75 - 7

Mathematical Programming 14 (1978) 161–169.
North-Holland Publishing Company

A CLASS OF METHODS FOR LINEAR PROGRAMMING

Markku KALLIO

Helsinki School of Economics, Helsinki, Finland

Evan L. PORTEUS

Stanford University, Stanford, CA, U.S.A.

Received 25 June 1975
Revised manuscript received 4 July 1977

A class of methods is presented for solving standard linear programming problems. Like the simplex method, these methods move from one feasible solution to another at each iteration, improving the objective function as they go. Each such feasible solution is also associated with a basis. However, this feasible solution need not be an extreme point and the basic solution corresponding to the associated basis need not be feasible. Nevertheless, an optimal solution, if one exists, is found in a finite number of iterations (under nondegeneracy). An important example of a method in the class is the reduced gradient method with a slight modification regarding selection of the entering variable.

Key words: Linear Programming Methods, Simplex Method, Reduced Gradient Method, Feasible Direction.

1. Introduction

Consider the linear program (LP):

$$\begin{aligned} &\text{find} && x \in \mathbf{R}^n \text{ to} \\ &\text{maximize} && cx, && \text{(LP1)} \\ &\text{subject to} && Ax = b, && \text{(LP2)} \\ &&& x \geq 0, && \text{(LP3)} \end{aligned}$$

where $c \in \mathbf{R}^n$, $A \in \mathbf{R}^{m \times n}$, and $b \in \mathbf{R}^m$. We assume that A has a full row rank. A *system solution* satisfies (LP2), a *homogeneous solution* satisfies $Ax = 0$, and a *feasible solution* satisfies (LP2) and (LP3). If x is (a) feasible (solution) and z is a homogeneous solution, then $x + \theta z$ is feasible as long as it is non-negative, for $\theta \in \mathbf{R}$. As θ increases, the objective function increases if and only if $cz > 0$. The simplex method [1] chooses as z one of the vectors corresponding to (changing the value of) a nonbasic variable, such that $cz > 0$. Our methods may choose as z a linear combination of such vectors, rather than just one. In particular, we may choose the direction given by the reduced gradient method (e.g., [4]). Both the simplex method and our methods select as a new feasible solution that one found by increasing θ (and the objective function) as much as possible (before violating the non-negativity constraint). We show that if the linear program has an optimal solution and is nondegenerate, then an optimal solution is found in a finite number of iterations by any of our methods.

2. Preliminaries

For convenience, we shall not distinguish between a variable, its index (subscript), and its value. For example, we say that β denotes the set of basic variables rather than the set of their indices, and a basic variable increases when its value increases.

Let $\alpha = \alpha(x) := \{i \mid x_i = 0\}$, so that α indexes the zero valued variables. In the simplex method, all the nonbasic variables would be in α, but such is not necessarily the case with our methods. Let β denote a basis index: a set of indices of m linearly independent columns of A. Assume, without loss, that variables are renumbered so that $\beta = \{1, 2, \ldots, m\}$, for convenience. Let $\sim\alpha$ and $\sim\beta$ denote the complements of α and β, respectively. Let a^j denote the jth column of A, for each j $(=1, 2, \ldots, n)$ and let B denote the basis matrix, consisting of the columns indexed by β. The vector $z^j \in \mathbf{R}^n$, which corresponds to the nonbasic variable $j \in \sim\beta$ is given componentwise by

$$z_i^j := \begin{cases} -(B^{-1}a^j)_i & \text{if } i \in \beta, \\ 1 & \text{if } i \in \sim\beta \text{ and } i = j, \\ 0 & \text{if } i \in \sim\beta \text{ and } i \neq j. \end{cases} \tag{1}$$

