数学建模算法与应用

第8章 时间序列

8.1 确定性时间序列分析方法

时间序列预测技术就是通过对预测目标自身时间序列的处理，来研究其变化趋势的。一个时间序列往往是以下几类变化形式的叠加或耦合。

（1）长期趋势变动。它是指时间序列朝着一定的方向持续上升或下降，或停留在某一水平上的倾向，它反映了客观事物的主要变化趋势。

（2）季节变动。

（3）循环变动。通常是指周期为一年以上，由非季节因素引起的涨落起伏波形相似的波动。

（4）不规则变动。通常它分为突然变动和随机变动。

通常用T_t表示长期趋势项，S_t表示季节变动趋势项，C_t表示循环变动趋势项，R_t表示随机干扰项。常见的确定性时间序列模型有以下几种类型。

（1）加法模型

$$y_t = T_t + S_t + C_t + R_t$$

（2）乘法模型

$$y_t = T_t \cdot S_t \cdot C_t \cdot R_t$$

（3）混合模型

$$y_t = T_t \cdot S_t + R_t, \quad y_t = S_t + T_t \cdot C_t \cdot R_t$$

其中y_t是观测目标的观测记录，均值$E(R_t)=0$，方差$\mathrm{Var}(R_t)=\sigma^2$。

如果在预测时间范围以内，无突然变动且随机变动的方差σ^2较小，并且有理由认为过去和现在的演变趋势将继续发展到未来时，可用一些经验方法进行预测，具体方法如下。

8.1.1 移动平均法

设观测序列为$y_1,\cdots,y_T$，取移动平均的项数$N<T$。一次移动平均值计算公式为

$$\begin{aligned} M_t^{(1)} &= \frac{1}{N}(y_t + y_{t-1} + \cdots + y_{t-N+1}) \\ &= \frac{1}{N}(y_{t-1} + \cdots + y_{t-N}) + \frac{1}{N}(y_t - y_{t-N}) \qquad (8.1) \\ &= M_{t-1}^{(1)} + \frac{1}{N}(y_t - y_{t-N}). \end{aligned}$$

二次移动平均值计算公式为

$$\begin{aligned} M_t^{(2)} &= \frac{1}{N}(M_t^{(1)} + \cdots + M_{t-N+1}^{(1)}) \\ &= M_{t-1}^{(2)} + \frac{1}{N}(M_t^{(1)} - M_{t-N}^{(1)}) \end{aligned} . \qquad (8.2)$$

当预测目标的基本趋势是在某一水平上下波动时，可用一次移动平均方法建立预测模型

$$\hat{y}_{t+1} = M_t^{(1)} = \frac{1}{N}(y_t + \cdots + y_{t-N+1}), \qquad (8.3)$$

其中$t = N, N+1, \cdots, T$。

其预测标准误差为

$$S=\sqrt{\frac{\sum_{t=N+1}^{T}(\hat{y}_t-y_t)^2}{T-N}}.$$

最近N期序列值的平均值作为未来各期的预测结果。一般N取值范围：$5 \le N \le 200$。当历史序列的基本趋势变化不大且序列中随机变动成分较多时，N的取值应较大一些。否则N的取值应小一些。在有确定的季节变动周期的资料中，移动平均的项数应取周期长度。选择最佳N值的一个有效方法是，比较若干模型的预测误差，预测标准误差最小者为好。

当预测目标的基本趋势与某一线性模型相吻合时，常用二次移动平均法，但序列同时存在线性趋势与周期波动时，可用趋势移动平均法建立预测模型

$$\hat{y}_{T+m} = a_T + b_T m，\ m = 1,2,\cdots,$$

其中$a_T = 2M_T^{(1)} - M_T^{(2)}$，$b_T = \dfrac{2}{N-1}(M_T^{(1)} - M_T^{(2)})$。

例 8.1　某企业 1 月～11 月份的销售收入时间序列如表 8.1 示。试用一次简单移动平均法预测第 12 月份的销售收入。

表 8.1　企业销售收入

月份 t	1	2	3	4	5	6
销售收入 y_t	533.8	574.6	606.9	649.8	705.1	772.0
月份 t	7	8	9	10	11	
销售收入 y_t	816.4	892.7	963.9	1015.1	1102.7	

解　分别取$N=4, N=5$的预测公式

$$\hat{y}_{t+1}^{(1)}=\frac{y_t+y_{t-1}+y_{t-2}+y_{t-3}}{4}, \quad t=4,5,\cdots,11,$$

$$\hat{y}_{t+1}^{(2)}=\frac{y_t+y_{t-1}+y_{t-2}+y_{t-3}+y_{t-4}}{5}, \quad t=5,\cdots,11.$$

当$N=4$时，预测值$\hat{y}_{12}^{(1)}=993.6$，预测的标准误差

$$S_1=\sqrt{\frac{\sum_{t=5}^{11}(\hat{y}_t^{(1)}-y_t)^2}{11-4}}=150.5.$$

当$N=5$时，预测值$\hat{y}_{12}^{(2)}=958.2$，预测的标准误差

$$S_2=\sqrt{\frac{\sum_{t=6}^{11}(\hat{y}_t^{(2)}-y_t)^2}{11-5}}=182.4.$$

计算结果表明，$N=4$时，预测的标准误差较小，所以选取$N=4$。预测第 12 月份的销售收入为 993.6。

移动平均法只适合做近期预测，而且是预测目标的发展趋势变化不大的情况。如果目标的发展趋势存在其它的变化，采用简单移动平均法就会产生较大的预测偏差和滞后。

8.1.2 指数平滑法

一次移动平均实际上认为最近N期数据对未来值影响相同，都加权$1/N$，而N期以前的数据对未来值没有影响，加权为 0。但是，二次及更高次移动平均的权数却不是$1/N$，且次数越高，权数的结构越复杂，但永远保持对称的权数，即两端项权数小，中间项权数大，不符合一般系统的动态性。

一般说来历史数据对未来值的影响是随时间间隔的增长而递减的。所以，更切合实际的方法应是对各期观测值依时间顺序进行加权平均作为预测值。指数平滑法可满足这一要求，而且具有简单的递推形式。

指数平滑法根据平滑次数的不同，又分为一次指数平滑法、二次指数平滑法和三次指数平滑法等，分别介绍如下。

1　一次指数平滑法

1）预测模型

设时间序列为 $y_1, y_2, \cdots, y_t, \cdots$，$\alpha$ 为加权系数，$0<\alpha<1$，一次指数平滑公式为

$$S_t^{(1)} = \alpha y_t + (1-\alpha) S_{t-1}^{(1)} = S_{t-1}^{(1)} + \alpha (y_t - S_{t-1}^{(1)}),\quad (8.4)$$

式（8.4）是由移动平均公式改进而来的。由式（8.1）知，移动平均数的递推公式为

$$M_t^{(1)} = M_{t-1}^{(1)} + \frac{y_t - y_{t-N}}{N},$$

以$M_{t-1}^{(1)}$作为y_{t-N}的最佳估计，则有

$$M_t^{(1)}=M_{t-1}^{(1)}+\frac{y_t-M_{t-1}^{(1)}}{N}=\frac{y_t}{N}+\left(1-\frac{1}{N}\right)M_{t-1}^{(1)},$$

令$\alpha=\dfrac{1}{N}$，以S_t代替$M_t^{(1)}$，即得式（8.4）

$$S_t^{(1)}=\alpha y_t+(1-\alpha)S_{t-1}^{(1)}.$$

为进一步理解指数平滑的实质，把式（8.4）依次展开，有

$$S_t^{(1)} = \alpha y_t + (1-\alpha)[\alpha y_{t-1} + (1-\alpha)S_{t-2}^{(1)}]$$
$$= \cdots = \alpha\sum_{j=0}^{\infty}(1-\alpha)^j y_{t-j}$$
，（8.5）

（8.5）式表明$S_t^{(1)}$是全部历史数据的加权平均，加权系数分别为$\alpha,\alpha(1-\alpha)$，$\alpha(1-\alpha)^2,\cdots$，显然有

$$\sum_{j=0}^{\infty}\alpha(1-\alpha)^j = \frac{\alpha}{1-(1-\alpha)} = 1,$$

由于加权系数符合指数规律，又具有平滑数据的功能，故称为指数平滑。

以这种平滑值进行预测，就是一次指数平滑法。预测模型为

$$\hat{y}_{t+1} = S_t^{(1)},$$

即

$$\hat{y}_{t+1} = \alpha y_t + (1-\alpha)\hat{y}_t, \quad (8.6)$$

也就是以第t期指数平滑值作为$t+1$期预测值。

2）加权系数的选择

在进行指数平滑时，加权系数的选择是很重要的。由式（8.6）可以看出，α的大小规定了在新预测值中新数据和原预测值所占的比重。α值越大，新数据所占的比重就愈大，原预测值所占的比重就愈小，反之亦然。若把式（8.6）改写为

$$\hat{y}_{t+1} = \hat{y}_t + \alpha(y_t - \hat{y}_t), \qquad (8.7)$$

则从式（8.7）可看出，新预测值是根据预测误差对原预测值进行修正而得到的。α的大小则体现了修正的幅度，α值愈大，修正幅度愈大；α值愈小，修正幅度也愈小。

若选取$\alpha = 0$，则$\hat{y}_{t+1} = \hat{y}_t$，即下期预测值就等于本期预测值，在预测过程中不考虑任何新信息；若选取$\alpha = 1$，则$\hat{y}_{t+1} = y_t$，即下期预测值就等于本期观测值，完全不相信过去的信息。这两种极端情况很难做出正确的预测。因此，α值应根据时间序列的具体性质在0～1之间选择。具体如何选择一般可遵循下列原则:

