To the Graduate Council:

I am submitting herewith a thesis written by Cathy W. Dyer entitled "The Effects of Matrix Scaling on the Design and Control of Multivariable Processes." I have examined the final copy of this thesis for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Master of Science, with a major in Chemical Engineering.

Duane D Bruns

Duane D. Bruns, Major Professor

We have read this thesis
and recommend its acceptance:

Charles F. Moore

J M Bailey

Accepted for the Council:

C W Minkel

Vice Provost
and Dean of The Graduate School

STATEMENT OF PERMISSION TO USE

In presenting this thesis in partial fulfillment of the requirements for a Master's degree at The University of Tennessee, Knoxville, I agree that the Library shall make it available to borrowers under rules of the Library. Brief quotations from this thesis are allowable without special permission, provided that accurate acknowledgment of the source is made.

Permission for extensive quotation from or reproduction of this thesis may be granted by my major professor, or in his absence, by the Head of Interlibrary Services when, in the opinion of either, the proposed use of the material is for scholarly purposes. Any copying or use of the material in this thesis for financial gain shall not be allowed without my written permission.

Signature Cathy W. Dyll

Date June 1987

THE EFFECTS OF MATRIX SCALING ON THE DESIGN AND CONTROL OF MULTIVARIABLE PROCESSES

A Thesis

Presented for the

Master of Science

Degree

The University of Tennessee, Knoxville

Cathy W. Dyer

June 1987

ACKNOWLEDGEMENTS

The author would like to thank the people on her committee for their assistance during the research. In particular, she would like to thank Dr. Duane Bruns for his guidance. She would also like to express her gratitude to the Measurement and Control Engineering Center for their financial support and encouragement during the research. She would like to especially thank Gim Gheng Khor for her resourcefulness and diligence during the simulation of the multivariable tuning regulator and the singular value analysis controller.

ABSTRACT

The motivation behind this work was to define matrix scaling so that process conditioning and process interaction applied to steady state process gain matrices would be put on a common basis. Since units are attached to the elements of a steady state gain matrix, it is often difficult to interpret what these elements mean on a physical basis.

In order to lower condition number, both empirical scaling methods and a numerical nonlinear search technique (NLST) were used. Much of this work was based on the scaling techniques used by Prasad (1). For analysis purposes, process steady state gain matrices were arbitrarily classified into three categories depending on the number and position of zeros in the matrix. These three categories were noninteracting, partially interacting, and full.

For noninteracting matrices of dimension 2x2, 3x3, and 4x4, geometric mean scaling and NLST yielded the same minimum condition number. Full matrices of dimension 2x2 had the same minimum condition when scaled either by geometric mean scaling or by a nonlinear least squares technique. By using an analytical method based on Tomlin's algorithm for geometric mean scaling, it can be shown that the minimum condition number for full, 2x2 matrices is the same for row followed by column scaling or for column followed by row scaling. Results of scaling using other empirical techniques such as arithmetic mean or equilibration were mixed. In general, NLST was the more effective method in finding minimum condition

number for all types of low order matrices. Much more work is needed to develop a framework for determining interaction for multivariable processes on a common basis.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

FIGURE PAGE

LIST OF SYMBOLS

A	any general matrix
a	any general vector
a_{ij}	scalar; one element in a matrix
c	controlled variable
diag	diagonal matrix
E	error matrix
G(s)	general transfer function matrix
g_{ij}	scalar transfer function between i^{-th} output and j^{-th} input
K	process steady state gain matrix
m	manipulated variable
r	input
U	left singular vectors, orthogonal matrix of SVD
V	right singular vectors, orthogonal matrix of SVD
y	output
κ	condition number; largest singular value divided by smallest nonzero singular value
Λ	relative gain array of a given process steady state matrix
σ_{ij}	elements of the relative gain array (RGA)
μ	eigenvalue
Σ	diagonal matrix of SVD
σ	element of Σ matrix
σ_*	minimum element of Σ matrix
ζ	disturbance input

CHAPTER 1

INTRODUCTION

Definition of Matrix Scaling

A matrix is scaled if it is pre-multiplied, post-multiplied or both pre- and post-multiplied by diagonal matrices. The general form of a scaled matrix will be shown as

$$K_{SC} = S_1 K S_2 \tag{1.1}$$

where S_1 and S_2 are diagonal matrices consisting of real, nonzero elements and K is the matrix to be scaled. The nonzero elements of S_1 and S_2 are called scaling factors. Pre-multiplication of K by S_1 performs row scaling while post-multiplication of K by S_2 does column scaling.

For mxn matrices, the following notation will be used for the scaling matrices:

$$S_1 = \text{diag}\ (S_1, \ldots, S_m) \tag{1.2}$$

$$S = \text{diag}\ (S_{m+1}, \ldots S_{m+n}) \tag{1.3}$$

For example consider an arbitrary 3x3 matrix

$$K_{sc} = \begin{bmatrix} S_1 & 0 & 0 \\ 0 & S_2 & 0 \\ 0 & 0 & S_3 \end{bmatrix} \begin{bmatrix} K_{11} & K_{12} & K_{13} \\ K_{21} & K_{22} & K_{23} \\ K_{31} & K_{32} & K_{33} \end{bmatrix} \begin{bmatrix} S_4 & 0 & 0 \\ 0 & S_5 & 0 \\ 0 & 0 & S_6 \end{bmatrix} \tag{1.4}$$

$$K_{sc} = \begin{bmatrix} S_1 K_{11} S_4 & S_1 K_{12} S_5 & S_1 K_{13} S_6 \\ S_2 K_{21} S_4 & S_2 K_{22} S_5 & S_2 K_{23} S_6 \\ S_3 K_{31} S_4 & S_3 K_{23} S_5 & S_3 K_{33} S_6 \end{bmatrix} \tag{1.5}$$

Categorization of Process Variables

Process variables can be categorized into the following two classes:

1. Independent variables (input variables) are process variables that are not influenced by the operation of the process but do affect its operation. These variables can be subdivided into manipulated or disturbance variables:
   a. Manipulated variables are adjustable variables.
   b. Disturbance variables are neither adjusted nor manipulated by a control system or human operator.
2. Dependent variables (output variables) refer to process variables that are affected by the operation or state of the process. Dependent variables can be classified further into controlled variables and intermediate or secondary variables:
   a. Controlled variables are typically those sought to be maintained at given set point.

b. Intermediate or secondary variables are those variables which are of secondary importance and may not be measured.

Notation Conventions

Typically, transfer function descriptions of processes are widely used in the chemical process industries. A process is often described in terms of either fitted experimental plant data or a model consisting of mass, component, energy, momentum, or other balances. A linear, time invariant (LTI) model is the most common and will be used throughout this work unless specified otherwise. For multiple-input multiple-output systems (MIMO), transfer function matrices are used. For a LTI system with two inputs and two outputs, the general transfer function would be

$$G(s) = \begin{bmatrix} g_{11}(s) & g_{12}(s) \\ g_{21}(s) & g_{22}(s) \end{bmatrix} \tag{1.6}$$

where $g_{ij}(s)$ is a scalar transfer function between the i-th output and the j-th input. This transfer function description is shown in Figure 1.1.

Another method of expressing the linearized process equations is in state space form as

$$\dot{x} = Ax + Bu \tag{1.7}$$

$$y = Cx + Du \tag{1.8}$$

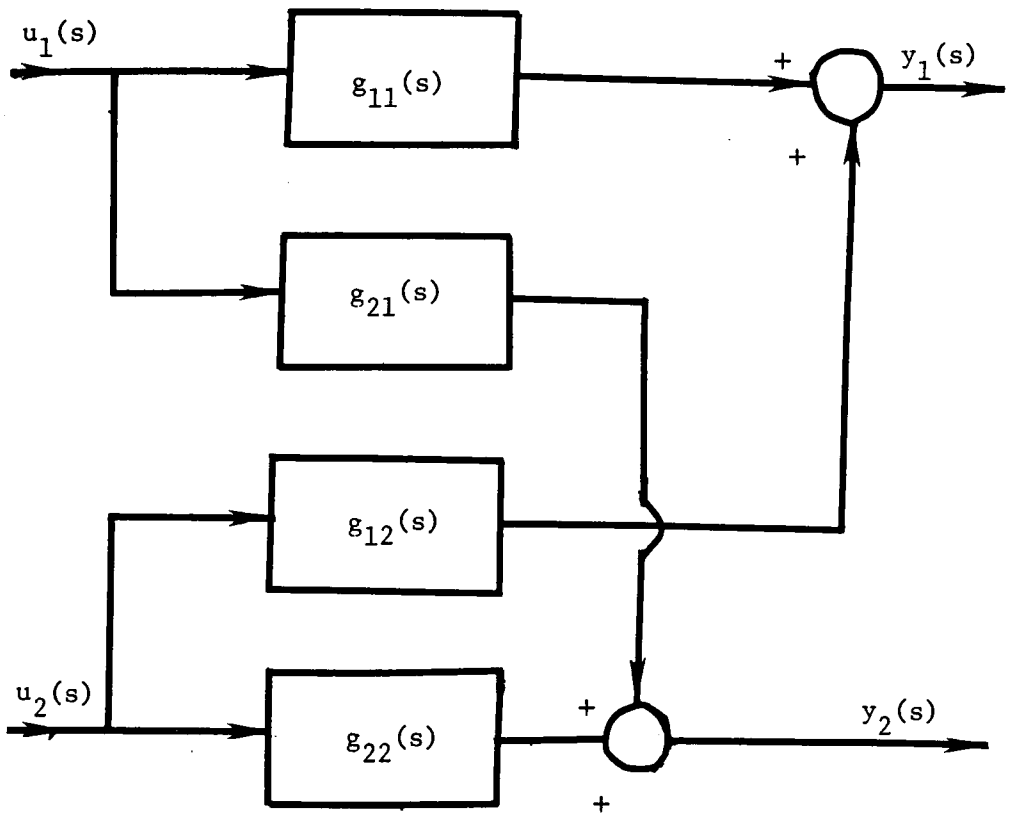


Figure 1.1. Process with Two Inputs and Two Outputs.

where u input vector of dimension mx1

x state vector of dimension nx1

y output vector of dimension px1

A system matrix of dimension nxn

B input or distribution matrix of dimension nxm

C output matrix of dimension pxn

D feedforward matrix of dimension pxm

The state space form for the closed loop system is shown in Figure 1.2.

Most of this work will be limited to steady state analysis. The process steady state gain matrix is obtained from the state space equations (1.7, 1.8) by setting $\dot{x}$=0.0.

$$x = -A^{-1}Bu \tag{1.9}$$

$$y = (-CA^{-1}B+D)u = Ku \tag{1.10}$$

$$K = -CA^{-1}B + D$$

Similarly the state equations (1.7, 1.8) can be written with the LaPlace operator as

$$sX(s) = AX(s) + BU(s) \tag{1.11}$$

$$Y(s) = CX(s) + DU(s) \tag{1.12}$$

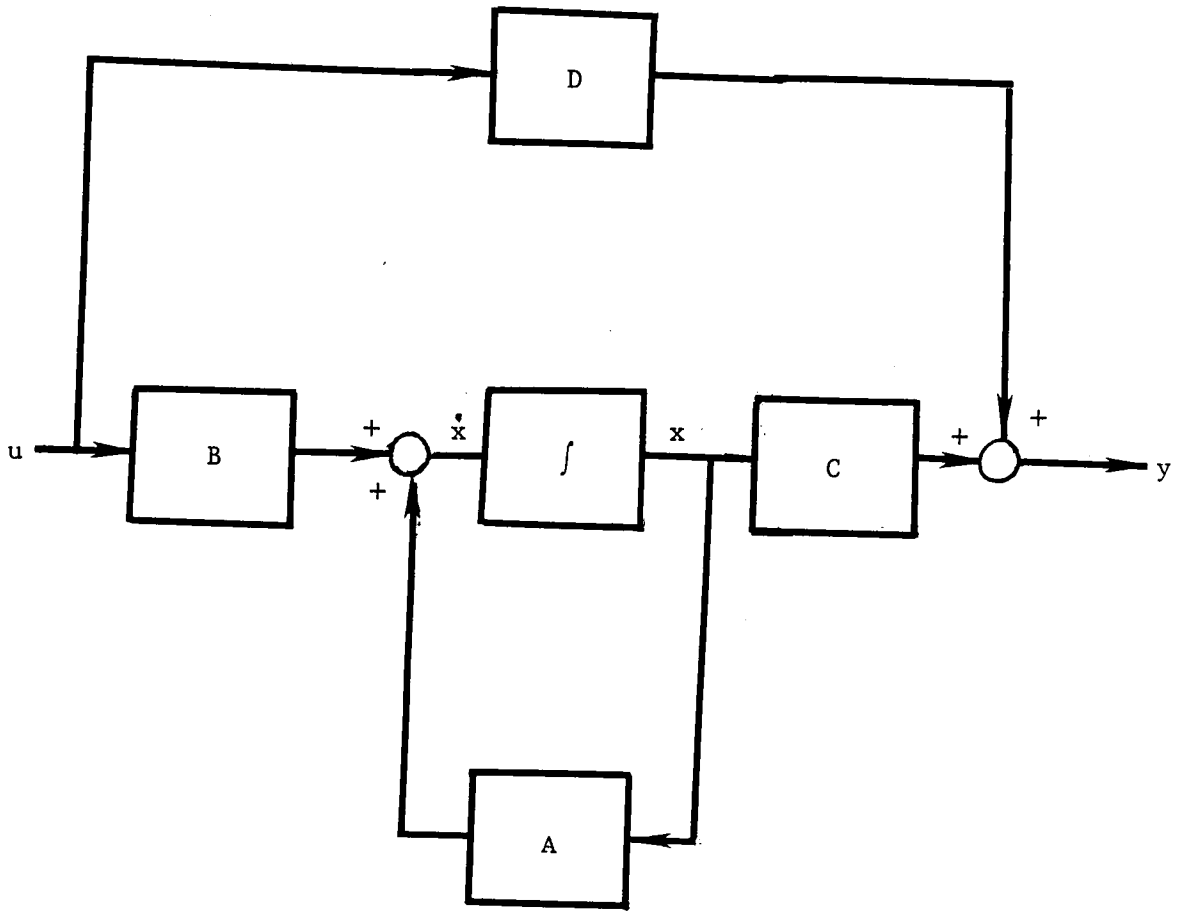


Figure 1.2. State Space Model for a Continuous Time Linear System.

Solving equation (1.11) for X and substitution into equation (1.12) gives

$$X(s) = (sI-A)^{-1}BU(s) \tag{1.13}$$

$$\dot{Y}(s) = [C(sI-A)^{-1}B + DU(s)] = G(s)U(s) \tag{1.14}$$

$$G(s) = [C(sI-A)^{-1}B + D]$$

The process transfer function matrix is represented by G(s). To determine the steady state gain matrix, set s=0.0 so that equation (1.14) becomes

$$Y(s) = [-CA^{-1}B + D](s)U(s) = KU(s) \tag{1.15}$$

which is the same result as equation (1.10). Equation (1.16) will hold as long as A^{-1} exists.

Further simplifications of equation (1.16) for the steady state process gain matrix are frequently made if the output matrix is the identity matrix. Therefore, equation (1.8) reduces in many cases to y=x. Under these same assumptions, equation (1.16) becomes

$$K = -A^{-1}B \tag{1.17}$$

This chapter contains a mathematical definition of matrix scaling that will be used throughout this work. The effect of

scaling will be examined in relation to the interpretation of process conditioning and interaction. Interaction of the process gain matrix will be viewed in terms of the possible loop pairings indicated by the relative gain array versus that of singular value analysis which will be described further in subsequent chapters.

CHAPTER 2

A BRIEF REVIEW OF THE MULTIVARIABLE CONTROL LITERATURE

Around 1869 scalar time response methods began with Maxwell's work on steam governors (4). The origin of scalar frequency response methods dates from 1932 with Nyquist's work. Around 1966 marks the beginning of multivariable, vector, frequency methods with the work of Rosenbrock (5).

Foss criticized the application of modern optimal control theory in the process industries (6). Several reasons for the lack of success in optimal control implementations are the limited knowledge of the actual processes and their often inherently nonlinear character. More success has been attained by controlling single-input single-output (SISO) systems with modern control theory than with multiple-input multiple-output (MIMO) systems. Problems with right hand plane transmission zeros and difficulty in assessing sensor failure were several reasons for the limited success of optimal control in MIMO systems.

Rijnsdorp and Seborg prepared a survey of advanced control applications in the process industries through 1976 (7). They tabulated thirty industrial applications of advanced control into the following nine categories:

1. dominant interaction;
2. decoupled or non-interacting control;
3. inverse and direct Nyquist array;

4. characteristic loci;
5. optimal control;
6. modal or pole placement;
7. Lyapunov functions;
8. incomplete state feedback;
9. adaptive control.

Decoupling and optimal control were cited in two-thirds of the applications (7). The other seven methods were divided among the ten remaining applications.

Interaction Measurement

Pairing a manipulated variable with a controlled variable and tuning each SISO loop independently is one simple method to control MIMO systems. When there is not much interaction among the process variables, SISO design techniques have been relatively successful. This approach to formalizing control structure is often termed multivariable single-input single-output (MVSISO).

Interaction is measured either at steady state or in terms of its dynamic response. Both time and frequency domain methods are used to examine dynamic interaction.

Relative Gain Array

A widely used measure of interaction is the relative gain array (RGA) (8). In general, the RGA is defined to be multiplication of corresponding elements of the steady state gain matrix by the transpose of its inverse (9). Therefore, the RGA is limited to

square, invertible systems. The RGA can be defined in relation to process control as a measure of the gain of a process without any control (the open loop gain) divided by the gain of the process with all loops perfectly controlled except for the one being studied (10).

Symbolically, the RGA, denoted by Λ, is represented as

$$\lambda_{ij} = \frac{\left(\frac{\partial y_i}{\partial u_j}\right)_u}{\left(\frac{\partial y_i}{\partial u_j}\right)_y}$$

$$\Lambda = \begin{bmatrix} \lambda_{11}, \cdots, \lambda_{1n} \\ \lambda_{1m}, \cdots, \lambda_{mn} \end{bmatrix} \qquad \text{where } m=n \qquad (2.1)$$

where u jth process input, u

y ith process output, y

u input variable

y controlled variable

For the case of a 2x2 matrix $\lambda_{11} = \frac{1}{1 - \frac{K_{21}K_{12}}{K_{11}K_{22}}}$

Some of the useful properties of RGA are listed (3, 12, 13). The properties are for nxn except for where noted as being listed as 2x2.

1. All rows and columns of the RGA matrix (Λ) add to one.
2. The elements of the RGA matrix are scale invariant and unitless.

3. A permutation of rows and columns in the steady state gain matrix will result in the same permutations in the RGA.
4. Large elements (Λ_{ij}) indicate that the process steady state matrix is nearly singular.
5. RGA elements of completely isolated subsystems are the same as the corresponding elements of the original system.
6. The RGA provides an estimate of the measure's sensitivity to parameter changes.
7. For 2x2 systems, the RGA will be the identity matrix if the steady state gain matrix is diagonal or triangular. For the diagonal matrix, the system is called decoupled while the triangular is referred to as one-way interactive.
8. For a 2x2 steady state gain matrix with only nonzero entries, the values of the RGA elements fit into
   a. If there are an odd number of positive elements in the gain matrix, then $\Lambda_{ij} \varepsilon (0,1)$ for i,j = 1,2.
   b. If there are an even number of positive elements in the gain matrix, then $\Lambda_{ij} \varepsilon (-\infty, 0) \cup (1, \infty)$ for i,j = 1,2.

The RGA pairing rules will be described later.

Dynamic Interaction Measures

Rijnsdorp (11), proposed one of the first interaction measures in the frequency domain for 2x2 systems. His interaction quotient

was

$$I = \frac{g_{12}g_{21}}{g_{11}g_{22}}\Bigg|_{s = jw} \tag{2.2}$$

where $g_{ij}(s)$ is an element of the plant transfer function matrix. He concludes that if the steady state (s=jw=0) value is near unity then interaction will cause poor control. Also, at frequencies when the magnitude is close to one and the phase is negative, the interaction leads to poor control. Good control can be achieved when his interaction quotient is constant and negative.

Davison (12, 13) devised measures of interaction that consider dynamic interactions. One measure is the performance index is

$$J_i = \max \int y_i(t)dt \text{ for } i = 1, \ldots, n$$

$$\text{for all } X(0) \text{ such that } \|x(0)\| = 1 \tag{2.3}$$

which is used to calculate the Interaction Index

$$I^i(K_1, K_2, \ldots) = \frac{J(i)^* - J(i)}{J(i)} \tag{2.4}$$

where J(i) is the performance for the controlled system using only the i-th control loop. $J(i)^*$ is the performance index of the system for all control loops operational. K_1, K_2 are controller gains.

In this representation the system must be described in state variable form.

A value of zero for the Interaction Index suggests there are no interaction effects for the loop under study. A negative value indicates that the system can be controlled by adding a control loop. Positive values indicate unfavorable interaction. As the positive index increases, the amount of interaction increases.

Davison (13) proposed an index for non-minimal phase systems:

$$D^r(K_s, K_t, \ldots) = \frac{DT}{DT+TC} \tag{2.5}$$

where DT is the dead time for the process to leave or cross through zero. TC is the dominant time constant in the process. K_s, K_t are the controller gains. The r-th controller is not connected to the system. Non-minimal phase refers to dead time or right hand plane zeros in the system's transfer function. His interaction index based on the non-minimal phase index is the

$$I^r(K_s, K_t, \ldots) = D(K_s,K_t)-D(0,0,0\ldots) \tag{2.6}$$

where r-th controller being studied is tuned out by setting the gain to zero. With values of the index classified as positive, negative, or zero these measures have the same interpretation as previously discussed.

The RGA has a serious shortcoming in that it does not address the dynamic aspects of a process. Witcher and McAvoy (14) expanded the RGA to the frequency domain for 2x2 systems. They demonstrated

that a measure by Nisenfeld and Shultz (15) was equivalent to Rijnsdorp's interaction measure at zero frequency.

Bristol (16) also generalized the RGA in the frequency domain to a general square transfer function matrix. This dynamic RGA is defined analogously to the steady state RGA. Thus, the transfer function is required to have an inverse that is a function of frequency. With this measure, systems with negative relative gain elements at zero frequency correspond to a non-minimum phase transfer function with either a right hand zero or pole in the closed loop transfer function.

RGA Pairing Rules

McAvoy (1) has analyzed steady state RGA extensively and has drawn up pairing rules for interpreting the RGA. His improved pairing rules, compared to original rules, are listed as:

1. Positive elements in the RGA matrix nearest to one should be considered first for pairing.
2. These pairings should be checked with Niederlinski's Theorem which states that for an open loop stable system, the closed loop system resulting from pairing u_1-y_1, u_2-y_2, . . . u_n-y_n is unstable if

$$\frac{|K|}{\prod_{i=1}^{n} K_{ii}} < 0 \qquad (2.7)$$

where $|K|$ is the determinant of the process steady state

gain matrix. If this condition is encountered, another positive pairing should be chosen with values closest to one.

3. Negative pairings should be avoided whenever possible.

Additional Guidelines on Interpreting the RGA

McAvoy also provides some secondary rules for interpreting the RGA.

1. Decoupling should be considered when a relative gain element (λ) is outside the range of $0.67<\lambda<1.50$ to avoid figure-8-interaction (Figure 2.1). Decoupling can be achieved by detuning loops, linear decoupling, or by changing variables. Linear decoupling involves linear combinations of the inputs or the outputs. Changing variables means that simple functions are formed of the original manipulated or controlled variables.
2. For $\lambda > 25$, alternative control strategies should be weighed over the RGA Single-Input-Single Output design.
3. For any value of relative gain, significant one-way interaction may occur. For 2x2 systems, McAvoy suggests that the criterion $|\frac{K_{ij}}{K_{ii}}| < \frac{1}{3}$ for all i,j $i \neq j$ be checked for a properly scaled gain matrix. All manipulated variables are properly scaled if they are between zero and one. If $|\frac{K_{ij}}{K_{ii}}| > \frac{1}{3}$, then one-way decoupling is advised.
4. To avoid implementation problems when using linear decouplers, the condition number can be determined from

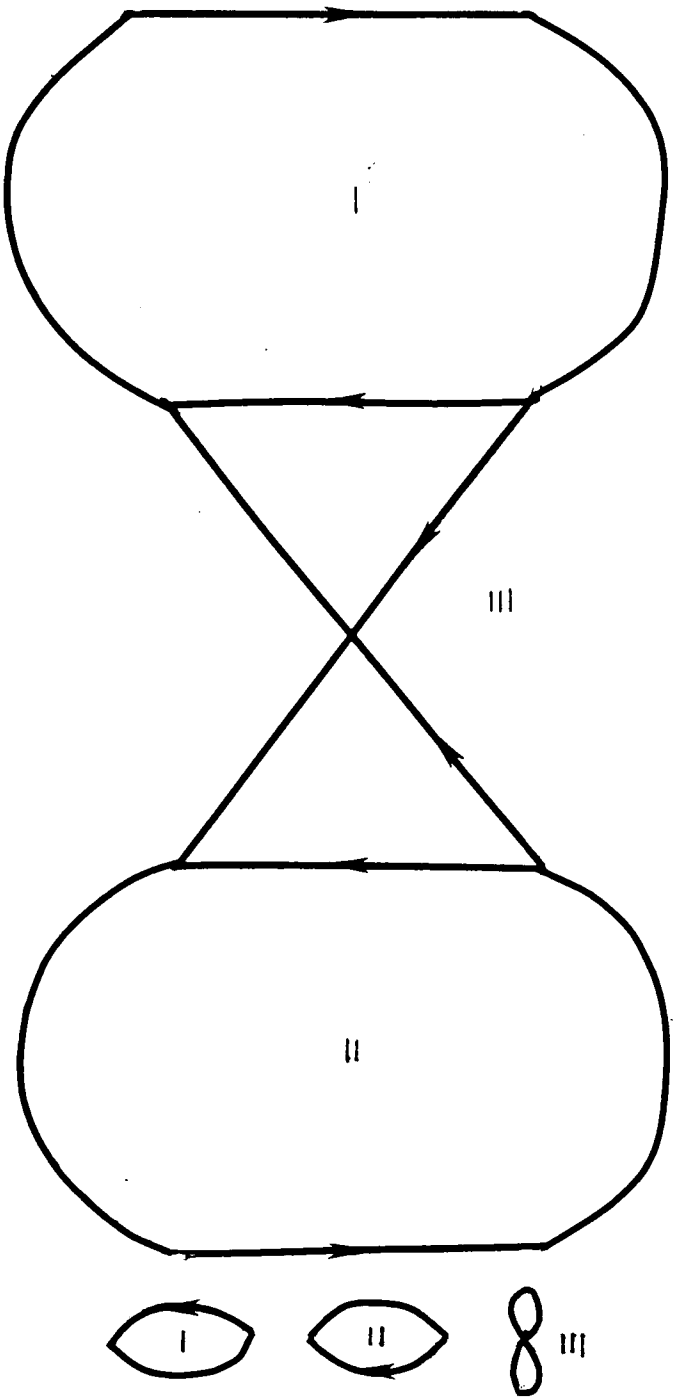


Figure 2.1. Figure -8- Type Interaction.

the SVD. McAvoy recommends that a condition number $\kappa \geq 50$ suggests that the given system is almost singular and decoupling will not be feasible.

In summary, the most common methods of modern multivariable control are listed. This work will involve noninteracting control strategies. Interaction was defined and guidelines were set out for the relative gain array.

CHAPTER 3

MATHEMATICAL BACKGROUND

The singular value decomposition can be described by the decomposition of an mxn matrix A into two orthogonal matrices and one diagonal matrix as follows (17):

$$\text{For } m > n$$
$$A = U \Sigma V^T \tag{3.1}$$

where A is an mxn matrix, U is an mxm orthogonal matrix, V is an nxn orthogonal matrix, Σ is $\begin{bmatrix} S & 0 \\ 0 & 0 \end{bmatrix}$ is an mxn matrix with $S = \text{diag}(\sigma_1, \sigma_2, \ldots, \sigma_n)$ and $\sigma_1 > \sigma_2 > \sigma_3 > \ldots > \sigma_n > 0$.

The singular values of A are the elements $\sigma_1, \sigma_2, \ldots, \sigma_n$. Singular values are defined as the positive square roots of the eigenvalues of A^TA. The columns of U, an orthogonal matrix, are called the left singular vectors while the columns of V, an orthogonal matrix, are the right singular vectors. The SVD exists for all matrices. The rank of A is equal to the number of its nonzero singular values. For a nonsquare matrix, the rank of A is equal to the min(m,n).

Vector and Matrix Norms

Vector and matrix norms are used with SVD to provide bounds relating manipulated variables to measured variables (18). The notion of vector norms will be described and extended to matrix norms. The purpose of vector norms is to determine the length of a

vector. Probably the most common vector measure is the Euclidean or l_2 -norm defined as

$$|| x ||_2 = (\Sigma|x_i|^2)^{\frac{1}{2}} \tag{3.2}$$

In the general case the sum of the magnitude of each element of the vector is raised to the p-th power and then the quantity is raised to the $\frac{1}{p}$ power in the following form:

$$|| x ||_p = (\Sigma|x_i|^p)^{\frac{1}{p}} \tag{3.3}$$

Besides the l_2 -norm, two other easily computed norms are the l_1- and l_∞ - norms. They are characterized as

$$|| x ||_1 = \Sigma|x_i| \qquad l_1 \text{ - norm} \tag{3.4}$$

$$|| x ||_\infty = \max |x_i| \qquad l_\infty \text{ - norm} \tag{3.5}$$

Some of the properties of these vector norms are

$$|| ax || = |a| \; || x || \tag{3.6a}$$

$$|| x + y || \leq || x || + || y || \tag{3.6b}$$

$$|| x || > 0 \text{ for all } x \neq 0 \text{ and } || x || = 0 \text{ for } x = 0 \tag{3.6c}$$

Matrix norms have properties similar to those of vector norms as well as a consistency requirement when matrices are multipled together. These properties for a matrix norm include

$$|| A || > 0 \text{ for all } A \neq 0 \text{ and } || A || = 0 \text{ iff } A = 0 \tag{3.7a}$$

$$|| bA || = |b| \; || A || \text{ for any scalar } b \tag{3.7b}$$

$$|| A + B || \leq || A || + || B || \tag{3.7c}$$

$$|| AB || \leq || A || \;\; || B || \tag{3.7d}$$

Norms of matrices are called induced, natural, or operator when they are based on the vector l-p norm and satisfy the above properties. They are defined as

$$|| A ||_p = \sup_{x \neq 0} \frac{|| Ax ||_p}{|| x ||_p} \quad \text{for all } x \tag{3.8}$$

where sup is the supremum or the least upper bound on $||Ax||$.

The one, two, infinity, and Euclidean matrix norms can be

$$|| A ||_1 = \max_j \left(\sum_{i=1}^{n} | a_{ij} | \right) \tag{3.9}$$

$$|| A ||_\infty = \max_i \left(\sum_{j=1}^{n} | a_{ij} | \right) \tag{3.10}$$

$$|| A ||_E = \left(\sum_{i,j}^{n} | a_{ij} |^2 \right)^{0.5} \tag{3.11}$$

$$|| A ||_2 = (\lambda \max A^T A)^{1/2} \tag{3.12}$$

It is important to note, that although the Euclidean norm is equivalent to the l_2 - norm for vectors, the Euclidean matrix norm is different from the l_2 - matrix norm. For example, the Euclidean norm for the identity matrix would be $\sqrt{2}$ while the l_2 - matrix norm is 2 for this case.

Properties of SVD

SVD is generally recognized as a reliable numerical method for determining the rank of a matrix (19). For finite precision, the zero threshold value below which singular values are assumed zero can be chosen. The value of the smallest singular value indicates how far a matrix is from matrices of lower rank.

The rank of a matrix indicates how close to singular it is. Thus, the rank shows how close the rows or columns of A are to being linearly dependent.

Condition number is defined formally as

$$\kappa(A) = || A || \; || A^{-1} || \qquad (3.13)$$

For square steady state gain matrices, the condition number is the largest singular value divided by the smallest nonzero singular value. Mathematically, the condition number is useful in examining the linear system Ax=b to show how errors in A and b be can be magnified in obtaining the solution. Generally the larger the condition number, the more nearly singular is the matrix A and the solution x becomes more sensitive to errors in A and b.

A third property of the SVD depends on the smallest singular value. The magnitude of the smallest singular value (σ_*) is a measure of the minimum distance to the nearest singular matrix. Therefore, it is a measure of the invertibility of the system. The size of the smallest singular value provides information about potential problems when implementing feedback control (20, 21).

