Closed form solution of Ridge Regression.

$$\min_{\beta} \frac{1}{2}\|y - X\beta\|_2^2 + \frac{\lambda}{2}\|\beta\|_2^2$$

$$f(\beta) = \frac{1}{2}\|y - X\beta\|_2^2 + \frac{\lambda}{2}\|\beta\|_2^2$$

$$f(\beta) = \frac{1}{2}(Y - \beta X)^T(Y - \beta X) + \frac{\lambda}{2}\beta^T\beta$$

$$\frac{\partial f(\beta)}{\partial \beta} = -X^T(Y - \beta X) + \lambda\beta$$

$$\frac{\partial f(\beta)}{\partial \beta} = 0 \iff -X^T(Y - \beta X) + \lambda\beta = 0$$

$$\iff \beta(X^TX + \lambda I_p) = X^TY$$

$$\iff \boxed{\beta = (X^TY)(X^TX + \lambda I_p)^{-1}}$$

$$f(\beta) = \frac{1}{2n} \sum_{i=1}^{n} (y_i - x_i \beta)^2 + \lambda |\beta|$$

a)
$$f(\beta) = \frac{1}{2n} \sum_{i=1}^{n} \left(y_i^2 - 2 x_i y_i \beta + x_i^2 \beta^2\right) + \lambda |\beta|$$

$$f(\beta) = \left(\frac{1}{2n} \sum_{i=1}^{n} y_i^2\right) - \left(\sum_{i=1}^{n} \frac{x_i y_i}{n} \beta\right) + \sum_{i=1}^{n} \frac{x_i^2 \beta^2}{2n} + \lambda |\beta|$$

We have:
$$\frac{\sum_{i=1}^{n} x_i^2}{n} = 1$$

And $\frac{1}{2n} \sum_{i=1}^{n} y_i^2 \sim$ constant

Therefore

$$\min_{\beta \in \mathbb{R}} f(\beta) = \min_{\beta \in \mathbb{R}} \left(\frac{1}{2} \beta^2 - \frac{1}{n} \beta \sum_{i=1}^{n} y_i x_i + \lambda |\beta|\right)$$

$$\frac{1}{n}\sum_{i=1}^{n} y_i x_i > 0.$$

...ssume:

$\beta < 0.$

$$\min_{\beta\in\mathbb{R}} \frac{1}{2}\beta^2 - \frac{1}{n}\beta\sum_{i=1}^{n} y_i x_i + \lambda|\beta|$$

$$= \min_{\beta\in\mathbb{R}} \frac{1}{2}\beta^2 - \frac{1}{n}\beta\sum_{i=1}^{n} y_i x_i - \lambda\beta$$

$$\frac{\partial}{\partial\beta} = 0 \Leftrightarrow \beta - \frac{\sum_{i=1}^{n} y_i x_i}{n} - \lambda = 0$$

$$\Leftrightarrow \beta = \frac{\sum_{i=1}^{n} y_i x_i}{n} + \lambda > 0 \text{ ↯}.$$

(Invalid !!!)

$\beta > 0$ while $\sum_{i=1}^{n} y_i x_i > 0$

$\boxed{\text{therefore: } \beta > 0}$

Case 1 a:

$$\frac{1}{n}\sum_{i=1}^{n} y_i x_i \le \lambda \quad ; \beta > 0$$

$$\min_{\beta\in\mathbb{R}} \frac{1}{2}\beta^2 - \frac{1}{n}\beta\sum_{i=1}^{n} y_i x_i + \lambda\beta$$

$$\geq \min_{\beta\in\mathbb{R}} \frac{1}{2}\beta^2 - \frac{1}{n}\beta\lambda + \lambda\beta \geq 0$$

$\boxed{\text{therefore optimum } \beta = 0}$

$$\cancel{\frac{\partial}{\partial\beta} = 0 \Leftrightarrow \beta - \frac{\lambda}{n} + \lambda = 0 \Leftrightarrow} \boxed{\beta = \lambda - \frac{\lambda}{n}}$$

$\frac{1}{n}\sum_{i=1}^{n} y_i x_i > \lambda.$

$\min\left(\frac{1}{2}\beta^2 - \frac{1}{n}\beta\sum_{i=1}^{n} y_i x_i + \lambda\beta\right)$

