

# 参数化最优控制问题的初值校正法

叶昌伦

贵州大学数学与统计学院, 贵州 贵阳

收稿日期: 2022年10月12日; 录用日期: 2022年11月8日; 发布日期: 2022年11月15日

---

## 摘要

针对带随机系数椭圆方程约束的最优控制问题, 我们提出一种结合共轭梯度法的初值校正新方法。该方法通过少量的样本求解最优控制, 然后得到的解作为抽取其它样本的初始迭代值, 从而减少整个系统求解最优控制平均值的成本。我们理论上说明该方法的合理性。两个数值例子表明我们的方法是充分优于未初值校正方法。

## 关键词

参数化最优控制, 共轭梯度法, 蒙特卡洛法

---

# An Initial Value Correction Method for the Parameterized Optimal Control Problem

Changlun Ye

Guizhou University, School of Mathematics and Statistics, Guiyang Guizhou

Received: Oct. 12<sup>th</sup>, 2022; accepted: Nov. 8<sup>th</sup>, 2022; published: Nov. 15<sup>th</sup>, 2022

---

## Abstract

For optimal control problems with random elliptic equation constraints, we propose an initial value correction new method combined with the conjugate gradient method. This method solves the optimal control through a small number of samples, and then the solution is used as the initial iteration value of extracting other samples, thus reducing the cost of solving the optimal control average for the whole system. We theoretically illustrate the rationality of the method. Two numerical examples show that our method is sufficiently superior to the no initial value correction method.

## Keywords

Parameterized Optimal Control, Conjugate Gradient Method, Monte Carlo Method

Copyright © 2022 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

最优控制问题广泛应用于工程和应用科学领域。对于确定性最优控制问题，存在着大量的数学理论和计算方法[1] [2] [3] [4] [5]。然而，边界条件、材料参数往往不能被精确地测量，所以随机模型更现实。与确定性最优控制问题一样，随机最优控制问题很难得到解析解，数值解是应用这些问题的一个很好的选择。在过去的几十年里，随机最优控制问题的数值研究已经引起了人们的关注[6] [7] [8] [9]。

对于随机系数的每一个固定实现，我们求解一个确定的最优控制系统。然后，估计最优控制的统计量(如：期望值、方差)。这使我们能够研究控制对偏微分方程约束中的随机波动的敏感性，并为受不确定输入影响的控制装置的设计提供实际信息。迄今为止，具有随机系数的椭圆偏微分方程已经得到了很好的研究，并且有大量有关这方面数值方法[10] [11]。具有随机系数的最优控制问题的研究相对少一些。[9]研究了受随机椭圆型偏微分方程约束的随机最优控制问题的最优解可测性。在传统有限元方法的框架下，[12]研究了随机最优控制问题最优解的存在唯一性。

众所周知，偏微分方程约束优化问题的数值解在计算上是昂贵的，因为它需要数值求解由最优性条件产生的偏微分方程系统的解：状态问题、伴随问题。特别是偏微分方程的参数随机化时，这使计算变得更具挑战性，并可能导致高维随机空间中的“维数诅咒”。目前已有一些求解随机最优控制问题的有效方法。

A. Borzi 和 G. Winckel 将随机配置法与多重网格结合求解具有随机系数抛物型和椭圆型最优控制问题[3]。O. Ghattas and A. Alexanderian, 针对均值方差风险规避最优控制问题，运用泰勒近似，简化原本最优控制模型[13]。蒙特卡罗方法(MC)是在随机空间中得到近似最优解的最直接的方案之一。然而，它的收敛率很低，并且抽取每个样本都必须解决一个完全确定性的最优控制问题，导致沉重的计算成本。为克服 MC 的不足，M. Hinze, F. Kuo, S. Vandewalle 等人使用多水平蒙特卡洛方法和拟蒙特卡洛方法模拟最优控制的均值，降低 MC 的计算成本[9] [14] [15]。但是他们没有降低单个样本成本，而是从方差减少的角度降低计算整个最优控制的样本量。

很多时候我们使用 MC 抽样时，没有充分利用已抽样本的信息，这也是导致计算花费增加一个原因。为更好利用已抽样本的信息，我们提出初值校正的方法，该方法通过少量的样本求解最优控制平均值，然后得到的平均值作为抽取其它样本的初始迭代值，从而减少求解最优控制统计量的成本。结合理论分析和数值实验，充分说明了方法实用性。

文章主要分为四部分，第一部分给出模型及一些基本的假设。第二部分给出最优控制问题的梯度及有限元离散模型。第三部分分析求解最优控制问题的误差，并提出基于初值校正的算法，分析算法的可靠性。第四部分给出两个数值算例，比较了初值校正与初值未校正算法的差异。

## 2. 模型及假设

我们用  $(\Omega, \mathcal{A}, P)$  表示一个概率空间，其中  $\Omega$  是一个样本空间， $\mathcal{A} \subset 2^\Omega$  是一个  $\sigma$  代数， $P: \mathcal{A} \rightarrow [0,1]$  是

