


Article

Existence of Generalized Augmented Lagrange Multipliers for Constrained Optimization Problems

Yue Wang ¹, Jinchuan Zhou ^{1,*}  and Jingyong Tang ²

¹ Department of Statistics, School of Mathematics and Statistics, Shandong University of Technology, Zibo 255049, China; wangyuezxyq@163.com

² School of Mathematics and Statistics, Xinyang Normal University, Xinyang 464000, China; tangjy@xynu.edu.cn

* Correspondence: jinchuanzhou@163.com

Received: 17 March 2020; Accepted: 22 April 2020; Published: 24 April 2020



Abstract: The augmented Lagrange multiplier as an important concept in duality theory for optimization problems is extended in this paper to generalized augmented Lagrange multipliers by allowing a nonlinear support for the augmented perturbation function. The existence of generalized augmented Lagrange multipliers is established by perturbation analysis. Meanwhile, the relations among generalized augmented Lagrange multipliers, saddle points, and zero duality gap property are developed.

Keywords: generalized augmented Lagrange multipliers; saddle points; duality theory

MSC: 90C26; 90C46

1. Introduction

This paper is concerned with the following nonlinear programming problem:

$$(P) \quad \begin{aligned} & \min_{x \in \Omega} f(x) \\ & \text{s.t. } g_i(x) \leq 0, i = 1, \dots, m, \\ & \quad h_j(x) = 0, j = 1, \dots, l, \end{aligned}$$

where Ω is a nonempty and closed subset in \mathbb{R}^n , $g_i(x) : \mathbb{R}^n \rightarrow \mathbb{R}$ for $i = 1, \dots, m$, and $h_j(x) : \mathbb{R}^n \rightarrow \mathbb{R}$ for $j = 1, \dots, l$ are continuous functions. For simplification of notation, let us denote $g(x) := (g_1(x), g_2(x), \dots, g_m(x))$, and $h(x) := (h_1(x), h_2(x), \dots, h_l(x))$. Note that the feasible region of (P) can be written as $\Omega \cap \mathcal{F}$, where

$$\mathcal{F} := \{x \mid g_i(x) \leq 0, i = 1, \dots, m; h_j(x) = 0, j = 1, \dots, l\}.$$

The classical Lagrangian function for the problem (P) is defined as

$$L(x, \lambda, \mu) := f(x) + \langle \lambda, g(x) \rangle + \langle \mu, h(x) \rangle, \quad (\lambda, \mu) \in \mathbb{R}_+^m \times \mathbb{R}^l.$$

A non-zero duality gap maybe arise for nonconvex optimization problems when using the above Lagrangian functions. Hence some modifications are necessary to overcome this difficulty, such as the augmented Lagrangian by introducing an augmented term, or the nonlinear Lagrangian by replacing the multiplier item and augmented term together by a nonlinear function. For example, the Hestenes–Powell–Rockafellar augmented Lagrangian [1–3], the cubic augmented

Lagrangian [4], Mangasarian's augmented Lagrangian [5,6], the exponential penalty function [7,8], the log-sigmoid Lagrangian [9], modified barrier functions [8,10], the p -th power augmented Lagrangian [11], and nonlinear augmented Lagrangian functions [12–15]. The other related discussion on augmented Lagrangians regarding special constrained optimization includes second-order cone programming [16,17], semidefinite programming [18–20], cone programming [21–23], semi-infinite programming [24,25], min-max programming [26], distributed optimization [27], mixed integer programming [28], stochastic mixed-integer programs [29], generalized Nash equilibrium problems [30], quasi-variational inequalities [31], composite convex programming [32], and sparse discrete problems [33].

The duality theory is closely related to the perturbation of primal problem. Precisely, for a given $(y, z) \in \mathbb{R}^m \times \mathbb{R}^l$, the perturbation problem of (P) is

$$(P_{(y,z)}) \quad \begin{aligned} & \min_{x \in \Omega} f(x) \\ & \text{s.t. } g_i(x) + y_i \leq 0, i = 1, \dots, m, \\ & \quad h_j(x) + z_j = 0, j = 1, \dots, l. \end{aligned}$$

Denote by $\text{val}(P)$ and $v(y, z) (:= \text{val}(P_{(y,z)}))$ the optimal values of (P) and $(P_{(y,z)})$, respectively. Clearly, $v(0, 0) = \text{val}(P)$. Denote by X^* the optimal solution set of problem (P) , and assume throughout the paper that the optimal value $\text{val}(P)$ is finite.

The augmented perturbation function is

$$v_r(y, z) := v(y, z) + r\sigma(y, z), \quad \forall (y, z) \in \mathbb{R}^m \times \mathbb{R}^l. \quad (1)$$

Here σ is called an augmenting function (see Section 2 below for details). Its properties are weakened from convex to level-bounded, or valley-at-zero. For example, in Rockafellar and Wets [34], a nonnegative convex augmenting function and the corresponding augmented Lagrangian dual problem of primal problem were introduced. A sufficient condition for the zero duality gap and a necessary and sufficient condition for the existence of an exact penalty representation were obtained. It was extended in [35] by replacing the convexity condition of the augmenting function with a level-boundedness condition. Using the theory of abstract convexity, a family of augmenting functions with almost peak at zero property and a class of corresponding augmented Lagrangian dual problems were introduced in [36]. Valley-at-zero property (similar to almost peak-at-zero property) was used in [37].

A vector (λ, μ) is said to be an augmented Lagrange multiplier for problem (P) (cf. [22,25]), if

$$v_r(y, z) \geq v_r(0, 0) + \langle \lambda, y \rangle + \langle \mu, z \rangle, \quad \forall (y, z) \in \mathbb{R}^m \times \mathbb{R}^l. \quad (2)$$

That means that (λ, μ) is a subgradient of $v_r(\cdot, \cdot)$ at $(y, z) = (0, 0)$. The set of all subgradients (λ, μ) is called the subdifferential of $v_r(y, z)$ at $(y, z) = (0, 0)$ and denoted by $\partial v_r(0, 0)$. Augmented Lagrange multipliers are an important concept in duality theory. Their existence is important for the global convergence analysis of primal-dual type algorithms based on the use of augmented Lagrangians [7,19,29,32,33]. In addition, augmented Lagrange multipliers are closely related to saddle points, the zero duality gap property, and exact penalty representation. Some results on the existence of augmented Lagrange multipliers are discussed for semi-infinite programming [25], cone programming [22,23], and eigenvalue composite optimization problems [38]. Moreover, CQ-free duality was proposed in the classical monograph [39] by Bonnans and Shapiro. The stronger results on CQ-free strong duality for semidefinite and general convex programming can be found in [40,41], and in more recent publications for semi-infinite, semidefinite, and copositive programming by Kostyukova and others [42,43]. Recently, Dolgopolik [44] studied the existence of augmented Lagrange multipliers for geometric constraint optimization by using the localization principle.

Recall that for convex programming, Lagrangian multiplier is a subgradient of perturbation function v at $u = 0$ in the sense of convex analysis; i.e.,

$$v(u) \geq v(0) + \langle \lambda, u \rangle, \quad \forall u.$$

For nonconvex programming, the Lagrangian multiplier can be used to estimate the subdifferential of the perturbation function at the origin. Precisely, for a minimization problem

$$\min f(x) + \theta(g(x)), \quad x \in X,$$

where $f : \mathbb{R}^n \rightarrow \mathbb{R}$, $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$, X is a closed set in \mathbb{R}^n , and $\theta : \mathbb{R}^m \rightarrow \bar{\mathbb{R}} := (\infty, +\infty]$ is proper, lsc, and convex. This model includes the constrained optimization problems (by letting θ be a indicator function) and composite optimization problems. Denote by S^* the solution set. For $\bar{x} \in S^*$, let

$$M(\bar{x}) := \{\lambda | 0 \in \nabla f(\bar{x}) + \lambda^T \nabla g(\bar{x}) + N_X(\bar{x}), \lambda \in \partial \theta(g(\bar{x}))\}$$

and

$$M^\infty(\bar{x}) := \{\lambda | 0 \in \lambda^T \nabla g(\bar{x}) + N_X(\bar{x}), \lambda \in N_{\text{dom}\theta}(g(\bar{x}))\}.$$

If X is regular and $M^\infty(\bar{x}) = \{0\}$ for every $\bar{x} \in S^*$, then

$$\partial v(0) \subset \bigcup_{\bar{x} \in S^*} M(\bar{x}), \quad \text{lip}v(0) \leq \max_{\substack{\lambda \in M(\bar{x}) \\ \bar{x} \in S^*}} \|\lambda\|. \tag{3}$$

It should be pointed out that the subdifferential that appeared in (3) is the limiting/Mordukhovich subdifferential, not a subdifferential in the sense of convex analysis. Here $M^\infty(\bar{x}) = \{0\}$ can be regarded as constraint qualification. In particular, if $\text{dom}\theta := \mathbb{R}^l \times \{0\}^l$ and $X := \mathbb{R}^n$, then this condition is *Mangasarian–Fromovitz constraint qualification*; if $\text{dom}\theta$ is a convex cone with nonempty interior and $X := \mathbb{R}^n$, then this condition is *Robinson’s constraint qualification*. The result (3) indicates that the Lagrangian multiplier provides an upper bound on the subdifferential of perturbation function and gives an estimate on the Lipschitz constant of perturbed function. It is very important for the convergence analysis of numerical algorithms.

Compared with the classical Lagrangian function, the augmented Lagrangian function has been successfully applied to study nonconvex programming. Hence an interesting question is how to use the augmented Lagrangian multiplier to study the subdifferential of v_τ , and further give an estimate on Lipschitz constant on v_τ . On subdifferentiability in nonconvex setting, Clarke’s pioneering work on generalized gradient opened the door to the study of general nonsmooth functions. Many concepts were introduced in the past few decades. Frequently used concepts include limiting/Mordukhovich subdifferential, Ioffe’s approximate and G-subdifferential, Michel and Penot’s subdifferential, Treiman’s linear subdifferential, Sussmann’s semidifferential, etc. Compared with the abstract subdifferential (pioneered by Warga), which is defined by a set of axioms, many subdifferentials have reasonable geometric explanations. For example, a convex subdifferential means a linear support, Frechét subdifferential means a smooth support, and a proximal subdifferential means a local quadratic support. The detailed discussion on other subdifferentials and their properties (particularly on calculus rules and the robust property) can be found in [34].

