



陈松蹊团队 年度总结

Song Xi Chen's Research Group
Annual Summary
www.songxichen.com



岁序更替,华章日新。 我们怀揣着对科研的热爱与 追求,结束了2024年收获颇 丰的旅程。让我们重新站在 这段旅程的起点,回顾往昔, 展望未来。 团队概览

学 术 交 流 果

研究项目

团

2024 团队在科学研究、社会服务方面积极探索,砥砺前行。



本科生

杨巧诗(北京大学) 封羽真、贾一轩 (清华大学)

硕士生

王昺皓、贾 尧、彭阿晓、 钱 磊、冯宇宁 (北大前沿交叉学院)

毕业生去向

王守霞(博后)上海财经大学 嘉(博士) 浙江大学 闫 夏(硕士)国家石油天然气管 网集团有限公司

指导老师

陈松蹊

(清华大学)

邱宇谋

(北大数学学院与统计中心)

郑小谷

(上海张江数学研究院

和新西兰大气和水文研究所资深科学家)

博士生

武、丘竞昆、童培峰、闫 晗、孙浩轩、李天宇、丁心如、 李 和、肖泽霖、詹皓翔、陆启隆、边明远、陈涵玥 (北京大学)

于丁一、陈新佑 (清华大学)

博士后

孙 杨(北大数学学院) 李虹霖 (北大数学学院)

The Pl 2024

陈松蹊教授再次入选2024 Stanford University全球2%顶尖科学家榜单



1月11日 广东省数据科学与技术交 叉应用重点实验室学术委 员会第二次会议·珠海



1月13日 第二届CCF中国数字经济50人论坛 高端峰会·杭州



4月20日 第二届中国青年统计学家论坛·徐州



5月23日 北京数字经济大讲堂·北京



7月12日 第二届全国统计与数据科学联合会议·昆明



8月29日 2024数博会"数字人才培养"论坛·贵阳



10月26日 陕西省第六届丝绸之路青 年学者论坛信息与智能发展大会ICAC ·西安



11月19日 统计学"101计划"工作启动会·北京

陈松蹊教授作为清华大学统计与数据科学系筹备和建设委员会的成员, 引领该校统计与数据科学系的建设和发展。



2024年7月10日清华大学成立统计与数据科学系 (陈松蹊、刘军、姜培学、李路明、蔺涛、严加安、徐宗本、林希虹)













具体地点请 关注后续通知

清华大学统计与数据科学系

扫描二维码

报名讲座

清华大学书院管理中心

主办:

中国科学院院士,数学家、统计学家,清华大学统计与 数据科学系教授,美国统计学会会士、国际数理统计学 会会士、美国科学促进会会士。

1983年本科毕业于北京师范大学数学系,1993年获澳 大利亚国立大学统计学专业博士学位。曾在北京经济学 院、澳大利亚联邦科学与工业研究组织(CSIRO)海洋 实验室、澳大利亚LaTrobe大学、新加坡国立大学、美 国爱荷华州立大学从事教学科研工作。2008至2021年 任北京大学光华管理学院商务统计与经济计量系主任、 联席系主任,2010至2019年任北京大学统计科学中心 创始主任。主要研究方向为超高维大数据统计推断、统 计学与地球科学交叉、经济统计学等。曾获2017年度 高等学校科学研究优秀成果奖(科学技术)自然科学奖 一等奖。

现任中国概率统计学会理事长,第十四届全国政协委员。







"总书记的信让我们也备受鼓舞"

部分海外归来的全国政协委员话拔尖创新人才培养与交叉学科建设

■背景新闻

陈

全

国

极

推

从海外归来扎根教坛 20年的姚期智院士 收到了习近平总书记的回信

记给中国科学院院士、清华大学教授机期 智田信,对姚斯智院士田関任此10年表。 公耕和,取取各种,表示运集回传, 报急,2024年,著名计算机科学专 家城期智进来了全取面国任数20年。姚 期智曾长期任数于美国著名高校,2004 年6月络去美国的济勇赦取回到中国定居 和工作, 现俗语名之世世里传来再发理理 20年来 动斯努力的图象主体系统

息科学领域被失创新人才、克善人工智能 柯城前疆布局、加快起枢世界关进水平作 出了重奏黄献。近日,姚斯智裕习近平总 书记写信, 汇报了回题任数20年来在特 蔡人才、科研创新等方面的情况。 田信中, 习近年总书记就斯施斯智说 士、坚守初心使命、发挥自身优势、带领

大家经验探索创新人才自主培养的模式。

人才培养和科技创新基地,为实现高水平

科技自立自强、建设教育强国科技强国作

麻省理工学院、斯坦福大学、加州大学他克 彩企业 存收新疆土安英高的任务 199 年当远为美国国家科学院院士,2000年获 得图灵英,是首位获得这类的华人学者。 2004年当选为中国科学院外籍院士。2003 年获得中国政府"友谊奖"。 2004年 在"特殊公司" 非军的东边 和杨振宁教授田圖定居并担任清华大学高等 研究中心荣誉主任的总装下,他期智数然决 定棒棒美国普林斯顿大学终身数投职位,会 联回到祖国,被聘为清华大学高等研究中心 教授、并先后担任清华信息科学技术国家实 学院一套港中文大学理论计算机科学研究中 心主任等。后创办交叉信息研究院、人工智 推动学科党及与前沿创新,打造高水平的 核学院并会任院长。2016年业务美国国籍 成为中国公民,正式转为中国科学院院士。

据清华大学官方限专介绍、从期智院士

于1945年12月出生于上海。1972年获哈佛

大学物理博士学位。1975年获得伊利诺伊

大学计算机科学博士学位。此后先后在美国

全国政协委员、香港科技大学(广州)副校长吴宏伟:

