Single and Multivariate Taylor Series

Just as polygons may be viewed as building blocks of 2-D geometry, polynomials can be used in the manner of Lego pieces to construct many differentiable functions. In particular, according to the Weierstrass approximation theorem, any continuous function $f:[a,b]\to R$ can be uniformly approximated to any degree of accuracy by a polynomial function. When this function is also infinitely differentiable and satisfies $\left|f^{(n)}(x)\right|\le M$ for some number M and for all x in the closed interval [a, b], we may use Taylor polynomials to achieve the desired approximation.[1]

Single-Variable Taylor Theorem

Suppose $f:[-a,a]\to R$ is continuously differentiable n + 1 times. We would like to find a polynomial of degree n that "looks like" the function f near x = 0.
It seems reasonable that this polynomial, p, will resemble f if

$$f(0)=p(0),\ f'(0)=p'(0),\ f''(0)=p''(0),...,\ f^{(n)}(0)=p^{(n)}(0).$$

Thus, if $p(x)=a_0+a_1x+a_2x^2+a_3x^3+...+a_nx^n$, we would expect to find that

$$f(0)=a_0=p(0),\ f'(0)=a_1=p'(0),\ f''(0)=2a_2=p''(0),\dots,\ f^{(n)}(0)=n!a_n=p^{(n)}(0).$$

In particular,

$$a_0=f(0),\ a_1=\frac{f'(0)}{1!},\ a_2=\frac{f''(0)}{2!},\dots,\ a_n=\frac{f^{(n)}(0)}{n!}.$$

It therefore appears that a polynomial of the form

$$p(x)=f(0)+\frac{f'(0)}{1!}x+\frac{f''(0)}{2!}x^2+\frac{f'''(0)}{3!}x^3+...+\frac{f^{(n)}(0)}{n!}x^n$$

is a good approximation to f near 0.
Similarly, if $c\in[-a,a]$, the polynomial

$$q(x)=f(c)+\frac{f'(c)}{1!}(x-c)+\frac{f''(c)}{2!}(x-c)^2+\frac{f'''(c)}{3!}(x-c)^3+...+\frac{f^{(n)}(c)}{n!}(x-c)^n$$

appears to be a good approximation of f near c.
But, intuition aside, how good is this approximation really? The answer is the content of Taylor's theorem:

[1] Not all continuous or even infinitely differentiable functions of the form $f:U\subseteq R\to R$ can be approximated by polynomials. Weierstrass theorem requires U to be compact (i.e. closed and bounded).

Theorem: Let $f: R \to R$ be n + 1 times continuously differentiable. Let $p(x) = f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \frac{f'''(a)}{3!}(x-a)^3 + ... + \frac{f^{(n)}(a)}{n!}(x-a)^n$ be the nth degree Taylor polynomial of f centered at x = a. Then, for $b \in R$, the error $E_n(b) = f(b) - p(b)$ is given by $E_n(b) = \int_a^b \frac{(b-t)^n}{n!} f^{(n+1)}(t)dt$. Furthermore, $|E_n(b)| \le \frac{(b-a)^{n+1}}{n!} M$, where $M = \max_{t\in[a,b]} |f^{(n+1)}(t)|$.

Proof: We will show that $f(b) = \sum_{i=0}^{n} \frac{f^{(i)}(a)}{i!}(b-a)^i + E_n(b)$. By the fundamental theorem of calculus, $f(b) = f(a) = \int_a^b f'(t)dt$. Using integration by parts with $u = f'(t)$ and $dv = 1$, we see that

$$\int_a^b f'(t)dt = uv\Big|_a^b - \int_a^b vdu = -(b-t)f'(t)\Big|_a^b - \int_a^b -(b-t)f''(t)dt \tag{1}$$

where we used the fact that $-(b-t)$ is an anti-derivative of 1.
Or, after simplifying (1),

$$\int_a^b f'(t)dt = f'(a)(b-a) + \int_a^b (b-t)f''(t)dt \tag{2}$$

Integrating by parts again, we get

$$\int_a^b f'(t)dt = f'(a)(b-a) + \frac{f''(a)}{2}(b-a)^2 + \int_a^b \frac{(b-t)^2}{2} f'''(t)dt \tag{3}$$