（1）如果时间序列波动不大，比较平稳，则α应取小一点，如 0.1～0.5，以减少修正幅度，使预测模型能包含较长时间序列的信息；

（2）如果时间序列具有迅速且明显的变动倾向，则α应取大一点，如 0.6～0.8，使预测模型灵敏度高一些，以便迅速跟上数据的变化。

3）初始值的确定

用一次指数平滑法进行预测，除了选择合适的α外，还要确定初始值$s_0^{(1)}$。初始值是由预测者估计或指定的。当时间序列的数据较多，比如在 20 个以上时，初始值对以后的预测值影响很少，可选用第一期数据为初始值。如果时间序列的数据较少，在 20 个以下时，初始值对以后的预测值影响很大，这时，就必须认真研究如何正确确定初始值。一般以最初几期实际值的平均值作为初始值。

例 8.2　某市 1976～1987 年某种电器销售额如表 8.2 所示。试预测 1988 年该电器销售额。

解　采用指数平滑法，并分别取$\alpha = 0.2, 0.5$和0.8进行计算，初始值

$$S_0^{(1)} = \frac{y_1 + y_2}{2} = 51,$$

即

$$\hat{y}_1 = S_0^{(1)} = 51.$$

按预测模型

$$\hat{y}_{t+1} = \alpha y_t + (1-\alpha)\hat{y}_t$$

计算各期预测值，列于表 8.2 中。

表 8.2 某种电器销售额及指数平滑预测值计算表(单位:万元)

年份	t	实际销售额 y_t	预测值 $\hat{y}_t$ $\alpha=0.2$	预测值 $\hat{y}_t$ $\alpha=0.5$	预测值 $\hat{y}_t$ $\alpha=0.8$
1976	1	50	51	51	51
1977	2	52	50.8	50.5	50.2
1978	3	47	51.04	51.25	51.64
1979	4	51	50.23	49.13	47.93
1980	5	49	50.39	50.06	50.39
1981	6	48	50.11	49.53	49.28
1982	7	51	49.69	48.77	48.26
1983	8	40	49.95	49.88	50.45
1984	9	48	47.96	44.94	42.09
1985	10	52	47.97	46.47	46.82
1986	11	51	48.77	49.24	50.96
1987	12	59	49.22	50.12	50.99

从表 8.2 可以看出，$\alpha = \mathbf{0.2, 0.5}$和$\mathbf{0.8}$时，预测值是很不相同的。究竟$\alpha$取何值为好，可通过计算它们的预测标准误差$S$，选取使$S$较小的那个$\alpha$值。预测的标准误差见表 8.3。

表 8.3　预测的标准误差

α	0.2	0.5	0.8
S	4.5029	4.5908	4.8426

计算结果表明$\alpha = 0.2$时，S较小，故选取$\alpha = 0.2$，预测 1988 年该电器销售额为$\hat{y}_{1988} = 51.1754$。

2　二次指数平滑法

一次指数平滑法虽然克服了移动平均法的缺点。但当时间序列的变动出现直线趋势时，用一次指数平滑法进行预测，仍存在明显的滞后偏差。因此，也必须加以修正。再作二次指数平滑，利用滞后偏差的规律建立直线趋势模型，这就是二次指数平滑法。

其计算公式为

$$\begin{cases} S_t^{(1)} = \alpha y_t + (1-\alpha) S_{t-1}^{(1)}, \\ S_t^{(2)} = \alpha S_t^{(1)} + (1-\alpha) S_{t-1}^{(2)}, \end{cases} \qquad (8.8)$$

式中 $S_t^{(1)}$ 为一次指数的平滑值； $S_t^{(2)}$ 为二次指数的平滑值。当时间序列 $\{y_t\}$，从某时期开始具有直线趋势时，可用直线趋势模型

$$\hat{y}_{t+m} = a_t + b_t m, \quad m = 1, 2, \cdots, \qquad (8.9)$$

$$\begin{cases} a_t = 2S_t^{(1)} - S_t^{(2)}, \\ b_t = \dfrac{\alpha}{1-\alpha}\left(S_t^{(1)} - S_t^{(2)}\right), \end{cases} \qquad (8.10)$$

进行预测。

例 8.3　我国 1965～1985 年的发电总量资料数据见表 8.4（单位：亿度），试用二次指数平滑法预测 1986 年和 1987 年的发电总量。

表 8.4　我国发电总量及一、二次指数平滑值计算表

年份	t	发电总量 y_t	一次平滑值	二次平滑值	y_{t+1} 的估计值
1965	1	676	676	676	
1966	2	825	720.7	689.4	676
1967	3	774	736.7	703.6	765.4
1968	4	716	730.5	711.7	784.0
1969	5	940	793.3	736.2	757.4
1970	6	1159	903.0	786.2	875.0
1971	7	1384	1047.3	864.6	1069.9
⋮	⋮	⋮	⋮	⋮	⋮
1985	21	4107	3523.1	3032.6	3916.6

$\alpha = 0.3$，初始值$S_0^{(1)}$和$S_0^{(2)}$都取序列的首项数值，即$S_0^{(1)} = S_0^{(2)} = 676$。计算$S_t^{(1)}, S_t^{(2)}$，列于表 8.4。得到

$$S_{21}^{(1)} = 3523.1, \quad S_{21}^{(2)} = 3032.6.$$

8.10)，可得$t = 21$时

$$a_{21} = 2S_{21}^{(1)} - S_{21}^{(2)} = 4013.7,$$

$$b_{21} = \frac{\alpha}{1-\alpha}(S_{21}^{(1)} - S_{21}^{(2)}) = 210.24,$$

$t=21$时直线趋势方程为

$$\hat{y}_{21+m}=4013.7+210.24m\,,$$

预测 1986 年和 1987 年的发电总量为（单位：亿度）

$$\hat{y}_{1986}=\hat{y}_{22}=\hat{y}_{21+1}=4223.95\,,$$

$$\hat{y}_{1987}=\hat{y}_{23}=\hat{y}_{21+2}=4434.19.$$

为了求各期的模拟值。可将式（8.10）代入直线趋势模型（8.9），并令$m=1$，则得

$$\hat{y}_{t+1}=(2S_t^{(1)}-S_t^{(2)})+\frac{\alpha}{1-\alpha}(S_t^{(1)}-S_t^{(2)}),$$

即

$$\hat{y}_{t+1}=\left(1+\frac{1}{1-\alpha}\right)S_t^{(1)}-\frac{1}{1-\alpha}S_t^{(2)}. \qquad (8.11)$$

令$t=1,2,\cdots,20$，由公式（8.11）可求出各期的模拟值。计算结果见表 8.4。

3　三次指数平滑法

当时间序列的变动表现为二次曲线趋势时，则需要用三次指数平滑法。三次指数平滑是在二次指数平滑的基础上，再进行一次平滑，其计算公式为

$$\begin{cases} S_t^{(1)} = \alpha y_t + (1-\alpha) S_{t-1}^{(1)}, \\ S_t^{(2)} = \alpha S_t^{(1)} + (1-\alpha) S_{t-1}^{(2)}, \\ S_t^{(3)} = \alpha S_t^{(2)} + (1-\alpha) S_{t-1}^{(3)}, \end{cases} \tag{8.12}$$

式中$S_t^{(3)}$为三次指数平滑值。

三次指数平滑法的预测模型为

$$\hat{y}_{t+m} = a_t + b_t m + C_t m^2, \quad m = 1, 2, \cdots, \qquad (8.13)$$

其中

$$\begin{cases} a_t = 3S_t^{(1)} - 3S_t^{(2)} + S_t^{(3)}, \\ b_t = \dfrac{\alpha}{2(1-\alpha)^2}[(6-5\alpha)S_t^{(1)} - 2(5-4\alpha)S_t^{(2)} + (4-3\alpha)S_t^{(3} \\ c_t = \dfrac{\alpha^2}{2(1-\alpha)^2}[S_t^{(1)} - 2S_t^{(2)} + S_t^{(3)}]. \end{cases}$$

(8.14)

例 8.4　某省 1978～1988 年全民所有制单位固定资产投资总额如表 8.5（单位：亿元）所示，试预测 1989 年和 1990 年固定资产投资总额。

表 8.5　某省全民所有制单位固定资产投资总额及一、二、三次指数平滑值计算表

年份	t	投资总额 y_t	一次平滑值	二次平滑值	三次平滑值	y_{t+1} 的估计值
1978	1	20.04	21.37	21.77	21.89	
1979	2	20.06	20.98	21.53	21.78	20.23
1980	3	25.72	22.40	21.79	21.78	19.56
1981	4	34.61	26.06	23.07	22.17	24.49
⋮	⋮	⋮	⋮	⋮	⋮	⋮
1988	11	232.26	151.77	101.28	68.43	196.26

