

**大学生创新训练计划**

**项目申报书**

**项目名称： 无人巡航艇中的位姿动态分析方法**

**项目负责人： 叶子怡**

**所在学院： 数 学 学 院**

**专业年级： 2019级 统计学**

**学 号： 2019141210008**

**手 机： 177 2097 2025**

**电子邮箱： 3282647072@qq.com**

**指导教师： 宋恩彬**

**项目起止年月： 2021年11月至2022年10月**

**项目参与学生人数： 5**

**四川大学教务处制**

2021年 10 月

填写说明

一、凡申报**四川大学“大学生创新训练计划”**必须填写本申报书。创新训练计划项目是本科生个人或团队，在导师指导下，自主完成创新性研究项目设计、研究条件准备和项目实施、研究报告撰写、成果（学术）交流等工作。

**二、“项目所属一级学科和代码”**参考《普通高等学校本科专业目录和专业介绍（2012年）》。

三、**“项目开展支撑平台”**指支撑本项目开展的国家级和省部级重点实验室（中心、平台等）、国家双创示范基地平台、教学实验中心（实验室）、企业、事业或其他单位等，表中填写平台名称，可以多个。

四、**“项目组成员”**人数原则上不超过五人，应排序。

五、**“项目成熟度**”请参考附件《项目成熟度量表》。

六、本书应该填写完整、内容详实、表达准确，数字一律填写阿拉伯数字。

七、报送申报书的电子文档至负责人所在学院。

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 项目名称 | 无人巡航艇中的位姿动态分析方法 | | | | | | |
| 申请类别 | √科学探索与工程技术类  □人文艺术与社会科学类 □交叉学科类 □企业命题组 | | | | | | |
| 申请经费 | 元 | | | 起止时间 | | 2021年11月  至2022年10月 | |
| 项目所属 一级学科和代码 |  | | | | | | |
| 项目开展 支撑平台 |  | | | | | | |
| 项目来源  （可多选） | □进课题组、进实验室、进科研团队参与的项目  □国家级和省部级重点实验室（中心、平台等）、国家双创示范基地平台支持申报项目  □指导老师的横向课题（对应企业命题组）  □揭榜企业类双创示范基地创新需求（对应企业命题组）  □交叉学科创新项目  □“青年红色筑梦之旅”计划项目  √基于前期研究实践成果、继续深入研究实践的创新项目  □其他 | | | | | | |
| 企业命题组项目合作情况（非企业命题组可不填） | 合作企业名称 | | - | | | | |
| 合作项目名称 | | - | | | | |
| 校内指导老师姓名 | | - | | | | |
| 企业联系人姓名 | | - | | | | |
| 企业联系人手机号 | | - | | | | |
| 所属重点支持领域（可不选） | 选择1项：K.社会事业与文化传承 | | | | | | |
| 负责人之前参与大创项目情况 | 格式如下，有则填，不限条目：  1. 负责人/团队成员，项目编号，项目名称，立项级别，项目类别，立项年份，结题成绩； | | | | | | |
| 项目成员之前参与大创项目情况 | 1. 李垚，团队成员，C2021117628，机构交易对股票价格的过度冲击研究，院级项目，创新训练，2021，待结题； | | | | | | |
| 项目负责人基本信息 | | | | | | | |
| 姓名 | 学号 | | 专业年级 | | | 所在学院 | |
| 叶子怡 | 2019141210008 | | 2019级 统计学 | | | 数学学院 | |
| 性别 | 手机 | | 电子邮箱 | | | 身份证号 | |
| 女 | 177 2097 2025 | | 3282647072  @qq.com | | | 350784200106231045 | |
| 项目组成员基本信息 | | | | | | | |
| 序号（含排序） | | 1 | 2 | | 3 | | 4 |
| 姓名/性别 | | 邵毅诚(男) | 叶仕文(男) | | 李垚(男) | | 张珊珊(女) |
| 学号 | | 2019141210035 | 2019141430208 | | 2018141482170 | | 2019141210046 |
| 专业年级 | | 2019级  统计学 | 2020级  数学类 | | 2019级  基地班 | | 2019级  大数据 |
| 所在学院 | | 数学学院 | 数学学院 | | 数学学院 | | 数学学院 |
| 手机 | | 133 5797 7095 | 187 8224 4909 | | 157 5648 8019 | | 152 0252 4369 |
| 电子邮箱 | | 757504160@qq.com | 496242502@qq.com | | liyaao@stu.scu.edu.cn | | 1195569177@qq.com |
| 身份证号 | | 320811200102061511 | 510182200011134837 | | 610114200004190010 | | 632122200108290062 |
| 签名 | |  |  | |  | |  |
| 指导教师1 基本信息  （非交叉学科类项目一般仅允许一位指导老师） | | | | | | | |
| 姓名 | 所在学院或单位 | | 研究方向 | | | 职称/职务 | |
| 宋恩彬 | 四川大学数学学院 | | 应用统计 | | | 教授 | |
| 性别/年龄 | 手机 | | 电子邮箱 | | | 签名 | |

|  |  |  |  |
| --- | --- | --- | --- |
| 男 42岁 | 130 1828 3193 | e.b.song@163.com |  |
| 指导教师2 基本信息  （交叉学科类项目需填写第二指导老师） | | | |
| 姓名 | 所在学院或单位 | 研究方向 | 职称/职务 |
| - | - | - | - |
| 性别/年龄 | 手机 | 电子邮箱 | 签名 |
| - | - | - | - |

