RICE UNIVERSITY

THE PRERESTORATIVE STEP IN THE SEQUENTIAL GRADIENT-RESTORATION
ALGORITHM FOR MATHEMATICAL PROGRAMMING PROBLEMS
WITH INEQUALITY CONSTRAINTS

by

AN-KUO WU

A THESIS SUBMITTED
IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE

Thesis Director's Signature

Angelo Miele

Houston, Texas
May 1975

ABSTRACT

The Prerestorative Step in the Sequential Gradient-Restoration

Algorithm for Mathematical Programming Problems

with Inequality Constraints

by

An-Kuo Wu

The problem of minimizing a function $f(x)$ subject to the constraint $\psi(x) \geq 0$ is considered. Here, f is a scalar, x is an n -vector, and ψ is a q -vector. The method employed is the sequential gradient-restoration algorithm with complete restoration. In this algorithm, the inequality constraint is transformed into equality constraint by suitable transformations.

The sequential gradient-restoration algorithm is composed of the alternate succession of gradient phases and restoration phases. In the gradient phase, one tries to improve the value of the function while avoiding excessive constraint violation. In the restoration phase, one tries to reduce the constraint error, while avoiding excessive change in the value of the function.

A modification of the sequential gradient-restoration algorithm is presented. This modification consists of inserting a prerestorative step prior to any iteration of the algorithm. The result of this modification is to reduce the constraint violating, while leaving unchanged the value of the function.

Five numerical examples are given to show the considerable beneficial effects associated with the modification presented here.

ACKNOWLEDGEMENT

The author is indebted to Dr. Angelo Miele for suggesting the topic and helpful discussions. This research was supported by the National Science Foundation, Grant No. GP-41158.

To my parents

TABLE OF CONTENTS

	Page
1. Introduction	1
2. Statement of the Problem	2
3. Review of the Sequential Gradient-Restoration Algorithm	5
4. Gradient Stepsize	8
5. Modification of the Algorithm	11
6. Experimental Conditions	12
7. Numerical Examples	14
8. Results and Conclusions	20
9. References	22

1. Introduction

In recent years, considerable attention has been given to iterative algorithms for minimizing a function $f(x)$ subject to the constraint $\varphi(x) = 0$, where f is a scalar, x an n -vector, and φ an q -vector. There are many algorithms available to solve the above problem. Among these, one of the best is the sequential gradient-restoration algorithm with complete restoration (Ref. 1). This algorithm is composed of the alternate succession of gradient phases and restoration phases. In the gradient phase, one tries to reduce the value of the function while avoiding excessive constraint violation. In the restoration phase, one tries to reduce the constraint error, while avoiding excessive change in the value of the function.

Originally, the sequential gradient-restoration algorithm was developed for equality constraint problems. However, it can also be employed for problems involving inequality constraints if suitable transformation are introduced (see, for instance, Ref. 2). Specifically, the inequalities $\psi_i(y) \geq 0$ are transformed into the equalities $\varphi_i \equiv \psi_i(y) - z_i^2 = 0$, where z_i are auxiliary variables.

In practice, there are cases in which the number of inequality constraints is rather large. With particular regard to these problems, it is desirable to find a way to improve the convergence characteristics of the sequential gradient-restoration algorithm.

In this paper, a modification of the sequential gradient-restoration algorithm is presented. It consists of inserting a prerestorative step prior to any iteration of the algorithm. In this prerestorative step, the auxiliary variables z_i are reset at the level $z_i = \sqrt{[\psi_i(y)]}$ if $\psi_i(y) > 0$, and they are left unchanged otherwise. The beneficial effects associated with this modification are illustrated by means of five numerical examples.

2. Statement of the Problem

We consider the problem of minimizing the function

$$g = g(y) \quad (1)$$

subject to the inequality constraint

$$\psi(y) \geq 0 \quad (2)$$

where g is a scalar, y an m -vector, and ψ a q -vector. Here, all vectors are column vectors.

