

Multivariate Calculus & Optimization

Lecture 14

Termeh Shafie

multivariate/multivariable/multidimensional calculus

$f(x)$ vs. $f(x, y), f(x, y, z), f(x, y, z, \dots)$

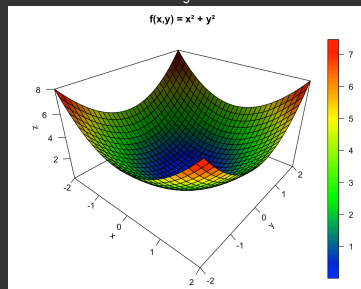
Multivariable (or multivariate) calculus extends single-variable calculus to functions of multiple variables. It includes:

- **Partial derivatives:** Differentiation with respect to one variable while keeping others constant.
- **Multiple integrals:** Double and triple integrals for computing areas, volumes, and more.
- **Vector calculus:** Topics like gradient, line/surface integrals.
- **Optimization:** Finding local maxima/minima of functions with/without constraints

$$f(x, y) = x^2 - y^2 \rightarrow$$

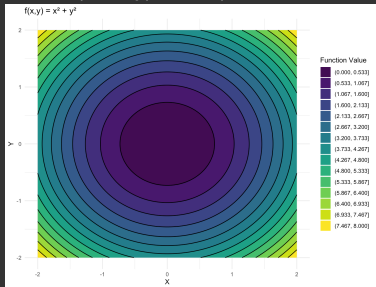
surface plot

3D visualization showing how a function behaves



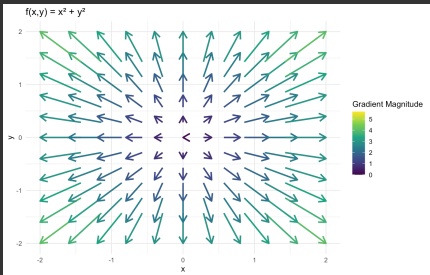
contour plot

2D visualization providing provide a top-down view of function values

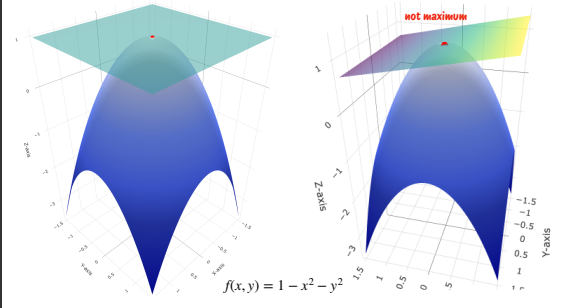


gradient plot

The gradient $\nabla f = (2x, 2y)$ shows the direction of steepest ascent



optimization



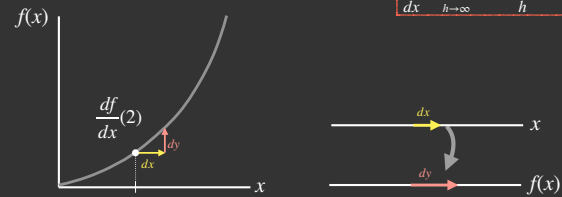
partial derivatives

A partial derivative is the derivative of a multivariable function with respect to one variable while treating all other variables as constants.

derivatives recap

Before: assume $f(x)$ with derivative $\frac{df}{dx}$. What does this mean?

$$\frac{df}{dx} = \lim_{h \rightarrow \infty} \frac{f(a+h) - f(a)}{h}$$



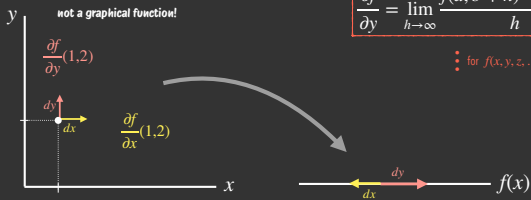
partial derivatives

Now: assume $f(x, y)$ with $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$

$$\frac{\partial f}{\partial x} = \lim_{h \rightarrow \infty} \frac{f(a+h, b) - f(a, b)}{h}$$

$$\frac{\partial f}{\partial y} = \lim_{h \rightarrow \infty} \frac{f(a, b+h) - f(a, b)}{h}$$

