Yokohama Publishers
ISSN 1349-8169 ONLINE JOURNAL

Pacific Journal of Optimization
Vol. 16, No. 2, pp. 315–327, 2020

LIMIT ANALYSIS FOR THE OPTIMAL VALUE OF A CLASS OF MINIMAX OPTIMIZATION PROBLEMS*

Fei Chen, Zhi Guo Feng† and K.F.C. Yiu

Abstract: A novel method is presented to analyze the limit for the optimal value of a class of minimax optimization problems with parameter. The minimax optimization problems can be transformed into semi-infinite programming (SIP) problems, and the optimal value series of the SIP problems with parameter are analyzed. First, we obtain that the optimal values of the cost function is monotonically decreasing as the parameter increases, and then the limit exists and can be computed by choosing a sufficiently large parameter. Next, we propose a novel method to obtain the limit of the optimal values by introducing a series of simplified subproblems. We derive the conditions and apply the fixed point theorem to prove that the function obtained by the proposed method is exactly the continuous limit function when the parameter tends to infinity, and the maximum value obtained by the proposed method is exactly the limit of optimal value series as the parameter tends to infinity. For illustration, numerical experiments are demonstrated to show the effectiveness and efficiency of the proposed method.

Key words: *minimax optimization, semi-infinite programming, fixed point theorem*

Mathematics Subject Classification: *49K35, 90C34*

1 Introduction

In real applications, the optimization problem is widely considered [2, 4, 5, 10]. Among the optimization problems, minimax optimization is a class of non-differentiable optimization problems, where the cost function is the worst case of several objectives. This kind of optimization problems has been widely used in engineering design and game theory. In many cases, the minimax optimization problem is equivalent to a semi-infinite programming problem [3,6,7]. In general, the algorithms for solving the semi-infinite programming problem include the discretization method [6, 7], gradient based method [1], cutting plane method [8, 9], exchange method [11] and so on. All these algorithms have their own characteristics. For example, the discretization method is very expensive, while it's also very stable and robust to solve any semi-infinite programming problems. The exchange method is very efficient, while it is only used in some special cases such as linear and convex problems. These methods can also be used to solve the minimax optimization problem.

*This paper is supported by the grant of National Science Foundation of Guangdong Province (No. 2020A1515010463), the grant of Chongqing Science and Technology Commission (No. cstc2017jcyjAX0161), the grant of Chongqing Normal University (No. 17XLB010), and program for scientific research start-up funds of Guangdong Ocean University. The third author is supported by RGC Grant PolyU. 152200/14E and PolyU Grant ZZGS.

†Corresponding author.

Minimax optimization plays important role in engineering design. In many engineering problems such as [2, 10], the maximum error depends on the frequency response function, where the frequency response function is related to the filter length, and the filter length represents the implementation complication and cost. Basically, if the filter length can be chosen arbitrary large, we hope that the optimized error can approach to the ideal zero error. However, this case is not always true. Hence, it's required to analyze the limit of optimal value series as the filter length increases. This can be treated as a guidance to decide whether the problem is designed according to a given performance index.

In this paper, we consider a class of minimax optimization problems with a parameter L, where L is the number of chosen basis functions. These problems can be transformed into semi-infinite programming problems. We first analyze the monotonicity of the optimal values and propose a novel method by formulating a series of simplified subproblems to find the limit of the optimal values as the parameter L tends to infinity. Then, we apply the fixed point theorem to prove that the maximum value obtained by the simplified subproblems is exactly the limit of the optimal values as the parameter L tends to infinity.

The paper is organized as follows. In section 2, we consider a class of minimax optimization problems and transform it into semi-infinite programming problems. In section 3, we first analyze the monotonicity of the optimal values. Then, we propose a new method to compute the limit of optimal values by introducing a series of simplified subproblems. We apply the fixed point theorem to verify that the maximum optimal value of the subproblems is exactly the limit of the optimal values of the semi-infinite programming problems. Numerical examples are illustrated in section 4 and conclusion is summarized in section 5.

2 Problem Formulation

We consider a class of minimax optimization problems in this paper which comes from the application problems in [2, 10]. The problem is formulated as follows:

$$(P) \qquad \begin{array}{l} \min\limits_{\boldsymbol{x}} \max\limits_{t\in\Omega} \quad g(\boldsymbol{H}(\boldsymbol{x},t)) \\ s.t. \boldsymbol{H}(\boldsymbol{x},t) = \boldsymbol{x}\boldsymbol{\varphi}(t), \forall t \in \Omega, \end{array}$$

where

$$\boldsymbol{x} = (\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_N)^T, \boldsymbol{x}_i = (x_{i1}, x_{i2}, \cdots, x_{iL}),$$

$$\boldsymbol{H} = (H_1, H_2, \cdots, H_N)^T, \boldsymbol{\varphi} = (\varphi_1, \varphi_2, \cdots, \varphi_L)^T,$$

$$H_i(\boldsymbol{x},t) = \sum_{k=1}^{L} x_{ik}\varphi_k(t)(i = 1, 2, \cdots, N).$$

The functions $\{\varphi_k(t), k = 1, 2, \cdots\}$ is a set of basis functions in continuous function space, and Ω is a given nonempty compact set in $\mathbb{R}$. Hence, for each i, $H_i(\boldsymbol{x},t)$ is expressed as the linear combination of the basis functions and L is the number of basis functions. The function $g \in C^2(\mathbb{R}^N, \mathbb{R})$ and has lower bound, that is, there exists a real number $M \in \mathbb{R}$ such that

$$g(\boldsymbol{H}(\boldsymbol{x},t)) \geq M, \quad \forall \boldsymbol{x} \in \mathbb{R}^{NL}, \forall t \in \Omega.$$

