



随机复杂系统设计方案排序问题的动态计量预算分配策略

肖辉¹, 魏志宏¹, 寇纲², 顾先明^{3*}

1. 西南财经大学管理科学与工程学院, 成都 611130

2. 西南财经大学大数据研究院, 成都 611130

3. 西南财经大学数学学院, 成都 611130

* 通信作者. E-mail: guxm@swufe.edu.cn

收稿日期: 2022-12-12; 修回日期: 2023-04-07; 接受日期: 2023-05-10; 网络出版日期: 2024-02-06

国家自然科学基金 (批准号: 71971176, 71910107002)、四川省中央引导地方科技发展专项项目 (批准号: 2022ZYD0006) 和中央高校基本科研业务费专项资金 (批准号: JBK2103010) 资助项目

摘要 仿真技术已被广泛地用于评估与优化随机复杂系统, 例如生产制造、网络通信、医疗健康等信息物理融合系统. 然而, 将仿真优化方法用于方案排序问题的研究甚少. 本文考虑只能通过系统仿真才能获得性能表现的 k 个备选方案排序问题, 在贝叶斯 (Bayes) 框架下, 构建以寻求最大化后验方案正确排序概率为目标的动态仿真预算优化分配模型. 通过将仿真预算分配决策过程表述成一个随机动态规划问题, 推导出一种动态计量预算分配策略, 并证明了该分配策略的渐进最优性. 数值实验以及案例分析的结果表明, 本文提出的算法显著地提高了方案排序问题的仿真优化效率.

关键词 仿真优化, 排序选优, 信息物理融合系统, 马尔科夫决策过程, 贝叶斯估计

1 引言

随着“工业 4.0”和“中国制造 2025”的实施, 信息物理融合系统 (cyber-physical system, CPS) 被认为是支撑信息化和工业化深度融合以及引领全球新一轮产业变革的核心技术体系^[1,2]. 数字孪生 (digital twin, DT) 的本质是在信息世界对物理世界的等价数字映射, 它是实现 CPS 的关键技术. 在 DT 技术中, 对系统进行仿真是其不可或缺的重要环节^[3,4]. 相较于数学解析模型, 系统仿真可以更加精确地对随机复杂系统进行建模^[5]. 然而系统仿真通常耗时且成本高昂, 特别是随着系统复杂性的增加, 仿真的成本会急剧上升^[6]. 同时, 由于系统的仿真输出具有随机性, 这使得决策者需要对每一个系统设计 (备选方案) 进行多次的仿真复制, 才能获得其系统性能表现统计意义上的精确估计^[7]. 因此, 如何优化仿真预算的分配以提高系统仿真的效率是仿真优化领域的重要课题之一.

引用格式: 肖辉, 魏志宏, 寇纲, 等. 随机复杂系统设计方案排序问题的动态计量预算分配策略. 中国科学: 信息科学, 2024, 54: 373–393, doi: 10.1360/SSI-2022-0451
Xiao H, Wei Z H, Kou G, et al. Dynamic computing budget allocation for design ranking in stochastic complex systems (in Chinese). Sci Sin Inform, 2024, 54: 373–393, doi: 10.1360/SSI-2022-0451

为了提高系统仿真的效率, 现有的文献提出了排序选优方法, 该方法通过优化仿真预算的分配, 从而有效地从备选方案中挑选出最优方案. 已有的排序选优方法大致可分为固定精度方法以及固定预算方法两大类^[8~10]. 固定精度方法旨在寻求一种可行的仿真预算分配规则以保证预先设定的正确选择概率 (probability of correct selection, PCS). 例如, 无差异区域 (indifferent-zone, IZ) 算法^[11,12] 等. 固定预算方法, 在给定仿真预算的情况下, 以最大化 PCS 或其他性能度量为目标来设计可行的仿真预算分配算法, 如最优计量预算分配 (optimal computing budget allocation, OCBA) 算法^[13] 以及动态仿真预算分配算法, 例如, 信息期望值 (expected value of information, EVI) 算法^[14], 知识梯度 (knowledge gradient, KG) 算法^[15], 期望改进 (expected improvement, EI) 算法^[16] 以及渐进最优分配 (asymptotically optimal allocation, AOA) 算法^[17] 等. 近年来, 排序选优方法已被拓展用以解决诸多仿真优化问题. 例如, 多目标的仿真优化问题^[18]、输入不确定性的仿真优化问题^[19]、多情景的仿真优化问题^[20]、随机仿真时间的仿真优化问题^[21]、最优子集的挑选问题^[22,23]、方案子集的排序问题^[24]、最简且足够好的方案挑选问题^[25]、存在数据删失的方案挑选问题^[26]、具有协变量的方案挑选问题^[27] 以及大规模仿真优化问题^[28,29].

现有的排序选优研究主要集中在“选优”问题, 很少有文献关注“排序”问题. 然而方案排序问题广泛地存在于诸多领域中. 例如, 在生物遗传学中, 通过对候选基因致病性进行排序有利于精确识别关键基因和深入研究遗传疾病的机理, 从而更好地治疗遗传性疾病^[30]. 在军事战略规划中, 对战时军事装备的维修需求优先级进行排序具有重要战略意义, 可大大提高战时军事装备保障效能^[31]. 在供应链管理领域中, 对订单以及配送进行排序, 可以降低订单最大延误时间以及配送成本^[32]. 在复杂网络领域中, 对复杂网络 (如, 交通网络、电力网络、互联网络等) 节点的重要性进行排序有助于识别网络中的重要性单位, 从而有效地控制复杂网络^[33]. 文献 [34,35] 基于 IZ 方法, 探究方差对仿真预算分配策略的影响, 给出了方案排序固定精度算法. 文献 [36] 基于大偏差理论以及 OCBA 方法, 综合考虑方差以及均值间差异对仿真预算分配决策过程的双重影响, 提出了方案全排序问题的最优计量预算分配 (optimal computing budget allocation for complete ranking, OCBA-CR) 算法. 文献 [37] 则进一步研究了考虑输入不确定性的方案全排序问题.

本文与文献 [36] 均研究如何优化固定仿真预算的分配, 使得方案正确排序的概率最大化问题. 文献 [36] 基于大偏差理论, 构建方案排序仿真优化问题, 推导出方案排序问题的仿真预算最优分配规则, 并根据该分配规则提出了 OCBA-CR 算法. 该算法是一个序贯分配算法, 它在每次迭代过程中将前一次迭代更新后的样本信息作为总体参数估计, 并带入仿真预算分配规则以指导本次迭代各备选方案的仿真复制过程. 仿真预算的分配决策是一个随机动态规划问题^[7]. 而 OCBA-CR 算法所使用的分配规则却是在假定完美信息 (如假定均值与方差已知) 下利用静态优化方法导出的结果. 因此, 在低置信情景 (即备选方案间均值接近且方差相对较大) 下, OCBA-CR 算法所采用的分配规则可能无法精准确定仿真预算的动态分配决策行为, 从而使得该算法出现性能表现不佳的情况^[17,38~40]. 基于上述分析, 本研究充分考虑仿真预算分配过程的动态决策特性, 采用随机动态规划方法来精准刻画仿真预算分配的动态决策过程, 并提出了渐进最优的方案全排序动态计量预算分配 (dynamic computing budget allocation for complete ranking, DCBA-CR) 算法. DCBA-CR 算法能够更加有效地捕捉仿真预算的动态分配决策行为并表现出优良且稳定的性能.

综上所述, 本文以最大化后验正确排序概率 (probability of correct ranking, PCR) 为目标, 把仿真预算分配决策过程表述为一个随机动态规划问题并将其等价转化成一个马尔科夫决策过程 (Markov decision process, MDP). 随后将后验 PCR 视为值函数, 遵循近似动态规划方法对后验 PCR 进行近似, 通过提前一步的形式, 正向地计算后验 PCR, 并提出了一种高效的动态计量预算分配策略以及相应的

表 1 符号表示
Table 1 Notation

Notation	Interpretation
k	The number of designs
T	Total number of simulation replications (budget)
\mathcal{O}^*	True designs ranking
$\hat{\mathcal{O}}_T$	Estimated designs ranking after allocating T simulation replications
$X_{i,t}$	The t -th simulation output for design i , where $i = 1, 2, \dots, k$
μ_i	Unknown mean of design i , i.e., $\mu_i = E(X_{i,t})$
σ_i^2	Known variance of design i , i.e., $\sigma_i^2 = \text{Var}(X_{i,t})$
$\mu_i^{(0)}$	Prior mean of μ_i
$(\sigma_i^{(0)})^2$	Prior variance of μ_i
$\mu_i^{(t)}$	Posterior mean of μ_i at the t -th simulation replications
$(\sigma_i^{(t)})^2$	Posterior variance of μ_i at the t -th simulation replications
t_i	The number of simulation replications allocated to design i after t simulation replications
$\bar{X}_{i,t}$	Sample mean of design i after the t -th simulation replications, i.e., $\bar{X}_{i,t} = \sum_{l=1}^{t_i} X_{i,l}/t_i$
$s_{i,t}^2$	Sample variance of design i after the t -th simulation replications, i.e., $s_{i,t}^2 = \sum_{l=1}^{t_i} (X_{i,l} - \bar{X}_{i,t})^2/(t_i - 1)$
α_i	Proportion of the total simulation budget allocated to design i , i.e., $\alpha_i = t_i/T$

DCBA-CR 算法. 最后证明了该动态分配策略能够渐进地收敛于最优仿真预算分配比例. 数值实验以及应用案例验证了本文所提出的 DCBA-CR 算法在一些低置信情景 (即备选方案间均值接近且方差相对较大) 以及高置信情景 (即备选方案间均值差距较大且方差相对较小) 下显著地提高了方案排序问题的仿真优化效率并明显地优于包括 OCBA-CR 算法在内的其他算法.

本文的结构如下: 第 2 节阐述了备选方案的排序问题, 给出了本文的基本假设, 并通过随机动态规划方法对备选方案的排序问题进行建模. 第 3 节提出了动态计量预算分配策略及其相应的算法, 并分析了该策略的一致性以及渐进最优性. 第 4 节提供了数值实验及其结果分析. 第 5 节对本文进行了总结.

2 问题描述

2.1 符号定义

本文使用的符号定义如表 1 所示.

2.2 模型建立

现考虑 k 个不同备选方案, 其中备选方案 i 的系统性能表现由其仿真输出的均值 $\mu_i \in \mathbb{R}, i \in \{1, 2, \dots, k\}$ 来度量, 且未知均值 μ_i 只能通过系统仿真进行估计. 由于系统仿真的随机性, 决策者需要大量的仿真预算 $T < \infty$ 对每个备选方案进行重复的仿真复制, 以获得其均值 μ_i 的精确估计. 定义备选方案均值的真实排序为 $\mathcal{O}^* \triangleq (\langle 1 \rangle, \langle 2 \rangle, \dots, \langle k \rangle)$, 其中 $\langle i \rangle, i \in \{1, 2, \dots, k\}$ 是均值 μ_i 的次序指标, 其满足 $\mu_{\langle 1 \rangle} > \mu_{\langle 2 \rangle} > \dots > \mu_{\langle k \rangle}$. 本研究的目标是针对备选方案的排序问题, 在固定仿真预算 T 下, 推导出一种渐进最优动态计量预算分配策略, 来最大化 k 个备选方案的正确排序概率. 在描述方案排序问题之前, 本文首先给出以下两个基本假设.

假设1 任意备选方案 $i \in \{1, 2, \dots, k\}$ 的仿真复制样本是独立同分布的并且任意两个不同备选方案 $i, j \in \{1, 2, \dots, k\}$ 之间的仿真复制也是相互独立的.

假设 1 是排序选优问题中常见的假设, 它在实践中也容易满足. 基于该假设有 $\mu_i = E(X_{i,t}), \sigma_i^2 = \text{Var}(X_{i,t}), i \in \{1, 2, \dots, k\}$.

假设2 存在 $\delta > 0$, 使得对任意的备选方案 $i \in \{1, 2, \dots, k-1\}$, 其系统性能表现都满足 $\mu_i - \mu_{i+1} > \delta$, 即本文不考虑系统性能表现差异小于 δ 的备选方案.

假设 2 是方案排序问题相关文献中经典的假设, 它使得每个备选方案之间的系统性能表现易于区分, 保证所有备选方案都可以按照分配规则获取仿真预算, 并且避免出现少数均值差异较小的备选方案占据全部仿真预算的极端情况. 从而确保随着仿真预算量趋近无穷大, 方案排序准确率可以趋近于 1^[36]. 此外, 假设 2 是参数 δ 的一个存在性假设, 即假设满足上述条件的 δ 存在即可. 在本文关于方案排序问题的研究以及求解算法实现过程中, 无须知道和设置 δ 的具体数值.

