Calculus for numerical analysis

Douglas Wilhelm Harder

April 2, 2025

Not unexpectedly, a number of students forget most of what they learned in their calculus course following a period of time between learning the material and using that material. This is a review of the concepts from your course in calculus that you will require in this course on numerical analysis.

Please note, despite being given these notes, topics that are used from calculus will never-the-less be summarized and described in class. Thus, while you should read these now, it is not expected that you have mastered all this material for the course lectures.

1 Properties of functions

Remember that a function is a mapping from one space into another. Most often, you have seen real-valued functions of a real variable, but you could have real-valued functions of two variables, such as $p(x,y) = 3x^2 + x - 4xy + 3y + 5y^2$. This "bivariate polynomial" maps the pair of real numbers (x,y) to a real number. The greatest-common-divisor function maps a pair of positive integers to a positive integer. The space we are mapping from is called the "domain" and the space we are mapping to is called the "co-domain", or "complementary domain". We write this as:

$$\begin{aligned} \sin &: \mathbf{R} \to \mathbf{R} \\ p(x,y) &: \mathbf{R} \times \mathbf{R} \to \mathbf{R} \\ \gcd &: \mathbf{Z}^+ \times \mathbf{Z}^+ \to \mathbf{Z}^+ \end{aligned}$$

where $\mathbf{R}$ represents the real numbers, integers are represented by $\mathbf{Z}$ and the positive integers (those greater than zero) are represented by $\mathbf{Z}^+$. The $\times$ represents the Cartesian product: all pairs (x,y) where x comes from the left-hand space, and y comes from the right-hand space.

The "range" of a function includes all entries in the co-domain such that there is a point in the domain that maps onto it. The range of the sine function is $[-1,1]$. The range of the bivariate polynomial p from above is $[-1,\infty)$ because $(-0.5,-0.5)$ is a global minimum and $p(-0.5,-0.5) = -1$. If you move away from this point in any direction, the value of $p(x,y)$ only increases, going to $+\infty$, which is why the range is $[-1,\infty)$. To see this, one can observe that we can rewrite this in a "completed square" form to get $p(x,y) = -4(x+0.5)(y+0.5) + 3(x+0.5)^2 + 5(y+0.5)^2 - 1$. Please note, you are not required to be able to deduce the range of such a polynomial: this is simply being presented as an interesting case.

The range of the exponential function is $(0,\infty)$, the range of the polynomial $q(x) = x^2$ is $[0,\infty)$, while the range of the polynomial $r(x) = x^3$ is $\mathbf{R}$; that is, the range is equal to the co-domain.

2 Continuous functions and the intermediate-value theorem

A function is continuous at a point x_0 if the limit of $f(x)$ as $x \to x_0$ equals $f(x_0)$. A real-valued function of a real variable $f(x)$ is continuous if it is continuous for every possible real argument x.

If a function is not continuous at a point x_0, we say that that function is discontinuous at x_0. There are three possibilities:

1. A removable discontinuity where $\lim_{x\to x_0^-} f(x) = \lim_{x\to x_0^+} f(x)$, regardless of the value of $f(x_0)$, where we have the option of prescribing $f(x_0)$ to be this common limit, so removing the discontinuity.
2. A jump discontinuity where $\lim_{x\to x_0^-} f(x) \neq \lim_{x\to x_0^+} f(x)$ and both are finite. The magnitude of the jump is the difference in these two limits.
3. An essential discontinuity if either limit $\lim_{x\to x_0^-} f(x)$ or $\lim_{x\to x_0^+} f(x)$ is either infinite or undefined.

The function $\frac{\sin(x)}{x}$ has a removable discontinuity at $x = 0$ as the limits from both the left and right equal 1. The function $\left\{ \begin{array}{cc} 0 & x < 0 \\ \mathrm{e}^{-x}\cos(x) & x \geq 0 \end{array} \right.$ has a jump discontinuity at $x = 0$. The complementary tangent function $(\cot(x))$ has an essential discontinuity at every multiple of π.

If a function f is continuous on an interval $[a, b]$, and if $\min\{f(a), f(b)\} \leq y \leq \max\{f(a), f(b)\}$, then there must exist a value $a \leq x \leq b$ such that $y = f(x)$.

3 One-to-one and onto

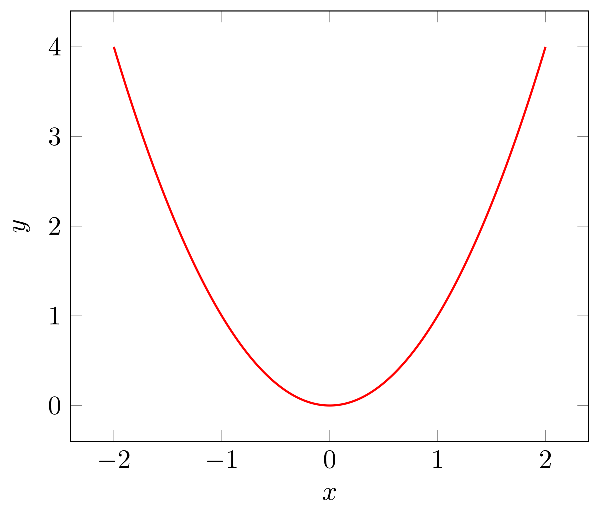
A function may have two important properties: First, a function may be "onto", meaning that its range is equal to the entire co-domain. Second, a function may be "one-to-one", meaning that each input maps to a unique point.

A function $y = f(x)$ is "onto" if for any horizontal line, the graph of that function crosses the horizontal line at least once. If there is just one y value for which there is no x such that $f(x) = y$, the function is not onto.

A function $y = f(x)$ is "one-to-one" if for any horizontal line, the graph of that function crosses that horizontal line at most once. If there is just one y value for which there are two (or more) x values, say $x_1 \neq x_2$ where $f(x_1) = f(x_2)$, then the function is not one-to-one.

