



基于神经网络的股票收益率预测研究

潘水洋, 刘俊玮, 王一鸣

(北京大学 经济学院, 北京 100871)

摘要: 在金融领域的资产定价模型修正过程中, 股市的非线性现象往往被选择性忽视, 未纳入模型框架, 现有模型亦无法刻画因子之间的非线性定价结构。为解决上述问题, 引入了机器学习领域中的神经网络模型, 以捕获市场组合收益率、市值、账面市值比三因子间的非线性定价结构, 并对股票收益率进行预测。将该模型与经典Fama-French三因子模型在样本外拟合优度、多空策略业绩表现上做了对比, 结果表明: 神经网络模型能精准捕获市场组合收益率、市值、账面市值比3个因子之间的非线性关系, 且在样本外拟合优度、多空策略业绩表现上均要优于传统三因子线性定价模型。

关键词: 机器学习; 非线性资产定价; 神经网络

中图分类号: O212

文献标志码: A

文章编号: 1008-9497(2019)05-550-06

PAN Shuiyang, LIU Junwei, WANG Yiming (School of Economics, Peking University, Beijing 100871, China)

Forecasting stock returns with artificial neural networks. Journal of Zhejiang University (Science Edition), 2019, 46 (5): 550-555

Abstract: In the process of correcting the asset pricing model, the stock market exhibits nonlinear phenomena. Nevertheless this widespread financial disparity is ignored selectively and not included in the model framework. The challenge is that the existing model cannot describe the nonlinear pricing structure among factors. In this paper, a neural network model in the field of machine learning is presented to capture the nonlinear pricing structure between the three factors of firm size, book-to-market ratio and return on the market portfolio. Based on the Chinese A-share market data, this model is fully compared with the classic Fama-French three-factor model. The empirical results show that the neural network model is better than the Fama-French three-factor linear pricing model in terms of goodness of fit outside the sample and performance of long-short strategy.

Key Words: machine learning; nonlinear asset pricing; neural network

0 引言

从计量经济学角度在对传统线性资产定价模型进行实证分析时,发现该模型的解释能力不足:或者回归截距项显著大于零,或者贝塔系数不显著等。造成这些问题的根源可能是解释变量遗漏,也可能是贝塔系数或随机项方差的时变性等。因此,针对

线性资产定价模型的修正出现了3个不同的努力方向:将失败归因于解释变量的遗漏,据此提出了多因子资产定价模型^[1-3];将失败归因于贝塔系数的时变性,据此提出了变参数(贝塔)资产定价模型^[4];将失败归因于随机项方差的时变性,据此提出了GARCH类资产定价模型等^[5]。然而上述3模型并没有从根本上解决资产定价模型解释能力不足的问题。首先,单纯依靠增加解释变量提高模型解释能

收稿日期: 2018-12-26.

基金项目: 国家社会科学基金青年项目(18CJY057).

作者简介: 潘水洋(1986—), ORCID: <http://orcid.org/0000-0003-3512-6503>, 男, 博士后, 主要从事实证资产定价与金融大数据分析研究. E-mail: panshuiyang@126.com.

力是一个无穷尽的过程。其次,由时变贝塔参数定价模型估计出的贝塔系数和市场组合的超额收益之间的协方差并不大,这意味着时变贝塔参数定价模型得到的收益和用CAPM模型计算出来的收益相仿^[6]。

