Coordinated Science Laboratory

UNIVERSITY OF ILLINOIS - URBANA, ILLINOIS

THE METHOD OF ADAPTIVE CONSTRAINED DESCENT

Stephen J. Kahne

REPORT R-154 **OCTOBER, 1962**

COORDINATED SCIENCE LABORATORY
UNIVERSITY OF ILLINOIS
URBANA, ILLINOIS

Contract DA-36-039-SC-85122
D/A Sub-Task 3-99-01-002

The research reported in this document was made possible by support extended to the University of Illinois, Coordinated Science Laboratory, jointly by the Department of the Army (Signal Corps), Department of the Navy (Office of Naval Research), and the Department of the Air Force (Office of Scientific Research) under Signal Corps Contract DA-36-039-SC-85122.

Abstract

The paper describes a method for determining a minimum of a scalar function of a vector variable. The vector variable may have vector inequality range constraints imposed. The technique ensures monotone descent and is known to be faster than the classical method of "steepest" descent in some cases. Moreover, it is not subject to the usual sensitivity problems which may arise using the steepest descent method.

Introduction

An important problem in numerical analysis is to determine the minimum of a scalar function of a vector variable. In many physical systems a minimum is required subject to the constraint that elements of the vector variable lie in a specified range. Thus, we may state the problem as follows:

Find

$$a = \min f(\underline{x}) \quad \text{where}$$

$$\underline{x}^- \leq \underline{x} \leq \underline{x}^+$$

underlined quantities represent n-vectors. A numerical procedure for the solution to this problem is described below. Convergence of the method to a minimum of the function is monotonic and has been shown to be faster than the classical steepest descent method which has achieved wide spread usage. The technique is adaptive in that, at each step in the iteration, the most sensitive element of the variable is "worked on" for the longest time.

Details of The Technique

Provision is made for giving the following information to the subroutine:

a) Current elements of vector variable (VAR)

b) Largest step size for each element (DVAR)

c) Dimension of vector variable (NVAR)

d) Ratio of new step size to old (GAMMA)

e) Sensitivity to be attained before moving to next element (PRECIS)

f) Major cycle sensitivity to be attained before giving special signal (THRESH)

g) Vector of minima for variables (VARMIN)

h) Vector of maxima for variables (VARMAX)

i) Current value of scalar function to be minimized (ERR)

j) Code number for program control (KEY)

The minimization process may be considered to be a series of n minor cycles making up each major cycle where n = NVAR. Each minor cycle reduces the sensitivity of the criterion function to one variable to a prescribed level. A major cycle is completed when one minor cycle for each variable is completed. A flag (KEY) signals to the calling program that one of three conditions exist. If KEY>0, the subroutine seeks the value of the criterion function (ERR); if KEY=0, a major cycle is about to begin; if KEY=-1, the decrease in criterion function has fallen below the parameter THRESH during the last major cycle. [It follows that KEY=0 is used for the initial entry to the subroutine.]

This routine is not fooled by functions whose gradient in the direction of one variable is very small. In this respect the method is better than the usual steepest descent calculation. In addition, the provision for variable constraints provides a feature not available in classical steepest descent procedures. Techniques for dealing with constrained minimization problems such as the penalty function approach and the gradient projection method do not have the advantage of simplicity which this scheme does in its logical approach to minimization.[1] However, the author is not in a position to evaluate the features of each method to provide a realistic comparison. A quantitive comparison which has been made shows that, for a problem requiring about one second for criterion function evaluation on a CDC 1604 computer the adaptive constrained descent technique (with constraints set very wide, i.e., ignored) took one third as long as the steepest descent approach to the same problem.

The reason for the increase in speed is due to the efficient use of the information carried in successive criterion function evaluations. Whereas the steepest descent method computes the total gradient of the function and calculates the step size to take in that direction, the present method is content to move in a direction along one variable at a time which produces a significant decrease in the criterion function (ERR). Not only does this increase efficiency of the method, but it also ensures that insensitive variables will not cause the routine to spend a large amount of time moving along this variable.

