Article

Multiplication Algorithms for Approximate Optimal Distributions with Cost Constraints

Lianyan Fu [1,*,†] **, Faming Ma** [2,†] **, Zhuoxi Yu** [1,*] **and Zhichuan Zhu** [1]

1   School of Mathematics and Statistics, Liaoning University, Shenyang 110036, China
2   School of Economics, Liaoning University, Shenyang 110036, China; 4032130446@smail.lnu.edu.cn
*   Correspondence: fulianyan@lnu.edu.cn (L.F.); yuzhuoxi@lnu.edu.cn (Z.Y.)
†   These authors contributed equally to this work.

Abstract: In this paper, we study the D- and A-optimal assignment problems for regression models with experimental cost constraints. To solve these two problems, we propose two multiplicative algorithms for obtaining optimal designs and establishing extended D-optimal (ED-optimal) and A-optimal (EA-optimal) criteria. In addition, we give proof of the convergence of the ED-optimal algorithm and draw conjectures about some properties of the EA-optimal algorithm. Compared with the classical D- and A-optimal algorithms, the ED- and EA-optimal algorithms consider not only the accuracy of parameter estimation, but also the experimental cost constraint. The proposed methods work well in the digital example.

Keywords: D-optimal; A-optimal; experimental design; Kuhn–Tucker method; experimental cost

MSC: 62K05

1. Introduction

For a given regression model, one can develop different experimental protocols with various regression design methods. In order to compare the merits of various programs, many criteria have been developed. Due to their good statistical properties, the D- and A-optimal criteria (Kiefer [1]) are widely used in experimental design in various fields, such as in clinical investigations, biological experiments, agricultural experiments, automatic control, mechanical engineering, etc. Kiefer and Wolfowitz [2] made seminal contributions to the field, as they were the first to extend the concept of the discrete design of experiments to the measurement space of continuous design, thus facilitating mathematical research and the application of optimization theory. Elfving [3] and Dette [4] also gave the basic knowledge and details of optimal design theory, based on which Silvey et al. [5–7] further developed the related theory of optimal design. In general, optimal design is also known as optimal allocation in the literature [8]. With the extensive and in-depth development of optimal design, multiplication algorithms for optimal design under different optimization criteria were proposed. For instance, Wynn [9] first proposed an algorithm for D-optimal design, which is now known as the W-algorithm. Fedorov [10] proposed the V-algorithm and proved its convergence. However, the above algorithms suffered from slow convergence and overly cumbersome computations. Atwood [11] proposed three modifications to the V-algorithm, which improved its convergence speed and accuracy. John and Draper [12] further improved Atwood's first suggestion and increased the efficiency of the algorithm. Silvey et al. [13] presented a multiplicative algorithm for computing D-optimal designs on a finite design, and this was a straightforward iterative algorithm with monotonic convergence. Kiefer and Wolfowitz [2] introduced the general equivalence theorem and proposed an algorithm for solving optimization problems. Yu [14] suggested the cocktail method, which combined the well-known vertex direction method (VDM) with a multiplicative algorithm to compute D-optimal approximation designs. Martin-Martin et al. [15] applied

a multiplicative algorithm to construct marginal and conditional restricted designs when certain factors were known in advance, i.e., they were given and not controlled. Harman and Pronzato [16] proposed a class of multiplicative algorithms for computing *D*-optimal designs with non-optimal support points removed. Gao et al. [8] constructed a series of multiplicative algorithms for finding the *D*- and *A*-optimal designs of regression models. Castro et al. [17] introduced a new method aimed at computing approximate *D*- and *A*-optimal designs for multivariate polynomial regression on compact (semi-algebraic) design spaces. Harman et al. [18] constructed a series of randomized exchange algorithms (REX) to obtain the *D*- and *A*-optimal approximation designs for regression models. Duan et al. [19] proposed a simple multiplicative approach to generating sequential experimental designs while focusing on *D*- and *A*-optimal criteria. Duan et al. [20] described two multiplicative algorithms for obtaining approximate *D*- and *A*-optimal designs for multi-dimensional linear regression on a large variety of design spaces.

Although a great deal of research has been conducted on optimal design and the algorithms for achieving it, little consideration has been given to the cost of experimentation [21–25]. The experimental cost refers to the various raw materials consumed in an experiment [26–28], and it is an essential factor that is not negligible in experiments. First, to take this factor into account, we construct two weighted objective functions and establish extended *D*-optimal (*ED*-optimal) and *A*-optimal (*EA*-optimal) criteria to make the objective function optimal. For the *ED*-optimal algorithm, the objective function is the difference between the logarithm of the determinant of the information matrix in the classical *D*-optimal criterion and the weighted average experimental cost. A design $\hat{w}$ is referred to as an *ED*-optimal design if it maximizes the objection function in the class $w \in \{w_i \geq 0, \text{and} \sum_{i=1}^{k} w_i = 1\}$ of all designs. For the *EA*-optimal algorithm, the objective function is the sum between the logarithm of the trace of the inverse matrix of the information matrix in the classical *A*-optimal criterion and the weighted average experimental cost. A design $\hat{w}$ is referred to as an *EA*-optimal design if it minimizes the objection function in the class $w \in \{w_i \geq 0, \text{and} \sum_{i=1}^{k} w_i = 1\}$ of all designs. Compared with the classical *D*- and *A*-optimal designs, the extended *D*- and *A*-optimal designs strike a balance between maximizing the determinant of the information matrix and minimizing the experimental cost. Second, to solve the two optimization problems, we obtain the general equivalence conditions for *ED*- and *EA*-optimality with the Kuhn–Duck method. On this basis, we propose multiplicative algorithms for computing *D*- and *A*-optimal designs with cost constraints in the regression model and provide proof of the monotonic convergence of the *ED*-optimal algorithm. Although we do not give proof of the properties associated with the *EA*-optimal algorithm, the simulation results strongly support the validity and reliability of the algorithm. Finally, we give conclusions and further indicate possible research directions.

In this study, we extend the classical *D*- and *A*-optimal algorithms to *D*- and *A*-optimal algorithms with cost constraints. Compared with the classical *D*- and *A*-optimal algorithms, the *ED*- and *EA*-optimal algorithms consider not only the accuracy of parameter estimation, but also the experimental cost constraint. The organization of this article is as follows. Section 2 presents the model and the extended optimality criteria. Section 3 shows the multiplication algorithms for determining the *EA*- and *ED*-optimal assignments and describes the algorithm-related properties. Section 4 presents simulations that show the good performance of the proposed algorithms. Section 5 presents the conclusions. Appendix A contains proof of the main results.

2. Model and Optimal Criteria

2.1. Model and Symbols

Consider the generalized regression model

$$Y(x) = \beta' f(x) + \varepsilon, \quad x \in \mathcal{X}, \tag{1}$$

where $f(x)$ is a known regressor associated with x, β is a vector of parameters, $\mathcal{X}$ is the design space, and ε is an error term with a mean of 0 and variance of σ^2. For different independent trials, the errors are independent. The p-dimensional vector x is taken from the design space $\mathcal{X}$. The experiment aims to obtain the regression model by selecting k mutually independent design points $x_1, x_2, \cdots, x_k$ and estimating the unknown parameters β.

