Calculus

1. **Limits, Series**
2. **Basic derivatives**
3. **Minima Maxima**
4. **Multivariate**
5. **Integral**
6. **Numerical Optimization**
7. **Matrix Calculus**

Saved memory full

Rule Name	Formula
Constant Rule	$\frac{d}{dx}(c) = 0$
Power Rule	$\frac{d}{dx}(x^n) = nx^{n-1}$
Constant Multiple Rule	$\frac{d}{dx}(c \cdot f(x)) = c \cdot f'(x)$
Sum Rule	$\frac{d}{dx}[f(x) + g(x)] = f'(x) + g'(x)$
Difference Rule	$\frac{d}{dx}[f(x) - g(x)] = f'(x) - g'(x)$
Product Rule	$\frac{d}{dx}[f(x)g(x)] = f'(x)g(x) + f(x)g'(x)$
Quotient Rule	$\frac{d}{dx}\left[\frac{f(x)}{g(x)}\right] = \frac{f'(x)g(x)-f(x)g'(x)}{g(x)^2}$
Chain Rule	$\frac{d}{dx}[f(g(x))] = f'(g(x)) \cdot g'(x)$

🧪 Common Derivatives:

Function	Derivative
x^n	nx^{n-1}
$\sin(x)$	$\cos(x)$
$\cos(x)$	$-\sin(x)$
$\tan(x)$	$\sec^2(x)$
e^x	e^x
$\ln($	

Calculus Cheat Sheet with Numerical Examples

---

1. Limits & Series

Definition:

- **Limit:** $\lim_{x \to a} f(x) = L$ means $f(x)$ approaches L as x approaches a.
- **Series:** Sum of infinite terms (e.g., $\sum_{n=1}^{\infty} a_n$).

Example (Limit):

Compute $\lim_{x \to 2} \frac{x^2-4}{x-2}$.
Solution:

$$\frac{x^2 - 4}{x - 2} = \frac{(x-2)(x+2)}{x-2} = x + 2 \quad (\text{for } x \neq 2)$$

$$\lim_{x \to 2}(x+2) = 4$$

Example (Series):

Does $\sum_{n=1}^{\infty} \frac{1}{n^2}$ converge?
Solution:
Yes (by **p-test**, since $p = 2 > 1$). Its sum is $\frac{\pi^2}{6}$.

---

2. Basic Derivatives

Definition:

$$f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}$$

Example:

Find $\frac{d}{dx}(3x^2 + \sin x)$.
Solution:

Example:

Find $\frac{d}{dx}(3x^2 + \sin x)$.
Solution:

$$\frac{d}{dx}(3x^2) = 6x, \quad \frac{d}{dx}(\sin x) = \cos x$$
$$\implies f'(x) = 6x + \cos x$$

---

3. Minima & Maxima

Key Idea:

- Find critical points where $f'(x) = 0$.
- Use **second derivative test**:
  - $f''(x) > 0$: Local minimum.
  - $f''(x) < 0$: Local maximum.

Example:

Find extrema of $f(x) = x^3 - 6x^2 + 9x$.
Solution:

1. $f'(x) = 3x^2 - 12x + 9 = 0 \implies x = 1, 3.$
2. $f''(x) = 6x - 12.$
   - At $x = 1$: $f''(1) = -6 < 0$ (**max**).
   - At $x = 3$: $f''(3) = 6 > 0$ (**min**).

---

4. Multivariate Calculus

Partial Derivatives:

$$\frac{\partial}{\partial x} f(x, y), \quad \frac{\partial}{\partial y} f(x, y)$$

Example:

Find $\frac{\partial}{\partial x}(x^2 y + \sin y)$.
Solution:

$$\frac{\partial}{\partial x}(x^2 y) = 2xy, \quad \frac{\partial}{\partial x}(\sin y) = 0$$
$$\implies \frac{\partial f}{\partial x} = 2xy$$

Gradient:

$$\nabla f = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \right)$$

---

5. Integrals

Definition:

$$\int_a^b f(x)\,dx = \text{Area under curve}$$

Example:

Compute $\int_0^2 (3x^2 + 1)\,dx$.
Solution:

$$\int (3x^2 + 1)\,dx = x^3 + x + C$$
$$\text{Evaluate at bounds: } (8 + 2) - (0 + 0) = 10$$

---

6. Numerical Optimization

Gradient Descent:

Iteratively move in the direction of steepest descent:

$$x_{n+1} = x_n - \alpha \nabla f(x_n)$$

2. $x_2 = 1.6 - 0.1 \cdot 3.2 = 1.28$

... converges to $x = 0$.