Clearly, the definition of an augmented Lagrangian multiplier given in (2) indicates that the augmented perturbation function is supported by a linear function at the origin. It corresponds to the subdifferential in the convex analysis. However, for a nonconvex setting, it is natural to consider whether a nonlinear support is available. Once it is done, we can establish and apply the duality theory in a more flexible environment. Define $\omega : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ such that $\omega(\eta) \rightarrow +\infty$ as $\eta \rightarrow +\infty$.

Definition 1. A vector (λ, μ) is said to be a generalized augmented Lagrange multiplier of (P) , if there exists $r \geq 0$ such that

$$v_r(y, z) \geq v_r(0, 0) + \phi_1(\lambda, y) + \phi_2(\mu, z), \quad \forall (y, z) \in \mathbb{R}^m \times \mathbb{R}^l, \tag{4}$$

where ϕ_i for $i = 1, 2$ possesses the following properties:

- (A₁) ϕ_i is continuous and $\phi_i(\cdot, 0) = 0$;
- (A₂) $\phi_i(x, y + z) \leq \phi_i(x, y) + \phi_i(x, z)$;
- (A₃) $\forall x \notin \mathcal{F}$, there exist a nonzero vector (u_0, v_0) and $\gamma < 0$ such that

$$\phi_1(\eta u_0, y) + \phi_2(\eta v_0, z) \leq \omega(\eta)\gamma,$$

whenever (y, z) satisfies $y + g(x) \leq 0, z + h(x) = 0$, and $\eta > 0$ is sufficiently large.

Since ϕ_i includes the inner product as special cases, (4) is an essential extension of (2) from linear support to nonlinear support.

As mentioned above, the augmented Lagrange multiplier is a subgradient (in the sense of convex analysis) of an augmented perturbation function at the origin. That means the augmented perturbation function has a linear support. The augmented Lagrange multiplier is extended in this paper to a new concept called the generalized augmented Lagrangian multiplier, in which a nonlinear support is allowed. The main aim of this paper is to study the existence of generalized augmented Lagrange multipliers. It helps us to better understand properties of an augmented perturbation function at the origin. Based on this nonlinear support, we need to re-investigate the corresponding duality theory, particularly by discussing the relations among generalized augmented Lagrange multipliers, saddle points, and the zero duality gap property. The existence of generalized augmented Lagrange multipliers is established by perturbation analysis of the primal problem.

We organize our paper as follows. Section 2 introduces the preliminaries. In Sections 3, we present the duality theory based on generalized augmented Lagrangians. Section 4 discusses the existence of generalized augmented Lagrange multipliers by perturbation analysis.

2. Preliminaries

In this section we clarify the notation, recall some background materials we need from duality theory, and develop some preliminary results.

Recall that

$$v_r(y, z) := v(y, z) + r\sigma(y, z), \quad \forall (y, z) \in \mathbb{R}^m \times \mathbb{R}^l.$$

where $\sigma : \mathbb{R}^{m+l} \rightarrow \mathbb{R}_+ := [0, +\infty)$ satisfies the following valley-at-zero property:

- (i) σ is continuous at 0 with $\sigma(0, 0) = 0$;
- (ii) $\inf\{\sigma(y, z) \mid \|(y, z)\| \geq \eta, y \in \mathbb{R}^m, z \in \mathbb{R}^l\} > 0$ for all $\eta > 0$.

The definition of the growth condition defined below was introduced in [23], as an extension of the one given in [3], where the augmenting function is restricted to be a quadratic function.

Definition 2. A function $v(y, z)$ is said to satisfy the growth condition with σ , if for any $\tau > 0$, there exist $a, c \in \mathbb{R}$ such that

$$v(y, z) \geq c - a\sigma(y, z), \quad \forall (y, z) \in \mathbb{R}^{m+l} \setminus \tau\mathbb{B}_{\mathbb{R}^{m+l}},$$

where $\mathbb{B}_{\mathbb{R}^{m+l}}$ denotes the closed unit ball in \mathbb{R}^{m+l} .

The dualizing parametrization function of the primal problem is defined as

$$F(x, y, z) := \begin{cases} f(x), & \text{if } x \in \Omega \text{ and } y + g(x) \leq 0, z + h(x) = 0, \\ +\infty, & \text{otherwise.} \end{cases} \tag{5}$$

For $(x, \lambda, \mu) \in \mathbb{R}^n \times \mathbb{R}_+^m \times \mathbb{R}^l$, the corresponding *generalized augmented Lagrangian* is

$$L(x, \lambda, \mu, r) := \inf \{ F(x, y, z) - \phi_1(\lambda, y) - \phi_2(\mu, z) + r\sigma(y, z) \mid (y, z) \in \mathbb{R}^{m+l} \}. \tag{6}$$

The generalized Lagrangian function is defined as

$$L_0(x, \lambda, \mu) := f(x) - \phi_1(\lambda, -g(x)) - \phi_2(\mu, -h(x)),$$

which reduces to the *classical Lagrangian* of (P) when $\phi_1(\lambda, y) = \langle \lambda, y \rangle$ and $\phi_2(\mu, z) = \langle \mu, z \rangle$.

If in particular $x \in \Omega$, the generalized augmented Lagrangian can be rewritten as

$$\begin{aligned} L(x, \lambda, \mu, r) &= \inf_{\substack{y+g(x) \leq 0 \\ z+h(x)=0}} \{ f(x) - \phi_1(\lambda, y) - \phi_2(\mu, z) + r\sigma(y, z) \} \\ &\geq \inf_{\xi_1 \leq 0, \xi_2=0} \{ f(x) - \phi_1(\lambda, -g(x)) - \phi_2(\mu, -h(x)) - \phi_1(\lambda, \xi_1) \\ &\quad - \phi_2(\mu, \xi_2) + r\sigma(\xi_1 - g(x), \xi_2 - h(x)) \} \\ &= \inf_{\xi_1 \leq 0, \xi_2=0} \{ L_0(x, \lambda, \mu) - \phi_1(\lambda, \xi_1) + r\sigma(\xi_1 - g(x), \xi_2 - h(x)) \}, \end{aligned} \tag{7}$$

where the inequality comes from (A_2) .

Definition 3. A solution $(x^*, \lambda^*, \mu^*) \in \Omega \times \mathbb{R}_+^m \times \mathbb{R}^l$ is said to be a *global saddle point* of the generalized augmented Lagrangian L for $r \geq 0$, if

$$L(x^*, \lambda, \mu, r) \leq L(x^*, \lambda^*, \mu^*, r) \leq L(x, \lambda^*, \mu^*, r), \quad \forall x \in \Omega, (\lambda, \mu) \in \mathbb{R}_+^m \times \mathbb{R}^l. \tag{9}$$

If the above inequalities hold for all $x \in \mathbb{B}_{\mathbb{R}^n}(x^*, \delta) \cap \Omega$, where $\mathbb{B}_{\mathbb{R}^n}(x^*, \delta)$ denotes the ball with center x^* and radius $\delta > 0$, then (x^*, λ^*, μ^*) is said to be a *local saddle point* of L .

The generalized augmented Lagrangian dual problem of (P) is defined as

$$(D) \quad \sup_{(\lambda, \mu, r) \in \mathbb{R}_+^m \times \mathbb{R}^l \times \mathbb{R}_+} \theta(\lambda, \mu, r),$$

where $\theta(\lambda, \mu, r)$ is the generalized augmented Lagrangian dual function given as

$$\theta(\lambda, \mu, r) := \inf \{ L(x, \lambda, \mu, r) \mid x \in \Omega \}. \tag{10}$$

Taking into account of (7) and (10), we have

$$\begin{aligned} &\inf_{(y,z) \in \mathbb{R}^m \times \mathbb{R}^l} \{ v_r(y, z) - \phi_1(\lambda, y) - \phi_2(\mu, z) \} \\ &= \inf_{(y,z) \in \mathbb{R}^m \times \mathbb{R}^l} \inf_{x \in \Omega} \{ f(x) - \phi_1(\lambda, y) - \phi_2(\mu, z) + r\sigma(y, z) \mid g(x) + y \leq 0, h(x) + z = 0 \} \\ &= \inf_{x \in \Omega} \inf_{y+g(x) \leq 0, z+h(x)=0} \{ f(x) - \phi_1(\lambda, y) - \phi_2(\mu, z) + r\sigma(y, z) \} \\ &= \inf_{x \in \Omega} L(x, \lambda, \mu, r) \\ &= \theta(\lambda, \mu, r). \end{aligned} \tag{11}$$

In addition, it also follows from (5) that

$$v(y, z) = \inf_{\substack{x \in \Omega \\ g(x)+y \leq 0, h(x)+z=0}} f(x) = \inf_{x \in \Omega} F(x, y, z).$$

It is well known that a zero duality gap between the problem (P) and its generalized augmented Lagrangian dual problem (D) holds if

$$\text{val}(P) = \sup_{(\lambda, \mu, r) \in \mathbb{R}_+^m \times \mathbb{R}^l \times \mathbb{R}_+} \theta(\lambda, \mu, r).$$

For $r \geq 0$, consider the following r -dual problem of (P), denoted by (D_r) ,

$$(D_r) \quad \sup_{(\lambda, \mu) \in \mathbb{R}_+^m \times \mathbb{R}^l} \theta(\lambda, \mu, r) = \sup_{(\lambda, \mu) \in \mathbb{R}_+^m \times \mathbb{R}^l} \inf_{x \in \Omega} L(x, \lambda, \mu, r).$$

Similarly, if for some fixed $r \geq 0$ such that

$$\text{val}(P) = \sup_{(\lambda, \mu) \in \mathbb{R}_+^m \times \mathbb{R}^l} \theta(\lambda, \mu, r),$$

then the zero duality gap property holds for the pair of problems (P) and (D_r) .

Define the optimal values of problems (D) and (D_r) by $\text{val}(D)$ and $\text{val}(D_r)$, respectively. It is clear that $\text{val}(D) = \sup_{r \in \mathbb{R}_+} \text{val}(D_r)$.

3. Duality Theory Based on Generalized Augmented Lagrangian Functions

In this section, we study the relationships among generalized augmented Lagrange multipliers, global saddle points, and the zero duality gap property between the primal problem and its generalized augmented Lagrangian dual problem. The related conclusions are given in Theorem 3 and Theorem 4.

Firstly, the *weak duality theorem* is given below, which shows that the dual problem provides a lower bound for (P).

