Bioinformatics

Discrete Mathematics and Optimisation

Solutions to Problem Sheet Extrema in Several Variables

Exercise 1.

(a) The gradient of f is $\nabla f(x,y) = (4x^3 - 16x, y^3 - 1)$. The critical points of f are real solutions of $\nabla f(x,y) = 0$, that is

$$\begin{aligned} &4x^3 - 16x = 0 \implies 4x(x^2-4) = 0 \implies x = 0 \text{ or } x = \pm 2, \\ &y^3 - 1 = 0 \implies y = \sqrt[3]{1} \implies y = 1, \text{ for } y \in \mathbb{R}. \end{aligned}$$

Thus the critical points of f are $(-2,1)$, $(0,1)$ and $(2,1)$.

(b) The Hessian of f at any point $(x,y) \in \mathbb{R}^2$ is given by

$$H_f(x,y) = \begin{pmatrix} 12x^2 - 16 & 0 \\ 0 & 3y^2 \end{pmatrix}.$$

In particular,

$$H_f(0,1) = \begin{pmatrix} -16 & 0 \\ 0 & 3 \end{pmatrix} \quad \text{and} \quad H_f(-2,1) = H_f(2,1) = \begin{pmatrix} 32 & 0 \\ 0 & 3 \end{pmatrix}.$$

We have $|H_f(-2,1)| = |H_f(2,1)| = 96 > 0$ and $32 > 0$. By Sylvester's criterion, this implies that $H_f(-2,1)$ and $H_f(2,1)$ are definite positive matrices. Thus $(-2,1)$ and $(2,1)$ are (strict) local minima of f.

On the other hand $|H_f(0,1)| = -48$, which implies that $(0,1)$ is a non-degenerate critical point of f and it is neither positive definite nor negative definite using Sylvester's criterion. Therefore it $(0,1)$ is a saddle point.

Exercise 2.

(a) We have $\nabla f(x,y) = (2x, 3y^2)$ and $\nabla g(x,y) = (2x, 4y^3)$. Since the point $(0,0)$ is solution of $\nabla f(0,0) = 0$ and $\nabla g(0,0) = 0$, it is a critical point of both f and g.

Furthermore, for $(x,y) \in \mathbb{R}^2$ we have

$$H_f(x,y) = \begin{pmatrix} 2 & 0 \\ 0 & 6y \end{pmatrix} \quad \text{and} \quad H_g(x,y) = \begin{pmatrix} 2 & 0 \\ 0 & 12y^2 \end{pmatrix}.$$

So that

$$H_f(0,0) = H_g(0,0) = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}.$$

And the determinant of both $H_f(0,0)$ and $H_g(0,0)$ is zero. This means that $(0,0)$ is a degenerate critical point of both f and g.

(b) As x and y appear in g solely as even powers, it is clear that $g(x, y) \geq 0 = g(0, 0)$ for any $(x, y) \in \mathbb{R}^2$. Thus $(0, 0)$ is a local minimum of g (it is actually even a global minimum in $\mathbb{R}^2$).

On the other hand, x appears as an odd power in f. So that $f(0, -\epsilon) = -\epsilon^3 < 0 = f(0, 0)$ for any $\epsilon > 0$. And $(0, 0)$ cannot be a local minimum of f.

Exercise 3.

(a) The gradient of f is

$$\nabla f(x, y, z) = (e^x - 2e^{-x-y-z}, e^y - 2e^{-x-y-z}, e^z - 2e^{-x-y-z}).$$

Given the symmetries of the function, notice that a critical point must satisfy $x = y = z$. Thus, we are looking for real solutions of $e^x - 2e^{-3x} = 0$, that is $(x, y, z) = \big(\ln(2)/4, \ln(2)/4, \ln(2)/4\big)$.

The Hessian of f at any point $(x, y, z) \in \mathbb{R}^3$ is given by

$$H_f(x, y, z) = \begin{pmatrix} e^x + 2e^{-x-y-z} & 2e^{-x-y-z} & 2e^{-x-y-z} \\ 2e^{-x-y-z} & e^y + 2e^{-x-y-z} & 2e^{-x-y-z} \\ 2e^{-x-y-z} & 2e^{-x-y-z} & e^z + 2e^{-x-y-z} \end{pmatrix}.$$

Denoting the determinant of the upper-left $i{-}by{-}i$ corner, i.e. the i-th leading principal minor, by Δ_i we have

$$\begin{aligned} \Delta_1 &= e^x + 2e^{-x-y-z} > 0, \\ \Delta_2 &= e^{x+y} + 2e^{-y-z} + 2e^{-x-z} > 0 \text{ and} \\ \Delta_3 &= e^{x+y+z} + 2e^{-z} + 2e^{-y} + 2e^{-x} > 0. \end{aligned}$$

And Sylvester's criterion implies that $H_f(x, y, z)$ is definite positive for any $(x, y, z) \in \mathbb{R}^3$. In particular, it is definite positive at $\big(\ln(2)/4, \ln(2)/4, \ln(2)/4\big)$, and that critical point is hence a (strict) local minimum of f.

