Glossary

1. **EY=** Expected value of Y
2. **X=** is the vector of independent variables (x1, x2, ...xp)
3. **$(X_j)^p_{j=1}$, =** this represents the same meaning as X= (x1, x2, ...xp) where j goes from 1 to p.
4. $Xj\beta$ **:=**(β1, β2....) represents the best regression parameters of the regression line.
5. **Y =** $\beta' X + \epsilon$**=** $\Sigma \beta j Xj + \epsilon$ This represents the regression equation.
6. **$E(Y - \beta' X)^2 = \frac{1}{N} \Sigma((Y - \beta' X)^2)$ =** Mean Squared error
7. **E(Y-β'X)*X =** First derivative of Mean squared error or normal equation
8. **(Y-β'X)** =ϵ= Residuals= **unexplained part** of the regression equation
9. **E*ϵ*X=** First derivative of Mean squared error Replace(**(Y-β'X) by**ϵ in equation 6)**.** Also called a **normal equation**.
10. **EY^2** Variance of Y and is given by: $\frac{1}{N}\Sigma(Y - \overline{Y})^2$, $\overline{Y}$ is the average value of Y
11. $\frac{1}{M}\Sigma(Yk - \widehat{\beta}Xk)^2$= MSEtest Where M is the number of observations in the test sample.
12. **Adjusted R^2**= 1- $(\frac{n}{n-p})\frac{En\widehat{\epsilon^2}}{EnY^2}$
13. X can be written as X=(D,W')', Where D is the target regressor, and W are the controls
14. $\tilde{V} = V - \gamma'_{VW} W, \quad \gamma_{VW} = \arg\min_{\gamma} \mathrm{E}(V - \gamma' W)^2,$ = The "**residual**" V˜ is found out by subtracting the part of V that is linearly predicted by W.
15. $\tilde{Y} = \beta_1 \tilde{D} + \beta_2' \tilde{W} + \tilde{\epsilon},$ In this, Y˜ represents the "**residual**" that is found out by subtracting the part of Y that is linearly predicted by W, D~ represents "**residual**" is found out by subtracting the part of D that is linearly predicted by W. and same for W~ and ϵ~.

This file is meant for personal use by rswetland@outlook.com only.
Sharing or publishing the contents in part or full is liable for legal action.

16. $$\min_{b \in \mathbb{R}^p} \quad \sum_i (Y_i - b'X_i)^2 + \lambda \cdot \sum_{j=1}^{p} |b_j|,$$ This is the objective function for **Lasso regression,** where we minimize the sum of squared prediction error plus a penalty given by the sum of the absolute values of the coefficients times a penalty level **λ**.

17. $$\hat{\beta}(\lambda) = \arg\min_{b \in \mathbb{R}^p} \sum_{i=1}^{n} (Y_i - b'X_i)^2 + \lambda \sum_j b_j^2.$$ This is the objective function for **Ridge regression,** where we minimize the sum of squared prediction error plus a penalty given by the sum of the Squared values of the coefficients times a penalty level **λ**.

$$\hat{\beta} = \arg\min_{b_1, \ldots, b_M} \sum_{i=1}^{n} \left(Y_i - \sum_{m=1}^{M} b_m 1(Z_i \in R_m) \right)^2$$

18. $\widehat{\beta}$ is equal to the average of Y_i's with Z_i's falling in the rectangle Rm. Whenever we have multiple observations falling under the same rectangle then the predicted value is the average of all the values in the rectangle. The rectangles or regions Rm are called nodes, and each node has a predicted value $\widehat{\beta}_m$ associated with it.

19. $$\text{CV-MSE}(\theta) = \frac{1}{K} \sum_{k=1}^{K} \text{MSE}_k(\theta).$$ This represents average **cross-validation MSE.** θ is a tuning parameter that minimizes the cross-validated MSE.

This file is meant for personal use by rswetland@outlook.com only.
Sharing or publishing the contents in part or full is liable for legal action.