

随机波动率 Hull-White 模型参数估计方法

江良¹, 林鸿熙²

(1. 莆田学院数学学院, 福建 莆田 351100; 2. 莆田学院商学院, 福建 莆田 351100)

摘要: 构建随机波动率的两因子模型, 应用两阶段半参数方法估计模型中的常系数参数, 使用核估计方法估计长期均值函数, 给出了两阶段估计方法的相容性和参数的渐近性性质. 实证结果表明了对比常系数模型, 引入长期均值函数模型将会改善似然函数估计值, 而且也能够很好地解释中央银行和政府已实施政策的有效性. 此外, 可以在不增加维数的条件下, 使用该模型对利率衍生品进行更有效地定价.

关键词: 长期均值; 随机波动率; 短期利率模型; 半参数估计; 核估计方法

中图分类号: F830.9; O212.7 文献标识码: A 文章编号: 1000-5781(2016)05-0633-10

doi: 10.13383/j.cnki.jse.2016.05.008

Parametric estimation of Hull-White model for stochastic volatility

Jiang Liang¹, Lin Hongxi²

(1. School of Mathematics, Putian University, Putian 351100, China;

2. School of Business, Putian University, Putian 351100, China)

Abstract: A two-factor model of stochastic volatility is established. A two-stage semi-parameter method is applied to estimate constant coefficient parameters of this model. Moreover, kernel estimator method is developed to estimate the long-term mean value function, by this method the consistency of the two-stage method and the asymptotic normality of parameters are obtained. The empirical results show that the likelihood function can be improved in the long-term mean value model rather than the constant coefficient model. Also, the model provides a good explanation for the effective policies implemented by the central bank and the government. Besides, the industries can use the above model for valuing interest-rate-derivative securities without increasing the dimension.

Key words: long-term average value; stochastic volatility; short term interest rate model; semi-parameter estimation; kernel estimation method

1 引言

利率是金融市场中非常重要的一个指标, 几乎所有的金融现象和活动都和利率相关. 因此构建合适的短期利率模型变得尤为重要. 构建短期利率模型可分为两类: 单因子模型; 多因子模型. 虽然实证表明了单因子模型也能很好地拟合市场的数据^[1,2], 但是单因子模型无法产生较复杂的收益率曲线的形状. 因此, 多因子短期利率模型应运而生^[3,4].

收稿日期: 2013-12-20; 修订日期: 2014-05-02.

基金项目: 国家自然科学基金资助项目(11471175); 福建省自然科学基金资助项目(2015J05012; 2016J01677); 莆田学院育苗基金资助项目(2014060; 2014061).

Litterman 等^[5]基于实证方法论述了随机波动率模型的必要性. Kim 等^[6]应用多因子模型对日本债券收益率数据进行分析发现当瞬时利率很小时, 两因子高斯模型拟合效果较好. Cheridito 等^[7]和 Pierre 等^[8]基于横截面数据相应给出多因模型的参数估计方法及实证的结果. 虽然, 上述的研究结果表明了可以通过引入状态变量改善模型的拟合效果, 但是相应的衍生品定价的维数也增加从而导致数值方法也变得更加困难, 而且相应的参数估计问题也变的困难. 如 Duffee 等^[9]指出了由于多因子模型的复杂性, 相应的参数估计识别问题也变得困难. 因此, 为了给出最优的统计方法, 选择了状态变量相互独立的仿射性结构模型.

为了给出简约的模型, Andersen 等^[10], Fong 等^[11], Longstaff 等^[12] 及 Surya^[13]构建了随机波动率两因子模型, 并通过实证分析说明了引入随机波动率是必要的. 然而在上述的模型中, 长期均值均是常数. 因此, Balduzzi 等^[14,15], 和 Chen^[16]研究了引入长期均值作为状态变量的短期利率模型, 并通过实证和数值分析说明了该模型不仅改变了短期利率期限结构的形状而且也改善了拟合效果. 范龙振等^[17]在漂移项中引入了储蓄存款利率作为状态变量来改善拟合效果. 但是, 由于引入新的状态变量, 相应的衍生品定价偏微分方程的维数也会增加, 从而导致数值方法变得比较复杂. 另一方面, 在文献[14,15]中, 长期均值几乎和瞬时利率呈现一样的震荡, 这种现象很难精确地刻画长期预期利率的趋势. 江良^[18]也通过核估计方法估计了 Hull-White 模型中的参数, 但该作者考虑的模型是单因子模型.

基于上面分析的原因, 本文用时间函数描述长期均值的变化, 考虑了随机波动率的短期利率模型, 并相应地给出了半参数估计方法. 该模型既能改善数据的拟合效果也不增加衍生品定价的维数, 并兼顾随机波动率良好的特性. 而且实证分析也表明了引入均值函数能更好地刻画中央银行和政府已实施政策的有效性(详细的分析见本文第 4 节). 此外, 考虑均值函数的模型也可改善衍生品的定价. 如 Hull 等^[2]研究了时间变量系数的单因子模型衍生品定价时, 他们发现其衍生品的定价和常数系数的两因子模型没有显著的差异. Grzelak 等^[19]引入 Hull-White 模型给出债券定价过程, 并指出引入 Hull-White 模型是必要的.

为了说明本文模型的有效性, 将使用似然函数来诊断模型. 由于模型中含有时间函数的参数, 因此将提出一种半参数估计方法来计算似然函数值. 根据 Härdle 等^[20]和 Ramsay 等^[21]对于半参数模型的估计方法一般可分为如下两类: 核估计方法; 正则化方法. 本文将选择核估计方法, 其主要原因是均值函数可通过核函数在每个节点上加权平均近似, 其近似的函数仅依赖于给定窗口. 此外, 由于瞬时波动率是不可观测的, 因此将通过两组不同到期日债券的线性组合来近似波动率过程^[16,22]. 本文模型是在风险中性测度下所构建的, 这就使得对利率衍生品的定价时可直接使用这些参数的估计值.

