

组队号#1965

2008 年 第一届“数学中国杯”

数学建模网络挑战赛

承 诺 书

我们仔细阅读了首届“数学中国杯”数学建模网络挑战赛的竞赛规则。

我们完全明白，在竞赛开始后参赛队员不能以任何方式（包括电话、电子邮件、网上咨询等）与队外的任何人（包括指导教师）研究、讨论与赛题有关的问题。

我们知道，抄袭别人的成果是违反竞赛规则的，如果引用别人的成果或其他公开的资料（包括网上查到的资料），必须按照规定的参考文献的表述方式在正文引用处和参考文献中明确列出。

我们郑重承诺，严格遵守竞赛规则，以保证竞赛的公正、公平性。如有违反竞赛规则的行为，我们将受到严肃处理。

我们允许数学中国网站(www.madio.net)公布论文，以供网友之间学习交流，数学中国网站以非商业目的的论文交流不需要提前取得我们的同意。

我们的参赛报名号为：

参赛队员（签名）：

队员 1：张健

队员 2：卜翔

队员 3：周玉香

参赛队教练员（签名）：

参赛队伍组别：大学组

组队号#1965

2008 年 第一届“数学中国杯”

数学建模网络挑战赛

编 号 专 用 页

参赛队伍的参赛号码：（请各个参赛队提前填写好）：1965

竞赛统一编号（由竞赛组委会送至评委团前编号）：

竞赛评阅编号（由竞赛评委团评阅前进行编号）：

组队号#1965

2008 年 第一届“数学中国杯” 数学建模网络挑战赛

题 目 沙尘暴与各项气象指标的关系分析关 键 词 多元回归分析 主成分分析 人工神经网络

摘 要：

问题一，本文作者采用统计回归模型，按一定方法从大量数据中选取适当数据，分析了沙尘暴频率与各自然因素的关系，同时提出了人工神经网络模型解决该问题的改进。

模型 I 多元线性回归模型
模型 II 多元非线性回归模型
改进 模型 III 神经网络模型

问题二，作者通过矩阵运算，进行数据标准化，消除了指标间的差异，并结合数量分类学知识，获得各指标的贡献率，以此得出一种评估沙尘天气气象等级的方法。

主要成分分析方法的应用

参赛队号 1965所选题目 C

参赛密码 _____
(由组委会填写)

组队号#1965

一、问题的重述

沙尘暴是一种灾害性天气，近年来频繁发生，对环境和人们的生产、生活都造成了影响甚至危害。对沙尘暴的发生条件需要加以研究，以便人们进行有效的预测和预防。请为沙尘暴的发生建立数学模型，具体问题如下：

问题一：对沙尘暴的发生次数进行分析，建立相应的数学模型，使人们可以根据相关气象数据的监测结果，估计出沙尘暴在某地区某段时间内发生的次数。

问题二：评估沙尘天气气象等级，预测在近期内观测站所在地发生沙尘暴的风险，并在高风险时发出警报信息。

二、问题的分析

沙尘暴这种灾害性天气，给国家和人民带来了巨大的损失，对沙尘暴的发生条件加以研究，就可以使人们能够进行有效的预防。极大的减少人们生活中的不便，财产上的损失，从这个目的出发，我们要尽量使过去的气象观测数据得到充分利用，并且借助数学方法建立相应的数学模型，以期能够做好合理的预测。

1. 数据分析：

(1) 合理利用已有数据。沙尘暴是由多方面的因素造成的，并且发生的频率不仅仅与气象因素有关，还与当地的地理环境密切相关（例如，土壤荒漠化也是引起沙尘暴的主要因素之一），为了更加准确的研究气象因素的影响，我们在选取数据时更多的考虑同一站点，同一季节数据。

(2) 充分利用数据。依据统计学的原理，数据越多，样本容量越大，越有利于提高分析预测的灵敏度。鉴于此我们选取了同一地区，地质条件基本相同的观测站作为数据来源，极大的扩大了数据的采集量，并且也具有较好的可靠性。

(3) 数据取舍。要得到有用的数学模型，对下一段时间内沙尘暴发生情况预测的合理性尤为重要，但是影响沙尘暴的气象因素有很多，根据主要分析的数学原理进行判断，求得各种气象因素对沙尘暴产生的影响贡献率将有助于我们分清主要方面，正确的数据取舍是较好预测的重要保证和前提。

2. 问题解决思路分析：

(1) 对问题 1 的分析。问题一要求根据相关气象数据的监测结果，估计出沙尘暴在某地区某段时间内发生的次数。这是一个较典型的统计回归问题，建立数学模型，根据已有的几种气象数据，合理选取，保证样本容量足够大，用数学软件对模型求解。

(2) 对问题 2 的分析。问题二要求预测在近期内观测站所在地发生沙尘暴的风险，由于要求近期的预测，而所给数据的最小时间周期是一个月，统计模型不能很好的满足近期实时性的要求，基于此我们用了大量的数据模拟，用指标离差最大化决策方法，找出各种气象因素对沙尘暴的影响强弱，从而力求更好的解决这一问题。

组队号#1965

三、模型的假设

- (1) 处在同一地区的观测站的地形地貌等外部条件基本相同。
- (2) 模型计算时只考虑给出的几种气象指标对沙尘暴的影响。
- (3) 每个观测站除给定的气象指标有变化外，其它环境因素在每年同一时间段内基本保持不变。