一个概率空间, 给定巴拿赫空间  $(X, \|\cdot\|_X)$ ,  $u \in L^p(\Omega, X)$  是强可测函数, 并满足  $u: \Omega \rightarrow X$ ,  $\|u\|_{L^p(\Omega, X)} < \infty$ , 其中:

$$\|u\|_{L^p(\Omega, X)} = \begin{cases} \left( \int_{\Omega} \|u\|_X^p dP(\omega) \right)^{\frac{1}{p}} & 1 \leq p < \infty, \\ \text{ess sup}_{\omega \in \Omega} \|u\|_X & p = \infty. \end{cases}$$

随机变量  $v$  的期望算子, 方差算子表示如下:

$$E[u] = \int_{\Omega} u dP(\omega), \quad \text{Var}(u) = E\left[(u - E[u])^2\right] = E[u^2] - E[u]^2.$$

## 2.1. 状态方程与目标函数

对于我们提出的方法并不假定任何特定的随机偏微分方程, 然而, 为了阐明方法的思想, 我们考虑状态方程是随机椭圆方程约束下的最优控制问题。

空间区域  $D = [0, 1]^2 \subset \mathbb{R}^2$ ,  $\partial D$  表示边界。在  $D$  上温度分布由状态  $y$  描述。控制  $u$  是  $D$  上的一个热源, 这里我们假设  $u$  与时间无关。热传导系数是一个随机场  $k: D \times \Omega \rightarrow \mathbb{R}: (x, \omega) \rightarrow k(x, \omega)$ 。整个系统用下面带随机系数的 PDE 描述:

$$\begin{aligned} -\nabla \cdot (k(x, \omega) \nabla y(x, \omega)) &= u(x), \quad (x, \omega) \in D \times \Omega \\ y(x, \omega) &= 0, \quad (x, \omega) \in \partial D \times \Omega \end{aligned} \quad (1)$$

$$k_{\min}(\omega) = \min_{x \in D} k(x, \omega), \quad k_{\max}(\omega) = \max_{x \in D} k(x, \omega).$$

我们给出以下假设:

A1:  $k_{\min}(\omega) > 0$ , a.s.,  $k_{\min}^{-1} \in L^p(\Omega)$ , 其中  $p \in [1, \infty)$ ,

A2:  $k \in L^p(\Omega, C^1(\bar{D}))$ , 其中  $p \in [1, \infty)$ ,

A3:  $u \in L^2(D)$ 。

进一步我们写出(1)的变分形式

$$(k(x, \omega) \nabla y(x, \omega), \nabla v) = (u, v), \quad v \in H_0^1(D).$$

由文献[11] [13]知, 上述方程几乎确定有解, 并且  $y \in L^2(\Omega, H^2(D)) \cap L^2(\Omega, H_0^1(D))$ 。

我们研究下面参数化的二次型目标函数

$$J = \frac{1}{2} \|y(x, \omega) - y_d(x)\|_{L^2(D)}^2 + \frac{\lambda}{2} \|u(x)\|_{L^2(D)}^2,$$

其他不同类型的目标函数可参考[14] [16], 注意到这里控制与参数  $\omega$  无关。  $\lambda$  表示正则化参数, 对每一个固定的参数  $\omega$ , 状态  $y$  与  $u$  满足方程(1)。

## 2.2. 随机场 $k(x, \omega)$

我们假设  $k$  是对数随机场, 即:  $k(x, \omega) = \exp(z(x, \omega))$ , 其中  $z$  是高斯随机场,  $E[z(x, \omega)] = 0$ , 根据文献[17], 高斯随机场的协方差定义为:

$$C(x_1, x_2) = \text{Cov}(z(x_1, \omega), z(x_2, \omega)) = \sigma^2 \exp\left(-\frac{\|x_1 - x_2\|_1}{a}\right),$$

其中  $\sigma^2$  是随机场的方差,  $a$  是相关长度。为得到  $z$  的样本, 通常方法有循环抽样[18], K-L 展开[19]。本文运用 K-L 开展

$z(x, \omega) = E[z(x, \omega)] + \sum_{n=1}^{\infty} \sqrt{\theta_n} \xi_n(\omega) f_n(x)$ . 有关 K-L 展开的详细内容可参考[18], 我们选择相关长度  $a = 0.3$ , 截断数  $n_{KL} = 500$ . 关于 K-L 展开抽样产生的误差不是文章关注的重点。图 1 给出不同方差下随机场的实现。