Now, we introduce a q -vector z of auxiliary variables z_i defined by

$$\psi_i(y) - z_i^2 = 0, \quad i = 1, 2, \dots, q \quad (3)$$

If $n = m+q$ is the total number of variables, we construct the n -vector¹

$$x = [y^T, z^T]^T \quad (4)$$

and the q -vector $\varphi(x)$ defined by

$$\varphi_i(x) = \psi_i(y) - z_i^2, \quad i = 1, 2, \dots, q \quad (5)$$

Then, the above problem can be recast as that of minimizing the function

$$f = f(x) \quad (6)$$

subject to the equality constraint

¹ The superscript T denotes the transpose of a matrix.

$$\varphi(x) = 0 \quad (7)$$

It is assumed that the first and second partial derivatives of the functions f and φ with respect to x exist and are continuous; it is also assumed that the constrained minimum exists.

2.1. First Order Conditions. From theory of maxima and minima, it is known that the above problem can be recast as that of minimizing the augmented function

$$F(x, \lambda) = f(x) + \lambda^T \varphi(x) \quad (8)$$

subject to the constraint (7). Here, the q -vector λ is the Lagrange multiplier.

If

$$F_x(x, \lambda) = f_x(x) + \varphi_x(x) \lambda \quad (9)$$

denotes the gradient of the augmented function, the optimum solution for x and λ must satisfy the relations

$$\varphi(x) = 0, \quad F_x(x, \lambda) = 0 \quad (10)$$

which are a system of $n+q$ equations in the $n+q$ components of x and λ . In Eqs. (9)-(10), the gradients f_x and F_x are n -vectors and the matrix φ_x is $n \times q$.

2.2. Approximate Solutions. Since the system (10) is generally nonlinear, approximate methods must be employed. In this connection, we introduce here the scalar performance indexes

$$P(x) = \varphi^T(x) \varphi(x) \quad (11)$$

and

$$Q(x, \lambda) = F_x^T(x, \lambda) F_x(x, \lambda) \quad (12)$$

which measure the errors in the constraint and the optimum condition, respectively.

Observe that $P = 0$ and $Q = 0$ for the exact solution, while $P > 0$ and/or $Q > 0$ for any approximation to the solution. When approximate methods are used, they must ultimately lead to the values of x and λ such that

$$P(x) \leq \epsilon_1 \quad (13)$$

and

$$Q(x, \lambda) \leq \epsilon_2 \quad (14)$$

where ϵ_1 and ϵ_2 are small, preselected numbers.

3. Review of the Sequential Gradient-Restoration Algorithm

The sequential gradient-restoration algorithm with complete restoration (SGRA-CR) is made of the alternate succession of gradient phases and restoration phases. Here, we give a concise summary of this algorithm.

3.1. Gradient Phase. The gradient phase is started providing

$$P(\mathbf{x}) \leq \epsilon_1 \quad (15)$$

It involves a single iteration, in which the augmented function is decreased, that is,²

$$F(\tilde{\mathbf{x}}, \lambda) < F(\mathbf{x}, \lambda) \quad (16)$$

In Eq. (16), the varied point $\tilde{\mathbf{x}}$ is computed with

$$\tilde{\mathbf{x}} = \mathbf{x} - \alpha \mathbf{F}_{\mathbf{x}}(\mathbf{x}, \lambda) \quad (17)$$

$$\varphi_{\mathbf{x}}^T(\mathbf{x}) \varphi_{\mathbf{x}}(\mathbf{x}) \lambda + \varphi_{\mathbf{x}}^T(\mathbf{x}) \mathbf{f}_{\mathbf{x}}(\mathbf{x}) = 0 \quad (18)$$

where α , the gradient stepsize, is a positive quantity.

In practice, certain limitation must be imposed on the magnitude of the displacement $\tilde{\mathbf{x}} - \mathbf{x}$ by limiting the constraint error $P(\tilde{\mathbf{x}})$ at the end of the gradient phase, for instance,

$$P(\tilde{\mathbf{x}}) \leq P_* \quad (19)$$

²The symbol $\mathbf{x}$ denotes the nominal point at the beginning of the gradient phase; the symbol $\tilde{\mathbf{x}}$ denotes the varied point at the end of the gradient phase.

For the computation of α , see Section 4.

3.2. Restoration Phase. The restoration phase is started providing

$$P(x) > \epsilon_1 \quad (20)$$

It involves several iterations, in each of which the constraint error is reduced, that is,³

$$P(\tilde{x}) < P(x) \quad (21)$$

In Eq. (21), the varied point $\tilde{x}$ is computed with

$$\tilde{x} = x - \beta \varphi_x(x) \sigma \quad (22)$$

$$\varphi_x^T(x) \varphi_x(x) \sigma - \varphi(x) = 0 \quad (23)$$

where β , the stepsize, is a positive quantity.

