

基于灰色系统 GM(1,1)模型分析 2019-nCoV

对意大利旅游经济的影响

1 研究目的

意大利旅游业发达，是世界第五大旅游国。此次疫情的爆发必定对其旅游行业产生很大的冲击。为了定量分析新冠肺炎对 2020 年意大利旅游收入产生的影响，本次报告使用灰色系统分析方法来进行有关分析。

2 数据来源

本次数据来源于 <https://zh.tradingeconomics.com/> 中有关意大利旅游收入的月度数据（以欧元计算）。为了避免 2008 年金融危机对本次分析产生影响，我们选取 2009 年-2020 年的数据。

3 模型介绍

灰色模型（Grey Model，简称 GM），即通过累加、累减等方式把原始数据变换成生成序列，再利用其中的规律进行建模预测的方法，可针对不同问题建立相应的灰色模型，如 GM(1,1)，GM(1,m)，GM(n,m)，分别用于一个自变量一个因变量，多个自变量一个因变量，多个自变量多个因变量。模型的优点是可以利用较少的数据进行预测，通过微分方程充分挖掘系统的本质，精度较高，可适用于中短期的预测。

其中，GM(1,1) 模型是灰色系统理论中的基本模型，它将无规律的原始数据累加生成后，使其变为有规律的生成数列，然后建立相应的微分方程模型，从而预测事物未来发展趋势和状态。

根据我们收集的历史数据可以看出，在正常情况下，全年的平均值能较好地反映出意大利旅游收入的变化规律，据此把预测分成两部分：

(1) 利用灰色理论建立 GM(1,1) 模型，由 2009-2019 年的平均值预测 2020 年平均值；

(2) 通过历史数据计算每个月的旅游收入与全年总值的关系，从而预测出正常情况下 2020 年前 4 个月的旅游收入。

GM(1,1) 建模步骤：

①数据的检验与处理

根据收集的数据计算出 2009-2019 年意大利旅游收入的平均值，即模型的原始数据列，记为 $X^{(0)} = (x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(n))$ ，计算数列的级比：

$$\lambda(k) = \frac{x^{(0)}(k-1)}{x^{(0)}(k)}, k = 2, 3, \dots, n$$

若所有级比都落在可容覆盖区间 $X = (e^{\frac{-2}{n+1}}, e^{\frac{2}{n+1}})$ 内，则数列 $X^{(0)}$ 可建立 GM(1,1) 模型。否则，需要对数据做适当的变换处理，如平移变换

$y^{(0)}(k) = x^{(0)}(k) + c, k = 1, 2, \dots, n$ ，取 c 使得数据列的级比都落在可容覆盖内。本文的原始序列级比都在可容覆盖区间 $(0.846, 1.181)$ 内，因此可以建立 GM (1,1) 模型。

②累加生成

令 $X^{(1)}$ 为 $X^{(0)}$ 的一次累加序列，则 $X^{(1)} = (x^{(1)}(1), x^{(1)}(2), \dots, x^{(1)}(n))$ ，其中，

$$x^{(1)}(k) = \sum_{i=1}^k x^{(0)}(i), k = 1, 2, \dots, n。$$

③均值生成

令 $Z^{(1)}$ 为 $X^{(1)}$ 的紧邻均值生成序列， $Z^{(1)} = (z^{(1)}(1), z^{(1)}(2), \dots, z^{(1)}(n))$ ，其中，

$$z^{(1)}(k) = \frac{1}{2}(x^{(1)}(k) + x^{(1)}(k+1)), k = 1, 2, 3, \dots, n。$$

④GM (1,1) 模型

GM (1,1) 模型的灰微分方程为： $X^{(0)}(k) + az^{(1)}(k) = b$ 。式中， a 为发展系数， b 为灰作用量，均为待定系数。

设待估参数向量 $\hat{\alpha} = [\hat{a}, \hat{b}]^T$ ，则灰微分方程的最小二乘估计参数列满足

$$\hat{\alpha} = [\hat{a}, \hat{b}]^T = (B^T B)^{-1} B^T Y，其中，$$

$$B = \begin{bmatrix} -z^{(1)}(2) & 1 \\ -z^{(1)}(3) & 1 \\ \vdots & \vdots \\ -z^{(1)}(n) & 1 \end{bmatrix}, \quad Y = \begin{bmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ \vdots \\ x^{(0)}(n) \end{bmatrix}$$

则此模型的时间响应函数为 $x^{(1)}(t) = (x^{(1)}(1) - \frac{b}{a})e^{-at} + \frac{b}{a}$ ，估计出参数 a 、 b 以后，可以利用此函数对原始序列进行预测。