---

7. Matrix Calculus

Gradient of a Quadratic Form:

For $f(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$:

$$\nabla f = (A + A^T)\mathbf{x}$$

Example:

Let $A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$, $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$.

$$f(\mathbf{x}) = 2x_1^2 + 2x_1x_2 + 2x_2^2$$

$$\nabla f = \begin{bmatrix} 4x_1 + 2x_2 \\ 2x_1 + 4x_2 \end{bmatrix}$$

Summary Table

Concept	Key Formula	Example
Limits	$\lim_{x \to a} f(x)$	$\lim_{x \to 0} \frac{\sin x}{x} = 1$
Derivatives	$\frac{d}{dx} x^n = nx^{n-1}$	$\frac{d}{dx} e^x = e^x$
Min/Max	$f'(x) = 0$, $f''(x)$ test	$f(x) = x^2$: min at $x = 0$
Multivariate	$\nabla f = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right)$	$\nabla(x^2 + y^2) = (2x, 2y)$
Integrals	$\int x^n \, dx = \frac{x^{n+1}}{n+1} + C$	$\int_0^1 x \, dx = \frac{1}{2}$
Optimization	$x_{n+1} = x_n - \alpha \nabla f$	Gradient descent on $f(x) = x^2$
Matrix Calc	$\nabla(\mathbf{x}^T A\mathbf{x}) = (A + A^T)\mathbf{x}$	Quadratic forms

---

🧮 Notation:

If $f(x, y)$ is a function of two variables:

- Partial derivative w.r.t. x:

$$\frac{\partial f}{\partial x}$$

- Partial derivative w.r.t. y:

$$\frac{\partial f}{\partial y}$$

📌 Example:

