基于高阶展开的测斜仪测量误差及不确定度分析

摘要

本文针对测斜仪数据的拟合问题，提出了一种基于高阶展开的综合拟合方案。通过对测斜仪数据的分段处理、模型拟合、残差分析以及蒙特卡洛模拟，系统地评估了不同模型在不同角度范围内的拟合效果和不确定度。研究结果表明，一次项模型在小角度范围内拟合效果良好，适用于对计算资源有限制且测量角度较小的场景；而三次项模型在更广泛的角度范围内具有更好的拟合精度，尤其是加入常数项后，能够更有效地补偿系统误差。同时，最小条件数方法在某些情况下相较于最小二乘法具有更好的拟合效果。通过蒙特卡洛模拟，量化了模型参数的不确定度，为测斜仪数据的高精度处理提供了理论依据和实践指导。最终选定的常数项、一次项和三次项模型在整体角度范围内展现了优越的拟合性能和稳定性。

关键词

测斜仪；高阶展开；不确定度；蒙特卡洛模拟；最小条件数方法

一、引言

（一）研究背景

在现代工程测量中，测斜仪作为一种高精度的测量仪器，被广泛应用于监测建筑物的倾斜、桥梁的挠度以及地质灾害中的边坡稳定性等场景[1]。然而，测斜仪的测量结果往往受到多种因素的影响，如仪器本身的系统误差、环境因素（温度、湿度等）以及操作过程中的误差等。这些误差的存在会降低测量结果的精度和可靠性，因此对测斜仪数据进行精确处理和误差分析具有重要的实际意义。

（二）研究目的和意义

本文的主要研究目的是通过对测斜仪数据进行高阶展开拟合，建立能够准确描述测斜仪测量误差的数学模型，并对其不确定度进行量化分析。通过构建高精度的误差模型，可以有效补偿测斜仪的测量误差，提高测量结果的准确性和可靠性。这对于工程测量中的高精度要求场景具有重要的应用价值，同时也为测斜仪的校准和优化提供了理论依据。

（三）研究方法概述

本文采用最小二乘法和最小条件数方法对测斜仪数据进行拟合，分别建立了一次项模型、一次项和常数项模型、一次项和三次项模型以及常数项、一次项和三次项模型。通过对不同模型的拟合效果进行评估和比较，确定了最适合测斜仪数据的高阶误差模型。同时，采用蒙特卡洛模拟方法对模型参数的不确定度进行量化分析，评估模型的稳定性和可靠性。

二、实验方法

（一）数据获取与预处理

实验数据来源于测斜仪的实测数据，包含正负两个方向的测量结果。数据预处理步骤如下：

1. 合并正负方向数据：将正负方向的数据合并，并按角度大小进行排序，确保数据的连续性和完整性。

2. 去除重复角度值的数据点：通过筛选和处理，去除重复角度值的数据点，确保数据的唯一性和准确性。

3. 分段处理：从 -1° 到 1° 开始，逐步扩展到 -30° 到 30°，每 1° 为一个分段，确保每段至少包含 5 个数据点。这种分段处理方式有助于详细分析不同角度范围内的测量误差特性。

（二）模型阶数确定

为了确定误差模型的阶数，我们对原始数据的残差进行了一系列分析。通过对残差进行求三阶导数，发现其曲线基本为常数，这表明误差具有三阶导数为常数的特性，符合泰勒展开的特性。因此，我们确定误差模型可以为三阶多项式，即包含常数项、一次项和三次项的模型。这一分析结果为后续的模型构建提供了理论依据，也验证了三阶模型在描述测斜仪测量误差方面的适用性和准确性。