|  |  |  |
| --- | --- | --- |
| **项目内容概述(限200字以内)** | | |
| 无人艇航行过程中,会受到外界各种因素，如：水域环境，海上风、浪、流等，导致翻艇现象。本项目旨在通过风、浪、船体状态等相关数据的采集，经过判别分析、非参数检验、滤波估计等统计学方法，对外界环境的影响，以及船本身传感器的影响进行动态分析，检测无人艇的航行稳定性状态，进而利用结论优化无人艇的动态位姿。 | | |
| **项目特色创新点概述（限100字以内）** | | |
| 1. 紧跟无人驾驶技术潮流；  2. 动态分析与实时监测；  3. 多元分析与因素融合；  4. 误差分析与不确定性评估；  5. 滤波处理与优化误差因素；  6. 贴合最新时代实际功用。 | | |
| **项目成熟度评估（项目成熟度请参考附件《项目成熟度量表》）**   |  | | --- | | 目前项目成熟度自评估为 级；  预期结题时项目成熟度达到 级。 | | |  |
| **项目组成员分工** | | |
| **姓名** | **承担工作内容** | |
| **叶子怡** | **合理分工，联系指导老师，模型动态分析** | |
| **邵毅诚** | **收集数据，模型动态分析，撰写文案** | |
| **叶仕文** | **查阅资料，整合信息，撰写文案** | |
| **李垚** | **数据处理，文案修改，文案审核** | |
| **张珊珊** | **数据处理，插值分析，撰写文案** | |