与此同时,国内外学者发现金融市场存在非线性特征,大量文献对此进行了实证检验和报道^[7-10]。股票市场具有非线性特征意味着随机游走与有效市场假设将失效,以此为理论假设基础的现代资本市场理论以及其他依赖正态分布或有限方差性质的金融计量学模型都将面临质疑。令人遗憾的是,在资产定价模型的修正过程中,股市具有非线性现象这一金融异象被选择性地忽视了,未被纳入模型框架,其挑战在于现有模型无法刻画因子间的非线性定价结构。为解决这一问题,笔者创新性地引入机器学习领域中的神经网络模型,以捕获市场组合收益率、市值和账面市值比三因子之间的非线性定价结构。对本模型的样本外拟合优度、多空策略业绩表现与经典Fama-French三因子模型做了比较。实证结果表明:神经网络模型能精准捕获市场组合收益率、市值和账面市值比3个因子之间的非线性关系。神经网络非线性模型在样本外拟合优度、多空策略业绩上均要优于传统三因子线性定价模型。本文的贡献主要体现在:(1)脱离了依靠增加解释变量个数提高线性资产定价模型解释力的传统实证资产定价的研究框架。(2)将机器学习领域中的神经网络模型运用于金融资产定价领域,实现定价因子之间的非线性关系捕获,丰富了资产定价的研究方法。(3)采用机器学习理论的非线性资产定价模型能够更加准确地预测我国股票投资组合的收益率,对投资者构建有效投资策略具有较强的借鉴意义。

1 神经网络非线性定价模型

与传统金融计量经济学方法相比,机器学习理论中的神经网络模型更关注预测结果,不关注估计量的渐进性质,因此将机器学习模型应用于资产定价领域,须针对研究的问题构造合适的实证框架^[11-13]。

学术界提出的随机贴现因子资产定价模型的一般表达式为

$$E_t(M_{t+1}R_{t+1})=1, \quad (1)$$

式(1)中, M_{t+1} 表示随机贴现因子,可表示为风险因子的函数。随机贴现因子与风险因子之间的函数关系可能为线性关系,也可能为非线性关系。当随机贴

现因子是风险因子的线性组合时,资产的收益可用线性因子模型(即贝塔模式)来刻画。当随机贴现因子不是风险因子的线性组合时,以上结果并不成立。从本文引言部分可知,大量实证文献均发现股票市场存在非线性特征,为了更好地刻画资产定价模型,实现理论与实践的完美结合,随机贴现因子应当满足: M_{t+1} 不仅能识别线性定价结构,而且也能识别非线性定价结构。本文通过引入机器学习中的神经网络模型来实现 M_{t+1} 的非线性映射。

1.1 神经网络模型

机器学习中的神经网络结构由输入层、隐含层、输出层组成。假设有 N 个任意样本集 $(x_j, y_j), j=1, 2, \dots, N$ 。其中,

$$x_j = [x_{j1}, x_{j2}, \dots, x_{jn}]^T \in \mathbb{R}^n, \quad (2)$$

$$y_j = [y_{j1}, y_{j2}, \dots, y_{jm}]^T \in \mathbb{R}^m, \quad (3)$$

通过神经网络可以自动寻找 x 与 y 之间的非线性映射关系。对于图1中具有 L 个隐含结点的单隐层神经网络的输出 t_j 可以表示为

$$\sum_{i=1}^L \theta_i g(w_i^T x_j + b_i) = t_j, \quad j=1, 2, \dots, N, \quad (4)$$

式(4)中, g 为激活函数,一般为 σ 函数。 $w_i = [w_{i1}, w_{i2}, \dots, w_{in}]^T$ 为第 i 个隐层单元的输入权重, b_i 为第 i 个隐层单元的偏置, $\theta_i = [\theta_{i1}, \theta_{i2}, \dots, \theta_{im}]^T$ 为第 i 个隐层单元的输出权重。 $t_j = [t_{j1}, t_{j2}, \dots, t_{jm}]^T$ 为神经网络预测输出值。上式的矩阵形式为

$$HB = H(w_1, w_2, \dots, w_L, b_1, b_2, \dots, b_L, x_1, x_2, \dots, x_N) B = T, \quad (5)$$

其中,

$$H = \begin{bmatrix} g(w_1^T x_1 + b_1) & \cdots & g(w_L^T x_1 + b_L) \\ \vdots & & \vdots \\ g(w_1^T x_N + b_1) & \cdots & g(w_L^T x_N + b_L) \end{bmatrix}_{N \times L}, \quad (6)$$

$$B = \begin{bmatrix} \theta_1^T \\ \vdots \\ \theta_L^T \end{bmatrix}_{L \times m}, \quad (7)$$

$$T = \begin{bmatrix} t_1^T \\ \vdots \\ t_L^T \end{bmatrix}_{N \times m}. \quad (8)$$

