Ch2 自回归移动平均模型

徐剑刚

自回归移动平均模型

- 时间序列分析方法是Box and Jenkins (1970)提出的，该法不考虑以经济或金融理论为依据的解释变量的作用，而是依据时间序列本身的变化规律，利用外推机制来描述时间序列。
- 必须注意的是，建立时间序列模型的前提是：*时间序列是平稳的*。

随机过程

- 由随机变量构成的一个有序序列称为随机过程，通常记为 $\{x(s,t), s\in S, t\in T\}$ S是样本空间，T为序数集。
- 对于每个t ($t\in T$)，$x(\bullet, t)$是样本空间S中的随机变量；
- 对于每个s($s\in S$)，$x(s,\bullet)$是随机过程在序数集T中的一次实现。一般将随机过程简称为过程，记为$\{x_t\}$或x_t 。
- 随机过程的一次观测结果称为时间序列，$\{x_t, t\in T\}$用表示。时间序列数据是所要研究变量的观测值按时间先后顺序排列的一组数据。如果我们把1997年1月1日至2002年12月31日间每个交易日收盘时的中信指数按时间先后排列起来，得到了中信指数时间序列。
- 通常，分析的数据是等时间间隔的，从而是一个离散的时间序列。
- 研究时间序列$\{x_t\}$的目的，就是分析x_t与其过去值$\{x_{t-1}, x_{t-2},\ldots\}$间的动态相关性。如果用线性模型分析，意味着$x_t$与其过去值$\{x_{t-1}, x_{t-2},\ldots\}$存在着线性关系。

滞后算子

- 滞后算子“*L*”是这样定义的 $Lx_t = x_{t-1}$
- Lx_t就是时间序列$\{x_t\}$在第t-1时刻的值x_{t-1}
- 一个滞后算子的多项式为 $\phi(L) = \phi_0 - \phi_1 L - \cdots \phi_p L^p = \phi_0 - \sum_{i=1}^{p} \phi_i L^i$

- 其中，ϕ_0=1，p是非负整数，为$\phi(L)$的阶数。将$\phi(L)$应用于序列x_t上，得 $\phi(L)x_t = x_t - \phi_1 x_{t-1} - \cdots \phi_p x_{t-1} = x_t - \sum_{i=1}^{p} \phi_i x_{t-i}$
- 在时间序列分析中，该方程常用来分析x_t与其过去值$\{x_{t-1}, x_{t-2}, \ldots\}$间的动态相关性。

滞后算子

- 假定c为常数，方程 $\phi(L)x_t=c$
- 称为p阶差分方程。如果c=0，那么，方程就是一个齐次方程。如果变量x_t满足差分方程(6.4)，称为方程的一个解。
- 不同的$\phi(L)$将描述x_t的不同的动态行为，常用$\phi(L)x_t=c$
- 差分方程来分析一个线性时间序列的动态结构。

平稳性

- 一个时间序列是随机变量按时间顺序排列的观测值，在经济和金融的应用中，我们仅能得到的是时间序列的一次实现，时间序列分析的目标就是从观测到的一次实现来对过程进行推断，常用的方法就是选择一个适当的模型来近似描述所研究的过程。
- 选择一个适当的模型，就涉及到评价样本数据的联合分布函数 $F(x_1, x_2, \cdots x_T) = \Pr(X_1 \le x_1, \cdots, X_T \le x_T)$
- 其中，T是样本容量，x_i是实数。通常$\{x_t\}$是一个观测序列。为了能更好地为时间序列构模，需要限制联合分布。进一步，为了预测，还要说明过程分布的一些关键性质，即时间不变性。

强平稳

- 在时间序列分析中，时间不变性是十分有用的，最常用的就是平稳性。
- 如果一个平稳过程的性质不随时间起点的变化而变化，也就是说，对于序数集*T*中的任何时间子集 $(t_1,t_2,\ldots,t_n)$
- 以及任何实数*k*，$(t_i+k)\in T, i=1,2,\ldots n$

$$F\left(x_{t_1},\ldots x_{t_n}\right)=F\left(x_{t_1+k},\ldots x_{t_n+k}\right)$$

- 称这个随机过程为强平稳过程。其中，*F*(•)表示*n*个随机变量的联合分布函数，这意味着该平稳过程所有存在的矩都不随时间的变化而变化。
- 强平稳表明了 x_{t_1} 和 x_{t_1+k} 的概率分布相同，
- $\{x_{t_1},x_{t_2}\}$ 的联合分布和 $\{x_{t_1+k},x_{t_2+k}\}$的联合分布相同，…，
- $\{x_{t_1},x_{t_2},\cdots x_{t_n}\}$ 的联合分布和 $\{x_{t_1},x_{t_2},\cdots x_{t_n}\}$ 的联合分布相同。

*m*阶平稳过程

- 强平稳的要求苛刻，因而引入较弱的条件
- 如果一个平稳过程*m*阶以下矩(包括*m*阶矩)的取值与时间无关，称随机过程为*m*阶平稳过程。
- 随机过程为*m*阶平稳过程并不要求 x_{t_1} 和 x_{t_1+k} 的概率分布相同，仅要求这两个分布的主要特征相同，只要求相等到*m*阶矩。

二阶平稳(弱平稳、协方差平稳）

- 只注重时间序列的一阶矩、二阶矩。
- 假设一个时间序列 $\{x_t\}_1^T$ 其T个均值为$E(x_1), E(x_2),\ldots, E(x_T)$，$T$个方差为$Var(x_1), Var(x_2),\ldots, Var(x_T)$，和$T(T\text{-}1)/2$个协方差为$Cov(x_i, x_j)$，$i\neq j$。
- 如果 $E(x_1)=\ldots=E(x_T)=E(x_t)=\mu<\infty$ $\quad Cov(x_i,x_j)=\sigma_{ij}<\infty$

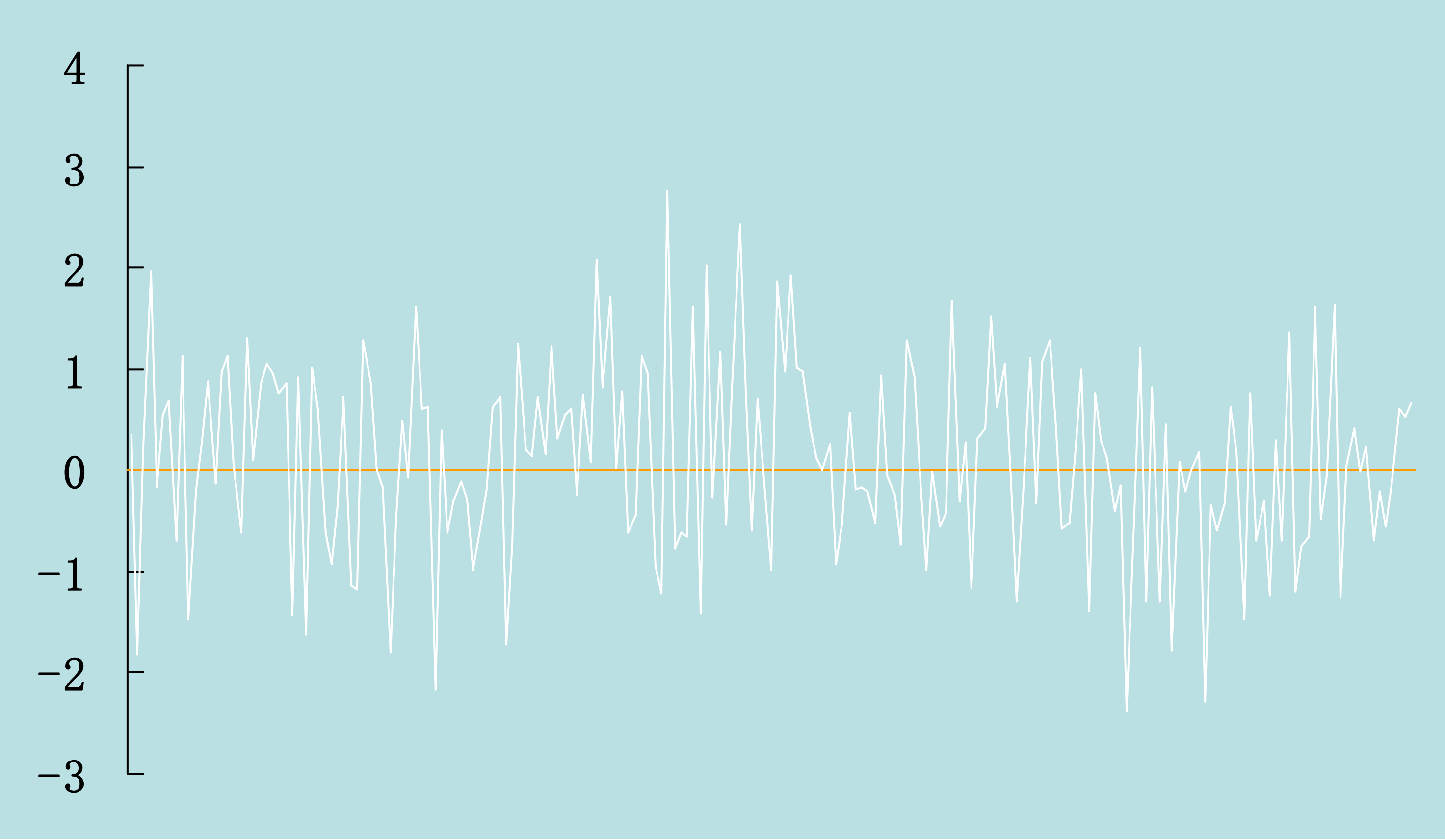
$$Var(x_1)=\ldots=Var(x_T)=Var(x_t)=\sigma^2<\infty$$

- 均与时间t无关，称x_t为二阶平稳过程，弱平稳、协方差平稳。
- 如果时间序列x_t的一阶矩、二阶矩具有时间不变性，那么，x_t是弱平稳的。当然，这里要求x_t的一阶矩、二阶矩都存在。
- 强平稳意味着过程的分布与时间无关，弱平稳意味着过程的二阶矩与时间无关。强平稳过程也是弱平稳的。
- 实际应用时，通常假设时间序列的分布是联合正态分布，这种假设出于统计上的方便性。因为，正态分布性质能为均值和二阶矩描述。
- 对于服从正态分布的时间序列，弱平稳就是强平稳。

白噪声

- 在二阶平稳过程中，白噪声序列$\{a_t\}$，其定义如下，
- (1)均值为0，即对于所有的t， $E(a_t)=0$
- (2)方差是常数，即对于所有的t， $E(a_t^2)=\sigma^2$
- (3)协方差为0，即对于$t \neq s$， $E(a_t a_s)=0$
- 也就是说，白噪声是均值为0、方差为σ^2的不相关序列。
- 白噪声相当于没有“记忆”过程，即过程第t时刻的值与所有过去直到t-1时刻的值(实际上也包括过程的未来值)都不相关。
- 白噪声过程滞后k期的自相关系数为0。应该指出的是，白噪声过程是人为的，在实际中过程的前后往往都存在着“记忆”。但是，白噪声为构造更复杂的模型提供了基本“元素”，因此，它在平稳过程理论中起着十分重要的作用。

