

近红外光谱测定溶液中甲醇的含量

吴卫红^{a,b} 王海水^{a*}

(^a 中国科学院长春应用化学研究所, 稀土化学与物理重点实验室 长春 130022; ^b 中国科学院研究生院 北京)

摘 要 测量了含微量甲醇(体积分数为0.04%~0.24%)的系列乙醇水溶液的近红外光谱,利用近红外光谱分析建立了预测甲醇含量的定量分析模型。比较了用外部检验法(Test Set-Validation)和交叉检验法(Cross-Validation)建立的数学模型,研究了使用外部检验法时,校正集和检验集样品数的改变对模型预测结果的影响。结果发现,当校正集样品数为15 检验集样品数为6(总样品数为21)时,使用外部检验法建立的数学模型预测结果较好,其校正集的均方根误差和检验集的预测均方根误差(分别为RMSEE和RMSEP)均较小(分别为0.011 5和0.010 5),而且很接近。结果表明,近红外光谱方法简单、准确而且实用。

关键词 近红外光谱,数学模型,定量分析,甲醇

中图分类号:O657.3

文献标识码:A

文章编号:1000-0518(2007)10-1101-04

现代近红外光谱分析技术的主要优点有:样品无须预处理、分析速度快、效率高、测量方便、可用于在线分析和无损分析。近红外光谱分析技术还适用于气、液、固体样品中的主要成分、微量甚至痕量成分的直接分析测定^[1-10]。近红外技术已广泛应用于食品、烟草、医药、化学、化工、石油等领域的定性和定量分析。近红外光谱主要研究的是样品分子基频振动的合频与倍频的信息特征,该光谱区信号的频率介于中红外区和可见光谱区之间。近红外光谱区主要反映含氢基团(C—H, O—H, S—H, N—H等)的特征振动信息,其信号强度比中红外区弱得多,因此样品不需要稀释,可以直接用来测量。由于吸收峰重叠,谱峰宽,精确进行近红外光谱带的归属是很困难的,所以需要借助于化学计量学来进行定性和定量分析^[11-14]。

目前,近红外光谱已经开始应用于白酒中乙醇含量的分析^[15,16]。甲醇是对人体有害的物质,据报道^[7],人服用甲醇中毒最低剂量为0.1 g/kg,经口摄入0.3~1.0 g/kg便可致死。在我国,白酒中甲醇含量测定的国标法(GB5009148-85)是品红亚硫酸法,该方法操作繁琐费时,不利于快速测定。本文尝试用近红外光谱技术对乙醇水溶液中的微量甲醇进行定量测试,希望找到一种快速方便的方法来测量白酒中甲醇的含量,为快速鉴别劣质白酒提供一种技术手段。

1 实验部分

1.1 试剂和仪器

采用乙醇、甲醇(均为分析纯试剂)和蒸馏水配制成实验样品21个。其中甲醇、乙醇、水的体积分数分别为0.04%~0.24%、11%~56%及43.44%~88.76%。

BRUKER VECTOR 22/N型傅立叶变换近红外光谱仪(德国),采用透射反射方式,扫描区间:4 000~12 000 cm^{-1} ,扫描次数:128次,分辨率:8 cm^{-1} 。定量建模算法为偏最小二乘法PLS。

1.2 数据处理

应用BRUKER OPUS/QUANT-2定量分析软件进行数据自动处理。分别用外部检验法和交叉检验法进行验证,确定最佳条件和主成分数,并且检验模型的预测能力。

2 结果与讨论

2.1 外部检验法中定标模型样品数对甲醇含量模型特征值的影响

实验发现,用外部检验法进行检验时,增加或是减少用于建立校正集(或检验集)的样品数时,对模型的预测能力有一定影响,即模型的特征值发生改变。表 1 为数学模型特征值随样品数的改变而发生的变化。