Because $Az^j = B(-B^{-1}a^j) + a^j = 0$, each z^j is a homogeneous solution. As mentioned before, the vectors z^j serve as directions of change for the simplex method. For the methods considered here, linear combinations of the vectors z^j will serve as such directions. For notational convenience, let $Z := (z^j) \in \mathbf{R}^{n\times(n-m)}$ denote the matrix of the vectors z^j and $w = (w_j) \in \mathbf{R}^{n-m}$ denote the vector of weights in such a linear combination, so that the direction of change is given by Zw. We shall index the components of w by the nonbasic variables (rather than the first $n-m$ integers). Thus $w = (w_j)$ for $j \in \sim\beta$. In particular, reference to w_j always carries the convention that $j \in \sim\beta$. For instance, when writing $w_j \geq 0$ for each $j \in \alpha$ below, we simply use α rather than the more cumbersome $\alpha \cap (\sim\beta)$. Taking eq. (1) into account, the component w_j indicates the rate of change for nonbasic variable j dictated by the direction Zw. Thus, where Zw is a direction in the space $\mathbf{R}^n$ of all variables, w is a direction in the space $\mathbf{R}^{n-m}$ of the nonbasic variables.

Lemma 1. *If x is a system solution and β is a basis index, then, for each system solution $\bar{x}$, there exists a (unique) $w \in \mathbf{R}^{n-m}$ such that*

$$\bar{x} = x + Zw. \tag{2}$$

If x and $\bar{x}$ are feasible solutions and $\alpha = \alpha(x)$, then $w_j \geq 0$ for each $j \in \alpha$.

Remark. Given a set of basic variables and a system solution x, any other system solution can be achieved by starting at x and adjusting the nonbasic variables. Among feasible solutions, the nonbasic variables equal to zero at x cannot be decreased while making this adjustment.

Proof. Let $S \subseteq \mathbf{R}^n$ be the subspace of homogeneous solutions. Because $\text{Rank}(A) = m$, it follows that $\text{Dim}(S) = n - m$. Clearly, the $n - m$ vectors z^j (for $j \in \sim\beta$) are linearly independent, and $z^j \in S$ for each j. Therefore, $\{z^j\}$ is a basis for S, and any $z \in S$ can be written as

$$z = Zw, \tag{3}$$

where $w \in \mathbf{R}^{n-m}$. Since $\bar{x}$ and x are system solutions, we have $A(\bar{x} - x) = 0$ and, hence, $\bar{x} - x \in S$. Eq. (2) then follows from eq. (3). The proof is completed by observing that if $w_j < 0$ for $j \in \alpha$, then $\bar{x}_j < 0$ whence $\bar{x}$ is not feasible.

Lemma 2. *If x is feasible, β is a basis index,*

$$cz^j \leq 0 \quad \textit{for all } j \in \alpha, \tag{4}$$

and

$$cz^j = 0 \quad \textit{for all } j \in \sim\alpha, \tag{5}$$

then x is optimal.

Remark. If increasing a zero valued nonbasic variable will not increase the objective function, and changing any other nonbasic variable will not change the objective function, then the current solution is optimal.

Proof. Suppose $\bar{x}$ is feasible. Then, by Lemma 1, we have the representation (2) with $w_j \geq 0$ for $j \in \alpha$. Premultiplying by c and applying eqs. (4) and (5) yields $c\bar{x} = cx + \sum_j cz^j w_j \leq cx$.

3. The admissible directions

An iteration starts with a feasible solution x and a basis index β and moves to another feasible solution $\bar{x}$ and basis index $\bar{\beta}$. We shall have $\bar{x} = x + \bar{\theta}Zw$ for some $\bar{\theta} \geq 0$ and some direction w. A particular method specifies a finite set $W(\alpha, \beta)$ which consists of the admissible directions w for any iteration where α indexes the zero valued variables and β is the basis index.

In the simplex method, $W(\alpha, \beta)$ consists of the unit vectors e^j that satisfy $cZe^j > 0$, whereas for the reduced gradient method, $W(\alpha, \beta)$ consists of a single element w given by

$$w_j = \begin{cases} -cz^j & \text{if } j \in \sim\alpha \text{ and } cz^j \leq 0, \\ 0 & \text{if } j \in \alpha \text{ and } cz^j \leq 0, \\ cz^j & \text{otherwise.} \end{cases} \tag{6}$$

That is, nonbasic variables are adjusted in proportion to their "reduced cost" unless they are currently zero and increasing them will not increase the objective function.

