Week 8 Homework

Math 510 — Christina Morgenstern

---

1. Compute the partial derivatives (with respect to both x and y) for the functions below.

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

$f_x(x, y) = 2xy + y^2 = \boldsymbol{y(2x + y)}$
$f_y(x, y) = x^2 + 2yx = \boldsymbol{x(x + 2y)}$

$g(x, y) = 4x^2 \sin(y) + xy$

$g_x(x, y) = \mathbf{8}\boldsymbol{x} \,\mathbf{sin}\, (\boldsymbol{y}) + \boldsymbol{y}$
$g_y(x, y) = 4x^2 \cos(y) + \boldsymbol{x} = \boldsymbol{x}(\mathbf{4}\boldsymbol{x}\,\mathbf{cos}(\boldsymbol{y}) + \mathbf{1})$

2. Given a function *f(x, y),* if I cut the graph with the plane x = 6, which partial derivative will the graph represent?

The correct notation is as follows:

$f_y(x, y) \; with \; x = 6$

If I am setting *x* to a constant (*x*=*6*), I am calculating the derivative with respect to *y*.

3. Let *P(w, m, c)* denote the profit (in dollars) for a company when they have *w* workers, *m* managers, and *c* computers. Explain what P_c *(w, m, c)* means in this case. What is the interpretation of P_c *(w, m, c)* > 0?

P_c *(w, m, c)* denotes the partial derivative of the function *P* with respect to *c* (computers), if the other variables *w* and *m* were kept constant. P_c describes the instantaneous rate of change of the function *P* if the number of workers and managers were kept constant.

P_c *(w, m, c)* > 0 means, that the slope of the tangent line is positive and thus the function is increasing at this point.

4. Under what circumstances might a linear approximation be a „bad" estimate for the value of a function?

If functions contain holes, gaps, jumps or sharp corners and are not differentiable at that point. In general, linear approximation is a bad estimate, if the function has a discontinuous behavior.

5. Let $f(x, y) = \sqrt{xy}$. Use linear approximation to estimate the value of *f(3.8, 4.1)*. (Hint: first figure out a nearby point that you can easily compute.)

$$f(x, y) = \sqrt{xy}$$

a. Figure out nearby point: $f(4, 4)$

$$f(4,4) = \sqrt{4 \times 4} = 4$$

b. Compute partial derivatives with respect to *x* and *y*:

$$f_x(x, y) = \frac{y}{2\sqrt{xy}}$$

$$f_y(x, y) = \frac{x}{2\sqrt{xy}}$$

c. Plug in values of nearby point and calculate derivates:

$$f_x(4,4) = \frac{4}{2\sqrt{16}} = \frac{4}{8} = \frac{1}{2}$$

$$f_y(4,4) = \frac{4}{2\sqrt{16}} = \frac{4}{8} = \frac{1}{2}$$

d. Plug in values into equation for linear approximation:

$$f(x_1, y_1) \approx f(x_0, y_0) + f_x(x_0, y_0)(x_1 - x_0) + f_y(x_0, y_0)(y_1 - y_0)$$

$$f(3.8, 4.1) \approx 4 + \frac{1}{2}(3.8 - 4) + \frac{1}{2}(4.1 - 4)$$

$$\boldsymbol{f(3.8, 4.1) \approx 3.95}$$

The estimated function value at $f(3.8, 4.1)$ is approximately 3.95.

6. Under what circumstances might gradient descent not work? Could it miss a point? Does your initial guess matter?

I guess, that gradient descent might be hard to apply to the opposite, to concave functions, where the error surface is bended upward.
Also, if the function contains many local minima there is a chance, that the algorithm might get stuck in a local minimum and not converge to the potential optimum. However, one might be able to circumvent this problem by applying different learning rates.

My initial guess of concave functions might not matter, because there is also gradient ascent which is similar to gradient descent but tries the opposite, finding the maximum in optimization.

7. Level curve of the function

$z = f(x, y) = -x^2 - y^2 = -8$ → Circle with center at (0,0) and radius $r = 2\sqrt{2}$

Circle Equation:
$(x - a)^2 + (y - b)^2 = r^2$

Rewrite $-x^2 - y^2 = -8$ in the form of the standard circle equation
$(x - 0)^2 + (y - 0)^2 = (2\sqrt{2})^2$

Therefore, the circle properties are:
$(a, b) = (0,0)$
$r = 2\sqrt{2}$

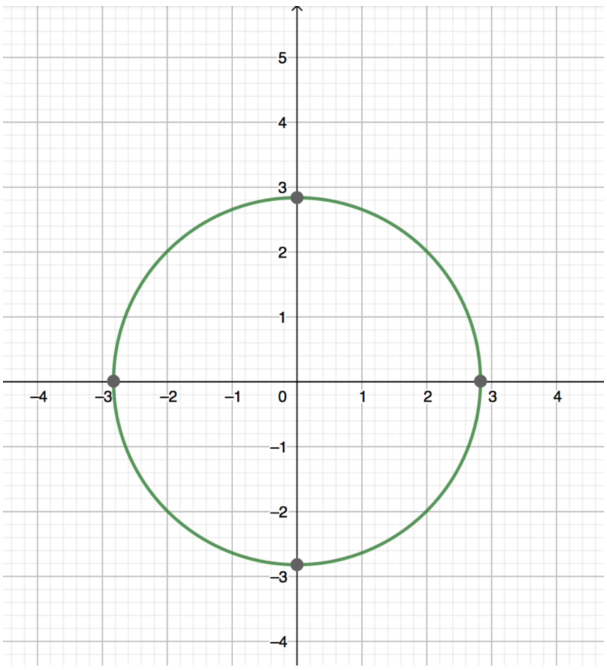

Fig. 1. Graph of the level curve of the function $f(x, y) = -x^2 - y^2$
Graphed using geogebra.org

