This is to certify that the
thesis entitled

PARAMETER ESTIMATION AND INTERPRETATION IN
SPATIAL AUTOREGRESSION MODELS

presented by

JIQIANG XU

has been accepted towards fulfillment
of the requirements for the

PhD. degree in Counseling, Educational Psychology and Special Education

Cassandra Book, EDU, Associate Dean

Major Professor's Signature

1998

Date

PARAMETER ESTIMATION AND INTERPRETATION IN SPATIAL AUTOREGRESSION MODELS

BY

JiQiang Xu

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Counseling, Educational Psychology
And Special Education

1998

ABSTRACT

This dissertation explores Spatial Autoregression Models (SAM). It gives the definition of the autocorrelation coefficient ρ in SAM, supplies the technique for the computational precision for parameter estimation in SAM, and makes the SAM model applicable in practice.

Based on SAM models, the autocorrelation coefficient ρ turns out to be the correlation coefficient between a matrix W and a vector Y, a new measurement in statistics for the social sciences; the classical Factor Analysis Method is also generalized to the non-variance - covariance matrices.

To the memory of my late mother and father

ACKNOWLEDGEMENTS

The author wishes to express his sincere appreciation to his dissertation committee, Drs. B. J. Becker, K. A. Frank, D. Gilliland and S. Raudenbush for their valuable advice, guidance and help in the preparation of the manuscript.

The author is specifically indebted to his advisor Dr. K. A. Frank for the suggestion of the dissertation topic which leads to the following outcomes.

The author is also thankful to the Measurement and Quantitative Methods (MQM) program, to the Department of Counseling, Educational Psychology and Special Education (CEPSE) for their encouragement and financial support.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS AND ABBREVIATIONS

MLE: Maximum Likelihood Estimation.

NR method: Newton – Raphson iteration method.

OLS method: Ordinary Least Square method.

SAM: Spatial Autoregression Model.

INTRODUCTION

The spatial autoregression model (SAM) represents the relationship between the vector Y representing attributes of subjects, and the association matrix W that represents the relationships among all subjects. In the model, the two main estimands, the autocorrelation coefficient ρ and variance of the random error term σ^2, are uniquely determined by the matrix W and the vector Y. But the commonly used techniques for the estimation of ρ produce many values outside of a sensible range for ρ, and even the interpretation of the parameter ρ itself is not yet well understood. All these facts restrict the use of SAM.

This dissertation explores the interpretation and estimation of ρ as well as σ^2, and emphasizes the importance of the extent of the consistency between W and Y which is captured by ρ. This estimation of ρ is especially useful in the social sciences where the data could be high in dimension (say around 10 or higher), and the data are possibly highly correlated.

Key words: MLE method, Newton-Raphson (NR) iteration, autocorrelation coefficient ρ.

Chapter 1

A GENERAL REVIEW IN THE LITERATURE OF SPATIAL AUTOREGRESSION MODELS (SAM)

Spatial Autoregression Models (SAM) are important in social sciences including in education. In this chapter, I first briefly introduce the parameters in SAM, and approaches researchers have applied to parameter estimation. I cite part of Ord's (1975) technical work on a maximum likelihood estimation (MLE) method, to which my new development in SAM is tightly related.

1.1. The importance of SAM model in education.

Researchers are interested in how students gain their academic achievement (an outcome vector Y). The variable Y could be caused by their personal variables (X_1) such as age, gender and race. Y could also be caused by their family variables (X_2) such as Social Economic Status (SES), income level and his / her mother's education. Again, Y could be caused by their school variables (X_3) such as private / public school, number of teachers in this school, etc. To answer the question how students gain their academic achievement, the multivariate regression is commonly used. When doing so, we assume that there is no mutual connection between individual students, or we are assuming that students' academic activities are mutually independent. For example, in Shavelson (1996), the first assumption for conducting the multivariate regression analysis (MRA) is "*Independence*: The scores for any particular subject are independent of the scores of all other subjects" (page 536). It is the same for conducting the analysis of variance (ANOVA) (page 378).

However, such kind of assumptions might be not appropriate. For example, Cressie (1993) said in his book focusing on the Spatial Models: "The notion that data close together, in time or space, are likely to be correlated (i.e., cannot be modeled as statistically independent) is a natural and social phenomena" (page 3). Also, Duke (1993) said in his article about Network Effects Models: "Researchers in the sociology of education have long recognized the importance of peer influences in shaping the academic achievements, aspirations, and educational attainments", and "… have also been very aware of the limitations of the methodologies they have employed in investigating peer influences" (page 465). Here, "Spatial Models" and "Network Effects Models" are exchangeable in the literature depending on the different topics.

When we start the educational quantitative research work based on the data set of Y, $\{X_1, X_2, X_3\}$ and $\{W_1, W_2, W_3\}$ like this, we usually conduct the multivariate regression treating Y as outcome and $\{X_1, X_2, X_3\}$ as predictors. Now new questions arise: what can we do with those $\{W_1, W_2, W_3\}$? Is there any kind of relation between Y and $\{W_1, W_2, W_3\}$? Or between $\{X_1, X_2, X_3\}$ and $\{W_1, W_2, W_3\}$? We think that there should be some kind of relation, strong or weak, between Y and $\{W_1, W_2, W_3\}$, also, between $\{X_1, X_2, X_3\}$ and $\{W_1, W_2, W_3\}$. Treating W, Y and X as individual variables, we hope to find some kind of relation between W and Y (or X).

The new questions make sense. Subjects are more or less mutually connected and influenced within a network (a matrix W). The connections could be geographical (W_1) such as the distances between the seats of pairs of students (so that W_1 is symmetrical). It could be personal (W_2) such as the level of friendship (thus W_2 is not necessarily symmetrical). The connection could also be social (W_3) such as the sports team membership of the same interest / club (then W_3 probably would be symmetrical).

Here is a practical example for teachers. In Frank (1995, 1996), the data collected from a group of 24 high school teachers included an association (weight) matrix W and some vectors Y. The author identified cohesive subgroups among those teachers based on W, the level of the frequency of their professional discussions. In these two articles, however, the relations between the weight matrix W and each of those vectors Y, such as the teachers' gender, race, year of teaching, and their moral agencies, were not considered due to the lack of available techniques. So that, we are not sure whether or not the weight matrix W, or the pattern of those teachers' professional communication, is more or less associated with some personal vector(s) in that school. Is the pattern of their professional communication mainly associated with individual's orientation to teaching? Or mainly because of their race and gender?

Practically, we are concerned that the subjects' variable (a vector Y) might be more related to one of the W matrices than to others. That is, the student achievement might be more related to the pattern of their friendships than to the geographical distances between their homes. Or, we are concerned that the pattern of subjects' friendship (a matrix W) might be more likely related to one of the X vectors than to others. That is, the pattern of subjects' friendship might be more associated with their gender than with their ages.

Theoretically, the first concern above regards a comparison of relations between one Y with different Ws; the second concern above means to compare the relations between one W with Y and different Xs. Both of those concerns are essentially focusing on the same topic: the relation between a matrix W and a vector Y which can be addressed through Spatial Autoregression Models (SAM).

Substantial developments have been made by Mead (1967, 1971), Ord (1975), Doreian (1981, 1982, 1989), Cressie (1993), Duke (1993), Marsden and Friedkin (1994), Leenders (1995) and others to the literature of Spatial Autoregression models (SAM). However, the remaining difficulties, both theoretical and technical, restrict the practical application of SAM in many instances.

This dissertation makes the SAM model applicable in practice. We will be able to measure the general relation between a matrix W and a vector Y, to compare the differences of relation between one W with different Y and Xs. Also, we may compare the differences of relation between one Y with different Ws. We'll find the essence of the relation between W and Y in SAM.

1.2. The MLE approach to SAM models.

The SAM model was initially applied to studies in geographical and agricultural economics such as Whittle (1954), Mead (1967), and Ord (1975). Working with the spatial autoregression model, researchers explored the relationship between a vector Y representing attributes of subjects, and a weight matrix W, or the association matrix in different contexts, describing the mutual relationship among those subjects.

In SAM, the subjects can range widely including plants, counties or persons; the relationship between subjects is often given in terms of geographical distance. For example, in Ripley (1981), possible subjects mentioned were trees, towns, birds' nests, imperfections of metals, galaxies and earthquakes. When the weight matrix W represents the mutual relationship between individuals in an organization, and Y represents an attribute of those individuals in the organization, we are applying the SAM model in a sociological and psychological sense. Then we may find the important application of SAM models in education.

Once the subjects are determined, the researcher needs to specify the strategy for choosing a measurement for the relationship, either geographical or interpersonal. In spatial autoregression models, the connections among subjects extend in all directions, so that even geographical distance might not be unique depending on the definition of "all" directions. For example, Anselin (1988) suggested geographical ways such as "short path" or "neighboring" for the definition (page 18). Mead (1967) introduced ways of making geometrical connections on a plane in different covering sizes (page 193). Both the above approaches are objective. When trying to apply the SAM model in social sciences, we might be dealing with the interpersonal relations in many different ways, either objective or subjective. For example, when a subject is making a decision, this decision could be influenced by the frequency of phone calls made to others in the social network, then this connection is objective. A subject's decision could also be influenced by his level of intention towards those he'd like to engage, then this connection is subjective. In general, we will deal with many different types of connections, which are represented by the weight matrix W.

SAM initially deals with effects through spatial autocorrelation. In Cliff and Ord (1973), the authors illustrated that concept: "If the presence of some quality in a county of a country makes its presence in neighboring counties more or less likely, we say that the phenomenon exhibits spatial autocorrelation" (page 1), although no direct definition of "autocorrelation" was given. Anselin (1988) acknowledged that spatial autocorrelation, or spatial dependence in his words, "is best known and acknowledged most often, particularly following the pathbreaking work of Cliff and Ord (1973)". The author also said: "it is generally taken to mean the lack of independence which is often present among observations in cross-sectional data sets" (page 8). Recently Leenders (1995) described spatial autocorrelation "either a variable or of an error is the situation where the observations of variables or the values of the error terms for different actors are not independent over time, through space, or across a network". All of the above give us an understanding of spatial autoregression models although a clear definition of autocorrelation is unavailable in the extant literature. We also notice that in recent decades, a lot of excellent research work has been done with the autocorrelation of error terms, but relatively less has been done with the autocorrelation of variables in spatial autoregression models.

1.2.1. The parameters ρ and σ^2 in the SAM model.

In Ord (1975), the weight matrix W (n×n) in SAM is assumed with entries $w_{ij} \geq 0$ ($i \neq j$) and $w_{ii} = 0$ for any i,j=1,2, ..., *n*. Ord's assumption is carried on in this dissertation from chapter 1 through chapter 4, and is developed to a more general case where $w_{ij} \geq 0$ is not required and $w_{ii} = a \neq 0$ is considered in chapter 5.

With $W = \{w_{ij}\}$, an (n×n) set of non-negative weights which represent the degree of association between the *j*th subject and the *i*th subject, and with $Y = \{y_i\}$, an (n×1) set of observed outcomes, the first order spatial autoregression model is

$$y_i = \alpha + \rho \sum_{j=1}^{n} w_{ij} y_j + \varepsilon_i \qquad (1.1)$$

where $\varepsilon \sim N(0, \sigma^2 I)$ with parameters α, ρ and σ^2.

Equation (1.1) can be reformulated in matrix notation by taking $\alpha = 0$ suggested by Ord (1975) (page121), as

$$Y = \rho WY + \varepsilon \qquad (1.2)$$

where W is the (n×n) matrix of weights and Y, ε are (n×1) vectors. We notice that W has entries $w_{ij} \geq 0$ ($i \neq j$) and $w_{ii} = 0$ for any i,j=1,2, ..., *n*.

This is the simplest first order spatial autoregression model in which the parameters ρ and σ^2 need to be estimated. We'll mostly focus on the relationship between W and Y, and the estimate of ρ as a function of W and Y.

For the SAM model, there are different methods of estimating the parameters ρ and σ^2. One of the commonly accepted methods is the maximum likelihood estimation (MLE) method. In MLE for the parameters ρ, the Newton-Raphson iteration is conducted repeatedly until convergence or divergence of parameters ρ is obtained by some rules. In each step for the SAM model, the researcher has to obtain an estimate of parameters ρ which is carried into a matrix for computation, then have the matrix inverted, and obtain a new estimate for the next step. Without a computer or when the dimension of the weight matrix W is high, say around 10 or higher, it is difficult to obtain a maximum likelihood estimate of parameters ρ and σ^2. So until the theoretical development of Ord (1975), this method was rarely considered practical due to computational difficulties.

1.2.2. The "once and for all" technique for the Newton-Raphson (NR) iteration in calculation.

In Ord (1975), writing A = A(ρ) = I - ρW which is in full rank, the log-likelihood function for ρ and σ^2, given Y is

$$l(\rho,\sigma^2) = -\left(\frac{n}{2}\right)\ln(2\pi\sigma^2) - \left(\frac{1}{2\sigma^2}\right)Y'A'AY + \ln|A|. \tag{1.3}$$

We may write the term $Y'A'AY$ as $\|AY\|^2$, and specifically write the term |A| = |detA| where "detA" is the determinant of A. Thus ln |A| is always defined. But in order to be consistent with Ord's article, I keep using ln |A|. A brief note is given later in this chapter.

Then the ML estimators are obtained as

$$\hat{\sigma}^2 = n^{-1}Y'A'AY = \left(\frac{1}{n}\right)\left[\left(Y - \hat{\rho}WY\right)'\left(Y - \hat{\rho}WY\right)\right], \tag{1.4}$$

and the ρ estimate is the maximizer of

$$l(\hat{\rho},\hat{\sigma}^2) = const - \left(\frac{n}{2}\right)\left[\ln(\hat{\sigma}^2) - \frac{2}{n}\ln|A|\right]. \tag{1.5}$$

Equivalently, it is to minimize

$$f(\hat{\rho}) = -\frac{2}{n}\ln|A| + \ln(\hat{\sigma}^2) = G(\hat{\rho}) + H(\hat{\rho}), \qquad (1.6)$$

where $G(\hat{\rho}) = -\frac{2}{n}\ln|A|$ and $H(\hat{\rho}) = \ln(\hat{\sigma}^2)$.

I separate the function f (ρ) into two parts G(ρ) and H(ρ) for the reasons given in chapter 3 (§3.1). This separation is a key step, which helps us to draw conclusions from the equation (1.6) by applying the Cauchy mean value theorem.

The possible minimizer of (1.6) would be the solution to the equation

$$f'(\rho) = 0 \qquad (1.7)$$

or those points on which f '(ρ) does not exist.

In the process of iterating to minimize f (ρ), | A |, the polynomial in ρ has to be evaluated afresh at each iteration. As Ord (1975) said, when *n* is large, or A is irregular, this process becomes computationally intensive. Thus Ord created a new computational procedure as follows.

Since $|A| = \prod_{i=1}^{n}(1-\rho\lambda_i)$ where $\lambda_i \in \{\lambda_i\}$, the eigenvalue set of W, then

$$\ln|A| = \ln\prod_{i=1}^{n}(1-\rho\lambda_i) = \sum_{i=1}^{n}\ln(1-\rho\lambda_i).$$

In the Newton-Raphson iteration process to find the minimizer of f (ρ) in (1.6), writing Y_L = WY, the iteration is taken as

$\rho_{r+1} = \rho_r - f'(\rho_r)/f''(\rho_r)$ where

$f'(\rho) = \left(\frac{2}{n}\right)\sum_{i=1}^{n}\lambda_i/(1-\rho\lambda_i) + 2\left(\rho(Y_L)'Y_L - Y'Y_L\right)/s^2$ and

$$f''(\rho)=\left(\frac{2}{n}\right)\sum_{i=1}^{n}(\lambda_i)^2/(1-\rho\lambda_i)^2+2(Y_L)'Y_L/s^2-4\left(\rho(Y_L)'Y_L-Y'Y_L\right)^2/s^4 \text{ where}$$

$$s^2\equiv s^2(\rho)=Y'Y-2\rho Y'Y_L+\rho^2(Y_L)'Y_L.$$

If ρ_r converges when $r\to\infty$, it converges to the ML estimate of the parameter ρ.

The advantage of Ord's technique of writing a determinant |A| into an algebraic product is obvious. Because we need to repeatedly calculate the value of |A| which changes in each step, we need to deal with a lot of matrix computation. But now with Ord's technique, whole matrix computations become simple algebra because the eigenvalues of W need to be evaluated only once. That is why this technique is called the "once and for all" technique.

1.2.3. A note in case the weight matrix W is asymmetrical.

We know that

$$\hat{\sigma}^2=n^{-1}Y'A'AY=\frac{1}{n}\left[Y'(I-\hat{\rho}W)'(I-\hat{\rho}W)Y\right]$$

$$=\frac{1}{n}\left[Y'Y-\hat{\rho}Y'WY-\hat{\rho}(WY)'Y+\hat{\rho}^2(WY)'(WY)\right].$$

When W′ = W, this form will be the same as that in Ord's Appendix A for the following computation. The function (1.6) is exactly equivalent to

$$f(\hat{\rho})=\left(-\frac{2}{n}\right)\sum_{i=1}^{n}\ln(1-\hat{\rho}\lambda_i)+\ln\left(Y'Y-2\hat{\rho}Y'Y_L+\hat{\rho}^2(Y_L)'Y_L\right)-\ln n. \quad (1.8)$$

where the author uses Y_L to represent WY.

But when W′ ≠ W, a possible case in social sciences, the above formula (1.8) and its derivatives need to be slightly adjusted. The new version of f (ρ) after adjustment should be

$$f(\hat{\rho}) = \left(-\frac{2}{n}\right)\sum_{i=1}^{n}\ln\left(1-\hat{\rho}\lambda_i\right)+\ln\left(Y'Y-\hat{\rho}Y'(W+W')Y+\hat{\rho}^2(WY)'WY\right)-\ln n. \qquad (1.8)'$$

The expression of the minimizer of f (ρ) in the Newton-Raphson iteration process is the same as above, namely

$$\rho_{r+1} = \rho_r - f'(\rho_r)/f''(\rho_r).$$

But others are now as follows:

$$f'(\rho)=\left(\frac{2}{n}\right)\sum_{i=1}^{n}\lambda_i/\left(1-\rho\lambda_i\right)+\left(2\rho(WY)'WY-Y'(W+W')Y\right)/s^2 \text{ and}$$

$$f''(\rho)=\left(\frac{2}{n}\right)\sum_{i=1}^{n}(\lambda_i)^2/\left(1-\rho\lambda_i\right)^2+2(WY)'WY/s^2-\left(2\rho(WY)'WY-Y'(W+W')Y\right)^2/s^4$$

where

$$s^2 \equiv s^2(\rho)=Y'Y-\rho Y'(W+W')Y+\rho^2(WY)'WY.$$

Here I use WY instead of Y_L to clarify the difference between W and W′.

In the following chapters, I am not going to emphasize the difference between symmetrical and asymmetrical Ws repeatedly. But in chapter 6, the data analysis involves an (8×8) asymmetrical weight matrix, and the formulas for computations are the new versions.

1.2.4. Problems when |A| is non-positive, and how to avoid the problems.

We know that $| A | = | I - \rho W |$ is an *n*th polynomial of ρ, and may have values either positive, or negative, or possibly zero when ρ ranges on $(-\infty, \infty)$. In some formulas of this chapter, e.g. (1.3), (1.5), the logarithm of $|A|$ is taken, this logarithm may make no sense if the value of $|A|$ is zero or negative. Also, in chapter 4, $|A|$ appears in a log-likelihood function which may also make no sense if the value $|A|$ is zero or negative. That is why I mentioned early in this chapter to write $|A|$ as $|\det A|$ to avoid such a technical problem. In the following, I treat $|A|$ as $|\det A|$ without further notification.

1.3. Methods other than MLE.

In Ord (1975), the MLE method is compared with the ordinary least square (OLS) method or generalized least square (GLS) method. Ord showed (page121-122) that the OLS estimator is inconsistent for a general weight matrix W, and is consistent only when W is triangular, a limited and non-interesting case.

1.3.1. Anselin's Bayesian method.

In Anselin (1988), the author introduced some other methods for comparison with MLE method. The author concluded (page 81): "the maximum likelihood approach to estimation and hypothesis testing in spatial process models is by far the better known methodological framework".

For the Bayesian method, the combination of prior information about the distribution of the parameters, namely the autoregressive coefficient ρ and the error variance σ^2 in the model, are needed. When we start to work on an SAM model, it might be difficult to get such information in advance. Without any prior information, however, this method is not appropriate (not a Bayesian one). As Anselin said (page 89-90), "following the standard approach in econometrics, diffuse prior densities for these parameters are expressed as:

$P(\sigma) \propto 1/\sigma, 0 < \sigma < +\infty$.

$P(\rho) \propto$ constant, $-1 < \rho < +1$."

There might be flaws in these priors. The second prior density means that ρ is uniformly distributed on the interval (-1, 1) which is the boundary for the model after the weight matrix W has been standardized as some authors suggested. However, the relationship between W and Y is not considered which can strongly affect the estimate ρ as I will show in the following chapters. Also, as we know, the absolute values of the maximum and minimum eigenvalues of the weight matrix W are not equal in most cases, so that $(1/\lambda_{min}, 1/\lambda_{max})$, the range of ρ, is not symmetrical to the origin in most cases. This second prior density seems not to be appropriate. Anyway, Anselin said (page 81) that in spatial models, the implementation of alternative approaches (other than MLE method) including the Bayesian method "has been rather limited".

1.3.2. Cressie's asymptotic property approach.

Cressie (1993) gives details of spatial modeling and parameter estimating regarding the lattice models in the chapters 6 and 7 of his book. The author said (page 458), for models in chapter 6 of his book where the first order spatial autoregression model is included, "estimation of model parameters is not always so straightforward. Of course, finite-sample properties, such as sufficiency, completeness, ancillarity, unbiasedness, minimum mean-square error … are still desirable; however, they are even more elusive than for the i.i.d. paradigm." As a conclusion, the author said that "methods of estimation are usually assessed via their asymptotic properties" while those finite-sample properties are not guaranteed.

Chapter 2

EXISTING PROBLEMS AND NEW PROBLEMS WHEN APPLYING SPATIAL AUTOREGRESSION MODELS TO THE SOCIAL SCIENCES

This chapter is mainly a list of problems: theoretical or technical; old or new in the literature of Spatial Autoregression Models (SAM).

When dealing with the agricultural and geographical data decades ago, the SAM model arose mainly as a theoretical work with limited practical meaning due to the computational difficulty and the lack of understanding of the model itself. When trying to apply spatial autoregression models in social network analysis in recent years, some new problems occurred in the model application. I list these old and new problems in the following, and am going to solve them in the following chapters.

2.1. The theoretical limitation, the old problems.

(a). The definition of the parameter ρ has not been clear.

- As cited in chapter 1, Cliff and Ord (1973) said that "If the presence of some quality in a county of a country makes its presence in neighboring counties more or less likely, we say that the phenomenon exhibits spatial autocorrelation". But no definition of autorcorrelation or illustration of the autocorrelation coefficient ρ were given.

- Mead (1967) wrote (page 191) that "λ is defined to be the competition coefficient" in his interplant competition model where the coefficient ρ was written as λ. In another paper (1971), the same author said (page 18) that "λ is a competition coefficient which, if the 'correct' form of f_j is found, would be expected to remain constant at a particular time in a particular environment". There, f_j is the element of the weight matrix (or the association matrix). We noticed that in Mead (1967), the author cited a discussion about the parameter ρ from another early paper, Kira et al. (1953). When dealing with a simplified setting, Kira et al. "found that most values of the (autocorrelation) coefficient were positive, which they interpreted as showing co-operative rather than competitive situations". This interpretation began to be considered as the meaning of the parameter ρ in a limited sense.
- Ord (1975) briefly says (page 120) that together with σ as the estimate of the standard deviation for the random error, "σ and ρ are parameters".
- In Doreian (1982), the author says (page 240) that "ρ is the spatial effects parameter". In his other paper (1989), the author says (page 285) that "ρ is the network autoregressive parameter".
- In Anselin (1988), on page 33, the first order spatial autoregression model is introduced with no statement dealing with the parameter ρ; on page 35, the author says that "ρ is the coefficient of the spatially lagged dependent variable (of WY)", giving no more details; on page 58, the author says that "ρ is a spatial autoregression coefficient".

- When tracing back the history of ρ, Leenders (1995) says (page 55) that "ρ is a scalar". The author also mentions (page 167) that ρ "should be considered a *descriptive* parameter rather than a governing parameter" with no further explaining on being "*descriptive*".
- Duke (1993) gives a statement about the parameter ρ by saying "ρ is the parameter representing the strength of the context effect". Also, no further statement was given in that paper.

From the above, we may find that these statements about parameter ρ are mainly in a sense of naming, rather than giving a definition. A new definition will be given in chapter 5 (§5.1).

(b). The necessity of the boundary for the ρ estimation is not clear.

The boundary of the estimate ρ is a key issue concerning the autoregression model in the literature. The estimated value of parameter ρ from a computer direct search frequently is large, and researchers have tried to find an appropriate boundary to restrict the divergence. The commonly accepted range ($1/\lambda_{min}$, $1/\lambda_{max}$) or similar ones can be found in Ord (1975), Doreian and Hunmon (1976), Anselin (1982). Recently, Leenders (1995) suggested that parameter ρ could truly take large values with no further reasoning. I will address this issue in chapter 4 (§4.3).

(c). The range and interpretation of the estimate of ρ have not been discussed.

The estimate ρ is the minimizer of the likelihood function f(ρ), so is a solution of the equation $f'(\rho) = 0$. We know that $f'(\rho) = 0$ is a non-algebraic equation with multiple roots, and an analytic expression of the estimate of ρ is difficult to obtain. We also know that the interpretation of the estimate of ρ was least talked about so far in the literature. The meaning of ρ being high or low, positive or negative, in which sense, and in what kind of scale, etc. has not been clearly stated. The range and interpretation of the estimate of ρ will be discussed in chapter 4 (§4.1).

2.2. Technical constraints, the new problems when applying the model to the social network analysis.

Because the matrix W could be asymmetrical, or the value w_{ij} of W might be negative, we are facing some new problems in the field.

(d). The matrix W might be asymmetrical, so that there might be (and most likely will be) complex eigenvalues. On the other hand, when dealing with the boundary problem, some researchers such as Ord (1975), Doreian (1981), Duke (1993) and Leenders (1995) suggested the "row normalization" for matrix W. Once W is row normalized by setting

$w_{ij}{*}=\frac{w_{ij}}{\sum_{j=1}^{n} w_{ij}}$ for any $i, j = 1, 2, ..., n$ so that $\sum_{j=1}^{n} w_{ij}^{*} = 1$ for any $i = 1, 2, ..., n$, the boundary

of the estimate ρ would be simply $|\rho| \leq 1$. However, the normalized matrix W* would become asymmetrical in most cases and cause the complex eigenvalue problem, although the original matrix W could be symmetrical. This issue is addressed in chapter 3 (§3.1.2).

(e). So far in the literature, the weight matrix is treated as $w_{ii} = 0$ for i=1,2, ..., n. In social sciences, however, the case $w_{ii} = a \neq 0$ could make sense but has not yet been considered. We need to explore the essence of the case $w_{ii} = a \neq 0$, and the relationship between the estimates ρ for cases $w_{ii} = 0$ and $w_{ii} = a \neq 0$. If we treat this "a" as a constant which may take zero, positive or even negative values, we are actually talking about W(a), a family of weight matrices in a more general sense. I will address this issue in chapter 5.

