

基于自适应模糊神经网络预测金属—EDTA 配合物稳定常数

多杰扎西

(青海师范大学 民族师范学院,青海 西宁 810008)

摘要:采用自适应模糊神经网络的方法,以金属离子外层主量子数(n)、电荷(Z)、半径(r)、适配价轨道数因子(w)及价电子结构因子(S)等为参数,关联金属—EDTA 配合物稳定常数。利用减法聚类算法确定模糊神经网络的结构,并结合模糊推理系统进行该网络参数的调整,网络仿真的结果是满意的。在此基础上,预测了 13 种金属—EDTA 配合物稳定常数。

关键词:自适应模糊神经网络;金属离子;EDTA 配合物;稳定常数

中图分类号:O655.29

文献标识码:A

文章编号:1008-858X(2014)03-0045-05

引 言

各类配合物的稳定常数的实验测定和预测历来都是配位化学的研究热点之一^[1]。由于配合物组成的特殊性,使得影响配合物稳定性既与金属离子和配体的本性有关,也与金属离子与配体二者的相互作用有关,因而实验测定或理论模型预测配合物的稳定常数比较困难。已有文献报道利用一种新型函数连接型神经网络研究金属离子—EDTA(乙二胺四乙酸)稳定常数^[2],自适应模糊神经网络预测金属离子性质也有报道^[3],但基于自适应模糊神经网络预测金属—EDTA 配合物稳定常数未见报道。本文基于自适应模糊神经网络,利用减法聚类算法确定模糊神经网络结构,通过对模糊推理系统进行网络参数调整,提高了网络的收敛速度和建模精度。应用于金属离子—EDTA 配合物稳定常数的预测,取得了满意的结果。

1 自适应模糊神经网络原理简介

自适应模糊神经网络,在功能上等效于一个具有 T-S 规则的模糊推理系统^[4]。主要包

括利用减法聚类算法确定自适应神经网络模糊推理系统(ANFIS: Adaptive Network-based Fuzzy Inference System)结构(一般为一个五层结构)、前提参数和结论参数的学习与调节等步骤。采用混合算法进行训练,通过对大量已知数据的学习,获得与输入、输出数据相匹配的隶属度函数和模糊规则,以反映数据的变化特点,完成将输入空间的数据映射到输出空间,从而实现自适应模糊神经网络建模。

2 预测金属—EDTA 配合物稳定常数的自适应模糊神经网络的建立

2.1 输入参数的选择

根据配位化学键理论对影响配合物稳定性的分析,选取金属离子外层电子主量子数(n)、电荷(Z)、半径(r)、适配价轨道数因子(w)及价电子结构因子(S)等共计 5 个参数作为自适应模糊神经网络的初始输入(表 1)。其中适配价轨道数因子(w)按文献中的方法计算^[2]。自定义价电子结构因子(S)按下式计算,

收稿日期:2014-04-04;修回日期:2014-04-22

作者简介:多杰扎西(1964-)男,教授,从事化学教育、分析化学教学与科研工作。

$$S = \frac{\sum_i e_i}{n} = \frac{E}{n}。$$

其中 E 是金属离子 f, d, s 各价层电子数之和; n 即金属离子外层电子的主量子数。价电子结构因子 S 可认为反映金属离子的静电势(分母隐含大小概念)。

容易理解,最外层电子主量子数 n 、离子电荷及半径表征金属离子的性质,而适配价轨道

数因子和价电子结构因子被认为能表征金属元素在成键中复杂的电子交换—相关作用和静电力作用,即能表征金属离子与配位体的相互作用。这样,对于指定配位体(性质是确定的)的配合物,选定的 5 个参数较全面地反映了影响配合物稳定性的诸多因素,作为研究金属离子—EDTA 配合物体系的特征变量,可用于建模。

表 1 金属离子的结构参数和 EDTA 配合物的 $\log K$ 值

Table 1 The structure parameters of metal ions and value of $\log K$ of EDTA complexes

