Lecture 4 — Calculus II

Henry Watson

Georgetown University

8/17/23

Morning challenge!

Take the derivative of the following functions (it is not necessary to simplify):

- $f(x) = x^3 + 6x^2 + 3$
- $f(x) = (x^3 + 5x) \times (x^2 - 2)$
- $f(x) = \frac{x^3+5x}{x^2-2}$
- $f(x) = \sqrt[5]{x^4 - 3x^2}$
- $f(x) = e^{5x^3+4x}$
- $f(x) = ax^2 + bx + c$, with a, b and c constants.

Agenda

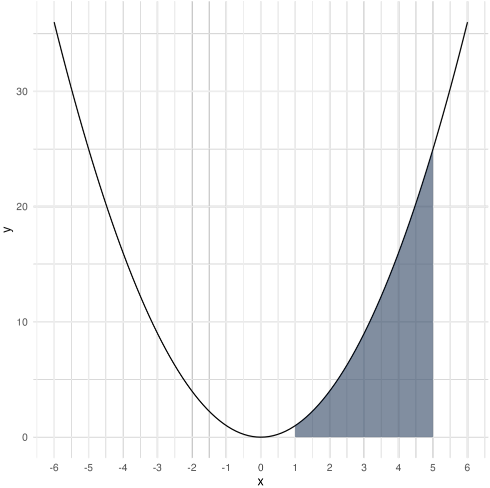
1. Integration

2. Multivariate calculus

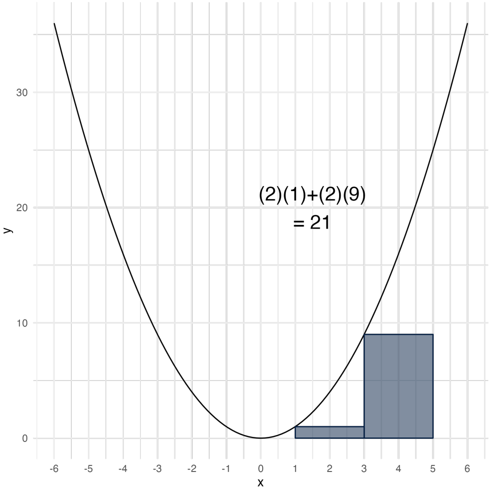
Agenda

1. Integration
2. Multivariate calculus

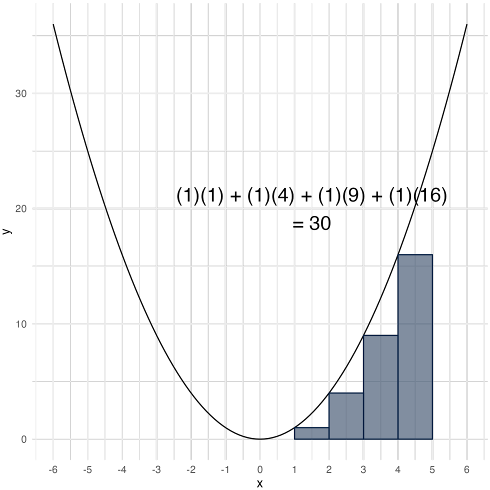
What is the area under the curve?



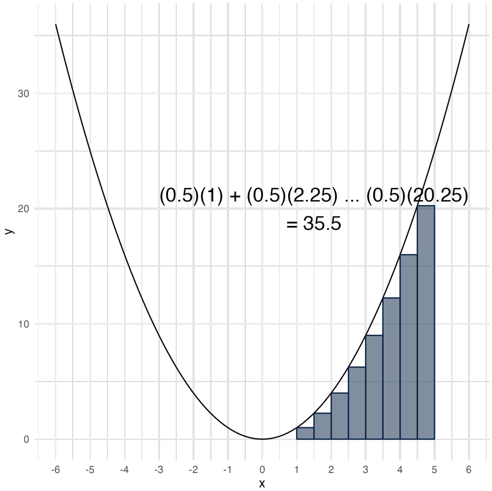
What is the area under the curve?



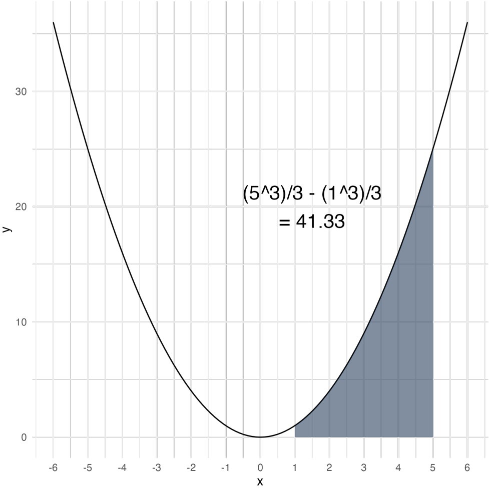
What is the area under the curve?

