

三次样条插值法计算药物在不同温度下的溶解度

查岭¹, 赵庆辉¹, 王英利²

(1. 周口市中心医院药学部, 河南 周口 466000; 2. 上海现代哈森(商丘)药业有限公司, 河南 商丘 476000)

摘要:目的 药物溶解度是药物的一个重要性质, 该文拟用插值法从有限温度下的溶解度计算任意温度下的溶解度。方法 选择 20 种电解质作为模型药物, 从文献中查得它们在不同温度下的溶解度, 并分别用三次样条插值法和幂函数展开法从这些模型药物在 0、10、30、40、60 和 80 °C 下溶解度计算得到它们在 20 °C 下的溶解度, 并与文献值作比较。结果 可用这两种方法计算药物在任意温度下的溶解度, 且三次样条插值法计算优于幂展开法。不仅如此, 插值法也可计算药物的其它一些重要性质, 如不同温度下的比热容、临界相对湿度等。结论 三次样条插值法在计算药物性质中可广泛应用。

关键词:三次样条插值; 幂函数展开; 温度; 溶解度

doi: 10.3969/j.issn.1009-6469.2017.05.009

Calculating drug solubility at different temperature by cubic spline interpolation method

ZHA Ling¹, ZHAO Qinghui¹, WANG Yingli²

(1. Department of Pharmacy, Zhoukou Central Hospital of Henan, Zhoukou, Henan 466000, China;

2. Shanghai Modern Hasen(Shangqiu) Pharmaceutical Co., Ltd., Shangqiu, Henan 476000, China)

Abstract: Objective To calculate the solubility at any temperature by cubic spline interpolation method according to the solubility at several temperatures. **Methods** Twenty drugs are selected as a model drugs and their solubility at different temperature is found from literature. The solubility of selected drugs at 20 °C is calculated by cubic spline interpolation method and Taylor series method according to the solubility at 0, 10, 30, 40, 60 and 80 °C, which is compared with the solubility from literature. **Results** The result shows the solubility at any temperature can be calculated by the two methods and the calculated solubility at 20 °C by cubic spline interpolation method, which is better than that from Taylor series method. Furthermore, the other property of drugs, such the specific heat and critical relative humidity at different temperature can be calculated by cubic spline interpolation method. **Conclusion** The interpolation method can widely be used in the calculation of drug property.

Key words: Cubic spline interpolation; Taylor series; Temperature; Solubility

药物的很多理化性质, 如在水中的溶解度、比热容等随温度的变化而变化, 而这些性质又是生产和理论研究中必须的。例如我们经常用到药物在某一特定温度下的溶解度, 但在文献上不可能查到药物在所有温度下的溶解度, 这时在某一特定温度下的溶解度就要从其它温度下的溶解度通过计算得到。如在兰氏化学手册上给出了一些物质在 0、10、20、30、40、60 和 80 °C 下溶解度^[1], 然而如果我们需 25 °C 下的溶解度数值时^[2], 兰氏化学手册上就没有给出了。比热容、溶剂的蒸气压也存在类似的情况。

1 研究思想介绍

1.1 3 种方法原理介绍

为了从部分温度数据点

下得到所有温度下某效应(如溶解度)的数据, 人们采用了各种方法, 这些方法的数学思想都非常明确, 分别简要介绍如下。

1.1.1 幂函数展开法 设 y 是 x 的函数, $y=f(x)$, 根据数学中的 Taylor 公式^[3], y 可用 x 的幂级数表示:

$$y = a_0 + a_1x + a_2x^2 + a_3x^3 + \cdots + a_nx^n + \cdots \quad (1)$$

物质的一些物理参数, 如摩尔热容与温度的关系, 就表示为如下幂函数的形式^[4]:

$$c_{p,m} = a + bT + cT^2 \quad (2)$$

1.1.2 参数拟合法 这种方法是一种半经验方法。通过建立模型, 经过一系列数学推导, 得出含有参数的物质特征关系式, 具体物质的参数可通过拟合实验数据获得, 从物质特征参数方程中可计算出物质在某一特定条件下的特性。如表示实际气