Proposition 1. *Let x be a feasible point of (P) and $(\lambda, \mu, r) \in \mathbb{R}_+^m \times \mathbb{R}^l \times \mathbb{R}_+$. Then*

$$\theta(\lambda, \mu, r) \leq \text{val}(P) \leq f(x).$$

Proof. Since x is feasible, i.e., $x \in \Omega$ and $g(x) \leq 0, h(x) = 0$, then $-g(x) \geq 0, -h(x) = 0$. So

$$L(x, \lambda, \mu, r) = \inf_{y \leq -g(x), z = -h(x)} \{f(x) - \phi_1(\lambda, y) - \phi_2(\mu, z) + r\sigma(y, z)\} \leq f(x), \tag{12}$$

where the inequality follows by letting $y = 0, z = 0$ and $\phi(\cdot, 0) = 0$. Hence

$$\theta(\lambda, \mu, r) = \inf_{x \in \Omega} L(x, \lambda, \mu, r) \leq f(x).$$

The arbitrariness of x ensures

$$\theta(\lambda, \mu, r) \leq \inf_{\substack{x \in \Omega \\ g(x) \leq 0, h(x) = 0}} f(x) = \text{val}(P). \tag{13}$$

□

Theorem 1. *Let $\sigma : \mathbb{R}^{m+l} \rightarrow \mathbb{R}_+$ and $(\lambda^*, \mu^*, r^*) \in \mathbb{R}_+^m \times \mathbb{R}^l \times \mathbb{R}_+$. Then (λ^*, μ^*) is a generalized augmented Lagrange multiplier of (P) with r^* if and only if (λ^*, μ^*, r^*) is an optimal solution of (D) and the zero duality gap property holds for problems (P) and (D).*

Proof. (Necessity). If (λ^*, μ^*) is a generalized augmented Lagrange multiplier of (P) with r^* , then

$$v_{r^*}(0, 0) = \inf_{(y, z) \in \mathbb{R}^m \times \mathbb{R}^l} \{v_{r^*}(y, z) - \phi_1(\lambda^*, y) - \phi_2(\mu^*, z)\},$$

where the above equation is due to Definition 1. According to (11), we have

$$val(P) = v_{r^*}(0,0) = \theta(\lambda^*, \mu^*, r^*).$$

This implies

$$val(P) = \theta(\lambda^*, \mu^*, r^*) \leq \sup_{(\lambda, \mu) \in \mathbb{R}_+^m \times \mathbb{R}^l} \theta(\lambda, \mu, r^*) \leq val(D) = \sup_{(\lambda, \mu, r) \in \mathbb{R}_+^m \times \mathbb{R}^l \times \mathbb{R}_+} \theta(\lambda, \mu, r) \leq val(P),$$

where the third inequality is due to (13). Hence, (λ^*, μ^*, r^*) is an optimal solution of (D) and

$$\sup_{(\lambda, \mu, r) \in \mathbb{R}_+^m \times \mathbb{R}^l \times \mathbb{R}_+} \theta(\lambda, \mu, r) = val(P).$$

(Sufficiency). Suppose (λ^*, μ^*, r^*) is an optimal solution of (D) and the zero duality gap property between (P) and (D) holds. Then

$$val(P) = \theta(\lambda^*, \mu^*, r^*) \leq val(D_{r^*}) \leq val(D) \leq val(P).$$

Hence

$$\theta(\lambda^*, \mu^*, r^*) = val(D_{r^*}) = val(P) = v(0,0) = v_{r^*}(0,0),$$

which together with (11) implies

$$v_{r^*}(0,0) = \theta(\lambda^*, \mu^*, r^*) = \inf_{(y,z) \in \mathbb{R}^m \times \mathbb{R}^l} \{v_{r^*}(y,z) - \phi_1(\lambda^*, y) - \phi_2(\mu^*, z)\}.$$

Therefore, (λ^*, μ^*) is a generalized augmented Lagrange multiplier of (P) with r^* . \square

From the proof of Theorem 1, we can see that (λ^*, μ^*) is an optimal solution of (D_{r^*}) and the zero duality gap property holds between (P) and (D_{r^*}) . It should be emphasized that the existence of generalized augmented Lagrange multipliers does not require that the primal problem (P) must be solvable. Indeed, in general, the optimal solution of a primal problem cannot be known in advance. The relation between the zero duality gap property and global saddle points is given below.

Theorem 2. Let $\sigma : \mathbb{R}^{m+l} \rightarrow \mathbb{R}_+$ and $r^* \geq 0$. Then (x^*, λ^*, μ^*) is a global saddle point of $L(x, \lambda, \mu, r^*)$ if and only if $val(P) = val(D_{r^*})$, and $x^* \in \Omega$, $(\lambda^*, \mu^*) \in \mathbb{R}_+^m \times \mathbb{R}^l$ are optimal solutions of (P) and (D_{r^*}) , respectively.

Proof. We first claim that

$$\sup_{(\lambda, \mu) \in \mathbb{R}_+^m \times \mathbb{R}^l} L(x, \lambda, \mu, r^*) = \begin{cases} f(x), & x \in \Omega \text{ and } g(x) \leq 0, h(x) = 0, \\ +\infty, & \text{otherwise.} \end{cases} \tag{14}$$

Consider the following two cases:

Case 1. x is infeasible. Then either $x \notin \Omega$ or $x \in \Omega$ while $x \notin \mathcal{F}$. If $x \notin \Omega$, from (5) and (6) we get

$$L(x, \lambda, \mu, r^*) = +\infty, \quad \forall (\lambda, \mu) \in \mathbb{R}_+^m \times \mathbb{R}^l. \tag{15}$$

If $x \in \Omega$, but $x \notin \mathcal{F}$, it follows from the property (A_3) that there exist nonzero $(\lambda_0, \mu_0) \in \mathbb{R}_+^m \times \mathbb{R}^l$ and $\gamma < 0$ such that

$$\phi_1(\eta \lambda_0, y) + \phi_2(\eta \mu_0, z) \leq \omega(\eta) \gamma, \tag{16}$$

whenever (y, z) satisfies $y + g(x) \leq 0, z + h(x) = 0$, and $\eta > 0$ sufficiently large. Hence

$$\begin{aligned} L(x, \eta\lambda_0, \eta\mu, r^*) &= \inf_{y+g(x)\leq 0, z+h(x)=0} \{f(x) - \phi_1(\eta\lambda_0, y) - \phi_2(\eta\mu, z) + r^*\sigma(y, z)\} \\ &\geq \inf_{y+g(x)\leq 0, z+h(x)=0} \{f(x) - \phi_1(\eta\lambda_0, y) - \phi_2(\eta\mu_0, z)\} \\ &\geq f(x) - \omega(\eta)\gamma, \end{aligned}$$

where the first inequality comes from the nonnegativity of σ , and the second inequality is due to (16). This together with $\gamma < 0$ further implies that

$$\sup_{(\lambda, \mu) \in \mathbb{R}_+^m \times \mathbb{R}^l} L(x, \lambda, \mu, r^*) \geq L(x, \eta\lambda_0, \eta\mu_0, r^*) \geq f(x) - \omega(\eta)\gamma \rightarrow +\infty, \text{ as } \eta \rightarrow +\infty;$$

i.e.,

$$\sup_{(\lambda, \mu) \in \mathbb{R}_+^m \times \mathbb{R}^l} L(x, \lambda, \mu, r^*) = +\infty. \tag{17}$$

Therefore, either $x \notin \Omega$ or $x \in \Omega, x \notin \mathcal{F}$, so it follows from (15) and (17) that

$$\sup_{(\lambda, \mu) \in \mathbb{R}_+^m \times \mathbb{R}^l} L(x, \lambda, \mu, r^*) = +\infty. \tag{18}$$

Case 2. x is feasible i.e., $x \in \Omega$ and $g(x) \leq 0, h(x) = 0$. In this case, it follows from (12) that for any $(\lambda, \mu) \in \mathbb{R}_+^m \times \mathbb{R}^l$,

$$L(x, \lambda, \mu, r^*) = \inf_{y \leq -g(x), z = -h(x)} \{f(x) - \phi_1(\lambda, y) - \phi_2(\mu, z) + r^*\sigma(y, z)\} \leq f(x). \tag{19}$$

According to the nonnegativity of σ , we also have

$$\sup_{(\lambda, \mu) \in \mathbb{R}_+^m \times \mathbb{R}^l} L(x, \lambda, \mu, r^*) \geq L(x, 0, 0, r^*) \geq f(x),$$

which together with (19) means that

$$\sup_{(\lambda, \mu) \in \mathbb{R}_+^m \times \mathbb{R}^l} L(x, \lambda, \mu, r^*) = f(x). \tag{20}$$

Putting (18) and (20) together yields the desired formula (14). Hence

$$\text{val}(P) = \inf_{x \in \Omega} \sup_{(\lambda, \mu) \in \mathbb{R}_+^m \times \mathbb{R}^l} L(x, \lambda, \mu, r^*).$$

On the other hand, note that the dual problem can be rewritten as

$$\text{val}(P) = \sup_{(\lambda, \mu) \in \mathbb{R}_+^m \times \mathbb{R}^l} \inf_{x \in \Omega} L(x, \lambda, \mu, r^*).$$

The desired result follows by applying the minimax relations theorem (Theorem 11.50 [34]). \square

Indeed, Theorem 2 shows that $\text{val}(P) = \text{val}(D)$, and $x^*, (\lambda^*, \mu^*, r^*)$ are optimal solutions of (P) and (D) respectively, provided that $\text{val}(D) = \text{val}(D_r)$, i.e., $\text{val}(D) = \sup_{r \in \mathbb{R}_+} \text{val}(D_r)$ by Proposition 1, and the maximum can be attained at some r . The converse statement obviously holds true. As just mentioned above, compared with the existence of augmented Lagrange multipliers, global saddle points require that the primal problem is solvable.

Theorem 3. Suppose that $\sigma : \mathbb{R}^{m+l} \rightarrow \mathbb{R}_+$ has a valley at zero, v satisfies the growth condition with σ , and

$$\liminf_{(y,z) \rightarrow (0,0)} v(y,z) < +\infty. \tag{21}$$

The following statements hold:

(i)

$$\sup_{(0,0,r) \in \mathbb{R}_+^m \times \mathbb{R}^l \times \mathbb{R}_+} \theta(0,0,r) = \liminf_{(y,z) \rightarrow (0,0)} v(y,z);$$

(ii) v is lower semi-continuous at the origin if and only if the zero duality gap property holds for problems (P) and (D).

