X1: 1.5, 1.2, 1.7, 1.8, 1.0, 1.1, 2.0
X2: 10.2, 10.4, 10.8, 11.5, 14.2, 9.4, 10.4
Weights: 0.5,1
Y: 5.6, 5.4, 5.7, 5.9, 4.9, 6.0, 5.5

<u>Lasso Regression</u>

Residual sum of Squares + λ * (Sum of the absolute value of the magnitude of coefficients)

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

Where,
λ denotes the amount of shrinkage.
λ = 0 implies all features are considered and it is equivalent to the linear regression where only the residual sum of squares is considered to build a predictive model
λ = ∞ implies no feature is considered i.e, as λ closes to infinity it eliminates more and more features
The bias increases with an increase in λ

<u>Ridge Regression</u>

This method performs L2 regularization. When the issue of multicollinearity occurs, least-squares are unbiased, and variances are large, this results in predicted values to be far away from the actual values.

$$Cost\ (W) = RSS(W) + \ \lambda * (sum\ of\ squares\ of\ weights)$$

$$= \sum_{i=1}^{N} \left\{ y_i \ - \sum_{j=0}^{M} w_j\, x_{ij} \right\}^2 + \lambda \sum_{j=0}^{M} w_j^2$$

Also, the formula for gradients:

$$w = w - \alpha \nabla_w J(w)$$

Q. Calculate the Ridge and Lasso regression cost and 1st update for the above given data frame. Here, Y is the predicted value and Xi are the provided attributes of the 7 data samples for lambda = 2, alpha = 0.001