Suppose that the total number of trials is N, the numbers of trial points repeated at various levels are $n_1, n_2, \cdots, n_k$, and $n_1 + \cdots + n_k = N$. Further, we have $w_1, w_2, \cdots, w_k$ as the probability of occurrence of the design points $x_1, x_2, \cdots, x_k$ with $w_i = \frac{n_i}{N} (i = 1, 2, \cdots, k)$ and $\sum_{i=1}^{k} w_i = 1$. Therefore, the information array is denoted as $H(w) = \sum_{i=1}^{k} w_i x_i x_i'$.

2.2. Optimal Criteria

In the context of generalized regression models, the attainment of reduced variance in parameter estimates and a concomitant decrease in the cost of the experimental procedure typically results in enhanced regression model accuracy and increased economic efficiency of the experiment.

2.2.1. *ED*-Optimal Criterion

In the classical D-optimal design, a design $\hat{w}$ is referred to as optimal if it maximizes $\log|H(w)|$ ($|\cdot|$ is the determinant of the matrix) while satisfying the constraints $\{w_i \geq 0, \text{and} \sum_{i=1}^{k} w_i = 1\}$. However, in the classical D-optimal design, people only consider the accuracy of parameter estimation and do not take into account the cost of the experiment. Therefore, we propose an objective function in the form of the difference between the logarithm of the determinant of the information matrix in the classical D-optimal criterion and the weighted average experimental cost, i.e.,

$$T(w) = \log|H(w)| - \sum_{i=1}^{k} w_i c_i, \tag{2}$$

where $H(w)$ is the information matrix [8], and $\sum_{i=1}^{k} w_i c_i$ denotes the cost of the experiment [27]. An increase in the determinant of the information matrix is positively associated with a reduction in the volume of the Wald-type joint confidence region of the model parameters, thereby contributing to higher estimation accuracy. Furthermore, a decrease in trial cost is observed, leading to more efficient utilization of various raw materials, thereby enhancing the economic efficacy of the experiment. Formally, a design $\hat{w}$ is referred to as an *ED*-optimal design if it maximizes the objection function (2) in the class $w \in \{w_i \geq 0, \text{and} \sum_{i=1}^{k} w_i = 1\}$ of all designs. That is,

$$\hat{w} = \arg\max\{T(w) : \text{subject to } w_i \geq 0, \text{and} \sum_{i=1}^{k} w_i = 1\}. \tag{3}$$

Compared with the classical D-optimal design, the extended D-optimal design strikes a balance between maximizing the determinant of the information matrix and minimizing the experimental cost.

2.2.2. *EA*-Optimal Criterion

In the classical A-optimal design, a design $\hat{w}$ is referred to as optimal if it minimizes $\log(\text{trace}(H^{-1}(w)))$ while satisfying the constraints $\{w_i \geq 0, \text{and} \sum_{i=1}^{k} w_i = 1\}$. Similarly, to take the experimental cost into account, we give a new objective function whose form is the sum between the logarithm of the trace of the inverse matrix of the information matrix in the classical A-optimal criterion and the weighted average experimental cost, i.e.,

$$G(w) = \log(\text{trace}(H^{-1}(w))) + \sum_{i=1}^{k} w_i c_i. \tag{4}$$

where $\log(\text{trace}(\boldsymbol{H}^{-1}(\boldsymbol{w})))$ is the criterion function of the classical *A*-optimal criterion [8] and $\sum_{i=1}^{k} w_i c_i$ is the experimental cost [27]. If the information matrix is non-degenerate, then the trace of the inverse matrix of the information matrix reflects the magnitude of the sum of the variances of the components of the maximum likelihood estimate of the parameter $\boldsymbol{\beta}$. A smaller trace of the inverse matrix of the information matrix and experimental cost implies that the sum of the variance of the estimated components is also smaller and that the experiment is more economically efficient. Formally, a design $\hat{\boldsymbol{w}}$ is referred to as an *EA*-optimal design if it minimizes the objection function (4) in the class $\boldsymbol{w} \in \{w_i \geq 0, \text{and} \sum_{i=1}^{k} w_i = 1\}$ of all designs. That is,

$$\hat{\boldsymbol{w}} = \arg\min\{G(\boldsymbol{w}) : \text{subject to } w_i \geq 0, \text{and} \sum_{i=1}^{k} w_i = 1\}. \tag{5}$$

Compared with the classical *A*-optimal design, the extended *A*-optimal design strikes a balance between minimizing the trace of the inverse matrix of the information matrix and minimizing the experimental cost.

3. Algorithms for Optimal Distributions

In this section, we present the iterative algorithm in the *ED*-optimal criterion and give proof of the monotonic convergence of the algorithm. Meanwhile, we propose an iterative algorithm under the *EA*-optimal criterion and draw conjectures about the monotonic convergence of the algorithm. The corresponding proofs of these theorems and lemmas are in Appendix A.

3.1. Algorithm for ED-Optimal Distribution

To find the *ED*-optimal design, we obtain the general equivalence condition with the Kuhn–Duck method, and the result is shown in Theorem 1.

Theorem 1. *$\hat{\boldsymbol{w}}$ is an ED-optimal assignment if and only if it satisfies the following conditions:*

$$\boldsymbol{x_i}' \boldsymbol{H}^{-1}(\hat{\boldsymbol{w}}) \boldsymbol{x_i} + \sum_{i=1}^{k} \hat{w}_i c_i = p + c_i \quad for \quad \hat{w}_i \neq 0, \tag{6}$$

$$\boldsymbol{x_i}' \boldsymbol{H}^{-1}(\hat{\boldsymbol{w}}) \boldsymbol{x_i} + \sum_{i=1}^{k} \hat{w}_i c_i \leq p + c_i \quad for \quad \hat{w}_i = 0. \tag{7}$$

Note that Equation (6) in Theorem 1 is a special case of the classical approximate *D*-optimal equivalence theorem (Kiefer and Wolfowitz [29]). Based on this, we propose a multiplicative algorithm for calculating the *ED*-optimal allocation, as shown below.

In Algorithm 1, m is the number of iterations, k is the number of design points, and p is the number of experimental variables. In general, we set $k \geq p$, and the parameter ζ is usually set to a very small value, which determines the criteria for stopping the algorithm. In part 4, we set $\zeta = 0.0001$. Compared to the *W*- and *V*-algorithms, our proposed algorithm has the advantage of running quickly and simplifying the computational process because it involves only simple matrix operations. Furthermore, Algorithm 1 maintains its validity in instances where the sample size *N* is exceedingly large, as it solely relies on the weight $\boldsymbol{w}$ and design point $\boldsymbol{x}$. Contrastingly, the VDM algorithm's computational process is reliant not only on the sample size *N*, but also on the presence of interactions between design points, resulting in a decline in its efficacy. Our proposed algorithm is also a multiplicative algorithm, so it also has excellent properties, such as simplicity, efficiency, and monotonic convergence.

Algorithm 1 Algorithm for *ED*-optimal assignment

Input: Design space $\mathcal{X}$ and stopping parameter ζ.
Output: *ED*-optimal assignment $\hat{\boldsymbol{w}}$.