Proof. (i). First, according to the condition (21) we show that

$$\sup_{(\lambda,\mu,r) \in \mathbb{R}_+^m \times \mathbb{R}^l \times \mathbb{R}_+} \theta(\lambda,\mu,r) = \liminf_{(y,z) \rightarrow (0,0)} v(y,z). \tag{22}$$

Assume that $\{(y^{(s)}, z^{(s)})\}$ is the sequence such that the liminf in (21) is attained; i.e.,

$$\lim_{s \rightarrow \infty} (y^{(s)}, z^{(s)}) = 0, \lim_{s \rightarrow \infty} v((y^{(s)}, z^{(s)})) = \liminf_{s \rightarrow 0} v(y,z). \tag{23}$$

Consider the following two cases:

Case 1. $\liminf_{(y,z) \rightarrow (0,0)} v(y,z) = -\infty$. For $(\lambda, \mu, r) \in \mathbb{R}_+^m \times \mathbb{R}^l \times \mathbb{R}_+$, it follows from (11) that

$$\begin{aligned} \theta(\lambda, \mu, r) &= \inf_{(y,z) \in \mathbb{R}^{m+l}} \{v_r(y,z) - \phi_1(\lambda, y) - \phi_2(\mu, z)\} \\ &\leq v(y^{(s)}, z^{(s)}) + r\sigma(y^{(s)}, z^{(s)}) - \phi_1(\lambda, y^{(s)}) - \phi_2(\mu, z^{(s)}), \end{aligned} \tag{24}$$

where the inequality comes from (1). Passing to limit (24), together with (23), we get

$$\theta(\lambda, \mu, r) \leq \lim_{s \rightarrow \infty} \{v(y^{(s)}, z^{(s)}) + r\sigma(y^{(s)}, z^{(s)}) - \phi_1(\lambda, y^{(s)}) - \phi_2(\mu, z^{(s)})\} = \liminf_{(y,z) \rightarrow (0,0)} v(y,z) = -\infty,$$

where the first equality comes from the continuity of σ and ϕ by (A1). Hence

$$\sup_{(\lambda,\mu,r) \in \mathbb{R}_+^m \times \mathbb{R}^l \times \mathbb{R}_+} \theta(\lambda,\mu,r) = \liminf_{(y,z) \rightarrow (0,0)} v(y,z).$$

Case 2. $\liminf_{(y,z) \rightarrow (0,0)} v(y,z) > -\infty$. Noting that $\theta(\lambda, \mu, r) \leq \liminf_{(y,z) \rightarrow (0,0)} v(y,z)$, then

$$\sup_{(\lambda,\mu,r) \in \mathbb{R}_+^m \times \mathbb{R}^l \times \mathbb{R}_+} \theta(\lambda,\mu,r) \leq \liminf_{(y,z) \rightarrow (0,0)} v(y,z). \tag{25}$$

Conversely, take k satisfying $\liminf_{(y,z) \rightarrow (0,0)} v(y,z) > k$. Then there exists $\tau > 0$ such that

$$v(y,z) + r\sigma(y,z) \geq v(y,z) \geq k, \quad \forall (y,z) \in \tau\mathbb{B}_{\mathbb{R}^{m+l}}, r \geq 0, \tag{26}$$

where the first inequality follows from the nonnegativity of σ . Since σ has a valley at zero, there exists $\varepsilon > 0$ such that

$$\sigma(y,z) \geq \varepsilon, \quad \forall (y,z) \in \mathbb{R}^{m+l} \setminus \tau\mathbb{B}_{\mathbb{R}^{m+l}}. \tag{27}$$

Using the growth condition of v with σ , for the above $\tau > 0$ there exist $a, c \in \mathbb{R}$ such that

$$v(y,z) \geq c - a\sigma(y,z), \quad \forall \tau > 0, (y,z) \in \mathbb{R}^{m+l} \setminus \tau\mathbb{B}_{\mathbb{R}^{m+l}}.$$

This together with (27) yields

$$v(y, z) + r\sigma(y, z) \geq c - a\sigma(y, z) + r\sigma(y, z) \geq c + (r - a)\varepsilon \geq k, \quad \forall (y, z) \in \mathbb{R}^{m+l} \setminus \tau\mathbb{B}_{\mathbb{R}^{m+l}}, r \geq r_0, \quad (28)$$

where $r_0 := \max\{a + (k - c)/\varepsilon, 0\} + 1$. From (26) and (28) we get

$$\theta(0, 0, r) = \inf_{(y,z) \in \mathbb{R}^{m+l}} \{v(y, z) + r\sigma(y, z)\} \geq k, \quad \forall r \geq r_0,$$

and

$$\sup_{r \in \mathbb{R}_+} \theta(0, 0, r) \geq k.$$

Since $k < \liminf_{(y,z) \rightarrow (0,0)} v(y, z)$ is arbitrary, then

$$\sup_{(\lambda, \mu, r) \in \mathbb{R}_+^m \times \mathbb{R}^l \times \mathbb{R}_+} \theta(\lambda, \mu, r) \geq \liminf_{(y,z) \rightarrow (0,0)} v(y, z).$$

Taking into account (25), we get in the last inequality

$$\sup_{(\lambda, \mu, r) \in \mathbb{R}_+^m \times \mathbb{R}^l \times \mathbb{R}_+} \theta(\lambda, \mu, r) = \liminf_{(y,z) \rightarrow (0,0)} v(y, z).$$

The statement (i) follows by letting $(\lambda, \mu) = (0, 0)$ in (24) and (25) and by a similar argumentation as above.

(ii). If v is lower semi-continuous at origin, then

$$\liminf_{(y,z) \rightarrow (0,0)} v(y, z) \geq v(0, 0) = \text{val}(P),$$

which together with (22) yields

$$\text{val}(P) \leq \liminf_{(y,z) \rightarrow (0,0)} v(y, z) = \sup_{(\lambda, \mu, r) \in \mathbb{R}_+^m \times \mathbb{R}^l \times \mathbb{R}_+} \theta(\lambda, \mu, r) \leq \text{val}(P).$$

Therefore, the zero duality gap property holds for (P) and (D) .

Conversely, according to (22), it is easy to see that the lower semi-continuity of v at the origin can be obtained if the zero duality gap property holds for problems (P) and (D) . \square

Corollary 1. Suppose that $\sigma : \mathbb{R}^{m+l} \rightarrow \mathbb{R}_+$ has a valley at zero, v satisfies the growth condition with σ , and

$$\liminf_{(y,z) \rightarrow (0,0)} v(y, z) < +\infty.$$

If v is lower semi-continuous at origin and $r^* \in \arg \sup_{r>0} \theta(0, 0, r)$, then the following statements hold:

- (i) $(0, 0)$ is a generalized augmented Lagrange multiplier;
- (ii) If the primal problem (P) has the optimal solution x^* , then $(x^*, 0, 0)$ and $(x^*, 0, 0, r^*)$ are saddle points of D_{r^*} and D , respectively.

Proof. The results follow immediately from Theorem 3. \square

Theorem 3 shows that the zero duality gap property is closely related with the lower semi-continuity of the perturbation function. In the definition of generalized augmented Lagrange multipliers, the inequality involved in (4) is required to be satisfied for all $(y, z) \in \mathbb{R}^{m+l}$, but Theorem 4 shows that this restriction can be weakened by just checking all (y, z) in some neighborhood of the

origin once some additional assumptions are imposed on augmented functions. In the following, we further require the ϕ satisfying the following property:

(A₄) For any $(\lambda, \mu) \in \mathbb{R}_+^m \times \mathbb{R}^l$, there exist $\rho > 0, \tau > 0$ such that

$$\rho\sigma(y, z) - \phi_1(\lambda, y) - \phi_2(\mu, z) \geq 0, \quad \forall (y, z) \in \mathbb{R}^{m+l} \setminus \tau\mathbb{B}_{\mathbb{R}^{m+l}}.$$

Theorem 4. Suppose that σ has a valley at zero and v satisfies the growth condition with σ . Then (P) has a generalized augmented Lagrange multiplier $(\lambda^*, \mu^*) \in \mathbb{R}_+^m \times \mathbb{R}^l$ if and only if there exists $r^* \in \mathbb{R}_+$ such that

$$v_{r^*}(y, z) \geq v_{r^*}(0, 0) + \phi_1(\lambda^*, y) + \phi_2(\mu^*, z), \quad \forall (y, z) \in \tau\mathbb{B}_{\mathbb{R}^{m+l}}. \tag{29}$$

Proof. (Necessity). The necessity is clear by the definition of generalized augmented Lagrange multiplier.

(Sufficiency). Since v satisfies the growth condition with σ , then for any $\tau > 0$ there exist $a, c \in \mathbb{R}$ such that

$$v(y, z) \geq c - a\sigma(y, z), \quad \forall (y, z) \in \mathbb{R}^{m+l} \setminus \tau\mathbb{B}_{\mathbb{R}^{m+l}}. \tag{30}$$

Since σ has a valley at zero, there exists $d > 0$ such that

$$\sigma(y, z) \geq d, \quad \forall (y, z) \in \mathbb{R}^{m+l} \setminus \tau\mathbb{B}_{\mathbb{R}^{m+l}}. \tag{31}$$

Combining the property (A₄) with (31) means that for any $(y, z) \in \mathbb{R}^{m+l} \setminus \tau\mathbb{B}_{\mathbb{R}^{m+l}}$ we have

$$\begin{aligned} &v_r(y, z) - v_r(0, 0) - \phi_1(\lambda^*, y) - \phi_2(\mu^*, z) \\ &\geq c - a\sigma(y, z) - v_r(0, 0) + r\sigma(y, z) - \phi_1(\lambda^*, y) - \phi_2(\mu^*, z) \\ &= c - \text{val}(P) + (r - a - \rho)\sigma(y, z) + [\rho\sigma(y, z) - \phi_1(\lambda^*, y) - \phi_2(\mu^*, z)] \\ &\geq c - \text{val}(P) + (r - a - \rho)\sigma(y, z) \\ &\geq c - \text{val}(P) + (r - a - \rho)d. \end{aligned}$$

Pick $r > \max \{r^*, a + \rho, a + \rho - [(c - \text{val}(P))/d]\}$. Then

$$v_r(y, z) - v_r(0, 0) - \phi_1(\lambda^*, y) - \phi_2(\mu^*, z) \geq 0, \quad \forall (y, z) \in \mathbb{R}^{m+l} \setminus \tau\mathbb{B}_{\mathbb{R}^{m+l}}. \tag{32}$$

It follows from (29) and (32) that

$$v_r(y, z) \geq v_r(0, 0) + \phi_1(\lambda^*, y) + \phi_2(\mu^*, z), \quad \forall (y, z) \in \mathbb{R}^{m+l}.$$

Hence (λ^*, μ^*) is a generalized augmented Lagrange multiplier of (P). \square

Here we list two classes of nonlinear functions satisfying the above assumptions (A₁)–(A₄).