白噪声过程的一次实现



自协方差函数和自相关函数

- $\gamma_k = Cov(x_t, x_{t-k})$ 称为$\{x_t\}$的自协方差函数。
- 对于每个*k*，γ_k是过程在相隔时间为*k*的一对值的协方差，称*k*为滞后阶数。
- 注意，γ_k只是*k*的函数，与观测值的时期*t*无关。因而，对于一个弱平稳过程来说，$\gamma_k = \gamma_{-k}$,
- 证明如下，将*t*换成*t+k*，有

$$\gamma_k = Cov(x_t, x_{t-k}) = Cov(x_{t-k}, x_t) = Cov(x_{t+k-k}, x_{t+k}) = Cov(x_t, x_{t+k}) = \gamma_{-k}$$

- 自相关函数(ACF)定义为，$\rho_k = \gamma_k / \gamma_0$
- 对于每个*k*，ρ_k是过程在相隔时间为*k*的一对值的相关关系，它可以作为x_t的一次实现与时移*k*后的同一次实现之间"相似"的度量。
- 注意，ρ_k只是*k*的函数，与观测值的时期*t*无关，而且，$\rho_k = \rho_{-k}$。自相关函数在建立自回归移动平均模型时非常重要。

线性时间序列

- 如果要生成一个不独立的序列，我们可以利用白噪声的线性组合，

$$x_t = a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \ldots = \sum_{j=0}^{\infty} \psi_j a_{t-j}$$

- 其中，ψ_0=1，{a_t}为白噪声序列。
- 实际上，上式将白噪声{a_t}的现行值和过去值的平均而生成观测序列x_t。依这种方式生成的过程称为线性过程，实际上是移动平均过程。
- 上式x_t的均值为0。如果x_t的均值不为0，只要在式的右边加上常数项μ即可。
- 如果x_t是弱平稳过程，那么其方差必存在，这样要求

$$\sum_{i=1}^{\infty} \psi_i^2 < \infty$$

x_t的一阶矩、二阶矩和自相关函数

- x_t的一阶矩即均值为0，$E(x_t)=0$，
- 方差为

$$\gamma_0 = E(x_t^2) = E(a_t + \psi_1 a_{t-1} + \ldots)^2$$
$$= (1+\psi_1^2+\psi_2^2\ldots)\sigma^2 = \sigma^2\sum_{j=0}^{\infty}\psi_j^2$$

- 自协方差函数为

$$\gamma_k = E(x_t x_{t-k}) = E(a_t + \psi_1 a_{t-1} + \ldots + \psi_k a_{t-k} + \ldots)(a_{t-k} + \psi_1 a_{t-k-1} + \ldots)$$
$$= \psi_k E(a_{t-k}^2) + \psi_{k+1}\psi_1 E(a_{t-k-1}^2) + \cdots = \sigma^2\sum_{k=0}^{\infty}\psi_j\psi_{j+k}$$

- 自相关函数为

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \frac{\sum_{j=0}^{\infty}\psi_j\psi_{j+k}}{\sum_{j=0}^{\infty}\psi_j}$$

沃尔德分解

- 任何线性时间序列模型都可表示为无限阶的移动平均过程，只不过不同的模型，对ψ_j权重的限制不同。
- 下面我们考虑时间序列分析中一个非常重要的定理，沃尔德(Wold, 1938)定理。沃尔德定理指任何平稳过程y_t可分解为两部分，

$$y_t = \mu_t + x_t$$

- 其中，x_t是线性过程，μ_t是确定性过程，
- 对于所有的t, s，μ_s和x_t不相关。确定性过程μ_t可由过程的过去值完全预测。

一阶自回归过程

$$x_t = a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \ldots = \sum_{j=0}^{\infty} \psi_j a_{t-j}$$

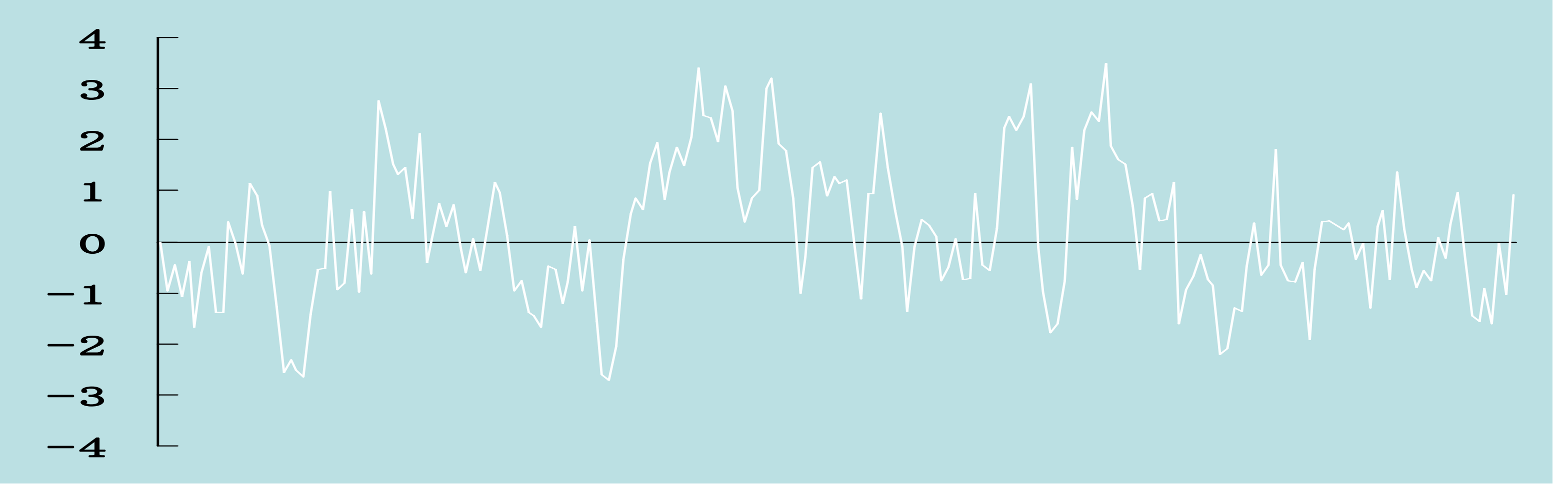
- 似乎相当复杂，但许多应用的模型可通过限定ψ_j权重而得到的，例如，令$\psi_j = \phi^j$，可写成

$$\begin{aligned} x_t &= a_t + \phi a_{t-1} + \phi^2 a_{t-2} + \cdots \\ &= a_t + \phi(a_{t-1} + \phi a_{t-2} + \cdots) \end{aligned}$$

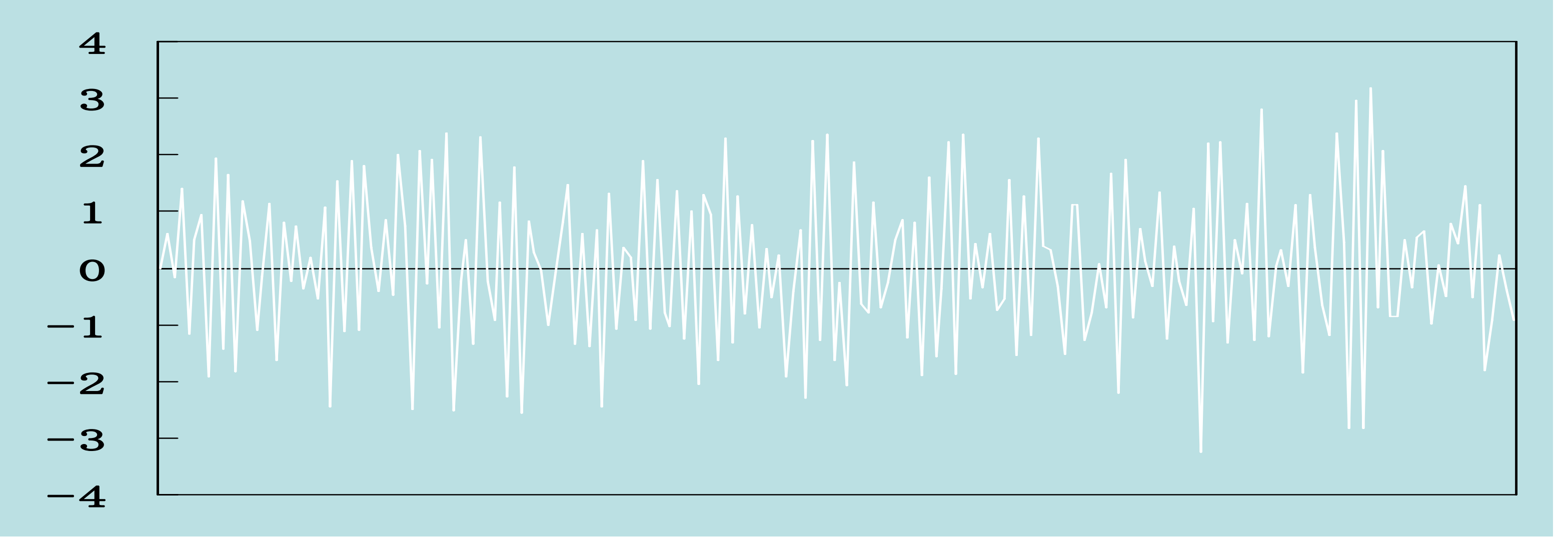
- 可得 $x_t = \phi x_{t-1} + a_t$
- 这就是一阶自回归过程，记为AR(1)。
- x_t部分地依赖于x_{t-1}，部分地依赖于“随机扰动”ε_t。或者说，x_t线性回归到x_{t-1}，ε_t是误差项。由于x_t对自身过去地回归，因而称之为“自回归过程”。

AR(1)过程的一次实现

- AR(1)过程x_t-0.8x_{t-1}=ε_t 的一次实现



- AR(1)过程x_t+0.8x_{t-1}=ε_t 的一次实现



平稳性

- 引入滞后算子*L*，AR(1)可表示为$(1-\phi L)x_t = a_t$

$$x_t = (1-\phi L)^{-1} a_t = (1+\phi L+\phi^2 L^2+\cdots)a_t$$
$$= a_t + \phi a_{t-1} + \phi^2 a_{t-2} + \cdots$$

- 只要$|\phi|$<1，上式是收敛的，x_t是平稳过程。
- AR(1)的特征方程为1－ϕ*L*=0，它的根为*L*＝1/ϕ，只要特征方程的根在单位圆外，AR(1)是平稳的。