1: Set the initial value of $\boldsymbol{w}$ to $\boldsymbol{w}(0) = (\frac{1}{k}, \cdots, \frac{1}{k})$.

2: For $m \geq 1$ and $i = 1, \cdots, k$,

$$w_i(m) = w_i(m-1) \frac{\boldsymbol{x_i}'\mathbf{H}^{-1}(\boldsymbol{w}(m-1))\boldsymbol{x_i} + \sum\limits_{i=1}^{k} w_i(m-1)c_i}{p + c_i},$$

$$\boldsymbol{H}(\boldsymbol{w}(m-1)) = w_1(m-1)\boldsymbol{x}_1\boldsymbol{x}_1' + w_2(m-1)\boldsymbol{x}_2\boldsymbol{x}_2' + \cdots + w_k(m-1)\boldsymbol{x}_k\boldsymbol{x}_k'.$$

3: Repeat step 2 until $\max\{|w_i(m) - w_i(m-1)|\} < \zeta$.

4: Output *ED*-optimal assignment $\hat{\boldsymbol{w}}$.

Later in this section, we prove that the proposed algorithm converges to *ED*-optimality. To obtain the convergence of the proposed algorithm, we first demonstrate that the sequence $\{T(\boldsymbol{w}(m))\}$ is monotonically convergent, and the result is given in Theorem 2. Second, based on the theoretical results for the sequence $\{T(\boldsymbol{w}(m))\}$, we further certify that sequence $\{\boldsymbol{w}(m)\}$ converges to the *ED*-optimal allocation, and the conclusion is given in Theorem 3.

Theorem 2. $T(\boldsymbol{w}(m)) - T(\boldsymbol{w}(m-1)) = \log|\boldsymbol{H}(\boldsymbol{w}(m))| - \log|\boldsymbol{H}(\boldsymbol{w}(m-1))| \geq 0$ *and* $\boldsymbol{w}(m) - \boldsymbol{w}(m-1) \to 0$.

Theorem 2 shows that as the number of iterations increases, the sequence $\{T(\boldsymbol{w}(m))\}$ monotonically increases and the distance between sequences $\{\boldsymbol{w}(m)\}$ tends to zero. Note that Appendix A presents proof that the sequence $\{T(\boldsymbol{w}(m))\}$ converges. Based on the content of Theorem 2, we can also obtain Theorem 3.

Theorem 3. *The sequence* $\{\boldsymbol{w}(m)\}$ *obtained with the proposed algorithm converges to ED-optimality.*

In Theorem 3, we will first show that the sequence $\{\boldsymbol{w}(m)\}$ converges and subsequently prove its convergence to *ED*-optimality.

3.2. Algorithm for EA-Optimal Distribution

Similarly, according to the Kuhn–Duck method, we can obtain the general equivalence condition for *EA*-optimality, as shown in Theorem 4.

Theorem 4. $\hat{\boldsymbol{w}}$ *is an EA-optimal assignment if and only if it satisfies the following conditions:*

$$\frac{trace(\boldsymbol{H}^{-1}(\hat{\boldsymbol{w}})\boldsymbol{x_i}\boldsymbol{x_i}'\boldsymbol{H}^{-1}(\hat{\boldsymbol{w}}))}{trace(\boldsymbol{H}^{-1}(\hat{\boldsymbol{w}}))} + \sum_{i=1}^{k} \hat{w}_i c_i = 1 + c_i \quad for \quad \hat{w}_i \neq 0, \tag{8}$$

$$\frac{trace(\boldsymbol{H}^{-1}(\hat{\boldsymbol{w}})\boldsymbol{x_i}\boldsymbol{x_i}'\boldsymbol{H}^{-1}(\hat{\boldsymbol{w}}))}{trace(\boldsymbol{H}^{-1}(\hat{\boldsymbol{w}}))} + \sum_{i=1}^{k} \hat{w}_i c_i \leq 1 + c_i \quad for \quad \hat{w}_i = 0. \tag{9}$$

Based on Equation (8), we propose a multiplicative algorithm to calculate the *EA*-optimal design, and Algorithm 2 is shown below.

Algorithm 2 Algorithm for EA-optimal distribution

Input: Design space $\mathcal{X}$ and stopping parameter ζ.
Output: ED-optimal assignment $\hat{w}$.
1: Set the initial value of w to $w(0) = (\frac{1}{k}, \cdots, \frac{1}{k})$.
2: For $m \geq 1$ and $i = 1, \cdots, k$,

$$w_i(m) = w_i(m-1) \frac{\frac{\text{trace}(H^{-1}(w(m-1))x_i x'_i H^{-1}(w(m-1))}{\text{trace}(H^{-1}(w(m-1)))} + \sum_{i=1}^{k} w_i(m-1)c_i}{1+c_i},$$

$$H(w(m-1)) = w_1(m-1)x_1 x'_1 + w_2(m-1)x_2 x'_2 + \cdots + w_k(m-1)x_k x'_k.$$

3: Repeat step 2 until $\max\{|w_i(m) - w_i(m-1)|\} < \zeta$.
4: Output EA-optimal assignment $\hat{w}$.

Although the theoretical basis of the EA-optimal algorithm is similar to that of the ED-optimal algorithm, we do not obtain relevant conclusions similar to those of Theorem 2 under the EA-optimal criterion. In Section 3, we describe extensive simulation studies that were performed to illustrate the EA-optimal algorithm's properties. In all cases, the algorithm was convergent and successfully found the optimal allocation, so the algorithm is feasible.

4. Numerical Illustrations

In this section, we describe some simulations that were performed to illustrate some of the excellent properties of the recommended algorithms. We calculated the ED- and EA-optimal allocations, as well as the values of the corresponding objective functions. In addition, we provide the convergence rates and the numbers of iterations of the algorithms. The simulation results showed that the proposed algorithms ran quickly and were estimated accurately. The algorithms are easy to program and are suitable for a variety of situations.

Initially, we generated a set of design points, denoted by $x_1, x_2, \cdots, x_k$, which were mutually independent and followed a uniform distribution over the interval $[-1, 1]$, i.e., U$(-1, 1)$. Furthermore, we specified the initial values of weights, $w_1, w_2, \cdots, w_k$, and the corresponding experimental costs $c_1, c_2, \cdots, c_k$, which were also independently drawn from a uniform distribution over the range $[0, 1]$, i.e., U$(0, 1)$. By taking Table A1 as an illustrative example, subject to the constraints $p = 5$, $k = 8$, and $p < k$, we obtained the values of $x_1, x_2, \cdots, x_8$ and the associated trial costs $c_1, c_2, \cdots, c_8$. Without loss of generality, we set the initial weights $w_1 = 0.125, w_2 = 0.125, \ldots, w_8 = 0.125$.

