



東北大學 秦皇島分校  
Northeastern University at Qinhuangdao

# 毕业论文

基于仿射变化的 NIR 多目标标定迁移方法研究

院 别	管理学院
专业名称	信息管理与信息系统
班级学号	1153114
学生姓名	王涛
指导教师	赵煜辉

2019 年 06 月 10 日



## 基于仿射变化的 NIR 多目标标定迁移方法研究

### 摘 要

标定迁移相比于重新标定在近红外光谱的实际运用中更有效。对于大多数的标定迁移方法来说，标准样本是必要的。但是，由于物质的组分具有挥发性以及反应性，保持样本的完整性是一个挑战。为了解决这一问题，本文提出了一种基于仿射不变量的多标签标定迁移方法。此方法能够在没有标准样本的情况下校准两台仪器之间的差异。本方法首先通过偏最小二乘法在主仪器上多标签回归模型，通过这一步我们可以获得到主仪器的预测值以及从仪器的伪标签。之后，就能够建立主仪器预测值之间的回归模型以及从仪器伪标签之间的回归模型。最后，再使用仿射变换和角度和偏差组成的增广矩阵校正伪标签。此方法通过在不同仪器上测量得到的玉米集和药片集来衡量其性能。通过结果可以得知，此方法能够显著较少预测误差，而且和其他方法相比，在大部分情况下能够实现最佳的预测误差。

**关键词：** 偏最小二乘法，多标签标定迁移，仿射变换，近红外光谱，无标样方法

## **Multi-label Calibration Transfer based on Affine Invariant for Modeling Different Near-infrared (NIR) Spectroscopy**

### **Abstract**

Calibration transfer is more effective for practical applications of near-infrared spectroscopy. Standardization samples are necessary for most of the transfer methods. However, maintaining the integrity of the standard samples is a challenge due to the volatility and reactivity of the components. To solve such a problem, a novel method named multi-label calibration transfer based on affine invariant (MLCTAI) is proposed in the work. MLCTAI can be employed to correct the difference of two instruments without standardization samples. MLCTAI firstly establishes multi-label regression model by Partial Least Squares (PLS2) on the master setting, by which the predicted values of the master and pseudo labels of the slave are obtained. Then, a regression model between the predicted labels and a regression model between the pseudo labels are established, and the angle of the slopes and the bias of the intercepts are calculated respectively. Finally, the affine transformation is adopted to correct the pseudo labels by an augmented matrix composed of the angle and the bias. The performance of the method was tested by corn and pharmaceuticals samples measured on different instruments. It can be seen from the results that MLCTAI can significantly reduce prediction errors and also can achieve the best RMSEPs in most cases compared with other methods.

**Keywords:** Partial least squares 2(PLS2), multi-label calibration transfer, Affine transformation, Near-infrared (NIR) spectroscopy, standard-free method

## 目 录

1	绪论 .....	1
1.1	研究背景 .....	1
1.1.1	近红外光谱分析技术 .....	1
1.1.2	迁移学习与标定迁移 .....	3
1.2	研究现状 .....	5
1.3	研究主要内容 .....	6
1.4	论文主要结构 .....	7
2	理论部分 .....	8
2.1	偏最小二乘法(PARTIAL LEAST SQUARE, PLS2) .....	8
2.2	仿射变换 .....	9
2.3	对比方法介绍 .....	11
2.3.1	有标样方法 .....	11
2.3.2	无标样方法 .....	13
2.4	基于仿射变化的近红外光谱多目标标定迁移方法研究 .....	14
2.5	MLCTAI 算法概要 .....	17
2.5.1	算法过程 .....	17
2.5.2	算法流程图 .....	18
3	实验 .....	20
3.1	实验环境 .....	20
3.2	实验数据 .....	20
3.3	最优参数选择 .....	21
3.4	实验结果评估 .....	22
4	实验结果 .....	23
4.1	最优参数 .....	23
4.2	MLCTAI 方法的结果 .....	24
4.3	相关方法比较结果 .....	25
4.3.1	无标样方法 .....	26

4.3.2 与有标样的方法比较 .....	30
4.4 工作展望.....	31
结 论.....	32
致 谢.....	33
参考文献.....	34
附 录.....	38
附录 A .....	38
附录 B .....	45

## 1 绪论

### 1.1 研究背景

#### 1.1.1 近红外光谱分析技术

1800 年, 赫歇尔于太阳光谱可见区域的红外部分的探测中发现了近红外光谱。近红外光(near infrared) 是位于可见光(Vis)和中红外(MIR)之间的电磁辐射波。作为人们在吸收光谱中发现的第一个非可见光区, 近红外光谱存在于光谱上 780-2526nm 到的区域。其又可以分为两个区域, 一个区域是近红外短波(780-1100nm), 另一个是近红外长波(1100-2526nm)。有机物中几种包含能量较高的化学键(主要是 NH、CH、OH)在中红外光谱区基频吸收的合频、倍频与差频三种不同的吸收带相互叠加形成了近红外光谱的光谱吸收带。早在 20 世纪 30 年代, 近红外光谱就得到了认可, 但其应用范围受当时科技水平的限制较为狭窄。到 20 世纪中, 各种不同的技术, 例如稳定的电源、光子探测器、信号放大器以及微型计算机等的发展, 极大地促进了近红外光谱的发展。到 20 世纪末期, 化学计量学发展迅猛, 近红外光谱分析也因其在光谱信息提取以及排除噪音干扰等几个方面有较好的效果, 将其引入进来。外加近红外光谱在测试中相比于其它方法展现出的特点, 其发展获得了更大的提升。

物质的结构检测以及化合物鉴别可以通过近红外光谱实现。在检测时, 分子内部原子之间发生相对振动, 分子也进行转动。当近红外光谱仪器发出的连续波长的红外光穿过物质, 物质分子中的基团会分辨红外光的频率是否和自身的振动频率或转动频率一致。当吸收了与自身频率相同的红外光, 分子获取能量, 能级发生跃迁, 从基态到能量较高的能级。分子在完成能级跃迁之后, 物质吸收在该处波长的光。分子吸收红光的情况被记录到近红外光谱仪中, 这就是近红外光谱图。

近红外光谱的横坐标使用波长或波数表示, 其代表了吸收峰的位置, 纵坐标则使用透光率或者吸光度表示, 代表了吸收强度。

当外界照射的电磁波的能量强度刚好达到物质中分子的某两个能级之间的能量差值, 那么分子可以通过吸收这个频率下电磁波的能量实现跃迁, 从而减小透射光强度。上述为产生近红外光谱的必要条件之一, 这也决定了波峰出现的位置。同时, 近红外光谱生成的必要条件之一是耦合作用存在于红外光与分子之间。分子获取能量之后进行跃迁, 必然会导致分子振动, 其偶极矩也会随之变化, 这也代表红外光的能量传递给了分

子，分子振动偶极矩则是能量传递的具体方式。每一次分子的跃迁，也就是每一次振动都会吸收相同频率的红外光，这样就会在近红外光谱上产生一个吸收峰。但总会存在一些没有偶极矩变化且红外非活性的振动分子；也会存在一些因振动频率相同发生简并的分子；还有一些分子振动频率过快或过慢导致超出检测范围等，这些因素使得实际红外谱图中的吸收峰数目损失一部分，远远不如理论结果。

近红外光谱分析技术可以划分为定性分析与定量分析。定性分析一般用于确定物质的成分，定量分析则是用于确定物质的含量。含氢基团倍频与合频振动频率不同是由不同种类的物质所含化学成分不同导致的。这一特性体现在近红外光谱上表现为峰位、峰数及峰强的差异。并且其图谱的差异是与样本化学成分的差异正相关的。我们便可以根据这一依据来区分组分区别较大的物质。但对于组分区别较小的物质，只有通过定量分析来确定其组分及其含量。定量分析在没有复杂样本制备过程的条件下，能够很快地同步进行物质成分多组分的快速分析，并且有很高的分析精度，还不会有化学污染。定量分析的主要过程如图 1.1 所示

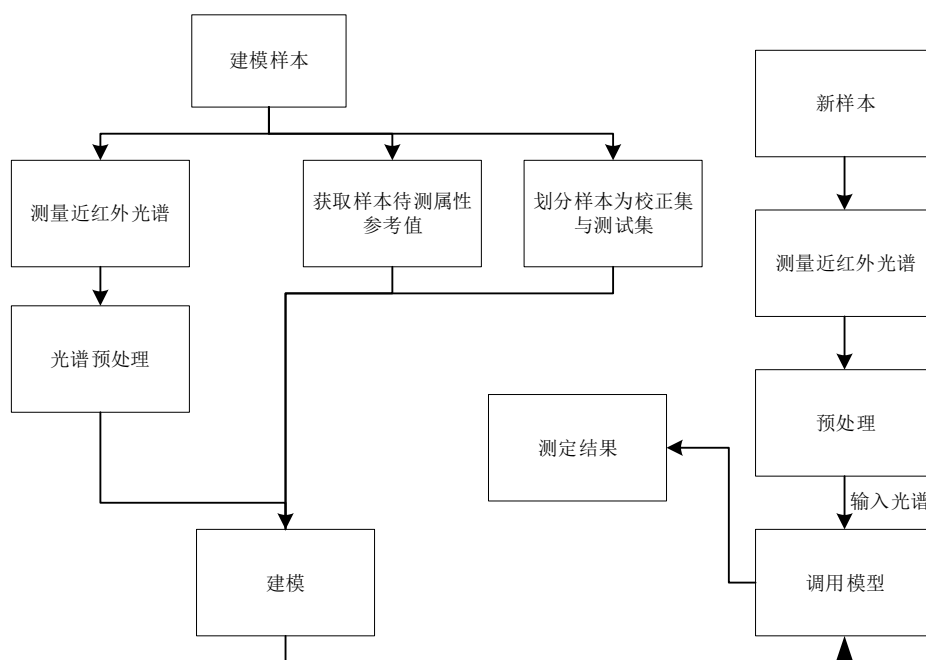


图 1.1 近红外光谱定量分析过程

近红外光谱还要求有一整套的技术流程。近红外光谱发挥较好的作用离不开下面三点的条件：(1)各项性能长期稳定且优秀的近红外光谱仪；(2)用于建模分析，且功能完备的化学计量学软件；(3)较好的模型。

近红外光谱在各方面的应用中展现出其快速、方便、高效、绿色、不破坏样本以及



不消耗化学试剂、不污染环境等优点。正因如此，近红外光谱在农业、医药以及食品等行业应用广泛。

(1) 农业。近红外光谱自发现以来，就一直应用于农产品领域。NIR 可以用于测定饲料中的各个营养成分的含量，包括蛋白质、纤维等。除此之外，还可以用来测定谷物中的水分、蛋白质、油、和淀粉等的含量。蔬菜中的成分含量也可以通过此方法来测定。不同的植物对于病虫害的抗性不同，常规的检测方法用于检测病虫害则会根据不同的情况采取不同的措施，不仅操作复杂，而且周期长，相比而言，近红外光谱就会省时省力，而且相当准确。Russell 等人通过近红外光谱对土壤进行测定从而来确定水稻在田间和温室中对氮的吸收能力<sup>[1]</sup>。

(2) 医学制药。近红外光谱能够快速、方便、准确地对药物进行鉴别和分析。不仅如此，药物的纯度鉴定也可以使用近红外光谱。与此同时，近红外光谱的非侵入式分析可以通过光导纤维以及传感技术，还有近红外光谱检测技术以及计算机网络技术等几门技术的结合而实现。生物体内不同的透明组织对近红外光具有不同的吸收与散射能力。据此我们可以使用近红外光谱区分不同的组织以及变化的组织。这一特性可以用于测量组织的某些光学参数或建立某些参数与光谱数据的关系。Afara 等人通过近红外光谱对关节软骨进行测量从而确定是否产生病变<sup>[2]</sup>。

(3) 食品。近红外光谱在食品方向的应用也十分的广泛，可以应用于酿造食品，如酒、酱油、醋等，也可以用于牛奶制品以及肉制品等方面，还适用于调味料、食用油、成分鉴别、真伪鉴别等方面。同时，近红外光谱可以用于分析各类奶制食品中的蛋白质、氨基酸和脂肪酸等的含量，并以此为根据评价其品质，同时可以应用于水果中各种成分的分析。蛋白粉中的蛋白质也可以通过近红外光谱检测，不仅安全无污染，而且不需要使用到任何的化学试剂产品<sup>[3]</sup>。

### 1.1.2 迁移学习与标定迁移

机器学习以及数据挖掘已经在回归、分类和聚类等领域取得较大的成绩<sup>[4]</sup>。作为机器学习的两大分类，传统机器学习与迁移学习是两个不同的概念。传统机器学习强调的从头开始，在每一次的学习过程中不依赖任何其他的方法与数据。而迁移学习则是通过根据以前的任务总结出相应的知识，并将其迁移到新的目标任务中。图 1.2 显示了两种学习方式的区别：

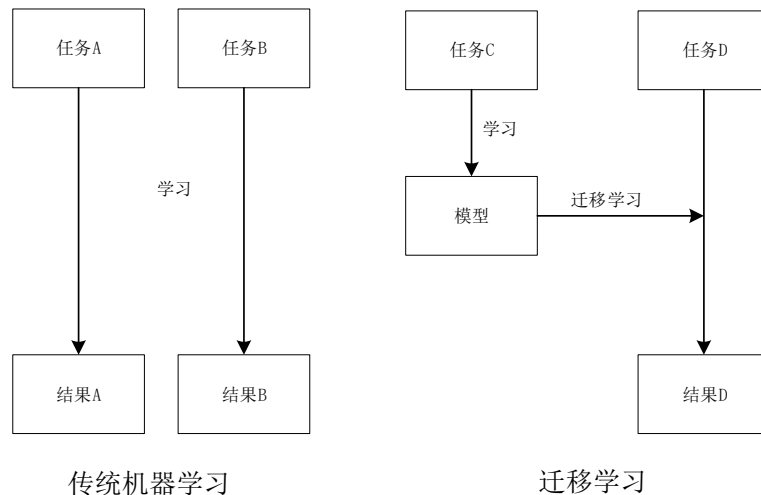


图 1.2 传统机器学习与迁移学习的区别

然而，许多传统的机器学习和数据挖掘方法对于数据的要求十分严苛，规定其训练集与测试集必须有相同的数据分布，而且必须来自相同的特征空间。在实际的应用中，当情形变得复杂时，条件瞬息万变，大多数方法不再适用，每次条件改变就需要大量的训练数据用于训练新的模型。在大部分情况下，这些数据十分昂贵，而且难以获取。在这样一个背景下，迁移学习就变得很有必要。

迁移学习属于机器学习的范畴，指一个在之前的任务中训练的模型被重新用在另一个任务中。NIPS-95 “Learning to learn” 研讨会上讨论了机器学习领域迁移学习的动机，其重点为需要终生使用机器学习方法来保留和重用以前学过的知识。自 1995 年以来，迁移学习以不同的名称受到了很多的关注，其使用过的名称多种多样，包含：终身学习、知识迁移、归纳迁移、多任务学习、知识整合、环境敏感学习、基于知识的归纳偏见，元学习，和增量/累积学习等。2005 年，国防高级研究计划局(DARPA)的信息处理技术办公室对迁移学习做了一个准确的定义：系统识别并将在以前任务中学习到的知识和技能应用于新任务的能力。

迁移学习一般迁移在类似但不相同的领域、任务和分布之间的知识。例如，一个打过篮球和一个没有练习过任何运动的人一起去学习足球。那么，练习过篮球的人会通过迁移学习，将练习篮球获得的身体的灵活性以及协调性用于足球，他将比没有练习过任何运动的人更快学会足球。同样，一个会玩吉他的人和一个没有学习过乐器的人一起去打架子鼓，那么，会玩鼓的那个人可以通过迁移学习，将其掌握的乐理知识应用到架子鼓上，能够更好地掌握住架子鼓的节奏和敲打的频率。因此，迁移学习研究的动机是，人们总是可以将以前学到的知识更快地应用于解决新问题。

迁移学习在化学计量学领域的应用就是标定迁移。标定迁移的兴起是由于近红外光谱的广泛应用，这也促进了多元标定技术的发展。但当现有的模型在新的环境下，或者在新的仪器上使用时，效果就会很差。但使用重新标定方法十分昂贵而且费时。在这样的环境下，标定迁移应运而生。标定迁移指在不同的仪器下或者不同的条件下对多元标定模型的迁移。标定迁移有两个主要的应用场景：(1) 在相似的样品和测量条件的多个仪器上使用相同的标定模型；(2) 使用相同的标定模型以适应新的测量和样品条件。

在化学计量学中，迁移方法一般是针对主从仪器的差异进行分析。为此，绝大多数办法都要求使用到一组样本称之为迁移标准，也就是标准样本，这是一组在两个仪器中测量得到的一组具有代表性的样本。这组样本通常提供了有关仪器之间的差异信息。据此可以开发出一些方法来纠正这个差异。这一组标准样本必须具有代表性和稳定性。

### 1.2 研究现状

近红外（NIR）光谱和多元标定方法的结合已成为监测化学和生物化学过程，食品和医药等方面的一种广泛使用的技术<sup>[5-7]</sup>。多元标定方法可以从近红外光谱中提取化学信息并构建效果较好的模型，较常用的两种方法是主成分回归(PCR)<sup>[8-10]</sup>和偏最小二乘法(PLS)<sup>[11-13]</sup>。然而，当精心构造的模型应用于新光谱或新的实验条件时，模型的性能就会下降。即使在理想情况下测量样本，原始仪器未捕获，且和新仪器相关的信息会导致预测效果很差。重新校准可用于解决此问题。但它在样品重新测量和模型构建方面十分昂贵且耗时。另一种可接受的解决方案是标定迁移，这指的是将一系列化学分析方法应用于单一的仪器上。

