

基于充电电压片段和核岭回归的 锂离子电池 SOH 估计

樊亚翔, 肖飞*, 许杰, 杨国润, 唐欣

(舰船综合电力技术国防科技重点实验室(海军工程大学), 湖北省 武汉市 430033)

State of Health Estimation of Lithium-ion Batteries Based on the Partial Charging Voltage Segment and Kernel Ridge Regression

FAN Yaxiang, XIAO Fei*, XU Jie, YANG Guorun, TANG Xin

(National Key Laboratory of Science and Technology on Vessel Integrated Power System (Naval University of Engineering),
Wuhan 430033, Hubei Province, China)

ABSTRACT: The online estimation of battery state-of-health (SOH) is an ever significant issue for the battery management system. Recently, due to its advantages such as model-free and flexibility, data-driven based methods are promising for online SOH estimation. Aiming at the problems of heavy computing burden and difficulty in implementing for microcontroller of the existing battery SOH estimating methods, a novel estimation approach based on the partial charging voltage segment and kernel ridge regression (KRR) for the SOH of lithium-ion batteries was proposed. KRR combines ridge regression with the kernel trick, which thus learns a non-linear function between the partial charging voltage segment and lithium-ion batteries SOH by the respective kernel and the data. The experimental results on two lithium-ion battery degradation datasets show that the proposed method can achieve fast and accurate SOH estimation, which can be further applied to existing BMS only by adopting partial charging voltage curve segments that could be easily obtained in actual working condition.

KEY WORDS: lithium-ion battery; state of health (SOH); partial charging voltage segment; kernel ridge regression (KRR); data-driven method

摘要: 电池健康状况的在线估计对于电池管理系统一直是一个非常重要的问题。近年来, 由于其具有灵活性和无模型优势, 基于数据驱动的方法在在线健康状态(state of health, SOH)估计领域展现出极大的潜力。文中针对现有的大部分

基于数据驱动的 SOH 估计方法存在计算量大以及较难在 BMS 微控制器中实现等问题, 提出一种采用片段充电曲线和核岭回归(kernel ridge regression, KRR)的锂离子电池 SOH 估计方法。KRR 是一种基于核方法的非线性回归算法, 通过将核技巧与岭回归结合, 能够建立充电电压片段和 SOH 之间的非线性映射关系。在 2 个公开锂离子电池老化数据集上的实验表明, 该方法只需采用实际电池使用工况中容易获得的充电电压片段, 就能够实现快速准确的 SOH 估计, 并且应用到现有的 BMS 微控制器中。

关键词: 锂离子电池; 电池健康状态; 片段充电电压数据; 核岭回归; 数据驱动方法

0 引言

锂离子电池因其成本低、能量密度高和使用寿命长等优点而被广泛应用。但是, 与许多化学、机械和电子系统一样, 锂离子电池的性能会因其电化学成分而降解而随时间和使用时间而下降, 从而导致容量和功率衰减。这种机制往往受不同因素(例如电池化学成分和制造及环境和工作条件)的影响, 是多种耦合老化机制的结果。为了确保电池老化后工作的安全性和可靠性, 需要对其健康状态进行评估。其中, 电池的健康状态(state of health, SOH)反映了电池相对于其寿命开始时的当前存储和供应能量/功率的能力, 该能力可以通过实际电池容量/内阻与其初始值之比计算获得。对于电池可用容量起着至关重要作用的应用, 例如电动汽车、储能电站、便携电子产品电源等, 通常将容量用于 SOH 表征。因此, 本文从容量的角度定义 SOH, 将 SOH 定义为实际电池容量与其初始值之比, 当 SOH 衰

基金项目: 国家自然科学基金项目(52007196, 51807199); 国防科技创新特区项目。

Project Supported by National Natural Science Foundation of China (52007196, 51807199); Project Supported by the National Defense Science Innovation Zone Project.

减到 80% 以下时则判定电池已达到最大使用寿命, 需要进行更换。

由于电池的 SOH 无法直接测出, 为了获取电池的 SOH, 大量的 SOH 估计方法已被提出。其中, 较为常见的是基于模型的方法, 通过设计等效电路模型来模拟电池内部的工作机理, 然后采用各种优化算法和观测器, 例如卡尔曼滤波^[1-2]、粒子滤波^[3]、滑模滤波器^[4]等以辨识模型参数和 SOH 状态。其中, 一种广泛使用的方法是使用电化学模型, 应用偏微分方程来模拟与老化密切相关的质量和电荷转移动力学^[5]。这类方法可以解决在线参数与状态同时估计问题, 但是依赖于建模精度, 而且观测器的精确性、稳定性与收敛性受初值的影响比较大。

由于其具有灵活性和无模型优势, 基于数据驱动的方法已经在学术界和工业界受到越来越广泛的关注^[6-7]。此类通过机器学习技术在离线阶段建立可测量电气量(或者电气特征量)同电池 SOH 之间的非线性映射关系, 然后在在线阶段通过新观测的电气量和训练完的机器学习模型即可计算获得电池 SOH。这类方法避免了基于模型的方法中的建模和参数辨识问题, 是一种实用性较强的电池 SOH 估计方法。对于这类方法而言, 特征输入和机器学习算法的选择是影响方法性能和实用的关键。文献[8-9]提出采用循环次数作为特征训练高斯过程回归(Gaussian process regression, GPR)以估计 SOH, 不仅取得了较高的估计精度而且能够对估计结果的不确定性进行量化。文献[10]提出从增量容量曲线中提取二阶高性能特征, 并且通过灰度相关分析来建立特征与 SOH 自己的映射关系以预测 SOH。通常, 锂离子电池充电过程中的电气量由于恒流恒压充电过程非常稳定, 已被广泛使用作为特征。文献[11]提出采用充电过程恒流充电时间、恒流充电容量、恒压充电时间、恒流充电容量 4 个值作为特征训练支持向量回归(support vector regression, SVR)模型以预测 SOH, 并在 NASA 电池老化数据集上验证了算法的精度。最近, 随着深度学习技术的飞速发展, 研究人员也开始把深度学习技术应用于锂电池 SOH 估计中。文献[12]提出将充电阶段的电压、电流及充电容量转换为大小为 25×3 的矩阵, 然后训练卷积神经网络(convolutional neural networks, CNN)以预测 SOH。文献[13]提出一种深度神经网络的方法, 通过循环次数预测 SOH 和剩余寿命, 取得了非常好的效果。从充放电曲线