Chapter 3

A NEW TECHNIQUE FOR ESTIMATION OF PARAMETERS ρ AND σ^2

TECHNICAL IMPROVEMENTS

In this chapter, I explore the construction of SAM model. Based on the understanding of the construction of SAM, I obtain two results regarding the parameter estimation of SAM. One is that in the interval $(1/\lambda_{min}, 1/\lambda_{max})$, the value of Y′WY and estimate ρ will take the same sign, which helps to understand the construction of solution space of the estimate ρ (§3.1). Another is the so-called "far end" method of initial value selection for the estimation of ρ, which helps to control the estimate ρ not "flying out" of the required boundary (§3.2).

3.1. The "Separation" of the log-likelihood function.

The key work here is to use a "separation" technique which helps to draw conclusions from the log-likelihood equation in the MLE method. That is, I separate the function $f'(\rho)$ into two parts: $g(\rho)$ and $h(\rho)$. In function (1.6), we had the function $f(\rho) = G(\rho) + H(\rho)$. Taking the derivative of function (1.6) gives

$f'(\rho) = G'(\rho) + H'(\rho) = g(\rho) + h(\rho)$ where

$$g(\rho) = \left(\frac{2}{n}\right)\sum_{i=1}^{n}\frac{\lambda_i}{(1-\rho\lambda_i)} \quad \text{and}$$

$$h(\rho) = 2 \cdot \frac{\rho (Y_L)' Y_L - Y'Y_L}{Y'Y - 2\rho Y'Y_L + \rho^2 (Y_L)' Y_L}.$$

with $Y_L = WY$.

Clearly, the value of g(ρ) is determined by ρ and W's eigenvalues $\{\lambda_i\}$, and the value of h(ρ) is determined by ρ, and both W and Y.

Now we have the derivative of the log-likelihood function as we had in chapter 1,

$$f'(\rho) = \left(\frac{2}{n}\right) \sum_{i=1}^{n} \frac{\lambda_i}{(1 - \rho\lambda_i)} + 2 \cdot \frac{\rho (Y_L)' Y_L - Y'Y_L}{Y'Y - 2\rho Y'Y_L + \rho^2 (Y_L)' Y_L}.$$

I discuss the solution of equation $f'(\rho) = 0$ which will be the solution to our parameter estimation.

3.1.1. Assuming W has all real eigenvalues.

Case 1. Assuming $Y'Y_L > 0$.

Applying the Cauchy mean value theorem shows that both $Y'Y_L$ and estimate ρ will have positive values in the boundary $(1/\lambda_{min}, 1/\lambda_{max})$.

[Table3A about here]

I briefly explain Table3A as follows. The second column from left is checking the value of function $f'(\rho)$ when $\rho = 0$. Since $g(0) = (2/n)\sum /\lambda_i = 0$ and $h(0) = -2(Y'Y_L)/ Y_L'Y_L < 0$, we have $f'(0) = g(0) + h(0) < 0$. The fourth and fifth columns from left are checking the value of function $f'(\rho)$ when $\rho = \rho^{*} = \min(\rho^{\nabla}, 1/\lambda_{max})$ where $\rho^{\nabla} = (Y'Y_L)/ Y_L'Y_L$, the OLS estimate. When $\rho^{\nabla} \le 1/\lambda_{max}$, we have

Table3A

The location of ρ, the solution of the equation $f'(\rho) = 0$.
Assuming $Y'Y_L > 0$ so that $\rho^{\nabla} = Y'Y_L / (Y_L)'Y_L > 0$.

	$\rho = 0$	There must be ρ^* such that $0 < \rho^* < \min(\rho^{\nabla}, 1/\lambda_{max})$	$\rho^* = \min(\rho^{\nabla}, 1/\lambda_{max})$ with $\rho^{\nabla} \leq 1/\lambda_{max}$	$\rho^* = \min(\rho^{\nabla}, 1/\lambda_{max})$ with $\rho^{\nabla} > 1/\lambda_{max}$
$g(\rho) = (2/n)\sum \lambda_i / (1-\rho\lambda_i)$	$g(0) = (2/n)\sum \lambda_i = 0$		$g(\rho^*) = g(\rho^{\nabla}) > 0$	$g(\rho^*) = g(1/\lambda_{max}) = +\infty$
$h(\rho) = 2(\rho(Y_L)'Y_L - Y'Y_L) / (Y'Y - 2\rho Y'Y_L + \rho^2(Y_L)'Y_L)$	$h(0) = -2Y'Y_L/Y'Y < 0$		$h(\rho^*) = h(\rho^{\nabla}) = 0$	$h(\rho^*) = h(1/\lambda_{max}) < 0$
$f'(\rho) = g(\rho) + h(\rho)$	$f'(0) = g(0) + h(0) < 0$	$f'(\rho^*) = 0$	$f'(\rho^*) = f'(\rho^{\nabla}) > 0$	$f'(\rho^*) = f'(1/\lambda_{max}) = +\infty$

$g(\rho^{\bullet}) = g(\rho^{\nabla}) > 0$, and $h(\rho^{\bullet}) = h(\rho^{\nabla}) = 0$ so that $f'(\rho^{\bullet}) = g(\rho^{\nabla}) + h(\rho^{\nabla}) > 0$. When $\rho^{\nabla} > 1/\lambda_{max}$, we have $g(\rho^{\bullet}) = g(1/\lambda_{max}) = +\infty$, and $h(\rho^{\bullet}) = h(1/\lambda_{max}) < 0$ so that $f'(\rho^{\bullet}) = g(1/\lambda_{max}) + h(1/\lambda_{max}) = +\infty$. At last, we have $f'(\rho^{\bullet}) > 0$ no matter $\rho^{\nabla} \leq 1/\lambda_{max}$ or $\rho^{\nabla} > 1/\lambda_{max}$.

Since $f'(\rho)$ is continuous on $(0, \rho^{\bullet})$ with $f'(0) < 0$ and $f'(\rho^{\bullet}-0) > 0$, we must have at least one point, namely ρ^{*}, within $(0, \rho^{\bullet})$ with $f'(\rho^{*}) = 0$ by Cauchy mean value theorem. That is, both $Y'Y_L$ and estimate ρ will have positive values in the interval $(1/\lambda_{min}, 1/\lambda_{max})$. Clearly to say, if $Y'Y_L > 0$ given W and Y, then the estimate $\rho > 0$. This conclusion is expressed in the third column from left in Table3A.

Case 2. Assuming $Y'Y_L < 0$.

In the similar way, we see that $Y'Y_L$ and the estimate of ρ will both be negative in the boundary $(1/\lambda_{min}, 1/\lambda_{max})$.

Case 3. Assuming $Y'Y_L = 0$.

In case $Y'Y_L = 0$, $Y'Y_L$ and the estimate of ρ will be both zeroes.

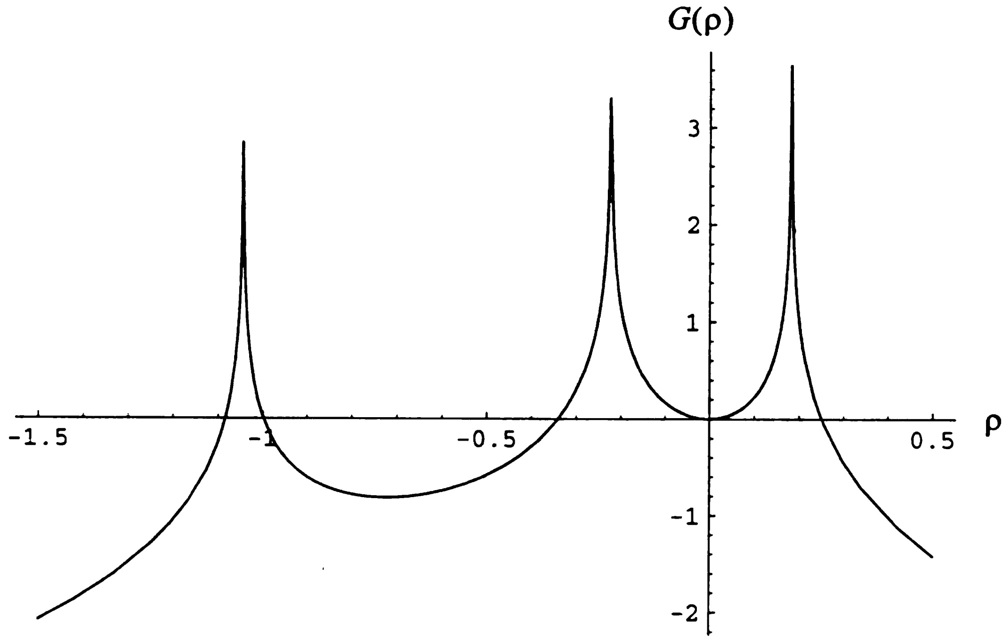
Now, we see that the sign of the estimate of ρ is uniquely determined by the sign of $Y'Y_L$ which is determined by the relationship between W and Y since $Y'Y_L = Y'WY$. Thus, we make the conclusion that the value of $Y'Y_L$ and estimate of ρ, the minimizer of the equation $f'(\rho) = 0$ will have the same sign in the boundary $(1/\lambda_{min}, 1/\lambda_{max})$. This gives three parts of the solution space of the estimate ρ taking zero, positive and negative values respectively.

An example is given in appendix (Appendix3A) where the dimensional number is 3. I choose $Y_1 = (1\ 1\ 1)'$ and $Y_2 = (1\ 0\ -1)'$ so that we will get $Y_1'Y_{1L} = Y_1'W\ Y_1 = 16 > 0$ and $Y_2'Y_{2L} = Y_2'W\ Y_2 = -6 < 0$ respectively. In the first case, the estimate of ρ will be positive, and in the second, the estimate of ρ will be negative. The separated parts of its log-likelihood function $G(\hat{\rho}) = -\frac{2}{n}\ln|A|$ and $H(\hat{\rho}) = \ln(\hat{\sigma}^2)$, and their corresponding derivatives are g(ρ) and h(ρ). In Figure3A-Figure3J, I show the likelihood functions and their corresponding derivatives for two different Y vectors in order to demonstrate how the positive / negative estimates of ρ will be located.

[Figure3A – Figure3J about here]

I briefly explain Figure3A – Figure3J as follows. Figure3A is the first part of the log-likelihood function, namely G(ρ) = -(2/n)ln(I - ρW). In the figure, the y axis represents G(ρ). Figure3B is the derivative of the first part of the log-likelihood function G(ρ), namely g(ρ) = G′(ρ). Both functions G(ρ) and g(ρ) are independent of Y. The second part of the log-likelihood function is related with Y. The shape of the figure will vary when the given Y varies. Figure3C is the second part of the likelihood function H(ρ) = $\ln(\sigma^2)$ when Y = (1 1 1)′, and Figure3D is the derivative of the second part of the likelihood function, namely h(ρ) = H′(ρ) when Y = (1 1 1)′.

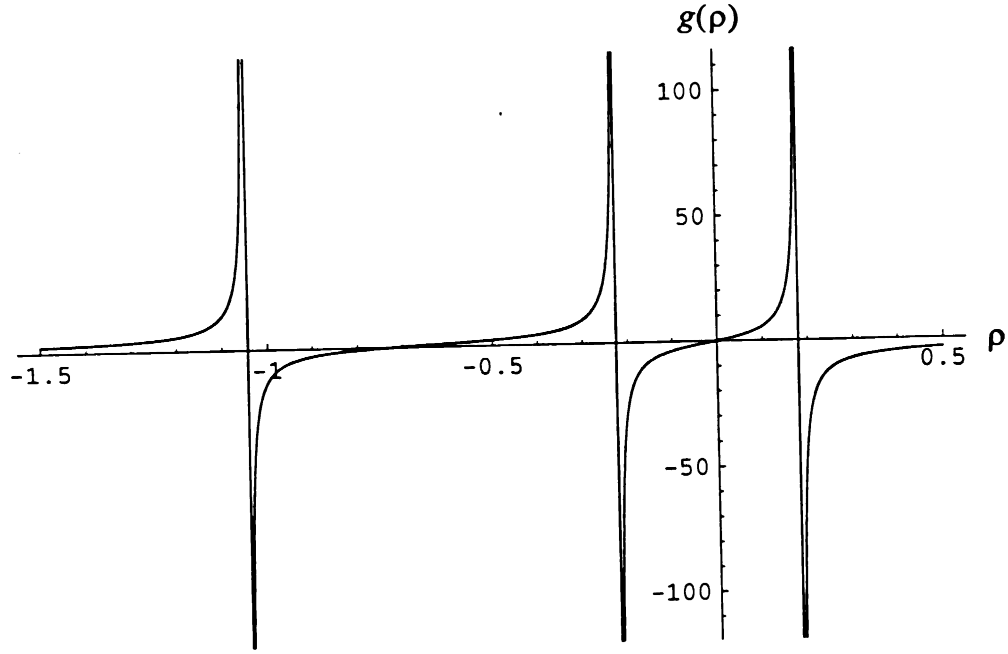
Now summing up part one and part two of the log-likelihood function, we get Figure3G, the figure of the likelihood function, namely f (ρ) when Y = (1 1 1)′, and Figure3H, the figure of the derivative of the likelihood function, namely f ′(ρ) when



$$G(\rho) = -\frac{2}{n}\ln(I - \rho W).$$

The first part of the log-likelihood function, n = 3.

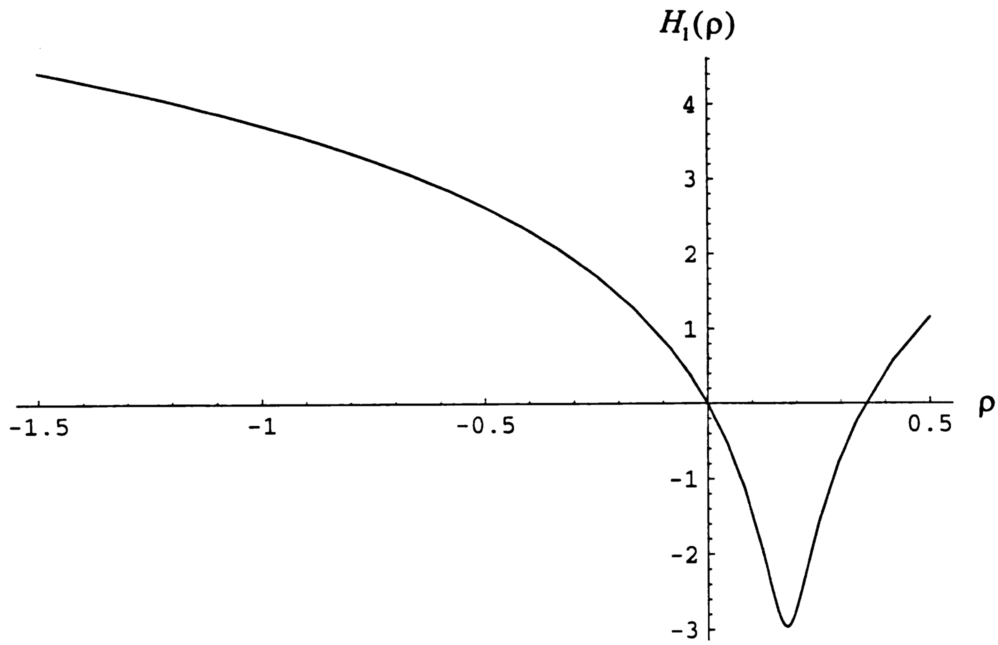
Figure3A



$$g(\rho) = G'(\rho) = \left(-\frac{2}{n}\ln(I - \rho W)\right)'.$$

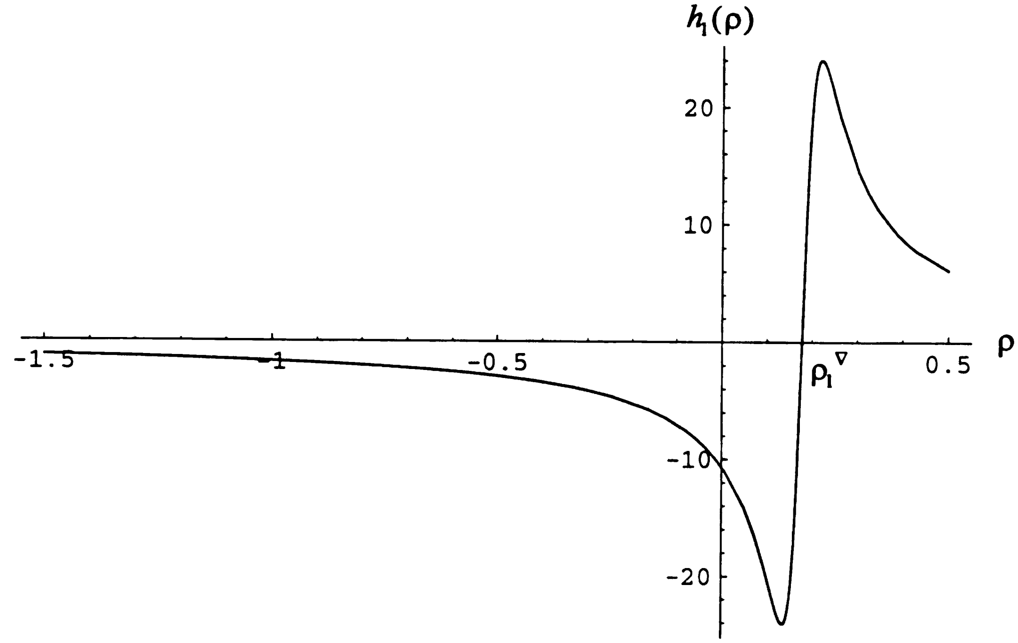
The derivative of the first part of the log-likelihood function, n = 3.

Figure3B



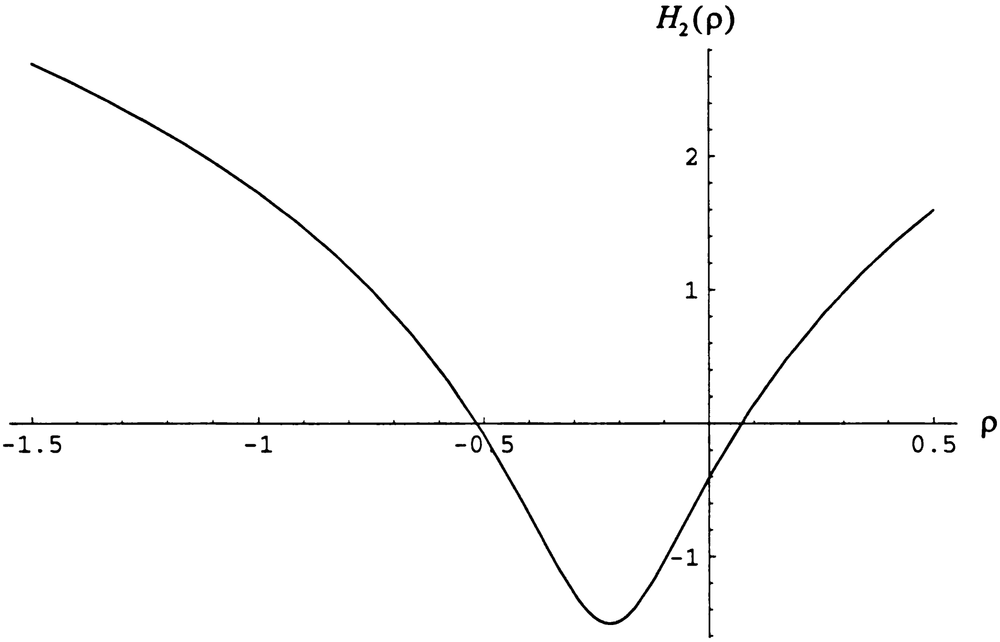
$H_1(\rho) = \ln(\sigma^2|_{y_1})$.
The second part of the log-likelihood function where $y_1 = (1\ 1\ 1)'$, n = 3.

Figure3C



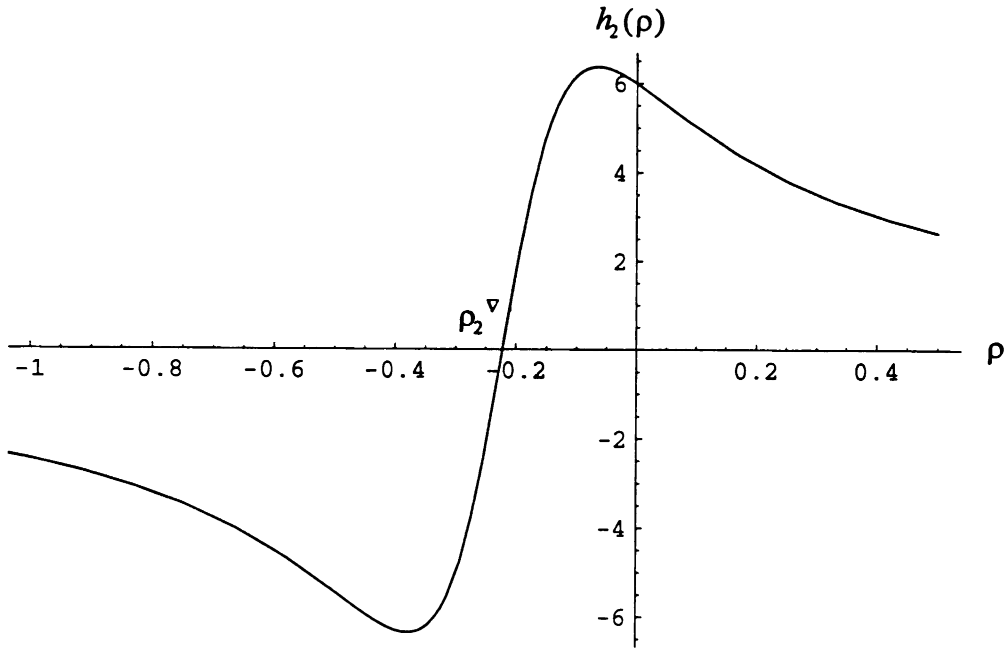
$h_1(\rho) = H_1'(\rho) = (\ln(\sigma^2|_{y_1}))'$.
The derivative of the second part of the log-likelihood function
Where $y_1 = (1\ 1\ 1)'$, n = 3.

Figure3D



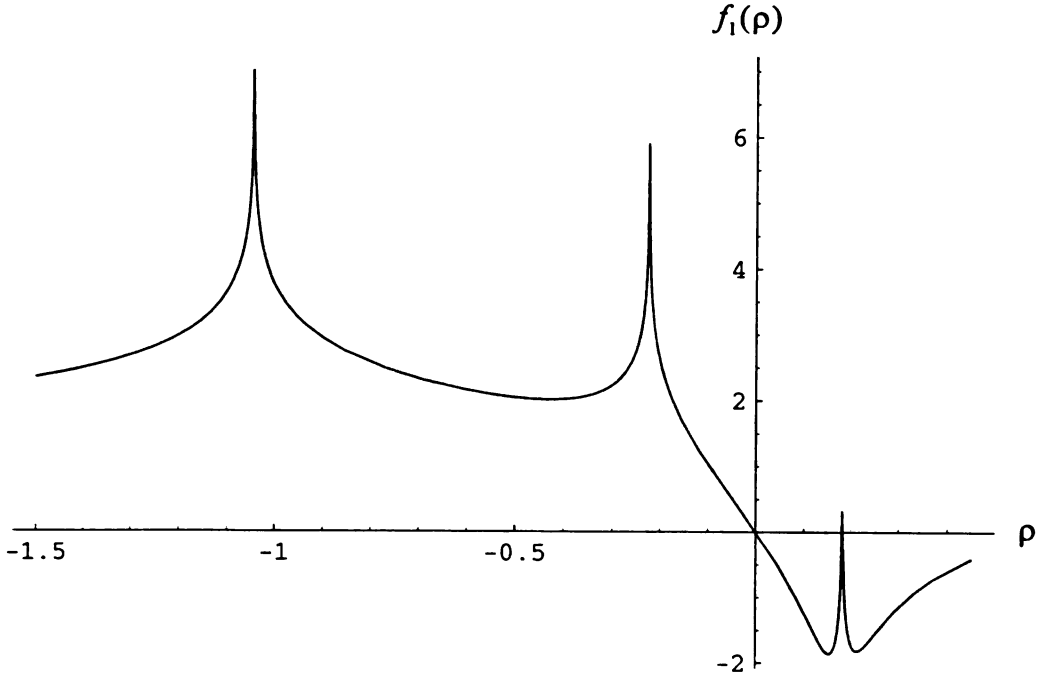
$H_2(\rho) = \ln(\sigma^2|_{y_2})$.
The second part of the log-likelihood function where $y_2 = (1\ 0\ \text{-}1)'$, $n = 3$.

Figure3E



$h_2(\rho) = H_2'(\rho) = (\ln(\sigma^2|_{y_2}))'$.
The derivative of the second part of the log-likelihood function
Where $y_2 = (1\ 0\ -1)'$, $n = 3$.

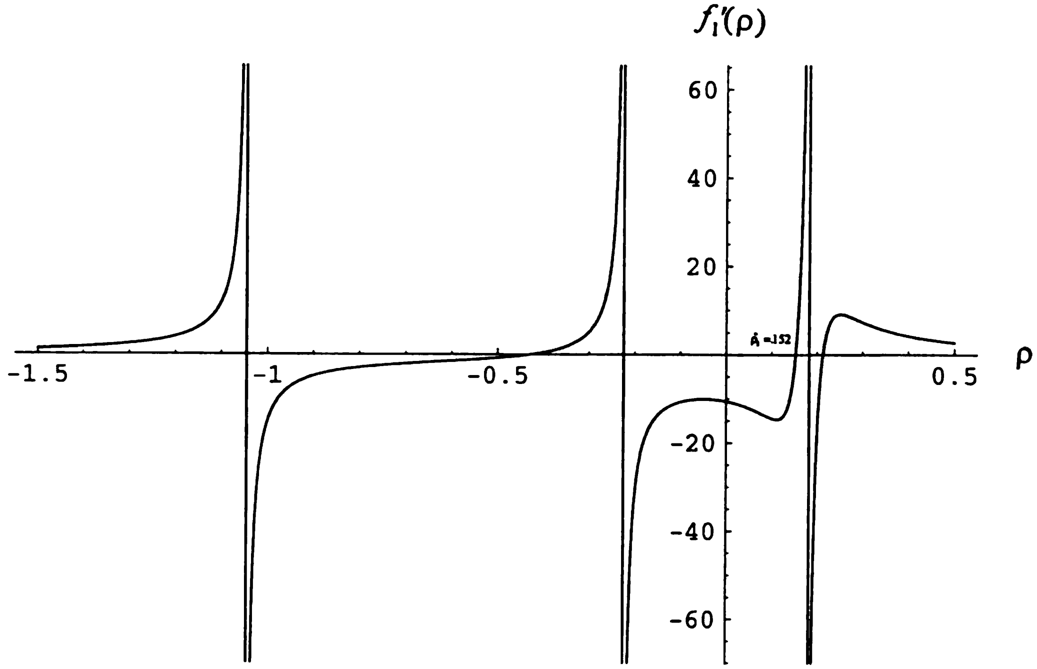
Figure3F



$$f_1(\rho) = G(\rho) + H_1(\rho) = -\frac{2}{n}\ln(I - \rho W) + \ln(\sigma^2\big|_{y_1}).$$

The log-likelihood function where $y_1 = (1\ 1\ 1)'$, $n = 3$.