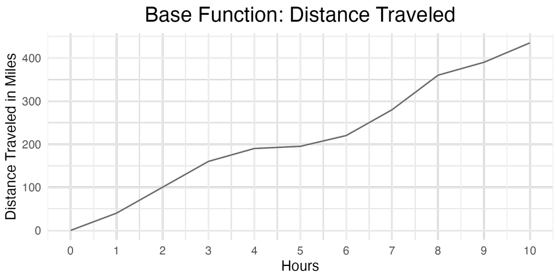


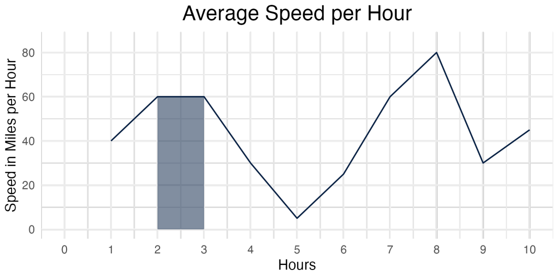
What is the area under the curve?



What is the area under the curve?




Base Function: Distance Traveled
Distance Traveled in Miles
400
300
200
100
0
0
1
2
3
4
5
6
7
8
9
10
Hours


Average Speed per Hour
Speed in Miles per Hour
80
60
40
20
0
0
1
2
3
4
5
6
7
8
9
10
Hours

Integrals

- Integrals tell us about the net effect of change, such as the net effect of a certain (changing) velocity over a certain period of time

Integrals

- Integrals tell us about the net effect of change, such as the net effect of a certain (changing) velocity over a certain period of time
- If the width of our rectangles is Δx and the height (value of the function) is $f(x)$, then we want to know $\lim_{\Delta \to 0} \sum_i f(x_i)\Delta x$

Integrals

- Integrals tell us about the net effect of change, such as the net effect of a certain (changing) velocity over a certain period of time
- If the width of our rectangles is Δx and the height (value of the function) is $f(x)$, then we want to know $\lim_{\Delta \to 0} \sum_i f(x_i)\Delta x$
- This is more commonly expressed as $\int_a^b f(x)dx$

Integrals

- Integrals tell us about the net effect of change, such as the net effect of a certain (changing) velocity over a certain period of time
- If the width of our rectangles is Δx and the height (value of the function) is $f(x)$, then we want to know $\lim_{\Delta \to 0} \sum_i f(x_i) \Delta x$
- This is more commonly expressed as $\int_a^b f(x) dx$
  - a and b bound the area under the curve on the left and right

Integrals

- Integrals tell us about the net effect of change, such as the net effect of a certain (changing) velocity over a certain period of time
- If the width of our rectangles is Δx and the height (value of the function) is $f(x)$, then we want to know $\lim_{\Delta \to 0} \sum_i f(x_i)\Delta x$
- This is more commonly expressed as $\int_a^b f(x)dx$
  - a and b bound the area under the curve on the left and right
  - A **Definite Integral** returns a value of the bounded area under the curve

Integrals

- Integrals tell us about the net effect of change, such as the net effect of a certain (changing) velocity over a certain period of time
- If the width of our rectangles is Δx and the height (value of the function) is $f(x)$, then we want to know $\lim_{\Delta \to 0} \sum_i f(x_i)\Delta x$
- This is more commonly expressed as $\int_a^b f(x)dx$
    - a and b bound the area under the curve on the left and right
    - A **Definite Integral** returns a value of the bounded area under the curve
    - $f(x)$ is the **integrand**

Integrals

- Integrals tell us about the net effect of change, such as the net effect of a certain (changing) velocity over a certain period of time
- If the width of our rectangles is Δx and the height (value of the function) is $f(x)$, then we want to know $\lim_{\Delta \to 0} \sum_i f(x_i)\Delta x$
- This is more commonly expressed as $\int_a^b f(x)dx$
  - a and b bound the area under the curve on the left and right
  - A **Definite Integral** returns a value of the bounded area under the curve
  - $f(x)$ is the **integrand**
  - dx is the **variable of integration**

Integrals

- Integrals tell us about the net effect of change, such as the net effect of a certain (changing) velocity over a certain period of time
- If the width of our rectangles is Δx and the height (value of the function) is $f(x)$, then we want to know $\lim_{\Delta \to 0} \sum_i f(x_i) \Delta x$
- This is more commonly expressed as $\int_a^b f(x)dx$
  - a and b bound the area under the curve on the left and right
  - A **Definite Integral** returns a value of the bounded area under the curve
  - $f(x)$ is the **integrand**
  - dx is the **variable of integration**
- An **Indefinite Integral** is an unbounded integral which returns the **antiderivative** function

