arXiv:2303.04746v1 [stat.ME] 8 Mar 2023

多目标最优回归设计的充分必要条件

Lucy L. Gao°*、Jane J. Ye‡、Shangzhi Zeng‡、Julie Zhou‡

° 统计系，英属哥伦比亚大学

‡ 数学与统计系，维多利亚大学

2023年3月9日

摘要

我们通常基于单个目标函数构造最优设计。为了更全面地反映实验目标，我们可以改为基于多个目标函数构造多目标最优设计。虽然已有算法用于求解多目标最优设计（例如，效率约束设计和最大最小最优设计），但如何验证算法所得解的最优性则远不明朗。本文针对离散设计空间上的效率约束设计和最大最小最优设计，给出了刻画其最优性的理论结果。我们展示了如何将这些结果与线性规划算法结合使用，以验证最优性。

关键词：最优性条件、效率、最大最小设计、线性规划、凸优化、稳健性

*作者按字母顺序排列。通讯作者：lucy.gao@stat.ubc.ca

1 引言

考虑将设计实验的输出建模为:

$$y_i = f(\mathbf{x}_i, \boldsymbol{\theta}) + \epsilon_i, \quad i = 1, \ldots, n, \tag{1}$$

其中， y_i 是设计点$\mathbf{x}_i \in S$ 处观测到的响应变量， $S \subseteq \mathbb{R}^p$时，$\boldsymbol{\theta} \in \mathbb{R}^q$ 是未知回归参数的向量，而 ϵ_i是相互独立的随机误差，满足 $\mathbb{E}[\epsilon_i] = 0$ 和 $\text{Var}(\epsilon_i) = \sigma^2$。最优设计通过选择 $\mathbf{x}_i$ 的取值，尽可能精确地回答相关实验问题。这个问题通常表述为单目标 最优设计问题，其中最优性是相对于通过将单个模型拟合到实验数据而获得的信息的某个汇总度量来定义的。例如，对于式 (1) 中特定选定的回归函数 $f(\cdot, \cdot)$ 和估计量 $\hat{\boldsymbol{\theta}}$，A-最优设计 $\hat{\theta}_1, \ldots, \hat{\theta}_q$ 可使 $\theta_1, \ldots, \theta_q$ 的平均方差最小。

然而，实验者的目标很复杂，无法用单目标最优设计准则充分概括。例如，实验者可能会拟合单个模型来进行推断和预测，而衡量推断能力与预测能力的单目标最优性准则之间几乎没有交集。此外，针对不同参数进行推断会回答不同的研究问题，而这些问题对实验者的重要程度各不相同（例如，主效应可能比交互项更重要，反之亦然）。单目标最优设计准则无法体现这种差异。还需考虑的是，实验者可能不确定 y_i 与$\mathbf{x}_i$之间关系的函数形式。因此，他们可能希望获得一种设计，使其对于式 (1) 形式的多个模型都具有良好的推断或预测能力，而非仅适用于式 (1) 形式的单个模型。

多目标最优设计结合了多种单目标最优设计准则。常见的形式包括复合形式，即针对一组用户指定的权重优化各准则的加权和；效率约束形式，即优化一个准则，同时要求设计相对于其他准则的效率高于用户指定的值；以及最大最小形式，即最大化一组最优性准则中的最小效率（Wong 1999，Wong & Zhou 2022）。我们重点关注效率约束和最大最小形式，因为复合形式中的权重实际意义难以解释。

许多论文研究了寻找效率约束和最大最小最优设计的算法（Huang & Wong 1998，Imhof & Wong 2000，Cheng & Yang 2019）。Wong & Zhou (2022) 提出了一种特别灵活的算法：他们将许多效率约束和最大最小最优设计问题表述为凸优化问题，然后使用现成的凸优化求解器（cvx, Grant & Boyd 2014）。

本文沿着 Wong & Zhou (2022) 的思路，将效率约束问题和最大最小问题表述为凸优化问题。随后，我们探讨如何验证通过 cvx 得到的效率约束最优设计或最大最小最优设计的最优性。我们完整刻画了离散设计空间上效率约束设计和最大最小效率设计的最优性。文献中已有关于效率约束最优设计的相关结果（例如，Cook & Wong 1994年和 Clyde & Chaloner 1996年）。我们对极小极大效率设计最优性的刻画似乎是新的，不过关于极小极大和最大最小单目标优化问题已有相关研究（例如，M¨uller & P´azman 1998年和 Dette 等 2007年）。

许多常用单目标最优设计准则（例如 D- 和 A-）的最优性刻画要求最优设计满足一组易于计算的不等式。相比之下，我们对效率约束设计和最大最小设计的最优性刻画提出，存在一组量，它们满足涉及最优设计的一组等式和不等式。人们认为，这类结果不适用于实际的最优性验证，因为尚不清楚如何高效地找到一组合适的量。以往关于效率约束最优设计问题的研究通过网格搜索和二分搜索寻找合适的一组量，但计算复杂度会随目标函数的数量呈指数级增长（Cheng & Yang 2019年）。值得注意的是，我们成功克服了这一挑战：事实证明，通过求解线性规划问题（Luenberger & Ye 2016），即可找到一组合适的量。线性规划问题是数学优化的基石，现成的软件可用于准确、高效地求解这类问题。

本文其余部分安排如下。第2节回顾与单目标最优性准则相关的概念，包括最优性的必要和充分条件。第3节和第4节介绍如何求解受效率约束的最优设计和最大最小最优设计，以及如何验证所得解的最优性。我们还推导了若干理论结果

。第5节将我们的方法应用于若干示例。结论见第6节。证明见附录。

2 单目标最优设计

我们考虑离散设计空间 $S_N = \{\mathrm{u}1、...、\mathrm{u}_N\} \subseteq S$，其中包含 N个设计点；u1、...、u_N和 S由用户指定。如果 S是连续设计空间，则 S_N 近似于 S。我们用 $\xi(\mathrm{w}) = \begin{pmatrix} \mathrm{u}_1 & \mathrm{u}_2 & \cdots & \mathrm{u}_N \\ w_1 & w_2 & \cdots & w_N \end{pmatrix}$ u u 表示 S_N 上的设计 ξ，其中 w 是一个 N-维向量，其第 i个元素 w_i 表示取值为 u_i的设计点所占的比例，其中 $i = 1、2、...、N$。令 $\Omega \equiv \left\{ \mathrm{w} \in \mathbb{R}^N : \sum_{i=1}^{N} w_i = 1、\ w_i \geq 0 \right\}$。

2.1 最优性准则

令$\mathbf{z}^f(\mathbf{x})$为一个 q-维向量，其第 j个元素为 $\frac{\partial f(\mathbf{x}\ \boldsymbol{\theta}\ )}{\partial \theta_j}\Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}^*}$，其中 $\boldsymbol{\theta}^*$ 是 $\boldsymbol{\theta}$ 的真实值。在回归函数为 $f(\cdot,\cdot)$ 的模型（1）中，设计为 $\xi(\mathbf{w})$ 时，$\boldsymbol{\theta}$ 的普通最小二乘估计量的渐近协方差矩阵与 $\mathcal{I}_f^{-1}(\mathbf{w})$ 成正比，其中

$$\mathcal{I}_f(\mathbf{w}) = \sum_{i=1}^{N} w_i \mathbf{z}_f(\mathbf{u}_i)\mathbf{z}_f^T(\mathbf{u}_i) \tag{2}$$

是在正态分布误差假设下，回归函数为$f(\cdot,\ \cdot)$的模型 (1) 的期望信息矩阵。如果 $f(\mathbf{x}\ \boldsymbol{\theta})$关于 $\boldsymbol{\theta}$是非线性的，那么 $\mathcal{I}_f(\mathbf{w})$可能取决于 $\boldsymbol{\theta}^*$。如果 $\mathcal{I}_f(\mathbf{w})$取决于 $\boldsymbol{\theta}^*$，那么优化涉及 $\mathcal{I}_f(\mathbf{w})$的设计准则会得到局部最优设计。在实践中，$\boldsymbol{\theta}^*$ 通常是未知的，因此我们必须用对其取值的“猜测”来代替，例如根据一个小规模先导研究得到的 $\boldsymbol{\theta}$ 估计值。

许多关于 S_N 的单目标最优设计准则可以转化为如下形式的凸优化问题：$\min_{\mathbf{w}\in\Omega} \Phi(\mathbf{w})$，其中 $\Phi(\mathbf{w}) = \phi(\mathcal{I}_f(\mathbf{w}))$，$\phi$ 是在所有正定矩阵集合上定义的凸函数；例如，参见表1。请注意，函数 $\Phi(\mathbf{w})$是凸函数与线性函数复合而成的凸函数。我们通过设计w 相对于最优设计的效率来衡量其质量，该效率记为 Eff(w)。

表1：求解 $\min_{\mathbf{w}\in\Omega}$ $\Phi(\mathbf{w})$ 的单目标最优性准则，其中$\Phi(\mathbf{w}) = \phi(\mathcal{I}_f(\mathbf{w}))$ 是定义在所有正定矩阵集合上的凸函数。我们用 $\lambda_{min}(\mathbf{M})$ 表示$\mathbf{M}$的最小特征值。

准则	D-	A-	c-, 当 $\mathbf{c}\in\mathbb{R}^q$ 时	L-, 当 $\mathbf{L}\in\mathbb{R}^{q\times q}$ 时 ′	E-
$\phi(\mathbf{M})$	–log det(M)	迹（$\mathbf{M}^{-1}$）	$\mathbf{c}^T\mathbf{M}^{-1}\mathbf{c}$	迹($\mathbf{L}^T\mathbf{M}^{-1}\mathbf{L}$)	$-\lambda_{min}(\mathbf{M})$
效率(w)	$\left(\frac{\exp\{\min_{\mathbf{w}'\in\Omega}\Phi(\mathbf{w}')\}}{\exp\{\Phi(\mathbf{w})\}}\right)^{1/q}$	$\frac{\min_{\mathbf{w}'\in\Omega}\Phi(\mathbf{w}')}{\Phi(\mathbf{w})}$	$\frac{\min_{\mathbf{w}'\in\Omega}\Phi(\mathbf{w}')}{\Phi(\mathbf{w})}$	$\frac{\min_{\mathbf{w}'\in\Omega}\Phi(\mathbf{w}')}{\Phi(\mathbf{w})}$	$\frac{\Phi(\mathbf{w})}{\min_{\mathbf{w}'\in\Omega}\Phi(\mathbf{w}')}$

基于 MATLAB 的软件包 CVX (Grant & Boyd 2014年)是一种用户友好的工具，可用于求解一类特殊的凸优化问题，其中包括表1所述的凸优化问题；有关 CVX 的更多详情，请参见第3.1节。此前， CVX 软件包已用于求解许多单目标最优设计问题；例如，参见 Gao & Zhou (2017年) 和 Wong & Zhou (2019年)。

