## Sociology 375 Exam 1 Fall 2014 Prof Montgomery

Answer all questions. 250 points possible.

1) [70 points] Consider the relation $A=\{(1,2),(1,3),(3,2),(3,4),(4,2)\}$ on the set of individuals $\mathrm{S}=\{1,2,3,4\}$.
a) Show how the relation A could be represented
i) as a (directed) graph
ii) as an adjacency matrix
iii) using infix notation
b) Determine (by computation or inspection)
i) the number of 2-paths between each pair of individuals
ii) the number of 3-paths between each pair of individuals
iii) the reachability matrix
iv) the distance matrix
c) State the matrix test for transitivity. Using this test, determine whether the relation A is transitive. If not, list any violations of transitivity.
2) [20 points] Suppose you are given a symmetric adjacency matrix that generates the reachability matrix R. Now consider the matrix $U$ composed of the unique rows of the reachability matrix. (Using matlab, this matrix is computed as $\mathrm{U}=$ unique( R , 'rows').) For purposes of social network analysis, what do we learn by counting the number of rows of U (i.e., by computing size( $\mathrm{U}, 1$ ) in matlab)? What do we learn by computing the row sums of $U$ (i.e., by computing sum $\left(\mathrm{U}^{\prime}\right)^{\prime}$ in matlab)? What answer would we get if we computed the column sums of U (i.e., $\operatorname{sum}(\mathrm{U})$ in matlab)? Briefly explain.
3) [50 pts] Consider the relation on $S=\{1,2,3,4,5\}$ given by the adjacency matrix

$$
\mathrm{A}=\left[\begin{array}{lllll}
0 & 1 & 1 & 1 & 0 \\
1 & 0 & 0 & 1 & 1 \\
1 & 0 & 0 & 1 & 0 \\
1 & 1 & 1 & 0 & 1 \\
0 & 1 & 0 & 1 & 0
\end{array}\right]
$$

a) Briefly discuss the concept of "regular equivalence" and then state the matrix test.
b) Suppose we partition $S$ into the subsets $\{1,3,4\}$ and $\{2,5\}$. Use the matrix test to show that this partition is not a regular equivalence. Then briefly discuss why the regular equivalence condition fails for some particular pair of individuals.
c) For this relation, the regular equivalence condition holds if all individuals are placed in the same equivalence class, and if each individual is placed in his/her own equivalence class. Find another partition of $S$ that satisfies the regular equivalence condition.
4) [30 points] Briefly describe the main findings in the two papers by Duncan Watts (American Journal of Sociology 1999, Science 2002) that we covered in lecture. [HINT: Three or four sentences about each paper should be sufficient. You might also want to use some diagrams to describe these findings. If so, make sure that they are properly labeled.]
5) [80 points] Consider the symmetric relation $A$ on the set of individuals $\{1, \ldots, 8\}$ which is represented by the following graph:


The A relation is also represented by the adjacency matrix A on the attached sheet, which also provides some relevant matlab computations.
a) Briefly interpret the R1, R2, R3, and R4 matrices from the attached sheet.
b) List the strong cliques (i.e., the 1-cliques) for the A relation.
c) List the 2-cliques for the A relation. How could you find these from the graph of the R 2 relation?
d) List the 3-cliques for the A relation.
e) The con matrix on the attached sheet gives the (local) connectivity between each pair of individuals. Briefly discuss two implications that follow from $\operatorname{con}(4,8)=2$.
f) What is the (global) connectivity level k for the entire A graph? List the k-cutset(s).
g) Given the connectivity level k from part (f), are there any subgraphs of A that are $\mathrm{k}+1$ connected? If not, explain why. If so, list the (nodes in the) $k+1$-component(s) and give one $\mathrm{k}+1$-cutset for each $\mathrm{k}+1$-component.

