3.2 1. Refomulate the loss function w.r.t. w_j.

$J(w_j) = \|w_j \cdot X_{\cdot,j} + X' \cdot w' - y\|_2^2 + \lambda|w_j| + \lambda\|w'\|_1$

If $X_{\cdot,j} = 0$, $J(w_j) = \|X'w' - y\|_2^2 + \lambda|w_j| + \lambda\|w'\|_1$

So. $w_j = 0$ is the minimizer of $J(w_j)$.

2. $J(w_j) = w_j^2\|X_{\cdot,j}\|_2^2 + 2w_j X_{\cdot,j}^T(X' \cdot w' - y) + \|X'w' - y\|_2^2 + \lambda|w_j| + \lambda\|w'\|_1$

3. If $w_j > 0$, $J'(w_j) = 2w_j\|X_{\cdot,j}\|_2^2 + 2X_{\cdot,j}^T(X' \cdot w' - y) + \lambda$.

If $w_j < 0$, $J'(w_j) = 2w_j\|X_{\cdot,j}\|_2^2 + 2X_{\cdot,j}^T(X' \cdot w' - y) - \lambda$

~~So if $w_j = 0$~~ $2X_{\cdot,j}^T(X' \cdot w' - y) + \lambda > 2X_{\cdot,j}^T(X' \cdot w' - y) - \lambda$

If both these terms are less than 0, $w_j = -\dfrac{2X_{\cdot,j}^T(X' \cdot w' - y) + \lambda}{\|X_{\cdot,j}\|^2} > 0$

If ~~both~~ they are bigger than 0. $w_j = -\dfrac{2X_{\cdot,j}^T(X'w' - y) - \lambda}{\|X_{\cdot,j}\|^2} < 0$.

If 0 is between them the $w_j = 0$.

4. ~~As shown~~ Shown in 3.

HW2

4 1. $J'(0;v)=\lim_{h\to 0}\frac{f(hv)-f(0)}{h}$

$$=\lim_{h\to 0}\left[\|hXv-y\|_2^2+\lambda h\|v\|_1-\|y\|_2^2\right]/h$$

~~$\|hXv-y\|_2^2+\|y\|_2^2<$~~

$\|hXv-y\|_2^2=(hXv-y)^T(hXv-y)$

$$=h^2v^TX^TXv-2hy^TXv+y^Ty.$$

$J'(0;v)=-2y^TXv+\lambda\|v\|_1$ 2. $J'(0;v)\geq 0\Rightarrow \lambda\geq\frac{2y^TXv}{\|v\|}$

~~3. $-2y^TXv\geq -2\|y^TX\|_\infty\|v\|_1$~~

~~Suppose $y^TX=u$, $-2y^TXv=-2\sum_{i=1}^m u_iv_i\geq -2\max(|u_i|)\cdot\sum_{i=1}^m\|v_i\|$~~

~~If $\lambda\geq 2\|y^TX\|_\infty$, $J'(0;v)\geq 0$.~~

3. Suppose $u=y^TX$.

$$\frac{2y^TXv}{\|v\|}=\frac{2\sum_{i=1}^m u_i\cdot v_i}{\sum_{i=1}^m\|v_i\|}\leq\frac{2\max(|u_i|)\sum_{i=1}^m|v_i|}{\sum_{i=1}^m|v_i|}=2\max(|u_i|)=2\|y^TX\|_\infty$$

if $\lambda\geq 2\|y^TX\|_\infty$. $w^*=0$ will be a global minimizer

4. $\|Xw+b1-y\|_2^2=w^TX^TXw+2w^TX^T(b\cdot 1-y)-2b1^Ty+y^Ty$

$J(hv)-J(0)=h^2v^TX^TXv+2hv^TX^T(b\cdot 1-y)+\lambda\cdot h\|v\|_1$

$J'(0;v)=2v^TX^T(b1-y)+\lambda\|v\|_1$

If $b\cdot 1=\bar{y}$, for $\lambda\geq 2\|X^T(y-\bar{y})\|_\infty$, $J'(0;v)\geq 0$. So $(w^*,b^*)=(0,\bar{y})$ is minimizer for J.

4.2 1. If a and b have different sign.

We can always find $a' \cdot b' \geq 0$ and $a'+b' = a+b$.

$x_1 \cdot a + x_2 \cdot b = x_1(a+b) = x_1 a' + x_2 \cdot b'$.

$\lambda|a'| + \lambda|b'| = \lambda|a+b| < \lambda|a| + \lambda|b|$

Thus $(a,b,r)^T$ can not be a minimizer.

For any c,d, if $c \cdot d \geq 0$ and $c+d = a+b$.

As the above analysis, $J(c,d) = J(a,b)$.

2. Note $J(\theta) = \|X_1 a + X_2 b + X_r \cdot r\|_2 + \lambda(a^2+b^2) + \lambda r^T r$.

$x_1 a + x_2 b = x_1(a+b)$, $a^2+b^2 \geq \frac{(a+b)^2}{2}$.

If $a \neq b$, we can always find $a' = \frac{a+b}{2} = b'$ different from a,b,

letting $J(a',b') \leq J(a,b)$.

So if (a,b) is a minimizer, a must ~~eq~~ equal b.