（三）模型拟合方法

本文采用了两种主要的拟合方法：最小二乘法和最小条件数方法。这两种方法在处理不同复杂度的模型时表现出各自的优势。

1. 最小二乘法

最小二乘法是一种经典的参数估计方法，通过最小化观测值与模型预测值之间的平方差来求解模型参数。其数学表达式为：

其中，为观测值， 为模型预测值，为模型参数向量。

通过对目标函数求导并令导数为零，可以得到正规方程：

解此方程可得参数估计值：

2. 最小条件数方法

最小条件数方法是一种通过优化条件数来提高矩阵数值稳定性的方法，它通过添加一个小的扰动项来改善矩阵的条件数，从而提高参数估计的稳定性。其目标函数为：

其中，是正则化参数，用于控制正则化项的权重。

通过对目标函数求导并令导数为零，得到修正后的正规方程：

解此方程可得参数估计值：

（四）模型结构

本文构建了以下四种模型：

1. 一次项模型：仅考虑角度的一次项对输出的影响，模型形式为：

2. 一次项和常数项模型：同时考虑角度的一次项和常数项，模型形式为：

3. 一次项和三次项模型：同时考虑角度的一次项和三次项，模型形式为：

4. 常数项、一次项和三次项模型：引入常数项，模型形式为：

这些模型通过逐步增加高阶项，能够更好地捕捉测斜仪数据的非线性特性。

（五）模型评估指标

为了量化不同模型的拟合效果，本文采用以下评估指标：

1. 均方误差（MSE）：衡量观测值与模型预测值之间的平均平方差，值越小表示模型拟合效果越好。

2. 决定系数：衡量模型对观测数据的解释能力，值越接近 1 表示模型拟合效果越好。

3. 均方根误差（RMSE）：均方误差的平方根，与 MSE 类似，值越小表示模型拟合效果越好。

（六）蒙特卡洛模拟

为了评估模型参数的不确定度，本文采用蒙特卡洛模拟方法生成大量模拟数据，并对模拟数据进行拟合，统计模型参数的分布特性。具体步骤如下：

1. 数据生成：基于原始数据的残差分布，生成模拟数据。利用残差的均值和标准差，通过正态分布生成模拟残差，并将其加到原始数据上，得到模拟数据。

2. 参数估计：对每组模拟数据进行拟合，得到模型参数的估计值。

3. 统计分析：计算模型参数的均值和标准差，评估参数的不确定度。

蒙特卡洛模拟步骤

1. 读取数据：读取原始数据和残差数据。

2. 生成模拟数据：基于残差的均值和标准差，生成模拟数据。

3. 模型拟合：对每组模拟数据进行模型拟合，得到模型参数。

4. 统计分析：计算模型参数的均值、标准差，并进行正态性检验。

三、结果与分析

（一）分段拟合结果

以 ±30° 分段为例，展示了不同模型的拟合结果，包括拟合曲线、残差分布、残差直方图和残差散点图。结果显示，随着模型复杂度的增加（从一次项到包含常数项、一次项和三次项），模型的拟合效果逐渐改善，值提高，MSE 值降低。

|  |  |
| --- | --- |
|  |  |
| 一阶模型在 ±30° 分段的拟合曲线对比图 | 一阶+常数项模型在 ±30° 分段的拟合曲线对比图 |
|  |  |
| 一阶+三阶模型在 ±30° 分段的拟合曲线对比图 | 一阶+三阶+常数项模型在 ±30° 分段的拟合曲线对比图 |

- 图 1：不同模型在 ±30° 分段的拟合曲线对比图，展示各模型对数据的拟合效果，三次项模型在广泛角度范围内的拟合效果对比图，展示其优越的拟合性能。。

|  |  |
| --- | --- |
|  |  |
| 一阶模型在 ±30° 分段的残差分布图 | 一阶+常数项模型在 ±30° 分段的残差分布图 |
|  |  |
| 一阶+三阶模型在 ±30° 分段的残差分布图 | 一阶+三阶+常数项模型在 ±30° 分段的残差分布图 |

- 图 2：各模型的残差散点图，展示残差与拟合值之间的关系，显示残差的大小和分布情况。

|  |  |
| --- | --- |
|  |  |
| 一阶模型在 ±30° 分段的残差直方图 | 一阶+常数项模型在 ±30° 分段的残差直方图 |
|  |  |
| 一阶+三阶模型在 ±30° 分段的残差直方图 | 一阶+三阶+常数项模型在 ±30° 分段的残差直方图 |

- 图 3：残差直方图，用于观察残差的分布形态。

（二）模型比较

对比了最小二乘法和最小条件数方法在不同模型结构下的拟合效果。发现对于简单的一次项模型，两种方法的拟合结果相近；而对于包含高次项和常数项的复杂模型，最小条件数方法在某些情况下能够获得更高的 值和更低的 MSE 值，表明其对数据的拟合能力更强。

|  |  |
| --- | --- |
|  |  |
| 一阶模型最小二乘法和最小条件数方法的 值对比图 | 一阶+常数项模型下最小二乘法和最小条件数方法的 值对比图 |
|  |  |
| 一阶+三阶模型下最小二乘法和最小条件数方法的 值对比图 | 一阶+三阶+常数项模型下最小二乘法和最小条件数方法的 值对比图 |