2.2 最优性的必要充分条件

表1中的所有准则都会导出凸目标函数，但这些目标函数并非处处可微。例如，E-最优设计准则会导出一个凸目标函数；当设计$\mathbf{w}$ 使得 $\mathcal{I}_f(\mathbf{w})$的最小特征值的几何重数大于1时，该函数不可微。因此，这种情况下的最优性条件依赖于次微分，它将导数的概念推广到凸函数类。我们将凸函数 Φ: $\mathbb{R}^N \to \mathbb{R}$ 在点 $\mathbf{w}$ 处的次微分记为 $\partial\Phi(\mathbf{w})$。以下结果介绍了次微分的基本性质（第2章，Mordukhovich & Nam，2013年）。

引理1. 设 Φ 和 Φ' 是定义在 $\mathbb{R}^N$上的两个取有限值的凸函数。那么，对于任意$\mathbf{w}\in\mathbb{R}^N$:

1. 如果 Φ 在 w 处可微，则 $\partial\Phi(w) = \{\nabla\Phi(w)\}$，其中 $\nabla\Phi(v)$ 是 N -向量，其 i 个分量为 $\frac{\partial\Phi}{\partial w_i}\big|_{\mathbf{w}=\mathbf{v}}$。*2.* 如果 $a \geq 0$，则 $\partial(a\Phi)(w) = a\partial\Phi(w) \equiv \{ag\colon\ g \in \partial\Phi(w)\}$。*3.* $\partial(\Phi+\Phi')(w) = \partial\Phi(w) + \partial\Phi'(w) \equiv \{g + g'\colon\ g \in \partial\Phi(w),\ g' \in \partial\Phi'(w)\}$.

令 $\mathbf{e}_i$ 表示一个 N-向量，其 i 个分量等于 1，所有其他分量均等于 0。以下结果刻画了离散设计空间上凸单目标最优设计准则的最优性。

定理 1. 假设 Φ: $\mathbb{R}^N \to \mathbb{R}$ 是一个凸函数。设 $\mathbf{w}^* \in \Omega$。那么，$\mathbf{w}^* \in \operatorname{argmin}_{\mathbf{w}\in\Omega} \Phi(\mathbf{w})$ 当且仅当

$$\exists\, \mathbf{g} \in \partial\Phi(\mathbf{w}^*) \; such\ that\ \mathbf{g}^T(\mathbf{w}^* - \mathbf{e}_i) \le 0 \; for\ all\ i = 1, 2, \ldots, N. \quad (3)$$

本小节接下来的内容将介绍一些结果，以帮助我们在特殊情形下评估条件 (3)：对于某个凸函数 ϕ，$\Phi(\mathbf{w}) = \phi(\mathcal{I}_f(\mathbf{w}))$。首先，若 ϕ 是凸函数，且在 $\mathcal{I}_f(\mathbf{w}^*)$ 处可微，则引理1表明$\partial\Phi(\mathbf{w}^*) = \{\nabla\Phi(\mathbf{w}^*)\}$，因此条件 (3) 可简化为：

$$\left[\nabla\Phi(\mathbf{w}^*)\right]^T (\mathbf{w}^* - \mathbf{e}_i) \le 0 \text{ for all } i = 1, 2, \ldots, N. \quad (4)$$

以下结果由矩阵链式法则（Petersen & Pedersen 2012 的第2.8.1节）得出，并刻画了 (4) 的左侧。

引理 2. 设 $\mathbf{w}^* \in \Omega$。若对于凸函数 ϕ 和回归函数 $f(\cdot,\cdot)$，$\Phi(\mathbf{w}) = \phi(\mathcal{I}_f(\mathbf{w}))$，且 ϕ 在 $\mathcal{I}_f(\mathbf{w}^*)$ 处可微，则：

$$\left[\nabla\Phi(\mathbf{w}^*)\right]^T (\mathbf{w}^* - \mathbf{e}_i) = d_{\phi,f}(\mathbf{u}_i, \mathbf{w}^*), \quad for\ all\ i = 1, 2, \ldots, N,$$

其中，对于所有 $i = 1$、2、.....、N，我们定义

$$d_{\phi,f}(\mathbf{u}_i, \mathbf{w}^*) \equiv trace\left(\left[\nabla\phi(\mathcal{I}_f(\mathbf{w}^*))\right]^T \left[\mathcal{I}_f(\mathbf{w}^*) - \mathbf{z}_f(\mathbf{u}_i)\mathbf{z}_f^T(\mathbf{u}_i)\right]\right), \quad (5)$$

并且，$\nabla\phi(\mathbf{M}^*)$ 是一个 $q \times q$ 矩阵，其$(j,\ j')th$ 项为 $\frac{\partial\phi}{\partial M_{jj'}}\Big|_{\mathbf{M}=\mathbf{M}^*}$。

由定理 1 和引理 2 可知，对于可微凸函数 ϕ，若 $\Phi(\mathbf{w}) = \phi(\mathcal{I}_f(\mathbf{w}))$，则刻画单目标最优性准则的最优性，只需检查对所有$i = 1$，2、.....、$N$，$d_{\phi,\ f}(\mathbf{u}_i,\ \mathbf{w}^*) \le 0$是否成立。此外，给定 $\nabla\phi(\mathbf{M}^*)$ 的公式后，$d_{\phi,\ f}(\mathbf{u}_i,\ \mathbf{w}^*)$ 很容易计算；参见表 2。

表 2：对于表1中的所有可微凸函数 ϕ，求解 $\min_{\mathbf{w}\in\Omega} \phi(\mathbf{A}_f(\mathbf{w}))$ 的最优性准则。$\nabla\phi(\mathbf{M})$ 的公式来自 Petersen & Pedersen（2012）。

Criteria	D-	A-	c-, for $\mathbf{c} \in \mathbb{R}^q$	L-, for $\mathbf{L} \in \mathbb{R}^{q\times q'}$
$\phi(\mathbf{M})$	$-\log\det(\mathbf{M})$	$\operatorname{trace}(\mathbf{M}^{-1})$	$\mathbf{c}^T\mathbf{M}^{-1}\mathbf{c}$	$\operatorname{trace}(\mathbf{L}^T\mathbf{M}^{-1}\mathbf{L})$
$\nabla\phi(\mathbf{M})$	$-\mathbf{M}^{-1}$	$-\mathbf{M}^{-2}$	$-\mathbf{M}^{-1}\mathbf{c}\mathbf{c}^T\mathbf{M}^{-1}$	$-\mathbf{M}^{-1}\mathbf{L}\mathbf{L}^T\mathbf{M}^{-1}$

结合定理 1、引理 2 和表1，可以得到离散设计空间上的各种经典等价定理；例如参见 Kiefer（1974）。

对于 E-最优性，$\Phi(\mathbf{w}) = \phi(\mathcal{I}_f(\mathbf{w}))$ 是一个不可微凸函数 ϕ (表1)，因此引理 2 并非总是适用。我们通过以下结果来解决这个问题。

引理 3. 设 $\Phi(\mathbf{w}) = -\lambda_{min}(\mathcal{I}_f(\mathbf{w}))$ 是一个回归函数$f(\cdot,\cdot)$。设 $\mathbf{w}^* \in \Omega$，且设 r^*为 $\lambda_{min}(\mathcal{I}_f(\mathbf{w}^*))$的几何重数。

1. 若 $r^* = 1$，则 $\partial\Phi(\mathbf{w}^*) = \{\nabla\Phi(\mathbf{w}^*)\}$，并且

$$[\nabla\Phi(\mathbf{w}^*)]^T(\mathbf{w}^* - \mathbf{e}_i) = [(\mathbf{v}^*)^T\mathbf{z}_f(\mathbf{u}_i)]^2 - \lambda_{min}(\mathcal{I}_f(\mathbf{w}^*)),\ \text{对所有}\ i = 1, 2, \ldots\ldots, N,$$

其中$\mathbf{v}^*$ 表示与 $\lambda_{min}(\mathcal{I}_f(\mathbf{w}^*))$相对应的任意单位特征向量。

2. 如果 $r^* > 1$，那么对于任意 $\mathbf{g} \in \partial\Phi(\mathbf{w}^*)$，存在 a_1、…、$a_{r^*} \geq 0$，使得 $\sum_{j=1}^{r^*} a_j = 1$，并且

$$\mathbf{g}^T(\mathbf{w}^* - \mathbf{e}_i) = d_{-\lambda_{min},f,\mathbf{a}}(\mathbf{u}_i, \mathbf{w}^*),\ \textit{for all}\ i = 1, 2, \ldots, N,$$

其中，对于所有 $i = 1$、2、…、N，我们定义

$$d_{-\lambda_{min},f,\mathbf{a}}(\mathbf{u}_i, \mathbf{w}^*) \equiv \sum_{j=1}^{r^*} a_j[(\mathbf{v}_j^*)^T\mathbf{z}_f(\mathbf{u}_i)]^2 - \lambda_{min}(\mathcal{I}_f(\mathbf{w}^*)), \qquad (6)$$

其中，$\mathbf{v}_1^*$、…、$\mathbf{v}_{r^*}^*$ 表示与 $\lambda_{min}(\mathcal{I}_f(\mathbf{w}^*))$ 对应的一组任意标准正交特征向量。

结合定理 1 与引理 3，可得离散设计空间上 E-最优性的等价定理（Kiefer，1974）。

3 效率约束下的最优设计

假设实验者主要关注优化某个特定的单目标最优性准则 Φ_1，同时又不希望其他准则 $\Phi_2, \ldots, \Phi_K$的效率损失过大，其中 $K \geq 2$。令 $\mathrm{Eff}_k(\mathbf{w})$ 表示设计 $\xi(\mathbf{w})$相对于准则 Φ_k的效率，其中

$k = 1, \ldots, K$。给定由实验者指定的常数 $m_2, \ldots, m_K \in (0, 1)$，定义在 S_N 上的效率约束最优设计 满足：

$$\left\{ \begin{array}{ll} \min\limits_{\mathbf{w} \in \mathbb{R}^{\mathbf{N}}} & \Phi_1(\mathbf{w}) \\ \text{subject to:} & \text{Eff}_k(\mathbf{w}) \geq m_k, \ k = 2, \ldots, K, \\ & \sum\limits_{i=1}^{N} w_i = 1, w_i \geq 0, \ i = 1, 2, \ldots, N. \end{array} \right\}. \tag{7}$$

