

DOI:10.3724/SP.J.1008.2008.00706

• 技术方法 •

## 一种无校正集的近红外光谱定量分析的新方法:相似区间最小值法

朱学义<sup>1,2</sup>, 亓云鹏<sup>1</sup>, 隋利娅<sup>3</sup>, 吴玉田<sup>1</sup>, 陆峰<sup>1\*</sup>

1. 第二军医大学药学院药物分析学教研室, 上海 200433

2. 解放军第12医院药械科, 疏勒 844200

3. 山东省药检所质保办, 济南 250012

**[摘要]** 目的:建立一种无需“校正集”的近红外光谱定量方法。方法:从待测化合物与混合物各一张光谱入手,提出一种相似区间最小值法无校正集分析法,即通过求导、差谱、相似性计算等手段筛选出混合物与待测化合物吸收最相似的区间,然后进行统计意义上的定量分析。结果:通过测定3个实际样品中的3种组分,得到平均回收率为100.9%,RSD=6.1%(n=9)。结论:相似区间最小值法无校正集分析法不需“校正集”,可得到较满意的定量分析结果,为近红外定量分析技术提供了一种新的可能,具有良好的应用前景。

**[关键词]** 近红外光谱法;定量分析;无校正集分析法

**[中图分类号]** R 917 **[文献标志码]** A **[文章编号]** 0258-879X(2008)06-0706-03

### A new calibration-free quantitative method for near-infrared spectroscopy

ZHU Xue-yi<sup>1,2</sup>, QI Yun-peng<sup>1</sup>, SUI Li-ya<sup>3</sup>, WU Yu-tian<sup>1</sup>, LU Feng<sup>1\*</sup>

1. Department of Drug Analysis, School of Pharmacy, Second Military Medical University, Shanghai 200433, China

2. Department of Medicine and Apparatus, No. 12 Hospital of PLA, Shule 844200

3. Department of Quality Control, Shandong Institute for Drug Control, Jinan 250012

**[ABSTRACT]** **Objective:** To develop a new quantitative method without calibration set for near-infrared spectroscopy (NIR). **Methods:** The new calibration-free NIR method was based on a spectrum of the sample and a spectrum of the mixture. The similar intervals between the spectra of the sample and the mixture were obtained through second derivation, differential spectra and calculation of similarity. Then statistical quantitative analysis was performed. **Results:** The contents of 3 components in 3 samples were determined with the new method; the average recovery was 100.9% and the relative standard deviation was 6.1% (n=9). **Conclusion:** Satisfactory results can be obtained by this new method without traditional calibration set. The new method provides an alternative for the quantitative analysis in NIRs and has a promising future.

**[KEY WORDS]** near-infrared spectroscopy; quantitative analysis; calibration-free method

[Acad J Sec Mil Med Univ, 2008, 29(6):706-708]

近红外光谱定量分析技术是近年来发展非常迅速的分析技术之一,具有快速、无损、环保、在线、高通量等优点,在石油化工<sup>[1]</sup>、农业<sup>[2]</sup>、药物分析<sup>[3]</sup>等领域有着广泛的应用,并取得了良好的经济效益和社会效益。应用近红外光谱技术进行定量分析的经典方式是先“建模”,其过程包括:(1)选择足够多的且有代表性的样品组成校正集;(2)通过现行标准方法测定校正样品的浓度或性质,得到“化学值”;(3)测定样品的近红外光谱;(4)采用定量校正方法建立数学模型。模型建立之后还有模型验证、模型更新与传递等过程,才有可

能建立起一个合适的模型,然后应用于样品实测。

但是,这种先建模后定量的方式更适用于石油、农业、食品等大生产模式的需要,在药物分析领域尚存在一定的局限性。比如,当样品量很少(只有几十个甚至十几个)时,则不能组成有代表性的校正集,不能建立有代表性的模型,即使建立了模型,实际应用的意义也不大;另外,当不能通过现行标准方法测定校正集样品,从而得到建模所需的“化学值”时,也无法完成模型的建立。另外,定量模型的建立需要花费大量的时间、人力和物力,成本比较高。

**[收稿日期]** 2007-10-11 **[接受日期]** 2008-03-25

**[基金项目]** 上海市科学技术委员会技术标准专项基金(05dz05006). Supported by Science & Technology Committee of Shanghai Municipal government(05dz05006).

