Multivariable Calculus for Continuous Probability

Preface

This short guide is meant for those who have experience in single variable calculus, but have not yet taken or have forgotten multivariable calculus. I am writing this specifically for students in CS 70 as a one-stop guide to any multivariable calculus needed in continuous probability, but this guide will not cover much on probability itself.
I would also like to emphasize that multivariable calculus works very similarly to single variable calculus, so if this is your first time seeing this material just trust your intuition and most of it should work fine.

1 Partial Derivatives

The main topic this guide will cover is multivariable integration, but before proceeding to anti-derivatives it is appropriate to start out with derivatives.

Consider a single variable function $f(x)$. The derivative is denoted by $\frac{\mathrm{d}f}{\mathrm{d}x}$. This essentially means *the change of $f(x)$ per a very small change in x*. Thus, the entire motivation of the derivate is finding the rate of change of a function by a variable.

Now let's consider the function $f(x, y) = x^3y^2$. This is a function of two variables, x and y, but let's imagine for the time being that y is a constant. Then we can find the derivative in terms of x normally.

$$\frac{\mathrm{d}f}{\mathrm{d}x} = 3x^2y^2$$

We can also do the same and calculate the derivative of f in terms of y by holding x constant.

$$\frac{\mathrm{d}f}{\mathrm{d}y} = 2x^3y$$

This is is exactly how you calculate the derivates of a multivariable function. These derivatives are called **partial derivatives**, and are written as $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$.

There are more nuances to partial derivatives regarding parametrization and finding the rate of change in any direction (not just along the x and y axes for example), but these topics will not show up any time in CS 70.

So in short, to calculate a partial derivate in terms of a variable, you simply hold all the other variables constant and differentiate normally.

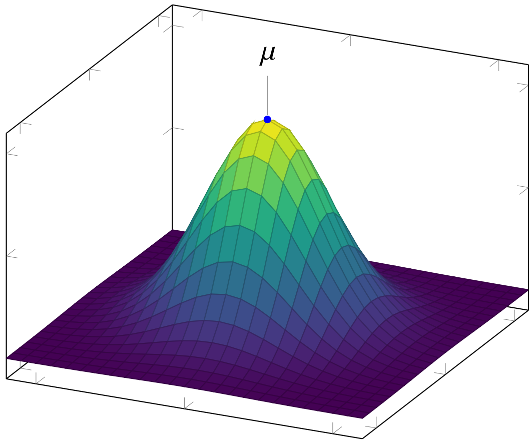
2 Multivariable Integrals

2.1 The Double Integral

We now get to the meat of the material, multivariable integrals. Once again, we will take our knowledge of single variable integrals to understand multivariable integrals.

As you may remember, the single variable integral represents the *area under a curve*. To do this, we first approximate the area of the curve with several rectangles, and progressively make these rectangles thinner and thinner until we get the exact area.

The same process applies for multivariable integrals, but instead of rectangles we use blocks (for a 3-dimensional space). Thus, as you may imagine, the double integral calculates the **volume under the curve**.



R This is a *Bivariate Gaussian*, or a joint distribution of two normal distributions, a canonical example of a multivariate continuous distribution.
What is the volume of this distribution?

Computing a double integral is very similar to computing a single integral. Just like partial derivatives, we hold one variable constant as we integrate the other.

Definition 2.1 — Double Integral
Let f be a function such that $f(x, y)$ is continuous over the region
$R = \{(x, y) \mid a \leq x \leq b,\ c \leq y \leq d\}$.
Then, the double integral over x and y (or the area A^a) is

$$V = \iint\limits_R f(x, y)\,\mathrm{d}A = \int_a^b \int_c^d f(x, y)\,\mathrm{d}y\,\mathrm{d}x = \int_c^d \int_a^b f(x, y)\,\mathrm{d}x\,\mathrm{d}y$$

[a] The area of a rectangle is $A = \textit{length} \times \textit{width}$. Because differentials represent infinitesimally small values, they behave linearly and thus $\mathrm{d}A = \mathrm{d}y\,\mathrm{d}x$

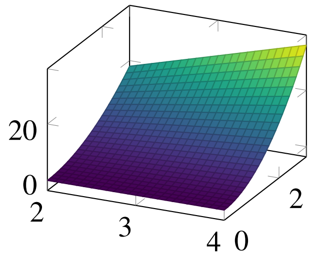
The integral is computed inside out. That is to say

$$\int_c^d \int_a^b f(x, y)\,dx\,dy = \int_c^d \left(\int_a^b f(x, y)\,dx \right) dy$$

■ **Example 2.2**

Evaluate

$$\int_2^4 \int_0^3 xy^2\,dy\,dx$$



Solution. We first integrate y while holding x constant, then integrate x normally.

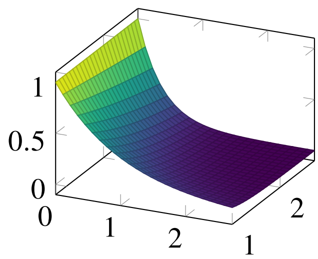
$$\begin{aligned}
\int_2^4 \int_0^3 xy^2\,dy\,dx &= \int_2^4 x\frac{y^3}{3}\Big|_0^3\,dx \\
&= \int_2^4 9x\,dx \\
&= 9\frac{x^2}{2}\Big|_2^4 \\
&= 9 \cdot 6 \\
&= 54
\end{aligned}$$

■

■ **Example 2.3**

Evaluate

$$\int_1^e \int_0^\infty \frac{1}{e^{yx}}\,dy\,dx$$



Solution.

$$\int_1^e \int_0^\infty \frac{1}{e^{yx}} \, dy \, dx = \int_1^e \frac{-1}{xe^{xy}}\Big|_0^\infty \, dx$$
$$= \int_1^e \left(0 - \left(\frac{-1}{x}\right)\right) dx = \int_1^e \frac{1}{x} \, dx$$
$$= \ln(x)\Big|_1^e = \ln(e) - \ln(1)$$
$$= 1$$

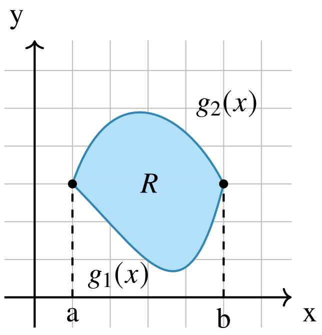
R Does the function within the bounds create a possible probability density function?