|  |
| --- |
| **一、项目简介** |
| **1. 研究内容**  1.1 概述（背景）  随着无人机、无人车技术的逐渐成熟，无人系统在海洋中的应用也逐渐广泛。水面无人艇是可以在海洋、河流等环境中自主完成任务的平台，是自动驾驶技术在水面环境中应用的最主要体现。相较于浮动平台、载人船舶、UUV等水下、水上设备，无人艇具有部署方便、操纵性强、成本低、覆盖范围广的优点。  同时，由于非载人的特点，无人艇能够适应更加危险的工作环境，在一些人类不可到达的场景中完成作业，极大程度上促进了人类探索事业的发展。所以，这一项目具有十分广阔的应用前景。  而为了充分满足航行与作业的要求，无人艇应该能够具备对环境的感知能力、分析能力与对自身在各种环境条件下在一定速度下是否能够保持稳定的判断能力。（需求）  在无人艇对自身位姿进行动态分析的过程中，需要对本体状态和外界环境进行感知进而进行自身位姿动态分析。位姿动态分析是保证无人艇稳定航行的关键，主要包括无人艇的姿态信息（横、纵倾角，横、纵摇，升沉幅度等）和运动状态信息（航向角、船速、位置等）。外界环境感知则是无人艇根据自身用途对特定信息的获取，主要包括水面条件、障碍物信息、气象条件等方面的综合信息。无人艇工作时的水面环境具有变化性、高动态等特点，所以无人艇自身的状态也会随着环境的不同发生变化，因此精确的位姿动态分析方法是无人艇能够安全执行任务的前提与基础。（重点解释与重点分析，挖掘问题的产生）  当无人艇在水面进行巡航时，受到的最大影响来自于海面上的风、浪、流以及水域本身的环境。所以本课题组此次将主要研究内容定位在——当已知精确的风力影响与无人艇自身行驶数据、缺乏精确海浪数据时如何通过数学手段对采集到的数据进行处理，对无人艇的位姿进行动态分析，既而判断出在当前水域、当前风浪条件与当前速度条件下，无人艇翻船的概率大小。（问题的要点，可以充实内容）  随着未来海洋探测、国防军事等领域的需求，无人艇势必会成为海上无人系统的重要组成部分，如何使无人艇具备准确的自身位姿动弹分析能力也必将成为研究重点。（展望）  **1.2 预想的总体流程**  **1.2.1 数据采集**  目前，已有措施已经能够做到让无人艇在水面航行时，对风向风速进行高精确度的测定，也能够获得无人艇的状态数据如船的时间、位置、倾斜角、俯仰角、偏航角等安全数据，指标等重要作业信息。而对于浪的数据，我们目前没有精确的数据采集方法，但可以通过螺旋桨的反作用力、溅起的浪花等数据侧面采集到关于浪的信息。  **1.2.2 数据预处理**  首先，对于数据的总体特征，对于几种常见的重要分布（对大数据样本，优先考虑正态分布），那么对于我们这个项目而言，我们先进行高斯性质的判定，决策参数/非参数类型。为此，选取合适的多元正态性的**稳健性检验方法**（充分照顾到检验方法对噪声的敏感度），同时也可通过插值拟合的方法，即在已知时间和位置的条件下，我们再测定每点的速度，进而得到一列较为连贯的数据集并对它进行拟合。  但要想提高精度，我们需要解决样本数据容量过低的问题，在考虑到经济成本的前提下，尽可能地去获取更多的、有效的样本数据信息。与此同时，本团队需要坚持统计学的分析原则与实事求是原则，结合相关领域从事科研人员，对于异常数据样本点进行决策、剔除，对于符合统计学规律，但与现实情况有所出入的推断结果，交由专业人员评估判断，分析决策。  **1.2.3 故障检测**  这一部分是要通过无人艇的自身数据来判断出无人艇的状态是否正常，即翻船的概率有多大。我们准备使用判别分析的方法，即收集一系列的数据集，比如{倾斜角、俯仰角、偏航角、是否翻船}，并且用贝叶斯算法、SVM或者卷积神经网络等机器学习方法，通过上述数据集进行训练，再用训练后的系统对实际使用中采集到的数据，即倾斜角、俯仰角、偏航角来对是否翻船进行预测与估计，优化无人艇的航行位姿。  **1.2.4 动态分析**  这个部分我们主要研究风和浪对船的影响。我们会先从几种方法入手，例如多维正态分布检验、非参数检验、滤波估计等方法，再在实践调研后，考虑受船本身传感器影响产生的噪声，选择一种方法或者尝试融合已有的多维正态分布检验方法得到一种适合本课题的方法对噪声的多维正态性进行检验。  **2.目的意义（摘录小标题+下方部分文本）**  **2.1 提高无人艇巡航作业的运行稳定性、安全性、成功率以及效率等**  目前，中国自主研发的智能无人艇已经在极浅水区、海洋岛礁等多种环境状态下得到成功应用，无人艇的应用前景已然显露出它的曙光。但是，随着人类探索需求的进一步扩大以及社会生产生活的进一步增长，我们更迫切地需要无人艇投入实际生产生活，科研探索等领域应用的是那些环境条件非常恶劣，人类很难前往勘探的地方，例如：南极洲等。  而在这种恶劣、不可测的条件环境中，无人艇如果能够对自身的位姿状态进行实时、准确的分析，可以有效地避免翻船导致的任务失败以及对物力资源、时间的浪费。特别是在救援任务中，执行任务的无人艇如果翻船，或许就意味着被救援人员获救的希望破灭。  **2.2 对国家战略储备方面作出一定的贡献**  在现在这个科技高度发达的时代，战争往往是顶尖科技技术的硬碰硬，而现在的作战部队更是越来越向无人化靠拢。比起智能无人机、智能无人坦克，智能无人艇等高端核心技术或更为复杂，更依赖于先进的科技支撑。  从上个世纪90年代开始，一些发达国家已将智能无人艇研发纳入军队武器装备现代化革新的重要项目。智能无人艇具备体积小、速度快、雷达反射面积不大等卓越性优势，可在江河湖海上灵活部署。  它被认为是一种作战用途广泛的新型智能水面作战力量。一个更加稳定、能保证人身作战安全的无人艇当然能在战场上发挥更大的作用，具有着举足轻重的战略性意义。  **2.3 推广无人艇的使用**  无人艇现在的使用并没有得到充分的普及率以及社会各界的广泛认可，当然与其相对高昂的成本，以及部分亟待解决的责任纠纷问题有关。  而若是我们能从以安全至上的原则出发，在最大程度上保证无人艇的运行安全的前提之上，尽最大可能地去减少无人艇的能源损耗，以及航行损耗等经济问题，等于变相降低了无人艇的运行成本，获得最大的实际效益。  若无人艇能得到普及与推广，那么将会最大可能地预防相关人员在进行海上作业时伤亡事件的发生，同时提高一些特定任务的完成效率。  **3.具体目标**  **3.1 研究过程性目标**   * 通过同济大学葛泉波研究员以及研究生学长学姐们提供的数据支持，以及无人艇相关理论专业知识与技术指导。能够对于无人艇运行的稳定性主要影响因素，科学地采用统计学方法进行评估与合理的参数估计，确定已有的采集数据进行预处理与优化。 * 对于异常样本点，能够结合本项目团队指导老师意见以及同济大学团队的相关专业意见，对整体样本数据进行一系列的综合性评价，充分保证研究对象的科学性，可靠性。 * 能够通过广泛实践调研，结合科学理论知识，做到坚持实事求是的重要原则，找到一套合适的统计推断、估计、预测方法体系（如：多元正态性检验的稳健方法的决策），以此为基本研究逻辑，搭建出能做到根据风、浪、无人艇自身运动数据动态分析，得出无人艇理想的稳定性运行情况，并且在此基础之上，不断追求达到更高的精确程度。 * 能在自制环境中得到验证，建立出一套符合实际运行情况的数学模型，科学模拟无人艇运行的稳定性状况，为无人艇的实际作业运行提供进一步的数据支撑与运行指导。   **3.2 研究应用展望（任意摘录）**   * 在日后该项目的实际操作环节当中，能够经得住实践的考验，并且根据实际情况，不断地优化模型内部的逻辑结构，进而有序地投入现实生活航行领域当中。 * 能够实时跟进时代最新的前沿发展动向，不断地根据无人艇以及其相关领域的发展动向，优化已建立的航行模型。并且，适时追踪应用场景的延拓方向，尽可能地加大应用的广泛性与普适性，提高项目成果的应用价值。 * 能够在一定程度上推动和加快无人技术相关领域的发展，进一步推进知识融合，产业融合，倡导多元化、多领域、多层次的元素相互融合，共同促进与发展，激发产业活力与产业创新潜能，提供发展动力。   本项目团队将尽己所能，为国内无人艇事业作出一定的实际应用贡献！  **4.国内外研究现状分析及评价**  **4.1 时代发展背景**  随着第三次工业革命信息技术产业的大发展以及第四次工业革命智能化产业发展的来临，无人装备制造领域取得了快速发展。无人艇（即：无人水面航行器(**Unman Surface Vehicle**)）作为无人装备制造领域的重要组成成分。同时，作为全世界范围内对于海洋资源探索的重要工具，无人艇及其相关领域有着充足的发展潜力与应用需求，对世界各国的综合实力以及科技进步有着极大的促进作用。  **4.2 研究现状分析**  目前与本次项目相契合的核心技术是**强扰动环境下的运动控制技术**。  针对海洋动力环境中风、浪、流对水面载具的影响，开发欠驱动运动控制算法，实现如高精度循线航行、高海况自航、自守位等功能。常规海洋调测多采用循线走航探测或定点剖面观测的方式作业。无人船体量小、推力有限，且属于典型的欠驱动运动体，因此容易受到海面动力因素的扰动，出现偏航距超限、航向不稳定、失速及大幅度摇动导致甲板上浪甚至倾覆的情况。因此在无人船运动控制方面，需要设计基于定位、姿态数据的前馈控制进行运动补偿，通过无人船实时位置、航速、艏向、横纵摇、升沉数据，判断风、流的恒定扰动以及波浪的周期性扰动的强度、方向，结合无人船惯性力、阻力、恢复力运动模型，提前干预推力矢量输出，确保无人船能够在以较高精度循线航行，或在高海况下通过艏向、推力调节确保无人船不失速及发生大幅度横摇。  ——节选自《一文看懂无人船现状及发展趋势》  无人艇实现自主感知、自主航行，首先要适应海洋环境特点，遵循相关海事规则。那么，无人驾驶首先从基础实现层面上来看，需要依靠导航、传感系统等的协同运作。  与自主航行直接相关的航行控制分系统又是无人系统的核心，它分为三个部分：  制导系统（Guidance system）。这个词借用导弹技术，主要是根据分配的任务、环境条件，以及导航信息，为船舶提供稳定、连续、优化的航行轨迹。  导航系统（Navigation system）。提供船舶的位置、航向、航速、加速度等实时的信息，感知周边态势。信息源来自艇上的各类传感器，如导航雷达、陀螺、卫星定位系统、光学系统、声呐、海流风流测量等等。  控制系统（Control System）。根据制导系统、导航系统提供的指令信息，对船舶航行实施自动控制，主要通过对主机、舵的控制。如果需要完成相关作战任务，需要具备对相关任务平台的远程控制能力。比如，水声通讯、水下搜索，水上水下目标打击等。  ——节选自《无人艇关键技术的现状与发展趋势》  船体内部控制系统的协同运作的研究仍然是我们目前这个时代的热门议题，在数学方面有重要分支：运筹学与控制论，在其中扮演着举足轻重的指导性地位，吸引着一代又一代数学人才踏上时代热潮。  国内外的卓越成果也是十分丰硕的：   * 哈尔滨工程大学在科技部“863项目”支持下开发了“天行一号”无人船。复合动力推进，航速超过50节。具备自主完成地形地貌测绘、水文气象信息采集能力。 * Ocean Infinite公司使用L3-ASV公司生产的C-Worker8作为水面通信节点和水下导航定位支持系统，通过与搭载声学探测设备的AUV组网协同搜索马航M370残骸。 * Sail Drone公司推出的风帆动力无人船：通过将原生海洋能作为直接动力实现超长航时，执行中、大尺度物理海洋观测任务。 * 法国ECA公司开发了“Inspector Mk”系列USV可执行浅水/极浅水测量和检查、濒海/沿海水文测量工作、海港/海上设备监测和保护、目标探测和分类等典型任务。   **4.3 评价**  无人艇的相关技术仍然是热门议题，但总体水平正迫切谋求上升当中。无人操作系统仍然无法独立地进行完整的作业任务，特型平台的构建以及智能系统的协同合作正逐渐成为无人系统方向发展的刚需。  无人系统是一个待挖掘的潜力股，我们应紧随时代步伐，做好时代新人！  参考消息：   1. 华东军民融合装备技术研究院, 2020年10月19日,《一文看懂无人船现状及发展趋势》； 2. 知乎作者 若谷（广州傲胜机器人科技有限公司 技术总监）, 2020年05月29日, 《无人艇关键技术的现状与发展趋势》. |

