1. a) i) The problem is

$$\begin{aligned}
\max_{p_k} \quad & \min_{i=1,\ldots,n} \frac{S_i}{I_i+\sigma_i} \\
s.t. \quad & 0 \le p_i \le P_i^{\max}, \quad i=1,\ldots,n, \\
& \sum_{k\in K_j} p_k \le P_j^{gp}, \quad j=1,\ldots,m, \\
& \sum_{k=1}^{n} G_{ik} p_k \le P_i^{rc}, \quad i=1,\ldots,n.
\end{aligned}$$

1 mark to the objective function, 1 mark for each constraint (four constraints). 1 mark extra mark for everything correct [6 marks]

ii) Since the problem is max $f(x)$, we want to be able to test when $f(x) \ge t$ and push this t as high as possible. Thus we look for a family of convex functions ϕ_t parametrised in t such that

(I) $f(x) \ge t \iff \phi_t(x) \le 0$ for each x.

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

Condition (II) ensures that if $\phi_t(x) \le 0$ (for all x), then $\phi_s(x) \le 0$ (for all x) for $s \le t$. This ensures consistency as if (I) holds for a high value t, it must hold for all lower values $s \le t$.
We want

$$\min_{i=1,\ldots,n} \frac{S_i}{I_i+\sigma_i} \ge t.$$

This is equivalent to

$$\frac{S_i}{I_i+\sigma_i} \ge t, \qquad \text{for all } i=1,\ldots,n.$$

From the positivity of $I_i+\sigma_i$

$$t(I_i+\sigma_i)-S_i \le 0, \qquad \text{for all } i=1,\ldots,n,$$

follows. So $\phi_t^i(x) = t(I_i+\sigma_i)-S_i$. Note that $s(I_i+\sigma_i)-S_i \le t(I_i+\sigma_i)-S_i$ for any $s \le t$ (again from the positivity of $I_i+\sigma_i$). Thus $\phi_t^i(x)$ satisfies both required properties.

3 marks for the derivation of Φ_t, 1 mark for showing it is nondecreasing. [4 marks]

b) i) The problem, as stated, is

$$\begin{aligned}
\max_{T,r,w} \quad & \alpha_4 T r^2 \\
s.t. \quad & \alpha_1 T r w^{-1} + \alpha_2 r + \alpha_3 r w \le C_{\max}, \\
& T_{\min} \le T \le T_{\max}, \\
& r_{\min} \le r \le r_{\max}, \\
& w_{\min} \le w \le w_{\max}, \\
& w \le \frac{1}{10} r.
\end{aligned}$$

This can be written in posynomial form as

$$
\begin{aligned}
\min_{T,r,w} \quad & \alpha_4^{-1} T^{-1} r^{-2} \\
s.t. \quad & \frac{\alpha_1}{C_{\max}} T r w^{-1} + \frac{\alpha_2}{C_{\max}} r + \frac{\alpha_3}{C_{\max}} r w \le 1, \\
& T_{\max}^{-1} T \le 1, \qquad T_{\min} T^{-1} \le 1, \\
& r_{\max}^{-1} r \le 1, \qquad r_{\min} r^{-1} \le 1, \\
& w_{\max}^{-1} w \le 1, \qquad w_{\min} w^{-1} \le 1, \\
& 10 w r^{-1} \le 1.
\end{aligned}
$$

1 mark for the objective function, 1 for the total cost, 1 mark for the three set constraints, 1 mark for the 10% *constraint, 1 mark for writing it in polynomial form.* [5 marks]

ii) Let $y_T = \log T$, $y_r = \log r$ and $y_w = \log w$. The geometric problem in convex form is

$$
\begin{aligned}
\min_{y_T, y_r, y_w} \quad & -y_T - 2y_r - \log \alpha_4 \\
s.t. \quad & \log\left(\frac{\alpha_1}{C_{\max}} e^{y_T + y_r - y_w} + \frac{\alpha_2}{C_{\max}} e^{y_r} + \frac{\alpha_3}{C_{\max}} e^{y_r + y_w} \right) \le 0, \\
& y_T - \log T_{\max} \le 0, \qquad -y_T + \log T_{\min} \le 0, \\
& y_r - \log r_{\max} \le 0, \qquad -y_r + \log r_{\min} \le 0, \\
& y_w - \log w_{\max} \le 0, \qquad -y_w + \log w_{\min} \le 0, \\
& y_w - y_r + \log 10 \le 0.
\end{aligned}
$$

1 mark for the objective function, 1 for the log sum exp, 1 mark for three left set constraints, 1 mark for three right set constraints, 1 mark for the 10% *constraint.* [5 marks]

2. a) i) Figure 2.1 shows the feasible set (the intersection of the two shaded disks) and some contour lines of the objective function. There is only one feasible point, $x^* = (1,0)$, so it is optimal for the primal problem, and we have $p^* = 1$.