(1) Let $\theta : \mathbb{R} \rightarrow \mathbb{R}$ be sublinear, continuous, and increasing with $\theta(0) = 0$.

Let

$$\phi(x, y) := \|x\|\theta(x^T y).$$

(1-1) $\phi(x, 0) = \|x\|\theta(0) = 0, \forall x$.

(1-2) For any x, y, z ,

$$\begin{aligned} \phi(x, y + z) &= \|x\|\theta(x^T(y + z)) = \|x\|\theta(x^T y + x^T z) \\ &\leq \|x\|\theta(x^T y) + \|x\|\theta(x^T z) = \phi(x, y) + \phi(x, z). \end{aligned}$$

(1-3) For any $x \notin \mathcal{F}$, then $(g(x), h(x)) \notin \mathbb{R}^m \times \{0\}^l$, i.e., $(0, 0) \notin \mathbb{R}^m \times \{0\}^l - (g(x), h(x))$. If $0 \notin \mathbb{R}^m - g(x)$, then according to convex set sperate theorem, there exist a nonzero vector u_0 and $\xi < 0$ such that $u_0^T y < \xi$ whenever $y + g(x) \leq 0$. Hence taking $v_0 := 0$ and $\gamma := \|u_0\|\theta(\xi)$, we have

$$\begin{aligned} \phi_1(\eta u_0, y) + \phi_2(\eta v_0, z) &= \phi_1(\eta u_0, y) = \eta \|u_0\| \theta(\eta u_0^T y) \\ &\leq \eta \|u_0\| \theta(\eta \xi) \leq \eta \|u_0\| \theta(\xi) = \omega(\eta) \gamma, \end{aligned}$$

where $\omega(\eta) := \eta$. Similarly, if $0 \notin \{0\}^l - h(x)$, there exists a nonzero vector v_0 and $\xi < 0$ such that $v_0^T z \leq \xi$ for $z = -h(x)$. Hence taking $u_0 := 0$ and $\gamma := \|v_0\|\theta(\xi)$, we have

$$\phi_1(\eta u_0, y) + \phi_2(\eta v_0, z) = \phi_2(\eta v_0, z) = \eta \|v_0\| \theta(\eta v_0^T z) \leq \eta \|v_0\| \theta(\xi) = \omega(\eta) \gamma.$$

(1-4) Let $\beta : \mathbb{R}_+ \rightarrow \mathbb{R}$ satisfy $\beta(t) > 0$ as $t > 0$. Assume that there exist $\alpha > 0, \tau > 0$ such that for all $t > \alpha$ and u with $\|u\| \geq \tau$, we have $\theta(t\|u\|) \leq \beta(t)\sigma(u)$.

For any (λ, μ) , letting $\varrho := \|\lambda\| + \|\mu\| + \alpha$ we have

$$\begin{aligned} \phi_1(\lambda, y) + \phi_2(\mu, z) &= \|\lambda\| \theta(\lambda^T y) + \|\mu\| \theta(\mu^T z) \leq \|\lambda\| \theta(\|\lambda\| \|y\|) + \|\mu\| \theta(\|\mu\| \|z\|) \\ &\leq \varrho \theta(\varrho \|y\|) + \varrho \theta(\varrho \|z\|) \leq 2\varrho \theta(\varrho \|(y, z)\|). \end{aligned}$$

As $\|(y, z)\| \geq \tau$, we have $2\varrho \theta(\varrho \|(y, z)\|) \leq 2\varrho \beta(\varrho) \sigma(y, z)$. Hence for all $\|(y, z)\| \geq \tau$ and $\rho := 2\varrho \beta(\varrho)$,

$$\rho \sigma(y, z) - \phi_1(\lambda, y) - \phi_2(\mu, z) \geq 0.$$

In particular, we can take θ as a piecewise linear function or a support function over a bounded closed interval, $w(t) := t, \beta(t) := t^2$, and $\sigma(u) := \|u\|$.

(2) Let $\theta : \mathbb{R}_+ \rightarrow \mathbb{R}$ satisfy $\theta(t) > 0$ if $t \neq 0$ and $\theta(t) \geq t^q$ as $t > 0$ are sufficiently large, where q is positive integer.

Define

$$\phi(x, y) := \theta(\|x\|)(x^T A y),$$

where A is a symmetric and invertible matrix.

(2-1) $\phi(x, 0) = 0, \forall x$.

(2-2) For any x, y, z ,

$$\phi(x, y + z) = \theta(\|x\|)(x^T A(y + z)) = \theta(\|x\|)(x^T A y + x^T A z) = \phi(x, y) + \phi(x, z).$$

(2-3) Similar to the argument given in (1-3), if $0 \notin \mathbb{R}^m - g(x)$, then according to convex set sperate theorem, there exists a nonzero vector \tilde{u}_0 and $\xi < 0$ such that $\tilde{u}_0^T y < \xi$ whenever $y + g(x) \leq 0$. Hence taking $u_0 := A^{-1}\tilde{u}_0, v_0 := 0$, and $\gamma := \theta(\|u_0\|)\xi$, we have

$$\begin{aligned} \phi_1(\eta u_0, y) + \phi_2(\eta v_0, z) &= \phi_1(\eta u_0, y) = \theta(\eta \|u_0\|) \eta u_0^T A y \\ &\leq \eta^{q+1} \|u_0\|^q \xi \leq \eta^q \theta(\|u_0\|) \xi = \omega(\eta) \gamma, \end{aligned}$$

whenever $\eta \geq \theta(\|u_0\|) / \|u_0\|^q$ and $\omega(\eta) := \eta^q$.

If $0 \notin \{0\}^l - h(x)$, there exists a nonzero vector \tilde{v}_0 and $\xi < 0$ such that $\tilde{v}_0^T z \leq \xi$ for $z = -h(x)$. Hence taking $u_0 := 0, v_0 := A^{-1}\tilde{v}_0$, and $\gamma := \theta(\|v_0\|)\xi$, we have

$$\begin{aligned} \phi_1(\eta u_0, y) + \phi_2(\eta v_0, z) &= \phi_2(\eta v_0, z) = \theta(\eta \|v_0\|) \eta v_0^T A z \\ &\leq \eta^{q+1} \|v_0\|^q \xi \leq \eta^q \theta(\|v_0\|) \xi \leq \omega(\eta) \gamma, \end{aligned}$$

whenever $\eta \geq \theta(\|v_0\|)/\|v_0\|^q$.

(2-4) Assume that there exists $\tau > 0$ such that $\sigma(u) \geq \|u\|$ as u with $\|u\| \geq \tau$. For any (λ, μ) , we have

$$\begin{aligned} \phi_1(\lambda, y) + \phi_2(\mu, z) &= \theta(\|\lambda\|)\lambda^T Ay + \theta(\|\mu\|)\mu^T Az \\ &\leq \theta(\|\lambda\|)(\|A\lambda\|\|y\|) + \theta(\|\mu\|)(\|A\mu\|\|z\|) \\ &\leq (\theta(\|\lambda\|)\|A\lambda\| + \theta(\|\mu\|)\|A\mu\|)\|(y, z)\| \\ &\leq (\theta(\|\lambda\|)\|A\lambda\| + \theta(\|\mu\|)\|A\mu\|)\sigma(y, z). \end{aligned}$$

Let $\rho := (\theta(\|\lambda\|)\|A\lambda\| + \theta(\|\mu\|)\|A\mu\|)$. Then for any (y, z) with $\|(y, z)\| \geq \tau$,

$$\rho\sigma(y, z) - \phi_1(\lambda, y) - \phi_2(\mu, z) \geq 0.$$

In particular, we can take $q := 2, \omega(t) := t^2, \theta(t) := t^3$, and $\sigma(u) := \|u\|^2$.

4. Existence of Generalized Augmented Lagrange Multipliers

In this section, we develop some sufficient conditions for the existence of generalized augmented Lagrange multipliers. Given $\varepsilon \geq 0$, define

$$W_1(\varepsilon) := \{x \in \Omega \mid \text{dist}(g(x), h(x); \mathbb{R}_-^m \times \{0\}^l) \leq \varepsilon\},$$

and

$$W_2(\varepsilon) := \{x \in \Omega \mid f(x) - v(0, 0) \leq \varepsilon\}.$$

Lemma 1. Suppose that $\sigma : \mathbb{R}^{m+l} \rightarrow \mathbb{R}_+$ has a valley at zero and

$$\inf_{x \in \Omega} \{L_0(x, \lambda^*, \mu^*) - \sup_{\xi_1 \leq 0} \phi_1(\lambda^*, \xi_1)\} > -\infty. \tag{33}$$

Then for any $\varepsilon > 0$, we have

$$\lim_{r \rightarrow +\infty} \inf_{x \in \Omega/W_1(\varepsilon)} L(x, \lambda^*, \mu^*, r) = +\infty, \tag{34}$$

and

$$\{x \in \Omega \mid L(x, \lambda^*, \mu^*, r) \leq v(0, 0)\} \subseteq W_1(\varepsilon) \cap W_2(\varepsilon), \tag{35}$$

whenever $r > 0$ is sufficiently large.

Proof. The proofs of (34) and (35) are given in parts (a) and (b), respectively.

