11 SOME PROPERTIES OF A NEW CONJUGATE GRADIENT METHOD

Y. H. Dai and Y. Yuan

State Key Laboratory of Scientific and Engineering Computing,
Institute of Computational Mathematics and Scientific/Engineering Computing,
Chinese Academy of Sciences, P. O. Box 2719, Beijing 100080, China

E-mail: dyh@cc.ac.cn, yyx@cc.ac.cn

Abstract: It is proved that the new conjugate gradient method proposed by Dai and Yuan [5] produces a descent direction at each iteration for strictly convex problems. Consequently, the global convergence of the method can be established if the Goldstein line search is used. Further, if the function is uniformly convex, two Armijo-type line searches, the first of which is the standard Armijo line search, are also shown to guarantee the convergence of the new method.

Keywords: unconstrained optimization, conjugate gradient, line search, convex, global convergence.

1 INTRODUCTION

Consider the unconstrained optimization problem

$$\min f(x), \quad x \in R^n, \tag{1.1}$$

where f is smooth and its gradient g is available. Conjugate gradient methods for solving (1.1) are iterative methods of the form

$$x_{k+1} = x_k + \alpha_k d_k, \tag{1.2}$$

$$d_k = \begin{cases} -g_k, & \text{for } k = 1; \\ -g_k + \beta_k d_{k-1}, & \text{for } k \geq 2, \end{cases} \tag{1.3}$$

where $\alpha_k > 0$ is a steplength obtained by a 1-dimensional line search and β_k is a scalar. The choice of β_k should be such that (1.2)–(1.3) reduces to the linear conjugate gradient method in the case when f is a strictly convex quadratic and α_k is the exact 1-dimensional minimizer. Well-known formulas for β_k are called the Fletcher-Reeves [10], Polak-Ribière-Polyak [23; 24], and Hestenes-Stiefel [15] formulas. Their convergence properties have been reported by many authors, including [2; 6; 7; 11; 14; 18; 22; 27]. Nice reviews of the conjugate gradient method can be seen in [11] and [20].

In [5], a new nonlinear conjugate gradient method is presented, which has the following formula for β_k:

$$\beta_k^{DY} = \|g_k\|^2 / d_{k-1}^T y_{k-1}. \tag{1.4}$$

It was shown in [5] that such a method can guarantee the descent property of each direction provided the steplength satisfies the Wolfe conditions (see [26]), namely,

$$f(x_k + \alpha_k d_k) - f(x_k) \leq \delta \alpha_k g_k^T d_k, \tag{1.5}$$

$$g(x_k + \alpha_k d_k)^T d_k \geq \sigma g_k^T d_k, \tag{1.6}$$

where $0 < \delta < \sigma < 1$. In this case, the global convergence of the method was also proved in [5] under some mild assumptions on the objective function. More exactly, we assume that f satisfies

Assumption 1.1 *(1) f is bounded below in the level set $\mathcal{L} = \{x \in \Re^n : f(x) \leq f(x_1)\}$; (2) In some neighborhood $\mathcal{N}$ of $\mathcal{L}$, f is continuously differentiable, and its gradient g is Lipschitz continuous, namely, there exists a constant $L > 0$ such that*

$$\|g(x) - g(y)\| \leq L\|x - y\|, \text{ for any } x, \ y \in \mathcal{N}. \tag{1.7}$$

In addition, based on this method, an algorithm using the Wolfe line search is explored in [5], which performs much better than the Polak-Ribière-Polyak method on the given 18 unconstrained optimization problems in [19].

In this paper, we will study the convergence properties of the new method for convex problems. We will prove that, without any line searches, the new method can also guarantee a descent direction at each iteration for strictly convex functions (see Theorem 2.1). Consequently, the global-convergence of the method is proved if the steplength is chosen by the Goldstein line search.

Further, if the function is uniformly convex, two Armijo-type line searches, the first of which is the standard Armijo line search, are also shown to guarantee the convergence of the new method (see Theorems 2.3 and 2.4). As a marginal note, the global and superlinear convergence of the BFGS method using the second Armijo-type line search for uniformly convex problems is also referred to (see §3). Some other remarks are also given in the last section.

