Q1

Consider optimization ~~prob~~ problem $\min f(x)$ where $Ax = b$ with f being a convex function and $A \in \mathbb{R}^{m \times n}$.
Point $x \in \mathbb{R}^n$ is optimal to $\min f(x)$ iff. It is feasible and $\exists \mu \in \mathbb{R}^m$ s.t. $\nabla f(x) = A^T \mu$.
Since this is a convex problem, by optimality condition, a feasible x is optimal iff $\nabla f^T(x)(y-x) \geq 0$ for $\forall y$ s.t. $Ay = b$.
Any y with $Ay = b$ can be written as $y = x + v$ where v is a point on the null space of A, or $Av = 0$, thus a feasible x is optimal if and only if $\nabla f^T(x) v \leq 0$.
Hence the optimality condition gives $\nabla f^T(x) v = 0$ for $\forall v$ s.t. $Av = 0$.
This means that $\nabla f^T(x) v = 0$ for $\forall v$ s.t. $Av = 0$.
So $\nabla f(x)$ is the orthogonal component of null space of A, which equals to the row space of A as equivalent column space of A^T.
Thus $\exists \mu \in \mathbb{R}^m$ s.t. $\nabla f(x) = A^T \mu$.

Q2

Since $f(x)$ is a strictly convex function, $f(Ax)$ is a convex function.
Since sum of convex functions is convex, $h(x)$ is convex function.
For any function f, define level set at α as $L_\alpha(f) = \{x \in dom(f) \mid f(x) \le \alpha\}$.
For $\forall$ convex function f, level set at α is convex because
$x, y \in L_\alpha(f)$, for $\forall \lambda \in [0,1]$,
$f(\lambda x + (1-\lambda) y) \le \lambda f(x) + (1-\lambda) f(y) \le \lambda \alpha + (1-\lambda)\alpha = \alpha$.
Let p denote the optimal value of a convex optimization problem with objective function f.
The set of optimal solutions to this problem is $L_p(f)$, so it's convex.
Thus, the set of all optimal solutions to a convex minimization problem is a convex set.
For the given problem, x^* & y^* are given solutions.
So $\frac{x^*+y^*}{2}$ is also a solution.
Suppose $Ax^* \neq Ay^*$,
$f(A(\frac{x^*+y^*}{2})) + g(\frac{x^*+y^*}{2}) < \frac{f(Ax^*)+f(Ay^*)}{2} + \frac{g(x^*)+g(y^*)}{2}$
Since g is convex and f is strictly convex.
So $h(\frac{x^*+y^*}{2}) < P$ where P is the optimal value of function h.
Contradiction.
So $Ax^* = Ay^*$.

Q3

$H_\mu(x)$ is continuous at $\|x\| = \mu$ since $\frac{\mu^2}{2\mu} = \mu - \frac{\mu}{2} = \frac{\mu}{2}$.

$\nabla\left(\frac{\|x\|^2}{2\mu}\right) = \frac{1}{2\mu} \cdot 2\|x\| = \frac{\|x\|}{\mu}$, $x = (x_1, x_2, \dots, x_n)$, $f(x) = \|x\| = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$

$$\nabla\left(\|x\| - \frac{\mu}{2}\right) = \nabla(\|x\|)$$
$$= \left(\partial/\partial x_1 f, \partial/\partial x_2 f, \dots, \partial/\partial x_n f\right)$$
$$= \left[\tfrac{1}{2}(x_1^2 + x_2^2 + \cdots + x_n^2)^{-\frac{1}{2}} \cdot 2x_1, \tfrac{1}{2}(x_1^2 + x_2^2 + \cdots + x_n^2)^{-\frac{1}{2}} \cdot 2x_2, \dots, \tfrac{1}{2}(x_1^2 + x_2^2 + \cdots + x_n^2)^{-\frac{1}{2}} \cdot 2x_n\right]$$
$$= \left(\frac{1}{\|x\|}x_1, \frac{1}{\|x\|}x_2, \dots, \frac{1}{\|x\|}x_n\right)$$
$$= \frac{x}{\|x\|}.$$

At $\|x\| = \mu$, $\frac{\|x\|}{\mu} = 1 = \left\|\frac{x}{\|x\|}\right\|$, so $\nabla H_\mu(x)$ is continuous. Thus $H_\mu \in C^{1,1}(\mathbb{R}^n)$.