2 随机波动率Hull-White短期利率模型

首先, 假设在风险中性测度下短期利率满足下面的模型(SVHW)

$$dr_t = (\theta(t) - ar_t)dt + \sqrt{V_t}dW_t^r, \quad (1)$$

$$dV_t = \alpha(\beta/\alpha - V_t)dt + \sigma_v\sqrt{V_t}dW_t^v, \quad (2)$$

其中 W_t^r, W_t^v 是一对标准布朗运动, $E[dW_t^r dW_t^v] = \rho dt$; $\alpha, \beta/\alpha$, 和 σ_v 分别表示随机波动率的速率, 均值和波动率.

在该模型中瞬时波动率满足 CIR 模型^[23]以至于使得波动率是非负的(假设 Feller 条件满足), 而且长期波动率是一个常数. 一个似乎更合理的假设是式(1)满足 CIR 随机过程, 然而归咎于相关系数 ρ , 此时债券价格无仿射性结构解^[24]. 但是相关系数又是一个非常重要的指标. 如, Balduzzi 等^[15]发现了利率和随机波动率水平变化是正相关的. Cheridito 等^[7]通过债券数据发现了相关系数也是正的(几乎接近1). 基于这些原因, 不考虑对于随机利率满足 CIR 动态过程的模型.

注意到, SVHW 模型包含一些其它的模型. 当波动率是常数时, 就是众所周知的 HW 模型^[2]; 当 $\theta(t)$ 是常数时, 模型变化为 FV 模型^[11]; 当波动率 V_t 和 $\theta(t)$ 都是常数时, 其模型为 Vasicek 模型^[25].

设 $P(t, T) = P(t, T, V_t, r_t)$ 为在 t 时刻到期日为 T 的债券价格. 设信息流 \mathcal{F}_t 是关于 V_t 和 r_t 在 t 之前的所有信息. 那么债券价格为 $P(t, T) = E \left[e^{-\int_t^T r_\tau d\tau} | \mathcal{F}_t \right]$, 其中 $E[\cdot]$ 为期望值算子.

根据计价单位转换原理^[24], 如果选取银行账户 $\bar{B}(t) = \exp(\int_0^t r_u du)$ 作为计价单位, 则 $P(t, T)/\bar{B}(t)$ 是一个鞅, 其微分为

$$d \left(e^{-\int_0^t r_\tau d\tau} P \right) = e^{-\int_0^t r_\tau d\tau} \cdot$$

$$\left(\left(\frac{\partial P}{\partial t} + \frac{1}{2} \sigma_v^2 V \frac{\partial^2 P}{\partial V^2} + \sigma_v \rho V \frac{\partial^2 P}{\partial r \partial V} + \frac{1}{2} V \frac{\partial^2 P}{\partial r^2} + (\beta - \alpha V) \frac{\partial P}{\partial V} + (\theta(t) - ar) \frac{\partial P}{\partial r} - rP \right) dt + \left(\sigma_v \sqrt{V} \frac{\partial P}{\partial V} dW_t^v + \frac{\partial P}{\partial r} dW_t^r \right) \right).$$

根据鞅的性质, 可得以下偏微分方程.

$$\frac{\partial P}{\partial t} + (\theta(t) - ar) \frac{\partial P}{\partial r} + \frac{1}{2} V \frac{\partial^2 P}{\partial r^2} + (\beta - \alpha V) \frac{\partial P}{\partial V} + \frac{1}{2} \sigma_v^2 V \frac{\partial^2 P}{\partial V^2} + \sigma_v \rho V \frac{\partial^2 P}{\partial r \partial V} - rP = 0, \quad (3)$$

其终端条件为 $P(T, T) = 1$.

根据 Duffie 等^[4]研究的结果, 方程式(3)存在仿射性结构解, 即 $P(t, T) = \exp(A(t, T) - rB(t, T) - VC(t, T))$, 其中 $A = A(t, T)$, $B = B(t, T)$ 及 $C = C(t, T)$ 满足下面的常微分方程组, 即

$$\begin{cases} B_t - aB + 1 = 0 \\ C_t - (\alpha + \sigma_v \rho B)C - \frac{1}{2} \sigma_v^2 C^2 - \frac{1}{2} B^2 = 0 \\ A_t - \theta(t)B - \beta C = 0, \end{cases} \quad (4)$$

其终端条件为 $A(T, T) = B(T, T) = C(T, T) = 0$. 显然在给定终端条件下 $B(t, T)$ 的解析解为 $B(t, T) = (1 - e^{-a(T-t)})/a$. 而常微分方程(4)是一个众所周知的 Riccati 方程. 由于该方程的系数是关于时间变量的函数, 因此一般没有解析解, 所以将使用数值方法求解. 若给定 C 和 B 的解, 则 A 的解为

$$A(t, T) = - \int_t^T \theta(t)B(t, T)dt - \beta \int_t^T C(t, T)dt.$$

3 参数校正方法

不失一般性, 假设债券价格是从当前时刻开始观察. 为了能给出参数估计方法, 首先需要处理隐含状态变量 V_t . 对于本文的问题, 将利用不同到期日可观测的债券价格数据近似随机波动率.

设 $R(t, T)$ 为在 t 时刻到期日为 T 的零息债券收益率, 则

$$R(t, T)(T - t) = -\ln P(t, T) = -A(t, T) + r_t B(t, T) + V_t C(t, T).$$

假设 $R(t, T_1)$ 和 $R(t, T_2)$ 是两个不同到期日收益率 ($T_1 \neq T_2$), 可得

$$R(t, T_1)(T_1 - t) = -\ln P(t, T) = -A(t, T_1) + r_t B(t, T_1) + V_t C(t, T_1), \quad (5)$$

$$R(t, T_2)(T_2 - t) = -\ln P(t, T) = -A(t, T_2) + r_t B(t, T_2) + V_t C(t, T_2). \quad (6)$$

从式(5)和式(6)可得

$$V_t(B_2 C_1 - B_1 C_2) = (B_2 R(t, T_1)(T_1 - t) - B_1 R(t, T_2)(T_1 - t)) + A_1 B_2 - A_2 B_1, \quad (7)$$

其中 $A_m = A(t, T_m)$, $B_m = B(t, T_m)$ 及 $C_m = C(t, T_m)$, $m = 1, 2$.

