Advanced Optimisation

1. a) i) The problem is

$$\begin{aligned} \max_{w,d,h} \quad & wdh \\ s.t. \quad & 2(wh+dh) \le A_{wall}, \\ & wd \le A_{floor}, \\ & \alpha \le \frac{h}{w} \le \beta, \\ & \gamma \le \frac{d}{w} \le \delta. \end{aligned}$$

The geometric problem in posynomial form is

$$\begin{aligned} \min_{w,d,h} \quad & w^{-1}d^{-1}h^{-1} \\ s.t. \quad & 2A_{wall}^{-1}(wh+dh) \le 1, \\ & A_{floor}^{-1}wd \le 1, \\ & \alpha wh^{-1} \le 1, \qquad \beta^{-1}hw^{-1} \le 1, \\ & \gamma wd^{-1} \le 1, \qquad \delta^{-1}dw^{-1} \le 1. \end{aligned}$$

[5 marks]

ii) The geometric problem in convex form is

$$\begin{aligned} \min_{y_w,y_d,y_h} \quad & -y_w - y_d - y_h \\ s.t. \quad & \log(e^{y_w+y_h} + e^{y_d+y_h}) + \log 2A_{wall}^{-1} \le 0, \\ & y_w + y_d - \log A_{floor} \le 0, \\ & y_w - y_h + \log\alpha \le 0, \qquad y_h - y_w - \log\beta \le 0, \\ & y_w - y_d + \log\gamma \le 0, \qquad y_d - y_w - \log\delta \le 0. \end{aligned}$$

[5 marks]

b) i) The problem is

$$\begin{aligned} \max_{x} \quad & \min_{j=1,\dots,m} (Ax)_j \\ s.t. \quad & 0 \le x \le 1, \\ & \mathbf{1}^\top x = k. \end{aligned}$$

[4 marks]

ii) The problem can be reformulated as the linear programme

$$\begin{aligned} \max_{x} \quad & t \\ s.t. \quad & t \le Ax, \\ & 0 \le x \le 1, \\ & \mathbf{1}^\top x = k. \end{aligned}$$

[3 marks]

iii) If at iteration k, $\tilde{x}_i^{(k)}$ is close to zero, then the weight $w_i^{(k+1)}$ is very large and the term $-w_i x_i$ causes a large decrease of the objective function unless x_i is pushed even closer to 0. Thus the heuristic produces a solution for which its components close to 0 are incentivised to become even closer to 0. This method does not produce a Boolean solution and it does not change much the values of the x_i which are close to 1. In thresholding, the resulting vector is Boolean. [3 marks]

2. a) The feasible set is the interval $[0,4]$. The (unique) optimal point is $x^* = 0$, and the optimal value is $p^* = 1$. [4 marks]

b) Figure 2.1 shows the objective function (solid/blue line), the constraint function (dashed/orange line), the feasible set (shaded/grey area). The plot also indicates p^* (red/dot). [4 marks]

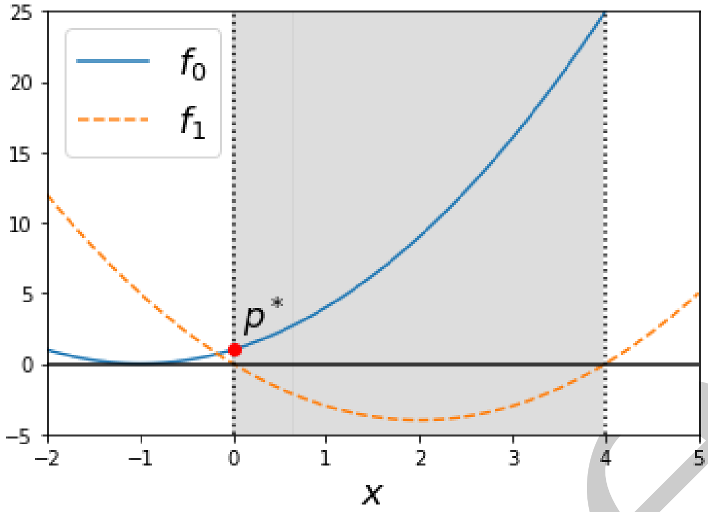

Figure 2.1 Plot for part b).

c) The Lagrangian is

$$L(x,\lambda) = x^2 + 2x + 1 + \lambda(x^2 - 4x) = (1+\lambda)x^2 + (2-4\lambda)x + 1.$$

