Geomatics Science and Technology 测绘科学技术, 2014, 2, 37-43
Published Online October 2014 in Hans. http://www.hanspub.org/journal/gst
http://dx.doi.org/10.12677/gst.2014.24007

Settlement Prediction Method for Turbine Foundation in Nuclear Power Station Based on Polynomial Regression Analysis

Yongwang Sun

Shenzhen Power Supply Planning Design Institute Co., Ltd., Shenzhen
Email: 530953818@qq.com

Received: Aug. 19th, 2014; revised: Sep. 15th, 2014; accepted: Sep. 23rd, 2014

Abstract

Based on five order polynomial regression, the prediction model was used to predict the subsidence of turbines plant. The results showed that the maximum relative error was 10% and the average relative error was 4.7%, indicating that this method was of good performance.

Keywords

Polynomial, Regression Analysis, Subsidence Prediction

基于多项式回归分析的核电站汽轮机基础沉降预测方法

孙永旺

深圳供电规划设计院有限公司，深圳
Email: 530953818@qq.com

收稿日期：2014年8月19日；修回日期：2014年9月15日；录用日期：2014年9月23日

摘　要

基于5阶多项式回归分析法建立汽机平台沉降预测模型，并应用该模型对岭澳二期核电厂汽轮机平台沉降监测数据进行了预测，结果表明，最大预测相对误差为10%，平均相对误差为4.7%，具有较好的实用性。

关键词

多项式，回归分析，沉降预测

1. 引言

汽机基础沉降预测是汽轮机厂房施工及运营期间的重要工作。在安装施工阶段，受施工活动影响，各时期荷载变化较大，由于各种因素影响，施工期间没有条件对各期荷载数据进行统计和分析，因此汽机平台沉降预测只能引入一个自变量，即监测时间。所进行的回归分析只能是一元回归分析。

在一元回归分析中，如果依变量 y 与自变量 x 的关系为非线性的，但是又找不到适当的函数曲线来拟合，则可以采用一元多项式回归。多项式回归的最大优点就是可以通过增加 x 的高次项对实测点进行逼近，直至满意为止。多项式回归可以处理相当一类非线性问题，它在回归分析中占有重要的地位，因为任一函数都可以分段用多项式来逼近。因此，在通常的实际问题中，不论依变量与其他自变量的关系如何，我们总可以用多项式回归来进行分析。

2. 多项式回归模型

2.1. 模型的建立

一元 m 次多项式回归方程为：

$$\hat{y} = b_0 + b_1 x + b_2 x^2 + \cdots + b_m x^m \tag{1}$$

令 $x_1 = x,\ x_2 = x^2,\ \cdots,\ x_m = x^m$，则回归方程转化为：

$$y_t = \beta_0 + \beta_1 x_{t1} + \beta_2 x_{t2} + \beta_3 x_{t3} + \cdots + \beta_p x_{tp} + \varepsilon_t, \quad (t = 1, 2, \cdots, n),\ \varepsilon_t \sim N\left(0, \sigma^2\right) \tag{2}$$

式中，t 表示观测值变量，共有 n 组观测值，p 表示因子个数，具体分析步骤如下：

多元线性回归的数学模型如(2)所示，用矩阵表示为：

$$y = x\beta + \varepsilon \tag{3}$$

式中，y 为维变形量的观测向量(因变量)，$y = \left(y_1, y_2, \cdots, y_n\right)^{\mathrm{T}}$；$x$ 是一个 $n \times \left(p+1\right)$ 矩阵，其形式为：

$$x = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1p} \\ 1 & x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix}$$

β 是回归系数向量，$\beta = \left(\beta_0, \beta_1, \cdots, \beta_p\right)^{\mathrm{T}}$；$\varepsilon$ 称残差，是服从正态分布的 n 维随机向量，$\varepsilon = \left(\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n\right)^{\mathrm{T}}$。

由最小二乘法可求得 β 的估值 β' 为：

$$\beta' = \left(x^{\mathrm{T}} x\right)^{-1} x^{\mathrm{T}} y \tag{4}$$

2.2. 模型的检验

1) 拟合优度检验

是指回归曲线对观测值的拟合程度。显然若观测点离回归曲线近，则拟合程度好；反之则拟合程度差。度量拟合优度的统计量是确定系数 R^2。

$$R^2 = \frac{\mathrm{ESS}}{\mathrm{TSS}} = 1 - \frac{\mathrm{RSS}}{\mathrm{TSS}} \tag{5}$$

