Notes on Taylor's formula
Math 511, Spring 2018

1 Taylor's formula in one dimension

We begin by recalling the standard Taylor formula (see e.g. Theorem 5.15 of Rudin)

Theorem 1.1. *Suppose $g : [a, b] \to \mathbb{R}$ is a C^{k-1} function $[a, b]$ and that $g^{(k-1)}(t)$ is differentiable for all $t \in (a, b)$. Then there exists $t_0 \in (a, b)$ such that*

$$g(b) = \sum_{l=0}^{k-1} \frac{g^{(l)}(a)}{l!}(b-a)^l + \frac{g^{(k)}(t_0)}{k!}(b-a)^k.$$

If $g^{(k)}(t)$ is also Riemann integrable, the remainder term can also be characterized as an integral, as in the following theorem.

Theorem 1.2. *Suppose $g : [a, b] \to \mathbb{R}$ differentiable up to order k on $[a, b]$ and that $g^{(k)}(t)$ is Riemann integrable on $[a, b]$. Then*

$$g(b) = \sum_{l=0}^{k-1} \frac{g^{(l)}(a)}{l!}(b-a)^l + \frac{(b-a)^k}{(k-1)!} \int_0^1 (1-t)^{k-1} g^{(k)}(a + t(b-a))\, dt \quad (1.1)$$

Note that the hypotheses of this theorem are satisfied when g is $C^k([a, b])$.

Proof of Theorem 1.2. To prove the theorem we induct on k. The $k = 1$ case is a consequence of the fundamental theorem of calculus:

$$g(b) - g(a) = \int_0^1 \frac{d}{dt}\left(g(a + t(b-a))\right) dt = (b-a) \int_0^1 g'(a + t(b-a))\, dt.$$

Now suppose that the formula holds for some $k \geq 1$. To see that this implies the $k+1$ case, it suffices to show that

$$\begin{aligned} \frac{(b-a)^k}{(k-1)!} \int_0^1 (1-t)^{k-1} g^{(k)}(a+t(b-a))\, dt = \\ \frac{g^{(k)}(a)}{k!}(b-a)^k + \frac{(b-a)^{k+1}}{k!} \int_0^1 (1-t)^k g^{(k+1)}(a+t(b-a))\, dt. \end{aligned} \tag{1.2}$$

To this end, observe that $\frac{1}{k}(1-t)^{k-1} = -\frac{d}{dt}(1-t)^k$. Integration by parts thus shows that the left hand side equals

$$\begin{aligned} &-\frac{(b-a)^k}{k!}(1-t)^k g^{(k)}(a+t(b-a))\Big|_{t=0}^{t=1} \\ &\qquad + \frac{(b-a)^k}{k!} \int_0^1 (1-t)^k \frac{d}{dt}\Big(g^{(k)}(a+t(b-a))\Big)\, dt. \end{aligned}$$

It is then verified that this equals the right hand side of (1.2). Indeed, the contribution of $t = 1$ to the first term vanishes and applying the chain rule as before to the integral yields the desired identity. □

Exercise. Show that the conclusion of Theorem 1.2 (i.e. identity (1.1)), holds in exactly the same way when when $[a, b]$ is replaced by $[b, a]$ for some $b < a$.

Exercise. Show that the remainder term in (1.1) (i.e. the integral) can be rewritten as

$$\frac{1}{(k-1)!} \int_a^b (b-s)^{k-1} g^{(k)}(s)\, ds.$$

2 Taylor's formula in higher dimensions

Let $\Omega \subset \mathbb{R}^n$ be an open set and $f : \Omega \to \mathbb{R}$. We say that f is k times continously differentiable, and denote $f \in \mathcal{C}^k(\Omega)$, if all the partial derivatives of order less than or equal to k exist and define continuous functions on Ω.

Given a dimension n, treated as fixed throughout any given discussion, we define a *multi-index* to be a vector $\alpha = (\alpha_1, \ldots, \alpha_n)$. In particular, any standard basis vector e_j in $\mathbb{R}^n$ defines a multi-index. We define the *order* of a multi-index to be

$$|\alpha| = \alpha_1 + \cdots + \alpha_n.$$

We further define the *factorial* of a multi-index as

$$\alpha! = \alpha_1!\, \alpha_2! \cdots \alpha_n!\,.$$

The purpose of multi-index notation is that it simplifies various formulas which arise when studying monomials in higher dimensions and partial derivatives in higher dimensions.