解　从图 8.1 可以看出，投资总额呈二次曲线上升，可用三次指数平滑法进行预测。

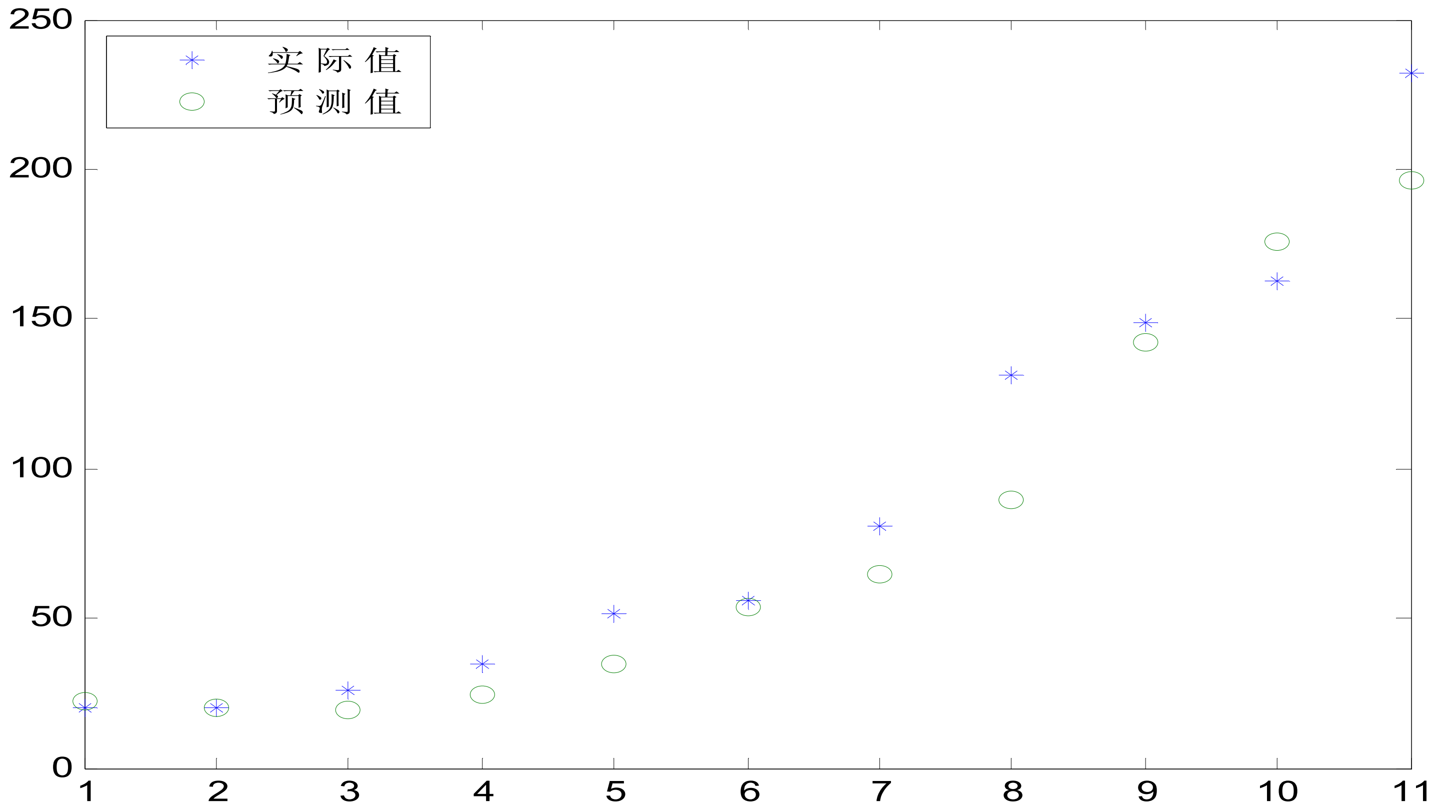

图 8.1　某省固定资产投资总额趋势图

取$\alpha = 0.3$，初始值

$$S_1^{(0)} = S_2^{(0)} = S_3^{(0)} = \frac{y_1 + y_2 + y_3}{3} = 21.94$$ 。

计算$S_t^{(1)}, S_t^{(2)}, S_t^{(3)}$列于表 8.5 中。得到

$$S_{11}^{(1)} = 151.77，\quad S_{11}^{(2)} = 101.28，\quad S_{11}^{(3)} = 68.43.$$

由公式（8.14），可得到当$t = 11$时

$$a_{11} = 219.91，\quad b_{11} = 38.38，\quad c_{11} = 1.62，$$

于是，得$t = 11$时预测模型为

$$\hat{y}_{11+m} = 219.91 + 38.38m + 1.62m^2.$$

预测 1989 年和 1990 年的固定资产投资总额为（单位：亿元）

$$\hat{y}_{1989} = \hat{y}_{12} = \hat{y}_{11+1} = a_{11} + b_{11} + c_{11} = 259.92,$$

$$\hat{y}_{1990} = \hat{y}_{13} = \hat{y}_{11+2} = a_{11} + 2b_{11} + 2^2 c_{11} = 303.16.$$

因为国家从 1989 年开始对固定资产投资采取压缩政策，这些预测值显然偏高，应作适当的修正，以消除政策因素的影响。

与二次指数平滑法一样，为了计算各期的模拟值，可将式（8.14）代入预测模型（8.13），并令$m=1$，则得

$$\hat{y}_{t+1}=\frac{3-3\alpha+\alpha^2}{(1-\alpha)^2}S_t^{(1)}-\frac{3-\alpha}{(1-\alpha)^2}S_t^{(2)}+\frac{1}{(1-\alpha)^2}S_t^{(3)} \tag{8.15}$$

令$t=0,1,2,\cdots,10$，公式（8.15）可求出各期的模拟值，见表 8.5。

指数平滑预测模型是以时刻t为起点，综合历史序列的信息，对未来进行预测的。选择合适的加权系数α是提高预测精度的关键环节。根据实践经验，α的取值范围一般以 0.1～0.3 为宜。α值愈大，加权系数序列衰减速度愈快，所以实际上α取值大小起着控制参加平均的历史数据个数的作用。α值愈大意味着采用的数据愈少。因此，可以得到选择α值的一些基本准则。

（1）如果序列的基本趋势比较稳，预测偏差由随机因素造成，则α值应取小一些，以减少修正幅度，使预测模型能包含更多历史数据的信息。

（2）如果预测目标的基本趋势已发生系统的变化，则α值应取得大一些。这样，可以偏重新数据的信息对原模型进行大幅度修正，以使预测模型适应预测目标的新变化。

另外，由于指数平滑公式是递推计算公式，所以必须确定初始值$S_0^{(1)},S_0^{(2)},S_0^{(3)}$。可以取前 3～5 个数据的算术平均值作为初始值。

8.1.3　差分指数平滑法

当时间序列的变动具有直线趋势时，用一次指数平滑法会出现滞后偏差，其原因在于数据不满足模型要求。因此，我们也可以从数据变换的角度来考虑改进措施，即在运用指数平滑法以前先对数据作一些技术上的处理，使之能适合于一次指数平滑模型，然后再对输出结果作技术上的返回处理，使之恢复为原变量的形态。差分方法是改变数据变动趋势的简易方法。下面讨论如何用差分方法来改进指数平滑法。

1．一阶差分指数平滑法

当时间序列呈直线增加时，可运用一阶差分指数平滑模型来预测。其公式如下

$$\nabla y_t = y_t - y_{t-1}, \tag{8.16}$$

$$\nabla \hat{y}_{t+1} = \alpha \nabla y_t + (1-\alpha)\nabla \hat{y}_t, \tag{8.17}$$

$$\hat{y}_{t+1} = \nabla \hat{y}_{t+1} + y_t, \tag{8.18}$$

其中的∇为差分记号。式（8.16）表示对呈现直线增加的序列作一阶差分，构成一个平稳的新序列；式(8.18)表示把经过一阶差分后的新序列的指数平滑预测值与变量当前的实际值迭加，作为变量下一期的预测值。对于这个公式的数学意义可作如下的解释。

因为

$$y_{t+1} = y_{t+1} - y_t + y_t = \nabla y_{t+1} + y_t, \qquad (8.19)$$

当采用按式（8.18）计算的预测值去估计式（8.19）中的y_{t+1}，从而式（8.19）等号左边的y_{t+1}也要改为预测值，亦即成为式（8.18)。

在前面我们已分析过，指数平滑值实际上是一种加权平均数。因此把序列中逐期增量的加权平均数（指数平滑值）加上当前值的实际数进行预测，比一次指数平滑法只用变量以往取值的加权平均数作为下一期的预测更合理。从而使预测值始终围绕实际值上下波动，从根本上解决了在有直线增长趋势的情况下，用一次指数平滑法所得出的结果始终落后于实际值的问题。

例 8.5　某工业企业 1977～1986 年锅炉燃料消耗量资料如表 8.6（单位：百吨）所示，试预测 1987 年的燃料消耗量。

表 8.6　某企业锅炉燃料消耗量的差分指数平滑法计算表（$\alpha = 0.4$）

年份	t	燃料消耗量 y_t	差分	差分指数平滑值	预测值
1977	1	24			
1978	2	26	2	2	26
1979	3	27	1	2	28
⋮	⋮	⋮	⋮	⋮	⋮
1987	11			2.49	46.49

解　由资料可以看出，燃料消耗量，除个别年份外，逐期增长量大体在 200 吨左右，即呈直线增长，因此可用一阶差分指数平滑模型来预测。我们取 $\alpha = \mathbf{0.4}$，初始值为差分序列首项值，计算结果列于表 8.6 中。

预测 1987 年燃料消耗量为 $\hat{y}_{1987} = \mathbf{46.49}$（百吨）。

2．二阶差分指数平滑模型

当时间序列呈现二次曲线增长时，可用二阶差分指数平滑模型来预测，计算公式如下

$$\nabla y_t = y_t - y_{t-1}, \tag{8.20}$$

$$\nabla^2 y_t = \nabla y_t - \nabla y_{t-1}, \tag{8.21}$$

$$\nabla^2 \hat{y}_{t+1} = \alpha \nabla^2 y_t + (1-\alpha)\nabla^2 \hat{y}_t, \tag{8.22}$$

$$\hat{y}_{t+1} = \nabla^2 \hat{y}_{t+1} + \nabla y_t + y_t, \tag{8.23}$$

其中∇^2表示二阶差分。

因为

$$
\begin{aligned}
y_{t+1} &= y_{t+1} - y_t + y_t = \nabla y_{t+1} + y_t \\
&= (\nabla y_{t+1} - \nabla y_t) + \nabla y_t + y_t \quad , \\
&= \nabla^2 y_{t+1} + \nabla y_t + y_t
\end{aligned}
$$