注释1 随着备选方案个数 k 的增大, 假设 2 的条件将逐渐难以被满足 (即 δ 值会越来越小). 这意味着本文所研究的方案排序方法主要适用于解决小规模到中等规模的方案排序问题.

令 $X_t = (X_{1,t}, X_{2,t}, \dots, X_{k,t}), t \leq T$ 为第 t 次仿真复制后所有备选方案的样本, 并且 X_t 服从 $Q(\cdot; \theta)$ 联合分布, 其中 $\theta \triangleq (\theta_1, \theta_2, \dots, \theta_k) \in \Theta$ 包含所有备选方案的未知参数, 特别地, 未知均值满足 $\mu = (\mu_{(1)}, \mu_{(2)}, \dots, \mu_{(k)}) \in \theta$. 设 $Q_i(\cdot; \theta_i)$ 为 $X_{i,t}$ 的边缘分布, 其中 θ_i 由备选方案 i 所有的未知参数构成. 于是根据假设 1, 则有 $Q(\cdot; \theta) = \prod_{i=1}^k Q_i(\cdot; \theta_i)$.

本文引入贝叶斯框架来量化未知参数 θ 的不确定性: 首先设未知参数 θ 服从前验分布 $F(\cdot; \zeta_0)$, 其中 ζ_0 由前验分布参数族中所有的超参数构成. 定义 t_i 为经历 t 次仿真预算分配后备选方案 i 获取的仿真复制次数, $X_i^{(t)} = (X_{i,1}, X_{i,2}, \dots, X_{i,t_i})$ 为备选方案 i 经历 t_i 次仿真复制后的样本信息, $\bar{X}_{i,t}$ 为 t 次仿真预算分配后方案 i 的样本均值, 即 $\bar{X}_{i,t} = \sum_{l=1}^{t_i} X_{i,l} / t_i, i = 1, 2, \dots, k$. 令信息集合为 $\mathcal{F}_t = \{X_1^{(t)}, X_2^{(t)}, \dots, X_k^{(t)}, \zeta_0\}$, 其中 $t \leq T, \mathcal{F}_0 \triangleq \zeta_0$. 根据贝叶斯法则, 以信息集合 \mathcal{F}_t 为条件的未知参数 θ 的后验分布形式如下:

$$F(d\theta | \mathcal{F}_t) = \frac{L(\mathcal{F}_t; \theta) F(d\theta; \zeta_0)}{\int_{\theta \in \Theta} L(\mathcal{F}_t; \theta) F(d\theta; \zeta_0)}, \quad (1)$$

其中 $L(\cdot; \theta)$ 是样本 X_t 的似然函数, 当 $F(\cdot | \mathcal{F}_t)$ 是连续分布时, $d \cdot$ 指的是勒贝格 (Lebesgue) 测度, 当 $F(\cdot | \mathcal{F}_t)$ 是离散分布时, $d \cdot$ 指的是计数测度. 相应地, 以信息集合 \mathcal{F}_t 为条件的 $X_{i,t+1}$ 的预测分布为

$$Q_i(dx_{i,t+1} | \mathcal{F}_t) = \frac{\int_{\theta \in \Theta} Q_i(dx_{i,t+1}; \theta_i) L(\mathcal{F}_t; \theta) F(d\theta; \zeta_0)}{\int_{\theta \in \Theta} L(\mathcal{F}_t; \theta) F(d\theta; \zeta_0)}. \quad (2)$$

注释2 若 θ 的前验分布 $F(\cdot; \zeta_0)$ 是样本分布 $Q(\cdot; \theta)$ 的共轭前验分布, 则 θ 的后验分布 $F(\cdot | \zeta_t)$ (其中, ζ_t 为其后验超参数) 与前验分布 $F(\cdot; \zeta_0)$ 属于同一参数族. 因此信息集合 \mathcal{F}_t 完全由 ζ_t 来确定, 并且 \mathcal{F}_t 与 ζ_t 的空间维数相同.

定义1 定义备选方案均值的后验排序表示如下:

$$\hat{O}_T(\mathcal{F}_T) \triangleq (\langle 1 \rangle_T, \langle 2 \rangle_T, \dots, \langle k \rangle_T), \quad (3)$$

其中, $\langle i \rangle_T$ 是 T 次仿真预算分配后, 备选方案后验均值 $\mu_i^{(T)}$ 的次序指标, $i = 1, 2, \dots, k$, 即其满足 $\mu_{\langle 1 \rangle_T}^{(T)} > \mu_{\langle 2 \rangle_T}^{(T)} > \dots > \mu_{\langle k \rangle_T}^{(T)}$.

本文使用后验正确排序概率 (probability of correct ranking, PCR_T) 来度量备选方案排序的精度, 其定义如下:

$$\text{PCR}_T = \mathbb{E} \left(\mathbb{I}\{\widehat{\mathcal{O}}_T(\mathcal{F}_T) = \mathcal{O}^*\} \middle| \mathcal{F}_T \right) = \mathbb{P} \left(\bigcap_{i=1}^{k-1} \{\mu_{\langle i \rangle_T} - \mu_{\langle i+1 \rangle_T} > 0\} \middle| \mathcal{F}_T \right), \quad (4)$$

其中, $\mathbb{I}\{\cdot\}$ 是指示函数, 即当事件成立时, 函数值等于 1 否则等于 0.

基于最优计量预算分配 (OCBA) 方法的观点, 决策者可以通过求解以下静态参数优化问题来指导仿真预算分配过程, 进而实现解决方案排序问题的目的:

$$\begin{aligned} \max_{\alpha_1, \alpha_2, \dots, \alpha_k} \quad & \mathbb{P} \left(\bigcap_{i=1}^{k-1} \{\bar{X}_{[i]_T} > \bar{X}_{[i+1]_T}\} \middle| \theta \right) \\ \text{s.t.} \quad & \sum_{i=1}^k \alpha_{[i]_T} = 1, \alpha_{[i]_T} > 0, i \in \{1, 2, \dots, k\}, \end{aligned} \quad (5)$$

其中, $[i]_T, i \in \{1, 2, \dots, k\}$ 通过对样本均值 $\bar{X}_{i,T}$ 进行排序获得, 即 $\bar{X}_{[1]_T, T} > \bar{X}_{[2]_T, T} > \dots > \bar{X}_{[k]_T, T}$, $\alpha_i = t_i/T$ 是 T 次仿真预算分配后, 备选方案 i 所获取的仿真预算比例. 一般情况下, 问题 (5) 的解不具有解析形式, 需要借助蒙特卡洛 (Monte Carlo) 仿真方法来获取. 文献 [36] 基于大偏差理论, 利用静态优化方法推导出了问题 (5) 的最优仿真预算分配规则, 然后根据该分配规则提出了方案全排序最优计量预算分配 (OCBA-CR) 算法. 然而仿真预算分配决策的本质是一个随机动态规划问题 [7]. 因此通过静态优化方法得到的 OCBA-CR 算法在低置信情景 (即备选方案间均值接近且方差相对较大) 下可能无法精准确定仿真预算的动态分配决策行为, 从而出现性能表现不佳的情况 [17, 38~40].

为了进一步提高方案排序问题的仿真预算分配效率, 在贝叶斯框架下, 本文充分考虑仿真预算分配决策过程的动态特性, 将该决策过程描述成一个随机动态规划问题, 具体过程如下:

定义 \mathcal{S}_t 为动态计量预算分配策略, 其具体形式为

$$\mathcal{S}_t(\cdot) = (S_1(\cdot), S_2(\cdot), \dots, S_t(\cdot)), \quad (6)$$

其中, $0 \leq t < T$, 基于信息集合 \mathcal{F}_{t-1} , $S_t(\mathcal{F}_{t-1}) = i \in \{1, 2, \dots, k\}$, 这意味着第 t 次仿真复制被分配给备选方案 i . 设 $t_i = \sum_{l=1}^t S_{i,l}(\mathcal{F}_{l-1})$, 其中 $S_{i,l}(\mathcal{F}_{l-1}) = \mathbb{I}\{S_l(\mathcal{F}_{l-1}) = i\}$. 现递归地给出动态计量预算分配策略 \mathcal{S}_T 的期望收益:

$$\begin{aligned} V_T(\mathcal{F}_T, \mathcal{S}_T) &\triangleq \mathbb{E} \left[\mathbb{I}\{\widehat{\mathcal{O}}_T(\mathcal{F}_T) = \mathcal{O}^*\} \middle| \mathcal{F}_T \right] = \int_{\theta \in \Theta} \mathbb{I}\{\widehat{\mathcal{O}}_T(\mathcal{F}_T) = \mathcal{O}^*\} F(d\theta | \mathcal{F}_T) \\ &= \mathbb{P} \left(\bigcap_{i=1}^{k-1} \{\mu_{\langle i \rangle_T} - \mu_{\langle i+1 \rangle_T} > 0\} \middle| \mathcal{F}_T \right), \end{aligned} \quad (7)$$

并且, 对 $0 \leq t < T$,

$$\begin{aligned} V_t(\mathcal{F}_t, \mathcal{S}_T) &\triangleq \mathbb{E} \left[V_{t+1}(\mathcal{F}_t \cup \{X_{i,t+1}\}; \mathcal{S}_T) \middle| \mathcal{F}_t \right] \Big|_{S_{t+1}(\mathcal{F}_t)=i} \\ &= \int_{\mathcal{X}_i} V_{t+1}(\mathcal{F}_t \cup \{x_{i,t+1}\}; \mathcal{S}_T) Q_i(dx_{i,t+1} | \mathcal{F}_t) \Big|_{S_{t+1}(\mathcal{F}_t)=i}, \end{aligned} \quad (8)$$

其中, \mathcal{X}_i 是 $X_{i,t+1}$ 的支撑集, $X_{i,t+1}$ 是将第 $t+1$ 次仿真预算分配给备选方案 i 进行仿真复制后得到的样本信息, $i \in \{1, 2, \dots, k\}$. 因此, 最优动态计量预算分配策略 \mathcal{S}_T^* 可被定义为该随机动态规划问题

的解:

$$\mathcal{S}_T^* \triangleq \arg \max_{\mathcal{S}_T} V_0(\zeta_0, \mathcal{S}_T). \quad (9)$$

上述随机动态规划问题可以看作是一个具有 $T + 1$ 时期的马尔科夫决策过程, 并且在整个 MDP 过程中, 决策者做出的行动只在 T 时期产生收益, 而在中间任意时期都不产生任何收益. 该 MDP 过程的具体结构如下:

ζ_0 : 初始状态;

\mathcal{F}_t : t 时期的状态, $0 \leq t \leq T$;

S_{t+1} : $t + 1$ 时期做出的行动, 在做出行动 S_{t+1} 之后状态 \mathcal{F}_t 转化到 \mathcal{F}_{t+1} , 其转化过程为

$$\mathcal{F}_t \rightarrow \mathcal{F}_{t+1} = \left\{ \mathcal{F}_t \cup \{X_{i,t+1}\} \right\} \Big|_{S_{t+1}(\mathcal{F}_t)=i}, \quad 0 \leq t < T. \quad (10)$$

$V_T(\mathcal{F}_T)$: 最终收益;

贝尔曼方程 (Bellman): 当 $t = T$ 时, 有

$$V_T(\mathcal{F}_T) \triangleq \mathbb{E} \left[\mathbb{I}\{\hat{\mathcal{O}}_T(\mathcal{F}_T) = \mathcal{O}^*\} \mid \mathcal{F}_T \right], \quad (11)$$

并且, 当 $0 \leq t < T$ 时, 有

$$V_t(\mathcal{F}_t) \triangleq \mathbb{E} \left[V_{t+1}(\mathcal{F}_t \cup \{X_{i,t+1}\}) \mid \mathcal{F}_t \right] \Big|_{S_{t+1}^*(\mathcal{F}_t)=i}, \quad (12)$$

其中,

$$S_{t+1}^*(\mathcal{F}_t) = \arg \max_{i=1,2,\dots,k} \mathbb{E} \left[V_{t+1}(\mathcal{F}_t \cup \{X_{i,t+1}\}) \mid \mathcal{F}_t \right]. \quad (13)$$

基于上述分析可知, 只需递归地求解贝尔曼方程 (11) 和 (12), 即可得到随机动态规划问题 (9) 的解 (即最优动态计量预算分配策略). 文献 [41] 的性质 1.3.1 为上述随机动态规划问题 (9) 的解与其对应的贝尔曼方程 (11) 和 (12) 最优解之间的等价性提供了理论支撑.

