Taylor Series with Remainder and Application to Stirling's Formula

John A. Gubner
Department of Electrical and Computer Engineering
University of Wisconsin–Madison

Abstract

Taylor's Theorem with remainder is stated and applied to the function $\ln(1-x)$ and used to show that

$$\left(1+\frac{x}{n}\right)^n \to e^x.$$

Other simple applications, a derivation of Stirling's Formula, and a simple proof of the theorem are also given.

If you find this writeup useful, or if you find typos or mistakes, please let me know at John.Gubner@wisc.edu

1. The Theorem and Examples

Taylor's Theorem with Remainder. *Suppose that f is n times continuously differentiable on $[a,b]$ for some integer $n \geq 0$, and assume that $f^{(n+1)}$ exists on (a,b). Fix any $x, x_0 \in [a,b]$. Then there exists a ξ between x and x_0 such that*

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k + \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-x_0)^{n+1}.$$

Example 1 (Taylor Series for $\ln(1-x)$). We begin by showing that if $f(x) = \ln(1-x)$ for $x < 1$, then $f'(x) = -1/(1-x)$, $f''(x) = -1/(1-x)^2$, $f'''(x) = -2/(1-x)^3$, and $f^{(iv)}(x) = -2\cdot 3/(1-x)^4$. In general, we have

$$f^{(k)}(x) = -\frac{(k-1)!}{(1-x)^k}.$$

Taking $x_0 = 0$, we have $f(0) = 0$, and so

$$\ln(1-x) = -\sum_{k=1}^{n} \frac{x^k}{k} - \frac{x^{n+1}}{(n+1)(1-\xi)^{n+1}}, \quad x < 1, \tag{1}$$

where ξ lies between 0 and x. Notice that $1-\xi$ is always positive, whether $0 < \xi < x < 1$ or $x < \xi < 0$. When $n = 1$, we have

$$\ln(1-x) = -x - \frac{x^2}{2(1-\xi)^2}, \quad x < 1.$$

Example 2 (Bounding the Remainder Term). In the above formula, if $|x|$ is small, then $|\xi|$ is also small because ξ lies between 0 and x. Hence, given any $0 < \varepsilon < 1$, if $|x| < 1$ is small enough, ξ will be so close to zero that $1/(1-\xi)^2$ is nearly equal to one, say that

$$1-\varepsilon < \frac{1}{(1-\xi)^2} < 1+\varepsilon,$$

and it follows that

$$\frac{x^2}{2}(1-\varepsilon) < \frac{x^2}{2(1-\xi)^2} < \frac{x^2}{2}(1+\varepsilon).$$

This allows upper and lower bounds on the approximation $\ln(1-x) \approx -x$; e.g.,

$$\frac{x^2}{2}(1+\varepsilon) < \ln(1-x) + x < \frac{x^2}{2}(1-\varepsilon). \tag{2}$$

Example 3 (Taylor Series for $\ln(1+t)$). If we put $x = -t$ in (1), we have

$$\ln(1+t) = -\sum_{k=1}^{n} \frac{(-t)^k}{k} - \frac{(-t)^{n+1}}{(n+1)(1-\xi)^{n+1}}, \quad t > -1,$$

where ξ lies between 0 and $-t$. With $n = 1$, we obtain

$$\ln(1+t) = t - \frac{t^2}{2(1-\xi)^2}. \tag{3}$$

The analog of (2) is

$$\frac{t^2}{2}(1+\varepsilon) < \ln(1+t) - t < \frac{t^2}{2}(1-\varepsilon). \tag{4}$$

This observation is useful in deriving Stirling's Formula in Section 3.

Example 4 (Taylor Series for $\ln t$). If we put $x = 1-t$ in (1), we have

$$\ln t = -\sum_{k=1}^{n} \frac{(1-t)^k}{k} - \frac{(1-t)^{n+1}}{(n+1)(1-\xi)^{n+1}}, \quad t > 0.$$

With $n = 1$, we have

$$\ln t = t - 1 - \frac{(t-1)^2}{2(1-\xi)^2}, \quad t > 0.$$

Dropping the last term yields the "log inequality,"

$$\ln t \le t - 1, \quad t > 0.$$

Remark. The log inequality can also be obtained directly by writing

$$\ln t = \int_1^t \frac{1}{x}\,dx \le \int_1^t 1\,dx = t - 1, \quad t \ge 1,$$

since $1/x \le 1$ for $x \ge 1$. A similar argument with $0 < t < 1$ yields the same inequality.