四、符号约定

x_1	月平均风速
x_2	大风日数
x_3	月平均气温
x_4	月降水量
x_5	月相对湿度
y	沙尘暴天数

五、模型的建立与求解

问题一的建立与求解

模型一：多元线性回归模型

为简便起见，我们先假定 x_1 、 x_2 、 x_3 、 x_4 、 x_5 等气象因素对沙尘暴频次的影响是线性的，并且各因素都是独立的，没有交互作用。除上述确定因素影响沙尘暴频次外，不确定因素也对沙尘暴频次有一定影响，对此我们引入不确定因数 ε 。这样可初步得到沙尘暴频次与各因素之间的多元线性回归模型方程：

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_5 + \varepsilon$$

下面以 52323 观测站 61-81 年为例。

直接用 MATLAB 统计工具箱的命令 regress 求解

组队号#1965

参数	参数估计值	参数置信区间
β_0	0	[0 0]
β_1	1.0246	[-1.0285 3.0777]
β_2	-3.5624	[-9.9731 2.8483]
β_3	0.6651	[-1.0761 2.4063]
β_4	0.0772	[-0.3382 0.4926]
β_5	0.1988	[-0.1984 0.5960]
$R^2=0.2518$ $F=1.2623$ $p=0.3278$		

显然，模型一与实际情况拟合程度很不理想，线性模型不够合理，无法如实地反映出沙尘暴频次随七个自然因素变化的真正趋势。

模型二：多元非线性回归模型

首先得到 y 分别与 x_1 、 x_2 、 x_3 、 x_4 、 x_5 的关系式，根据这些关系式得出以 x_1 、 x_2 、 x_3 、 x_4 、 x_5 为回归变量(自变量)的函数。

注意：为了得到 y 分别与 x_1 、 x_2 、 x_3 、 x_4 、 x_5 的关系而不受其他因素的影响，我们采用了这样一种方法：对于适当长度的区间内的变量如 x_1 ，对所有对应的 y 值求数学期望值。由于其他变量此时可看作随机，这样做可使它们对 y 正面或负面的影响抵消，只有 x_1 与 y 有函数关系。

下面以宁夏 81-05 年为例，来求解模型。

1. 沙尘暴天数 y 与月平均风速 x_1 关系

有数据可知，月平均风速 x_1 大于 4 的概率较小，随机性较大，所以取 $1 \leq x_1 \leq 4$ 。

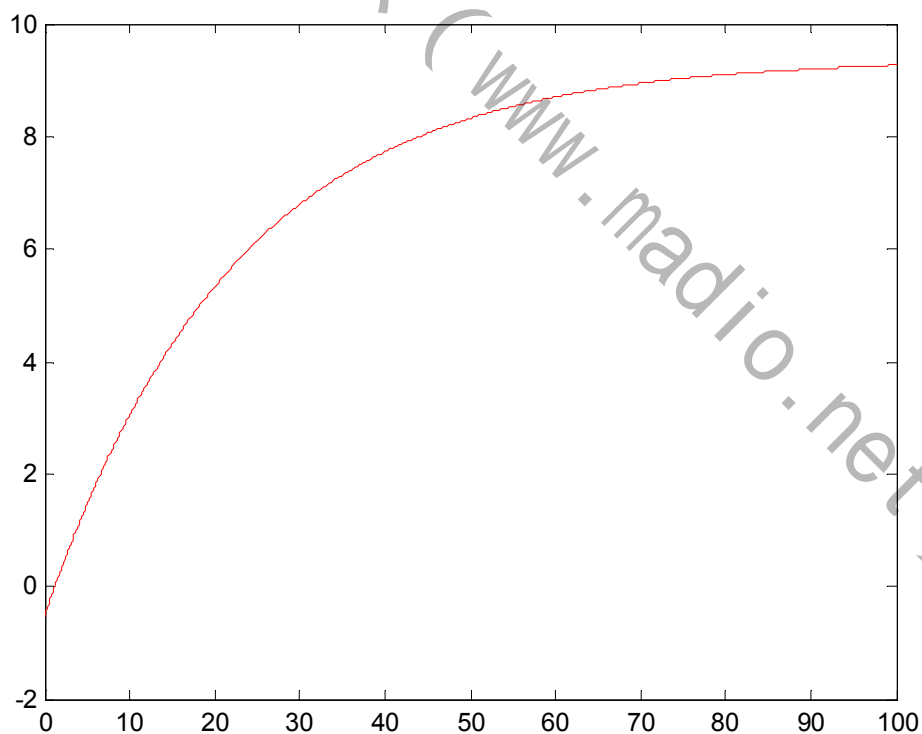
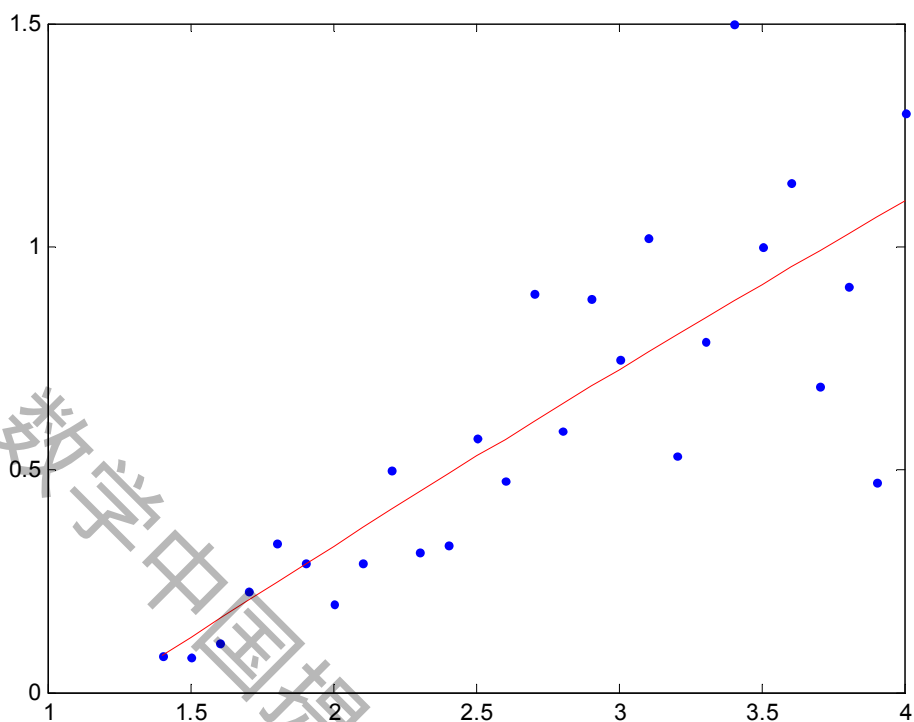
用 MATLAB 作散点图，从图中分析出随着 x_1 增加， y 的值增长趋于平缓，用指数曲线拟合可得到指数函数模型