It can be seen that the function $\max\limits_{t\in\Omega} g(\boldsymbol{H}(\boldsymbol{x},t))$ in Problem (P) is nonsmooth, which can not be solved directly. In general, this problem can be transformed into a smooth optimization problem. For this, let $y = \max_{t\in\Omega} g(\boldsymbol{H}(\boldsymbol{x}, t))$, the problem (P) is equivalent to the problem as follows:

$$P(L) \qquad \begin{array}{l} \min_{y,\boldsymbol{x}} \quad y \\ s.t. \begin{cases} g(\boldsymbol{H}(\boldsymbol{x}, t) - y \le 0, \quad \forall t \in \Omega, \\ \boldsymbol{H}(\boldsymbol{x}, t) = \boldsymbol{x}\boldsymbol{\varphi}(t), \end{cases} \end{array}$$

where Ω is a compact set with infinite points. For any point t in Ω, $g(\boldsymbol{H}(\boldsymbol{x}, t)) - y$ is a constraint function. Therefore, $g(\boldsymbol{H}(\boldsymbol{x}, t)) - y$ contains infinite constraint functions. Then, the minimax optimization problem (P) is transformed into a semi-infinite programming (SIP) problem $P(L)$. Note that both the cost function and the constraint function for each t are smooth, gradient based method can be used to solve the problem $P(L)$.

Obviously, the problem (P) (or $P(L)$) is related to the parameter L. Then, if the parameter L changes, the corresponding optimal solution and optimal value also change. In general, the parameter L is related to the implementation cost in real applications. If the parameter L increases, then the implementation cost also increases. Hence, we optimize the problem such that the performance can also be improved. It is necessary to analyze the change rule between the optimal value and the parameter L.

3 Optimal Solution Analysis

3.1 Optimal solution of the problem

Denote the optimal solution and optimal value of the problem $P(L)$ with the parameter L by y^{*L} and $\boldsymbol{x}^{*L}$, respectively. Obviously, we have

$$y^{*L} = \max_{t\in\Omega} g(\boldsymbol{H}(\boldsymbol{x}^{*L}, t)).$$

First, we have the following result.

Lemma 3.1. *The optimal value series $\{y^{*L} : L = 1, 2, \cdots\}$ of Problem $P(L)$ is monotonically decreasing and there exists a limit as L tends to infinity.*

Proof. Suppose that A_L is the feasible set of the problem $P(L)$ related to y, that is,

$$A_L = \{y \in \mathbb{R} \mid \exists \boldsymbol{x} \in \mathbb{R}^{NL}, s.t.\ g(\boldsymbol{H}(\boldsymbol{x}, t)) - y \le 0, \forall t \in \Omega\}.$$

To prove that $\{y^{*L} : L = 1, 2, \cdots\}$ is monotonically decreasing as L increases, it's required to show that $A_L \subset A_{L+1}$ as L increases, i.e., $\forall y \in A_L$, we have $y \in A_{L+1}$.

Since $\varphi_k(t)$ is a basis function and the set of basis functions can be spanned into a space. Denote $\Delta_L = \{\varphi_1(t), \varphi_2(t), \cdots, \varphi_L(t)\}$, then Δ_L is a set with respect to L. Obviously, we have $\Delta_L \subset \Delta_{L+1}$, and then $span(\Delta_L) \subset span(\Delta_{L_{k+1}})$, where

$$span(\Delta_{L_k}) = \{\overline{H}_i(t) : \overline{H}_i(t) = \boldsymbol{x}_i\boldsymbol{\varphi}(t) = \sum_{k=1}^{L} x_{ik}\varphi_k(t), i = 1, \ldots, N, \forall \boldsymbol{x}_i \in \mathbb{R}^L\},$$

$$\overline{\boldsymbol{H}} = (\overline{H}_1, \overline{H}_2, \cdots, \overline{H}_N).$$

Since $H_i(\boldsymbol{x}, t) \in span(\Delta_{L_k}), i \in \{1, 2, \cdots, N\}$, then for any $y \in A_L$, there exists a vector $\boldsymbol{x}^{(L)} \in \mathbb{R}^{NL}$, such that $g(\boldsymbol{H}(\boldsymbol{x}^{(L)}, t)) - y \le 0$. Note that

$$\boldsymbol{H}(\boldsymbol{x}^{(L)}, t) \in span^N(\Delta_L) \subset span^N(\Delta_{L+1}).$$

In particular, we choose one vector $\boldsymbol{x}^{(L+1)}$ as

$$\boldsymbol{x}^{(L+1)} = \left(\boldsymbol{x}_1^{(L+1)}, \boldsymbol{x}_2^{(L+1)}, \cdots, \boldsymbol{x}_N^{(L+1)}\right)^T,$$

$$\boldsymbol{x}_i^{(L+1)} = \left(x_i^{(L+1)}, x_i^{(L+1)}, \cdots, x_i^{(L+1)}, 0\right)^T,$$

then, we obtain

$$\boldsymbol{H}(\boldsymbol{x}^{(L+1)}, t) = \boldsymbol{H}(\boldsymbol{x}^{(L)}, t),$$

and

$$g(\boldsymbol{H}(\boldsymbol{x}^{(L+1)}, t)) - y = g(\boldsymbol{H}(\boldsymbol{x}^{(L)}, t)) - y \leq 0, \quad \forall t \in \Omega.$$

Hence, $y \in A_{L+1}$. Note that y is arbitrary, we have $A_L \subset A_{L+1}$.

Obviously, the optimal value series of the cost function is monotonically decreasing as L increases.

Moreover, g has a lower bound and the optimal value is monotonically decreasing as L increases, there must be a limit of the optimal value series $\{y^{*L}\}$. The proof completes. □

Since the optimal value series is monotonically decreasing, its limit must be the best value, which has the guidance sense in practice. That is, if the designed performance is better than the limit, it is impossible to find a suitable L such that performance is satisfied, and then the design can be ignored. If the designed performance is poorer than but very close to the limit, it is required to choose a sufficient large L to implement the design. If the designed performance is poorer than and not close to the limit, then the design is normal and can be implemented easily.

By Lemma 3.1, the limit $\lim\limits_{L\to+\infty} y^{*L}$ can be approximated by the optimal value y^{*L} when L is a sufficiently large value. However, if L increases, the number of variables also increases and the computation of the optimal value becomes very expensive. Then, efficient method for computing the limit is required.