式中：

$\mathrm{TSS} = \sum_{i=1}^{n}\left(y_i - \bar{y}\right)^2$ (称为总离差平方和)；

$\mathrm{ESS} = \sum_{i=1}^{n}\left(\hat{y}_i - \bar{y}\right)^2$ (称为回归平方和)；

$\mathrm{RSS} = \sum_{i=1}^{n}\left(y_i - \hat{y}\right)^2$ (称为残差平方和)；

该统计量越接近于 1，模型的拟合优度越高。

2) 方程的显著性检验(F 检验)

在实际问题中，事先不知道自变量和因变量之间是否存在线性关系，在求线性回归方程之前，(5)式的线性模型只是一种假设。方程的显著性检验，旨在对模型中被解释变量与解释变量之间的线性关系在总体上是否显著成立作出推断。

如果因变量 y 与自变量 $x_1, x_2, \cdots, x_p$ 之间不存在线性关系，则模型(2)中的 β 为 0 向量，即有原假设：

$$H_0\text{:}\ \ \beta 1 = 0, \cdots, \beta_p = 0$$

在此假设作为(5)的约束条件，求得统计量

$$\mathrm{F} = \frac{\mathrm{ESS}/p}{\mathrm{RSS}/\left(n - p - 1\right)} \tag{6}$$

式中：

$\mathrm{ESS} = \sum_{i=1}^{n}\left(\hat{y}_i - \bar{y}\right)^2$ (称为回归平方和)；

$\mathrm{RSS} = \sum_{i=1}^{n}\left(y_i - \hat{y}\right)^2$ (称为残差平方和)；

$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$

在原假设成立时，统计量 F 应服从 $\mathrm{F}\left(p, n - p - 1\right)$ 分布，故在选择显著性水平 α 后，可用下式检验原假设：

$$p\left\{\left|\mathrm{F}\right| \geq \mathrm{F}_{1-\alpha, p, n-1} \middle| H0\right\} = \alpha \tag{7}$$

对回归方程的有效性进行检验，假如(6)成立，即认为在显著性水平 α 下，y 对 $x_1, x_2, \cdots, x_p$ 有显著的线性关系，回归方程是显著的。

3. 工程应用

岭澳核电站二期采用CPR1000技术半速汽轮机，为输出功率为1118.79 MWe的压水堆核电机组，汽轮

机采用额定功率为1118.79 MW、1500 r/m、单轴、中间再热四缸六排汽凝汽式汽轮机。汽轮机本体部分由一台高中压缸和两台低压缸组成，布置在+16.20 m层汽机平台上。根据半速机的特点，在汽机柱和汽机平台分别设置了2组24个监测点，定期进行沉降监测。下表1为3MX232沉降点的部分观测数据。

应用5次多项式进行回回归分析，具体步骤如下：

3.1. 模型的建立

根据上表数据，将时距作为自变量，实测值作为因变量，利用matlab软件进行多项式回归分析。一般来说，多项式阶数越高，预测精度越高，在本例中，通过实验发现，将多项式从5阶升高到6阶后，精度的提高已经不是非常显著，因此本例采用5阶多项式。matlab代码如下：

```
x=[0 149 158 191 221 254 277 321 384 436 463 496 527 558 583
  586 588 589 591 627 638 667 707 714 732 739 748 760 774 778]
y=[0 1.8 1 1.8 1.5 1.8 1.7 4.4 4.3 5.2 5.6 5.4 6.1 4.6
  5.7 3.6 5.6 5.8 6 7.2 7 8.4 8 8.9 9.2 9.1 9 8.4 8.1 8]
[p,s]=polyfit(x,y,5)
u=0:1:800;
v=polyval(p,u);
w=polyval(p,x);
s=[814]
f=polyval(p,s) '预测值
w;
plot(x,y,'o:',u,v)
xlabel('时间(d)')
ylabel('沉降量(mm)')
title('5阶多项式回归曲线')
```