(a) For any fixed $x \in \Omega/W_1(\varepsilon)$, it follows from the definition of $W_1(\varepsilon)$ that

$$\text{dist}(g(x), h(x); \mathbb{R}_-^m \times \{0\}^l) > \varepsilon,$$

which implies that for any (ξ_1, ξ_2) with $\xi_1 \leq 0, \xi_2 = 0$ we have

$$\inf_{x \in \Omega/W_1(\varepsilon)} \|(\xi_1, \xi_2) - (g(x), h(x))\| \geq \inf_{x \in \Omega/W_1(\varepsilon)} \text{dist}(g(x), h(x); \mathbb{R}_-^m \times \{0\}^l) \geq \varepsilon.$$

According to the valley-at-zero property of σ , for any $x \in \Omega/W_1(\varepsilon)$, there exists $\zeta > 0$ such that

$$\sigma(\xi_1 - g(x), \xi_2 - h(x)) \geq \zeta. \tag{36}$$

It follows from (8) that

$$\begin{aligned} L(x, \lambda^*, \mu^*, r) &\geq \inf_{\xi_1 \leq 0, \xi_2 = 0} \{L_0(x, \lambda^*, \mu^*) - \phi_1(\lambda^*, \xi_1) + r\sigma(\xi_1 - g(x), \xi_2 - h(x))\} \\ &\geq \inf_{\xi_1 \leq 0} \{L_0(x, \lambda^*, \mu^*) - \phi_1(\lambda^*, \xi_1) + r\zeta\} \\ &= L_0(x, \lambda^*, \mu^*) - \sup_{\xi_1 \leq 0} \phi_1(\lambda^*, \xi_1) + r\zeta, \end{aligned}$$

where the second inequality comes from (36). This implies that

$$\inf_{x \in \Omega/W_1(\varepsilon)} L(x, \lambda^*, \mu^*, r) \geq \inf_{x \in \Omega} \{L_0(x, \lambda^*, \mu^*) - \sup_{\xi_1 \leq 0} \phi_1(\lambda^*, \xi_1)\} + r\zeta.$$

Passing to limit in the above inequality, we get

$$\lim_{r \rightarrow +\infty} \inf_{x \in \Omega/W_1(\varepsilon)} L(x, \lambda^*, \mu^*, r) \geq \inf_{x \in \Omega} \{L_0(x, \lambda^*, \mu^*) - \sup_{\xi_1 \leq 0} \phi_1(\lambda^*, \xi_1)\} + \lim_{r \rightarrow +\infty} r\zeta = +\infty,$$

where the equality comes from the fact that $\inf_{x \in \Omega} \{L_0(x, \lambda^*, \mu^*) - \sup_{\xi_1 \leq 0} \phi_1(\lambda^*, \xi_1)\}$ is finite by (33). Hence, (34) is true.

(b) First prove that

$$\{x \in \Omega \mid L(x, \lambda^*, \mu^*, r) \leq v(0, 0)\} \subseteq W_1(\varepsilon).$$

We argue it by contradiction. If there exist $\varepsilon_0 > 0$, $r_k \rightarrow \infty$, and $x_k \in \Omega$ such that

$$L(x_k, \lambda^*, \mu^*, r_k) \leq v(0, 0), \quad x_k \notin W_1(\varepsilon_0),$$

then

$$L(x_k, \lambda^*, \mu^*, r_k) \geq \inf_{x \in \Omega/W_1(\varepsilon_0)} L(x, \lambda^*, \mu^*, r_k).$$

Passing to limit in the above inequality, we get

$$\liminf_{k \rightarrow \infty} L(x_k, \lambda^*, \mu^*, r_k) \geq \lim_{k \rightarrow \infty} \inf_{x \in \Omega/W_1(\varepsilon_0)} L(x, \lambda^*, \mu^*, r_k) = +\infty,$$

where the equality comes from part (a). Clearly, this contradicts the finiteness of $v(0, 0)$.

Next, we claim that

$$\{x \in \Omega \mid L(x, \lambda^*, \mu^*, r) \leq v(0, 0)\} \subseteq W_2(\varepsilon).$$

Suppose, on the contrary, that there exist $\varepsilon_0 > 0$, $r_k \rightarrow \infty$ and $x_k \in \Omega$ such that

$$L(x_k, \lambda^*, \mu^*, r_k) \leq v(0, 0), \quad x_k \notin W_2(\varepsilon_0). \tag{37}$$

From (7) and (37), we conclude that there exist $y_k + g(x_k) \leq 0, z_k + h(x_k) = 0$ such that

$$\begin{aligned} v(0, 0) + \frac{\varepsilon_0}{2} &\geq L(x_k, \lambda^*, \mu^*, r_k) + \frac{\varepsilon_0}{2} \\ &\geq f(x_k) - \phi_1(\lambda^*, y_k) - \phi_2(\mu^*, z_k) + r_k\sigma(y_k, z_k). \end{aligned} \tag{38}$$

By the property (A_4) , for above (λ^*, μ^*) , there exist $\rho > 0, \tau > 0$ such that

$$\rho\sigma(y, z) - \phi_1(\lambda^*, y) - \phi_2(\mu^*, z) \geq 0, \quad \forall (y, z) \in \mathbb{R}^{m+l} \setminus \tau\mathbb{B}_{\mathbb{R}^{m+l}}. \tag{39}$$

Further using valley-at-zero property of σ , for above $\tau > 0$ there exists $d_1 > 0$ such that

$$\sigma(y, z) \geq d_1, \quad \forall (y, z) \in \mathbb{R}^{m+l} \setminus \tau\mathbb{B}_{\mathbb{R}^{m+l}}. \tag{40}$$

Now, let us prove that $(y_k, z_k) \rightarrow 0$. Let us consider the following cases:

Case 1. There exists an infinite subset $N_1 \subseteq N$ such that $\|(y_k, z_k)\| \geq \tau$ for all $k \in N_1$. Note that

$$\begin{aligned} v(0,0) + \varepsilon_0 &\geq f(x_k) - \phi_1(\lambda^*, y_k) - \phi_2(\mu^*, z_k) + \rho\sigma(y_k, z_k) + (r_k - \rho)\sigma(y_k, z_k) \\ &\geq f(x_k) + (r_k - \rho)\sigma(y_k, z_k) \\ &\geq f(x_k) + (r_k - \rho)d_1, \end{aligned} \tag{41}$$

where the second inequality comes from the assumption (39), and the third step is due to (40). The right side in (41) can be arbitrary large as $N_1 \ni k \rightarrow \infty$, which contradicts the finiteness of $v(0,0)$.

Case 2. $\|(y_k, z_k)\| \leq \tau$ as k sufficiently large. Since ϕ_1 and ϕ_2 are continuous by the property A_1 , for above (λ^*, μ^*) and $\tau > 0$, we can find $d_2 \in \mathbb{R}$ such that

$$\phi_1(\lambda^*, y) + \phi_2(\mu^*, z) \leq d_2, \quad \forall (y, z) \in \mathbb{R}^{m+l} \setminus \tau\mathbb{B}_{\mathbb{R}^{m+l}}.$$

Then

$$\begin{aligned} v(0,0) + \varepsilon_0 &\geq f(x_k) - \phi_1(\lambda^*, y_k) - \phi_2(\mu^*, z_k) + r_k\sigma(y_k, z_k) \\ &\geq f(x_k) - d_2 + r_k\sigma(y_k, z_k) \\ &\geq -d_2 + r_k\sigma(y_k, z_k). \end{aligned}$$

Due to the boundedness of (y_k, z_k) , we have

$$\sigma(y_k, z_k) \leq \frac{v(0,0) + \varepsilon_0 + d_2}{r_k} \rightarrow 0, \quad \text{as } k \rightarrow \infty,$$

which in turn implies that $(y_k, z_k) \rightarrow 0$ by the valley-at-zero property of σ .

Applying $(y_k, z_k) \rightarrow 0$ into (38) yields $v(0,0) + \frac{\varepsilon_0}{2} \geq f(x_k)$. It contradicts $x_k \notin W_2(\varepsilon_0)$ by the definition of $W_2(\varepsilon)$. Therefore, (35) holds. \square

Remark 1. Note that $\phi_2(\mu^*, \xi_2)$ is not used in the assumption (33). The reason is $\phi_2(\mu^*, \xi_2) = 0$ as $\xi_2 = 0$, since the perturbation for equality constraint is restricted to the subspace $\{0\}^l$.

Theorem 5. Suppose that $\sigma : \mathbb{R}^{m+l} \rightarrow \mathbb{R}_+$ has a valley at zero and

$$\inf_{x \in \Omega} \{L_0(x, \lambda^*, \mu^*) - \sup_{\xi_1 \leq 0} \phi_1(\lambda^*, \xi_1)\} > -\infty.$$

For any $x^* \in X^*$, (x^*, λ^*, μ^*) is a local saddle point of $L(x, \lambda^*, \mu^*, r)$ for some $r^* > 0$ and there exist a bounded subset $\Omega \subset \mathbb{R}^n$ and $\varepsilon_0 > 0$ such that

$$\{x \in \Omega \mid \text{dist}(g(x), h(x); \mathbb{R}_-^m \times \{0\}^l) \leq \varepsilon_0, f(x) - v(0,0) \leq \varepsilon_0\} \subset \Omega. \tag{42}$$

Then (λ^*, μ^*) is a generalized augmented Lagrange multiplier of (P).

Proof. According to the relationship among the generalized augmented Lagrange multiplier, the zero duality gap property, and global saddle points established in Theorems 1 and 2, we only need to justify that (x^*, λ^*, μ^*) is a global saddle point of $L(x, \lambda, \mu, r)$.

According to the definition of local saddle points, there exists $\delta > 0$ such that

$$L(x^*, \lambda, \mu, r^*) \leq L(x^*, \lambda^*, \mu^*, r^*) \leq L(x, \lambda^*, \mu^*, r^*), \quad \forall x \in \mathbb{B}_{\mathbb{R}^n}(x^*, \delta) \cap \Omega, (\lambda, \mu) \in \mathbb{R}_+^m \times \mathbb{R}^l. \tag{43}$$

It follows by invoking (14) and the first inequality in (43) that

$$L(x^*, \lambda^*, \mu^*, r^*) = f(x^*). \tag{44}$$

By the monotonicity of $L(x^*, \lambda^*, \mu^*, r)$ in r , we also have

$$L(x^*, \lambda^*, \mu^*, r^*) \leq L(x^*, \lambda^*, \mu^*, r) \leq f(x^*), \quad \forall r \geq r^*, \tag{45}$$

where the second inequality comes from (12). Combining (44) and (45) implies

$$L(x^*, \lambda^*, \mu^*, r) = f(x^*), \quad \forall r \geq r^*, \tag{46}$$

which together with (12) again yields

$$L(x^*, \lambda^*, \mu^*, r) = f(x^*) \geq L(x^*, \lambda, \mu, r), \quad \forall r \geq r^*, \quad (\lambda, \mu) \in \mathbb{R}_+^m \times \mathbb{R}^l. \tag{47}$$