No.	M(z)	n	Z	r	w	S	$\log K_{M-EDTA}$			绝对误差 ΔE
							实验值 ^[5]	本文值	文献值 ^[2]	
1	Ag(1)	4	1	1.26	-2	2.500 0	7.32	7.324 5	7.301 0	0.004 5
2	Al(3)	2	3	0.50	-2	0.000 0	16.10	16.099 6	15.827 7	-0.000 4
3	Am(3)	6	3	0.99	3	1.000 0	18.10	18.054 8	17.900 3	-0.045 2
4	Ba(2)	5	2	1.35	3	0.000 0	7.78	7.775 7	7.293 8	-0.004 3
5	Be(2)	1	2	0.31	-2	0.000 0	9.30	9.301 7	9.545 5	0.001 7
6	Bi(3)	6	3	0.96	-3	4.333 3	22.80	22.800 8	22.804 8	0.000 8
7	Ca(2)	3	2	0.99	3	0.000 0	11.00	11.000 3	11.385 8	0.000 3
8	Cd(2)	4	2	0.97	-2	2.500 0	16.40	16.414 8	17.255 7	0.014 8
9	Ce(3)	5	3	1.03	3	0.200 0	16.80	16.542 7	16.544 6	-0.257 3
10	Cf(3)	6	3	0.98	3	1.500 0	19.00	19.015 5	19.516 6	0.015 5
11	Cm(3)*	6	3	0.99	3	1.166 7	18.40	18.475	18.449 4	0.075 0
12	Co(2)	3	2	0.74	-1	2.333 3	16.30	16.252 9	15.774 8	-0.047 1
13	Cr(2)	3	2	0.84	1	1.333 3	13.60	13.588 5	13.395 4	-0.011 5
14	Cr(3)*	3	3	0.64	1	1.000 0	23.00	23.097	23.002 4	0.097 0
15	Cu(2)	3	2	0.72	-2	3.000 0	18.70	18.740 6	18.186 2	0.040 6
16	Dy(3)	5	3	0.91	3	1.800 0	18.00	18.192 4	17.989 4	0.192 4
17	Er(3)*	5	3	0.88	3	2.200 0	18.18	18.498	18.623 8	0.318 0
18	Eu(3)	5	3	0.95	3	1.200 0	17.90	17.230 7	16.964 9	-0.669 3
19	Fe(2)	3	2	0.76	0	2.000 0	14.30	14.366 0	15.605 8	0.066 0
20	Fe(3)	3	3	0.64	0	1.666 7	24.20	24.208 5	25.203 2	0.008 5
21	Ga(3)	3	3	0.62	-2	3.333 3	20.20	20.209 8	20.792 0	0.009 8
22	Gd(3)	5	3	1.10	-2	1.400 0	17.20	17.199 7	17.504 2	-0.000 3
23	Hg(2)	5	2	1.10	-2	4.800 0	21.80	21.801 6	21.813 4	0.001 6
24	Ho(3)	5	3	0.89	3	2.000 0	18.10	18.126 1	18.312 7	0.026 1
25	In(3)	4	3	0.81	-2	2.500 0	24.90	24.888 8	24.093 5	-0.011 2
26	La(3)	5	3	1.06	3	0.000 0	16.30	16.429 0	16.677 9	0.129 0
27	Mg(2)*	2	2	0.65	-2	0.000 0	8.64	8.666 1	7.548 1	0.026 1
28	Li(1)	1	1	0.68	-2	0.000 0	2.79	2.790 7	2.600 0	0.000 7
29	Lu(3)	5	3	0.85	3	2.800 0	19.80	19.657 3	19.708 7	-0.142 7
30	Mn(2)	3	2	0.80	0	1.666 7	13.80	13.754 0	13.311 7	-0.046 0

(续表1)