Integrals

- Integrals tell us about the net effect of change, such as the net effect of a certain (changing) velocity over a certain period of time
- If the width of our rectangles is Δx and the height (value of the function) is $f(x)$, then we want to know $\lim_{\Delta \to 0} \sum_i f(x_i)\Delta x$
- This is more commonly expressed as $\int_a^b f(x)dx$
  - a and b bound the area under the curve on the left and right
  - A **Definite Integral** returns a value of the bounded area under the curve
  - $f(x)$ is the **integrand**
  - dx is the **variable of integration**
- An **Indefinite Integral** is an unbounded integral which returns the **antiderivative** function
  - When you take the derivative of the indefinite integral function, you reproduce the integrand $f(x)$

Fundamental theorem of calculus

- $\int_a^b f(x)dx = F(b) - F(a)$

Fundamental theorem of calculus

- $\int_a^b f(x)dx = F(b) - F(a)$
- Area under the curve = Antiderivative (Upper bound) - Antiderivative (Lower bound)

Fundamental theorem of calculus

- $\int_a^b f(x)dx = F(b) - F(a)$
- Area under the curve = Antiderivative (Upper bound) - Antiderivative (Lower bound)
- Differentiation and integration are inverse operations of one another

Reverse the rules of differentiation

- Suppose $f(x) = x^3$. What is the antiderivative $F(x)$?

Reverse the rules of differentiation

- Suppose $f(x) = x^3$. What is the antiderivative $F(x)$?
- Reverse the power rule of differentiation: $(x^n)' = nx^{n-1}$

Reverse the rules of differentiation

- Suppose $f(x) = x^3$. What is the antiderivative $F(x)$?
- Reverse the power rule of differentiation: $(x^n)' = nx^{n-1}$
- For integration: $\int x^n dx = \frac{x^{n+1}}{n+1}$

Reverse the rules of differentiation

- Suppose $f(x) = x^3$. What is the antiderivative $F(x)$?
- Reverse the power rule of differentiation: $(x^n)' = nx^{n-1}$
- For integration: $\int x^n dx = \frac{x^{n+1}}{n+1}$
- $F(x) = \frac{x^4}{4} + C$

Reverse the rules of differentiation

- Suppose $f(x) = x^3$. What is the antiderivative $F(x)$?
- Reverse the power rule of differentiation: $(x^n)' = nx^{n-1}$
- For integration: $\int x^n dx = \frac{x^{n+1}}{n+1}$
- $F(x) = \frac{x^4}{4} + C$
- Check your answer: $\frac{dF(x)}{dx} = \frac{\cancel{4}x^3}{\cancel{4}} = x^3$ ✓✓

Do not forget the C!

- When you take the derivative of a constant, it is zero

Do not forget the C!

- When you take the derivative of a constant, it is zero
- There is not enough information in a derivative for us to reverse engineer what that constant was in the original function

Do not forget the C!

- When you take the derivative of a constant, it is zero
- There is not enough information in a derivative for us to reverse engineer what that constant was in the original function
- So whenever you take an antiderivative, you must include $+C$ at the end to note that there may be a constant.

Do not forget the C!

- When you take the derivative of a constant, it is zero
- There is not enough information in a derivative for us to reverse engineer what that constant was in the original function
- So whenever you take an antiderivative, you must include $+C$ at the end to note that there may be a constant.
- When we took the antiderivative of x^3, we don't know if the constant should be $\frac{x^4}{4} + 1$ or $\frac{x^4}{4} + 100$
- Taking the derivative of either of those would return x^3 because the derivative of constants is 0

Do not forget the C!

- When you take the derivative of a constant, it is zero
- There is not enough information in a derivative for us to reverse engineer what that constant was in the original function
- So whenever you take an antiderivative, you must include $+C$ at the end to note that there may be a constant.
- When we took the antiderivative of x^3, we don't know if the constant should be $\frac{x^4}{4} + 1$ or $\frac{x^4}{4} + 100$
- Taking the derivative of either of those would return x^3 because the derivative of constants is 0
- Therefore, we say $F(x) = \frac{x^4}{4} + C$

