Multidimensional Scaling for Selecting Small Groups in College Courses

Iain PARDOE

Many college courses use group work as a part of the learning and evaluation process. Class groups are often selected randomly or by allowing students to organize groups themselves. However, if it is desired to control some aspect of the group structure, such as increasing schedule compatibility within groups, multidimensional scaling can be used to form such groups. This article describes how this has been adopted in an undergraduate statistics course. Resulting groups have been more homogeneous with respect to student schedules than groups selected randomly—an example from winter quarter 2004 increased correlations between student schedules from a mean of .29 before grouping to a within-group mean of .50. Further, the exercise allows opportunities to discuss a wealth of statistical concepts in class, including surveys, association measures, multidimensional scaling, and statistical graphics.

KEY WORDS: Cooperative learning; Ordination; Perceptual mapping; Principal coordinate analysis; Teaching.

1. INTRODUCTION

I have used group work in my undergraduate statistics course for a number of years, and have found it to be a useful method for improving student learning by raising student interest and increasing class participation. As well as working together during class time, students work extensively together in groups outside of class on homework assignments and projects. In managing small groups in the classroom—in my case groups of 3–5 students in a class of approximately 60—I have experimented with various ad-hoc methods for selecting the groups, each of which have had their drawbacks. For example, randomly selecting groups has led to frequent student complaints that they have difficulties meeting as a group outside of class time due to incompatible schedules. On the other hand, allowing students to self-select groups has tended to produce groups of friends in which there is very little diversity (gender, age, and ethnicity, as well as academic ability). In an attempt to balance the conflicting goals of selecting groups whose members have mostly similar schedules while at the same time maintaining group diversity, I have developed a method for using multidimensional scaling (MDS) to accomplish this task.

The use of small groups in college courses stems from the concept of cooperative learning, whereby groups of, say, three to five students work together as a team to solve a problem or complete an assignment [see Garfield (1993); Giraud (1997); Keeler and Steinhorst (1995); and Magel (1998) for examples in the field of statistics]. The National Council of Teachers of Mathematics (2000) and the National Research Council (1989) advocated cooperative learning in elementary and secondary education, while Johnson, Johnson, and Smith (1991) and Garfield (1993) extolled the virtues of cooperative learning in the college classroom. Johnson et al. (1991) showed that when students work together, they often accomplish more, and at a higher level, than they could individually. Garfield (1993, par. 8) cited published research that suggests that "the use of small group learning activities leads to better group productivity, improved attitudes, and sometimes, increased achievement."

With reference to the question of how to select cooperative learning groups, Garfield (1993, par. 18) noted that "the instructor may allow students to self-select groups or groups may be formed by the instructor to be either homogeneous or heterogeneous on particular characteristics (e.g., grouping together all students who received A's on the last quiz, or mixing students with different majors)." The remainder of this article describes how to use MDS to select groups to be homogeneous on student schedules. The method also enables inclusion of further criteria for group selection, such as making sure that each group has at least one member with a particular skill. The method is described in sufficient detail that it can be applied to any course in which cooperative learning groups are used, and can easily be adapted to work with characteristics other than student schedules.

The next section briefly describes MDS and how it can be applied to the problem of grouping students with similar schedules. The following section presents results from my undergraduate course in the winter quarter of 2004 and an evaluation of how well the method worked. Presenting and evaluating the results also provides valuable opportunities for covering a number of statistical concepts in class, including surveys, association measures, multidimensional scaling, and statistical graphics. The final section contains a discussion.

2. MULTIDIMENSIONAL SCALING FOR SELECTING GROUPS

MDS is a series of methods for displaying a set of objects in low-dimensional space (often 2D) that reflects similarities between the objects (see Kruskal and Wish 1978 for an overview). MDS results can be used to create a 2D map where the physical distances between the objects on the map are meant to correspond closely with the measured object similarities. The input to the procedure is a matrix representing pairwise similarities

Iain Pardoe is Assistant Professor, Department of Decision Sciences, Charles H. Lundquist College of Business, University of Oregon, Eugene, OR 97403–1208 (E-mail: ipardoe@lcbmail.uoregon.edu). The author thanks Chuck Kalnbach, Ron Bramhall, and Anne Forrestel of the Lundquist College of Business Leadership and Communications Center for helpful discussions, and the editor and an associate editor for helpful comments on an earlier version of this article.

