ZERO OR NEAR-TO-ZERO LAGRANGE MULTIPLIERS IN LINEARLY CONSTRAINED NONLINEAR PROGRAMMING

L. F. ESCUDERO

We discuss in this work the using of Lagrange multipliers estimates in linearly constrained nonlinear programming algorithms and the implication of zero or near-to-zero Lagrange multipliers. Some methods for estimating the tendency of the multipliers are proposed in the context of a given algorithm.

1. INTRODUCTION

The linearly constrained nonlinear programming (LCNP) problem is

minimize $F(X)$ $X \in F \subseteq R^n$ (1.1)

where

$F \underline{\Delta} \{X | \bar{b} \geq AX \geq \underline{b}, \quad U \geq X \geq l\}$ (1.2)

where A is an $m.n$ matrix, $m<n$, and $F(X)$ is a general nonlinear twice continuously differenciable function, at least, for feasible points such that for all $\bar{X} \in F$ the level sets

$L(\bar{X}) \underline{\Delta} \{X \in F, \quad F(X) \leqslant F(\bar{X})\}$ (1.3)

are bounded. Let M be the set of constraints, E be the set of equality constraints (such that $i \in E$ if $\bar{b}_i = \underline{b}_i$), and J be the set of variables. Let $\overset{*}{A}$ be the $\overset{*}{t}.n$ matrix of <u>active</u> constraints at a local optimal point, say $\overset{*}{X}$ and $\overset{*}{b}$ the $\overset{*}{t}$-vector of right-hand-side corresponding to $\overset{*}{A}$ (i.e., $\overset{*}{A}\overset{*}{X} = \overset{*}{b}$), such that $\overset{*}{t} = |\overset{*}{W}|$ where $\overset{*}{W}$ is the set of active constraints and $\overset{*}{b}_i = \bar{b}_i \vee \underline{b}_i$ for $i \in \overset{*}{W}$. Let $\overset{*}{V}$ be the set of active variables at $\overset{*}{X}$, such that $j \in \overset{*}{V}$ if $\overset{*}{X}_j = U_j \vee l_j$ and $r = |\overset{*}{V}|$. Let $\overset{*}{I}$ be the $\overset{*}{r}.n$ matrix of active bounds at $\overset{*}{X}$, such that it is the $n.n$ identity matrix I from where the row related to variable $j \notin \overset{*}{V}$ has been deleted.

We shall define vector $g(X) \equiv g$ as the vector whose j-th element is $\delta F(X)/\delta X$ and the Hessian matrix $G(X) \equiv G$ as the symmetric matrix whose (i,j)-th element is $\delta^2 F(X)/\delta X_i \delta X_j$.

The algorithm /2/ concerned with this paper is assumed to generate a <u>sequence of feasible estimates</u> $\{X^{(k)}\}$ of $\overset{*}{X}$ (<u>weak local minimum</u>) by obtaining a <u>stepdirection</u> $d^{(k)}$ and a <u>steplength</u> $\alpha^{(k)}$ such that $X^{(k)} = X^{(k-1)} + d^{(k)}\alpha^{(k)}$ and $\lim X^{(k)} = \overset{*}{X}$ for $k \to \infty$, where $F(\overset{*}{X}) \leqslant F(\bar{X})$ for $\overset{*}{X}, \bar{X} \in F$ and $||\overset{*}{X} - \bar{X}|| \leqslant \sigma$ being σ small enough to define the neighbourhood of $\overset{*}{X}$.

Note that $AX - Y = \underline{b}$, $\bar{b} - \underline{b} \geqslant Y \geqslant 0 \leftrightarrow \bar{b} \geqslant AX \geqslant \underline{b}$; then $i \in \overset{*}{W}$ for $\overset{*}{Y}_i = 0 \vee \bar{b}_i - \underline{b}_i$. The X-variables are termed <u>structural</u>; the Y-variables are termed <u>slack</u>.

Because the constraints are a linear system, the properties of linear subspaces make it possible to state a simple characterization of all feasible moves from a feasible point. Consider the move between two feasible points $\overset{*}{X}$ and $\bar{X}$ along the manifold defined by the sets $\overset{*}{W}$ an $\overset{*}{V}$; by linearity $\overset{*}{A}(\overset{*}{X} - \bar{X}) = 0$ and $\overset{*}{I}(\overset{*}{X} - \bar{X}) = 0$ since $\overset{*}{A}\overset{*}{X} = \overset{*}{b}$, $\overset{*}{A}\bar{X} = \overset{*}{b}$ and $\overset{*}{X}_j = \bar{X}_j$ $\forall j \in \overset{*}{V}$ and, then,

$\overset{*}{A}d = 0, \quad \overset{*}{I}d = 0$ (1.4)

where d is the stepdirection from $\overset{*}{X}$ to $\bar{X}$ such that $\bar{X} = \overset{*}{X} + \alpha d$. Any vector d for which (1.4)

- L.F. Escudero, Centro de Investigación UAM-IBM. P) Castellana, 4 - Madrid-1

- Article rebut el Juny de 1982.

holds is a feasible stepdirection from $\overset{*}{X}$ — with respect to the above manifold; it is also termed <u>active stepdirection</u>; it will be descent if $F(\bar{X})<F(\overset{*}{X})$. Steplength α is required to be $0<\alpha\leq\alpha_m$, where α_m defines the maximum allowed steplength such that $\bar{X}$ is still feasible and $F(\bar{X})$ is descent.

Let us define a <u>non-active stepdirection</u> d as the feasible stepdirection such that some constraint or bound is removed from the sets $\overset{*}{W}$ and $\overset{*}{V}$, respectively; a feasible stepdirection d is non-active if $\exists i\in M-E\cap\overset{*}{W}$ for which $A_i d>0$ if $\overset{*}{Y}_i=0$, $A_i d<0$ if $\overset{*}{Y}_i=\bar{b}_i-\underline{b}_i$, or $\exists j\in\overset{*}{V}$ for which $d_j>0$ if $\overset{*}{X}_j=l_j$, $d_j<0$ if $\overset{*}{X}_j=U_j$.

The paper is organized as follows. Sec 2 describes the optimality conditions for $\overset{*}{X}$ being a weak local optimum in (1.1)-(1.2), so that the degenerate sets of active constraints and bounds are defined. Sec. 3 motivates the analysis of these sets, such that it states the reasons for analyzing the zero or near-to-zero Lagrange multipliers estimates. Sec. 4, briefly, describes the formulas that are being used in a given algorithm to obtain Lagrange multipliers estimates. Sec. 5 outlines the deactivating process in that algorithm. Sec. 6, finally, is devoted to some procedures that are proposed to get some insight on the tendency of zero or near-to-zero Lagrange multipliers estimates for optimal or 'quasi-optimal' solutions in the manifold defined by sets $\overset{*}{W}$ and $\overset{*}{V}$.

2. OPTIMALITY CONDITIONS

The necessary optimality conditions for $\overset{*}{X}$ being a weak local minimum are as follows. See equivalent conditions in /12/, /9/, /11/ and /8/ among others. As we state them below, they help to analyze the risk of using Lagrange multipliers estimates of nonbasic (structural and slack) variables, mainly when they are zero or near-to-zero, at a given optimal or 'quasi-optimal' point $\overset{*}{X}$ in the manifold defined by $\overset{*}{W}$ and $\overset{*}{V}$.