- 图 4：不同模型结构下，最小二乘法和最小条件数方法的 值对比图。

|  |  |
| --- | --- |
|  |  |
| 一阶模型最小二乘法和最小条件数方法的 MSE 值对比图 | 一阶+常数项模型下最小二乘法和最小条件数方法的 MSE 值对比图 |
|  |  |
| 一阶+三阶模型下最小二乘法和最小条件数方法的 MSE 值对比图 | 一阶+三阶+常数项模型下最小二乘法和最小条件数方法的 MSE 值对比图 |

- 图 5：不同模型结构下，最小二乘法和最小条件数方法的 MSE 值对比图。

（三）参数变化趋势

绘制了不同角度范围内，模型参数（如 ）的变化趋势图。观察到模型参数随着角度范围的扩大呈现出一定的变化规律，这可能与测斜仪在不同角度范围内的测量特性有关。

|  |  |
| --- | --- |
|  |  |
| 一阶+三阶模型参数 随角度范围变化的趋势图 | 一阶+三阶+常数项模型参数 随角度范围变化的趋势图 |

- 图 6：模型参数 随角度范围变化的趋势图。

|  |  |
| --- | --- |
|  |  |
| 一阶模型参数 随角度范围变化的趋势图 | 一阶+常数项模型参数 随角度范围变化的趋势图 |
|  |  |
| 一阶+三阶模型参数 随角度范围变化的趋势图 | 一阶+三阶+常数项模型参数 随角度范围变化的趋势图 |

- 图 7：模型参数 随角度范围变化的趋势图。

|  |  |
| --- | --- |
|  |  |
| 一阶+三阶模型参数 随角度范围变化的趋势图 | 一阶+三阶+常数项模型参数角度范围变化的趋势图 |

- 图 8：模型参数随角度范围变化的趋势图。

（四）蒙特卡洛模拟结果

通过对模拟数据的拟合，得到了模型参数的分布直方图、正态概率图，并计算了参数的均值和标准差。结果显示，模型参数近似服从正态分布，且最小条件数方法得到的参数分布与最小二乘法相比具有更小的标准差，表明其参数估计更为稳定。

|  |
| --- |
|  |
|  |
|  |

- 图 9：模型参数的概率分布直方图。

蒙特卡洛模拟详细结果

1. 参数分布：绘制了模型参数的分布直方图，显示参数近似服从正态分布。

2. 正态性检验：采用 Kolmogorov-Smirnov 检验，验证参数分布的正态性。结果显示，所有参数的 p 值均大于 0.05，表明参数分布与正态分布无显著差异。（p 值是在假设检验中用来衡量观测结果显著性的指标，表示在原假设成立的情况下，出现当前样本结果或更极端结果的概率。在 Kolmogorov-Smirnov 检验中，p 值用于判断样本数据是否服从正态分布。）

3. 不确定度评估：计算了模型参数的均值和标准差，量化了参数的不确定度。

（五）小角度拟合效果

一次项模型在小角度范围内拟合效果良好，适用于对计算资源有限制且测量角度较小的场景。其原因在于小角度情况下测斜仪的测量误差与角度之间往往呈现出较好的线性关系，一次项模型足以捕捉主要的变化趋势。

|  |  |
| --- | --- |
|  |  |
| 一阶模型最小二乘法和最小条件数方法的拟合效果图 | 一阶+常数项模型下最小二乘法和最小条件数方法的 拟合效果图 |

- 图 10：小角度范围内（如 ±10°），一次项模型的拟合效果图，展示其良好的线性拟合能力。

（六）三次项模型的普遍性及常数项的作用

三次项模型在更广泛的角度范围内具有更好的拟合精度，尤其是加入常数项后，能够更有效地补偿系统误差。常数项可以对测斜仪的零点漂移等系统误差进行有效补偿，使得模型能够更精准地贴合实际测量数据。通过比较不同模型的 值，发现加入常数项后的模型在各个角度范围内的 值均显著提高，表明常数项的引入显著提升了模型的拟合能力。