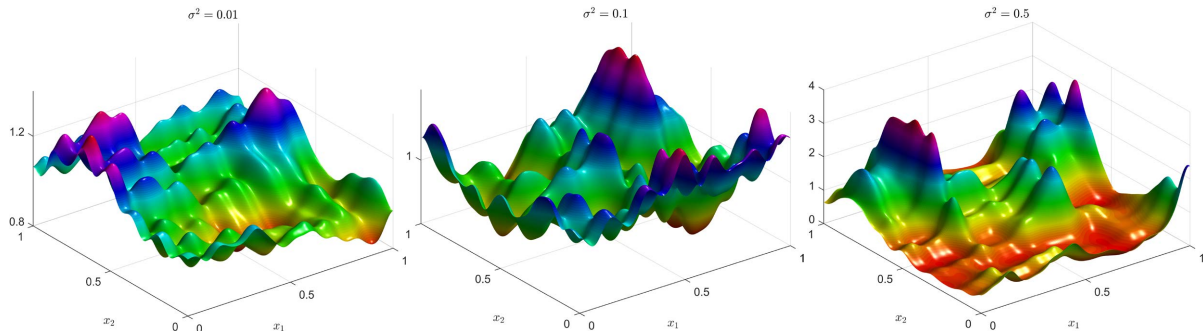


Figure 1. Realization of random field  
图 1. 随机场的实现

### 3. 最优控制的梯度及离散最优控制

为估计最优控制期望, 我们采用 MC 进行抽样。固定随机参数  $\omega$ , 得到一个参数化的最优控制问题:

$$\begin{aligned} \min_{u \in L^2(D)} J^\omega(u) &= \frac{1}{2} \|y^\omega - y_d\|_{L^2(D)}^2 + \frac{\lambda}{2} \|u\|^2, \\ -(k^\omega \nabla y^\omega, \nabla v) &= (u, v), \quad v \in H_0^1(D), \end{aligned}$$

其中  $y^\omega$  表示依赖随机参数  $\omega$ , 标准的证明方法[1] [9], 我们可以知道最优控制有唯一解  $u^{*\omega}$ , 同时可以得到最优控制的梯度:

$$\begin{aligned} (k^\omega y^\omega, \nabla v) &= (u, v), \quad v \in H_0^1(D), \\ (k^\omega \nabla p^\omega, \nabla v) &= (y^\omega - y_d, v), \quad v \in H_0^1(D) \\ \nabla J^\omega(u) &= \lambda u + p^\omega. \end{aligned}$$

下面我们离散这个系统, 设  $T_h$  是  $D$  的三角形剖分, 最大网格尺寸  $h = \max_{T \in T_h} \text{diam}(T)$ , 并且  $\bar{D} = \bigcup_{T \in T_h} \bar{T}$ .

假定三角形剖分是拟一致的, 我们定义有限元空间:

$$V_h = \{v_h \in C^0(\bar{D}) : v_h \text{ 在每个三角形单元 } T \in T_h \text{ 是线性多项式, 且 } v_h|_{\partial D} = 0\}.$$

得到下面有限元离散形式

$$\begin{aligned} (k^\omega \nabla y_h^\omega, \nabla v) &= (u, v), \quad v \in V_h, \\ (k^\omega \nabla p_h^\omega, \nabla v) &= (y_h^\omega - y_d, v), \quad v \in V_h \\ \nabla J^\omega(u) &= \lambda u + p_h^\omega. \end{aligned}$$

进一步我们得到矩阵形式:

$$\begin{aligned} A_h y^\omega &= M_h u, \\ A_h p^\omega &= M_h (y^\omega - y_d), \\ \nabla J^\omega(u) &= \lambda u + p^\omega, \end{aligned}$$

其中  $h$  表示离散参数,  $\mathbf{A}_h$  表示刚度矩阵依赖于  $\omega$ ,  $\mathbf{M}_h$  是对称的质量矩阵独立于  $\omega$ 。

#### 4. 数值方法

对于上述最优控制问题, 常见的数值优化方法是梯度下降法, 而在实际应用中, 如果我们需要较高精度数值, 需要更细的步长, 及很小的正则化参数, 这使得整个离散最优控制系统很病态, 增加最优控制的数值求解花费。

下面, 我们统计上采用蒙特卡洛方法(MC), MC 方法它的收敛率不受维数的影响, 但是抽样成本昂贵, 因为每次抽样都要求解一个最优控制系统, 因此我们希望通过一些技巧减少 MC 抽样成本。优化算法采用共轭梯度法, 分析获取控制期望所产生的成本。

对于每一个固定样本  $\omega$ , 离散最优控制问题的数值解我们用  $u_{k,h}^{*,\omega}$  表示, 其中  $k$  优化算法的迭代次数, 因此:

$$E[u_{k,h}^*] \approx E[u^*].$$

我们定义

$$E_N[u_{k,h}^*] = \frac{u_{k,h}^{*,\omega_1} + \dots + u_{k,h}^{*,\omega_N}}{N},$$

其中  $N$  表示样本量,  $\{u_{k,h}^{*,\omega_i}\}_{i=1}^N$  独立同分布,  $E_N[u_{k,h}^*]$  是  $E[u_{k,h}^*]$  的近似。进一步。我们给出求解最优控制  $E[u^*]$  的均方误差:

$$\begin{aligned} \|E_N[u_{k,h}^*] - E[u^*]\|_{L^2(\Omega;L^2(D))}^2 &= \|E_N[u_{k,h}^*] - E[u_{k,h}^*] + E[u_{k,h}^*] - E[u^*]\|_{L^2(\Omega;L^2(D))}^2 \\ &= \|E_N[u_{k,h}^*] - E[u_{k,h}^*]\|_{L^2(\Omega;L^2(D))}^2 + \|E[u_{k,h}^*] - E[u^*]\|_{L^2(\Omega;L^2(D))}^2. \end{aligned}$$