No.	M(z)	n	Z	r	w	S	logK _{M-EDTA}			绝对误差 △E
							实验值 ^[5]	本文值	文献值 ^[2]	
31	Na(1)	2	1	0.95	-2	0.000 0	1.66	1.659 9	1.838 1	-0.000 1
32	Nd(3)	5	3	1.00	3	0.600 0	16.60	16.626 9	16.982 2	0.026 9
33	Ni(2)*	3	2	0.72	-1	2.666 7	18.50	17.465	16.998 2	-1.035 0
34	Pb(2)	6	2	1.20	-3	4.333 3	18.30	18.298 8	18.292 9	-0.001 2
35	Pd(2)	4	2	0.86	-1	2.000 0	18.50	18.494 6	18.383 3	-0.005 4
36	Pm(3)*	5	3	0.98	3	0.800 0	17.40	16.766	16.672 4	-0.634 0
37	Pr(3)	5	3	1.01	3	0.400 0	16.50	16.558 0	16.531 0	0.058 0
38	Pu(3)	6	3	1.00	3	0.833 3	18.10	18.117 6	17.432 8	0.017 6
39	Ra(2)	6	2	1.40	3	0.000 0	7.40	7.404 3	7.476 4	0.004 3
40	Sc(3)	3	3	0.81	3	0.000 0	23.10	23.098 7	22.367 3	-0.001 3
41	Sm(3)	5	3	0.96	3	1.000 0	16.40	16.902 1	16.796 6	0.502 1
42	Sn(2)	5	2	1.02	-3	2.400 0	22.10	22.100 5	22.087 9	0.000 5
43	Sr(2)	4	2	1.03	3	0.000 0	8.80	8.802 0	8.807 3	0.002 0
44	Tb(3)	5	3	0.92	3	1.600 0	17.60	17.756 6	17.740 3	0.156 6
45	Ti(3)	3	3	0.77	2	0.333 3	21.30	21.300 4	22.787 6	0.000 4
46	Tl(3)	5	3	0.95	-2	4.800 0	22.50	22.498 9	22.520 9	-0.001 1
47	Tm(3)	5	3	0.87	3	2.400 0	19.40	18.858 5	18.960 6	-0.541 5
48	V(2)	3	2	0.88	1	1.000 0	12.70	12.709 5	12.477 0	0.009 5
49	V(3)	3	3	0.74	2	0.666 7	25.90	25.898 8	25.633 9	-0.001 2
50	Y(3)	4	3	0.93	3	0.000 0	18.30	18.305 7	19.170 2	0.005 7
51	Yb(3)	5	3	0.86	3	2.600 0	18.70	19.224 9	19.025 6	0.524 9
52	Zn(2)	3	2	0.74	-2	3.333 3	16.40	16.364 5	16.073 2	-0.035 5
53	Pu(4)	6	4	0.90	3	0.666 7	17.66	17.669 0	18.028 8	0.009 0
54	Pu(6)	6	6	0.81	3	0.333 3	17.66	17.657 6	17.587 4	-0.002 4
55	Th(4)	6	4	0.99	3	0.000 0	23.20	23.201 9	23.037 8	0.001 9
56	U(4)*	6	4	0.93	3	0.333 3	17.50	19.449 6	18.567 3	1.949 6
57	V(5)	3	5	0.59	3	0.000 0	18.05	18.053 6	18.081 7	0.003 6
58	Zr(4)	4	4	0.79	3	0.000 0	19.40	19.389 7	19.316 5	-0.010 3
1	Co(3)	3	3	0.63	0	2.000 0		23.858	24.667 9	
2	Au(1)	5	1	1.37	-2	4.800 0		16.035	14.753 3	
3	Ac(3)	6	3	1.11	3	0.000 0		19.458	15.838 9	
4	Pa(3)	6	3	1.05	3	0.333 3		18.912	16.833 2	
5	Np(3)	6	3	1.01	3	0.666 7		18.200	16.987 9	
6	Bk(3)	6	3	0.98	3	1.333 3		18.474	18.955 4	
7	Es(3)	6	3	0.98	3	1.666 7		19.616	20.076 4	
8	Fm(3)	6	3	0.97	3	1.833 3		19.648	20.627 9	
9	Md(3)	6	3	0.96	3	2.000 0		19.643	21.203 5	
10	No(3)	6	3	0.95	3	2.166 7		19.601	21.799 9	
11	Lr(3)	6	3	0.94	3	2.333 3		19.529	22.418 9	
12	Pa(5)	6	5	0.90	3	0.000 0		21.416	22.219 2	
13	Np(5)	6	5	0.88	3	0.333 3		20.131	19.370	

