Onderzoeksrapport nr. : 7608

A MODIFIED REDUCED GRADIENT METHOD
FOR DUAL POSYNOMIAL PROGRAMMING¹

by

J.G. ECKER²

W. GOCHET³

Y. SMEERS⁴

1. This research was supported in part by the National Science Foundation, Grant No. MCS75-09443 A01.
2. Department of Mathematical Sciences, Rensselaer Polytechnic Institute, Troy, New York 12181
3. Department of Applied Economics, K.U.L. and C.O.R.E., Catholic University of Louvain, Louvain, Belgium.
4. Department of Engineering, U.C.L. and C.O.R.E., Catholic University of Louvain, Louvain, Belgium.

Wettelijk Depot : D/1976/2376/17.

1. INTRODUCTION.

In Ref. 1, Beck and Ecker present an algorithm for the solution of the dual of a posynomial programming problem. The algorithm is in essence Zangwill's convex simplex method, see Ref. 2, with some important modifications required because of two basic problems :

- (i) the nondifferentiability of the objective function at a point where some of the variables are equal to zero, and
- (ii) the fact that each "block of dual variables" must have all components positive or all components equal to zero at optimality.

Both these problems will be more fully discussed in section 2.

Essentially the modifications in Ref. 1 are based on the following : whereas in the convex simplex method only one independent (nonbasic) variable is allowed to change at every iteration, Beck and Ecker show how a complete block of variables can be allowed to move to or away from zero at certain iterations. Computational experience has shown, as was expected, that the iterations involving block changes result in significantly greater increases in the objective function value than the (standard convec simplex) iterations involving a change in a single nonbasic variable.

In this paper, we show how Wolfe's reduced gradient method (Ref. 3) can be modified to overcome the difficulties (i) and (ii) above. The reduced gradient method allows for changes in all nonbasic variables at each iteration and we will show below how to modify the method so that block changes can be permitted at certain iterations. An important advantage of this approach (over using the convex simplex method as the underlying algorithm) is that on those iterations not involving block changes the reduced gradient method permits all the nonbasic variables to be changed simultaneously. In developing the required modifications, we use some of the basic ideas in Ref. 1. A main difference, however, lies in the criteria that are used for a block of variables to move to or away from zero. In particular, with respect to moving a block of variables away from zero, a

direction will be constructed that gives the largest increase in the objective function value, at least locally. The procedure of Beck and Ecker uses a direction of strict increase, in general different from the locally best direction. For a short but clear description of the reduced gradient method the reader is referred to Ref. 4.

2. THE DUAL OF A POSYNOMIAL GEOMETRIC PROGRAM.

Associated with any posynomial geometric program is the following dual program, Ref. 5, expressed here with the concave logarithmic form of the objective function :

$$\begin{aligned} \max \quad V(x) &= \sum_{k=0}^p \sum_{i \in [k]} x_i \log(c_i \Lambda_k / x_i) \\ \text{subject to} \quad &A^T x = 0 \\ &\sum_{i \in [0]} x_i = 1 \\ &x \geq 0 \end{aligned}$$

where Λ_k denotes $\sum_{i \in [k]} x_i$ and $x_i \log(c_i \Lambda_k / x_i)$ is defined to be zero when $x_i = 0$, and $[k] \equiv \{m_k, m_k+1, \dots, n_k\}$, $m_0 = 1$, $m_1 = n_0+1$, $\dots$, $n_p = m_{p-1}+1$, $n_p = n$. It is clear that the objective function is nondifferentiable at a point where $x_i = 0$ for some i . Furthermore it is a property of any finite optimal solution that for a given k , either $x_i = 0$, $\forall i \in [k]$ or $x_i > 0$, $\forall i \in [k]$, see Ref. 5. As in Ref. 1, we assume, without loss of generality, that this problem is canonical, Ref. 5, p. 169, so that there exists a strictly positive feasible vector which can be used as the starting point for the algorithm. Finding such a vector requires, at most, the solution of one linear program. The components G_{x_i} of the gradient of the objective function can be easily computed as

$$G_{x_i}(x) = \log(c_i \Lambda_k / x_i), \quad i \in [k], \quad k = 0, 1, \dots, p.$$

We denote the number of equalities in the constraint set by m and assume

that every collection of m columns from the constraint matrix is linearly independent.

3. THE MODIFIED REDUCED GRADIENT METHOD.

In this section we give a description of the reduced gradient method with the modifications that are necessary to overcome the difficulties mentioned above. The justification for these modifications will be given in section 4.