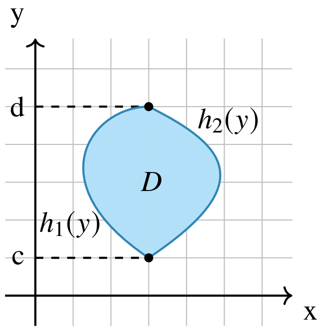
■

2.2 Double Integrals Over General Regions

You might've noticed that so far, we have only been able to integrate over rectangles defined by the bounds of our integrals. So how do we integrate over any general region? The answer lays in setting the integration bounds.

If the region R that we wish to integrate lies between two continuous functions, we can set the bounds of our integral to those functions to integrate over R.





R The functions can be either functions of x or functions of y.

Notice that $g_1(x)$ and $g_2(x)$ output values for y. If we want to integrate a function $f(x, y)$ over the region R, we use the bounding functions g_1 and g_2 as the bounds for our y integral.

Likewise, $h_1(y)$ and $h_2(y)$ output values for x, so if we want to integrate a function $f(x, y)$ over the region D, we use the bounding functions h_1 and h_2 as the bounds for our x integral.

Proposition 2.4 — Double Integral over General Regions Let $f(x, y)$ be a function that is continuous over a region R that lies between two functions of x, such that $R = \{(x, y) \mid a \le x \le b, g_1(x) \le y \le g_2(x)\}$ where $g_1(x)$ and $g_2(x)$ completely bound R. Then

$$V = \iint_R f(x, y)\, \mathrm{d}A = \int_a^b \int_{g_1(x)}^{g_2(x)} f(x, y)\, \mathrm{d}y\, \mathrm{d}x$$

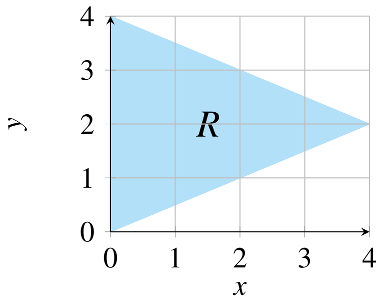
Likewise, let $f(x, y)$ be a function that is continuous over the region D that lies between two functions of y, such that $D = \{(x, y) \mid h_1(y) \le x \le h_2(y), c \le y \le d\}$ where $h_1(y)$ and $h_2(y)$ completely bound D. Then

$$V = \iint_D f(x, y)\, \mathrm{d}A = \int_c^d \int_{h_1(y)}^{h_2(x)} f(x, y)\, \mathrm{d}x\, \mathrm{d}y$$

Example 2.5

Let R be the blue region shown on the right. Evaluate

$$\iint_R xy\, \mathrm{d}A$$

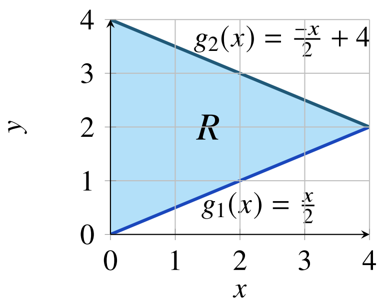


Solution.

The region lies between the functions

$$g_1(x) = \frac{x}{2}$$
$$g_2(x) = \frac{-x}{2} + 4$$

and between $x = 0$ and $x = 4$.



We set up our integrals accordingly.

$$\iint_R yx\,dA = \int_0^4 \int_{\frac{x}{2}}^{\frac{-x}{2}+4} xy\,dy\,dx$$
$$= \int_0^4 x\frac{y^2}{2}\Big|_{\frac{x}{2}}^{\frac{-x}{2}+4} dx = \int_0^4 x\frac{(\frac{-x}{2}+4)^2-(\frac{x}{2})^2}{2}\,dx$$
$$= \int_0^4 x\frac{16-4x}{2}\,dx = \int_0^4 8x-2x^2\,dx$$
$$= 4x^2 - \frac{2x^3}{3}\Big|_0^4 = 64 - \frac{128}{3}$$
$$= \frac{64}{3}$$

■

3 Joint Continuous Distribution CheatSheet

Double integrals are all the multivariable calculus you'll be expected to know. Any questions involving multivariable calculus will primarily focus on setting up the double integrals to check your understanding of probability; you will not have to do any crazy calculations. Nonetheless, here is a cheatsheet for joint continuous distributions.

Cheatsheet

Let $f_{XY}(x,y)$ be the PDF of a joint continuous distribution of random variables X and Y with corresponding marginal PDFs $f_X(x)$ and $f_Y(y)$.

$$f_X(x) = \int f_{XY}(x,y)\,dy$$
$$f_Y(y) = \int f_{XY}(x,y)\,dx$$
$$E(x) = \int x f_X(x)\,dx$$
$$E(Y) = \int y f_Y(y)\,dy$$
$$Var(X) = \int x^2 f_X(x)\,dx - (E(x))^2$$
$$Var(Y) = \int y^2 f_Y(y)\,dy - (E(y))^2$$
$$Cov(X,Y) = \iint xy f_{XY}\,dx\,dy - E(X)\,E(Y)$$