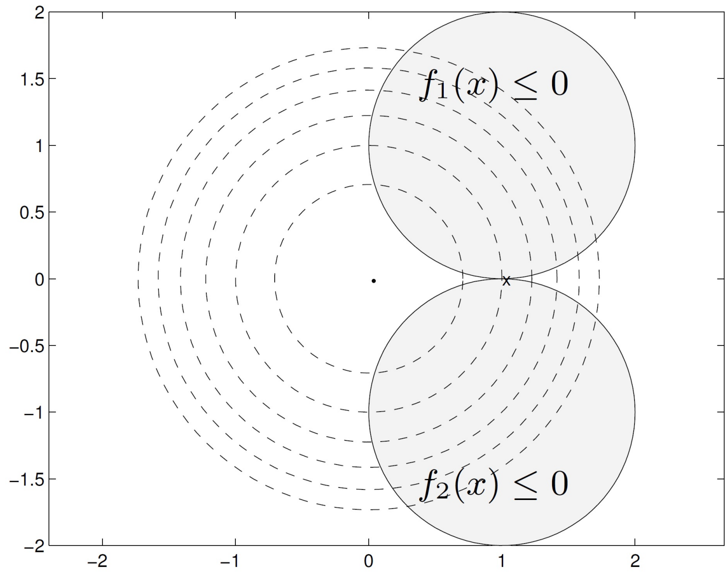

Figure 2.1 Plot for part 2.a).

1 mark for the plot, 1 mark for the optimal point, 1 mark for the optimal value. [3 marks]

ii) The KKT conditions are

$$\begin{aligned} (x_1-1)^2+(x_2-1)^2 &\leq 1, & (x_1-1)^2+(x_2+1)^2 &\leq 1, \\ \lambda_1 &\geq 0, & \lambda_2 &\geq 0, \\ 2x_1+2\lambda_1(x_1-1)+2\lambda_2(x_1-1) &= 0, & 2x_2+2\lambda_1(x_2-1)+2\lambda_2(x_2+1) &= 0, \\ \lambda_1((x_1-1)^2+(x_2-1)^2-1) &= 0, & \lambda_2((x_1-1)^2+(x_2+1)^2-1) &= 0. \end{aligned}$$

At $x^* = (1,0)$, these conditions reduce to

$$\lambda_1 \geq 0, \qquad \lambda_2 \geq 0, \qquad 2 = 0, \qquad \lambda_1 = \lambda_2,$$

which, because of the third condition, do not hold.

1 mark for the KKT conditions, 1 mark for checking the conditions. [2 marks]

iii) The Lagrange dual function is

$$g(\lambda_1,\lambda_2) = \inf_{x_1,x_2} L(x_1,x_2,\lambda_1,\lambda_2).$$

The Lagrangian L is

$$\begin{aligned} &L(x_1,x_2,\lambda_1,\lambda_2) \\ &= x_1^2+x_2^2+\lambda_1((x_1-1)^2+(x_2-1)^2-1)+\lambda_2((x_1-1)^2+(x_2+1)^2-1) \\ &= (\lambda_1+\lambda_2+1)x_1^2+(\lambda_1+\lambda_2+1)x_2^2-2(\lambda_1+\lambda_2)x_1-2(\lambda_1-\lambda_2)x_2+\lambda_1+\lambda_2. \end{aligned}$$

- If $\lambda_1+\lambda_2+1 < 0$, then for $x_1 \to +\infty$ or $x_2 \to +\infty$, we have $g \to -\infty$.

- If $\lambda_1 + \lambda_2 + 1 = 0$, then

$$L(x_1, x_2, \lambda_1, \lambda_2) = 2x_1 + 4\lambda_2 x_2 + 2x_2 - 1,$$

and for e.g. $x_1 \to -\infty$, we have $g \to -\infty$.