Now, we establish the first inequality in (9). To complete the proof, it remains to show that

$$L(x^*, \lambda^*, \mu^*, r) \leq L(x, \lambda^*, \mu^*, r), \quad \forall x \in \Omega \setminus \mathbb{B}_{\mathbb{R}^n}(x^*, \delta), \tag{48}$$

whenever r is sufficiently large. Suppose on the contrary that we can find $r_k \rightarrow +\infty$ and $x_k \in \Omega \setminus \mathbb{B}_{\mathbb{R}^n}(x^*, \delta)$ such that

$$L(x_k, \lambda^*, \mu^*, r_k) < L(x^*, \lambda^*, \mu^*, r_k). \tag{49}$$

Hence, applying (46) into (49) and together with the fact $x^* \in X^*$ yields

$$L(x_k, \lambda^*, \mu^*, r_k) < f(x^*) = v(0, 0), \tag{50}$$

which means that x_k belongs to the set $\{x \in \Omega \mid L(x, \lambda^*, \mu^*, r_k) \leq v(0, 0)\}$. Taking into account of (35) in Lemma 1, we obtain that for any $\varepsilon \in (0, \varepsilon_0)$, $x_k \in W_1(\varepsilon) \cap W_2(\varepsilon)$, which further implies that $x_k \in \Lambda$ by (42). We can assume without loss of generality that x_k converges to \bar{x} . According to the continuity of $f(x)$, $g(x)$ and $h(x)$, together with the closedness of $W_1(\varepsilon)$ and $W_2(\varepsilon)$, we obtain that $\bar{x} \in W_1(\varepsilon) \cap W_2(\varepsilon)$. Therefore, $\bar{x} \in W_1(0) \cap W_2(0)$ by the arbitrariness of $\varepsilon > 0$, which further implies that $\bar{x} \in X^*$. By assumption, $(\bar{x}, \lambda^*, \mu^*)$ is also a local saddle point of $L(x, \lambda, \mu, r)$ for some $\bar{r} > 0$; i.e., there exists $\bar{\delta} > 0$ such that

$$L(\bar{x}, \lambda, \mu, \bar{r}) \leq L(\bar{x}, \lambda^*, \mu^*, \bar{r}) \leq L(x, \lambda^*, \mu^*, \bar{r}), \quad \forall x \in \mathbb{B}_{\mathbb{R}^n}(\bar{x}, \bar{\delta}) \cap \Omega, \quad (\lambda, \mu) \in \mathbb{R}_+^m \times \mathbb{R}^l. \tag{51}$$

Similar to the above argument, it follows from (44) that

$$L(\bar{x}, \lambda^*, \mu^*, \bar{r}) = f(\bar{x}) = v(0, 0) = \text{val}(P). \tag{52}$$

Since $x_k \in \mathbb{B}_{\mathbb{R}^n}(\bar{x}, \bar{\delta})$ and $r_k \geq \bar{r}$ for k large enough, from (51) and (52), it follows

$$L(x_k, \lambda^*, \mu^*, r_k) \geq L(x_k, \lambda^*, \mu^*, \bar{r}) \geq L(\bar{x}, \lambda^*, \mu^*, \bar{r}) = f(\bar{x}) = f(x^*) = v(0, 0),$$

which contradicts (50). This justifies (48).

By the fact (43), (47) and (48), we conclude that (x^*, λ^*, μ^*) is a global saddle point of $L(x, \lambda, \mu, r)$ for r large enough. Therefore, (λ^*, μ^*) is a generalized augmented Lagrange multiplier of (P) . \square

Theorem 6. Suppose that $\sigma : \mathbb{R}^{m+1} \rightarrow \mathbb{R}_+$ has a valley at zero and

$$\inf_{x \in \Omega} \{L_0(x, \lambda^*, \mu^*) - \sup_{\xi_1 \leq 0} \phi_1(\lambda^*, \xi_1)\} > -\infty.$$

Let x^* be the unique global optimal solution of (P). If (x^*, λ^*, μ^*) is a local saddle point of $L(x, \lambda, \mu, r)$ for some $r \geq 0$, and there exists $\varepsilon_0 > 0$ such that

$$\{x \in \Omega \mid \text{dist}(g(x), h(x); \mathbb{R}^m \times \{0\}^l) \leq \varepsilon_0\} \subset \Lambda, \tag{53}$$

where Λ is a bounded subset in \mathbb{R}^n , then (λ^*, μ^*) is a generalized augmented Lagrange multiplier of (P).

Proof. For $\varepsilon_0 > 0$, it follows from (34) in Lemma 1 that there exists $r_1 > 0$ such that

$$\inf_{x \in \Omega/W_1(\varepsilon_0)} L(x, \lambda^*, \mu^*, r_1) \geq f(x^*).$$

That is to say, for any $x \in \Omega \setminus W_1(\varepsilon_0)$, we have

$$L(x, \lambda^*, \mu^*, r_1) \geq f(x^*), \quad \forall x \in \Omega \setminus W_1(\varepsilon_0). \tag{54}$$

To complete the proof, we next need to show that there exists $r_2 > 0$ such that

$$L(x, \lambda^*, \mu^*, r_2) \geq f(x^*), \quad \forall x \in W_1(\varepsilon_0). \tag{55}$$

Suppose on the contrary that there exist $r_k \rightarrow \infty$ and $\{x_k\} \subset W_1(\varepsilon_0)$ such that

$$f(x^*) > L(x_k, \lambda^*, \mu^*, r_k). \tag{56}$$

According to (53) and Λ being bounded, $W_1(\varepsilon_0)$ is bounded, which further implies that $\{x_k\}$ has at least a cluster point \bar{x} . We assume without loss of generality that x_k converges \bar{x} .

We now claim that \bar{x} is a feasible point of (P). If \bar{x} is not feasible, then $\text{dist}(g(x), h(x); \mathbb{R}^m \times \{0\}^l) > 2m_0$ for some $m_0 > 0$. Therefore, $\text{dist}(g(x_k), h(x_k); \mathbb{R}^m \times \{0\}^l) > m_0$ as k sufficiently large. It in turn implies

$$L(x_k, \lambda^*, \mu^*, r_k) \geq \inf_{x \in \Omega/W_1(m_0)} L(x, \lambda^*, \mu^*, r_k).$$

Taking the limits on both sides yields

$$\liminf_{k \rightarrow \infty} L(x_k, \lambda^*, \mu^*, r_k) \geq \lim_{k \rightarrow \infty} \inf_{x \in \Omega/W_1(m_0)} L(x, \lambda^*, \mu^*, r_k) = +\infty, \tag{57}$$

where the equality comes from Lemma 1. Combining (56) with (57) together yields a contradiction to the finiteness of $f(x^*)$. This justifies the feasibility of \bar{x} for (P).

By hypothesis, (x^*, λ^*, μ^*) is a local saddle point of $L(x, \lambda, \mu, r)$ for some $r \geq 0$; then there exists a neighborhood $\mathbb{B}_{\mathbb{R}^n}(x^*, \delta)$ such that

$$L(x^*, \lambda, \mu, r) \leq L(x^*, \lambda^*, \mu^*, r) \leq L(x, \lambda^*, \mu^*, r), \quad \forall x \in \mathbb{B}_{\mathbb{R}^n}(x^*, \delta) \cap \Omega, (\lambda, \mu) \in \mathbb{R}_+^m \times \mathbb{R}^l. \tag{58}$$

Putting (14), (58), and the monotonicity of $L(x^*, \lambda^*, \mu^*, r)$ with respect to r together means that for any $r' \geq r$,

$$\begin{aligned} f(x^*) &= \sup_{(\lambda, \mu) \in \mathbb{R}_+^m \times \mathbb{R}^l} L(x^*, \lambda, \mu, r) = L(x^*, \lambda^*, \mu^*, r) \leq L(x^*, \lambda^*, \mu^*, r') \\ &\leq \sup_{(\lambda, \mu) \in \mathbb{R}_+^m \times \mathbb{R}^l} L(x^*, \lambda, \mu, r') = f(x^*). \end{aligned}$$

That is,

$$f(x^*) = L(x^*, \lambda^*, \mu^*, r'), \quad \forall r' \geq r. \tag{59}$$

Taking into account of (56), we obtain that

$$L(x^*, \lambda^*, \mu^*, r) = f(x^*) > L(x_k, \lambda^*, \mu^*, r_k) \geq L(x_k, \lambda^*, \mu^*, r),$$

where the last step is due to the monotonicity of $L(x, \lambda^*, \mu^*, r)$ with respect to r . Thus, it follows from (58) that $x_k \notin \mathbb{B}_{\mathbb{R}^n}(x^*, \delta)$ whenever k is sufficiently large; i.e., $\bar{x} \neq x^*$. Since x^* is the unique global optimal solution of (P) and invoking that \bar{x} is feasible, we have

$$f(\bar{x}) - f(x^*) > 0.$$

Define $\varrho := (f(\bar{x}) - f(x^*)) / 2 > 0$. Using (7) and (56), there exist $y_k \leq -g(x_k), z_k = -h(x_k)$ such that

$$\begin{aligned} f(x^*) + \varrho &> L(x_k, \lambda^*, \mu^*, r_k) + \varrho \\ &\geq f(x_k) - \phi_1(\lambda^*, y_k) - \phi_2(\mu^*, z_k) + r_k \sigma(y_k, z_k). \end{aligned}$$

Similarly to the argument given in Lemma 1, we conclude from the property (A_4) that $(y_k, z_k) \rightarrow 0$. Hence $f(x^*) + \varrho \geq f(x_k)$. Passing to limit, we get $f(x^*) + \varrho \geq f(\bar{x})$, which together with the definition of ϱ implies that $f(x^*) \geq f(\bar{x})$. It is clearly the case that \bar{x} is also an optimal solution. Hence $\bar{x} = x^*$, since the optimal solution is unique. This justifies (55).

Let $r^* := \max\{r_1, r_2\}$. Taking into account of (54) and (55), we obtain that

$$L(x, \lambda^*, \mu^*, r^*) \geq f(x^*) = L(x^*, \lambda^*, \mu^*, r^*), \quad \forall x \in \Omega, \tag{60}$$

where the equation comes from (59). Hence, according to (11) and (60),

$$\inf_{(y,z) \in \mathbb{R}^{m+l}} \{v_{r^*}(y, z) - \phi_1(\lambda^*, y) - \phi_2(\mu^*, z)\} = \inf_{x \in \Omega} L(x, \lambda^*, \mu^*, r^*) = f(x^*) = v(0, 0),$$

where the last step comes from that x^* is the optimal solution of (P). This further implies

$$v_{r^*}(y, z) \geq v_{r^*}(0, 0) + \phi_1(\lambda^*, y) + \phi_2(\mu^*, z), \quad \forall (y, z) \in \mathbb{R}^{m+l}.$$

Hence (λ^*, μ^*) is a generalized augmented Lagrange multiplier of (P). \square

The existence of generalized augmented Lagrange multipliers is established in two different scenarios: one is applicable to the case of unique solution while another is applicable to the case of multiple optimal solutions.