Rules of integration

Table 7.1: List of Rules of Integration

Fundamental theorem of calculus	$\int_a^b f(x)dx = F(b) - F(a)$
Rules for bounds	$\int_a^b f(x)dx = -\int_b^a f(x)dx$ $\int_a^a f(x)dx = 0$ $\int_a^b f(x)dx = \int_a^c f(x)dx + \int_c^b f(x)dx$ for $c \in [a,b]$
Linear rule	$\int (af(x) + bg(x))dx = a\int f(x)dx + b\int g(x)dx$
Integration by substitution	$\int_a^b f(g(u))g'(u)du = \int_{g(a)}^{g(b)} f(x)dx$
Integration by parts	$\int f(x)g'(x)dx = f(x)g(x) - \int f'(x)g(x)dx$
Power rule 1	$\int x^n dx = \frac{x^{n+1}}{n+1} + C$ if $n \neq -1$
Power rule 2	$\int x^{-1}dx = \ln\|x\| + C$
Exponential rule 1	$\int e^x dx = e^x + C$
Exponential rule 2	$\int a^x dx = \frac{a^x}{\ln(a)} + C$
Logarithm rule 1	$\int \ln(x)dx = x\ln(x) - x + C$
Logarithm rule 2	$\int \log_a(x)dx = \frac{x\ln(x)-x}{\ln(a)} + C$
Trigonometric rules	$\int \sin(x)dx = -\cos(x) + C$ $\int \cos(x)dx = \sin(x) + C$ $\int \tan(x)dx = -\ln(\|\cos(x)\|) + C$
Piecewise rules	Split definite integral into corresponding pieces

More simple examples

Example 2: Calculate the area under the curve x^2 from 3 to 9

More simple examples

Example 2: Calculate the area under the curve x^2 from 3 to 9
From the fundamental theorem of calculus we know that the answer is given by:

$$\int_3^9 f(x)dx = F(9) - F(3)$$

More simple examples

Example 2: Calculate the area under the curve x^2 from 3 to 9
From the fundamental theorem of calculus we know that the answer is given by:

$$\int_3^9 f(x)dx = F(9) - F(3)$$

The next step is to determinate $F(x)$. Using the power rule, we know that $\int x^n dx = \frac{x^{n+1}}{n+1}$ so the antiderivative is $\frac{x^3}{3}$.

More simple examples

Example 2: Calculate the area under the curve x^2 from 3 to 9
From the fundamental theorem of calculus we know that the answer is given by:

$$\int_3^9 f(x)dx = F(9) - F(3)$$

The next step is to determinate $F(x)$. Using the power rule, we know that $\int x^n dx = \frac{x^{n+1}}{n+1}$ so the antiderivative is $\frac{x^3}{3}$.
Replacing that in the fundamental theorem:

More simple examples

Example 2: Calculate the area under the curve x^2 from 3 to 9
From the fundamental theorem of calculus we know that the answer is given by:

$$\int_3^9 f(x)dx = F(9) - F(3)$$

The next step is to determinate $F(x)$. Using the power rule, we know that $\int x^n dx = \frac{x^{n+1}}{n+1}$ so the antiderivative is $\frac{x^3}{3}$.
Replacing that in the fundamental theorem:

$$\int_3^9 f(x)dx = \frac{9^3}{3} - \frac{3^3}{3}$$

More simple examples

Example 2: Calculate the area under the curve x^2 from 3 to 9
From the fundamental theorem of calculus we know that the answer is given by:

$$\int_3^9 f(x)dx = F(9) - F(3)$$

The next step is to determinate $F(x)$. Using the power rule, we know that $\int x^n dx = \frac{x^{n+1}}{n+1}$ so the antiderivative is $\frac{x^3}{3}$.
Replacing that in the fundamental theorem:

$$\int_3^9 f(x)dx = \frac{9^3}{3} - \frac{3^3}{3}$$
$$\int_3^9 f(x)dx = 243 - 9$$

More simple examples

Example 2: Calculate the area under the curve x^2 from 3 to 9
From the fundamental theorem of calculus we know that the answer is given by:

$$\int_3^9 f(x)dx = F(9) - F(3)$$

The next step is to determinate $F(x)$. Using the power rule, we know that $\int x^n dx = \frac{x^{n+1}}{n+1}$ so the antiderivative is $\frac{x^3}{3}$.
Replacing that in the fundamental theorem:

$$\int_3^9 f(x)dx = \frac{9^3}{3} - \frac{3^3}{3}$$
$$\int_3^9 f(x)dx = 243 - 9$$
$$\int_3^9 f(x)dx = 234 \checkmark\checkmark$$

Integration by substitution

- When the integral contains a complicated composite function, integration by substitution can help!