事实上, 在本文的上半部分并没有对备选方案 $i \in \{1, 2, \dots, k\}$ 的仿真输出 X_i 的分布 $Q_i(\cdot; \theta_i)$ 指定具体分布形式, 这是因为上述的所有分析对于 $Q_i(\cdot; \theta_i)$ 为一般分布的情况均适用. 下面为了便于对随机动态规划问题 (9) 的最优动态计量预算分配策略进一步地分析, 本文给出以下假设:

假设3 对所有备选方案 $i \in \{1, 2, \dots, k\}$, 都有 X_i 服从正态分布 $N(\mu_i, \sigma_i^2)$, 其中均值 μ_i 未知, 方差 σ_i^2 已知.

基于假设 3, 可知 μ_i 服从共轭正态前验分布, 即, $\mu_i \sim N(\mu_i^{(0)}, (\sigma_i^{(0)})^2)$. 同时根据式 (1) 和 (2), 可得 μ_i 的后验分布服从正态分布 $N(\mu_i^{(t)}, (\sigma_i^{(t)})^2)$, 其中

$$\mu_i^{(t)} = \left(\sigma_i^{(t)} \right)^2 \left(\frac{\mu_i^{(0)}}{(\sigma_i^{(0)})^2} + \frac{t_i \bar{X}_{i,t}}{\sigma_i^2} \right), \quad (\sigma_i^{(t)})^2 = \left(\frac{1}{(\sigma_i^{(0)})^2} + \frac{t_i}{\sigma_i^2} \right)^{-1}. \quad (14)$$

特别地, 若共轭正态前验分布满足 $\sigma_i^{(0)} \rightarrow \infty$, 则有 $\mu_i^{(t)} = \bar{X}_{i,t}$, 并称该前验分布为无信息前验分布; 若均值 μ_i 以及方差 σ_i^2 均未知, 则 μ_i 服从正态 - 伽马 (Normal-Gamma) 共轭前验分布 [42]. 在实际应用中假设 3 通常是成立的, 因为可以使用组平均数来作为系统的单个仿真输出, 根据中心极限定理可知这些仿真输出近似服从正态分布 [7, 43].

3 方案排序问题的渐进最优动态计量预算分配策略

本节遵循近似动态规划的范式, 基于单个特征值函数近似方法, 推导出一种有效的渐进最优动态计量预算分配策略. 具体地, 与文献 [17] 做法类似, 本文摒弃逆向归纳方法, 而以提前一步最大化近似值函数 (approximation of value function, VFA) 的形式, 正向地求解贝尔曼方程 (11) 和 (12), 从而最终解决随机动态规划问题 (9). 值得注意的是, 这种提前一步正向地计算近似值函数的方法能够有效地避免 MDP 问题中常见的维数灾难问题.

3.1 值函数近似

假设任意一次仿真预算分配都是最后一次, 则值函数 $V_t(\mathcal{F}_t)$ 可表示为

$$V_t(\mathcal{F}_t) = \mathbb{P} \left(\bigcap_{i=1}^{k-1} \left\{ \mu_{\langle i \rangle_t}^{(t)} - \mu_{\langle i+1 \rangle_t}^{(t)} > 0 \right\} \middle| \mathcal{F}_t \right). \quad (15)$$

基于信息集合 \mathcal{F}_t , 以及 μ_i 服从后验分布 $N(\mu_i^{(t)}, (\sigma_i^{(t)})^2)$, 有

$$\begin{aligned} & (\mu_{\langle 1 \rangle_t} - \mu_{\langle 2 \rangle_t}, \dots, \mu_{\langle i \rangle_t} - \mu_{\langle i+1 \rangle_t}, \dots, \mu_{\langle k-1 \rangle_t} - \mu_{\langle k \rangle_t}) \\ & \sim MN \left(\mu_{\langle 1 \rangle_t}^{(t)} - \mu_{\langle 2 \rangle_t}^{(t)}, \dots, \mu_{\langle i \rangle_t}^{(t)} - \mu_{\langle i+1 \rangle_t}^{(t)}, \dots, \mu_{\langle k-1 \rangle_t}^{(t)} - \mu_{\langle k \rangle_t}^{(t)}, \Sigma \right), \end{aligned}$$

其中, $i \in \{1, 2, \dots, k\}$, 协方差矩阵 $\Sigma = \Gamma' \Lambda \Gamma = (a_{ij})_{(k-1) \times (k-1)}$, “'” 为矩阵的转置变换,

$$\begin{aligned} \Gamma &= [\text{tridiag}(-1, 1, 0); -e_{k-1}] \in \mathbb{R}^{k \times (k-1)}, \\ \Lambda &= \text{diag} \left((\sigma_{\langle 1 \rangle_t}^{(t)})^2, (\sigma_{\langle 2 \rangle_t}^{(t)})^2, \dots, (\sigma_{\langle k \rangle_t}^{(t)})^2 \right), \end{aligned}$$

并且 $e_{k-1} \in \mathbb{R}^{k-1}$ 是第 $k-1$ 个元素为 1 的单位向量. 由矩阵 Γ 的列满秩性质可知, 协方差矩阵 Σ 是对称正定的. 因此, 根据 Cholesky 分解定理, 则存在上三角矩阵 $U = (u_{ij})_{(k-1) \times (k-1)}$ 使得 $\Sigma = \Gamma' \Lambda \Gamma = U' U$, 并且对任意的 $i = 1, 2, \dots, k-1$, a_{ij} 满足以下形式:

$$\begin{cases} a_{ij} = u_{1i}u_{1j} + u_{2i}u_{2j} + \dots + u_{ji}u_{jj}, & i < j, \\ a_{ii} = u_{1i}^2 + u_{2i}^2 + \dots + u_{ii}^2. \end{cases}$$

因此, 值函数 $V_t(\mathcal{F}_t)$ 可以被重新表示为

$$\begin{aligned} V_t(\mathcal{F}_t) &= \mathbb{P} \left(\bigcap_{i=1}^{k-1} \left\{ \mu_{\langle i \rangle_t} - \mu_{\langle i+1 \rangle_t} > 0 \right\} \middle| \mathcal{F}_t \right) \\ &= \mathbb{P} \left\{ \mu_{\langle 1 \rangle_t} - \mu_{\langle 2 \rangle_t} > 0, \dots, \mu_{\langle k-1 \rangle_t} - \mu_{\langle k \rangle_t} > 0 \middle| \mathcal{F}_t \right\} \\ &= \mathbb{P} \left\{ \sum_{l=1}^i u_{l,i} Z_l > \mu_{\langle i+1 \rangle_t}^{(t)} - \mu_{\langle i \rangle_t}^{(t)}, i = 1, 2, \dots, k-1 \right\} \\ &= \frac{1}{(\sqrt{2\pi})^{k-1}} \int \dots \int_{\sum_{i=1}^i u_{l,i} Z_l > \mu_{\langle i+1 \rangle_t}^{(t)} - \mu_{\langle i \rangle_t}^{(t)}} \prod_{l=1}^{k-1} \exp \left(-\frac{z_l^2}{2} \right) d_{z_1}, \dots, d_{z_{k-1}} \\ &= \frac{1}{(\sqrt{2\pi})^{k-1}} \int \dots \int_R \exp \left(-\frac{\sum_{l=1}^{k-1} z_l^2}{2} \right) d_{z_1}, d_{z_2}, \dots, d_{z_{k-1}}, \end{aligned} \quad (16)$$

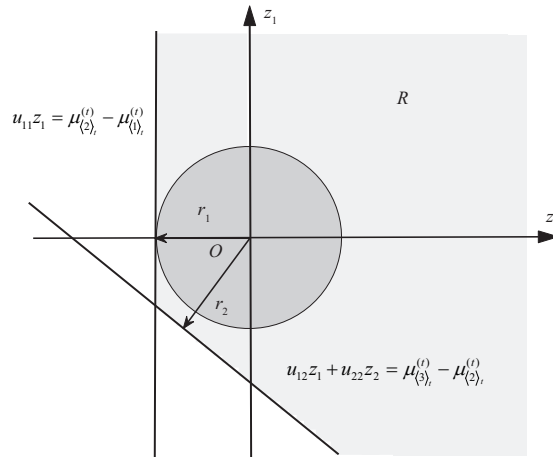


图 1 用被积函数 $\exp(-\frac{(z_1^2+z_2^2)}{2})$ 在圆形区域中的积分来近似其在整个积分区域 R 上的积分

Figure 1 Using the integral of the integrand function $\exp(-\frac{(z_1^2+z_2^2)}{2})$ in the circular region to approximate its integral on the entire integral region R

其中, $Z_l, l \in \{1, 2, \dots, k-1\}$ 是相互独立的标准正态分布随机变量, 积分区域 R 由以下超平面围成:

$$\sum_{l=1}^i u_{l,i} z_l > \mu_{(i+1)_t}^{(t)} - \mu_{(i)_t}^{(t)}, \quad i = 1, 2, \dots, k-1. \quad (17)$$

由于多维标准正态分布的密度函数相对于到原点的距离以指数阶速度衰减, 因此可以采用区域 R 中以原点为中心的内切超球大小来近似刻画其在整个积分区域 R 上的积分值 (即值函数 $V_t(\mathcal{F}_t)$). 使用内切超球大小来近似 $V_t(\mathcal{F}_t)$ 的好处在于超球与其半径的大小序列等价. 因此最大化值函数 $V_t(\mathcal{F}_t)$ 问题可被等价转化为上述内切超球半径的最大化问题.

基于上述分析, 在本研究的剩余部分将主要考虑解决最大化上述内切超球半径问题. 为了便于称谓, 不妨将上述内切超球半径直接称作值函数 $V_t(\mathcal{F}_t)$ 的单个特征 VFA [44] 并将其定义为

$$\tilde{V}_t(\mathcal{F}_t) = \min\{r_1^2(\mathcal{F}_t), r_2^2(\mathcal{F}_t), \dots, r_{k-1}^2(\mathcal{F}_t)\}, \quad (18)$$

其中, $i \in \{1, 2, \dots, k-1\}$,

$$r_i^2(\mathcal{F}_t) = \frac{(\mu_{(i)_t}^{(t)} - \mu_{(i+1)_t}^{(t)})^2}{(\sigma_{(i)_t}^{(t)})^2 + (\sigma_{(i+1)_t}^{(t)})^2}. \quad (19)$$

以 $k=3$ 为例, 图 1 中直观地展示了对值函数 $V_t(\mathcal{F}_t)$ 进行单个特征 VFA 的具体过程, 其中区域 R 是二维标准正态密度函数的积分区域. 如图 1 可知, 值函数 $V_t(\mathcal{F}_t)$ 的近似值是半径为 r_1 的最大内切圆面积.

3.2 渐进最优动态计量预算分配策略

假设 $t+1$ 步是最后一步, 现在第 t 步对第 j 个备选方案进行仿真复制, 则以提前一步的形式正向地计算 $\tilde{V}_t(\mathcal{F}_t)$, 其具体形式如下:

$$\tilde{V}_t(\mathcal{F}_t, j) = \mathbb{E} \left[\tilde{V}_{t+1}(\mathcal{F}_t \cup \{X_{j,t+1}\}) \middle| \mathcal{F}_t \right], \quad (20)$$

其中, $0 \leq t < T, j \in \{1, 2, \dots, k\}$. 为了简化式 (20) 的计算, 本文基于文献 [45] 提出的确定性等价近似方法, 对 $\tilde{V}_t(\mathcal{F}_t, j)$ 做进一步的近似:

$$\hat{V}_t(\mathcal{F}_t, j) = \tilde{V}_{t+1}(\mathcal{F}_t, \mathbb{E}[X_{j,t+1} | \mathcal{F}_t]), \quad (21)$$

其中, $\hat{V}_t(\mathcal{F}_t, j)$ 有以下 3 种情况.

