Math 116 – Spring 2009 – Taylor Polynomials using Integration by Parts

The extension from local linear approximation to local quadratic approximation was made in connection with understanding the trapezoid and midpoint rules for numerical integration. Viewing the linear approximation as having constant derivative then suggests extending to the case of **a linear approximation for $f'(x)$ near $x = c$**, which means a quadratic approximation of the original function. In this handout, the further extension to a local approximation by a polynomial of any given order, known as a **Taylor polynomial** is developed. The error made by such an approximation is given in several equivalent forms. The use of Taylor polynomials is then illustrated in several settings.

The local linear approximation to a differentiable function $f(x)$ near $x = c$ is given by $f(c) + f'(c) \cdot (x - c)$. If this approximation is made to $f'(x)$, namely $f'(c) + f''(c) \cdot (x - c)$, and then integrating back to approximate $f(x)$ using the value at $x = c$ generates a quadratic approximation:

$$f(x) \approx P_2(x) = f(c) + f'(c) \cdot (x - c) + f''(c) \cdot \frac{(x - c)^2}{2}.$$

Example: The function $f(x) = \sqrt{1 + x}$ can be differentiated easily, with the following results: $f'(x) = \frac{1}{2}(1 + x)^{-\frac{1}{2}}$ and $f''(x) = -\frac{1}{4}(1 + x)^{-\frac{3}{2}}$. Using the location $c = 0$ where $f(0) = 1$, $f'(0) = \frac{1}{2}$, $f''(0) = -\frac{1}{4}$, leads to the approximation:

$$\sqrt{1 + x} \approx 1 + \frac{1}{2}\,x - \frac{x^2}{8}.$$

To get a sense of the accuracy of this approximation, consider two numerical examples we know: $\sqrt{1.21} = 1.1$ which uses $x = 0.21$ and closer in $\sqrt{1.0201} = 1.01$ which uses $x = 0.0201$. Direct calculation of the quadratic approximation at these values of x yield approximations as follows:

$$\sqrt{1.21} \approx 1.099487500 \quad \text{and} \quad \sqrt{1.0201} \approx 1.009999499$$

which are clearly quite accurate and better than the linear approximation alone.

Note that reducing x by a factor of 0.1 reduced the error by a factor of about 0.001, suggesting that the error in this level of approximation is proportional to x^3.

Error in Quadratic Approximation in the Limit: An application of L'Hopital's rule repeatedly helps us get a feeling for the error in the quadratic approximation $T_2(x)$ and confirms the apparent cubic size of the remainder as follows:

$$\lim_{x\to c}\frac{f(x)-P_2(x)}{(x-c)^3}=\lim_{x\to c}\frac{f(x)-(f(c)+f'(c)\cdot(x-c)+f''(c)\cdot\frac{(x-c)^2}{2})}{(x-c)^3}$$

$$=\lim_{x\to c}\frac{f'(x)-(f'(c)+f''(c)\cdot(x-c))}{3\,(x-c)^2}=\lim_{x\to c}\frac{f''(x)-f''(c)}{3\cdot 2\,(x-c)}=\frac{f^{(3)}(c)}{3!}.$$

The above calculation tells us how to create a cubic local approximation by putting our limit value back under the denominator of $(x-c)^3$, leading to a general recipe by extension:

The Taylor polynomial $P_n(x)$ for a function $f(x)$ of order n near the location $x=c$ is given by the expression:

$$P_n(x)=f(c)+f'(c)\cdot(x-c)+\frac{1}{2}f''(c)\cdot(x-c)^2+\ldots+\frac{1}{n!}f^{(n)}(c)\cdot(x-c)^n$$

This polynomial is designed so that it has the same value and same value for its first n derivatives at the location $x=c$. The n-th derivative is constant. Notice that the polynomial will nearly always be different for different functions f and also for different locations c. Therefore we might write $P_n(x,f,c)$, but we rarely indicate that dependence.

The key properties of the Taylor polynomial $P_n(x)$ are:

1. $P_n(x)$ has the same value at $x=c$ as the function f it approximates, along with the first n derivatives also, each evaluated at the point $x=c$.
2. The $n+1$-st derivative of P_n is zero for all x.
3. As a function, P_n is defined for all x values.
4. $\lim_{x\to c}\frac{f(x)-P_n(x)}{(x-c)^{n+1}}=\frac{f^{(n+1)}(c)}{(n+1)!}$ assuming the differentiability of f.