Integration by substitution

- When the integral contains a complicated composite function, integration by substitution can help!
- $\int_a^b = f(g(u))g'(u)du = \int_{g(a)}^{g(b)} f(x)d(x)$

Integration by substitution

- When the integral contains a complicated composite function, integration by substitution can help!
- $\int_a^b = f(g(u))g'(u)du = \int_{g(a)}^{g(b)} f(x)d(x)$
- Look in the composite function for a piece which would be the derivative of another piece
  - $g(u)$ and $g'(u)$ where u is some piece of the composite function

Example

- Calculate $\int_1^2 2x(x^2+1)^3 dx$

Example

- Calculate $\int_1^2 2x(x^2+1)^3 dx$
- Integration by Substitution: $\int_a^b = f(g(u))g'(u)du = \int_{g(a)}^{g(b)} f(x)d(x)$

Example

- Calculate $\int_1^2 2x(x^2+1)^3 dx$
- Integration by Substitution: $\int_a^b = f(g(u))g'(u)du = \int_{g(a)}^{g(b)} f(x)d(x)$
- What could we use for u here? Looking for a piece of the function that also has its derivative in the function

Example

- Calculate $\int_1^2 2x(x^2+1)^3 dx$
- Integration by Substitution: $\int_a^b = f(g(u))g'(u)du = \int_{g(a)}^{g(b)} f(x)d(x)$
- What could we use for u here? Looking for a piece of the function that also has its derivative in the function
  - $u = x^2 + 1$
  - $\frac{du}{dx} = 2x$
  - $u = x^2 + 1$

Example

- Calculate $\int_1^2 2x(x^2+1)^3 dx$
- Integration by Substitution: $\int_a^b = f(g(u))g'(u)du = \int_{g(a)}^{g(b)} f(x)d(x)$
- What could we use for u here? Looking for a piece of the function that also has its derivative in the function
  - $u = x^2 + 1$
  - $\frac{du}{dx} = 2x$
  - $u = x^2 + 1$
- Substituting in our original expression: $\int_1^2 u^3 du$

Example

- Calculate $\int_1^2 2x(x^2+1)^3 dx$
- Integration by Substitution: $\int_a^b = f(g(u))g'(u)du = \int_{g(a)}^{g(b)} f(x)d(x)$
- What could we use for u here? Looking for a piece of the function that also has its derivative in the function
  - $u = x^2 + 1$
  - $\frac{du}{dx} = 2x$
  - $u = x^2 + 1$
- Substituting in our original expression: $\int_1^2 u^3 du$
- However, the bounds were defined for x, not for u. So we have to adapt them to u.

Example

- Calculate $\int_1^2 2x(x^2+1)^3 dx$
- Integration by Substitution: $\int_a^b = f(g(u))g'(u)du = \int_{g(a)}^{g(b)} f(x)d(x)$
- What could we use for u here? Looking for a piece of the function that also has its derivative in the function
  - $u = x^2 + 1$
  - $\frac{du}{dx} = 2x$
  - $u = x^2 + 1$
- Substituting in our original expression: $\int_1^2 u^3 du$
- However, the bounds were defined for x, not for u. So we have to adapt them to u.
  - Lower bound: $(1)^2 + 1 = 2$
  - Upper bound: $(2)^2 + 1 = 5$

Example

- Calculate $\int_1^2 2x(x^2+1)^3 dx$
- Integration by Substitution: $\int_a^b = f(g(u))g'(u)du = \int_{g(a)}^{g(b)} f(x)d(x)$
- What could we use for u here? Looking for a piece of the function that also has its derivative in the function
  - $u = x^2 + 1$
  - $\frac{du}{dx} = 2x$
  - $u = x^2 + 1$
- Substituting in our original expression: $\int_1^2 u^3 du$
- However, the bounds were defined for x, not for u. So we have to adapt them to u.
  - Lower bound: $(1)^2 + 1 = 2$
  - Upper bound: $(2)^2 + 1 = 5$
- Finally, the right expression to calculate is: $\int_2^5 u^3 du$

Example

- Calculate $\int_1^2 2x(x^2+1)^3 dx$
- Integration by Substitution: $\int_a^b = f(g(u))g'(u)du = \int_{g(a)}^{g(b)} f(x)d(x)$
- What could we use for u here? Looking for a piece of the function that also has its derivative in the function
    - $u = x^2 + 1$
    - $\frac{du}{dx} = 2x$
    - $u = x^2 + 1$
- Substituting in our original expression: $\int_1^2 u^3 du$
- However, the bounds were defined for x, not for u. So we have to adapt them to u.
    - Lower bound: $(1)^2 + 1 = 2$
    - Upper bound: $(2)^2 + 1 = 5$
- Finally, the right expression to calculate is: $\int_2^5 u^3 du$
- Final answer: $\int_2^5 u^3 du = \frac{5^4}{4} - \frac{2^4}{4} = 152.25$

Agenda

Multivariate calculus

- Calculus with more than one variable in the function

Multivariate calculus

- Calculus with more than one variable in the function
- $f(x,y,z) = \frac{xy^2}{zy+x} - \frac{y}{x+4} + 1$

Multivariate calculus

- Calculus with more than one variable in the function
- $f(x,y,z) = \frac{xy^2}{zy+x} - \frac{y}{x+4} + 1$
- $f(x_1, x_2, x_3) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3$

Partial derivatives

Partial derivative: the instantaneous rate of change in y (marginal effect) due to one variable while holding the others constant

Partial derivatives

Partial derivative: the instantaneous rate of change in y (marginal effect) due to one variable while holding the others constant

Treat every variable other than x as a constant, then take the derivative with respect to x.