2 MAIN RESULTS

In this section, we assume that f satisfies Assumption 1.1 and $\mathcal{L}$ is a convex set. In this case, we say that f is convex on $\mathcal{L}$ if

$$(g(x) - g(y))^T (x - y) \geq 0; \quad \text{for any } x, y \in \mathcal{L}; \tag{2.1}$$

and that f is strictly convex on $\mathcal{L}$ if

$$(g(x) - g(y))^T (x - y) > 0, \quad \text{for any } x, y \in \mathcal{L},\ \S \neq \dagger. \tag{2.2}$$

We also say that f is uniformly convex on L if there exists a constant $\eta > 0$ such that

$$(g(x) - g(y))^T (x - y) \geq \eta \|x - y\|^2, \quad \text{for all } x,\ y \in \mathcal{L}. \tag{2.3}$$

Note that f has a unique minimizer on $\mathcal{L}$ if f is uniformly convex, whereas there is possibly no any minimizer of f on $\mathcal{L}$ if f is only a strictly convex function. To show this, a 1-dimensional example can be drawn from [16], which is

$$f(x) = e^{-x}, \quad x \in R^1. \tag{2.4}$$

In the following theorems, we always assume that

$$\|g_k\| \neq 0, \quad \text{for all } k, \tag{2.5}$$

for otherwise, a stationary point has already been found.

Theorem 2.1 *Suppose that x_1 is a starting point for which Assumption 1.1 holds. Consider the method (1.2)–(1.3) where β_k is given by (1.4). Then if f is strictly convex on $\mathcal{L}$, we have that for all $k \geq 1$,*

$$g_k^T d_k < 0. \tag{2.6}$$

Proof. (2.6) clearly holds due to $d_1 = -g_1$. Suppose that (2.6) holds for some $k-1$. Since f is strictly convex, we have from (1.2) and (2.2) that

$$d_{k-1}^T y_{k-1} > 0. \tag{2.7}$$

Multiplying (1.3) with g_k and applying (1.4), we obtain

$$g_k^T d_k = \frac{\|g_k\|^2}{d_{k-1}^T y_{k-1}} g_{k-1}^T d_{k-1}, \tag{2.8}$$

which with the induction supposition and (2.7) implies that $g_k^T d_k < 0$. Thus by induction, (2.6) holds for all $k \geq 1$. □

Thus we have proved that the new method without any line searches can provide a descent direction for strictly convex problems unless the gradient norm at the current point is zero. We now conclude that if, further, the steplength α_k is chosen by the Goldstein line search, there exists at least a subsequence of $\{\|g_k\|\}$ generated by the new method converges to zero. The Goldstein line search, first presented by Goldstein [12], accepts a steplength $\alpha_k > 0$ if it satisfies

$$\delta_1 \alpha_k g_k^T d_k \leq f(x_k + \alpha_k d_k) - f_k \leq \delta_2 \alpha_k g_k^T d_k, \tag{2.9}$$

where $0 < \delta_2 < 1/2 < \delta_1 < 1$.

Theorem 2.2 *Suppose that x_1 is a starting point for which Assumption 1.1 holds. Consider the method (1.2)–(1.3) where β_k is given by (1.4). Then if f is strictly convex on $\mathcal{L}$ and if α_k is chosen by the Goldstein line search, we have that* $\liminf \|g_k\| = 0$.

Proof. First, it follows by the mean value theorem and (1.7) that

$$\begin{aligned} f(x_k + \alpha_k d_k) - f_k &= \int_0^1 g(x_k + t\alpha_k d_k)^T (\alpha_k d_k) dt \\ &= \alpha_k g_k^T d_k + \alpha_k \int_0^1 [g(x_k + t\alpha_k d_k) - g_k]^T d_k dt \\ &\leq \alpha_k g_k^T d_k + \frac{1}{2} L \alpha_k^2 \|d_k\|^2. \end{aligned} \tag{2.10}$$