 DOI: 10.1198/000313004X5680

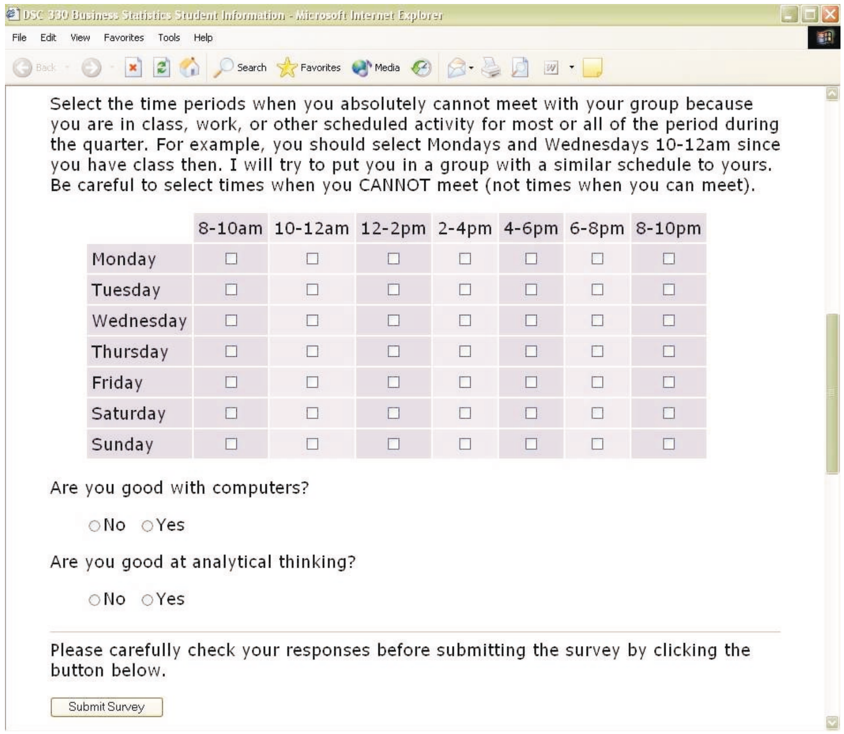

Figure 1. Online survey for collecting student information including schedule availabilities.

or proximities between the objects. This matrix can be obtained directly, for example in marketing surveys where consumers compare products by ranking or rating their similarities, or indirectly, by calculating an association measure for the objects based on the values of one or more covariates. For this application, the goal is to identify groups of students that have similar schedules, so the latter approach using calculated associations between students based on their schedules is appropriate.

For a metric MDS analysis (Torgerson 1958)—when the similarity measures are interval- or ratio-scaled—an eigenvalue decomposition of (a transformation of) the matrix of similarity measures provides the coordinate locations of each object on the map. This approach derives from the geometric relationship between the scalar product of a pair of vectors (representing two objects in the dataset) and the distance between the objects. The eigenvalue decomposition solves for the scalar products between all pairs of objects, which in turn represent the map coordinates of the objects. Similar to a principal components analysis of a covariance matrix, the first eigenvector represents the first coordinate axis (which exhibits the largest variance in distances), the second eigenvector represents the second coordinate axis, and so on. When, as is often the case, the first two eigenvalues account for the bulk of the variance in the distances, a 2D solution is adequate for mapping the objects effectively.

When the similarities are ordinal measures, a nonmetric MDS analysis (Shepard 1962) is more appropriate. Nonmetric MDS places the objects on the map to preserve a monotonic relationship between the observed similarities and the distances calculated from the map. A measure of "badness of fit," or "stress," summarizes how far a proposed solution is from the desired monotonic relationship, and then a numerical algorithm reconfigures the coordinates of the objects to iteratively minimize the stress measure. Although a goal of both metric and nonmetric MDS analyses is to produce maps that effectively represent similarities between objects, they operate on different data scales and use fundamentally different approaches to accomplish this task.

Because it is possible to calculate a ratio-scaled measure of association between student schedules, as shown in the following, a metric MDS approach is appropriate in this application. The process for mapping students using a metric MDS based on their schedules begins with collecting relevant data. Schedule information can be obtained in a number of different ways, including pen-and-paper surveys, e-mail (with or without an attached spreadsheet to enter data), and online surveys. In a statistics class there is clearly an opportunity here to engage the class in a discussion of data collection techniques and surveying. For the undergraduate statistics course that I teach, I ask the students to complete an online survey before the second class of the quarter. I use WebSurveyor software—see www.websurveyor.com—which makes this process particularly easy, resulting in a spreadsheet containing the student data. Figure 1 shows a screenshot of the part of the survey that is relevant to this article (I also collect contact information so that once groups are formed, students have all the information they need to work together effectively).

