某面包房经营数据的研究及预测

摘要

当今时代，随着学校的不断发展，越来越多的餐饮类商家也在逐步被引入学校内部， 满足学校师生多元化的美食需求。本文基于题目所给的数据信息，建立数学模型进行分析，从而对面包房日后的经营数据进行预测，为其长期进驻校园提供数据参考。

针对问题一，本文通过对表格数据的观察，发现销售额的增长总体呈现上升趋势，我们针对麦A店的销售数据建立GM（1，1）灰色预测模型，其中预测采用了 Matlab分析，从而以“十一”假期前的销售数据来估算出假期这五天的营业额。

针对问题二，本文通过对表格中所给的数据进行分析，发现数据无明显规律，便采用线性拟合对营业额进行预测，通过构建多项式函数的方法确定日期与营业额的关系。最后，将日期带入函数关系式，进而得到11月前三日的预测值为，判断其是否能超过3000元。

**关键词：** **GM（1，1）灰色预测** **线性拟合 非线性回归**

1、问题重述

学校后勤上学期引进了多家品牌店，让“吃在学校” 的同学们倍感幸福。几位老师也去“探店”了，从某面包房（称作“麦A店”）获得了部分经营数据，为该店开业初期一段时间内的营业额（附件1）。为了让该品牌店能进驻校园并长期的开下去，后勤集团想邀请你们团队帮忙。希望你们充分应用所掌握的数学知识，结合建模、编程求解等。

1. 附件1中有自9月9日开业至11月1日之前的营业额，但中间缺失了“十一”假期的5天数据，请通过多种合适方法估算出这5天的营业额，并检验其方法的准确性。
2. 依据附件1数据，建立数学模型对“麦A店”11月1日、2日和3日的营业额进行预测，并根据模型确定营业额能否突破3000元，如果能，请预测出首次突破3000元的日期；如果不能，请根据模型分析说明理由。

2、问题分析

2.1 对问题一的分析

问题一要求利用已有的营业额对此后短时间的营业额进行预测，对于已有数据进行分析，**发现得不出明显规律且数据量较少**。进行**级比检验**后，发现原始数据无法通过检验，对此，本文对原始数据进行平移变换处理，将所有数据在原来的基础上分别增加**10**，生成新的数据，进行**级比检验**后，发现数据满足**GM（1，1）灰色预测模型**的要求。因此，利用**GM（1，1）灰色预测模型进行**解答。

2.2 对问题二的分析

问题二本质上是利用已有的部分营业额对11月份开始三天的销售数据进行预测，因为这三天是对已知的大时间范围的延续，因此本文建立非线性回归中的多项式回归模型对营业额进行预测，将已有数据输入matlab中，利用编写的代码进行线性拟合，从而对11月份开始三天的营业额进行预测，得出预测值判断其其是否超过3000元。

3、模型假设

（1）假设本题所提供的营业额数据均真实有效，具有计算分析价值。

（2）假设本题中所涉及的某面包店不会遭遇不可抗力影响。

（3）假设该面包店的营业情况不会受季节变化等因素影响。

4、符号说明

|  |  |
| --- | --- |
| 符号 | 符号意义 |
|  | 第k天的营业额 |
|  | 前k天的营业额之和 |
|  | 相对误差 |
|  | 级比偏差 |
|  | 级比检验 |
|  |  |
|  |  |

5、模型的建立与求解

5.1GM（1，1）灰色预测模型的建立与求解

数据的预处理：

题目已给出该面包店9月9日至9月29日的每一日的营业额，将已知的数据在此列出。

|  |  |
| --- | --- |
| 日期 | 营业额 |
| 2022年9月9日 | 160 |
| 2022年9月10日 | 175 |
| 2022年9月11日 | 145 |
| 2022年9月12日 | 155 |
| 2022年9月13日 | 190 |
| 2022年9月14日 | 228 |
| 2022年9月15日 | 257 |
| 2022年9月16日 | 281 |
| 2022年9月17日 | 316 |
| 2022年9月18日 | 288 |
| 2022年9月19日 | 305 |
| 2022年9月20日 | 380 |
| 2022年9月21日 | 396 |
| 2022年9月22日 | 459 |
| 2022年9月23日 | 520 |
| 2022年9月24日 | 505 |
| 2022年9月25日 | 510 |
| 2022年9月26日 | 590 |
| 2022年9月27日 | 620 |
| 2022年9月28日 | 690 |
| 2022年9月29日 | 749 |

首先将2022年9月9日的营业额记为，则2022年9月10日的营业额为，后续数据处理方式相同。将原始数据进行级比检验，引入进行检验，其中：

经过级比检验后，发现数据并不符合灰色预测的要求，故进行平移变换。经研究发现，将原始数据分别增加10时，数据恰好满足灰色预测模型要求，故将原数据分别添加10作为预测所用数据。