The above relation and the first inequality in (2.9) imply that

$$\alpha_k \geq c \frac{|g_k^T d_k|}{\|d_k\|^2}, \tag{2.11}$$

where $c = \frac{2(1-\delta)}{L}$. Because f is bounded below, we have from (2.9) that

$$\sum_{k \geq 1} \alpha_k |g_k^T d_k| < \infty. \tag{2.12}$$

Thus by (2.11) and (2.12), it follows that

$$\sum_{k\geq 1} \frac{(g_k^T d_k)^2}{\|d_k\|^2} < \infty. \tag{2.13}$$

We now proceed by contradiction and assume that $\liminf_{k\to\infty} \|g_k\| \neq 0$. Then there exists a constant $\tau > 0$ such that for all $k \geq 1$,

$$\|g_k\| \geq \tau. \tag{2.14}$$

Noting that the fraction in (2.8) is just the formula (1.4), we also have that

$$\beta_k = \frac{g_k^T d_k}{g_{k-1}^T d_{k-1}}. \tag{2.15}$$

(1.3) can be re-written as

$$d_k + g_k = \beta_k d_{k-1}. \tag{2.16}$$

Squaring both sides of the above equation, we get that

$$\|d_k\|^2 = \beta_k^2 \|d_{k-1}\|^2 - 2g_k^T d_k - \|g_k\|^2. \tag{2.17}$$

Dividing both sides by $(g_k^T d_k)^2$ and applying (2.15),

$$\frac{\|d_k\|^2}{(g_k^T d_k)^2} = \frac{\|d_{k-1}\|^2}{(g_{k-1}^T d_{k-1})^2} - \frac{2}{g_k^T d_k} - \frac{\|g_k\|^2}{(g_k^T d_k)^2}. \tag{2.18}$$

On the other hand, if we denote

$$l_{k-1} = \frac{g_k^T d_{k-1}}{g_{k-1}^T d_{k-1}}, \tag{2.19}$$

(2.8) is equal to

$$g_k^T d_k = \frac{1}{(l_{k-1} - 1)} \|g_k\|^2. \tag{2.20}$$

Substituting this into (2.18), we can get that

$$\frac{\|d_k\|^2}{(g_k^T d_k)^2} = \frac{\|d_{k-1}\|^2}{(g_{k-1}^T d_{k-1})^2} + \frac{1 - l_{k-1}^2}{\|g_k\|^2}. \tag{2.21}$$

Summing this expression and noting that $d_1 = -g_1$, we obtain

$$\frac{\|d_k\|^2}{(g_k^T d_k)^2} \leq \sum_{i=1}^{k} \frac{1}{\|g_i\|^2}. \tag{2.22}$$

Then we have from (2.22) and (2.14) that

$$\frac{\|d_k\|^2}{(g_k^T d_k)^2} \le \frac{k}{\tau^2}, \tag{2.23}$$

which implies that

$$\sum_{k\ge 1} \frac{(g_k^T d_k)^2}{\|d_k\|^2} = \infty. \tag{2.24}$$

Thus (2.13) and (2.24) give a contradiction, which concludes the proof. □

The following theorem is given to the standard Armijo line search. This line search, first studied by Armijo [1], is to determine the smallest integer $m \ge 0$ such that, if one defines

$$\alpha_k = \lambda^m, \tag{2.25}$$

then

$$f(x_k + \alpha_k d_k) - f_k \le \delta \alpha_k g_k^T d_k. \tag{2.26}$$

Here λ and δ are any positive parameters less than 1.

Theorem 2.3 *Suppose that x_1 is a starting point for which Assumption 1.1 holds. Consider the method (1.2)–(1.3) where β_k is given by (1.4). Then if f is uniformly convex on $\mathcal{L}$ and if α_k is chosen by the Armijo line search, there exists a constant $c_1 > 0$ such that for all $k \ge 1$,*

$$g_k^T d_k \le -c_1 \|g_k\|^2. \tag{2.27}$$

((2.27) is called in [11] the sufficient descent condition holds.) Further, we have that $\lim \|g_k\| = 0$.

Proof. It follows from Theorem 2.1 that (2.6) holds for all $k \ge 1$. Similarly to (2.10), we can prove by the mean value theorem and (2.3) that

$$f(x_k + \alpha_k d_k) - f(x_k) \ge \alpha_k g_k^T d_k + \frac{1}{2}\eta \alpha_k^2 \|d_k\|^2. \tag{2.28}$$

Then from (2.6), (2.26) and (2.28), we get that

$$\alpha_k \le c_2 \frac{|g_k^T d_k|}{\|d_k\|^2}, \tag{2.29}$$

where $c_2 = \frac{2(1-\delta)}{\eta}$ is constant. Besides it, the Lipschitz condition (1.7) gives

$$|g_{k+1}^T d_k - g_k^T d_k| \le \|g_{k+1} - g_k\| \|d_k\| \le \alpha_k \|d_k\|^2. \tag{2.30}$$

Thus by (2.6) and (2.30),

$$l_k - 1 = \frac{g_{k+1}^T d_k - g_k^T d_k}{g_k^T d_k} \geq \frac{\alpha_k \|d_k\|^2}{g_k^T d_k} \geq -Lc_2. \tag{2.31}$$

Since we also have that $l_k - 1 < 0$ due to (2.6) and (2.3), it follows from this, (2.31) and (2.20) that (2.27) holds with $c_1 = \frac{1}{Lc_2}$.

