Second order directional derivative of optimal solution function in parametric programming problem

Bhuvnesh Khatana and Geetanjali Panda

[a] *Indian Institute of Technology, Kharagpur, 721302, India.*

Abstract

In this paper, the second-order directional derivative of the optimal value function and the optimal solution function are obtained for a strongly stable parametric problem with non-unique Lagrange multipliers. Some properties of the Lagrange multipliers are proved. It is justified that the second-order directional derivative of the optimal solution function for the parametric problem can be obtained by solving a suitable convex quadratic programming problem corresponding to an appropriate set of multipliers.

Keywords: Bilevel programming problems, parametric programming problem, sensitivity analysis, variational inequalities

AMS subject classifications. 65K10, 90C26, 90C33, 90C53

1. Introduction

In this paper, we consider the following smooth parametric optimization problem,

$$(P_x): \qquad \min_{y} \; f(x,y) \text{ s.t. } \; g(x,y) \leq 0, \;\; h(x,y) = 0,$$

where $f : \mathbb{R}^n \times \mathbb{R}^m \mapsto \mathbb{R}$, $g : \mathbb{R}^n \times \mathbb{R}^m \mapsto \mathbb{R}^p$, and $h : \mathbb{R}^n \times \mathbb{R}^m \mapsto \mathbb{R}^q$ are continuously differentiable functions. The feasible set of P_x is

$$\Omega_x := \{y \in \mathbb{R}^m \;:\; g(x,y) \leq 0, h(x,y) = 0\}.$$

The optimal solution of P_x, denoted by $y(x)$, is known as the optimal solution function at x, and $y \in \Omega_x$ is the feasible point of P_x. Furthermore, the optimal value of P_x at x, is denoted by $\phi(x)$,

$$\phi(x) := \min_{y} \{f(x,y) \text{ s.t. } g(x,y) \leq 0, h(x,y) = 0\},$$

which is known as the optimal value function at x. The sensitivity analysis of the parametric problem is the study of the behavior of parametric problem with respect to change in the parameter x.

The sensitivity analysis of a parametric programming problem has been a topic of interest because of its application to optimization theory, including semi-infinite optimization problems, Min-Max problems, Bilevel optimization problems, etc. A detailed study of the sensitivity analysis of parametric optimization problems can be found in the monographs by Bonnans and Shapiro [2], Fiacco [5], and Luderer et al. [9].

Sensitivity analysis of the optimal solution often assumes the strongly stable property of the optimal solution of the parametric problem to ensure continuity of the optimal solution function. The additional assumption of the constant rank condition ensures the Lipschitz continuity of the optimal solution function.

Some different sets of assumptions are seen in Shapiro [11, 12], to ensure the Lipschitz continuity of the optimal solution function and optimal value function, where the parametric problem is not necessarily uniquely solved in the neighborhood of the reference point.

Some of the existing contributions on first-order properties of P_x include the contributions by Fiacco [5],Gauvin and Janin [6], Shapiro [12], Auslender and Cominetti [1], Dempe [3], Ralph and Dempe [10], and Stechlinski et al. [13], among many others. Initial work on the sensitivity analysis of the optimal solution function of P_x does not consider changes in the active index set of parametric problems under parameter perturbations, relying instead on relatively strong assumptions, such as the linear independence constraint qualification (LICQ) and the strict complementarity condition. In such cases, the optimal solution function is continuously differentiable, and the implicit function theorem is directly used on the KKT (Karush-Kuhn-Tucker) optimality system to determine the sensitivity of the optimal solution function (see Fiacco [5]). Some contributions in this area in the absence of linear independent constraint qualification include Gauvin and Janin [6], Shapiro [12], Dempe [3], and Ralph and Dempe [10]. Gauvin and Janin [6] have studied the directional differentiability, Lipschitz continuity, and Hölder continuity properties of the optimal solution function under the second-order sufficient condition and Robinson's constraint qualification. Shapiro [12] obtained the first order directional derivative for the optimal solution and second order directional derivative for value function through an optimization problem corresponding to a suitable subset of multipliers. Dempe [3] and Ralph and Dempe [10] investigated a suitable multiplier set for computing the first-order directional derivative of the optimal solution function by solving a quadratic optimization problem.

In this paper, we focus on the study of the second-order directional differentiability of the optimal solution function in cases where the Lagrange multipliers associated with the parametric problem are not necessarily unique. Furthermore, the first-order results from Dempe [3] and Ralph and Dempe [10] are extended by identifying a subset of Lagrange multipliers that produce the second-order directional derivative of the optimal solution function by solving a quadratic programming problem.

This work is structured into several sections. In Section 2, some basic notations are provided along with existing results on the first-order properties of the optimal solution function. Second-order properties of the optimal value function are studied in Section 3. Using the results of Section 3, some properties of Lagrange multipliers of P_x are studied in Section 4. The main results on the second-order properties of the optimal solution function are studied in Section 5 using the results of Sections 3 and 4.

2. Preliminaries

For $x \in \mathbb{R}^n$, let $L(x, y, u, v) := f(x, y) + u^T g(x, y) + v^T h(x, y)$ denote the Lagrange function for P_x at x, where $u \in \mathbb{R}^p$, $v \in \mathbb{R}^q$ are Lagrange multiplier vectors associated with the inequality constraint $g(x, y) \leq 0$ and the equality constraint $h(x, y) = 0$. KKT necessary optimality conditions for P_x at a feasible point $(x, y(x))$ are

$$\nabla_y L(x, y(x), u, v) = 0,\ u^T g(x, y(x)) = 0,\ u \geq 0$$

for some $(u, v) \in \mathbb{R}^p \times \mathbb{R}^q$.

Some basic notations are stated below, which are used in subsequent sections at several places.

Notations

- $\nabla F := \begin{pmatrix} \nabla F_1 \\ \nabla F_2 \\ \vdots \\ \nabla F_q \end{pmatrix} \in \mathbb{R}^{q \times p}$, where each $\nabla F_i \in \mathbb{R}^{1 \times p}$ is the gradient of a function $F_i : \mathbb{R}^p \to \mathbb{R}$.

- $\nabla^2 F := \begin{pmatrix} \nabla^2 F_1 \\ \nabla^2 F_2 \\ \vdots \\ \nabla^2 F_q \end{pmatrix} \in \mathbb{R}^{q \cdot p \times p}$, where $\nabla^2 F_i \in \mathbb{R}^{p \times p}$ is the Hessian of F_i.

  Given $d \in \mathbb{R}^p$, denote $d^T \nabla^2 F d := \begin{pmatrix} d^T \nabla^2 F_1 d \\ d^T \nabla^2 F_2 d \\ \vdots \\ d^T \nabla^2 F_q d \end{pmatrix}$.

- $z := \begin{pmatrix} x \\ y \end{pmatrix}$ and the operator ∇^2_{zz} is defined as $\nabla^2_{zz} := \begin{pmatrix} \nabla^2_{xx} & \nabla^2_{xy} \\ \nabla^2_{yx} & \nabla^2_{yy} \end{pmatrix}$.
- N_p:= Neighborhood of a point p.
- $\Lambda_k := \{1, 2, \cdots, k\}$ is the index set of size k.
- Hessian of L with respect to (x, y, u, v) is denoted by $\widehat{\nabla}^2 L$ and defined as $\widehat{\nabla}^2 L := \begin{pmatrix} \nabla^2_{xx} L & \nabla^2_{xy} L & \nabla^2_{xu} L & \nabla^2_{xv} L \\ \nabla^2_{yx} L & \nabla^2_{yy} L & \nabla^2_{yu} L & \nabla^2_{yv} L \\ \nabla^2_{ux} L & \nabla^2_{uy} L & \nabla^2_{uu} L & \nabla^2_{uv} L \\ \nabla^2_{vx} L & \nabla^2_{vy} L & \nabla^2_{vu} L & \nabla^2_{vv} L \end{pmatrix}$,

  and Hessian of L with respect to (x, y) is denoted by $\nabla^2_{zz} L$.
- $I_x(y) := \{i \in \Lambda_p : g_i(x, y) = 0\}$ is the active set at $y \in \Omega_x$.
- $J(u) := \{i \in I_x(y) : u_i > 0\}$ is the set of indices with positive Lagrange multipliers in the vector u.
- The set of Lagrange multipliers of P_x is denoted by $U_x(y(x))$, which is,

$$U_x(y(x)) := \{(u, v) \in \mathbb{R}^p \times \mathbb{R}^q \ : \ \nabla_y L(x, y(x), u, v) = 0, \ u^T g(x, y(x)) = 0, \ u \geq 0\},$$

  and $EU_x(y(x))$ denotes the vertex set of $U_x(y(x))$.

The following assumptions at $(x, y(x))$ are used in several places to study the second-order properties of P_x in this paper.

Assumptions:

- (**A1**) The functions f, g and h are thrice continuously differentiable at $(x, y(x))$.
- (**A2**) Mangasarian-Fromovitz constraint qualification (MFCQ) holds at $(x, y(x))$. That is, the family $\left(\nabla_y h_i(x, y(x))\right)_{i \in \Lambda_q}$ is linearly independent, and there exists a nonzero vector $d_y \in \mathbb{R}^m$ such that

$$\nabla_y g_i(x, y(x)) d_y < 0, \ i \in I_x(y(x)), \quad \nabla_y h_i(x, y(x)) d_y = 0, \ i \in \Lambda_q.$$

- (**A3**) For all $(u, v) \in U_x(y(x))$, and for all nonzero vectors $d_y \in \mathbb{R}^m$ satisfying

$$\nabla_y g_i(x, y(x)) d_y = 0, \ i \in J(u), \ \nabla_y h_i(x, y(x)) d_y = 0, \ i \in \Lambda_q,$$

  we have $d_y^T \nabla^2_{yy} L(x, y(x), u, v) d_y > 0$, where $J(u) := \{i \in I_x(y(x)) : u_i > 0\}$ is the set of indices with positive Lagrange multipliers in the vector u.
- (**A4**) Constant rank constraint qualification ($CRCQ$) holds at $(x, y(x))$. That is, there exists a neighborhood $N_{(x,y)}$ of (x, y) such that for any subset W of $I_x(y)$, the family of gradient vectors $\left((\nabla_y g_i(x', y'))_{i \in W}, (\nabla_y h_i(x', y'))_{i \in \Lambda_q}\right)$ has the same rank for all $(x', y') \in N_{(x,y)}$.
- **(A5)** P_x is a convex programming problem for every x.

It is well known that under Assumption $A2$, $U_x(y(x))$ is a non-empty, convex, and bounded polyhedral set, which is not necessarily a singleton set. Further, Assumptions $A1-A3$ ensure that $y(x)$ is a strict local minimum point of P_x. Finally, Assumptions $A1-A4$ provide the Lipschitz continuity of the optimal solution function (see, Liu [8], Ralph and Dempe [10]). The assumption $A5$ ensures that $y(\cdot)$ is the global optimal solution function. In fact, under Assumptions $A1-A5$, $y(\cdot)$ is the uniquely determined global optimal solution function.

The optimal solution function $y(\cdot)$ of P_x is said to be first-order directionally differentiable at $x \in \mathbb{R}^n$ along the direction $d_x \in \mathbb{R}^n$, if

$$y'(x; d_x) := \lim_{t \downarrow 0} \frac{y(x + td_x) - y(x)}{t}$$

exists finitely. Similarly, the first-order directional derivative of the optimal value function $\phi(\cdot)$ along the direction d_x at x is

$$\phi'(x; d_x) := \lim_{t \downarrow 0} \frac{\phi(x + td_x) - \phi(x)}{t}$$

In the following theorem, we collect some of the existing results which were originally stated in the context of a parametric problem with inequality constraints($\min_y \ f(x, y)$ s.t. $g(x, y) \leq 0$,) and assemble them here in the context of the problem P_x.

Theorem 2.1. *(i). Under Assumptions $A1-A3$, $y(\cdot)$ is directionally differentiable at x. For a given vector d_x, there exists $(u, v) \in U_x(y(x))$ such that $y'(x; d_x)$ is the unique solution of the following quadratic programming problem :*

$$\begin{aligned} QP_{(u,v)}(x; d_x): \quad \min_{d_y} \quad & \frac{1}{2} d_y^T \nabla^2_{yy} L(x, y(x), u, v) d_y + d_y^T \nabla^2_{xy} L(x, y(x), u, v) d_x \\ & s.t.\ d_y \in K^1_u(x; d_x), \end{aligned}$$

where $K^1_u(x; d_x) := \left\{ d_y \in \mathbb{R}^m \ : \ (d_x, d_y) \in K^1_u(x) \right\}$, and $K^1_u(x)$ is the critical cone at $(x, y(x))$ with respect to $(u, v) \in U_x(y(x))$

$$K^1_u(x) := \left\{ (d_x, d_y) \in \mathbb{R}^n \times \mathbb{R}^m \ : \ \begin{array}{ll} \nabla_x g_i(x, y(x)) d_x + \nabla_y g_i(x, y(x)) d_y = 0, & i \in J(u), \\ \nabla_x g_i(x, y(x)) d_x + \nabla_y g_i(x, y(x)) d_y \leq 0, & i \in I_x(y(x)) \setminus J(u), \\ \nabla_x h_i(x, y(x)) d_x + \nabla_y h_i(x, y(x)) d_y = 0, & i \in \Lambda_q \end{array} \right\}.$$

(ii). Under Assumptions $A1-A4$, $y'(x; d_x)$ is the unique optimal solution of $QP_{(u,v)}(x; d_x)$ for any $(u, v) \in S^1(d_x)$, where

$$S^1(d_x) := \{(u, v) \in U_x(y(x)) \ : \ (u, v) \text{ solves } P^1_x\},$$

$$(P^1_x :) \max_{(u,v) \in U_x(y(x))} \sum_{i \in I_x(y(x))} u_i \nabla_x g_i(x, y(x)) d_x + \sum_{i \in \Lambda_q} v_i \nabla_x h_i(x, y) d_x.$$

(iii). Under Assumptions $A1-A3$, $y(\cdot) : N_x \to N_y$ is a continuous map, where N_x and N_y are open neighborhoods about x and y respectively.