为了简化, 在实际应用中对于表达式(7)通过线性组合来近似其动态的过程, 即

$$V_t \approx \alpha_0 + \alpha_1 (B_2 R(t, T_1)(T_1 - t) - B_1 R(t, T_2)(T_2 - t)), \quad (8)$$

其中 α_0 和 α_1 是常数. 显然式(8)仅依赖于参数 α_0, α_1 及 a . 比较动态过程(2), 辅助过程(8)简化参数估计过程.

下面将通过式(1)和辅助过程式(8)建立似然函数并给出半参数估计方法.

设 $\Delta t = t_i - t_{i-1} (i = 1, 2, \dots, N)$, 其中 t_N 是最大的观察值及 $t_0 = 0$. 式(1)的欧拉离散形式为

$$r_{i+1} - r_i \approx (\theta_i - ar_i)\Delta t + \sqrt{V_i} \sqrt{\Delta t} z_{i+1}, \quad i = 0, 1, 2, \dots, N-1, \quad (9)$$

其中 $r_i = r_{t_i}, V_i = V(t_i), \theta_i = \theta(t_i)$ 及 z_{i+1} 是一个服从标准正态分布的随机数. 若给定 V_i , 式(9)的离散式是有偏的. 尽管在给定 V 的条件下, 随机微分方程(1)有解析表达式^[24], 但是, 如果 Δt 足够小, 有偏的现象将会减少. 实际上, Glasserman^[26]论述了式(9)的离散式是可行的. 相应式(8)的离散过程为

$$V_i = \alpha_0 + \alpha_1(B_2 R(t_i, T_1)(T_2 - t_i) - B_1 R(t_i, T_2)(T_2 - t_i)). \quad (10)$$

根据式(9)和式(10), 似然函数为

$$L(\theta, \boldsymbol{\eta}) = \ln \prod_{i=0}^{N-1} f(r_{i+1}|r_i, \boldsymbol{\eta}, \theta) = - \sum_{i=0}^{N-1} \left(\frac{1}{2} \ln V_i^2 + \frac{(r_{i+1} - r_i(1 - a\Delta t) - \theta_i \Delta t)^2}{2V_i^2 \Delta t} \right) + c, \quad (11)$$

其中 c 是不依赖于 θ_i 和 $\boldsymbol{\eta}$ 的常数, $\boldsymbol{\eta} = (a, \alpha_0, \alpha_1)$, $f(\cdot|\cdot, \cdot, \cdot)$ 是条件密度函数, r_i 是 $\mathbf{r} = (r_0, r_1, \dots, r_N)$ 的分量, $i = 0, 1, 2, \dots, N$.

为了估计 $\theta(t)$ (其离散值为 θ_i), 引入众所周知 Nadaraya-Watson 核函数估计^[20]. 那么 $\theta(t)$ 可近似为

$$\theta_h(t) = \sum_{i=0}^{N-1} W_i(t) Y_i, \quad (12)$$

其中 $Y_i = r_{i+1} - r_i(1 - a\Delta t)$, $W_i(t)$ 是权重函数 $W_i(t) = K_h(t - t_i) \left(\sum_{i=0}^{N-1} K_h(t - t_i) \right)^{-1}$, h 是窗口, $K_h(s) = h^{-1}K(sh^{-1})$, 其中 $K(\cdot)$ 是核函数. 选取合适的 h 值将会改善 $\theta(t)$ 的光滑性. 若 $h \rightarrow 0$, $\theta_h(t) \rightarrow Y_i$, 这隐含着对于 $\theta(t)$ 是过分估计. 若 $h \rightarrow \infty$, $\theta_h(t) \rightarrow N^{-1} \sum_{i=0}^{N-1} Y_i$, 意味着 $\theta(t)$ 是常数. 一般来讲, h 选取比核函数选取更加重要, 如 Härdle 等^[20]指出不同核函数选取最优窗口 h 之间相差一项常数因子. 因此本文将选取高斯核函数.

下面将给出基于时间序列数据的两因子模型极大似然估计方法. 基于式(11)和式(12), 似然函数近似为

$$\ell(\theta, \boldsymbol{\eta}) = \ell(\theta_h, \boldsymbol{\eta} | \mathbf{r}, h) = - \sum_{i=0}^{N-1} \left(\frac{1}{2} \ln V_i^2 + \frac{(r_{i+1} - r_i(1 - a\Delta t) - \theta_h(t_i)\Delta t)^2}{2V_i^2 \Delta t} \right), \quad (13)$$

其中 V_i 用式(10)来近似计算.

注意到, 若直接最大化似然函数式(13), 其过程比较复杂, 因此基于 Simar 等^[27]的思想, 采用两阶段估计方法. 其核心思想是把上述问题化为两个简单的问题: 一个非参数问题; 一个全参数问题, 其算法如下:

步骤 1 给定一个初值 $\boldsymbol{\eta}_0$;

步骤 2 求优化问题(13), 即

$$\hat{\theta}_h(t, \boldsymbol{\eta}) = \operatorname{argmax} \ell(\theta_h(t), \boldsymbol{\eta}); \quad (14)$$

步骤 3 根据步骤 2 的估计结果 $\hat{\theta}_h(t, \boldsymbol{\eta})$, 求下面的优化问题

$$\hat{\boldsymbol{\eta}} = \operatorname{argmax}_{\boldsymbol{\eta} \in \Omega(\boldsymbol{\eta})} \ell(\hat{\theta}_h(t, \boldsymbol{\eta}_0), \boldsymbol{\eta}), \quad (15)$$

其中 $\Omega(\boldsymbol{\eta})$ 是参数 $\boldsymbol{\eta}$ 所在的领域;

步骤 4 重复步骤 2 和步骤 3 直到收敛.