各种各样的标定迁移方法已经被开发了出来，其中一些方法在[14-16]中进行了全面讨论。对于大多数方法，必须使用一组标准样品在两个仪器上构建迁移模型。这就是有标样的方法。例如，斜率和偏差校正(SBC)<sup>[17-19]</sup>，分段直接标准化(PDS)<sup>[20-21]</sup>和基于典型相关分析(CCA)<sup>[22]</sup>的标定迁移方法。

根据不同的预测值之间存在线性关系，SBC拟合了标准样本响应值和预测值之间的线性关系。然后这个线性关系就用做校正预测值。由wang等人提出的PDS方法是一种常用的方法，且经常被用作其他算法的对比算法<sup>[23]</sup>。它基于这样的假设：光谱之间的相关数据仅限于一个小区域。此外，PDS的重点在于确定窗口大小和标准样本的数量。但是，对于每个窗口移动，它需要在模型构造中进行大量计算。此外，基于CCA的标定迁移方法非常适合关联两个光谱。首先，通过典型相关分析<sup>[24-25]</sup>获得两个仪器的典型向

量。然后，它们被应用在普通最小二乘回归(LS)中构建标定迁移模型<sup>[26-27]</sup>。最后再通过此模型校正光谱的差异。

然而，标准样品非常昂贵，并且在一些实际应用中难以或甚至不可能获得标准样品。在这种情况下，特别需要建立没有标准样品的转移模型。这类方法，如多元散射校正(MSC)<sup>[28-30]</sup>和迁移成分回归(TCR)<sup>[31-33]</sup>等，在参考文献[16]中进行了细致的讨论。

作为一种经典的预处理方法，MSC 适用于化学相似光谱之间的校正。它具有处理光谱来减少或消除差异并增强光谱吸收信息的能力。然而，在许多情况下，整个波长范围内光谱的斜率和偏差的变化较大，MSC 方法此时具有较差的性能。其他常用的无标样方法还有广义最小二乘法(GLS)<sup>[34]</sup>，正交信号校正(OSC)<sup>[35-37]</sup>和有限脉冲响应滤波(FIR)<sup>[38-39]</sup>等。

TCR 是迁移成分分析(TCA)和普通最小二乘法(OLS)的组合。由香港科技大学杨强教授团队提出的 TCA 方法主要将两个仪器的数据投影到一个共同的低维再生核 Hilbert 空间，其中源域和目标域之间的距离应尽可能接近，同时最大限度地保留它们的内部属性。然而，TCR 在进行计算时需要从仪器的参考值，并且特征空间的确定涉及复杂且昂贵的计算。其他具有类似效果的方法包括核主成分分析(KPCA)<sup>[40-42]</sup>，协方差算子逆回归(COIR)<sup>[43]</sup>等。

### 1.3 研究主要内容

本文提出了一种名为基于仿射变换的多标签标定迁移方法(MLCTAI)。此方法是无标样的方法，旨在通过仿射变换校正仪器差异。它创造性地寻求浓度之间的线性势。趋势通常反映为直线，并且可以通过最小二乘回归 (LS) 拟合。同时，在 MLCTAI 方法中，数据集必须包含多种物质的浓度值。因此，主仪器的预测值包含多个标签。由于从仪器缺少响应值，因此由主仪器构建的 PLS2 模型将应用于从仪器模型的建立，并计算伪预测值以构建从仪器线性模型。但是，从主仪器视角观察从仪器的数据时会出现偏差。偏差具体表现为从仪器模型相对于主仪器模型的角度和截距差异。仿射变换则是通过旋转和平移来校正两个仪器模型的偏差，并且最终获得校正的预测值。

MLCTAI 的过程可大致分为以下两个步骤。

#### 1. 建立主从仪器模型

首先在主仪器上根据近红外光谱以及相应的浓度值建立 PLS2 模型，并计算出回归系数。其次，根据回归系数计算出主仪器的预测值，提取出其中的两种成分用于建立线

性模型。同样的，根据回归系数计算出从仪器的伪预测值，提取出其中同样的两种成分建立线性模型。

### 2. 计算偏差并进行迁移

根据前面得到的两个线性模型，计算出两个模型的偏差。偏差具体体现为两条拟合直线的斜率和截距的差异。接着就可以通过仿射变换公式将从仪器的伪预测值进行变换，这便矫正了主从仪器之间的差异。

## 1.4 论文主要结构

本文一共五章，其结构及内容概括如下：

第一章介绍了本文课题的相关知识领域，以及介绍本文的研究内容和结构。

第二章介绍了与本文相关的理论知识，包括交叉验证、偏最小二乘法、仿射变换、5种对比算法以及基于仿射变化的近红外光谱多目标标定迁移方法的理论介绍。

第三章介绍了本文的实验部分，包括实验环境、实验数据、最优参数的选取以及实验结果评估的参数。

第四章是对本文实验结果的展示，包括实验得到的最优参数、MLCTAI方法测试结果以及同其他算法对比得到的结果。

第五章是本文的结论，并对未来研究工作进行展望。

## 2 理论部分

### 2.1 偏最小二乘法(PARTIAL LEAST SQUARE 2, PLS2)

偏最小二乘法（PLS）是一类广泛运用的回归方法，旨在通过潜变量对光谱数据( $\mathbf{X}$ )与响应值( $\mathbf{Y}$ )之间的关系进行建模<sup>[44]</sup>。PLS 中的潜变量涵盖光谱数据和响应值的最大相关信息，可以通过最大化  $\mathbf{X}$  的得分和  $\mathbf{Y}$  的得分的协方差来求得，具体计算公式如下：

$$\begin{aligned} H &= \arg \max_{\mathbf{w}, \mathbf{c}} \text{cov}(\mathbf{X}\mathbf{P}, \mathbf{Y}\mathbf{Q}) \\ \text{subject to } & \|\mathbf{P}\|_2 = 1, \|\mathbf{Q}\|_2 = 1 \end{aligned} \quad (1)$$

$\mathbf{P}$  和  $\mathbf{Q}$  表示  $\mathbf{X}$  和  $\mathbf{Y}$  的载荷矩阵。拉格朗日乘子法可以用于这个公式的求解。

当响应值包含多种成分时，经典的最小二乘法需要多次进行建模，然而 PLS2 方法只需建模一次即可。对于  $\mathbf{X}$  和  $\mathbf{Y}$  矩阵的每一种成分，可以求得一种线性关系。对于多组分的  $\mathbf{X}$  矩阵，可以获取到一组线性关系：

$$\begin{cases} \mathbf{Y}_1 = \mathbf{X}\boldsymbol{\beta}_1 + \mathbf{E}_1 \\ \mathbf{Y}_2 = \mathbf{X}\boldsymbol{\beta}_2 + \mathbf{E}_2 \\ \dots \\ \mathbf{Y}_k = \mathbf{X}\boldsymbol{\beta}_k + \mathbf{E}_k \end{cases} \quad (2)$$

将  $\mathbf{Y}_1, \mathbf{Y}_2, \dots, \mathbf{Y}_k$  一列挨着一列放入到一个  $m \times k$  的矩阵中，同时也对  $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \dots, \boldsymbol{\beta}_k$  与  $\mathbf{E}_1, \mathbf{E}_2, \dots, \mathbf{E}_k$  采取相同的处理方式。最终，建立起 PLS2 模型：

$$\mathbf{Y} = \mathbf{X}\mathbf{B} + \mathbf{E} \quad (3)$$

其中， $\mathbf{Y} = [\mathbf{Y}_1, \mathbf{Y}_2, \dots, \mathbf{Y}_k]$  为模型的响应值， $\mathbf{B} = [\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \dots, \boldsymbol{\beta}_k]$  代表 PLS2 模型的回归系数。 $\mathbf{E} = [\mathbf{E}_1, \mathbf{E}_2, \dots, \mathbf{E}_k]$  是  $\mathbf{Y}$  经过 PLS2 处理之后的残差矩阵。PLS2 方法的目标就是计算出回归系数，并且使  $\mathbf{E}$  尽可能的小，甚至是  $\mathbf{0}$ 。

一种用于计算回归系数的著名 PLS 算法是非线性迭代偏最小二乘法(Nonlinear iterative partial least squares, NIPALS)。该算法并不能一次性得到主成分，它是一个迭代的过程。对于每一次迭代，我们都可以计算得到一组  $\mathbf{W}_i$ ， $\mathbf{T}_i$ ， $\mathbf{P}_i$  和  $\mathbf{Q}_i$ 。之后使用  $\mathbf{X} - \mathbf{T}_i\mathbf{P}_i$  与  $\mathbf{Y} - \mathbf{T}_i\mathbf{Q}_i$  分别代替  $\mathbf{X}$  和  $\mathbf{Y}$ ，再进行下一次迭代。如此迭代下去，直到收敛。

对这些迭代过程进行组合，便得到了公式(2.4)：

$$\begin{aligned} \mathbf{X} &= \mathbf{T}\mathbf{P}^T + \mathbf{G} \\ \mathbf{Y} &= \mathbf{T}\mathbf{Q}^T + \mathbf{F} \end{aligned} \quad (4)$$

其中  $\mathbf{T}=[\mathbf{T}_1, \mathbf{T}_2, \dots, \mathbf{T}_j]$  代表  $\mathbf{X}$  的得分矩阵， $\mathbf{P}=[\mathbf{P}_1, \mathbf{P}_2, \dots, \mathbf{P}_j]$  和  $\mathbf{Q}=[\mathbf{Q}_1, \mathbf{Q}_2, \dots, \mathbf{Q}_j]$  分别代表  $\mathbf{X}$  和  $\mathbf{Y}$  的载荷矩阵， $\mathbf{G}$  和  $\mathbf{F}$  则是  $\mathbf{X}$  和  $\mathbf{Y}$  的残差矩阵。 $j$  表示通过交叉验证得到的最佳潜变量数。

最终，PLS2 模型的回归系数可以通过公式(2.5)求出：

$$\mathbf{B}=\mathbf{W}(\mathbf{P}^T\mathbf{W})^{-1}\mathbf{Q}^T \quad (5)$$

这里  $\mathbf{W}=[\mathbf{W}_1, \mathbf{W}_2, \dots, \mathbf{W}_j]$  表示权重矩阵。

## 2.2 仿射变换

作为一种几何变换，仿射变换可以将物体从一个空间映射到另一个空间，同时保证自身结构不变。通常，仿射变换会改变一些属性，例如方向，大小和位置，但它还是保留了一部分属性。仿射变换具有平行线转换为平行线和有限点映射到有限点的一般特征<sup>[45]</sup>。仿射变换是大量的变换的组合。二维仿射变换常见的基础变换有：缩放，平移，旋转，反射等。

在本文中，我们主要讨论的是二维仿射变换中的旋转和平移。图 2.1 可以更加形象地理解仿射变换。

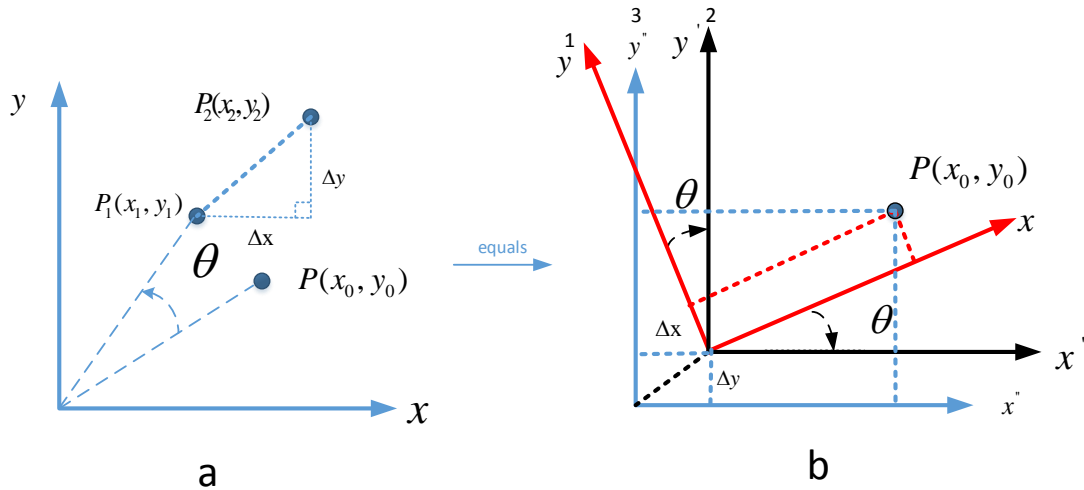


图 2.1 仿射变换的旋转和平移

在图 2.1(a)中，初始点  $P$  的坐标是  $(x_0, y_0)$ ， $P_1(x_1, y_1)$  则是点  $P$  经过旋转之后得到的点。坐标系中点的逆时针旋转等于坐标系的顺时针旋转而该点不动。点  $P$  在新的坐标系(坐标系 2)中的位置就相当于  $P_1$  在原始坐标系(坐标系 1)中的位置。之后，再对  $P_1$  进行  $x$  轴和  $y$  轴的平移，最终得到了  $P_2(x_2, y_2)$  点，这也相当于点  $P_1$  不动，对坐标系 2 进行相反的平移到坐标系 3。为一次性包含上述的变化，引进变换矩阵  $\mathbf{M}'$ ：

$$\mathbf{M}' = \begin{bmatrix} \cos \theta & \sin \theta & 0 \\ -\sin \theta & \cos \theta & 0 \\ \Delta x & \Delta y & 1 \end{bmatrix} \quad (6)$$

这里的 $\theta$ 代表旋转的角度， $\Delta x$ 和 $\Delta y$ 分别表示P点在x轴和y轴上平移的距离。

对p点进行仿射变换的公式可以写成如下形式：

$$\begin{bmatrix} x_2 & y_2 & 1 \end{bmatrix} = \begin{bmatrix} x & y & 1 \end{bmatrix} \begin{bmatrix} \cos \theta & \sin \theta & 0 \\ -\sin \theta & \cos \theta & 0 \\ \Delta x & \Delta y & 1 \end{bmatrix} \quad (7)$$

这个矩阵公式也可以改写成下面的形式：

$$\begin{aligned} x_2 &= x \cos \theta - y \sin \theta + \Delta x \\ y_2 &= x \sin \theta + y \cos \theta + \Delta y \end{aligned} \quad (8)$$

接下来讨论在本实验中使用到的用于校正伪预测值的仿射变换。根据公式(2.6)，构建出伪预测值的增广矩阵 $\tilde{\mathbf{U}}_{ij}$ 和变换矩阵 $\mathbf{M}$ ：

$$\begin{cases} \tilde{\mathbf{U}}_{ij} = [\tilde{\mathbf{Y}}_i^s, \tilde{\mathbf{Y}}_j^s, \mathbf{1}] \\ \mathbf{M} = \begin{bmatrix} \lambda_i \cos \Delta \theta & \lambda_i \sin \Delta \theta & 0 \\ -\lambda_j \sin \Delta \theta & \lambda_j \cos \Delta \theta & 0 \\ 0 & \Delta b & 1 \end{bmatrix} \end{cases} \quad (9)$$

其中， $\mathbf{1}$ 表示元素全为1的列向量。 $\lambda_i$ 和 $\lambda_j$ 表示伪预测值的缩放因子，可以通过下面的公式进行计算：

$$\begin{cases} \lambda_i = \left[ \frac{\tilde{\mathbf{Y}}_i^s - \min(\tilde{\mathbf{Y}}_i^s)}{\max(\tilde{\mathbf{Y}}_i^s) - \min(\tilde{\mathbf{Y}}_i^s)} + \min(\tilde{\mathbf{Y}}_i^s) \right] * (\max(\tilde{\mathbf{Y}}_i^s) - \min(\tilde{\mathbf{Y}}_i^s)) \\ \lambda_j = \left[ \frac{\tilde{\mathbf{Y}}_j^s - \min(\tilde{\mathbf{Y}}_j^s)}{\max(\tilde{\mathbf{Y}}_j^s) - \min(\tilde{\mathbf{Y}}_j^s)} + \min(\tilde{\mathbf{Y}}_j^s) \right] * (\max(\tilde{\mathbf{Y}}_j^s) - \min(\tilde{\mathbf{Y}}_j^s)) \end{cases} \quad (10)$$

其中， $\tilde{\mathbf{Y}}_i^s$ 和 $\tilde{\mathbf{Y}}_j^s$ 分别表示从仪器的伪预测值。

最终，伪预测值通过下面的公式校正：

$$\hat{\mathbf{U}}_{ij} = \tilde{\mathbf{U}}_{ij} \mathbf{M} \quad (11)$$

公式中的 $\hat{\mathbf{U}}_{ij} = [\hat{\mathbf{Y}}_i^s, \hat{\mathbf{Y}}_j^s, \mathbf{1}]$ 表示已校正的预测值的增广矩阵。



## 2.3 对比方法介绍

### 2.3.1 有标样方法

#### 1. 分段直接标准化(Piecewise direct standardization, PDS)

分段直接标准化方法是一种比较常用的有标样方法，经常作为新技术的对比方法。与其他标准化方法相比，其优点在于局部性和多元性，能够纠正强度差异，波长移动和峰值展宽。对于近红外光谱，光谱相关的部分被限制在较小的一个区域内。分段直接标准化方法正是基于此提出的。分段直接标准化中用于建立线性关系的窗口也因此而产生。