Four Fundamental Subspaces

Four subspaces can be used to describe a matrix that is contained in a linear equation of the form y=Ax. In a control context this equation states a matrix A is an operator that maps x onto y so matrix A relates inputs and outputs. The subspaces are Im(A), Im($A^{\perp}$), Ker(A), and Ker($A^{\perp}$) where $A^{\perp}$ is a matrix orthogonal to matrix A. Im refers to the image of the matrix, and Ker is the nullspace or the Kernel of the matrix. The Kernel is associated with the inputs, while the image relates to the outputs. The following two equations relate the four fundamental subspaces as follows:

$$\begin{aligned} \mathrm{Im}(A^{\perp}) &= \mathrm{Ker}\,(A^{T}) \\ \mathrm{Im}(A^{T}) &= \mathrm{Ker}\,(A^{\perp}) \end{aligned} \tag{3.14}$$

where A^{T} is the transpose of A.

These subspaces are related to the SVD of any matrix (22, 23)

$$\begin{aligned} A &= U\Sigma V^{T} \\ A &= [U_1 U_2] \begin{bmatrix} S & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \end{bmatrix} \end{aligned} \tag{3.15}$$

Image (A) represents the columns of U_1 and all the possible outputs of Image ($A^{\perp}$) symbolizes the columns of U_2 and contains all the values of y which are not possible outputs. It is the orthogonal complement of image A. Kernel (A) contains all inputs x that map into the nullspace of y. It represents columns of V_2. Kernel (A^{T})

is the orthogonal complement of Kernel (A) that contains all the inputs x which lead to nonzero values of the output y. It represents the columns of V_1. Refer to Figure 3.1.

Orthogonal and Diagonal Matrices

Since SVD involves the decomposition of a matrix into three component matrices, one diagonal and two orthogonal matrices, some of the important characteristics of these types of matrices will be described.

An orthogonal matrix P is a nonsingular, real, square matrix with orthonormal columns and rows (17). It has the following properties:

1. $P^TP=I$
2. $PP^T=I$

The first property implies that columns of P form an orthonormal set of vectors while the second property states that the rows form an orthonormal set. An orthonormal set of vectors means that they are both normalized and mutually orthogonal.

3. If P is orthogonal, the det P=1. P is called properly orthogonal if its determinant is +1 and improperly orthogonal if its determinant is -1.
4. If two matrices P and Q are orthogonal, then their product is also orthogonal.
5. An orthogonal matrix preserves both angles and lengths when it is used for defining a linear transformation. If P is orthonormal, it has all the properties of an

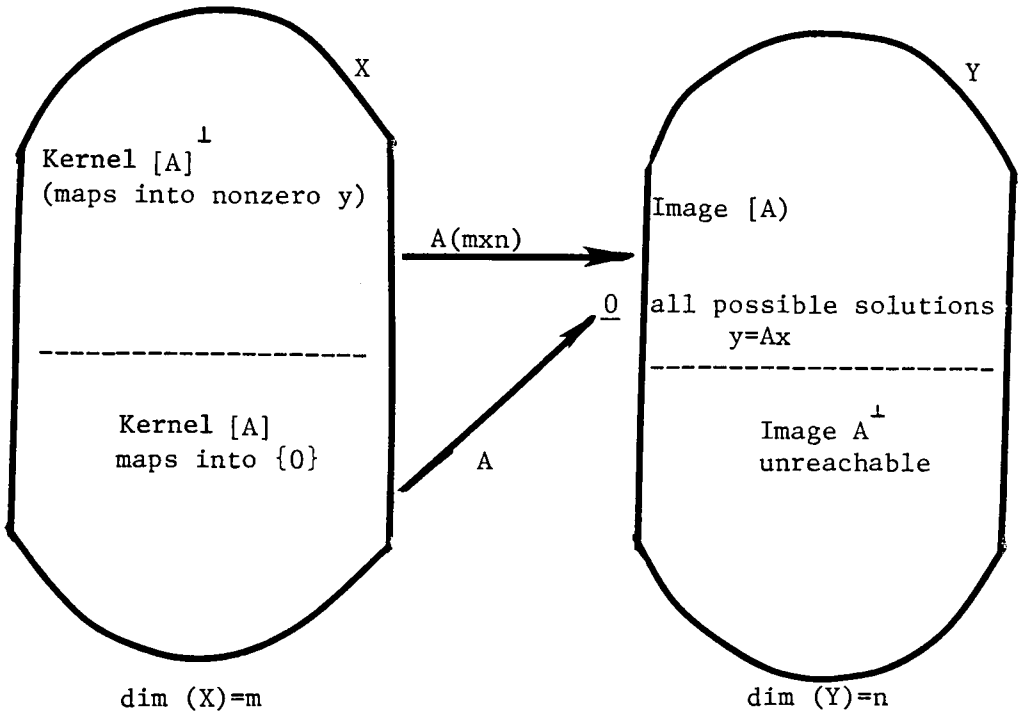


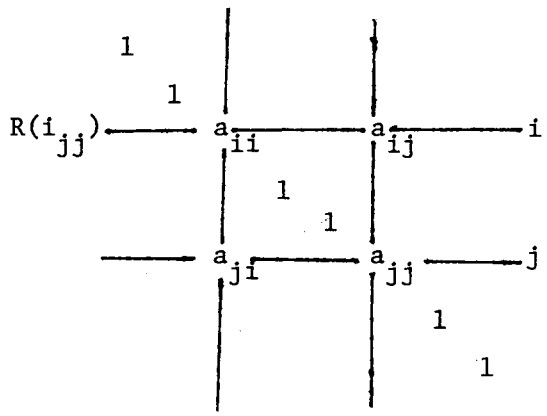
Figure 3.1. Four Fundamental Subspaces that Map X onto Y.

orthogonal matrix and it also satisfies for all x the following properties:

$$|| Px ||_2 = || x ||_2 \tag{3.16a}$$

$$|| P ||_2 = 1 \tag{3.16b}$$

Orthogonal matrices provide computational advantages when used in conjunction with the l_2 - norm. For example, the inverse of an orthogonal matrix is its transpose. "The l_2 - norm for an orthogonal matrix is one, and the modulus of all real and imaginary eigenvalues of the matrix is one." (9) One type of orthogonal matrix is the elementary rotation matrix with the general form (23)



(3.17)

where $a_{ij} = \cos$

$a_{ij} = -\sin$

$a_{ji} = \sin$

$a_{jj} = \cos$

This is generally called a Givens reflection. If $a_{ij} = \sin\theta$ and $a_{ji} = -\sin\theta$ with $a_{ii} = a_{jj} = \cos\theta$, the form is a Givens rotation.

A diagonal matrix is defined as a matrix whose off-diagonal entries are all zero. Pre-multiplying a mxn matrix A by a diagonal matrix Sm involves multiplying the i-th row of A by the i-th diagonal element of S_m. Post-multiplying A by Sn, a diagonal matrix, involves multiplying the j-th column of A by the j-th diagonal element of Sn. Thus Sm and Sn are called scaling matrices.

$$\begin{bmatrix} S_{m1} & 0 & 0 \\ 0 & S_{m2} & 0 \\ 0 & 0 & S_{m3} \end{bmatrix} \begin{bmatrix} A_{11} & A_{21} \\ A_{12} & A_{22} \\ A_{13} & A_{23} \end{bmatrix} \begin{bmatrix} S_{n1} & 0 \\ 0 & S_{n2} \end{bmatrix} = \tag{3.18}$$

$$\begin{bmatrix} S_{m1} A_{11} S_{n1} & S_{m1} A_{21} S_{n2} \\ S_{m2} A_{12} S_{n1} & S_{m2} A_{22} S_{n2} \\ S_{m3} A_{13} S_{n1} & S_{m3} A_{23} S_{n2} \end{bmatrix}$$

Diagonal matrices will be employed in scaling procedures for both square and nonsquare matrices. Although matrix multiplication is generally not commutative, it is for the case for multiplying diagonal matrices. The order of multiplication of diagonal matrices is not important.

Pseudoinverse

The pseudoinverse or generalized inverse is useful for solving problems that are underdetermined, singular, or overdetermined. The pseudoinverse is defined as

$$A^{+} = V\Sigma^{+}U^{T} \tag{3.19}$$

where Σ^{+} is $\begin{bmatrix} \frac{1}{\sigma} & \sigma_i > 0 \\ 0 & \sigma_i = 0 \end{bmatrix}$

A^{+} is the pseudoinverse

The solution found with the pseudoinverse is the minimum norm of all possible solutions that minimizes the norm $||Ax-b||$. The pseudoinverse is identical to the standard inverse for full rank, square matrices, and the solution represents the unique solution for a set of linear equations. For the underdetermined case, the solution calculated using the pseudoinverse is one vector or solution out of infinitely many with a minimum norm. The overdetermined case can be useful in solving the pseudoinverse that yields the least square estimate for the solution. See Figure 3.2.

A few of the properties of the pseudoinverse are listed as follows (24):

1. A^{+} is an nxm matrix. A^{+} originates with the b vector in R^{m} and produces a vector x in R^{n}.
2. The column space of A^{+} is the row space of A and vice versa. Therefore, the rank A^{+}=rank A.
3. Usually $AA^{+}=I$ because A may not have a right-inverse. However, AA^{+} always equals the projection of P onto the column space

$$AA^{+}b = A\bar{x} = Pb = p \tag{3.20}$$

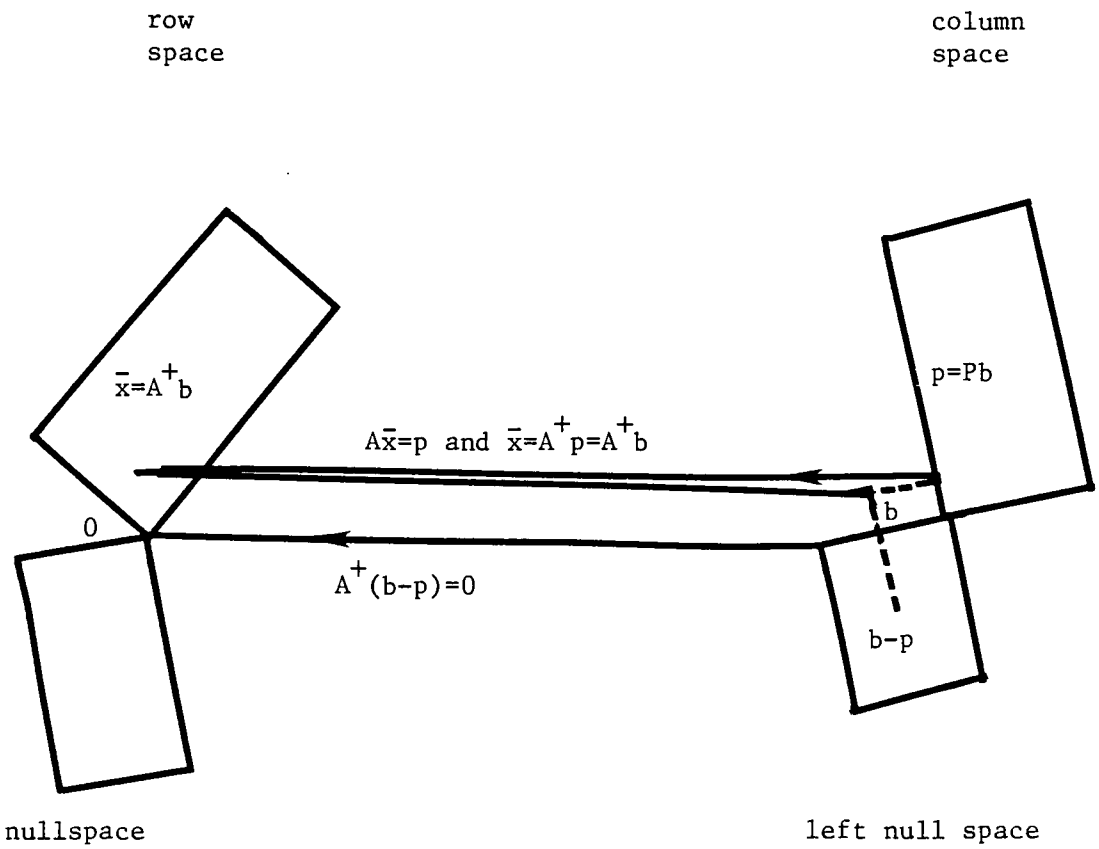
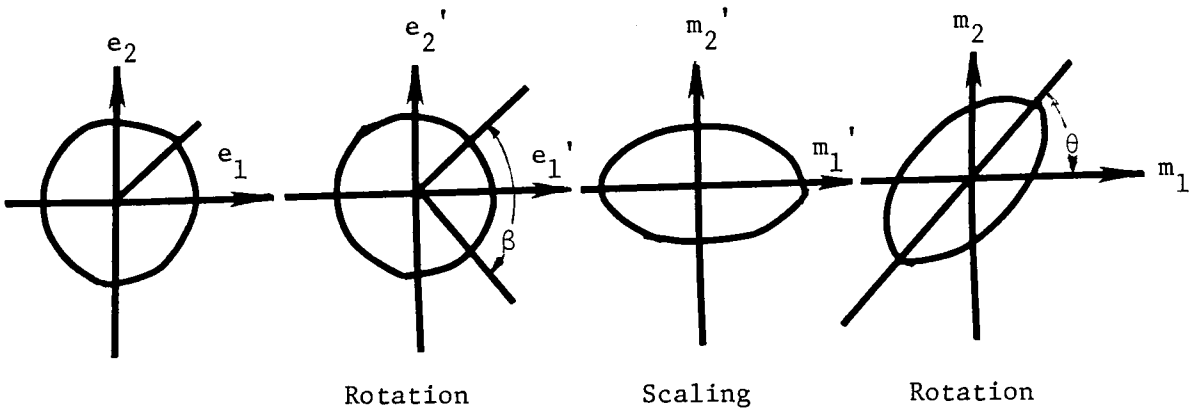


Figure 3.2. The Relationship of a Matrix A and Its Pseudoinverse A^{+}.

Geometrical Interpretation of SVD

The manipulated variables for a specific steady state gain matrix is considered to be a unit ball of inputs or the set of all possible unit length vectors for the process (25, 26, 27). A single vector mapped through a matrix results in a transformed vector. For mapping a unit ball of vectors, the mapping will be an ellipsoid. The ellipsoid will range from well-shaped representing a well-conditioned process (low condition number) to a line, a degenerate shape representing a poorly conditioned process (high condition number). The length of the major axis of the ellipse is related to the largest singular value while the length of the minor axis corresponds to the smallest singular value. Size, shape, and orientation of the ellipse are determined by the original matrix. Geometrical linear transformations of orthogonal matrices can be considered to be simple rotations in space. The component matrices in the SVD may be viewed as three transformations for the input vector. The orthogonal matrix V^T gives a rotation of the input. The second matrix Σ scales the rotated vector. The third matrix U rotates the output. This transformation is illustrated in Figure 3.3. The coordinate system for U and the coordinate system for V^T are both orthogonal. However, one of these coordinate systems may be rotated with respect to the other.



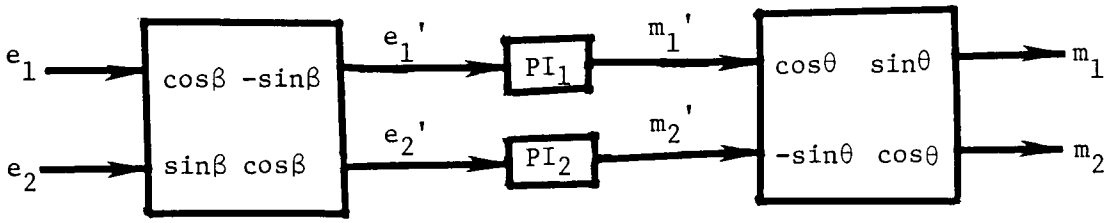


Figure 3.3. Geometric Interpretation of Rotations and Scaling for a 2x2 SVA Controller.

Condition Number

Condition number is defined for any nonsingular mxn matrix A as (23).

$$\kappa_\beta(A) = || A || \; || A^{-1} ||_\beta \qquad (3.21)$$

where $||\kappa||_\beta$ is any matrix norm

In the case of the l_2 - matrix norm, the condition number is usually defined as

$$\kappa_2(A) = \frac{\sqrt{\mu_1}}{\sqrt{\mu_n}} = \frac{\sigma_1}{\sigma_n} \qquad (3.22)$$

where μ_1 is the largest eigenvalue

μ_n is the smallest nonzero eigenvalue

σ_1 is the largest singular value

σ_n is the smallest nonsingular singular value

The condition number provides a bound for the relative error in x due to perturbation in A or b. Condition number implies how close to singular a matrix is or how much a matrix has to be perturbed to be singular. A small value of condition number implies that perturbations in A or b introduce only small perturbations in x. In this case, the problem is considered well-conditioned. However, when the condition number is large (i.e. K $\geq$50 (2)), small perturbations in A or b can cause large perturbations in x for linear systems. In this case, the systems are considered ill-conditioned. In application to process control, "the condition

number quantifies the sensitivity of the system with respect to uncertainties in the matrix (modeling errors)" (22).

Other measures of conditioning are weaker than the l_2 - condition number. For example, the determinant is a poor measure of conditioning since it depends on scaling as well as the order of the matrix and is limited to square systems (24).

When A^{-1} does not exist, then condition number is considered to be infinite since the SVD will have a zero singular value. A perfectly conditioned system has a condition number of 1 since

$$1 = || I || = || AA^{-1} || \leq || A || \; || A^{-1} || = \kappa(A) \qquad (3.23)$$

Also condition number yields information about the rank since if at least one singular value is zero then the matrix is rank deficient.

When conditioning is considered in relation to process control, the magnitude of the condition number of the process steady state gain matrix is related to the difficulty of the control problem. This difficulty is associated with the magnitude of change needed in the manipulated variables to move the output variables in various directions. If a process has a low condition number, the outputs can be moved in any direction with the same magnitude of manipulated variables. A process with a higher condition number is more difficult to change or move in some output directions. The process output is easiest to change in a direction associated with the largest singular value by making an input change in the input

direction associated with the largest singular value. A process output is difficult to change in the output direction that is associated with the smallest singular value. A direct relationship between conditioning and ease of control is a too simplistic concept. Other factors, such as interaction, should be considered and judgments have to be made concerning the relative weights of various factors for the given system.

The condition number applied to least square problems is somewhat more complicated than that for square systems (28). For a least squares system, the condition number consists of one to three factors depending on whether or not the vector b lies in the row space of A which has linearly independent columns.

$$\frac{||x-\bar{x}||_2}{||x||_2} < 2\kappa \frac{||E_1||_2}{||A||_2} + 4\kappa^2 \frac{||E_2||_2}{||A||_2} \frac{||b_2||_2}{||b_1||_2} + 8\kappa^3 \frac{||E||_2^2}{||A||_2} \tag{3.24}$$

where $K = ||A||_2 \ ||A^+||_2$. E_1 (b_1) and $E_2(b_2)$ are projections of E(b) onto the row space of A and onto the orthogonal complement of the row space of A.

Numerical Computation of the SVD

Computation of the singular value decomposition involves robust numerical procedures (29). Golub and Kahan devised a computational method for calculating SVD (30). An ALGOL program was translated by Golub and Reinsch for computing the SVD in FORTRAN (31). A somewhat modified version of the SVD routine was used in EISPACK (32).

The QR algorithm is used extensively in the computation of SVD (23, 33, 34). This algorithm is used in subroutine packages such as IMSL, EISPAC, and LINPACK. Two major steps are involved in computing the SVD. First, the given matrix A is transformed into an upper Hessenberg, bidiagonal or tridiagonal form. The given matrix is transformed into an upper bidiagonal form by Householder transformations. All the elements of the first column are zeroed except for the diagonal element. Then elements of the first row are zeroed except for the diagonal and superdiagonal elements. After applying orthogonal matrices, a matrix remains with only nonzero diagonal and superdiagonal elements.

The second step, based on a variation of the QR algorithm of Frances (35), involves finding the singular values of this bidiagonal matrix. This procedure proceeds iteratively as a series of applications of orthogonal rotations and shifts to zero the superdiagonal row.

Rutishauser proposed the algorithm using the LR transformation to solve the algebraic eigenvalue problem for symmetric band matrices (36). A sequence of matrices is formed in which each matrix is decomposed into a lower triangular (L_i) and upper triangular forms.

$$A_i = L_i R_i \tag{3.25}$$

These L_i and R_i matrices are multiplied in reverse order to form

$$A_{i+1} = R_i L_i \tag{3.26}$$

$$A_{i+1} = L_i^{-1} A_i L_i \tag{3.27}$$

Assuming L_i are nonsingular, A_{i+1} and A_i are related by a similarity transformation. This algorithm is very difficult to implement. It is not always numerically stable because of the triangular decomposition and involves a considerable amount of computation. Francis (35, 37) and Kubalovskaya (38) suggest that the L matrix is replaced by a unitary matrix which results in the QR algorithm.

$$A_i = Q_i R_i \tag{3.28}$$

$$A_{i+1} = Q_i^{-1} A_i Q_i \tag{3.29a}$$

$$= R_i Q_i \text{ for } i=1, 2, 3, \ldots \tag{3.29b}$$

Implementation of the QR algorithm is only possible when the matrix is pretreated to be in upper Hessenberg, bidiagonal, or tridiagonal form. Inclusion of origin shifts in this algorithm facilitates convergence of the algorithm and avoids convergence problems if distinct eigenvalues have the same modulus.

Typically, SVD programs are written for the case when m>n. For n>m, the transpose of the original matrix can be taken and the roles of U and V are interchanged from that of the SVD of the original

matrix. For the SVD programs in LINCON, IMSL, and EISPACK, the SVD for the underdetermined case can also be found by squaring off the original matrix with zeros. Least squares problems can be solved more efficiently with LINPACK using SQRDC and SQRSL. SQRDC performs the QR decomposition that is then acted upon by SQRSL which provides various factors for the solution of the least squares problem. For m>n, the routine DSVDC does the SVD so that the form of the diagonal matrix is

$$U^T AV = \begin{bmatrix} \Sigma \\ 0 \end{bmatrix} \tag{3.30}$$

In the case of m<n, this subroutine performs the SVD so that the matrix is

$$U^T AV = [\Sigma \ 0] \tag{3.31}$$

Matrix and vector properties related to SVD have been introduced in this chapter. Most of the theorems that will be presented in Chapter 5 can be proved using the properties of orthogonal and diagonal matrices. The effect of conditioning on square process gain matrices of low order will be discussed in Chapter 6. The geometrical interpretation of SVD can be used in the development of the singular value analysis controller (9, 18).

CHAPTER 4

LITERATURE REVIEW OF MATRIX SCALING

The scaling problem is discussed widely in mathematics because of the limitations on accuracy in numerical algorithms that are due to machine computations. A properly scaled process gain matrix is one in which information can be obtained about controlling the process. Scaling in this thesis involves scaling the process gain matrix for a minimum condition. This scaled gain matrix needs to reflect the physical character of the original unscaled gain matrix as well as be mathematically related to the original gain matrix. Scaling can be defined as the "variable transformation (that) converts the variables from units that typically reflect the physical nature of the problem to units that display certain desirable properties during the minimization process" (39). Bauer stressed the importance of a minimal condition number to decrease error in numerical computations (40).

Bauer defined condition number of a matrix as the "ratio of its least upper and greatest lower bounds with respect to a pair of matrix norms" (41). For studying a process steady state matrix, the l_2 -norm does not exist. Various numerical and empirical methods are suggested for optimizing the scaling of a matrix. For instance, A. van der Sluis (42) recommends row and/or column equilibration. Analytical optimal scaling procedures exist for the l_1 - and l_∞ - norms for simultaneous row and column scaling to minimize condition

number for irreducible matrices (43). A matrix A is called fully indecomposable if there do not exist permutation matrices P and Q of the form

$$PAQ = \begin{bmatrix} R & S \\ 0 & T \end{bmatrix} \tag{4.1}$$

No universally agreed upon criteria exist for a "well-scaled" matrix. J. A. Tomlin considered a matrix with elements of nearly the same magnitude to be well-scaled. He discussed three empirical scaling techniques that were applied for linear programming problems. These empirical scaling techniques include:

1. Equilibration

   All the elements in each row (or column) are divided by the largest element in each row (or column). In the scaled matrix, the largest element of each row (or column) is of order unity.

2. Geometric Mean

   The square root of the product of the maximum element times the minimum element in each row (or column) is calculated. Each original element is divided by this factor corresponding to a given row (or column). Row followed by column scaling refers to row factors being taken from the original matrix, column factors are then obtained from the row scaled matrix, similarly, column followed by row scaling involves column scaling factors

being taken from the original matrix. Row scaling factors are then obtained from the column scaled matrix.

3. Arithmetic Mean

   The elements of each row are divided by the arithmetic mean for that row. Elements for each column are treated similarly.

Refer to Table 4.1 for the algorithm of each of these empirical scaling techniques.

The order of the operation in each scaling technique may affect the condition number. Also a combination of these scaling techniques could be used. Equilibration can be appended to matrices previously geometric mean scaled or arithmetic mean scaled. By varying the order of the row or column scaling, 14 scalings are possible with the three techniques of equilibration, geometric mean, and arithmetic mean scaling. These combinations of empirical scaling techniques are listed as follows:

1. Arithmetic mean RC;
2. Arithmetic mean CR;
3. Equilibration RC;
4. Equilibration CR;
5. Geometric mean RC;
6. Geometric mean CR;
7. Geometric mean RC, Arithmetic mean RC;
8. Geometric mean RC, Arithmetic mean CR;
9. Geometric mean CR, Arithmetic mean RC;

Table 4.1. Types of Empirical Scaling Techniques

Method	Algorithm
1. Arithmetic mean	$\frac{\text{matrix element}}{\text{arithmetic mean}}$
2. Equilibration	$\frac{\text{matrix element}}{\text{maximum element}}$
3. Geometric mean	$\frac{\text{matrix element}}{\sqrt{\lvert\text{max. element}\rvert \times \lvert\text{min. element}\rvert}}$

10. Geometric mean CR, Arithmetic mean CR;
11. Geometric mean RC, Equilibration RC;
12. Geometric mean RC, Equilibration CR;
13. Geometric mean CR, Equilibration RC;
14. Geometric mean CR, Equilibration CR;

where RC is row followed by column scaling,

CR is column followed by row scaling.

Bauer derives upper bounds of the minimal condition number obtained by optimal scaling (41). For the l_2 -norm, optimal scaling is characterized by S_1AS_2 and $(S_1AS_2)^{-1}$ having equal row sums of absolute values. Using the representation for a linear equation Ax=b, row scaling of this system is the transformation of A by a diagonal matrix S_1 on the left while column scaling is the transformation of A by a diagonal matrix S_2 on the right.

The l_1 - and l_∞ - norms can be used to bound the l_2 -norm. Golub and van Loan (44) list several useful relations between these norms for $A \epsilon R^{mxn}$

$$|| A ||_2 < \sqrt{|| A ||_1 || A ||_\infty} \tag{4.2}$$

$$\frac{1}{\sqrt{n}} || A ||_\infty < || A ||_2 < \sqrt{m} || A ||_\infty \tag{4.3}$$

$$\frac{1}{\sqrt{m}} || A ||_1 < || A ||_2 < \sqrt{n} || A ||_1 \tag{4.4}$$

where m > n

Various empirical scalings can vary the form of a matrix without changing the nature of the problem in the context of linear programming. Therefore, numerous attempts have been made to determine techniques to well-scale a matrix that are also efficient (e.g. Bauer (41, 45) Tomlin (46), Curtis and Reid (47), Fulkerson and Wolfe (48), Haming (49), Orchard-Hays (50), and Saunders and Schneider (51)). Fulkerson and Wolfe (48) examined the ratios of nonzero elements in a nonnegative matrix. They defined a measure of goodness of scaling for a matrix using a maximum ratio. Saunders and Schneider (51) extended the work of Fulkerson and Wolfe for matrices with negative elements so that absolute values are considered in scaling. Curtis and Reid (47) applied a least square approach rather than a maximal element approach to scaling.

Most scaling techniques are designed for only square systems. Rothblum and Schneider (52) have developed a concept for scaling nonnegative, rectangular matrix F

$$F_{sc} = EFG \qquad (4.5)$$

where $F \neq 0$

E and G are nonsingular, nonnegative, diagonal matrices of the appropriate dimension. F_{sc} is a scaled nonsingular, nonnegative matrix.

They defined a measure of goodness of scaling as the maximum ratio of nonzero elements of the given matrix. A minimal value of the goodness of scaling can be determined in terms of cyclic

products associated with a directed graph corresponding to the matrix for the set of all scalings of a particular matrix. In other words, the scaling problem is converted into a linear problem in which extreme points of the polytope are characterized.

$$\alpha(S) = \max_{L_{ij} > 0, L_{kp} > 0} \frac{S_i\, L_{ij}\, S_j^{-1}}{S_k\, L_{kp}\, S_p^{-1}} \tag{4.6}$$

where α is a scalar

$S \varepsilon S_+^n$ (set of square, nonnegative, nonsingular, diagonal matrices)

Analogously, the measure for a rectangular matrix F of dimension $n_1 x n_2$ is

$$\beta = \inf_{E \varepsilon S_+^{n_1},\ G \varepsilon S_+^{n_2}} \max_{F_{ij} > 0,\ F_{kp} > 0} \frac{E_i\, F_{ij}\, G_j^{-1}}{E_k\, F_{kp}\, G_p^{-1}} \tag{4.7}$$

where β is a scalar

inf (infinum) is the greatest lower bound

Rosenblum and Schneider developed a method for characterizing the scalars α and β for both square and rectangular matrices.

David Nash (53) generalized the Eckard-Young Principal Axis Theorem which can be applied for scaling rectangular matrices. This theorem states that "If $\{A_i\}$ is a set of complex mxn matrices such that $A_i A_j^*$ and $A_i^* A_j^*$ are Hermitean for all i and j with i = j, then there exist unitary matrices P and Q such that for each i the matrix PA_iQ is real and diagonal. (A rectangular matrix $[a_{pq}]$ is diagonal if $a_{pq}=0$ when $p \neq q$.)" When there is only one matrix A in the set A_i, classical singular value decomposition results.

The algorithm used in this work geometric mean scaling procedure of the form S_1AS_2 was based on work from Tomlin (46) and Gill (39). Gill recommends that a geometric mean program terminate either after a fixed number of iterations or when the condition number converges in some predefined manner. In the geometric mean scaling program used in this thesis, the test for convergence involves comparing condition numbers for the previous and present values of row followed by column scaling or column followed by row scaling or both.

For linear systems, the problem of scaling involves replacing a given matrix A with S_1AS_2 such that S_1 and S_2 are diagonal matrices that perform row and column scaling, respectively, on A (54). Restating the scaling problem involves showing the equivalence of the following two equations:

$$Ax = b \tag{4.7}$$

$$(S_1AS_2)(S_2^{-1}x) = S_1b \tag{4.8}$$

Although two mathematically equivalent formulations of a problem are truly identical, the numerical solutions may not be the same because of rounding errors. Several sources of error in the given matrix include measurement error, rounding error or truncation from computation, and subsequent computation performed on A when the computed solution is the exact solution of a slightly perturbed form involving A+E. The matrix E consists of error elements for each of

the corresponding elements of A. The computed solution depends on how close to singular the original matrix A is.

A system with a large condition number is considered ill-conditioned with respect to inversion if $\kappa(A)$ is 10^{K} when the linear solution is calculated in decimal arithmetic (t-digit) it will be accurate to only t-k figures. For the condition number to be significant, the elements of the error matrix corresponding to the given system must be about the same size. Any useful scaling strategy should involve estimating the absolute size of the errors in the given matrix. Then this matrix should be scaled so that these estimated errors are nearly the same magnitude as possible.

Various types of scaling procedures exist to precondition a given matrix so that the elements are within at most a few orders of magnitude (55). Even if an algorithm is numerically stable, rounding errors of larger elements can become so large when compared to smaller elements that the final calculated results are corrupted. Some of these scaling techniques for preconditioning systems of linear equations are optimal scaling (41) and equilibration (56, 57), balancing for the eigenvalue problem (58), and methods for the generalized eigenvalue problem (55). Much of the variation in types of scaling results from methods and purposes of scaling depending on the particular problem. The routine BALANC is a numerically reliable algorithm contained in the software package EISPACK for state scaling by minimizing a 1_{∞} - or row sum norm for $A=D_x^{-1} AD_x$. An analytical solution for optimal scaling that minimizes the

condition number associated with the l_2 -norm is only possible when both A and A^{-1} have a checkerboard sign distribution (41). Williams (55) has developed a scaling algorithm that minimizes the Gerschgorin discs of a given matrix that has tight eigenvalue bounds that are computationally inexpensive. This method can be used when scaling over a frequency range to place the poles of the system in the unit circle.