**[作者简介]** 朱学义,硕士生, E-mail:xyz8917@sina.com

\* 通讯作者(Corresponding author). Tel:021-25070389-83, E-mail:fenglufeng@hotmail.com

本文提出一种无需校正集的近红外定量分析的新方法——相似区间最小值法, 仅从一张混合物光谱和一张化合物光谱入手, 通过求导、差谱、相似等数学手段, 就可以达到初步定量的目的, 从而避开大样品量和“化学值”的要求, 更加体现了近红外光谱用于定量分析时的“快速、高效”的优势。

1 原理

图 1 为含有一定量待测化合物的混合物近红外光谱图, 图 2 为图 1 的二阶导数谱(纵坐标无单位, 下同)。图 3 为图 2 的局部放大图。

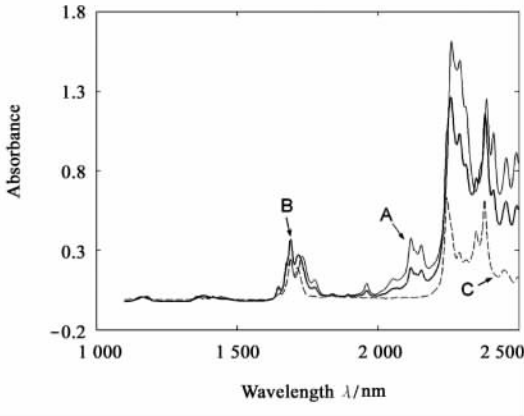


图 1 化合物 A、混合物 B 和干扰物质 C 的近红外光谱图

Fig 1 NIR spectra of compound A, mixture B and interference C

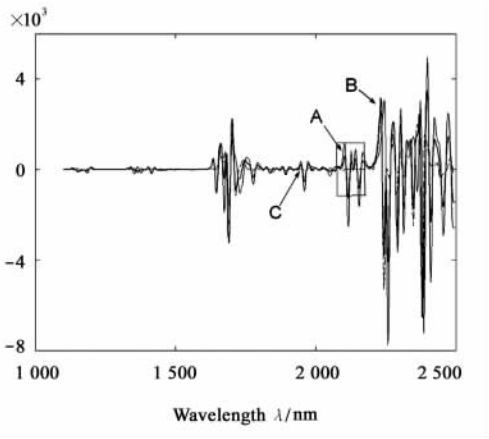


图 2 化合物 A、混合物 B 和干扰物质 C 的近红外二阶导数光谱图

Fig 2 Second derivative NIR spectra of compound A, mixture B and interference C

混合物的光谱吸收可以用下式表示:

$$B = A \times i\% + C \quad (1)$$

其中, A 为化合物二阶导数光谱, B 为混合物二阶导数光谱, C 为干扰物质二阶导数光谱, i 为化合物在混合物中的含量。由图 3 可见: 在区间 ab 中, C 接近于 0,  $A \times i\%$  接近于 B; 即在混合物中, 在区间 ab 中, 化合物所受其他物质的干扰最小。

在区间 ab 中, 只可能得到化合物的吸收光谱 A 和混合

物的吸收光谱 B。用  $S_{ab}(A, B)$  来表示在区间 ab 中, A 和 B 之间的相似程度, 此处  $S_{ab}(A, B)$  接近于 1。

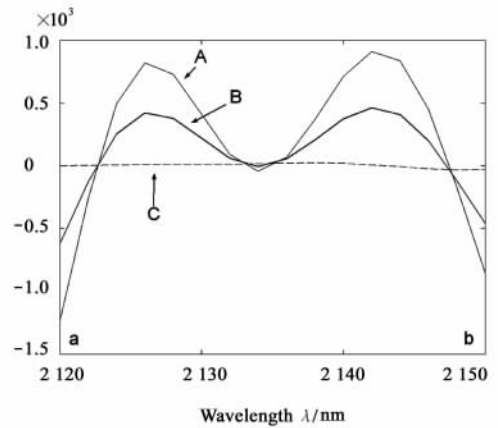


图 3 图 2 的局部放大图

Fig 3 Partial enlargement of Fig 2

用混合物与一定含量的化合物做差谱, 含量范围为 1%~1 000%, 其中必然有一条差谱是由混合物与所求含量的化合物得到的。差谱由下式表示:

$$D_i = B - A \times I\%, \quad (I = 1, 2, 3 \dots 1\ 000) \quad (2)$$

其中,  $D_i$  为混合物与一定含量化合物的差谱; B 为混合物光谱; A 为化合物光谱, I 为含量因子。图 4 为化合物、混合物和一条差谱 ( $I = 1\ 000$ ) 的二阶导数光谱图。

