Informative Lagrange Multipliers in the Nonlinear Parametric Programming Model

Tao Jie  
University of Shanghai for Science and Technology

Gao Yan  (✉ gaoyan@usst.edu.cn)  
University of Shanghai for Science and Technology

Research Article

Keywords: Parametric nonlinear programming, Informative Lagrange multiplier, Shadow price, Penalty function method

Posted Date: November 29th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-2309493/v1

License: ☒ This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Additional Declarations: No competing interests reported.
Informative Lagrange Multipliers in the Nonlinear Parametric Programming Model

Tao Jie, Gao Yan

Abstract  The shadow price expresses the marginal cost with respect to the variation of constraints, and it is extremely useful in the sensitivity analysis of nonlinear programming models. However, the shadow price may fail to exist in particular parametric programming models, and the informative Lagrange multipliers are proposed to supplement the theory of the shadow price. The traditional analysis of informative Lagrange multipliers is based on the right hand side perturbation model, in which the resource constraints are assumed to be arbitrarily violated, and the variation of resources are measured by the perturbations. In the line of traditional analysis, the minimum norm Lagrange multiplier is proved to be informative since it expresses the rate of cost improvement per unit constraints violation along the steepest descent direction. However, the internal cause of the resource variations is neglected, and the minimum norm Lagrange multiplier may fail to be informative when the perturbations are not only on the right hand side of the constraints. In this paper, we extend the classical constraint violation condition to the generalized constraint violation condition, which captures the characteristic of the problem structure of nonlinear parametric programming models. Based on the generalized constraint violation condition, we provide sufficient conditions for the minimum norm Lagrange multiplier to be informative. Furthermore, we propose a kind of penalty function method to derive the informative LM in fully parametric programming models, which means that the perturbations are not only on the right hand side of the constraints. Finally, we use examples to support our theoretic results.

Keywords  Parametric nonlinear programming; Informative Lagrange multiplier; Shadow price; Penalty function method

1 INTRODUCTION

The shadow price is interpreted as the marginal cost with respect to the increment of resources, and it is expressed as the directional derivative of the optimal value function with respect to the perturbation on the constraints. In the economic area, shadow prices have long been recognized as Lagrange multipliers (hereafter referred to as, LM). Indeed, the LM expresses the sensitivity
information when the LM is unique. For example, under the linear independence constraint qualification, second order sufficient optimality condition, together with the strict complementarity slackness condition, the differential stability of the LM can be obtained [c.f. Fiacco and McCormick (1990)]. Subsequent work can be referred to as Jittorntrum (1984); Robinson (1982). However, the equivalence of the shadow price and LM may fail to hold when LMs are multiple. This observation raises the natural question of identifying LMs that convey sensitivity information.

The first step in this direction is taken by Aucamp and Steinberg (1982); Akgül (1984). They point out that multiple LMs may exist when the optimal solution of a linear programming model is degenerate. They also propose the notion of buying / selling shadow prices, which denote the price of an extra unit of resources that deserves to buy or to sell. Similar studies can also be found in Jansen et al. (1997). These results, however, are limited in linear programming models, and therefore it is naturally to consider whether these results can be extended to nonlinear programming models.

Gauvin and Tolle (1977) study the differential properties of the optimal value function under the assumption of Mangasarian - Fromovitz constraint qualification in the case of right hand side (hereafter referred to as, RHS) perturbations. In Gauvin and Tolle (1977), sufficient conditions to guarantee the existence of the Dini derivatives of the optimal value function is provided, and based on this conditions, buying / selling SPs are extended to general nonlinear programming models [c.f. Gauvin (1980)]. In the subsequent work, these results are further extended to the fully parametric nonlinear programming models; that is, the perturbations are not only on the RHS of the constraints [e.g., see Wolkowicz (2020) for the “restricted” LM when the nonlinear program is unstable; see Kim and Cho (1988), Crema (1995), Cho and Kim (1992) for the average shadow prices]. A significant step forward for developing the sensitivity interpretation of LMs is taken by Bertsekas and Ozdaglar (2002), in which the notion of informative LM is proposed. The informative LM provides a significant amount of sensitivity information by indicating which constraints to violate in order the effect a cost improvement. The salient property of the informative LM is consistent with the classical sensitivity interpretation of LMs. There is, however, discrepancy between informative LMs and classical shadow prices. The shadow price, according to its definition indicated in our previous discussion, expresses the rate of improvement of the optimal value function with respect to the increment of the constraints, and it follows that the existence of the shadow price is on the condition that the partial derivatives of the optimal value function exists (or at least the optimal value function is Lipschitz continuous). The informative LM, however, does not require this assumption [c.f. Tao and Gao (2021b)]. Therefore, it is not surprising that at least one informative LM exists with hardly any assumptions.

Bertsekas and Ozdaglar (2002) show that an LM with the minimum Euclidean norm is informative, and it can be used to measure the scarcity of resources when the classical shadow price of the nonlinear programming fails to exist. The minimum norm LM (hereafter referred to as, MNLM) expresses the rate of improvement per unit constraint violation, along the maximum improvement direction $d_{\text{MNLM}}$. The existence of $d_{\text{MNLM}}$ is guaranteed by the “constraint violation condition”. We refer to the literature Bertsekas et al. (2003) for more details. The MNLM is extremely useful in nonconvex optimization models since it provides a simple way to identify the “bottlenecks” of resources supplement when the classical SPs fail to exist [c.f. Tao and Gao (2021a)].

However, a practical drawback of MNLM is that it only uses the RHS vector of the constraints to measure the variation of resources, while neglecting the reason for these variations. To illustrate the importance of fully variations on the optimization model in practical applications, we refer to the example presented in the 3rd section in Bonyadi et al. (2014). In fact, the variations of resources are caused by external factors (represented by the perturbations in the model), such as the perturbations in the problem data [c.f. Bonyadi et al. (2014), Crema (2018)]. These perturbation factors need to be considered in the modelling stage, rather than arbitrarily assumed on the RHS of the constraints. For example, the line of analysis on the MNLM assumes that the supplement of all resources in the constraints can be violated, which ignores the possibility of scarce resources. The reason is that the supplement of scarce resources is limited, and hence it is not reasonable to
assume the increment of such resources. Therefore it is natural to ask whether MNLM still conveys the sensitivity information when the cause of the resource variations is taken into consideration.

In this research, we focus attention primarily on the fully parametric nonlinear programming models. We will first extend the classical constraint violation condition to a more general version which captures the characteristic of the problem structure of fully parametric mathematical programming models. Then we use a counterexample to show that the MNLM may not be informative when the perturbation vector is not only on the RHS of the constraints, and then we will propose sufficient conditions to ascertain the equivalence of MNLM and informative LMs. Furthermore, we will propose a kind of penalty function method to derive an informative LM in fully parametric nonlinear programming models. To the best of our knowledge, there was no similar work done in the parametric nonlinear programming domain that involves a characterization of informative LMs in fully parametric programming models.

1.1 Structure of the Paper

In Section 2, we present some preliminary results on the set of LMs and constraint violation conditions. In Section 3, we propose that the internal cause of the resource variations is neglected in the classical analysis of the informative LMs, and the perturbation factors need to be considered in the modelling stage. Furthermore, we extend the classical constraint violation condition to a generalized version which captures the characteristic of the problem structure of fully parametric mathematical programming models. In Section 4, we show that the MNLM may fail to be informative, and propose the sufficient condition for an MNLM to be informative in the fully parametric programming model. In Section 5, we propose a kind of the penalty function method to derive an informative LM. In Section 6, several examples are used to show the validity of our proposed method. Finally, Section 7 contains some concluding remarks.

2 Preliminaries

2.1 Notations

Unless particularly specified, the norm used in this paper is the Euclidean norm; \( \mathbb{R}^n \) denotes the \( n \)-dimensional Euclidean space; \( \cdot \rightarrow \cdot \) denotes the limit of a certain sequence; \( g^+ = \max\{0, g\} \); \( e_i \) represents the vector whose elements are all equal to zero except the \( i \)th element being one; \( \nabla \) denotes the gradient operator; let \( \{y_k\} \) be a sequence in \( \mathbb{R}^n \), \( (y_k)_i \) denotes the \( i \)th element of \( y_k \), \( i = 1, \ldots, n \); \( 0_{m \times n} \) represents a zero valued \( m \times n \) matrix, and \( I_n \) represents an \( n \)-dimensional identity matrix; Let \( R \) be a set and \( x \) is some vector, \( \text{dist}(x, R) \) denotes the distance from \( x \) to \( R \), i.e., \( \text{dist}(x, R) = \inf_{z \in R} \|x - z\| \); \( B(x, r) \) denotes an open ball centered at \( x \) with radius \( r \), and its closure is denoted as \( \overline{B(x, r)} \).

