最优化方法在天气预报中的应用

冯耀煌　杨　旭
（辽宁省气象科学研究所）

提　要

本文根据最优化方法原理，提出了选择非线性最优预报因子和建立非线性预报方程的办法。试验表明：用此方法选出来的非线性预报因子和建立的方程比线性的好。最后还指出了此方法可推广到预报量离散的情况，建立非线性的判别方程，同时可用于长、中、短期天气预报。

一、前　言

最优化方法是解决最优问题的一种数值计算方法，它随着电子计算机的发展而迅速发展起来，现已广泛应用于空间技术、工程设计、无线通讯、农业规划等许多方面。对于在天气预报中如何应用最优化方法，这还是一个值得探讨的问题。天气预报也有一个寻优的问题，例如如何选择最优的预报因子？如何选择合适的参数建立最优的预报方程？本文就是试用最优化方法来解决这些问题，经过试验，取得了较好的效果，为今后寻找最优预报因子和建立最优的预报方程开辟了一条新的途径。

二、用最优化方法寻找预报因子

1.问题的提出

统计天气预报效果的好坏，虽然与预报方法有关，但与因子选择的关系更为密切，可以说，统计天气预报的基础是因子的选择。因此如何选择符合实际情况的和与预报量相关密切的因子是非常重要的工作。在过去，判断预报因子好坏一般是计算预报量 y 与预报因子 x 的相关系数，其公式为

$$R=\frac{\sum_{i=1}^{N}(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum_{i=1}^{N}(x_i-\bar{x})^2\sum_{i=1}^{N}(y_i-\bar{y})^2}} \tag{1}$$

其中 $\bar{x}=\frac{1}{N}\sum_{i=1}^{N}x_i$，$\bar{y}=\frac{1}{N}\sum_{i=1}^{N}y_i$。在给定的显著性水平$\alpha$下，若 $R>R_\alpha$（临界值），则认为因子通过显著性检验，是比较好的，否则就认为是不好的。但这只是在线性假设前提下得出的结论，如果因子本身与预报量是非线性的关系，那么，这种办法就不能选出真正的好因子，并且会把一些非线性的好因子漏掉。要选出真正符合实际情况的好因子，那只有计算非线性相关系数，用非线性相关系数的好坏来决定因子的取舍。这样不但可以防止漏掉有用的非线性因子，而且原来选出的线性因子变为非线性后，相关系数还可以提高。这是因为因子和预报量的关系，实际上都是非线性的。

2.用最优化方法计算非线性相关系数

以前考虑非线性因子时，只考虑 x^2、$\sqrt{x}$、$1/x$等几种特殊情况，现在要考虑一般情况x^a($-\infty<a<+\infty$)，这时计算 y与x^a的非线性相关系数公式为

$$R'=\frac{\sum_{i=1}^{N}(x_i^a-\bar{x})(y_i-\bar{y})}{\sqrt{\sum_{i=1}^{N}(x_i^a-\bar{x})^2\sum_{i=1}^{N}(y_i-\bar{y})^2}} \tag{2}$$

其中 $\bar{x}=\frac{1}{N}\sum_{i=1}^{N}x_i^a$，$\bar{y}=\frac{1}{N}\sum_{i=1}^{N}y_i$。现在问

题是如何选定a，使$|R'|$为最大。因为（2）式为非线性的，同时只有一个变量a，因此这是一个非线性规划单变量函数的寻优问题。即要用最优化方法求出a，使目标函数

$$f(a)=-\left|\frac{\sum_{i=1}^{N}(x_i^a-\bar{x})(y_i-\bar{y})}{\sqrt{\sum_{i=1}^{N}(x_i^a-\bar{x})^2\sum_{i=1}^{N}(y_i-\bar{y})^2}}\right|$$

$=\min$。

a的求法有两种，一种是黄金分割法，另一种是外推内插法。这两种方法各有优缺点，黄金分割法的计算次数较少，计算过程比较简单，但要给出a的存在区间。外推内插法是先用外推法求出最优点存在的区间，然后用内插法求出二次多项式的极值点，这个方法虽然迭代时间可能多一点，但只要给出初始点和初始步长，便可一次找到最优点及最小的目标函数值。因此我们选用后一种方法，具体方法参见文献〔1〕。求出a后代入（2）式便得R'，这样求出的R'便是Y与X^a的非线性相关系数，再经过显著性检验，如$R'>R_\alpha$，则可选X^a为预报因子。