(i) $\overset{*}{X}\in F$ (feasible)

(ii) The <u>reduced gradient</u> vector, say $\overset{*}{h}$ of $F(X)$ vanishes, such that

$$\overset{*}{h}=\overset{*}{Z}^t\overset{*}{g}=0 \qquad (2.1)$$

where $\overset{*}{Z}$ is a $n.(n-\overset{*}{t}-\overset{*}{r})$ full column rank matrix, whose columns form the null basis of the range of matrix $(\overset{*}{A}^t,\ \overset{*}{I}^t)$ and, then,

$$\overset{*}{A}\overset{*}{Z}=0,\quad \overset{*}{I}\overset{*}{Z}=0 \qquad (2.2)$$

Based on (1.4) and (2.2), we may note that any linear combination of the columns of $\overset{*}{Z}$ give an active stepdirection d,

$$d=\overset{*}{Z}d_S \qquad (2.3)$$

where d_S is a $(n-\overset{*}{t}-\overset{*}{r})$-vector termed reduced stepdirection (or <u>superbasic stepdirection</u>), such that a vector d that cannot be expressed by (2.3) is not an active stepdirection in the manifold defined by $\overset{*}{W}$ an $\overset{*}{V}$.

Any point at which the reduced gradient h vanishes (2.1) is termed <u>constrained stationary point</u>. To see that, let us examine the Taylor-series expansion of F(X) about $\overset{*}{X}$ along an active stepdirection d in the manifold defined by $\overset{*}{W}$ an $\overset{*}{V}$:

$$F(\overset{*}{X}+\alpha d) = F(\overset{*}{X}) + \alpha d_S^t\overset{*}{Z}^t\overset{*}{g} + 1/2\alpha^2 d_S^t\overset{*}{Z}^t G(\overset{*}{X}+\theta\alpha d)\overset{*}{Z}d_S + 0(||d||_2) \qquad (2.4)$$

where θ satisfies $0\leq\theta\leq1$. Suppose that $\overset{*}{X}$ is a local minimum, but $d_S^t\overset{*}{Z}^t\overset{*}{g}<0$; then, there must exist $\bar{\alpha}>0$ such that $\alpha d_S^t\overset{*}{Z}^t\overset{*}{g}$ + $1/2\alpha^2 d_S^t\overset{*}{Z}^t G(\overset{*}{X}+\theta\alpha d)\overset{*}{Z}d_S<0$ for all $0<\alpha\leq\bar{\alpha}$ and, then, $F(\overset{*}{X}+\alpha\overset{*}{Z}d_S)<F(\overset{*}{X})$. Similarly, it can be shown that $\overset{*}{X}$ is non-optimal if $d_S^t\overset{*}{Z}^t\overset{*}{g}>0$. Therefore, $d_S^t\overset{*}{Z}^t\overset{*}{g}$ must be zero in order for $\overset{*}{X}$ to be a minimum. Thus, a necessary condition for $\overset{*}{X}$ to be a local minimum in the manifold defined by $\overset{*}{W}$ and $\overset{*}{V}$ is that $d_S^t\overset{*}{Z}^t\overset{*}{g}$ must vanish for every d_S, which implies (2.1) must hold.

The result (2.1) implies that $\overset{*}{g}$ must be a linear combination of the rows of $\overset{*}{A}$ and $\overset{*}{I}$,

$$\overset{*}{g} = \overset{*}{A}^t\overset{*}{\mu} + \overset{*}{I}^t\overset{*}{\lambda} \qquad (2.5)$$

for some vectors $\overset{*}{\mu}$ and $\overset{*}{\lambda}$; they are termed the <u>Lagrange multipliers</u> of the active constraints and bounds, respectively. Note that (2.5) is equivalent to (2.1) since any n-vector can be expressed as a linear combination of the columns of matrices ($\overset{*}{A}^t$, $\overset{*}{I}^t$) and $\overset{*}{Z}$, - and hence

$$\overset{*}{g} = \overset{*}{A}^t\overset{*}{\mu} + \overset{*}{I}^t\overset{*}{\lambda} + \overset{*}{Z}g_Z$$

for some vectors $\overset{*}{\mu}$, $\overset{*}{\lambda}$ and g_Z. Premultiplying $\overset{*}{g}$ by $\overset{*}{Z}^t$ and using (2.1) and (2.2), it results

$$0 = \overset{*}{Z}^t\overset{*}{g} = \overset{*}{Z}^t\overset{*}{A}^t\overset{*}{\mu} + \overset{*}{Z}^t\overset{*}{I}^t\overset{*}{\lambda} + \overset{*}{Z}^t\overset{*}{Z}g_Z = \overset{*}{Z}^t\overset{*}{Z}g_Z \quad (2.7)$$

Since $\overset{*}{Z}$ is a full column rank matrix, $\overset{*}{Z}^t\overset{*}{Z}$ is nonsingular and, then, (2.7) only holds for $g_Z=0$ such that, by using (2.2), it finally - results that (2.5) holds. By simple substitu tion, and using (2.2), we may see that (2.5) also implies (2.1).

(iii) Uniqueness of the Lagrange multi--- pliers.

Let us partition matrix $\overset{*}{A}$ and gradient $\overset{*}{g}$ --- such that $\overset{*}{A}=(\overset{**}{BS},\overset{*}{N})$ and $\overset{*}{g}=(\overset{*}{g}^t_{BS},\overset{*}{g}^t_N)^t$, where $\overset{*}{N}$ is a $\overset{*}{t}.\overset{*}{r}$ matrix defined by sets $\overset{*}{W}$ and $\overset{*}{V}$ and $\overset{*}{g}_N$ is the gradient of set $\overset{*}{V}$. Based in (2.5), $\overset{*}{\lambda}$ can be written,

$$\overset{*}{\lambda} = \overset{*}{g}_N - \overset{*}{N}^t\overset{*}{\mu} \quad (2.8)$$

such that $\overset{*}{\mu}$ satisfies the linear system

$$\overset{*}{g}_{BS} = (\overset{**}{BS})^t\mu \quad (2.9)$$

Point $\overset{*}{X}$ does not require A to be a full row rank matrix, but the uniqueness of vectors $\overset{*}{\mu}$ and $\overset{*}{\lambda}$ require $\overset{**}{BS}$ to have that property.

Assume that $\overset{*}{\mu}_1$ and $\overset{*}{\mu}_2$ satisfy (2.9). Then,

$$\overset{*}{g}_{BS} = (\overset{**}{BS})^t\overset{*}{\mu}_1 = (\overset{**}{BS})^t\overset{*}{\mu}_2,\ (\overset{**}{BS})^t(\overset{*}{\mu}_1-\overset{*}{\mu}_2)=0 \quad (2.10)$$

If the rows of $\overset{**}{BS}$ are linearly independent, $(\overset{*}{\mu}_1-\overset{*}{\mu}_2)=0$ and then,

$\overset{*}{\mu}_1=\overset{*}{\mu}_2$.

In any case, computational stability in the algorithms that obtain the sequence $\{X^{(k)}\}\to\overset{*}{X}$ requires $\{A_i\}$ to be linearly independent for $i\in M$.

There are several ways to characterize matrix Z /9/, /13/, /11/, /3/, /5/, /6/; depending - on how big $\overset{*}{t}$ and $\overset{*}{r}$ are expected, the most --- atractive ways to obtain it are based ------- (although it is not explicitly calculated) - on variable-reduction and QR-factorization - of matrix $(BS)^t$.

(iv) The sign of Lagrange multipliers --- must be as follows:

$\overset{*}{\mu}_i \gtrless 0$ for $i\in E$ (equality constraint).

$\overset{*}{\mu}_i=0$ for $i\notin W$ (non-active inequality cons traint).

$\overset{*}{\mu}_i\geq 0$ for $i\in M\text{-}E\cap\overset{*}{W}$ such that $\overset{*}{Y}_i=0$ (active inequality constraint whose associa-- ted slack variable has the value zero)

$\overset{*}{\mu}_i\leq 0$ for $i\in M\text{-}E\cap\overset{*}{W}$ such that $\overset{*}{Y}_i=\bar{b}_i-\underline{b}_i$ (ac tive inequality constraint whose asso ciated slack variable takes its upper bound).

