1. 平稳序列：(i) 如果时间序列 $\{X_t\}=\{X_t: t\in\mathbb{N}\}$ 满足 (1) $\forall t\in\mathbb{N}$, $EX_t^2<\infty$;
(2) $\forall t\in\mathbb{N}$, $EX_t=\mu$; (3) $\forall t,s\in\mathbb{N}$, $E[(X_t-\mu)(X_s-\mu)]=\gamma_{t-s}$, 则称 X_t 是平稳时间序列，简称为平稳序列。称实数列 $\{\gamma_t\}$ 为 $\{X_t\}$ 的自协方差函数；称 $\rho_k=\frac{\gamma_k}{\gamma_0}$ 为自相关系数。

EX_t 和 $Var X_t$ 均为与 t 无关的常数；$\forall s,t\in\mathbb{Z}, k\in\mathbb{Z}$, $cov(X_t,X_s)=Cov(X_{t+k},X_{s+k})=\gamma_{s-t}$，即协方差结构平移不变。

(ii) $\{\gamma_t\}$ 具有如下性质：(1) 对称性：$\gamma_k=\gamma_{-k}$ $\forall k\in\mathbb{Z}$.

(2) 非负定性：$\forall n\in\mathbb{N}^+$, $\Gamma_n=\begin{bmatrix}\gamma_0 & \gamma_1 & \cdots & \gamma_{n-1}\\ \gamma_1 & \gamma_0 & \cdots & \gamma_{n-2}\\ \vdots & \vdots & \ddots & \vdots\\ \gamma_{n-1} & \gamma_{n-2} & \cdots & \gamma_0\end{bmatrix}$ 是非负定阵.

(3) $|\gamma_k|\le\gamma_0$ $\forall k\in\mathbb{Z}$ 成立.

2. 白噪声：设 $\{\varepsilon_t\}$ 是一个平稳序列，如果对任何 $s,t\in\mathbb{N}$, $E\varepsilon_t=\mu$, $Cov(\varepsilon_t,\varepsilon_s)=\begin{cases}\sigma^2, & t=s,\\ 0, & t\neq s.\end{cases}$ 就称 $\{\varepsilon_t\}$ 是一个白噪声，记作 $WN(\mu,\sigma^2)$. 设 $\{\varepsilon_t\}$ 是 $WN(\mu,\sigma^2)$.

3. 正交平稳序列：对于平稳序列 $\{X_t\}$ 和 $\{Y_t\}$,

(1) 若对于任何 $s,t\in\mathbb{Z}$, $E(X_tY_s)=0$. 则称 $\{X_t\}$ 和 $\{Y_t\}$ 正交；

(2) 若对于任何 $s,t\in\mathbb{Z}$, $Cov(X_t,Y_s)=0$. 则称 $\{X_t\}$ 和 $\{Y_t\}$ 不相关.

对于零均值的平稳序列，正交性和不相关性等价.

1101C-08 201412·2500

4. 绝对可和：若实数列 $\{a_j\}$ 满足 $\sum_{j=-\infty}^{\infty}|a_j|<\infty$，就称 $\{a_j\}$ 绝对可和。

对于绝对可和的实数列 $\{a_j\}$，定义零均值白噪声 $\{\varepsilon_t\}$ 的无穷滑动和 $X_t=\sum_{j=-\infty}^{\infty}a_j\varepsilon_{t-j}$，$t\in\mathbb{Z}$，则 $\{X_t\}$ 是平稳序列。

一般线性平稳序列只要求 $\{a_j\}$ 平方可和，即只要求 $\sum_{j=-\infty}^{\infty}a_j^2<\infty$。

5. Hilbert 空间：用 $L^2(X)$ 表示平稳序列 $\{X_t\}$ 中随机变量有限线性组合的全体：

$L^2(X)=\{\sum_{j=1}^{k}a_jX(t_j)\mid a_j\in\mathbb{R},t_j\in\mathbb{Z},1\le j\le k,k\in\mathbb{N}^+\}$。在 $L^2(X)$ 上定义内积 $\langle X,Y\rangle=E(XY)$。

引入距离 $\|Y-X\|=\langle Y-X,Y-X\rangle^{1/2}$，则 $L^2(X)$ 是内积、距离空间。

用 L^2 表示二阶矩有限的随机变量的全体：$L^2=\{X:EX^2<\infty\}$，则 L^2 也是内积、距离空间，$L^2(X)\subset L^2$。

