Linear Regression Models

Segment 5 – Model Selection

Topic 1 – Stepwise Model Selection

Sudarsan N.S. Acharya (sudarsan.acharya@manipal.edu)

Topics

1. Model Selection
2. Forward Stepwise Selection
3. Backward Stepwise Selection
4. Issues with Stepwise Selection

Model Selection

Model Selection

- Models can be compared based on performance metrics such as

Model Selection

- Models can be compared based on performance metrics such as the adjusted R^2 statistic,

Model Selection

- Models can be compared based on performance metrics such as the adjusted R^2 statistic, prediction error using cross validation,

Model Selection

- Models can be compared based on performance metrics such as the adjusted R^2 statistic, prediction error using cross validation, or model complexity scores such as AIC, BIC (to be discussed in detail in the next video).

Model Selection

- Models can be compared based on performance metrics such as the adjusted R^2 statistic, prediction error using cross validation, or model complexity scores such as AIC, BIC (to be discussed in detail in the next video).
- Smaller the model complexity scores, better is the model.

Model Selection

- Models can be compared based on performance metrics such as the adjusted R^2 statistic, prediction error using cross validation, or model complexity scores such as AIC, BIC (to be discussed in detail in the next video).
- Smaller the model complexity scores, better is the model.
- In general, we would like to have models that

Model Selection

- Models can be compared based on performance metrics such as the adjusted R^2 statistic, prediction error using cross validation, or model complexity scores such as AIC, BIC (to be discussed in detail in the next video).
- Smaller the model complexity scores, better is the model.
- In general, we would like to have models that (1) are simple

Model Selection

- Models can be compared based on performance metrics such as the adjusted R^2 statistic, prediction error using cross validation, or model complexity scores such as AIC, BIC (to be discussed in detail in the next video).
- Smaller the model complexity scores, better is the model.
- In general, we would like to have models that (1) are simple (2) generalize well

Model Selection

- Models can be compared based on performance metrics such as the adjusted R^2 statistic, prediction error using cross validation, or model complexity scores such as AIC, BIC (to be discussed in detail in the next video).
- Smaller the model complexity scores, better is the model.
- In general, we would like to have models that (1) are simple (2) generalize well (3) are interpretable.

Forward Stepwise Selection

Choosing a model complexity score (AIC/BIC) for model selection,

Forward Stepwise Selection

Choosing a model complexity score (AIC/BIC) for model selection,

Forward Stepwise Selection

Choosing a model complexity score (AIC/BIC) for model selection,

1. start with an intercept-only model.

Forward Stepwise Selection

Choosing a model complexity score (AIC/BIC) for model selection,

1. start with an intercept-only model.
2. Add the single predictor that leads to the greatest reduction in the chosen model score.

Forward Stepwise Selection

MANIPAL ACADEMY of HIGHER EDUCATION (Institution of Eminence Deemed to be University)

Choosing a model complexity score (AIC/BIC) for model selection,

1. start with an intercept-only model.
2. Add the single predictor that leads to the greatest reduction in the chosen model score.
3. Repeat step-2 until there is no improvement in the model complexity score.

Forward Stepwise Selection

MANIPAL ACADEMY of HIGHER EDUCATION
(Institution of Eminence Deemed to be University)

Choosing a model complexity score (AIC/BIC) for model selection,

1. start with an intercept-only model.
2. Add the single predictor that leads to the greatest reduction in the chosen model score.
3. Repeat step-2 until there is no improvement in the model complexity score.

Computational complexity is *quadratic* in the number of predictors.

Backward Stepwise Selection

Backward Stepwise Selection

Choosing a model complexity score (AIC/BIC) for model selection,

Backward Stepwise Selection

Choosing a model complexity score (AIC/BIC) for model selection,

1. start with a model that includes *all* predictors.

Backward Stepwise Selection

Choosing a model complexity score (AIC/BIC) for model selection,

1. start with a model that includes *all* predictors.
2. Delete the single predictor that leads to the greatest reduction in the chosen model score.

Backward Stepwise Selection

Choosing a model complexity score (AIC/BIC) for model selection,

1. start with a model that includes *all* predictors.
2. Delete the single predictor that leads to the greatest reduction in the chosen model score.
3. Repeat step-2 until there is no improvement in the model complexity score.

Backward Stepwise Selection

Choosing a model complexity score (AIC/BIC) for model selection,

1. start with a model that includes *all* predictors.
2. Delete the single predictor that leads to the greatest reduction in the chosen model score.
3. Repeat step-2 until there is no improvement in the model complexity score.

Computational complexity is again *quadratic* in the number of predictors.

Backward Stepwise Selection

Choosing a model complexity score (AIC/BIC) for model selection,

1. start with a model that includes *all* predictors.
2. Delete the single predictor that leads to the greatest reduction in the chosen model score.
3. Repeat step-2 until there is no improvement in the model complexity score.

Computational complexity is again *quadratic* in the number of predictors. However, backward stepwise cannot be used when the number of samples is less than or equal to the number of predictors as the design matrix will have *linearly dependent* columns.

Issues with Stepwise Selection

Issues with Stepwise Selection

- It is not guaranteed to find the best possible model out of all 2^p models containing subsets of the p predictors.

Issues with Stepwise Selection

- It is not guaranteed to find the best possible model out of all 2^p models containing subsets of the p predictors.
- The same data is used throughout in both forward and backward stepwise selection approaches.

Issues with Stepwise Selection

- It is not guaranteed to find the best possible model out of all 2^p models containing subsets of the p predictors.
- The same data is used throughout in both forward and backward stepwise selection approaches.
- Typically, this will result in an underestimate of the true prediction error of the model on unseen data.

Issues with Stepwise Selection

- It is not guaranteed to find the best possible model out of all 2^p models containing subsets of the p predictors.
- The same data is used throughout in both forward and backward stepwise selection approaches.
- Typically, this will result in an underestimate of the true prediction error of the model on unseen data.
- The train-test-validation split of the dataset or cross validation can be used to address this.

Issues with Stepwise Selection

- It is not guaranteed to find the best possible model out of all 2^p models containing subsets of the p predictors.
- The same data is used throughout in both forward and backward stepwise selection approaches.
- Typically, this will result in an underestimate of the true prediction error of the model on unseen data.
- The train-test-validation split of the dataset or cross validation can be used to address this.
- Bottom line:

Issues with Stepwise Selection

- It is not guaranteed to find the best possible model out of all 2^p models containing subsets of the p predictors.
- The same data is used throughout in both forward and backward stepwise selection approaches.
- Typically, this will result in an underestimate of the true prediction error of the model on unseen data.
- The train-test-validation split of the dataset or cross validation can be used to address this.
- Bottom line: separate test data from training data before model-building.

Summary

- Goals of model selection
- Differences between forward and backward stepwise model selection
- Issues with stepwise model selection