Remark 3.2. For the choice of L to estimate the limit by the optimal value y^{*L}, it is required to analyze the estimated error between the limit and the optimal value y^{*L}. However, the estimation depends on the choice of basis functions. Then, there is no uniform formula of the estimation for different basis functions. Basically, the analysis of power series is easy, while the analysis of trigonometric series is relatively difficult. Furthermore, since there is an optimization operation, the function $\boldsymbol{H}(\boldsymbol{x}^{(L*)}, t)$ is not necessary the partial sum of the expansion of $\lim\limits_{L\to+\infty} \boldsymbol{H}(\boldsymbol{x}^{(L*)}, t)$. Then, the estimation of the error can be very large. For this, we estimate the limit by gradually increasing the parameter L in numerical experiment and choose L when the improvement of the performance is sufficiently small.

3.2 Limit value analysis

In order to find the limit of the problem $P(L)$, we propose a novel method by treating the function $\overline{\boldsymbol{H}}$ as the variables and solve the problem directly. That is, for each $t \in \Omega$, we formulate a subproblem as follows:

$$P(t) \qquad \begin{aligned} &\min_{y,\overline{\boldsymbol{H}}} \quad y \\ &s.t. \begin{cases} g(\overline{\boldsymbol{H}}, t) - y \leq 0, \\ \overline{\boldsymbol{H}} = (\overline{H}_1, \overline{H}_2, \cdots, \overline{H}_N). \end{cases} \end{aligned}$$

It can be seen that for each $t \in \Omega$, the problem $P(t)$ is a general optimization problem, where $\overline{\boldsymbol{H}}(t)$ is the decision vector. Denote the optimal value of the problem $P(t)$ by y_t^*. Then for any $t \in \Omega$, the optimal solution $\overline{\boldsymbol{H}}^*(t)$ and the corresponding optimal value can be obtained by solving Problem $P(t)$. Then, we set

$$y^* = \max_{t \in \Omega} y_t^*. \tag{3.1}$$

Note that the optimization of $\overline{\boldsymbol{H}}(t)$ has ignored the effect of the parameter L, therefore, y^* must be better than all the optimal values of the problem $P(L)$. To find the conditions such that y^* is the limit of the optimal value series $\{y^{*L} : L = 1, 2, \cdots\}$, we first have the following lemma.

Lemma 3.3. *If y_t^* is the optimal value of Problem $P(t)$ for any $t \in \Omega$, then the constraint must be active, that is, $g(\overline{\boldsymbol{H}}^*_t, t) = y_t^*$.*

Proof. Assume that the constraint is not active, that is, $g(\overline{\boldsymbol{H}}^*_t, t) - y_t^* < 0$. Define a new number y_0^* by $y_0^* = g(\overline{\boldsymbol{H}}^*_t, t)$, then we obtain

$$y_t^* > g(\overline{\boldsymbol{H}}^*_t, t) = y_0^*,$$

$$g(\overline{\boldsymbol{H}}^*_t, t) - y_0^* = 0.$$

Then, $(y_0^*, \overline{\boldsymbol{H}}^*_t)$ is also a feasible solution of Problem $P(t)$. However, y_0^* is better than y_t^*, which contradicts the optimality of y_t^*. Hence, the assumption does not hold and we have $g(\overline{\mathbf{H}}^*_{\mathbf{t}}, t) - y_t^* = 0$. The proof completes. □

By Lemma 3.3, if Problem $P(t)$ is optimized, the constraint must be active. Hence, $(y_t^*, \overline{\boldsymbol{H}}^*_t)$ is also the optimal solution of the following problem with equality constraint.

$$\widetilde{P}(t) \qquad \begin{aligned} &\min_{y, \overline{\boldsymbol{H}}} \quad y \\ &s.t. \begin{cases} g(\overline{\boldsymbol{H}}, t) - y = 0, \\ \overline{\boldsymbol{H}} = (\overline{H}_1, \overline{H}_2, \cdots, \overline{H}_N). \end{cases} \end{aligned}$$

The KKT optimality condition satisfies the following condition:

$$(\nabla_{\boldsymbol{w}} f(\boldsymbol{w}_0))^T = \mathbf{0}, \tag{3.2}$$

where $\boldsymbol{w} = (\lambda, y, \overline{\boldsymbol{H}})$, $f(\boldsymbol{w}) = y + \lambda(g(\overline{\mathbf{H}}, t) - y)$ is the Lagrange function, λ is a multiplier, and $\boldsymbol{w}_0 = (\lambda_0, y_0, \overline{\boldsymbol{H}}_0)$ is the optimal value of Problem $\widetilde{P}(t)$ at $t_0 \in \Omega$. By direct computations, (3.2) becomes

$$\boldsymbol{F}(t_0, \boldsymbol{w}_0) = \mathbf{0},$$

where

$$\boldsymbol{F}(t, \boldsymbol{w}) = \begin{pmatrix} 1 - \lambda \\ g(\overline{\boldsymbol{H}}, t) - y \\ \lambda \left(\nabla_{\overline{\boldsymbol{H}}} g(\overline{\boldsymbol{H}}, t)\right)^T \end{pmatrix}.$$

Note that $\boldsymbol{F}$ is a $(N+2)$-dimensional vector function, its gradient with respect to $\boldsymbol{w}$ can be computed as

$$\nabla \boldsymbol{F} = \begin{pmatrix} -1 & 0 & \mathbf{0} \\ 0 & -1 & \nabla_{\overline{\boldsymbol{H}}} g(\overline{\boldsymbol{H}}, t) \\ \left(\nabla_{\overline{\boldsymbol{H}}} g(\overline{\boldsymbol{H}}, t)\right)^T & 0 & \lambda \nabla^2_{\overline{\boldsymbol{H}}} g(\overline{\boldsymbol{H}}, t) \end{pmatrix}.$$

Then, by direct computations, we obtain

$$\det(\nabla \boldsymbol{F}) = \lambda \det\left(\nabla^2_{\overline{\boldsymbol{H}}} g(\overline{\boldsymbol{H}}, t)\right).$$

Based on the discussion above, it is obvious that the optimality condition is a group of equations for each t. Then, the optimal solution $\boldsymbol{w}$ can be determined by the equations (3.2) for any t in Ω. Hence, the optimal solution $\boldsymbol{w}$ is a vector function of t. To discuss whether this function is continuous, we have the following theorem.