根据计算结果，得出回归方程为：

$$\hat{y} = -9.229\mathrm{e}-013x^5 + 1.88\mathrm{e}-009x^4 - 1.358\mathrm{e}-006x^3 + 0.000409x^2 - 0.03315x + 0.16226 \quad (8)$$

利用模型(8)求出预测值以及残差，结果见表2，拟合曲线见图1。

Table 1. Settlement data of turbine foundation
表 1. 汽机基础沉降数据

观测时间	*时距* d	*实测值* m	*观测时间*	*时距* d	***实测值* m**
08-3-16	0	0	09-6-23	463	5.6
08-8-12	149	1.8	09-7-26	496	5.4
08-8-21	158	1.4	09-8-26	527	6.1
08-9-23	191	1.8	09-9-26	558	4.6
08-10-23	221	1.5	09-10-21	583	5.7
08-11-25	254	1.8	09-10-24	586	3.6
08-12-18	277	1.7	09-10-26	588	5.6
09-2-1	321	4.4	09-10-27	589	5.8
09-4-5	384	4.3	09-10-29	591	6
09-5-27	436	5.2	09-12-4	627	7.2

Table 2. Results of settlement forecasting by polynomial regression method
表 2. 多项式回归法预测结果

观测时间	*时距* d	*实测值* m	*预测值* m	*偏差* mm	***相对误差*** **%**
09-12-15	638	7	6.9937	−0.0063	0.09%
10-1-13	667	8.4	7.5603	−0.8397	10%
10-2-22	707	8	8.2477	0.2477	3.10%
10-3-1	714	8.9	8.3438	−0.5562	6.25%
10-3-19	732	9.2	8.5388	−0.6612	7.19%
10-3-26	739	9.1	8.5902	−0.5098	5.60%
10-4-4	748	9	8.6323	−0.3677	4.09%
10-4-16	760	8.4	8.6402	0.2402	2.86%
10-4-30	774	8.1	8.5678	0.4678	5.78%
10-5-4	778	8	8.5288	0.5288	6.61%
10-6-9	814	7.7	7.724	0.024	0.31%

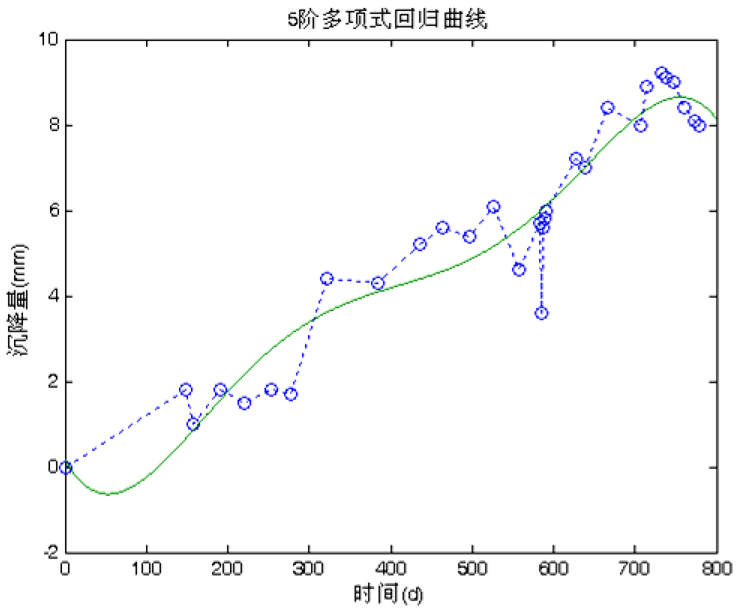

Figure 1. Settlement prediction curve by polynomial regression method
图 1. 多项式回归法沉降量预测曲线

3.2. 模型检验

1) 拟合优度检验，根据计算，$R^2 = 0.9185$，说明模型拟合程度较好。

2) 回归方程检验(F 检验)

F 的统计量为 43.18，取 $\alpha = 0.05$，查 F 分布表可得自由度为 (k) 和 $(n-k-1)$ 的临界值为 $F_{0.05}(6,30-6-1) = 2.57$。显然，F 的统计量大于临界值，说明 5 个自变量与因变量成显著关系。