|  |
| --- |
| **二、研究技术路线及可行性分析** |
| 1. **研究技术路线**     图1 研究技术路线脉络图  **1.1 数据采集**  首先我们需要采集无人艇航行过程中风的数据，得到风向和风速。  然后，通过螺旋桨的反作用力、溅起的浪花等数据侧面采集到浪的信息。  最后，我们通过观测船体本身的性质，即船的航行时间、航行位置、倾斜角、俯仰角、偏航角等物理量以及船体是否侧翻，得出无人艇航行过程中风、浪对船体的作用效果以及作用强度。  **1.2 数据预处理**  已知{时间}和{位置}，需要测定每点的{速度}。一方面，需要对数据进行清洗，剔除野值。另一方面，需要通过插值、拟合解决采样频率过低的问题。  **1.2.1 野值的处理**  假设变量服从正态分布，我们选取正态分布的参考线法识别异常点。  根据正态分布的定义，数据点落在偏离均值值内的概率为68.2%，落在偏离均值的域内的概率为95.4%，落在偏离均值域内的概率为99.6%。  由于数据点落在区域以外的概率不足5%，即为小概率事件，因此我们将这类数据点归为异常值点，数据落在区域以外的概率为0.04%，为极小概率事件，因此将该类数据点归为极端异常值点。从而找出并剔除异常值点和极端异常值点。  **1.2.2 插值**  插值能够根据有限个点，补插连续函数，使得这条连续曲线通过全部给定的离散数据点，在数据预处理中有重要作用。以下是各个插值方法的对比：  **1.2.2.1 拉格朗日插值**  给定n+1个插值节点，让插值多项式在这n+1个节点处的函数值等于被插函数的函数值，即构造插值多项式去逼近所有的数据。  若被插函数是n次多项式函数且n+1阶连续可导，则插值余项趋于0，插值结果很精确。但如果被插函数只有低阶导数时，不能保证插值函数能收敛到被插函数。  当n趋于无穷大时，扰动后的插值函数和没有进行扰动的插值函数距离越来越无法控制，即高次Lagrange插值的稳定性很差。  因此我们不选该种插值方法。  **1.2.2.2 牛顿插值**  因为插值多项式具有唯一性，所以Newton插值和Lagrange多项式是同一多项式的不同表现形式，但是牛顿插值比拉格朗日插值好在每增加一个插值节点，不用重新计算插值基函数，从而大大优化了算法，节省了时间。  对比拉格朗日插值，牛顿插值法当自变量很小的时候，由于差商是函数值做差以及差商做差，有效位数相减且累积以后会得到较大的损失和误差。  **1.2.2.3 Hermite插值**  Hermite插值和上述两个插值最大的区别是其不仅在每个自变量有相应的函数值约束条件，而且还有一阶导数值的约束条件。这样就能控制插值多项式的极性，而不至于和被插函数有相反的极性。  若被插函数是2n+1次多项式函数且2n+2阶连续可导，则余项趋于0，插值结果很精确。但如果被插函数只有低阶导数时，其余项的也和拉格朗日插值余项一样不一定收敛。  **1.2.2.4 分段插值和样条插值**  由于前几个插值技术在高次时数值计算稳定性不太好，一个数据的扰动可能会导致整个插值结果的大幅度改变，因此我们采取分段低次插值多项式以及在此基础上的样条插值。  在普通多项式回归中，受到全局性影响。为了解决这个问题，我们可以把数据分布分成不同的几个部分，然后针对每一部分作线性函数或非线性的低阶多项式函数，可以较好地反映每段和总体的函数法。  分段插值时采用分段三次Hermite插值，因为分段线性插值在插值节点是尖点，是不可微的，没有良好的光滑性。因此我们在插值节点给出函数值和导数值的约束条件。  三次样条指的是具有一组约束（连续性、一阶和二阶连续性）的分段多项式。样条插值在分段插值的基础上更加光滑。因为它要求在插值点处插值函数连续，插值函数一阶导连续，插值函数二阶导连续。三次分段Hermite插值函数则比较稳定，但不如样条标线的那样光滑。  **1.2.3 拟合**  **1.2.3.1 线性回归**  线性回归是最简单常见的统计模型，由于我们所要做的是时间—位置的回归，无人艇的速度不一定是匀速，是非线性的，且线性回归无法充分利用采集的信息，因此线性回归精度不好，不是合适的模型。因此我们对线性回归作一点改进，也就是多项式回归。  **1.2.3.2 多项式回归**  多项式回归是线性回归扩展到自变量和因变量之间的非线性关系的一种方法，是用多项式代替线性模型。多项式回归的缺点也不少。一方面，随着等式变得越来越复杂，函数的数量也会逐渐增加，这就导致我们很难对它们进行处理。另一方面，即便是在这么简单的一维数据集上，幂越高，曲线经过的信号点越多，形状也越诡异，这时模型已经出现过拟合倾向。而且它在本质上是非局部的，如果我们改变训练集上一个点的Y值，这会影响多项式对远处某点的拟合情况。因此，为了避免在整个数据集上使用高阶多项式，我们可以用多个不同的低阶多项式函数作为替代。  **1.2.3.3 回归样条法**  回归样条一般能比多项式回归更好。对比多项式必须要用高次多项式灵活地拟合整个数据集，回归样条在保留非线性函数的灵活性的同时，依靠节点保证了整体的稳定性。回归样条比多项式样条更为稳定，还可以根据节点的变化来调整灵活性，还可以防止过拟合。  回归样条法就是把训练集分成多个区间，在每个区间上对训练集做不同的回归模型。为了让模型既灵活到可以拟合各种形状的曲线，又可以防止过拟合，因此我们需要样条基。  记样条基为,..., ，然后得到模型：    我们把数据分布分成不同的几个部分，然后针对每一部分拟合线性函数或非线性的低阶多项式函数的方法叫做分段多项式。分段多项式每一段上的基函数都是形如的形式，根据分段的每段的函数特征去确定节点从而确定多项式的次数。节点越多，约束条件就越多，分段多项式的次数也就越高，多项式越灵活。  如何怎么选择节点呢？一种可行的方法是选择数据分布中的剧烈变化区域作为节点；第二种方法则是在数据变化复杂的地方多设置节点，在看起来更稳定的地方少设置节点，但为了简便还是会截取长度相同的区间。另外，平均分配相同样本点个数是第三种常用的方法。  如何设置约束条件呢？第一，因为是分段的，所以一个必要的约束条件就是两个多项式在节点上是连续的。第二，为了拟合函数的光滑性，在节点上增加导数值约束条件，即两个多项式的一阶导数必须相同。  值得注意的是，每增加一个约束条件，就会释放一个自由度，从而降低分段多项式拟合的复杂性。