## matlab computations for question 5



1a) $[15 \mathrm{pts}]$ i)

ii) $\mathrm{A}=\left[\begin{array}{llll}0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0\end{array}\right]$
iii) $1 \mathrm{~A} 2,1 \mathrm{~A} 3,3 \mathrm{~A} 2,3 \mathrm{~A} 4,4 \mathrm{~A} 2$
b) $[36 \mathrm{pts}]$ i) $\mathrm{A}^{2}=\left[\begin{array}{llll}0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0\end{array}\right] \quad$ bii) $\mathrm{A}^{3}=\left[\begin{array}{llll}0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0\end{array}\right]$
biii) reachability $(\mathrm{A})=\left[\begin{array}{llll}1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1\end{array}\right]$
biv) distance $(\mathrm{A})=\left[\begin{array}{cccc}0 & 1 & 1 & 2 \\ \infty & 0 & \infty & \infty \\ \infty & 1 & 0 & 1 \\ \infty & 1 & \infty & 0\end{array}\right]$
c) [19 pts] The matrix test for transitivity is $\mathrm{A}-\mathrm{A}^{2} \# \geq 0$.

$$
\text { Here, } \mathrm{A}-\mathrm{A}^{2} \#=\left[\begin{array}{llll}
0 & 1 & 1 & 0 \\
0 & 0 & 0 & 0 \\
0 & 1 & 0 & 1 \\
0 & 1 & 0 & 0
\end{array}\right]-\left[\begin{array}{llll}
0 & 1 & 0 & 1 \\
0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0
\end{array}\right]=\left[\begin{array}{cccc}
0 & 0 & 1 & -1 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 \\
0 & 1 & 0 & 0
\end{array}\right]
$$

and thus the relation A is not transitive. Note that 1 A 3 and 3 A 4 but not 1 A 4 .
2) [20 pts] If the adjacency matrix is symmetric, then reachability is an equivalence relation and will generate equivalence classes (called "weak cliques"). Thus, each unique row of $R$ characterizes one of the weak cliques. The number of rows of $U$ gives the number of weak cliques, and the rows sums give the number of individuals in each weak clique. Recognizing that equivalence classes partition the set of individuals, each individual belongs to exactly one weak clique. Thus, every column of $U$ contains exactly one 1 (and the matlab command sum( U ) would produce a row vector of 1 's).

3a) [15 pts] A regular equivalence is a partition of individuals into equivalence classes such that everyone in each equivalence class has ties to and from the same equivalence classes (i.e., to and from the same types of individuals). Regular equivalence is a weaker condition than structural equivalence, which requires that everyone in each equivalence class has ties to and from the same individuals (not merely the same types of individuals). Given an adjacency matrix A and equivalence relation E, the relation E is a regular equivalence when $(\mathrm{AE}) \#=(\mathrm{EA}) \#$.
b) [20 pts] Applying the matrix test,

$$
\begin{aligned}
& (A E) \#=\left(\left[\begin{array}{lllll}
0 & 1 & 1 & 1 & 0 \\
1 & 0 & 0 & 1 & 1 \\
1 & 0 & 0 & 1 & 0 \\
1 & 1 & 1 & 0 & 1 \\
0 & 1 & 0 & 1 & 0
\end{array}\right]\left[\begin{array}{lllll}
1 & 0 & 1 & 1 & 0 \\
0 & 1 & 0 & 0 & 1 \\
1 & 0 & 1 & 1 & 0 \\
1 & 0 & 1 & 1 & 0 \\
0 & 1 & 0 & 0 & 1
\end{array}\right]\right) \#=\left[\begin{array}{lllll}
1 & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 & 1 \\
1 & 0 & 1 & 1 & 0 \\
1 & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 & 1
\end{array}\right] \\
& (E A) \#=\left(\left[\begin{array}{lllll}
1 & 0 & 1 & 1 & 0 \\
0 & 1 & 0 & 0 & 1 \\
1 & 0 & 1 & 1 & 0 \\
1 & 0 & 1 & 1 & 0 \\
0 & 1 & 0 & 0 & 1
\end{array}\right]\left[\begin{array}{lllll}
0 & 1 & 1 & 1 & 0 \\
1 & 0 & 0 & 1 & 1 \\
1 & 0 & 0 & 1 & 0 \\
1 & 1 & 1 & 0 & 1 \\
0 & 1 & 0 & 1 & 0
\end{array}\right]\right) \#=\left[\begin{array}{lllll}
1 & 1 & 1 & 1 & 1 \\
1 & 1 & 0 & 1 & 1 \\
1 & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 & 1 \\
1 & 1 & 0 & 1 & 1
\end{array}\right]
\end{aligned}
$$