单隐层神经网络对输入的 N 个样本集 $(x_j, y_j), j=1, 2, \dots, N$ 进行自我学习,通过最小化输出值与实际值的误差平方和估计参数 $\widehat{w}_i, \widehat{b}_i, \widehat{\theta}_i$,得到 x 与 y 之间的非线性映射关系:

$$\min_{w, b, \theta} \|Y - T\| = \min_{w, b, \theta} \|Y - H(w_i, b_i) B\|, \quad (9)$$

式(9)等价于最小化损失函数:

$$E = \sum_{j=1}^N \left(\sum_{i=1}^L y_j - \theta_i g(\mathbf{w}_i^\top \mathbf{x}_j + b_i) \right)^2. \quad (10)$$

式(10)无解析解,采用梯度下降法可得到模型参数估计。

1.2 神经网络非线性定价结构辨识

将神经网络模型应用于资产定价的一个重要原因是该模型能刻画股票市场的非线性现象。假设市场存在 M 只股票,股票的定价因子有 N 个,在 t 时刻,拥有股票因子数据集 \mathbf{X}_t 及股票收益率 \mathbf{y}_t :

$$\mathbf{X}_t = \begin{bmatrix} x_{11}^t & \cdots & x_{1N}^t \\ \vdots & & \vdots \\ x_{M1}^t & \cdots & x_{MN}^t \end{bmatrix},$$

$$\mathbf{y}_t = [y_1^t \quad \cdots \quad y_M^t]^\top.$$

资产定价的实质就是找到因子数据集 \mathbf{X}_t 与股票收益率 \mathbf{y}_t 之间的映射关系 F :

$$\mathbf{y}_t = F(\mathbf{X}_t). \quad (11)$$

线性资产定价模型:假设因子数据集 \mathbf{X}_t 与股票收益率 \mathbf{y}_t 之间存在线性关系,此时式(10)变为

$$\mathbf{y}_t = \mathbf{X}_t \boldsymbol{\beta}. \quad (12)$$

神经网络非线性资产定价模型:假设因子数据集 \mathbf{X}_t 与股票收益率 \mathbf{y}_t 之间存在非线性关系:

$$\mathbf{y}_t = F(\mathbf{X}_t, \mathbf{X}_{t-1}, \mathbf{X}_{t-2}, \cdots | \mathbf{W}, \mathbf{b}, \boldsymbol{\theta}). \quad (13)$$

式(13)中, $\mathbf{W}, \mathbf{b}, \boldsymbol{\theta}$ 为神经网络模型中的待估计参数, F 为式(5)所描述的神经网络非线性映射。对比式(12)和(13)发现,与传统线性资产定价模型相比,神经网络资产定价模型能捕获到因子数据集 \mathbf{X}_t 与股票收益率 \mathbf{y}_t 之间的非线性映射关系,非线性映射通过以下激活函数实现:

$$f(x) = \frac{1}{1 + e^{-x}}. \quad (14)$$

只要因子数据集 \mathbf{X}_t 与股票收益率 \mathbf{y}_t 之间存在非线性定价结构,含有多个隐含层的神经网络结构从理论上就能够辨识隐藏在数据背后的非线性定价结构。下文将同时运用 Fama 三因子线性定价模型和神经网络非线性资产定价模型对我国 A 股市场进行实证分析,通过比较两模型的样本外拟合优度和多空策略业绩来检验神经网络定价模型的有效性。

2 实证分析

选取 Fama-French 三因子资产定价模型作为基准模型,通过对比分析检验神经网络非线性定价模型的有效性。该模型在国内外股票市场有大量的实证检验^[1-3,14-15]。