同样，用$\nabla^2 y_{t+1}$的估计值代替$\nabla^2 y_{t+1}$得到式（8.23）。差分方法和指数平滑法的联合运用，除了能克服一次指数平滑法的滞后偏差之外，对初始值的问题也有显著的改进。因为数据经过差分处理后，所产生的新序列基本上是平稳的。这时，初始值取新序列的第一期数据对于未来预测值不会有多大影响。

其次，它拓展了指数平滑法的适用范围，但是，对于指数平滑法存在的加权系数α的选择问题，以及只能逐期预测问题，差分指数平滑模型也没有改进。

8.1.4 具有季节性特点的时间序列的预测

这里提到的季节，可以是自然季节，也可以是某种产品的销售季节等。显然，在现实的经济活动中，表现为季节性的时间序列是非常多的。比如，空调、季节性服装的生产与销售所产生的数据等。对于季节性时间序列的预测，要从数学上完全拟合其变化曲线是非常困难的。但预测的目的是为了找到时间序列的变化趋势，尽可能地做到精确。从这个意义上讲，可以有多种方法，下面介绍其中一种，即所谓季节系数法。季节系数法的具体计算步骤如下

（1）收集m年的每年各季度或者各月份（每年n个季度）的时间序列样本数据a_{ij}。其中，i表示年份的序号（$i=1,2,\cdots,m$），j表示季度或者月份的序号（$j=1,2,\cdots,n$）。

（2）计算每年所有季度或所有月份的算术平均值$\bar{a}$，即

$$\bar{a}=\frac{1}{k}\sum_{i=1}^{m}\sum_{j=1}^{n}a_{ij}，\ k=mn.$$

（3）计算同季度或同月份数据的算术平均值

$$\bar{a}_{\cdot j}=\frac{1}{m}\sum_{i=1}^{m}a_{ij}，\ j=1,2,\cdots,n。$$

（4）计算季度系数或月份系数$b_j = \bar{a}_{\cdot j} / \bar{a}$。

（5）预测计算。当时间序列是按季度列出时，先求出预测年份（下一年）的年加权平均

$$y_{m+1} = \frac{\sum_{i=1}^{m} w_i y_i}{\sum_{i=1}^{n} w_i},$$

式中，$y_i = \sum_{j=1}^{n} a_{ij}$为第$i$年的年合计数；$w_i$为第$i$年的权数，按自然数列取值，即$w_i = i$。再计算预测年份的季度平均值$\bar{y}_{m+1}$：$\bar{y}_{m+1} = y_{m+1} / n$。最后，预测年份第$j$季度的预测值为

$$y_{m+1,j} = b_j \bar{y}_{m+1}.$$

例 8.6　某商店某类商品 1999—2003 年各季度的销售额如表 8.7（单位：元）所示。试预测 2004 年各季度的销售额。

表 8.7　某商店某类商品 1999-2003 年各季度销售额

季度 年份	1	2	3	4
1999	137920	186742	274561	175433
2000	142814	198423	265419	183521
2001	131002	193987	247556	169847
2002	157436	200144	283002	194319
2003	149827	214301	276333	185204

8.2　平稳时间序列模型

这里的平稳是指宽平稳，其特性是序列的统计特性不随时间的平移而变化，即均值和协方差不随时间的平移而变化。

8.2.1 时间序列的基本概念

定义 8.1 给定随机过程$\{X_t, t \in T\}$。固定t，X_t是一个随机变量，设其均值为μ_t，当t变动时，此均值是t的函数，记为$\mu_t = E(X_t)$，称为随机过程的均值函数。

固定t，设X_t的方差为σ_t^2。当t变动时，这个方差也是t的函数，记为

$$\sigma_t^2 = \mathrm{Var}(X_t) = E[(X_t - \mu_t)^2].$$

称为随机过程的方差函数。方差函数的平方根σ_t称为随机过程的标准差函数，它表示随机过程X_t对于均值函数μ_t的偏离程度。

定义 8.2 对随机过程$\{X_t, t\in T\}$，取定$t,s\in T$，定义其自协方差函数为

$$\gamma_{t,s} = \mathrm{Cov}(X_t, X_s) = E[(X_t-\mu_t)(X_s-\mu_s)].$$

为刻画$\{X_t, t\in T\}$在时刻t与s之间的相关性,还可将$\gamma_{t,s}$标准化，即定义自相关函数

$$\rho_{t,s} = \frac{\gamma_{t,s}}{\sqrt{\gamma_{t,t}}\sqrt{\gamma_{s,s}}} = \frac{\gamma_{t,s}}{\sigma_t\sigma_s}.$$

因此，自相关函数$\rho_{t,s}$是标准化自协方差函数。

定义 8.3 设随机序列$\{X_t, t=0,\pm1,\pm2,\cdots\}$满足

（1）$E(X_t)=\mu=$常数；

（2）$\gamma_{t+k,t}=\gamma_k$（$k=0,\pm1,\pm2,\cdots$）与t无关。

则称X_t为平稳随机序列（平稳时间序列），简称平稳序列。

定义 8.4　设平稳序列$\{\varepsilon_t, t=0,\pm1,\pm2,\cdots\}$的自协方差函数$\gamma_k$是

$$\gamma_k = \sigma^2\delta_{k,0} = \begin{cases} 0, & k \neq 0, \\ \sigma^2, & k = 0, \end{cases}$$

其中$\delta_{k,0} = \begin{cases} 1, & k = 0 \\ 0, & k \neq 0 \end{cases}$，则称该序列为平稳白噪声序列。

平稳白噪声序列的方差是常数σ^2，因为$\gamma_k = 0(k \neq 0)$，则ε_t的任意两个不同时点之间是不相关的。平稳白噪声序列是一种最基本的平稳序列。

对平稳序列X_t的样本$X_1, X_2, \cdots, X_n$，可以用样本均值估计随机序列的均值

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} X_t = \bar{X}.$$

样本自协方差函数有如下两种形式

$$\hat{\gamma}_k = \frac{1}{n}\sum_{t=1}^{n-k}(X_{t+k} - \bar{X})(X_t - \bar{X}), \quad 0 \le k \le n-1,$$

或者

$$\hat{\gamma}_k^* = \frac{1}{n-k}\sum_{t=1}^{n-k}(X_{t+k} - \bar{X})(X_t - \bar{X}), \quad 0 \le k \le n-1,$$

且$\hat{\gamma}_{-k} = \hat{\gamma}_k$（$\hat{\gamma}_{-k}^* = \hat{\gamma}_k^*$）。

样本自相关函数

$$\hat{\rho}_k = \frac{\hat{\gamma}_k}{\hat{\gamma}_0}, \quad 0 \le k \le n-1,$$

或者

$$\hat{\rho}_k^* = \frac{\hat{\gamma}_k^*}{\hat{\gamma}_0}, \quad 0 \le k \le n-1.$$

定义 8.5　设$\{\varepsilon_t, t=0,\pm1,\pm2,\cdots\}$是零均值平稳白噪声，$\mathrm{Var}(\varepsilon_t)=\sigma_\varepsilon^2$。若$\{G_k, k=1,2,\cdots\}$是一数列，满足

$$\sum_{k=0}^{\infty}|G_k|<+\infty，\ G_0=1. \tag{8.24}$$

定义随机序列

$$X_t=\sum_{k=0}^{\infty}G_k\varepsilon_{t-k}, \tag{8.25}$$

则X_t称为随机线性序列。在条件（8.24）下，可证式（8.25）中的X_t是平稳序列。若零均值平稳序列X_t能表示为式（8.25）的形式，这种形式称为传递形式，$\{G_k, k=0,1,2,\cdots\}$称为 Green 函数。

定义 8.6 设$\{X_t, t=0,\pm1,\pm2,\cdots\}$是零均值平稳序列，从时间序列预报的角度引出偏相关函数的定义。如果已知$\{X_{t-1}, X_{t-2},\cdots, X_{t-k}\}$的值，要求对$X_t$做出预报。此时，可以考虑由$\{X_{t-1}, X_{t-2},\cdots, X_{t-k}\}$对$X_t$的线性最小均方估计，即选择系数$\phi_{k,1},\phi_{k,2},\cdots,\phi_{k,k}$，使得

$$\min \delta = E[(X_t - \sum_{j=1}^{k}\phi_{k,j}X_{t-j})^2].$$

将δ展开，得

$$\delta=\gamma_0-2\sum_{j=1}^{k}\phi_{k,j}\gamma_j+\sum_{j=1}^{k}\sum_{i=1}^{k}\phi_{k,j}\phi_{k,i}\gamma_{j-i},$$

令$\dfrac{\partial\delta}{\partial\phi_{k,j}}=0$，$j=1,2,\cdots,k$，得

$$-\gamma_j+\sum_{i=1}^{k}\phi_{k,i}\gamma_{j-i}=0,\quad j=1,2,\cdots,k.$$

两端同除γ_0并写成矩阵形式，可知$\phi_{k,j}$应满足下列线性方程组

$$\begin{bmatrix} 1 & \rho_1 & \cdots & \rho_{k-1} \\ \rho_1 & 1 & \cdots & \rho_{k-2} \\ \vdots & \vdots & & \vdots \\ \rho_{k-1} & \rho_{k-2} & \cdots & 1 \end{bmatrix} \begin{bmatrix} \phi_{k,1} \\ \phi_{k,2} \\ \vdots \\ \phi_{k,k} \end{bmatrix} = \begin{bmatrix} \rho_1 \\ \rho_2 \\ \vdots \\ \rho_k \end{bmatrix}. \tag{8.26}$$

式（8.26）称为 Yule－Walker 方程。我们称$\{\phi_{k,k}, k=1,2,\cdots\}$为$X_t$的偏相关函数。

下面介绍一种重要的平稳时间序列—ARMA 时间序列。ARMA 时间序列分为三种类型：

（1）AR 模型，即自回归序列（Auto Regressive Model）;

（2）MA 序列，即移动平均序列（Moving Average Model）;