Because (AE) $\# \neq(\mathrm{EA}) \#, \mathrm{E}$ is not a regular equivalence. Consider individuals 2 and 3. From the matrix tests, we see that 2AE3 but not 2EA3. In words, 2 is adjacent to someone in 3's equivalence class, but there is no one in 2's class who is adjacent to 3 . Thus, E is not a regular equivalence.
c) $[15 \mathrm{pts}]$ Some possible answers are $\{\{1,2,3,5\},\{4\}\}$ or $\{\{1,2\},\{3,4,5\}\}$ or $\{\{1,4\},\{2,5\},\{3\}\}$ or $\{\{1,2,3\},\{4,5\}\}$ or $\{\{1,3\},\{2,4\},\{5\}\}$. You didn't need to draw the graph of A, but it may have helped to draw it as

4) [30 pts] In his 1999 AJS paper, Watts showed how "small-world" graphs (which have a high clustering coefficient and low average distance) can be obtained by randomly "rewiring" some of the edges of a highly structured "caveman" graph (which is highly clustered with high average distance). Intuitively, a little random rewiring creates "shortcuts" across the caveman graph, greatly decreasing average distance but only slightly decreasing the clustering coefficient.

In the 2002 Science paper, Watts et al studied the conditions under which networks are "searchable." In the classic small-world experiment conducted by Travers and Milgram, some "starter" individuals were able to forward a package to a "target" individual through a chain of acquaintances, each of whom followed the simple rule of forwarding the package to an acquaintance similar to the target (e.g., in the same city or holding the same occupation). To determine the types of networks that can be "searched" in this fashion, Watts et al first posit a "cognitive hierarchy" (which allows them to measure "social distance" between individuals based on their characteristics), and then introduce a homophily parameter governing the extent to which individuals tend to form links to others with low social distance. Numerical experiments reveal that chains (between a randomly chosen starter and target) are most likely to be completed when homophily is high enough (so that low social distance is associated with low network distance) but too high (because complete homophily would eliminate all paths between starters and targets with high social distance).

5a) [5 pts] Each Rt matrix gives reachability at step t
b) $[16 \mathrm{pts}]\{1,2,3\},\{2,3,4\},\{4,5,6\},\{5,6,7,8\}$
c) $[17 \mathrm{pts}]\{1,2,3,4\},\{2,3,4,5,6\},\{4,5,6,7,8\}$. The 2 -cliques of the A relation are the strong cliques (i.e., 1 -cliques) of the R 2 relation.
d) $[8 \mathrm{pts}]\{1,2,3,4,5,6\},\{2,3,4,5,6,7,8\}$
e) [12 pts] First, you would need to remove 2 nodes to disconnect 4 and 8. (In particular, you would need to remove the nodes in the cutset $\{5,6\}$.) Second, 4 and 8 are connected by 2 node-independent paths. (For instance, the paths $4-6-8$ and $4-5-8$.)
f) [ 6 pts] The global connectivity is given by the minimum element of the con matrix. Thus, the global connectivity of the graph of A equals 1 . The (unique) 1-cutset is $\{4\}$.
g) [16 pts] There are two different 2-components: the 2 -component $\{1,2,3,4\}$ has 2 -cutset $\{2,3\}$; the other 2 -component $\{4,5,6,7,8\}$ has 2 -cutset $\{5,6\}$.

Sociology 375 Exam 2 Fall 2014 Prof Montgomery
Answer all questions. 250 points possible.

1) [30 points] Consider a social relation $A$ on a set of 5 individuals. The corresponding adjacency matrix (along with its eigenvectors and eigenvalues) is given by

| $A=$ |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: |
|  |  |  |  |  |
| 0 | 1 | 0 | 0 | 0 |
| 1 | 0 | 1 | 0 | 0 |
| 0 | 1 | 0 | 1 | 0 |
| 0 | 0 | 1 | 0 | 1 |
| 0 | 0 | 0 | 1 | 0 |