Fama-French 三因子资产定价模型中,股票收益率可由市场组合收益率、市值、账面市值比进行解释。根据此理论,构建三因子线性定价预测模型:

$$r_{it} = \beta_0 + \beta_1 R_{m,t-1} + \beta_2 \text{Size}_{i,t-1} + \beta_3 \text{Btop}_{i,t-1} + e_{it}, \quad (15)$$

式(15)中, r_{it} 表示股票 i 在时间 t 时的收益率; $R_{m,t-1}$ 表示在时间 $t-1$ 时按市值加权形成的市场组合收益率; $\text{Size}_{i,t-1}$ 表示股票 i 在时间 $t-1$ 时的流通市值; $\text{Btop}_{i,t-1}$ 表示股票 i 在时间 $t-1$ 时的账面市值比; e_{it} 表示三因子定价线性预测模型预测误差; $\beta_0, \beta_1, \beta_2, \beta_3$ 为线性定价模型参数。对市场组合收益率、市值、账面市值比等定价因子在每个横截面上均减去均值并除以标准差,进行归一化处理。

为检验神经网络资产定价模型的有效性,采用样本外模型预测值的拟合优度 R^2 、Diebold-Mariano 检验以及构建多空策略业绩表来检验模型的有效性,基准模型为 Fama-French 三因子线性定价模型。评判模型优劣的标准有:

(1) 样本外拟合度 R^2 , R^2 越大,模型越优。

(2) Diebold-Mariano 检验。

(3) 多空投资策略。

在模型输入变量保持不变的前提下,如果以上 3 条均较优,就有足够的理由认定神经网络资产定价模型能够识别隐藏在定价因子数据下的非线性关系,非线性定价模型优于线性定价模型。

2.1 模型样本外拟合度比较

选取 1997 年 1 月至 2017 年 12 月沪深两市 A 股股票的月度收益率、收盘价、合并财务报表中流通股、总股本、所有者权益合计数据,剔除 *ST、已退市、上市不足 1 年的股票。收盘价与流通股数量相乘得到股票流通市值;收盘价与总股本相乘得到股票总市值;所有者权益合计除以股票总市值得到账面市值比。市场组合收益率采用中证全 A 指数收益率。所有因子在横截面上减去均值除以标准差做归一化处理。本文的数据来源于 wind 资讯数据库。

将 1997 年 1 月至 2007 年 12 月的数据作为模型训练数据,用于估计 Fama-French 三因子线性定价模型参数和神经网络非线性资产定价模型参数。将 2008 年 1 月至 2017 年 12 月的数据作为样本外验证数据,用于检验模型样本外的预测能力。估计模型参数后,从 2008 年 1 月开始,在每月月初,将股票最新的市值因子数据、账面市值比因子数据、中证全 A 指数月度收益率数据分别输入 Fama-French 三因子线性定价预测模型、神经网络非线性资产定价预测

模型,得到每个模型下个月股票收益率的预测值,再计算模型样本外的拟合优度 R^2 ,拟合优度计算公式为

$$R^2 = 1 - \frac{\sum_{(i,t)} (r_{i,t} - \widehat{r_{i,t}})^2}{\sum_{(i,t)} r_{i,t}^2}, \tag{15}$$

式(15)中, $r_{i,t}$ 表示第*i*只股票在*t*时刻的实际收益率, $\widehat{r_{i,t}}$ 表示第*i*只股票在*t*时刻的预测收益率。表1为Fama-French模型和神经网络模型股票收益率样本的外拟合优度,其中NN1~NN5分别表示采用1~5层隐含层神经网络结构非线性定价模型,输入层均为3个节点(分别代表市场组合收益率、市值、账面市值比3个因子),输出层均为10个节点(用于股票收益率预测分组)。NN1隐含层含有32个节点,NN2隐含层分别含有32,16个节点,NN3隐含层分别含有32,16,8个节点,NN4隐含层分别含有32,16,8,4个节点,NN5隐含层分别含有32,16,8,4,2个节点。采用非线性激活函数式(14)实现。由表1知,不同隐含层数神经网络模型样本的外拟合优度均高于Fama-French模型样本。