在分段直接标准化中，主仪器上标准样本的测量值和在主仪器上测量的窗口的测量值有关。对于每一个波长，都可以使用偏最小二乘法(PLS)或者主成分回归(PCR)建立线性关系如下：

$$\mathbf{X}_j^m = \mathbf{X}_j^{sw} \mathbf{b}_j \quad (12)$$

其中  $\mathbf{X}_j^m$  表示主仪器标准样本在波长  $j$  处的光谱， $\mathbf{X}_j^{sw}$  表示从仪器在波长为  $j$  时窗口上的光谱。 $\mathbf{b}_j$  则表示在波长为  $j$  时的迁移系数向量。

对于每一次移动窗口，都对应了一个新的模型，每个模型都可以求出一个迁移系数向量。将所有的模型组合到一起，将所得到的  $n$  个迁移系数向量按照顺序放入一个对角矩阵的对角线上，就得到了分段直接标准化中的带状对角变换矩阵  $\mathbf{T}$ ：

$$\mathbf{T} = \text{diag}(\mathbf{b}_1^T, \mathbf{b}_2^T, \dots, \mathbf{b}_n^T) \quad (13)$$

其中  $n$  表示波长的数量。

最后，通过公式(2.13)，利用变换矩阵  $\mathbf{T}$  将未知样品的响应值  $\mathbf{X}_u$  投影到原始测量空间，就可以使用已有的模型测量其属性值。

$$\mathbf{X}_{trans} = \mathbf{X}_u \mathbf{T} \quad (14)$$

PDS 方法根据其假设设定一个窗口值，每一次窗口移动都会伴随着建模。因而，在效果极佳的同时，无法避免的缺点是复杂的计算。

#### 2. 斜率偏差校正(slope and bias correction, SBC)

斜率偏差校正假设在主从仪器上标准样本的预测值之间存在线性关系。当在新测量的仪器中，所有数据之间存在的仪器差异系统而简单时，斜率偏差校正方法具有较好的效果。当系统差异更复杂时，此方法不再适用。

斜率偏差校正方法首先在主仪器上建立一个 PLS 模型，求得回归系数  $\mathbf{m}$ 。

$$\mathbf{y} = \mathbf{X}_m \mathbf{m} + \mathbf{e} \quad (15)$$

$\mathbf{e}$  表示 PLS 模型的残差矩阵。

对两种仪器上测量的标准样品的属性值进行了首次预测, 得到主从仪器标准样本的预测值。

$$\begin{aligned} \mathbf{y}_m^s &= \mathbf{X}_m^s \mathbf{m} \\ \mathbf{y}_s^s &= \mathbf{X}_s^s \mathbf{m} \end{aligned} \quad (16)$$

其中,  $\mathbf{X}_m^s$  和  $\mathbf{X}_s^s$  分别表示主从仪器上的标准样本,  $\mathbf{y}_m^s$  和  $\mathbf{y}_s^s$  则是表示标准样本的预测值。

利用最小二乘法(ordinary least squares)或者正交最小二乘法(orthogonal least squares), 可以求得两个预测值之间的线性关系, 由此可以得到此线性关系的斜率 *slope* 和截距 *bias*。最终通过斜率和截距校正新样本的预测值。

$$\mathbf{y}_{corr} = slope \cdot y_{new} + bias \quad (17)$$

### 3. 基于典型相关分析的标定迁移方法

典型相关分析(canonical correlation analysis, CCA)中的关键中间量是综合变量, 本方法正是通过此变量之间的相关关系来反映两组数据整体之间的相关关系。此方法的基本原理有两步: 一是从原始数据中提取出具有代表性的综合变量(分别是两组数据中各数据的线性组合), 之后再利用两个综合变量的关系来体现出两组数据整体之间的相关性。最终通过典型相关分析得到的是两个典型向量。

对于典型相关分析的标定迁移方法, 首先应该在主仪器的标定集上建立一个 PLS 模型, 求得回归系数  $\mathbf{b}_{cca}$ 。

$$X_{cal}^m, \mathbf{y}_{cal}^m, folds, max\_comp \xrightarrow{PLS} \mathbf{b}_{cca} \quad (18)$$

接下来通过典型相关分析分别对主从仪器的标样建立模型, 得到两个典型向量  $\mathbf{T}_m$  与  $\mathbf{T}_s$ :

$$\begin{aligned} \mathbf{X}_{std}^m, \mathbf{y}_{std}^m &\xrightarrow{CCA} \mathbf{T}_m \\ \mathbf{X}_{std}^s, \mathbf{y}_{std}^s &\xrightarrow{CCA} \mathbf{T}_s \end{aligned} \quad (19)$$

之后通过下面的公式获得主从仪器的典型变量

$$\begin{aligned} \mathbf{L}_m &= \mathbf{X}_{std\_cen}^m \mathbf{T}_m \\ \mathbf{L}_s &= \mathbf{X}_{std\_cen}^s \mathbf{T}_s \end{aligned} \quad (20)$$

其中， $\mathbf{X}_{std\_cen}^m$  与  $\mathbf{X}_{std\_cen}^s$  分别表示经过中心化之后的  $\mathbf{X}_{std}^m$  和  $\mathbf{X}_{std}^s$ 。

然后通过公式(2.21)获取一个系数，此系数用于对原始数据进行变换得到预测值：

$$\mathbf{coef} = \mathbf{T}_s \cdot \mathbf{L}_m \cdot \mathbf{L}_s \cdot \mathbf{b}_{cca} \quad (21)$$

最终，预测值可以通过下面的公式获得：

$$\mathbf{y}_{pre} = \mathbf{y}_{test} \cdot \mathbf{coef} \quad (22)$$

### 2.3.2 无标样方法

#### 1. 多元散射校正(multiplicative scatter/signal correction, MSC)

多元散射校正法是一种常用的预处理方法，它可以保证在没有标样的情况下保证更好的匹配光谱响应值。经过它处理之后，光谱数据中的散射情况可以得到很大的缓解，因而也使得与成分含量相关的光谱吸收信息得到加强。

多元散射校正的大致过程如下图所示。

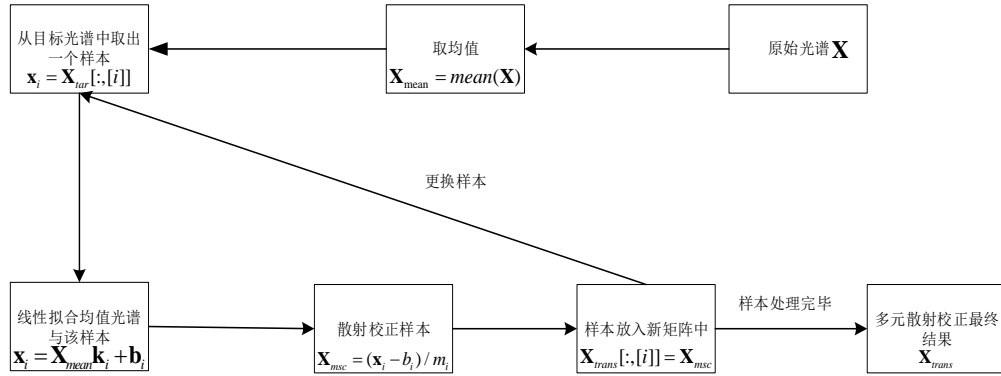


图 2.2 多元散射校正处理光谱过程

在多元标定技术中，首先应该求出原始光谱的均值，这个均值将作为标准。当输入需要进行处理的样本后，光谱会根据输入的样本数进行相应次数的循环，每次循环处理一个样本。每一次循环会将该样本与均值进行线性拟合，求得该样本相对于均值的线性平移量(截距)与倾斜偏移量(回归系数)。之后就可以使用这两个量来校正原始光谱，用每个样本的原始光谱减平移量，再除以偏移量，光谱相对倾斜的基线就得到修正。于此同时，样本的成分含量对应的光谱吸收信息却在处理过程中没有改变，这也大大提高了光谱的信噪比。

#### 2. 迁移成分回归(transfer component regression, TCR)

迁移成分分析(transfer component analysis, TCA)最初是针对域适应(domain adaptation)提出的一种基于特征的迁移学习方法。域适应中存在着一个问题就是如何减

少源域和目标域之间的差异。迁移成分分析则是根据此问题提出的，它将源域和目标域中的数据一起映射到一个再生核希尔伯特空间中，在这个空间中，最小化源域和目标域的差异。

迁移成分分析通过引入最大均值差异(Maximum Mean Discrepancy, MMD)来度量分布之间的差异。

取  $f: \mathbf{X} \rightarrow \mathbf{H}$  为非线性映射，其可以将数据从原始域内映射到高维的再生核希尔伯特空间， $\mathbf{X}_s = \{\mathbf{x}_{s_1}, \mathbf{x}_{s_2}, \dots, \mathbf{x}_{s_m}\}$  与  $\mathbf{X}_t = \{\mathbf{x}_{t_1}, \mathbf{x}_{t_2}, \dots, \mathbf{x}_{t_n}\}$  为源域和目标域数据， $\mathbf{X}'_s = \{\mathbf{x}'_{s_i} | \mathbf{x}'_{s_i} = f(\mathbf{x}_{s_i}), i=1, 2, \dots, m\}$  与  $\mathbf{X}'_t = \{\mathbf{x}'_{t_i} | \mathbf{x}'_{t_i} = f(\mathbf{x}_{t_i}), i=1, 2, \dots, n\}$  为通过  $f$  映射之后的数据。最大均值差异通过度量两个域之间的经验均值的距离平方实现：

$$dist(\mathbf{X}'_s, \mathbf{X}'_t) = \left\| \frac{1}{m} \sum_{i=1}^m \mathbf{x}'_{s_i} - \frac{1}{n} \sum_{i=1}^n \mathbf{x}'_{t_i} \right\|_{\mathbf{H}}^2 \quad (23)$$

通过引入核学习(kernel learning)，可以将映射  $f$  解出。

迁移成分分析实质上是一种降维与提取特征的方法，和主成分分析(PCA)比较类似。通过迁移成分分析处理，我们可以得到高维数据中的特征。将之与回归分析方法进行结合，就是迁移成分回归。

## 2.4 基于仿射变化的近红外光谱多目标标定迁移方法研究

在 MLCTAI 算法中，给定主仪器的输入输出  $(\mathbf{X}^m, \mathbf{Y}^m)$ ，从仪器的输入  $\mathbf{X}^s$ ，这个算法的目的是求出从仪器的预测值  $\tilde{\mathbf{Y}}^s$ 。当物质浓度处于同一范围内，由于两台仪器之间存在差异，所以在主仪器视角下观测的光谱数据与从仪器视角的观测值不同。因此，对于预测值中的每一个成分，观测结果也是不同的。

我们通过预测值之间的线性趋势来寻找主从仪器之间的偏差，所以此方法针对的浓度值有两种。通过上面的分析可知，在从仪器上面测得的两个预测值之间的系数和主仪器之间也存在一定的偏差。我们需要校正主从仪器之间的偏差，求得和理想值更为接近的从仪器的预测值。

假定主仪器的响应值包含  $k$  种物质，那么从  $k$  种物质中取出一种物质，有  $k$  种取法。对  $i=0, 1, 2, \dots, k-1$  分别进行实验，会得到  $k$  组结果。当  $i$  取某一个值时，对应的  $j$  可以取除了  $i$  之外，从  $0$  到  $k-1$  的任何值，一共可以得到  $k-1$  组结果。

首先，需要从  $k$  列值中取出两列。根据图 2.3 所表示的方法从  $(\hat{\mathbf{Y}}^m, \tilde{\mathbf{Y}}^s)$  中提取出任

意两列  $(\hat{\mathbf{Y}}_i^m, \hat{\mathbf{Y}}_j^m)$  和  $(\tilde{\mathbf{Y}}_i^s, \tilde{\mathbf{Y}}_j^s)$ 。

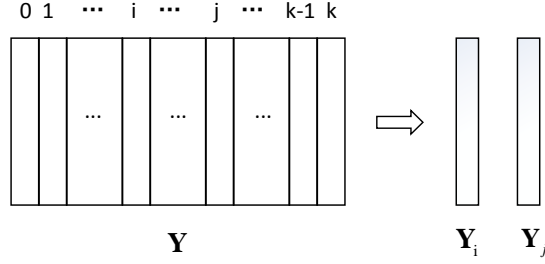


图 2.3 从  $\mathbf{Y}$  中任取两列

在对偏差进行校正之前，为了减少预测误差，我们需要对主从仪器的两组预测值进行预处理，也就是归一化处理。归一化实质就是将数据变换到[0,1]范围内，以此保持相同的尺度。下面给出了归一化公式：

$$\mathbf{Y}_i = (\mathbf{Y}_i - \min(\mathbf{Y}_i)) / (\max(\mathbf{Y}_i) - \min(\mathbf{Y}_i)) \quad (24)$$

经过预处理之后的浓度预测值  $(\hat{\mathbf{Y}}_i^m, \hat{\mathbf{Y}}_j^m)$  与  $(\tilde{\mathbf{Y}}_i^s, \tilde{\mathbf{Y}}_j^s)$ ，可以通过最小二乘法(Least Squares)在主从仪器下构建模型：

$$\begin{aligned} \hat{\mathbf{Y}}_j^m &= \hat{\mathbf{Y}}_i^m \tan \theta^m + b^m \\ \tilde{\mathbf{Y}}_j^s &= \tilde{\mathbf{Y}}_i^s \tan \theta^s + b^s \end{aligned} \quad (25)$$

其中， $\tan \theta^m$  为主仪器的两个预测值之间的回归系数， $b^m$  为回归直线的偏差。 $\tan \theta^s$  为主仪器视角下求得的从仪器的“有偏见”的系数， $b^s$  对应“有偏见”的偏差。

通过下面的步骤对上述斜率和截距的偏差  $\Delta\theta$  和  $\Delta b$  进行求解：

- (1) 首先对  $(\hat{\mathbf{Y}}_{i\_norm}^m, \hat{\mathbf{Y}}_{j\_norm}^m)$  线性回归，得到主仪器视角下主仪器的夹角和偏差。

主仪器 PLS2 模型的回归系数如下：

$$\mathbf{B}^m = \mathbf{W}_m (\mathbf{P}_m^T \mathbf{W}_m)^{-1} \mathbf{Q}_m^T \quad (26)$$

由主仪器模型的回归系数，求主仪器光谱的预测值公式如下：

$$\hat{\mathbf{Y}}^m = \mathbf{X}^m \mathbf{B}^m \quad (27)$$

之后根据图 2.3 的方式，从  $\hat{\mathbf{Y}}^m$  中提取出第  $i$  列和第  $j$  列，并进行归一化预处理，得到  $(\hat{\mathbf{Y}}_{i\_norm}^m, \hat{\mathbf{Y}}_{j\_norm}^m)$ 。

根据预处理之后的两个浓度的预测值，利用最小二乘法计算相应的斜率。公式如下：

$$\tan \theta_{ij}^m = ((\hat{\mathbf{Y}}_{i\_norm}^m)^T \hat{\mathbf{Y}}_{i\_norm}^m)^{-1} \hat{\mathbf{Y}}_{i\_norm}^m \hat{\mathbf{Y}}_{j\_norm}^m \quad (28)$$

$\tan \theta_{ij}^m$  为主仪器视角下，主仪器的 i 与 j 两个浓度预测值之间的回归系数。

根据回归系数，就可以求得对应的偏差：

$$\mathbf{b}_{ij}^m = \hat{\mathbf{Y}}_{j\_norm}^m - \hat{\mathbf{Y}}_{i\_norm}^m * \tan \theta_{ij}^m \quad (29)$$

(2) 然后对  $(\hat{\mathbf{Y}}_{i\_norm}^m, \hat{\mathbf{Y}}_{j\_norm}^m)$  进行线性回归，得到主仪器视角下从仪器的夹角和偏差。

根据主仪器的 PLS2 回归系数，求得从仪器的伪预测值如下：

$$\tilde{\mathbf{Y}}^s = \mathbf{X}^s \mathbf{B}^m \quad (30)$$

同样，从  $\tilde{\mathbf{Y}}^s$  中提取出第 i 列和第 j 列，并进行归一化预处理，得到  $(\tilde{\mathbf{Y}}_{i\_norm}^s, \tilde{\mathbf{Y}}_{j\_norm}^s)$ 。

和公式(2.28)相同，计算相应的斜率。公式如下：

$$\tan \theta_{ij}^s = ((\tilde{\mathbf{Y}}_{i\_norm}^s)^T \tilde{\mathbf{Y}}_{i\_norm}^s)^{-1} \tilde{\mathbf{Y}}_{i\_norm}^s \tilde{\mathbf{Y}}_{j\_norm}^s \quad (31)$$

$\tan \theta_{ij}^s$  为主仪器视角下，从仪器的 i 与 j 两个浓度预测值之间的回归系数。

根据回归系数，就可以求得对应的偏差：

$$\mathbf{b}_{ij}^s = \tilde{\mathbf{Y}}_{j\_norm}^s - \tilde{\mathbf{Y}}_{i\_norm}^s * \tan \theta_{ij}^s \quad (31)$$