中提取了充电截止电压、放电截止电压、恒流充电时间、恒压充电时间等 8 个特征量, 文献[14]等人训练了一个 3 层的前向传播神经网络, 并结合马尔科夫链以模拟电池老化过程的长时间响应, 实现 SOH 估计。

受以上启发, 本研究团队在之前的工作中也提出直接将完整的充电曲线作为特征, 并且设计了一种 CNN 模型^[15-16]估计锂离子电池 SOH。方法直接对传感器测量的电压、电流、温度进行处理, 非常直观, 在 2 个公开数据集上的实验结果验证了算法的优越性。但是, 当把这些算法部署于以微控制器为主的 BMS 中, 遇到了 2 个问题: 1) 这些方法把完整的充电数据(即 SOC 从 0%~100% 的充电数据)直接作为输入或者提取特征作为输入以估计 SOH, 但是在电池实际使用的过程中, 受工况影响充电过程的初始 SOC 状态和结束 SOC 状态很难固定下来的, 所获得的充电数据往往是一个片段(类似 SOC 从 20% 到 80% 或者 SOC 从 50% 到 100%), 因此需要针对实际的片段充电数据设计 SOH 估计算法; 2) 之前提出的 SOH 估计方法都是在 PC 平台上进行实现的, 对于计算资源的消耗很大, 尤其是基于深度学习的算法, 因此将方法迁移到面向移动和嵌入式的电池管理系统中比较困难, 寻求计算代价小的 SOH 估计方法是硬件实现的关键。

鉴于此, 本文从锂离子电池实际工作情况以及常用电池管理系统(battery management system, BMS)微控制器的计算能力 2 个出发点, 针对锂离子电池片段充电电压数据, 引入一种在线计算量小的机器学习方法——核岭回归(kernel ridge regression, KRR), 提出一种新的 SOH 估计方法。根据对公开的锂离子电池老化数据分析可知, 锂离子电池恒流-恒压充电过程具有固定形式且只有部分时间段内 dQ/dV 变化比较激烈, 因此可以选择此时间片段的电压估计 SOH。充电电压片段均是随时间变化的一维时序数据且均对应一个唯一的 SOH 估算值。而 KRR 方法是一种改进的岭回归模型, 利用核函数将原始数据映射到高维空间, 映射后的数据在高维空间往往呈线性关系, 具有较强的泛化性能。KRR 不仅能够挖掘准确出充电电压片段与 SOH 之间的映射关系, 而且在测试阶段计算量较小, 能够在以微控制器为主的 BMS 系统上实现。在 2 个公开数据集上进行实验结果表明, 本文提出的方法精度高, 泛化能力强。

1 片段充电电压数据分析

对于锂离子电池 SOH 估计来说, 特征选择是关键, 它不仅影响了算法的计算量, 还能够决定 SOH 估计的精度。对于 SOH 估计问题, 常见的特征输入^[6]主要可分为模型依赖特征、直接测量特征以及二阶处理特征。其中, 模型依赖特征主要是指欧姆阻抗、极化阻抗以及极化容量等, 这些特征依赖于复杂模型不适应于实时应用, 目前很少使用。直接测量特征往往就是直接采用 BMS 中的传感器记录的循环次数、电压、电流、温度等; 而二阶处理特征则指的是通过直接测量特征进行简单计算获得二阶特征, 例如 IC/DV 曲线^[17]、 dV/dt 曲线^[18]等。