情况 1: 若 $j = 1$, 则

$$\hat{V}_t(\mathcal{F}_t, 1) = \min \left[\frac{(\mu_{\langle 1 \rangle_t}^{(t)} - \mu_{\langle 2 \rangle_t}^{(t)})^2}{(\sigma_{\langle 1 \rangle_t}^{(t+1)})^2 + (\sigma_{\langle 2 \rangle_t}^{(t)})^2}, \min_{i \neq 1} \left\{ \frac{(\mu_{\langle i \rangle_t}^{(t)} - \mu_{\langle i+1 \rangle_t}^{(t)})^2}{(\sigma_{\langle i \rangle_t}^{(t)})^2 + (\sigma_{\langle i+1 \rangle_t}^{(t)})^2} \right\} \right]. \quad (22)$$

情况 2: 若 $j \in \{2, 3, \dots, k-1\}$, 则

$$\hat{V}_t(\mathcal{F}_t, j) = \min \left[\frac{(\mu_{\langle j-1 \rangle_t}^{(t)} - \mu_{\langle j \rangle_t}^{(t)})^2}{(\sigma_{\langle j-1 \rangle_t}^{(t)})^2 + (\sigma_{\langle j \rangle_t}^{(t+1)})^2}, \frac{(\mu_{\langle j \rangle_t}^{(t)} - \mu_{\langle j+1 \rangle_t}^{(t)})^2}{(\sigma_{\langle j \rangle_t}^{(t+1)})^2 + (\sigma_{\langle j+1 \rangle_t}^{(t)})^2}, \min_{i \neq j-1, i \neq j} \left\{ \frac{(\mu_{\langle i \rangle_t}^{(t)} - \mu_{\langle i+1 \rangle_t}^{(t)})^2}{(\sigma_{\langle i \rangle_t}^{(t)})^2 + (\sigma_{\langle i+1 \rangle_t}^{(t)})^2} \right\} \right]. \quad (23)$$

情况 3: 若 $j = k$, 则

$$\hat{V}_t(\mathcal{F}_t, k) = \min \left[\frac{(\mu_{\langle k-1 \rangle_t}^{(t)} - \mu_{\langle k \rangle_t}^{(t)})^2}{(\sigma_{\langle k-1 \rangle_t}^{(t)})^2 + (\sigma_{\langle k \rangle_t}^{(t+1)})^2}, \min_{i \neq k-1} \left\{ \frac{(\mu_{\langle i \rangle_t}^{(t)} - \mu_{\langle i+1 \rangle_t}^{(t)})^2}{(\sigma_{\langle i \rangle_t}^{(t)})^2 + (\sigma_{\langle i+1 \rangle_t}^{(t)})^2} \right\} \right]. \quad (24)$$

因此, 我们可以得到问题 (9) (即方案全排序问题) 的动态计量预算分配 (DCBA-CR) 策略:

$$\hat{S}_{t+1}^*(\mathcal{F}_t) = \arg \max_{j=1,2,\dots,k} \hat{V}_t(\mathcal{F}_t, j). \quad (25)$$

注释 3 $\tilde{V}_t(\mathcal{F}_t)$ 是值函数 $V_t(\mathcal{F}_t)$ 的单个特征 VFA (需注意的是 $V_t(\mathcal{F}_t)$ 与 $\tilde{V}_t(\mathcal{F}_t)$ 的数值大小并没有直接近似关系, 但两者的大小序列等价), $\tilde{V}_t(\mathcal{F}_t, j)$ 为 $\tilde{V}_t(\mathcal{F}_t)$ 的提前一步正向计算形式. 此外, 为了进一步简化计算 $\tilde{V}_t(\mathcal{F}_t, j)$, 式 (21) 利用确定性等价近似方法得到了 $\tilde{V}_t(\mathcal{F}_t, j)$ 的近似值函数 $\hat{V}_t(\mathcal{F}_t, j)$. 基于以上对 $V_t(\mathcal{F}_t)$, $\tilde{V}_t(\mathcal{F}_t)$ 以及 $\hat{V}_t(\mathcal{F}_t)$ 三者关系的总结可知, DCBA-CR 策略 (25) 通过每次最大化 $\hat{V}_t(\mathcal{F}_t, j), j \in \{1, 2, \dots, k\}$, 从而增大 $\tilde{V}_t(\mathcal{F}_t, j)$, 进而最终实现逐步最大化 $V_t(\mathcal{F}_t)$ (即后验 PCR) 的目的.

实际上, 使用 DCBA-CR 策略对仿真预算进行分配, 能够实现一致性以及渐进最优性, 具体情况见以下定理.

定理 1 当 $t \rightarrow \infty$ 时, 则

$$\lim_{t \rightarrow \infty} \hat{O}_t(\mathcal{F}_t) \rightarrow \mathcal{O}^*, \quad \text{a.s.} \quad (26)$$

证明 首先证明按照 DCBA-CR 策略 (25) 对仿真预算进行分配, 当 $t \rightarrow \infty$ 时, 几乎处处 (almost surely, a.s.) 伴随着每一个备选方案所获取的仿真复制次数 $t_i \rightarrow \infty$. 现定义,

$$D = \left\{ i \in \{1, 2, \dots, k\} : \text{方案 } i \text{ 被分配的仿真复制次数几乎处处有限次, 即 } \lim_{t \rightarrow \infty} t_i < \infty, \text{ a.s.} \right\}.$$

下面, 利用反证法来证明 $D = \emptyset$. 假设 $\{1, 2, \dots, k\} \cap D \neq \emptyset$, 则对任意 $j \in D$ 都有

$$\lim_{t \rightarrow \infty} (\sigma_{\langle j \rangle_t}^{(t)})^2 > 0, \quad \lim_{t \rightarrow \infty} [(\sigma_{\langle j \rangle_t}^{(t)})^2 - (\sigma_{\langle j \rangle_{t+1}}^{(t+1)})^2] > 0, \quad (27)$$

以及存在 $w \in D$ 使得

$$\lim_{t \rightarrow \infty} \left[\widehat{V}_t(\mathcal{F}_t; w) - \widetilde{V}_t(\mathcal{F}_t) \right] > 0 \quad \text{a.s.}, \quad (28)$$

并且, 对任意 $j' \in \{1, 2, \dots, k\} - D$, 其中 “-” 为集合的差运算, 都满足

$$\lim_{t \rightarrow \infty} (\sigma_{\langle j' \rangle_t}^{(t)})^2 = 0, \quad \lim_{t \rightarrow \infty} \left[(\sigma_{\langle j' \rangle_t}^{(t)})^2 - (\sigma_{\langle j' \rangle_{t+1}}^{(t+1)})^2 \right] = 0, \quad (29)$$

以及

$$\lim_{t \rightarrow \infty} \left[\widehat{V}_t(\mathcal{F}_t; j') - \widetilde{V}_t(\mathcal{F}_t) \right] = 0 \quad \text{a.s.} \quad (30)$$

显然式 (28) 和 (30) 与 DCBA-CR 策略 (25) 矛盾. 因此, 则有 $\lim_{t \rightarrow \infty} t_i = \infty, \text{a.s.}, i \in \{1, 2, \dots, k\}$. 遵循大数定律 (law of large numbers, LLN), 定理 1 得证.

注释 4 定理 1 证明了使用 DCBA-CR 策略对仿真预算进行分配能够实现一致性, 即, 当 $t \rightarrow \infty$ 时, 由仿真估计得来的方案排序 $\widehat{\mathcal{O}}_t(\mathcal{F}_t)$ 与真正的方案排序 \mathcal{O}^* 一致. 此外, 由定理 1 的证明可知, 当 $t \rightarrow \infty$ 时, 则有 $t_i \rightarrow \infty, \text{a.s.}, i \in \{1, 2, \dots, k\}$. 因此, 根据 LLN 可得 $\lim_{t \rightarrow \infty} \mu_i^{(t)} = \mu_i, \lim_{t \rightarrow \infty} (\sigma_i^{(t)})^2 = \sigma_i^2/t_i$.

基于上述分析, 为了便于对 DCBA-CR 策略相关性质的进一步讨论, 现使用 μ_i 以及 σ_i^2/t_i 来分别替代式 (22)~(24) 中的 $\mu_i^{(t)}, (\sigma_i^{(t)})^2$. 因此 DCBA-CR 策略 (25), 即 $\widehat{S}_{t+1}^*(\mathcal{F}_t) = \arg \max_{j=1, 2, \dots, k} \widehat{V}_t(\mathcal{F}_t, j)$ 中 $\widehat{V}_t(\mathcal{F}_t, j)$ 可重写为以下形式.

情况 1: 若 $j = 1$, 则

$$\widehat{V}_t(\mathcal{F}_t, 1) = \min \left[\frac{(\mu_{\langle 1 \rangle} - \mu_{\langle 2 \rangle})^2}{\sigma_{\langle 1 \rangle}^2/t_{\langle 1 \rangle} + \sigma_{\langle 2 \rangle}^2/t_{\langle 2 \rangle}}, \min_{i \neq 1} \left\{ \frac{(\mu_{\langle i \rangle} - \mu_{\langle i+1 \rangle})^2}{\sigma_{\langle i \rangle}^2/t_{\langle i \rangle} + \sigma_{\langle i+1 \rangle}^2/t_{\langle i+1 \rangle}} \right\} \right]. \quad (31)$$

情况 2: 若 $j \in \{2, 3, \dots, k-1\}$, 则

$$\widehat{V}_t(\mathcal{F}_t, j) = \min \left[\frac{(\mu_{\langle j-1 \rangle} - \mu_{\langle j \rangle})^2}{\sigma_{\langle j-1 \rangle}^2/t_{\langle j-1 \rangle} + \sigma_{\langle j \rangle}^2/t_{\langle j \rangle} + 1}, \frac{(\mu_{\langle j \rangle} - \mu_{\langle j+1 \rangle})^2}{\sigma_{\langle j \rangle}^2/t_{\langle j \rangle} + 1 + \sigma_{\langle j+1 \rangle}^2/t_{\langle j+1 \rangle}}, \min_{i \neq j-1, i \neq j} \left\{ \frac{(\mu_{\langle i \rangle} - \mu_{\langle i+1 \rangle})^2}{\sigma_{\langle i \rangle}^2/t_{\langle i \rangle} + \sigma_{\langle i+1 \rangle}^2/t_{\langle i+1 \rangle}} \right\} \right]. \quad (32)$$

情况 3: 若 $j = k$, 则

$$\widehat{V}_t(\mathcal{F}_t, k) = \min \left[\frac{(\mu_{\langle k-1 \rangle} - \mu_{\langle k \rangle})^2}{\sigma_{\langle k-1 \rangle}^2/t_{\langle k-1 \rangle} + \sigma_{\langle k \rangle}^2/t_{\langle k \rangle} + 1}, \min_{i \neq k-1} \left\{ \frac{(\mu_{\langle i \rangle} - \mu_{\langle i+1 \rangle})^2}{\sigma_{\langle i \rangle}^2/t_{\langle i \rangle} + \sigma_{\langle i+1 \rangle}^2/t_{\langle i+1 \rangle}} \right\} \right]. \quad (33)$$

事实上, 使用 DCBA-CR 策略 (25) 对所有备选方案进行 $t \rightarrow \infty$ 次仿真预算分配后, 备选方案获取的仿真预算分配比例序列 $(\alpha_1^{(t)}, \alpha_2^{(t)}, \dots, \alpha_k^{(t)})$ 的聚点都满足文献 [36] 提出的方案排序最优仿真预算分配条件, 该性质的具体内容以及证明过程如以下所示.

定理 2 定义文献 [36] 基于大偏差理论得到的最优仿真预算分配比例为 $(\alpha_1^*, \dots, \alpha_k^*)$, 其满足以下最优分配条件:

$$\begin{aligned} G_1(\alpha_{\langle 1 \rangle}^*, \alpha_{\langle 2 \rangle}^*) &= \dots = \min \left\{ G_{i-1}(\alpha_{\langle i-1 \rangle}^*, \alpha_{\langle i \rangle}^*), G_i(\alpha_{\langle i \rangle}^*, \alpha_{\langle i+1 \rangle}^*) \right\} \\ &= \dots = G_{k-1}(\alpha_{\langle k-1 \rangle}^*, \alpha_{\langle k \rangle}^*), \quad i = 2, \dots, k-2, \end{aligned} \quad (34)$$

$$\sum_{i=1}^k \alpha_i^* = 1, \alpha_i^* > 0, \quad \forall i = 1, 2, \dots, k,$$

其中, $\langle i \rangle, i = 1, 2, \dots, k$ 是 μ_i 的排序指标, 即 $\mu_{\langle 1 \rangle} > \mu_{\langle 2 \rangle} > \dots > \mu_{\langle k \rangle}$,

$$G_i(\alpha_{\langle i \rangle}^*, \alpha_{\langle i+1 \rangle}^*) = \frac{(\mu_{\langle i \rangle} - \mu_{\langle i+1 \rangle})^2}{\sigma_{\langle i \rangle}^2 / \alpha_{\langle i \rangle}^* + \sigma_{\langle i+1 \rangle}^2 / \alpha_{\langle i+1 \rangle}^*}, \quad (35)$$

则有, $(\alpha_1^{(t)}, \alpha_2^{(t)}, \dots, \alpha_k^{(t)})$ 的所有聚点都满足等式 (34).

