Lecture 4 - Calculus II

Theodore Landsman

Georgetown University

August 18th - 2022

Morning challenge!

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

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

Agenda

Integration

Multivariate calculus

Agenda

Integration

Multivariate calculus

What is the area under the curve?

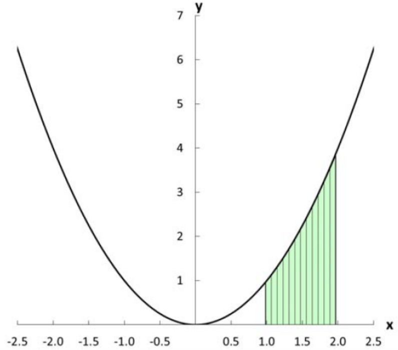
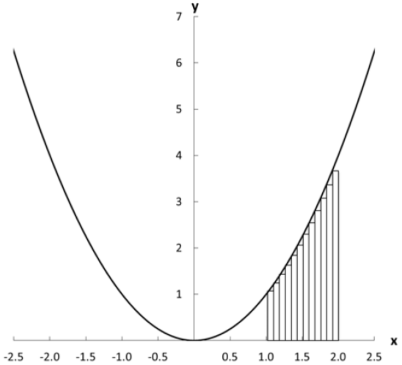


Figure 7.1: Area under $y = x^2$ from $x = 1$ to $x = 2$



Integrals

Definite Integral: An operator that helps us find the area under a curve.

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

What area? The area of the region bounded by f(x) and the *x* axis on the domain *a* to *b*.

The $f(x)$, or more generally, the expression multiplying the *dx*, is known as the *integrand*. The *dx* tells you the variable of integration.

When unbounded, it is called **indefinite integral** and represents the antiderivative:

- The definite integral returns a value, the area under the curve.
- The indefinite integral returns a function that, when differentiated, reproduces the integrand.

Fundamental theorem of calculus

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

Area under the curve = Antiderivative (Upper bound) - Antiderivative (Lower bound)

Antiderivative

The antiderivative is the function whose derivative is $f(x)$.

$$F(x) = \int f(x)dx$$

This is the same as an "indefinite integral".

Differentiation and antidifferentiation are inverse operations of one another.

$$\frac{dF(x)}{dx} = f(x)$$

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.

Example: Suppose $f(x) = 1$. What is its antiderivative?

Following the rules of differentiation, we know that $F(x)$ at least contains x, because $\frac{dx}{dx} = 1$.

However, we do not know anything about the constant. It could be $F(x) = x+1$ or $F(x) = x+100$. And due to we do not know, we say:

$$F(x) = x + c$$

More simple examples

Example 1: Suppose $f(x) = x^3$. What is its antiderivative?

Following the rules of differentiation, we know that $\frac{dx^4}{dx} = 4x^3$ and this is similar to the x^3 that we are looking for.

So if we cancel the 4, we will get our result.

Therefore: $F(x) = \frac{x^4}{4} + c$ is the antiderivative!

Check your answer: $\frac{dF(x)}{dx} = \frac{\cancel{4}x^3}{\cancel{4}} = x^3$ ✓✓

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	$\int x^{-1}dx = \ln\lvert x\rvert + C$
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	$\int \sin(x)dx = -\cos(x) + C$ $\int \cos(x)dx = \sin(x) + C$ $\int \tan(x)dx = -\ln(\lvert\cos(x)\rvert) + C$
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

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 our initial expression:

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

$$\int_3^9 f(x)dx = 243 - 9$$

$$\int_3^9 f(x)dx = 234 \checkmark\checkmark$$

Integration by substitution

Integration by substitution is attempted whenever the integral contains a composite function that one cannot integrate easily.

