This course requires a familiarity with basic calculus, some multivariate calculus, linear algebra, and some basic notions of metric topology. We will review these concepts throughout the class as they are needed, starting with calculus. Calculus is needed to prove that numerical algorithms are correct and converge at an appropriate rate. It is also essential in understanding probability distributions over continuous variables.

These notes will also introduce the concept of a Taylor series, which is a widely used concept in numerical analysis that will appear many times during this course.

1 Convergence and limits

Convergence describes when a sequence of numbers “approaches” another number arbitrarily closely, or when a function \( f(x) \) “approaches” a value \( y \) at some point \( x \) where \( f(x) \) is not well-defined. To define these notions rigorously requires some rather subtle math from the field of real analysis. (These subtleties are crucial to deal appropriately with the strange things that can happen when dealing with infinities).

\textbf{Cauchy series.} An infinite sequence of real numbers \( a_1, a_2, a_3, \ldots \) is called \textit{Cauchy convergent} when the distance between subsequent numbers becomes arbitrarily small as the index increases. More precisely, for any real number \( \epsilon > 0 \), there exists an index \( N \) such that \( |a_m - a_n| < \epsilon \) for all \( m, n > N \).

This is significant because \( \epsilon \) can be chosen arbitrarily close to 0, and all numbers in the sequence with index \( > N \) will be closer than \( \epsilon \). (Note that \( N \) is dependent on \( \epsilon \).)

Examples:
• The sequence \( a_n = 1/n \) for \( n = 1, 2, \ldots \) is Cauchy convergent.

• \( a_n = e^{-n} \) is Cauchy convergent.

• \( a_n = (\sin(0.15n))/n \) is Cauchy convergent.

• Counterexample: \( a_n = n \mod 2 \) is not Cauchy convergent.

**Limit of a sequence.** \( L \) is said to be the *limit* of a convergent sequence \( a_1, a_2, \ldots \) if for any real number \( \epsilon > 0 \), there exists an index \( N \) such that \(|a_n - L| < \epsilon\) for all \( n > N \). If a sequence \( a_1, a_2, \ldots \) has a limit \( L \), then we write

\[
\lim_{n \to \infty} a_n = L
\]  

Sequences that are not Cauchy convergent do not have limits. But not all Cauchy convergent sequences from a certain domain have a limit in that domain. For example, consider the domain of positive numbers \( \mathbb{R}^+ = \{x| x > 0\} \). The sequence \( a_n = 1/n \) consists of numbers from \( \mathbb{R}^+ \), but there is no number in \( \mathbb{R}^+ \) that is a limit of the sequence.

(A side note: a set in which all Cauchy sequences have a limit in the set is known as *compact*.)

**Limit of a real-valued function.** The limit of a real-valued function is a similar concept, but requires a slightly different definition. Let \( f(x) \) be a function on the reals, and let \( c \) be a real number. Then \( L \) is the *limit of \( f(x) \) at \( c \)* if \( f(x) \) becomes arbitrarily close to \( L \) as \( x \) approaches \( c \) (from both sides). More precisely, for any \( \epsilon > 0 \), there exists a \( \delta \) such that \(|f(x) - L| \leq \epsilon\) for all \( x \) that satisfy \( 0 < |x - c| < \delta \). If a limit exists, we write this as

\[
\lim_{x \to c} f(x) = L
\]  

Note that in this definition, \( \epsilon \) can be chosen arbitrarily close to zero, and that \( \delta \) is chosen depending on the choice of \( \epsilon \). Note also that \( f(c) \) does not have to be defined. For example, if \( f(x) = (\sin x)/x \), then \( f(0) \) is not defined because it evaluates to 0/0. But, you can show that \( f(x) \) approaches 1 as \( x \) approaches 0 (at the moment it is not obvious how one would prove this).

**One-sided and asymptotic limits.** There are also the notions of one-sided limits, in which \( f(x) \) is required to approach \( L \) only as \( x \) increases toward \( c \) (resp., decreases toward \( c \)). Limits can also be *asymptotic* as \( x \) approaches
The precise definition of an asymptotic limit is similar to the definition of the limit of a sequence.

**Continuity.** A real function \( f(x) \) is continuous over the interval \( I \) if and only if \( \lim_{x \to c} f(x) = f(c) \) for all \( c \in I \). (More specifically, this is known as pointwise continuity. There are other subtle definitions of continuity as well.)

**Minimum and maximum.** For a set \( S \subseteq \mathbb{R} \), a minimum of the set is an element \( x \in S \) such that \( x \leq y \) for all \( y \in S \). A similar definition holds for the maximum, but with \( x \) and \( y \) switched.

Any finite set has a maximum and a minimum, but infinite sets do not necessarily have either a minimum or a maximum. For example, the interval \( S = (0, 1] \) has a maximum, namely 1, but it has no minimum, because no matter what value \( x \) is picked, you can pick another value in \( S \) that is smaller, for example, \( x/2 \). Hence, we must use an alternate definition for the intuition that 0 is the tightest lower bound for \( S \). This gives rise to the definition of infimum and supremum.

**Infimum and supremum.** The infimum of a set \( S \subseteq \mathbb{R} \) is the largest number \( x \in \mathbb{R} \) such that \( x \leq y \) for all \( y \in S \). The supremum of a set \( S \subseteq \mathbb{R} \) is the smallest number \( x \in \mathbb{R} \) such that \( y \leq x \) for all \( y \in S \).