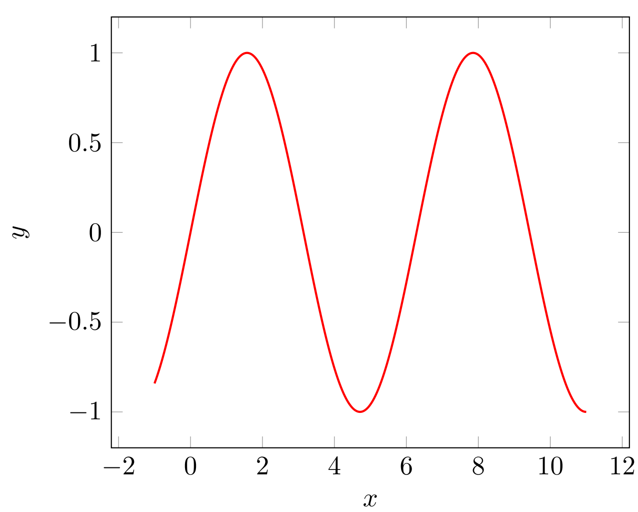
For example, the function $y = x^2$ is neither one-to-one nor is it onto. It is not one-to-one because there are two x-values that cross the graph when $y = 4$ (specifically, when $x = 2$ and $x = -2$). It is not onto because the graph does not cross the horizontal line $y = -2$.

The polynomial $y = x^2$

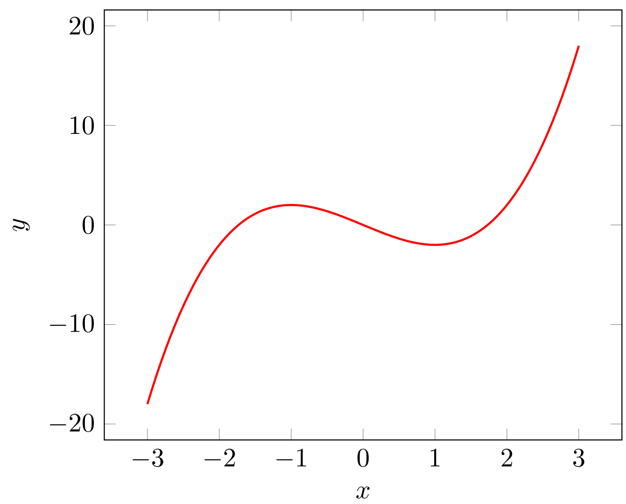


For example, the function $y = \sin(x)$ is also not one-to-one nor is it onto. It is not one-to-one because there are infinitely many values that cross the graph when $y = 0$ (specifically, when $x = n\pi$ for any integer n). It is not onto because the graph does not cross the horizontal line $y = 2$.

The sine function

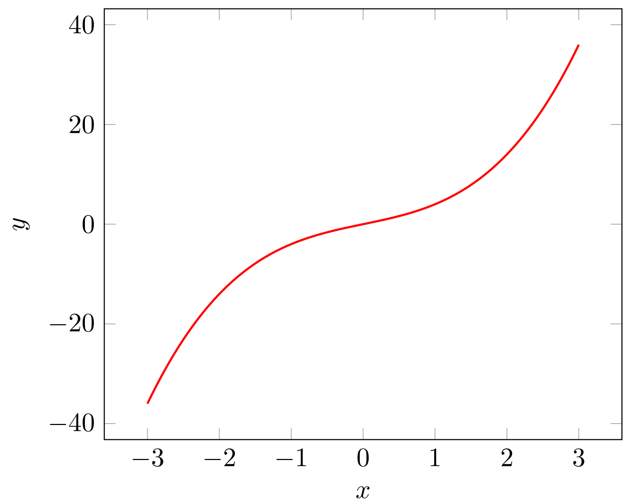


The function $y = x^3 - 3x$ is onto, but it is not one-to-one, because if $y = 2$ or $y = -2$, there are two x-values such that $x^3 - 3x = \pm 2$, and if $-2 < y < 2$, then there are three x-values such that $x^3 - 3x = y$.



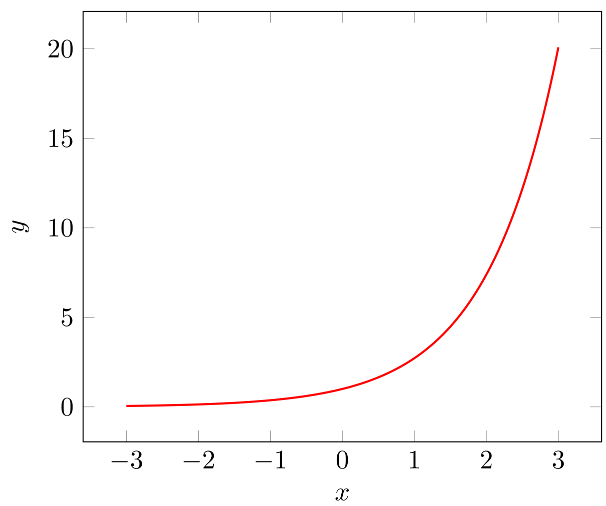
The polynomial $x^3 - 3x$

The function $y = x^3 + 3x$, however, is both one-to-one and onto.



The polynomial $x^3 + 3x$

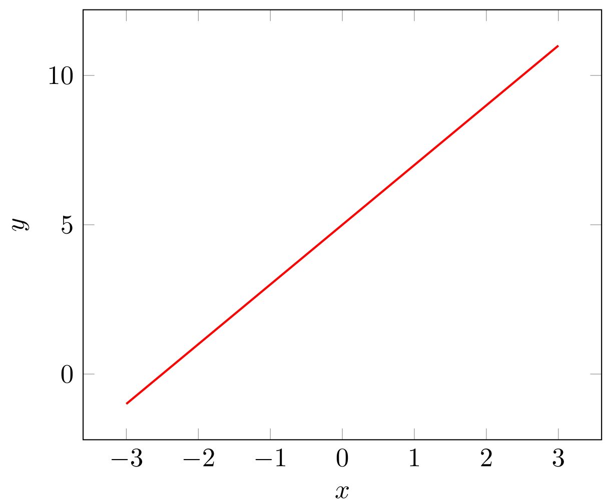
The function $y = \mathrm{e}^x$ is one-to-one, for if $y > 0$, then there is one unique solution such that $y = \mathrm{e}^x$, specifically, $x = \ln(y)$. It is not onto, because if $y \leq 0$, there is no solution such that $y = \mathrm{e}^x$.



