

Project Report

ELECTRIC MOTOR TEMPERATURE DETECTION

Submitted by:

Anushtha-01201032017

Mahima Sharu-01601032017

Sakshi-02001032017

Under the Supervision of

Mr. Rishabh Kaushal

Assistant Professor

Department of Information Technology



Department of Information Technology

Indira Gandhi Delhi Technical University for Women

Kashmere Gate, Delhi – 110006

CONTENTS

1.Introduction

1.1 Problem Statement

1.2 Objective

2.Proposed Methodology

2.1Dataset

2.2 Attribute Description

2.3 Data Pre-processing

2.3.1 Data Files Combining

2.3.2 Data Cleaning

2.4 Feature Computation

3.Data Exploration

3.1 Metrics

3.2 Visualizations

3.2.1 Identification of missing values

3.2.2 Correlation between variables

3.2.3 Relation between motor speed and rotor temp

3.2.4 Relation between torque and current q- component(i_q)

4. Algorithms

4.1 Implementation of Linear Regression

4.2 Implementation of the Principal Component Regression

4.3 Implementation of the Polynomial Regression

5. Regularization

5.1 regularization in linear regression

5.1a ridge regression

5.1b lasso regression

5.1c elastic net regression

6. Future work

CHAPTER1

Introduction:

The Permanent Magnet Synchronous Motor (PMSM) is an AC synchronous motor whose field excitation is provided by permanent magnets, and has a sinusoidal Back EMF waveform. With permanent magnets the PMSM can generate torque at zero speed. The main purpose of the data set's recording is to be able to model the stator temperature of a PMSM in real-time. In addition, precise thermal modelling gets more and more important with the rising relevance of functional safety

1.1 Problem Statement

Due to the intricate structure of an electric traction drive, direct measurement with thermal sensors is not possible for stator temperatures, sensor outage or even just deterioration can't also be administered properly.

Being able to have strong estimators for the stator temperature helps the automotive industry to manufacture motors with less material and enables control strategies to utilize the motor to its maximum capability. A precise torque estimate leads to more accurate and adequate control of the motor, reducing power losses and eventually heat build-up.

However, can we build a model that determines the value of motor temperature to a higher accuracy so that adequate measures be taken for the longer durability of the motor?

1.2 Objective

Predicting stator temperature and torque to determine the longevity of motor.

CHAPTER -2

Proposed Methodology

2.1 Data Set

Number of Instances- 998071

Number of Attributes- 13

Type- Categorical

ATTRIBUTE NAME	ATTRIBUTE VALUE
AMBIENT	FLOAT
COOLANT	FLOAT
u_d	FLOAT
u_q	FLOAT
MOTOR SPEED	FLOAT
TORQUE	FLOAT
I_d	FLOAT
I_q	FLOAT
PM	FLOAT
STATOR_YOKE	FLOAT
STATOR_TOOTH	FLOAT
STATOR_WINDING	FLOAT
PROFILE ID	INT

2.2 Attribute Description

- ambient - Ambient temperature as measured by a thermal sensor located closely to the stator.
- Coolant - Coolant temperature. The motor is water cooled. Measurement is taken at outflow.
- u_d - Voltage d-component
- u_q - Voltage q-component
- motor_speed - Motor speed
- torque - Torque induced by current.
- i_d - Current d-component
- i_q - Current q-component
- pm - Permanent Magnet surface temperature representing the rotor temperature. This was measured with an infrared thermography unit.
- stator_yoke - Stator yoke temperature measured with a thermal sensor.
- stator_tooth - Stator tooth temperature measured with a thermal sensor.
- stator_winding - Stator winding temperature measured with a thermal sensor.
- profile_id - Each measurement session has a unique ID. Make sure not to try to estimate from one session onto the other as they are strongly independent.

2.3 DataPre-processing

2.3.1 Data Collection

The data set comprises several sensor data collected from a permanent magnet synchronous motor (PMSM). Each profile is actually an independent experiment that was conducted and the measures were recorded during each experiment.

2.3.2 Data Cleaning

No row contained NULL values or Missing values so, no processing was required here.

2.4 Feature Computation

Direct Features:

These features are already present in the dataset as attributes, so no computation is required as such.

CHAPTER -3

Data Exploration

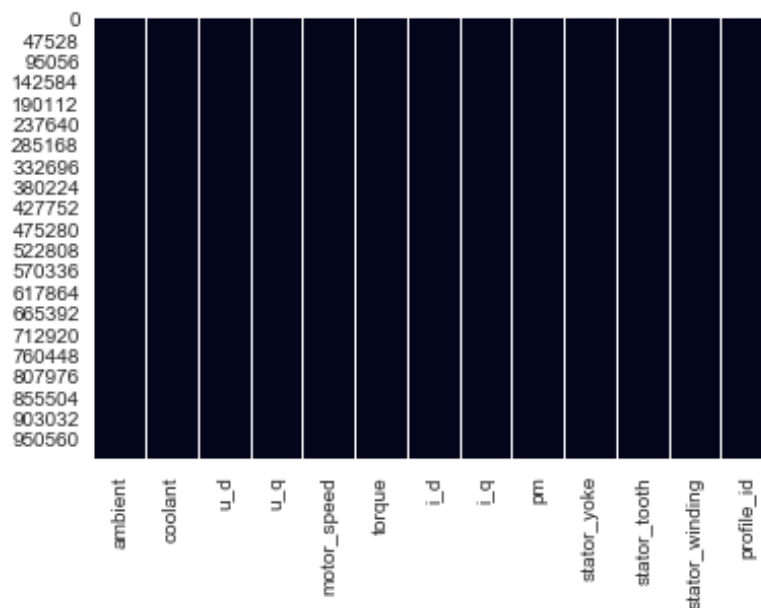
3.1 Metrics

ATTRIBUTE NAME	MEAN
AMBIENT	-0.003905
COOLANT	0.004723
u_d	0.004780
u_q	-0.005690
MOTOR SPEED	-0.006336
TORQUE	-0.003333
I_d	0.006043
I_q	-0.003194
PM	-0.004396
STATOR_YOKE	0.000609
STATOR_TOOTH	-0.002208
STATOR_WINDING	-0.003935
PROFILE ID	50.732001