当期望的最小效率 $m_2, \ldots, m_K$ 较大时，所有设计都可能无法满足（7）中的约束。

3.1 凸优化问题

假设对于所有 $k = 1, \ldots, K,$， $\Phi_k(\mathbf{w}) = \phi_k(\mathcal{I}_{f_k}(\mathbf{w}))$，其中 $\phi_1, \ldots, \phi_K$ 是从表1中选取的连续凸函数，且（2）中的 $\mathcal{I}_{f_k}(\mathbf{w})$ 是模型（1）的期望信息矩阵，该模型的回归函数为 $f_k(\mathbf{x}\ \boldsymbol{\theta})$，对应设计 $\xi(\mathbf{w})$ 下的 $\boldsymbol{\theta} \in \mathbb{R}^{q_k}$ 。我们利用表1中 $\text{Eff}_k(\mathbf{w})$ 的定义，将（7）改写为以下问题：

$$\left\{ \begin{array}{ll} \min\limits_{\mathbf{w} \in \mathbb{R}^{\mathbf{N}}} & \Phi_1(\mathbf{w}) \\ \text{subject to:} & \Phi_k(\mathbf{w}) \leq h_k(m_k), \ k = 2, \ldots, K, \\ & \sum\limits_{i=1}^{N} w_i = 1, \ w_i \geq 0, \ i = 1, 2, \ldots, N. \end{array} \right\}, \tag{8}$$

其中，我们将 $h_k(m)$ 定义为

$$h_k(m) = \begin{cases} \left(\min\limits_{\mathbf{w}' \in \Omega} \Phi_k(\mathbf{w}') \right) - q_k \log(m), & \text{if } \Phi_k(\mathbf{w}) = -\log\det(\mathcal{I}_{f_k}(\mathbf{w})), \\ m \left(\min\limits_{\mathbf{w}' \in \Omega} \Phi_k(\mathbf{w}') \right), & \text{if } \Phi_k(\mathbf{w}) = -\lambda_{min}(\mathcal{I}_{f_k}(\mathbf{w})), \\ \frac{1}{m} \left(\min\limits_{\mathbf{w}' \in \Omega} \Phi_k(\mathbf{w}') \right), & \text{otherwise.} \end{cases} \tag{9}$$

对于所有 $k = 1$，2、.....， K, $\Phi_k(\mathbf{w}) = \phi_k(\mathcal{I}_{f_k}(\mathbf{w}))$ 都是凸函数，因为ϕ_k是凸函数，而 $\mathcal{I}_{f_k}$ 是线性函数。因此，(8) 是一个凸优化问题。请注意，我们的表述不同于 Wong and Zhou（2022），因为对于 D-最优性，我们使用 $\Phi_k(\mathbf{w}) = -\log\det(\mathcal{I}_{f_k}(\mathbf{w}))$，而不是 $\Phi_k(\mathbf{w}) = [\det(\mathcal{I}_{f_k}(\mathbf{w})]^{-1/q_k}$。

事实上，(8) 是一个凸优化问题，可以通过 `CVX`(Grant & Boyd 2014年) 求解。它是一个基于 MATLAB 的程序包，可用于一类特殊的优化问题；关于这一类问题的详细信息，参见 Grant &Boyd (2008)。 `CVX` 会自动将 (8) 转换为数值凸优化求解器（例如 SDPT3 或 SeDuMi）可求解的形式，然后将数值结果转换回原始形式。

3.2 必要充分条件

在最小效率不等式约束可以被严格满足的假设下，以下结果刻画了问题 (8) 的最优性。

定理 2. 假设存在 $\mathbf{w} \in \Omega$ 满足 $Eff_k(\mathbf{w}) > m_k$，对所有 $k = 2, \ldots, K$ 均成立。设 $\mathbf{w}^*$ 是问题 (8) 的一个可行解。那么，$\mathbf{w}^*$是问题 *(8)* 的解，当且仅当存在 $\eta_2, \ldots, \eta_K \geq 0$ 使得

1. $\eta_k \left(\Phi_k(\mathbf{w}^*) - h_k(m_k)\right) = 0$ *for all* $k = 2, \ldots, K$, *and*
2. $\mathbf{w}^* \in \arg\min_{\mathbf{w} \in \Omega} \left[\Phi_1(\mathbf{w}) + \sum_{k=2}^{K} \eta_k \Phi_k(\mathbf{w})\right]$.

定理 2 与 Lee (1988)、Cook & Wong (1994) 以及 Clyde & Chaloner (1996) 中的结果相关。

现在我们将讨论如何利用第 2.2 节中的结果改写定理 2。首先，假设 $\phi_1, \ldots, \phi_K$ 均对应于 D-、 A-、 c- 或 L-最优性。于是，根据定理 1 和引理 2，我们可以将定理 2 中的条件 2 替换为：

$$d_{\phi_1, f_1}(\mathbf{u}_i, \mathbf{w}^*) + \sum_{k=2}^{K} \eta_k d_{\phi_K, f_K}(\mathbf{u}_i, \mathbf{w}^*) \leq 0, \text{ for all } i = 1, 2, \ldots, N, \qquad (10)$$

对于式 (5) 中定义的 $d_{\phi,\ f}(\mathbf{u}_i,\ \mathbf{w}^*)$。表 2 给出了 $\nabla\phi_k(\mathbf{M})$ 的公式。

否则，我们必须结合引理1和引理 3 应用定理 1，以改写定理 2 中的条件 2。下面通过一个例子加以说明。

例 1. 假设 $\phi_1(\mathbf{M}) = -\lambda_{min}(\mathbf{M})$，且 $\phi_2, \ldots, \phi_K$ 均对应于 D-、 A-、 c-或 L-最优性。于是，由定理 1 和引理1可知，定理 2 中的条件 2 成立，当且仅当存在$\mathbf{g} \in \partial\Phi_1(\mathbf{w}^*)$使得

$$\mathbf{g}^T \left(\mathbf{w}^* - \mathbf{e}_i\right) + \sum_{k=2}^{K} \eta_k d_{\phi_k, f_k}(\mathbf{u}_i, \mathbf{w}^*) \leq 0 \text{ for all } i = 1, 2, \ldots, N, \qquad (11)$$

对于 (5) 中定义的 $d_{\phi,f}(\mathbf{u}_i, \mathbf{w}^*)$。令 r^* 为 $\lambda_{min}(\mathcal{I}_{f_1}(\mathbf{w}^*))$ 的几何重数。由引理 3 还可得：

- 情况1（$r^* = 1$）：定理 2 中的条件 2 成立，当且仅当

$$[(\mathbf{v}^*)^T \mathbf{z}_{f_1}(\mathbf{u}_i, \mathbf{w}^*)]^2 - \lambda_{min}(\mathcal{I}_{f_1}(\mathbf{w}^*)) + \sum_{k=2}^{K} \eta_k d_{\phi_k, f_k}(\mathbf{u}_i, \mathbf{w}^*) \leq 0, \forall\, i = 1, \ldots, N, \tag{12}$$

  其中，$\mathbf{v}^*$ 表示 $\lambda_{min}(\mathcal{I}_{f_1}(\mathbf{w}^*))$的任意单位特征向量。

- 情况2（$r^* > 1$）：定理 2 中的条件 2 成立，当且仅当存在 $a_1, \ldots, a_{r^*} \geq 0$ 使得 $\sum_{j=1}^{r^*} a_j = 1$ 并且

$$d_{-\lambda_{min}, f_1, \mathbf{a}}(\mathbf{u}_i, \mathbf{w}^*) + \sum_{k=2}^{K} \eta_k d_{\phi_k, f_k}(\mathbf{u}_i, \mathbf{w}^*) \leq 0,\ \forall\, i = 1, \ldots, N, \tag{13}$$

  其中，$d_{-\lambda_{min}, f, \mathbf{a}}(\mathbf{u}_i, \mathbf{w}^*)$定义见（6）。

3.3 通过线性规划验证最优性

假设我们通过数值求解（8）得到了$\mathbf{w}^*$，其中 $\phi_1, \ldots, \phi_K$分别对应 D-、 A-、 c-或 L-最优性。根据定理 2，如果能找到 $\eta_2, \ldots, \eta_K \geq 0$ 使得满足定理 2 中的条件 1–2，我们就知道$\mathbf{w}^*$ 是最优的。然而，$\mathbf{w}^*$ 是一个近似数值解，因此不太可能精确满足条件 1–2。我们转而通过寻找 $\eta_2, \ldots, \eta_K \geq 0$ 来检查$\mathbf{w}^*$ 是否“足够接近”最优，使得：

$$\eta_k \left(\Phi_k(\mathbf{w}^*) - h_k(m_k)\right) \leq \delta, \quad k = 2, \ldots, K, \tag{14}$$

$$-\eta_k \left(\Phi_k(\mathbf{w}^*) - h_k(m_k)\right) \leq \delta, \quad k = 2, \ldots, K, \tag{15}$$

$$d_{\phi_1, f_1}(\mathbf{u}_i, \mathbf{w}^*) + \sum_{k=2}^{K} \eta_k d_{\phi_k, f_k}(\mathbf{u}_i, \mathbf{w}^*) \leq \delta, \quad i = 1, 2, \ldots, N, \tag{16}$$

其中，δ 是一个较小的正常数（例如 $\delta = 10^{-4}$）。这里，(14)–(15) 放宽了定理 2 中的条件 1，(16) 放宽了定理 2 中的条件 2，而 δ控制着我们对“足够接近”最优的定义。类似思想也见于单目标最优设计（Wong & Zhou 2019年）。

我们提出求解以下优化问题:

$$\left\{ \begin{array}{ll} \min\limits_{\boldsymbol{\eta}\in\mathbb{R}^{K-1}} & \mathbf{1}_{K-1}^T\boldsymbol{\eta} \\ \text{subject to: } \boldsymbol{\eta}\geq \mathbf{0}_{K-1}, & \mathbf{B}_1^T\boldsymbol{\eta}\leq \mathbf{b}_1, \;\; \mathbf{C}_1\boldsymbol{\eta}\leq \delta\mathbf{1}_{K-1}, -\mathbf{C}_1\boldsymbol{\eta}\leq \delta\mathbf{1}_{K-1}. \end{array} \right\}, \quad (17)$$

其中，$\leq$和 $\geq$表示逐分量不等式，$\mathbf{1}_{K-1}$是每个元素均等于 1 的（$K-1$）-向量，$\mathbf{B}_1$是一个（$K-1$）$\times N$矩阵，其第（k,i）个元素为$d_{\phi_{k+1}f_{k+1}}(\mathbf{u}_i, \mathbf{w}^*)$；其中，对于 $k=1,2,\ldots,K-1$，$d_{\phi f}(\mathbf{u}_i, \mathbf{w}^*)$定义于式 (5) 中；$\mathbf{b}_1$是一个 N-向量，其第 i 个元素等于 $\delta - d_{\phi_1 f_1}(\mathbf{u}_i, \mathbf{w}^*)$，且 $\mathbf{C}_1 = \text{diag}(\Phi_2(\mathbf{w}^*) - h_2(m_2)\ldots, \Phi_K(\mathbf{w}^*) - h_K(m_K))$。

