



AgEcon SEARCH
RESEARCH IN AGRICULTURAL & APPLIED ECONOMICS

The World's Largest Open Access Agricultural & Applied Economics Digital Library

This document is discoverable and free to researchers across the globe due to the work of AgEcon Search.

Help ensure our sustainability.

Give to AgEcon Search

AgEcon Search
<http://ageconsearch.umn.edu>
aesearch@umn.edu

*Papers downloaded from **AgEcon Search** may be used for non-commercial purposes and personal study only. No other use, including posting to another Internet site, is permitted without permission from the copyright owner (not AgEcon Search), or as allowed under the provisions of Fair Use, U.S. Copyright Act, Title 17 U.S.C.*

Econ.

Erasmus University Rotterdam

ECONOMETRIC INSTITUTE

Netherlands school of economics

Report 7606/E

SOLUTION OF ECONOMETRIC EQUATION SYSTEMS
BY MEANS OF A MODIFIED GAUSS-SEIDEL PROCEDURE

by A. Kunstman and T. Kloek

GIANNINI FOUNDATION OF
AGRICULTURAL ECONOMICS
LIBRARY
WITHDRAWN

JUL 19 1976

March 1976

SOLUTION OF ECONOMETRIC EQUATION SYSTEMS
BY MEANS OF A MODIFIED GAUSS-SEIDEL PROCEDURE

by A. Kunstman and T. Kloek

Contents

	Page
Abstract	1
1. Introduction and summary	1
2. Derivation of the method	4
3. Application to linear systems	10
4. Application to nonlinear systems	17
References	23

Abstract

The Gauss-Seidel method for solving equation systems is modified by using certain weighted averages of current Gauss-Seidel values and corresponding values from the preceding iteration step. This modified Gauss-Seidel method is shown to converge for all linear systems. It is relatively efficient for systems with many zeros in the coefficient matrix. It has also been successfully applied to some nonlinear econometric equation systems.

1. INTRODUCTION AND SUMMARY

The starting point of our approach is the well-known Gauss-Seidel (hereafter G.S.) method for solving linear equation systems.¹ This method can be described as follows. Consider a system of n linear equations in n unknowns, which can be written as

$$(1.1) \quad Ax = b$$

¹ For a treatment of iterative methods for the solution of systems of equations an extensive literature is available. See for example Faddeev and Faddeeva (1963), Hildebrand (1974) or Varga (1962). For an application of this method to the solution of econometric systems see Fromm and Klein (1969), Klein and Evans (1969) or Ball *et al.* (1975).

where A is a given $n \times n$ matrix of full rank, b a given n -vector and x an n -vector of unknowns. Without loss of generality² we can assume that $a_{ii} = 1$ ($i = 1, \dots, n$). The matrix A of (1.1) is split up as follows.

$$(1.2) \quad A = L - U$$

where

$$(1.3) \quad L = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ a_{21} & 1 & 0 & \dots & 0 \\ a_{31} & a_{32} & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & a_{n3} & \dots & 1 \end{bmatrix}$$

Then (1.1) can be rewritten as

$$(1.4) \quad Lx - Ux = b$$

G.S. consists of a sequence of computed values $x^{(1)}, \dots, x^{(T)}$, depending on an initial guess $x^{(0)}$, where $x^{(t)}$ is obtained by

$$(1.5) \quad x^{(t)} = L^{-1}b + L^{-1}Ux^{(t-1)}$$

The sequence $x^{(1)}, \dots, x^{(t)}, \dots$ converges to a finite limit if

$$(1.6) \quad (L^{-1}U)^t \rightarrow 0 \quad \text{as } t \rightarrow \infty$$

as can easily be verified. For an arbitrary system this convergence is

²The normalized system (1.1) can be obtained from a general linear system $A^*x = b^*$ by means of

$$D^{-1}A^*x = D^{-1}b^*$$

where D is a diagonal matrix with typical element a_{ii}^* , provided that $a_{ii}^* \neq 0$ for all i . This condition can always be met by an appropriate ordering of the equations of the system. Normalization implies that for every variable x_i we have an explicit expression in terms of the other variables ($x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n$).

not guaranteed automatically, because convergence depends on the magnitude of the eigenvalues of $L^{-1}U$.³ Reordering of the equations of the original system and a new normalization in the way of (1.1) can lead to a convergent process, but there are no general rules with respect to performing such a reordering.

For that reason it has been our goal to develop a modification of G.S., which enables the application of the method to every linear system, even in cases where the original G.S. procedure leads to a divergent process. An additional advantage is that in the case of a convergent G.S. process, the speed of convergence can be improved upon. As far as non-linear systems of equations are concerned, in a number of cases G.S. has been proven a useful tool as well. For those cases our modified procedure (henceforth indicated as M.G.S.) can be applied, as well, with the same useful property that convergence is guaranteed.

In principle, M.G.S. can be applied to every normalized system. For a number of these systems, however, the method is inefficient, compared with, for example, straightforward inversion. This is especially the case for large systems $Ax = b$ with a full matrix A .

A large class of econometric systems, however, possesses properties that make them very well suited for the application of M.G.S. These properties are:

- (1) Normalization can be performed, even in the case of non-linearities;⁴
- (2) Even in large systems the endogenous variables are explained by only a few other endogenous variables (in terms of the linear system (1.1): the matrix A contains a great number of zero off-diagonal elements);
- (3) Many or all equations of the system are linear. In such systems the relative efficiency of the method is much greater (as will be shown by one of our experiments).