Number of unique profiles IDs in dataset: 52

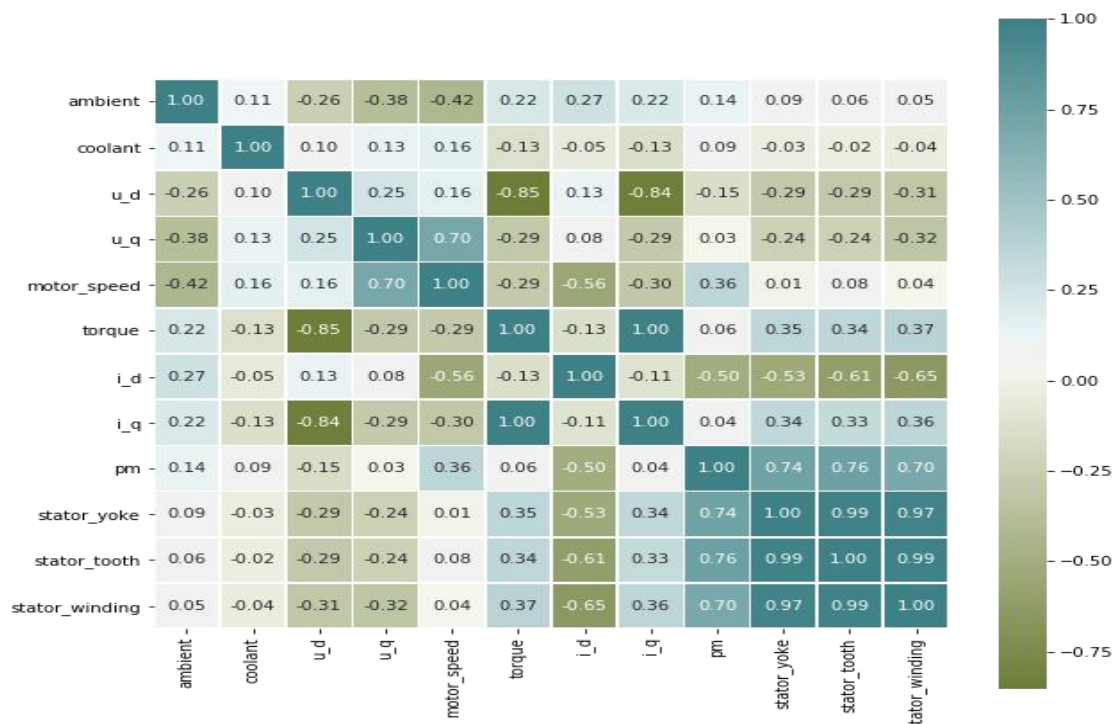
3.2 Visualization

3.2.1 Identification of Missing values:



The figure above shows the heatmap for different attributes in the dataset. Heatmap is created for the identification of any possible missing values in dataset. As seen, there are no missing values in the dataset, had there been any missing values, they would be shown with different colour.

3.2.2 Correlation between variables



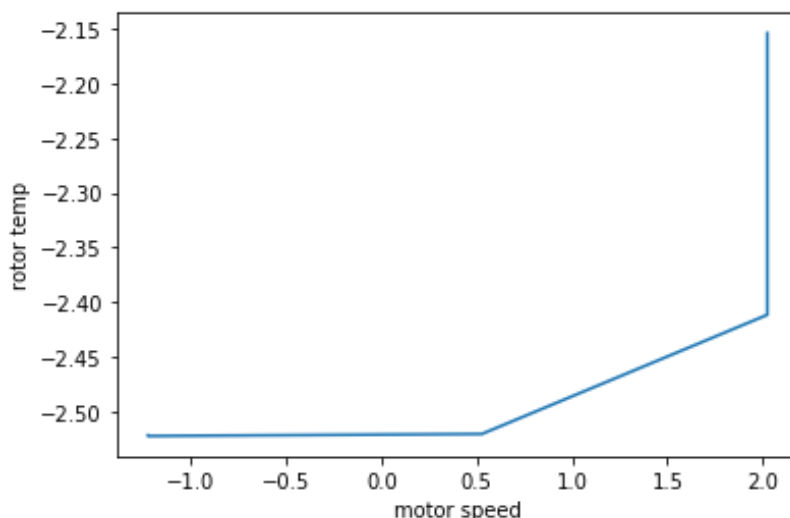
The figure here is a correlation matrix between different variables. It shows how two variables are related to each other. Dark shades represent positive correlation while lighter shades represent negative correlation. The numerical values are the values by which features are related to each other.

Observations: 1. There is an almost perfect linear relation between the torque and i_q . This suggests that we can use a linear regression model to estimate the torque using i_q .

2. The temperatures in stator yoke, tooth and winding are highly correlated since they represent the temperature in the stator and they should be pretty similar (there is no isolation in the stator between those three parts).