Partial derivatives

Partial derivative: the instantaneous rate of change in y (marginal effect) due to one variable while holding the others constant
Treat every variable other than x as a constant, then take the derivative with respect to x.

$$f(x,y,z) = 3x^2y + zy^3 + ln(y) - x$$
$$\frac{df(x,y,z)}{dx} = 6xy - 1$$

Partial derivatives

Partial derivative: the instantaneous rate of change in y (marginal effect) due to one variable while holding the others constant
Treat every variable other than x as a constant, then take the derivative with respect to x.

$$f(x,y,z) = 3x^2y + zy^3 + ln(y) - x$$
$$\frac{df(x,y,z)}{dx} = 6xy - 1$$

You can take a partial derivative for any variable in the function:

$$\frac{df(x,y,z)}{dy} = 3x^2 + 3zy^2 + \frac{1}{y}$$

Partial derivatives

Partial derivative: the instantaneous rate of change in y (marginal effect) due to one variable while holding the others constant
Treat every variable other than x as a constant, then take the derivative with respect to x.

$$f(x,y,z) = 3x^2y + zy^3 + ln(y) - x$$
$$\frac{df(x,y,z)}{dx} = 6xy - 1$$

You can take a partial derivative for any variable in the function:

$$\frac{df(x,y,z)}{dy} = 3x^2 + 3zy^2 + \frac{1}{y}$$

Key: The bottom of the derivative notation tells you which variable you will be taking the derivative with respect to!

Gradients

A vector of all possible first order derivatives.

For $f(x, y, z)$ the gradient $\nabla = \begin{bmatrix} dx \\ dy \\ dz \end{bmatrix}$

Gradients

A vector of all possible first order derivatives.

For $f(x, y, z)$ the gradient $\nabla = \begin{bmatrix} dx \\ dy \\ dz \end{bmatrix}$

Simply take the first order partial derivative with respect to each variable and arrange as a vector!

$$f(x, y, z) = x^2 + 5xy + z^3$$
$$\frac{df(x,y,z)}{dx} = 2x + 5y$$
$$\frac{df(x,y,z)}{dy} = 5x$$
$$\frac{df(x,y,z)}{dz} = 3z^2$$
$$\nabla = \begin{bmatrix} 2x + 5y \\ 5x \\ 3z^2 \end{bmatrix}$$

Mixed partial derivatives

- Take the derivative first with respect to one variable, then take the second derivative with respect to another variable.

Mixed partial derivatives

- Take the derivative first with respect to one variable, then take the second derivative with respect to another variable.
- Indicated by $\frac{d^2 f}{dxdy}$ or $\frac{\partial^2 f}{\partial x \partial y}$ or the notation f_{xy}

Mixed partial derivatives

- Take the derivative first with respect to one variable, then take the second derivative with respect to another variable.
- Indicated by $\frac{d^2 f}{dxdy}$ or $\frac{\partial^2 f}{\partial x \partial y}$ or the notation f_{xy}

$$f(x,y,z) = 3x^2y + zy^3 + ln(y) - x$$

First take the derivative with respect to x

Mixed partial derivatives

- Take the derivative first with respect to one variable, then take the second derivative with respect to another variable.
- Indicated by $\frac{d^2f}{dxdy}$ or $\frac{\partial^2 f}{\partial x \partial y}$ or the notation f_{xy}

$$f(x,y,z) = 3x^2y + zy^3 + ln(y) - x$$

First take the derivative with respect to x

$$\frac{df(x,y,z)}{dx} = 6xy - 1$$

Then take the derivative with respect to y

Mixed partial derivatives

- Take the derivative first with respect to one variable, then take the second derivative with respect to another variable.
- Indicated by $\frac{d^2f}{dxdy}$ or $\frac{\partial^2 f}{\partial x \partial y}$ or the notation f_{xy}

$$f(x,y,z) = 3x^2y + zy^3 + ln(y) - x$$

First take the derivative with respect to x

$$\frac{df(x,y,z)}{dx} = 6xy - 1$$

Then take the derivative with respect to y

$$\frac{d^2f(x,y,z)}{dxdy} = 6x$$

The Hessian matrix

Hessians are used in optimization of multivariate functions. They tell us how a function behaves in multiple dimensions.