青年科学家尽可在祖国施展抱负

港。缘于年少时的梦想,吴宏伟的归国之路 走得异常坚定。 "我在上中学的时候。就和我的老师 讲,我的人生理想就是建设祖国。因为年

在心底的發揮在图外深语財神事課列檢驗 发,促使我国来报效祖国的梦想宣加清晰。" 果宏伟园忆起自己曾经在外求学生活的 日子。"那时总有人因为看到我们是亚洲人 的长档,就间我们是不是日本人、韩国人 在生活上,全习上各方面确实都有领得到的 渴望被尊重的感觉。种种生活上的小事也激 发了我本就热烈的归国情和内心深处的民族

"1995年回到祖国、才直的心安。" 草 宏伟有些动容。"那时国家对各个领域的人 才都很無缺,包括我所从事的专业。因此, 2000年我也被北京大学地质系邀请做兼职 教授,参与周末工程地质硕士班课程的讲 报。" 网忆起那段日子、早安伟分享了一个 动人的小故事。"那时和我对接的教授和我 這載,因为他们没有经费支付我每周的往返 机票和住宿等花铺。我就和他说,课时曹我

这样的转变,吴宏伟见证着,也参与着。 1995年, 他从刘桥大学回国就在香港科技大学任教职, 一直着力开展国际化人才培养工 作。"在我看来,香港科技大学这样的大学一 定要扛起打响中国'招牌'的责任。我们不但 一步提升中国在科研界的学术试验和 创建有 影响力的中国学术体系,以引领更高水平的创

如何实践这一理想?吴宏伟介绍,香港科 技大学广州校区从创建之初便课定了走交叉学 料建设设备情格。"可以说 香港彩越大垒户 州校区所有的学生都是建立在学科交叉的模式 下伯義的。日報为止 我们也可以称谓于某世 界上最大规模的学科交叉培养院校。"吴宏伟 自信且写定.

在他看来, 开展交叉学科建设道路, 首先 批開賽明确学科交叉建设的目的是什么? "应 该是解决达到国际水平,满足国内累求的重大 工程科学问题。"吴宏伟表示,交叉学科建设 是提升我国学术话语权的必重项、也是推动各 种真实问题解决的必需项。"以 AI 发展为例 能解决其发展所出现的现实问题,需要计算



中国科学技技士。1993年,在澳大利亚获博 士孝位。并先后于澳大利亚、美国、新加坡任 新夏、2008年回恩、平业室大学林颢、保存 从事框套建数指统计研究的同时,以需求大年 污染防治的重大需求为出发点,在数学地球物 理领域作出了前沿交叉或果。



異志傳: 全面软协委员, 客港科技大学上 水及环境工程学系讲席教授。研究生说说长荣 "州校区副校长。李湛工植料孝说院士,英国 皇直工程详读十 (中型等)。1993年、京菜图 获得士学位,1995年回国,任春港科技大学 数取。主要从事非地和土力学、生态差土学及 土-结构作用领域的研究工作。



歌阳宏伟: 全国政协委员、农工党浙江省 泰会副主委,民建实社宣布各副主任。美国医 举与生物工程资会士(中国得)。2003年,1 新加坡拼搏十举位并担任排度科学家: 2007 辛回国,任浙江六学医学院教授。他奉先开展 结庆关节软膏运织工程与再生治疗,开拓中国 要丝材料或为医用材料。

免过度研究。有一些研究命题。即使

不计成本接去继译文 訂能和根據真

有突破。甚至还会对科研生态产生负

南影响 电转导性价格轮的"肉

全国政协委员、中国科学院院士陈松蹊

生活中。作为中国人,不论身在何 处,一直心系祖国,关注祖国方方面 面的发展。也希望能有机会为国家

发展贡献力量。2008年,恰巧有这

样的机会,我没有犹豫、就回到了国

人民政协报: 如今十几年已

过,您怎么看得自己曾经的选择?

陈松璐:这是一个"完全正确"的选择!在国外。我的生活确

实看起来舒适安逸, 拿到终身教授

的职位后、要做的事情就是接条学

生、上课,做做自己感兴趣的科

研,一直到退休的生活,完全可以

预别。正所谓"人各有志",能够

的生活。而正是这样的参与。让我

人民政协报: 您的理想在田田

陈检理:实现啦。我们国家"讲求

集由力量办大事"可以开展多领域

结合的"大科研"项目,攻克世界级推

额。填补相关研究领域空白。从愈自

身际从事的结计学研究夹着,可以将

很多国计民生问题。总之,祖国为我

真正参与到国家发展中才是我向往

觉得人生舞台很大。

启去成了吗?

▶▶▶ "无悔当初的选择"

人民政协和: 他了习近年总书 记站中国科学院院士、清华大学教 技迹期智的回信,您是不是也一样 陈松厚:多明、非常感动会深 受赦難。姚先生回国后,为我们国

宏色型发展的人工智能領域投票了 大批人才,作出了极高的贡献。 相信习近干地书记的这封信, 也是写给我们所有从海外归国奉献 于祖国的科研和教育工作者的。它 是对我们已有工作的肯定。更是对 我们要进一步做好为党育人。为国 育才工作的鞭策。激励大家众志成 械,为教育强国。科技强国。人才

人民政协程: 我们了解到您曾 后并后在澳大利亚、新加林、英国 等国家开展教学科研工作,为什么 会说接回图? 十几年前依洗接回图 **时候是怎样的一番店坐?** 陈松蕻:民族感情已经进入自