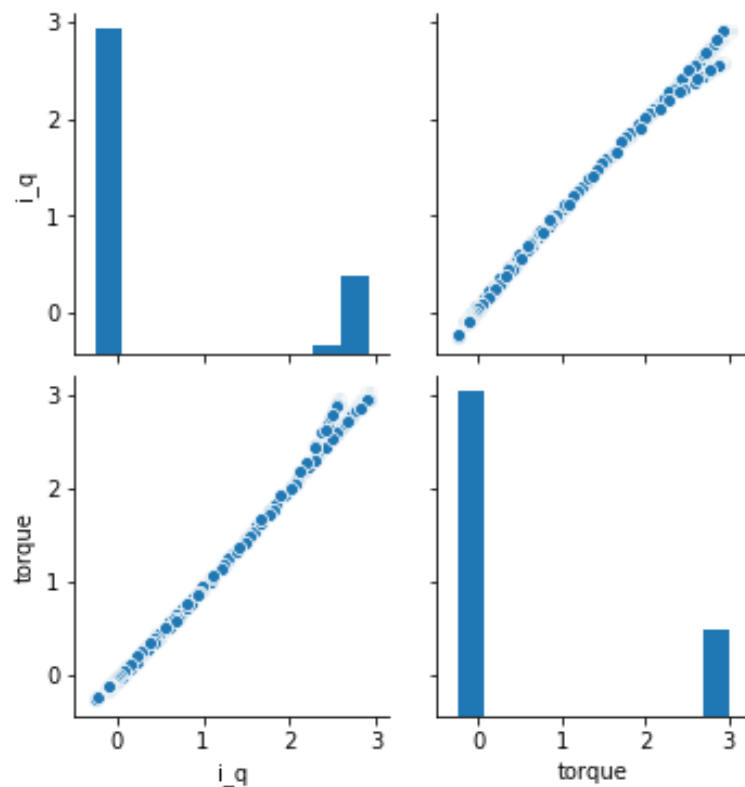
3. The coolant seems to be highly correlated with the stator yoke. It is understandable because the coolant runs through the stator yoke to cool it down.

3.2.3 Motor Speed and rotor temperature



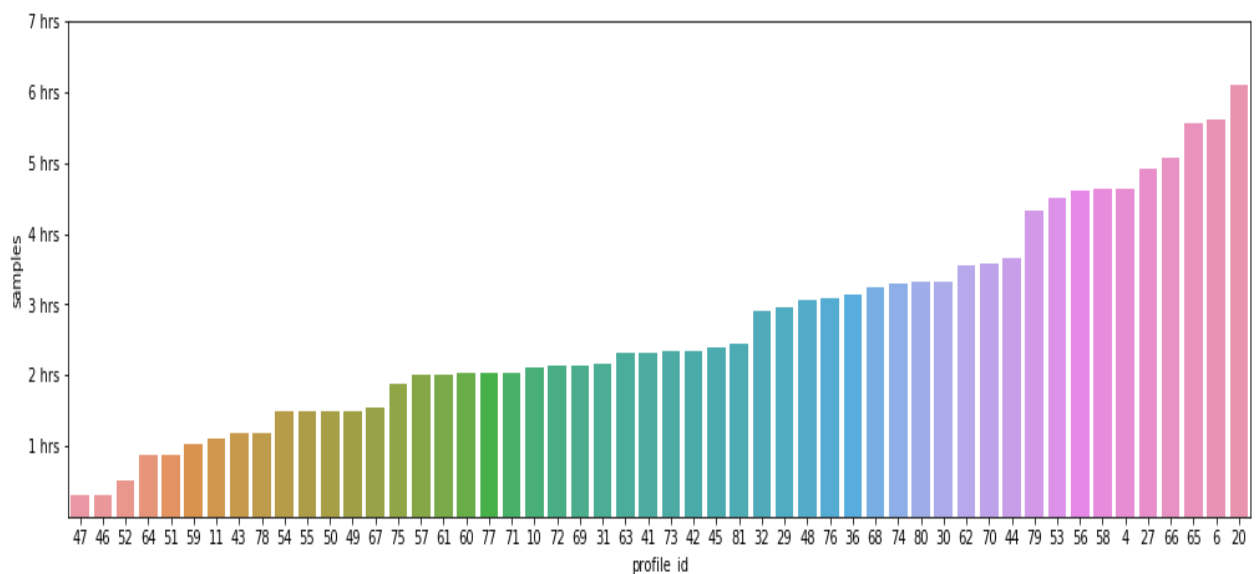
Here, this figure shows how motor speed affects change in rotor temperature. Rotor temp is the surface temperature of the motor in which changes are observed as shown in the figure relative to changes taking place in speed of motor which will generate heat loss. It shows that motor temp initially remains constant with gradual increase in speed, then rises linearly relative to increase in speed for some time before it keeps on rising on constant speed, hence causing heating of motor.

3.2.4 Relation between torque and current q- component(i_q)



The above figure is a pair plot between torque and current component clearly showing the linear relationship between the two variables. This implies that linear regression model can be used to estimate the torque using the measure of the current i_q .

3.2.5 Measurement of session length



The plot above shows that all measurement sessions range from 20 minutes to around 6 hours. The two short session ids "46" and "47" might be not very representative as temperatures inside electric motors need time to vary.

Chapter 4

Algorithms

4.1 Implementation of Linear Regression.

Linear regression is a linear approach to modelling the relationship between a dependent variable and one or more independent variables. The case of one independent variable is called simple linear regression. For more than one independent variable, the process is called **multiple linear regression**.

