Q. 6.6.5.

Given:
$n=2$
$p=2$
$x_{11} = x_{12}$
$x_{21} = x_{22}$
$y_1 + y_2 = 0$
$x_{11} + x_{21} = 0$
$x_{12} + x_{22} = 0$
$\hat{\beta}_0 = 0$

a). Ridge regression, to minimize:

$$\sum_{i=1}^{n} \left(y_i - \beta_0 - \sum_{j=1}^{p} \beta_j x_{ij} \right)^2 + \lambda \sum_{j=1}^{p} \beta_j^2$$

$$\Rightarrow \sum_{i=1}^{2} \left(y_i - \beta_0 - \sum_{j=1}^{2} \beta_j x_{ij} \right)^2 + \lambda \sum_{j=1}^{2} \beta_j^2$$

$$\Rightarrow \left(y_1 - \hat{\beta}_0 - \hat{\beta}_1 x_{11} - \hat{\beta}_2 x_{12} \right)^2 + \left(y_2 - \hat{\beta}_0 - \hat{\beta}_1 x_{21} - \hat{\beta}_2 x_{22} \right)^2 + \lambda \left(\hat{\beta}_1^2 + \hat{\beta}_2^2 \right)$$

$\Rightarrow$ as given $x_{11} = x_{12}$ & $x_{22} = x_{21}$

we get :-

$$\left[y_1 - \beta_0 - x_{11}(\hat{\beta}_1 + \hat{\beta}_2)\right]^2 + \left[y_2 - \hat{\beta}_0 - x_{22}(\hat{\beta}_1 + \hat{\beta}_2)\right]^2$$
$$+ \lambda\left(\hat{\beta}_1^{\,2} + \hat{\beta}_2^{\,2}\right)$$

Now, $\hat{\beta}_0 = 0$

$$\Rightarrow \left[y_1 - x_{11}(\hat{\beta}_1 + \hat{\beta}_2)\right]^2 + \left[y_2 - x_{22}(\hat{\beta}_1 + \hat{\beta}_2)\right]^2$$
$$+ \lambda\left(\hat{\beta}_1^{\,2} + \hat{\beta}_2^{\,2}\right) \longrightarrow \text{equation } (1).$$

b). Prove $\hat{\beta}_1 = \hat{\beta}_2$

By taking derivatives of the equation ① above with respect to $\hat{\beta}_1$ & $\hat{\beta}_2$ and setting them equal to 0. :-

Taking derivative wrt $\hat{\beta}_1$

$$\frac{\partial(\hat{\beta}_1, \hat{\beta}_2)}{\partial(\hat{\beta}_1)} = 0 \quad \left[\text{let equation 1 be } f(\hat{\beta}_1 \hat{\beta}_2)\right] \Rightarrow$$

$$2\left[y_1 - x_{11}(\hat{\beta}_1 + \hat{\beta}_2)\right] \cdot \left[0 - x_{11} - 0\right] +$$

$$2\left[y_2 - x_{22}(\hat{\beta}_1 + \hat{\beta}_2)\right] \cdot \left[0 - x_{22} - 0\right] +$$

$$2 \cdot \lambda \cdot \hat{\beta}_1$$

$$= \left[2y_1 - 2\hat{\beta}_1 x_{11} - 2\hat{\beta}_2 x_{11}\right] \cdot (-x_{11}) +$$

$$\left[2y_2 - 2\hat{\beta}_1 x_{22} - 2\hat{\beta}_2 x_{22}\right] \cdot (-x_{22}) + 2\lambda\hat{\beta}_1 = 0$$