$\cancel{= \left(\frac{1}{2}\beta^2 - \frac{1}{n}\beta\lambda + \lambda\beta\right)}$

$\frac{\partial}{\partial\beta} = \beta - \frac{\sum y_i x_i}{n} + \lambda = 0 \iff \boxed{\beta = \frac{\sum y_i x_i}{n} - \lambda}$

Case 2: $\frac{1}{n}\sum_{i=1}^{n} y_i x_i = 0.$

$\frac{1}{2}\beta^2 - \frac{1}{n}\beta\sum_{i=1}^{n} y_i x_i + \lambda|\beta|$

$= \frac{1}{2}\beta^2 + \lambda|\beta|$

$\boxed{\text{If } \beta = 0 \Rightarrow f(\beta) = 0}$ (true)

If $\beta < 0$:

$\frac{\partial f}{\partial\beta} = \frac{\partial\left(\frac{1}{2}\beta^2 - \lambda\beta\right)}{\partial\beta} = \beta - \lambda = 0$

$\iff \beta = \lambda$ (Invalid) since $\lambda \geq 0$

Therefore $\beta = 0$ when $\sum_{i=1} y_i x_i = 0$

If $\beta > 0$

$\frac{\partial f}{\partial\beta} = \frac{\partial\left(\frac{1}{2}\beta^2 + \lambda\beta\right)}{\partial\beta} = \beta + \lambda = 0$

$\iff \beta = -\lambda$ (Invalid) since $\lambda \geq 0$

$$\frac{1}{n}\sum_{i=1}^{n} y_i x_i < 0.$$

$\beta > 0$:

$$f(\beta) = \frac{1}{2}\beta^2 - \frac{\beta}{n}\sum_{i=1}^{n} y_i x_i + \lambda\beta$$

$$\frac{\partial f}{\partial(\beta)} = \beta - \frac{\sum_{i=1}^{n} y_i x_i}{n} + \lambda = 0.$$

$$\Leftrightarrow \frac{\partial f}{\partial \beta} = 0 \Leftrightarrow \beta = \lambda - \frac{\sum_{i=1}^{n} y_i x_i}{n} \quad (\cancel{\text{hold true}})$$

(Invalid)!!

If $\beta < 0$:

$$f(\beta) = \frac{1}{2}\beta^2 - \frac{\beta}{n}\sum_{i=1}^{n} y_i x_i - \lambda\beta$$

$$\frac{\partial f}{\partial(\beta)} = \beta - \frac{\sum_{i=1}^{n} y_i x_i}{n} - \lambda = 0$$

$$\Leftrightarrow \beta = \sum_{i=1}^{n} \frac{y_i x_i}{n} + \lambda \quad (\text{hold true})$$

therefore $\boxed{\beta \leq 0 \text{ while } \frac{\sum_{i=1}^{n} y_i x_i}{n} < 0}$

(3a)

$$\frac{1}{n}\sum_{i=1}^{n} y_i x_i \geq -\lambda. \; ; \quad \beta \leq 0$$

$$-\frac{1}{n}\beta\sum_{i=1}^{n} y_i x_i - \lambda\beta \geq 0 \quad \forall \beta \leq$$

$$f(\beta) = \frac{1}{2}\beta^2 - \frac{1}{n}\beta\sum_{i=1}^{n} y_i x_i$$

$$\frac{1}{2}\beta^2 \Rightarrow \boxed{\min f(\beta) = 0 \text{ when } \beta = 0}$$

$$\frac{1}{n}\sum_{i=1}^{n} y_i x_i < -\lambda \; ; \quad \beta \le 0$$

$$f(\beta) = \frac{1}{2}\beta^2 - \frac{\beta}{n}\sum_{i=1}^{n} y_i x_i - \lambda\beta$$

$$\frac{\partial f(\beta)}{\partial \beta} = \beta - \frac{\sum y_i x_i}{n} - \lambda = 0$$

$$\Leftrightarrow \boxed{\beta = \frac{\sum y_i x_i}{n} + \lambda}$$

1) optimal solution:

$$\hat{\beta}_{optimal} = \text{sign}(\hat{\beta})\left(|\hat{\beta}| - \lambda\right)_+$$

where $\hat{\beta}$ is the least square estimatee

of $\sum_{i=} (y_i - x_i \beta)^2$