强国建设提供坚实力量。

脉,自己有时都会感到惊讶。那样 一份强烈的对国家对民族的感情会 影响着自己在大事小情上的一切所 开始在澳大利亚联邦科学院海洋交 验室任続计师。1993年9月份,北 京与恶尼共同竞争2000年夏季奥 运会的申办权。恶兄成功了。瞬 间,澳洲同事们兴奋异常。但是作 为一个中国人的栽却伤心郁闷。我 意识到,我和我的祖国那么深况、 紧密地联系在一起。

人民政协报: 基于您的实现来 看,十会年来,我们国家在前沿科 可以说,这种和祖国命运网想 技创新及人才培养等工作上有哪些 **阿喜的感觉充斥在我在国外的日常**

造成很多科研工作者为了写文 章而去河文章的问题,往往就忽视了现实生活中的真问题。 陈松瑛: 变化很大, 更确切地

高度重视。十余年来, 国家对人才 促使科研环境也在不断优化。以考 核标准为例,当前在人才引进过程 中,除了审核文章发表质量及数量 外,还会逐步用国际标准未考核评 价数值、具值比加、在云边里、会 基于来自世界的同行的评议结果。 更客观和全面地评定人才。同时 科研单位也专门设置了秘书岗位 以让年轻教师摆揽琐碎的日常事务 惠林。凡此种种、朝体现出绘图在

科研环境创设方面的进步。 ▶▶▶要让学生成为学 科交叉的联结者、受益者

人民政协报: 北次田信中。 习近平从书记特别提到空里营料建 设。党的二十大报告中也曾特别指 出、加强基础营料、新兴营料、农 **又学料建设。您作为"数学+** "统计+"的科学家,怎样看持我 国的交叉学科建设? 陈松祺:在我看来,思考交叉学

数学和统计学方法,应用于大气系统 建模与预测 科研数据集中成一条 料建设证额向某手基础研查与应用 研究的综合视像。特别是对于理科 以通过过于气氛接的现金 延伸解决 们搭建了广阔的舞台。我置身其中, 虽然还会有各种挑战,需要忘我付

在我看来,提倡交叉学科建设 就是关注基础研究与应用研究的结 结、为驱动、自然而然地有机结合起 说,是进步很大。最重要的原因在 于以习近于国主为核心的专中央的 总之,就是基础研究也要面向应 用,从实践中发现问题、超级界定命 题,开展研究,也让研究在实践中检 验。以此形成良性循环。

> 的超验?基于您的菌标说野来看,我 们有哪些工作待定基? 践和证,在分香中,最好的立节 学料建设一定是体现在学生培养上。

直白地说,国外在这方面有些好的经 验值得我们坐习。学生们会通过基础 **谑程及联合导师选择的方式,基于自** 己的科研主题开展"交叉"。既保证 了学生的基础学力。又满足了提升学 生综合发展需要。 我们现有的行政体制确实会比较

限制交叉学科的发展,主要体现在人 事制度上。首先、教练的人事编制的 - 小坐除,在工资待遇,结协者枯 教研学习、教授会投票等方面都无法 实现"交叉"。我们要有受活的人事 制度越勤於循鍵全時任即。因討、全 生的指导模式也比较单一,现在大多 教学生只能有一位指导教练 傩以直 公司 基金·ETF

两会前瞻|专访全国政协委员、院士陈松蹊:目前尤为缺 乏高质量再分析科学数据集 急需强化数据分析人才培养力

[建投专题] 直击2024年全国两会-两会前籍 + 2024-03-01 11:19 星期五 財联社记者 王全宝

①目前,中国公共数据获取渠道不畅;科学研究过度依赖国外公共数据集;目前尤为缺乏高质量的再分析科 学数据集

②统计学在数据分析人才培养主力军作用未得到充分发挥;统计学未被列入基础学科;师资严重短缺;政府 部门缺乏统计与数据分析高级专业人才。

财联社3月1日讯(记者王全宝 赵相锋)"尽快优化有关政策,实现公共数据的开放共享,从而促进非公共机构的 数据流动和价值创造,引导中国数据要素市场的建立和良性循环。"2024全国两会即将召开前夕,第十四届全国 政协委员、中国科学院院士陈松蹊在接受财联社专访时建议。

作为数学家、统计学家,陈松蹊对数据要素领域有深入的探索,对数据领域的科研和应用方面的痛点难点有着深 刻的理解。今年全国两会期间,关于数据方面,陈松蹊委员带来了关于"共享公共数据实现科学数据自立自强" 和"加强数据分析人才培养力度"两份提案。为此,财联社结合数据要素相关问题对陈松蹊委员进行专访。











财联社: 两会前瞻 | 专访全国政协委员、院士陈松蹊: 目前尤为缺乏高质 量再分析科学数据集;急需强化数据分析人才培养力度

https://www.cls.cn/detail/1608013

中国信息报:全国政协委员陈松蹊:加强数据分析人才培养力度

http://www.zgxxb.com.cn/pc/layout/202403/07/node 02.html#co ntent 39357

全国政协委员、中国科学院院士陈松蹊:建议加强数据分析人才培养力度 http://epaper.bjnews.com.cn/html/2024-03/05/content 841621.htm?div=-1