The order of discussion is as follows. In section 2 the mathematical formulation is presented. Section 3 contains the results of the application

³ Economists can easily demonstrate a divergent G.S. procedure by drawing a cobweb diagram. For the conditions for convergence see the references mentioned in footnote (1). Special conditions on the matrix A ensue that G.S. always converges to the solution. This is, for example, the case if A is symmetric or Hermitian and positive-definite. See Faddeev and Faddeeva (1963).

⁴ For econometric systems with nonlinearities the use of G.S. is widespread. See, for example, Fromm and Taubman (1967), Fromm and Klein (1969), and Ball *et al.* (1975).

of the method to a number of linear models and a comparison is made with straightforward inversion. Finally, in section 4, the method is used to solve two systems of non-linear equations.

The results, which are qualitatively the same for the linear and the non-linear examples, can be summarized as follows. First, when G.S. was applied to a specific order of solution, this led to a divergent process. Without any difficulty M.G.S. could be applied, leading to a convergent process. Second, in an alternative order of solution, G.S. led to a convergent process. However, the speed of convergence could be considerably improved by application of M.G.S.

2. DERIVATION OF THE METHOD

Consider the following system of two linear equations in two unknowns

$$(2.1) \quad \begin{aligned} x_1 + a_{12}x_2 &= b_1 \\ a_{21}x_1 + x_2 &= b_2 \end{aligned}$$

It is supposed that this system has a unique solution, leading to the necessary and sufficient condition $a_{12}a_{21} \neq 1$. Writing system (2.1) in the way of (1.4) we obtain

$$(2.2) \quad \begin{bmatrix} 1 & 0 \\ a_{21} & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} - \begin{bmatrix} 0 & -a_{12} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

Application of G.S. (eq. (1.5)) leads to

$$(2.3) \quad \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ a_{21} & 1 \end{bmatrix}^{-1} \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ a_{21} & 1 \end{bmatrix}^{-1} \begin{bmatrix} 0 & -a_{12} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1(t-1) \\ x_2(t-1) \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 \\ -a_{21} & 1 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} + \begin{bmatrix} 0 & -a_{12} \\ 0 & a_{12}a_{21} \end{bmatrix} \begin{bmatrix} x_1(t-1) \\ x_2(t-1) \end{bmatrix}$$

So

$$(2.4) \quad x_{1g}(t) = b_1 - a_{12}x_2(t-1)$$

$$(2.5) \quad x_{2g}^{(t)} = b_2 - a_{21}b_1 + a_{12}a_{21}x_2^{(t-1)}$$

where the subscript g indicates the computation by means of a G.S. step. A sufficient condition for convergence of the sequence generated by (2.5) to a finite limit is

$$|a_{12}a_{21}| < 1$$

However, this condition is not necessarily satisfied. In that case G.S. leads to a divergent process.⁵

Now we consider a modified method. Given some value $x_2^{(t-1)}$ (2.5) is replaced by

$$(2.6) \quad \begin{aligned} x_2^{(t)} &= h_2 x_{2g}^{(t)} + (1 - h_2) x_2^{(t-1)} \\ &= \{1 - h_2(1 - a_{12}a_{21})\} x_2^{(t-1)} + h_2(b_2 - a_{21}b_1) \end{aligned}$$

that is, by a weighted sum of x_{2g}^t and $x_2^{(t-1)}$. A sufficient condition for the convergence of the sequence generated by (2.6) to some finite limit $x_2^{(\infty)}$ is

$$(2.7) \quad |1 - h_2(1 - a_{12}a_{21})| < 1$$

It will be clear from (2.7) that not only the speed of convergence depends on the value of h_2 , but also that convergence always can be obtained by an appropriate choice of h_2 .⁶ Now the speed of convergence is as high as

⁵ It can easily be shown that a convergent process can be reached by computing x_1 by means of equation 2, that is by another ordering and normalization of (2.1). In that case we have

$$\begin{aligned} (1/a_{12})x_1 + x_2 &= b_1/a_{12} \\ x_1 + (1/a_{21})x_2 &= b_2/a_{21} \end{aligned}$$

A sufficient condition for convergence is $|(1/a_{12})(1/a_{21})| < 1$; this condition is automatically satisfied if $|a_{12}a_{21}| > 1$.

⁶ Fromm and Klein (1969) indicate the possibility of choosing an h such that the speed of convergence is increased. Klein and Evans (1969) mention the fact that an appropriate choice of the weighting factor h_2 in a system of two equations in two unknowns "is capable of converting a divergent in a convergent path".

possible if h_2 is chosen such that

$$1 - h_2(1 - a_{12}a_{21}) = 0$$

or

$$(2.8) \quad h_2 = \frac{1}{1 - a_{12}a_{21}}$$

leading to the solution of x_2 in the first step of the iterative procedure.⁷

The computation of h_2 by (2.8) is not a good starting point, since its generalization for a system of n equations will turn out to be computationally inefficient. There is, however, another way to compute h_2 , that has a useful generalization for systems of n equations. Starting with an initial value $x_2^{(0)}$ we obtain from (2.5)

$$(2.9) \quad x_{2g}^{(1)} = b_2 - a_{21}b_1 + a_{12}a_{21}x_2^{(0)}$$

and

$$(2.10) \quad x_{2g}^{(2)} = b_2 - a_{21}b_1 + a_{12}a_{21}x_{2g}^{(1)}$$

Subtracting (2.9) from (2.10) leads to

$$(2.11) \quad x_{2g}^{(2)} - x_{2g}^{(1)} = a_{12}a_{21}(x_{2g}^{(1)} - x_2^{(0)})$$

or

$$(2.12) \quad a_{12}a_{21} = (x_{2g}^{(2)} - x_{2g}^{(1)}) / (x_{2g}^{(1)} - x_2^{(0)})$$

