Problem 1

First calculate the Hessian of the function f,

$$\begin{aligned}
f(x) &= \prod_{i=1}^{n} x_i^{\alpha_i} \\
\frac{\partial f}{\partial x_i} &= \prod_{j=1}^{n} x_j^{\alpha_j} \left(\alpha_i \cdot \frac{1}{x_i} \right) = \frac{\alpha_i}{x_i} f(x) \\
\frac{\partial^2 f}{\partial x_i \partial x_j} &= \begin{cases} \dfrac{\alpha_i(\alpha_i - 1)}{x_i^2} f(x), & i = j \\ \dfrac{\alpha_i \alpha_j}{x_i x_j} f(x), & i \neq j \end{cases}
\end{aligned}$$

Now given any vector $y \in \mathbb{R}^n$,

$$\begin{aligned}
y^T \nabla^2 f(x) y &= \sum_{i,j=1}^{n} y_i \frac{\partial^2 f}{\partial x_i \partial x_j} y_j \\
&= f(x) \left(\sum_{i=1}^{n} \frac{\alpha_i(\alpha_i - 1) y_i^2}{x_i^2} + \sum_{i \neq j} \frac{\alpha_i \alpha_j y_i y_j}{x_i x_j} \right) \\
&= f(x) z^T B z
\end{aligned}$$

where $z \in \mathbb{R}^n, B \in \mathbb{R}^{n \times n}$ are vector and matrix such that $z_i = y_i / x_i$ and

$$B_{ij} = \begin{cases} \alpha_i(\alpha_i - 1), & i = j \\ \alpha_i \alpha_j, & i \neq j \end{cases}$$

Note that B is diagonally dominant since $\sum_{j=1}^{n} \alpha_j \leq 1$,

$$|B_{ii}| - \sum_{j \neq i} |B_{ij}| = \alpha_i \left(1 - \alpha_i - \sum_{j \neq i} \alpha_j \right) = \alpha_i \left(1 - \sum_{j=1}^{n} \alpha_j \right) \geq 0$$

Now that B have negative diagonal entries $\alpha_i(\alpha_i - 1)$, it's a negative semi-definite matrix, $z^T B z \leq 0$. Therefore $y^T \nabla^2 f(x) y = f(x) z^T B z \leq 0$ and we proved the Hessian of f is negative semi-definite and f is concave. □

Problem 2

Suppose $\theta \in [0,1]$, $\mu = 1-\theta$; use the convexity of g,

$$\begin{aligned}
g(\theta x_1 + \mu x_2, y) &\le \theta g(x_1, y) + \mu g(x_2, y), \forall y \\
g(\theta x_1 + \mu x_2, y) - t &\le \theta (g(x_1, y) - t) + \mu (g(x_2, y) - t), \forall y, t \\
\max\{g(\theta x_1 + \mu x_2, y) - t, 0\} &\le \theta \max\{g(x_1, y) - t, 0\} + \mu \max\{g(x_2, y) - t, 0\}, \forall y, t \\
\mathbb{E}\max\{g(\theta x_1 + \mu x_2, y) - t, 0\} &\le \mathbb{E}\left(\theta \max\{g(x_1, y) - t, 0\} + \mu \max\{g(x_2, y) - t, 0\}\right) \\
&= \theta \mathbb{E}\max\{g(x_1, y) - t, 0\} + \mu \mathbb{E}\max\{g(x_2, y) - t, 0\}, \forall t \\
t + \frac{1}{1-\beta}\mathbb{E}\max\{g(\theta x_1 + \mu x_2, y) - t, 0\} &\le \theta\left(t + \frac{1}{1-\beta}\mathbb{E}\max\{g(x_1, y) - t, 0\}\right) \\
&\quad + \mu\left(t + \frac{1}{1-\beta}\mathbb{E}\max\{g(x_2, y) - t, 0\}\right), \forall t
\end{aligned}$$

