Advanced Optimisation

Information for candidates:

- A function $f : \mathbb{R}^n \to \mathbb{R}$ with $\textbf{dom}\, f = \mathbb{R}^n_{++}$ defined as

  $$f(x) = c x_1^{a_1} x_2^{a_2} \cdots x_n^{a_n},$$

  with $c > 0$ and $a_i \in \mathbb{R}$, is called a monomial. A sum of monomials of the form

  $$f(x) = \sum_{k=1}^{K} c_k x_1^{a_{1k}} x_2^{a_{2k}} \cdots x_n^{a_{nk}},$$

  with $c_k > 0$, is called a posynomial.

- A geometric programme (GP) in posynomial form is described by

  $$\begin{array}{ll} \min & f_0(x) \\ s.t. & f_i(x) \le 1, \quad i = 1, \dots, m, \\ & h_i(x) = 1, \quad i = 1, \dots, p, \end{array}$$

  where f_0, ..., f_m are posynomials and h_1, ..., h_p are monomials. A GP in posynomial form can be transformed into convex form by considering the change of variables $y_i = \log x_i$, so $x_i = e^{y_i}$, and then taking the logarithm of all the functions.

1. a) Consider n transmitters with nonnegative powers p_1, ..., p_n, transmitting to n receivers. Let G_{ij} be the nonnegative gain from transmitter j to receiver i. The signal power at receiver i is then $S_i = G_{ii} p_i$, while the interference power at receiver i is $I_i = \sum_{k \neq i} G_{ik} p_k$. The signal to interference plus noise ratio, denoted SINR, at receiver i is given by the ratio of signal power over the sum of the interference power and the self noise $\sigma_i > 0$ of receiver i. The objective of the problem is to maximise the minimum SINR ratio, over all receivers. There are a number of constraints on the powers that must be satisfied in addition to the nonnegativity constraint already stated. First, there is a maximum allowable power $P_i^{\max}$ for each transmitter. Second, the transmitters are partitioned into groups, with each group sharing the same power supply, so there is a total power constraint for each group of transmitter. This constraint is formulated as follows. We have disjoint subsets K_1, ..., K_m with $K_1 \cup \cdots \cup K_m = \{1, \ldots, n\}$. For each group K_j the total associated transmitter power cannot exceed P_j^{gp}. Finally, we have a limit $P_k^{rc} > 0$ on the total received power at each receiver.

i) Formulate the SINR maximisation problem. [6 marks]

ii) The formulated problem is quasiconvex. Note that if f is quasiconvex, then the problem $\max f(x)$ can be approximated by a family of convex feasibility problems based on a family of convex functions ϕ_t parametrised in t such that

(I) $f(x) \geq t \iff \phi_t(x) \leq 0$ for each x;

(II) $\phi_t(x)$ is nondecreasing in t.

Determine ϕ_t for the SINR maximisation problem and show that the properties (I) and (II) are satisfied.
Hint: the function ϕ_t is a vector with n components ϕ_t^i. [4 marks]

b) A heated fluid at temperature T (degrees above ambient temperature) flows in a pipe with fixed length and circular cross section of radius r. A layer of insulation, with thickness $w \ll r$, surrounds the pipe to reduce heat loss through the pipe walls. The design variables in this problem are T, r, and w. The rate of heat loss is proportional to Tr/w, so over a fixed period, the energy cost due to heat loss is given by $\alpha_1 Tr/w$. The cost of the pipe, which has a fixed wall thickness, is approximately proportional to the total material, i.e., it is given by $\alpha_2 r$. The cost of the insulation is also approximately proportional to the total insulation material, i.e., $\alpha_3 rw$. The total cost is the sum of these three costs. The rate of heat flow down the pipe is entirely due to the flow of the fluid, which has a fixed velocity, i.e., it is given by $\alpha_4 Tr^2$. The constants α_i, $i = 1, 2, 3, 4$, are all positive, as are the variables T, r, and w.

i) Formulate the problem of maximisation of the total heat flow down the pipe, subject to an upper limit $C_{\max}$ on total cost, and the constraints $T \in [T_{\min}, T_{\max}]$, $r \in [r_{\min}, r_{\max}]$ and $w \in [w_{\min}, w_{\max}]$. The condition $w \ll r$ is codified by the constraint that the thickness w is less than or equal to 10% of the radius r. Express this problem as a minimisation geometric program in posynomial form. [5 marks]

ii) Express this problem as a geometric program in convex form.