To show Lipshitz continuous with $L = \frac{1}{\mu}$:

Case 1: For x, y, $\|x\|, \|y\| \le \mu$,

$$\|\nabla H_\mu(x) - \nabla H_\mu(y)\| = \left\|\frac{\|x\|}{\mu} - \frac{\|y\|}{\mu}\right\| = \left\|\frac{\|x\| - \|y\|}{\mu}\right\| \le \frac{1}{\mu}\|x - y\|$$

Case 2: For x & y, $\|x\|$ & $\|y\| > \mu$

$$\|\nabla H_\mu(x) - \nabla H_\mu(y)\| = \left\|\frac{x}{\|x\|} - \frac{y}{\|y\|}\right\| = \|1 - 1\| = 0 \le \frac{1}{\mu}\|x - y\|$$

since $\left\|\frac{x}{\|x\|}\right\| = 1$.

Case 3: For x & y, $\|x\| \le \mu$, $\|y\| \ge \mu$.

$$\|\nabla H_\mu(x) - \nabla H_\mu(y)\| = \left\|\frac{\|x\|}{\mu} - \frac{y}{\|y\|}\right\| = \left\|\frac{\|x\|}{\mu} - 1\right\| = \frac{1}{\mu}\|\|x\| - \mu\|$$
$$= \frac{1}{\mu}(\mu - \|x\|) \text{ since } \|x\| \le \mu$$
$$\le \frac{1}{\mu}(\|y\| - \|x\|) \text{ since } \|y\| > \mu$$
$$\le \frac{1}{\mu}\|y - x\| = \frac{1}{\mu}\|x - y\|$$

Case 4: For x, y, $\|x\| > \mu$, $\|y\| \le \mu$, same as case 3.

Thus $\|\nabla H_\mu(x) - \nabla H_\mu(y)\| \le \frac{1}{\mu}\|x - y\|$, $H_\mu \in C^{1,1}_{1/\mu}(\mathbb{R}^n)$.

Q4

(i) $f(x) = x_1^2 + 2x_1x_2 + 2x_2^2 + x_3^2 + 3x_1 - 4x_2$

$= (x_1+x_2)^2 + x_2^2 + x_3^2 + 3x_1 - 4x_2$

$= (x_1+x_2)^2 + (x_2-2)^2 + x_3^2 + 3|x_1| - 4$

$= (A_1x+b_1)^2 + (A_2x+b_2)^2 + (A_3x+b_3)^2 + 3\cdot|A_4x| - 4$

All terms are either affine or positive scalar times 1-norm or a constant, so they are convex. Since sum perserves convexity, $f(x)$ is convex.

$g_1 = \sqrt{2x_1^2 + x_1x_2 + 4x_2^2 + 4} + \frac{(x_1 - x_2 + x_3 + 1)^2}{x_1+x_2} - 6$

$= \sqrt{(x_1^2 + x_1x_2 + \frac{1}{4}x_2^2) + x_1^2 + \frac{15}{4}x_2^2 + 4} + \frac{\|x_1 - x_2 + x_3 + 1\|^2}{x_1+x_2} - 6$

$= \sqrt{(x_1^2 + x_1x_2 + \frac{1}{4}x_2^2) + x_1^2 + \frac{15}{4}x_2^2 + 4} + \frac{\|x_1 - x_2 + x_3 + 1\|^2}{x_1+x_2} - 6$

$= \|(x_1 + \frac{1}{2}x_2, x_1, \frac{\sqrt{15}}{2}x_1, 2)\| + \text{quad-over-lin}(x_1 - x_2 + x_3 + 1, x_1 + x_2) - 6$

Similar to $f(x)$, g_1 is convex since it's sum of convex terms.

$g_2(x) = (1\,0\,0)\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix} + \begin{pmatrix}1\\0\\0\end{pmatrix}$

$g_3(x) = (0\,1\,0)\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix} + \begin{pmatrix}0\\1\\0\end{pmatrix}$

$g_4(x) = (0\,0\,1)\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix} + \begin{pmatrix}0\\0\\1\end{pmatrix}$

$g_i(x) \le 0$, $x_i \ge 1$, so $g_i(x)$ is affine, or convex.

Thus objective and constraint functions are convex, so P is convex.

CVX gives $x^* = (1,1,1)^T$

Q4

(ii) $f(x) = |2x_1+3x_2+x_3| + x_1^2+x_2^2+x_3^2 + \sqrt{2x_1^2+4x_1x_2+7x_2^2+10x_2+6}$

$= \text{norm}([(2x_1+3x_2+x_3),1]) + \sum_{i=1}^{3} x_i^3 + \sqrt{2(x_1^2+2x_1x_2+x_2^2)+(5x_1^2+10x_2+5)+1}$

$= \text{norm}([(2x_1+3x_2+x_3),1]) + \sum_{i=1}^{3} x_i^3 + \text{norm}([\sqrt{2}(x_1+x_2), \sqrt{5}(x_2,1)); 1])$.