表 1 甲醇溶液数目对数学模型特征值的影响

Table 1 Effects of sample amounts on eigenvalue of mathematic models

Number of Sample		RMSEE	RMSEP	R^2	R_c^2	Rank
Calibration	Validation					
11	10	0.020 7	0.014 0	91.66	87.11	2
13	8	0.018 5	0.011 4	94.61	89.03	2
15	6	0.011 5	0.010 5	96.25	95.71	6
16	5	0.018 1	0.006 72	98.43	88.91	4
17	4	0.012 4	0.004 1	99.33	94.85	5

一个理想的数学模型需要具备如下特点:2 类 R^2 数值互相接近且接近 100, RMSEP 和 RMSEE 相互接近且越小越好, Rank 合适。本文 Rank 也即因子数的上限是 10, 因子数的大小对模型质量的影响比较大, 进而影响模型的预测能力。因子数太大会导致过度拟合, 增加噪音, 对定标模型的预测能力特别好而对检验模型的预测能力下降, 从而降低模型的预测能力。因子数太小则会出现拟合不足, 模型不能真实反映全部特性。从表 1 中可以看出, 综合 5 个特征值(RMSEE, RMSEP, R^2 , R_c^2 , Rank)考虑, 采用校正集样品数为 15, 检验集样品数为 6 比较合适。

2.2 外部检验法和交叉检验法建立的甲醇含量数学模型比较

一个用于定量分析的数学模型建立后需要评价这个模型的预测可靠性, 通常有 2 种检验方法可用: 交叉检验和外部检验。本文分别用这 2 种检验方法对建立的模型进行了检验。图 1 和图 2 为分别采用外部检验法和交叉检验法所建模型的真实值与预测值关系曲线, 横坐标是真实值, 纵坐标是预测值。

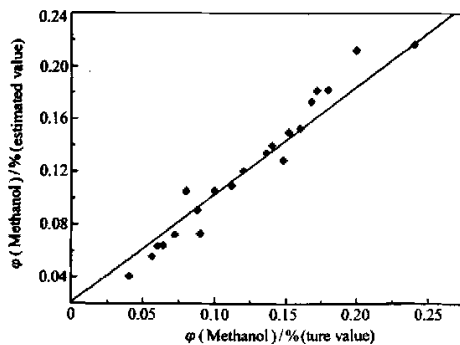


图 1 采用交叉检验法时检验集中甲醇含量真实值与预测值关系曲线

Fig. 1 Plot of methanol true value vs estimated value via Cross Validation

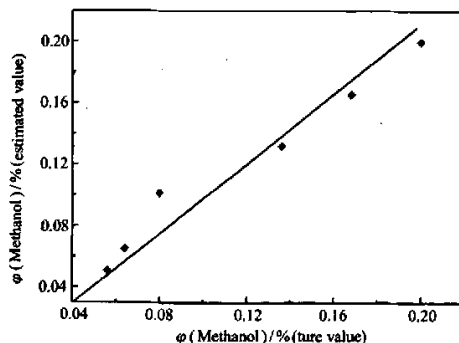


图 2 采用外部检验法时检验集中甲醇含量真实值与预测值关系曲线

Fig. 2 Plot of methanol true value vs estimated value via Test Set Validation

从图 1 和图 2 结果可知, 用交叉检验法和外部检验法进行验证的结果显示, 甲醇含量的预测值与真实值之间线性关系良好。

用外部检验法和交叉检验法进行验证所显示的模型对甲醇含量的预测值和其真实值结果分别列于表 2 和表 3。

表 2 外部检验法中甲醇含量的真实值与预测值

Table 2 Result of predicting content of methanol with Test Set Validation

Methanol	$\varphi(\text{methanol})/\%$		Value of error/%	Spec Res
	true value	estimated value		
3 [*]	0.080	0.101 5	26.8	0.001 13
7 [*]	0.064	0.065 22	1.91	0.001 5
12 [*]	0.200	0.199 7	0.141	0.002 16
14 [*]	0.056	0.051 03	8.88	0.001 67
18 [*]	0.168	0.165 6	1.44	0.001 49
21 [*]	0.136	0.132 1	2.88	0.000 925