CO-PI



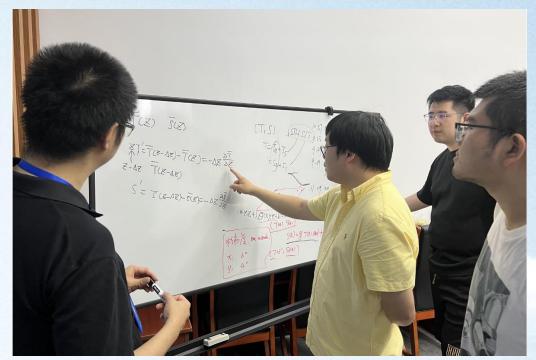
邱宇谋副教授

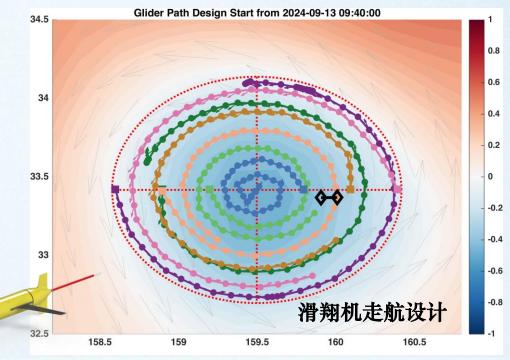
2023年7月入职北京大学数学科学院 概率统计系及统计科学中心,长聘副 教授。研究方向为高维数据分析,因 果推断,缺失数据,统计方法在精准 农业、法医与流行病领域的应用。



邱宇谋参加统计海洋联合课题研究

- Qiu, Y.* and Gu, Y. (2024). F Statistics for High-Dimensional Inference of Linear Model. Accepted by the Bernoulli Journal.
- Qiu, Y.* and Guo, B. (2024). Minimax Detection Boundary and Sharp Optimal Test for Gaussian Graphical Model. Journal of the Royal Statistical Society Series B 86 1221–1242.
- Lee, H., Qiu, Y.*, Carriquiry, A. and Ommen, D. (2024). Density-based Matching Rule: Optimality, Estimation, and Application in Forensic Problems. The Annals of Applied Statistics 18770-793.
- Wang, H., Qiu, Y.*, Guo, H., Yin, Y. and Liu, P. (2024). Information-incorporated Gene Network Construction with FDR Control. Bioinformatic 40,btae 125.





学术会议

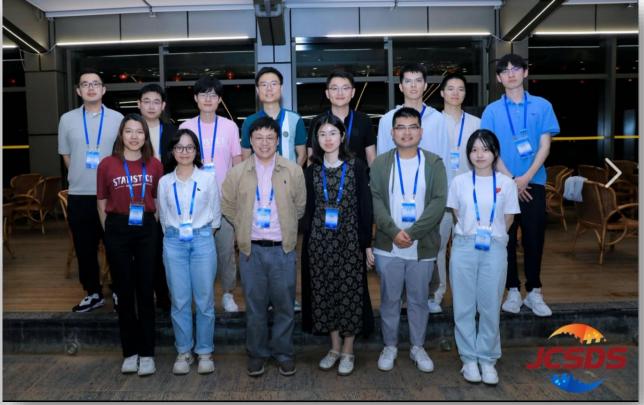
第二 届全国统计与数据科学联合会议

昆明









学生荣誉



苏 武

荣获"第八届北大-清华统计论坛优秀海报奖"



顾 嘉

荣获"第八届北大-清华统计论坛优秀毕业生奖"



童培峰

已在癫痫信号处理方向以第一作者身份发表了五篇 SCI期刊和1篇人工智能顶会论文。

今年3月,北大长沙研究院引进了脑机智能与应用统计中心,童培峰担任主任助理,快速组建起了一支完整的研发团队,专心做技术攻关和产业转化。

本年度毕业生



王守霞 (博士后) 上海财经大学助理研究员



顾嘉(博士) 浙江大学数据科学研究中心 百人计划研究员(助理教授)

研究方向:分布式统计推断,联邦学习,在线推断, 流行病学建模,数据融合

- Gu, J. and Chen, S.X. (2024) Statistical Inference for Decentralized Federated Learning, The Annals of Statistics, 2024, Vol. 52, No. 6, 2931–2955.
- Gu, J. and Chen, S.X. (2024) Distributed Statistical Inference under Heterogeneity, Journal of Machine Learning Research 24(2023)1-57.



闫 夏(硕士) 国家石油天然气管网集团有限公司 科技数字化工程师

研究方向: 半参数建模、函数型数据、高维统计及数据同化

- ◆ 博士后基金面上项目,2023M730090,非平稳函数型时间序列的估计、检验及应用,2023/04-2024/7,结题,主持
- ◆ 国家自然科学基金重大项目,12292983,高分辨率集合Kalman滤波及相关滤波方法与碳源汇融合算法,2023/01-2027/12,在研,参与(项目骨干)
- Hao-Xuan Sun, Shouxia Wang, Xiaogu Zheng, Song Xi Chen(2024)High-dimensional Ensemble Kalman Filter with Localization, Inflation and Iterative Updates. Quarterly Journal of the Royal Meteorological Society, QJ-24-0005.R2.
- Shouxia Wang, Tao Huang, Jinhong You & Ming-Yen Cheng (2022). Robust Inference for Nonstationary Time Series with Possibly Multiple Changing Periodic Structures. Journal of Business & Economic Statistics. 40(4):1718–1731.