$$\begin{bmatrix} \frac{d^2 f(x,y,z)}{dx^2} & \frac{d^2 f(x,y,z)}{dxdy} & \frac{d^2 f(x,y,z)}{dxdz} \\ \frac{d^2 f(x,y,z)}{dydx} & \frac{d^2 f(x,y,z)}{dy^2} & \frac{d^2 f(x,y,z)}{dydz} \\ \frac{d^2 f(x,y,z)}{dzdx} & \frac{d^2 f(x,y,z)}{dzdy} & \frac{d^2 f(x,y,z)}{dz^2} \end{bmatrix}$$

The Hessian matrix

Hessians are symmetric matrices

$$\begin{bmatrix} \frac{d^2f(x,y,z)}{dx^2} & \frac{d^2f(x,y,z)}{dxdy} & \frac{d^2f(x,y,z)}{dxdz} \\ \frac{d^2f(x,y,z)}{dydx} & \frac{d^2f(x,y,z)}{dy^2} & \frac{d^2f(x,y,z)}{dydz} \\ \frac{d^2f(x,y,z)}{dzdx} & \frac{d^2f(x,y,z)}{dzdy} & \frac{d^2f(x,y,z)}{dz^2} \end{bmatrix}$$

And this is because $\frac{d^2f(x,y,z)}{dxdy} = \frac{d^2f(x,y,z)}{dydx}$

Group exercise

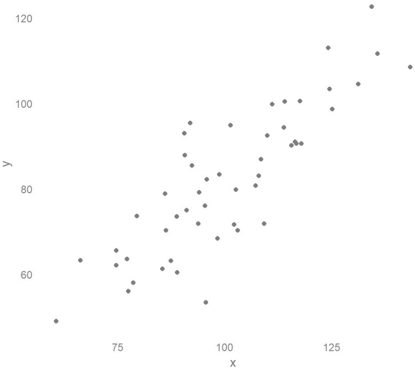
Using this function:

$$f(x,y,z) = x + y + z + x^2y^2z^2$$

Find the gradient (∇) and the Hessian.

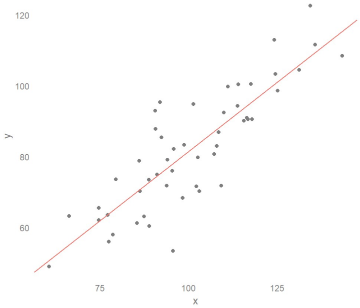
Application - Ordinary Least Squares (OLS)

- Here we have a relationship between X and Y
- The relationship isn't perfect; there is an error term ε_i
- $Y_i = \alpha + \beta X_i + \varepsilon_i$



Application - Ordinary Least Squares (OLS)

- Our goal is to identify a linear relationship between X and Y that minimizes the squared "residuals"
- Residual: the difference between the actual point Y_i and our estimation of that point $\hat{Y}_i$
- Why squared residuals? So that positive and negative residuals don't cancel out!



Deriving OLS

We want to minimize the difference between the actual point and our estimation of that point

$$min \sum (Y_i - \hat{Y}_i) \tag{1}$$

Deriving OLS

We want to minimize the difference between the actual point and our estimation of that point

$$min \sum (Y_i - \hat{Y}_i) \tag{1}$$

We can define our estimated point in terms of the intercept α and the slope β

$$min \sum (Y_i - (\alpha + X_i\beta))^2 \tag{2}$$

Deriving OLS

We want to minimize the difference between the actual point and our estimation of that point

$$min\sum(Y_i - \hat{Y}_i) \tag{1}$$

We can define our estimated point in terms of the intercept α and the slope β

$$min\sum(Y_i - (\alpha + X_i\beta))^2 \tag{2}$$

Start by taking the partial derivative with respect to α
Chain Rule: $(g(f(x))' = g'(f(x))f'(x)$

$$\frac{\partial}{\partial\alpha} = \sum(2)(Y_i - \alpha - X_i\beta)(-1) \tag{3}$$

Deriving OLS

To minimize the equation, we set the derivative equal to 0 (this will return a global minimum due to the known shape of the function)

Chain Rule: $(g(f(x))' = g'(f(x))f'(x)$

$$0 = \sum (2)(Y_i - \alpha - X_i\beta)(-1) \tag{4}$$

Deriving OLS

To minimize the equation, we set the derivative equal to 0 (this will return a global minimum due to the known shape of the function)

Chain Rule: $(g(f(x))' = g'(f(x))f'(x)$

$$0 = \sum(2)(Y_i - \alpha - X_i\beta)(-1) \tag{4}$$

Essential rules of summations:

$\sum X_i = N\bar{X}$

$\sum X = NX$

Deriving OLS

To minimize the equation, we set the derivative equal to 0 (this will return a global minimum due to the known shape of the function)

