

## 拟合实验数据的新方程——非整数幂多项式方程

盛景云, 方维平, 王跃敏, 郑泉兴, 杨意泉

(厦门大学化学系, 福建, 厦门, 361005)

**摘要:** 将实验数据拟合成方程在科学研究和工程计算上具有十分重要的作用。目前, 拟合实验数据最常用的是整数幂多项式  $f(x) = c_0 + c_1x + c_2x^2 + \dots + c_nx^n$ , 但是用此多项式拟合各类数据时有时误差较大。本文提出一个新方程, 即双系列非整数幂多项式  $g(x) = c_0 + c_1x^a + \dots + c_kx^{ka} + c_{k+1}x^{(k+1)b} + \dots + c_nx^{nb}$ , 式中  $a, b$  为参数,  $c_i (i = 0, 1, 2, \dots, n)$  为待定系数。在拟合各类实验数据时, 新方程总是优于整数幂多项式。

**关键词:** 拟合; 实验曲线; 实验数据; 方程; 多项式

**中图分类号:** O 657.31

**文献标识码:** A

**文章编号:** 1001-4160(2004)05-725-728

### A modified equation for correlating experimental data ——nonintegral power polynomial equation

SHENG JingYun, FANG WeiPing, WANG YueMin, ZHENG QuanXing and YANG YiQuan

(Department of Chemistry, Xiamen University, Xiamen, 361005, Fujian, China)

**Abstract:** It is important to correlate experimental data into a continuous function in scientific research and technical development. The equation mostly used for this purpose is polynomial:  $f(x) = c_0 + c_1x + c_2x^2 + \dots + c_nx^n$ . But the correlating results of this polynomial are sometimes undesirable. In this paper, a modified equation for this usage is proposed:  $g(x) = c_0 + c_1x^a + \dots + c_kx^{ka} + c_{k+1}x^{(k+1)b} + \dots + c_nx^{nb}$ , where  $a$  and  $b$  are parameters and  $c_i (i = 0, 1, 2, \dots, n)$  are undetermined coefficients. When this new equation is used for correlating different kinds of experimental data (curve), smaller errors will always be obtained in comparison with the normal polynomial  $f(x)$ .

**Key words:** correlation, experimental curve, experimental data, equation, polynomial

Sheng JY, Fang WP, Wang YM, Zhen QX and Yang YQ. A modified equation for correlating experimental data——nonintegral power polynomial equation. Computers and Applied Chemistry, 2004, 21(5):725-728.

#### 1 前言

众所周知, 将实验曲线(实验数据点组成的曲线)拟合成方程具有重要的作用, 如利用所得方程进行内插和优化计算; 进行数学运算特别是微分和积分, 可推导出其他表现物理规律的方程。

至今为止, 拟合实验数据时最常使用的方程是整数幂多项式。但用此多项式拟合实验曲线时, 如果相应级数发散或收敛速度慢, 则拟合误差较大。此外, 大多数研究者针对某一具体的实验结果进行拟合, 或者通过拟合某一科学技术领域的实验数据提出专门方程,<sup>[1-8]</sup> 但这类方程只能应用于相关领域, 未能推广应用。

本文的目的是针对目前在实验曲线拟合方面所存

在的问题, 提出一个应用较为普遍, 形式简单, 容易进行各种数学运算特别是微积分计算, 且待定系数容易确定和拟合精度较高的方程。

#### 2 新方程的提出

我们发现, 在整数幂有限多项式

$$f(x) = c_0 + c_1x + c_2x^2 + \dots + c_nx^n \quad (1)$$

中, 对  $x$  进行简单的幂函数变换, 可以在多项式项数不变的情况下使数据拟合精度明显提高。对  $x$  进行幂函数变换, 实际上是将连续函数近似表达成  $x$  的非整数幂多项式