Second, in accordance with the proposed Algorithm 1, the stopping criterion was set to $\max\{|w_i(n) - w_i(n-1)|\} < 0.0001$ for $i = 1, \cdots, k$, and we obtained the ED-optimal allocations and the values of the objective function. Table A1 shows the results of the simulation. In Table A2, keeping the value of k constant, by adjusting the number of experimental variables p, we obtained the corresponding computational results. As seen in Tables A1 and A2, the ED-optimal algorithm converged, and the sum of weights was 1 in each case, which meant that the algorithm could accurately estimate the optimal weights. Table A5 shows the average number of runs and the average time (in seconds) for 50 simulation cycles, with standard deviations in parentheses, indicating that the proposed ED-optimal algorithm converged quickly. When the value of k was fixed, the number of iterations and time decreased as p increased.

Finally, similarly to the above computational procedure, we found the EA-optimal assignments, the corresponding objective function values, and the average number of operations and time for 50 simulations. Taking Table A3 as an example, when $p = 5$ and $k = 8$, we first generated the corresponding design points, costs, and initial weights. According to Algorithm 2, we could obtain the EA-optimal weights and the objective function values. It can be seen from Table A3 that the proposed Algorithm 2 converged and the sum of the weights was 1 in each case, indicating the superiority of the EA-optimal algorithm

for parameter estimation. Moreover, when p was kept constant, the value of the objective function (4) became smaller as k increased. According to the EA-optimal criterion, the estimation of the parameters was more accurate. Thus, this also verified the conjecture that we mentioned above. As shown in Table A4, when k remained constant, the value of the objective function (4) became larger as p increased. The results in Table A6 show the average number of iterations and the average elapsed time (standard deviation in parentheses) for 50 simulations of the proposed Algorithm 2, indicating that the algorithm ran rapidly. In addition, when the value of p was kept constant, the number of iterations and time increased with k.

5. Conclusions

In this article, we studied the extended D- and A-optimal assignment problems in a regression model, i.e., the D- and A-optimal allocation issues with cost constraints. To obtain optimal designs, we extended the classical D- and A-optimal algorithms to D- and A-optimal algorithms with cost constraints. Compared with the classical D- and A-optimal designs, the extended D- and A-optimal designs strike a balance between maximizing the precision of parameter estimation and minimizing the experimental cost. For the ED-optimal algorithm, the objective function is the difference between the logarithm of the determinant of the information matrix in the classical D-optimal criterion and the weighted average experimental cost. We provided proof of the monotonicity and convergence of the algorithm. For the EA-optimal algorithm, the objective function is the sum between the logarithm of the trace of the inverse matrix of the information matrix in the classical A-optimal criterion and the weighted average experimental cost. Although there is no relevant proof in the literature concerning the nature of the algorithm, the simulation results show that the algorithm performs well. It is worth noting that while this study focuses on D- and A-optimal designs for linear models, the ideas of the proposed algorithms can be generalized to other standards and nonlinear models.

Author Contributions: Conceptualization, L.F. and F.M.; methodology, L.F.; software, F.M.; validation, F.M., Z.Y., and Z.Z.; formal analysis, L.F.; writing—original draft preparation, L.F., F.M., Z.Y., and Z.Z.; writing—review and editing, L.F., F.M., Z.Y., and Z.Z.; visualization, L.F., F.M., Z.Y., and Z.Z.; supervision, L.F., F.M., Z.Y., and Z.Z.; funding acquisition, L.F., F.M., Z.Y., and Z.Z. All authors have read and agreed to the published version of the manuscript.

Data Availability Statement: These algorithms were implemented in R, and the programs are available from the authors upon request.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study, in the collection, analyses, or interpretation of data, in the writing of the manuscript, or in the decision to publish the results.

Appendix A

Proof Theorem 1. According to the Kuhn–Tucker condition (Kuhn and Tucker [30]), $\hat{w}$ is the optimal solution of (2) if and only if

$$0 \geq \sum_{i=1}^{k} \frac{\partial T}{\partial w_i}(w_i - \hat{w}_i)$$

$$= \sum_{i=1}^{k} \left[\frac{\partial \log |H(w)|}{\partial w_i} - c_i \right](w_i - \hat{w}_i)$$

$$= \sum_{i=1}^{k} \left[\text{trace}(x_i x_i' H^{-1}(\hat{w})) - c_i \right](w_i - \hat{w}_i)$$

$$= \sum_{i=1}^{k} \left[x_i' H^{-1}(\hat{w}) x_i - c_i \right](w_i - \hat{w}_i)$$

$$= \sum_{i=1}^{k} \left[x_i' H^{-1}(\hat{w}) x_i - c_i \right] w_i - \sum_{i=1}^{k} \hat{w}_i \left[x_i' H^{-1}(\hat{w}) x_i - c_i \right]$$

$$= \sum_{i=1}^{k} \left[x_i' H^{-1}(\hat{w}) x_i - c_i \right] w_i - \sum_{i=1}^{k} \text{trace}\left(\hat{w}_i x_i x_i' H^{-1}(\hat{w}) \right) + \sum_{i=1}^{k} \hat{w}_i c_i$$

$$= \sum_{i=1}^{k} \left[x_i' H^{-1}(\hat{w}) x_i - c_i \right] w_i - \text{trace}\left(\sum_{i=1}^{k} \hat{w}_i x_i x_i' H^{-1}(\hat{w}) \right) + \sum_{i=1}^{k} \hat{w}_i c_i$$

$$= \sum_{i=1}^{k} \left[x_i' H^{-1}(\hat{w}) x_i - c_i \right] w_i - p + \sum_{i=1}^{k} \hat{w}_i c_i,$$

for all $w(w_i \geq 0, \text{and} \sum_{i=1}^{k} w_i = 1)$, which implies (6) and (7). $\square$

To prove the monotonicity of the sequence $\{T(w(m))\}$, we give Lemmas A1–A3.

Lemma A1. *For* $m \geq 1$, $w_1(m)c_1 + w_2(m)c_2 + \cdots + w_k(m)c_k = w_1(m-1)c_1 + w_2(m-1)c_2 + \cdots + w_k(m-1)c_k$.

Proof. For $m \geq 1$, $w_1(m)c_1 + w_2(m)c_2 + \cdots + w_k(m)c_k$,

$$= \sum_{i=1}^{k} w_i(m-1) x_i' H^{-1}(w(m-1)) x_i + \sum_{i=1}^{k} w_i(m-1)c_i - p$$

$$= \text{trace}\left(H(w(m-1)) H^{-1}(w(m-1)) \right) + \sum_{i=1}^{k} w_i(m-1)c_i - p$$

$$= w_1(m-1)c_1 + w_2(m-1)c_2 + \cdots + w_k(m-1)c_k.$$

According to step 3 in Algorithm 1, we can get

$$w_1(m)c_1 + w_2(m)c_2 + \cdots + w_k(m)c_k = \sum_{i=1}^{k} w_i(m-1) x_i' H^{-1}(w(m-1)) x_i$$

$$+ \sum_{i=1}^{k} w_i(m-1)c_i - p.$$

Further, we have

$$\sum_{i=1}^{k} w_i(m-1)\boldsymbol{x_i'}\boldsymbol{H}^{-1}(\boldsymbol{w}(m-1))\boldsymbol{x_i} = \sum_{i=1}^{k} w_i(m-1)\text{trace}\Big(\boldsymbol{x_i}\boldsymbol{x_i'}\boldsymbol{H}^{-1}(\boldsymbol{w}(m-1))\Big)$$

$$= \text{trace}\Big(\sum_{i=1}^{k} w_i(m-1)\boldsymbol{x_i}\boldsymbol{x_i'}\boldsymbol{H}^{-1}(\boldsymbol{w}(m-1))\Big) = \text{trace}\Big(\boldsymbol{H}(\boldsymbol{w}(m-1))\boldsymbol{H}^{-1}(\boldsymbol{w}(m-1))\Big).$$