3．Y与X^a的非线性相关系数的几何意义

我们首先看一下Y与X相关系数的几何意义，如图1所示，R＞0为正相关（图1a），R＜0为负相关（图1b）。但由于变量

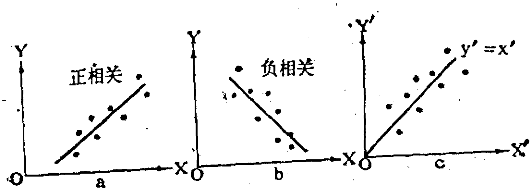

图1　线性相关系数几何意义

X或Y加（减）一个数或乘（除）一个数后相关系数仍不变，所以经过坐标变换后，不论正相关还是负相关，Y与X相关系数的大小反映了变换后点集$\{(x_i',y_i'),i=1,2,\cdots\cdots,N\}$在$y'=x'$的直线附近的情况（图1c）。相关系数大就说明所有这些点离直线$y'=x'$比较近，否则就比较远。同样可以看出，Y与X^a的非线性相关系数就反映了变换后点集$\{(x_i',y_i'),i=1,2,\cdots\cdots,N\}$在曲线$y'=x'^a$附近的情况，如果非线性相关系数大，就反映了这些点离曲线$y'=x'^a$比较近，否则就比较远。根据a的不同取值，Y与X^a的各种非线性相关情况如图2所示。

4．计算实例

现在给出预报量Y及12个预报因子X_i（i=1,2,……，12）的22次观测数据（表1）。用外推内插法求出Y与各个X的最优点

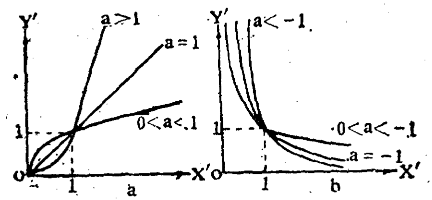

图2　非线性相关系数几何意义

a. $Y'=X'^a(a>0)$，　b. $Y'=X'^a(a<0)$

a及其最大的目标函数值（非线性相关系数），现把计算结果和原来线性相关系数列表比较（表2）。

从表2可以看出，Y与X^a(a为最优点)的非线性相关系数都比原来Y与X（a=1）的线性相关系数要高，大约平均提高20%左右。提高最多的有：X_9由0.1857上升到0.4078，X_5由0.4914上升到0.6505，X_8由0.4858上升到0.6341。同时原来通过显著性检验（$R_{0.01}=0.537$）只有X_1，现在X_1、X_2、X_3、X_5、X_8、X_{10}均通过$R_{0.01}$的显著性检验。这样选出来的因子为非线性因子，它将为建立非线性预报方程打下基础。

三、用最优化方法建立预报方程

1．建立多元线性回归方程

多元线性回归的数学模型为

$$Y=\beta_0+\beta_1X_1+\beta_2X_2+\cdots\cdots+\beta_KX_K+\varepsilon \tag{3}$$

式中Y为预报量，$X_i(i=1,2,\cdots\cdots,K)$为预报因子，$\varepsilon$为误差，$\beta_0$、$\beta_i$（i=1,2,……,K）为回归系数。要得到的估计值$\hat\beta_0$、$\hat\beta_i$，就是要使残差平方和最小，即