The exponential function e^x

The linear polynomial $y = ax + b$ for $a \neq 0$ is both one-to-one and onto, because for any given y, there is one and only one x that is mapped onto it, specifically, $x = \frac{y-b}{a}$.

The linear polynomial $2x + 5$



If a function $y = f(x)$ is both one-to-one and onto, it is "invertible," meaning that we find a function $x = f^{-1}(y)$ so that $x = f^{-1}(f(x))$ for all x and $y = f(f^{-1}(y))$ for all y.

4 Invertibility

A function that is one-to-one and onto a given domain and co-domain may be "inverted" where if $f : D \to R$, we may define $f^{-1} : R \to D$ where if $y = f(x)$, we define $x = f^{-1}(y)$. Here are some common examples:

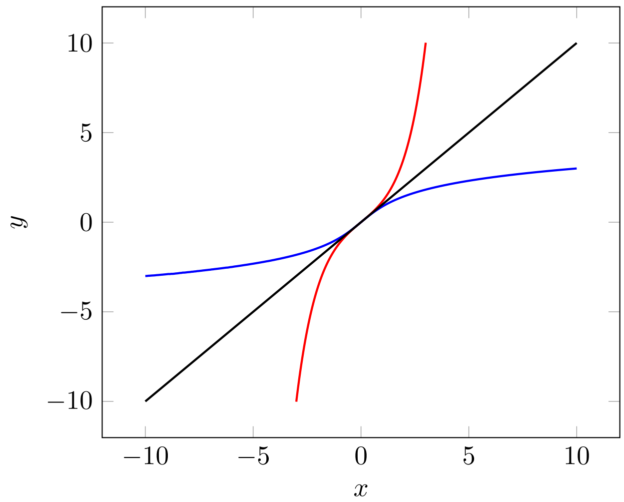
- The linear polynomial $p(x) = ax + b$ with $a \neq 0$ is one-to-one and onto, and its inverse is $x = p^{-1}(y) = \frac{y-b}{a}$.
- The function $f(x) = x^3$ is one-to-one and onto, and its inverse is $x = f^{-1}(y) = \sqrt[3]{y}$.
- The function $y = \sinh(x) = \frac{\mathrm{e}^x - \mathrm{e}^{-x}}{2}$ is one-to-one and onto, and its inverse is $x = \sinh^{-1}(y) = \ln(x + \sqrt{x^2 + 1})$, although the proof of this is beyond the scope of this course, but discussed at Wikipedia.

It is possible to restrict the domain and co-domain to make a function one-to-one and onto. For example:

- The polynomial $p(x) = x^2$ is neither one-to-one nor onto, but if we restrict the domain and co-domain to $p : [0, \infty) \to [0, \infty)$, it is both one-to-one and onto, and $p^{-1}(x) = \sqrt{x}$.
  Similarly, it is one-to-one and onto if we restrict the domain to be $p : (\infty, 0] \to [0, \infty)$, in which case, the inverse is $p^{-1}(x) = -\sqrt{x}$.
- The exponential function e^x is one-to-one but not onto, but if we restrict the co-domain to the positive real numbers, so $\mathrm{e}^{\cdot} : \mathbf{R} \to (0, \infty)$, it is also onto, so we may define its inverse, the natural logarithm, $\ln : (0, \infty) \to \mathbf{R}$ where $x = \ln(y)$ if and only if $y = \mathrm{e}^x$.
- The cosine function is neither one-to-one nor onto, but if we restrict the domain and co-domain to be $\cos : [0, \pi] \to [-1, 1]$, we may define an inverse $\cos^{-1} : [-1, 1] \to [0, \pi]$ where $x = \cos^{-1}(y)$ for $-1 \leq y \leq 1$ if and only if $y = \cos(x)$ and $0 \leq x \leq \pi$.
- The tangent function is onto but not one-to-one, but if we restrict the domain to be $\tan : \left(-\frac{\pi}{2}, \frac{\pi}{2}\right) \to \mathbf{R}$, then it is also one-to-one, and we may define $\tan^{-1} : \mathbf{R} \to \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$, where $x = \tan^{-1}(y)$ if and only if $y = \tan(x)$ and $-\frac{\pi}{2} < x < \frac{\pi}{2}$.

It is not always possible to write the inverse of a function in a nice format; indeed, the natural logarithm was defined to be the inverse of the exponential function.

The graph of the inverse $x = f^{-1}(y)$ of a function $y = f(x)$ is the reflection of the graph of the latter through the line $y = x$. For example, The hyperbolic sine function and its inverse



5 Differentiation

Given a real-valued function of a real variable f, a point x_0 and a positive value h, the slope between the points $(x_0, f(x_0))$ and $(x_0 + h, f(x_0 + h))$ is given by

$$\frac{f(x_0 + h) - f(x_0)}{(x_0 + h) - x_0} = \frac{f(x_0 + h) - f(x_0)}{h}.$$

The derivative of f at x_0 is the slope of the function f at x_0, and is written as $\frac{\mathrm{d}}{\mathrm{d}x} f(x_0)$ or $\frac{\mathrm{d}f(x_0)}{\mathrm{d}x}$. The function that is the derivative of f is written as $\frac{\mathrm{d}f}{\mathrm{d}x}$. It can be found by taking the limit as $h \to 0$ of the above slope, if it exists:

$$\frac{\mathrm{d}}{\mathrm{d}x} f(x_0) = \lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h}.$$

The derivative may also be written as f' or $f^{(1)}$ and the derivative evaluated at a point x_0 may be written as $f'(x_0)$ or $f^{(1)}(x_0)$.