Our approach will encompass, but not be limited to, these two examples. Indeed, there are several ways of specifying $W(\alpha, \beta)$, the set of admissible directions. Each specification produces a method: we are therefore treating a class of methods. We seek to show that any such method will produce an optimal solution in a finite number of iterations, under non-degeneracy assumptions. To do this, we require that the finite set $W(\alpha, \beta)$ satisfy the following conditions:

1*. If $w \in W(\alpha, \beta)$, then
 (i) $w_j \geq 0$ for all $j \in \alpha$,
 (ii) $cZw > 0$,
 (iii) $cz^j w_j > 0$ for all j such that $w_j \neq 0$.

2*. If $W(\alpha, \beta)$ is empty, then no $w \in \mathbf{R}^{n-m}$ satisfies 1*(i)–(iii).

1* states conditions on the direction vector for the nonbasic variables. If 1*(i) is not satisfied, then, by Lemma 1, a positive movement in that direction results in an infeasible solution. Thus we think of 1*(i) as a feasibility condition for the direction. This guarantees that movement in an admissible direction will be limited by either

(1) a previously positive nonbasic variable decreasing to zero or

(2) a basic variable decreasing to zero.

1*(ii) simply requires that movement in an admissible direction must increase the objective function. With only 1*(i) and 1*(ii), one can easily construct examples where "zig-zagging" (and non-finite convergence) occurs. Condition 1*(iii) is to prevent "zig-zagging". It requires that movement due to any of the nonbasic variables being adjusted, taken individually, must increase the objective function. Using 1*(iii), we need only require $w \neq 0$ to get 1*(ii).

2* is the optimality condition, shown by the following result:

Lemma 3. *If x is feasible, $\alpha = \alpha(x)$, and β is a basis index such that $W(\alpha, \beta)$ is empty, then x is optimal.*

Proof. If $W(\alpha, \beta)$ is empty, then $cz^j > 0$, for any j, contradicts 2*. The same is true if $j \in \sim\alpha$ and $cz^j < 0$. Thus, $cz^j \leq 0$ for all j and $cz^j = 0$ for $j \in \sim\alpha$; that is, eqs. (4) and (5) hold, and, by Lemma 2, the current solution is optimal.

4. The class of methods

Having discussed the admissible directions for an iteration, we are ready to see how the basis is updated.

If a nonbasic variable decreases to zero, then, to save on computations, we simply leave the basis index unchanged. If not, then at least one basic variable has decreased to zero. We arbitrarily select one of these to be the leaving variable l. For the entering variable, there may be many candidates: any variable e is a candidate if it is currently positive ($\bar{x}_e > 0$) and $\bar{\beta} = \beta \cup \{e\} - \{l\}$ is a legitimate basis index. We have not been able to achieve our results by allowing the selection to be carried out arbitrarily. Rather, we require that the nonbasic variables be ordered and that the first candidate in the list be selected as the entering variable. The leaving variable is then placed at the end of the list, to update the ordering of the nonbasic variables. We shall clarify the role of the ordering later. For now, we simply summarize the steps of a method.

Given a specification of $W(\alpha, \beta)$ for all α and β, the form of the resulting method is as follows:

1°. Initialization: specify β, a feasible solution x, and $\alpha = \alpha(x)$.

2°. Specify direction: select an element w from $W(\alpha, \beta)$. If $W(\alpha, \beta)$ is empty, then stop (x is optimal by Lemma 3).

3°. Find new feasible solution: determine $\bar{\theta} = \max\{\theta \mid x + \theta Z w \geq 0\}$ and replace x by $x + \bar{\theta} Z w$. If $\bar{\theta} = \infty$, then stop ((LP) is unbounded).

4°. Update α and β.

(i) if there exists $j \in \sim\alpha$ such that $x_j = 0$, update α and return to 2° (without a basis change);

(ii) otherwise, update α and pick any $l \in \alpha \cap \beta$ as the leaving variable. Pick the first element $e \in \sim\beta$ such that $e \in \sim\alpha$ and $\bar{\beta} := \beta \cup \{e\} - \{l\}$ is a basis index. Replace β by $\bar{\beta}$, place l at the end of $\sim\beta$, and return to 2°.

