Problems marked with a (*) are a bit more complex and can be skipped at a first read.

If you don't have a lot of time focus on the Problems/subquestions marked with (♡).

6.1. BONUS PROBLEM. Let $f \in C^2(\mathbb{R}^n)$ be a convex function.

(a) Show that $z \in \mathbb{R}^n$ is a critical point of f if and only if z is a global minimizer.

(b) Provide an example of such an f in some $\mathbb{R}^n$ with $n > 1$, which is always nonnegative, but does not have a minimum point. That is to say

$$f(x) > \inf_{\mathbb{R}^n} f \geq 0 \text{ for all } x \in \mathbb{R}^n.$$

You can use all the Theorems seen in class.

6.2. The signature of a 2×2 matrix. Despite the definition, it is not necessary to compute the eigenvalues of a matrix to find its signature[1]. Prove that for a 2×2 matrix M we have the following simple rule to determine the signature in terms of the $\det M$ and $\mathrm{Tr} M$:

- If $\det M > 0, \mathrm{Tr} M > 0$ then M is positive definite,
- If $\det M > 0, \mathrm{Tr} M < 0$ then M is negative definite,
- If $\det M < 0$ then M is indefinite,
- If $\det M = 0$, then M is degenerate.

6.3. Isoperimetric triangles. Among all the triangles with perimeter equal to 2, find the ones with the largest area. You may give for granted Heron's formula, which gives the area of a triangle in terms of the length of its sizes x, y, z:

$$A = \sqrt{p(p-x)(p-y)(p-z)}, \qquad \text{with } 2p := x + y + z,$$

so that in our case $p = 1$.

6.4. Barycenter (♡). Let $y_1, \ldots, y_k \in \mathbb{R}^n$ be given. Show that there is exactly one point for which

$$f(x) = \|x - y_1\|^2 + \cdots + \|x - y_k\|^2, \quad x \in \mathbb{R}^n$$

is minimal and determine this point.

6.5. Linear regression I (♡). You study the house market in Zürich over a year in which N houses are sold. You keep track of the size of the houses $x_1, \ldots, x_N$ and the respective sale prices $y_1, \ldots, y_N$. Now you would like would like to find "the" function $f \colon \mathbb{R} \to \mathbb{R}$ that gives

$$\text{sale price} = f(\text{size of the house}),$$

[1] Ask ChatGPT about the Principal Minor Theorem

and you make the (not unreasonable) assumption that f is affine, i.e., $f_{a,b}(x) = ax + b$ for some coefficients $a, b \in \mathbb{R}$. Among all such functions find (in terms of the data you collected) the value of the parameters a, b that minimizes the average quadratic error

$$E(a,b) := \sum_{i=1}^{N} (y_i - f_{a,b}(x_i))^2, \qquad a, b \in \mathbb{R}.$$

6.6. Convex functions (♡). Decide whether the following functions f_i are convex in the convex domain $U_i \subset \mathbb{R}^n$. Try to find, in each case, the simplest argument, you can almost always avoid lengthy computations.

1. $f_1(x,y) = x^2 + y^2 - 4y$ defined in $U_1 = \mathbb{R}^2$
2. $f_2(x,y) = x^2 + y^2 - y^4$ defined in $U_2 = \{(x,y) \in \mathbb{R}^2 : x^2 + y^2 < \frac{1}{10000}\}$
3. $f_3(x,y) = x^2 + y^2 - 4xy$ defined in $U_3 = \mathbb{R}^2$
4. $f_4(x,y) = x^2 + y^2 - 4xy$ defined in $U_4 = \{(x,y) \in \mathbb{R}^2 : 0 < 10x < |y|\}$
5. $f_5(x) = \phi(g(x)), x \in U_5$ where $g \in C^2(U_5)$ is any convex function in $U_5 \subset \mathbb{R}^n$ and $\phi \in C^2(\mathbb{R})$ is any convex and increasing function.
6. $f_6(x,y) = (1 + x^2 + y^2)^{1/2}$ defined in $U_6 = \mathbb{R}^2$
7. $f_7(x,y) = -(1 + x^2 + y^2)^{-1/2}$ defined in $U_7 = \mathbb{R}^2$
8. $f_8(x) = \sum_{i=1}^{n} |x_i|^p$ in $U_8 = \mathbb{R}^n$, where $p \geq 1$ is some fixed exponent.
9. $f_9(x) = \max\{\phi(x), \psi(x)\}$ where $\phi, \psi \in C(U_9)$ are any pair of convex functions defined in some open set $U_6 \subset \mathbb{R}^n$.
10. $f_{10}(x) = |x|$ defined in $U_{10} = \mathbb{R}^n$.
11. $f_{11}(x) = \phi(|x|)$ in $U_{11} = B_1 \subset \mathbb{R}^n$, where $\phi \in C(\mathbb{R})$ is any convex function.

