1 (a) because f is covnex and

$$\frac{b-x}{b-a}+\frac{x-a}{b-a}=1$$

so we have

$$\frac{b-x}{b-a}f(a)+\frac{x-a}{b-a}f(b)\geq f(\frac{b-x}{b-a}a+\frac{x-a}{b-a}b)=f(\frac{ab-ax+bx-ab}{b-a})=f(x)$$

(b) because f is covnex and $x\in(a,b)$,there must exist $\theta(0\leq\theta\ \leq 1)$ so that

$$x=\theta a+(1-\theta)b$$

for the first inequality,we have

$$\frac{f(x)-f(a)}{x-a}=\frac{f(\theta a+(1-\theta)b)-f(a)}{\theta a+(1-\theta)b-a}\leq\frac{\theta f(a)+(1-\theta)f(b)-f(a)}{(1-\theta)(b-a)}=\frac{f(b)-f(a)}{b-a}$$

so we proved

$$\frac{f(x)-f(a)}{x-a}\leq\frac{f(b)-f(a)}{b-a}$$

the second ineequality is same as the first one

$$\frac{f(b)-f(x)}{b-x}=\frac{f(b)-f(\theta a+(1-\theta)b)}{b-\theta a-(1-\theta)b}\geq\frac{f(b)-\theta f(a)-(1-\theta)f(b)}{\theta(b-a)}=\frac{f(b)-f(a)}{b-a}$$

2 (a) suppose f is a piecewise function,it's expression is

$$f=\begin{cases}-x^2,x\leq 0\\ x^2,x>0\end{cases}$$

obviouslly,f is not convex,then let $C=[1,2]$,so C is a convex set,the expression of f on C is

$$f=x^2,x\in[1,2]$$

so f is convex on C

(b) let f be a piecewise function and it's expression is

$$f=\begin{cases}-x^2,x\leq 0\\ x^2,x>0\end{cases}$$

as we can see,f is a non-convex function,now let $C=[1,2)\cup(2,3]$,then
C is a non-convex set too.$\forall x,y\in C$,we always have

$$f(y)\geq f(x)+\nabla f(x)^T(y-x)$$

so f is convex on C.

3 suppose$X_1, X_2 \in X$,$0 \leq \theta \leq 1$,then we need to prove that $\theta X_1 + (1-\theta)X_2 \in X$ because X is the global minimizers of f,so we have

$$f(X_1) = f(X_2) \leq f(u), \forall u \in R^n$$

because f is convex,so

$$f(\theta X_1 + (1-\theta)X_2)) \leq \theta f(X_1) + (1-\theta)f(X_2) = \theta f(X_1) + (1-\theta)f(X_1) = f(X_1)$$

because $f(X_1) \leq f(u), \forall u \in R^n$,so we have

$$f(\theta X_1 + (1-\theta)X_2)) = f(X_1) \leq f(u), \forall u \in R^n$$

so $\theta X_1 + (1-\theta)X_2 \in X$,and we have proved X is convex.

4 because $-log(x)$ is convex,so

$$-log(\frac{x_1 + x_2 + \cdots + x_m}{m}) \leq -\frac{1}{m}(log(x_1) + log(x_2) + \cdots + log(x_m)) = -\frac{1}{m}log(x_1 \cdots x_m)$$

multiply by -1 and take the exponential of both sides yields and we can get

$$\sqrt[m]{x_1 \cdots x_m} \leq \frac{x_1 + x_2 + \cdots + x_m}{m}$$

For the only m, there is only one set of $x_1, x_2 \cdots, x_m$ that can holds the equation and it occurs only if $x_1 = x_2 = \cdots = x_m$

5 let $f(x) = \underset{y \in S}{inf}||x - y||$,suppose $x_1, x_2 \in S, \theta \in [0,1]$,$\forall x \in R^n$,we have

$$\begin{aligned} f(\theta x_1 + (1-\theta)x_2) &= \underset{y \in S}{inf}||\theta x_1 + (1-\theta)x_2 - y|| \\ &\leq \underset{y \in S}{inf}(\theta||x_1 - y|| + (1-\theta)||x_2 - y||) \\ &\leq \theta||x_1 - y|| + (1-\theta)||x_2 - y||) \\ &= \theta f(x_1) + (1-\theta)f(x_2) \end{aligned}$$

so $dist(x, S)$ is convex.