Example 5 (Convergence of $(1 + x/n)^n$ to e^x). We show that $n\ln(1+x/n) \to x$, from which it follows that

$$(1 + x/n)^n \to e^x$$

since the exponential is a continuous function. In (3), let $t = x/n$ so that

$$\ln(1 + x/n) = x/n - \frac{(x/n)^2}{2(1-\xi_n)^2},$$

where we write ξ_n because it depends on n. Since ξ_n lies between $-x/n$ and zero, as n becomes large, $\xi_n \to 0$. Multiplying the above display by n yields

$$n\ln(1 + x/n) = x - \frac{x^2/n}{2(1-\xi_n)^2}.$$

Hence,

$$\Big| n\ln(1 + x/n) - x \Big| = \frac{x^2/n}{2(1-\xi_n)^2}.$$

The denominator tends to two, and the numerator tends to zero, and so the quotient tends to $0/2 = 0$.

Example 6 (Generalization of Example 5). A similar argument shows that if $x_n \to x$, then $n\ln(1 + x_n/n) \to x$. It then follows that

$$(1 + x_n/n)^n \to e^x.$$

To see that $n\ln(1+x_n/n) \to x$, write

$$\begin{aligned}\left|n\ln(1+x_n/n) - x\right| &\le \left|n\ln(1+x_n/n) - x_n\right| + |x_n - x| \\ &\le \frac{x_n^2/n}{2(1-\xi_n)^2} + |x_n - x|.\end{aligned}$$

Since x_n converges, it is bounded, and so x_n^2/n tends to zero.

2. Proof of the Taylor's Theorem with Remainder

An easy proof can be carried out by regarding the polynomial approximation as a function of both x and x_0; i.e., put

$$\begin{aligned}p_n(x,x_0) &:= \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k \\ &= f(x_0) + f'(x_0)(x-x_0) + \cdots + \frac{f^{(n)}(x_0)}{n!}(x-x_0)^n.\end{aligned}$$

Note that $p_n(t,t) = f(t)$ since the terms with derivatives are multiplied by zero.

We must prove the existence of ξ such that

$$f(x) = p_n(x,x_0) + \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-x_0)^{n+1}. \tag{5}$$

Consider the function

$$h(t) := p_n(x,t) + \frac{f(x) - p_n(x,x_0)}{(x-x_0)^{n+1}}(x-t)^{n+1}.$$

Observe that

$$h(x_0) = p_n(x,x_0) + f(x) - p_n(x,x_0) = f(x),$$

and

$$h(x) = p_n(x,x) + \frac{f(x) - p_n(x,x_0)}{(x-x_0)^{n+1}} \cdot 0 = p_n(x,x) = f(x)$$

as well. By Rolle's Theorem [2, p. 107, Th. 5.8], there is a point ξ between x and x_0 such that $h'(\xi) = 0$. Now

$$\begin{aligned}h'(t) &= \frac{\partial}{\partial t} p_n(x,t) - \frac{f(x) - p_n(x,x_0)}{(x-x_0)^{n+1}}(n+1)(x-t)^n \\ &= \frac{f^{(n+1)}(t)}{n!}(x-t)^n - \frac{f(x) - p_n(x,x_0)}{(x-x_0)^{n+1}}(n+1)(x-t)^n,\end{aligned}$$

because the partial derivative results in a telescoping sum that simplifies (use the derivative-of-product rule for for $k = 1, \dots, n$, but not for $k = 0$). Setting $t = \xi$ and $h'(\xi) = 0$ and canceling common factors, we have

$$\frac{f^{(n+1)}(\xi)}{n!} = \frac{f(x) - p_n(x, x_0)}{(x - x_0)^{n+1}}(n+1),$$

from which (5) follows. □

3. Derivation of Stirling's Formula

Recall that

$$\Gamma(x) := \int_0^\infty t^{x-1} e^{-t}\, dt, \quad x > 0.$$

Using integration by parts, it is easy to show that $\Gamma(x+1) = x\Gamma(x)$. We derive a limiting form for $\Gamma(x+1)$ and divide the result by x to obtain Stirling's Formula,