Bonvin and Mellichamp (59) have devised a scaling method based on dynamic models of large-scale systems. Their method can be used to analyze systems with linear or nonlinear, lumped or distributed parameter models in which the input and state variables are analyzed. The purpose of their scaling technique is to "normalize state and input variables so as to eliminate the arbitrariness associated with the choice of units but to keep the relative strength of input/state coupling." They argue that scaling based on steady-state values is insufficient because it does not include the range of variables. Preconditioning variables on the basis of their ranges is also inadequate since arbitrary gain characteristics are introduced and it is subjective. Bonvin and Mellichamp propose scaling input and state variables on a common basis with respect to physical arguments. Their scaling factors are valid only in the region near an operating point since they depend implicitly on steady state values.

Lau (21, 60) has proposed a scaling method which he argues reflects the physical basis of the system. He developed an

algorithm to reformulate the scaling problem into two minimization problems, namely variable normalization and equation equilibration. He observed that linear dependency is related to condition number. As the linear dependency between rows and columns increases, the matrix is more likely to be singular and have a higher condition number. He calculated scaling factors to minimize an index of both linear row and column dependence which therefore minimizes condition number. First the variables are normalized to limit variability of the system's variables. Then the equations are scaled so that the maximum term in each equation is of comparable magnitude. This method is not dependent on the order of scaling so it overcomes a deficiency inherent in the empirical scaling methods.

His scaling method will be summarized by the following steps (21):

For any nxm matrix K

1. The measure of linear dependency between any two rows is

$$M_{ij} = \frac{\left[\sum_{k=1}^{m} k_{ik}\, k_{jk}\right]^2}{\sum_{k=1}^{m} k_{ik}^2 \sum_{k=1}^{n} k_{jk}^2} = \frac{\langle r_j, r_j \rangle^2}{\langle r_i, r_i \rangle \langle r_j, r_j \rangle} \qquad (4.9)$$

2. The total measure of row linear dependency will be given as

$$M = \sum_{i=2}^{n} \sum_{j=1}^{i-1} M_{ij} \qquad (4.10)$$

3. The original matrix is pre- and post-multiplied by diagonal matrices S_1 and S_2

$$K_{sc} = S_1 \, K \, S_2 \tag{4.11}$$

4. The diagonal matrices S_1 and S_2 are found such that M is minimized

$$\bar{M} = \sum_{i=2}^{n} \sum_{j=1}^{i=1} \bar{M}_{ij} \tag{4.12}$$

$$\bar{M}_{ij} = \frac{\left[\sum_{k=1}^{m} k_{ik} \, k_{jk} \, S_{2k}^2\right]^2}{\sum_{k=1}^{m} k_{ik} \, S_{2k}^2 \sum_{k=1}^{m} k_{jk} \, S_{2k}^2} = \frac{\langle \bar{r}_i, \bar{r}_j \rangle^2}{\langle \bar{r}_i, \bar{r}_i \rangle \langle \bar{r}_j, \bar{r}_j \rangle} \tag{4.13}$$

$$S_2 = \text{diag} \, (S_{2k}) \tag{4.14}$$

5. Differentiating M with respect to S_{21}

$$\frac{\partial \bar{M}}{\partial S_{21}} = \sum_{i=2}^{n} \sum_{j=1}^{i=1} [2 \langle \bar{r}_i, \bar{r}_j \rangle \langle \bar{r}_i, \bar{r}_i \rangle \langle r_j, r_j \rangle k_{i1} \, k_{j1}$$

$$- \langle r_i, r_j \rangle^2 [\langle \bar{r}_i, \bar{r}_j \rangle k_{j1}^2 + \langle \bar{r}_j, \bar{r}_j \rangle k_{i1}^2] /$$

$$\langle \bar{r}_i, \bar{r}_i \rangle^2 \langle \bar{r}_j, \bar{r}_j \rangle^2 \quad = 0 \text{ for } 1 = 1,2, \ldots, m \tag{4.15}$$

6. A first order numerical method such as Davidson-Fletcher Powell of Steepest Descent can be used to solve the first order system of nonlinear equations.
7. Then a similar procedure is applied to the columns as well as the rows.

The problem of scaling is still unclear as to how it is defined. This thesis examines scaling with the objective of minimizing condition number. The algorithms for empirical scaling techniques have been listed and will be used extensively in Chapter 6. An overview of some of the representative work on scaling since the early 1960s has been presented in the chapter.

CHAPTER 5

MATHEMATICAL DEFINITIONS AND THEOREMS RELATING TO SVD

In order to examine some of the properties of SVD and scaling, the following Definitions (1-10) and Theorems (1-15) from Prasad (1) will be stated as well as several additional theorems (60). Theorems relating to condition number and those classifying steady state gain matrices into three types of matrices will be utilized later. A set of matrices will be denoted by an offset letter while an element of a set will be listed as a capital letter.

Definitions

1. An identity matrix I is a square matrix having positive ones as diagonal elements and zeros as off-diagonal elements. It satisfies the relationship with another square matrix A of the same dimension that AI=IA=A. No units are attached to the elements of an identity matrix.
2. A square matrix will be called J if it consists of positive ones along the secondary diagonal in the positions $(J_{1n}, J_{2(n-1)}, \cdots J_{(n-1)2}, J_{n1})$ and the remaining elements are all zero. Elements of a J matrix do not have units.
3. A set of square matrices is denoted i_{NI} if it has the same form as an identity matrix but it contains signed ones along the diagonal. Units are incorporated in the elements of a matrix in i_{NI} since these matrices refer to

steady state gain matrices. For the case of 2x2 matrices there are four $I_{NI} \in i_{NI}$

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} \tag{5.1}$$

Similarly a set of nxn matrices is denoted j_{NI} if it has the form of a J matrix except with signed ones along the secondary diagonal. Elements of j_{NI} have units. There are four examples of j_{NI} matrices of dimension 2x2

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix} \tag{5.2}$$

4. $\mathcal{U}_{NI}$ and $\mathcal{V}_{NI}$ are a set of nonsingular, square, orthonormal matrices that have only n nonzero elements so that one element in each row and each column has one and only one plus or minus one element. The remaining elements in these matrices are all zeros. A matrix that is an element of one of these sets, $\mathcal{U}_{NI}$ and $\mathcal{V}_{NI}$ will be denoted U_{NI} or V_{NI}, respectively.
5. A square, nonsingular matrix will be called orthogonal ($A^{\perp}$) if $A^T = A^{-1}$.
6. Scaling for mxn matrices will be defined as

$$A_{SC} = S_1 A S_2 \tag{5.3}$$

where S_1 and S_2 are square, diagonal matrices, A_{SC} is of dimension mxn, A is of dimension mxn, S_1 is of dimension mxm, and S_2 is of dimension nxn. S_1 scales the rows of A while S_2 scales the columns of A.

7. S_1 and S_2 can be redefined in terms of the square scaling matrices S_R and S_C, respectively. One element in S_1 and S_2 is set to be one so that both S_R and S_C have one less scaling parameter than S_1 and S_2. For example

$$S_1 = \text{diag}\ (s_1, s_2) \tag{5.4a}$$

$$S_2 = \text{diag}\ (s_3, s_4) \tag{5.4b}$$

$$S_R = \text{diag}\ \left(\frac{s_1}{s_2}\ ,\ 1\right) \tag{5.4c}$$

$$S_c = \text{diag}\ \left(\frac{s_3}{s_4}\ ,\ 1\right) \tag{5.4d}$$

$$K_{SC} = \begin{bmatrix} s_1 & 0 \\ 0 & s_2 \end{bmatrix} \begin{bmatrix} k_{11} & k_{12} \\ k_{21} & k_{22} \end{bmatrix} \begin{bmatrix} s_3 & 0 \\ 0 & s_4 \end{bmatrix} \tag{5.4e}$$

$$K_{SCR} = \begin{bmatrix} S_R & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} k_{11} & k_{12} \\ k_{21} & k_{22} \end{bmatrix} \begin{bmatrix} S_c & 0 \\ 0 & 1 \end{bmatrix} \tag{5.4f}$$

Classification of Square Matrices

The following three types of matrices will be assumed to be nonsingular and square.

8. The matrix A will be considered noninteracting (A_{NI}) if it has only n nonzero elements ordered so that each row and each column has one and only one nonzero element. In each case if A is a process steady state gain matrix, each input will affect one and only one output.
9. A is said to be partially interacting A_{PI} if it has one or more zero elements and at least n + 1 nonzero elements. In this case an input u can affect one or more output(s) y if A represents the steady state gain matrix ($A \varepsilon A_{PI}$).
10. The matrix A will be called full or interacting if it contains only nonzero elements. In this case each input will affect each output.

Any matrices that do not fall into any of the above classifications are singular.

Theorems about SVD, Scaling, and Condition Number

Theorem 1. The SVD of a matrix A is equal to $A=U*\Sigma*V^T$. If rows of A are interchanged, the corresponding rows of U will be interchanged. Also, if the columns of A are interchanged, the corresponding columns of V^T will be interchanged.

Theorem 2. The singular values of A_{NI}, a square noninteracting matrix are the same as the absolute

values of the nonzero elements of A_{NI}.

Theorem 3. The necessary and sufficient condition for A to be an nxn noninteracting matrix, A_{NI}, is that its U and V from the SVD are elements of $\mathcal{U}_{NI}$ and $\mathcal{V}_{NI}$ defined in Definition 4.

Theorem 4. Only 2n-2 scaling factors are necessary to scale a square matrix instead of 2n when investigating the effect of scaling on U, V, or the condition number of the matrix under scaling.

This minimum number of scaling factors was used extensively for partial scaling of a matrix with a numerical nonlinear search technique (NLST) to minimize condition number. In the case of a 3x3 matrix, only four scaling factors (2n-2) were necessary to investigate the effect of scaling instead of six scaling factors.

Theorem 5. K_{SRC} (Equation 5.4f) can be scaled by a scalar so that one of its elements is one.

Theorem 6. If a square, noninteracting matrix is scaled, the diagonal scaling matrices S_1 and S_2 are associated only with the diagonal matrix Σ for the SVD of the given matrix.

Theorem 7. If any matrix A is multiplied by a positive scalar s, this scalar will be multiplied with the each of the diagonal elements of the Σ matrix of the SVD. In effect, scalar multiplication of the matrix A does not affect U, V, or the condition number.

Multiplication of a matrix A by a negative scalar reverses the signs of the U matrix of the SVD compared with the U matrix of the SVD from scaling a matrix with a positive scalar of the same magnitude.

Theorem 8. A square, full or partially interacting matrix can not be scaled to form a noninteracting matrix.

Theorem 9. Any square noninteracting matrix A will be a noninteracting matrix after scaling.

Theorem 10. The necessary and sufficient condition for an orthogonal matrix is that $AA^T = I$.

Theorem 11. If a square matrix can be scaled to be orthogonal, then it will have a condition number of one.

Theorem 12. If $A_{SC}^{-1} = A_{SC}^{T}$, then A can be scaled to be orthogonal. A has a condition number of 1.

Theorem 13. Any nxn noninteracting matrix can be scaled to have a condition number of one.

Theorem 14. Any matrix with all its singular values equal but not equal to one can be scaled to have a condition number of one. For example

$$K = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \qquad \sigma_1 = \sigma_2 = 1.414$$

$$\kappa = 1$$

$$KK^T = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \neq I \qquad K_{SC_{ij}} = \frac{K_{ij}}{\sigma_i}$$

$$K_{SC} = \begin{bmatrix} 0.707 & 0.707 \\ 0.707 & -0.707 \end{bmatrix} \qquad \sigma_1 = \sigma_2 = 1$$

$$\kappa = 1$$

$$K_{SC}(K_{SC})^T = \begin{bmatrix} 0.707 & 0.707 \\ 0.707 & -0.707 \end{bmatrix} \begin{bmatrix} 0.707 & 0.707 \\ 0.707 & -0.707 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I$$

Theorem 15. Any square matrix with all of its singular values equal to one has a condition number of one and is orthogonal. For example

$$K = \begin{bmatrix} 0.5 & 0.5 & -0.5 & -0.5 \\ 0.5 & 0.5 & 0.5 & 0.5 \\ -0.5 & 0.5 & 0.5 & -0.5 \\ -0.5 & 0.5 & -0.5 & 0.5 \end{bmatrix} \qquad \sigma_1 = \sigma_2 = \sigma_3 = \sigma_4 = 1$$

$$\kappa = 1$$

$$KK^T = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = I$$

Grosdidier (61) et al., have several theorems concerning the relationship between condition number and the RGA. One theorem states that the minimum condition number for a matrix with dimension 2x2 is

Theorem 16 $$\kappa_{min} = || \Lambda ||_1 + \sqrt{|| \Lambda ||_1^2 - 1} \qquad (5.5)$$

$$\kappa_{min} = 1 \text{ if and only if } || \Lambda ||_1 = 1 \qquad (5.6)$$

$$\text{i.e. } \kappa_{min} \leq 2 || \Lambda ||_1$$

$$\kappa_{min} = 2 || \Lambda ||_1 \text{ as } || \Lambda ||_1 \rightarrow \infty \qquad (5.7)$$

where $|| \Lambda ||_1$ is the 1_1- norm of the RGA

This theorem could have been expressed as easily in terms of the 1_∞ -norm for 2x2 systems. If one element of the RGA for a 2x2 matrix is known, then the other three elements are determined since the sum of each row or the sum of each column is one.

A closed-loop system is considered sensitive if it is unstable for small amounts of model/plant mismatch. Process steady state gain matrices with an odd number of negative elements tend to be well behaved and are not sensitive. When significant interaction is present, good control can usually be obtained using multivariable compensation such as steady state decoupling.

The value of $2 ||\Lambda||_1$ is an upper bound for κ_{min}. Large elements in the RGA suggest that $||\Lambda||_1$ is large. This implies that $_{min}$ is relatively large. Therefore, a 2x2 system with large RGA elements is sensitive to modeling errors and is difficult to control with any control strategy. These conclusions are supported by Shinskey's work (10) on the relationship of λ_{11} and decoupler error sensitivity on the stability of decoupled 2x2 systems. The amount of decoupler error necessary to destabilize the closed-loop system

tends to decrease as $|\lambda_{11}|$ increases for systems with $\lambda_{11} < 0$ and $\lambda_{11} > 1$. Closed-loop stability of steady state process gain matrices with $0 < \lambda_{11} < 1$ was not affected by decoupler error.

Another theorem of Grosidier et al. (61) was based on their numerical results is stated as follows:

Theorem 17. "No system larger than 3x3 with all nonzero entries can have a minimum condition number of unity." This theorem suggests that as the size of an optimally scaled matrix grows, κ_{min} increases in magnitude. Thus, the matrix tends to be more sensitive to modeling errors.

Theorem 17 will be shown to be incorrect by refuting their proof in Appendix D and by listing examples of 7 cases of 4x4 full matrices that are orthogonal and thus have a condition number of one. Refer to Table 5.1.

Theorem 18. The relative changes in the elements of a square arbitrary gain matrix (g_{ij}), inverse gain matrix ($\hat{g}_{ij}$) and RGA elements (λ_{ij}) are related by the following expressions:

$$\frac{d\lambda_{ij}}{\lambda_{ij}} = (1-\lambda_{ij}) \frac{dg_{ij}}{g_{ij}} \tag{5.8}$$

$$\frac{d\lambda_{ij}}{\lambda_{ij}} = \frac{\lambda_{ij}-1}{\lambda_{ij}} \frac{d\hat{g}_{ji}}{\hat{g}_{ji}} \tag{5.9}$$

Table 5.1. Real, Symmetric Matrices with a Condition Number of One

4 x 4 Matrices

$$\begin{pmatrix} 1 & 1 & -1 & -1 \\ 1 & 1 & 1 & 1 \\ -1 & 1 & 1 & -1 \\ -1 & 1 & -1 & 1 \end{pmatrix} \qquad \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix}$$

$$\begin{pmatrix} 1 & -1 & 1 & -1 \\ -1 & 1 & 1 & -1 \\ 1 & 1 & 1 & 1 \\ -1 & -1 & 1 & 1 \end{pmatrix} \qquad \begin{pmatrix} 1 & -1 & -1 & 1 \\ -1 & 1 & -1 & 1 \\ -1 & -1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix}$$

$$\begin{pmatrix} 1 & -1 & 1 & 1 \\ -1 & 1 & 1 & 1 \\ 1 & 1 & 1 & -1 \\ -1 & -1 & -1 & 1 \end{pmatrix} \qquad \begin{pmatrix} 1 & 1 & 1 & -1 \\ 1 & 1 & -1 & 1 \\ 1 & -1 & 1 & 1 \\ -1 & 1 & 1 & 1 \end{pmatrix}$$

$$\begin{pmatrix} -1 & 1 & 1 & 1 \\ 1 & -1 & 1 & 1 \\ 1 & 1 & -1 & 1 \\ 1 & 1 & 1 & -1 \end{pmatrix}$$

Equation (5.8) shows that as the RGA elements become larger than one, the λ_{ij} become more sensitive to changes in the elements of the process gain matrix (g_{ij}). For large values of λ_{ij}, the relative change in λ_{ij} and g_{ij} are approximately equal as shown in equation (5.9). However, for large values of λ_{ij}, there is small relative change in g_{ij} that causes large relative change in $\hat{g}_{ji}$. This suggests an error sensitive system.

Units on SVD

Questions concerning the consistency and distribution of units for the steady state gain matrix and the component matrices of SVD are not considered at all in the mathematical literature on SVD. Prasad (1) states that the distribution of units in SVD is probably unique because of the ordering of operations. He shows that the units on U, Σ, V^T must be of a form that the units of the U elements are associated with the output units, units of the V^T elements are associated with the input units, and the Σ matrix is unitless.

The distribution of units on SVD will be considered for a matrix of dimension 2x2.

Assuming a linear system

$$y = Kx \tag{5.10}$$

The SVD of the process steady state gain matrix is then

$$K = U\Sigma V^T \tag{5.11}$$

Substituting the SVD of the process steady state gain matrix in equation (5.10) and writing the equation on an element by element basis yields

$$y_1 = (u_{11}\sigma_1 v_{11}^T + u_{11}\sigma_2 v_{21}^T)x_1 + (u_{11}\sigma_1 v_{12}^T + u_{11}\sigma_2 v_{22}^T)x_2 \qquad (5.12)$$

$$y_2 = (u_{21}\sigma_1 v_{11}^T + u_{22}\sigma_2 v_{21}^T)x_1 + (u_{21}\sigma_1 v_{12}^T + u_{22}\sigma_2 v_{22}^T)x_2 \qquad (5.13)$$

By writing the above two equations, as four equations in terms of inputs and outputs

$$\frac{[y_1]}{[x_1]} = (u_{11}\sigma_1 v_{11}^T + u_{11}\sigma_2 v_{21}^T) \qquad (5.14)$$

$$\frac{[y_1]}{[x_2]} = (u_{11}\sigma_1 v_{12}^T + u_{11}\sigma_2 v_{22}^T) \qquad (5.15)$$

$$\frac{[y_2]}{[x_1]} = (u_{21}\sigma_1 v_{11}^T + u_{22}\sigma_2 v_{21}^T) \qquad (5.16)$$

$$\frac{[y_2]}{[x_2]} = (u_{21}\sigma_1 v_{12}^T + u_{22}\sigma_2 v_{22}^T) \qquad (5.17)$$

$$\text{Let } k_{11} = u_{11}\sigma_1 v_{11}^T = \frac{[y_1]}{[x_1]} \qquad (5.18)$$

$$k_{11} = u_{11}\sigma_2 v_{21}^T = \frac{[y_1]}{[x_1]} \qquad (5.19)$$

$$k_{12} = (u_{11}\sigma_1 v_{12}^T) = \frac{[y_1]}{[x_2]} \tag{5.20}$$

$$k_{12} = (u_{11}\sigma_2 v_{22}^T) = \frac{[y_1]}{[x_2]} \tag{5.21}$$

$$k_{21} = (u_{21}\sigma_1 v_{11}^T) = \frac{[y_2]}{[x_1]} \tag{5.22}$$

$$k_{21} = (u_{22}\sigma_2 v_{21}^T) = \frac{[y_2]}{[x_1]} \tag{5.23}$$

$$k_{22} = (u_{21}\sigma_1 v_{12}^T) = \frac{[y_2]}{[x_2]} \tag{5.24}$$

$$k_{22} = (u_{22}\sigma_2 v_{22}^T) = \frac{[y_2]}{[x_2]} \tag{5.25}$$

From equations 5.14 - 5.17, it appears that units can be put in four places on $[y_1]$, $[y_2]$, $[x_1]$, and $[x_2]$. In equations 5.12 - 5.13, there are ten places for units. Four units could be placed on the u_{ij}, four on the v_{ij}, and two on the σ_i. This could be treated as a problem of eight equations with the unknowns (y_1, y_2, x_1, x_2) or as a problem for four equations and eight unknowns.

The most likely distribution of units can be shown by considering equations 5.14 and 5.22. For this case the units are distributed so that

$[u_{ij}]$ is related to $[y_1]$ or that the first row of U has the same units is y_1

$[u_{2j}]$ is related to $[y_2]$ or that the second row of U has the same units as y_2

$[v_{1j}]^T$ corresponds to $\frac{1}{[x_1]}$ or that the first column of V^T has the same units as $\frac{1}{[x_1]}$

$[v_{2j}]^T$ is related to $\frac{1}{[x_2]}$ or that the second column of V^T has units on $\frac{1}{[x_2]}$

$[\sigma_i]$ is unitless.

Definitions and theorems concerning SVD have been presented. Definitions 8-10 classify a square, nonsingular matrix into three types. This classification scheme was used in examining the effects of scaling on low order, square process gain matrices. Theorem 4 concerning the minimum number of scaling factors needed to scale a matrix will be verified with NLST so that partial scaling agrees with full scaling. Scalar multiplication by a positive scalar that was discussed in Theorem 7 implies that a matrix can be scaled in an infinite number of ways and that the effect of the scalar will be on the diagonal matrix of the SVD.

CHAPTER VI

MATRIX SCALING STUDIES

With the objective of minimizing condition number, empirical scaling techniques and a nonlinear least squares method were applied to low order, square process gain matrices. These techniques were applied to linearized steady state models in terms of process analysis. In general, the nonlinear least squares technique (NLST) was more effective in determining the "apparent" minimum condition number than any of the empirical techniques. However for full, 2x2 matrices and for NI matrices of dimension 2x2, 3x3, and 4x4, the geometric mean scaling technique yielded the same minimum condition number as NLST. Theorem 13 states that a condition number of 1 can be found for any square noninteracting matrix by scaling. This is achieved by geometric mean scaling. A new theorem presented later in this chapter shows that minimum condition number can be reached for full, 2x2 matrices with geometric mean scaling.

Description of a Nonlinear Least Squares Search Technique

The PATERN search subroutine, an optimization technique for minimizing or maximizing a unimodal function of up to 1000 independent variables with nonlinear constraints was used to seek the lowest condition number and the scaling factors necessary to scale the matrix to this condition number. This method is one of the direct search techniques so it does not evaluate any gradient of the function. It is quite easy to use; however, the primary

disadvantage of this technique as with all numerical search algorithms is that there is no guarantee that a global minimum can be found. If the function contains ridges and valleys, a local minimum can appear to be a global minimum. Documentation on this search technique is given by Wilde (61) and Moore (62). Four routines are needed to use PATERN for scaling as illustrated in Figure 6-1. The main program defines the search parameters and calls PATERN. Subroutine PROC contains the objective function that is evaluated when given the values of the search parameters. Another subroutine BOUNDS checks if the search parameters have violated any constraints. The scaling factors or search parameters were never to be set to zero. For each run, the initial step size for each search parameter, number of passes through PATERN, form of the final output, and the initial values of the scaling factors are specified by the programmer. A pass through PATERN is a complete cycle in which the minimum condition number is found for the matrix scaled from the initial steady state gain matrix to within the step size given for the scaling factors. For each pass made through PATERN the search step decreased by a factor of ten. Three passes through PATERN were specified before the lowest condition number was found. The condition number was considered optimal if the same value was obtained by decreasing the step by a factor of ten with other parameters held constant. If the initial step size was 1x10E-01 and three passes were used, the final step size was 1x10E-04 for a given type of scaling. Typically, the range of

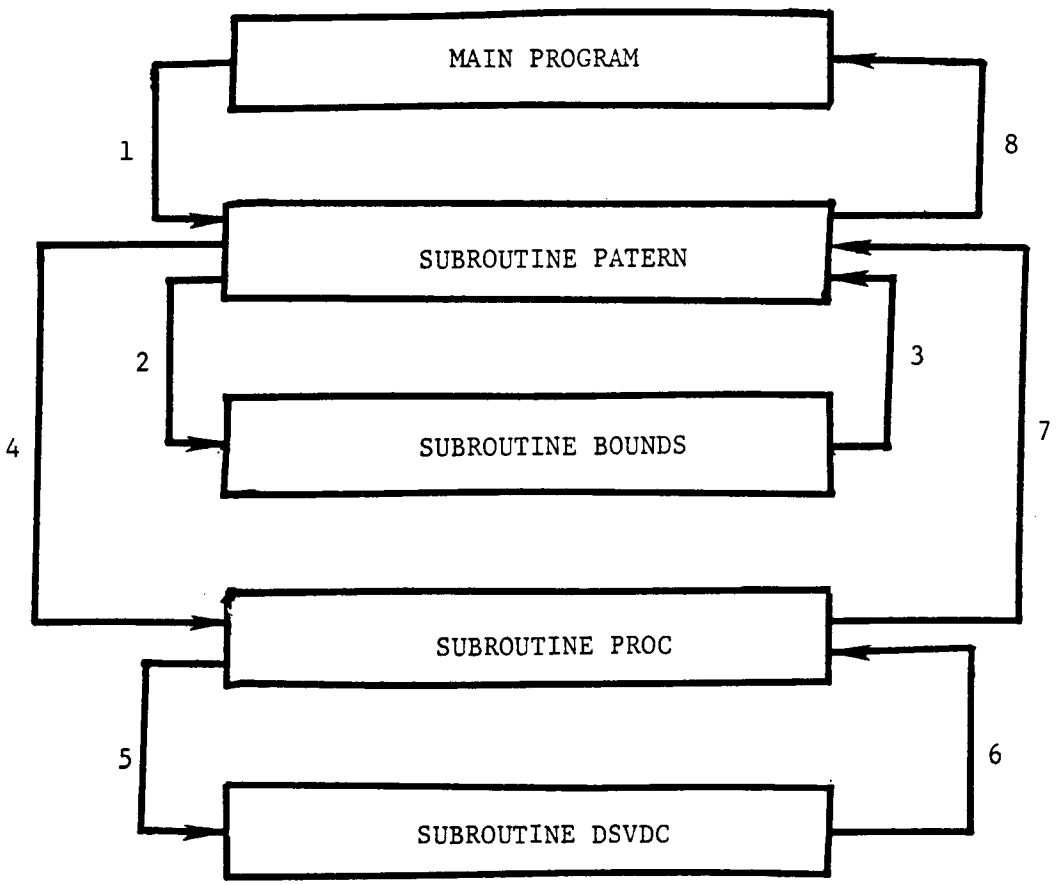


Figure 6.1. Flow Diagram for Nonlinear Least Squares Scaling Technique.

scaling factors used as starting points was 0.1 to 2.0 with the initial step search size set at 0.1. If the condition number obtained from PATERN depended on the search starting point, then the search is probably being carried out over a non-unimodal surface. If the same optimum condition numbers are obtained from different starting points, this indicates a unimodal surface although it is not guaranteed.

Results from Geometric Mean and NLST Scaling

Geometric mean scaling which was first discussed in Chapter 4 was used in this thesis with Tomlin's algorithm. Row followed by column scaling refers to row scaling factors taken from the original matrix. Column scaling factors are then generated from the row scaled matrix. For NI matrices of dimension 2x2, 3x3, and 4x4 and for full, 2x2 matrices, geometric mean scaling provided a lower condition number than that of the unscaled matrix with both the first RC scaling and the first CR scaling giving the same condition number. A "zero criterion" was set at 1xE-08 for the smaller singular values. That is, if a singular value was smaller than 1xE-08, it was considered to be zero. This value is larger than 1.3878D-17, machine zero on the VAX/11-780, but it is much less than most values used for engineering variables. In the geometric mean scaling of a square matrix containing one or more zeros in any row or column, a procedure was made to ignore the zero elements. That is, if a zero(s) exits in a row or column, the element of that row or colum which has the smallest absolute value, is used as the

minimum element in the scaling factor. The main purpose of this was to avoid division by zero in developing the scaling factors. The zero element(s) would remain in the same position(s) of the matrix for successive scaling iterations of the matrix. Refer to Table B.1 in the Appendix for the main program and subroutines of the iterative geometric mean scaling program and to Figure B.2 for a flow chart of the iterative geometric mean scaling procedure.

For NI, 2x2 matrices, geometric mean scaling gave the minimum condition number on the first iteration of either row followed by column scaling or by column followed by row scaling. This is shown in Table 6.1. The scaling factors for these geometric mean scaled matrices are listed in Appendix E.1. The scaling factors for these NI matrices depend on to the absolute value of the nonzero elements.

These geometric mean scalings can be explained by Theorem 13 which states that any square matrix can be scaled to have a condition number of one. The results of scaling NI, 2x2 matrices by NLST are shown in Table 6.1. The scaling factors for these matrices are presented in Appendix F.1. The U and V matrices for NI matrices scaled by geometric mean scaling were elements of $\mathcal{U}_{NI}$ and $\mathcal{V}_{NI}$. According to Theorem 3, the necessary and sufficient condition for a square matrix to be noninteracting is that $U \epsilon \mathcal{U}_{NI}$ and $V \epsilon \mathcal{V}_{NI}$. Theorem 2 states that the singular values for a square, noninteracting matrix are the same as the absolute values of the nonzero elements of the process steady state gain matrix. For

Table 6.1. Effects of Two Scaling Methods on Condition Number for 2 x 2 Matrices

Gain Matrix	Type	κ_{un}	κ_{sc}	CNBND	% Difference
Iterative Geometric Mean					
$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$	NI	1.00000	1.0000	2.00000	0.00
$\begin{pmatrix} 1 & 0 \\ 3 & 4 \end{pmatrix}$	PI	6.34233	2.39326	2.00000	-41.78
$\begin{pmatrix} 0.58 & -0.45 \\ 0.35 & -0.48 \end{pmatrix}$	FULL	7.23819	7.06946	7.21092	-2.33
NLST					
$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}^{*}$	NI	1.00000	1.000000	2.00000	0.00
$\begin{pmatrix} 1 & 0 \\ 3 & 4 \end{pmatrix}^{+}$	PI	6.34233	1.548701	2.00000	-75.58
$\begin{pmatrix} 0.58 & -0.45 \\ 0.35 & -0.48 \end{pmatrix}^{*}$	FULL	7.23819	7.069464	7.21092	-2.33

*SCAL41 k_1 = 0.1 to 2, STEP = 0.01, NPASS = 3, IO = 1, NP = 2.

+SCAL21 k_2 = 0.2 to 4, STEP = 0.001, NPASS = 3, IO = 1, NP = 2.

CNBND = 2 max[$\|\Lambda\|_1$ $\|\Lambda\|_\infty$]. % Difference = $\frac{\kappa_{un} - \kappa_{sc}}{\kappa_{un}} \times 100.$

example, the process steady state matrix

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 6 & 0 & 0 \\ 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}$$

has the singular values of 6, 5, 3, and 1 after the SVD is performed.