In particular, when integrating by substitution, look for whether there is a function and something that looks like it could be the function's derivative.

$$\int_a^b = f(g(u))g'(u)du = \int_{g(a)}^{g(b)} f(x)d(x)$$

Example

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

- Answer: Let's define $u = x^2 + 1$
- Then, $\frac{du}{dx} = 2x$. Therefore, we could say: $du = 2xdx$
- 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

Integration

Multivariate calculus

Multivariate calculus

It is just calculus with more than one variable in the function.

$$f(x,y,z) = \frac{xy^2}{zy+x} - \frac{y}{x+4} + 1$$

$$f(K,L) = K^{\alpha} L^{1-\alpha}$$

$$f(x_1,x_2,x_3) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3$$

Partial derivatives

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 $\nabla = \begin{bmatrix} dx \\ dy \\ dz \end{bmatrix}$

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.

Indicated by $\frac{d^2f}{dxdy}$ or $\frac{\partial^2 f}{\partial x \partial y}$ or the notation f_{xy}

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

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

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

The Hessian matrix

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

This is basically the second derivative when your function exists 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

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

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

Group exercise

Using this function:

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

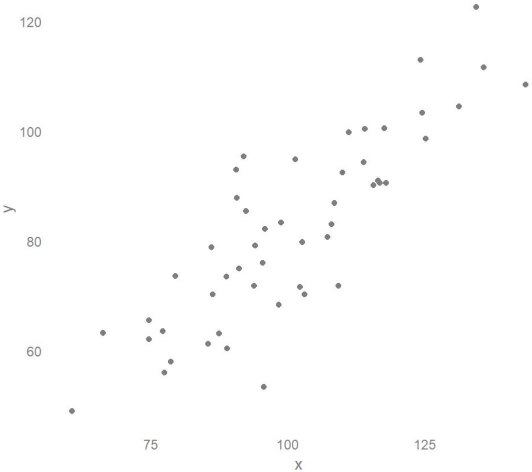
Find the gradient (∇) and the Hessian.

Ovid ♡

Application - Ordinary Least Squares (OLS)

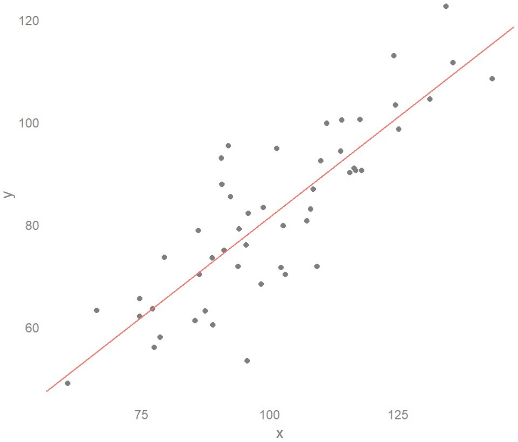
Suppose that you want to understand the following relationship that was generated by the following equation (data generation process):

$Y_i = \alpha + \beta X_i + \varepsilon_i$



Application - Ordinary Least Squares (OLS)

OLS help us to identify the linear equation that minimizes the "residuals" squared: $\hat{Y}_i = \hat{\alpha} + \hat{\beta X_i}$. What is a residual? The difference between the actual point (Y_i) and our estimation of that point $(\hat{Y}_i)$



Application - Ordinary Least Squares (OLS)

What we want is to minimize the residuals:

$$MinS = Min\sum_{i=1}^{n}(e_i^2) = Min\sum_{i=1}^{n}((Y_i - \hat{Y}_i)^2) = Min\sum_{i=1}^{n}((Y_i - \hat{\alpha} - \hat{\beta}X_i)^2)$$

FOC:

$$\frac{\partial S}{\partial \hat{\alpha}} = -2\sum_{i=1}^{n}(Y_i - \hat{\alpha} - \hat{\beta}X_i) = 0 \tag{1}$$

$$\frac{\partial S}{\partial \hat{\beta}} = -2\sum_{i=1}^{n}X_i(Y_i - \hat{\alpha} - \hat{\beta}X_i) = 0 \tag{2}$$