自协方差函数

- 可由AR(1)过程的一阶矩、二阶矩来分析其统计特性

$$E\left((x_t-\phi x_{t-1})x_{t-k}\right)=E\left(a_t x_{t-k}\right)$$

$$\gamma_k-\phi\gamma_{k-1}=E\left(a_t x_{t-k}\right)=\sum_{j=0}^{\infty}\phi^j a_t a_{t-k-j}$$

- 由于a_t是白噪声，对于$k+j>0$，则 $E\left(a_t a_{t-k-j}\right)=0$
- 因此，当$k=0$时，则 $\gamma_0-\phi\gamma_{-1}=\sigma^2=\gamma_0-\phi\gamma_1$
- 当$k=1$时，则 $\gamma_1-\phi\gamma_0=0$ 过程的方差为

$$\gamma_0=\frac{\sigma^2}{1-\phi^2}$$

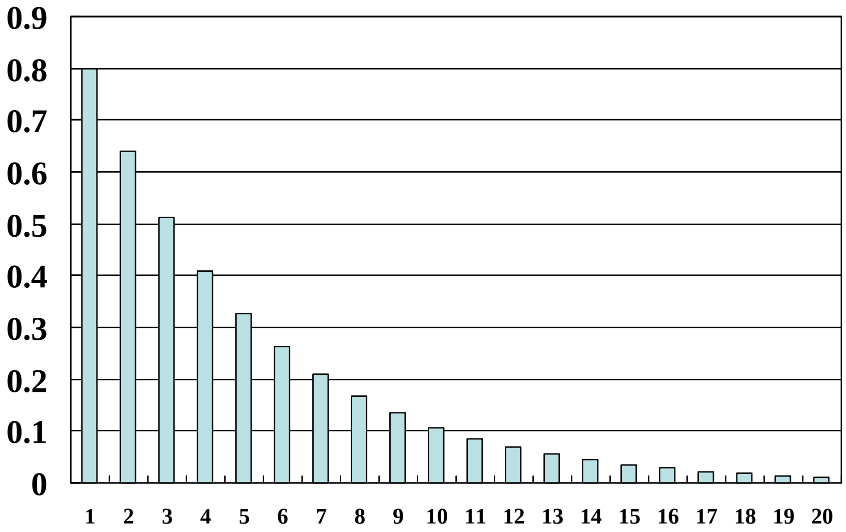
- 当$k>0$时，自协方差函数

$$\gamma_k-\phi\gamma_{k-1}=0$$

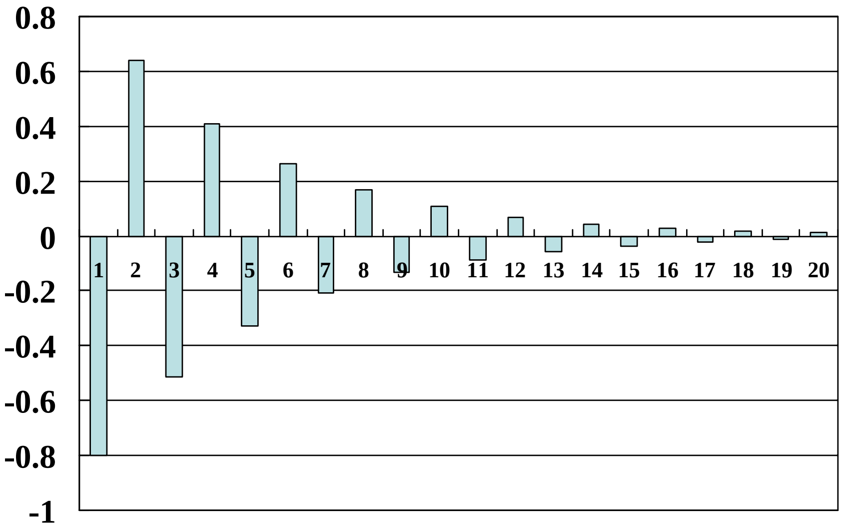
，$k=1,2,\ldots$

自相关系数ACF

AR(1)过程x_t-0.8x_{t-1}=a_t的ACF



AR(1)过程x_t+0.8x_{t-1}=a_t的ACF



- AR(1)模型(ε_t为零均值单位方差的正态随机过程)的自相关系数。
- 当ϕ=0.8>0时，自相关系数以指数形式衰减;
- 当ϕ=-0.8<0时，自相关系数以正负相间方式衰减。
- 而且，$|\phi|$ 越接近于1，衰减速度越慢。

二阶自回归过程

- x_t不仅与x_{t-1}有关，而且与x_{t-2}有关。AR(2)过程为

$$x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + a_t \qquad \left(1-\phi_1 L-\phi_2 L^2\right)x_t = a_t$$

- 平稳性
- 要使x_t平稳，须对ϕ_1、ϕ_2作限制。
- 对于AR(1)过程，如果平稳，特征方程的根必须在单位圆外，同样，若要AR(2)过程平稳，其特征方程 $1-\phi_1 L-\phi_2 L^2 = 0$
- 若要AR(2)过程平稳，必须满足$|g_1|<1$，$|g_2|<1$。而

$$g_{1,2} = \left[\phi_1 \pm \sqrt{\left(\phi_1^2 + 4\phi_2\right)}\right]/2$$

- g_1、g_2要么是一对复根，要么是实根。这时，AR(2)过程平稳，ϕ_1、ϕ_2必须满足

$$\phi_1 + \phi_2 < 1, -\phi_1 + \phi_2 < 1, \left|\phi_2\right| < 1$$

- 对于复根，要求

$$\phi_1^2 + 4\phi_2 < 0$$

自协方差函数、自相关函数

- AR(2)过程的自协方差函数，

$$E(x_t x_{t-k}) - \phi_1 E(x_{t-1} x_{t-k}) - \phi_2 E(x_{t-2} x_{t-k}) = E(a_t x_{t-k})$$

- a_t仅与x_t有关，而与x_{t-j}(j>0)无关，当k>0时，

$$\gamma_k - \phi_1 \gamma_{k-1} - \phi_2 \gamma_{k-2} = E(a_t x_{t-k}) = 0$$ ，k＝1,2,…

- 当k＝0时，有 $$\gamma_0 - \phi_1 \gamma_{-1} - \phi_2 \gamma_{-2} = \sigma^2 = \gamma_0 - \phi_1 \gamma_1 - \phi_2 \gamma_2$$
- 两边除以γ_0，可得AR(2)过程的自相关函数

$$\rho_k - \phi_1 \rho_{k-1} - \phi_2 \rho_{k-2} = 0, k = 1,2,\cdots$$

- 当k＝1时 $$\rho_1 - \phi_1 \rho_0 - \phi_2 \rho_{-1} = 0 = \rho_1 - \phi_1 \rho_0 - \phi_2 \rho_1$$
- 当k＝2时 $$\rho_2 - \phi_1 \rho_1 - \phi_2 \rho_0 = 0$$
- 对于k>2，

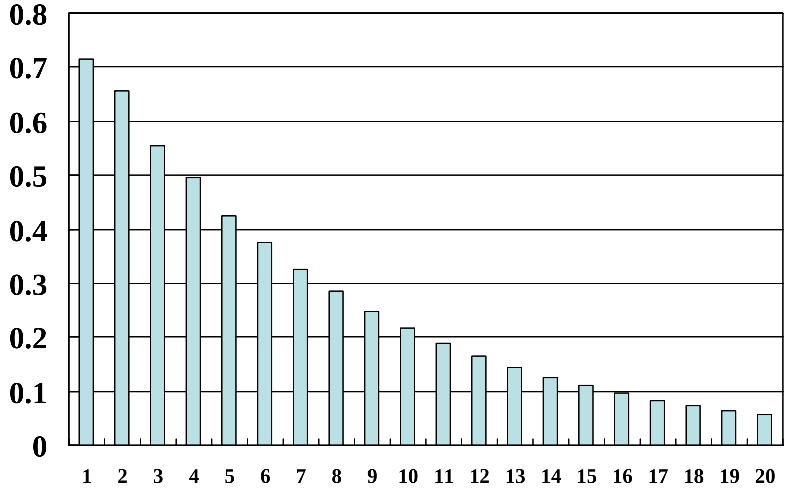
$$\rho_1 = \frac{\phi_1}{1 - \phi_2}$$

$$\rho_2 = \frac{\phi_1^2}{1 - \phi_2} + \phi_2$$

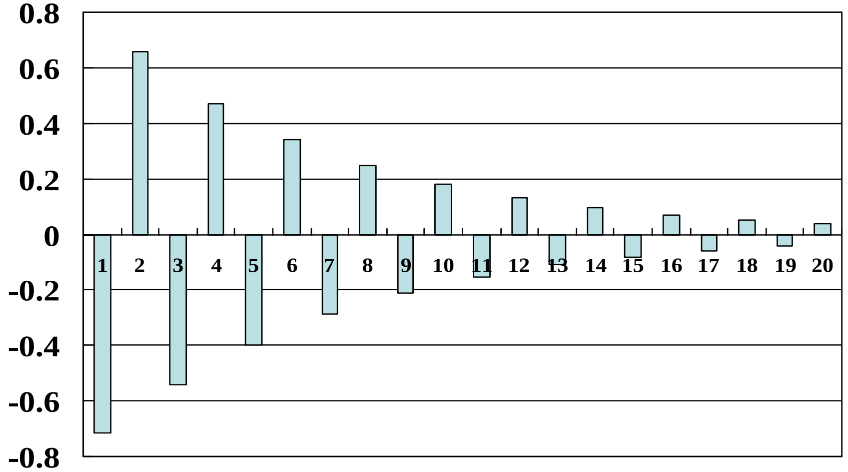
$$\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2}$$

