Week 8 - Introduction to Regression Models

Week 8 - Regression - What is Regression - What problems does this solve? - Linear Regression - Simple on paper implementation - Slightly more complex R implementation

Introduction:

Welcome to this lecture on linear regression in the context of mathematical modeling. Linear regression is a fundamental statistical and mathematical technique used to model the relationship between one or more independent variables and a dependent variable. In this lecture, we will explore the principles of linear regression, its applications, and how it fits into mathematical modeling. Understanding Linear Regression:

Linear regression aims to establish a linear relationship between independent variables (predictors) and a dependent variable (response). It assumes that this relationship can be expressed as a linear equation: Y = + X + X + … + X + , where: Y is the dependent variable. X , X , …, X are the independent variables. , , , …, are the coefficients. represents the error term, accounting for unexplained variability. Simple Linear Regression:

In simple linear regression, there is only one independent variable (X) and one dependent variable (Y). The relationship is expressed as Y = + X + . The goal is to find the best-fitting line that minimizes the sum of squared errors (residuals) between the predicted and actual values of Y. Multiple Linear Regression:

Multiple linear regression extends the concept to multiple independent variables: Y = + X + X + … + X + . It allows modeling complex relationships between multiple predictors and the dependent variable. Applications of Linear Regression:

Linear regression is widely used in various fields, including: Economics: Modeling economic factors and predicting outcomes. Finance: Predicting stock prices and risk assessment. Medicine: Predicting patient outcomes based on medical variables. Engineering: Predicting performance and optimizing processes. Social Sciences: Analyzing social and behavioral data. Model Evaluation:

Evaluating a linear regression model is essential to assess its quality and predictive power. Common evaluation metrics include: R-squared (R^2): Measures the proportion of variance explained by the model. Mean Squared Error (MSE) and Root Mean Squared Error (RMSE):

Measure the average squared error of predictions. Residual plots: Visualize the distribution of residuals to check for patterns or anomalies. Assumptions in Linear Regression:

Linear regression assumes several key assumptions: Linearity: The relationship between independent and dependent variables is linear. Independence: Residuals are independent of each other. Homoscedasticity: Residuals have constant variance across all levels of independent variables. Normality: Residuals follow a normal distribution. Advanced Techniques:

Linear regression can be extended with advanced techniques such as ridge regression, lasso regression, and elastic net regression to handle multicollinearity and prevent overfitting. Conclusion:

Linear regression is a powerful mathematical modeling technique that models relationships between independent and dependent variables using linear equations. It is widely applicable in various fields for prediction, analysis, and decision-making. Understanding the assumptions and techniques of linear regression is crucial for effective modeling and interpretation. Q&A Session:

Please feel free to ask any questions related to linear regression in mathematical modeling.

Theoretical Questions

Practical Questions

To complete for Week 9

- ☐ Install R
- ☐ Install R Studio
- ☐ Install the following libraries
- ☐ Create a R Markdown File
- ☐ Save your R Markdown File as a pdf