

Multivariate Analysis

Homework #1 (2024-2025), Questions

Q1. Derivatives of traces of matrix products

This is useful for solving the minimization equations for regression and related procedures. Let A be an $M \times N$ matrix, U an $N \times M$ matrix, V an $N \times M$ matrix, W an $M \times M$ matrix, and Z an $N \times N$ matrix. Assume all entries are real, so that the adjoint is equal to the transpose.

Compute the partial derivatives with respect to each element $a_{j,k}$ of each of the following, and assemble into a matrix: $tr(UA)$, $tr(AV)$, $tr(A^*WA)$, and $tr(AZA^*)$.

Q2: Standard regression

Setup for standard regression: You are given a matrix of measured data, Y , of n rows and k columns: each row is a measurement, univariate if $k = 1$ and multivariate if $k > 1$. You are also given a set of p regressors, a matrix X , of n rows and p columns. The goal is to fit the measured data with $Y^{fit} = XB$, where B , the matrix to be found, is the set of regression coefficients. Standard regression determines B in the model $Y^{fit} = XB$ by minimizing $R_{std}^2 = tr((Y - Y^{fit})^*(Y - Y^{fit}))$.

Use the results of Q1 to find B by setting $\frac{d}{dB} R_{std}^2 = 0$.

Q3: Ridge regression

Ridge regression uses the same set-up as standard regression, but puts a premium on finding models in which the regression coefficients are small (in addition to how well they fit the data). The extent to which the criterion of "small regression coefficients" is weighted is given by a parameter k . More formally, ridge regression determines B by minimizing $R_{ridge}^2 = R_{std}^2 + k \cdot tr(B^*B) = tr((Y - Y^{fit})^*(Y - Y^{fit})) + k \cdot tr(B^*B)$. Use the results

of Q1 and Q2 to find B by setting $\frac{d}{dB} R_{ridge}^2 = 0$.