显然, 在给定参数 $\boldsymbol{\eta}$ 条件下, 优化问题(14)转化为一个标准的非参数估计问题, 即

$$\hat{\theta}_h(\boldsymbol{\eta}) = \operatorname{argmin} \sum_{i=0}^{N-1} (Y_i - \theta_h(t_i))^2. \quad (16)$$

在算法步骤 3 中, 将使用 Ait-Sahalia 等^[22] 的极大似然估计方法. 注意在使用算法时, 需要回答两个主要的问题: 步骤 2 和步骤 3 是否相容的; 步骤 3 的渐近性质.

接下去部分, 将回答这两个问题. 首先考虑算法步骤 2, 若给定真实的 η^* , 当 $h \rightarrow 0$, $\theta_h(t, \eta^*)$ 能否依概率收敛到相应的真实值 θ^* . 为了证明其相容性, 需要下面的假设.

假设 1 在给定 η 条件下, 假设 Y_i 的真实密度函数为 $f(Y|\eta)$, 其均值和方差满足下面的条件, 即

$$0 < E[Y] < \infty, \text{ 且 } \text{Var}(Y) < \infty,$$

其中 $\text{Var}(\cdot)$ 表示方差.

假设 1 说明了在使用数据时, 相应的均值和方差是有界的. 现实中, 瞬时利率的时间序列数据的均值和方差一定是可满足上面条件的. 而且从假设 1 可以得出二阶矩是有界的, 即 $E[Y^2] = \text{Var}(Y) + (E[Y])^2 < \infty$. 这个条件保证 $\theta(\eta, t)$ 是均方收敛的, 从而可以得到定理 1 的结果. 根据假设 1, 给出下面相容性定理, 其证明过程可参考文献[28]中的性质 3.1.1.

定理 1 假设 1 成立. 如果 $Nh \rightarrow \infty$ 和 $h \rightarrow 0$ 那么有

$$\theta_h(t, \eta) \xrightarrow{P} \hat{\theta}(\eta), \tag{17}$$

其中 \xrightarrow{P} 表示依概率收敛.

定理 1 说明了, 若给定真实值 η^* , 通过优化问题(14)所得 θ 值将概率收敛到 θ^* . 因此, 需要进一步考虑优化问题(15)的相容性问题. 此时, 需要一些正则性假设.

假设 2 假设 $\Omega(\eta)$ 是紧的. $\theta^* = \theta^*(t)$ 及 η^* 是最大似然函数(13)的真实解且是 $\Omega(\eta)$ 的内点.

假设 2 表明了, 在算法中对于常系数取值必须是有界的, 而在实际应用中取有界的常系数是显然的. 另一方面假设解是 $\Omega(\eta)$ 的内点是数学上技术处理.

假设 3 定义下面的极限存在且一致收敛, 即

$$d(\eta; \theta) = \lim_{N \rightarrow +\infty} d_N(\eta; \theta) = \lim_{\frac{1}{N}} (\ell(\eta, \theta) - \ell(\eta^*, \theta)), \quad \forall \theta \in \Omega(\theta).$$

进一步假设, 当 $\theta = \theta^*$, $d(\eta; \theta^*) = 0$ 解是唯一, 即 $\eta = \eta^*$.

假设 3, 实际是 Kullback 等^[29] 相对熵的概念. 若假设 $\ell(\eta^*, \theta^*)$ 的真实密度函数为 $f(r|\eta^*, \theta^*) > 0$, 根据中心极限定理直接获得

$$d(\eta) = - \int \ln \frac{f(r|\eta, \theta^*)}{f(r|\eta^*, \theta^*)} f(r|\eta^*, \theta^*) dr.$$

基于文献[29]中引理3.1, $d(\eta) \leq 0$, 而且等号成立当且仅当 $\eta = \eta^*$.

假设 4 假设给定 η , $\ell(\eta, \theta)$ 关于 $\theta(t)$ 是 Lipschitz 连续且 Fréchet 导数存在并有界. 另一方面, 给定 θ , $\nabla \ell(\eta, \theta)$ 向量的每个元素存在且 Lipschitz 连续. 矩阵 $\nabla^2 \ell(\eta, \theta)$ 每个元素有定义并 Lipschitz 连续, 而且该矩阵是可逆的.

在上面的正则性假设条件下, 下面定理说明了算法中式(15)的相容性性质.

定理 2 设定理 1 中的所有假设及假设 4 成立. 假设 $\hat{\theta}$ 是通过算法中步骤 2 估计得到. 下面的相容性成立, 即

$$\sup_{\eta \in \Omega(\eta)} (d_N(\eta; \hat{\theta}) - d_N(\eta; \theta^*)) \xrightarrow{P} 0. \tag{18}$$

若假设 3 成立, 那么有下面相容性估计

$$\hat{\eta} \xrightarrow{P} \eta^*. \tag{19}$$

证明 根据定义

$$\begin{aligned} d_N(\boldsymbol{\eta}; \hat{\boldsymbol{\theta}}) - d_N(\boldsymbol{\eta}; \boldsymbol{\theta}^*) &= \frac{1}{N} (\ell(\boldsymbol{\eta}, \hat{\boldsymbol{\theta}}) - \ell(\boldsymbol{\eta}^*, \hat{\boldsymbol{\theta}}) - \ell(\boldsymbol{\eta}, \boldsymbol{\theta}^*) + \ell(\boldsymbol{\eta}^*, \boldsymbol{\theta}^*)) \\ &= \frac{1}{N} (\ell(\boldsymbol{\eta}, \hat{\boldsymbol{\theta}}) - \ell(\boldsymbol{\eta}, \boldsymbol{\theta}^*)) + \frac{1}{N} (\ell(\boldsymbol{\eta}^*, \boldsymbol{\theta}^*) - \ell(\boldsymbol{\eta}, \boldsymbol{\theta}^*)). \end{aligned}$$

根据假设 4 及定理 1, 显然式(18)成立. 基于文献[30]中定理 5.7, 式(18)成立隐含着式(19)成立. 证毕. 最后, 给出算法中式(15)渐近性. 首先, 需要给出 Fisher 信息矩阵.