(iv). Under Assumptions $A1-A3$, we have $V(d_x) \subseteq S^1(d_x) \subseteq U_x(y(x))$, where

$$V(d_x) := \left\{ (u, v) \in U_x(y(x)) \ : \ \begin{array}{c} \exists (u^k, v^k) \in U_{x^k}(y(x^k)), \ \{(u^k, v^k)\} \to (u, v) \\ \text{for } \{t_k\} \downarrow 0, \ t_k > 0, \text{ and } x^k := x + t_k d_x \ \forall k. \end{array} \right\}$$

(v). $y(\cdot)$ is second-order directionally differentiable under Assumptions $A1-A4$.

Proof. Proof of these results is straightforward from some existing results. The proof of (i) follows from Theorem 1, Dempe [3] and Shapiro [12] for P_x. The proof of (ii) follows from Theorem 10 of Ralph and Dempe [10] for P_x. The proof of (iii) follows from Theorem 7.2 of Kojima [7] for P_x. The proof of (iv) follows from Lemma 2.3 of Dempe [3] for P_x. The proof of (v) follows from Theorem 3.6 of Liu [8] for P_x. □

Lemma 2.2 of Dempe [3] shows that the set $K_u^1(x, d_x)$ represents the set of Lagrange multiplier vectors of P_x^1, which is non-empty if and only if $(u, v) \in S^1(d_x)$. Thus the quadratic program $QP_{(u,v)}(x, d_x)$ has a non-empty feasible set if and only if $(u, v) \in S^1(d_x)$, and $S^1(d_x)$ being the solution set of a linear programming problem P_x^1, is not necessarily a singleton set. Hence, Theorem 2.1(i) does not explicitly determine the directional derivative $y'(x; d_x)$ as it is unclear which $(u, v) \in U_x(y(x))$ should be used to derive $y'(x; d_x)$ from $QP_{(u,v)}(x, d_x)$. Nevertheless, with the additional Assumption $A4$, Theorem 2.1(ii) show that for any $(u, v) \in S^1(d_x)$, the first-order directional derivative $y'(x; d_x)$ is the solution of $QP_{(u,v)}(x; d_x)$.

Lemma 2.1. *Consider the parametric problem with equality constraints as*

$$\min_y f(x, y) \ s.t. \ h(x, y) = 0,$$

where $f : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^m$ *and* $h : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^q$. *Let* $\bar{y}(\cdot) \in \mathbb{R}^m$ *be the optimal solution function, and* $\bar{v}(\cdot) \in \mathbb{R}^q$ *be the Lagrange multiplier vector to the above problem. Suppose Assumptions* $A1, A3$ *are satisfied at* $(x, \bar{y}(x))$ *and the family* $(\nabla_y h_i(x, \bar{y}(x)))_{i \in \Lambda_q}$ *is linearly independent. Then* $(\bar{y}(.), \bar{v}(.)) \in C^2$ *in a neighborhood of* x.

Proof. Proof of this lemma follows directly from the proof of Corollary 3.2.5. of Shapiro [12] as a special case. □

The objective of the present contribution is to study some second-order properties of P_x, which is carried out in three stages: first, the second-order properties of the optimal value function $\phi(\cdot)$ are investigated, then properties of Lagrange multipliers are proved and next, the existence of second order directional derivative of optimal solution function is justified by solving a suitable quadratic programming problem.

Let the first-order directional derivative $y'(x; d_x)$ exist at x in the direction $d_x \in \mathbb{R}^n$. The second-order directional derivative of the optimal solution function $y(\cdot)$ exists at x in the direction d_x if the limit

$$y''(x; d_x) := \lim_{t \downarrow 0} \frac{y(x + td_x) - y(x) - ty'(x; d_x)}{t^2}$$

exists finitely.
Similarly, the second order directional derivative of the optimal value function $\phi(\cdot)$ along the direction d_x at x is

$$\phi''(x; d_x) := \lim_{t \downarrow 0} \frac{\phi(x + td_x) - \phi(x) - t\phi'(x; d_x)}{t^2}.$$

3. Second order properties of optimal value function

The following lemma is used to justify the second-order properties of the optimal value function $\phi(\cdot)$.

Lemma 3.1. *The set* $K_u^1(x; d_x)$ *remains the same irrespective of the selection of* $(u, v) \in S^1(d_x)$.

Proof. The proof of this result is similar to Lemma 8 of Ralph and Dempe [10] in the context of P_x. □

Since $K_u^1(x; d_x)$ does not depend on (u, v) if $(u, v) \in S^1(d_x)$, we denote $K_u^1(x; d_x) = K^1(x; d_x)$ if $(u, v) \in S^1(d_x)$ in the following derivations.

Theorem 3.1. *Suppose Assumptions $A1 - A5$ hold at $(x, y(x))$. Then $\phi'(x; d_x)$ and $\phi''(x; d_x)$ exist and*

$$\phi'(x; d_x) = \max_{(u,v)\in U_x(y(x))} \nabla_x L(x, y(x), u, v) d_x,$$

$$\phi''(x; d_x) = \max_{(u,v)\in S^1(d_x)} 0.5 \begin{pmatrix} d_x^T & y'(x; d_x)^T \end{pmatrix} \nabla^2_{zz} L(x, y(x), u, v) \begin{pmatrix} d_x \\ y'(x; d_x) \end{pmatrix}.$$

Proof. Suppose Assumptions $A1 - A3$ and $A5$ hold. From Theorem 4.1 and Relation 4.7 in Shapiro [12], it can be concluded that first order directional derivative of $\phi(\cdot)$ along the path $x(t) = x + td_x$ exists and

$$\phi'(x; d_x) = \max_{(u,v)\in EU_x(y(x))} \nabla_x L(x, y(x), u, v) d_x.$$

$EU_x(y(x))$ is the vertex set of $U_x(y(x))$. As the above problem is a linear programming problem, the solution is attained at a vertex. Hence

$$\phi'(x; d_x) = \max_{(u,v)\in EU_x(y(x))} \nabla_x L(x, y(x), u, v) d_x = \max_{(u,v)\in U_x(y(x))} \nabla_x L(x, y(x), u, v) d_x.$$

This proves first part.
From Theorem 4.2 of Shapiro [12], the second order directional derivative of $\phi(\cdot)$ along $x(t) = x + td_x$ exists, and is computed as

$$\phi''(x; d_x) := \min_{d_y \in \bigcup\limits_{U_x(y(x))} K_u^1(x; d_x)} \max_{(u,v)\in ES^1(d_x)} 0.5 \begin{pmatrix} d_x^T & d_y^T \end{pmatrix} \nabla^2_{zz} L(x, y(x), u, v) \begin{pmatrix} d_x \\ d_y \end{pmatrix}, \quad (1)$$

where $ES^1(d_x)$ is the vertex set of $S^1(d_x)$. From Lemma 2.2 of Dempe [3], $K_u^1(x; d_x)$ is non-empty if and only if $(u, v) \in S^1(d_x)$. From Lemma 3.1, $K_u^1(x; d_x)$ is independent of (u, v) if $(u, v) \in S^1(d_x)$. Hence,

$$\bigcup_{(u,v)\in U_x(y(x))} K_u^1(x; d_x) = \bigcup_{(u,v)\in S^1(d_x)} K_u^1(x; d_x) = K^1(x; d_x)$$

Since the inner maximization problem in Expression (1) is a linear programming problem in the variables (u, v) for fixed d_y, and the solution is attained at the vertex set, hence from (1),

$$\phi''(x; d_x) = \min_{d_y\in K^1(x;d_x)} \max_{(u,v)\in S^1(d_x)} 0.5 \begin{pmatrix} d_x^T & d_y^T \end{pmatrix} \nabla^2_{zz} L(x, y(x), u, v) \begin{pmatrix} d_x \\ d_y \end{pmatrix}$$

Since $\begin{pmatrix} d_x^T & d_y^T \end{pmatrix} \nabla^2_{zz} L(x, y(x), u, v) \begin{pmatrix} d_x \\ d_y \end{pmatrix}$ is a convex function in d_y for fixed (u, v) over the polyhedral set $K^1(x; d_x)$, and concave function in (u, v) for fixed d_y over the compact polyhedral set $S^1(d_x)$ therefore using min-max Theorem 3 of Fan [4], we obtain

$$\min_{d_y\in K^1(x;d_x)} \max_{(u,v)\in S^1(d_x)} \begin{pmatrix} d_x^T & d_y^T \end{pmatrix} \nabla^2_{zz} L(x, y(x), u, v) \begin{pmatrix} d_x \\ d_y \end{pmatrix}$$
$$= \max_{(u,v)\in S^1(d_x)} \min_{d_y\in K^1(x;d_x)} \begin{pmatrix} d_x^T & d_y^T \end{pmatrix} \nabla^2_{zz} L(x, y(x), u, v) \begin{pmatrix} d_x \\ d_y \end{pmatrix}.$$

Hence,

$$\phi''(x; d_x) = \max_{(u,v)\in S^1(d_x)} \min_{d_y\in K^1(x;d_x)} \begin{pmatrix} d_x^T & d_y^T \end{pmatrix} \nabla^2_{zz} L(x, y(x), u, v) \begin{pmatrix} d_x \\ d_y, \end{pmatrix}$$
$$= \max_{(u,v)\in S^1(d_x)} \begin{pmatrix} d_x^T & y'(x; d_x)^T \end{pmatrix} \nabla^2_{zz} L(x, y(x), u, v) \begin{pmatrix} d_x \\ y'(x; d_x) \end{pmatrix},$$

as from Theorem 2.1 (ii), $y'(x; d_x)$ is the optimal solution of $QP_{(u,v)}(x; d_x)$ and $K^1(x; d_x)$ is the feasible set of $QP_{(u,v)}(x; d_x)$ for each $(u, v) \in S^1(d_x)$.

Hence, the result follows. □

4. Properties of Lagrange multipliers

The set of Lagrange multipliers (u,v) at $(x,y(x))$ of P_x is

$$U_x(y(x)) = \left\{(u,v) \in \mathbb{R}^p \times \mathbb{R}^q \ : \nabla_y L(x,y(x),y,v) = 0,\ u^T g(x,y) = 0,\ u \geq 0\right\}.$$

Under Assumption $A2$, $U_x(y(x))$ is a bounded polyhedral set. Hence the set of vertices of $U_x(y(x))$, which is,

$$EU_x(y(x)) = \{(u,v) \ : \ (u,v) \text{ is the vertex of } U_x(y(x))\}.$$

is a finite set. Let the cardinality of $EU_x(y(x))$ be $\Lambda(x)$.

Let x^k be the points in the neighborhood of x in the direction d_x such that $x^k = x + t_k d_x$, where $\{t_k\}$ is a sequence of positive real numbers converging to 0. P_{x^k} is the parametric problem at x^k, $U_{x^k}(y(x^k))$ is the set of Lagrange multipliers of P_{x^k} at $(x^k, y(x^k))$, and $EU_{x^k}(y(x^k))$ is the vertex set of $U_{x^k}(y(x^k))$.

Consider the set of points $(u,v) \in U_x(y(x))$, for which there is sequence $\{(u^k,v^k)\}$, where $(u^k,v^k) \in U_{x^k}(y(x^k))$ such that $\{(u^k,v^k)\} \to (u,v)$, as

$$V(d_x) := \left\{ (u,v) \in U_x(y(x)) \ : \ \begin{array}{c} \exists (u^k,v^k) \in U_{x^k}(y(x^k)),\ \{(u^k,v^k)\} \to (u,v) \\ \text{for } \{t_k\} \downarrow 0,\ t_k > 0, \text{ and } x^k := x + t_k d_x\ \forall k \end{array} \right\}.$$

For given $d_z = \begin{pmatrix} d_x \\ y'(x;d_x) \end{pmatrix}$ at x, consider a linear programming problem as

$$(P_x^2 :) \quad \max_{(u,v)\in S^1(d_x)} \sum_{i\in I_x(y(x))} 0.5 u_i d_z^T \nabla^2_{zz} g_i(x,y(x)) d_z + \sum_{i\in\Lambda_q} 0.5 v_i d_z^T \nabla^2_{zz} h_i(x,y(x)) d_z.$$

Denote $S^2(d_x)$ as the solution set of this problem. That is,

$$S^2(d_x) := \{(u,v) \in S^1(d_x) \ : \ (u,v) \text{ solves } P_x^2\}.$$

Clearly $S^2(d_x) \subseteq S^1(d_x)$ and $S^2(d_x)$ is a non-empty bounded convex set because $S^1(d_x)$ is a non-empty bounded convex set.

Lemma 4.1. *Suppose Assumptions A1 and A2 hold at $(x,y(x))$. For a given vector $d_z = \begin{pmatrix} d_x \\ y'(x;d_x) \end{pmatrix}$, the following system in d_y is consistent if and only if $(u,v) \in S^2(d_x)$.*

$$\left.\begin{array}{l} 0.5 d_z^T \nabla^2_{zz} g_i(x,y(x)) d_z + \nabla_y g_i(x,y(x)) d_y = 0,\ i \in J(u), \\ 0.5 d_z^T \nabla^2_{zz} g_i(x,y(x)) d_z + \nabla_y g_i(x,y(x)) d_y \leq 0,\ i \in I_x(y(x);d_x) \setminus J(u), \\ 0.5 d_z^T \nabla^2_{zz} h_i(x,y(x)) d_z + \nabla_y h_i(x,y(x)) d_y = 0, i \in \Lambda_q. \end{array}\right\} \quad (2)$$

Additionally, the solution set of this system remains invariant irrespective of the selection of $(u,v) \in S^2(d_x)$.