（3）ARMA 序列，即自回归移动平均序列（Auto Regressive Moving Average Model）。

1．AR(p)序列

设$\{X_t, t=0,\pm1,\pm2,\cdots\}$是零均值平稳序列，满足下列模型

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \cdots + \phi_n X_{t-p} + \varepsilon_t \qquad (8.27)$$

其中ε_t是零均值、方差是σ_ε^2的平稳白噪声。则称X_t是阶数为p的自回归序列，简记为AR(p)序列，而

$$\phi = [\phi_1, \phi_2, \cdots, \phi_p]^T$$

称为自回归参数向量，其分量$\phi_j, j=1,2,\cdots,p$称为自回归系数。

引进后移算子对描述式（8.27）比较方便。算子$\boldsymbol{B}$定义如下

$$BX_t \equiv X_{t-1}, \quad B^k X_t \equiv X_{t-k}.$$

记算子多项式

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p,$$

则式（8.27）可以改写为

$$\phi(B) X_t = \varepsilon_t.$$

2．MA(q)序列

设$\{X_t, t=0,\pm1,\pm2,\cdots\}$是零均值平稳序列，满足下列模型

$$X_t=\varepsilon_t-\theta_1\varepsilon_{t-1}-\theta_2\varepsilon_{t-2}-\cdots-\theta_q\varepsilon_{t-q}, \qquad (8.28)$$

其中ε_t是零均值、方差是σ_ε^2的平稳白噪声，则称X_t是阶数为q的移动平均序列，简记为MA(q)序列，而

$$\theta=[\theta_1,\theta_2,\cdots,\theta_q]^T$$

称为移动平均参数向量，其分量$\theta_j, j=1,2,\cdots,q$称为移动平均系数。

在工程上，一个平稳白噪声发生器通过一个线性系统，如果其输出是白噪声的线性叠加，那么这一输出服从 MA 模型。

对于线性后移算子$\boldsymbol{B}$，有

$$\boldsymbol{B\varepsilon_t \equiv \varepsilon_{t-1}},\quad \boldsymbol{B^k\varepsilon_t \equiv \varepsilon_{t-k}},$$

再引进算子多项式

$$\boldsymbol{\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q},$$

则式（8.28）可以改写为

$$\boldsymbol{X_t = \theta(B)\varepsilon_t}.$$

3． ARMA(p,q)序列

设$\{X_t, t=0,\pm1,\pm2,\cdots\}$是零均值平稳序列，满足下列模型

$$X_t-\phi_1X_{t-1}-\cdots-\phi_pX_{t-p}=\varepsilon_t-\theta_1\varepsilon_{t-1}-\cdots-\theta_q\varepsilon_{t-q}, \tag{8.29}$$

其中ε_t是零均值、方差是σ_ε^2的平稳白噪声，则称X_t是阶数为 p,q 的自回归移动平均序列，简记为ARMA(p,q)序列。当$q=0$时，它是AR(p)序列；当$p=0$时，它为MA(q)序列。

应用算子多项式$\phi(B),\theta(B)$，式（8.29）可以写为

$$\phi(B)X_t=\theta(B)\varepsilon_t.$$

对于一般的平稳序列$\{X_t,t=0,\pm1,\pm2,\cdots\}$，设其均值$E(X_t)=\mu$，满足下列模型

$$(X_t-\mu)-\phi_1(X_{t-1}-\mu)-\cdots-\phi_p(X_{t-p}-\mu)$$

$$=\varepsilon_t-\theta_1\varepsilon_{t-1}-\cdots-\theta_q\varepsilon_{t-q}$$

（8.30）

其中ε_t是零均值、方差是σ_ε^2的平稳白噪声，利用后移算子$\phi(B),\theta(B)$，式（8.30）可表为

$$\phi(B)(X_t-\mu)=\theta(B)\varepsilon_t.$$

关于算子多项式$\phi(B),\theta(B)$，通常还要作下列假定

（1）$\phi(B)$和$\theta(B)$无公共因子，又$\phi_p \neq 0$，$\theta_q \neq 0$；

（2）$\phi(B)=0$的根全在单位圆外，这一条件称为模型的平稳性条件；

（3）$\theta(B)=0$的根全在单位圆外，这一条件称为模型的可逆性条件。

8.2.2 ARMA 模型的构建及预报

在实际问题建模中，首先要进行模型的识别与定阶，即要判断是**AR(*p*)**，**MA(*q*)**，**ARMA(*p*,*q*)**模型的类别，并估计阶数 p,q。其实，这都归结到模型的定阶问题。当模型定阶后，就要对模型参数 $\phi=[\phi_1,\phi_2,\cdots,\phi_p]^T$ 及 $\theta=[\theta_1,\theta_2,\cdots,\theta_q]^T$ 进行估计。定阶与参数估计完成后，还要对模型进行检验，即要检验 ε_t 是否为平稳白噪声。若检验获得通过，则 ARMA 时间序列的建模完成。作为时间序列建模之后的一个重要应用，我们还要讨论 ARMA 时间序列的预报。

1　ARMA 模型的构建

1）ARMA 模型定阶的 AIC 准则

AIC 准则又称 Akaike 信息准则，是由日本统计学家 Akaike 于 1974 年提出的。AIC 准则是信息论与统计学的重要研究成果，具有重要的意义。

$\mathbf{ARMA}(\boldsymbol{p},\boldsymbol{q})$序列 AIC 定阶准则为：选 $\boldsymbol{p},\boldsymbol{q}$，使得

$$\mathbf{min} \quad \mathbf{AIC} = \boldsymbol{n}\ln\hat{\boldsymbol{\sigma}}_{\varepsilon}^{2} + 2(\boldsymbol{p}+\boldsymbol{q}+1), \qquad (8.31)$$

其中，$\boldsymbol{n}$是样本容量，$\hat{\boldsymbol{\sigma}}_{\varepsilon}^{2}$是$\boldsymbol{\sigma}_{\varepsilon}^{2}$的估计与$\boldsymbol{p}$和$\boldsymbol{q}$有关。若当$\boldsymbol{p}=\hat{\boldsymbol{p}}$，$\boldsymbol{q}=\hat{\boldsymbol{q}}$时，式（8.31）达到最小值，则认为序列是$\mathbf{ARMA}(\hat{\boldsymbol{p}},\hat{\boldsymbol{q}})$。

当ARMA(p,q)序列含有未知均值参数μ时，模型为

$$\phi(B)(X_t-\mu)=\theta(B)\varepsilon_t,$$

这时，未知参数个数为$k=p+q+2$，AIC 准则为：选取p,q，使得

$$\min \quad \mathrm{AIC}=n\ln\hat{\sigma}_{\varepsilon}^{2}+2(p+q+2). \tag{8.32}$$

实际上, 式（8.31）与式（8.32）有相同的最小值点$\hat{p},\hat{q}$。

2）ARMA 模型的参数估计

ARMA 模型的参数估计有矩估计，逆函数估计法，最小二乘估计，条件最小二乘估计，最大似然估计等方法，这里我们就不给出各种估计的数学原理和参数估计表达式，直接使用 Matlab 工具箱给出相关的参数估计。

3）ARMA 模型检验的χ^2检验

若拟合模型的残差记为$\hat{\varepsilon}_t$，它是ε_t的估计。例如，对AR(p)序列，设未知参数的估计是$\hat{\phi}_1,\hat{\phi}_2,\cdots,\hat{\phi}_p$，则残差

$$\hat{\varepsilon}_t = X_t - \hat{\phi}_1 X_{t-1} - \cdots - \hat{\phi}_p X_{t-p}，\quad t = 1,2,\cdots,n,$$

（设$X_0 = X_{-1} = \cdots = X_{1-p} = 0$）.

记

$$\eta_k = \frac{\sum_{t=1}^{n-k} \hat{\varepsilon}_t \hat{\varepsilon}_{t+k}}{\sum_{t=1}^{n} \hat{\varepsilon}_t^2}, \quad k = 1,2\cdots,L,$$

其中L为$\hat{\varepsilon}_t$自相关函数的拖尾数，Ljung－Box 的χ^2检验统计量是

$$\chi^2 = n(n+2)\sum_{k=1}^{L} \frac{\eta_k^2}{n-k}. \tag{8.33}$$

检验的假设是

$$H_0: \rho_k = 0，当k \le L；H_1：对某些k \le L，\rho_k \ne 0.$$

在H_0成立时，若样本容量n充分大，χ^2近似于$\chi^2(L-r)$分布，其中r是估计的模型参数个数。

χ^2**检验法：**给定显著性水平α，查表得上α分位数$\chi^2_\alpha(L-r)$，则当$\chi^2 > \chi^2_\alpha(L)$时拒绝$H_0$，即认为$\varepsilon_t$非白噪声，模型检验未通过；而当$\chi^2 \le \chi^2_\alpha(L-r)$时，接受$H_0$，认为$\varepsilon_t$是白噪声，模型通过检验。

Matlab 中作 Ljung－Box 检验的函数为 lbqtest。

2 ARMA(p,q)序列的预报

时间序列的m步预报，是根据$\{X_k, X_{k-1}, \cdots\}$的取值对未来$k+m$时刻的随机变量$X_{k+m}(m>0)$做出估计。估计量记作$\hat{X}_k(m)$，它是$X_k, X_{k-1}, \cdots$的线性组合。

1）AR(p)序列的预报

AR(p)序列的预报递推公式

$$
\begin{cases}
\hat{X}_k(1)=\phi_1 X_k+\phi_2 X_{k-1}+\cdots+\phi_p X_{k-p+1}, \\
\hat{X}_k(2)=\phi_1 \hat{X}_k(1)+\phi_2 X_k+\cdots+\phi_p X_{k-p+2}, \\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
\hat{X}_k(p)=\phi_1 \hat{X}_k(p-1)+\phi_2 \hat{X}_k(p-2)+\cdots+\phi_{p-1}\hat{X}_k(1)+\phi_p X_k, \\
\hat{X}_k(m)=\phi_1 \hat{X}_k(m-1)+\phi_2 \hat{X}_k(m-2)+\cdots+\phi_p \hat{X}_k(m-p), m>p.
\end{cases}
$$