$$y = -9.8907e^{-0.0447x_1} + 9.3719 \quad (1)$$

由 (1) 式得： x_1 趋于无穷时， y 无限趋近 9.3719，即最大月沙尘暴天数为一定值而非无穷，且每个月的天数是固定的，也不可能是无穷，所以这是合理的。

令 $y=0$ ，得 $x=1.2053$ 。这是沙尘暴发生的月平均风速的阈值，达到该阈值才能发生沙尘暴现象。

组队号#1965

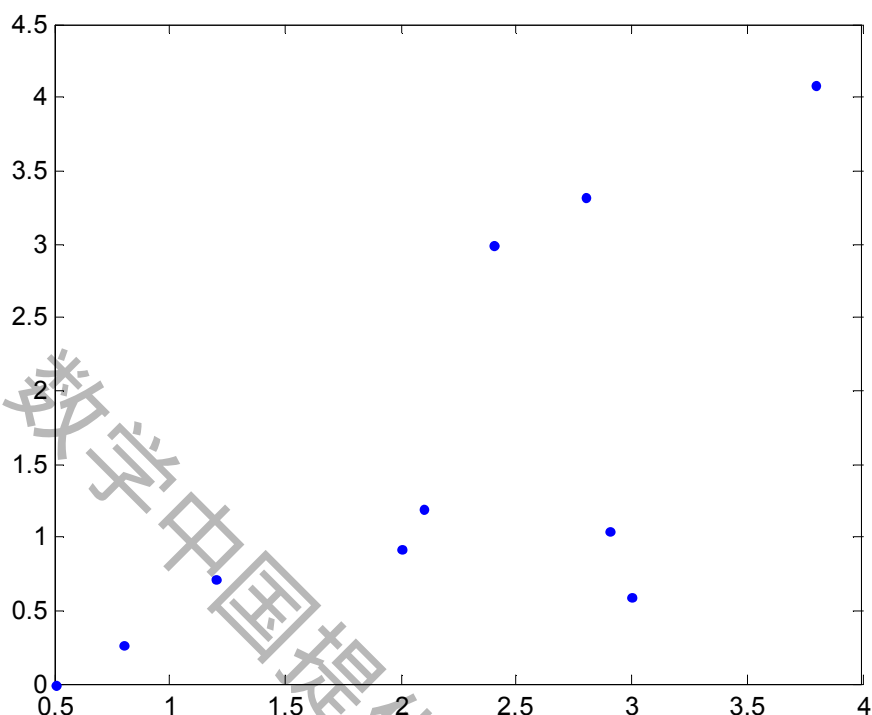


2. 沙尘暴天数 y 与大风天数 x_2 关系

大风天数有明显的地域性与时域性，某一地区多年同一月份的大风天数几乎不变。大风天气是形成大范围沙尘暴的动力因素。但是，沙尘暴发生频次与大风天数的分布在空间上并不完全一致，表明沙尘暴的发生频次除了受到区域大风日数的影响外，同样受