With partially interacting, 2x2 matrices, it was observed that geometric mean scaling did not obtain the minimum condition number as illustrated by one example in Table 6.1. Using NLST on this type of matrix would yield the minimum condition number as shown in Table 6.2. Since it was not always obvious on which iteration of geometric mean scaling the "apparent" minimum condition number would converge, PI matrices were studied in terms of their base matrix. A class of matrices will be defined that consists of only the elements 0 and ±1. These matrices will be referred to as base matrices. Any matrix can be said to have a base matrix although not all matrices can be scaled to base matrix form. For any given matrix, the positive elements are replaced by positive ones, the negative elements are replaced by negative ones, and the zeros remain in the same positions to put the matrix into base matrix form. Geometric mean scaling of this matrix has no effect in lowering condition number since the unscaled condition number was the same as the RC and CR scaled condition numbers. This occurs as the geometric mean

Table 6.2. Effects of Condition Number of Partially Interacting Matrices Scaled by a Nonlinear-Least Squares Technique

Gain Matrix	κ_{un}	CNBND	κ_{sc}	% Difference
2 x 2 Matrices				
$\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$	2.61803	2.00000	1.55929	-40.44
k_1 = 0.1 to 2, STEP = 0.01, NPASS = 3, IO = 1				
$\begin{pmatrix} 1 & 0 \\ 3 & 4 \end{pmatrix}$	6.34233	2.00000	1.54870	-75.58
k_2 = 0.2 to 4, STEP = 0.001, IO = 1, NPASS = 3				
3 x 3 Matrices				
$\begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$	2.61803	2.00000	1.67778	-35.91
k_1 = 0.1 to 2, STEP = 0.01, IO = 1, NPASS = 3, NP = 4				
$\begin{pmatrix} 0 & 4 & 5 \\ 3 & 0 & 0 \\ 0 & 7 & 0 \end{pmatrix}$	2.85354	2.00000	1.34661	-52.81
k_1 = 0.1 to 2, STEP = 0.01, IO = 1, NPASS = 3, NP = 4				
4 x 4 Matrices				
$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$	4.04892	2.00000	1.15168	-71.56
k_1 = 0.1 to 2, STEP = 0.1, IO = 1, NPASS = 3, NP = 6				
$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 6 & 7 & 0 \\ 0 & 0 & 5 & 2 \\ 0 & 0 & 0 & 3 \end{pmatrix}$	10.0943	2.00000	1.84537	-81.72
k_1 = 0.1 to 2, STEP = 0.001, NPASS = 3, IO = 1, NP = 6				

CNBND = 2 max $[\|\Omega\|_1 \ \|\Omega\|_\infty]$.

% Difference = $\dfrac{\kappa_{un} - \kappa_{sc}}{\kappa_{sc}} \times 100$.

scaling factors for a base matrix are all one. This effect is shown in Table 6.1. The elements of U and V for the unscaled and scaled base matrix remained the same. When the matrix was not base scaled, after NLST scaling the matrices were symmetrical so that $||U_{11}|| = ||U_{22}||$, $||U_{12}|| = ||U_{21}||$, $||V_{11}|| = ||V_{22}||$ and $||V_{12}|| = ||V_{21}||$.

An iterative geometric mean scaling was developed to lower the condition number of PI matrices of low order. In the iterative geometric mean scaling, the RC scaling was alternated with the CR scaling until the convergence criteria for minimum condition number were reached.

Since the general objective of geometric mean scaling for 2x2 matrices is to make all the elements approximately one, a full matrix can be scaled with four equations and four unknowns.

$$\begin{bmatrix} s_1 & 0 \\ 0 & s_2 \end{bmatrix} \begin{bmatrix} k_{11} & k_{12} \\ k_{21} & k_{22} \end{bmatrix} \begin{bmatrix} s_3 & 0 \\ 0 & s_4 \end{bmatrix} = \begin{bmatrix} s_1 k_{11} s_3 & s_1 k_{12} s_4 \\ s_2 k_{21} s_3 & s_2 k_{22} s_4 \end{bmatrix}$$

$$\approx \begin{bmatrix} \pm 1 & \pm 1 \\ \pm 1 & \pm 1 \end{bmatrix} \qquad (6.1)$$

by scalar scaling

$$\begin{bmatrix} \frac{s_1 s_3}{s_2 s_4} k_{11} & \frac{s_1}{s_2} k_{12} \\ \frac{s_3}{s_4} k_{21} & k_{22} \end{bmatrix} \approx \frac{1}{s_2 s_4} \begin{bmatrix} \pm 1 & \pm 1 \\ \pm 1 & \pm 1 \end{bmatrix} \qquad (6.2a)$$

$$\begin{bmatrix} a_1 a_2 k_{11} & a_1 k_{12} \\ a_2 k_{21} & k_{22} \end{bmatrix} \approx \frac{1}{a_3} \begin{bmatrix} \pm 1 & \pm 1 \\ \pm 1 & \pm 1 \end{bmatrix}. \qquad (6.2b)$$

where $\frac{s_1}{s_2} = a_1$ $\quad \frac{s_3}{s_4} = a_2$ $\quad \frac{1}{s_2 s_4} = a_3$

Rewritten in this form there are four equations and three unknowns. Therefore, the form of all elements of a matrix equal to ±1 can not be reached.

For a PI, 2x2 matrix to become a base matrix

$$\begin{bmatrix} k_{11} & 0 \\ k_{21} & k_{22} \end{bmatrix} = \begin{bmatrix} \pm 1 & 0 \\ \pm 1 & \pm 1 \end{bmatrix} \qquad (6.3)$$

by equation 6.2b, this equation becomes

$$\begin{bmatrix} a_1 a_2 k_{11} & 0 \\ a_2 k_{21} & k_{22} \end{bmatrix} = \frac{1}{a_3} \begin{bmatrix} \pm 1 & 0 \\ \pm 1 & \pm 1 \end{bmatrix} \qquad (6.4)$$

Generally an nxn, PI matrix has n-1 row scaling factors, n-1 column scaling factors, and one scalar scaling. Therefore such a matrix can have at most 2n-1 nonzero elements and be scaled to its base matrix form. In other words, a PI matrix of dimension nxn which has at least $n^2 - (2n-1)$ zeros can be scaled to its base matrix. For all the examples in Table 6.2 that can be scaled to a base matrix reach that form to within three decimal places by the

last iteration of geometric mean scaling. The matrix

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 6 & 7 & 0 \\ 0 & 4 & 5 & 2 \\ 8 & 0 & 0 & 3 \end{bmatrix}$$

with eight nonzeros can not reach the base scaled form so that by the thirty-fifth CR scaling the matrix is

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0.0174 & 0.9829 & 0 \\ 0 & 0.9829 & 1.0174 & 1.00054 \\ 0.00038 & 0 & 0 & 0.99962 \end{bmatrix}$$

As noted previously for full, 2x2 matrices the order of row and column scaling when using geometric mean scaling did not make any difference since column followed by row scaling gave the same condition number as row followed by column scaling.

In general, the minimum condition number could always be found for 2x2, full matrices with either GM or NLST scaling. If the condition number of the GM scaled matrix was not one, the elements of U and V are from the set of $\mathcal{U}_I$ and $\mathcal{V}_I$ matrices. The $\mathcal{U}_I$ and $\mathcal{V}_I$ matrices are both defined as

$$\begin{bmatrix} \pm 0.707 & \pm 0.707 \\ \pm 0.707 & \pm 0.707 \end{bmatrix}$$

However if the original matrix scaled by geometric mean forms a matrix with equal singular values ($\kappa = 1$), then the elements of the scaled orthogonal matrices U_{sc} and V_{sc} are not both elements of $\mathcal{U}_I$ and $\mathcal{V}_I$. There appear to be no set forms for the U and V matrices when there are multiple eigenvalues in the process gain matrix. An example of this case is

$$\underset{\text{GAIN}}{\begin{bmatrix} 1 & 1 \\ 100 & -100 \end{bmatrix}} \quad \underset{U_{uns}}{\begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}} \quad \underset{V_{uns}}{\begin{bmatrix} -0.707 & -0.707 \\ +0.707 & -0.707 \end{bmatrix}} \qquad \kappa = 100$$

$$\underset{\text{RC1}}{\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}} \quad \underset{U_{sc}}{\begin{bmatrix} -0.707 & -0.707 \\ -0.707 & +0.707 \end{bmatrix}} \quad \underset{V_{sc}}{\begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}} \qquad \kappa_{sc} = 1$$

$$\underset{\text{CR1}}{\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}} \quad \underset{U_{sc}}{\begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}} \quad \underset{V_{sc}}{\begin{bmatrix} -0.707 & -0.707 \\ +0.707 & -0.707 \end{bmatrix}} \qquad \kappa_{sc} = 1$$

where uns is unscaled

sc is scaled by geometric mean.

For this case the U_{sc} of the first row followed by column iteration is related to U matrices of the original matrix, first column followed by row iteration, and subsequent iterations by a Givens rotation.

$$\underset{\text{Givens rotation}}{\begin{bmatrix} +0.707 & -0.707 \\ +0.707 & +0.707 \end{bmatrix}} \quad \underset{U_{RC1}}{\begin{bmatrix} -0.707 & -0.707 \\ -0.707 & +0.707 \end{bmatrix}} = \underset{U_{uns}}{\begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}} \tag{6.5}$$

The analytical formulation for full, 2x2 matrices adapted from Tomlin's algorithm will be shown for both row followed by column scaling or by column followed by row scaling.

ROW FOLLOWED BY COLUMN GEOMETRIC MEAN SCALING (RK)C

$$\begin{bmatrix} \frac{1}{\sqrt{|k_{11}||k_{12}|}} & 0 \\ 0 & \frac{1}{\sqrt{|k_{21}||k_{22}|}} \end{bmatrix} \begin{bmatrix} \pm k_{11} & \pm k_{12} \\ \pm k_{21} & \pm k_{22} \end{bmatrix} = \begin{bmatrix} \frac{\pm k_{11}}{\sqrt{|k_{11}||k_{12}|}} & \frac{\pm k_{12}}{\sqrt{|k_{11}||k_{12}|}} \\ \frac{\pm k_{21}}{\sqrt{|k_{21}||k_{22}|}} & \frac{\pm k_{22}}{\sqrt{|k_{21}||k_{22}|}} \end{bmatrix} \tag{6.6}$$

$$\begin{bmatrix} \frac{\pm k_{11}}{\sqrt{|k_{11}||k_{12}|}} & \frac{\pm k_{12}}{\sqrt{|k_{11}||k_{12}|}} \\ \frac{\pm k_{21}}{\sqrt{|k_{21}||k_{22}|}} & \frac{\pm k_{22}}{\sqrt{|k_{21}||k_{22}|}} \end{bmatrix} \begin{bmatrix} \frac{\sqrt[4]{k_{11}k_{12}k_{21}k_{22}}}{\sqrt{|k_{11}||k_{21}|}} & 0 \\ 0 & \frac{\sqrt[4]{k_{11}k_{12}k_{21}k_{22}}}{\sqrt{|k_{12}||k_{22}|}} \end{bmatrix}$$

$$= a \begin{bmatrix} \frac{\pm 1}{\sqrt{|k_{12}||k_{21}|}} & \frac{\pm 1}{\sqrt{|k_{11}||k_{22}|}} \\ \frac{\pm 1}{\sqrt{|k_{22}||k_{11}|}} & \frac{\pm 1}{\sqrt{|k_{12}||k_{21}|}} \end{bmatrix} \tag{6.7}$$

where $a = \sqrt[4]{k_{11}k_{12}k_{21}k_{22}}$

COLUMN FOLLOWED BY ROW GEOMETRIC MEAN SCALING R(KC)

$$\begin{bmatrix} \pm k_{11} & \pm k_{12} \\ \pm k_{21} & \pm k_{22} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{|k_{11}||k_{21}|}} & 0 \\ 0 & \frac{1}{\sqrt{|k_{12}||k_{22}|}} \end{bmatrix} = \begin{bmatrix} \frac{\pm k_{11}}{|k_{11}||k_{21}|} & 0 \\ 0 & \frac{\pm k_{22}}{|k_{12}||k_{22}|} \end{bmatrix} \tag{6.8}$$

$$\begin{bmatrix} \frac{\sqrt[4]{k_{11}k_{12}k_{21}k_{22}}}{\sqrt{|k_{11}||k_{12}|}} & 0 \\ 0 & \frac{\sqrt[4]{k_{11}k_{12}k_{21}k_{22}}}{\sqrt{|k_{21}||k_{22}|}} \end{bmatrix} \begin{bmatrix} \frac{\pm k_{11}}{\sqrt{|k_{11}||k_{21}|}} & 0 \\ 0 & \frac{\pm k_{22}}{\sqrt{|k_{12}||k_{22}|}} \end{bmatrix}$$

$$= a \begin{bmatrix} \frac{\pm 1}{\sqrt{|k_{12}||k_{21}|}} & \frac{\pm 1}{\sqrt{|k_{11}||k_{22}|}} \\ \frac{\pm 1}{\sqrt{|k_{12}||k_{22}|}} & \frac{\pm 1}{\sqrt{|k_{12}||k_{21}|}} \end{bmatrix} \tag{6.9}$$

where $a = \sqrt[4]{k_{11}k_{12}k_{21}k_{22}}$

Another scheme for geometric mean scaling would be to use both the row scaling factors and the column scaling factors from the given process steady state matrix. This scheme will be called simultaneous row and column scaling. This result will be shown in the following equation:

$$\begin{bmatrix} \left(\frac{1}{\sqrt{|k_{11}||k_{12}|}}\right)\left(\pm k_{11}\right)\left(\frac{1}{\sqrt{|k_{11}||k_{21}|}}\right) & \left(\frac{1}{\sqrt{|k_{11}||k_{12}|}}\right)\left(\pm k_{12}\right)\left(\frac{1}{\sqrt{|k_{12}||k_{22}|}}\right) \\ \left(\frac{1}{\sqrt{|k_{21}||k_{22}|}}\right)\left(\pm k_{21}\right)\left(\frac{1}{\sqrt{|k_{11}||k_{21}|}}\right) & \left(\frac{1}{\sqrt{|k_{21}||k_{22}|}}\right)\left(\pm k_{22}\right)\left(\frac{1}{\sqrt{|k_{12}||k_{22}|}}\right) \end{bmatrix} =$$

$$\begin{bmatrix} \frac{\pm 1}{\sqrt{k_{12}k_{21}}} & \frac{\pm 1}{\sqrt{k_{11}k_{22}}} \\ \frac{\pm 1}{\sqrt{k_{11}k_{22}}} & \frac{\pm 1}{\sqrt{k_{12}k_{21}}} \end{bmatrix} \tag{6.10}$$

Equations 6.7 and 6.9 illustrate that the scaled matrix is the same independent of the order of geometric mean scaling. Row followed by column scaling gives the same condition number as column followed by row scaling for full, 2x2 matrices. These two equations are related to equation 6.10 by a scalar. According to Theorem 7, any of these scaling procedures will lead to the same condition number since the U, Σ, and V matrices are unchanged by scalar scaling.

Grosdidier et al. (61), outlined a proof for scaling that yields a minimum condition number for full 2x2 matrices. The scaling factors for this scaling method are shown as:

$$S_1 = \begin{bmatrix} s_1 & 0 \\ 0 & s_1\left|\frac{k_{11}k_{12}}{k_{21}k_{22}}\right|^{\frac{1}{2}} \end{bmatrix} \tag{6.11}$$

$$S_2 = \begin{bmatrix} s_2 & 0 \\ 0 & s_2\left|\frac{k_{11}k_{21}}{k_{12}k_{22}}\right|^{\frac{1}{2}} \end{bmatrix} \tag{6.12}$$

where s_1 and s_2 are positive scalars

If S_1 of equation 6.11 is multiplied by $\dfrac{1}{s_1(\sqrt{|k_{11}||k_{12}|})}$ and S_2 is multiplied by $\dfrac{\sqrt[4]{k_{11}k_{12}k_{21}k_{22}}}{s_2\sqrt{|k_{11}||k_{21}|}}$, then Tomlin's row followed by column scaled matrices are obtained in equations 6.13 and 6.14, respectively.

$$S_1 = \begin{bmatrix} \frac{1}{\sqrt{|k_{11}||k_{12}|}} & 0 \\ 0 & \frac{1}{\sqrt{|k_{21}||k_{22}|}} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & \sqrt{\frac{|k_{11}||k_{12}|}{|k_{21}||k_{22}|}} \end{bmatrix} \tag{6.13}$$

$$S_2 = a \begin{bmatrix} \frac{\sqrt[4]{k_{11}k_{12}k_{21}k_{22}}}{\sqrt{|k_{11}||k_{21}|}} & 0 \\ 0 & \frac{\sqrt[4]{k_{11}k_{12}k_{21}k_{22}}}{\sqrt{|k_{12}||k_{22}|}} \end{bmatrix} \tag{6.14}$$

$$\text{Let } a = \sqrt[4]{k_{11}k_{12}k_{21}k_{22}}$$

Thus Tomlin's row followed by column scaling is proportional to simultaneous row followed by column scaling. These results can be summarized by the following new theorems.

Theorem 6.1. For full, 2x2 matrices geometric mean scaling gives the minimum condition number.

Theorem 6.2. For full, 2x2 matrices scaled by geometric mean scaling, the same minimum condition number can be obtained by row followed by column scaling as by column followed by row scaling.

Since simultaneous row and column scaling is related to Tomlin's geometric mean scaling by a scalar the same minimum condition number can be obtained with this scaling method.

Geometric mean scaling for noninteracting, 3x3 matrices provided an optimum condition number after at most one iteration. The GM/RC scaling after one iteration yielded the same condition number as the GM/CR scaling, and the condition number was equal to one. When the SVD was performed on the RC scaled matrix or CR

scaled matrix their U matrices were elements of $\mathcal{U}_{NI}$ and their V matrices were elements of $\mathcal{V}_{NI}$.

For PI, 3x3 matrices that are not base scaled, geometric mean scaling would lower the condition number although the nonlinear least squares technique would provide an even lower condition number. Using this scaling technique elements of the geometric mean scaled matrix had the same absolute values although the order was different.

In the case of full, 3x3 matrices, geometric mean scaling usually decreased the condition number. The scaling factors that correspond to these matrices are presented in Appendix E.2. Unlike the case for full, 2x2 matrices, no full, 3x3 matrices that consisted of elements equal to $|\frac{\sqrt{3}}{3}|$ were found that were also orthogonal. Either the full matrix had a $\kappa=2$ if it contained multiple singular values or $\kappa=\infty$ if the matrix were rank deficient.

It can be shown that a 3x3 matrix with all elements equal to $|\frac{\sqrt{3}}{3}|$ can not be orthogonal. Of the following eleven forms of matrices of this type, that the first two are singular and the other nine have multiple eigenvalues.

$$\begin{bmatrix} 0.577 & -0.577 & -0.577 \\ -0.577 & 0.577 & 0.577 \\ -0.577 & 0.577 & 0.577 \end{bmatrix} \quad \begin{bmatrix} 0.577 & -0.577 & 0.577 \\ -0.577 & 0.577 & -0.577 \\ 0.577 & -0.577 & 0.577 \end{bmatrix}$$

Table 6.3. Effects of Two Scaling Methods on Condition Number for 3 x 3 Matrices

Gain Matrix	Type	κ_{un}	κ_{sc}	CNBND	% Difference
Iterative Geometric Mean (SCAL 33)					
$\begin{pmatrix} 3 & 0 & 0 \\ 0 & 7 & 0 \\ 0 & 0 & 5 \end{pmatrix}$	NI	2.33333	1.00000	2.00000	+57.14
$\begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$	PI	2.61803	2.61803	2.00000	0
$\begin{pmatrix} 1 & 1 & -0.1 \\ 0.1 & 2 & -1 \\ -2 & -3 & 1 \end{pmatrix}$	FULL	42.25010	49.6327	24.41509	-17.47
NLST (SCAL 33)					
$\begin{pmatrix} 3 & 0 & 0 \\ 0 & 7 & 0 \\ 0 & 0 & 5 \end{pmatrix}^*$	NI	2.33333	1.00023	2.00000	+57.13
$\begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}^*$	PI	2.61803	1.677784	2.00000	+35.91
$\begin{pmatrix} 1 & 1 & -0.1 \\ 0.1 & 2 & -1 \\ -2 & -3 & 1 \end{pmatrix}^*$	FULL	42.25010	24.37407	24.41509	+42.31

*k_1 = 0.1 to 2, STEP = 0.01, NPASS = 3, IO = 1, NP = 4.

CNBND = 2 max[$\|\Lambda\|_1$ $\|\Lambda\|_\infty$].

% Difference = $\frac{\kappa_{un} - \kappa_{sc}}{\kappa_{un}}$ x 100.

$$\begin{bmatrix} 0.577 & -0.577 & 0.577 \\ -0.577 & 0.577 & 0.577 \\ 0.577 & 0.577 & 0.577 \end{bmatrix} \quad \begin{bmatrix} 0.577 & 0.577 & -0.577 \\ 0.577 & 0.577 & 0.577 \\ -0.577 & 0.577 & 0.577 \end{bmatrix}$$

$$\begin{bmatrix} 0.577 & 0.577 & 0.577 \\ 0.577 & 0.577 & -0.577 \\ 0.577 & -0.577 & 0.577 \end{bmatrix} \quad \begin{bmatrix} -0.577 & 0.577 & 0.577 \\ 0.577 & -0.577 & 0.577 \\ 0.577 & 0.577 & -0.577 \end{bmatrix}$$

$$\begin{bmatrix} 0.577 & 0.577 & -0.577 \\ 0.577 & -0.577 & 0.577 \\ -0.577 & 0.577 & 0.577 \end{bmatrix} \quad \begin{bmatrix} 0.577 & -0.577 & -0.577 \\ -0.577 & 0.577 & -0.577 \\ -0.577 & -0.577 & 0.577 \end{bmatrix}$$

$$\begin{bmatrix} 0.577 & 0.577 & -0.577 \\ 0.577 & -0.577 & 0.577 \\ 0.577 & 0.577 & 0.577 \end{bmatrix} \quad \begin{bmatrix} 0.577 & 0.577 & 0.577 \\ 0.577 & -0.577 & 0.577 \\ -0.577 & 0.577 & 0.577 \end{bmatrix}$$

$$\begin{bmatrix} 0.577 & 0.577 & -0.577 \\ 0.577 & 0.577 & 0.577 \\ -0.577 & 0.577 & 0.577 \end{bmatrix}$$

For noninteracting matrices, the general trend persists that minimum condition number is achieved for geometric mean scaling for the first iteration of RC or CR scaling as shown in Table 6.4. Scaling factors for these matrices are shown in Appendix E.3. If the partial interacting matrix were not base scaled, the condition number usually decreased with geometric mean scaling. The condition number could be lowered below that of geometric mean scaling by NSLT. Scaling factors for matrices scaled by NLST are given in Appendix F.2. Elements of U_{sc} and V_{sc} by geometric mean scaling contained the same elements when base scaled. However, the elements of U_{sc} and V_{sc} appeared to show no relationship to each other when not base scaled.

In the case of 4x4, full matrices, geometric mean scaling usually lowered the condition number although NLST could lower it

Table 6.4. Effects of Two Scaling Methods on Condition Number for 4x4 Matrices

Gain Matrix	Type	κ_{un}	κ_{sc}	CNBND	% Difference
Iterative Geometric Mean (SCALT4)					
$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 6 & 0 & 0 \\ 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}$	NI	6.00000	1.00000	2.00000	+83.3
$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 6 & 7 & 0 \\ 0 & 0 & 5 & 2 \\ 0 & 0 & 0 & 3 \end{bmatrix}$	PI	10.09430	2.76998	2.00000	+72.6
$\begin{bmatrix} 3.3 & 5.6 & 7.6 & 4.5 \\ 4.5 & 7.6 & 5.6 & 3.3 \\ 12.4 & 3.6 & 7.5 & 4.7 \\ 4.7 & 7.5 & 3.6 & -12.4 \end{bmatrix}$	FULL	10.5099	9.6755	8.68788	17.34
NLST (SCAL4)					
$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 6 & 0 & 0 \\ 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}$	NI	6.00000	1.00166	2.00000	+83.31
$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 6 & 7 & 0 \\ 0 & 0 & 5 & 2 \\ 0 & 0 & 0 & 3 \end{bmatrix}$	PI	10.09430	1.52190	2.00000	+84.92
$\begin{bmatrix} 3.3 & 5.6 & 7.6 & 4.5 \\ 4.5 & 7.6 & 5.6 & 3.3 \\ 12.4 & 3.6 & 7.5 & 4.7 \\ 4.7 & 7.5 & 3.6 & -12.4 \end{bmatrix}$	FULL	10.5099	8.292711	8.68788	17.34

k_1 = 0.1 to 2, STEP = 0.01, NPASS = 3, IO = 1, NP = 6.

CNBND = 2 max[$\|\Omega\|_1$ $\|\Omega\|_\infty$].

% Difference = $\frac{\kappa_{un} - \kappa_{sc}}{\kappa_{un}} \times 100.$

more. Scaling factors for NLST scaled matrices of dimension 4x4 are shown in Appendix F.3. There is no discernible pattern in the U and V matrices that were scaled by geometric mean or by NLST.

The basic algorithm for geometric mean scaling can be extended to nonsquare systems. For example, 3x2 systems were briefly considered and scaled in the following manner:

$$K_{sc} = \begin{bmatrix} s_1 & 0 & 0 \\ 0 & s_2 & 0 \\ 0 & 0 & s_3 \end{bmatrix} \begin{bmatrix} k_{11} & k_{21} \\ k_{12} & k_{22} \\ k_{12} & k_{23} \end{bmatrix} \begin{bmatrix} s_4 & 0 \\ 0 & s_5 \end{bmatrix} \qquad (6.16)$$

Scaling factors for empirical scaling methods or for NLST have the only constraint that they can not be zero.

Grosdidier et al. (61) plotted smallest singular value or σ_* versus $||\Lambda||_1$ for a large number of 3x3 and 4x4 matrices with random coefficients between -100 and +100. They stated in a theorem that no full matrix of dimension greater than 3x3 can have a condition number of one. An effort was made to test their conjecture. A series of real, symmetric matrices were generated with most elements equal to +1 and with varying combinations of -1. For example, the following combinations were possible in the 4x4 case: a set of 3 pairs out of a set of 6 pairs of symmetric elements equal to -1 are interchanged

$$\binom{6}{1} = 6$$

$$\binom{6}{2} = 15$$

$$\binom{6}{3} = 20$$

$\binom{6}{4} = 15$

$\binom{6}{5} = 6$

$\binom{6}{6} = 1$

$\overline{63}$ total combinations

3 pairs $1,2 \rightleftarrows 2,1;\ 2,3 \rightleftarrows 3,2;\ 1,4 \rightleftarrows 4,1$

6 possible pairs $1,2 \rightleftarrows 2,1;\ 2,3 \rightleftarrows 3,2;\ 3,4 \rightleftarrows 4,3;\ 1,3 \rightleftarrows 3,1;$
$1,4 \rightleftarrows 4,1;\ 2,4 \rightleftarrows 4,2$

$$\begin{bmatrix} 1 & -1 & 1 & -1 \\ -1 & 1 & -1 & 1 \\ 1 & -1 & 1 & 1 \\ -1 & 1 & 1 & 1 \end{bmatrix} \quad \kappa = 1$$

The rank deficient cases for these matrices were eliminated. Then, the SVD was performed on the remaining matrices. Tentatively, no cases of full orthogonal matrices were found for these real, symmetric matrices of dimension 3x3 and 5x5. However, seven cases were found for $\kappa = 1$ for 4x4 matrices as shown in Table 5.1. Their theorem is not quite correct and a proof by counterexample using properties of orthogonal matrices will be shown in Appendix D. For the 4x4 matrices the condition number of the absolute value of the elements of the U matrix were not the same as the corresponding elements of the V matrix. For real, symmetric matrices with elements equal to -1 along the diagonal or the secondary diagonal of dimension 3x3, 5x5, and 6x6, the absolute values of the elements of

the U matrix were the same as the elements of the V matrix within the corresponding column. Table 6.5 summarizes the information about unscaled condition number for these matrices.

A series of full, square matrices with all elements equal to one except the diagonal elements being -1 were run. The same matrix could be rotated 180° either clockwise or counterclockwise and have the same decomposition except for the signs on the elements of corresponding U and V matrices. For matrices of dimension 4x4 to 20x20, the following formulas could be generated for their respective condition number and singular values. Matrices with dimension greater than 20x20 were not run.

$$\kappa = \frac{\sigma_1}{\sigma_n} = \frac{n-2}{2}$$

$$\sigma_1 = n-2$$

$$\sigma_{2,\ldots,}\sigma_n = 2$$

Comparison of Empirical Scaling Techniques

Five forms of empirical scaling namely arithmetic mean (AM), geometric mean (GM), geometric mean followed by equilibration (GM/EQ), and arithmetic mean followed by equilibration (AM/EQ) were combined in various order to form 14 combinations. When Prasad (1) examined five full transfer functions, he found that arithmetic mean scaling with row followed by column scaling (AM/RC) tended to increase the condition number in comparison with the unscaled condition number. This is not thought to be a general result.

Table 6.5. Properties of Selected Real, Symmetric Matrices with all Elements ±1

Gain Matrix	κ_{un}	Singular Values	U and V *
$\begin{pmatrix} -1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \end{pmatrix}$	2	$\sigma_1 = \sigma_2 = 2.00,\ \sigma_3 = 1.00$.	$\lvert u_{kj} \rvert = \lvert v_{ij} \rvert$
$\begin{pmatrix} -1 & 1 & 1 & 1 \\ 1 & -1 & 1 & 1 \\ 1 & 1 & -1 & 1 \\ 1 & 1 & 1 & -1 \end{pmatrix}$	1	$\sigma_1 = \sigma_2 = \sigma_3 = \sigma_4 = 2.00$	$\lvert u_{ij} \rvert \neq \lvert v_{ij} \rvert$
$\begin{pmatrix} -1 & 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & 1 & 1 \\ 1 & 1 & -1 & 1 & 1 \\ 1 & 1 & 1 & -1 & 1 \\ 1 & 1 & 1 & 1 & -1 \end{pmatrix}$	1.5	$\sigma_1 = 3.00,\ \sigma_2 = \sigma_3 = \sigma_4 = \sigma_5 = 2.00$	$\lvert u_{kj} \rvert = \lvert v_{ij} \rvert$
	3.24	$\sigma_1 = \sigma_2 = 3.24,\ \sigma_3 = \sigma_4 = 1.24,\ \sigma_5 = 1.00$	
	4.48	$\sigma_1 = 3.39,\ \sigma_2 = 2.76,\ \sigma_3 = 2,\ \sigma_4 = 1.13,$ $\sigma_5 = 0.0757$	
	6.34	$\sigma_1 = 3.56,\ \sigma_2 = \sigma_3 = \sigma_4 = \sigma_5 = 2.00,\ \sigma_6 = 0.56$	
$\begin{pmatrix} -1 & 1 & 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & 1 & 1 & 1 \\ 1 & 1 & 1 & -1 & 1 & 1 \\ 1 & 1 & 1 & 1 & -1 & 1 \\ 1 & 1 & 1 & 1 & 1 & -1 \end{pmatrix}^{+}$			
	2	$\sigma_1 = 4.00,\ \sigma_2 = \sigma_3 = \sigma_4 = \sigma_5 = \sigma_6 = 2.00$	$\lvert u_{kj} \rvert = \lvert v_{ij} \rvert$
	4.04	$\sigma_1 = 3.60,\ \sigma_2 = 3.24,\ \sigma_3 = 2.49$ $\sigma_4 = 2.00,\ \sigma_5 = 1.24,\ \sigma_6 = 0.890$	
	5.41	$\sigma_1 = 3.76,\ \sigma_2 = 3.06,\ \sigma_3 = \sigma_4 = \sigma_5 = 2.00,\ \sigma_6 = 0.695$	

*Applies only to real, symmetric matrices with all elements +1 except for -1 along the diagonal or secondary diagonal.

+Only 34 cases were run from combinations of $\binom{15}{6}$, $\binom{15}{7}$, and $\binom{15}{8}$.

Nine empirical scaling techniques were examined on selected PI and full matrices of dimension 3x3. Refer to Tables 6.6-6.9. Arithmetic mean increased the condition number in four of the six examples. However for the first example of Table 6.6, AM/CR decreased the condition number more than by any of the other empirical methods. Geometric mean increased the condition number in three of the examples. On the basis of this limited number of examples, it is not apparent that geometric mean is as effective in lowering the condition number as it appears to be for 2x2 matrices. In all cases the lowest condition number was obtained using NLST. When the minimum condition number using NLST is less than the CNBND, it appears that this bound is fairly tight.