$Q=\sum_{i=1}^{N}(y_i-\hat{y}_i)^2=\min$，其中$\hat{y}_i=\hat\beta_0$

表 1　预报量和预报因子的数据

N	Y	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	X_{11}	X_{12}
1	-80.04	82.3	47.7	15.8	30.5	3.1	22.0	22.0	12.3	76.3	77.4	49.3	23.6
2	27.91	73.2	3.9	18.2	16.7	5.0	22.0	24.3	14.5	71.4	81.7	67.8	47.3
3	-13.03	83.3	209.8	19.2	8.1	4.2	20.4	22.4	14.5	115.9	82.6	99.9	35.4
4	45.36	67.2	30.8	15.5	123.6	3.7	22.7	21.6	16.8	98.1	90.7	96.1	75.1
5	56.00	39.8	40.3	19.4	138.5	4.2	22.3	23.6	15.4	114.6	64.5	103.4	20.2
6	122.99	71.7	4.5	19.6	45.2	4.9	20.8	24.2	15.0	112.0	54.0	77.9	74.1
7	52.94	64.8	30.2	18.1	28.0	8.4	19.2	21.0	12.6	109.8	110.7	79.0	66.9
8	-242.23	101.8	171.4	14.4	2.8	2.9	19.7	20.1	12.8	101.7	79.6	64.1	131.9
9	-9.16	92.2	7.8	17.7	1.0	4.6	21.6	23.8	13.9	111.6	70.1	94.6	30.5
10	47.92	69.0	19.9	18.2	39.2	4.5	22.0	22.0	14.5	106.3	61.8	59.5	31.0
11	-187.00	80.9	29.7	12.4	22.5	3.9	18.8	19.9	11.7	60.8	101.4	60.5	139.6
12	76.07	67.9	5.6	15.7	38.7	5.5	21.0	20.9	14.6	81.0	66.0	71.8	53.2
13	90.15	52.3	2.0	17.9	34.4	6.7	21.3	20.8	16.7	102.5	80.6	77.7	52.8
14	167.23	74.2	10.3	16.1	265.4	7.8	23.9	21.8	17.3	89.0	77.8	92.6	50.8
15	-50.79	89.3	22.7	15.1	3.6	4.2	19.1	19.2	15.1	81.0	81.2	84.7	63.7
16	15.09	66.3	11.2	16.4	4.9	5.1	20.5	23.3	15.6	104.2	77.1	104.6	132.6
17	-68.89	78.6	70.0	16.3	46.3	5.8	22.5	21.9	18.5	80.3	87.9	99.8	74.5
18	-88.20	90.0	88.8	15.5	60.2	5.6	21.4	21.0	12.6	98.0	100.7	84.4	38.5
19	-38.81	76.5	10.5	15.8	4.2	4.4	23.4	21.9	13.0	100.9	74.7	97.4	163.6
20	0.42	90.8	3.7	15.1	3.7	9.2	23.0	22.1	16.6	68.7	75.1	63.5	93.6
21	10.32	88.2	60.2	15.3	25.8	8.3	23.1	21.3	14.9	73.7	96.6	92.3	35.7
22	174.28	72.9	1.2	15.7	39.1	6.8	22.7	23.0	14.5	76.8	54.3	82.3	17.4

表2　线性相关系数(R)与非线性相关系数(R′)比较表

因子	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	X_{11}	X_{12}
R	-0.5914	-0.5350	0.5330	0.4573	0.4914	0.4503	0.4769	0.4858	0.1857	-0.4569	0.2695	-0.4845
R′	-0.6840	-0.6459	-0.5972	0.4917	-0.6505	-0.4762	-0.5077	-0.6341	-0.4078	0.5398	-0.3457	-0.4974
a	5.7780	0.0165	-4.2043	0.3720	-3.2631	-12.9015	-7.1740	-9.2969	-11.7489	-4.0314	-3.3689	1.8460

$+\hat{\beta}_1 x_{i1}+\hat{\beta}_2 x_{i2}\cdots+\hat{\beta}_K x_{iK}$，这是一个非线性规则寻优问题，其求解方法有两种：

（1）间接寻优方法(也称解析法)　这种方法就是利用求导数寻求函数极值的方法，在这里就是对残差平方和（Q）求偏导数并令其为零，最后解正规方程组得到回归系数估计值，这也就是我们以前通常用的最小二乘法。

（2）直接寻优方法(也称搜索法)　这是一种数值方法，利用函数在某一局部区域的性质或一些已知点的数值，来确定下一步计算的点，这样一步步搜索逼近，最后达到最优点。在这里就是要建立多变量的目标函数

$$f(\hat{\beta}_0,\hat{\beta}_1,\cdots,\hat{\beta}_K)=\sum_{i=1}^{N}(y_i-\hat{\beta}_0-\hat{\beta}_1 x_{i1}-\cdots-\hat{\beta}_K x_{iK})^2=\min,$$

然后给出一个初值，用多变量函数的寻优方法迭代求出回归系数的估计值。

关于多变量函数的寻优问题，也有多种方法，但为了便于在计算机上实现，我们选用了单纯形加速法。现在举一个例子，数据

仍用表 1 中预报量Y及X_1、X_5、X_{10}、X_{12}等 4 个预报因子（这是逐步回归精选出的因子）。用上面两种方法建立线性预报方法，从计算结果看出，两者基本一样。因为前一种方法计算时间较短，所以建立线性回归方程时仍用最小二乘法来求回归系数的估计值。