For $\lambda > -1$, the Lagrangian reaches its minimum at $\tilde{x} = \frac{2\lambda - 1}{1+\lambda}$. For $\lambda \leq -1$, the Lagrangian is unbounded from below. Thus the dual function is

$$g(\lambda) = \begin{cases} \frac{(5-4\lambda)\lambda}{1+\lambda} & \lambda > -1 \\ -\infty & \lambda \leq -1. \end{cases}$$

[4 marks]

d) The Lagrange dual problem is

$$\begin{array}{ll} \max\limits_{\lambda} & \frac{(5-4\lambda)\lambda}{1+\lambda} \\ s.t. & \lambda \geq 0. \end{array}$$

The dual optimum occurs at $\lambda^* = \frac{1}{2}$, with $d^* = 1$. Figure 2.2 shows the dual function (solid/blue line) and the maximum $d^* = p^*$ (red/dot). We can directly observe that strong duality holds (as it must as Slater's constraint qualification is satisfied). [4 marks]

e) The perturbed problem is infeasible for $u < -4$, since $\inf_x (x^2 - 4x) = -4$. For $u \geq -4$, the feasible set is the interval

$$\left[2 - \sqrt{4+u}, 2 + \sqrt{4+u}\right],$$

given by the two roots of $x^2 - 4x - u = 0$. For $-1 \leq u \leq 5$ the optimiser is $x^*(u) = 2 - \sqrt{4+u}$ (the closest point in the feasible set to the minimiser of f_0,

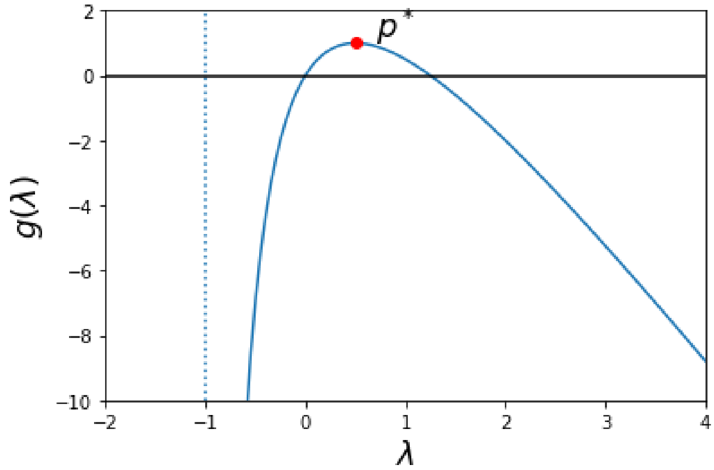

Figure 2.2 Plot for part d).

i.e. $x^* = -1$). For $u \geq 5$, the optimiser is the unconstrained minimiser of f_0, i.e. $x^*(u) = -1$ and $p^*(u) = 0$. In summary,

$$p^*(u) = \begin{cases} \infty & u < -4 \\ u - 6\sqrt{4+u} + 13 & -4 \leq u \leq 5 \\ 0 & u \geq 5. \end{cases}$$

[4 marks]

3. a) i) Take two points x_1 and x_2 in the ball, i.e. $||x_1 - x_c||_2 \leq r$ and $||x_2 - x_c||_2 \leq r$. We want to show that any convex combination is inside the ball. For any $0 \leq \theta \leq 1$ we have

$$\begin{aligned} ||\theta x_1 + (1-\theta)x_2 - x_c||_2 &= ||\theta(x_1 - x_c) + (1-\theta)(x_2 - x_c)||_2 \\ &\leq \theta||x_1 - x_c||_2 + (1-\theta)||x_2 - x_c||_2 \\ &\leq r, \end{aligned}$$

where we have used the triangle inequality. [3 marks]

ii) Note that

$$\begin{aligned} ||x-a||_2^2 \leq ||x-b||_2^2 &\iff (x-a)^\top(x-a) \leq (x-b)^\top(x-b) \\ &\iff x^\top x - 2a^\top x + a^\top a \leq x^\top x - 2b^\top x + b^\top b \\ &\iff 2(b-a)^\top x \leq b^\top b - a^\top a. \end{aligned}$$

Therefore, the result follows by taking $c = 2(b-a)$ and $d = b^\top b - a^\top a$. [3 marks]

b) i) The Hessian of f is

$$\nabla^2 f(x) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix},$$

which is neither positive semidefinite nor negative semidefinite. Therefore f is neither convex nor concave. [3 marks]

ii) We have seen in the module that this function (quadratic over linear) is convex. We can prove this by computing the Hessian

$$\nabla^2 f(x) = \begin{bmatrix} \frac{2}{x_2} & -\frac{2x_1}{x_2^2} \\ -\frac{2x_1}{x_2^2} & \frac{2x_1^2}{x_2^3} \end{bmatrix} = \frac{2}{x_2}\begin{bmatrix} 1 \\ -\frac{2x_1}{x_2} \end{bmatrix}\begin{bmatrix} 1 & -\frac{2x_1}{x_2} \end{bmatrix}$$

