Mathematics Department Stanford University
Math 51H – Mean value theorem, Taylor's theorem and integrals

If f is a C^1 real valued function on an open set $U \subset \mathbb{R}^n$, we have for any x, h, i, with $\|h\|$ sufficiently small, that

$$f(x+he_i) - f(x) = hf'(x+\theta he_i)$$

for some $\theta \in (0,1)$. To see this, let

$$F(t) = f(x+the_i),\ t \in [0,1].$$

The chain rule shows that F is differentiable, with

$$F'(t) = h(D_i f)(x+the_i);$$

the composition of continuous functions being continuous, F is actually C^1. Thus, the mean value theorem gives

$$F(1) - F(0) = F'(\theta)$$

for some $\theta \in (0,1)$. Substituting in F, F' yields

$$f(x+he_i) - f(x) = h(D_i f)(x+\theta he_i). \tag{1}$$

Here is a different way of doing the same. Let's suppose $h > 0$; $h < 0$ is similar, and $h = 0$ is automatically true for any θ. Then let $G : [0,h] \to \mathbb{R}$ be defined by

$$G(s) = f(x+se_i),\ s \in [0,h].$$

Then by the chain rule

$$G'(s) = (D_i f)(x+se_i).$$

On the other hand by the mean value theorem

$$G(h) - G(0) = hG'(c),\ c \in (0,h).$$

Substituting in G

$$f(x+he_i) - f(x) = h(D_i f)(x+ce_i).$$

Letting $\theta = c/h$, so $\theta \in (0,1)$, we get

$$f(x+he_i) - f(x) = h(D_i f)(x+\theta he_i)$$

again.

Let's also write this out using the fundamental theorem of calculus: if ϕ is C^1 on $[a,b]$ then

$$\phi(b) - \phi(a) = \int_a^b \phi'(s)\,ds.$$

Note that a component-by-component check shows that the fundamental theorem of calculus is also valid for $\mathbb{R}^m$-valued f, so we allow such f from now on. Applying this for F, we get

$$f(x+he_i) - f(x) = F(1) - F(0) = \int_0^1 F'(t)\,dt = \int_0^1 h(D_i f)(x+the_i)\,dt,$$

i.e.

$$f(x+he_i) - f(x) = h\int_0^1 (D_i f)(x+the_i)\,dt.$$

So what we have in comparison with (1) is that $(D_i f)(x+\theta he_i)$ is replaced by $\int_0^1 (D_i f)(x+the_i)\,dt$.

Now let us work with he_i replaced by a vector $h \in \mathbb{R}^n$. For this purpose we should make sure that the line segment between x and $x+h$ is contained in the domain of definition of F, i.e. U; a convenient way of phrasing this is to make the stronger assumption $B_\rho(x) \subset U$, and $\|h\| < \rho$. So let $F(t) = f(x+th)$, $t \in [0,1]$. By the fundamental theorem of calculus,

$$F(1) - F(0) = \int_0^1 F'(t)\,dt,$$

and by the chain rule

$$F'(t) = \sum_{i=1}^n h_i (D_i f)(x+th).$$

Substituting in,

$$f(x+h) - f(x) = \int_0^1 \sum_{i=1}^n h_i (D_i f)(x+th)\,dt = \sum_{i=1}^n h_i \int_0^1 (D_i f)(x+th)\,dt.$$

Notice that writing $(D_i f)(x+th) = (D_i f)(x) + ((D_i f)(x+th) - (D_i f)(x))$ gives

$$f(x+h) - f(x) = \sum_{i=1}^n h_i (D_i f)(x) + \sum_{i=1}^n h_i \int_0^1 ((D_i f)(x+th) - (D_i f)(x))\,dt$$

which is a more precise remainder term than in the definition of differentiability (using continuity of the derivative) since we have an explicit error term. Indeed we have that give $\varepsilon > 0$ there is $\delta > 0$ such that $\|h\| < \delta$ implies $\|(D_i f)(x+th) - (D_i f)(x)\| < \varepsilon$ for all i (indeed, for each $\varepsilon > 0$ a $\delta_i > 0$ exists for this statement for $D_i f$, let $\delta > 0$ be the minimum of these finitely many positive numbers), so for $\|h\| < \delta$,

$$\Big\| \sum_{i=1}^n h_i \int_0^1 ((D_i f)(x+th) - (D_i f)(x))\,dt \Big\| \le \sum_{j=1}^n |h_i| \varepsilon \le n\varepsilon \|h\|.$$

