

数据拟合中光滑参数的优化

王丽 王文剑 姜高霞

(山西大学计算机与信息技术学院 太原 030006)

摘要 数据的函数化是函数数据分析(Functional Data Analysis, FDA)的基础,也是区别于其它分析方法的关键步骤。数据拟合作为数据函数化的主要方法,通常可转化为损失函数和正则项的优化问题,其中的光滑参数就起着权衡损失和过拟合风险的作用。在光滑参数的选择方法中,广义交叉验证(Generalized Cross-Validation, GCV)是一种通用而且较好的参数选择方法,然而 GCV 是对离散值进行计算,欲得到较准确的光滑参数仍需做大量的计算。针对此问题,提出拟合优化和差分两种求解策略以提高最优光滑参数的求解效率,并在算法精度及效率方面进行了比较分析。在模拟数据和真实数据上的实验结果表明:两种策略与常用的网格法相比,求解效率有较大提高,且算法精度方面几乎相同,此外差分求解策略在精度方面略优于拟合优化求解策略,而拟合优化求解策略的效率更高。

关键词 光滑参数, 广义交叉验证, 差分求解策略, 拟合优化求解策略

中图分类号 TP311 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2015.9.043

Optimization for Smoothing Parameter in Process of Data Fitting

WANG Li WANG Wen-jian JIANG Gao-xia

(School of Computer and Information Technology, Shanxi University, Taiyuan 030006, China)

Abstract Data functionalizing is the basis of functional data analysis (FDA) and important step differed from other analysis methods. As the main approach of data functionalizing, data fitting usually can be converted into an optimization problem including loss function and the regularization term, and smoothing parameter plays a compromising role in weighing loss and the risk of over fitting. Generalized cross-validation (GCV) is a general and better parameter selection way, but massive calculation may be needed in order to get a more accurate smoothing parameter because GCV is calculated on discrete values. Aiming at this problem, the fitting optimization and the finite difference solution strategies were proposed to improve the solution efficiency of selection of the optimal smoothing parameter, and their precision and efficiency were compared and analyzed. The experiment results on simulated and real data sets demonstrate that the two proposed strategies are greatly improved in efficiency compared with the conventional grid method with almost the same precision. The finite difference solution strategy is better than the fitting optimization solution strategy in terms of algorithm precision, and the latter is more efficient.

Keywords Smoothing parameter, Generalized cross-validation, Finite difference solution strategy, Fitting optimization solution strategy

1 引言

函数型数据(Functional Data, FD)的概念起初由加拿大统计学家 Ramsay 以及 Dalian 于 1982 年提出^[1-3]。函数型数据是指在某个连续集上的一组测量,这个连续集通常指时间,也可以是心理空间、物理空间等^[4]。在实际应用中,函数型数据常见于生物、心理学、气象学、经济学等众多领域^[5-8],如不同行业的经济数据,不同地域的温度、降水量数据,能源(如煤炭、石油、汽油等)的消耗量,气象、物理、生物、心理研究领域的各种实验数据等。当这些离散观测数据的连续集取得非常

密集时,数据的函数特征往往会很明显,但函数特征是数据的内在结构,并不是外在表现形式。为了动态地对这些数据进行分析预测,挖掘出数据背后更多的信息,连续地展示比较深层的动态规律性^[9],就需要将离散观测数据视为一个整体,转化为连续的函数,即进行数据的函数化处理。数据函数化是 FDA 的基础,也是区别于其它传统分析方法的关键步骤。

目前数据的函数化通常采用函数插值和数据拟合两种方法实现。当观测数据比较准确,或者不考虑观测误差的影响时,常采用函数插值方法,但由于通常的观测数据都存在误差

到稿日期:2014-07-20 返修日期:2014-11-29 本文受国家自然科学基金(61273291, 71031006),山西省回国留学人员科研资助项目(2012 008),中国民航信息技术科研基地开放基金(CAAC-ITRB-201305)资助。

王丽(1991—),女,硕士生,主要研究方向为机器学习、函数型数据分析, E-mail: 18434303025@163.com;王文剑(1968—),女,博士,教授,博士生导师,CCF 会员,主要研究方向为机器学习、计算智能、图像处理等, E-mail: wjwang@sxu.edu.cn(通信作者);姜高霞(1987—),男,博士生,主要研究方向为函数型数据分析及应用。