I ask the students to "select the time periods when you absolutely cannot meet with your group because you are in class, work, or other scheduled activity for most or all of the period during the quarter." I use two-hour time periods between 8:00 a.m. and 10:00 p.m. for all days including weekends; this seems to provide a reasonable compromise between survey burden and information quality. The resulting spreadsheet records a row of zeros and ones for each student showing when they are available and unavailable to meet for group work outside of class time. This spreadsheet also records a unique identification number for each student, and binary indicators for whether a student con-

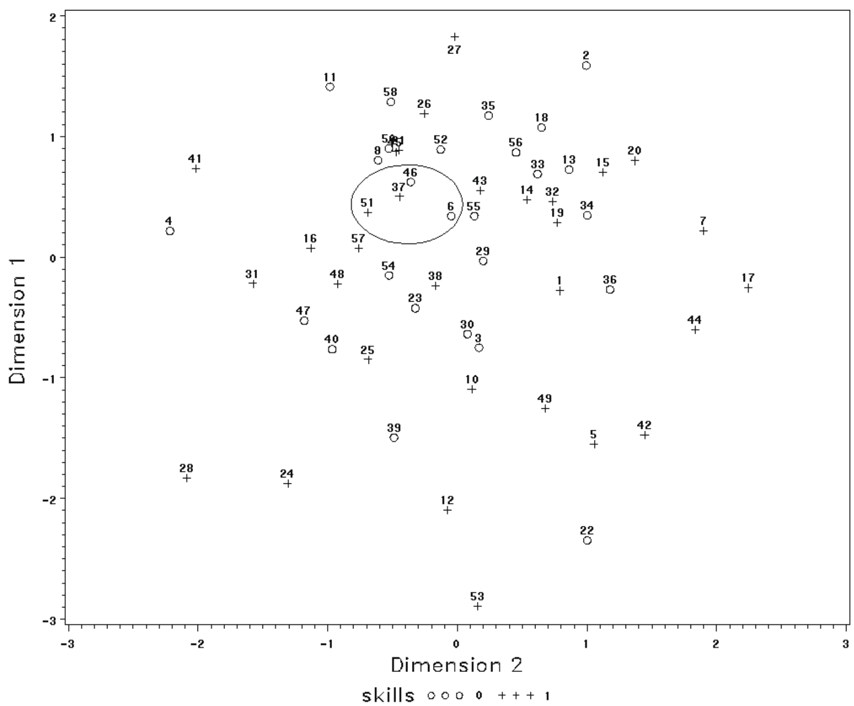

Figure 2. Perceptual map showing the students labeled by identification number and marked according to whether they have computing/analytical skills (circles = no; crosses = yes). Physical distances between the students on the map are meant to correspond closely with the similarities between their schedules. The ellipse shows an example of a selected group.

siders themselves to be "good with computers" and "good at analytical thinking."

Administering the survey and explaining its purpose provides a number of opportunities for discussing statistical concepts in class, for example, how to measure the similarity of one student's schedule with another. This naturally leads to the notion of an association measure, in this case for data that can be summarized in a two-by-two contingency table: the two row categories are the counts of the available/unavailable designations for one student while the column categories do the same for another student. For example, a typical table for a pair of students is the following:

		Student A	
		Available	*Unavailable*
Student	*Available*	27	7
B	*Unavailable*	9	6

One way to measure association in this table is to look at a scaled difference between concordant pairs of time periods (where students match and are either both available or both unavailable) and discordant pairs (where students do not match)—Kendall's τ_b (Kendall 1945) is one such quantity:

$$\tau_b = \frac{P - Q}{((P + Q + X_0)(P + Q + Y_0))^{1/2}},$$

where P is the number of concordant pairs ($27 \times 6 = 162$ for the table above), Q is the number of discordant pairs ($7 \times 9 = 63$), X_0 is the number of pairs tied on "Student A" ($27 \times 9 + 7 \times 6 = 285$), Y_0 is the number of pairs tied on "Student B" ($27 \times 7 + 9 \times 6 = 243$). Thus, τ_b for the pair of students in the table above is $99/(510 \times 468)^{1/2} = .203$. For two-by-two tables, Kendall's τ_b is algebraically the same as the usual (Pearson product-moment) correlation between each student's availabilities, and so this might usefully lead into a discussion of bivariate correlation in more general situations.

