Perturbation Approach to Sensitivity Analysis in Mathematical Programming[1]

E. Castillo[2], A. J. Conejo[3], C. Castillo[4], R. Mínguez[5] and D. Ortigosa[6]

Communicated by M. A. Simmaan

[1]The authors are indebted to the Ministry of Science and Education of Spain, Projects CICYT DPI2002-04172-C04-02 and CICYT DPI2003-01362, and to the Fulbright Commission for partial support. The authors are also grateful to the referees for their comments, which allowed an improvement of the quality of the paper.

[2]Professor, Department of Applied Mathematics and Computational Sciences, University of Cantabria, Santander, Spain.

[3]Professor, Department of Electrical Engineering, University of Castilla-La Mancha, Ciudad Real, Spain.

[4]Assistant Professor, Grupo de Puertos y Costas, CEAMA, University of Granada, Granada, Spain.

[5]Assistant Professor, Department of Applied Mathematics, University of Castilla-La Mancha, Ciudad Real, Spain; and Visiting Scientist, Department of Social Statistics, Cornell University, Ithaca, New York, USA.

[6]Ph.D. Student, Department of Mathematics, University of La Rioja, Logroño, Spain.

Abstract

This paper presents a perturbation approach for performing sensitivity analysis of mathematical programming problems. Contrary to standard methods, the active constraints are not assumed to remain active if the problem data are perturbed, nor the partial derivatives are assumed to exist. In other words, all the elements, variables, parameters, Karush–Kuhn–Tucker multipliers and the objective function values may vary provided that optimality is maintained, and the general structure of a feasible perturbation (which is a polyhedral cone) is obtained. This allows determining: (a) the local sensitivities, (b) whether or not partial derivatives exist, and (c) if the directional derivative for a given direction exists. A method for the simultaneous obtention of the sensitivities of the objective function optimal value and the primal and dual variable values with respect to data is given. Three examples illustrate the concepts presented and the proposed methodology. Finally, some relevant conclusions are drawn.

Key Words: Local sensitivity, mathematical programming, duality, polyhedral cone.

1 Introduction

Sensitivity analysis consists of determining "how" and "how much" specific changes in the parameters of an optimization problem influence the optimal objective function value and the point (or points) where the optimum is attained.

The problem of sensitivity analysis in nonlinear programming has been discussed by several authors, as, for example, Vanderplaats (Ref. 1), Sobiesky et al. (Ref. 2), Enevoldsen (Ref. 3), Roos, Terlaky and Vial (Ref. 4), Bjerager and Krend (Ref. 5), etc. There are at least three ways of deriving equations for the unknown sensitivities: (a) the Lagrange multiplier equations of the constrained optimum (see Sobiesky et al. (Ref. 2)), (b) differentiation of the Karush–Kuhn–Tucker conditions to obtain the sensitivities of the objective function with respect to changes in the parameters (see Vanderplaats (Ref. 1), Sorensen and Enevoldsen (Ref. 6) or Enevoldsen (Ref. 3)), and (c) the extreme conditions of a penalty function (see Sobiesky et al. (Ref. 2)).

The existing methods for sensitivity analysis may present four main limitations:

(i) They provide the sensitivities of the objective function value and the primal variables values with respect to parameters, but not the sensitivities of the dual variables with respect to parameters.

(ii) For different cases there are diverse methods for obtaining each of the sensitivities (optimal objective function value or primal variable values with respect to parameters), but there is no integrated approach providing all the sensitivities at once.

(iii) They assume the existence of partial derivatives of the objective function or the optimal solutions with respect to the parameters; however, this is not always the case. In fact, there are cases in which partial or directional derivatives fail to exist. In addition, most techniques reported in the literature do not distinguish between right and left derivatives. Ross, Terlaky and Vial (Ref. 4) state: "It is surprising that in the literature on sensitivity analysis it is far from common to distinguish between left- and right-shadow prices". By left- and right-shadow prices they mean left- and right-derivatives of the objective function with respect to parameters at the current optimal value.

(iv) They assume that the active constraints remain active, which implies that there is no need to distinguish between equality or inequality constraints, because all the active constraints can be considered as equality constraints, and inactive constraints will remain inactive for small changes in the parameters. As a consequence, the set of possible changes (perturbations) has (locally) the structure of a linear space.

The aim of this paper is twofold: (a) to perform a general analysis, without assuming neither the existence of partial derivatives, nor that the active inequality constraints remain active, and (b) to give an integrated approach that gives all sensitivities (objective function value, primal and dual variables values with respect to parameters) at once. As a consequence, a distinction between equality and inequality constraints is necessary, because the active inequality constraints may become inactive, but equality constraints are always active. Since we deal with local sensitivity in this paper, we do not need to consider inactive constraints because if they are inactive, they will remain inactive after a differential change. So, in what follows we could consider only active inequality constraints.

The above considerations lead to a set of feasible changes that has the structure of a cone for the most general case. Furthermore, this analysis allows determining whether or not partial or directional derivatives with respect to parameters exist.

This paper is structured as follows. In Section 2 we derive the general structure of feasible changes for a Karush–Kuhn–Tucker solution to remain a Karush–Kuhn–Tucker solution. In Section 3 a method for determining the set of all feasible perturbations is provided. In Section 4 a method for dealing with directional and partial derivatives is introduced. In Section 5 one regular non-degenerate, one regular degenerate and one non-regular illustrative examples are presented (regularity and degeneracy are precisely defined below). In Section 6 some conclusions are provided.

2 Sensitivity Analysis

In this section we analyze the sensitivity of the optimal solution of a nonlinear programming problem to changes in the data values. Many authors, as those already mentioned, have studied different versions of this problem. Some of them have dealt with the linear programming problem and discussed the effect of changes of (i) the cost coefficients, (ii) the right hand sides of the constraints or (iii) the constraint coefficients on either (a) the optimal value of the objective function or (b) the optimal solution. A similar analysis has been done for nonlinear problems. However, these authors have dealt only with changes that keep invariant the set of active constraints.

In the following, we deal with the calculation of partial derivatives with respect to parameters without forcing the set of active constraints to remain active.

It is relevant to note that to perform a local sensitivity analysis, the objective function (1) and the nonlinear constraints in (2)-(3) of the general problem below can be replaced by the corresponding quadratic approximations that share tangent hyperplanes and Hessian matrices, at the solution point.

Consider the following Nonlinear Programming Problem (NLPP):

$$\min_{x} \quad z = f(x, a), \tag{1}$$

$$\text{s.t.} \quad h(x, a) = 0, \tag{2}$$

$$g(x, a) \leq 0, \tag{3}$$

where $f : \mathbb{R}^n \times \mathbb{R}^p \to \mathbb{R}$, $h : \mathbb{R}^n \times \mathbb{R}^p \to \mathbb{R}^\ell$, $g : \mathbb{R}^n \times \mathbb{R}^p \to \mathbb{R}^m$ with $h(x, a) = (h_1(x, a), \ldots, h_\ell(x, a))^T$ and $g(x, a) = (g_1(x, a), \ldots, g_m(x, a))^T$ are functions over the feasible region $S(a) = \{x | h(x, a) = 0, g(x, a) \leq 0\}$ and $f, h, g \in C^2$.

Let x^* be a local solution of problem (1)-(3) and a regular point of the constraints. If J is the set of indices j for which $g_j(x^*, a) = 0$, a local solution x^* is a regular point of the constraints $h(x, a) = 0$ and $g(x, a) \leq 0$ if the gradient vectors $\nabla_x h_k(x^*, a)$, $\nabla_x g_j(x^*, a)$, $k = 1, \ldots, \ell; j \in J$ are linearly independent. Note that other regularization conditions are possible, see Luenberger (Ref. 7).

