Problem 1.

Consider the unconstrained optimisation problem

$$\min_{x,y} f(x, y),$$

where $f\colon \mathbb{R}^2 \to \mathbb{R}$ is defined as

$$f(x, y) = x^3 + 7x^2 - 2xy + y^2 + 3x + 6y.$$

a) Find all local and global minimisers of f.
(10 points)

b) Determine whether the function f is convex.
(5 points)

c) Perform one step of Newton's method with backtracking Armijo line search starting from the point $(x, y) = (0, 0)$. Use the parameters $c_1 = 0.1$ (sufficient decrease parameter) and $\rho = 0.5$ (contraction factor).
(10 points)

Solution.

a) We start by computing the gradient of f: We have

$$\nabla f(x, y) = \begin{pmatrix} 3x^2 + 14x - 2y + 3 \\ -2x + 2y + 6 \end{pmatrix}.$$

The first order optimality condition $\nabla f(x, y)$ thus becomes

$$\begin{aligned} 3x^2 + 14x - 2y + 3 &= 0, \\ -2x + 2y + 6 &= 0. \end{aligned}$$

From the second equation we get

$$-2y = -2x + 6 \qquad \text{or} \qquad y = x - 3.$$

Inserting this in the first equation, we obtain

$$3x^2 + 12x + 9 = 0,$$

which has the two solutions $x = -1$ and $x = -3$. The corresponding values of y are $y = -4$ and $y = -7$. Thus we have the critical points

$$(x^*, y^*) = (-1, -4) \qquad \text{and} \qquad (x^*, y^*) = (-3, -7).$$

The Hessian of f is

$$H_f(x, y) = \begin{pmatrix} 6x + 14 & -2 \\ -2 & 2 \end{pmatrix}.$$

For $x = -1$, this is positive definite (trace and determinant are both positive), while it is indefinite for $x = -3$ (trace and determinant are both negative). Thus (x^*, y^*) is the unique strict local minimum of f.

Since $\lim_{x \to -\infty} f(x, 0) = -\infty$ (as the third order term dominates the lower order terms asymptotically), the function f is unbounded below and thus no global minimum exists.

b) In the previous part, we have seen that the Hessian of f can be indefinite. Therefore f is not convex.

c) At the point $(0, 0)$ we have

$$\nabla f(0, 0) = \begin{pmatrix} 3 \\ 6 \end{pmatrix}$$

and

$$H_f(0, 0) = \begin{pmatrix} 14 & -2 \\ -2 & 2 \end{pmatrix}.$$

The search direction for the line search is therefore

$$p = -H_f(0, 0)^{-1} \nabla f(0, 0) = \begin{pmatrix} 14 & -2 \\ -2 & 2 \end{pmatrix}^{-1} \begin{pmatrix} 3 \\ 6 \end{pmatrix} = \frac{1}{24} \begin{pmatrix} 2 & 2 \\ 2 & 14 \end{pmatrix} \begin{pmatrix} 3 \\ 6 \end{pmatrix} = -\frac{1}{4} \begin{pmatrix} 3 \\ 15 \end{pmatrix}.$$

Moreover,

$$\langle p, \nabla f(0, 0) \rangle = -\frac{99}{4}$$

and $f(0, 0) = 0$. The Armijo condition with $\alpha = 1$ and $c_1 = 0.1$ thus requires that

$$f(-3/4, -15/4) \leq -\frac{99}{40}.$$

We compute

$$\begin{aligned} f(-3/4, -15/4) &= -\frac{27}{64} + \frac{63}{16} - \frac{90}{16} + \frac{225}{16} - \frac{9}{4} - \frac{90}{4} \\ &= \frac{-27 + 252 - 360 + 900 - 144 - 1440}{64} = \frac{1152 - 1971}{64} = -\frac{819}{64}. \end{aligned}$$

Since obviously

$$f(-3/4, -15/4) = -\frac{819}{64} < -\frac{99}{40} = f(0,0) - c_1\langle p, \nabla f(0,0)\rangle,$$

Armijo's condition is satisfied and the step length $\alpha = 1$ is accepted.

Problem 2.

