COMP6237 Linear Regression and Maximum Likelihood Estimation

January 21, 2021

Abstract

Problem sheet for lecture on linear regression and maximum likelihood estimation. Please attempt to solve/answer before first tutorial on linear regression (in which this problem sheet will be discussed). Worked solutions will be published after the tutorial.

1 Linear Regression I

A data set is constructed by taking 100 samples from a normal distribution with mean $\mu = 5$ and standard deviation $\sigma = 2$ to construct a random variable $X_i, i = 1,, 100$. Then, a 2nd random variable $Y_i, i = 1,, 100$ is constructed by taking the values of the corresponding X_i and adding one half of a third random variate drawn from a normal distribution with mean 5 and standard deviation 2 and thus a set of 100 pairs (X_i, Y_i) is obtained. Find the parameters of a linear regression of Y on X (both by doing the numerical experiment and by calculating the result analytically).

2 Linear Regression II

Some person wants to conduct a least squares regression on a data set of N (X, Y) pairs, but attaches varying importance to deviations of various (X, Y) pairs to the line of best fit. The relative importance of deviations of pair (X_i, Y_i) are given by a function $f(i) = f_i$. Find an expression for the line of best fit generated by this procedure.

3 Linear Regression III

Consider the regression problem in which we have pairs of (x_i, y_i), $x_i \in R^d$ and $y_i \in R$ (see page 27/28 of the lecture slides). Show that the (augmented) parameter vector $\tilde{w}$ can be obtained from the data matrix $\tilde{X}$ and the vector of y-values y via $\tilde{w} = (\tilde{X}^T \tilde{X})^{-1} \tilde{X}^T y$.

4 Maximum Likelihood Estimation I

Repeated coin tossing of an (unfair) coin produces 100 heads up and 120 tails up. Find a maximum likelihood estimate for the probability that a coin toss will result in heads up.

5 Maximum Likelihood Estimation II

A number of observations x_i, $i = 1, ..., N$ are known to have been sampled from an exponential distribution $P(x) \sim \exp(-\lambda x)$. Find a maximum likelihood estimate for λ.