5. Conclusions

In this paper, we studied the generalized augmented Lagrangian multiplier, which is an extension of the augmented Lagrangian multiplier from linear support to nonlinear support for an augmented perturbation function. Some sufficient conditions for the existence of generalized augmented Lagrangian multipliers were developed. In particular, the relationships among global saddle points, generalized augmented Lagrangian multipliers, and the zero duality gap property between the primal problem and its generalized augmented Lagrangian dual problem were established. Several interesting topics are left for further investigation. For example, one is developing some necessary and sufficient conditions for the existence of generalized augmented Lagrangian multipliers by using the localization principle; another is studying the generalized differentiation of support functions from the subdifferential view.

Author Contributions: All authors contributed equally and significantly in writing this article. All authors have read and agreed to the published version of the manuscript.

Funding: This work is supported by National Natural Science Foundation of China (11771255, 11801325) and Young Innovation Teams of Shandong Province (2019KJ1013).

Acknowledgments: The authors are gratefully indebted to the anonymous referees for their valuable suggestions that helped us greatly improve the original presentation of the paper.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Birgin, E.G.; Martinez, J.M. *Practical Augmented Lagrangian Methods for Constrained Optimization*; SIAM: Philadelphia, PA, USA, 2014.
2. Curtis, F.E.; Jiang, H.; Robinson, D.P. An adaptive augmented Lagrangian method for large-scale constrained optimization. *Math. Program.* **2015**, *152*, 201–245. [[CrossRef](#)]
3. Rockafellar, R.T. Augmented Lagrange multiplier functions and duality in nonconvex programming. *SIAM J. Control Optim.* **1974**, *12*, 268–285. [[CrossRef](#)]
4. Kiwiel, K.C. On the twice differentiable cubic augmented Lagrangian. *J. Optim. Theory Appl.* **1996**, *88*, 233–236. [[CrossRef](#)]
5. Mangasarian, O.L. Unconstrained Lagrangians in nonlinear programming. *SIAM J. Control Optim.* **1975**, *12*, 772–791. [[CrossRef](#)]
6. Wu, H.X.; Luo, H.Z. Saddle points of general augmented Lagrangians for constrained nonconvex optimization. *J. Glob. Optim.* **2012**, *53*, 683–697. [[CrossRef](#)]
7. Tseng, P.; Bertsekas, D.P. On the convergence of the exponential multiplier method for convex programming. *Math. Program.* **1993**, *60*, 1–19. [[CrossRef](#)]
8. Wang, C.Y.; Li, D. Unified theory of augmented Lagrangian methods for constrained global optimization. *J. Glob. Optim.* **2009**, *44*, 433–458. [[CrossRef](#)]
9. Polyak, R.A. Log-sigmoid multipliers method in constrained optimization. *Ann. Oper. Res.* **2001**, *101*, 427–460. [[CrossRef](#)]
10. Polyak, R.A. Modified barrier functions: Theory and methods. *Math. Program.* **1992**, *54*, 177–222. [[CrossRef](#)]
11. Wu, H.X.; Luo, H.Z. A note on the existence of saddle points of p-th power Lagrangian for constrained nonconvex optimization. *Optimization* **2012**, *61*, 1231–1245. [[CrossRef](#)]
12. Wang, C.Y.; Yang, X.Q.; Yang, X.M. Unified nonlinear Lagrangian approach to duality and optimal paths. *J. Optim. Theory Appl.* **2007**, *135*, 85–100. [[CrossRef](#)]
13. Burachik, R.S.; Iusem, A.N.; Melo, J.G. Duality and exact penalization for general augmented Lagrangians. *J. Optim. Theory Appl.* **2010**, *147*, 125–140. [[CrossRef](#)]
14. Wang, C.Y.; Yang, X.Q.; Yang, X.M. Nonlinear augmented Lagrangian and duality theory. *Math. Oper. Res.* **2012**, *38*, 740–760. [[CrossRef](#)]
15. Wang, C.; Liu, Q.; Qu, B. Global saddle points of nonlinear augmented Lagrangian functions. *J. Glob. Optim.* **2017**, *68*, 125–146. [[CrossRef](#)]
16. Zhang, L.W.; Gu, J.; Xiao, X.T. A class of nonlinear Lagrangians for nonconvex second-order cone programming. *Comput. Optim. Appl.* **2011**, *49*, 61–99. [[CrossRef](#)]
17. Zhou, J.C.; Chen, J.S. On the existence of saddle points for nonlinear second-order cone programming problems. *J. Glob. Optim.* **2015**, *62*, 459–480. [[CrossRef](#)]
18. Fukuda, E.H.; Lourenco, B.F. Exact augmented Lagrangian functions for nonlinear semidefinite programming. *Comput. Optim. Appl.* **2018**, *71*, 457–482. [[CrossRef](#)]
19. Sun, D.F.; Sun, J.; Zhang, L.W. The rate of convergence of the augmented Lagrangian method for nonlinear semidefinite programming. *Math. Program.* **2008**, *114*, 349–391. [[CrossRef](#)]
20. Zhao, X.Y.; Sun, D.F.; Toh, K.-C. A Newton-CG augmented Lagrangian method for semidefinite programming. *SIAM J. Optim.* **2010**, *20*, 1737–1765. [[CrossRef](#)]
21. Dolgopolik, M.V. Augmented Lagrangian functions for cone constrained optimization: the existence of global saddle points and exact penalty property. *J. Glob. Optim.* **2018**, *71*, 237–296. [[CrossRef](#)]
22. Shapiro, A.; Sun, J. Some properties of the augmented Lagrangian in cone constrained optimization. *Math. Oper. Res.* **2004**, *29*, 479–491. [[CrossRef](#)]
23. Zhou, Y.Y.; Zhou, J.C.; Yang, X.Q. Existence of augmented Lagrange multipliers for cone constrained optimization problems. *J. Glob. Optim.* **2014**, *58*, 243–260. [[CrossRef](#)]
24. Burachik, R.S.; Yang, X.Q.; Zhou, Y.Y. Existence of augmented Lagrange multipliers for semi-infinite programming problems. *J. Optim. Theory Appl.* **2017**, *173*, 471–503. [[CrossRef](#)]

25. Rückmann, J.-J.; Shapiro, A. Augmented Lagrangians in semi-infinite programming. *Math. Program.* **2009**, *116*, 499–512. [[CrossRef](#)]
26. Wang, C.Y.; Zhou, J.C.; Xu, X.H. Saddle points theory of two classes of augmented Lagrangians and its applications to generalized semi-infinite programming. *Appl. Math. Optim.* **2009**, *59*, 413–434. [[CrossRef](#)]
27. Chatzipanagiotis, N.; Dentcheva, D.; Zavlanos, M.M. An augmented Lagrangian method for distributed optimization. *Math. Program.* **2015**, *152*, 405–434. [[CrossRef](#)]
28. Feizollahi, M.J.; Ahmed, S.; Sun, A. Exact augmented Lagrangian duality for mixed integer linear programming. *Math. Program.* **2017**, *161*, 365–387. [[CrossRef](#)]
29. Boland, N.; Christiansen, J.; Dandurand, B.; Eberhard, A.; Oliveira, F. A parallelizable augmented Lagrangian method applied to large-scale non-convex-constrained optimization problems. *Math. Program.* **2019**, *175*, 503–536. [[CrossRef](#)]
30. Kanzow, C.; Steck, D. Augmented Lagrangian methods for the solution of generalized Nash equilibrium Problems. *SIAM J. Optim.* **2016**, *26*, 2034–2058. [[CrossRef](#)]
31. Kanzow, C.; Steck, D. Quasi-variational inequalities in Banach spaces: theory and augmented Lagrangian methods. *SIAM J. Optim.* **2019**, *29*, 3174–3200. [[CrossRef](#)]
32. Liu, Y.F.; Liu, X.; Ma, S.Q. On the nonergodic convergence rate of an inexact augmented Lagrangian framework for composite convex programming. *Math. Oper. Res.* **2019**, *44*, 632–650. [[CrossRef](#)]
33. Teng, Y.; Yang, L.; Song, X.L.; Yu, B. An augmented Lagrangian proximal alternating method for sparse discrete optimization problems. *Numer. Algorithms* **2020**, *83*, 833–866. [[CrossRef](#)]
34. Rockafellar, R.T.; Wets, J.-B. *Variational Analysis*; Springer: New York, NY, USA, 1998.
35. Huang, X.X.; Yang, X.Q. A unified augmented Lagrangian approach to duality and exact penalization. *Math. Oper. Res.* **2003**, *28*, 533–552. [[CrossRef](#)]
36. Rubinov, A.M.; Huang, X.X.; Yang, X.Q. The zero duality gap property and lower semicontinuity of the perturbation function. *Math. Oper. Res.* **2002**, *27*, 775–791. [[CrossRef](#)]
37. Burachik, R.S.; Rubinov, A. Abstract convexity and augmented Lagrangians. *SIAM J. Optim.* **2007**, *18*, 413–436. [[CrossRef](#)]
38. Kan, C.; Song, W. Second-order conditions for existence of augmented Lagrange multipliers for eigenvalue composite optimization problems. *J. Glob. Optim.* **2015**, *63*, 77–97. [[CrossRef](#)]
39. Bonnans, J.F.; Shapiro, A. *Perturbation Analysis of Optimization Problems*; Springer: New York, NY, USA, 2000.
40. Ramana, M.; Tuncel, L.; Wolkowicz, H. Strong duality for semidefinite programming. *SIAM J. Optim.* **1997**, *7*, 641–662. [[CrossRef](#)]
41. Borwein, J.M.; Wolkowicz, H. Characterization of optimality for the abstract convex program with finite-dimensional range. *J. Aust. Math. Soc.* **1981**, *30*, 390–411. [[CrossRef](#)]
42. Kostyukova, O.I.; Tchemisova, T.V. Optimality conditions for convex semi-infinite programming problems with finitely representable compact index sets. *J. Optim. Theory Appl.* **2017**, *175*, 76–103. [[CrossRef](#)]
43. Kostyukova, O.I.; Tchemisova, T.V. Optimality criteria without constraint qualification for linear semidefinite problems. *J. Math. Sci.* **2012**, *182*, 126–143. [[CrossRef](#)]
44. Dolgopolik, M.V. Existence of augmented Lagrange multipliers: Reduction to exact penalty functions and localization principle. *Math. Program.* **2017**, *166*, 297–326. [[CrossRef](#)]