The following example from Grosdidier et al. (61) illustrates how sensitive a system can be with a 5% difference between the model and the plant gain matrices.

$$K_m = \begin{bmatrix} 1 & 1.05 \\ 0.95 & 1 \end{bmatrix} \tag{6.17}$$

$$K_p = \begin{bmatrix} 1 & 1.05 \\ 1 & 1 \end{bmatrix} \tag{6.18}$$

where K_m is the model transfer function

K_p is the plant transfer function

Table 6.6. Effect of Iterative Geometric Mean Scaling on Condition Number for 3 x 3 Matrices

Gain Matrix	No.	Type	κ_{un}	CNBND	κ_{sc}	% Difference
$\begin{bmatrix} 7 & 0 & 0 \\ 2 & 4 & 0 \\ 2.3 & 2.3 & 2.1 \end{bmatrix}$	1	PI	4.5219	2.0000	2.6356	41.71
$\begin{bmatrix} -2 & 1.5 & 1 \\ 1.5 & -1 & 2 \\ 1 & 2 & 1.5 \end{bmatrix}$	2	FULL	2.08167	2.1026	2.08167*	0.00
$\begin{bmatrix} 1 & -2 & -2 \\ 0 & 1 & 1 \\ 1 & -2 & -1.5 \end{bmatrix}$	3	PI	17.9443	18.0000	20.8840*	-16.38
$\begin{bmatrix} 1 & 20 & 0.5 \\ 0 & 1 & 0.5 \\ 1 & 1 & 1 \end{bmatrix}$	4	PI	56.1934	2.2000	3.9413	+92.99
$\begin{bmatrix} 1.5 & 4.8 & 4.8 \\ 7.9 & 8.8 & 7.4 \\ 6.9 & 9.0 & 3.4 \end{bmatrix}$	5	FULL	11.2024	10.8850	13.5221*	-20.71
$\begin{bmatrix} 1 & 1 & -0.1 \\ 0.1 & 2 & -1 \\ -2 & -3 & 1 \end{bmatrix}$	6	FULL	42.2501	24.4151	49.6327	-17.47

*Value at termination of program. CNBND = $2\|\Lambda\|_1 \|\Lambda\|_\infty$. % Difference = $\frac{\kappa_{un} - \kappa_{sc}}{\kappa_{un}} \times 100$.

Table 6.7. Effect of Different Scaling Techniques on Condition Number for 3 x 3 Matrices

Gain Matrix No.	κ_{AMRC} (% Difference)	κ_{AMCR} (% Difference)	κ_{EQRC} (% Difference)	κ_{EQCR} (% Difference)	κ_{NLST} (% Difference)
1	2.13131 (+95.05)	1.85933 (+95.68)	3.37360 (+92.17)	1.93667 (+95.50)	1.11185 (+97.42)
2	11.76350 (-465.10)	11.76350 (-465.10)	2.08167 (+0.00)	2.08167 (+0.00)	2.08167 (+0.00)
3	22.16040 (-23.50)	22.16040 (-23.50)	21.29720 (-18.69)	21.29720 (-18.69)	17.94427 (+0.00)
4	3.31458 (+94.10)	2.94074 (+94.77)	7.52868 (+91.94)	3.87494 (+93.10)	2.19941 (+96.09)
5	12.41870 (-10.86)	12.46350 (-11.26)	11.29980 (+0.87)	11.15680 (+0.41)	10.79591 (+3.63)
6	43.14910 (-2.09)	78.13752 (-84.94)	25.63400 (-39.33)	25.63400 (-39.33)	24.37407 (+42.31)

Table 6.8. Effect of Geometric Mean Followed by an Empirical Scaling Technique on Condition Number for 3 x 3 Matrices

Gain Matrix No.	$\kappa_{GM/AMRC}$ (% Difference)	$\kappa_{GM/AMCR}$ (% Difference)	$\kappa_{GM/EQRC}$ (% Difference)	$\kappa_{GM/EQCR}$ (% Difference	κ_{NLST} % Difference)
1	2.90328 (+93.26)	2.90223 (+93.26)	4.02395 (+90.66)	4.02395 (+90.66)	1.18520 (+97.40)
4	2.21072 (+96.07)	2.21072 (+96.07)	2.72396 (+95.15)	2.62396 (+95.15)	2.19941 (+96.09)

Table 6.9. Percent Change in Condition Number for Scaled 3 x 3 Matrices

Empirical Methods		% Difference					
		No. 1	2	3	4	5	6
		Type PI	FULL	PI	PI	FULL	FULL
AM							
	RC	+95.05	-465.10	-23.50	+94.10	-10.86	-2.09
	CR	+95.68	-465.10	-23.50	+94.77	-11.26	-84.94
EQ							
	RC	+92.17	0	-18.69	+91.94	-0.87	-39.33
	CR	+95.50	0	-18.69	+93.10	+0.41	-39.33
GM		+41.71	$+1.44 \times 10^{-3}$	-16.38	+92.99	-20.71	+17.47
GM/AM							
	RC	+93.26	-	-	+96.07	-	-
	CR	+93.26	-	-	+96.07	-	-
GM/EQ							
	RC	+90.66	-	-	+95.15	-	-
	CR	+90.66	-	-	+95.15	-	-
Nonlinear Least Squares Technique		+97.42	+0.00	+0.00	+96.09	+3.63	+42.31

$$\% \text{ Difference} = \frac{\kappa_{un} - \kappa_{sc}}{\kappa_{un}} \times 100.$$

AM = arithmetic mean
EQ = equilibration
GM = iterative geometric mean (RC)
RC = row followed by column
CR = column followed by row

	λ_{11}	κ_{un}	κ_{min}	$\frac{\kappa_{un}-\kappa_{min}}{\kappa_{min}} \times 100$
K_m	400	1602	1598	0.0250
K_p	-70	82.04	81.99	0.062

Perturbations

Symmetric perturbations of two elements, four elements, and six elements of a 3x3 singular matrix had unscaled condition numbers varying from 6143.25 to infinity. These perturbed matrices are shown in Table 6.10. The perturbations of these matrices ranged from 0.625% to 2.5%. These same eight matrices when scaled by iterative geometric mean had condition numbers ranging from 5845.84 to infinity. The range of condition number obtained for the perturbed matrices illustrates how sensitive a matrix can be to slight changes.

Types of Scaling and Scaling Factors Using NLST

Prasad examined three types of scalings using the nonlinear least squares search technique and applied these methods to five full, 2x2 steady state process gain matrices. For individual row or column scaling, only one scaling factor is changed in the row scaling matrix or in the column scaling matrix. All of the other scaling factors are not one. Partial scaling refers to scaling the

Table 6.10. Variations in Condition Number from Perturbations of a Singular 3 x 3 Gain Matrix

No.	Gain Matrix	κ_{un}	κ_{GM}	% Difference
1	$\begin{bmatrix} 1 & 2.05 & 3 \\ 3.95 & 5 & 6.05 \\ 7 & 7.95 & 9 \end{bmatrix}$	6143.3	5845.8	+4.84
2	$\begin{bmatrix} 1 & 2.05 & 3 \\ 3.95 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$	13565	12997	+4.19
3	$\begin{bmatrix} 1 & 2.05 & 3.05 \\ 3.95 & 5 & 6.05 \\ 6.95 & 7.95 & 9 \end{bmatrix}$	24770	23699	+4.32
4	$\begin{bmatrix} 1 & 2 & 3.05 \\ 4 & 5 & 6 \\ 6.95 & 8 & 9 \end{bmatrix}$	24526	23578	+3.87
5	$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6.05 \\ 7 & 7.95 & 9 \end{bmatrix}$	121970	116076	+4.83
6	$\begin{bmatrix} 1 & 2 & 3.05 \\ 4 & 5 & 6.05 \\ 6.95 & 7.95 & 9 \end{bmatrix}$	∞	∞	0
7	$\begin{bmatrix} 1 & 2.05 & 3.05 \\ 3.95 & 5 & 6 \\ 6.95 & 8 & 9 \end{bmatrix}$	∞	∞	0
8	$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$	∞	∞	0

$$\% \text{ Difference} = \frac{\kappa_{un} - \kappa_{sc}}{\kappa_{un}} \times 100.$$

minimum number of rows and columns in a matrix simultaneously as defined by Theorem 4. Full matrix scaling varies all the scaling elements in the scaling matrices and thus scales all the rows and columns simultaneously. For a 2x2 system, only one element out of four is changed in the scaling matrices by individual scaling. One row and one column are scaled simultaneously in partial scaling of 2x2 systems by changing one element in each scaling matrix. Full scaling would vary all four elements of the scaling matrices.

The following notation will be applied to equation 1.1 for a scaled 2x2 system.

$$K_{sc} = \begin{bmatrix} s_1 & 0 \\ 0 & s_2 \end{bmatrix} \begin{bmatrix} k_{11} & 0 \\ k_{21} & k_{22} \end{bmatrix} \begin{bmatrix} s_3 & 0 \\ 0 & s_4 \end{bmatrix} \qquad (6.19)$$

Therefore, there are four combinations of partial scalings for a 2x2 matrix $(s_1,s_3;s_1,s_4;s_2,s_3;s_2,s_4)$.

For the 2x2 systems he studied, Prasad found that partial scaling was required to scale a system optimally based on Theorems 4 and 5. His numerical examples illustrated these results.

Condition Number Surface

Some information about the conditioning of a process can be determined from the condition number surface of the process under scaling. Lau obtained sensitivity of operability information by plotting contours of the condition number on reactor volume versus conversion of one component and by plotting minimum singular value on reactor volume versus conversion for the value of the steady

state gain matrix over a region defined by state and constraint equations (22). The order of scaling by the empirical methods of GM/EQ,EQ, or EQ/GM affected his contours of sensitivity and invertibility profoundly. With his method of variable normalization and equation equilibration, the sensitivity changed less than a factor of two in his example of a first-order endothermic reaction in a CSTR.

Prasad found that there was variation in condition numbers of up to 99.4% for five transfer functions using partial and full scaling. He proposed that these effects were due to a rough condition number surface. However it appears that he was mistaken by using step sizes that were too small. His initial and final step size for partial scaling were 1x10E-3 and 1x10E-6. For full scaling his step sizes were 1x10E-2 and 1x10E-5. Three dimensional plots were generated for the same five full, 2x2 transfer functions listed in Table C.1 under partial scaling. Figures 6.2-6.6 illustrate these plots and their associated unscaled condition numbers. For each of these transfer functions, partial scaling was performed over certain ranges. The data were plotted on axes corresponding to $S_{r_1} - \frac{1}{\kappa} - S_{c_1}$ in most instances. By using $\frac{1}{\kappa}$ instead of κ, the condition number varies from 0 to 1. In Figure 6.4 which scales the transfer function of Weischedel and McAvoy's Column C the surface of the plot is smooth. The point corresponding to $\frac{1}{\kappa} = 0.0138$ is near the "hump" of the plot which is in the optimal operating region. The plot of $S_{r_1} - |U_{11}| - S_{c_1}$ shows a surface that is smooth over

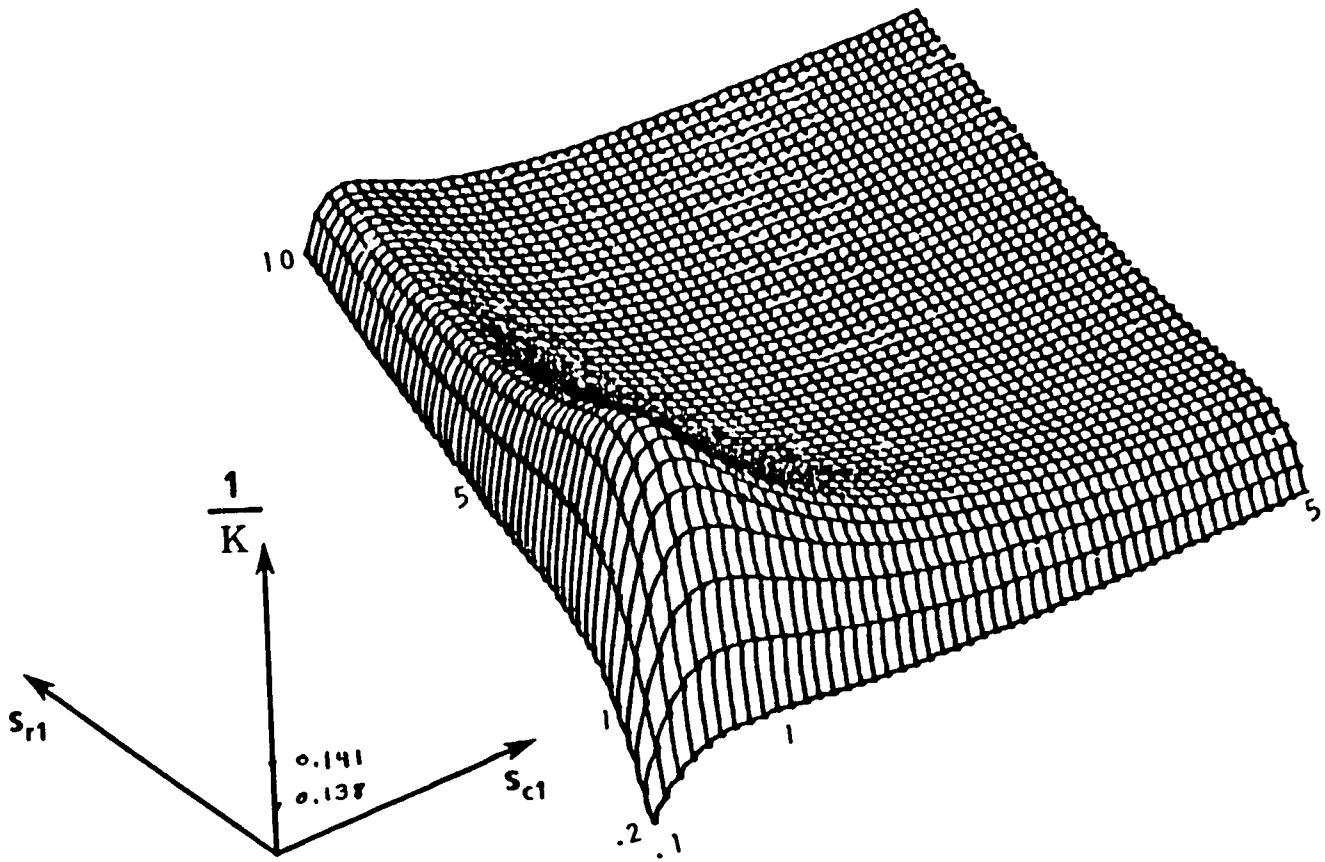


Figure 6.2. The Reciprocal of Condition Number ($\frac{1}{\kappa}$) as a Function of Scaling Factors (S_{r_1} and S_{c_1}) for a Transfer Function of McAvoy's Column A.

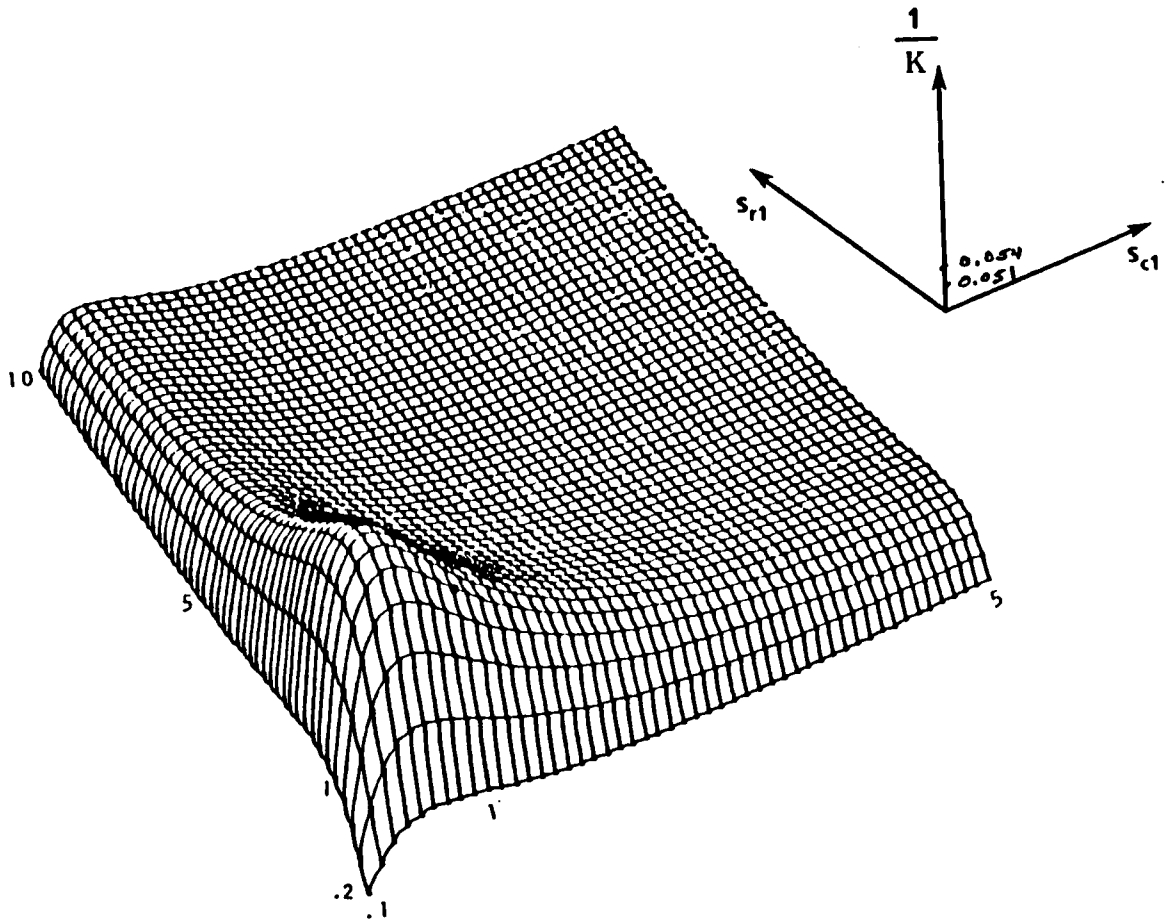


Figure 6.3. The Reciprocal of Condition Number ($\frac{1}{K}$) as a Function of Scaling Factors (S_{r_1} and S_{c_1}) for a Transfer Function of McAvoy's Column B.

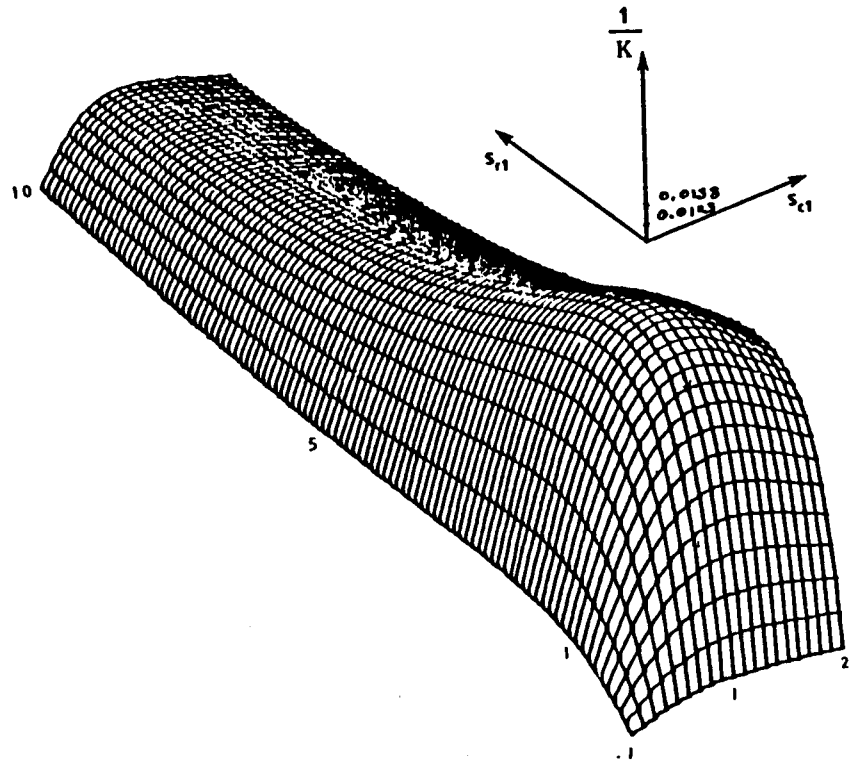


Figure 6.4. The Reciprocal of Condition Number ($\frac{1}{\kappa}$) as a Function of Scaling Factors (S_{r_1} and S_{c_1}) for a Transfer Function of Weideschel and McAvoy's Column C.

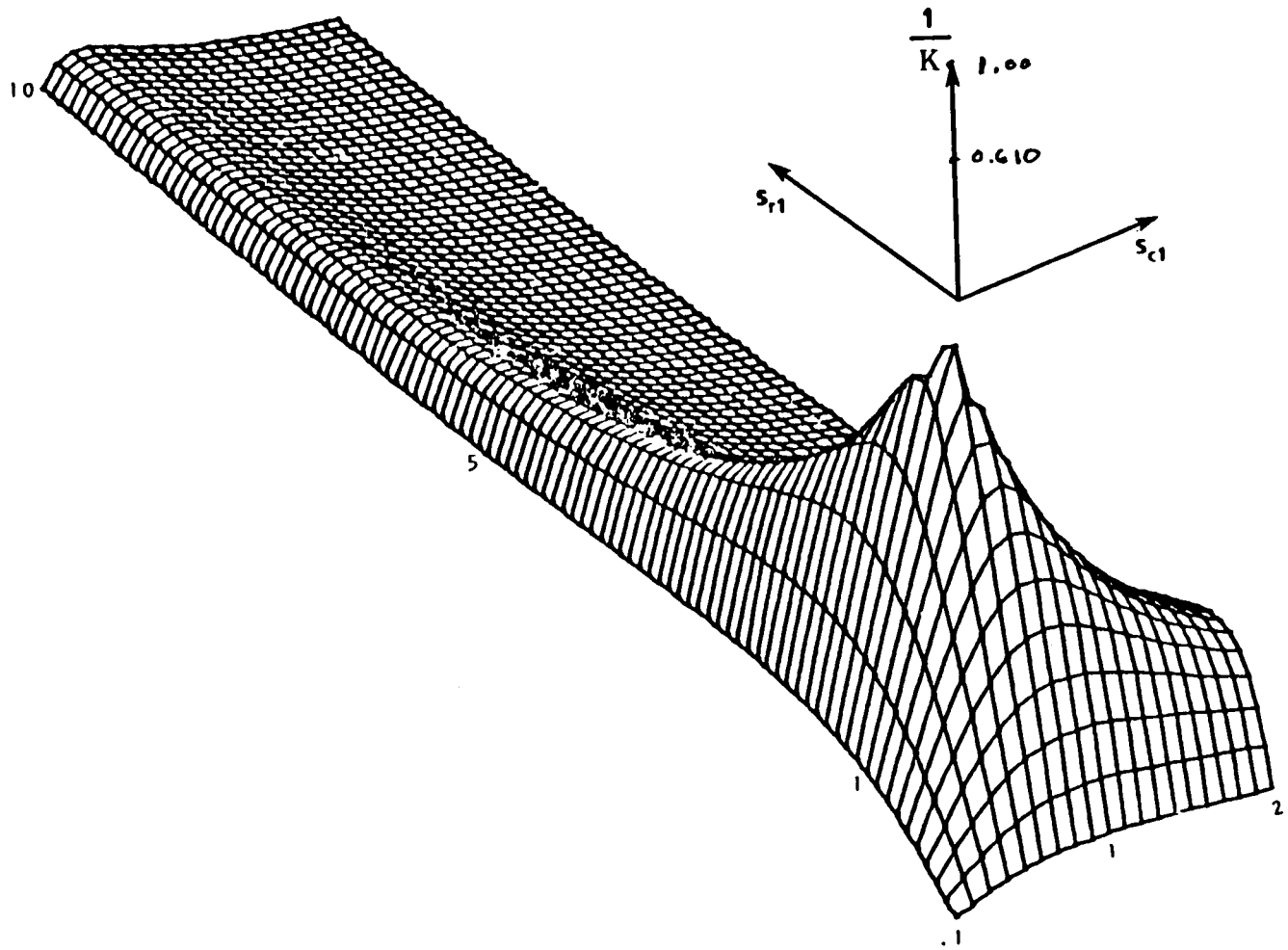


Figure 6.5. The Reciprocal of Condition Number ($\frac{1}{\kappa}$) as a Function of Scaling Factors (S_{c_1} and S_{c_1}) for a Transfer Function of Wood and Berry.

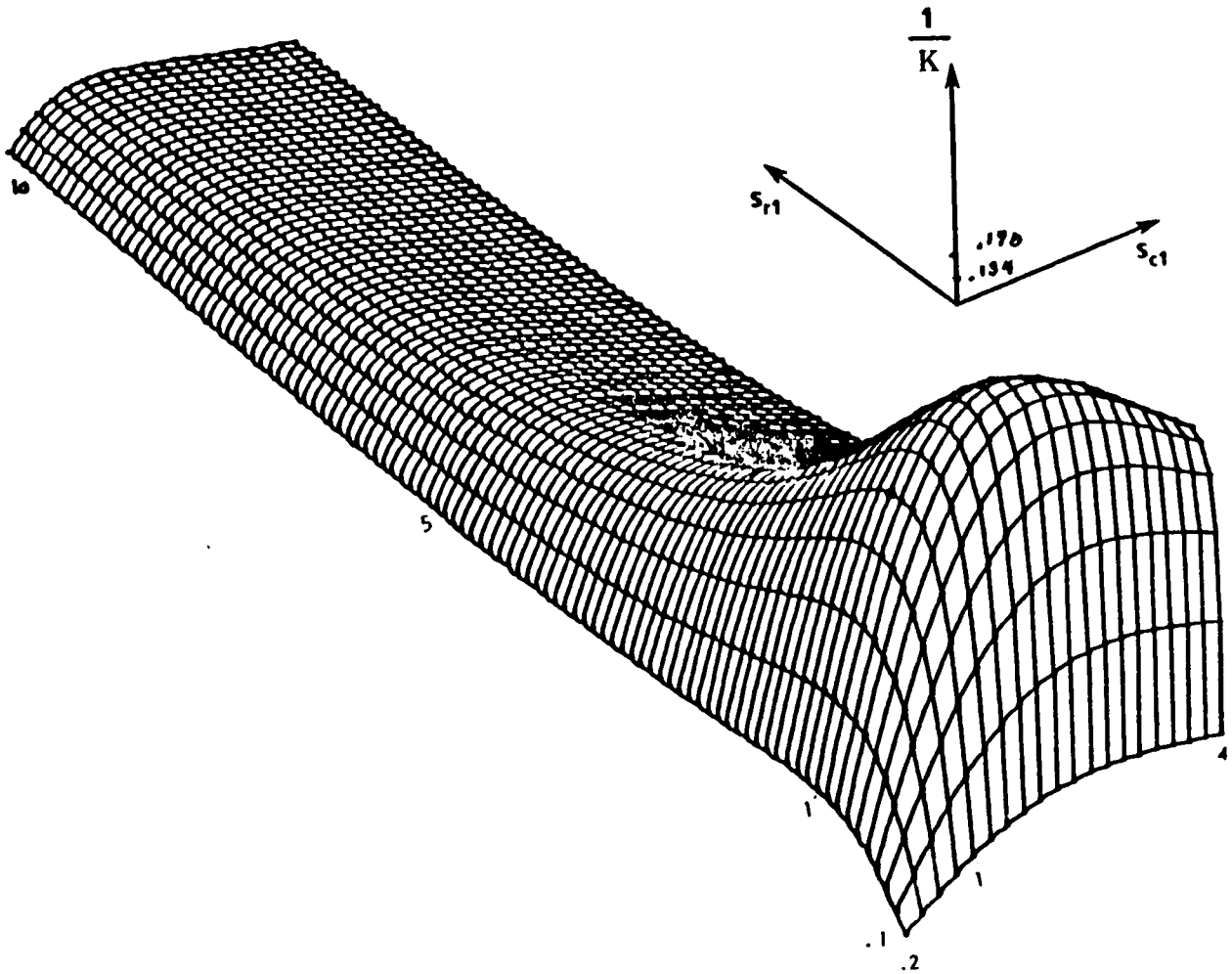


Figure 6.6. The Reciprocal of Condition Number ($\frac{1}{K}$) as a Function of Scaling Factors (S_{r_1} and S_{c_1}) for a Transfer Function of Lau.

the range of partial scaling which is illustrated in Figure 6.7. The row scaling factor S_{r_1} ranges from 0.1 to 10 while the column scaling factor S_{c_1} ranges from 0.1 to 2.

Partial versus Full Scaling

According to Theorem 4, a square gain matrix scaled using NLST will have approximately the same minimum condition number as one fully scaled. This theorem was supported by numerous examples of process gain matrices that were minimally scaled by varying the number of scaling factors required for partial or full scaling while the other parameters were held constant. For these selected full gain matrices, the minimum condition numbers for partial and full scaling were as follows:

GAIN	TYPE OF SCALING	κ_{min}
$\begin{bmatrix} 0.58 & -0.45 \\ 0.35 & -0.48 \end{bmatrix}$	PARTIAL	7.069464
	FULL	7.069464
	STEP = 0.01, NPASS = 3, IO = 1	
$\begin{bmatrix} 1 & 1 & -0.1 \\ 0.1 & 2 & -1 \\ -2 & -3 & 1 \end{bmatrix}$	PARTIAL	24.37407
	FULL	24.37407
	STEP = 0.01, NPASS = 3, IO = 1	

SVA Pairing

Pairing of manipulated variables with measured variables can be done with the SVD of the process gain matrix. A singular value is associated with a right and left singular vector. The left singular vector U corresponds to the process outputs while the right singular

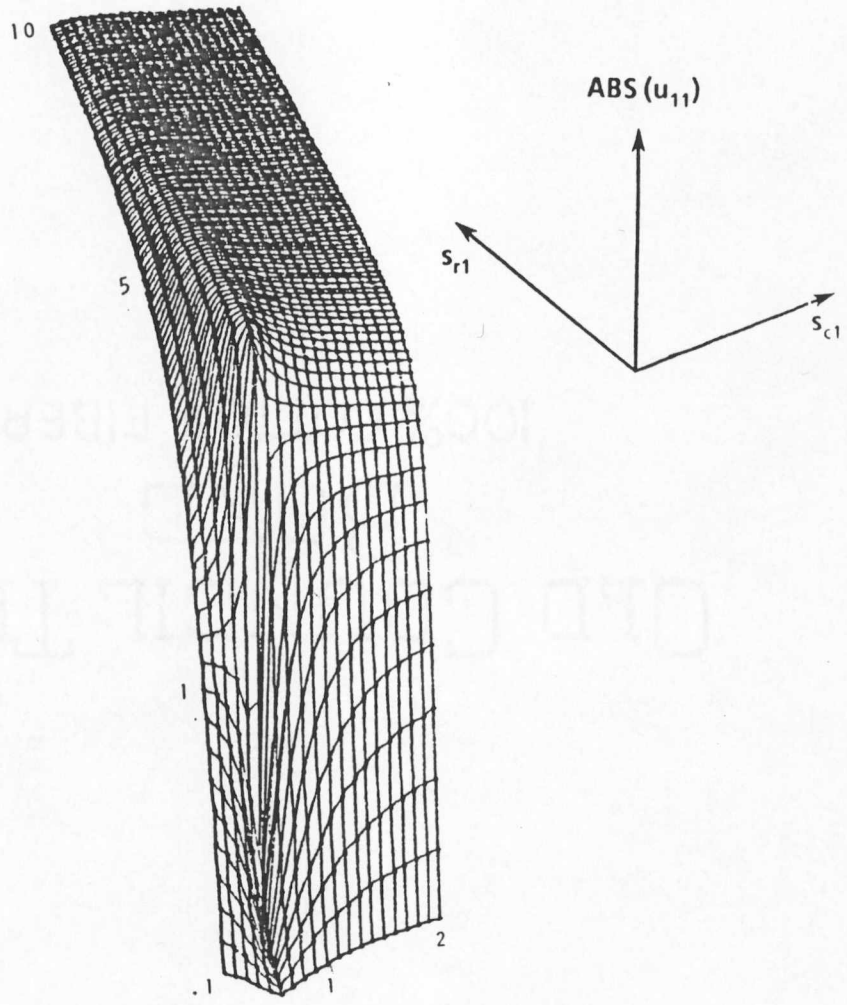


Figure 6.7. The Absolute Value of U_{11} as a Function of Scaling Factors (S_{r_1} and S_{c_1}) for a Transfer Function of Lau.