Proof. Recall the optimization problem P_x^1. The feasible set $U_x(y(x))$ of P_x^1 is a non-empty, bounded, polyhedral set under Assumption $A2$. Using the definition of $U_x(y(x))$, the optimization problem P_x^1 can be expressed as,

$$\begin{aligned} (P_x^1) \quad \max_{(u,v)} & \sum_{i\in I_x(y(x))} u_i \nabla_x g_i(x,y(x)) + \sum_{i\in\Lambda_q} v_i \nabla_x h_i(x,y(x)) \\ \text{s.t. } & \nabla_y f(x,y(x)) + \sum_{i\in I_x(y(x))} u_i^T \nabla_y g_i(x,y(x)) + \sum_{i\in\Lambda_q} v_i \nabla_y h_i(x,y(x)) = 0, \\ & u_i \geq 0,\ i \in I_x(y(x)). \end{aligned}$$

As $S^1(d_x)$ is the solution set to P_x^1, so any $(u,v) \in S^1(d_x)$ satisfies the following KKT optimality conditions for P_x^1.

$$\begin{aligned}&\nabla_x g_i(x,y(x))d_x + \nabla_y g_i(x,y(x))d_y + s_i = 0,\ i \in I_x(y(x)),\\ &s_i \geq 0, \quad u_i s_i = 0,\ i \in I_x(y(x)),\end{aligned}$$

where $d_r \in \mathbb{R}^m$ and s_i are the Lagrange multipliers associated with the constraint $\nabla_y f(x,y(x)) + \sum_{i\in I_x(y(x))} u_i^T \nabla_y g(x,y(x)) + \sum_{i\in\Lambda_q} v_i \nabla_y h(x,y(x)) = 0$ and $u_i \geq 0,\ i \in I_x(y(x))$ respectively. Hence, any $(u,v) \in S^1(d_x)$ satisfies

$$\begin{aligned}&\nabla_y f(x,y(x)) + \sum_{i\in I_x(y(x))} u_i^T \nabla_y g_i(x,y(x)) + \sum_{i\in\Lambda_q} v_i \nabla_y h_i(x,y(x)) = 0,\\ &u_i\left(\nabla_x g_i(x,y(x))d_x + \nabla_y g_i(x,y(x))d_r\right) = 0,\ i \in I_x(y(x)),\\ &u_i \geq 0, \quad \nabla_x g_i(x,y(x))d_x + \nabla_y g_i(x,y(x))d_r \leq 0, \quad i \in I_x(y(x)).\end{aligned}$$

As P_x^1 is a linear programming problem, $(u,v) \in S^1(d_x)$ satisfies the above optimality conditions for any fixed d_r. Observe that $y'(x;d_x) \in K_u^1(x;d_x)$ and $K_u^1(x;d_x)$ is the set of Lagrange multiplier for P_x^1 from Lemma 2.2 of Dempe [3]. Therefore $d_r = y'(x;d_x)$ satisfies the above system. Hence $S^1(d_x)$ can be explicitly expressed as

$$S^1(d_x) = \left\{ (u,v) \in \mathbb{R}^p \times \mathbb{R}^q \ : \ \begin{aligned}&\nabla_y f(x,y(x)) + \sum_{i\in I_x(y(x))} u_i^T \nabla_y g(x,y(x)) + \sum_{i\in\Lambda_q} v_i \nabla_y h(x,y(x)) = 0,\\ &u_i = 0,\ \text{if}\ \ \nabla_x g_i(x,y(x))d_x + \nabla_y g_i(x,y(x))y'(x;d_x) < 0,\\ &u_i \geq 0,\ \text{if}\ \nabla_x g_i(x,y(x))d_x + \nabla_y g_i(x,y(x))y'(x;d_x) = 0.\end{aligned} \right\}$$

Recall the optimization problem (P_x^2), which can be expressed as follows using the set $S^1(d_x)$.

$$\begin{aligned}(P_x^2 :) \quad &\max_{(u,v)} \sum_{i\in I_x(y(x))} 0.5u_i d_z^T \nabla_{zz}^2 g_i(x,y(x))d_z + \sum_{i\in\Lambda_q} 0.5 v_i d_z^T \nabla_{zz}^2 h_i(x,y(x))d_z\\ &\text{s.t.}\ \nabla_y f(x,y(x)) + \sum_{i\in I_x(y(x))} u_i^T \nabla_y g_i(x,y(x)) + \sum_{i\in\Lambda_q} v_i \nabla_y h_i(x,y(x)) = 0, &(3a)\\ &\quad u_i = 0,\ \text{if}\ \ \nabla_x g_i(x,y(x))d_x + \nabla_y g_i(x,y(x))y'(x;d_x) < 0, &(3b)\\ &\quad u_i \geq 0,\ \text{if}\ \nabla_x g_i(x,y(x))d_x + \nabla_y g_i(x,y(x))y'(x;d_x) = 0. &(3c)\end{aligned}$$

The KKT optimality conditions of P_x^2 are

$$\begin{aligned}&0.5d_z^T \nabla_{zz}^2 g_i(x,y(x))d_z + \nabla_y g_i(x,y(x))d_y + w_i = 0,\ i \in I_x(y(x)), &(4a)\\ &0.5d_z^T \nabla_{zz}^2 h_i(x,y(x))d_z + \nabla_y h_i(x,y(x))d_y = 0,\ i \in \Lambda_q, &(4b)\\ &w_i \in \mathbb{R}, \ \ i \in \{i \in I_x(y(x)) \ : u_i = 0,\ \nabla_x g_i(x,y(x))d_x + \nabla_y g_i(x,y(x))y'(x;d_x) < 0\}, &(4c)\\ &w_i = 0,\ i \in \{i \in I_x(y(x);d_x) \ : u_i > 0\}, &(4d)\\ &w_i \geq 0, \ \ i \in \{i \in I_x(y(x);d_x) \ : u_i = 0\}, &(4e)\end{aligned}$$

where d_y is Lagrange multiplier associated with the constraint (3a), w_i is the Lagrange multiplier corresponding to the constraints (3b)-(3c), and (u,v) is the optimal solution of P_x^2. Using the value of w_i from (4c)-(4e) in (4a), we obtain

$$\left.\begin{aligned}&0.5d_z^T \nabla_{zz}^2 g_i(x,y(x))d_z + \nabla_y g_i(x,y(x))d_y \in \mathbb{R},\\ &\qquad i \in \{i \in I_x(y(x)) \ : u_i = 0, \nabla_x g_i(x,y(x))d_x + \nabla_y g_i(x,y(x))y'(x;d_x) < 0\},\end{aligned}\right\} \quad (5a)$$

$$0.5d_z^T \nabla_{zz}^2 g_i(x,y(x))d_z + \nabla_y g_i(x,y(x))d_y = 0,\ \{i \in I_x(y(x);d_x) \ : u_i > 0\} = J(u), \quad (5b)$$

$$0.5d_z^T \nabla_{zz}^2 g_i(x,y(x))d_z + \nabla_y g_i(x,y(x))d_y \leq 0,\ \{i \in I_x(y(x);d_x) \ : u_i = 0\} = I_x(y(x);d_x) \setminus J(u), \quad (5c)$$

$$0.5d_z^T \nabla_{zz}^2 h_i(x,y(x))d_z + \nabla_y h_i(x,y(x))d_y = 0,\ i \in \Lambda_q. \quad (5d)$$

Then, it can be easily verified that (5a)-(5d) and System (2) are the same. As $S^1(d_x)$ is a compact non-empty set, the feasible set of the linear programming problem P_x^2 is non-empty. Hence the set of optimal solutions of P_x^2 is non-empty.

To prove the if and only if part, consider $(u,v) \in S^2(d_x)$. Since $S^2(d_x)$ is the set of optimal solutions to P_x^2, and (4a)–(4e) are the KKT conditions for P_x^2, it follows that (4a)–(4e) are consistent. Therefore, (5a)–(5d) are also consistent and hence, System (2) is consistent for given d_z.
Conversely, suppose System (2) is consistent. Then (5a)–(5d) are consistent for some $(u,v) \in U_x(y(x))$. Let d_y^* be the solution of (5a)-(5d) for given d_z. Consider $w_i^* = -\big(0.5\, d_z^T \nabla^2_{zz} g_i(x,y(x))\, d_z + \nabla_y g_i(x,y(x))\, d_y^*\big)$. Then, (d_y^*, w^*) satisfies (4a)–(4e) for given d_z. Since (4a)-(4e) are the KKT conditions for P_x^2, it follows that $(u,v) \in S^2(d_x)$.

This proves the first part of the lemma. Since P_x^2 is a linear optimization problem, the set of multipliers in System (2) of P_x^2, is not dependent on the optimal solutions $(u,v) \in S^2(d_x)$. Hence, the second part of this lemma holds true. □

Theorem 4.1. *Suppose Assumptions $A1 - A5$ hold at $(x, y(x))$. Then*

$$V(d_x) \subseteq S^2(d_x) \subseteq S^1(d_x) \subseteq U_x(y(x)).$$

Proof. From Theorem 2.1(iv), it follows that $V(d_x) \subseteq S^1(d_x) \subseteq U_x(y(x))$. Clearly, $S^2(d_x) \subseteq S^1(d_x)$ as $S^2(d_x)$ is the set of optimal solution to P_x^2, for given $d_z = \begin{pmatrix} d_x \\ y'(x;d_x) \end{pmatrix}$. Next to show that $V(d_x) \subseteq S^2(d_x)$. Assume, on the contrary, that there is $(u^0,v^0) \in V(d_x) \setminus S^2(d_x)$. Then $(u^0,v^0) \in S^1(d_x)$ as $V(d_x) \subseteq S^1(d_x)$. Since $(u^0,v^0) \notin S^2(d_x)$ and $S^2(d_x)$ is the set of optimal solutions of P_x^2, there exists real number $\delta > 0$ satisfying

$$\begin{aligned}\sum_{i \in I_x(y(x))} & 0.5 u_i^0 d_z^T \nabla^2_{zz} g_i(x,y(x)) d_z + \sum_{i \in \Lambda_q} 0.5 v_i^0 d_z^T \nabla^2_{zz} h_i(x,y(x)) d_z \\ & < \max_{(u,v) \in S^1(d_x)} \sum_{i \in I_x(y(x))} 0.5 u_i d_z^T \nabla^2_{zz} g_i(x,y(x)) d_z + \sum_{i \in \Lambda_q} 0.5 v_i d_z^T \nabla^2_{zz} h_i(x,y(x)) d_z - \delta.\end{aligned}$$

The term $0.5 d_z^T \nabla^2_{zz} f(x,y(x)) d_z$ is independent of the multipliers (u,v). Adding this term to both sides of the above expression and substituting the value of L, we obtain

$$0.5 d_z^T \nabla^2_{zz} L(x,y(x),u^0,v^0) d_z < \max_{(u,v) \in S^1(d_x)} 0.5 d_z^T \nabla^2_{zz} L(x,y(x),u,v) d_z - \delta.$$

Substituting the value of $\phi''(x;d_x)$ from Theorem 3.1 in the above inequality,

$$0.5 d_z^T \nabla^2_{zz} L(x,y(x),u^0,v^0) d_z + \delta < \phi''(x;d_x). \tag{6}$$

Since $(u^0,v^0) \in V(d_x)$, there exists some $(u^k,v^k) \in U_{x^k}(y(x^k))$ such that $\{(u^k,v^k)\} \to (u^0,v^0)$. From the continuity of $y(\cdot)$ (Theorem 2.1(iii)), we can conclude that $\{y(x^k)\} \to y(x)$.
Further, from Theorem 2.1(v), since $y(\cdot)$ is second-order directionally differentiable under Assumptions $A1 - A4$, we have $\lim_{k\to\infty} \frac{y(x^k)-y(x)}{t_k} = y'(x;d_x)$ and $\lim_{k\to\infty} \frac{y(x^k)-y(x)-t_k y'(x;d_x)}{t_k^2} = y''(x;d_x)$.
Selecting $t_k = ||(u^k,v^k) - (u^0,v^0)||$, $\lim_{k\to\infty} \frac{(u^k,v^k)-(u^0,v^0)}{t_k} = (d_u,d_v)$ possibly over a subsequence.
Hence, $y(x^k) = y(x) + t_k y'(x;d_x) + t_k^2 y''(x;d_x) + o(t_k^2)$, $u^k = u^0 + t_k d_u + o(t_k)$, and $v^k = v^0 + t_k d_v + o(t_k)$ for some subsequence of $\{x_k\}$.

From Taylor's expansion of L about $(x, y(x), u^0, v^0)$,

$$
\begin{aligned}
L(x^k,&y(x^k), u^k, v^k) \\
&= L(x, y(x), u^0, v^0) + t_k\Big(\nabla_x L(x, y(x), u^0, v^0)d_x + \nabla_y L(x, y(x), u^0, v^0)y'(x; d_x) \\
&\qquad\qquad + \nabla_u L(x, y(x), u^0, v^0)d_u + \nabla_v L(x, y(x), u^0, v^0)d_v\Big) \\
&+ t_k^2\left(0.5\begin{pmatrix} d_z^T & d_u^T & d_v^T\end{pmatrix}\widehat{\nabla}^2 L(x, y(x), u^0, v^0)\begin{pmatrix} d_z \\ d_u \\ d_v\end{pmatrix} + \nabla_y L(x, y(x), u^0, v^0)y''(x; d_x)\right. \\
&\left.+ \nabla_u L(x, y(x), u^0, v^0)\left(\lim_{k\to\infty}\frac{u^k - u^0 - t_k d_u}{t_k^2}\right) + \nabla_v L(x, y(x), u^0, v^0)\left(\lim_{k\to\infty}\frac{v^k - v^0 - t_k d_v}{t_k^2}\right)\right) + o(t_k^2).
\end{aligned} \tag{7}
$$

Using Taylor's expansion of $\nabla_u L$ about $(x, y(x), u^0, v^0)$ up to first order,

$$
\nabla_u L(x^k, y(x^k), u^k, v^k) = \nabla_u L(x, y(x), u^0, v^0) + t_k\Big(d_x^T\nabla_{ux}^2 L(x, y(x), u^0, v^0) + y'(x; d_x)^T\nabla_{uy}^2 L(x, y(x), u^0, v^0)\Big) + o(t_k)
$$

Operating $t_k d_u$ on both sides in the above expression,

$$
\begin{aligned}
t_k\nabla_u L(x^k, y(x^k), u^k, v^k)d_u =& t_k\nabla_u L(x, y(x), u^0, v^0)d_u + \\
& t_k^2\Big(d_x^T\nabla_{ux}^2 L(x, y(x), u^0, v^0)d_u + y'(x; d_x)^T\nabla_{uy}^2 L(x, y(x), u^0, v^0)d_u\Big) + o(t_k^2).
\end{aligned}
$$

Rearranging the terms in the above expression yields

$$
\begin{aligned}
t_k\nabla_u L(x, y(x), u^0, v^0)d_u =& t_k\nabla_u L(x^k, y(x^k), u^0, v^0)d_u \\
& - t_k^2\Big(d_x^T\nabla_{ux}^2 L(x, y(x), u^0, v^0)d_u + y'(x; d_x)^T\nabla_{uy}^2 L(x, y(x), u^0, v^0)d_u\Big) + o(t_k^2).
\end{aligned} \tag{8}
$$