全球二氧化碳多源反演算法

Published in partnership with CECCR at King Abdulaziz University

3

ttps://doi.org/10.1038/s41612-024-00824-w

A new global carbon flux estimation methodology by assimilation of both in situ and satellite CO₂ observations

Check for updates

Wu Su¹, Binghao Wang¹, Hanyue Chen², Lin Zhu³, Xiaogu Zheng⁴,5 & Song Xi Chen ® 2,6,7 ⊠



苏武,第一作者 北大前沿交叉学院大数据 中心23级博士生

王昺皓

北大学前沿交叉学院大数 据中心22级硕士生

陈涵玥

北大统计科学中心22级博 士牛 团队基于多源异质观测反演碳通量研究 在《npj Climate and Atmospheric Science》 在线发表

文章提出了一种新的全球碳通量估算方法,通过多源数据同化系统(Multi-observation Carbon Assimilation System,MCAS),同时整合卫星和地面观测数据,以改进碳通量反演的准确性。由于两类观测数据间的异质性,两类观测在同一个地点可能呈现出不同的趋势,导致碳通量计算结果不稳定。现有的方法尚未解决这一问题,国际最权威的碳通量反演产品CarbonTracker直到如今还仅用原位数据进行反演。本研究使用改进的集合卡尔曼滤波算法(Ensemble Kalman Filter),通过不同的膨胀因子处理多源观测数据间的异质性,解决了以往单一数据源同化系统不确定性较大的问题。

研究论文《Statistical Inference for Four-regime Segmented Regression Models》发表在 The Annals of Statistics

分段回归模型(segmented regression models)作为一类介于全局参数模型和非参数模型之间的回归模型,兼具着模型可解释性和稳健性的优点。但因其目标函数的非规则性,分段回归模型的优化和统计推断面临着更多的挑战。本文考虑了一类较为灵活的分段回归模型,不仅在模型形式上对现有文献进行了推广,并且对于优化算法、渐进性质的刻画和推断方法都做出了突破和改进。

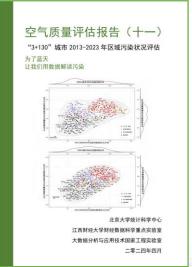


闫晗,第一作者 北大光华管理学院 2020级博士生

文章提出的模型在北京空气污染的实际数据集中得到良好应用效果。通过分段回归建模,文章发现PM2.5和气象变量(如风速、温度和湿度)的关系在不同季节可由三到四段的分段模型刻画,分别对应着污染的形成、传输和清洗的不同阶段,表明文章提出的分段模型在有良好拟合效果的同时具有物理可解释性。



《空气质量评估报告(11)》



《空气质量评估报告(11)》提供了一个全面深入的视角,分析从首个"大气十条"收官以来空气质量的动态,聚焦2017年至2023年的监测数据,清晰展示了疫情这一特殊时期前后的空气质量变化。

团队推出 空气质量评估平台

课题组成员

李虹霖 博士后

北大数学学院

刘俊讲师

江西财经大学财经数据科学重点实验室 **王 佳** 讲师

江西财经大学财经数据科学重点实验室 **郭斌** 教授

西南财经大学统计研究中心、统计学院 其他成员: 詹皓翔、陈涵玥、边明远、 冯宇宁、杨巧诗、闫 夏、彭阿晓、 孙浩轩



No.

数据同化方法

高维集合卡尔曼滤波同化方法

研究论文《结合局部化、膨胀因子和迭代更新的高维集合卡尔曼滤波方法》发表在 Quarterly Journal of the Royal Meteorological Society

文章提出使用统计相合的高维协方差矩阵估计量替代传统集合卡尔曼滤波方法的样本协方差阵,并对其关键参数,局部化距离尺度,给出优化方案。这一替换将能够解决高维情形下的预报误差协方差阵估计问题。同时,一种基于极大似然估计的膨胀因子和迭代更新方法被引入,用以对抗模型偏差。由此,文章建立了一种结合局部化、膨胀因子和迭代更新的高维集合卡尔曼滤波(HD-EnKF)算法。





孙浩轩,第一作者 (北大前沿交叉学院2022级博士生)

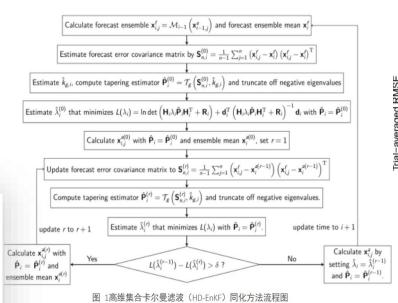
王守霞

北大数学学院博士后

(现上海财经大学统计与管理学院助理研究员)

郑小谷

上海张江数学研究院和新西兰大气和水文研究所资深科学家



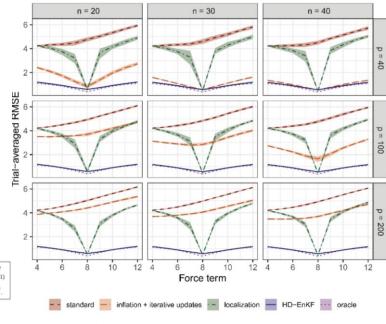


图 2试验平均RMSE及其5%-95%分位数关于不同强迫项F,维数p=q=40,100,200和集合数n=20,30,40

去中心化 联邦学习

研究论文《非中心化联邦学习统计推断》 (Statistical Inference for Decentralized Federated Learning) 在统计学期刊 The Annals of Statistics 发表