(3) 最后，求得主从仪器两组系数之间的夹角和偏差。公式如下：

$$\begin{aligned} \Delta \theta &= \theta_{ij}^m - \theta_{ij}^s \\ \Delta \mathbf{b} &= \mathbf{b}_{ij}^m - \mathbf{b}_{ij}^s \end{aligned} \quad (32)$$

因为仿射变换的变换中心是坐标原点，，所以每个样本点在进行旋转之前，需要调整旋转中心到坐标原点，旋转后在转换到原坐标系中。利用前面计算到的夹角和偏差，对伪预测值进行转换。转换公式如下：

$$\hat{\mathbf{Y}}_i^s = \tilde{\mathbf{Y}}_i^s * \sin \Delta \theta + (\tilde{\mathbf{Y}}_j^s - \mathbf{b}_{ij}^s) * \cos \Delta \theta + \mathbf{b}_{ij}^m \quad (33)$$

当 i 取某一个值时，j 对应可以取除 i 之外的  $k-1$  个值。因此，可以得到  $k-1$  组结果。

对这些结果取均值，最终得到从仪器第 i 个组分的预测值：

$$\hat{\mathbf{Y}}_{i\_one}^s = (\sum \tilde{\mathbf{Y}}_i^s * \sin \Delta \theta + (\tilde{\mathbf{Y}}_j^s - \mathbf{b}_{ij}^s) * \cos \Delta \theta + \mathbf{b}_{ij}^m) / (k-1) \quad (34)$$

其中，j 从 0 取到  $k-1$ ，且 j 不等于 i。

对应 i 从 0 到  $k-1$ ，一共可以求得 k 个  $\hat{\mathbf{Y}}_{i\_one}^s$ 。将所有的结果放置到一个矩阵中，就可以求得从仪器被校正之后的预测值：

$$\hat{\mathbf{Y}}^s = [\hat{\mathbf{Y}}_{0\_one}^s, \hat{\mathbf{Y}}_{1\_one}^s, \dots, \hat{\mathbf{Y}}_{K-1\_one}^s] \quad (35)$$

最终，根据从仪器的输入  $\mathbf{X}^s$  与校正之后的预测值  $\hat{\mathbf{Y}}^s$ ，可以求得从仪器模型的回归系数：

$$\mathbf{B}^s = (\mathbf{X}^s)^{-1} [\hat{\mathbf{Y}}_{0\_one}^s, \hat{\mathbf{Y}}_{1\_one}^s, \dots, \hat{\mathbf{Y}}_{K-1\_one}^s] \quad (36)$$

## 2.5 MLCTAI 算法概要

### 2.5.1 算法过程

本算法中，需要用到的数据集包括三部分，分别为主仪器的光谱及对应的标签  $(\mathbf{X}_{cal}^m, \mathbf{Y}_{cal}^m)$ 、从仪器的光谱  $\mathbf{X}_{cal}^s$  和从仪器光谱的测试集  $\mathbf{X}_{test}^s$ 。本算法一共有三步，分别是主仪器的建模，仿射变换模型的建立和预测阶段。

1. 主仪器 PLS2 建模。对主仪器光谱的校准集进行 PLS2 建模，可以求得 PLS2 模型中相应的系数  $\mathbf{B}^m$ 。

$$(\mathbf{X}_{cal}^m, \mathbf{Y}_{cal}^m) \xrightarrow{\text{PLS2}} \mathbf{B}^m \quad (37)$$

2. 仿射变换模型的建立。这一部分用到的数据有  $(\mathbf{X}_{cal}^m, \mathbf{X}_{cal}^s)$ 。我们的目的是计算主仪器的预测值之间的回归系数和对应的从仪器的系数之间的夹角和偏差。

(1) 根据主仪器 PLS 模型求得的参数，计算在主仪器视角下观测到的主仪器的预测值和从仪器的伪预测值。

$$\begin{aligned} \hat{\mathbf{Y}}_{cal}^m &= \mathbf{X}_{cal}^m \mathbf{B}^m \\ \tilde{\mathbf{Y}}_{cal}^s &= \mathbf{X}_{cal}^s \mathbf{B}^m \end{aligned} \quad (38)$$

(2) 根据图 2.3 从  $(\hat{\mathbf{Y}}_{cal}^m, \tilde{\mathbf{Y}}_{cal}^s)$  抽取第  $i$  列和第  $j$  列，分别得到  $(\hat{\mathbf{Y}}_{i\_cal}^m, \hat{\mathbf{Y}}_{j\_cal}^m)$  与  $(\tilde{\mathbf{Y}}_{i\_cal}^s, \tilde{\mathbf{Y}}_{j\_cal}^s)$ 。为了使主仪器和从仪器的预测值之间具有相同的尺度，需要分别对主从仪器的预测值进行归一化处理，使每一个值都处在 0 到 1 之间。

$$\begin{aligned} (\hat{\mathbf{Y}}_{i\_cal}^m, \hat{\mathbf{Y}}_{j\_cal}^m) &\xrightarrow{\text{normalization}} (\hat{\mathbf{Y}}_{i\_norm}^m, \hat{\mathbf{Y}}_{j\_norm}^m) \\ (\tilde{\mathbf{Y}}_{i\_cal}^s, \tilde{\mathbf{Y}}_{j\_cal}^s) &\xrightarrow{\text{normalization}} (\tilde{\mathbf{Y}}_{i\_norm}^s, \tilde{\mathbf{Y}}_{j\_norm}^s) \end{aligned} \quad (39)$$

(3) 分别计算主从仪器预测值之间的回归系数，求得回归系数的角度和偏差。

$$\begin{aligned} (\hat{\mathbf{Y}}_{i\_norm}^m, \hat{\mathbf{Y}}_{j\_norm}^m) &\xrightarrow{\text{least square}} \theta_{ij}^m, \mathbf{b}_{ij}^m \\ (\tilde{\mathbf{Y}}_{i\_norm}^s, \tilde{\mathbf{Y}}_{j\_norm}^s) &\xrightarrow{\text{least square}} \theta_{ij}^s, \mathbf{b}_{ij}^s \end{aligned} \quad (40)$$

(4) 根据求得的参数  $(\theta_{ij}^m, \mathbf{b}_{ij}^m)$  与  $(\theta_{ij}^s, \mathbf{b}_{ij}^s)$ ，计算主从仪器之间的夹角和偏差。

$$\begin{aligned} \Delta\theta &= \theta_{ij}^m - \theta_{ij}^s \\ \Delta\mathbf{b} &= \mathbf{b}_{ij}^m - \mathbf{b}_{ij}^s \end{aligned} \quad (41)$$

3. 预测阶段。此部分使用到从仪器光谱的测试样本。根据第二部分求得的一组夹角和偏差，将从仪器光谱的测试样本仿射变换到主仪器，求得更为接近测量值的预测值。

(1) 首先根据主仪器 PLS 标定模型参数求得主仪器视角下的预测值：

$$\tilde{\mathbf{Y}}_{\text{test}}^s = \mathbf{X}_{\text{test}}^s \mathbf{B}^m \quad (42)$$

(2) 根据图 2.3 的方式从  $\tilde{\mathbf{Y}}_{\text{test}}^s$  中提取出第  $i$  列和第  $j$  列 ( $\tilde{\mathbf{Y}}_{i\_test}^s, \tilde{\mathbf{Y}}_{j\_test}^s$ )。并进行归一化操作，使之和仿射变换建模过程有相同的尺度。

$$(\tilde{\mathbf{Y}}_{i\_test}^s, \tilde{\mathbf{Y}}_{j\_test}^s) \xrightarrow{\text{normalization}} (\tilde{\mathbf{Y}}_{i\_norm\_test}^s, \tilde{\mathbf{Y}}_{j\_norm\_test}^s) \quad (43)$$

(3) 根据公式(2.10)与公式(2.11)，用于仿射变换的增广矩阵  $\tilde{\mathbf{U}}_{\text{test}}$  以及变换矩阵  $\mathbf{M}$ ， $(\tilde{\mathbf{Y}}_{i\_norm\_test}^s, \tilde{\mathbf{Y}}_{j\_norm\_test}^s)$  的缩放因子在下面的公式中给出：

$$\tilde{\mathbf{U}}_{\text{test}} = [\tilde{\mathbf{Y}}_{i\_norm\_test}^s, \tilde{\mathbf{Y}}_{j\_norm\_test}^s, -\mathbf{b}_{ij}^s, \mathbf{1}] \quad (44)$$

$$\mathbf{M} = \begin{bmatrix} \lambda_i \cos \Delta\theta & \lambda_i \sin \Delta\theta & 0 \\ -\lambda_j \sin \Delta\theta & \lambda_j \cos \Delta\theta & 0 \\ 0 & \mathbf{b}_{ij}^m & 1 \end{bmatrix} \quad (45)$$

$$\begin{cases} \lambda_i = \left[ \frac{\tilde{\mathbf{Y}}_{i\_norm\_test}^s - \min(\tilde{\mathbf{Y}}_i^s)}{\max(\tilde{\mathbf{Y}}_i^s) - \min(\tilde{\mathbf{Y}}_i^s)} + \min(\tilde{\mathbf{Y}}_i^s) \right] * (\max(\tilde{\mathbf{Y}}_i^s) - \min(\tilde{\mathbf{Y}}_i^s)) \\ \lambda_j = \left[ \frac{\tilde{\mathbf{Y}}_{j\_norm\_test}^s - \min(\tilde{\mathbf{Y}}_j^s)}{\max(\tilde{\mathbf{Y}}_j^s) - \min(\tilde{\mathbf{Y}}_j^s)} + \min(\tilde{\mathbf{Y}}_j^s) \right] * (\max(\tilde{\mathbf{Y}}_j^s) - \min(\tilde{\mathbf{Y}}_j^s)) \end{cases} \quad (46)$$

(4) 根据下面的公式，对  $(\tilde{\mathbf{Y}}_{i\_norm\_test}^s, \tilde{\mathbf{Y}}_{j\_norm\_test}^s)$  进行旋转平移，平移到主仪器，得到一个增广矩阵  $\hat{\mathbf{U}}_{\text{corr}}$ ：

$$\hat{\mathbf{U}}_{\text{corr}} = \tilde{\mathbf{U}}_{\text{test}} \mathbf{M} \quad (47)$$

其中， $\hat{\mathbf{U}}_{\text{corr}} = [\hat{\mathbf{Y}}_{i\_norm\_test}^s, \hat{\mathbf{Y}}_{j\_norm\_test}^s, \mathbf{1}]$ 。 $(\hat{\mathbf{Y}}_{i\_norm\_test}^s, \hat{\mathbf{Y}}_{j\_norm\_test}^s)$  表示已经校正的从仪器预测值。

(5) 对  $(\hat{\mathbf{Y}}_{i\_norm\_test}^s, \hat{\mathbf{Y}}_{j\_norm\_test}^s)$  逆归一化，得到主仪器预测值对应的尺度。

$$(\hat{\mathbf{Y}}_{i\_norm\_test}^s, \hat{\mathbf{Y}}_{j\_norm\_test}^s) \xrightarrow{\text{inverse normalization}} (\hat{\mathbf{Y}}_{i\_test}^s, \hat{\mathbf{Y}}_{j\_test}^s) \quad (48)$$

此处的  $(\hat{\mathbf{Y}}_{i\_test}^s, \hat{\mathbf{Y}}_{j\_test}^s)$  为最终的预测值。

## 2.5.2 算法流程图

图 2.4 表示 MLCTAI 算法的具体实现过程。



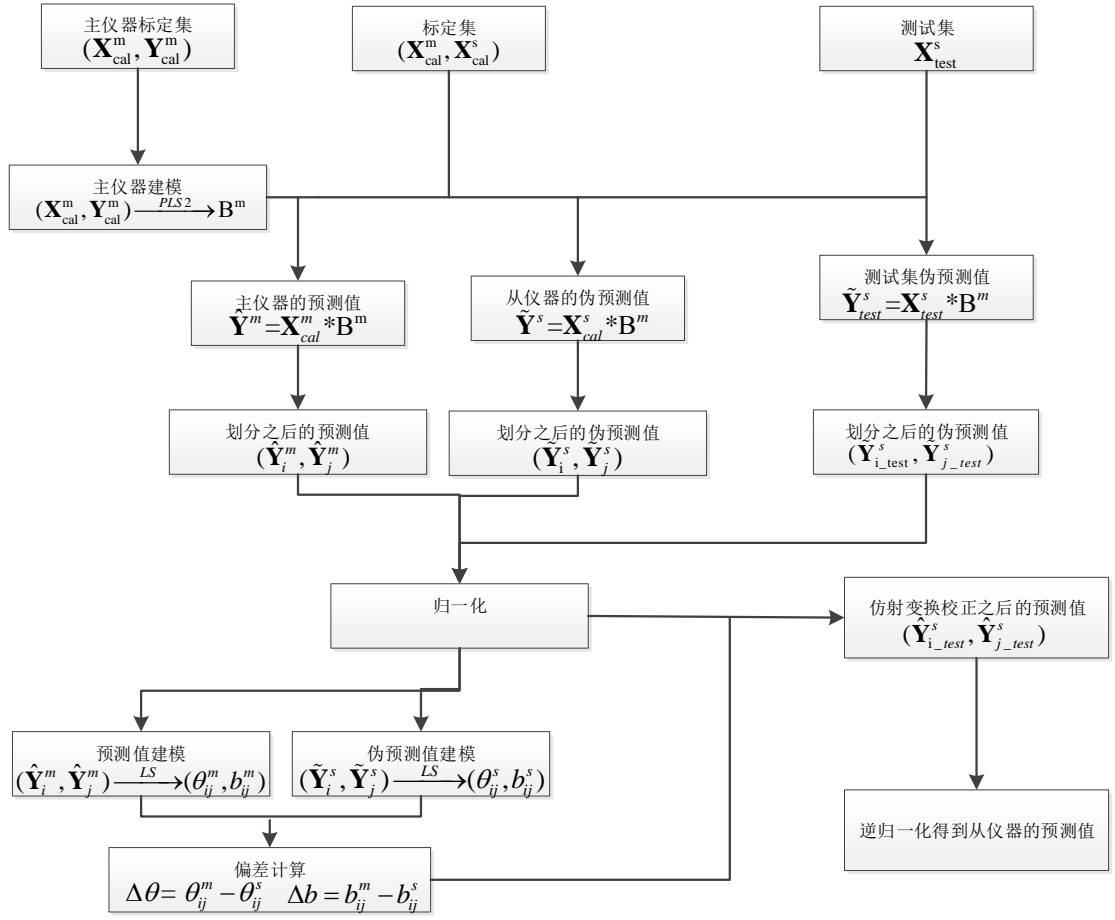


图 2.4 MLCTAI 方法的具体实现过程

### 3 实验

#### 3.1 实验环境

实验所用的所有程序均在个人电脑上实现，使用的语言是 python，python 版本为 2.7.15。个人电脑配置为：Intel(R) Core(TM) i5-5200U CPU @2.20GHz 2.19GHz，运行内存(RAM)为 8.00GB。电脑系统为 windows 8.1 64 位操作系统。

#### 3.2 实验数据

在本实验中，使用到两个公开可用的数据集来检验 MLCTAI 方法的性能。第一个使用的数据集是玉米数据集，该数据集包含了波长在 1100-2400nm 范围内，且波长间隔为 2nm,700 频道的三台仪器上测量得到的 80 个样本。每一个样本里面都包含了四个成分：水分，油，蛋白质，淀粉。在本实验中，四种成分都参与实验。在 m5spec 仪器上面测量的光谱作为主仪器，在 mp6spec 仪器上面测量的光谱作为从仪器。图 3.1(A)(B)显示了主从仪器的近红外光谱。为体现主从仪器的不同，特意从主从仪器样本里面选择了第 41 个样本，图 3.1(C)显示了主从仪器该样本的不同之处。数据集通过 KS 算法<sup>[46]</sup>被划分为拥有 64 个样本的校准集和拥有 16 个样本的测试集。第二个使用到的数据集是一个公开可获取的药片数据集。它由在透射模式下的两台 Multitab 光谱仪(FossNIRsystems, Silverspring, MD)测量得到的 655 个药片剂样品的 1308 个光谱组成，其波长范围为 600 至 1898nm，间隔为 2nm。它一共包含三组数据，分别是 155 个标定样本,460 个测试样本，还有 40 个验证样本。在本实验中，使用到的是 test 仪器(test\_1,test\_2)上测量得到的 460 个测试样本(test\_Y)。在每一个样本里面，包含了三个成分：weight、hardness、assay。在本实验中，只有 hardness 与 assay 参与实验。在 test\_1 仪器上测量的光谱作为主仪器，在 test\_2 仪器上测量的光谱作为从仪器。图 4(D)(E)展示了主从仪器的光谱。和玉米集相同，为体现光谱的不同之处，特地选择了主从仪器光谱中的第 126 个样本，图 4(F)显示了这两个样本。根据 KS 算法进行划分，368 个样本被划分为标定集，92 个样本被划分为测试集。