或者噪声,因此数据拟合是数据函数化的主要方法。通过数据拟合得到的近似函数并不要求近似函数在每个观测点必须与观测数据相同,而是使整体误差较小,能较好地体现观测数据。

为避免观测数据存在的误差或噪声带来的影响,同时使拟合曲线较为光滑,数据拟合常通过光滑基函数的线性组合来表示函数型数据。基函数法将离散的观测数据转变成函数^[10],对于周期函数常利用傅里叶基(Fourier Basis)进行拟合,非周期函数则利用样条基(Spline Basis)进行拟合,而在实际中最常用的样条基函数是B样条基函数^[11]。基函数个数 k 的取值决定了对数据的拟合程度; k 越大,对数据的拟合效果越好,但可能会拟合到噪声或者测量误差; k 值过小,则可能会损失拟合函数的某些重要特征。

当基函数确定时,扩张系数 C_k ^[12]的值就决定了近似函数。对 C_k 的估计方法有很多,最早发展起来的光滑处理方法是小二乘法,它也是最基本的方法^[13]。它通过最小化最小二乘模型确定系数 C_k ,然而此方法对 k 的取值并没有明确的选择方法。Ramsay和Silverman^[14]将正则化的思想引入其中,提出了粗糙惩罚光滑法。粗糙惩罚光滑法对数据粗糙度进行了正则化惩罚,而光滑参数(Smoothing parameters)则决定对数据粗糙度的惩罚程度。当利用粗糙惩罚法代替最小二乘法求取近似函数时,根据基函数个数 k 的取值决定光滑性就转变为根据光滑参数 λ 决定光滑性^[12]。图1显示了不同光滑参数 λ 对拟合函数的影响; λ 过小,对拟合函数的粗糙度惩罚较小,易产生过拟合现象;反之则对拟合函数的粗糙度惩罚较大,拟合易接近于直线,不能体现原始数据的变化特征^[15]。所以光滑参数的选取对数据拟合起着关键作用。

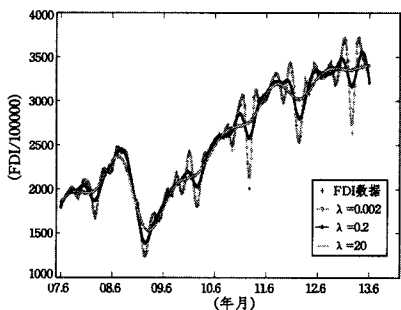


图1 FDI数据不同光滑参数示意图

Stone提出的交叉验证法(Cross-Validation, CV)可以求取光滑参数 λ ,其应用范围十分广泛,但这种方法会导致过拟合问题,而且计算复杂度较高,也不具有稳定性^[12]。为了避免上述问题,Craven和Wahba^[16]提出了广义交叉验证法(GCV)。通过最小化广义交叉验证得到的光滑参数比交叉验证法更可靠^[17]。该方法得到了很广泛的应用,然而它是对离散的光滑参数值进行计算,欲得到最优光滑参数仍需做大量的计算,并利用在算法精度方面较准确的常规网格法来获取,这是一个较为复杂的过程。针对此问题,本文提出了拟合优化和差分两种求解策略,避免了通过大量实验求取GCV所带来的计算复杂、计算量大等问题,并通过实例说明两种求解策略的效果。

2 光滑参数的求解策略

本文将求解最优光滑参数问题归结为一维无约束优化问

题。根据优化方法设计思想的不同,大致分为解析方法、启发式方法和数值迭代方法。解析方法一般针对特定的优化问题;启发式方法(如遗传算法)多用于大型、复杂的问题;数值迭代方法是实际中较为常用的策略。对于光滑参数的优化问题采用数值迭代方法。由于变化较为简单,可结合高效的牛顿迭代法求解。然而牛顿法需要一、二阶导数的信息,而离散化的GCV无法直接得到导数信息,故本文提出拟合优化求解策略和差分求解策略,同时与常用的网格法进行比较分析,避免了网格法^[18]计算量庞大、效率不高等问题。

2.1 拟合优化求解策略

拟合优化求解策略的基本思想为:对离散的 $GCV(\lambda)$ 再次数据拟合,即对 $GCV(\lambda)$ 通过光滑B样条基函数拟合以求取函数形式,在此基础上利用函数形式即可求取一、二阶导数信息,通过高效的牛顿迭代法求取最优光滑参数值。其主要步骤如下:

Step 1 对观测数据 $\{(x_i, y_i), i=1, 2, \dots, n\}$ (n 为样本量)通过以下处理方法得到数据 $\{(\lambda_i, GCV_i), i=1, 2, \dots, n\}$ 。

Step 1.1 对观测数据 $\{(x_i, y_i), i=1, 2, \dots, n\}$ 建立模型 $y_i = f(x_i)$,并利用光滑基函数的线性组合:

$$x(t_j) = \sum_{k=1}^K C_k \varphi_k(t_j) \quad (1)$$

来表示函数型数据,其中 $\varphi_k(t_j)$ ($k=0, 1, \dots, K$)为基函数, k 为基函数的个数, C_k 为扩张系数。

Step 1.2 通过构造B样条基粗糙惩罚模型 $PENSSR$, $\langle x|y \rangle = \sum_j \{y_j - x(t_j)\}^2 + \lambda \times PEN_2(x)$ 进行数据拟合,其中

$$PEN_2(x) = \int \{D^2 x(s)\}^2 ds = \|D^2 x\|^2 \quad (2)$$

式中, $D^m x(t)$ 表示 $x(t)$ 的 m 阶导函数。

Step 1.3 通过最小化广义交叉验证模型

$$GCV(\lambda) = \frac{\frac{1}{n} \sum_{i=1}^n [y_i - \hat{x}_\lambda(t_i)]^2}{[1 - \text{trace}(S_\lambda)/n]^2} \quad (3)$$

求取Step 1.2模型中的光滑参数 λ ,以得到数据 $\{(\lambda_i, GCV_i), i=1, 2, \dots, n\}$,其中 $S_\lambda = \Phi(\Phi' \Phi + \lambda R)^{-1} \Phi'$, Φ 表示全体基函数。

Step 2 对数据 $\{(\lambda_i, GCV_i), i=1, 2, \dots, n\}$ 通过光滑B样条基函数再次数据拟合,得到函数形式 $g(\lambda)$ 。

Step 3 求解函数 $g(\lambda)$ 的一阶导数 g' 和二阶导数 g'' 。

Step 4 通过牛顿迭代法求取最优光滑参数 λ 。其具体算法如下:

Step 4.1 初始化。确定起始点 λ_i 和误差阈值 ϵ 。

Step 4.2 迭代前的准备。假设第 k 步光滑参数为 λ_k ,计算 $g'(\lambda_k)$ 和 $g''(\lambda_k)$ 。

Step 4.3 进行迭代。 $\lambda_{k+1} = \lambda_k - \frac{g'(\lambda_k)}{g''(\lambda_k)}$ 。

Step 4.4 停止准则。重复Step 4.2和Step 4.3,直到 $|\lambda_{k+1} - \lambda_k| < \epsilon$ 。

由上述算法可以看出,每迭代1次需要计算1次 $g(\lambda)$ 的一阶导数 g' 和二阶导数 g'' 。

2.2 差分求解策略

差分求解策略的基本思想为:以差分思想代替求取函数形式从而获取导数信息,对离散的 $GCV(\lambda)$ 数据直接利用差分方法求取最优光滑参数。这种求解策略的主要步骤如下:

Step 1 与拟合优化求解策略的Step 1相同。

Step 2 对数据 $\{(\lambda_i, GCV_i), i=1, 2, \dots, n\}$ 直接通过差分方法求取。其具体算法如下:

Step 2.1 初始化。确定起始点 λ_i 和误差阈值 ϵ 。

Step 2.2 迭代前的准备。假设第 k 步光滑参数为 λ_k , 计算 $GCV'(\lambda_k)$ 和 $GCV''(\lambda_k)$, 其中

$$GCV_1'(\lambda_k) = \frac{GCV(\lambda_k + \delta) - GCV(\lambda_k)}{\delta} \quad (4)$$

$$GCV_2'(\lambda_k) = \frac{GCV(\lambda_k) - GCV(\lambda_k - \delta)}{\delta} \quad (5)$$

$$GCV''(\lambda_k) = \frac{GCV_1'(\lambda_k) - GCV_2'(\lambda_k)}{\delta} \quad (6)$$