2.2 Basic Definitions

Consider the following economic system:

\[
\begin{align*}
\min & \quad f(x, y) \\
\text{s.t.} & \quad h_i(x, y) = 0, i = 1, \ldots, p, \\
& \quad g_j(x, y) \leq 0, j = 1, \ldots, q, \\
& \quad (x, y) \in \mathbb{R}^n \times \mathbb{R}^m.
\end{align*}
\]  

(1)
Throughout this paper, we assume that \( f, h_i, g_j : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}, \ i = 1, \ldots, p, \ j = 1, \ldots, q \) are continuously differentiable functions, which represent the cost function and the resource constraints, respectively. \( x \) and \( y \) denotes the decision variable and the perturbation vector respectively. Note that if the vector \( y \) only perturbs the RHS of the constraints, Problem (1) will be turned into the following RHS perturbed model

\[
\begin{align*}
\min_x f(x) \\
\text{s.t.} \quad h_i(x) &= y_i, i = 1, \ldots, p, \\
g_j(x) &\leq y_{j+p}, j = 1, \ldots, q.
\end{align*}
\]

(2) Let \( R \)

In comparison, Problem (1) is referred to as the fully parametric nonlinear programming model. For notation simplicity, we denote by \( f_{\text{inf}}(\cdot) \) the optimal value function of Problem (1). For some perturbation vector \( y \in \mathbb{R}^m \), let \( R(y) \) and \( P(y) \) denote the set of feasible solutions and optimal solutions, respectively, i.e.,

\[
R(y) = \{ x \in \mathbb{R}^n | h_i(x, y) = 0, i = 1, \ldots, p, g_j(x, y) \leq 0, j = 1, \ldots, q \},
\]

\[
P(y) = \{ x \in R(y) | f(x, y) \leq f_{\text{inf}}(y) \}.
\]

It is obvious that both \( R(y) \) and \( P(y) \) are multifunctions. For completeness, we include the continuity properties of multifunctions, and we refer to Bank et al. (1982); Berge (1997); Hogan (1973) for more discussion.

**Definition 1** A multifunction \( R : \mathbb{R}^m \to \mathbb{R}^n \) is

1. **closed** at \( y^* \in \mathbb{R}^m \) if for each \( y_k \to y^* \), \( x_k \in R(y_k) \), \( x_k \to x^* \), it follows that \( x^* \in R(y^*) \).

2. **upper semicontinuous** (or simply u.s.c) at \( y^* \) if for each open set \( G \supseteq R(y^*) \), there exists a \( \delta > 0 \) such that for all \( y \in B(y^*; \delta) \), \( G \supseteq R(y) \) holds.

3. **lower semicontinuous** (or simply l.s.c) at \( y^* \) if for each open set \( G \cap R(y^*) \neq \emptyset \), there exists a \( \delta > 0 \) such that for all \( y \in B(y^*; \delta) \), \( G \cap R(y) \neq \emptyset \) holds.

4. **Hausdorff upper semi-continuous** (or simply H-u.s.c) at \( y^* \) if for each \( \epsilon > 0 \) there exists a \( \delta > 0 \) such that \( R(y) \subseteq B(R(y^*), \epsilon) \), for all \( y \in B(y^*, \delta) \).

5. **Hausdorff lower semi-continuous** (or simply H-l.s.c) at \( y^* \) if for each \( \epsilon > 0 \) there exists a \( \delta > 0 \) such that \( R(y^*) \subseteq B(R(y), \epsilon) \), for all \( y \in B(y^*, \delta) \).

The following relations are immediate consequences of the above definitions:

1. Let \( R(y^*) \) be a relatively compact set. \( R \) is H-l.s.c at \( y^* \iff R \) is l.s.c at \( y^* \).

2. Let \( R \) be uniformly bounded near \( y^* \), then \( R(y^*) \) is compact and \( R \) is u.s.c at \( y^* \iff R \) is closed at \( y^* \).

Let \( \mathcal{U} \) and \( \mathcal{B} \) denote the solvability set and the feasibility set respectively, i.e.,

\[
\mathcal{U} = \{ y | P(y) \neq \emptyset \},
\]

\[
\mathcal{B} = \{ y | R(y) \neq \emptyset \}.
\]

Now we introduce some definitions and develop some preliminary results. The Lagrange dual problem of Problem (1) is defined as follows:

\[
\max_{\lambda \in \mathbb{R}^m, \mu \geq 0} \min_{x \in \mathbb{R}^n} \{ L(x, \lambda, \mu, y) \},
\]

(3) where \( L : \mathbb{R}^{n+p+q+m} \to \mathbb{R} \) is the Lagrange function of Problem (1)

\[
L(x, \lambda, \mu, y) = f(x, y) + \sum_{i=1}^{p} \lambda_i h_i(x, y) + \sum_{j=1}^{q} \mu_j g_j(x, y).
\]
For \((x^*, y^*) \in P(y^*) \times U\), a Lagrange multiplier \((\lambda^*, \mu^*) \in \mathbb{R}^{p+q}\) of Problem (1) at \(x^* \in P(y^*)\) is a vector satisfying the following conditions,

\[
\nabla_x f(x^*, y^*) + \sum_{i=1}^{p} \lambda_i^* \nabla_x h_i(x^*, y^*) + \sum_{j=1}^{q} \mu_j^* \nabla_x g_j(x^*, y^*) = 0 \tag{4}
\]

\[
\mu^* \geq 0, \tag{5}
\]

\[
\mu_i^* = 0, i \notin A(x^*), \tag{6}
\]

where \(A(x^*) = \{i|g_i(x^*, y^*) = 0, i \in \{1, \ldots, q\}\}\). For notational convenience, we assume \(A(x^*) = \{j_1, \ldots, j_r\}\) in the subsequent sections.

Bertsekas et al. (2003) propose the notion of informative LMs based on the RHS perturbed mathematical programming problem, i.e., Problem (2). In Bertsekas et al. (2003), an LM satisfying the constraint violation (CV) condition is called the informative LM. Let us now recall the definition of the CV condition (in a slightly different but more complete formulation), given in Bertsekas et al. (2003).

**Definition 2** Let \((\lambda^*, \mu^*) \in \mathbb{R}^{p+q}\) be an LM of Problem (2) corresponding to the optimal solution \(x^*\) with \(y_i = 0, i = 1, \ldots, p + q\). Furthermore, let \(I^* = \{i|\lambda_i^* \neq 0, i = 1, \ldots, m\}\), \(J^* = \{j \in A(x^*)|\mu_j^* > 0\}\), and \(R(y) = \{x|h_i(x) = y_i, i = 1, \ldots, p, g_j(x) \leq y_{j+p}, j = 1, \ldots, q\}\) be the set of the feasible solution of Problem (2). Assume \(I^* \cup J^* \neq \emptyset\). If there exists a sequence \(\{y_k\} \subseteq \mathbb{R}^{p+q}\) in the neighbourhood of \(y^*\) such that

1. \(f_{\text{inf}}(y_k) < f(x^*)\),
2. for all \(i \in I^*\), \(\lambda_i^*(y_k) > 0\), and for all \(j \in J^*\), \(\mu_j^*(y_k)_{j+p} > 0\), where \((y_k)_i\) denotes the \(i\)th element of the vector \(y_k\),
3. for all \(i \notin I^* \cup J^*\), \(((y_k)_i)| = o(\omega(k))\),

where \(\omega(k) = \min_{i \in I^* \cup J^*} |(y_k)_i|\).

Then \((\lambda^*, \mu^*)\) is called the **informative LM** of Problem (2), and the conditions restricted on \(\{y_k\}\) is called the **constraint violation condition**.

We will develop a few guiding principles for understanding the significance of the informative LMs.

1. We first note that if in addition, the sequence \(\{y_k\}\) (in Definition 2) satisfies \(R(y_k) \neq \emptyset\) for all \(k\), then the condition “\((1) f_{\text{inf}}(y_k) < f(x^*)\)” in Definition 2 can be weakened to the condition that “\((1)\)” there exists \(x_k \in R(y_k)\) such that \(f(x_k) < f(x^*)\) for all \(k\), as shown in Bertsekas et al. (2003). The discussion in the subsequent sections is more heavily weighted towards the condition “\((1)\)” instead of the condition “\((1)\)” in Definition 2. We also remark that if the feasibility set mapping \(R(\cdot)\) is l.s.c at \(y^*\), then \(R(y_k) \neq \emptyset\) for any \(y_k\) in the neighbourhood of \(y^*\).

2. The CV condition does not assume that Problem (2) is feasible for any vector \(y \in \mathbb{R}^{p+q}\) near the origin. Roughly, we only consider a particular sequence of constraint violations \(y_k\) with \(R(y_k) \neq \emptyset\). In particular, if \((y_k)_i = g_i(x_k) - g_i(x^*)\) satisfies the CV condition, then the corresponding LM is informative. Moreover as would become clear below, the condition “\(x_k \in R(y_k)^-\)” is not sufficient, and it can be further weakened to dist\((x_k, R(y_k))\) \(\rightarrow 0\) as \(k \rightarrow \infty\).

3. We can construct the constraint violation sequence \(\{x_k\}\) by using the first order information of the cost function and resource constraints. If there exists a sequence \(\{d_k\} \subseteq \mathbb{R}^n\) with \(d_k \rightarrow 0\) (as \(k \rightarrow \infty\)) such that

   (i) \(\sum_{d_k \parallel \nabla f(x^*)} d_k < 0\),

   (ii) \(\sum_{d_k \parallel \nabla h_i(x^*)} d_k > 0\), for all \(i \in I^* = \{i \in I^*|\lambda_i^* > 0\}\),
(iii) \( \nabla h_i(x^*)^T d_k \rightarrow 0 \), for all \( i \in I^* \), \( h_i(x^*) < 0 \),

(iv) \( \nabla g_j(x^*)^T d_k \rightarrow 0 \), for all \( j \notin J^* \),

(v) \( \nabla g_j(x^*)^T d_k \geq 0 \), for all \( j \in J^* \),

(vi) \( \nabla g_j(x^*)^T d_k \leq 0 \), for all \( j \notin J^* \).