文章探讨了基于最通用的非中心化联邦学习(DFL,图2)算法的异质性M估计的统计推断问题,该算法囊括了许多SGD算法作为特例。文章推导了估计量空间平均轨迹的均方误差(MSE)上界以及各客户端本地估计之间的共识误差(consensus error)上界,还证明了基于DFL算法的Polyak-Ruppert(PR)平均估计量的渐近正态性。研究表明,当联邦学习系统规模不是特别大的情形下(相对于每个客户端本地样本量T),PR-估计量的渐近方差和全样本M-估计量相同,这意味着PR平均估计器是统计渐近有效(efficient)的。

https://doi.org/10.1214/24-AOS2452

© Institute of Mathematical Statistics, 2024



顾嘉,第一作者 2024年7月博士毕业于北大统计科学中心, 现为浙江大学数据科学研究中心助理教授。

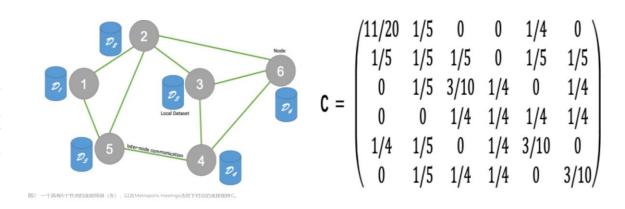
STATISTICAL INFERENCE FOR DECENTRALIZED

FEDERATED LEARNING

By Jia Gui^{1,a} and Song Xi Chen^{2,b}

¹Center for Data Science, Zhejiang University, ^agujia@zju.edu.cn ²Department of Statistics and Data Science, Tsinghua University, ^bsschen@tsinghua.edu.cn

This paper considers decentralized Federated Learning (FL) under heterogeneous distributions among distributed clients or data blocks for the Mestimation. The mean squared error and consensus error across the estimators from different clients via the decentralized stochastic gradient descent algorithm are derived. The asymptotic normality of the Polyak–Ruppert (PR) averaged estimator in the decentralized distributed setting is attained, which shows that its statistical efficiency comes at a cost as it is more restrictive on the number of clients than that in the distributed M-estimation. To overcome the restriction, a one-step estimator is proposed which permits a much larger number of clients while still achieving the same efficiency as the original PR-averaged estimator in the nondistributed setting. The confidence regions based on both the PR-averaged estimator and the proposed one-step estimator are constructed to facilitate statistical inference for decentralized FL.



为了在更大规模的非中心化联邦学习系统中进行统计推断,文章提出了一种计算高效、统计有效的一步更新(one-step update)估计量。该一步更新估计量以步长较小的PR平均估计量作为初始估计量,并通过一个修正项来改进其统计效率。对于损失函数是否光滑(具有二阶可微性)的不同情形,文章针对性地给出了置信域的构造方法,并且建立了对应的理论保证。



EEG-X智能脑电分析平台

数据赋能脑科学,癫痫诊疗效率有望大幅提升



脑科学

自主研发的全流程癫痫智能诊疗云台——EEG-X智能脑电分析平台在首届"医工融合·尖山论道——湖南省智慧健康与脑机智能创新发展论坛"上亮相。面向脑科学临床问题,分析并提取脑部数据蕴含的丰富信息,完成大脑的数字孪生,用数据辅助临床决策,对提高脑部疾病的临床诊断效率,扩大优质医疗资源的覆盖范围,惠及更广大的患者群体具有关键意义。

童培峰为第一作者的两篇研究论文在线发表:

- 1、《偏差调整的时空脑电图/脑磁图溯源定位和统 计推断》
 - (Debiased Estimation and Inference for Spatial –Temporal EEG/MEG Source Imaging)

发表于医学影像顶级期刊

《IEEE Transactions on Medical Imaging》

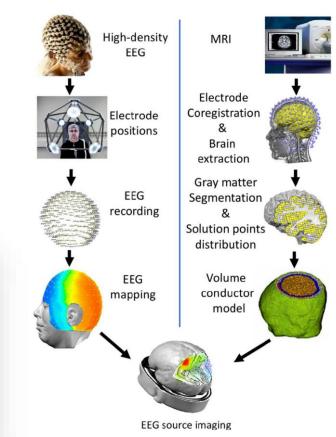
- 2、《基于Transformer的儿童Rolandic癫痫患者的 间歇期痫性放电检测算法》
 - (Detection of interictal epileptiform discharges using transformer based deep neural network for patients with self-limited epilepsy with centrotemporal spikes)

发表于医学工程权威期刊

《Biomedical Signal Processing and Control》



在算法核心研发者、北京大学童培峰博士的介绍和演示下,与会者见证了EEG-X平台如何高效处理复杂的脑电信号,通过量化统计分析与直观的可视化呈现,助力医生快速准确地解读脑电数据,并一键生成专业的读图报告,显著提高了癫痫疾病的临床诊疗效率。





2024发表论文成果



已接收的论文:

- 1) Zhang, SY, S.X. Chen and Yumou Qiu (2024) Mean Tests For High-dimensional Time Series, Statistica Sinica, to appear.
- 2) Tong, P.F., Chen, S.X. and Tang, C.Y. (2024) Multivariate calibrations with auxiliary information, Statistica Sinica, to appear. DOI:10.5705/ss.202023.0151.
- 3) Zheng, Xiangyu and Chen, S.X.(2024)Segmented Linear Regression Trees, Acta Mathematica Sinica, to appear.
- 4) Pei Feng Tong, Bosi Dong, Xiangdong Zeng, Lei Chen, Song Xi Chen. (2025) Detection of interictal epileptiform discharges using transformer based deep neural network for patients with self-limited epilepsy with centrotemporal spikes, Biomedical Signal Processing and Control, Volume 101, 107238, to appear.