均方收敛：$\forall\xi_n\in L^2$，$\xi_0\in L^2$。若 $\lim_{n\to\infty}\|\xi_n-\xi_0\|=\lim_{n\to\infty}E|\xi_n-\xi_0|^2=0$，则称 ξ_n 在 L^2 中收敛到 ξ_0，记作 $\xi_n\xrightarrow{L^2}\xi_0$；若当 $n,m\to\infty$ 时 $\|\xi_n-\xi_m\|\to0$，称 $\{\xi_n\}$ 是 L^2 中的基本列或 Cauchy 列。

完备的内积空间又称为 Hilbert 空间。L^2 是 Hilbert 空间。

用 $\overline{L^2}(X)$ 表示 L^2 中含 $L^2(X)$ 的最小闭子空间，则 $\overline{L^2}(X)$ 是 Hilbert 空间，称为由平稳序列 $\{X_t\}$ 生成的 Hilbert 空间。

(*) 验证 $Y_t=\sum_{k=-\infty}^{\infty}\varphi_k(t)$ 是平稳序列的步骤，整体分为两大步：

(1) 证明 Y_t 的存在性——作为极限而存在；

(2) 证明 Y_t 的平稳性——$EY_t=\mu$，$EY_t^2<\infty$，$\mathrm{cov}(Y_t,Y_s)=\gamma_{t-s}$。

1101C-08 201412·2500

中国科学技术大学

其中(ii)又可分为以下几步:

(i) 构造$\mathring{L}^2(X)$进而构造$\overline{L}^2(X)$; (ii) 根据题目已知条件(通常在所证结论中)构造辅助序列$\{\varphi_n(t)\}$,证明$\{\varphi_n(t)\}$是$\overline{L}^2(X)$中的Cauchy列; (iii) 定义$Y_t=\lim_{n\to\infty}\varphi_n(t)$.

6. 谱密度:设平稳序列$\{X_t\}$有自协方差函数$\{\gamma_k\}$.

(1) 如果有$[-\pi,\pi]$上的单调不减右连续函数$F(\lambda)$,使得$\gamma_k=\int_{-\pi}^{\pi}e^{ik\lambda}dF(\lambda)$, $F(-\pi)=0$, $k\in\mathbb{Z}$,就称$F(\lambda)$是$\{X_t\}$或$\{\gamma_k\}$的谱分布函数,简称为谱函数.

(2) 如果有$[-\pi,\pi]$上的非负函数$f(\lambda)$使得$\gamma_k=\int_{-\pi}^{\pi}f(\lambda)e^{ik\lambda}d\lambda$,则称对$\gamma_k=\int_{-\pi}^{\pi}e^{ik\lambda}f(\lambda)d\lambda$, $k\in\mathbb{Z}$,称$f(\lambda)$是$\{X_t\}$或$\{\gamma_k\}$的谱密度函数. 谱密度.

Herglotz定理:平稳序列的谱唯一存在.

7. 推移算子:对任何时间序列$\{X_t\}$和无穷级数$\psi(z)=\sum_{j=-\infty}^{\infty}b_jz^j$,只要级数$\sum_{j=-\infty}^{\infty}b_jX_{t-j}$在某种意义下收敛(a.s.收敛、依概率收敛、均方收敛),就定义$\psi(\mathscr{B})=\sum_{j=-\infty}^{\infty}b_j\mathscr{B}^j$, $\psi(\mathscr{B})X_t=\sum_{j=-\infty}^{\infty}b_j\mathscr{B}^jX_t=\sum_{j=-\infty}^{\infty}b_jX_{t-j}$. 称$\mathscr{B}$是时间$t$的向后推移算子,简称为推移算子.

8. AR(p)模型:若$\{\varepsilon_t\}$是$WN(0,\sigma^2)$,实数$a_1,a_2,\cdots,a_p(a_p\neq0)$使得多项式$A(z)$的零点都在单位圆外:$A(z)=1-\sum_{j=1}^{p}a_jz^j\neq0$, $|z|\leq1$. 就称p阶差分方程$X_t=\sum_{j=1}^{p}a_jX_{t-j}+\varepsilon_t$, $t\in\mathbb{Z}$是一个p阶自回归模型,简称为AR(p)模型.

1101C-08 201412·2500

中国科学技术大学

满足AR(p)模型的平稳时间序列$\{X_t\}$称为平稳解或AR(p)序列. AR(p)模型可改写为$A(\mathscr{B})X_t=\varepsilon_t, t\in Z$.