Then the sequence \( \{ x_k \} \) with \( x_k = x^* + d_k \) satisfies the CV condition, and \( (\lambda^*, \mu^*) \) is an informative LM. Since for each \( k \),

\[
(f(x)_i - f(x^*)) = \nabla f(x^*)^T d_k + o(\|d_k\|),
\]

\[
(y_k)_i = h_i(x_k) - h_i(x^*) = \nabla h_i(x^*)^T d_k + o(\|d_k\|), i = 1, \ldots, p,
\]

\[
(y_k)_{i+p} = g_j(x_k) - g_j(x^*) = \nabla g_j(x^*)^T d_k + o(\|d_k\|), j = 1, \ldots, q,
\]

then by the assumptions on \( d_k \), it follows that \( \{ x_k \} \) satisfies the CV condition. Furthermore, by the first-order optimality condition \( \nabla f(x^*) = -\left( \sum_{i \in I^*} \lambda_i^* \nabla h_i(x^*) + \sum_{j \in J^*} \mu_j^* \nabla g_j(x^*) \right) \), we have \( f(x_k) - f(x^*) = -\left( \sum_{i \in I^*} \lambda_i^* h_i(x_k) + \sum_{j \in J^*} \mu_j^* g_j(x_k) \right) + o(\|x_k - x^*\|) \), and hence \( (\lambda^*, \mu^*) \) expresses the sensitivity information of the resource constraints.

Carrying this idea one step further, we will show that \( x_k \in R(y_k) \) \(^1\) is not sufficient to ascertain the CV condition. It can be replaced by some weaker conditions in some special nonlinear programming models. Suppose that \( x_k \) and \( y_k \) satisfy the following conditions

(i) \( f(x_k) < f(x^*) \),

(ii) for all \( i \in I^* \), \( \lambda_i^*(y_k)_i > 0 \), and for all \( j \in J^* \), \( \mu_j^*(y_k)_{i+p} > 0 \),

(iii) for all \( i \notin I^* \cup J^* \), \( |(y_k)_i| = o(\omega(k)) \).

However, we assume \( \text{dist}(x_k, R(y_k)) \rightarrow 0 \) (as \( k \rightarrow \infty \)) instead of \( x_k \in R(y_k) \) for each \( k \). We will show that the CV condition still holds for \( \{ x_k \} \).

**Proposition 1** Suppose the feasible solution set \( R(y) \) is nonempty, uniformly compact and l.s.c near \( y^* \). Furthermore, suppose there exists a compact set \( K \) such that for all \( y \) in the neighbourhood of \( y^* \), \( P(y) \cap K \neq \emptyset \). Let \( y_k \rightarrow y^* \) and the sequence \( \{ x_k \} \) satisfies \( \text{dist}(x_k, R(y_k)) \rightarrow 0 \) (as \( k \rightarrow \infty \)). Then the CV condition holds for \( \{ y_k \} \).

**Proof** First we will show that \( R \) is H-u.s.c at \( y_k \) for all \( k \). Since \( h_i \) and \( g_j \) are continuously differentiable functions, \( i = 1, \ldots, p \), \( j = 1, \ldots, q \), then it follows that \( R \) is closed at \( y_k \) for any \( k \), which together with the uniformly compactness assumption, yields that \( R \) is H-u.s.c at \( y_k \) and \( R(y_k) \) is compact for all \( k \).

Furthermore, notice that \( R \) is l.s.c. and uniformly compact near \( y^* \), it follows that \( R \) is H-l.s.c near \( y^* \). Therefore, by using the fact that \( \text{dist}(x_k, R(y_k)) \rightarrow 0 \) (as \( k \rightarrow \infty \)), we have for \( y_{k,j} \rightarrow y_k \) (as \( j \rightarrow \infty \)), \( x_k \in R(y_k) \) holds for sufficiently large \( j \).

Then we will show that the optimal value function \( f_{\inf} \) is lower semi-continuous at \( y_k \) for all \( k \). Let us fix some \( k \), and let \( y_{k,i} \rightarrow y_k \) (as \( i \rightarrow \infty \)). By the assumption that \( P(y) \cap K \neq \emptyset \), we have there exists \( x_{k,i} \in P(y_{k,i}) \cap K \). Let us select a subsequence of \( \{ x_{k,i} \} \) (namely \( \{ x_{k,i} \} \) with \( x_{k,i} \in P(y_{k,i}) \)) such that \( x_{k,i} \rightarrow x_k \) (as \( i \rightarrow \infty \)), and

\[
\lim_{i \rightarrow \infty} f(x_{k,i}) = \lim_{i \rightarrow \infty} f(x_{k,i}).
\]

\(^1\) Refer to Condition (1') in the CV condition.
Since $R$ is closed at $y_k$ and $K$ is compact, it follows that $\bar{x}_k \in R(y_k) \cap K$. Therefore we have

$$\liminf_{j \to \infty} f_{\inf}(y_{kj}) = \liminf_{j \to \infty} f(x_{kj}) = \lim_{l \to \infty} f(x_{kl}) = f(\bar{x}_k) \geq f(\bar{y}_k),$$

where the first equality is owing to $x_{kj} \in P(y_{kj})$ and the second equality follows from Eq. (7).

Finally by the lower semi-continuity property of $f_{\inf}$ and $x_k \in R(y_k)$, we have for sufficiently large $j$,

$$f_{\inf}(y_k) \leq f_{\inf}(y_{kj}) \leq f(x_k) < f(x^*),$$

which completes the proof. □

The key assumption for Proposition 1 is that $R$ is l.s.c near $y^*$, and it is satisfied if the constraints are convex and weakly analytic in $x$ [c.f. Bank et al. (1982)].

3 Generalized Constraint Violation Conditions

The line of analysis of the CV condition is based on Problem (2) in which the variations of the resource constraints are measured by $y$, while the reason for this variation is neglected. In this section, we will extend the constraint violation condition to the generalized constraint violation (GCV) condition, which is used to derive informative LMs of Problem (1).

In Problem (2), the variation of resources is quantified as $y$. However, the reason for this variation is neglected. In fact, the variations of resources are caused by external factors (represented by the “price” for this resource tends to $\infty$), such as the perturbations in the problem data. As an example, consider the constraint function of the $i$th resource $g_i(x, y)$. Suppose the perturbation vector moves along the direction $d_y \in \mathbb{R}^m$ from $y^*$ to $y = y^* + d_y$, then the variation of $g_i(x, y)$ at $(x^*, y^*)$ (caused by the perturbation $y$) can be measured as

$$\delta(y) = - (g_i(x^*, y^*) - g_i(x^*, y^* - d_y))$$

$$= -\nabla_y g_i(x^*, y^*)^T d_y + o(||d_y||).$$

(8)

Hence, in general the value of the perturbation $y$ is not equal to the variation of resources. We note, however, if $g_i(x, y) = g_i(x) - y_i$ (as in Problem (2)) and $y_i^* = 0$, then the variation of $g_i(x, y)$ at $(x^*, y^*)$ is expressed as $y_i$. In this case (fully developed in classical literatures, e.g., Bertsekas and Ozdaglar (2002); Bertsekas et al. (2003)), the value of the perturbation $y$ is equal to the variation of the resource. Therefore, the perturbation vector $y$ is responsible for the resource variation, but it cannot be used to measure the variation. This allows a better understanding of the role of perturbation vectors. Suppose that the constraint imposed on the $i$th resource is $g_i(x)$ without any perturbations, which implies the scarcity in the supplement of the $i$th resource, it follows that the “price” for this resource tends to $\infty$. Consequently, the classical treatment of the resource violations may have serious conceptual limitations, since the resource constraints are always assumed to be violated in the classical line of analysis [c.f. Bertsekas and Ozdaglar (2002)]. As will be clear in the following sections, classical treatment of constraint violations may lead to an incorrect informative LM.

Let us now consider the role of the decision variable $x$ in Problem (1), and restrict our attention on the active constraints at $x^* \in P(y^*)$. We note that the inactive inequality constraints are inconsequential, since the subsequent analysis focuses in a small neighbourhood of $x^*$, within which these constraints remain inactive. For a given perturbation $y$, let

$$\tilde{R}(y) = \{ x | h_i(x, y) = 0, g_j(x, y) = 0, i = 1, \ldots, p, j \in \mathcal{A}(x^*) \}$$
be the “active feasibility set” of Problem (1) at \( y = y^* \). Suppose \( x(y) \in \tilde{R}(y) \) and \( x(y^*) = x^* \), then

\[
\begin{align*}
    h_i(x(y), y) - h_i(x^*, y^*) = \nabla_x h_i(x^*, y^*)^T (x(y) - x(y^*)) + \nabla_y h_i(x^*, y^*) (y - y^*) \\
    + o(\|x(y) - x(y^*)\|, \|y - y^*\|), \quad i = 1, \ldots, p, \\
    g_j(x(y), y) - g_j(x^*, y^*) = \nabla_x g_j(x^*, y^*)^T (x(y) - x(y^*)) + \nabla_y g_j(x^*, y^*) (y - y^*) \\
    + o(\|x(y) - x(y^*)\|, \|y - y^*\|), \quad j \in A(x^*).
\end{align*}
\]

Notice that if \( x(y) \) is locally Lipschitz near \( y^* \), then by using the fact that \( x(y) \in \tilde{R}(y) \), the variation of resources can also be measured as (calculated up to the first order)

\[
\delta(y)_i = \nabla_x h_i(x^*, y^*)^T (x(y) - x(y^*)), \quad i = 1, \ldots, p, \tag{9}
\]

and

\[
\delta(y)_{p+i} = \nabla_x g_j(x^*, y^*)^T (x(y) - x(y^*)), \quad j \in A(x^*) \tag{10}
\]

By comparing Eq. (9) and (10) with Eq. (8), we see that the quantity of resource variations can also be expressed by the change of decision variable \( x(y) \). Since the improved cost \( f(x(y), y) \) is required to check in the CV condition, the expression of resource variations can benefit from Eq. (9) and (10). Furthermore we note that the assumption \( x(y) \in \tilde{R}(y) \) can be replaced by \( \text{dist}(x(y), \tilde{R}(y)) = o(\|y - y^*\|) \).