证明 首先由波尔查诺 - 魏尔斯特拉斯 (Bolzano-Weierstrass) 定理^[46]可知, $(\alpha_1^{(t)}, \alpha_2^{(t)}, \dots, \alpha_k^{(t)})$ 至少存在一个聚点. 现定义 $(\hat{\alpha}_1, \hat{\alpha}_2, \dots, \hat{\alpha}_k)$ 为 $(\alpha_1^{(t)}, \alpha_2^{(t)}, \dots, \alpha_k^{(t)})$ 的任意一聚点, 其满足 $\sum_{i=1}^k \hat{\alpha}_i = 1, \hat{\alpha}_i > 0$ 对任意的 $i = 1, 2, \dots, k$. 不失一般性, 假设 $(\alpha_1^{(t)}, \alpha_2^{(t)}, \dots, \alpha_k^{(t)})$ 收敛到 $(\hat{\alpha}_1, \hat{\alpha}_2, \dots, \hat{\alpha}_k)$. 首先要证明对任意的 $i \in \{1, 2, \dots, k\}$ 都有 $\hat{\alpha}_i > 0$. 观察式 (35) 可得

$$\begin{aligned} \lim_{t \rightarrow \infty} \left. \frac{\partial G_i(x, \alpha_{\langle i+1 \rangle}^{(t)})}{\partial x} \right|_{x=\alpha_{\langle i \rangle}^{(t)}} &= \lim_{t \rightarrow +\infty} \left[\frac{(\mu_{\langle i \rangle} - \mu_{\langle i+1 \rangle})^2}{\sigma_{\langle i \rangle}^2 / (t_{\langle i \rangle} + 1) + \sigma_{\langle i+1 \rangle}^2 / t_{\langle i+1 \rangle}} - \frac{(\mu_{\langle i \rangle} - \mu_{\langle i+1 \rangle})^2}{\sigma_{\langle i \rangle}^2 / t_{\langle i \rangle} + \sigma_{\langle i+1 \rangle}^2 / t_{\langle i+1 \rangle}} \right] \\ &= \lim_{t \rightarrow +\infty} t \left[\frac{(\mu_{\langle i \rangle} - \mu_{\langle i+1 \rangle})^2}{\sigma_{\langle i \rangle}^2 / (\alpha_{\langle i \rangle}^{(t)} + 1/t) + \sigma_{\langle i+1 \rangle}^2 / \alpha_{\langle i+1 \rangle}^{(t)}} - \frac{(\mu_{\langle i \rangle} - \mu_{\langle i+1 \rangle})^2}{\sigma_{\langle i \rangle}^2 / \alpha_{\langle i \rangle}^{(t)} + \sigma_{\langle i+1 \rangle}^2 / \alpha_{\langle i+1 \rangle}^{(t)}} \right] \\ &= \lim_{t \rightarrow +\infty} \left(\frac{\sigma_{\langle i \rangle}}{\alpha_{\langle i \rangle}^{(t)}} \right)^2 \frac{(\mu_{\langle i \rangle} - \mu_{\langle i+1 \rangle})^2}{(\sigma_{\langle i \rangle}^2 / \alpha_{\langle i \rangle}^{(t)} + \sigma_{\langle i+1 \rangle}^2 / \alpha_{\langle i+1 \rangle}^{(t)})^2}, \end{aligned} \quad (36)$$

以及

$$\begin{aligned} \lim_{t \rightarrow \infty} \left. \frac{\partial G_i(\alpha_{\langle i \rangle}^{(t)}, x)}{\partial x} \right|_{x=\alpha_{\langle i+1 \rangle}^{(t)}} &= \lim_{t \rightarrow +\infty} \left[\frac{(\mu_{\langle i \rangle} - \mu_{\langle i+1 \rangle})^2}{\sigma_{\langle i \rangle}^2 / t_{\langle i \rangle} + \sigma_{\langle i+1 \rangle}^2 / (t_{\langle i+1 \rangle} + 1)} - \frac{(\mu_{\langle i \rangle} - \mu_{\langle i+1 \rangle})^2}{\sigma_{\langle i \rangle}^2 / t_{\langle i \rangle} + \sigma_{\langle i+1 \rangle}^2 / t_{\langle i+1 \rangle}} \right] \\ &= \lim_{t \rightarrow +\infty} t \left[\frac{(\mu_{\langle i \rangle} - \mu_{\langle i+1 \rangle})^2}{\sigma_{\langle i \rangle}^2 / \alpha_{\langle i \rangle}^{(t)} + \sigma_{\langle i+1 \rangle}^2 / (\alpha_{\langle i+1 \rangle}^{(t)} + 1/t)} - \frac{(\mu_{\langle i \rangle} - \mu_{\langle i+1 \rangle})^2}{\sigma_{\langle i \rangle}^2 / \alpha_{\langle i \rangle}^{(t)} + \sigma_{\langle i+1 \rangle}^2 / \alpha_{\langle i+1 \rangle}^{(t)}} \right] \\ &= \lim_{t \rightarrow +\infty} \left(\frac{\sigma_{\langle i+1 \rangle}}{\alpha_{\langle i+1 \rangle}^{(t)}} \right)^2 \frac{(\mu_{\langle i \rangle} - \mu_{\langle i+1 \rangle})^2}{(\sigma_{\langle i \rangle}^2 / \alpha_{\langle i \rangle}^{(t)} + \sigma_{\langle i+1 \rangle}^2 / \alpha_{\langle i+1 \rangle}^{(t)})^2}. \end{aligned} \quad (37)$$

采用反证法, 假设存在一备选方案 $i \in \{1, 2, \dots, k\}$ 满足以下任意一种情况:

(i) 若 $i \neq k$ 并且 $\hat{\alpha}_i > 0, \hat{\alpha}_{i+1} = 0$, 则有

$$\lim_{t \rightarrow \infty} \left. \frac{\partial G_i(x, \alpha_{\langle i+1 \rangle}^{(t)})}{\partial x} \right|_{x=\alpha_{\langle i \rangle}^{(t)}} = 0, \quad \lim_{t \rightarrow \infty} \left. \frac{\partial G_i(\alpha_{\langle i \rangle}^{(t)}, x)}{\partial x} \right|_{x=\alpha_{\langle i+1 \rangle}^{(t)}} > 0. \quad (38)$$

(ii) 若 $i = k$ 并且 $\hat{\alpha}_{i-1} = 0, \hat{\alpha}_i > 0$, 则有

$$\lim_{t \rightarrow \infty} \left. \frac{\partial G_{i-1}(x, \alpha_{\langle i \rangle}^{(t)})}{\partial x} \right|_{x=\alpha_{\langle i-1 \rangle}^{(t)}} > 0, \quad \lim_{t \rightarrow \infty} \left. \frac{\partial G_i(\alpha_{\langle i-1 \rangle}^{(t)}, x)}{\partial x} \right|_{x=\alpha_{\langle i \rangle}^{(t)}} = 0. \quad (39)$$

式 (38) 和 (39) 均与 DCBA-CR 策略 (25) 存在矛盾. 因此, 对任意 $i \in \{1, 2, \dots, k\}$, 都有 $\hat{\alpha}_i > 0$. 现证明 $(\hat{\alpha}_1, \hat{\alpha}_2, \dots, \hat{\alpha}_k)$ 满足等式 (34). 仍然采用反证法, 假设 $(\hat{\alpha}_1, \hat{\alpha}_2, \dots, \hat{\alpha}_k)$ 不满足等式 (34), 则可能发生以下 4 种情况.

情况 1:

$$G_1(\hat{\alpha}_{\langle 1 \rangle}, \hat{\alpha}_{\langle 2 \rangle}) > \min \{G_{i-1}(\hat{\alpha}_{\langle i-1 \rangle}, \hat{\alpha}_{\langle i \rangle}), G_i(\hat{\alpha}_{\langle i \rangle}, \hat{\alpha}_{\langle i+1 \rangle})\}, \quad (40)$$

其中 $i \in \{2, \dots, k-2\}$.

情况 2:

$$\min \{G_{i-1}(\hat{\alpha}_{\langle i-1 \rangle}, \hat{\alpha}_{\langle i \rangle}), G_i(\hat{\alpha}_{\langle i \rangle}, \hat{\alpha}_{\langle i+1 \rangle})\} > \min \{G_{j-1}(\hat{\alpha}_{\langle j-1 \rangle}, \hat{\alpha}_{\langle j \rangle}), G_j(\hat{\alpha}_{\langle j \rangle}, \hat{\alpha}_{\langle j+1 \rangle})\}, \quad (41)$$

其中 $i \neq j$ 和 $i, j \in \{2, \dots, k-2\}$.

情况 3:

$$\min \{G_{i-1}(\hat{\alpha}_{\langle i-1 \rangle}, \hat{\alpha}_{\langle i \rangle}), G_i(\hat{\alpha}_{\langle i \rangle}, \hat{\alpha}_{\langle i+1 \rangle})\} > G_{k-1}(\hat{\alpha}_{\langle k-1 \rangle}, \hat{\alpha}_{\langle k \rangle}), \quad (42)$$

其中 $i \in \{2, \dots, k-2\}$.

情况 4:

$$G_1(\hat{\alpha}_{\langle 1 \rangle}, \hat{\alpha}_{\langle 2 \rangle}) > G_{k-1}(\hat{\alpha}_{\langle k-1 \rangle}, \hat{\alpha}_{\langle k \rangle}). \quad (43)$$

上述 (40)~(43) 任意一种情况成立, 都与 $(\alpha_1^{(t)}, \alpha_2^{(t)}, \dots, \alpha_k^{(t)})$ 收敛于 $(\hat{\alpha}_1, \hat{\alpha}_2, \dots, \hat{\alpha}_k)$ 的假设矛盾. 现以情况 1 为例来证明上述事实, 其余情况的证明过程比较类似, 在此省略. 假设存在 $i \in \{1, 2, \dots, k\}$ 使得情况 1 成立, 根据 G_1, G_i 以及 G_{i-1} 在 $(0, 1) \times (0, 1)$ 的连续性, 则存在 $N_0 \in \mathbb{N}^+$ 使得对任意的 $t > N_0$, 都有

$$G_1(\alpha_{\langle 1 \rangle}^{(t)}, \alpha_{\langle 2 \rangle}^{(t)}) > \min \left\{ G_{i-1} \left(\alpha_{\langle i-1 \rangle}^{(t)}, \alpha_{\langle i \rangle}^{(t)} \right), G_i \left(\alpha_{\langle i \rangle}^{(t)}, \alpha_{\langle i+1 \rangle}^{(t)} \right) \right\} \quad (44)$$

成立. 于是, 由 $\alpha_{\langle i \rangle}^{(t)}, \alpha_{\langle i+1 \rangle}^{(t)}$ 以及 $G_i(\alpha_{\langle i \rangle}^{(t)}, \alpha_{\langle i+1 \rangle}^{(t)})$ 的定义, 可得

$$\frac{(\mu_{\langle 1 \rangle} - \mu_{\langle 2 \rangle})^2}{\sigma_{\langle 1 \rangle}^2/t_{\langle 1 \rangle} + \sigma_{\langle 2 \rangle}^2/t_{\langle 2 \rangle}} > \min \left\{ \frac{(\mu_{\langle i-1 \rangle} - \mu_{\langle i \rangle})^2}{\sigma_{\langle i-1 \rangle}^2/t_{\langle i-1 \rangle} + \sigma_{\langle i \rangle}^2/t_{\langle i \rangle}}, \frac{(\mu_{\langle i \rangle} - \mu_{\langle i+1 \rangle})^2}{\sigma_{\langle i \rangle}^2/t_{\langle i \rangle} + \sigma_{\langle i+1 \rangle}^2/t_{\langle i+1 \rangle}} \right\}. \quad (45)$$

因此, 遵循 DCBA-CR 策略 (25) 以及式 (31)~(33), 则有备选方案 $\langle i \rangle$ 将被分配仿真预算, 而备选方案 $\langle 1 \rangle$ 将被终止仿真复制, 这与 $(\alpha_1^{(t)}, \alpha_2^{(t)}, \dots, \alpha_k^{(t)})$ 收敛到 $(\hat{\alpha}_1, \hat{\alpha}_2, \dots, \hat{\alpha}_k)$ 矛盾, 因此 $(\hat{\alpha}_1, \hat{\alpha}_2, \dots, \hat{\alpha}_k)$ 满足式 (34), 定理 2 得证.

当等式 (34) 有唯一解时, DCBA-CR 策略能实现渐进最优性, 具体如下:

推论 1 假设等式 (34) 存在唯一解为 $(\alpha_1^*, \dots, \alpha_k^*)$, 则有

$$\lim_{t \rightarrow \infty} \alpha_i^{(t)} \rightarrow \alpha_i^* \quad \text{a.s.}, \quad (46)$$

其中, $\forall i \in \{1, 2, \dots, k\}, \alpha_i^{(t)} = t_i/t$ 为经历 t 次仿真预算分配后备选方案 i 所获取的仿真预算比例.

算法 1 DCBA-CR

输入: Total number of designs k , total number of simulation replications T , the initial simulation replication number n_0 .