自相关系数ACF

AR(2)过程$x_t-0.5x_{t-1}-0.3x_{t-2}=a_t$



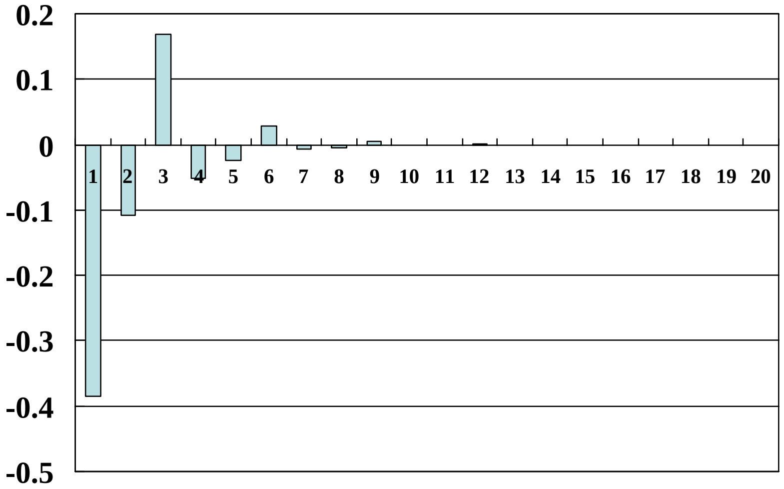
AR(2)过程$x_t-0.5x_{t-1}+0.3x_{t-2}=a_t$



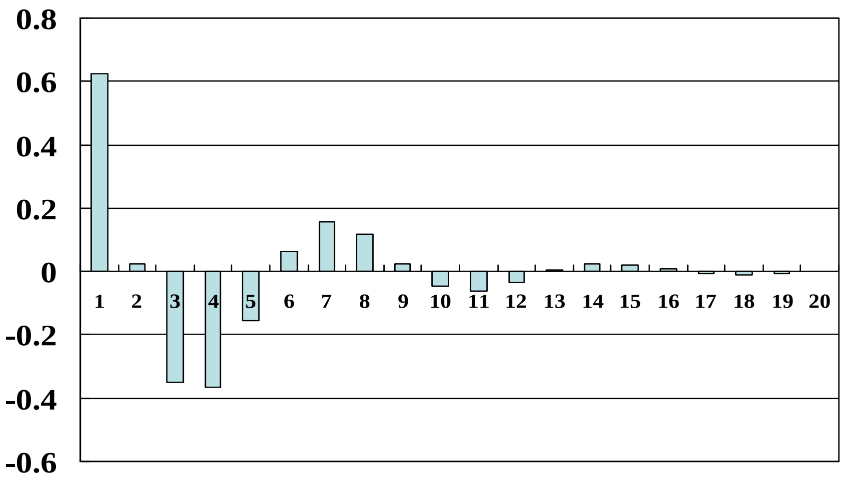
- g_1、g_2是实根，自相关系数以单调或正负相间的指数形式衰减;

自相关系数ACF

AR(2)过程$x_t+0.5x_{t-1}+0.3x_{t-2}=a_t$的ACF



AR(2)过程$x_t-x_{t-1}+0.5x_{t-2}=a_t$的ACF



- 当g_1、g_2是复根时，自相关系数以正弦波方式衰减。

p阶自回归过程

- 一般，p阶自回归过程AR(p)为

$$x_t - \phi_1 x_{t-1} - \phi_2 x_{t-2} - \ldots - \phi_p x_{t-p} = a_t$$

- 利用滞后算子，

$$\left(1 - \phi_1 L - \phi_2 L^2 - \ldots - \phi_p L^p\right) x_t = \phi(L) x_t = a_t$$

- 其中， $\phi(L) = 1 - \phi_1 L - \phi_2 L^2 - \ldots - \phi_p L^p$
- 若要将x_t表示为a_t的线性组合，即

$$x_t = \Psi(L) a_t$$

- 由 $\phi(L)\Psi(L) = 1$
- 可得 $\Psi(L)$ 的系数。

平稳性、自协方差函数

- 使x_t平稳，特征方程 $\phi(L)=0$ 的根在单位圆外
- AR(p)过程自协方差函数

$$\gamma_k - \phi_1\gamma_{k-1} - \cdots - \phi_p\gamma_{k-p} = 0$$

- AR(p)过程的自相关函数

$$\rho_k = \phi_1\rho_{k-1} + \ldots + \phi_p\rho_{k-p}$$

偏自相关函数

- 自回归过程意味着自相关系数呈衰减方式，如果用AR过程建模，却难以确定阶数p，引入偏自相关函数(PACF)。
- 设k阶自回归方程为：

$$x_t = \Phi_{k1} x_{t-1} + \ldots + \Phi_{kk} x_{t-k} + a_t$$

- Φ_{kk}是最后一个回归系数，Φ_{kk}可以看成是滞后期k的函数，称Φ_{kk}为偏自相关系数。因为偏自相关系数Φ_{kk}表示排除了中间变量x_{t-1}，x_{t-2}，…，x_{t-k+1}的影响后，x_t与x_{t-k}的自相关系数。
- 两边乘以x_{t-k}，两边取期望，然后除以γ_0，可得

$$\rho_j = \Phi_{k1}\rho_{j-1} + \Phi_{k2}\rho_{j-2} + \cdots + \Phi_{kk}\rho_{j-k}, j = 1,2,\cdots k$$

偏自相关系数

- 对于AR(p)过程，当$k \leqslant p$时，$\Phi_{kk} \neq 0$；当$k>p$时，$\Phi_{kk}=0$。
- 因此，偏自相关系数在滞后期p以后具有截尾特性，确定p。
- 例如，对于AR(1)过程，当$k=1$时，$\Phi_{11} \neq 0$，而当$k>1$时，$\Phi_{kk}=0$。所以，AR(1)过程的偏自相关系数在$k=1$出现峰值。而$k>1$，偏自相关系数为0，用偏自相关系数图可以确定自回归过程的阶数。
- 对于AR(2)，$\Phi_{11}=\rho_1$ $\Phi_{22}=\frac{\rho_2-\rho_1^2}{1-\rho_1^2}$ 当$k>2$时，$\Phi_{kk}=0$。
- 对于AR(p)，$\Phi_{11} \neq 0$，$\Phi_{22} \neq 0$，…，$\Phi_{pp} \neq 0$，当$k>p$时，$\Phi_{kk}=0$。
- 实证检验时，为了确定自回归过程的阶数，通常要计算Φ_{kk}的标准差，在自回归阶数等于p的假设下，$p+1$及更高阶的偏自相关系数的估计量 $\hat{\Phi}_{kk}$ ($k>p$)服从分布$N(0, 1/T)$，T为样本容量。
- 对于$k>p$，原假设：$\hat{\Phi}_{kk}=0$ 。在原假设成立的条件下，统计量 $\sqrt{T}\hat{\Phi}_{kk}$
- 服从标准正态分布$N(0, 1)$。

股价指数天报酬的自相关系数偏自相关系数

滞后阶数	上证指数		成份股指数		中信指数	
	自相关系数	偏自相关系数	自相关系数	偏自相关系数	自相关系数	偏自相关系数
1	-0.012	-0.012	0.033	0.033	0.042	0.042
2	-0.032	-0.032	-0.010	-0.011	-0.025	-0.027
3	0.019	0.018	0.049***	0.049***	0.040	0.042
4	0.057**	0.056**	0.048***	0.045***	0.026	0.022
5	-0.015	-0.012	-0.010	-0.012	-0.007	-0.007
6	-0.002	0.001	-0.005	-0.006	-0.034	-0.034
7	0.018	0.016	0.000	-0.005	0.019	0.020
8	-0.040	-0.042	-0.037	-0.038	0.019	0.016
9	-0.044	-0.043	-0.022	-0.018	-0.008	-0.006
10	0.009	0.005	0.006	0.007	-0.024	-0.022

主要货币汇率天报酬自相关系数偏自相关系数

滞后阶数	欧元/美元		英镑/美元		日元/美元	
	自相关系数	偏自相关系数	自相关系数	偏自相关系数	自相关系数	偏自相关系数
1	0.031***	0.031***	0.055*	0.055*	0.035***	0.035**
2	-0.021	-0.022	-0.006	-0.010	0.015	0.014
3	-0.023	-0.022	-0.006	-0.005	-0.050*	-0.051*
4	0.033***	0.034***	0.040**	0.041**	0.014	0.017
5	0.006	0.003	0.020	0.016	0.003	0.003
6	-0.021	-0.02	-0.018	-0.02	-0.014	-0.018
7	0.003	0.006	-0.030	-0.027	0.008	0.011
8	-0.011	-0.013	0.013	0.015	0.006	0.006
9	0.034***	0.034***	0.047*	0.044**	0.004	0.001
10	0.017	0.016	0.019	0.015	0.036**	0.037**

移动平均过程

- 另一种常见的过程就是移动平均过程 (moving average)，记为MA(q)，其中q为MA的阶数。
- 我们先介绍MA(1)、MA(2)过程，然后说明MA(q)，我们主要说明MA过程的均值、可逆性、自协方差函数、自相关函数、偏自相关函数。

一阶移动平均过程

$$x_t = a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \ldots = \sum_{j=0}^{\infty} \psi_j a_{t-j}$$

- 令的权重系数$\psi_1 = -\theta$，$\psi_j = 0$，可得 $x_t = a_t - \theta a_{t-1} = (1-\theta L) a_t$
- 这就是一阶移动平均过程，记为MA(1)。
- MA过程与AR过程不同，它是$\{a_t\}$过程的现在值a_t和过去值a_{t-1}的线性组合，所以，a_t仅影响x_t的1个未来值。
- 也就是说，相隔时间大于1的x_t中所包含的a_t已全部更新，所以，相隔时间超过1的两个x_t是不相关的，因此，MA(1)过程的2阶及以上的自相关系数为0，自相关系数具有截尾特性。
- 由于$\{x_t\}$是白噪声$\{a_t\}$的线性组合，显然，无论ψ_j取值如何，$\{x_t\}$总是平稳的。
- 对于MA(1)过程，相差一期变量是相关，相差一期以上的变量是不相关，过程的记忆仅为一期。

可逆性

- 与MA(1)过程相关的一个重要概念就是可逆性，也就是，如果MA(1)过程可逆，那么，MA(1)过程可表示为

$$(1-\theta L)^{-1}x_t = a_t$$

- 即无限阶AR过程。此时，要求特征方程的根

$$\Theta(L) = 1-\theta L = 0$$

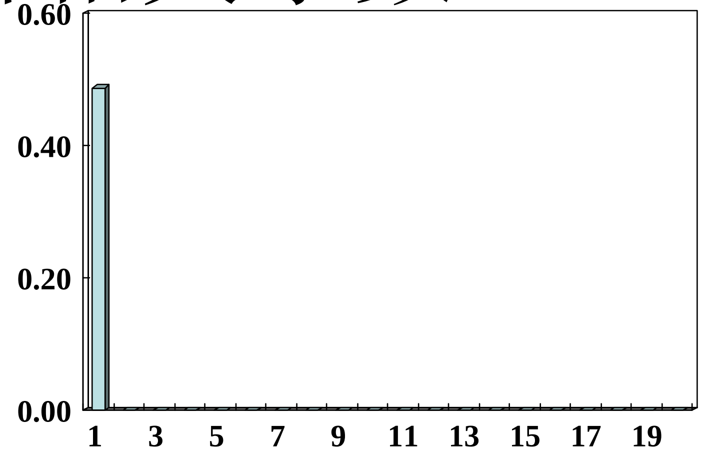
- 在单位圆外，此时，$|\theta| < 1$。

自协方差函数

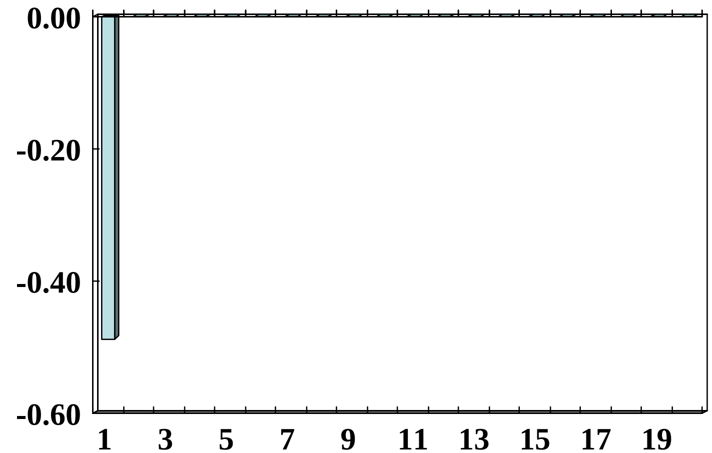
- MA(1)过程的均值为0，方差为 $Var(x_t)=(1+\theta^2)\sigma^2$
- 自协方差函数为，
- $$\gamma_1 = E(x_t x_{t-1}) = E\left[(a_t - \theta a_{t-1})(a_{t-1} - \theta a_{t-2})\right] = -\theta\sigma^2$$
- $$\gamma_k = E(x_t x_{t-k}) = 0$$ k>1
- 自相关函数表示为 $$\rho_1 = \frac{-\theta}{1+\theta^2}$$
- $$\rho_k = 0$$ k>1
- 对于MA(1)过程，从自相关函数上可以发现，MA(1)过程的自相关函数具有截尾特征。