组队号#1965

到表征区域特征的降水量、温度、地形、地貌、植被覆盖等条件的影响。



我们建立简单的线性模型

$$y = \beta_0 + \beta_1 x_1 + \varepsilon \quad (2)$$

3. 沙尘暴天数 y 与月平均气温 x_3 关系

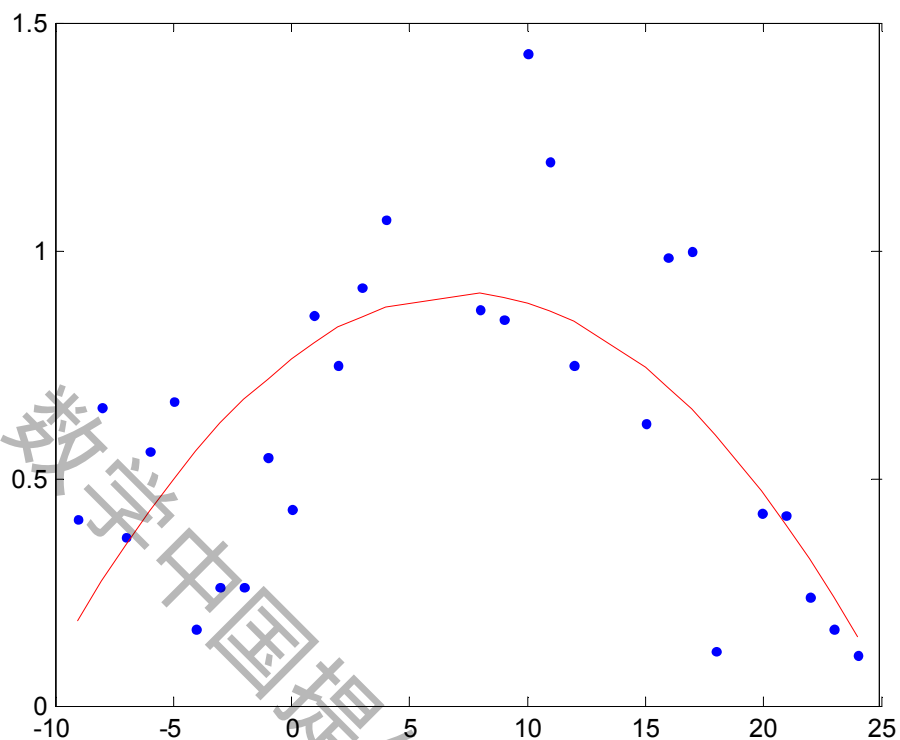
由图建立二次模型，MATLAB 拟合可得

$$y = -0.0027x_3^2 + 0.0394x_3 + 0.7611 \quad (3)$$

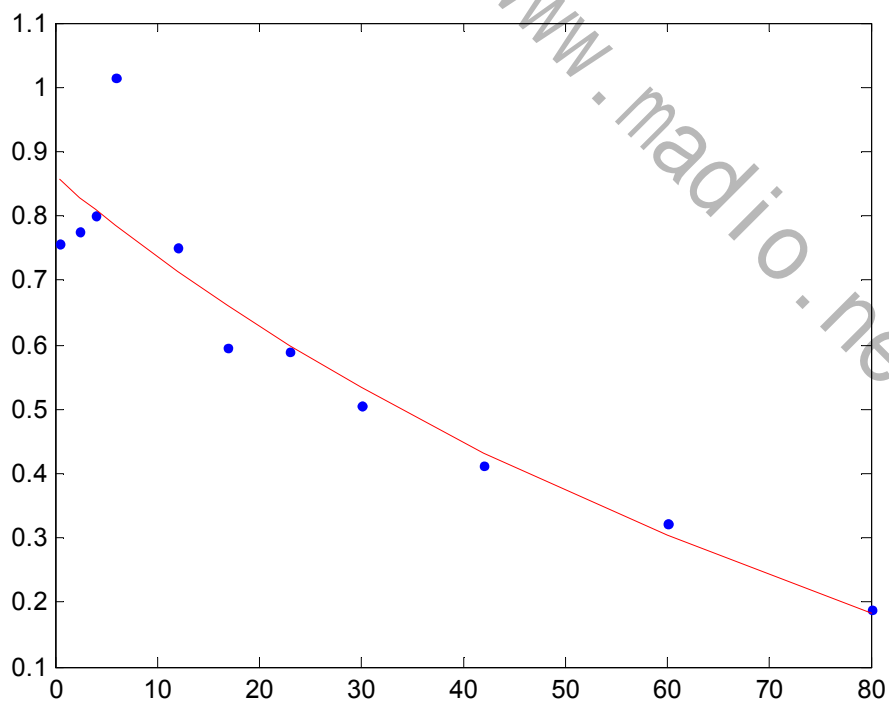
随着温度的升高，沙尘暴天数先增加，后减少。 $x_3=7.3$ 时，沙尘暴天数最多。

月平均气温对沙尘暴天数的二次关系并不是直接作用的，而是间接的。夏季增温加快了地表的蒸发，土壤水分减少，到了冬春季风力达到一定条件，易形成沙尘暴天气。春季的3月份起，气温开始明显回升，而且回升的速度快。温度快速回升易导致下层大气强烈受热，造成空气不稳定，这将促使对流作用及湍流作用的加强，有利于风速的加大和上升气流的产生，为风沙活动提供有利条件。