As a first example, consider a monomial in n-dimensions $x_1^{\alpha_1} \cdots x_n^{\alpha_n}$. Taking $\alpha = (\alpha_1, \ldots, \alpha_n)$, we may denote this monomial as

$$x^\alpha = x_1^{\alpha_1} \cdots x_n^{\alpha_n}.$$

As a second example, suppose $f \in \mathcal{C}^k(\Omega)$ as defined above. By Clairaut's theorem, if $l \leq k$, we know that any partial derivative of f of order l is independent of the order of differentiation. We may therefore unambiguously write

$$\partial^\alpha f(x) = \frac{\partial^{|\alpha|} f(x)}{\partial x_1^{\alpha_1} \cdots \partial x_n^{\alpha_n}}$$

but it is understood that any other order of partial differentiation yields the same function.

Theorem 2.1 (Taylor's formula in higher dimensions)**.** *Let $\Omega \subset \mathbb{R}^n$ be an open set and $f : \Omega \to \mathbb{R}$ be in $\mathcal{C}^k(\Omega)$. Then for any $x, x+y \in \Omega$ such that the line segment between them lies in Ω, we have that*

$$f(x+y) = \sum_{|\alpha|<k} \frac{y^\alpha}{\alpha!} \partial^\alpha f(x) + R_f^k(x,y)$$

where $R_f(x,y)$ is a remainder term that can be characterized in one of two ways:

$$R_f^k(x,y) = \sum_{|\alpha|=k} \frac{y^\alpha}{\alpha!} \partial^\alpha f(c) \qquad \textit{(classical form)} \tag{2.1}$$

$$R_f^k(x,y) = \sum_{|\alpha|=k} \frac{y^\alpha}{\alpha!} \int_0^1 k(1-t)^{k-1} \partial^\alpha f(x+ty)\, dt \qquad \textit{(integral form)} \tag{2.2}$$

where in the former case, c is some point on the line segment between x and $x+y$.

The crux of the proof of this theorem lies in the following lemma.

Lemma 2.2. *Let f, Ω be as in Theorem 2.1, and define $g : [0,1] \to \Omega$ by*

$$g(t) = f(x+ty).$$

Then $g \in C^k([0,1])$ *satisfies*

$$g^{(k)}(t) = k! \sum_{|\alpha|=k} \frac{y^\alpha}{\alpha!} \partial^\alpha f(x+ty). \tag{2.3}$$

To see that Theorem 2.1 follows as a corollary to the lemma, we apply Theorem 1.2 to g in order to obtain (2.2). The identity (2.1) follows from applying Theorem 1.1 to g, then realizing $x + t_0 y$ yields a point on the line segment between x and $x + ty$.

Proof of Lemma 2.2. That $g \in C^k([0,1])$ is a consequence of the chain rule, and is implicit in the argument below. We prove (2.3) by induction, noting that the $k = 0$ case is clear.

We thus assume the formula to be true for some k and show that it implies the $k+1$ case, when $k \geq 0$. If $f \in \mathcal{C}^{k+1}(\Omega)$, we have

$$\begin{aligned} g^{(k+1)}(t) = \frac{d}{dt}\left(g^{(k)}(t)\right) &= \frac{d}{dt}\left(k! \sum_{|\alpha|=k} \frac{y^\alpha}{\alpha!} \partial^\alpha f(x+ty)\right) \\ &= k! \sum_{|\alpha|=k} \frac{y^\alpha}{\alpha!} \frac{d}{dt}\left(\partial^\alpha f(x+ty)\right) \\ &= k! \sum_{|\alpha|=k} \frac{y^\alpha}{\alpha!} \left(\sum_{j=1}^{n} y_j \partial^{\alpha+e_j} f(x+ty)\right) \\ &= k! \sum_{|\alpha|=k} \sum_{j=1}^{n} \frac{y^{\alpha+e_j}}{\alpha!} \partial^{\alpha+e_j} f(x+ty), \end{aligned} \tag{2.4}$$

where the fourth identity follows from the chain rule. We now have to reindex the sum on the right as a sum over multi-indices of order $k+1$. As sets, observe that

$$\{\alpha + e_j : |\alpha| = k, 1 \leq j \leq n\} = \{\beta : |\beta| = k+1\}.$$