Based on the preceding analysis above, we extend the notion of constraint violation condition to the fully parametric program. For completeness we introduce some notations. For a given \( d_y \in \mathbb{R}^m \), there exists a sequence \( \{y_k\} \) such that \( y_k \rightarrow y^* \), and \( \frac{y_k - y^*}{\|y_k - y^*\|} \rightarrow \frac{d_y}{\|d_y\|} \). Let

\[
\delta(y)_{i} = -\nabla_y h_i(x^*, y^*)^T d_y, \quad i = 1, \ldots, p,
\]

and

\[
\delta(y)_{p+l} = -\nabla_y g_j(x^*, y^*)^T d_y, \quad l = 1, \ldots, r,
\]

denote the resource variations induced by the perturbation vector \( y \) from \( y^* \) along the direction \( d_y \). Furthermore, set

\[
\delta(y)_{i}^{k} = \nabla_y h_i(x^*, y^*)^T \frac{y_k - y^*}{\|y_k - y^*\|}, \quad i = 1, \ldots, p, \tag{11}
\]

and

\[
\delta(y)_{p+l}^{k} = \nabla_y g_j(x^*, y^*)^T \frac{y_k - y^*}{\|y_k - y^*\|}, \quad l = 1, \ldots, r, \tag{12}
\]

and it is obvious that \( \delta(y)_{i}^{k} \rightarrow \delta(y)_{i} \) as \( k \rightarrow \infty \), \( i = 1, \ldots, p + r \).

**Definition 3** Let \( (\lambda^*, \mu^*) \in \mathbb{R}^{p+q} \) be an LM of Problem (1) corresponding to \( x^* \in P(y^*) \). Furthermore, let \( I^* = \{i | \lambda^*_i \neq 0, i = 1, \ldots, p\} \), \( J^* = \{j | \in A(x^*) | \mu^*_j > 0\} \). Assume \( I^* \cup J^* \neq \emptyset \). If there exists a continuous function \( x(\cdot): \mathcal{U} \rightarrow \mathbb{R}^n \) such that

1. \( x(y^*) = x^* \);
2. \( x(y) \) is locally Lipschitz continuous near \( y^* \),

and if in addition, there exists a sequence \( \{y_k\} \) with \( y_k \rightarrow y^* \) such that

\[
\frac{\text{dist}(x(y_k), \tilde{R}(y_k))}{\|y_k - y^*\|} \rightarrow 0 \quad \text{as} \quad k \rightarrow \infty; \tag{3}
\]

\[
\nabla_x f(x^*, y^*)^T (x(y_k) - x(y^*)) \rightarrow 0; \tag{4}
\]

for all \( i \in I^* \), \( \lambda^*_i \delta(y)_{i}^{k} > 0 \); and for all \( l \in J^* \), \( \mu^*_l \delta(y)_{p+l}^{k} > 0 \); and for all \( i \notin I^* \cup J^* \), \( |\delta(y)_{i}^{k}| = o(\omega(k)) \),

where \( \delta(y)_{i}^{k} \) (\( i = 1, \ldots, p + r \)) is defined by Eq. (11) and (12), and

\[
\omega(k) = \min_{i \in I^* \cup J^*} |\delta(y)_{i}^{k}|.
\]
then \((\lambda^*, \mu^*)\) is called the informative LM of Problem (1). Furthermore, the conditions restricted on \(x(\cdot)\) and \(\{y_k\}\) are called the generalized constraint violation condition (GCV condition).

Notice that if we set \(x_k = x(y_k)\) with \(y_k \neq y^*\) converging to the variation direction \(\frac{d_k}{\|d_k\|}\), then \(x_k\) satisfies the CV condition. Therefore, the “generalized” constraint violation in this sense is an extension of the CV condition. We also note that the third item of the GCV condition implies that the constraints of Problem (1) at \(x(y_k)\) are permitted to be slightly violated to the extent \(o(\|y_k - y^*\|)\). This is not surprising in view of the active feasibility set \(\hat{R}(y_k)\) may be empty for some \(y\) in the neighbourhood of \(y^*\).

We then explain why we use \(\sum f(x^*, y^*)^T (x(y_k) - x(y^*)) < 0\) in the fourth condition in Definition 3 to replace the argument \(f(x(y), y) < f(x^*, y^*)\). Let us consider the variations of the objective function around \(y^*\). By the second assumption on \(x(y)\) in Definition 3, we have for \(y\) in the neighbourhood of \(y^*\), the objective function \(f(x, y)\) at \(x(y)\) is expressed as

\[
f(x(y), y) = f(x^*, y^*) + \nabla_x f(x^*, y^*)^T (x(y) - x^*) + \nabla_y f(x^*, y^*)^T (y - y^*)
+ o(||x(y) - x(y^*)||, ||y - y^*||),
= f(x^*, y^*) + \nabla_x f(x^*, y^*)^T (x(y) - x^*) + \nabla_y f(x^*, y^*)^T (y - y^*)
+ o(||y - y^*||).
\]

(13)

From Eq. (13), the variations of the objective function are decomposed into two parts: one is the variations caused by exogenous variable, i.e.,

\[
\nabla_y f(x^*, y^*)^T (y - y^*),
\]

and the other is caused by the endogenous variable, i.e.,

\[
\nabla_x f(x^*, y^*)^T (x(y) - x^*),
\]

We note that the former variation should be excluded from consideration, since once the perturbation \(y\) is fixed, the variation of the objective function \(\nabla_y f(x^*, y^*)^T (y - y^*)\) is identically constant, and is not influenced by the adjustment of the decision variables.

4 Sufficient Conditions for Minimum Norm Lagrange Multiplier to be Informative

The MNLM may not convey the sensitivity information in fully parametric programming models. As an example, consider the following problem [c.f. Gauvin and Dubau (1984)]

\[
\begin{align*}
\min f(x, y) &= -x_2 \\
\text{s.t.} \ g_1(x, y) &= x_2 - x_1^2 \leq 0, \\
g_2(x, y) &= x_2 + x_1^2 - y \leq 0.
\end{align*}
\]

At \(y^* = 0\), the unique optimum is \(x^* = (0, 0)^T\) with the set of LMs

\[
M(x^*, y^*) = \{(\mu_1, \mu_2) | \mu_1 + \mu_2 = 1, \mu_1 \geq 0, \mu_2 \geq 0\}.
\]

Obviously the minimum norm LM is \(\mu_1^* = \mu_2^* = \frac{1}{2}\), implying that there exists some \(y\) and some corresponding feasible solution \(x(y)\) such that the increment of the constraints \(g_1(x, y)\) and \(g_2(x, y)\) may lead to the cost improvement (calculated up to first order); that is,

\[
\nabla_y g_i(x^*, y^*)^T (y - y^*) > 0, i = 1, 2,
\]

and

\[
\nabla_x f(x^*, y^*)^T (x(y) - x^*) < 0.
\]
However, \( \nabla_y g_1(x^*, y^*)^T (y - y^*) \) is identically zero, implying that MNLM is not informative.

This counterexample inspires us to develop the sufficient conditions to guarantee a certain LM to be informative. To get a sense of the main idea, consider the following problem

\[
\min_{d \in \mathbb{R}^n} \nabla_x f(x^*, y^*)^T d + \frac{1}{2} \sum_{i=1}^{p} (\nabla_x h_i(x^*, y^*)^T d)^2 + \frac{1}{2} \sum_{j \in A(x^*)} ((\nabla_x g_j(x^*, y^*)^T d)^+)^2.
\] (15)

The Lagrange dual of Problem (15) is

\[
\begin{align*}
\min_{(\lambda, \mu) \in \mathbb{R}^{p+2}} & \frac{1}{2} \| (\lambda, \mu) \|^2 \\
\text{s.t.} & \quad \nabla_x f(x^*, y^*) + \sum_{i=1}^{p} \lambda_i \nabla_x h_i(x^*, y^*) + \sum_{j=1}^{q} \mu_j \nabla_x g_j(x^*, y^*) = 0 \\
& \quad \mu \geq 0.
\end{align*}
\] (16)

Assume the set of LMs at \( x^* \in P(y^*) \) is nonempty. Then by Proposition 6.4.1 and 6.4.2 in Bertsekas et al. (2003), strong duality holds for Problem (15) and (16), and the optimal solution sets of both problems are also nonempty. Furthermore, let \( d^* \) and \( (\lambda^*, \mu^*) \) be the optimal solution of Problem (15) and (16), respectively, then by the optimality condition we have

\[
\lambda_i^* = \nabla_x h_i(x^*, y^*)^T d^*, \quad i = 1, \ldots, p,
\] (17)

\[
\mu_j^* = (\nabla_x g_j(x^*, y^*)^T d^*)^+, \quad j \in A(x^*),
\] (18)

and

\[
\nabla_x f(x^*, y^*)^T d^* = -\| (\lambda^*, \mu^*) \|^2.
\] (19)

We will use the optimal solution \( d^* \) of Problem (15) to construct the sequence of active feasible solution \( x(y) \), and build the relationship between the MNLM and the informative LM. To facilitate notation, we assume \( A(x^*) = \{j_1, \ldots, j_r\} \).