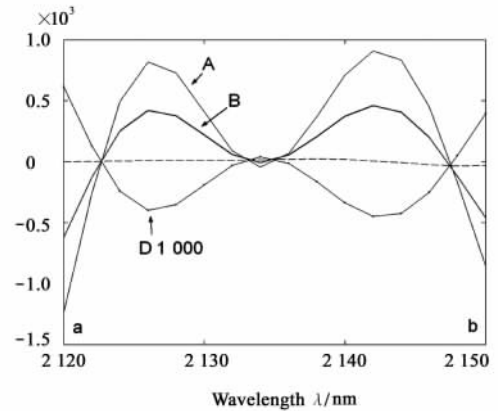


图 4 化合物 A、混合物 B 和第 1 000 条差谱  $D_{1000}$  ( $I = 1\ 000$ ) 的二阶导数光谱图

Fig 4 Second derivative spectra of compound A, mixture B and the 1 000<sup>th</sup> differential spectrum,  $D_{1000}$

用  $S_{ab}(A, D_i)$  来表示每一个差谱与化合物的相似程度。我们可以发现, 在区间 ab 中, 随着 I 的增大,  $S_{ab}(A, D_i)$  逐渐减小, 当 I 接近于 i 时,  $S_{ab}(A, D_i)$  接近于 0, 绝对值最小; 随后当  $I > i$  时,  $S_{ab}(A, D_i)$  变为负值, 绝对值逐渐加大。

如果我们在混合物中能够找到类似区间 ab 的特征区间并利用上述特征求得 i, 那么就实现了对含有一定量化合物的混合物进行定量分析的目的。

采用窗口大小为 m, 沿波长轴连续移动, 分别求待测化

合物与混合物的相似系数  $S_{mj}(A, B)$ 、待测物质与各个差谱的相似系数  $S_{mj}(A, D_I)$ , ( $I=1, 2, \dots, 1000$ ),  $J$  为窗口  $m$  中点的波长。同时满足下列两个条件的区间即为计算含量值需要选择的特征区间 ( $Z_j$ ):

$$S_{mj}(A, B) > n_1, (n_1 = 0 \sim 1); S_{mj}(A, D_{1000}) < n_2, (n_2 = -1 \sim 0) \quad (3)$$

其中  $j$  为特征区间中点的波长,  $n_1$ 、 $n_2$  分别为相似系数的选择界限。实际上, 每一个特征区间都可以求得一个预测值  $i_j$ , 它们的均值即为混合物中化合物的含量预测值。

## 2 方法和结果

近红外光谱来自于 Windig 课题组 (Analytical Technology Division, Eastman Kodak Company, Rochester, New York) [4], 选取其中的 3 个组分: 二氯甲烷、二氯丙烷、丙酮。用 3 个组分的原始光谱模拟 3 个混合物作为模拟样品; 选取 3 个实际样品作为待测样品, 测定 3 个所选组分在混合物中的含量。

方法程序由 Matlab6.5 编程实现。

图 5 列出了 3 种组分和真实样品 1 的近红外光谱图。

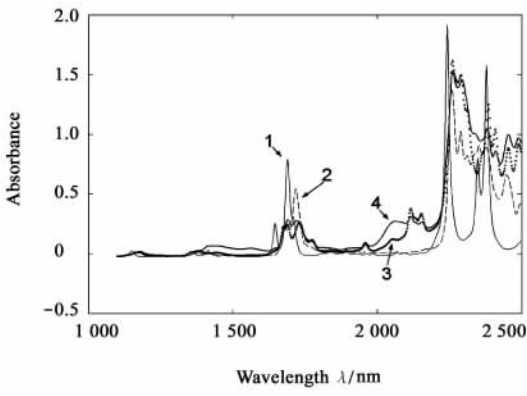


图 5 三种组分和真实样品 1 的近红外光谱

Fig 5 NIR spectra of 3 components and real sample 1

1: Methylene chloride; 2: Dichloropropane; 3: Acetone; 4: Real sample 1

表 1 为模拟样品的测定结果, 表 2 为实际样品的测定结果。由表可见, 总体回收率均在 90%~110%, 在模拟样品 3 中, 即使二氯甲烷的含量低至 7%, 回收率仍然较好。