实际上, 电压、电流、温度是任何 BMS 易于获取的观测量。图 1 给出了不同 SOH 下充电电压

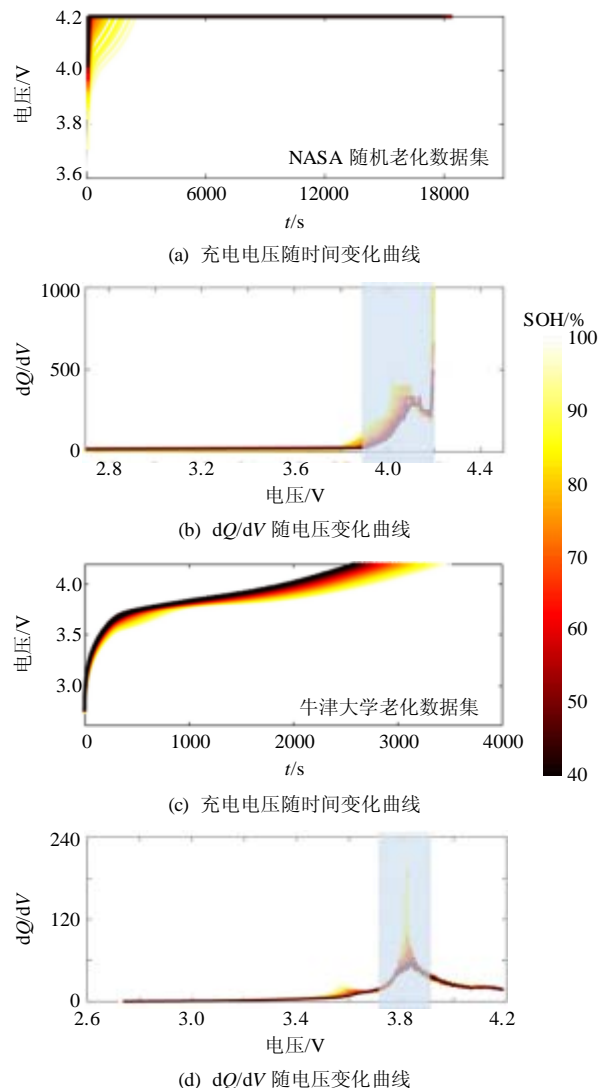


图 1 不同 SOH 下充电电压随时间变化和增量容量曲线

Fig. 1 Charge voltage vs. time and dQ/dV vs. voltage curve under different SOH

充电电压随时间变化和 dQ/dV 随电压变化曲线, 其中电池寿命衰减由白色→黄色→红色→黑色的渐变色表示, 图 1(a)、(c)中两条曲线代表两块电池不同老化程度时电池的恒流恒压充电电压变化, 而图 1(b)、(d)则是两块电池对应的增量容量 (incremental capacity curve, IC) 曲线。可以发现, 充电过程中电池电压曲线因充电方式为固定的恒流-恒压方式而相对稳定, 且随着锂离子电池的老化存在规律性的变化。具体来说, 从图 1(a)、(c)中可以发现, 当电池使用时间短, SOH 较大时(曲线黄色部分), 充电电压上升至截止电压较缓慢, 充电电流保持恒定电流的时间较长, 这符合电池越新则内部化学反应时间越长、容量越大的实质。而当电池使用过一段时间, SOH 减小后(曲线红黑色部分), 则呈现出完全相反的现象, 充电电压迅速上升至截止电压。这也是电压曲线可以作为直接测量特征进行 SOH 估计的原因。

实际上, 电池工作时受放电工况影响充电曲线很难是完整的(即 SOC 为 0%~100%), 所获得的充电数据往往是一个片段(类似 SOC 从 20%到 80%或者 SOC 从 50%到 100%)。而且, 通过观察电池的增量容量曲线(图 1(b)、(d))可以发现, 两块电池只有小部分电压区域内 dQ/dV 变化比较激烈(蓝色阴影部分), 而大部分电压区域内的 dQ/dV 变化比较缓慢。进一步, 通过观测不同 SOH 下的增量容量曲线可以发现, 当 SOH 越大, 相同电压时 dQ/dV 变化愈发激烈(增量容量曲线呈黄色)。反之, 当电池老化, SOH 较小时, 相同电压时 dQ/dV 变化平缓(增量容量曲线呈黑色)。

因此, 可以看出该电压片段与 SOH 存在非常明显的规律, 可以选择这段电压片段作为估计 SOH 的直接输入^[19]。实际上, 选取合适的充电电压片段就存在 2 个参数需要设定, 第 1 个是初始电压 V_0 , 第 2 个是输入电压片段数据的长度 N (即从 V_0 开始总共选取的数据点个数)。其中, 第 2 个参数直接影响到在线计算效率, 越小越好。而且, 电压片段的选取直接决定了该次充放电循环工况是否能够估计出 SOH。因此, 在 3.4 节中重点讨论了 V_0 和 N 的选取。

2 基于片段充电电压和 KRR 的锂电池 SOH 估计方法

针对 1 节提到的片段充电电压数据, 本文提出

了一种基于核岭回归(kernel ridge regression, KRR)的锂离子电池 SOH 估计方法。核岭回归方法是一种改进的岭回归模型,利用核函数将原始数据映射到高维空间,映射后的数据在高维空间往往呈线性关系,具有较强的泛化性能。

给定训练样本 $D = \{(\mathbf{x}_1, y_1), (\mathbf{x}_2, y_2), \dots, (\mathbf{x}_m, y_m)\}$, 其中 $\mathbf{x}_i \in \mathbf{R}^N$ 为片段充电电压数据, y_i 为 SOH, N 为片段充电电压数据维度, m 为片段充电电压数据维度, 期望学得一个线性回归模型, 使得:

$$y \approx \mathbf{w}^T \mathbf{x} \quad (1)$$

式中 \mathbf{w} 为待确定的模型参数。则可以转换为一个最小二乘问题:

$$\arg \min_{\mathbf{w}} \|\mathbf{w}^T \mathbf{x} - y\|_2^2 \quad (2)$$

但是对于简单的线性回归模型, 其对输入 \mathbf{x} 的噪声非常敏感, 当输入 \mathbf{x} 有微小的误差, 输出变量 y 就会变化很大。一个简单的解决方法就是采用线性岭回归(linear ridge regression, LRR), 限制 \mathbf{w} 的大小, 给式(2)加上一个 L2 范数正则化约束, 即为

$$\arg \min_{\mathbf{w}} \{\|\mathbf{w}^T \mathbf{x} - y\|_2^2 + \lambda \|\mathbf{w}\|_2^2\} \quad (3)$$

式中 $\lambda > 0$ 为平衡损失和正则项之间的系数。实际上, 线性岭回归可以看成是一种使用 L2 范数正则化的线性最小二乘法, 其具有较好的稳定性和泛化能力。但是, 其主要是针对线性问题而设计的, 无法有效解决自变量之间的非线性情况。

为了解决这个问题, 一种常见的方法是使用非线性映射将样本映射到高维空间, 然后在高维空间中学习模型。其中, 核函数是一种广泛使用的方法^[20], 它通过定义内核函数隐性地学习样本在高维空间中的相似性。把核函数和线性岭回归结合即是核岭回归(kernel ridge regression, KRR), 即可表示为:

$$\arg \min_{\mathbf{w}} \left\{ \frac{1}{n} \sum_{i=1}^n \|f_i - y_i\|_2^2 + \lambda \|f\|_H^2 \right\} \quad (4)$$

$$f_i = \sum_{j=1}^n \alpha_j \kappa(\mathbf{x}_j, \mathbf{x}_i) \quad (5)$$

式中: $\|\cdot\|_H$ 是一个希尔伯特空间范数; $\kappa(\cdot, \cdot)$ 表示核函数。给定一个 $n \times n$ 的核矩阵 K , 则该问题则可以转换为一个线性系统如下:

$$(K + \lambda n I) \alpha = y \quad (6)$$

$$\tilde{y} = \sum_{i=1}^n \alpha_i \kappa(\mathbf{x}_i, \tilde{\mathbf{x}}) \quad (7)$$

式中: K 为由训练数据核函数组成的核矩阵 $K_{ij} = \kappa(\mathbf{x}_i, \mathbf{x}_j)$; α 为未知向量。在本文中, 选择常见的高斯核作为核函数, 即为

$$\kappa(\mathbf{x}_i, \mathbf{x}_j) = \exp\left(-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|^2}{2\sigma^2}\right) \quad (8)$$