* 核对数据集

2.2 自适应模糊神经网络的实现

借助于 MATLAB 的模糊逻辑工具箱提供的有关函数^[6], 主要有 `anfis`、`genfis2`、`evalfis` 等, 通过编制程序可以实现自适应模糊神经网络建模。本文预测金属—EDTA 配合物稳定常数的自适应模糊神经网络已在 lenovo PC 机 (CPU: 2.2GHz) 调试通过, 程序运行过程大致如下。

取表 1 中 58 个已知 $\log K$ 值为样本, 其中的 51 个为训练集, 而挑选出有代表性的 $C_m(3)$ 等 7 个 (以星号标记) 样例为核对数据集, 监测训练过程。因网络的输入变量数较多, 为了减少网络的模糊规则, 压缩网络规模, 先利用减法聚类算法, 以类半径为 0.5 生成初始的模糊神经网络结构。在此基础上, 调用 `anfis` 函数, 根据上述原理, 利用自适应神经网络混合学习算

法来确定模糊推理系统的参数。经过对输入、输出数据以初始步长为 0.22 的 130 次迭代, 即时训练均方根误差达到最小 (Minimal TRN-RMSE = 0.1616), 核对误差不再下降 (图 1), 停止训练, 得到经过调整的模糊神经网络模型, 其网络参数得到优化。自适应模糊推理系统信息 (ANFIS info): 节点数 (Number of nodes), 80; 线性参数数目 (Number of linear parameters), 36; 非线性参数数目 (Number of nonlinear parameters), 60; 总参数数目 (Total number of parameters), 96; 训练数据对 (Number of training data pairs), 51; 校核数据对 (Number of checking data pairs), 7; 模糊规则数目 (Number of fuzzy rules), 6。58 个样例的 $\log K$ 计算值与 13 个未知样的预测结果见表 1。

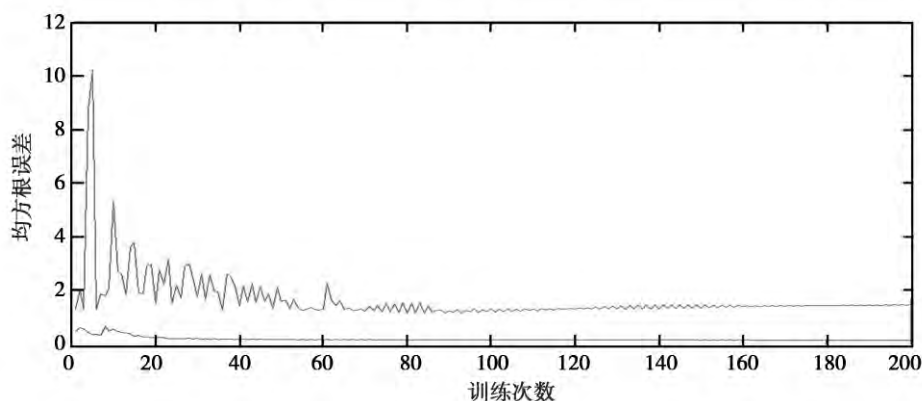


图 1 网络训练误差图

Fig. 1 Training errors of the network

3 结果与讨论

表 1 列出了自适应模糊神经网络模型的计算结果。51 个样例计算结果与文献实验值相比, 均方根误差仅为 0.16882, 7 个测试样例计算结果与文献值、实验值相比, 均方根误差为 0.8776。作为比较, 表 1 也列出了文献中采用函数连接型神经网络对同一样本的计算结果。其中 51 个样例计算结果与文献值、实验值相比, 均方根误差达 0.49282, 7 个测试样例计算结果均方根误差为 0.8714。可见自适应模糊