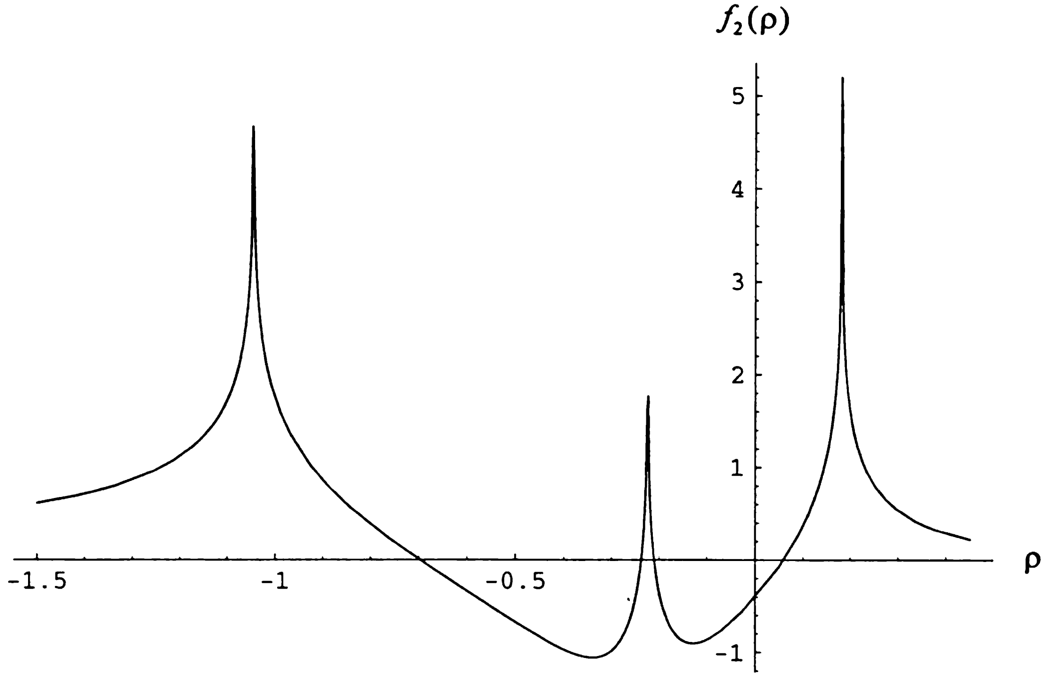
Figure3G



$$f_1'(\rho) = g(\rho) + h_1(\rho) = \left(-\frac{2}{n}\ln(I - \rho W)\right)' + (\ln(\sigma^2|_{y_1}))'.$$

The derivative of the log-likelihood function where $y_1 = (1\ 1\ 1)'$, $n = 3$.

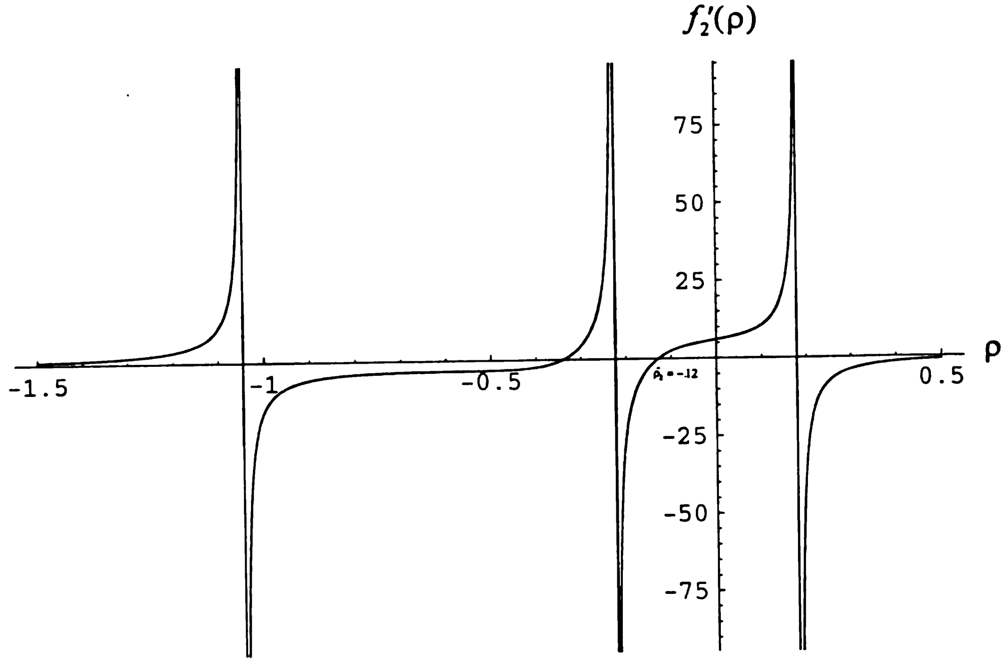
Figure3H



$$f_2(\rho) = G(\rho) + H_2(\rho) = -\frac{2}{n}\ln(I - \rho W) + \ln(\sigma^2\big|_{y_2}).$$

The log-likelihood function where $y_2 = (1\ 0\ \text{-}1)'$, n = 3.

Figure3I



$$f_2'(\rho) = g(\rho) + h_2(\rho) = \left(-\frac{2}{n}\ln(I - \rho W)\right)' + (\ln(\sigma^2|_{y_2}))'.$$

The derivative of the log-likelihood function where $y_2 = (1\ 0\ \text{-}1)'$, n = 3.

Figure3J

$Y = (1\ 1\ 1)'$.

Similarly, Figure3E is the second part of the likelihood function $H(\rho) = \ln(\sigma^2)$ when $Y = (1\ 0\ {-1})'$, and Figure3F is the derivative of the second part of the likelihood function, namely $h(\rho) = H'(\rho)$ when $Y = (1\ 0\ {-1})'$. We also have Figure3I, the figure of the likelihood function $f(\rho)$ when $Y = (1\ 0\ {-1})'$, and Figure3J, the figure of the derivative of the likelihood function $f'(\rho)$ when $Y = (1\ 0\ {-1})'$.

When $Y_1 = (1\ 1\ 1)'$, we see in Figure3H that the curve of $f'(\rho)$ meets the ρ axis three times. The one which is within the boundary $(1/\lambda_{min}, 1/\lambda_{max})$ is what we want to find, the estimate of ρ. It is positive as we have predicted in Table3A. Similarly, when $Y_2 = (1\ 0\ {-1})'$, we see in Figure3J that the estimate ρ is negative.

3.1.2. Assuming that W is asymmetrical and has pairs of conjugate complex eigenvalues.

Duke (1993) has shown that $\sum_{i=1}^{n} \frac{\lambda_i^2}{(1-\rho\lambda_i)^2}$ represents a real number for the variance matrix of parameters including ρ. Duke did not mention the derivatives of $g(\rho)$ and $h(\rho)$ which are used for computation, nor the functions $f'(\rho)$ and $f''(\rho)$ which are used in the whole process of iterations. The conjugate complex eigenvalue problem was also touched by Leenders (1995) but not completed.

In the case of conjugate complex eigenvalues, algebraic calculation shows that $f'(\rho)$ and $f''(\rho)$ remain real in each step of each iteration (see Appendix3B), then with an initial value ρ° real, the next ρ estimate generated from the iteration will be real, and so forth in all the following steps. If this process of iteration converges, it converges to a real value, the estimate of ρ.

3.1.3. The solution space of the estimate ρ.

As we know, any matrix W (even if complex) is similar to a Jordan canonical standard form that is not necessarily diagonal. That is, by an orthogonal transformation, any matrix W is similar to a matrix Λ with diagonal blocks of Jordan matrices. Here, a Jordan matrix has its principal diagonal constant while the next diagonal contains ones, and all the remaining elements zeros. When the order of these blocks of Jordan matrices and the other eigenvalues is fixed, the orthogonal transformation matrix P is unique if W is non-singular.

The above consideration about the Jordan canonical standard form is mostly mathematical, but we need to talk a little about it here. As we know, when W is similar to a Jordan canonical standard form, the corresponding Λ might not be diagonal. But we may have found that Λ being diagonal or not is not a restriction to our previous work, and not to the following work as well.

At last, we may find that there is almost no restriction to the weight matrix. It can be symmetrical or not; sparse or not; singular or not. The only restriction to W might be that W is real in practice. But theoretically and mathematically, even this restriction is not necessary.

When W is real and symmetrical, we know that W is similar to a real diagonal matrix, $W = P'\Lambda P$ where Λ is diagonal and P is orthogonal with $P'P = I$. When W is non-singular and the order of its eigenvalues is fixed, the corresponding P is unique. Now assuming W is real and symmetrical, we have the corresponding

$$\Lambda = \begin{pmatrix} L_1 & 0 & 0 \\ 0 & L_2 & 0 \\ 0 & 0 & L_3 \end{pmatrix}$$ where L_1 is diagonal with all elements positive (all of W′s positive eigenvalues) and with its index set I_1; L_3 is also diagonal with all elements negative (all of W′s negative eigenvalues) and with its index set I_3; all elements of L_2 are zeroes (all of W′s nil eigenvalues) and with its index set I_2. Seemingly, the dimension number of $L_1 \cup L_2 \cup L_3$ is *n*. In the following discussion, it's convenient to put all eigenvalues of L_1 and L_3 in a descending order.

In order to decide the sign of the estimate of ρ, we need to solve for

$$Y'Y_L > 0 \tag{3.1.1}$$

$$Y'Y_L = 0 \tag{3.1.2}$$

$$Y'Y_L < 0 \tag{3.1.3}$$

respectively.

It's easy to see that $Y'Y_L = Y'WY = Y'P\Lambda P'Y = (P'Y)'\Lambda(P'Y)$. Let $P'Y = Z$, so that $Y = PZ$ because $P'P=I$. The transformation between Y and Z is orthogonal and $Y'Y_L$ is transformed into $Z'\Lambda Z$. Corresponding to $\Lambda = \begin{pmatrix} L_1 & 0 & 0 \\ 0 & L_2 & 0 \\ 0 & 0 & L_3 \end{pmatrix}$, I write $Z = (Z_1\ Z_2\ Z_3)'$ where Z_i is the subset of vector Z and get

$$Y'Y_L = Z'\Lambda Z = \begin{pmatrix} Z_1 & Z_2 & Z_3 \end{pmatrix} \begin{pmatrix} L_1 & 0 & 0 \\ 0 & L_2 & 0 \\ 0 & 0 & L_3 \end{pmatrix} \begin{pmatrix} Z_1 \\ Z_2 \\ Z_3 \end{pmatrix} = \sum_{i=1}^{3} Z_i{}' L_i Z_i .$$

Recalling that L_2 is a nil block, L_1 and L_3 are both diagonal with all positive and all negative eigenvalues respectively, I transform the equations (3.1.1), (3.1.2) and (3.1.3) into:

$$S_Z(\lambda, Z) = \sum_{i \in I_1} \lambda_i \cdot z_i^2 + \sum_{i \in I_3} \lambda_i \cdot z_i^2 > 0 \tag{3.2.1}$$

$$S_Z(\lambda, Z) = \sum_{i \in I_1} \lambda_i \cdot z_i^2 + \sum_{i \in I_3} \lambda_i \cdot z_i^2 = 0 \tag{3.2.2}$$

$$S_Z(\lambda, Z) = \sum_{i \in I_1} \lambda_i \cdot z_i^2 + \sum_{i \in I_3} \lambda_i \cdot z_i^2 < 0 \tag{3.2.3}$$

where z_i ($i \in I_1$, I_3) are the elements of Z′s subset Z_1, Z_3 respectively.

Clearly, now we get the solution space of the estimate ρ when ρ equals zero from (3.2.2) which is actually an *n*-dimensional conicoid. The "inside" and "outside" of the conicoid would be the solution spaces from inequalities (3.2.1) and (3.2.3) respectively. Thus, we get the construction of the solution space of the estimate ρ from a weight matrix W given any observation vector Y. By the orthogonal transformation, we transform W (and Y) from W and Y space into Λ (and Z) in Λ and Z space where each axis Z_i (weighted by the corresponding eigenvalue λ_i) represents the *i*th factor. We notice that Λ is diagonal.

To make the space visible, we draw a conicoid in a three dimensional space (see Figure3K). This figure is based on the data from the example 1 (see Appendix3A) where W is a 3×3 matrix on X or Z_1 ($\lambda_1 = 5.50963$), Y or Z_2 ($\lambda_2 = -4.55287$) and Z or Z_3 ($\lambda_3 = -.956762$) coordinate system within a range of ($0 \le X \le 2, -2 \le Y \le 2$).

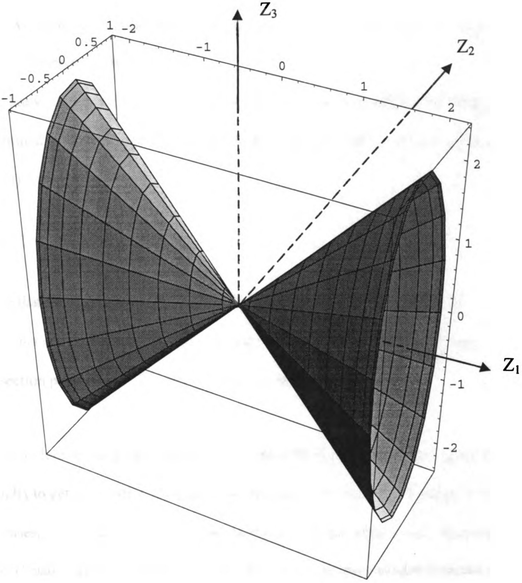
[Figure3K about here]

Figure3K gives the construction of the estimate of ρ in the Z space based on the weight matrix W given in example one. When a given Y is transformed into Z space, and if the corresponding Z is exactly on the surface of the conicoid, we get estimate of $\rho = 0$. If the transformed Z is "inside" or "outside" the cone, we will obtain estimates of $\rho > 0$ and < 0 respectively.

When the dimensional number *n* is larger than 3, we are unable to draw a real figure, but the construction will be the same as an *n*-dimensional conicoid.

3.2. The "Far End" method of the initial value selection.

3.2.1. Ord (1975) suggested using $Y'Y_L / Y'Y$ as an initial value for the Newton - Raphson iteration. In §3.1, we obtained the graphical impression that the estimate of ρ would be pretty close to $Y'Y_L / Y_L'Y_L$, not $Y'Y_L / Y'Y$. It is possible that the latter can be larger than $1 / \lambda_{max}$, and out of boundary $(1/\lambda_{min}, 1/\lambda_{max})$ before any iteration, producing an immediate problem. It could also be the case that the latter is negatively larger than $1/\lambda_{min}$, and out of the boundary as well. Both cases will be verified in chapter 6 when I run a data simulation. On the other hand, it is also possible that although we selected an initial value within the sensible range $(1/\lambda_{min}, 1/\lambda_{max})$, we may not get to a "good" place within the boundary, and the estimate ρ in the next steps may go out of the boundary and take huge values easily. It is "flying out" or "falling out" of the boundary, a difficulty bothering researchers for long. In the following, I talk about the initial value selection problem in detail, and suggest a "far end" method.



$$5.50963Z_1^2 - 4.55287Z_2^2 - .956762Z_3^2 = 0.$$

Figure3K

3.2.2. A new method of the initial value selection based on the construction of W and the relationship between W and Y.

We know that for $f'(\rho) = g(\rho) + h(\rho)$, the derivative of the likelihood function, $g(\rho)$ is determined by W only, and $h(\rho)$ is determined by both W and Y. Looking at the equation for $h(\rho) = 0$, or

$$h(\rho) = 2 \cdot \frac{\rho (Y_L)' Y_L - Y'Y_L}{Y'Y - 2\rho Y'Y_L + \rho^2 (Y_L)' Y_L} = 0,$$

we see that the numerator of the left side is linear in the parameter ρ while the denominator is a quadratic form of ρ, so that the equation has one and only one intersection point $\rho^\nabla = Y'Y_L / (Y_L)'Y_L$ at which $h(\rho) = 0$.

We may look at the graphs Figure3D and Figure3F from the example 1 given in the appendix to get an intuitive impression. In the graphs, we see $\rho_1^\nabla > 0$ and $\rho_2^\nabla < 0$. But in both cases, we see that when $\rho > \rho^\nabla$, we have $h(\rho) > 0$, and when $\rho < \rho^\nabla$, $h(\rho) < 0$. There are two "peaks" symmetrical to the intersection point. Simple calculation shows that the peaks have width $\{(Y'Y)((Y_L)'Y_L) - (Y'Y_L)^2\}^{1/2} / ((Y_L)'Y_L)$, and height $((Y_L)'Y_L) / \{(Y'Y)((Y_L)'Y_L) - (Y'Y_L)^2\}^{1/2}$. The product of the width and height is a constant one. In case the width is small, the peak would be close to the intersection point $\rho^\nabla = Y'Y_L / (Y_L)'Y_L$ and is 'steep', just as $h_1(\rho)$ shown in Figure3D. Otherwise, when the width is not small, the peak appears relatively far from the intersection point ρ^∇ and is 'flat', just as $h_2(\rho)$ shown in Figure3F. Notice that the scales in the figures are different due to the limitation of the computer figuring function.

Assuming $Y'Y_L > 0$ so that $\rho^\nabla > 0$, we know that the solution of the estimate ρ to the equation $f'(\rho) = 0$ is located between 0 and ρ^∇. If the width is small, then the peaks would be close to ρ^∇, and are relatively 'steep'. We see that there would be a twist around the peak area. Once an estimate ρ_r from an iteration falls into this peak area, and it happens that the twist is "sharp", the tangent taken for the next iteration by the Newton tangent method may cut the ρ axis far away, or "fly out" easily. Our best choice (based on intuition) is to take the initial value around .90 of the value $\min(\rho^\nabla, 1/\lambda_{max})$, or even .95 of the value. That is, to take it between 0 and $\rho^\bullet = \min(\rho^\nabla, 1/\lambda_{max})$ and very close to the latter. The selection is the same for the case $Y'Y_L < 0$. We may call this selection as "far end" method.

At last, the estimate of parameter ρ is obtained, and the estimate of parameter σ^2 follows via formula (1.4) from chapter 1.

The verification of the difference between Ord's and my new initial value settings is in chapter 6.

Chapter 4

A TRANSFORMATION FROM THE W, Y SYSTEM TO THE Λ, Z SYSTEM

A DEVELOPMENT OF THE SAM MODEL

In this chapter, I will address problems regarding the parameter ρ, both the technical aspect regarding the estimation of ρ, and the theoretical aspect regarding the function of ρ in the model. There is a new understanding about the boundary regularity $(1/\lambda_{min}, 1/\lambda_{max})$ for the parameter ρ.

4.1. A phenomenon: Why does the estimate ρ approach an extreme value easily?

One big issue regarding the SAM model is the interpretation of parameter ρ, either the real effect of ρ in the model, or the estimate of ρ. In the literature, researchers have mainly focused on the fact that Newton-Raphson maximum likelihood estimates may take extreme values. Efforts were made to restrict this estimate ρ within specified boundaries.

Different boundaries were suggested by Ord (1975), Doreian and Hunmon (1976), Anselin (1982), and Leenders (1995). Recently, some began to consider the situation that the real parameter ρ may take any real values while carrying the idea of finding the boundary for estimate ρ. Commenting on the boundary problem, Leenders (1995) first said that "... the 'appropriate' regularity condition is that ρ estimate may attain any value, except $1/\lambda_i$, i=1,2, ..., g for λ_i real" (page100, Appendix to chapter 3). Then after citing different versions of restrictions in the literature such as "$-1/|\lambda_{max}| \leq \hat{\rho} \leq 1/\lambda_{max}$" from Ord (1975), also from Doreian and Hunmon (1976), and "$1/\lambda_{min} < \hat{\rho} < 1$" from Anselin (1982), Leenders emphasized that the case when the ρ estimate falls out of the boundary "is not even a rare case", and said that "Substantively, however, it may be difficult or even impossible to interpret values of $\hat{\rho}$ that falls outside of the unity interval" as a conclusion (page71).

We know that theoretically, there are *n* solutions of $\hat{\rho}$ to the equation $f'(\rho) = 0$ as we see in Figure3H and Figure3J intuitively. By using the "separation" and "far end" methods I introduced in last chapter, we may find any of these solutions without technical difficulties. But we have a problem of making choices. That is, which one(s) of $\hat{\rho}$ should we choose to be the solution to the model (1.2)? We know that one of the *n* solutions is located in the range $(1/\lambda_{min}, 1/\lambda_{max})$. We also know that $\rho^{\nabla} = Y'Y_L / (Y_L)'Y_L$ is the global minimizer of the random error term of the model (1.2), and ρ^{∇} can be located in or out of the range $(1/\lambda_{min}, 1/\lambda_{max})$ based on both W and Y.

We may address the issue of boundary now. The phenomenon of $\hat{\rho}$ "flying out of the boundary" comes from two different situations.

Situation 1. If ρ^{∇} is located in the range $(1/\lambda_{min}, 1/\lambda_{max})$, then we may start the iteration of ML estimation with an initial value which is ideally located in the range $(1/\lambda_{min}, 1/\lambda_{max})$. The random error term is to be minimized in a global sense. But even doing so, we often get a large value of $\hat{\rho}$ that is far out of the range as a result. If this is what we may call "flying out", this is actually a computational error as I said in §3.2.2., and is easy to avoid by applying the "far end" technique.

Situation 2. If ρ^{∇} is not located in the range $(1/\lambda_{min}, 1/\lambda_{max})$ but another one, say $(1/\lambda_i, 1/\lambda_{i+1})$ where one of $1/\lambda_i$ or $1/\lambda_{i+1}$ could be either $-\infty$ or $+\infty$, then we have a problem of making choices. First choice is, we may prefer to choose the $\hat{\rho}$ which is located in the same range as ρ^{∇} is. By doing so, we have the random error term minimized in a global sense, but the $\hat{\rho}$ is out of the boundary $(1/\lambda_{min}, 1/\lambda_{max})$ and could be very large. Noticing that the similar computational error of "flying out" mentioned in situation 1 may also happen here, we still need to avoid this computational error. Second choice is, we may prefer to choose the $\hat{\rho}$ which is located in the boundary $(1/\lambda_{min}, 1/\lambda_{max})$ so that the $\hat{\rho}$ could not be large, but the random error term is minimized only in a local sense because ρ^{∇} is not in the boundary as we know. In social network analysis, we prefer to make the second choice. The reason is given later in this chapter.

4.2. The technical reason.

Here is the technical reason as to why the estimate of ρ approaches an extreme value: the relationship between W and Y may cause the estimate of ρ to fly out in computation. It's just a technical error.

In the above, we have mentioned that the estimate of ρ may "fly out" of the range $(1/\lambda_{min}, 1/\lambda_{max})$ caused by the twist from the relationship between W and Y. What is more, when W's dimension is relatively large, and the matrix W is not very sparse, it is usual that the eigenvalues λ_{max} and λ_{min} might be relatively large in the absolute value sense, then the range $(1/\lambda_{min}, 1/\lambda_{max})$ would be pretty small, and the "flying out" would occur rather easily. To avoid this technical error, we need only to select an initial value by picking from "far end" as we have said in §3.2.2.

4.3. The theoretical reason.

Here is the theoretical reason as to why the estimate of ρ takes an extreme value: a problem of making choices.

The original autoregression model is

$$Y = \rho WY + \varepsilon,$$

where the error term ε is assumed to have an identically independent (i.i.d.) normal distribution $N(0, \sigma^2 I)$. Now, an orthogonal transformation from W, Y system to Λ, Z system will help to find a rule in the new Λ, Z system. That is, the standard deviation of each *i*th factor Z is adjusted by both the autocorrelation ρ, and the corresponding *i*th eigenvalue λ_i. Let's explore it.

As I have said in chapter 3 (§3.1.3), mathematically, any square matrix W can be transformed into a Jordan canonical standard form which might not always be diagonal with an orthogonal transformation matrix P. But if the order of the eigenvalues of W is fixed, the orthogonal matrix P is unique when W is non-singular. With an orthogonal transformation matrix P for $W = P'\Lambda P$, and writing $PY = Z$, the SAM model is transformed from the W, Y system to the Λ, Z system as

$$Z = \rho\Lambda Z + P\varepsilon.$$

It's well known that the new error term Pε remains i.i.d. normal, and we may rewrite it as ε for the convenience without confusion.

We know that the likelihood function for the model in W, Y system is

$$L(\rho, \sigma^2) = \frac{|A|}{\left(\sqrt{2\pi\sigma^2}\right)^n} e^{-\frac{1}{2\sigma^2}Y'A'AY} \tag{4.1}$$

where A = | I - ρW |.

When it is transformed into the Λ, Z system, the likelihood function is now

$$L(\rho, \sigma^2) = \frac{|I - \rho\Lambda|}{\left(\sqrt{2\pi\sigma^2}\right)^n} e^{-\frac{1}{2\sigma^2}Z'(I-\rho\Lambda)'(I-\rho\Lambda)Z} . \tag{4.2}$$

As I have said in §1.2.4, we may simply rewrite |A| as $\left(|A|^2\right)^{1/2}$ to avoid the negative value problem in both (4.1) and (4.2), and the likelihood function (4.2) becomes

$$L(\rho, \sigma^2) = \frac{\prod_{i=1}^{n}\left((1-\rho\lambda_i)^2\right)^{1/2}}{\left(\sqrt{2\pi\sigma^2}\right)^n} \exp\left(-\frac{\sum_{i=1}^{n} z_i(1-\rho\lambda_i)^2 z_i}{2\sigma^2}\right)$$

$$= \prod_{i=1}^{n}\left(\frac{1}{\sqrt{2\pi\left(\sigma/(1-\rho\lambda_i)\right)^2}} \exp\left(-\frac{z_i^2}{2\left(\sigma/(1-\rho\lambda_i)\right)^2}\right)\right). \tag{4.3}$$

Now we see that for $z_i \sim N(0, \sigma^2/(1-\rho\lambda_i)^2)$, i=1,2,..., *n*, this likelihood function is a product of densities of all {z_i (the *i*th factor)} which are independent but not identical because their variances are different from each other. I discuss the variance $\sigma_i^2 = \sigma^2/(1-\rho\lambda_i)^2$ below.

(1) ρ cannot be a reciprocal of any eigenvalue of W, (namely $1/\lambda_i$) for any index i, or the function (4.3) will make no sense.

(2) When ρ approaches $(1/\lambda_i)$ for any index i, the corresponding variance σ_i^2 approaches infinity, the function (4.3) is converging to zero in a limit sense. This process can be seen in different ways. In mathematics, it is simply a process of limitation. In physics, this is the case when the factor z_i is becoming a white noise, while all other factors are dysfunctioned. In social sciences, this might be the case in which only the factor Z_i is becoming a dominating figure with its variance $\sigma_i^2 = \sigma^2/(1-\rho\lambda_i)^2$ approaching infinity while all other factors are muted with their variances $\sigma_i^2 = \sigma^2/(1-\rho\lambda_i)^2$ remaining finite.

(3) When $\rho = 0$, all z_i have the same variance σ^2, so that all z_i are i.i.d. normally distributed.

(4) Now $\rho \neq 0$. First for the case $\rho > 0$, assuming W has a positive eigenvalue set $\{\lambda_1, \lambda_2, \ldots, \lambda_k\}$ in a descending order.

(i) For all negative eigenvalues λ_j, we have $1 < 1-\rho\lambda_j$, so that all the corresponding variances σ_j^2 are shrunk.

(ii) When $0 < \rho < 1/\lambda_1$ so that $0 < \rho < 1/\lambda_i$ for $i = 1,2, \ldots, k$, we have $1 > 1-\rho\lambda_i > 0$ for all i. In such a case, the variance σ_i^2 is enlarged for all $i = 1,2, \ldots, k$.

(iii) When $1/\lambda_1 < \rho < 1/\lambda_2$, then for $i = 2,3, \ldots, k$, we have $1 > 1-\rho\lambda_i > 0$ and the variance σ_i^2 is enlarged for $i = 2,3, \ldots, k$ only. The variance σ_1^2 behaves as follows.