[5 marks]

2. This question covers two independent topics. Part a) is about duality and part b) is about multi-objective optimisation.

a) Consider the QCQP

$$\begin{aligned} \min \quad & x_1^2 + x_2^2 \\ s.t. \quad & (x_1-1)^2 + (x_2-1)^2 \leq 1, \\ & (x_1-1)^2 + (x_2+1)^2 \leq 1, \end{aligned}$$

with variable $x \in \mathbb{R}^2$.

i) Sketch the feasible set and some level sets of the objective function. Find, graphically, the optimal point x^* and optimal value p^*. [3 marks]

ii) Write the KKT conditions for the problem. Do there exist Lagrange multipliers λ_1 and λ_2 that prove that x^* is optimal? [2 marks]

iii) Obtain the Lagrange dual function $g(\lambda_1, \lambda_2)$.
Hint: divide the analysis in the three cases $\lambda_1 + \lambda_2 + 1 > 0$, $\lambda_1 + \lambda_2 + 1 = 0$ *and* $\lambda_1 + \lambda_2 + 1 < 0$. [5 marks]

iv) Solve the Lagrange dual problem. Does strong duality hold? Use this result to provide further insight on the KKT conditions obtained in part ii).
Hint: if g is symmetric, i.e. $g(\lambda_1, \lambda_2) = g(\lambda_2, \lambda_1)$, *then* $\lambda_1 = \lambda_2 = \lambda$.

[3 marks]

b) Figure 2.1 shows the optimal trade-off curve (solid black line) and the set of achievable values (the shaded area) for the bi-criterion optimisation problem

$$\min_{w.r.t. \mathbb{R}_+^2} \quad (||Ax-b||_2^2, ||x||_2^2),$$

for some $A \in \mathbb{R}^{100\times 10}$, $b \in \mathbb{R}^{100}$.

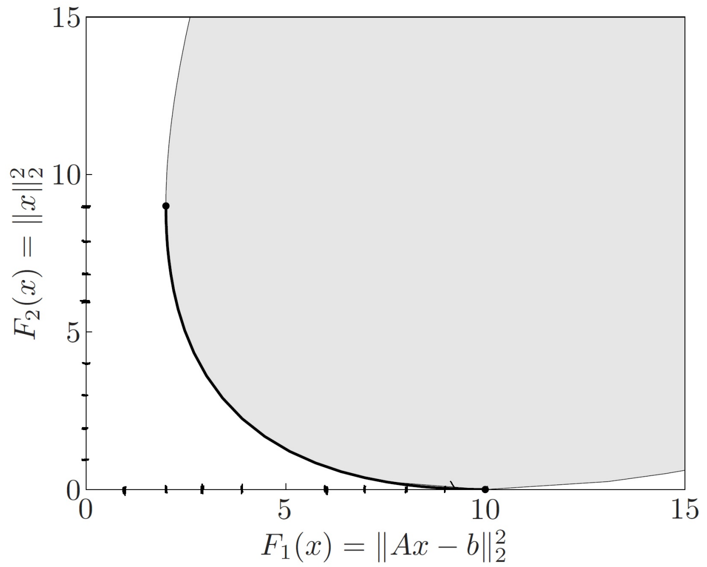

Figure 2.1 Trade-off curve and achievable values for part 2.b).

We denote by x_{ls} the solution of the least-squares problem

$$\min \quad ||Ax-b||_2^2.$$

Answer the following questions using information from the plot.
Hint: since reading a plot has some degree of uncertainty, you can say, e.g. "about 4" if the value is "close to 4" but you are unsure if it is exactly "4".

i) Estimate $||x_{ls}||_2$.

ii) Estimate $||Ax_{ls}-b||_2^2$.

iii) Estimate $||b||_2$.

iv) Estimate the optimal value of the problem

$$\begin{aligned} \min \quad & ||Ax-b||_2^2 \\ s.t. \quad & ||x||_2^2 = 1. \end{aligned}$$

v) Estimate the optimal value of the problem

$$\begin{aligned} \min \quad & ||Ax-b||_2^2 \\ s.t. \quad & ||x||_2^2 \leq 1. \end{aligned}$$

vi) Estimate the optimal value of the problem

$$\min \quad ||Ax-b||_2^2 + ||x||_2^2.$$

Justify your answer.

vii) What is the rank of A? Justify your answer.

