doi：10.13413/j.cnki.jdxblxb.2024228

Bayes 推断和神经网络求解美式回望期权的隐含波动率

陶 李[1]，朱本喜[2]，钱译缘[2]，徐嘉琪[2]
（1. 海南大学 国际商学院，海口 570228；2. 吉林大学 数学学院，长春 130012）

摘要：首先，用原始对偶活跃集方法求解期权定价正问题，将相应的数值解作为监督学习的输出，然后用训练好的神经网络替代期权定价正问题模型. 其次，结合 Bayes 推断与神经网络进行 Metropolis-Hastings 采样，求解隐含波动率反问题. 该方法减少了采样过程中正问题计算量庞大的问题，从而加速了反问题求解过程.

关键词：隐含波动率；Bayes 推断；神经网络；替代模型
中图分类号：O241.8 **文献标志码**：A **文章编号**：1671-5489(2024)06-1363-07

Solving Implied Volatility of American Lookback Options by Bayesian Inference and Neural Network

TAO Li[1], ZHU Benxi[2], QIAN Yiyuan[2], XU Jiaqi[2]
(1. *College of International Business, Hainan University, Haikou* 570228, *China*;
2. *College of Mathematics, Jilin University, Changchun* 130012, *China*)

Abstract: Firstly, we used the original dual active set method to solve the forward problem of option pricing, with the corresponding numerical solutions as the output for supervised learning, and then replaced the forward problem model of option pricing with a well-trained neural network. Secondly, we combined Bayesian inference with neural networks for Metropolis-Hastings sampling to solve the inverse problem of implied volatility. This method reduced the problem of large computational complexity of the forward problem during the sampling process, thereby accelerating the solution process for the inverse problem.

Keywords: implied volatility; Bayesian inference; neural network; surrogated model

隐含波动率是期权的一个重要指标，它反映了市场对于标的价格的变动幅度. 投资者可以利用它来预测标的价格的未来走势和供求关系. 求解期权的隐含波动率，可视为解决期权定价模型的反问题. Achdou 等[1]使用最小二乘公式和下降算法实现了对波动率的求解；Burkovska 等[2]用缩基法代替经典的有限元法，从而减少了求解的复杂度；Kutner[3]利用基于二次近似法的数值方法，计算了美式期权的隐含波动率. 传统数值方法虽然能保证求解精度，但当模型为非线性或高维时，通常需要耗费大量计算时间，计算效率较低. 近年来，深度学习的快速发展提供了解决高维问题的新思路. Han 等[4]采用了一种基于深度学习的方法，处理一般的高维抛物型偏微分方程；Liu 等[5]建立了一个校准

收稿日期：2024-06-05.

第一作者简介：陶 李(1999—)，女，汉族，硕士，从事金融学的研究，E-mail：ttaoli99@163.com. **通信作者简介**：钱译缘(1996—)，女，汉族，博士，从事期权定价和随机控制优化的研究，E-mail：qianyy21@mails.jlu.edu.cn.

基金项目：国家自然科学基金面上项目(批准号：12271207).

神经网络框架，计算了隐含波动率. 上述基于深度学习求解高维偏微分方程的方法，在保证精度的情况下，显著减少了计算所需时间. 本文提出一个利用神经网络替代正问题的模型，并结合 Bayes 推断的数值方法，快速求解隐含波动率.

1 正问题描述和数值算法

设 V 是期权价格，S 是当前股票价格，J 是股票价格在过去到当前时间的最大值，t 是到期日与当前时刻的差值，σ 是波动率，r 是无风险利率，q 是股息收益率. 美式回望看跌期权价格 $V=V(S,J,t;\sigma)$ 满足如下自由边界问题[6]：

$$\mathscr{L}V:=\frac{\partial V}{\partial t}+\frac{1}{2}\sigma^2S^2\frac{\partial^2V}{\partial S^2}+(r-q)S\frac{\partial V}{\partial S}-rV=0,\quad B(t)<S\leqslant J<+\infty,\quad 0\leqslant t<T,\tag{1}$$