There are a few rules you must know to calculate the derivative:

1. $\frac{\mathrm{d}}{\mathrm{d}x}c = 0$, or the slope of a constant function is zero.
2. $\frac{\mathrm{d}}{\mathrm{d}x}x^n = nx^{n-1}$.
3. $\frac{\mathrm{d}}{\mathrm{d}x}cf(x) = c\frac{\mathrm{d}}{\mathrm{d}x}f(x)$.
4. $\frac{\mathrm{d}}{\mathrm{d}x}(f(x) + g(x)) = \frac{\mathrm{d}}{\mathrm{d}x}f(x) + \frac{\mathrm{d}}{\mathrm{d}x}g(x)$.
5. $\frac{\mathrm{d}}{\mathrm{d}x}\sin(x) = \cos(x)$.
6. $\frac{\mathrm{d}}{\mathrm{d}x}\cos(x) = -\sin(x)$.
7. $\frac{\mathrm{d}}{\mathrm{d}x}\mathrm{e}^{\lambda x} = \lambda\mathrm{e}^{\lambda x}$.
8. $\frac{\mathrm{d}}{\mathrm{d}x}f(g(x)) = f'(g(x))\frac{\mathrm{d}}{\mathrm{d}x}g(x)$.

From Rules 3 and 4, we may deduce that

$$\frac{\mathrm{d}}{\mathrm{d}x}(\alpha f(x) + \beta g(x)) = \alpha\frac{\mathrm{d}}{\mathrm{d}x}f(x) + \beta\frac{\mathrm{d}}{\mathrm{d}x}g(x).$$

This is called "linearity" and is the same property shared with matrices:

$$A(\alpha\mathbf{u}_1 + \beta\mathbf{u}_2) = \alpha A\mathbf{u}_1 + \beta A\mathbf{u}_2.$$

Rule 7 has the same property as an eigenvalue and eigenvector: the derivative of an exponential function is a scalar multiple of that same function, just like $A\mathbf{u} = \lambda\mathbf{u}$.

We define the n^{th} derivative of f as the derivative being applied n times to f, and it is written as $\frac{\mathrm{d}^n f}{\mathrm{d}x^n}$ or $f^{(n)}(x)$.

6 Taylor series

A function that is n times differentiable at a point x_0 if f and all derivatives up to $\frac{\mathrm{d}^n f}{\mathrm{d}x^n}$ are all continuous at x_0. If f is n times differentiable on an interval $[x_0, x]$, then we can approximate $f(x)$ by considering the following "n^{th}-order Taylor series expansion around x_0":

$$f(x) \approx f(x_0) + f'(x_0)(x - x_0) + \frac{1}{2}f''(x_0)(x - x_0)^2 + \frac{1}{3!}f'''(x_0)(x - x_0)^3 + \cdots + \frac{1}{n!}f^{(n)}(x_0)(x - x_0)^n.$$

This is a polynomial in x of degree n. This means that functions can be approximated around x_0 by polynomials, and the higher the degree of the polynomial, the better the approximation.

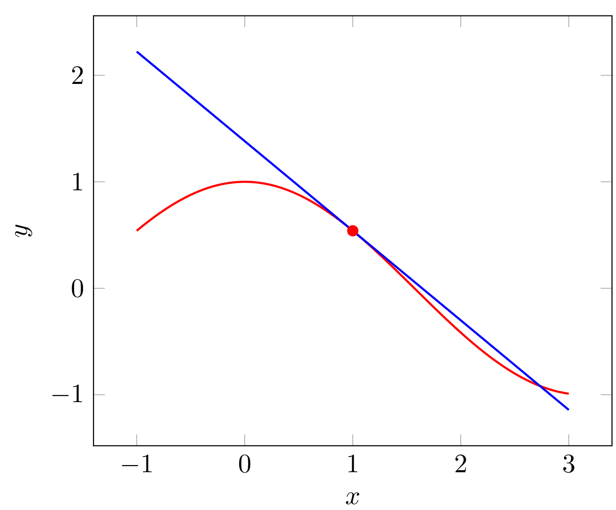
For uniformity, we will write this the notation $f^{(k)}(x)$ for the k^{th} derivative of f, yielding

$$\begin{aligned}
f(x) &\approx \sum_{k=0}^{n} \frac{1}{k!} f^{(k)}(x_0)(x - x_0)^k \\
&= \frac{1}{0!}f^{(0)}(x_0)(x - x_0)^0 + \frac{1}{1!}f^{(1)}(x_0)(x - x_0)^1 + \frac{1}{2!}f^{(2)}(x_0)(x - x_0)^2 + \cdots + \frac{1}{n!}f^{(n)}(x_0)(x - x_0)^n \\
&= f(x_0) + f^{(1)}(x_0)(x - x_0) + \frac{1}{2}f^{(2)}(x_0)(x - x_0)^2 + \cdots + \frac{1}{n!}f^{(n)}(x_0)(x - x_0)^n.
\end{aligned}$$

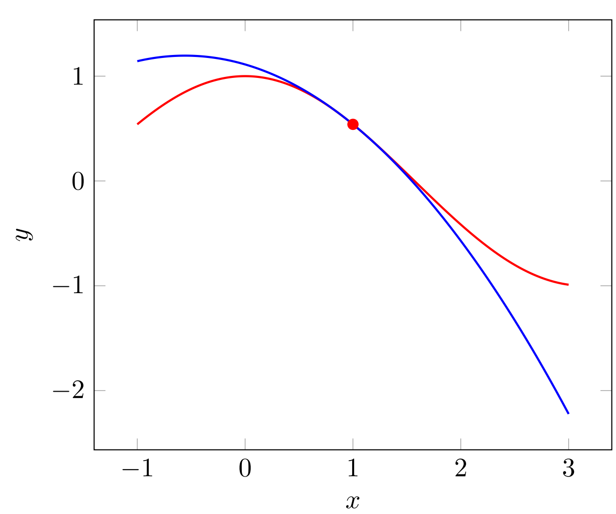
Recall that the zeroeth derivative of $f(x)$ is just $f(x)$, so $f^{(0)}(x) = f(x)$. This is called the n^{th}-order Taylor series approximation of $f(x)$.