体的 Van der vaals 方程^[5]:

$$(p + \frac{a}{v_m^2})(v_m - b) = RT \quad (3)$$

1.1.3 插值法 插值法有多种方法,最常用的是三次样条插值(cubic spline interpolation)^[6],它是通过一系列数据点 (x_i, y_i) 的一条连续光滑曲线,由曲线上横坐标 x 值可求得纵坐标 y 的值。

以上3种方法各有特点:幂函数展开法和插值法中的参数缺乏物理意义,参数法中的参数有其明显的物理意义;幂函数展开法和插值法是纯数学方法的应用,参数法则是一种半经验半理论方法;幂函数展开法的数学基础是 Taylor 公式,插值法是把物质某性质看作是某个参数的连续光滑函数;在数学计算上,参数方法多是求解非线性方程,虽然计算机技术的发展给求解非线性方程带来方便,但总的来说,求解非线性方程的还是较为困难的工作;幂函数展开法和插值法则是通过求解线性方程组来得到系数,计算较为简单^[6-10]。

1.2 本研究关注点——三次样条插值法 人们对物质特性的计算大都用的是幂函数展开法或参数拟合法,很少用到插值法。在本研究中,我们尝试用插值法计算物质的特性参数。在水中的溶解度无疑是重要性质之一,在生产和理论研究中都重要应用,溶解度数据在文献中又较为完整,便于计算和验证,因此我们选用溶解度作为对象进行讨论。

本研究选用20种常见电解质,通过文献查得它们在0、10、30、40、60和80℃的溶解度数据,然后用插值方法计算出它们在20℃的溶解度,并把计算的结果同文献值作比较,结果表明,计算得到的溶解度与实验值相对误差大都在5%以内,表明可用三次样条插值法由有限温度下的溶解度计算任意温度下的溶解度。另外,我们也比较了三次样条插值法和幂展开法计算的溶解度,结果表明,三次样条插值法优于幂展开法。

我们之所以选择电解质作为模型药物是因为电解质理论发展得较早,我们可以从文献上查得其溶解度。从计算过程来说,电解质与药物之间并没有本质不同,我们当然也可用插值法计算药物的溶解度。

本研究的计算采用 Matlab 7.6 程序完成^[7-8]。

1.3 三次样条插值法数学原理^[6,8] 给定区间 $[a, b]$ 上有 $n+1$ 个节点: $a = x_0 < x_1 < \dots < x_n = b$ 和这些节点的函数值 $f(x_i) = y_i, i = 0, 1, \dots, n$ 。若函数 $S(x)$ 满足 $S(x) = y_i, i = 0, 1, \dots, n; S(x)$ 在每个小区间 $[x_i, x_{i+1}]$ 上是不高于三次的多项式, $S(x)$ 在区间 $[a, b]$ 上有两阶连续导数,则称 $S(x)$ 为 $f(x)$ 在

区间 $[a, b]$ 上的三次样条插值函数。要构造三次样条插值函数 $S(x)$:

$$S(x) = S_i(x) = a_i x^3 + b_i x^2 + c_i x + d_i \quad (4)$$

需要 $4n$ 个条件,插值条件 $S(x) = y_i$ 提供了 $n+1$ 个条件, $S(x)$ 连续、一阶、二阶导数连续共提供了 $3n-3$ 个条件:

$$S(x_i - 0) = S(x_i + 0), S'(x_i - 0) = S'(x_i + 0), S''(x_i - 0) = S''(x_i + 0), i = 1, 2, \dots, n-1。$$

除了这 $3n-3$ 还少两个条件,这两个条件的不同便构成了不同的插值方法,在 Matlab 中的 Spline 命令中使 $a_1 = a_2, a_{n-1} = a_n$,增加这两个条件后,三次样条插值函数 $S(x)$ 就可唯一确定了。

将以上条件代入式(4),我们便得到了 $4n$ 元一次方程组,通过解这个方程组,我们就可以得到 $S(x)$ 的表达式。随着计算机技术的发展,上述计算过程极为简单。