⑤模型检验

在利用 GM (1,1) 模型进行预测前，需要进行模型精度和误差的检验，来保证预测的准确性。

(1)残差检验

对估计值和实际值的残差进行逐点检验，并计算原始序列 $x^{(0)}(i)$ 与 $\hat{x}^{(0)}(i)$ 的

绝对残差序列 $\Delta^{(0)}(i) = |x^{(0)}(i) - \hat{x}^{(0)}(i)|$ ，进而计算其相对残差序列 $\Phi_i = \frac{\Delta^{(0)}(i)}{x^{(0)}(i)}$ ，最

后得出整个模型的平均相对误差 $\overline{\Phi} = \frac{1}{n} \sum_{i=1}^n \Phi_i$ ，将其与给定的残差水平进行比较，

若小于给定的残差水平，则称模型为残差合格模型。

(2)后验差检验

对残差分布的统计特性进行检验。首先，分别计算原始序列和残差序列的平均值 $\bar{X} = \frac{1}{n} \sum_{k=1}^n x^{(0)}(k)$ 、方差 $S_1 = \sqrt{\frac{1}{n} \sum_{k=1}^n [x^{(0)}(k) - \bar{X}]^2}$ 、残差的均值 $\bar{E} = \frac{1}{n-1} \sum_{k=2}^n E(k)$ 、

残差的方差 $S_2 = \sqrt{\frac{1}{n-1} \sum_{k=2}^n [E(k) - \bar{E}]^2}$ 、后验差比值 $C = \frac{S_2}{S_1}$ 、小概率误差

$P = P\{|E(k) - \bar{E}| < 0.6745 S_1\}$ ，最后，根据相应标准的临界值的大小，综合评价模型的拟合效果，常用的模型精度判断等级见表 1。

表 1 GM (1,1) 模型拟合的精度要求

检验指标	优	合格	勉强合格	不合格
平均相对误差	<0.01	<0.05	<0.1	>0.2
C 值	<0.35	<0.45	<0.5	>0.65

⑥预测

根据历史数据计算出 2020 年每个月的旅游收入占全年总值的比例 u_i ，即

$u_i = \sum_{j=1}^{11} a_{ij} / \sum_{i=1}^{12} \sum_{j=1}^{11} a_{ij}$ ，在得到 2020 年的预测值后，可算出 2020 年每个月的旅游收入

预测值。

4 程序介绍

#构建 GM(1,1) 函数

```
gm11<-function(x0, t){
  x1<-cumsum(x0)
  b<-numeric(length(x0)-1)
  n<-length(x0)-1
  for(i in 1:n){
    b[i]<--(x1[i]+x1[i+1])/2
  }
  b}
```

#得到 GM(1,1) 参数估计值并输出

```
D<-numeric(length(x0)-1)
```

```

D[]<-1
B<-cbind(b,D)
BT<-t(B)
M<-solve(BT%*%B)
YN<-numeric(length(x0)-1)
YN<-x0[2:length(x0)]
alpha<-M%*%BT%*%YN
alpha2<-matrix(alpha,ncol=1)
a<-alpha2[1]
u<-alpha2[2]
cat("GM(1,1) 参数估计值: ",'\n',"发展系数-a=", -a," ", "灰色作用量
u=", u,'\n','\n')
#计算 x(0) 的模拟值
y<-numeric(length(c(1:t)))#t 为给定的预测个数
y[1]<-x1[1]
for(w in 1:(t-1)) {
  y[w+1]<-(x1[1]-u/a)*exp(-a*w)+u/a
}
#输出 x(0) 的模拟值
xy<-numeric(length(y))
xy[1]<-y[1]
for(o in 2:t) {
  xy[o]<-y[o]-y[o-1]
}
cat("x(0) 的模拟值: ",'\n',xy,'\n','\n')
#计算残差 e
e<-numeric(length(x0))
for(l in 1:length(x0)) {
  e[l]<-x0[l]-xy[l] #得残差
}
#计算相对误差
e2<-numeric(length(x0))
for(s in 1:length(x0)) {
  e2[s]<-(abs(e[s])/x0[s]) #得相对误差
}
cat("残差平方和=", sum(e^2), '\n')
cat("平均相对误差=", sum(e2)/(length(e2)-1)*100, "%", '\n')
cat("相对精度=", (1-(sum(e2)/(length(e2)-1)))*100, "%", '\n', '\n')
#后验差比值检验

avge<-mean(abs(e)); esum<-sum((abs(e)-avge)^2); evar=esum/(length(e)-1);
se=sqrt(evar)

avgx0<-mean(x0); x0sum<-sum((x0-avgx0)^2); x0var=x0sum/(length(x0)); sx=

