1.研究背景及意义

锂离子电池具备高能量密度、循环寿命长的优点，是目前储能技术研发重点之一。为遏制环境污染、土地沙漠化、能源紧缺等问题，国家大力发展新型能源，大力挖掘绿色能源的开发，大力支持能源汽车的生产，以尽快实现“碳达峰、碳中和”的目标，促进经济转型，促进社会可持续发展。“十三五”指出为迎合国际上环保理念的趋势，坚持可持续发展理念，需要进一步创新发展电动汽车，推进我国汽车产业转型升级。电池的SOH和RUL是锂离子电池的重要参数，也是锂离子电池的重要研究方向，能准确辨识电池的SOH和RUL对新能源汽车的安全性和操作性具有重要意义[6-8]。由于锂离子电池的组成结构，目前无法利用传感技术直接得到电池实时的SOH，因此开发电池的SOH估计和RUL预测的模型有助于推动新能源产业的发展。锂离子电池具有使用寿命长、能量密度高、低自放电率、无记忆效应以及宽温度范围等优点，被广泛用做各类机器的储能部件，如：新能源汽车、移动电话、笔记本电脑等设备。

与镍氢等其它储能设备相比，锂离子电池具有容量大、循环充放电寿命长、无记忆效应、内阻小等优点，被广泛应用于航天、新能源汽车、海上轮船等大型设备。

锂离子电池是一个时变的电化学系统，工作过程中内部反应复杂多变，在动态系统下处理各种物质和能量，会产生固体电解质界面膜增长、锂离子沉降、集流体腐蚀、隔膜损伤和电解液氧化等副反应。这些副反应阻碍了电池正负极的嵌锂和脱锂过程，导致电池的性能衰减，宏观上表现为容量减少和内阻增加，最终降低了电池的使用寿命。在长时间使用过程中，因外部环境和内部电极材料腐蚀、隔膜老化等原因，锂离子电池性能会逐渐退化和失效。电池的突然失效容易导致电子设备失去能量来源而停止工作，进而造成事故。锂离子电池的可靠性、安全性和寿命已成为制约其成功应用的至关重要问题，因此如何准确地预测锂离子电池的SOH及电池的剩余使用寿命是目前锂电池研究亟待解决的问题和难题。

锂离子电池的老化失效会导致设备的停电和断电，进而导致事故的发生。因此及时、准确地预测电池的SOH及RUL有利于及时规划和管理电池，有利于及时更换失效电池，提前预防事故的发生。由于锂电池内部的化学反应以及外部环境的影响，电池的使用寿命随着时间的推移逐渐衰退老化甚至失效。锂电池的衰弱退化无疑增加了许多电子设备的维修成本，而且电池的突然失效易导致新能源汽车等大型设备失去工作，甚至导致重大事故的发生。

综上所述，准确的电池寿命预测对于提高能源利用率、保护环境、降低成本、增强可靠性以及推动电池技术的进步都具有重要意义。它对于改善各个领域的设备和系统性能，以及推动可持续发展和资源管理都具有深远的社会和经济价值。因此，电池寿命预测研究的重要性在现代社会愈发显著。

2.研究内容

研究带有噪声的电池寿命预测相对于没有噪声的方法具有显著优势：这种研究更接近实际应用场景，因为在现实生活中，电池性能受到各种噪声源的干扰，如温度变化、充电和放电速率的波动等。因此，噪声预测更准确地反映了电池在复杂环境下的真实运行情况。带有噪声的电池寿命预测能够提前发现电池可能出现问题的征兆，因为噪声通常是由电池性能的微小变化引起的。这有助于及早采取维护措施，避免电池失效或损坏，从而延长其寿命。另外，通过噪声分析，可以更好地了解电池性能的波动性，为资源的更有效利用提供了机会。这有助于降低不必要的电池更换和维护成本，提高能源系统的可维护性和性能。

2.1 数据集噪声相同

在目前提供的公共数据集包括美国宇航局埃姆斯卓越预测中心的NASA数据集和马里兰大学高级生命周期工程中心提供的CALCE数据集，都是来源于同一个电池批次，因此，研究数据集的噪声是同一个分布的情况是将电池寿命的预测模型推广到现实的必不可少的步骤。

2.2 不同数据集噪声不同

在提高模型的鲁棒性，将其推广至更广的数据集范围时会考虑到，不同的数据集来源不同导致数据集的噪声分布状况不同，此时，需要解决不同的数据集噪声所服从的分布不一致的条件下如何对该数据集进行处理能够降低噪声对数据集的影响得到更精准的预测精度。

3.研究方法

3.1 降噪算法

将输入特征进行降噪分解。现有的常用输入特征来源于，电池循环的充放电电流电压曲线、容量增量曲线、电池充放电时的内阻、充放电温度变化数据及这些数据的统计特征。现有常用的信号降噪方法包括：小波变换、卡尔曼滤波、模态分解、去噪编码器等。我们根据每种算法的特点和使用场景进行对比分析，选出最合适的降噪算法。

3.2 预测算法

现有的针对电池寿命预测的模型数不胜数，不同的机器学习模型适应不同的健康因子，在面对有噪声的电池数据集的时候，如何选择合适的人工智能方法，并选用合适的处理方式对其进行优化是本研究的主要难点。