**Theorem 1** Let

\[
\delta_i = \nabla_x h_i(x^*, y^*)^T d^*, \quad i = 1, \ldots, p,
\] (20)

\[
\delta_{p+i} = (\nabla_x g_j(x^*, y^*)^T d^*)^+, \quad j_i \in A(x^*),
\] (21)

where \( d^* \) is the optimal solution of Problem (15). Furthermore, let

\[
A = \begin{pmatrix}
\nabla_y h_1(x^*, y^*)^T \\
\vdots \\
\nabla_y h_p(x^*, y^*)^T \\
\nabla_y g_{j_1}(x^*, y^*)^T \\
\vdots \\
\nabla_y g_{j_r}(x^*, y^*)^T
\end{pmatrix},
\]

and \( \delta = (\delta_1, \ldots, \delta_{p+r}) \). Suppose that \( \delta \) is in the range of the matrix A (i.e., \( \delta \in \text{Range}(A) \)). Then there exists a continuous function \( x(y) \) satisfying the GCV condition, and the optimal solution of Problem (16) is an informative LM.
Proof Since by assumption \( \delta \in \text{Range}(A) \), then there exists a vector \( d_y \in \mathbb{R}^m \) such that \( \delta = Ad_y \). Therefore there exists a sequence \( \{y_k\} \) with \( y_k \to y^* \) such that \( \frac{y_k - y^*}{\|y_k - y^*\|} \to \frac{d_y}{\|d_y\|} \), and it follows that

\[
\delta_i = -\nabla_y h_i(x^*, y^*)^T \frac{y_k - y^*}{\|y_k - y^*\|} + o(1), \quad i = 1, \ldots, p,
\]

(22)

and

\[
\delta_{p+i} = -\nabla_y g_{ji}(x^*, y^*)^T \frac{y_k - y^*}{\|y_k - y^*\|/\|d_y\|} + o(1), \quad ji \in A(x^*).
\]

(23)

Thus, we set \( x(y) = x^* + \|y - y^*\| d_y \), and we will show that \( x(y) \) satisfies the GCV condition.

Firstly, it is easy to see that \( x(y^*) = x^* \) and \( x(y) \) is locally Lipschitz near \( y^* \), and the first two conditions of GCV satisfies. Then we will show that the distance from \( x(y) \) to \( \hat{R}(y) \) is \( o(||y - y^*||) \).

By using \textbf{Eq}s (20) to (23), we have for each \( i = 1, \ldots, p \),

\[
\frac{h_i(x(y_k), y_k)}{\|y_k - y^*\|/\|d_y\|} = \nabla_x h_i(x^*, y^*)^T d^* + \nabla_y h_i(x^*, y^*)^T \frac{y_k - y^*}{\|y_k - y^*\|/\|d_y\|} + o(1),
\]

(24)

\[
= \delta_i + \nabla_y h_i(x^*, y^*)^T \frac{y_k - y^*}{\|y_k - y^*\|/\|d_y\|} + o(1),
\]

(25)

\[
= o(1),
\]

(26)

and for each \( ji \in A(x^*) \),

\[
\frac{g_{ji}(x(y_k), y_k)}{\|y_k - y^*\|/\|d_y\|} = \nabla_x g_{ji}(x^*, y^*)^T d^* + \nabla_y g_{ji}(x^*, y^*)^T \frac{y_k - y^*}{\|y_k - y^*\|/\|d_y\|} + o(1),
\]

(27)

\[
= \delta_{p+i} + \nabla_y g_{ji}(x^*, y^*)^T \frac{y_k - y^*}{\|y_k - y^*\|/\|d_y\|} + o(1),
\]

(28)

\[
= o(1),
\]

(29)

implying that the distance from \( x(y_k) \) to \( \hat{R}(y) \) is \( o(||y_k - y^*||) \).

Then we will show that the sequence \( x(y) \) satisfies the rest conditions of GCV. Since by the continuity property of \( x(y) \) and \( y_k \to y^* \), we have for each \( k \)

\[
\nabla_x f(x^*, y^*)^T d^* = \nabla_x f(x^*, y^*)^T \frac{y_k - y^*}{\|y_k - y^*\|/\|d_y\|} + o(1).
\]

(30)

Using the \textbf{Eq}. (17), (18) and (19), we have

\[
\nabla_x f(x^*, y^*)^T \frac{(x(y_k) - x(y^*))}{\|y_k - y^*\|} = \nabla_x f(x^*, y^*)^T \frac{d^*}{\|d_y\|} + o(1),
\]

(31)

\[
\leq \frac{||\lambda^*||^2}{\|d_y\|^2} + o(1),
\]

(32)

Similarly, we have for \( i \in I^* \) \( \{i \in \{1, \ldots, p\} | \lambda^*_i \neq 0 \} \),

\[
\lambda^*_i \delta_i = (\lambda^*_i)^2 > 0,
\]

(33)

and for \( ji \in J^* \) \( \{ji \in A(x^*) | \mu^*_ji > 0 \} \),

\[
\mu^*_ji \delta_{p+i} = (\mu^*_ji)^2 > 0,
\]

(34)

For \( i \notin I^* \), and any \( \epsilon > 0 \),

\[
\delta_i = \nabla_x h_i(x^*, y^*)^T d^* = 0,
\]

(35)

and for \( ji \in A(x^*) \setminus J^* \),

\[
\delta_{p+i} = (\nabla_x g_{ji}(x^*, y^*)^T d^*)^+ = 0.
\]

(36)

From the relations above, we see that \( x(y) \) and the sequence \( \{y_k\} \) can be used to establish the GCV condition for \( (\lambda^*, \mu^*) \), so \( (\lambda^*, \mu^*) \) is an informative LM. 

\(\square\)
Note that as a corollary of Theorem 1, if
\[
\text{rank} \begin{pmatrix}
\nabla_y h_1(x^*, y^*)^T \\
\vdots \\
\nabla_y h_p(x^*, y^*)^T \\
\nabla_y g_{j_1}(x^*, y^*)^T \\
\vdots \\
\nabla_y g_{j_m}(x^*, y^*)^T
\end{pmatrix} = m,
\]
then the optimal solution of Problem (16) is informative. This corollary explains why the MNLM of Problem (2) is an informative LM. For active constraints in Problem (2), the partial derivatives of the constraint functions with respect to \(y\) at \((x^*, y^*)\) is
\[
\nabla_y h_i(x^*, y^*)^T = -e_i, i = 1, \ldots, p,
\]
and
\[
\nabla_y g_{j_i}(x^*, y^*)^T = -e_{p+i}, j_i \in A(x^*),
\]
where \(e_i\) is the unit vector with the \(i\)th element being one and the rest elements being zero. Therefore, the matrix \(A\) is equal to the \(m + r\) - dimensional identity matrix, and the corollary follows.

5 Penalty Function Method to Compute the Informative Lagrange Multipliers

In this section, we propose a penalty function method to compute the informative LM in Problem (1). Throughout this section, we assume the perturbation vector is varied from \(y^*\) to \(\tilde{y}\). Furthermore, we assume the resource variations induced by the perturbation \(y = \tilde{y}\) are measured as
\[
\delta(\tilde{y})_i = -\nabla_y h_i(x^*, y^*)^T (\tilde{y} - y^*), i = 1, \ldots, p,
\]
and
\[
\delta(\tilde{y})_{p+l} = -\nabla_y g_{j_i}(x^*, y^*)^T (\tilde{y} - y^*), l = 1, \ldots, r,
\]
where \(\delta(\tilde{y})_{p+l} > 0, l = 1, \ldots, r\). For notational convenience, we introduce the following notation. For any \(d \in \mathbb{R}^n\) and \(y \in \mathbb{R}^m\), we set
\[
a^h(d)_i = \nabla_x h_i(x^*, y^*)^T d, i = 1, \ldots, p,
a^g(d)_j = \nabla_x g_{j_i}(x^*, y^*)^T d, j = 1, \ldots, q.
\]
For any perturbation vector \(y\), consider the following problem
\[
\min_d q(d) = \nabla_x f(x^*, y^*)^T d + \frac{1}{2} \sum_{i=1}^{p} (a^h(d)_i)^2 + \frac{1}{2} \sum_{j \in A(x^*)} (a^g(d)_j)^2
\]
\text{s.t.} \quad a^h(d)_i - \delta(\tilde{y})_i = 0, i = 1, \ldots, p, \\
\quad (a^g(d)_j)^+ - \delta(\tilde{y})_j = 0, j_i \in A(x^*). 
\]
(24)

We assume the optimal solution set of Problem (24) is nonempty, and moreover there always exists a minimum point \(d^*(\tilde{y})\) of Problem (24) such that \(\nabla q(d^*(\tilde{y})) = 0\). This assumption is essential in developing informative LMs, and it is assumed throughout this section.