□

Lemma A2. *Let* $\boldsymbol{H}(w_1, \cdot, w_k) = \sum_{i=1}^{k} w_i \boldsymbol{x_i}\boldsymbol{x_i'}$, *with* $w_i \geq 0$, $\boldsymbol{x_i}\boldsymbol{x_i'}$ *being a nonnegative definite matrix on* $\mathcal{X}$ *and* $|\boldsymbol{H}(w_1, \cdot, w_k)| \neq 0$; *thenm*

$$\log|\boldsymbol{H}(w_1^*, \cdot, w_k^*)| - \log|\boldsymbol{H}(w_1, \cdot, w_k)| \geq \sum_{i=1}^{k} w_i \, trace\Big(\boldsymbol{H}^{-1}(w_1, \cdots, w_k)\boldsymbol{x_i}\boldsymbol{x_i'}\Big)\log\frac{w_i^*}{w_i}$$

for $w_1^* \geq 0, \cdots, w_k^* \geq 0$.

Proof. See Gao et al. [8]. □

Lemma A3. *Suppose that* $\boldsymbol{w}^* = (w_1^*, \cdot, w_k^*)$ *and* $\boldsymbol{w} = (w_1, \cdot, w_k)$, *which satisfy* $\sum_{i=1}^{k} w_i^* = \sum_{i=1}^{k} w_i$, *are two probability vectors in* $\mathbb{R}^k$; *then,*

$$\sum_{i=1}^{k} |w_i^* - w_i| \leq \Big[2\sum_{i=1}^{k} w_i^* \log\frac{w_i^*}{w_i}\Big]^{\frac{1}{2}}.$$

Proof. See Kullback [31]. □

Proof Theorem 2.

$$T(\boldsymbol{w}(m)) - T(\boldsymbol{w}(m-1))$$

$$= \Big[\log|\boldsymbol{H}(\boldsymbol{w}(m))| - \sum_{i=1}^{k} w_i(m)c_i\Big] - \Big[\log|\boldsymbol{H}(\boldsymbol{w}(m-1))| - \sum_{i=1}^{k} w_i(m-1)c_i\Big]$$

$$= \log|\boldsymbol{H}(\boldsymbol{w}(m))| - \log|\boldsymbol{H}(\boldsymbol{w}(m-1))| \quad \text{(by Lemma 1)}$$

$$\geq \sum_{i=1}^{k} w_i(m-1)\boldsymbol{x_i'}\boldsymbol{H}^{-1}(\boldsymbol{w}(m-1))\boldsymbol{x_i}\log\frac{w_i(m)}{w_i(m-1)} \quad \text{(by Lemma 2)}$$

$$= \sum_{i=1}^{k} \Big[w_i(m)(p+c_i) - w_i(m-1)\sum_{i=1}^{k} w_i(m-1)c_i\Big]\log\frac{w_i(m)}{w_i(m-1)}$$

$$= \sum_{i=1}^{k} w_i(m)(p+c_i)\log\frac{w_i(m)}{w_i(m-1)} - s\sum_{i=1}^{k} w_i(m-1)\log\frac{w_i(m)}{w_i(m-1)}$$

$$= \sum_{i=1}^{k} t_i w_i(m)\log\frac{t_i w_i(m)}{t_i w_i(m-1)} + s\sum_{i=1}^{k} w_i(m-1)\log\frac{w_i(m-1)}{w_i(m)}$$

$$\geq \frac{1}{2}\Big[\sum_{i=1}^{k} t_i |w_i(m) - w_i(m-1)|\Big]^2 + \frac{s}{2}\Big[\sum_{i=1}^{k} |w_i(m) - w_i(m-1)|\Big]^2 \quad \text{(by Lemma 3)}$$

$$\geq \frac{1}{2}(\eta^2 + s)\Big[\sum_{i=1}^{k} |w_i(m) - w_i(m-1)|\Big]^2 \geq 0,$$

where $t_i = p + c_i$, s$= \sum_{i=1}^{k} w_i(m-1)c_i$ and $\eta = \min\{t_i\}$ for $i = 1, \cdots, k$. □

To obtain the convergence of the sequence $\{T(\boldsymbol{w}(m))\}$, we give Lemma A4.

Lemma A4. $\boldsymbol{H} \in \mathbb{R}^{p \times p}$ *is a nonnegative matrix; then,*

$$|\boldsymbol{H}| \leq \prod_{i=1}^{p} h_{ii},$$

where h_{ij} is the (i,j)-th element in $\boldsymbol{H}$.

Proof. See Anderson [32]. □

$$T(\boldsymbol{w}) = \log|\boldsymbol{H}(\boldsymbol{w})| - \sum_{i=1}^{k} w_i c_i \leq \log|\boldsymbol{H}(\boldsymbol{w})| \leq \sum_{i=1}^{p} \log|h_{ii}|,$$

that is, the sequence $\{T(\boldsymbol{w}(m))\}$ is uniformly bounded and monotonically increasing, so it is convergent. We have

$$0 = \lim_{m \to \infty} T(\boldsymbol{w}(m)) - T(\boldsymbol{w}(m-1)) = \lim_{m \to \infty} \log|\boldsymbol{H}(\boldsymbol{w}(m))| - \log|\boldsymbol{H}(\boldsymbol{w}(m-1))|$$

$$\geq \lim_{m \to \infty} \frac{1}{2}(\eta^2 + s)\left[\sum_{i=1}^{k} |w_i(m) - w_i(m-1)|\right]^2 \geq 0,$$

which implies $\boldsymbol{w}(m) - \boldsymbol{w}(m-1) \to 0$.

Proof Theorem 3. The sequence $\{\boldsymbol{w}(m)\}$ is convergent, and Gao et al. [8] gave the proof. Let $\hat{\boldsymbol{w}} = \lim_{m \to \infty} \boldsymbol{w}(m)$; then,

$$\hat{w}_i = \lim_{m \to \infty} w_i(m) = \lim_{m \to \infty} w_i(m-1) \frac{\boldsymbol{x}_i' \boldsymbol{H}^{-1}(\boldsymbol{w}(m-1))\boldsymbol{x}_i + \sum\limits_{i=1}^{k} w_i(m-1)c_i}{p + c_i}$$

$$= \hat{w}_i \frac{\boldsymbol{x}_i' \boldsymbol{H}^{-1}(\hat{\boldsymbol{w}})\boldsymbol{x}_i + \sum\limits_{i=1}^{k} \hat{w}_i c_i}{p + c_k},$$

which implies

$$\boldsymbol{x}_i' \boldsymbol{H}^{-1}(\hat{\boldsymbol{w}})\boldsymbol{x}_i + \sum_{i=1}^{k} \hat{w}_i c_i = p + c_i \quad for \quad \hat{w}_i \neq 0$$

and

$$\boldsymbol{x}_i' \boldsymbol{H}^{-1}(\hat{\boldsymbol{w}})\boldsymbol{x}_i + \sum_{i=1}^{k} \hat{w}_i c_i \leq p + c_i \quad for \quad \hat{w}_i = 0.$$

□

So, the sequence $\{\boldsymbol{w}(m)\}$ converges to *ED*-optimality.

