

STAT505 Assessment #11

1. J&W Exercise 10.18. The data are found in “paper.dat” with columns $y_1 - y_4$ and $z_1 - z_4$ (described in Exercise 7.26).

The data contain measurements on characteristics of pulp fibers and the paper made from them. To correspond with the notation in this chapter, let the paper characteristics be (these are the y 's)

$x_1^{(1)}$ = breaking length

$x_2^{(1)}$ = elastic modulus

$x_3^{(1)}$ = stress at failure

$x_4^{(1)}$ = burst strength

and the pulp fiber characteristics be (these are the z 's)

$x_1^{(2)}$ = arithmetic fiber length

$x_2^{(2)}$ = long fiber fraction

$x_3^{(2)}$ = fine fiber fraction

$x_4^{(2)}$ = zero span tensile

Determine the sample canonical variates and their correlations. Are the first canonical variates good summary measures of their respective sets of variables? Explain. Test for the significance of the canonical relations with $\alpha = .05$. Interpret the significant canonical variables.

2. For a sample of 50 sales representatives, measures of sale performance and scores on tests of reasoning and intellectual are recorded. In this analysis we look at how two “sales” variables 1) sales growth and 2) sales profitability relate to three “test” variables 1) mechanical reasoning, 2) abstract reasoning, and 3) mathematics.
 - (a) What is the value of the first canonical correlation?
 - (b) Write the equations that give the first (unstandardized) pair of canonical variables. You can round the coefficients to three decimal places.
 - (c) Suppose that a sales person has data values of sales growth = 93, profitability = 96, mechanical reasoning = 12, abstract reasoning = 10 and math ability = 20. Compute the canonical variable values for the first pair of unstandardized canonical variables. Don't worry about centering.
 - (d) What is the result of the test that all canonical correlations equal 0? What is the p -value and conclusion?
 - (e) What is the result of the test that the second canonical correlation equals 0? What is the p -value and conclusion?
 - (f) What are the correlations between the sales variables (the x -variables) and the first canonical variable that is constructed from them?
 - (g) What are the correlations between the test variables (the y -variables) and the first canonical variable that is constructed from them?
 - (h) Why are there only two canonical correlations in this situation?
 - (i) Write a brief conclusion about the relationship between the two sets of variables.