常用的回归样条法如下，增加的约束条件各不相同。  **1.2.3.3.1 三次样条**  三次样条指的是具有约束（连续性、一阶和二阶连续性）的分段多项式。  **1.2.3.3.2 自然样条**  相比三次样条，自然三次样条在边界区域增加了一个线性约束。  这是因为拟合数据分布的多项式函数在数据边界地带往往是不稳定的，边界区域的已知数据少，函数曲线常常会过拟合。增加一个线性约束后，多项式能够更平滑地扩展到边界节点之外。  **1.3 故障检测（判别分析）**  在无人艇的航行过程中，需要进行故障检测，通过判别分析判定艇的航行状态，检测艇的状态稳不稳定，是否可能翻船。将是否翻船的判断问题转化为二分类问题。将是否翻船转化成虚拟变量y，若未翻船，将y转化为0；若翻船，将y转化为1。将倾斜角、俯仰角和偏航角作为自变量。  首先收集数据集{倾斜角、俯仰角、偏航角、是否翻船}，再将数据集分为训练集和测试集，把数据集中80%的数据用来训练模型，用剩下的20%的数据进行对模型进行检验。最后，利用得到的模型，在已知{倾斜角、俯仰角、偏航角}时，对{是否翻船}进行预测。  下面是二分类问题的2种解法对比。  **1.3.1 线性概率模型（LPM）**  利用最小二乘法进行OLS回归得到，若y<0.5，认为没有翻船；反之认为翻船。  缺点：鲁棒性很差，容易受到噪点的影响，使回归模型在训练集上表现很差。这主要是由于线性回归在整个实数域内敏感度一致造成的。此外，预测值会出现y>1和y<0这种不现实的情况。  **1.3.2 Logistic回归**  逻辑回归就是一种减小预测范围，将预测值限定为[0,1]间的一种回归模型。具体的实现步骤如下：  令，代入sigmoid函数中，将h理解成y=1发生的概率。下面对参数w进行极大似然估计，得到  利用梯度下降法，，得到  其中表示学习速率，迭代至收敛即可求得。  若>0，则概率，认为不翻船。若<0,则概率，认为翻船。  **1.4 动态分析**  在动态分析中，需要分析风和浪对船的影响，并对受船本身传感器影响产生的噪声进行多维正态分布的检验以及进一步的统计推断。  **1.4.1 多维正态分布检验（重点）**  因为无人艇航行过程中受噪声影响，很多时候我们需要判断噪声是否服从多维正态分布。但是目前暂无广泛认可的多维正态分布检验方法，团队通过查阅资料，找到了从定义高维情形下的偏度峰度出发的检验方法。  将偏度峰度概念从一维情形向高维情形拓展，Mardia、Theilen、Srivastava分别定义了三种偏度峰度，即M、T、S型多维偏度峰度。基于这三种多维偏度峰度，能够进行多维正态分布的检验。除此之外，还基于S型多维正态分布检验，构造了MJB1、MJB2、MJB3三种多维JB统计量，综合考虑了偏度和峰度的影响，比简单的M、T、S更加精确。  因为多维情形下偏度峰度渐进服从特定分布（比如正态分布、卡方分布等），所以收敛速度不一，当样本数n不够大时，检验效果不够好。此时可以通过蒙特卡洛方法，求出样本数较少情形下的分位点。  从降维的角度入手也可以实现多维正态分布的检验。主成分检验法是一种效果很好的检验方法，步骤为先利用主成分分析降维，再用一维的正态分布检验方法。  **1.4.1.1 M型多维偏度检验法**  定义M型多维偏度为：    M型样本多维偏度为：    构造的统计量为：    **1.4.1.2 M型多维峰度检验法**  定义M型多维峰度为：    M型样本多维偏度为：    构造的统计量为：    **1.4.1.3 T型多维偏度和峰度检验法**  T型偏度峰度相对于M型，主要的不同在于先用协方差阵S标准化得到Y，之后的统计量构造中不包含S，即多了一个标准化步骤，思路和效果与M型差不多。具体步骤如下：  先对所得样本进行标准化：        定义T型样本多维偏度为：    T型多维偏度检验法构造的统计量为：    定义T型样本多维峰度为：    T型多维峰度检验法构造的统计量为：    **1.4.1.4 S型多维偏度和峰度检验法**  先对所得样本进行正交变换：  定义S型样本多维偏度为：    S型多维偏度检验法构造的统计量为：    定义S型样本多维偏度为：    S型多维峰度检验法构造的统计量为：    利用样本协方差阵S的特征值和特征向量，对X进行正交变换，得到的。所以组成的矩阵的行向量两两正交，最后用正交化得到的构造S型偏度峰度。  **1.4.1.5 多维JB统计量**  MJB1的统计量是自由度为p的卡方分布与标准正态分布的平方相加，由卡方分布的可加性，得到自由度为p+1的卡方分布。    MJB2是MJB1的修正，把期望和方差的实际值带入了。    MJB3是MJB2的修正，使它的方差等于自由度为p+1的卡方分布的方差，即（2p+2）。      **1.4.1.6 主成分检验法（作为备用方法，这个方法对噪声敏感度较高，可能不实用）**  除了之前讨论的多维偏度峰度检验法，目前采用比较多的方法是利用主成分分析降维，转化成一维正态分布的检验问题。  因为多维正态分布的样本协方差阵S不一定是对角阵，我们可以通过正交变换把协方差阵变成对角阵，样本矩阵X变成主成分矩阵Y，Y的每一行是一个正交变换后的样本，第j列是所有样本的第j个分量的取值。  因为Y的协方差矩阵是对角阵，所以Y满足多维正态分布当且仅当Y的每一列满足一维正态分布。之后只需要对Y的每一列用JB、AJB检验法即可。  **1.4.2 参数估计和非参数估计**  对噪声进行多维正态分布检验之后，还需要进一步的讨论：如果呈多维正态分布，参数怎么估计？如果不呈多维正态分布，呈什么分布？这就需要利用参数估计和非参数检验来解决。  **1.4.2.1** **多维正态分布下的参数估计**  根据极大似然定理，若，则估计得到的参数如下：  **1.4.2.2 非参数检验**  在实际问题中，我们不知道总体的分布，这时在进行参数估计和假设检验之前，要进行非参数检验。非参数估计主要分为两类：一类是分布函数的拟合检验，用于确定总体的分布类型；另一类是在总体分布未知时，两总体关系（分布是否相同、是否相互独立）的检验。  **1.4.2.2.1 分布函数的拟合检验**  以下主要介绍2种检验方法：皮尔逊卡方检验和科尔莫戈夫（Dn）检验。两者各有其优缺点。皮尔逊卡方检验适用的范围很广，对离散型或连续型，一维或多维，含有未知参数或不含的数据均适用。