Step 2.3 进行迭代。 $\lambda_{k+1} = \lambda_k - \frac{GCV'(\lambda_k)}{GCV''(\lambda_k)}$ 或者 $\lambda_{k+1} = \lambda_k - \frac{GCV_2'(\lambda_k)}{GCV''(\lambda_k)}$ 。

Step 2.4 停止准则。重复 Step 2.2 和 Step 2.3, 直到

$$|\lambda_{k+1} - \lambda_k| < \epsilon。$$

由上述算法可以看出, 每迭代 1 次需要计算 3 次 GCV, 分别是: $GCV(\lambda_k)$ 、 $GCV(\lambda_k + \delta)$ 、 $GCV(\lambda_k - \delta)$ 。

3 仿真实验

本文对模拟测试数据集以及真实测试数据集的实验均在 Matlab 环境下实现, 其中真实数据(居民消费价格指数(CPI)、邮电业务总量_当期值、进出口总值_当期值)均来源于国家统计局(<http://www.stats.gov.cn>)。采用本文方法选取最优光滑参数, 并与网格法在算法精度方面进行比较。所得结果如表 1 和表 2 所列。

表 1 模拟数据实验结果

模拟数据集	初始值设定	算法	迭代次数	GCV 计算次数或对 $g(\lambda)$ 的求导次数	运行时间 (s)	光滑参数	误差百分比 (%)
模拟数据集 1 $\sin(x)/x$	$x_{\min}+2$	网格法	81	81	7.83	15.3	
		差分法	2	9	0.83	15.3	0
		牛顿迭代法	2	6	0.06	15.3	0
		差分法	1	6	0.51	15.3	0
	$(x_{\min}+x_{\max})/2$	牛顿迭代法	1	4	0.03	15.3	0
		差分法	2	9	0.63	15.3	0
		牛顿迭代法	2	6	0.03	15.3	0
		差分法	2	9	0.63	15.3	0
模拟数据集 2 $1+2 \times x+3 \times x^2+4 \times x^3$	$x_{\min}+1$	网格法	39	39	1.43	8.0	
		差分法	2	9	0.15	8.0	0
		牛顿迭代法	1	4	0.02	8.1	1.25
		差分法	1	6	0.12	8.0	0
	$(x_{\min}+x_{\max})/2$	牛顿迭代法	0	2	0.02	8.1	1.25
		差分法	2	9	0.17	8.0	0
		牛顿迭代法	1	4	0.01	8.1	1.25
		差分法	2	9	0.17	8.0	0
模拟数据集 3 $\log(x)/x$	$x_{\min}+0.1$	网格法	32	32	1.14	0.60	
		差分法	0	3	0.07	0.60	0
		牛顿迭代法	0	2	0.01	0.61	1.67
		差分法	0	3	0.07	0.60	0
	$(x_{\min}+x_{\max})/2$	牛顿迭代法	0	2	0.02	0.61	1.67
		差分法	1	6	0.10	0.60	0
		牛顿迭代法	0	2	0.02	0.61	1.67
		差分法	1	6	0.10	0.60	0

表 2 真实数据实验结果

真实数据集	初始值设定	算法	迭代次数	GCV 计算次数或对 $g(\lambda)$ 的求导次数	运行时间 (s)	光滑参数	误差百分比 (%)
真实数据集 1 居民消费价格指数(CPI)	$x_{\min}+0.05$	网格法	31	31	1.72	0.60	
		差分法	1	6	0.24	0.60	0
		牛顿迭代法	1	4	0.01	0.60	0
		差分法	1	6	0.23	0.60	0
	$(x_{\min}+x_{\max})/2$	牛顿迭代法	0	2	0.01	0.60	0
		差分法	1	6	0.30	0.60	0
		牛顿迭代法	1	4	0.02	0.60	0
		差分法	1	6	0.30	0.60	0
真实数据集 2 邮电业务总量_当期值	$x_{\min}+0.1$	网格法	33	33	2.06	0.25	
		差分法	2	9	0.28	0.25	0
		牛顿迭代法	2	6	0.02	0.25	0
		差分法	0	3	0.12	0.25	0
	$(x_{\min}+x_{\max})/2$	牛顿迭代法	1	4	0.02	0.25	0
		差分法	2	9	0.27	0.25	0
		牛顿迭代法	2	6	0.03	0.25	0
		差分法	2	9	0.27	0.25	0
真实数据集 3 进出口总值_当期值	$x_{\min}+2$	网格法	53	53	3.13	9.3	
		差分法	1	6	0.17	9.3	0
		牛顿迭代法	1	4	0.03	9.3	0
		差分法	2	9	0.24	9.3	0
	$(x_{\min}+x_{\max})/2$	牛顿迭代法	2	6	0.02	9.3	0
		差分法	0	3	0.26	9.3	0
		牛顿迭代法	2	6	0.03	9.3	0
		差分法	2	9	0.26	9.3	0