```

```

sqrt(x0var)
cv<-se/sx #得验差比值
cat("后验差比值检验:",'\n',"C 值=",cv,'\n')#对后验差比值进行检验,与
一般标准进行比较判断预测结果好坏。
if(cv < 0.35){
  cat("C 值<0.35, GM(1,1) 预测精度等级为: 好", '\n', '\n')
}else{
  if(cv<0.5){
    cat("C 值属于[0.35, 0.5), GM(1,1) 模型预测精度等级为: 合格", '\n',
'\n')
  }else{
    if(cv<0.65){
      cat("C 值属于[0.5, 0.65), GM(1,1) 模型预测精度等级为: 勉强合格",
'\n', '\n')
    }else{
      cat("C 值>=0.65, GM(1,1) 模型预测精度等级为: 不合格", '\n', '\n')
    }
  }
}
}
##画出输入序列 x0 的预测序列及 x0 的比较图像
plot(xy, col='blue', type='b', pch=16, xlab='时间序列', ylab='值')
points(x0, col='red', type='b', pch=4)
legend('topleft', c('预测值','原始值'), pch=c(16,4), lty=1,
col=c('blue','red'))
}
#预测意大利月均旅游收入
x0<-c(2404.70, 2438.09, 2574.22, 2671.31, 2757.81, 2853.37, 2962.99, 3029.93,
3262.90, 3476.03, 3691.85)
gm11(x0, length(x0)+1)

```

5 结果分析

5.1 模型求解

由 2009-2019 年意大利旅游收入的月度数据,可计算得年平均值数列及一次累加值数列:

$X^{(0)}=(2404.70, 2388.09, 2574.22, 2671.31, 2757.81, 2853.37, 2962.99, 3029.93, 3262.90, 3476.03, 3691.85)$

$X^{(1)}=(2404.70, 4842.79, 7417.01, 10088.32, 12846.14, 15699.51, 18662.49, 21692.43, 24955.33, 24831.35, 32133.21)$

从而得到其紧邻均值生成序列:

$Z^{(1)}=(3623.75, 6129.90, 8752.67, 11467.23, 14272.82, 17181.00, 20177.46, 23323.88, 266693.34, 30277.28, 32123.21)$

求得其发展系数和灰作用量: $a=-0.0445$, $b=2251.087$, 则此模型的时间响应

函数为

$$x^{(1)}(t)=52990.92e^{0.0445t}-50586.22$$

在利用上式进行预测之前，对其模型精度和误差进行检验。在软件中计算得到，残差检验中 $\Phi=0.0155$ ，在 0.01 与 0.05 之间，模型精度合格；而在后验差检验中， $C=0.0898$ ，小于 0.35，模型精度为优。因此，可以用此模型对 2020 年意大利月均旅游收入进行预测。

根据上式预测得到： $\hat{X}^{(0)}(11)=3763.723$ （百万），即预测所得 2020 年全年旅游收入为 45164.68（百万）。预测结果如图 1 所示。

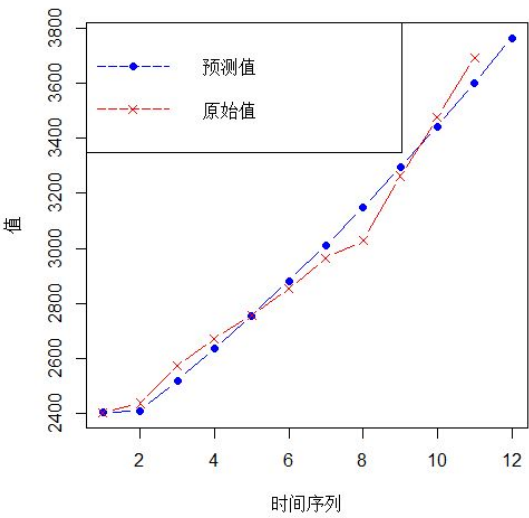


图 1 意大利 2020 年旅游收入预测

同时，2020 年第 i 个月的指标值占全年总值的比例为 $u=(0.0511,0.0459,0.0606,0.0759,0.0934,0.1056,0.1363,0.1313,0.1095,0.0868,0.0539,0.0496)$ ，月度预测值与月度真实值比较结果见表 1。

表 1 月度预测值与月度真实值比较（单位：百万欧元）

月份	1 月	2 月	3 月	4 月
预测值	2308.43	2075.94	2737.21	3425.87
实际值	2263.71	1820.95		

5.2 结果分析

根据上表中的预测结果，意大利 1 月份的旅游收入应为 2308.43 百万欧元，但实际收入却为 2263.71 百万欧元，二者相差 44.72 百万欧元，差距可能并不明显；而预测意大利 2 月份的旅游收入应为 2075.94 百万欧元，实际收入为 1820.96 百万欧元，差距达到了约 255 百万欧元，说明新冠肺炎疫情对意大利旅游产生的冲击在 2 月份更大。这可能是由于在 1 月份人们对新冠肺炎疫情不够了解所造成的，在 1 月 23 日武汉封城后，各国都采取了相应的防范措施，因此对意大利旅游情况的冲击加大。

另外，对于 3、4 月份，由于数据没有更新，无法获得其实际值，后期会进

行一定的补充。