(a) If $1/\lambda_1 < \rho < 2/\lambda_1$, then $0 > 1-\rho\lambda_1 > -1$, the variance $\sigma_1^2 = \sigma^2/(1-\rho\lambda_i)^2$ is enlarged.

(b) If $2/\lambda_1 < \rho$, then $-1 > 1-\rho\lambda_1$, the variance $\sigma_1^2 = \sigma^2/(1-\rho\lambda_i)^2$ is shrunk.

From (a) and (b), we see that for σ_1^2, when ρ is increasing away from $1/\lambda_1$, the variance σ_1^2 was larger than σ^2 but decreasing to σ^2, and is then decreasing and shrunk less than σ^2.

(iv) The discussion for the following $\lambda_2, \ldots, \lambda_k$ would be the same as for λ_1 in part (ii) and (iii).

(v) For the case $\rho < 0$. We will get the similar series of results as in (i-iv).

Let's summarize the above (1) - (4) and talk about the change of the variance σ_i^2 in the following.

Case 1.

When $\rho = 0$, for all weights λ_i, positive or negative, their variances $\sigma_i^2 = \sigma^2/(1-\rho\lambda_i)^2$ remain the same as σ^2. Neither enlargement, nor shrinkage would occur. This is a "fair and natural" status.

Case 2.

When $0 < \rho < 1/\lambda_1$ where λ_1 is the λ_{max}, for all negative weights, their variances are shrunk. The larger the eigenvalue is (in an absolute value sense), the greater the shrinkage will be.

When $0 < \rho < 1/\lambda_1$, for all positive weights, their variances are enlarged. The larger the eigenvalue is, the more the enlargement will be.

This case shows a kind of "bias". In this case, the factors with negative weights contribute less than it should, and the factors with positive weights contribute more than it should.

We notice that in classical factor analysis, variances are proportional to the corresponding eigenvalues where all eigenvalues are non-negative. But in SAM models, variances are linked to both positive and negative weights, especially to those whose absolute values are large.

Case 3.

When $1/\lambda_1 < \rho < 1/\lambda_2$ where λ_1 is the λ_{max}, for all negative weights, their variances are shrunk (more than they were in case 2). The larger the eigenvalue is (in an absolute value sense), the more the shrinkage will be.

When $1/\lambda_1 < \rho < 1/\lambda_2$, for all positive weights, their variances are adjusted differently. For all positive weights $\lambda_2, \lambda_3, \ldots, \lambda_k$, their variances are enlarged (more than they were in case 2). The larger the eigenvalue is, the more the enlargement will be.

When $1/\lambda_1 < \rho < 1/\lambda_2$, for the positive weight λ_1, its variance is first being enlarged, then shrunk, and shrunk more and more when ρ is becoming larger and larger.

This case shows a kind of "bias and manipulation". In this case, the factors with negative weights contribute much less than they did in case 2. Those factors located right of ρ with positive weights will contribute much more than they did in case 2, while for Z_1, the factor located left of ρ with positive weight corresponding to λ_1, its variance is first being enlarged, then being shrunk. That is, Z_1 will contribute first more then less and less.

Case 4.

When ρ is becoming larger and larger, similar to the case 3, the variances of all negative weights are shrunk more and more.

The counts of factors with positive weights located right of ρ will be less and less, the variances of these remaining positive weights are enlarged much more. Their contributions are artificially increasing.

The counts of factors with positive weights located left of ρ will be greater and greater, the variances of these positive weights are first enlarged but then shrunk much more. Their contributions are artificially first increasing then decreasing.

This case also shows a kind of "bias and manipulation" in a more serious situation. It is biased more and more against the original weights.

Case 5,6,7.

These are cases for $0 < \rho$. They are similar to the case 2,3,4 for $\rho > 0$, and I don't write in details.

Seemingly, all the above cases are essentially a problem of choices. We may take any value of parameter ρ in the model. Taking any non-zero ρ means a kind of bias. When this ρ is becoming larger and larger (in an absolute value sense), the weights of this system are much more seriously adjusted, and less original. We may choose an extreme value for ρ for a special reason, but it is mostly less meaningful.

One example might be the case of a congress meeting where each congressman is represented by a factor. Each one is with his eigenvalue either positive or negative depending on the case. Senior ones carry large eigenvalues and large variances, and influence the system more than those junior ones who carry small eigenvalues and influence the system less. In case parameter ρ goes out, say positively, of the boundary, then some special situation happens. Assuming the parameter ρ is now between the reciprocals of the largest and the second largest eigenvalues, then not only the variances of all negative eigenvalues are being shrunk, the variance of the largest eigenvalue is also being shrunk. The variance of the second largest eigenvalue, and the following ones, are being enlarged at the same time. Back to the congress meeting, such a situation means that not only the influences of all congressmen from negative side are shrunk, but the influence of the first leader from positive side (represented by the first factor) is also shrunk, while the influence of the second leader from positive side (represented by the second factor) is enlarged, and so are those following positive ones. It might be because the first leader from positive side is constrained, or his role is ignored, either one is causing the system unstable. When the parameter ρ moves rightwards further and further, this situation is becoming more and more serious. At last, it could be such a case that the variances of all factors with negative eigenvalues and almost all factors with positive eigenvalues are shrunk, while the variances of the remaining one or two factors with very small positive eigenvalues are highly enlarged. In the congress, it means that only one or two very junior persons are making decision while all others, senior and junior ones are constrained. We should believe that this kind of decision could not be very much stable, and the system itself is not stable at all.

Chapter 5

THE DEFINITION OF PARAMETER ρ AND THE W(a) FAMILY OF THE WEIGHT MATRICES

In this chapter, I first give parameter ρ a general definition as an indicator of consistency between W and Y in the model. Second, I develop a W(a) family that is a general format of the weight matrix W with a constant "a" as the principal diagonal elements. The assumptions of $w_{ij} \geq 0$ ($i \neq j$) and $w_{ii} = 0$ carried on in chapter one through chapter four are not required now. Specifically, we now consider W's principal diagonal elements "a" as a constant, taking zero or non-zero values. Based on the understanding of the W(a) family in SAM models, I try to extend the concept of the classical Factor Analysis method to a more general level.

5.1. The definition of the parameter ρ.

In Xu (1996), an early paper working on spatial autoregression models, I described "The autocorrelation parameter ρ in the model indicates the extent of the members' communication-cooperation within this group and is important to estimate." Now I give the parameter ρ a definition below.

In spatial autoregression models (SAM), the parameter ρ is an indicator of the extent of consistency between a weight matrix W and an observation vector Y. With λ_{max} and λ_{min} as the maximum and minimum eigenvalues of W, if W and Y are more consistent, the estimate of ρ goes positively higher, and intends to reach $1/\lambda_{max}$, the maximum of the range of estimate of ρ. If W and Y are less consistent, the estimate of ρ goes negatively higher, and intends to reach $1/\lambda_{min}$, the minimum of the range of estimate of ρ. The word of "consistency" is in an *n*-dimensional Euclidean distance sense based on SAM.

Now we may consider the parameter ρ as a correlation coefficient between a general matrix W which is not necessarily a variance – covariance matrix, and a vector Y. It seems to be a new measurement in statistics for the social sciences.

5.2. The geometrical reasoning of the definition of the parameter ρ.

The transformed Z space as in chapter 4 is spanned by factors $Z_1, Z_2, ..., Z_n$ with the weights of the corresponding eigenvalues $\lambda_1, \lambda_2, ..., \lambda_n$ respectively. Notice that the Z space has its combined direction Z^* which is essentially the combination of all these axes weighted by their directional weights $\{\lambda_i\}$, the corresponding eigenvalues.

When an observation Y° is transformed into Z°, the estimate ρ is actually a projection of Z° on Z^*, the combined direction of $\{Z_i\}$ in the Z space. If the value of this projection is positive and high, we see that Z° is consistent with Z^*, the combined direction of Z space, and we say that Y° is positively consistent to the matrix W. If the projection is negative and high, we say that Y° is negatively consistent with the matrix W. That is why we may consider the parameter ρ as a correlation coefficient between a vector Y and a weight matrix W, a new concept in statistics.

Z space is a multidimensional space. It could happen that a vector Y is exactly transformed onto one axis Z_i in Z space. But in most cases, the corresponding Z from Y is usually "between" those axis Z_i. So that mathematically, I would prefer to call this new technique as a "non-eigenvector analysis".

5.3. A comparison of SAM with linear models.

In a sense, the function and behavior of estimate of ρ seems to be similar to the estimate *r* of the correlation coefficient between two vectors X and Y. This estimate ρ is substantively a kind of correlation coefficient, but it is between a vector Y and a matrix W.

In Table5A, the second column refers to a correlation coefficient analysis between two standardized variables X and Y. The model is

$$Y_i = rX_i + \varepsilon_i,\ \varepsilon_i \sim N(0, \sigma^2),\ i=1,2,\ldots, n$$

with $cov(\varepsilon_i, \varepsilon_j) = 0$ when $i \neq j$. The parameters r and σ^2, the variance of the random error term need to be estimated.

The third column in Table5A refers to a spatial autoregression model with a weight matrix W and an observation vector Y. The model is

$$Y = \rho WY + \varepsilon,\ \varepsilon \sim N(0, \sigma^2 I).$$

The parameters ρ and σ^2, the variance of random error term need to be estimated.

All values a, c, B, D in Table5A are constants.

[Table5A about here]

I briefly compare the differences between these two models as below. Some more comparisons can be seen in chapter 7 (§7.1).

- The solution of SAM model is not an algebraic one, so that an analytic formula for the solution of $\hat{\rho}$ is not available.
- The range of r in linear regression model is [-1, 1] when X and Y are standardized. But in SAM model, the range of $\hat{\rho}$ is $(1/\lambda_{min}, 1/\lambda_{max})$ which can not be standardized to [-1, 1].
- The properties are easy to verify.

5.4. The W(a) family.

In order to consider the estimate of ρ from weight matrix W in which all elements of the principal diagonal are a non-zero constant "a", we need to develop the W(a) family, a more general concept of the weight matrix W for the following reasons.

As we know, w_{ii} is usually treated as zero in the literature. This property guarantees that the summation of W′s eigenvalues equals to zero, or $\Sigma\lambda_i = 0$ which helps us to obtain the conclusion in chapter 4 that "the values of estimate of ρ and Y′WY have the same sign within the boundary", and to obtain the construction of the distribution of the estimate of ρ in space. At the same time, the importance of the condition of $w_{ii} = 0$, and the case where $w_{ii} \neq 0$ might not be emphasized enough in the literature especially in social sciences. For instance, Leenders (1995) says (page 54) that "An entry w_{ij} of W denotes

Table5A

	estimate r	estimate ρ
model	$Y_i = rX_i + \varepsilon_i$ $\varepsilon_i \sim$ iid normal $cov(X_i, \varepsilon_i) = 0$	$Y = \rho WY + \varepsilon$ $\varepsilon \sim N(0, \sigma^2 I)$ $cov(Y, \varepsilon) \neq 0$
definition	correlation coefficient between a vector Y and a vector X	correlation coefficient between a vector Y and a matrix W
analytic formula	available	not available
range	[-1, +1]	$(1/\lambda_{min}, 1/\lambda_{max})$
unit and scale	no scale	no scale
sign	"+" means positively consistent and vice versa	"+" means positively consistent and vice versa
property	$r(aX+C,by+D) = r(X,Y)$. $\lvert r \rvert \leq 1$. $\lvert r \rvert = 1$ if $P(Y=bX) = 1$.	$\rho(aY,bW) = (1/b)\rho(Y,W)$ $1/\lambda_{min} < \rho < 1/\lambda_{max}$ $\rho=1/\lambda_{min}$ if $P(Y=Y_{\lambda min})=1$, $\rho=1/\lambda_{max}$ if $P(Y=Y_{\lambda max})=1$.

the influence actor j has on actor i". Then we may ask: does an entry $w_{ii} = 0$ mean the actor has nil influence on him/herself? In other cases, if an entry w_{ij} represents the level of trust subject j gives subject i, or w_{ij} is a measure of psychological links such as the "intention" of communicating with others, then it's hard to say that $w_{ii} = 0$ means that the ith ego trusts himself at a nil level, or the ith subject intends to communicate with himself not at all. Both statements seem to be not very persuasive, and we need to work on the case where $w_{ii} = a \neq 0$ for i=1,2,..., n. This turns out to be a W(a) family problem.

W(a) is a weight matrix, real and possibly symmetrical. The letter "a" denotes that the elements on W's diagonal equal a constant "a". This "a" can be positive, and possibly negative which may have less practical meaning. When a = 0, the matrix W(0) is the weight matrix W we have been working on till now. When a = 1, and if the absolute values of W's all entries w_{ij} are less than or equal to 1, this W(1) has all eigenvalues non-negative, and can represent a variance-covariance matrix for a data set. In other words, all variance-covariance matrices belong to a subfamily of W(1). We obtain an important generalization of W.

In previous chapters, we have been working on W(0). For W(0), we obtained an orthogonal matrix P, eigenvalue set $\{\lambda_i\}$ for the diagonal matrix Λ, and the maximum, minimum eigenvalues λ_{max}, λ_{min} etc. Now we rewrite them in the "W(a) family" format with "a" as a constant. That is, for a given W(a), we have the corresponding P(a), $\{\lambda_i(a)\}$, Λ(a), and $\lambda_{max}(a)$, $\lambda_{min}(a)$ etc. If a = 0, we are back to the simplified version.

Easy linear algebra shows that we always have the following properties:

1. $P(a) = P(0)$. That is, the transformation matrices are identical for the $W(a)$ family no matter what value of "a" is.

2. $\{\lambda_i(a)\} = \{\lambda_i(0) + a\}$. That is, for $W(a)$ and $W(0)$, we have $\lambda_i(a) = \lambda_i(0) + a$ for $i=1,2, \ldots, n$. Seemingly, we have $\lambda_{max}(a) = \lambda_{max}(0) + a$ and $\lambda_{min}(a) = \lambda_{min}(0) + a$.

3. $\Lambda(a) = \Lambda(0) + aI$ when the order of the eigenvalues are fixed. We always put them in a descending order in our discussion. This is actually a direct result from 2.

Easy computation also shows the following properties (4-6) which I will apply in chapter 6 when I calculate the ratios of the estimate ρ to the right bound $1/\lambda_{max}$ or to the left bound $1/\lambda_{min}$.

4. Let $W^* = kW$ where k is a non-zero scalar, then for $W = P'\Lambda P$ with eigenvalue set $\{\lambda_i\}$, we have $W^* = P'(k\Lambda)P$ with eigenvalue set $\{\lambda_i^*\} = \{k\lambda_i\}$. Specifically, we have $\lambda_{max}^* = k\lambda_{max}$.

5. In the SAM models, if an estimate of ρ is obtained over a given Y, then the corresponding estimate ρ^* over the same given Y and the new $W^* = kW$ will be ρ/k.

6. When k varies, the new estimate $\rho^* = \rho/k$ will vary. But we have

$$\rho^* / (1/\lambda_{max}^*) = (\rho/k) / (1/k\lambda_{max}) = \rho / (1/\lambda_{max}).$$

So, this ratio is invariant against scalar k, just as the projection of Y on W is invariant in §5.2.

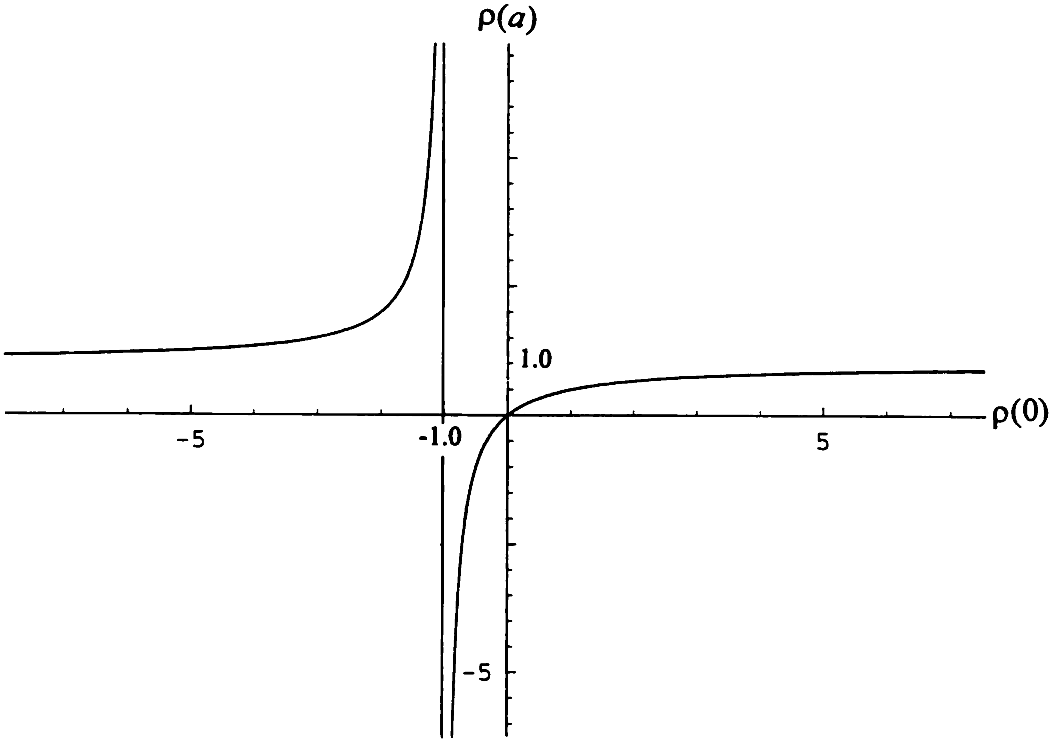
Now I consider the following situation. With a given Y, when dealing with two different weight matrices, we would obtain different ρ estimates. I claim that a direct comparison of these two estimates ρ might be not appropriate. The reason is, consider a first matrix with fairly large eigenvalues λ_{max} and λ_{min} that would cause a small boundary. Then the obtained estimate of ρ could not be a large one even though it might be very close to the value $1/\lambda_{max}$ which means the ratio of the obtained estimate ρ to the value $1/\lambda_{max}$ is large. Suppose the second matrix has comparatively small eigenvalues λ_{max} and λ_{min} which would cause a large boundary. The obtained estimate ρ might be relatively large in an absolute value meaning compared with the first ρ estimate. However, it is also possible that the second ρ estimate is actually not very much close to the value $1/\lambda_{max}$, and the corresponding ratio is small. This is exactly what happens in chapter 6 when I conduct a data analysis and compare ρ estimates from two different but related (24×24) weight matrices. The invariance of the ratio will help for our comparison in §6.2.2.5.

Applying the above properties 1 - 3, I get the following relationship easily:

$$\rho(a) = \frac{\rho(0)}{1 + a\rho(0)} \tag{5.1}$$

[Figure5A about here.]

The relationship between ρ(a) and ρ(0) from (5.1) can be seen in Figure5A. It is a hyperbola. In the figure, the X axis represents ρ(0) which is the ρ estimate obtained from a given Y and W(0), and the Y axis represents ρ(a) which is the ρ estimate obtained from the same given Y and W(a) that is the same as W(0) except its principal diagonal elements "a". There is a one-to-one correspondence between ρ(a) and ρ(0). Taking derivative to equation (5.1) gives



$$\rho(a) = \frac{\rho(0)}{1 + a\rho(0)}\Big|_{a=1}.$$

The X axis represents the estimate of ρ(0) from Y and W(0).
The Y axis represents the estimate of ρ(a) from Y and W(a).

Figure5A

$$\rho'(a)\Big|_{\rho(0)} = \frac{1}{(1+a\rho(0))^2} \tag{5.2}$$

The expression of (5.2) shows that this correspondence is strictly monotonely increasing since the right side of (5.2) is always positive. In summary, we see that the curve meets the origin, and the slope of the curve at the origin equals 1 which is easy to verify, and is for any value of "a", so that the two estimates ρ(a) and ρ(0) are almost equal around the origin for any value "a". When parting from origin, the curve gradually begins to change in slope.

When ρ(0) approaches the value of $-\frac{1}{a}$ from right side, we see that the hyperbola approaches negative infinity. The hyperbola approaches positive infinity when ρ(0) approaches $-\frac{1}{a}$ from left side. In such a case, the one-to-one relationship is preserved, but the range jumps dramatically. Easy mathematical work shows that when a = 1, the hyperbola of (5.1) is symmetrical to the line of ρ(a) = -ρ(0). So that, if the value of "a" is around 1, we may have the ideal mapping between ρ(a) and ρ(0). That is, ρ(a) and ρ(0) are fairly approximately close around zero.

5.5. The dynamic system in a social setting of the trio: estimate ρ, W and Y.

By the definition of parameter ρ given in §5.1, ρ is the indicator of the level of consistency between W and Y, or the estimate of ρ is a function of W and Y. A new question now is, can we see the model in other way? Can we treat Y as the function of W and ρ when minimizing the random error terms?

5.5.1. When W is fixed, and parameter ρ is given, what is the expected Y?

This is actually to solve a stochastic equation as

$Y=(I - \rho W)^{-1}\varepsilon.$

The solution of Y appears not to be meaningful since this is only a transformed random error vector. But, as we know from chapter 4, the solution Y is a subspace in a super-conic shape in a multiple space. It can be seen as the best fit of Y to the weight matrix W. Of course, it contains infinitively many solutions, but when some marginal conditions are available, or some of the prior information about the solution Y is available, then the subspace of solution Y will be specified. In this sense, we are looking for the "best fit" of Y to W at the given ρ level.

5.5.2. A note to equation (5.1)

For the moment, we consider the value of "a" non-negative. When "a" is negative, it is less practically meaningful and we are not going to talk about that here. We will talk about the case when "a" is negative including "$a \to -\infty$" in chapter 7 where we consider it as an extension of the classical factor analysis technique in a more general sense.

We see that when $a \to +\infty$, we have from (5.1) that $\rho(a) \to 0$ if $\rho(0) \neq 0$. Also, when a $\to 0$, we have $\rho(a) \to \rho(0)$ for any $\rho(0)$. When the value of "a" varies, we may obtain information from these subjects as they relate with others in the network setting.

When a = 0, these egos (y_i) are simple information processors. They don't have ideas from themselves, but make decisions totally based on other subjects' attitudes. We see that the case "a = 0" means these subjects are totally objective.

When a > 0, these subjects in the model become more than simple information processors. The model effect is not purely objective. These egos (y_i) make decisions not only based on others' information but also based on their own. When "a" is positively increasing, the weights of these subjects' own information are increasing. That is, the positively increasing "a" indicates that these subjects are more and more self-centering.

The extreme case when "$a = +\infty$" means these subjects are totally subjective. It's easy to understand that in matrix W, $a \rightarrow +\infty$ actually means that all w_{ij} (for any $i \neq j$) are approaching 0, and we find that these egos (y_i) don't want to accept information from any others for their own consideration. They are purely subjective.

Practically, subjects in a social network are mostly between two extremes. Neither "$a = 0$" nor "$a = +\infty$" are practical and ideal. We prefer to have "a" at an appropriate level between 0 and $+\infty$. We see that when a is close to 0, these egos are more likely cooperative and less self-centered; when a is becoming larger and larger, these egos are more likely self-centered and less cooperative.

Chapter 6

DATA ANALYSIS

In this chapter, I conduct a data analysis. There is a data simulation in part I and an example in part II. I use my new Newton-Raphson technique in both parts. The result is the realization of the definition I gave in chapter 5 to the parameter ρ in the SAM models.

6.0. The source of data.

In chapter 4 of Frank (1996), a weighted data set collected in January 1993 was introduced. The initial motivation for using this data set was to identify cohesive subgroups from professional discussions among teachers from a high school named "Our Hamilton High" located in the Chicago area. In this set, a group of 24 teachers of the school were surveyed. In Frank (1996) (page106), the collected data include these teachers' gender, race, years of teaching in the school, and their level of moral agency which was one of four measures "based on each teacher's extent of agreement (1 = strongly disagree to 4 = strongly agree) with items referring to the teacher's sentiments and orientation towards teaching". Another variable p, the level of the frequency of their professional discussions, was collected such that "each teacher listed the five (or fewer) teachers with whom he or she had most often discussed professional matters during the 1992-1993 school year. The teachers were asked to weight the frequency of discussions (1 = less than once a month, 2 = two to three times a month, 3 = once or twice a week, and 4 = almost daily)" (page 99) (see Appendix6A). Based on the levels of the frequency of their discussions, an association (weight) matrix W (24×24) was constructed in the following way. If the *i*th teacher indicated engaging in professional discussion with the *j*th teacher at level p, the frequency of discussions (p might take values either 1, 2, 3 or 4), then the entry $w_{ij} = p$. If the *i*th teacher didn't talk with the *j*th teacher, then $w_{ij} = 0$. Noticing that the conversations were not mutually initiated, we observe that the weight matrix W is asymmetrical.

Now I will use the information mentioned above to conduct my data analysis in both part I and part II. Other available information collected from that school included the teachers' subject field (the courses they were teaching) and their office numbers, but I do not use them here.

6.1. Part I. Data simulation.

6.1.1. The goal of the data simulation.

The goal of the data simulation is to check the level of goodness and efficiency of my new technique and to evaluate the effects of using different initial values in the estimation. Ord (1975) suggested using the initial value $(Y'Y_L / Y'Y)$, and I suggest using $(Y'Y_L / Y_L'Y_L)$ as I have said in chapter 3 (§3.2.1)..

6.1.2. Simulating data in a spatial autoregression model (SAM).

In an initial SAM model

$$Y = \rho WY + \varepsilon \tag{6.1}$$

where $\varepsilon \sim N(0, \sigma^2 I)$, the weight matrix W is known, Y is observed. The researcher estimates the parameter ρ by optimizing the likelihood function from model (6.1).

In this part, I first create a multivariate random error term which is i.i.d. normal. With a given weight matrix W, a pre-chosen value of the parameter ρ∘ (I denote the value as ρ∘ here to tell the difference from the original ρ) within the boundary of $(1/\lambda_{min}, 1/\lambda_{max})$, I calculate Y using the formula

$$Y = (I - \rho_\circ W)^{-1} \varepsilon \tag{6.1$'$}$$

which is from model (6.1).

With the calculated Y as observed, I start the process of estimating the parameter ρ by optimizing the likelihood function from model (6.1) to get the estimate ρ (which would be different from the pre-chosen ρ∘).

For the same pre-chosen ρ∘, I repeat the above steps a number of times by using different random errors I created, so that the calculated Y would be different. Then I obtain a series of parameter ρ estimates. By getting the distribution of the obtained series of estimate of ρ and comparing with the pre-chosen ρ∘, I may find the extent of efficiency of the maximum likelihood technique for the parameter estimation.

I would also repeat the above steps a number of times using other different pre-chosen ρ∘ values to find the extent of efficiency of my new technique when the pre-chosen ρ∘ is ranging within the preferred boundary of $(1/\lambda_{min}, 1/\lambda_{max})$ from left side to right side.

Again, I would use the same series of multivariate random error terms I created before, and the same pre-chosen ρ∘ values to repeat all the previous procedure. But this time, I use Ord's (1975) initial value ($Y'Y_L$ / $Y'Y$) instead of mine. Thus, I would have two series of estimates ρ for each pre-chosen ρ°. By comparing the two different series, we may have a better understanding of my new technique.

6.1.3. The plan of the work.

In an early version, I used a subgroup A(4×4) of the W matrix from Frank (1996) (see Appendix6A) and found important differences. It was found that using Ord's initial value setting, we got about 23% of estimates ρ within the required range ($1/\lambda_{min}$, $1/\lambda_{max}$) and the percentage of violations to the requirement is 77%, a fairly large percentage (see Appendix6B). While using my initial value setting with the "far end" technique, we got all 100% of estimates of ρ being within the required range ($1/\lambda_{min}$, $1/\lambda_{max}$). Seemingly, the new initial value selection method gives much better results.