GAIN

$$\begin{bmatrix} -0.788 & -0.615 \\ -0.615 & 0.788 \end{bmatrix}$$

RGA

$$\begin{bmatrix} 0.621 & 0.379 \\ 0.379 & 0.621 \end{bmatrix} \qquad \begin{matrix} 0 \rightarrow I \\ 1 \quad 1 \\ 2 \quad 2 \end{matrix}$$

SVA

$$U \qquad \Sigma \qquad V^T$$

$$\begin{bmatrix} -0.788 & -0.615 \\ -0.615 & 0.788 \end{bmatrix} \begin{bmatrix} 1.43581 & 0 \\ 0 & 1.43581 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad \begin{matrix} 0 \rightarrow I \\ 1 \quad 1 \\ 2 \quad 2 \end{matrix}$$

where 0 is the output variable
I is the input variable

Comparison of RGA and SVA Pairing

Selected examples of 2x2, 3x3, and 4x4 process gain matrices were scaled using the iterative geometric mean scaling and with NLST for minimum condition number. Table 6.11 shows the pairings of manipulated and measured variables from the RGA, unscaled SVA, geometric mean scaled SVA, and the NLST scaled SVA for each of the selected gain matrices. For these PI, 2x2 matrices, the SVA pairings do not have a physical interpretation because the second manipulative variable is not connected to any of the measured variables of the process gain matrix. It was hypothesized that the pairings for the RGA would be the same as those for the NLST scaled SVA for all types of matrices, but this prediction will be shown not to hold.

For noninteracting matrices of this table, the RGA pairings are in agreement with those of the NLST scaled SVA and the GM scaled

Table 6.11. Comparison of RGA and SVA for Scaled and Unscaled Matrices

Gain	Type	RGA	SVA_{un}	SVA_{GM}	$SVA_{\kappa min}$
		2 x 2			
		(0→I)	(U→V)	(U→V)	(U→V)
$\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}$	NI	1-1	1-1	1-1	1-1
		2-2	2-2	2-2	2-2
$\begin{pmatrix}1 & 0\\ 1 & 1\end{pmatrix}$	PI	1-1	2-1	2-1	1-1*
		2-2	1-2	1-2	2-2
$\begin{pmatrix}1 & 0\\ 3 & 4\end{pmatrix}$	PI	1-1	2-2	2-1	1-1
		2-2	1-1	1-2	2-2
$\begin{pmatrix}0.58 & -0.45\\ 0.35 & -0.48\end{pmatrix}$	FULL	1-1	None	None	None
		2-2			
		3 x 3			
$\begin{pmatrix}3 & 0 & 0\\ 0 & 7 & 0\\ 0 & 0 & 5\end{pmatrix}$	NI	1-1	2-2	1-1	1-1
		2-2	3-3	2-2	2-2
		3-3	1-1	3-3	3-3
$\begin{pmatrix}0 & 1 & 1\\ 1 & 0 & 0\\ 0 & 1 & 0\end{pmatrix}$	PI	2-1	1-2	1-2	1-2*
		3-2	2-3	2-1	2-1
		1-3	3-3	3-3	3-3*
$\begin{pmatrix}0 & 4 & 5\\ 3 & 0 & 0\\ 0 & 7 & 0\end{pmatrix}$	PI	2-1	1,3-2	1-2	1-2,3
		3-2	1,3-3	2-1	2-1
		1-3	2-1	3-3	3-2,3
$\begin{pmatrix}1. & 1. & -0.1\\ 0.1 & 2. & -1.\\ -2. & -3. & 1.\end{pmatrix}$	FULL	3-1	3-2	1-1,2,3	3-2
		1-2	2-1	2-3	2-1
		3-3	1-3	3-2	3-2

Table 6.11 (continued)

Gain	Type	RGA	SVA_{un}	SVA_{GM}	$SVA_{\kappa min}$
		4 x 4			
$\begin{pmatrix}1&0&0&0\\0&6&0&0\\0&0&5&0\\0&0&0&3\end{pmatrix}$	NI	1-1	2-2	1-1	1-1
		2-2	3-3	2-2	2-2
		3-3	4-4	3-3	3-3
		4-4	1-1	4-4	4-4
$\begin{pmatrix}1&0&0&0\\0&1&1&0\\0&0&1&1\\0&0&0&1\end{pmatrix}$	PI	1-1	3-3	3-3	3-3
		2-2	2-4	2-4	1-1
		3-3	1-1	1-1	2-2
		4-4	4-2	4-2	3-2
$\begin{pmatrix}1&0&0&0\\0&6&7&0\\0&0&5&2\\0&0&0&3\end{pmatrix}$	PI	1-1	2-3	3-3	3-3
		2-2	3-4	2-4	2,4-4
		3-3	4-2	1-1	1-1
		4-4	1-1	3-2	3-3
$\begin{pmatrix}3.3&5.6&7.6&4.5\\4.5&7.6&5.6&3.3\\-12.4&3.6&7.5&4.7\\4.7&7.5&3.6&-12.4\end{pmatrix}$	FULL	3-1	3,4-1,4	1,2-3,4	1,2-2,3
		2-2	1,2-2,3	1,2-1,4	3,4-3,4
		1-3	3,4-3,4	3,4-1,4	3,4-1,4
		4-4	1,2-2,3	1,2-2,3	1,2-2,3

*Corresponding elements of $|u_{ij}|$ and $|v_{ij}|$ have a percent difference $\leq$ 14.12%.

+The comma denotes that several inputs or outputs are paired with corresponding outputs or inputs.

vector V corresponds to the process inputs. Permutations in the manipulated and measured variables of the process gain matrix are associated with interchanges in the rows of U and V. Thus, columns of the process gain matrix are used for pairing. The largest singular value is paired with the elements with the largest absolute value in the first columns of U and V. Then the next largest singular value is paired with the elements of U and V that have the next largest absolute value in the second columns of these matrices. Quantitative or semi-quantitative rules do not exist for SVA as for RGA. McAvoy's rules for RGA pairing were listed in Chapter 2. Some of the main rules for RGA pairing are summarized as follows:

1. RGA elements close to one are first to be considered for pairing.
2. The stability of the resulting pairing should be checked using Niederlinski's theorem. If this theorem shows that if the pairing $r_1 - y_1$, $r_2 - y_2$,..., $r_n - y_n$ for the closed loop system is unstable if

$$\frac{|K|}{\prod_{i=1}^{n} K_{ii}} < 0 \qquad (6.20)$$

   where $|K|$ is the determinant of the steady state gain matrix if this pairing is unstable, then another positive pairing is chosen.
3. Whenever possible negative pairings should be avoided. For example the RGA and SVA pairing for this matrix is

SVA. The unscaled SVA matrices have the same inputs and outputs paired as those for the NI, RGA matrices except for the order of the pairings. This is an expected result since condition number is equal to one when these matrices are scaled by geometric mean or by NLST. Two by two, PI matrices have the same pairing by RGA as by NLST scaled SVA in both the base scaled and not base scaled forms whereas the results for unscaled SVA and GM scaled SVA differ. The pairing by RGA and the NLST scaled SVA for the remaining examples of 3x3 and 4x4 matrices are not the same. For most full 2x2 matrices, there is no pairing after geometric or NLST scaling since $U \varepsilon \mathcal{U}_I$ and $V \varepsilon \mathcal{V}_I$. Thus, it appears that there is no relationship between RGA and NLST scaled SVA when the matrix is scaled for minimum condition number except for 2x2, PI or NI, square matrices that are of any dimension.

Comparison of Interaction Measures with An Extruder

The transfer function of a 40 mm diameter screw extruder used for blow injection of LDPE (65) was analyzed to determine if there were relationships between pairings of the RGA, unscaled SVA, and minimally scaled SVA. The steady state gain matrix consisted of four manipulated variables and two controlled variables as is shown below:

$$\begin{array}{c} \\ P_m \\ T_m \end{array} \begin{array}{c} \begin{array}{cccc} \theta & \Omega & Tw & Tw_5 \end{array} \\ \begin{bmatrix} 0.15 & 0.11 & -0.132 & -0.01 \\ 0.045 & 0.118 & +0.928 & 0.445 \end{bmatrix} \end{array}$$

where T_m melt temperature at the die inlet, °C

P_m melt pressure at the die inlet, $\frac{MN}{m^2}$

θ die restrictor angle, deg

Ω screw speed, rpm

T_w barrel wall temperature, °C

T_{w_5} set point temperature at the die, °C

Six combinations of 2x2 matrices were analyzed from the steady state gain matrix. These subsystems are listed as follows:

GAIN NO.

$$\begin{array}{c} \\ P_m \\ T_m \end{array} \begin{array}{c} \begin{array}{cc} \Omega & T_w \end{array} \\ \begin{bmatrix} 0.11 & -0.132 \\ 0.118 & 0.928 \end{bmatrix} \end{array} \quad 1$$

$$\begin{array}{c} \\ P_m \\ T_m \end{array} \begin{array}{c} \begin{array}{cc} \theta & T_{w_5} \end{array} \\ \begin{bmatrix} 0.15 & -0.01 \\ 0.045 & 0.445 \end{bmatrix} \end{array} \quad 2$$

$$\begin{array}{c} \\ P_m \\ T_m \end{array} \begin{array}{c} \begin{array}{cc} T_w & T_{w_5} \end{array} \\ \begin{bmatrix} -0.132 & -0.01 \\ 0.928 & 0.445 \end{bmatrix} \end{array} \quad 3$$

GAIN NO.

$$\begin{array}{c} \\ P_m \\ T_m \end{array} \begin{array}{c} \begin{array}{cc} \Omega & T_{w_5} \end{array} \\ \begin{bmatrix} 0.11 & -0.01 \\ 0.118 & 0.445 \end{bmatrix} \end{array} \quad 4$$

$$\begin{array}{c} \\ P_m \\ T_m \end{array} \begin{array}{c} \begin{array}{cc} \theta & T_w \end{array} \\ \begin{bmatrix} 0.15 & -0.132 \\ 0.045 & 0.928 \end{bmatrix} \end{array} \quad 5$$

$$\begin{array}{c} \\ P_m \\ T_m \end{array} \begin{array}{c} \begin{array}{cc} \theta & \Omega \end{array} \\ \begin{bmatrix} 0.15 & 0.11 \\ 0.045 & 0.118 \end{bmatrix} \end{array} \quad 6$$

By choosing square subsystems of the gain matrix, the RGA which is limited to square systems that can be directly compared with the SVA. The unscaled SVA obtained from the gain matrix indicates that

the second controlled variable (Tm) should be paired with the third manipulated variable (Tw) and that the first controlled variable (Pm) should be paired with the first manipulated variable (θ). The comparison of possible loop pairings of the RGA, unscaled SVA, and minimally scaled SVA is shown in Table 6.12.

Interpretation of these results is difficult. One case (3) of the unscaled SVA pairings disagrees with that of the RGA. Three cases of the minimally scaled SVA pairings correspond with the RGA.

Conclusions and Recommendations

With the objective of minimizing condition number, NSLT was found to be more effective than GM for 3x3 and 4x4 process gain matrices. For 2x2, full process gain matrices, GM was as efficient in lowering the condition number as NLST. There appears to be no relationship between RGA pairings and the SVA pairings that correspond to the minimally scaled matrix except for NI, 2x2, 3x3, and 4x4 matrices and PI, 2x2 matrices. The minimum condition number is somewhat difficult to determine with certainty. Using NLST an apparent minimum can be local or global. Grosdidier, Morari, and Holt's bound of 2 $\max[||\Lambda||_1 ||\Lambda||_\infty]$ was used as a guide for minimum condition number as well as running the NLST consecutively while decreasing the step size by a factor of ten with the other parameters held constant.

The question can be raised whether the optimal or minimum condition number is a suitable criterion for scaling. Grosdidier et al. state that "optimal scaling ensures that the least conservative

Table 6.12. Comparison of Interaction Measures for a Plastics Extruder

No.	RGA	RGA_{int}	SVA_{un}	$SVA_{un\ int}$	$SVA_{\kappa min}$	$SVA_{\kappa min\ int}$
1	y_1-m_1 y_2-m_2	$P_m-\Omega$ T_m-T_w	y_2-m_2 y_1-m_1	T_m-T_w $P_m-\Omega$	None	-
2	y_1-m_2 y_2-m_1	$T_m-T_{w_5}$ $P_m-\Omega$	y_1-m_2 y_2-m_1	$T_m-T_{w_5}$ $P_m-\Omega$	y_2-m_1 y_1-m_2	$P_m-\Omega$ $T_m-T_{w_5}$
3	y_1-m_1 y_2-m_2	P_m-T_w $T_m-T_{w_5}$	y_2-m_1 y_1-m_2	T_m-T_w $P_m-T_{w_5}$	y_2-m_1 y_1-m_2	$P_m-T_{w_5}$ T_m-T_w
4	y_1-m_2 y_2-m_1	$T_m-T_{w_5}$ $P_m-\theta$	y_1-m_2 y_2-m_1	$T_m-T_{w_5}$ $P_m-\theta$	y_1-m_2 y_2-m_1	$T_m-T_{w_5}$ $P_m-\theta$
5	y_1-m_1 y_2-m_2	$P_m-\theta$ T_m-T_w	y_2-m_2 y_1-m_1	T_m-T_w $P_m-\theta$	None	-
6	y_2-m_1 y_1-m_2	$P_m-\theta$ $T_m-\Omega$	y_2-m_1 y_1-m_2	$P_m-\theta$ $T_m-\Omega$	None	-

int = interpretation; y = output variable; m = manipulated variable.

value of the condition number is obtained and it should not be given a physical interpretation." A minimally scaled matrix may be mathematically related to the original unscaled matrix although the physical system that it represents may be altered by the scaling. By Theorem 7 which involves the scalar multiplication of a matrix, positive scaling can be performed in an infinite number of ways. Which scaling is best for a particular process gain matrix is a poorly understood problem.

More effort needs to be placed on the relationship of conditioning and interaction. A noninteracting process gain matrix that has a low condition number does not present a control problem. In most instances that type of structure could be controlled in the open loop. A noninteracting process gain matrix with a high condition number can be scaled to a condition number of one. Such a process could be controlled with a gain compensation system. A process that is highly interacting but with a low condition number can be controlled with an SVA controller (9) or a structural compensator (22) that decouples the interaction in the process. A process that is highly interacting and has a high condition number will need both a structural compensator and a gain compensator (22) and will be quite difficult to control over the entire frequency range.

It is recommended that the iterative geometric mean scaling and the NLST could be extended to nonsquare process gain matrices. These two programs would be easier to use and understand if they

were made interactive and if they used structured programming. A different algorithm for the nonlinear least squares technique that is more efficient (i.e. requires less storage space in computer memory) is needed for scaling process gain matrices with dimension greater than 4x4.

Extending the SVD to the frequency domain would provide more information about robustness and the dynamic behavior of the system.

Further work is needed to examine the interaction present in multivariable systems on a consistent basis. The RGA and SVD provide a rough quantitative means for obtaining information about a process steady state gain matrix.

BIBLIOGRAPHY

BIBLIOGRAPHY

1. Prasad, J., "Singular Value Analysis and Scaling Applications for Multivariable Process Design and Control," Ph.D. Dissertation, University of Tennessee, Knoxville, Tennessee (1984).

2. McAvoy, T. J., Interaction Analysis: Principles and Applications, Instruments Society of America, Research Triangle Park, North Carolina (1983).

3. Jensen, N., Fisher, D. G., and S. L. Shah, "Interaction Analysis in Multivariable Control Systems," AIChE Journal, 32, 6, 959-970 (1986).

4. Macfarlane, A. G. J., "A Survey of Some Recent Results in Linear Multivariable Feedback Theory," Automatica, 8, 455-492 (1972).

5. Rosenbrock, H. H., "On the Design of Linear Multivariable Control Systems," Proc. 3rd IFAC Congress, London 1A1-1A16 (1966).

6. Foss, A. S., "Critique of Chemical Process Control Theory," AIChE J., 13, 209-214 (1973).

7. Rijnsdorp, J. E., and D. E. Seborg, "A Survey of Experimental Applications of Multivariable Control to Process Control Problems," Chemical Process Control, A. S. Foss and M. M. Denn, eds., AIChE Symposium Series, 72 (159), 112-123 (1976).

8. Bristol, E. H., "On a New Measure of Interaction for Multivarible Process Control," IEEE Trans. on Auto Control, AC-11, 133-134 (1966).

9. Smith, C. R., "Multivariable Control Using Analysis Procedures," Ph.D. Dissertation, University of Tennessee, Knoxville, Tennessee (1981).

10. Shinskey, F. G., Process Control Systems, McGraw-Hill, New York (1979).

11. Rijnsdorp, J. E., "Interaction in Two-variable Control Systems for Distillation Columns I," Automatica, 1, 15 (1965).

12. Davison, E. J., and F. T. Mann, "Interaction Index for Multivariable Control Systems," Automatica, 791-799 (1969).

13. Davison, E. J., "A Nonminimum Phase Index and its Application to Interacting Multivariable Control Systems," Automatica, 5, 791-799 (1969).

14. Witcher, M. F., and T. J. McAvoy, "Interacting Control Systems: Steady State and Dynamic Measurement of Interaction," ISA Trans., 16(3), 35-41 (1977).

15. Nisenfeld, A. E., and H. M. Schultz, "Interaction Analysis Applied to Control System Design," Instrumentation Technology, 52-57 (April 1971).

16. Bristol, E. H., "RGA 1977: Dynamic Effects of Interaction," Dec. Cont. Conf., New Orleans, FP4, 1096-1106 (1977).

17. Noble, B., and J. W. Daniel, Applied Linear Algebra, Prentice-Hall, Inc., Englewood Cliffs, New Jersey (1977).

18. Bruns, D. D., and C. R. Smith, "Singular Value Analysis: A Geometrical Structure for Multivariable Control," Presented at AIChE Winter Meeting, Orlando, Florida (February 1982).

19. Klema, V. C., and A. J. Laub, "The Singular Value Decomposition: Its Computation and Some Applications," IEEE Trans. Auto. Contr., AC-25, 164-176 (1980).

20. Morari, M., "Flexibility and Resiliency of Process Systems," Proc. Int. Symp. Proc. Sys. Eng., 223, Kyoto (1982).

21. Lau, H., "Studies on the Control of Integrated Chemical Processes," Ph.D. Dissertation, University of Minnesota, Minneapolis, Minnesota (1982).

22. Young, D. M., and R. T. Gregory, A Survey of Numerical Mathematics, Vol. 2, Addison-Wesley, California (1973).

23. Strang, G., Linear Algebra and its Applications, Academic Press, New York (1976).

24. Moore, B. C., "Singular Value Analysis of Linear Systems, Parts I, II," Systems Control Report. 7601 and 7602, Dept. of Electrical Engineering, University of Toronto, Ontario (July 1978).

25. Anton, H., Elementary Linear Algebra, John Wiley & Sons, New York (1981).

26. Moore, C. F., "Applications of Singular Value Decomposition to the Design, Analysis, and Control of Industrial Processes," IEEE Proceedings, 2, 643-650 (1986).

27. Stewart, G. W., Introduction to Matrix Computations, Academic Press, New York (1973).

28. Golub, G. H., and J. H. Wilkinson, "Ill-Conditioned Eigensystems and the Computations of the Jordan Canonical Form," SIAM Review, 576-613 (1976).

29. Golub, G. H., and W. Kahan, "Calculating the Singular Values and Pseudo-Inverse of a Matrix," SIAM J. Numer. Anal., 2, 205-224 (1965).

30. Forsythe, G. E., Malcolm, M. A., and C. B. Moler, Computer Methods for Mathematical Computations, Prentice-Hall, Inc., Englewood Cliffs, New Jersey (1977).

31. Smith, B. T., Matrix Eigensystem Routines-EISPAC Guide, Springer-Verlag, New York (1970).

32. Golub, G. H., and C. Reinsch, "Singular Value Decomposition and Least-Squares Solutions," in Handbook for Automatic Computations, Vol. II (J. H. Wilkinson and C. Reinsch eds.), 134-151 Springer-Verlag, New York (1971).

33. Lawson, C. L., and R. J. Hanson, Solving Least Squares Problems, Prentice-Hall, New Jersey (1974).

34. Francis, J. G. F., "The QR Transformation: A Unitary Analogue to the LR Transformation-Part 1," Computer Journal, 4, 265-271 (1961).

35. Rutishauser, H., "Solution of Eigenvalue Problems with the LR Transformation," in Further Contributions to the Solution of Simultaneous Linear Equations and the Determination of Eigenvalues, Nat. Bur. of STD. AMS 49, 921 (1958).

36. Francis, J. G. F., "The QR Transformation-Part 2," Computer Journal, 4, 332-345 (1966).

37. Kublanovskaya, V. M., "On Some Algorithms for the Solution of the Complete Eigenvalue Problem," Zh. vych. mat., 1, 555-570, 921 (1961).

38. Gill, P., Murray, W., and M. H. Wright, Practical Optimization, Academic Press, Inc., New York (1981).

39. Bauer, F. L., "Optimal Scaling of Matrices and the Importance of the Minimal Condition," Proc. IFIP Congress, North Holland (1962).

40. Bauer, F. L., "Optimally Scaled Matrix," Numer. Math., 5, 73-87 (1983).

41. van der Sluis, A., "Condition Numbers and Equilibration of Matrices, Numer. Math., 14, 14-23 (1969).

42. Businger, P. A., "Matrices Which Can Be Optimally Scaled," Numer. Math., 12, 346-348 (1968).

43. Golub, G. H., and C. F. van Loan, "An Analysis of the Total Least Squares Problem," SIAM J. Numer. Anal., 17(6) 883-893 (1980).

44. Bauer, F., "Remarks on Optimally Scaled Matrices," Numer. Math., 13, 1-3 (1969).

45. Tomlin, J. A., "On Scaling Linear Programming Problems," Mathematical Programming Study, 4, 146-166 (1975).

46. Curtis, A. R., and J. K. Reid, "On the Automatic Scaling of Matrices for Gaussian Elimination," Journal of the Institute of Mathematics and Its Applications, 10, 118-124 (1972).

47. Fulkerson, D. R., and P. Wolfe, "An Algorithm for Scaling Matrices," SIAM Review, 4, 142-146 (1962).

48. Hamming, R. W., Introduction to Applied Numerical Analysis, McGraw-Hill, New York (1971).

49. Orchard-Hays, W., Advanced Linear Programming Computing Techniques, McGraw-Hill, New York (1969).

50. Saunders, B. D., and H. Schneider, "Cones, Graphs, and Optimal Scalings of Matrices," Linear and Multilinear Algebra, 8, 121-135 (1975).

51. Rothblum, U. G., and H. Schneider, "Characterization of Optimal Scalings of Matrices," Mathematical Programming, 19, 121-136 (1980).

52. Nash, D., "The Generalized Eckart-Young Principal Axis Theorem," Linear and Multilinear Algebra, 2, 353-355 (1975).

53. Dongarra, J. J., Bunch, J. R., Moler, C. B., and G. W. Stewart, LINPACK Users' Guide, Philadelphia, PA SIAM (1979).

54. Williams, T. W. C., "Obtaining Minimal Gershgorin Discs by Scaling the States," Int. J. of Control, 42(5), 1155-1173 (1985).

55. Wilkinson, J. H., The Algebraic Eigenvalue Problem, Oxford University Press (1956).

56. Forsythe, G., and Moler, C. B., Computer Solution of Linear Algebraic Systems, Prentice-Hall, Inc., Englewood Cliffs, New Jersey (1967).

57. Partlett, B. N., and C. Reinsch, "Balancing a Matrix for Calculation of Eigenvalues and Eigenvectors," Numer. Math., 13, 293-304 (1969).

58. Bonvin, D., and D. A. Mellichamp, "Rescaling State and Input Variables for the Purpose of Analyzing Process Models," Proc. IEEE, 1, 626-631 (1985).

59. Lau, H., and K. F. Jensen, "Evaluation of Changeover Control Policies by Singular Value Analysis-Effects of Scaling," AIChE J., 31(1), 135-146 (1985).

60. Grosdidier, P., Morari, M., and B. R. Holt, "Closed-Loop Properties From Steady-State Gain Information," Ind. Eng. Chem. Fundam., 24, 221-235 (1985).

61. Wilde, D. J., Optimum Seeking Methods, Prentice-Hall, Inc., Englewood Cliffs, New Jersey (1964).

62. Moore, C. F., Smith, C. L., and P. W. Murill, "Multidimensional Optimization using Patern Search," IBM Share Library, PID No. 6448 (1968).

APPENDICES

APPENDIX A

FURTHER WORK ON GEOMETRIC MEAN SCALING

APPENDIX A

FURTHER WORK ON GEOMETRIC MEAN SCALING

Another scheme for the iterative geometric mean scaling for optimizing condition number was briefly considered. The scaling algorithm for row scaling was as follows:

first row

$GM = (a_{11}, a_{12}, a_{13}, \ldots, a_{1n})$

first column

$GM = (a_{11}, a_{21}, a_{31}, \ldots, a_{m1})$

Preliminary results were inconclusive for this algorithm for 3x3 matrices.

A procedure for generating full, nxn orthogonal matrices would be useful as an analytical tool. The primary steps generating for a 4x4 matrix are listed below:

1. Nine elements would be designated randomly.
2. Each row and/or column would be made orthonormal.
3. The matrix would be tested to determine if it were orthogonal ($AA^T=I$).
4. If the matrix were orthogonal, then it would be printed out. If it were not orthogonal, steps one through four would be repeated.

APPENDIX B

FLOWCHART AND TABLE FOR ITERATIVE GEOMETRIC MEAN SCALING AND FLOWCHART FOR NLST SCALING

Table B.1. Sections of an Iterative Geometric Mean Scaling Technique

Sections	Purpose
MAIN PROGRAM	
SUBROUTINES	
GMRC	ROW FOLLOWED BY COLUMN GEOMETRIC MEAN SCALING
GMCR	COLUMN COLLOWED BY ROW GEOMETRIC MEAN SCALING
ARETRN	RETURNS THE ORIGINAL GAIN MATRIX
ACHANG	STORES THE ORIGINAL GAIN MATRIX
DSVDC	PERFORM SINGULAR VALUE DECOMPOSITION
SVDOUT	PRINTS OUTPUT OF DSVDC