$$g(x) = c_0 + c_1x^a + c_2x^{2a} + \dots + c_nx^{na} \quad (2)$$

式中,  $a$  为参数, 当  $a = 1$  时, 式(2)还原成式(1)。

以下举例说明这种近似表达的合理性。当  $a >$

收稿日期: 2003-11-28; 修回日期: 2004-02-20

作者简介: 盛景云 (1979—), 女, 福建人, 硕士, 物理化学专业; 导师: 方维平。

联系人: 方维平。

0 时,式(2)可写成

$$g(x) = g(0) + c_1 x^a + c_2 x^{2a} + \dots + c_n x^{na} \quad (3)$$

分别求  $g(x)$  的一阶至  $n$  阶导数在  $x=1$  处的值,有

$$g^{(1)}(1) = a c_1 + 2 a c_2 + \dots + n a c_n$$

$$g^{(2)}(1) = a(a-1)c_1 + 2a(2a-1)$$

$$c_2 + \dots + na(na-1)c_n$$

⋮

$$g^{(n)}(1) = a(a-1)(a-2)\dots(a-n+1)c_1 +$$

$$2a(2a-1)(2a-2)\dots(2a-n+1)c_2$$

$$+ \dots + na(na-1)(na-2)$$

$$\dots(na-n+1)c_n \quad (4)$$

由此方程组可确定系数  $c_i (i=1, 2, \dots, n)$ 。为简便起见,设  $n=2$ ,可解得

$$c_1 = \frac{1}{a^2} [(2a-1)g^{(1)}(1) - g^{(2)}(1)]$$

$$c_2 = \frac{1}{2a^2} [(1-a)g^{(1)}(1) + g^{(2)}(1)] \quad (5)$$

在此条件下,式(3)可表为

$$g(x) = g(0) + \frac{1}{a^2} [(2a-1)g^{(1)}(1) - g^{(2)}(1)]$$

$$x^a + \frac{1}{2a^2} [(1-a)g^{(1)}(1) + g^{(2)}(1)]x^{2a} \quad (6)$$

相应地,式(1)变成

$$f(x) = f(0) + f^{(1)}(0)x + \frac{f^{(2)}(0)}{2!}x^2 \quad (7)$$

为了表明式(6)与式(7)逼近连续函数精度的区别,以函数  $y = (1+x)^{1.5}$  为例进行计算。此时,式(7)变为

$$y_1 \equiv f(x) = 1 + 1.5x + \frac{1.5 \times 0.5}{2}x^2 \quad (8)$$

而式(6)变为

$$y_2 \equiv g(x) = 1 + \frac{1}{a} [(2a-1) \times 1.5 \times 2^{0.5}$$

$$- 1.5 \times 0.5 \times 2^{-0.5}]x^a + \frac{1}{2a^2}$$

$$[(1-a) \times 1.5 \times 2^{0.5} + 1.5$$

$$\times 0.5 \times 2^{(-0.5)}]x^{2a}$$

(9)

计算结果列于表1。表1中ARD(平均相对偏差)由下式定义(设  $Y_i$  代表  $y_{1,i}$  或  $y_{2,i}$ )

$$ARD = \text{mean}(RD)_i = \text{mean} \left| \frac{Y_i - y_i}{y_i} \right| \quad (10)$$

$i=1, 2, \dots, 11$ 。以下各表的ARD定义与此类似。

表1 式(6)与式(7)近似表达  $y = (1+x)^{1.5}$  的效果对比

Table 1 Difference of Eq. (6) and Eq. (7) in expressing  $y = (1+x)^{1.5}$

$x$	0	0.2	0.4	0.6	0.8	1.0	1.2	1.4	1.6	1.8	2.0	ARD/%
$y$	1.000	1.314	1.656	2.024	2.415	2.828	3.263	3.718	4.192	4.685	5.196	/
$y_1$	1.000	1.315	1.660	2.035	2.440	2.875	3.340	3.835	4.360	4.915	5.500	2.157
$y_2(a=1)$	1.000	1.329	1.679	2.050	2.442	2.856	3.291	3.747	4.224	4.723	5.243	0.904
$y_2(a=1.03)$	1.000	1.317	1.664	2.034	2.426	2.840	3.274	3.730	4.207	4.705	5.224	0.358
$y_2(a=1.05)$	1.000	1.308	1.653	2.022	2.414	2.827	3.263	3.718	4.195	4.693	5.211	0.122
$y_2(a=1.07)$	1.000	1.300	1.642	2.010	2.402	2.816	3.250	3.706	4.183	4.680	5.197	0.425

此例表明,用非整数幂多项式逼近连续函数时,效果可能优于整数幂多项式。

进一步发现,用双系列的非整数幂多项式

$$g(x) = c_0 + c_1 x^a + \dots + c_k x^{ka} + c_{k+1} x^{(k+1)b} + c_{k+2} x^{(k+2)b} + \dots + c_n x^{nb} \quad (11)$$