Proof Theorem 4. According to the Kuhn–Tucker condition (Kuhn and Tucker [30]), $\hat{\boldsymbol{w}}$ is the optimal solution of (4) if and only if

$$0 \leq \sum_{i=1}^{k} \frac{\partial G}{\partial w_i}(w_i - \hat{w}_i)$$

$$= \sum_{i=1}^{k} \left[\frac{\partial \log(\text{trace}(\boldsymbol{H}^{-1}(\boldsymbol{w})))}{\partial w_i} + c_i\right](w_i - \hat{w}_i)$$

$$= -\sum_{i=1}^{k} \left[\frac{\text{trace}(\boldsymbol{H}^{-1}(\hat{\boldsymbol{w}})\boldsymbol{x}_i\boldsymbol{x}_i'\boldsymbol{H}^{-1}(\hat{\boldsymbol{w}}))}{\text{trace}(\boldsymbol{H}^{-1}(\hat{\boldsymbol{w}}))} - c_i \right] w_i$$

$$+ \frac{1}{\text{trace}(\boldsymbol{H}^{-1}(\hat{\boldsymbol{w}}))} \sum_{i=1}^{k} \text{trace}(\boldsymbol{H}^{-1}(\hat{\boldsymbol{w}})\hat{w}_i\boldsymbol{x}_i\boldsymbol{x}_i'\boldsymbol{H}^{-1}(\hat{\boldsymbol{w}})) - \sum_{i=1}^{k} \hat{w}_i c_i$$

$$= -\sum_{i=1}^{k} \left[\frac{\text{trace}(\boldsymbol{H}^{-1}(\hat{\boldsymbol{w}})\boldsymbol{x}_i\boldsymbol{x}_i'\boldsymbol{H}^{-1}(\hat{\boldsymbol{w}}))}{\text{trace}(\boldsymbol{H}^{-1}(\hat{\boldsymbol{w}}))} - c_i \right] w_i$$

$$+ \frac{1}{\text{trace}(\boldsymbol{H}^{-1}(\hat{\boldsymbol{w}}))} \text{trace}\left(\boldsymbol{H}^{-1}(\hat{\boldsymbol{w}}) \left(\sum_{i=1}^{k} \hat{w}_i\boldsymbol{x}_i\boldsymbol{x}_i' \right) \boldsymbol{H}^{-1}(\hat{\boldsymbol{w}}) \right) - \sum_{i=1}^{k} \hat{w}_i c_i$$

$$= -\sum_{i=1}^{k} \left[\frac{\text{trace}(\boldsymbol{H}^{-1}(\hat{\boldsymbol{w}})\boldsymbol{x}_i\boldsymbol{x}_i'\boldsymbol{H}^{-1}(\hat{\boldsymbol{w}}))}{\text{trace}(\boldsymbol{H}^{-1}(\hat{\boldsymbol{w}}))} - c_i \right] w_i + 1 - \sum_{i=1}^{k} \hat{w}_i c_i$$

for all $\boldsymbol{w}(w_i \geq 0, \text{and } \sum_{i=1}^{k} w_i = 1)$, which implies (8) and (9).
□

Table A1. Design points, costs, *ED*-optimal weights, and values of objective function (2) when *p* = 5.

Design Points (k = 8)	Costs (k = 8)	Weights (k = 8)	Design Points (k = 12)	Costs (k = 12)	Weights (k = 12)
(0.92,0.99,−0.77,−0.20,−0.03)	0.97	0.0831	(0.60,0.46,0.98,−0.07,0.48)	0.32	0.0039
(−0.56,0.58,0.58,0.05,−0.07)	0.30	0.1428	(−0.86,0.04,−0.34,0.89,0.44)	0.61	0.1021
(0.45,−0.64,−0.46,0.34,−0.95)	0.43	0.1486	(−0.42,−0.76,0.99,−0.57,0.80)	0.84	0.1537
(−0.75,−0.96,−0.52,0.29,−0.25)	0.40	0.1300	(0.76,0.56,0.86,−0.57,0.75)	0.66	0.1753
(0.75,0.28,0.99,−0.29,0.78)	0.38	0.0929	(0.26,0.32,−0.79,0.78,0.17)	0.18	0.0052
(0.87,0.32,−0.32,0.99,−0.89)	0.43	0.1814	(0.39,−0.48,0.55,−0.24,0.29)	0.76	0.0000
(−0.54,−0.68,−0.05,0.38,−0.98)	0.86	0.0813	(0.95,−0.99,−0.45,0.11,−0.73)	0.45	0.1670
(−0.64,0.88,−0.69,−0.69,0.11)	0.65	0.1399	(−0.55,−0.26,−0.19,−0.95,0.89)	0.21	0.2027
			(−0.35,−0.36,−0.28,0.33,0.85)	0.47	0.0714
			(−0.19,−0.48,0.64,−0.90,0.09)	0.59	0.0000
			(0.77,0.04,−0.92,0.95,−0.31)	0.32	0.1186
			(−0.75,−0.06,0.05,−0.82,0.03)	0.84	0.0000
T(*w*)	−7.2778				−5.8847

When *p* = 5, *k* = 8, or *k* = 12, we first generated the design points, costs, and initial weights. According to Algorithm 1, we could obtain the *ED*-optimal weights and the corresponding objective function values.

Table A2. Design points, costs, *ED*-optimal weights, and values of objective function (2) when *k* = 10.

Design Points (p = 3)	Costs (p = 3)	Weights (p = 3)	Design Points (p = 6)	Costs (p = 6)	Weights (p = 6)
(−0.86,0.31,0.75)	0.16	0.0002	(0.23,−0.30,−0.35,−0.98,−0.59,−0.16)	0.47	0.1556
(−0.44,−0.44,−0.65)	0.04	0.0000	(−0.54,−0.22,−0.45,0.15,0.30,0.35)	0.97	0.1362
(0.17,−0.94,0.80)	0.04	0.0927	(0.60,−0.49,−0.94,0.21,−0.35,−0.31)	0.76	0.0851
(−0.43,−0.87,0.70)	0.55	0.1759	(−0.82,1.00,0.04,0.67,−0.66,0.91)	0.83	0.1298
(−0.80,0.96,−0.68)	0.86	0.1610	(−0.71,0.61,−0.67,0.65,0.14,0.85)	0.74	0.0815
(−0.69,0.30,−0.73)	0.27	0.0000	(0.45,−0.70,−0.70,0.69,−0.59,−0.54)	0.36	0.1175
(0.26,0.24,−0.29)	0.28	0.0000	(−0.22,−0.67,0.88,0.50,−0.39,0.60)	0.52	0.1643
(0.21,−0.15,−0.80)	0.36	0.0000	(−0.33,0.25,−0.17,0.49,0.48,−0.04)	0.44	0.0001
(0.36,0.62,0.81)	0.10	0.3208	(0.29,−0.41,−0.21,−0.92,0.30,0.23)	0.41	0.1296
(−0.93,0.63,0.82)	0.61	0.2494	(0.41,−0.93,0.01,−0.36,−0.39,−0.16)	0.69	0.0001
T(*w*)	−2.508				−10.2531

When *k* = 5, *p* = 3, or *p* = 6, we first generated the design points, costs, and initial weights. According to Algorithm 1, we could obtain the *ED*-optimal weights and the corresponding objective function values.