$\doteq -2$

$+$

$\neq$

$$= -2y_1x_{11} + 2\hat{\beta}_1 x_{11}^2 + 2\hat{\beta}_2 x_{11}^2 - 2y_2x_{22} + 2\hat{\beta}_1 x_{22}^2 + 2\hat{\beta}_2 x_{22}^2 + 2\lambda\hat{\beta}_1 = 0$$

$$\left[\because\ y_1 + y_2 = 0 \Rightarrow y_1 = -y_2,\quad x_{11} = -x_{21},\quad x_{12} = -x_{22}\right]$$

$$\Rightarrow 4\hat{\beta}_1 x_{11}^2 + 4\hat{\beta}_2 x_{11}^2 - 4y_1x_{11} + 2\lambda\hat{\beta}_1 = 0$$

$$\Rightarrow \hat{\beta}_1\left(4x_{11}^2 + 2\lambda\right) - 4y_1x_{11} + 4x_{11}^2\hat{\beta}_2 = 0$$

we get
$$\boxed{\hat{\beta}_1 = \frac{2y_1x_{11} - 2x_{11}^2\hat{\beta}_2}{\lambda + 2x_{11}^2}}$$

— equation 2

Similarly,

by doing $\dfrac{\partial(\hat{\beta}_1, \hat{\beta}_2)}{\partial \hat{\beta}_2} = 0$

we get

$$\boxed{\hat{\beta}_2 = \frac{2y_1x_{11} - 2x_{11}^2\beta_1^2}{2\lambda + 2x_{11}^2}}$$

– equation (3).

Now let $c = \dfrac{2y_1x_{11}}{\lambda + 2x_{11}^2}$

& $K = \dfrac{-2x_{11}^2}{\lambda + 2x_{11}^2}$

So from equation (2) & (3) we get:

$$\hat{\beta}_1 = C + K\beta_2 \quad \text{— (4)}$$

$$\hat{\beta}_2 = C + K\hat{\beta}_1 \quad \text{— (5)}$$

Now, substituting value of β_1 in (5)

$$\hat{\beta}_2 = c + k\hat{\beta}_1$$

$$\hat{\beta}_2 = c + k\left(c + k\hat{\beta}_2\right)$$

$$\hat{\beta}_2 = c + kc + k^2\hat{\beta}_2$$

$$\hat{\beta}_2 - k^2\hat{\beta}_2 = c + kc$$

$$\hat{\beta}_2\left(1 - k^2\right) = c\left(k+1\right)$$

$$\boxed{\hat{\beta}_2 = \frac{c\,\cancel{(k+1)}}{(1-k)\cancel{(1+k)}} = \frac{c}{1-k}} \text{ - eq (6)}$$

Now substituting value of $\hat{\beta}_2$ in (4).

$$\hat{\beta}_1 = c + k \cdot \left(\frac{c}{1-k}\right)$$

$$= \frac{c\left(1-k\right) + kc}{1-k}$$

$$\boxed{\hat{\beta}_1 = \frac{c - c\cancel{k} + kc}{1-k} = \frac{c}{1-k}} \text{ eq (7).}$$

we see that eq 6 & 7 are same.

$\therefore \boxed{\hat{\beta}_1 = \hat{\beta}_2}$

Hence, proved.

c)

The lasso optimization problem seeks to minimize :

$$(y_1 - \hat{\beta}_1 x_{11} - \hat{\beta}_2 x_{12})^2 + (y_2 - \hat{\beta}_1 x_{21} - \hat{\beta}_2 x_{22})^2$$
$$+ \lambda (|\hat{\beta}_1| + |\hat{\beta}_2|)$$

$[\because \hat{\beta}_0 = 0]$

Also, further simplifying we get

$$2(y_1 - \hat{\beta}_1 x_{11} - \hat{\beta}_2 x_{11})^2 + \lambda(|\hat{\beta}_1| + |\hat{\beta}_2|)$$

equation (A) .

d).

Taking partial derivatives with respect to $\hat{\beta}_1$ & $\hat{\beta}_2$ & setting them to 0, we get:

$$4x_{11}\left[y_1 - x_{11}(\hat{\beta}_1 + \hat{\beta}_2)\right] = \pm\lambda$$

The sign of RHS depends on the sign of $\hat{\beta}_1$ and $\hat{\beta}_2$.

If $\hat{\beta}_1$ & $\hat{\beta}_2$ are +ve, the sign is +ve

If they are -ve, sign is -ve.

This equation represents the boundary of the laso constraint and hence the lasso optimization problem has many possible solutions.