- 1: **Initialize:** $t = 0$ and perform n_0 simulation replications for each design;
- 2: **while** $t < T - k \times n_0$, **do**
- 3: **Update:** Calculate the posterior variances $(\sigma_i^{(t)})^2$ and posterior means $\mu_i^{(t)}$ with sample means $\bar{X}_{i,t}$ and sample variances $s_{i,t}^2$, for $i = 1, 2, \dots, k$, according to (14);
- 4: **Allocation:** Calculate $\hat{S}_{t+1}^*(\mathcal{F}_t) = \arg \max_{j=1,2,\dots,k} \hat{V}_t(\mathcal{F}_t, j)$, $i = 1, 2, \dots, k$, according to (22) ~ (24) as the allocated design;
- 5: **Simulation:** Run an additional simulation for the allocated design;
- 6: $t \leftarrow t + 1$;
- 7: **end while**

输出: Result of ranking of all designs.

3.3 动态计量预算分配算法

为了验证 DCBA-CR 策略的有效性, 基于式 (22)~(25), 本文提出了方案全排序动态计量预算分配 (DCBA-CR) 算法. 由于备选方案 i 的 μ_i 未知并且服从已知前验分布 $N(\mu_i^{(0)}, (\sigma_i^{(0)})^2)$, 其中 $i \in \{1, 2, \dots, k\}$, $\mu_i^{(0)}$ 是前验均值, $(\sigma_i^{(0)})^2$ 是前验方差. 因此, 在初始阶段需要对每个备选方案进行初始仿真复制 n_0 次并基于初始仿真输出样本以及前验信息来估计出 $\hat{S}_{n_0+1}^*(\mathcal{F}_{n_0}) = \arg \max_{j=1,2,\dots,k} \hat{V}_{n_0}(\mathcal{F}_{n_0}, j)$. 后续地, 每次对仿真预算进行分配时, 算法都会计算 $\hat{S}_{t+1}^*(\mathcal{F}_t) = \arg \max_{j=1,2,\dots,k} \hat{V}_t(\mathcal{F}_t, j)$ 并将本次的仿真预算分配给 $\hat{S}_{t+1}^*(\mathcal{F}_t)$. 此外, 由 DCBA-CR 策略的定义可知, 每次将仿真预算分配给备选方案 $\hat{S}_{t+1}^*(\mathcal{F}_t)$ 时, $\tilde{V}_t(\mathcal{F}_t)$ 的增量最大. 这意味着在 DCBA-CR 策略下, $\tilde{V}_t(\mathcal{F}_t)$ 关于 t 具有单增性质. 又由值函数 $V_t(\mathcal{F}_t)$ 与 $\tilde{V}_t(\mathcal{F}_t)$ 的大小序列等价, 因此基于 DCBA-CR 策略开发的 DCBA-CR 算法可以达到逐步最大化 $V_t(\mathcal{F}_t)$ (即后验 PCR) 的目的. 关于 $\hat{S}_{t+1}^*(\mathcal{F}_t)$ 的详细计算步骤如下.

步骤 1: 计算出样本均值 $\bar{X}_{i,t}$ (即 $\bar{X}_{i,t} = \sum_{l=1}^{t_i} X_{i,l}/t_i$) 和样本方差 $s_{i,t}^2$ (即 $s_{i,t}^2 = \sum_{l=1}^{t_i} (X_{i,l} - \bar{X}_{i,t})^2 / (t_i - 1)$), 其中, t_i 是经过 t 次仿真预算分配后, 备选方案 i 所获取的仿真复制次数;

步骤 2: 根据式 (14) 计算出后验均值 $\mu_i^{(t)}$ 以及后验方差 $(\sigma_i^{(t)})^2$;

步骤 3: 根据式 (22)~(24), 计算出 $\hat{S}_{t+1}^*(\mathcal{F}_t) = \arg \max_{j=1,2,\dots,k} \hat{V}_t(\mathcal{F}_t, j)$;

注意, 在步骤 2 中, 我们使用样本方差 $s_{i,t}^2$ 代替总体方差 σ_i^2 , $i \in \{1, 2, \dots, k\}$ 来计算后验均值以及方差. 因此, DCBA-CR 算法如算法 1 所示.

4 数值实验

本节通过一系列数值实验, 来验证 DCBA-CR 算法的有效性. 将 DCBA-CR 算法与均等分配 (equal allocation, EA) 算法、方差比例 (proportional to variance, PTV) 算法、方案子集排序最优计量预算分配 (optimal computing budget allocation for subsets ranking, OCBA-SR) 算法以及方案排序最优计量预算分配 (OCBA-CR) 算法进行对比实验, 具体如下.

- EA 算法是最基本的仿真预算分配算法, 它常被用作其他算法的比较基准, 该算法均分仿真预算 T 来估计每个备选方案的均值 μ_i , $i = 1, 2, \dots, k$, 即每个备选方案大约被分配 T/k 次仿真复制.

- PTV 算法也是排序选优文献中常被用于比较的算法, 它在初始阶段对每个备选方案都进行 n_0 次仿真复制, 并且在每期仿真预算分配过程中都以样本方差来估计总体方差, 该方法规定每个备选方案 i 的仿真预算分配比例与其样本方差 s_i^2 成反比, 即 $\alpha_i/s_i^2 = \alpha_j/s_j^2, \forall i, j \in \{1, 2, \dots, k\}$. 特别地, 当每个备选方案的方差均相等时, PTV 算法与 EA 算法等价^[11].

• OCBA-SR 算法由文献 [24] 提出, 该算法的仿真预算分配规则为

$$\frac{\alpha_{h_1}}{\alpha_{h_2}} = \frac{\sigma_{h_1}^2 / \min\{(c_{g_1} - \bar{X}_{h_1})^2, (c_{g_1-1} - \bar{X}_{h_1})^2\}}{\sigma_{h_2}^2 / \min\{(c_{g_2} - \bar{X}_{h_2})^2, (c_{g_2-1} - \bar{X}_{h_2})^2\}}, \quad (47)$$

其中, $\forall h_1, h_2 \in \{1, 2, \dots, k\}$, 常数 c_{g_i} 以及 $c_{g_{i-1}}$ 与 h_i 相邻, $i = 1, 2$. 并且 $c_{g_i} = (\bar{X}_{g_i} + \bar{X}_{g_{i+1}})/2$, $\bar{X}_{g_i} = \max_{j \in S_{g_i}} \bar{X}_j$, $\bar{X}_{g_{i+1}} = \min_{j \in S_{g_{i+1}}} \bar{X}_j$, $g_i \in \{1, 2, \dots, n-1\}$, n 为方案子集 S_{g_i} 的个数. 特别地, 当 $n = k$ (即每个方案子集中只有 1 个备选方案) 时, OCBA-SR 算法本质上是对备选方案进行排序.

• OCBA-CR 算法由文献 [36] 提出, 它的仿真预算分配规则为

$$\frac{\alpha_i}{\alpha_j} = \frac{\sigma_i^2 / \min\{(\bar{X}_{i+1} - \bar{X}_i)^2, (\bar{X}_i - \bar{X}_{i-1})^2\}}{\sigma_j^2 / \min\{(\bar{X}_{j+1} - \bar{X}_j)^2, (\bar{X}_j - \bar{X}_{j-1})^2\}}, \quad (48)$$

其中, $\forall i, j \in \{1, 2, \dots, k\}$.

注释5 当使用 OCBA-SR 算法对备选方案进行排序 (即需要排序子集的个数 $n = k$) 时, 其与 OCBA-CR 算法等价. 具体原因如下: 根据 OCBA-SR 算法的定义, 则有 $g_i = h_i$, $c_{g_i} = (\bar{X}_{h_i} + \bar{X}_{h_{i+1}})/2$ 以及 $c_{g_{i-1}} = (\bar{X}_{h_{i-1}} + \bar{X}_{h_i})/2$, $i = 1, 2$. 现将 $g_i, c_{g_i}, c_{g_{i-1}}$ 以及 $n = k$ 带入 OCBA-SR 算法的仿真预算分配比例式 (47) 中, 可得

$$\frac{\alpha_{h_1}}{\alpha_{h_2}} = \frac{\sigma_{h_1}^2 / \min\{(c_{g_1} - \bar{X}_{h_1})^2, (c_{g_1-1} - \bar{X}_{h_1})^2\}}{\sigma_{h_2}^2 / \min\{(c_{g_2} - \bar{X}_{h_2})^2, (c_{g_2-1} - \bar{X}_{h_2})^2\}} = \frac{\sigma_{h_1}^2 / \min\{(\bar{X}_{h_1} - \bar{X}_{h_{1+1}})^2, (\bar{X}_{h_{1-1}} - \bar{X}_{h_1})^2\}}{\sigma_{h_2}^2 / \min\{(\bar{X}_{h_2} - \bar{X}_{h_{2+1}})^2, (\bar{X}_{h_{2-1}} - \bar{X}_{h_2})^2\}}, \quad (49)$$

其中, $\forall h_1, h_2 \in \{1, 2, \dots, k\}$. 结合式 (48) 可知, OCBA-SR 算法与 OCBA-CR 算法等价.

为了客观地对上述算法进行对比实验, 在每组数值实验中, 本文设置每期仿真预算增量为 $\Delta = 1$, 每个备选方案的初始仿真复制次数为 $n_0 = 20$, 独立重复实验的次数为 10000 次 (其中, 在每次进行独立实验之前, 首先要通过备选方案均值的前验分布来随机生成 1 组 $\mu = \{\mu_1, \mu_2, \dots, \mu_k\}$ 作为本次实验中备选方案的均值), 并且在实验中使用后验 PCR 以及 PCR 来分别衡量 DCBA-CR 算法, OCBA-SR 算法, OCBA-CR 算法, PTV 算法和 EA 算法的性能表现 (其中, 后验 PCR 和 PCR 通过分别计算 10000 次独立实验中后验均值排序等于真实均值排序事件发生的频率值和样本均值排序等于真实均值排序事件发生的频率值来获取).

此外, 在所有实验 1~7 中, OCBA-SR 算法、OCBA-CR 算法、PTV 算法以及 EA 算法对每期仿真预算增量 $\Delta = 1$ 的分配方式, 具体如下: 为了方便叙述, 首先设 t (其中, $0 \leq t < T$) 次仿真预算分配后, 方案 i 被分配的仿真预算量为 $t_i, i \in \{1, 2, \dots, k\}$, 现使用上述算法对 $t+1$ 次仿真预算进行分配.

(1) OCBA-CR 算法、OCBA-SR 算法以及 PTV 算法.

Step 1: 使用样本均值以及样本方差来代替总体均值以及总体方差的形式, 通过上述算法的最优仿真预算分配规则计算出每个备选方案 i 应被分配的仿真预算分配比例 β_i , 定义 $t'_i = (t+1)\beta_i, i \in \{1, 2, \dots, k\}$;

Step 2: 令 $a_i = t'_i - t_i, i \in \{1, 2, \dots, k\}$, 取 $i^* = \arg \max_{i=1, 2, \dots, k} a_i$, 然后将第 $t+1$ 次仿真预算分配给备选方案 i^* , 至此第 $t+1$ 次仿真预算分配完成.

(2) EA 算法.

Step 1: 计算出每个备选方案 i 应被分配的仿真预算量为 $t'_i = (t+1)/k, i \in \{1, 2, \dots, k\}$, 然后对 t'_i 进行取整得到 t_i^\dagger ;

Step 2: 令 $b = (t+1) - \sum_{i=1}^k t_i^\dagger$, 如果 $b < 0$, 则将 $t+1$ 次仿真预算分配给备选方案 k , 否则将 $t+1$ 次仿真预算分配给备选方案 b , 至此第 $t+1$ 次仿真预算分配完成.

表 2 低置信情景下, 数值算例的参数设置

Table 2 Parameter setting for numerical examples under low confidence scenario

Experiment	k	T	Prior distribution of mean μ_i	Variance of output distribution σ_i^2
1	10	20000	$N(0, 3^2)$	i^2
2	10	20000	$N(0, (i/\sqrt{10})^2)$	$(i/\sqrt{2})^2$
3	10	20000	$N(0, ((11-i)/\sqrt{3})^2)$	$(3(11-i)/2)^2$

表 3 高置信情景下, 数值算例的参数设置

Table 3 Parameter setting for numerical examples under high confidence scenario

Experiment	k	T	Prior distribution of mean μ_i	Variance of output distribution σ_i^2
4	20	5000	$N(5i, 0.7i)$	7^2
5	20	5000	$N(5i, 0.7i)$	$(0.7i)^2$
6	20	5000	$N(5i, 0.7i)$	$(15 - 0.7i)^2$
7	60	5000	$N(0.5i^2, 3^2)$	$(\sqrt{61-i})^2$

4.1 算法性能对比实验

为了比较 DCBA-CR 算法、OCBA-SR 算法、OCBA-CR 算法、PTV 算法以及 EA 算法的性能, 本文考虑使用一系列低置信情景 (即备选方案间均值接近且方差相对较大) 下的参数设置以及高置信情景 (即备选方案间均值差距较大且方差相对较小) 下的参数设置进行数值实验, 具体参数设置情况如表 2 和 3 所示.