Let's go now one order further in this expansion. Namely, we write

$$F(1) - F(0) = \int_0^1 F'(t)\,dt = \int_0^1 1 \cdot F'(t)\,dt = (t-1)F'(t)|_0^1 - \int_0^1 (t-1)F''(t)\,dt,$$

where we integrated by parts, using $t-1$ as an antiderivative of 1; we are making this choice so that the $t = 1$ boundary term cancels. Thus,

$$F(1) = F(0) + F'(0) + \int_0^1 (1-t)F''(t)\,dt. \tag{2}$$

Now, as above $F'(t) = \sum_{i=1}^n h_i (D_i f)(x+th)$, so applying the chain rule again,

$$F''(t) = \sum_{i=1}^n h_i \sum_{j=1}^n h_j (D_j D_i f)(x+th) = \sum_{i=1}^n \sum_{j=1}^n h_i h_j (D_j D_i f)(x+th)$$

Substitution into (2) yields

$$f(x+h) = f(x) + \sum_{i=1}^n h_i (D_i f)(x) + \sum_{i,j=1}^n h_i h_j \int_0^1 (1-t)(D_i D_j f)(x+th)\,dt. \tag{3}$$

This is Taylor's theorem with second order integral remainder. Rewriting as before, using $\int_0^1 (1-t)\,dt = -\frac{1}{2}(1-t)^2|_0^1 = \frac{1}{2}$,

$$f(x+h) = f(x) + \sum_{i=1}^n h_i (D_i f)(x) + \frac{1}{2}\sum_{i,j=1}^n h_i h_j (D_i D_j f)(x) \\ + \sum_{i,j=1}^n h_i h_j \int_0^1 (1-t)[(D_i D_j f)(x+th) - (D_i D_j f)(x)]\,dt.$$

Again, using the continuity of the partials, denoting the last term by $E(x,h)$, given $\varepsilon > 0$ there is $\delta > 0$ such that $\|h\| < \delta$ implies $\|D_iD_jf(x+h) - D_iD_jf(x)\| < \varepsilon$ (for each i, j, there is a $\delta_{ij} > 0$ with this property, and then take δ as the minimum of these), and then

$$\begin{aligned}\|E(x,h)\| &\leq \sum_{i,j=1}^{n} \|h\|^2 \int_0^1 (1-t)\|(D_iD_jf)(x+th) - (D_iD_jf)(x)\|\,dt \\ &\leq \sum_{i,j=1}^{n} \|h\|^2 \int_0^1 (1-t)\varepsilon\,dt = n^2\|h\|^2 \frac{\varepsilon}{2}(1-t)^2|_0^1 = n^2\frac{\varepsilon}{2}\|h\|^2 < \|h\|^2 n^2 \varepsilon\end{aligned}$$

follows, so given $\varepsilon' > 0$, choosing $\varepsilon = \varepsilon'/n^2$, this $\delta > 0$ yields

$$\lim_{h\to 0} \|h\|^{-2}\|E(x,h)\| = 0.$$

Assuming F is C^k, proceeding inductively, we get

$$F(1) = \sum_{j=0}^{k-1} \frac{1}{j!} D^j F(0) + \frac{1}{(k-1)!} \int_0^1 (1-t)^{k-1} D^k F(t)\,dt \tag{4}$$

from (2), as is straightforward to check. Also, using the chain rule inductively

$$D^\ell F(t) = \sum_{i_1=1}^{n} \cdots \sum_{i_\ell=1}^{n} h_{i_1} \ldots h_{i_\ell} (D_{i_1} \ldots D_{i_\ell} f)(x+th).$$

Hence we conclude Taylor's theorem with an integral remainder formula as above:

Theorem 1 *If $x \in \mathbb{R}^n$, $\rho > 0$, $f : B_\rho(x) \to \mathbb{R}^m$ is C^k then for $\|h\| < \rho$,*

$$\begin{aligned}f(x+h) =& \sum_{j=0}^{k-1} \sum_{i_1=1}^{n} \cdots \sum_{i_j=1}^{n} \frac{1}{j!} h_{i_1} \ldots h_{i_j} (D_{i_1} \ldots D_{i_j} f)(x) \\ &+ \sum_{i_1=1}^{n} \cdots \sum_{i_k=1}^{n} \frac{1}{(k-1)!} h_{i_1} \ldots h_{i_k} \int_0^1 (1-t)^{k-1} (D_{i_1} \ldots D_{i_k} f)(x+th)\,dt.\end{aligned}$$