其中 $B(t)$ 为未知自由边界，收益函数为 $F(S,J,T)=J-S$ 且 $J=\max\limits_{0\leqslant\zeta\leqslant t}S_\zeta$，波动率函数为 $\sigma(t)=\sum\limits_{i=1}^{d}\sigma_ia_i(t)$. 这里，$\{a_i(t)\}_{i=1}^d$ 是 $L^2([0,T])$ 有限维逼近空间的一组已知基底(例如 $\{\sin(it)\}_{i=0}^{d-1}$，$\{e^{it}\}_{i=0}^{d-1}$)，$\{\sigma_i\}_{i=1}^d$ 为参数. 则相应的线性互补模型为

$$\begin{cases}\mathscr{L}V\cdot(V-G)=0, & JX\leqslant S\leqslant J<+\infty,\quad 0\leqslant t<T,\\ \mathscr{L}V\leqslant 0,\quad V\geqslant G, & JX\leqslant S\leqslant J<+\infty,\quad 0\leqslant t<T,\\ F(S,J,T)=(J-S), & JX\leqslant S\leqslant J<+\infty,\end{cases}\tag{2}$$

其中 $G=J-S$ 是赎回价格，X 满足：

$$X=\frac{-r+q+\sigma^2/2-\sqrt{(-r+q+\sigma^2/2)^2+2r\sigma^2}}{-r+q-\sigma^2/2-\sqrt{(-r+q+\sigma^2/2)^2+2r\sigma^2}}.$$

利用变换 $x=\ln\dfrac{J}{S}$，$\tau=T-t$ 和 $u(x,\tau)=\dfrac{V(S,J,t)}{S}$，线性互补模型(2)可改写为

$$\begin{cases}(\kappa(x)u_\tau-(\gamma\kappa(x)u_x)_x+q\kappa(x)u)\cdot(u-g)=0, & 0\leqslant x<L,\quad 0<\tau\leqslant T,\\ \kappa(x)u_\tau-(\gamma\kappa(x)u_x)_x+q\kappa(x)u\geqslant 0,\quad u-g\geqslant 0, & 0\leqslant x<L,\quad 0<\tau\leqslant T,\\ u(x,0)=g(x), & 0\leqslant x\leqslant L,\\ u_x(0,\tau)=0,\quad u(L,\tau)=g(L), & 0<\tau\leqslant T,\end{cases}\tag{3}$$

其中 $\gamma=\dfrac{\sigma^2}{2}$，$\kappa(x)=e^{(q-r-\gamma)x/\gamma}$，$g(x)=e^x-1$，$L=\ln\dfrac{1}{X}$.

设

$$H_L^1(\Omega):=\{v\in H^1(\Omega);\ v\geqslant g(x),\ v(L)=g(L)\},$$
$$H_{E0}^1(\Omega):=\{v\in H^1(\Omega);\ v\geqslant g(x),\ v(L)=0\},\qquad \Omega=[0,L].$$

引理 1 设 $u\in L^2(0,T;H^2(\Omega))$，$u_\tau\in L^2(0,T;L^2(\Omega))$，且 $u(0,\tau)>g(0)$，$0<\tau\leqslant T$. 则 u 是模型(3)的解当且仅当 u 是下列变分不等式问题的解：

$$(\kappa(x)u_\tau,v-u)+(\gamma\kappa(x)u_x,v_x-u_x)+(q\kappa(x)u,v-u)\geqslant 0,\quad \forall v\in H_{E0}^1(\Omega),\quad 0<\tau\leqslant T,$$

且 $u(x,0)=g(x)$.

令时间和空间的剖分分别为

$$\mathscr{T}_t:\ 0=\tau_0<t_1<\cdots<t_M=T,\quad \Delta t=\frac{T}{M},\quad t_j=j\Delta t,\quad j=0,1,\cdots,M,$$

$$\mathscr{T}_h:\ 0=x_0<x_1<\cdots<x_N=L,\quad h=\frac{L}{N},\quad x_i=ih,\quad i=0,1,\cdots,N,$$

其中 M 和 N 为正整数. 设 $V^h(\Omega)\subset H^1(\Omega)$ 为分片线性函数空间，取有限元试探函数空间和检验函数空间分别为

$$V_L^h(\Omega):=\{v_h\in V_h(\Omega);\ v(x_n)\geqslant g(x_n),\ n=0,1,\cdots,N-1,\ v_h(L)=g(L)\}=\mathrm{span}\{\varphi_i(x)\}_{i=1}^N,$$
$$V_{E0}^h(\Omega):=\{v_h\in V_h(\Omega);\ v(x_n)\geqslant g(x_n),\ n=0,1,\cdots,N-1,\ v_h(L)=0\}.$$

则半离散格式为：求 $u_h(\tau)\in V_L^h(\Omega)$，使得 $u_h(x,0)=g_I(x)$，且

$(\kappa(x)u_{h\tau},v_h-u_h)+(\gamma\kappa(x)u_{hx},v_{hx}-u_{hx})+(q\kappa(x)u_h,v_h-u_h)\geqslant 0,\quad \forall v_h\in V_{E0}^h(\Omega),\quad 0<\tau\leqslant T,$

其中 $g_I(x)$为 $g(x)$在 $V_L^h(\Omega)$中的插值.