In a similar manner, from Taylor expansion on $\nabla_v L$ about $(x, y(x), u^0, v^0)$ up to first order, and then operating $t_k d_v$ on both sides and rearranging the terms, we obtain

$$
\begin{aligned}
t_k\nabla_v L(x, y(x), u^0, v^0)d_v =& t_k\nabla_v L(x^k, y(x^k), u^0, v^0)d_v \\
& - t_k^2\Big(d_x^T\nabla_{vx}^2 L(x, y(x), u^0, v^0)d_v + y'(x; d_x)^T\nabla_{vy}^2 L(x, y(x), u^0, v^0)d_v\Big) + o(t_k^2).
\end{aligned} \tag{9}
$$

Next, using the expression for $\widehat{\nabla}^2 L$(see the notations)

$$
\begin{aligned}
&\begin{pmatrix} d_z^T & d_u^T & d_v^T\end{pmatrix}\widehat{\nabla}^2 L(x, y(x), u^0, v^0)\begin{pmatrix} d_z \\ d_u \\ d_v\end{pmatrix} \\
&\quad= d_z^T\nabla_{zz}^2 L(x, y(x), u^0, v^0)\, d_z + 2\, d_x^T\nabla_{ux}^2 L(x, y(x), u^0, v^0)\, d_u + 2\, y'(x; d_x)^T\nabla_{uy}^2 L(x, y(x), u^0, v^0)\, d_u \\
&\quad+ 2\, d_x^T\nabla_{vx}^2 L(x, y(x), u^0, v^0)\, d_v + 2\, y'(x; d_x)^T\nabla_{vy}^2 L(x, y(x), u^0, v^0)\, d_v \\
&\qquad+ d_u^T\nabla_{uu}^2 L(x, y(x), u^0, v^0)\, d_u + 2\, d_u^T\nabla_{vu}^2 L(x, y(x), u^0, v^0)\, d_v + d_v^T\nabla_{vv}^2 L(x, y(x), u^0, v^0)\, d_v \\
&= d_z^T\nabla_{zz}^2 L(x, y(x), u^0, v^0)\, d_z + 2\, d_x^T\nabla_{ux}^2 L(x, y(x), u^0, v^0)\, d_u + 2\, y'(x; d_x)^T\nabla_{uy}^2 L(x, y(x), u^0, v^0)\, d_u \\
&\qquad+ 2\, d_x^T\nabla_{vx}^2 L(x, y(x), u^0, v^0)\, d_v + 2\, y'(x; d_x)^T\nabla_{vy}^2 L(x, y(x), u^0, v^0)\, d_v \text{ (since last three terms of} \\
&\qquad\qquad\text{the above expression vanish)}
\end{aligned} \tag{10}
$$

Using (8), (9) and (10) in (7), and then simplifying the resulting Expression (7), we get

$$\begin{aligned}
&L(x^k, y(x^k), u^k, v^k) - L(x, y(x), u^0, v^0) \\
&= t_k\Big(\nabla_x L(x, y(x), u^0, v^0)d_x + \nabla_y L(x, y(x), u^0, v^0)y'(x; d_x) + \nabla_u L(x^k, y(x^k), u^0, v^0)d_u + \nabla_v L(x^k, y(x^k), u^0, v^0)d_v\Big) \\
&\quad + t_k^2\Big(0.5d_z^T \nabla^2_{zz} L(x, y(x), u^0, v^0)d_z + \nabla_y L(x, y(x), u^0, v^0)y''(x; d_x) \\
&\quad + \nabla_u L(x, y(x), u^0, v^0)\Big(\lim_{k\to\infty} \frac{u^k - u^0 - t_k d_u}{t_k^2}\Big) + \nabla_v L(x, y(x), u^0, v^0)\Big(\lim_{k\to\infty} \frac{v^k - v^0 - t_k d_v}{t_k^2}\Big)\Big) + o(t_k^2).
\end{aligned} \tag{11}$$

Since $J(u) \subseteq J(u^k) \subseteq I_{x^k}(y(x^k)) \subseteq I_x(y(x))$ for large k, the following relation holds for large k.

$$\nabla_u L(x, y(x), u^0, v^0)(u^k - u^0) = \sum_{i\in I_x(y(x))} g_i(x, y(x))(u_i^k - u_i^0) = 0.$$

Dividing both side of the above expression by t_k and taking limit $k \to \infty$,

$$\nabla_u L(x, y(x), u^0, v^0)d_u = g(x, y(x))^T d_u = 0.$$

$$\begin{aligned}
\text{Next,}\ \ &\nabla_u L(x, y(x), u^0, v^0)\left(\lim_{k\to\infty} \frac{u^k - u^0 - t_k d_u}{t_k^2}\right) \\
&= \lim_{k\to\infty}\left(\frac{\nabla_u L(x, y(x), u^0, v^0)(u^k - u^0)}{t_k^2} - \frac{\nabla_u L(x, y(x), u^0, v^0)d_u}{t_k}\right) \\
&= 0
\end{aligned} \tag{12}$$

Further, observe that

$$\left.\begin{aligned}
&\nabla_y L(x, y(x), u^0, v^0) = 0, \\
&\nabla_u L(x^k, y(x^k), u^0, v^0)d_u = \lim_{k\to\infty} \sum_{i\in I_{x^k}(y(x^k))} g_i(x^k, y(x^k))\frac{u_i^k - u_i^0}{t_k} = 0, \\
&\nabla_v L(x^k, y(x^k), u^0, v^0)d_v = h(x^k, y(x^k))^T d_v = 0, \\
&\nabla_v L(x, y(x), u^0, v^0)\lim_{k\to\infty} \frac{v^k - v^0 - t_k d_v}{t_k^2} = h(x, y(x))^T \lim_{k\to\infty} \frac{v^k - v^0 - t_k d_v}{t_k^2} = 0.
\end{aligned}\right\} \tag{13}$$

Using (12)-(13) in (11),

$$\begin{aligned}
\phi(x^k) - \phi(x) &= f(x^k, y(x^k)) + {u^k}^T g(x^k, y(x^k)) - f(x, y(x)) - u^T g(x, y(x))\ (as\ {u^k}^T g(x^k, y(x^k)) = u^T g(x, y(x)) = 0) \\
&= L(x^k, y(x^k), u^k, v^k) - L(x, y(x), u^0, v^0) \\
&= t_k \nabla_x L(x, y(x), u^0, v^0)d_x + 0.5t_k^2 d_z \nabla^2_{zz} L(x, y(x), u^0, v^0)d_z + o(t_k^2).
\end{aligned}$$

From Theorem 3.1,
$\phi'(x; d_x) = \max_{(u,v)\in U_x(y(x))} \nabla_x L(x, y(x), u, v)d_x = \nabla_x L(x, y(x), u^0, v^0)d_x$ as $(u^0, v^0) \in S^1(d_x)$ is the solution of P_x^1 and

$$\begin{aligned}
\phi''(x; d_x) &= \lim_{k\to\infty} \frac{\phi(x^k) - \phi(x) - t_k\phi'(x; d_x)}{t_k^2} \\
&= 0.5d_z^T \nabla^2_{zz} L(x, y(x), u^0, v^0)d_z
\end{aligned}$$

This contradicts (6). Hence, the result follows. □

4.1. Convergence of the sequence $\{(u^k, v^k)\}$

Let (u^k, v^k) be Lagrange multipliers of P_{x^k}. We justify the convergence of the sequence $\{(u^k, v^k)\}$ to some (u, v) over a subsequence of $\{x_k\}$, where (u, v) is the Lagrange multipliers of P_x. That is, we need to justify that $V(d_x)$ is non-empty. In the next result, we will justify the existence of such a sequence $\{(u^k, v^k)\}$. From Theorem 4.1, $V(d_x) \subseteq S^2(d_x)$. If $(u, v) \notin V(d_x)$ and $(u, v) \in S^2(d_x)$ then we will justify later in Lemma 4.3 that there exits a sequence of perturbed multipliers $(\bar{u}^k, \bar{v}^k)$ converging to (u, v).

Lemma 4.2. *Let $x^k = x + t_k d_x$ for some sequence of positive real numbers $\{t_k\}$ converging to* 0*, and Assumptions $A1 - A4$ hold at $(x, y(x))$.*

1. *$V(d_x)$ is a non-empty set.*
2. *Suppose $(u, v) \in V(d_x)$. There exists $(u^k, v^k) \in U_{x^k}(y(x^k))$ for each k such that $\lim_{k\to\infty} \frac{(u^k,v^k)-(u,v)}{t_k} = (d_u, d_v)$ for some $(d_u, d_v) \in \mathbb{R}^p \times \mathbb{R}^q$, and $\lim_{k\to\infty} \frac{(u^k,v^k)-(u,v)-t_k(d_u,d_v)}{t_k^2}$ exists.*

Proof. **(1):**

Observe that $(x^k, y(x^k))$ lies in the neighborhood of (x, y) for large k. As Assumptions $A1 - A3$ hold, so from continuity of optimal solution function $y(\cdot)$, we have $\{(x^k, y(x^k))\} \to (x, y(x))$.

From Assumption $A2$ and $A4$, the family $\left(\nabla_y h_i(x^k, y(x^k))\right)_{i\in\Lambda_q}$ is linearly independent at $(x^k, y(x^k))$ for large k.

First, we justify the result for the set of extreme points of $U_{x^k}(y(x^k))$ and then extend it to $U_{x^k}(y(x^k))$. $EU_{x^k}(y(x^k))$ is the set of extreme points of $U_{x^k}(y(x^k))$. Let $\left(u^{*k}, v^{*k}\right) \in EU_{x^k}(y(x^k))$ at x^k and $\Lambda(x^k)$ be the cardinality of $EU_{x^k}(y(x^k))$. Any point $(u^k, v^k) \in U_{x^k}(y(x^k))$ can be expressed as the convex combination of the vertices of $U_{x^k}(y(x^k))$. Hence

$$(u^k, v^k) = \sum_{j=1}^{\Lambda(x^k)} \alpha_j^{*k} \left(u^{*k}, v^{*k}\right)^j, \text{where } \sum_{j=1}^{\Lambda(x^k)} \alpha_j^{*k} = 1,\ \alpha_j^{*k} \geq 0. \tag{14}$$

We claim that $\left\{\left(u^{*k}, v^{*k}\right)\right\} \to (u^*, v^*)$ for some $(u^*, v^*) \in EU_x(y(x))$.

Since $\left(u^{*k}, v^{*k}\right) \in EU_{x^k}(y(x^k))$, the family of gradients $\left(\left(\nabla_y g_i(x^k, y(x^k))\right)_{i\in J(u^{*k})}, \left(\nabla_y h_i(x^k, y(x^k))\right)_{i\in K(v^{*k})}\right)$ is linearly independent for large k, where $K(v^{*k}) = \{i \in \Lambda_q\ :\ v_i^{*k} \neq 0\} \subseteq \Lambda_q$. We may assume that the sets $J(u^{*k})$ and $K(v^{*k})$ remain the same by choosing a subsequence of $\{x^k\}$. Denote these sets by J^* and K^* respectively. Thus,

$$\nabla_y f(x^k, y(x^k)) + \sum_{i\in J^*} u_i^{*k} \nabla_y g_i(x^k, y(x^k)) + \sum_{i\in K^*} v_i^{*k} \nabla_y h_i(x^k, y(x^k)) = 0. \tag{15}$$

By the continuity of $\nabla_y f, \nabla_y g$, and $\nabla_y h$ and boundedness property of $U_x(y(x))$, letting $k \to \infty$ in (15), we obtain

$$\nabla_y f(x, y(x)) + \sum_{i\in J^*} u_i^* \nabla_y g_i(x, y(x)) + \sum_{i\in K^*} v_i^* \nabla_y h_i(x, y(x)) = 0,$$

for some $(u^*, v^*) \in U_x(y(x))$. From Assumption $A4$, the family of vectors $\left((\nabla_y g_i(x, y(x)))_{i\in J^*}, (\nabla_y h_i(x, y(x)))_{i\in K^*}\right)$ is linearly independent, so $(u^*, v^*) \in EU_x(y(x))$.

Since $\{\alpha_j^{*k}\}$ is a bounded sequence for each j, $\{\alpha_j^{*k}\} \to \alpha_j^*$ for some subsequence such that $\sum_{j=1}^{\Lambda(x)} \alpha_j^* = 1$. Hence, $\left\{(u^{*k}, v^{*k})\right\} \to (u^*, v^*)$ and $(u^*, v^*) \in EU_x(y(x))$ over a subsequence of $\{x^k\}$.

Next, taking $k \to \infty$ both sides in the expression (14) over an appropriate subsequence of $\{x_k\}$, we obtain

$$\lim_{k\to\infty} (u^k, v^k) = \sum_{j=1}^{\Lambda(x)} \alpha_j^* (u^*, v^*)^j \triangleq (u, v)\ (say), \tag{16}$$

As $(u^*, v^*)^j \in EU_x(y(x))$ and (u, v) is the convex combination of the points of $EU_x(y(x))$, so $(u, v) \in U_x(y(x))$. This proves the first part of the lemma.

(2)

Consider the following optimization problem $P^*(x)$ corresponding to the index set J^* in x

$$(P^*(x)): \quad \min_y f(x, y) \text{ s.t. } g_i(x, y) = 0,\ i \in J^*, \quad h_i(x, y) = 0,\ i \in K^*.$$

Let $y^*(x)$ be the optimal solution of $P^*(x)$, $u^*(x)$ and $v^*(x)$ be the Lagrange multipliers associated with constraints $g_i(x, y) = 0, i \in J^*$ and $h_i(x, y) = 0, i \in K^*$ at x, respectively.