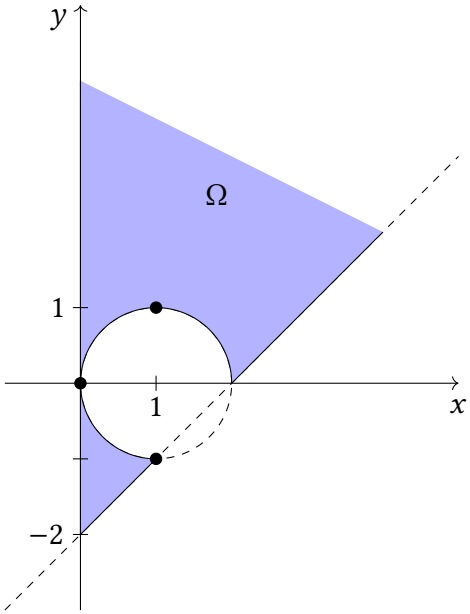
Consider the optimisation problem

$$\min_{x,y} (x-1)^2 + (y+2)^2$$

subject to the constraint $(x, y) \in \Omega$, where $\Omega \subset \mathbb{R}^2$ is given by the inequalities

$$x \geq 0, \qquad y \geq x - 2, \qquad \text{and} \qquad (x-1)^2 + y^2 \geq 1.$$

Here is a sketch of the set Ω (it continues infinitely to the top right):



a) Determine for each point in Ω whether the LICQ holds.
(5 points)

b) Determine the tangent cone $T_\Omega(x, y)$ and the cone of linearised feasible directions $\mathcal{F}(x, y)$ to the set Ω in the points $(x, y) = (1, -1)$, $(x, y) = (1, 1)$, and $(x, y) = (0, 0)$. (The points are marked in the sketch.)
(10 points)

c) Find all KKT points and all local minimisers for this optimisation problem. Use the second order necessary / sufficient conditions whenever applicable.
(15 points)

Solution.

As a preparation, we define

$$f(x, y) = (x-1)^2 + (y+2)^2$$

and

$$c_1(x, y) = x, \qquad c_2(x, y) = y - x + 2, \qquad c_3(x, y) = (x-1)^2 + y^2 - 1,$$

and we compute

$$\nabla f(x, y) = \begin{pmatrix} 2(x-1) \\ 2(y+2) \end{pmatrix}, \quad \nabla c_1(x, y) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \nabla c_2(x, y) = \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \quad \nabla c_3(x, y) = \begin{pmatrix} 2(x-1) \\ 2y \end{pmatrix}.$$

a) We see immediately that $\nabla c_1(x, y) \neq 0$ and $\nabla c_2(x, y) \neq 0$ for all x, y. Moreover, the only point where $\nabla c_3(x, y) = 0$ is the point $(x, y) = (1, 0)$, which is not feasible. Thus LICQ holds whenever none or only a single constraint is active.

It remains to check the points where more than one constraint is active:

- At $(x, y) = (-2, 0)$, the constraints c_1 and c_2 are active. Since they are obviously linearly independent, LICQ holds.
- At $(x, y) = (1, -1)$, the constraints c_2 and c_3 are active. We have $\nabla c_3(1, -1) = (0, -2)^T$ and $\nabla c_2(1, -1) = (-1, 1)^T$, which are obviously linearly independent. Thus, again, LICQ holds.
- At $(x, y) = (2, 0)$, the constraints c_2 and c_3 are active. We have $\nabla c_3(0, 2) = (-2, 0)^T$ and $\nabla c_2(1, -1) = (-1, 1)^T$, which are obviously linearly independent. Again, LICQ holds.
- Finally, at $(x, y) = (0, 0)$, the constraints c_1 and c_3 are active and $\nabla c_1(x, y) = (1, 0)^T$ and $\nabla c_3(x, y) = (-2, 0)^T$. These are linearly independent, and thus LICQ fails.