进一步，在时间方向采取向后 Euler 格式，则全离散格式为

$(\kappa(x)((1+q\Delta\tau)u_h^m-u_h^{m-1}),v_h-u_h^m)+\Delta\tau(\gamma\kappa(x)u_{hx}^m,v_{hx}-u_{hx}^m)\geqslant 0,\quad \forall v_h\in V_{E0}^h(\Omega),\quad m=1,2,\cdots,M,$

其相应的矩阵形式为

$$(\boldsymbol{D}\boldsymbol{U}^m,\boldsymbol{V}-\boldsymbol{U}^m)\geqslant(\boldsymbol{W}^m,\boldsymbol{V}-\boldsymbol{U}^m),\quad \forall\boldsymbol{V}\geqslant\boldsymbol{F}_1,\quad m=1,2,\cdots,M,\tag{4}$$

其中 $\boldsymbol{D}=\gamma\Delta\tau\boldsymbol{A}+(1+q\Delta\tau)\boldsymbol{B}$，$\boldsymbol{W}^m=\boldsymbol{B}\boldsymbol{U}^{m-1}-\boldsymbol{F}_2$，$\boldsymbol{V}=(v_0,v_1,2,\cdots,v_{N-1})^{\mathrm{T}}$，$\boldsymbol{U}^m=(u_0^m,u_1^m,\cdots,u_{N-1}^m)^{\mathrm{T}}$，$\boldsymbol{A}=(A_{i,j})_{N\times N}$，$\boldsymbol{B}=(B_{i,j})_{N\times N}$，$i,j=0,1,\cdots,N-1$，$\boldsymbol{F}_1=(g(x_0),\cdots g(x_{N-1}))^{\mathrm{T}}$，$\boldsymbol{F}_2=(0,\cdots,0,g(L)\Delta\tau(q\boldsymbol{B}_{N-1,N}+\gamma\boldsymbol{A}_{N-1,N}))^{\mathrm{T}}$，$A_{i,j}=(\kappa(x)\varphi_j',\varphi_i')$，$B_{i,j}=(\kappa(x)\varphi_j,\varphi_i)$，$i,j=0,1,\cdots,N-1$.

算法 1[7-8]　原始对偶活跃集方法.

输入：$\boldsymbol{U}^0,\boldsymbol{F}_1,\boldsymbol{W}^m$；

输出：$\boldsymbol{U}^m,\boldsymbol{X}_l^m,\boldsymbol{X}_r^m$.

for $m=1:M$, do

$k=0$，$\rho>0$，$\varepsilon>0$，$\boldsymbol{U}^{(0)}=\max\{\boldsymbol{U}^{m-1},\boldsymbol{F}_1\}$，$\boldsymbol{U}^{(1)}=0$，$\lambda^{(0)}\geqslant 0$.

while $\|\boldsymbol{U}^{(k+1)}-\boldsymbol{U}^{(k)}\|_\infty<\varepsilon$, do

① 令活跃集为 $IS^{(k+1)}$，非活跃集为 $AS^{(k+1)}$：

$IS^{(k+1)}=\{i\in N,\lambda_i^{(k)}+\rho(G_i^m-U_i^{(k)})\leqslant 0\}$，

$AS^{(k+1)}=\{i\in N,\lambda_i^{(k)}+\rho(G_i^m-U_i^{(k)})>0\}$，

$N=\{0,1,\cdots,N-1\}$.

② 求解下列方程，得到 $\boldsymbol{U}^{(k+1)}$ 和 $\lambda^{(k+1)}$：

$\boldsymbol{D}\boldsymbol{U}^{(k+1)}-\lambda^{(k+1)}=\boldsymbol{W}^m$，

$\boldsymbol{U}^{(k+1)}=\boldsymbol{G}_i^m$，$i\in AS^{(k+1)}$，

$\lambda_i^{(k+1)}=0$，$i\in IS^{(k+1)}$.

end while

$\boldsymbol{U}^m=\boldsymbol{U}^{(k+1)}$，

$l^m=\min(IS^{(k+1)})$，$r^m=\max(IS^{(k+1)})$，

$x_l^m=x_{l^m}$，$x_r^m=x_{r^m}$.

end for

本文采用算法 1 对式(4)进行数值求解. 为方便，上述正问题整体求解过程，可以抽象为如下算子：

$$\boldsymbol{V}=T(\boldsymbol{\sigma}),\tag{5}$$

其中 $\boldsymbol{\sigma}=(\sigma_1,\cdots,\sigma_d)^{\mathrm{T}}$ 为波动率参数向量，$\boldsymbol{V}=(\boldsymbol{U}^0,\cdots,\boldsymbol{U}^M)$为期权离散矩阵，$T$ 为数值求解算子.