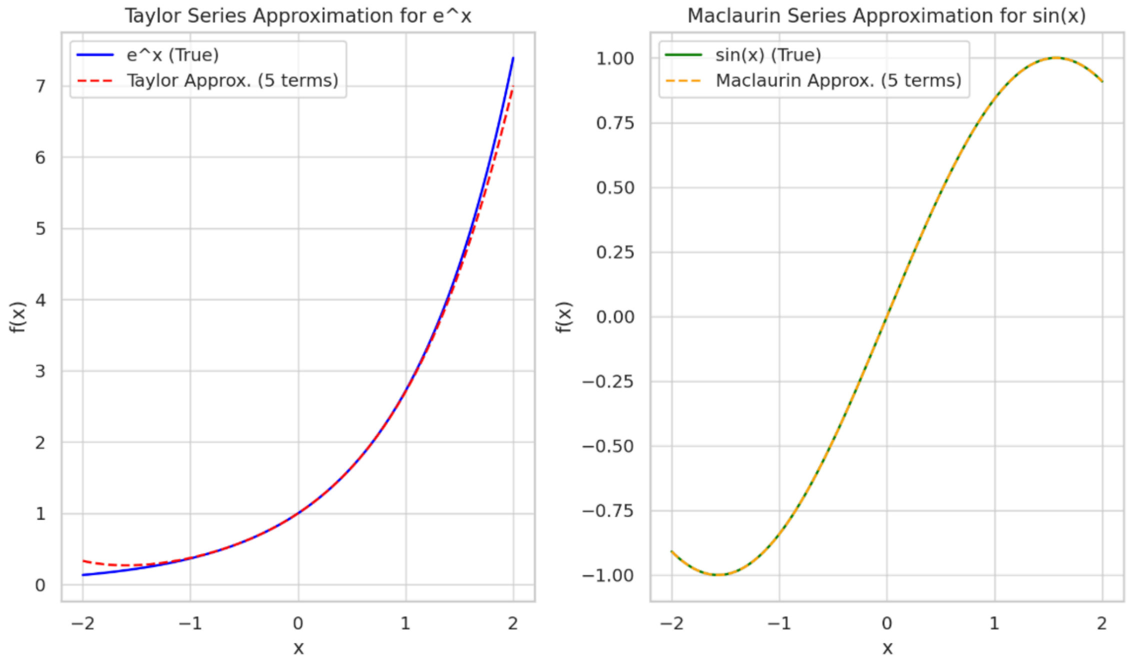
Let $f(x, y) = x^2y + 3xy^2$

- $\frac{\partial f}{\partial x} = 2xy + 3y^2$
- $\frac{\partial f}{\partial y} = x^2 + 6xy$

👉 When you take ∂/∂x, treat y **as a constant** (and vice versa).

📊 Use in ML:

- **Gradient Descent** (optimizing loss functions)
- **Backpropagation** in Neural Networks
- **Surface analysis** in optimization



Here are the visualizations for both the Taylor and Maclaurin series:

1. **Taylor Series Approximation for e^x**: The blue curve represents the true function e^x, while the red dashed curve shows the approximation from the Taylor series expansion (using 5 terms) around $x = 0$.
2. **Maclaurin Series Approximation for $\sin(x)$**: The green curve represents the true function $\sin(x)$, and the orange dashed curve is the approximation from the Maclaurin series (5 terms).

Maclaurin Series vs Taylor Series

Both Maclaurin and Taylor series are ways of approximating a function using polynomials, but they differ in the points at which they approximate the function.

Taylor Series:

- **Definition**: The Taylor series of a function $f(x)$ about a point a is an infinite sum of terms calculated from the values of the function's derivatives at a.

  Formula:

  $$f(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \frac{f^{(3)}(a)}{3!}(x-a)^3 + \dots$$

- **General Use**: The Taylor series can be used for any point a where the function is differentiable.

  Example: The Taylor series of e^x around $a = 0$ is:

  $$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots$$

Maclaurin Series:

- **Definition**: The Maclaurin series is a special case of the Taylor series where the expansion is done around $a = 0$. In other words, it's a Taylor series centered at zero.

  Formula:

  $$f(x) = f(0) + f'(0)x + \frac{f''(0)}{2!}x^2 + \frac{f^{(3)}(0)}{3!}x^3 + \dots$$

- **General Use**: The Maclaurin series is useful when the function is known to be well-behaved near $x = 0$.

  Example: The Maclaurin series for $\sin(x)$ is:

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

Series & Sequences

- **Sequence**: A sequence is an ordered list of numbers, typically represented as $a_1, a_2, a_3, \ldots$. Each number in the sequence is called a "term." A sequence can be finite or infinite.

  Example:

  $$1, 2, 3, 4, \ldots \quad \text{(This is a simple sequence of natural numbers.)}$$

- **Series**: A series is the sum of the terms of a sequence. For example, the sum of the sequence $1, 2, 3, \ldots$ would be the series $1 + 2 + 3 + \ldots$.

  Example:

  $$S = \sum_{n=1}^{\infty} \frac{1}{2^n} = \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \cdots = 1$$

  This is an infinite geometric series.

---

1. Limits & Series

Limits:

- **Intuition**: Limits describe how a function behaves as its input approaches a specific point. For example, as x approaches 2, $f(x) = x^2$ approaches 4.

  $$\lim_{x \to 2} x^2 = 4$$

Series:

- **Intuition**: A series is the sum of terms of a sequence. The **Taylor series** expands a function into an infinite sum based on its derivatives.

  Example: The **Maclaurin series** for e^x:

  $$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots$$

---

2. Basic Derivatives

Intuition: The derivative represents the slope or rate of change of a function.

Example: For $f(x) = x^2$, the derivative is:

$$f'(x) = 2x$$

- At $x = 3$, the slope is $2(3) = 6$.

---

3. Minima & Maxima

Intuition: To find the peaks or valleys of a function, set the derivative to 0 (critical points) and use the second derivative to classify them.