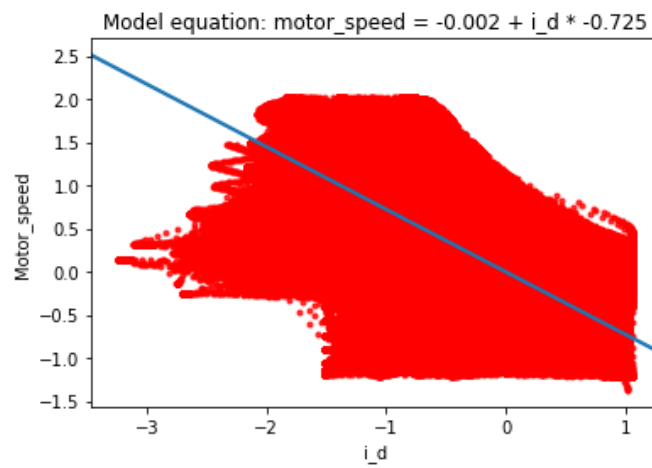
Linear regression was the first type of regression analysis to be studied rigorously, and to be used extensively in practical applications.^[4] This is because models which depend linearly on their unknown parameters are easier to fit than models which are non-linearly related to their parameters and because the statistical properties of the resulting estimators are easier to determine.

Linear regression models are often fitted using the least squares approach, but they may also be fitted in other ways, such as by minimizing the "lack of fit" in some other norm (as with least absolute deviations regression), or by minimizing a penalized version of the least squares cost function as in ridge regression (L^2 -norm penalty) and lasso (L^1 -norm penalty). Conversely, the least squares approach can be used to fit models that are not linear models. Thus, although the terms "least squares" and "linear model" are closely linked, they are not synonymous.

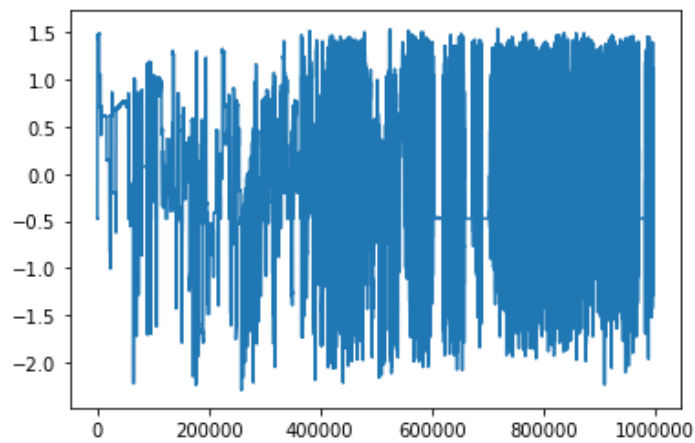
In the Electric Motor Temperature Database, the data set consists of multiple measurement sessions. The motor is excited by hand-designed driving cycles denoting a reference motor speed and a reference torque. Currents in d/q-coordinates (columns "id" and iq") and voltages in d/q-coordinates (columns "ud" and "uq") are a result of a standard control strategy trying to follow the reference speed and torque. Columns "motor_speed" and "torque" are the resulting quantities achieved by that strategy, derived from set currents and voltages.

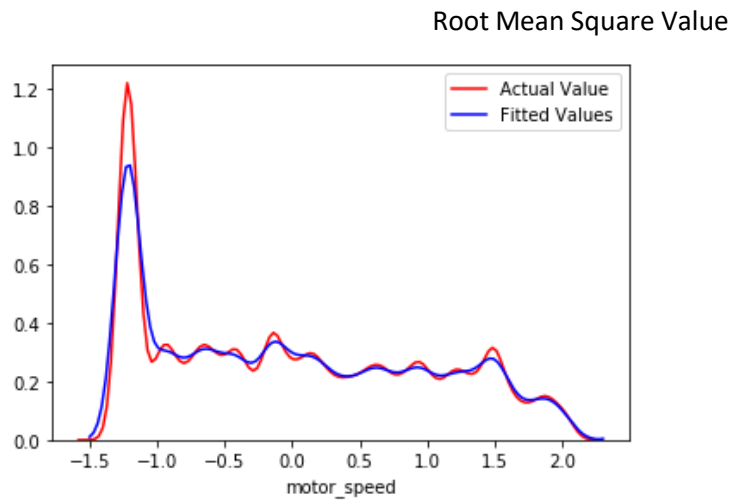
Graphical Representation of the Results:

Model Visualisation



Residual Form





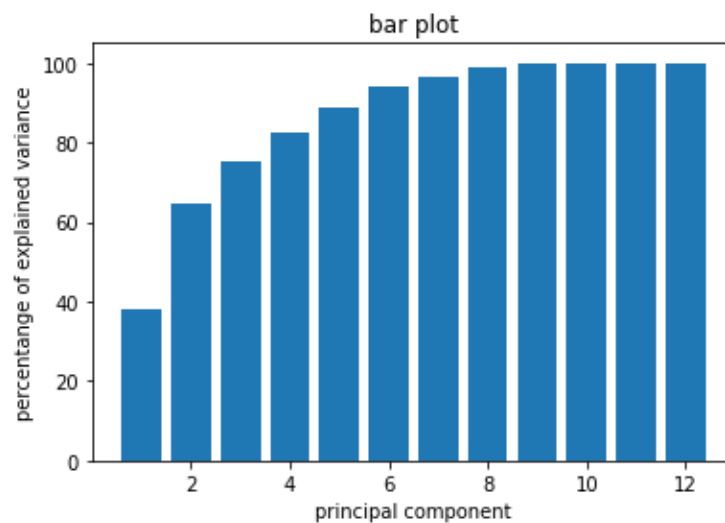
Description:

From the above graphical representations which shows the model visualisation, residual error and root mean squared error for different configurations.