2　反问题描述和 Bayes 算法

在式(5)中，如果已知期权离散矩阵 $\boldsymbol{V}$ 和 $\boldsymbol{G}$，求波动率参数向量 $\boldsymbol{\sigma}$，则称为反问题. 该类问题是有确定性的，通常也是病态的，一般用 Tikhonov 正则化方法[9]求解该类问题. 但从统计学的角度，更关注的是真解的分布，而不再是一个点. 式(5)的统计模型版本可写为

$$\boldsymbol{V}=T(\boldsymbol{\sigma})+\boldsymbol{\delta},\tag{6}$$

其中 $\boldsymbol{\delta}=\{\delta_{ij}\}_{i=1,2,\cdots,N}^{j=1,2,\cdots,M+1}$，$\delta_{ij}\sim\delta\cdot N(0,1)$，$N(0,1)$是均值为 0、方差为 1 的标准正态分布.

2.1　Bayes 公式

Bayes 公式为

$$\pi_V(\boldsymbol{\sigma}):=\pi(\boldsymbol{\sigma}\,|\,\boldsymbol{V})=\frac{\pi(\boldsymbol{V}\,|\,\boldsymbol{\sigma})\pi_0(\boldsymbol{\sigma})}{\int_Z\pi(\boldsymbol{V}\,|\,\boldsymbol{\sigma})\pi_0(\boldsymbol{\sigma})\mathrm{d}\boldsymbol{\sigma}},\tag{7}$$

其中 $\pi_{\boldsymbol{V}}(\boldsymbol{\sigma})$ 和 $\pi_0(\boldsymbol{\sigma})$ 分别是波动率后验和先验的概率密度函数. 于是有

$$\pi_{\boldsymbol{V}}(\boldsymbol{\sigma})=\frac{\pi_\eta(\boldsymbol{V}-T(\boldsymbol{\sigma}))\pi_0(\boldsymbol{\sigma})}{\int_Z \pi_\eta(\boldsymbol{V}-T(\boldsymbol{\sigma})\pi_0(\boldsymbol{\sigma})\mathrm{d}\boldsymbol{\sigma}}\propto \pi_{\boldsymbol{\delta}}(\boldsymbol{V}-T(\boldsymbol{\sigma}))\pi_0(\boldsymbol{\sigma}), \tag{8}$$

其中$\propto$表示两者成比例的关系，$\pi_{\boldsymbol{\delta}}(\cdot)$为最大似然估计，并通常考虑先验为正态分布.

在线性正演模型和高斯先验的假设下，最大似然估计等价于 Tikhonov 正则化[10]. 从确定性的角度可见，通过优化框架只能得到反问题相关的一个点估计. 而从统计学的角度，可根据 Bayes 公式获得更多的统计信息. 而且这个后验分布，使得反演层参数的刻画更全面并包含不确定性[11]. 本文采用 Metropolis-Hastings(MH)算法，从后验分布 $\pi_{\boldsymbol{V}}(\boldsymbol{\sigma})$ 中进行采样.

2.2 Metropolis-Hastings 算法

Markov Chain Monte Carlo (MCMC)算法是一类用于从复杂分布中抽样的统计方法，其中 MH 算法是常用的 MCMC 方法之一[12].

算法 2[12] Metropolis-Hastings 算法.

输入：后验分布 $\pi_{\boldsymbol{V}}(\boldsymbol{\sigma})$，一个选定分布 q，最大样本数 M_1 和一个初始点 $\boldsymbol{\sigma}_0$；

输出：后验样本集合 $\{\boldsymbol{\sigma}_m\}_{m=1}^{M_1}$.