观察数值实验图 2~8 发现, 在所有的实验中, DCBA-CR 算法的性能表现均明显优于 OCBA-CR 算法、OCBA-SR 算法、PTV 算法以及 EA 算法. 根据注释 5 可知, 当使用 OCBA-CR 算法对备选方案进行排序时, 该算法与 OCBA-SR 算法等价. 因此, 在所有的实验中 OCBA-CR 算法与 OCBA-SR 算法的性能表现基本一致. 此外, 当每个备选方案的方差均相等时, 根据 PTV 算法的定义可知, 该算法与 EA 算法等价. 因此, 在设置所有备选方案方差均相等的实验 4 中, PTV 算法与 EA 算法的性能表现无明显差异.

在数值实验 1~3 中, 由前验分布所随机产生的部分备选方案均值通常差距较小甚至接近相等. 在这种低置信情景下, 根据式 (47) 和 (48) 可知 OCBA-CR 算法以及 OCBA-SR 算法几乎将所有的仿真预算分配给均值非常接近的少数备选方案, 这使得其他备选方案因仿真复制次数的严重不足而产生非平稳样本均值估计, 从而最终导致 OCBA-CR 算法以及 OCBA-SR 算法无法有效地对备选方案进行正确排序. 此外, 由于 PTV 算法以及 EA 算法的仿真预算分配规则不含有任何样本均值信息, 这使得两种算法的仿真预算分配过程不会出现上述少数备选方案占据近乎全部仿真预算的极端情况. 因此在实验 1~3 中 OCBA-CR 算法以及 OCBA-SR 算法的性能表现不佳甚至明显劣于 PTV 算法以及 EA 算法. 与实验 1~3 不同, 实验 4~7 的前验分布所随机产生的备选方案均值通常差异相对较大. 这种高置信情况能够满足文献 [36] 中关于不同备选方案间均值具有一定差距的假设, 这使得相较于 PTV 算法、EA 算法、OCBA-CR 算法以及 OCBA-SR 算法更能有效地对仿真预算进行合理分配. 因此在实验 4~7 中 OCBA-CR 算法以及 OCBA-SR 算法的性能表现明显优于 PTV 算法以及 EA 算法.

观察表 4 可知, 在实验 1~7 中, DCBA-CR 算法实现 PCR 达到 0.50, 0.50, 0.50, 0.75, 0.70, 0.75, 0.75 时需消耗的仿真预算相较于 OCBA-CR 算法节约的仿真预算分别为 > 10671 , > 8547 , > 4345 , 1063, 1206, 1991, 1222. 综合上述对实验结果的分析可知, 本文提出 DCBA-CR 算法显著地提高了方

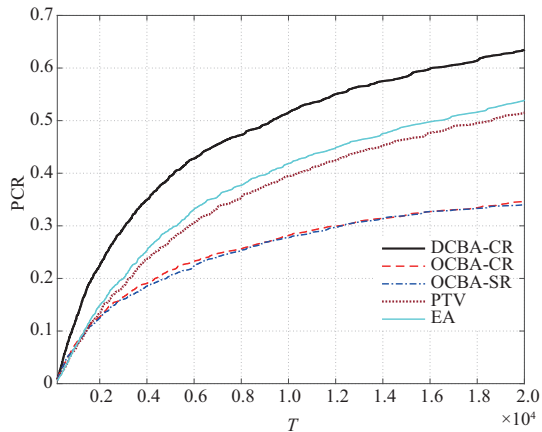


图 2 实验 1 的比较结果

Figure 2 Comparison result for experiment 1

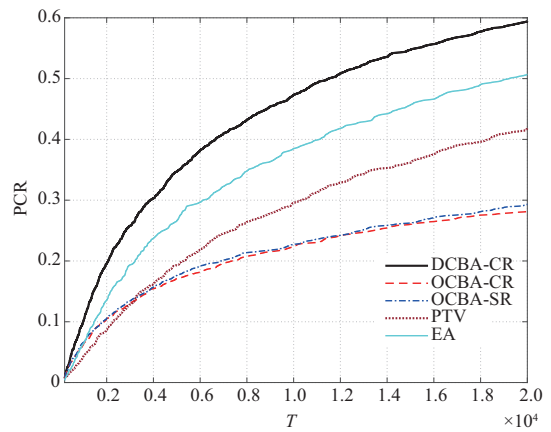


图 3 实验 2 的比较结果

Figure 3 Comparison result for experiment 2

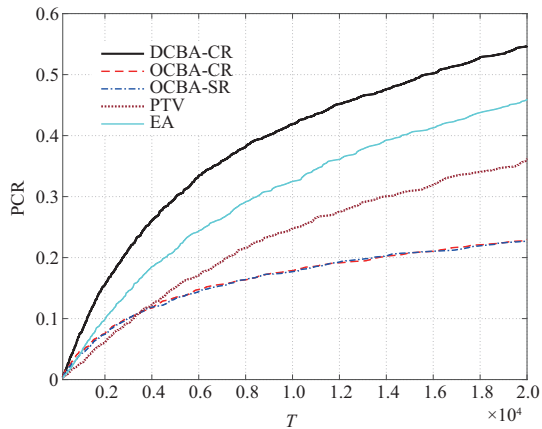


图 4 实验 3 的比较结果

Figure 4 Comparison result for experiment 3

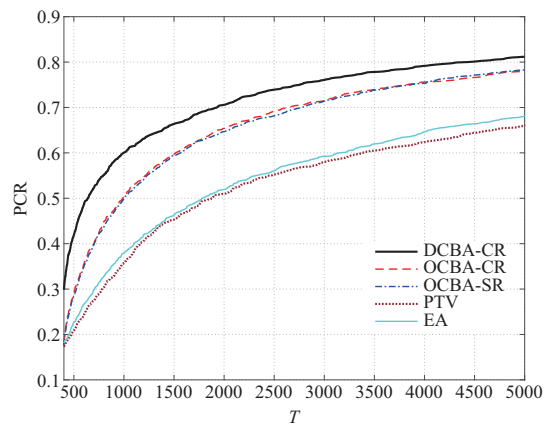


图 5 实验 4 的比较结果

Figure 5 Comparison result for experiment 4

案排序问题的仿真效率.

4.2 应用案例

本文考虑军事战略规划领域中的一个应用实例. 现代信息化战争的高强度以及复杂多变特性, 使得其对战时装备维修保障决策的时效性以及准确性提出了更高的要求^[47~49]. 在战时装备维修保障决策过程中, 战时装备维修需求优先级的评估与排序是维修保障决策的重要前提与基础^[31]. 然而, 一方面由于战时装备维修保障需求优先级的一些重要影响因素通常具有不确定性的显著特点(如坦克战时修理工时近似服从艾拉姆伽(Erlang)分布^[50]), 这使得战时装备维修需求优先级服从某种随机分布, 另一方面现代化战争的复杂多变特性使得一般无法使用数学解析模型对其进行建模与分析. 因此, 对战时装备维修需求优先级精确评估与排序面临着巨大挑战. 文献^[51]指出, 军事仿真系统是战时装备维修需求的有效建模与分析工具. 然而, 由于仿真系统具有随机性, 这使得需要对每种军事装备进行多次

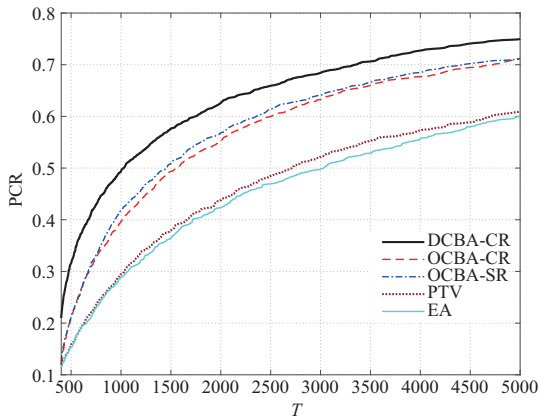


图 6 实验 5 的比较结果

Figure 6 Comparison result for experiment 5

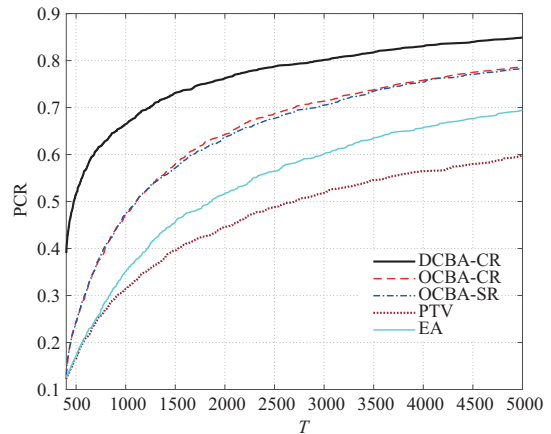


图 7 实验 6 的比较结果

Figure 7 Comparison result for experiment 6

表 4 PCR 达到一些水平时需要消耗的仿真预算次数

Table 4 Simulation budget required for reaching some PCR values

Experiment	PCR	DCBA-CR	OCBA-SR	OCBA-CR	PTV	EA
1	0.50	9329	> 20000	> 20000	18576	16537
2	0.50	11453	> 20000	> 20000	> 20000	19240
3	0.50	15655	> 20000	> 20000	> 20000	> 20000
4	0.75	2760	3866	3823	> 5000	> 5000
5	0.70	3330	4440	4536	> 5000	> 5000
6	0.75	1765	3822	3756	> 5000	> 5000
7	0.75	2087	3282	3309	> 5000	> 5000

的系统仿真复制才能获得其战时维修需求优先级的精确估计^[6,7]. 因此, 如何合理分配有限的仿真预算以增强战时装备维修保障决策的时效性和提高战时装备维修需求优先级排序的准确性, 具有重大战略意义. 本文假设第 i 类军事装备的战时维修需求优先级概率分布为 $\gamma_i \sim N(\mu_i, \sigma_i^2), i \in \{1, 2, \dots, k\}$, 其中, 未知参数 μ_i 为均值, σ_i^2 为已知方差, 并且 μ_i 服从前验分布 $N(\mu_i^{(0)}, (\sigma_i^{(0)})^2)$. 根据贝叶斯法则, 可得均值 μ_i 的后验分布为 $\mu_i \sim N(\mu_i^{(t)}, (\sigma_i^{(t)})^2)$, 其中后验均值 $\mu_i^{(t)}$, 后验方差 $(\sigma_i^{(t)})^2$ 由式 (14) 计算可得. 问题寻求的目标是最大化军事装备的战时维修需求优先级正确排序的概率.

本实验考虑 $k = 15$ 类军事装备的战时维修优先级排序, 并且假设第 i 类军事装备的战时维修需求优先级分布为 $\gamma_i \sim N(\mu_i, \sigma_i^2), i \in \{1, 2, \dots, k\}$, 其中, $\sigma_i = 10, \mu_i$ 的前验分布为 $N(7i + 30, 3^2)$. 此外, 本实验与实验 1~7 的实验设计一致, 即将 DCBA-CR 算法与 OCBA-SR 算法、OCBA-CR 算法、PTV 算法以及 EA 算法, 这 4 种算法进行对比实验. 图 9 展示了 DCBA-CR 算法应用于军事装备的战时维修需求优先级排序问题的实验结果.

如图 9 所示, DCBA-CR 算法的性能表现最优, OCBA-CR 以及 OCBA-SR 算法次之, 为了达到 PCR = 0.80, 上述 5 种算法需要消耗的仿真预算量, 分别为 DCBA-CR 算法需要 1295 次, OCBA-CR 算法需要 1930 次, OCBA-SR 算法需要 2023 次, PTV 算法需要 4770 次, EA 算法需要 4840 次.