在本实验中，两个数据集的浓度数据均相同，而且均在同一条件下进行测量，测量得到之后均在同一条件下使用各种方法进行验证。

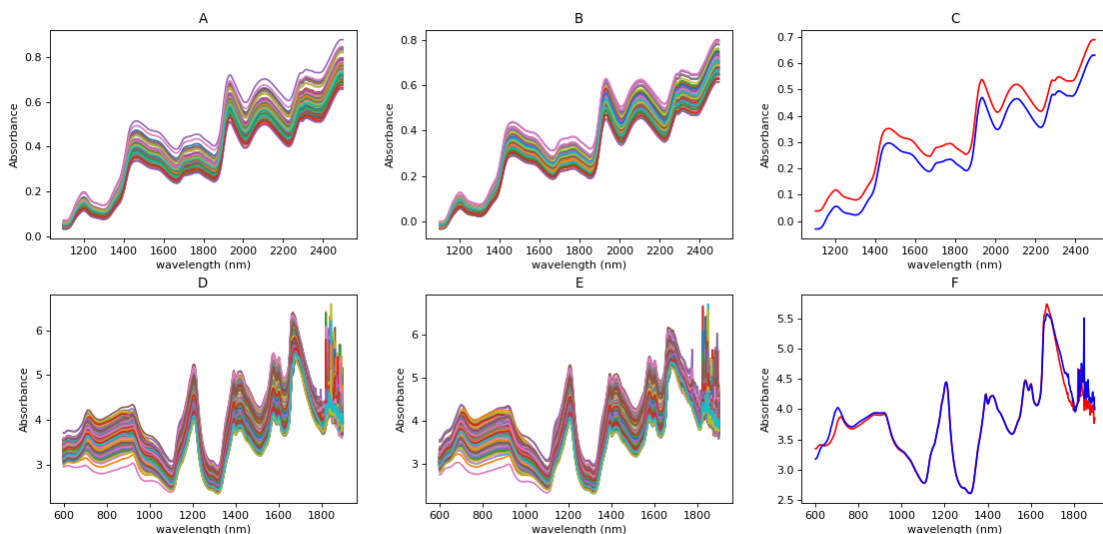


图 3.1 玉米集(A)m5spec,(B)mp6spec 仪器和药片集(D)test\_1,(E)test\_2 仪器的光谱图以及(C)玉米集第 41 个样本,(F)药片集第 126 个样本的光谱

### 3.3 最优参数选择

MLCT\_AI 方法中的潜变量数量变化范围为 1 到 15。最优的潜变量数则通过 10 折交叉验证获取。在每一次进行交叉验证的时候,都会得到一个交叉验证误差(RMSECV)。交叉验证误差最小时的潜变量数就是最佳的。

根据 MLCTAI 方法浓度值不止一列的特性,对每一个成分,MLCTAI 最终的实验结果数量等于成分数量减一,因为该成分要与除去自己的所有成分进行实验。所以对玉米集而言,一种成分有三个结果。对于药片集,一种成分有两个结果。为了从里面选出最佳的结果,从两个数据集的从仪器标定集中选出 50%-75%部分的数据作为确认集,计算其预测误差,将其作为判断的标准。从所有结果中,选择出确认集误差最小的一组成分组合,此组合将用于 MLCTAI 方法在此成分下的结果。与 MLCT\_AI 进行对比的方法一共有五个。其中一个为迁移成分回归(TCR)。在 TCR 方法中,TCA 空间的维度变化范围为 1 到 20。TCA 的最佳空间维度在变换到公共空间里面的两个矩阵之间的距离最小时取得。其余的四个方法中,SBC,MSC,CCA,在 PLS 交叉验证部分,潜变量数的变化范围也是 1 到 15。这三种方法均使用 10 折交叉验证来获取最佳潜变量数,和 MLCT\_AI 相同。PDS 方法在确定最佳窗口值时使用的是 10 折交叉验证的 PLS 模型,但在移动窗口创建的 PLS 模型中使用的是 5 折交叉验证的 PLS 模型。PDS 方法的窗口大小是从 3 变到 16,每次窗口大小增加 2。

对三种有标样的对比方法也进行了一系列的测试并从中选出最佳的结果。三种有标样的方法的标样数量范围从 10 到 30。标样的划分方法和数据集的划分方式相同，均使用 KS 算法划分。对三种方法的每个标样数量均进行测试，从 20 组中选出训练误差最小的一组作为方法最终的结果。

### 3.4 实验结果评估

在本模型中，一共采用三种评估方式来对结果以及参数进行评估。

1. 均方根误差(root mean square error, RMSE)。均方根误差被用做参数选择与模型评估。在本实验的方法里面用到了三种均方根误差。RMSECV 表示交叉验证的误差，通常用作选择最佳潜变量。RMSEC 表示训练集的训练误差。而 RMSEP 则是测试集的预测误差。均方根误差的计算方法如下：

$$RMSE = \sqrt{\frac{\sum (y - \hat{y})^2}{n}} \quad (49)$$

其中， $y$  表示测量值， $\hat{y}$  表示  $y$  的预测值， $n$  表示样本的数量。

2. 改善率(h)。引入这个评价指标是为更好地与其他方法进行对比，对比 MLCT\_AI 方法相对于其他的模型的预测误差改善情况，其计算公式如下：

$$h = (1 - \frac{RMSEP}{RMSEP_o}) \times 100\% \quad (50)$$

其中， $RMSEP$  代表 MLCT\_AI 方法的预测误差， $RMSEP_o$  则是其他方法的预测误差。

3. 秩和检验。为检测 MLCT\_AI 方法的结果是否与其他方法的结果有显著性差异，引入了 Wilcoxon 符号秩和检验，其置信度为 95%。当检验结果小于 0.05 时，表明有显著性差异，结果大于 0.05，则没有显著性差异。

## 4 实验结果

### 4.1 最优参数

所有算法在构造主仪器回归模型时，均使用 PLS 建模，MLCTAI 方法使用的是 PLS2 方法。PDS 方法在测试时，使用 PLS 模型计算变换函数。

在 MLCTAI 的实验中，用于决定一个成分的几组预测误差优劣的确认集被使用到了，也对其进行了测试，表 4.1 展示了两个数据集的确认集预测误差。表中第一列以及第六列表示 MLCTAI 中的第一个成分，第一行则表示第二个成分。从表中可以看出，玉米集最终用于实验的组合是：moisture-oil, oil-starch, protein-starch, starch-protein。药片集最终用于实验的组合是：hardness-assay, assay-hardness。这几个组合将用于后面的所有 MLCTAI 方法的实验中。

表 4.1 两个数据集的确认集预测误差

玉米	moisture	oil	protein	starch	药片	weight	hardness	assay
moisture		0.25979	0.30317	0.30025	hardness	0.96047		0.96044
oil	0.10798		0.08947	0.08328				
protein	0.25158	0.18610		0.16204	assay	5.18919	4.61600	
starch	0.41449	0.38431	0.36472					

表 4.2 显示六种方法在两个数据集的几种成分下的最佳潜变量数量(TCR 的最佳子空间维度)。当用玉米集的四个成分进行测试时，PDS 方法的最佳窗口值为 3, 5, 3, 3。当用药片集的两个成分进行测试时，PDS 方法的最佳窗口数均为 13, 15, 3。对于有标样的三个对比方法，其在实验中使用到的最佳标样数量也展示在表 4.2 中。其中，上标 a 表示有标样方法在该成分下的最佳标样数。

表 4.2 所有方法最佳潜变量数量(TCR 的最佳子空间数)和有标样方法的标样数量

		MLMCTAI	TCR	MSC	PDS		SBC		CCA	
玉米	moisture	14	5	14	14	29 <sup>a</sup>	14	23 <sup>a</sup>	14	23 <sup>a</sup>
	oil	14	6	15	15	27 <sup>a</sup>	15	27 <sup>a</sup>	15	21 <sup>a</sup>
	protein	14	10	15	15	23 <sup>a</sup>	15	14 <sup>a</sup>	15	30 <sup>a</sup>
	starch	14	9	15	15	23 <sup>a</sup>	15	30 <sup>a</sup>	15	30 <sup>a</sup>
药片	hardness	5	6	3	3	21 <sup>a</sup>	3	28 <sup>a</sup>	3	11 <sup>a</sup>

续表 4.2 所有方法最佳潜变量数量(TCR 的最佳子空间数)和有标样方法的标样数量

	MLMCTAI	TCR	MSC	PDS		SBC		CCA	
assay	5	7	5	5	29 <sup>a</sup>	5	14 <sup>a</sup>	5	10 <sup>a</sup>

注意：上标 a 表示该方法在该成分下的最佳标样数

## 4.2 MLCTAI 方法的结果

对于两个数据集，本节探索了 MLCTAI 方法的校正效果。在划分得到从仪器测试集的伪预测值之后，未校正的值直接使用伪预测值与测试集的测量值求解预测误差，校正的预测值则是先使用仿射变换转化伪预测值，再求解预测误差。图 5 展示了玉米集的 moisture, oil, protein, starch 成分，药片集的 hardness 与 assay 成分校正与未校正的预测值之间的关系。在表 4.3 中直接展示了伪预测值与 MLCTAI 方法的预测误差。

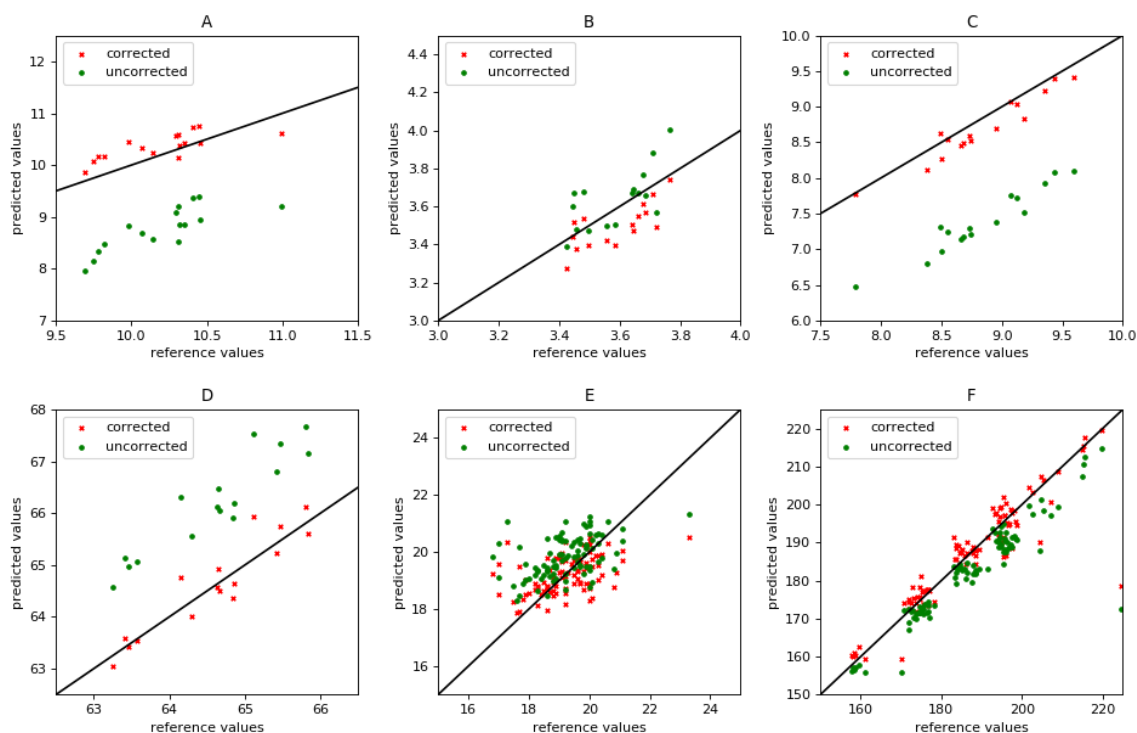


图 4.1 玉米集(A)moisture, (B)oil, (C)protein, (D)starch 与药片集(E)hardness, (F)assay 校正与未校正的预测值分布关系

从图中可以看出仿射变换使得校正后的预测值分布更加靠近直线，而未校正的预测值远离直线。从表中也可以知道，伪预测值校正之后的预测误差更小。二者同时说明仿射变换可以校正仪器之间的偏差。

表 4.3 伪预测值与 MLCTAI 方法的预测误差

	玉米集				药片集	
	moisture	oil	protein	starch	hardness	assay
uncorrected	1.43390	0.12563	1.44941	1.62845	1.11665	7.62779
corrected	0.27613	0.12485	0.18232	0.33714	0.92166	6.18287

注意：uncorrected 表示未校正的预测误差，corrected 表示已校正的预测误差

#### 4.3 相关方法比较结果

在本实验中使用到 5 个对比方法，MSC、TCR 是无标样的标定迁移方法，SBC、PDS、CCA 是有标样的标定迁移方法。本实验分别使用玉米数据集(corn dataset)与药片数据集(Pharmaceutical dataset)对 MLCTAI 方法与其余的五种方法的预测性能与预测稳定性进行测试。预测性能包括预测误差、改善率与 Wilcoxon 符号秩和检验。预测稳定性则是比较各个方法的预测值分布情况。因 MLCTAI 方法没有训练误差(从仪器缺少响应值)，故将确认集的预测值画到分布图之中。

表 4.4 展示了 MLCTAI 方法与其余五种对比方法的训练误差与预测误差。表 4.5 展示了五种比较方法的改善率和 Wilcoxon 秩和检验的结果。为了研究 MLCTAI 和五种比较方法的预测稳定性，两个数据集的标定集与测试集校正预测值的散点图如图 4.2 到图 4.7 所示。每一张图表示一种成分，图 4.2 到图 4.5 表示玉米集的四种成分，图 4.6 与图 4.7 表示药片集的后两种成分。直线和点之间的距离可用于判断预测结果是否良好，距离越近，结果越好。

表 4.4 MLCTAI 方法与其余对比方法的训练误差与预测误差

		无标样方法			有标样方法		
		MLCTAI	TCR	MSC	PDS	SBC	CCA
玉米	moisture	RMSEC		0.53472	1.92840	0.27809	0.35210
		RMSEP	0.27613	0.38057	1.66890	0.23703	0.33044
	oil	RMSEC		0.14135	1.21647	0.11071	0.08148
		RMSEP	0.12485	0.12551	1.23209	0.11194	0.08471
	protein	RMSEC		0.25852	1.77294	0.28609	0.22428
		RMSEP	0.18232	0.29475	1.80087	0.28193	0.18829
	starch	RMSEC		1.27479	1.89165	0.72842	0.39667
		RMSEP	0.33714	0.62664	1.93129	0.78977	0.37538

续表 4.4 MLCTAI 方法与其他对比方法的训练误差与预测误差

		MLCTAI	TCR	MSC	PDS	SBC	CCA
药片	hardness	RMSEC	0.90771	1.23382	0.89227	0.92382	1.24778
		RMSEP	0.92166	0.94503	1.42026	0.95934	0.95363
	assay	RMSEC	4.75850	9.81891	3.795474	6.81395	10.98652
		RMSEP	6.18287	6.37243	7.84073	5.93626	7.58534

注意：因 MLCTAI 方法中从仪器没有响应值，故 MLCTAI 方法没有训练误差(RMSEC)。

表 4.5 MLCTAI 方法与其他对比方法的改善率和 Wilcoxon 秩和检验结果

		玉米				药片	
		moisture	oil	protein	starch	hardness	assay
TCR	h(%)	27.44272	0.52510	38.14375	46.19778	2.471993	2.97478
	p	0.02289	0.21460	0.917632	0.67912	6.41E-15	0.000444
MSC	h(%)	83.45451	89.86667	89.87591	82.54281	35.10589	21.1442
	p	0.00044	0.00044	0.000438	0.00044	8.15E-17	8.7E-17
PDS	h(%)	-16.49488	-11.52643	35.33151	57.31070	3.927406	-4.15421
	p	0.16267	0.00320	0.005234	0.91763	8.15E-17	0.090295
SBC	h(%)	16.43721	-47.38482	3.173579	10.18556	3.351898	18.48923
	p	0.87672	0.00378	0.004455	0.27753	8.15E-17	8.15E-17
CCA	h(%)	-71.83142	14.31612	49.18478	47.71810	12.39262	45.65548
	p	0.05572	0.40804	0.836139	0.40804	1.69E-13	5.16E-15

4.3.1 无标样方法

首先是对预测性能的对比。对于玉米集，所有成分下，MLCTAI 方法的误差小于 TCR 与 MSC 方法的误差。本方法相对于两个方法均有提高，其中 MSC 提高较大，提高比例分别为：83.45%，89.86%，89.87%，82.54%，TCR 只有在 oil 成分下提高率较低。在 95% 的置信度下，MSC 方法与本方法在所有成分下均存在显著性差异，TCR 方法只有在 moisture 成分下与 MLCTAI 方法存在显著性差异，其余成分均不存在显著性差异。对药片集，MLCTAI 方法相比于其余两种无标样方法，误差是最小的，比 TCR 略小一些，比 MSC 小很多。在改善率方面，两种成分下 TCR 的改善率都较低，而 MSC 方法的改善率较高，为：35.10%、21.14%。在 95% 的置信度与两种成分下，两种方法与 MLCTAI 方法均存在着显著性差异。综合上述，就预测性能而言，MLCTAI 无疑比两种无标样的方



法更优秀。

其次是对预测稳定性的对比。从图 4.2-图 4.7(A)(B)(C)看来，对玉米集而言，在四种成分下，MLCTAI 方法的预测值分布较为稳定，均匀地分布在直线周围。对于 TCR 方法，在四种成分下，其预测分布点比较散，分布结果均不如 MLCTAI 方法。对于 MSC 方法，当成分是 moisture 时，分布比较松散，其余成分下，预测分布比较紧致。但四种成分下的预测值都远离真实值，其结果远远不如 MLCTAI。对药片集，MLCTAI 方法的预测值分布较为靠近直线，其中 TCR 方法的预测值分布和 MLCTAI 比较接近，MSC 在两种成分下的分布都比较分散，不如 MLCTAI。从上面的分析可得知，MLCTAI 方法预测值分布一直都较为稳定，TCR 方法时好时坏，MSC 方法则一直都很差。因而，就预测性能而言，MLCTAI 比两种无标样的方法更好。