式中 σ 为高斯核的带宽, 控制着函数的平滑度、噪声容忍度和泛化程度, 通常设置为 $\sigma = 0.1$ 。在训练阶段, 算法的目标数通过式(6)求解获得 α 。在测试阶段, 给定新观测数据 $\tilde{\mathbf{x}}$, 算法通过式(7)来对 SOH 进行预测。表 1 给出了核岭回归的整个算法流程。

表 1 基于核岭回归的 SOH 估计方法流程

Table 1 The SOH estimation method flow based on kernel ridge regression

序号	步骤
输入	已知充电电压片段——SOH 数据对 (\mathbf{x}_i, y_i) 用于训练; 另外 $(\tilde{\mathbf{x}}_j, \tilde{y}_j)$ 用于测试, 需要求取参数 α , 其中 $i \in \{1, 2, \dots, m\}$, $j \in \{1, 2, \dots, n\}$
输出	预测 SOH 结果 $\{\tilde{y}_j\}_{j=1}^n$
1	选择核函数, 构造一个大小 $n \times n$ 的核函数矩阵 K
2	for $i \in 1:m$ 进行
3	for $j \in 1:n$ 进行
4	计算数矩阵 K 中的各元素, $K[i][j] \leftarrow \kappa(\mathbf{x}_i, \mathbf{x}_j)$
5	通过式(5)中 $(K + \lambda n I) \alpha = y$ 求取 α
6	for $i \in 1:m$ 进行 $\tilde{y}_j \leftarrow \sum_{i=1}^n \alpha_i \kappa(\mathbf{x}_i, \tilde{\mathbf{x}}_j)$

3 实验与分析

为了验证本文所提出的 SOH 估计方法的效果, 本节基于 2 个锂离子电池实测数据集进行实验。另外, 本节还对输入电压片段区间选择进行了探讨, 以验证所提出估计方法的适用范围。最后, 对该算法的计算复杂度进行了说明。

3.1 锂离子电池老化数据集介绍及评价指标设置

本文采用 NASA 锂电池随机使用数据集^[22]和牛津电池老化数据集^[27]进行实验。以上 2 个锂离子电池数据集除了记录锂电池使用过程中的充电电压数据, 并且每隔一段循环后通过放电阶段安时积分的方法对电池容量进行测定以获取实际 SOH, 因此适用于本节实验。

NASA 锂离子电池随机数据集所使用的电池为 LG Chem 18650 圆柱形电池, 其额定容量为 2.1A·h, 常规工作电压范围为 3.2~4.2V。该数据集共对 28 块相同的锂离子电池进行测试, 并根据不同实验工况将测试电池分为 7 组, 最小容量衰减到 0.80A·h。

虽然每组实验在放电时采用不同的随机工况，但是其充电时均为相同的恒流恒压工况。该数据集可用于进行锂离子电池 SOH 估算研究的完整恒流-恒压充电曲线总共 842 条。唯一不足的是，该数据集充电时记录的电压采样点时间间隔过大(30s)，失去了很多细节信息。牛津大学电池老化数据集采用了 8 块 Kokam 的钴酸锂离子袋装电池，该电池的额定容量为 740mA·h。与 NASA 随机老化数据集不同的是，其所有数据均在 40℃ 恒定环境温度下，通过使用 ARTEMIS 市区行驶工况对锂离子电池重复进行 2C(1.48A) 恒定电流放电并重新充电获得，容量最小衰减到 0.43A·h。充电过程中，电压的采集频率为 1Hz，共获得 517 条完整充电电压曲线用于 SOH 估计。2 个数据集集中的部分电池容量衰减曲线如图 2 所示，可看出电池的容量随着使用时间(循环次数)衰减。

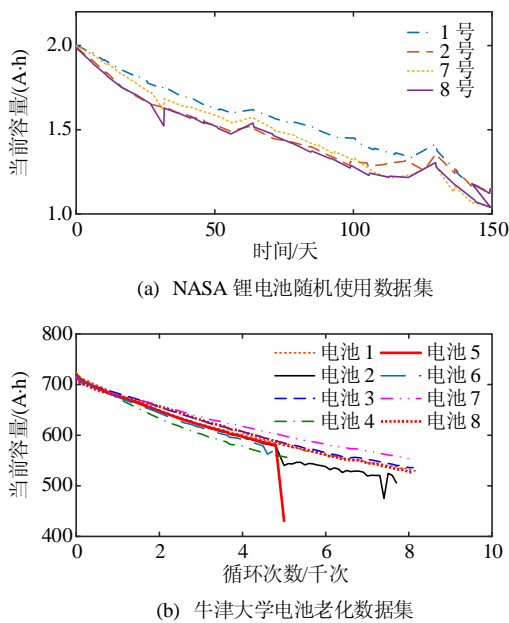


图 2 部分电池容量衰减曲线

Fig. 2 Partial battery capacity degraded curve

另外，为了对本文提出的 SOH 方法性能进行定量评估，本文采用平均绝对值误差(mean absolute error, MAE)和最大误差(max error, MAX)作为评价指标。如式(9)、(10)所示。

$$MAE = \frac{1}{n} \sum_{j=1}^n |y_j - y'_j| \quad (9)$$

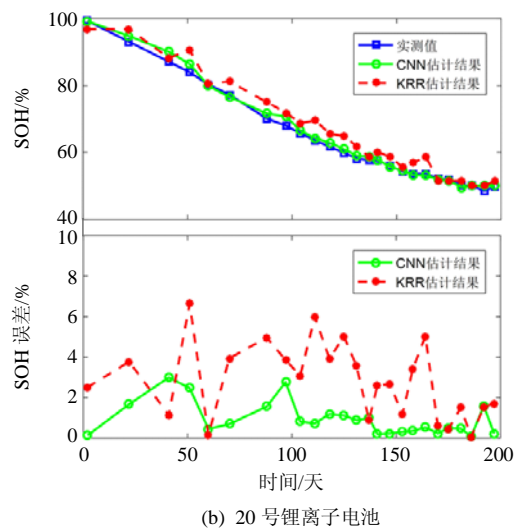
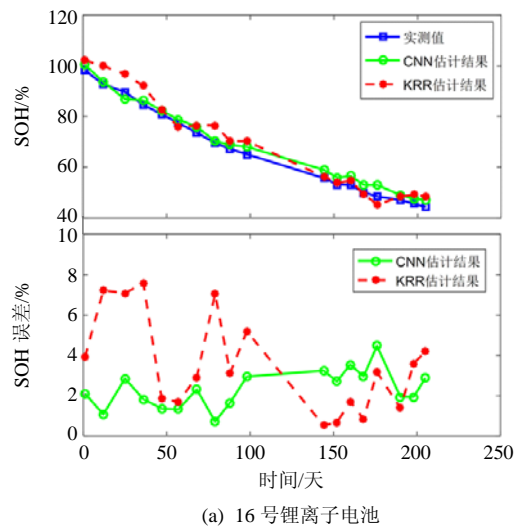
$$MAX = \max |y_j - y'_j| \quad (10)$$