Then, there exists a pair of vectors $\lambda^* \in \mathbb{R}^\ell$ and $\mu^* \in \mathbb{R}^m$ such that (Bazaraa, Sherali and Shetty (Ref. 8) or Luenberger (Ref. 7)):

$$\nabla_x f(x^*, a) + \sum_{k=1}^{\ell} \lambda_k^* \nabla_x h_k(x^*, a) + \sum_{j=1}^{m} \mu_j^* \nabla_x g_j(x^*, a) = 0_n \tag{4}$$

$$h_k(x^*, a) = 0; \quad k = 1, 2, \cdots, \ell, \tag{5}$$

$$g_j(x^*, a) \leq 0; \quad j = 1, 2, \cdots, m, \tag{6}$$

$$\mu_j^* \, g_j(x^*, a) \;=\; 0; \;\; j = 1, 2, \ldots, m, \tag{7}$$

$$\mu^* \;\geq\; 0_m \tag{8}$$

As is well known, the vectors λ^* and μ^* are called the *Karush–Kuhn–Tucker multipliers.* Conditions (4)-(8) are denominated Karush-Kuhn-Tucker (KKT) conditions. Conditions (5)–(6) are called *the primal feasibility* conditions. Condition (7) is known as the *complementary slackness condition.* Condition (8) requires the nonnegativity of the multipliers of the inequality constraints, and is referred to as the *dual feasibility conditions.*

Furthermore, the Hessian of the Lagrangian at x^*, μ^* and λ^*

$$\nabla_{xx} f(x^*, a) + \sum_{k=1}^{\ell} \lambda_k^* \, \nabla_{xx} h_k(x^*, a) + \sum_{j=1}^{m} \mu_j^* \, \nabla_{xx} g_j(x^*, a) \tag{9}$$

is assumed to be positive definite on the subspace that is orthogonal to the subspace spanned by the gradients of the constraint functions.

Then, for λ and $\mu \geq 0$ near λ^* and μ^*, the dual function is defined as

$$\phi(\lambda, \mu) = \underset{x}{\text{minimum}} \; \left[f(x, a) + \sum_{k=1}^{\ell} \lambda_k \, h_k(x, a) + \sum_{j=1}^{m} \mu_j \, g_j(x, a) \right] \tag{10}$$

where the minimum is taken locally near x^*, and the dual problem is

$$\max_{\lambda, \mu \geq 0} \quad \phi(\lambda, \mu) \tag{11}$$

whose solution is λ^*, μ^*.

We aim at determining the sensitivity of the optimal solution (x^*, λ^*, μ^*, z^*) of (4) - (8) to changes in the parameters, i.e., we perturb or modify x^*, a, λ^*, μ^*, z^* in such a way that the KKT conditions still hold. Thus, to obtain the sensitivity equations we differentiate (1) and (4)-(8), as follows:

$$
\begin{aligned}
(\nabla_x f(x^*,a))^T dx + (\nabla_a f(x^*,a))^T da - dz &= 0 && (12)\\
\left(\nabla_{xx} f(x^*,a) + \sum_{k=1}^{\ell} \lambda_k^* \nabla_{xx} h_k(x^*,a) + \sum_{j=1}^{m} \mu_j^* \nabla_{xx} g_j(x^*,a)\right) dx & \\
+\left(\nabla_{xa} f(x^*,a) + \sum_{k=1}^{\ell} \lambda_k^* \nabla_{xa} h_k(x^*,a) + \sum_{j=1}^{m} \mu_j^* \nabla_{xa} g_j(x^*,a)\right) da & \\
+\nabla_x h(x^*,a) d\lambda + \nabla_x g(x^*,a) d\mu &= 0_n && (13)
\end{aligned}
$$

$$
\begin{aligned}
(\nabla_x h(x^*,a))^T dx + (\nabla_a h(x^*,a))^T da &= 0_\ell && (14)\\
(\nabla_x g_j(x^*,a))^T dx + (\nabla_a g_j(x^*,a))^T da &= 0, \text{ if } \mu_j^* \neq 0, j \in J && (15)\\
(\nabla_x g_j(x^*,a))^T dx + (\nabla_a g_j(x^*,a))^T da &\leq 0, \text{ if } \mu_j^* = 0, j \in J && (16)\\
-d\mu_j &\leq 0, \text{ if } \mu_j^* = 0, j \in J && (17)\\
d\mu_j \left[(\nabla_x g_j(x^*,a))^T dx + (\nabla_a g_j(x^*,a))^T da\right] &= 0, \text{ if } \mu_j^* = 0, j \in J && (18)
\end{aligned}
$$

where all the matrices are evaluated at the optimal solution, and redundant constraints have been removed. More precisely, the constraints (15)-(18) are simplifications of the constraints that result directly from differentiating (6)- (8), i.e., from

$$(\nabla_x g_j(x^*,a))^T dx + (\nabla_a g_j(x^*,a))^T da \leq 0, \quad j \in J, \qquad (19)$$

and

$$(\mu_j^* + d\mu_j)\left(g_j(x^*, a) + dg_j(x^*, a)\right) = \mu_j^* dg_j(x^*, a) + d\mu_j\left(g_j(x^*, a) + dg_j(x^*, a)\right), \quad j \in J. \tag{20}$$

Since all these inequality constraints are active, we have $g_j(x^*, a) = 0;\ \forall j \in J$ and then (20) results in (15) for $\mu_j^* \neq 0$, and in (18) for $\mu_j^* = 0$.

Finally, since (15) implies (19), for $\mu_j^* \neq 0$, (19) must be written only for $\mu_j^* = 0$, i.e., (16).

Note that constraint (15) forces the constraints $g_j(x^*, a) = 0$ whose multipliers are different from zero ($\mu_j^* \neq 0$) to remain active, constraint (16) allows the optimal point to move inside the feasible region, constraint (17) forces the Lagrange multipliers to be greater or equal to zero, and (18) forces the new point to hold the *complementary slackness condition* for $\mu_j^* = 0$. This last constraint is a second order constraint that implies that one of the constraints (16) or (17) has to be an equality constraint.

In matrix form, the system (12)-(17) can be written as:

$$M\delta p = \left[\begin{array}{c|c|c|c|c} F_x & F_a & 0 & 0 & -1 \\ \hline F_{xx} & F_{xa} & H_x^T & G_x^T & 0 \\ \hline H_x & H_a & 0 & 0 & 0 \\ \hline G_x^1 & G_a^1 & 0 & 0 & 0 \end{array}\right] \left[\begin{array}{c} dx \\ \hline da \\ \hline d\lambda \\ \hline d\mu \\ \hline dz \end{array}\right] = 0 \tag{21}$$

$$N\delta p = \left[\begin{array}{c|c|c|c|c} G_x^0 & G_a^0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & -I_{m_J}^0 & 0 \end{array}\right] \delta p \leq 0 \tag{22}$$

where $m_J = \text{card}(J)$ is the number of active inequality constraints and the meaning of matrices M and N becomes clear from the system (12)-(17), and the submatrices are defined below (corresponding dimensions in parenthesis)