In practice, we first consider the value $\beta = 1$. If Ineq. (21) is satisfied, the value $\beta = 1$ is acceptable. Otherwise, the previous value of β must be replaced by some smaller value (for instance, using a bisection process) until Ineq. (21) is satisfied.

It is emphasized that the restoration phase involves several iterations. The last restorative iteration is characterized by the fact that the varied point $\tilde{x}$ satisfies not only Ineq. (21) but also Ineq. (13).

3.3. Complete Cycle. A complete cycle of the algorithm includes a

³ The symbol x denotes the nominal point at the beginning of a restorative iteration; the symbol $\tilde{x}$ denotes the varied point at the end of a restorative iteration. For the first restorative iteration, the position vector x is identical with the position vector available at the end of the gradient phase.

gradient phase and a restoration phase. If x_1 , x_2 denote the beginning point and the final point of a cycle,⁴ a complete cycle terminates when the descent property

$$f(x_2) < f(x_1) \quad (24)$$

is satisfied. Otherwise, the gradient stepsize α must be bisected as many times as need until, after restoration, Ineq. (24) is satisfied.

⁴ The initial point of a cycle x_1 is identical with the nominal point x of the gradient phase. The final point of a cycle x_2 is identical with the varied point $\tilde{x}$ of the last restorative iteration.

4. Gradient Stepsize

In this section, we use a very simple search scheme to determine the stepsize α for a given $\mathbf{x}$ and λ . If one denotes

$$\mathbf{p} = \mathbf{F}_{\mathbf{x}}(\mathbf{x}, \lambda) \quad (25)$$

Eq. (17) becomes

$$\tilde{\mathbf{x}} = \mathbf{x} - \alpha \mathbf{p} \quad (26)$$

This is a one-parameter family of varied points for which the augmented function and the constraint error take the form

$$\begin{aligned} F(\tilde{\mathbf{x}}, \lambda) &= F(\mathbf{x} - \alpha \mathbf{p}, \lambda) = \tilde{F}(\alpha) \\ P(\tilde{\mathbf{x}}) &= P(\mathbf{x} - \alpha \mathbf{p}) = \tilde{P}(\alpha) \end{aligned} \quad (27)$$

Then, we employ a quadratic interpolation scheme to determine the optimum value α_* , that is, the value such that

$$\tilde{F}_{\alpha}(\alpha_*) = 0 \quad (28)$$

Let $\tilde{F}(\alpha)$ be represented in the quadratic form

$$\tilde{F}(\alpha) = k_0 + k_1 \alpha + k_2 \alpha^2 \quad (29)$$

where k_0, k_1, k_2 are coefficients to be determined. Let $Q(\mathbf{x}, \lambda)$ denote the error in optimality conditions at the nominal point. Let the reference stepsize α_0 be defined by

$$\begin{aligned}
\alpha_0 &= 1/\sqrt{[Q(x, \lambda)]} & \text{if} & \quad Q(x, \lambda) > Q_* \\
\alpha_0 &= 1/\sqrt{Q_*} & \text{if} & \quad Q(x, \lambda) \leq Q_*
\end{aligned} \tag{30}$$

where Q_* is a preselected number.

Consider the succession stepsize

$$0, \alpha_0, 2\alpha_0, 4\alpha_0, 8\alpha_0, 16\alpha_0, \dots \tag{31}$$

Denote by α_1 the smallest stepsize in the sequence (31) such that the following inequality is satisfied:

$$\tilde{F}(\alpha_1) \geq \tilde{F}(0) + \tilde{F}'_\alpha(0) \alpha_1/2 \tag{32}$$

If this is the case, then we determine the coefficients of the quadratic form (29) with the relations

$$\begin{aligned}
k_0 &= \tilde{F}(0) \\
k_1 &= \tilde{F}'_\alpha(0) \\
k_2 &= (1/\alpha_1^2) [\tilde{F}(\alpha_1) - \tilde{F}(0) - \alpha_1 \tilde{F}'_\alpha(0)]
\end{aligned} \tag{33}$$

With the coefficients computed according to (33), the optimum stepsize α_* is determined by

$$\alpha_* = -k_1/2k_2 \tag{34}$$

and it is such that

$$0 \leq \alpha_* \leq \alpha_1 \tag{35}$$

providing Ineq. (32) is satisfied.

