[Problem 2.14]

Proof of thm 2.2:

IF: assume $g(t)=f(x+tv)$ is convex for $\forall x \in \text{dom} f$, $v \in \mathbb{R}^n$, then for $\forall \theta \in [0,1]$ and $\forall x, y \in \text{dom} f$, $\forall t_1\ t_2 \in \mathbb{R}$, we have

$$\begin{cases} u = (t_2 x - t_1 y) / (t_2 - t_1) \\ v = (x-y)/(t_1 - t_2) \end{cases}$$

then we have

$$g(\theta t_1 + (1-\theta)t_2) = f\left[u + (\theta t_1 + (1-\theta)t_2)v\right]$$
$$= f(\theta x + (1-\theta)y)$$
$$\leq \theta g(t_1) + (1-\theta) g(t_2)$$
$$= \theta f(u+t_1 v) + (1-\theta) f(u + t_2 v)$$
$$= \theta f(x) + (1-\theta) f(y)$$

so by definition f is convex.

ONLY IF: suppose now $f(x)$ is convex, then for $\forall \theta \in [0,1]$ and $\forall\ t_1\ t_2 \in \mathbb{R}$

$$g(\theta t_1 + (1-\theta)t_2) = f(x + [\theta t_1 + (1-\theta)t_2]v)$$

$$= f[(\theta + (1-\theta))x + \theta t_1 v + (1-\theta)t_2 v]$$

$$\leq f[\theta(x + t_1 v) + (1-\theta)(x + t_2 v)]$$

$$\leq \theta f(x + t_1 v) + (1-\theta) f(x + t_2 v)$$

$$= \theta g(t_1) + (1-\theta) g(t_2)$$

so by definition $g(t)$ is convex for $\forall x \in \operatorname{dom} f$

Proof of thm 2.3:

ONLY IF: suppose $f(x)$ is convex,

then for $\forall\, \theta \in (0,1]$

$$\theta f(y) + (1-\theta) f(x) \geq f(\theta y + (1-\theta)x)$$

$$f(y) \geq \frac{f(\theta y + (1-\theta)x) - (1-\theta) f(x)}{\theta}$$

$$= f(x) + \frac{f(\theta y + (1-\theta)x) - f(x)}{\theta}$$

$$= f(x) + \frac{1}{\theta} \nabla f(\xi)^T \theta (y-x)$$

for $\xi \in [x, (1-\theta)x + \theta y)]$ for $\forall\, \theta \in [0,1]$

thus $\xi = x$ and we have

$$f(y) \geq f(x) + \nabla f^T(x)(y-x) \text{ for } \forall\, x, y \in \operatorname{dom} f$$

IF: assume for $\forall x, y \in \text{dom} f$ we have

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

then for $\forall \theta \in [0,1]$ we have

$$f(\theta y + (1-\theta)x) \geq f(\theta' y + (1-\theta')x) + \nabla f(\theta' y + (1-\theta')x)^T (y-x)(\theta - \theta')$$

using thm 2.2, this means

$$g(\theta) \geq g(\theta') + g'(\theta')(\theta - \theta')$$

$\Rightarrow$ for $\forall t, \theta$, $l = \lambda t + (1-\lambda)\theta$ we have

$$g(t) \geq g(l) + g'(l)(t - l)$$

$$g(\theta) \geq g(l) + g'(l)(\theta - l)$$

$$\Rightarrow \lambda g(t) + (1-\lambda) g(\theta) \geq g(l) = g(\lambda t + (1-\lambda)\theta)$$

So $g(\cdot)$ is convex, according to thm 2.2

$f(\cdot)$ is also convex.

<u>Proof of thm 2.4:</u>

We first prove the 1-dimensional case:

If f is convex, then for $\forall y > x \in \mathrm{dom} f$

$$f(y) \geq f(x) + f'(x)(y-x)$$
$$f(x) \geq f(y) + f'(y)(x-y)$$

$$\Rightarrow \quad f'(x)(y-x) \leq f(y) - f(x) \leq f'(y)(y-x)$$

$$\Rightarrow \quad \frac{f'(y) - f'(x)}{y-x} \geq 0$$

Let $y \to x$ we have $f''(x) \geq 0$.

Now suppose $f''(x) \geq 0$, then

$$f(y) = f(x) + f'(x)(y-x) + \frac{1}{2} f''(z)(y-x)^2$$
$$\geq f(x) + f'(x)(y-x)$$

$\Rightarrow$ $f(x)$ is convex.

Now consider multi-dimension case, f is convex if and only if $g(t) = f(x+ty)$ is convex and

$$g''(t) = y^T \nabla^2 f(x+ty) y$$

so $g''(t) > 0 \iff \nabla^2 f(x+ty) \succeq 0$ $\square$.

[Problem 2.15]

$$\begin{pmatrix} P & A^T \\ A & 0 \end{pmatrix} \begin{pmatrix} x^* \\ v^* \end{pmatrix} = \begin{pmatrix} -q \\ b \end{pmatrix}$$

If P and $AP^{-1}A^T$ are invertible, then

$$\begin{pmatrix} P & A^T \\ A & 0 \end{pmatrix}^{-1} = \begin{pmatrix} P^{-1} - P^{-1}A^T(AP^{-1}A^T)^{-1}AP^{-1}, & P^{-1}A^T(AP^{-1}A^T)^{-1} \\ (AP^{-1}A^T)^{-1}AP^{-1}, & -AP^{-1}A^T \end{pmatrix}$$

so $x^* = q\left[-P^{-1} + P^{-1}A^T(AP^{-1}A^T)^{-1}AP^{-1}\right] + bP^{-1}A^T(AP^{-1}A^T)^{-1}$

If null(P) is non-trival, P is not invertible and the solution does not exists if

$$\begin{pmatrix} -q \\ b \end{pmatrix} \in \operatorname{span} \begin{pmatrix} P & A^T \\ A & 0 \end{pmatrix}$$

otherwise the solution doesn't exist.

[Problem 2.16]

From the dual problem we know that

$$p^*(u, v) \geq p^*(0,0) - \lambda^{*T} u - \nu^{*T} v$$

then let $v=0$, $u = te_i$, we have

$$\frac{p^*(te_i, 0) - p^*(0,0)}{t} \geq -\lambda_i^* \quad \text{for } t>0$$

so as $t \to 0$ we have

$$\frac{\partial p^*(0,0)}{\partial u_i} \geq -\lambda_i^*$$

for $t<0$ we get $\frac{\partial p^*(0,0)}{\partial u_i} \leq -\lambda_i^*$, so we have

$$\frac{\partial p^*(0,0)}{\partial u_i} = -\lambda_i^*$$

and the same argument can be applied to

$$\frac{\partial p^*(0,0)}{\partial v_i} = -\nu_i^*$$