Theorem 3.4. *Suppose that* $\boldsymbol{F}(t_0, \boldsymbol{w}_0) = 0$*,* $\det\left(\nabla^2_{\overline{\boldsymbol{H}}} g(\overline{\boldsymbol{H}}_0, t_0)\right) \neq 0$*. If* $\boldsymbol{F}$ *and* $\nabla \boldsymbol{F}$ *are continuous in* $U \times V \subset \mathbb{R} \times \mathbb{R}^{N+2}$*, which is a neighborhood of the point* $(t_0, \boldsymbol{w}_0)$*. Then, there exists a neighborhood* $U_0 \times V_0 \subset U \times V$ *of* $(t_0, \boldsymbol{w}_0)$ *, and a unique continuous function* $\boldsymbol{\psi} : U_0 \to V_0$ *such that* $\boldsymbol{F}(t, \boldsymbol{\psi}(t)) = 0$*,* $t \in U_0$*.*

Proof. First, we denote the closed interval with center t_0 and distance r by $\overline{B}(t_0, r)$, and the closed ball with center $\boldsymbol{w}_0$ and distance δ by $\overline{B}(\boldsymbol{w}_0, \delta)$. We denote all continuous vector functions which are defined in $\overline{B}(t_0, r)$ and taken value in $\overline{B}(\boldsymbol{w}_0, \delta)$ by $C(\overline{B}(t_0, r), \overline{B}(\boldsymbol{w}_0, \delta))$.

Since the constraint is active, we have $\lambda \neq 0$, and

$$\det(\nabla \boldsymbol{F}(t_0, \boldsymbol{w}_0)) = \lambda \det\left(\nabla^2_{\overline{\boldsymbol{H}}} g(\overline{\boldsymbol{H}}_0, t_0)\right) \neq 0.$$

Then, the matrix $\nabla \boldsymbol{F}(t_0, \boldsymbol{w}_0)$ is invertible and we have

$$(\nabla \boldsymbol{F}(t_0, \boldsymbol{w}_0))^{-1} \nabla \boldsymbol{F}(t_0, \boldsymbol{w}_0) = I. \tag{3.3}$$

Define a mapping $T : \boldsymbol{\psi} \to T\boldsymbol{\psi}$ as follows.

$$(T\boldsymbol{\psi})(t) = \boldsymbol{\psi}(t) - (\nabla \boldsymbol{F}(t_0, \boldsymbol{w}_0))^{-1} \boldsymbol{F}(t, \boldsymbol{\psi}(t)), t \in \overline{B}(t_0, r).$$

First, we define a norm in $C(\overline{B}(t_0, r), \overline{B}(\boldsymbol{w}_0, \delta))$. For any two vector functions $\boldsymbol{\psi}(t)$ and $\boldsymbol{\phi}(t)$ on $C(\overline{B}(t_0, r), \overline{B}(\boldsymbol{w}_0, \delta))$, its norm can be defined by

$$\|\boldsymbol{\psi}(t), \boldsymbol{\phi}(t)\|_\infty = \max_{i \in \{1,2,\cdots,N+2\}} \max_{t \in \overline{B}(t_0, r)} |\psi_i(t) - \phi_i(t)|.$$

It follows by (3.3) and the continuity of $\nabla \boldsymbol{F}$ that $\forall \epsilon > 0$, there exist $r > 0, \delta > 0$ and continuous functions $\boldsymbol{\varphi}^1, \cdots, \boldsymbol{\varphi}^{N+2} \in C(\overline{B}(t_0, r), \overline{B}(\boldsymbol{w}_0, \delta))$, such that

$$\left\| I - (\nabla \boldsymbol{F}(t_0, \boldsymbol{w}_0))^{-1} \boldsymbol{G}(\boldsymbol{\varphi}^1, \cdots, \boldsymbol{\varphi}^{N+2}) \right\|_\infty < \epsilon,$$

where $\|.\|_\infty$ is the maximum of the absolute value of all matrix elements, and

$$\boldsymbol{G}(\boldsymbol{\varphi}^1, \cdots, \boldsymbol{\varphi}^{N+2}) = \begin{pmatrix} \nabla F_1(t, \boldsymbol{\varphi}^1) \\ \vdots \\ \nabla F_{N+2}(t, \boldsymbol{\varphi}^{N+2}) \end{pmatrix}.$$

Hence, for any two vector functions $\boldsymbol{\psi}(t), \boldsymbol{\phi}(t)$ in $C(\overline{B}(t_0, r), \overline{B}(\boldsymbol{w}_0, \delta))$, we have

$$T\boldsymbol{\psi} - T\boldsymbol{\phi} = \boldsymbol{\psi}(t) - \boldsymbol{\phi}(t) - \left[(\nabla \boldsymbol{F}(t_0, \boldsymbol{w}_0))^{-1} (\boldsymbol{F}(t, \boldsymbol{\psi}(t)) - \boldsymbol{F}(t, \boldsymbol{\phi}(t)))\right].$$

Based on the mean value theorem, there exists $\eta_i \in (0, 1)$ such that

$$\boldsymbol{\varphi}^i(t) = \eta_i \boldsymbol{\psi}(t) + (1 - \eta_i)\boldsymbol{\phi}(t),$$

and

$$F_i(t,\boldsymbol{\psi}(t)) - F_i(t,\boldsymbol{\phi}(t)) = \nabla F_i(t,\boldsymbol{\varphi}^i(t))(\boldsymbol{\psi}(t)-\boldsymbol{\phi}(t)).$$

