MA 230 Exam 3

1. (25 pt) Find and classify the critical points (determine whether they are a local minimum, local maximum, or saddle point) for $f(x,y) = x + y + \frac{x+y}{xy}$.

Solution:

$$\nabla f(x,y) = \begin{bmatrix} 1 - \frac{1}{x^2} \\ 1 - \frac{1}{y^2} \end{bmatrix}$$

Note that $xy \neq 0$ because the function isn't defined on those points. This means that points that satisfy $xy = 0$ are automatically not critical points. Only points such that $\nabla f(x,y) = \vec{0}$ or undefined are critical points. In this case, the points that make the gradient undefined are not critical points. This leaves these four points as critical points: $(1,1), (1,-1), (-1,1), (-1,-1)$.

$$\nabla^2 f(x,y) = \begin{bmatrix} \frac{2}{x^3} & 0 \\ 0 & \frac{2}{y^3} \end{bmatrix}$$

Using the second derivative test four times, we have that $(1,1)$ is a local minimum, $(-1,1), (1,-1)$ are saddle points, and $(-1,-1)$ is a local maximum.

2. (a) (8 pt) Let $f(x, y) > 0$ for all x and y. Show that the critical points of $f(x, y)$ are exactly the same as the critical points of $\sqrt{f(x, y)}$.

(b) (12 pt) The distance between a curve $z = f(x, y)$ and a point (x_0, y_0, z_0) is shown below.

$$d(x, y) = \sqrt{(x - x_0)^2 + (y - y_0)^2 + (f(x, y) - z_0)^2}$$

Using part (a), find the minimum distance between the plane $x + y + z = 1$ and the point $(2, -1, -2)$ and show that it is the global minimum.

(c) (15 pt) Let us now define our point parametrically along the curve $(2t, -t, -2t)$. Using the Envelope Theorem, find $x^*(t)$, $y^*(t)$ which minimize the distance from the plane $x + y + z = 1$ as functions of t.

Solution:

(a) The critical points of $f(x, y)$ are points such that the gradient at the point is equal to the zero vector or undefined.

Let $F(x, y) = \sqrt{f(x, y)}$. Then the gradient of F is below.

$$\nabla F = \frac{1}{2\sqrt{f(x, y)}} \begin{bmatrix} f_x \\ f_y \end{bmatrix} = \frac{1}{2\sqrt{f(x, y)}} \nabla f$$

Because $\frac{1}{2\sqrt{f(x,y)}}$ is never zero or undefined, the only source of critical points comes from ∇f. This means that both f and F share the same critical points.

(b) Using part (a), it is clear that the critical points of $d(x, y)$ and $d^2(x, y)$ are the same. It is much easier to find and classify the critical points of $d^2(x, y)$, so that is what will be done. Let $f(x, y) = d^2(x, y) = (x - x_0)^2 + (y - y_0)^2 + (g(x, y) - z_0)^2$ and $g(x, y) = 1 - x - y$. This can be simplified to the following:

$$\begin{aligned} f(x, y) &= (x - x_0)^2 + (y - y_0)^2 + (1 - x - y - z_0)^2 \\ &= (x - 2)^2 + (y + 1)^2 + (x + y - 3)^2 \\ &= 2x^2 + 2y^2 - 10x - 4y + 2xy + 14 \end{aligned}$$

$$\nabla f = \begin{bmatrix} 4x + 2y - 10 \\ 4y + 2x - 4 \end{bmatrix} \quad , \quad \nabla^2 f = \begin{bmatrix} 4 & 2 \\ 2 & 4 \end{bmatrix}$$

The partial derivatives are defined everywhere, so they are both set equal to zero. After some algebra, the only critical point is $x = \frac{8}{3}$ and $y = -\frac{1}{3}$.

Note that $f_{xx} > 0$ and $\det \nabla^2 f > 0$ for all x and y. This means that f is strictly convex and the point $(\frac{8}{3}, -\frac{1}{3})$ is a global minimum.

Finally, the minimal distance is $d(\frac{8}{3}, -\frac{1}{3}) = \sqrt{(\frac{2}{3})^2 + (\frac{2}{3})^2 + (\frac{2}{3})^2} = \boxed{\dfrac{2}{\sqrt{3}}}$.

(c) With unknown parameter t, we redefine our objective function to be

$$\begin{aligned} f(x, y, t) &= (x - (2t))^2 + (y - (-t))^2 + (z - (-2t))^2 \\ &= (x - 2t)^2 + (y + t)^2 + (1 - x - y + 2t)^2 \\ &= 2x^2 + 2y^2 - 2x - 2y + 2xy - 8tx - 2ty + 9t^2 + 4t + 1. \end{aligned}$$

We then get our gradient to be

$$\nabla f = \begin{bmatrix} 4x - 2 + 2y - 8t \\ 4y - 2 + 2x - 2t \end{bmatrix} = \vec{0}.$$

Manipulating the bottom equation, we get

$$x = 1 + t - 2y.$$

Substituting into the top equation, we get

$$2 - 6y - 4t = 0.$$

Thus, we conclude that $\boxed{y^*(t) = \dfrac{1 - 2t}{3}}$ and $\boxed{x^*(t) = \dfrac{1 + 7t}{3}}$.

