ADVANCED OPTIMISATION

The average of this module in the year 2021-2022 was 70.2 (median 76.16). The average in the exam was 65.4 (median 70.0).

Information for candidates:

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

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

  where $c > 0$ 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

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

  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.

- Consider two functions $h : \mathbb{R}^k \to \mathbb{R}$ and $g : \mathbb{R}^n \to \mathbb{R}^k$. Let $f = h \circ g : \mathbb{R}^n \to \mathbb{R}$, be their composition defined by $f(x) = h(g(x)) = h(g_1(x), \dots, g_k(x))$, $\mathbf{dom}\ f = \{x \in \mathbf{dom}\ g : g(x) \in \mathbf{dom}\ h\}$.

  - f is convex if h is convex, $\tilde{h}$ is nondecreasing in each argument and g_i are convex;
  - f is convex if h is convex, $\tilde{h}$ is nonincreasing in each argument and g_i are concave;

  where $\tilde{h}$ indicates the extended-value extension of h.

1. a) Consider a box with width w, height h and depth d. The total later surface (of the four walls) must not exceed A_{wall}. The total floor surface (the bottom) must not exceed A_{floor}. The ratio between height and width must be in the set $[\alpha, \beta]$. The ratio between depth and width must be in the set $[\gamma, \delta]$. Our objective is to maximise the volume of the box subject to these constraints.

i) Formulate the considered problem in posynomial form.

[5 marks]

ii) Let $y_w = \log w$, $y_d = \log d$, $y_h = \log h$. Transform the problem formulated in part a.i) into a GP in convex form.
Hint: see "Information for candidates". [5 marks]

b) Consider the problem of trying to protect a parade route by placing a limited number of guards at strategic points. The parade route is discretized into m points. There are n possible guard locations with associated decision variable $x \in \{0,1\}^n$, where $x_i = 1$ if and only if a guard is placed at location i. Associated with guard location i is a coverage vector $a_i = \begin{bmatrix} a_{1i} & \cdots & a_{mi} \end{bmatrix}^\top \in \mathbb{R}^m$, which describes how well a guard placed at location i would 'cover' each point j, with $j = 1, \ldots, m$, in the parade route. We assume that guard coverage is additive, so that the vector describing the total coverage of every parade route point is given by $Ax \in \mathbb{R}^m$, where $A \in \mathbb{R}^{m \times n}$ has a_i as its i-th column. This implies that the total coverage of point j is $(Ax)_j$, where $(Ax)_j$ indicates the j-th element of Ax. The parade route is only as secure as its least well-covered point. Our goal is to place k guards to maximize the minimum coverage over the points in the route.

i) Formulate the problem as a convex maximisation problem using the function $\min_{j=1,\ldots,m}\{\cdot\}$ to form the objective function and relaxing the Boolean constraint. [4 marks]

ii) Reformulate the problem as a linear programme that has just a vector $t \in \mathbb{R}^m$ as objective function by adding a new constraint. [3 marks]

iii) Consider now a heuristic called "Iterated Weighted ℓ_1 Heuristic" which tries to recover the solution of the original Boolean problem. The heuristic consists in solving a sequence of convex problems in which the objective function is $t - w^\top x$ (instead of simply t as in b.ii)). The algorithm starts by solving the problem with $w^{(0)} = 0$ finding an optimal point $x^{(0)}$ to the relaxed problem formulated in b.ii). Let $\tilde{x}^{(k)}$ be the optimal solution at iteration k. Then the weight is updated as

$$w_i^{(k+1)} = \frac{1}{0.0001 + \tilde{x}_i^{(k)}} \qquad i = 1, \ldots, n,$$

and the problem is solved again for $k = 1, 2, \ldots, N$. Explain what this heuristic causes and how it differs from thresholding. [3 marks]

2. Consider the optimisation problem

$$\begin{aligned} \min \quad & x^2+2x+1 \\ s.t. \quad & x(x-4) \leq 0, \end{aligned}$$

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

a) Determine the feasible set, the optimal value and the optimal solution.
[4 marks]

b) Plot the objective function and the constraint function versus x (in the same plot). In the plot, clearly indicate the feasible set, the optimal point and the optimal value.
[4 marks]

c) Determine the Lagrange dual function g.
Hint: there are two cases to consider depending on the value of λ. [4 marks]

d) State the dual problem. Find the dual optimal value and the dual optimal solution λ^*. Sketch the Lagrange dual function g and indicate its maximum value on the plot. Does strong duality hold? [4 marks]

e) Let $p^*(u)$ denote the optimal value of the perturbed problem

$$\begin{aligned} \min \quad & x^2+2x+1 \\ s.t. \quad & x(x-4) \leq u, \end{aligned}$$

as a function of the parameter u. Determine $p^*(u)$. [4 marks]