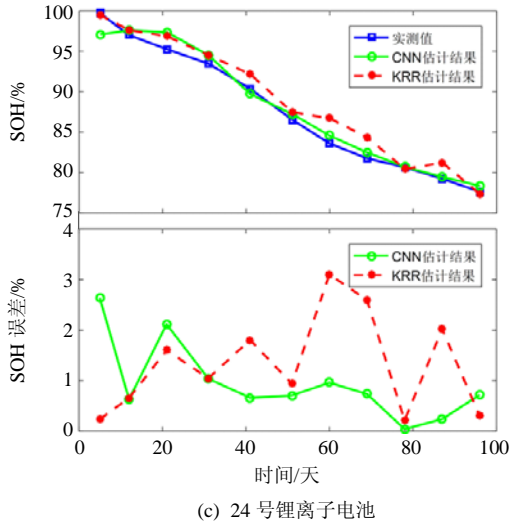
式中：总共有 n 个待测样本； y_j 、 y'_j 分别表示 SOH 估计结果和实际值。

3.2 NASA 锂离子电池随机使用数据集估计结果

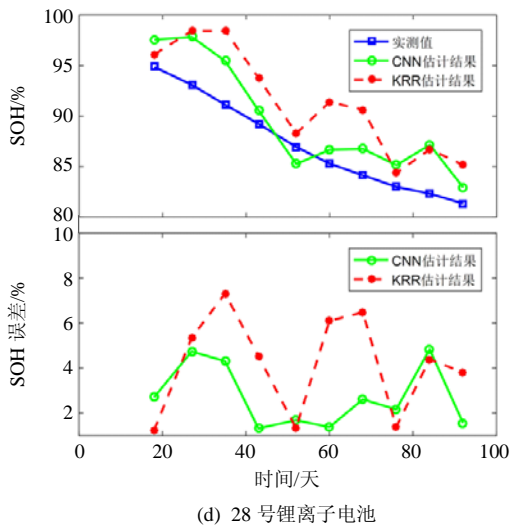
对于 NASA 锂离子电池随机数据集，选取其中 16 号锂离子电池(25℃ 低倍率放电)、20 号锂离子电池(25℃ 高倍率放电)、24 号锂离子电池(40℃ 低倍率放电)以及 28 号锂离子电池(40℃ 高倍率放电)的恒流-恒压充电阶段的数据作为测试样本，并将剩下的 24 块锂离子电池实测数据。

对于输入的片段电压充电曲线，初始电压 V_0 为 3.9V，而输入片段电压的长度 $N=40$ (由于该数据集数据采样间隔为 30s， $N=40$ 为 1200s 的充电时间)。这相当于完整充电曲线 1/3 的区间，相对于完整的 SOC 区间该区间电池正常使用时更容易覆盖。而核岭回归中式(4)正则化参数 λ 设置为 0.00001。为了评价提出本文方法的有效性，本文选取了本团队之前提出的基于 CNN 的 SOH 方法进行了对比。值得注意的是，CNN 方法采用的是完整充电电压曲线作为输入进行 SOH 估计。两种方法 SOH 估计结果如图 3 所示，定量结果如表 2 所示。





(c) 24号锂离子电池



(d) 28号锂离子电池

图3 NASA 锂离子电池随机使用数据集 SOH 估计曲线
Fig. 3 SOH estimation curves of NASA-Randomized battery usage data set

表2 NASA 锂电池随机使用数据集 SOH 性能评估结果
Table 2 Performance evaluation results of SOH estimation for NASA-randomized battery usage data set %

方法	评价标准	16号	20号	24号	28号	全局量
CNN	MAE	1.38	1.07	1.32	1.44	1.26
	MAX	4.87	2.83	2.56	4.64	4.87
KRR	MAE	2.61	2.05	2.28	2.22	2.29
	MAX	7.73	6.94	2.71	7.49	7.73

通过图3和表2可以看出,对于NASA锂离子电池随机使用数据集,KRR方法取得了不错的估计结果。4块电池估计结果MAE和MAX分别为2.61%、2.05%、2.28%、2.22%和7.73%、6.94%、2.71%、7.49%。但是,相比于CNN方法,4块电池KRR方法取得的SOH估计结果性能均有一定下降。这主要有2个原因:1)该数据集充电电压数据的记录间隔为30s,失去了很多细节信息,影响

表3 牛津大学电池老化数据集 SOH 性能评估结果

Table 3 Performance evaluation results of SOH estimation for oxford battery degradation dataset %

方法	评价标准	Cell 4	Cell 8	全局量
CNN	MAE	1.18	0.74	0.99
	MAX	2.31	2.97	2.97
KRR	MAE	0.21	0.25	0.23
	MAX	0.51	0.76	0.76

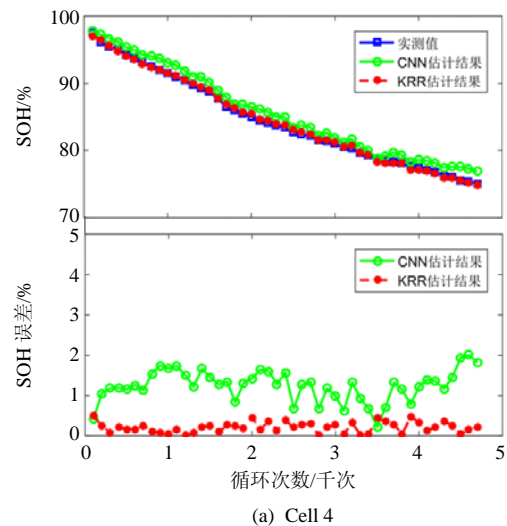
了SOH估计性能;2)CNN方法将温度信息也考虑在内,而KRR方法未考虑温度信息。尽管如此,KRR方法的MAE均小于5%,可以满足目前BMS的需求。