We now proceed by contradiction and assume that (2.14) always holds for some constant $\tau > 0$. Under Assumption 1.1 on f, it can be shown (for example, see [3]) that if the steplength α_k is chosen by the Armijo line search, either

$$\alpha_k = 1 \tag{2.32}$$

or

$$\alpha_k \geq c_3 \frac{|g_k^T d_k|}{\|d_k\|^2} \tag{2.33}$$

holds for every k, where $c_3 > 0$ is some constant. If there exists an infinite subsequence, $\{k_i\}$ say, such that (2.32) holds. Then summing (2.26) over the iterates and noting that f is bounded below, we have that

$$\lim_{i \to \infty} g_{k_i}^T d_{k_i} = 0. \tag{2.34}$$

This, (2.14) and (2.27) clearly give a contradiction. Thus (2.33) must hold for all sufficient large k. In this case, similarly to the proof of Theorem 2.2, we have that (2.24) and (2.13) hold, which contradict each other. Therefore we must have $\lim \|g_k\| = 0$. □

In the following, we turn our attention to another Armijo-type line search and re-establish the global convergence of the new method. Given any parameters $\lambda \in (0, 1)$ and $\delta > 0$, this line search is to determine the smallest integer $m \geq 0$ such that, if one defines

$$\alpha_k = \lambda^m, \tag{2.35}$$

then

$$f(x_k + \alpha_k d_k) - f_k \leq -\delta \alpha_k^2 \|d_k\|^2. \tag{2.36}$$

Such a line search is a simplified version of those proposed in [17] and [13], in connection with no-derivative methods for unconstrained optimization. Note also that based on the line searches proposed in [17] and [13], a new line search technique was designed in [14] which guarantees the global convergence of the Polak-Ribière-Polyak conjugate gradient method. For the clarity in notation, we call the line search (2.35)–(2.36) as the second Armijo-type line search.

Theorem 2.4 *Suppose that x_1 is a starting point for which Assumption 1.1 holds. Consider the method (1.2)–(1.3) where β_k is given by (1.4). If f is uniformly convex on $\mathcal{L}$ and if α_k is chosen by the second Armijo-type line search, then (2.27) holds for some constant $c_1 > 0$ and all $k \geq 1$. Further, we have that* $\lim \|g_k\| = 0$.

Proof. It follows from (2.28) and (2.36) that

$$\alpha_k \leq \frac{1}{\frac{1}{2}\eta + \delta} \frac{|g_k^T d_k|}{\|d_k\|^2}. \tag{2.37}$$

Therefore, similar to the proof of Theorem 2.3, we can show that (2.27) holds for some constant $c_1 > 0$.

Because $\|g_k\|$ is bounded, (2.27) implies that

$$\|d_k\| \geq c_1 \|g_k\|. \tag{2.38}$$

If $\alpha_k < 1$, the line search implies that

$$f(x_k + \lambda^{-1}\alpha_k d_k) - f_k > -\delta \lambda^{-2} \alpha_k^2 \|d_k\|^2. \tag{2.39}$$

On the other hand, similar to (2.10), we have that

$$f(x_k + \lambda^{-1}\alpha_k d_k) - f_k \leq \lambda^{-1}\alpha_k g_k^T d_k + \frac{1}{2} L \lambda^{-2} \alpha_k^2 \|d_k\|^2. \tag{2.40}$$

Combining (2.39) and (2.40), we can see that (2.33) holds with $c_3 = \frac{2\lambda}{L+2\delta}$. Thus it follows from (2.36), (2.38) and (2.27) that

$$\begin{aligned} f(x_k) - f(x_{k+1}) &\geq \min\left[\delta \|d_k\|^2, \delta c_3^2 \frac{(g_k^T d_k)^2}{\|d_k\|^2}\right] \\ &\geq \min\left[\delta c_1^2 \|g_k\|^2, \delta c_3^2 \frac{(g_k^T d_k)^2}{\|d_k\|^2}\right]. \end{aligned} \tag{2.41}$$