We will look at first, second, third and fourth-order Taylor series approximations of $\cos(x)$ at $x_0 = 1$:

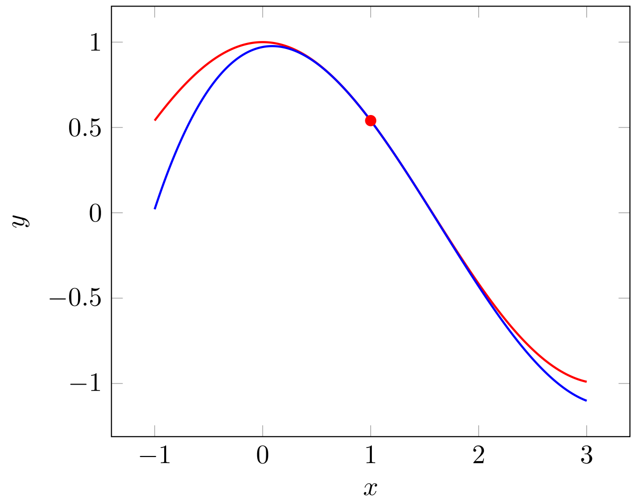
The first-order Taylor series of $\cos(x)$ at $x = 1$



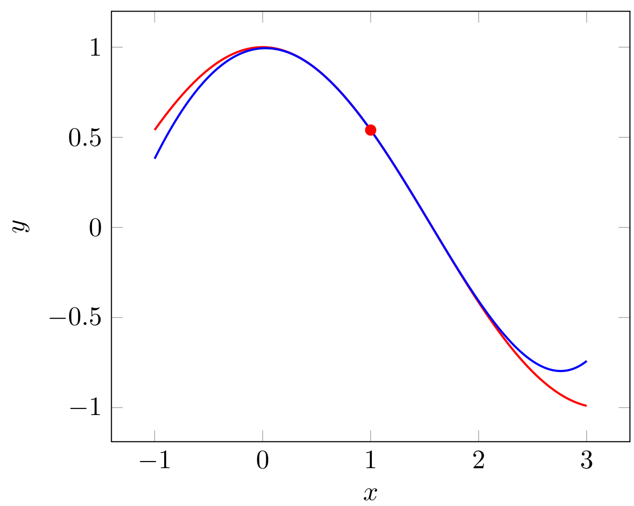
The second-order Taylor series of $\cos(x)$ at $x = 1$



The third-order Taylor series of $\cos(x)$ at $x = 1$

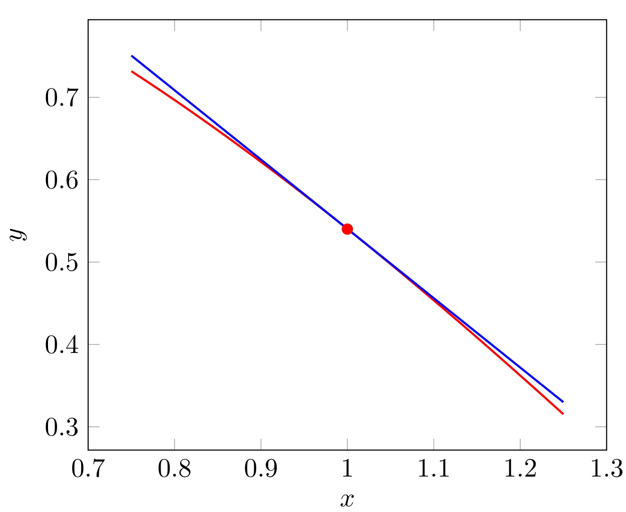


The fourth-order Taylor series of $\cos(x)$ at $x = 1$



You will see that each subsequent higher-order Taylor series approximation is more accurate over a longer region than the previous approximation. However, even if you zoom in on the first-order Taylor series, you see it becomes quite reasonable approximation of $\cos(x)$ so long as the point x is not too far away from $x_0 = 1$:

The first-order Taylor series of $\cos(x)$ at $x = 1$



Thus, one may ask how good such an approximation? Fortunately, there is even an error term that tells us approximately how large the error will be.

If f is $n + 1$ times differentiable on an interval $[x_0, x]$, then we can change Taylor series into an equality

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

where ξ is between x_0 and x. The problem is, beyond knowing that there is this ξ, but if you are looking for a proof that this is the case, here is an argument for this:

1. On the interval $[x_0, x]$, the function cannot grow above the Taylor series faster than:

$$\frac{1}{(n+1)!}\left(\max_{x_0 \le \xi \le x} f^{(n+1)}(\xi)\right)(x - x_0)^{n+1}.$$

   Similarly, it cannot grow below than the minimum value faster than:

$$\frac{1}{(n+1)!}\left(\min_{x_0 \le \xi \le x} f^{(n+1)}(\xi)\right)(x - x_0)^{n+1}.$$

2. By continuity, there are two points $x_0 \le \xi_{\min}, \xi_{\max} \le x$ such that $\max_{x_0 \le \xi \le x} f^{(n+1)}(\xi) = f^{(n+1)}(\xi_{\max})$ and $\min_{x_0 \le \xi \le x} f^{(n+1)}(\xi) = f^{(n+1)}(\xi_{\min})$.
3. Therefore, by the intermediate-value theorem, there must be some point ξ between $\xi_{\max}$ and $\xi_{\min}$ such that the above must be an equality, and because $\xi_{\max}$ and $\xi_{\min}$ are between x_0 and x, then so must $x_0 \le \xi \le x$.

The symbol "ξ" (pronounced 'k-psi') is the Greek equivalent of "x".

6.1 The applicability of Taylor series

The usefulness of the Taylor series lies not only in providing a polynomial approximation but also in revealing when such an approximation breaks down. Specifically, it highlights situations where relying on a Taylor series is inappropriate—most notably when the system experiences a sudden change or *shock*.

For example, suppose the function $y(t)$ describes the motion of an object with momentum. In the absence of external impulses, if we know the object's position, velocity, and acceleration at time t_0, the Taylor series allows us to estimate its future position $y(t)$ for $t > t_0$. The series builds this estimate based on the assumption that the system behaves smoothly and predictably over the interval.

