CALCULUS IN NON-LINEAR PROGRAMMING PROBLEM

By

KANTI SWARUP and R. K. GUPTA

In the recent paper [1] Wilde has derived with the help of calculus, the necessary conditions for a relative minimum of non-linear differentiable objective function of non-negative values constrained by non-linear differentiable inequalities.

It is possible to establish the sufficiency of these conditions *by using only well known theorems of differential calculus* provided we assume further that the objective function is such that it can be transformed into a strictly convex function under given constraints. (Assuming sufficient differentiability, an anisotropically convex function may be transformed into a convex function by application of the theory of invariance of quadratic forms [4].) The results are expressed entirely in terms of partial derivatives.

According to Wilde a non-linear programming problem is to find a set of non-negative variables, $x_j\,(j=1, 2,\cdots J)$ that satisfy a set of differentiable constraints:

$$f_i\,(x_j)\geqslant 0 \qquad (i=1, 2, \cdots I) \tag{1}$$

and that also minimise a differentiable non-linear objective function $g(x_i)$.

Let $f_i\,(x_j)$ represent I functions mapping J real variables x_j into real variables f_i. This transformation may be expressed by the following I equations in the $I+J$ variables f_i and x_j.

$$f_i-f_i\,(x_j)=0 \quad \left.\begin{matrix} i=1, \cdots I \\ j=1, \cdots J \end{matrix}\right\} \tag{2}$$

Let these equations have real solution (f_i^o, x_j^o) such that

$$f_i^o-f_i^o\,(x_j)=0 \tag{3}$$

The differentiability of $f_i\,(x_j)$ guarantees that the I functions $f_i=f_i\,(x_j)$ regarded as functions of all $I+J$ independent variables are differentiable in the neighbourhood of the point (f_i^o, x_j^o).

From among these $I+J$ independent variables, let us select any I say $x_l\,(l=1, \cdots, p)$ and $f_q\,(q=p+1, \cdots, I)$ where $1 \leqslant p \leqslant I$

If Jacobian

$$|J\,(f_i^o, x_j^o)| \neq 0 \tag{4}$$

then the set of equations (2) can be solved for I variables x_l and f_q in terms of remaining J variables which to be definite, we shall call $F_r\,(r=1,2,\cdots p)$ and $X_s\,(s=p+1,\cdots J)$

Now we consider differentiable function $g(x_j)$ as in [1]

$$g(x_j)=G(F_r, X_s) \tag{5}$$

at any point (F_r, X_s) in the vicinity of (F_r^o, X_s^o) such that F_r, f_q, x_l and X_s satisfy equation (2).

Let the value of G at the point (F_r^o, X_s^o) be designated by G^o. Using the approximation theorem about (F_r^o, X_s^o), we have

$$\begin{aligned} g-G^o= & \sum_{r=1}^{p}\left(\frac{\partial G}{\partial F_r}\right)^o(F_r-F_r^o)+\sum_{s=p+1}^{J}\left(\frac{\partial G}{\partial X_s}\right)^o(X_s-X_s^o) \\ & +\frac{1}{\underline{|2}}\Bigg[\sum_{r=1}^{p}\left(\frac{\partial^2 G}{\partial F_r^2}\right)^o(F_r-F_r^o)^2+\sum_{\substack{l,k=1\\ l\neq k}}^{p}\left(\frac{\partial^2 G}{\partial F_l\,\partial F_k}\right)^o(F_l-F_l^o)(F_k-F_k^o) \\ & +\sum_{r=1}^{p}\sum_{s=p+1}^{J}\left(\frac{\partial^2 G}{\partial F_r\,\partial X_s}\right)^o(F_r-F_r^o)(X_s-X_s^o)+\sum_{s=p+1}^{J}\left(\frac{\partial^2 G}{\partial X_s^2}\right)^o(X_s-X_s^o)^2 \\ & +\sum_{\substack{i,j=p+1\\ i\neq j}}^{J}\left(\frac{\partial^2 G}{\partial X_i\,\partial X_j}\right)^o(X_i-X_i^o)(X_j-X_j^o)\Bigg]+O\{(F_r-F_r^o)^3, (X_s-X_s^o)^3\} \end{aligned} \tag{6}$$

The symbol $\left(\frac{\partial G}{\partial F_k}\right)^o$ for some index $k\,(1\leqslant k\leqslant p)$ represents partial derivative of G with respect to F_k at (F_r^o, X_s^o).