MA (1)过程的自相关系数。

- 当θ=-0.8时，一阶自相关系数大于0，其他的自相关系数为0，MA(1)过程的自相关函数具有截尾特征。
- 当θ=0.8时，一阶自相关系数小于0，其他的自相关系数为0，MA(1)过程的自相关函数具有截尾特征。
- 回顾AR(1)过程的偏自相关系数，其特性刚好与MA(1)过程的自相关系数特性一样，具有截尾特征



MA(1)过程$x_t=a_t+0.8a_{t-1}$的ACF



MA(1)过程$x_t=a_t-0.8a_{t-1}$的ACF

二阶移动平均过程

$$x_t = a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \ldots = \sum_{j=0}^{\infty} \psi_j a_{t-j}$$

- 令的权重系数$\psi_1 = -\theta_1$， $\psi_2 = -\theta_2$ $x_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} = \left(1 - \theta_1 L - \theta_2 L^2\right) a_t$
- 这就是二阶移动平均过程，记为MA(2)，
- MA过程与AR过程不同，它是$\{a_t\}$过程的现在值a_t和过去值a_{t-1}、 a_{t-2}线性组合，所以，a_t仅影响x_t的2个未来值。
- 因此，MA(2)过程的3阶及以上的自相关系数为0，自相关系数具有截尾特性。对于MA(2)过程，相差二期变量是相关，相差二期以上的变量是不相关，过程的记忆仅为二期。

可逆性

- 如果MA(2)过程可逆，要求特征方程 $\Theta(L) = 1 - \theta_1 L - \theta_2 L^2 = 0$
- 的根在单位圆外，此时对参数的限制如下，$|\theta_2| < 1$，$\theta_1 + \theta_2 < 1$，$\theta_2 - \theta_1 < 1$。

自协方差函数、自相关函数

- MA(2)过程的均值为0，方差为 $Var(x_t)=\gamma_0=\left(1+\theta_1^2+\theta_2^2\right)\sigma^2$
- 自协方差函数为

$$\gamma_1=E(x_t x_{t-1})=E\left[(a_t-\theta_1 a_{t-1}-\theta_2 a_{t-2})(a_{t-1}-\theta_1 a_{t-2}-\theta_2 a_{t-3})\right]=\left(\theta_1\theta_2-\theta_1\right)\sigma^2$$

$$\gamma_2=-\theta_2\sigma^2$$

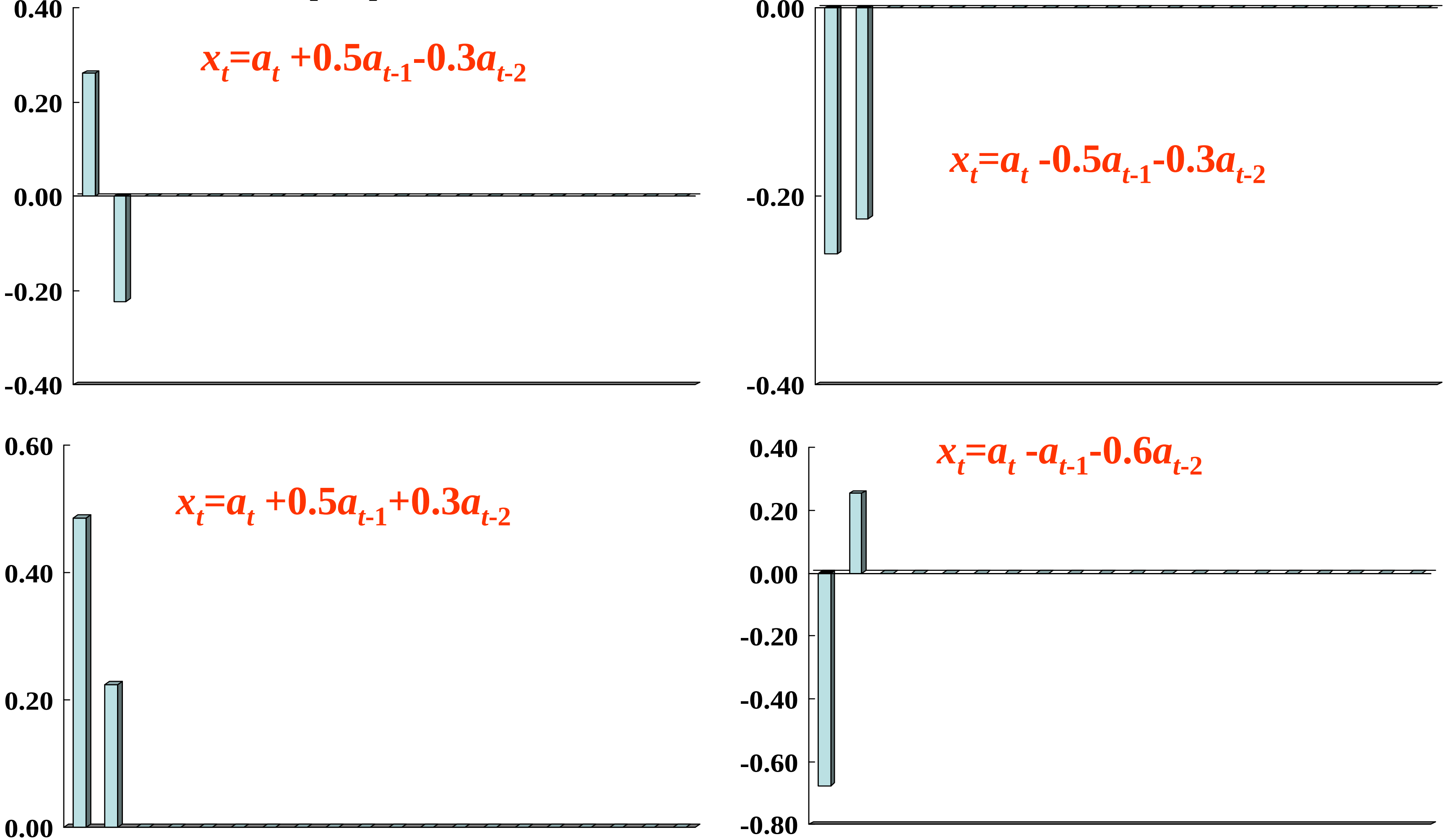
- $\gamma_k=E(x_t x_{t-k})=0,$ *k*>2
- 自相关函数表示为

$$\rho_1=\frac{-\theta_1+\theta_1\theta_2}{1+\theta_1^2+\theta_2^2}$$

$$\rho_2=\frac{-\theta_2}{1+\theta_1^2+\theta_2^2}$$

- $\rho_k=0$ *k*>2
- 对于MA(2)过程，从自相关函数上可以发现，(k>2)。可见，MA(2)过程的自相关函数也具有截尾特征。

MA (2)过程的自相关系数。



- 对于MA(2)过程，滞后1、2阶自相关系数不为0，而其他的滞后大于2的自相关系数为0，具有截尾特征。

q阶移动平均过程

$$x_t = a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \ldots + \psi_q a_{t-q}$$

$$x_t = \left(1 - \theta_1 L - \theta_2 L^2 - \ldots - \theta_q L^q\right) a_t$$

- 称为q阶移动平均过程，记为MA(q)。其中，a_t为白噪声。
- MA过程是$\{a_t\}$过程的现在值a_t和q个过去值a_{t-1}的线性组合，所以，a_t仅影响x_t的q个未来值。
- 也就是说，相隔时间大于q的x_t中所包含的a_t已全部更新，所以，相隔时间超过q的两个x_t是不相关的，因此，MA(1)过程的q阶以上的自相关系数为0，自相关系数具有截尾特性。
- 任何一个q阶移动平均过程 $\{x_t\}$是由q+1个白噪声的线性加权构成，所以任何一个移动平均过程是平稳的。
- 对于MA(q)过程，相差q期变量是相关，相差q期以上的变量是不相关，过程的记忆为q期。

可逆性

- MA过程可逆性的条件是特征方程

$$\Theta(L) = \left(1 - \theta_1 L - \theta_2 L^2 - \ldots - \theta_q L^q\right) = 0$$

- 的全部根的绝对值必须大于1，这样，MA(q)可以转换成AR(∞)过程。因为

$$x_t = \Theta(L) a_t$$

- 平稳，如果 $\Theta(L)^{-1} x_t = a_t$ 变得不平稳，显然失去可逆性。

自协方差函数、自相关函数

- MA(q)过程的均值为0，方差为 $Var(x_t)=\gamma_0=\sigma^2(1+\theta_1^2+\theta_2^2+\ldots+\theta_q^2)$
- 对于$k\le q$，自协方差函数为：

$$Cov(x_t,x_{t-k})=E(a_t-\theta_1 a_{t-1}-\ldots-\theta_k a_{t-k}-\theta_q a_{t-q})(a_{t-k}-\theta_1 a_{t-k-1}-\ldots-\theta_q a_{t-k-q})$$

- $$\gamma_k=\sigma^2(-\theta_k+\theta_1\theta_{k+1}+\ldots+\theta_{q-k}\theta_q)$$

$$\gamma_k=E(x_t x_{t-k})=0$$ $k>q$

- 自相关函数表示为

$$\rho_k=\frac{-\theta_k+\theta_1\theta_{k+1}+\cdots+\theta_{q-k}\theta_q}{1+\theta_1^2+\theta_2^2+\cdots+\theta_q^2},k=1,2,\cdots,q$$