但是它依赖区间的划分，并且可能接受不真的原假设。科尔莫戈夫检验解决了这一问题，但是它只适用于连续型数据，且一般不含有未知参数。  **1.4.2.2.1.1 皮尔逊卡方检验**  设分成m个小区间，分布函数中未知参数个数为l,构造统计量：  为样本观察值落入第i个小区间的频数，为理论频数。  **1.4.2.2.1.2 科尔莫戈夫（Dn）检验**  构造统计量：  其中为样本经验分布函数，为理论分布函数。  查科尔莫戈夫分布分位表可知它的分位点。  **1.4.2.2.2 非参数检验两总体是否有相同分布**  在两样本独立的假设下，非参数检验中的相同性检验有很多种。下面介绍3种主要检验方法，分别是斯米尔诺夫检验、秩和检验、游程检验。  **1.4.2.2.2.1 斯米尔诺夫检验**  构造统计量：  其中、为两样本经验分布函数。  查斯米尔诺夫检验临界表可知它的分位点。  **1.4.2.2.2.2 秩和检验**  设是总体的样本，设是总体的样本，首先求两样本的秩。把,从小到大排序，得到，记的秩为i。如果它们有相同的样本点，每个点取平均秩（如1,4,4,5的秩为1,2.5,2.5,4）。将第一个样本的秩和相加得到，将第二个样本的秩和相加得到。  查秩和检验表即可对进行假设检验。  **1.4.2.2.2.3 游程检验**  设是总体的样本，设是总体的样本，首先求两样本的秩。把, 从小到大排序，得到，若来源于第一个样本，记为0，反之记为1。称连续出现同一个样本的样品段为1个游程，总游程数为R。  查游程总数检验表可对R进行假设检验。  **1.4.2.2.3 非参数独立性检验**  构造统计量：  表示样品落入横向第i个小区间和纵向第k个小区间构成的小区域的个数。。  **1.4.3 滤波估计**  传感器采集数据时，会受到各种各样的干扰产生异常值，导致不能够显示出真实信息。滤波的作用就是从包含干扰信息的观测值中，尽可能还原出真实的信息。对于含有不确定信息的动态系统，我们可以通过卡尔曼滤波消除噪声的影响，估计动态系统的状态。  **1.4.3.1 卡尔曼滤波**  卡尔曼滤波器就是根据上一时刻的状态，预测当前时刻的状态，将预测的状态与当前时刻的测量值进行加权，加权后的结果才认为是当前的实际状态。  卡尔曼滤波假设观测值服从多元正态分布，所以观测值的每个分量都会服从一元正态分布，均值为该分量的中心，方差体现该分量的不确定性。根据前一个状态的观测值，我们可以用协方差阵提取观测值各个分量之间的相关关系，然后用状态转移矩阵对当前状态进行预测。每个预测步骤之后，都加入一定的不确定性。  最后得到了两块正态分布，一块基于前一状态给出当前状态的概率分布P1，另一块基于当前时刻的传感器读数给出当前状态的概率分布P2。将P1和P2融合，生成一个新的高斯分布，其中新的正态分布的均值和方差是两个独立的高斯分布的相关参量的加权，这个加权就是卡尔曼增益。  下面是卡尔曼滤波的具体实现步骤：    这里的x为状态向量，通过左乘一个矩阵F，再加上外部的影响u，得到预测的状态向量x'。这里的F成为状态转移矩阵。    该公式中P为状态协方差矩阵，表示系统的不确定程度。这里的Q表示过程噪声，工程上，我们一般将Q设置为单位矩阵参与运算。    这个公式计算的是实际观测到的测量值z与预测值x'之间差值y。将状态向量左乘一个矩阵H，才能与测量值进行相应的运算，这个H被称为测量矩阵。      这两个公式求的是卡尔曼增益K，即y值的权值。第一个公式中的R是测量噪声矩阵，表示测量值与真值之间的差值。一般情况下，传感器的厂家会提供该值。      第一个公式完成了当前状态向量x的更新，不仅考虑了上一时刻的预测值，而且考虑了当前时刻的测量值，和整个系统的噪声，第二个公式更新了系统的不确定度P，用于下一个周期的运算，其中I为与状态向量同维度的单位矩阵。  **2.可行性分析**  **2.1 技术可行性分析**  同济大学葛泉波教授提供技术上的支持，通过在小型无人船上安装传感器、做小型仿真实验等为团队提供必要的数据来源。无人巡航艇的位姿状态受到海上的风、流影响，如何准确地测量无人艇航行过程中的各类数据，这与自动化专业的知识密切相关。葛泉波教授的2名研究生薛子建、程惠茹同学的高度参与，为团队提供了可靠的专业背景，使数据分析更加科学合理。  另外，为实现无人巡航艇的位姿状态分析，团队利用多种统计软件进行数据处理，并且尝试经过调研寻找更加适用本问题的统计方法。  以判断无人艇航行过程中的噪声是否呈多维正态分布为例，虽然目前多维正态分布检验还没有广泛认可的方法，但是通过查阅相关资料，团队解决了多维正态分布检验的软件实现问题。  通过STATA实现M型偏度峰度检验法、Doornik-Hansen检验，Henze-Zirkler检验。  下载工具包后，通过SPSS实现7种检验方法，包括M型偏度峰度检验法、S型偏度峰度检验法、Doornik-Hansen检验，Henze-Zirkler检验，Royston检验，Shapiro-Wilk检验，Small检验。此外，也为多元正态性提供了卡方与马氏距离的绘图。  通过MATLAB求主成分矩阵，再利用主成分检验方法，将多维正态分布的检验问题转化为多个一维正态分布的检验问题。  在充分了解多维情形下偏度、峰度的三种定义下，编程实现Mardia、Theilen、Srivastava三种偏度峰度检验法，并且利用1型多维JB统计量MJB1、2型多维JB统计量MJB2、3型多维JB统计量MJB3综合多维情形下偏度、峰度检验法，得到了更好的结果。  **2.2 组织可行性**  团队成员学习成绩优异，对统计知识有浓厚的兴趣。团队成员不仅在所学专业中名列前茅，而且在完成课业的基础上，旁听研究生的课程《滤波估计》、《凸优化》等。  团队成员具备扎实的理论知识，能够利用所学解决问题。团队成员来自统计专业、统计（大数据）专业、数学基地班以及数学大类，在具备概率论、数理统计、多元统计、泛函分析等必备知识的同时，能够优势互补，互相启发。  团队成员掌握必要的软件实现技能，能够熟练运用常用软件STATA、MATLAB、SPSS实现数据处理。  团队成员有科研训练的经验。团队成员目前参与了3个小火花项目和3个大创项目。其中，有2个小火花项目已结题，分别为《护士法立法公众意向分析——10万份问卷调查的大数据分析》和《key多项式的组合研究》。 |