Continuing in this fashion, we see that

$$\int_a^b f'(t)dt = \sum_{i=1}^{n} \frac{f^{(i)}(a)}{i!}(b-a)^i + \int_a^b \frac{(b-t)^{(n)}}{n!} f^{(n+1)}(t)dt \tag{4}$$

Thus,

$$f(b) = \sum_{i=0}^{n} \frac{f^{(i)}(a)}{i!}(b-a)^i + \int_a^b \frac{(b-t)^{(n)}}{n!} f^{(n+1)}(t)dt. \tag{5}$$

It follows that $E_n(b) = f(b) - p(b) = \int_a^b \frac{(b-t)^{(n)}}{n!} f^{(n+1)}(t)dt$ as desired.
Finally, to estimate the error $E_n(b)$, observe that for $t \in [a,b]$, $|b-t| \le |b-a|$.

Thus,

$$\left|E_n(b)\right| = \left|\int_a^b \frac{(b-t)^n}{n!} f^{(n+1)}(t)dt\right| \le \int_a^b \frac{\left|b-t\right|^n}{n!}\left|f^{(n+1)}(t)\right|dt \le \int_a^b \frac{\left|b-a\right|^n}{n!} Mdt = \frac{\left|b-a\right|^{n+1}}{n!} M \qquad \blacktriangledown$$

As a consequence of Taylor's theorem, it follows that for x near a,
$f(x) = \sum_{i=0}^{n} \frac{f^{(i)}(a)}{i!}(x-a)^i + E_n(x)$, where $E_n(x) = \int_a^x \frac{(x-t)^n}{n!} f^{(n+1)}(t)dt$ is a continuous function of x (Why?).

Example: Let $f(x) = e^{x-1}$. If f is approximated by a 5th degree Taylor polynomial centered at x = 1, estimate the error when this polynomial is used to approximate the value of f(1.5).

Solution: We wish to estimate $\left|E_5(1.5)\right|$. By Taylor's theorem, we know that

$$\left|E_5(1.5)\right| = \left|\int_1^{1.5} \frac{(1.5-t)^5}{5!} f^{(6)}(t)dt\right| \le \frac{\left|1.5-1\right|^6}{5!} \max_{t\in[1,1.5]}\left|f^{(6)}(t)\right| = \frac{1}{2^6 5!} e^{1.5} < \frac{3^2}{2^6 5!} < 0.0036.$$

In other words, the estimation is accurate up to two decimal places.

Multivariate Taylor Theorem

Suppose that $f : R^m \to R$ has continuous mixed partials up to the (n + 1)th order. We would like to find a polynomial in m variables that is a "good" approximation to f near x = a $\in R^m$. We can reduce this problem to the single-variable case by defining $g : R \to R$ by $g(t) = f(a + t(x-a))$. That is, for any fixed x, we can parameterize a path $l(t) = a + t(x-a)$ and compose it with the function f. Since g(0) = f(a) and g(1) = f(x), the nth order Taylor approximation of the multivariate function f(x) about x = a must be the same as the nth degree Taylor approximation of g(1) about t = 0. In particular,

$$f(x) = g(1) = g(0) + \frac{g'(0)}{1!}1 + \frac{g''(0)}{2!}1^2 + \dots + \frac{g^{(n)}(0)}{n!}1^n + E_n(1).$$

Where $E_n(1) = \int_0^1 \frac{(t-1)^n}{n!} g^{(n+1)}(t)dt$.