从整体效果来看，无论是预测性能还是预测稳定性，MLCTAI 都优于其它两个无标样方法。

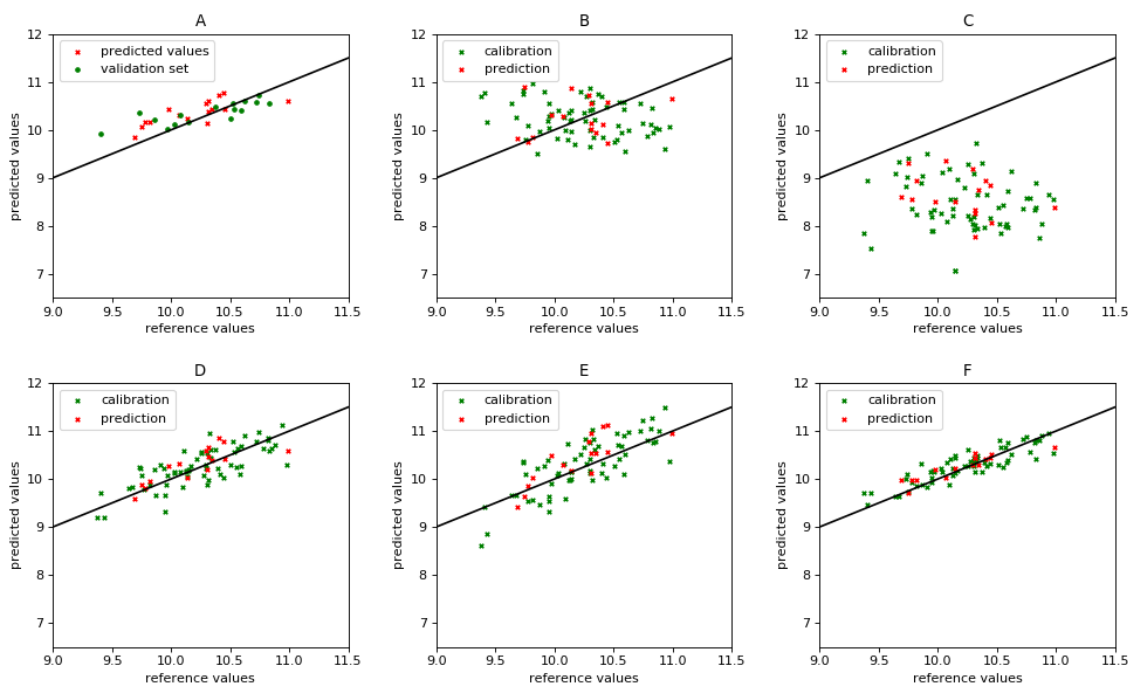


图 4.2 (A)MLCTAI, (B)TCR, (C)MSC, (D)PDS, (E)SBC, (F)CCA 方法在玉米集中 moisture 成分下的预测值分布

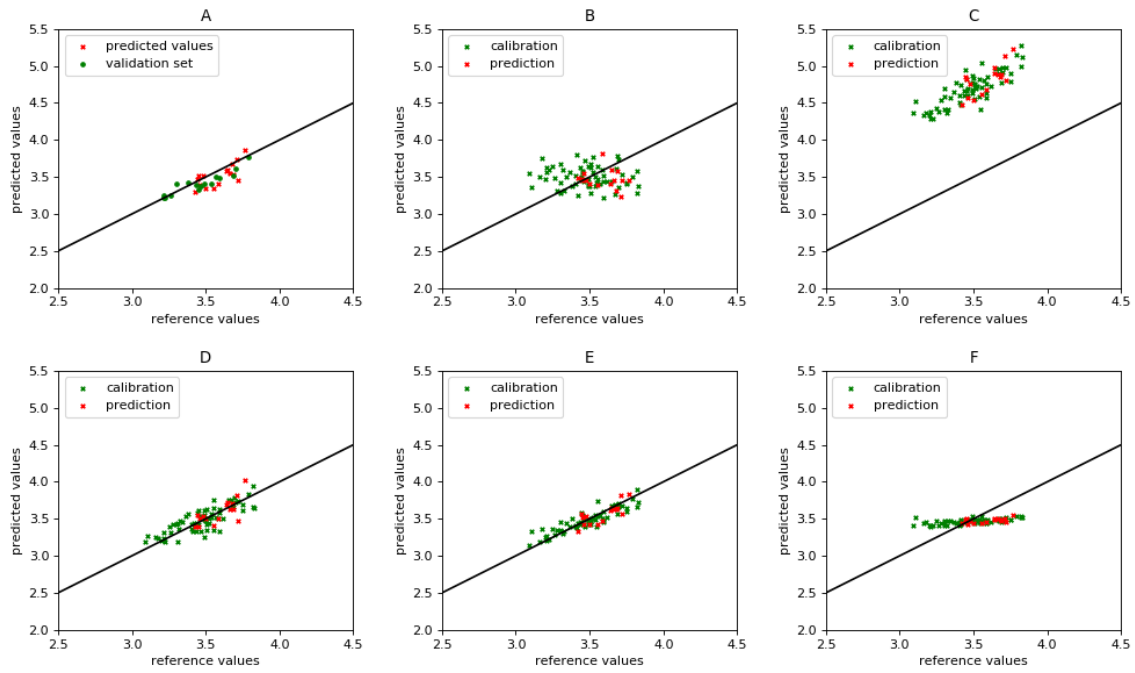


图 4.3 (A)MLCTAI, (B)TCR, (C)MSC, (D)PDS, (E)SBC, (F)CCA 方法在玉米集中 oil 成分下的预测值分布

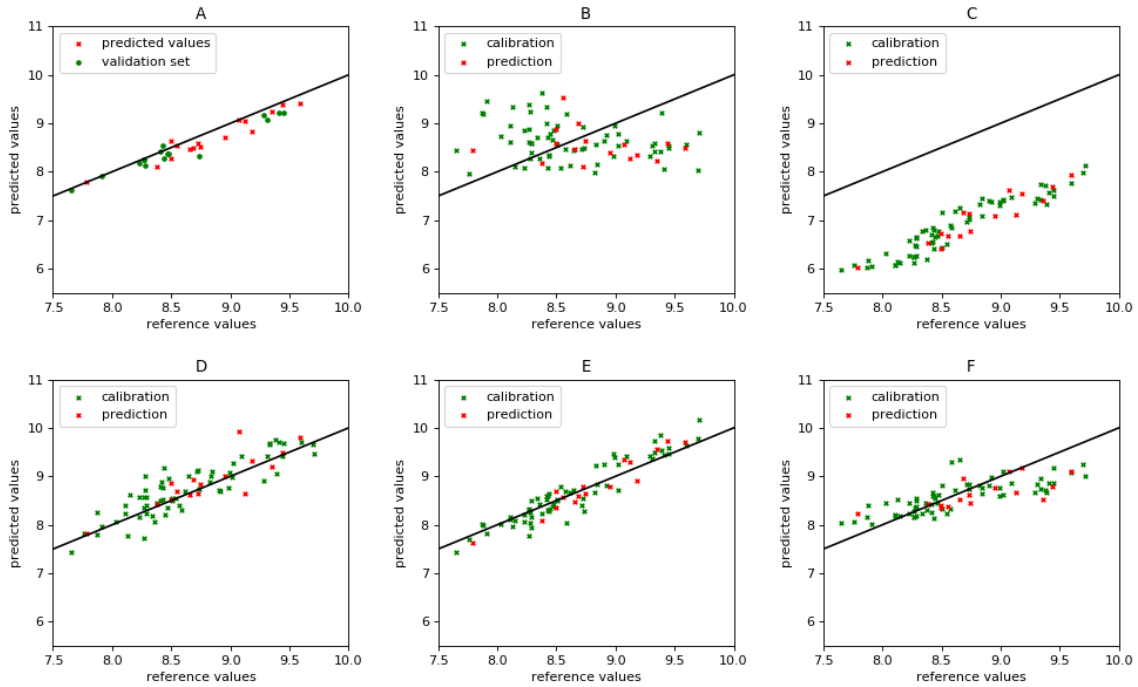


图 4.4 (A)MLCTAI, (B)TCR, (C)MSC, (D)PDS, (E)SBC, (F)CCA 方法在玉米集中 protein 成分下的预测值分布

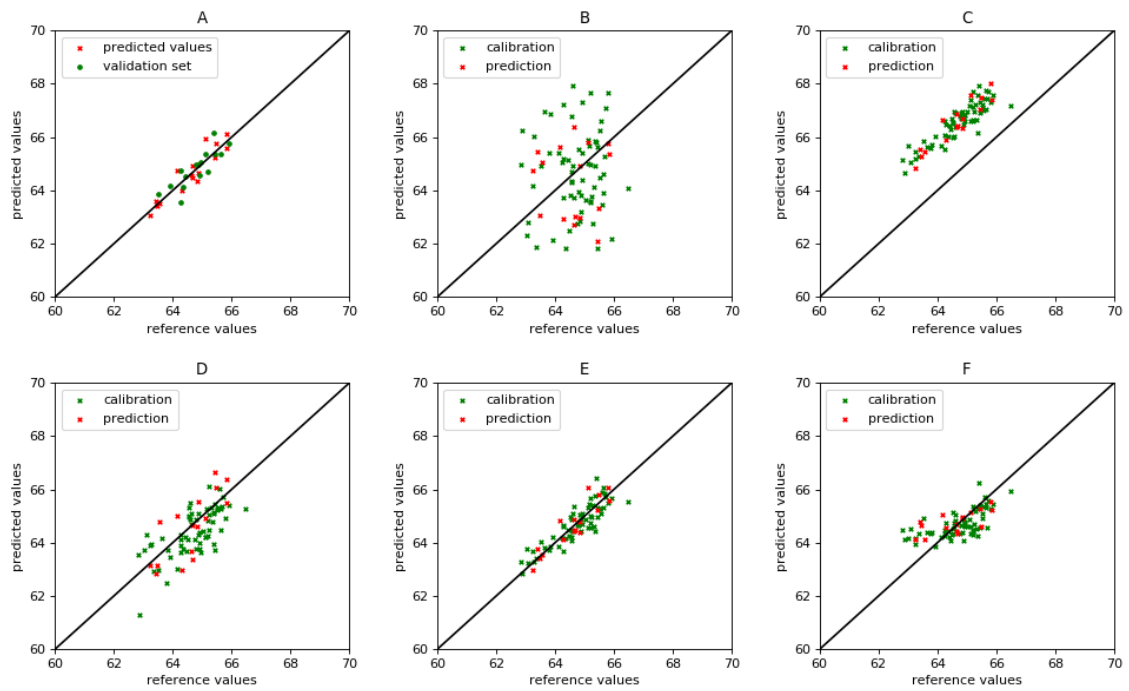


图 4.5 (A)MLCTAI, (B)TCR, (C)MSC, (D)PDS, (E)SBC, (F)CCA 方法在玉米集中 starch 成分下的预测值分布

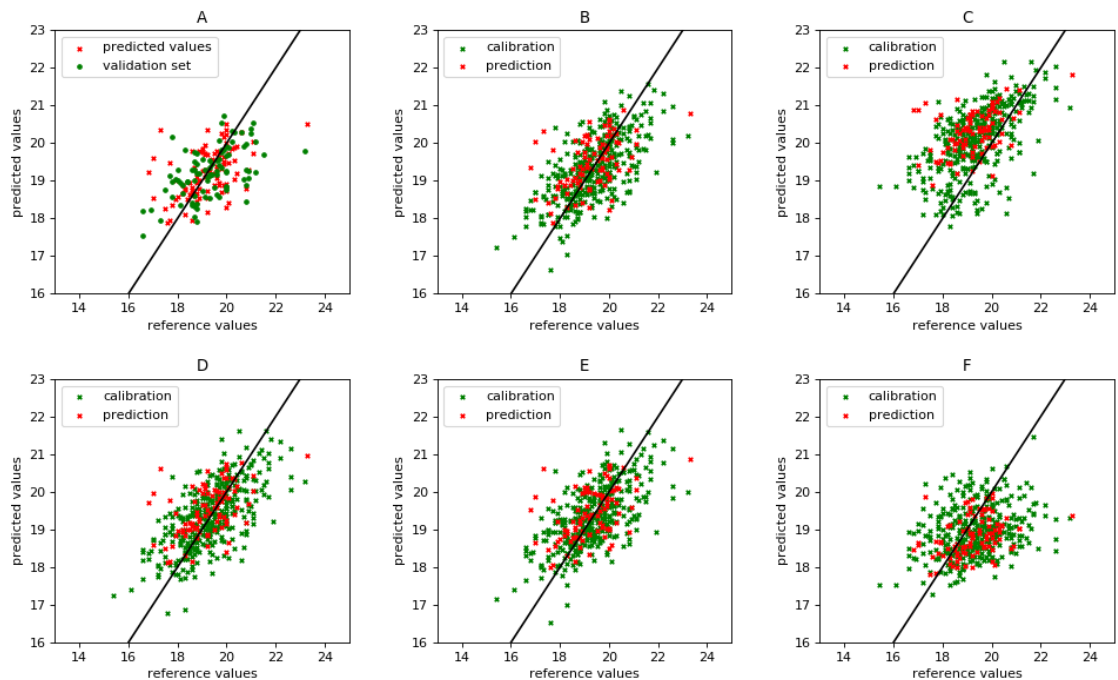


图 4.6 (A)MLCTAI, (B)TCR, (C)MSC, (D)PDS, (E)SBC, (F)CCA 方法在药片集中 hardness 成分下的预测值分布

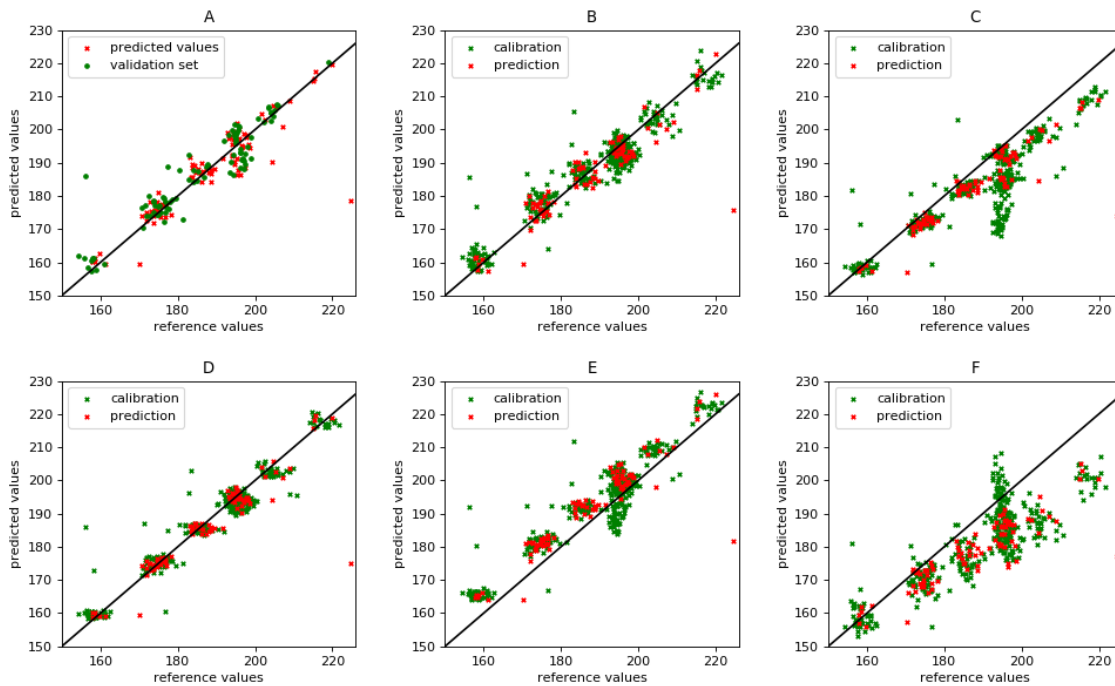


图 4.7 (A)MLCTAI, (B)TCR, (C)MSC, (D)PDS, (E)SBC, (F)CCA 方法在药片集中 assay 成分下的预测值分布

#### 4.3.2 与有标样的方法比较

首先是预测性能的比较。对于玉米集，当成分为 moisture 时，SBC 方法的预测误差是远大于 MLCTAI 方法的，其余两种方法的预测误差小于 MLCTAI。但 PDS 与 CCA 方法和 MLCTAI 之间不存在显著性差异。当成分为 oil 时，MLCTAI 方法的预测误差略大于 CCA 方法的误差。MLCTAI 方法与 PDS、SBC 方法之间存在显著性差异，与 CCA 方法之间则不存在显著性差异。当成分为 protein 时，MLCTAI 方法的预测误差略大于 SBC，远远大于 PDS 和 CCA。MLCTAI 相对于 CCA 的改善率较高。MLCTAI 方法相对于 PDS 与 SBC 存在显著性差异。当成分为 starch 时，MLCTAI 方法的预测误差也是最小的。MLCTAI 相对于 PDS、CCA 的改善率较高，为 57.31%、47.71%。MLCTAI 方法与三者之间不存在显著性差异。对于药片集，当成分为 hardness 时，MLCTAI 方法的预测误差最小，相对于其他三种方法改善率不高。但在 95%的置信度下，均存在显著性差异。当成分为 assay 时，MLCTAI 方法的预测误差略大于 PDS，但不存在显著性差异。SBC 与 CCA 方法的预测误差均比 MLCTAI 大，且存在显著性差异。综上所述，虽然有少数几组结果 MLCTAI 不如其他方法，但 MLCTAI 方法整体的预测性能是比三种有标样方法优秀的。