Application - Ordinary Least Squares (OLS)

Reordering (1) we have:

$$
\begin{aligned}
-2\sum_{i=1}^{n}(Y_i - \hat{\alpha} - \hat{\beta}X_i) = 0 \\
\sum_{i=1}^{n} Y_i = \hat{\alpha}n + \hat{\beta}\sum_{i=1}^{n} X_i \\
\hat{\alpha} = \frac{\sum_{i=1}^{n} Y_i - \hat{\beta}\sum_{i=1}^{n} X_i}{n} \\
\hat{\alpha} = \overline{Y} - \hat{\beta}\overline{X}
\end{aligned}
\tag{3}
$$

Application - Ordinary Least Squares (OLS)

Reordering (2) and replacing $\hat{\alpha}$:

$$
\begin{aligned}
-2\sum_{i=1}^{n} X_i(Y_i - \hat{\alpha} - \hat{\beta}X_i) &= 0 \\
\sum_{i=1}^{n} X_iY_i &= \hat{\alpha}\sum_{i=1}^{n} X_i + \hat{\beta}\sum_{i=1}^{n} X_i^2 \\
\sum_{i=1}^{n} X_iY_i &= (\overline{Y} - \hat{\beta}\overline{X})\sum_{i=1}^{n} X_i + \hat{\beta}\sum_{i=1}^{n} X_i^2 \\
\sum_{i=1}^{n} X_iY_i - \overline{Y}\sum_{i=1}^{n} X_i &= -\hat{\beta}\overline{X}\sum_{i=1}^{n} X_i + \hat{\beta}\sum_{i=1}^{n} X_i^2 \\
\sum_{i=1}^{n} X_iY_i - \overline{Y}\sum_{i=1}^{n} X_i &= \hat{\beta}(\sum_{i=1}^{n} X_i^2 - \overline{X}\sum_{i=1}^{n} X_i) \\
\hat{\beta} &= \frac{\sum_{i=1}^{n} X_iY_i - \overline{Y}\sum_{i=1}^{n} X_i}{\sum_{i=1}^{n} X_i^2 - \overline{X}\sum_{i=1}^{n} X_i}
\end{aligned}
\tag{4}
$$

Application - Ordinary Least Squares (OLS)

We had: $\hat{\beta} = \frac{\sum_{i=1}^{n} X_i Y_i - \overline{Y} \sum_{i=1}^{n} X_i}{\sum_{i=1}^{n} X_i^2 - \overline{X} \sum_{i=1}^{n} X_i}$

Special property that could be used for X and Y:

$$\overline{Y} = \frac{\sum_{i=1}^{n} Y_i}{n} \tag{5}$$

So, we can rearrange our $\hat{\beta}$:

$$\begin{aligned} \hat{\beta} &= \frac{\sum_{i=1}^{n} X_i Y_i - \frac{\sum_{i=1}^{n} X_i \sum_{i=1}^{n} Y_i}{n}}{\sum_{i=1}^{n} X_i^2 - \frac{\sum_{i=1}^{n} X_i \sum_{i=1}^{n} X_i}{n}} \\ \hat{\beta} &= \frac{n \sum_{i=1}^{n} X_i Y_i - \sum_{i=1}^{n} X_i \sum_{i=1}^{n} Y_i}{n \sum_{i=1}^{n} X_i^2 - (\sum_{i=1}^{n} X_i)^2} \end{aligned} \tag{6}$$

Application - Ordinary Least Squares (OLS)

Also, from (5), we know:

$$\sum_{i=1}^{n} Y_i = \frac{n\sum_{i=1}^{n} Y_i}{n} = n\overline{Y} \tag{7}$$

And we had:

$\hat{\beta} = \frac{n\sum_{i=1}^{n} X_iY_i - \sum_{i=1}^{n} X_i \sum_{i=1}^{n} Y_i}{n\sum_{i=1}^{n} X_i^2 - (\sum_{i=1}^{n} X_i)^2}$