Instead of the coefficients a_{12} and a_{21} we thus can use the first differences $(x_{2g}^{(2)} - x_{2g}^{(1)})$ and $(x_{2g}^{(1)} - x_2^{(0)})$ to compute the weighting factor h_2 . This result, rather trivial for a system of two equations, will turn out to be applicable for a system of n equations, as well.

So for a system of two linear equations in two unknowns it has been proved that M.G.S. always in a few steps leads to the solution.⁸ Starting with some value $x_2^{(0)}$, G.S. is used to compute $x_{1g}^{(1)}$, $x_{2g}^{(1)}$, $x_{1g}^{(2)}$, and $x_{2g}^{(2)}$,

⁷ If $h_2 = 1/(1 - a_{12}a_{21})$, formula (2.6) leads to $x_2^{(t)} = (b_2 - a_{21}b_1)/(1 - a_{12}a_{21})$, which is nothing but the solution obtained by straightforward inversion.

⁸ This is even the case in the situation where $a_{12}a_{21} = -1$ leading to a G.S.-process that neither converges, nor diverges.

by means of formulas (2.4) and (2.5), respectively. Then h_2 is computed by means of (2.8), using relation (2.12). Finally, (2.6) is used to compute x_2 , whereas from (2.4) we obtain $x_1 = x_{1g}^{(3)}$.

We next generalize this procedure for the n -dimensional system (1.1). This system is partitioned as follows

$$(2.13) \quad \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

where A , x , and b are conformably partitioned, such that A_{11} is of order $(i-1) \times (i-1)$, A_{22} of order 1×1 and A_{33} of order $(n-i) \times (n-i)$. Now a procedure for the solution of the subvectors x_1 and x_2 (where x_2 consists of only one element) is defined in three stages.

First, let the subvector x_3 be given and let $x_2 = x_2^{(t-1)}$. Then, from (2.13) we take the subsystem

$$(2.14) \quad A_{11}x_1 + A_{12}x_2 + A_{13}x_3 = b_1$$

leading to the solution

$$(2.15) \quad x_1^{(t)} = A_{11}^{-1}[b_1 - A_{12}x_2^{(t-1)} - A_{13}x_3] = b_1^* - A_{11}^{-1}A_{12}x_2^{(t-1)}$$

where $b_1^* = A_{11}^{-1}[b_1 - A_{13}x_3]$. It will be clear that the existence of A_{11}^{-1} has to be assumed, which is slightly more restrictive⁹ than the original existence assumption for A^{-1} .

Second, from (2.13) we take the subsystem

$$(2.16) \quad A_{21}x_1 + A_{22}x_2 + A_{23}x_3 = b_2$$

Let x_3 be given as before. Then $x_2^{(t)}$ is computed from (2.16) as follows

$$(2.17) \quad x_{2g}^{(t)} = b_2 - A_{21}x_1^{(t)} - A_{23}x_3 = b_2^* - A_{21}x_1^{(t)}$$

where $b_2^* = b_2 - A_{23}x_3$, and where $x_1^{(t)}$ comes from (2.15). It has to be

⁹ To our knowledge, econometric equation systems (normalized, with $a_{ii} = 1$ for all i) satisfy this condition as a rule.

realized that A_{22} is a diagonal element of A , equal to 1.

Third, $x_2^{(t)}$ is computed as a weighted sum; compare (2.6). We replace (2.17) by

$$(2.18) \quad \begin{aligned} x_2^{(t)} &= h_i x_{2g}^{(t)} + (1 - h_i) x_2^{(t-1)} \\ &= h_i \{b_2^* - A_{21} x_1^{(t)}\} + (1 - h_i) x_2^{(t-1)} \end{aligned}$$

After substitution of (2.15) we obtain

$$(2.19) \quad \begin{aligned} x_2^{(t)} &= h_i b_2^* - h_i A_{21} (b_1^* - A_{11}^{-1} A_{12} x_2^{(t-1)}) + (1 - h_i) x_2^{(t-1)} \\ &= h_i (b_2^* - A_{21} b_1^*) + (1 - h_i + h_i A_{21} A_{11}^{-1} A_{12}) x_2^{(t-1)} \end{aligned}$$