2 方法与结果

2.1 例举 为了验证上述猜想,我们从兰氏化学手册上查得了20种常见无机物在0、10、20、30、40、60和80℃下溶解度,结果见表1。在本研究中,我们用0、10、30和40℃下的溶解度数据计算20℃时的溶解度并把计算结果与20℃的溶解度作了比较。

2.2 三次样条插值法计算溶解度 插值法实质是解多元一次方程组,将表1中的温度作为自变量,相应的溶解度作为因变量,用 Matlab 7.6 计算,整个计算过程非常简单,只需要输入几条命令就行了。计算的思路如下:先得到每种物质的样条插值多项式,再由样条插值多项式计算20℃时的溶解度。现以6点法为例具体做法如下。

在 Matlab 命令窗口输入温度 x 和对应的溶解度 y 后,再输入`pp = spline(x, y); ppval(pp, 20)`即可。如用三次样条插值6点法计算 MgCl_2 在20℃的溶解度,首先在 Matlab 命令窗口中输入温度 $x = [0 \ 10 \ 30 \ 40 \ 60 \ 80]$;再输入对应温度下的溶解度 $y = [52.9 \ 53.6 \ 55.8 \ 57.5 \ 61 \ 66.1]$;接着输入`pp = spline(x, y);`,得到计算 MgCl_2 在20℃的三次样条插值系数矩阵,最后用`ppval(pp, 20)`即可得到 MgCl_2 在20℃的溶解度54.5080。即在 Matlab 命令窗口中输入代码:

```
x = [0 10 30 40 60 80]; y = [52.9 53.6 55.8 57.5 61 66.1]; pp = spline(x, y); ppval(pp, 20), 给出的结果为 54.5080。
```

也可编一个小程序,一次计算出所有结果,代码如下:

表 1 20 种常见物质在不同温度下的溶解度(g/100 g 水)

物质名称	0 ℃	10 ℃	20 ℃	30 ℃	40 ℃	60 ℃	80 ℃
MgCl ₂	52.9	53.6	54.6	55.8	57.5	61	66.1
Li ₂ CO ₃	1.54	1.43	1.33	1.26	1.17	1.01	0.85
LiI	151	157	165	171	179	202	435
NH ₄ Cl	29.4	33.2	37.2	41.4	45.8	55.3	65.6
NH ₄ NO ₃	118	150	192	242	297	421	580
(NH ₄) ₂ SO ₄	70.6	73	74.5	78	81	88	95
CaCl ₂ ·6H ₂ O	59.5	64.7	74.5	100	128	137	147
I ₂	0.014	0.02	0.029	0.039	0.052	0.1	0.225
KBr	53.6	59.5	65.3	70.6	75.4	85.5	94.9
K ₂ CO ₃	105	108	111	114	117	127	140
KCl	28	31.2	34.2	37.2	40.1	45.8	51.3
KI	128	136	144	153	162	167	192
KNO ₃	13.9	21.2	31.6	45.3	61.3	106	167
KIO ₃	4.6	6.27	8.08	10.3	12.6	18.3	24.8
AgNO ₃	122	167	216	265	311	440	585
NaBr	80.2	85.2	90.8	98.4	107	118	120
NaH ₂ PO ₄	56.5	69.8	86.9	107	133	173	211
NaCl	35.7	35.8	35.9	36.1	36.4	37.1	38
NaNO ₃	73	80.8	87.6	94.9	102	122	148
ZnCl ₂	342	363	395	437	452	488	541

表 2 用三次样条插值法计算的溶解度与文献值比较(20 ℃)(g/100 g 水)