Assume that, at the beginning of an iteration, the m largest variables are the basic variables denoted by y and the remaining variables, the nonbasic ones denoted by z (ties are broken arbitrarily). Assume further that all basic variables are positive (nondegeneracy assumption) and that a nonbasic variable is zero if all the variables in the same block are nonbasic and equal to zero. Notice that at the beginning of the first iteration these assumptions are satisfied (since all variables are positive). Moreover, the algorithm is constructed in such a way that they remain satisfied.

Assume that the system of dual constraint equations is rewritten as

$$y + Tz = b$$

where y is the $(m \times 1)$ vector of basic variables

z is the $((n-m) \times 1)$ vector of nonbasic variables

T is a $(m \times (n-m))$ matrix

b is a $(m \times 1)$ vector of constraints.

Furthermore, let

$$K = \{0, 1, 2, \dots, p\}$$

$$K_0 = \{k \mid x_i \text{ is nonbasic and zero for every } i \in [k]\}$$

$$K_1 = \{k \mid x_i \text{ is nonbasic and positive for every } i \in [k]\}$$

$$M_i = \sum_j G_{y_j}(x) t_{ji}$$

One iteration of our modified reduced gradient algorithm can then be described as follows.

Step 1.

$$(a) \text{ Let } \Delta z_i = \frac{c_i e^{-M_i}}{\sum_{[k]} c_i e^{-M_i}} \quad \text{all } i \in [k] \text{ and } k \in \bar{K}_0$$

$$\text{where } \bar{K}_0 = \{k \in K_0 \mid \sum_{[k]} c_i e^{-M_i} > 1\}$$

$$(b) \text{ Let } \Delta z_i = 0, \quad \text{all } i \in [k] \text{ and } k \in K_0 \setminus \bar{K}_0$$

$$(c) \text{ Let } \Delta z_i = -z_i, \quad \text{all } i \in [k] \text{ and } k \in \bar{K}_1$$

$$\text{where } \bar{K}_1 = \{k \in K_1 \mid r_i \cong_{z_i} G_{z_i}(x) - \sum_j G_{y_j}(x) t_{ji} < 0, \text{ all } i \in [k]\}$$

$$(c') \text{ An alternative here is to set } \Delta z_i = -z_i, \quad i \in [k] \text{ and } k \in K_1$$

$$\text{whenever } -p_k z_{[k]}^T r_{[k]} \geq q_k r_{[k]}^T r_{[k]}$$

$$\text{with } q_k \text{ a fixed number, } 0 < q_k \leq 1,$$

$$\text{and the scalar } p_k = \frac{\|r_{[k]}\|}{\|z_{[k]}\|}$$

$$\text{where } r_{[k]} = (r_i), \quad i \in [k] \text{ and } z_{[k]} = (z_i), \quad i \in [k]$$

$$(d) \text{ Let } \Delta z_i = r_i = G_{z_i}(x) - \sum_j G_{y_j}(x) t_{ji} \quad \text{for } i \in [k], \quad k \in K \setminus \{K_0 \cup \bar{K}_1\}$$

Step 2. If $\Delta z_i = 0$, all $i \in [k]$ and $k \in K$, stop; the current point is optimal. Otherwise, compute

$$\Delta y = -T \Delta z$$

Step 3. Find $\alpha_1, \alpha_2, \alpha_3$ such that

$$(a) \quad \alpha_1 = \max \{\alpha \mid y + \alpha \Delta y \geq 0\}$$

$$(b) \alpha_2 = \max \{ \alpha \mid z + \alpha \Delta z \geq 0 \}$$

$$(c) \alpha_3 = \max \{ V'(x + \alpha \Delta x) : 0 \leq \alpha \leq \min(\alpha_1, \alpha_2) \}$$

$$(d) \text{ Let } \bar{x} = x + \alpha_3 \Delta x$$

Step 4. Pivot in the system $y + Tx = b$ so that the m largest components of $\bar{x}$ constitute the basic variables. Update the index sets K_0 and K_1 and return to step 1.

4. JUSTIFICATION AND THEORY.

Step 1(d) is the direction for the nonbasic variables as proposed by the reduced gradient method. Notice that all $z_i, i \in [k], k \in K \setminus \{K_0 \cup \bar{K}_1\}$ are strictly positive so that the reduced gradient components r_i are well defined. Moreover we need not worry about the nonnegativity condition on the z_i in the selection of the direction.

As will be explained below steps 2 and 3 guarantee that a variable will never become zero unless the whole block of variables is equal to zero. Hence step 1(a) and (b) define all direction vectors for blocks of variables that are zero at the current point. The justification for these steps is based on the following result.

Theorem

Assume that, at the current point, a block of nonbasic variables z_i for $i \in [k^*]$, is equal to zero. Then :

(i) Any direction $\Delta z_i, i \in [k^*]$, satisfying

$$\begin{cases} \Delta v_i < c_i e^{-M_i} \\ \sum_{i \in [k^*]} \Delta v_i = 1 \\ [k^*] \Delta v_i > 0 \end{cases} \text{ where } v_i = \frac{\Delta z_i}{\Delta z_{[k^*]}} \text{ with } \Delta z_{[k^*]} = \sum_{i \in [k^*]} \Delta z_i$$

is a direction of increase.