组队号#1965



4. 沙尘暴天数 y 与月降水量 x_4 关系



由图建立双曲线模型，MATLAB 拟合得

组队号#1965

$$y = \frac{268.5959}{x_4 + 142.6114} - 1.0231 \quad (4)$$

由式(4)得, x_4 趋于无穷时, y 趋向与 -1.0231 近似等于 0。 $x_5=0$ 时, $y=0.86$ 。

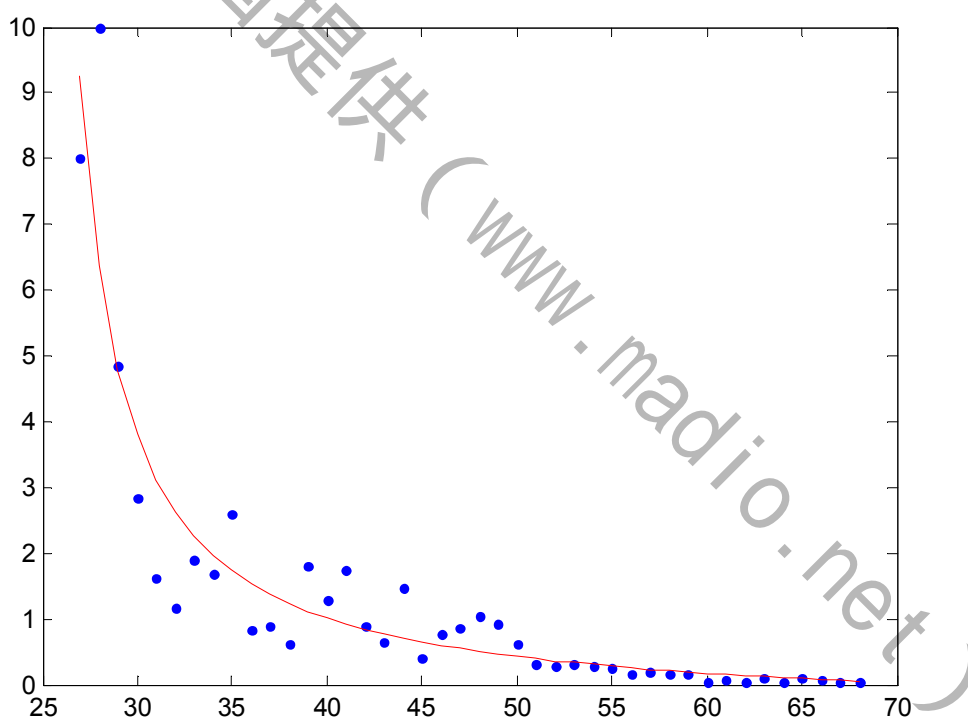
当降水量充沛时,有利于植被生长,地表植被覆盖度相对较高,对沙尘暴发生的抑止作用强,沙尘暴发生的次数和强度相应会降低。相反,当持续干旱时,地表植被覆盖度将大大减少,裸露土地面积增加,加之气温的影响,土壤含水量也将减小,土质疏松,沙尘暴发生的频次将提高。

5 沙尘暴天数 y 与月相对湿度 x_5 关系

由图建立双曲线模型, MATLAB 拟合得

$$y = \frac{22.8305}{x_5 - 24.6549} - 0.4703 \quad (5)$$

月相对湿度 x_5 有月降水量 x_4 密切联系, 所以得出的关系式也近似。



6. 综合上面的分析, 建立如下的回归模型

$$y = \beta_0 + \beta_1 e^{-0.0447x_1} + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_3^2 + \frac{\beta_5}{x_4 + 142.6114} + \frac{\beta_6}{x_5 - 24.6549} + \varepsilon \quad (6)$$

利用 MATLAB 统计工具箱求解, x 取每四年的期望值, y 为四年沙尘暴总天数, 得

组队号#1965

参数	参数估计值	参数置信区间
β_0	522	[-226 1270]
β_1	404.5	[083 726]
β_2	13.7	[3 24]
β_3	-231.4	[-350 -113]
β_4	11.8	[6 18]
β_5	7377	[-32806 47560]
β_6	5580.1	[3874 7286]
$R^2=0.9105$ $F=28.8327$ $p=0.0000$		

结果显示， $R^2=0.9105$ 指因变量的 91.05%可由模型确定， F 值远远超过 F 检验的临界值， p 远小于 α ，因而模型二从整体上来看是可用的。

问题二的建立与求解

我们的目的是求得不同气象因子对沙尘暴产生作用的大小，从上述 n 个指标中寻找影响沙尘暴的主要方面。如果起主要作用的刚好就在某几个指标中，那么只需将这几个指标提出来进行分析，主要影响沙尘暴的指标就找到了。所以我们下面要做的是找出主要方面，分三个步骤：

1. 取样（足够的样本容量为准）

先从所给的数据中抽取 t 次记录数值，这些数据可以组成一个 $t \times n$ 的矩阵（ t 表示数值个数， n 表示气象因子个数，），得到原始数值矩阵为

$$\begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ x_{t1} & x_{t2} & \cdots & x_{tn} \end{bmatrix}$$

2. 数据的标准化：（消除各气象指标单位不同的影响）

在原始数据中，每个指标的数据来源不同，单位的选取不同，因而数值的大小及变化幅度不同。标准中的这种差异，不仅毫无意义，还会干扰分析。为了取得正确的分析结果，有必要对原始数据进行变换，以消除这种没有意义的差别。

先求每个指标的平均值和离差平均和
平均值：

组队号#1965

$$\bar{x}_j = \frac{1}{t} \sum_{k=1}^t x_{kj}, \quad j=1, 2, \dots, n;$$

离差平方和：

$$d_j = \sqrt{\sum_{k=1}^t (x_{kj} - \bar{x}_j)^2}, \quad j=1, 2, \dots, n;$$

再做变换

$$h_{ij} = \frac{x_{ij} - \bar{x}_j}{d_j}, \quad i=1, 2, \dots, t; \quad j=1, 2, \dots, n;$$

得变换后的原始数值矩阵为

$$\begin{bmatrix} h_{11} & h_{12} & \cdots & h_{1n} \\ h_{21} & h_{22} & \cdots & h_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ h_{n1} & h_{n2} & \cdots & h_{nn} \end{bmatrix}$$