To summarise, LICQ holds at every point except $(x, y) = (0, 0)$.

b) At $(1, -1)$, the constraints c_2 and c_3 are active and LICQ holds. Moreover $\nabla c_2(1, -1) = (1, -1)^T$ and $\nabla c_3(1, -1) = (0, -2)^T$. Thus

$$T_\Omega(1, -1) = \mathcal{F}(1, -1) = \left\{(p, q) \in \mathbb{R}^2 : p - q \geq 0 \text{ and } -2q \geq 0\right\}.$$

At $(1, 1)$, only the constraint c_3 is active, LICQ holds, and $\nabla c_3(1, 1) = (0, 2)^T$. Thus

$$T_\Omega(1, 1) = \mathcal{F}(1, 1) = \left\{(p, q) \in \mathbb{R}^2 : 2q \geq 0\right\}.$$

Finally, at $(0,0)$ the constraints c_1 and c_3 are active and $\nabla c_1(0,0) = (1,0)^T$ and $\nabla c_2(0,0) = (-2,0)^T$. However, LICQ fails and thus we can at first only infer that

$$\mathcal{F}(0,0) = \{(p,q) : p \geq 0 \text{ and } -2p \geq 0\} = \mathbb{R} \cdot (0,1).$$

Since, locally around $(0,0)$, the y-axis is feasible, we can conclude that

$$\mathbb{R} \cdot (0,1) \subset T_\Omega(0,0).$$

From the inclusion $T_\Omega(0,0) \subset \mathcal{F}(0,0)$ we thus obtain that

$$\mathcal{F}(0,0) = \mathbb{R} \cdot (0,1) \subset T_\Omega(0,0) \subset \mathcal{F}(0,0),$$

which in fact shows that

$$T_\Omega(0,0) = \mathcal{F}(0,0) = \mathbb{R} \cdot (0,1).$$

c) The KKT conditions for this problem are

$$\begin{aligned}
2(x-1) &= \lambda - \mu + 2\nu(x-1),\\
2(y+2) &= \mu + 2\nu y,\\
x &\geq 0,\\
y &\geq x-2,\\
(x-1)^2 + y^2 &\geq 1,\\
\lambda &\geq 0,\\
\mu &\geq 0,\\
\nu &\geq 0,\\
\lambda x &= 0,\\
\mu(y-x+2) &\geq 0,\\
\nu\big((x-1)^2 + y^2 - 1\big) &\geq 0.
\end{aligned}$$

For finding the KKT points of the problem, we consider all possible active sets:

- If no constraints are active, we have that $\lambda = \mu = \nu = 0$ and thus

  $$2(x-1) = 0 \qquad \text{and} \qquad 2(y+2) = 0.$$

  The solution $(x,y) = (1,-2)$ of this system is infeasible, and thus we obtain no KKT point.
- If only c_1 is active, we have that $\mu = \nu = 0$ and $x = 0$ and thus

  $$2(0-1) = \lambda \qquad \text{and} \qquad 2(y+2) = 0.$$

  The first equation implies that $\lambda = -2 < 0$, and thus we have no KKT point.

- If only c_2 is active, we have that $\lambda = \nu = 0$ and $y = x - 2$. Thus we obtain the equations

$$2(x-1) = -\mu \qquad \text{and} \qquad 2(y+2) = \mu.$$

Summing these two equations and using that $y = x - 2$, we obtain that

$$0 = 2(x-1) + 2x = 4x - 2,$$

and thus

$$x = \frac{1}{2}, \qquad y = -\frac{3}{2}, \qquad \mu = 1.$$

This point is feasible and thus a KKT point.

- If only c_3 is active, we have that $\lambda = \mu = 0$ and $(x-1)^2 + y^2 = 1$. We obtain the equations

$$2(x-1) = 2\nu(x-1) \qquad \text{and} \qquad 2(y+2) = 2\nu y.$$

The first equation implies that either $x = 1$ or $\nu = 1$ (or both). In the case $\nu = 1$, the second equation implies that $2(y+2) = y$, which is clearly impossible. Thus $x = 1$, which in turn implies that $y = \pm 1$. We therefore obtain the two candidates $(x, y) = (1, \pm 1)$, both of which are feasible. At $(x, y) = (1, -1)$, the constraint c_2 is feasible as well, and thus we will discuss this point later. For $y = 1$, the second equation reads $6 = 2\nu$, and thus $\nu = 3$. We therefore have the KKT point $(x, y) = (1, 1)$ with Lagrange multiplier $\nu = 3$.

- If c_1 and c_2 are active, we have that $(x, y) = (0, -2)$ and $\nu = 0$, and we obtain the equations

$$-2 = \lambda - \mu \qquad \text{and} \qquad 0 = \mu.$$

The only solution of this system is $\mu = 0$ and $\lambda = -2 < 0$, and thus we have no KKT point.