## 2 Derivatives

The derivative \( f'(a) \) of a real function \( f(x) \) at \( x = a \) is given by the limit

\[
f'(a) = \frac{df}{dx}(a) = \lim_{h \to 0} \frac{f(a + h) - f(a)}{h}.
\]  

(3)

Strictly speaking, we require the limit to exist as \( h \) approaches zero from both above and below. If the limit exists, we say \( f \) is differentiable at \( a \). The notation \( f'(x) \) and \( \frac{df}{dx}(a) \) are two ways of writing the same thing and should be understood to be equivalent.

**Differentiation operator.** The derivative of \( f \) refers to the entire function \( f' \), and the \( \cdot' \) operation is known as differentiation. This operator is also written \( \frac{d}{dx} \).

**Continuity and Differentiability.** Differentiable functions are continuous, but the converse is not necessarily true (e.g., \( f(x) = |x| \)). If \( f \) is continuous, we call it \( C0 \) continuous. If its derivative exists, and \( f' \) is continuous, then
we say $f$ is $C^1$ continuous. If the second derivative of $f$ exists, then $f$ is $C^2$
continuous, and so on for higher derivatives. If all of $f$’s higher derivatives
exist then it is called infinitely differentiable.

2.1 Frequently used differentiation rules

You are responsible for knowing the derivatives of elementary functions (poly-
nomials, sin, cos, exp, and log). We will often use the following differentiation
rules:

- Chain rule: if $h(x) = f(g(x))$, then $h'(x) = f'(g(x))g'(x)$.
- Multiplication rule: if $h(x) = f(x)g(x)$, then $h'(x) = f'(x)g(x) + f(x)g'(x)$.
- Inverse rule: if $h(x) = f^{-1}(x)$, then $h'(x) = 1/f'(f^{-1}(x))$.
- L’Hôpital’s rule: if $\lim_{x \to c} f(x) = \lim_{x \to c} g(x) = 0$ or $\pm\infty$ and $g'(x) \neq 0$
on an open interval containing $c$, then $\lim_{x \to c} f(x)/g(x) = \lim_{x \to c} f'(x)/g'(x)$
if the limit is well defined.

3 Taylor series

The Taylor series is an important tool in numerical analysis and will be used
multiple times throughout this course.

**Taylor series.** The Taylor series of a function $f(x)$ in the neighborhood of
$a$ is given by the following infinite series:

$$g_\infty(x) = f(a) + \sum_{n=1}^{\infty} f^{(n)}(a) \frac{(x - a)^n}{n!}$$

(4)

where $f^{(n)}(a)$ is the $n$’th derivative of $f$ at $a$. For a certain class of functions
(analytic functions), $g_\infty$ is exactly $f$. For general functions, however, equality
does not hold.

**kth order Taylor polynomial.** It is common to use a truncation of the
Taylor series to approximate $f$. Consider truncating the summation to the
first $k$ derivatives:

$$g_k(x) = f(a) + \sum_{n=1}^{k} f^{(n)}(a) \frac{(x - a)^n}{n!}.$$
This is known as the Taylor polynomial of degree \( k \). For example, the 1st order Taylor polynomial has a graph that is simply a line through the point \((a, f(a))\) with the slope \( f'(a) \). The 2nd order Taylor polynomial is a quadratic curve that is tangent to the graph of \( f \) at \( a \), and also has the same second derivative as \( f \) at \( a \).

It is natural to ask how well the Taylor polynomial approximates a given function? For analytic functions, \( g_k \) approaches \( f \) as \( k \) grows larger. Note that the error is not uniform, and typically grows as \( x \) moves farther from \( a \). This pattern is typical even for non-analytic functions as well. Taylor’s theorem gives us a way of quantifying the approximation error.

**Taylor’s theorem.** Let \( f(x) \) be \( k \) times differentiable at \( x = a \). Then there exists a function \( h(x) \) such that

\[
f(x) = f(a) + \sum_{n=1}^{k} \frac{f^{(n)}(a)(x-a)^n}{n!} + h(x)(x-a)^n
g_k(x)
\]

and \( \lim_{x \to a} h(x) = 0 \).

Moreover, the exact form of \( h \) can be derived more precisely under the assumption that \( f(x) \) is \( k + 1 \) times differentiable on the interval between \( x \) and \( a \). In this case, it can be proven that

\[
f(x) - g_k(x) = f^{(k+1)}(\xi)(x-a)^{k+1}/(k+1)!
g_k(x)
\]

for some value \( \xi \) on the interval between \( x \) and \( a \). In other words, if \( f \) is smooth (its higher derivatives are small) then the Taylor approximation provides a good fit in a neighborhood \( a \).

### 4 Exercises

1. Prove the fact that if a sequence is not Cauchy convergent, then it does not have a limit.

2. Is \( \sin(1/x) \) continuous on \((0, \infty)\)? Is it differentiable? Does it have a limit as \( x \to 0 \)?

3. Use L’Hôpital’s rule to find the limit of \( (\sin x)/x \) as \( x \to 0 \).
4. Find the 2nd-order Taylor approximation of the function $1/(x + 1)$ centered at the point $x = 0$.

5. What is the Taylor series of a degree $k$ polynomial?