$$\boxed{\lim_{x\to\infty} \frac{\Gamma(x)}{x^{x-1/2}e^{-x}} = \sqrt{2\pi}.}$$

We begin with

$$\Gamma(x+1) = \int_0^\infty t^x e^{-t}\, dt = \int_0^\infty \exp[x \ln t - t]\, dt \tag{6}$$

and follow the method of Diaconis and Freedman [1]. The big-picture plan is to make the change of variable $v = (t - x)/\sqrt{x}$ or $t = x + \sqrt{x}\, v$ so that

$$\Gamma(x+1) = \sqrt{x} \int_{-\sqrt{x}}^\infty \exp[x \ln(x + \sqrt{x}v) - (x + \sqrt{x}v)]\, dv.$$

Then write the argument of exp as

$$\begin{aligned} x \ln\bigl(x + \sqrt{x}v\bigr) - \bigl(x + \sqrt{x}v\bigr) &= x\{\ln(x + \sqrt{x}v) - (1 + v/\sqrt{x}\,)\} \\ &= x\{\ln(1 + v/\sqrt{x}\,) + \ln x - (1 + v/\sqrt{x}\,)\} \end{aligned}$$

so that[1]

$$\Gamma(x+1) = x^{x+1/2} e^{-x} \int_{-\sqrt{x}}^\infty e^{-xg\left(v/\sqrt{x}\right)}\, dx, \tag{7}$$

[1] Since $\Gamma(x+1) = x\Gamma(x)$, the formula for $\Gamma(x)$ will involve $x^{x-1/2}$ instead of $x^{x+1/2}$.

where

$$g(t) := t - \ln(1+t). \tag{8}$$

The remaining work is to show that this last integral tends to $\sqrt{2\pi}$ as $x \to \infty$. This will be done by breaking the range of integration into three parts so that one of them tends to $\sqrt{2\pi}$ while the others tend to zero.

3.1. Motivating the Change of Variable in Two Steps

The integrand on the right in (6) is maximized at $t = x$, which we take to be positive since we will later divide by x. Since the maximum occurs at $t = x$, we first make the change of variable $u = t - x$ in (6) to obtain

$$\Gamma(x+1) = \int_{-x}^{\infty} \exp[x\ln(u+x) - (u+x)]\,du. \tag{9}$$

Write

$$\begin{aligned} x\ln(u+x) - (u+x) &= x\{\ln(u+x) - (1+u/x)\} \\ &= x\Big\{\ln\frac{u+x}{x} + \ln x - (1+u/x)\Big\} \\ &= x\{\ln(1+u/x) + \ln x - (1+u/x)\} \\ &= x\{\ln(1+u/x) - u/x)\} + x\ln x - x. \end{aligned}$$

Hence,

$$\begin{aligned} \Gamma(x+1) &= x^x e^{-x} \int_{-x}^{\infty} \exp[x\{\ln(1+u/x) - u/x)\}]\,du \\ &= x^x e^{-x} \int_{-x}^{\infty} e^{-xg(u/x)}\,du, \quad \text{where } g \text{ is defined in (8).} \end{aligned}$$

By (3), we see that

$$xg(u/x) = \frac{u^2/x}{2(1-\xi)},$$

which still has x in the numerator. So we make the change of variable $v = u/\sqrt{x}$ and obtain (7).

3.2. Analysis of the Integral

Below we will break up the range of integration in (7) into the three intervals $(-\sqrt{x}, -L)$, $[-L, L]$, and (L, ∞), and use the approximation

$$xg\big(v/\sqrt{x}\,\big) \approx v^2/2, \quad \text{for } v \in [-L, L],$$

along with the fact that $\int_{-\infty}^{\infty} e^{-v^2/2}\,dv = \sqrt{2\pi}$. But first, fix $0 < \varepsilon < 1$, and choose L so large that

$$\int_{-L\sqrt{1\pm\varepsilon}}^{L\sqrt{1\pm\varepsilon}} e^{-v^2/2}\,dv \approx \int_{-\infty}^{\infty} e^{-v^2/2}\,dv = \sqrt{2\pi} \tag{10}$$

and such that $e^{-L^2/2}/L$ is negligible. Choose $x > L^2$ (i.e., $\sqrt{x} > L$) so large that if $|\xi| < L/\sqrt{x}$, then

$$1-\varepsilon < \frac{1}{(1-\xi)^2} < 1+\varepsilon.$$

Multiplying through by $-v^2/2$ yields

$$-\frac{v^2}{2}(1-\varepsilon) > -\frac{v^2}{2(1-\xi)^2} > -\frac{v^2}{2}(1+\varepsilon).$$