差分法和牛顿迭代法对于不同的测试数据在不同的初始点到达最优值的迭代过程如图 2—图 7 所示。将上述两种方法求得的最优光滑参数代入 B 样条基粗糙惩罚模型,从而得到近似函数。由于近似函数曲线效果近似相同,因此通过表 1 和表 2 对两种方法的算法精度以及算法效率进行比较分析。

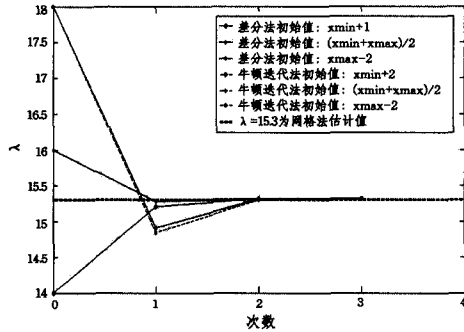


图 2 算法在 $(\sin(x)/x)$ 上的迭代过程

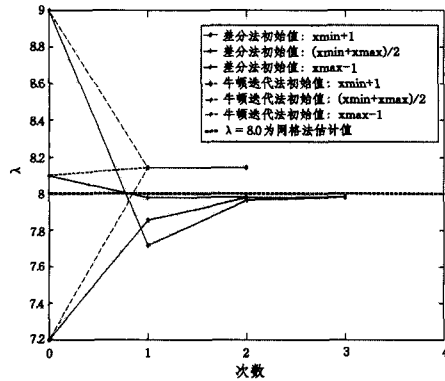


图 3 算法在 $(1+2x+3x^2+4x^3)$ 上的迭代过程

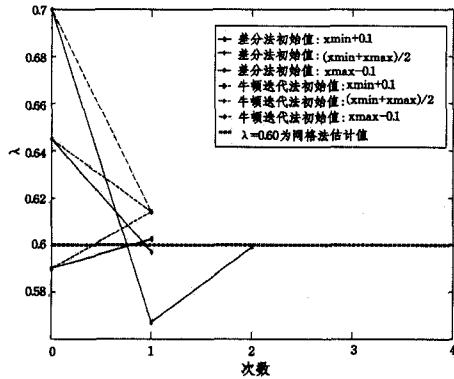


图 4 算法在 $(\log(x)/x)$ 上的迭代过程

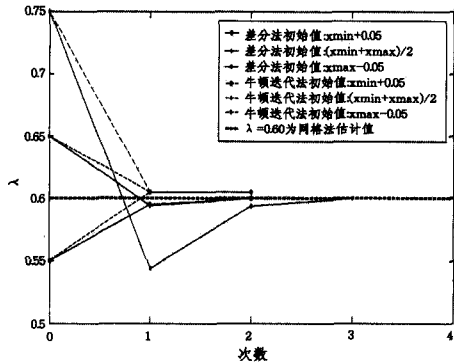


图 5 算法在 CPI 数据集上的迭代过程

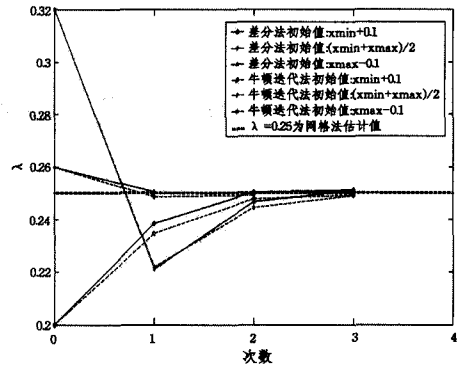


图 6 算法在邮电业务总量_当期值上的迭代过程

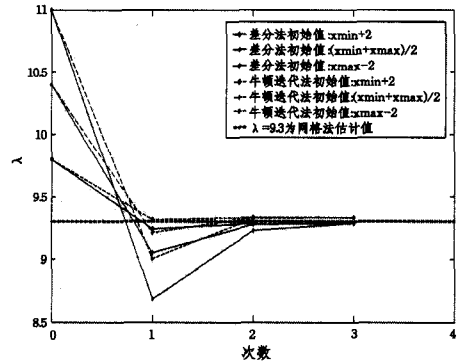


图 7 算法在进出口总值_当期值上的迭代过程

从算法精度方面考虑,由误差百分比可知,从整体上说,对于不同的测试数据集以及初始值,差分法均可达到最优值;而不同的测试数据集以及初始值对牛顿迭代法则有影响:对模拟数据集 1 以及真实数据集中不同的初始值能达到最优值,而对模拟数据集 2、3 中不同的初始值不能够都达到最优值,尽管误差百分比比较小。这表明牛顿迭代法对初始值较敏感,对初始值的设定有较高的要求。所以在算法精度方面,差分法的算法精度高于牛顿迭代法,差分求解策略略优于拟合优化求解策略。

从算法效率方面考虑,由迭代次数及运行时间可知:牛顿迭代法以及差分法的算法效率明显高于网格法,说明拟合优化和差分求解两种策略极大地提高了求解最优光滑参数的效率;同时牛顿迭代法明显高于差分法,这说明了差分法因对 GCV 进行计算而耗费了时间,而牛顿迭代法首先得到函数形式,极大地提高了算法效率。所以在算法效率方面,牛顿迭代法优于差分法,拟合优化求解策略优于差分求解策略。

结束语 本文提出了数据拟合中光滑参数的两种优化策略:拟合优化求解策略与差分求解策略,可以快速地求出最优光滑参数。与常用的网格法相比,两种策略求解效率有较大提高,且算法精度方面几乎相同,为下一步的函数型数据分析奠定了基础。

参考文献

- [1] 杨威. 函数型回归模型的成分选取[D]. 吉林: 东北师范大学, 2009
Yang Wei. Variable selection in functional regression model [D]. Jilin: Northeast Normal University, 2009

- factorization[C]//Proceedings of the 27th International Conference on Machine Learning(ICML-10). 2010;711-718
- [11] Koren Y. Collaborative filtering with temporal dynamics [J]. Communications of the ACM, 2010, 53(4): 89-97
- [12] Xiong L, Chen X, Huang T K, et al. Temporal Collaborative Filtering with Bayesian Probabilistic Tensor Factorization[C] // SDM. 2010; 211-222
- [13] Xiang L, Yuan Q, Zhao S, et al. Temporal recommendation on graphs via long-and short-term preference fusion[C]// Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM, 2010; 723-732