Since Assumption $A3$ holds for $P^*(x)$, and the family $\left((\nabla_y g_i(x, y(x)))_{i \in J^*}, (\nabla_y h_i(x, y(x)))_{i \in K^*}\right)$ is linearly independent, Lemma 2.1 is applied to $P^*(x)$. Hence, there are unique vector functions $y^*(\cdot), u^*(\cdot)$, and $v^*(\cdot) \in C^2$, satisfying KKT optimality conditions for $P^*(x)$ in the neighborhood of x.
For x^k in the neighborhood of x,

$$\nabla_y f(x^k, y^*(x^k)) + \sum_{i \in J^*} (u^*(x^k))_i \nabla_y g_i(x^k, y^*(x^k)) + \sum_{i \in K^*} (v^*(x^k))_i \nabla_y h_i(x^k, y^*(x^k)) = 0.$$

From expression (15) and above expression, we have

$$u_i^{*k} = (u^*(x^k))_i,\ \forall i \in J^*,\ v_i^{*k} = (v^*(x^k))_i,\ \forall i \in K^* \text{ and } \ y(x^k) = y^*(x^k) \text{ for large } \ k. \tag{17}$$

As $u^*(\cdot), v^*(\cdot) \in C^2$, expanding these about x up-to second order yields

$$(u^*(x^k))_i = (u^*(x))_i + t_k d^*_{u_i} + t_k^2 d^{1^*}_{u_i} + o(t_k^2),\ i \in J^*, \tag{18}$$
$$(v^*(x^k))_i = (v^*(x))_i + t_k d^*_{v_i} + t_k^2 d^{1^*}_{v_i} + o(t_k^2),\ i \in K^* \tag{19}$$

for some $d_u^* \triangleq \begin{cases} d^*_{u_i} & i \in J^* \\ 0 & i \notin J^* \end{cases}$, $d_v^* \triangleq \begin{cases} d^*_{v_i} & i \in K^* \\ 0 & i \notin K^* \end{cases}$, $d_u^{1^*} \triangleq \begin{cases} d^{1^*}_{u_i} & i \in J^* \\ 0 & i \notin J^* \end{cases}$, $d_v^{1^*} \triangleq \begin{cases} d^{1^*}_{v_i} & i \in K^* \\ 0 & i \notin K^* \end{cases} \in \mathbb{R}^q$.
From the above discussion, one can conclude that $J(u^{*k}) = J^*$ for large k. Hence, $u_i^{*k} > 0$ for $i \in J^*$ and $u_i^{*k} = 0$ for $i \notin J^*$. As $\{u^{*k}\} \to u^*$ so $J(u^*) \subseteq J(u^{*k})$ for large k. Thus, $J(u^*) \subseteq J^*$ for large k. Then $u_i^* \geq 0$ for $i \in J^*$, and $u_i^* = 0$ for $i \notin J^*$.
From (17) and (18) ,

$$\lim_{k \to \infty} \frac{u^{k^*} - u^*}{t_k} = \begin{cases} \lim_{k \to \infty} \frac{\left(u^*(x^k)\right)_i - \left(u^*(x)\right)_i}{t_k} & i \in J^* \\ 0 & i \notin J^* \end{cases} = \begin{cases} d^*_{u_i} & i \in J^* \\ 0 & i \notin J^* \end{cases} = d_u^*,$$

where $\left(u^*(x^k)\right)_i$ and $\left(u^*(x)\right)_i$ denote the i^{th} component of vectors $u^*(x^k)$ and $u^*(x)$ respectively. Next,

$$\lim_{k \to \infty} \frac{u^{*k} - u^* - t_k d_u^*}{t_k^2} = \begin{cases} \lim_{k \to \infty} \frac{\left(u^*(x^k)\right)_i - \left(u^*(x)\right)_i - t_k d^*_{u_i}}{t_k^2} & i \in J^* \\ 0 & i \notin J^* \end{cases} = \begin{cases} d^{1^*}_{u_i} & i \in J^* \\ 0 & i \notin J^* \end{cases} = d_u^{1^*}.$$

Similarly, from (17) and (19), we obtain

$$\lim_{k \to \infty} \frac{v^{*k} - v^*}{t_k} = d_v^* \text{ and } \lim_{k \to \infty} \frac{v^{*k} - v^* - t_k d_v^*}{t_k^2} = d_v^{1^*}$$

Since $\{(u^{*k}, v^{*k})\} \to (u^*, v^*)$, where $(u^{*k}, v^{*k}) \in EU_{x^k}(y(x^k))$ and $(u^*, v^*) \in EU_x(y(x))$ therefore $\Lambda(x^k) \leq \Lambda(x)$ for large k. Consider the vector $(u^k, v^k) \in U_{x^k}(y(x^k))$, which is expressed in the convex combination of vertices of $U_{x^k}(y(x^k))$ as

$$(u^k, v^k) = \sum_{j=1}^{\Lambda(x)} \alpha_j^* \left(u^{*k}, v^{*k}\right)^j \text{ where } \sum_{j=1}^{\Lambda(x)} \alpha_j = 1,\ \alpha_j^* \geq 0, \tag{20}$$

where α_j^* is the limit point of the sequence $\{\alpha_j^k\}^*$ from the expression (16). Then, we obtain

$$\begin{aligned}
\lim_{k\to\infty} \frac{(u^k, v^k) - (u, v)}{t_k} &= \lim_{k\to\infty} \sum_{j=1}^{\Lambda(x)} \alpha_j^* \frac{\left(u^{*k}, v^{*k}\right)^j - (u^*, v^*)^j}{t_k} \text{ (from (16) and (20))} \\
&= \sum_{j=1}^{\Lambda_x} \alpha_j (d_u^*, d_v^*)^j,
\end{aligned}$$

and

$$\begin{aligned}
&\lim_{k\to\infty} \frac{(u^k, v^k) - (u, v) - t_k \sum_{j=1}^{\Lambda(x)} \alpha_j^* (d_u^*, d_v^*)^j}{t_k^2} \\
&= \lim_{k\to\infty} \sum_{j=1}^{\Lambda(x)} \alpha_j^* \frac{\left(u^{*k}, v^{*k}\right)^j - (u^*, v^*)^j - t_k (d_u^*, d_v^*)^j}{t_k^2} \text{ (from (16), and (20))} \\
&= \sum_{j=1}^{\Lambda(x)} \alpha_j^* (d_u^{1^*}, d_v^{1^*})^j.
\end{aligned}$$

This proves the lemma. □

Let $x^k = x + t_k d_x$, where $\{t_k\} \downarrow 0$ is sequence of positive real numbers, and $G(x^k; d_x) \in \mathbb{R}^p$ be defined as

$$G_i(x^k; d_x) := \begin{cases} g_i(x^k, y(x^k)) & \text{if } i \in \{i \in \Lambda_p \ : \ \exists (u, v) \in S^2(d_x), u_i > 0\} \\ 0 & \text{otherwise} \end{cases},$$

where $S^2(d_x)$ is computed at x. Consider the following problem $\bar{P}_{x^k}$, which is obtained from P_{x^k} replacing $g(x^k, y) \leq 0$ by $g(x^k, y) \leq G(x^k; d_x)$.

$$(\bar{P}_{x^k}) \ : \ \min_y f(x^k, y) \text{ s.t. } g(x^k, y) \leq G(x^k; d_x),\ h(x^k, y) = 0.$$

The necessary KKT optimality conditions for $\bar{P}_{x^k}$ are

$$\left.\begin{array}{l} \nabla_y f(x^k, y) + u^T \nabla_y g(x^k, y) + v^T \nabla_y h(x^k, y) = 0, \quad u^T \left(g(x^k, y) - G(x^k; d_x)\right) = 0, \\ u \geq 0, \quad g(x^k, y) - G(x^k; d_x) \leq 0,\ h(x^k, y) = 0. \end{array}\right\} \tag{21}$$

Let $\bar{U}_{x^k}(y(x^k); d_x)$ denote the set of Lagrange multipliers of $\bar{P}_{x^k}$,

$$\bar{U}_{x^k}(y(x^k); d_x) := \{(\bar{u}^k, \bar{v}^k) \ : \ (x^k, y(x^k), \bar{u}^k, \bar{v}^k) \text{ satisfy (21)}\}.$$

We say $(\bar{u}^k, \bar{v}^k) \in \bar{U}_{x^k}(y(x^k); d_x)$ as the perturbed multipliers of the perturbed KKT system (21).

Lemma 4.3. *Suppose $y(\cdot) : N_x \to N_y$ is directionally differentiable and Assumptions $A1, A3$ and $A4$ are satisfied at $(x, y(x))$. For each $(u, v) \in S^2(d_x)$ and the sequence $\{x^k\}$ such that $x^k = x + t_k d_x$, where $\{t_k\} \downarrow 0$ is the sequence of positive real numbers, there exists $(\bar{u}^k, \bar{v}^k) \in \bar{U}_{x^k}(y(x^k); d_x)$ such that $\{(\bar{u}^k, \bar{v}^k)\} \to (u, v)$.*

Proof. The proof of this lemma follows in the proof of Lemma 9 of Ralph and Dempe [10], after changing the role of $S(x;d)$ by $S^2(d_x)$, and $M^*(x^k;d)$ by $\bar{U}_{x^k}(y(x^k);d_x))$. □

From the above lemma, it is clear that $\bar{U}_{x^k}(y(x^k);d_x) \neq \emptyset$ for some x^k in the neighborhood of x, and (21) is satisfied at $(y(x^k), \bar{u}^k, \bar{v}^k)$. Hence

$$\left.\begin{aligned} &\nabla_y f(x^k, y(x^k)) + {\bar{u}^k}^T \nabla_y g(x^k, y(x^k)) + {\bar{v}^k}^T \nabla_y h(x^k, y(x^k)) = 0, \quad {\bar{u}^k}^T \left(g(x^k, y(x^k)) - G(x^k; d_x)\right) = 0,\\ &\bar{u}^k \geq 0, \quad g(x^k, y(x^k)) - G(x^k; d_x) \leq 0, \; h(x^k, y(x^k)) = 0. \end{aligned}\right\} \tag{22}$$

5. Second order properties of optimal solution function

The main difficulty in obtaining the second-order directional derivative $y''(x;d_x)$ is to determine the set $V(d_x)$, which is difficult to obtain explicitly without imposing restrictive assumptions on P_x. It is justified in Theorem (2.1) that, $y'(x;d_x)$ is the optimal solution of $QP_{(u,v)}(x;d_x)$ under Assumptions $A1-A3$ for any $(u,v) \in S^1(d_x)$. In this section, we extend this concept to compute the second-order directional derivative by solving a suitable quadratic optimization problem.

Let $(u,v) \in V(d_x)$. As $\{x^k\} \to x$, $\{y(x^k)\} \to y(x)$ from the continuity of $y(\cdot)$ (Theorem (2.1) (iii)). From Theorem 2.1(vi), $y(\cdot)$ is second-order directionally differentiable at x. Hence,

$$\lim_{k\to\infty} \frac{y(x^k) - y(x)}{t_k} = y'(x;d_x) \text{ and } \lim_{k\to\infty} \frac{y(x^k) - y(x) - t_k d_y}{t_k^2} = y''(x;d_x).$$

From Lemma 4.2, for $(u,v) \in V(d_x)$, there exists $(u^k, v^k) \in U_{x^k}(y(x^k))$ such that $\{(u^k, v^k)\}$ converges to (u,v) possibly over a subsequence of $\{x_k\}$, and

$$\lim_{k\to\infty} \frac{(u^k, v^k) - (u,v)}{t_k} = (d_u, d_v) \text{ (say) and } \lim_{k\to\infty} \frac{(u^k, v^k) - (u,v) - t_k(d_u, d_v)}{t_k^2} = (d_u^1, d_v^1) \text{ (say)},$$

for some $(d_u, d_v), (d_u^1, d_v^1) \in \mathbb{R}^p \times \mathbb{R}^q$. Hence, for some subsequence of $\{x^k\}$,

$$\begin{aligned} &y(x^k) = y(x) + t_k y'(x;d_x) + t_k^2 y''(x;d_x) + o(t_k^2),\\ &u^k = u + t_k d_u + t_k^2 d_u^1 + o(t_k^2),\\ &\text{and } v^k = v + t_k d_v + t_k^2 d_v^1 + o(t_k^2). \end{aligned}$$

Since $(u^k, v^k) \in U_{x^k}(y(x^k))$, the following KKT optimality conditions for P_{x^k} hold at $(x^k, y(x^k))$,

$$\nabla_y L(x^k, y(x^k), u^k, v^k) = 0, \tag{23a}$$

$$g_i(x^k, y(x^k)) = 0, \; i \in J(u^k), \quad h_i(x^k, y(x^k)) = 0, \; i \in \Lambda_q, \tag{23b}$$

$$g_i(x^k, y(x^k)) \leq 0, \quad i \notin J(u^k), \tag{23c}$$

$$u_i^k g_i(x^k, y(x^k)) = 0, \quad i \in \Lambda_p. \tag{23d}$$

Since $(u,v) \in U_x(y(x))$,

$$\nabla_y L(x, y(x), u, v) = 0, \; u_i \geq 0, \; i \in \Lambda_p \; g_i(x, y(x)) = 0, \; i \in I_x(y(x)), \; h_i(x, y(x)) = 0, \; i \in \Lambda_q. \tag{23e}$$

Using Taylor's expansion up-to second order on $\nabla_y L$ about $(x, y(x), u, v)$, and on $g_i,\ i \in \Lambda_p,\ h_i,\ i \in \Lambda_q$ about $(x, y(x))$, we obtain

$$\begin{aligned}
&\nabla_y L(x^k, y(x^k), u^k, v^k)^T \\
&\qquad = \nabla_y L(x, y(x), u, v)^T + t_k \Big(\nabla^2_{xy} L(x, y(x), u, v) d_x + \nabla^2_{yy} L(x, y(x), u, v) y'(x; d_x) \\
&\qquad\qquad + \nabla^2_{uy} L(x, y(x), u, v) d_u + \nabla^2_{vy} L(x, y(x), u, v) d_v \Big) \\
&\qquad + t_k^2 \left(0.5 \begin{pmatrix} d_z^T & d_u^T & d_v^T \end{pmatrix} \widehat{\nabla}^2 \big(\nabla_y L(x, y(x), u, v)^T\big) \begin{pmatrix} d_z \\ d_u \\ d_v \end{pmatrix} + \nabla^2_{yy} L(x, y(x), u, v) y''(x; d_x) \right. \\
&\qquad\qquad \left. + \nabla^2_{uy} L(x, y(x), u, v) d_u^1 + \nabla^2_{vy} L(x, y(x), u, v) d_v^1 \right) + e_m o(t_k^2),
\end{aligned} \tag{24a}$$

where e_m as the vector of dimension m whose each component is 1 and

$$\left.\begin{aligned}
g_i(x^k, y(x^k)) &= g_i(x, y(x)) + t_k \big(\nabla_x g_i(x, y(x)) d_x + \nabla_y g_i(x, y(x)) y'(x; d_x) \big) \\
&\quad + t_k^2 \Big(0.5 d_{zz}^T \nabla^2_{zz} g_i(x, y(x)) d_z + \nabla_y g_i(x, y(x)) y''(x; d_x) \Big) + o(t_k^2),\ i \in \Lambda_p, \\
h_i(x^k, y(x^k)) &= h_i(x, y(x)) + t_k \big(\nabla_x h_i(x, y(x)) d_x + \nabla_y h_i(x, y) y'(x; d_x) \big) \\
&\quad + t_k^2 \Big(0.5 d_z^T \nabla^2_{zz} h_i(x, y) d_z + \nabla_y h_i(x, y(x)) y''(x; d_x) \Big) + o(t_k^2),\ i \in \Lambda_q
\end{aligned}\right\} \tag{24b}$$

respectively.