表1 模型样本外拟合优度

Table1 Goodness of fit outside sample

Fama-French	NN1	NN2	NN3	NN4	NN5
0.21	0.31	0.32	0.37	0.35	0.29

2.2 Diebold-Mariano 检验

用2008年1月至2017年12月的数据作为样本外验证数据,采用预先训练好的模型预测股票下一期收益率,在此基础上,用Diebold和Mariano对不

同模型两两之间进行配对检验,比较不同模型之间的预测能力,结果如表2所示。从表2第2行可以看出,具有不同隐含层的神经网络非线性资产定价模型样本的外预测能力均要好于Fama-French三因子线性定价模型。隐含层数不同的神经网络非线性资产定价模型之间无显著性差异。

表2 Diebold-Mariano 配对检验

Table 2 Diebold-Mariano pairing test

	NN1	NN2	NN3	NN4	NN5
Fama-French	2.01**	2.25**	3.01***	2.53**	1.99**
NN1		0.83	1.11	0.43	0.51
NN2			0.1	0.21	-1.01
NN3				-0.43	-1.61
NN4					-0.32

注 **,**分别表示在5%,1%水平下显著。

2.3 多空策略业绩比较

为了进一步检验神经网络模型的有效性,本文基于Fama-French模型、神经网络模型构造多空策略组合:Fama-French因子线性组合预测策略和神经网络因子非线性组合预测策略。

Fama-French因子线性组合预测策略构造过程:每个月初,Fama-French模型基于最新因子截面数据预测股票下一期收益率,然后将股票收益率预测结果按降序排列等分为10组,做多第1组股票,做空第10组股票。

神经网络因子非线性组合预测策略构造过程:每个月初,神经网络模型基于最新因子截面数据预测股票下一期收益率,然后将股票收益率预测结果按降序排列等分为10组,做多第1组股票,做空第10组股票。

表3 分组收益

Table 3 Group's return

Fama-French					NN1				
组别	年化收益/%	年化波动率/%	夏普比率	最大回撤/%	组别	年化收益/%	年化波动率/%	夏普比率	最大回撤/%
第1组	14.98	23.92	0.59	46.80	第1组	19.86	25.34	0.86	46.77
第2组	13.46	27.36	0.56	50.56	第2组	13.93	24.01	0.70	46.76
第3组	11.44	23.39	0.51	45.96	第3组	11.69	24.23	0.67	49.91
第4组	9.47	23.18	0.66	43.51	第4组	8.17	24.07	0.51	50.15
第5组	10.27	23.53	0.44	48.05	第5组	6.62	24.17	0.46	50.67
第6组	6.44	22.96	0.39	47.10	第6组	5.06	24.45	0.34	50.33
第7组	3.64	23.48	0.27	51.37	第7组	2.26	24.17	0.25	52.00
第8组	-1.98	23.64	0.04	52.07	第8组	-0.49	23.91	0.16	51.77
第9组	2.34	23.19	0.22	49.21	第9组	-1.13	25.06	0.04	58.06
第10组	-1.86	23.34	-0.18	57.63	第10组	-3.61	25.56	-0.24	57.16

表 4 分组收益
Table4 Group's return

NN2					NN3				
组别	年化收益/%	年化波动率/%	夏普比率	最大回撤/%	组别	年化收益/%	年化波动率/%	夏普比率	最大回撤/%
第 1 组	19.98	24.88	0.86	46.66	第 1 组	22.70	24.60	0.96	44.64
第 2 组	14.71	23.88	0.72	46.02	第 2 组	16.65	24.00	0.77	44.56
第 3 组	12.21	24.22	0.69	49.29	第 3 组	14.04	23.60	0.68	46.50
第 4 组	9.45	23.61	0.52	49.68	第 4 组	11.05	23.37	0.57	46.04
第 5 组	7.92	23.84	0.46	50.15	第 5 组	9.31	23.26	0.50	47.84
第 6 组	5.22	24.41	0.34	50.30	第 6 组	7.22	23.58	0.42	49.83
第 7 组	3.01	23.83	0.26	50.81	第 7 组	4.72	23.51	0.32	51.01
第 8 组	0.31	23.42	0.18	50.46	第 8 组	2.44	23.65	0.22	51.50
第 9 组	-0.92	25.06	-0.05	56.63	第 9 组	1.03	24.19	0.08	53.60
第 10 组	-4.32	25.42	-0.22	57.11	第 10 组	-7.73	25.23	-0.19	56.25