（8.34）

由此可见，$\hat{X}_k(m)(m\geq 1)$ 仅仅依赖于X_t的k时刻以前的p个时刻的值X_k, X_{k-1}，$\cdots, X_{k-p+1}$。这是AR(p)序列预报的特点。

2）MA(q)与ARMA(p,q)序列的预报

关于MA(q)序列$\{X_t, t=0,\pm1,\pm2,\cdots\}$的预报，有

$$\hat{X}_k(m)=0\text{，}\quad m>q.$$

因此，只需要讨论$\hat{X}_k(m)$，$m=1,2,\cdots,q$。为此，定义预报向量

$$\hat{X}_k^{(q)}=[\hat{X}_k(1),\hat{X}_k(2),\cdots,\hat{X}_k(q)]^T, \qquad (8.35)$$

所谓递推预报是求$\hat{X}_k^{(q)}$与$\hat{X}_{k+1}^{(q)}$的递推关系，对MA(q)序列，有

$$\hat{X}_{k+1}(1)=\theta_1\hat{X}_k(1)+\hat{X}_k(2)-\theta_1X_{k+1},$$

$$\hat{X}_{k+1}(2)=\theta_2\hat{X}_k(1)+\hat{X}_k(3)-\theta_2X_{k+1},$$

$$\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\hat{X}_{k+1}(q-1)=\theta_{q-1}\hat{X}_k(1)+\hat{X}_k(q)-\theta_{q-1}X_{k+1},$$

$$\hat{X}_{k+1}(q)=\theta_q\hat{X}_k(1)-\theta_qX_{k+1}.$$

从而得

$$\hat{X}_{k+1}^{(q)}=\begin{bmatrix}\theta_1 & 1 & 0 & \cdots & 0\\ \theta_2 & 0 & 1 & \cdots & 0\\ \vdots & \vdots & \vdots & & \vdots\\ \theta_{q-1} & 0 & 0 & \cdots & 1\\ \theta_q & 0 & 0 & 0 & 0\end{bmatrix}\hat{X}_k^{(q)}-\begin{bmatrix}\theta_1\\ \theta_2\\ \vdots\\ \theta_q\end{bmatrix}X_{k+1}. \tag{8.36}$$

递推初值可取$\hat{X}_{k_0}^{(q)}=0$（k_0较小）。因为模型的可逆性保证了递推式渐近稳定，即当n充分大后，初始误差的影响可以逐渐消失。

对于ARMA(p,q)序列，

$$\hat{X}_k(m)=\phi_1\hat{X}_k(m-1)+\phi_2\hat{X}_k(m-2)+\cdots+\phi_p\hat{X}_k(m-p)$$

，$m>p$.

因此，只需要知道$\hat{X}_k(1),\hat{X}_k(2),\cdots,\hat{X}_k(p)$，就可以递推算得$\hat{X}_k(m)$，$m>p$。仍定义预报向量（8.35）。令

$$\phi_j^*=\begin{cases}\phi_j, & j=1,2,\cdots,p,\\ 0, & j>p.\end{cases}$$

可证下列递推预报公式

$$\hat{X}_{k+1}^{(q)}=\begin{bmatrix} -G_1 & 1 & 0 & \cdots & 0 \\ -G_2 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ -G_{q-1} & 0 & 0 & \cdots & 1 \\ -G_q+\phi_q^* & \phi_{q-1}^* & \phi_{q-2}^* & \cdots & \phi_1^* \end{bmatrix}\hat{X}_k^{(q)}+\begin{bmatrix} G_1 \\ G_2 \\ \vdots \\ G_{q-1} \\ G_q \end{bmatrix}X_{k+1}+\begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ \sum_{j=q+1}^{p}\phi_j^* X_{k+q+1-j} \end{bmatrix} \tag{8.37}$$

这里G_j满足$X_t=\sum_{j=0}^{\infty}G_j\varepsilon_{t-j}$，式（8.37）中第三项当$p\le q$时为零。由可逆性条件保证，当$k_0$较小时，可令初值$\hat{X}_{k+1}^{(q)}=0$。

在实际中，模型参数是未知的。若已建立了时间序列的模型，则理论模型中的未知参数用其估计替代，再用上面介绍的方法进行预报。

8.3　时间序列的 Matlab 相关工具箱及命令

Matlab 时间序列的相关命令在系统辨识（system identification）工具箱，计量经济学（econometrics）工具箱，金融（financial）工具箱。

在系统辨识工具箱，有观察数据的获取命令idinput；数据预处理命令 detrend，idfilt，idresamp；模型结构的选择命令 struc,arxstruc,ivstruc 和 selstruc；参数估计命令 ar，arx，armax，ivx 等；模型预报与仿真的命令 compare，pe，predict，sim 等。

。

在计量经济学工具箱，有模型参数获取或设置命令 garchget，garchset；建模和仿真等命令 garchfit，garchpred，garchsim。

在金融工具箱，有构造时间序列数组的命令 fints，ascii2fts；用户图形界面解法命令 ftsgui，ftstool。

注：可以到 http://www.mathworks.com/help/网站下载相关工具箱用户使用手册的 pdf 帮助文件。

8.3.1 系统辨识工具箱相关命令的使用

例 8.7 随机产生时间序列

$$X_t + 0.6X_{t-1} + 0.2X_{t-2} = \varepsilon_t$$

的 10000 个观测值，其中的后 10 个数据见表 8.7，利用这 10000 个数据估计其模型参数，并预测第 10001, 10002 和 10003 个值。

表8.7 模拟观测数据

t	9991	9992	9993	9994	9995	9996	9997	9998	9999	10000
x_t	0.7752	0.2128	-1.2496	0.7664	-0.4751	0.9895	0.0470	1.6276	0.0101	-0.2287

注：表 8.7 中的数据是打开 Matlab 后第一次运行的数据，因为 Matlab 使用的是伪随机数，实际上是 Matlab 中已经排好序的数据。

解　首先利用Matlab软件辨识得到AR(2)模型

$$y_t = -0.5786 y_{t-1} - 0.1888 y_{t-2} + \varepsilon_t,$$

利用（8.34）式计算的3个预测值分别为

$$\hat{x}_{10001} = -0.5786 x_{10000} - 0.1888 x_{9999} = 0.1304,$$

$$\hat{x}_{10002} = -0.5786 \hat{x}_{10001} - 0.1888 x_{10000} = -0.0323,$$

$$\hat{x}_{10003} = -0.5786 \hat{x}_{10002} - 0.1888 \hat{x}_{10001} = -0.0060.$$

Matlab的计算结果显示为

$$A(q)y(t) = e(t),$$

$$A(q) = 1 + 0.5786\, q^{-1} + 0.1888\, q^{-2}.$$

例 8.8（续 8.7）　随机产生时间序列

$$X_t + 0.6X_{t-1} + 0.2X_{t-2} = \varepsilon_t$$

的 10000 个观测值，实际计算验证用 AIC 准则定阶的正确性。

解　　序列的真阶是 $p=2$。分别用 **AR(1)**，**AR(2)**，**AR(3)** 拟合，（由于是随机模拟，每次的运行结果都是不一样的）其中的一次运行结果的AIC值分别为0.0083，-0.0325，-0.0324，显示应为 **AR(2)** 序列。

例8.9（续例8.7） 随机产生时间序列 $X_t + 0.6X_{t-1} + 0.2X_{t-2} = \varepsilon_t$ 的10000个观测值，利用这10000个数据估计其模型参数，并用 χ^2 检验法进行模型检验。

检验结果为h=0，模型通过了检验。

例8.10　随机产生时间序列 $X_t=\varepsilon_t-0.6\varepsilon_{t-1}-0.2\varepsilon_{t-2}$ 的10000个观测值，其中的后10个数据见表8.8，利用这10000个数据估计其模型参数。

表8.8　模拟观测数据

t	9991	9992	9993	9994	9995	9996	9997	9998	9999	10000
x_t	0.0803	0.3235	-0.9669	0.0592	-0.2652	0.8577	0.5457	1.8538	0.9961	0.1029

解　利用Matlab软件求得的MA(2)模型为

$$y_t = \varepsilon_t - 0.5773\varepsilon_{t-1} - 0.2195\varepsilon_{t-2}$$

例 8.11　模拟产生下列时间序列的 10000 个观测数据，

$$X_t - 0.8X_{t-1} = \varepsilon_t - 0.4\varepsilon_{t-1}, \quad \varepsilon_t \sim N(0,1),$$

建立适当的ARMA(p,q)模型，并用χ^2检验法进行模型检验。

解　我们通过枚举法利用 AIC 准则，来对模型定阶，选取 AIC 取值最小的$\boldsymbol{p,q}$作为模型的阶数。由于是随机模拟，每一次的运行结果都是不一样的，我们这里就不给出具体的答案。

8.3.2　计量经济学工具箱相关命令的使用

本节主要介绍一下 GARCH 模型，GARCH 表示广义自回归条件异方差（Generalized Auto Regressive Conditional Heteroscedasticity）。GARCH 模型分为均值方程与方差方程两部分。

均值方程形式如下

$$y_t = C + \sum_{i=1}^{R} \varphi_i y_{t-i} + \varepsilon_t + \sum_{j=1}^{M} \theta_j \varepsilon_{t-j} + \sum_{k=1}^{N_x} \beta_k X(t,k), \tag{3.38}$$

其中$\{\varphi_i\}$是自回归系数，$\{\theta_j\}$是移动平均系数；X是解释回归矩阵，它的每一个列是一个时间序列，$X(t,k)$表示该矩阵第t行第k列的数据。

GARCH(P,Q)模型方差方程形式如下

$$\sigma_t^2 = K + \sum_{i=1}^{P} G_i \sigma_{t-i}^2 + \sum_{j=1}^{Q} A_j \varepsilon_{t-j}^2 , \tag{8.39}$$

其中系数满足下列约束条件

$$\sum_{i=1}^{P} G_i + \sum_{j=1}^{Q} A_j < 1 ,$$

$K > 0$，$G_i > 0$，$A_j > 0$，$i = 1, 2, \cdots, P$；$j = 1, 2, \cdots, Q$.