Therefore, we'll succeed in our endeavor to represent f as an nth order Taylor polynomial as soon as we are able to compute $g^{(k)}(0)$ in terms of f. Let's see if we can deduce a pattern:
By the chain rule,

$$g'(t) = \frac{d}{dt}\left(f(a + t(x-a))\right) = (x_1 - a_1)\frac{\partial f}{\partial x_1}(a + t(x-a)) + \dots + (x_m - a_m)\frac{\partial f}{\partial x_m}(a + t(x-a)). \quad (1)$$

Where (1) can be written more compactly as

$$(x-a)\bullet\nabla f \tag{2}$$

Where ∇f is evaluated at $l(t)=a+t(x-a)$.

Similarly,

$$g''(t)=\frac{d}{dt}\left[(x_1-a_1)\frac{\partial f}{\partial x_1}(a+t(x-a))+\ldots+(x_m-a_m)\frac{\partial f}{\partial x_m}(a+t(x-a))\right]=$$

$$(x_1-a_1)\frac{d}{dt}\left[\frac{\partial f}{\partial x_1}(a+t(x-a))\right]+\ldots+(x_m-a_m)\frac{d}{dt}\left[\frac{\partial f}{\partial x_m}(a+t(x-a))\right]. \tag{3}$$

Observe that each of the $\frac{\partial f}{\partial x_i}(a+t(x-a))$ is a function of one variable t that comes about as a result of composition of a multivariate function of m variables (namely $\frac{\partial f}{\partial x_i}(x_1,\ldots,x_m)$) and a path function (namely $l(t)=a+t(x-a)$). Thus, $\frac{d}{dt}\left[\frac{\partial f}{\partial x_i}(a+t(x-a))\right]$ can be written in the same form as (2). In particular,

$$\frac{d}{dt}\left[\frac{\partial f}{\partial x_i}(a+t(x-a))\right]=(x-a)\bullet\nabla\frac{\partial f}{\partial x_i} \tag{4}$$

Where, again, $\nabla\frac{\partial f}{\partial x_i}$ is evaluated at $l(t)=a+t(x-a)$

Therefore,

$$g''(t)=\sum_{i=1}^{m}(x_i-a_i)\left[(x-a)\bullet\nabla\frac{\partial f}{\partial x_i}\right]. \tag{5}$$

Or,

$$g''(t)=\sum_{i=1}^{m}(x_i-a_i)\left[\sum_{j=1}^{m}(x_j-a_j)\frac{\partial^2 f}{\partial x_j\partial x_i}(a+t(x-a))\right]=$$

$$\sum_{i=1}^{m}\sum_{j=1}^{m}(x_i-a_i)(x_j-a_j)\frac{\partial^2 f}{\partial x_j\partial x_i}(a+t(x-a)) \tag{6}$$

By labeling $l(t)=a+t(x-a)$ with u in expression (6), we get

$$\sum_{i=1}^{m}\sum_{j=1}^{m}(x_i - a_i)(x_j - a_j)\frac{\partial^2 f}{\partial x_j \partial x_i}(u) \tag{7}$$

Notice that (7) resembles the square of a sum of m terms (i.e. $\left(\sum_{i=1}^{m} A_i\right)^2 = \sum_{i=1}^{m}\sum_{j=1}^{m} A_j A_i$). To make further use of this analogy, denote by C_m^∞ the space of all functions on m variables that have continuous mixed partials of every order. Define $T : C_m^\infty \to C_m^\infty$ by $T(q) = (x-a) \bullet \nabla q$. In other words, since we regard x as fixed, we may think of q as a function of some variable u = $(u_1, ..., u_m)$. Observe that T is a linear transformation from the space C_m^∞ to itself.[2] Let $\varphi : C_m^\infty \to R$ be defined by $\varphi(q) = q(a)$. Since, by (2), $g'(t) = (x-a) \bullet \nabla f$, it follows that $g'(0) = (x-a) \bullet \nabla f(a) = \varphi(T(f))$.[3] Similarly, by (5),