for $m=1:M_1$

提议一个候选点：$\boldsymbol{\sigma}_c\sim q(\cdot\,|\boldsymbol{\sigma}_{m-1})$；

计算接受概率：$\alpha(\boldsymbol{\sigma}_c,\boldsymbol{\sigma}_{m-1})=\min\left\{1,\frac{\pi_{\boldsymbol{V}}(\boldsymbol{\sigma}_c)q(\boldsymbol{\sigma}_{m-1}\,|\,\boldsymbol{\sigma}_c)}{\pi_{\boldsymbol{V}}(\boldsymbol{\sigma}_{m-1})q(\boldsymbol{\sigma}_c\,|\,\boldsymbol{\sigma}_{m-1})}\right\}$；

生成一个样本 $u\sim\mathcal{U}(0,1)$；

if $u<\alpha(\boldsymbol{\sigma}_c,\boldsymbol{\sigma}_{m-1})$

$\boldsymbol{\sigma}_m=\boldsymbol{\sigma}_c$；

else

$\boldsymbol{\sigma}_m=\boldsymbol{\sigma}_{m-1}$；

end if

end for

虽然 MH 采样算法可以计算真解的后验分布，但无论是 Tikhonov 正则化方法，还是 MH 采样算法，都需要求解大量的正问题(5). 虽然每个正问题求解只需几秒钟，但计算大量正问题，会增加整个算法的计算时间. 本文用神经网络的方法构造替代模型，以减少整体算法的计算量.

2.3 神经网络替代模型

全连接神经网络(FNN)是一种特殊结构的神经网络，它由一系列全连接层组成，对非线性函数可以有效地逼近. 全连接神经网络由$(L+1)$层网络组成，其中前一层的节点 x^{l-1} 与下一层的所有节点 x^l 都有连接. 在每个节点上，有两个函数依次复合构成. 第一个是一次函数 $\mathscr{L}^l(x^{l-1})=W^lx^{l-1}+b^l$：$\mathbb{R}^{n_{l-1}}\to\mathbb{R}^{n_l}$，第二个是激活函数 σ^l：$\mathbb{R}^{n_l}\to\mathbb{R}^{n_l}$(例如 Sigmoid,tanh,ReLU)，其中 W^l 表示权重，b^l 表示偏置. 图 1 为 FNN 的示意图. 为简便，令 $\Theta=\{W^l,b^l\}_{l=1}^L$ 表示 FNN 的所有网络参数，则可以定义一个由 Θ 参数化的映射 $\mathscr{FNN}$：$\mathbb{R}^m\to\mathbb{R}^n$：

$$y=\mathscr{FNN}(z;\Theta)=\mathscr{L}^L\circ\sigma^{L-1}\circ\mathscr{L}^{L-1}\circ\cdots\circ\sigma^1\circ\mathscr{L}^1(z). \tag{9}$$

用式(9)求解正问题(5)，即 $z=\boldsymbol{\sigma}$，$y=\boldsymbol{V}$，$\mathscr{FNN}\approx T$，$m=d$，$n=N\times(M+1)$. 设训练集为 $\{(\boldsymbol{\sigma}_i,\boldsymbol{V}_i)\}_{i=1}^{n_1}$，测试集为 $\{(\boldsymbol{\sigma}_i,\boldsymbol{V}_i)\}_{i=1}^{n_2}$，$n^*=n_1+n_2$ 表示样本集的总数. 对于监督学习，考虑平方损失函数，即

$$\mathrm{Loss}(\Theta)=\frac{1}{2n_1}\sum_{i=1}^{n_1}\|\boldsymbol{V}_i-\mathscr{FNN}(\boldsymbol{\sigma}_i;\Theta)\|^2. \tag{10}$$