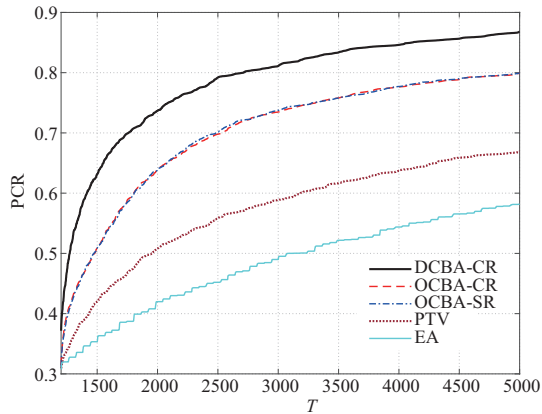


图 8 实验 7 的比较结果

Figure 8 Comparison result for experiment 7

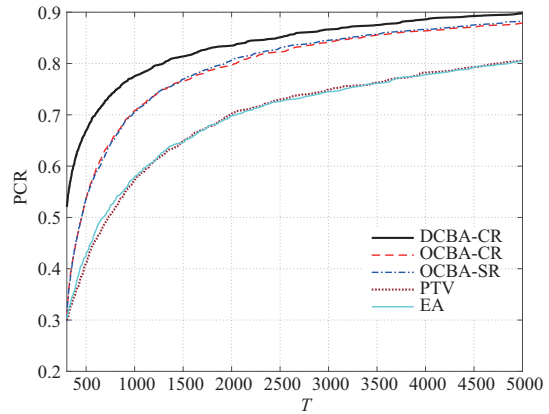


图 9 应用案例的比较结果

Figure 9 Comparison result for an application case

5 总结

本文研究了生产制造、网络通信、医疗健康等诸多信息物理系统中常见的方案排序仿真优化问题。在贝叶斯框架下, 基于随机动态规划方法, 把仿真预算分配决策问题描述成随机动态规划问题并将其转化为马尔科夫决策过程, 遵循值函数近似方法, 以最大化后验正确排序概率 (PCR) 为目标, 推导出渐进最优动态计量预算分配 (DCBA-CR) 策略, 并且证明该策略能够实现一致性以及渐进最优性。此外, 本文给出了 DCBA-CR 算法来实现 DCBA-CR 策略。7 组人工算例以及 1 个应用案例的数值实验结果表明, DCBA-CR 算法与现有的方案排序算法相比, 显著地提高了备选方案排序问题的仿真效率。未来研究方向可以考虑不同备选方案仿真输出之间存在相关性问题的渐进最优动态计量预算分配策略研究。

参考文献

- 1 China Electronics Standardization Institute. White Paper: Cyber-Physical System. 2017 [中国电子技术标准化研究院. 信息物理系统白皮书. 2017. <http://www.cesi.cn/images/editor/20171010/20171010133255806.pdf>]
- 2 Guan X H, Zhao Q C, Jia Q S, et al. Information Physical Integration of Energy Systems. Beijing: Science Press, 2016 [管晓宏, 赵千川, 贾庆山, 等. 信息物理融合能源系统. 北京: 科学出版社, 2016]
- 3 Grieves M. Digital twin: manufacturing excellence through virtual factory replication. White Paper, Florida Institute of Technology. 2015
- 4 China Electronics Standardization Institute. White Paper of Digital Twin Application. 2020 [中国电子技术标准化研究院. 数字孪生应用白皮书. 2020. <http://www.cesi.cn/images/editor/20201118/20201118163619265.pdf>]
- 5 Jia Q S. Enhancing ordinal optimization: a theory study and applications. Dissertation for Ph.D. Degree. Beijing: Tsinghua University, 2006. 1-5 [贾庆山. 增强序优化理论研究及应用. 博士学位论文. 北京: 清华大学, 2006. 1-5]
- 6 Chen C H, Lee L H. Stochastic Simulation Optimization: An Optimal Computing Budget Allocation. Singapore: World Scientific Publishing, 2010. 7-9
- 7 Peng Y J. Simulation resource sharing and dynamic decision in statistical selection problems. Dissertation for Ph.D. Degree. Shanghai: Fudan University, 2014. 1-8 [彭一杰. 统计选择问题中的资源共享与动态决策. 博士学位论文. 上海: 复旦大学, 2014. 1-8]
- 8 Gabillon V, Ghavamzadeh M, Lazaric A. Best arm identification: a unified approach to fixed budget and fixed confidence. In: Proceedings of the 25th International Conference on Neural Information Processing Systems, 2012. 3212-3220

- 9 Hunter S R, Nelson B L. Parallel ranking and selection. In: *Advances in Modeling and Simulation*. New York: Springer, 2017
- 10 Hong L J, Fan W W, Luo J. Review on ranking and selection: a new perspective. *Front Eng Manag*, 2021, 8: 321–343
- 11 Rinott Y. On two-stage selection procedures and related probability-inequalities. *Commun Stat-Theor Methods*, 1978, 7: 799–811
- 12 Fan W W, Hong L J, Nelson B L. Indifference-zone-free selection of the best. *Operations Res*, 2016, 64: 1499–1514
- 13 Chen C H, Lin J W, Yücesan E, et al. Simulation budget allocation for further enhancing the efficiency of ordinal optimization. *Discrete Event Dynamic Syst*, 2000, 10: 251–270
- 14 Chick S E, Branke J, Schmidt C. Sequential sampling to myopically maximize the expected value of information. *Inform J Computing*, 2010, 22: 71–80
- 15 Frazier P I, Powell W B, Dayanik S. A knowledge-gradient policy for sequential information collection. *SIAM J Control Optim*, 2008, 47: 2410–2439
- 16 Ryzhov I O. On the convergence rates of expected improvement methods. *Operations Res*, 2016, 64: 1515–1528
- 17 Peng Y J, Chong E K P, Chen C H, et al. Ranking and selection as stochastic control. *IEEE Trans Automat Contr*, 2018, 63: 2359–2373
- 18 Lee L H, Chew E P, Teng S Y. Computing budget allocation rules for multi-objective simulation models based on different measures of selection quality. *Automatica*, 2010, 46: 1935–1950
- 19 Gao S Y, Xiao H, Zhou E L, et al. Robust ranking and selection with optimal computing budget allocation. *Automatica*, 2017, 81: 30–36
- 20 Cakmak S, Zhou E L, Gao S Y. Contextual ranking and selection with Gaussian processes. In: *Proceedings of the Winter Simulation Conference*, 2021. 1–12
- 21 Jia Q S. Efficient computing budget allocation for simulation-based optimization with stochastic simulation time. *IEEE Trans Automat Contr*, 2013, 58: 539–544
- 22 Chen C H, He D H, Fu M C, et al. Efficient simulation budget allocation for selecting an optimal subset. *Inform J Computing*, 2008, 20: 579–595
- 23 Gao S Y, Chen W W. A new budget allocation framework for selecting top simulated designs. *IIE Trans*, 2016, 48: 855–863
- 24 Zhang J Q, Li Z Z, Wang C, et al. Approximate simulation budget allocation for subset ranking. *IEEE Trans Contr Syst Technol*, 2016, 25: 358–365
- 25 Jia Q S, Zhou E L, Chen C H. Efficient computing budget allocation for finding simplest good designs. *IIE Trans*, 2013, 45: 736–750
- 26 Liu H T, Xiao H, Lee L H, et al. A convergent algorithm for ranking and selection with censored observations. *IIE Trans*, 2023, 55: 523–535
- 27 Ding L, Hong L J, Shen H H, et al. Technical note-knowledge gradient for selection with covariates: consistency and computation. *Naval Res Logistics*, 2022, 69: 496–507
- 28 Zhong Y, Hong L J. Knockout-tournament procedures for large-scale ranking and selection in parallel computing environments. *Operations Res*, 2022, 70: 432–453
- 29 Zhong Y, Liu S, Luo J, et al. Speeding up Paulson’s procedure for large-scale problems using parallel computing. *Inform J Computing*, 2022, 34: 586–606
- 30 Moreau Y, Tranchevent L C. Computational tools for prioritizing candidate genes: boosting disease gene discovery. *Nat Rev Genet*, 2012, 13: 523–536
- 31 Dong Z W, Jia X S, Hu Q W, et al. Research on equipment maintenance support requirement priority assessment based on multiple attribute decision making in wartime. *Fire Control & Command Control*, 2011, 36: 156–159 [董泽委, 贾希胜, 胡起伟, 等. 多属性决策的战时集群装备维修保障需求优先级评估. *火力与指挥控制*, 2011, 36: 156–159]
- 32 Chen Z L, Pundoor G. Order assignment and scheduling in a supply chain. *Operations Res*, 2006, 54: 555–572
- 33 Liu J Q, Li X R, Dong J C. A survey on network node ranking algorithms: representative methods, extensions, and applications. *Sci China Tech Sci*, 2021, 64: 451–461
- 34 Beirlant J, Dudewicz E J, van der Meulen E C. Complete statistical ranking of populations, with tables and applications. *J Comput Appl Math*, 1982, 8: 187–201
- 35 Bishop T A. Designing simulation experiments to completely rank alternatives. In: *Proceedings of the Winter*

- Simulation Conference, 1978. 203–205
- 36 Xiao H, Lee L H, Ng K M. Optimal computing budget allocation for complete ranking. *IEEE Trans Automat Sci Eng*, 2013, 11: 516–524
- 37 Xiao H, Gao F, Lee L H. Optimal computing budget allocation for complete ranking with input uncertainty. *IISE Trans*, 2020, 52: 489–499
- 38 Peng Y J, Chen C H, Chong E K P, et al. A review of static and dynamic optimization for ranking and selection. In: *Proceedings of the Winter Simulation Conference*, 2018. 1909–1920
- 39 Li H D, Lam H, Peng Y J. Efficient learning for clustering and optimizing context-dependent designs. *Operations Res*, 2022. doi: 10.1287/opre.2022.2368
- 40 Shi Z S, Peng Y J, Shi L Y, et al. Dynamic sampling allocation under finite simulation budget for feasibility determination. *Inform J Computing*, 2022, 34: 557–568
- 41 Bertsekas D. *Dynamic Programming and Optimal Control: Volume 1*. Belmont: Athena Scientific, 1995. 20–21
- 42 DeGroot M H. *Optimal Statistical Decisions*. New York: McGraw Hill, 1970. 169–171
- 43 Bechhofer R E, Santner T J, Goldsman D M. Design and analysis of experiments for statistical selection, screening and multiple comparisons. *Technometrics*, 1995, 38: 289–290
- 44 Powell W B. *Approximate Dynamic Programming: Solving the Curses of Dimensionality*. Hoboken: John Wiley & Sons, 2007. 235–239
- 45 Bertsekas D P. Dynamic programming and suboptimal control: a survey from ADP to MPC. *Eur J Control*, 2005, 11: 310–334
- 46 Rudin W. *Principles of Mathematical Analysis*. New York: McGraw Hill, 1964. 51–52
- 47 Niu T L, Wang J, Du Y B, et al. Research on sequencing decision-making of priority for equipment repair tasks in battlefield. *Fire Control & Command Control*, 2011, 36: 127–130 [牛天林, 王洁, 杜燕波, 等. 装备战场抢修任务优先度排序决策研究. *火力与指挥控制*, 2011, 36: 127–130]
- 48 Liu B. A requirement computing model of equipment maintenance resource in non-warfare action. *J Naval Aviat Univ*, 2013, 28: 333–336 [刘波. 非战争军事行动中装备维修保障资源需求计算模型. *海军航空大学学报*, 2013, 28: 333–336]
- 49 Liu W, Xu H, Shi Q. Research of equipment maintenance level classification in wartime. *Ord Ind Autom*, 2007, 26: 23–24 [刘伟, 徐航, 石全. 战时装备维修等级划分研究. *兵工自动化*, 2007, 26: 23–24]
- 50 Xu H, Zhu Y F, Chen C L. Research on battlefield damaged armored materiel maintenance man hour simulation and its distribution. *J Syst Simul*, 2006, 18: 2945–2947 [徐航, 朱一凡, 陈春良. 战伤装甲装备修理工时仿真及其分布规律研究. *系统仿真学报*, 2006, 18: 2945–2947]
- 51 Wang Z C. Development and application of simulation technology. *Strategic Study CAE*, 2003, 5: 40–44 [王子才. 仿真技术发展及应用. *中国工程科学*, 2003, 5: 40–44]

Dynamic computing budget allocation for design ranking in stochastic complex systems

Hui XIAO¹, Zhihong WEI¹, Gang KOU² & Xianming GU^{3*}

1. *School of Management Science and Engineering, Southwestern University of Finance and Economics, Chengdu 611130, China;*

2. *School of Big Data, Southwestern University of Finance and Economics, Chengdu 611130, China;*

3. *School of Mathematics, Southwestern University of Finance and Economics, Chengdu 611130, China*

* Corresponding author. E-mail: guxm@swufe.edu.cn

Abstract Simulation has been widely used to evaluate and optimize stochastic complex systems such as manufacturing, telecommunication, and healthcare systems, which are known as cyber-physical systems. However, research on design ranking in simulation optimization is limited. This research considers the complete ranking of k designs, whose performance can only be evaluated through simulation. Using the Bayesian framework, an optimal dynamic computing budget allocation model is developed, which maximizes the posterior probability of correct rankings. The computing budget allocation model is reformulated as a stochastic dynamic programming problem. This paper derives the optimal dynamic computing budget allocation rule and confirms its asymptotic optimal property. Numerical experiments and a case study demonstrate that the proposed simulation procedure considerably increased the simulation optimization efficiency for design ranking problems.

Keywords simulation optimization, ranking and selection, cyber-physical system, Markov decision process, Bayesian estimation