6.7. Multiple choice. Among the following statements about convex functions mark those (and only those) which are always true.

(a) If $f \in C^1(U)$ is convex in some open convex set $U \subset \mathbb{R}^n$ and f has a local maximum at $z \in U$, then $\nabla f \equiv 0$ in U.

(b) If $f \in C^1(U)$ is convex in some open convex set $U \subset \mathbb{R}^n$ and f has a global maximum at $z \in U$, then $\nabla f \equiv 0$ in U.

(c) Assume $f_n \in C^2(\mathbb{R})$ is a sequence of convex functions that converge pointwise to some $f \colon \mathbb{R} \to \mathbb{R}$. Is f necessarily convex?

(d) There exists a convex function $f \in C^\infty(\mathbb{R}^2)$ such that

$$f(x) = 1 - 2x_1 + x_2^3 + O(|x|^4) \text{ as } |x| \to 0.$$

(e) There exists a convex function $f \in C^\infty(\mathbb{R}^2)$ such that

$$f(x) = 1 - 2x_1 + x_2^4 + O(|x|^4) \text{ as } |x| \to 0.$$

(f) A convex set is not necessarily connected.

6.8. Multiple choice. The Hessian matrix of $f \in C^2(\mathbb{R}^n)$ is positive semidefinite at a critical point x_0 of f, i.e.,

$$\langle v, Hf(x_0)v\rangle \geq 0 \text{ for all } v \in \mathbb{R}^n.$$

Which of the following statements necessarily hold? (There may be more than one).

(a) x_0 is a strict local minimum of f.

(b) x_0 is a local minimum of f.

(c) x_0 is not a local maximum of f.

(d) None of the above statements.

6.9. Minimization. The function $f\colon \mathbb{R}^2 \to \mathbb{R}$ is given by $f(x, y) = 2x^2 + y^2 - x$. Determine the extrema of f on...

(a) ...the unit circle $\mathbb{S}^1 = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}$;

(b) ...the closed unit disk $\mathbb{D} = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq 1\}$.

6.10. Lagrange Multipliers (♡). Consider the function $f(x, y, z) = 3x - y + 2z$ and the set

$$M = \{(x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 + z^2 = 1, x + y = 0\}.$$

Determine the extrema of f on M and their nature.

6.11. Critical Points. Let $f : \mathbb{R}^2 \to \mathbb{R}$ be the function $f(x, y) = (ax^2 + by^2)e^{-x^2-y^2}$ with real parameters $a, b \in \mathbb{R}$. Find all critical points and determine their nature with the Hessian test, depending on a, b.

Hints:

6.2 Use the spectral Theorem and the properties:

$$\det(AB) = \det(A)\det(B), \quad \mathrm{Tr}(AB) = \mathrm{Tr}(BA).$$

6.3 Minimize A^2 instead of A. You can use the method of Lagrange multipliers.

6.5 Do not get distracted by the setting, you after all you have to minimize a quadratic polynomial of a, b....

6. Solutions

Solution of 6.1:

(a) If $z \in \mathbb{R}^n$ is a global minimum point then $\nabla f(z) = 0$ by Proposition 11.4. If $\nabla f(z)$ then by Proposition 11.24 it holds

$$f(y) \geq f(z) + \nabla f(z) \cdot (y - z) = f(z) \text{ for all } y \in \mathbb{R}^n,$$

which means — by definition — that z is a global minimizer.