- 而对于$k>q$，$\rho_k=0$ 。

自回归与移动平均过程的比较

- 平稳的AR(p)过程可转换成无限阶MA(∞)过程，

$$\Phi(L)x_t = a_t \Rightarrow x_t = \Phi(L)^{-1}a_t$$

- • 一个可逆的MA(q)可以转换成AR(∞)。
- AR仅考虑平稳性，只要其特征方程的根在单位圆外，但不考虑可逆性。
- MA(q)仅考虑可逆性，只要其特征方程的根在单位圆外，不考虑平稳性。
- 关于AR过程和MA过程的自相关系数和偏自相关系数比较

	AR(p)	MA(q)
自相关系数ρ_k	$\rho_k \neq 0$，*指数衰减*	k>q，$\rho_k = 0$
偏自相关函数Φ_{kk}	$k>p$，$\Phi_{kk}=0$	$\Phi_{kk} \neq 0$

- 如果$\rho_k \neq 0$，$\Phi_{kk} \neq 0$，且均拖尾时，需用AR过程与MA过程相结合的自回归移动平均过程。

自回归移动平均模型

- 更一般的过程就是用*p*阶自回归过程和*q*阶移动平均过程的混合模型来描述，称为自回归移动平均模型，简记为ARMA(*p*,*q*)，x_t满足

$$x_t - \phi_1 x_{t-1} - \cdots - \phi_p x_{t-p} = a_t - \theta_1 a_{t-1} - \cdots - \theta_q a_{t-q}$$

- 其中，ϕ_i、θ_j为常数(i=1,2,…,p，j=1,2, …,q)，a_t为白噪声。

ARMA(1,1)

- 在自回归移动平均模型中，最简单是ARMA(1,1)，x_t满足

$$x_t - \phi x_{t-1} = a_t - \theta a_{t-1}$$

- 若用滞后算子表示，

$$(1-\phi L)x_t = (1-\theta L)a_t$$

- 平稳性和可逆性
- 如果 $1-\phi L = 0$ 的根在单位圆外，$|\phi|<1$，那么x_t是平稳的。式可表示为MA(∞)。
- 如果可逆性条件成立，即$|\theta|<1$，那么x_t是可逆的，式可表示为AR(∞)。

自协方差函数

- ARMA(1,1)过程的均值为0，式两边乘以x_{t-k}，两边取期望，

$$E\left[(x_t-\phi x_{t-1})x_{t-k}\right]=E\left[(a_t-\theta a_{t-1})x_{t-k}\right]$$

- 即 $\gamma_k-\phi\gamma_{k-1}=E\left[(a_t-\theta a_{t-1})x_{t-k}\right]$
- 当k=0时，

$$\gamma_0-\phi\gamma_1=\sigma^2-\theta(\phi-\theta)\sigma^2$$

- 当k=1时，$\gamma_1-\phi\gamma_0=-\theta\sigma^2$
- 那么，ARMA(1,1)过程的方差为

$$\gamma_0=\frac{1+\theta^2-2\phi\theta}{1-\phi^2}\sigma^2$$

- ARMA(1,1)过程的自协方差函数为

$$\gamma_1=\frac{(1-\phi\theta)(\phi-\theta)}{1-\phi^2}\sigma^2$$

- $$\gamma_k-\phi\gamma_{k-1}=0$$ ，k>1

自相关函数

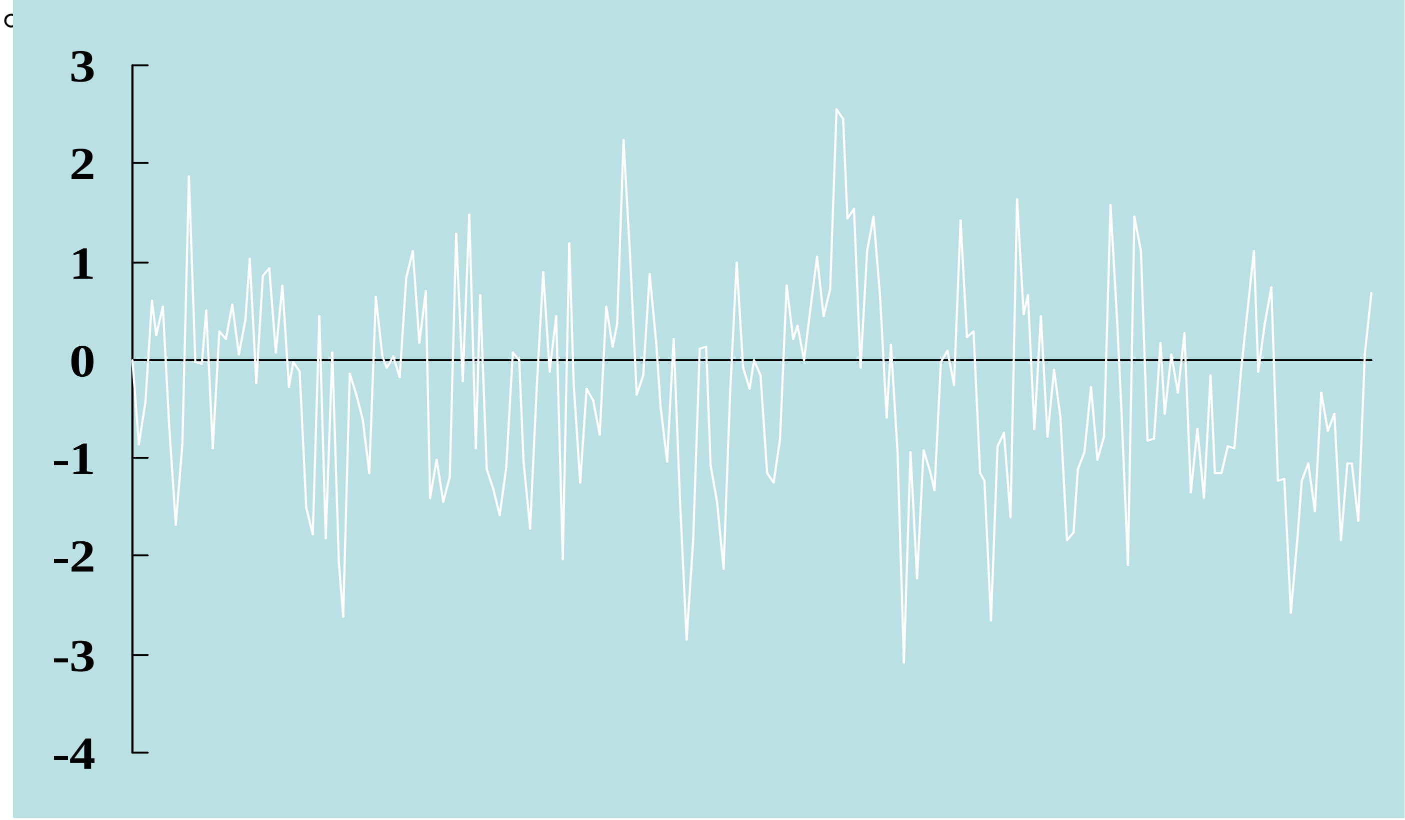
$$\rho_1 = \frac{(1-\phi\theta)(\phi-\theta)}{1+\theta^2-2\phi\theta} = \phi - \frac{\theta\sigma^2}{Var(x_t)} \neq \phi$$

- 当$k>1$时，$\rho_k = \phi\rho_{k-1}$，
- 但是偏自相关系数并未截尾。ARMA(1,1)的自相关系数类同于AR(1)的自相关系数，自相关系数以指数衰减。但与AR(1)有所不同，ARMA(1,1)的自相关系数衰减是从ρ_1开始，而不是从$\rho_0=1$开始，而且$\rho_1\neq\phi$，
- 大多数金融时间序列的ϕ，θ是正的，且$\phi>\theta$。若$\phi-\theta$很小，那么，$\rho_1\neq\phi$，ρ_1远小于ϕ。

ARMA(1,1)过程的一次实现

$$x_t = 0.85x_{t-1} + a_t - 0.5a_{t-1}$$

- 图为ARMA(1,1)模型
- (a_t为零均值单位方差的正态随机过程)的某次实现的200个观测值。



ARMA(*p*,*q*)过程

- 前面假设ARMA过程的均值为0，现考虑更一般ARMA(*p*,*q*)
- $$x_t - \phi_1 x_{t-1} - \cdots - \phi_p x_{t-p} = c + a_t - \theta_1 a_{t-1} - \cdots - \theta_q a_{t-q}$$
- *c*为常数。假设ARMA(*p*,*q*)过程的均值为μ， $$\mu = \frac{c}{1-\phi_1-\cdots-\phi_p}$$
- 若用滞后算子表示

$$(1-\phi_1 L-\phi_2 L^2-\cdots-\phi_p L^p)(x_t-\mu)=\left(1-\theta_1 L-\theta_2 L^2-\cdots-\theta_q L^q\right)a_t$$

$$\phi(L)=1-\phi_1 L-\phi_2 L^2-\ldots-\phi_p L^p,\Theta(L)=1-\theta_1 L-\theta_2 L^2-\ldots-\theta_q L^q$$

- ARMA(*p*,*q*) $$\phi(L)x_t = c+\Theta(L)a_t$$
- 平稳性和可逆性
- 如果$\phi(L)$=0的根在单位圆外，那么，x_t是平稳的。式可表示为MA(∞)。
- 如果Θ (L)=0的根在单位圆外，那么，x_t是可逆的，式可表示为AR(∞)。

自协方差函数

- 假设c、μ都为0，式两边乘以x_{t-k}，两边取期望，可得

$$E\left[(x_t-\phi_1 x_{t-1}-\cdots\phi_p x_{t-p})x_{t-k}\right]=E\left[(a_t-\theta_1 a_{t-1}-\cdots-\theta_q a_{t-q})x_{t-k}\right]$$

- 利用

$$E(x_t a_{t-k})=\begin{cases}\sigma^2 & k=0\\ \phi_k\sigma^2 & k>0\\ 0 & k<0\end{cases}$$

- 可得

$$\gamma_k-\gamma_1\rho_{k-1}-\cdots-\gamma_p\rho_{k-p}=\begin{cases}(1-\theta_1\phi_1-\cdots-\theta_q\phi_q)\sigma^2 & k=0\\ -(\theta_k-\theta_{k+1}\phi_1-\cdots-\theta_q\phi_{q-k})\sigma^2 & k=1,\cdots,q\\ 0 & k\geq q+1\end{cases}$$

- 其中，$\phi_0=0$，$\theta_j=0(j>q)$。

自相关函数

- 自相关函数ρ_k满足
- $$\rho_k - \phi_1 \rho_{k-1} - \cdots \phi_p \rho_{k-p} = 0$$ $k>q$
- 自相关函数满足p阶差分方程$\phi(L)\rho_k$=0 ($k>q$)，而$\rho_1,\ldots,\rho_q$作为初始条件。对于式，当$k=q+1,\ldots, q+p$时，可得到广义Yule-Walker方程，

$$\begin{bmatrix} \rho_{q+1} \\ \rho_{q+2} \\ \vdots \\ \rho_{q+p} \end{bmatrix} = \begin{bmatrix} \rho_q & \rho_{q-1} & \rho_{q-2} & \cdots & \rho_{q+2-p} & \rho_{q+1-p} \\ \rho_{q+1} & \rho_q & \rho_{q-1} & \cdots & \rho_{q+3-p} & \rho_{q+2-p} \\ \vdots & \vdots & & & & \vdots \\ \rho_{q+p-1} & \rho_{q+p-2} & \rho_{q+p-3} & & \rho_{q+1} & \rho_q \end{bmatrix} \begin{bmatrix} \phi_1 \\ \phi_2 \\ \vdots \\ \phi_p \end{bmatrix}$$