⋮ for $f(x, y, z, \dots)$



partial derivatives

A partial derivative is the derivative of a multivariable function with respect to one variable while treating all other variables as constants.

example

Assume following function: $f(x, y) = x^2y + 3xy^3$

Partial Derivative with Respect to x : $\frac{\partial f}{\partial x} = \frac{\partial}{\partial x}(x^2y + 3xy^3) \implies \frac{\partial f}{\partial x} = 2xy + 3y^3$

Partial Derivative with Respect to y : $\frac{\partial f}{\partial y} = \frac{\partial}{\partial y}(x^2y + 3xy^3) \implies \frac{\partial f}{\partial y} = x^2 + 9xy^2$

the symmetry of second partial derivatives

example

$$\begin{array}{ccc} & f(x, y) = x^2y + 3xy^3 & \\ & \swarrow \quad \searrow & \\ \frac{\partial f}{\partial x} = 2xy + 3y^3 & & \frac{\partial f}{\partial y} = x^2 + 9xy^2 \end{array}$$

the symmetry of second partial derivatives

example

$$\begin{array}{ccc} & f(x, y) = x^2y + 3xy^3 & \\ & \swarrow \quad \searrow & \\ \frac{\partial f}{\partial x} = 2xy + 3y^3 & & \frac{\partial f}{\partial y} = x^2 + 9xy^2 \\ \swarrow \quad \searrow & & \swarrow \quad \searrow \\ \frac{\partial^2 f}{\partial x^2} = \frac{\partial}{\partial x} \left(\frac{\partial f}{\partial x} \right) & \frac{\partial^2 f}{\partial xy} = \frac{\partial}{\partial x} \left(\frac{\partial f}{\partial y} \right) \\ = 2y & = 2x + 9y^2 \end{array}$$

the symmetry of second partial derivatives

example

$$\begin{array}{c} f(x, y) = x^2y + 3xy^3 \\ \swarrow \quad \searrow \\ \frac{\partial f}{\partial x} = 2xy + 3y^3 \qquad \frac{\partial f}{\partial y} = x^2 + 9xy^2 \\ \swarrow \quad \searrow \qquad \swarrow \quad \searrow \qquad \swarrow \quad \searrow \qquad \swarrow \quad \searrow \\ \frac{\partial^2 f}{\partial x^2} = \frac{\partial}{\partial x} \left(\frac{\partial f}{\partial x} \right) \quad \frac{\partial^2 f}{\partial xy} = \frac{\partial}{\partial x} \left(\frac{\partial f}{\partial y} \right) \quad \frac{\partial^2 f}{\partial yx} = \frac{\partial}{\partial y} \left(\frac{\partial f}{\partial x} \right) \quad \frac{\partial^2 f}{\partial y^2} = \frac{\partial}{\partial y} \left(\frac{\partial f}{\partial y} \right) \\ = 2y \qquad \qquad \qquad = 2x + 9y^2 \qquad \qquad \qquad = 2x + 9y^2 \qquad \qquad \qquad = 18xy \end{array}$$