拟合实验数据时,在项数相同或减少的情况下,其拟合误差可进一步降低。当  $a=b$  时,式(11)还原为式(2);当  $a=b=1$  时,式(11)还原为式(1)。在实际拟合中,给定一对  $(a, b)$  数,对应一个ARD,通过试算,总可找到使ARD最小的  $a$  和  $b$ 。

新方程的一个显著优点是,式(11)是一个显方程多项式,在应用此方程计算其他物理量时,容易进行各种符号数学运算特别是进行积分或微分,且可用最小二乘法方便地确定系数  $c_i$ 。

### 3 应用举例

#### 3.1 CO<sub>2</sub> 压缩因子实验数据的拟合

应用 MATLAB 6.1 软件进行拟合(以下两例同此),结果列于表2。表2中实验数据  $Z_{\text{exp}}$  来自文献<sup>[9]</sup>第172页,  $Z_{c_1}$  表示式(1)拟合值,  $Z_{c_2}$  表示式(2)最佳拟合值( $T=0^\circ\text{C}$ 时,  $a=-0.5$ ;  $T=25^\circ\text{C}$ 时,  $a=-0.5$ ),  $Z_{c_3}$  表示式(11)最佳拟合值( $T=0^\circ\text{C}$ 时,  $a=-0.8, b=0.2$ ;  $T=25^\circ\text{C}$ 时,  $a=-0.6, b=-0.2; k=7$ ),  $Z_{c_4}$  表示 Virial 方程  $Z = 1 + B_1 P + B_2 P^2 + \dots + B_n P^n$  的结果。

由表2可见,新方程式(11)及式(2)的拟合精度总是高于整数幂多项式和 Virial 方程。当曲线变化尖锐( $T=0^\circ\text{C}$ )时,新方程的拟合效果尤为显著。

3.2 乙缩醛蒸气压(kPa)的拟合

以乙缩醛 ( $T_c = 265.2^\circ\text{C}$ ,  $P_c = 2978\text{kPa}$ ) 为例,  $P_{\text{exp}}$  为实验值(来自文献[10]第625页),  $P_{c_1}$  为式(1)拟合结果,  $P_{c_2}$  为式(2)拟合值 ( $n = 3, a = 1.5; n$

$= 4, a = 2.2; n = 5, a = 1.7$ ),  $P_{c_3}$  为式(11)拟合值 ( $n = 3, k = 2, a = 2.1, b = 3; n = 4, k = 3, a = 2.2, b = 2.6; n = 5, k = 4, a = 1.5, b = 2.8$ )。

表2 式(1)式(2)式(11)和 Virial 方程的拟合结果对比( $n = 8$ )

Table 2 Correlated results of  $\text{CO}_2$  compressibility factor using Eq. (1), Eq. (2), Eq. (11) and Virial equation ( $n = 8$ )

$P(\text{MPa})$	$T = 0^\circ\text{C}$					$T = 25^\circ\text{C}$				
	$Z_{\text{exp}}$	$Z_{c_1}$	$Z_{c_2}$	$Z_{c_3}$	$Z_{c_4}$	$Z_{\text{exp}}$	$Z_{c_1}$	$Z_{c_2}$	$Z_{c_3}$	$Z_{c_4}$
0.1	0.9999	1.0997	0.9999	0.9999	0.9799	1.0933	1.0864	1.0933	1.0933	1.0043
0.5	0.9711	0.9805	0.9711	0.9711	0.9025	1.0701	1.0711	1.0701	1.0701	1.0145
1	0.9335	0.8562	0.9335	0.9335	0.8118	1.0401	1.0437	1.0401	1.0401	1.0133
2	0.8004	0.6416	0.8003	0.8004	0.6505	0.9412	0.9676	0.9412	0.9412	0.9736
5	0.0709	0.2663	0.0717	0.0709	0.3143	0.7259	0.6669	0.7258	0.7259	0.6998
10	0.2444	0.1538	0.2396	0.2445	0.1372	0.2290	0.2887	0.2296	0.2290	0.2774
20	0.3810	0.4013	0.3984	0.3817	0.3903	0.4492	0.3952	0.4458	0.4478	0.3876
30	0.5694	0.5721	0.5544	0.5655	0.5945	0.6188	0.6744	0.6260	0.6245	0.6898
40	0.7405	0.7291	0.7293	0.7430	0.7076	0.7979	0.7557	0.7948	0.7933	0.7410
50	0.9065	0.9149	0.9072	0.9096	0.9267	0.9683	0.9889	0.9646	0.9646	0.9970
60	1.0684	1.0658	1.0793	1.0687	1.0626	1.1335	1.1283	1.1341	1.1354	1.1261
80	1.3825	1.3827	1.3955	1.3772	1.3828	1.4583	1.4586	1.4616	1.4633	1.4587
100	1.6844	1.6844	1.6725	1.6868	1.6844	1.7670	1.7670	1.7655	1.7640	1.7670
ARD/%	/	27.753	1.124	0.168	34.006	/	5.153	0.254	0.222	6.157