|  |
| --- |
| **三、研究基础（对项目的参与动机、已有知识储备、相关研究和训练基础）** |
| **1.项目参与动机**  **1.1 巩固学科知识，提升数据处理、建模分析的能力。**  我们相信这个项目的完成过程将充分夯实我们在相关课程中学习到的各类知识。同时在查找文献、搜集数据和建模分析等过程中，我们希望可以锻炼到大家在科研过程中解决各类问题的能力。  **1.2 将所学知识应用到实际，培养科研相关能力。**  当前我们虽然已经学习了概率论、数理统计、数学建模等专业课程，但是实际应用的直观感受能力仍有待提高。我们希望通过思考海上的风、浪和流对海上船舶的航行状态所能产生的影响，尝试理解这些影响的产生过程、机理,并分析对这个机理的建模过程中存在哪些不确定性、误差，进而提升科研能力。  **1.3 希望为相关业界做出实际贡献。**  如果结果足够完美，我们希望通过本项目向无人驾驶有关方向做出切实贡献。力争以新知识、新成果求新突破、新成绩。  **2.已有知识储备**  小组全体成员已修读高等代数、概率论、数理统计、数学建模与实验等课程，部分成员在修多元统计、回归分析、数值分析等统计学专业课程，同时项目组负责人叶子怡在旁听、自修最优化、滤波分析等课程，这使得我们在本项目中所能应用到的知识体系更加完备。  **3.相关研究**  在四川大学数学学院“小火花”科研训练中，我项目组成员参与过《护士法立法公众意向分析——10万份问卷调查的大数据分析》、《基于机器学习的垃圾短信自动识别》这两项相关研究。带着以往在研究工作中的经验，我们相信我们能够在本项目中做到求真务实、提高质量、紧跟进度、协调推进。  **4.训练基础**  本项目组成员掌握STATA、MATLAB、SPSS等软件的运用，同时也有java、Python等编程技能。有了这些基础，相信我们将能够做到：注重质量，积极推进建模工作优质化；各司其职，积极推进编程工作有序化。 |
| **四、研究计划和进度（就文献查询、社会调查、方案设计、实验研究、数据处理、研制开发、撰写论文或研究报告、结题和答辩、成果推广、论文发表、专利申请等工作逐项计划时间，时间节点精确到月份）** |
| 1. **筹备期**   在筹备期我们需要让小组成员对我们的这一项目有充分的认识，进而对小组进行的任务分配、组织大家查找相关资料、文献，以及填写申报书。   1. **前期**   组织大家进行多次讨论，头脑风暴、集思广益，从课堂所学以及课外了解到的信息出发，争取建立良好、有效的数学模型来解决问题。力争找到一种适用于船舶航行中噪声的多维正态分布检验，并找到其特征。   1. **中期**   利用计算机软件处理数据、分析数据，在这一过程中我们也将不断改进、优化模型，使它更能贴近实况。   1. **后期**   对于分析情况进行最终的思考、讨论，将误差降至最小。   |  |  | | --- | --- | | **研究计划和进度表** | | | 计划时间 | 计划进度 | | 2021.9——2021.10 | 对小组进行初步的任务分配、组织大家查找相关资料、文献，以及填写申报书。 | | 2021.10——2022.3 | 小组根据所收集到的信息，从所学知识出发对情况进行分析，建立合适的数学模型。 | | 2022.3——2022.5 | 通过统计学、计算机方法进行数据处理与分析。 | | 2022.5——2022.7 | 对分析结果进行讨论，进而得出结论。 | | 2022.7——2022.10 | 撰写论文、准备答辩。 | |