the symmetry of second partial derivatives

example

$$\begin{array}{c} f(x, y) = x^2y + 3xy^3 \\ \swarrow \quad \searrow \\ \frac{\partial f}{\partial x} = 2xy + 3y^3 \qquad \frac{\partial f}{\partial y} = x^2 + 9xy^2 \\ \swarrow \quad \searrow \qquad \swarrow \quad \searrow \qquad \swarrow \quad \searrow \qquad \swarrow \quad \searrow \\ \frac{\partial^2 f}{\partial x^2} = \frac{\partial}{\partial x} \left(\frac{\partial f}{\partial x} \right) \quad \frac{\partial^2 f}{\partial xy} = \frac{\partial}{\partial x} \left(\frac{\partial f}{\partial y} \right) \quad \frac{\partial^2 f}{\partial yx} = \frac{\partial}{\partial y} \left(\frac{\partial f}{\partial x} \right) \quad \frac{\partial^2 f}{\partial y^2} = \frac{\partial}{\partial y} \left(\frac{\partial f}{\partial y} \right) \\ = 2y \qquad \qquad \qquad = 2x + 9y^2 \qquad \qquad \qquad = 2x + 9y^2 \qquad \qquad \qquad = 18xy \end{array}$$

the symmetry of second partial derivatives

Schwarz's theorem

If the second partial derivatives are continuous, the order of differentiation is not important and we therefore have:

$$\frac{\partial^2 f}{\partial xy} = \frac{\partial^2 f}{\partial yx}$$

gradient

The gradient of a scalar function $f(x_1, x_2, \dots, x_n)$ is a vector field that points in the direction of the greatest rate of increase of f .

For a function $f: \mathbb{R}^n \rightarrow \mathbb{R}$, the gradient is denoted as:

$$\nabla f = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \dots, \frac{\partial f}{\partial x_n} \right)$$

where each component is a partial derivative of f with respect to one of the variables.

Direction: The gradient points in the direction of the steepest ascent of f .

Magnitude: The magnitude $\|\nabla f\|$ represents the rate of the steepest increase.

Zero Gradient: If $\nabla f = 0$, the point is a critical point (possible max, min, or saddle point).

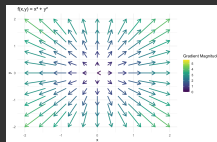
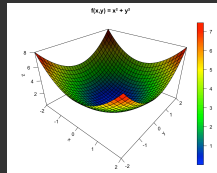
gradient

The gradient captures all the partial derivative information of a multivariable function.

example

For $f(x, y) = x^2 + y^2$, the gradient is:

$$\nabla f = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \right) = (2x, 2y)$$



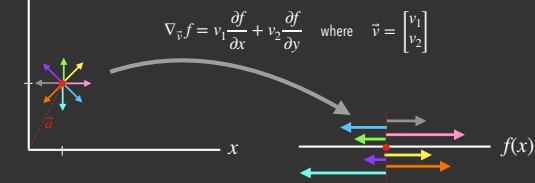
gradient

$$f(x, y) = x^2 + y^2 \implies \nabla f = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix} = \begin{bmatrix} 2x \\ 2y \end{bmatrix}$$

directional derivatives

$$\nabla_{\vec{v}} f(\vec{a}) = \lim_{h \rightarrow 0} \frac{f(\vec{a} + h \cdot \vec{v}) - f(\vec{a})}{h}$$

not a graphical function!



gradient

Zero Gradient:

If $\nabla f = 0$, the point is a critical point (max, min, or saddle point)

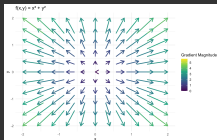
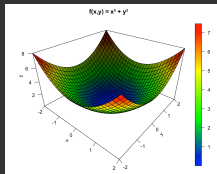
example

For $f(x, y) = x^2 + y^2$, the gradient is:

$$\nabla f = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \right) = (2x, 2y)$$

Find minimum (we see from image):

$$\nabla f = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix} = \begin{bmatrix} 2x \\ 2y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \vec{0}$$



gradient: saddle point

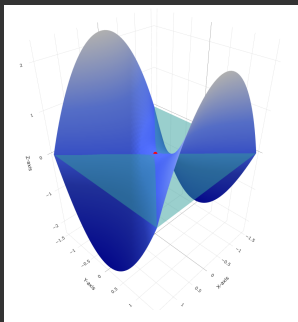
example

For $f(x, y) = x^2 - y^2$, the gradient is:

$$\nabla f = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \right) = (2x, -2y)$$

$$\nabla f = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix} = \begin{bmatrix} 2x \\ -2y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \vec{0}$$