物质名称	文献值	4 点法	相对偏差/%(4 点法)	6 点法	相对偏差/%(6 点法)	第 4 列与第 6 列差
MgCl ₂	54.6	54.5333	-0.12	54.508	-0.17	0.05
Li ₂ CO ₃	1.33	1.3417	0.88	1.3429	0.97	-0.09
LiI	165	163.6667	-0.81	163.1227	-1.14	0.33
NH ₄ Cl	37.2	37.2	0.00	37.2018	0.00	0
NH ₄ NO ₃	192	192.1667	0.09	192.1871	0.10	-0.01
(NH ₄) ₂ SO ₄	74.5	75.4	1.21	75.3939	1.20	0.01
CaCl ₂ ·6H ₂ O	74.5	78.55	5.44	77.4393	3.95	1.49
I ₂	0.029	0.0283	-2.41	0.0284	-2.07	-0.34
KBr	65.3	65.2333	-0.10	65.2764	-0.04	-0.06
K ₂ CO ₃	111	111	0.00	111.0644	0.06	-0.06
KCl	34.2	34.25	0.15	34.2491	0.14	0.01
KI	144	144.3333	0.23	144.0184	0.01	0.22
KNO ₃	31.6	31.8	0.63	31.8911	0.92	-0.29
KIO ₃	8.08	8.18	1.24	8.1912	1.38	-0.14
AgNO ₃	216	215.8333	-0.08	216.7301	0.34	-0.42
NaBr	90.8	91.2	0.44	91.046	0.27	0.17
NaH ₂ PO ₄	86.9	86.2833	-0.71	85.8031	-1.26	0.55
NaCl	35.9	35.9167	0.05	35.911	0.03	0.02
NaNO ₃	87.6	87.9667	0.42	88.0331	0.49	-0.07
ZnCl ₂	395	401	1.52	402.6595	1.94	-0.42
平均相对偏差 ^a	-	-	0.83	-	0.82	-

注：^a 相对偏差 = $\frac{\text{计算值} - \text{文献值}}{\text{文献值}} \times 100\%$ ；平均相对偏差 = $\frac{\sum |\text{相对偏差}|}{n} \times 100\%$ 。

表3 用幂函数展开法计算的溶解度与文献值比较(20℃)(g/100g水)

物质名称	文献值	二次多项式	相对偏差/%	三次多项式	相对偏差/%	四次多项式	相对偏差/%
MgCl ₂	54.6	54.5906	-0.02	54.6507	0.09	54.5951	0.01
Li ₂ CO ₃	1.33	1.347	1.28	1.3421	0.91	1.3399	-0.74
LiI	165	137.4124	-16.72	176.4271	6.93	167.1204	-1.29
NH ₄ Cl	37.2	37.1884	-0.03	37.1873	-0.03	37.1946	0.01
NH ₄ NO ₃	192	193.0131	0.53	193.9502	1.02	192.5071	-0.26
(NH ₄) ₂ SO ₄	74.5	75.4254	1.24	75.2687	1.03	75.3818	-1.18
CaCl ₂ ·6H ₂ O	74.5	90.4034	21.35	85.6779	15.00	81.669	-9.62
I ₂	0.029	0.0182	-37.24	0.0308	6.21	0.0287	1.03
KBr	65.3	64.9808	-0.49	65.1081	-0.29	65.1438	0.24
K ₂ CO ₃	111	110.1661	-0.75	110.6851	-0.28	110.8396	0.14
KCl	34.2	34.1953	-0.01	34.2303	0.09	34.2473	-0.14
KI	144	143.64	-0.25	147.3199	2.31	145.4183	-0.98
KNO ₃	31.6	30.8104	-2.50	31.6524	0.17	31.635	-0.11
KIO ₃	8.08	8.1347	0.68	8.122	0.52	8.1515	-0.88
AgNO ₃	216	210.3353	-2.62	212.7052	-1.53	213.861	0.99
NaBr	90.8	94.0144	3.54	91.5259	0.80	91.4948	-0.77
NaH ₂ PO ₄	86.9	91.3394	5.11	88.0505	1.32	87.3587	-0.53
NaCl	35.9	35.9296	0.08	35.9179	0.05	35.9252	-0.07
NaNO ₃	87.6	86.3018	-1.48	87.5306	-0.08	87.778	-0.20
ZnCl ₂	395	399.1211	1.04	403.8172	2.23	399.1384	-1.05
平均相对偏差 ^a	-	-	4.85	-	2.04	-	1.01

注：^a 相对偏差 = $\frac{\text{计算值} - \text{文献值}}{\text{文献值}} \times 100\%$ ；平均相对偏差 = $\frac{\sum |\text{相对偏差}|}{n} \times 100\%$ 。