$\overset{*}{\lambda}_j=0$ for $j\notin\overset{*}{V}$ (non-active variable).

$\overset{*}{\lambda}_j\geq 0$ for $j\in\overset{*}{V}$ such that $\overset{*}{X}_j=l_j$ (active va-- riable at its lower bound).

$\overset{*}{\lambda}_j\leq 0$ for $j\in\overset{*}{V}$ such that $\overset{*}{X}_j=U_j$ (active va-- riable at its upper bound).

The set $D_1\cup D_2$, where $D_1\triangleq\{i\in M\text{-}E\cap\overset{*}{W}$ for ----- $\overset{*}{\mu}_i=0\}$ and $D_2\triangleq\{j\in\overset{*}{V}$ for $\overset{*}{\lambda}_j=0\}$ is termed -- <u>degenerate</u> set of active constraints and ---- bounds.

The reason for condition (iv) is as fo----- llows: Since a non-active stepdirection is -- also feasible, point $\overset{*}{X}$ will not be a local mi nimum if $F(\bar{X})<F(\overset{*}{X})$ such that $\bar{X}=\overset{*}{X}+\alpha d$ and d is a non-active stepdirection; to avoid this --- possibility we must add a condition that ensu res $\overset{*}{g}^t d\not<0$ for every non-active stepdirection and, then, $F(\bar{X})\not<F(\overset{*}{X})$ where $F(\bar{X})$ can be ----- written

$$F(\overset{*}{X}+\alpha d) = F(\overset{*}{X}) + \alpha\overset{*}{g}^t d + 0(||d||) \quad (2.11)$$

such that for $\overset{*}{g}^t d<0$ there is always a small scalar $\alpha>0$ for which

$F(\overset{*}{X}+\alpha d)<F(\overset{*}{X})$.

Note that the optimality of $\overset{*}{X}$ requires that (2.5) holds; then, it results,

$$\overset{*}{g}^t d = \overset{*}{\mu}^t\overset{*}{A}d + \overset{*}{\lambda}^t\overset{*}{I}d \quad (2.12)$$

Since $A_i d=0$ for $i \in E$ (see that $\overset{*}{\mu}_i$ is not restricted in sign), $\overset{*}{g}{}^t d$ can be written,

$$\overset{*}{g}{}^t d = \sum_{i \in M-E \cap \overset{*}{W}} \overset{*}{\mu}_i A_i d + \sum_{j \in \overset{*}{V}} \overset{*}{\lambda}_j d_j \geqslant 0 \qquad (2.13)$$

Condition (2.13 holds if $\overset{*}{\mu}_i \geqslant 0$ for $\overset{*}{Y}_i=0$, $\mu_i \leqslant 0$ for $\overset{*}{Y}_i=\bar{b}_i-\underline{b}_i$, $\overset{*}{\lambda}_j \geqslant 0$ for $\overset{*}{X}_j=1_j$, and $\overset{*}{\lambda}_j \leqslant 0$ for $\overset{*}{X}_j=U_j$ since, otherwise, it is always possible to find a stepdirection d such that $A_i d=0$, $\overset{*}{I}d=0$, $A_\ell d>0$ if $\overset{*}{Y}_\ell=0$ for $i \in \overset{*}{W}$, $\ell \in M-E \cap \overset{*}{W}$, $i \neq \ell$, or $A_i d=0$, $\overset{*}{I}d=0$, $A_\ell d<0$ if $\overset{*}{Y}_i=\bar{b}_i-\underline{b}_i$ for $i \in \overset{*}{W}$, $\ell \in M-E \cap \overset{*}{W}$, $i \neq \ell$, or $\overset{*}{A}d=0$, $d_j=0$, $d_k>0$ for $j,k \in \overset{*}{V}$, $\overset{*}{X}_k=1_k$, $j \neq k$, or $\overset{*}{A}d=0$, $d_j=0$, $d_k<0$ for $j,k \in \overset{*}{V}$, $\overset{*}{X}_k=U_k$, $j \neq k$. If $\overset{*}{A}$ is regular (its rows are linearly independent) then the above non-active stepdirection could be very easily found by using its pseudo-inverse matrix; the sign of the Lagrange multipliers may be also proved by using the Farkas'lemma by without requiring the regularity assumption on matrix $\overset{*}{A}$ (see /8/).

(v) Positive semi-definiteness of the Hessian matrix.

The reduced Hessian matrix $\overset{*}{H}$ must be positive semi-definite, where

$$\overset{*}{H} \equiv \overset{*}{Z}{}^t \overset{*}{G} \overset{*}{Z} \qquad (2.14)$$

To see that, let us examine the Taylor-series expansion of F(X) about $\overset{*}{X}$ along an active stepdirection d such that, by using (2.1) and (2.3), it can be written

$$F(\overset{*}{X}+\alpha d) = F(\overset{*}{X}) + 1/2\alpha^2 d_S^t \overset{*}{Z}{}^t G(\overset{*}{X}+\theta\alpha d)\overset{*}{Z} d_S + + 0(||d||_2) \qquad (2.15)$$

If G is indefinite, by continuity $\overset{*}{G}$ will be indefinite for $\alpha>0$ being small enough, such that by definition $\exists d_S$ for which $d_S^t \overset{*}{Z}{}^t G(\overset{*}{X}+\theta\alpha d)\overset{*}{Z} d_S<0$ and, then, $\overset{*}{X}$ is not a local minimum.

Note that the above condition is equivalent to require that matrix $\overset{*}{G}$ be positive semi-definite but only for the active stepdirections in the manifold defined by $\overset{*}{W}$ and $\overset{*}{V}$.

If the degenerate set $D_1 \cup D_2$ is not empty then the positive semi-definiteness property of $\overset{*}{G}$ must be extended to the non-active step direction d for which it holds $A_i d>0 \wedge \overset{*}{Y}_i=0 \wedge i \in D_1$, or $A_i d<0 \wedge Y_i=\bar{b}_i-\underline{b}_i \wedge i \in D_1$, or $d_j>0 \wedge \overset{*}{X}_j=1_j \wedge j \in D_2$, or $d_j<0 \wedge \overset{*}{X}_j=U_j \wedge j \in D_2$. To see that, recall that $\overset{*}{g}{}^t d$ (2.12) vanishes for active stepdirections d (where $\overset{*}{A}d=0$, $\overset{*}{I}d=0$), but also for non-active stepdirections d where the operators vector ρ in $\overset{*}{A}d\rho 0$, $\overset{*}{I}d\rho 0$ includes the feasible inequality in the constraints and bounds whose Lagrange multipliers $\overset{*}{\mu}_i$ and $\overset{*}{\lambda}_i$ are zero.

Conditions (i)-(ii) and (iv)-(v) are <u>necessary conditions</u> for local optimality; if Hessian matrix is required in (v) to be positive definite then they are <u>sufficient conditions</u>.

3. MOTIVATION FOR ZERO OR NEAR-TO-ZERO LAGRANGE MULTIPLIERS ANALYSIS.

Point $\overset{*}{X}$ is a local minumum in the manifold defined by $\overset{*}{W}$ and $\overset{*}{V}$ if conditions (i)-(ii) are satisfied and matrix $\overset{*}{H}$ (2.14) is positive definite for all d_S. To test if $\overset{*}{X}$ is also the solution of problem (1.1)-(1.2), it is required to analyze the sign of the active constraints and bounds Lagrange multipliers, such that if $\overset{*}{\mu}_i$ for $\forall i \in M-E \cap \overset{*}{W}$ (or $\overset{*}{\lambda}_j$ for $\forall j \in \overset{*}{V}$) have not the apropiate signs then constraint i (or bound j) must be deactivated (that is, a non-active stepdirection must be obtained). Note that e.g. if $\overset{*}{\mu}_i<0$ for $i \in M-E \cap \overset{*}{W}$ and $\overset{*}{Y}_i=0$, the feasible stepdirection d that deactivates constraint i (and, then, $A_i d>0$ such that constraint i is delected from $\overset{*}{W}$) is descent (see (2.11) and (2.13)); viceversa, a descent stepdirection is also non-active (and, then, feasible). On the other hand, e.g. if $\overset{*}{\mu}_i>0$ for $i \in M-E \cap \overset{*}{W}$ and $\overset{*}{Y}_i=0$, any non-active stepdirection is non-descent; viceversa, a descent -- stepdirection is non-feasible (and, then, -- $A_i d<0$).