(b) The Hessian of f is definite positive at every point of the domain of f, this implies that f is a convex function. It has thus a unique global minimum, and the local minimum $\big(\ln(2)/4, \ln(2)/4, \ln(2)/4\big)$ is hence a global minimum.

Exercise 4.

(a) The gradient of f is $\nabla f(x, y) = (4x^3 - 4y, -4x + 4y^3)$. Thus the critical points of f are $P_1 = (0, 0)$, $P_2 = (1, 1)$ and $P_3 = (-1, -1)$. The Hessian of f at a point $(x, y) \in \mathbb{R}^2$ is given by

$$H_f(x, y) = \begin{pmatrix} 12x^2 & -4 \\ -4 & 12y^2 \end{pmatrix}.$$

So that

$$H_f(P_1) = \begin{pmatrix} 0 & -4 \\ -4 & 0 \end{pmatrix} \quad \text{and} \quad H_f(P_2) = H_f(P_3) = \begin{pmatrix} 12 & -4 \\ -4 & 12 \end{pmatrix}.$$

Using Sylvester's criterion, we see that $H_f(P_2)$ and $H_f(P_3)$ are definite positive. This means that P_2 and P_3 are (strict) local minima of f. We also have that $H_f(P_1)$ is not definite as it's first principal minor is 0. Therefore as $|H_f(P_1)| = 16 \neq 0$, P_1 is non-degenerate and not definite and therefore P_1 is a saddle point of f.

(b) First observe that

$$4\max(x^4, y^4) \geq f(x,y) \geq \max(x^4, y^4) - 4\max(x^2, y^2).$$

Furthermore, whenever $|(x,y)| = \sqrt{x^2+y^2} \to \infty$ we have both $\max(x^4, y^4) \to \infty$ and

$$\max(x^4, y^4) - 4\max(x^2, y^2) = \max(x^2, y^2)(\max(x^2, y^2) - 4) \to \infty.$$

Thus $f(x,y) \to \infty$, as $|(x,y)| \to \infty$.

(c) By (a), $|P_2| = |P_3| = 1$ are the only local minima of f in $\mathbb{R}^3$.

Furthermore, (b) implies that there are global minima as the function tends to ∞ in all directions and so it is bounded below. Thus the global minima must be local minima and as $f(P_2) = f(P_3)$ we have that both P_2 and P_3 are global minima.

Exercise 5. We need to maximise the function

$$\begin{aligned} f(x) &= \mathbb{P}(\text{we observe } k_1, k_2, k_3 \mid \theta = x) \\ &= \binom{k_1+k_2+k_3}{k_1}(x^2)^{k_1} \cdot \binom{k_2+k_3}{k_2}(2x(1-x))^{k_2} \cdot \binom{k_3}{k_3}((1-x)^2)^{k_3}. \end{aligned}$$

We note that this function achieves a maximum at the same point where the log of this function achieves a maximum. We set

$$g(x) = \log(f(x)) = \text{Constant} + 2k_1\log(x) + k_2\log(x) + k_2\log(1-x) + 2k_3\log(1-x),$$

so that

$$g'(x) = \frac{2k_1+k_2}{x} - \frac{k_2+2k_3}{1-x}.$$

This means that g (and hence f) has a critical point at $x = (2k_1+k_2)/(2k_1+2k_2+2k_3)$. One can then readily to verify that this is indeed a global maximum.

Exercise 6.

(a) The gradient of f is $\nabla f(x,y) = (3x^2 - 12y, -12x + 24y^2)$. The critical points of f are real solutions of $\nabla f(x,y) = (0,0)$, given by $P_1 = (0,0)$ and $P_2 = (2,1)$.

(b) The Hessian matrix of f at a point $(x,y) \in \mathbb{R}^2$ is given by

$$H_f(x,y) = \begin{pmatrix} 6x & -12 \\ -12 & 48y \end{pmatrix}.$$

In particular,

$$H_f(P_1) = \begin{pmatrix} 0 & -12 \\ -12 & 0 \end{pmatrix} \quad \text{and} \quad H_f(P_2) = \begin{pmatrix} 12 & -12 \\ -12 & 48 \end{pmatrix}.$$

And we have $|H_f(P_1)| = -144 < 0$. Using Sylvester's criterion, this implies that $H_f(P_1)$ is indefinite so that P_1 is a saddle-point of f.

While $|H_f(2,1)| = 144 \cdot 5 > 0$ and $12 > 0$ imply by Sylvester's criterion that $H_f(P_2)$ is definite positive. Thus P_2 is a (strict) local minimum of f.