- If $\lambda_1 + \lambda_2 + 1 > 0$, the L reaches its minimum for

$$x_1 = \frac{\lambda_1 + \lambda_2}{\lambda_1 + \lambda_2 + 1}, \qquad x_2 = \frac{\lambda_1 - \lambda_2}{\lambda_1 + \lambda_2 + 1}.$$

In summary,

$$g(\lambda_1, \lambda_2) = \begin{cases} \dfrac{\lambda_1 + \lambda_2 - (\lambda_1 - \lambda_2)^2}{\lambda_1 + \lambda_2 + 1}, & \lambda_1 + \lambda_2 + 1 > 0, \\ -\infty, & \text{otherwise.} \end{cases}$$

1 mark for the Lagrangian, 1 mark for each case, 1 mark the dual function. [5 marks]

iv) The Lagrange dual problem is

$$\begin{array}{ll} \max & \dfrac{\lambda_1 + \lambda_2 - (\lambda_1 - \lambda_2)^2}{\lambda_1 + \lambda_2 + 1} \\ s.t. & \lambda_1 \geq 0, \qquad \lambda_2 \geq 0. \end{array}$$

Since g is symmetric, the optimum (if it exists) occurs with $\lambda_1 = \lambda_2$. The dual function then simplifies to

$$g(\lambda) = \frac{2\lambda}{2\lambda + 1}.$$

Computing the derivative, we see that this function is always increasing. Thus, $g(\lambda) \to 1$ as $\lambda \to +\infty$. We have $d^* = p^* = 1$, but the dual optimum is not attained. Despite strong duality holds, the KKT conditions fail because the dual optimum is not attained.

1 mark for the simplified function, 1 mark for optimal value, 1 mark for the justification. [3 marks]

b) i) $||x_{ls}||_2 = 3$.

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

iii) $||b||_2 = \sqrt{10}$.

iv) About 5.5.

v) About 5.5.

vi) About 6. In fact, from the curve we see that if $||Ax - b||_2^2 = 2$, then $||x||_2^2 = 9$, so 11, if $||Ax - b||_2^2 = 3$, then $||x||_2^2 = 4$, so 7, if $||Ax - b||_2^2 = 4$, then $||x||_2^2 = 2.1$, so 6.1, if $||Ax - b||_2^2 = 5$, then $||x||_2^2 = 1.5$, so 6.5.

vii) The rank of A is 10 since the least square solution is unique.

1 mark for each question. [7 marks]

3. a) i) We use the second-order condition.

$$f(x) = \log x, \quad f'(x) = \frac{1}{x}, \quad f''(x) = -\frac{1}{x^2}.$$

Since $f''(x) < 0$ for all $x \in \mathbb{R}_{++}$, f is concave. [1 mark]

ii) We use the second-order condition.

$$f(x) = x\log x, \quad f'(x) = \log x + 1, \quad f''(x) = \frac{1}{x}.$$

Since $f''(x) > 0$ for all $x \in \mathbb{R}_{++}$, f is convex. [1 mark]

iii) We use Jensen's inequality. Let $f(x) = \max\{x_1, x_2\}$ and $0 \leq \theta \leq 1$. Then

$$\begin{aligned} f(\theta x + (1-\theta)y) &= \max\{\theta x_1 + (1-\theta)y_1, \theta x_2 + (1-\theta)y_2\} \\ &\leq \max\{\theta x_1, \theta x_2\} + \max\{(1-\theta)y_1, (1-\theta)y_2\} \\ &= \theta \max\{x_1, x_2\} + (1-\theta)\max\{y_1, y_2\} \\ &= \theta f(x) + (1-\theta) f(y). \end{aligned}$$

[1 mark]

iv) We use the second-order condition.

$$f(x) = \log(e^{x_1} + e^{x_2}), \quad \nabla f(x) = \begin{bmatrix} \dfrac{e^{x_1}}{e^{x_1} + e^{x_2}} \\ \dfrac{e^{x_2}}{e^{x_1} + e^{x_2}} \end{bmatrix},$$

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

The term $\dfrac{e^{x_1+x_2}}{(e^{x_1} + e^{x_2})^2}$ is always positive. The matrix is positive semidefinite. This can be shown in several ways, e.g. with the test on the minors (first 1 and second 0), computing the eigenvalues (2 and 0) or showing that the matrix can be written as $\begin{bmatrix} -1 & 1 \end{bmatrix}^\top \begin{bmatrix} -1 & 1 \end{bmatrix}$. Since $\nabla^2 f(x) \geq 0$ for all $x \in \mathbb{R}^2$, f is convex.