3.3. 预测精度分析

衡量预测精度的指标有相对误差和平均相对误差。两者的计算公式分别为：

$$\rho = \frac{\hat{y} - y}{y} \times 100\%$$

$$\bar{\rho}_1 = \frac{1}{n} \sum \rho_1$$

从表 1 可以看出，5 阶多项式进行汽机平台沉降预测，最大相对误差为 29.06%。平均相对误差为 10.2%，说明用 5 阶多项式进行汽机平台沉降预测，具有较高的精度。

4. 结论

为了验证模型精度，作者用将多项式回归分析和 GM 模型在本案中经行了对比[7] [8]，GM 模型预测曲线如图 2 所示。

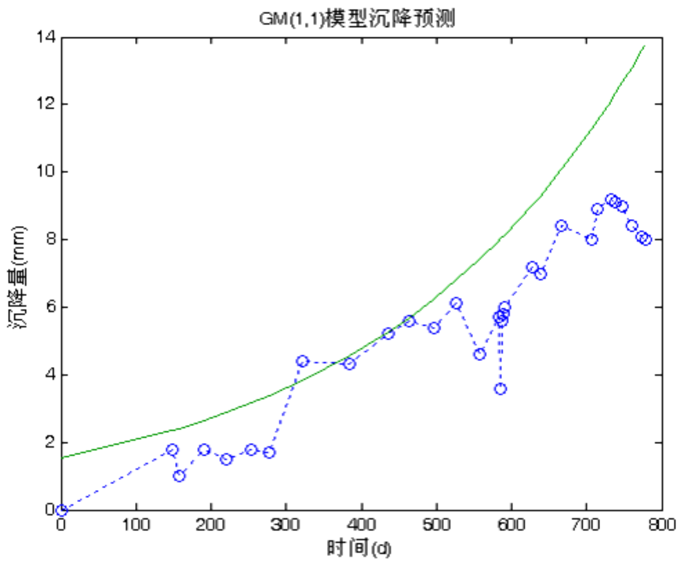

Figure 2. Settlement prediction curve by GM method
图 2. GM 模型沉降预测曲线

从两者对比来看，多项式回归模型在本案中预测效果好于 GM 模型。

但是此种方法只能建立一个自变量和因变量的对应关系。如果要详细分析各种因素对沉降量的影响，沉降监测过程中应该引入时间、荷载等多种自变量，利用多元回归分析建立各个自变量和应变量之间的关系，才能更好的为施工和安全决策服务。

参考文献 (References)

[1] 工程测量规范 (2005) GB50026-2007.

[2] 建筑变形测量规程 (1998) JGJ/T 8-971998-06-01.

[3] 火力发电厂土建结构设计技术规定 (1993) DL 5022-1993.

[4] 王黎明, 陈颖, 杨楠 (2008) 应用回归分析. 复旦大学出版社, 上海.

[5] 张俊中, 宋蕾, 等 (2004) 多元回归分析模型在变形监测中的应用. *河南工程学院学报*, **3**, 22-25.

[6] 张勤, 张启能, 等 (2003) 生物统计学. 中国农业大学出版社, 北京.

[7] 邓聚龙 (1987) 灰色系统基本方法. 华中理工大学出版社, 武汉.

[8] 肖南, 赵来军, 党林立 (2003) 现代数值计算方法. 北京大学出版社, 北京.

[9] 吴清海 (2006) 非等间距灰色模型在沉降预测中的应用. *淮海工学院学报*(*自然科学版*), **4**, 12
[10] 张满 (2007) 沉降观测周期的确定方法. *现代测绘*, **6**, 28, 30.
[11] 郭志强 (2007) 对高层建筑施工中的沉降观测的探讨. *勘察*、*测绘与测试技术*, **11X**, 100-101.
[12] 陈伟清 (2007) 建筑物沉降变形分析与预测技术应用. *勘察科学技术*, **3**, 53-55.
[13] 曹玉明 (2007) 论电厂建(构) 筑物沉降观测的质量控. *岩土工程·勘测*, **3**, 30-32.
[14] 丁召武 (2007) 论电厂汽机岛基础沉降及安装的相关问题. *广东科技*, **10X**, 177-178.