(ii) There exists at least one direction of increase iff

$$\sum_{[k^x]} c_i e^{-M_i} > 1.$$

The best direction of increase locally is given by

$$\Delta z_i = \frac{c_i e^{-M_i}}{\sum_{[k^x]} c_i e^{-M_i}}, \quad i \in [k^x]$$

Proof.

(i) Though the reduced gradients $r_i, i \in [k^x]$, are not defined when $z_i = 0, i \in [k^x]$, along any direction $\Delta z_i > 0, i \in [k^x]$, r_i is well defined and equal to

$$r_i = \log \frac{c_i \Delta z_i [k^x]}{\Delta z_i} - \sum_j G_j(x) t_{ji}, \quad i \in [k^x]$$

In fact, given the direction $\Delta z_i > 0, i \in [k^x]$, the gradients remain constant along the line $\lambda \Delta z_i, i \in [k^x]$ for $\lambda > 0$. Hence $\Delta z_i, i \in [k^x]$, is a direction of increase whenever

$$\Delta z_i > 0, \quad \forall i \in [k^x],$$

$$\text{and } r_i > 0, \quad \forall i \in [k^x].$$

These conditions are easily seen to be equivalent to those stated in (i) of the theorem.

(ii) We first note that $\Delta z_i, i \in [k^x]$, is a direction of increase iff the inner product between this direction and the reduced gradient at the current point is positive; that is, iff

$$\sum_{[k^x]} \Delta z_i \{ \log (c_i \Delta z_i [k^x] / \Delta z_i) - M_i \} > 0.$$

There exists a direction such that the above strict inequality holds iff the maximum inner product over all possible normalized directions is positive. To find the direction of best increase locally, we solve the problem:

$$\text{Max}_{\Delta z_{[k^x]}} \sum_{[k^x]} \Delta z_i \{ \log(c_i \Delta z_{[k^x]} / \Delta z_i) - M_i \}$$

$$\text{subject to } \sum_{[k^x]} \Delta z_i = 1$$

$$\Delta z_i > 0, \text{ for each } i \in [k^x].$$

Neglecting the positivity condition, the solution can be easily obtained as

$$\Delta z_i^x = \frac{c_i e^{-M_i}}{\sum_{[k^x]} c_i e^{-M_i}}, \quad i \in [k^x]$$

which is strictly positive and satisfies the conditions of (i). The maximum value computed by using this solution Δz_i^x equals $\log \sum_{[k^x]} c_i e^{-M_i}$.

In particular, this shows that there is a direction of increase iff $\sum_{[k^x]} c_i e^{-M_i} > 1$ and the proof is complete.

Step 1(a) of the algorithm uses the direction $\Delta z_i^x, i \in [k^x]$, of best local increase whenever $\sum_{[k^x]} c_i e^{-M_i} > 1$. This is a so called block increase away from zero since all variables in such a block will move away from zero. If there does not exist a direction of increase for blocks that are currently at zero, then step 1(b) of the algorithm keeps those variables at zero.

We now justify step 1(c) of the algorithm. In order to do that we shall prove that in step 3, a nonbasic variable will never go to zero whenever the direction specified by the reduced gradient (or any other direction for that matter) is used unless by chance that direction happens to be the direction that satisfies $r_i = -z_i$ for all i in the block in question. Since this is very unlikely and remembering that, at an optimal solution,

either all variables in a block or none are zero, step 1(c) tries to overcome this difficulty. Indeed, whenever certain conditions are satisfied, the direction in step 1(c) is the direction in which all the (nonbasic) variables in block k can move to zero at the same time. The condition states that, in order to select this direction, the reduced gradients for the variables in that block must be all negative. This means that the objective value increases when all the variables in that block are decreased in the direction of the reduced gradient. The direction specified in step 1(c) also decreases all the variables in that block but is different from the reduced gradient direction. However, it can easily be seen to be a direction of increase since

$$\sum_{[k]} (-z_i) r_i > 0 \text{ whenever } k \in \bar{K}_1.$$

This procedure is suggested in the modified convex simplex method of Beck and Ecker, Ref. 1.

The alternative step 1(c') has the same objective. It states that the direction to zero for a block of variables will be selected whenever the inner product between that direction and the reduced gradient (being a measure of the increase in the objective function locally when the direction $-z_i$ is selected) is at least a fraction of the norm of the reduced gradient (being a measure of the increase in the objective function locally when the direction r_i is selected). The scalar p_k is a normalization constant to make a meaningful comparison possible. The extreme case $q_k=1$ corresponds to always selecting the reduced gradient as the direction in step 1(c) while $q_k=0$ corresponds to selecting $\Delta z_i = -z_i$ provided it is a direction of increase. The rest of the algorithm follows exactly the reduced gradient method.