(b) $f(x) := e^{x_1+x_2}$ works. Clearly $f(x) > 0$ for all $x \in \mathbb{R}^2$, and $\lim_{t\to\infty} f(-t,-t) = \lim_{t\to\infty} e^{-2t} = 0$, so $\inf_{\mathbb{R}^2} f = 0$. Furthermore f is convex since for all unit vectors v we have

$$\partial_v f(x) = e^{x_1+x_2}(v_1 + v_2),$$

and

$$\partial_{vv} f(x) = e^{x_1+x_2}(v_1 + v_2)^2 \geq 0.$$

Solution of 6.2: By the Spectral Theorem there is $O \in \mathbb{R}^{2\times 2}$ such that

$$OMO^T = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}, \qquad OO^T = \mathbf{1} \text{ (in particular } \det O = \pm 1),$$

with $\lambda_1, \lambda_2 \in \mathbb{R}$. Thus we can see the (perhaps well-known) fact that

$$\det M = \det(OMO^T) = \lambda_1\lambda_2, \quad \mathrm{Tr} M = \mathrm{Tr}(MO^TO) = \mathrm{Tr}(OMO^T) = \lambda_1 + \lambda_2.$$

It is immediate now to check that

- $\lambda_1\lambda_2 > 0, \lambda_1 + \lambda_2 > 0$ implies that $\lambda_1 > 0$ and $\lambda_2 > 0$.
- $\lambda_1\lambda_2 > 0, \lambda_1 + \lambda_2 < 0$ implies that $\lambda_1 < 0$ and $\lambda_2 < 0$.
- $\lambda_1\lambda_2 < 0$ implies that one of the λ_i is positive and the other negative.
- $\lambda_1\lambda_2 = 0$ implies that one of them is zero.

Solution of 6.3: The set open of possible (x, y, z) that represent sides of a triangle is given by

$$U := \{(x, y, z) \in \mathbb{R}^3 : x > 0, y > 0, z > 0, x + y > z, y + z > x, z + x > y\}.$$

We seek to minimize the area, which is the same as minimizing the its square, which by Heron's formula is given by

$$f(x, y, z) = A^2 = (1 - x)(1 - y)(1 - z), \qquad (x, y, z) \in U$$

subject to the constraint $x + y + z = 2$. By Weierstrass Theorem f admits a maximum point $(x_0, y_0, z_0) \in \overline{U} \cap \{x + y + z = 2\}$.

Case 1: (x_0, y_0, z_0) lies on ∂U. Notice that any point in the closure $\overline{U}$ needs to satisfy

$$U := \{(x, y, z) \in \mathbb{R}^3 : x \geq 0, y \geq 0, z \geq 0, x + y \geq z, y + z \geq x, z + x \geq y\},$$

as can be seen taking a converging sequence of elements of U. Thus, an element on the boundary $\partial U = \overline{U} \setminus U$, must satisfy at least one among the following (otherwise it would lie in the interior)

$$x = 0, \quad y = 0, \quad z = 0, \quad x + y = z, \quad y + z = x, \quad z + x = y.$$

In all these cases we have that our triangle degenerates into a segment (one of the sides is zero!) and so f must vanish at these points, which must be minimum points. Let us check this:

- If (say) $x = 0$, then $y \geq z$ and $z \geq y$, so $z = y$. Since the perimeter is 2 we find $y = z = 1$, which leads to $f(0, 1, 1) = 0$. The other cases are symmetric.
- If (say) $x + y = z$, then $2 - z = z$, thus $z = 1$ and so $f(x, y, 1) = 0$. The other cases are symmetric.

We conclude that the "battle for the maximum" is fought in the interior of U.