What's the solution? Is it max or min?



the Hessian

The Hessian matrix is a square matrix of second-order partial derivatives of a scalar-valued function $f(x_1, x_2, \dots, x_n)$ is

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

- The Hessian matrix is a square matrix of second-order partial derivatives of a scalar-valued function
- The Hessian provides a way to classify critical points (where the gradient is zero):
 - If the Hessian is positive definite ($H > 0$), the critical point is a local minimum.
 - If the Hessian is negative definite ($H < 0$), the critical point is a local maximum.
 - If the Hessian has both positive and negative eigenvalues, the critical point is a saddle point.

global min/max

example

For Function: $f(x, y) = 4 - x^2 - y^2$ the Hessian is given by

$$H = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} \end{bmatrix} = \begin{bmatrix} -2 & 0 \\ 0 & -2 \end{bmatrix}$$

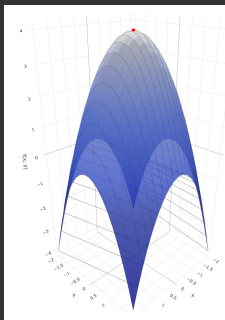
Both eigenvalues are negative (-2, -2), so H is negative definite

\implies Local max at $(0, 0)$

Is it global?

Since $f(x, y) \rightarrow -\infty$ as $|x|, |y| \rightarrow \infty$, the function is unbounded and has only one maximum, which must be global

$\implies (0, 0)$ is the global maximum, and $f(0, 0) = 4$



Hessian

The Hessian provides a way to classify critical points (where the gradient is zero).

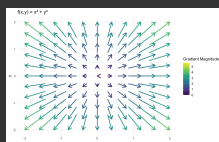
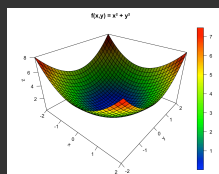
example

For $f(x, y) = x^2 + y^2$, the Hessian is:

$$H = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$

Diagonal matrix with positive values indicating function is convex

Eigenvalues?



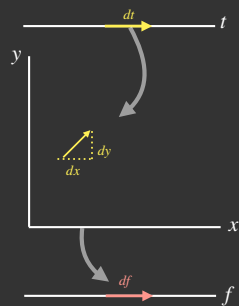
chain rule: multivariable functions

If $f(x(t), y(t))$, and both x and y are functions of a single variable t , then:

$$\frac{df(x(t), y(t))}{dt} = \frac{\partial f}{\partial x} \frac{dx}{dt} + \frac{\partial f}{\partial y} \frac{dy}{dt}$$

chain rule: multivariable functions

$$\frac{df(x(t), y(t))}{dt} = \frac{\partial f}{\partial x} \frac{dx}{dt} + \frac{\partial f}{\partial y} \frac{dy}{dt}$$



$$\textcircled{1} \quad dx = \frac{dx}{dt} dt \quad dy = \frac{dy}{dt} dt$$

$$\textcircled{2} \quad df_{dx} = \frac{\partial f}{\partial x} dx \quad df_{dy} = \frac{\partial f}{\partial y} dy$$

$$\rightarrow df = \frac{\partial f}{\partial x} \frac{dx}{dt} dt + \frac{\partial f}{\partial y} \frac{dy}{dt} dt$$

$$\frac{df}{dt} = \frac{\partial f}{\partial x} \frac{dx}{dt} + \frac{\partial f}{\partial y} \frac{dy}{dt}$$

the Jacobian

If we have a function $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$ mapping an n -dimensional input to an m -dimensional output,