$$
\begin{aligned}
F_{x(1\times n)} &= \left(\nabla_x f(x^*, a)\right)^T, && (23)\\
F_{a(1\times p)} &= \left(\nabla_a f(x^*, a)\right)^T, && (24)\\
F_{xx(n\times n)} &= \nabla_{xx} f(x^*, a) + \sum_{k=1}^{\ell} \lambda_k^* \nabla_{xx} h_k(x^*, a) + \sum_{j=1}^{m_J} \mu_j^* \nabla_{xx} g_j(x^*, a), && (25)\\
F_{xa(n\times p)} &= \nabla_{xa} f(x^*, a) + \sum_{k=1}^{\ell} \lambda_k^* \nabla_{xa} h_k(x^*, a) + \sum_{j=1}^{m_J} \mu_j^* \nabla_{xa} g_j(x^*, a), && (26)\\
H_{x(\ell\times n)} &= \left(\nabla_x h(x^*, a)\right)^T, && (27)\\
H_{a(\ell\times p)} &= \left(\nabla_a h(x^*, a)\right)^T, && (28)\\
G_{x(m_J\times n)} &= \left(\nabla_x g(x^*, a)\right)^T, && (29)\\
G_{a(m_J\times p)} &= \left(\nabla_a g(x^*, a)\right)^T, && (30)
\end{aligned}
$$

where G_x^0 and G_a^0 refer to the submatrices of G_x and G_a, respectively, associated with the null μ-multipliers of active constraints, G_x^1 and G_a^1 refer to the submatrices of G_x and G_a, respectively, associated with the non-null μ-multipliers of active constraints, and $-I_{m_J}^0$ is the negative of a unit matrix after removing all rows $j \in J$ such that $\mu_j^* \neq 0$.

In order to consider the second order condition (18) the system (21)-(22) has to be modified extracting from (22) and adding to (21) the row associated with either the term

G^0 or $-I^0_{m_J}$ for each constraint such that $\mu^*_j = 0;\ j \in J$. The interpretation is simple, we add into (21) the term related to G^0 for the constraints we want to remain active after the perturbation, or the term associated with $-I^0_{m_J}$ for the constraints we want to allow to become inactive. Note that 2^{m_0} combinations (systems) are possible, where m_0 in the number of constraints whose $\mu^*_j = 0$. In what follows we initially consider the system (21)-(22) and later we take into account (18).

3 Determining the Set of All Feasible Perturbations

Conditions (21)-(22) define the set of feasible perturbations $\delta p = (dx, da, d\lambda, d\mu, dz)^T$, i.e., for moving from one KKT solution to another KKT solution.

Since (21)-(22) constitute an homogeneous linear system of equalities and inequalities in $dx, da, d\lambda, d\mu$ and dz, its general solution is a polyhedral cone (see Padberg (Ref. 9), Castillo, Cobo et al. (Ref. 10) and Castillo, Jubete et al. (Ref. 11)):

$$\delta p = \sum_{i=1}^{t} \rho_i v_i + \sum_{j=1}^{q} \pi_j w_j, \tag{31}$$

where $\rho_i \in \mathbb{R}; i = 1, 2, \cdots, t$, and $\pi_j \in \mathbb{R}^+; j = 1, 2, \cdots, q$, and v_i and w_j are vectors that generate the linear space and the proper cone parts of the polyhedral cone, respectively.

It should be noted that since a linear space is a particular case of a cone, one can obtain a linear space as the solution of a homogeneous system of linear inequalities.

The vertex cone representation (31) of the feasible perturbations can be obtained using the Γ-algorithm (see Padberg (Ref. 9) and Castillo et al. (Ref. 12)) that is known to be computational intensive for large problems. However, one can obtain first the solution of (21) (the corresponding null space), and then use the Γ-algorithm to incorporate the constraints in (22), which are only a reduced number (active inequality constraints with null μ-multipliers) or none. Note that the null space computation is a standard procedure whose associated computational burden is similar to that of solving a linear homogeneous system of N equations, $O(N^3)$ (Ref. 13).

Nevertheless, as we shall see, the obtention of the vertex cone representation (31), though convenient, could be unnecessary.

Once (31) is known, all feasible perturbations become available. Note that if we want to take into account (18) all possible combinations of the system (21)-(22) must be solved so that several solutions (31) exist. Any selection of $\rho_i \in \quad i = 1, 2, \ldots, t$ and $\pi_j \in \mathbb{R}^+; j = 1, 2, \ldots, q$ in any solution leads to a feasible perturbation and all of them can be obtained in this form.

4 Discussion of Directional and Partial Derivatives

Conditions (21)-(22) can be written as

$$U\left[\begin{array}{c|c|c|c} dx & d\lambda & d\mu & dz \end{array}\right]^T = Sda \quad (32)$$

$$V\left[\begin{array}{c|c|c|c} dx & d\lambda & d\mu & dz \end{array}\right]^T \leq Tda \quad (33)$$

where the matrices U, V, S and T are:

$$U = \left[\begin{array}{c|c|c|c} F_x & 0 & 0 & -1 \\ \hline F_{xx} & H_x^T & G_x^T & 0 \\ \hline H_x & 0 & 0 & 0 \\ \hline G_x^1 & 0 & 0 & 0 \end{array}\right], \quad S = -\left[\begin{array}{c} F_a \\ \hline F_{xa} \\ \hline H_a \\ \hline G_a^1 \end{array}\right], \quad (34)$$

$$V = \left[\begin{array}{c|c|c|c} G_x^0 & 0 & 0 & 0 \\ \hline 0 & 0 & -I_{m_J}^0 & 0 \end{array}\right], \quad T = -\left[\begin{array}{c} G_a^0 \\ \hline 0 \end{array}\right]. \quad (35)$$

Note that as system (32)-(33) comes from (21)-(22) and due to condition (18), several systems (32)-(33) corresponding to the different combinations may exist.

An optimal point (x^*, λ^*, μ^*, z^*) can be classified as follows:

Regular Point: The solution (x^*, λ^*, μ^*, z^*) is a regular point if the gradient vectors of the active constraints are linearly independent. Under this circumstance, the optimal point can be nondegenerate or degenerate:

(i) **Nondegenerate:** The Lagrange multipliers μ^* of active inequality constraints are different from zero, there is no matrix V and U^{-1} exists.

(ii) **Degenerate:** The Lagrange multipliers μ^* of active inequality constraints are different from zero, there is no matrix V and U^{-1} does not exist. Alternatively, some of the Lagrange multipliers multipliers of active inequality constraints in μ^* are equal to zero and matrix U^{-1} does not exist because U is not a square matrix.

Nonregular Point: The gradient vectors of the active constraints are linearly dependent. Note that the KKT conditions do not characterize adequately this case because there are infinite Lagrange value combinations that hold. However, the method also provides the sensitivities for given values of the Lagrange multipliers. In this case no difference is made between non-degenerate and degenerate cases because matrix U is never invertible.

Note that the most common situation occurs when we have a regular non-degenerate point. The cases of regular degenerate and non-regular points are exceptional. However, since we deal with a set of parametric optimization problems (we use parameters a), normally there exist particular values for the parameters such that these two cases occur as important transition situations.

Expressions (32) and (33) allow determining: (a) Directional derivatives if they exist. (b) Partial derivatives if they exist. (c) All partial derivatives at once when they exist.

Note that existence means that there is a feasible perturbation where the KKT conditions still hold. We deal with all these problems in the following subsections.

4.1 Determining Directional Derivatives

To check if a directional derivative exists, we replace da by the corresponding unit vector and solve all possible combinations of the system (32)-(33). If it exists (existence) at least for one of the combinations and the solution is unique (uniqueness) the directional derivative exists.