如果我们能够找到满足 (17) 的解 $\boldsymbol{\eta}^*$，那么就知道 $\mathbf{w}^*$ 和 $\boldsymbol{\eta}^*$ 共同满足 (14)–(16)。这意味着定理 2 中的条件（近似地）成立，因此 $\mathbf{w}^*$ 确实是 (8) 的最优解。此外，(17) 是一个线性规划问题 （Luenberger & Ye 2016）：其目标函数和约束均为线性的。因此，我们只需使用现成的线性规划求解器（例如 Optimization Toolbox 中的 `linprog` 函数）即可求解 (17)，该工具箱属于 `MATLAB`。

如果 ϕ_1、.....、ϕ_K 中有一个或多个对应于 E-最优性，那么下面的例子将说明，我们仍然可以通过线性规划验证定理 2 中的条件。

例 1（续）。考虑第 3.2 节中的例 1，其中 $\phi_1(\mathbf{M}) = -\lambda_{min}(\mathbf{M})$，且 ϕ_2、.....、ϕ_K 均对应于 D-、 A-、 c- 或 L-最优性。回想一下，我们将 r^* 定义为 $\lambda_{min}(\mathcal{I}_{f_1}(\mathbf{w}^*))$ 的几何重数。我们先前已经证明，若 $r^*=1$，则定理 3 中的条件 2 等价于 (12)，而 (12) 定义了 η_2、.....、η_K 中的一组 N 线性等式。因此，$_K$ 我们可以在满足 (14)–(15) 中定义的 $2(K-1)$ 个线性不等式以及以下 (12) 的放宽形式的条件下，最小化 $\sum \eta_k$，

$$[(\mathbf{v}^*)^T\mathbf{z}_{f_1}(\mathbf{u}_i,\mathbf{w}^*)]^2 - \lambda_{min}(\mathcal{I}_{f_1}(\mathbf{w}^*)) + \sum_{k=2}^{K}\eta_k d_{\phi_k,f_k}(\mathbf{u}_i,\mathbf{w}^*) \leq \delta, \forall\, i=1,\ldots,N.$$

如果我们能够找到这个线性规划问题的一个解，那么我们就知道$\mathbf{w}^*$ 是一个最优解。

我们还在第3.2节证明了，如果 $r^*>1$，那么条件 2 成立，当且仅当存在 $a_1,\ldots,a_{r^*}\geq 0$和 $\eta_2,\ldots,\eta_K\geq 0$，使得$_{r^*}\sum a_j = 1$ 且式（13）成立。因此，定理 2 表明，$\mathbf{w}^*$ 对于$_{j=1}$是最优的

（8）成立，当且仅当存在 $a_1, \ldots, a_{r^*} \geq 0$ 和 $\eta_2, \ldots, \eta_K \geq 0$，使得 $\eta_k(\Phi_k(\mathbf{w}^*) - h_k(m_k)) = 0$ 对所有 $k = 2, \ldots, K$ 都成立，且式（13）成立。注意，式（13）定义了关于 $\eta_2, \ldots, \eta_K$ 和 $a_1, \ldots, a_{r^*}$ 的 N 个线性不等式。因此，我们 K r^* 可以在满足式（14）–（15）中定义的 $2(K-1)$ 个线性不等式以及下列对式（13）的松弛条件的约束下，使 $\sum \eta_k + \sum a_j$ 最小化，

$$d_{-\lambda_{min}, f_1, \mathbf{a}}(\mathbf{u}_i, \mathbf{w}^*) + \sum_{k=2}^{K} \eta_k d_{\phi_k, f_k}(\mathbf{u}_i, \mathbf{w}^*) \leq \delta, \ \forall \ i = 1, \ldots, N.$$

再一次，如果我们能够找到这个线性规划问题的解，那么我们就知道 $\mathbf{w}^*$ 是一个最优解。

4 最大最小最优设计

假设实验者需要一种设计，使其对所有 K 单目标最优性准则都具有合理的效率。我们将这个最大最小设计问题表述为：

$$\left\{ \begin{array}{ll} \max\limits_{\mathbf{w} \in \mathbb{R}^N} & \min\{\mathrm{Eff}_1(\mathbf{w}), \ldots, \mathrm{Eff}_K(\mathbf{w})\} \\ \text{subject to:} & \sum\limits_{i=1}^{N} w_i = 1, \ w_i \geq 0, \ i = 1, 2, \ldots, N. \end{array} \right\}. \tag{18}$$

4.1 凸优化问题

问题 (18) 很难直接求解，因为目标函数包含一个最小化运算。不过，我们可以将 (18) 等价地表述为：

$$\left\{ \begin{array}{ll} \max\limits_{\mathbf{w} \in \mathbb{R}^N, t \in \mathbb{R}} & 1/t \\ \text{subject to:} & \mathrm{Eff}_k(\mathbf{w}) \geq 1/t, \ k = 1, \ldots, K, \\ & t \geq 0, \ w_i \geq 0, i = 1, \ldots, N, \ \sum_{i=1}^{N} w_i = 1 \end{array} \right\}. \tag{19}$$

通过引入一个额外的优化变量（t），该表述消除了目标函数中的最小化运算。此外，当 $\Phi_k(\mathbf{w}) = \phi_k(\mathcal{I}_{f_k}(\mathbf{w}))$ 中的 ϕ_1，……，ϕ_K 取自

表1中的凸函数时，(19) 等价于：

$$\left\{\begin{array}{ll}\min\limits_{\mathbf{w}\in\mathbb{R}^N,t\in\mathbb{R}} & t \\ \text{subject to: } \Phi_k(\mathbf{w}) \le h_k(1/t),\ k=1,\ldots,K, & \\ \qquad\qquad t\ge 0, \sum_{i=1}^N w_i = 1,\ w_i \ge 0, i=1,\ldots,N. & \end{array}\right\}, \tag{20}$$

其中， $h_k(\cdot)$ 的定义见式 (9)。这是一个凸优化问题，因为 $h_k(1/t)$ 是关于 t 的凹函数，而 $\Phi_k(\mathbf{w})$ 是 $\mathbf{w}$ 的凸函数。此外，我们可以使用 `CVX` 求解式 (19)，因为该问题符合 Grant & Boyd (2008) 描述的 `CVX` 建模框架。请注意，我们的形式更为一般，因为我们允许用户对表1中的准则任意组合、“混搭”使用。

4.2 最优性的充要条件

以下结果刻画了式 (20) 的最优性。

定理 3。 假设$(\mathbf{w}^*, t^*)$ 是问题(20)的可行解。那么，$(\mathbf{w}^*, t^*)$ 是式 *(20)* 的解，当且仅当存在满足以下条件的 $\eta_1, \ldots, \eta_K \ge 0$：

1. $\sum\limits_{k=1}^K \eta_k \left(\frac{d}{d_t} h_k(1/t)\Big|_{t=t^*} \right) = 1.$
2. $\eta_k(\Phi_k(\mathbf{w}^*) - h_k(1/t^*)) = 0$ *for all* $k = 1, 2, \ldots, K.$
3. $\mathbf{w}^* \in \arg\min\limits_{\mathbf{w}\in\Omega} \left\{ \sum\limits_{k=1}^K \eta_k \Phi_k(\mathbf{w}) \right\}.$

现在我们讨论如何利用第 2.2 节中的结果改写定理 3。首先，假设 $\phi_1, \ldots, \phi_K$ 均对应于 *D*- 、 *A*- 、 *c*- 或 *L*- 最优性。于是，由定理 1 和引理 1可知，我们可以将定理 3 中的条件 3 替换为：

$$\sum_{k=1}^K \eta_k d_{\phi_k, f_k}(\mathbf{u}_i, \mathbf{w}^*) \le 0 \text{ for all } i = 1, 2, \ldots, N, \tag{21}$$

其中， $d_k(\mathbf{u}_i, \mathbf{w}^*)$ 由式（5）给出， $\nabla\phi_k(\mathbf{M})$ 的公式见表 2。否则，我们需要结合引理1和引理 3 应用定理 1，以改写条件 3。下面举例说明。

例 2. 假设 ϕ_1(M)、 $= -\lambda_{min}$(M) 以及 $\phi_2, \ldots, \phi_K$ 均对应于 D-、 A-、 c- 或 L- 最优性。于是，由定理 1 和引理 1 可知，定理 3 中的条件 3 成立，当且仅当存在$\mathbf{g} \in \partial\Phi_1(\mathbf{w}^*)$ 使得

$$\eta_1 \mathbf{g}^T (\mathbf{w}^* - \mathbf{e}_i) + \sum_{k=2}^{K} \eta_k d_{\phi_k, f_k}(\mathbf{u}_i, \mathbf{w}^*) \leq 0, \ \forall\ i = 1, 2, \ldots, N, . \tag{22}$$

令 r^* 为 $\lambda_{min}(\mathcal{I}_{f_1}(\mathbf{w}^*))$ 的几何重数。然后，根据引理 3，我们可以考虑两种情况:

- 情况1（$r^* = 1$）：定理 3 中的条件 3 成立，当且仅当

$$\eta_1[(\mathbf{v}^*)^T \mathbf{z}_{f_1}(\mathbf{u}_i, \mathbf{w}^*)]^2 - \eta_1 \lambda_{min}(\mathcal{I}_{f_1}(\mathbf{w}^*)) + \sum_{k=2}^{K} \eta_k d_{\phi_k, f_k}(\mathbf{u}_i, \mathbf{w}^*) \leq \delta, \forall\, i = 1, \ldots, N, \tag{23}$$

  where $\mathbf{v}^*$ denotes an arbitrary unit eigenvector corresponding to $\lambda_{min}(\mathcal{I}_{f_1}(\mathbf{w}^*))$.