Note that the Lagrange multipliers take the first-order rate of change in the objective function (1.1) due to a change in the right-hand-side of the related constraint or bound; see (2.13). A caution has to be made since - the magnitude of the Lagrange multipliers is not invariant to scaling changes.

We may see that the sign is more important - than the magnitude of the Lagrange multi--- pliers. Note that if they are not obtained - with exact arithmetic, the 'computed' value of e.g. $\overset{*}{\mu}_i$ for $i \in M\text{-}E \cap \overset{*}{W}$ and $\overset{*}{Y}_i=0$ may be --- 'sligthly' negative when the 'exact' value - is positive and, then, once the constraint - is chosen to be deactivated, the related -- non-active stepdirection is non-descent; if the 'computed' value is positive when the -- exact value is negative, then a premature -- termination of the algorithm may occur ----- without reaching the optimum $\overset{*}{X}$. Any computer algorithm works, by its own nature, with finite precision and the results are subject - to unstabilities due to cancellation and --- rounding errors in intermediate operations - /11/.

An additional difficulty arises in the pre-- sence of zero computed value of some Lagrange multiplier since, in that case, there is more uncertainty on the sign of its exact value. - Recall that if it is zero, the positive definiteness property of matrix $\overset{*}{H}$ is not enough - to guarantee (together with the other condi-- tions) that $\overset{*}{X}$ is a local minimum.

Some algorithms deactivate constraints and - bounds even before the local optimum in a given manifold is reached such that, once a --- 'quasi-optimal' solution is obtained, estimates of the Lagrange multipliers are calcula-- ted and, based on them, the desactivating pro cess is executed; see /13/ among others. ---- These estimations introduce a new uncertainty on the sign of the 'exact' Lagrange multi---- pliers.

4. LAGRANGE MULTIPLIERS ESTIMATES.

See in /10/, /3/, several methods to obtain Lagrange multipliers estimates. We use two - types of formulas, the so-termed first-order estimates, see in /3/ the motivation for not using second-order estimates.

Following a traditional approach /13/, let - the active constraints matrix, say A be par-titioned as

$$\bar{A}d = (\bar{B},\bar{S},\bar{N}) \begin{pmatrix} d_B \\ d_S \\ d_N \end{pmatrix} = 0 \qquad (4.1)$$

where the basic stepdirection d_B is used to satisfy the constraints set, the superbasic stepdirection d_S is allowed to vary to minimize F(X) (1.1) and the nonbasic stepdirec-- tion d_N is zero, such that set $\bar{V}$ is fixed at any of their bounds. Here $\bar{B}\bar{S} \equiv (\bar{B},\bar{S})$ and $\bar{B}$ is a $\bar{t}.\bar{t}$ nonsingular matrix. At each iteration, the problem then becomes determining vector $d=(d_B^t,d_S^t,d_N^t)^t$ so that it is feasible-descent. Since $d_N=0$ and d_S is allowed to be free, it results

$$d_B = -\bar{B}^{-1}\bar{S}d_S$$

such that the variable-reduction characterization of matrix Z can be written

$$Z = \begin{pmatrix} -B^{-1}S \\ I \\ 0 \end{pmatrix} \qquad (4.3)$$

so that (2.3) holds.

The cuadratic approximation of the uncons--- trained reduced problem of minimizing F(X) - in the manifold $\bar{W}$ and $\bar{V}$ as a function of the current superbasic set of variables d_S can - be written

$$\text{minimize } \bar{h}^t d_S + 1/2 d_S^t \bar{H} d_S \qquad (4.4)$$

where $\bar{h}$ and $\bar{H}$ are given by (2.1) and (2.14), respectively. Note that $\bar{h}$ can also be written

$$\bar{h} = \bar{g}_S - \bar{S}^t\bar{\mu}_B \qquad (4.5)$$

where $\bar{\mu}_B$ solves the linear system

$$\bar{g}_B = \bar{B}^t\mu_B \qquad (4.6)$$

such that $\bar{g}_{BS} \equiv (\bar{g}_B^t, \bar{g}_S^t)^t$ where $\bar{g}_B$ and $\bar{g}_S$ are - the basic and superbasic gradients, respectively. Theoretically, the algorithm continues till $||\bar{h}||=0$ or the superbasic set is empty and, then, the deactivating process is executed by analyzing the Lagrange multipliers (except if it has been decided to do so in - the presence of 'quasi-optimal' solutions so that Lagrange multipliers estimates are obtained).

Let assume that $\overset{*}{X}$ is an optimal point in the manifold $\overset{*}{W}$ and $\overset{*}{V}$. Then, the Lagrange multipliers vector $\overset{*}{\mu}$ is obtained by solving system (2.9); let $\overset{*}{\mu}_{BS}$ be the solution. Since -- $||\overset{*}{h}||=0$ or the superbasic set is empty, it - is clear that $\overset{*}{\mu}_{BS}=\overset{*}{\mu}_B$(4.6). Thus, it is not - required to solve (2.9) since $\overset{*}{\mu}_B$ is updated at each iteration to solve problem (4.4) --- (see /12/, /2/). Note also that $\overset{*}{\mu}_B$ is the negative of the LP simplex multipliers. The -- Lagrange multipliers vector $\overset{*}{\lambda}$ of the active (nonbasic) structural variables is obtained by using formula (2.8); note also that $\overset{*}{\lambda}$ and $\overset{*}{h}$ are the negative of the LP reduced cost -- vectors related to the 'nonbasic' and 'superbasic' associated LP subproblems, respectively.

When a 'quasi-optimal' solution, say $\bar{X}$ is obtained in a given manifold defined by $\bar{W}$ and $\bar{V}$, $\bar{\mu}_{BS} \neq \bar{\mu}_B$ and $\bar{\lambda}_{BS} \neq \bar{\lambda}_B$. Let us term $\bar{\mu}_B$ and $\bar{\lambda}_B$ as <u>basic-based</u> active constraints and bounds Lagrange multipliers estimates, respectively; and $\bar{\mu}_{BS}$ and $\bar{\lambda}_{BS}$ as <u>basic-superbasic-based</u> active constraints and bounds Lagrange multipliers estimates, respectively. Of course, - estimates based on the basic-superbasic set are generally more accurate than those based in the basic set; see in /11/ a good discussion on the subject.

The motivation for obtaining 'quasi-optimal' solutions and, then, interrupting the minimization of the unconstrained reduced nonlinear problem is based on the assumption that it is likely that current sets $\bar{W}$ and $\bar{V}$ are - not the optimal sets $\overset{*}{W}$ and $\overset{*}{V}$ in problem ---- (1.1)-(1.2) and, then, it could be beneficial to analyze if it is worthy to delete some active constraint or bound before reaching the optimum in the reduced problem, but after reaching a solution close to that optimum.