One can obtain first the solution of (32) (the corresponding null space), and then use the Γ-algorithm to incorporate the constraints in (33), which are only a reduced number (active inequality constraints with null μ-multipliers).

4.2 Partial Derivatives

A partial derivative is a special case of directional derivative. The partial derivative of u with respect to a_k means the increment in u due to a unit increment in a_k and null increments in $a_r, r \neq k$. Then, in a feasible perturbation δp that contains a unit component da_k together with null values for components $da_i, \forall i \neq k$, the remaining perturbation components contain the corresponding right-derivatives (sensitivities) with respect to a_k, that is:

$$\delta p = \left(\frac{dx_1}{da_k^+}, \ldots, \frac{dx_n}{da_k^+}, 0, \ldots, 0, 1, 0, \ldots, 0, \frac{d\lambda_1}{da_k^+}, \ldots, \frac{d\lambda_p}{da_k^+}, \frac{d\mu_1}{da_k^+}, \ldots, \frac{d\mu_{m_J}}{da_k^+}, \frac{dz}{da_k^+} \right)^T \tag{36}$$

Similarly, a feasible perturbation of the form

$$\delta p = \left(\frac{dx_1}{da_k^-}, \ldots, \frac{dx_n}{da_k^-}, 0, \ldots, 0, -1, 0, \ldots, 0, \frac{d\lambda_1}{da_k^-}, \ldots, \frac{\lambda_p}{da_k^-}, \frac{d\mu_1}{da_k^-}, \ldots, \frac{d\mu_{m_J}}{da_k^-}, \frac{dz}{da_k^-} \right)^T \tag{37}$$

contains as the remaining components all the left-derivatives with respect to a_k. If both exist, and coincide in absolute value but not in sign, the corresponding partial derivative exists.

The partial derivative is obtained solving the directional derivatives for da_k and $-da_k$, respectively, and checking if both exist, and coincide in absolute value but not in sign. If the answer is positive the corresponding partial derivative exists.

Note that this procedure also allows to know if there are directional derivatives for any arbitrary vector da in both directions da and $-da$.

4.3 Obtaining All Sensitivities at Once

If the solution $(x^*,\ \lambda^*,\ \mu^*,\ z^*)$ is a nondegenerate regular point, then the matrix U is invertible and the system (32)-(33) is unique and it becomes

$$\left[\ dx \mid d\lambda \mid d\mu \mid dz\ \right]^T = U^{-1} S\ da\ , \tag{38}$$

where (33) is satisfied trivially since V does not exist.

Several partial derivatives can be simultaneously obtained if the vector da in (38) is replaced by a matrix including several vectors (columns) with the corresponding unit directions. In particular, replacing da by the unit matrix I_p in (38) all the partial derivatives are obtained. The matrix with all partial derivatives becomes:

$$\left[\begin{array}{cccc} dx/da & d\lambda/da & d\mu/da & dz/da \end{array}\right]^T = U^{-1}S \tag{39}$$

For any vector da the derivatives in both directions da and $-da$ are obtained simultaneously.

In some particular cases the system (38) can be easily solved by decomposition. For example, if H_x or G_x are square invertible matrices, one gets $dx = H_x^{-1}H_a da$ or $dx = (G_x)^{-1} G_a da$, respectively and one can proceed to solve $d\lambda, d\mu$ and dz, using the remaining equations.

It should be noted that the previous study has been done assuming that one is interesting in calculating the directional or partial derivatives of x, λ, μ and z with respect to a. However, one can think of calculating the derivatives with respect to x, λ, μ or z, of the corresponding variables, or even the derivatives with respect to a combination of components of x, a, λ, μ and z, of the remaining components. Thus, the applicability of the above relations is much more important that this simple example of directional or partial derivatives. However, a detailed analysis of all these possibilities is out of the scope of this paper.

5 Illustrative Examples

In this section we illustrate the theory developed in Sections 2, 3 and 4 by its application to a regular non-degenerate (it is the most common case), a regular degenerate and a non-regular examples.

5.1 Regular Nondegenerate Example

In this section we apply the above method to the problem of estimating the parameters of a uniform distribution based on a sample using the method of moments with constraints.

Consider the uniform random variable family with densities of the form $f(y; a, b) = 1/(b-a);\ \ a \le y \le b$, with mean $(a+b)/2$ and variance $(b-a)^2/12$.

To estimate the parameters a and b based on a random sample, we use the constrained method of moments, that consists of solving the optimization problem:

$$\min_{a,b} \quad z = \left((a+b)/2 - \bar{y}\right)^2 + \left((b-a)^2/12 - \sigma^2\right)^2 \tag{40}$$

$$\text{s. t.} \quad a - y_{\min} \le 0 : \ \mu_1 \tag{41}$$

$$y_{\max} - b \le 0 : \ \mu_2 \tag{42}$$

where $\bar{y}$ and σ^2 are the sample mean and variance, respectively, μ_1 and μ_2 are the corresponding dual variables, and $y_{\min}$ and $y_{\max}$ are the minimum and maximum values of the sample, respectively.

In this case, the system (12)-(17) becomes:

$$\begin{aligned}
0 &= \left[\frac{a+b}{2} - \bar{y} - \frac{b-a}{3}\left(\frac{(b-a)^2}{12} - \sigma^2\right)\right] da + \left[\frac{a+b}{2} - \bar{y} + \frac{b-a}{3}\left(\frac{(b-a)^2}{12} - \sigma^2\right)\right] db \\
&\quad -\frac{1}{n}\sum_{i=1}^{n}\left[a+b-2\bar{y}+4(y_i-\bar{y})\left(\frac{(b-a)^2}{12} - \sigma^2\right)\right] dy_i - dz \qquad (43) \\
0 &= \left[\frac{1}{2} + \frac{(b-a)^2}{18} + \frac{1}{3}\left(\frac{(b-a)^2}{12} - \sigma^2\right)\right] da + \left[\frac{1}{2} - \frac{(b-a)^2}{18} - \frac{1}{3}\left(\frac{(b-a)^2}{12} - \sigma^2\right)\right] db \\
&\quad -\frac{1}{n}\sum_{i=1}^{n}\left(1 - \frac{2(b-a)(y_i-\bar{y})}{3}\right) dy_i + d\mu_1 \qquad (44) \\
0 &= \left[\frac{1}{2} - \frac{(b-a)^2}{18} - \frac{1}{3}\left(\frac{(b-a)^2}{12} - \sigma^2\right)\right] da + \left[\frac{1}{2} + \frac{(b-a)^2}{18} + \frac{1}{3}\left(\frac{(b-a)^2}{12} - \sigma^2\right)\right] db \\
&\quad -\frac{1}{n}\sum_{i=1}^{n}\left(1 + \frac{2(b-a)(y_i-\bar{y})}{3}\right) dy_i - d\mu_2 \qquad (45) \\
0 &= -db + dy_n \qquad (46)
\end{aligned}$$

Note that data is ordered increasingly so that $y_{\max} = y_n$ and $y_{\min} = y_1$. For the sake of a detailed illustration, we have chosen a sample of size $n = 5$ from a uniform parent function with $a = 0$ and $b = 1$, and data $y_1 = 0.2$, $y_2 = 0.3$, $y_3 = 0.4$, $y_4 = 0.5$, $y_5 = 0.95$. We have selected such a small sample size to be able to present the whole mathematical structure of the solution.