---

1. Leitmann, G., ed., "Optimization Techniques", Academic Press, 1962, Ch. 6.

Further observations

Experience with the method has shown it to be superior to both other methods tried namely the classical steepest descent technique and the "one-at-a-time" procedure. Of course neither of the classical approaches can handle the vector variable constraints without considerable modification. The problem of local minima of a function is a difficulty common to all minimization schemes. The method of adaptive constrained descent is no exception. The minimum which is selected (in a multi-minima function) is related to the starting value of the vector variable and the initial step size vector (DVAR) and the usual precautions must be taken if one seeks either the "nearest" minimum value or the "lowest" minimum value.

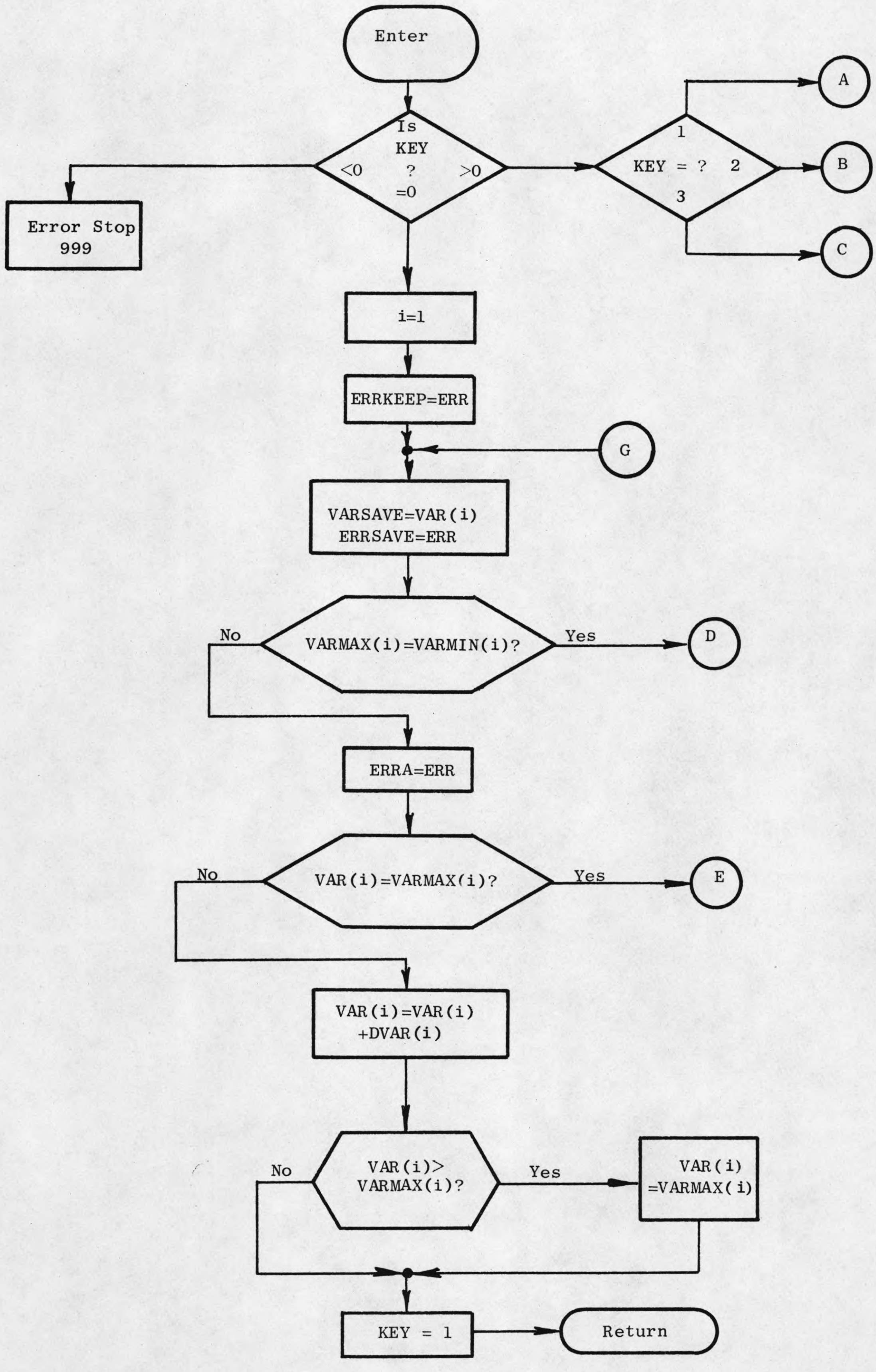
A significant feature of this method is the way in which the vector variable constraints are incorporated into the iterative procedure. In complex physical systems currently being studied in universities and industrial laboratories the presence of such constraints is very frequently required for meaningful analysis and design.