Then the computation of $x_1^{(t)}$ and $x_2^{(t)}$ as an iterative step can be summarized as

$$(2.20) \quad \begin{bmatrix} x_1^{(t)} \\ x_2^{(t)} \end{bmatrix} = \begin{bmatrix} b_1^* \\ h_i (b_2^* - A_{21} b_1^*) \end{bmatrix} + \begin{bmatrix} 0 & -A_{11}^{-1} A_{12} \\ 0 & 1 - h_i + h_i A_{21} A_{11}^{-1} A_{12} \end{bmatrix} \begin{bmatrix} x_1^{(t-1)} \\ x_2^{(t-1)} \end{bmatrix}$$

as follows from (2.15) and (2.19). The sequence generated by (2.20) converges to a finite limit if

$$(2.21) \quad \begin{bmatrix} 0 & -A_{11}^{-1} A_{12} \\ 0 & 1 - h_i + h_i A_{21} A_{11}^{-1} A_{12} \end{bmatrix}^n \rightarrow 0 \text{ as } n \rightarrow \infty$$

which results in the sufficient condition

$$(2.22) \quad |1 - h_i + h_i A_{21} A_{11}^{-1} A_{12}| < 1$$

It will be clear that always a value for h_i can be found such that (2.22) is satisfied. As can be seen from (2.22) the speed of convergence is as high as possible for a value of h_i such that

$$(2.23) \quad 1 - h_i + h_i A_{21} A_{11}^{-1} A_{12} = 0$$

or

$$(2.24) \quad h_i = \frac{1}{1 - A_{21}A_{11}^{-1}A_{12}}$$

As indicated earlier, the computation of h_i in this way is inefficient, because it depends on A_{11}^{-1} . But the way in which h_2 was computed (see equations (2.8) and (2.12)) can easily be generalized. Consider formula (2.19), and suppose that initially h_i is taken equal to 1, so that a G.S. step results. Then we obtain for $x_{2g}^{(1)}$ and $x_{2g}^{(2)}$

$$(2.25) \quad x_{2g}^{(1)} = b_2^* - A_{21}b_1^* + A_{21}A_{11}^{-1}A_{12}x_2^{(0)}$$

$$(2.26) \quad x_{2g}^{(2)} = b_2^* - A_{21}b_1^* + A_{21}A_{11}^{-1}A_{12}x_{2g}^{(1)}$$

Subtracting (2.25) from (2.26) results in

$$(2.27) \quad x_{2g}^{(2)} - x_{2g}^{(1)} = A_{21}A_{11}^{-1}A_{12}(x_{2g}^{(1)} - x_2^{(0)})$$

or

$$(2.28) \quad A_{21}A_{11}^{-1}A_{12} = (x_{2g}^{(2)} - x_{2g}^{(1)}) / (x_{2g}^{(1)} - x_2^{(0)})$$

So the value for $A_{21}A_{11}^{-1}A_{12}$, and therefore h_i , can be obtained from two unweighted steps in the computation of x_2 . Once h_i has been computed the solution for x_2 and x_1 , given x_3 , can be obtained by means of (2.18) and (2.15).

Solution of the system as a whole is carried out in a recursive way, by successively putting $i = 2, \dots, n$. Once the optimal h_i has been computed in step i , it need not be recomputed in step $i+1$, since it is a constant, as follows from (2.24). The procedure will be illustrated in detail in the next section.

Finally, it follows from the construction of the solution procedure that under the condition mentioned the procedure always converges to the solution $A^{-1}b$ of $Ax = b$. The proof is by induction and makes use of the following steps:

(i) a system of two linear equations can be made convergent to a finite solution by means of an appropriate weighting factor h_2 ; compare (2.8) and (2.12);

(ii) given the possibility to obtain a solution for the vector x_1 from the first $i-1$ equations, a stepwise procedure for the solution of

x_1 and x_2 (the i -th element of x) given x_3 can be made convergent to a finite limit by means of an appropriate weighting factor h_i ; compare (2.24) and (2.28).

3. APPLICATION TO LINEAR SYSTEMS

In this section we shall describe the results of the application of the M.G.S. method to linear systems. First, the method will be illustrated by means of two very simple systems, namely of three and five equations, respectively.

Consider the following system of three linear equations in three unknowns.

$$(3.1) \quad \begin{aligned} (i) \quad & x_1 + 2x_2 + 3x_3 = 60 \\ (ii) \quad & -2x_1 + x_2 + 2x_3 = 10 \\ (iii) \quad & 4x_1 - 3x_2 + x_3 = 20 \end{aligned}$$

The solution to this system is

$$x_1 = x_2 = x_3 = 10.$$

as easily can be verified. Application of G.S. in such a way that x_i is computed by means of the i^{th} equation results in a divergent process, as can be seen from table 3.1.

TABLE 3.1. APPLICATION OF G.S. TO SYSTEM (3.1)

Iteration $t =$	$x_1^{(t)}$	$x_2^{(t)}$	$x_3^{(t)}$
0		1.0	1.0
1	55.0	118.0	154.0
2	-638.0	-1574.0	-2150.0
3	9658.0	23626.0	32266.0
4	-143990.0	-352502.0	-481526.0
5	2149642.0	5262346.0	7188490.0

To apply M.G.S. two points are to be noted. (i) The variable x_i can be considered only after a solution for x_1, \dots, x_{i-1} has been reached, that holds for the equations 1, ..., $i-1$, given the values of x_1, \dots, x_n .

(ii) After the computation of $x_i^{(t)}$, the solution for x_1, \dots, x_{i-1} from equations 1, ..., $i-1$ has to be adapted, given the new value $x_i^{(t)}$.