Since this A(4×4) is rather small, now I use a combination of subgroup A and subgroup B from Frank (1996) for my example. Teacher ID 17 and ID 8 in subgroup B were removed for simplicity, and I rewrite this matrix as A(8×8). A is asymmetrical and contains a pair of conjugate eigenvalues. Its maximum eigenvalue is 6.87348 and the minimum eigenvalue is –3.43211. The boundary from the reciprocal of the maximum and minimum eigenvalues then is (-.291366, .145487) = (B_L, B_R) where B_L and B_R represent left bound and right bound of the required boundary.

I use nine initial values as the pre-chosen ρ° in the following way. They are listed from left to right as

$$\{\rho^\circ_i\} = \{.90B_L, 70B_L, .50B_L, .30B_L, 0, .30B_R, .50B_R, .70B_R, .90B_R\} \qquad (6.2)$$

That is,

$$\{\rho^\circ_i\} = \{-.26223, -.20396, -.14568, -.0874, 0, .04365, .07274, .10184, .13094\} \qquad (6.2)'$$

We see that the left four are negative, and right four are positive. The middle one is zero.

With a seed =3837, mean value = 0 and variance = 1, I use SAS programming to generate a series of random vectors (8×1) as the random error term ε_i for i = 1 to n. For simplicity, I chose the number n =30, not a large one. But even with n at this level, we may find that the obtained estimates ρ are fairly normally distributed around the pre-chosen ρ° values respectively.

By using each of these nine pre-chosen initial values from (6.2), I calculate Y_i for i = 1 to 30 based on (6.1)′. With A known, the Y_i calculated as observed, I obtain nine series of estimates ρ_i (i = 1 to 30) for each of these nine initially pre-chosen values. The distributions of these nine series of estimates of $\{\rho_i\}$ (i = 1 to 30) are easy to obtain. Actually, all the above steps of the plan was conducted in my SAS program 1 (see Appendix6C). The flow chart of the program is shown in Appendix6D.

I also prepared another SAS program 2 (see Appendix6E). Program 1 and 2 are substantially the same except one minor difference. In program 1, the obtained estimate of the parameter ρ is restricted within the preferred boundary, namely ($1/\lambda_{min}$, $1/\lambda_{max}$). If the estimate is out of the boundary, I treat it as missing so that the number of obtained estimates of the parameter ρ is usually less than 30 (see Appendix6F). But in program 2, this restriction is taken off, so that the total 30 estimates will be printed out which can be located anywhere on the real line, within or out of the boundary (see Appendix6G).

6.1.4. Discussion of the results.

In general, the new technique gives most of the estimate ρ located within the boundary. Generally but not necessarily, when the pre-chosen $\rho_\circ$ is close to the left bound, some of the estimates are out of the left side of the boundary; when the pre-chosen $\rho_\circ$ is close to the right bound, some of the estimates are out of the right side of the boundary.

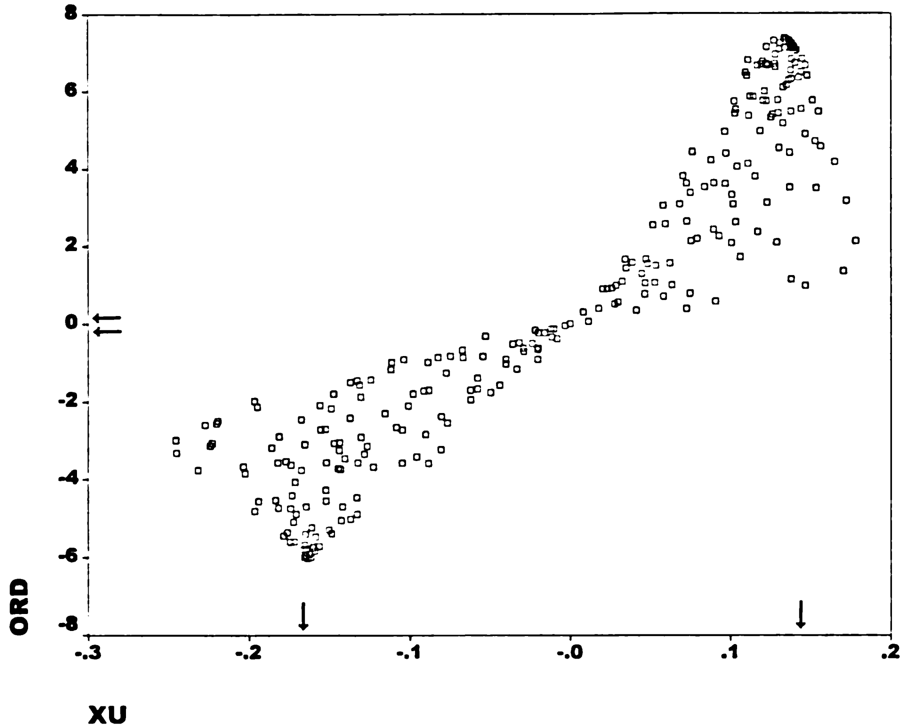
6.1.5. Comparisons between different strategies of the initial value selections.

As I planned above, I re-conduct the same program again using Ord's suggested initial values. Now I use $(Y'Y_L / Y'Y)$ instead of $(Y'Y_L / Y_L'Y_L)$ as the initial value. I found that the difference in results is obvious. A large portion of estimates of ρ flew out of the boundary when I used Ord's method of the initial value selection.

Figure6A helps to understand the difference between these two strategies of initial value selection. Remember I got a total of 270 8-dimensional y vectors in the above simulation. I used both Ord's and my strategies of initial value selection to get initial values for iteration. Now, each of the 270 y vectors corresponds to a dot in Figure6A with the coordinate (x, y) = (Xu's initial value, Ord's initial value). Noticing that the scales for the horizontal and vertical axes in Figure6A are not equal, I also give individual histograms to compare the two strategies of the initial value selection in detail (Figure6B and Figure6C).

[Figure6A, 6B and 6C about here]

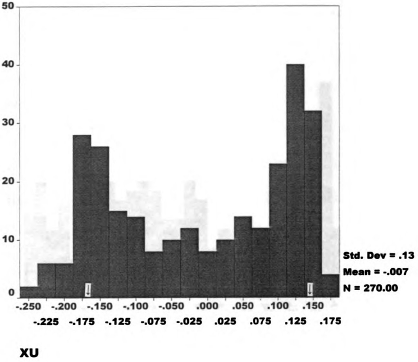
Figure6B is obtained from the initial values using my method. We may find that when using my method, the initial values distributed pretty well within the boundary. That is,



For each dot in the graph,
the X coordinate is Xu's initial value, ranging (-.245376, .178482),
the Y coordinate is Ord's initial value, ranging (-6.0202, 7.363602).

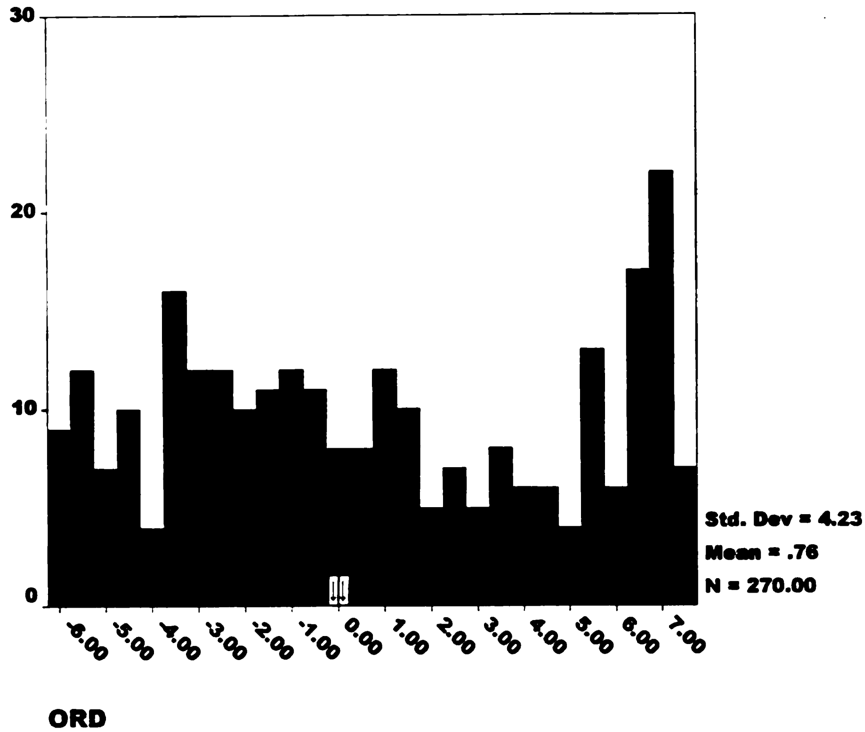
The required boundary is (-.1692563, .145487) which is between two arrows on each coordinate.

Figure6A



Distribution of Xu's initial values.
The required boundary is (-.1692563, .145487) which is between two arrows.
About 15.93% of Xu's initial values are out of boundary.

Figure6B



Distribution of Ord's initial values.
The required boundary is (-.1692563, .145487) which is between two arrows.
About 97.78% of Ord's initial values are out of boundary.

Figure6C

about 84% of the estimates based on my initial values are located within the boundary, about 5% are out from right side and 10% are out from left side. We know from chapter 3 that the Cauchy mean value theorem plays an important role in locating the estimate ρ. With an initial value within the boundary, the estimate would be also located within the same boundary, an ideal result.

Figure6C is obtained from the initial values using Ord's method. Noticing the scales of Figure6B and Figure6C are different, we may find that when using Ord's method, the initial values are not distributed ideally. That is, about only 2% of that initial values are located within the boundary, 50% are out from right side and about 47% are out from left side. Once the initial value is out of the boundary in the beginning, by the Cauchy mean value theorem, the final estimate is most likely to be located in the same interval, thus a poor result occurs.

We may change the seed value when creating the random error term, and the result might be a little different, but the situation of "large portion being out of the boundary" will remain the same when using Ord's method of initial value selection.

6.2. Part II. A practical example.

6.2.1. The goal of the practical work.

Here I apply the new technique to work on a practical data set. The theoretical base is, the parameter ρ in a spatial autoregression model is an indicator of the level of consistency between W, the weight matrix and Y, the observed real outcome. We want to compare the different levels of consistency between W, and four observations (Y).

6.2.2. Steps of the practical work.

6.2.2.1. To obtain the eigenvalues of W.

I use the whole weight matrix W (24×24) from Frank (1996) (see Appendix6A). This W matrix is asymmetrical, and I used Fortran programming to get $\{\lambda_i\}$, the eigenvalue set of W. (see Appendix6H).

6.2.2.2. In this school, "sex", "race", "year of teaching" and "moral agency" are four variables obtained from these teachers. Writing them in a vector format Y (24×1), I use each of these Ys in SAM model to estimate ρ, the correlation coefficient between W and each Y. In such a way, I may find an answer to the question: among those variables "sex", "race", "year of teaching" and "moral agency", to which one(s) is the pattern matrix W most highly consistent?

Coding sex, I have male = 1 and female = 2.

Coding race, I choose two different versions W1NW2 and NB1B2.

In W1NW2, I code white =1, non-white = 2.

In NB1B2, I code non-black = 1, black = 2.

The reason why I do so is because of the existence of the two "others". They might be coded as non-white, or as non-black. We may even code them with the middle value 1.5 which might be a little strange but make some sense. There are only two "others" in this school. I found that their level of moral agency are relatively low, and different coding technique may affect the analysis result especially when dealing with moral agency. A comparison shows that black teachers had higher moral agency than whites (and others) with a mean difference around -.25 which is not significant with p value around .078 (see Appendix6I).

Note the value of moral agency for teacher ID 24 who is a white male is missing. In such a case, I may redo the analysis at a dimension number 23 (one unit less than 24) level, and may still obtain general relationship between W and Y. But it would not be consistent with analyses based on other variables (sex, race and year of teaching). To be consistent, I assign teacher ID 24 a value of moral agency .21391 from teacher ID 11 who was similar to ID 24 in terms of race, gender, and year of teaching. Also, I assign teacher ID 24 another value 0.19042 that is the average value of moral agency over all white males. As a contrast, I assign him a value of moral agency –.01690 from teacher ID 18 who is a black male, teaching the same course of physical education for similar years as a contrast.

6.2.2.3. The real work of estimation using model (6.1).

With W given and Y_{race}, Y_{sex} (with two coding methods), Y_{year} and Y_{moral} (with $moral_{24}$ = .21391, as well as $moral_{24}$ = .19042 and $moral_{24}$ = -.01690), I estimate those ρs. Using Ord's formula (1975, page 124), I also calculate the standard deviation of those ρs. Ord's formula might be problematic, and a brief discussion is in §6.4.

[Table6A about here]

The result is shown in Table6A. We find that both ρ_{sex} and ρ_{race} are a little larger than .06, ρ_{year} is slightly below .06. These three values of ρ_{sex}, ρ_{race} and ρ_{year} are pretty close.

Table6A

W's entries are 0, 1, 2, 3, 4.

	Sex	NB1B2 Non-black =1 Black=2	W1NW2 White=1 Non-white =2	Year of teaching	Moral (ID 24= .21391)	Moral (ID 24= .19042)	Moral (ID 24= -.01690)
Estimate ρ^*	.06051 (.01909)	.06147 (.01784)	.06340 (.01498)	.05910 (.02076)	.04653 (.02932)	.04644 (.02936)	.04503 (.02990)
% of Estimate ρ to the $(1/\lambda_{max})$	85.08%	90.24%	89.13%	83.09%	65.42%	65.29%	63.31%

* Figures in parentheses are standard deviations.

λ_{max} = 14.059360 and $1/\lambda_{max}$ = .0711270.

λ_{min} = -7.701774 and $1/\lambda_{min}$ = -.1298402.

All of these ρ_{sex}, ρ_{race} and ρ_{year} are seemingly larger than ρ_{moral} which is at around .046 level, no matter ID 24 takes which value for his moral agency. All estimates are within two standard deviations of one another, even if using the smallest standard deviation of .015. We notice that all ρ_{sex}, ρ_{race} and ρ_{year} are selection effects, effects of people choosing to have professional discussion to others like themselves. That is, people may decide the level they choose subjects to talk with just because of that person's gender, or race, or because of that person's seniority. The ρ_{moral} is the only effect which could be considered as influence.

6.2.2.4 Discussion of the table.

I make the following observations:

- W is positively associated with race with ρ_{race} at .06 level (89%). It might mean that when these teachers choose subjects to initiate a professional discussion, the race is a main element to consider. They prefer to choose subjects of same race to have professional discussion.
- W is also positively associated with sex with ρ_{sex} = .0605 (85%). The levels of the associations of W with sex and with race are close although sex is slightly lower. It might mean that when these teachers choose subjects to initiate a discussion, sex is also a main element to consider. They prefer to choose subjects of same sex to have professional discussion. It is not clear whether or not sex and race are associated among those teachers. A correlation coefficient matrix of all variables including sex, year, two different versions of race, three different versions of moral agency shows that there is no strong evidence to show that sex and race are highly correlated with r

= .299 and .293 for two different versions of races (W1NW2 and NB1B2) (see Appendix6J).

- With ρ_{year} = .059 (83%), W is also positively associated with years of teaching but at a slightly lower level compared with race and sex. This is a relative comparison. It does not confirm that year of teaching is not important compared with race or with sex. In next chapter (§6.2.5), I would explore further.
- With ρ_{moral} at .046 level (65%), Moral agency is less consistent with W than three other Ys in the analyses. There is not much difference among the three different values I assigned to the ID 24.

To summarize, we get a range of 60% ~ 90% of the ratio (estimate of ρ) / ($1/\lambda_{max}$) in the ρ estimation.

6.2.3. Compare W(0,1,2,3,4) with W(0,1).

Dealing with the weight matrix W expressing connections between the regions in spatial regression models, Ripley (1981) suggested (page 98) that "This can be just a binary matrix giving 1 if the two regions have a common boundary, 0 otherwise, or it could depend on the length of the common boundary, the distances between the regions or transport costs between them". Comparing the binary measure with the length of the common boundary, we see that the author is actually talking about the coding or re-coding of the weight matrix W. Ripley also said that "Bartels (1979) suggests that simple binary weights have proved as adequate as more complex schemes". Let's try to explore this idea.

From now on, we consider another weight matrix W* which is related to the weight matrix W I used in the above. In §6.2.1, the weight matrix has cell ranging from 0 to 4. Now I dichotomize the value as 0:0, 1→4:1. That is, I re-code W into W* as $w^*_{ij} = 0$ if $w_{ij} = 0$, and $w^*_{ij} = 1$ either $w_{ij} = 1,2,3$ or 4. We know that the level of w_{ij} indicates the frequency of the teachers' discussions. When W is re-coded into W* in such a way, the different level of frequency is reduced to a simple "yes or no" level, a binary one.

I repeat the same work for W* as I did for W in §6.2.2.3. The eigenvalues of W* are different from those of W (see Appendix 6K).

Using W*, I re-estimate the autoregression coefficient ρ by using the same Y values: Y_{race} , Y_{sex} , Y_{year} and Y_{moral}, and obtained the results shown in a table (Table6B). Interesting enough, we find that for the W*, the ratios of ρ_{sex}, ρ_{race} (two methods of coding) and ρ_{year} to $(1/\lambda_{max})$ are similar to the corresponding values obtained from W. We notice that W and W* have different eigenvalue sets. The λ_{max} and λ_{min} from W is 14.059360 and –7.701774; while the λ_{max} and λ_{min} from W* is 4.051115 and –2.083131. It is appropriate to use the ratio of estimate ρ to the bound $1/\lambda_{max}$ or $1/\lambda_{min}$ to compare the results as we introduced in §5.4 and applied in §6.2.2.4. The results are shown in Table6B. The ratios of ρ_{moral} to $(1/\lambda_{max})$ are relatively lower in W* than in W: they were at around 65% levels from W, but now they are at around 51% levels from W*. The ratio of ρ_{race} to $(1/\lambda_{max})$ and that of ρ_{sex} to $(1/\lambda_{max})$ are remaining at the same levels in W* as they were in W.

[Table6B about here]

Table6B

W's entries are 0, 1 only.

	Sex	NB1B2 Non-black =1 Black=2	W1NW2 White=1 Non-white =2	Year of teaching	Moral (ID 24= .21391)	Moral (ID 24= .19042)	Moral (ID 24= -.01690)
Estimate ρ *	.2097 (.05746)	.2198 (.04475)	.2274 (.03344)	.2033 (.06418)	.1284 (.1010)	.1281 (.1011)	.1232 (.1022)
% of Estimate ρ to the $(1/\lambda_{max})$	84.97%	89.05%	92.10%	82.36%	52.04%	51.89%	49.90%

* Figures in parentheses are standard deviations.

λ_{max} = 4.051115 and $1/\lambda_{max}$ = .2468456.

λ_{min} = -2.083131 and $1/\lambda_{min}$ = -.4800466.

How can we interpret this difference between W* and W regarding ρ_{moral}? We know that in W*, the different levels (0, 1, 2, 3, 4) of frequency of professional discussion are simplified to two levels (0, 1). When the percentage of ρ_{moral} from W and Y is higher than that from W* and Y, it demonstrates that the teachers' moral agency level is associated with how frequently they initiate a professional discussion. The relatively higher ρ_{moral} from W and Y compared to that from W* and Y indicates that the frequency of professional discussion is positively associated with teachers' moral agency level. When different levels of frequency are replaced by a simple "yes / no" choice, professional discussion is not that positively and highly associated with teachers' moral agency. This pleasant conclusion seems to be acceptable in a common sense.

As we have said in chapter 5 (§5.2), the parameter ρ plays a role as the correlationship coefficient between a vector Y and a matrix W, then different strategies of coding / re-coding either Y or W may lead to changes to the value of parameter ρ. Specifically to the matrices W and W*, one is having categorical entries from 0 to 4 and another is re-coded in a binary way, we have found a difference between these two matrices. There might be other differences if we re-code the matrix W in still other ways. But here we are not going to do further.

6.2.4. The cause – effect relationship between W and Y.

In Holland and Leinhardt (1981), the relationship described is Y as a function of W. But now we may emphasize that there might not be a direct cause-effect relationship between W and Y. We just say that W and Y are associated. That is, it is possible that Y is the cause of W. It is also possible that W becomes the cause of Y. For example, when

W is the association or pattern matrix among a group of subjects, and Y is their gender, then usually W should not be considered as a cause of Y since by no means the subjects' behavior can be the reason that they have their own gender. However, in a social behavior sense, it might be better to assign some subject(s) a "social gender" which is different from his / her biological one.

6.3. About the formulas for the standard deviation of ρ.

In Table5A and Table5B, I calculate the standard deviation of ρ using Ord's formula. There are different versions of formulas regarding the standard error term of parameters in SAM in the literature. I list some of them below.

Doreian, P. (1982) (page249) with 4 parameters.

Duke, J. B. (1993) (page 472) with 3 parameters.

Both of the above formulas can be simplified to a formula with 2 parameters as below.

Ord, K. (1975) (page 124) with 2 parameters.

Doreian, P. (1981) (page 367) with 2 parameters.

The details of Ord's formula can be seen in Doreian, P. (1981). Actually, all those formulas are essential the same. I notice the following facts as comments.

1. Those formulas are asymptotic;
2. The importance of the relation between W and Y was not emphasized in the formulas;
3. The fact that the boundary of ρ is not symmetrical to the origin when dealing with W(0) was not considered in the formulas;

4. Those formulas are functions of ρ and ρ is multiple-valued. It is not clear how to pick out the value(s) of ρ from a group of many for the functions.

Because of the reasons above, it might be problematic to apply Ord's formula to calculate the standard deviations of ρ in my work. So are the comparisons with the standard deviation of ρ. Since there is no other choices, Ord's formula is used, and further consideration is necessarily expected.

Chapter 7

DISCUSSION

In previous chapters, we have obtained a new understanding of the SAM model itself, and a new technique for parameter estimation in SAM model. Here I am going to compare the new understanding with some other multivariate analysis methods in statistics (§7.2 and §7.3.1 - §7.3.5) and plans for the further consideration.

7.1. The comparison of the new technique with the OLS technique.

In the following models, Y is the outcome, X is the predictor, W is the weight matrix, r is the correlation coefficient, ρ is the autocorrelation coefficient, and ε is the i.i.d. normal random error. I want to make comparison of the new Newton-Raphson (NR) technique of parameter ρ estimation in the SAM model with the ordinary least square (OLS) technique of the correlation coefficient parameter r estimation.

(a). $Y = rX + \varepsilon$

This is a general linear model with constant setting to 0 for convenience.

(b). $Y = \rho WY + \varepsilon$

This is a general autoregression model with Y appears in both sides.

(c). $Y = \rho(WY) + \varepsilon$

This is a general autoregression model but treated as a linear model by computing WY and using it as X, so that actually, Y appears at left side only.

Comments.

Model (a) is a typical regression model with no intercept.

Model (b). If Y is one of the eigenvector of W, then the error term ε happens to be zero. But practically, it's unlikely especially when the dimensional number of W is high.

Model (c) is an abuse of the spatial autoregression model $Y = \rho WY + \varepsilon$ by treating (WY) as an observation X and solving for the normal equation for ρ and σ^2 estimates. Calling this estimates of ρ and σ^2 as a "normal" solution, this "normal" solution is different from the solution we obtained in §3.2. We have to notice the fact that the "normal" σ^2 estimate is always smaller than that from §3.2. Can we say that this smaller estimate of the error term is better? Then why bother working on the spatial autoregression model?

The answer is simple: we must keep Y equal in both sides of the equity of the spatial autoregression model, whereas the "normal" solution violated this requirement although models (b) and (c) look like the same. Let's talk a little more below.

In model (a) above, we are trying to get the solution of $\hat{Y} = r\mathrm{X}$ where r is obtained by minimizing the error term ε. In model (b), we are trying to get the solution of $\hat{Y} = \rho\hat{Y}$ where ρ is obtained by optimizing the likelihood function from model (b). We notice that in model (b), $\hat{Y}$ appears in both sides of the equity. Now in model (c), when we abuse the model by writing WY = X, we are actually trying to get the solution of $\hat{Y} = \rho\mathrm{X}$, or

$\hat{Y} = \rho(WY)$ by minimizing the error term ε. We notice that in the left side of this expression of $\hat{Y} = \rho(WY)$, we have $\hat{Y}$, the predicted Y whereas in the right side of this expression, we have Y, the original observation. Seemingly, $\hat{Y}$ is not necessarily to be the same as Y is, so that the requirement of the model (b) in which the same Y must appear in both sides of the equity is violated.

7.2. The comparison of the new technique with factor analysis: the extended Factor Analysis Method.

Substantively, the new NR technique is pretty close to the classical factor analysis. I am going to talk a little more about the comparisons of the classical factor analysis with my new technique in SAM models.

The history of factor analysis technique can be traced back to the beginning of the century with the early work starting at one or two factor levels. In Harman (1960), the author said that "a principal objective of factor analysis is to attain a parsimonious description of observed data". The author also said that "while the goal of complete description cannot be reached theoretically, it may be approached practically in a limited field of investigation where a relatively small number of variables is considered exhaustive" (page 5). In factor analysis, we are going to transform the original data set into a new factor space, and use less factors to express the original data set via the so-called "factor extraction" step. We know that the main goal is to reduce the dimension number, but when doing so, the original data set will lose some information, while applying my new Newton-Raphson (NR) technique, we will lose less information.

Actually, in factor analysis, the variance-covariance matrix D is a specification of W(a) as I said in chapter 5. Here D is semi-positively definite so that all eigenvalues of D are non-negative with a total summation equal n, the dimension number. Some eigenvalues are big or much bigger than 1 whereas some are small, close to 0 or even equal 0 if the rank of matrix D is less than n, the dimension number. The step of "factor extraction" means that those factors whose corresponding eigenvalues are at the "bottom level", namely those which are close to be 0 or equal 0, should not be considered for the factor list. Only those factors whose corresponding eigenvalues are at the "top level", namely those which are the biggest ones, should be considered for the communality. We know that those "top level" factors are important and should be considered because of their large variation. But we can not say that "bottom level" ones are not important, and should not be considered in the data analysis. When we are dealing with the power balancing as we do in the SAM model, these negative factors (especially those factors corresponding to the negatively large eigenvalues) must be considered. That is, in SAM models, both those factors corresponding to the positively large and negatively large eigenvalues play the same important role. They are equally important.

In a broad sense, the technique of classical factor analysis can be applied for many types of matrices especially for those with principal diagonal elements equal based on the SAM model. This means that the matrices we are working with in factor analysis can be not only variance-covariance matrices, but also more general ones. We surely assume all elements on the principal diagonal equal, because a meaningful result can be obtained based on this assumption.

In the SAM model

$Y = \rho W(a)Y + \varepsilon$

where $\varepsilon \sim N(0, \sigma^2 I)$, if "a" equals zero, this is exactly what we were talking about in previous chapters. If "a" is not zero, then we may standardize the model as

$Y = (a\rho)(W/a)Y + \varepsilon.$

So that the new weight matrix will have all elements on the principal diagonal one while the parameter ρ does not substantially change.

Now, if all entries of the standardized weight matrix W/a are less than or equal to one in an absolute value meaning, this W/a is really a variance-covariance matrix. All its eigenvalues will be positive with some zeros possibly. But if some of the entries are bigger than one in an absolute value meaning, or in other words in case the value of "a" is decreasing to zero, then W will have some negative eigenvalues, and will have more if "a" continues decreasing. When "a" is negatively large enough, we will have all eigenvalues negative. The above statement can be verified easily applying the W(a) family properties 1-3 from §5.4.