3.3 牛津大学电池老化数据集估计结果

为了进一步验证提出的SOH估计方法的适用性,本节在牛津电池老化数据集上对该方法进行了验证。其中,将4号和8号电池数据作为测试样本,其余6块电池作为训练样本。

对于该数据集,输入的片段电压充电曲线参数设置为初始电压 V_0 为3.8V,而输入片段电压的长度 $N=600$ (约为600s的充电时间),这相当于完整充电曲线六分之一的区间。其他实验设置与NASA锂离子电池随机使用数据集一致。SOH估计结果曲线和定量结果分别如图4和表3所示。

从图4和表3中可以看出,相对于基于CNN的SOH估计方法,本文提出的基于充电电压片段和KRR的方法取得了极佳的SOH估计效果,证明了本文所提出的方法同样能够在牛津电池老化数据集上实现准确的SOH估计。具体来说,本文提出方法在Cell 4和Cell 8两块电池分别取得了0.21%MAE、0.51%MAX和0.25%MAE、0.76%MAX,远低于现有BMS中5%MAE的SOH设计



(a) Cell 4

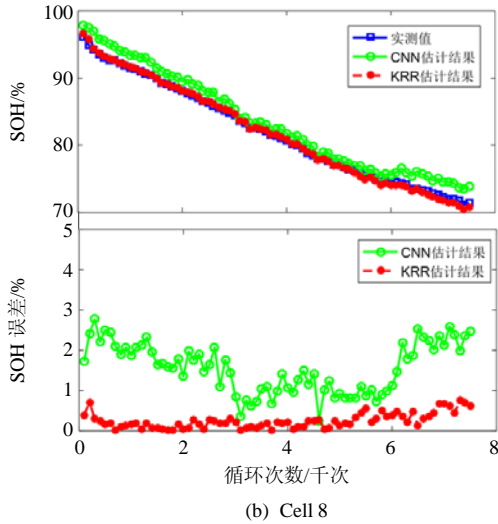


图 4 牛津大学电池老化数据集 SOH 估计曲线
Fig. 4 SOH estimation curves of Oxford battery degradation dataset

表 4 2 种方法算法复杂度对比

Table 4 Computation complexity comparison of the two methods

方法	CNN	KRR
训练复杂度	$O(m^2 knlN)$	$O(m^3)$
测试复杂度	$O(knlN)$	$O(N^2)$
参数个数	$(kN+1)nl$	N

k 为卷积核大小, n 为卷积核个数, l 为卷积层数, N 为输入数据维度, m 为训练样本数, N 为电压片段中选取的数据点个数

需求。相对于 NASA 锂离子电池随机使用数据集, 在牛津大学电池老化数据集上取得了较高的 SOH 结果精度, 这主要是因为该数据集该数据集充电电压数据的记录间隔为 1s, 其片段充电数据中包含的信息更为丰富。

3.4 电压片段区间选取对于 SOH 估计结果的影响

对于本文提出的方法而言, 电压片段区间选取对于 SOH 估计结果精度有着较大的影响。1 节提到, 由初始电压 V_0 和电压片段数据的长度 N 决定了电压片段区间选择。为了分析以上电压片段区间选取对本文所提出的算法的影响, 本节分别选择不同的 V_0 和 N , 采用遍历法实现锂离子电池的 SOH 估计, 其他实验设置与 3.2 和 3.3 节保持一致, 2 个数据集的结果如图 5 所示, 其中评价指标为全局 MAE。

图 5(a)给出了电压片段区间选取对于 NASA 锂离子电池随机使用数据集的 SOH 估计影响。其中, 水平坐标分别为初始电压和电压片段数据长度, 其取值范围分别是 [3.3V, 4.2V] 和 [10, 90] (即 [300s, 2700s] 的时间范围)。图中黄色以及绿色区域表示全

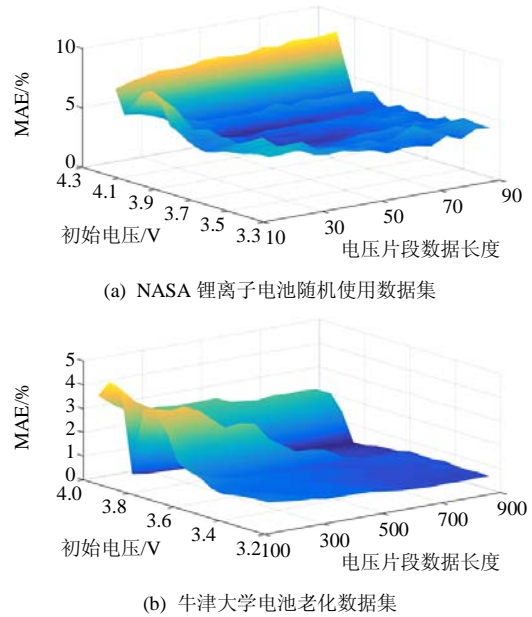


图 5 电压片段区间选取对于 SOH 估计结果的影响
Fig. 5 The influence of the initial voltage and length of segment on the SOH estimation error.

