Homework 3

Simon Holm
AI503: Calculus
Teacher: Shan Shan

Table of Contents

1. Problem 1

A drug is injected into a patient's blood vessel. The function $c = f(x, t)$ represents the concentration of the drug at a distance x mm in the direction of the blood flow measured from the point of injection and at time t seconds since the injection. What are the units of the following partial derivatives? What are their practical interpretations? What do you expect their signs to be?

- $\frac{\partial c}{\partial x}$
- $\frac{\partial c}{\partial t}$

1.1. Solution

- For $\frac{\partial c}{\partial x}$:
  - Units: concentration per distance (e.g., mm/s/mm)
  - Interpretation: Rate of change of drug concentration with respect to distance along the blood vessel
  - Expected sign: Negative (concentration decreases as distance from injection point increases)
- For $\frac{\partial c}{\partial t}$:
  - Units: concentration per time (e.g., mm/s/s)
  - Interpretation: Rate of change of drug concentration with respect to time
  - Expected sign: Negative (concentration decreases over time as drug is metabolized/cleared)

2. Problem 2

Is there a function f which has the following partial derivatives? If so what is it? Are there any others?

Let

$$f_x(x, y) = 4x^3y^2 - 2y^4$$
$$f_y(x, y) = 2x^4y - 12xy^3$$

2.1. Solution

If we intergrate the functions $f_{x(x,y)}$ and $f_{y(x,y)}$

$$F_x = \int f_x dx = x^4y^2 - 2y^4x + c(y)$$
$$F_y = \int f_y dy = x^4y^2 - 3y^4x + c(x)$$

Since $F_x \neq F_y$, f is **not** a function.

3. Problem 3

Find the rate of change of $f(x,y) = x^2 + y^2$ at the point $(1,2)$ in the direction of the vector $\vec{u} = \begin{pmatrix} 0.6 \\ 0.8 \end{pmatrix}$.

3.1. Solution

We start by calculating the gradient of f:

$$\nabla f(x,y) = \left(f_x, f_y\right) = (2x, 2y)$$

since we know that $\|u\| = 1$, we can find the directional derivative in the direction of u:

If it was not a unit vector $\underbrace{\vec{v_u} = \frac{1}{\|v\|}\vec{v}}_{\text{devide each component of v with the norm}}$

$$f_{\vec{u}}(x,y) = \nabla f \cdot \vec{u} = (2x, 2y) \cdot (0.6, 0.8) = 2x(0.6) + 2y(0.8)$$

At the point (1, 2):

$$f_{\vec{u}}(1,2) = 2(1)(0.6) + 2(2)(0.8) = 1.2 + 3.2 = 4.4$$

Thus, the rate of change of f at the point (1, 2) in the direction of the vector $\vec{u}$ is 4.4.

3.2. Another solution

For $f(x,y)$, **unit** vector $\begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$ and point (a,b)

$$f_{\vec{u}}(a,b) = \lim_{h\to 0} \frac{f(a+hu_1, b+hu_2) - f(a,b)}{h}$$

Then insert point and vector in to equation $f(x,y) = x^2 + y^2$

$$\begin{aligned} f_{\vec{u}}(a,b) &= \lim_{h\to 0} \frac{f((1+h(0.6)),(2+h(0.8))) - f(1,2)}{h} \\ &= \lim_{h\to 0} \frac{(f(1+1.2h+0.36h^2)+(4+3.2h+0.64h^2)) - 5}{h} \\ &= \lim_{h\to 0} \frac{4.4h+h^2}{h} = \lim_{h\to 0} \frac{h(4.4+h)}{h} = 4.4 \end{aligned}$$

4. Problem 4

Let $f(x, y) = x^2y^3$. At the point $P(-1, 2)$, find a vector that is

(a) in the direction of maximum rate of change
(b) in the direction of minimum rate of change
(c) in a direction in which the rate of change is zero

4.1. Solution

To find the required vectors, we first need to compute the gradient of the function $f(x, y) = x^2y^3$:

$$\nabla f(x, y) = \left(f_x, f_y\right) = \left(2xy^3, 3x^2y^2\right)$$

At the point $P(-1, 2)$, we have:

$$\nabla f(-1, 2) = \left(2(-1)\left(2^3\right), 3(-1)^2\left(2^2\right)\right) = (-16, 12)$$

Now we can find the required vectors:

(a) The direction of maximum rate of change is given by the gradient vector itself (by definition):

$$\text{Max rate of change direction} = \nabla f(-1, 2) = \vec{v} = \begin{pmatrix} -16 \\ 12 \end{pmatrix}$$

(b) The direction of minimum rate of change is in the opposite direction of the gradient:

$$\text{Min rate of change direction} = -\nabla f(-1, 2) = \vec{v} = \begin{pmatrix} 16 \\ -12 \end{pmatrix}$$

(c) We now look for the normal vector to the gradient, which will give us a direction in which the rate of change is zero. One such vector can be found by swapping the components of the gradient and changing one sign:

We can find this because $f_u\vec{a} = 0 \Longleftrightarrow \nabla f(\vec{a}) \cdot \vec{u} = 0$ when the vectors are orthogonal.

so..

$$(-16, 12) \cdot (x, y) = 0$$

$$-16x + 12y = 0$$

We can se that..

$$16x = 12y$$

and then $x = \frac{12y}{16} = \frac{3}{4}y$

Thus, the normal vector is

$$\vec{v} = \begin{pmatrix} 3 \\ 4 \end{pmatrix}$$

.

5. Problem 5

Find a unit vector normal to the surface S given by $z = x^2y^2 + y + 1$ at the point (0, 0, 1).

5.1. Solution

We start by stating that $f(x, y, z) = x^2y^2 + y + 1 - z = 0$.

Then, we compute the gradient of F:

$$\nabla f(x, y, z) = \left(f_x(x, y, z), f_y(x, y, z), f_z(x, y, z)\right) = (2xy^2, 2x^2y + 1, -1)$$

At the point (0, 0, 1), we have:

$$\nabla f(0, 0, 1) = (0, 1, -1) = \vec{v}$$

Now we convert this vector into a **unit** vector:

$$\vec{v}_u = \frac{1}{\|\vec{v}\|}\vec{v} = \frac{1}{\sqrt{0^2+1^2+(-1)^2}}(0, 1, -1) = \frac{1}{\sqrt{2}}(0, 1, -1)$$

The **unit** vector is then

$$\vec{v}_u = \left(0, \frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right)$$

6. Problem 6

A student was asked to find the directional derivative of $f(x, y) = x^2e^y$ at the point $(1, 0)$ in the direction of $v = 4i + 3j$. The student's answer was

$$f_{v(1,0)} = \nabla f(1, 0) \cdot v = \frac{8}{5}i + \frac{3}{5}j$$

At a glance, how do you know this is wrong?

6.1. Solution

The student's answer is incorrect because there is no way that those two expressions are equal. the directional derivative should be a scalar value, not a vector.

The correct approach is to first normalize the vector v to get the unit vector u in the direction of v:

$$\|v\| = \sqrt{4^2 + 3^2} = 5$$

$$\vec{u} = \left(\frac{4}{5}, \frac{3}{5}\right)$$

The correct answer would be $f_{v(1,0)} = \nabla f(1, 0) \cdot \vec{u} = 2 \cdot \left(\frac{4}{5}\right) + 1 \cdot \left(\frac{3}{5}\right) = \frac{8}{5} + \frac{3}{5} = \frac{11}{5} = 2.2$.

7. Problem 7

Find the equation of the tangent plane at the given point.

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

at the point $(0, 3, 9)$

7.1. Solution

We start by calculating the gradient of the function $f(x, y) = \ln(x + 1) + y^2$:

$$\nabla f(x, y) = \left(f_x, f_y\right) = \left(\frac{\partial f}{\partial x} \ln(x + 1) + y^2, \frac{\partial f}{\partial y} \ln(x + 1) + y^2\right)$$

$$\nabla f(x, y) = \left(\frac{1}{x+1}, 2y\right)$$

now we find the gradient at the point (0, 3):

$$\nabla f(0, 3) = \left(\frac{1}{0+1}, 2 \cdot (3)\right) = (1, 6)$$

The equation of the tangent plane at the point (x_0, y_0, z_0) is given by:

$$z - z_0 = f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0)$$

Substituting the values we have:

$$z - 9 = 1(x - 0) + 6(y - 3)$$

Simplifying this, we get:

$$z - 9 = x + 6y - 18$$

$$z = x + 6y - 9$$

8. Problem 8

Are the following statements true or false?