然后是预测稳定性的比较。对玉米集，从图 4.2-图 4.7(A)(D)(E)(F)可知，在四种成分下，MLCTAI 方法的预测值分布较为稳定，均匀地分布在直线四周。PDS 方法在成分为 moisture 与 oil 时的分布和 MLCTAI 基本一致，其余成分下，分布略微松散，不如 MLCTAI 方法均匀。SBC 方法在成分为 moisture 时，其预测值分布较 MLCTAI 松散。其余成分下，和 MLCTAI 一样，分布较为靠近直线。CCA 方法在成分为 moisture 时，其预测分布十分好，基本超过了 MLCTAI。在成分为 oil 时，其分布更加规律，但其趋向偏离了直线，故效果不如 MLCTAI。其余两种成分下，CCA 预测分布相比 MLCTAI 较为松散。对药片集，PDS 方法在两种成分下的分布基本和 MLCTAI 一直，差距很小。SBC 方法在成分为 hardness 时和 MLCTAI 方法的分布基本相同，当成份为 assay 时略微有些偏离直线。CCA 方法在两种成分下的分布都较为松散，不如 MLCTAI。综上所述，对比方法在有的成分下分布结果很好，但有的成分下分布结果较差，相比而言 MLCTAI 的稳定性是最好的。

在与有标样的方法的对比中，当成分为 oil 时，MLCTAI 方法的预测误差小于 PDS 与 SBC，且还存在显著性差异，MLCTAI 方法在这两个情况下是不如有标样方法的。但从整体而言，MLCTAI 方法的预测性能都较好，而且预测稳定性也很好。而且对于有标样的方法，标样的采集与保存比较困难。总体而言，MLCTAI 方法相比于这三种有标样方法有很大优势。

#### 4.4 工作展望

虽然本文提出的 MLCTAI 算法已经取得了较好的预测效果，但是仍然存在不足。在上述实验中，某些实验结果虽然比其他的方法的结果好，但是其结果本身偏大，是否可以通过某些方式，比如预处理方式，将数据进行处理，之后再进行实验，可以使结果适当小于现有结果。还存在几组比较特殊的实验结果，在该实验下，MLCTAI 方法的性能不如其他的方法，比如在玉米集的 oil 成分下，MLCTAI 方法的预测性能远远不如 PDS 与 SBC。那么是否可以通过一些处理方式，能够保证在其他的结果基本保持一致的同时，能够减小当前的这几组结果，让它优于 PDS 与 SBC。同时还有其他的一些部分也有不足之处，这将是下一步工作研究的方向。

## 结    论

本文提出了无标样的基于仿射变换的多标签标定迁移方法(MLCTAI)，其探索仪器上预测值的不同成分间的线性趋势。五种对比方法和 MLCTAI 方法在同样的两个数据集和同样的实验条件下进行实验。通过结果可以看出，MLCTAI 方法可以通过仿射变换减小预测误差。相比于其余五种方法，MLCTAI 方法在预测性能与预测稳定性方面展现出更加优越的能力。因而，MLCTAI 方法可以用于校正仪器间的偏差。

## 致 谢

首先我要感谢我的导师赵煜辉老师，在毕业设计过程中，赵老师十分耐心地对我的论文的方向进行指导，严格要求我，并给我的实验思路以及行文内容编排布局提出了大量的宝贵建议，给了我非常大的支持和帮助。

其次，我还要感谢所有支持和帮助我的老师和同学们。整个大学四年间，是你们陪伴着我成长。在大学，我收获的不仅仅是知识，还有深深的友谊。转眼间，我就要结束我的大学生涯了。但我的人生还很长很长，未来，我会走到不同的地方，会经历不一样的风景，但无论如何，大学四年的时光，和你们在一起的日子，会深深地刻在我的心中。

最后，再次感谢我身边的所有人，是你们激励我一起前行，大家一起加油！

## 参考文献

- [1] Russell C A , Angus J F , Batten G D , et al. The potential of NIR spectroscopy to predict nitrogen mineralization in rice soils[J]. *Plant and Soil*, 2002, 247(2):243-252.
- [2] Afara I , Singh S , Oloyede A . Application of near infrared (NIR) spectroscopy for determining the thickness of articular cartilage[J]. *Medical Engineering & Physics*, 2013, 35(1):88-95.
- [3] Abdo S , Freel K , Handley E , et al. Determination of Protein Content by NIR Spectroscopy in Protein Powder Mix Products[J]. *Journal of Aoac International*, 2016, 99(2):-.
- [4] Pan S J , Yang Q . A Survey on Transfer Learning[J]. *IEEE Transactions on Knowledge and Data Engineering*, 2010, 22(10):1345-1359.
- [5] Du W , Chen Z P , Zhong L J , et al. Maintaining the predictive abilities of multivariate calibration models by spectral space transformation[J]. *Analytica Chimica Acta*, 2011, 690(1):64-70.
- [6] Cen H, He Y. Theory and application of near infrared reflectance spectroscopy in determination of food quality[J]. *Trends in Food Science & Technology*, 2007, 18(2):0-83.
- [7] Roggo Y, Chalus P, Maurer L, et al. A review of near infrared spectroscopy and chemometrics in pharmaceutical technologies[J]. *Journal of Pharmaceutical & Biomedical Analysis*, 2007, 44(3):683-700.
- [8] T. Næs, H. Martens. Principal component regression in NIR analysis: Viewpoints, background details and selection of components[J]. *Journal of Chemometrics*, 1988, 2(2):155-167.
- [9] Geladi P, Esbensen K. Regression on multivariate images: Principal component regression for modeling, prediction and visual diagnostic tools[J]. *Journal of Chemometrics*, 1991, 5(2):97-111.
- [10] Jolliffe I T. A Note on the Use of Principal Components in Regression[J]. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 1982, 31(3):300-303.
- [11] Wold S, Michael Sjöström, Eriksson L. PLS-regression: a basic tool of chemometrics[J]. *Chemometrics & Intelligent Laboratory Systems*, 2001, 58(2):109-130.
- [12] Rosipal R, Nicole Krämer. Overview and Recent Advances in Partial Least Squares[C]// *International Statistical and Optimization Perspectives Workshop "Subspace, Latent Structure and Feature Selection"*. Springer, Berlin, Heidelberg, 2005.
- [13] Geladi P, Kowalski B R. Partial Least-Squares Regression: A Tutorial[J]. *Analytica*



- Chimica Acta, 1986, 185(1):1-17.
- [14] Workman J J. A Review of Calibration Transfer Practices and Instrument Differences in Spectroscopy[J]. Applied Spectroscopy, 2018.
- [15] Brown S D. Transfer of Multivariate Calibration Models ☆ [J]. Comprehensive Chemometrics, 2009, 64(2):181-192.
- [16] Malli B , Birlutiu A , Natschl?Ger T . Standard-free calibration transfer - An evaluation of different techniques[J]. Chemometrics and Intelligent Laboratory Systems, 2017, 161:49-60.
- [17] Bouveresse E, Hartmann C, Massart D L. Standardization of Near-Infrared Spectrometric Instruments[J]. Analytical Chemistry, 1996, 68(6):982-990.
- [18] Cooper J B, Larkin C M, Abdelkader M F. Calibration transfer of near-IR partial least squares property models of fuels using virtual standards[J]. Journal of Chemometrics, 2011, 25(9):496-505.
- [19] Pereira C F, Pimentel M F, Galvão RK, et al. A comparative study of calibration transfer methods for determination of gasoline quality parameters in three different near infrared spectrometers[J]. Analytica Chimica Acta, 2008, 611(1):41-47.
- [20] Bouveresse E, Massart D L. Improvement of the piecewise direct standardization procedure for the transfer of NIR spectra for multivariate calibration[J]. Chemometrics & Intelligent Laboratory Systems, 1996, 32(2):201-213.
- [21] Kowalski B R, Wang Y. Calibration Transfer and Measurement Stability of Near-Infrared Spectrometers[J]. Applied Spectroscopy, 1992, 46(5):764-771.
- [22] Fan W, Liang Y, Yuan D, et al. Calibration model transfer for near-infrared spectra based on canonical correlation analysis[J]. Analytica Chimica Acta, 2008, 623(1):22-29.
- [23] Wang Y, Veltkamp D J, Kowalski B R. Multivariate instrument standardization[J]. Analytical Chemistry, 1991, 63(23):2750-2756.
- [24] Zheng K, Zhang X, Iqbal J, et al. Calibration transfer of near - infrared spectra for extraction of informative components from spectra with canonical correlation analysis[J]. Journal of Chemometrics, 2015, 28(10):773-784.
- [25] Thompson B. Canonical correlation analysis.[M]// Encyclopedia of Statistics in Behavioral Science. 2005.
- [26] Leng L, Zhang T, Kleinman L, et al. Ordinary least square regression, orthogonal regression, geometric mean regression and their applications in aerosol science[C]// 2007.
- [27] Cochrane D, Orcutt G H. Application of Least Squares Regression to Relationships

- Containing Auto- Correlated Error Terms[J]. Publications of the American Statistical Association, 1949, 44(245):32-61.
- [28] Geladi P D , Macdougall D B , Martens H . Linearization and Scatter-Correction for NIR Reflectance Spectra of Meat[J]. Applied Spectroscopy, 1985, 39(3):491-500.
- [29] Isaksson T , Tormod Næs. The Effect of Multiplicative Scatter Correction (MSC) and Linearity Improvement in NIR Spectroscopy[J]. Applied Spectroscopy, 1988, 42(7):1273-1284.
- [30] Helland I S , Næs T , Isaksson T . Related versions of the multiplicative scatter correction method for preprocessing spectroscopic data[J]. Chemometrics and Intelligent Laboratory Systems, 1995, 29(2):233-241.
- [31] Pan S J , Tsang I W , Kwok J T , et al. Domain Adaptation via Transfer Component Analysis[J]. IEEE Transactions on Neural Networks, 2011, 22(2):199-210.
- [32] Matasci G , Volpi M , Kanevski M , et al. Semisupervised Transfer Component Analysis for Domain Adaptation in Remote Sensing Image Classification[J]. IEEE Transactions on Geoscience and Remote Sensing, 2015, 53(7):3550-3564.
- [33] Pan S J , Kwok J T , Yang Q . Transfer Learning via Dimensionality Reduction[C]// Proceedings of the Twenty-Third AAAI Conference on Artificial Intelligence, AAAI 2008, Chicago, Illinois, USA, July 13-17, 2008. DBLP, 2008.
- [34] Fu Q , Wang J , Lin G , et al. Short-Wave Near-Infrared Spectrometer for Alcohol Determination and Temperature Correction[J]. Journal of Analytical Methods in Chemistry, 2012, 2012(1):728128.
- [35] Wold S , Antti H , Lindgren F , et al. Orthogonal Signal Correction of Near-Infrared Spectra[J]. Chemometrics and Intelligent Laboratory Systems, 1998, 44(1):175-185.
- [36] Jonas Sjöblom, Svensson O , Josefson M , et al. An evaluation of orthogonal signal correction applied to calibration transfer of near infrared spectra[J]. Chemometrics and Intelligent Laboratory Systems, 1998, 44(1):229-244.
- [37] Preys S , Roger J M , Boulet J C . Robust calibration using orthogonal projection and experimental design. Application to the correction of the light scattering effect on turbid NIR spectra[J]. Chemometrics and Intelligent Laboratory Systems, 2008, 91(1):28-33.
- [38] Jenkins W , Leon B . The use of residue number systems in the design of finite impulse response digital filters[M]// Residue number system arithmetic: modern applications in digital signal processing. IEEE Press, 1986.
- [39] Kramer K E , Morris R E , Rose-Pehrsson S L . Comparison of two multiplicative signal

- correction strategies for calibration transfer without standards[J]. Chemometrics and Intelligent Laboratory Systems, 2008, 92(1):33-43.
- [40] Scholkopf B . Kernel principal component analysis[C]// Proc. 7th International Conference on Artificial Neural Networks, Jan. 1997. Springer, Berlin, Heidelberg, 1997.
- [41] Fukumizu K , Bach F R , Jordan M I . Dimensionality Reduction for Supervised Learning with Reproducing Kernel Hilbert Spaces[J]. Journal of Machine Learning Research, 2004, 5:73-99.
- [42] Schölkopf B , Smola A , Müller, Klaus-Robert. Nonlinear Component Analysis as a Kernel Eigenvalue Problem[J]. Neural Computation, 1998, 10(5):1299-1319.
- [43] Kim M , Pavlovic V . Central Subspace Dimensionality Reduction Using Covariance Operators[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2011, 33(4):657-670.
- [44] Rosipal R , Nicole Krämer. Overview and Recent Advances in Partial Least Squares[C]// International Statistical and Optimization Perspectives Workshop "Subspace, Latent Structure and Feature Selection". Springer, Berlin, Heidelberg, 2005.
- [45] Feng Z W , Shan L S , Wuhan, et al. Translation, rotation and scaling changes in image registration based affine transformation model[J]. Infrared and Laser Engineering, 2001.
- [46] Wu W , Walczak B , Massart D L , et al. Artificial neural networks in classification of NIR spectral data: Design of the training set[J]. Chemometrics and Intelligent Laboratory Systems, 1996, 33(1):35-46.

## 附 录

### 附录 A

#### A Review on Affine Transformation

##### ABSTRACT

The growth of multimedia applications, Fractal image compression techniques is very important for the efficient transmission and storage of digital data. The purpose of this paper, we described fractal image compression using affine transformation correspond to translations, scaling, rotations and reflections in metric space.

##### 1. INTRODUCTION

In geometry, an affine transformation in latin “affinis” means “related” is a transformation introduced by Euler in 1748 in his book “Introduction in analysis in infinitorum then 1827 August Mobius wrote on affine geometry.

In geometry, an affine transformation is a function that maps an object from an affine space to another and which preserve structures. Indeed, an affine transformation preserves lines or distance ratios but change the orientation, size or position of the object. The set of affine transformation is composed of various operations. Translations which modify object position in the image. Homothetic transformations composed of the contraction and dilatation of an object, both scaling operations. The transvection (shear mapping) which shifts every point of an object. Rotation which to allows to rotate an object according to its axis [3]. Affine transformation improved of compression ratio of fractal image compression. The improvement of compression rates is done by applying the lossy compression techniques on the parameters of the affine transformation of the fractal compressed image.

##### 2. AFFINE TRANSFORMATION

The use of homogeneous coordinates is the central point of affine transformation which allow us to use the mathematical properties of matrices to perform transformations. So to transform an image, we use a matrix  $\mathbf{T} \in M_3(R)$  providing the changes to apply

$$\mathbf{T} = \begin{bmatrix} a_{11} & a_{12} & t_x \\ a_{21} & a_{22} & t_y \\ p_x & p_y & 1 \end{bmatrix}$$

The vector  $[t_x, t_y]$  represent the translation vector according the canonical vectors. the vector  $[p_x, p_y]$  represents the projection vector on the basis. The square matrix composed by the  $a_{ij}$  elements is the affine transformation matrix [3] [5]. An affine transformation  $T: R^2 \rightarrow R^2$  is a transformation of the form  $T: Ax + B$  defined by

$$T \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} + \begin{pmatrix} t_x \\ t_y \\ 1 \end{pmatrix}$$

Where the parameter  $(a_{11}, a_{12}, a_{21}, a_{22})$  form the linear part which determines the rotation, skew and scaling and the parameters  $(t_x, t_y)$  are the translation distances in x and y directions, respectively [1] [2]

$$\begin{cases} \begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = T \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} + \begin{pmatrix} t_x \\ t_y \\ 1 \end{pmatrix} \\ x' = xa_{11} + ya_{12} + t_x \\ y' = xa_{21} + ya_{22} + t_y \end{cases}$$

The general affine transformation can be defined with only six parameters:

$\theta$  : the rotation angle.

$t_x$  : the x component of the translation vector.

$t_y$  : the y component of the translation vector.

$S_x$  : the x component of the scaling vector.

$S_y$  : the y component of the scaling vector.

$Sh_x$  : the x component of the shearing vector.

$Sh_y$  : the y component of the shearing vector [3]

In the other words, The Fractal is made up of the union of several copies of itself, where each copy is transformed by a function  $T_i$ , such a function is 2D affine transformation, so the IFS is defined by a finite number of affine transformation which characterized by Translation,

scaling, shearing and rotation.

