# SOLUTIONS to Exercises from Optimization

1. Use the bisection method to find the root correct to 6 decimal places: $3x^3 + x^2 = x + 5$

   SOLUTION: For the root finding algorithm, we need to rewrite the equation as:

   $$3x^3 + x^2 - x - 5 = 0$$

   Then we need an interval on which the sign of the function changes. We see that $f(0) = -5$, $f(1) = -2$ and $f(2) = 21$. Therefore, we'll use the interval $[1, 2]$.

   Using the Matlab function that is included in the notes, in the command window we would type:

```
>> f=inline('3*x^3+x^2-x-5');
>> rt=bisect(f,1,2,1e-6)
Finished after 19 iterates

rt =

    1.1697
```

2. If we have $f(0, 1) = 3$ and $\nabla f((0, 1)) = [1, -2]$, then the linearization will be

   $$3 + (x - 0) - 2(y - 1)$$

   Substituting $x = 1/2, y = 1/2$ gives us $9/2$.

3. If $f = x^2 y + 3y$, we compute the linearization at $(1, 1)$:

   $$f(1, 1) = 1 + 3 = 4 \qquad \nabla f = [2xy, x^2 + 3]\Big|_{(1,1)} = [2, 4]$$

   Therefore, the linearization is:

   $$4 + 2(x - 1) + 4(y - 1)$$

   (Which you can leave in this form).

4. Let $f = xy + y^2$. At the point $(2, 1)$, in which direction is $f$ increasing the fastest? How fast is it changing?

   SOLUTION: The function will increase the fastest in the direction of the gradient. The value of the directional derivative is how fast the function is changing (it takes the place of the regular derivative in that regard). The answers then are the gradient evaluated at $(2, 1)$ and the magnitude of the gradient at $(2, 1)$:

   $$\nabla f(2, 1) = [1, 4] \qquad \|\nabla f\| = \sqrt{1^2 + 4^2} = \sqrt{17}$$

5. Use the second derivatives test to classify the critical points of the following:

(a) $x^2 + xy + y^2 + y$
Compute the gradient and the Hessian. Solve for where the gradient is zero (the critical points), then check the Hessian:

$$\nabla f = [2x + y, x + 2y + 1]$$

The only critical point is $x = 1/3, y = -2/3$. Compute the Hessian:

$$Hf = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \quad \Rightarrow \quad \begin{array}{rl} f_{xx}f_{yy} - f_{xy}^2 & = 4 - 1 = 3 > 0 \\ f_{xx} & > 0 \end{array} \quad \Rightarrow \quad \text{Local Min}$$

(b) $xy(1 - x - y)$
Same technique as the last problem, but we have a lot more critical points. First, the gradient and Hessian are:

$$\nabla f = [y - y^2 - 2xy, \quad x - x^2 - 2xy] \qquad Hf = \begin{bmatrix} -2y & 1 - 2x - 2y \\ 1 - 2x - 2y & -2x \end{bmatrix}$$

To find the critical points, we can factor the first equation:

$$y(1 - y - 2x) = 0$$

so that $y = 0$ or $y = 1 - 2x$. Substitute these into the second equation to get the corresponding $x$. First, if $y = 0$, then

$$x - x^2 = 0 \quad \Rightarrow \quad x = 0, x = 1$$

If $y = 1 - 2x$, then

$$x(1 - x - 2y) = 0 \quad \Rightarrow \quad x(3x - 1) = 0$$

so $x = 0$ (and $y = 1$) or $x = 1/3$ (and $y = 1/3$). All told, we have four critical points:

$$(0, 0), (1, 0), (0, 1), (1/3, 1/3)$$

Putting each into the Hessian:

$$Hf(0, 0) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

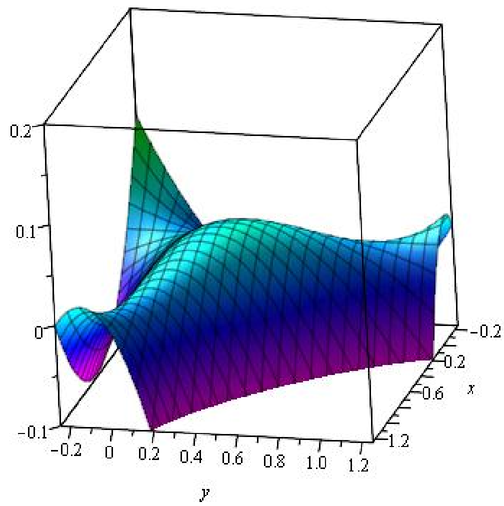
The determinant is negative- We have a saddle point here. Next point will be $(1, 0)$:

$$Hf(1, 0) = \begin{bmatrix} 0 & -1 \\ -1 & -2 \end{bmatrix}$$

The determinant is negative here, too- Another saddle. Next point is $(0,1)$ and we'll get a saddle there, too. The final point is $(1/3,1/3)$:

$$Hf(1/3,1/3) = \begin{bmatrix} -2/3 & -1/3 \\ -1/3 & -2/3 \end{bmatrix}$$

The determinant is positive, and $f_{xx} < 0$, so we have a LOCAL MAX. Here's a sketch of the surface from Maple