Returning to the group formation task, statistical software can then be used to perform an MDS analysis. I use SAS software, although many other common software packages could be used. The SAS code used in this article is available at lcb1.uoregon.edu/ipardoe/research.htm. In particular, I first use `proc corr` to produce a matrix of correlations representing pairwise similarities between students' schedules. I then transform these correlations into differences by subtracting from one. These differences are now a ratio scale ranging from zero (identical schedules) to two (completely mismatched schedules). I next use `proc mds` (with option level = ratio) to perform a metric MDS analysis. Depending on the type of course, it may be appropriate to devote some class time at this point to delve deeper into MDS; for example, discussing differences between metric MDS and nonmetric MDS.

Output from the MDS procedure can then be used to create a 2D map where the physical distances between the students on the map are meant to correspond closely with the similarities between their schedules. Figure 2 provides an example of such a map for the 58 students in the undergraduate statistics course that I taught in winter quarter 2004. Students are labeled by their assigned identification numbers, and I have also used different plotting symbols to distinguish students claiming to have computing or analytical skills from those that do not. All that remains is to use this map to determine the groups. I have found that it is usually sufficient to print the map out, and then, by eye, delineate boundaries between groups such that there are four or five students in each group, and each group has at least one member claiming to have computing or analytical skills. A more formal clustering technique might be used at this point—providing another opportunity to introduce a statistical concept into class—but I have found little need to venture beyond informally eyeballing the map.

3. RESULTS FROM WINTER QUARTER 2004

Because the map in Figure 2 represents similarities between students' schedules (at least to the degree that the MDS analysis has been successful at representing high-dimensional data in two-dimensional space), using the map to select groups ought to produce more homogeneous groups than selecting groups randomly. For example, I used the map to place students 6, 37, 46, and 51 together (see Figure 2)—they should be expected to have more similar schedules than four students selected purely at random. In fact their within-group correlations (the $\binom{4}{2} = 6$ correlations of pairs of students within the group) ranged between .5 and .8, whereas the $\binom{58}{2} = 1{,}653$ correlations across the whole class went as low as $-.5$.

To evaluate the success of the group selection process, I therefore considered this question in more detail. One way to summarize the effectiveness of the process, and which also provides a useful lead-in to discussing graphical summaries in class, is to compare the empirical distribution of schedule correlations for all pairs of students with the empirical distribution of correlations within the final selected groups. Figure 3 displays the resulting histograms using common axis scales to ease comparison.

The distribution of correlations clearly shifts to the right, so within-group correlations tend to be higher than correlations across the class as a whole. In terms of numerical summary statistics, the within-group correlations in the lower histogram average to .50, while the all-pairs correlations in the upper histogram have a mean of .29. There are relatively more correlations at the high end of the scale in the lower (grouped) histogram, and so in the most homogeneous groups (which tend to be the ones comprised of very tightly clustered students in Figure 2), schedules match remarkably well. However, while the very low end has been removed (no within-group correlations below $-.1$), there are some students with schedules that match few others in the class, making it difficult to place them in a homogeneous group. For example, it would have been difficult to place student 28 in a homogeneous group. The optimal group from this student's perspective would have had a minimum correlation of .1, and they actually ended up in the group with the lowest within-group correlation of $-.1$; neither are particularly homogeneous groups.

This last issue motivates another discussion question for potential use in class—why not just try to find the optimal group for each student? Apart from being extremely time consuming (particularly as the class size grows), what is optimal for one student may not be optimal for another—for example, placing student 28 in their optimal group would have changed the group membership of a number of other groups and greatly reduced their homogeneity.

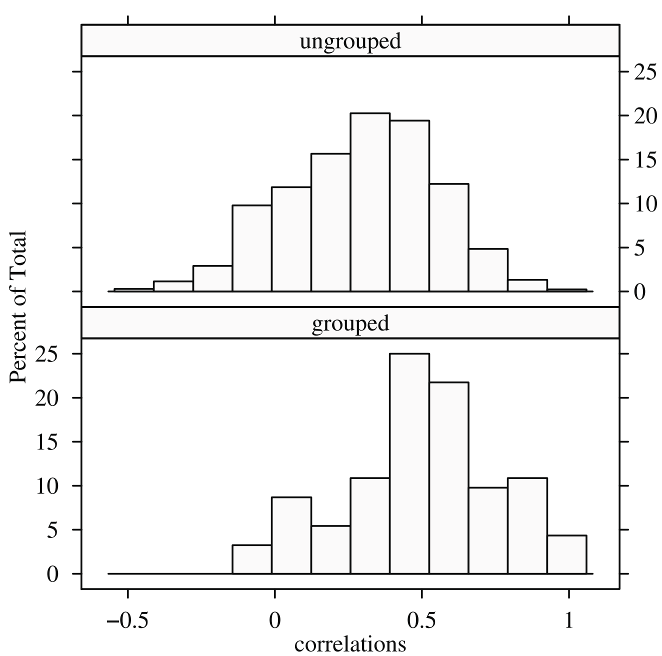

Figure 3. Histograms of correlations before and after forming groups.

