Psychology 594 - Fall 2013 - Instructions for Final
Larry Hubert & Ehsan Bokhari & Justin Kerm

Your final is expected to be typed. You may turn in a hardcopy version or submit an electronic version via email to kern4@illinois.edu. If you email it, write in the subject line "Multivariate Take-home Final" Please do not include any MATLAB code or output. The instructions below indicate what should be reported in your final for each question.

1. Start by describing the data you selected from Michael Lee's website. Also, indicate any pre-processing steps that were needed. For example, did you have to convert a similarity matrix into a dissimilarity matrix? If so, describe the process that was used. NOTE: You may use any dataset except the abstract numbers data (`abstractnumbers.mat`) and the adolescent drug use data (`druguse.mat`).

   (a) Describe how you used the functions (e.g., "First, I input the abstract numbers dissimilarity matrix into the `squareform.m` function to create a vector containing the unique entries in the dissimilarity matrix. Next, I input this vector of dissimilarities into the ... "). Also, include the figure that is produced by the `dendrogram.m` function.

   (b) For all of these questions, describe (in your own words) what you are inputting into each of the functions and what outputs you get from each function. I want to ensure that you understand what each function is doing. See the below instructions for the additional output that you should report from each function.

       i. `order.m`:
          Report the obtained optimal ordering of your objects.

          `ultrafit.m`:
          Report the variance-accounted-for by the fit ultrametric matrix.

          `ultrafnd.m`:
          Report the number of random starting permutations that you used and the variance-accounted-for by the best ultrametric matrix.

       ii. `ultrafnd_confit.m`:
          Report the variance-accounted-for by the ultrametric matrix fit to the dissimilarity matrix, the variance-accounted-for by the anti-Robinson matrix fit to the dissimilarity matrix, and the variance-accounted-for by the ultrametric matrix fit to the anti-Robinson matrix.

          `ultrafnd_confnd.m`:
          Report the same things that you reported for the previous problem. Also, report the obtained optimal ordering of the objects.

       iii. `partitionfnd_averages.m` followed by `partitionfit.m`:
          Report the variance-accounted-for by the fit proximity matrix. Also, describe the three memberships (partitions) that were most influential in forming the fit proximity matrix (i.e., describe the three input memberships that resulted in the largest

obtained weights).

`partitionfnd_diameters.m` followed by `partitionfit.m`:
Report the same things that you reported for the previous problem.

iv. `cent_ultrafnd confit.m`:
Report the variance-accounted-for by the fit proximity matrix.

`cent_ultrafnd_confnd.m`:
Report the variance-accounted-for by the fit proximity matrix and the obtained optimal ordering of the objects.

v. `atreefit.m`:
Report the variance-accounted-for by the fit proximity matrix.

`atreefnd.m`:
Report the number of random starting permutations that you used and the variance-accounted-for by the best proximity matrix.

`atreedec.m`:
There is nothing additional to report here.

`ultraorder.m`:
Report the obtained optimal ordering of your objects according to the additive tree's ultrametric component.

vi. `consec_subsetfit.m`:
Report the variance-accounted-for by the fit proximity matrix. Also, describe the three memberships (partitions) that were most in influential in forming the fit proximity matrix.

`consec_subsetfit alter.m`:
Report the same things that you reported for the previous problem.

2. NO SCRIPT FILE. Instead, include the five figures created in the script file, along with a description of each.

3. Report which model (continuous vs. categorical) you choose and why. Also, interpret the chosen model to say a few things about the structure of the Supreme Court.