6. Use Matlab and our Newton's Method program to find two critical points of $f$ correct to three decimal places.

$$f(x,y) = x^4 + y^4 - 4x^2y + 2y$$

SOLUTION: First, we need a function file that will output the value of $f$, the gradient of $f$, and the Hessian of $f$ (although we won't use the value of $f$). Here's one script file as an example:

```
function [f,df,hf]=testfuncEx4(A)
%  Function for Exercise 4
%    We assume the input to be a vector, A=(x,y)
x=A(1); y=A(2);  %Not necessary, but will make it easier to read below.
f=x^4+y^4-4*x^2*y+2*y;
df=[4*x^3-8*x*y; 4*y^3-4*x^2+2];  %Newton's Method needs column (not row)
hf=[12*x^2-8*y, -8*x
    -8*x, 12*y^2];
```

Now that we have our function, take a look at the inputs to the multivariate Newton's Method code, and then we'll write a script file:

```
% Example Script for HW 4:
%  I'll make 5 runs with random starting points

out=zeros(2,5);  %To store output from Newton's Method

for K=1:5
    A=2*rand(2,1)-1;  %ordered pairs in the right domain
                      % (x,y) are in (-1,1) each.  This is
                      %  the initial guess for Newton.

    out(:,K)=MultiNewton('testfuncEx4',A,50,1e-3);
end
```

In the command window, I get the following when I run the script:

```
>> Ex4
Newton used 12 iterations
Newton used 9 iterations
Newton used 4 iterations
Newton used 5 iterations
Newton used 7 iterations
>> out

out =

   -0.0000   -1.5919    0.7192    0.0000   -0.0000
   -0.7937    1.2670    0.2587   -0.7937   -0.7937
```

7. Exercise: Compute the next step of steepest descent from the example in the last section.

SOLUTION: We recall that the function is

$$f(x,y) = 4x^2 - 4xy + 2y^2$$

And we now use the point $(0,1)$. From our previous computation, we already have the general form of the gradient, so evaluating it at $(0,1)$ gives $[-4,4]$. Therefore, for some value of $h$, we will set

$$\vec{x}_2 = \vec{x}_1 - h\nabla f(\vec{x}_1) = \begin{bmatrix} 0 \\ 1 \end{bmatrix} - h \begin{bmatrix} -4 \\ 4 \end{bmatrix} = \begin{bmatrix} 4h \\ 1-4h \end{bmatrix}$$

To find the optimal step size $h$, we evaluate the gradient along this path to obtain the function $\phi(t)$:

$$\phi(t) = f\left(\begin{bmatrix} 4t \\ 1-4t \end{bmatrix}\right)$$

(By the way, this problem is easy enough that you could do this in Maple to find the minimum directly- You might plot the $\phi(t)$ function for yourself).

$\phi'(t)$ (setting it to zero):

$$\phi'(t) = -\nabla f(4t, 1-4t) \cdot \nabla f(0,1) = 320t - 32$$

Therefore, the derivative is zero for $t = 1/2$, and the second derivative is positive, so this is indeed where the minimum occurs.

Now, using that step size:

$$\vec{x}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} - \frac{1}{10}\begin{bmatrix} -4 \\ 4 \end{bmatrix} = \frac{1}{5}\begin{bmatrix} 2 \\ 3 \end{bmatrix}$$

8. This may be obvious- We set

$$\vec{x}_{i+1} = \vec{x}_i - h\nabla f(\vec{x}_i)$$

where the $h$ was the solution to:

$$\nabla f(\vec{x}_i - t\nabla f(\vec{x}_i) \cdot \nabla f(\vec{x}_i) = 0$$

Therefore, this equation translates to saying:

$$\nabla f(\vec{x}_{i+1}) \cdot \nabla f(\vec{x}_i) = 0$$

Therefore, we are always "zig-zagging" around the domain when we perform gradient descent.