- Jia Gu, Song Xi Chen.(2024) Statistical Inference for Decentralized Federated Learning, The Annals of Statistics, 2024, Vol. 52, No. 6, 2931–2955.
- Han Yan, Song Xi Chen.(2024) Statistical Inference for Four-regime Segmented Regression Models. The Annals of Statistics, 2024, Vol. 52, No. 6, 2668–2691.
- Su, W., Wang, B., Chen, H.et al.(2024) A new global carbon flux estimation methodology by assimilation of both in situ and satellite CO2 observations.npj Clim Atmos Sci 7, 287 (2024). https://doi.org/10.1038/s41612-024-00824-w.
- Pei Feng Tong, Haoran Yang, Xinru Ding, Yuchuan Ding, Xiaokun Geng, Shan An, Guoxin Wang, and Song Xi Chen. (2024) Debiased Estimation and Inference for Spatial—Temporal EEG/MEG Source Imaging, IEEE Transactions on Medical Imaging.
- Hao-Xuan Sun, Shouxia Wang, Xiaogu Zheng, Song Xi Chen(2024)High-dimensional Ensemble Kalman Filter with Localization, Inflation and Iterative Updates. Quarterly Journal of the Royal Meteorological Society, QJ-24-0005.R2.
- Zhang, Fei-Chi; Xie, Jin-Han; Chen, Song xi; Zheng Xiaojing. (2024) Robust Relation of Streamwise Velocity Autocorrelation in Atmospheric Surface Layers Based on an Autoregressive Movingaverage Model. Journal of Fluid Mechanics, Volume 981.
- Hanyue Chen; Song Xi Chen; Mu Mu. (2024) A Statistical Review on the Optimal Fingerprinting Approach in Climate Change Studies. Climate Dynamics, 62(2):1439-1446.
- Zheng, X-Y. and Chen, S.X. (2024) Dynamic synthetic control method for evaluating treatment effects in auto-regressive processes. Journal Of The Royal Statistical Society Series *B*,86(1): 155-176.
- Tong, P.F., Zhan, H.X. and Chen, S.X. (2024) Ensembled Seizure Detection based on Small Training Samples, IEEE Transaction on Signal Processing, 72: 1-14.DOI: 10.1109/TSP.2023.3333546.





10月20日团队秋游:探访八大处



12月6日我们的导师: 郑小谷教授75岁生日



9月17日团在青岛崂山实验室欢度中秋

国家自然科学基金重大研究计划

"西太平洋地球系统多圈层相互作用数据集成研究"





2024年4月1日,由北京大学牵头承担,中国海洋大学、青岛海地质研究所、江西财经大学、华东师范大学参与的国家自然科学基金重大研究计划"西太平洋地球系统多圈层相互作用"集成项目"西太平洋多圈层相互作用数据集成研究"启动会在山东青岛举行。

集成项目负责人陈松蹊院士重点讲述了项目在物 理与数据双驱动建模、大气模式的人工智能算法、 碳源汇计算的国际融合算法、极端天气的数值模 式预报等方向预期实现的亮点成果。

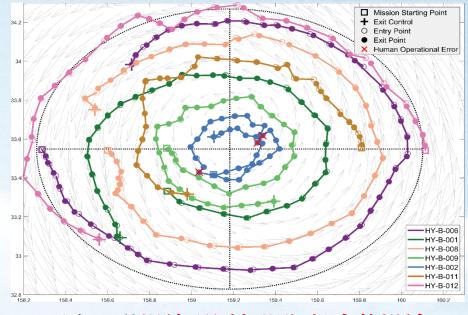




Chasing an Ocean Eddy with Gliders and Statistics

助力西太科学观察 2024年9月14-22日

共享航次计划2023年度重大科学考察项目 (编号: NORC2024-584)



西太平洋涡旋观测与滑翔机走航设计

(左,中)团队在青岛崂山国家实验室, (右)团队完成任务回到北京







国自然重大项目

融汇海量观测数据的大气系统建模与预报中的关键数学问题与算法 (在研)



2023.01—2027.12

2024年,项目第二年, 获得相关成果共38项, 项目进展顺利。

2024年6月13日项目半年度会议

经面

本项目建立面向儿童癫痫和自闭症的智能分析系统以及辅助诊断决策支持系统,推动儿童脑发育障碍性疾病的研究和应用向前发展。

关于非侵入式脑电溯源定位和脑电棘波/尖波检测的两篇研究论文发表

ARTICLE 1

ARTICLE 2





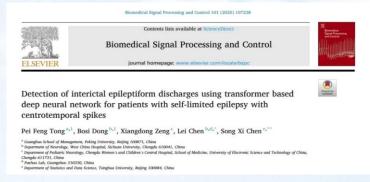


IEEE TRANSACTIONS ON MEDICAL IMAGING, VOL. XX, NO. XX, XXXX 2024

Debiased Estimation and Inference for Spatial-Temporal EEG/MEG Source Imaging

Pei Feng Tong, Haoran Yang, Xinru Ding, Yuchuan Ding, Xiaokun Geng, Shan An, Senior Member, IEEE, Guoxin Wang, Senior Member, IEEE and Song Xi Chen

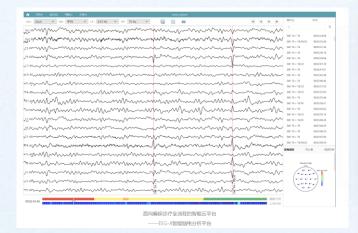
第一作者为北大光华管理学院2020级博士生童培峰



童培峰与四川大学华西医院神经内科董博思博士为共同第一作者

自主研发EEG-X智能脑电分析平台

融合了高精度的人工智能和统计分析算法以及高质量的临床数据与诊疗经验,为医生提供了前所未有的智能辅助诊断工具。



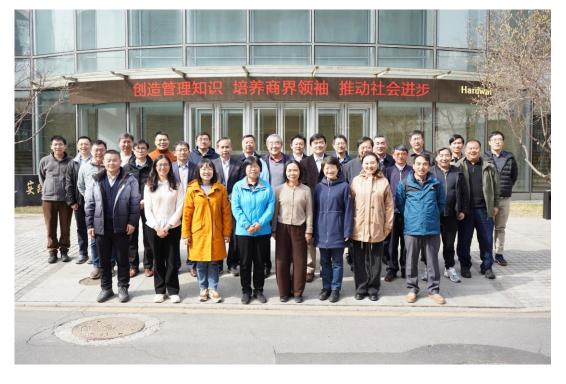
展望未来,EEG-X智能脑电分析平台将秉持普惠医疗的理念,提高院内脑电检查的诊断效率,推动针对癫痫疾病的一流诊疗能力向基层医疗机构推广,从而改善优质医疗资源分布不平衡不充分的现状,帮助更多罹患癫痫疾病的患者及时获得诊治,实现医疗价值、社会价值和民生价值的有机统一。