Note that the initial feasible solution x need not be an extreme point of the feasible region, and the basic solution associated with the initial specification of β need not be feasible. Indeed, x and β can be chosen independently.

5. Proof of finite convergence in the nondegenerate case

We assume that (LP) is nondegenerate: for any basis B, all elements of $B^{-1}b$ are nonzero. We shall show (in Lemma 4 below) that the basis change in 4°(ii) can always be carried out. Then, it will be clear that all steps can be carried out. The process stops only if an unbounded solution is detected (in 3°) or the optimality condition (in 2°) is satisfied, and, if (LP) has a bounded optimal solution, then one is found in a finite number of iterations (Theorem 8).

Lemma 4. *Suppose* (LP) *is nondegenerate, x is any system solution, β is a basis index, and $\alpha = \alpha(x)$. For any $l \in \alpha \cap \beta$ there exists $e \in (\sim\alpha) \cap (\sim\beta)$ such that $\bar{\beta} = \beta \cup \{e\} - \{l\}$ is a legitimate new basis index; i.e., the columns $\{a^j \mid j \in \bar{\beta}\}$ are linearly independent.*

Remark. Note that x need not be the feasible solution found at a given iteration. Indeed, this result will be used for the case when x is an optimal solution.

Proof. Let $\delta := (\sim\alpha) \cap (\sim\beta)$. Since $l \in \alpha$, we have $x_l = 0$. Thus Lemma 1 yields

$$x_l = (B^{-1}b)_l + \sum_{j \in \delta} z_l^j x_j = 0.$$

By nondegeneracy, $(B^{-1}b)_l \neq 0$. Therefore $\delta \neq \emptyset$, and there exists $e \in \delta$ such that $z_l^e \neq 0$. By definition, $z_l^e = -(B^{-1}a^e)_l$, and the set $\{B^{-1}a^i \mid i \in \beta - \{l\}\}$ consists of all unit vectors in $\mathbf{R}^m$ except the lth. Therefore, the vectors $\{B^{-1}a^i\}$, $i \in \beta \cup \{e\} - \{l\}$ are linearly independent and so are the vectors $\{a^i\}$, $i \in \beta \cup \{e\} - \{l\}$, because B is nonsingular.

Lemma 5. *If* (LP) *is nondegenerate and x is a system solution, then at most $n - m$ of the components of x are equal to zero.*

Proof. Suppose the contrary. Let δ be a set having less than m elements such that $\sum_{i \in \delta} x_i a^i = b$. Let δ' be a largest set of linearly independent columns indexed by δ. Then, because we can express the linearly dependent columns as linear combinations of the independent columns, there exists $x_i' \in R$, for each $i \in \delta'$, such that $\sum_{i \in \delta'} x_i' a^i = b$. Because $\operatorname{Rank}(A) = m$, there exists a nonempty set $\rho \subseteq \sim\delta'$ such that $\delta' \cup \rho$ is a basis index. Let B denote the corresponding basis. Then, because $B^{-1}b$ is unique, $(B^{-1}b)_i = x_i'$ for $i \in \delta'$ and $(B^{-1}b)_i = 0$ for $i \in \rho$. This contradicts the nondegeneracy assumption.

Lemma 6. *If* (LP) *is nondegenerate, $\{x^k\}$ is a sequence of feasible solutions for* (LP) *which converges to $\bar{x}$, and $\bar{\alpha} = \alpha(\bar{x})$, then there exists N such that, for $k \geq N$, $x_j^k = 0$ implies $j \in \bar{\alpha}$, for all j.*

Proof. Here, $\bar{x}$ is feasible for (LP) since each x^k is feasible and the set of feasible solutions is closed. By Lemma 5, some of the components of $\bar{x}$ must be nonzero, so $\sim\bar{\alpha}$ is nonempty. Thus, we can introduce

$$\epsilon := \min_{j \in \sim\bar{\alpha}} \bar{x}_j > 0.$$

Let $D := \{x \in \mathbf{R}^n \mid \|x - \bar{x}\| < \epsilon\}$, where any norm, such as any L_p norm, that satisfies $|x_j| \leq \|x\|$ for all j will suffice. Then, for any $x \in D$, we have

$$\epsilon > \|x - \bar{x}\| \geq |x_j - \bar{x}_j|$$

and therefore

$$x_j > \bar{x}_j - \epsilon \geq 0 \quad \text{for all } j \in \sim\bar{\alpha}. \tag{7}$$