- If c_1 and c_3 are active, we have that $(x, y) = (0, 0)$ and $\mu = 0$, and we obtain the equations

$$-2 = \lambda - 2\nu \qquad \text{and} \qquad 4 = 0.$$

Obviously, these are unsolvable.

- If c_2 and c_3 are active, we have the two points $(x, y) = (1, -1)$ and $(x, y) = (2, 0)$ and $\lambda = 0$. For $(x, y) = (1, -1)$ we obtain the equations

$$0 = -\mu \qquad \text{and} \qquad 2 = \mu - 2\nu.$$

This implies that $\mu = 0$ and $\nu = -1 < 0$. Thus the point $(1, -1)$ is no KKT point.
For $(x, y) = (2, 0)$ we obtain the equations

$$2 = -\mu + 2\nu \qquad \text{and} \qquad 4 = \mu.$$

Thus $\mu = 4$ and $\nu = 3$ and $(1, -1)$ is a KKT point.

Thus we have the KKT points $(1/2, -3/2)$ with Lagrange multipliers $(\lambda, \mu, \nu) = (0, 1, 0)$, the point $(1, 1)$ with Lagrange multipliers $(\lambda, \mu, \nu) = (0, 0, 3)$, and the point $(1, -1)$ with Lagrange multipliers $(\lambda, \mu, \nu) = (0, 4, 3)$.

The Hessian of the Lagrangian is

$$H_{\mathcal{L}}(x, y; \lambda, \mu, \nu) = \begin{pmatrix} 2 - 2\nu & 0 \\ 0 & 2 - 2\nu \end{pmatrix}.$$

At the KKT point $(1/2, -3/2)$, the Hessian of the Lagrangian is positive definite (as $\nu = 0$), and we have a strict local minimum. At the KKT point $(1, -1)$, the critical cone is trivial as we have two linearly independent active constraints with non-zero Lagrange parameters. Thus the second order sufficient conditions are trivially satisfied. At the KKT point $(1, 1)$, the Hessian of the Lagrangian is negative definite and the critical cone is non-trivial.[1] Thus this point is no local minimiser.

Finally, we consider the point $(x, y) = (0, 0)$. Although LICQ does not hold at this point, we still have that $T_\Omega(0, 0) = \mathcal{F}(0, 0)$, and thus the KKT conditions are necessary optimality conditions also at that point. Since $(0, 0)$ is no KKT point, we can conclude that it is no local minimiser.[2]

To summarise, we have the following KKT points and local minimisers:

- The KKT point and local minimiser $(1/2, -3/2)$.
- The KKT point and local minimiser $(1, -1)$.
- The KKT point $(1, 1)$, which is no local minimiser.

Problem 3.

Assume that $f : \mathbb{R}^d \to \mathbb{R}$ is continuously differentiable and bounded below. We consider the gradient descent method

$$x_{k+1} = x_k - \alpha_k \nabla f(x_k)$$

for the minimisation of f, where the step lengths $\alpha_k > 0$ are chosen such that the Armijo condition with fixed parameter $c_1 > 0$ is satisfied in each step. Assume moreover that there exists $\varepsilon > 0$ such that

$$\alpha_k \geq \varepsilon > 0 \qquad \text{for all } k.$$

Show that

$$\lim_{k \to \infty} \nabla f(x_k) = 0.$$

[1] Specifically, the critical cone equals $C(1, 1) = \{(p, 0) : p \in \mathbb{R}\}$.

[2] Alternatively, one can of course also argue that $f(0, -1/k) \leq f(0, 0)$ for all $k \in \mathbb{N}$.

(10 points)

Solution.

By assumption, the function f is bounded below, and thus there exists $C \in \mathbb{R}$ such that $f(x) \geq C$ for all $x \in \mathbb{R}^d$. By Armijo's condition, we obtain that

$$f(x_{k+1}) - f(x_k) \leq -c_1 \alpha_k \|\nabla f(x_k)\|^2.$$

Alternative 1: Thus

$$\sum_{k=0}^{\infty} c_1 \alpha_k \|\nabla f(x_k)\|^2 \leq \sum_{k=0}^{\infty} \big(f(x_k) - f(x_{k+1})\big) \leq f(x_0) - C.$$

Since $\alpha_k \geq \varepsilon > 0$ for all k, it follows that

$$\sum_{k=0}^{\infty} \|\nabla f(x_k)\|^2 \leq \frac{f(x_0) - C}{c_1 \varepsilon} < \infty.$$

This implies in particular that $\lim_{k\to\infty} \nabla f(x_k) = 0$.