- 情况2（$r^* > 1$）：定理 3 中的条件 3 成立，当且仅当存在

  存在 $a_1, \ldots, a_{r^*} \geq 0$ 使得 $\sum_{j=1}^{r^*} a_j = 1$ 并且

$$\eta_1 d_{-\lambda_{min}, f_1, \mathbf{a}}(\mathbf{u}_i, \mathbf{w}^*) + \sum_{k=2}^{K} \eta_k d_{\phi_k, f_k}(\mathbf{u}_i, \mathbf{w}^*) \leq 0 \text{ for all } i = 1, 2, \ldots, N,$$

  其中 $d_{-\lambda_{min},\ f_1,\ \mathbf{a}}(\mathbf{u}_i,\ \mathbf{w}^*)$ 在式 (6) 中定义。

4.3 通过线性规划验证最优性

假设我们通过数值求解式（20）（例如借助 `CVX`）得到了一个候选解 $(\mathbf{w}^*, t^*)$，其中 $\phi_1, \ldots, \phi_K$ 均对应于 D-、 A-、 c- 或 L-最优性。基于第4.2节的结果，我们希望找到

$\eta_1, \ldots, \eta_K \geq 0$，使得:

$$\sum_{k=1}^{K} \eta_k \left(\frac{d}{dt} h_k(1/t)\Big|_{t=t^*} \right) = 1, \tag{24}$$

$$\eta_k(\Phi_k(\mathbf{w}^*) - h_k(1/t^*)) \leq \delta, \quad 1 = 2, \ldots, K, \tag{25}$$

$$-\eta_k(\Phi_k(\mathbf{w}^*) - h_k(1/t^*)) \leq \delta, \quad 1 = 2, \ldots, K, \tag{26}$$

$$\sum_{k=1}^{K} \eta_k d_{\phi_k, f_k}(\mathbf{u}_i, \mathbf{w}^*) \leq \delta, \quad i = 1, 2, \ldots, N, \tag{27}$$

其中 δ 是一个较小的正常数。这里，我们放宽了定理 3 中的条件 2 和 3，因为$\mathbf{w}^*$ 是一个近似解，做法与第 3.3 节类似。我们通过使用 MATLAB 中的 `linprog` 函数求解以下线性规划问题来实现这一目标:

$$\left\{ \begin{array}{l} \min\limits_{\boldsymbol{\eta} \in \mathbb{R}^K} \ \mathbf{1}_K^T \boldsymbol{\eta} \\ \text{subject to: } \boldsymbol{\eta} \geq \mathbf{0}_K, \ \mathbf{b}_2^T \boldsymbol{\eta} = 1, \ \mathbf{B}_2^T \boldsymbol{\eta} \leq \delta \mathbf{1}_N, \mathbf{C}_2 \boldsymbol{\eta} \leq \delta \mathbf{1}_K, -\mathbf{C}_2 \boldsymbol{\eta} \leq \delta \mathbf{1}_K. \end{array} \right\}, \tag{28}$$

其中，$\mathbf{b}_2$ 是一个 K 维向量，其第 k 个分量等于 $\left(\frac{d}{dt} h_k(1/t)\Big|_{t=t^*} \right)$，其中 h_k 由式 (9) 定义；$\mathbf{B}_2$是一个 $K \times N$ 矩阵，其第 (k, i) 个元素等于 $d_{\phi_k, f_k}(\mathbf{u}_i, \mathbf{w}^*)$，其中 $d_{\phi, f}(\mathbf{u}_i, \mathbf{w}^*)$ 由式 (5) 定义；$\mathbf{C}_2 =$ 为对角矩阵，$\mathrm{diag}(\Phi_1(\mathbf{w}^*) - h_1(m_1) \ldots, \Phi_K(\mathbf{w}^*) - h_K(m_K))$。

当 $\phi_1, \ldots, \phi_K$ 中的一个或多个对应于 E-最优性时，我们仍然可以将定理 3 中的条件 3 改写为一组线性不等式；例如，见例 2。因此，我们仍然可以使用线性规划验证定理 3 中的条件。关于如何建立线性规划问题的细节从略，因为其思路与第 3.3 节中的例 1 类似。

5 应用

在以下三个应用中，按照第 3.3 节和第 4.3 节所述通过线性规划验证最优性时，我们设定 $\delta = 10^{-4}$ 。任何大于 10^{-6} 的 δ 取值都会得到相同的结果。所有计- 算均在一台配备 10 核和 16 GB 内存的 2021 M1 Macbook Pro 上完成。我们在`https://github.com/lucylgao/multi-objective-paper-code-2022` 提供了用于复现所有数值结果的 MATLAB 代码。

应用 1。考虑形式为 (1) 的四参数室模型，其中 $p = 1$、$q = 4$、$f(x, \boldsymbol{\theta}) = \theta_1 e^{-\theta_2 x} + \theta_3 e^{-\theta_4 x}$，以及 $S = [0, 15]$，其中响应 y_i 表示药物在各室中的浓度水平，而 x 表示采样时间。该模型已在针对各种最优性准则（包括多目标准则）的最优设计中得到研究（Huang & Wong 1998年，Cheng & Yang 2019年）。

我们寻求满足效率约束的最优设计，以求解式 (7)，其中 $\Phi_k(\mathbf{w}) = \phi_k(\mathcal{I}_f(\mathbf{w}))$，$k = 1$ 取值为 2、3。与 Cheng & Yang (2019) 一样，我们令 ϕ_1 对应于具有 $\mathbf{L} = \text{diag}\left(\frac{1}{\theta_1}, \frac{1}{\theta_2}, \frac{1}{\theta_3}, \frac{1}{\theta_4}\right)$ 的 L-最优性，ϕ_2 对应于 D-最优性，而 ϕ_3 对应于具有 $\mathbf{L} = \left[\int_2^{10} \mathbf{z}_f(x, \boldsymbol{\theta}^*)\mathbf{z}_f^T(x, \boldsymbol{\theta}^*)dx\right]^{\frac{1}{2}}$ 的 L-最优性，其中 $\mathbf{z}_f(x, \boldsymbol{\theta}) = (e^{-\theta_2 x}, -\theta_1 x e^{-\theta_2 x}, e^{-\theta_4 x}, -\theta_3 x e^{-\theta_4 x})^\top$，且 $\boldsymbol{\theta}^* = (5.25, 1.34, 1.75, 0.13)^\top$。我们将连续设计空间 S 离散化，形成具有 $u_i = 15(i-1)/(N-1)$ 的 S_N，其中 $i = 1, \ldots, N$。

首先，我们通过对每个 $k = 1, 2, 3$，使用 `CVX`求解 $\min_{\mathbf{w}\in\Omega} \Phi_k(\mathbf{w})$，得到单目标最优设计。然后，我们使用 `CVX` 求解带有 $m_2 = 0.9$、$m_3 = 0.8$ 和 $N = 501$ 的式 (7)，得到一个记为 $\mathbf{w}^{*m}$的解。我们在表3中报告了单目标最优设计和 $\mathbf{w}^{*m}$。在 $\mathbf{w}^{*m}$处的效率与 Cheng & Yang（2019年）报告的结果接近。

表3：应用 1 中的单目标最优设计，以及带有 $m_2 = 0.9, m_3 = 0.8$ 的效率约束最优设计。

ϕ_1-最优点（权重）	ϕ_2-最优点（权重）	ϕ_3-最优点（权重）	效率约束点（权重）
0（0.0591）	0（0.2500）	0 (0.1339)	0 (0.1339)
0.6300 (0.1315)	0.6600 (0.2500)	0.9600 (0.0663)	0.6600 (0.1513)
2.9400 (0.3126)	2.8800 (0.2500)	3.300 (0.4502)	3.0300 (0.2481)
13.2900 (0.4968)	11.0100 (0.2441)	9.7500 (0.2231)	3.0600 (0.0942)
	11.0400 (0.0059)	9.7800 (0.2502)	10.8300 (0.0807)
			10.8600 (0.2918)

接下来，我们使用 `linprog` MATLAB 函数求解带有 $\delta = 10^{-4}$ 的 (17)，如第 3.3 节所述，从而验证定理 2 中关于$\mathbf{w}^{*m}$ 的最优性条件。求解 (17) 得到 $\eta_2^* = 36.4870$ 和 $\eta_3^* = 5.0767$。由于我们

得到了一个解，我们知道 $\mathbf{w}^{*m}$ 确实是受效率约束的最优设计（定理 2）。图 1 显示，对于所有 $i = 1, 2, \ldots, N$，$d_{\phi_1,f}(u_i, \mathbf{w}^{*m}))+\sum \eta_k^* d_{\phi_k,f}(u_i, \mathbf{w}^{*m})) \leq_{k=2}\delta$；这直观地表明，对于 $\mathbf{w}^{*m}$，η_2^* 和 η_3^*，定理 2 中的条件 2 得到了满足。图 1 还显示，对于所有 $k = 1, 2, 3$，$d_{\phi_k,f}(u_i, \mathbf{w}^{*m})$ 并非处处非负。因此，$\mathbf{w}^{*m}$ 并非使 Φ_1, Φ_2 或 Φ_3 达到最小值的单目标最优设计（定理 1 和引理 2）。

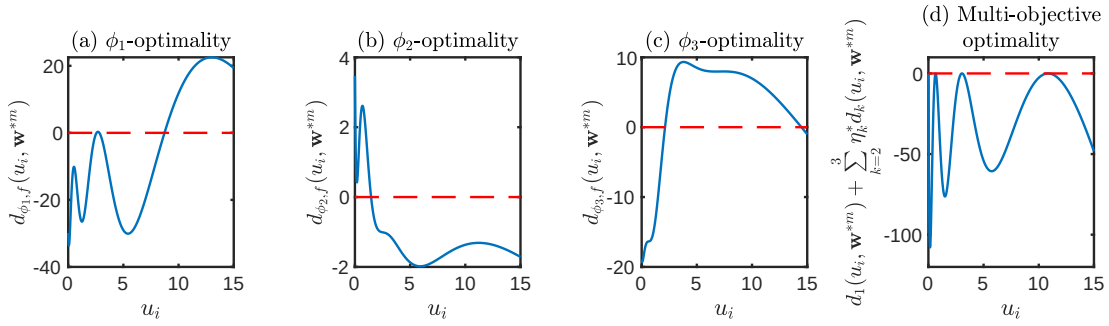

图 1：在应用 1 中，对于 $(m_2, m_3) = (0.9, 0.8)$，面板 (a)–(c) 展示了 $d_{\phi_k,f}(u_i, \mathbf{w}^{*m}))$ 在 $k = 1, 2, 3$ 时的图，面板 (d) 则展示了 $d_{\phi_1,f}(u_i, \mathbf{w}^{*m}))+\sum_{k=2}^{3} \eta_k^* d_{\phi_k,f}(u_i, \mathbf{w}^{*m}))$。在面板 (d) 中，红色虚线表示水平线 $y = \delta$，对应 $\delta = 10^{-4}$.