$g''(0) = \sum_{i=1}^{m}(x_i - a_i)\left[(x-a) \bullet \nabla\left(\frac{\partial f}{\partial x_i}(a)\right)\right] = (x-a) \bullet \left[\nabla\left(\sum_{i=1}^{m}(x_i - a_i)\frac{\partial f}{\partial x_i}(a)\right)\right] =$

$(x-a) \bullet \nabla\big((x-a) \bullet \nabla f(a)\big) = \varphi(T(T(f))) = \varphi(T^2(f))$. More generally, it can be shown by induction that $g^{(k)}(0) = \varphi(T^k(f))$.

It follows that,

$$f(x) = \sum_{k=0}^{n}\frac{\varphi(T^k(f))}{k!} + E_n(1) \tag{8}$$

We will have more to say about the error $E_n(1)$ later on. For now, let us try to re-express (8) free from the functions φ and T. With a slight abuse of notation, we may represent T by

$$(x-a) \bullet \nabla = (x_1 - a_1)\frac{\partial}{\partial u_1} + ... + (x_m - a_m)\frac{\partial}{\partial u_m}. \tag{9}$$

With this notation, T^k becomes

$$\big((x-a) \bullet \nabla\big)^k = \left((x_1 - a_1)\frac{\partial}{\partial u_1} + ... + (x_m - a_m)\frac{\partial}{\partial u_m}\right)^k \tag{10}$$

---

[2] To verify this, pick any two functions f, h in C_m^∞ and any real scalar c. Then

$T(q+ch) = (x-a) \bullet \nabla(q+ch) = (x-a) \bullet (\nabla q + c\nabla h) = (x-a) \bullet \nabla q + c(x-a) \bullet \nabla h =$
$T(q) + cT(h)$.

[3] By hypothesis, the function f with which we are currently dealing has continuous mixed partials up to the (n+1)th order. Therefore, strictly speaking, f is not in the domain of T. However, we will pay a heavy price if we insist on being too rigorous here.

Where $T^k(f(u)) = \left((x-a)\bullet\nabla\right)^k f(u) = \left((x_1 - a_1)\frac{\partial}{\partial u_1} + ... + (x_m - a_m)\frac{\partial}{\partial u_m} \right)^k f(u)$ (11)

In other words, the nth degree Taylor polynomial of f may be written simply as

$$P_n(x_1,....,x_m) = \sum_{k=0}^{n} \frac{((x-a)\bullet\nabla)^k f(u)}{k!}\Bigg|_{u=a} \quad (12)$$

To finally apply the analogy that exists between function composition and scalar multiplication, notice that we can rewrite (11) as

$$T^k(f) = \sum_{i_1=1}^{m} ... \sum_{i_k=1}^{m} (x_{i_1} - a_{i_1})...(x_{i_k} - a_{i_k}) \frac{\partial^k f}{\partial u_{i_k} ... \partial u_{i_1}} (u_1,...,u_m) \quad (13)$$

By the equality of mixed partials, (13) behaves just like an ordinary kth power of a sum of m real numbers.

Computing Powers of T

In this section, we would like to fully exploit the analogy between the powers of the linear transformation T and powers of ordinary sums of numbers. If $f : R^m \to R$ has all of its mixed partials up to the (n + 1)th order defined in an open ball $B_\delta(a)$ and if all of the (n +1)th order mixed partials are continuous at x = a, how might we compute $((x-a)\bullet\nabla)^k(f)$? To answer this question, start with the case m = 2. For simplicity of notation, we'll denote $x - a = (x_1 - a_1, x_2 - a_2)$ by (h_1, h_2). With this notation, $(x-a)\bullet\nabla$ becomes $(h_1\frac{\partial}{\partial u_1} + h_2\frac{\partial}{\partial u_2})$.
Using your knowledge about the powers of sums, observe that