3. 主成分分析（即找出各气象因子对沙尘暴产生的贡献率）

我们的目的是从上述 n 个指标中寻找影响沙尘暴的主要方面。如果起主要作用的刚好就再某几个指标中，那么只需将这几个指标提出来进行分析，主要矛盾就抓住了。为此，用一个向量来体现多维事物的主要方面，以指标向量的线性组合来体现多个指标的综合效果，得主成分向量

$$F_i = a_{1i}H_1 + a_{2i}H_2 + \cdots + a_{ni}H_n, \quad i=1, 2, \dots, n.$$

其中 $H_j = (h_{1j}, h_{2j}, \dots, h_{nj})^T$

如此构造的主要方面，考虑到它们各自所代表的实际意义时是相互独立的，所谓相互独立，用数学方法表示就是两个不同向量互相垂直，即 $F_i^T \bullet F_j = 0 (i \neq j)$ 。若把 F_j 写成矩阵与向量乘积的形式，则

$$F_i = (H_1, H_2, \dots, H_n)(a_{1i}, a_{2i}, \dots, a_{ni})^T = H(a_{1i}, a_{2i}, \dots, a_{ni})^T,$$

垂直条件可写成

组队号#1965

$$F_i^T \bullet F_j = (a_{1i}, a_{2i}, \dots, a_{ni}) \bullet H^T H (a_{1i}, a_{2i}, \dots, a_{ni})^T.$$

记 $R = H^T H$ ，可以证明， R 就是原始数据各指标的相关系数矩阵，记为

$$\begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ r_{21} & r_{22} & \cdots & r_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ r_{n1} & r_{n2} & \cdots & r_{nn} \end{bmatrix}$$

其中 $r_{11} = r_{22} = \cdots = r_{nn} = 1$ ，且 R 是一个对称矩阵。用 MATLAB 计算可得矩阵 R 的特征值，从而得各因素的贡献率及累计贡献率。

	特征值	贡献率	累计贡献率
月平均风速	0.3159	6.32%	6.32%
大风发生	0.4934	9.87%	16.91%
平均气温	0.8445	16.89%	33.80%
平均降水	1.5385	30.77%	64.57%
相对湿度	1.8077	36.15%	100%

表 1 宁夏地区-各因素特征值、贡献率及累计贡献率

	特征值	贡献率	累计贡献率
月平均风速	0.3226	6.54%	6.54%
大风发生	0.6223	12.45%	18.99%
平均气温	0.9717	19.43%	38.42%
平均降水	1.1258	22.52%	60.94%
相对湿度	1.9576	39.15%	100%

表 2 蒙古地区-各因素特征值、贡献率及累计贡献率

	特征值	贡献率	累计贡献率
月平均风速	0.2765	5.53%	5.53%
大风发生	0.3062	6.12%	11.65%
平均气温	0.8469	16.94%	28.59%
平均降水	1.3975	27.95%	56.54%
相对湿度	2.1729	43.46%	100.00%

组队号#1965

表 3 甘肃地区-各因素特征值、贡献率及累计贡献率

	特征值	贡献率	累计贡献率
月平均风速	0.2249	4.50%	4.50%
大风发生	0.5981	11.96%	16.46%
平均气温	1.1176	22.35%	38.81%
平均降水	1.2161	24.32%	63.13%
相对湿度	1.8432	36.86%	100%

表 4 青海地区-各因素特征值、贡献率及累计贡献率

主要成分分析方法小结：

由于第一主成分（相对湿度）的贡献率较高，所以降水在沙尘暴的成因中占有举足轻重的地位，在沙尘暴的预报与监测过程中，应密切注意预报与监控。

强风是起沙尘的动力，是沙尘暴形成的决定因素。第三大主成分（风）的贡献率显然低于第一主成分，也就是说它所代表的意义不及第一主成分重要，但是比率仍较高，亦有必要进行分析。

设函数 Z 为沙尘天气气象等级函数， Z 值越大表示沙尘天气出现的风险越高，反之亦然。设 h_1 为月平均风速转化后的与单位无关量， h_2 为月大风发生次数转化后的与单位无关量， h_3 为月平均气温转化后的与单位无关量， h_4 为月平均降水转化后的与单位无关量， h_5 为相对湿度转化后的与单位无关量， p_1 为月平均风速的贡献率， p_2 为月大风发生次数的贡献率， p_3 为月平均气温的贡献率， p_4 为月平均降水的贡献率， p_5 为相对湿度的贡献率。

$$Z = h_1 \cdot p_1 + h_2 \cdot p_2 + h_3 \cdot p_3 + h_4 \cdot p_4 + h_5 \cdot p_5$$