Example: For $f(x) = x^2$:

- $f'(x) = 2x$, set $2x = 0$, so $x = 0$ is a critical point.
- $f''(x) = 2$ (positive), so $x = 0$ is a **local minimum**.

---

4. Multivariate Calculus

Intuition: In multivariate calculus, we take partial derivatives to find how a function changes with respect to each variable, holding others constant.

Example: For $f(x, y) = x^2 + y^2$, the partial derivatives are:

$$\frac{\partial f}{\partial x} = 2x, \quad \frac{\partial f}{\partial y} = 2y$$

The **gradient** $\nabla f = (2x, 2y)$ points in the direction of maximum increase.

---

5. Integral

Intuition: An integral represents the area under a curve (accumulation of quantities).

Example: To find the area under $f(x) = x$ from 0 to 2:

$$\int_0^2 x\,dx = \left[\frac{x^2}{2}\right]_0^2 = \frac{4}{2} - 0 = 2$$

Matrix calculus is an extension of regular calculus that deals with **derivatives of functions involving vectors and matrices**. It's crucial in machine learning, especially in optimization (like gradient descent).

🔢 Matrix Calculus

It's the extension of regular calculus to **vectors and matrices**, essential in **machine learning, deep learning, and optimization**.

📌 Core Concepts:

1. Gradient (∇)

- For scalar function $f(\mathbf{x})$, where $\mathbf{x}$ is a vector:

$$\nabla f = \left[\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \ldots, \frac{\partial f}{\partial x_n}\right]^T$$

2. Jacobian (J)

- For vector-valued function $\mathbf{f}(\mathbf{x})$:
  A matrix of partial derivatives:

$$J_{ij} = \frac{\partial f_i}{\partial x_j}$$

3. Hessian (H)

- For second-order derivatives of scalar function:

$$H_{ij} = \frac{\partial^2 f}{\partial x_i \partial x_j}$$

🧠 Common Matrix Derivatives:

Expression	Derivative w.r.t. $\mathbf{x}$
$\mathbf{a}^T\mathbf{x}$	$\mathbf{a}$
$\mathbf{x}^T\mathbf{x}$	$2\mathbf{x}$
$\mathbf{x}^T A\mathbf{x}$	$(A + A^T)\mathbf{x}$ (if A not symmetric)
$\log(\det(X))$	$(X^{-1})^T$
$\mathrm{tr}(AX)$	A^T

📊 Use in Machine Learning:

- Backpropagation
- Gradient descent (vectorized)
- Loss function optimization
- Multivariate regression

Saved memory full

🧮 Example:

Let

$$f(\mathbf{x}) = \mathbf{x}^T A \mathbf{x}$$

Where

$$\mathbf{x} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \quad A = \begin{bmatrix} 3 & 1 \\ 1 & 2 \end{bmatrix}$$

✅ Step 1: Compute the function value

$$f(\mathbf{x}) = \begin{bmatrix} 1 & 2 \end{bmatrix} \begin{bmatrix} 3 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 & 2 \end{bmatrix} \begin{bmatrix} 3*1+1*2 \\ 1*1+2*2 \end{bmatrix} = \begin{bmatrix} 1 & 2 \end{bmatrix} \begin{bmatrix} 5 \\ 5 \end{bmatrix} = 1*5+2*5 = 15$$

✅ Step 2: Compute the Gradient

The gradient of $f(\mathbf{x}) = \mathbf{x}^T A \mathbf{x}$ is:

$$\nabla f = (A + A^T)\mathbf{x}$$

Since A is symmetric:

$$\nabla f = 2A\mathbf{x} = 2 \begin{bmatrix} 3 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = 2 \begin{bmatrix} 3*1+1*2 \\ 1*1+2*2 \end{bmatrix} = 2 \begin{bmatrix} 5 \\ 5 \end{bmatrix} = \begin{bmatrix} 10 \\ 10 \end{bmatrix}$$

🎯 Result:

- $f(\mathbf{x}) = 15$
- $\nabla f(\mathbf{x}) = \begin{bmatrix} 10 \\ 10 \end{bmatrix}$

✅ Indefinite Integral

It **does not have limits.**
It gives a **family of functions** (general solution).

