

$$\begin{aligned} \min_{r, w} \quad & \frac{1}{2} \|r\|_2^2 + \lambda \|w\|_1 \\ \text{s.t.} \quad & r = Xw - y \end{aligned}$$

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

Thus, recall that, $g(\mu) = \inf_{w, r} \mathcal{L}(w, r, \mu)$

$$= \inf_w h(r) - \lambda \sup_w \left(\frac{1}{\lambda} (X^T \mu)^T w - \|w\|_1 \right) + \mu^T y$$

We've shown: $r^* = -\mu^*$ for the 1st part

For the 2nd part: $\sup_w \frac{1}{\lambda} (X^T \mu)^T w - \|w\|_1 = \left\| \frac{X^T \mu}{\lambda} \right\|_1^*$

When considering $\|\cdot\|_1^*$:

$$\begin{aligned} \bullet \text{ for any } y \in \mathbb{R}^m, \forall x \in \mathbb{R}^n, \quad y^T x - \|x\|_1 &= \sum_{i=1}^n y_i x_i - \sum_{i=1}^n |x_i| \\ &= \sum_{i=1}^n y_i \operatorname{sgn}(x_i) |x_i| - \sum_{i=1}^n |x_i| \\ &= \sum_{i=1}^n |x_i| (y_i \operatorname{sgn}(x_i) - 1) \end{aligned}$$

This quantity explodes whenever $y_i > 1$ or $-y_i < 1$
(if $\operatorname{sgn}(x_i) = +1$) (if $\operatorname{sgn}(x_i) = -1$)

So y needs to be s.t. $\|y\|_\infty \leq 1$

• Then $y^T x - \|x\|_1 \leq \sum_{i=1}^n |x_i| (|y_i| - 1)$ because $y_i \operatorname{sgn}(x_i) \leq |y_i|$

Equality holds when $\forall i \in [1, n], |x_i| (|y_i| - 1) = 0$

$$\text{ie } x_i = 0 \text{ or } |y_i| - 1 = 0$$

• In our case, w^* satisfies $w_i^* \left(\frac{|X^T \mu^*|_i}{\lambda} - 1 \right) = 0 \quad \forall i \in [1, d]$

$$\text{ie } \boxed{w_i^* \left(\left(\frac{X^T \mu^*}{\lambda} \right)_i - 1 \right) = 0} \text{ ie either } w_i^* = 0 \text{ or } \left(\frac{X^T \mu^*}{\lambda} \right)_i - 1 = 0$$