有统计学知识，选取宁夏 53614, 53615, 53705 所有观测站点得预测：

当 $Z > 0$ 时有发生沙尘暴的危险，人们要减少出门的次数

当 $Z > 10$ 时危害程度升级属于高风险期，应即使发出警报信息

六. 模型的评价与推广

模型优点

(1) 采用的数学模型有成熟的理论基础，可信度高。

组队号#1965

(2) 建立的数学模型都有相应的专用软件支持，算法简便，编程实现简单，推广容易。

(3) 利用数学工具，通过 matlab 编程的方法，严格地对模型求解，具有科学性。

(4) 在预测时，从三种不同角度分别进行分析预测，使预测结果更加准确，贴近实际。

模型缺点

(1) 在建模的过程中，不能考虑净全辐射，总辐射，直接辐射，月蒸发量等各个方面对预测的影响，只能忽略部分因素，得到大概的结果。

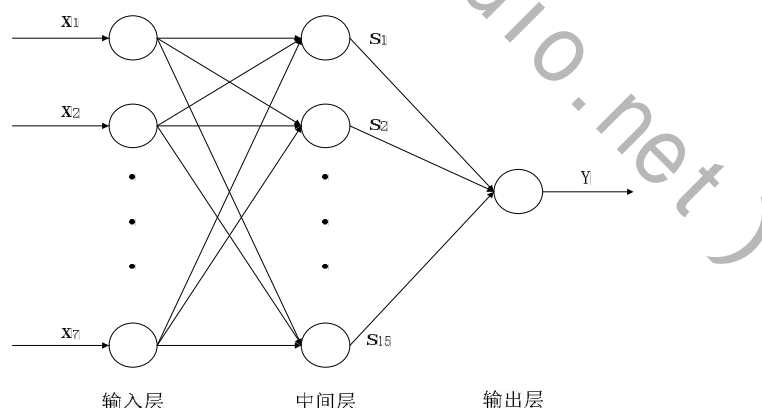
(2) 对问题 2 的解决没有更加深入

模型改进

(1) 经典的神经网络模型

沙尘暴与各种自然因素更接近于非线性关系，考虑到回归方法在处理非线性问题上繁琐性和精度差等缺陷，而人工神经网络模拟生物大脑的结构和功能而建立起来的人工智能识别方法，有大量神经元广泛互联而成的复杂网络结构，神经元之间的相互作用通过连接权值来体现。它建立起的数学模型能更精确的逼近非线性输入与输出之间的映射，消除了回归法处理非线性问题时的缺点。同时它具有很强的容错性、学习性、自适应性，特别适用于解决因果关系复杂的非确定性推理问题，我们可以引入人工神经网络模型。

采用误差反向传播方法（BP）算法建立模型。以自然指标的 6 个因素作为输入，构建 6 节点的输入层；以沙尘暴天数为输出，构建单节点的输出层；中间层选择 n 节点，搭建一个的三层神经网络，其中门限值选择为零。则网络连接形式如图一所示



(2) 此模型是建立在沙尘暴的发生只与个别气象指标的基础上，有较大的缺憾，根据常识，沙尘暴的发生与沙源（土壤沙化状况）有较大关系，但是在此模型下没有得到充分的体现。

组队号#1965

模型的推广

(1) 预测问题是一个普遍性的问题，尤其是经典的神经网络模型，更能得到广泛的应用，可以应用于各种类似的预测问题中。

七. 参考文献

[1]湖北省大学生数学建模竞赛专家组，数学建模（本科组），武汉，华中科技大学出版社，2006年2月。

[2]马国顺等，Meteorological factors analysis of the cause of the sandstorm, www.fpt.192.168.cn, 2008, 4, 13。

[3]姜启源 谢金星 叶俊，数学模型（第三版），北京，高等教育出版社，2003年8月。

附录：

%问题一：模型一

```
>> a=xlsread('a');
>> x(:,1)=ones(20,1)
>> for i=1:20
for j=1:5
x(i,j+1)=sum(a(12*i-11:12*i,j+1))/12;
end
y(i,1)=sum(a(12*i-11:12*i,6));
end
>> [b,bint,r,rint,stats]=regress(y,x(:,1:6))
b =
```

```
0
1.0246
-3.5624
0.6651
0.0772
0.1988
```

bint =

```
0 0
-1.0285 3.0777
-9.9731 2.8483
-1.0761 2.4063
-0.3382 0.4926
-0.1984 0.5960
```


组队号#1965

r =

1. 8488
0. 3888
-1. 1577
-0. 8879
0. 4929
-0. 0158
-0. 3540
-0. 2709
-1. 3253
-0. 9065
1. 2346
1. 3811
-1. 1523
0. 6871
1. 9190
-0. 0138
1. 9136
-1. 0626
-1. 4782
-1. 2408

rint =

-0. 6344	4. 3320
-2. 2289	3. 0064
-3. 5948	1. 2794
-3. 0710	1. 2951
-1. 9866	2. 9723
-2. 7867	2. 7551
-2. 7005	1. 9924
-2. 9723	2. 4305
-3. 6002	0. 9495
-3. 3997	1. 5868
-1. 0797	3. 5489
-1. 0474	3. 8095
-3. 5700	1. 2655
-1. 9433	3. 3174
-0. 6101	4. 4482
-2. 3999	2. 3723
-0. 4283	4. 2556

组队号#1965

```
-3.4251    1.2999
-3.3378    0.3815
-3.7739    1.2922
```

stats =

```
0.2518    1.2623    0.3278    1.7432
```

%问题一：模型二

%feng.m

```
function y=feng(b,x)
y=b(1)*exp(b(2)*x)+b(3);
```

```
>>c=xlsread('风速')
>>plot(c(:,1),c(:,2),'.')
>>[b,r,J]=nlinfit(c(:,1)',c(:,2)', 'feng', [-30,1,30])
```

b =

```
-300.2267   -0.0013   299.7717
```

r =

Columns 1 through 12

```
-0.0101   -0.0557   -0.0607    0.0149    0.0833   -0.0002   -0.1322
-0.0789    0.0935   -0.1331   -0.1561    0.0477
```

Columns 13 through 24

```
-0.0872    0.2922   -0.0564    0.2037    0.0266    0.2594   -0.2653
-0.0484    0.6247    0.0856    0.1895   -0.3049
```