表3 式(1)式(2)式(11)拟合乙缩醛蒸气压的结果对比

Table 3 Correlating results of the saturated vapor pressure of aldehyde acetal using Eq. (1), Eq. (2) and Eq. (11)

$T/^\circ\text{C}$	20	40	60	80	100	120	140	160	180	200	220	240	ARD/%
$P_{\text{exp}}$	2.938	8.423	20.87	46.09	92.41	171.0	296.3	485.7	760.0	1144	1667	2366	0
$P_{c_1}$	-13.05	20.981	37.01	53.08	87.24	157.5	282.0	478.7	765.6	1161	1682	2348	67.442
$n = 3$ $P_{c_2}$	3.064	8.270	20.68	46.06	92.62	171.4	296.4	485.4	759.5	1144	1668	2366	0.648
$P_{c_3}$	2.938	8.263	20.79	46.13	92.56	171.2	296.4	485.6	759.7	1144	1668	2366	0.239
$n = 4$ $P_{c_1}$	3.603	7.358	20.36	46.52	93.29	171.6	296.1	484.7	759.0	1144	1669	2365	3.378
$P_{c_2}$	3.011	8.338	20.86	46.11	92.40	171.0	296.3	485.7	760.0	1144	1667	2366	0.301
$P_{c_3}$	2.941	8.321	20.91	46.19	92.44	171.0	296.2	485.6	760.1	1144	1667	2366	0.152
$n = 5$ $P_{c_1}$	2.834	8.686	20.85	45.84	92.26	171.2	296.6	485.7	759.7	1144	1667	2366	0.647
$P_{c_2}$	2.951	8.392	20.88	46.10	92.39	171.0	296.3	485.7	760.0	1144	1667	2366	0.080
$P_{c_3}$	2.940	8.418	20.88	46.08	92.39	171.0	296.3	485.7	760.0	1144	1667	2366	0.018

由表3可见,当  $n = 3$  时式(11)的拟合误差仍然比  $n = 5$  时式(1)的拟合误差小。

3.3 乙酸异戊酯液体粘度( $\mu\text{Pa} \cdot \text{s}$ )的拟合

以乙酸异戊酯液体为例,  $\eta_{\text{exp}}$  为实验值(来自文献[10]第465页),  $\eta_{c_1}$  为式(1)拟合值,  $\eta_{c_2}$  为式

(2)拟合值 ( $n = 4, a = 0.97; n = 5, a = -0.57; n = 6, a = -0.32$ ),  $\eta_{c_3}$  为式(11)拟合值 ( $n = 4, k = 2, a = 1.66, b = 0.60; n = 5, k = 3, a = 1.10, b = 0.83; n = 6, k = 4, a = -0.20, b = -0.44$ )。

表4 式(1)式(2)式(11)拟合乙酸异戊酯液体粘度的对比

Table 4 Correlation of the viscosity of isoamyl acetate using Eq. (1), Eq. (2) and Eq. (11)

$T/^\circ\text{C}$	-60	-20	20	60	100	140	180	220	260	300	ARD/%
$n = 4$ $\eta_{\text{exp}}$	4.360	1.720	0.871	0.520	0.347	0.250	0.191	0.132	0.094	0.062	0
$\eta_{c_1}$	4.284	1.921	0.796	0.395	0.329	0.334	0.269	0.118	-0.012	0.112	33.032
$\eta_{c_2}$	4.285	1.922	0.776	0.408	0.345	0.335	0.258	0.112	-0.004	0.111	32.588
$\eta_{c_3}$	4.364	1.651	0.949	0.536	0.323	0.230	0.187	0.151	0.104	0.052	9.123
$n = 5$ $\eta_{c_1}$	4.345	1.778	0.806	0.507	0.390	0.273	0.157	0.107	0.131	0.051	12.885
$\eta_{c_2}$	4.360	1.720	0.871	0.524	0.333	0.255	0.196	0.143	0.095	0.051	3.840
$\eta_{c_3}$	4.361	1.714	0.885	0.508	0.346	0.258	0.188	0.131	0.094	0.062	1.487
$n = 6$ $\eta_{c_1}$	4.467	0.564	-0.092	0.259	0.514	0.395	0.117	0.022	0.194	0.039	60.262
$\eta_{c_2}$	4.356	1.720	0.871	0.520	0.346	0.253	0.186	0.135	0.094	0.062	0.661
$\eta_{c_3}$	4.360	1.720	0.871	0.520	0.346	0.253	0.186	0.134	0.094	0.062	0.654