由此可得，对经过数据处理后的新数据进行整理，建立如下表格

|  |  |
| --- | --- |
| 序号 | 营业额 |
| 1 | 170 |
| 2 | 185 |
| 3 | 155 |
| 4 | 165 |
| 5 | 200 |
| 6 | 238 |
| 7 | 267 |
| 8 | 291 |
| 9 | 326 |
| 10 | 298 |
| 11 | 315 |
| 12 | 390 |
| 13 | 406 |
| 14 | 469 |
| 15 | 530 |
| 16 | 515 |
| 17 | 520 |
| 18 | 600 |
| 19 | 630 |
| 20 | 700 |
| 21 | 759 |

其中，新数据将替代原始数据进行使用。数据无明显规律，将其累加生成序列，其计算公式为：

将所得到的数据进行整理和绘图，结果如下：

由图可知，近似为一条指数曲线，因此可用一个指数曲线的表达式来逼近这个新序列，相应可构建一阶常微分方程来求解拟合指数曲线的函数表达式，因此，设：

但已知的数据是离散的，同时为了进一步消除数据随机性，将改写为，将改写为，考虑原方程的数据较少等因素，因此将改为取前后两个时刻的均值。因此，对原式进行均值生成和改写，得：

其中，和为已知数据，函数还有两个未知参数，就可用最小二乘法求出未知参数，即建立拟合函数，求出当拟合函数求的值与已知数据的平方差最小时，便得到未知参数的取值。同时设方程矩阵形式

即：

由此可知当取得最小值时，可得到U的值。求解可得：

可通过此式求出，将所求出的参数带入原微分方程，求出的值。将带入即可得到2022年9月30日的营业额的预测值。

此后，对模型分别进行相对误差检验和级比偏差检验，将数据分别带入以下公式进行检验：

相对误差检验：

级比偏差检验：

其中，，分别都小于0.2，则说明模型适用。

将数据和代码输入matlab中可解出所需预测的营业额的值，经过平移变换转换回去后，可得出预测数据如下：

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 814.4 | 881.5 | 954.0 | 1032.5 | 1117.3 |

5.2基于多项式方程的非线性回归模型的建立与求解

首先，我们将“十一”假期之后的营业额整理，并在此将数据列出，如下：

|  |  |  |  |
| --- | --- | --- | --- |
| 序号 | 营业额 | 序号 | 营业额 |
| 1（2022年10月6日） | 1100 | 15 | 1580 |
| 2 | 1111 | 16 | 1688 |
| 3 | 1200 | 17 | 1700 |
| 4 | 1225 | 18 | 1737 |
| 5 | 1280 | 19 | 1670 |
| 6 | 1300 | 20 | 1695 |
| 7 | 1345 | 21 | 1798 |
| 8 | 1380 | 22 | 1742 |
| 9 | 1489 | 23 | 1650 |
| 10 | 1499 | 24 | 1805 |
| 11 | 1521 | 25 | 1790 |
| 12 | 1490 | 26 | 1718 |
| 13 | 1526 | 27 | 1768 |
| 14 | 1631 |  |  |