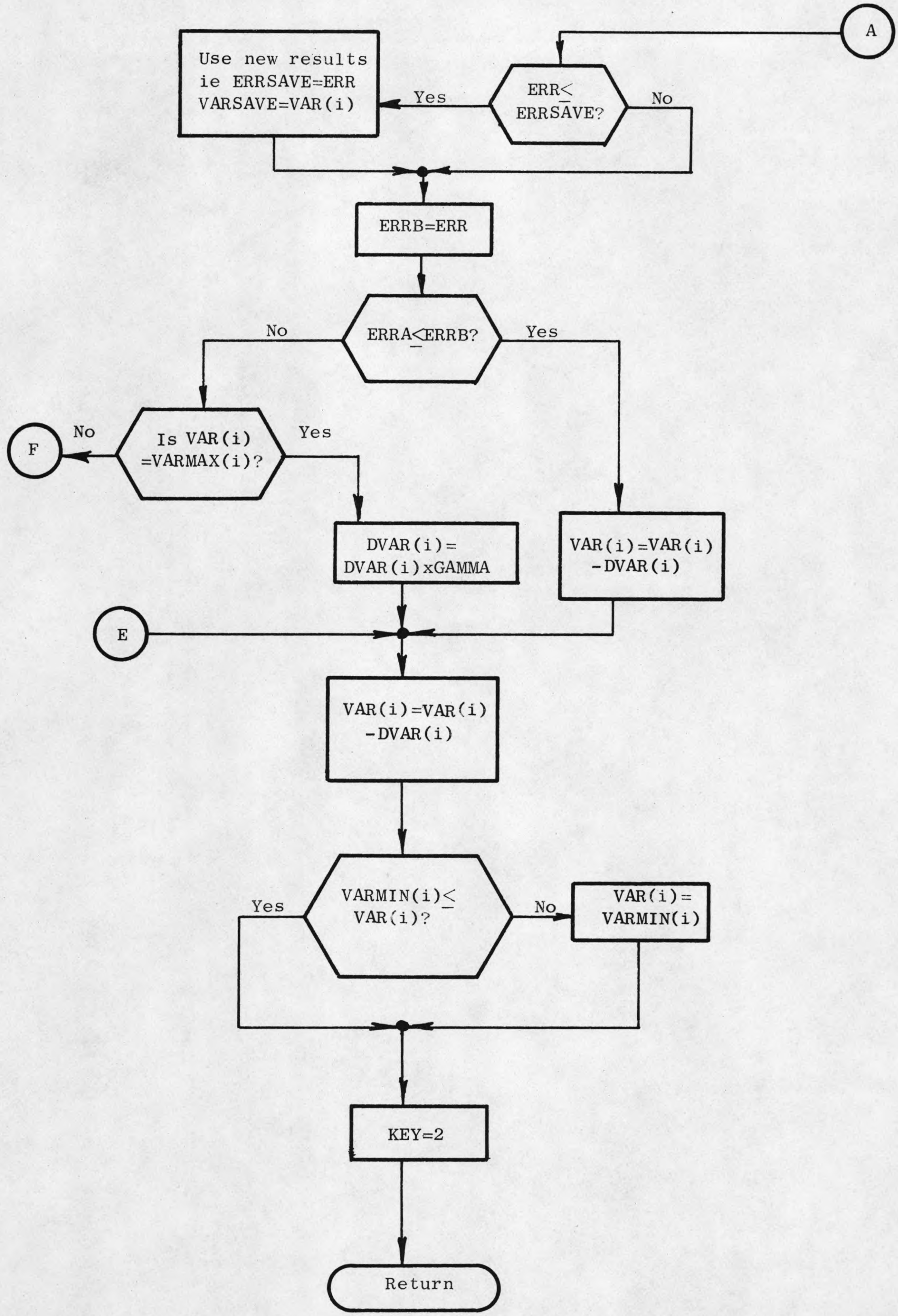
It is not suggested that the logic of the minimization technique presented in this note is new. However, in the area of system theory where such a procedure finds much use, it is observed that very little has been written which describes efficient minimization procedures for the most general and most complicated type of minimization, namely one in which the dimension of the vector variable is very large (above ten) and some of the variables have inequality range constraints imposed. It is in the spirit of describing and documenting a practical numerical technique that this note is written. The author would appreciate reports of experience with this scheme.