假设 5 假设下列的极限存在且一致收敛, 即 $\lim_{N \rightarrow +\infty} (\nabla^2 \ell(\boldsymbol{\eta}^*, \boldsymbol{\theta}^*)) / N = -I(\boldsymbol{\eta}^*, \boldsymbol{\theta}^*)$.

定理 3 设假设 1 ~ 假设 5 成立, 那么有下面的渐近性,

$$\sqrt{N}(\hat{\boldsymbol{\eta}} - \boldsymbol{\eta}^*) \xrightarrow{D} \mathcal{N}(0, I(\boldsymbol{\eta}^*, \boldsymbol{\theta}^*)^{-1}),$$

其中记号 \xrightarrow{D} 表示依分布收敛.

证明 根据极大似然估计方法, 优化问题(13)在极值点上梯度为零, 即

$$0 = \nabla \ell(\boldsymbol{\eta}, \hat{\boldsymbol{\theta}}) \Big|_{\boldsymbol{\eta}=\hat{\boldsymbol{\eta}}} = \nabla \ell(\boldsymbol{\eta}, \hat{\boldsymbol{\theta}}) \Big|_{\boldsymbol{\eta}=\boldsymbol{\eta}^*} + \nabla^2 \ell(\boldsymbol{\eta}, \hat{\boldsymbol{\theta}}) \Big|_{\boldsymbol{\eta}=\hat{\boldsymbol{\eta}}} (\hat{\boldsymbol{\eta}} - \boldsymbol{\eta}^*),$$

其中 $\hat{\boldsymbol{\eta}} \in U(\boldsymbol{\theta}^*) \subset \Omega(\boldsymbol{\eta})$.

通过简单的代数运算, 重写上面等式为

$$\sqrt{N}(\hat{\boldsymbol{\eta}} - \boldsymbol{\eta}^*) = \left(N^{-1} \nabla^2 \ell(\hat{\boldsymbol{\eta}}, \hat{\boldsymbol{\theta}}) \right)^{-1} (1/\sqrt{N}) \nabla \ell(\boldsymbol{\eta}^*, \hat{\boldsymbol{\theta}}).$$

为了证明上面等式依分布收敛, 需要分开处理这个问题. 首先, 有

$$\frac{1}{\sqrt{N}} \nabla \ell(\boldsymbol{\eta}^*, \hat{\boldsymbol{\theta}}) = \frac{1}{\sqrt{N}} \left(\nabla \ell(\boldsymbol{\eta}^*, \boldsymbol{\theta}^*) + \nabla \ell(\boldsymbol{\eta}^*, \hat{\boldsymbol{\theta}}) - \nabla \ell(\boldsymbol{\eta}^*, \boldsymbol{\theta}^*) \right) \xrightarrow{P} \frac{1}{\sqrt{N}} \nabla \ell(\boldsymbol{\eta}^*, \boldsymbol{\theta}^*).$$

最后依概率收敛是根据定理 1 结论. 根据极大似然定义及中心极限定理, 上面极限是服从零均值方差为 $I(\boldsymbol{\eta}^*, \boldsymbol{\theta}^*)$ 的正态分布.

现在给出 Hesse 矩阵的估计, 即

$$\frac{1}{N} \nabla^2 \ell(\hat{\boldsymbol{\eta}}, \hat{\boldsymbol{\theta}}) = \frac{1}{N} \left((\nabla^2 \ell(\hat{\boldsymbol{\eta}}, \hat{\boldsymbol{\theta}}) - \nabla^2 \ell(\boldsymbol{\eta}^*, \hat{\boldsymbol{\theta}})) + (\nabla^2 \ell(\boldsymbol{\eta}^*, \hat{\boldsymbol{\theta}}) - \nabla^2 \ell(\boldsymbol{\eta}^*, \boldsymbol{\theta}^*)) + \nabla^2 \ell(\boldsymbol{\eta}^*, \boldsymbol{\theta}^*) \right).$$

显然根据定理 1 及定理 2, 上面等式依概率收敛到 $I(\boldsymbol{\eta}^*, \boldsymbol{\theta}^*)$.

证毕.

4 实证分析

由于我国债券数据的不完整性, 因此将使用美国债券数据. 根据 Durham^[31] 的结果, 选取每周交易 3 个月到期日的债券收益率近似无风险短期利率, 其时间从 2000-01-07—2011-12-30, 总的数据为 629(数据来源于 <http://www.ustreas.gov>). Balduzzi 等^[15] 论述了使用不同的期限债券近似波动率是没有显著的差别, 因此本文将使用到期日为 1 年和 2 年的债券数据, 时间步长为 1/52. 图 1 描述了市场数据. 江良^[18] 使用同样的数据来比较 Vasicek 模型和 Hull-White 模型的似然函数值, 而且也得出 Hull-White 模型的似然函数值稍微地改善的结果, 但为了比较模型的有效性, 这些结果也将列出来.

基于 Härdle 等^[20] 的结果, 通过拇指原则, 对于高斯核函数, 最优宽度 $h = \bar{\sigma} (4/(3N))^{1/5}$, 其中 $\bar{\sigma}$ 是分布函数的标准方差, 而且 $\bar{\sigma}$ 可近似为

$$\bar{\sigma} = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (Y_i - \bar{Y})^2}, \quad (20)$$

其中 $\bar{Y} = \frac{1}{N} \sum_{i=1}^N Y_i$.

为了区分不同 h 值所对应估计值, 设 h_1, h_2 分别表示通过式(20)及式(21)近似获得, 而且也设 h_3 为 $\bar{\sigma} = 1$ 近似.