Case 2: (x_0, y_0, z_0) lies in the interior of U. The method of Lagrange multipliers applies with $g(x, y, z) := x + y + z - 1$ and, since ∇g is never zero, gives the system

$$\begin{cases} (1-y)(1-z) = \lambda, \\ (1-z)(1-x) = \lambda, \\ (1-x)(1-y) = \lambda \\ x + y + z = 2, \\ (x, y, z) \in U, \lambda \in \mathbb{R}. \end{cases}$$

We solve this system. First we claim that no one among x, y, z can be equal to one. For example assume $z = 1$, then necessarily $\lambda = 0$ and then also $(1-x)(1-y) = 0$ so $x = 1$ or $y = 1$. Since $x + y + z = 2$ we find that one among x, y is zero, contradicting $(x, y, z) \in U$.

Now proceed taking the difference of the first two equation, finding

$$0 = (1 - z)(x - y),$$

which, since $z \neq 1$, leads to $x = y$. If we take now the difference between the second two equations we find

$$(1 - x)(y - z) = 0,$$

and, since $x \neq 1$, we find $y = z$. Thus the only solution of this system needs to satisfy

$$x = y = z, x + y + z = 2 \Rightarrow x = y = z = \tfrac{2}{3},$$

since we assumed $(x_0, y_0, z_0) \in U$ we know that it solves the system, hence

$$(x_0, y_0, z_0) = (\tfrac{2}{3}, \tfrac{2}{3}, \tfrac{2}{3}), f(\tfrac{2}{3}, \tfrac{2}{3}, \tfrac{2}{3}) = 1/27$$

which corresponds to the equilateral triangle.

Solution of 6.4: We start observing that f is a sum of convex functions hence it is convex. Furthermore f is smooth (it is a polynomial). Since f is convex, a point is a global minimizer if and only if it is a critical point. We prove that there is a unique critical point, proving simultaneously that there exist a minimum point, and that it is unique.

We obtain for the partial derivatives of f

$$\partial_{x_j} f(x) = 2\sum_{i=1}^{k} (x_j - (y_i)_j)\,.$$

For x to be a critical point, $\partial_{x_j} f(x) = 0$ for all j, hence $x_j = \frac{1}{k}\sum_{i=1}^{k}(y_i)_j$. Direct calculation yields

$$\partial_{x_l}\partial_{x_j} f(x) = 2k\delta_{lj},$$

thus H is diagonal with entries $2k$ for all $x \in \mathbb{R}^n$. We obtain that H is positive definite, hence at the critical point, there is a minimum, and the point defining the minimum is uniquely determined by the equality $x_j = \frac{1}{k}\sum_{i=1}^{k}(y_i)_j$.

Solution of 6.5: We are asked to minimize the error function

$$E : (a,b) \mapsto \sum_{i=1}^{N}(ax_i + b - y_i)^2,$$

given the fixed parameters $x_1, \dots, x_N, y_1, \dots, y_N$. E is the sum of squares of affine functions, so it is convex and $C^\infty(\mathbb{R}^2)$. Thus, any critical point will be a global minimizer, such a point is found solving

$$\begin{cases} 0 = \partial_1 E(a,b) = 2\sum_{i=1}^{n}(ax_i + b - y_i)x_i \\ 0 = \partial_2 E(a,b) = \sum_{i=1}^{n}(ax_i + b - y_i) \end{cases}$$

which is equivalent to the linear system

$$\begin{bmatrix} \sum_{i=1}^n x_i^2 & \sum_{i=1}^n x_i \\ \sum_{i=1}^n x_i & n \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^n x_i y_i \\ \sum_{i=1}^n y_i \end{bmatrix}.$$

Whose solution is

$$a = \frac{n\sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{n\sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2}, \quad b = \frac{-\sum_{i=1}^n x_i \sum_{i=1}^n x_i y_i + \sum_{i=1}^n x_i^2 \sum_{i=1}^n y_i}{n\sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2},$$

where we used

$$\begin{bmatrix} \sum_{i=1}^n x_i^2 & \sum_{i=1}^n x_i \\ \sum_{i=1}^n x_i & n \end{bmatrix}^{-1} = \frac{1}{(\sum_{i=1}^n x_i^2)\cdot n - (\sum_{i=1}^n x_i)^2} \begin{bmatrix} n & -\sum_{i=1}^n x_i \\ -\sum_{i=1}^n x_i & \sum_{i=1}^n x_i^2 \end{bmatrix}$$