Applying the penalty function method to Problem (24), it yields
\[
\min_d L_c(d) = \nabla_x f(x^*, y^*)^T d + \frac{1}{2} \sum_{i=1}^{p} (a^h(d)_i)^2 + \frac{1}{2} \sum_{j \in A(x^*)} (a^g(d)_j)^2 + \frac{c}{2} P(d)
\]
(25)
where
\[ P(d) = \sum_{i=1}^{p} \left( a^h(d_i) - \delta(y)_i \right)^2 + \sum_{j \in A(x^*)} (a^g(d)_j - \delta(y)_j)^2, \]
and \( c \) is the penalty parameter. Note that if \( c = 0 \), then Eq. (25) will fall into the category of
Eq. (15). As will be clear in Proposition 3, the penalty term \( cP(d) \) converges to zero as \( c \to \infty \). Bearing this in mind, for sufficiently large penalty parameter \( c \) in Problem (25), \( a^g(d)_j \) is positive
\((l = 1, \ldots, r)\) and the dual problem derived as
\[
\min_{\lambda, \mu} \frac{1}{2} \|(\lambda, \mu)\|^2 - \frac{c}{2(\epsilon + 1)} \left( \sum_{i=1}^{p} (\lambda_i - \delta(y)_i)^2 + \sum_{j \in A(x^*)} (\mu_j - \delta(y)_j)^2 \right)
\]
s.t. \( \nabla_x f(x^*, y^*) + \sum_{i=1}^{p} \lambda_i \nabla_x h_i(x^*, y^*) + \sum_{j=1}^{q} \mu_j \nabla_x g_j(x^*, y^*) = 0 \)
\( \mu \geq 0 \).
Assume the set of LMs at \( x^* \in P(y^*) \) (i.e., \( M(x^*) \)) is nonempty. Then by the nonlinear Farkas’
\( x \) and \( (\lambda_*(\tilde{y}), \mu_*(\tilde{y})) \) be the
\( \lambda_*(\tilde{y}) = c \left( a^h(d_*(\tilde{y}))_i - \delta(\tilde{y})_i \right) + a^h(d_*(\tilde{y}))_i, \ i = 1, \ldots, p. \]
\( \mu_*(\tilde{y})_j = c \left( a^h(d_*(\tilde{y}))_j - \delta(\tilde{y})_p+i \right) + (a^g(d_*(\tilde{y}))_j)^+, \ j \in A(x^*), \]
\( \mu_*(y)_j = 0, j \notin A(x^*), \)
\[
\nabla_x f(x^*, y^*)^T d_*(\tilde{y}) = -\|(\lambda_*(\tilde{y}), \mu_*(\tilde{y}))\|^2 + c \left( \sum_{i=1}^{p} (a^h(d_*(\tilde{y}))_i - \delta(\tilde{y})_i)^2 + \sum_{i=1}^{p} (a^g(d_*(\tilde{y}))_j - \delta(\tilde{y})_j)^2 \right) +
\]
\[
c \left( \sum_{i=1}^{p} a^h(d_*(\tilde{y}))_i (a^h(d_*(\tilde{y}))_i - \delta(\tilde{y})_i) + \sum_{i=1}^{p} (a^g(d_*(\tilde{y}))_j)^+(a^g(d_*(\tilde{y}))_j - \delta(\tilde{y})_j) \right).
\]
The penalty function method to compute the informative LM is developed as follows.

**Algorithm 1** Penalty Function Method

Given start point \( d_0 \in \mathbb{R}^n \), tolerance \( \epsilon > 0 \), \( c_1 > 0 \), \( \gamma > 1 \). Set \( k = 1 \).

Repeat
1. Solving Problem (25) with \( c_k \) and \( d_{k-1} \) as the initial point. Let the minimum point be \( d_k \).
2. Set \( c_{k+1} = \gamma c_k \), \( k = k+1 \).

Until \( c_k P(d_k) \leq \epsilon \).

Output \( d(\tilde{y}) = d_k \).

We first derive some preliminary results based on the penalty function method in the following
two propositions.

**Proposition 2** Let \( \{d_k\} \) be a sequence generated by the penalty function method. Then we have
\[
L_{c_{k+1}}(d_{k+1}) \geq L_{c_k}(d_k),
\]
\[
P(d_{k+1}) \leq P(d_k),
\]
\[
L_{c_{k+1}}(d_{k+1}) - P(d_{k+1}) \geq L_{c_k}(d_k) - P(d_k).
\]
Proof Noticing that $c_{k+1} > c_k > 0$, it follows that

\[
L_{c_{k+1}}(d_{k+1}) = \nabla_x f(x^*, y^*)^T d_{k+1} + \frac{1}{2} \sum_{i=1}^{p} \left( a^h(d_{k+1}) \right)^2 + \frac{1}{2} \sum_{j \in A(x^*)} \left( a^g(d_{k+1}) \right)^2 + \frac{c_{k+1}}{2} P(d_{k+1}) \\
\geq \nabla_x f(x^*, y^*)^T d_{k+1} + \frac{1}{2} \sum_{i=1}^{p} \left( a^h(d_{k+1}) \right)^2 + \frac{1}{2} \sum_{j \in A(x^*)} \left( a^g(d_{k+1}) \right)^2 + \frac{c_k}{2} P(d_{k+1}) \\
= L_{c_k}(d_{k+1}) \geq L_{c_k}(d_k),
\]

which proves Eq. (31). Furthermore, since by assumption that $d_k$ and $d_{k+1}$ minimize $L_{c_k}(d)$ and $L_{c_{k+1}}(d)$, respectively, we have

\[
L_{c_k}(d_{k+1}) \geq L_{c_k}(d_k), \\
L_{c_{k+1}}(d_k) \geq L_{c_{k+1}}(d_{k+1}).
\]

Then, by combining the two inequalities above, we have

\[
(c_{k+1} - c_k) (P(d_k) - P(d_{k+1})) \geq 0,
\]

which yields Eq. (32).

Eq. (33) directly follows from Eq. (31) and (32). \qed

Proposition 3 Let $\{d_k\}$ be a sequence generated by the penalty function method, and $d^*(\tilde{y})$ be the global minimum of Problem (24). Assume $c_k \to \infty$ as $k \to \infty$, and the sequence $\{d_k\}$ has a limit point $d_{\infty}$, then

\[
\lim_{k \to \infty} P(d_k) = 0, \\
\lim_{k \to \infty} c_k P(d_k) = 0.
\]

Proof Without loss of generality, we assume $d_k \to d_{\infty}$ as $k \to \infty$. First, it is implied from Proposition 2 that $L_{c_k}(d_k)$ is monotonically nondecreasing with respect to $k$. Furthermore, since $d_k$ minimizes Problem (25) with $c = c_k$, we have

\[
L_{c_k}(d_k) \leq L_{c_k}(d(\tilde{y})) = q(d(\tilde{y})),
\]

hence $L_{c_k}(d_k)$ converges to limits $L_{\infty}$. Furthermore, we notice that the sequence $\{L_{c_k}(d_k) - P(d_k)\}$ (or $\{q(d_k)\}$) is also monotonically nondecreasing, and for all $k,$

\[
L_{c_k}(d_k) - P(d_k) \leq L_{c_k}(d_k) \leq L_{c_k}(d(\tilde{y})) = q(d(\tilde{y})),
\]

it follows that $\{q(d_k)\}$ also converges to limits $q_{\infty}$. Therefore we have

\[
\lim_{k \to \infty} c_k P(d_k) = \lim_{k \to \infty} c_k (L_{c_k}(d_k) - q(d_k)) = L_{\infty} - q_{\infty}.
\]

Since $c_k \to \infty$, then we have $\lim_{k \to \infty} P(d_k) = 0$, and Eq. (34) follows. Furthermore, by continuity of the function $P(d)$ and $q(d)$, we have

\[
0 = \lim_{k \to \infty} P(d_k) = P(d_{\infty}),
\]

\[
q(d_{\infty}) = \lim_{k \to \infty} q(d_k) \leq q(d(\tilde{y})),
\]

then by the assumption on $d^*(\tilde{y})$, it follows that $d_{\infty}$ is equal to $d^*(\tilde{y})$, and hence

\[
\lim_{k \to \infty} c_k P(d_k) = L_{\infty} - q_{\infty} = 0.
\]

\qed
The preceding two propositions imply that the term
\[ c_k \left( a^h(d_{cs}(y))_i - \delta(y)_i \right), \ i = 1, \ldots, p \]
in Eq. (27), and the term
\[ c_k \left( a^q(d_{cs}(y))_{ji} - \delta(y)_{ji} \right), \ ji \in A(x^*) \]
will all tend to zero as \( k \to \infty \). In other words, taking the output vector \( d(y) \) of the penalty function method into Eq. (27) to (30), it yields the Lagrange multipliers
\[ \lambda(\hat{y})_i = a^h(d(\hat{y}))_i, \ i = 1, \ldots, p, \]  
\[ \mu(\hat{y})_{ji} = (a^q(d(\hat{y}))^+_{ji}, j_i \in A(x^*), \]  
\[ \mu(y)_j = 0, j \notin A(x^*), \]  
and the cost improvement (calculated up to first order)
\[ \nabla_x f(x^*, y^*)^T d(\hat{y}) = -\|\lambda(\hat{y}), \mu(\hat{y})\|^2. \]  
We note that by Eq. (34) in Proposition 3, the resources variations of each constraint (i.e., \( \delta(y)_i, i = 1, \ldots, p + q \)) can be expressed by its partial derivative with respect to the decision vector \( x \); that is,
\[ \delta(y)_i = a^h(d(y))_i, i = 1, \ldots, p, \]  
\[ \delta(y)_{p+i} = (a^q(d(y))^+_{ji}, j_i \in A(x^*), \]  
then it follows by Eq. (36), (37) and (38) that the sign of each LM (obtained by the penalty function method) is consistent with the variation of resources. Next we will use the output vector \( d(y) \) of the penalty function method, to construct the function \( x(\cdot) \) that satisfying the GCV condition in Section 3. We will first prove that the optimal solution of Problem (25) \( d(y) \) is locally Lipschitz near \( y^* \).

**Theorem 2** Let \( P^E(y) \) and \( U^E \) respectively denotes the set of the optimal solution and the set of the feasible solution of Problem (25). For \( y \in U^E \cup B(y^*; r) \) with \( r > 0 \), and let \( d(y) \in P^E(y) \). Then \( d(y) \) is locally Lipschitz continuous near \( y^* \).