$$(h_1\frac{\partial}{\partial u_1} + h_2\frac{\partial}{\partial u_2})(f) = h_1\frac{\partial f}{\partial u_1} + h_2\frac{\partial f}{\partial u_2}; \quad (1)$$

$$(h_1\frac{\partial}{\partial u_1} + h_2\frac{\partial}{\partial u_2})^2(f) = h_1^2\frac{\partial^2 f}{\partial u_1^2} + 2h_1h_2\frac{\partial^2 f}{\partial u_1\partial u_2} + h_2^2\frac{\partial^2 f}{\partial u_2^2}; \quad (2)$$

$$(h_1\frac{\partial}{\partial u_1} + h_2\frac{\partial}{\partial u_2})^3(f) = h_1^3\frac{\partial^3 f}{\partial u_1^3} + 3h_1^2h_2\frac{\partial^3 f}{\partial u_1^2\partial u_2} + 3h_1h_2^2\frac{\partial^3 f}{\partial u_1\partial u_2^2} + h_2^3\frac{\partial^3 f}{\partial u_2^3}. \quad (3)$$

For those of you who know a little bit of combinatorics,

$$(h_1 \frac{\partial}{\partial u_1} + h_2 \frac{\partial}{\partial u_2})^k (f) = \sum_{i=0}^{k} \binom{k}{i} h_1^i h_2^{k-i} \frac{\partial^k f}{\partial u_1{}^i \partial u_2{}^{k-i}} \qquad (4)$$

Where $\binom{k}{i} = \frac{k!}{i!(k-i)!}$.

Similarly, for $m > 2$, $(x-a) \bullet \nabla = (h_1 \frac{\partial}{\partial u_1} + h_2 \frac{\partial}{\partial u_2} + ... + h_m \frac{\partial}{\partial u_m})$.

By the multinomial theorem,

$$((x-a) \bullet \nabla)^k (f) = \sum_{\substack{i_1, i_2, ..., i_m; \\ i_1 + i_2 + ... + i_m = k}} \binom{k}{i_1 \quad i_2 \quad \cdots \quad i_m} h_1^{i_1} h_2^{i_2} ... h_m^{i_m} \frac{\partial^k f}{\partial u_1{}^{i_1} \partial u_2{}^{i_2} ... \partial u_m{}^{i_m}} \qquad (5)$$

Where the sum is taken over all non-negative integer valued vectors $(i_1, i_2, ..., i_m)$, the sum of whose coordinates equals to k.

Finally, recall that the kth term in the nth order multivariate Taylor expansion is $\frac{\varphi(T^k(f))}{k!}$. In other words,

$$\frac{\varphi(T^k(f))}{k!} = \frac{1}{k!} \sum_{\substack{i_1, i_2, ..., i_m; \\ i_1 + i_2 + ... + i_m = k}} \binom{k}{i_1 \quad i_2 \quad \cdots \quad i_m} h_1^{i_1} h_2^{i_2} ... h_m^{i_m} \frac{\partial^k f}{\partial u_1{}^{i_1} \partial u_2{}^{i_2} ... \partial u_m{}^{i_m}} (a_1, a_2, ..., a_n)$$

Now we are ready to solve a few examples:

Example: Let $f(x, y) = e^{x+y}$. Compute the second-order Taylor polynomial $P_2(x, y)$, about the point (x, y) = (0, 0).

Solution: First, re-define f in terms of a new variable: $f(u, v) = e^{u+v}$. We wish to compute $\varphi(T^0(f)) + \varphi(T(f)) + \frac{\varphi(T^2(f))}{2}$, where $\varphi(q(u, v)) = q(0,0)$. Now, $\varphi(T^0(f)) = f(0,0) = e^0 = 1$;

$\varphi(T(f)) = (x, y) \bullet \nabla f(0,0) = x \frac{\partial f}{\partial u}(0,0) + y \frac{\partial f}{\partial v}(0,0) = x + y$;

$\varphi(T^2(f)) = (x \frac{\partial}{\partial u} + y \frac{\partial}{\partial v})^2 (f(0,0)) = x^2 \frac{\partial^2 f}{\partial u^2}(0,0) + 2xy \frac{\partial^2 f}{\partial u \partial v}(0,0) + y^2 \frac{\partial^2 f}{\partial v^2}(0,0) =$
$x^2 + 2xy + y^2$.