3. a) Prove the following facts.

i) The Euclidean ball $B(x_c, r) = \{x : ||x - x_c||_2 \le r\}$ is convex. Name any inequality that you use. [3 marks]

ii) Let a and b be distinct points in $\mathbb{R}^n$. Show that the set of all points that are closer in Euclidean norm to a than b, i.e. $\{x : ||x-a||_2^2 \le ||x-b||_2^2\}$ is a halfspace.
Hint: show that it can be written as $\{x : c^\top x \le d\}$ *for some* c *and* d.

[3 marks]

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

i) $f(x_1, x_2) = x_1 x_2$ on $\mathbb{R}^2_{++}$. [3 marks]

ii) $f(x_1, x_2) = \frac{x_1^2}{x_2}$ on $\mathbb{R} \times \mathbb{R}_{++}$. [3 marks]

c) Consider the basic norm approximation problem

$$\min_x \phi(Ax - b)$$

where $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, $x \in \mathbb{R}^n$ and ϕ is a penalty function. Explain how the residuals r_i, where $r = Ax - b$, tend to be distributed in the following cases:

i) ϕ is the ℓ_1-norm.

ii) ϕ is the ℓ_2-norm.

iii) ϕ is the ℓ_∞-norm.

[4 marks]

d) Formulate the problem

$$\begin{array}{ll} \min\limits_x & ||Ax - b||_1 \\ \text{s.t.} & ||x||_\infty \le 1, \end{array} \tag{3.1}$$

where $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, and $x \in \mathbb{R}^n$, as a linear programme. Explain in detail the relation between problem (3.1) and the obtained linear programme.
[4 marks]

4. The log barrier is based on the approximation $\widehat{I_-}(u) = -\frac{1}{t}\log(-u)$ of the indicator function $I_-(u)$ of the non-positive reals. We can also construct barriers from other approximations, which in turn yield generalizations of the central path and barrier method. Let $h : \mathbb{R} \to \mathbb{R}$ be a twice differentiable, closed, increasing convex function, with **dom** $h = -\mathbb{R}_{++}$ (this implies $h(u) \to \infty$ as $u \to 0$). One such function is the standard logarithm $h(u) = -\log(-u)$; another example is $h(u) = -\frac{1}{u}$ (for $u < 0$). Now consider the optimisation problem (without equality constraints)

$$\begin{array}{ll} \min\limits_{x} & f_0(x) \\ \text{s.t.} & f_i(x) \leq 0, \qquad i = 1, \dots, m, \end{array}$$

where f_i are twice differentiable. We define the h-barrier for this problem as

$$\phi_h(x) = \sum_{i=1}^{m} h(f_i(x)),$$

with domain $\{x : f_i(x) < 0, i = 1, \dots, m\}$. When $h(u) = -\log(-u)$, this is the usual *log barrier*; when $h(u) = -\frac{1}{u}$, then ϕ_h is called the *inverse barrier*. We define the h-central path as

$$x^*(t) = \operatorname{argmin} t f_0(x) + \phi_h(x),$$

where $t > 0$ is a parameter (we assume that for each t, the minimiser exists and is unique).

a) Explain why $t f_0(x) + \phi_h(x)$ is convex in x, for each $t > 0$.
Hint: see the "Information for candidates". [4 marks]

b) Consider the case $h(u) = -\frac{1}{u}$.

i) Let $z = x^*(t)$ be the minimiser of $t f_0(x) + \phi_h(x)$. Find the dual feasible $\lambda^*(t)$ as function of z. Find the associated duality gap. [8 marks]

ii) Recall that we can interpret the central path conditions as a continuous deformation of the KKT complementary condition. Using $\lambda^*(t)$ found in part b.i), determine the deformed complementary condition. Show that as $t \to \infty$, the original complementary condition is recovered. Explain any difference with respect to the modified complementary condition of the standard log barrier. [4 marks]

c) Consider now the case $h(u) \neq -\frac{1}{u}$. The duality gap is

$$\frac{1}{t}\sum_{i=1}^{m} h'(f_i(z))(-f_i(z)).$$

Determine all the convex functions h for which the duality gap depends only on t and m, and no other problem data (i.e. not on f_i). Compare these functions with the standard log barrier. [4 marks]