1 mark for the Hessian, 1 mark for the conclusion. [2 marks]

b) i) Let $f(x) = ax + b$. As a function of x, $yx - ax - b$ is bounded if and only if $y = a$, in which case it is constant. Therefore,

$$f^*(y) = \begin{cases} -b, & y = a, \\ +\infty, & \text{otherwise.} \end{cases}$$

[2 marks]

ii) Let $f(x) = -\log x$, with $x \in \mathbb{R}_{++}$. The function $xy + \log x$ is unbounded above if $y \geq 0$. For $y < 0$, computing the gradient, the maximum is reached at $x = -1/y$. Therefore,

$$f^*(y) = \begin{cases} -\log(-y) - 1, & y < 0, \\ +\infty, & y \geq 0. \end{cases}$$

[2 marks]

c) i) Let $f(x) = x\log x$ on $\mathbb{R}_{++}$. Then

$$f'(x) = 1 + \log x, \quad f''(x) = \frac{1}{x}, \quad f'''(x) = -\frac{1}{x^2}.$$

Thus

$$\frac{|f'''(x)|}{f''(x)^{\frac{3}{2}}} = \frac{\frac{1}{x^2}}{\frac{1}{x^{\frac{3}{2}}}} = \frac{1}{\sqrt{x}}.$$

Note that as $x \to 0^+$, this is unbounded. Thus, for x small enough we have that $\frac{1}{\sqrt{x}} \leq 2$ fails. This is the case, for instance, for $x = \frac{1}{5}$.

1 mark for the formula, 1 mark for the conclusion. [2 marks]

ii) Let $g(x) = x\log x - \log x$ on $\mathbb{R}_{++}$. Then

$$g'(x) = -\frac{1}{x} + 1 + \log x, \quad g''(x) = \frac{1}{x^2} + \frac{1}{x}, \quad g'''(x) = -\frac{2}{x^3} - \frac{1}{x^2}.$$

Thus

$$h(x) = \frac{|g'''(x)|}{g''(x)^{\frac{3}{2}}} = \frac{2+x}{(1+x)^{\frac{3}{2}}}.$$

Note that $h(0) = 2$ and $h'(x) = -\frac{2+\frac{x}{2}}{(1+x)^{\frac{5}{2}}}$. Recall that we are on $\mathbb{R}_{++}$, thus this last function is always negative. In summary, $h(0) = 2$ and $h(x)$ is strictly decreasing, showing that

$$\frac{|g'''(x)|}{g''(x)^{\frac{3}{2}}} \leq 2.$$

Hence, the function is self-concordant on $\mathbb{R}_{++}$.

1 mark for h, 1 mark for h', 1 mark for the conclusion. [3 marks]

d) i) The centrality equation is given by

$$0 = t\nabla f_0(x^*(t)) + \nabla\phi(x^*(t)) + A^\top \hat{v}.$$

Differentiating the centrality equation yields

$$\nabla f_0(z) + \left(t\nabla^2 f_0(z) + \nabla^2\phi(z)\right)\frac{dx^*}{dt} = 0.$$

Thus, the tangent to the central path at $x^*(t)$ is given by

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

[2 marks]

ii) To show that $f_0(x^*(t))$ decreases as t increases, we show that

$$\frac{df_0(x^*(t))}{dt} < 0.$$

In fact,

$$\begin{aligned}\frac{df_0(x^*(t))}{dt} &= \nabla f_0(x^*(t))^\top \frac{dx^*}{dt} \\ &= -\nabla f_0(x^*(t))^\top \left(t\nabla^2 f_0(x^*(t)) + \nabla^2\phi(x^*(t))\right)^{-1}\nabla f_0(x^*(t)) < 0,\end{aligned}$$

since it is a negative quadratic form. [2 marks]

iii) We compute $\hat{x}$ as

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

Comparing this with x^n, we notice that the difference is in the terms $t\nabla^2 f_0(x^*(t))$ (for the predictor-corrector method) vs $\mu t\nabla^2 f_0(x^*(t))$ (for Newton's method). Computing x^n and $\hat{x}$ for $f_0(x) = c^\top x$ yields

$$\hat{x} = x^n = x^*(t) - (\mu - 1)t \left(\nabla^2 \phi(x^*(t))\right)^{-1} c,$$

i.e. they are the same. It follows that whenever $\nabla^2 f_0(x^*(t)) = 0$ (e.g. for LPs and QPs), the two methods are identical. [2 marks]