All terms convex, $f(x)$ convex.

$g_1(x) = {x_1^2+1}/{x_2} + 2x_1^2+5x_2^2+10x_3^2+4x_1x_2+2x_1x_3+2x_2x_3-7$

$= \text{quad-over-lin}(x_1,x_2) + \text{quad-over-lin}(1,x_2) + (x_1^2+4x_1x_2+4x_3)^2$

$+ (x_1^2+2x_1x_3+x_3^2) + (x_3^2+2x_2x_3+x_3^2) + 8x_3^2 - 7$

All terms convex, $g_1(x)$ convex.

$g_2(x) = (-1\ 0\ 0)(x_1\ x_2\ x_3)^T$, affine, $g_2(x) \le 0$, $x_1 \ge 0$

$g_3(x) = (0\ -1\ 0)(x_1\ x_2\ x_3)^T + \begin{pmatrix}0\\1\\0\end{pmatrix}$, affine, $g_3(x) \le 0$, $x_2 \ge 1$.

$2x_1^2+4x_1x_2+7x_2^2+10x_2+6$

$= 2(x_1^2+2x_1x_2+x_2^2)+(5x_2^2+10x_2+5)+1$

$= \sqrt{2}(x_1+x_2)^2 + \sqrt{5}(x_2+1)^2+1$

≥ 0.

Thus all objective and constraint functions are convex, so P is convex.

CVX gives $x^* = (0, 1, -0.4317)^T$

Q4

(iii) $f(x) = \dfrac{x_1^4 + 2x_1^2 x_2^2 + x_2^4}{x_1^2 + 2x_1x_2 + x_2^2} + \sqrt{x_3^2+1}$

$= \dfrac{(x_1^2 + x_2^2)^2}{(x_1+x_2)^2} + \text{norm}([x_3; 1])$

$=$ square_pos (quad_over_line([x_1; x_2], x_1+x_2)) + norm([x_3; 1])

All term are convex, $f(x)$ is convex.

$g_1(x) = x_1^2 + x_2^2 + 2x_3^2 + 2x_1x_2 + 2x_1x_3 + 2x_2x_3 - 100$

$= (x_1\ x_2\ x_3)\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} - 100$

$= x^T Q x - 100$ where $Q \succeq 0$.

$g_1(x)$ is convex.

$h_1(x) = (1\ 1\ 1)\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} - 2$, affine, $h_1(x) = 0$ then $x_1 + x_2 + x_3 = 2$.

$g_2(x) = (-1\ {-1}\ 0)(x_1\ x_2\ x_3)^T + 1$, affine, $g_2(x) \le 0$ so $x_1 + x_2 \ge 1$.

So all objective and constraints are convex, so P is convex.

CVX gives $x^* = (0.6125, 0.6125, 0.7749)^T$

Q4

(iv) $f(x) = \frac{x_1^4}{x_2^2} + x_2^4/x_1^2 + 2x_1x_2 + |x_1+5| + |x_2+5| + |x_3+5|$

$= (x_1^2/x_2)^2 + (x_2^2/x_1)^2 + 2x_1x_2 + \text{norm}(x_1+5,1) + \text{norm}(x_2+5,1) + \text{norm}(x_3+5,1)$

$= (x_1^2/x_2 + x_2^2/x_1)^2 + \text{norm}(x_1+5,1) + \text{norm}(x_2+5,1) + \text{norm}(x_3+5,1)$

$=$ square_pos (quad_over_lin(x_1,x_2) + quad_over_lin(x_2,x_1)) +

norm$(x+5,1)$ + norm$(x_2+5,1)$ + norm$(x_3+5,1)$

All terms are convex, $f(x)$ is convex.

$g_1(x) = ((x_1^2+x_2^2+x_3^2+1)^2+1)^2 + x_1^4 + x_2^4 + x_3^4 - 200$

$=$ square_pos(square_pos(sum_square(x), 1), 1) $+ x_1^4 + x_2^4 + x_3^4 - 200$

All terms convex, $g_1(x)$ ~~are~~ convex.