This system can be solved uniquely if the determinant is nonzero, that is if

$$n\sum_{i=1}^n x_i^2 - \Big(\sum_{i=1}^n x_i\Big)^2 \neq 0,$$

but this is always a positive number by the QM-AM inequality, unless the $x_1 = \ldots = x_N$ are all equal. In this case the system becomes

$$\begin{bmatrix} x_1^2 & x_1 \\ x_1 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \frac{\sum_{i=1}^n y_i}{n} \begin{bmatrix} x_1 \\ 1 \end{bmatrix},$$

which is solved by $a = 0$ and $b = \frac{\sum_{i=1}^n y_i}{n}$.

Solution of 6.6:

- f_1 is convex since its Hessian matrix is $2\mathbf{1}_{2\times 2}$ which is positive definite.

  Similarly,

  $$Hf_2(x,y) = \begin{bmatrix} 2 & 0 \\ 0 & 2 - 12y^2 \end{bmatrix},$$

  which is positive definite if and only if $2 - 12y^2 > 0$, and this is true in the given domain U (but not in the whole $\mathbb{R}^2$...).

- f_3 and f_4 are not convex since their restriction to the line $t \mapsto (t,t)$ is **not** convex

  $$f_3(t,t) = f_4(t,t) = -2t^2.$$

- We observe that $|\cdot|$ is convex (like any other norm), this can be checked with the definition:

  $$|tx + (1-t)y| \le t|x| + (1-t)|y|$$

  were we used the positive homogeneity and the triangular inequality. This proves that f_{10} is convex.

- We show that also f_5 is convex, with a computation using the chain rule, first

  $$\partial_v(\phi(g(x))) = \phi'(g(x))\partial_v g(x),$$

  and then

  $$\partial_{vv}(\phi(g(x))) = \underbrace{\phi''(g(x))}\geq 0(\partial_v g(x))^2 + \underbrace{\phi'(g(x))\partial_{vv}g(x)}\geq 0 \geq 0.$$

  It immediately follows that f_6 is convex, since $t \mapsto (1+t^2)1/2$ is convex (Analysis I) and so is $x \mapsto |x|$.

  Similarly f_8 is convex because it is a sum of convex functions. Since $x \mapsto x_i$ is convex and so is $t \mapsto |t|^p$, then each $|x_i|^p$ is convex.

- f_7 is not convex since its restriction to the line $t \mapsto (t,t)$ is the function $t \mapsto -(1+2t^2)-1/2$ which is not convex (Analysis I).

- f_9 is always convex, just write both the convexity inequalities

$$\phi(tx+(1-t)y) \le t\phi(x)+(1-t)\phi(y) \le tf_6(x)+(1-t)f_6(x),$$
$$\psi(tx+(1-t)y) \le t\psi(x)+(1-t)\psi(y) \le tf_6(x)+(1-t)f_6(x),$$

thus the same bound holds for $\max\{\phi(tx+(1-t)y), \psi(tx+(1-t)y)\}$.

- It is not necessarily convex, take $f_{11}(x) := e^{-|x|}$ which corresponds to $\phi(t) = e^{-t}$. f_{11} cannot be convex since it reaches its absolute maximum value (i.e., 1) in the interior point $x = 0$ and it is not constant.

Solution of 6.7:

(a) False. Take $f(x)+\max\{x_1, \frac{1}{2}\}^4$, $U = B_1 \subset \mathbb{R}^n$ which is convex, has a local maximum at $x = 0$ (it is locally constant there), but its gradient does not vanish everywhere.