|  |  |
| --- | --- |
| **五、项目研究支撑条件（项目所依托的重点实验室（中心、平台）、双创平台、课题组等各类单位能提供的直接支持项目开展的软硬件设施和其他校内外资源）** | |
| 本项目由千岛湖科学研究院提供设备、技术以及测试实验上的支持，包括提供无人艇以及无人艇控制系统平台的开发，并且可根据项目的实际需求进行测试仪器的加装并收集一系列的实验数据以供分析。研究院有专家导师以及课题组团队对项目研究内容进行指导，在多次调查、研讨后确定了本项目能解决的实际问题，并且该问题的解决有一定的重要意义。因此本项目所需的运营基本条件都已具备，能确保科研训练的顺利进行。 | |
|  |  |
| **六、预期成果形式（可多选）** | |
| （1）□SCI论文 篇  （2）□核心期刊论文 篇  （3）□会议论文 篇  （4）□内部编印期刊论文 篇  （5）□授权发明专利 项  （6）□申请发明专利 项  （7）□创新类竞赛获奖  （8）□创业类竞赛获奖  （9）□其他 名称： | |
|  |  |
| **七、项目经费概算（按申报项目目标任务需要进行预算，经费执行情况将与结题考核成绩挂钩）** | |
| **1.经费来源（单位：元）**  申请项目专项经费  **2.经费支出（单位：元）**  （1）仪器设备费  （2）耗材费  （3）测试加工费  （4）国内会务及差旅费  （5）国外会务及差旅费  （6）文献/知识产权事务费  （7）办公费（含文印、办公用品等）  （8）其他费用  **合计** | |

|  |
| --- |
| **八、评审情况** |
| **指导教师意见：** |
|  |
| **指导教师（签名）： 年 月 日** |
| **学院推荐意见：** |
|  |
| **主管院长签名： 年 月 日** |
| **学校专家评审意见：** |
|  |
| **组长签名： 年 月 日** |
| **学校认定意见及批准经费：** |
|  |
| **学校负责人签名： 年 月 日** |