We make arrangement for the whole process: from "a" equals to positive infinity, to positively large values, to positive 1, then less than 1 but still positive, to zero, and then negatively increasing until "a" reaches negative infinity. In whole this process, the corresponding eigenvalue sets are stable, with only a constant "a" adjustment. At the same time, the important fact is, their corresponding orthogonal matrix P and factors remain the same when W is non-singular.

Now we see, in a classical factor analysis work, we pick out the top level factors with large variation, and delete the bottom ones with the least variation. This is because "a" equals one. When "a" takes value zero, we know that both top level factors and bottom level factors must be considered because both of them carry large variation now, although the proportion of the variations are not equal as the positive and negative eigenvalues are not symmetrical in an absolute value meaning. It is now those factors whose eigenvalues are at the middle level, or close to zero, that should be removed because they carry little variation only. This process will continue to work when "a" moves in either direction. Say, when "a" is becoming negatively large, then those bottom eigenvalues and factors are becoming leading as they carry the major part of the variation. Those top ones should be out of consideration as the proportion of the variation they carry is now small. All we have found here is, the whole set of factors is in a dynamic process when the value "a" is varying from $+\infty$ to $-\infty$. Here I want to emphasize that the bottom ones need to be considered because they may have potential importance while the top ones always do. In general, top ones and bottom ones are more likely important comparing with those in the middle places.

A question might be asked: does it make sense when we are talking about the value of "a" being $+\infty$ or $-\infty$? The answer would be "yes". We see that in a network, if those subjects are purely objective, we have $a = 0$ and obtain the results before. When "a" is positively increasing, those subjects are becoming more self-confirmation centered, more subjective, and less influenced by others. When $a = +\infty$, this network becomes a small world of purely self-confirmation centered subjects. When "a" is negative, and negatively increasing, those subjects are becoming more self-negation centered, more likely influenced by others. When $a = -\infty$, this network becomes a small world of purely self-negation centered subjects. Both of the above positive / negative extreme examples might be found in some special societies.

7.3. The next stage work.

7.3.1. We may apply the Newton-Raphson technique for the autocorrelation coefficient parameter ρ for the multilinear regression. We know that Doreian (1981, 1982, 1989), Duke (1993), Leenders (1995) were working on the model:

$$Y = \rho WY + XB + \varepsilon$$

where ε is the random error term. This is a linear model combined with the autoregressive information. With the newly developed Newton-Raphson technique for the autocorrelation coefficient parameter ρ, we may apply this new technique in practice without technical difficulties.

7.3.2. The original SAM model is time-independent, but it can be viewed as a result from a long period of negotiation and balancing. That is, we may understand that the original model is the result from the time-dependent models such as

$$Y_{t+1} = \rho WY_t + \varepsilon$$

in which the observed Y is time-dependent but the information of interaction W is time-independent. When $t \to \infty$, we get the SAM model. More generally, the model can be

$$Y_{t+1} = \rho W_t Y_t + \varepsilon$$

in which the observed Y and the information of interaction W are both time-dependent. If W_t converges to a matrix W, we get back $Y = \rho\ WY + \varepsilon$ if Y also converges.

7.3.3. Cressie (1993) introduced the stationary processes in a plane prescribed by Whittle (1954). With a set of translation operators defined on a plane, and a translation function of operators based on the connection matrix W, Cressie shows that the SAM model is exactly a spatial analogue of an autoregressive time-series model (page406). It follows that we may directly apply the newly developed technique working on the translated model without difficulties. Similar approaches were discussed by Bartels (1979) and others.

7.3.4. We see that the above time-series models are more advanced, and have appeared in a wide range of application, namely autoregressive moving average (ARMA) models. It is worth to explore the possibility to apply my new technique of the parameter ρ estimation into the time-dependent model studies.

7.4. Conclusion.

There is some limitation to the application of the new technique of the parameter ρ estimation in SAM models. So far, we can only work on those matrices W(a) with the diagonal elements identical. When dealing with W(0), the summation of W(0)'s all eigenvalues $\{\lambda_i\}$ is zero so that we have $\lambda_{min} < 0$ and $\lambda_{max} > 0$. However, we have no technique to have the maximum and minimum eigenvalues equal in an absolute value sense. Consequently, the corresponding boundary $(\lambda_{min}, \lambda_{max})$ is not symmetrical to the origin, and we can not standardize the boundary to be [-1, 1] as we do for the correlation coefficient *r* between two vectors. This would cause the interpretation less intuitive, and a little computationally complicated.

We now have the newly developed technique to estimate the parameters ρ and σ^2 in SAM models. We also have the definition and interpretation of parameter ρ estimate.

SAM models represent the relationship between a vector Y and a weight matrix W. When there is association among a group of subjects, we always have chances to explore the relationship between the pattern of the mutual motion and observations obtained from the group of subjects.

APPENDICES

Appendix3A

Example 1

$$W=\begin{pmatrix}0 & 1 & 3\\ 1 & 0 & 4\\ 3 & 4 & 0\end{pmatrix},\lambda=\begin{pmatrix}\lambda_1\\ \lambda_2\\ \lambda_3\end{pmatrix}=\begin{pmatrix}5.50963\\ -4.55287\\ -0.956762\end{pmatrix},Y_1=\begin{pmatrix}1\\ 1\\ 1\end{pmatrix},Y_2=\begin{pmatrix}1\\ 0\\ -1\end{pmatrix}.$$

$$g(\rho)=\left(\frac{2}{n}\right)\sum_{i=1}^{n}\frac{\lambda_i}{(1-\rho\lambda_i)}$$

$$=\left(\frac{2}{3}\right)\cdot\left[\frac{5.50963}{1-5.50963\rho}+\frac{-4.455287}{1+4.455287\rho}+\frac{-0.956762}{1+0.956762\rho}\right].$$

$$h_i(\rho)=2\cdot\frac{\rho(Y_{iL})'Y_{iL}-Y_i'Y_{iL}}{Y_i'Y_i-2\rho Y_i'Y_{iL}+\rho^2(Y_{iL})'Y_{iL}},(i=1,2)\ \text{so that}$$

$$h_1(\rho)=2\cdot\frac{\rho(Y_{1L})'Y_{1L}-Y_1'Y_{1L}}{Y_1'Y_1-2\rho Y_1'Y_{1L}+\rho^2(Y_{1L})'Y_{1L}}=\frac{2(90\rho-16)}{(3-32\rho+\rho^2)}\ \text{and}$$

$$h_2(\rho)=2\cdot\frac{\rho(Y_{2L})'Y_{2L}-Y_2'Y_{2L}}{Y_2'Y_2-2\rho Y_2'Y_{2L}+\rho^2(Y_{2L})'Y_{2L}}=\frac{2\cdot(27\rho+6)}{(2+12\rho+27\rho^2)}.$$

Solve for $f_i'(\rho)=g(\rho)+h_i(\rho)=0$ (i = 1,2), we get

$\hat{\rho}_1\approx .152$ and

$\hat{\rho}_2\approx -.12$.

Appendix3B

The calculation of the conjugate eigenvalues

We need only check $f'(\rho)$ and $f''(\rho)$ with one pair of conjugate complex eigenvalues here. Assume that

$\lambda_{1,2} = u \pm iv$ with $v \neq 0$, then for any real ρ,

$$\sum_{i=1}^{2} \frac{\lambda_i}{1-\rho\lambda_i} = \frac{u+iv}{1-\rho u - i\rho v} + \frac{u-iv}{1-\rho u + i\rho v} = \frac{2[u(1-\rho u)-\rho v^2]}{(1-\rho u)^2+(\rho v)^2} \text{ and}$$

$$\sum_{i=1}^{2} \frac{\lambda_i^{\,2}}{(1-\rho\lambda_i)^2} = \frac{(u+iv)^2}{(1-\rho u - i\rho v)^2} + \frac{(u-iv)^2}{(1-\rho u + i\rho v)^2}$$

$$= \frac{2\{(u^2-v^2)[(1-\rho u)^2-(\rho v)^2]-4\rho u v^2(1-\rho u)\}}{[(1-\rho u)^2+(\rho v)^2]^2}.$$

Both terms involving the eigenvalues remain real, and so do $f'(\rho)$ and $f''(\rho)$ as well, since their remaining components are real.

Appendix6A

The raw data matrix W(24×24)

```
N            Group And Actor ID
24       |AAAA|BBBBBB|CCCCCCCC|DDDDDD|
         |    |      |        |      |
         |    |      |        |      |
         |    |22 21 | 111 11 |1*1121|
Group  ID|7129|326078|55043694|8*1312|
---------+----+------+--------+------+
 1 A    7|A213|......|........|...1..|
 1 A    1|4A3.|......|.4......|......|
 1 A    2|33A.|......|........|......|
 1 A    9|433A|......|........|......|
---------+----+------+--------+------+
 2 B   23|.2..|B443..|........|......|
 2 B   22|.1..|4B....|...4....|....2.|
 2 B    6|....|4.B...|........|......|
 2 B   20|....|33.B..|........|...1..|
 2 B   17|..3.|3..3B.|........|.3..2.|
 2 B    8|....|....1B|........|......|
---------+----+------+--------+------+
 3 C    5|....|......|C...3.33|.3....|
 3 C   15|.4..|..4...|.C.4..4.|4.....|
 3 C   10|....|......|33C.4.3.|..4...|
 3 C   14|.4..|.4....|444C....|......|
 3 C    3|3...|.4....|4.44C...|......|
 3 C   16|.1..|.....4|3.2.3C..|......|
 3 C   19|....|......|444..4C4|......|
 3 C    4|....|......|3..3.44C|......|
---------+----+------+--------+------+
 4 D   18|.1..|......|.1......|D.1...|
 4 D   **|....|......|4.3.....|3D4...|
 4 D   11|....|......|4..4...4|44D...|
 4 D   13|..3.|...1..|........|...D3.|
 4 D   21|....|.3....|........|.343D.|
 4 D   12|.1..|.1....|........|.3..3D|
```

The value in each cell represents the extent to which teacher in the row indicated engaging in professional discussions with the teacher in the column (a value of 4 is almost daily and a value of 1 is less than once a month).

Appendix6B

The comparison of the results from different initial value selections

Initial value of ρ	% of the estimate of ρ being within $(1/\lambda_{min}, 1/\lambda_{max})$	% of improvement from last step
Ord's $\rho^{\circ} = Y'Y_L / Y'Y$	23	
$\rho^{\circ}_1 = 0.0$	59	36
$\rho^{\circ}_2 = Y'Y_L / (Y_L)'Y_L$	89	30
$\rho^{\circ}_3 =$ .75min or .75max *	96	7
$\rho^{\circ}_4 =$ .90min or .90max *	100	4

* min = $\min(\rho^{\nabla}, 1/\lambda_{max})$ if $Y'Y_L > 0$ and

max = $\max(\rho^{\nabla}, 1/\lambda_{min})$ if $Y'Y_L < 0$.

Appendix6C

A SAS program (1)

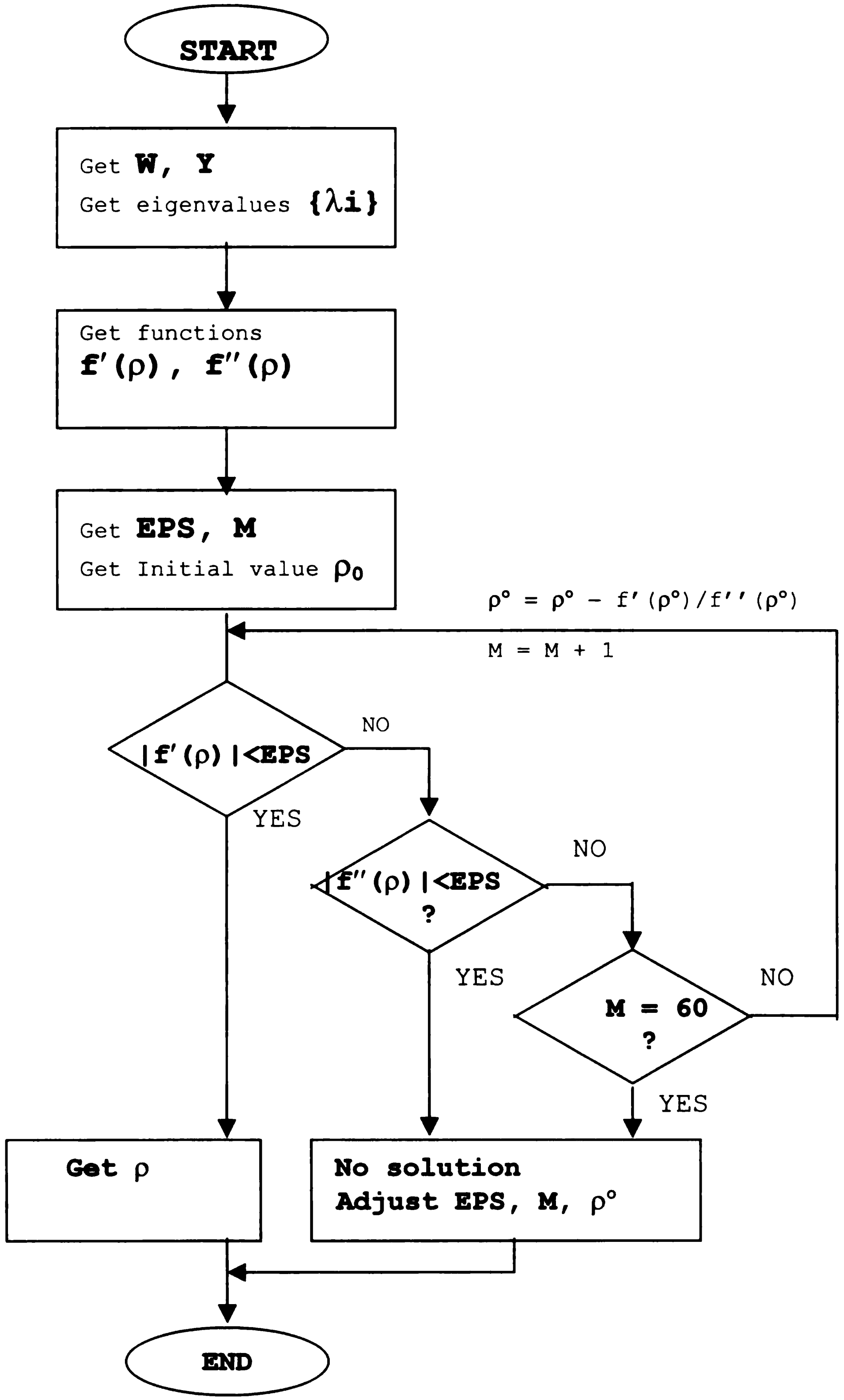
```
*All estimates will be within the boundary, or treated as missing.
options nocenter;
proc iml;

n=8;
nr=30;
nk=9;

w={0 2 1 3 0 0 0 0,
   4 0 3 0 0 0 0 0,
   3 3 0 0 0 0 0 0,
   4 3 3 0 0 0 0 0,
   0 2 0 0 0 4 4 3,
   0 1 0 0 4 0 0 0,
   0 0 0 0 4 0 0 0,
   0 0 0 0 3 3 0 0};

I8=I(8);

*print w;
*print I8;

zr={6.87348 6.80376 0 -.895568 -1.72068 -1.72068 -3.43211 -5.9082}`;
*print zr;
zc={0 0 0 0 .687622 -.687622 0 0}`;
*print zc;

BR=.145487;
*print BR;
BL=-.1692563;
*print BL;

rr0={-.15233 -.11848 -.084628 -.050777 0 .043646 .072743 .10184
.130938};
print rr0;
r0=j(1,1,0);

ee=j(n,nr,0);
e0=j(n,1,0);
seed=3837;
do i=1 to n;
   do j=1 to nr;
      ee(|i,j|)=0+1*normal(seed);
      *if abs(ee(|i,j|))>1 then ee(|i,j|)=.;
   end;
end;
*print ee;

lamr=j(n,1,0);
lamc=j(n,1,0);
```

```
y=j(n,nr,0);    *y=0;

a=j(n,nk,0);
b=j(n,nk,0);
c=j(n,nk,0);

a=1-zr*rr0;
*print a;
b=zc*rr0;
*print b;
c=a#a+b#b;
*print c;

dcc=j(1,nr,0);
ecc=j(1,nr,0);
gcc=j(1,nr,0);
yyc=j(n,1,0);

Rho=j(nr,nk,0);

ah=j(n,1,0);
bh=j(n,1,0);
h1=j(n,1,0);
h0=0;
j0=0;
fh=0;
fj=0;

XX=0;
af=j(n,1,0);
bf=j(n,1,0);
f1=j(n,1,0);

aaf=j(n,1,0);
bbf=j(n,1,0);
ff1=j(n,1,0);

XX=0;
start GetRho;
   m=0;
   eps=1.0E-06;
   if abs(gc)<eps then XX=.;

   repeat:
   m=m+1;
   if abs(X0)<eps then XX=X0;
   if abs(X0)>eps & abs(gc)>eps then do;
              af=(1-X0#zr)#zr-X0#zc#zc;
              bf=(1-X0#zr)##2+(X0#zc)##2;
              f1=af/bf;
              s2=dc-2*X0#ec+X0#X0#gc;
              ff=2*f1[+]/n+2*(X0#gc-ec)/s2;
           if abs(ff)<eps then XX=X0;
           if abs(ff)>eps then do;
                 aaf=((1-X0#zr)#zr-X0#zc#zc)##2+zc##2;
                 bbf=((1-X0#zr)##2+(X0#zc)##2)##2;
                 ff1=aaf/bbf;
```

```
               s2=dc-2*X0#ec+X0#X0#gc;
               sff=2*ff1[+]/n+2*gc/s2-4*(X0*gc-ec)/s2/s2;
            if abs(sff)<eps then XX=.;
            if m=60 then XX=.;

            if abs(sff)>eps & m<60 then X1=X0-ff/sff;
               X0=X1;
               goto repeat;
        end;
     end;
finish GetRho;

**********************************************************************
ak=j(n,1,0);
bk=j(n,1,0);
ck=j(n,1,0);

II=j(n,n,0);
WI=j(n,n,0);
IW=j(n,n,0);

do k=1 to nk;

   r0=rr0[k];
   ak=a[,k];
   bk=b[,k];
   ck=c[,k];

   IW=inv(I8-W*r0);
*print IW;

do j=1 to nr;

   e0=ee[,j];
*print e0;
   dc=j(1,1,0);
   ec=j(1,1,0);
   gc=j(1,1,0);

   yyc=IW*e0;
*print yyc;

   dcc[j]=yyc`*yyc;
   ecc[j]=yyc`*(w*yyc);
   gcc[j]=yyc`*(w`*w*yyc);


   dc=dcc[j];
   *print dc;
   ec=ecc[j];
   *print ec;
   gc=gcc[j];
   *print gc;

      X0=.9995*ec/gc;
      run GetRho;
      Rho(|j,k|)=XX;
```

```
      if Rho(|j,k|)<BL then Rho(|j,k|)=.;
      if Rho(|j,k|)>BR then Rho(|j,k|)=.;
    end;
end;
print Rho;

varnames={rho1 rho2 rho3 rho4 rho5 rho6 rho7 rho8 rho9};

create norms from Rho (|colname=varnames|);
append from Rho;
close norms;
quit;

proc means data=norms; var _all_;
output out=normdata
mean=mrho1 mrho2 mrho3 mrho4 mrho5 mrho6 mrho7 mrho8 mrho9
var =vrho1 vrho2 vrho3 vrho4 vrho5 vrho6 vrho7 vrho8 vrho9;

proc print data=normdata; var _all_;
   title'Empirical Distribution of Rho estimation';
proc univariate plot data=norms; var _all_;

run;
```

Appendix6D - A Flow Chart



Appendix6E

A SAS program (2)

```
*All estimates will be within the boundary, or treated as missing.
options nocenter;
proc iml;

n=8;
nr=30;
nk=9;

w={0 2 1 3 0 0 0 0,
   4 0 3 0 0 0 0 0,
   3 3 0 0 0 0 0 0,
   4 3 3 0 0 0 0 0,
   0 2 0 0 0 4 4 3,
   0 1 0 0 4 0 0 0,
   0 0 0 0 4 0 0 0,
   0 0 0 0 3 3 0 0};

I8=I(8);

*print w;
*print I8;

zr={6.87348 6.80376 0 -.895568 -1.72068 -1.72068 -3.43211 -5.9082}`;
*print zr;
zc={0 0 0 0 .687622 -.687622 0 0}`;
*print zc;

BR=.145487;
*print BR;
BL=-.1692563;
*print BL;

rr0={-.15233 -.11848 -.084628 -.050777 0 .043646 .072743 .10184
.130938};
print rr0;
r0=j(1,1,0);

ee=j(n,nr,0);
e0=j(n,1,0);
seed=3837;
do i=1 to n;
   do j=1 to nr;
      ee(|i,j|)=0+1*normal(seed);
      *if abs(ee(|i,j|))>1 then ee(|i,j|)=.;
   end;
end;
*print ee;

lamr=j(n,1,0);
lamc=j(n,1,0);
```

```
y=j(n,nr,0);    *y=0;

a=j(n,nk,0);
b=j(n,nk,0);
c=j(n,nk,0);

a=1-zr*rr0;
*print a;
b=zc*rr0;
*print b;
c=a#a+b#b;
*print c;

dcc=j(1,nr,0);
ecc=j(1,nr,0);
gcc=j(1,nr,0);
yyc=j(n,1,0);

Rho=j(nr,nk,0);

ah=j(n,1,0);
bh=j(n,1,0);
h1=j(n,1,0);
h0=0;
j0=0;
fh=0;
fj=0;

XX=0;
af=j(n,1,0);
bf=j(n,1,0);
f1=j(n,1,0);

aaf=j(n,1,0);
bbf=j(n,1,0);
ff1=j(n,1,0);

XX=0;
start GetRho;
   m=0;
   eps=1.0E-06;
   if abs(gc)<eps then XX=.;

   repeat:
   m=m+1;
   if abs(X0)<eps then XX=X0;
   if abs(X0)>eps & abs(gc)>eps then do;
              af=(1-X0#zr)#zr-X0#zc#zc;
              bf=(1-X0#zr)##2+(X0#zc)##2;
              f1=af/bf;
              s2=dc-2*X0#ec+X0#X0#gc;
              ff=2*f1[+]/n+2*(X0#gc-ec)/s2;
           if abs(ff)<eps then XX=X0;
           if abs(ff)>eps then do;
                  aaf=((1-X0#zr)#zr-X0#zc#zc)##2+zc##2;
                  bbf=((1-X0#zr)##2+(X0#zc)##2)##2;
                  ff1=aaf/bbf;
```

```
               s2=dc-2*X0#ec+X0#X0#gc;
               sff=2*ff1[+]/n+2*gc/s2-4*(X0*gc-ec)/s2/s2;
            if abs(sff)<eps then XX=.;
            if m=60 then XX=.;

            if abs(sff)>eps & m<60 then X1=X0-ff/sff;
               X0=X1;
               goto repeat;
        end;
     end;
finish GetRho;

**********************************************************************
ak=j(n,1,0);
bk=j(n,1,0);
ck=j(n,1,0);

II=j(n,n,0);
WI=j(n,n,0);
IW=j(n,n,0);

do k=1 to nk;

   r0=rr0[k];
   ak=a[,k];
   bk=b[,k];
   ck=c[,k];

   IW=inv(I8-W*r0);
*print IW;

do j=1 to nr;

   e0=ee[,j];
*print e0;
   dc=j(1,1,0);
   ec=j(1,1,0);
   gc=j(1,1,0);

   yyc=IW*e0;
*print yyc;

   dcc[j]=yyc`*yyc;
   ecc[j]=yyc`*(w*yyc);
   gcc[j]=yyc`*(w`*w*yyc);


   dc=dcc[j];
   *print dc;
   ec=ecc[j];
   *print ec;
   gc=gcc[j];
   *print gc;

      X0=.9995*ec/gc;
      run GetRho;
      Rho(|j,k|)=XX;
```

```
      if Rho(|j,k|)<BL then Rho(|j,k|)=.;
*     if Rho(|j,k|)>BR then Rho(|j,k|)=.;
    end;
end;
print Rho;

varnames={rho1 rho2 rho3 rho4 rho5 rho6 rho7 rho8 rho9};

create norms from Rho (|colname=varnames|);
append from Rho;
close norms;
quit;

proc means data=norms; var _all_;
output out=normdata
mean=mrho1 mrho2 mrho3 mrho4 mrho5 mrho6 mrho7 mrho8 mrho9
var =vrho1 vrho2 vrho3 vrho4 vrho5 vrho6 vrho7 vrho8 vrho9;

proc print data=normdata; var _all_;
   title'Empirical Distribution of Rho estimation';
proc univariate plot data=norms; var _all_;

run;
```

Appendix6F

A SAS output from program (1)

```
        RR0
   -0.15233  -0.11848 -0.084628 -0.050777         0  0.043646
:  0.072743   0.10184  0.130938



        RHO
  -0.093043 -0.035922 -0.004561 0.0225311 0.0578763         .
:         .         . 0.1328653

  -0.152541 -0.112992 -0.070801 -0.030338 0.0239434 0.0649262
: 0.0899643 0.1133124 0.1342584

   -0.14385         .         . -0.088461 -0.064811 -0.039001
: -0.012986 0.0306876 0.1094544

          .         .         .         .         . -0.067272
: -0.025934 0.0351362 0.1132646

  -0.157897 -0.138627 -0.118766 -0.093317 -0.039522 0.0184154
: 0.0587202 0.0975185 0.1322099

          .         .         .         .         . -0.044463
: -0.004264 0.0516837 0.1204356

  -0.144655 -0.103632 -0.076379 -0.057642 -0.031859 -0.001434
: 0.0268882 0.0627848 0.1035768

  -0.029778 0.0225825         .         .         .         .
:         .         . 0.1330539

  -0.150079 -0.102038 -0.053859 -0.014943 0.0284294 0.0593437
: 0.0797859         . 0.1254203

  -0.151816         .         . -0.101576 -0.077795 -0.050537
:  -0.02298 0.0207714 0.1023134

  -0.076481 -0.006553   0.01267 0.0262203 0.0476586 0.0704985
: 0.0888963 0.1101619 0.1335608

          .         . -0.088569  -0.05954 -0.007212 0.0415269
: 0.0738519  0.104953 0.1336305

          .         .  -0.11541 -0.083249 -0.024229 0.0301886
: 0.0653107 0.0981717         .

          . -0.130321  -0.10168 -0.069988 -0.017111 0.0317415
: 0.0646765 0.0969821         .

  -0.153238 -0.118188 -0.077435 -0.034343 0.0264485  0.070925
: 0.0961682 0.1181363 0.1371757
```

```
         . -0.091266 -0.049254 -0.016049  0.026863 0.0621555
: 0.0858966 0.1099969 0.1342313

         .         . -0.094694 -0.069018  -0.02249 0.0246738
: 0.0592215 0.0954299 0.1310436

  -0.127586  -0.05403 -0.012775 0.0168752 0.0534261 0.0815975
: 0.0996438 0.1173267 0.1347104

  -0.145711 -0.089183 -0.034598 0.0095697 0.0586682  0.090647
: 0.1087552 0.1248541 0.1382911

  -0.067926  -0.00001 0.0251898 0.0448484 0.0720933 0.0947013
: 0.1094199 0.1237517 0.1374021

  -0.154478 -0.113364 -0.062899  -0.01334 0.0479039 0.0870594
: 0.1075293 0.1242923 0.1375327

         . -0.090035  -0.05806 -0.036156 -0.013602 0.0024012
: 0.0158846 0.0406473         .

         .         .         . -0.096187 -0.041528 0.0142268
: 0.0531585 0.0919801 0.1294309

  -0.055392 -0.039437 -0.024876 -0.007934 0.0215263 0.0516421
: 0.0749359 0.1013587 0.1309394

  -0.157611 -0.136778  -0.11481 -0.088274 -0.037288 0.0155642
: 0.0534711 0.0920938 0.1296356

  -0.084654 -0.029044 -0.007296 0.0137483 0.0466771 0.0758931
:  0.095576  0.115246         .

         .         .         .         . -0.053363 -0.016234
: 0.0148812 0.0523493 0.0940635

  -0.152127 -0.108562 -0.060382 -0.014673 0.0431066 0.0821144
:  0.103535 0.1217698  0.136639

  -0.161596 -0.145931 -0.127747 -0.104224 -0.053415 0.0053668
: 0.0485434 0.0912364 0.1303182

  -0.140211 -0.067346 -0.006049 0.0374953 0.0805432  0.105402
: 0.1181952 0.1285706 0.1371136
```