The corresponding parameter estimates using the constrained method of moments proposed above are:

$$\hat{a} = -0.00468, \quad \hat{b} = 0.95. \qquad (47)$$

The values of the dual variables are $\mu_1 = 0$; $\mu_2 = 0.0053$, showing that the minimum and maximum values of the sample, $y_1 = 0.2$ and $y_5 = 0.95$, force the constraint (41) to become inactive and the constraint (42) to become active, respectively. Since (41) is inactive, we ignore it in what follows.

In this case the matrices U and S become (see (34)):

$$U = \begin{bmatrix} 0.0000 & 0.0053 & 0 & -1 \\ 0.5534 & 0.4465 & 0 & 0 \\ 0.4465 & 0.5534 & -1 & 0 \\ 0.0000 & 0.0053 & 0 & 0 \end{bmatrix}, \; S = \begin{bmatrix} -0.0007 & 0.0000 & 0.0006 & 0.0013 & 0.0043 \\ 0.2344 & 0.2217 & 0.2090 & 0.1962 & 0.1390 \\ 0.1657 & 0.1784 & 0.1911 & 0.2039 & 0.2611 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & -1.0000 \end{bmatrix} \tag{48}$$

and the matrices V and T do not exist.

Since U is invertible, the problem is non-degenerate, and from (39) we obtain the partial derivatives:

$$\begin{bmatrix} \partial a/\partial y_1 & \cdots & \partial a/\partial y_5 \\ \partial b/\partial y_1 & \cdots & \partial b/\partial y_5 \\ \partial \mu_2/\partial y_1 & \ldots & \partial \mu_2/\partial y_5 \\ \partial z/\partial y_1 & \cdots & \partial z/\partial y_5 \end{bmatrix} = U^{-1}S = \begin{bmatrix} 0.4235 & 0.4005 & 0.3775 & 0.3545 & -0.5560 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 1.0000 \\ 0.0235 & 0.0005 & -0.0225 & -0.0455 & 0.0440 \\ 0.0007 & 0.0001 & -0.0006 & -0.0013 & 0.0010 \end{bmatrix} \tag{49}$$

5.2 Regular Degenerate Example

Consider the following simple nonlinear programming problem:

$$\min_{x_1, x_2} \quad f(x) = a_1 x_1^2 + x_2^2 \tag{50}$$

$$\text{s.t.} \quad h(x) = x_1 x_2^2 - a_2 = 0 : \; \lambda \tag{51}$$

$$g(x) = -x_1 + a_3 \leq 0 : \; \mu \tag{52}$$

with λ, μ the multipliers corresponding to the constraints (51) and (52), respectively.

The solution of this problem for the particular case $a_1 = a_3 = 1$ and $a_2 = 2$ is:

$$x_1^* = 1, \quad x_2^* = \sqrt{2}, \quad \lambda^* = -1, \quad \mu^* = 0, \quad z^* = 3. \tag{53}$$

A vector of changes $\delta p = (dx_1, dx_2, da_1, da_2, da_3, d\lambda, d\mu, dz)^T$ must satisfy the system (12)-(17), which for this example becomes (in matrix form):

$$M\delta p = \left(\begin{array}{rr|rrr|r|r|r} 2 & 2\sqrt{2} & 1 & 0 & 0 & 0 & 0 & -1 \\ \hline 2 & -2\sqrt{2} & 2 & 0 & 0 & 2 & -1 & 0 \\ -2\sqrt{2} & 0 & 0 & 0 & 0 & 2\sqrt{2} & 0 & 0 \\ \hline 2 & 2\sqrt{2} & 0 & -1 & 0 & 0 & 0 & 0 \end{array} \right) \delta p = 0 \tag{54}$$

$$N\delta p = \left(\begin{array}{rr|rrr|r|r|r} -1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \end{array} \right) \delta p \le 0. \tag{55}$$

In this case, matrix U has no inverse because is not a square matrix, the gradients of the constraints are linearly independent and one of the Lagrange multipliers is null; so, we have a regular degenerate case.

Note that we have not considered (18) yet. If we want (i) the inequality constraint to remain active the first equation in (55) should be removed and included in (54), whereas if we want (ii) the inequality constraint to be allowed to become inactive then the second equation in (55) should be removed and included in (54). The corresponding solutions are the cones

$$\begin{pmatrix} dx_1 \\ dx_2 \\ da_1 \\ da_2 \\ da_3 \\ d\lambda \\ d\mu \\ dz \end{pmatrix} = \rho_1 \begin{bmatrix} 1 \\ 0 \\ -2 \\ 2 \\ 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} + \rho_2 \begin{bmatrix} 0 \\ \frac{1}{\sqrt{2}} \\ 1 \\ 2 \\ 0 \\ 0 \\ 0 \\ 3 \end{bmatrix} + \pi \begin{bmatrix} 0 \\ -\frac{1}{\sqrt{2}} \\ 2 \\ -2 \\ 0 \\ 0 \\ 6 \\ 0 \end{bmatrix} \text{ and } \begin{pmatrix} dx_1 \\ dx_2 \\ da_1 \\ da_2 \\ da_3 \\ d\lambda \\ d\mu \\ dz \end{pmatrix} = \rho_1 \begin{bmatrix} 1 \\ 0 \\ -2 \\ 2 \\ 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} + \rho_2 \begin{bmatrix} 0 \\ \frac{1}{\sqrt{2}} \\ 1 \\ 2 \\ 0 \\ 0 \\ 0 \\ 3 \end{bmatrix} + \pi \begin{bmatrix} 1 \\ 0 \\ -2 \\ 2 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \quad (56)$$

respectively, where $\rho_1, \rho_2 \in \mathbb{R}$ and $\pi \in \mathbb{R}^+$, that give all feasible perturbations. Note, for example, that the component associated with $d\mu = (6\pi_1 \text{ or } 0)$ is always positive for (8) to hold, whereas the component related to the equality constraint $d\lambda = (\rho_1 \text{ or } \rho_1 + \pi)$ can be positive or negative.

In order to study the existence of directional derivatives with respect to a_1 we use the directions $da = (\,1 \quad 0 \quad 0\,)^T$ and $da = (\,-1 \quad 0 \quad 0\,)^T$, and solve the two possible combinations of (32)-(33) that lead to:

$$\frac{dp^*}{da_1^+} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 2 \\ 1 \end{bmatrix}; \quad \frac{dp^*}{da_1^-} = [\emptyset], \quad \text{and} \quad \frac{dp^*}{da_1^+} = [\emptyset]; \quad \frac{dp^*}{da_1^-} = \begin{bmatrix} \frac{1}{3} \\ -\frac{1}{3\sqrt{2}} \\ \frac{1}{3} \\ 0 \\ -1 \end{bmatrix}, \quad (57)$$

respectively, where $dp^*/da_1 = (\, dx_1/da_1 \; dx_2/da_1 \; d\lambda/da_1 \; d\mu/da_1 \; dz/da_1\,)$ and $[\emptyset]$ means that there is no solution, which implies that both directional derivatives exist (existence and uniqueness) but only the partial derivative of z with respect a_1 exists $\dfrac{\partial z}{\partial a_1} = 1$. For the remaining variables the directional derivatives do not coincide in absolute value, therefore, the corresponding partial derivatives do not exist. Note that in the right-derivative the solution point remains the same but the Lagrange multiplier μ associated with the inequality constraint becomes different from zero. For the left-derivative the solution point changes and the inequality constraint becomes inactive, note that the Lagrange multiplier associated with the equality constraint $h(x)$ changes but it is sufficient for getting a new optimal solution whereas the one related to the inequality constraint remains equal to zero.