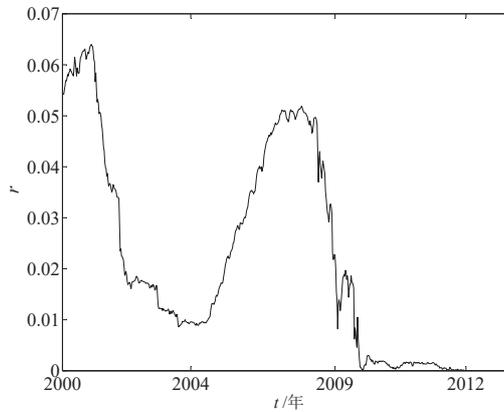


图 1 2000-01-07—2011-12-30 市场数据.

Fig. 1 Market data from 2000-01-07 to 2011-12-30.

表 1 列出了不同模型的参数估计值. 对 Vasicek 模型和 FV 模型, 长期均值分为 $\theta/a = 0.005 1$ 和 $\theta/a = 0.022 8$. 当考虑均值函数时, 基于 HW 模型, 显然窗口 $h = h_1$ 时, 似然值最大; 对于 SVHW 模型, 选取 $h = h_2$ 时, 似然函数值最大. 对比表 1 中的似然函数值, Vasicek 模型拟合效果最差, SVHW ($h = h_2$) 拟合效果最好. 而且观察表 1 中的数据, 基于 FV 模型的似然函数值明显比单因子模型的似然函数值要大. 这一结果表明了两因子模型比单因子模型能更好地刻画市场数据; 另一方面, 对于单因子模型考虑随机波动率比考虑均值函数影响要大, 如 FV 模型与 Vasicek 模型的似然值比较, 其似然比值为 1.378 7; HW 模型与 Vasicek 模型的似然比值为 1.140 8. 而对于 FV 模型, 考虑均值函数是必要的, 如: 当 $h = h_2$, SVHW 模型的似然函数值比 FV 模型的似然函数值要大, 大约为 1 700.

表 1 不同模型的参数估

Table 1 Parameters estimates for the different model

Model	θ	a	σ_r	α_0	α_1	l
Vasicek	0.001 3	0.255 3	0.011 0	—	—	2 506.5
FV	0.0021	0.093 4	—	0.006 4	0.290 3	3 455.7
HW ($h_1 = 0.016 4$)	—	0.271 1	0.006 3	—	—	2 859.4
HW ($h_2 = 0.079 7$)	—	0.280 6	0.009 8	—	—	2 756.2
HW ($h_3 = 0.292 5$)	—	380 6	0.063 6	—	—	2 542.2
SVHW ($h_1 = 0.016 4$)	—	0.093 9	—	0.006 5	0.394 8	3 455.5
SVHW ($h_1 = 0.080 4$)	—	0.292 7	—	0.000 3	0.011 1	5 134.3
SVHW ($h_1 = 0.292 5$)	—	0.093 8	—	0.006 5	0.294 9	3 448.8

图 2 描述了 HW 模型和 SVHW 对于不同窗口 h 选取所对应 $\theta(t)$ 的数值结果. 观察图 2 中的图形, 若长期均值过度地拟合, 导致数值结果出现激烈震荡. 这种现象很难给出参数的预测值和相应的衍生品定价, 而且业界也不喜欢不稳定的数值结果. 结合表 1 中的数据, 显然对于单因子模型, 若考虑数值结果的稳定性, 均值函数的模型对于拟合效果只是稍微地改善. 这就意味着对于实践者若使用单因子模型, 就应该权衡拟合的效果和数值稳定性. 进一步观察图 2 可知, 有一些数值解是负的. 对于高斯模型, 短期利率取负的概率是大于零, 因此负的 $\theta(t)$ 值是有可能的. 而且负的值表明了投资者对已实施政策的态度是消极的.

观察图 2, 考虑均值函数有力地解释了 2008 年发生次贷危机前后, 投资者对于政府和美联储救市行为的态度变化. 但是对于常系数模型, 这种现象是很难解释的. 如在 2008 年发生次贷危机时, 不管美联储不断下调利率, 但是投资者的长期预期利率还是下降. 而在 2008 年 10 月份, 通过美国政府救市, 投资者的长期预期利率值有明显的上升, 这种现象持续到 2009 年年初, 投资者对整个预期收益率又回落. 而后联邦储备

局通过购买国债救市,这就导致投资进一步对将来的预期收益率恢复了信心.因此,考虑均值函数不仅对数据拟合效果有了改善,而且为中央银行和政府对于实施政策的有效性提供了较好的参考依据.