Consider again model (3.1) starting with $x_2^{(0)} = x_3^{(0)} = 1.0$ we obtain from equations (3.1.i) and (3.1.ii) the following sequence of values, obtained by G.S. steps.¹⁰

$$\begin{aligned} x_{1g}^{(1)} &= 55.0 & x_{2g}^{(1)} &= 118.0 \\ x_{1g}^{(2)} &= -179.0 & x_{2g}^{(2)} &= -350.0 \end{aligned}$$

Now three successive unweighted values of x_2 are available, which enables us to compute h_2 :

$$\begin{aligned} h_2 &= 1 / \left\{ 1 - \frac{x_{2g}^{(2)} - x_{2g}^{(1)}}{x_{2g}^{(1)} - x_2^{(0)}} \right\} \\ &= 1 / \left\{ 1 - \frac{-468}{117} \right\} = 0.2 \end{aligned}$$

according to formulas (2.8) and (2.12). Then $x_2^{(2)}$ is recomputed as a weighted sum of two successive values

$$\begin{aligned} x_2^{(2)} &= h_2 x_{2g}^{(2)} + (1 - h_2) x_{2g}^{(1)} \\ &= (0.2)(-350.0) + (0.8)(118.0) \\ &= 24.40 \end{aligned}$$

whereas finally we obtain from equation (3.1.i)

$$x_1^{(3)} = 8.20$$

It has to be noted that $x_1^{(3)} = 8.20$ and $x_2^{(2)} = 24.40$ is the solution to equations (3.1.i) and (3.1.ii), given $x_3^{(0)} = 1.00$.

Now $x_{3g}^{(1)}$ is computed by means of equation (3.1.iii), given $x_1^{(3)}$ and $x_2^{(2)}$:

¹⁰ Here the subscript g is introduced again to distinguish between G.S. and M.G.S.

$$x_{3g}^{(1)} = 60.40$$

Then the process turns back to the computation of new values for x_1 and x_2 :

$$x_{1g}^{(4)} = -170.0$$

$$x_2^{(3)} = h_2 x_{2g}^{(3)} + (1 - h_2) x_2^{(2)} = -70.64$$

where $x_{2g}^{(3)}$ is obtained by means of equation (3.1.ii)

$$x_{2g}^{(3)} = 10.0 + 2x_{1g}^{(4)} - 2x_3^{(1)} = -450.80$$

and finally

$$x_1^{(5)} = 20.08$$

Again it can easily be verified that $x_1^{(5)} = 20.08$ and $x_2^{(3)} = -70.64$ is the solution to (3.1.i) and (3.1.ii), given $x_{3g}^{(1)} = 60.40$. Now a new value for x_3 can be computed

$$x_{3g}^{(2)} = -272.24$$

Three successive values for x_3 are available, leading to the weighting factor h_3

$$h_3 = 1 / \left\{ 1 - \frac{x_{3g}^{(2)} - x_{3g}^{(1)}}{x_{3g}^{(1)} - x_3^{(0)}} \right\} = 0.1515$$

and to the ultimate solution for x_3

$$\begin{aligned} x_3^{(2)} &= h_3 x_{3g}^{(2)} + (1 - h_3) x_{3g}^{(1)} \\ &= (0.1515)(-272.24) + (0.8485)(60.40) \\ &= 10.00 \end{aligned}$$

after which the process returns for the last time to equations (3.1.i), (3.1.ii) and (3.1.i) for the computation of x_{1g} , x_2 and x_1 . The whole sequence is given in table 3.2.

TABLE 3.2. APPLICATION OF M.G.S. TO SYSTEM (3.1)

Elementary step	h_i	x_1	x_2	x_3
0		--	1.00	1.00
1		55.00		
2			118.00	
3		-179.00		
4	$\left\{ \begin{array}{l} h_2 = 0.2000 \end{array} \right.$		-350.00 ^a	
5			24.40 ^b	
6		8.20		60.40
7		-170.00		
8			-450.80 ^a	
9			-70.64 ^b	
10	$\left\{ \begin{array}{l} h_3 = 0.1515 \end{array} \right.$	20.08		-272.24 ^a
11				10.00 ^b
12		171.28	332.56 ^a	
13			10.00 ^b	

^a Unweighted value, obtained by a G.S. step.

^b Weighted value.

The iterative scheme for the system of three equations rapidly leads to a solution. If the computation of one value $x_i^{(t)}$ is considered as one elementary step, where the computation of the unweighted values $x_{ig}^{(t)}$ and their weighted counterpart $x_i^{(t)}$ are taken together, a number of 13 elementary steps is needed, apart from the weighting factors h_i . As will be clear, however, the number of steps grows rapidly, when the number of equations becomes larger, due to the fact that every time a new variable x_i , say, has been computed, the values of x_1, \dots, x_{i-1} have to be adapted. In general, for a system of n equations the number of elementary steps is given by

$$(3.2) \quad t = 2^{n+1} - 3$$

This leads to the conclusion that for large systems the method is far less efficient than ordinary matrix inversion.