在工具箱中可以通过命令 garch 指定模型的结构，garch 的语法为

model = garch(P,Q)

model = garch(Name,Value)

其中 Name 为属性，Value 是对应的属性值。

在工具箱中可以通过命令 estimate 对模型中的参数进行估计，estimate 的语法为

EstMdl = estimate(Mdl,y)

其中输入参数 Mdl 指定模型的结构，y 为时间序列的样本观测值。输出参数 EstMdl 是模型的参数估计值。

例 8.12　模拟产生下列时间序列的 10000 个观测数据，

均值方程 $y_t = 0.8y_{t-1} + \varepsilon_t + 0.4\varepsilon_{t-1}$，　（8.40）

方差方程 $\sigma_t^2 = 0.01 + 0.2\sigma_{t-1}^2 + 0.3\varepsilon_{t-1}^2$，　（8.41）

利用 GARCH 模型估计模型的参数，用 χ^2 检验法进行模型检验，并预测第 10001, 10002 和 10003 个值。

解　对比（8.38）式和（8.40）式，（8.39）式和（8.41）式，属性参数$C=0$，$AR=0.8$，$MA=0.4$，$K=1$，$GARCH=0.2$，$ARCH=0.3$。由于是随机模拟，每一次运行的预测值都是不一样的。我们这里就不给出计算结果了。

例 8.13(续例 8.7)　使用例 8.7 的数据建立 GARCH 预测模型。

解　这里不对模型进行定阶，只拟合一个GARCH模型，说明相关命令的用法。计算结果为

$$y_t = 0.0016 - 0.4862 y_t + \varepsilon_t,$$

$$\sigma_t^2 = 0.5329 + 0.4607\sigma_{t-1}^2 + 0.0160\varepsilon_{t-1}^2.$$

3个预测值分别为

$$\hat{x}_{10001} = 0.1128,\quad \hat{x}_{10002} = -0.0532,\quad \hat{x}_{10003} = 0.0275.$$

可以看出 3 个预测值与例 8.7 的预测值有差异，其原因是模型的结构不同。

8.3.3　金融工具箱的使用

1　金融工具箱的相关命令

金融工具箱有时间序列数组的创建命令 fints，把 ASCII 文件的内容保存为 Matlab 中的 fints 型时间序列变量的命令 ascii2fts。

例 8.14　利用 fints 函数创立日期型数组。

该工具箱还有把fints型数据转化为矩阵数据的函数fts2mat，抽出时间序列某些数据的命令extfield，合并多个时间序列的命令merge，计算相关系数的命令corrcoef，计算偏相关系数的命令parcorr，计算自相关系数的命令autocorr等。

2　金融工具箱的用户图形界面命令

用户图形界面解法有两个命令ftsgui，ftstool。

1. ftsgui用户图形界面功能介绍

在Matlab命令窗口输入如下代码

```
>> ftsgui
```

时间序列主窗口显示如图8.1所示。

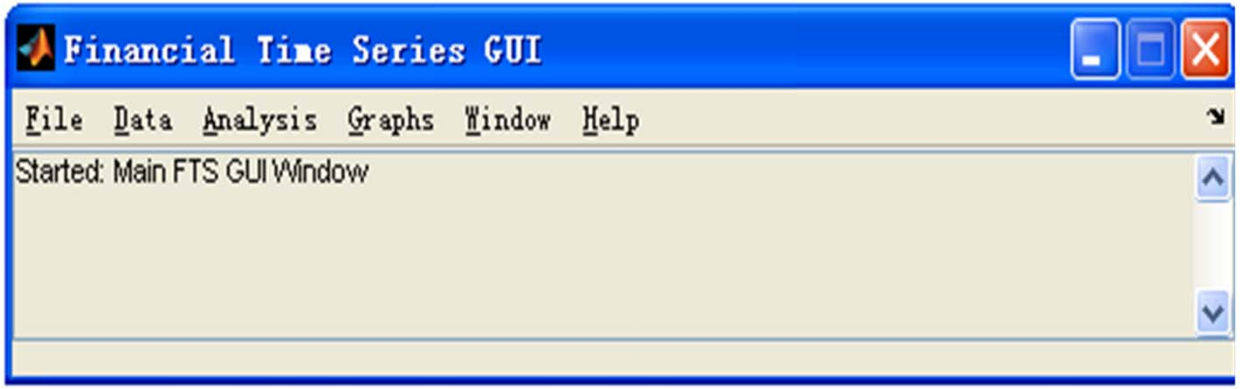

图 8.1　MATLAB 时

时间序列用户图形界面

主窗口有6个菜单选项，分别是File、Data、Analysis、Graphs、Windows、Help。举例来说，Matlab自带有迪斯尼公司的股价数据文件，数据文件名是disney.mat，其中含有迪斯尼股价的时间、开盘价、最高价、最低价、收盘价等内容，而且有缺失的数据，缺失的数据内容为NaN符号。

在\matlab\R2014b\toolbox\finance\findemos可以找到该文件。调入disney.mat文件后，同时还出现另外3个窗口。这是因为disney文件中含有3个fints型的数组dis，dis_nv和q_dis，3个窗口分别对应于不同的数组。

2. ftstool用户图形界面功能介绍

ftstool 用户图形界面可以实现数据获取，转换，合并和画图的一些功能。

8.4 ARIMA 序列与季节性序列

前面介绍了平稳时间序列的建模与预报，在实际中遇到的时间序列往往有三个特性：趋势性、季节性与非平稳性。本节主要采用 Box－Jenkins 方法，即差分方法，有时还要用时间序列的变换方法，消除其趋势性、季节性，使得变换后的序列是平稳序列，并假设为 ARMA 序列，再用上面介绍的方法去研究。

8.4.1　ARIMA 序列及其预报

我们先看一个例子。考虑研究时间序列 $\{X_t, t=0,\pm1,\pm2,\cdots\}$，满足

$$X_t - 1.5X_{t-1} + 0.5X_{t-2} = \varepsilon_t, \qquad (8.42)$$

改写为 $\phi(B)X_t = \varepsilon_t$，其中 $\phi(B) = 1 - 1.5B + 0.5B^2$。$\phi(B) = 0$ 的根是 $B_1 = 1$，$B_2 = 2$，其中 $B_1 = 1$ 在单位圆周上，并非在单位圆外，即原序列非平稳，因而不是 AR(2) 序列。但若改写式（8.42）为

$$(X_t - X_{t-1}) - 0.5(X_{t-1} - X_{t-2}) = \varepsilon_t,$$

令

$$\nabla \boldsymbol{X}_t = \boldsymbol{X}_t - \boldsymbol{X}_{t-1} = \boldsymbol{W}_t, \tag{8.43}$$

则有

$$\boldsymbol{W}_t - \mathbf{0.5}\boldsymbol{W}_{t-1} = \boldsymbol{\varepsilon}_t,$$

$\boldsymbol{W}_t$是**AR(1)**序列。式（8.43）定义的运算∇称为 1 阶向后差分运算。经过这样的 1 阶差分运算，原来的非平稳序列$\boldsymbol{X}_t$转化为平稳序列$\boldsymbol{W}_t$。

由上例可见，差分运算可以使一类非平稳序列（即带有趋势性的序列）平稳化。如果 1 阶差分还不能使时间序列平稳化，还可以进行 2 阶差分，3 阶差分，直至第d阶差分，最后将序列化为平稳序列。

1 阶差分

$$\nabla X_t = X_t - X_{t-1} = (1-B)X_t,$$

2 阶差分

$$\nabla^2 X_t = X_t - 2X_{t-1} + X_{t-2} = (1-B)^2 X_t,$$

一般地，d阶差分

$$\nabla^d X_t = (1-B)^d X_t,$$

其中∇^d称为d阶差分算子，

$$\nabla^d \equiv (1-B)^d = 1 - \binom{d}{1}B + \binom{d}{2}B^2 + \cdots + (-1)^{d-1}\binom{d}{d-1}B$$

.