Alternative 2: In particular, it follows that $f(x_{k+1}) < f(x_k)$ as long as $\nabla f(x_k) \neq 0$. Since f is bounded below, the decreasing sequence $\{f(x_k)\}_{k\in\mathbb{N}}$ converges, and thus the differences $f(x_k) - f(x_{k+1})$ tend to 0. Thus

$$0 = \lim_{k\to\infty} \big(f(x_k) - f(x_{k+1})\big) \geq \limsup_{k\to\infty} c_1 \alpha_k \|\nabla f(x_k)\|^2 \geq c_1 \varepsilon \limsup_{k\to\infty} \|\nabla f(x_k)\|^2,$$

which proves that $\lim_{k\to\infty} \nabla f(x_k) = 0$.

Problem 4.

Define the functions $f_1, f_2 \colon \mathbb{R}^2 \to \mathbb{R}$,

$$f_1(x, y) = (x-1)^2 + (y-2)^2, \qquad f_2(x, y) = x^2 + (y+2)^2.$$

Find all Pareto-optimal solutions of the multi-criteria optimisation problem

$$\min_{(x,y)\in\mathbb{R}^2} \big(f_1(x, y), f_2(x, y)\big).$$

(10 points)

Solution.

The functions f_1 and f_2 are strictly convex (and coercive), and thus the Pareto-optimal solutions are all the minimisers of the weighted sum problems

$$\min_{(x,y)\in\mathbb{R}^2} \lambda f_1(x, y) + (1 - \lambda) f_2(x, y)$$

for $0 \leq \lambda \leq 1$. Since these are strictly convex and differentiable problems, the unique solutions are found by setting the gradient to 0. This results in the equations

$$\begin{aligned} 2\lambda(x - 1) + 2(1 - \lambda)x &= 0, \\ 2\lambda(y - 2) + 2(1 - \lambda)(y + 2) &= 0. \end{aligned}$$

A brief computation shows that the solution of this system is the point

$$(x_\lambda, y_\lambda) = \big(\lambda, 4\lambda - 2\big).$$

Problem 5.

Assume that $f\colon \mathbb{R}^d \to \mathbb{R}$ is convex, continuous, and bounded below, and that $A \in \mathbb{R}^{m\times d}$ with $m < d$ has full rank. We define the function $g\colon \mathbb{R}^m \to \mathbb{R}$,

$$g(y) := \inf_{\substack{x\in\mathbb{R}^d \\ Ax=y}} f(x).$$

Show that g is convex!

(10 points)

Solution.

Alternative 1:

For fixed $y \in \mathbb{R}^m$ consider the optimisation problem

$$\min_x f(x) \qquad \text{s.t. } Ax = y.$$

Since A has full rank m, it follows that there exists some $x \in \mathbb{R}^d$ such that $Ax = y$. As a consequence, strong duality holds and thus the value of this optimisation problem is the same as the value of the dual problem. The Lagrangian of this problem is

$$\mathcal{L}(x, \lambda) = f(x) - \langle \lambda, Ax - y \rangle,$$

and thus the dual objective function is

$$q(\lambda) = \inf_x \big(f(x) - \langle \lambda, Ax - y\rangle\big) = \langle \lambda, y\rangle - \sup_x \big(\langle A^T\lambda, x\rangle - f(x)\big) = \langle \lambda, y\rangle - f^*(A^T\lambda).$$

As a consequence,

$$g(y) = \inf_{\substack{x\in\mathbb{R}^d \\ Ax=y}} f(x) = \sup_{\lambda\in\mathbb{R}^m} \big(\langle \lambda, y\rangle - f^*(A^T\lambda)\big).$$

Now we note that, for fixed $\lambda \in \mathbb{R}^m$, the function

$$y \mapsto \langle \lambda, y\rangle - f^*(A^T\lambda)$$

is affine (and in particular convex). Thus g is the supremum of convex functions and thus convex itself.

Alternative 2:

Since A has full rank m, it follows that the equation $Ax = y$ has for each $y \in \mathbb{R}^m$ a solution $x \in \mathbb{R}^d$. Thus g is well-defined. Moreover, because f is bounded below, it follows that $g(y) < -\infty$ for all $y \in \mathbb{R}^m$.