Note that

$$\begin{aligned} & \|\boldsymbol{\varphi}^i(t) - \boldsymbol{w}_0\|_\infty \\ = \; & \|\eta_i(\boldsymbol{\psi}(t)-\boldsymbol{w}_0) + (1-\eta_i)(\boldsymbol{\phi}(t)-\boldsymbol{w}_0)\|_\infty \\ \leq \; & \eta_i\|\boldsymbol{\psi}(t)-\boldsymbol{w}_0\|_\infty + (1-\eta_i)\|\boldsymbol{\phi}(t)-\boldsymbol{w}_0\|_\infty \\ \leq \; & \eta_i\delta + (1-\eta_i)\delta = \delta, \end{aligned}$$

we have $\boldsymbol{\varphi}^i(t) \in C(\overline{B}(t_0,r), \overline{B}(\boldsymbol{w}_0,\delta))$, $\forall i = 1,2,\cdots,N+2$. Then,

$$\begin{aligned} & (\nabla \boldsymbol{F}(t_0,\boldsymbol{w}_0))^{-1}(\boldsymbol{F}(t,\boldsymbol{\psi}(t)) - \boldsymbol{F}(t,\boldsymbol{\phi}(t))) \\ = \; & (\nabla \boldsymbol{F}(t_0,\boldsymbol{w}_0))^{-1}\begin{pmatrix} \nabla F_1(t,\boldsymbol{\varphi}^1) \\ \vdots \\ \nabla F_{N+2}(t,\boldsymbol{\varphi}^{N+2}) \end{pmatrix}(\boldsymbol{\psi}(t)-\boldsymbol{\phi}(t)) \\ = \; & (\nabla \boldsymbol{F}(t_0,\boldsymbol{w}_0))^{-1}\boldsymbol{G}(\varphi^1,\cdots,\boldsymbol{\varphi}^{N+2})(\boldsymbol{\psi}(t)-\boldsymbol{\phi}(t)). \end{aligned}$$

Hence, we have

$$\begin{aligned} & \|T\boldsymbol{\psi} - T\boldsymbol{\phi}\|_\infty \\ = \; & \left\|\left(I - (\nabla \boldsymbol{F}(t_0,\boldsymbol{w}_0))^{-1}\boldsymbol{G}(\boldsymbol{\varphi}^1,\cdots,\boldsymbol{\varphi}^{N+2})\right)(\boldsymbol{\psi}(t)-\boldsymbol{\phi}(t))\right\|_\infty \\ \leq \; & (N+2)\left\|I - (\nabla \boldsymbol{F}(t_0,\boldsymbol{w}_0))^{-1}\boldsymbol{G}(\boldsymbol{\varphi}^1,\cdots,\boldsymbol{\varphi}^{N+2})\right\|_\infty \cdot \|(\boldsymbol{\psi}(t)-\boldsymbol{\phi}(t))\|_\infty \\ \leq \; & (N+2)\epsilon\|(\boldsymbol{\psi}(t)-\boldsymbol{\phi}(t))\|_\infty. \end{aligned}$$

Set $\epsilon < \frac{1}{2(N+2)}$, it implies that

$$\|T\boldsymbol{\psi} - T\boldsymbol{\phi}\|_\infty < \frac{1}{2}\|\boldsymbol{\psi}-\boldsymbol{\phi}\|_\infty.$$

Therefore, T is a contraction mapping in the subspace $C(\overline{B}(t_0,r), \overline{B}(\boldsymbol{w}_0,\delta))$.

Furthermore, it follows by (3.3) and the continuity of $\boldsymbol{F}(t,\boldsymbol{w})$ that there exists $r_1 > 0$ such that if $0 < r < r_1$, we have for any $t \in \overline{B}(t_0,r)$ that

$$\begin{aligned} & \left\|(\nabla \boldsymbol{F}(t_0,\boldsymbol{w}_0))^{-1}\boldsymbol{F}(t,\boldsymbol{w}_0)\right\|_\infty \\ = \; & \left\|(\nabla \boldsymbol{F}(t_0,\boldsymbol{w}_0))^{-1}\boldsymbol{F}(t,\boldsymbol{w}_0) - \boldsymbol{F}(t_0,\boldsymbol{w}_0)\right\|_\infty \\ < \; & \tfrac{\delta}{2}. \end{aligned} \tag{3.4}$$

For convenience, we also denote the constant function defined in $\overline{B}(t_0, r)$ and equals to $\boldsymbol{w}_0$ by $\boldsymbol{w}_0$. Obviously, $\boldsymbol{w}_0 \in C(\overline{B}(t_0, r), \overline{B}(\boldsymbol{w}_0, \delta))$, then

$$T\boldsymbol{w}_0 = \boldsymbol{w}_0 - \left(\nabla \boldsymbol{F}(t_0, \boldsymbol{w}_0)\right)^{-1} \boldsymbol{F}(t, \boldsymbol{w}_0). \tag{3.5}$$

By (3.4) and (3.5), we have

$$\|T\boldsymbol{w}_0 - \boldsymbol{w}_0\|_\infty = \left\| \left(\nabla \boldsymbol{F}(t_0, \boldsymbol{w}_0)\right)^{-1} \boldsymbol{F}(t, \boldsymbol{w}_0) \right\|_\infty < \frac{\delta}{2}.$$

Thus, for any $\boldsymbol{\psi} \in C(\overline{B}(t_0, r), \overline{B}(\boldsymbol{w}_0, \delta))$ which satisfies $\boldsymbol{\psi}(t_0) = \boldsymbol{w}_0$, if $0 < r < r_1$, we obtain

$$\begin{aligned} \|T\boldsymbol{\psi} - \boldsymbol{w}_0\|_\infty &\le \|T\boldsymbol{\psi} - T\boldsymbol{w}_0\|_\infty + \|T\boldsymbol{w}_0 - \boldsymbol{w}_0\|_\infty \\ &\le \tfrac{1}{2}\|\boldsymbol{\psi} - \boldsymbol{w}_0\| + \tfrac{\delta}{2} \\ &\le \tfrac{\delta}{2} + \tfrac{\delta}{2} = \delta. \end{aligned}$$

Let $X = C(\overline{B}(t_0, r), \overline{B}(\boldsymbol{w}_0, \delta))$, then $(X, \|.\|_\infty)$ is a *Banach* space, which is complete. Note that T is a contraction mapping and $TX \subset X$. It follows by the fixed point theorem that there exists a unique $\boldsymbol{\psi} \in X$ such that $T\boldsymbol{\psi} = \boldsymbol{\psi}$, i.e.,

$$\boldsymbol{\psi}(t) - \left(\nabla \boldsymbol{F}(t_0, \boldsymbol{w}_0)\right)^{-1} \boldsymbol{F}(t, \boldsymbol{\psi}(t)) = \boldsymbol{\psi}(t), \quad \forall t \in \overline{B}(t_0, r).$$