>> [eigvec, eigval] = eig(A)

| eigvec $=$ |  |  |  |  |
| ---: | ---: | ---: | ---: | ---: |
|  |  |  |  |  |
| 0.2887 | -0.5000 | 0.5774 | 0.5000 | 0.2887 |
| -0.5000 | 0.5000 | 0.0000 | 0.5000 | 0.5000 |
| 0.5774 | 0.0000 | -0.5774 | 0.0000 | 0.5774 |
| -0.5000 | -0.5000 | -0.0000 | -0.5000 | 0.5000 |
| 0.2887 | 0.5000 | 0.5774 | -0.5000 | 0.2887 |

eigval $=$

| -1.7321 | 0 | 0 | 0 | 0 |
| :---: | ---: | ---: | ---: | ---: |
| 0 | -1.0000 | 0 | 0 | 0 |
| 0 | 0 | -0.0000 | 0 | 0 |
| 0 | 0 | 0 | 1.0000 | 0 |
| 0 | 0 | 0 | 0 | 1.7321 |

a) State the equation for the Bonacich centrality measure. Using the A matrix above, compute this measure given $\alpha=1$ and $\beta=0$. Given these parameter values, which individual(s) are the most "central"?
b) What is the substantive interpretation of the parameter $\beta$ in the Bonacich centrality measure? For the example above, what is the upper bound for $\beta$ ? What is the ranking of individuals as $\beta$ approaches this upper bound? How do you know this?
2) [40 points] Consider a kinship system in a society with three clans characterized by the matrices

$$
\mathrm{W}=\left[\begin{array}{lll}
0 & 1 & 0 \\
0 & 0 & 1 \\
1 & 0 & 0
\end{array}\right] \quad \mathrm{C}=\left[\begin{array}{lll}
1 & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & 1
\end{array}\right]
$$

where $\mathrm{W}(\mathrm{i}, \mathrm{j})=1$ indicates that a man in clan i must marry a woman in clan j , and $\mathrm{C}(\mathrm{i}, \mathrm{j})=1$ indicates that a man in clan i has children in clan j . Is this society g -balanced? If so, show that the g-balance condition is satisfied, and briefly discuss the implications. If not, explain why the g-balance condition doesn't hold, and briefly discuss the implications.
3) [60 points] Suppose that six actors $(1,2,3,4,5,6)$ attended three events (A, B, C), with the attendance pattern given by the matrix

$$
P=\left[\begin{array}{lll}
0 & 0 & 1 \\
0 & 1 & 0 \\
0 & 1 & 1 \\
1 & 1 & 0 \\
1 & 1 & 1 \\
1 & 0 & 1
\end{array}\right]
$$

a) In his paper on duality, Breiger (1974) shows how an actor-by-event matrix might be used to derive an actor-by-actor matrix and an event-by-event matrix. Give the formula for each matrix. Then, using the P matrix above, compute the actor-by-actor matrix and interpret the elements of this matrix.
b) We might say that actor i "contains" actor j if the set of events attended by i contains the set of events attended by j . Given the P matrix above, report the adjacency matrix for the strict containment relation on the set of actors, and then draw the Hasse diagram. [HINT: There are simple matrix procedures for deriving the containment matrix and then the Hasse matrix. But to save time, you can simply complete each step by inspection (without doing all the matrix algebra) if you prefer.]
c) Below is a Galois lattice (oriented so that the universal set of events is at the top). Explain why this lattice could not have been generated by the P matrix above.

d) Construct the Galois lattice that is generated by the P matrix above. Use reduced labeling for both actors and events.
4) [60 points] Using the P matrix from question 3, we may compute the correlation coefficient matrix X (which results from comparing each pair of columns in P ). The X matrix is given below, along with its eigenvectors and eigenvalues.

```
>> \(\mathrm{X}=\operatorname{corrcoef}(\mathrm{P})\)
\(X=\)
    \(\begin{array}{rrr}1.0000 & 0 & 0\end{array}\)
        \(0 \quad 1.0000-0.5000\)
        \(0-0.50001 .0000\)
\(\gg\) [eigvec, eigval] \(=\operatorname{eig}(X)\)
eigvec \(=\)
\begin{tabular}{rrr}
0 & 1.0000 & 0 \\
-0.7071 & 0 & -0.7071 \\
-0.7071 & 0 & 0.7071
\end{tabular}
eigval \(=\)
\begin{tabular}{rrr}
0.5000 & 0 & 0 \\
0 & 1.0000 & 0 \\
0 & 0 & 1.5000
\end{tabular}
```

