Math 241 C1H - More Examples of Max/Min Problems

Example I: Least squares linear regression. (Taken from Susan Jane Colley's Vector Calculus text, pg. 271-274).

Suppose we have some pairs of data (x_1, y_1), (x_2, y_2), ..., (x_n, y_n), and we suspect that there is a linear relationship between betwen the xs and the ys, namely, that

$$y_i \approx mx_i + b$$

for some m and b. How do we find the best m and b?

We do this by minimizing the function

$$f(m,b) = \sum_{i=1}^{n}(y_i - (mx_i + b))^2 = \sum_{i=1}^{n} y_i^2 - 2m\sum_{i=1}^{n} x_iy_i - 2b\sum_{i=1}^{n} y_i + \sum_{i=1}^{n}(mx_i + b)^2.$$

Here, the variables are m and b, the slope and y-intercept of the line. We find that

$$\begin{aligned}\frac{\partial f}{\partial m} &= -2\sum_{i=1}^{n} x_iy_i + \sum_{i=1}^{n} 2(mx_i + b)x_i \\ &= -2\sum_{i=1}^{n} x_iy_i + 2m\sum_{i=1}^{n} x_i^2 + 2b\sum_{i=1}^{n} x_i.\end{aligned}$$

Then,

$$\begin{aligned}\frac{\partial f}{\partial b} &= -2\sum_{i=1}^{n} y_i + \sum_{i=1}^{n} 2(mx_i + b) \\ &= -2\sum_{i=1}^{n} y_i + 2m\sum_{i=1}^{n} x_i + 2nb.\end{aligned}$$

The critical point (m, b) thus satisfies the pair of equations

$$\begin{aligned}(\sum x_i^2)m + (\sum x_i)b &= \sum x_iy_i \\ (\sum x_i)m + nb &= \sum y_i.\end{aligned}$$

This is a system of two linear equations in m and b. The solution is

$$\begin{aligned}m &= \frac{n\sum x_iy_i - (\sum x_i)(\sum y_i)}{n\sum x_i^2 - (\sum x_i)^2} \\ b &= \frac{(\sum x_i^2)(\sum y_i) - (\sum x_i)(\sum x_iy_i)}{n\sum x_i^2 - (\sum x_i^2)}.\end{aligned}$$

Now, we'll check if this critical point is a local min or a local max. Differentiating again, we get that

$$\frac{\partial^2 f}{\partial m^2} = 2\sum x_i^2$$
$$\frac{\partial^2 f}{\partial m \partial b} = 2\sum x_i$$
$$\frac{\partial^2 f}{\partial b^2} = 2n.$$

Clearly $\frac{\partial^2 f}{\partial m^2} > 0$ if not all the x_is are zero. Next, the determinant of

$$Hf = \begin{bmatrix} \frac{\partial^2 f}{\partial m^2} & \frac{\partial^2 f}{\partial m \partial b} \\ \frac{\partial^2 f}{\partial b \partial m} & \frac{\partial^2 f}{\partial b^2} \end{bmatrix}$$

is

$$4n\sum x_i^2 - 4(\sum x_i)^2.$$

Is this positive or negative?

We will use the Cauchy-Schwarz inequality (see Exercise 1.3.10) to show that it is positive, and hence that the critical point found above is a local (and hence global) min.

Let $\vec{v}_1 = (x_1, x_2, \ldots, x_n)$ and let $\vec{v}_2 = (1, 1, 1, \ldots, 1)$ be vectors in $\mathbb{R}^n$. Then, the Cauchy-Schwarz inequality

$$|\vec{v}_1 \cdot \vec{v}_2| \leq \|\vec{v}_1\| \|\vec{v}_2\|$$

says that

$$|\sum x_i| \leq \sqrt{\sum x_i^2}\sqrt{n}.$$

Squaring, we get

$$(\sum x_i)^2 \leq n(\sum x_i^2).$$

In particular, $4n(\sum x_i^2) - 4(\sum x_i)^2 \geq 0$. We have equality when the vectors are parallel, which would mean that all the x_is are the same. This doesn't really make sense.

Example II: (From one of my own research papers).

We wish to maximize the function

$$f(x, y) = y^{12} e^{-4\pi y} \prod_{n=1}^{\infty} (1 - 2\cos(2\pi n x) e^{-2\pi n y} + e^{-4\pi n y})^{24}.$$

on $\{(x, y) : -1 \leq x \leq 0, x^2 + y^2 \geq 1, (x+1)^2 + y^2 \geq 1\}$.

This function is an infinite product (which is sort of like an infinite sum). I will not attempt to explain why I was interested in doing this, although you can come to office hours and ask me.

1. Because of the $e^{-4\pi y}$, the function decays exponentially as $y \to \infty$. Computing the value at $(-1/2, \sqrt{3}/2)$, we get

$$f(-1/2, \sqrt{3}/2) \approx 0.0000419041...$$

Because of the exponential decay, the maximum occurs when $y \leq 0.8676$.

2. We have $f(x, y) = f(1 - x, y)$. This implies that $f_x(-1/2, y) = 0$. It is also possible to show that $f_{xx}(x, y) < 0$ for all (x, y) in the domain. This implies that $f_x(x, y) \leq 0$ with equality only when $x = -1/2$. This implies that any critical points are on the line $x = -1/2$.

3. The function satisfies

$$f\left(\frac{-x^2 - y^2 - x}{x^2 + y^2}, \frac{y}{x^2 + y^2}\right) = f(x, y).$$

(This is not at all obvious). Differentiating with respect to y and setting $x = -1/2$, $y = \sqrt{3}/2$ shows that $f_y(-1/2, \sqrt{3}/2) = 0$. Also, it is possible to show that $f_{yy}(-1/2, y) < 0$ for any $y \geq \sqrt{3}/2$. This implies that the only critical point is at $(-1/2, \sqrt{3}/2)$.

4. As we showed, $f_{xx}(-1/2, \sqrt{3}/2)$ and $f_{yy}(-1/2, \sqrt{3}/2)$ are both negative. Thus, $D(-1/2, \sqrt{3}/2) = f_{xx}f_{yy} - f_{xy}^2 < 0$. Hence, this point is a local maximum. Since it is the only critical point, it is the global maximum. (For technical reasons, in this case it is not necessary to carefully check the values on the boundary of the domain).

The value at this maximum point is

$$f(-1/2, \sqrt{3}/2) = \left(\frac{\sqrt{2}\pi}{3c^3}\right)^{24},$$

where $c = \int_0^\infty x^{-1/3} e^{-x}\, dx \approx 1.35412....$