Note that estimation $\bar{\mu}_B$ is already obtained. Estimation $\bar{\mu}_{BS}$ is based on the $\bar{Q}\bar{R}$-factorization of matrix $\bar{B}\bar{S}$. While solving the unconstrained reduced problem, eihter a basic or a superbasic variable may strike a bound during the search. If a superbasic variable strikes a bound, it is made nonbasic, the dimension of the manifold is reduced by one, and the - search continues. If a basic variable strikes a bound, the basic variable is exchanged with an appropiate superbasic variable and the resulting superbasic variable is made nonbasic. The estimation $\bar{\mu}_B$ is updated at each iteration since $\bar{h}$ (4.5) requires it; but $\bar{\mu}_{BS}$ is - only obtained when it is required to execute the deactivating process.

Estimation $\bar{\mu}_{BS}$ is obtained by minimizing the square of the euclidean length

$$||\bar{g}_{BS} - (\bar{B}\bar{S})^t \mu_{BS}||_2 \tag{4.7}$$

For obtaining $\bar{\mu}_{BS}$ (<u>minimal linear least square solution</u>) we use an implementation of the Gram-Schmidt QR-factorization of a matrix -- where the number of rows is greater than the number of columns; see the motivation and details in /1/, /4/, but the matrices involved are as follows.

Let $\bar{Q}$ be a $(n-\bar{t}-\bar{r}).\bar{t}$ orthogonal matrix and $\bar{R}$ a $\bar{t}.\bar{t}$ nonsingular upper triangular matrix -- with identity diagonal such that

$$(\bar{B}\bar{S})^t = \bar{Q}\bar{R} \tag{4.8}$$

It can be shown that the vector $\bar{\mu}_{BS}$ that minimize problem (4.7) is also the vector that satisfies the system

$$\bar{R}\mu_{BS} = \bar{Q}^t \bar{g}_{BS} \qquad (4.9)$$

Then, it is required to calculate $\bar{Q}$ and $\bar{R}$ -- for obtaining $\bar{\mu}_{BS}$; they are updated each time a basic-superbasic (structural or slack) variable is made nonbasic or a nonbasic (structural or slack) variable is deactivated ; -- they are calculated anew the first time that μ_{BS} is used, or after a given number of updatings so that some unstability due to too-many intermediate operations is avoided /1/, - /4/.

5. DEACTIVATING PROCESS.

Assume that $\overset{*}{X}$ is an 'optimal' or 'quasi-optimal' solution in the manifold defined by $\overset{*}{W}$ - and $\overset{*}{V}$; assume also that the given algorithm selects, at each time, only one candidate -- nonbasic (structural or slack) variable to - be deactivated . Note that not all nonbasic variables are candidates for being desactivated; e.g. if $\overset{*}{X}$ is only 'quasi-optimal', 'unsafe' nonbasic variables do not belong to the candidate set if an anti-zigzagging strategy is to be used. See /2/,/5/. Let C_1 and C_2 be candidate sets of structural nonbasic variables and active inequality constraints, respectively. Note that each slack variable is associated with an inequality constraint.

Let γ be an indicator such that $\gamma=1$ means -- the basic superbasic-based estimate $\overset{*}{\mu}_{BS}$ (4.6) is allowed for $\overset{*}{X}$; otherwise ($\gamma=0$), only the basic-based estimate $\overset{*}{\mu}_B$ (4.9) can be used.

Since $\overset{*}{\mu}_{BS}=\overset{*}{\mu}_{BS}$ for $\overset{*}{X}$ being optimal in the manifold $\overset{*}{W}$ and $\overset{*}{V}$, indicator is set to zero - for the given iteration even if basic-superbasic-based Lagrange multipliers estimates - are allowed.

Case $\gamma=0$.

The structural nonbasic variable to be deac-- tivated is the variable, say $k \in \overset{*}{V}$ with the most favorable basic-based Lagrange multi--- plier estimate $\overset{*}{\lambda}_{Bk}$ such that

$$|\overset{*}{\lambda}_{B_k}| = \max\{|\overset{*}{\lambda}_{Bj}| \,\big|\, \overset{*}{\lambda}_{Bj}<-\varepsilon_1 \wedge \overset{*}{X}_j = 1_j,\ \overset{*}{\lambda}_{Bj}>\varepsilon_2 \wedge \overset{*}{X}_j = U_j\ \ j \in C_1\} \qquad (5.1)$$

where ε_1 is a given small positive tolerance (tipically, $\varepsilon_1=10^{-4}$) that is intended to give lower priority to variables with zero or near-to-zero Lagrange multipliers estimates.

If k=0 then the active inequality constraint, say $k \in M-E \cap \overset{*}{W}$ with the most favorable basic-based Lagrange multiplier estimate $\overset{*}{\mu}_{B_k}$ is to - be deactivated , such that

$$|\overset{*}{\mu}_{B_k}| = \max\{|\overset{*}{\mu}_{B_i}| \,|\, \overset{*}{\mu}'_{B_i}<-\varepsilon_1 \wedge \overset{*}{Y}_i = 0,\ \overset{*}{\mu}_{B_i}>\varepsilon_i \wedge \overset{*}{Y}_i = \bar{b}_i - \underline{b}_i\ \ i \in C_2\} \qquad (5.2)$$

for the same tolerance ε_1.

Assume that k=0. If the following condition does not hold

$$(\exists j|-\varepsilon_1 \leq \overset{*}{\lambda}_{B_j} \leq \varepsilon_1\ \ j \in C_1) \vee (\exists i|-\varepsilon_1 \leq \overset{*}{\mu}_{B_i} \leq \varepsilon_1\ \ i \in C_2) \qquad (5.3)$$

the action to be taken depends on the character of point $\overset{*}{X}$: if it is a 'quasi-optimal' - solution in the manifold $\overset{*}{W}$ and $\overset{*}{V}$, the next - iteration obtains the related superbasic --- stepdirection; otherwise, it is assumed that $\overset{*}{X}$ is also an optimal solution in problem --- (1.1)-(1.2). If (5.3) holds, let us redefine sets D_1 and D_2 (see sec. 2) such that D_1 takes the subset of C_1 for which the first --- part of (5.3) holds and D_2 takes the subset of C_2 for which the second part of (5.3) --- holds. Thus, set $D_1 \cup D_2$ defines the zero or near-to-zero Lagrange multipliers estimates.

Case $\gamma=1$.

Note that estimates $\overset{*}{\mu}_{BS}$ and $\overset{*}{\mu}_B$ (and, then, - $\overset{*}{\lambda}_{BS}$ and $\overset{*}{\lambda}_B$) are to be used. Generally, $\overset{*}{\mu}_{BS}$ -

is more accurate than $\overset{*}{\mu}_B$ since it uses more information; but, it makes sense to select a candidate nonbasic (structural or slack) variable if among other requirements /5/, <u>both estimates agree in the appropiate sign</u>. In any case, note that the reduced gradient --- (with structural and, although provisionally. slack elements) only uses the basic-based estimate; see (4.5).

The structural nonbasic variable to be deactivated is the variable, say $k \in \overset{*}{V}$ with the most favorable basic-superbasic-based Lagrange multiplier estimate $\overset{*}{\lambda}_{BS_k}$ whose basic-based estimate $\overset{*}{\lambda}_{B_k}$ agrees in sign or, at least, is zero or near-to-zero, such that

$$|\overset{*}{\lambda}_{BS_k}| = \max\{|\overset{*}{\lambda}_{BS_j}| \,\Big|\, \overset{*}{\lambda}_{BS_j} < -\varepsilon_1 \wedge \overset{*}{\lambda}_{B_j} \leq \varepsilon_1 \wedge \overset{*}{X}_j = l_j, \quad \overset{*}{\lambda}_{BS_j} > \varepsilon_1 \wedge \overset{*}{\lambda}_{B_j} \geq -\varepsilon_1 \wedge \overset{*}{X}_j = U_j \quad j \in C_1\} \tag{5.4}$$

If k=0 the active inequality constraint to be desactivated is the constraint, say ---- $k \subset M-E \cap \overset{*}{W}$ with the most favorable basic-superbasic-based Lagrange multiplier estimate --- $\overset{*}{\mu}_{BS_k}$ whose estimation $\overset{*}{\mu}_{B_k}$ agrees in sign or, at least, is zero or near-to-zero, such that

$$|\overset{*}{\mu}_{BS_k}| = \max\{|\overset{*}{\mu}_{BS_i}| \,\Big|\, \overset{*}{\mu}_{BS_i} < -\varepsilon_1 \wedge \overset{*}{\mu}_{B_i} \leq \varepsilon_1 \wedge \overset{*}{Y}_i = 0, \quad \overset{*}{\mu}_{BS_i} > \varepsilon_1 \wedge \overset{*}{\mu}_{B_i} \geq -\varepsilon_1 \wedge \overset{*}{Y}_i = \bar{b}_i - \underline{b}_i \quad i \in C_2\} \tag{5.5}$$

for the same tolerance ε_1 used above.

