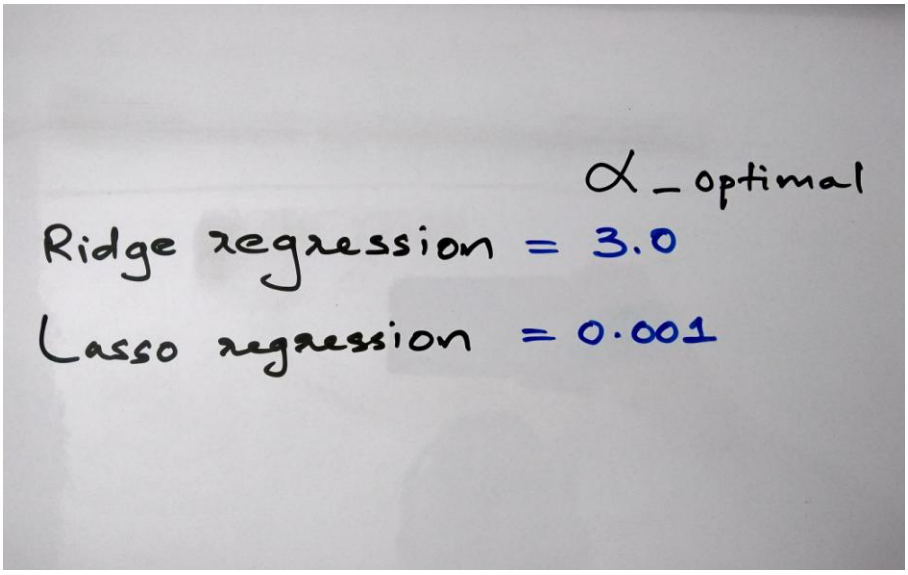
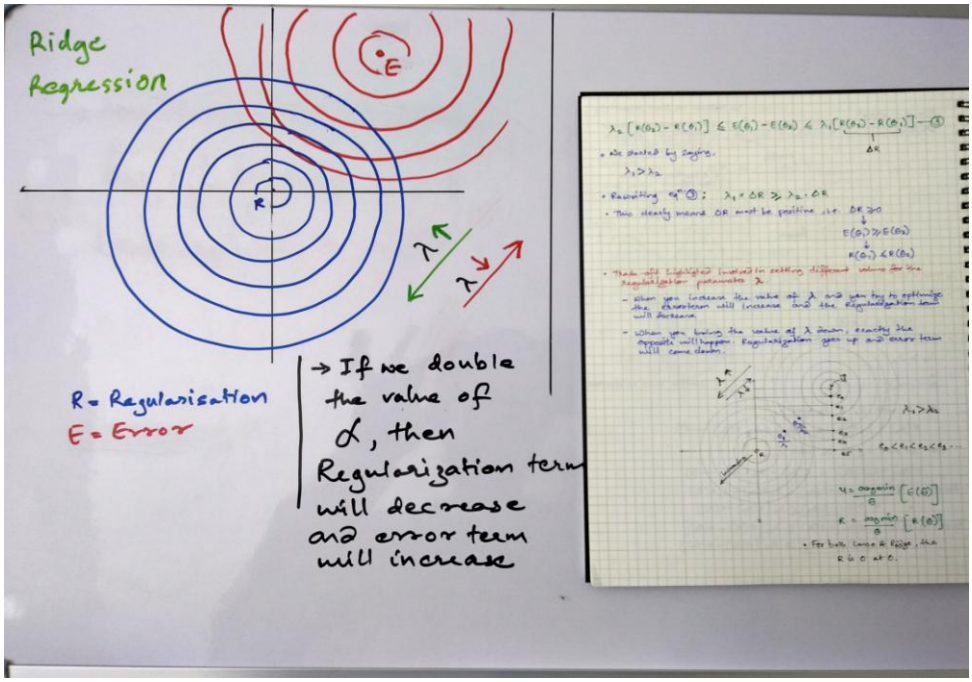
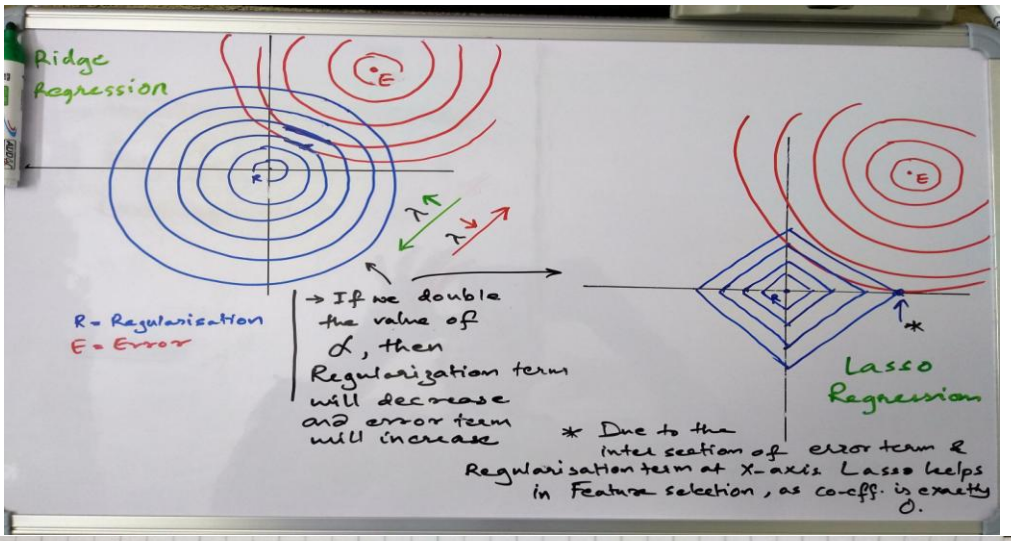
1. **What is the optimal value of alpha for ridge and lasso regression? What will be the changes in the model if you choose double the value of alpha for both ridge and lasso? What will be the most important predictor variables after the change is implemented?**