By (8) and (4), we see that $-xg\big(v/\sqrt{x}\,\big)$ is equal to the middle quotient above where ξ lies between $-v/\sqrt{x}$ and zero. In the integral

$$\int_{-L}^{L} e^{-xg\left(v/\sqrt{x}\right)}\,dv,$$

$|v|/\sqrt{x} < L/\sqrt{x}$, and so it satisfies

$$\int_{-L}^{L} e^{-v^2(1+\varepsilon)/2}\,dv \le \int_{-L}^{L} e^{-xg\left(v/\sqrt{x}\right)}\,dv \le \int_{-L}^{L} e^{-v^2(1-\varepsilon)/2}\,dv.$$

These upper and lower bounds are approximately $\sqrt{2\pi}$ on account of (10). It remains to consider the other two intervals of integration mentioned earlier. Put $h_x(v) := xg\big(v/\sqrt{x}\,\big)$, and note that

$$h_x'(v) = xg'\big(v/\sqrt{x}\,\big)/\sqrt{x} = \sqrt{x}\left[1 - \frac{1}{1+v/\sqrt{x}}\right] = \frac{v}{1+v/\sqrt{x}}$$

is increasing for $v > 0$ with $h_x'(v) \to v$ as $x \to \infty$. Note also that since $g(t) \to \infty$ as $t \to \infty$,[2] $h_x(v) \to \infty$ as $v \to \infty$. We can now write

$$\int_{L}^{\infty} e^{-xg\left(v/\sqrt{x}\right)}\,dv = \int_{L}^{\infty} e^{-h_x(v)}\,dv$$

[2] Since $g'(t) = 1 - 1/(1+t) = t/(1+t)$, it follows that

$$g(t) - g(1) = \int_1^t g'(s)\,ds = \int_1^t \frac{s}{1+s}\,ds \ge \int_1^t \frac{1}{1+s}\,ds = \ln(1+t) - \ln 2,$$

which tends to infinity as t increases.

$$
\begin{aligned}
&= \int_L^\infty \frac{h_x'(v)}{h_x'(v)} e^{-h_x(v)}\, dv \\
&\le \int_L^\infty \frac{h_x'(v)}{h_x'(L)} e^{-h_x(v)}\, dv \\
&= \frac{1}{h_x'(L)} \int_L^\infty h_x'(v) e^{-h_x(v)}\, dv \\
&= \frac{1}{h_x'(L)} \left(-e^{-h_x(v)}\right)\Big|_{v=L}^{v=\infty} \\
&= \frac{e^{-h_x(L)}}{h_x'(L)}.
\end{aligned}
$$

As $x \to \infty$, $h_x'(L) \to L$, while $h_x(L) = xg\big(L/\sqrt{x}\big) \to L^2/2$. Hence, the above integral tends to $e^{-L^2/2}/L$, which is negligible since L is large. We can similarly treat

$$
\begin{aligned}
\int_{-\sqrt{x}}^{-L} e^{-xg\left(v/\sqrt{x}\right)}\, dv &= \int_{-\sqrt{x}}^{-L} e^{-h_x(v)}\, dv \\
&= \int_{-\sqrt{x}}^{-L} \frac{h_x'(v)}{h_x'(v)} e^{-h_x(v)}\, dv \\
&\le \frac{1}{h_x'(-L)} \int_{-\sqrt{x}}^{-L} h_x'(v) e^{-h_x(v)}\, dv \\
&= \frac{-e^{-h_x(-L)}}{h_x'(-L)}.
\end{aligned}
$$

Note that the denominator is negative.

References

[1] P. Diaconis and D. Freedman, “An elementary proof of Stirling’s formula,” *Amer. Math. Monthly*, vol. 93, no. 2, pp. 123–125, Feb. 1986.

[2] W. Rudin, *Principles of Mathematical Analysis*, 3rd ed. New York: McGraw-Hill, 1976.