Then,

$$\left(\nabla \boldsymbol{F}(t_0, \boldsymbol{w}_0)\right)^{-1} \boldsymbol{F}(t, \boldsymbol{\psi}(t)) = 0, \quad \forall t \in \overline{B}(t_0, r).$$

Hence, we have

$$\boldsymbol{F}(t, \boldsymbol{\psi}(t)) = 0, \quad \forall t \in \overline{B}(t_0, r).$$

Note that $\boldsymbol{\psi}(t)$ is unique, we obtain $\boldsymbol{\psi}(t_0) = \boldsymbol{w}_0$. The proof completes. □

In order to discuss the continuous property of $\boldsymbol{\psi}(t)$ for all $t \in \Omega$, we have the following theorem.

Theorem 3.5. *Suppose that* $\boldsymbol{F}(t, \boldsymbol{w}(t)) = 0$, $\det\left(\nabla^2_{\overline{\boldsymbol{H}}} g(\overline{\boldsymbol{H}}, t)\right) \neq 0$, $\forall t \in \Omega$. *If* $\boldsymbol{F}$ *and* $\nabla \boldsymbol{F}$ *are continuous in a neighborhood of the point* $(t, \boldsymbol{w}(t))$ *for any* $t \in \Omega$. *Then,* $\boldsymbol{w}(t)$ *must be continuous in* Ω *and is the unique function which satisfies* $\boldsymbol{F}(t, \boldsymbol{w}(t)) = 0, \forall t \in \Omega$.

Proof. It follows by Theorem 3.4 that for any $v \in \Omega$, there exist $r_v > 0$, and a unique continuous function $\boldsymbol{\psi}^v(t)$, which is defined in $\overline{B}(v, r_v)$, such that $\boldsymbol{F}(v, \boldsymbol{\psi}^v(t)) = 0, \forall t \in \overline{B}(v, r_v)$. Then, the equations $\boldsymbol{F}(v, \boldsymbol{\psi}^v(t)) = 0$ is also satisfied in the open interval $B(v, r_v)$.

For the set of open intervals $\{B(v, r_v)_i, v \in \Omega\}$, it is a cover of the set Ω. Since Ω is compact, there exists a finite subset $v_i, i = 1, 2, \cdots, m_v$, such that the set $\{B(v_i, r_{v_i}) : i = 1, \cdots, m_v\}$ is also a cover of Ω.

For each interval $B(v_i, r_{v_i})$, it follows by the uniqueness of $\boldsymbol{\psi}(t)$ which satisfies $\boldsymbol{F}(t, \boldsymbol{\psi}(t)) = 0$ that $\boldsymbol{\psi}(t) = \boldsymbol{w}(t), \forall t \in B(v_i, r_{v_i})$. Since $\boldsymbol{w}(t)$ is continuous in each interval $B(v_i, r_{v_i})$, $\boldsymbol{w}(t)$ is also continuous in the union of the set $\{B(v_i, r_{v_i}) : i = 1, \cdots, m_v\}$. Hence, $\boldsymbol{w}(t)$ is also continuous in Ω. The proof completes. □

Then, by the theorems above, we can show that the limit can be obtained by the proposed method, which is stated as follows.

Theorem 3.6. *Assume that the conditions in Theorem 3.4 hold for any t in Ω, then,*

$$\lim_{L\to+\infty} y^{*L} = \max_{t\in\Omega} y_t^*.$$

Proof. For each $t \in \Omega$, we can obtain the optimal cost function value y_t^* and the optimal solution $\overline{\boldsymbol{H}}_t^*$ of Problem $P(t)$, where the constraint is also active. Let $\overline{\boldsymbol{H}}^*(t) = \overline{\boldsymbol{H}}_t^*$, we can define a vector function $\overline{\boldsymbol{H}}^*(t)$ for all $t \in \Omega$ such that $g(\overline{\boldsymbol{H}}^*(t)) - y_t^* = 0$, where $\overline{\boldsymbol{H}}^*(t) = (\overline{H}_1^*(t), \overline{H}_2^*(t), \cdots, \overline{H}_N^*(t))$.

It follows by Theorem 3.5 that we have $\overline{\boldsymbol{H}}^*(t) \in C^N(\Omega)$. Then, there exists a series $\{\boldsymbol{H}^{(L)}(t) : L = 1, 2, \ldots\}$ with each $\boldsymbol{H}^{(L)}(t)$ in $span^N(\Delta_L)$ such that

$$\lim_{L\to+\infty} \boldsymbol{H}^{(L)}(t) = \overline{\boldsymbol{H}}^*(t), \quad \forall t \in \Omega.$$

Define a new problem:

$$\begin{aligned} &\min \quad y \\ &s.t.\ g(\boldsymbol{H}^{(L)}(t)) - y \le 0, \forall t \in \Omega. \end{aligned}$$

This problem is only a special case of Problem $P(L)$. Suppose that the optimal value of the problem above is y^{L*}. Then, by the optimality of Problem $P(L)$, we have $y^{L*} \ge y^{*L}$. Note that the constraint must be active at some $t \in \Omega$, it follows by the continuity of the function g that we have

$$\lim_{L\to+\infty} 0 = \lim_{L\to+\infty} \left(\max_{t\in\Omega} g(\boldsymbol{H}^{(L)}(t)) - y^{L*} \right) = \max_{t\in\Omega} g(\overline{\boldsymbol{H}}^*(t)) - \lim_{L\to+\infty} y^{L*}.$$

Hence,

$$\lim_{L\to+\infty} y^{L*} = \max_{t\in\Omega} y_t^*.$$

Since $y^{L*} \ge y^{*L}$, then

$$\lim_{L\to+\infty} y^{*L} \le \lim_{L\to+\infty} y^{L*} = \max_{t\in\Omega} y_t^*. \tag{3.6}$$