表 5 分组收益
Table 5 Group's return

NN4					NN5				
组别	年化收益/%	年化波动率/%	夏普比率	最大回撤/%	组别	年化收益/%	年化波动率/%	夏普比率	最大回撤/%
第 1 组	21.60	27.57	0.92	46.92	第 1 组	17.36	28.06	0.75	50.89
第 2 组	13.68	24.84	0.75	44.94	第 2 组	12.42	25.08	0.64	48.53
第 3 组	13.46	26.06	0.66	49.15	第 3 组	10.54	26.87	0.61	53.55
第 4 组	8.95	23.80	0.55	49.73	第 4 组	6.06	27.14	0.48	53.00
第 5 组	6.95	24.69	0.46	48.70	第 5 组	4.74	24.57	0.38	52.43
第 6 组	4.27	24.62	0.41	54.08	第 6 组	2.01	26.33	0.26	51.95
第 7 组	2.28	24.07	0.29	55.12	第 7 组	0.71	24.65	0.23	53.93
第 8 组	2.38	24.90	0.18	53.65	第 8 组	-1.17	25.83	0.11	52.95
第 9 组	-2.02	25.24	-0.06	54.34	第 9 组	-2.95	25.72	-0.03	58.90
第 10 组	-5.93	25.87	-0.21	58.60	第 10 组	-3.28	27.85	-0.31	59.97

表 3~表 5 为各个模型策略的分组表现。从策略分组表现看,由于 Fama-French 分组策略未能捕获到因子之间的非线性定价结构,导致第 5 组收益率高于第 4 组,第 9 组收益率高于第 8 组,组间收益率并未呈现递减关系。与此相对应的是,神经网络分组策略能够捕获到因子间的非线性定价结构,其预测准确率要好于 Fama-French 分组策略:从第 1 组到第 10 组,收益率表现出严格的递减规律,且第 1 组(多头)收益率均高于 Fama-French 三因子模型第 1 组,第 10 组(空头)收益率均低于 Fama-French 三因子模型第 10 组。

表 6 为各模型做多第 1 组、做空第 10 组的多空策略表现。从表 6 中可以看出,神经网络模型多空策略年化收益、年化波动率、夏普比率、最大回撤均要好于 Fama-French 模型。进一步表明,神经网络模型捕获到了因子之间的非线性关系,策略表现要优于 Fama-French 模型。

3 结 论

针对股市具有非线性特征,在资产定价模型的修正过程中易被选择性忽视,创新性地将机器学习领域中的神经网络模型引入其中,捕获市场组合收益率、市值和账面市值比三因子间的非线性定价结构。新模型与经典 Fama-French 三因子模型在样本外拟合优度、多空策略业绩表现上进行了对比。实证结果表明,神经网络模型能精准捕获市场组合收益率、市值、账面市值比 3 个因子之间的非线性关系。神经网络非线性模型在样本外拟合优度、多空策略业绩上均要好于传统三因子线性定价模型。

3.1 中国 A 股市场存在典型的非线性特征,借助机器学习理论,能够精准捕获定价因子之间的非线性关系。

3.2 机器学习是以数据驱动为核心来构建模型的,并

依靠交叉验证选择最优模型,与传统计量经济学方法相比,机器学习更关注预测结果,不关注估计量的渐进性质,因此,将机器学习模型与传统计量模型进行对比时,须针对研究的问题构造合适的实证框架,本文做了初步探索。

表6 分组多空收益
Table 6 Group long and short returns

策略	多空组合年化收益/%	多空组合收益年化波动率/%	夏普比率	最大回撤/%
Fama-French	16.42	12.19	1.84	7.95
NN1	21.20	9.29	2.84	7.71
NN2	22.51	10.52	2.91	6.92
NN3	25.13	9.93	3.04	7.01
NN4	24.07	10.23	2.94	7.23
NN5	19.17	11.10	2.23	7.65

参考文献(References):

[1] FAMA E F, FRENCH K R. The cross-section of expected stock returns[J]. **The Journal of Finance**, 1992, 47(2):427-465.