Variable	N	Mean	Std Dev	Minimum
RHO1	20	-0.1250335	0.0407659	-0.1615960
RHO2	21	-0.0805083	0.0493466	-0.1459314
RHO3	23	-0.0575235	0.0447200	-0.1277471
RHO4	26	-0.0349216	0.0473422	-0.1042238
RHO5	27	0.0055903	0.0462534	-0.0777951
RHO6	28	0.0343597	0.0479813	-0.0672717
RHO7	28	0.0616695	0.0424770	-0.0259340

RHO8	27	0.0915260	0.0324845	0.0207714
RHO9	26	0.1274066	0.0123797	0.0940635

Variable	Maximum
RHO1	-0.0297784
RHO2	0.0225825
RHO3	0.0251898
RHO4	0.0448484
RHO5	0.0805432
RHO6	0.1054020
RHO7	0.1181952
RHO8	0.1285706
RHO9	0.1382911

OBS	_TYPE_	_FREQ_	MRHO1	MRHO2	MRHO3	MRHO4
1	0	30	-0.12503	-0.080508	-0.057524	-0.034922

OBS	MRHO5	MRHO6	MRHO7	MRHO8	MRHO9
1	.0055903	0.034360	0.061669	0.091526	0.12741

OBS	VRHO1	VRHO2	VRHO3	VRHO4	VRHO5
1	.0016619	.0024351	.0019999	.0022413	.0021394

OBS	VRHO6	VRHO7	VRHO8	VRHO9
1	.0023022	.0018043	.0010552	.00015326

Univariate Procedure

Variable=RHO1

Moments

N	20	Sum Wgts	20
Mean	-0.12503	Sum	-2.50067
Std Dev	0.040766	Variance	0.001662
Skewness	1.149984	Kurtosis	-0.04093
USS	0.344243	CSS	0.031575
CV	-32.604	Std Mean	0.009116
T:Mean=0	-13.7165	Pr>|T|	0.0001
Num ^= 0	20	Num > 0	0
M(Sign)	-10	Pr>=|M|	0.0001
Sgn Rank	-105	Pr>=|S|	0.0001

Quantiles(Def=5)

100% Max	-0.02978	99%	-0.02978
75% Q3	-0.08885	95%	-0.04259

```
  50% Med  -0.14518        90%  -0.06166
  25% Q1   -0.15289        10%  -0.15775
   0% Min   -0.1616         5%  -0.15975
                            1%   -0.1616
 Range      0.131818
 Q3-Q1      0.064041
 Mode        -0.1616


               Extremes

    Lowest    Obs      Highest    Obs
   -0.1616(      29) -0.08465(      26)
   -0.1579(       5) -0.07648(      11)
  -0.15761(      25) -0.06793(      20)
  -0.15448(      21) -0.05539(      24)
  -0.15324(      15) -0.02978(       8)


Missing Value          .
Count                 10
% Count/Nobs       33.33


   Stem Leaf                     #  Boxplot
     -2 0                        1     |
     -4 5                        1     |
     -6 68                       2     |
     -8 35                       2  +-----+
    -10                             |     |
    -12 8                        1  |  +  |
    -14 884332206540            12  *-----*
    -16 2                        1     |
        ----+----+----+----+
    Multiply Stem.Leaf by 10**-2
```

Univariate Procedure

Variable=RHO1

```
                     Normal Probability Plot
   -0.03+                                             * +++++
        |                                        * +++++
        |                                   *  *++++
        |                                ** ++++
        |                            +++++
        |                       +++++   *
        |           *   *  *+**+**** ***
   -0.17+      *      +++++
         +----+----+----+----+----+----+----+----+----+----+
             -2        -1         0        +1        +2
```

Univariate Procedure

```
Variable=RHO2

                    Moments

 N                 21  Sum Wgts         21
 Mean        -0.08051  Sum        -1.69068
 Std Dev     0.049347  Variance   0.002435
 Skewness    0.638473  Kurtosis   -0.65491
 USS         0.184815  CSS        0.048702
 CV          -61.2938  Std Mean   0.010768
 T:Mean=0    -7.47641  Pr>|T|       0.0001
 Num ^= 0          21  Num > 0           1
 M(Sign)         -9.5  Pr>=|M|      0.0001
 Sgn Rank      -112.5  Pr>=|S|      0.0001


               Quantiles(Def=5)

 100% Max  0.022582       99%  0.022582
  75% Q3   -0.03944       95%  -0.00001
  50% Med  -0.09127       90%  -0.00655
  25% Q1   -0.11336       10%  -0.13678
   0% Min  -0.14593        5%  -0.13863
                           1%  -0.14593
 Range     0.168514
 Q3-Q1     0.073928
 Mode      -0.14593


                    Extremes

    Lowest    Obs     Highest    Obs
  -0.14593(    29) -0.03592(       1)
  -0.13863(     5) -0.02904(      26)
  -0.13678(    25) -0.00655(      11)
  -0.13032(    14) -0.00001(      20)
  -0.11819(    15) 0.022582(       8)


Missing Value        .
Count                9
% Count/Nobs     30.00


   Stem Leaf                      #  Boxplot
      2 3                         1     |
      0                                 |
     -0 70                        2     |
     -2 969                       3  +-----+
     -4 4                         1  |     |
     -6 7                         1  |     |
     -8 109                       3  *--+--*
    -10 833942                    6  +-----+
    -12 970                       3     |
    -14 6                         1     |
        ----+----+----+----+
```

```
    Multiply Stem.Leaf by 10**-2
```

Univariate Procedure

Variable=RHO2

```
                        Normal Probability Plot
    0.03+                                                  *  ++++
        |                                                ++++
        |                                            *+*++
   -0.03+                                       *  *+*++
        |                                      *+++
        |                                  +++*
   -0.09+                              +++***
        |                          *+***  **
        |                     *  *+*+
   -0.15+                *   ++++
         +----+----+----+----+----+----+----+----+----+----+
             -2        -1         0        +1        +2
```

Univariate Procedure

Variable=RHO3

Moments

N	23	Sum Wgts	23
Mean	-0.05752	Sum	-1.32304
Std Dev	0.04472	Variance	0.002
Skewness	0.177019	Kurtosis	-0.99104
USS	0.120103	CSS	0.043997
CV	-77.7421	Std Mean	0.009325
T:Mean=0	-6.1689	Pr>|T|	0.0001
Num ^= 0	23	Num > 0	2
M(Sign)	-9.5	Pr>=|M|	0.0001
Sgn Rank	-127	Pr>=|S|	0.0001

Quantiles(Def=5)

100% Max	0.02519	99%	0.02519
75% Q3	-0.01277	95%	0.01267
50% Med	-0.06038	90%	-0.00456
25% Q1	-0.09469	10%	-0.11541
0% Min	-0.12775	5%	-0.11877
		1%	-0.12775

Range	0.152937
Q3-Q1	0.08192
Mode	-0.12775

Extremes

Lowest	Obs	Highest	Obs

```
  -0.12775(      29)  -0.0073(      26)
  -0.11877(       5) -0.00605(      30)
  -0.11541(      13) -0.00456(       1)
  -0.11481(      25)  0.01267(      11)
  -0.10168(      14)  0.02519(      20)


Missing Value          .
Count                  7
% Count/Nobs       23.33


   Stem Leaf                       #  Boxplot
      2 5                          1     |
      0 3                          1     |
     -0 3765                       4  +-----+
     -2 55                         2  |     |
     -4 849                        3  |  +  |
     -6 76130                      5  *-----*
     -8 59                         2  +-----+
    -10 9552                       4     |
    -12 8                          1     |
        ----+----+----+----+
    Multiply Stem.Leaf by 10**-2
```

Univariate Procedure

Variable=RHO3

```
                       Normal Probability Plot
    0.03+                                                  +*++
        |                                             ++*++
        |                                       **+*+*
        |                                    *+*++
   -0.05+                                +***
        |                           +*****
        |                       ++**
        |                 *  ++*+*  *
   -0.13+             *+++++
         +----+----+----+----+----+----+----+----+----+----+
             -2        -1         0        +1        +2
```

Univariate Procedure

Variable=RHO4

Moments

N	26	Sum Wgts	26
Mean	-0.03492	Sum	-0.90796
Std Dev	0.047342	Variance	0.002241
Skewness	0.046978	Kurtosis	-1.32498
USS	0.08774	CSS	0.056032

```
CV           -135.567  Std Mean    0.009285
T:Mean=0     -3.76126  Pr>|T|        0.0009
Num ^= 0           26  Num > 0            7
M(Sign)            -6  Pr>=|M|       0.0290
Sgn Rank       -113.5  Pr>=|S|       0.0021


           Quantiles(Def=5)

100% Max  0.044848        99%  0.044848
 75% Q3    0.00957        95%  0.037495
 50% Med  -0.03234        90%   0.02622
 25% Q1   -0.08325        10%  -0.09619
  0% Min  -0.10422         5%  -0.10158
                           1%  -0.10422
Range     0.149072
Q3-Q1     0.092819
Mode      -0.10422


                 Extremes

   Lowest     Obs     Highest     Obs
 -0.10422(      29) 0.016875(      18)
 -0.10158(      10) 0.022531(       1)
 -0.09619(      23)  0.02622(      11)
 -0.09332(       5) 0.037495(      30)
 -0.08846(       3) 0.044848(      20)


Missing Value          .
Count                  4
% Count/Nobs       13.33


   Stem Leaf                      #  Boxplot
      4 5                         1     |
      2 367                       3     |
      0 047                       3  +-----+
     -0 65538                     5  |     |
     -2 640                       3  *--+--*
     -4 8                         1  |     |
     -6 090                       3  |     |
     -8 63883                     5  +-----+
    -10 42                        2     |
        ----+----+----+----+
    Multiply Stem.Leaf by 10**-2
```

Univariate Procedure

Variable=RHO4

```
                       Normal Probability Plot
    0.05+                                         ++++*
        |                                    *+*+*
```

```
          |                                                 ***++
          |                                           *****+*+
   -0.03+                                       **+++
          |                                   ++**
          |                               +++*  *
          |                           *+*+***
   -0.11+      *   ++*+
            +----+----+----+----+----+----+----+----+----+----+
                 -2        -1         0         +1        +2
```

Univariate Procedure

Variable=RHO5

Moments

N	27	Sum Wgts	27
Mean	0.00559	Sum	0.150939
Std Dev	0.046253	Variance	0.002139
Skewness	-0.12845	Kurtosis	-1.29253
USS	0.056468	CSS	0.055624
CV	827.3815	Std Mean	0.008901
T:Mean=0	0.628024	Pr>|T|	0.5355
Num ^= 0	27	Num > 0	14
M(Sign)	0.5	Pr>=|M|	1.0000
Sgn Rank	32	Pr>=|S|	0.4524

Quantiles(Def=5)

100% Max	0.080543	99%	0.080543
75% Q3	0.047659	95%	0.072093
50% Med	0.021526	90%	0.058668
25% Q1	-0.03729	10%	-0.05342
0% Min	-0.0778	5%	-0.06481
		1%	-0.0778
Range	0.158338		
Q3-Q1	0.084947		
Mode	-0.0778		

Extremes

Lowest	Obs	Highest	Obs
-0.0778(	10)	0.053426(	18)
-0.06481(	3)	0.057876(	1)
-0.05342(	29)	0.058668(	19)
-0.05336(	27)	0.072093(	20)
-0.04153(	23)	0.080543(	30)

Missing Value	.
Count	3
% Count/Nobs	10.00

```
 Stem Leaf                     #  Boxplot
    8 1                        1     |
    6 2                        1     |
    4 3788389                  7  +-----+
    2 24678                    5  *-----*
    0                             |  +  |
   -0 747                      3  |     |
   -2 7242                     4  +-----+
   -4 3320                     4     |
   -6 85                       2     |
      ----+----+----+----+
  Multiply Stem.Leaf by 10**-2
```

Univariate Procedure

Variable=RHO5

```
                        Normal Probability Plot
  0.09+                                                   +++*
      |                                              ++++*
      |                                        ****+** *
      |                                   *****+++
  0.01+                                  +++++
      |                              ++***
      |                          +*****
      |                     +*+**
 -0.07+     *  ++*+
       +----+----+----+----+----+----+----+----+----+----+
           -2        -1         0        +1        +2
```

Univariate Procedure

Variable=RHO6

Moments

N	28	Sum Wgts	28
Mean	0.03436	Sum	0.962071
Std Dev	0.047981	Variance	0.002302
Skewness	-0.52076	Kurtosis	-0.64462
USS	0.095216	CSS	0.06216
CV	139.6444	Std Mean	0.009068
T:Mean=0	3.78927	Pr>|T|	0.0008
Num ^= 0	28	Num > 0	22
M(Sign)	8	Pr>=|M|	0.0037
Sgn Rank	139	Pr>=|S|	0.0006

Quantiles(Def=5)

100% Max	0.105402	99%	0.105402
75% Q3	0.073409	95%	0.094701

```
 50% Med   0.036634        90%  0.090647
 25% Q1    0.003884        10%  -0.04446
  0% Min   -0.06727         5%  -0.05054
                            1%  -0.06727

Range      0.172674
Q3-Q1      0.069525
Mode       -0.06727
```

```
                    Extremes

     Lowest     Obs      Highest     Obs
   -0.06727(      4) 0.082114(       28)
   -0.05054(     10) 0.087059(       21)
   -0.04446(      6) 0.090647(       19)
     -0.039(      3) 0.094701(       20)
   -0.01623(     27) 0.105402(       30)
```

```
Missing Value          .
Count                  2
% Count/Nobs        6.67
```

```
   Stem Leaf                    #  Boxplot
     10 5                       1     |
      8 22715                   5     |
      6 25016                   5  +-----+
      4 229                     3  |     |
      2 502                     3  *--+--*
      0 25468                   5  +-----+
     -0 61                      2     |
     -2 9                       1     |
     -4 14                      2     |
     -6 7                       1     |
        ----+----+----+----+
    Multiply Stem.Leaf by 10**-2
```

Univariate Procedure

Variable=RHO6

```
                          Normal Probability Plot
    0.11+                                       ++++   *
        |                                 * **+*   *
        |                             *****++
    0.05+                          ***+++
        |                        ***+
        |                   *****
   -0.01+               +*+*
        |           ++++*
        |       +++*   *
   -0.07+   +++*
         +----+----+----+----+----+----+----+----+----+----+
              -2        -1         0        +1        +2
```

```
Univariate Procedure

Variable=RHO7

                    Moments

 N                   28  Sum Wgts          28
 Mean          0.061669  Sum         1.726746
 Std Dev       0.042477  Variance    0.001804
 Skewness      -0.75971  Kurtosis    -0.44793
 USS           0.155204  CSS         0.048716
 CV            68.87855  Std Mean    0.008027
 T:Mean=0      7.682367  Pr>|T|        0.0001
 Num ^= 0            28  Num > 0           24
 M(Sign)             10  Pr>=|M|       0.0002
 Sgn Rank           189  Pr>=|S|       0.0001


               Quantiles(Def=5)

 100% Max  0.118195        99%  0.118195
  75% Q3   0.095872        95%   0.10942
  50% Med  0.069581        90%  0.108755
  25% Q1   0.037716        10%  -0.01299
   0% Min  -0.02593         5%  -0.02298
                            1%  -0.02593
 Range     0.144129
 Q3-Q1     0.058156
 Mode      -0.02593


                    Extremes

    Lowest    Obs     Highest    Obs
  -0.02593(     4) 0.103535(     28)
  -0.02298(    10) 0.107529(     21)
  -0.01299(     3) 0.108755(     19)
  -0.00426(     6)  0.10942(     20)
  0.014881(    27) 0.118195(     30)


Missing Value          .
Count                  2
% Count/Nobs        6.67


   Stem Leaf                   #  Boxplot
     10 048998                 6     |
      8 069066                 6  +-----+
      6 5545                   4  *--+--*
      4 93399                  5  |     |
      2 7                      1  +-----+
      0 56                     2     |
     -0 34                     2     |
     -2 63                     2     |
```

```
    ----+----+----+----+
Multiply Stem.Leaf by 10**-2
```

Univariate Procedure

Variable=RHO7

```
                         Normal Probability Plot
 0.11+                                                  **+*   *    *
     |                                           ******+
     |                                       *****++
     |                                  *****+
     |                               ++*+
     |                          ++++* *
     |                     +++++ * *
-0.03+   +++*   *
      +----+----+----+----+----+----+----+----+----+----+
          -2        -1         0        +1        +2
```

Univariate Procedure

Variable=RHO8

Moments

N	27	Sum Wgts	27
Mean	0.091526	Sum	2.471203
Std Dev	0.032485	Variance	0.001055
Skewness	-0.94129	Kurtosis	-0.37333
USS	0.253616	CSS	0.027436
CV	35.4921	Std Mean	0.006252
T:Mean=0	14.64031	Pr>|T|	0.0001
Num ^= 0	27	Num > 0	27
M(Sign)	13.5	Pr>=|M|	0.0001
Sgn Rank	189	Pr>=|S|	0.0001

Quantiles(Def=5)

100% Max	0.128571	99%	0.128571
75% Q3	0.117327	95%	0.124854
50% Med	0.098172	90%	0.124292
25% Q1	0.062785	10%	0.035136
0% Min	0.020771	5%	0.030688
		1%	0.020771
Range	0.107799		
Q3-Q1	0.054542		
Mode	0.020771		

Extremes

Lowest	Obs	Highest	Obs

```
  0.020771(      10)  0.12177(       28)
  0.030688(       3) 0.123752(       20)
  0.035136(       4) 0.124292(       21)
  0.040647(      22) 0.124854(       19)
  0.051684(       6) 0.128571(       30)


Missing Value          .
Count                  3
% Count/Nobs       10.00


   Stem Leaf                     #  Boxplot
     12 24459                    5     |
     11 003578                   6  +------+
     10 15                       2  |      |
      9 1225788                  7  *--+--*
      8                             |      |
      7                             |      |
      6 3                        1  +------+
      5 22                       2     |
      4 1                        1     |
      3 15                       2     |
      2 1                        1     |
        ----+----+----+----+
    Multiply Stem.Leaf by 10**-2


Univariate Procedure

Variable=RHO8

                     Normal Probability Plot
   0.125+                                  +** *  *   *
        |                             *****
        |                          ***++
        |                   ********++
        |                      +++
   0.075+                   +++
        |                ++*
        |             ++* *
        |          +++ *
        |       ++*   *
   0.025+   ++*+
         +----+----+----+----+----+----+----+----+----+----+
             -2        -1         0        +1        +2


Univariate Procedure

Variable=RHO9

                 Moments

 N                  26  Sum Wgts          26
```

```
 Mean         0.127407  Sum         3.312571
 Std Dev       0.01238  Variance    0.000153
 Skewness     -1.53732  Kurtosis    1.311749
 USS          0.425875  CSS         0.003831
 CV           9.716681  Std Mean    0.002428
 T:Mean=0     52.47697  Pr>|T|        0.0001
 Num ^= 0           26  Num > 0           26
 M(Sign)            13  Pr>=|M|       0.0001
 Sgn Rank        175.5  Pr>=|S|       0.0001


             Quantiles(Def=5)

 100% Max  0.138291        99%  0.138291
  75% Q3    0.13471        95%  0.137533
  50% Med  0.132538        90%  0.137402
  25% Q1    0.12542        10%  0.103577
   0% Min  0.094064         5%  0.102313
                            1%  0.094064
 Range     0.044228
 Q3-Q1      0.00929
 Mode      0.094064


                  Extremes

    Lowest     Obs      Highest     Obs
  0.094064(      27) 0.137114(       30)
  0.102313(      10) 0.137176(       15)
  0.103577(       7) 0.137402(       20)
  0.109454(       3) 0.137533(       21)
  0.113265(       4) 0.138291(       19)


Missing Value          .
Count                  4
% Count/Nobs       13.33


   Stem Leaf                     #  Boxplot
     13 5777788                  7  +-----+
     13 00112334444             11  *-----*
     12 59                        2  +--+--+
     12 0                         1     |
     11                                 |
     11 3                         1     |
     10 9                         1     0
     10 24                        2     0
      9
      9 4                         1     *
        ----+----+----+----+
    Multiply Stem.Leaf by 10**-2


Univariate Procedure
```

```
Variable=RHO9

                       Normal Probability Plot
  0.1375+                                +**+* * *    *
        |                      ********+**
        |                  ** *  ++++
  0.1225+                 *  ++++
        |               +++++
        |           ++++ *
  0.1075+       ++++   *
        |   ++++   * *
        |+++
  0.0925+     *
         +----+----+----+----+----+----+----+----+----+----+
             -2        -1         0        +1        +2
```

Appendix6G

A SAS output from program (2)

```
        RRO
   -0.15233   -0.11848 -0.084628 -0.050777          0  0.043646
:  0.072743    0.10184  0.130938



        RHO
  -0.093043 -0.035922 -0.004561 0.0225311 0.0578763 0.2526557
: 0.2173169 0.1853992 0.1328653

  -0.152541 -0.112992 -0.070801 -0.030338 0.0239434 0.0649262
: 0.0899643 0.1133124 0.1342584

   -0.14385 -0.210056 -0.214727 -0.088461 -0.064811 -0.039001
: -0.012986 0.0306876 0.1094544

  -0.201988 -0.230823 -0.237128 -0.235618 -0.226637 -0.067272
: -0.025934 0.0351362 0.1132646

  -0.157897 -0.138627 -0.118766 -0.093317 -0.039522 0.0184154
: 0.0587202 0.0975185 0.1322099

   -0.18748 -0.215079 -0.226746 -0.228666 -0.223148 -0.044463
: -0.004264 0.0516837 0.1204356

  -0.144655 -0.103632 -0.076379 -0.057642 -0.031859 -0.001434
: 0.0268882 0.0627848 0.1035768

  -0.029778 0.0225825 0.1462347 0.4329411 0.3431254  0.272335
: 0.2303099 0.1920124 0.1330539

  -0.150079 -0.102038 -0.053859 -0.014943 0.0284294 0.0593437
: 0.0797859 0.1462349 0.1254203

  -0.151816  -0.20018 -0.207394 -0.101576 -0.077795 -0.050537
:  -0.02298 0.0207714 0.1023134

  -0.076481 -0.006553   0.01267 0.0262203 0.0476586 0.0704985
: 0.0888963 0.1101619 0.1335608

  -0.196689 -0.214034 -0.088569  -0.05954 -0.007212 0.0415269
: 0.0738519  0.104953 0.1336305

    -0.1854  -0.20148  -0.11541 -0.083249 -0.024229 0.0301886
: 0.0653107 0.0981717 0.1687808

  -0.182849 -0.130321  -0.10168 -0.069988 -0.017111 0.0317415
: 0.0646765 0.0969821 0.1681896

  -0.153238 -0.118188 -0.077435 -0.034343 0.0264485  0.070925
: 0.0961682 0.1181363 0.1371757
```

```
  -0.192002 -0.091266 -0.049254 -0.016049  0.026863 0.0621555
: 0.0858966 0.1099969 0.1342313

  -0.216725 -0.208973 -0.094694 -0.069018  -0.02249 0.0246738
: 0.0592215 0.0954299 0.1310436

  -0.127586  -0.05403 -0.012775 0.0168752 0.0534261 0.0815975
: 0.0996438 0.1173267 0.1347104

  -0.145711 -0.089183 -0.034598 0.0095697 0.0586682  0.090647
: 0.1087552 0.1248541 0.1382911

  -0.067926  -0.00001 0.0251898 0.0448484 0.0720933 0.0947013
: 0.1094199 0.1237517 0.1374021

  -0.154478 -0.113364 -0.062899  -0.01334 0.0479039 0.0870594
: 0.1075293 0.1242923 0.1375327

  -0.194392 -0.090035  -0.05806 -0.036156 -0.013602 0.0024012
: 0.0158846 0.0406473  0.213422

  -0.181668 -0.198527 -0.202976 -0.096187 -0.041528 0.0142268
: 0.0531585 0.0919801 0.1294309

  -0.055392 -0.039437 -0.024876 -0.007934 0.0215263 0.0516421
: 0.0749359 0.1013587 0.1309394

  -0.157611 -0.136778  -0.11481 -0.088274 -0.037288 0.0155642
: 0.0534711 0.0920938 0.1296356

  -0.084654 -0.029044 -0.007296 0.0137483 0.0466771 0.0758931
:  0.095576  0.115246 0.1462383

   -0.20465 -0.227138 -0.226753  -0.22218 -0.053363 -0.016234
: 0.0148812 0.0523493 0.0940635

  -0.152127 -0.108562 -0.060382 -0.014673 0.0431066 0.0821144
:  0.103535 0.1217698  0.136639

  -0.161596 -0.145931 -0.127747 -0.104224 -0.053415 0.0053668
: 0.0485434 0.0912364 0.1303182

  -0.140211 -0.067346 -0.006049 0.0374953 0.0805432  0.105402
: 0.1181952 0.1285706 0.1371136
```

Variable	N	Mean	Std Dev	Minimum
RHO1	30	-0.1481504	0.0472468	-0.2167251
RHO2	30	-0.1198988	0.0738956	-0.2308235
RHO3	30	-0.0830843	0.0878433	-0.2371277
RHO4	30	-0.0387161	0.1155761	-0.2356177
RHO5	30	0.0014760	0.0974320	-0.2266374
RHO6	30	0.0495687	0.0741633	-0.0672717
RHO7	30	0.0724791	0.0580952	-0.0259340

RHO8	30	0.0998283	0.0403738	0.0207714
RHO9	30	0.1336401	0.0218096	0.0940635

Variable	Maximum
RHO1	-0.0297784
RHO2	0.0225825
RHO3	0.1462347
RHO4	0.4329411
RHO5	0.3431254
RHO6	0.2723350
RHO7	0.2303099
RHO8	0.1920124
RHO9	0.2134220

OBS	_TYPE_	_FREQ_	MRHO1	MRHO2	MRHO3	MRHO4
1	0	30	-0.14815	-0.11990	-0.083084	-0.038716

OBS	MRHO5	MRHO6	MRHO7	MRHO8	MRHO9
1	.0014760	0.049569	0.072479	0.099828	0.13364

OBS	VRHO1	VRHO2	VRHO3	VRHO4	VRHO5
1	.0022323	.0054606	.0077165	0.013358	.0094930

OBS	VRHO6	VRHO7	VRHO8	VRHO9
1	.0055002	.0033750	.0016300	.00047566

Univariate Procedure

Variable=RHO1

Moments

N	30	Sum Wgts	30
Mean	-0.14815	Sum	-4.44451
Std Dev	0.047247	Variance	0.002232
Skewness	0.918047	Kurtosis	0.284406
USS	0.723192	CSS	0.064735
CV	-31.8911	Std Mean	0.008626
T:Mean=0	-17.1748	Pr>|T|	0.0001
Num ^= 0	30	Num > 0	0
M(Sign)	-15	Pr>=|M|	0.0001
Sgn Rank	-232.5	Pr>=|S|	0.0001