Table A3. Design points, costs, *EA*-optimal weights, and values of objective function (4) when *p* = 5.

Design Points (*k* = 8)	Costs (*k* = 8)	Weights (*k* = 8)	Design Points (*k* = 12)	Costs (*k* = 12)	Weights (*k* = 12)
(−0.99,−0.95,0.64,−0.47,0.56)	0.40	0.0020	(0.14,0.45,0.11,0.63,0.00)	0.02	0.0000
(−0.42,−0.54,0.19,0.47,−0.05)	0.92	0.0000	(−0.12,−1.00,0.24,0.31,0.75)	0.27	0.0000
(0.01,−0.13,0.39,0.00,0.60)	0.05	0.3003	(−0.03,0.85,−0.56,−0.84,−0.63)	0.96	0.0046
(−0.88,−0.25,−0.77,0.48,0.94)	0.05	0.1719	(0.07,−0.75,0.52,−0.26,0.55)	0.44	0.0000
(0.51,0.46,−0.85,0.98,0.20)	0.49	0.1488	(0.45,−0.17,0.59,−0.90,0.07)	0.24	0.0017
(−0.68,−0.19,0.47,0.10,−0.65)	0.83	0.1320	(−0.83,0.80,0.17,0.27,0.21)	0.61	0.0496
(0.81,0.97,0.01,−0.77,−0.81)	0.23	0.0909	(−0.88,−0.30,0.94,−0.15,0.33)	0.73	0.1503
(0.67,−0.78,0.10,0.04,−0.20)	0.42	0.1540	(−0.53,−0.91,−0.20,−0.69,−0.20)	0.01	0.2335
			(0.96,0.66,0.87,−0.72,−0.27)	0.63	0.0338
			(0.85,0.38,0.62,0.56,−0.90)	0.40	0.1730
			(−0.51,0.41,−0.60,0.51,−0.80)	0.07	0.1868
			(−0.40,0.94,0.25,−0.91,−0.73)	0.03	0.1665
$G(\hat{w})$	3.7949				3.0554

When *p* = 5, *k* = 8, or *k* = 12, we first generated the design points, costs, and initial weights. According to Algorithm 2, we could obtain the *EA*-optimal weights and the corresponding objective function values.

Table A4. Design points, costs, *EA*-optimal weights, and values of objective function (4) when *k* = 10.

Design Points (*p* = 3)	Costs (*p* = 3)	Weights (*p* = 3)	Design Points (*p* = 6)	Costs (*p* = 6)	Weights (*p* = 6)
(−0.80,−0.97,−0.26)	0.22	0.2756	(−0.86,−0.88,−0.47,0.76,−0.99,0.90)	0.30	0.1015
(0.14,−0.67,−0.50)	0.86	0.0000	(0.00,0.47,−0.65,0.47,0.95,0.99)	0.02	0.1511
(0.19,0.66,−0.21)	0.81	0.0000	(0.72,0.33,0.91,−0.35,0.44,0.59)	0.34	0.0932
(0.46,−0.30,0.28)	0.98	0.0000	(0.77,0.99,−0.14,0.10,−0.87,0.46)	0.29	0.1367
(0.32,−0.28,−0.02)	0.58	0.0000	(0.29,−0.70,−0.31,−0.45,−0.18,0.35)	0.24	0.2404
(1.00,−0.54,−0.74)	0.20	0.3158	(−0.33,0.78,−0.39,−0.25,0.26,−0.24)	0.75	0.0468
(0.59,0.83,0.57)	0.15	0.0676	(−0.29,−0.01,−0.64,−0.04,−0.77,0.71)	0.58	0.0000
(0.32,−0.08,0.77)	0.33	0.0007	(−0.32,0.80,−0.51,−0.17,−0.47,−0.63)	0.55	0.0756
(0.77,−0.05,−0.17)	0.00	0.0000	(0.75,0.83,−0.32,0.59,−0.22,0.97)	0.70	0.0015
(0.17,−0.51,0.83)	0.70	0.3403	(−0.34,0.28,0.34,−0.54,−0.47,0.41)	0.96	0.1533
$G(\hat{w})$	2.2659				3.6571

When *k* = 10, *p* = 3, or *p* = 6, we first generated the design points, costs, and initial weights. According to Algorithm 2, we could obtain the *EA*-optimal weights and the corresponding objective function values.

Table A5. The average number of iterations and average elapsed time of the *ED*-optimal algorithm.

k	*p*	Average No. of Iterations (s.d.)	Average Elapsed Time in Sec. (s.d.)
10	4	76.6(46.0)	0.024(0.015)
	5	52.7(27.8)	0.019(0.011)
	8	21.6(11.1)	0.008(0.009)
20	4	93.8(55.5)	0.079(0.054)
	5	71.0(28.6)	0.062(0.033)
	8	43.4(9.0)	0.038(0.011)
	10	35.1(7.2)	0.031(0.011)
	15	18.3(6.8)	0.017(0.010)
30	4	99.8(58.5)	0.127(0.079)
	5	70.5(23.0)	0.096(0.037)
	8	53.7(13.6)	0.065(0.015)
	10	39.1(6.5)	0.050(0.011)
	15	27.4(3.0)	0.040(0.011)
	20	18.6(3.4)	0.033(0.010)
	25	9.0(4.1)	0.022(0.013)

Table A5. *Cont.*

k	p	Average No. of Iterations (s.d.)	Average Elapsed Time in Sec. (s.d.)
40	4	104.7(53.2)	0.179(0.092)
	5	84.5(34.7)	0.137(0.060)
	8	60.5(16.4)	0.107(0.028)
	10	45.3(7.3)	0.078(0.014)
	15	31.0(2.7)	0.062(0.010)
	20	23.2(1.5)	0.056(0.009)
	25	18.3(2.8)	0.050(0.010)
	30	10.7(3.6)	0.038(0.013)

The table gives the average number of iterations and the average running time in seconds for 50 simulations of the proposed Algorithm 1, with the standard deviation in parentheses.