Columns 25 through 27

```
-0.1224   -0.5999    0.1905
```

J =

组队号#1965

1.0e+003 *

0.0010	-0.4196	0.0010
0.0010	-0.4495	0.0010
0.0010	-0.4794	0.0010
0.0010	-0.5093	0.0010
0.0010	-0.5391	0.0010
0.0010	-0.5690	0.0010
0.0010	-0.5989	0.0010
0.0010	-0.6288	0.0010
0.0010	-0.6586	0.0010
0.0010	-0.6885	0.0010
0.0010	-0.7183	0.0010
0.0010	-0.7481	0.0010
0.0010	-0.7779	0.0010
0.0010	-0.8078	0.0010
0.0010	-0.8376	0.0010
0.0010	-0.8674	0.0010
0.0010	-0.8972	0.0010
0.0010	-0.9269	0.0010
0.0010	-0.9567	0.0010
0.0010	-0.9865	0.0010
0.0010	-1.0163	0.0010
0.0010	-1.0460	0.0010
0.0010	-1.0757	0.0010
0.0010	-1.1055	0.0010
0.0010	-1.1352	0.0010
0.0010	-1.1649	0.0010
0.0010	-1.1947	0.0010

```
>>[YY,delta]=nlpredci('feng',c(:,1)',b,r,J);
>>plot(c(:,1),c(:,2),'k+',c(:,1),YY,'r')
```

```
%shui.m
function y=shui(b,x)
y=b(1)./(x+b(2))+b(3);
```

```
>>c=xlsread('降水')
>>plot(c(:,1),c(:,2),'.')
>>[b,r,J]=nlinfit(c(:,1)',c(:,2)', 'shui',[1,10,0]);
>>[YY,delta]=nlpredci('feng',c(:,1)',b,r,J);
>>plot(c(:,1),c(:,2),'k+',c(:,1),YY,'r')
```

组队号#1965

```
>>a=xlsread('a');
>>x(:,1)=ones(24,1);
>>x(:,2)=exp(a(:,1).*(-0.0447));
>>x(:,3)=a(:,2);
>>x(:,4)=a(:,3);
>>x(:,5)=a(:,3).^2;
>>x(:,6)=(a(:,4)+142.6114).\1;
>>x(:,7)=(a(:,5)-24.6549).\1;
>>y=a(:,6);
>>[b,bint,r,rint,stats]=regress(y,x)
```

b =

1.0e+003 *

```
0.5220
0.4045
0.0137
-0.2314
0.0118
7.3770
5.5801
```

bint =

1.0e+004 *

```
-0.0226    0.1270
0.0083     0.0726
0.0003     0.0024
-0.0350   -0.0113
0.0006     0.0018
-3.2806    4.7560
0.3874     0.7286
```

r =

```
14.1592
-9.1754
2.8113
-3.7089
-4.3734
```

组队号#1965

-12.9801
7.4742
-16.9706
3.6994
2.8218
2.6243
-1.9029
5.1892
-9.6165
9.7091
5.6123
-6.9199
-2.2883
2.3346
-4.1170
8.4820
14.9366
-7.9212
0.1203

rint =

-2.0604 30.3788
-26.4154 8.0646
-14.4382 20.0608
-21.9134 14.4955
-22.3020 13.5552
-30.4989 4.5388
-1.6149 16.5633
-32.0047 -1.9366
-13.3716 20.7703
-14.4682 20.1117
-16.2271 21.4758
-19.3038 15.4980
-13.8744 24.2528
-28.8345 9.6014
-8.3678 27.7859
-11.8938 23.1185
-24.3311 10.4914
-20.0408 15.4642
-11.6660 16.3351
-16.5807 8.3466
-9.7306 26.6947

组队号#1965

-1.5628	31.4359
-24.9234	9.0810
-16.8879	17.1285

stats =

0.9105	28.8327	0.0000	91.0857
--------	---------	--------	---------

问题 2:

程序:

function f11(x)

A=x;

[m,n]=size(A);

B=sum(A)/m;

for i=1:n

C(:,i)=A(:,i)-B(i);

end

D=sqrt(sum(C.^2));

for i=1:n

E(:,i)=C(:,i)/D(i)

end

T=E.'*E;

[V,M]=eig(T)

测试结果:

xlsread('1.xls')

ans =

4.0000	3.8000	8.4000	12.3000	48.0000
3.5000	2.8000	14.8000	26.9000	34.0000
3.3000	2.4000	2.6000	0.5000	25.0000
2.6000	0.4000	-2.7000	0.1000	38.0000
4.6000	3.8000	9.8000	16.0000	43.0000
3.5000	3.8000	11.0000	13.5000	32.0000
2.7000	2.4000	0.4000	5.7000	41.0000
3.1000	1.1000	16.6000	26.4000	41.0000

>> f11(ans)

V =

0.1011	0.7346	0.0405	0.4467	0.4989
0.0409	-0.6579	0.2195	0.5846	0.4191
-0.7136	-0.0168	0.1878	-0.4187	0.5290
0.6812	-0.1093	0.0063	-0.5280	0.4952
-0.1220	-0.1239	-0.9565	0.0674	0.2245

M =

组队号#1965

0.0271	0	0	0	0
0	0.1820	0	0	0
0	0	0.9375	0	0
0	0	0	1.1994	0
0	0	0	0	2.6540

数学中国提供 (www.madio.net)