

explain these topics one by one try to cover hidden aspects of them
Multivariable Calculus
Partial derivatives
Gradients, Jacobians, Hessians
Double & triple integrals
Vector calculus (divergence, curl, line integrals — optional depth)

Great — let's build **Multivariable Calculus** from the ground up and go *beyond textbook explanations*, showing the hidden structure behind each concept.

I'll cover them **one by one**, with intuition + geometry + formulas + why they matter in ML/physics.

1 PARTIAL DERIVATIVES — Hidden View

What they usually tell you

A partial derivative is the derivative of a multivariable function w.r.t. one variable while keeping others constant.

Example:

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

$$\frac{\partial f}{\partial x} = 2xy, \quad \frac{\partial f}{\partial y} = x^2 + 3$$

Hidden aspects (the deeper layer)

(A) Partial derivative = slicing a surface

A multivariable function $f(x, y)$ is a **surface**.

A partial derivative is the slope when you slice this surface with a plane:

- $\frac{\partial f}{\partial x}$ = slope along the **x-direction**, slice parallel to xz-plane
- $\frac{\partial f}{\partial y}$ = slope along the **y-direction**, slice parallel to yz-plane

This is why it feels like 1D derivative inside a higher dimension.

(B) They depend on coordinate choice

Hidden fact:

Partial derivatives are *not invariant* — they depend on how you choose the axes.

If you rotate the coordinate system, partial derivatives **change**, because "holding y constant" means something different.

This is why physicists prefer **gradient vectors** (coordinate independent).

(C) Mixed partials symmetry (Clairaut's theorem)

For smooth functions:

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x}$$

This symmetry turns out to be fundamental:

- makes Hessian matrices symmetric
 - enables optimization algorithms (Newton's method, convexity tests)
-

(D) Partial derivatives \neq total derivative

For a function $f(x(t), y(t))$, movement is along a *curve*.

Total derivative:

$$\frac{df}{dt} = f_x \frac{dx}{dt} + f_y \frac{dy}{dt}$$

Partial derivatives alone do **not** predict motion along arbitrary directions — that job belongs to gradients.

2 GRADIENTS, JACOBIANS, HESSIANS

Gradient — The Direction of Maximum Increase

Definition

For $f : \mathbb{R}^n \rightarrow \mathbb{R}$:

$$\nabla f = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{bmatrix}$$

Hidden aspects

(A) Gradient is perpendicular to level sets

For $f(x, y) = c$, the gradient is **normal** to the curve.

This is extremely important:

- in ML \rightarrow gradient points toward steepest ascent
 - in physics \rightarrow gradient is force direction
 - in PDEs \rightarrow heat flows perpendicular to isotherms
-

(B) Magnitude = steepest slope

The length $\|\nabla f\|$ = how steep the climb is.

(C) Directional derivative

In direction unit vector u :

$$D_u f = \nabla f \cdot u$$

This shows gradient encodes *all possible slopes* in all directions.

JACOBIAN — Derivative of a Vector Function

If $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$,

$$J_F = \begin{bmatrix} \frac{\partial F_1}{\partial x_1} & \dots & \frac{\partial F_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial F_m}{\partial x_1} & \dots & \frac{\partial F_m}{\partial x_n} \end{bmatrix}$$

Hidden aspects

(A) Jacobian \approx best linear transformation

Every nonlinear mapping locally behaves like a matrix:

$$F(x + \Delta x) \approx F(x) + J_F(x)\Delta x$$

This is the foundation of:

- backpropagation
 - linearization in control theory
 - manifold transformations
-

(B) Jacobian determinant = volume scaling

If $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$, then:

$$|\det(J_F)| = \text{how volumes scale under the mapping}$$

Used in:

- change of variables in integrals
 - normalizing flows (ML)
 - coordinate transformations in physics
-

HESSIAN — Curvature of a Function

For $f : \mathbb{R}^n \rightarrow \mathbb{R}$,

$$H_f = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \dots \\ \vdots & \vdots & \end{bmatrix}$$

Hidden aspects

(A) Hessian = shape of the bowl

The Hessian tells whether you're at:

- **minimum** (positive definite)
 - **maximum** (negative definite)
 - **saddle** (mixed signs)
-

(B) Eigenvalues = curvature in principal directions

If $\lambda_i > 0 \rightarrow$ curved upward

If $\lambda_i < 0 \rightarrow$ curved downward

This is how second-order optimization methods decide step direction.

3 DOUBLE & TRIPLE INTEGRALS

Double integrals (over area)

$$\iint_R f(x, y) dA$$

Interpretation:

(A) Volume under a surface

Standard meaning.

(B) Integrating density

If f = density, then double integral gives:

- mass
 - charge
 - probability
-

Hidden aspect: Region transformation

If you transform coordinates:

$$(x, y) = T(u, v)$$

then integrals transform as:

$$\iint_R f(x, y) dx dy = \iint_{T^{-1}(R)} f(T(u, v)) |\det(J_T)| du dv$$

Jacobian determinant appears again — because it scales area.

This is the core of:

- polar coordinates
 - change of basis
 - probability distributions
 - multivariate Gaussians
-

Triple integrals

Same idea but 3D:

$$\iiint_V f(x, y, z) dV$$

Used for:

- physical mass
 - flux
 - probability in 3D
 - center of mass
 - electric/magnetic fields
-

4 VECTOR CALCULUS — Divergence, Curl, Line Integrals

Let a vector field be:

$$\vec{F} = \langle P, Q, R \rangle$$

Divergence — “How much flows out of a point”

$$\nabla \cdot \vec{F} = P_x + Q_y + R_z$$

Hidden meaning:

- divergence $> 0 \rightarrow$ source (fluid expands)
 - divergence $< 0 \rightarrow$ sink (fluid compresses)
 - divergence $= 0 \rightarrow$ incompressible field
-

Curl — “Amount of swirling”

$$\nabla \times \vec{F}$$

It measures rotation of a vector field around a point.

Hidden meaning:

- curl = 0 → conservative field
- potential function exists
- line integral path-independent

Used heavily in electromagnetism & fluid dynamics.

Line integrals

$$\int_C \vec{F} \cdot d\vec{r}$$

Meaning: work done by force along a path.

Hidden structure:

- If curl = 0, path does not matter
 - If curl ≠ 0, path strongly matters
-

◆ READY FOR NEXT STEP?

I can now do ANY of the following if you want:

- ✓ Explain each topic with diagrams
- ✓ Do solved examples for each
- ✓ Show how they connect to ML (gradient descent, Hessians in optimization, Jacobians in deep nets)
- ✓ Show intuition using geometric animations
- ✓ Give 50 practice problems with solutions
- ✓ Build a full “Multivariable Calculus → Deep Learning Math” roadmap

Tell me which direction you want to go next!