Remark. In practice, an upper limit must be imposed on the stepsize either directly or indirectly through the constraint error. In the latter case, the stepsize must be consistent with the inequality

$$\tilde{P}(\alpha) \leq P_* \quad (36)$$

where P_* is a positive number.

In this connection, let α_2 denote the smallest stepsize in the sequence (31) such that Ineq. (36) is violated. If this is the case, then we renounce to optimizing the stepsize and we determine α_* through the relation

$$\alpha_* = \alpha_2/2 \quad (37)$$

5. Modification of the Algorithm

Now, we present a modification of the sequential gradient-restoration algorithm, designed to improve the convergence characteristics. This modification consists of inserting a prerestorative step prior to any iteration of the algorithm. The result of this modification is to reduce the constraint violation.

As Eq. (4) states, the position vector x includes the original vector y and the auxiliary vector z . Let y, z denote quantities evaluated at the beginning of the prerestorative step and let $\bar{y}, \bar{z}$ denote quantities evaluated at the end of the prerestorative step. We propose that the prerestorative step be done as follows:

$$\bar{y} = y \quad (38)$$

and

$$\bar{z}_i = z_i \quad \text{if} \quad \psi_i(y) \leq 0, \quad i = 1, 2, \dots, q \quad (39)$$

$$\bar{z}_i = \sqrt{[\psi_i(y)]} \quad \text{if} \quad \psi_i(y) > 0, \quad i = 1, 2, \dots, q \quad (40)$$

so that

$$f(\bar{x}) = f(x), \quad P(\bar{x}) \leq P(x) \quad (41)$$

In this way, the prerestorative step reduces the constraint error, while leaving the value of the function unchanged.

6. Experimental Conditions

In order to evaluate the proposed modification, five numerical examples were solved using SGRA-CR both in the form given in Section 3 and with the modification given in Section 5. The algorithms were programmed in FORTRAN IV, and the numerical results were obtained using the IBM 370/155 computer and double-precision arithmetic.

Starting Point. For all of the examples, the nominal point chosen to start the algorithm is defined by

$$y_1 = \dots = y_m = z_1 = \dots = z_q \quad (42)$$

In addition, when possible and simple to compute, the original m-vector y is chosen consistently with the inequality constraint.

Reference Stepsize. For the gradient phase, the reference stepsize α_0 is computed with

$$\begin{aligned} \alpha_0 &= 1/\sqrt{[Q(x, \lambda)]} & \text{if} & \quad Q(x, \lambda) > 1 \\ \alpha_0 &= 1 & \text{if} & \quad Q(x, \lambda) \leq 1 \end{aligned} \quad (43)$$

Stepsize Limitation. For the gradient phase, the following limitation was imposed on the stepsize:

$$\tilde{P}(\alpha) \leq 1 \quad (44)$$

Convergence. The convergence of an algorithm is defined through the inequalities

$$P(x) \leq 10^{-8}, \quad Q(x, \lambda) \leq 10^{-4} \quad (45)$$

Nonconvergence. Conversely, nonconvergence is defined by means of the inequalities

$$N > 100 \quad (46-1)$$

or

$$N_b > 20 \quad (46-2)$$

where N is the total number of iterations and N_b is the number of stepsize bisections needed to satisfy the basic descent properties of the algorithm, namely, Ineqs. (16), (21), and (24).

7. Numerical Examples

In this section, five numerical examples are presented. In two examples, the function to be minimized is linear; in the other examples, it is nonlinear. The constraints are given by inequality relations, linear or nonlinear.