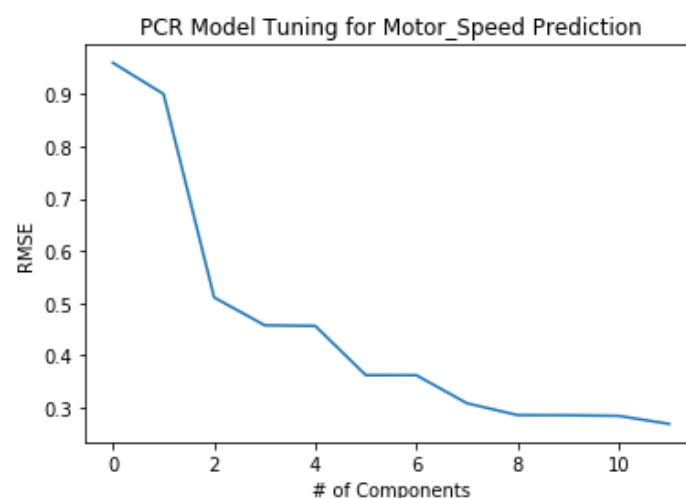
4.2 Implementation of Principal Component Regression.

Principal Components Regression is a technique for analyzing multiple regression data that suffer from multicollinearity. When multicollinearity occurs, least squares estimates are unbiased, but their variances are large so they may be far from the true value. By adding a degree of bias to the regression estimates, principal components regression reduces the standard errors. It is hoped that the net effect will be to give more reliable estimates.

Bar plot for percentage of explained variance vs principal component-----



Following the usual notation, suppose our regression equation may be written in matrix form as $\mathbf{Y} = \mathbf{XB} + \mathbf{e}$ where \mathbf{Y} is the dependent variable, \mathbf{X} represents the independent variables, \mathbf{B} is the regression coefficients to be estimated, and \mathbf{e} represents the errors or residuals.



10 component is good for the model because RMSE value is the smallest for this component number. That's why there is no need to tune the model.

Description:

From the above graphical representations which shows the Bar Plot and PCR(Principal Component Regression) Model Tuning for the motor speed prediction.

4.3 Implementation of Polynomial Regression:

Polynomial Regression is a form of linear regression in which the relationship between the independent variable x and dependent variable y is modeled as an n th degree polynomial. Polynomial regression fits a nonlinear relationship between the value of x and the corresponding conditional mean of y , denoted $E(y | x)$.

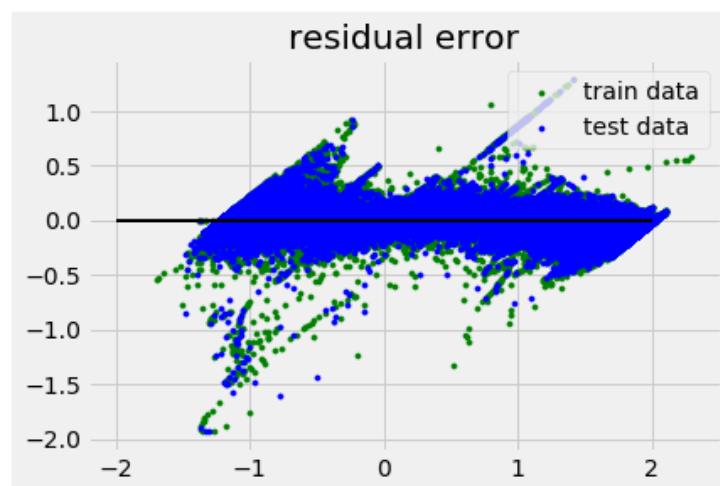
The basic goal of regression analysis is to model the expected value of a dependent variable y in terms of the value of an independent variable x . In simple regression, we used following equation –

$$y = a + bx + e$$

Residue Error:

As residuals are the difference between any data point and the regression line, they are sometimes called “**errors**.” Error in this context doesn’t mean that there’s something wrong with the analysis; it just means that there is some unexplained difference. In other words, the residual is the error that isn’t explained by the regression line.

Graph for the implementation of Polynomial Regression for residue model:



Description:

From the above graphical representations which shows the residual error based on the Polynomial Regression.

5. Regularization

5.1 Regularization in Linear Regression.

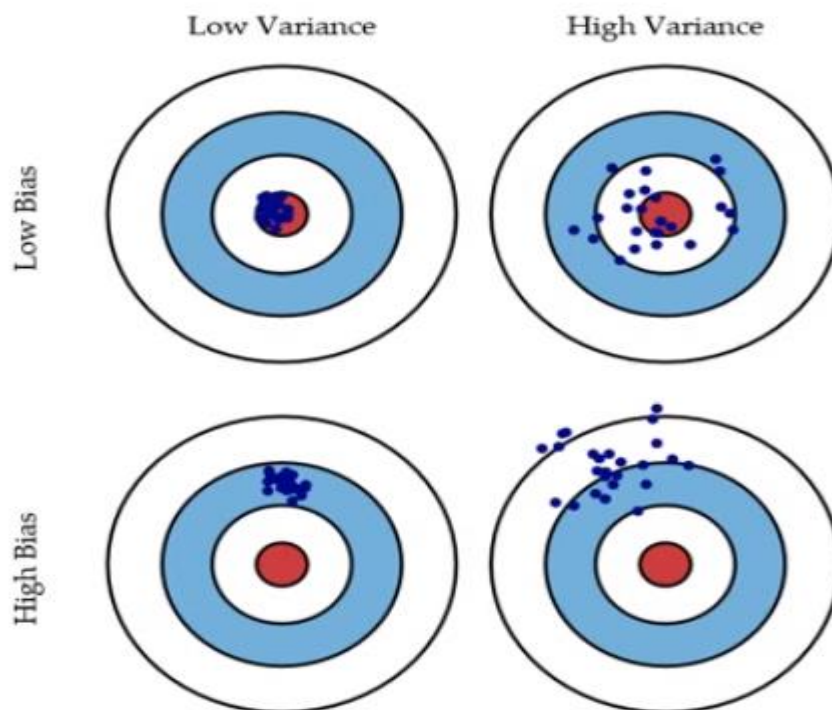
This is a form of regression, that constraints / regularizes or shrinks the coefficient estimates towards zero. In other words ,**this technique discourages learning a more complex or flexible model,so as to avoid the risk of overfitting.**

Bias and Variance in Regression models

Bias- Bias is the simplifying assumptions made by a model to make the target function easier to learn.

Variance- Variance is the amount that the estimate of the target function will change if different training data was used.

There could be 4 for bias/variance in regression model



1. As variance increases, the spread of our data point increases which result in less accurate prediction.

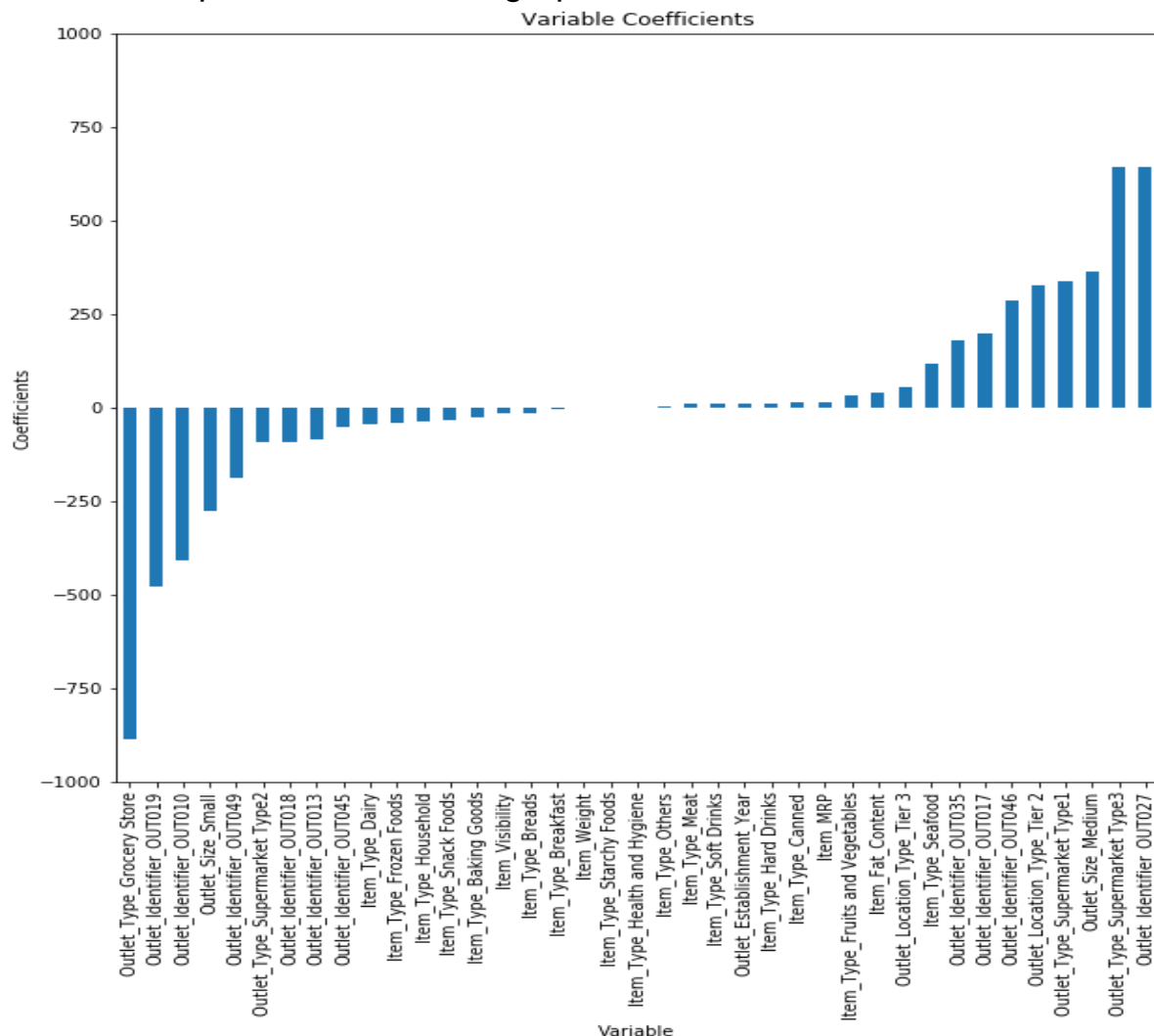
2. As Bias increases the error between our predicted value and the observed values increases.

Overfitting—As we add more and more parameters to our model, its complexity increases, which results in increasing variance and decreasing bias

To overcome overfitting there are two ways-

- 1.Reduce the model complexity
- 2.Regularization

In regularization , what we do is normally we keep the same number of features ,but reduce the magnitude of the coefficients. To do this one can use the plot the coefficient graph of all these variables.



Above show coeff. vs variable graph

We have two techniques to overcome Over-fitting problem

5.1a Ridge Regression for linear regression

Ridge regression is L2 norm.

Ridge regression adds “squared magnitude” of coefficient as penalty term to the loss function.

$$\sum_{i=1}^n \left(y_i - \beta_0 - \sum_{j=1}^p \beta_j x_{ij} \right)^2 + \lambda \sum_{j=1}^p \beta_j^2 = \text{RSS} + \lambda \sum_{j=1}^p \beta_j^2$$

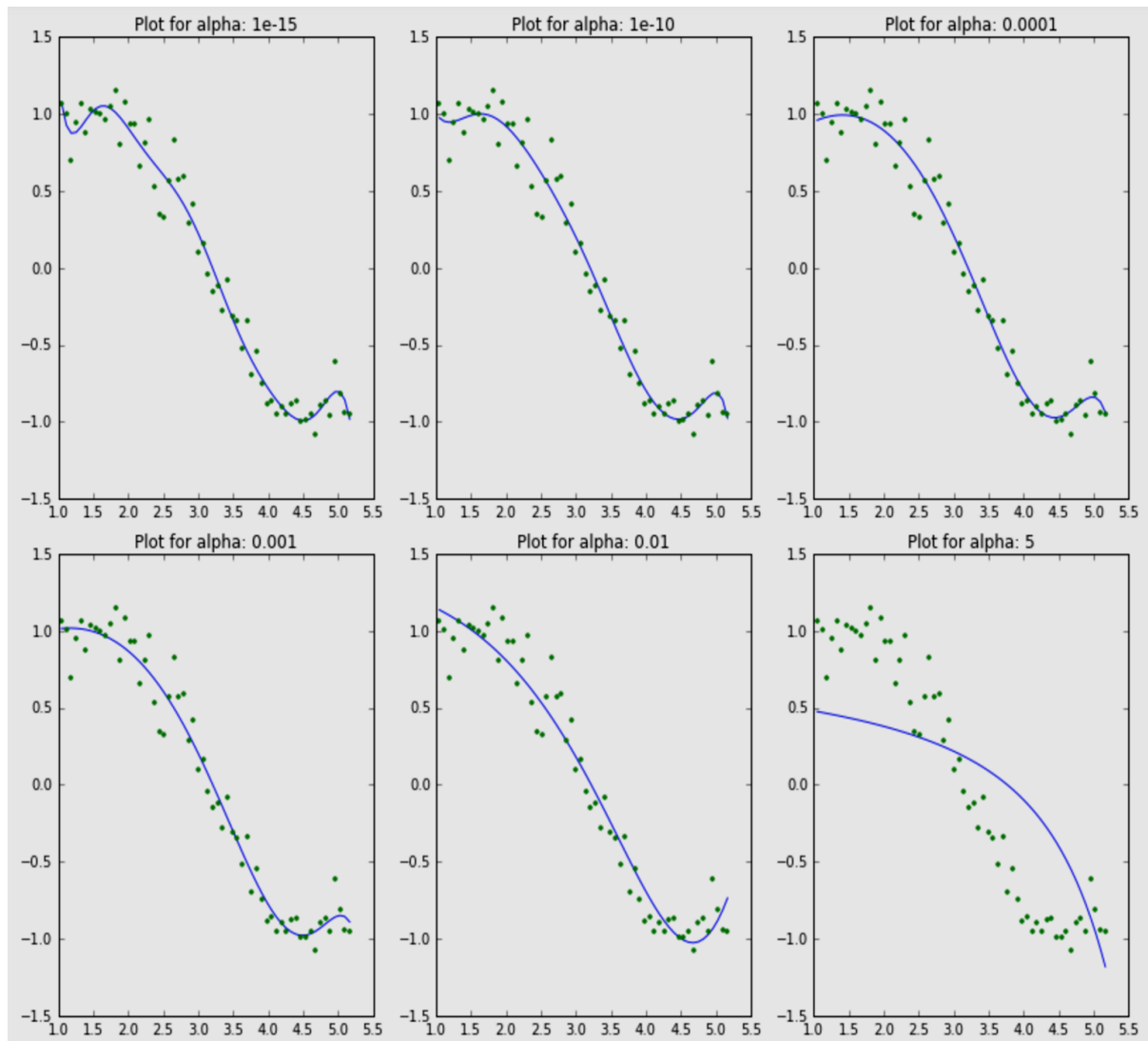
Objective = loss + α * (sum of square of coefficients)

Here, α (alpha) is the perimeter which balances the amount of emphasis given to minimizing loss v/s minimum sum of square of coefficients.

Alpha can take various values. Here we define alpha which is a penalty factor:

When,

1. The objective becomes same as simple linear regression. We'll get the same coefficients as simple linear regression.
2. The coefficients will be zero. Because of infinite weight on the square of coefficients, anything less than zero will make the objective infinite
3. The magnitude of alpha will decide the weight given to different parts of the objective. The coefficients will be somewhere between 0 and 1 for simple linear regression.



The above plots for the best fit line, we see that as the value of alpha increases the model complexity reduces. Higher the alpha, bigger the penalty though higher values of alpha reduces overfitting, significantly high values can cause underfitting as well(

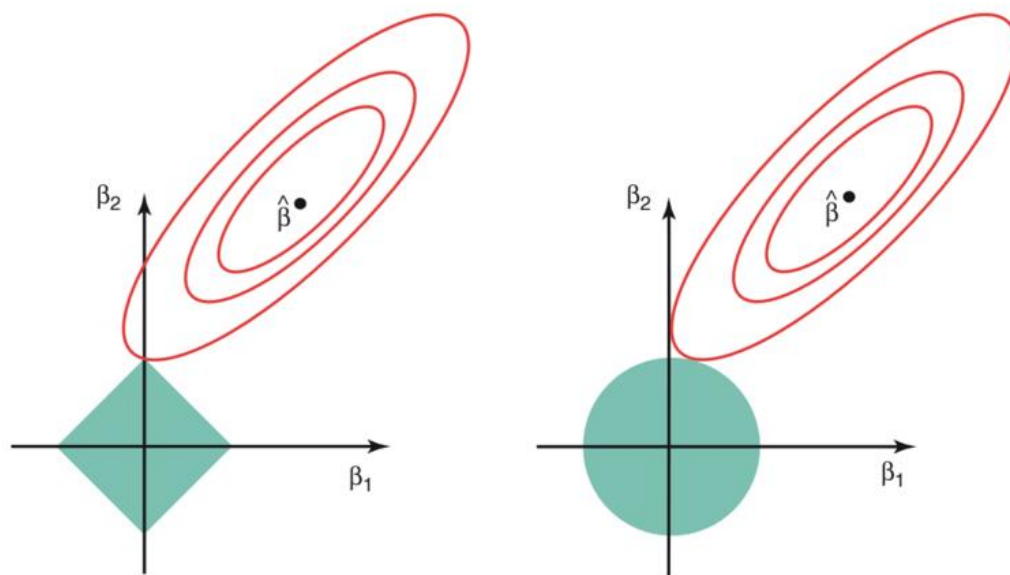
In lasso, the magnitude of the coefficient are shrunk to a small magnitude but they are never zero. It shrinks the parameters, therefore it is mostly used to prevent multicollinearity. It reduces the model complexity by **coefficient shrinkage**.

5.1b Lasso Regression

LASSO full form is least absolute shrinkage selector operator. It is quite similar to ridge expression. LASSO adds “absolute value of magnitude” of coefficient as penalty term to the loss function. It uses $|\beta_j|$ (modulus) instead of square of β , as its penalty. In statistics this is known as the L1 norm.

Consider there are 2 parameters in a given problem. Then according to above formulation, the **ridge regression is expressed by $\beta_1^2 + \beta_2^2 \leq s$** .

Similarly, **for lasso, the equation becomes, $|\beta_1| + |\beta_2| \leq s$** . This implies that *lasso coefficients have the smallest RSS(loss function) for all points that lie within the diamond given by $|\beta_1| + |\beta_2| \leq s$* .



The above image shows the constraint functions (green areas), for lasso (left) and ridge regression (right), along with contours for RSS (red ellipse). Since ridge regression has a circular constraint with no sharp points, this intersection will not generally occur on an axis, and so the ridge regression coefficient estimates will be exclusively non-zero. However, the lasso constraint has corners at each of the axes, and so the ellipse will often intersect the constraint region at an axis. When

this occurs, one of the coefficients will equal zero. In higher dimensions (where parameters are much more than 2), many of the coefficient estimates may equal zero simultaneously.

What does Regularization achieve?

A standard least squares model tends to have some variance in it, i.e. this model won't generalize well for a data set different than its training data. ***Regularization, significantly reduces the variance of the model, without substantial increase in its bias.*** So the tuning parameter λ , used in the regularization techniques described above, controls the impact on bias and variance. As the value of λ rises, it reduces the value of coefficients and thus reducing the variance. ***Till a point, this increase in λ is beneficial as it is only reducing the variance (hence avoiding overfitting), without losing any important properties in the data.*** But after certain value, the model starts losing important properties, giving rise to bias in the model and thus underfitting. Therefore, the value of λ should be carefully selected.

5.1 c Elastic Net Regression

It combines the power of L1 and L2 regularisation Elastic regression generally works well when we have a big dataset. Consider that we have a bunch of correlated independent variables in a dataset, then elastic net will simply form a group consisting of these correlated variables. Now if any one of the variable of this group is a strong predictor (meaning having a strong relationship with dependent variable), then we will include the entire group in the model building, because omitting other variables (like what we did in lasso) might result in losing some information in terms of interpretation ability, leading to a poor model performance.

The equation looks like these

$$L_{enet}(\hat{\beta}) = \frac{\sum_{i=1}^n (y_i - x_i' \hat{\beta})^2}{2n} + \lambda \left(\frac{1-\alpha}{2} \sum_{j=1}^m \hat{\beta}_j^2 + \alpha \sum_{j=1}^m |\hat{\beta}_j| \right),$$

Elastic Net Regression Equation

where α is the mixing parameter between ridge ($\alpha = 0$) and lasso ($\alpha = 1$) and λ .

$$\alpha = a + b$$

$$\lambda = a / (a+b)$$

here a and b are weights assigned to L1 and L2 term respectively and set in a such a way that they control trade off between L1 and L2

$$a * (\text{L1 term}) + b * (\text{L2 term})$$

Programmatically we use `l1_ratio` as a parameter of a function, which alone decides the type of regression (lasso, ridge, Elastic Net) Let α (or $a+b$) = 1, and now consider the following cases:

- If `l1_ratio` = 1, therefore if we look at the formula of `l1_ratio`, we can see that `l1_ratio` can only be equal to 1 if $a=1$, which implies $b=0$. Therefore, it will be a lasso penalty.
- Similarly if `l1_ratio` = 0, implies $a=0$. Then the penalty will be a ridge penalty.
- For `l1_ratio` between 0 and 1, the penalty is the combination of ridge and lasso.