接下来，我们改变 m_2 和 m_3，观察结果如何变化；表 4 给出了具有代表性的结果。对于 $m_2 = 0.90$ 和 $m_3 = 0.70$，我们发现$\eta_3^* = 0$；这是因为 $\Phi_3(\mathbf{w}^{*m}) < h_3(m_3)$，且 η_2^*，η_3^* 满足定理 2 中的条件 1。类似地，对于 $m_2 = 0.70$ 和 $m_3 = 0.70$，我们发现 $\eta_2^* = \eta_3^* = 0$，因为 $\Phi_k(\mathbf{w}^{*m}) < h_k(m_k)$（即 $\text{Eff}_k(\mathbf{w}^{*m}) > m_k$），当 $k = 2, 3$ 时成立；这意味着多目标最优设计 $\mathbf{w}^{*m}$ 也是使 Φ_1 最大化的单目标最优设计（定理 1 和引理 2）。对于 $m_2 = 0.90$ 和$m_3 = 0.90$，不存在可行解。

计算 $(m_2, m_3) = (0.9, 0.8)$ 的最优设计，对于 $N = 101, 501, 1001$ 分别耗时 19.5、25.6 和 36.2 秒。验证效率约束设计的最优性耗时不到一秒。

应用 2. 临床剂量探索研究中常用几种剂量-反应模型。我们考虑 Bretz 等人（2010）提出的四种形式为 (1) 的竞争回归模型，以构造最大最小最优设计，并验证定理 3 中关于最优设计的充分必要条件。这四种不同的模型为:

表 4：对于应用 1，不同 $(m_2,\ m_3)$ 下的效率以及 η_2^* 和 η_3^*。

情形 $(m_2,\ m_3)$	(0.90, 0.80)	(0.90, 0.70)	(0.70, 0.70)	(0.90, 0.90)
$\eta_2^*,\ \eta_3^*$	36.4870, 5.0767	7.2923, 0	0, 0	NA
$\text{Eff}_1(\mathbf{w}^{*m})$	0.8694	0.9360	1.000	NA
$\text{Eff}_2(\mathbf{w}^{*m}),\ \text{Eff}_3(\mathbf{w}^{*m})$	0.9000, 0.8000	0.9000, 0.7035	0.7317, 0.7746	NA

（i）线性模型：$f_1(x,\ \boldsymbol{\theta}) = \theta_{11} + \theta_{12}x$；（ii）Emax I 模型：$f_2(x,\ \boldsymbol{\theta}) = \theta_{21} + \theta_{22}x/(\theta_{23} + x)$；（iii）Emax II 模型：$f_3(x,\ \boldsymbol{\theta}) = \theta_{31} + \theta_{32}x/(\theta_{33} + x)$；（iv）逻辑斯蒂模型：$f_4(x,\ \boldsymbol{\theta}) = \theta_{41} + \theta_{42}/(1 + \exp[(\theta_{43} - x)/\theta_{44}])$，其中 $x \in [0, 500]$ (μg) 表示剂量水平。设 S_N 包含 $[0, 500]$中 $N = 501$ 个等距网格点。与 Bretz 等人（2010）一样，我们假设 Emax I、Emax II 和逻辑斯蒂模型的真实参数值分别为 (60, 294, 25)、(60, 340, 107.14) 和 (49.62, 290.51, 150, 45.51)。（线性模型的信息矩阵不依赖于其真实参数值。）

我们令 $\Phi_k(\mathbf{w}) = \phi_k(\mathcal{I}_{f_k}(\mathbf{w}))$，其中 ϕ_k 对应于所有 $k = 1, \ldots, 4$ 的 D-最优性（定义见表1），然后通过 CVX求解问题 (20)，得到记为 $\mathbf{w}^{*mm}$ 的解。S_N 上的单目标 D-最优设计和最大最小 D-最优设计见表 5。我们发现，$t^* = 1.1712$，对于 $k = 1$,2,4，$\text{Eff}_k(\mathbf{w}^{*mm}) = 0.8538$，且 $\text{Eff}_3(\mathbf{w}^{*mm}) = 0.8547$。通过 MATLAB 函数 `linprog` 求解问题 (28)，得到 $\eta_1^* = 0.1983, \eta_2^* =$0.1291, $\eta_3^* = 0$ 和 $\eta_4^* = 0.0968$。由于我们得到了一个解，因此知道$\mathbf{w}^{*mm}$ 是最大最小 D-最优设计（定理 3）。

图 2 展示了 $d_{\phi_k,\ f_k}(u_i,\ \mathbf{w}^{*mm})$ 在 $k = 1$，.....，4 以及$\sum_4 \eta_k^* d_{\phi_k,\ f_k}(u_j, \mathbf{w}^{*mm})$ 下的图。图 2(e) 确认定理$_{k=1}$3 中的条件 3 得到满足。图 2(a)–(d) 表明，对于四个模型中的任何一个，$\mathbf{w}^{*mm}$ 都不是单目标D-最优设计（定理 1 和引理 2）。

计算最优设计分别耗时 18.2、31.0 和 48.7 秒，对应于$N = 101$, 501，1001。验证最大最小设计的最优性耗时不到一秒。

应用3：考虑形式为（1）的线性模型，其中 $p = 2, q = 5$,

表 5：应用 2 的最优设计。

线性模型 点（权重）	Emax I 点（权重）	Emax II 点（权重）	逻辑斯蒂 点（权重）	最大最小 点（权重）
0 (0.5000)	0 (0.3333)	0 (0.3333)	0 (0.2500)	0 (0.2406)
500 (0.5000)	22 (0.3333)	75 (0.3333)	114 (0.2500)	19 (0.1806)
	500 (0.3333)	500 (0.3333)	204 (0.1316)	112 (0.1314)
			205 (0.2500)	204 (0.1070)
			500 (0.2500)	205 (0.0178)
				500 (0.3225)

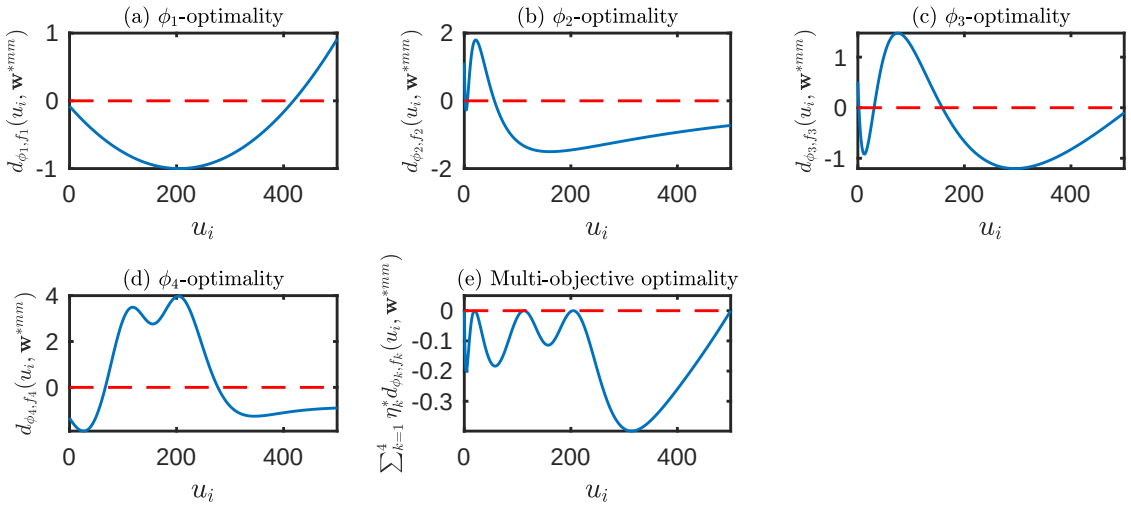

图 2：对于应用 2，我们展示了 (a)–(d) $d_{\phi_k,\ f_k}(u_i,\ \mathbf{w}^{*mm})$ 对$_4k = 1$，2，3，4 的图，以及 (e) $\sum \eta_k^* d_{\phi_k,\ f_k}(u_i,\ \mathbf{w}^{*mm})$ 的图。在面板 (e) 中，红色虚线$_{k=1}$表示水平线 $y = \delta$，对应于 $\delta = 10^{-4}$.

$f\ (\mathbf{x};\ \boldsymbol{\theta})\ = \theta_1 + x_1\theta_2 + x_2\theta_3 + x_1x_2\theta_4 + x_2^2\theta_5$，以及 $S = \{0, 1\} \times [-1, 1]$。我们令$S^1_{N/2}$ 包含 $[-1, 1]$ 上的201个等间距点，并将设计空间 S 离散化以构成 $S_N = \{0, 1\} \times S^1_{N/2}$，其中 $N = 402$。这里，信息矩阵不依赖于真实参数值。

我们令 $\Phi_k(\mathbf{w}) = \phi_k(\mathcal{I}_f(\mathbf{w}))$，其中 $k = 1$、2、3 对应于 ϕ_1——

最优性，ϕ_2 对应于 E-最优性，ϕ_3 对应于 c-最优性，其中 $\mathbf{c} = (0、0、0、1、0)^T$，然后按照第4.1节所述求解问题 (20)，以获得最大最小最优设计；我们将该解记为 $\mathbf{w}_{lm}^{*mm}$。S_N 上的单目标最优设计和最大最小最优设计见表6。我们发现，$t^* = 1.2979$、$\text{Eff}_1(\mathbf{w}_{lm}^{*mm}) = 0.9298$ 和 $\text{Eff}_1(\mathbf{w}_{lm}^{*mm}) = 0.7705$ 对于 $k = 2、3$ 成立。在这种情形下，$\lambda_{min}(\mathcal{I}_f(\mathbf{w}_{lm}^{*mm}))$ 的几何重数等于 1。因此，为了验证$\mathbf{w}_{lm}^{*mm}$,3的最优性，我们使用 `linprog` MATLAB 函数，在满足 (23) 中的$_{k=1}$线性不等式约束以及 (24)–(26) 中的线性等式和不等式约束的条件下，使 $\sum \eta_k$ 最小化。我们得到一个解 $\eta_1^* = 0、\ \eta_2^* = 0.2445$ 和 $\eta_3^* = 0.0151$。因此，我们得到的最大最小设计是最优的。

表6：应用3的最优设计。

A-最优 点 [权重]	E-最优 点 [权重]	c-最优 点 [权重]	最大最小 点 [权重]
(0, -1) [0.1859]	(0, -1) [0.2069]	(0, -1) [0.2500]	(0, -1) [0.1926]
(1, -1) [0.1399]	(1, -1) [0.1379]	(1, -1) [0.2500]	(1, -1) [0.1926]
(0, 0) [0.2287]	(0, 0) [0.2414]	(0, 1) [0.2500]	(0, 0) [0.1679]
(1, 0) [0.1197]	(0, 1) [0.0690]	(1, 1) [0.2500]	(1, 0) [0.0616]
(0, 1) [0.1859]	(1, 1) [0.2069]		(0, 1) [0.1926]
(1, 1) [0.1399]			(1, 1) [0.1926]

计算最优设计分别耗时 11.2、14.5 和 23.2 秒，适用于$N/2 = 101$、201、*401*。验证最大最小设计的最优性耗时不到一秒。

6 结论

本文展示了如何使用凸优化在离散设计空间上求解多目标最优设计问题，以及如何使用线性规划验证设计的最优性。我们的方法可应用于效率约束或最大最小最优设计问题，这些问题结合了表1中的任意单目标准则。