Example: $f(x)=\sin(x)$, $n=5$, $c=0$. Here the function value $f(c)$ is $\sin(0)=0$, while its first derivative is $\cos(x)$ which is 1 at $x=0$. The second derivative is 0, the third derivative is -1, and this pattern continues forever, so we find the polynomial with $n=5$ to be:

$$P_5(x)=x-\frac{x^3}{3!}+\frac{x^5}{5!}$$

We know that the sine function never goes above height 1 or below -1, but the polynomial surely does, so the approximation is not very good for large values of x.

Taylor's theorem with Remainder describes how to approximate a function near some location using a polynomial of degree n (or less) for some positive number n and expresses the error in some fashion. One way to derive the formula with an expression for the error/remainder is by using integration by parts plus the Fundamental theorem.

Taylor's theorem (with remainder) describes quantitatively the difference between $f(x)$ and $P_n(x)$, which we write as $f(x) = P_n(x) + R_n(x)$ (R_n is called the remainder). One formula for this is an integral expression (there are lots of others too), which we find using integration by parts. **This is not the most common presentation, and not the one in our text.** But it isn't very technical either, once we see the trick.

Taylor's theorem says that if f has $n+1$ continuous derivatives on an interval containing x and a, then $f(x) = P_n(x) + R_n(x)$ where

$$R_n(x) = \int_c^x \frac{(x-t)^n}{n!} \cdot f^{(n+1)}(t)\, dt.$$

The main point is that if c and x are close together and the derivative function $f^{(n)}(t)$ is reasonable in size, then $R_n(x)$ will be small for large n. In particular, a use of the mean value theorem and a substitution in the integration variable shows that the following equation holds:

$$R_n(x) = f^{(n+1)}(t_*)\,\frac{(x-c)^{n+1}}{(n+1)!}$$

for some value t_* that lies between x and c.

Proof: For $n = 0$, we find the formula becomes

$$f(x) = f(c) + R_0(x), \quad R_0(x) = \int_c^x f'(t)\, dt.$$

This is the Fundamental Theorem of Calculus!!

For $n = 1$, we want to show

$$f(x) = f(c) + f'(c) \cdot (x-c) + \int_c^x (x-t) \cdot f''(t)\, dt.$$

We note that the linear term $f'(c) \cdot (x-c)$ is the value of the function $(x-t)f'(t)$ as a function of t when $t = c$. This suggests our integration by parts trick.

From the case $n = 0$ (Fund. Theorem of Calculus), we know

$$f(x) = f(c) + \int_c^x f'(t)\, dt$$

so we really need to show that

$$\int_c^x f'(t)\, dt = f'(c) \cdot (x-c) + \int_c^x (x-t) \cdot f''(t)\, dt.$$

This follows from integration by parts as follows:

$$\int_c^x f'(t)\,dt = f'(t)\cdot(t-x)|_{t=a}^{t=x} - \int_c^x (t-x)f''(t)\,dt$$

using integration by parts with $u = f'(t), dv = dt, v = t - x$ (**not plain** t). Now notice that the term for v, namely $t - x$, is zero when $t = x$, so the only term from the product uv is when $t = c$, which was our guiding observation earlier. The minus sign for the term with $t = c$ balances the other minus sign, and the two minus signs in the integral term become a plus also.

Finally, to do the case of general n, use induction and integration by parts in a similar fashion. Thus we want to express

$$R_{n-1}(x) = \int_c^x \frac{(x-t)^{n-1}}{(n-1)!}\cdot f^{(n)}(t)\,dt$$

in terms of the sum of the degree n term in the polynomial plus $R_n(x)$. Again using integration by parts we obtain the desired result as follows:

$$R_{n-1}(x) = \int_c^x \frac{(x-t)^{n-1}}{(n-1)!}\cdot f^{(n)}(t)\,dt = -f^{(n)}(t)\cdot\frac{(x-t)^n}{n!}|_{t=a}^{t=x} - \int_c^x -f^{(n+1)}(t)\frac{(x-t)^n}{n!}\,dt$$

which used integration by parts with $u_n = f^{(n)}(t), dv_n = \frac{(x-t)^{n-1}}{(n-1)!}\,dt$. Noting the value at $t = x$ is zero, but at $t = c$ we get the next term in the Taylor polynomial, and the integral is the remainder term $R_n(x)$ once we cancel two minus signs multiplied.