局 MAE 较大, 估计性能较差, 而蓝色区域表示全局 MAE 较小, 估计性能较好。可以总结出, SOH 估计性能和 2 个参数不是简单的单调关系。当初始电压选取较大时 ($V_0 > 4.1V$) 或者电压片段数据的长度选择较小时 ($N < 30$), SOH 估计性能差, 这主要是因为充电截止电压区域电压为一条直线或者电压片段过短导致有效特征信息过少。通过定量分析, 深蓝色区域取得了最小的估计 MAE = 2.29%, 其中 $V_0 = 3.9V$, $N = 40$, 对应值图 1(b) 可以发现此片段为 DQ/DV 变化开始变得激烈的区域, 因此包含较多的有效特征信息。

相似地, 图 5(b) 给出了电压片段区间选取对于牛津大学电池老化数据集的 SOH 估计影响。其初始电压和电压片段数据取值范围分别是 [3.2V, 4V] 和 [100, 900] (即 [100s, 900s] 的时间范围)。同 NASA 锂离子电池随机使用数据集一样, 当初始电压选取较大时 ($V_0 > 3.9V$) 或者电压片段数据的长度选择较小时 ($N < 300$), SOH 估计性能差。对于该数据集来说, SOH 取得精度最高 (MAE = 0.23%) 的电压片段区间为 $V_0 = 3.8V$, $N = 600$, 对应值图 1(d) 可以发现此为片段为 DQ/DV 变化最激烈的区域。

3.5 算法复杂性分析

本节对本文提出的方法计算复杂度进行了分析。表 4 给出了本文提出的 SOH 估计方法和本研究团队之前提出的基于 CNN 的 SOH 估计方法计算复杂度对比。

通过表4可知,无论是训练阶段还是测试阶段KRR方法的计算复杂度均远小于CNN方法。具体来说,对于核岭回归算法而言,其训练阶段的计算复杂度为 $O(m^3)^{[21]}$,其中 m 为训练样本数,虽然当训练样本数较大时计算消耗较大,可以依赖计算资源丰富的PC服务器进行计算。而根据式(7)可以推导出其测试阶段的计算复杂度 $O(N^2)$, N 为电压片段中选取的数据点个数,当选取的电压片段选取的数据点个数较少时能够在微控制器中实现。此外,相比于CNN方法中大量的参数 $((kN+1)nl)$,KRR方法中仅需要 N ,对存储空间没有太高需求,也更适合在现有的TI28335、STM32F103等BMS微控制器中实现。

4 结论

本文针对锂电池的实际使用工况,提出了一种基于充电电压片段和支持矢量回归的锂电池SOH估计方法。在2个公开数据集上进行了实验分析,通过遍历法找到了用于估计SOH的最优充电电压片段,并将实验结果与之前本研究团队提出的基于CNN的SOH方法进行了对比,证明了该方法的可行性。相对于基于CNN的SOH估计方法,该方法对于充电电压数据记录的精度要求高,但是计算复杂度非常小,更适用于在现有的TI28335、STM32F103等BMS微控制器中实现。下一步将会把温度信息也考虑在内,针对不同温度下的锂电池充电电压设计更为精确的SOH估计算法。

参考文献

- [1] COUTO L D, SCHORSCH J, JOB N, et al. State of health estimation for lithium ion batteries based on an equivalent-hydraulic model: an iron phosphate application[J]. *Journal of Energy Storage*, 2019, 21: 259-271.
- [2] ASHWIN T R, MCGORDON A, JENNINGS P A. Electrochemical modelling of Li-ion battery pack with constant voltage cycling[J]. *Journal of Power Sources*, 2017, 341: 327-339.
- [3] LIU Chang, WANG Yujie, CHEN Zonghai. Degradation model and cycle life prediction for lithium-ion battery used in hybrid energy storage system[J]. *Energy*, 2019, 166: 796-806.
- [4] NING Bo, CAO Binggang, WANG Bin, et al. Adaptive sliding mode observers for lithium-ion battery state estimation based on parameters identified online[J]. *Energy*, 2018, 153: 732-742.
- [5] RAMADESIGAN V, NORTHROP P W C, DE S, et al. Modeling and simulation of lithium-ion batteries from a systems engineering perspective[J]. *Journal of the Electrochemical Society*, 2012, 159(3): R31-R45.
- [6] LI Yi, LIU Kailong, FOLEY A M, et al. Data-driven health estimation and lifetime prediction of lithium-ion batteries: a review[J]. *Renewable and Sustainable Energy Reviews*, 2019, 113: 109254.
- [7] 刘昊天, 王萍, 程泽. 一种编解码器模型的锂离子电池健康状态估算[J]. *中国电机工程学报*, 2021, 41(5): 1851-1860.
LIU Haotian, WANG Ping, CHENG Zhe. A novel method based on encoder-decoder framework for li-ion battery state of health estimation[J]. *Proceedings of the CSEE*, 2021, 41(5): 1851-1860(in Chinese).
- [8] RICHARDSON R R, OSBORNE M A, HOWEY D A. Gaussian process regression for forecasting battery state of health[J]. *Journal of Power Sources*, 2017, 357: 209-219.
- [9] RICHARDSON R R, OSBORNE M A, HOWEY D A. Battery health prediction under generalized conditions using a Gaussian process transition model[J]. *Journal of Energy Storage*, 2019, 23: 320-328.
- [10] LIA Xiaoyu, WANG Zhenpo, ZHANG Lei, et al. State-of-health estimation for Li-ion batteries by combing the incremental capacity analysis method with grey relational analysis[J]. *Journal of Power Sources*, 2019, 410-411: 106-114.
- [11] DENG Yuanwang, YING Hejie, E Jiaqiang, et al. Feature parameter extraction and intelligent estimation of the State-of-Health of lithium-ion batteries[J]. *Energy*, 2019, 176: 91-102.
- [12] SHEN Sheng, SADOUGHI M, CHEN Xiangyi, et al. A deep learning method for online capacity estimation of lithium-ion batteries[J]. *Journal of Energy Storage*, 2019, 25: 100817.
- [13] KHUMPROM P, YODO N. A data-driven predictive prognostic model for lithium-ion batteries based on a deep learning algorithm[J]. *Energies*, 2019, 12(4): 660.
- [14] DAI Houde, ZHAO Guangcai, LIN Mingqiang, et al. A novel estimation method for the state of health of lithium-ion battery using prior knowledge-based neural network and markov Chain[J]. *IEEE Transactions on Industrial Electronics*, 2019, 66(10): 7706-7716.
- [15] 李超然, 肖飞, 樊亚翔, 等. 基于卷积神经网络的锂离子电池SOH估算[J]. *电工技术学报*, 2020, 35(19): 4106-4119.
LI Chaoran, XIAO Fei, FAN Yaxiang, et al. An approach to lithium-ion battery SOH estimation based on convolutional neural network[J]. *Transactions of China*