2. 建立多元非线性回归方程

多元非线性回归数学模型为

$$Y=\beta_0+\beta_1X_1^{a_1}+\beta_2X_2^{a_2}+\cdots\cdots+\beta_KX_K^{a_k}+\varepsilon \quad (4)$$

式中β_0、β_i $(i=1,2,\cdots\cdots,K)$为回归系数，a_i $(i=1,2,\cdots\cdots,K)$为预报因子X_i的幂。要求出这些参数的估计值也有两种方法：

（1）间接法（两步法） 首先用上述寻找非线性预报因子的方法求出预报量Y与因子非线性相关系数最大时的a_i。然后，用$X_i^{a_i}$作为预报因子，建立多元非线性回归方程。

$$\hat{y}=\hat{\beta}_0+\hat{\beta}_1x_1^{a_1}+\hat{\beta}_2x_2^{a_2}+\cdots\hat{\beta}_Kx_K^{a_k} \quad (5)$$

（2）直接法（一步法） 就是一步直接建立非线性回归方程，基本上同建立线性回归方程的直接法，不同的主要是这时的目标函数

$$f(\hat{\beta}_0,\hat{\beta}_1,\cdots,\hat{\beta}_K,a_1,\cdots,a_K)=\sum_{i=1}^{N}(y_i-\hat{\beta}_0-\hat{\beta}_1x_{i1}^{a_1}\cdots-\hat{\beta}_Kx_{iK}^{a_k})=\min,$$

给出适当的初值，用单纯形加速法便可求出这些参数。

现在还是以建立线性回归方程的预报量和预报因子的数据为例，在上面用最优化方法寻找预报因子中，已求出非线性因子分别为$X_1^{5.778}$、$X_5^{-3.2631}$、$X_{10}^{-4.0314}$、$X_{12}^{1.846}$，用上面两种办法建立非线性预报方程，所得结果基本上也是一样的。从上机时间来看，还是采取上机时间较短的第一种方法。

3. 多元非线性回归方程与多元线性回归方程比较

由于前面建立预报方程的预报量和预报因子都相同，而建立多元线性回归的两种方法结果基本一样，建立多元非线性回归的两种方法结果也完全相同，因此，这里只作多元非线性回归与多元线性回归两种方法的比较。

（1）预报方程比较

线性回归预报方程

$$\hat{y}=341.1393-3.0682X_1+28.7802X_5-2.7924X_{10}-0.5544X_{12};$$

非线性回归预报方程

$$\hat{y}=63.108-39.7479\times10^{-11}X_1^{5.778}-47.1841\times10^{7}X_5^{-3.2631}+13.3036\times10^{8}X_{10}^{-4.0314}-760.8142\times10^{5}X_{12}^{1.846}$$

（2）两个方程检验结果比较

从表 3 可以看出，无论统计检验、实况检验和试报检验，用最优化方法找出的非线性因子建立的非线性回归方程效果都比原来的逐步回归效果要好。

表3 两个方程检验结果比较

方法	Q	R	F	S	E	试报	误差
线性	44189.31	0.8914	16.4440	50.9840	36.2915	68.1315	16.8685
非线性	42481.28	0.8959	17.2760	49.9890	30.8504	74.3534	10.6466

注：Q为残差平方和，R为复相关系数，F为F检验值，S为标准差，E为平均绝对误差（下同）。

四、最优化方法在长期预报中应用试验

我们选取预报量Y为全省 3—5 月降水距平百分率，待选因子为前一年 11—12 月北半球 500hPa 月平均场球谐系数C_n^m（振幅）、Q_n^m（位相），共 60 个因子做了两个试验：

①用一般逐步回归选出的 4 个因子X_4、X_{21}、X_{44}、X_{46}建立线性回归方程与建立非线性回归方程比较。

②用计算非线性相关系数方法选出 4 个

表 4　选出因子的线性相关系数与非线性相关系数

样本	项目＼因子	X_4	X_8	X_{12}	X_{21}	X_{38}	X_{44}	X_{46}
1952—1984 (33)	R_1	－0.2899	0.2049	－0.1736	0.4834	0.3257	0.3588	0.3435
	R_a	－0.2940	0.3989	0.4551	0.4933	0.4091	0.3686	0.3500
	a	1.4878	3.9688	－3.8872	0.6020	0.2203	0.3758	1.5288
1952—1985 (34)	R_1	－0.3063	0.2174	－0.1647	0.4951	0.3087	0.3151	0.3036
	R_a	－0.3089	0.3959	0.4375	0.5046	0.4032	0.3274	0.3181
	a	1.3888	3.8750	－4.1631	0.6068	0.2122	0.2544	1.8221