- [14] Rendle S, Freudenthaler C, Schmidt-Thieme L. Factorizing personalized markov chains for next-basket recommendation[C]// Proceedings of the 19th International Conference on World Wide Web. ACM, 2010; 811-820
- [15] Ma H, Yang H, Lyu M R, et al. Sorec: social recommendation using probabilistic matrix factorization[C]// Proceedings of the 17th ACM Conference on Information and Knowledge Management. ACM, 2008; 931-940
- [16] Yuan Q, Chen L, Zhao S. Factorization vs. regularization: fusing heterogeneous social relationships in top-n recommendation[C]// Proceedings of the fifth ACM Conference on Recommender Systems. ACM, 2011; 245-252
- [17] Ma H, Zhou D, Liu C, et al. Recommender systems with social regularization[C]// Proceedings of the fourth ACM International Conference on Web Search and Data Mining. ACM, 2011; 287-296
- [18] Jamali M, Ester M. A transitivity aware matrix factorization model for recommendation in social networks[C]// Proceedings of the Twenty-Second international joint conference on Artificial Intelligence. 2011; 2644-2649
- [19] Shen Y, Jin R. Learning personal+social latent factor model for social recommendation [C] // Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM, 2012; 1303-1311
- [20] Ye M, Liu X, Lee W C. Exploring social influence for recommendation: a generative model approach [C] // Proceedings of the 35th International ACM SIGIR Conference on Research and Development in Information Retrieval. ACM, 2012; 671-680
- [21] Salakhutdinov R, Mnih A. Probabilistic Matrix Factorization [OL]. [http://www. cs. torontov. edu/~amnih/papers/pmf. pdf](http://www.cs.torontov. edu/~amnih/papers/pmf. pdf)
- [22] 王越, 程昌正. 协同过滤算法在电影推荐中的应用[J]. 四川兵工学报, 2014, 35(5): 86-88
- Wang Yue, Cheng Chang-zheng. Application of Collaborative Filtering Algorithms in Movie Recommendation[J]. Journal of Sichuan Ordnance, 2014, 35(5): 86-88