Because $\{x^k\}$ converges to $\bar{x}$, there exists N such that $k \geq N$ implies $x^k \in D$. Then, by eq. (7), $x_j^k > 0$ for all $j \in \sim\bar{\alpha}$; that is, $x_j^k = 0$ implies that $j \in \bar{\alpha}$.

Lemma 7. *If* (LP) *is nondegenerate, then, in* $n - m$ *iterations, either a basis change occurs or the algorithm stops.*

Proof. A basis change does not occur if a nonbasic variable, say x_j, decreases to zero. If this is the case, then, by 1*(iii), $cz^j < 0$. Thus, by 1*(i) and 1*(iii), $w_j = 0$ during the following iterations—and consequently, the nonbasic variable x_j remains at zero—as long as no basis change occurs.

Thus, the algorithm cannot proceed beyond $n - m$ iterations without a basis change.

Theorem 8. *If* (LP) *is nondegenerate and it has a bounded optimal solution, then any method of the class solves* (LP) *in a finite number of iterations.*

Proof. Under our assumptions, the process stops only if an optimal solution is found. Suppose, to the contrary, that the process does not stop. Then a sequence $\{x^k, k = 0, 1, \ldots\}$ of feasible solutions is generated, where x^0 is the initial feasible solution. Let d^k and θ^k denote the direction and distance, respectively, for iteration k, for each $k (= 1, 2, \ldots)$. Thus $x^k = x^{k-1} + \theta^k d^k$ for each k, so that $\theta^k = \|x^k - x^{k-1}\| / \|d^k\|$. Hence,

$$\begin{aligned} x^k &= x^{k-1} + \|x^k - x^{k-1}\| d^k / \|d^k\| = \cdots \\ &= x^0 + \sum_{j=1}^{k} \|x^j - x^{j-1}\| d^j / \|d^j\|. \end{aligned} \tag{8}$$

Now let

$$\mu := \inf_{(\alpha,\beta)} \left[\min_{w \in W(\alpha,\beta)} cZw / \|Zw\| \right].$$

By 1*(ii), because each $W(\alpha, \beta)$ is a finite set, and since there are only a finite number of distinct (α, β) combinations, we conclude that μ must be positive. Because (LP) is bounded, there exists M such that $cx < M$ for all feasible x. Select arbitrary k. Then, by eq. (8) and the definition of μ, we get

$$\begin{aligned} M > cx^k &= cx^0 + \sum_{j=1}^{k} \|x^j - x^{j-1}\| cd^j / \|d^j\| \\ &\geq cx^0 + \sum_{j=1}^{k} \|x^j - x^{j-1}\| \mu. \end{aligned}$$

Since $\mu > 0$,

$$\sum_{j=1}^{k} \|x^j - x^{j-1}\| \leq (M - cx^0)/\mu \quad \text{for all } k,$$

and therefore $\{x^k\}$ is a Cauchy sequence. Let $\bar{x}$ denote its limit and $\bar{\alpha} = \alpha(\bar{x})$.

By Lemma 6, there is an N such that, for each iteration $k \geq N$, a variable l leaving the basis must be an element of $\bar{\alpha}$. If we can show that such a variable cannot re-enter the basis, then there can only be a finite number of basis changes beyond iteration N. Using Lemma 7 gives us the contradiction needed to prove the result.