Now let $y, z \in \mathbb{R}^m$, and let $0 < \lambda < 1$. Since f is bounded below, there exist for every $\varepsilon > 0$ some x_y^ε, $x_z^\varepsilon \in \mathbb{R}^m$ such that

$$Ax_y^\varepsilon = y \qquad \text{and} \qquad Ax_z^\varepsilon = z,$$

and

$$f(x_y^\varepsilon) \le g(y) + \varepsilon \qquad \text{and} \qquad f(x_z^\varepsilon) \le g(z) + \varepsilon.$$

Now the convexity of f implies that

$$f(\lambda x_y^\varepsilon + (1-\lambda)x_z^\varepsilon) \le \lambda f(x_y^\varepsilon) + (1-\lambda) f(x_z^\varepsilon) \le \lambda g(y) + (1-\lambda) g(z) + \varepsilon.$$

By construction, we have

$$A(\lambda x_y^\varepsilon + (1-\lambda)x_z^\varepsilon) = \lambda y + (1-\lambda)z.$$

Thus

$$g(\lambda y + (1-\lambda)z) = \inf_{\substack{x\in\mathbb{R}^d \\ Ax=\lambda y+(1-\lambda)z}} f(x) \le f(\lambda x_y^\varepsilon + (1-\lambda)x_z^\varepsilon) \le \lambda g(y) + (1-\lambda) g(z) + \varepsilon.$$

Since $\varepsilon > 0$ was arbitrary, this implies that

$$g(\lambda y + (1-\lambda)z) \le \lambda g(y) + (1-\lambda) g(z).$$

Since $y, z \in \mathbb{R}^m$ and $0 < \lambda < 1$ were arbitrary, this proves the convexity of g.

Alternative 3:

Let $y, z \in \mathbb{R}^m$, and let $0 < \lambda < 1$. Then we can write

$$g\big(\lambda y + (1-\lambda)z\big) = \inf_{\substack{x\in\mathbb{R}^d \\ Ax=\lambda y+(1-\lambda)z}} f(x).$$

If $x_1 \in \mathbb{R}^d$ satisfies $Ax_1 = y$ and $x_2 \in \mathbb{R}^d$ satisfies $Ax_2 = z$, then $x := \lambda x_1 + (1-\lambda)x_2$ satisfies $Ax = \lambda y + (1-\lambda)z$. As a consequence, we have the inclusion

$$\big\{x \in \mathbb{R}^d : x = \lambda x_1 + (1-\lambda)x_2,\ Ax_1 = y,\ Ax_2 = z\big\} \subseteq \big\{x \in \mathbb{R}^d : Ax = \lambda y + (1-\lambda)z\big\}$$

Therefore,

$$\inf_{\substack{x\in\mathbb{R}^d \\ Ax=\lambda y+(1-\lambda)z}} f(x) \le \inf_{\substack{x\in\mathbb{R}^d \\ x=\lambda x_1+(1-\lambda)x_2 \\ Ax_1=y \\ Ax_2=z}} f(x) = \inf_{\substack{x_1,\,x_2\in\mathbb{R}^d \\ Ax_1=y \\ Ax_2=z}} f(\lambda x_1 + (1-\lambda)x_2),$$

Now we can use the convexity of f to estimate

$$\begin{aligned} g(\lambda y + (1-\lambda)z) &= \inf_{\substack{x_1,\,x_2\in\mathbb{R}^d \\ Ax_1=y \\ Ax_2=z}} f(\lambda x_1 + (1-\lambda)x_2) \\ &\le \inf_{\substack{x_1,\,x_2\in\mathbb{R}^d \\ Ax_1=y \\ Ax_2=z}} \big(\lambda f(x_1) + (1-\lambda) f(x_2)\big) \\ &= \lambda \inf_{\substack{x_1\in\mathbb{R}^d \\ Ax_1=y}} f(x_1) + (1-\lambda) \inf_{\substack{x_2\in\mathbb{R}^d \\ Ax_2=z}} f(x_2) = \lambda g(y) + (1-\lambda) g(z), \end{aligned}$$

which shows that g is convex. Here the equality between the second and third line holds, because there is no coupling between the variables x_1 and x_2 in either the constraint or the objective function of the problem in the second line.

Problem 6.