From (23a) and (23e), $\lim\limits_{k \to \infty} \frac{\nabla_y L(x^k, y(x^k), u^k, v^k)^T - \nabla_y L(x, y(x), u, v)^T}{t_k} = 0$. Hence, from (24a),

$$\nabla^2_{xy} L(x, y(x), u, v) d_x + \nabla^2_{yy} L(x, y(x), u, v) y'(x; d_x) + \nabla^2_{uy} L(x, y(x), u, v) d_u + \nabla^2_{vy} L(x, y(x), u, v) d_v = 0. \tag{25a}$$

Since $\{u^k\}$ converge to u therefore $J(u) \subseteq J(u^k)$ for large k, $g_i(x^k, y(x^k)) = 0,\ i \in J(u)$ in (23b). Next, using (23e), $\lim\limits_{k \to \infty} \frac{g_i(x^k, y(x^k)) - g_i(x, y(x))}{t_k} = 0$ for $i \in J(u)$. Hence, from (24b),

$$\nabla_x g_i(x, y(x)) d_x + \nabla_y g_i(x, y(x)) y'(x; d_x) = 0,\ i \in J(u). \tag{25b}$$

Furthermore, using (23c) and (23e), we have $\lim\limits_{k \to \infty} \frac{g_i(x^k, y(x^k)) - g_i(x, y(x))}{t_k} \leq 0$ for $i \in I_x(y(x)) \setminus J(u)$. Hence, from (24b),

$$\nabla_x g_i(x, y(x)) d_x + \nabla_y g_i(x, y(x)) y'(x; d_x) \leq 0,\ i \in I_x(y(x)) \setminus J(u). \tag{25c}$$

Similarly, using (23b) and (23e), $\lim\limits_{k \to \infty} \frac{h_i(x^k, y(x^k)) - h_i(x, y(x))}{t_k} = 0$ for $i \in \Lambda_q$. Hence, from (24b),

$$\nabla_x h_i(x, y(x)) d_x + \nabla_y h_i(x, y(x)) y'(x; d_x) = 0,\ i \in \Lambda_q. \tag{25d}$$

Since $u_i = 0$ for $i \in I_x(y(x)) \setminus J(u)$ and $u_i^k \geq 0$ for $i \in I_x(y(x))$, we have $d_{u_i} = \lim\limits_{k \to \infty} \frac{u_i^k - u_i}{t_k} \geq 0$ for $i \in I_x(y(x)) \setminus J(u)$. Since $I_{x^k}(y(x^k)) \subseteq I_x(y(x))$ for large k therefore $u_i = 0$ and $u_i^k = 0$ for $i \notin I_x(y(x))$ and large k. Hence $d_{u_i} = 0$ for $i \notin I_x(y(x))$.

If $d_{u_i} = \lim\limits_{k\to\infty} \frac{u_i^k - u_i}{t_k} > 0$ for $i \in I_x(y(x)) \setminus J(u)$ then $u_i^k > 0$ for large k. In that case $i \in J(u^k)$, and $g_i(x^k, y(x^k)) = 0$ in (23b) for large k.
This implies

$$\lim_{k\to\infty} \frac{g_i(x^k, y(x^k)) - g_i(x, y(x))}{t_k} = \nabla_x g_i(x, y(x))d_x + \nabla_y g_i(x, y(x))y'(x; d_x) = 0,$$

whenever $d_{u_i} > 0$ and $i \in I_x(y(x)) \setminus J(u)$.
On the other hand, if $\nabla_x g_i(x, y(x))d_x + \nabla_y g_i(x, y(x))y'(x; d_x) < 0$ for $i \in I_x(y(x)) \setminus J(u)$ then from (24b), $g_i(x^k, y(x^k)) < 0$ for large k. In that case, $u_i^k = 0$ from (23d). Hence, $d_{u_i} = \lim\limits_{k\to\infty} \frac{u_i^k - u_i}{t_k} = 0$ whenever $\nabla_x g_i(x, y(x))d_x + \nabla_y g_i(x, y(x))y'(x; d_x) < 0$ for $i \in I_x(y(x)) \setminus J(u)$. Thus,

$$\left.\begin{array}{l} d_{u_i}(\nabla_x g_i(x, y(x))d_x + \nabla_y g_i(x, y(x))y'(x; d_x)) = 0, i \in I_x(y(x)) \setminus J(u), \\ d_{u_i} = 0, \quad i \notin I_x(y(x)), \quad d_{u_i} \geq 0, \; i \in I_x(y(x)) \setminus J(u). \end{array}\right\} \tag{25e}$$

It can be readily verified that (25a)-(25e) are in fact KKT optimality conditions of the convex quadratic programming problem $QP_{(u,v)}(x; d_x)$. Hence, $y'(x; d_x)$ uniquely solves this system for given d_x, where d_{u_i}, $i \in I_x(y(x))$ and d_{v_i}, $i \in \Lambda_q$ are the Lagrange multipliers associated with constraints in $QP_{(u,v)}(x; d_x)$.

We further use the following notations of the index sets:

$$\widehat{J}(u; d_u) := \{i \in I_x(y(x)) \setminus J(u) \; : \; d_{u_i} > 0\} \cup J(u).$$

$$I_x(y(x); d_x) := \{i \in I_x(y(x)) : \nabla_x g_i(x, y(x))d_x + \nabla_y g_i(x, y(x))y'(x; d_x) = 0\}.$$

Proceeding in a similar manner as in the formulation of (25a)-(25e), using the index sets $\widehat{J}(u; d_u)$ and $I_x(y(x); d_x)$, we obtain

$$0.5 \begin{pmatrix} d_z^T & d_u^T & d_v^T \end{pmatrix} \widehat{\nabla}^2(\nabla_y L(x, y(x), u, v)^T) \begin{pmatrix} d_z \\ d_u \\ d_v \end{pmatrix} + \nabla^2_{yy} L(x, y(x), u, v)y''(x; d_x) + \nabla^2_{uy} L(x, y(x), u, v)d_u^1 + \nabla^2_{vy} L(x, y(x), u, v)d_v^1 = 0, \tag{26a}$$

$$0.5 d_z^T \nabla^2_{zz} g_i(x, y(x))d_z + \nabla_y g_i(x, y(x))y''(x; d_x) = 0, i \in \hat{J}(u; d_u), \tag{26b}$$

$$0.5 d_z^T \nabla^2_{zz} g_i(x, y(x))d_z + \nabla_y g_i(x, y(x))y''(x; d_x) \leq 0, \; i \in I_x(y(x); d_x) \setminus \hat{J}(u; d_u), \tag{26c}$$

$$0.5 d_z^T \nabla^2_{zz} h_i(x, y)d_z + \nabla_y h_i(x, y(x))y''(x; d_x) = 0, \; i \in \Lambda_q, \tag{26d}$$

$$\left.\begin{array}{l} d_{u_i}^1 \left(0.5 d_z^T \nabla^2_{zz} g_i(x, y(x))d_z + \nabla_y g_i(x, y(x))y''(x; d_x)\right) = 0, \\ d_{u_i}^1 \geq 0, \; i \in I_x(y(x); d_x) \setminus \hat{J}(u; d_u), \; d_{u_i}^1 = 0, i \notin I_x(y(x); d_x). \end{array}\right\} \tag{26e}$$

(d_u, d_v) satisfying (25a)–(25e) is not necessarily unique since it corresponds to the Lagrange multipliers for $QP_{(u,v)}(x, d_x)$. However, $(y''(x; d_x), d_u^1, d_v^1)$ satisfies the system (26a)–(26e) for given d_x regardless of any solution (d_u, d_v) in (25a)–(25e).

From Theorem 2.1 (i), $y'(x; d_x)$ is the solution of $QP_{(u,v)}(x; d_x)$ for given $(u, v) \in S^2(d_x) \subseteq S^1(d_x)$ and $d_{u_i}, i \in I_x(y(x))$ and $d_{v_i}, i \in \Lambda_q$ are the associated Lagrange multipliers. Consider the vector $d_u \in \mathbb{R}^p$ as

$$d_u = \begin{cases} d_{u_i} & i \in I_x(y(x)) \\ 0, & \textit{otherwise} \end{cases},$$

and a set $M_{(u,v)}$ collecting the vectors (d_u, d_v) for some fixed $(u,v) \in S^2(d_x)$. $M_{(u,v)}$ is a convex polyhedral set.
Further, denote the index set $K(v) = \{i \in \Lambda_q \ : \ v_i \neq 0\}$, and let $\overline{EM}_{(u,v)}$ be obtained from $M_{(u,v)}$ so that the family

$$\Big((\nabla_y g_i(x, y(x))_{i \in J(u) \cup J(d_u)}, \big(\nabla_y h_i(x, y(x))\big)_{i \in K(v) \cap K(d_v)} \Big)$$

is linear independent. As $(y'(x; d_x), d_u, d_v)$ satisfies the KKT optimality conditions (25a)-(25e) for $(u,v) \in S^2(d_x) \cap EU_x(y(x))$ and $(d_u, d_v) \in \overline{EM}_{(u,v)}$, so $\lim_{k\to\infty} \frac{(\bar{u}^k, \bar{v}^k) - (u,v)}{t_k} = (d_u, d_v)$ possibly over a subsequence of $\{x_k\}$ and $(\bar{u}^k, \bar{v}^k) \in \bar{U}_{x^k}(y(x^k); d_x)$.

Lemma 5.1. *Suppose Assumptions $A1 - A4$ hold at $(x, y(x))$, and $(y'(x; d_x), d_u, d_v)$ satisfies the conditions* (25a)-(25e) *for $(u,v) \in S^2(d_x) \cap EU_x(y(x))$ and $(d_u, d_v) \in \overline{EM}_{(u,v)}$.*
Then, $d_y := y''(x; d_x)$ satisfies the following system:

$$\left.\begin{aligned} &0.5 d_z^T \nabla^2_{zz} g_i(x, y(x)) d_z + \nabla_y g_i(x, y(x)) d_y = 0, \ \ i \in \hat{J}(u; d_u), \\ &0.5 d_z^T \nabla^2_{zz} g_i(x, y(x)) d_z + \nabla_y g_i(x, y(x)) d_y \leq 0, \ \ i \in I_x(y(x); d_x) \setminus \hat{J}(u; d_u), \\ &0.5 d_z^T \nabla^2_{zz} h_i(x, y(x)) d_z + \nabla_y h_i(x, y(x)) d_y = 0, \ \ i \in \Lambda_q. \end{aligned}\right\} \tag{27}$$

Proof. Since $y''(x; d_x)$ satisfies (26b)-(26d) for each $(u,v) \in V(d_x)$ therefore $y''(x; d_x)$ satisfies System (2) for each $(u,v) \in V(d_x)$. From Lemma 4.1, System (2) remains unchanged irrespective of the selection $(u,v) \in S^2(d_x)$. Hence, $y''(x; d_x)$ satisfies System (2) for each $(u,v) \in S^2(d_x)$ as $V(d_x) \subseteq S^2(d_x)$.
Next to show that $y''(x; d_x)$ satisfies (27) for $(u,v) \in S^2(d_x) \cap EU_x(y(x))$ and $(d_u, d_v) \in \overline{EM}_{(u,v)}$. To prove the lemma, it is sufficient to prove that
$0.5 d_z^T \nabla^2_{zz} g_i(x, y(x)) d_z + \nabla_y g_i(x, y(x)) y''(x; d_x) = 0$ for $i \in \hat{J}(u; d_u) \setminus J(u)$.

From Lemma 4.3, for $(u,v) \in S^2(d_x)$ there is a sequence of positive real numbers $\{t_k\}$ converging to 0, $x^k = x + t_k d_x$ and $(\bar{u}^k, \bar{v}^k) \in \bar{U}_{x^k}(y(x^k); d_x)$, so that $\{(\bar{u}^k, \bar{v}^k)\} \to (u,v)$. Since, $(y'(x; d_x), d_u, d_v)$ satisfies the conditions (25a)-(25e) for $(u,v) \in S^2(d_x) \cap EU_x(y(x))$ and $(d_u, d_v) \in \overline{EM}_{(u,v)}$, $\lim_{k\to\infty} \frac{(\bar{u}^k, \bar{v}^k) - (u,v)}{t_k} = (d_u, d_v)$ possibly over a subsequence of $\{x_k\}$ and $(\bar{u}^k, \bar{v}^k) \in \bar{U}_{x^k}(y(x^k); d_x)$.