Therefore, if the theorem is not true, there exists a constant $c_4 > 0$ such that

$$f(x_k) - f(x_{k+1}) \geq c_4 \min\left[1, \frac{(g_k^T d_k)^2}{\|d_k\|^2}\right] \tag{2.42}$$

for all k. Because $f(x_k)$ is bounded below, we have that

$$\sum_{k \geq 1} \frac{(g_k^T d_k)^2}{\|d_k\|^2} < \infty. \tag{2.43}$$

The above inequality and the proof of Theorem 2.2 implies that $\lim \|g_k\| = 0$. This completes our proof. □

3 SOME REMARKS

The Goldstein line search and the Armijo line search were designed respectively by Goldstein [12] and Armijo [1] to ensure the global convergence of the steepest descent method. Under these line searches, it was shown in [25] and [3] that there are the global and superlinear convergences of the BFGS method for uniformly convex problems. One can see without difficulty that these results also apply to the second Armijo-type line search (2.35)–(2.36). For the latter case, by [3], it suffices to note that if $\alpha_k = 1$ for some k, we have from this and (2.36) that the relation (3.9) in [3] holds with $\eta = \delta(\beta')^{-2}$.

Assume that the line search conditions are (1.2)–(1.3). It was shown in [4] and [7] that if the parameter $\sigma \in (0,1)$ is specifically chosen, the Fletcher-Reeves method and the Polak-Ribière-Polyak method may fail due to producing an uphill direction even if f is a 1-dimensional function in the form

$$f(x) = \frac{t}{2}x^2, \quad x \in R^1, \tag{3.1}$$

where $t > 0$ is some constant. In [9], another conjugate gradient method was proposed which can provide the descent property if the steplength satisfies (1.2)–(1.3) in which $\sigma \in (0,1)$ is any. This method, called conjugate descent method, has the following formula for β_k,

$$\beta_k^{CD} = \|g_k\|^2/(-d_k^T g_k). \tag{3.2}$$

However, the convergence of the conjugate descent method can only be obtained (see [8]) under the line search conditions (1.5) and

$$\sigma g_k^T d_k \leq g(x_k + \alpha_k d_k)^T d_k \leq 0, \tag{3.3}$$

where also $0 < \delta < \sigma < 1$. For any constant $\sigma_1 > 0$, a convex example is given in [8] which shows that the conjugate descent method needs not converge if the line search conditions are (1.5) and

$$\sigma g_k^T d_k \leq g(x_k + \alpha_k d_k)^T d_k \leq -\sigma_1 g_k^T d_k. \tag{3.4}$$

The new method has the nice property of providing a descent search direction for any nonzero steplength provided that the objective function is strictly convex. For general functions, one can show that either d_k^{CD} or d_k^{DY} is a descent direction, where d_k^{CD} and d_k^{DY} are search directions generated by the conjugate descent method and the new method respectively. Therefore it is possible to construct an *ad hoc* efficient method by combining the conjugate descent method and the new method.

Acknowledgment

This work was supported by the National Natural Science Foundation of China.

References

[1] L. Armijo, *Minimization of functions having Lipschitz continuous first partial derivatives*, Pacific Journal of Mathematics, 16 (1) (1966), pp. 1-3

[2] M. Al-Baali, *Descent property and global convergence of the Fletcher-Reeves method with inexact line search,* IMA J. Numer. Anal. 5 (1985), pp. 121-124

[3] R. H. Byrd and J. Nocedal, *A Tool for the Analysis of Quasi-Newton Methods with Application To Unconstrained Minimization*, SIAM J. Numer. Anal. 26(3), (1989), pp. 727-739.

[4] Y. H. Dai, *Analyses of Nonlinear Conjugate Gradient Methods*, Ph.D. thesis, Institute of Computational Mathematics and Scientific/Engineering Computing, Chinese Academy of Sciences, 1997. (in Chinese)

[5] Y. H. Dai and Y. Yuan, *A Nonlinear conjugate Gradient Method with Nice Global Convergence Properties*, Research report ICM-95-038, Institute of Computational Mathematics and Scientific/Engineering Computing, Chinese Academy of Sciences, 1995.