On the other hand, for each L, we have $\overline{\boldsymbol{H}}^{*L} \in span^N(\Delta_L) \subset C^N(\Omega)$. Note that the optimization of Problem $P(t)$ in the space $C^N(\Omega)$ is optimal, we have

$$y^{*L} \ge \max_{t\in\Omega} y_t^*.$$

Then, we obtain

$$\lim_{L\to+\infty} y^{*L} \ge \max_{t\in\Omega} y_t^*. \tag{3.7}$$

Thus, by (3.6) and (3.7), we have

$$\lim_{L\to+\infty} y^{*L} = \max_{t\in\Omega} y_t^*. \tag{3.8}$$

The proof completes. □

The importance of Theorem 3.6 is that we can have two ways to compute the limit of the optimal values. The first is to increase the parameter L, where a sufficiently large value should be chosen to estimate the limit. However, the number of variables is very large and the optimization is very expensive. The second is to solve the subproblem $P(t)$ for every t and find the maximum. Since the parameter L is ignored, the complexity of each subproblem has been greatly reduced and the limit can be obtained efficiently.

4 Numerical Experiments

In the section, the relation (3.8) can be verified by the following two examples, where the computations were implemented in Matlab.

Example 4.1. We consider the minimax optimization problem as follows:

$$\min_{\boldsymbol{x}\in\mathbb{R}^{NL}} \max_{t\in\Omega} \left|\boldsymbol{A}^T(t)\boldsymbol{H}(\boldsymbol{x},t) - G_d(t)\right|^2 + \sum_{k=1}^{N} H_k^2(\boldsymbol{x},t)$$
$$s.t. \begin{cases} A_k(t) = \cos((2k-1)t+2), k = 1,2,\cdots,N, \\ H_k(\mathbf{x},t) = \sum\limits_{i=1}^{L} x_{ki}t^{i-1}, k = 1,2,\cdots,N, \end{cases}$$

where $N = 4$ and

$$\Omega = \Omega_1 \cup \Omega_2,\ \Omega_1 = [0.1\pi, 0.3\pi],\ \Omega_2 = [0.5\pi, \pi],$$

$$G_d(t) = \begin{cases} 1, & t \in \Omega_1, \\ 0, & t \in \Omega_2. \end{cases}$$

This example can be considered as an interpolation problem, where the power functions are used as the basis functions and $H_k(\mathbf{x},t)$ is the polynomial function. It can be seen that this minimax optimization problem is equivalent to the semi-infinite programming problem as follow:

$$\min_{z,\boldsymbol{x}} z$$
$$s.t. \left|\boldsymbol{A}^T(t)\boldsymbol{H}(\boldsymbol{x},t) - G_d(t)\right|^2 + \sum_{k=1}^{N} H_k^2(\boldsymbol{x},t) \le z.$$

First, we choose the parameter L from 1 to 10 and solve the problem. The optimal values can be depicted in Figure 1. It can be seen that the optimal value series is monotonically decreasing as L increases.

Next, it's not difficult to verify that the conditions in Theorem 3.5 hold. Then, we can apply the proposed method to compute the limit of the optimal values. For this, we compute the optimal value by solving the subproblems for each t, then we take the maximum and obtain the limit, which can also be depicted in Figure 1. It is obvious that the optimal value series approaches to the limit as L tends to infinity. Hence, the relation (3.8) has been verified. Note that, the running time of computing the limit is only 0.5772 seconds. However, it costs 21.3721 seconds to obtain the maximum $z^* = 0.3401$ in the case that $L = 6$, which shows that the proposed method is efficient.

Example 4.2. In acoustic signal processing [10], the linear phase FIR filters are designed such that the actual response fits to a given ideal response. This problem can be transformed into a minimax optimization problem as follow:

$$\min_{\boldsymbol{x}\in\mathbb{R}^{NL}} \max_{f\in\Omega} \left|\boldsymbol{A}^T(f)\boldsymbol{H}(\boldsymbol{x},f) - G_d(f)\right|^2 + 2\sum_{k=1}^{N} H_k^2(\boldsymbol{x},f)$$
$$s.t. \begin{cases} A_k(f) = \cos((2k-1)f+2), k = 1,2,\cdots,N, \\ H_k(\mathbf{x},f) = \sum\limits_{i=1}^{L} x_{ki}\cos(if), k = 1,2,\cdots,N, \end{cases}$$

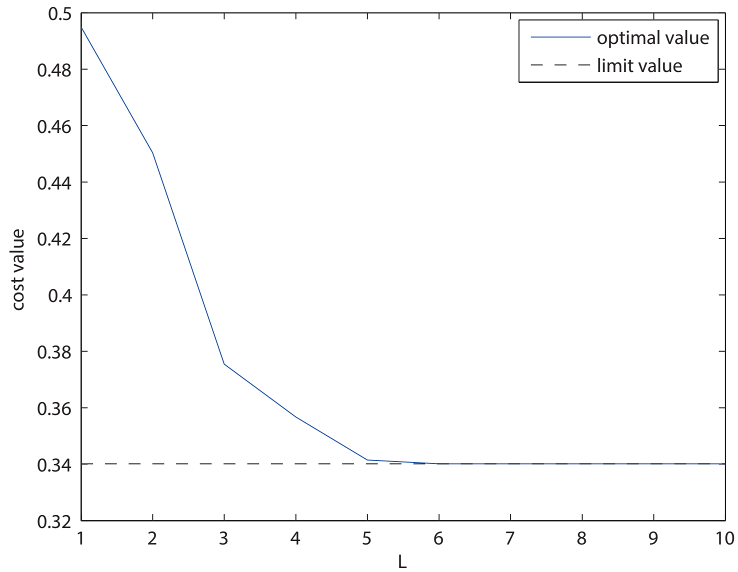

Figure 1: The optimal values with different parameter L in Example 1.

where $N = 5$ and

$$\Omega = \Omega_1 \cup \Omega_2,\ \Omega_1 = [0.05\pi, 0.5\pi],\ \Omega_2 = [0.6\pi, \pi],$$

$$G_d(f) = \begin{cases} 1, & f \in \Omega_1, \\ 0, & f \in \Omega_2. \end{cases}$$