Now we have assumed that equations (2) are such that the transformed objective function $G(F_r, X_s)$ is strictly convex differentiable function over the convex set "S" of feasible solutions x_j.

Therefore

$$\begin{aligned} L=\frac{1}{\underline{|2}}\Bigg[& \sum_{r=1}^{p}\left(\frac{\partial^2 G}{\partial F_r^2}\right)^o(F_r-F_r^o)^2+\sum_{r=1}^{p}\sum_{s=p+1}^{J}\left(\frac{\partial^2 G}{\partial F_r\,\partial X_s}\right)^o(F_r-F_r^o)(X_s-X_s^o) \\ & +\sum_{\substack{l,k=1\\ l\neq k}}^{p}\left(\frac{\partial^2 G}{\partial F_l\,\partial F_k}\right)^o(F_l-F_l^o)(F_k-F_k^o)+\sum_{s=p+1}^{J}\left(\frac{\partial^2 G}{\partial X_s^2}\right)^o(X_s-X_s^o)^2 \\ & +\sum_{\substack{i,j=p=1\\ i\neq j}}^{J}\left(\frac{\partial^2 G}{\partial X_i\,\partial X_j}\right)^o(X_i-X_i^o)(X_j-X_j^o)\Bigg] \end{aligned} \tag{7}$$

(@ If (f_i^o, x_j^o) is a basic feasible solution)

@ For definition of basic feasible, non-degenerate and relative minimum etc., reference may be had to Wilde's paper.

$$L=\frac{1}{\underline{|2}}\Big[\sum_{r=1}^{p}\Big(\frac{\partial^2 G}{\partial F_r}\Big)^o F_r^2+\sum_{r=1}^{p}\sum_{s=p+1}^{J}\Big(\frac{\partial^2 G}{\partial F_r\,\partial X_s}\Big)^o F_r\,X_s+\sum_{\substack{l,\,k=1\\ l\neq k}}^{p}\Big(\frac{\partial^2 G}{\partial F_l\,\partial F_k}\Big)^o F_l\,F_k$$

$$+\sum_{s=p+1}^{J}\Big(\frac{\partial^2 G}{\partial X_s^2}\Big)^o X_s^2+\sum_{\substack{i,\,j=p+1\\ i\neq j}}^{J}\Big(\frac{\partial^2 G}{\partial X_i\,\partial X_j}\Big)^o X_i\,X_j\Big] \tag{8}$$

is a positive definite form, {[2] Appendix}

i. e. $L > 0,$

for all values of F_r, X_s except when

$$F_r=0$$
$$X_s=0$$

then $L=0.$ (9)

Theorem :

For a basic non-degenerate feasible solution to give a relative minimum, it is necessary and sufficient that

$$\left.\begin{aligned}&\Big(\frac{\partial G}{\partial F_r}\Big)^o \geqslant 0 \qquad (r=1,\cdots p)\\ &\Big(\frac{\partial G}{\partial X_s}\Big)^o \geqslant 0 \qquad (s=p+1,\cdots J)\end{aligned}\right\} \tag{10}$$

The non-degeneracy of the solution implies that it will remain feasible for sufficiently small changes in F_r and X_s.

Necessary condition has been proved by Wilde in Lemma {1} of his paper.

To establish the sufficiency of these conditions we have to show that G_o is a relative-minimum,
when

$$\Big(\frac{\partial G}{\partial F_r}\Big)^o \geqslant 0 \qquad (r=1,\cdots p)$$

$$\Big(\frac{\partial G}{\partial X_s}\Big)^o \geqslant 0 \qquad (s=p+1,\cdots J)$$

Proof: The proof is divided into two parts.

1st Part : At least one of the first order partial derivatives is not zero.

2nd Part : All the first order partial derivatives are zero.