Consider $i \in \hat{J}(u; d_u) \setminus J(u)$. Then $u_i = 0$, and $d_{u_i} > 0$. Hence, $\bar{u_i}^k > 0$ for large k. From Lemma (4.3), $\bar{U}_{x^k}(y(x^k); d_x) \neq \emptyset$, and (22) is satisfied in $(y(x^k), \bar{u}^k, \bar{v}^k)$. Since $u_i = 0$ therefore $G_i(x^k; d_x) = 0$.
Hence, from the complementarity constraints $\bar{u_i}^k (g_i(x^k, y(x^k)) - G_i(x^k; d_x)) = 0$ in (22), we obtain $g_i(x^k, y(x^k)) = 0$ for large k.
Since $\hat{J}(u; d_u) \subseteq I_x(y(x); d_x) \subseteq I_x(y(x))$ therefore $g_i(x,y) = 0$ and $\nabla_x g_i(x, y(x)) d_x + \nabla_y g_i(x, y(x)) y'(x; d_x) = 0$ for $i \in \hat{J}(u; d_u) \setminus J(u)$. Hence,

$$\lim_{k\to\infty} \frac{g_i(x^k, y(x^k)) - g_i(x, y(x)) - t_k(\nabla_x g_i(x, y(x)) d_x + \nabla_y g_i(x, y(x)) y'(x; d_x))}{t_k^2} = 0.$$

Then, using Taylor's expansion of g_i for $i \in \hat{J}(u; d_u) \setminus J(u)$ about (x,y) as in (24b), we get

$$0.5 d_z^T \nabla^2_{zz} g_i(x, y(x)) d_z + \nabla_y g_i(x, y(x)) y''(x; d_x) = 0.$$

This proves the result. □

In the following theorem, the second-order directional derivative of the optimal solution function is computed using the second-order approximation of the parametric problem.

Theorem 5.1. *Suppose Assumptions $A1-A5$ are satisfied at $(x, y(x))$. Then, for each $(u,v) \in S^2(d_x) \cap EU_x(y(x))$, $y''(x;d_x)$ is the unique solution of the following quadratic programming problem.*

$$(QCP_{(u,v)}(x;d_x)): \quad \min_{d_y} \quad \left(0.5\begin{pmatrix} d_z^T & d_u^T & d_v^T \end{pmatrix} \widehat{\nabla}^2 \left(\nabla_y L(x,y(x),u,v)^T\right) \begin{pmatrix} d_z \\ d_u \\ d_v \end{pmatrix}\right)^T d_y + {d_y}^T \nabla^2_{yy} L(x,y(x),u,v) d_y$$

$$\text{s.t. } 0.5 d_z^T \nabla^2_{zz} g_i(x,y(x)) d_z + \nabla_y g_i(x,y(x)) d_y = 0, \; i \in \hat{J}(u;d_u), \tag{28a}$$

$$0.5 d_z^T \nabla^2_{zz} g_i(x,y(x)) d_z + \nabla_y g_i(x,y(x)) d_y \leq 0, \; i \in I_x(y(x);d_x) \setminus \hat{J}(u;d_u), \tag{28b}$$

$$0.5 d_z^T \nabla^2_{zz} h_i(x,y(x)) d_z + \nabla_y h_i(x,y(x)) d_y = 0, \; i \in \Lambda_q. \tag{28c}$$

where $d_z = \begin{pmatrix} d_x \\ y'(x;d_x) \end{pmatrix}$ *and* $(y'(x;d_x), d_u, d_v)$ *solves* (25a)-(25e) *for* $(d_u, d_v) \in \overline{EM}_{(u,v)}$ *at* $(x, y(x), u, v)$ *.*

Proof. As Assumption $A3$ holds for P_x at $(x,y(x))$, $QCP_{(u,v)}(x;d_x)$ satisfies the second order sufficient optimality conditions. Hence, it has a unique solution.

From Lemma 5.1, $y''(x;d_x)$ satisfies (27), which is same as (28a)-(28c). Hence $QCP_{(u,v)}(x;d_x)$ has non-empty feasible set. We show that $y''(x;d_x)$ solves $QCP_{(u,v)}(x;d_x)$ uniquely for any $(u,v) \in S^2(d_x)$. Consider the following two cases.

Case(i) $(u,v) \in V(d_x) \cap EU_x(y(x))$: The system (26a)-(26e) represents the KKT optimality conditions of $QCP_{(u,v)}(x;d_x)$ for given (u,v) and $(y''(x;d_x), d_u^1, d_v^1)$ satisfies this system. $d^1_{u_i}$, $i \in I_x(y(x);d_x)$ and $d^1_{v_i}$, $i \in \Lambda_q$ are the Lagrange multipliers associated with the constraints of $QCP_{(u,v)}(x;d_x)$. Hence, $y''(x;d_x)$ is the optimal solution of $QCP_{(u,v)}(x;d_x)$.
Case(ii) $(u,v) \in (S^2(d_x) \setminus V(d_x)) \cap EU_x(y(x))$: To prove $y''(x;d_x)$ solves $QCP_{(u,v)}(x;d_x)$, it is sufficient to show that the following variational inequality holds at $y''(x;d_x)$ for (u,v) and feasible d_y satisfying (28a)-(28c).

$$\left\langle 0.5\begin{pmatrix} d_z^T & d_u^T & d_v^T \end{pmatrix} \widehat{\nabla}^2 \left(\nabla_y L(x,y(x),u,v)^T\right) \begin{pmatrix} d_z \\ d_u \\ d_v \end{pmatrix} + \nabla^2_{yy} L(x,y(x),u,v) y''(x;d_x) \quad d_y - y''(x;d_x) \right\rangle \geq 0. \tag{29}$$

As $(y'(x;d_x), d_u, d_v)$ solves (25a)-(25e) for $(d_u, d_v) \in \overline{EM}_{(u,v)}$ at $(x, y(x), u, v)$, $\lim_{k\to\infty} \frac{(\bar{u}^k, \bar{v}^k)-(u,v)}{t_k} = (d_u, d_v)$ for some appropriate subsequence of $\{x_k\}$ and $(\bar{u}^k, \bar{v}^k) \in \bar{U}_{x^k}(y(x^k);d_x)$. Hence, from (22) for each k in the subsequence, it follows that

$$\nabla_y L(x^k, y(x^k), \bar{u}^k, \bar{v}^k) = \nabla_y f(x^k, y(x^k)) + \sum_{i\in\Lambda_p} \bar{u_i}^k \nabla_y g_i(x^k, y(x^k)) + \sum_{i\in\Lambda_q} \bar{v_i}^k \nabla_y h_i(x^k, y(x^k)) = 0. \tag{30}$$

As $(u,v) \in S^2(d_x) \subseteq U_x(y(x))$,

$$\nabla_y L(x, y(x), u, v) = \nabla_y f(x, y(x)) + \sum_{i\in\Lambda_p} u_i \nabla_y g_i(x, y(x)) + \sum_{i\in\Lambda_q} v_i \nabla_y h_i(x, y(x)) = 0. \tag{31}$$

Using Taylor's expansion of $\nabla_y L$ up-to second order about $(x, y(x), u, v)$(see (24a)) and using (25a), (30), and (31), we can obtain

$$0.5\begin{pmatrix} d_z^T & d_u^T & d_v^T \end{pmatrix} \widehat{\nabla}^2 \left(\nabla_y L(x,y(x),u,v)^T\right) \begin{pmatrix} d_z \\ d_u \\ d_v \end{pmatrix} + \nabla^2_{yy} L(x,y(x),u,v) y''(x;d_x)$$

$$+ \lim_{k\to\infty} \sum_{i\in I_x(y(x))} \frac{\bar{u}_i^k - u_i - t_k d_{u_i}}{t_k^2} \nabla_y g_i(x,y(x))^T + \lim_{k\to\infty} \sum_{i\in \Lambda_q} \frac{\bar{v}_i^k - v_i - t_k d_{v_i}}{t_k^2} \nabla_y h_i(x,y(x))^T = 0$$

Hence,

$$\begin{aligned}
&\left\langle 0.5\begin{pmatrix} d_z^T & d_u^T & d_v^T \end{pmatrix} \widehat{\nabla}^2 \left(\nabla_y L(x,y(x),u,v)^T\right) \begin{pmatrix} d_z \\ d_u \\ d_v \end{pmatrix} + \nabla^2_{yy} L(x,y(x),u,v) y''(x;d_x) \quad d_y - y''(x;d_x) \right\rangle \\
&= \left\langle -\lim_{k\to\infty} \sum_{i\in I_x(y(x))} \frac{\bar{u}_i^k - u_i - t_k d_{u_i}}{t_k^2} \nabla_y g_i(x,y(x))^T - \lim_{k\to\infty} \sum_{i\in \Lambda_q} \frac{\bar{v}_i^k - v_i - t_k d_{v_i}}{t_k^2} \nabla_y h_i(x,y(x))^T \quad d_y - y''(x;d_x) \right\rangle \\
&= -\lim_{k\to\infty} \sum_{i\in I_x(y(x))} \frac{\bar{u}_i^k - u_i - t_k d_{u_i}}{t_k^2} \left\langle \nabla_y g_i(x,y(x))^T \quad d_y - y''(x;d_x) \right\rangle \\
&\qquad - \lim_{k\to\infty} \sum_{i\in \Lambda_q} \frac{\bar{v}_i^k - v_i - t_k d_{v_i}}{t_k^2} \left\langle \nabla_y h_i(x,y(x))^T \quad d_y - y''(x;d_x) \right\rangle .
\end{aligned} \tag{32}$$

Since $y''(x;d_x)$ satisfies (28a) for $i \in \hat{J}(u;d_u)$,

$$0.5 d_z^T \nabla^2_{zz} g_i(x,y(x)) d_z + \nabla_y g_i(x,y(x)) y''(x;d_x) = 0.$$

Similarly, for any d_y satisfying (28a) for $i \in \hat{J}(u;d_u)$, we have

$$0.5 d_z^T \nabla^2_{zz} g_i(x,y(x)) d_z + \nabla_y g_i(x,y(x)) d_y = 0.$$

From last two relations,

$$\left\langle \nabla_y g_i(x,y(x))^T \quad d_y - y''(x;d_x) \right\rangle = 0, \quad \forall i \in \hat{J}(u;d_u).$$

Again, for $d_y,\ y''(x;d_x)$ satisfying (28c), it follows that $0.5 d_z^T \nabla^2_{zz} h_i(x,y(x)) d_z + \nabla_y h_i(x,y(x)) d_y = 0$ and $0.5 d_z^T \nabla^2_{zz} h_i(x,y(x)) d_z + \nabla_y h_i(x,y(x)) y''(x;d_x) = 0$ for $i \in \Lambda_q$. Hence,

$$\left\langle \nabla_y h_i(x,y(x))^T \quad d_y - y''(x;d_x) \right\rangle = 0, \quad i \in \Lambda_q.$$

For $i \in I_x(y(x)) \setminus \hat{J}(u;d_u)$, it follows that $u_i = 0, d_{u_i} = 0$. Then, inserting last two expression in (32), we obtain

$$\begin{aligned}
&\left\langle 0.5\begin{pmatrix} d_z^T & d_u^T & d_v^T \end{pmatrix} \widehat{\nabla}^2 \left(\nabla_y L(x,y(x),u,v)^T\right) \begin{pmatrix} d_z \\ d_u \\ d_v \end{pmatrix} + \nabla^2_{yy} L(x,y(x),u,v) y''(x;d_x) \quad d_y - y''(x;d_x) \right\rangle \\
&\qquad = -\lim_{k\to\infty} \sum_{I_x(y(x))\setminus \hat{J}(u;d_u)} \frac{\bar{u}_i^k}{t_k^2} \left\langle \nabla_y g_i(x,y(x))^T \quad d_y - y''(x;d_x) \right\rangle
\end{aligned} \tag{33}$$

Let $\nabla_x g_i(x,y(x)) d_x + \nabla_y g_i(x,y(x)) y'(x;d_x) < 0$ for $i \in I_x(y(x)) \setminus \hat{J}(u;d_u))$.
Using this inequality in the Taylor's expansion of g_i about $(x,y(x))$ as in (24b), we obtain $g_i(x^k, y(x^k)) < 0$ for large k. Since $u_i = 0$ for $i \in I_x(y(x)) \setminus \hat{J}(u;d_u)$, it follows that $G_i(x^k;d_x) = 0$ for $i \in I_x(y(x)) \setminus \hat{J}(u;d_u)$.

Then from the complementarity constraints $\bar{u}_i^k \left(g_i(x^k, y(x^k)) - G_i(x^k; d_x)\right) = 0$, in (22), we get $\bar{u}_i{}^k = 0$ for large k. Hence,

$$\frac{\bar{u}_i^k}{t_k^2} = 0, \quad i \in \{i \in I_x(y(x)) \setminus \hat{J}(u; d_u) : \ \nabla_x g_i(x, y(x))d_x + \nabla_y g_i(x, y(x))y'(x; d_x) < 0\} \tag{34}$$

for large k.

Next, consider the other case for which $i \in I_x(y(x)) \setminus \hat{J}(u; d_u)$ and $\nabla_x g_i(x, y(x))d_x + \nabla_y g_i(x, y(x))y'(x; d_x) = 0$. Thus, $i \in I_x(y(x); d_x) \setminus \hat{J}(u; d_u)$. Since $y''(x; d_x)$ satisfies (28b), we have

$$0.5 d_z^T \nabla_{zz}^2 g_i(x, y(x))d_z + \nabla_y g_i(x, y(x))y''(x; d_x) \leq 0, \quad i \in I_x(y(x); d_x) \setminus \hat{J}(u; d_u).$$

Here raises two cases for $i \in I_x(y(x); d_x) \setminus \hat{J}(u; d_u)$.