注：R_1为 Y 与 X 的线性相关系数，R_a为 Y 与 X^a 的非线性相关系数。
X_4：前一年11月份北半球500hPa 高度平均场振幅 C^2_4
X_8：前一年11月份北半球500hPa 高度平均场振幅 C^3_7
X_{12}：前一年11月份北半球500hPa 高度平均场振幅 C^4_{10}
X_{21}：前一年12月份北半球500hPa 高度平均场振幅 C^2_8
X_{38}：前一年11月份北半球500hPa 高度平均场位相 Q^3_7
X_{44}：前一年11月份北半球500hPa 高度平均场位相 Q^5_9
X_{46}：前一年12月份北半球500hPa 高度平均场位相 Q^1_3

因子 X_8、X_{12}、X_{21}、X_{38} 建立非线性回归方程与①比较。

每次试验做两次：①资料为1952—1984年（33个样本），试报1985年；②资料为1952—1985年（34个样本），试报1986年。

首先选出因子的线性相关系数与非线性相关系数比较列于表 4。

从表 4 看出，非线性相关系数都比线性相关系数高，而且用线性方法漏选的因子（如 X_{12}），用非线性方法可选出来，且非线性相关系数较高。两次试验求出的线性与非线性回归方程从略，现只把对方程的检验结果比较列于表 5。

表 5　两次试验检验结果比较

样本	方法	因　子	Q	R	F	S	E	试　报	误　差
1952—1984 (33)	线　性	4,21,44,46	8987.283	0.7543	9.2397	17.9158	12.8087	3.6099	16.3901
	非线性	4,21,44,46	8357.080	0.7741	10.4643	17.2762	11.7839	5.7191	14.2809
	非线性	8,12,21,38	7189.930	0.8097	13.3304	16.0122	12.4413	7.2673	12.7327
1952—1985 (34)	线　性	4,21,44,46	9189.836	0.7524	9.4589	17.8014	12.8430	35.0869	33.0869
	非线性	4,21,44,46	8550.265	0.7722	10.7087	17.1708	11.8697	29.5621	27.5621
	非线性	8,12,21,38	7315.698	0.8091	13.7393	15.8829	12.4303	17.9168	15.9168

注：1985年实况为20，1986年实况为2。

五、结束语

1.用最优化方法解决统计天气预报中的一些非线性问题，这是一种尝试。经过试验证明是成功的，用这种方法可以找到非线性相关很好的预报因子，用这种因子建立的非线性回归方程都比逐步回归选出来的因子建立的线性回归方程要好。因此，这为寻找最优预报因子和建立最优预报方程开辟了一条新的途径，而且更接近于非线性的实际情况。

2.这种最优化方法还可以推广到预报量是离散型的情形，这时只要将目标函数改变一下，便可用于寻找离散型预报量的非线性因子以及建立非线性的判别方程。同时这种方法不论长、中、短期预报都可用。

3.关于解的唯一性问题，如果目标函数

是凸函数，那它的最优解就是唯一的，但由于目标函数是否为凸函数不容易验证，因此，一般我们求出的为局部最优解。在实际工作中还发现，初值不一样或者步长不一样，它的解有些也不一样，所以可以用多个初值或多种步长进行迭代，求出多个局部的最优解再选出其中最好的。

4.本文所用的最优化方法在文献〔1〕中有详细的叙述，FORTRAN 源程序在书中也有，但要改变目标函数，本文所用的程序都在 IBM 微机上通过。

参考文献

〔1〕范鸣玉、张莹等编著，最优化技术基础，清华大学出版社，1982年11月。

〔2〕朱伯承，统计天气预报，上海科学技术出版社，1981年4月。

An application of the optimizing method to the weather forecasting

Feng Yaohuang Yang Xu

(Institute of Meteorological science, Liaoning Province)

Abstract

In this paper, a method to select some nonlinear predictors and to build a nonlinear forecasting equation was preposed based on optimization principles. The experiments show both the nonlinear predictors and corresponding equation based on this method are better than those linear equations. Finally, the authors indicate that this method can be extended to building of nonlinear discriminatory equation when predictands are discrete. This method is applicable to various forecastings of long-, medium- and short-ranges.