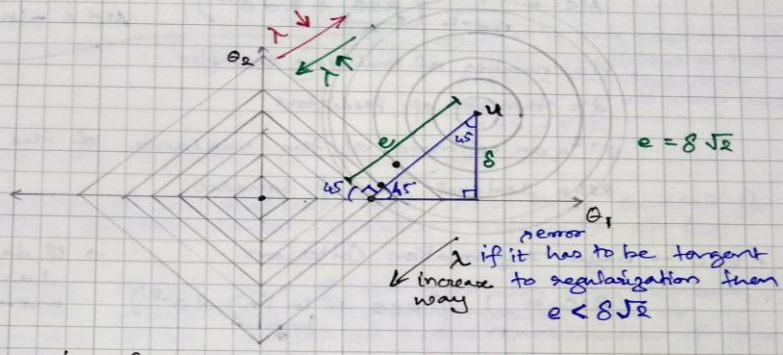






for LASSO:

$$R = \min_{\theta} [R(\theta)] \Rightarrow R(\theta) = \sum |\theta_i|$$



λ if it has to be tangent to regularization then error increase way

$e < \delta\sqrt{2}$

- number of non-zero components in Lasso $m \leq \frac{\delta_1 \cdot \|u\|_1}{\lambda}$

This is largest eigenvalue of $x^T \cdot x$ → [feature1 ... featur / datapoint1 / datapnt n]

- As you keep increasing λ, the number of non-zero components starts decreasing.

What kind of model will you get if both coefficient values have extremely low value. i.e. near to origin?

- A simple model which is likely to underfit the data.
- When both coefficients are close to zero, you are somewhere near the origin, where the error contours are high. but the regularization contours are very low.
- Such a model is likely to underfit the data since the error is quite high. (though it is 'simple' since the coefficients are small.

-> After doubling the alpha value for Ridge :

	Variable	Coeff
7	OverallQual	0.296213
25	1stFlrSF	0.233147
56	Total_SF	0.212063
28	GrLivArea	0.207526
8	OverallCond	0.207032
22	TotalBsmtSF	0.176506
18	BsmtFinSF1	0.170162
43	GarageArea	0.118287
26	2ndFlrSF	0.117468
36	TotRmsAbvGrd	0.115015

-> After doubling the alpha value for Lasso:

	Variable	Coeff
56	Total_SF	0.744581
7	OverallQual	0.583186
28	GrLivArea	0.443558
42	GarageCars	0.162434
8	OverallCond	0.160306
18	BsmtFinSF1	0.142666
189	SaleType_New	0.103521
24	CentralAir	0.091293
61	MSZoning_RL	0.084225
43	GarageArea	0.075877

2. **You have determined the optimal value of lambda for ridge and lasso regression during the assignment. Now, which one will you choose to apply and why?**

   As the Business goal is to get important features that affect the House price ,and the model is supposed to handle the variation, lot of features with smaller coefficient would increase in bias. Hence I would go for Lasso with 0.001 as alpha value as it pulls the less information value variables to 0. Making model dependent upon a small set of features. This is a generalized model.
   Also after doubling the alpha for lasso would make it more generalized, but then it might happen that we may loose some important variable.

3. **After building the model, you realised that the five most important predictor variables in the lasso model are not available in the incoming data. You will now have to create another model excluding the five most important predictor variables. Which are the five most important predictor variables now?**
   -> Now after dropping the top5 variables, the five most important predictor variables are: (some are derived features)
   - 1stFlrSF
   - 2ndFlrSF
   - TotalBsmtSF
   - GarageArea
   - Functional

4. **How can you make sure that a model is robust and generalisable? What are the implications of the same for the accuracy of the model and why?**
   To make model robust and generalisable by:
   - creating derived features
   - handling outliers
   - transformation of features
   - selection of features based on business understanding
   - Selection of minimalistic variables

   To make model more generalisable, it is important to get a good balance between bias-variance trade off. If the model is over-fit , It will definitely increase the accuracy and will perform good on the data it is trained on.
   But, In case scenarios such as Surprise Housing, A lot of features might go missing in the input dataset, Which might lead to a high variance for a good accuracy scoring model. Hence A model build with generalisable approach may lead to poor accuracy but it is more powerful in handling variance.