Thus, $P_2(x, y) = 1 + x + y + \frac{x^2 + 2xy + y^2}{2} = 1 + (x + y) + \frac{(x + y)^2}{2}$ (Looks familiar?)

Example: Let $f(x,y) = Sin(xy)$. Compute the second-order Taylor polynomial $P_2(x,y)$ about (x, y) = $(1, \pi/2)$.

Solution: Just like you did before, re-define the function in terms of a new variable: $f(u,v) = Sin(uv)$. Before we go on, it would be helpful to have all the mixed partials up to the second order computed and evaluated at the point $(1, \pi/2)$:

0-order partials

$f(1,\pi/2) = Sin(\pi/2) = 1$

1-order partials

$\frac{\partial f}{\partial u}(u,v) = vCos(uv)$ so $\frac{\partial f}{\partial u}(1,\pi/2) = \frac{\pi}{2}Cos(\pi/2) = 0$

$\frac{\partial f}{\partial v}(u,v) = uCos(uv)$ so $\frac{\partial f}{\partial v}(1,\pi/2) = Cos(\pi/2) = 0$

2-order partials

$\frac{\partial^2 f}{\partial u^2}(u,v) = -v^2 Sin(uv)$ so $\frac{\partial^2 f}{\partial u^2}(1,\pi/2) = -\left(\frac{\pi}{2}\right)^2 Sin(\pi/2) = -\left(\frac{\pi}{2}\right)^2$

$\frac{\partial^2 f}{\partial v \partial u}(u,v) = Cos(uv) - uvSin(uv)$ so $\frac{\partial^2 f}{\partial v \partial u}(1,\pi/2) = Cos(\pi/2) - \left(\frac{\pi}{2}\right)Sin(\pi/2) = -\frac{\pi}{2}$

$\frac{\partial^2 f}{\partial u \partial v}(u,v) = Cos(uv) - uvSin(uv)$ so $\frac{\partial^2 f}{\partial u \partial v}(1,\pi/2) = Cos(\pi/2) - \left(\frac{\pi}{2}\right)Sin(\pi/2) = -\frac{\pi}{2}$

$\frac{\partial^2 f}{\partial v^2}(u,v) = -u^2 Sin(uv)$ so $\frac{\partial^2 f}{\partial v^2}(1,\pi/2) = -Sin(\pi/2) = -1$

Now, $\varphi(T^0(f)) = f(1,\pi/2) = 1$;

$\varphi(T(f)) = (x-1, y-\pi/2) \bullet \nabla f(1,\pi/2) = (x-1)\frac{\partial f}{\partial u}(1,\pi/2) + (y-\pi/2)\frac{\partial f}{\partial v}(1,\pi/2) =$
$(x-1)\cdot 0 + (y-\pi/2)\cdot 0 = 0$;

$\varphi(T^2(f)) = (x-1)^2 \frac{\partial^2 f}{\partial u^2}(1,\pi/2) + 2(x-1)(y-\pi/2)\frac{\partial^2 f}{\partial u \partial v}(1,\pi/2) + (y-\pi/2)^2 \frac{\partial^2 f}{\partial v^2}(1,\pi/2) =$
$-\left(\frac{\pi}{2}\right)^2 (x-1)^2 - 2\frac{\pi}{2}(x-1)(y-\pi/2) - (y-\pi/2)^2$

Thus, $P_2(x,y) = 1 - \frac{\left(\frac{\pi}{2}\right)^2 (x-1)^2 + \pi(x-1)(y-\pi/2) + (y-\pi/2)^2}{2}$

Another form of the error function

Recall that for a function $g: R \to R$ that is n + 1 times continuously differentiable, $E_n(b) = \int_a^b \frac{(b-t)^n}{n!} g^{(n+1)}(t)dt$. We will now show that $E_n(b) = \frac{g^{(n+1)}(c)}{(n+1)!}$ for some c $\in$ (a, b). This form of the error will be useful in the next chapter.