Example 7.1. Minimize the function

$$f = -y_1 + y_2 + y_3 - y_4 \quad (49)$$

subject to the inequality constraints

$$\begin{aligned} y_1(1-y_1) &\geq 0 \\ y_2 &\geq 0 \\ y_3 &\geq 0 \\ y_4(1-y_4) &\geq 0 \\ 1 - y_1 y_2 &\geq 0 \\ 1 - y_1 y_3 - y_2^2 &\geq 0 \\ 2 - y_1 y_4 - y_2 y_3 &\geq 0 \\ 1 - y_2 y_4 - y_3^2 &\geq 0 \\ 1 - y_3 y_4 &\geq 0 \end{aligned} \quad (50)$$

Introduce the auxiliary variables z_i defined by

$$\begin{aligned} y_1(1-y_1) - z_1^2 &= 0 \\ y_2 - z_2^2 &= 0 \\ y_3 - z_3^2 &= 0 \\ y_4(1-y_4) - z_4^2 &= 0 \\ 1 - y_1 y_2 - z_5^2 &= 0 \end{aligned} \quad (51)$$

$$\begin{aligned}
1 - y_1 y_3 - y_2^2 - z_6^2 &= 0 \\
2 - y_1 y_4 - y_2 y_3 - z_7^2 &= 0 \\
1 - y_2 y_4 - y_3^2 - z_8^2 &= 0 \\
1 - y_3 y_4 - z_9^2 &= 0
\end{aligned} \tag{51}$$

Then, the previous problem can be recast as that of minimizing (49) subject to the equality constraints (51). The function (49) admits the relative minimum $f = -2.00$ at the point defined by

$$\begin{aligned}
y_1 &= 1.00, & y_2 &= 0.00, & y_3 &= 0.00, & y_4 &= 1.00 \\
z_1 &= 0.00, & z_2 &= 0.00, & z_3 &= 0.00, & z_4 &= 0.00 \\
z_5 &= 1.00, & z_6 &= 1.00, & z_7 &= 1.00, & z_8 &= 1.00 \\
z_9 &= 1.00 \\
\lambda_1 &= -1.00, & \lambda_2 &= -1.00, & \lambda_3 &= -1.00, & \lambda_4 &= -1.00 \\
\lambda_5 &= 0.00, & \lambda_6 &= 0.00, & \lambda_7 &= 0.00, & \lambda_8 &= 0.00 \\
\lambda_9 &= 0.00
\end{aligned} \tag{52}$$

The nominal point chosen for starting the algorithms is

$$y_1 = \dots = y_4 = z_1 = \dots = z_9 = 0.20 \tag{53}$$

Example 7.2. Minimize the function

$$f = y_1^2 + y_2^2 - 16y_1 - 10y_2 \tag{54}$$

subject to the inequality constraints

$$\begin{aligned}
11 - y_1^2 + 6y_1 - 4y_2 &\geq 0 \\
y_1 y_2 - 3y_2 - \exp(y_1 - 3) + 1 &\geq 0 \\
y_1 &\geq 0 \\
y_2 &\geq 0
\end{aligned} \tag{55}$$

Introduce the auxiliary variables z_i defined by

$$\begin{aligned}
11 - y_1^2 + 6y_1 - 4y_2 - z_1^2 &= 0 \\
y_1 y_2 - 3y_2 - \exp(y_1 - 3) + 1 - z_2^2 &= 0 \\
y_1 - z_3^2 &= 0 \\
y_2 - z_4^2 &= 0
\end{aligned} \tag{56}$$

Then, the previous problem can be recast as that of minimizing (54) subject to the equality constraints (56). The function (54) admits the relative minimum $f = -79.81$ at the point defined by

$$\begin{aligned}
y_1 &= 5.24, & y_2 &= 3.75 \\
z_1 &= 0.00, & z_2 &= 0.00, & z_3 &= 2.29, & z_4 &= 1.94 \\
\lambda_1 &= -0.81, & \lambda_2 &= -0.33, & \lambda_3 &= 0.00, & \lambda_4 &= 0.00
\end{aligned} \tag{57}$$

The nominal point chosen for starting the algorithms is

$$y_1 = y_2 = z_1 = \dots = z_4 = 3 \tag{58}$$

Example 7.3. Minimize

$$f = (y_1 - 1)^2 + (y_2 - 2)^2 \quad (59)$$

subject to the inequality constraints

$$-5 \leq y_1 \leq 5, \quad -5 \leq y_2 \leq 5 \quad (60)$$

Introduce the auxiliary variables z_i defined by

$$\begin{aligned} y_1 + 5 - z_1^2 &= 0, & y_2 + 5 - z_2^2 &= 0 \\ 5 - y_1 - z_3^2 &= 0, & 5 - y_2 - z_4^2 &= 0 \end{aligned} \quad (61)$$