表 3 交叉检验法中甲醇含量的真实值与预测值

Table 3 Result of predicting content of methanol with cross Validation

Methanol	$\varphi(\text{methanol})/\%$		Value of error/%	Spec Res
	true value	estimated value		
1 [*]	0.04	0.040 67	1.68	0.104
2 [*]	0.06	0.063 61	6.02	0.063 4
3 [*]	0.08	0.105 3	31.6	0.056 6
4 [*]	0.12	0.120 5	0.4	0.090 5
5 [*]	0.14	0.139 6	0.256	0.064 8
6 [*]	0.18	0.182 3	1.28	0.147
7 [*]	0.064	0.064 16	0.25	0.083 9
8 [*]	0.148	0.128 5	13.2	0.076 1
9 [*]	0.152	0.149 6	1.61	0.031 3
10 [*]	0.09	0.072 93	19.0	0.232
11 [*]	0.1	0.105 8	5.8	0.207
12 [*]	0.2	0.211 8	5.9	0.109
13 [*]	0.24	0.216 4	9.83	0.157
14 [*]	0.056	0.055 82	0.327	0.179
15 [*]	0.072	0.072 19	0.269	0.068
16 [*]	0.112	0.109 3	2.38	0.066 5
17 [*]	0.16	0.152 9	4.41	0.057 4
18 [*]	0.168	0.173 2	3.10	0.065 9
19 [*]	0.172	0.181 6	5.58	0.056 6
20 [*]	0.088	0.090 92	3.32	0.176
21 [*]	0.136	0.134 2	1.33	0.097 9

从图 1、图 2 看出,无论是进行外部验证还是交叉验证,甲醇含量的真实值与预测值之间相关性良好($R^2 > 92$),而从表 2 及表 3 中则得到了进一步的证实,除了个别结果误差较大之外,甲醇含量的预测值与其真实值是很接近的,而且从光谱残留值来看,这个值越小说明实验结果的重现性越好。因此,用近红外光谱分析法定量分析甲醇含量时,结果准确,也即用近红外光谱法建立的数学模型预测得到的乙醇水溶液中甲醇的含量与其真实值接近,这说明近红外光谱分析方法可用于溶液中甲醇含量的快速测定。

参 考 文 献

- 1 Sauvage L, Frank D, Stearne J, Millikan M B. *Anal Chim Acta* [J], 2002, 458: 223
- 2 GAO Jun(高俊), XU Yong-Ye(徐永业), YAO Cheng(姚成). *Chinese J Appl Chem* (应用化学) [J], 2005, 22(12): 1 390
- 3 SHI Yong-Gang(史永刚), FENG Xin-Lu(冯新泸), SUN Ping(孙萍). *Chinese J Spectrosc Lab* (光谱实验室) [J], 2005, 22(3): 575
- 4 TANG Fu-Wei(唐福伟), JIN Yi-Feng(金一丰), QIAN Jian-Fang(钱建方), WANG Xin-Rong(王新荣). *Chinese J Spectrosc Lab* (光谱实验室) [J], 2005, 22(4): 683
- 5 WANG Guo-Qing(王国庆), WANG Fang(王芳), CHEN Da(陈达), SU Qing-De(苏庆德), SHAO Xue-Guang(邵学

