

COMP6237 Logistic regression and model reduction

February 12, 2021

Abstract

Problem sheet for lecture on logistic regression and non-linear data. Please attempt to solve/answer before the tutorial on logistic regression (in which this problem sheet will be discussed). Worked solutions will be published after the tutorial.

1 Logistic Regression and model reduction I

Explore predicting Oscar success of movies using the data set <https://www.southampton.ac.uk/mb1a10/stats/filmData.txt> discussed in the lecture. Build logistic regression models to predict movie success based on all predictors given in the data set. Explore model reduction – which of the predictor (box office takings, critics score, length, budget, country of origin) should be included in the best model?

2 Logistic Regression II

A data set has been collected to relate the age of a learner to the outcome of driving tests. Carrying out logistic regression, somebody obtains a slope of $w=0.01$ and an intercept of $b = 0.1$. What are the chances of a 100 year old applicant to pass the test?

3 Logistic Regression III

Somebody collects a data set to analyze examination outcomes (discriminating between fail, pass, and repeat) of students on a three year Bsc degree and carries out multinomial logistic regression to predict the outcome dependent on the year of study. Results give: (i) intercept (fail)=1 slope (fail)=-1 and (ii) intercept (pass)=3 slope (pass)=-1/2. What is the chance of a student having to repeat the 3rd year?

4 Model Reduction

Consider the ridge regression problem (slide 34 of the lectures). Derive an expression for the optimal (augmented) weight vector w . In the formulation for ridge regression on the slide also the bias term in w is penalized. This is not always desirable. How would the procedure (and the result derived above) have to be modified to avoid this penalization?

5 Transforming data

Consider the problem of kernel regression (slide 49 of the lecture slides). Derive the expression for the optimal weight vector w given a transformation ϕ .