(a) Ridge regression seeks to minimize:

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

Given that $n = 2,\ p = 2,\ x_{11} = x_{12} = x_1$

$x_{21} = x_{22} = x_2$

$x_{11} + x_{21} = 0$

$x_{12} + x_{22} = 0$

$\hat{\beta}_0 = 0$

from (1)

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

$$\Rightarrow \text{Min } (y_1 - \hat{\beta}_1 x_1 - \hat{\beta}_2 x_1)^2 + (y_2 - \hat{\beta}_1 x_2 - \hat{\beta}_2 x_2)^2 + \lambda(\hat{\beta}_1^2 + \hat{\beta}_2^2)$$

(b) $\min (y_1 - \hat{\beta}_1 x_1 - \hat{\beta}_2 x_2)^2 + (y_2 - \hat{\beta}_1 x_2 - \hat{\beta}_2 x_2) + \lambda(\hat{\beta}_1^2 + \hat{\beta}_2^2)$

derivating wrt to $\hat{\beta}_1$:

$$\Rightarrow y_1^2 + (\hat{\beta}_1 x_1 + \hat{\beta}_2 x_1)^2 - 2y_1(\hat{\beta}_1 x_1 + \hat{\beta}_2 x_1)$$
$$+ y_2^2 + (\hat{\beta}_1 x_2 + \hat{\beta}_2 x_2) - 2y_2(\hat{\beta}_1 x_2 + \hat{\beta}_2 x_2)$$
$$+ \lambda(\hat{\beta}_1^2 + \hat{\beta}_2^2)$$

$$\Rightarrow y_1^2 + (\hat{\beta}_1 x_1)^2 + (\hat{\beta}_2 x_1)^2 + 2\hat{\beta}_1 x_1 \hat{\beta}_2 x_1 - 2y_1 \hat{\beta}_1 x_1 - 2y_1 \hat{\beta}_2 x_1$$
$$+ y_2^2 + (\hat{\beta}_1 x_2)^2 + (\hat{\beta}_2 x_2)^2 + 2\hat{\beta}_1 x_2 \cdot \hat{\beta}_2 x_2 - 2y_2 \beta_1 x_2 - 2y_2 \hat{\beta}_2 x_2$$
$$+ \lambda\hat{\beta}_1^2 + \lambda\hat{\beta}_2^2 \quad - (11)$$

Now taking derivative & setting to '0'

$$\Rightarrow 2\hat{\beta}_1 x_1 + 2x_1^2\hat{\beta}_2 - 2y_1x_1 + 2\hat{\beta}_1 x_2 + 2\hat{\beta}_2 x_2^2 - 2y_2x_2 + 2\lambda\hat{\beta}_1 = 0$$

$$\Rightarrow \hat{\beta}_1(2x_1 - 2x_1y_1 + 2x_2 + 2\lambda) = 2x_2y_2 - 2x_1^2\hat{\beta}_2 - 2\hat{\beta}_2 x_2^2 + 2y_1x_1$$

$$\hat{\beta}_1 = \frac{2x_2y_2 - 2x_1^2\hat{\beta}_2 - 2\hat{\beta}_2 x_2^2 + 2x_1y_1}{2x_1 + 2x_2 + 2\lambda \cancel{- 2x_1y_1}} \quad \text{— (iii)}$$

Now taking derivative of (ii) and setting to '0' wrt to $\hat{\beta}_2$ we get.

$$\Rightarrow 2\hat{\beta}_2 x_1 + 2\hat{\beta}_2 x_2 + 2\lambda\hat{\beta}_2 = 2x_2y_2 - 2\hat{\beta}_1 x_2^2 - 2\hat{\beta}_1 x_1^2 + 2x_1y_1$$

$$\hat{\beta}_2(2x_1 + 2x_2 + 2\lambda) = 2x_1y_1 + 2x_2y_2 - 2\hat{\beta}_1(x_1^2 + x_2^2)$$

$$\hat{\beta}_2 = \frac{2x_1y_1 + 2x_2y_2 - 2\hat{\beta}_1(x_1^2 + x_2^2)}{2x_1 + 2x_2 + 2\lambda} \quad \text{— (iv)}$$

Subtracting (iii) and (iv)

$$\hat{\beta}_1 - \hat{\beta}_2 = \frac{2x_1y_1 + 2x_2y_2 - 2\hat{\beta}_2(x_1^2 + x_2^2)}{2x_1 + 2x_2 + 2\lambda} - \frac{2x_1y_1 + 2x_2y_2 - 2\hat{\beta}_1(x_1^2 + x_2^2)}{2x_1 + 2x_2 + 2\lambda}$$

$$= \frac{(x_1^2 + x_2^2)}{2x_1 + 2x_2 + 2\lambda} \times (\hat{\beta}_1 - \hat{\beta}_2)$$

The symmetry between the LHS & RHS suggests

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

(c) The Lasso optimisation problem is similar to that of the ridge and minimizes

$$(Y_1 - \hat{\beta}_1 x_1 - \hat{\beta}_2 x_1)^2 + (Y_2 - \hat{\beta}_1 x_2 - \hat{\beta}_2 x_2)^2$$
$$+ \lambda(|\beta_1| + |\beta_2|)$$

(d) The Lasso constraint is

$$(Y_1 - \hat{\beta}_1 x_1 - \hat{\beta}_2 x_1)^2 + (Y_2 - \hat{\beta}_1 x_2 - \hat{\beta}_2 x_2)^2$$

subject to $|\hat{\beta}_1| + |\hat{\beta}_2| \leq s$

→ This means that the Lasso constraint takes the form of a diamond centered at the origin of the plane $(\hat{\beta}_1 . \hat{\beta}_2)$ & the intersection the axes at a distance s from origin using the constraints of the problem, we have

$$2\left(y_1 - (\hat{\beta}_1 + \hat{\beta}_2) x\right]^2 \geq 0 \quad -(1)$$

to minimize this $\hat{\beta}_1 + \hat{\beta}_2 = \frac{y_1}{x_1}$ (setting (1) to '0')

This is a line which is parallel to the edge of the diamond of the constraint. So, the edge $\hat{\beta}_1 + \hat{\beta}_2 = s$ is a solution to the optimization problem.

similarly.

$\hat{\beta}_1 + \hat{\beta}_2 = -s$ is also a solution with $\hat{\beta}_1 + \hat{\beta}_2 \leq 0$