- First, if $0.5 d_z^T \nabla_{zz}^2 g_i(x, y(x))d_z + \nabla_y g_i(x, y(x))y''(x; d_x) < 0$ then using second order Taylor's expansion of g_i about $(x, y(x))$ as in (24b), we obtain that $g_i(x^k, y(x^k)) < 0$. Since $u_i = 0$ for $i \in I_x(y(x); d_x) \setminus \hat{J}(u; d_u)$, it follows that $G_i(x^k; d_x) = 0$ for $i \in I_x(y(x); d_x) \setminus \hat{J}(u; d_u)$. Hence, from the complementarity constraints $\bar{u}_i^k \left(g_i(x^k, y(x^k)) - G_i(x^k; d_x)\right) = 0$ in (22), $\bar{u}_i^k = 0$ for large k. Thus,

  $$\frac{\bar{u}_i^k}{t_k^2} = 0, \quad i \in \{i \in I_x(y(x); d_x) \setminus \hat{J}(u; d_u) : \ 0.5 d_z^T \nabla_{zz}^2 g_i(x, y(x))d_z + \nabla_y g_i(x, y(x))y''(x; d_x) < 0\} \tag{35}$$

  for large k.
- In the other case, $i \in I_x(y(x); d_x) \setminus \hat{J}(u; d_u)$ and

  $$0.5 d_z^T \nabla_{zz}^2 g_i(x, y(x))d_z + \nabla_y g_i(x, y(x))y''(x; d_x) = 0.$$

  Since d_y satisfies (28b), for $i \in I_x(y(x); d_x) \setminus \hat{J}(u; d_u)$, we have

  $$0.5 d_z^T \nabla_{zz}^2 g_i(x, y(x))d_z + \nabla_y g_i(x, y(x))d_y \leq 0.$$

  From last two relations, we obtain $\left\langle \nabla_y g_i(x, y(x))^T \quad d_y - y''(x; d_x) \right\rangle \leq 0$.
  Hence, for $i \in \{i \in I_x(y(x); d_x) \setminus \hat{J}(u; d_u) : \ 0.5 d_z^T \nabla_{zz}^2 g_i(x, y(x))d_z + \nabla_y g_i(x, y(x))y''(x; d_x) = 0\}$, we deduce that

  $$-\lim_{k \to \infty} \frac{\bar{u}_i^k}{t_k^2} \left\langle \nabla_y g_i(x, y(x))^T \quad d_y - y''(x; d_x) \right\rangle \geq 0. \tag{36}$$

  Using (34),(35), and (36), one can see that the right hand side of (33) is nonnegative . Thus, (29) holds. Hence, $y''(x; d_x)$ solves $QCP_{(u,v)}(x; d_x)$ for any $(u, v) \in S^2(d_x)$.

□

We conclude this section with an example. Different steps of the theoretical results to compute $y''(x; d_x)$ as the solution of $QCP_{(u,v)}(x; d_x)$ is explained here.

Example 5.1.

$$\begin{aligned} &\min_y (y-5)^2 \\ &s.t.\ x_2^2 + y x_1^2 - 3 \leq 0, \\ &\quad x_1 + 2y - 7 \leq 0. \end{aligned}$$

Suppose $x^0 = (1,0)^T$ and $d_x = (0,1)^T$ are given. The Lagrange function of the problem is

$$L(x,y,u) = (y-5)^2 + u_1(x_2^2 + yx_1^2 - 3) + u_2(x_1 + 2y - 7),$$

where u_1 and u_2 are the Lagrange multipliers associated with constraints $x_2^2+yx_1^2-3 \le 0$ and $x_1+2y-7 \le 0$, respectively. The KKT optimality conditions for the example are

$$\begin{aligned}
&\nabla_y L(x,y,u) = 2(y-5) + u_1(x_1^2) + 2u_2 = 0,\\
&u_1 \ge 0, x_2^2 + yx_1^2 - 3 \le 0\ , \quad u_1(x_2^2 + yx_1^2 - 3) = 0,\\
&u_2 \ge 0,\ x_1 + 2y - 7 \le 0, \quad u_2(x_1 + 2y - 7 \le 0) = 0.
\end{aligned}$$

The solution to the example is $y(x^0) = 3$. The set of Lagrange multipliers at x^0 is

$$U_{x^0}(y(x^0)) = \left\{ (u_1,u_2)\ :\ \begin{matrix} u_1 + 2u_2 = 4, \\ u_1 \ge 0, u_2 \ge 0. \end{matrix} \right\} = \text{convex hull of} \left\{ \begin{pmatrix} 4 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \end{pmatrix} \right\},$$

which is the line segment joining $(4,0)$ and $(0,2)$. Computation of the set $S^1(d_x)$ at x^0 is as follows.

$$\begin{aligned}
S^1(d_x) &= \operatorname*{arg\,max}_{(u_1,u_2)\in U_{x^0}(y(x^0))} \quad u_1 \nabla_x g_1(x^0, y(x^0)) d_x + u_2 \nabla_x g_2(x^0, y(x^0)) d_x \\
&= \operatorname*{arg\,max}_{(u_1,u_2)\in U_{x^0}(y(x^0))} \quad u_1.(2x_1^0 y(x^0) d_{x_1} + 2x_2^0 d_{x_2}) + u_2.(2d_{x_1}) \\
&= \operatorname*{arg\,max}_{(u_1,u_2)\in U_{x^0}(y(x^0))} \quad u_1.(0) + u_2.(0) \\
&= U_{x^0}(y(x^0)).
\end{aligned}$$

Since all the assumptions of Theorem 2.1(i) are satisfied at $(x^0, y(x^0))$, therefore, for any $(u_1,u_2) \in S^1(d_x)$ solution of $QP_{(u_1,u_2)}(x^0; d_x)$ provides $y'(x^0; d_x)$ at x^0. At $(u_1,u_2) = \big(4,0\big)$ and $d_x = (0,1)^T$, we have $J(u) = \{1\}$ and $I_{x^0}(y(x^0)) = \{1,2\}$.

$$\begin{aligned}
(QP_{(u_1,u_2)}(x^0; d_x)): \quad &\min_{d_y} \quad \frac{1}{2} d_y^T \nabla^2_{yy} L(x^0, y(x^0), u) d_y + d_y^T \nabla^2_{xy} L(x^0, y(x^0), u) d_x \\
&\text{s.t. } \nabla_x g_1(x^0, y(x^0)) d_x + \nabla_y g_1(x^0, y(x^0)) d_y = 0, \\
&\quad \nabla_x g_2(x^0, y(x^0)) d_x + \nabla_y g_2(x^0, y(x^0)) d_y \le 0.
\end{aligned}$$

After substituting the value of L, g_1 and g_2, and d_x, x^0 and $y(x^0)$

$$\begin{aligned}
&\min_{d_y} d_y^2 + 2u_1 x_1^0 d_{x_1} d_y \\
&\text{s.t. } 2y(x^0) x_1^0 d_{x_1} + 2x_2^0 d_{x_2} + (x_1^0)^2 d_y = 0, \\
&\quad d_{x_1} + 2d_y \le 0.
\end{aligned}$$

Putting the value of d_x, x^0 and $y(x^0)$, the solution to the above problem is $d_y^* = y'(x^0; d_x) = 0$ and Lagrange multiplier vector $d_u = (0,0)^T$. Next step is to compute the multiplier set $S^2(d_x)$ by solving P_x^2. Here,

$d_z = \begin{pmatrix} d_x \\ y'(x^0; d_x) \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$, and $S^2(d_x)$ is obtained by solving P_x^2 at x^0 as

$$\begin{aligned} S^2(d_x) &= \underset{u \in S^1(d_x)}{\arg\max} \;\; u_1.d_z^T \nabla^2_{zz} g_1(x^0, y(x^0)) d_z + u_2.d_z^T \nabla^2_{zz} g_2(x^0, y(x^0)) d_z \\ &= \underset{u \in S^1(d_x)}{\arg\max} \;\; u_1. \begin{pmatrix} d_{x_1} \\ d_{x_2} \\ y'(x^0; d_x) \end{pmatrix}^T \begin{bmatrix} 6 & 0 & 2 \\ 0 & 2 & 0 \\ 2 & 0 & 0 \end{bmatrix} \begin{pmatrix} d_{x_1} \\ d_{x_2} \\ y'(x^0; d_x) \end{pmatrix} + u_2 \begin{pmatrix} d_{x_1} \\ d_{x_2} \\ y'(x^0; d_x) \end{pmatrix}^T \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{pmatrix} d_{x_1} \\ d_{x_2} \\ y'(x^0; d_x) \end{pmatrix} \\ &= \underset{u \in S^1(d_x)}{\arg\max} \;\; u_1(6d_{x_1}^2 + 4y'(x^0; d_x) d_{x_1} + 2d_{x_2}^2) + u_2.0 \\ &= \underset{u \in S^1(d_x)}{\arg\max} \;\; 2u_1 \\ &= (4, 0) \end{aligned}$$

$y''(x^0; d_x)$ is the solution of $QCP_{(u_1,u_2)}(x^0; d_x)$ for $(u_1, u_2) = (4, 0) \in S^2(d_x)$. Here we have,

$$\begin{aligned} 0.5 \begin{pmatrix} d_z^T & d_u^T \end{pmatrix} \widehat{\nabla}^2 \left(\nabla_y L(x, y(x), u, v)^T \right) \begin{pmatrix} d_z \\ d_u \end{pmatrix} &= \begin{pmatrix} d_{x_1} \\ d_{x_2} \\ y'(x^0; d_x) \\ d_{u_1} \\ d_{u_2} \end{pmatrix}^T \begin{pmatrix} u_1 & 0 & 0 & x_1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ x_1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} d_{x_1} \\ d_{x_2} \\ y'(x^0; d_x) \\ d_{u_1} \\ d_{u_2} \end{pmatrix} \\ &= u_1 d_{x_1}^2 + 2x_1 d_{x_1} d_{u_1} \\ &= 0, \\ 0.5 d_z^T \nabla^2_{zz} g_1(x^0, y(x^0)) d_z &= 3d_{x_1}^2 + 2y'(x^0; d_x) d_{x_1} + d_{x_2}^2 \\ &= 1, \\ \text{and } 0.5 d_z^T \nabla^2_{zz} g_2(x^0, y(x^0)) d_z &= 0. \end{aligned}$$

Here $J(u; d_u) = 1$ and $I_{x^0}(y(x^0); d_x) = \{1, 2\}$. Hence $QCP_{(u_1,u_2)}(x^0; d_x)$ takes the following form

$$\begin{aligned} (QCP_{(u_1,u_2)}(x^0; d_x)) : \quad \min_{d_y} \;& \left(0.5 \begin{pmatrix} d_z^T & d_u^T \end{pmatrix} \widehat{\nabla}^2 \left(\nabla_y L(x, y(x), u, v)^T \right) \begin{pmatrix} d_z \\ d_u \end{pmatrix} \right)^T d_y + {d_y}^T \nabla^2_{yy} L(x, y(x), u) d_y \\ \text{s.t. } & 0.5 d_z^T \nabla^2_{zz} g_1(x, y(x)) d_z + \nabla_y g_1(x, y(x)) d_y = 0, \\ & 0.5 d_z^T \nabla^2_{zz} g_2(x, y(x)) d_z + \nabla_y g_2(x, y(x)) d_y \leq 0. \end{aligned}$$

Substituting values of L, g_1 and g_2, the above problem reduces to

$$\min_{d_y} 2d_y^2 \;\; \text{s.t. } 1 + d_y = 0, \quad d_y \leq 0.$$

The solution to the above problem is $y''(x^0; d_x) = -1$.

Consider the path $x(t) = x^0 + td_x = (1, 0)^T + t(0, 1)^T = (1, t)^T$ for $t \geq 0$. From the first and second constraints, $y \leq 3 - t^2$ and $y \leq 3$. This reduces to $y \leq 3 - t^2$ as $3 - t^2 \leq 3$ for $t \geq 0$. Thus $y((x(t)) = 3 - t^2$. In that case, we have $\lim_{t \downarrow 0} \frac{y(x(t)) - y(x^0)}{t} = 0$ and $\lim_{t \downarrow 0} \frac{y(x(t)) - y(x^0) - ty'(x^0; d_x)}{t^2} = -1$.

6. Conclusion

This work contributes a practical method for computing the second-order directional derivative of smooth parametric programming problems. It investigates the case where the associated Lagrange multipliers in the

KKT optimality conditions of the parametric problem are not necessarily unique. To derive the directional derivatives, one requires the Lagrange multiplier subset $V(d_x) \subseteq U_x(y(x))$, which cannot be computed explicitly without imposing restrictive assumptions. In the existing literature, the inclusion of the set $V(d_x)$ as $S^1(d_x)$ was obtained under Assumptions $A1$–$A4$, which provides the first-order directional derivative by solving a quadratic programming problem. In this work, the idea is extended to the second order, and we identify a subset $S^2(d_x)$ of multipliers from $S^1(d_x)$. For each multiplier in $S^2(d_x)$, the second-order derivative $y''(x; d_x)$ is obtained by solving a quadratic programming problem $QCP_{(u,v)}(x; d_x)$, which constitutes the main result of this work. This study essentially extends the results established in Ralph and Dempe [10] and Dempe [3] using the second order approximation of the parametric problem.The findings are significant in optimization theory and can be applied to derive second-order optimality conditions for bilevel programming problems.

References

[1] Auslender, A. and Cominetti, R. (1990). First and second order sensitivity analysis of nonlinear programs under directional constraint qualification conditions. *Optimization*, 21(3):351–363.

[2] Bonnans, J. F. and Shapiro, A. (2013). *Perturbation analysis of optimization problems*. Springer Science & Business Media.

[3] Dempe, S. (1993). Directional differentiability of optimal solutions under slater's condition. *Mathematical programming*, 59:49–69.

[4] Fan, K. (1952). Fixed-point and minimax theorems in locally convex topological linear spaces. *Proceedings of the National Academy of Sciences*, 38(2):121–126.

[5] Fiacco, A. V. (1983). Introduction to sensitivity and stability analysis in non linear programming.

[6] Gauvin, J. and Janin, R. (1988). Directional behaviour of optimal solutions in nonlinear mathematical programming. *Mathematics of Operations Research*, 13(4):629–649.

[7] Kojima, M. (1980). Strongly stable stationary solutions in nonlinear programs. In *Analysis and computation of fixed points*, pages 93–138. Elsevier, the University of Wisconsin–Madison.

[8] Liu, J. (1995). Sensitivity analysis in nonlinear programs and variational inequalities via continuous selections. *SIAM Journal on Control and Optimization*, 33(4):1040–1060.

[9] Luderer, B., Minchenko, L., and Satsura, T. (2013). *Multivalued analysis and nonlinear programming problems with perturbations*, volume 66. Springer Science & Business Media.

[10] Ralph, D. and Dempe, S. (1995). Directional derivatives of the solution of a parametric nonlinear program. *Mathematical programming*, 70(1):159–172.

[11] Shapiro, A. (1988a). Perturbation theory of nonlinear programs when the set of optimal solutions is not a singleton. *Applied Mathematics and Optimization*, 18(1):215–229.

[12] Shapiro, A. (1988b). Sensitivity analysis of nonlinear programs and differentiability properties of metric projections. *SIAM Journal on Control and Optimization*, 26(3):628–645.

[13] Stechlinski, P., Khan, K. A., and Barton, P. I. (2018). Generalized sensitivity analysis of nonlinear programs. *SIAM Journal on Optimization*, 28(1):272–301.