When using (5.4) and (5.5) it is suggested to give the lowest priority to the candidate nonbasic (structural or slack) variables for which the basic-based Lagrange multipliers estimates are zero or near-to-zero.

When a candidate nonbasic (structural or --- slack) variable is deactivated with zero or near-to-zero basic-based Lagrange multiplier estimate, its related element in the new reduced gradient $\bar{h}$ (4.5) will be $\varepsilon_1 + \varepsilon_2$ if -- $\overset{*}{X}_j = U_j$ or $\overset{*}{Y}_i = \bar{b}_i - \underline{b}_i$ respectively, and $-(\varepsilon_1 + \varepsilon_2)$ if $\overset{*}{X}_j = l_j$ or $\overset{*}{Y}_i = 0$, respectively. Note that $\bar{h}$ will be used for obtaining the superbasic -- stepdirection of the next iteration. ε_2 is a given small positive tolerance (typically, - $\varepsilon_2 = 10^{-4}$) used for the required perturbation in the apropiate sign of the zero or -- near-to-zero basic-based Lagrange multipliers estimates.

Assume that k=0. If the following condition does not hold

$$(\exists j \,|\, -\varepsilon_1 \leq \overset{*}{\lambda}_{BS_j} \leq \varepsilon_1 \quad j \in C_1) \vee (\exists i \,|\, -\varepsilon_1 \leq \overset{*}{\mu}_{BS_i} \leq \varepsilon_1 \quad i \in C_2\} \tag{5.6}$$

the next iteration obtains the superbasic -- stepdirection; note that $\overset{*}{X}$ is quasi optimal. Let $D_1 \cup D_2$ define the set of zero or near-to-zero basic-superbasic-based Lagrange multipliers estimates as it was defined for $\gamma=0$, such that their related basic-based estimates are favorable or, at least, zero or near-to-zero.

<u>6. ESTIMATING THE TENDENCY OF THE LAGRANGE MULTIPLIERS ESTIMATES.</u>

When conditions (5.3) for $\gamma=0$ and (5.6) for $\gamma=1$ hold, the ambiguity on the Lagrange multipliers estimates does not allow to select a variable from set $D_1 \cup D_2$ to be deactivated, , except if some perturbation on the active -- bounds of the X- and Y-variables is produced; in this way, an estimation on the tendency of the zero or near-to-zero Lagrange multipliers estimates could be obtained.

Before describing the procedure and since -- the structural variables may be classified - according to the linearity of the terms of - the objective function, let us make the following partition of these variables in pure linear, linear with variable-coefficient and nonlinear; see in /7/ how to use, in a given algorithm, this and other types of variables partition . A variable is <u>pure linear</u> if its coefficient is constant in all terms; a variable is <u>linear with-variable-coefficient</u> - if, for a given value of the other variables that are used in the same term, it is a linear function of the given variable; and a - variable is <u>nonlinear</u> if, for a given value of the other variables in the same term, it is a nonlinear function of the given variable. An example is as follows: $F(X) = 4X_1 + X_2 \log X_3$; variable X_1 is pure linear, X_2 is linear with variable-coefficient, and X_3 is nonlinear. Let L define the set of linear variables and P the set of pure linear variables.

The estimation of the tendency of the zero - or near-to-zero Lagrange multipliers estimates is only obtained for the basic-superbasic-based estimates; but, note that $\overset{*}{\mu}_{BS}=\overset{*}{\mu}_B$ (and, then, $\overset{*}{\lambda}_{BS}=\overset{*}{\lambda}_B$) for $\gamma=0$.

The tests to select a variable to be deactivated from set $D_1 \cup D_2$ could be as follows.

Test 1. Select a (structural) variable from set D_1.

It is only performed if $J/\overset{*}{V}=P$; i.e., the whole basic and superbasic set of variables is pure linear; then, estimation $\overset{*}{\mu}_{BS}$ will not - be modified by the perturbation to be produced in test 1.

Let us produce an small perturbation on the active bound of variable, say j for -------- $j \in D_1 \cap (J/L)$; let $\tilde{g}_j$ be the element in the objective function gradient related to variable j evaluated at the perturbed solution $\tilde{X}$, such that $\tilde{X}_\ell=\overset{*}{X}_\ell$ for $\ell \in J$, $\ell \neq j$, $\tilde{X}_j=\overset{*}{X}_j+\varepsilon_2$ if $\overset{*}{X}_j=l_j$ -- and $\overset{*}{X}_j-\varepsilon_2$ if $\overset{*}{X}_j=U_j$. Note that only element $\tilde{g}_j$ is required at each perturbation, being unchanged the other elements of $\overset{*}{g}$. Note also - that variable j is nonlinear, since the gradient elements of linear variables are constant (and then, the solution perturbation -- has not any effect).

The perturbed basic-superbasic-based estimate $\overset{*}{\lambda}_{BS_j}$ can be written (see (2.8))

$$\tilde{\lambda}_{BS_j} = \overset{*}{\lambda}_{BS_j} - \overset{*}{g}_j + \tilde{g}_j \qquad (6.1)$$

If the following condition holds, variable j is to be deactivated (and, then, set k=j).

$$(\tilde{\lambda}_{BS_j}-\overset{*}{\lambda}_{BS_j}<-\varepsilon_3 \,|\, \overset{*}{X}_j=l_j) \vee (\tilde{\lambda}_{BS_j}-\overset{*}{\lambda}_{BS_j}>\varepsilon_3 \,|\, \overset{*}{X}_j=U_j) \qquad (6.2)$$

Otherwise, it seems that the tendency of $\overset{*}{\lambda}_{BS_j}$ is not favorable and then, variable j is not to be deactivated. ε_3 is a given small positive tolerance, such that any number with -- magnitude less than ε_3 will be discarded (i.e. set to zero) as being insignificance in any - circunstance; typically, $\varepsilon_3=10^{-12}$.

The first variable $j \in D_1 \cap (J/L)$ that satisfies (6.2) is the variable, say k to be deactivated.

Test 2. Select a (structural) variable from set D_1.

If k=0 (i.e., test 1 did not select any variable) test 2 will be used; it is not performed if $J/\overset{*}{V}=P$; i.e., the whole basic and - superbasic set of variables is pure linear - since, otherwise, the results would be very similar to those obtained in test 1.

Let us perturb simultaneously the active --- bounds of the whole set $D_1 \cap (J/L)$ with the same criterion used in test 1; gradients $\tilde{g}_B$ -- for $\gamma=0$ and $\tilde{g}_{BS}$ for $\gamma=1$ are, alternatively, evaluated at the perturbed solution $\tilde{X}$ and, - then, the perturbed basic-based estimate $\tilde{\mu}_B$ is obtained for $\gamma=0$ by solving system (4.6) with $\overset{*}{g}_B$ being substituted by $\tilde{g}_B$ and the perturbed basic-superbasic-based estimate $\tilde{\mu}_{BS}$ - is obtained for $\gamma=1$ by solving problem (4.7) being $\overset{*}{g}_{BS}$ substituted by $\tilde{g}_{BS}$.