Chain Rule: $(g(f(x))' = g'(f(x))f'(x)$

$$0 = \sum(2)(Y_i - \alpha - X_i\beta)(-1) \tag{4}$$

Essential rules of summations:

$\sum X_i = N\bar{X}$

$\sum X = NX$

$$0 = N\bar{Y} - N\alpha - \beta N\bar{X} \tag{5}$$

Deriving OLS

To minimize the equation, we set the derivative equal to 0 (this will return a global minimum due to the known shape of the function)

Chain Rule: $(g(f(x))' = g'(f(x))f'(x)$

$$0 = \sum(2)(Y_i - \alpha - X_i\beta)(-1) \tag{4}$$

Essential rules of summations:

$\sum X_i = N\bar{X}$

$\sum X = NX$

$$0 = N\bar{Y} - N\alpha - \beta N\bar{X} \tag{5}$$

$$\alpha = \bar{Y} - \beta\bar{X} \tag{6}$$

Deriving OLS

Take the partial derivative of the original equation with respect to β

$$min \sum (Y_i - (\alpha + X_i \beta))^2 \tag{7}$$

Deriving OLS

Take the partial derivative of the original equation with respect to β

$$min \sum (Y_i - (\alpha + X_i \beta))^2 \tag{7}$$

$$\frac{\partial}{\partial \beta} = \sum (2)(Y_i - \alpha - X_i \beta)(-X_i) \tag{8}$$

Deriving OLS

Take the partial derivative of the original equation with respect to β

$$min \sum (Y_i - (\alpha + X_i \beta))^2 \tag{7}$$

$$\frac{\partial}{\partial \beta} = \sum (2)(Y_i - \alpha - X_i \beta)(-X_i) \tag{8}$$

Set the partial derivative equal to 0

$$0 = \sum X_i (Y_i - \alpha - X_i \beta) \tag{9}$$

Deriving OLS

Substitute in $\alpha = \bar{Y} - \beta\bar{X}$

$$0 = \sum X_i(Y_i - \bar{Y} - \beta\bar{X} - X_i\beta) \tag{10}$$

$$0 = \sum X_i Y_i - X_i\bar{Y} + X_i\beta\bar{X} - \beta X_i^2 \tag{11}$$

Deriving OLS

Substitute in $\alpha = \bar{Y} - \beta\bar{X}$

$$0 = \sum X_i(Y_i - \bar{Y} - \beta\bar{X} - X_i\beta) \tag{10}$$

$$0 = \sum X_i Y_i - X_i\bar{Y} + X_i\beta\bar{X} - \beta X_i^2 \tag{11}$$

Apply distributive property of summations

$$0 = \sum X_i Y_i - \sum X_i\bar{Y} + \sum X_i\beta\bar{X} - \sum \beta X_i^2 \tag{12}$$

The terms $\sum X_i$ simplify to $N\bar{X}$, and we can pull β out in front of the summation

Deriving OLS

$$0 = \sum X_i Y_i - N\bar{X}\bar{Y} + \beta N\bar{X}\bar{X} - \beta \sum X_i^2 \tag{13}$$

Deriving OLS

$$0 = \sum X_i Y_i - N\bar{X}\bar{Y} + \beta N\bar{X}\bar{X} - \beta \sum X_i^2 \tag{13}$$

To solve for β, move all of the β terms to one side of the equation

$$\beta \sum X_i^2 - \beta N\bar{X}^2 = \sum X_i Y_i - N\bar{X}\bar{Y} \tag{14}$$

$$\beta(\sum X_i^2 - N\bar{X}^2) = \sum X_i Y_i - N\bar{X}\bar{Y} \tag{15}$$

$$\beta = \frac{\sum X_i Y_i - N\bar{X}\bar{Y}}{\sum X_i^2 - N\bar{X}^2} \tag{16}$$

Deriving OLS

This formula is more conventionally written as

$$\beta = \frac{\sum(X_i - \bar{X})(Y_i - \bar{Y})}{\sum(X_i - \bar{X})^2} \tag{17}$$

Deriving OLS

This formula is more conventionally written as

$$\beta = \frac{\sum(X_i - \bar{X})(Y_i - \bar{Y})}{\sum(X_i - \bar{X})^2} \tag{17}$$

Numerator:

$$\sum X_i Y_i - \sum X_i \bar{Y} - \sum \bar{X} Y_i + \sum \bar{X}\bar{Y} \tag{18}$$

$$\sum X_i Y_i - N\bar{X}\bar{Y} \tag{19}$$

Denominator:

$$\sum X_i^2 - \sum 2X_i\bar{X} + \sum \bar{X}^2 \tag{20}$$

$$\sum X_i^2 - N\bar{X}^2 \tag{21}$$