多目标最优设计为增强对参数和/或模型设定错误的稳健性提供了天然契机：我们

可以纳入基于一系列 $\boldsymbol{\theta}^*$猜测值构建的目标函数，和/或基于多个模型的信息矩阵构建的目标函数。序贯多目标最优设计或许能进一步提升稳健性，因为我们可以分阶段选择设计点和权重，并利用各阶段的数据，为下一阶段目标函数中采用的参数和/或模型的选择提供依据。这或许是一个富有前景的未来研究方向。

本文能够得出这些结果并提出相应算法，是因为在模型 (1) 下，普通最小二乘估计量的渐近协方差矩阵的逆矩阵是 $\mathbf{w}^*$ 的线性函数；参见公式 (2)。将本文的结果和算法扩展到具有类似性质的其他模型和估计量并不困难。例如，我们可以允许 (1) 中的误差向量具有异方差或分块对角协方差结构，并使用广义最小二乘估计量。另一个例子是具有典则连接函数的广义线性模型，在其中我们使用最大似然估计量估计 $\boldsymbol{\theta}$ 。

涉及一组未知参数的最优性充分必要条件，也出现在本文所讨论的多目标设计问题之外的情形中。例如，任何涉及凸但不可微目标函数的单目标最优设计问题（如单目标 E-最优性）都属于这种情形。当这些条件在未知参数中定义了线性等式与不等式时，我们可以像本文一样使用线性规划对其进行验证。

我们的结果和算法的一个局限在于假设设计空间是离散的。未来研究的一个重要方向是，在连续设计空间下推导结果并开发算法。

致谢

这项研究工作部分得到了加拿大自然科学与工程研究委员会探索资助的支持。

参考文献

Bretz, F., Dette, H. & Pinheiro, J. C. (2010), ‘临床剂量探索研究中最优设计的实际考量’, *Statistics in Medicine* **29**(7-8),731–742.

Cheng, Q. & Yang, M. (2019), '关于多目标最优设计', *Journal of Statistical Planning and Inference* **200**, 87–101. Clarke, F. H. (1983), *Optimization and nonsmooth analysis*, Wiley, 纽约 . Clyde, M. & Chaloner, K. (1996), ' 多目标设计问题中约束设计与加权设计的等价性 ', *Journal of the American Statistical Association* **91**(435), 1236– 1244. Cook, R. D. & Wong, W. K. (1994), '关于约束最优设计与复合最优设计的等价性', *Journal of the American Statistical Association* **89**(426), 687–692. Dette, H., Haines, L. M. & Imhof, L. A. (2007), '回归模型的最大最小设计与贝叶斯最优设计', *Statistica Sinica* **17**(2), 463–480. Gao, L. L. & Zhou, J. (2017), '基于二阶最小二乘估计量的D-最优设计', *Statistical Papers* **58**(1), 77–94. Grant, M. & Boyd, S. (2014), 'CVX：用于规范凸规划的 MATLAB 软件，2.1版', http://cvxr.com/cvx. Grant, M. C. & Boyd, S. P. (2008), 非光滑凸规划的图实现, *in*V. Blondel, S. Boyd & M. Kimura, eds, '学习与控制近期进展（向 M. Vidyasagar 致敬）', Springer, pp. 95–110. Huang, Y.-C. & Wong, W.-K. (1998), '多目标最优设计的序贯构造', *Biometrics* **54**(4), 1388–1397. Imhof, L. & Wong, W. K. (2000), '寻找极大极小效率设计的图解法', *Biometrics* **56**(1), 113–117. Kiefer, J. (1974), '最优设计的一般等价理论（近似理论）', *The Annals of Statistics* **2**(5), 849–879. Lee, C. M.-S. (1988), '约束最优设计', *Journal of Statistical Planning and Inference***18**(3), 377–389. Lewis, A. (1999), '特征值的非光滑分析', *Mathematical Programming* **84**, 1–24. Luenberger, D. G. & Ye, Y. (2016), *Linear and nonlinear programming (4th ed.)*, Springer, Cham. Mordukhovich, B. S. & Nam, N. M. (2013), *An easy path to convex analysis and applications*, 数学与统计综合讲座, Springer, Cham. M¨uller, C. H. & P´azman, A. (1998), '充分必要条件在极大极小有效设计中的应用', *Metrika* **48**(1), 1–19. Petersen, K. B. & Pedersen, M. S. (2012), '矩阵手册'. 20121115版.**网址：** *http://www2.compute.dtu.dk/pubdb/pubs/3274-full.html*Wong, W. K. (1999), '多目标设计策略的最新进展', *Statistica Neer-landica* **53**(3), 257–276. Wong, W. K. & Zhou, J. (2019), '用于构造各种最优回归设计的基于CVX的算法', *Canadian Journal of Statistics* **47**(3), 374–391.

Wong, W. K. & Zhou, J. (2022年), '使用 CVX 构建用于生物医学研究的多重目标最优设计', *Journal of Computational and Graphical Statistics*(刚刚获接受，可在线查阅：https://doi.org/10.1080/10618600.2022.2104858。).

附录

一般凸优化理论

本节回顾凸优化中关于最优性条件的一般结果。

命题 A1. *Let* $\Phi: \mathbb{R}^N \mapsto \mathbb{R}$ *be a convex function and* C *be a closed convex set. Then,* $\mathbf{w}^* \in \operatorname{argmin}_{\mathbf{w}\in C} \Phi(\mathbf{w})$ *if and only if there exists* $\mathbf{g} \in \partial\Phi(\mathbf{w}^*)$ *such that*

$$\mathbf{g}^T(\mathbf{w}-\mathbf{w}^*) \geq 0, \quad \textit{for all } \mathbf{w} \in C, \tag{29}$$

where $\partial\Phi(\mathbf{w}^*) \equiv \{\mathbf{g} \in \mathbb{R}^N: \Phi(\mathbf{w}) - \Phi(\mathbf{w}^*) \geq \mathbf{g}^T(\mathbf{w}-\mathbf{w}^*) \ \forall \mathbf{w} \in \mathbb{R}^N\}$.

命题 A1 是 Mordukhovich和Nam（2013年）定理4.14的直接推论。以下结果刻画了受约束凸优化问题的最优性。

命题 A2. *Define the following convex optimization problem:*

$$\min_{\mathbf{w}\in C} \quad \Phi_0(\mathbf{w}) \qquad \textit{subject to: } \Phi_l(\mathbf{w}) \leq 0, \ l = 1, \ldots, L, \tag{30}$$

where Φ_0, ..., Φ_L: $\mathbb{R}^N \to \mathbb{R}$ *are convex functions and* C *is a closed convex set. Suppose that Slater's condition holds, i.e. there exists* $\mathbf{w}' \in C$ *such that* $\Phi_l(\mathbf{w}') < 0$ *for all* $l = 1$, ..., L. *Then, a feasible solution* $\mathbf{w}^*$ *of* (30) *solves* (30) *if and only if there exists* η_1, ..., $\eta_L \geq 0$ *such that* $\eta_l\Phi_l(\mathbf{w}^*) = 0$ *for all* $l = 1$, 2, ..., L *and* $\mathbf{w}^* \in \arg\min_{w\in C} \left\{\Phi_0(\mathbf{w}) + \sum_{l=1}^{L} \eta_k\Phi_l(\mathbf{w})\right\}$.

Proof. 我们假设 Slater 条件成立。因此，Mordukhovich 和 Nam（2013年）的定理4.18表明，当且仅当存在乘子 $\eta_1,..,\eta_L \geq 0$ 使得 $\eta_l\Phi_l(\mathbf{w}^*) = 0$ 对所有 $l = 1, 2, \ldots, L$ 都成立时，$\mathbf{w}^*$ 是问题（30）的最优解，并且

$$0 \in \partial\Phi_0(\mathbf{w}^*) + \sum_{l=1}^{L} \eta_l\partial\Phi_l(\mathbf{w}^*) + \left\{\mathbf{g} \in \mathbb{R}^N : \mathbf{g}^T\mathbf{w}^* \geq \mathbf{g}^T\mathbf{w} \ \forall \ \mathbf{w} \in C\right\}. \tag{31}$$

因此，根据引理1，（31）等价于：

$$\exists \mathbf{g} \in \partial\left(\Phi_0 + \sum_{l=1}^{L} \eta_l\Phi_l\right)(\mathbf{w}^*) \ \text{ such that } \ \mathbf{g}^T(\mathbf{w}-\mathbf{w}^*) \geq 0 \quad \text{for all } \mathbf{w} \in C. \tag{32}$$

最后，根据命题 A1，(32) 成立当且仅当 $\mathbf{w}^* \in \arg\min_{\mathbf{w}\in C} \left\{\Phi_0(\mathbf{w}) + \sum_{l=1}^{L} \eta_l\Phi_l(\mathbf{w})\right\}$。

□

B 定理 1 的证明

令 $\mathbf{w}^* \in \Omega$。由于 $\Phi(\mathbf{w}) = \phi(\mathcal{I}_f(\mathbf{w}))$ 是凸的，且 Ω 是闭凸集，根据命题 A1，$\mathbf{w}^* \in \operatorname{argmin}_{\mathbf{w}\in\Omega} \Phi(\mathbf{w})$ 当且仅当

$$\exists \mathbf{g} \in \partial\Phi(\mathbf{w}^*) \text{ such that } \mathbf{g}^T(\mathbf{w} - \mathbf{w}^*) \geq 0, \text{ for all } \mathbf{w} \in \Omega. \tag{33}$$

假设（33）成立。由于对所有 $i = 1、2、\ldots\ldots、N$，$\mathbf{e}_i \in \Omega$，我们有（3）成立。现在假设（3）成立。由于所有 $\mathbf{w} \in \Omega$ 对所有 $i = 1、2、\ldots\ldots、N$ 都有 $w_i \geq 0$，因此（3）意味着存在 $\mathbf{g} \in \partial\Phi(\mathbf{w}^*)$，使得对所有 $\mathbf{w} \in \Omega$，$\sum_{i=1}^{N} w_i \mathbf{g}^T(\mathbf{e}_i - \mathbf{w}^*) \geq 0$。注意到 $\mathbf{g}^T(\mathbf{w} - \mathbf{w}^*) = \sum_{i=1}^{N} w_i \mathbf{g}^T(\mathbf{e}_i - \mathbf{w}^*)$，证明完毕。

□

C 引理 3 的证明

我们首先建立 $(-\Phi)(\mathbf{w}) = \lambda_{min}(\mathcal{I}_f(\mathbf{w}))$ 的性质，因为这使我们能够利用关于最小特征值函数 $\lambda_{min}(\mathbf{M})$ 的现有理论结果。