The first term of the cost function is the error between the actual response function and the second term is the regularization to prevent the singularity of the problem. It can be seen that the cosine functions are used as the basis functions. This minimax optimization problem is translated into the semi-infinite programming problem as follow:

$$\begin{aligned} &\min_{z,\boldsymbol{x}} z \\ &s.t. \left| \boldsymbol{A}^T(f)\boldsymbol{H}(\boldsymbol{x}, f) - G_d(f) \right|^2 + 2\sum_{k=1}^{N} H_k^2(\boldsymbol{x}, f) \le z. \end{aligned}$$

To solve this problem, we first verify that the conditions in Theorem 3.5 hold. Then, we treat the function $\boldsymbol{H}(\boldsymbol{x}, f)$ as the variables first and obtain the limit as 0.4545. Next, we choose the parameter L from 1 to 12 and obtain the corresponding optimal values by solving the problems. The optimal values and the limit value can be depicted in Figure 2. It can be seen that the optimal value series is monotonically decreasing and approaches to the limit as L tends to infinity. Thus, the relation (3.8) is true and the proposed method is effective. Furthermore, we apply the proposed method to compute the limit with only 0.3276 seconds, while the running time to obtain the maximum is 51.7143 seconds in the case that $L = 8$. Hence, the proposed method is more efficient.

5 Conclusion

In this paper, we considered a class of minimax optimization problems with parameter L, which can be transformed into a series of equivalent semi-infinite programming problems

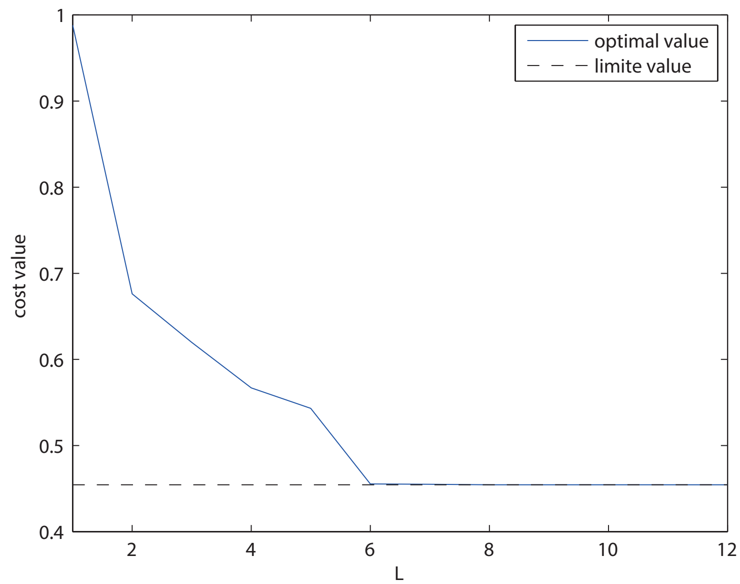

Figure 2: The cost function value with different parameter L in Example 2.

and can be solved by gradient based method. We analyze the limit of the optimal values as the parameter tends to infinity. To compute the limit, we proposed a novel method by decomposing the problem into a series of simplified subproblems. Then, we computed the maximum of these optimal values to obtain the limit. Furthermore, we deduced the condition and applied the fixed point theorem to support the theoretical basis of the proposed method. Finally, we verified the effectiveness and efficiency of the proposed method by numerical examples which show that the suitable L value is different with different basis functions.

References

[1] Z.G. Feng, K.L. Teo and V. Rehbock, A smoothing approach for semi-infinite programming with projected new-type algorithm, *J. Ind. Manag. Optim.* 5 (2009) 141–151.

[2] Z.G. Feng, K.F.C. Yiu and S.Y. Wu, Design of sparse filters by a discrete filled function technique, *Circuits Systems Signal Process.* 37 (2018) 4279–4294.

[3] R. Hettich and K.O. Kortanek, Semi-infinite programming: theory, methods, and application, *SIAM Review* 35 (1993) 380–429.

[4] C. Liu, Z. Gong, H.W.J. Lee and K.L. Teo, Robust bi-objective optimal control of 1,3-propanediol microbial batch production process, *J. Proc. Cont.* 78 (2019) 170–182.

[5] C. Liu, R. Loxton, Q. Lin and K.L. Teo, Dynamic optimization for switched time-delay systems with state-dependent switching conditions, *SIAM J. Control Optim.* 56 (2018) 3499–3523.

[6] M. Lopez and G. Still, Semi-infinite programming, *European J. Oper. Res.* 180 (2007) 491–518.

[7] A. Shapiro, Semi-infinite programming, duality, discretization and optimality conditions, *Optimization* 58 (2009) 133–161.

[8] T.J. Shiu and S.Y. Wu, Relaxed cutting plane method with convexification for solving nonlinear semi-infinite programming, *Comput. Optim. Appl.* 53 (2012) 91–113.

[9] S.X. Wang and Y.X. Yuan, Feasible method for convex semi-infinite programming, *SIAM J. Optim.* 25 (2015) 2537–2560.

[10] K.F.C. Yiu, X.Q. Yang, S. Nordholm and K.L. Teo, Near-field broadband beamformer design via multidimensional semi-infinite linear programming techniques, *IEEE Trans. Spee. Audio Proc.* 11 (2003) 725–732.

[11] L.P. Zhang, S.Y. Wu and M.A. Lopez, A new exchange method for convex semi-infinite programming, *SIAM J. Optim.* 20 (2010) 2959–2977.

Manuscript received 19 December 2018
revised 24 April 2019
accepted for publication 22 May 2019

Fei Chen
School of mathematics science, Chongqing normal university
Chongqing, P.R. China
E-mail address: 835938789@qq.com

Zhi Guo Feng
Faculty of mathematics and computer science, Guangdong ocean university
Zhanjiang, P.R. China
E-mail address: 18281102@qq.com

K.F.C. Yiu
Department of applied mathematics, Hongkong polytechnic university
Hongkong, P.R. China
E-mail address: macyiu@polyu.edu.hk