Then, the previous problem can be recast as that of minimizing the function (59)

subject to the equality constraints (61). The function (59) admits the relative minimum $f = 0$ at the point defined by

$$\begin{aligned} y_1 &= 1.00, & y_2 &= 2.00 \\ z_1 &= 2.45, & z_2 &= 2.65, & z_3 &= 2.00, & z_4 &= 1.73 \\ \lambda_1 &= 0.00, & \lambda_2 &= 0.00, & \lambda_3 &= 0.00, & \lambda_4 &= 0.00 \end{aligned} \quad (62)$$

The nominal point chosen for starting the algorithms is

$$y_1 = y_2 = z_1 = z_2 = z_3 = z_4 = 1 \quad (63)$$

Example 7.4. Minimize

$$f = 0.01(y_1 - 1)^2 + (y_2 - y_1^2)^2 \quad (64)$$

subject to the inequality constraint

$$y_1 \leq -1 \quad (65)$$

Introduce the auxiliary variables z_1 defined by

$$-y_1 - 1 - z_1^2 = 0 \quad (66)$$

Then, the previous problem can be recast as that of minimizing the function (64) subject to the equality constraint (66). The function (64) admits the relative minimum $f = 0.04$ at the point defined by

$$\begin{aligned} y_1 &= -1.00, & y_2 &= 1.00 \\ z_1 &= 0.00 \\ \lambda_1 &= -0.04 \end{aligned} \quad (67)$$

The nominal point chosen for starting the algorithms is

$$y_1 = y_2 = z_1 = -2 \quad (68)$$

Example 7.5. Minimize

$$f = -(2y_1 + y_2) \quad (69)$$

subject to the inequality constraints

$$\begin{aligned} y_1(1 - y_1) &\geq 0 \\ y_2(1 - y_2) &\geq 0 \\ 1 - 2y_1y_2 &\geq 0 \end{aligned} \quad (70)$$

Introduce the auxiliary variables z_i defined by

$$\begin{aligned} y_1(1 - y_1) - z_1^2 &= 0 \\ y_2(1 - y_2) - z_2^2 &= 0 \\ 1 - 2y_1y_2 - z_3^2 &= 0 \end{aligned} \tag{71}$$

Then, the previous problem can be recast as that of minimizing (69) subject to the equality constraints (71). The function (69) admits the relative minimum $f = -2.50$ at the point defined by

$$\begin{aligned} y_1 &= 1.00 , & y_2 &= 0.50 \\ z_1 &= 0.00 , & z_2 &= 0.50 , & z_3 &= 0.00 \\ \lambda_1 &= -1.50 , & \lambda_2 &= 0.00 , & \lambda_3 &= -0.50 \end{aligned} \tag{72}$$

The nominal point chosen for starting the algorithms is

$$y_1 = y_2 = z_1 = z_2 = z_3 = 0.50 \tag{73}$$

8. Results and Conclusions

The examples described in Section 7 were solved with the sequential gradient-restoration algorithm with complete restoration (SGRA-CR); both the standard algorithm of Section 3 and the modification proposed in Section 5 were employed. The numerical results are presented in Table 1, where the number of iterations for convergence N is shown. Also shown is the cumulative number of iterations for convergence ΣN for the five examples considered. From this table, it is clear that the modification proposed in this thesis improves considerably the convergence characteristics of the sequential gradient-restoration algorithm.

Remark. After completing this work, three additional numerical examples were solved on IBM 370/155 for functions involving up to 9 variables, subject to up to 15 inequality constraints. The qualitative and quantitative results obtained were consistent with those presented here.

Table 1. Number of iterations for convergence N for SGRA-CR.

	Standard algorithm	Modified algorithm
Example 7.1	17	8
Example 7.2	28	12
Example 7.3	8	1
Example 7.4	19	11
Example 7.5	16	11
ΣN	88	43

References

1. MIELE, A., HUANG, H.Y., and HEIDEMAN, J.C., Sequential Gradient-Restoration Algorithm for the Minimization of Constrained Functions, Ordinary and Conjugate Versions, Journal of Optimization Theory and Applications, Vol. 4, No. 4, 1969.
2. MIELE, A., TIETZE, J.L., and LEVY, A.V., Comparison of Several Gradient Algorithms for Mathematical Programming Problems, Rice University, Aero-Astronautics Report No. 94, 1972.