The scope of this work does not cover the -- procedures for solving system (4.6), nor problem (4.7); in any case, inverse matrix $\overset{*}{B}^{-1}$ is not obtained in the first case and system (4.9) is not explicitly solved in the second case. The updated $\overset{**}{LU}$ and $\overset{**}{QR}$ factorizations - of matrices $\overset{*}{B}$ and $(\overset{**}{BS})^t$, respectively are - used; see /1,2,4/.

The perturbed estimate $\tilde{\lambda}_{BS_j}$ for $j \in D_1 \cap (J/L)$ - is obtained by using in (2.8) the perturbed element $\tilde{g}_j$ evaluated at the perturbed solution $\tilde{X}$ and the perturbed estimate $\tilde{\mu}_{BS}$; note that $\tilde{\mu}_{BS}=\tilde{\mu}_B$ for $\gamma=0$.

The first variable $j \in D_1 \cap (J/L)$ that satisfies (6.2) is the variable, say k to be deactivated.

Test 3. Select an (active inequality) constraint from set D_2.

We will obtain directly, in test t3, the basic-superbasic-based estimate $\tilde{\mu}_{BS}$ of the ---

constraints set $\overset{*}{W}$ that solves the problem --

$$\min || \tilde{g}_{BS} - (\overset{**}{BS})^t \mu_{BS} ||_2^2 \qquad (6.3)$$

for $\gamma=1$, or solves the system

$$\tilde{g}_B = \overset{*}{B}^t \mu_B \qquad (6.4)$$

for $\gamma=0$, where $\tilde{\mu}_{BS}=\tilde{\mu}_B$ for $\gamma=0$,

$(\overset{**}{BS})^t = \overset{**}{QR}$, $\tilde{g}_B \equiv g(\overset{*}{X}_B+\tilde{d}_B)$ and $\tilde{g}_{BS} \equiv g(\overset{*}{X}_{BS}+\tilde{d}_{BS})$, such that $\overset{*}{X}_B$ and $\overset{*}{X}_{BS}$ take the optimal or -quasi-optimal values of the basic set and basic and superbasic set of variables, respectively, and $\tilde{d}_B$ and $\tilde{d}_{BS}$ are the related deviations from points $\overset{*}{X}_B$ and $\overset{*}{X}_{BS}$.

Deviations $\tilde{d}_{BS}$ satisfies the right-hand-side perturbed original problem, where the nonbasic variables are fixed to their values $\overset{*}{X}_N$ and only the active constraints set $\overset{*}{W}$ is considered, such that

$$\begin{aligned}\overset{**}{BS}(\overset{*}{X}_{BS}+\tilde{d}_{BS}) &= \overset{*}{b}-\overset{**}{N}\overset{*}{X}_N+\varepsilon_2 I_\varepsilon \\ &\equiv \overset{*}{b}_{BS}+\varepsilon_2 I_\varepsilon \qquad (6.5)\end{aligned}$$

where $\overset{*}{X}_j=1_j \vee U_j$ for $j \in \overset{*}{V}$, and $\overset{*}{b}_i=\bar{b}_i \vee \underline{b}_i$ for $i \in \overset{*}{W}$. Note that $\overset{*}{b}_{BS}$ takes the right-hand-side vector to be perturbed in system (6.5), where - $\overset{**}{BS}$ takes the constraints matrix such that -- $\overset{**}{BS}_{rc}$ takes the (r,c)-th element of matrix $\overset{**}{BS}$ for $r=1,\dots,\overset{*}{t}$ and $c=1,\dots,n-\overset{*}{t}-\overset{*}{r}$. Let i(r) define the constraint $i \in \overset{*}{W}$ related to index r in (6.5); similarly, j(c) defines the basic or superbasic variable $j \in J/\overset{*}{V}$ related to index c in (6.5). I_ε defines a $\overset{*}{t}$-vector such that $I_{\varepsilon_r}=-1$ for r such that $i(r) \in D_2 \wedge \overset{*}{Y}_i=\bar{b}_i-\underline{b}_i$, $I_{\varepsilon_r}=+1$ for r such that $i(r) \in D_2 \wedge \overset{*}{Y}_i=0$, and --- $I_{\varepsilon_r}=0$ for r such that $i(r) \notin D_2$.

In a similar way, deviation $\tilde{d}_B$ solves the -- problem

$$\begin{aligned}\overset{*}{B}(\overset{*}{X}_B+\tilde{d}_B) &= \overset{*}{b}-\overset{**}{S}\overset{*}{X}_S-\overset{**}{N}\overset{*}{X}_N+\varepsilon_2 I_\varepsilon \\ &\equiv \overset{*}{b}_B+\varepsilon_2 I_\varepsilon \qquad (6.6)\end{aligned}$$

such that $\overset{*}{B}_{rc}$ takes the (r,c)-th element of matrix $\overset{*}{B}$ for $r=1,\dots,\overset{*}{t}$ and $c=1,\dots,\overset{*}{t}$. Let -- $i(r) \in \overset{*}{W}$ define the constraint related to in-- dex r in (6.6); similarly, j(c) defines the basic variable related to index c in (6.6). Vector I_ε in (6.6) is defined as in (6.5).

Note that constraint $i \in M-E \cap \overset{*}{W}$ will be only perturbed if its Lagrange multiplier estimate $\overset{*}{\mu}_{B_i}$ is zero or near-to-zero for $\gamma=0$, or its Lagrange multiplier estimate $\overset{*}{\mu}_{BS_i}$ is zero or near-to-zero and $\overset{*}{\mu}_{B_i}$ is not non-favorable (i.e., it is favorable or, at least, zero or near-to-zero) for $\gamma=1$.

For obtaining the perturbed gradient $\tilde{g}_B$ in - (6.4) and $\tilde{g}_{BS}$ in (6.3) such that the above - approach be practical, it is required a fast procedure for obtaining $\tilde{d}_B$ in (6.6) and $\tilde{d}_{BS}$ in (6.5).

<u>Case $\gamma=0$.</u>

The deviation $\tilde{d}_B$ that satisfies (6.6) for -- perturbation ε_2 in $\overset{*}{b}_{B_r}$ for $i(r) \in D_2$ is such - that

$$\begin{aligned}\tilde{d}_B &= \varepsilon_2 \overset{*}{B}^{-1} I_\varepsilon \\ &\equiv \varepsilon_2 \Big(\sum_{r | i(r) \in D_2 \wedge \overset{*}{Y}_i=0} \overset{*}{B}_r^{-1} - \\ &\quad - \sum_{r | i(r) \in D_2 \wedge \overset{*}{Y}_i=\bar{b}_i-\underline{b}_i} \overset{*}{B}_r^{-1} \Big) \qquad (6.7)\end{aligned}$$

where $\overset{*}{B}_r^{-1}$ takes the r-th column in matrix -- $\overset{*}{B}^{-1}$, such that $\tilde{d}_{B_c}$ takes the deviation from the solution $\overset{*}{X}_{j(c)}$. Note that inverse basic - matrix $\overset{*}{B}^{-1}$ is not explicitly calculated, since its triangular factors $\overset{*}{L}$ (lower) and $\overset{*}{U}$ --- (upper) are fresh anew or kept updated in the previous iterations /2/. In any case, $\overset{*}{B}^{-1}$ is postmultiplied by a vector for solving (6.7) and premultiplied by a vector for solving --- (6.4).