$g_2(x) = \max\{x_1^2 + 4x_1x_2 + 9x_2^2, x_1, x_2\} - 40$

$= \max\{\max\{(x_1+2x_2)^2 + 5x_2^2, x_1\}, x_2\} - 40$

$g_2(x)$ convex since max is non-decreasing.

$g_3(x) = (-1\ \ 0\ \ 0)(x_1 x_2 x_3)^T + 1$, $g_3(x) \le 0$, $x_1 \ge 1$, affine.

$g_4(x) = (0\ -1\ \ 0)(x_1 x_2 x_3)^T + 1$, $g_4(x) \le 0$, $x_2 \ge 1$, affine

Thus all objective and constraint functions are convex, so P is convex.

CVX gives $x^* = (1\ \ 1\ \ -0.7833)^T$

Q5

$\Rightarrow$ Assume x^* is stationary point of (P).

Since f is conti. diff. on C, $C = \{x : a^T x = 1\} \in \mathbb{R}^n$.

$f(y) - f(x^*) \approx \nabla f(x^*)^T (y - x^*) \geq 0$ for y close to x & $x, y \in C$ since x^* is a stationary point. So x^* is a local minimum.

By KKT conditions theorem, $\exists \lambda \in \mathbb{R}$ s.t. $\nabla f(x^*) + \lambda a = 0$

$$\begin{cases} \partial/\partial x_1 f(x^*) + \lambda a_1 = 0 \\ \vdots \\ \partial/\partial x_n f(x^*) + \lambda a_n = 0 \end{cases}, \quad \begin{cases} \partial/\partial x_1 f(x^*) = -\lambda a_1 \\ \vdots \\ \partial/\partial x_n f(x^*) = -\lambda a_n \end{cases}$$

Since $a \in \mathbb{R}^n_{++}$, $a_i > 0$ $\forall i = 1 \cdots n$,

$$\begin{cases} \partial/\partial x_1 f(x^*)/a_1 = -\lambda \\ \vdots \\ \partial/\partial x_n f(x^*)/a_n = -\lambda \end{cases}, \text{ or } \frac{\partial/\partial x_1 f(x^*)}{a_1} = \cdots = \frac{\partial/\partial x_n f(x^*)}{a_n}$$

$\Leftarrow$ Assume $\dfrac{\partial/\partial x_1 f(x^*)}{a_1} = \cdots = \dfrac{\partial/\partial x_n f(x^*)}{a_n} = \lambda$

Then $\nabla f(x^*) = \lambda \cdot a$ and $\nabla f(x^*)^T (x - x^*) = \lambda a^T (x - x^*) = \lambda (1 - 1) = 0$

So $\nabla f(x^*)^T (x - x^*) = 0$ $\forall x \in C$ where $C = \{x : a^T x = 1\}$

Thus $\nabla f(x^*)^T (x - x^*) \geq 0$ and x^* is a stationary point.

Q6

(i) $f = 2x_1^2 + 3x_2^2 + 4x_3^2 + 2x_1x_2 - 2x_1x_3 - 8x_1 - 4x_2 - 2x_3$

$$\begin{cases} d/dx_1 \, f = 4x_1 + 2x_2 - 2x_3 - 8 = 0 \\ d/dx_2 \, f = 6x_2 + 2x_1 - 4 = 0 \\ d/dx_3 \, f = 8x_3 - 2x_1 - 2 = 0 \end{cases}$$

solve get $(\frac{43}{17}, -\frac{3}{17}, \frac{15}{17})$, not feasible, x_2 must ≥ 0.

so plug in $x_2 = 0$ and solve again get

$(\frac{17}{7}, 0, \frac{6}{7}) = \bar{x}$

Thus $\bar{x}$ is an optimal solution.

Q6

(ii) Since $P_C(x) = P_{\mathbb{R}^3_+}(x) = [x]_+ = ([x_1]_+, [x_2]_+, [x_3]_+)$
where $[x_i]_+ = x_i$ if $x_i \geq 0$ and 0 if $x_i < 0$.
Since $\nabla f(\vec{x}) = 2Q\vec{x} + 2b$, $|\max(q_i)| = 4$

$$\|\nabla f(\vec{x}) - \nabla f(\vec{y})\| = \|2Q\vec{x} - 2Q\vec{y}\|$$
$$= \|2Q(\vec{x} - \vec{y})\|$$
$$\leq 8\|\vec{x} - \vec{y}\|$$

So $L = 8$ and $t = \frac{1}{8}$.
The 24th iteration gives the optimal solution
$f_x = -10.5714$ and $x = (2.4286, 0, 0.8571)$.