从表 4 可见,在此例中,若用整数幂多项式(1)拟合,其平均相对偏差并不随着  $n$  的增加而单调减小,而式(2)及式(11)总是收敛的。

#### References

- 1 Yu Yinan. Nonlinear fitting method to calculate the rate constant of the acetic ether saponification. *Transaction of Donghua University*, 2003, 28(3):93-95.
- 2 Nasrifar Kh, Ayatollahi Sh and Moshfeghian. A compressed liquid density correlation. *Fluid Phase Equilibrium*, 2000, 168(2):149-163.
- 3 Gao Wuzi, Robert L, Robinson Jr and Khaled AM. Improved correlations for heavy n-paraffin physical properties. *Fluid Phase Equilibrium*, 2001, 179(1-2):207-216.
- 4 Cao Yongsheng, Chen Yiwei, Zhu Jinlin and Cheng Rongming. Application of MATLAB programming in chemometrics. *Computers and Applied Chemistry*, 2003, 20(3):274-276.
- 5 Hao Pingjiao and Li Shiyu. Applications of MATLAB in chemical engineering calculating. *Computers and Applied Chemistry*, 2000, 17(4):371-374.
- 6 Nie Zhaoguang, Liu Dianhua, Ying Weiyong and Fang Dingye. Model of intrinsic kinetics of dimethyl ether directly synthesis from syngas containing  $N_2$  over mixed catalyst. *Computers and Applied Chemistry*, 2003, 20(5):662-666.
- 7 Cheng Jiang, Huang Chaoyun and Yang Zhuoru. Process computation for chemical absorption with MATLAB. *Computers and Applied Chemistry*, 2003, 20(1):31-34.
- 8 Lin Jun and Gu Zhengni. Vapor-Liquid equilibria of the ternary system hexane + methyl cyclopentane + dibutyl-o-phthalate. *Computers and Applied Chemistry*, 2003, 20(5):615-617.
- 9 Liu Guangqi, Ma Lianxiang and Liu Jie. Handbook of the data of substance properties in chemistry and chemical engineering (inorganic volume). Beijing:Chemical Industry Press, 2002.
- 10 Liu Guangqi, Ma Lianxiang and Liu Jie. Handbook of the Data of Substance Properties in Chemistry and Chemical Engineering (organic volume). Beijing:Chemical Industry Press, ed 1, 2002.

#### 附中文参考文献

- 1 余逸男. 乙酸乙酯皂化反应实验数据的非线性拟合. *东华大学学报(自然科学版)*, 2003, 28(3):93-95.
- 2 曹永生, 陈奕卫, 朱金林, 成荣明. MATLAB 程序设计在化学计量学中的应用. *计算机与应用化学*, 2003, 20(3):274-276.
- 3 郝平娇, 李士雨. 浅谈 MATLAB 在化工计算中的应用. *计算机与应用化学*, 2000, 17(4):371-374.
- 4 聂兆广, 刘殿华, 应卫勇, 房鼎业. 多功能混合催化剂上含氮合成气一步法制二甲醚的本征动力学模型. *计算机与应用化学*, 2003, 20(5):662-666.
- 5 程江, 黄超云, 杨卓如. MATLAB 用于化学吸收过程动态模拟. *计算机与应用化学*, 2003, 20(1):31-34.
- 6 林军, 顾正桂. 正己烷-甲基环戊烷-邻苯二甲酸二丁酯三元体系汽液平衡研究. *计算机与应用化学*, 2003, 20(5):615-617.
- 7 刘光启, 马连湘, 刘杰. 化学化工物性数据手册(无机卷). 北京:化学工业出版社, 2002.
- 8 刘光启, 马连湘, 刘杰. 化学化工物性数据手册(有机卷). 北京:化学工业出版社, 2002.