- 如果已得到相关系数ρ_k，就可以求得ϕ_k，$k=1,2,\ldots,p$。

自回归单整移动平均模型

- 上面我们假设x_t是平稳时间序列，如果x_t是非平稳时间序列，因为时间序列分析重点在于平稳时间序列，所以我们要将非平稳的时间序列转化为平稳的时间序列。这需要用差分的方法。

差分

- 时间序列$\{x_t\}$一阶差分定义为

$$\Delta x_t = x_t - x_{t-1} = (1-L)x_t$$

- 如果时间序列$\{x_t\}$的均值为μ

$$(1-L)(x_t-\mu) = (1-L)x_t - (1-L)\mu = (1-L)x_t$$

- 一阶差分后，已消除了均值，因此，对于一阶差分序列Δx_t，我们已得不到有关于x_t的均值信息。
- 时间序列$\{x_t\}$二阶差分定义为 $\Delta^2 x_t = \Delta(x_t - x_{t-1}) = x_t - 2x_{t-1} + x_{t-2} = (1-L)^2 x_t$
- 由于

$$\Delta^2(x_t + c + \beta t) = (1-L)^2 x_t + (1-L)^2 c + (1-L)^2(\beta t)$$
$$= (1-L)^2 x_t + \beta t - 2(\beta(t-1)) + \beta(t-2) = (1-L)^2 x_t$$

- 经过二次差分后，我们已得不到有关于x_t的均值和时间趋势的信息。
- 一般将d次差分算子记为$\Delta^d x_t$或$\Delta^d x_t = (1-L)^d x_t$。
- 对于时间序列x_t，如果必须经过d次差分之后才能变换成一个平稳的可逆的ARMA过程，而当进行d-1次差分之后仍是一个非平稳的过程，则称此过程具有d阶单整性，记x_t为I(d)，平稳过程则记为I(0)。

.随机游动或ARIMA(0,1,0)模型

- 如果时间序列x_t服从 $x_t = x_{t-1} + a_t$
- 称x_t是随机游动，或ARIMA(0,1,0)模型。
- 实际上，一阶差分序列Δx_t是白噪声。
- 模型也可以写成 $x_t = \phi x_{t-1} + a_t \qquad \phi = 1$
- 由于ϕ=1，特征方程1－ϕL=0的根在单位圆上，模型有一个单位根。
- 随机游动模型是非平稳的，可以将模型写为

$$x_t = (1-L)^{-1} a_t = a_t + a_{t-1} + a_{t-2} + \cdots$$

- 可见，x_t是白噪声a_t的现在值和过去值之和。因此，a_t任何过去值a_{t-k}(也称为冲击)对x_t的影响不会衰减或消失的，从而是持续的，其中，k>0。
- 那么，x_t的方差 $Var(x_t) = t\sigma^2$
- 是t 的线性递增函数。

有漂移的随机游动模型

- 更一般的，如果时间序列x_t服从

$$x_t = c + x_{t-1} + a_t$$

- 称x_t服从有漂移的随机游动，c称为漂移项。假定过程从t=0开始，那么，

$$x_t = (1-L)^{-1}c + (1-L)^{-1}a_t = ct + \sum_{i=0}^{t} a_{t-i}$$

- 常数项c 是时间趋势的斜率，它说明了平稳时间序列与非平稳时间序列的区别。
- 在平稳时间序列中，常数项c仅影响均值，
- 非平稳时间序列，常数项c是原序列中时间趋势的斜率。
- 当常数项不等于0时，有漂移的随机游动过程与纯随机游动过程是完全不同的。

指数平滑模型

- 预测中最常用的模型是指数平滑模型，模型的一般形式为(假定常数项为0)

$$x_t - x_{t-1} = a_t - \theta a_{t-1} \qquad |\theta| < 1$$

- ARIMA(0,1,1)模型.
- 由于模型是可逆的，可表示为

$$x_t = (1-\theta)x_{t-1} + \theta(1-\theta)x_{t-2} + \theta^2(1-\theta)x_{t-3} + \cdots + a_t = \sum_{i=1}^{\infty}\theta^{i-1}(1-\theta)x_{t-i} + a_t$$

- 注意到 $\sum_{i=1}^{\infty}\theta^{i-1}(1-\theta) = (1-\theta)\sum_{i=1}^{\infty}\theta^{i-1} = 1$
- x_t的现在值是其过去值的加权平均再加上随机误差a_t。权重依指数形式衰减。

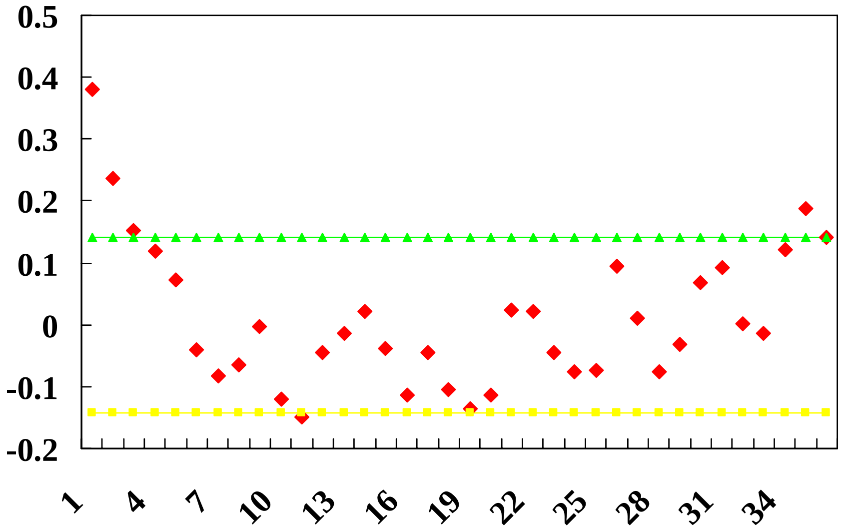
ARIMA过程的自相关系数、偏自相关系数

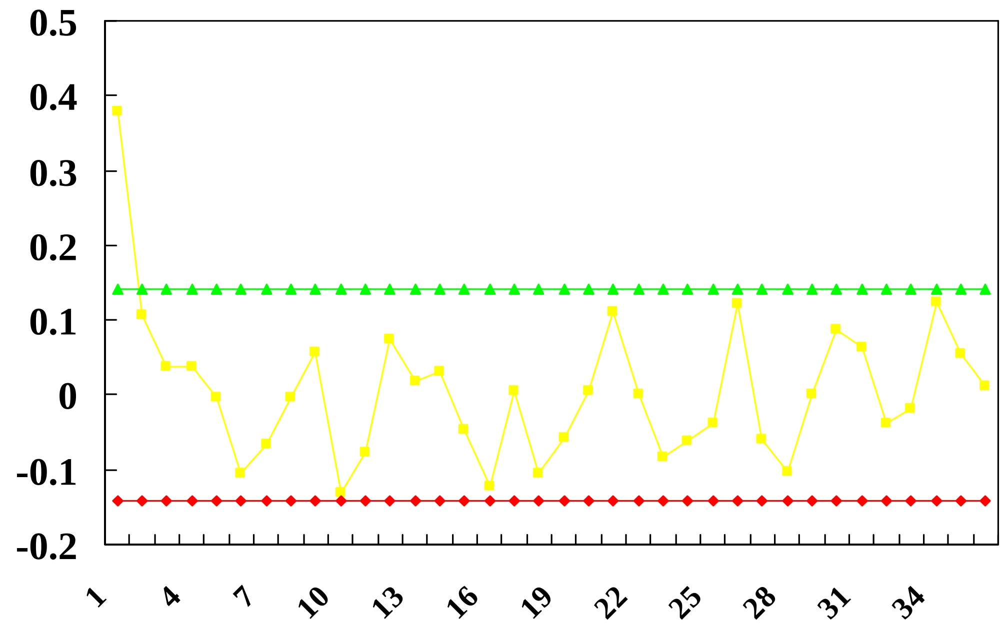
模型	自相关系数	偏自相关系数
ARIMA(1,1,1)	缓慢地线性衰减	
AR(1)	若ϕ_1>0，指数衰减；	若ϕ_1>0，k=1有正峰值，然后截尾；
	若ϕ_1<0，正负交替指数衰减；	若ϕ_1<0，k=1有负峰值，然后截尾
AR(2)	若θ_1>0，k=1有正峰值，然后截尾；	若θ_1>0，指数衰减；
	若θ_1<0，k=1有负峰值，然后截尾	若θ_1<0，正负交替指数衰减
MA(1)	指数或正弦衰减	k=1，2时有两个峰值，然后截尾
MA(2)	k=1，2时有两个峰值，然后截尾	指数或正弦衰减
ARMA(1,1)	k=1时有峰值，然后指数衰减	k=1时有峰值，然后指数衰减
ARMA(2,1)	k=1时有峰值，然后指数或正弦衰减	k=1，2时有两个峰值，然后指数衰减
ARMA(1,2)	k=1，2时有两个峰值，然后指数衰减	k=1时有峰值，然后指数或正弦衰减
ARMA(2,2)	k=1，2时有两个峰值，然后指数衰减	k=1，2时有两个峰值，然后指数衰减

识别ARMA模型的阶数

- 序列的自相关系数与偏自相关系数图可以为识别模型阶数*p*、*q*提供信息。
- 如果自相关系数图表示为拖尾衰减特征，而偏自相关系数图表现为*p*期后截尾特征，该过程是一个*p*阶自回归过程。
- 如果自相关系数图*q*期出现截尾，而偏自相关系数图呈拖尾衰减特征，该过程为MA(q)
- 如果自相关系数图和偏自相关系数图均表现了拖尾衰减特征，说明这是一个混合形式的随机过程。为识别*p*、*q*，可以认为该过程由MA和AR两部分叠加而成。
- ARMA(*p*,*q*)过程的自相关系数图在*q-p*滞后期后指数衰减和正弦衰减，而ARMA(*p*,*q*)过程的偏自相关系数图在*q-p*滞后期后指数衰减和正弦衰减。
- 当自相关系数图出现一个峰值后呈指数衰减，而偏自相关系数图出现一个峰值后呈指数或正弦衰减，说明是一个ARMA(1,1)过程，当自相关系数图出现两个峰值后呈指数衰减，而偏自相关系数图出现一个峰值后呈指数或正弦衰减，说明这是一个ARMA(1,2)过程。