[6] Y. H. Dai and Y. Yuan, *Further studies on the Polak-Ribière-Polyak method*, Research report ICM-95-040, Institute of Computational Mathematics and Scientific/Engineering Computing, Chinese Academy of Sciences, 1995.

[7] Y. H. Dai and Y. Yuan, *Convergence properties of the Fletcher-Reeves method*, IMA J. Numer. Anal. Vol. **16** No. 2 (1996), pp. 155-164.

[8] Y. H. Dai and Y. Yuan, *Convergence Properties of the Conjugate Descent Method*, Mathematical Advances, No. **6** (1996), pp. 552-562.

[9] R. Fletcher (1987), *Practical Methods of Optimization vol. 1: Unconstrained Optimization*, John Wiley & Sons (New York).

[10] R. Fletcher and C. Reeves, *Function minimization by conjugate gradients*, Comput. J. **7** (1964), pp. 149-154.

[11] J. C. Gilbert and J. Nocedal, *Global convergence properties of conjugate gradient methods for optimization*, SIAM J. Optimization. Vol. **2** No. 1 (1992), pp. 21-42.

[12] A. A. Goldstein, *On steepest descent*, SIAM J. on Control A, Vol. 3, No. 1, (1965), pp. 147-151.

[13] L. Grippo, F. Lampariello, and S. Lucidi, *Global convergence and stabilization of unconstrained minimization methods without derivatives*, Journal of Optimization Theory and Applications 56 (1988) pp. 385-406

[14] L. Grippo and S. Lucidi, *A globally convergent version of the Polak-Ribière conjugate gradient method,* Math. Prog. 78 (1997), pp. 375-391.

[15] M. R. Hestenes and E. L. Stiefel, *Methods of conjugate gradients for solving linear systems*, J. Res. Nat. Bur. Standards Sect. **5**, 49 (1952), 409-436.

[16] J. C. Lagarias, J. A. Reeds, M. H. Wright and P. E. Wright, *Convergence Properties of the Nelder-Mead Simplex Algorithm in Low Dimensions*, Technical Report 96-4-07, Computing Sciences Research Center, Bell Laboratories, 1996.

[17] R. De Leone, M. Gaudioso, and L. Grippo, *Stopping criteria for linesearch methods without derivatives*, Mathematical Programming 30 (1984) 285-300.

[18] G. H. Liu, J. Y. Han and H. X. Yin, *Global convergence of the Fletcher-Reeves algorithm with an inexact line search*, Report, Institute of Applied Mathematics, Chinese Academy of Sciences, 1993.

[19] J. J. Morè, B. S. Garbow and K. E. Hillstrom, *Testing unconstrained optimization software*, ACM Transactions on Mathematical Software **7** (1981), pp. 17-41.

[20] J. L. Nazareth, *A view of conjugate gradient-related algorithms for nonlinear optimization*, in: Linear and Nonlinear Conjugate Gradient-Related Methods, L. Adams and J. L. Nazareth, eds., SIAM, Philadelphia, 1996, pp. 149-163.

[21] J. Nocedal, *Theory of Algorithm for Unconstrained Optimization*, Acta Numerica (1991), pp. 199-242.

[22] M. J. D. Powell, *Nonconvex minimization calculations and the conjugate gradient method*, in: Lecture Notes in Mathematics vol. **1066**, Springer-Verlag (Berlin) (1984), pp. 122-141.

[23] E. Polak and G. Ribière, *Note sur la convergence de directions conjugées*, Rev. Francaise Informat Recherche Opertionelle, 3e Année **16** (1969), pp. 35-43.

[24] B. T. Polyak, *The conjugate gradient method in extremem problems*, USSR Comp. Math. and Math. Phys. **9** (1969), pp. 94-112.

[25] J. Werner, *Über die globale konvergenz von Variable-Metric Verfahren mith nichtexakter Schrittweitenbestimmung*, Numer. Math. 31 (1978), pp. 321-334.

[26] P. Wolfe, *Convergence conditions for ascent methods*, SIAM Review **11** (1969), pp. 226-235.

[27] G. Zoutendijk, *Nonlinear Programming, Computational Methods*, in: Integer and Nonlinear Programming (J. Abadie, ed.), North-Holland (Amsterdam), 1970, pp. 37-86.