设$\{X_t, t=0,\pm1,\pm2,\cdots\}$是非平稳序列。若存在正整数$d$，使得$\nabla^d X_t = W_t$，而$\{W_t, t=0,\pm1,\pm2,\cdots\}$是$\mathbf{ARMA}(p,q)$序列，则称$X_t$是$\mathbf{ARIMA}(p,d,q)$序列。这时，$X_t$满足

$$\phi(B)\nabla^d X_t = \theta(B)\varepsilon_t. \tag{8.44}$$

若$\nabla^d X_t$为平稳序列，但均值$\mu \neq 0$，则$\nabla^d X_t - \mu$为平稳零均值序列，满足

$$\phi(B)(\nabla^d X_t - \mu) = \theta(B)\varepsilon_t，\ t > d， \tag{8.45}$$

此时，称X_t为一般$\mathbf{ARIMA}(p,d,q)$序列，若μ未知，可用$\nabla^d X_t$的平均值$\bar{x}$估计。

若X_t的观测样本是$X_1, X_2, \cdots, X_n$，经过 1 阶差分后，数据减少为$n-1$个；2 阶差分以后，数据为$n-2$个；一般地，d阶差分以后，数据为$n-d$个。由d阶差分$\nabla^d X_t$复原数据，需要给定初值$X_1, X_2, \cdots, X_d$。

在确定模型时，往往采用下面的方法。先对X_t的样本$X_1,X_2,\cdots,X_n$，计算样本自相关函数与样本偏相关函数。如果是截尾的或者是拖尾的（即被负指数控制的），说明已服从 ARMA 模型。若自相关函数与偏相关函数至少有 1 个不是截尾的或拖尾的，说明X_t不是平稳的，可以作 1 阶差分$\nabla X_t, t=2,3,\cdots,n$，并求其样本自相关函数与样本偏相关函数，再用上述方法讨论。这样，直至判断$\nabla^d X_t$是平稳序列为止。在实际计算中，若遇到样本自相关函数或样本偏相关函数的图形虽然下降，但下降很慢，应认为是非平稳序列，需作差分运算。

若初值$X_1, X_2, \cdots, X_d$已知，由

$$W_t = \nabla^d X_t, \quad t = d+1, \cdots, n,$$

可以复原X_t。给出$d = 1,2$时的复原公式

（1）$d = 1$

$$X_t = X_1 + \sum_{j=1}^{t-1} W_{j+1} = X_k + \sum_{j=1}^{t-k} W_{j+k}, \quad t > k \geq 1. \tag{8.46}$$

（2）$d=2$

$$X_t = X_2 + (t-2)(X_2 - X_1) + \sum_{j=1}^{t-2}(t-j-1)W_{j+2}$$

$$= X_k + (t-k)(X_k - X_{k-1}) + \sum_{j=1}^{t-k}(t-j-k+1)W_{j+k}$$

$$t > k \geq 2. \tag{8.47}$$

设X_t是ARIMA(p,d,q)序列，则当$p=0$时，称为IMA(d,q)序列；当$q=0$时，称为ARI(p,d)序列。

设 X_t 是 $\mathrm{ARIMA}(p,d,q)$ 序列，则当 $p=0$ 时，称为 $\mathrm{IMA}(d,q)$ 序列；当 $q=0$ 时，称为 $\mathrm{ARI}(p,d)$ 序列。

下面简单介绍 ARIMA 序列的预报。

设$\{X_t, t=0,\pm1,\pm2,\cdots\}$是ARIMA$(p,d,q)$序列，仅讨论$d=1$与$d=2$的情形，这在实用中是最常见的。

（1）当$d=1$时，$\nabla X_t = W_t$，有

$$\nabla \hat{X}_k(m) = \hat{W}_k(m),$$

即

$$\hat{X}_k(m) - \hat{X}_k(m-1) = \hat{W}_k(m).$$

由此得

$$\hat{X}_k(m) = \hat{X}_k(m-1) + \hat{W}_k(m) = X_k + \sum_{j=1}^{m} \hat{W}_k(j). \tag{8.48}$$

（2）当$d=2$时，$\nabla^2\hat{X}_k(m)=\hat{W}_k(m)$，即

$$\hat{X}_k(m)-2\hat{X}_k(m-1)+\hat{X}_k(m-2)=\hat{W}_k(m).$$

复原$\hat{X}_k(m)$，可得

$$\hat{X}_k(m)=X_k+m(X_k-X_{k-1})+\sum_{j=1}^{m}(m+1-j)\hat{W}_k(j).$$

（8.49）

下面用实例说明如何实现时间序列的建模与预报。

例 8.15　某化工生产过程每 2 小时的浓度读数(逐行排列）如表 8.11 所示。

（1）对这一生产过程建模;

（2）对这一生产过程进行 10 步预测。

表 8.11　化工生产过程浓度数据

17.0	16.6	16.3	16.1	17.1	16.9	16.8	17.4	17.1	17.0
16.7	17.4	17.2	17.4	17.4	17.0	17.3	17.2	17.4	16.8
17.1	17.4	17.4	17.5	17.4	17.6	17.4	17.3	17.0	17.8
17.5	18.1	17.5	17.4	17.4	17.1	17.6	17.7	17.4	17.8
17.6	17.5	16.5	17.8	17.3	17.3	17.1	17.4	16.9	17.3
17.6	16.9	16.7	16.8	16.8	17.2	16.8	17.6	17.2	16.6
17.1	16.9	16.6	18.0	17.2	17.3	17.0	16.9	17.3	16.8
17.3	17.4	17.7	16.8	16.9	17.0	16.9	17.0	16.6	16.7
16.8	16.7	16.4	16.5	16.4	16.6	16.5	16.7	16.4	16.4
16.2	16.4	16.3	16.4	17.0	16.9	17.1	17.1	16.7	16.9
16.5	17.2	16.4	17.0	17.0	16.7	16.2	16.6	16.9	16.5
16.6	16.6	17.0	17.1	17.1	16.7	16.8	16.3	16.6	16.8
16.9	17.1	16.8	17.0	17.2	17.3	17.2	17.3	17.2	17.2
17.5	16.9	16.9	16.9	17.0	16.5	16.7	16.8	16.7	16.7
16.6	16.5	17.0	16.7	16.7	16.9	17.4	17.1	17.0	16.8
17.2	17.2	17.4	17.2	16.9	16.8	17.0	17.4	17.2	17.2
17.1	17.1	17.1	17.4	17.2	16.9	16.9	17.0	16.7	16.9
17.3	17.8	17.8	17.6	17.5	17.0	16.9	17.1	17.2	17.4
17.5	17.9	17.0	17.0	17.0	17.2	17.3	17.4	17.4	17.0
18.0	18.2	17.6	17.8	17.7	17.2	17.4			

解　通过计算自相关函数和偏相关函数，确定取 $d=1$。利用 AIC 和 BIC 准则定阶，取 $\mathbf{ARIMA(1,1,1)}$ 模型。参数的估计为

$$\hat{\phi}_1 = 0.2425，\hat{\theta}_1 = -0.8383.$$

得到模型为

$$(1-0.2425B)(1-B)X_t = (1-0.8383B)\varepsilon_t.$$

（2）经计算，其 10 步预报值见表 8.12。

表 8.12　10 步预报值

步数	**1**	**2**	**3**	**4**	**5**
预报值	**17.4717**	**17.4891**	**17.4933**	**17.4943**	**17.4946**
步数	**6**	**7**	**8**	**9**	**10**
预报值	**17.4946**	**17.4946**	**17.4947**	**17.4947**	**17.4947**

注：这一问题也可取简化模型IMA(1,1)

$$(1-B)X_t = 0.0034 + (1-0.7062B)\varepsilon_t.$$

这时，其 10 步于预报值及预报标准差见表 8.13。

表 8.13　10 步预报值（IMA(1,1)）

步数	1	2	3	4	5
预报值	17.5174	17.5208	17.5243	17.5277	17.5312
步数	6	7	8	9	10
预报值	17.5346	17.5381	17.5415	17.5450	17.5484

8.4.2　季节性序列及其预报

在不少实际问题中，时间序列有很明显的周期规律性，例如气温、雨量、用电量等。由季节性因素或其它周期因素引起的周期性变化的时间序列，我们称为季节性时间序列，相应的模型为季节性模型。例如，以电力负荷为例，如果X_t表示时刻t的用电负荷量，以1小时为采样间隔，显然X_t将包含24小时的周期性变化。一般，上午有一个用电高峰，晚上又有一个用电高峰；中午有一个用电较少的低谷，深夜则出现一天用点量最少的低谷。

一般地，对周期s的序列，可先进行差分运算

$$\nabla_s X_t = (1 - B^s) X_t,$$

$$\nabla_s^d = (1 - B^s)^d X_t,$$

然后再进行 ARIMA 建模。

例 8.16　测得某地区一口井 7 年的地下水埋深数据如表 8.14 所示。试预报第 8 年全年的地下水埋深。

表 8.14　井水埋深数据

年份 \ 月份	1	2	3	4	5	6	7	8	9	10	11	12
1	9.40	8.81	8.65	10.01	11.07	11.54	12.73	12.43	11.64	11.39	11.1	10.85
2	10.71	10.24	8.48	9.88	10.31	10.53	9.55	6.51	7.75	7.8	5.96	5.21
3	6.39	6.38	6.51	7.14	7.26	8.49	9.39	9.71	9.65	9.26	8.84	8.29
4	7.21	6.93	7.21	7.82	8.57	9.59	8.77	8.61	8.94	8.4	8.35	7.95
5	7.66	7.68	7.85	8.53	9.38	10.09	10.59	10.83	10.49	9.21	8.66	8.39
6	8.27	8.14	8.71	10.43	11.47	11.73	11.61	11.93	11.55	11.35	11.11	10.49
7	10.16	9.96	10.47	11.70	10.1	10.37	12.47	11.91	10.83	10.64	10.29	10.34

解　（1）首先进行时间序列模型定阶。因为数据有下降趋势，又有 12 个月的季节性，故对数据作下列差分运算

$$\boldsymbol{W_t = \nabla\nabla_{12} X_t}.$$

对$\boldsymbol{W_t}$进行 ARMA 模型拟合。用选取的$\boldsymbol{p,q}$的各种阶数进行试算，用 AIC 和 BIC 准则进行定阶，确定选取$\boldsymbol{p = 1, q = 1}$。

（2）建立模型并进行预测。得到的模型为

$$(1+0.7020B)(1-B)(1-B^{12})X_t=(1+0.9882B)\varepsilon_t,$$

算得第 8 年全年预报值如表 8.15 所示。

表 8.15　12 个月地下水埋深预报值

步数	1	2	3	4	5	6
预报值	**10.3719**	**9.9178**	**10.6062**	**11.7110**	**10.1989**	**10.4072**
步数	**7**	**8**	**9**	**10**	**11**	**12**
预报值	**12.5505**	**11.9601**	**10.9014**	**10.6964**	**10.3570**	**10.3996**