Example of estimation: For the function $g(x) = \cos(2\,x)$ we find $g'(x) = -2\,\sin(2\,x)$ and $g''(x) = -4\,\cos(2\,x)$ and so on, so for any t_* we can guarantee that

$$|g^{(n+1)}(t_*)| \le 2^{n+1}$$

and using this in the remainder formula for this example, we find:

$$|R_n| \le \frac{2^{n+1}\,|x-c|^{n+1}}{(n+1)!}$$

Using Taylor Polynomial Approximation to Approximate π

With the combination of Taylor polynomials and integration, Isaac Newton gave a very accurate estimate for the number π. He did this by computing a particular area, one-sixth of a circle of radius 1/2, approximately but very accurately.

To keep the algebra simple, Newton chose the circle with diameter through the points $(0,0)$ and $(1,0)$. This has radius 1/2 and center at $(\frac{1}{2},0)$, and thus equation $y^2+(x-\frac{1}{2})^2 = (\frac{1}{2})^2$. After expansion or by direct reasoning about roots, this becomes:

$$y^2 + x^2 - x = 0$$

and thus the equation for the upper semicircle is

$$y = \sqrt{x - x^2} = \sqrt{x(1-x)}.$$

The point $(\frac{1}{4}, \frac{\sqrt{3}}{4})$ on the circle cuts off a slice of the circle of angle $\pi/3$ with the side along the x-axis from the center to $(0,0)$, so the area is $\pi/24$, since it is 1/6 of the circle of radius 1/2. At the same time, the area is composed of a triangle, whose area is $\sqrt{3}/32$ (look at base, height). Thus the area under the curve from 0 to 1/4 is precisely $\pi/24 - \sqrt{3}/32$. Newton noted that if he could find an accurate approximation to this area and a good approximate value for the triangular area involving $\sqrt{3}$, he would have a good estimate for π.

So the estimation of π is reduced to calculating accurately the integral

$$\int_0^{\frac{1}{4}} \sqrt{x(1-x)}\,dx.$$

Newton did this by using the Taylor polynomial with enough terms to approximate the factor $\sqrt{1-x}$ and then multiplying by $\sqrt{x}$. Each term then involves integration of a fractional power, easily done, and combining them together gives the desired estimate.

Let $g(x) = \sqrt{1-x}$. This is the composite function $u^{\frac{1}{2}}$ where $u = 1-x$, so it has derivatives that are fairly painless to find, and on the interval $0 \le x \le \frac{1}{4}$ the factor $1-x$ ranges from 1 to $\frac{3}{4}$ and stays far from the value 0 where the differentiability fails. By the chain rule, the $1-x$ term contributes a factor of -1 for each differentiation in u, so the Taylor polynomial for $g(x)$ is the well-known expansion

$$g(x) = \sqrt{1-x} = 1 - \frac{1}{2}\,x - \frac{1}{8}\,x^2 - \frac{1}{16}\,x^3 - \frac{5}{128}\,x^5 - \ldots$$

and then multiplying by the square root term we get the terms to integrate:

$$\sqrt{x(1-x)} = x^{\frac{1}{2}}\,g(x) = x^{\frac{1}{2}} - \frac{1}{2}\,x^{\frac{3}{2}} - \frac{1}{8}\,x^{\frac{5}{2}} - \frac{1}{16}\,x^{\frac{7}{2}} - \frac{5}{128}\,x^{\frac{11}{2}} - \ldots$$

and then integrating we get the expansion for the area under the curve, namely

$$\int_0^{\frac{1}{4}} \sqrt{x(1-x)}\,dx = \int_0^{\frac{1}{4}} [x^{\frac{1}{2}}\,g(x) = x^{\frac{1}{2}} - \frac{1}{2}\,x^{\frac{3}{2}} - \frac{1}{8}\,x^{\frac{5}{2}} - \frac{1}{16}\,x^{\frac{7}{2}} - \frac{5}{128}\,x^{\frac{9}{2}} - \ldots]\,dx$$
$$= \frac{2}{3}\,x^{\frac{3}{2}} \;-\frac{1}{5}\,x^{\frac{5}{2}} \;-\frac{1}{28}\,x^{\frac{7}{2}} - \frac{1}{72}\,x^{\frac{9}{2}} - \frac{5}{11\cdot 64}x^{\frac{11}{2}} \;-\ldots|_0^{\frac{1}{4}}$$