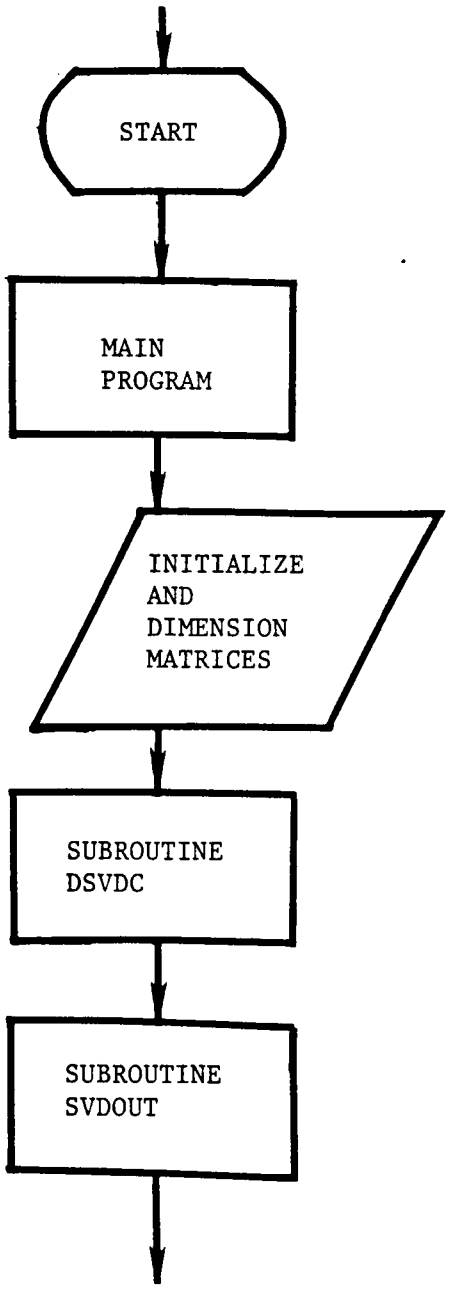
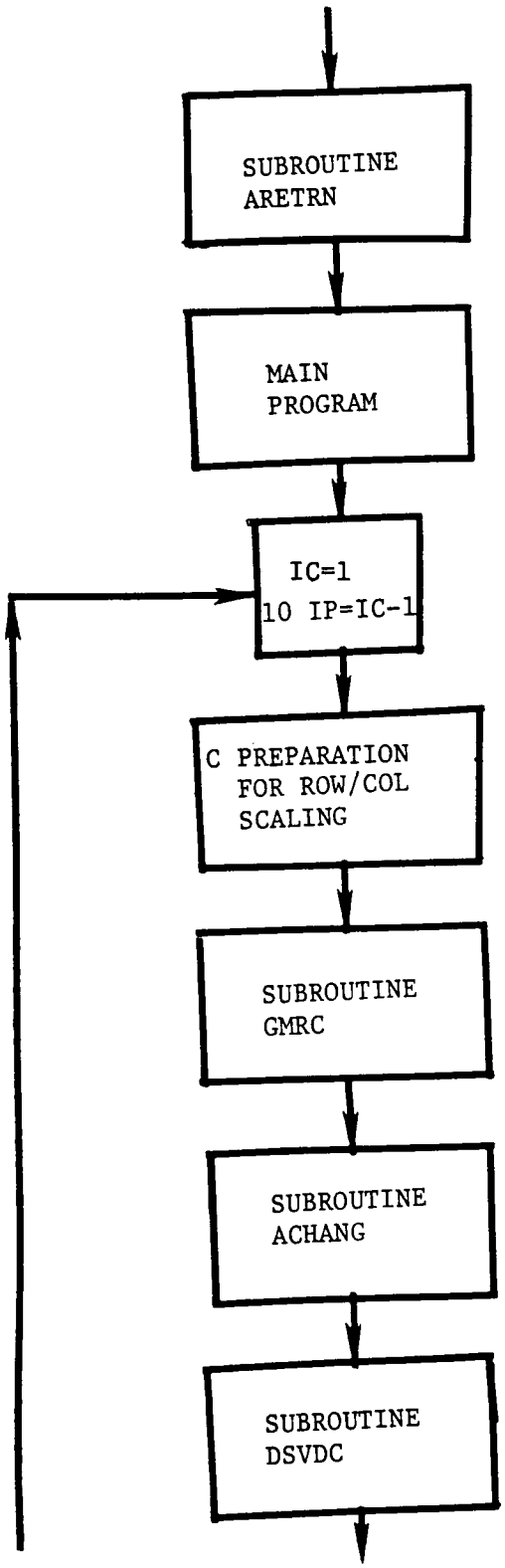
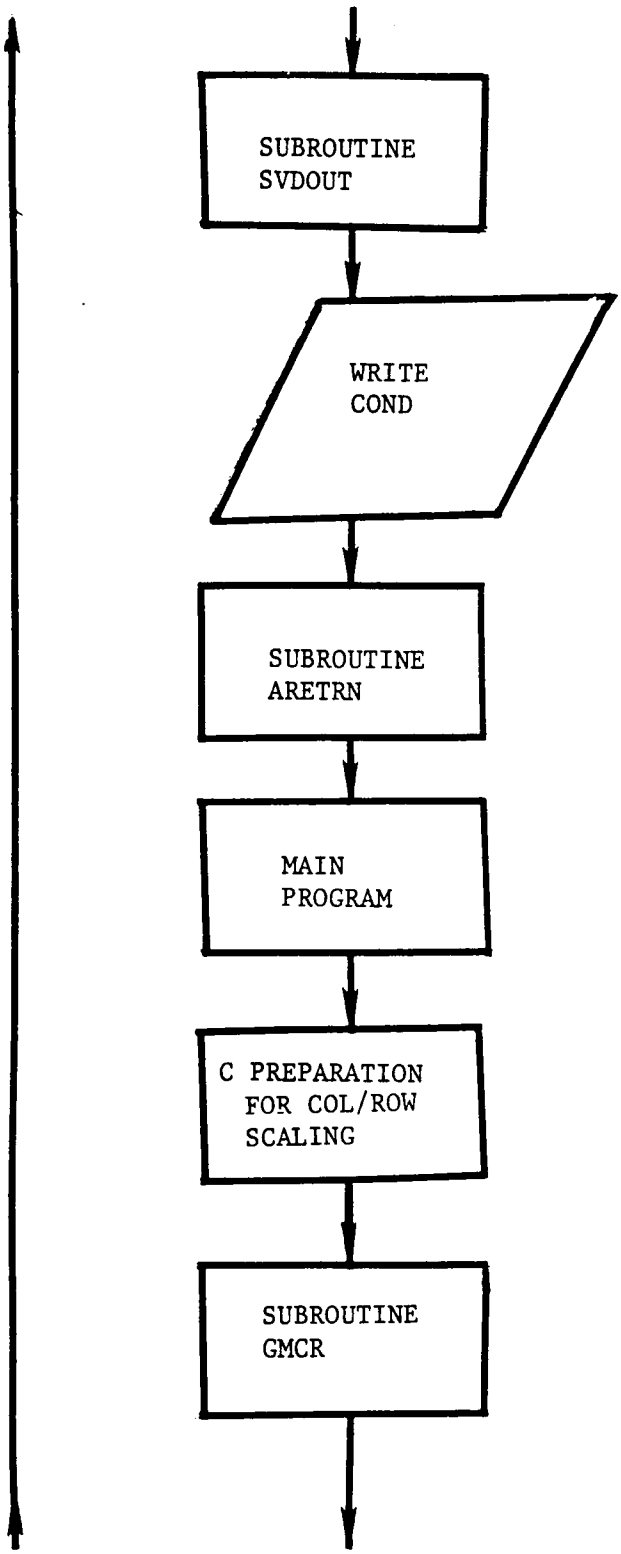
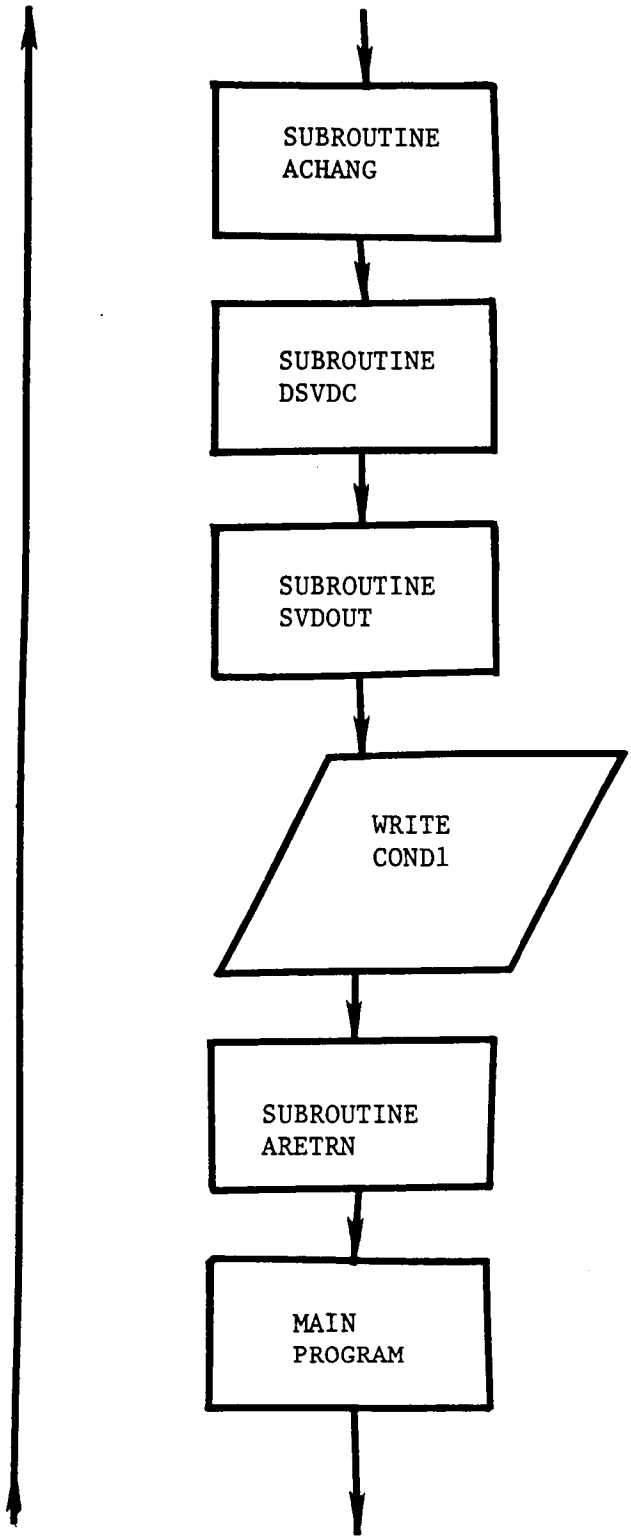
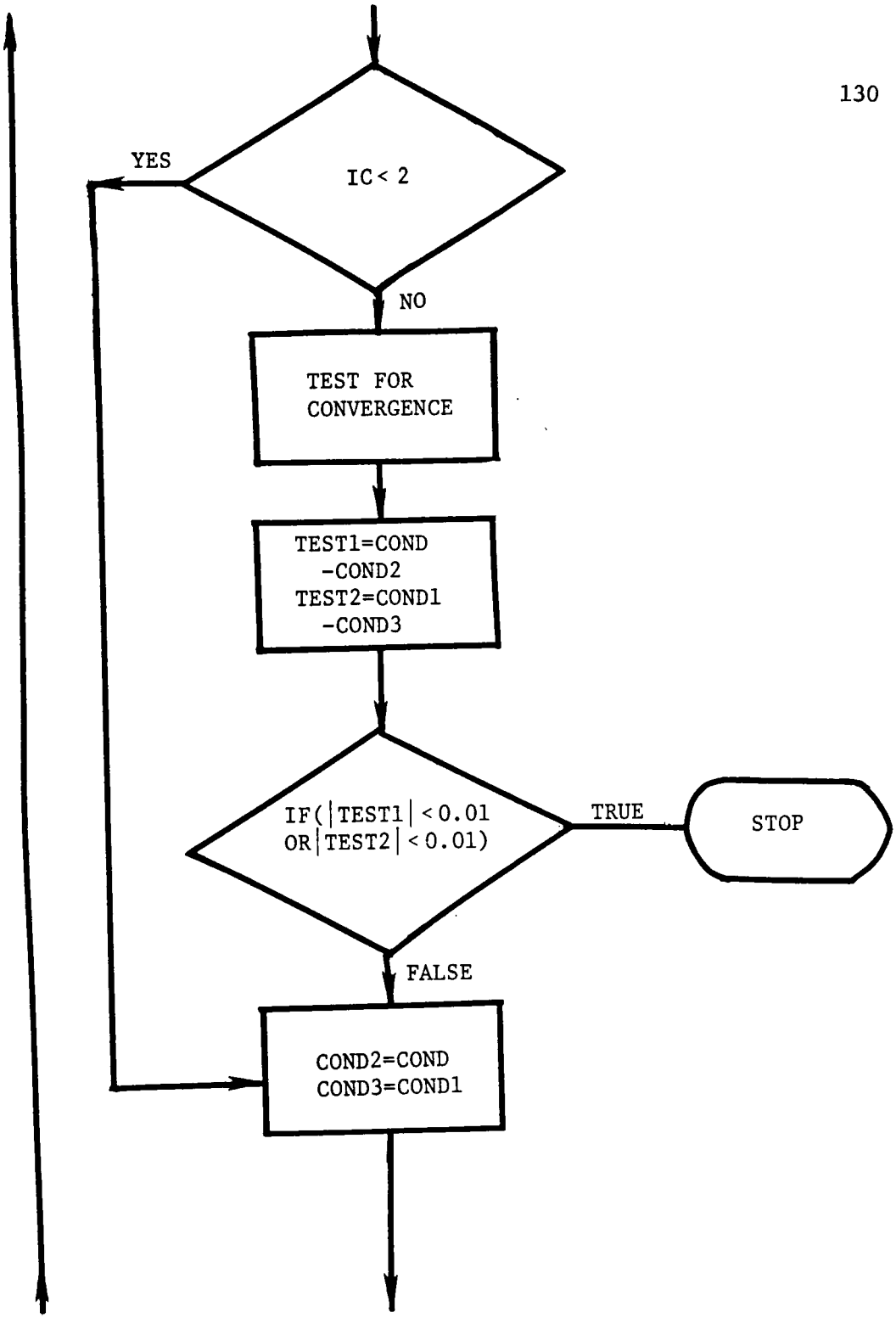


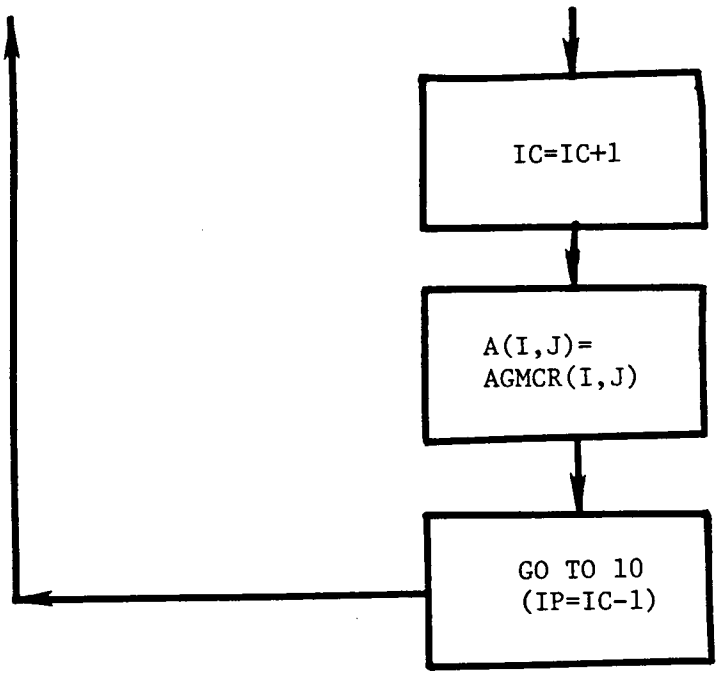
Figure B.2. A Simplified Flowchart of Iterative Geometric Mean Scaling.


SUBROUTINE
ARETRN
MAIN
PROGRAM
IC=1
10 IP=IC-1
C PREPARATION
FOR ROW/COL
SCALING
SUBROUTINE
GMRC
SUBROUTINE
ACHANG
SUBROUTINE
DSVDC


SUBROUTINE
SVDOUT
WRITE
COND
SUBROUTINE
ARETRN
MAIN
PROGRAM
C PREPARATION
FOR COL/ROW
SCALING
SUBROUTINE
GMCR


SUBROUTINE
ACHANG
SUBROUTINE
DSVDC
SUBROUTINE
SVDOUT
WRITE
COND1
SUBROUTINE
ARETRN
MAIN
PROGRAM


IC < 2
YES
NO
TEST FOR CONVERGENCE
TEST1=COND -COND2
TEST2=COND1 -COND3
IF(|TEST1| < 0.01 OR|TEST2| < 0.01)
TRUE
STOP
FALSE
COND2=COND
COND3=COND1


IC=IC+1
A(I,J)=
AGMCR(I,J)
GO TO 10
(IP=IC-1)

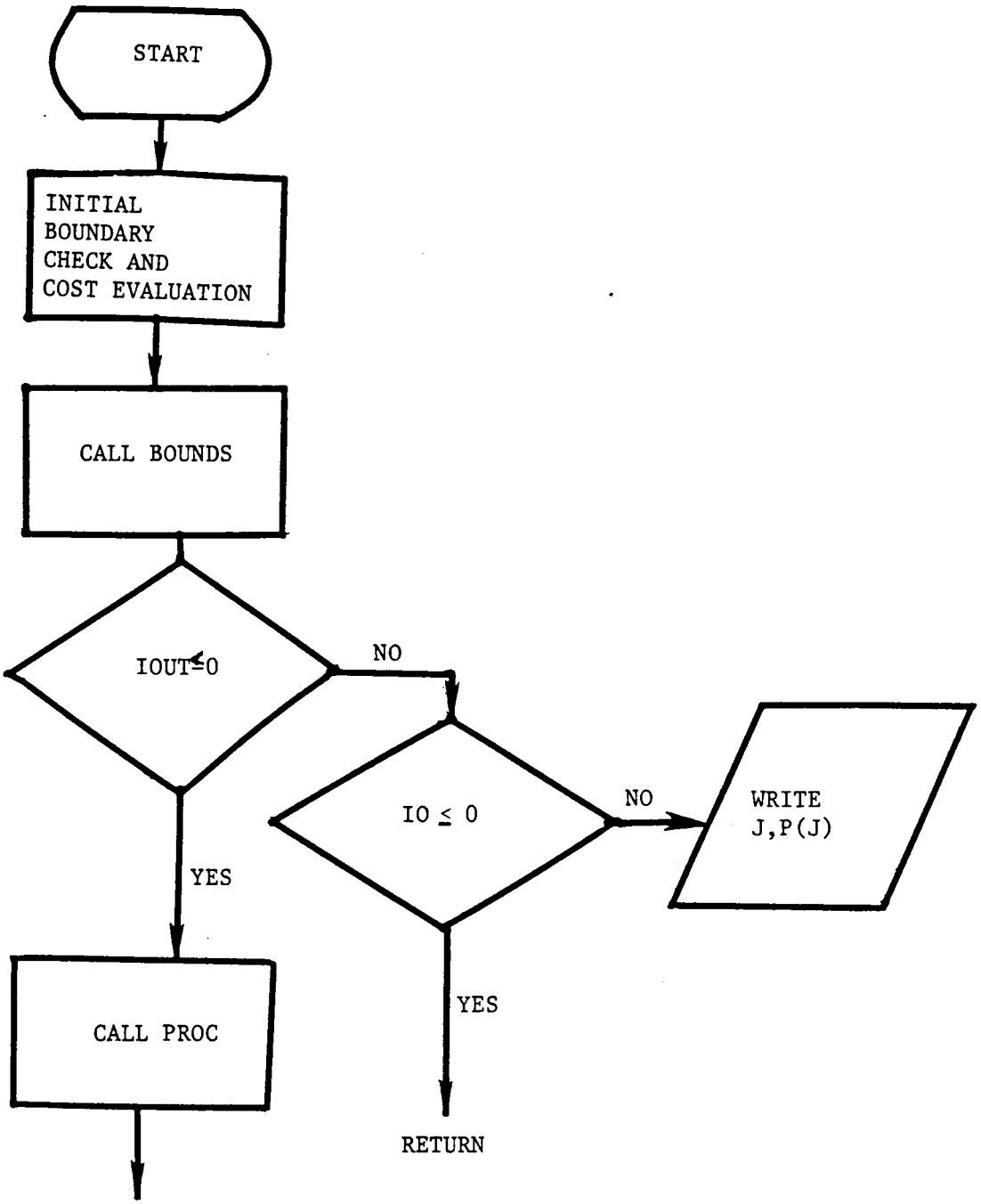
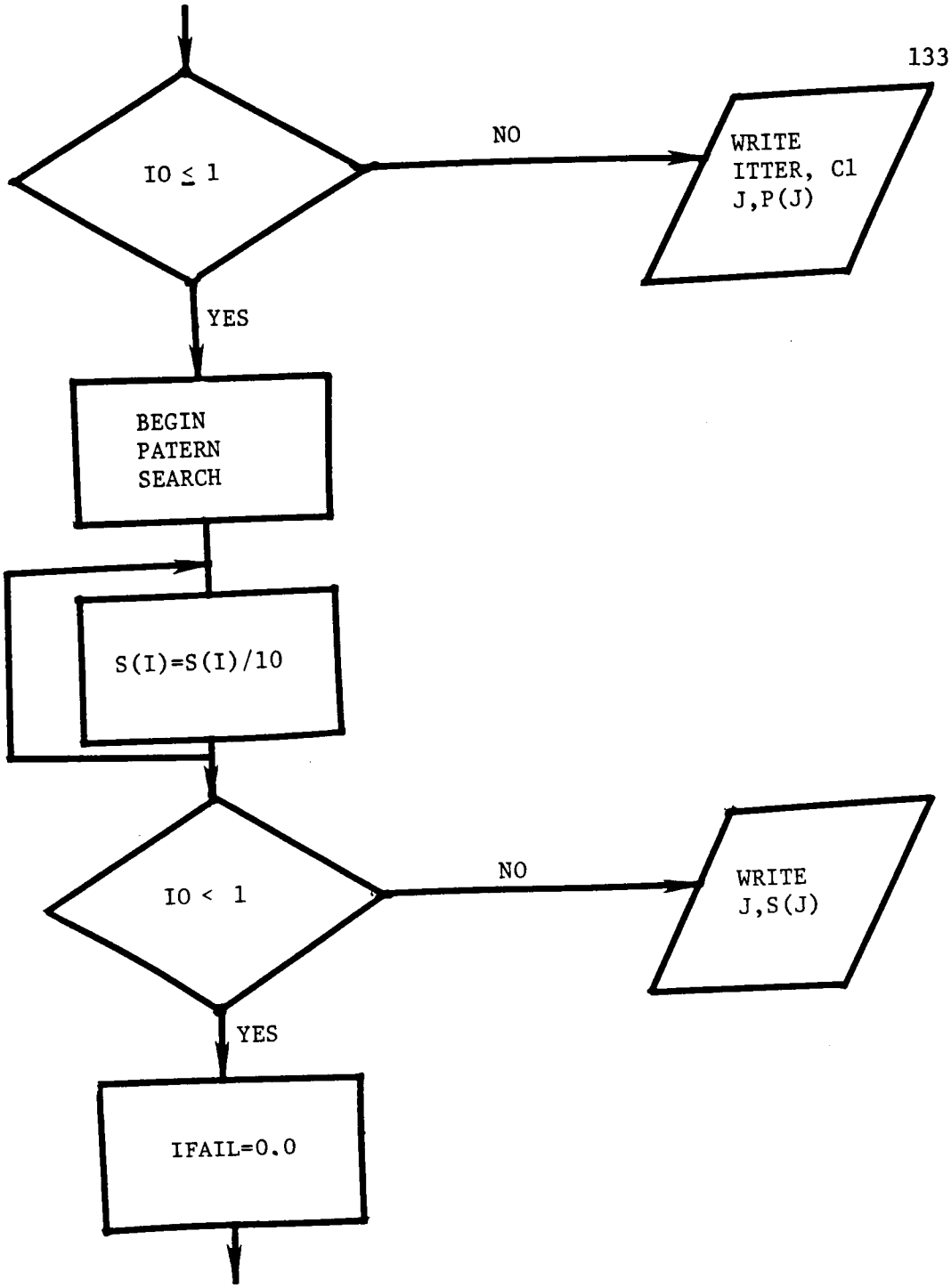
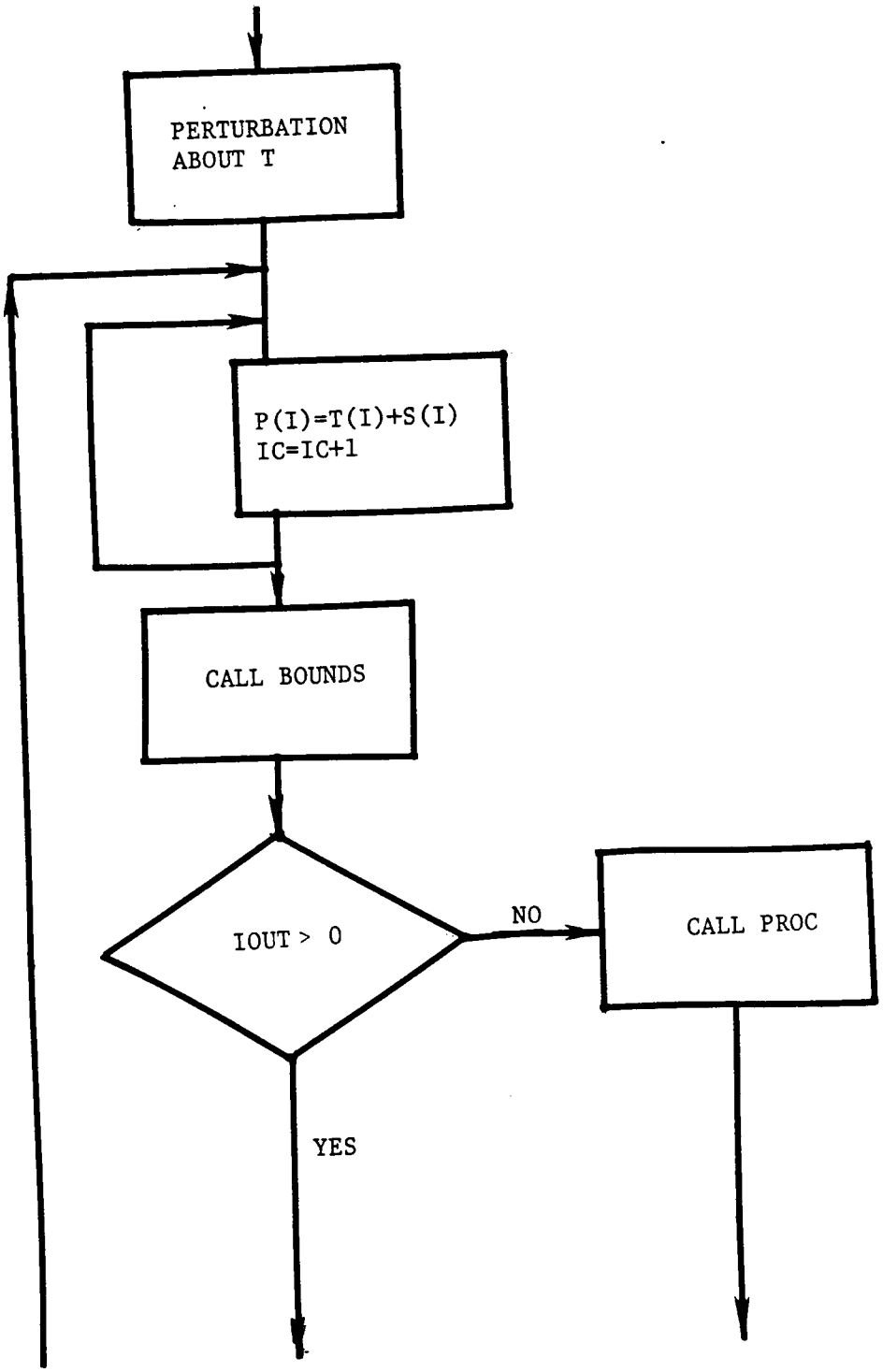
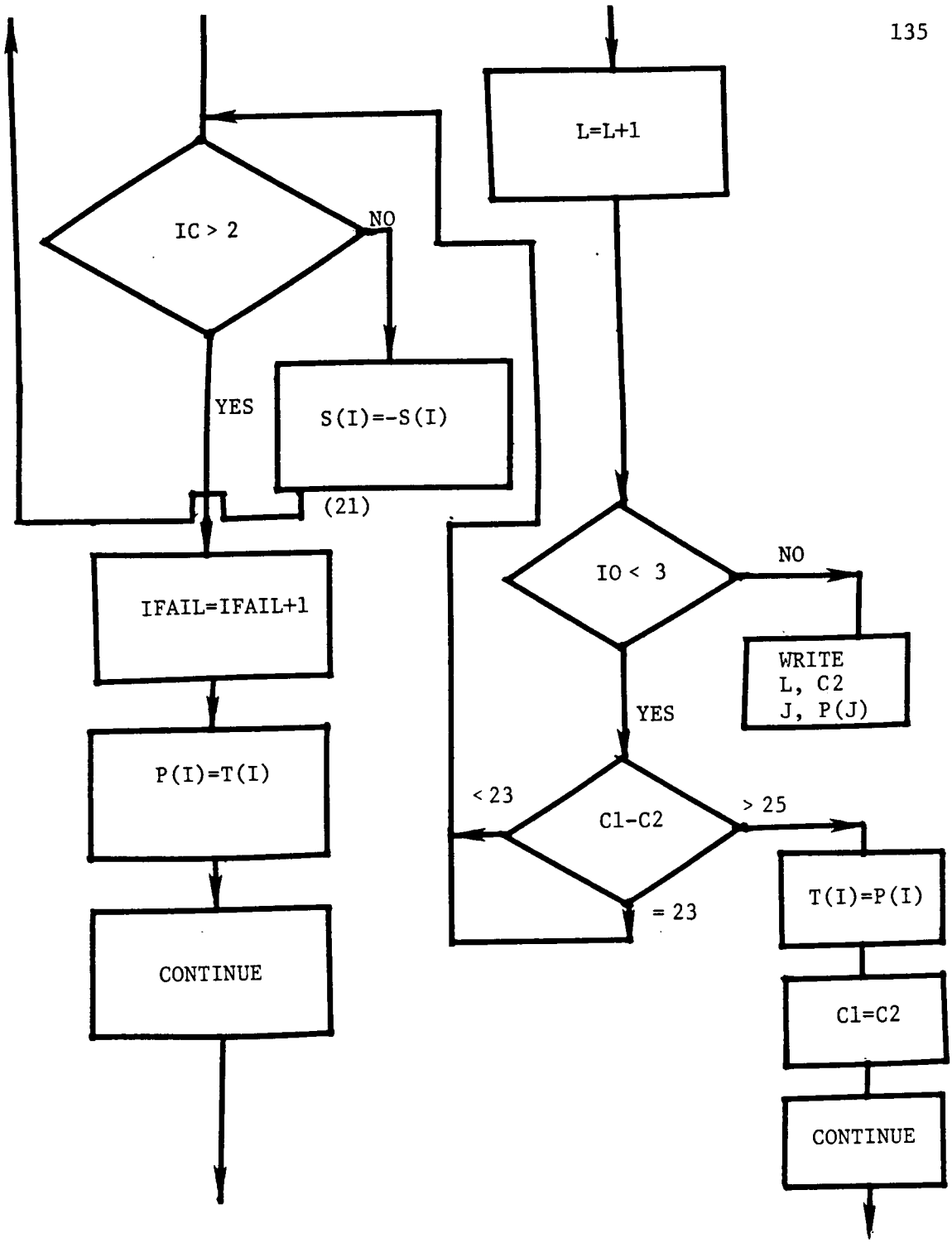
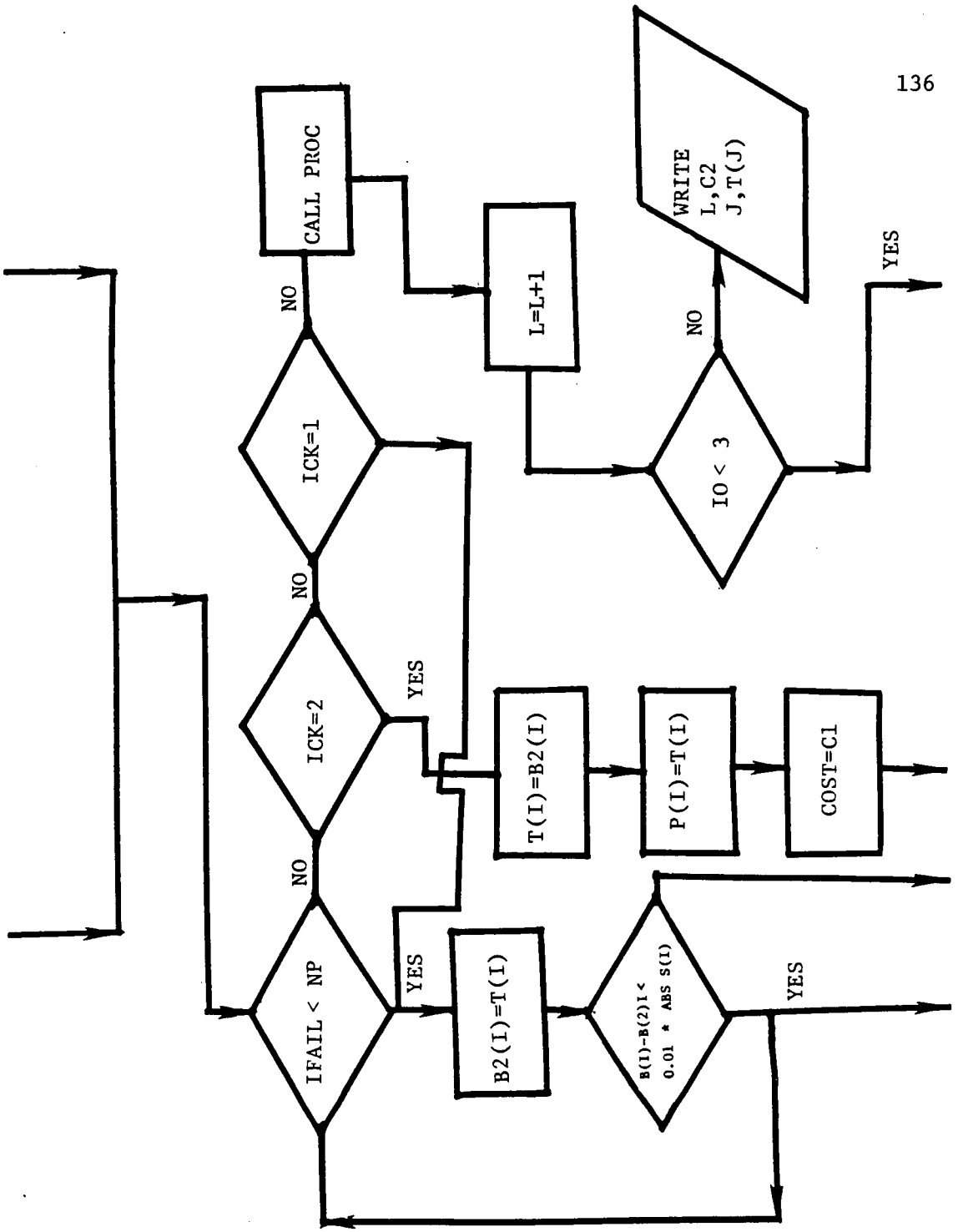


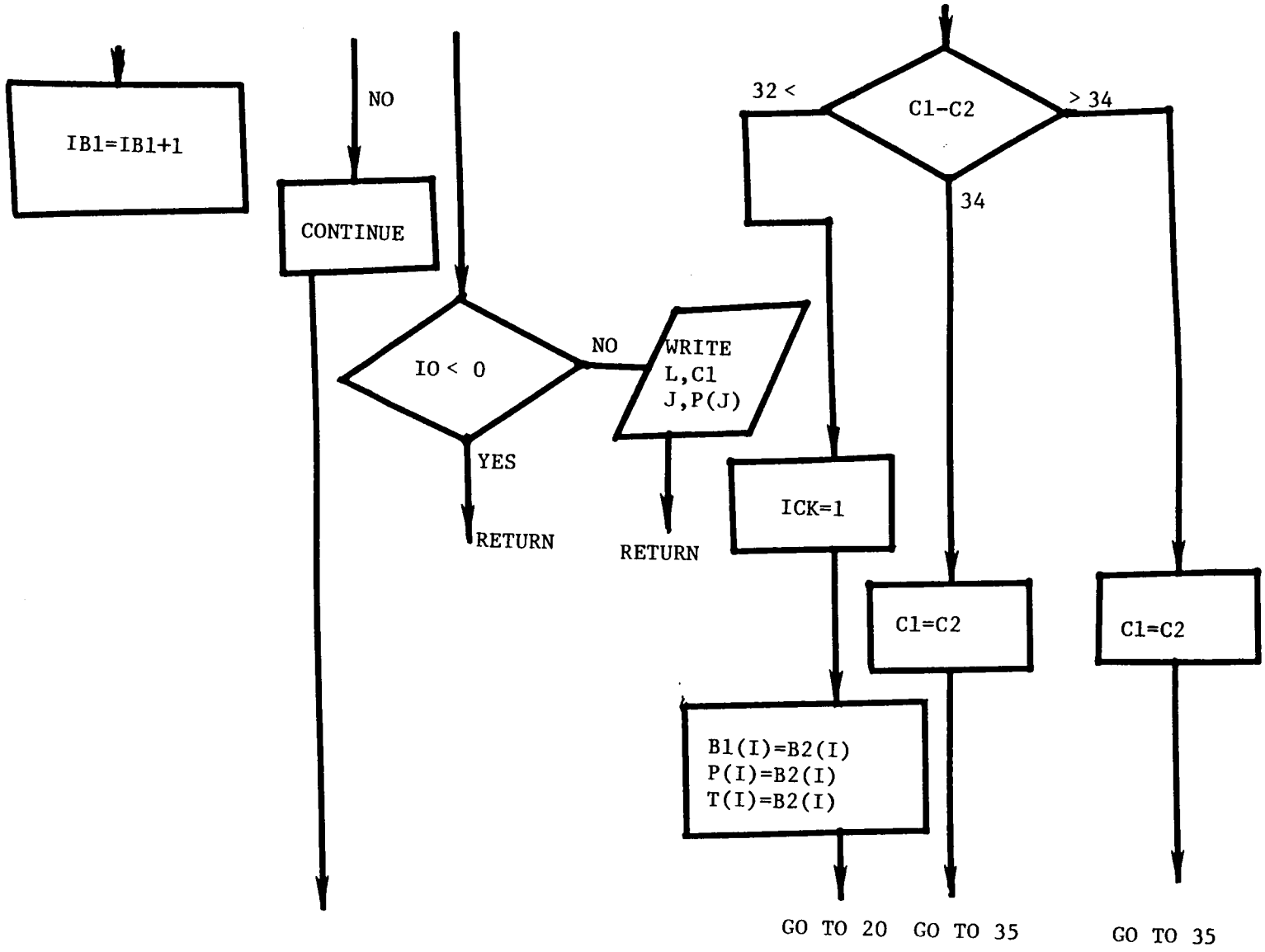
Figure B.3. A Flowchart for Subroutine PATERN.