Notation:

$$\int f(x)\,dx$$

Output:

$$F(x) + C$$

Where C is a constant of integration.

Example:

$$\int x^2\,dx = \frac{x^3}{3} + C$$

✅ Definite Integral

It **has limits** (from a to b).
It gives a **number** (area under the curve).

Notation:

$$\int_a^b f(x)\,dx$$

Output:

$$F(b) - F(a)$$

Example:

$$\int_1^2 x^2\,dx = \left[\frac{x^3}{3}\right]_1^2 = \frac{8}{3} - \frac{1}{3} = \frac{7}{3}$$

💡 In Short:

Feature	Indefinite Integral	Definite Integral
Limits of integration	❌ None	✅ Yes (from a to b)
Output	Function + constant C	Number (area)
Use	General antiderivative	Area, total accumulation

1. Jacobian Matrix

The **Jacobian** is a matrix of first-order partial derivatives of a vector-valued function. It describes the rate of change of each component of the output with respect to each input.

Notation:

If $\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$ is a vector-valued function of $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$, the Jacobian matrix J is:

$$J = \begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_1}{\partial x_2} \\ \frac{\partial y_2}{\partial x_1} & \frac{\partial y_2}{\partial x_2} \end{bmatrix}$$

Example:

Let $\mathbf{y} = \begin{bmatrix} x^2 + y \\ x + y^2 \end{bmatrix}$, where x and y are inputs. The Jacobian matrix J would be:

$$J = \begin{bmatrix} \frac{\partial(x^2+y)}{\partial x} & \frac{\partial(x^2+y)}{\partial y} \\ \frac{\partial(x+y^2)}{\partial x} & \frac{\partial(x+y^2)}{\partial y} \end{bmatrix} = \begin{bmatrix} 2x & 1 \\ 1 & 2y \end{bmatrix}$$

2. Hessian Matrix

The **Hessian** is a square matrix of second-order partial derivatives of a scalar function. It represents the curvature of the function and helps understand its behavior around a point.

Notation:

For a scalar function $f(x, y)$, the Hessian matrix H is:

$$H = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} \end{bmatrix}$$

Example:

Let $f(x, y) = x^2 + 3xy + y^2$. The Hessian matrix H would be:

$$H = \begin{bmatrix} \frac{\partial^2(x^2+3xy+y^2)}{\partial x^2} & \frac{\partial^2(x^2+3xy+y^2)}{\partial x \partial y} \\ \frac{\partial^2(x^2+3xy+y^2)}{\partial y \partial x} & \frac{\partial^2(x^2+3xy+y^2)}{\partial y^2} \end{bmatrix} = \begin{bmatrix} 2 & 3 \\ 3 & 2 \end{bmatrix}$$

$$H = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} \end{bmatrix}$$

Where:

- $\frac{\partial^2 f}{\partial x^2}$ is the second derivative of f with respect to x.
- $\frac{\partial^2 f}{\partial y^2}$ is the second derivative of f with respect to y.
- $\frac{\partial^2 f}{\partial x \partial y}$ and $\frac{\partial^2 f}{\partial y \partial x}$ are the mixed partial derivatives.

Example

Let’s take a simple scalar function:

$$f(x, y) = x^2 + 3xy + y^2$$

Now, let's compute the Hessian matrix for this function:

1. **First derivatives:**
   - $\frac{\partial f}{\partial x} = 2x + 3y$
   - $\frac{\partial f}{\partial y} = 3x + 2y$
2. **Second derivatives:**
   - $\frac{\partial^2 f}{\partial x^2} = 2$ (the second derivative of f with respect to x)
   - $\frac{\partial^2 f}{\partial y^2} = 2$ (the second derivative of f with respect to y)
   - $\frac{\partial^2 f}{\partial x \partial y} = 3$ (the mixed derivative, first x then y)
   - $\frac{\partial^2 f}{\partial y \partial x} = 3$ (the mixed derivative, first y then x)

The **Hessian matrix** will thus be:

$$H = \begin{bmatrix} 2 & 3 \\ 3 & 2 \end{bmatrix}$$