The directional derivatives with respect to a_2 are obtained using the directions $da = \begin{pmatrix} 0 & 1 & 0 \end{pmatrix}^T$ and $da = \begin{pmatrix} 0 & -1 & 0 \end{pmatrix}^T$, and solving the two possible combinations of (32)-(33) leading to:

$$\frac{dp^*}{da_2^+} = [\emptyset], \quad \frac{dp^*}{da_2^-} = \begin{bmatrix} 0 \\ -\frac{1}{2\sqrt{2}} \\ 0 \\ 1 \\ -1 \end{bmatrix} \quad \text{and} \quad \frac{dp^*}{da_2^+} = \begin{bmatrix} \frac{1}{6} \\ \frac{1}{3\sqrt{2}} \\ \frac{1}{6} \\ 0 \\ 1 \end{bmatrix}, \quad \frac{dp^*}{da_2^-} = [\emptyset], \tag{58}$$

respectively, which implies that both directional derivatives exist (existence and uniqueness) but only the partial derivative of z with respect a_2 exists $\dfrac{\partial z}{\partial a_2} = 1$. For the remaining variables the partial derivatives do not exist. Note that in the right-derivative the solution point changes and the inequality constraint becomes inactive. The gradients of the objective and equality constraint remain with the same direction but different magnitude whereas for the left derivative the solution point changes as well but the inequality constraint remains active with Lagrange multiplier different from zero. Note that the inequality constraint forces the new solution point to move along its limit.

Analogously, the directional derivatives with respect to a_3 are obtained using the directions $da = (\,0 \quad 0 \quad 1\,)^T$ and $da = (\,0 \quad 0 \quad -1\,)^T$, and solving the two possible combinations of (32)-(33) leading to:

$$\frac{dp^*}{da_3^+} = \begin{bmatrix} 1 \\ -\frac{1}{\sqrt{2}} \\ 1 \\ 6 \\ 0 \end{bmatrix}, \quad \frac{dp^*}{da_3^-} = [\emptyset] \quad \text{and} \quad \frac{dp^*}{da_3^+} = [\emptyset], \quad \frac{dp^*}{da_3^-} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \tag{59}$$

respectively, which implies that both directional derivatives exist (existence and uniqueness) but only the partial derivative of z with respect a_3 exists $\dfrac{\partial z}{\partial a_3} = 0$. For the remaining variables the partial derivatives do not exist. Note that in the right-derivative the solution point changes but the inequality constraint remains active with Lagrange multiplier different from zero. The inequality constraint forces the solution point to move to the right. For the left-derivative the solution point does not change and the inequality constraint becomes inactive.

5.3 Nonregular Example

Consider the following simple nonlinear programming problem:

$$\underset{x_1, x_2}{\text{min}} \quad f(x) = x_1^2 + x_2^2 \tag{60}$$

$$\text{s.t.} \quad h(x) = -x_1 + a_1 = 0 \tag{61}$$

$$g_1(x) = -x_1 - x_2 + 2a_1 \le 0 \tag{62}$$

$$g_2(x) = a_2x_1 - x_2 \le 0 \tag{63}$$

with λ , μ_1, μ_2 the multipliers corresponding to the constraints (61)-(63).

The solution of this problem for the particular case $a_1 = a_2 = 1$ is:

$$x_1 = x_2 = 1; \;\; \mu_1 = \frac{4-\lambda}{2}; \;\; \mu_2 = \frac{\lambda}{2}. \tag{64}$$

Note that the dual problem has infinite solutions. Since the two inequality constraints are active, they will remain active or inactive in a neighborhood of the optimum depending on the values of the Lagrange multipliers. Then, a vector of changes $\delta p = (dx_1, dx_2, da_1, da_2, d\lambda, d\mu_1, d\mu_2, dz)^T$ must satisfy the system (12)-(17).

For all possible cases, the M matrix in (21) can be obtained from the following matrix

$$M = \left[\begin{array}{rr|rr|r|rr|r} 2 & 2 & 0 & 0 & 0 & 0 & 0 & -1 \\ \hline 2 & 0 & 0 & \mu_2 & -1 & -1 & 1 & 0 \\ 0 & 2 & 0 & 0 & 0 & -1 & -1 & 0 \\ \hline -1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ \hline -1 & -1 & 2 & 0 & 0 & 0 & 0 & 0 \\ 1 & -1 & 0 & 1 & 0 & 0 & 0 & 0 \end{array}\right] \tag{65}$$

by removing the rows corresponding to the null μ-multipliers, and the matrix N in (22) can be obtained from the matrix

$$N = \left[\begin{array}{rr|rr|r|rr|r} -1 & -1 & 2 & 0 & 0 & 0 & 0 & 0 \\ 1 & -1 & 0 & 1 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \end{array}\right] \tag{66}$$

by removing the rows corresponding to the non null μ-multipliers.

We analyze the only possible two different cases (see Equation (64)):

Case 1: $\mu_1, \mu_2 \neq 0$. For example $\lambda = 2;\ \ \mu_1 = 1;\ \ \mu_2 = 1$. In this case, the matrix U is singular because the gradients of the active constraints are not linearly independent; so, we have a non-regular case. Since all μ-multipliers are non-null, the N matrix does not exist and the system (21)-(22), using expression (65), becomes

$$M\delta p = \left(\begin{array}{rr|rr|r|rr|r} 2 & 2 & 0 & 0 & 0 & 0 & 0 & -1 \\ \hline 2 & 0 & 0 & 1 & -1 & -1 & 1 & 0 \\ 0 & 2 & 0 & 0 & 0 & -1 & -1 & 0 \\ \hline -1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ \hline -1 & -1 & 2 & 0 & 0 & 0 & 0 & 0 \\ 1 & -1 & 0 & 1 & 0 & 0 & 0 & 0 \end{array}\right) \delta p = 0. \qquad (67)$$

Note that in this example there is no need to consider (18) because the Lagrange multipliers are different from zero. The solution is the linear space

$$\begin{pmatrix} dx_1 \\ dx_2 \\ da_1 \\ da_2 \\ d\lambda \\ d\mu_1 \\ d\mu_2 \\ dz \end{pmatrix} = \rho_1 \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \\ 0 \\ 2 \\ 0 \\ 4 \end{bmatrix} + \rho_2 \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 2 \\ -1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \rho_1 \\ \rho_1 \\ \rho_1 \\ 0 \\ 2\rho_2 \\ 2\rho_1 - \rho_2 \\ \rho_2 \\ 4\rho_1 \end{bmatrix},\ \ \rho_1, \rho_2 \in \mathbb{R} \qquad (68)$$

that gives all feasible perturbations. Note that the vector associated with ρ_2 corresponds to the feasible changes in the Lagrange multipliers owing to the linearly dependence of the constraint gradients.