Quantiles(Def=5)

100% Max	-0.02978	99%	-0.02978

```
 75% Q3    -0.14021        95%  -0.05539
 50% Med   -0.15289        90%   -0.0722
 25% Q1     -0.1854        10%  -0.19934
  0% Min   -0.21673         5%  -0.20465
                            1%  -0.21673
Range      0.186947
Q3-Q1      0.045189
Mode       -0.21673
```

```
                Extremes

   Lowest     Obs     Highest    Obs
 -0.21673(      17) -0.08465(      26)
 -0.20465(      27) -0.07648(      11)
 -0.20199(       4) -0.06793(      20)
 -0.19669(      12) -0.05539(      24)
 -0.19439(      22) -0.02978(       8)
```

```
   Stem Leaf                     #  Boxplot
     -2 0                        1     0
     -4 5                        1     0
     -6 68                       2     0
     -8 35                       2     |
    -10                                |
    -12 8                        1     |
    -14 884332206540            12  +--+--+
    -16 2                        1  |     |
    -18 7427532                  7  +-----+
    -20 752                      3     |
        ----+----+----+----+
    Multiply Stem.Leaf by 10**-2
```

Univariate Procedure

Variable=RHO1

```
                     Normal Probability Plot
  -0.03+                                                 *   +++
       |                                             *   ++++
       |                                         *  *++++
  -0.09+                                      **++++
       |                                   +++++
       |                               ++++  *
  -0.15+                          ***********
       |                       ++*+
       |                   *  ******
  -0.21+      *    *+*+++
        +----+----+----+----+----+----+----+----+----+----+
             -2        -1         0        +1        +2
```

Univariate Procedure

```
Variable=RHO2

                  Moments

 N                 30  Sum Wgts          30
 Mean         -0.1199  Sum         -3.59696
 Std Dev     0.073896  Variance    0.005461
 Skewness    0.104961  Kurtosis    -0.96073
 USS         0.589628  CSS         0.158356
 CV          -61.6316  Std Mean    0.013491
 T:Mean=0    -8.88704  Pr>|T|        0.0001
 Num ^= 0          30  Num > 0            1
 M(Sign)          -14  Pr>=|M|       0.0001
 Sgn Rank      -229.5  Pr>=|S|       0.0001


             Quantiles(Def=5)

 100% Max  0.022582         99%  0.022582
  75% Q3   -0.06735         95%  -0.00001
  50% Med  -0.11318         90%   -0.0178
  25% Q1   -0.20018         10%  -0.21456
   0% Min  -0.23082          5%  -0.22714
                             1%  -0.23082
 Range     0.253406
 Q3-Q1     0.132833
 Mode      -0.23082


                  Extremes

   Lowest     Obs     Highest     Obs
  -0.23082(      4) -0.03592(      1)
  -0.22714(     27) -0.02904(     26)
  -0.21508(      6) -0.00655(     11)
  -0.21403(     12) -0.00001(     20)
  -0.21006(      3) 0.022582(      8)


   Stem Leaf                     #  Boxplot
      2 3                        1     |
      0                                |
     -0 70                       2     |
     -2 969                      3     |
     -4 4                        1     |
     -6 7                        1  +-----+
     -8 109                      3  |     |
    -10 833942                   6  *-----*
    -12 970                      3  |  +  |
    -14 6                        1  |     |
    -16                             |     |
    -18 9                        1  |     |
    -20 540910                   6  +-----+
    -22 17                       2     |
        ----+----+----+----+
    Multiply Stem.Leaf by 10**-2
```

Univariate Procedure

Variable=RHO2

```
                       Normal Probability Plot
    0.03+                                            +*+
        |                                          ++
        |                                       *+*
        |                                  **+*+
        |                                 *++
        |                               +*+
        |                            +***
        |                        *****
        |                     ***+
        |                    *++
        |                 +++
        |               ++  *
        |           *+*+****
   -0.23+     *   *++
         +----+----+----+----+----+----+----+----+----+----+
             -2        -1         0        +1        +2
```

Univariate Procedure

Variable=RHO3

Moments

N	30	Sum Wgts	30
Mean	-0.08308	Sum	-2.49253
Std Dev	0.087843	Variance	0.007716
Skewness	0.034	Kurtosis	0.442815
USS	0.430867	CSS	0.223777
CV	-105.728	Std Mean	0.016038
T:Mean=0	-5.18049	Pr>|T|	0.0001
Num ^= 0	30	Num > 0	3
M(Sign)	-12	Pr>=|M|	0.0001
Sgn Rank	-197.5	Pr>=|S|	0.0001

Quantiles(Def=5)

100% Max	0.146235	99%	0.146235
75% Q3	-0.02488	95%	0.02519
50% Med	-0.07359	90%	0.004054
25% Q1	-0.11877	10%	-0.22074
0% Min	-0.23713	5%	-0.22675
		1%	-0.23713
Range	0.383362		
Q3-Q1	0.09389		
Mode	-0.23713		

```
                 Extremes

   Lowest     Obs      Highest     Obs
 -0.23713(       4) -0.00605(       30)
 -0.22675(      27) -0.00456(        1)
 -0.22675(       6)  0.01267(       11)
 -0.21473(       3)  0.02519(       20)
 -0.20739(      10) 0.146235(        8)


  Stem Leaf                    #  Boxplot
     1 5                       1     0
     1
     0
     0 13                      2     |
    -0 321110                  6  +-----+
    -0 9988766655             10  *--+--*
    -1 32210                   5  +-----+
    -1                               |
    -2 433110                  6     |
       ----+----+----+----+
   Multiply Stem.Leaf by 10**-1
```

Univariate Procedure

Variable=RHO3

```
                        Normal Probability Plot
   0.175+
        |                                             *+++++
        |                                         +++++
        |                                   ++++++*+*
  -0.025+                             +****** *
        |                       ********
        |                  *****++
        |             ++++++
  -0.225+      *++++*+* * **
        +----+----+----+----+----+----+----+----+----+----+
             -2        -1         0        +1        +2
```

Univariate Procedure

Variable=RHO4

```
                 Moments

 N                30  Sum Wgts         30
 Mean       -0.03872  Sum        -1.16148
 Std Dev    0.115576  Variance   0.013358
 Skewness   2.034541  Kurtosis   9.414231
 USS        0.432345  CSS        0.387377
 CV         -298.522  Std Mean   0.021101
```

```
T:Mean=0    -1.83478  Pr>|T|        0.0768
Num ^= 0          30  Num > 0            8
M(Sign)           -7  Pr>=|M|       0.0161
Sgn Rank      -140.5  Pr>=|S|       0.0023


           Quantiles(Def=5)

100% Max  0.432941         99%  0.432941
 75% Q3    0.00957         95%  0.044848
 50% Med  -0.03525         90%  0.031858
 25% Q1   -0.08846         10%   -0.1632
  0% Min  -0.23562          5%  -0.22867
                            1%  -0.23562
Range     0.668559
Q3-Q1      0.09803
Mode      -0.23562


                Extremes

   Lowest     Obs     Highest    Obs
 -0.23562(      4) 0.022531(      1)
 -0.22867(      6)  0.02622(     11)
 -0.22218(     27) 0.037495(     30)
 -0.10422(     29) 0.044848(     20)
 -0.10158(     10) 0.432941(      8)


  Stem Leaf                     #  Boxplot
     4 3                        1     *
     3
     3
     2
     2
     1
     1
     0
     0 1122344                  7  +-----+
    -0 43321111                 8  *--+--*
    -0 99987766                 8  +-----+
    -1 000                      3     |
    -1                                |
    -2 432                      3     0
       ----+----+----+----+
   Multiply Stem.Leaf by 10**-1


Univariate Procedure

Variable=RHO4

                         Normal Probability Plot
  0.425+                                                  *
       |
       |
```

```
          |                                                *     +
          |                                                 ++++
          |                                             ++++
          |                                         ++++
          |                                    +++++
          |                                +++*****  *  *  *
          |                            *******
          |                   *********
          |              *  *++++
          |             ++++
  -0.225+      *  ++*+*
          +----+----+----+----+----+----+----+----+----+----+
               -2        -1         0        +1        +2
```

Univariate Procedure

Variable=RHO5

Moments

N	30	Sum Wgts	30
Mean	0.001476	Sum	0.044279
Std Dev	0.097432	Variance	0.009493
Skewness	0.681362	Kurtosis	5.641961
USS	0.275362	CSS	0.275297
CV	6601.247	Std Mean	0.017789
T:Mean=0	0.082973	Pr>|T|	0.9344
Num ^= 0	30	Num > 0	15
M(Sign)	0	Pr>=|M|	1.0000
Sgn Rank	18.5	Pr>=|S|	0.7104

Quantiles(Def=5)

100% Max	0.343125	99%	0.343125
75% Q3	0.047659	95%	0.080543
50% Med	0.007157	90%	0.065381
25% Q1	-0.03952	10%	-0.0713
0% Min	-0.22664	5%	-0.22315
		1%	-0.22664
Range	0.569763		
Q3-Q1	0.087181		
Mode	-0.22664		

Extremes

Lowest	Obs	Highest	Obs
-0.22664(	4)	0.057876(	1)
-0.22315(	6)	0.058668(	19)
-0.0778(	10)	0.072093(	20)
-0.06481(	3)	0.080543(	30)
-0.05342(	29)	0.343125(	8)

```
 Stem Leaf                     #  Boxplot
    3 4                        1     *
    2
    2
    1
    1
    0 55556678                 8  +-----+
    0 223334                   6  *--+--*
   -0 444322211                9  +-----+
   -0 8655                     4     |
   -1
   -1
   -2 32                       2     0
      ----+----+----+----+
  Multiply Stem.Leaf by 10**-1
```

Univariate Procedure

Variable=RHO5

```
                       Normal Probability Plot
 0.325+                                             *
      |
      |                                              +++++
      |                                         +++++
      |                                    +++++
      |                              +++++** * * *
      |                         *********
      |                 *********
      |           * * **+++
      |          +++++
      |     +++++
-0.225++++++*   *
       +----+----+----+----+----+----+----+----+----+----+
           -2        -1         0        +1        +2
```

Univariate Procedure

Variable=RHO6

Moments

N	30	Sum Wgts	30
Mean	0.049569	Sum	1.487061
Std Dev	0.074163	Variance	0.0055
Skewness	1.363486	Kurtosis	3.184025
USS	0.233217	CSS	0.159506
CV	149.6171	Std Mean	0.01354
T:Mean=0	3.660828	Pr>|T|	0.0010
Num ^= 0	30	Num > 0	24
M(Sign)	9	Pr>=|M|	0.0014
Sgn Rank	168.5	Pr>=|S|	0.0001

```
            Quantiles(Def=5)

 100% Max   0.272335       99%  0.272335
  75% Q3    0.081598       95%  0.252656
  50% Med   0.046584       90%  0.100052
  25% Q1    0.005367       10%  -0.04173
   0% Min   -0.06727        5%  -0.05054
                            1%  -0.06727
 Range      0.339607
 Q3-Q1      0.076231
 Mode       -0.06727


                  Extremes

   Lowest     Obs      Highest     Obs
 -0.06727(       4) 0.090647(       19)
 -0.05054(      10) 0.094701(       20)
 -0.04446(       6) 0.105402(       30)
   -0.039(       3) 0.252656(        1)
 -0.01623(      27) 0.272335(        8)


  Stem Leaf                      #  Boxplot
     2 57                        2     0
     2
     1
     1 1                         1     |
     0 566677888999             12  +--+--+
     0 011222334                 9  +-----+
    -0 4420                      4     |
    -0 75                        2     |
       ----+----+----+----+
   Multiply Stem.Leaf by 10**-1


Univariate Procedure

Variable=RHO6

                     Normal Probability Plot
  0.275+                                          *   *
       |                                               +++++
       |                                          +++++++
       |                                   +++++++*
       |                              *********** *
       |                         *********
       |                    *+*+**++
 -0.075+              *+++*++
        +----+----+----+----+----+----+----+----+----+----+
            -2        -1         0        +1        +2


Univariate Procedure
```

```
Variable=RHO7

                     Moments

 N                   30  Sum Wgts          30
 Mean          0.072479  Sum         2.174372
 Std Dev       0.058095  Variance    0.003375
 Skewness      0.733946  Kurtosis    1.740269
 USS           0.255473  CSS         0.097876
 CV            80.15438  Std Mean    0.010607
 T:Mean=0      6.833345  Pr>|T|        0.0001
 Num ^= 0            30  Num > 0           26
 M(Sign)             11  Pr>=|M|       0.0001
 Sgn Rank         218.5  Pr>=|S|       0.0001


                Quantiles(Def=5)

 100% Max    0.23031         99%    0.23031
  75% Q3    0.099644         95%   0.217317
  50% Med   0.074394         90%   0.113808
  25% Q1    0.048543         10%   -0.00862
   0% Min   -0.02593          5%   -0.02298
                              1%   -0.02593
 Range      0.256244
 Q3-Q1        0.0511
 Mode       -0.02593


                     Extremes

    Lowest     Obs      Highest      Obs
  -0.02593(      4)  0.108755(       19)
  -0.02298(     10)   0.10942(       20)
  -0.01299(      3)  0.118195(       30)
  -0.00426(      6)  0.217317(        1)
  0.014881(     27)   0.23031(        8)


   Stem Leaf                        #   Boxplot
     22 0                           1      0
     20 7                           1      0
     18
     16
     14
     12
     10 048998                      6   +-----+
      8 069066                      6   |     |
      6 5545                        4   *--+--*
      4 93399                       5   +-----+
      2 7                           1      |
      0 56                          2      |
     -0 34                          2      |
     -2 63                          2      |
        ----+----+----+----+
    Multiply Stem.Leaf by 10**-2
```

Univariate Procedure

Variable=RHO7

```
                       Normal Probability Plot
    0.23+                                             *
        |                                         *     ++++
        |                                            +++
        |                                         +++
        |                                     ++++
        |                                  +++
        |                              +++*** * *
        |                           ******
        |                       ****
        |                  *****
        |                +*++
        |             ++**
        |          +*+*
   -0.03+     *++++*
        +----+----+----+----+----+----+----+----+----+----+
            -2        -1         0        +1        +2
```

Univariate Procedure

Variable=RHO8

Moments

N	30	Sum Wgts	30
Mean	0.099828	Sum	2.99485
Std Dev	0.040374	Variance	0.00163
Skewness	0.019259	Kurtosis	0.443616
USS	0.346242	CSS	0.047271
CV	40.44321	Std Mean	0.007371
T:Mean=0	13.543	Pr>|T|	0.0001
Num ^= 0	30	Num > 0	30
M(Sign)	15	Pr>=|M|	0.0001
Sgn Rank	232.5	Pr>=|S|	0.0001

Quantiles(Def=5)

100% Max	0.192012	99%	0.192012
75% Q3	0.12177	95%	0.185399
50% Med	0.103156	90%	0.137403
25% Q1	0.091236	10%	0.037892
0% Min	0.020771	5%	0.030688
		1%	0.020771
Range	0.171241		
Q3-Q1	0.030533		
Mode	0.020771		

```
                  Extremes

   Lowest     Obs      Highest    Obs
 0.020771(      10) 0.124854(      19)
 0.030688(       3) 0.128571(      30)
 0.035136(       4) 0.146235(       9)
 0.040647(      22) 0.185399(       1)
 0.051684(       6) 0.192012(       8)
```

```
  Stem Leaf                     #  Boxplot
    18 52                       2     0
    16
    14 6                        1     |
    12 24459                    5  +-----+
    10 15003578                 8  *--+--*
     8 1225788                  7  +-----+
     6 3                        1     |
     4 122                      3     0
     2 115                      3     0
       ----+----+----+----+
   Multiply Stem.Leaf by 10**-2
```

Univariate Procedure

Variable=RHO8

```
                       Normal Probability Plot
   0.19+                                                 *    *++++
       |                                                +++++
       |                                           ++++*
       |                                      ++**** *
   0.11+                                 *******
       |                          *******+
       |                        +*+++
       |                   ++*+**
   0.03+     *++++*+*
        +----+----+----+----+----+----+----+----+----+----+
             -2        -1         0        +1        +2
```

Univariate Procedure

Variable=RHO9

```
                  Moments

 N                 30  Sum Wgts         30
 Mean         0.13364  Sum        4.009202
 Std Dev      0.02181  Variance   0.000476
 Skewness    1.599801  Kurtosis   5.713664
 USS         0.549584  CSS        0.013794
 CV          16.31965  Std Mean   0.003982
```

```
T:Mean=0    33.56214  Pr>|T|        0.0001
Num ^= 0          30  Num > 0           30
M(Sign)           15  Pr>=|M|       0.0001
Sgn Rank       232.5  Pr>=|S|       0.0001


           Quantiles(Def=5)

100% Max  0.213422       99%  0.213422
 75% Q3   0.137176       95%  0.168781
 50% Med  0.133307       90%  0.157214
 25% Q1   0.129431       10%  0.106516
  0% Min  0.094064        5%  0.102313
                          1%  0.094064
Range     0.119359
Q3-Q1     0.007745
Mode      0.094064


                Extremes

   Lowest    Obs     Highest    Obs
 0.094064(      27) 0.138291(      19)
 0.102313(      10) 0.146238(      26)
 0.103577(       7)  0.16819(      14)
 0.109454(       3) 0.168781(      13)
 0.113265(       4) 0.213422(      22)


   Stem Leaf                     #  Boxplot
     21 3                        1     *
     20
     19
     18
     17
     16 89                       2     *
     15
     14 6                        1     |
     13 001123344445777788      18  +--+--+
     12 059                      3  +-----+
     11 3                        1     0
     10 249                      3     0
      9 4                        1     *
        ----+----+----+----+
    Multiply Stem.Leaf by 10**-2


Univariate Procedure

Variable=RHO9

                       Normal Probability Plot
   0.215+                                                  *
        |
        |
        |                                                     ++++
```

```
     |                                          +++++
     |                                      +*+*
0.155+                                 +++++
     |                            +++++    *
     |                    ****************
     |                ****++++
     |               *+++
     |         *+*+*+
0.095+     *++++
      +----+----+----+----+----+----+----+----+----+----+
          -2        -1         0        +1        +2
```

Appendix6H

The eigenvalues of W(0,1,2,3,4)

W's entries are 0, 1, 2, 3, 4.

14.059360	0.000000E+00
8.023599	0.000000E+00
6.534783	0.000000E+00
4.083396	0.000000E+00
3.894535	-1.389182
3.894535	1.389182
1.310192	0.000000E+00
9.841095E-01	-1.732912
9.841095E-01	1.732912
5.363135E-01	0.000000E+00
-7.958040E-08	0.000000E+00
-4.903842E-01	0.000000E+00
-1.171220	-1.956163
-1.171220	1.956163
-1.336884	-1.309866
-1.336884	1.309866
-2.525492	-1.615801
-2.525492	1.615801
-4.415877	0.000000E+00
-4.827520	-2.857574
-4.827520	2.857574
-4.899200	0.000000E+00
-7.075458	0.000000E+00
-7.701774	0.000000E+00

Appendix6I

The comparison of moral levels

T-Test

Group Statistics

	W1NW2	N	Mean	Std. Deviation	Std. Error Mean
MORAL1	1.00	16	.1510794	.2560826	6.40E-02
	2.00	8	.2155950	.4052280	.1432697

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means				
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference
MORAL1	Equal variances assumed	1.477	.237	-.478	22	.637	-6.5E-02	.1348340
	Equal variances not assumed			-.411	9.891	.690	-6.5E-02	.1569231

T-Test

Group Statistics

	NB1B2	N	Mean	Std. Deviation	Std. Error Mean
MORAL1	1.00	18	.1092361	.2874029	6.77E-02
	2.00	6	.3626300	.3040383	.1241231

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means				
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference
MORAL1	Equal variances assumed	.044	.836	-1.845	22	.078	.2533939	.1373046
	Equal variances not assumed			-1.792	8.208	.110	.2533939	.1414053

T-Test

Group Statistics

	W1NW2	N	Mean	Std. Deviation	Std. Error Mean
MORAL2	1.00	16	.1496112	.2557655	6.39E-02
	2.00	8	.2155950	.4052280	.1432697

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means				
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference
MORAL2	Equal variances assumed	1.494	.234	-.490	22	.629	-6.60E-02	.1347570
	Equal variances not assumed			-.421	9.883	.683	-6.60E-02	.1568908

T-Test

Group Statistics

	NB1B2	N	Mean	Std. Deviation	Std. Error Mean
MORAL2	1.00	18	.1079311	.2869526	6.76E-02
	2.00	6	.3626300	.3040383	.1241231

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means				
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference
MORAL2	Equal variances assumed	.048	.828	-1.857	22	.077	.2546989	.1371428
	Equal variances not assumed			-1.802	8.197	.108	.2546989	.1413545

Appendix6J

The race×sex×moral table

Correlations

		SEX	YEAR	MORAL1	MORAL2	MORAL3
Pearson Correlation	SEX	1.000	-.252	.024	.026	.050
	YEAR	-.252	1.000	.160	.157	.126
	MORAL1	.024	.160	1.000	1.000**	.988**
	MORAL2	.026	.157	1.000**	1.000	.991**
	MORAL3	.050	.126	.988**	.991**	1.000
	W1NW2	.299	.029	.101	.104	.123
	NB1B2	.293	.099	.366	.368	.382
Sig. (2-tailed)	SEX	.	.235	.913	.903	.815
	YEAR	.235	.	.455	.464	.556
	MORAL1	.913	.455	.	.000	.000
	MORAL2	.903	.464	.000	.	.000
	MORAL3	.815	.556	.000	.000	.
	W1NW2	.156	.892	.637	.629	.566
	NB1B2	.165	.645	.078	.077	.066
N	SEX	24	24	24	24	24
	YEAR	24	24	24	24	24
	MORAL1	24	24	24	24	24
	MORAL2	24	24	24	24	24
	MORAL3	24	24	24	24	24
	W1NW2	24	24	24	24	24
	NB1B2	24	24	24	24	24

Appendix6J (Continued)

The race×sex×moral table

Correlations

		W1NW2	NB1B2
Pearson Correlation	SEX	.299	.293
	YEAR	.029	.099
	MORAL1	.101	.366
	MORAL2	.104	.368
	MORAL3	.123	.382
	W1NW2	1.000	.816**
	NB1B2	.816**	1.000
Sig. (2-tailed)	SEX	.156	.165
	YEAR	.892	.645
	MORAL1	.637	.078
	MORAL2	.629	.077
	MORAL3	.566	.066
	W1NW2	.	.000
	NB1B2	.000	.
N	SEX	24	24
	YEAR	24	24
	MORAL1	24	24
	MORAL2	24	24
	MORAL3	24	24
	W1NW2	24	24
	NB1B2	24	24

**. Correlation is significant at the 0.01 level (2-tailed).

Appendix6K

The eigenvalues of W(0,1)

W's entries are 0, 1 only.

4.051115	0.000000E+00
2.651598	0.000000E+00
1.883632	0.000000E+00
1.548894	-3.286612E-01
1.548894	3.286612E-01
7.357295E-01	0.000000E+00
3.613770E-01	-6.469532E-01
3.613770E-01	6.469532E-01
2.966202E-01	-6.964666E-02
2.966202E-01	6.964666E-02
5.441884E-07	0.000000E+00
-2.522140E-01	-7.226394E-01
-2.522140E-01	7.226394E-01
-3.105704E-01	0.000000E+00
-6.678435E-01	-1.496080E-01
-6.678435E-01	1.496080E-01
-6.870127E-01	-5.055384E-01
-6.870127E-01	5.055384E-01
-1.397442	-9.904619E-01
-1.397442	9.904619E-01
-1.533783	0.000000E+00
-1.899674	-9.174101E-02
-1.899674	9.174101E-02
-2.083131	0.000000E+00

REFERENCES

Anselin, L. (1988) Spatial Economics: Methods and Models. Kluwer Academic Publishers.

Bartels, C. P. A. (1979) Operational Statistical Methods for Analyzing Spatial Data. Exploratory and Explanatory Statistical Analysis of Spatial Data. C. P. A. Bartels and R. H. Ketellapper. (Eds.) Martinus Nijhoff, Boston, Mass. pp. 5-50.

Cliff, A.D. and Ord, J.K. (1973) Spatial Autocorrelation. London. Pion Ltd.

Cressie, N. A. C. (1993) Statistics for Spatial Data (Revised Edition). John Wiley & Sons, Inc. New York.

Doreian, P. (1981) Estimating Linear Models with Spatially Distributed Data. Sociological methodology. (Leinhardt, S. Ed.) pp. 359-388. S.F. Jessey-Bass.

Doreian, P. (1982). Maximum Likelihood Methods for Linear Models. Spatial Effect and Spatial Disturbance Terms. Social Methods & Research. Vol.10 No.3, Feb. 1982. pp. 245-269. Sage Publications, Inc.

Doreian, P. (1989) Two Regions of Network Effects Autocorrelation. pp. 280-295.

Doreian, P. and Hunmon, N. P. (1976) Modeling Social Processes. New York: Elsevier.

Duke, J.B. (1993) Estimation of the Network Effects Model in a Large Data Set. Social Methods & Research. Vol.21 No.4, May 1993. pp. 465-481. Sage Publications, Inc.

Frank, K. (1995) Identifying Cohesive Subgroups. Social Networks. 17 (1995), pp. 27-56.

Frank, K. (1996) Maping Interactions within and between Cohesive Subgroups. Social Networks. 18 (1996), pp. 93-119.

Harman, Harry H. (1960) Modern Factor Analysis. The University of Chicago Press.

Holland, Paul W. and Leinhardt, Samuel (1981) An Exponential Family of Probability Distributions for Directed Graphs. Journal of the American Statistical Association. 1981 March, Volume 78, Number 373. pp. 33-50.

Kira, T. K., Ogawa, H., and Sakazaki, N. (1953) Intraspecific competition among higher plants. I: Competition-yield-density interrelationship in regularly dispersed populations. *Journal of the Institute of Polytechnics, Osaka City University*, Series D, 4, 1-16.

Leenders, R.A. (1995) Models for Network Dynamics: A Markovian Framework. Journal of Mathematical Sociology. 1995, pp. 1-21.

Marsden, P.V. and Friedkin, N.E. (1994) Network Studies of Social Influence. Social Psychology and Diffusion.

Mead, R. (1967) A Mathematical Model for the Estimation of Interplant Competition. Biometrics, 23 (June 1967), pp. 189-205.

Mead, R. (1971) Statistical Ecology. (Patil, G.P. etc. ed.) Vol. 2, 1971, pp. 13-30.University Park: Pennsylvania State Univ. Press.

Ord, K. (1975) Estimation Methods for Models of Spatial Interaction. Journal of the American Statistical Association. 1975 March, pp. 120-126.

Ripley, B. D. (1981) Spatial Statistics. John Wiley & Sons, Inc. New York.

Shavelson, R. J. (1996) Statistical Reasoning for the Behavioral Sciences. (3rd ed.) ALLYN and BACON, Boston.

Whittle, P. (1954) On Stationary Processes in the Plane. Biometrika, 41 (Parts 3 and 4), pp. 434-449.

Xu, J. (1996) Solution Methods for Spatial Linear Models, An unpublished apprenticeship paper. College of Education, Michigan State University.