6 let $x, y \in R^n$,so we have

$$\begin{aligned} (\theta f + (1-\theta)g)^* &= sup\{xy - (\theta f(x) + (1-\theta)g(y))\} \\ &\leq \theta sup\{xy - f(x)\} + (1-\theta)sup\{xy - g(y)\} \\ &= \theta f^* + (1-\theta)g^* \end{aligned}$$

7 suppose $x \in C$,if $C \subset D$,then

$$S_C(x) = sup\{x^T y | \forall y \in C\}, S_D(x) = sup\{x^T y | \forall y \in D\}$$

$\forall (x, y) \in C$,we always have $x^T y \leq S_C(x)$,because $C \subset D$,so $x^T y \leq S_D(x)$,and we have

$S_C(x) \leq S_D(x)$.The appearance of '$<$ 'is due to the elements in D are more than C,so the minimum of $S_D(x)$ is $S_C(x)$.

conversely, for any point $x \in R^n$,$S_C(x)$ is the smallest among the set of supporting hyperplanes (supporting half-spaces) for subset C, where the hyperplanes contain x. Similarly, $S_D(x)$ is the smallest among the set of supporting hyperplanes for subset D that contain x.

Based on the given condition, we have $S_C(x) \leq S_D(x)$, which means that for any x, the supporting hyperplanes of C do not extend farther from the origin than the supporting hyperplanes of D, with the direction of the hyperplane's normal vector as an indicator. In other words, each point in C is located on the same side of the corresponding supporting hyperplanes in D or is closer to the origin. This is because the normal vectors of the supporting hyperplanes point outward, and therefore, points in C cannot be inside the supporting hyperplanes of D.

This implies that every point in C is in D because they all stay within D. Thus, according to the definition, C is a subset of D.

This completes the proof. Therefore, if $S_C(x) \leq S_D(x)$, then C is a subset of D.

8 (a) For each i, we have$f^*(y) = y^T x - f(x) = y_i x_i - \underset{i=1,2,\cdot,n}{max}\ x_i$. This is a linear function of x_i with slope y_i,the supremum of $y^T x - f(x)$ over x is achieved when the slope y_i is equal to the maximum slope among all $y_j, j = 1, \cdots, n$. That is, when $y_i = max\ y_j$,If max $y_j > 0$, then the supremum is infinite, since we can make x_i arbitrarily large. If max $y_j \leq 0$, then the supremum is zero, since we can choose $x_i = 0$,Therefore, the conjugate function is

$$f^*(y) = \begin{cases} 0, max\ y_j \leq 0 \\ \infty, max\ y_j > 0 \end{cases}$$

(b) we know that $f^*(y) = sup\{xy - f(x)\}$,let $xy - f(x) = 0$,we have

$$y = |x|^{p-1}$$

which mean $y \geq 0$,so $dom_f = \{y \geq 0\}$.so we can confirm that when $x = y^{\frac{1}{p-1}}$,$xy - f(x)$ reaches its maximum,so we have

$$f^*(y) = y^{\frac{p}{p-1}} - \frac{y^{\frac{p}{p-1}}}{p} = \frac{p-1}{p} y^{\frac{p}{p-1}} = \frac{y^q}{q}$$

because $y \geq 0$,so $f^*(x) = \frac{|x|^q}{q}$,and we proved that

$$(\frac{|x|^p}{p})^* = \frac{|x|^q}{q}$$

(c) let $(xy - f(x))' = 0$ and we can know that $x = -\sqrt{\frac{1}{y}}$,then we can get

$$f^*(y) = \begin{cases} -2\sqrt{-y}, y \leq 0 \\ +\infty, y > 0 \end{cases}$$

(d) let $(xy - f(x))' = 0$ and we can know that $x = -\frac{1}{y}$,then we can get

$$f^*(y) = \begin{cases} -ln(-y) - 1, y < 0 \\ +\infty, y > 0 \end{cases}$$

9 (a) transform the function $f(x)$ into Moreau-Yosida regularization form:

$$f_\mu(x) = inf\{||y||_1 + \frac{1}{2\mu}||y - x||^2\} + \frac{\alpha}{2}||x||^2$$

where μ is a positive constant. Then, we can calculate the conjugate function of $f_\mu(x)$

$$f_\mu^*(y) = \sup_x\{y^T x - f_\mu(x)\}$$

substituting $f_\mu(x)$ into the above equation,we get

$$f_\mu^*(y) = \sup_x\{y^T x - \inf_z[||z||_1 + \frac{1}{2\mu}||z - x||^2] - \frac{\alpha}{2}||x||^2\}$$

we can find the optimal solution by taking the derivative.$\forall i$,let $u_i = \frac{y_i}{\alpha}$,then we have

$$u_i = \begin{cases} -\frac{1}{2\mu}, y_i < -\alpha \\ [-\frac{1}{2\mu}, \frac{1}{2\mu}], -\alpha \leq y_j \leq \alpha \\ \frac{1}{2\mu}, y_i > \alpha \end{cases}$$

therefore,we have

$$f^*(y) = f_0^*(y) = \begin{cases} 0, ||y||_1 \leq \alpha \\ +\infty, ||y||_1 > \alpha \end{cases}$$