函数 $(-\Phi)(\mathbf{w})$ 在 $\mathbb{R}^N$中的每一点都不可微。此外，次微分的概念不适用于 $(-\Phi)(\mathbf{w})$，因为 $(-\Phi)(\mathbf{w})$ 是 *concave* 而不是凸函数。不过，它具有 *Clarke subdifferential* (Clarke 1983)，将梯度的概念推广到了局部 Lipschitz 连续函数这一类。定义在 $\mathbb{R}^N$ 上的局部 Lipschitz 连续函数 $h(\mathbf{w})$ 的 Clarke 次微分定义如下：

$$\partial^C h(\mathbf{w}) = \operatorname{co}\left(\left\{\mathbf{v} \in \mathbb{R}^N : \exists \{\mathbf{w}_k\}_{k=1}^{\infty} \text{ s.t. } \lim_{k\to\infty} \mathbf{w}_k \text{ exists}, \nabla h(\mathbf{w}_k) \text{ exists, and } \lim_{k\to\infty} \nabla h(\mathbf{w}_k) = \mathbf{v}\right\}\right),$$

wh其中 $\operatorname{co}(S)$ 是集合 S的凸包，即所有包含 S的凸集的交集。

通过观察 $(-\Phi)(\mathbf{w})$ 是由不可微凹函数 λ_{min} 与线性函数 $\mathcal{I}_f$复合而成，我们可以刻画 Clarke 次微分 $\partial^C(-\Phi)(\mathbf{w})$。因此，根据 Clarke 次微分链式法则（Clarke 1983 中的定理 2.3.10），$\mathbf{g}_c \in \partial^c(-\Phi)(\mathbf{w}^*)$ 当且仅当存在 $\mathbf{M} \in \partial^c \lambda_{\min}(\mathcal{I}_f(\mathbf{w}^*))$，使得对任意 $\mathbf{w} \in \mathbb{R}^N$，

$$\mathbf{g}_c^T \mathbf{w} = \operatorname{trace}\left(\mathbf{M}\left(\sum_{i=1}^{N} w_i \mathbf{z}_f(\mathbf{u}_i)\mathbf{z}_f^T(\mathbf{u}_i)\right)\right). \tag{34}$$

此外，Lewis（1999）的推论 10 表明

$$\partial^c \lambda_{min}(\mathcal{I}_f(\mathbf{w}^*)) = \left\{\sum_{j=1}^{r^*} a_j \mathbf{v}_j^* [\mathbf{v}_j^*]^T : a_j \geq 0, \sum_{j=1}^{r^*} a_j = 1\right\}, \tag{35}$$

回想一下，在引理 3 的陈述中，我们将 $\mathbf{v}_1, \ldots, \mathbf{v}_{r^*}$ 定义为与 $\lambda_{min}(\mathcal{I}_f(\mathbf{w}^*))$相对应的任意一组 r^*线性无关单位特征向量，其中 r^* 是几何

重数。由 (2) 和 (34)–(35) 中 $\mathcal{I}_f(\mathbf{w}^*)$的定义可知，若 $\mathbf{g}_c \in \partial^c(-\Phi)(\mathbf{w}^*)$，则存在 $a_1, \ldots, a_{r^*} \geq 0$，使得 $\sum_{j=1}^{r^*} a_j = 1$ 且 $\mathbf{g}_c^T(\mathbf{e}_i - \mathbf{w}^*) = \sum_{j=1}^{r^*} a_j([\mathbf{v}_j^*]^T \mathbf{z}_f(\mathbf{u}_i))^2 - \lambda_{\min}(\mathcal{I}_f(\mathbf{w}^*))$，对所有 $i = 1, 2, \ldots, N.$ 均成立。此外，$\partial\Phi(\mathbf{w}^*) = \partial^c\Phi(\mathbf{w}^*) = -[\partial^c(-\Phi)(\mathbf{w}^*)],$，其中第一个等式由 Clarke 1983 的命题 2.2.7 得出，第二个等式由 Clarke 1983 的命题 2.3.1 得出。因此，对于任意 $\mathbf{g} \in \partial\Phi(\mathbf{w}^*)$，我们知道 $-\mathbf{g} \in \partial^c(-\Phi)(\mathbf{w}^*)$。由此，存在 $a_1, \ldots, a_{r^*} \geq 0$，使得 $\sum_{j=1}^{r^*} a_j = 1$ 且 $-\mathbf{g}^T(\mathbf{e}_i - \mathbf{w}^*) = \sum_{j=1}^{r^*} a_j([\mathbf{v}_j^*]^T \mathbf{z}_f(\mathbf{u}_i))^2 - \lambda_{\min}(\mathcal{I}_f(\mathbf{w}^*))$，对所有 $i = 1, 2, \ldots, N.$均成立

□

D 定理 2 的证明

由于我们假设存在满足 $\mathrm{Eff}_k(\mathbf{w}) > m_k$ for all $k = 2, \ldots, K$ 的 $\mathbf{w} \in \Omega$，因此问题 (8) 满足 Slater 条件。此外，对于所有 $k = 1, 2, \ldots, K$，$\Phi_k(\mathbf{w}) = \phi_k(\mathcal{I}_{f_k}(\mathbf{w}))$ 是凸函数，而 Ω 是凸集。因此，根据命题 A2，$\mathbf{w}^* \in \Omega$ 是 (8) 的解，当且仅当存在 $\eta_2, \ldots, \eta_K \geq 0$，使得 $\eta_k(\Phi_k(\mathbf{w}^*) - h_k(m_k)) = 0$for all $k = 2, 3, \ldots, K$，并且 $\mathbf{w}^* \in \arg\min_{\mathbf{w}\in\Omega} \left\{ \Phi_1(\mathbf{w}) + \sum_{k=2}^{K} \eta_k(\Phi_k(\mathbf{w}) - h_k(m_k)) \right\}$. 注意到$\sum_{k=2}^{K} \eta_k h_k(m_k)$ 与 $\mathbf{w}$ 无关，证明即告完成。

□

E 定理 3 的证明

我们首先确认，以下对式（20）的重述

$$\min_{\mathbf{w}\in\Omega, t\geq 0} \; t \qquad \text{subject to: } \Phi_k(\mathbf{w}) \leq h_k(1/t),\ k = 1, \ldots, K, \tag{36}$$

满足命题 A2 中的条件。由于 Ω 和 $\mathbb{R}$ 是闭凸集，$\Omega \times \mathbb{R}$ 是闭凸集。此外，对于所有 $k = 1, 2, \ldots, K$，$\Phi_k(\mathbf{w}) = \phi_k(\mathcal{I}_{f_k}(\mathbf{w}))$ 是凸函数。我们还知道，对于所有 $k = 1, 2, \ldots, K$，式（9）中的 $h_k(1/t)$ 对所有 t 都是凹函数，因为对于任意正定矩阵 $\mathbf{M}$，$-\lambda_{\min}(\mathbf{M})/t$ 是关于 t 的凹函数，t 是线性函数，且 $q_k \log(t)$ 是凹函数。为证明 Slater 条件成立，我们需要找到 $\mathbf{w}' \in \Omega$ 和 $t' > 0$，使得对于所有 $k = 1, 2, \ldots, K$，$\Phi_k(\mathbf{w}') < h_k(1/t')$。根据表1中效率函数 $\mathrm{Eff}_k(\mathbf{w}')$ 的定义，$\Phi_k(\mathbf{w}') < h_k(1/t')$ 当且仅当 $\mathrm{Eff}_k(\mathbf{w}') > 1/t'$，且对于所有 $\mathbf{w}' \in \Omega$，$\mathrm{Eff}_k(\mathbf{w}') > 0$。因此，选择 $\mathbf{w}' = \arg\min \Phi_1(\mathbf{w})$ 且 $\mathbf{w}\in\Omega t' = 2/\left(\min_{k=1,2,\ldots,K} \mathrm{Eff}_k(\mathbf{w}') \right)$，即可满足 Slater 条件。

现在，我们可以将命题 A2 应用于式 (36)，得到以下结果：一个可行解

$(\mathbf{w}^*, t^*)$ 是式 (36) 的解，当且仅当存在 ν，η_1，……，$\eta_K \geq 0$ 满足：

$$\nu t^* = 0, \tag{37}$$

$$\eta_k(\Phi_k(\mathbf{w}^*) - h_k(1/t^*)) = 0 \text{ for all } k = 1, 2, \ldots, K, \tag{38}$$

$$(\mathbf{w}^*, t^*) \in \underset{\mathbf{w}\in\Omega, t\in\mathbb{R}}{\arg\min} \left\{ t - \nu t + \sum_{k=1}^{K} \eta_k(\Phi_k(\mathbf{w}) - h_k(1/t)) \right\}. \tag{39}$$

式 (39) 中的优化问题是可分离的。因此，式 (39) 可改写为

$$\mathbf{w}^* \in \underset{\mathbf{w}\in\Omega}{\arg\min} \left\{ \sum_{k=1}^{K} \eta_k \Phi_k(\mathbf{w}) \right\}, \tag{40}$$

$$t^* \in \underset{t\in\mathbb{R}}{\arg\min} \left\{ t - \nu t - \sum_{k=1}^{K} \eta_k h_k(1/t) \right\}. \tag{41}$$

由于 $g(t) = t - \nu t - \sum_{k=1}^{K} \eta_k h_k(1/t)$ 是凸函数，我们可以将式 (41) 改写为

$$1 - \nu - \sum_{k=1}^{K} \eta_k \left[\frac{d}{dt} h_k(1/t) \Big|_{t=t^*} \right] = 0. \tag{42}$$

这意味着，存在 ν，η_1，……，$\eta_K \geq 0$ 满足式 (37)–(39)，当且仅当存在η_1，……，$\eta_K \geq 0$ 满足式 (38)、(40) 以及

$$1 - \sum_{k=1}^{K} \eta_k \left[\frac{d}{dt} h_k(1/t) \Big|_{t=t^*} \right] \geq 0, \qquad t^* \left(1 - \sum_{k=1}^{K} \eta_k \left[\frac{d}{dt} h_k(1/t) \Big|_{t=t^*} \right] \right) = 0. \tag{43}$$

由于式 (38) 是定理 3 中的条件 2，而式 (40) 是定理 3 中的条件 3，因此还需证明式 (43) 等价于定理 3 中的条件 1。只需证明 $t^* > 0$。回顾一下，式 (36) 等价于式 (19)，因此，式 (36) 中 t 的最优值是由最优设计达到的最大最小效率的倒数。效率的取值范围为 0 到 1，因此 t 的最优值必须大于 1，即 $t^* > 1$。

□