Let $a \in \mathbb{R}^d$ be fixed and consider the constrained optimisation problem

$$\min_{x\in\mathbb{R}^d} \sum_{i=1}^{d} \exp(x_i) \qquad \text{s.t.} \begin{cases} x \le 0, \\ \langle a, x\rangle \le 1. \end{cases} \tag{1}$$

a) Formulate the KKT conditions for (1). Moreover, show that (1) has no solution if there exists any $1 \le i \le d$ such that $a_i \ge 0$.
(5 points)

b) Compute the Lagrangian dual of the optimisation problem (1). Does the dual problem have a solution?
(10 points)

Solution.

a) The Lagrangian of the problem is $\mathcal{L}\colon \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R} \to \mathbb{R}$,

$$\mathcal{L}(x, \lambda, \mu) = \sum_{i=1}^{d} \exp(x_i) + \sum_{i=1}^{d} \lambda_i x_i - \mu\big(1 - \langle a, x\rangle\big),$$

which we can also write as

$$\mathcal{L}(x, \lambda, \mu) = -\mu + \sum_{i=1}^{d} \big(\exp(x_i) + \lambda_i x_i + \mu a_i x_i. \tag{2}$$

The KKT conditions for (1) are therefore

$$\begin{aligned}
\exp(x_i) &= -\lambda_i - \mu a_i, && i = 1, \dots, d,\\
x_i &\le 0, && i = 1, \dots, d,\\
\langle a, x\rangle &\le 1,\\
\lambda_i &\ge 0, && i = 1, \dots, d,\\
\mu &\ge 0,\\
\lambda_i x_i &= 0, && i = 1, \dots, d,\\
\mu\big(1 - \langle a, x\rangle\big) &= 0.
\end{aligned}$$

Moreover, as the constraints are linear, these are necessary optimality conditions.

Assume now that x^* is a solution of (1), and let λ^*, μ^* be corresponding Lagrange multipliers. Since $\exp(t) > 0$ for all $t \in \mathbb{R}$, we obtain in particular that

$$-\lambda_i^* - \mu^* a_i > 0 \qquad \text{for all } i = 1, \dots, d.$$

Since $\lambda_i^* \ge 0$ and $\mu^* \ge 0$, this implies that

$$\mu^* a_i < -\lambda_i^* \le 0,$$

which is only possible if $a_i < 0$ for all $i = 1, \dots, d$. Thus, if there exists i such that $a_i \ge 0$, then there exists no KKT point and thus no solution of the problem.

b) Recalling the Lagrangian of the problem (see (2)), we see that the dual objective function is

$$q(\lambda,\mu) = -\mu + \inf_{x\in\mathbb{R}^d}\Big(\sum_{i=1}^{d}\big(\exp(x_i) + \lambda_i x_i + \mu a_i x_i\big)\Big) = -\mu + \sum_{i=1}^{d}\inf_{x_i\in\mathbb{R}}\big(\exp(x_i) + \lambda_i x_i + \mu a_i x_i\big).$$

Denote now

$$f(t) = \exp(t).$$

Then

$$\inf_{x_i\in\mathbb{R}}\big(\exp(x_i) + \lambda_i x_i + \mu a_i x_i\big) = -\sup_{x_i\in\mathbb{R}}\big(-(\lambda_i+\mu a_i)x_i - \exp(x_i)\big) = -f^*(-\lambda_i - \mu a_i).$$

Moreover (see the attached formula sheet)

$$f^*(s) = \begin{cases} -s\log s - s & \text{if } s < 0,\\ 0 & \text{if } s = 0,\\ +\infty & \text{if } s > 0. \end{cases} \tag{3}$$

Thus the dual problem is

$$\max_{\substack{\lambda\in\mathbb{R}^d\\ \mu\in\mathbb{R}}} -\mu - \sum_{i=1}^{d} f^*(-\lambda_i - \mu a_i) \qquad \text{s.t. } \lambda \geq 0 \text{ and } \mu \geq 0,$$

where f^* is as defined in (3).

The primal objective function $x \mapsto \sum_i \exp(x_i)$ is obviously bounded below (by 0). Moreover, we only have linear equality and inequality constraints, and the feasible set is non-empty (e.g. the vector 0 is feasible). Thus Slater's constraint qualification holds. As a consequence, strong duality applies and the dual problem admits a solution.