这里误差分为两部分, 前面部分表示 MC 抽样产生的统计误差, 而后面部分表示有限元误差和共轭梯度迭代产生的误差。我们先估计前面部分:

$$\|E_N[u_{k,h}^*] - E[u_{k,h}^*]\|_{L^2(\Omega;L^2(D))}^2 = \frac{1}{N} \int_D \text{Var}(u_{k,h}^*) dx.$$

后面部分, 我们简单运用三角不等式将有限元离散和共轭梯度法误差分开:

$$\|E[u_{k,h}^*] - E[u^*]\|_{L^2(\Omega;L^2(D))}^2 \leq 2 \left( \|u_h^* - u^*\|_{L^2(\Omega;L^2(D))}^2 + \|u_{k,h}^* - u_h^*\|_{L^2(\Omega;L^2(D))}^2 \right)$$

根据有限元误差估计[11][9], 我们知道  $\|u_h^* - u^*\|_{L^2(\Omega;L^2(D))} \leq C_1 h^2$ , 同时我们知道共轭梯度法是线性收敛,

所以  $\|u_{k,h}^* - u_h^*\|_{L^2(\Omega;L^2(D))} \leq C_2 \exp(-k)$ , 因此我们得到求解最优控制的误差估计:

$$\|E_N[u_{k,h}^*] - E[u^*]\|_{L^2(\Omega;L^2(D))} \leq \frac{1}{\sqrt{N}} \left( \int_D \text{Var}(u_{k,h}^*) dx \right)^{\frac{1}{2}} + C_1 h^2 + C_2 \exp(-k).$$

为总的误差小于  $\epsilon$  ( $\epsilon < 1$ ), 我们需要选择

$$N = O(\epsilon^{-2}), \quad h = O(\sqrt{\epsilon}), \quad k = O\left(\log \frac{1}{\epsilon}\right).$$

因此总的花费为  $O\left(\epsilon^{-2} \log \frac{1}{\epsilon}\right)$ 。仔细地讲我们至少需要求解  $O\left(\epsilon^{-2} \log \frac{1}{\epsilon}\right)$  次维度为  $O(\sqrt{\epsilon})$  的线性方程

组(非线性方程组), 成本是非常高的, 尤其整个离散最优控制系统比较病态时( $\lambda, h$  比较小时)。

为减少优化方法迭代次数, 下面提出一种后验方法选取优化方法的初始迭代值, 减少整个最优控制的花费。

首先在给出算法之前, 我们给出一些准备, 记

$$(\mathbf{u}, \mathbf{v})_{M_h} = \mathbf{u}^T M_h \mathbf{v}, \quad \|\mathbf{u}\|_{M_h} = \sqrt{(\mathbf{u}, \mathbf{u})_{M_h}}.$$

第  $k$  次迭代用到的变量我们用  $(k)$  上标表示,  $\nabla \mathbf{J}^{(k)}$  表示迭代  $k$  次计算的梯度, 共轭梯度法的搜索方向

$$\mathbf{d}^{(k)} = -\nabla \mathbf{J}^{(k)} + \alpha^{(k)} \mathbf{d}^{(k-1)}, \text{ 其中 } \mathbf{d}^{(0)} = \nabla \mathbf{J}^{(0)}, \quad (2)$$

关于  $\alpha^{(k)}$ , 我们使用戴袁公式

$$\alpha^{(k)} = \frac{\|\nabla \mathbf{J}^{(k)}\|_{M_h}^2}{(\mathbf{d}^{(k-1)}, \nabla \mathbf{J}^{(k)} - \nabla \mathbf{J}^{(k-1)})_{M_h}}. \quad (3)$$

关于戴袁公式更多内容可参考文献[19]。控制更新公式为  $\mathbf{u}^{(k+1)} = \mathbf{u}^{(k)} + s^{(k)} \mathbf{d}^{(k)}$ , 其中  $s^{(k)}$  表示步长。下面是共轭梯度算法:

---

#### 算法 1 共轭梯度法

---

- 1: 输入  $\epsilon$ ,  $k_{\max}$ ,  $\mathbf{u}^{(0)}$  %~给定迭代误差, 最大迭代次数, 迭代初始控制
  - 2: 根据  $\mathbf{u}^{(0)}$  计算梯度  $\nabla \mathbf{J}^{(0)}$
  - 3: **for**  $k=0, \dots, k_{\max}-1$  **do**
  - 4: **if**  $\|\nabla \mathbf{J}^{(k)}\|_{M_h} \leq \epsilon$  则
  - 5: **return**  $\mathbf{u}^{(k)}$
  - 6: **end if**
  - 7: 根据公式(2), (3)决定方向  $\mathbf{d}^{(k)}$ , 并通过二次逼近选取步长  $s^{(k)}$
  - 8:  $\mathbf{u}^{(k+1)} = \mathbf{u}^{(k)} + s^{(k)} \mathbf{d}^{(k)}$
  - 9: 根据  $\mathbf{u}^{(k+1)}$  计算梯度  $\nabla \mathbf{J}^{(k+1)}$
  - 10: **end for**
- 