由北京大学、中国人民大学、 南开大学、中国科学院数学与 系统研究院牵头

设计、收集、汇总全国七十多所 高校调研问卷并进行数据分析



3月18日统计学科调查研究项目咨询会



对统计学科进行全面摸底, 了解统计学科现状和社会需求, 为数学及其他学科调研做出范例。

2024.01——2024.12

国自然面上项目-变系数流行病学模型的统计推断 (结题) 2021.01—2024.12

Domestic Reports

- ₹3月9号疫情分析简报
- ₹3月4号疫情分析简报
- ₹3月2号疫情分析简报
- ₹2月29号疫情分析简报
- ₹2月24号疫情分析简报

国内疫情报告

- 4、Zhu Yuru; Gu Jia; Qiu Yumou; Chen Songxi;, The Annals of Applied Statistics.
- 5、Zhu Yuru; Gu Jia; Qiu Yumou; Chen Songxi, Research.
- 6、Zheng Xiangyu; Chen Songxi, Journal of the Royal Statistical Society Series B: Statistical Methodology.
- 7. Ying Zhang; Song Xi Chen; Le Bao, Environmetrics.
- 8. Chen Songxi; Qiu Yumou; Zhang Shuyi, The Annals of Statistics.
- 9. Chen Songxi; Guo Bin; Qiu Yumo, Journal of Econometrics.

1、 Yan, Han; Zhu, Yuru; Gu, Jia, et al., Proceedings of the Royal Society A-Mathematical Physical and Engineering Sciences.

- 2、顾嘉; 陈松蹊; 董倩; 邱宇谋, 统计研究.
- 3、Zheng, X-Y.; Guo, B.; He, J.; Chen, S.X., Environmetrics.

International Reports

Apr 23rd International 4月23日国际疫情报告 Report

Apr 3rd International R 4月3日国际疫情报告 eport

Mar 29th International Report 3月29日国际疫情报告

Mar 21st International R3月20日国际疫情报告 eport

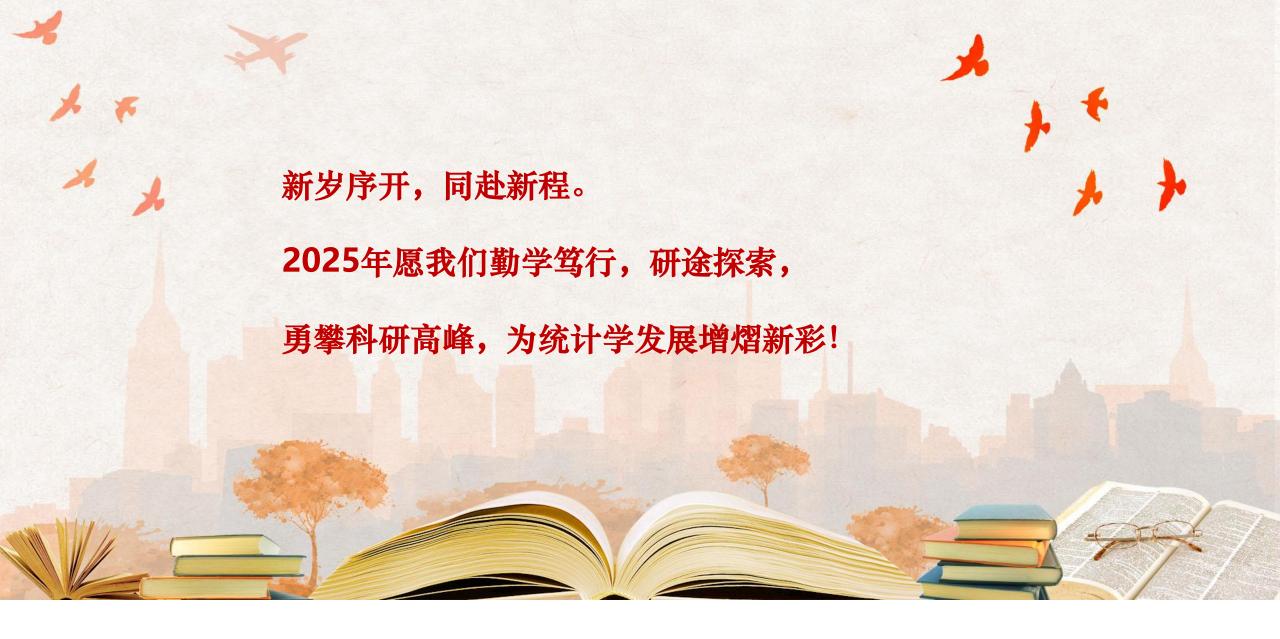
Mar 16th International 3月16日国际疫情报告 Report

Mar 13 International 3月13日国际疫情报告 Report

Mar 6th International R 3月6日国际疫情报告eport

Mar 4th International 3月4日国际疫情报告 Report

国际疫情报告



让我们携手共进,一路"统"行!