Using (7) on the previous expression:

$$\begin{aligned}
\hat{\beta} &= \frac{n\sum_{i=1}^{n} X_iY_i - n^2\overline{XY}}{n\sum_{i=1}^{n} X_i^2 - n^2\overline{X}^2} \\
\hat{\beta} &= \frac{n(\sum_{i=1}^{n} X_iY_i - n\overline{XY})}{n(\sum_{i=1}^{n} X_i^2 - n\overline{X}^2)} \\
\hat{\beta} &= \frac{\sum_{i=1}^{n} X_iY_i - n\overline{XY}}{\sum_{i=1}^{n} X_i^2 - n\overline{X}^2}
\end{aligned} \tag{8}$$

Application - Ordinary Least Squares (OLS)

For simplicity, let's define:

$\phi = \sum_{i=1}^{n} X_i Y_i - n\overline{XY}$

$\varphi = \sum_{i=1}^{n} X_i^2 - n\overline{X}^2$.

In ϕ we can add and subtract $n\overline{XY}$:

$$\phi = \sum_{i=1}^{n} X_i Y_i - n\overline{XY} - n\overline{XY} + n\overline{XY} \tag{9}$$

Using (7) we have:

$$\phi = \sum_{i=1}^{n} X_i Y_i - \overline{Y} \sum_{i=1}^{n} X_i - \overline{X} \sum_{i=1}^{n} Y_i + \sum_{i=1}^{n} \overline{XY} \tag{10}$$

Application - Ordinary Least Squares (OLS)

We had:

$\phi = \sum_{i=1}^{n} X_i Y_i - \overline{Y} \sum_{i=1}^{n} X_i - \overline{X} \sum_{i=1}^{n} Y_i + \sum_{i=1}^{n} \overline{XY}$

Finally, we know that the summation is a linear operator, and therefore:

$$\begin{aligned}
\phi &= \sum_{i=1}^{n} (X_i Y_i - \overline{Y} X_i - \overline{X} Y_i + \overline{XY}) \\
\phi &= \sum_{i=1}^{n} (X_i (Y_i - \overline{Y}) - \overline{X} (Y_i - \overline{Y})) \\
\phi &= \sum_{i=1}^{n} (X_i - \overline{X})(Y_i - \overline{Y}))
\end{aligned} \tag{11}$$

Application - Ordinary Least Squares (OLS)

We repeat this procedural for φ, adding and subtracting $n\overline{X}^2$:

$$
\begin{aligned}
\varphi &= \sum_{i=1}^{n} X_i^2 - n\overline{X}^2 \\
\varphi &= \sum_{i=1}^{n} X_i^2 - n\overline{X}^2 + n\overline{X}^2 - n\overline{X}^2 \\
\varphi &= \sum_{i=1}^{n} X_i^2 - \overline{X}\sum_{i=1}^{n} X_i + n\overline{X}^2 - \overline{X}\sum_{i=1}^{n} X_i \\
\varphi &= \sum_{i=1}^{n} X_i^2 - 2\overline{X}\sum_{i=1}^{n} X_i + n\overline{X}^2 \\
\varphi &= \sum_{i=1}^{n} X_i^2 - \sum_{i=1}^{n} 2\overline{X}X_i + \sum_{i=1}^{n} \overline{X}^2 \\
\varphi &= \sum_{i=1}^{n} (X_i^2 - 2\overline{X}X_i + \overline{X}^2) \\
\varphi &= \sum_{i=1}^{n} (X_i - \overline{X})^2
\end{aligned}
\tag{12}
$$

Application - Ordinary Least Squares (OLS)

Replacing ϕ and φ in (10), we get the traditional form of β:

$$\hat{\beta} = \frac{\sum_{i=1}^{n}(X_i - \overline{X})(Y_i - \overline{Y}))}{\sum_{i=1}^{n}(X_i - \overline{X})^2} \tag{13}$$