**Proof** Without loss of generality, we assume \( A(x^*) = \{j_1, \ldots, j_r\} \). First we notice \( d(y) \) solves Problem (25), which can be reformulated as
\[ \min_{(d, t, s, u, v, l) \in \mathbb{R}^{n+r+p}} \nabla_x f(x^*, y^*)^T d + \frac{1}{2} \sum_{i=1}^p t_i^2 + \frac{1}{2} \sum_{l=1}^r s_l^2 + \frac{\xi}{2} \sum_{i=1}^p u_i^2 + \frac{\xi}{2} \sum_{l=1}^r v_l^2 \]  
\[ \nabla_x h_i(x^*, y^*)^T d \leq t_i, i = 1, \ldots, p \]  
\[ \nabla_x g_{py}(x^*, y^*)^T d \leq s_l, l = 1, \ldots, r \]  
\[ \nabla_x h_i(x^*, y^*)^T d \leq u_i + \delta(y)_i, i = 1, \ldots, p \]  
\[ \nabla_x g_{py}(x^*, y^*)^T d \leq v_l + \delta(y)_{p+l}, l = 1, \ldots, r \]  
\[ s_l \geq 0, l = 1, \ldots, r, \]
where \( t_i, s_l, u_i, v_l \) are auxiliary variables. Furthermore, using the following transformation
\[ d = d_1 - d_2, d_1 \geq 0, d_2 \geq 0, \]
\[ t = t_1 - t_2, t_1 \geq 0, t_2 \geq 0, \]
\[ u = u_1 - u_2, u_1 \geq 0, u_2 \geq 0, \]
\[ v = v_1 - v_2, v_1 \geq 0, v_2 \geq 0, \]
\[ x = (d_1, d_2, t_1, t_2, s, u_1, u_2, v_1, v_2), \]
\[ C = \begin{pmatrix} 0_{2n \times 2n} & I_{(2p+r)} \end{pmatrix} \sqrt{A_{(p+r)}}, \]

\[ w^T = (\nabla_x f(x^*, y^*)^T, -\nabla_x f(x^*, y^*)^T, 0_{1 \times (4p+3r)}), \]

\[ A = \begin{pmatrix} \nabla_x h_1(x^*, y^*)^T, -\nabla_x h_1(x^*, y^*)^T, -e_1^T, e_1^T, 0_{1 \times r} 0_{1 \times p} 0_{1 \times p} 0_{1 \times r} 0_{1 \times r} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
\nabla_x h_p(x^*, y^*)^T, -\nabla_x h_p(x^*, y^*)^T, -e_p^T, e_p^T, 0_{1 \times r} 0_{1 \times p} 0_{1 \times p} 0_{1 \times r} 0_{1 \times r} \\
\nabla_x g_{j_1}(x^*, y^*)^T, -\nabla_x g_{j_1}(x^*, y^*)^T, 0_{1 \times p} 0_{1 \times p} 0_{1 \times r} -e_1^T, e_1^T 0_{1 \times r} 0_{1 \times r} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
\nabla_x g_{j_p}(x^*, y^*)^T, -\nabla_x g_{j_p}(x^*, y^*)^T, 0_{1 \times p} 0_{1 \times p} 0_{1 \times r} -e_p^T, e_p^T 0_{1 \times r} 0_{1 \times r} \\
\nabla_x g_{j_{p+1}}(x^*, y^*)^T, -\nabla_x g_{j_{p+1}}(x^*, y^*)^T, 0_{1 \times p} 0_{1 \times p} 0_{1 \times r} -e_1^T, e_1^T 0_{1 \times r} 0_{1 \times r} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
\nabla_x g_{j_{p+r}}(x^*, y^*)^T, -\nabla_x g_{j_{p+r}}(x^*, y^*)^T, 0_{1 \times p} 0_{1 \times p} 0_{1 \times r} -e_r^T, e_r^T 0_{1 \times r} 0_{1 \times r} \\
\end{pmatrix}, \]

\[ b = \begin{pmatrix} \delta(y)_1 \\
\vdots \\
\delta(y)_{p+r} \end{pmatrix}, \]

Problem (42) can be written as a standard quadratic programming

\[
\begin{align*}
\min_x & \quad \frac{1}{2} x^T C x + w^T x \\
\text{s.t.} & \quad Ax \leq b \\
x & \geq 0.
\end{align*}
\] (43)

with \( C \) being positive semidefinite, and the constraint’s right hand side vector \( b \) being \( y \)-perturbed. By using the equivalence between the quadratic programming and the linear complementarity problem\( (\text{LCP}) \), the solution set of Problem (43) is equivalent to the solution set of the following \( \text{LCP} \):

\[ Mz + q \geq 0, \quad z \geq 0, \quad z^T(Mz + q) = 0, \] (44)

with the vector \( q \) perturbed by \( y \). Here \( M = \begin{pmatrix} C & A^T \\ -A & 0 \end{pmatrix}, q^T = (w^T, b^T), z^T = (x^T, \xi^T), \) and \( \xi \) is the multiplier corresponding to the constraint \( Ax \leq b \). Let \( P^{\text{LCP}}(y) \) denotes the solution set of Problem (44), then according to Theorem 7.2.1 in Cottle et al. (2009) we have for any \( y \in U^E \cup B(y^*; r) \) there exists a constant \( L > 0 \) such that

\[ P^{\text{LCP}}(y) \subseteq P^{\text{LCP}}(y^*) + L\|y - y^*\|B(0, 1) \]

Since \( P^{\text{LCP}}(y^*) = \{z(y^*) = (x^*, \xi)(x^*, \xi) \text{satisfies Problem (44)}\} \), then for any \( z(y) = (x(y), \xi(y)) \) we have \( \|z(y) - z(y^*)\| \leq L\|y - y^*\| \), and it follows that \( \|x(y) - x^*\| \leq L\|y - y^*\| \), which shows the locally Lipschitz continuity of \( x(y) \). Since \( x(y) = (d_1, d_2, t_1, t_2, s, u_1, u_2, v_1, v_2) \) and the \( d(y) = d_1 - d_2, \)
it follows that \( d(y) \) is also locally Lipschitz continuous with respect to \( y \in U^E \cap B(y^*, r) \). \( \square \)

We then use the optimal solution \( d(\bar{y}) \) of Problem (25) with \( y = \bar{y} \) to construct the function \( x(\cdot) \) and the sequence \( \{y_k\} \) satisfying the GCV condition.

**Theorem 3** Let \( x^* \in P(y^*) \), and \( y_k \to y^* \) such that \( \frac{y_k - y^*}{\|y_k - y^*\|} \to \frac{\bar{y} - y^*}{\|\bar{y} - y^*\|} \). Set \( x(y) = x^* + \frac{\|y - y^*\|}{\|\bar{y} - y^*\|} d(\bar{y}) \), where \( d(\bar{y}) \) is the optimal solution of Problem (25) with \( y = \bar{y} \). Then the function \( x(\cdot) \) and the sequence \( \{y_k\} \) satisfy the GCV condition, and the optimal solution of Problem (26) is an informative LM.

**Proof** In order to prove \( x(\cdot) \) and \( \{y_k\} \) satisfy the GCV condition, we need to show that \( x(\cdot) \) and \( \{y_k\} \) satisfy the assumptions in Definition 3. It is easy to see that \( x(y^*) = x^* \), and the first condition of GCV satisfies. Then we prove that \( x(\cdot) \) is locally Lipschitz near \( y^* \). By Theorem 2, there exists an \( L \) such that
\[
\|d(y)\| \leq L\|y - y^*\|, \forall y \in B(y^*, r),
\]
for some \( r > 0 \), and it follows that
\[
\|x(y) - x(y^*)\| = \frac{\|d(\bar{y})\|\|y - y^*\|}{\|y - y^*\|} \leq L\|y - y^*\|,
\]
implying \( x(\cdot) \) is locally Lipschitz near \( y^* \). Furthermore, by the locally Lipschitz continuity of \( x(\cdot) \) and the assumptions on \( \{y_k\} \) in Theorem 3, we have for \( i = 1, \ldots, p, \)
\[
\frac{h_i(x(y_k), y_k)}{\|y_k - y^*\|/\|y - y^*\|} = \nabla_x h_i(x^*, y^*)^T \frac{x(y_k) - x^*}{\|y_k - y^*\|/\|y - y^*\|} + \nabla_y h_i(x^*, y^*)^T \frac{y_k - y^*}{\|y_k - y^*\|/\|y - y^*\|} + o(1),
\]
where the second equality follows \( \frac{y_k - y^*}{\|y_k - y^*\|} \to \frac{\bar{y} - y^*}{\|\bar{y} - y^*\|} \) as \( k \to \infty \), and the last equality holds since \( a^h(d(\bar{y}))_i = \delta(\bar{y})_i \). Similarly, we also have for all \( j_i \in A(x^*), \frac{g_j(x(y_k), y_k)}{\|y_k - y^*\|/\|y - y^*\|} \to 0(\text{as} k \to \infty) \), which implies that the distance from \( x(y_k) \) to \( \hat{R}(y_k) \) is \( o(\|y_k - y^*\|) \).