In order to study the existence of partial derivatives with respect to a_1 we use the directions $da = (1 \quad 0)^T$ and $da = (-1 \quad 0)^T$, that imply (see (68)) $\rho_1 = 1$ and

$\rho_1 = -1$, respectively, and

$$\begin{pmatrix} dx_1 \\ dx_2 \\ da_1 \\ da_2 \\ d\lambda \\ d\mu_1 \\ d\mu_2 \\ dz \end{pmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \\ 0 \\ 2 \\ 0 \\ 4 \end{bmatrix} + \rho_2 \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 2 \\ -1 \\ 1 \\ 0 \end{bmatrix}, \quad \begin{pmatrix} dx_1 \\ dx_2 \\ da_1 \\ da_2 \\ d\lambda \\ d\mu_1 \\ d\mu_2 \\ dz \end{pmatrix} = \begin{bmatrix} -1 \\ -1 \\ -1 \\ 0 \\ 0 \\ -2 \\ 0 \\ -4 \end{bmatrix} + \rho_2 \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 2 \\ -1 \\ 1 \\ 0 \end{bmatrix}, \tag{69}$$

which implies that the following partial derivatives exist: $\dfrac{\partial x_1}{\partial a_1} = \dfrac{\partial x_2}{\partial a_1} = 1$, and $\dfrac{\partial z}{\partial a_1} = 4$ because they are unique. However, the partial derivatives $\dfrac{\partial \lambda}{\partial a_1}$, $\dfrac{\partial \mu_1}{\partial a_1}$ and $\dfrac{\partial \mu_2}{\partial a_1}$ do not exist, because the corresponding $d\lambda$, $d\mu_1$, $d\mu_2$ are not unique (they depend on the arbitrary real number ρ_2).

Alternatively, it is possible to consider the direction in which the desired partial derivative is looked for, $da = (\,1 \quad 0\,)^T$, and solve (32)-(33) with da and $-da$ leading to:

$$\frac{dp^*}{da_1^+} = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 2 \\ 0 \\ 4 \end{bmatrix} + \rho_2 \begin{bmatrix} 0 \\ 0 \\ 2 \\ -1 \\ 1 \\ 0 \end{bmatrix}, \quad \frac{dp^*}{da_1^-} = \begin{bmatrix} -1 \\ -1 \\ 0 \\ -2 \\ 0 \\ -4 \end{bmatrix} + \rho_2 \begin{bmatrix} 0 \\ 0 \\ 2 \\ -1 \\ 1 \\ 0 \end{bmatrix}, \tag{70}$$

where $dp^*/da_1 = (\, dx_1/da_1 \; dx_2/da_1 \; d\lambda/da_1 \; d\mu_1/da_1 \; d\mu_2/da_1 \; dz/da_1 \,)$.

As (33) does not exist in this case, this condition holds strictly and (70) provides the partial derivatives if the solution is unique. The partial derivatives obtained coincide with the ones obtained from (69) , i.e., $\partial x_1/\partial a_1 = \partial x_2/\partial a_1 = 1$, and $\partial z/\partial a_1 = 4$, whereas the partial derivatives $\partial \lambda/\partial a_1$, $\partial \mu_1/\partial a_1$ and $\partial \mu_2/\partial a_1$ do not exist, because they are not unique (they depend on the arbitrary real number ρ_2).

Since in (68) $da_2 = 0$, the partial derivatives with respect to a_2 do not exist. The same result can be obtained considering the direction $da = (\,0 \quad 1\,)^T$ in (32), that has no solution, that is, no derivative exist with respect to da_2.

Note that in this case the active constraints remain active (all μ multipliers are positive). This implies that the cone degenerates to a linear space.

Case 2: $\mu_1 = 0$; $\mu_2 \neq 0$. For example, $\lambda = 4$; $\mu_1 = 0$; $\mu_2 = 2$. In this case, the matrix U is singular because the gradients of the active constraints are not linearly independent; so, we also have a non-regular case. The system (21) - (22), using expression (65) and (66), becomes:

$$M\delta p = \left(\begin{array}{rr|rr|r|rr|r} 2 & 2 & 0 & 0 & 0 & 0 & 0 & -1 \\ \hline 2 & 0 & 0 & 2 & -1 & -1 & 1 & 0 \\ 0 & 2 & 0 & 0 & 0 & -1 & -1 & 0 \\ \hline -1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ \hline 1 & -1 & 0 & 1 & 0 & 0 & 0 & 0 \end{array}\right) \delta p = 0 \;, \tag{71}$$

$$N\delta p = \left(\begin{array}{rr|rr|r|rr|r} -1 & -1 & 2 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 \end{array}\right) \delta p \leq 0 \;. \tag{72}$$

Note that we have not considered (18) yet. If we want (i) the inequality constraint $g_1(x)$ to remain active the first equation in (72) should be removed and included in (71), whereas if we want (ii) the inequality constraint to be allowed to become inactive then the second equation in (72) should be removed and included in (71).

The corresponding solutions are the cones

$$\begin{pmatrix} dx_1 \\ dx_2 \\ da_1 \\ da_2 \\ d\lambda \\ d\mu_1 \\ d\mu_2 \\ dz \end{pmatrix} = \rho \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \\ 4 \\ 0 \\ 2 \\ 4 \end{bmatrix} + \pi \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ -2 \\ 1 \\ -1 \\ 0 \end{bmatrix} \quad \text{and} \quad \begin{pmatrix} dx_1 \\ dx_2 \\ da_1 \\ da_2 \\ d\lambda \\ d\mu_1 \\ d\mu_2 \\ dz \end{pmatrix} = \rho \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \\ 4 \\ 0 \\ 2 \\ 4 \end{bmatrix} + \pi \begin{bmatrix} -1 \\ 1 \\ -1 \\ 2 \\ 4 \\ 0 \\ 2 \\ 0 \end{bmatrix}, \tag{73}$$

respectively, where $\rho \in \mathbb{R}$ and $\pi \in \mathbb{R}^+$. Analogously to the previous case, the vector associated with π for the first hypothesis corresponds to the feasible changes in the Lagrange multipliers owing to the linearly dependence of the constraint gradients but only positive increments are allowed because as $\mu_1 = 0$, a negative increment would imply a negative multiplier which is incompatible with KKT conditions.

Note that constraint $g_1(x)$ it is not necessary for getting the optimal solution (64), this means that it could be removed and the same optimal solution would still remain.

In order to study the existence of partial derivatives with respect to a_1 we use the directions $da = (1 \quad 0)^T$ and $da = (-1 \quad 0)^T$, that imply considering (73)-left,

$\rho = 1$ and $\rho = -1$, respectively, leading to:

$$\begin{pmatrix} dx_1 \\ dx_2 \\ da_1 \\ da_2 \\ d\lambda \\ d\mu_1 \\ d\mu_2 \\ dz \end{pmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \\ 4 \\ 0 \\ 2 \\ 4 \end{bmatrix} + \pi \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ -2 \\ 1 \\ -1 \\ 0 \end{bmatrix}, \quad \begin{pmatrix} dx_1 \\ dx_2 \\ da_1 \\ da_2 \\ d\lambda \\ d\mu_1 \\ d\mu_2 \\ dz \end{pmatrix} = \begin{bmatrix} -1 \\ -1 \\ -1 \\ 0 \\ -4 \\ 0 \\ -2 \\ -4 \end{bmatrix} + \pi \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ -2 \\ 1 \\ -1 \\ 0 \end{bmatrix}, \tag{74}$$

and considering (73)-right, $\rho = 1, \pi = 0$ and $\rho = -1, \pi = 0$, respectively, dealing to:

$$\begin{pmatrix} dx_1 \\ dx_2 \\ da_1 \\ da_2 \\ d\lambda \\ d\mu_1 \\ d\mu_2 \\ dz \end{pmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \\ 4 \\ 0 \\ 2 \\ 4 \end{bmatrix}, \quad \begin{pmatrix} dx_1 \\ dx_2 \\ da_1 \\ da_2 \\ d\lambda \\ d\mu_1 \\ d\mu_2 \\ dz \end{pmatrix} = \begin{bmatrix} -1 \\ -1 \\ -1 \\ 0 \\ -4 \\ 0 \\ -2 \\ -4 \end{bmatrix}, \tag{75}$$

respectively, which imply that the following partial derivatives exist: $\partial x_1/\partial a_1 = \partial x_2/\partial a_1 = 1$ and $\partial z/\partial a_1 = 4$, because they are unique and have the same absolute value and different sign. However, the partial derivatives $\partial\lambda/\partial a_1$, $\partial\mu_1/\partial a_1$ and $\partial\mu_2/\partial a_1$ do not exist, because the corresponding $d\lambda$, $d\mu_1$, $d\mu_2$ are not unique (they depend on the arbitrary real number π in (74)).