FIRST PART

Let $\Big(\frac{\partial G}{\partial F_k}\Big)^o > 0$ $(1 \leq k \leq p)$ while all other first order derivatives may be zero.

From (6)

$$g-G^o=\left(\frac{\partial G}{\partial F_k}\right)^o F_k+O\left(F_r^2, X_s^2\right) \tag{11}$$

A point (F_r, X_s) considered can be taken in such a small neighbourhood of $(0, 0)$ i. e. (f_i^o, x_j^o) so that $O(F_r^2, X_s^2)$ can be neglected in comparison to $F_k \neq 0$

therefore $\quad g-G^o > 0$

or $\quad g > G^o$

In case $F_k=0$ then $O(F_r^2, X_s^2)$ can no longer be neglected and then

$$g-G^o=L+O\left(F_r^3, X_s^3\right) \tag{12}$$

The result is also obvious when more first order partial derivatives are not zero.

SECOND PART

$$\left(\frac{\partial G}{\partial F_r}\right)^o=0 \qquad (r=1, \cdots p)$$

$$\left(\frac{\partial G}{\partial X_s}\right)^o=0 \qquad (s=p+1, \cdots J)$$

Now

$$g(F_r, X_s)-G^o=L+O\left(F_r^3, X_s^3\right) \tag{13}$$

Thus for (12) and (13) together, we have L as a positive definite form.

Therefore $L > 0$ for all feasible (F_r, X_s) except when

$$F_r=0 \qquad (r=1, \cdots p)$$

$$X_s=0 \qquad (s=p+1, \cdots J)$$

then $\quad L=0$

in that case our feasible solution coincides with (f_i^o, x_j^o) which is a *unique* basic feasible solution. Thus for all (F_r, X_s) other than (f_i^o, x_j^o)

$$L > 0$$

$O(F_r^3, X_s^3)$ can be neglected in comparison to $O(F_r^2, X_s^2)$ which occurs in L.

Therefore in some neighbourhood of (f_i^o, x_j^o)

$$g-G^o > 0$$

Thus G^o is a minimum (relative).

Hence for any feasible solution (f_i, x_j) in some neighbourhood of non-degenerate basic feasible solution (f_i^o, x_j^o)

$$g > G^o$$

i.e. G^o is a relative minimum.

As in [1] we have

$$\left.\begin{array}{ll} y_i = \begin{cases} \dfrac{\partial G}{\partial f_i} & i=1, \cdots\cdots\cdots p \\ 0 & i=p+1, \cdots\cdots I \end{cases} \\ h_j = \begin{cases} 0 & j=1, \cdots\cdots\cdots p \\ \dfrac{\partial G}{\partial x_j} & j=p+1, \cdots\cdots J \end{cases} \end{array}\right\} \tag{14}$$

Hence our theorem :

For a non-degenerate basic feasible solution (f_i^o, x_j^o) to give a relative-minimum (which is also absolute minimum) the necessary and sufficient conditions are

$$\left.\begin{array}{ll} y_i^o \geqslant 0 & i=1, \cdots I \\ h_j^o \geqslant 0 & j=1, \cdots J \end{array}\right\} \tag{15}$$

ACKNOWLEDGEMENT

We are extremely indebted to Prof. R. S. Varma, D. Sc., F. N. I., Head of the Department of Mathematics and Dean, Faculty of Mathematics, University of Delhi, Delhi-7 for his encouragement and valuable help throughout the preparation of this paper.

REFERENCES

[1]. Wilde, J. Douglass "Differential Calculus in Non-Linear Programming" Operations Res. Nov. and Dec. 1962, Vol. 10, No. 6.

[2]. Karlin, Samuel. "Mathematical Methods and theory in Games, Programming and Economics". Reading Mass, 1959.

[3]. H. W. Kuhn & A. W. Tucker. "Non-Linear Programming" Proc. 2nd Berkeley Symposium on Math. Statistics and Probability. Aug. 1950, pp. 481–492.

[4]. M. A. Hanson. "Duality Theorem in Non-Linear Programming with Non-Linear Constraints" Australian Journal of Statistics, vol. 3, No. 2, 1961, pp. 64–71.

(University of Dehli)

(Received March 10, 1966)