1 Mastery set [25 point]

A. Since each $\partial(a_i^T x + b_i) = \{a_i\}$,
$$\partial f(x) = \text{conv}\left(\{a_i : a_i^T x + B = f(x)\}\right).$$
Since each $\partial f_i(A_i x) = A_i^T \partial f_i(A_i x)$,
$$\partial f(x) = \text{conv} \bigcup_{i : f_i(A_i x) = f(x)} A_i^T \partial f_i(A_i x).$$
Since each $\partial \|x\|_{p_i} = \left\{\text{argmax}_{\|y\|_{q_i}=1} y^T x\right\}$ (where q_i is the Hölder conjugate of p_i), each $\partial \|A_i x\|_{p_i} = \left\{A_i^T \text{argmax}_{\|y\|_{q_i}=1} y^T A_i x\right\}$, so
$$\partial f(x) = \text{conv} \bigcup_{i : \|A_i x\|_{p_i} = f(x)} \left\{A_i^T \operatorname*{argmax}_{\|y\|_{q_i}=1} y^T A_i x\right\}.$$

B. This is true under any strictly convex norm $\|\cdot\|$ (including $\|\cdot\|_p, p \in (1,\infty)$), since the projection is
$$\operatorname*{argmin}_{y \in C} \|x - y\|,$$
which is a strictly convex optimization problem and hence has a unique solution.

In the case that C is nonconvex, the projection need not be unique. Consider, for instance, $C = \{-1, 1\} \subseteq \mathbb{R}$. Then, $\|0 - (-1)\| = \|0 - 1\|$, and so the projection is not unique.

C. In general the projection operator onto a convex set need not be differentiable. Consider, for instance, $C = (-\infty, 0] \subseteq \mathbb{R}$. C is clearly convex, but the projection operator is
$$P_C(x) = \begin{cases} x & \text{if } x \leq 0 \\ 0 & \text{else} \end{cases},$$
which is not differentiable at $x = 0$.

D. A linear program is always convex. The objective function is linear, and hence convex. The constraint set is an intersection of half-planes, which are always convex, and hence the constraint set is convex.

E. Let
$$\begin{aligned} f_1(x) &:= ax + b, \\ f_2(x) &:= e^x, \\ f_3(x) &:= x \log x, \\ f_4(x) &:= \|x\|^2/2. \end{aligned}$$
Clearly,
$$f_1^*(y) = \max_x xy - ax - b = \max_x (y - a)x - b = \begin{cases} b & \text{if } y = a \\ +\infty & \text{else} \end{cases}.$$
If
$$0 = \frac{d}{dx} xy - e^x = y - e^x,$$
then, for $y > 0$, $x = \log y$. Thus
$$f_2^*(y) = \max_x xy - e^x = \begin{cases} +\infty & \text{if } y < 0 \\ 0 & \text{if } y = 0 \\ y \log y - y & \text{if } y > 0 \end{cases}.$$
If, for $x > 0$,
$$0 = \frac{d}{dx} xy - x \log x = y - \log x - 1,$$
then, $x = e^{y-1}$. Thus,
$$f_3^*(y) = \max_x xy - x \log x = ye^{y-1} - (y-1)e^{y-1} = e^{y-1}.$$

If x maximizes $x^T y - \|x\|^2/2$, then, letting $c := \|x\|$, x maximizes $x^T y - c^2/2$, and hence $x^T y = c\|y\|_*$, where $\|\cdot\|_*$ denotes the dual norm of $\|\cdot\|$. Thus,
$$f_4^*(y) = \max_{c \in \mathbb{R}} c\|y\|_* - c^2/2.$$
If
$$0 = \frac{d}{dc} c\|y\|_* - c^2/2 = \|y\|_* - c,$$
then $c = \|y\|_*$, and so
$$f_4^*(y) = \|y\|_*^2/2.$$