(b) True. Since z is a local maximum $\nabla f(z) = 0$, but f is convex so

$$f(x) \ge f(z) + \nabla f(z)\cdot(y-z) = f(z) \text{ for all } x \in U,$$

but on the other hand $f(x) \le f(z)$ for all $x \in U$, because z is a local maximum. We conclude that f is necessarily constant, thus its gradient vanish identically.

(c) True, pick any $x, y \in U$ and $t \in [0,1]$, by definition we known

$$f_n(tx+(1-t)y) \le tf_n(x)+(1-t)f_n(y) \text{ for all } n \in \mathbb{N},$$

if we keep x, y, t fixed and let $n \to \infty$ we find the convexity inequality for f at those points. Since x, y, t were arbitrary we conclude that f had to be convex.

(d) False. A convex function lies above its tangent plane at a point, thus in this case

$$f(x) \ge 1 - 2x_1 \text{ for all } x \in \mathbb{R}^n,$$

(we find it from the Taylor expansion at $x = 0$). But then we infer that for some small $\rho > 0$ and large $M > 0$ it must hold

$$1 - 2x_1 + x_2^3 + M|x|^4 \ge f(x) \ge 1 - 2x_1 \text{ for all } |x| < \rho,$$

which reshuffling terms is

$$x_2^3 \ge -M|x|^4 \text{ for all } |x| < \rho,$$

which is impossible if we take $x = (0, -r, 0, \dots, 0)$ and let $r \downarrow 0$.

(e) True, $f = 1 - 2x_1 + x_2^4$ is itself convex.

(f) False, by definition it is pathwise connected by straight segments, so it is connected.

Solution of 6.8:

(a) False, $f \equiv 0$ is a counterexample.

(b) False, $f = x_1^3$ at $x_0 = 0$ is a counterexample.

(c) False, $f = -x_1^4$ at $x_0 = 0$ is a counterexample.

(d) True because all the other statements are indeed false.

Solution of 6.9:

(a) We use Lagrange multipliers. To begin with, we observe that $\mathbb{S}^1 = g^{-1}(0)$, where $g\colon \mathbb{R}^2 \to \mathbb{R}$ is given by $g(x, y) = x^2 + y^2 - 1$. Hence, a Lagrange function $L\colon \mathbb{R}^3 \to \mathbb{R}$ for our problem is given by

$$L(x, y, \lambda) = f(x, y) - \lambda g(x, y)$$

We proceed as in Example 11.7. According to 11.6, at a local extremum $p = (x, y) \in g^{-1}(0)$, the following equations must be satisfied:

$$0 = \partial_x L(x, y, \lambda) = 2x(2 - \lambda) - 1, \tag{1}$$
$$0 = \partial_y L(x, y, \lambda) = y(2 - 2\lambda), \tag{2}$$
$$0 = \partial_\lambda L(x, y, \lambda) = -(x^2 + y^2 - 1). \tag{3}$$

We make a case distinction based on equation (2):

(i) If $y = 0$: Then $x^2 = 1$ by 3, i.e., $x = \pm 1$. Thus, there are local extrema at $(1, 0)$ and $(-1, 0)$ with function values $f(1, 0) = 1$ and $f(-1, 0) = 2$.

(ii) If $\lambda = 1$: Then $x = \frac{1}{2}$ by 1. Substituting into 3 yields $y = \pm\frac{\sqrt{3}}{2}$. Hence, there are additional local extrema at $(\frac{1}{2}, \frac{\sqrt{3}}{2})$ and $(\frac{1}{2}, -\frac{\sqrt{3}}{2})$ with function values $f(\frac{1}{2}, \frac{\sqrt{3}}{2}) = f(-\frac{1}{2}, -\frac{\sqrt{3}}{2}) = \frac{3}{4}$.

Comparing function values at all local extrema, we see that the function f on $\mathbb{S}^1$ has a global maximum at $(-1, 0)$ and two global minima at $(\frac{1}{2}, \frac{\sqrt{3}}{2})$ and $(\frac{1}{2}, -\frac{\sqrt{3}}{2})$.