|  |  |
| --- | --- |
|  |  |
| 一阶+常数项模型与一阶模型差值图 | 一阶+三阶+常数项模型与一阶+三阶模型差值图 |

- 图 11：加入常数项前后的模型 值对比图，直观显示常数项对模型拟合能力的提升。

（七）最终模型确定

通过对不同角度范围的分段拟合结果进行综合分析，最终选定常数项、一次项和三次项模型作为最终模型。该模型在广泛的角度范围内展现出优越的拟合性能，其 值显著高于其他模型，且通过蒙特卡洛模拟得到的参数估计稳定，不确定度在可接受范围内。这表明常数项、一次项和三次项模型不仅能够有效捕捉测斜仪数据的非线性特性，还能提供可靠的测量结果，适用于各种工程测量场景。

|  |
| --- |
| 图表, 折线图  AI 生成的内容可能不正确。 |
| 一阶+三阶+常数项模型在 ±30° 分段的拟合曲线对比图 |

- 图 12：最终选定模型在整体角度范围内的拟合效果图，展示其优越的拟合性能和稳定性。

四、讨论与结论

（一）模型选择

研究表明，常数项、一次项和三次项模型能够更准确地拟合测斜仪数据，适应测斜仪在不同角度范围内的非线性特性。在实际应用中，应根据数据特点和测量精度要求选择合适的模型结构。一次项模型在小角度范围内拟合效果良好，适用于对计算资源有限制且测量角度较小的场景；而三次项模型在更广泛的角度范围内具有更好的拟合精度，尤其是加入常数项后，能够更有效地补偿系统误差。

（二）拟合方法比较

最小条件数方法在处理复杂模型时表现出一定的优势，尤其是在提高拟合精度和稳定性方面。然而，该方法的计算复杂度相对较高，对于大规模数据处理可能需要更多的计算资源和时间。

（三）不确定度评估

蒙特卡洛模拟为模型参数的不确定度评估提供了一种有效手段。通过模拟，可以量化模型参数的分布特性，为测量结果的不确定度评定提供依据。结果显示，模型参数的不确定度在可接受范围内，表明所提出的拟合方法具有较高的可靠性。

（四）误差模型分析

对原始数据的误差进行分析，发现误差具有中心对称性，符合泰勒展开的特性。通过构建三阶误差模型（不包含偶次项），验证了模型的可行性和准确性。一阶模型适用于小角度范围，而三阶模型在更广泛的角度范围内表现出更好的拟合效果。

（五）加入常数项后拟合度更好的原因

1. 模型更灵活：无常数项的模型必须通过坐标原点，而有常数项的模型可以自由地在垂直方向上平移，更好地适应数据的整体趋势。

2. 数据的实际情况：实际测量数据往往受到各种因素的影响，常数项可以用来补偿这种偏差，使模型更贴近实际数据。

3. 评估指标的改善：加入常数项后，模型可以更好地捕捉数据的整体趋势，从而提高 值，降低 MSE 和 RMSE 值。

（六）最小二乘法和最小条件数法得到的 k 值不同但 R² 相同的原因

1. 模型复杂度和参数解释的差异：不同的参数组合可能都能很好地解释数据的变异。

2. 正则化的影响：最小条件数法引入了正则化，这会改变参数估计值，但并不直接针对 优化。

3. 数据特性：数据可能对某些参数的变化不敏感。

4. 数值优化的差异：不同的优化算法可能收敛到不同的解。

（七）验证和进一步分析

1. 检查残差：比较两种方法的残差分布和残差大小。

2. 可视化结果：绘制两种方法的拟合曲线和数据点，观察它们在不同区域的拟合情况。

3. 敏感性分析：对模型参数进行敏感性分析，观察参数变化对模型输出的影响。

4. 交叉验证：使用交叉验证评估两种方法的泛化能力，观察它们在不同数据子集上的表现。

（八）最小二乘法和最小条件数法的不同之处

1. 对结果的影响：最小二乘法对数据中的噪声较为敏感，而最小条件数法通过添加正则化项，提高了参数估计的稳定性。

2. 适用场景：最小二乘法适用于数据噪声较小、设计矩阵条件数良好的情况；最小条件数法适用于数据噪声较大或设计矩阵条件数较差的情况。

综上所述，本文所提出的基于高阶展开的测斜仪数据拟合方案，通过详细的实验分析和结果评估，为测斜仪数据的高精度处理提供了一套完整的方法体系，具有重要的理论价值和实际意义。最终选定的常数项、一次项和三次项模型在整体角度范围内展现了优越的拟合性能和稳定性，为测斜仪数据的精确处理提供了可靠的解决方案。同时，通过加入常数项和使用最小条件数法，进一步提高了模型的拟合精度和稳定性。