$x = [0 \ 10 \ 30 \ 40 \ 60 \ 80]$;
 $y = [52.9 \ 53.6 \ 55.8 \ 57.5 \ 61 \ 66.1; 1.54 \ 1.43 \ 1.26 \ 1.17 \ 1.01 \ 0.85; \dots]\%$ 输入表中所列物质在对应温度下的溶解度
for $n = 1:20$; $a = y(n,:)$; $pp = \text{spline}(x,a)$; $z(1,n) = \text{ppval}(pp,20)$; end;
给出表中所列物质在 20℃溶解度分别结果为:
[54.5080、1.3429、163.1227、37.2018、192.1871、75.3939、77.4393、0.0284、65.2764、111.0644、34.2491、144.0184、31.8911、8.1912、216.7301、91.0460、85.8031、35.9110、88.0331、402.6595]。

我们分别用四个数据点(0、10、30、40℃)和七个数据点(0、10、30、40、60、80℃)计算 20℃时的溶解度,并将其与文献中的值作比较,结果见表 2、图 1。

2.3 幂函数展开法计算溶解度 作为对比,我们用幂函数展开法^[9-10]计算了选择电解质在 20℃下的溶解度,方法是先用 0、10、30、40、60、80℃下的溶解度对温度进行最小二乘拟合,得到多项式的系数,再将 20℃代入多项式便得到了电解质在 20℃时的溶解度,并将其与文献中的值作比较,结果见表 3、图 2。

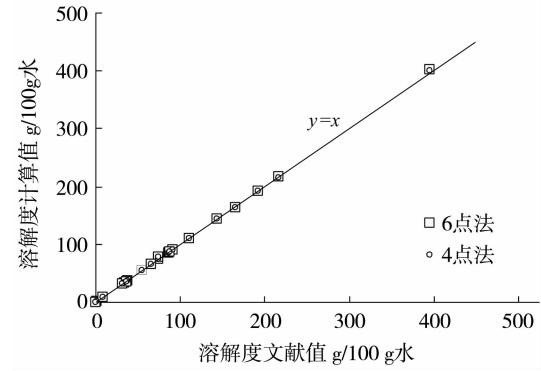


图1 用三次样条插值法计算的溶解度(20℃)

用 Matlab 计算,整个过程也非常简单,同样也是几条命令就可以了。具体做法如下:

在 Matlab 命令窗口输入温度 x 和对应的溶解度 y 后,再输入 $p = \text{polyfit}(x,y,n)$; $\text{polyval}(p,20)$ 即可(n 是多项式的次数)。如用幂函数展开法(4 次多项式)计算 MgCl_2 在 20℃的溶解度,首先在 Matlab 命令窗口中输入温度 $x = [0 \ 10 \ 30 \ 40 \ 60 \ 80]$; ,再输入对应温度下的溶解度 $y = [52.9 \ 53.6 \ 55.8 \ 57.5 \ 61 \ 66.1]$; ,接着输入 $p = \text{polyfit}(x,y,4)$; ,得到计算 MgCl_2 溶解度的最小 2 乘拟合系数矩阵,最后

用 $\text{polyval}(p, 20)$ 即可得到 MgCl_2 在 $20\text{ }^\circ\text{C}$ 的溶解度 54.5951。即中 Matlab 命令窗口中输入以下代码:

$x = [0\ 10\ 30\ 40\ 60\ 80]; y = [52.9\ 53.6\ 55.8\ 57.5\ 61\ 66.1]; p = \text{polyfit}(x, y, 4); \text{polyval}(p, 20)$, 给出的结果为 54.5951, 同样, 也可一次计算出表 3 中所列物质的溶解度。

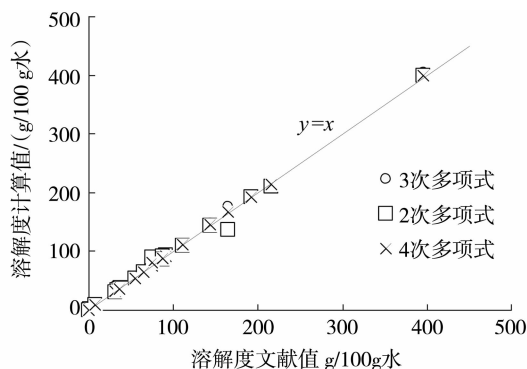


图2 用幂函数展开法计算的溶解度 ($20\text{ }^\circ\text{C}$)