(b) Since we have already examined the boundary $\mathbb{S}^1 = \partial\mathbb{D}$ in part (a), we only need to check the interior for extrema. To do this, we compute

$$Df(x, y) = (4x - 1, 2y).$$

At a local extremum $p = (x, y)$ of f, we must have

$$0 = Df(x, y) = (4x - 1, 2y),$$

i.e., $p = (x, y) = (\frac{1}{4}, 0)$. Evidently, p lies in the interior of the unit disk $\mathbb{D}^\circ = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 < 1\}$. We compute the Hessian matrix

$$H_f(x, y) = \begin{bmatrix} 4 & 0 \\ 0 & 2 \end{bmatrix},$$

which is positive definite, implying a local minimum at p.

The function value of f at p is $f(p) = f(\frac{1}{4}, 0) = -\frac{1}{8}$. Comparing this function value with those from part (a), we see that f on the closed unit disk $\mathbb{D}$ has a global minimum at $(\frac{1}{4}, 0)$ and a global maximum at $(-1, 0)$.

Solution of 6.10: The Lagrange function corresponding to f and M is given by

$$L(x, y, z, \lambda_1, \lambda_2) = 3x - y + 2z - \lambda_1(x^2 + y^2 + z^2 - 1) - \lambda_2(x + y).$$

At a local extremum $(x, y, z) \in M$, the following equations must be satisfied:

$$\begin{aligned}
\partial_x L(x, y, z, \lambda_1, \lambda_2) &= 3 - 2\lambda_1 x - \lambda_2 = 0\\
\partial_y L(x, y, z, \lambda_1, \lambda_2) &= -1 - 2\lambda_1 y - \lambda_2 = 0\\
\partial_z L(x, y, z, \lambda_1, \lambda_2) &= 2 - 2\lambda_1 z = 0\\
\partial_{\lambda_1} L(x, y, z, \lambda_1, \lambda_2) &= -(x^2 + y^2 + z^2 - 1) = 0\\
\partial_{\lambda_2} L(x, y, z, \lambda_1, \lambda_2) &= -(x + y) = 0
\end{aligned}$$

From $-x - y = 0$ follows $y = -x$, hence $z^2 = 1 - 2x^2 \implies z = \pm\sqrt{1 - 2x^2}$. From $2 - 2\lambda_1 z = 0$ follows $\lambda_1 = \frac{1}{z} = \frac{1}{\pm\sqrt{1-2x^2}}$. Moreover, we have

$$3 - 2\lambda_1 x = -1 + 2\lambda_1 x \implies 1 = \lambda_1 x.$$

So, $z = x = -y$ and hence $x^2 = 1 - 2x^2 \iff 3x^2 = 1 \iff x = \pm\frac{1}{\sqrt{3}}$. By substitution into f, we get

$$f\left(\tfrac{1}{\sqrt{3}}(1, -1, 1)\right) = \tfrac{6}{\sqrt{3}} = 2\sqrt{3}, \quad f\left(-\tfrac{1}{\sqrt{3}}(1, -1, 1)\right) = -\tfrac{6}{\sqrt{3}} = -2\sqrt{3}.$$

Thus, we have found all extrema of f on M, and f attains a maximum at $\frac{1}{\sqrt{3}}(1, -1, 1)$ and a minimum at $-\frac{1}{\sqrt{3}}(1, -1, 1)$.

Solution of 6.11: Firstly, if $a = b = 0$, f is the zero function. All points are critical and global maxima and minima simultaneously.

