Lecture 1 - Linear Regression

Linear regression is a supervised learning algorithm that models the relationship between a dependent variable and one or more independent variables by fitting a linear equation to observed data.

Simple linear regression uses one predictor: y = mx + b. Multiple linear regression uses two or more predictors to model complex relationships.

The cost function for linear regression is the Mean Squared Error (MSE): J(theta) = (1/2m) sum(h(x) - y)^2. Errors are squared to make all errors positive and to penalise large deviations more than small ones.

R-squared (R^2) measures the proportion of variance explained by the model, ranging from 0 to 1. A value close to 1 indicates a strong fit.

Key Concepts

Gradient descent: An optimization algorithm that iteratively adjusts parameters in the direction of steepest descent of the loss function.

Learning rate: Controls the step size. Too high causes divergence; too small leads to slow convergence.

Overfitting: When the model fits training data too closely and fails to generalize. Use regularization (Ridge, Lasso) to prevent this.