- Electrotechnical Society, 2020, 35(19): 4106-4119(in Chinese).
- [16] FAN Yaxiang, XIAO Fei, LI Chaoran, et al. A novel deep learning framework for state of health estimation of lithium-ion battery[J]. Journal of Energy Storage, 2020, 32: 101741.
- [17] BERECIBAR M, DEVRIENDT F, DUBARRY M, et al. Online state of health estimation on NMC cells based on predictive analytics[J]. Journal of Power Sources, 2016, 320: 239-250.
- [18] YANG Duo, ZHANG Xu, PAN Rui, et al. A novel Gaussian process regression model for state-of-health estimation of lithium-ion battery using charging curve[J]. Journal of Power Sources, 2018, 384: 387-395.
- [19] 周頔, 宋显华, 卢文斌, 等. 基于日常片段充电数据的锂电池健康状态实时评估方法研究[J]. 中国电机工程学报, 2019, 39(1): 105-111.
ZHOU Di, SONG Xianhua, LU Wenbin, et al. Real-time SOH estimation algorithm for lithium-ion batteries based on daily segment charging data[J]. Proceedings of the CSEE, 2019, 39(1): 105-111(in Chinese).
- [20] 周志华. 机器学习[M]. 北京: 清华大学出版社, 2016: 126-129.
ZHOU Zhihua. Machine learning[M]. Beijing: Tsinghua University Press, 2016: 126-129(in Chinese).
- [21] ZHANG Yuchen, DUCHI J, WAINWRIGHT M. Divide and conquer kernel ridge regression: a distributed algorithm with minimax optimal rates[J]. The Journal of Machine Learning Research, 2015, 16(1): 3299-3340.



樊亚翔

在线出版日期: 2021-01-11。

收稿日期: 2020-09-25。

作者简介:

樊亚翔(1989), 男, 工学博士, 助理研究员, 研究方向为深度学习和智能状态监测及故障诊断等, fanyaxiang@126.com;

*通信作者: 肖飞(1977), 男, 教授, 博士生导师, 研究方向为电力电子及电气传动等, xfeyninger@qq.com。

(责任编辑 吕鲜艳)

State of Health Estimation of Lithium-ion Batteries Based on the Partial Charging Voltage Segment and Kernel Ridge Regression

FAN Yaxiang, XIAO Fei*, XU Jie, YANG Guorun, TANG Xin
(Naval University of Engineering)

KEY WORDS: lithium-ion battery; state of health (SOH); partial charging voltage segment; kernel ridge regression (KRR); data-driven method

Lithium-ion batteries are being extensively used as the energy storage element to store and transform the electric energy. Nevertheless, the online estimation of battery state-of-health is an ever significant issue for the battery management system. Recently, due to its advantages such as model-free and flexibility, data-driven based methods are promising for online state-of-health estimation. The entire estimation process includes four steps: data collection, extract the features, train a machine learning model, online testing. Among them, feature extraction and model selection are the two key steps.

Aiming at the problems of heavy computing burden and difficulty in implementing for microcontroller of the existing battery SOH estimating methods, a novel estimation approach based on the partial charging voltage segment and kernel ridge regression for the state-of-health (SOH) of lithium-ion batteries is proposed. Compared with the other features, partial charging voltage segment is much more impractical for the reason that it need not the fixed initial SOC and the fully charged. Kernel Ridge Regression (KRR) combines Ridge regression with the kernel trick, which thus learns a non-linear function between the partial charging voltage segment and lithium-ion batteries SOH by the respective kernel and the data.

Let $\mathbf{x} = \{\mathbf{x}_i\}$ and $y = \{y_i\}$ be the input and output of the data, respectively. y_i can be the SOH, $\mathbf{x}_i = [x_i^1, x_i^2, \dots, x_i^n]$ can be the measured partial charging voltage from experimental data, then the KRR model can be represented as below:

$$\arg \min_w \left\{ \frac{1}{n} \sum_{i=1}^n \|f_i - y_i\|_2^2 + \lambda \|f\|_H^2 \right\} \quad (1)$$

$$f_i = \sum_{j=1}^n \alpha_j \kappa(\mathbf{x}_j, \mathbf{x}_i) \quad (2)$$

λ is a positive parameter that controls α ; $\kappa(\mathbf{x}_j, \mathbf{x}_i)$ is the kernel function; α is the unknown solution vector; $\|\cdot\|_H$ is a Hilbert space norm. Then solve linear equation for α as below:

$$(K + \lambda nI)\alpha = y \quad (3)$$

And obtain SOH with the partial charging voltage as below:

$$\tilde{y} = \sum_{i=1}^n \alpha_i \kappa(\mathbf{x}_i, \hat{\mathbf{x}}) \quad (4)$$

In the experiment, NASA-Randomized Battery Usage Data Set and Oxford Battery Degradation Dataset are utilized to validate the proposed method. Mean Absolute Error (MAE) and Max Error (MAX) are chose to quantitatively evaluate the performance of the proposed method.

The experimental results in Table 1 and 2 show that the proposed method can achieve high SOH estimation performances with average value of MAE separately below 2.29% and 0.23% in NASA dataset and Oxford dataset. In detail, it is clear that the performance of KRR is superior to CNN in Oxford dataset, and worse than CNN in NASA dataset. That is due to the reason that the KRR method ignore the influence of temperature on SOH. However, the computational complexity of KRR method is very small, and it is more suitable for implementation in existing BMS microcontrollers such as TI28335 and STM32F103. To extend this work, the future research will concentrate on designing the proposed method in the prototype hardware implementation for the BMS.

Table 1 Results of SOH estimation in NASA dataset %

Method	Criteria	16#	20#	24#	28#	Total
CNN	MAE	1.38	1.07	1.32	1.44	1.26
	MAX	4.87	2.83	2.56	4.64	4.87
KRR	MAE	2.61	2.05	2.28	2.22	2.29
	MAX	7.73	6.94	2.71	7.49	7.73

Table 2 Results of SOH estimation in Oxford dataset %

Method	Criteria	Cell 4	Cell 8	Total
CNN	MAE	1.18	0.74	0.99
	MAX	2.31	2.97	2.97
KRR	MAE	0.21	0.25	0.23
	MAX	0.51	0.76	0.76