For critical points $(x, y) \in \mathbb{R}^2$, we compute the gradient

$$\nabla f = \begin{bmatrix} 2x(a - ax^2 - by^2)e^{-x^2-y^2} \\ 2y(b - ax^2 - by^2)e^{-x^2-y^2} \end{bmatrix} = 0.$$

To classify the critical points, we compute the Hessian matrix $H(x, y)$:

$$\begin{bmatrix} 2(a - 5ax^2 - by^2 + 2ax^4 + 2bx^2y^2) & 4xy(ax^2 - a + by^2 - b) \\ 4xy(ax^2 - a + by^2 - b) & 2(b - 5by^2 - ax^2 + 2by^4 + 2ax^2y^2) \end{bmatrix} e^{-x^2-y^2}.$$

- If $a = 0, b \neq 0$, then
$$\nabla f = \begin{bmatrix} -2xby^2e^{-x^2-y^2} \\ 2y(b - by^2)e^{-x^2-y^2} \end{bmatrix} = 0.$$
  From the second coordinate, we deduce $y = 0$ or $y = -1$ or $y = 1$. If $y = 0$, x can be arbitrary satisfying the condition of the first coordinate, otherwise $x = 0$.
  So, the critical points are $(x, 0)$ with $x \in \mathbb{R}$ arbitrary, as well as $(0, -1)$ and $(0, 1)$.
  For $a = 0$:
$$H(x, 0) = \begin{bmatrix} 0 & 0 \\ 0 & 2b \end{bmatrix} e^{-x^2}.$$

The matrix is singular, we also observe this because the critical points lie on the entire line $\mathbb{R} \times \{0\}$.

With $H(0,-1) = H(0,1) = \begin{bmatrix} -2b & 0 \\ 0 & -4b \end{bmatrix} e^{-1}$, $(0,-1)$ and $(0,1)$ are both minima (if $b < 0$) and maxima (if $b > 0$).

- The case $a \neq 0, b = 0$ follows symmetrically to (degenerate) critical points $(0, y)$ with $y \in \mathbb{R}$ arbitrary, as well as minima $(-1, 0)$ and $(1, 0)$ if $a < 0$ and otherwise maxima.

- The case $a \neq 0, b \neq 0$: If $x = 0$, similarly $y = 0$ or $y = 1$ or $y = -1$. Analogously for $y = 0$. We obtain critical points $(0,0), (0,1), (0,-1), (1,0), (-1,0)$. If $x \neq 0$ and $y \neq 0$, then $a - ax^2 - by^2 = 0 = b - ax^2 - by^2$. Subtracting these equations, we find $a = b$. We handle these extra points separately (see below). We compute

$$H(0,0) = \begin{bmatrix} 2a & 0 \\ 0 & 2b \end{bmatrix}, H(0,\pm 1) = \begin{bmatrix} 2(a-b) & 0 \\ 0 & -4b \end{bmatrix} e^{-1}, H(\pm 1, 0) = \begin{bmatrix} -4a & 0 \\ 0 & 2(b-a) \end{bmatrix} e^{-1}$$

  - If $a > b > 0$, we have a minimum at $(0,0)$, two saddle points at $(0, \pm 1)$, and two maxima at $(\pm 1, 0)$.
  - If $b > a > 0$, we have a minimum at $(0,0)$, two maxima at $(0, \pm 1)$, and two saddle points at $(\pm 1, 0)$.
  - If $a > 0 > b$, we have a saddle point at $(0,0)$, two minima at $(0, \pm 1)$, and two maxima at $(\pm 1, 0)$.
  - If $b > 0 > a$, we have a saddle point at $(0,0)$, two maxima at $(0, \pm 1)$, and two minima at $(\pm 1, 0)$.
  - If $0 > a > b$, we have a maximum at $(0,0)$, two minima at $(0, \pm 1)$, and two saddle points at $(\pm 1, 0)$.
  - If $0 > b > a$, we have a maximum at $(0,0)$, two saddle points at $(0, \pm 1)$, and two minima at $(\pm 1, 0)$.

- In the case $a = b$, we find additional critical points: Since $a \neq 0$, $1 - x^2 - y^2 = 0$. Thus, (x, y) lies on the circle with radius 1.

  For points with $x^2 + y^2 = 1$ and with $a = b$, the Hessian matrix is

$$H(x,y) = \begin{bmatrix} -4ax^2 & -4axy \\ -4axy & -4ay^2 \end{bmatrix} e^{-1}.$$

  It has determinant 0, being singular, as expected, since the entire circle $x^2 + y^2 = 1$ is a critical set.