3.2 TLS和OLS算法

TLS算法：

针对有误差的数据矩阵和数据向量，需满足, 特征矩阵 , 标签向量 , 噪声 and , 模型样本是bn维向量 。

总体最小二乘算法即在满足的前提下用约束优化问题叙述为:

（3-1）

可以将上述目标优化函数转化成以下两步，即:

（3-2）

构建拉格朗日函数：

, （3-3）

利用KKT条件，将问题转化：

(3-4)

(3-5)

(3-6)

根据 (3-4) 和 (3-5)可以得到：

, , (3-7)

根据(3-6) 和 (3-7)可以得到:

, , (3-8)

根据(3-6) 和 (3-8)可以得到:

(3-9)

将（3-8）和（3-9）带入原始目标函数得到：

（3-10）

则原始的问题： （3-1）可以转化为：

（3-11）

解决方法：

式（3-11）利用homogenization argument，总体最小二乘问题将等价地转换为：

-> （3-12）

其中，

为解决问题我们首先忽略约束条件使 , 那么问题转化为问题

（3-13）

我们知道z的大小不会对结果造成任何影响, 因此我们设置z的大小 ，则原始问题转化为：

（3-14）

经过推导, →,将解决的问题转化为：

（3-15）

根据KKT条件:

（3-16）

其中 是的特征向量，我们可以得到，问题中 是 的最小特征向量。

问题的解 是的最小特征向量所对应的最小特征值：

（3-17）

OLS算法：

最小二乘估计： （3-18）

（3-19）

（3-20）

但数据都带有一定噪声，假设噪声满足零均值的高斯分布

因此， （3-21）

则,（3-22）

使用MLE求解W：

首先是最大似然函数：

（3-23）

再求W：

（3-24）

传统的TLS和LS算法优于本身特性能够对带有噪声的电池数据集进行预测，但并不能应对不同数据集的噪声分布不同这种情况，因此需要对传统的OLS和TLS进行改进。

3.3 EM算法

给定数据集，假设样本间相互独立，我们想要拟合模型到数据的参数。根据分布得到似然函数：（3-25）

我们可以最大化似然函数以求解，但当存在隐变量时，似然函数表示为：

（3-26）

对于某个样本i，我们使用表示样本i隐含变量z的分布概率，则

（3-27）

引入，并利用Jensen不等式：

（（3-28））

似然函数变换为：

（3-29）

我们得到了的形式，

不等式取等号：当为常数时，不等式可以取等号，对等式变换并两边求和

,,

（3-30）

E-Step：求条件期望，本质是求出来自哪种类别的概率，为简化将log中的去掉，（3-31）

M-Step: 求（3-32），更新参数

在研究不同数据集的噪声分布不同时，由于所训练的模型系数和噪声标准差都是未知，即同时存在两个位置变量，EM算法求解隐变量的思想符合本研究的解决思路。

4.预期目标

（1）使用合适的降噪算法对电池数据集进行去噪效果明显，比其他降噪方法有更高的预测精度。

（2）使用特征工程对分解得到的分量进行处理之后的预测结果比未使用结果有提升。

（3）使用TLS\_EM和OLS\_EM这两种方法相较于传统的TLS和LS有更高的预测精度，且在噪声增强的情况下比传统算法优势明显。

5.难点及存在问题

（1）降噪方法的选择。降噪算法种类数不胜数，如何选择合适算法是本研究的难点。电池预测中选择合适的降噪方法是一个难点，主要因为电池系统的复杂性和不确定性。首先，电池的性能受多种因素影响，包但不仅限于充放电速率、温度、老化程度等。这导致电池系统产生的数据具有高度非线性和时变性质，难以简单建模。其次，电池预测受到外部环境和使用条件的影响，如温度变化、负载波动等，这些因素引入了随机性和不确定性。传统的降噪方法往往难以捕捉到这些复杂的非确定性变化，从而影响预测的准确性。此外，电池系统中可能存在噪声干扰，如传感器误差、测量噪声等，这些噪声源对预测模型的训练和性能评估都构成挑战。选择适当的降噪方法需要兼顾对系统复杂性、外部环境变化和噪声的适应性，以确保预测模型在真实工作条件下具有稳健性和准确性。因此，电池预测中的降噪问题需要综合考虑多方面因素，这增加了该任务的难度。

（2）预测模型的改进。传统的机器学习方法所适应的输入特征并不一定能够适应经降噪分解之后的已经改变的分量，故需要对其进行改进处理。如何使用合适的处理手段也是本研究的难点之一。

（3）模型方法的健壮性。本研究中出现的电池预测健康因子是电池地100次循环和第150次循环的电压-放电曲线面积差，目前公开的数据集中只有一个数据集包含本研究所需要的健康因子，其他公开的电池数据集中，马里兰大学的数据集和美国航天局的数据集并没有包含所需信息，故模型的健壮性有待扩展。

6.工作的初步计划

|  |  |
| --- | --- |
| **时间阶段** | **具体工作** |
| 2024.01~2023.02 | 查阅资料，阅读文献，复现代码。 |
| 2024.03~2023.05 | 实现所提出的方法，并选取部分数据集进行实验，验证方法的性能。 |
| 2023.06~2023.10 | 根据实验结果调整方法的细节，并进行大规模数据集的实验。 |
| 2024.10~2024.11 | 与现有方法进行对比，进一步优化算法。 |
| 2024.12~2025 | 撰写毕业论文，准备毕业答辩。 |