Problem 5.1

$$\min \sum_i (y_i - a - \sum_j x_{ij}\beta_j)^2$$

$$\text{s.t.} \quad \sum_j |\beta_j| \le t$$

(a) for optimal β_j^*, we take derivative of object function w.r.t. a should be equal to zero when $a = \hat{a}$

so $\sum_i (y_i - \hat{a} - \cancel{\beta}\sum_j x_{ij}\beta_j) = 0$

$$\sum_i y_i - n\hat{a} - \sum_i \sum_j x_{ij}\beta_j = 0$$

~~more obvis~~ divided by n, we have

$$\bar{y} - \hat{a} - \sum_j \frac{x_{ij}}{n}\beta_j = 0$$

since every column of X, which $x_{i1}, x_{i2}, \cdots x_{ik}$ [crossed out]) have been stadardized, so $E[x_{i1}] = E[x_{i2}] \cdots = 0$

$$\therefore \cancel{\bar{y} = \hat{a}} \quad \sum_j \frac{x_{ij}}{n}\beta_j = 0$$

$$\therefore \quad \hat{a} = \bar{y}$$

(b) find t_0 that constraint is binding if and only if $0 \le t \le t_0$.

constraints are binding means if change the constraints, the optimal solution changes together.

we know $\hat{\beta} = \cancel{(X^TX)^{-1}} X^+ Y$, so let $t_0 = |X^+ Y|$

if $t < t_0$, the $\hat{\beta}$ will change. if $t > t_0$, $\hat{\beta}$ will stay same.

And when $t < t_0$, $\sum_j |\hat{\beta}_j| = t$, since it is the ~~optimal~~ local ~~solution~~ optimal solution, thus global solution.

(c) if $d^* = p^*$, so

$$f_0(x^*) = g(\lambda^*, v^*) = \inf\left(f_0(x^*) + \sum_i \lambda_i^* f_i(x) + \sum_j v_j^* h_j(x)\right)$$

$$\Rightarrow \quad f_0(x^*) \le f_0(x^*) + \sum_i \lambda_i^* f_i(x) + \sum_j v_j^* h_j(x)$$

if only if $\lambda_i^* f_i(x) = 0$, the inequality can hold.

(d) Here, relative interior is $\sum_j |\beta_j| < t$,

which exactly satisfies Slater's condition

$\exists x \in \text{relint}(D)$, such that $f_i(x) < 0$, when $f_i(x) = \sum_j |\beta_j| - t$

and $Ax = b$ (No equality constraint here)

So, It is strong dual. (c) tells that strong duality must have <u>"Active Constraints"</u>

(e) This actubley L1 Regression (or Lasso Reg).

We can see this like the Lagrange Function of (5.1) when $t=0$ (which satisfies $t \geq 0$)

so (5.1), (5.2) actually have the same solution.

Problem 5.3

$\mathcal{L}(x,v) = x^TQ\cancel{A}x + v^T(Ax-b)$

$\|x\|_Q = \|x\|_Q^2$

$\mathcal{L}(x,v) = x^TQx + v^T(Ax-b)$

$\frac{\partial L}{\partial x} = 2Qx^* + A^Tv^* = 0 \qquad \frac{\partial L}{\partial v} = Ax^* - b = 0$

$x^* = A^{-1}b$

$\cancel{x^* = -1}$

$-2(A^T)^{-1}QA^{-1}b$

$x^* = -\frac{1}{2}Q^{-1}A^Tv \qquad v^* = -2\cancel{A}^TQx^* = \cancel{2A^T}$

$\inf_x \mathcal{L}(x,v) = \frac{1}{4}(Q^{-1}A^Tv)^TQ(Q^{-1}A^Tv) + v^T(-\frac{1}{2}AQ^{-1}A^Tv - b)$

$= \frac{1}{4}v^TA(Q^{-1})^TQQ^{-1}A^Tv \cancel{=} (-\frac{1}{2}v^TAQ^{-1}A^Tv - b^Tv)$

$= -\frac{1}{4}v^TAQ^{-1}A^Tv - b^Tv.$

$g(v)$

$\therefore \frac{\partial g}{\partial v} = -\frac{1}{2}AQ^{-1}A^Tv^* - b = 0$

$\cancel{v = 2b}$

$AQ^{-1}A^Tv^* = -2b$

$v^* = -2(AQ^{-1}A^T)^{-1}b$