which evaluates to the sum

$$\frac{2}{3}\cdot\frac{1}{8} - \frac{1}{5}\cdot\frac{1}{32} - \frac{1}{28}\cdot\frac{1}{128} \;-\frac{1}{72}\cdot\frac{1}{512} \;-\frac{5}{11\cdot 64}\cdot\frac{1}{2048} \;-\ldots$$

which works out to the decimal .076777310628 if we truncate after the 13/2 power. In Maple, I worked with 20 decimal places and worked out the expansion for the polynomial approximation to $g(x)$ to the ninth power. The upshot is an approximation to π as $24[\frac{\sqrt{3}}{32}+$ computed area] which has an error of less than 0.3×10^{-8}, which is a relative error of less than 10^{-9}.

Note that Newton's success was related to the wise choice of area, the evaluation point which simplified the many square roots to estimate, and accurate computation of the square root of 3, along with a minor amount of computational fortitude.

A similar integral with the derivative of the arcsine, namely $(1-x^2)^{-\frac{1}{2}}$ could work, but you need to stay away from $x=-1$ and $x=1$. You would also need some ingenuity to get a series for that square root, but it could be done by using the substitution $u=x^2$ and finding the series in u near $u=0$ for the appropriate square root function. It turns out to be less accurate for the equivalent work of calculation. I'm not sure if Newton tried that.

Later we will see how Euler did a different estimation of π using the Taylor polynomials for the inverse tangent function along with some major ingenuity.

Using Taylor Polynomial Approximation to Revisit Calculus Topics

Second Derivative Test for Local Max/Min: At a point where $f'(c) = 0$ and $f''(c) \neq 0$, the quadratic approximation determines the values near $x = c$, with the result that $f''(c) > 0$ leads to a parabolic approximation opening upward, while $f''(c) < 0$ leads to a parabolic downward approximation.

Second Derivative Positivity in General and Convexity: At a point where $f'(c) \neq 0$ and $f''(c) \neq 0$, the quadratic term is added or subtracted from the linear approximation, leading to the result that the sign of $f''(c)$ determines whether the graph of f lies above or below the graph of the tangent line locally, which is one way concavity or convexity is defined.

Error Terms for Simpson's Rule: For the Trapezoid and Midpoint Approximations for integration, the quadratic Taylor polynomial described how the methods lead to error and then Simpson's Rule was designed to integrate quadratic terms correctly. For $(x-c)^3$ on an interval with center at $x = c$, the integral is 0 and Simpson's Rule does it correctly (indeed, so do the Trapezoid and Midpoint separately). So according to Taylor polynomial approximation, the next term to look at is $(x-c)^4$. On a single interval of length $h = \frac{b-a}{n}$, the actual integral of the term $(x-c)^4$ is $(x-c)^5/5$ evaluated at the endpoints, which are $x = c \pm h/2$, leading to the exact value $h^5/80$. Simpson's Rule would use the weighted sum of the values of the fourth power at the endpoints and at the center, leading to an approximation value of $h^5/48$, so the error is:

$$\text{exact value } - \text{Simpson's } = h^5(1/80 - 1/48) = -h^5/120$$

which should be multiplied by $f^{(4)}(t_*)/4!$. Many books use $2n$ where this description used n so the fraction in the estimate will look different. As usual, when adding the pieces together, a Riemann sum for the integral of the derivative comes into play or an extreme estimate using the maximal value of the derivative can be used.

Final Remarks

Taylor polynomials provide a local approximation by polynomials. The theme of approximation has many variations. One is to do global approximation, either based on values at selected points or integrals. These methods include Fourier series, various orthogonal polynomial expansions, and using values at points and polynomials, the so-called Newton polynomials which are a kind of extension of the Taylor polynomial idea where derivatives become differences of function values.

Other functions can be used locally. One approach is to use ratios of polynomials and again match the function and derivatives at $x = c$. This is called Padé approximation and to whet your appetite, one approximation to the function e^x near $x = 0$ is given by

$$P_{2,2}(x) = \frac{x^2 + 6x + 12}{x^2 - 6x + 12}$$

which if you explore this approximation on a computer graphically and numerically does a better job than the Taylor polynomial of order 4.