Problem 3

1. The Lagrangian and the Lagrange daul are

$$\begin{aligned}
L(x, y, \nu) &= \|y\|_2 + \gamma\|x\|_1 + \nu^T(Ax - b - y) \\
g(\nu) &= \min_{x,y} \|y\|_2 + \gamma\|x\|_1 + \nu^T(Ax - b - y) \\
&= -\nu^T b + \min_x \nu^T Ax + \gamma\|x\|_1 + \min_y \|y\|_2 - \nu^T y \\
&= -\nu^T b - \gamma \max_x \left(\left(-\frac{1}{\gamma}A^T\nu\right)^T x - \|x\|_1\right) - \max_y \left(\nu^T y - \|y\|_2\right) \\
&= -\nu^T b - \chi_S(\nu)
\end{aligned}$$

where $S = \left\{\left\|-\frac{1}{\gamma}A^T\nu\right\|_1^* \le 1\right\} \cup \{\|\nu\|_2^* \le 1\} = \{\|A^T\nu\|_\infty \le \gamma\} \cup \{\|\nu\|_2 \le 1\}$ and the convex indicator function $\chi_S(\nu)$ will send the value $g(\nu)$ to $-\infty$ if $\nu \notin S$.
2. Since $f_0(x) = \|Ax - b\|_2 + \gamma\|x\|_1$ is minimized with x^* and $Ax^* - b \ne 0$, we have the subgradient

$$\partial f_0(x) = \frac{A^T(Ax^* - b)}{\|Ax^* - b\|_2} - \gamma sgn(x^*) = A^T r - \gamma sgn(x^*) \ni 0$$

Here $sgn(x_i^*) = \begin{cases} x_i^*/|x_i^*|, & x_i^* \ne 0 \\ [-1, 1], & x_i^* = 0 \end{cases}$ is the set valued function. We see now $|(A^T r)_i| \in |\gamma sgn(x_i^*)| \subseteq [0, \gamma]$ regardlessly and $\|A^T r\|_\infty \le \gamma$. If we dot product the above equation with x^*, we see that

$$r^T Ax^* - \gamma sgn(x^*)^T x^* = r^T Ax^* - \gamma\|x\|_1 \ni 0$$

Since $sgn(x^*)^T x^* = \{\|x\|_1\}$ becomes a single ton set, we have $r^T A x^* - \gamma\|x\|_1 = 0$.
3. From above we see

$$a_i^T r + \gamma sgn(x_i^*) = 0 \text{ if } x_i^* \neq 0$$

WLOG suppose $\|a_1\|_2 < \gamma$, then by Cauchy's inequality $a_i^T r \leq \|a_i\|_2\|r\|_2 < \gamma$ and this can't make sense of the above equality hence $x_1^* = 0$. □

Problem 4

We need a formula for

$$\sup_{C_i a_i \preceq d_i} \pm(a_i^T x - b)$$

when $x \in \mathbb{R}^n$ is fixed. This is in fact the optimal value of the following LP

$$\begin{array}{ll} \text{minimize} & \mp (a_i^T x - b_i) \\ \text{subject to} & C_i a_i \preceq d_i \end{array}$$

with variable a_i. Since there're only affine constraints, the Slater's condition is satisfied whenever the problem is feasible. The Lagrangian and the Lagrange dual of the above problem with negative sign are

$$\begin{aligned} L(a_i, \lambda) &= -a_i^T x + b_i + \lambda^T (C_i a_i - d_i) \\ g(\lambda) &= \begin{cases} b_i - \lambda^T d_i, & -x + C_i^T \lambda = 0 \\ \infty, & \text{otherwise} \end{cases} \end{aligned}$$

Hence the optimum is $q^* = b_i - x^T C_i^+ d_i = p^*$ if the problem is feasible, that is, when $P_i \neq \emptyset$, which is assumed. Similarly

$$\sup_{C_i a_i \preceq d_i} -a_i^T x + b_i = -b_i + x^T C_i^+ d_i$$

The problem can now be formulated

$$\begin{array}{ll} \text{minimize} & t^T t \\ \text{subject to} & -b_i + x^T C_i^+ d_i \leq t_i \text{ for } i = 1, \cdots, m \\ & b_i - x^T C_i^+ d_i \leq t_i \text{ for } i = 1, \cdots, m \end{array}$$

□