3 讨论

表2、图1是三次样条插值法计算的溶解度与文献给出的溶解度比较的结果。如果计算值与文献值相等,则图中数据点在参考线 $y=x$ 上。由图和表可见,无论是4点法或6点法计算得溶解度都与文献值非常接近。为了进一步证明两者的等效性,我们对4点法和6点法得到的相对偏差用单样本双侧 t 检验进行了检验,得到的为 t 统计量为0.512 9, 小于临界值 $t_{0.05, 19} = 2.093\ 0, P > 0.05$, 表明在4点法和6点法得到的结果差异无统计学意义。由表2还可以看出两者的平均相对偏差分别为0.83%和0.82%, 两者计算的准确度也是非常相近的。

虽然4点法和6点法计算的结果基本一致,但如果由实验得到数据,两者的工作量是不一样的,4点法需要4个点的实验数据,6点法则需要6个点的实验数据,后都的实验量远大于前者,建议用4点法较好。

表3、图2是幂函数展开法计算的溶解度与文献给出的溶解度比较的结果。同理如果计算值与文献值相等,则图中数据点在参考线 $y=x$ 上。图2中的数据点大体分布在参考线 $y=x$ 上,表明幂函数展开法计算的溶解度与文献给出的溶解度也非常接近。由图和表中还可以看出,虽然图2中的数据点大体分布在参考线 $y=x$ 上,但和图1相比则明显

偏离直线,表明幂函数展开法得到的结果没有三次样条插值法得到的结果好。从表3我们可以看出幂展开的次数越高,平均相对偏差越小,但这并不意味着拟合的次数越高越好,因为拟合的次数最高为数据点数目减1,且拟合次数增高,实验中的随机误差也引入拟合公式中。

综上所述,我们可用三次样条插值法计算药物在任意温度下的溶解度,且计算结果要比幂展开法计算要精确。在三次样条插值法计算过程中,有4个数据点就够了,需要的数据点也较少,在物质参数估算中有一定优越性。

应该说明的是,插值法作为一种数学方法在科学研究上有广泛应用,我们在本研究中仅讨论了它在溶解度计算上的应用。药物的许多性质,如任意温度下的热容、液体在任意温度下的蒸汽压、任意温度下的临界相对湿度等都可从有限实验数据中通过计算得到。另外,Matlab是一个非常优秀的数学软件,被广泛应用到科学研究中,我们希望通过本研究能使该软件在医药领域得到更广泛的应用。

参考文献

- [1] James G. Speight, LANGE'S HANDBOOK OF CHEMISTRY [M]. 16th Edition, MCGRAW-HILL, 2004.
- [2] WANG Y, ZHAN X, XIANG C, et al. Calculating the critical relative humidity from the solubility of electrolyte according to extended non-random two liquid model [J]. Chem Pharm Bull, 2009, 57 (9): 943-947.
- [3] 同济大学应用数学系. 高等数学(上册) [M]. 北京: 高等教育出版社, 2002.
- [4] 傅献彩. 物理化学(上册) [M]. 北京: 人民教育出版社, 2005.
- [5] 张三慧. 大学物理学(第二册) [M]. 北京: 清华大学出版社, 1999.
- [6] 任雪娟, 张耀存, 马开玉. 数值计算方法 [M]. 北京: 气象出版社, 2007.
- [7] 樊天锁, 芮兵. 样条插值的 MATLAB 实现 [J]. 佳木斯大学学报(自然科学版), 2011, 29(2): 238-240.
- [8] 耿爱成. Matlab 在三次样条函数教学中的应用 [J]. 价值工程, 2016, 35(18): 181-182.
- [9] 孙祥, 徐流美, 吴清. MATLAB 7.0 基础教程 [M]. 北京: 清华大学出版社, 2005.
- [10] 施阳. MATLAB 语言工具箱: TOOLBOX 实用指南 [M]. 西安: 西北工业大学出版社, 1998.

(收稿日期: 2016-10-09, 修回日期: 2016-12-03)