本文主要目标是通过适当的优化算法，找到最优参数 Θ^*，即

$$\Theta^*=\arg\min_{\Theta}\mathrm{Loss}(\Theta). \tag{11}$$

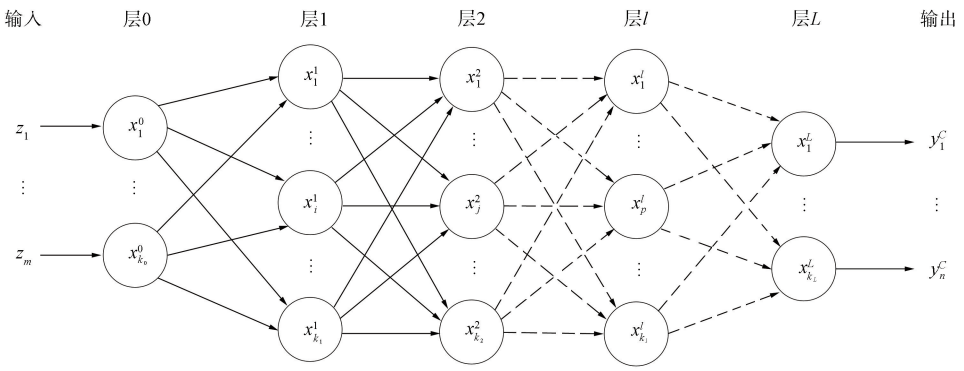


图 1　全连接神经网络结构示意图

Fig. 1　Schematic diagram of fully connected neural network

解决问题(10)的常见优化算法目前已有很多，如梯度下降法、牛顿法和随机梯度下降法等. 对本文所考虑的问题，通常 n_1 很大，导致计算梯度的代价很高. 为减少相应的计算量，这里使用最小批量梯度下降方法. 该方法将训练集划分为小批次，用于计算损失函数误差和梯度，并更新 Θ，综合了随机梯度下降法和梯度下降法的优势.

关于替代模型，本文有如下收敛性结果. 假设：

(H_1) 散度函数 f 满足

$$|f(t)| \leqslant \begin{cases} C_1 |\log t|, & 0 < t < 1, \\ C_2 t \log t, & t \geqslant 1, \end{cases}$$

其中 C_1 和 C_2 是依赖于 f 的两个常数.

引理 2[11]　对 $\forall \varepsilon > 0$，存在 $C = C(\varepsilon) > 0$，使得对任意 $\boldsymbol{\sigma} \in \mathbb{R}^d$，有

$$\| T(\boldsymbol{\sigma}) - \mathcal{FNN}(\boldsymbol{\sigma}; \Theta) \| \leqslant C \exp\{\varepsilon \| \boldsymbol{\sigma} \|^2\} \Phi_{\mathrm{T}}(\Theta),$$

其中 $\Phi_T(\cdot)$表示与 T 相关的任何函数，只需满足 $\lim\limits_{\sharp\{\Theta\} \to \infty} \Phi_{\mathrm{T}}(\Theta) = 0$，$\sharp\{\Theta\}$表示神经网络参数的个数.

设 μ_0 为先验测度，μ 为离散算子 $T(\cdot)$对应的真实模型后验测度，μ_F 为神经网络 $\mathcal{FNN}(\cdot;\Theta)$替代的后验测度，下列定理给出了在 f 散度下两者误差的上界.

定理 1　假设(H_1)成立，设测度 $\mu, \mu_F \ll \mu_0$，$T(\cdot)$和 $\mathcal{FNN}(\cdot;\Theta)$都是 Lipschitz 函数. 则测度 μ 和 μ_F 的误差在 f 散度下是收敛的，即存在常数 $C > 0$，使得

$$D_f(\mu \| \mu_F) := \int_{\mathbb{R}^d} f\left(\frac{\mathrm{d}\mu_1 / \mathrm{d}\mu_0}{\mathrm{d}\mu_2 / \mathrm{d}\mu_0}\right) \frac{\mathrm{d}\mu_2}{\mathrm{d}\mu_0} \mathrm{d}\mu_0(\sigma) \leqslant C \Phi_G(\Theta). \tag{12}$$

定理 1 的证明类似文献[11]，首先利用正演模型和替代模型是 Lipschitz 函数的性质，以及相应逼近的假设，并基于 Fernique 定理，得到归一化常数的估计. 然后，根据散度函数 f 具有分段上界的假设，以及正演模型和替代模型的假设，将 f 散度进行分段估计，从而可得结论. 定理 1 表明，当本文替代神经网络模型与正演模型足够逼近时，在 f 散度意义下相应的替代后验测度收敛到真实模型的后验测度，故替代模型有一定的理论保障.

3　数值实例

下面进行数值实验($d = 1, 3$)验证本文提出算法的有效性. 对于正问题(5)，选择 $T = 1$，$r = 0.03$，$q = 0.01$，并将转换后的截断长度设为 $L = 1.6$，以确保截断区域足够大. 对于离散化问题，时间和空间分割分别为 $M = 150$ 和 $N = 100$.

用 MH 和 MH+FNN 算法针对式(6)进行采样，并设先验分布 $\pi_0(\boldsymbol{\sigma}) \sim N_d(-1.8 \times \mathbf{1}_d, 0.1 \times \boldsymbol{I}_d)$，取密度 $q(\cdot;\boldsymbol{\sigma}) \sim N_d(\boldsymbol{\sigma}, 10^{-6} \times \boldsymbol{I}_d)$. 此外，采样噪声振幅为 $\delta = 0.012$. 在产生 $\boldsymbol{\sigma}$ 的若干样本后，可利用算法 1 对现性互补问题(4)进行求解，得到正问题的若干组解(即期权离散矩阵 $\boldsymbol{V}$)，并作为监督学习目标泛函中的真解. 然后，以 $\boldsymbol{\sigma}$ 的若干样本为输入参数，正问题的若干组解为输出，经过优化训练，得到神经网络 FNN 的参数，并用该网络代替正问题的求解.