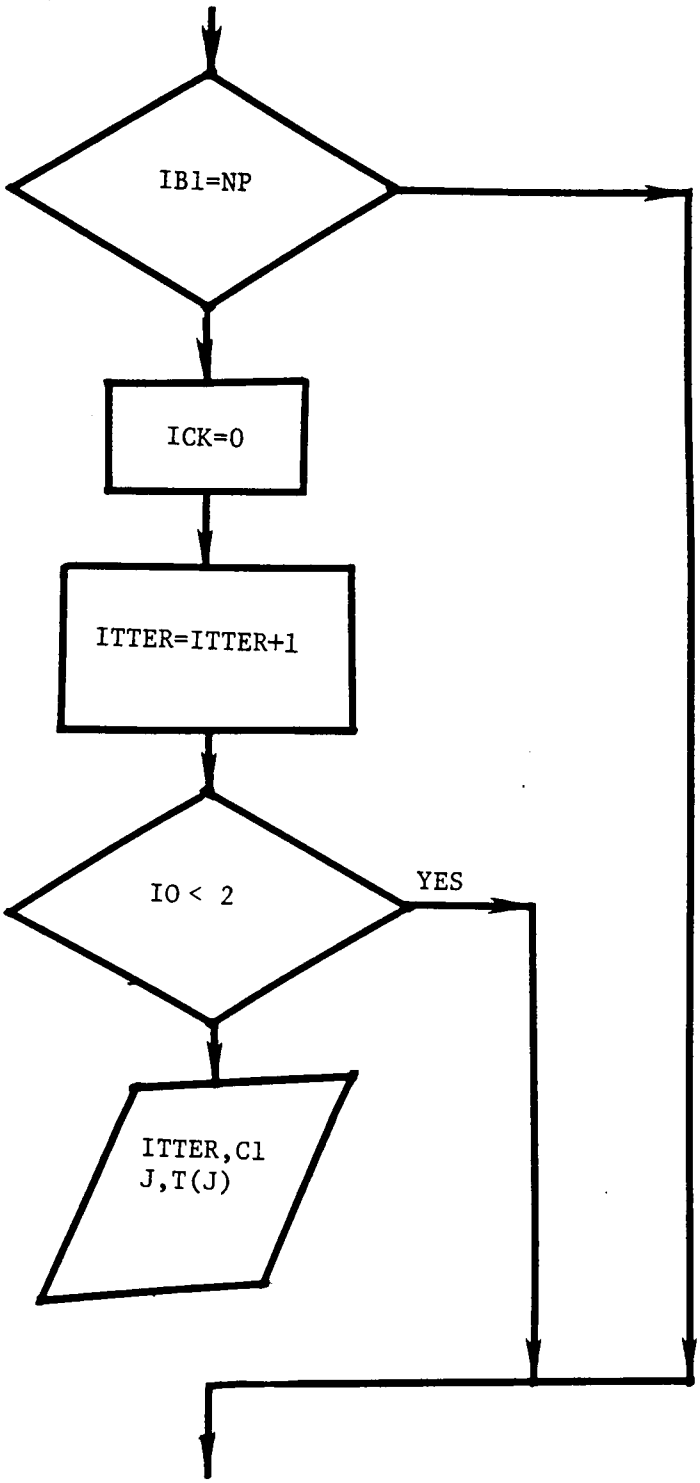

IO ≤ 1
NO
WRITE
ITTER, C1
J,P(J)
YES
BEGIN
PATERN
SEARCH
S(I)=S(I)/10
IO < 1
NO
WRITE
J,S(J)
YES
IFAIL=0.0

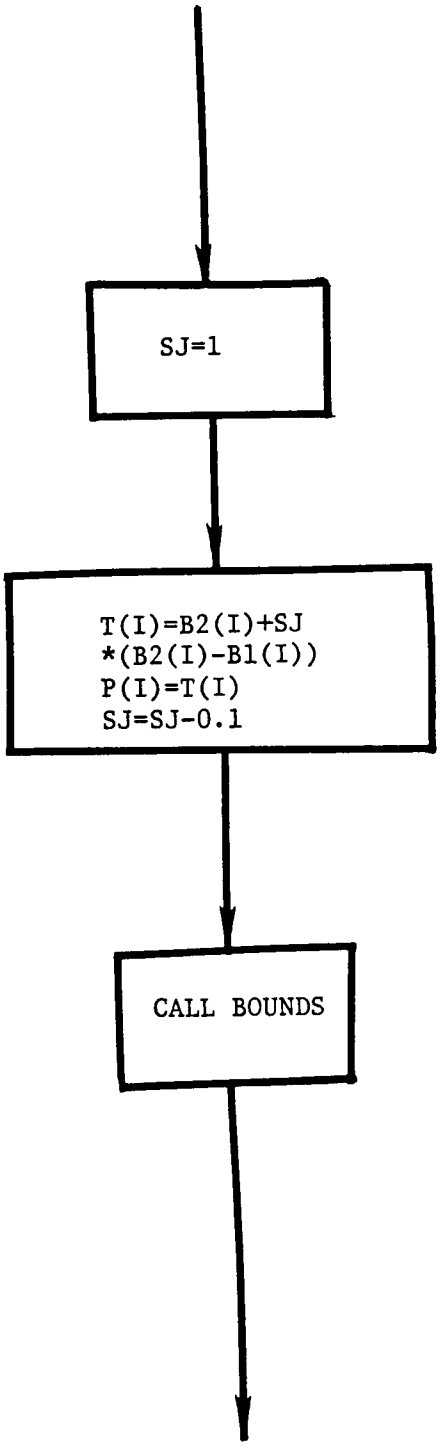

PERTURBATION
ABOUT T
P(I)=T(I)+S(I)
IC=IC+1
CALL BOUNDS
IOUT > 0
NO
CALL PROC
YES

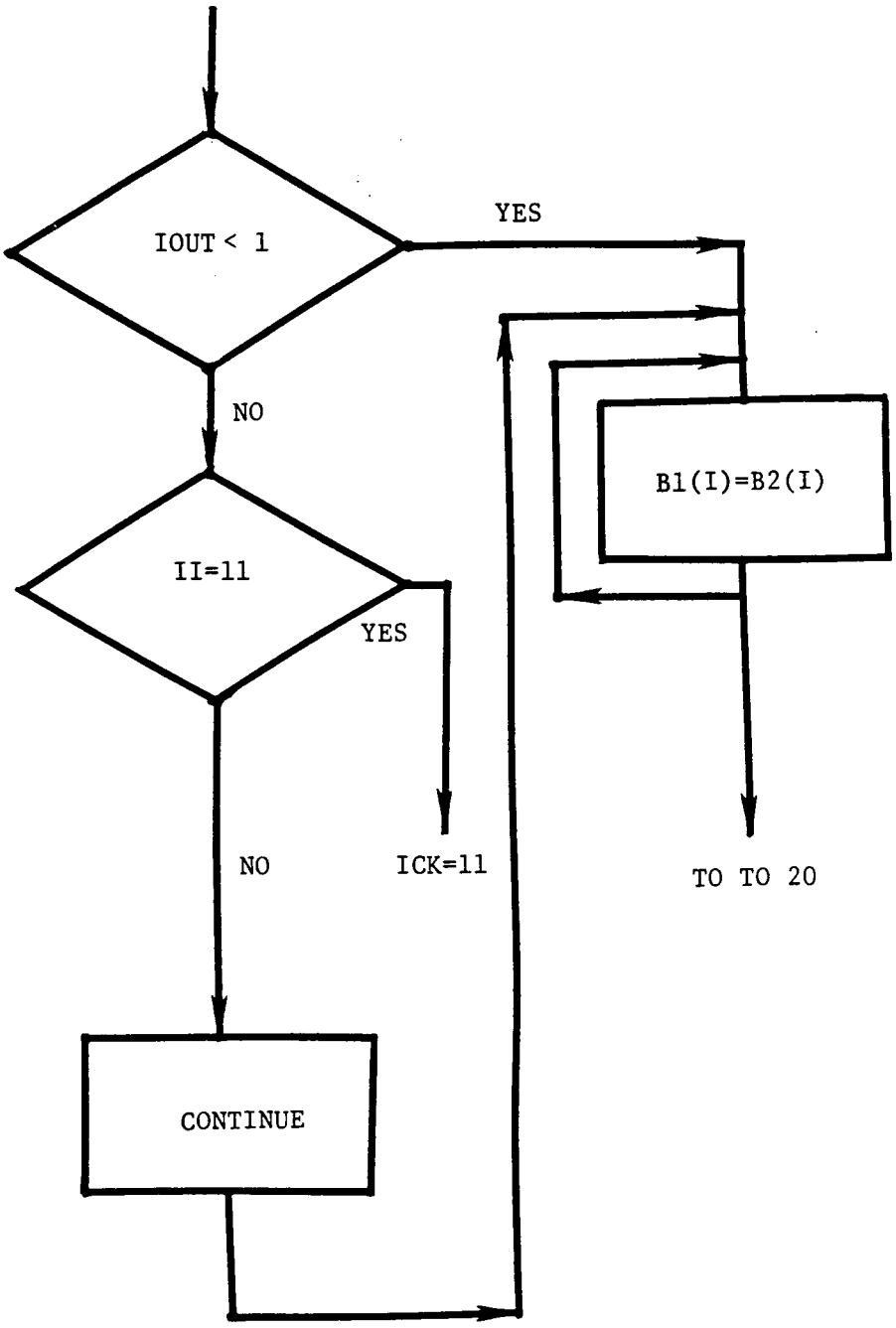

L=L+1
IC > 2
NO
YES
S(I)=-S(I)
(21)
IFAIL=IFAIL+1
P(I)=T(I)
CONTINUE
IO < 3
NO
WRITE
L, C2
J, P(J)
YES
< 23
C1-C2
> 25
= 23
T(I)=P(I)
C1=C2
CONTINUE


CALL PROC
NO
ICK=1
NO
ICK=2
NO
IFAIL < NP
YES
B2(I)=T(I)
B(1)-B(2)I <
0.01 * ABS S(I)
YES
YES
T(I)=B2(I)
P(I)=T(I)
COST=C1
L=L+1
NO
IO < 3
YES
WRITE
L,C2
J,T(J)


IB1=IB1+1
NO
CONTINUE
IO < 0
NO
WRITE
L,C1
J,P(J)
YES
RETURN
RETURN
32 <
C1-C2
> 34
34
ICK=1
C1=C2
C1=C2
B1(I)=B2(I)
P(I)=B2(I)
T(I)=B2(I)
GO TO 20
GO TO 35
GO TO 35


IB1=NP
ICK=0
ITTER=ITTER+1
IO < 2
YES
ITTER,C1
J,T(J)


SJ=1
T(I)=B2(I)+SJ
*(B2(I)-B1(I))
P(I)=T(I)
SJ=SJ-0.1
CALL BOUNDS


IOUT < 1
YES
NO
B1(I)=B2(I)
II=11
YES
NO
ICK=11
TO TO 20
CONTINUE

APPENDIX C

FOUR TRANSFER FUNCTION MODELS

Table C-1. Four Transfer Function Models

	Transfer Function Models
Weischedel and McAvoy	
Column A	$\begin{bmatrix} \frac{0.58e^{-s}}{(6.3s+1)(1.96s+1)} & \frac{-0.45e^{-s}}{(5.68s+1)(3.0s+1)} \\ \frac{0.35e^{-1.3s}}{(5.0s+1)(0.67s+1)} & \frac{-0.48e^{-s}}{(4.7s+1)(0.36s+1)} \end{bmatrix}$
Column B	$\begin{bmatrix} \frac{0.562e^{-s}}{(7.74s+1)(7.74s+1)} & \frac{-0.516e^{-1.5s}}{(7.1s+1)(7.1s+1)} \\ \frac{0.344e^{-1.5s}}{(15.8s+1)(0.5s+1)} & \frac{-0.393e^{-s}}{(13.8s+1)(0.4s+1)} \end{bmatrix}$
Column C	$\begin{bmatrix} \frac{0.471e^{-s}}{(30.7s+1)(30.7s+1)} & \frac{-0.495e^{-2s}}{(28.5s+1)(28.5s+1)} \\ \frac{0.749e^{-1.7s}}{(57s+1)(57s+1)} & \frac{-0.832e^{-s}}{(50.5s+1)(50.5s+1)} \end{bmatrix}$
Wood and Berry	$\begin{bmatrix} \frac{12.8e^{-s}}{1+16.7s} & \frac{-18.9e^{-3s}}{1+21.0s} \\ \frac{6.6e^{-7s}}{1+10.9s} & \frac{-19.4e^{-3s}}{1+14.4s} \end{bmatrix}$

APPENDIX D

PROOF THAT A SQUARE, ORTHOGONAL MATRIX HAS A CONDITION NUMBER OF ONE

APPENDIX D

PROOF THAT A SQUARE, ORTHOGONAL MATRIX HAS A CONDITION NUMBER OF ONE

Theorem 11: If a square matrix is scaled to be orthogonal, then it will have a condition number of one.

Proof: Let K = K

where is orthogonal

By singular value decomposition

$K = U \quad V^T$

$KV = U$

$U^TKV = U^{-1}KV$ since $U^T = U^{-1}$

$U^TU \quad VV^T = U^{-1}KV$ since $U^TU = I$, $VV^T = I$

$= \quad U^{-1}KV$

where U^{-1},K, and V are all

Since each matrix, U^{-1},K, and V are , then their product is orthogonal, is and diagonal.

By definition an orthogonal matrix has a condition number of one.

APPENDIX E

SCALING FACTORS FOR GEOMETRIC MEAN SCALED MATRICES

Table E.1. The Decomposition of Process Gain Matrices and Their Geometric Mean Scaled Matrices with Their Scaling Factors

Type	2 x 2 Matrix
NI	Gain
	$\begin{pmatrix} 1.000 & 0.000 \\ 0.000 & 1.000 \end{pmatrix}$
	U: $\begin{pmatrix} -1.000 & 0.000 \\ 0.000 & 1.000 \end{pmatrix}$ V: $\begin{pmatrix} -1.000 & 0.000 \\ 0.000 & 1.000 \end{pmatrix}$
	GM RC(1)
	$\begin{pmatrix} 1.000 & 0.000 \\ 0.000 & 1.000 \end{pmatrix}$
	U_{sc}: $\begin{pmatrix} -1.000 & 0.000 \\ 0.000 & 1.000 \end{pmatrix}$ V_{sc}: $\begin{pmatrix} -1.000 & 0.000 \\ 0.000 & 1.000 \end{pmatrix}$
	S_1^{-1}: $\begin{pmatrix} 1.000 & 0.000 \\ 0.000 & 1.000 \end{pmatrix}$ S_2^{-1}: $\begin{pmatrix} 1.000 & 0.000 \\ 0.000 & 1.000 \end{pmatrix}$
PI	Gain
	$\begin{pmatrix} 1.000 & 0.000 \\ 3.000 & 4.000 \end{pmatrix}$
	U: $\begin{pmatrix} -0.12218 & -0.99251 \\ -0.99251 & 0.12218 \end{pmatrix}$ V: $\begin{pmatrix} -0.61541 & -0.78821 \\ -0.78821 & 0.61541 \end{pmatrix}$
	GM RC(1)
	$\begin{pmatrix} 1.07457 & 0.00000 \\ 0.930605 & 1.00000 \end{pmatrix}$

Table E.1 (continued)

Type	2 x 2 Matrix
PI (continued)	U_{sc} $\begin{bmatrix} -0.57659 & -0.81704 \\ -0.81704 & 0.57659 \end{bmatrix}$ V_{sc} $\begin{bmatrix} -0.86040 & -0.50948 \\ -0.50948 & 0.86040 \end{bmatrix}$ S_1^{-1} $\begin{bmatrix} 1.00000 & 0.00000 \\ 0.00000 & 3.46410 \end{bmatrix}$ S_2^{-1} $\begin{bmatrix} 0.930605 & 0.00000 \\ 0.000000 & 1.15470 \end{bmatrix}$
FULL	Gain $\begin{bmatrix} 0.5800 & -0.4500 \\ 0.3500 & -0.4800 \end{bmatrix}$ U $\begin{bmatrix} -0.77996 & -0.62583 \\ -0.62583 & 0.77996 \end{bmatrix}$ V $\begin{bmatrix} -0.71774 & -0.69632 \\ 0.69632 & -0.71774 \end{bmatrix}$ GM RC(1) $\begin{bmatrix} 1.15305 & -0.867267 \\ 0.867267 & -1.15305 \end{bmatrix}$ U_{sc} $\begin{bmatrix} -0.70711 & -0.70711 \\ -0.70711 & 0.70711 \end{bmatrix}$ V_{sc} $\begin{bmatrix} -0.70711 & -0.70711 \\ 0.70711 & -0.70711 \end{bmatrix}$ S_1^{-1} $\begin{bmatrix} 0.510882 & 0.00000 \\ 0.000000 & 0.409878 \end{bmatrix}$ S_2^{-1} $\begin{bmatrix} 0.984602 & 0.00000 \\ 0.000000 & 1.01564 \end{bmatrix}$

Table E.2. The Decomposition of Process Gain Matrices and Their Geometric Mean Scaled Matrices with Their Scaling Factors (3 x 3 Matrix)

Type	3 x 3 Matrix

NI

Gain

$$\begin{bmatrix} 3.00000 & 0.00000 & 0.00000 \\ 0.00000 & 7.00000 & 0.00000 \\ 0.00000 & 0.00000 & 5.00000 \end{bmatrix}$$

U

$$\begin{bmatrix} 0.00000 & 0.00000 & -1.00000 \\ -1.00000 & 0.00000 & 0.00000 \\ 0.00000 & 1.00000 & 0.00000 \end{bmatrix}$$

V

$$\begin{bmatrix} 0.00000 & 0.00000 & -1.00000 \\ -1.00000 & 0.00000 & 0.00000 \\ 0.00000 & 1.00000 & 0.00000 \end{bmatrix}$$

GM RC(1)

$$\begin{bmatrix} 1.00000 & 0.00000 & 0.00000 \\ 0.00000 & 1.00000 & 0.00000 \\ 0.00000 & 0.00000 & 1.00000 \end{bmatrix}$$

U_{sc}

$$\begin{bmatrix} -1.00000 & 0.00000 & 0.00000 \\ 0.00000 & -1.00000 & 0.00000 \\ 0.00000 & 0.00000 & 1.00000 \end{bmatrix}$$

V_{sc}

$$\begin{bmatrix} -1.00000 & 0.00000 & 0.00000 \\ 0.00000 & -1.00000 & 0.00000 \\ 0.00000 & 0.00000 & 1.00000 \end{bmatrix}$$

S_1^{-1}

$$\begin{bmatrix} 3.00000 & 0.00000 & 0.00000 \\ 0.00000 & 7.00000 & 0.00000 \\ 0.00000 & 0.00000 & 5.00000 \end{bmatrix}$$

S_2^{-1}

$$\begin{bmatrix} 1.00000 & 0.00000 & 0.00000 \\ 0.00000 & 1.00000 & 0.00000 \\ 0.00000 & 0.00000 & 1.00000 \end{bmatrix}$$

PI

Gain

$$\begin{bmatrix} 0.00000 & 1.00000 & 1.00000 \\ 1.00000 & 0.00000 & 0.00000 \\ 0.00000 & 1.00000 & 0.00000 \end{bmatrix}$$

U

$$\begin{bmatrix} 0.85605 & 0.00000 & -0.52573 \\ 0.00000 & -1.00000 & 0.00000 \\ 0.52573 & 0.00000 & 0.85065 \end{bmatrix}$$

V

$$\begin{bmatrix} 0.00000 & -1.00000 & 0.00000 \\ 0.85065 & 0.00000 & 0.52573 \\ 0.52573 & 0.00000 & -0.85065 \end{bmatrix}$$

Table E.2 (continued)

Type	3 x 3 Matrix

PI (continued)

GM RC(1)

$$\begin{bmatrix} 0.00000 & 1.00000 & 1.00000 \\ 1.00000 & 0.00000 & 0.00000 \\ 0.00000 & 1.00000 & 0.00000 \end{bmatrix}$$

U_{sc}

$$\begin{bmatrix} 0.85065 & 0.00000 & -0.52573 \\ 0.00000 & -1.00000 & 0.00000 \\ 0.52573 & 0.00000 & 0.85065 \end{bmatrix}$$

V_{sc}

$$\begin{bmatrix} 0.00000 & -1.00000 & 0.00000 \\ 0.85065 & 0.00000 & 0.52573 \\ 0.52573 & 0.00000 & -0.85065 \end{bmatrix}$$

S_1^{-1}

$$\begin{bmatrix} 1.00000 & 0.00000 & 0.00000 \\ 0.00000 & 1.00000 & 0.00000 \\ 0.00000 & 0.00000 & 1.00000 \end{bmatrix}$$

S_2^{-1}

$$\begin{bmatrix} 1.00000 & 0.00000 & 0.00000 \\ 0.00000 & 1.00000 & 0.00000 \\ 0.00000 & 0.00000 & 1.00000 \end{bmatrix}$$

FULL

Gain

$$\begin{bmatrix} 0.34000 & 6.00000 & 98.0000 \\ 0.25000 & 5.40000 & 2.70000 \\ 0.12000 & 67.0000 & 34.0000 \end{bmatrix}$$

U

$$\begin{bmatrix} -0.86541 & -0.49487 & 0.07848 \\ -0.10788 & 0.03107 & -0.99368 \\ -0.48931 & 0.86841 & 0.08027 \end{bmatrix}$$

V

$$\begin{bmatrix} -0.32968 & -0.09574 & -0.93923 \\ -0.33459 & 0.94212 & 0.02141 \\ -0.88282 & -0.32131 & 0.34263 \end{bmatrix}$$

GM RC(1)

$$\begin{bmatrix} 1.09075 & 0.323486 & 3.09132 \\ 3.06869 & 1.11394 & 0.325872 \\ 0.326456 & 3.06320 & 0.909479 \end{bmatrix}$$

U_{sc}

$$\begin{bmatrix} -0.57850 & -0.77647 & 0.24989 \\ -0.59727 & 0.19460 & -0.77807 \\ -0.55552 & 0.59936 & 0.57634 \end{bmatrix}$$

V_{sc}

$$\begin{bmatrix} -0.59579 & -0.02128 & -0.80286 \\ -0.57528 & 0.70887 & 0.40812 \\ -0.56043 & -0.70502 & 0.43458 \end{bmatrix}$$

Table E.2 (continued)

Type	3 x 3 Matrix					
FULL (continued)	S_1^{-1}			S_2^{-1}		
	5.77235	0	0	.0499275	0	0
	0	2.83549	0	0	4.95590	0
	0	0	2.83549	0	0	14.2680

Table E.3. The Decomposition of Process Gain Matrices and Their Geometric Mean Scaled Matrices with Their Scaling Factors (4 x 4 Matrix)

Type: NI — 4 x 4 Matrix

Gain

$$\begin{bmatrix} 1.00000 & 0.00000 & 0.00000 & 0.00000 \\ 0.00000 & 1.00000 & 0.00000 & 0.00000 \\ 0.00000 & 0.00000 & 1.00000 & 0.00000 \\ 0.00000 & 0.00000 & 0.00000 & 1.00000 \end{bmatrix}$$

U

$$\begin{bmatrix} -1.00000 & 0.00000 & 0.00000 & 0.00000 \\ 0.00000 & -1.00000 & 0.00000 & 0.00000 \\ 0.00000 & 0.00000 & -1.00000 & 0.00000 \\ 0.00000 & 0.00000 & 0.00000 & 1.00000 \end{bmatrix}$$

V

$$\begin{bmatrix} -1.00000 & 0.00000 & 0.00000 & 0.00000 \\ 0.00000 & -1.00000 & 0.00000 & 0.00000 \\ 0.00000 & 0.00000 & -1.00000 & 0.00000 \\ 0.00000 & 0.00000 & 0.00000 & 1.00000 \end{bmatrix}$$

GM RC(1)

$$\begin{bmatrix} 1.00000 & 0.00000 & 0.00000 & 0.00000 \\ 0.00000 & 6.00000 & 0.00000 & 0.00000 \\ 0.00000 & 0.00000 & 5.00000 & 0.00000 \\ 0.00000 & 0.00000 & 0.00000 & 3.00000 \end{bmatrix}$$

U_{sc}

$$\begin{bmatrix} 0.00000 & 0.00000 & 0.00000 & -1.00000 \\ -1.00000 & 0.00000 & 0.00000 & 0.00000 \\ 0.00000 & -1.00000 & 0.00000 & 0.00000 \\ 0.00000 & 0.00000 & 1.00000 & 0.00000 \end{bmatrix}$$

V_{sc}

$$\begin{bmatrix} 0.00000 & 0.00000 & 0.00000 & -1.00000 \\ -1.00000 & 0.00000 & 0.00000 & 0.00000 \\ 0.00000 & -1.00000 & 0.00000 & 0.00000 \\ 0.00000 & 0.00000 & 1.00000 & 0.00000 \end{bmatrix}$$

Table E.3 (continued)

Type	4 x 4 Matrix

NI (cont'd)

S_1^{-1}

$$\begin{bmatrix} 6.00000 & 0.00000 & 0.00000 & 0.00000 \\ 0.00000 & 5.00000 & 0.00000 & 0.00000 \\ 0.00000 & 0.00000 & 3.00000 & 0.00000 \\ 0.00000 & 0.00000 & 0.00000 & 1.00000 \end{bmatrix}$$

S_2^{-1}

$$\begin{bmatrix} 1.00000 & 0.00000 & 0.00000 & 0.00000 \\ 0.00000 & 1.00000 & 0.00000 & 0.00000 \\ 0.00000 & 0.00000 & 1.00000 & 0.00000 \\ 0.00000 & 0.00000 & 0.00000 & 1.00000 \end{bmatrix}$$

PI

Gain

$$\begin{bmatrix} 1.00000 & 0.00000 & 0.00000 & 0.00000 \\ 0.00000 & 6.00000 & 7.00000 & 0.00000 \\ 0.00000 & 0.00000 & 5.00000 & 2.00000 \\ 0.00000 & 0.00000 & 0.00000 & 3.00000 \end{bmatrix}$$

U

$$\begin{bmatrix} 0.00000 & 0.00000 & 0.00000 & -1.00000 \\ -0.90021 & -0.36394 & -0.23909 & 0.00000 \\ -0.43454 & 0.71543 & 0.54711 & 0.00000 \\ -0.02807 & 0.59641 & -0.80219 & 0.00000 \end{bmatrix}$$

V

$$\begin{bmatrix} 0.00000 & 0.00000 & 0.00000 & -1.00000 \\ -0.53508 & -0.54257 & -0.64753 & 0.00000 \\ -0.83950 & 0.25582 & 0.47936 & 0.00000 \\ -0.09444 & 0.80011 & -0.59238 & 0.00000 \end{bmatrix}$$

GM RC(1)

$$\begin{bmatrix} 1.00000 & 0.00000 & 0.00000 & 0.00000 \\ 0.00000 & 1.00000 & 0.82657 & 0.00000 \\ 0.00000 & 0.00000 & 0.120990 & 0.795271 \\ 0.00000 & 0.00000 & 0.00000 & 0.125743 \end{bmatrix}$$

Table E.3 (continued)

Type	4 x 4 Matrix

GM RC(1)

$$\begin{pmatrix} 0.595861 & 1.23660 & 1.67824 & 0.812538 \\ 0.812538 & 1.67824 & 1.23660 & 0.595861 \\ -1.67824 & 0.595861 & 1.24138 & 0.636108 \\ 0.636108 & 1.24138 & 0.595861 & -1.67824 \end{pmatrix}$$

U_{sc}

$$\begin{pmatrix} -0.63484 & -0.09766 & -0.31140 & 0.70033 \\ -0.63484 & 0.09766 & -0.31140 & -0\ 70033 \\ -0.31140 & -0.70033 & 0.63484 & -0.09766 \\ -0.31140 & 0.70033 & 0.63484 & 0.09766 \end{pmatrix}$$

V_{sc}

$$\begin{pmatrix} -0.16184 & 0.67699 & -0.68834 & 0.20417 \\ -0.68834 & 0.20417 & 0.16184 & -0.67699 \\ -0.68834 & -0.20417 & 0.16184 & 0.67699 \\ -0.16184 & -0.67699 & -0.68834 & -0.20417 \end{pmatrix}$$

APPENDIX F

SCALING FACTORS FOR NLST

SCALED MATRICES

Table F.1. Minimum Condition Numbers and Scaling Factors for Selected NLST Scaled Matrices

Gain	Type	K_{min}	k_1	k_2	k_3	k_4	k_5	k_6
				2 x 2 Matrices				
$\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}$	NI	1.000000		1.000000		1.000000		
		SCAL41	k_4 = 0.1 to 2, STEP = 0.01, NPASS = 3, IO = 1, NP = 2					
$\begin{pmatrix}1 & 0\\ 3 & 4\end{pmatrix}$	PI	1.548701	0.1330999	1.630029				
		SCAL21	k_4 = 0.1 to 2, STEP = 0.001, NPASS = 3, IO = 1, NP = 2					
$\begin{pmatrix}0.58 & -0.45\\ 0.35 & -0.48\end{pmatrix}$	F	7.069464	0.8023002	1.031700				
		SCAL41	k_1 = 0.1 to 2, STEP = 0.01, NPASS = 3, IO = 1, NP = 2					
				3 x 3 Matrices				
$\begin{pmatrix}3 & 0 & 0\\ 0 & 7 & 0\\ 0 & 0 & 5\end{pmatrix}$	NI	1.000005	1.662463	0.8448415	1.002530	0.8454717		
		SCAL33	k_1 = 0.1 to 2, STEP = 0.001, NPASS = 3, IO = 1, NP = 4					
$\begin{pmatrix}0 & 1 & 1\\ 1 & 0 & 0\\ 0 & 1 & 0\end{pmatrix}$	PI	1.71219	0.5631000	0.7710029	0.8850015	0.535006		
		SCAL33	k_1 = 0.1 to 2, STEP = 0.001, NPASS = 3, IO = 1, NP = 4					
$\begin{pmatrix}1 & 1 & -0.1\\ 0.1 & 2 & -1\\ -2 & -3 & 1\end{pmatrix}$	F	24.40575	2.486917	1.151420	0.4418792	0.4064716		
		SCAL33	k_1 = 0.1 to 2, STEP = 0.001, NPASS = 3, IO = 1, NP = 4					

Table F.1. (continued)

Gain	Type	k_{min}	k_1	k_2	k_3	k_4	k_5	k_6
				4 4 4 Matrices				
$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 6 & 0 & 0 \\ 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}$	NI	1.005478	1.529548	1.963662	0.7156268	0.63938232	0.7691277	0.7782292
		SCAL4	k_1 = 0.1 to 2, STEP = 0.001, NPASS = 3, IO = 1, NP = 6					
$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 6 & 7 & 0 \\ 0 & 0 & 5 & 2 \\ 0 & 0 & 2 & 3 \end{pmatrix}$	PI	3.278774	3.707491	1.252012	0.1394854	3.034115	0.7549554	0.8122481
		SCAL4	k_1 = 0.1 to 2, STEP = 0.001, NPASS = 3, IO = 1, NP = 6					
$\begin{pmatrix} 3.3 & 5.6 & 7.6 & 4.5 \\ 4.5 & 7.6 & 5.6 & 3.3 \\ -12.4 & 3.6 & 7.5 & 4.7 \\ 4.7 & 7.5 & 3.6 & -12.4 \end{pmatrix}$	F	9.293558	2.270036	2.270134	1.026994	1.336014	1.005996	1.337013
			k_1 = 0.1 to 2, STEP = 0.01, NPASS = 3, IO = 1, NP = 6					

VITA

Cathy W. Dyer was born in Nashville, Tennessee, and attended public school in the same city. By August 1976, she graduated with a B.A. in Psychology from George Peabody College for Teachers. After several jobs that were low paying and uninteresting, she changed career fields and graduated from Nashville State Technical Institute in 1979. For approximately two years, she worked as a laboratory technician at Oak Ridge National Laboratory. She graduated with a B. S. in Chemical Engineering from The University of Tennessee, Knoxville in March 1984, and a M.S. in June 1987.