However, if the object collides with the ground, hits an obstacle, or is struck by another object with significant momentum between t_0 and t, the motion is no longer smooth. The original position, velocity, and acceleration no longer capture the system's behavior after the impact. In such cases, the Taylor approximation fails because the underlying assumptions—smoothness and continuity of derivatives—are violated.

This is why understanding the error in a Taylor series approximation is essential. The error term tells us how far our estimate might stray from the true value and helps us judge whether the approximation is valid over the interval of interest. Large errors signal that the system may have undergone a change—such as a shock—that the Taylor series cannot account for, alerting us to reconsider our modeling approach.

6.2 The error of a first-order Taylor series

For a first-order Taylor series, the error is

$$f(x) = f(x_0) + f^{(1)}(x_0)(x - x_0) + \frac{1}{2}f^{(2)}(\xi)(x - x_0)^2.$$

The exact value of ξ is unknown, but it must lie between x_0 and x, so $x_0 \le \xi \le x$.

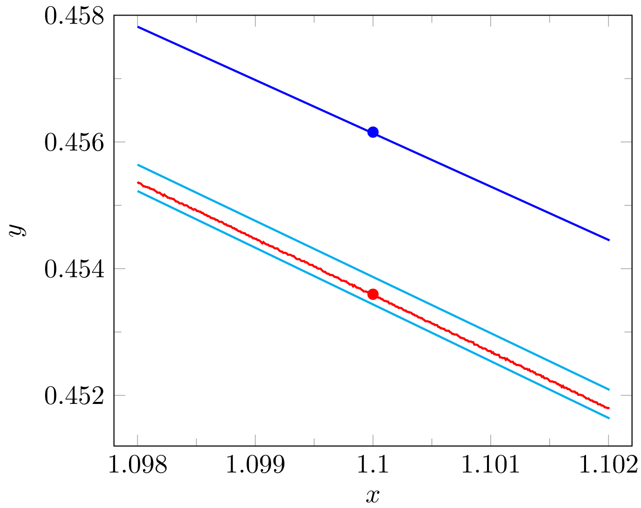
When f is the cosine function, we have $\cos(1.1) \approx \cos(1) - \sin(1) \cdot 0.1$. In reality, $\cos(1.1) = 0.4535961214$ and $\cos(1) - \sin(1) \cdot 0.1 = 0.4561552074$. However, the second derivative of $\cos(x)$ is $-\cos(x)$, and the range is $[-cos(1), -cos(1.1)] \approx [-0.5403, -0.4536]$, and therefore the error of the approximation $\cos(1) - \sin(1) \cdot 0.1$ must fall somewhere between $\frac{1}{2}(-0.4536)0.1^2 = -0.002702$ and $\frac{1}{2}(-0.5403)0.1^2 = -0.002268$. The actual error is

$$\cos(1.1) - (\cos(1) - \sin(1) \cdot 0.1) = -0.0025590860,$$

which falls between the two calculated values.

Below is a plot of the Taylor series approximation (dark blue), the actual function (red), and the Taylor series plus the minimum and maximum values of the second derivative times plus or minus the upper and lower bounds on the error, observing that the actual value of $\cos(1.1)$ does fall between these error bounds.

The first-order Taylor series of $\cos(x)$ at $x_0 = 1$ at $x = 1.1$ with the upper and lower error bounds



When f is the function such that $f(x) = 4xe^{-2x}\sin(3x)$, and we want to approximate $f(1.65)$ given the expansion at $x = 1.6$, we have

$$f(1.65) \approx f(1.6) + f^{(1)}(1.6) \cdot 0.05 = -0.2598775359 + 0.05 \cdot 0.4258113190 = -0.2385869700.$$

The correct value is $f(1.65) = -0.2365892999$, so the error is 0.00199767005.

The minimum and maximum values of the second derivative on the interval $[1.6, 1.65]$ are $[1.277, 1.761]$, so multiplying each of these by $\frac{1}{2}0.05^2$, we get that the error must lie in $[0.001597, 0.002201]$, and we see that the actual error does fall in this interval.

Important: We do not calculate the error of the Taylor series. Instead, the fact we know the bounds of the error gives us confidence that we can use the Taylor series in our approximations when it is appropriate, and that if we do use it, so long as, in this case, the second derivative does not get too large in magnitude between x_0 and x.

6.3 The error of a second-order Taylor series

For a second-order Taylor series, the error is

$$f(x) = f(x_0) + f^{(1)}(x_0)(x - x_0) + \frac{1}{2}f^{(2)}(x_0)(x - x_0)^2 + \frac{1}{6}f^{(3)}(\xi)(x - x_0)^3.$$

When f is the cosine function, we have

$$\cos(1.1) \approx \cos(1) - \sin(1) \cdot 0.1 - \frac{1}{2}\cos(1) \cdot 0.1^2.$$

As above, $\cos(1.1) = 0.4535961214$ and $\cos(1) - \sin(1) \cdot 0.1 - \frac{1}{2}\cos(1) \cdot 0.1^2 = 0.4534536959$. The third derivative of $\cos(x)$ is $\sin(x)$, and the range is $[\sin(1), \sin(1.1)] \approx [0.8415, 0.8912]$, and therefore the error of the approximation $\cos(1) - \sin(1) \cdot 0.1$ must fall somewhere between .0001402 .0001485 $\frac{1}{6}(0.8415)0.1^3 = 0.0001402$ and $\frac{1}{6}(0.8912)0.1^3 = -0.0001485$. The actual error is

$$\cos(1.1) - (\cos(1) - \sin(1) \cdot 0.1 - \frac{1}{2}\cos(1) \cdot 0.1^2) = 0.0001424,$$

which falls between the two calculated values.