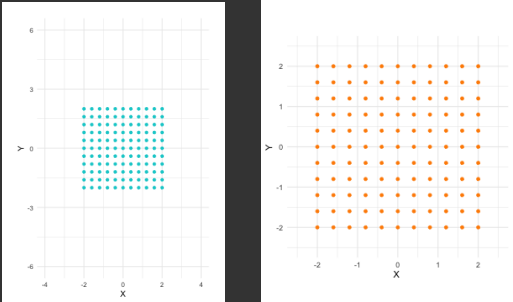
$$f(x_1, x_2, \dots, x_n) = \begin{bmatrix} f_1(x_1, x_2, \dots, x_n) \\ f_2(x_1, x_2, \dots, x_n) \\ \vdots \\ f_m(x_1, x_2, \dots, x_n) \end{bmatrix}, \text{ then the Jacobian matrix contains all first-order partial derivatives of } f:$$

$$J = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \dots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \dots & \frac{\partial f_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \frac{\partial f_m}{\partial x_2} & \dots & \frac{\partial f_m}{\partial x_n} \end{bmatrix}$$

the Jacobian

- The Jacobian matrix is a matrix of all first-order partial derivatives of a vector-valued function. It generalizes the concept of a derivative to multiple variables and dimensions.
- Measures how a function transforms space:
It describes the local scaling, rotation, or shearing of a function.
- Useful in nonlinear transformations
- The Jacobian determinant represents the factor by which the transformation stretches or squishes the n -dimensional volumes around a certain input.

linear and non-linear transformation



linear and non-linear transformation

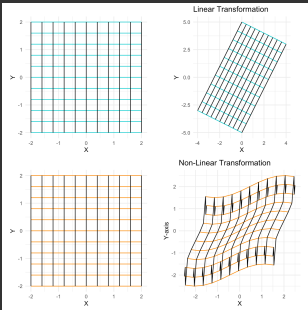
Linear:

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = A \cdot \begin{bmatrix} x \\ y \end{bmatrix} \text{ where } A = \begin{bmatrix} 1 & 1 \\ -1 & 2 \end{bmatrix}$$

Non-Linear:

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} f(x, y) \\ g(x, y) \end{bmatrix} = \begin{bmatrix} x + 0.5 \sin(y) \\ y + 0.5 \sin(x) \end{bmatrix}$$

but locally linear when we zoom in!



the Jacobian

example

Let f be a transformation from \mathbb{R}^2 to \mathbb{R}^2 with the following Jacobian matrix:

$$J = \begin{bmatrix} 3x^2 - 4 & 0 \\ 0 & 3y^2 - 4 \end{bmatrix}$$

What is the determinant of f ? How will f stretch or squish the space around the point $(1, -1)$?

where the Hessian, gradient and Jacobian meet

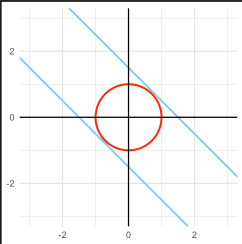
- The **gradient** points in the direction of steepest ascent.
- The **Jacobian** describes how the components of a vector function change with respect to changes in input variables
- The **Hessian** describes the local curvature of a scalar function

Matrix	Purpose	Function Type	Size
Gradient ∇f	First-order derivatives	$f: \mathbb{R}^n \rightarrow \mathbb{R}$	$n \times 1$
Jacobian J	First-order derivatives of vector functions	$f: \mathbb{R}^n \rightarrow \mathbb{R}^m$	$m \times n$
Hessian H	Second-order derivatives	$f: \mathbb{R}^n \rightarrow \mathbb{R}$	$n \times n$

- Gradient is the Jacobian of a scalar function $f: \mathbb{R}^n \rightarrow \mathbb{R} : \nabla f = J$
- Hessian is the Jacobian of the Gradient $\nabla f: H = J_{\nabla f}$

for your awareness: constrained optimization

optimize $f(x, y)$ subject to $g(x, y) = k$



$f(x, y) = 2x + y$
 $g(x, y) = x^2 + y^2 = 1$