ARMA(1,1)过程自相关系数、偏自相关系数





ARMA模型的建模

- 构建ARMA模型一般涉及三个步骤：模型识别、模型估计和诊断检验
- 三个步骤不断地重复，一直到得到合适的模型为止。
- 实际上，这三个步骤不仅应用于时间序列的分析，而且应用于其他模型的构建，如ARCH模型等。

模型识别

- ARMA模型识别是指从样本数据中确定模型的阶数，主要有两种方法，
- 一是从样本的自相关系数、偏自相关系数确定AR和MA的阶数；二为利用信息准则方法来确定模型的阶数。

样本自相关函数和偏自相关函数

- 给定样本的自相关系数ρ_k、偏自相关系数Φ_{kk}，Box-Jenkins的方法就是将样本自相关系数、样本偏自相关系数与理论的ACF、PACF相比，然后选取合适阶数p、q，利用最小二乘法或最大似然估计法估计模型的参数，再诊断检验模型的残差，这样的过程一直持续到得到合适的模型为止。
- 建立ARMA模型第一步，是估计观测数据的均值、方差、样本自相关系数、偏自相关系数，。
- 请注意：自相关系数用于判别MA的阶数，而偏自相关系数用于判别AR的阶数。

case

- 以1953年4月至2003年2月的月度美国10年国债收益率与3个月国库券利率之差(简称为利差)为例
- 前10阶自相关系数在1%显著性水平下显著大于0，且呈指数衰减的方式，故可以考虑AR模型，而从偏自相关系数来看，前4阶以及滞后7、9、10均在5%显著性水平下显著不等于0，也就是说并未出现截尾的形态。因而考虑用ARMA模型。

滞后阶数*k*	自相关系数	偏自相关系数	LB(*k*)
1	0.949*	0.949*	542.48
2	0.875*	-0.267*	1003.8
3	0.813*	0.158*	1403.1
4	0.753*	-0.108**	1745.9
5	0.698*	0.073***	2041.1
6	0.652*	0.005	2298.8
7	0.617*	0.094**	2530.3
8	0.589*	-0.008	2741.3
9	0.552*	-0.099**	2927.0
10	0.500*	-0.128*	3079.7

case

- ARMA(2,1)模型，估计结果如下：

$$x_t = 0.1253 + 0.6127\, x_{t-1} + 0.2967\, x_{t-2} + a_t - 0.6727\, a_{t-1}$$

- (3.5) (7.1) (3.5) (9.8)
- 调整的R^2=0.9148，D.W.＝1.99，括号内为t统计值拟合的参数。
- 利差样本的均值 $\hat{\mu} = 0.1253 /(1 - \hat{\phi}_1 - \hat{\phi}_2) = 1.383$
- 平稳性条件， $\hat{\phi}_1 + \hat{\phi}_2 = 0.9094, \hat{\phi}_2 - \hat{\phi}_1 = -0.3160, \hat{\phi}_2 = 0.2967$
- 而且，$\hat{\phi}_1^2 + 4\hat{\phi}_2 = 1.562 > 0$ 有两个实根，0.93、-0.32。
- 下面考虑所选的模型是否适当。先分析模型的残差是否是白噪声，利用自相关系数，前8阶的自相关系数在5%显著性水平下，均与0无显著差异，也可利用LB统计量，例如，LB(8)＝5.928，在5%显著性水平下小于χ^2(8)的临界值。
- 另外，我们也可考虑用ARMA(2,2)模型拟合数据，估计结果如下：

$$x_t = 0.1380 + 0.4153\, x_{t-1} + 0.4752\, x_{t-2} + a_t - 0.8757\, a_{t-1} - 0.0979\, a_{t-2}$$

- (5.5) (1.7) (2.2) (3.6) (0.8)
- 但是，加入的a_{t-2}项的估计系数在统计意义上不具显著性，用ARMA(1,2)较适当。

信息准则

- 选择模型阶数*p*、*q*的另一种方法，利用信息准则的方法，其基本形式为

$$crit(m) = -2\ln(\hat{\sigma}_a^2) + f(T,m)$$

- *m*指模型，*T*为样本容量，$\hat{\sigma}_a^2$ 是模型残差方差的最大似然估计，*f*(*T*,*m*)为模型*m*中样本容量*T*和估计参数个数的函数。
- 式的第一项反映模型拟合优度，第二项是“惩罚函数”，主要惩罚模型的高阶。
- 式在对模型的拟合残差和模型的阶数之间作出选择，以体现对残差和模型阶数两者的重要性的不同侧重。在选择适当模型时，选择使准则函数最小的阶数。

信息准则

- 确定ARMA(*p*,*q*)模型阶数常用的准则函数包括
- (1) 赤池(Akaike,1974)提出了赤池信息准则(AIC)，
- $$AIC(p,q)=\ln\hat{\sigma}_a^2+2(p+q)/T$$
- (2)许瓦尔兹(Schwarz,1978)提出的许瓦尔兹信息准则(SC)，

$$BC(p,q)=\ln\hat{\sigma}_a^2+(p+q)\ln(T)/T$$

- (3)汉南和奎恩(Hannan and Quinn, 1979) 的汉南和奎恩信息准则

$$HQ(p,q)=\ln\hat{\sigma}_a^2+(p+q)\ln[\ln(T)]/T$$

- 一般，阶数的选择就是利用，

$$AIC(p_1,q_1)=\min_{p,q}AIC(p,q)$$

$$BC(p_1,q_1)=\min_{p,q}BC(p,q)$$

- 得到的p_1，q_1就是模型的阶数。
- 关于AIC和SC的比较，一般AIC选择的模型阶数较真实模型的高，而SC选择的模型阶数与真实模型的是一致的

上证指数天报酬ARMA模型阶数的选择

	q p	0	1	2	3	4	5
AIC	0	3.8173	3.8180	3.8172	3.8176	3.8163	3.8177
	1	3.8185	3.8186	3.8176	3.8178	3.8175	3.8187
	2	3.8190	3.8197	3.8005	3.8157	3.8185	3.8201
	3	3.8201	3.8193	3.7986	3.8157	3.8121	3.8141
	4	3.8178	3.8183	3.8130	3.8159	3.8099	3.8087
	5	3.8191	3.8192	3.8080	3.8098	3.8109	3.8109

SC	0	3.8290	3.8253	3.8282	3.8322	3.8346	3.8397
	1	3.8258	3.8295	3.8323	3.8361	3.8394	3.8443
	2	3.8300	3.8344	3.8188	3.8377	3.8841	3.8494
	3	3.8346	3.8376	3.8205	3.8415	3.8413	3.8470
	4	3.8360	3.8402	3.8391	3.8391	3.8428	3.8453
	5	3.8401	3.8448	3.8419	3.8419	3.8474	3.8512

模型估计、诊断检验

- 在时间序列分析中，估计模型的参数常用最大似然估计方法
- 三、诊断检验
- 时间序列建模的第3个步骤就是模型的诊断检验，如同回归分析，也注重对模型残差的分析，因而，考虑残差图、残差序列的相关性以及异常点的检测。

- 我们可以画出残差序列图，便于检验异常点、残差序列的相关性、异方差，以及检验残差序列的正态性。残差序列图是时间序列分析的组成部分。
- 残差序列相关性也可用残差序列自相关系数、偏自相关系数检验。如果模型适当，那么，残差是白噪声。这样，可以利用上面所述的方法，如检验自相关系数是否显著异于0，还有利用LB统计量从总体上检验序列是否存在相关性。

ARMA模型的预测

- 间序列分析的主要目的之一就是预测，通常利用最小均方误差的准则来获得点预测。当然预测也有风险，就是为了自变量未来值的不确定性和所用模型参数的不确定性。
- 例如，模型参数是估计值，而不是真实的参数，对于预测，我们不考虑第二个不确定性，认为估计的参数是真实的参数。

预测

- 预测开始之时是*T*，预测*T+l*时刻的值，这种预测称为*l*步预测，预测值，$\hat{x}_T(l)$
- 就是给定x_{T-1}, x_{T-2},…和模型*f*的条件下，x_{T+l}的条件期望，即

$$\hat{x}_T(l) = E(x_{T+l} | x_T, x_{T-1}, \cdots, f)$$

- *l*步预测误差为 $e_T(l) = x_{T+l} - \hat{x}_T(l)$
- *l*步预测误差方差为

$$Var(e_T(l)) = Var(x_{T+l} - \hat{x}_T(l))$$

预测评价

- **1、预测误差平均值：**

$$\bar{e}=\frac{1}{l}\sum_{\tau=1}^{l}e_T(\tau)$$

如果 $\bar{e}$ 不接近于零，则表明预测值同实际值之间存在系统性偏差。

- **2、预测误差的标准差：**

$$SD=\sqrt{\frac{1}{l}\sum_{\tau=1}^{l}e_T^2(\tau)}$$

若**SD**较大，则说明预测的误差较大，

- 平均绝对误差率

$$MAPE=\frac{1}{l}\sum_{\tau=1}^{l}\left|\frac{e_T(\tau)}{x_{T+\tau}}\right|\times 100\%$$

- 这是一相对指标，但不适宜可以取零值的经济变量

AR(1)模型的预测

- AR(1)模型为 $x_t - \phi x_{t-1} = a_t a$
- l步预测， $\hat{x}_T(l) = \phi^l x_T$
- l步预测误差为 $e_T(l) = a_{T+l} + \phi a_{T+l-1} + \cdots + \phi^{l-1} a_{T+1}$
- l步预测误差方差为 $Var(e_T(l)) = (1+\phi^2+\cdots+\phi^{2(l-1)})\sigma_a^2$
- 当l趋向于无穷大时 $\hat{x}_T(l) \to 0$
- 一般，当l趋向于无穷大时， $\hat{x}_T(l)$ 趋向于x_t的均值。
- 由于$|\phi| < 1$ l步预测误差方差收敛于x_t的方差。$\sigma_a^2 / (1-\phi^2)$
- 总之，对于平稳的AR(1)模型，由于序列的自相关系数依指数衰减至0，意味着x_t在T时刻的值对于x_{T+l} (l较大)没有实质的影响。

MA(1)模型 预测

- MA(1)模型为

$$x_t = a_t - \theta a_{t-1} \qquad |\theta| < 1$$

- *l*步预测，

$$\hat{x}_T(l) = \begin{cases} -\theta a_T & l = 1 \\ 0 & l > 1 \end{cases}$$

- 由于MA(1)模型记忆为一期，因而，一期后，记忆消失。
- *l*步预测误差为，

$$e_T(l) = \begin{cases} a_{T+l} & l = 1 \\ a_{T+l} - \theta a_{T+l-1} & l > 1 \end{cases}$$

- *l*步预测误差方差为，

$$Var(e_T(l)) = \begin{cases} \sigma_a^2 & l = 1 \\ (1+\theta^2)\sigma_a^2 & l > 1 \end{cases}$$