When f is the function such that $f(x) = 4xe^{-2x}\sin(3x)$, and we want to approximate $f(1.65)$ given the expansion at $x = 1.6$, we have

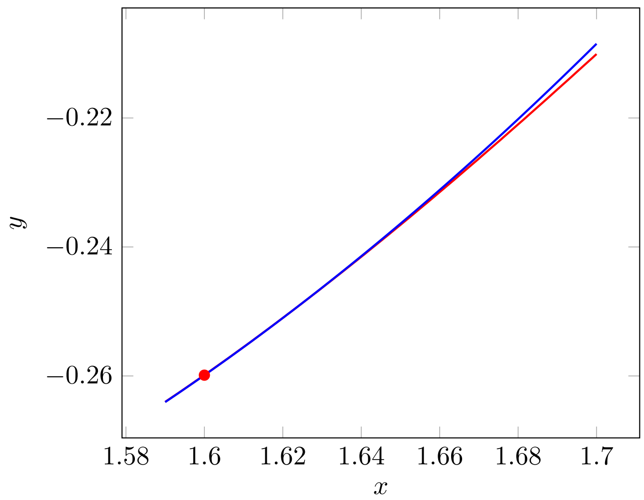
$$\begin{aligned} f(1.65) &\approx f(1.6) + f^{(1)}(1.6) \cdot 0.05 + \frac{1}{2}f^{(2)}(1.6) \cdot 0.05^2 \\ &= -0.2598775359 + 0.4258113190 \cdot 0.05 + \frac{1}{2}1.760762325 \cdot 0.05^2 \\ &= -0.2363860171. \end{aligned}$$

The correct value is $f(1.65) = -0.2365892999$, so the error is -0.0002032828.

The minimum and maximum values of the third derivative on the interval $[1.6, 1.65]$ are $[-0.0002047, -0.0001972]$ so multiplying each of these by $\frac{1}{6}0.05^3$, we get that the error must lie in $[-0.0002047, -0.0001972]$, and we see that the actual error does fall in this interval.

You can see this approximation here:

The second-order Taylor series of $4xe^{-2x}\sin(3x)$ at $x = 1.6$



In this graph, you will see that the second-order Taylor series does match the concavity of the function $4xe^{-2x}\sin(3x)$ close to $x = 1.6$; however, as we move away from that point, the two functions diverge due to the non-quadratic behavior of the function.

Important: Students often ask, "If we can calculate the error, why not use it to find the exact answer?" The key is that we don't know the exact value of ξ where $x_0 \le \xi \le x$, and in practice, higher derivatives may be difficult or impossible to compute. The value of the error term, then, is not to give an exact correction, but to provide confidence that the approximation is close—provided the higher derivatives (like the third derivative in a second-order approximation) are not too large. The Taylor series guarantees that, under such conditions, the approximation should be reasonably accurate.

For example, if you are constantly accelerating a car between the time $t_0 = 0$ s and $t = 3$ s, then knowing the position, speed and that constant rate of acceleration at time $t_0 = 0$ s will allow you to estimate your position reasonably accurately at time $t = 3$ s; however, if between those times you "slam" your foot on the accelerator when $t = 1$ s (not suggested!), the significant resulting change in the acceleration will render the Taylor series approximation useless: the error will be too large.

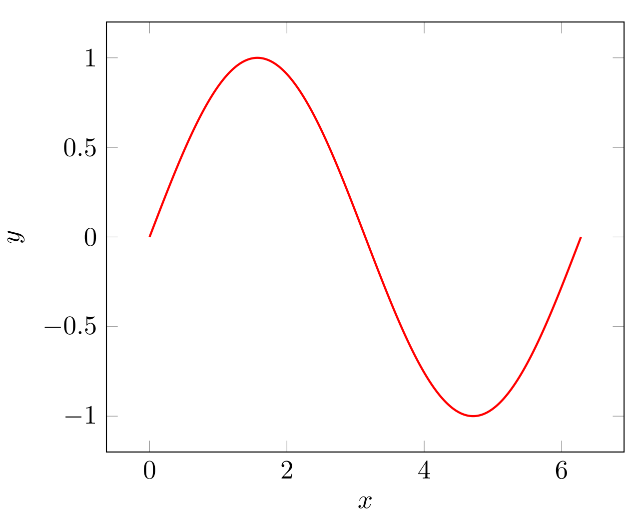
If you were designing a driver-assist program (for example, breaking for an obstacle), then you could use a second-order Taylor series to estimate where the car will be in the next ten seconds, but if after three seconds, the driver then changes the depression of the accelerator or if the car is struck by another vehicle or object, also changing the acceleration, the previous prediction must be flagged as invalid, and it may have to be recomputed. A small change in acceleration may, however, be ignored.

7 Integration

If f is a function defined on the interval $[a, b]$, then the integral of $f(x)$ from $x = a$ to $x = b$ is the area under the curve, where if the function is negative, the "area" is also considered to be negative. This area is written as $\int_a^b f(x)\mathrm{d}x$.

For example, it happens that $\int_0^\pi \sin(x)\mathrm{d}x = 2$, while $\int_\pi^{2\pi} \sin(x)\mathrm{d}x = -2$, and $\int_0^{2\pi} \sin(x)\mathrm{d}x = 0$, as the two areas cancel out.

The trigonometric function $\sin(x)$ on $[0, 2\pi]$

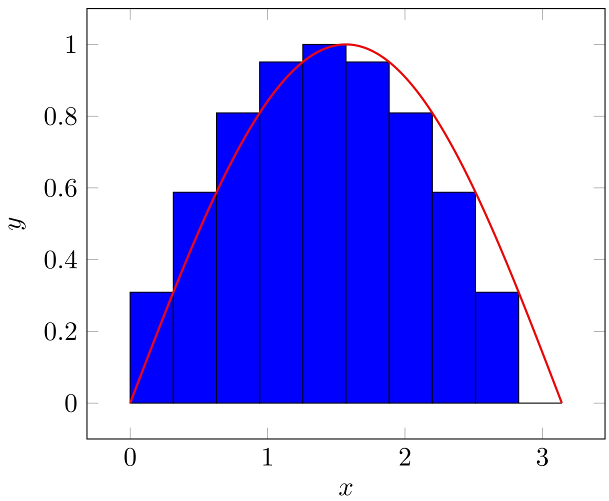


For almost all cases relevant to engineering, we can define the integral as follows: Given a positive integer n, let $h = \frac{b-a}{n}$ and define $x_k = a + kh$, so $x_0 = a$ and $x_n = b$. Then $f(x_k)h$ is the area of a rectangle of width h of height $f(x_k)$, and therefore if we sum all the areas of all these rectangles, then that sum should be an approximation of the integral, so

