(P) $\min_{w} \frac{1}{2}\|Xw-y\|_2^2 + \lambda\|w\|_1$ λ fixed, >0

$\iff \min_{r,w} \frac{1}{2}\|r\|_2^2 + \lambda\|w\|_1$

s.t $r = Xw - y$

The Lagrangian is: $\mathcal{L}(w,r,\mu) = \frac{1}{2}\|r\|_2^2 + \lambda\|w\|_1 + \mu^T(r - Xw + y)$

$= \frac{1}{2}r^Tr + \mu^Tr + \lambda\|w\|_1 - \mu^TXw + \mu^Ty$

$g(\mu) = \inf_{r,w} \mathcal{L}(w,r,\mu) = \inf_{r} \underbrace{\frac{1}{2}r^Tr + \mu^Tr}_{h(r)} + \lambda \inf_{w}(\|w\|_1 - \frac{1}{\lambda}\mu^TXw) + \mu^Ty$

$= \inf_{r} h(r) - \lambda \sup_{w}\left(\frac{1}{\lambda}(X^T\mu)^Tw - \|w\|_1\right) + \mu^Ty$ $\lambda > 0$

h is convex, differentiable: $\nabla h(\tilde{r}) = 0 \iff \tilde{r} = -\mu$

$h_{min} = h(\tilde{r}) = \frac{1}{2}\mu^T\mu - \mu^T\mu = -\frac{1}{2}\mu^T\mu$

$g(\mu) = -\frac{1}{2}\mu^T\mu - \lambda\|\cdot\|_1^*\left(\frac{1}{\lambda}X^T\mu\right) + \mu^Ty$

See HW2, $\|\cdot\|_1^*(y) = \begin{cases} 0 & \text{if } \|y\|_\infty \leq 1 \\ +\infty & \text{otherwise.} \end{cases}$

Hence, (P) is equivalent to $\max_{\mu} -\frac{1}{2}\mu^T\mu + \mu^Ty$

st $\|\frac{1}{\lambda}X^T\mu\|_\infty \leq 1$

ie $\boxed{\begin{array}{ll} \min_{\mu} & \frac{1}{2}\mu^T\mu - \mu^Ty \\ \text{s.t} & \|\frac{1}{\lambda}X^T\mu\|_\infty \leq 1 \end{array}}$ (Dual)

However, $\|\frac{1}{\lambda}X^T\mu\|_\infty \leq 1 \iff \max_{i=1,\dots,d} |(X^T\mu)_i| \leq \lambda$

$\iff \begin{cases} \forall i \in [1;d], (X^T\mu)_i \leq \lambda \\ \forall i \in [d+1, 2d], (-X^T\mu)_i \leq \lambda \end{cases}$