取神经网络结构为$\{d,50,50,50,50,151\times100\}$，$d=1,3$. 使用 Adam 优化器对网络进行训练，每次迭代使用128个样本的小批量，并进行5 000次循环的训练，而权重初始化通过 Glorot 均匀初始化. 这里，每次使用原始对偶活跃集方法，对式(6)进行一次求解，可得一个期权离散矩阵$\boldsymbol{V}$，即为一个样本. 取总样本数$n^*=20\ 000$，其中训练集的样本数$n_1=16\ 000$，测试集的样本数$n_2=4\ 000$. 所有的数值实验都在一台 Intel Core i7 的计算机工具包 MATLAB R2020b 上实现.

例1 考虑如下一维的波动率：

$$\sigma(t)=\sigma a(t),\qquad a(t)=0.3\mathrm{e}^t. \tag{13}$$

表1列出了当$\sigma=0.15,0.3,0.45$时，Bayes 方法(MH)和 Bayes+替代模型(MH+FNN)的数值结果，其中时间包括线下时间和线上时间. 这里，FNN 的实验输出的是期权价格V，但 MH 和 MH+FNN 实验输出的是波动率σ. 由表1可见，两种方法的精度近似，但后者花费的时间只有前者的1/3左右，故神经网络作为替代模型，加速效果显著.

表1 当$d=1$时，MH 和 MH+FNN 算法的数值结果

Table 1 Numerical results of MH and MH+FNN algorithms when $d=1$

σ	MH			MH+FNN		
	时间/s	均值	$10^2\times$标准差	时间/s	均值	$10^2\times$标准差
0.15	2 341.5+0	0.149	0.45	824.6+6.3	0.153	0.549
0.30	2 361.2+0	0.298	0.52	835.9+6.6	0.292	0.553
0.45	2 487.1+0	0.447	0.57	847.5+6.7	0.441	0.622

表2列出了 MH+FNN 算法固定每层神经元为50时，不同隐藏层的数值结果. 表3列出了 MH+FNN 算法固定隐藏层为3时，不同神经元个数的数值结果. 表2和表3的数值结果表明，随着隐藏层和神经元个数的增加，精度会有些提高，但会增加一定的时间. 对于增加样本，有类似的结果，这里不再列出.

表2 当每层神经元为50时不同隐藏层的数值结果

Table 2 Numerical results of different hidden layers when each layer of neurons is 50

隐藏层层数	时间/s	均值σ
2	826.2	0.148，0.298，0.447
3	917.3	0.149，0.299，0.448
4	975.6	0.151，0.299，0.449

表3 当隐藏层为3时不同神经元的数值结果

Table 3 Numerical results of different neurons when hidden layer is 3

神经元个数	时间/s	均值σ
25	825.1	0.148，0.297，0.447
50	1 018.9	0.149，0.298，0.448
75	1 105.2	0.151，0.299，0.449

例2 考虑含3个参数的波动率，即

$$\sigma(t)=\sum_{i=1}^{3}\sigma_i a_i(t),$$

其中$a_1(t)=1$，$a_2(t)=t$，$a_3(t)=3^{-t}$，波动率的真解为$\sigma_1=0.1$，$\sigma_2=0.15$，$\sigma_3=0.2$. 表4列出了$(\sigma_1,\sigma_2,\sigma_3)=(0.1,0.15,0.2)$时，Bayes 方法(MH)和 Bayes+替代模型(MH+FNN)的数值结果. 与一维的结果类似，MH+FNN 的计算结果和 MH 类似，但时间上要快2/3左右.

表4 当$d=3$，$\boldsymbol{\sigma}=(0.1,0.15,0.2)$时，MH 与 MH+FNN 算法的数值结果

Table 4 Numerical results of MH and MH+FNN algorithms when $d=3$，$\boldsymbol{\sigma}=(0.1,0.15,0.2)$

算法	时间/s	均值	$10^2\times$标准差
MH	2 647.5+0	(0.122，0.146，0.195)	0.64
MH+FNN	863.2+7.2	(0.124，0.145，0.192)	0.69

图2给出了 MH 和 MH+FNN 算法采样的相关图. 由图2可见：(A)，(C)，(F)是自相关图，其两

条线基本重合，说明自相关度非常高；(B),(D),(E)是协相关图，其各自的两个环线也基本重合，说明协相关度很高. 实验结果表明，两种采样效果一致.