Then we will show that the sequence \( \{x(y_k)\} \) satisfies the rest conditions of GCV. Since by **Eq.** (39), we have
\[
\nabla_x f(x^*, y^*)^T (x(y_k) - x(y^*)) = \nabla_x f(x^*, y^*)^T d(\bar{y}) = -\frac{\|\lambda(\bar{y}, u(\bar{y}))\|^2}{\|\bar{y} - y^*\|} < 0.
\]

Similarly, by **Eq.** (36), (37), (40) and (41), we have for \( i \in I(\bar{y}) = \{i \in \{1, \ldots, p\} | \lambda(\bar{y})_i \neq 0\}, \)
\[
\mu(\bar{y})_i, \delta(\bar{y})_i = \lambda(\bar{y})_i^2 > 0,
\]
and for \( j \in J(\bar{y}) = \{j_i \in A(x^*) | \mu(\bar{y})_j_i > 0\}, \)
\[
\mu(\bar{y})_j_i, \delta(\bar{y})_{j_i} = \mu(\bar{y})_j_i^2 > 0,
\]
For \( i \notin I(\bar{y}), \)
\[
\delta(\bar{y})_i = a^h(d(\bar{y}))_i = 0,
\]
and for \( j_i \in A(x^*) \setminus J(\bar{y}), \)
\[
\delta(\bar{y})_{j_i} = (a^h(d(\bar{y}))_{j_i})^+ = 0.
\]
From the relations above, we see that the sequence \( \{x(y_k)\} \) can be used to establish the GCV condition for \( \lambda^*, \mu^* \), so \( (\lambda^*, \mu^*) \) is an informative LM. \( \square \)
Let us remark that Theorem 3 may fail without the assumption which is introduced at the beginning of this section. First we will show that Theorem 3 fails when the solution set of Problem (24) is empty. Consider the preceding example (i.e., Eq. (14)). Suppose the perturbation $y$ in Eq. (14) is changed from $y^* = 0$ to $\tilde{y} = 1$, and the variation of the two constraints are $\delta(\tilde{y})_1 = 0$ and $\delta(\tilde{y})_2 = 1$. By using the penalty function method to solve Eq. (14) we obtain $d(\tilde{y})_1 \in \mathbb{R}$ and $d(\tilde{y})_2 = \frac{1}{2}$, from which the corresponding LM is

$$\mu(\tilde{y})_1 = (\nabla_x g_1(x^*, y^*)^T d(\tilde{y}))^+ = \frac{1}{2},$$

and

$$\mu(\tilde{y})_2 = (\nabla_x g_2(x^*, y^*)^T d(\tilde{y}))^+ = \frac{1}{2},$$

which is not an informative LM. The reason is that the value of penalty term $P(d(\tilde{y}))$ is

$$\left(0, 1\right) \left(\frac{d(\tilde{y})_1}{d(\tilde{y})_2}\right)^2 + \left(0, 1\right) \left(\frac{d(\tilde{y})_1}{d(\tilde{y})_2} - \tilde{y}\right)^2 = \frac{1}{2},$$

which implies that the distance from $\{x(y_k)\}$ (i.e., $x(y_k) = x^* + \frac{\parallel y_k - y^* \parallel}{\parallel y_k - y^* \parallel} d(\tilde{y})$) to $\tilde{R}(y_k)$ is larger than $o(\parallel y_k - y^* \parallel)$. Indeed, put $x(y_k) = \left(\frac{\parallel y_k \parallel}{\parallel y_k \parallel} d(\tilde{y})_1\right)$ into the two constraints in Eq. (14), it follows that

$$g_1(x(y_k), y_k) = \frac{1}{2} \parallel y_k \parallel - (d(\tilde{y})_1 y_k)^2,$$

and

$$g_2(x(y_k), y_k) = \frac{1}{2} \parallel y_k \parallel + (d(\tilde{y})_1 y_k)^2 - y_k,$$

neither of which is equal to $o(\parallel y_k \parallel)$.

There are some preprocessing techniques can be applied in case where the assumption of Theorem 3 fails. Taking Eq. (14) as an example. Since the constraint $g_1$ is stable (not perturbed by vector $y$), and it implies that any cost improvement is independent of the variation of $g_1$, and the corresponding LM with respect to $g_1$ must be 0, and $g_1$ can be removed from further analysis in the penalty function. In this case, $\delta(\tilde{y}) = 1$ is in the range of $(\nabla_x g_2(x^*, y^*)^T)$, and the assumption of Theorem 25 satisfies. It follows that the penalty function of Eq. (14) is

$$P(d(\tilde{y})) = \left(0, 1\right) \left(\frac{d(\tilde{y})_1}{d(\tilde{y})_2} - \tilde{y}\right)^2,$$

and the optimal solution of Eq. (25) with $y = \tilde{y}$ is

$$d(\tilde{y})_1 \in \mathbb{R},$$

and

$$d(\tilde{y})_2 = \frac{c + 1}{c + 2} \to 1 \text{ (as } c \to \infty).$$

Therefore the corresponding LM is

$$\mu(\tilde{y})_1 = 0,$$

and

$$\mu(\tilde{y})_2 = (\nabla_x g_2(x^*, y^*)^T d(\tilde{y}))^+ = 1,$$

which is informative.
Next we show that Theorem 3 fails if the gradient of \( q(d) \) in Problem (24) at \( d^*(\tilde{y}) \) is not equal to zero. Consider the following problem

\[
\begin{align*}
\min_{x} & \quad -x \\
\text{s.t.} & \quad x - y \leq 0,
\end{align*}
\]

where \( x \) is the decision variable and \( y \) is the perturbation. Suppose \( y^* = 0 \), then \( x^* = 0 \) is the unique minimum at \( y = y^* \). Let the perturbation \( \tilde{y} > 0 \) and \( \tilde{y} \neq 1 \), it follows that the variation of the constraint is \( \delta(\tilde{y}) = \tilde{y} \). Applying the penalty function method to the example, it yields that the optimal solution \( d(\tilde{y}) = \tilde{y} \). However, \( \mu = (\nabla_x g(x^*, y^*)^T d(\tilde{y}^*))^+ = \tilde{y} \) is not an LM. The reason is that the assumption \( \nabla q(d^*) = 0 \) of Problem (24) is violated in this example.

6 Examples

In this section, we apply our proposed penalty function method to compute informative LMs to some examples.

The first example shows that the informative LM may exist for some particular perturbation even if the shadow price is not existed. This example is extracted from Bank et al. (1982).

\[
\begin{align*}
\min_{x} & \quad x \\
\text{s.t.} & \quad y_1 x = 0 \\
& \quad x + 1 - y_2 \geq 0,
\end{align*}
\]

where \( x \) is the decision variable, \( y = (y_1, y_2)^T \) is the perturbation vector. The optimal solution at \( y^* = (0, 0)^T \) is \( x^* = -1 \), and the optimal value function

\[
f_{\text{inf}}(y) = \begin{cases}
-1 + y_2, & \text{if } y_1 = 0, y_2 \in \mathbb{R} \\
0, & \text{if } y_1 \neq 0, y_2 \geq 1, \\
-\infty, & \text{if } y_1 \neq 0, y_2 < 1,
\end{cases}
\]

which is not continuous at \( y^* = (0, 0)^T \), and it follows that the classical shadow prices are not existed. Now we apply the penalty function method to compute the informative Lagrange multiplier at \( x^* \). Suppose the variation of the perturbation vector is \( y_1 \neq 0 \) and \( y_2 > 0 \), then

\[
\begin{align*}
\delta_1 &= -\nabla_y h(x^*, y^*)^T (y - y^*) = y_1, \\
\delta_2 &= -\nabla_y g(x^*, y^*)^T (y - y^*) = y_2,
\end{align*}
\]

and

\[
\nabla_x h(x^*, y^*) = 0, \quad \nabla_x g(x^*, y^*) = -1.
\]

Since

\[
\begin{pmatrix}
\delta_1 \\
\delta_2
\end{pmatrix} \notin \text{Range} \begin{pmatrix}
\nabla_x h(x^*, y^*) \\
\nabla_x g(x^*, y^*)
\end{pmatrix},
\]

it follows that the informative Lagrange multiplier at \( x^* \) is not existed when the perturbation is \( y_1 \neq 0 \) and \( y_2 \in \mathbb{R} \). However, assume the perturbation is \( y_1 = 0 \) and \( y_2 \in \mathbb{R} \), then the assumption \( \delta \in \text{Range} \begin{pmatrix}
\nabla_x h(x^*, y^*) \\
\nabla_x g(x^*, y^*)
\end{pmatrix} \) satisfies. Applying the penalty function method proposed in Section 5, the corresponding LM is \( (0, 1)^T \) which is informative.

The second example is extracted from Gauvin and Dubeau (1984):

\[
\begin{align*}
\min_{x} & \quad -x \\
\text{s.t.} & \quad g(x) - y \leq 0,
\end{align*}
\]
where $x$ is the decision variable and $y$ is the perturbation, and

$$g(x) = \begin{cases} 
-(x + \frac{1}{2})^2 + \frac{5}{4} & \text{if } x \leq 0, \\
e^{-x} & \text{if } x \geq 0.
\end{cases}$$

Since the optimal value function is not continuous at $y^* = 0$, it follows that the shadow price does not exist at $(x^*, y^*)^T = (0, 0)^T$. However, for any perturbation $y > 0$, we have $\delta_y = y > 0$. Applying the penalty function method, we obtain the optimal direction $d(y) = \frac{x}{\sqrt{y}}$ and the corresponding LM is $\nabla_xg(x^*)^T d(y) = y$, which is informative.

7 Conclusion

The major insights from our analysis are:

(1) The internal cause of the resource variations is neglected in the classical analysis of the informative LM. The perturbation factors need to be considered in the modelling stage.

(2) The MNLM may fail to be informative in full parametric programming models.

(3) We provide sufficient conditions for an MNLM to be informative.

(4) Our proposed penalty function method can be applied to derive an informative LM in fully parametric programming models.

Acknowledgements This research is supported by National Natural Science Foundation of China grant No. 72071130, Humanities and Social Sciences of the Ministry of Education grant No. 22YJC630125, and Shanghai Office of Philosophy and Social Science grant No. 2020G016-BGL377.

References