我们使用 MC 估计控制期望时, 每次抽样时都需要调用算法 1, 抽取所有的样本所用的初值往往都是一样的。这没有充分使用已抽样的样本信息。对此我们提出一种初值校正的方法。

---

#### 算法 2 初值校正求解期望算法

---

- 1: 输入  $\epsilon$ ,  $\mathbf{u}^{(0)}$ ,  $N$ ,  $N_0$  %~给定迭代误差, 初始迭代值, 迭代初始控制, 总抽样数, 校正样本数。注意  $N_0 \ll N$ 。
  - 2: 调用算法 1 抽样计算  $E_{N_0} [\mathbf{u}_{k,h}^*]$ 。
  - 3:  $\mathbf{u}^{(0)} = E_{N_0} [\mathbf{u}_{k,h}^*]$ , 作为  $N - N_0$  样本的初始迭代值。
  - 4: 调用算法 1 抽样计算  $E_{N-N_0} [\mathbf{u}_{k,h}^*]$ , %注意此时优化算法的迭代初值已被  $E_{N_0} [\mathbf{u}_{k,h}^*]$  代替。
  - 5:  $E_N [\mathbf{u}_{k,h}^*] = E_{N_0} [\mathbf{u}_{k,h}^*] + \frac{N - N_0}{N} (E_{N-N_0} [\mathbf{u}_{k,h}^*] - E_{N_0} [\mathbf{u}_{k,h}^*])$
-

通过少量的样本求解最优控制，然后将迭代算法的初值用平均值代替，这个平均值或许能够靠近每个样本的精确值，从而减少样本的迭代次数。

更精确的讲，根据切比雪夫不等式

$$P\left(\|u_{kh}^* - E[u_{kh}^*]\|_{L^2(D)} > \epsilon\right) \leq \frac{E\left[\|u_{kh}^* - E[u_{kh}^*]\|_{L^2(D)}^2\right]}{\epsilon^2} = \frac{\int_D \text{Var}(u_{kh}^*) dx}{\epsilon^2}.$$

很容易地看出，当方差小时，大多数的样本的精确值是靠近平均值。使得大多数的样本的迭代次数减少。当方差比较大时，算法 2 的或许效果不是那么明显，可以从后面的实验看出这一点。但是这样校正初值总是比没有校正初值的效果好，因为许多情况下，样本总是在平均值附近扰动。下面我们给出了两个数值例子，表明算法 2 的实用性。

## 5. 数值实验

这部分，我们给出两个例子，并比较了初值校正(aNCG-MC)和初值未校正算法(NCG-MC)。设期望状态

$$y_d(x) = \begin{cases} 1 & x \in [0.25, 0.75] \times [0.25, 0.75], \\ 0 & \text{其他}. \end{cases}$$

关于随机场的参数： $a=0.3$ ， $n_{KL}=500$ ，我们数值实验在 MATLAB2016a 上进行 (Intel(R)Core(TM)i7-6700HQ CPU@2.60GHz 2.59GHz)，关于求解线性代数方程，使用 matlab 中的反斜杠算子，两个算例的数值结果见图 2、图 3。

### 5.1. 算例 1

这里我们考虑第 3 节的最优控制问题，设置空间离散步长  $h = \frac{1}{2^5}$ ，迭代误差  $\epsilon = 1e-5$ ，初始迭代  $\mathbf{u}^{(0)} = 1$ ，抽样数  $N = 200$ ，校正样本量  $N_0 = 20$ ，设置不同的正则化参数  $\lambda = 1e-4, 1e-6, 1e-8$ ，选取步长

$$s^{(k)} = \frac{-\left(\mathbf{A}_h^{-1} \mathbf{M}_h \mathbf{d}^{(k)}, \mathbf{y}^{(k-1)} - \mathbf{y}_d\right)_{\mathbf{M}_h} - \lambda \left(\mathbf{d}^{(k)}, \mathbf{u}^{(k-1)}\right)_{\mathbf{M}_h}}{\left\|\mathbf{A}_h^{-1} \mathbf{M}_h \mathbf{d}^{(k)}\right\|_{\mathbf{M}_h}^2 + \lambda \left\|\mathbf{d}^{(k)}\right\|_{\mathbf{M}_h}^2},$$

其中  $\mathbf{A}_h^{-1}$  表示刚度矩阵  $\mathbf{A}_h$  的逆，通过算法 2 求解，结果见表 1，表里的数据表示抽取 200 个样本总的迭代数。我们固定  $\lambda = 1e-6$ ，设置不同的迭代误差  $\epsilon = 1e-3, 1e-4, 1e-5$ ，见表 2 的结果。

**Table 1.** Example 1: aNCG-MC and NCG-MC ( $\epsilon = 1e-5$ )

**表 1.** 算例 1: aNCG-MC 与 NCG-MC ( $\epsilon = 1e-5$ )

$\sigma^2$	$\lambda = 1e-4$			$\lambda = 1e-6$			$\lambda = 1e-8$		
	aNCG-MC	NCG-MC	节省	aNCG-MC	NCG-MC	节省	aNCG-MC	NCG-MC	节省
0.01	839	1189	29.44%	1562	4205	62.85%	2313	10430	77.82%
0.1	1113	1344	17.19%	2479	4192	40.86%	4247	11378	62.67%
0.5	1220	1337	8.75%	3097	3965	21.89%	6316	11323	44.22%

从实验可以看出，初值校正是优于初值未校正，尤其是当  $\lambda$  很小时。但当方差较大时，样本差异大时，初值校正的方法效果稍差一些，这时我们可以考虑分类的思想，将样本相似的聚在一起，减少整体

样本的差异性。

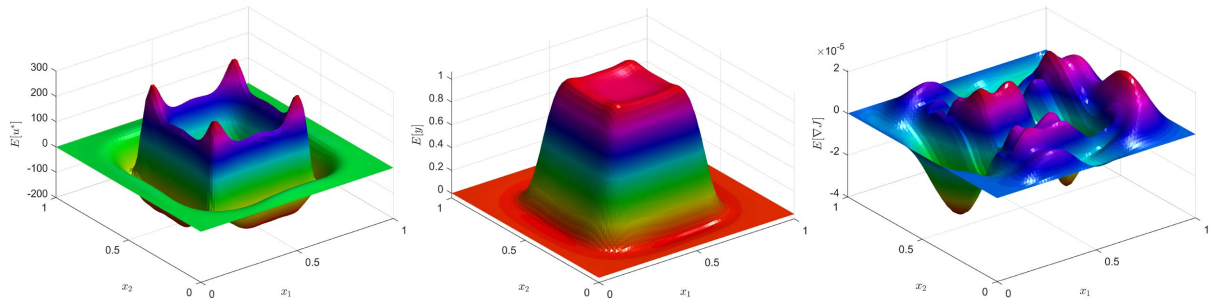


Figure 2. Example 1 Numerical results ( $\lambda = 1e-6, \epsilon = 1e-5, \sigma^2 = 0.5, h = \frac{1}{2^6}$ )

图 2. 算例 1 数值结果 ( $\lambda = 1e-6, \epsilon = 1e-5, \sigma^2 = 0.5, h = \frac{1}{2^6}$ )

Table 2. Example 1: aNCG-MC and NCG-MC ( $\lambda = 1e-6$ )

表 2. 算例 1: aNCG-MC 与 NCG-MC ( $\lambda = 1e-6$ )

$\sigma^2$	$\epsilon = 1e-3$			$\epsilon = 1e-4$			$\epsilon = 1e-5$		
	aNCG-MC	NCG-MC	节省	aNCG-MC	NCG-MC	节省	aNCG-MC	NCG-MC	节省
0.01	400	2001	80.01%	551	2341	76.46%	824	2179	62.18%
0.5	320	410	21.95%	717	1316	45.52%	1678	2111	20.51%

### 5.2. 算例 2

我们考虑下面的半线性状态方程

$$-\nabla \cdot (k(x, \omega) \nabla y(x, \omega)) + f(y(x, \omega)) = u(x), \quad (x, \omega) \in D \times \Omega,$$

$$y(x, \omega) = 0, \quad (x, \omega) \in \partial D \times \Omega,$$

其中  $f$  表示非线性反应项。设置  $f(x) = 5 + x^3$ ，关于上述方程解的存在唯一性可参考文献[3] [20]。下面我们给出这个问题的梯度

$$\begin{aligned} (k^\omega \nabla y^\omega, \nabla v) + (f(y^\omega), v) &= (u, v), \quad v \in H_0^1(D), \\ (k^\omega \nabla p^\omega, \nabla v) + (f'(y^\omega) p^\omega, v) &= (y^\omega - y_d, v), \quad v \in H_0^1(D), \\ \nabla J^\omega(u) &= \lambda u + p^\omega, \end{aligned} \tag{4}$$

其中  $f'$  表示  $f$  的导数。用  $\mathcal{N}$  表示求解状态变量的算子，即  $y_\omega = \mathcal{N}(u)$ ，离散(4)可得

$$\begin{aligned} \mathbf{y}^\omega &= \mathcal{N}(\mathbf{u}) \\ \mathbf{C}_h \mathbf{p}^\omega &= \mathbf{M}_h (\mathbf{y}^\omega - \mathbf{y}_d), \\ \nabla \mathbf{J}^\omega(\mathbf{u}) &= \lambda \mathbf{u} + \mathbf{p}^\omega, \end{aligned}$$

其中  $\mathbf{C}_h = \mathbf{A}_h + \mathbf{B}_h$ ， $\mathbf{B}_h$  表示伴随方程左端第二项的离散有限元矩阵。取迭代步长

$$s^{(k)} = \frac{-\left(\nabla \mathbf{J}^{(k)}, \mathbf{d}^{(k)}\right)_{\mathbf{M}_h}}{\left(1 - \mathbf{p}^{(k)} \otimes f''(\mathbf{y}^{(k-1)}), \left(\mathbf{C}_h^{-1} \mathbf{d}^{(k)}\right) \otimes \left(\mathbf{C}_h^{-1} \mathbf{d}^{(k)}\right)\right)_{\mathbf{M}_h} + \lambda \|\mathbf{d}^{(k)}\|_{\mathbf{M}_h}^2},$$



其中  $\mathbf{a} \otimes \mathbf{b}$  表示  $\mathbf{a}, \mathbf{b}$  对应元素相乘组成的列向量,  $\mathbf{C}_h^{-1}$  表示  $\mathbf{C}_h$  的逆,  $f''$  表示  $f$  的二阶导。我们采用牛顿法求解状态方程, 其他的参数设置均与问题一相同, 结果见表 3、表 4。

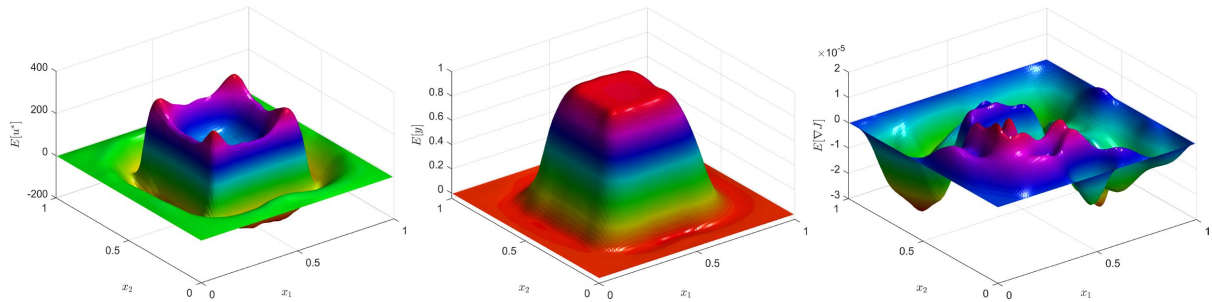


Figure 3. Example 2 Numerical results ( $\lambda = 1e-6, \epsilon = 1e-5, \sigma^2 = 0.5, h = \frac{1}{2^6}$ )

图 3. 算例 2 数值结果( $\lambda = 1e-6, \epsilon = 1e-5, \sigma^2 = 0.5, h = \frac{1}{2^6}$ )

Table 3. Example 2: aNCG-MC and NCG-MC ( $\epsilon = 1e-5$ )

表 3. 算例 2: aNCG-MC 与 NCG-MC ( $\epsilon = 1e-5$ )

$\sigma^2$	$\lambda = 1e-4$			$\lambda = 1e-6$			$\lambda = 1e-8$		
	aNCG-MC	NCG-MC	节省	aNCG-MC	NCG-MC	节省	aNCG-MC	NCG-MC	节省
0.01	522	1202	56.57%	435	2666	83.68%	441	7321	93.98%
0.1	596	1211	50.78%	777	2913	73.33%	4247	11378	62.67%
0.5	646	1138	43.23%	1232	2949	58.22%	2900	8546	66.07%

Table 4. Example 2: aNCG-MC and CG-MC ( $\lambda = 1e-6$ )

表 4. 算例 2: aNCG-MC 与 CG-MC ( $\lambda = 1e-6$ )

$\sigma^2$	$\epsilon = 1e-3$			$\epsilon = 1e-4$			$\epsilon = 1e-5$		
	aNCG-MC	NCG-MC	节省	aNCG-MC	NCG-MC	节省	aNCG-MC	NCG-MC	节省
0.01	200	600	66.67%	200	1597	87.48%	436	2643	83.50%
0.5	203	578	64.88%	297	1649	81.99%	1212	2980	59.33%

## 6. 结论

本文通过初值校正的方法减少每次抽样的迭代次数, 从而减少整个 MC 抽样成本。两个数值例子表明我们的方法是充分优于未初值校正方法。初值校正方法可结合其他的优化方法(原始对偶有效集、牛顿法等等), 或许有更好的效果。也可以结合多水平蒙特卡洛方法或拟蒙特卡洛方法, 进一步减少计算最优控制的成本。

## 基金项目

项目名称: 随机最优控制问题的高效 Monte Carlo 有限元法; 合同编号: 国家自然科学基金项目(11961008)。

## 参考文献

- [1] Hinze, M., Pinnau, R., Ulbrich, M. and Ulbrich, S. (2009) Optimization with PDE Constraints. Mathematical Modeling: Theory and Applications (MMTA, Volume 23), Springer, New York.

- 
- [2] Liu, W., Ma, H., Tang, T. and Yan, N. (2004) A Posteriori Error Estimates for Discontinuous Galerkin Time-Stepping Method for Optimal Control Problems Governed by Parabolic Equations. *SIAM Journal on Numerical Analysis*, **42**, 1032-1061. <https://doi.org/10.1137/S0036142902397090>
- [3] Tröltzsch, F. (2010) Optimal Control of Partial Differential Equations: Theory, Procedures, and Applications. American Mathematical Society, Providence. <https://doi.org/10.1090/gsm/112>
- [4] Meidner, D. and Vexler, B. (2008) A Priori Error Estimates for Space-Time Finite Element Discretization of Parabolic Optimal Control Problems. Part I: Problems without Control Constraints. *SIAM Journal on Control and Optimization*, **47**, 1150-1177. <https://doi.org/10.1137/070694016>
- [5] Gong, W. and Yan, N. (2011) Mixed Finite Element Method for Dirichlet Boundary Control Problem Governed by Elliptic PDEs. *SIAM Journal on Control and Optimization*, **49**, 984-1014. <https://doi.org/10.1137/100795632>
- [6] Kunoth, A. and Schwab, C. (2013) Analytic Regularity and GPC Approximation for Control Problems Constrained by Linear Parametric Elliptic and Parabolic PDEs. *SIAM Journal on Control and Optimization*, **51**, 2442-2471. <https://doi.org/10.1137/110847597>
- [7] Borzi, A. (2010) Multigrid and Sparse-Grid Aschemes for Elliptic Control Problems with Random Coefficients. *Computing and Visualization in Science*, **13**, 153-160. <https://doi.org/10.1007/s00791-010-0134-4>
- [8] Gong, B., Sun, T., Shen, W. and Liu, W. (2016) A Priori Error Estimate of Stochastic Galerkin Method for Optimal Control Problem Governed by Random Parabolic PDE with Constrained Control. *International Journal of Computational Methods*, **13**, Article ID: 1650028. <https://doi.org/10.1142/S0219876216500286>
- [9] Ahmad Ali, A., Ullmann, E. and Hinze, M. (2017) Multilevel Monte Carlo Analysis for Optimal Control of Elliptic PDEs with Random Coefficients. *SIAM/ASA Journal on Uncertainty Quantification*, **5**, 466-492. <https://doi.org/10.1137/16M109870X>
- [10] Babuška, I., Nobile, F. and Tempone, R. (2010) A Stochastic Collocation Method for Elliptic Partial Differential Equations with Random Input Data. *SIAM Review*, **52**, 317-355. <https://doi.org/10.1137/100786356>
- [11] Gunzburger, D., Webster, G. and Zhang, G. (2014) Stochastic Finite Element Methods for Partial Differential Equations with Random Input Data. *Acta Numerica*, **23**, 521-650. <https://doi.org/10.1017/S0962492914000075>
- [12] Hou, L., Lee, J. and Manouzi, H. (2011) Finite Element Approximations of Stochastic Optimal Control Problems Constrained by Stochastic Elliptic PDEs. *Journal of Mathematical Analysis and Applications*, **384**, 87-103. <https://doi.org/10.1016/j.jmaa.2010.07.036>
- [13] Borzi, A. and Winkel, G. (2009) Multigrid Methods and Sparse-Grid Collocation Techniques for Parabolic Optimal Control Problems with Random Coefficients. *SIAM Journal on Scientific Computing*, **31**, 2172-2192. <https://doi.org/10.1137/070711311>
- [14] Barel, A.V. and Vandewalle, S. (2017) Robust Optimization of PDEs with Random Coefficients Using a Multilevel Monte Carlo Method. *SIAM/ASA Journal on Uncertainty Quantification*, **7**, 174-202. <https://doi.org/10.1137/17M1155892>
- [15] Guth, P.A., Kaarnioja, V., Kuo, F., Schillings, C. and Sloan, I.H. (2021) A Quasi-Monte Carlo Method for an Optimal Control Problem under Uncertainty. *SIAM/ASA Journal on Uncertainty Quantification*, **9**, 354-383. <https://doi.org/10.1137/19M1294952>
- [16] Alexanderian, A., Petra, N., Stadler, G. and Ghattas, O. (2017) Mean-Variance Risk-Averse Optimal Control of Systems Governed by PDEs with Random Parameter Fields Using Quadratic Approximations. *SIAM/ASA Journal on Uncertainty Quantification*, **5**, 1166-1192. <https://doi.org/10.1137/16M106306X>
- [17] Cliffe, K.A., Giles, M.B., Scheichl, R. and Teckentrup, A.L. (2011) Multilevel Monte Carlo Methods and Applications to Elliptic PDEs with Random Coefficients. *Computing and Visualization in Science*, **14**, Article No. 3. <https://doi.org/10.1007/s00791-011-0160-x>
- [18] Lord, G.J., Powell, C.E. and Shardlow, T. (2014) An Introduction to Computational Stochastic PDEs. Cambridge University Press, Cambridge. <https://doi.org/10.1017/CBO9781139017329>
- [19] Dai, Y.-H. and Yuan, Y. (1999) A Nonlinear Conjugate Gradient Method with a Strong Global Convergence Property. *SIAM Journal on Optimization*, **10**, 177-182. <https://doi.org/10.1137/S1052623497318992>
- [20] Evans, L.C. (2010) Partial Differential Equations. Second Edition, American Mathematical Society, Providence. <https://doi.org/10.1090/gsm/019/02>