Math 253 2017: Homework 4

1. **Best quadratic approximation.** Consider a function of two variables $f(x, y)$ and a point (x_0, y_0). Let $A = f(x_0, y_0)$, $B = f_x(x_0, y_0)$, $C = f_y(x_0, y_0)$, $E = f_{xx}(x_0, y_0)$, $F = f_{xy}(x_0, y_0)$, and $G = f_{yy}(x_0, y_0)$. Now define another function $g(x, y)$ by

$$g(x, y) := A + \begin{pmatrix} B & C \end{pmatrix} \begin{pmatrix} x - x_0 \\ y - y_0 \end{pmatrix} + \frac{1}{2} \begin{pmatrix} x - x_0 & y - y_0 \end{pmatrix} \begin{pmatrix} E & F \\ F & G \end{pmatrix} \begin{pmatrix} x - x_0 \\ y - y_0 \end{pmatrix}. \tag{1}$$

   (a) Expand the linear algebra to get an expression for $g(x, y)$ with neither matrices nor vectors.

   (b) Compute $\vec{\nabla} g$ and the "Hessian matrix of g" denoted by $\mathbf{H}(g) = \begin{pmatrix} g_{xx} & g_{xy} \\ g_{xy} & g_{yy} \end{pmatrix}$. Evaluate g, $\vec{\nabla} g$, and $\mathbf{H}(g)$ at (x_0, y_0). Confirm that these agree with $f(x_0, y_0)$, $\vec{\nabla} f(x_0, y_0)$, and $\mathbf{H}(f)(x_0, y_0)$. Thus we see that g is the best quadratic approximation to f at the point (x_0, y_0).[1]

   (c) Explain without calculation why we expect $g(x, y)$ to have exactly one critical point.

[1] This is analogous to the best linear approximation covered earlier in the course. The idea can be generalized to find the Taylor series expansion by adding cubic terms, quartic terms, ad infinitum, provided the function is smooth enough.

(d) Show that the critical point of $g(x, y)$ is the point $\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} = -\begin{bmatrix} E & F \\ F & G \end{bmatrix}^{-1} \begin{pmatrix} B \\ C \end{pmatrix} + \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$.

(e) Use the previous result to give an explicit formula for x_1 and y_1 in terms of B, C, E, F, G, x_0 and y_0.

(f) Under what conditions on A, B, C, E, F, G can we guarantee g has a minimum?

(g) Let $f(x, y) = \sin(\pi x) + y^2 - y + xy$, and suppose $x_0 = 3/2$ and $y_0 = 1$. Find $g(x, y)$. Plot both f and g on the same axes using technology. Label (x_0, y_0) on your plot. Find (x_1, y_1); display it in decimal form and label it on your plot. Attach the plot.

2. **Numerical Optimization.** Let $f(x, y)$ be an unknown function but suppose we have a "black box": we can feed in (x_0, y_0) and it will evaluate the function and its first and second partial derivatives: $A = f(x_0, y_0)$, $B = f_x(x_0, y_0)$, $C = f_y(x_0, y_0)$, $E = f_{xx}(x_0, y_0)$, $F = f_{xy}(x_0, y_0)$, and $G = f_{yy}(x_0, y_0)$. We seek a local minimum of $f(x, y)$.[2] The black box—let's call it the "Phantom Power 98"—is connected to the internet at https://www.math.ubc.ca/~cbm/math253/2017/fp98.

   (a) Suppose we start at $x_0 = -3$ and $y_0 = 11$. Query the Phantom Power 98 and give the matrix form of $g_0(x, y)$, the best quadratic approximation to f (hint: use the previous question).

   (b) Find the minimizer (x_1, y_1) of g_0. What is the minimum value of g_0?

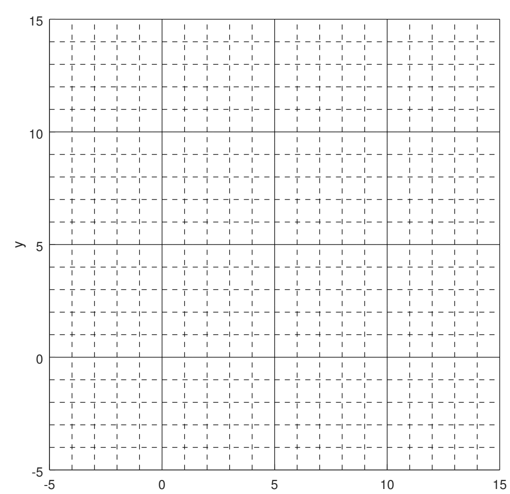
   (c) Query the Phantom Power 98 to find $f(x_1, y_1)$. Explain why this is different from $g_0(x_1, y_1)$.

   (d) Now construct $g_1(x, y)$ using the partial derivatives of f at (x_1, y_1).

[2] This question explores a basic form of the "BFGS algorithm" (sadly not "Big Friendly GiantS" nor "Big F*%king GunS" but rather "Broyden–Fletcher–Goldfarb–Shanno"). BFGS is used to solve massively-many-variable optimization problems—where do we drill the next well? How to price these airline tickets? When should I sell my bitcoin? The bottleneck is $\mathbf{H}^{-1}$: BFGS approximates it and improves it during each (x_k, y_k) update; numerical linear algebra makes the world go round!

(e) Find the minimizer (x_2, y_2) of g_1.

(f) Continue in this manner until you reach a (local) minimum of f (how will you know you're at a local minimum of f?) On an xy-plane, sketch the sequence of points (x_k, y_k), connecting them with line segments.



(g) Suppose at some step we had got to the point $(x_k, y_k) = (99, 88)$. What can you say about the critical point of g_k? So maybe that is not a good choice for (x_{k+1}, y_{k+1})... Instead, if we just want to move "downhill", which direction should we move in to search for a reasonable (x_{k+1}, y_{k+1})?

5 marks 3. (a) Suppose $z = f(x, y)$ describes the surface of a mountain. The values of some derivatives are known at certain points:

point	f_x	f_y	f_{xx}	f_{yy}	f_{xy}
A	0	1	4	-3	-2
B	0	0	4	3	-4
C	10	2	-5	3	2
E	0	0	-3	-3	2
F	0	0	5	5	5
G	0	0	3	4	3

For each situation, identify a suitable point **and *briefly* explain:** e.g., "X, because it is a local minimum." **Justify your answers with calculations.**

i. Mika wants to hike to point with view in every direction (where the mountain does not obscure her view). Where should she go?

Answer:

ii. Xiaofei is a biologist who wants to re-introduce a species of frog into a wet environment. Where is water mostly likely to form a pool?

Answer:

iii. Donnie wants to have picnic at a place where the ground is flat and dry. He thinks a peak would be too windy. Where should he go?

Answer:

iv. Akshat also wants to have a picnic where the ground is flat, but he likes surprises. Where should he go?

Answer: