29 March Assignment

June 18, 2023

[ ]:
```
Q1. What is Lasso Regression, and how does it differ from other regression techniques?
```

ANS-

[ ]:
```
Lasso Regression is a regression analysis method that performs both variable selection and regularization. It uses absolute values in
the penalty function instead of squares, which leads to penalizing (or equivalently constraining the sum of the absolute values of the
estimates) values that cause some of the parameter estimates to turn out exactly zero. This makes Lasso Regression useful for feature
selection.

Lasso Regression is similar to Ridge Regression in that both attempt to minimize the sum of squared residuals (RSS) along with some penalty
term. However, Lasso Regression constrains or regularizes the coefficient estimates of the model by zeroing out some features’ coefficients.
```

[ ]:
```
```

[ ]:
```
Q2. What is the main advantage of using Lasso Regression in feature selection?
```

ANS-

[ ]:
```
Lasso regression is useful for feature selection because it provides a principled way to reduce the number of features in a model.
By adding the L1 regularization term, Lasso regression can shrink the coefficients towards zero, and when   is sufficiently large, some
coefficients are driven to exactly zero. This property of Lasso makes it useful for feature selection, as the variables with zero coefficients
are effectively removed from the model. Lasso was designed to improve the interpretability of machine learning models by reducing the number
of features.

The main advantage of a LASSO regression model is that it has the ability to set the coefficients for features it does not consider
interesting to zero. This means that the model does some automatic feature selection to decide which features should and should not be
```

```
included on its own. Reduced overfitting is another advantage of LASSO␣
  ↪regression.
```

[ ]:

[ ]:
```
Q3. How do you interpret the coefficients of a Lasso Regression model?
```

ANS-

[ ]:
```
In Lasso Regression model, the coefficients are interpreted as the log odds for␣
  ↪a 1 unit change in the coefficient while holding all other
coefficients constant. Lasso regression performs L1 regularization, which adds␣
  ↪a penalty equal to the absolute value of the magnitude of
coefficients. This type of regularization can result in sparse models with few␣
  ↪coefficients; Some coefficients can become zero and eliminated
from the model. The coefficients can be used to understand the impact of each␣
  ↪feature on the target variable, and also help in feature
selection.
```

[ ]:

[ ]:
```
Q4. What are the tuning parameters that can be adjusted in Lasso Regression,␣
  ↪and how do they affect the
model's performance?
```

ANS-

[ ]:
```
In Lasso Regression, the tuning parameter ( ), sometimes called a penalty␣
  ↪parameter, controls the strength of the penalty term in
ridge regression and lasso regression. It is basically the amount of shrinkage,␣
  ↪where data values are shrunk towards a central point,
like the mean. The tuning parameter is used to control the trade-off between␣
  ↪fitting the model well to the training data and keeping
the model simple. A larger value of   will result in more shrinkage and a␣
  ↪simpler model.
```

[ ]:

[ ]:
```
Q5. Can Lasso Regression be used for non-linear regression problems? If yes,␣
  ↪how?
```

ANS-

[ ]:
```
Yes, Lasso Regression can be used for non-linear regression problems. One way␣
  ↪to do this is by using Gaussian basis functions and imposing
a weighted lasso penalty on a nonlinear regression model. This can help select␣
  ↪the number of basis functions effectively and reduce some
```

```
unknown parameters in linear regression models toward exactly zero.

Another way is to fit a lasso regressor to the whole lot, multiplying out your␣
↪brackets giving you 2m+2 coefficients. Then by performing a
change of variables you can make this a linear regression problem again.
```

[ ]:

```
Q6. What is the difference between Ridge Regression and Lasso Regression?
```

ANS-

```
Both Ridge Regression and Lasso Regression are regularization methods that␣
↪minimize the sum of squared residuals along with some penalty term.
The difference is that Ridge Regression takes the square of the coefficients␣
↪and never sets them to absolute zero, while Lasso Regression
takes the magnitude and tends to make coefficients to absolute zero. Lasso␣
↪Regression can be used for automatic feature selection and removing
insignificant variables from the model. However, Lasso Regression may struggle␣
↪with some types of data, such as when the number of predictors
is greater than the number of observations, or when there are highly collinear␣
↪variables. In these cases, Elastic Net, which combines the
regularization of both Lasso and Ridge, may be better. Ridge Regression may␣
↪perform better when many predictor variables are significant
and have similar coefficients.
```

[ ]:

```
Q7. Can Lasso Regression handle multicollinearity in the input features? If␣
↪yes, how?
```

ANS-

```
Yes, Lasso Regression can handle multicollinearity in the input features. In␣
↪fact, Lasso Regression is one of the methods that can handle
multicollinearity. It does this by shrinking the coefficients of correlated␣
↪variables towards zero1. This is because Lasso Regression uses
L1 regularization which adds a penalty term to the loss function that is␣
↪proportional to the absolute value of the coefficients.
This penalty term forces some of the coefficients to be zero, which effectively␣
↪removes some of the features from the model.
This way, Lasso Regression can handle multicollinearity by selecting only one␣
↪feature from a group of correlated features.
```

[ ]:

[ ]:
```
Q8. How do you choose the optimal value of the regularization parameter
(lambda) in Lasso Regression?
```

ANS-

[ ]:
```
To choose the optimal value of the regularization parameter (lambda) in Lasso
Regression, there are two main approaches:

Choose   such that some information criterion, e.g., AIC or BIC, is the
smallest.
Perform cross-validation and select the value of   that minimizes the
cross-validated sum of squared residuals (or some other measure).
Alternatively, you can use root finding algorithms like Brent's method or
golden section search to directly optimize   .
```