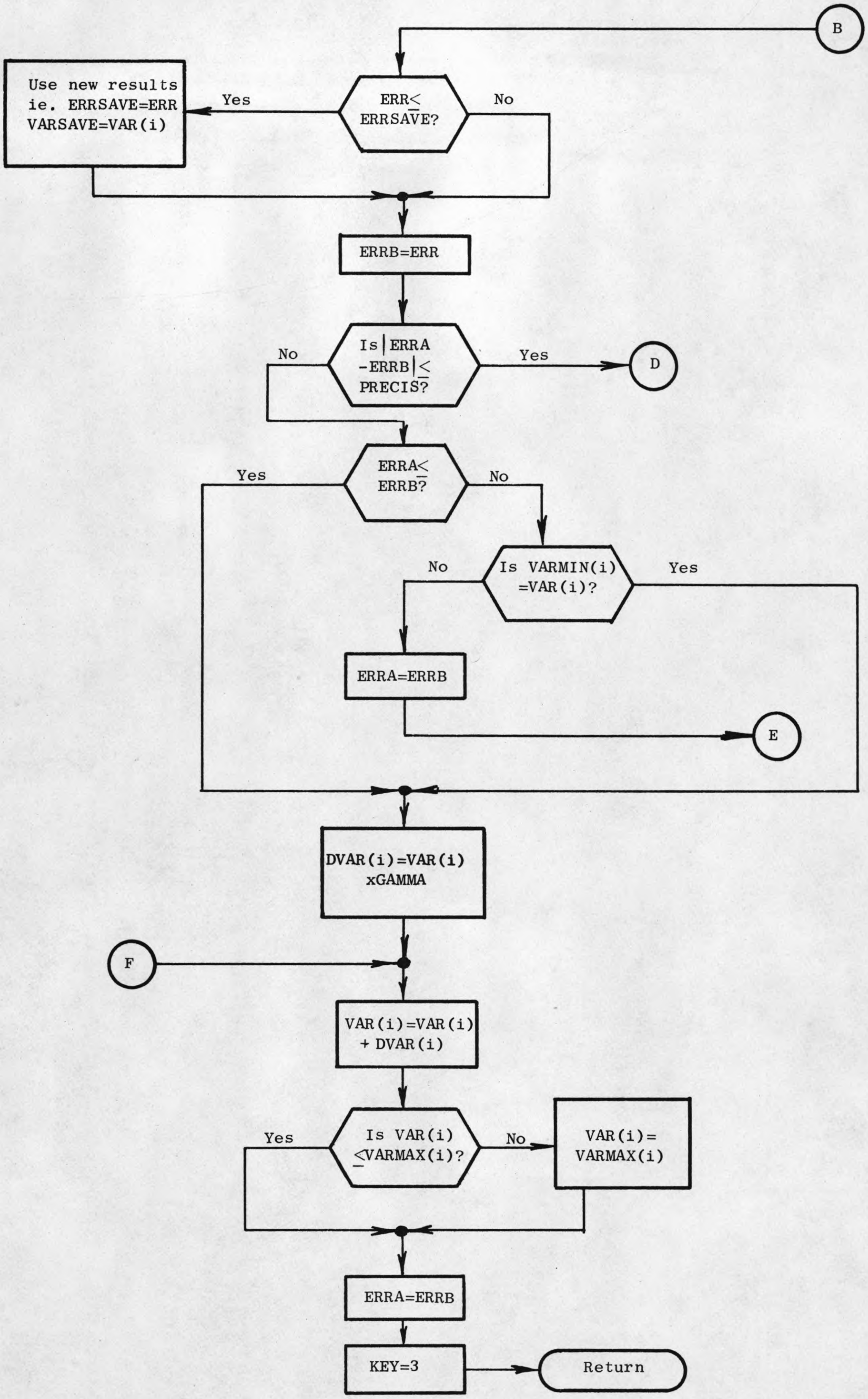
Flow Chart

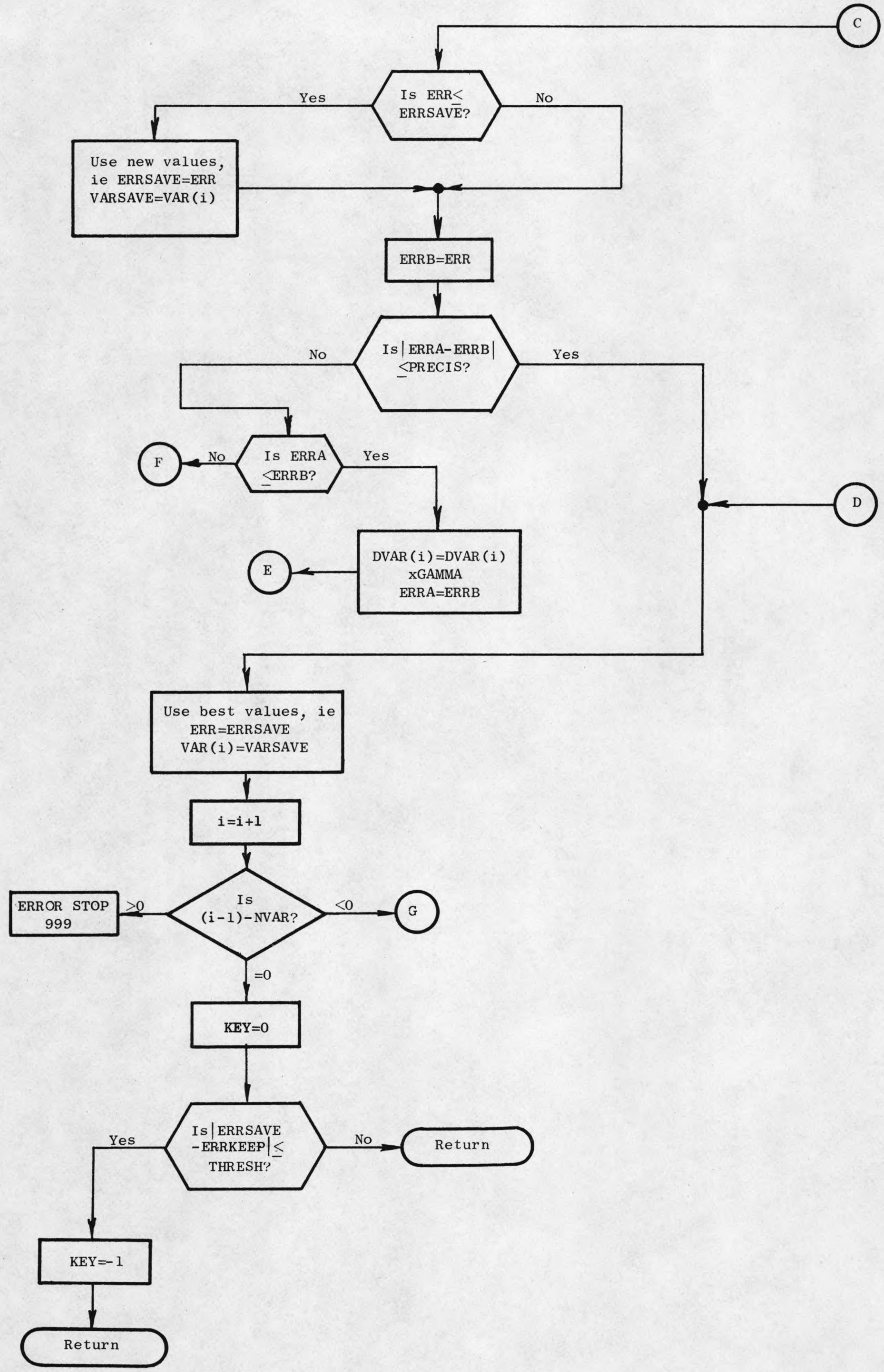
A detailed flow chart of the adaptive constrained descent subroutine is shown in the accompanying figures and a FORTRAN listing is available upon request.