There are, however, situations where the method becomes more attractive. Consider as an illustration the following system.

$$\begin{aligned}
 (3.3) \quad & \begin{aligned}
 & \text{(i)} \quad x_1 + 9x_5 = 10 \\
 & \text{(ii)} \quad 2x_1 + x_2 + 3x_3 = 6 \\
 & \text{(iii)} \quad 3x_1 - 4x_2 + x_3 = 0 \\
 & \text{(iv)} \quad 5x_1 + 2x_2 - 4x_3 + x_4 = 4 \\
 & \text{(v)} \quad 4x_1 - 2x_2 + x_3 + 3x_4 + x_5 = 1
 \end{aligned}
 \end{aligned}$$

The solution to this system is

$$x_i = 1 \quad i = 1, 2, \dots, 5$$

as can easily be verified. Application of G.S. such that x_i is computed by equation i for every i leads to a divergent process. For a system of five linear equations in five unknowns without zero coefficients M.G.S. needs $t = 2^6 - 3 = 61$ elementary steps, according to (3.2). However, in the case of system (3.3) effective use can be made of the fact that a number of coefficients is zero. It has to be noted that after the computation of $x_i^{(t)}$ all the variables x_1, \dots, x_{i-1} are to be adapted. It will be clear, however, that such an adaptation can be omitted as soon as the values of x_1, \dots, x_{i-1} are not affected by the value of x_i . Such cases arise in the system under consideration, where the variables x_2 and x_3 do not appear in equation (3.3.i), whereas x_4 is not contained in the first three equations (3.3.i - 3.3.iii). This reduces the number of elementary steps to 19, as can be seen from table 3.3, in which the results of the computation are summarized. The way in which in this simple system a considerable saving of computations could be achieved indicates the way in which larger systems with a great number of zero coefficients are to be tackled.

To investigate such a system a matrix A of order 25×25 was constructed in the following way.

- (a) The diagonal elements a_{ii} were put equal to 1.
- (b) For every element a_{ij} ($i \neq j$) a random number was chosen from a rectangular distribution between 0 and 1. For a number ≤ 0.95

TABLE 3.3. APPLICATION OF M.G.S. TO SYSTEM (3.3)

Elementary step	h_i	x_1	x_2	x_3	x_4	x_5
0	--	--	5.00	5.00	5.00	5.00
1		-35.00				
2			61.00			
3				349.00		
4			-971.00			
5	{ $h_3 = 0.0769$			-3779.00 ^a		
6				31.46 ^b		
7			-18.38			
8					341.62	
9						1097.62
10		-9868.54	19648.69			
11				8352.15		
12			-5313.37			
13					93382.02	
14	{ $h_5 = -0.0036$					300642.33 ^a
15						1.00 ^b
16		1.00	-25052.44			
17				1.00		
18			1.00			
19					1.00	

^a Unweighted value^b Weighted value

the element a_{ij} was put equal to zero, whereas for a number > 0.95 the value of the element was obtained by means of a random drawing from a normal distribution, with zero mean and standard deviation 2.

So about 95% of the off-diagonal elements are zero, not so unlikely a situation in an econometric model. This matrix was put in the system.

(3.4)

$$Ax = 1$$

where $\mathbf{1}$ is a column vector of unit elements. By means of computer routines we tried to solve (3.4) in the following three ways.

- (a) Straightforward inversion;
- (b) G.S.
- (c) M.G.S.

However, to apply methods (b) and (c) as efficiently as possible, the rows and columns of A were interchanged by means of a simple heuristic algorithm in such a way that as much as possible zero elements appeared above and right from the main diagonal. Furthermore the non-zero elements above and right from the main diagonal were shifted to the right as far as possible. One restriction was imposed upon these operations: the diagonal elements of the original matrix were maintained on the diagonal. It should be noted that for the application of M.G.S. a vector of indices was constructed, indicating the order in which the computations are to be performed. In this way we avoided to test over and over again the non-diagonal elements.

The results of our computations,¹¹ applied to 10 different systems $Ax = \mathbf{1}$, are as follows. For the computation of 10 solutions by means of straightforward inversion about 65 sec. were needed.¹² M.G.S. needed about 85 sec. for the computation of the same solutions. So for sparse matrices M.G.S. is only slightly less efficient if compared with straightforward inversion. For the comparison of G.S. and M.G.S. see table 3.4. In five of the ten cases G.S. did not lead to a solution because of a divergent process (nrs. 2, 3, 4, 8, 10). In two other cases (6 and 9) the results of the computations were exactly the same for both methods. This was due to the fact that the matrix A , after the interchange of rows and columns, was triangular, as can be seen from the number of elementary operations of M.G.S., which was 25.¹³ Finally, in three cases (1, 5 and 7) both methods lead to the same result, apart from differences after the 5th decimal. However, G.S. is far less efficient, as can be seen from a comparison of the number of elementary operations.

¹¹ Performed by means of the I.B.M. 1130 computer of the Econometric Institute.

¹² The inversion routine is based on the well-known Gauss-Jordan reduction technique.

¹³ The number of elementary operations for G.S. was 50, because an additional round had to be performed in order to test the rate of convergence.

TABLE 3.4. NUMBERS OF OPERATIONS NEEDED FOR THE SOLUTION
OF 10 SYSTEMS OF ORDER 25 BY MEANS OF G.S. AND M.G.S.

System nr.	G.S. ^a		M.G.S. Number of elementary operations ^c
	Number of Iterations ^b	Number of elementary operations	
1	83	2075	327
2	--	--	30
3	--	--	345
4	--	--	282
5	54	1350	38
6	2	50	25
7	28	700	43
8	--	--	352
9	2	50	25
10	--	--	80

^a The computations were terminated when $|x_i^{(t)} - x_i^{(t-1)}| \leq |10^{-7} x_i^{(t-1)}|$ for all i . The results obtained in this way were up to at least 5 decimals equal to those obtained by M.G.S.

^b By one iteration we mean the computation of a new vector x .

^c According to our earlier definition an elementary operation is the computation of one element x_i .