- 广). *Spectrosc Spect Anal*(光谱学与光谱分析)[J], 2004, 24(12):1 540
- 6 LI Dai-Xi(李代禧), WU Zhi-Yong(吴智勇), XU Duan-Jun(徐端钧), XU Yuan-Zhi(徐元植). *Chinese J Anal Chem*(分析化学)[J], 2004, 32(8):1 070
- 7 ZHAO Suo-Lao(赵锁劳), PENG Yu-Kui(彭玉魁). *Chinese J Anal Chem*(分析化学)[J], 2002, 30(8):978
- 8 LIU Li-Li(刘荔荔), XING Wang-Xing(邢旺兴), JIA Nuan(贾暖), ZHU Li(朱力), WU Yu-Tian(吴玉田). *Chinese J Mater Med*(中国中药杂志)[J], 2003, 28(2):178
- 9 QIN Fang-Li(覃方丽), MIN Shun-Geng(闵顺耕), LI Ning(李宁). *Spectrosc Spect Anal*(光谱学与光谱分析)[J], 2003, 23(6):1 090
- 10 LÜ Li-Na(吕丽娜), ZHANG Yue(张玥), ZHOU Ding-Wen(周定文). *J Tianjin Univ*(天津大学学报)[J], 2004, 37(12):1 093
- 11 XU Guang-Tong(徐广通), YUAN Hong-Fu(袁红福), LU Wan-Zhen(陆婉珍). *Spectrosc Spect Anal*(光谱学与光谱分析)[J], 2000, 20(2):135
- 12 ZHENG Yong-Mei(郑咏梅), ZHANG Jun(张军), CHEN Xing-Dan(陈星旦), SHEN Xuan-Guo(申铨国). *Spectrosc Spect Anal*(光谱学与光谱分析)[J], 2004, 24(9):1 047
- 13 ZHAO Li-Li(赵丽丽), ZHAO Long-Lian(赵龙莲), LI Jun-Hui(李军会), ZHANG Lu-Da(张录达), YAN Yan-Lu(严衍禄). *Spectrosc Spect Anal*(光谱学与光谱分析)[J], 2004, 24(1):41
- 14 ZHANG Zhong-Hu(张中湖), CHAI Yi-Feng(柴逸峰), ZHANG Jun-Ren(张君仁). *Chinese J Pharmaceuticals*(中医药工业杂志)[J], 2005, 36(5):296
- 15 ZHANG Ling(张玲), YU Jian(于健). *Food Ind*(食品工业)[J], 2006, (4):27
- 16 WANG Zhen-Zuo(王贞佐), LU Jia-Hui(逯家辉), GUO Wei-Liang(郭伟良), CHEN Shu-Yan(陈淑彦), WEI Wei(韦韦), TENG Li-Rong(滕利荣). *Sci Tech Food Ind*(食品工业科技)[J], 2005, 26(11):163
- 17 ZHANG Jiong(张炯), WAN Wei-Guo(万伟国), XU Mai-Ling(徐麦玲), ZOU He-Jian(邹和建). *J Lab Med*(环境与职业医学)[J], 2002, 19(3):180

Content of Methanol in Solution Measured via Near-infrared Spectra

WU Wei-Hong^{a, b}, WANG Hai-Shui^{a*}

(^aKey Laboratory of Rare Earth Chemistry and Physics, Changchun Institute of Applied Chemistry,
Chinese Academy of Sciences, Changchun, 130022;

^bGraduate School of the Chinese Academy of Sciences, Beijing)

Abstract Near-infrared spectra of ethanol solutions were obtained and the mathematic models for quantitative analysis of trace methanol in solution were set up. The mathematic models established by test set-validation and by cross-validation were compared, and the effect of the change of the numbers of samples used in calibration model and validation model of test set-validation on the mathematic models was evaluated. It was found that the predicted results of the mathematic model are good when the number of samples is 15 in calibration model and that in validation model is 6 (the total is 21). In that case, the value of the root mean square error of the estimation of calibration and the root mean square error of prediction of validation (RMSEE and RMSEP, respectively) are very small (0.011 5 and 0.010 5 respectively) and almost the same. Our results indicate that near infrared spectroscopy can be used to analyze the content of methanol in ethanol solution very well.

Keywords near-infrared spectroscopy, mathematic model, quantitative analysis, methanol