Since x_2 is in $\mathbb{R}_{++}$, the Hessian is positive semidefinite and f is convex. [3 marks]

c) i) The ℓ_1-norm penalty function produces many residuals which are very small or exactly zero. This is due to the fact that this norm continues to weigh small residuals in the same way as bigger residuals and there is always an incentive to push a residual to zero (while for the ℓ_2-norm, for instance, the incentive decreases quadratically). Thus, the ℓ_1-norm penalty function tends to produce *sparse* solutions which may be of interest for physical or numerical reasons.

ii) The ℓ_2-norm penalty puts very small weight on small residuals, but large weight on large residuals. The approximation has many modest, but non-zero, residuals, and relatively few larger ones.

iii) As the minimisation of the ℓ_∞-norm tries to minimize the maximum residual (in absolute value), it naturally shows an accumulation of residuals around the maximum value. That is because if a residual is larger than the maximum, then there is an incentive to push it down. However, if a residual is below the maximum, there is no incentive to move it around. So residuals between the maximum and minimum are more or less evenly spread. [4 marks]

d) The equivalent LP is

$$\begin{aligned} \min_{t,x} \quad & \mathbf{1}^\top t \\ \text{s.t.} \quad & -t \preccurlyeq Ax - b \preccurlyeq t, \\ & -\mathbf{1} \preccurlyeq x \preccurlyeq \mathbf{1}. \end{aligned}$$

Assume x is fixed and optimise over t. The first constraint in the LP says that $-t_k \leq a_k^\top x - b_k \leq t_k$ for all k. Note that each of these constraints depends only on one t_k and that the objective function of the LP is the sum of the t_k. Hence, the objective function is separable and the optimum over t_k is achieved by $t_k = |a_k^\top x - b_k|$.

Clearly, if x is fixed, the optimal value over t of the LP is $||Ax - b||_1$. Thus, optimising over t and x simultaneously is equivalent to the original problem.

For the constraint $||x||_\infty \leq 1$ note that if $|x_k| \leq 1$ for all k, then $1 \geq \max_k |x_k| = ||x||_\infty$. [4 marks]

4. a) The function h is increasing and its domain extends infinitely in the negative direction. Hence, $\tilde{h}$ is increasing. Thus, the composition rules show that $\phi_h(x)$ is convex in x, since $\tilde{h}$ is increasing and convex, and the f_i are convex. Then $tf_0(x) + \phi_h(x)$ is convex because it is a positive linear combination of convex functions. [4 marks]

b) i) Let $z = x^*(t)$ be the minimiser of $tf_0(x) + \phi_h(x)$. Thus, the minimiser satisfies $t\nabla f_0(z) + \nabla\phi_h(z) = 0$. Expanding we get

$$t\nabla f_0(z) + \sum_{i=1}^{m} \frac{1}{f_i^2(z)} \nabla f_i(z).$$

This shows that z minimises the Lagrangian $f_0(x) + \sum_{i=1}^{m} \lambda_i(t) f_i(x)$ for

$$\lambda_i^*(t) = \frac{1}{tf_i^2(z)}, \qquad i = 1, \ldots, m.$$

The associated dual function is

$$g(\lambda^*) = f_0(z) + \sum_{i=1}^{m} \lambda_i^*(t) f_i(z) = f_0(z) + \sum_{i=1}^{m} \frac{1}{tf_i(z)}.$$

Hence, the duality gap is

$$\frac{1}{t} \sum_{i=1}^{m} \frac{1}{f_i(z)}.$$

[8 marks]

ii) The complementary condition is

$$\lambda_i(t) f_i(z) = \frac{1}{tf_i^2(z)} f_i(z) = \frac{1}{tf_i(z)}.$$

As $t \to \infty$, $\lambda_i(t) f_i(z) \to 0$. Differently from the complementary condition of the log barrier, the deformation depends on $f_i(x)$. In particular, if the central path is close to violate a constraint, then the complementary condition may be very different from the original condition.

[4 marks]

c) The only way for the given expression not to depend on problem data (except t and m) is for $-uh'(u)$ to be constant. This means $h'(u) = -\frac{a}{u}$ for some constant a, so $h(u) = -a\log(-u) + b$ for some constant b. Since h must be convex and increasing, we need $a > 0$. Thus, h gives rise to a scaled, offset, log barrier. In particular, the central path associated with h is the same as for the standard log barrier. [4 marks]