a) Find the $(3 \times 1)$ column vector $v$ such that the outer product $v v^{\prime}$ gives the best leastsquares approximation of the X matrix. [HINT: The inner product $\mathrm{v}^{\prime} \mathrm{v}$ is equal to an eigenvalue.]
b) Using the $v$ vector from part (a), compute the estimated matrix $E_{1}=v v^{\prime}$, and then compute the residual matrix $\mathrm{R}_{1}=\mathrm{X}-\mathrm{E}_{1}$. Use the X matrix to compute the total sum of squares (TSS), and then use the $\mathrm{R}_{1}$ matrix to compute the residual sum of squares $\left(\mathrm{RSS}_{1}\right)$. The difference between TSS and $\mathrm{RSS}_{1}$ is the explained sum of squares $\left(E S S_{1}\right)$. How is the $\mathrm{ESS}_{1}$ related to an eigenvalue? [HINT: If you're correct, you should be able to easily verify your answer from the matlab computations supplied above.]
c) Find the $(3 \times 1)$ column vector $u$ such that the outer product $E_{2}=u u^{\prime}$ gives the best least-squares approximation of the $\mathrm{R}_{1}$ matrix. [HINT: The inner product $\mathrm{u}^{\prime} \mathrm{u}$ is equal to an eigenvalue.]
d) What is the interpretation of the vectors $v$ and $u$ for factor analysis? Show how they are used to produce a scatterplot. [NOTE: Your scatterplot doesn't need to be perfect, but you should label the points and axes appropriately, and indicate numerical coordinates. The labels for the points are taken from question 2.]
5) [60 points] The attached sheet gives adjacency matrices for a positive relation $P$ and a negative relation N on a set of 20 individuals. Following the version of balance theory we covered in class, these relations are symmetric (in matlab notation, $\mathrm{P}=\mathrm{P}^{\prime}$ and $\mathrm{N}=\mathrm{N}^{\prime}$ ) and disjoint (in matlab notation, the matrix $\mathrm{P} \& N$ would be a $20 \times 20$ matrix of 0 s). Further assume that the signed graph of $(\mathrm{P}, \mathrm{N})$ is complete in the sense that every edge between distinct nodes is either positive or negative. I further assume that both $P$ and $N$ are antireflexive. Thus, the matrix $\mathrm{P}+\mathrm{N}$ is a $20 \times 20$ matrix with 0 s on the main diagonal and 1 s everywhere else (as shown on the attached sheet).
a) Given 20 individuals, what is the total number of dyads in a dyad census? What is the total number of triads in a triad census?
b) Given the maintained assumptions on symmetry and completeness, there are 2 types of dyads and 4 types of triads in the signed graph of ( $\mathrm{P}, \mathrm{N}$ ). Use the computations below to conduct a dyad census and then a triad census. [HINT \#1: Do not attempt to use the P and N matrices directly. They are provided merely to help you conceptualize the problem. The dyad and triad censuses can be conducted using only the computations provided below.] [HINT \#2: If you've correctly computed these censuses, the total number of dyads and triads should match your answers to part (a).]
c) Based on the dyad census, compute the expected number of triads of each type following the logic of Davis and Leinhardt. [HINT: To simplify your calculations, you can follow Davis and Leinhardt's assumption that sampling occurs with replacement.]
d) According to balance theory, which types of triads are prohibited? Comparing the actual numbers of triads to the expected numbers of triads, does this data set seem to strongly support balance theory? Briefly discuss.

## matlab computations for problem 5

```
>> trace(P*P)
ans = 272
>> trace(N*N)
ans = 108
>> trace(P*P*P)
ans = 2496
```

```
>> trace( }\textrm{P}*\textrm{P}*\textrm{N}
```

>> trace( }\textrm{P}*\textrm{P}*\textrm{N}
ans = 1000
ans = 1000
>> trace(P*N*N)
>> trace(P*N*N)
ans = 400
ans = 400
>> trace(N*N*N)
>> trace(N*N*N)
ans = 144

```
ans = 144
```