4. THE APPLICATION OF THE METHOD: THE CASE OF A NONLINEAR MODEL

For the application of our method to nonlinear systems we used two economic models, taken from the existing literature.

The first one, a small Keynesian model,¹⁴ runs as follows

$$\begin{aligned}
 (4.1) \quad & (1) C_w = WN \\
 & (2) C_r/P = 10 + 0.6(P_y - WN)/P \\
 & (3) I/P = 30 \\
 & (4) P_y = C_w + C_r + I \\
 & (5) y = 50 + 7N - 0.02 N^2 \\
 & (6) W/P = 7 - 0.04 N
 \end{aligned}$$

where

C_r = consumption of profit recipients

¹⁴ This model was taken from Ackley (1961).

C_w = consumption of wage earners
 I = investment
 N = employment
 P = price level
 W = wage level
 y = real national income

The wage level is taken to be autonomous, and it is supposed to be 5. Then the solution to the model is

$C_r = 70$ $P = 1$
 $C_w = 250$ $W = 5$
 $I = 30$ $y = 350$
 $N = 50$

There are a number of alternative ways in which the model can be solved by means of an iterative procedure. Two of these alternatives were picked out for illustrative purposes. The first one demonstrates the way in which the convergence speed of a convergent G.S.-procedure can be increased by means of M.G.S. The solution scheme which needs no further explanation is given in table 4.1.

TABLE 4.1. ITERATIVE SCHEME FOR THE SOLUTION OF MODEL 4.1.

Equation	Variable					
	N	P	C_w	I	C_r	y
5	x					x
6	x	x				
1	x		x			
3		x		x		
2	x	x			x	x
4		x	x	x	x	x

Starting with $y = y^{(0)} = 300$, the results of the iterations are summarized in table 4.2. For M.G.S. the weighting factor h was computed at the end of iteration 2, and thereafter maintained throughout the whole process. It would have been possible to compute a new h after some iterations, but we learned from some tests that not much was gained in that way.¹⁵

¹⁵ Only one h , related to the last equation of the iterative scheme, has to be computed, because the rest of the system is recursive.

TABLE 4.2. APPLICATION OF ORDINARY AND MODIFIED
GAUSS-SEIDEL TO THE ITERATIVE SCHEME
OF TABLE 4.1.

Iteration	Value of y^a	
	G.S.	M.G.S.
2	313.0705	356.8439
4	323.0316	350.6771
6	330.4806	350.0640
8	335.9683	350.0060
9	338.1271	349.9982
10	339.9648	350.0006
20	348.1957	
30	349.6829	
40	349.9445	
45	349.9768	
46	349.9805	
47	349.9836	

^a The iterative process was terminated when the relative change of $y^{(t)}$ was less than 10^{-5} .

As can be seen from table 4.2 G.S. terminates after 47 iterations, whereas M.G.S. only needs 10 iterations.¹⁶

The second alternative shows a situation in which G.S. did not converge to the solution. The iterative scheme is given in table 4.3. The iterations were started with $I = I^{(0)} = 50$. The G.S. process was stopped after a few iterations, because the values of $I^{(t)}$ increased rapidly, as can be seen from table 4.4. On the other hand, even in this case M.G.S. converges rapidly to the solution, that is obtained after 17 iterations up to three decimal places accurate. These last results are not presented here, however, because in this case convergence could be speeded up by the computation of a new value for h after a number of iterations. In our example h was recomputed after 4 iterations, which lead to a reduction in the number of iterations from 17 to 11, as can be seen from the results as presented in table 4.4.

¹⁶ For this case the number of iterations of G.S. and M.G.S. can be compared directly, because in both procedures every iteration covers all the equations.

TABLE 4.3. ALTERNATIVE ITERATIVE SCHEME
FOR THE SOLUTION OF MODEL 4.1

Equation	Variable					
	P	N	C _w	y	C _r	I
3	x					x
6	x	x				
1		x	x			
5		x		x		
2	x	x		x	x	
4	x		x	x	x	x

TABLE 4.4. ORDINARY AND MODIFIED GAUSS-SEIDEL,
APPLIED TO THE ITERATIVE SCHEME OF TABLE 4.3

Iteration	Value of I ^a	
	G.S.	M.G.S.
1	150.00000	150.00000
2	950.00000	35.71429
3	7728.94741	32.44898
4	65346.53825	31.17278
5	555095.63334	35.26590
6	--	30.17174
8	--	30.00565
10	--	30.00019
11	--	30.00004

^a The iterative process was terminated when the relative change of $I^{(t)}$ was less than 10^{-5} .

As a second example use was made of the Klein III model, a system of 16 non-linear equations in 16 variables.¹⁷ For the specification of the entire model we refer to the publication mentioned. For our purposes it has to be noted that the model can be written as a blockrecursive system. The first block consists of 2 equations, the second block contains 10 equations, whereas finally 4 equations constitute the third block.

¹⁷ This model was taken from Klein (1950). Our solution is for 1940.

We concentrate on the second block. The iterative scheme is given in table 4.5. The order of the equations in this scheme leads to an almost recursive system. Note that both equations and variables have been reordered in the same way, so that all diagonal (unit) elements have been maintained on the diagonal. A procedure like this one probably has been referred to by Klein and Evans.¹⁸ Note, however, that in a simultaneous model "the main lines of economic causation" are not uniquely defined. This last point is stressed, because as will be clear from our results, G.S. applied to the submodel of table 4.5 leads to a divergent process, which seems to be in contradiction with the opinion of Klein and Evans. The difficulties arise because there remain a number of equations which can be ordered in different ways. For the latter ordering there are no simple rules. In table 4.5, for example, the last two rows (and the corresponding columns) could have been interchanged, without disturbing the recursive character of the remaining equations.

