HW3 AA & CCA

0859605 陳冠景

Q: Run canonical correlation analysis for PC ability and three achievement score.

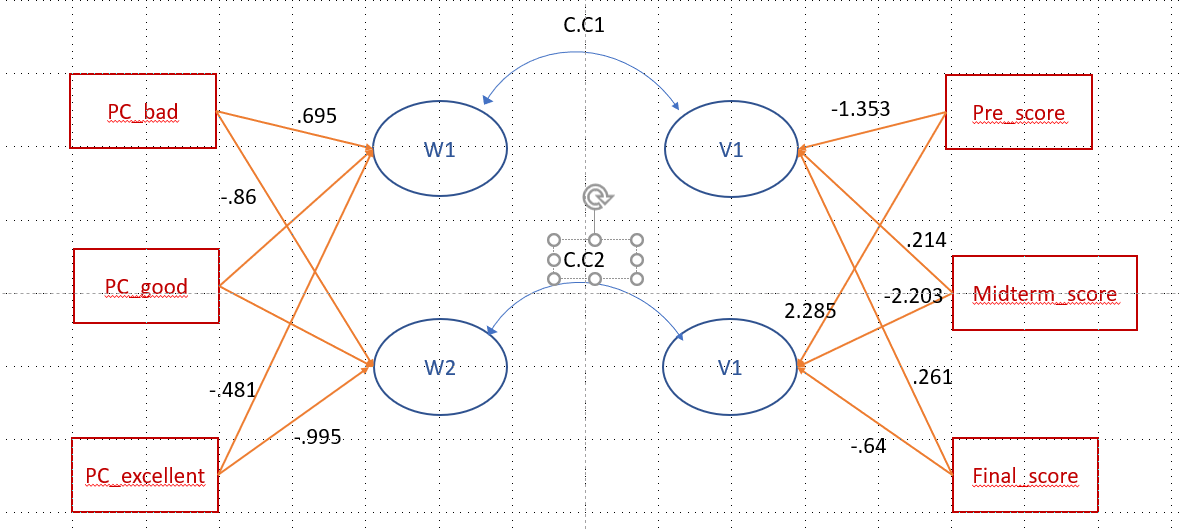
Since the variable , PC\_Ability is categorical, we need to transform it to numeric variable by one hot encoding. With this kind of transformation , we only need two levels to represent the PC\_Ability variable sets, since the last one is decided when the other 2 levels are known.

As a result, the bad level and the excellent level are chosen in variable sets of PC\_ability and are used to determine how much portion of variance in the PC\_Ability is explained by the score set(pre\_score,midterm\_score,final\_score).

Use cancor(X,Y) funtion in R to compute the significance of canonical correlation. The result shows that first canonical correlation plus second canonical correlation are significant (CanR =.915, Wilk’sλ=.14 ,*F* = 5.569 , *p* = .002),which means the PC\_Ability and the scores are significantly related. The second canonical correlation is not reported duo to its insignificant wilk’s test result and low explained variance (<1%). Using SEM to compute whether the separate canonical correlation coefficient is significant or not.

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Table 1. | | | | | | | | | | | | |
| The canonical correlation analysis of PC ability and achievement scores. | | | | | | | | | | | | |
|  | Canonical Variate 1 | | | |  | Canonical Variate 2 | | | | | |  |
| Constructs | *Coef* | *rs* | *rs2* | *rc* |  | *Coef* | *rs* | | *rs2* | | *rc* |  |
| PC\_Ability |  |  |  |  |  |  |  | |  | |  |  |
| Bad | .695 | -.9 | .81 |  |  | -.86 | -.435 | | .189 | |  |  |
| excellent | -.481 | .777 | .604 |  |  | -.995 | -.629 | | .395 | |  |  |
|  |  |  |  | .915 |  |  |  | |  | | .366 |  |
| Scores |  |  |  |  |  |  |  | |  | |  |  |
| Pre\_score | -1.353 | .98 | .96 |  |  | 2.285 | -.148 | | .022 | |  |  |
| Midterm\_score | .214 | .859 | .738 |  |  | -2.203 | -.505 | | .255 | |  |  |
| Final\_score | .261 | .542 | .294 |  |  | -.64 | -.351 | | .123 | |  |  |
|  |  |  |  |  |  |  |  | |  | |  |  |
| Explained Variance(%) |  | 5.176  (97.09%) | |  |  |  | | .155  (2.91%) | |  | |  |
| Wilk’s Lambda  (*F*,*p*-value) |  |  | .14(5.569, .002) | |  |  | |  | |  | |  |

*Note. Coef* = standardized coefficient, *rs*=structure coefficient , *rs2* = squared structure coefficient , *rc* =canonical correlation coefficient. Explained variance is based on eigenvalues.



The figure above shows the correlation relationship between 2 sets of variables.

The redundancy index for PC ability and scores can represent what percent of the total variance of a set of variables could be explained by the variance of other set of variables.

The squared of CR2(.134) multiply by the explained variance of the variate(2.91%) to get the RI value , .039 for PC ability, which represents 3.9% PC ability could be explained by the variate of a sets of achievement scores. (but CR2 is not significant.)

|  |  |  |
| --- | --- | --- |
| Table 2. | | |
| Redundancy index explained by the opposite canonical variate. | | |
|  | CanR1 | CanR2 |
| PC\_ability | .593 | .039 |
| Achievement scores | .557 | .018 |