FLOW CHART FOR
ADAPTIVE CONSTRAINED DESCENT
MINIMIZATION PROCEDURE


Enter
Is KEY ?
<0
>0
=0
Error Stop 999
KEY = ?
1
2
3
A
B
C
i=1
ERRKEEP=ERR
G
VARSAVE=VAR(i)
ERRSAVE=ERR
VARMAX(i)=VARMIN(i)?
No
Yes
D
ERRA=ERR
VAR(i)=VARMAX(i)?
No
Yes
E
VAR(i)=VAR(i)
+DVAR(i)
VAR(i)>
VARMAX(i)?
No
Yes
VAR(i)
=VARMAX(i)
KEY = 1
Return


A
ERR≤ ERRSAVE?
Yes
No
Use new results ie ERRSAVE=ERR VARSAVE=VAR(i)
ERRB=ERR
ERRA≤ERRB?
No
Yes
Is VAR(i) =VARMAX(i)?
No
F
Yes
DVAR(i)= DVAR(i)xGAMMA
VAR(i)=VAR(i) -DVAR(i)
E
VAR(i)=VAR(i) -DVAR(i)
VARMIN(i)≤ VAR(i)?
Yes
No
VAR(i)= VARMIN(i)
KEY=2
Return


B
Use new results
ie. ERRSAVE=ERR
VARSAVE=VAR(i)
Yes
ERR≤
ERRSAVE?
No
ERRB=ERR
Is |ERRA
-ERRB|≤
PRECIS?
No
Yes
D
ERRA≤
ERRB?
Yes
No
Is VARMIN(i)
=VAR(i)?
No
Yes
ERRA=ERRB
E
DVAR(i)=VAR(i)
xGAMMA
F
VAR(i)=VAR(i)
+ DVAR(i)
Is VAR(i)
≤VARMAX(i)?
Yes
No
VAR(i)=
VARMAX(i)
ERRA=ERRB
KEY=3
Return


C
Is ERR≤ ERRSAVE?
Yes
No
Use new values, ie ERRSAVE=ERR VARSAVE=VAR(i)
ERRB=ERR
Is |ERRA-ERRB| ≤PRECIS?
No
Yes
Is ERRA ≤ERRB?
No
Yes
F
D
DVAR(i)=DVAR(i) xGAMMA ERRA=ERRB
E
Use best values, ie ERR=ERRSAVE VAR(i)=VARSAVE
i=i+1
Is (i-1)-NVAR?
>0
ERROR STOP 999
<0
G
=0
KEY=0
Is |ERRSAVE -ERRKEEP| ≤ THRESH?
Yes
No
Return
KEY=-1
Return