(上接第 229 页)

- [2] 曾玉钰, 翁金钟. 函数数据聚类分析方法探析[J]. 统计与信息论坛, 2007, 22(5): 10-14
- Zeng Yu-yu, Weng Jin-zhong. Initiative research of the clustering approach about functional data [J]. Statistics and Information Forum, 2007, 22(5): 10-14
- [3] 丁晗. 基于函数型数据分析的高中学习成绩评价与预测[D]. 吉林: 东北师范大学, 2009
- Ding Han. High school academic performance evaluation and prediction based on functional data analysis [D]. Jilin: Northeast Normal University, 2009
- [4] 李树良. 心电图形状分析的统计方法[D]. 上海: 华东师范大学, 2011
- Li Shu-liang. Electrocardiogram (ecg) shape analysis of statistical method [D]. Shanghai: East China Normal University, 2011
- [5] Florindo J B, Backes A R, de Castro M, et al. A comparative study on multiscale fractal dimension descriptors [J]. Pattern Recognition Letters, 2012, 33(6): 798-806
- [6] Song J J, Deng W, Lee H J, et al. Optimal classification for time-course gene expression data using functional data analysis [J]. Computational Biology and Chemistry, 2008, 32(6): 426-432
- [7] Park C, Koo J Y, Kim S, et al. Classification of gene functions using support vector machine for time-course gene expression data[J]. Computational Statistics & Data Analysis, 2008, 52(5): 2578-2587
- [8] Suhaila J, Jemain A A, Hamdan M F, et al. Comparing rainfall patterns between regions in Peninsular Malaysia via a functional data analysis technique[J]. Journal of Hydrology, 2011, 411(3): 197-206
- [9] 朱建平. 基于模型参数展开的函数回归及其应用[J]. 商业经济与管理, 2009, 208(2): 81-85
- Zhu Jian-ping. The functional regression model based on basis function expansion of model parameter and its application [J]. Journal of business economics, 2009, 208(2): 81-85
- [10] 靳刘蕊. 函数型数据分析方法及应用研究[D]. 厦门: 厦门大学, 2008
- Jin Liu-rui. The study on the methods of functional data analysis and their application [D]. Xiamen: Xiamen University, 2008
- [11] 李静, 田卫东. 基于 B 样条隶属函数的模糊推理系统[J]. 计算机应用, 2011, 31(2): 490-492
- Li Jing, Tian Wei-dong. Fuzzy inference system based on B-spline membership function [J]. Journal of Computer Applications, 2011, 31(2): 490-492
- [12] 李红广. 基于 B 样条基粗糙惩罚的某些约束函数型数据光滑方法研究[D]. 上海: 华东师范大学, 2008
- Li Hong-guang. Research on methods of functional data with some constraints Based on b-spline basis with roughness penalty [D]. Shanghai: East China Normal University, 2008
- [13] 王剑. 线性回归系数的 Stein 估计[D]. 武汉: 华中科技大学, 2007
- Wang Jian. Stein estimation for linear regression coefficient [D]. Wuhan: Huazhong University of Science and Technology, 2007
- [14] Ramsay J O, Silverman B W. Functional Data Analysis (second edition)[M]. New York: Springer, 2005
- [15] 姜高霞, 王文剑. 经济周期波动的函数型时序分解方法[J]. 统计与信息论坛, 2014, 29(3): 22-28
- Jiang Gao-xia, Wang Wen-jian. Functional decomposition model of time series on business cycle analysis [J]. Statistics and Information Forum, 2014, 29(3): 22-28
- [16] Graven P, Wabba G. Smoothing noisy data with spline functions [J]. Number Math, 1979, 31(4): 377-403
- [17] Ramsay J O, Silverman B W. Applied functional data analysis: methods and case studies [M]. Springer, 2002
- [18] 顾坚, 刘伟. 面向 NUMA 集群的代数多重网格算法优化[J]. 计算机科学, 2014, 4(2): 114-118
- Gu Jian, Liu Wei. Optimizing algebraic multigrid on NUMA-based cluster system [J]. Computer Science, 2014, 41(2): 114-118