Table A6. The average number of iterations and average elapsed time of the EA-optimal algorithm.

k	p	Average No. of Iterations (s.d.)	Average Elapsed Time in Sec. (s.d.)
10	4	52.4(24.2)	0.052(0.031)
	5	47.0(22.2)	0.044(0.022)
	8	17.7(7.5)	0.019(0.012)
20	4	91.2(52.6)	0.144(0.079)
	5	81.3(41.3)	0.154(0.065)
	8	42.2(10.5)	0.088(0.026)
	10	30.5(7.5)	0.062(0.017)
	15	15.2(6.3)	0.030(0.014)
30	4	94.6(48.6)	0.270(0.145)
	5	86.8(30.8)	0.252(0.102)
	8	55.4(15.0)	0.164(0.041)
	10	40.8(7.5)	0.124(0.029)
	15	24.6(4.2)	0.084(0.018)
	20	14.3(3.8)	0.064(0.019)
	25	9.0(1.9)	0.046(0.014)
40	4	112.7(52.4)	0.425(0.186)
	5	89.8(38.2)	0.379(0.143)
	8	56.5(12.1)	0.231(0.061)
	10	46.1(8.7)	0.19(0.04)
	15	27.8(4.0)	0.135(0.021)
	20	19.3(2.6)	0.112(0.018)
	25	12.6(3.1)	0.093(0.023)
	30	8.7(2.1)	0.082(0.021)

The table gives the average number of iterations and the average running time in seconds for 50 simulations of the proposed Algorithm 2, with the standard deviation in parentheses.

References

1. Kiefer, J. General equivalence theory for optimum designs (approximate theory). *Ann. Stat.* **1974**, *2*, 849–879. [CrossRef]
2. Kiefer, J.; Wolfowitz, J. Optimum designs in regression problems. *Ann. Math. Stat.* **1959**, *30*, 271–294. [CrossRef]
3. Elfving, G. Optimum allocation in linear regression theory. *Ann. Math. Stat.* **1952**, *23*, 255–262. [CrossRef]
4. Dette, H.; Studden, W.J. Canonical Moments and Optimal Design-First Application. In *The Theory of Canonical Moments with Applications in Statistics, Probability, and Analysis*; John Wiley & Sons: New York, NY, USA, 1997; pp. 128–159.
5. Silvey, D. *Optimal Design*; Chapman and Hall: New York, NY, USA, 1980.
6. Box, G.E.P.; Draper, N.R. *Empirical Model-building and Response Surfaces*; John Wiley & Sons: New York, NY, USA, 1987; pp. 229–231.
7. Gilmour, S.G.; Trinca, L.A. Optimum design of experiments for statistical inference. *J. R. Stat Soc. Ser. C Appl. Stat.* **2012**, *61*, 345–401. [CrossRef]
8. Gao, W.; Chan, P.S.; Ng, H.K.T.; Lu, X. Efficient computational algorithm for optimal allocation in regression models. *J. Comput. Appl. Math.* **2014**, *261*, 118–126. [CrossRef]
9. Wynn, H.P. The sequential generation of D-optimum experimental designs. *Ann. Math. Stat.* **1970**, *41*, 1655–1664. [CrossRef]
10. Fedorov, V.V. Continuous Optimal Designs (Statistical Methods). In *Theory of Optimal Experiments*; Academic Press: New York, NY, USA, 1972; pp. 64–153.

11. Atwood, C.L. Sequences converging to D-optimal designs of experiments. *Ann. Stat.* **1973**, *1*, 342–352. [CrossRef]
12. John, R.C.S.; Draper, N.R. D-optimality for regression designs: A review. *Technometrics* **1975**, *17*, 15–23. [CrossRef]
13. Silvey, S.D.; Titterington, D.H.; Torsney, B. An algorithm for optimal designs on a design space. *Commun. Stat. Theory Methods* **1978**, *7*, 1379–1389. [CrossRef]
14. Yu, Y. D-optimal designs via a cocktail algorithm. *Stat. Comput.* **2011**, *21*, 475–481. [CrossRef]
15. Martin-Martin, R.; Torsney, B.; López-Fidalgo, J. Construction of marginally and conditionally restricted designs using multiplicative algorithms. *Comput. Stat. Data Anal.* **2007**, *51*, 5547–5561. [CrossRef]
16. Harman, R.; Pronzato, L. Improvements on removing nonoptimal support points in D-optimum design algorithms. *Stat. Probab. Lett.* **2007**, *77*, 90–94. [CrossRef]
17. De Castro, Y.; Gamboa, F.; Henrion, D.; Hess, R.; Lasserre, J.-B. Approximate optimal designs for multivariate polynomial regression. *Ann. Stat.* **2008**, *47*, 127–155. [CrossRef]
18. Harman, R.; Filová, L.; Richtárik, P. A randomized exchange algorithm for computing optimal approximate designs of experiments. *J. Am. Stat. Assoc.* **2020**, *115*, 348–361. [CrossRef]
19. Duan, J.; Gao, W.; Ng, H.K.T. Efficient computational algorithm for optimal continuous experimental designs. *J. Comput. Appl. Math.* **2019**, *350*, 98–113. [CrossRef]
20. Duan, J.; Gao, W.; Ma, Y.; Ng, H.K.T. Efficient computational algorithms for approximate optimal designs. *J. Stat. Comput. Simul.* **2022**, *92*, 764–793. [CrossRef]
21. Yang, M.; Biedermann, S.; Tang, E. On optimal designs for nonlinear models: A general and efficient algorithm. *J. Am. Stat. Assoc.* **2013**, *108*, 1411–1420. [CrossRef]
22. Jones, B.; Allen-Moyer, K.; Goos, P. A-optimal versus D-optimal design of screening experiments. *J. Qual. Technol.* **2021**, *53*, 369–382. [CrossRef]
23. Torsney, B.; Martín-Martín, R. Multiplicative algorithms for computing optimum designs. *J. Stat. Plan Inference* **2009**, *139*, 3947–3961. [CrossRef]
24. Yu, Y. Monotonic convergence of a general algorithm for computing optimal designs. *Ann. Stat.* **2010**, *38*, 1593–1606. [CrossRef]
25. Goudarzi, M.; Khazaei, S.; Jafari, H. D-optimal designs for linear mixed model with random effects of Dirichlet process. *Commun. Stat. Simul. Comput.* **2021**, 1–10. [CrossRef]
26. Harman, R.; Bachratá, A.; Filová, L. Construction of efficient experimental designs under multiple resource constraints. *Appl. Stoch. Models Bus. Ind.* **2016**, *32*, 3–17. [CrossRef]
27. Harman, R.; Benková, E. Barycentric algorithm for computing d-optimal size-and cost-constrained designs of experiments. *Metrika* **2017**, *80*, 201–225. [CrossRef]
28. Coetzer, R.; Haines, L.M. The construction of D- and I-optimal designs for mixture experiments with linear constraints on the components. *Chemometr. Intell. Lab. Syst.* **2017**, *171*, 112–124. [CrossRef]
29. Kiefer, J.; Wolfowitz, J. The equivalence of two extremum problems. *Can. J. Math.* **1960**, *12*, 363–366. [CrossRef]
30. Khun, H.W.; Tucker, A.W. *Non-Linear Programming, Proceeding Second Berkeley Symposium Mathematical Statistic and Probability (ed) Nyman*; University of California Press: Berkeley, CA, USA, 1951; pp. 481–492.
31. Kullback, S. A lower bound for discrimination information in terms of variation (corresp.). *IEEE Trans. Inf. Theory* **1967**, *13*, 126–127. [CrossRef]
32. Anderson, T.W. *An Introduction to Multivariate Statistical Analysis*, 3rd ed.; John Wiley and Sons: New York, NY, USA, 2003.