3. We are given a series of data points (x_i, y_i) to be $(0,0)$, $(0,-1)$, $(2,0)$, and $(0,3)$. We want to find a line of best-fit, $y = mx + b$, that minimizes the squared error from the data. We define this error function as

$$E(m,b) = \sum_{i=1}^{4} (y_i - mx_i - b)^2.$$

(a) (2 pt) Verify that $E(m,b) = 4b^2 + 4m^2 + 4bm - 4b + 10$ for the given data.

(b) (13 pt) Find the values of m, b which minimize the error function, and find that minimum value. Show that the point is a local minimum.

(c) (20 pt) Let us now try to find the line of best-fit through Ridge Regression. This is a technique which uses the same error function, but has an additional constraint that values m, b lie on a circle centering the origin, or $m^2 + b^2 = k$. Solve for the values of m, b which minimize the error function with this constraint when $k = 1$, find the minimum value, and show the point is a local minimum.

(d) (5 pt) Estimate the change in the value of the error function if the constraint was to change to $k = 1.1$.

Solution:

(a) Plugging in the given data points, our error function is

$$\begin{aligned} E(m,b) &= \left((-b)^2 + (-1-b)^2 + (-2m-b)^2 + (3-b)^2\right) \\ &= \boxed{4b^2 + 4m^2 + 4bm - 4b + 10}. \end{aligned}$$

(b) Our gradient will be

$$\nabla E = \begin{bmatrix} 8m + 4b \\ 8b + 4m - 4 \end{bmatrix} = \vec{0}.$$

From here, we can easily get that $\boxed{m = -\frac{1}{3}}$ and $\boxed{b = \frac{2}{3}}$. To verify this is a local minimum, we need to find the Hessian:

$$\nabla^2 E = \begin{bmatrix} 8 & 4 \\ 4 & 8 \end{bmatrix}.$$

Since $\det(\nabla^2 E) > 0$, we can clearly see that our stationary point is a local minimum. We conclude that our minimum error is $\boxed{E\left(-\frac{1}{3}, \frac{2}{3}\right) = \frac{26}{3}}$.

(c) We begin by forming the Lagrangian

$$L(m, b, \lambda) = 4b^2 + 4m^2 + 4bm - 4b + 10 - \lambda(m^2 + b^2 - 1).$$

We now take its partial derivatives:

$$\frac{\partial L}{\partial m} = 4b + 8m - 2m\lambda = 0 \implies \lambda = \frac{2b+4m}{m}.$$

$$\frac{\partial L}{\partial b} = 8b + 4m - 4 - 2b\lambda = 0 \implies \lambda = \frac{4b+2m-2}{b}.$$

$$\frac{\partial L}{\partial \lambda} = m^2 + b^2 - 1 = 0 \implies m^2 + b^2 = 1.$$

Equating the formulas for the Lagrange multiplier, we get

$$\begin{aligned} \frac{2b+4m}{m} = \frac{4b+2m-2}{b} &\implies 2b^2 + 4bm = 4bm + 2m^2 - 2m \\ &\implies 2b^2 = 2m^2 - 2m \\ &\implies b^2 = m^2 - m. \end{aligned}$$

Plugging into our constraint, we now get that

$$m^2 + (m^2 - m) = 1 \implies 2m^2 - m - 1 = 0 \implies m = \frac{1 \pm \sqrt{1+8}}{4} = -\frac{1}{2}, 1.$$

This also means that $b = \pm\frac{\sqrt{3}}{2}, 0$. This means our stationary points are $(1,0)$, $\left(-\frac{1}{2}, \frac{\sqrt{3}}{2}\right)$, and $\left(-\frac{1}{2}, -\frac{\sqrt{3}}{2}\right)$. We now need to determine which point is the local minimum. We can immediately rule out $(1,0)$, since it does not satisfy the equality constraints. To compare the other two, we need to create the Hessian of the Lagrangian:

$$\nabla^2 L = \begin{bmatrix} 8 - 2\lambda & 4 \\ 4 & 8 - 2\lambda \end{bmatrix}.$$

For $\left(-\frac{1}{2}, \frac{\sqrt{3}}{2}\right)$, we get $\lambda = 4 - 2\sqrt{3}$. Since $det(\nabla^2 L\left(-\frac{1}{2}, \frac{\sqrt{3}}{2}\right) = 32 > 0$, we conclude that it is the local minimum. Therefore, we conclude that the minimum point is $\boxed{\left(-\frac{1}{2}, \frac{\sqrt{3}}{2}\right)}$ and the minimum value is $E\left(-\frac{1}{2}, \frac{\sqrt{3}}{2}\right) = \boxed{14 - 3\sqrt{3}}$.

(d) We can approximate the change in the value to be

$$\Delta E \approx \lambda(1.1 - 1) = (4 - 2\sqrt{3})\left(\frac{1}{10}\right) = \boxed{\frac{2-\sqrt{3}}{5}}.$$