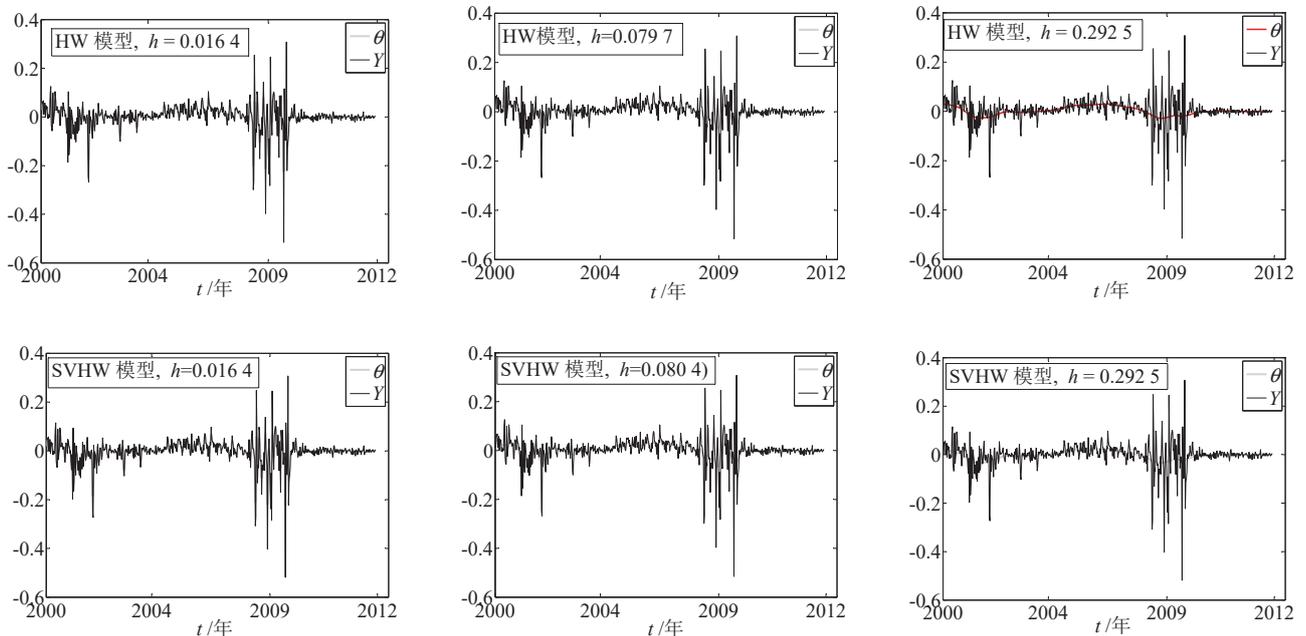


图2 基于 HW 模型和 SVHW 模型对于不同窗口 h 的 $\theta(t)$ 数值结果

Fig. 2 The numerical results $\theta(t)$ for the different bandwidth h based on HW model and SVHW model

观察表1中的数据 and 图2,当 $h = h_1$ 时,基于 SVHW 模型, $\theta(t)$ 函数的拟合效果最好,但是不稳定,其似然函数值明显比 $h = h_2$ 时似然函数值要小.这就说明了引入随机波动率和均值函数是有必要的.

图3显示了 SVHW 模型对不同窗口选取所对应的瞬时波动率.

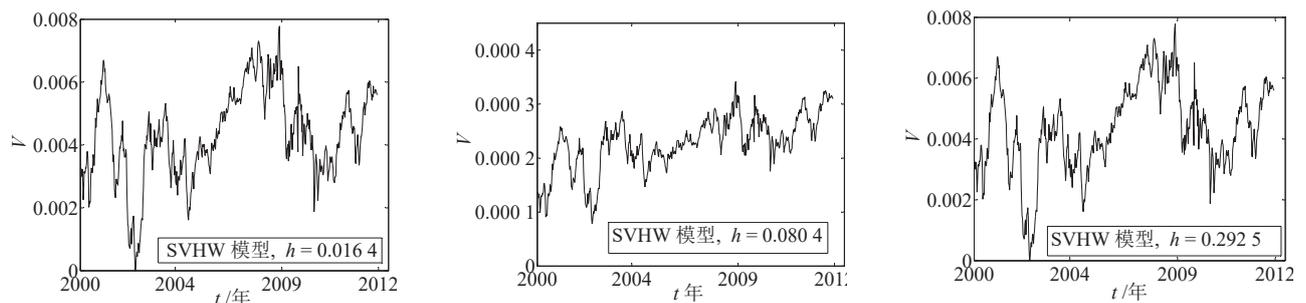


图3 基于 SVHW 模型对于不同窗口 h 所对应的瞬时波动率 (V)

Fig. 3 The spot volatilities (V) by the different bandwidth h based on SVHW Model,

图4画出了 FV 模型的瞬时波动率.观察图3和图4,当 $h = h_1, h_3$ 时,其瞬时波动率没有明显的差异,这种现象和表1中的似然函数值保持一致,因此当选取 $h = h_1, h_3$ 时,考虑长期均值是没有必要的.然而当 $h = h_2$ 时,其图像有着显著的差异,而且似然函数值也明显改善.比较 SVHW ($h = h_2$) 模型的瞬时波动率, FV 模型的瞬时波动率被高估了.

表2列出了水平短期利率和波动率的相关系数估计值.观察表2中数据,对于 SVHW 模型,当 $h = h_1, h_3$ 时,相关性系数估计值和 FV 模型所得的估计值没有显著的差异.对于 SVHW ($h = h_2$) 模型,相关性系数估计值最小,但所对应的似然值是最大的(看表1中的数据).此外,观察图1、图3和图4,瞬时利率和波动率运动的趋势是相同的.这种现象表明它们之间的相关系数一定为正的(参考表2中的数据).

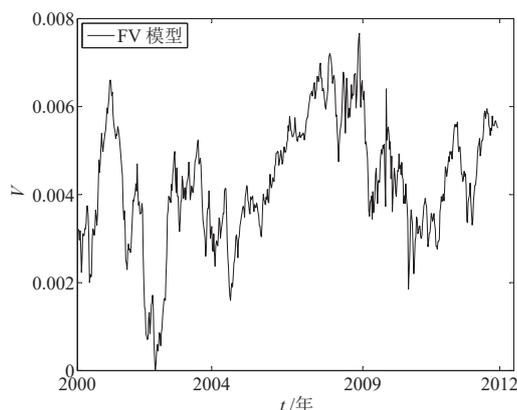


图 4 基于 FV 模型的短期波动率

Fig. 4 Spot volatility for FV Model

表 2 短期利率和短期波动率水平相关性估计值

Table 2 Correlation coefficient estimates between the short interest rate and the spot volatility

	FV model		SVHW Model		
	dV		dV(h ₁)	dV(h ₂)	dV(h ₃)
dr	0.239 4		0.239 2	0.147 0	0.239 2

5 结束语

众所周知, 不断地合理引入状态变量, 其相应的模型拟合效果将变得更好. 但是基于模型的衍生品定价的维数也会增加, 这就导致衍生品定价数值方法变得更加困难. 因此本文考虑随机波动率和均值函数的两因子模型. 为了表明模型的合理性, 本文选取时间序列数据, 采用核函数估计方法并对随机波动率通过一个辅助的过程近似进行参数估计. 进一步给出了估计方法的相容性和渐近性性质.

实证分析表明了, 考虑随机波动率和均值函数是必要的, 而且均值函数的假设对于系数校正效果有明显的改善. 但是对于单因子模型可能考虑均值函数需要根据实际应用的情况而定. 另一方面, 考虑均值函数能够很好地解释一些金融市场的行为, 特别是给政府和中央银行对已实施政策的有效性提供了很好的参考依据. 因此可以断言, SVHW 模型比单因子模型及 FV 模型更加有效. 对于业界, 可以在不增加状态变量的情况下, 使用更好的模型进行衍生品定价. 在进一步研究中, 将考虑基于债券价格数据(横截面数据) 两因子的参数估计方法并给出 SVHW 模型的衍生品定价.