(a) If $f(x, y)$ has $f_y(x, y) = 0$ then f must be a constant.

(b) If f is a symmetric two-variable function, that is $f(x, y) = f(y, x)$ then $f_{x(x,y)} = f_{y(x,y)}$.

(c) For $f(x, y)$, if $\frac{f(0.01,0)-f(0,0)}{0.01} > 0$, then $f_x(0, 0) > 0$.

8.1. Solution

(a) False, $f_y(x, y) = 0$ only says that there is a point where the the gradient is equal to 0

(b) False, by contradiction $f(x, y) = x^2y^2$ for this:

$$f_x(x, y) = 2xy^2$$

and

$$f_{y(x,y)} = 2x^2y$$

this is not the same, therefore it is a contradiction.

(c) False. We know that

$$\lim_{h\to 0} \frac{f(0+h,0)-f(0,0)}{h} = f_{x(0,0)}$$

in this instance $h = 0.01$ (fixed) so we cant say anything for certain.

9. Problem 9

Given a function defined as following.

$$L(\theta_1, \theta_2, b) = -\frac{1}{m} \sum_{i=1}^{m} \left[y^i \log(\hat{y}^i) + (1 - y^i) \log(1 - \hat{y}^i) \right]$$

where

$$\hat{y}^i = f(x_1^i, x_2^i)$$

and

$$f(x_1, x_2) = \frac{1}{1 + e^{-(\theta_1 x_1 + \theta_2 x_2 + b)}}$$

Suppose $x_1^i, x_2^i, y^i \in \mathbb{R}, i = 1, ..., m$ are given Use the fact that

$$\sigma(z) = \frac{1}{1 + e^{-z}}, \qquad \sigma'(z) = \sigma(z)(1 - \sigma(z))$$

to prove the following is true

$$\frac{\partial L}{\partial \theta_1} = \frac{1}{m} \sum_{i=1}^{m} \left(\hat{y}^{(i)} - y^{(i)} \right) x_1^{(i)}$$
$$\frac{\partial L}{\partial \theta_2} = \frac{1}{m} \sum_{i=1}^{m} \left(\hat{y}^{(i)} - y^{(i)} \right) x_2^{(i)}$$
$$\frac{\partial L}{\partial b} = \frac{1}{m} \sum_{i=1}^{m} \left(\hat{y}^{(i)} - y^{(i)} \right)$$

.

9.1. Solution

To simplify $L = \sum_{i=1}^{m} l^i$ where $l^i = [y^i \log(\hat{y}^i) + (1-y^i)\log(1-\hat{y}^i)]$

(a) $\dfrac{\partial L}{\partial \theta_1} = \dfrac{1}{m}\sum_{i=1}^{m}(\hat{y}^{(i)} - y^{(i)})x_1^{(i)}$

So we need to find $\frac{\partial l}{\partial \theta_1}$

$$\frac{\partial l}{\partial \theta_1} = -\left[\frac{\partial}{\partial \theta_1} y \log(\hat{y}) + \frac{\partial}{\partial \theta_1}(1-y)\log(1-\hat{y})\right]$$
$$= -\left[y\frac{1}{\hat{y}}\frac{\partial \hat{y}}{\partial \theta_1} + (1-y)\frac{1}{1-\hat{y}}\frac{\partial(1-\hat{y})}{\partial \theta_1}\right] \qquad \text{(apply ing chainrule)}$$

since $\frac{\partial(1-\hat{y})}{\partial \theta_1} = \frac{\partial 1}{\partial \theta_1} - \frac{\partial \hat{y}}{\partial \theta_1} = 0 - \frac{\partial \hat{y}}{\partial \theta_1}$

$$= -\left[y\frac{1}{\hat{y}}\frac{\partial \hat{y}}{\partial \theta_1} + (1-y)\frac{1}{1-\hat{y}}\frac{\partial \hat{y}}{\partial \theta_1}\right]$$
$$= -\left[y\frac{1}{\hat{y}}\frac{\partial \hat{y}}{\partial \theta_1} + \frac{1-y}{1-\hat{y}}\frac{\partial \hat{y}}{\partial \theta_1}\right]$$

Then factor out $\frac{\partial \hat{y}}{\partial \theta_1}$

$$= -\left(\frac{y}{\hat{y}} + \frac{1-y}{1-\hat{y}}\right)\cdot\frac{\partial \hat{y}}{\partial \theta_1}$$