The gradient vector:

$\nabla f = [f_x(x, y) \quad f_y(x, y)]$

$\nabla f = [-2x - y^2 \quad -x^2 - 2y]$

Plug in values $(\sqrt{8}, 0)$, $(2,2)$ and $(0, \sqrt{8})$ to obtain the following gradient vectors, respectively:

$\nabla f_u = [-5.7 \quad -8]$

$\nabla f_v = [-8 \quad -8]$

$\nabla f_w = [-8 \quad -5.7]$

I graphed the gradient vectors $\nabla f_u, \nabla f_v, \nabla f_w$ starting from A$(\sqrt{8}, 0)$, $B(2,2)$ and C $(0, \sqrt{8})$, respectively. The vectors represent the direction of the rate of change.

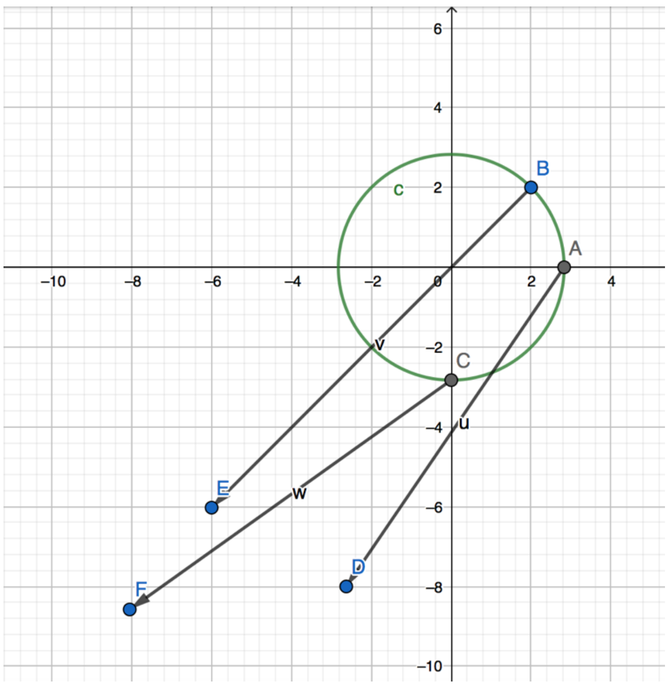

Fig. 2. Graph of the level curve of the function $f(x, y)$with the gradient vectors *u*(AD), *v*(BE) and *w*(CF)
Graphed using geogebra.org

And a 3D graphical representation of the gradient vectors $\nabla f_u, \nabla f_v, \nabla f_w$ starting from A($\sqrt{8}$, 0), $B(2,2)$ and $C\ (0, \sqrt{8})$, respectively:

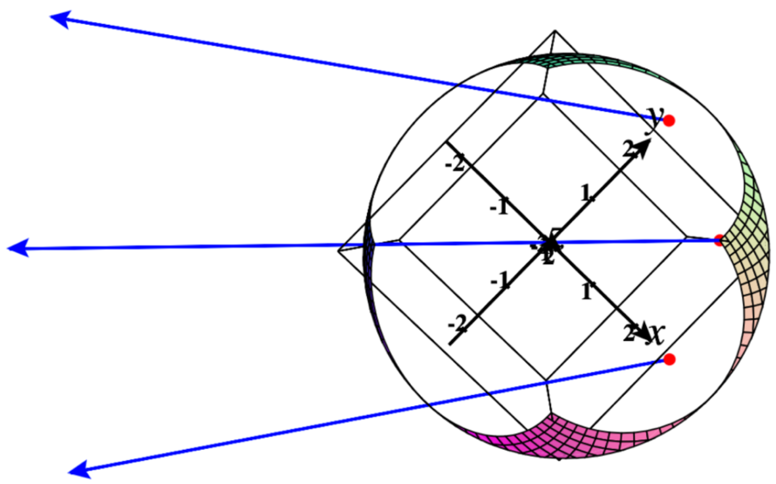

Fig. 3. Graph of the level curve of the function $f(x, y)$ with the gradient vectors $\nabla f_u, \nabla f_v, \nabla f_w$ (*blue arrows*) starting from A($\sqrt{8}$, 0), $B(2,2)$ and $C\ (0, \sqrt{8})$ (red dots)
Graphed using CalcPlot3D
(https://www.monroecc.edu/faculty/paulseeburger/calcnsf/CalcPlot3D/)