1. let $h(x,y) = f(x-y)d(y)$. Then $\log h(x,y) = \log f(x-y)d(y) = \log f(x-y) + \log d(y)$.
We know $\log f(x-y)$ and $\log d(y)$ are concave, so, their sum $\log h(x,y)$ is also concave.
Then we use the theorem that the integral of a log-concave function is also log-concave.
We thus know $\int h(x,y)dy = (f*g)(x)$ is log-concave.

2. a). show $f(x) = \frac{p(x)}{q(x)}$ is quasi-convex, where $p(x)$ is convex and $p(x) \geq 0$, $q(x)$ is concave and $q(x) > 0$.

This is equivalent to show that for $\lambda \in [0,1]$, $x, y \in dom(f)$,

$$f(\lambda x + (1-\lambda)y) \leq \max(f(x), f(y))$$

- If $\max(f(x), f(y)) = f(x)$, then we need to show that $\frac{p(\lambda x+(1-\lambda)y)}{q(\lambda x+(1-\lambda)y)} \leq \frac{p(x)}{q(x)}$.
~~This means we~~ ~~Well~~, lhs $\leq \frac{\lambda p(x) + (1-\lambda) p(y)}{\lambda q(x) + (1-\lambda) q(y)}$ ~~need to show~~, so we now need to show $\frac{\lambda p(x) + (1-\lambda)p(y)}{\lambda q(x) + (1-\lambda) q(y)} - \frac{p(x)}{q(x)} \leq 0$.
This is equivalent to show $\lambda p(x)q(x) + (1-\lambda) p(y) q(x) - \lambda p(x) q(x) - (1-\lambda) q(y) p(x) \leq 0$
since the denominator is positive.
This is equivalent to show $(1-\lambda)(p(y)q(x) - q(y)p(x)) \leq 0$, which is true because
$\frac{p(x)}{q(x)} \geq \frac{p(y)}{q(y)} \Rightarrow p(y)q(x) \leq p(x)q(y) \Rightarrow p(y)q(x) - q(y)p(x) \leq 0$ ✓
- Similarly, if $\max(f(x), f(y)) = f(y)$, we can carry out the same calculation to show that $\frac{p(\lambda x+(1-\lambda)y)}{q(\lambda x+(1-\lambda)y)} \leq \frac{p(y)}{q(y)}$.

Hence, $f(\lambda x+(1-\lambda)y) \leq \max(f(x), f(y))$, so, $f(x) = \frac{p(x)}{q(x)}$ is quasi-convex.

2. (continued). (b). $f(x) \le t \iff \frac{p(x)}{q(x)} \le t \iff p(x) \le t q(x) \iff p(x) - t q(x) \le 0$

(c). We want to show $\phi_t(x) = p(x) - t q(x)$ is convex. Then for $x, y \in dom(f)$, $\lambda \in [0,1]$,

$$\phi_t(\lambda x + (1-\lambda) y) = p(\lambda x + (1-\lambda) y) - t q(\lambda x + (1-\lambda) y) \le \lambda p(x) + (1-\lambda) p(y) - t(\lambda q(x) + (1-\lambda) q(y))$$
$$= \lambda (p(x) - t q(x)) + (1-\lambda)(p(y) - t q(y))$$
$$= \lambda \phi_t(x) + (1-\lambda) \phi_t(y)$$

Thus, $\phi_t(x)$ is convex.

For $t_1 < t_2$, $\phi_{t_1}(x) - \phi_{t_2}(x) = p(x) - t_1 q(x) - (p(x) - t_2 q(x)) = (t_2 - t_1) q(x) > 0$

Hence, $\phi_t(x)$ is decreasing in t.

3. $f(x)$ is quasi-linear, so $\{x \mid f(x) \le \alpha\}$ (sublevel set) and its complement $\{x \mid f(x) > \alpha\}$ are both convex set for any $\alpha \in \mathbb{R}$. Furthermore, we know the sublevel set should be nested, which means $\{x \mid f(x) \le \alpha_1\} \subseteq \{x \mid f(x) \le \alpha_2\}$ for $\alpha_1 < \alpha_2$.

Hence, we know the sublevel set ~~can~~ is a halfspace and the norm vector stays the same as α changes $\{x \mid f(x) \le \alpha\} = \{x \mid a^T x \le b(\alpha)\}$, where $b(\alpha)$ is non-decreasing.