This can be simplified by:

$$\frac{y}{\hat{y}} - \frac{1-y}{1-\hat{y}} = \frac{y(1-\hat{y})}{\hat{y}(1-\hat{y})} - \frac{\hat{y}(1-y)}{\hat{y}(1-y)} = \frac{y(1-\hat{y})-(\hat{y}(1-y))}{\hat{y}(1-y)}$$
$$\frac{y-y\hat{y}-(\hat{y}-\hat{y}y)}{\hat{y}(1-y)} = \frac{y-y\hat{y}-\hat{y}+\hat{y}y}{\hat{y}(1-y)} = \frac{y-\hat{y}}{\hat{y}(1-y)}$$

So now

$$\frac{\partial l}{\partial \theta_1} = -\frac{y-\hat{y}}{\hat{y}(1-y)}\frac{\partial \hat{y}}{\partial \theta_1}.$$

Since $\sigma(x) = \frac{1}{1+e^{-x}}, \sigma'(x) = \sigma(x)(1-\sigma(x))$ and by chain-rule

Then for $f(z) = \frac{1}{1+e^{z}}$

$$\frac{\partial \hat{y}}{\partial \theta_1} \Rightarrow \frac{\partial \hat{y}}{\partial z}\frac{\partial z}{\partial \theta_1} = \hat{y}(1-\hat{y})\cdot x_1$$

So now

$$\frac{\partial l}{\partial \theta_1} = -\frac{y-\hat{y}}{\hat{y}(1-y)}\frac{\partial \hat{y}}{\partial \theta_1} = \frac{y-\hat{y}}{\hat{y}(1-y)}\hat{y}(1-\hat{y})\cdot x_1 =$$
$$\frac{\partial l}{\partial \theta_1}(\hat{y}-y)x_1 \Longrightarrow \boldsymbol{\frac{\partial L}{\partial \theta_1} = \frac{1}{m}\sum_{i=1}^{m}(\hat{y}^{(i)} - y^{(i)})x_1^{(i)}}$$

(b) $\dfrac{\partial L}{\partial \theta_2} = \dfrac{1}{m}\sum_{i=1}^{m}(\hat{y}^{(i)} - y^{(i)})x_2^{(i)}$

Following the same procedure as (a) but differentiating w.r.t. θ_2 instead of θ_1

So.

$$\frac{\partial \hat{y}}{\partial \theta_2} \Rightarrow \frac{\partial \hat{y}}{\partial z}\frac{\partial z}{\partial \theta_2} = \hat{y}(1-\hat{y})\cdot x_2$$

So now

$$\frac{\partial l}{\partial \theta_2} = -\frac{y-\hat{y}}{\hat{y}(1-y)}\frac{\partial \hat{y}}{\partial \theta_2} = \frac{y-\hat{y}}{\hat{y}(1-y)}\hat{y}(1-\hat{y})\cdot x_2 =$$
$$\frac{\partial l}{\partial \theta_2}(\hat{y}-y)x_2 \Longrightarrow \boldsymbol{\frac{\partial L}{\partial \theta_2} = \frac{1}{m}\sum_{i=1}^{m}(\hat{y}^{(i)} - y^{(i)})x_2^{(i)}}$$

(c) $\dfrac{\partial L}{\partial b} = \dfrac{1}{m}\sum_{i=1}^{m}\big(\hat{y}^{(i)} - y^{(i)}\big)$

Following the same procedure as (a) but differentiating w.r.t. b instead of θ_1

So.

$$\tfrac{\partial \hat{y}}{\partial b} \Rightarrow \tfrac{\partial \hat{y}}{\partial z}\tfrac{\partial z}{\partial b} = \hat{y}(1-\hat{y})$$

So now

$$\tfrac{\partial l}{\partial b} = -\tfrac{y-\hat{y}}{\hat{y}(1-y)}\tfrac{\partial \hat{y}}{\partial b} = \tfrac{y-\hat{y}}{\hat{y}(1-y)}\hat{y}(1-\hat{y}) =$$
$$\tfrac{\partial l}{\partial b}(\hat{y}-y)x_2 \Longrightarrow \boldsymbol{\tfrac{\partial L}{\partial b} = \tfrac{1}{m}\sum_{i=1}^{m}\big(\hat{y}^{(i)} - y^{(i)}\big)}$$