Arguing as above using the continuity of the partial derivatives, if we write

$$(D_{i_1} \ldots D_{i_k} f)(x+th) = (D_{i_1} \ldots D_{i_k} f)(x) + ((D_{i_1} \ldots D_{i_k} f)(x+th) - (D_{i_1} \ldots D_{i_k} f)(x)),$$

using $\int_0^1 (1-t)^{k-1}\,dt = -\frac{1}{k}(1-t)^k|_0^1 = \frac{1}{k}$, we get

$$\begin{aligned}f(x+h) =& \sum_{j=0}^{k} \sum_{i_1=1}^{n} \cdots \sum_{i_j=1}^{n} \frac{1}{j!} h_{i_1} \ldots h_{i_j} (D_{i_1} \ldots D_{i_j} f)(x) + E_k(x,h), \\ E_k(x,h) =& \sum_{i_1=1}^{n} \cdots \sum_{i_k=1}^{n} \frac{1}{(k-1)!} h_{i_1} \ldots h_{i_k} \int_0^1 (1-t)^{k-1} \Big((D_{i_1} \ldots D_{i_k} f)(x+th) - (D_{i_1} \ldots D_{i_k} f)(x+th)\Big)\,dt,\end{aligned}$$

and

$$\lim_{h\to 0} \|h\|^{-k}\|E_k(x,h)\| = 0.$$

Note: if$f$ is real valued, one can get an alternative version of Taylor's theorem: there exists $\theta \in (0,1)$ such that

$$f(x+h) = \sum_{j=0}^{k-1} \sum_{i_1=1}^{n} \cdots \sum_{i_j=1}^{n} \frac{1}{j!} h_{i_1} \dots h_{i_j} (D_{i_1} \dots D_{i_j} f)(x) + \sum_{i_1=1}^{n} \cdots \sum_{i_k=1}^{n} \frac{1}{k!} h_{i_1} \dots h_{i_k} (D_{i_1} \dots D_{i_k} f)(x+\theta h).$$

Notice that this is just the statement that there exists $\theta \in (0,1)$ such that

$$\frac{1}{k} \sum_{i_1=1}^{n} \cdots \sum_{i_k=1}^{n} h_{i_1} \dots h_{i_k} (D_{i_1} \dots D_{i_k} f)(x+\theta h) = \sum_{i_1=1}^{n} \cdots \sum_{i_k=1}^{n} h_{i_1} \dots h_{i_k} \int_0^1 (1-t)^{k-1} (D_{i_1} \dots D_{i_k} f)(x+th)\, dt,$$

i.e. if we let

$$\phi(t) = \sum_{i_1=1}^{n} \cdots \sum_{i_k=1}^{n} h_{i_1} \dots h_{i_k} (D_{i_1} \dots D_{i_k} f)(x+th)$$

then there exists $\theta \in (0,1)$ such that

$$\phi(\theta) = \int_0^1 k(1-t)^{k-1} \phi(t)\, dt.$$

To see this, let $I = \int_0^1 k(1-t)^{k-1}\phi(t)\, dt$, and note that there must exist $t_1, t_2 \in [0,1]$ such that $\phi(t_1) \leq I$ and $\phi(t_2) \geq I$ for if say $\phi(t) > I$ for all $t \in [0,1]$ then

$$I = \int_0^1 k(1-t)^{k-1}\phi(t)\, dt > \int_0^1 k(1-t)^{k-1} I\, dt = -(1-t)^k|_0^1 I = I,$$

which is a contradiction. Having found such t_1, t_2, if ϕ at either one of them is I, we are done, otherwise suppose $t_1 < t_2$ (with $t_1 > t_2$ analogous), and use the intermediate value theorem: a continuous real-valued function on $[t_1, t_2]$ with $\phi(t_1) < \phi(t_2)$ attains all values in $[\phi(t_1), \phi(t_2)]$, in particular attains the value I in the interval (t_1, t_2); proving the claim, and completing the proof of the mean value form of Taylor's theorem.