### 2.1 Translation

A translation is a function that moves every point with constant distance in a specified direction by the vector  $\mathbf{T} = [\mathbf{T}_x \quad \mathbf{T}_y]$  which provide the orientation and the distance according to axis x and y respectively [3]. The mathematical transformation is the following [4] [6]:

$$\left\{ \begin{array}{l} T = \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \\ x' = x + t_x \\ y' = y + t_y \end{array} \right.$$

### 2.2 Transvection

A transvection is a function that shifts every point with constant distance in a basis direction x or y [3]. The mathematical vertical and horizontal transvection respectively is the following [4] [6]:

$$\left\{ \begin{array}{l} T = \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & s_v & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \\ x' = x + ys_v \\ y' = y \end{array} \right.$$

$$\left\{ \begin{array}{l} T = \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ s_h & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \\ x' = x \\ y' = y + xs_h \end{array} \right.$$

### 2.3 Rotation

Rotation a circular transformation around a point or an axis. A rotation around an axis defined by the normal vector to the image plane located at the center of the image [3]. The mathematical rotation about clockwise and counter clockwise is the following [4] [6]:

$$\left\{ \begin{array}{l} T = \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} \cos \theta & \sin \theta & 0 \\ -\sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \\ x' = x \cos \theta + y \sin \theta \\ y' = -x \sin \theta + y \cos \theta \end{array} \right.$$

$$\left\{ \begin{array}{l} T = \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \\ x' = x \cos \theta - y \sin \theta \\ y' = x \sin \theta + y \cos \theta \end{array} \right.$$

## 2.4 Scaling

Scaling is a linear transformation that enlarge or shrinks objects by a scale factor that is the same in the direction [3]. The mathematical transformation is the following [4] [6]:

$$\left\{ \begin{array}{l} T = \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} s_x & 0 & 0 \\ 0 & s_y & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \\ x' = xs_x \\ y' = ys_y \end{array} \right.$$

## REFERENCES

- [1] Barnsely M. F., “Fractal Everywhere”, Acadmic Press, San Diego, 1993.
- [2] Barnsely M. F. and Hurd L.P., “Fractal Image Compression”, AK Peters Ltd, Wellesely, Massachurets, 1992.
- [3] Coste, “Affine Transformation, Landmarks registration”, Nonlinear warping, October, 2012.
- [4] David F. Rogers and J. Alan Admas, “Mathematical elements for computer graphics”, 2nd Ed, NewYork, 1990.
- [5] Yan Bin Jia, “Transformaions in Homogeneous coordinates”, Notes 2014.
- [6] University of Texas at Austin, coputer graphics, 2010.

## 中文译文 A

### 仿射变换综述

#### 摘要

多媒体应用的发展，分形图像压缩技术对于数字数据的高效传输和存储非常重要。

本文的目的，我们描述了使用仿射变换的分形图像压缩对应于度量空间中的平移，缩放，旋转和反射。

## 1. 介绍

在几何学中，拉丁语“affinis”中的仿射变换意味着“相关”，欧拉于 1748 年在他的著作“Introduction in analysis in infinitorum”中对其进行介绍。然后 1827 年 8 月莫比乌斯在仿射几何写下了仿射变换。

在几何中，仿射变换是将对象从仿射空间映射到另一个并保留结构的函数。实际上，仿射变换保留了线或距离比，但改变了对象的方向，大小或位置。仿射变换集由各种变换组成。改变图像中对象位置的平移。由物体的收缩和膨胀组成的同构变换，两者都是缩放操作。横向（剪切映射）移动物体的每个点。旋转允许根据轴旋转物体。仿射变换提高了分形图像压缩的压缩比。通过将有损压缩技术应用于分形压缩图像的仿射变换的参数来完成压缩率的改进。

## 2. 仿射变换

齐次坐标的使用是仿射变换的中心点，它允许我们使用矩阵的数学属性来执行变换。因此，为了变换图像，我们使用可提供更改的矩阵  $\mathbf{T} \in M_3(R)$  来实现：

$$\mathbf{T} = \begin{bmatrix} a_{11} & a_{12} & t_x \\ a_{21} & a_{22} & t_y \\ p_x & p_y & 1 \end{bmatrix}$$

根据典型向量， $[t_x, t_y]$  表示平移向量， $[p_x, p_y]$  代表了基础上的投影向量。由  $a_{ij}$  构成的平方矩阵是仿射变换矩阵。仿射变换  $T: R^2 \rightarrow R^2$  是一种这样的变换形式： $T: Ax + B$ ，其通过下面的公式实现：

$$T \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} + \begin{pmatrix} t_x \\ t_y \\ 1 \end{pmatrix}$$

其中参数  $(a_{11}, a_{12}, a_{21}, a_{22})$  构成了决定旋转，倾斜和缩放的线性部分。参数  $(t_x, t_y)$  分别是 x 方向和 y 方向的平移量。



$$\begin{cases} \begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = T \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} + \begin{pmatrix} t_x \\ t_y \\ 1 \end{pmatrix} \\ x' = xa_{11} + ya_{12} + t_x \\ y' = xa_{21} + ya_{22} + t_y \end{cases}$$

一般仿射变换只需六个参数即可定义：

$\theta$  : 旋转角

$t_x$ : 旋转向量的 x 成分

$t_y$  : 旋转向量的 y 成分

$S_x$  : 缩定向量的 x 成分

$S_y$  : 缩定向量的 y 成分

$Sh_x$  : 剪切向量的 x 成分

$Sh_y$  : 剪切向量的 y 成分

换句话说，分形是由它自身的几个复制品组成的，其中每个复制品都由一个函数  $T_i$  变换，这样一个函数就是 2D 仿射变换。因此，IFS 是由有限个仿射变换定义的，其特征是平移、缩放、剪切和旋转。

## 2.1 平移

平移是一个函数，它通过向量  $\mathbf{T} = [\mathbf{T}_x \quad \mathbf{T}_y]$  在一个指定的方向上移动每一个点，它分别根据 x 轴和 y 轴提供方向和距离。

$$\begin{cases} T = \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \\ x' = x + t_x \\ y' = y + t_y \end{cases}$$

## 2.2 缩并

缩并是在基础方向 x 或 y 上以恒定距离移动点的函数。数学上的垂直和水平缩并分别如下：

$$\left\{ \begin{array}{l} T = \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & s_v & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \\ x' = x + ys_v \\ y' = y \end{array} \right.$$

$$\left\{ \begin{array}{l} T = \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ s_h & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \\ x' = x \\ y' = y + xs_h \end{array} \right.$$

### 2.3 旋转

旋转是围绕一个点或一个轴形成的圆形结构。围绕由法线矢量定义的轴旋转到位于图像中心的图像平面。关于顺时针和逆时针方向的数学旋转如下：

$$\left\{ \begin{array}{l} T = \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} \cos \theta & \sin \theta & 0 \\ -\sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \\ x' = x \cos \theta + y \sin \theta \\ y' = -x \sin \theta + y \cos \theta \end{array} \right.$$

$$\left\{ \begin{array}{l} T = \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \\ x' = x \cos \theta - y \sin \theta \\ y' = x \sin \theta + y \cos \theta \end{array} \right.$$

### 2.4 缩放

缩放是一种线性变换，它用相同方向上的比例因子放大或缩小物体。数学变换如下：

$$\left\{ \begin{array}{l} T = \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} s_x & 0 & 0 \\ 0 & s_y & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \\ x' = xs_x \\ y' = ys_y \end{array} \right.$$

## 附录 B

### **A Review of Calibration Transfer Practices and Instrument Differences in Spectroscopy**

#### **ABSTRACT**

Calibration transfer for use with spectroscopic instruments, particularly for near-infrared, infrared, and Raman analysis, has been the subject of multiple articles, research papers, book chapters, and technical reviews. There has been a myriad of approaches published and claims made for resolving the problems associated with transferring calibrations; however, the capability of attaining identical results over time from two or more instruments using an identical calibration still eludes technologists. Calibration transfer, in a precise definition, refers to a series of analytical approaches or chemometric techniques used to attempt to apply a single spectral database, and the calibration model developed using that database, for two or more instruments, with statistically retained accuracy and precision. Ideally, one would develop a single calibration for any particular application, and move it indiscriminately across instruments and achieve identical analysis or prediction results. There are many technical aspects involved in such precision calibration transfer, related to the measuring instrument reproducibility and repeatability, the reference chemical values used for the calibration, the multivariate mathematics used for calibration, and sample presentation repeatability and reproducibility. Ideally, a multivariate model developed on a single instrument would provide a statistically identical analysis when used on other instruments following transfer. This paper reviews common calibration transfer techniques, mostly related to instrument differences, and the mathematics of the uncertainty between instruments when making spectroscopic measurements of identical samples. It does not specifically address calibration maintenance or reference laboratory differences.

#### **INTRODUCTION**

One may discuss two main scenarios regarding the subject of calibration transfer to include: (1) the concept of using the same calibration model over time on multiple instruments using similar samples and measurement conditions; or (2) for using the same calibration model to adapt to new measurement and sample conditions, such as changes in natural product phenotype, chemical composition, temperature, sample presentation geometry, variation in instrumentation methods, and the like. The first scenario may be referred to simply as “calibration transfer between instruments” and the second, “calibration maintenance.” This review specifically addresses calibration transfer between instruments, and within that scenario more specifically the subject of calibration transfer involving instrument differences and methods to identify and

compensate for those differences. The topic of calibration maintenance is a significant topic of its own and is not addressed here in the allotted space for this article. In addition, in the scope of this review, limited equations are given for the most well-used transfer algorithms, with more extensive equations shown for instrument evaluation methods.

It is noted that multiple papers addressing calibration maintenance and the issues involved have been published. One paper discusses the concepts of calibration space and its relationship to a specific calibration model. The convolution of the initial sample matrix variables and instrument environment variables determine the conditions for the primary model, and as these conditions change a dynamic modeling maintenance approach is required [1].

Calibration transfer in its purest form allows a calibration to be moved from one instrument to another with statistically retained accuracy and precision [2,3]. Ideally, a model developed on one instrument would provide the statistically identical analysis on the same sample set when used on other instruments, without bias or slope adjustments, without additional product samples for standardization, and without product recalibration [4].

A variety of textbooks and papers exist describing the general methods and history of calibration transfer. Among these are references describing the details of the background chemometrics, as well as the instrumentation and data types involved [5,6]. One comprehensive calibration transfer review paper presents an overview of the different mathematical methods used for calibration transfer and a critical assessment of their validity and applicability [7]. The focus of that review paper is on a description of methods used for transfer of calibrations for near-infrared (NIR) spectra. A second early, yet prominent, review of calibration transfer thoroughly describes chemometric methods used for most aspects of calibration transfer [8]. That review classifies transfer methods as being represented by three main approaches, to include: (1) developing more robust calibrations, derived by combining data from multiple instruments and by using spectral pretreatments of that data; (2) by modifying calibrations using bias and slope adjustments (the traditional method); and (3) by adjusting spectra, using such methods as piecewise direct standardization (PDS) combined with a patented method that requires spectra from transfer sample sets measured on each instrument involved in the transfer experiment [9].

A comprehensive book chapter presents a review of calibration transfer methods and a critical assessment of the more common methods in terms of applicability and capabilities [10]. This previously published review presents an overview of calibration transfer methods and a critical assessment of the more common methods in terms of applicability and capabilities.

There is a focus on specific methods for calibration transfer using NIR spectra, and it is noted that the methodology described is widely applicable to other analytical measurement techniques. Comprehensive reviews of chemometric methods, including calibration transfer, have been published since 1980 as part of the Fundamental Reviews in Analytical Chemistry [11]. One of the most cited of such fundamental chemometrics reviews, which delves into multiple topics related to calibration and calibration transfer methods, is found in Lavine et al [12].

#### REFERENCE

1. P. Shahbazikhah, J.H. Kalivas. "A Consensus Modeling Approach to Update a Spectroscopic Calibration". *Chemom. Intell. Lab. Syst.* 2013. 120: 142–153.
2. J.J. Workman Jr. "The Essential Aspects of Multivariate Calibration Transfer". *40 Years of Chemometrics—From Bruce Kowalski to the Future*. Washington D.C.: American Chemical Society, 2015, pp.257–260.
3. H. Mark, J. Workman. "Calibration Transfer, Part I". *Spectroscopy*. 2013. 28(2): 24–37.
4. J.J. Workman Jr. "First Principles of Instrument Calibration". *NIR news*. 2016. 27(3): 12–15.
5. H. Mark, J. Workman. *Chemometrics in Spectroscopy*, 1st ed. Boston, MA: Elsevier/Academic Press, 2010.
6. J.J. Workman. "Calibration Transfer". In: J. Workman, editor. *The Concise Handbook of Analytical Spectroscopy: Physical Foundations, Techniques, Instrumentation, and Data Analysis*. New Jersey; Singapore: World Scientific Publishing–Imperial College Press, 2016, Volume 3: Pp. 185–188. Volume 4: Pp. 159–162.
7. R.N. Feudale, N.A. Woody, H. Tan, A.J. Myles, et al. "Transfer of Multivariate Calibration Models: A Review". *Chemom. Intell. Lab. Syst.* 2002. 64(2): 181–192.
8. T. Fearn. "Standardization and Calibration Transfer for Near Infrared Instruments: A Review". *J. Near Infrared Spectrosc.* 2001. 9(4): 229–244.
9. J. Shenk. M. Westerhaus. *Optical Instrument Calibration System*. US Patent 4866644. Filed 1986. Issued 1989.
10. S.D. Brown. "Transfer of Multivariate Calibration Models: A Review". *Chemom. Intell. Lab. Syst.* 2002. 64(2): 181–192.
11. B.R. Kowalski. "Chemometrics". *Anal. Chem.* 1980. 52(5): 112R–122R.
12. B. Lavine, J. Workman. "Chemometrics". *Anal. Chem.* 2008. 80(12): 4519–4531.

## 中文译文 B

### 光谱学标定迁移实践与仪器差异研究综述

#### 摘要

可以用于光谱仪器，特别是用于近红外，红外和拉曼分析的标定迁移，已成为多篇文章，研究论文，书籍章节和技术评论的主题。为解决与标定迁移相关的问题，已经发布了无数的方法和声明；然而，使用相同标定模型的两个或多个仪器在一段时间内获得相同结果的能力仍未达到技术人员的要求。在精确定义中，标定迁移是指用于尝试应用单个光谱数据库的一系列分析方法或化学计量学技术，以及使用该数据库为两个或更多仪器开发的标定模型，具有统计学上保留的准确度和精确度。理想情况下，一个人会为任何特定的应用开发一个单一的标定模型，并且不加区分地在仪器之间迁移，并获得相同的分析或预测结果。这种精密校准传递涉及许多技术方面，涉及测量仪器的重现性和重复性，以及用于标定模型的参考化学值，一种用于校准的多元数学，以及样品表示的可重复性和重现性。理想情况下，在单一仪器上开发的多变量模型在迁移后用于其他仪器时将提供统计上相同的分析。本文回顾了常见的标定迁移技术，主要涉及仪器差异，以及仪器在进行相同样品的光谱测量时的不确定的数学计算。它没有专门针对迁移维护或参考实验室差异。

#### 介绍

可以讨论有关标定迁移主题的两种主要场景，包括：(1) 在使用相似的样品和测量条件的多个仪器上使用相同的标定模型的概念；或(2)使用相同的校准模型以适应新的测量和样品条件，如天然产物表型、化学成分、温度、样品表示几何学等的变化，仪器方法的变化，等等。第一种情况可以简单地称为“仪器之间的标定迁移”，第二种是“迁移维护”。本综述专门讨论了仪器之间的标定迁移，在这种情况下，更具体地涉及仪器差异和方法的校准转移主题包括仪器差异以及如何识别并校正这些差异。迁移维护这一主题本身就是一个重要的主题，本文在分配的空间中没有讨论这个主题。此外，在本综述的范围内，给出了最常用的迁移算法的有限方程，并为仪器评价方法提供了更广泛的方程。

值得注意的是，已经发表了多篇关于标定迁移和所涉问题的论文。其中一篇讨论了标定空间的概念及其与特定标定模型的关系。初始样本矩阵变量和仪器环境变量的卷积决定了初始模型的条件，并且随着这些条件的变化，动态建模维护也随之发生变化。随

着这些条件的改变，需要使用动态建模维护方法。

标定迁移以其最纯粹的形式允许标定模型从一种仪器迁移到另一种仪器，并在统计上保持精度和精度。理想情况下，在一种仪器上开发的模型将在其他仪器上使用时，在同一样本集上提供统计上相同的分析，没有偏差或斜率调整，没有额外的标准化产品样本，也没有产品重新校准。

各种教科书和论文描述了标定迁移的一般方法和历史。其中包括描述背景化学计量学细节的参考文献，以及所涉及的仪器和数据类型。一篇全面的标定迁移综述论文概述了用于标定迁移的各种数学方法，并对其有效性和适用性进行了批判性评估。该综述的重点是描述近红外光谱(NIR)光谱的标定迁移方法。第二个早期的，但却是突出的，校准转移的综述，彻底描述了化学计量学方法用于校准转移的大部分方面。该综述将转移方法分为三种主要方法，包括：(1)通过组合来自多个仪器的数据并利用该数据的光谱预处理，发展更稳健的标定模型；(2)利用偏差和斜率调整来修正校准(传统方法)；(3)通过调整光谱，采用分段直接标准化(PDS)等方法，并结合专利方法，要求从包含在迁移实验中的每个仪器上测量的转移样本集中提取光谱。

一个全面的书籍章节介绍了标定迁移方法的综述，并从适用性和能力的角度对比较常用的方法进行了批判性评估。这篇先前发表的评论概述了标定迁移方法，并从适用性和能力的角度对比较常见的方法进行了批判性评估。重点介绍了利用近红外光谱进行标定迁移的具体方法，并指出所述方法广泛适用于其他分析测量技术。自 1980 年以来，包括标定迁移在内的化学计量学方法的全面评论已作为分析化学基础评论的一部分出版。这类基本化学计量学综述中引用最多的一个是 Lavine 等人提出的一篇综述，它深入探讨了与校准和校准转移方法有关的多个主题。