[7 marks]

3. a) Prove whether the following functions are convex, concave or neither.

i) $f(x) = \log x$ in $\mathbb{R}_{++}$. [1 mark]

ii) $f(x) = x \log x$ in $\mathbb{R}_{++}$. [1 mark]

iii) $f(x) = \max\{x_1, x_2\}$ in $\mathbb{R}^2$. [1 mark]

iv) $f(x) = \log(e^{x_1} + e^{x_2})$ in $\mathbb{R}^2$. [2 marks]

b) Let $f : \mathbb{R}^n \to \mathbb{R}$. Recall that the conjugate of the function f is defined as

$$f^*(y) = \sup_{x \in \textbf{dom}\, f} \left(y^\top x - f(x) \right),$$

i) Let $f(x) = ax + b$, with $x \in \mathbb{R}$. Compute $f^*(y)$. [2 marks]

ii) Let $f(x) = -\log x$, with $x \in \mathbb{R}_{++}$. Compute $f^*(y)$. [2 marks]

c) Recall that a convex function $f : \mathbb{R} \to \mathbb{R}$ is self-concordant if $|f'''(x)| \leq 2 f''(x)^{3/2}$ for all $x \in \textbf{dom}\, f$.

i) Establish whether the function $f(x) = x \log x$ on $\mathbb{R}_{++}$ is self-concordant. [2 marks]

ii) Establish whether the function $g(x) = x \log x - \log x$ on $\mathbb{R}_{++}$ is self-concordant.
Hint: let $h(x) = \frac{|g'''(x)|}{g''(x)^{3/2}}$. *Compute* $h(0)$, *compute* $h'(x)$ *and use this information to study self-concordance for* $x > 0$.

[3 marks]

d) Consider the convex problem

$$\begin{array}{ll} \min & f_0(x) \\ s.t. & f_i(x) \leq 0, \quad i = 1, \ldots, m, \\ & Ax = b, \end{array}$$

and the barrier $\phi(x) = -\sum_{i=1}^{m} \log(-f_i(x))$.

i) Let $\frac{dx^*(t)}{dt}$ be the tangent to the central path at the point $x^*(t)$. Show that the explicit expression for the tangent is

$$\frac{dx^*}{dt} = -\left(t \nabla^2 f_0(x^*(t)) + \nabla^2 \phi(x^*(t)) \right)^{-1} \nabla f_0(x^*(t)).$$

[2 marks]

ii) Exploiting the tangent, show that $f_0(x^*(t))$ decreases as t increases.

[2 marks]

iii) In the standard barrier method, $x^*(\mu t)$ is computed using Newton's method, starting from the initial point $x^*(t)$. One alternative that has been proposed is to make an approximation or prediction $\hat{x}$ of $x^*(\mu t)$, and then start Newton's method for computing $x^*(\mu t)$ from $\hat{x}$. The idea is that this should reduce the number of Newton steps, since $\hat{x}$ is (presumably) a better initial point than $x^*(t)$. This method of centering is called a *predictor-corrector method*, since it first makes a prediction of what $x^*(\mu t)$ is, then corrects the prediction using Newton's method. The most widely used predictor is the first-order predictor, based on the tangent to the central path. This predictor is given by

$$\hat{x} = x^*(t) + \frac{dx^*(t)}{dt}(\mu t - t).$$

Let instead $x^n = x^*(t) + \Delta x_{nt}$ be the Newton update, namely

$$x^n = x^*(t) - (\mu - 1)t \left(\mu t \nabla^2 f_0(x^*(t)) + \nabla^2 \phi(x^*(t))\right)^{-1} \nabla f_0(x^*(t)).$$

Compute $\hat{x}$ and compare the formulas of $\hat{x}$ and x^n commenting on any difference or similarity. Then compute $\hat{x}$ and x^n for the special case in which $f_0 = c^T x$ and comment on any difference or similarity. Finally, describe what condition a generic f_0 must satisfy for $\hat{x}$ and x^n to be the same.

[2 marks]