Suppose variable l leaves the basis at iteration $k \geq N$. To show that l cannot re-enter the basis, we show that at each basis change beyond iteration k, there exists a nonbasic variable e in $\sim\bar{\alpha}$ such that

(i) variable e is currently positive,

(ii) $\beta \cup \{e\} - \{l\}$ is a basis index, and

(iii) e precedes l in the ordering. Conditions (i) and (ii) guarantee that e is an eligible candidate to enter the basis, and (iii) means that e would be chosen in preference to l. By Lemma 6, (i) holds for all elements of $\sim\bar{\alpha}$ after iteration N. Condition (ii) follows directly from Lemma 4, where now $\bar{x}$ plays the role of the system solution, the basis index for the current iteration that of β, the current leaving variable that of l, and $\bar{\alpha}$ that of α. Finally, (iii) follows from the fact that when l leaves the basis (immediately thereafter being the last variable having left the basis and thus, at the end of the list of nonbasic variables) all other nonbasic elements in $\sim\bar{\alpha}$ precede l and they remain so because elements of $\sim\bar{\alpha}$ cannot leave the basis after iteration N.

6. Discussion

We have required that each of our methods retain an ordering of the nonbasic variables. Our ordering puts leaving variables at the end of the list of nonbasic variables, so that a chronological ordering is established: preference for selecting the entering variable is given to those nonbasic variables that have been nonbasic for the greatest number of consecutive iterations. (The initial ordering can be arbitrary.) The ordering plays a simple role, and only in the proof of Theorem 8. There, after sufficiently many iterations, only eventually zero variables (those that will equal zero in the limit) can leave the basis. Since our ordering will put them at the end of the list, we can show that they will never re-enter the basis. This result then yields the desired finite convergence.

Actually another ordering could be used. We could give preference to nonbasic variables with the largest current values. Since after sufficiently many iterations, the eventually zero variables will have sufficiently small values, we could again guarantee that such variables would not re-enter the basis.

Thus, our ordering plays the role of a sufficient condition: it allows us to prove Theorem 8. However, we have not been able to show that the role played by the ordering is necessary. Nevertheless, we are not alone in addressing this role. For example, Huard [2] provides an alternative rule for handling the nonbasic variables that also guarantees finite convergence for the reduced gradient method in the nondegenerate case.

The constrained gradient method [3] follows the steps of our methods. However, for an admissible direction w our condition 1*(iii) is replaced by the requirement that the rate of improvement in the objective function in the direction Zw is maximal under the requirement 1*(i). A simple geometric examination translates this requirement for w to:

$$\begin{aligned} &\text{minimize} \quad (c - Zw)^{\mathrm{T}}(c - Zw), \\ &\text{subject to} \quad w_j \geq 0 \quad \text{for } j \in \alpha. \end{aligned}$$

If the solution of this (quadratic programming) problem does not satisfy 1*(ii), then (by an application of Lemma 3) the current solution is optimal. One can easily construct an example where the direction chosen by the constrained gradient method does not satisfy 1*(iii) and therefore the method does not belong to our class. However, Lemke [3] does provide a proof of finite convergence of this method, in the nondegenerate case.

Appendix. An example where the constrained gradient method does not satisfy 1*(iii)

Consider

$$\begin{aligned} &\text{maximize} \quad (1, 1, -3)x, \\ &\text{subject to} \quad (1, 2, 1)\, x = 4, \quad x \geq 0. \end{aligned}$$

Suppose the initial feasible solution $x^0 = (1, 1, 1)$ and the initial basis index $\beta = \{1\}$. Then we have:

$$\begin{aligned} z^2 &= (-2, 1, 0), \qquad cz^2 = -1, \\ z^3 &= (-1, 0, 1), \qquad cz^3 = -4. \end{aligned}$$

The direction chosen by the constrained gradient method is $w = (w_2, w_3) = (1, -3)$, so that $z = Zw = (1, 1, -3) = c$. Now $cz^2 w_2 = -1$ contradicts 1*(iii) (but $cz^2 w_3 = 12$ does not).

References

[1] G.B. Dantzig, *Linear programming and extensions* (Princeton University Press, Princeton, NJ, 1963).

[2] P. Huard, "Convergence of the reduced gradient method", presented in the Nonlinear Programming Symposium, Madison (1974).

[3] C.E. Lemke, "The constrained gradient method of linear programming", *Journal of the Society of Industrial Applied Mathematics* 9 (1961) 1–17.

[4] P. Wolfe, "Methods of nonlinear programming", in: J. Abadie, ed., *Nonlinear programming* (Wiley, New York, 1967) Ch. 6.