$$\int_a^b f(x)\mathrm{d}x \approx \sum_{k=1}^{n} f(x_k)h.$$

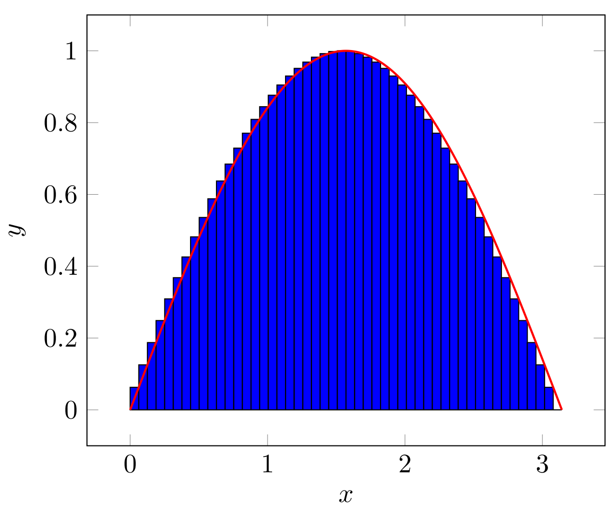
We call this a "Riemann sum." For example, if $n = 10$, the Riemann sum approximation of $\int_0^\pi \sin(x)\mathrm{d}x$ is the area of the blue rectangles shown here:

The Riemann sum approximation of $\sin(x)$ from 0 to π with $n = 10$.



You will notice that sometimes the area is underestimated, and at others it is overestimated. If $n = 50$, we have the area shown here in blue:

The Riemann sum approximation of $\sin(x)$ from 0 to π with $n = 50$.



The integral is the limit as $n \to \infty$, so

$$\int_a^b f(x)\mathrm{d}x = \lim_{n\to\infty} \sum_{k=1}^{n} f(x_k)h.$$

8 Partial derivatives and the gradient

Consider real-valued function f of two independent variables x and y. At a point (x_0, y_0), we can calculate the slope between this point and the point $(x_0 + h, y_0)$, which is in the direction of x, by

$$\frac{f(x_0+h, y_0) - f(x_0, y_0)}{(x_0+h) - x_0} = \frac{f(x_0+h, y_0) - f(x_0, y_0)}{h}.$$

The partial derivative at (x_0, y_0) in the direction of x is written as $\frac{\partial f}{\partial x}$ and can be found by taking the limit of the above slope:

$$\frac{\partial}{\partial x} f(x_0, y_0) = \lim_{h\to 0} \frac{f(x_0+h, y_0) - f(x_0, y_0)}{h}.$$

Similarly, we can find the partial derivative with respect to y (that is, the slope in the direction of y) as follows:

$$\frac{\partial}{\partial y} f(x_0, y_0) = \lim_{h\to 0} \frac{f(x_0, y_0+h) - f(x_0, y_0)}{h}.$$

To find the partial derivative of a function with respect to x, treat all ys as constant and simply differentiate with respect to x, with a similar calculation for finding the partial derivative with respect to y. Thus, if $f(x,y) = x^3 - 4x^2y + 2xy^2 + 5y^3$, then

$$\frac{\partial}{\partial x} f(x,y) = 3x^2 - 8xy + 2y^2 \text{ and } \frac{\partial}{\partial y} f(x,y) = -4x^2 + 4y + 15y^2.$$

The gradient of this real-valued function of two real variables is the two-dimensional vector containing the two partial derivatives, and it is written as $\vec{\nabla} f = \begin{pmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{pmatrix}$.

At a point (x_0, y_0), the it happens that the gradient gives the direction of maximum increase, and and the tangent plane at that point is defined as

$$p(x_0 + \Delta x, y_0 + \Delta y) = f(x_0, y_0) + \vec{\nabla} f(x_0, y_0) \cdot \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix},$$

where $\cdot$ is the dot product, so if you were to move in the direction $(\Delta x, \Delta y)$ at the same rate of increase as you experienced at (x_0, y_0), then your change in height would be exactly $\vec{\nabla} f(x_0, y_0) \cdot \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix}$.

Thus, if Δx and Δy are very small, we have an equivalent statement to the Taylor series:

$$f(x_0 + \Delta x, y_0 + \Delta y) \approx f(x_0, y_0) + \vec{\nabla} f(x_0, y_0) \cdot \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix},$$

For example, if $f(x,y) = x^3 - 4x^2y + 2xy^2 + 5y^3$, then $f(1,2) = 41$ and

$$\vec{\nabla} f(x,y) = \begin{pmatrix} 3x^2 - 8xy + 2y^2 \\ -4x^2 + 4y + 15y^2 \end{pmatrix}$$

and thus the gradient at $(1,2)$ is

$$\vec{\nabla} f(1,2) = \begin{pmatrix} -5 \\ 64 \end{pmatrix}.$$

We note that if we move in the direction of $(0.9, 2.1)$ that

$$f(0.9, 2.1) \approx 41 + \begin{pmatrix} -5 \\ 64 \end{pmatrix} \cdot \begin{pmatrix} -0.1 \\ 0.1 \end{pmatrix} = 47.9,$$

and this is close to $f(0.9, 2.1) = 48.168$, but the direction did not change as much as when we moved the same distance in the direction of the gradient $\begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix} = \begin{pmatrix} -0.0110149795 \\ 0.1409917381 \end{pmatrix}$, which yields the maximum change at

$$f(0.9889850205, 1.1409917381) \approx 41 + \begin{pmatrix} -5 \\ 64 \end{pmatrix} \cdot \begin{pmatrix} -0.01101497954 \\ 0.1409917381 \end{pmatrix} = 50.07854614.$$

9 Acknowledgments

Many thanks to Zarra Sarker for finding various mistakes and typos, as well as making valuable suggestions to make the material clearer for a second-year engineering student.