## Sociology 375 Exam $2 \quad$ Fall $2014 \quad$ Solutions

1a) [14 pts] The formula may be written $c=\alpha\left[A+\beta A^{2}+\beta^{2} A^{3}+\ldots\right] 1$, where 1 is understood as a column vector of 1 s . Given $\alpha=1$ and $\beta=0$, the centrality vector is simply given by the row sums of the A matrix. For the A matrix in this problem,

$$
c(1,0)=\left[\begin{array}{l}
1 \\
2 \\
2 \\
2 \\
1
\end{array}\right]
$$

which implies that actors 2,3 , and 4 are most "central."
b) [16 pts] The parameter $\beta$ reflects the weight placed on longer paths (with weight $\beta^{\mathrm{t}-1}$ on paths of length t ). The upper bound for $\beta$ is given by the reciprocal of the largest eigenvalue of the A matrix. For this problem, this upper bound is $1 / 1.7321=0.577$. If $\beta$ was set above this value, the infinite sum $\left[A+\beta A^{2}+\beta^{2} A^{3}+\ldots\right]$ would not converge. As $\beta$ approaches its upper bound, the centrality vector is proportional to the leading eigenvector (given by the final column of the eigvec matrix). Thus, actor 3 is the most central (no longer tied with actors 2 and 4).
2) $[40 \mathrm{pts}]$ Given $\mathrm{W}^{\circ} \mathrm{W}=\mathrm{WW}$, the W and C matrices generate the group G with elements $\{\mathrm{W}, \mathrm{C}, \mathrm{WW}\}$ and multiplication table

|  | C | W | WW |
| :--- | :--- | :--- | :--- |
| C | C | W | WW |
| W | W | WW | C |
| WW | WW | C | W |



As shown, the graph of W and C on the set of clans is isomorphic to the graph of the multiplication table of the group. Thus, the society is $g$-balanced. This implies that individuals have consistent obligations to others (every path between 2 clans has the same sign), and that every individual has a consistent self-concept (the sign of every cycle equals the identity element of the group).

3a) [15 pts] The actor-by-actor matrix is $\mathrm{PP}^{\mathrm{T}}$; the event-by-event matrix is $\mathrm{P}^{\mathrm{T}} \mathrm{P}$. For the present problem,

$$
\mathrm{PP}^{\mathrm{T}}=\left[\begin{array}{llllll}
1 & 0 & 1 & 0 & 1 & 1 \\
0 & 1 & 1 & 1 & 1 & 0 \\
1 & 1 & 2 & 1 & 2 & 1 \\
0 & 1 & 1 & 2 & 2 & 1 \\
1 & 1 & 2 & 2 & 3 & 2 \\
1 & 0 & 1 & 1 & 2 & 2
\end{array}\right]
$$

Element $\left(\mathrm{PP}^{\mathrm{T}}\right)(\mathrm{i}, \mathrm{j})$ gives the number of events attended jointly by actors i and j .
b) [20 pts] Strict containment relation

Hasse diagram

$$
\mathbf{C}=\left[\begin{array}{llllll}
0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 \\
1 & 1 & 1 & 1 & 0 & 1 \\
1 & 0 & 0 & 0 & 0 & 0
\end{array}\right]
$$


c) [5 pts] This Galois lattice implies that actor 6 attends only event A (rather than events A and C ).
d) [20 pts] The correct Galois lattice is


4a) [15 pts] The vector $v$ is given by the leading eigenvector, multiplied by the square root of the corresponding eigenvalue. Thus,

$$
\mathrm{v}=\sqrt{1.5}\left[\begin{array}{c}
0 \\
-0.7071 \\
0.7071
\end{array}\right]=\left[\begin{array}{c}
0 \\
-0.866 \\
0.866
\end{array}\right]
$$

