P1: ① f is convex ② $\sum_{i=1}^{m} \lambda_i = 1$ to prove $f(\sum_{i=1}^{m} \lambda_i x_i) \le \sum_{i=1}^{m} \lambda_i f(x_i)$.

When $m=1$, then $\overset{\lambda_i=1}{f(x) = f(x) * 1}$ proved.

when $m=2$ $f(\sum_{i=1}^{m} \lambda_i x_i) = f(\lambda_1 x_1 + \lambda_2 x_2) \le \lambda_1 f(x_1) + \lambda_2 f(x_2)$ (by definition of convex)

$$f(\lambda_1 x_1 + (1-\lambda_1) x_2) \le \lambda_1 f(x_1) + (1-\lambda_1) f(x_2)$$

when $m \ge 2$

suppose ~~that~~ $f(\sum_{i=1}^{m} \lambda_i x_i) \le \sum_{i=1}^{m} \lambda_i f(x_i)$ is ture.

then for $m+1=k$ ~~$f(\sum_{i=1}^{k} \lambda_i x_i) \le \sum_{i=1}^{m} \lambda f(x_i)$ is still ture. (we need to proff)~~

$$f(\sum_{i=1}^{k} \lambda_i x_i) = f((1-\lambda_k) \sum_{i=1}^{m} \frac{\lambda_i x_i}{1-\lambda_k} + \lambda_k x_k) \le (1-\lambda_k) f(\sum_{i=1}^{k-1} \frac{\lambda_i x_i}{1-\lambda_k}) + \lambda_k f(x_k)$$

with that $(1-\lambda_k) + \lambda_k = 1$

$$(1-\lambda_k) f(\sum_{i=1}^{k-1} \frac{\lambda_i x_i}{1-\lambda_k}) + \lambda_k f(x_k) = (1-\lambda_k) f(\sum_{i=1}^{m} \frac{\lambda_i x_i}{1-\lambda_k}) + \lambda_k f(x_k)$$

$$(1-\lambda_k) f(\sum_{i=1}^{m} \frac{\lambda_i}{1-\lambda_k} x_i) + \lambda_k f(x_k) \le (1-\lambda_k) \sum_{i=1}^{m} \frac{\lambda_i}{1-\lambda_k} f(x_i) + \lambda_k f(x_k)$$

$$= \sum_{i=1}^{m} \lambda_i f(x_i) + \lambda_k f(x_k)$$

then $f(\sum_{i=1}^{k} \lambda_i x_i) \le \sum_{i=1}^{k} \lambda_i f(x_i)$ get proved. $= \sum_{i=1}^{k} \lambda_i f(x_i)$

P2. (i) ~~for λ_i in dom~~

choose any two point x and y. from $\bigcap_{i=1}^{m} \mathrm{dom} f_i$ and set $\theta \in (0,1]$

then for Any f_i, ~~we in dom f_i~~ (convex)

We have. $f_i(\theta x + (1-\theta) y) \le \theta f_i(x) + (1-\theta) f_i(y)$ and (as $\lambda_i \in R_+$) $\lambda_i f_i(\theta x + (1-\theta) y) \le \lambda_i \theta f_i(x) + (1-\theta)\lambda_i f_i(y)$

the. we summarize all the function

$$\sum_{i=1}^{m} \lambda_i f_i(\theta x + (1-\theta) y) \le \sum_{i=1}^{m} \lambda_i \theta f_i(x) + \sum_{i=1}^{m} (1-\theta) \lambda_i f_i(y) \qquad f = \sum_{i=1}^{m} \lambda_i f_i$$

$$f(\theta x + (1-\theta) y) \le \theta f(x) + (1-\theta) f(y) \Rightarrow f \text{ is convex on } \mathrm{dom}(f) := \bigcap_{i=1}^{m} \mathrm{dom} f_i$$

P2 (ii) as f is convex. $g(x) = AX + b$ is an affine function

so that for $g(x_1+x_2)$ we have $g(x_1+x_2) = A(x_1+x_2)+b$.

So let set $x_0 = \theta x + (1-\theta)y$ $(\theta \in (0,1])$ then $g(x_0) = g(\theta x + (1-\theta)y) = A(\theta x + (1-\theta)y) + b$

$g(x_0) = \theta(Ax+b) + (1-\theta)(Ay+b) = \theta g(x) + (1-\theta)g(y)$

so for g we have $g(\theta x + (1-\theta)y) = \theta g(x) + (1-\theta)g(y)$

as f is convex. $g(x)$ and $g(y) \in \text{dom}(f)$ for $\theta \in [0,1]$

we have $f(\theta g(x) + (1-\theta)g(y)) \le \theta f(g(x)) + (1-\theta)f(g(y))$ set $f \circ g = h$

$$\begin{aligned} f(\theta g(x) + (1-\theta)g(y)) &\le \theta f(g(x)) + (1-\theta)f(g(y)) \\ \parallel \quad & \\ f(g(\theta x + (1-\theta)y)) &\le \theta f(g(x)) + (1-\theta)f(g(y)) \\ h(\theta x + (1-\theta)y) &\le \theta h(x) + (1-\theta)f(y) \end{aligned}$$

$\Rightarrow h(x)$ is convex $\Rightarrow f \circ g$ is convex.

P3 $f(x) = x^T Q x + b^T x + c$ Q is symmetrix matrix, $f(x)$ smooth with $2\|Q\|$

for smoothness. $\Leftrightarrow \|\nabla f(x) - \nabla f(y)\|_* \le L\|x-y\|$ for all $x, y \in D$.

P3 c $f(x) = x^T Q x + b^T x + c$ Q is symmetrix matrix

we need to prove that $f(x)$ is smooth with $2\|Q\|$

$\Leftrightarrow$ $\|\nabla f(y) - \nabla f(x)\| \le L\|y-x\|_2$ $\forall y, z \in R^n$ $L = 2\|Q\|$

$\nabla f(x) = 2Qx + b$ $\|\nabla f(y) - \nabla f(x)\| = \|2Q(y-x)\| = 2\|Q(y-x)\| \le \underbrace{2 \cdot \|Q\|}_{L}\|(y-x)\|$

So $f(x)$ is smooth with $\|2Q\| = L$.

P4 projected gradient decent
$$y_{t+1} = x_t - \eta \nabla f(x_t)$$
$$x_{t+1} = \Pi_X(y_{t+1}) \qquad (\eta > 0)$$

f is convex-differentiable $f: X \to R$ for some $t \ge 0$ $x_{t+1} = x_t$

to prove that x_t is a minimizer of f over closed-convex f.

for $x \in X$ we have $\langle \Pi_X(y) - y, x - \Pi_X(y) \rangle \ge 0$

then when $x_{t+1} = x_t = \Pi_X(y_{t+1})$ $x_t = y_{t+1} + \eta \nabla f(x_t) \Leftrightarrow y_{t+1} = x_t - \eta \nabla f(x_t)$

$\Downarrow$

$x_t = \Pi_X(y_{t+1}) = \Pi_X(x_t - \eta \nabla f(x_t))$ and $y = x_t - \eta \nabla f(x_t)$, $\Pi_X(y) = x_t$

then for all $x \in X$ we have $\langle x_t - (x_t - \eta \nabla f(x_t)), (x - x_t) \rangle \ge 0$

$\Leftrightarrow \langle \nabla f(x_t), (x - x_t) \rangle \ge 0$ as f is convex so x_t is a minimal of X