神经网络模型预测精度高。该网络对金属—EDTA 配合物的 $\log K$ 的拟合和预测是成功的。所选 5 个输入参数较全面地反映了影响配合物稳定的诸多因素, 用于网络建模是合理的。

本文建立的自适应模糊神经网络模型, 一方面由于自定义价电子结构因子 S 计算的 E 部分本身包含 f 、 d 、 s 因素, S 本身又隐含静电势概念, 与电负性 X_p 意义相当, 因此可适当地压缩输入变量, 即可将原 8 个原始变量 (n 、 f 、 d 、 s 、 Z 、 r 、 X_p 、 w) 调整为 5 个输入变量 (n 、 z 、 r 、 w 、 S), 仍采用快速单次减法聚类算法, 分割输入空间, 确定初始模糊神经网络结构, 从而较好地解决

了模糊规则数目可能出现的组合爆炸问题,网络模型仅用6条模糊规则就能完成从输入到输出的模糊推理。另一方面,又由于采用核对数据监测训练过程,避免了过训练,从而使所建立的网络模型稳定性高,泛化能力强。可以看到,本文的训练均方根误差与拟合和预测均方根误差相差不大,而一般人工神经网络训练均方根误差可达到很低,但拟合和预测均方根误差却很大,往往在万倍左右^[7]。表1下部列出了目前尚缺实验数据的部分过渡金属-EDTA配合物的稳定常数的预测值,可作为相关应用时的参考。

4 结 语

基于自适应模糊神经网络对金属-EDTA配合物的稳定常数的拟合和预测是成功的。所选5个输入参数较全面地反映了影响配合物稳定的诸多因素,用于网络建模是合理的。自适应模糊神经网络作为一个很有特色的神经网络,可应用于复杂的非线性函数映射研究,有望

成为元素和化合物构效关系研究的有效辅助手段。

参考文献:

- [1] 杨兴华,印春生,潘忠孝. 镧系金属离子 Ln^{3+} 定量构效关系的神经网络研究[J]. 高等学校化学学报, 2001, 22(9):1466-1471.
- [2] 杨兴华,印春生,潘忠孝,等. 人工神经网络用于金属离子-乙二胺四乙酸络合物稳定性研究[J]. 分析化学, 1999, 27(11):1268-1273.
- [3] 吴启勋,宋萍. 自适应模糊神经网络预测阳离子标准的研究[J]. 计算机与应用化学, 2007, 24(10):1421-1424.
- [4] 田丽,段争光,金礼力. 自适应模糊神经网络在用电预测中的应用[J]. 测控技术, 2004, 23(6):78-79.
- [5] John A D. Lange's Handbook of Chemistry (fifteenth Edition) (中译本) [M]. 北京:北京科学出版社, 2003, Section 8.
- [6] 吴晓莉,林哲辉. MATLAB 辅助模糊系统设计 [M]. 西安:西安电子科技大学出版社, 2002:131.
- [7] 吴启勋,胡树青. 贝叶斯正则化神经网络预测金属晶体结合能的研究[J]. 计算机与应用化学, 2004, 21(4):604-608.

Prediction of the Stabilities of Metal - EDTA Complex with Adaptive Fuzzy Neural Network

DUOJIE Zhaxi

(Nationality Teacher's College Qinghai Normal University, Xining 810008, China)

Abstract: An adaptive fuzzy neural network was applied to study the relationships between the structural parameters of metal ions and their stability constants of EDTA complexes, associated with the structural parameters such as principal quantum number in the outer-shell (n), electric charges (Z), ionic radii (r), the factor of adaptive valence orbit number (w) and the structural factor in the valence layer (S). Subtractive clustering algorithm was used to confirm the structure of fuzzy neural network, and combined fuzzy inference systems to process regulation of the network parameters. The simulation results were satisfactory, and stability constants data of 13 metallic ions without experimental data were predicted.

Key words: Adaptive Fuzzy Neural Network; Metallic ions; EDTA complexes; Stability constants