TABLE 4.5. ITERATIVE SCHEME FOR THE SOLUTION OF THE KLEIN III MODEL

Equation ^a	Variable									
	X	I	H	r	v	D ₁	C	R ₁	Y	p
49	x									x
39	x	x								x
40	x		x							x
45				x					x	
44				x	x				x	
42				x		x			x	
41							x		x	
37				x	x			x		
34		x	x			x	x		x	
35	x							x	x	x

^a See for the numbering of the equations Klein (1950), pages 108-110.

¹⁸ Klein and Evans (1969) state that for the application of G.S. there is "no simple rule for showing in advance how to normalize and order equations in a model to find convergent algorithms leading to solutions. We do know, in many applied cases, that convergent patterns easily can be found, and the one used in the present model does, in fact, converge readily. If one traces the main lines of economic causation through a model and then normalizes and orders equations to reproduce this pattern of causation, it is likely that a convergent algorithm will result."

The results of our computation are given in table 4.6. The results obtained by means of G.S. need no further comments. With respect to M.G.S. the following can be remarked. Starting with an initial value $p^{(0)}$, values for X , I and H can be obtained. Then, given an initial value $Y^{(0)}$, values for r , v , D_1 , C and R_1 are computed, ultimately leading to a new value for Y . So the equations 45, 44, 42, 41, 37 and 34 form an inner loop, leading to values for r , v , D_1 , C , R_1 and Y by means of an iterative process, given the values of X , I , H and p . For the computation of the variables within the inner loop use is made of a weighting factor h_9 . Once the inner loop has converged, a new value for p is obtained. Then the process returns to equation 49 and the same procedure starts again. For the computation of p also use is made of a weighting factor (h_{10}). It has to be remarked, that the number of iterations could be reduced here by recomputing h_{10} during the procedure. In this case h_{10} was recomputed after every second step, leading to a reduction in the number of steps for the computation of p from 19 to 13.

TABLE 4.6. G.S. AND M.G.S., APPLIED TO THE ITERATIVE SCHEME OF TABLE 4.5 ^{a,b}

G.S.		M.G.S.		
Iteration	p	Iteration	Y	p
1	1.2557	1	62.8700	1.9021
2	14.2053	2	188.8467	1.3412
5	0.0532	3	101.4189	0.8113
10	0.0837	4	38.8034	1.2778
15	0.2627	5	92.8980	0.9417
20	0.1800	6	52.4107	1.2234
30	4.0940	7	85.8172	1.0807
40	0.1600	8	68.2117	1.1942
50	0.1602	9	82.0965	1.1686
60	-2.7910	10	78.8916	1.1883
70	0.1354	11	81.3581	1.1875
80	0.7093	12	81.2579	1.1881
90	0.1342	13	81.3346	1.1881
100	0.0654			

^a For M.G.S. only the number of iterations for the computation of p are counted. For the computation of Y two iterations were sufficient in all cases except the first time, where three iterations were needed.

^b See next page.

It can be concluded, that for the solution of non-linear systems M.G.S. can be a useful tool in all cases where G.S. would be applicable, but not automatically leads to a convergent process. Instead of searching for another ordering such that G.S. converges, use can be made of M.G.S., where only a reordering of the equations is helpful as far as it depends on the recursiveness of parts of the system.

References

- [1] Ackley, G. (1961), Macroeconomic Theory, The Macmillan Company, New York.
- [2] Ball, R.J., B.D. Boatwright, T. Burns, P.W.M. Lobban, and G.W. Miller (1975), "The London Business School Quarterly Econometric Model of the UK Economy", from G.A. Renton (editor), Modelling the Economy, Heinemann Educational Books Limited, London.
- [3] Faddeev, D.K., and V.N. Faddeeva (1963), Computational Methods of Linear Algebra, W.H. Freeman & Comp., San Francisco and London.
- [4] Fromm, G., and L.R. Klein (1969), "Solutions of the Complete System", from J.S. Duessenberry, G. Fromm, L.R. Klein and E. Kuh (editors), The Brookings Model: Some Further Results, North-Holland Publishing Company, Amsterdam/London, Rand McNally & Company, Chicago.
- [5] Fromm, G., and P. Taubman (1967), Policy Simulations with an Econometric Model, North-Holland Publishing Company, Amsterdam/London, Rand McNally & Company, Chicago.
- [6] Hildebrand, F.B. (1974), Introduction to Numerical Analysis, McGraw Hill, New York.
- [7] Klein, L.R. (1950), Economic Fluctuations in the United States 1921-1941, John Wiley & Sons, Inc., New York, Chapman & Hall, Limited, London.
- [8] Klein, L.R., and M.K. Evans (1969), Econometric Gaming, The Macmillan Company, Collier-Macmillan Limited, London.
- [9] Varga, R. (1962), Matrix Iterative Analysis, Prentice Hall, Inc., Englewood Cliffs, New Jersey.

^b The computations were terminated when the relative change of both $y^{(t)}$ and $p^{(t)}$ was less than 10^{-6} .