首先使用皮尔逊相关系数对数据进行分析：

皮尔逊系数：两个变量之间的皮尔逊相关系数被定义为两个变量之间的协方差和标准差的商，公式由此得出：

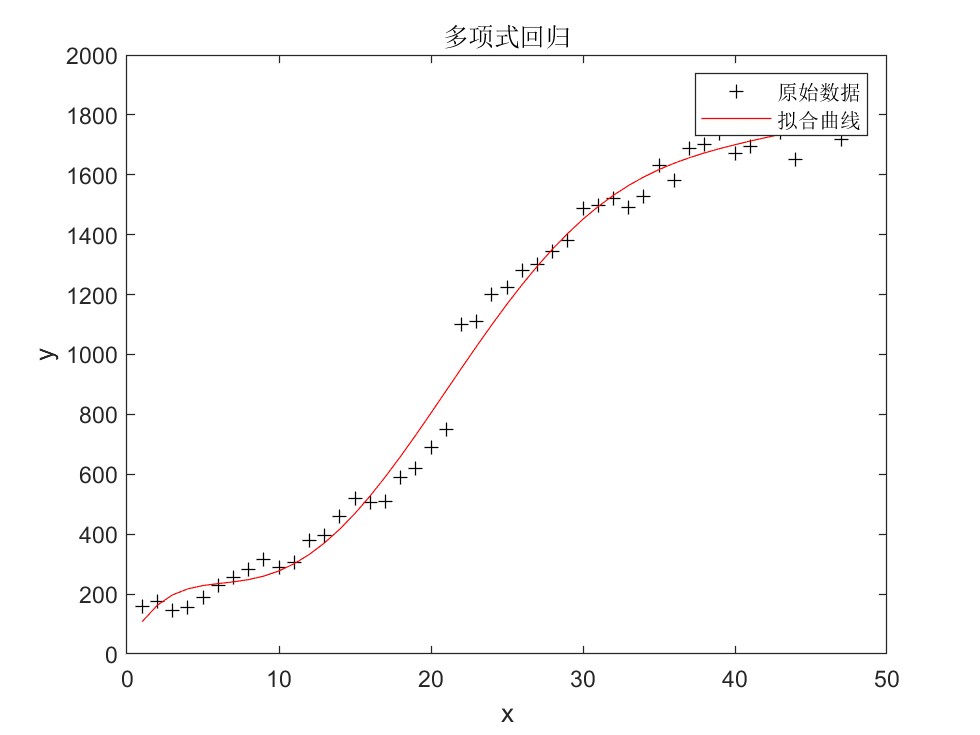
再估算样本的标准差以及协方差，便可得出皮尔逊相关系数：

r 也可由样本点的标准分数均值估计，得到与上式等价的表达式：

其中，，，分别都是是对样本的标准分数、样本平均值和样本标 准差。通过 matlab 进行分析相关系数的相关性比较，可根据下表：

|  |  |  |
| --- | --- | --- |
| 相关性 | 负 | 正 |
| 无相关性 | -0.09to0.0 | 0.0to0.09 |
| 弱相关性 | -0.3to-0.1 | 0.1to0.3 |
| 中相关性 | -0.5to-0.3 | 0.3to0.5 |
| 强相关性 | -1.0to-0.5 | 0.5to1.0 |

根据上述数据和“十一”假期之前的营业额，将所有数据进行线性拟合，并将附件一中的销售额数据导入 MATLAB，进行代码求解，得到了这样的图像结果：



横轴 x 为日期变化，竖轴 y 为销售额变化，由计算及编程可以得出该曲线及其函数关系：

同样的，进行计算将11月1日，11月2日，11月3日这三天的日期分别代入函数进行计算，得出的结果分别为：11月1日营业额:1845 ，11月2日营业额:1905 ，11月3日销售额:1945。

6、附录

附录一：问题一预测“十一”假期五天的营业额

%输入历史数据

history\_data = [170 185 155 165 200 238 267 291 326 298 315 390 406 469 530 515 520 600 630 700 759];

n = length(history\_data);

X0 = history\_data&apos;;% 累加生成序列

history\_data\_agg = cumsum(history\_data);

X1 = history\_data\_agg&apos;;% 计算数据矩阵B和数据向量Y

B = zeros(n - 1, 2);

Y = zeros(n - 1, 1);

for i = 1:n – 1

B(i, 1) = -0.5 \* (X1(i) + X1(i + 1));

B(i, 2) = 1;

Y(i) = X0(i + 1);

end% 计算GM(1,1)微分方程的参数a和u

A = (B&apos; \* B) \ B&apos; \* Y;

a = A(1);

u = A(2);% 建立灰色预测模型

XX0 = zeros(n, 1);

XX0(1) = X0(1);

for i = 2:n

XX0(i) = (X0(1) - u / a) \* (1 - exp(a)) \* exp(-a \* (i - 1));

end% 模型精度的后验差检验

e = mean(X0 - XX0);% 求历史数据平均值

aver = mean(X0);% 求历史数据方差

s12 = var(X0);% 求残差方差

s22 = var(X0 - XX0 - e);% 求后验差比值

C = s22 / s12;% 求小误差概率

count = sum(abs((X0 - XX0) - e) < 0.6754 \* sqrt(s12));

P = count / n;

if (C < 0.35 && P > 0.95)% 预测精度为一级

m = 5; % 请输入需要预测的年数

fprintf(&apos;往后 %d 各年负荷为：\n&apos;, m);

f = zeros(m, 1);

for i = 1:m

f(i) = (X0(1) - u / a) \* (1 - exp(a)) \* exp(-a \* (i + n - 1));

disp(f(i));

end

else

disp(&apos;灰色预测法不适用&apos;);

end

附录二：问题二预测11月初三天的营业额

x=1:1:48;

y=[160 175 145 155 190 228 257 281 316 288 305 380 396 459 520 505 510 590 620 690 749 1100 1111 1200 1225 1280 1300 1345 1380 1489 1499 1521 1490 1526 1631 1580 1688 1700 1737 1670 1695 1798 1742 1650 1805 1790 1718 1768];

[p,s]=polyfit(x,y,6);

t=polyconf(p,x,s);

plot(x,y,'k+',x,t,'r');

xlabel('x');

ylabel('y');

title('多项式回归');

legend('原始数据','拟合曲线');