Lemma (Generalized Mean-Value Theorem): Let $f, g: [a,b] \subset R \to R$ be continuous on [a, b] and differentiable on (a, b). If, in addition, g'(t) $\neq$ 0 on (a, b) then $\frac{f(b)-f(a)}{g(b)-g(a)} = \frac{f'(c)}{g'(c)}$ for some scalar c in (a, b).

Proof: Let $H: [a,b] \subset R \to R$ be defined by H(t) = (f(b) – f(a))g(t) – (g(b) – g(a))f(t). Then H is continuous on [a, b] and differentiable on (a, b). Thus, by the mean-value theorem, there exists a scalar c in (a, b) such that $H'(c) = \frac{H(b)-H(a)}{b-a}$. But H(a) = H(b) so H'(c) = 0. In particular, 0 = H'(c) = (f(b) – f(a))g'(c) – (g(b) – g(a))f'(c), from which the statement of the lemma follows.

▼

Theorem: Let $g: R \to R$ be n + 1 times continuously differentiable in an open interval I with t = a as its center. Then the nth order Taylor approximation error at b $\in$ I, $E_n(b)$, may be represented in the form $\frac{g^{(n+1)}(c)}{(n+1)!}(b-a)^{n+1}$ for some c $\in$ (a, b) $\in$ I

Proof: Let $q(t) = (t-a)^{n+1}$. Then q is continuous on [a, b] and differentiable on (a, b). Furthermore, $q, q', q'', ..., q^{(n)}$ are never 0 on the interval (a, b). If $p_n(t)$ is the nth degree Taylor polynomial of g, then $E_n(t) = g(t) - p_n(t)$, is continuous on [a, b] and differentiable on (a, b) (as are $E_n', E_n'', ..., E_n^{(n)}$). Notice that $q(a) = q'(a) = ... = q^{(n)}(a) = 0$. Similarly, $E_n(a) = E_n'(a) = ... = E_n^{(n)}(a) = 0$. Therefore, by the generalized mean-value theorem,

$$\frac{E_n(b)}{q(b)} = \frac{E_n(b) - E_n(a)}{q(b) - q(a)} = \frac{E_n'(c_1)}{q'(c_1)} \quad (1)$$

Where a < c_1 < b.
Applying the generalized mean-value theorem again, we see that

$$\frac{E_n'(c_1)}{q'(c_1)} = \frac{E_n'(c_1) - E_n'(a)}{q'(c_1) - q'(a)} = \frac{E_n''(c_2)}{q''(c_2)} \quad (2)$$

Where a < c_2 < c_1 < b.
Continuing in this fashion, we see that

$$\frac{E_n(b)}{q(b)} = \frac{E_n^{(n+1)}(c_{n+1})}{q^{(n+1)}(c_{n+1})} = \frac{g^{(n+1)}(c_{n+1})}{(n+1)!} \tag{3}$$

Where a < c_{n+1} <…< c_2 < c_1 < b

Thus, after multiplying equation (3) by q(b) and re-naming c_{n+1} by c, we get the desired result.

▼

The theorem above implies that for $f : R^m \to R$ and $g : R \to R$ given by g(t) = f(a + t(x – a)), $E_n(1)$ is the same as $\frac{g^{(n+1)}(c)}{(n+1)!}$ for some c in the interval (0, 1). In terms of the function f, this is the same as $\left.\frac{((x-a)\bullet\nabla)^{n+1} f(u)}{(n+1)!}\right|_{u=x_0}$ where $x_0 = a + c(x-a)$.