<u>Case $\gamma=1$</u>

The deviation $\tilde{d}_{BS}$ that satisfies (6.5) for an small enough perturbation ε_2 in $\overset{*}{b}_{BS_r}$ for ---- $i(r) \in D_2$ can be written

$$\tilde{d}_{BS} = \varepsilon_2 (\overset{**}{BS})^{+t} I_\varepsilon \qquad (6.8)$$

where

$$(\overset{**}{BS})^+ \equiv (\overset{**}{BS}(\overset{**}{BS})^t)^{-1} \overset{**}{BS} \qquad (6.9)$$

is the $\overset{*}{t}.(n-\overset{*}{t}-\overset{*}{r})$ <u>pseudo-inverse</u> matrix of the $(n-\overset{*}{t}-\overset{*}{r}).\overset{*}{t}$ matrix $(\overset{**}{BS})^t$.

In effect /4,2/, from (6.5) it results

$$\overset{**}{BS}(\overset{*}{X}_{BS}+\tilde{d}_{BS}) = \overset{*}{b}_{BS}+\varepsilon_2 I_\varepsilon$$

$$= \overset{*}{b}_{BS}+\varepsilon_2\overset{**}{BS}(\overset{**}{BS})^t(\overset{**}{BS}(\overset{**}{BS})^t)^{-1}\ I_\varepsilon$$

$$= \overset{*}{b}_{BS}+\overset{**}{BS}\varepsilon_2(\overset{**}{BS})^{+t}I_\varepsilon$$

$$= \overset{**}{BS}(\overset{*}{X}_{BS}+\varepsilon_2(\overset{**}{BS})^{+t}I_\varepsilon \qquad (6.10)$$

and, finally,

$$\tilde{d}_{BS} = \varepsilon_2\ \Big(\sum_{r|i(r)\in D_2\wedge\overset{*}{Y}_i=0} (\overset{**}{BS})^{+t}_r \ -$$

$$-\sum_{r|i(r)\in D_2\wedge\overset{*}{Y}_i=\bar{b}_i-\underline{b}_i} (\overset{**}{BS})^{+t}_r\Big) \qquad (6.11)$$

where $(\overset{**}{BS})^+_r$ takes the r-th row in matrix -- $(\overset{**}{BS})^+$, such that $\tilde{d}_{BS_c}$ takes the deviation from the solution $\overset{*}{X}_{j(c)}$ for $j(c)\in J/\overset{*}{V}$.

Since $(\overset{**}{BS})^t = \overset{**}{QR}$ (4.8) and $(\overset{**}{BS})^+$ can be expressed by (6.9), it results

$$\overset{*}{R}(\overset{**}{BS})^+ = \overset{*}{Q}^t \qquad (6.12)$$

such that the $\overset{*}{t}$-vector $(\overset{**}{BS})^+_c$ for $c=1,\dots,n-\overset{*}{t}-\overset{*}{r}$ can be written

$$(\overset{**}{BS})^+_{\overset{*}{t}c} = \overset{*}{Q}_{c\overset{*}{t}} \qquad (6.13a)$$

$$(\overset{**}{BS})^+_{rc} = \overset{*}{Q}_{cr} - \sum_{\ell=r+1}^{\overset{*}{t}} \overset{*}{R}_{r\ell}(\overset{**}{BS})^+_{\ell c} \quad \text{for}$$

$$r=\overset{*}{t}-1,\dots,1 \qquad (6.13b)$$

Note that not all rows $(\overset{**}{BS})^+_r$ are required, - but only rows $r=1,\dots,\overset{*}{t}$ such that $i(r)\in D_2$; - then, it could be possible that matrix $(\overset{**}{BS})^+$ is not required to be completely calculated.

Once obtained $\overset{*}{X}_B+\tilde{d}_B$ for $\gamma=0$ and $\overset{*}{X}_{BS}+\tilde{d}_{BS}$ -- for $\gamma=1$, perturbed estimates $\tilde{\mu}_B$ and $\tilde{\mu}_{BS}$ are obtained from (6.4) and (6.3), respectively. Setting $\tilde{\mu}_{BS}=\tilde{\mu}_B$ for $\gamma=0$, the first active --- inequality constraint i D_2 for which the --- following condition holds is to be deactivated (and, then, set k=i).

$$(\tilde{\mu}_{BS_i}-\overset{*}{\mu}_{BS_i}<-\varepsilon_3|\overset{*}{Y}_i=0)\vee(\tilde{\mu}_{BS_i}-\overset{*}{\mu}_{BS_i}>\varepsilon_3|\overset{*}{Y}_i=\bar{b}_i-\underline{b}_i) \qquad (6.14)$$

such that a priority is given for $\gamma=1$ to the constraint with the maximum absolute value in its estimate $\overset{*}{\mu}_{B_i}$ if it is favorable.

If k=0 it seems that the tendency on $\overset{*}{\mu}_{BS}$ is not favorable; the action to be taken depends on the character of $\overset{*}{X}$. If it is a 'quasi-optimal' solution in the manifold $\overset{*}{W}$ and $\overset{*}{V}$, the -- next iteration obtains the related superbasic stepdirection. If $\overset{*}{X}$ is an optimal point in -- that manifold, it is also assumed that it is an optimal point in problem (1.1)-(1.2).

7. CONCLUSION

The influence of the degenerate sets of active inequality constraints and bounds in the optimality conditions has been analyzed for local optimal points in linearly constrained nonlinear programming (LCNP) problems. Some procedures have been described to get some insight on the tendency of zero or near-to-zero Lagrange multipliers estimates for non-basic (structural and slack) variables in -- the frame of a given LCNP algorithm.

8. REFERENCES

/1/ ESCUDERO, L.F., "An implementation of -- the QR factorization for solving overdetermined systems of linear equations". QUESTIIO 4 (1980) 89-94.

/2/ ESCUDERO, L.F., "An algorithm for large-scale quadratic programming problems and its extensions to the linearly constrained case", IBM Madrid Scientific Center, report SCR-01.81, Madrid, 1981.

/3/ ESCUDERO, L.F., "Lagrange multipliers estimates for constrained minimization", QUESTIIO 5 (1981) 173-186.

/4/ ESCUDERO, L.F., "QR factorization and -- its updatings", IBM Madrid Scientific Center, report SCR-01.82, Madrid, 1982. See also, "On QR factorization updatings," -- Trabajos de Estadística e Investigación Operativa (in printing).

/5/ ESCUDERO, L.F., "Dirección de búsqueda - en programación nolineal en gran escala con condiciones lineales y escasa densidad . Método Newton-Truncado reducido," Anales de Ingenieria Mecanica 1 (1982).

See also "On diagonally-preconditioning the Truncated-Newton method for superscale linearly constrained nonlinear - programming, IBM Madrid Scientific Center, report SCR-04-82, Madrid, 1982.

/6/ ESCUDERO, L.F., "Programación matemática", División de Matemáticas, Facultad de Ciencias, Universidad Autónoma de Madrid, Madrid, 1982.

/7/ ESCUDERO, L.F., "Strategies selection on linearly constrained nonlinear programming" IBM Madrid Scientific Center, re-- port SCR-02.82, Madrid, 1982.

/8/ Fletcher, R., "Practical methods for optimization (Wiley, London, 1981).

/9/ GILL, P.E., and MURRAY, W., "Numerical - methods for constrained optimization" -- (Academic Press, London, 1974).

/10/ GILL, P.E. and MURRAY, W., "The computation of Lagrange multipliers estimates for constrained minimization", Mathematical Programming 17 (1979) 32-60.

/11/ GILL, P.E., MURRAY, W. and WRIGHT,M.H., "Practical optimization"(Academic Press, London, 1981).

/12/ MANGASARIAN, O., "Nonlinear programming" (McGraw-Hill, New York, 1969).

/13/ Murtagh, B.A. and SAUNDERS, M.A.,"Large-scale linearly constrained optimization" Mathematical Programming 14 (1978) 41-72.