Note that $\mathrm{v}^{\prime} \mathrm{v}=1.5=$ largest eigenvalue.
b) $[25 \mathrm{pts}]$ The estimated matrix is thus given by $\mathrm{E}_{1}=\left[\begin{array}{ccc}0 & 0 & 0 \\ 0 & 0.75 & -0.75 \\ 0 & -0.75 & 0.75\end{array}\right]$
and the residual matrix is given by $\mathrm{R}_{1}=X-\mathrm{E}_{1}=\left[\begin{array}{ccc}1 & 0 & 0 \\ 0 & 0.25 & 0.25 \\ 0 & 0.25 & 0.25\end{array}\right]$.
Squaring each element of the X matrix and summing over all elements, we obtain

$$
\mathrm{TSS}=3 * 1^{2}+2 *(-0.5)^{2}=3.5
$$

Similarly, using the $\mathrm{R}_{1}$ matrix, we obtain $\mathrm{RSS}_{1}=1^{2}+4^{*}(0.25)^{2}=1.25$
Thus, $\mathrm{ESS}_{1}=3.5-1.25=2.25$, which equals the square of the largest eigenvalue, $(1.5)^{2}$
c) [10 pts] The vector $u$ is given by the eigenvector associated with the second largest eigenvalue, multiplied by the square root of that eigenvalue. Thus,

$$
u=\sqrt{1}\left[\begin{array}{l}
1 \\
0 \\
0
\end{array}\right]=\left[\begin{array}{l}
1 \\
0 \\
0
\end{array}\right]
$$

Note that $u^{\prime} u=1=$ second largest eigenvalue.
d) [10 pts] The vector v gives the "factor loadings" for the first dimension (coordinates for x axis of scatterplot); the vector u gives factor loadings for the second dimension (coordinates for y axis of scatterplot). The scatterplot should show event A plotted at $(0,1)$, event B plotted at $(-.866,0)$, and event C plotted at $(.866,0)$.

5a) [10 pts] There are $20 * 19 / 2=190$ dyads and $20 * 19 * 18 / 6=1140$ triads.
b) [20 pts] Each dyad is either positive or negative. (Recall that there are no null dyads.)
number of positive dyads $=\operatorname{trace}\left(\mathrm{P}^{*} \mathrm{P}\right) / 2=136$
number of negative dyads $=\operatorname{trace}\left(\mathrm{N}^{*} \mathrm{~N}\right) / 2=54$
(note that $136+54=190$ as required)
The 4 types of triads are those originally identified by Heider: triads can have 3, 2, 1 , or 0 positive edges (and hence $0,1,2$, or 3 negative edges respectively).
number of triads with 3 positive edges $=\operatorname{trace}\left(\mathrm{P}^{*} \mathrm{P} * \mathrm{P}\right) / 6=416$
number of triads with 2 positive edges $=\operatorname{trace}\left(\mathrm{P}^{*} \mathrm{P} * \mathrm{~N}\right) / 2=500$
number of triads with 1 positive edge $=\operatorname{trace}\left(\mathrm{P}^{*} \mathrm{~N}^{*} \mathrm{~N}\right) / 2=200$
number of triads with 0 positive edges $=\operatorname{trace}\left(\mathrm{N}^{*} \mathrm{~N} * \mathrm{~N}\right) / 6=24$
(note that $416+500+200+24=1140$ as required)
c) [20 pts] Given the dyad census, we may construct triads "randomly" following the logic of Davis and Leinhardt (1972). As a first step, we determine the probability that each dyad is either positive or negative:
$\mathrm{p}=$ probability of positive dyad $=136 / 190=.716$
$\mathrm{n}=$ probability of negative dyad $=54 / 190=.284$
Using these probabilities, the expected probability of each type of triad is given by
probability of triad with 3 positive edges $=p^{3}=.367$
probability of triad with 2 positive edges $=3 \mathrm{p}^{2} \mathrm{n}=.437$
probability of triad with 1 positive edge $=3 \mathrm{pn}^{2}=.173$
probability of triad with 0 positive edges $=\mathrm{n}^{3}=.023$
Multiplying each of these probabilities by the number of triads, we obtain
expected number of triads with 3 positive edges $=.367 * 1140=418.4$
expected number of triads with 2 positive edges $=.437 * 1140=498.2$
expected number of triads with 1 positive edge $=.173 * 1140=197.2$
expected number of triads with 0 positive edges $=.023 * 1140=26.2$
d) [10 pts] In balance theory, the triads with 2 and 0 positive edges are prohibited. For this data set, the expected and actual numbers of triads are very close. (Given that I originally generated the P and N matrices randomly, this result was anticipated.) If balance theory was actually operating, the actual number of triads would presumably be much lower than the expected number of triads for the prohibited types of triads.