4. DISCUSSION

When making extensive use of small groups in college courses, the way in which the groups are selected can impact operational aspects, such as how well groups are able to schedule times to meet, as well as personal characteristics, such as student diversity within groups. To avoid problems that can arise when groups are randomly assigned or students are allowed to self-select, the process described in this article uses multidimensional scaling for selecting groups that aims to increase scheduling homogeneity within groups without reducing diversity. The approach has worked successfully in classes in which I have adopted it, and student feedback has been very favorable. In theory, within group diversity remains as high as for random group selection, although to the extent that "similar students" have similar schedules, this may not strictly hold in practice. The approach can also take into account additional student characteristics, such as making sure each group contains at least one student with a particular skill. The method as described uses Pearson correlations (equivalent to Kendall's τ_b in this case) to calculate student schedule similarities, but it should be possible to adapt the method to use alternative association measures; for example, if it was desired to consider matches on times when students can meet as more important than matches on times when students cannot meet. An added benefit, when applied to statistics courses, is that the approach offers many opportunities for exploring statistical concepts in class.

I have focused on using the approach to enhance within-group schedules, but the approach is general enough to consider other characteristics too, such as student grades or majors (as mentioned by Garfield 1993), or skills. For example, suppose it is desired to form groups in which students have a wide variety of skills and competencies (such as analytical, writing, organizing, public speaking, leadership, computing, detail-oriented, good with "big-picture," etc.). Scores (from 1 to 10, say) on each of these skills can be obtained via a student survey, and then a metric MDS analysis can be based on pairwise correlations between student scores. In this case, the goal would be to group students with negative correlations so that each group has a good mix of skills. Thus, the appropriate differences matrix to be analyzed would comprise "correlations plus one"—values close to zero would correspond to students with very different skills who could benefit from being placed in the same group, while values close to two would correspond to students with similar skills who could benefit from being placed in different groups.

The general approach does involve some start-up costs. These include: collecting the relevant information from students (e.g., scheduling information provided through an online survey); analyzing the data (e.g., using `proc mds` in SAS to create a map of students in which those close together have similar schedules and those far apart have very different schedules); using the map to select groups (for example, informally clustering the students by eye). Once the survey and SAS code have been written however, it is easy to implement the approach for any class where homogeneous groups need to be formed.

[Received June 2004. Revised August 2004.]

REFERENCES

Garfield, J. (1993), "Teaching Statistics Using Small-Group Cooperative Learning," *Journal of Statistics Education* [on-line], 1. Available at www.amstat.org/publications/jss/.

Giraud, G. (1997), "Cooperative Learning and Statistics Instruction," *Journal of Statistics Education* [on-line], 5. Available at www.amstat.org/publications/jss/.

Johnson, D., Johnson, R., and Smith, K. (1991), "Cooperative Learning: Increasing College Faculty Instructional Productivity," in *ASHE-ERIC Higher Education Report No. 4*, Washington, DC: The George Washington University, pp. TKKK–TKKK.

Keeler, C. M., and Steinhorst, R. K. (1995), "Using Small Groups to Promote Active Learning in the Introductory Statistics Course: A Report From the Field," *Journal of Statistics Education* [on-line], 3. Available at www.amstat.org/publications/jss/.

Kendall, M. G. (1945), "The Treatment of Ties in Rank Problems," *Biometrika*, 33, 239–251.

Kruskal, J. B., and Wish, M. (1978), "Multidimensional Scaling," in *Sage University Paper on Quantitative Applications in the Social Sciences*, Newbury Park, CA: Sage Publications.

Magel, R. C. (1998), "Using Cooperative Learning in a Large Introductory Statistics Class," *Journal of Statistics Education* [on-line], 6. Available at www.amstat.org/publications/jss/.

National Council of Teachers of Mathematics (2000), *Principles and Standards for School Mathematics*, Reston, VA: NCTM.

National Research Council (1989), *Everybody Counts: A Report to the Nation on the Future of Mathematics Education*, Washington, DC: National Academy Press.

Shepard, R. N. (1962), "The Analysis of Proximities: Multidimensional Scaling with an Unknown Distance Function," *Psychometrika*, 27, 125–140; 219–246.

Torgerson, W. S. (1958), *Theory and Methods of Scaling*, New York: Wiley.