COORDINATE DESCENT FOR UNREGULARISED REG.

j^{th} DIMENSION / COORDINATE

$$\hat{y}_i = \theta_0 x_i^0 + \theta_1 x_i^1 + \dots \theta_j x_i^j + \dots \theta_d x_i^d$$

$$e_i = y_i - \hat{y}_i = y_i - (\theta_0 x_i^0 + \dots \theta_j x_i^j + \dots \theta_d x_i^d)$$

$$\sum_{i=1}^{N} e_i^2 = RSS = \sum_{i=1}^{N} \left(y_i - (\theta_0 x_i^0 + \dots \theta_j x_i^j + \dots \theta_d x_i^d)\right)^2$$

$$\frac{\partial RSS(\theta_j)}{\partial \theta_j} = 0 \Rightarrow 2 \sum_{i=1}^{N} \left(y_i - (\theta_0 x_i^0 + \dots \theta_j x_i^j + \dots)\right)(-x_i^j)$$

$$2 \sum_{i=1}^{N} \left(y_i - (\theta_0 x_i^0 + \dots \theta_d x_i^d)\right)(-x_i^j) + 2 \sum_{i=1}^{N} \theta_j (x_i^j)^2 = 0$$

$\hat{y}_i$ (WITHOUT θ_j)

$$\theta_j = \frac{\sum_{i=1}^{N} \left(y_i - (\theta_0 x_i^0 + \dots \theta_d x_i^d)\right)(x_i^j)}{\sum_{i=1}^{N} (x_i^j)^2}$$

$$\theta_j = \frac{\sum_{i=1}^{N} \left(y_i - (\theta_0 x_i^0 + \ldots \theta_d x_i^d)\right)\left(x_i^j\right)}{\sum_{i=1}^{N} \left(x_i^j\right)^2}$$

$$= \frac{\rho_j}{z_j}$$

$$\rho_j = \sum_{i=1}^{N} x_i^j \left(y_i - \hat{y}_i^{(-j)}\right)$$

$-j$ means all but 'j'th dimension used for predicting $\hat{y}_i$

$$z_j = \sum_{i=1}^{N} \left(x_i^j\right)^2$$

COORDINATE DESCENT FOR LASSO

MINIMIZE $\sum e_i^2 + \delta^2 \{ |\theta_0| + |\theta_1| + \dots |\theta_j| + \dots |\theta_d| \}$

RSS FOR UNREGULARISED CASE

$$\frac{\partial (\text{LASSO OBJECTIVE})}{\partial \theta_j} = -2p_j + 2\theta_j z_j + \delta^2 \frac{\partial}{\partial \theta_j} |\theta_j| \quad \text{①}$$

$$\frac{\partial}{\partial \theta_j} |\theta_j| = \begin{cases} 1 & \theta_j > 0 \\ [-1, 1] & \theta_j = 0 \\ -1 & \theta_j < 0 \end{cases}$$

CASE I $(\theta_j > 0)$

Eq^n (1) $= 0$

$$-2p_j + 2\theta_j z_j + \delta^2 \times 1 = 0$$

$$\Rightarrow \quad \theta_j = \frac{p_j - \delta^2/2}{z_j}$$

FOR $\theta_j > 0$; $p_j - \delta^2/2 > 0$

$\Rightarrow$ WHENEVER $\boxed{p_j > \delta^2/2 \quad \text{SET} \quad \theta_j = \frac{p_j - \delta^2/2}{z_j}}$

Case III $\theta_j < 0$

$$\boxed{p_j < -\frac{\delta^2}{2} \quad \text{SET} \quad \theta_j = \frac{p_j + \delta^2/2}{z_j}}$$

Case II

$$\theta_j = 0$$

$$\frac{\partial}{\partial \theta_j}(\text{LASSO OBJECTIVE}) = -2\rho_j + \cancelto{0}{2\theta_j z_j} + \delta^2 \cancelto{[-1,1]}{\frac{\partial}{\partial \theta_j}} |\theta_j|$$

$$= \left[-2\rho_j - \delta^2,\ -2\rho_j + \delta^2\right]$$

Now 0 lies b/w -1 and 1

$$-2\rho_j - \delta^2 \le 0 \quad \text{AND} \quad -2\rho_j + \delta^2 \ge 0$$

$$\boxed{-\frac{\delta^2}{2} \le \rho_j \le \delta^2/2 \implies \text{SET } \theta_j = 0}$$

SUMMARY FOR COORDINATE DESCENT

$$\theta_j = \left[\begin{array}{ll} \dfrac{\rho_j + \frac{\delta^2}{2}}{z_j} & \text{if } \rho_j < -\frac{\delta^2}{2} \\ 0 & \text{if } -\frac{\delta^2}{2} \leq \rho_j \leq \frac{\delta^2}{2} \\ \dfrac{\rho_j - \frac{\delta^2}{2}}{z_j} & \text{if } \rho_j > \frac{\delta^2}{2} \end{array}\right]$$