If $b(\alpha)$ is strictly increasing, then we can rewrite the sublevel set as $\{x \mid b^{-1}(a^T x) \le \alpha\}$, where we showed $f(x) = g(a^T x)$, where $g = b^{-1}$ is monotone (increasing).

If $b(\alpha)$ is non-decreasing, let $b^{-1}(y) = \inf_x \{x \mid b(x) = y\}$

4. See next page.

4. If we write down the dual of the primal problem, it's

$$\begin{aligned} \max \quad & -b^T z \\ \text{s.t.} \quad & A^T z + c = 0 \\ & z \geq 0 \end{aligned}$$

- If the dual problem is feasible, ~~and bounded~~ then it's also bounded, because otherwise, the primal is infeasible, which is impossible since A is square and non-singular. Then, let the optimal solution for the dual be z^*, we know $\begin{cases} A^T z^* + c = 0 \\ z^* \geq 0 \end{cases} \Rightarrow A^{-T}(-c) \geq 0$
By strong duality, we also know $p^* = d^* = -b^T z^* = b^T A^{-T} c$

- If the dual problem is infeasible, then the primal is either infeasible or unbounded. In this problem, the primal is always feasible so the primal is unbounded, $p^* = -\infty$

Therefore, $p^* = \begin{cases} b^T A^{-T} c, & \text{if } A^{-T} c \leq 0 \\ -\infty, & \text{otherwise} \end{cases}$

5. a). If $A \succ 0$, ~~the~~ we can first make a change of variables $y = A^{1/2} x$, and write $\tilde{c} = A^{-\frac{1}{2}} c$. With this new variable, the problem becomes $\begin{aligned} \min \quad & \tilde{c}^T y \\ \text{s.t.} \quad & y^T y \leq 1 \end{aligned}$, which means minimizing a linear function over the unit ball.
The optimal solution is $y^* = \frac{-\tilde{c}}{\|\tilde{c}\|_2}$

In general when $A \notin S_+^n$, we can make a change of variables based on eigendecomposition $A = Q \operatorname{diag}(\lambda) Q^T = \sum_{i=1}^{n} \lambda_i q_i q_i^T$. We define $y = Qx$, $b = Qc$ and express the problem as

$$\begin{aligned} \min \quad & \sum_{i=1}^{n} b_i y_i \\ \text{s.t.} \quad & \sum_{i=1}^{n} \lambda_i y_i^2 \leq 1 \end{aligned}$$

- If $\lambda_i > 0, \forall i$, the problem is reduced to what talked about above,
- $\lambda_n < 0$, the problem is unbounded below. By making $y_n \to +\infty$ (or $-\infty$), we can make every point feasible
- $\lambda_n = 0$, If for some i, $b_i \neq 0$ and $\lambda_i = 0$, the problem is unbounded below
- $\lambda_n = 0$, and $b_i = 0$ for all i with $\lambda_i = 0$, we can reduce the problem to smaller one with all $\lambda_i > 0$.

5. (continued) (b). Make a change of variables $y = A^{1/2}(x - x_c)$, $x = A^{-\frac{1}{2}}y + x_c$ and solve the new equivalent problem $\min \ c^T A^{-\frac{1}{2}} y + c^T x_c$

$$\text{s.t.} \quad y^T y \le 1$$

The solution is $y^* = -\left(\frac{1}{\|A^{-\frac{1}{2}}c\|_2}\right) A^{-\frac{1}{2}} c$, $x^* = x_c - \left(\frac{1}{\|A^{-\frac{1}{2}}c\|_2}\right) A^{-1} c$

(c). If $B \succeq 0$, then the optimal value is obviously zero (since $x^T B x \ge 0$ for all x, with equality if $x = 0$)

In general, we can make a change of variables $y = A^{\frac{1}{2}} x$ with $A \succ 0$. The problem then becomes $\min \ y^T A^{-\frac{1}{2}} B A^{\frac{1}{2}} y$.

$$\text{s.t.} \quad y^T y \le 1$$

If the constraint $y^T y \le 1$ is active at the optimum $(y^{*T} y^* = 1)$, then the optimal value is $\lambda_{min}(A^{-\frac{1}{2}} B A^{-\frac{1}{2}})$ because $\lambda_{min} = \inf\limits_{x^T x = 1} x^T A x$.

If $y^{*T} y < 1$, then it must be at a point where the gradient of objective ~~is~~ function vanishes, i.e., $By = 0$, in which the optimal value is zero.