[2] FAMA E F, FRENCH K R. A five-factor asset pricing model[J]. **Journal of Financial Economics**, 2015, 116(1):1-22.

[3] ZHANG L. The investment CAPM[J]. **European Financial Management**, 2017, 23(4): 545-603. DOI: 10.1111/eufm.12129

[4] ADRIAN T, FRANZONI F. Learning about Beta: Time-varying factor loadings, expected returns, and the conditional CAPM[J]. **Journal of Empirical Finance**, 2009, 16(4): 537-556. DOI: 10.2139/ssrn.391562

[5] HARVEY C R, LIU Y, ZHU H Q. ...and the cross-section of expected returns[J]. **The Review of Financial Studies**, 2016, 29(1): 5-68. DOI: 10.2139/ssrn.2249314

[6] LEWELLEN J, NAGEL S. The conditional CAPM does not explain asset-pricing anomalies[J]. **Journal of Financial Economics**, 2006, 82(2): 289-314. DOI: 10.2139/ssrn.441341

[7] DAVIDSON J, TERÄSVIRTA T. Long memory and nonlinear time series[J]. **Journal of Econometrics**, 2002, 110(2):105-112.

[8] 郝清民. 中国股市收益率长记忆性R/S非线性分析[J]. **管理工程学报**, 2007, 21(2): 115-117. DOI: 10.3969/j.issn.1004-6062.2007.02.024

HAO Q M. R/S nonlinear analysis on the long memory of Chinese stock markets return[J]. **Journal of Industrial Engineering Management**, 2007, 21(2): 115-117.

[9] 李红权,汪寿阳,马超群. 股价波动的本质特征是什么?——基于非线性动力学分析视角的研究[J]. **中国管理科学**, 2008, 16(5):1-8.

LI H Q, WANG S Y, MA C Q. What' s the nature of volatility in stock prices? -Based on the nonlinear dynamical analysis principle[J]. **Chinese Journal of Management Science**, 2008, 16(5):1-8.

[10] PANAS E. Estimating fractal dimension using stable distributions and exploring long memory through ARFIMA models in Athens stock exchange[J]. **Applied Financial Economics**, 2001, 11(4): 395-402. DOI:10.1080/096031001300313956

[11] ATHEY S. Beyond prediction: Using big data for policy problems[J]. **Science**, 2017, 355(6324):483-485. DOI:10.1126/science.aal4321

[12] VARIAN R. Big data: New tricks for econometrics[J]. **Journal of Economic Perspectives**, 2014, 28(2): 3-28. DOI:10.1257/jep.28.2.3

[13] VARIAN H R. Causal inference in economics and marketing[J]. **Proceedings of the National Academy of Sciences**, 2016, 113(27):7310-7315. DOI:10.1073/pnas.1510479113

[14] HOU K W, XUE C, ZHANG L. Digesting anomalies: An investment approach[J]. **The Review of Financial Studies**, 2015, 28(3): 650-705. DOI: 10.2139/ssrn.2152674

[15] 赵胜民,闫红蕾,张凯. Fama-French五因子模型比三因子模型更胜一筹吗? 来自中国A股市场的经验证据[J]. **南开经济研究**, 2016, 15(2):41-59.

ZHAO S M, YAN H L, ZHANG K. Does Fama-French five factor model outperform three factor model? Evidence from China's A-share market[J]. **Nankai Economic Studies**, 2016, 15(2):41-59.