It remains to be shown that no variable, either basic or nonbasic, will go to zero without the whole block, to which it belongs, going to zero. To show that we need the assumption that all variables are bounded.

The only place in the algorithm where a variable could conceivably go to zero without the corresponding block of variables going to zero, is in step 3.c, where the maximum is searched of the objective function in the direction Δx .

For $0 \leq \alpha < \alpha_3$, it holds that, for a given direction Δx ,

$$\frac{dV}{d\alpha} = \sum_{i \in N \setminus (K_0 \setminus \bar{K}_0)} G_{z_i}(x + \alpha \Delta x) \Delta z_i + \sum_j G_{y_j}(x + \alpha \Delta x) \Delta y_j$$

where N is the set of nonbasic variables. $\frac{dV}{d\alpha}$ is a continuous function of α for $0 \leq \alpha < \alpha_3$ and a given direction Δx provided we set

$$G_{z_i}(x + \alpha \Delta x) \Big|_{\alpha=0} = \log [c_i \Delta z_{[k]} / \Delta z_i] \quad \text{whenever } z_i = 0, i \in [k]$$

which is the (constant) gradient along the given direction Δz_i , $i \in [k]$.

Since Δx is a direction of increase, it holds that $\frac{dV}{d\alpha} \Big|_{\alpha=0} > 0$.

Suppose now that, as $\alpha \rightarrow \alpha_3$, there exists proper subsets S_k of $[k]$, $k \in K' \subseteq K$ such that

$$\begin{aligned} z_i &\rightarrow 0, & i \in S_k \\ y_j &\rightarrow 0, & j \in S_k \end{aligned}$$

It is possible, of course, that S_k contains either only basic variables y_j or only nonbasic variables z_i . It is immediate from the expressions of $G_{z_i}(\cdot)$ and $G_{y_j}(\cdot)$ that

$$G_{z_i}(x + \alpha \Delta x) \rightarrow +\infty \quad \text{as } \alpha \rightarrow \alpha_3 \quad \text{for } i \in S_k, k \in K' \subseteq K$$

$$G_{y_j}(x + \alpha \Delta x) \rightarrow +\infty \quad \text{as } \alpha \rightarrow \alpha_3 \quad \text{for } j \in S_k, k \in K' \subseteq K.$$

Since for variables x_i going to zero it must be the case that $\Delta x_i < 0$, it follows that

$$\sum_{k \in K'} \sum_{S_k} G_{z_i} (x + \alpha \Delta x) \Delta z_i + \sum_{k \in K'} \sum_{S_k} G_{y_j} (x + \alpha \Delta x) \Delta y_j \rightarrow -\infty \text{ as } \alpha \rightarrow \alpha_3.$$

All other terms in the expression for $\frac{dV}{d\alpha}$ are bounded because of the assumption that all variables are bounded. Hence

$$\frac{dV}{d\alpha} \rightarrow -\infty \text{ as } \alpha \rightarrow \alpha_3.$$

This contradicts the fact that the maximum of V in the direction Δx occurs at $\alpha = \alpha_3$. Hence no subsets S_k as defined above can exist.

5. CONCLUSION.

Modifications to the reduced gradient method have been developed that allow for the solution of the dual of a polynomial programming problem. In order to overcome the difficulties associated with the nondifferentiability of the objective function, we have shown how, on certain iterations of the reduced gradient method, blocks of variables can be allowed to move to or away from zero. On the iterations not involving block changes, standard reduced gradient iterations are performed. We expect the method we have proposed here to be computationally superior to the modified concave simplex method, see Ref. 1.

REFERENCES

- 1 BECK, P.A., and ECKER, J.G., "A Modified Concave Simplex Algorithm for Geometric Programming", Journal of Optimization Theory and Applications, Vol. 15, pp. 189-202, 1975.
- 2 ZANGWILL, W.I., "The Convex Simplex Method", Management Science, Vol. 14, No. 3, 1967.
- 3 WOLFE, P., "Methods of Nonlinear Programming", pp. 97-131, Nonlinear Programming edited by J. Abadie, Interscience, John Wiley and Sons, New York, N.Y., 1967.
- 4 LUENBERGER, David G., Introduction to Linear and Nonlinear Programming, Addison-Wesley, Reading, Massachusetts, 1973.
- 5 DUFFIN, R.J., PETERSON, E.L., and ZENER C., Geometric Programming-Theory and Applications, John Wiley and Sons, New York, N.Y., 1967.