To see this, first note that every member of the set on the left is a member of the set on the right. Moreover, any member $\beta = (\beta_1, \ldots, \beta_n)$ of the set on the right satisfies $\beta_j > 0$ for some $j = 1, \ldots, n$ and hence can be written as $\alpha + e_j$ for some $|\alpha| = k$. So the sum in (2.4) can indeed be rewritten as a sum over the multi-indices β of order $k+1$. But this requires setting $\beta = \alpha + e_j$ above, meaning that the factor of $1/\alpha!$ is replaced by $1/(\beta - e_j)!$. The one catch is that this factorial is meaningful only when $\beta - e_j$ is a multi-index itself, that is, $\beta_j > 0$.

So given $|\beta| = k+1$, consider

$$B_\beta = \{\beta - e_j : \beta_j > 0\} = \{\alpha : |\alpha| = k \text{ and } \alpha = \beta - e_j \text{ for some } j = 1, \dots n\}.$$

Hence the sum in (2.4) can be rewritten as

$$k! \sum_{|\beta|=k+1} \sum_{\alpha \in B_\beta} \frac{y^\beta}{\alpha!} \partial^\beta f(x+ty) = k! \sum_{|\beta|=k+1} \frac{y^\beta}{\beta!} \partial^\beta f(x+ty) \sum_{\alpha \in B_\beta} \frac{\beta!}{\alpha!}.$$

The induction is finished if we can show $\sum_{\alpha \in B_\beta} \frac{\beta!}{\alpha!} = k+1$, which follows from

$$\sum_{\alpha \in B_\beta} \frac{\beta!}{\alpha!} = \sum_{j=1}^{n} \beta_j = |\beta| = k+1.$$

Indeed, either $\beta_j > 0$ meaning that $\beta = \alpha + e_j$ for some $\alpha \in B_\beta$, or $\beta_j = 0$, in which case the contribution is not counted in the first expression and contributes trivially in the second. $\square$

Remark 2.3. We conclude these notes by remarking that whenever $f \in \mathcal{C}^2(\Omega)$ and $c \in \Omega$, we have the identity

$$\sum_{|\alpha|=2} \frac{y^\alpha}{\alpha!} \partial^\alpha f(c) = \frac{1}{2} (\nabla^2 f(c) y) \cdot y$$

where $\nabla^2 f(c)$ denotes the Hessian matrix (and not the Laplacian as some texts do)

$$\nabla^2 f(c) = \begin{bmatrix} D^2_{11} f(c) & D^2_{12} f(c) & \cdots & D^2_{1n} f(c) \\ D^2_{21} f(c) & D^2_{22} f(c) & \cdots & D^2_{2n} f(c) \\ \vdots & \vdots & \cdots & \vdots \\ \vdots & \vdots & \cdots & \vdots \\ D^2_{n1} f(c) & D^2_{n2} f(c) & \cdots & D^2_{nn} f(c) \end{bmatrix}.$$

To see this, note that since $D^2_{ji} f(c) y_j y_i = D^2_{ij} f(c) y_i y_j$,

$$\begin{aligned} \frac{1}{2} (\nabla^2 f(c) y) \cdot y &= \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} D^2_{ij} f(c) y_i y_j \\ &= \sum_{i=1}^{n} \frac{1}{2} D^2_{ii} f(c) y_i y_i + \sum_{i=1}^{n} \sum_{j=i+1}^{n} D^2_{ij} f(c) y_i y_j. \end{aligned}$$

Indeed, the last expression here rewrites the second one as a sum over diagonal and off-diagonal terms. By the observed symmetry in i, j, the off-diagonal terms are counted twice in the second expression, but only once in the last expression, cancelling the factor of $1/2$. But the right hand side here can be rewritten as a sum over multi-indices $\sum_{|\alpha|=2} \frac{y^\alpha}{\alpha!} \partial^\alpha f(c)$, using that $1/\alpha! = 1$ when $\alpha = e_i + e_j$ for $i \neq j$ and $1/\alpha! = 1/2$ when $\alpha = 2e_i$ for some $i = 1, \dots, n$.