Alternatively, it is possible to consider the directions in which the desired directional derivatives are looked for, $da = (1 \quad 0)^T$ and $da = (-1 \quad 0)^T$, and solve the two possible combinations of (32)-(33) leading to:

$$\frac{dp^*}{da_1^+} = \begin{bmatrix}1\\1\\4\\0\\2\\4\end{bmatrix} + \pi \begin{bmatrix}0\\0\\-2\\1\\-1\\0\end{bmatrix}, \quad \frac{dp^*}{da_1^-} = \begin{bmatrix}-1\\-1\\-4\\0\\-2\\-4\end{bmatrix} + \pi \begin{bmatrix}0\\0\\-2\\1\\-1\\0\end{bmatrix}, \; \pi \in \mathbb{R}^+, \tag{76}$$

and

$$\frac{dp^*}{da_1^+} = \begin{bmatrix}1\\1\\4\\0\\2\\4\end{bmatrix}, \quad \frac{dp^*}{da_1^-} = \begin{bmatrix}-1\\-1\\-4\\0\\-2\\-4\end{bmatrix}, \tag{77}$$

where the same results as in (74) and (75) are obtained.

In order to study the existence of partial derivatives with respect to a_2, we use the directions $da = (0 \quad 1)^T$ and $da = (0 \quad -1)^T$, that imply considering (73)-right, that $\rho = \pi = 1/2$ and as the value of π can just be positive, it is not possible to get $da_2 = -1$ neither from (73)-left nor from (73)-right and then no partial derivatives exist with respect to a_2. Therefore, as $\pi > 0$ only right-derivatives can exist:

$$\delta p = [0 \quad 1 \quad 0 \quad 1 \quad 4 \quad 0 \quad 2 \quad 2]^T \,, \tag{78}$$

which implies: $\partial x_1/\partial a_2^+ = 0$, $\partial x_2/\partial a_2^+ = 1$, $\partial\lambda/\partial a_2^+ = 4$, $\partial\mu_1/\partial a_2^+ = 0$, $\partial\mu_2/\partial a_2^+ = 2$ and $\partial z/\partial a_2^+ = 2$ because they are unique.

Alternatively, if we try to solve (32)-(33) using $da = (\,0 \quad 1\,)^T$ and $da = (\,0 \quad -1\,)^T$:

$$\frac{dp^*}{da_2^+} = \pi \begin{bmatrix} 0 \\ 0 \\ -2 \\ 1 \\ -1 \\ 0 \end{bmatrix}, \; \frac{dp^*}{da_2^-} = \pi \begin{bmatrix} 0 \\ 0 \\ -2 \\ 1 \\ -1 \\ 0 \end{bmatrix} \quad \text{and} \quad \frac{dp^*}{da_2^+} = \begin{bmatrix} 0 \\ 1 \\ 4 \\ 0 \\ 2 \\ 2 \end{bmatrix}, \; \frac{dp^*}{da_2^-} = [\emptyset], \tag{79}$$

where the same results as in (78) are obtained.

Note that in this example constraint g_1 becomes inactive.

6 Conclusions

In the context of sensibility analysis in nonlinear programming, the main conclusions derived from the work reported in this paper are the following:

(i) There is no need to assume that the active constraints remain active after small perturbations in the neighborhood of a solution point, the method allows active constraints with multipliers equal to zero to become inactive after the perturbation.

(ii) The theory of polyhedral cones allows deriving the most general feasible perturbation.

(iii) The proposed method allows determining whether or not a given directional or partial derivative exists.

(iv) The method allows calculating left and right derivatives in all directions if they exist.

(v) The feasible directions define directions with existing directional derivatives.

(vi) For regular non-degenerate cases (the most common in real world applications), the proposed method allows obtaining simultaneously the partial derivatives of the objective function and the primal and dual variables with respect to the data in an elegant and neat way, just solving a system of linear equations. For the degenerate regular and non-regular cases, alternative procedures with higher computational burden are provided.

References

1. Vanderplaats, G. N., *Numerical Optimization Techniques for Engineering Design*, McGraw-Hill, New York, New York, USA, 1984.

2. Sobieski, J. S., Barthelemy, J. F. and Riley, K. M., *Sensitivity of Optimal Solutions of Problems Parameters*, AIAA Journal. Vol. 20, pp. 1291–1299, 1982.

3. Enevoldsen, I., *Sensitivity Analysis of Reliability-Based Optimal Solution*, ASCE Journal of Engineering Mechanics. Vol. 120, pp. 198–205, 1994.

4. Roos, C., Terlaky, T. and Vial, J. P., *Theory and Algorithms for Linear Optimization: An Interior Point Approach*, John Wiley & Sons, Chichester, UK, 1997.

5. Bjerager, P. and Krenk, S., *Parametric Sensitivity in First Order Reliability Theory*, ASCE Journal Engineering Mechanics. Vol. 115, pp. 1577–1582, 1989.

6. Sorensen, J. D. and Enevoldsen, I., *Sensitivity Analysis in Reliability-Based Shape Optimization*, Optimization and Artificial Intelligence in Civil and Structural Engineering, Vol. 221, Edited by B. H. V. Topping, Kluwer Academic Publishers, Dordrech, The Netherlands, 1992.

7. Luenberger, D. G., *Linear and Nonlinear Programming*, Second Edition, Addison-Wesley, Reading, Massachusetts, USA, 1989.

8. Bazaraa, M. S., Sherali, H. D. and Shetty, C. M., *Nonlinear Programming. Theory and Algorithms*, Second Edition, John Wiley & Sons, New York, New York, USA, 1993.

9. Padberg, M., *Linear Optimization and Extensions*, Springer, Berlin, Germany, 1995.

10. Castillo, E., Cobo, A., Jubete, F., Pruneda, R. E. and Castillo, C., *An Orthogonally Based Pivoting Transformation of Matrices and Some Applications*, SIAM Journal on Matrix Analysis and Applications, Vol. 22, pp. 666–681, 2000.

11. Castillo, E., Jubete, F., Pruneda, E. and Solares, C., *Obtaining Simultaneous Solutions of Linear Subsystems of Equations and Inequalities*, Linear Algebra and its Applications, Vol. 346, pp. 131–154, 2002.

12. Castillo, E., Cobo, A., Jubete, F. and Pruneda, R. E., *Orthogonal Sets and Polar Methods in Linear Algebra: Applications to Matrix Calculations, Systems of Equations and Inequalities, and Linear Programming*, John Wiley & Sons, New York, New York, USA, 1999.

13. Press, W. H., Teukolsky, S. A., Vetterling, W. T. and Flannery, B. P., *Numerical Recipes in C*, Cambridge University Press, New York, New York, USA, 1992.