表 1 回收率测定结果

Tab 1 Recoveries of 3 components in simulated samples

Simulated samples	Component	Amount $\varphi$ /(%)	Found $\varphi$ /(%)	Recovery (%)
1	Methylene chloride	20	18.7	93.5
	Dichloropropane	60	55.9	93.17
	Acetone	20	20.3	101.5
2	Methylene chloride	15	14.7	97.67
	Dichloropropane	70	70.5	100.71
	Acetone	15	15.5	103.33
3	Methylene chloride	7	7.3	109.29
	Dichloropropane	85	83.5	98.24
	Acetone	8	8.6	107.5

表 2 样品测定结果

Tab 2 Recoveries of 3 components in real samples

Real samples	Component	Amount $\varphi$ /(%)	Found $\varphi$ /(%)	Recovery (%)
1	Methylene chloride	10.21	11.2	109.70
	Dichloropropane	22.38	21.5	96.07
	Acetone	34.84	32.3	92.71
2	Methylene chloride	10.12	9.7	95.85
	Dichloropropane	22.43	23.9	106.56
	Acetone	22.51	23.8	105.73
3	Methylene chloride	11.56	11.5	99.45
	Dichloropropane	22.08	21.3	96.47
	Acetone	34.25	36.1	105.40

## 3 讨论

近红外光谱原谱比较平缓, 重叠很严重, 特征显现不突出。二阶导数光谱既能反映原谱的信息, 特征性又比较强, 所以在实际应用过程中采用近红外光谱的二阶导数谱进行算法运算, 求导的方法采用移动窗口多项式[5]法。

目前衡量光谱的相似程度主要有 4 种方法: 欧氏距离法、相关系数法、夹角余弦法和光谱信息散度法。根据文献[6], 夹角余弦法准确反映光谱相似情况, 所以本实验选择夹角余弦法作为光谱相似程度衡量的方法。

当在选择区间内待测化合物与其他干扰物质的吸收很相似, 即存在光谱共线性时, 方法的准确性会受到影响, 这也是采用一维方法解析灰色体系时都要面临的问题。

本方法的结果受参数, 特别是公式(3)中  $n_1$  和  $n_2$  的影响比较大, 如对实际样品 1 中二氯甲烷组分的分析中, 当  $n_1 = 0.85, n_2 = -0.90$  时, 求得含量为 11.2%; 当  $n_1 = 0.80, n_2 = -0.90$  时, 求得含量为 11.8%; 当  $n_1 = 0.90, n_2 = -0.90$  时, 求得含量为 12.0%。因此如何提高方法的稳定性还值得进一步深入研究。

本研究仅从一张混合物近红外光谱和一张化合物红外光谱入手, 运用求导、差谱与计算相似性等数学手段对混合物中所含化合物进行了定量分析的尝试, 并取得了理想的结果, 从而避开了“建模”周期长、成本大的缺点, 特别是当建模条件不能满足时, 为近红外光谱定量分析技术的实际应用开辟了一条新的道路。

## [参考文献]

[1] 顾洁, 王艳斌, 袁洪福, 陆婉珍. 甲基叔丁基醚含量对近红外光谱法测定汽油族组成的影响[J]. 石油化工, 2005, 34: 1208-1211.

[2] 张晓芳, 俞信, 阎吉祥, 倪勇. 近红外反射技术开放式检测棉籽中水分和油含量的研究[J]. 光谱学与光谱分析, 2007, 27: 473-476.

[3] 范积平, 张贞良, 张柳瑛, 吴玉田. 近红外光谱法测定药用大黄中 4 种蒽醌类成分[J]. 第二军医大学学报, 2005, 26: 1194-1195.

[4] Windig W, Stephenson D A. Self-modeling mixture analysis of second-derivative near-infrared spectra data using the SIMPLISMA approach [J]. Anal Chem, 1992, 64: 2735-2742.

[5] 梁逸曾, 杜一平. 分析化学计量学[M]. 重庆: 重庆大学出版社, 2004: 91.

[6] 杜培军, 陈云浩, 方涛, 唐宏. 基于光谱特征的高光谱遥感影像检索[J]. 光谱学与光谱分析, 2005, 25: 1171-1175.

[本文编辑] 尹茶