参考文献:

- [1] 杨宝臣, 李 彪. 基于最优动态利率模型的认股权证定价研究. 系统工程学报, 2009, 24(3): 264–271.
Yang B C, Li B. Research on warrant pricing based on optimal dynamical interest rate model. Journal of Systems Engineering, 2009, 24(3): 264–271. (in Chinese)
- [2] Hull J, White A. Pricing interest-rate-derivative securities. The Review of Financial Studies, 1990, 3(4): 573–392.
- [3] Dai Q, Singleton K. Specification analysis of affine term structure models. The Journal of Finance, 2000, 55(5): 1943–1978.
- [4] Duffie D, Kan R. A yield-factor model of interest rates. Mathematical Finance, 1996, 6(4): 379–406.
- [5] Litterman R, Scheinkman J A. Volatility and yield curve. Journal of Fixed Income, 1991, 1(1): 49–53.
- [6] Kim D H, Singleton K J. Term structure models and the zero bound: An empirical investigation of Japanese yields. Journal of Economics, 2012, 170(1): 32–49.
- [7] Cheridito P, Filipovic D, Kimmel R L. Market price of risk specifications for affine models: Theory and evidence. Journal of Financial Economics, 2007, 83(1): 123–170.

- [8] Pierre C D, Goldsterin R S, Jones C S. Can interest rate volatility be extracted from the cross section of bond yields. *Journal of Financial Economics*, 2009, 94(1): 47–66.
- [9] Duffee G R, Stanton R H. Estimation of dynamic term structure models. *Quarterly Journal of Financ*, 2012, 2(2), 1–51.
- [10] Andersen T G, Lund J. Estimating continuous-time stochastic volatility methods of the short term interest rate. *Journal of Econometrics*, 1997, 77(2), 343–377.
- [11] Fong G H, Vasicek O A. Fixed-income volatility management. *Journal of Portfolio Management*, 1991, 7(4): 41–46.
- [12] Longstaff F A, Schwartz E. Interest rate volatility and the term structure: A two-factor general equilibrium model. *Journal of Finance*, 1992, 47(4): 1259–1282.
- [13] Surya B A. Two-dimensional Hull-White model for stochastic volatility and its nonlinear filtering estimation. Singapore: *Procedia Computer Science*, 2011, 4, 1431–1440.
- [14] Balduzzi P, Das S, Foresi S. The central tendency: A second factor in bond yield. *Review of Economics and Statistics*, 1998, 80(1): 62–72.
- [15] Balduzzi P, Das S R, Foresi S, et al. A simple approach to three-factor affine models of the term structure. *Journal of Fixed Income*, 1996, 6(3): 43–53.
- [16] Chen L. *Interest Rate Dynamics, Derivatives Pricing, and Risk Management*. Heidelberg: Springer Verlag, 1996.
- [17] 范龙振, 程南雁. 一类均值为跳跃-均值回复过程的利率模型. *系统工程学报*, 2011, 26(3): 298–305.
Fan L Z, Chen N Y. Interest rate model with jump-mean reverting stochastic mean. *Journal of Systems Engineering*, 26(3): 298–305. (in Chinese)
- [18] 江 良. 基于时间序列数据 Hull-White 模型半参数估计方法. *莆田学院学报*, 2013, 20(5): 1–4
Jiang Liang. Semi-parametric estimation for Hull-White model via time series data. *Journal of Putian University*, 2013, 20(5): 1–4. (in Chinese)
- [19] Grzelak L A, Oosterlee C W, Weeren S V. Extension of stochastic volatility equity models with the Hull - White interest rate process. *Quantitative Finance*, 2012, 12(1): 89–105.
- [20] Härdle W, Müller M, Sperlich S, et al. *Nonparametric and Semiparametric Models an Introduction*. Berlin: Springer, 2004.
- [21] Ramsay J O, Silverman B W. *Functional Data Anaysis*. New York: Springer Verlag, 2005.
- [22] Ait-Sahalia Y, Kimmel R. Maximum likelihood estimation of stochastic volatility models. *Journal of Financial Economics*, 2007, 83(2): 413–452.
- [23] Cox J C, Ingersoll J E, Ross S A. A theory of the term structure of interest rates. *Econometrica*, 1985, 53(2): 385–407.
- [24] Brigo D, Mercurio F. *Interest-rate Models: Theory and Practice*. Berlin: Springer Verlag, 2006.
- [25] Vasicek O A. An equilibrium characterization of the term structure. *Journal of Financial Economics*, 1977, 5(2): 177–188.
- [26] Glasserman P. *Monte Carlo methods in Financial Engineering*. New York: Springer Verlag, 2003.
- [27] Simar L, Wilson P W. Estimation and inference in two-stage, semi-parametric models of production process. *Journal of Econometrics*, 2007, 136(1): 31–64.
- [28] Härdle W. *Applied Nonparametric Regression*, *Econometric Society Monographs*. Cambridge: Cambridge University Press, 1990.
- [29] Kullback S, Leibler R A. On information and sufficiency. *Annals of Mathematical Statistics*, 1951, 22 (1): 79 – 86.
- [30] Van Der Vaart A W. *Asymptotic Statistics*. Combridge: Combridge University Press, 1998.
- [31] Durham G B. Likelihood-based specification analysis of continuous-time models of the short-term interest rate. *Journal of Financial Economics*, 2003, 70(3): 463–487.

作者简介:

江 良 (1978—), 男, 福建莆田人, 博士, 副教授, 研究方向: 金融工程和金融计算, Email: ptjiang@163.com;

林鸿熙 (1969—), 男, 福建莆田人, 硕士, 教授, 研究方向: 演化博弈论及其应用, Email: linhongxi@163.com.