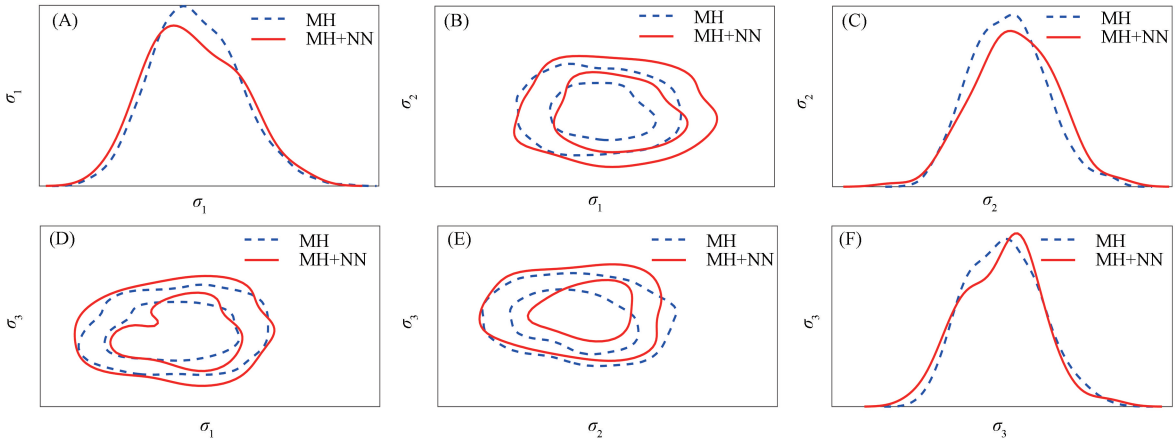


图 2　MH 和 MH＋FNN 算法两种后验分布的相关图

Fig. 2　Correlation diagrams of two posterior distributions for MH and MH＋FNN algorithms

参　考　文　献

[1] ACHDOU Y, INDRAGOBY G, PIRONNEAU O. Volatility Calibration with American Options [J]. Methods Appl Anal, 2004, 11(4): 533-556.

[2] BURKOVSKA O, GLAU K, MAHLSTEDT M, et al. Complexity Reduction for Calibration to American Options [J]. J Comput Finance, 2019, 23(1): 25-60.

[3] KUTNER G W. Determining the Implied Volatility for American Options Using the QAM [J]. Financ Rev, 1998, 33(1): 119-130.

[4] HAN J Q, JENTZEN A, E W N. Solving High-Dimensional Partial Differential Equations Using Deep Learning [J]. Proc Natl Acad Sci USA, 2018, 115(34): 8505-8510.

[5] LIU S Q, LEITAO Á, BOROVYKH A, et al. On a Calibration Neural Network to Extract Implied Information from American Options [J]. Appl Math Finance, 2020, 28(5): 449-475.

[6] BLACK F, SCHOLES M. The Pricing of Options and Corporate Liabilities [J]. J Polit Econ, 1973, 81(3): 637-654.

[7] GAO Y, SONG H M, WANG X S, et al. Primal-Dual Active Set Method for Pricing American Better-of Option on Two Assets [J]. Commun Nonlinear Sci Numer Simulat, 2020, 80: 104976-1-104976-15.

[8] SONG H M, WANG X S, ZHANG K, et al. Primal-Dual Active Set Method for American Lookback Put Option Pricing [J]. East Asian J Appl Math, 2017, 7(3): 603-614.

[9] BOUCHOUEV I, ISAKOV V. The Inverse Problem of Option Pricing [J]. Inverse Probl, 1997, 13(5): L11-L17.

[10] LEHTINEN M S, PÄIVÄRINTA L, SOMERSALO E. Linear Inverse Problems for Generalised Random Variables [J]. Inverse Probl, 1989, 5(4): 599-612.

[11] GAO Y, LIU H Y, WANG X C, et al. A Bayesian Scheme for Reconstructing Obstacles in Acoustic Waveguides [J]. J Sci Comput, 2023, 97(3): 53-1-53-45.

[12] HASTINGS W K. Monte Carlo Sampling Methods Using Markov Chains and Their Applications [J]. Biometrika, 1970, 57(1): 97-109.

（责任编辑：李　琦）