qquad (\boldsymbol{x} \in B; \sum_{i=1}^n s_i = m).$$

Since $|h_i| \leq \|\boldsymbol{h}\|$ and hence by (1.3.3) as well as the above inequality we have

$$|r(\boldsymbol{h})| \leq K \|\boldsymbol{h}\|^m \qquad (\boldsymbol{h} \in B) \tag{1.3.4}$$

for $K = MC$, where M is the number of n-tuples of non-negative integers which sum to m.

Remark 1.3.5. The term $r(\boldsymbol{h})$ is often called the remainder term. What we have shown, via (1.3.4) is that $r(\boldsymbol{h})/\|\boldsymbol{h}\|^{m-1} \to 0$ as $\boldsymbol{h} \to \boldsymbol{0}$.

In view of the above discussion, we have proved the following version of Taylor's theorem for several variables.

Theorem 1.3.6. *Let U be an open subset of $\mathbf{R}^n$ and $f \colon U \to \mathbf{R}$ a $\mathscr{C}^m$ map for some $m \geq 1$. Let $\boldsymbol{a} \in U$. Then in an open neighbourhood W of $\boldsymbol{0}$, $W + \boldsymbol{a} \subset U$, we have*

$$f(\boldsymbol{a} + \boldsymbol{h}) = \sum_{k=0}^{m-1} \sum_{s_1 + \dots + s_n = k} \frac{(D_1^{s_1} \dots D_n^{s_n} f)(\boldsymbol{a})}{s_1! \dots s_n!} h_1^{s_1} \dots h_n^{s_n} + r(\boldsymbol{h})$$

where $r(\boldsymbol{h})/\|\boldsymbol{h}\|^{m-1} \to 0$ as $\boldsymbol{h} \to \boldsymbol{0}$. The remainder term $r(\boldsymbol{h})$ is of the form

$$r(\boldsymbol{h}) = \sum_{s_1 + \dots + s_n = m} \frac{(D_1^{s_1} \dots D_n^{s_n} f)(\boldsymbol{a} + \theta \boldsymbol{h})}{s_1! \dots s_n!} h_1^{s_1} \dots h_n^{s_n}$$

for some $\theta \in (0, 1)$.

1.4. **The Hessian matrix.** Let U be open in $\mathbf{R}^n$, $\boldsymbol{a}$ a point in U, and f a real valued map on U such that f'' exists at $\boldsymbol{a}$, i.e. f' exists in a neighbourhood of $\boldsymbol{a}$ and $(f')'$ exists at $\boldsymbol{a}$. The *Hessian matrix* of f at $\boldsymbol{a}$ is the $n \times n$ matrix $H(f)(\boldsymbol{a})$ whose $(i, j)^{\text{th}}$ entry is $D_{ij} f(\boldsymbol{a})$, $1 \leq i, j \leq n$. If f is $\mathscr{C}^2$, equality of order differentiation of mixed partials shows that $H(f)(\boldsymbol{x})$ is symmetric at every $\boldsymbol{x} \in U$.

Examples 1.4.1. Let $f(x, y, z) = x^2 - 4y^2 + 2z^2 - 8yz + 14zx$. Then $H(f)$ is defined on all on $\mathbf{R}^3$ and is

$$H(f)(x, y, z) = \begin{bmatrix} 2 & 0 & 14 \\ 0 & -8 & -8 \\ 14 & -8 & 4 \end{bmatrix}$$

which is $2A$, where A is the matrix associated with the quadratic form f.

Another example is $f(x, y) = x^3 y - 3xy + y^3$. In this case $H(f)$ is defined on all of $\mathbf{R}^2$ and

$$H(f)(x, y) = \begin{bmatrix} 6xy & 3x^2 - 3 \\ 3x^2 - 3 & 6y \end{bmatrix}.$$

The following theorem is the one that is most used in working out local extrema.

Theorem 1.4.2. *Let* U *be open in* $\mathbf{R}^n$*,* $\boldsymbol{a}$ *a point in* U *and* $f\colon U \to \mathbf{R}$ *a* $\mathscr{C}^3$ *function. Let* $H = H(f)(\boldsymbol{a})$ *be the Hessian matrix of* f *at* $\boldsymbol{a}$*. Then there exists an open neighbourhood* W *of* $\mathbf{0}$ *such that for* $\boldsymbol{x} \in W$ *we have*

$$f(\boldsymbol{a}+\boldsymbol{x}) = f(\boldsymbol{a}) + \langle \boldsymbol{\nabla} f(\boldsymbol{a}), \boldsymbol{x}\rangle + \frac{1}{2}\boldsymbol{x}^t H \boldsymbol{x} + r(\boldsymbol{x})$$

where $r(\boldsymbol{x})/\|\boldsymbol{x}\|^2 \to 0$ *as* $\boldsymbol{x} \to \mathbf{0}$*.*

Proof. This is just a special case of Theorem 1.3.6. ☐

About these notes. This lecture was supposed to be given on March 23, 2020. Classes got suspended on March 17 because of the coronavirus COVID 19 pandemic, and all teaching moved online. These course notes are a reasonably faithful record of the lectures given (before the shutdown) at the Chennai Mathematical Institute (CMI) in the January-April semester of 2019-20. The course is Analysis II, a core course for first year undergraduates at CMI. For more material related to this course, visit `https://www.cmi.ac.in/~pramath/teaching.html#ANA2`. If you have comments on these notes, or on related course material, please send an email to pramath@cmi.ac.in.

References

[R] W. Rudin, *Principles of Mathematical Analysis*, (Third Edition), McGraw-Hill, New Delhi, 1976.