9. Wold系数：设多项式$A(z)$的互异根是$z_1, z_2, \cdots, z_k$. 则对$1<\rho<\min\{|z_j|\}$. $[A(z)]^{-1}=\frac{1}{A(z)}$是$\{z: |z|\le\rho\}$内的解析函数. 从而有Taylor级数$A^{-1}(z)=\sum_{j=0}^{\infty}\psi_j z^j$, $|z|\le\rho$. 由级数在$z=\rho$的绝对收敛性得到：当$j\to\infty$时, $|\psi_j\rho^j|\to 0$. 于是由$\psi_j=o(\rho^{-j})$, $j\to\infty$知$\{\psi_j\}$绝对可和. 而且$\min\{|z_j|\}$越大, $\psi_j\to 0$越快.

定义$A^{-1}(\mathscr{B})=\sum_{j=0}^{\infty}\psi_j\mathscr{B}^j$, 得$X_t=A^{-1}(\mathscr{B})A(\mathscr{B})X_t=A^{-1}(\mathscr{B})\varepsilon_t, t\in Z$.

即$X_t=A^{-1}(\mathscr{B})\varepsilon_t=\sum_{j=0}^{\infty}\psi_j\varepsilon_{t-j}$, $t\in Z$决定的平稳序列是AR(p)序列. $\{\psi_j\}$称为平稳序列$\{X_t\}$的Wold系数.

因果形式, 具体表达式需要确定$\{\psi_j\}$的值, 参见习题

10. AR(p)序列$\{X_t\}$的谱密度函数$f(\lambda)=\frac{\sigma^2}{2\pi}\left|\sum_{j=0}^{\infty}\psi_j e^{ij\lambda}\right|^2=\frac{\sigma^2}{2\pi|A(e^{i\lambda})|^2}$, $A(z)$是AR(p)模型的特征多项式. 另：平稳序列$\{X_t\}$的自协方差函数$\{\gamma_k\}$绝对可和：$\sum_{k=-\infty}^{\infty}|\gamma_k|<\infty$, 则$\{X_t\}$有谱密度$f(\lambda)=\frac{1}{2\pi}\sum_{k=-\infty}^{\infty}\gamma_k e^{-ik\lambda}=\frac{1}{2\pi}\sum_{k=-\infty}^{\infty}\gamma_k\cos(k\lambda)=$

$\frac{1}{2\pi}[\gamma_0+2\sum_{k=1}^{\infty}\gamma_k\cos(k\lambda)]$.

补充：由自回归系数$a_1, a_2, \cdots, a_p$推导Wold系数$\{\psi_k\}$的公式：$\psi_0=1$, $\psi_k=\sum_{j=1}^{p}a_j\psi_{k-j}$, $k\ge 1$.

11. Yule-Walker方程：对$X_t=\sum_{j=1}^{p}a_jX_{t-j}+\varepsilon_t$, $t\in Z$ (p阶自回归模型的基本定义). 则

$$\begin{bmatrix}X_t\\X_{t+1}\\\vdots\\\vdots\\X_{t+n-1}\end{bmatrix}=\begin{bmatrix}X_{t-1}&X_{t-2}&\cdots&X_{t-n}\\X_t&X_{t-1}&\cdots&X_{t+1-n}\\\vdots&\vdots&&\vdots\\\vdots&\vdots&&\vdots\\X_{t+n-2}&X_{t+n-3}&\cdots&X_{t-1}\end{bmatrix}a_n+\begin{bmatrix}\varepsilon_t\\\varepsilon_{t+1}\\\vdots\\\vdots\\\varepsilon_{t+n-1}\end{bmatrix}$$

其中$a_n=(a_{n,1}, a_{n,2}, \cdots, a_{n,n})^T=(a_1, a_2, \cdots, a_p, 0, \cdots, 0)^T$.

1101C-08 201412·2500

定义 $\Gamma_n=\begin{pmatrix}\gamma_0 & \gamma_1 & \cdots & \gamma_{n-1}\\ \gamma_1 & \gamma_0 & \cdots & \gamma_{n-2}\\ \vdots & \vdots & & \vdots\\ \gamma_{n-1} & \gamma_{n-2} & \cdots & \gamma_0\end{pmatrix}$ 和 $\gamma_n=\begin{pmatrix}\gamma_1\\ \gamma_2\\ \vdots\\ \gamma_n\end{pmatrix}$，则有下列Yule-Walker方程：

$\gamma_n=\Gamma_n a_n;\ \gamma_0=\gamma_n^T a_n+\sigma^2,\ n\geq p$ ~~特别地~~，当 $p=2$ 时，方程简化为 $\begin{pmatrix}\gamma_0 & \gamma_1\\ \gamma_1 & \gamma_0\end{pmatrix}\begin{pmatrix}a_1\\ a_2\end{pmatrix}=\begin{pmatrix}\gamma_1\\ \gamma_2\end{pmatrix}$

化简为自相关系数形式即 $\begin{pmatrix}1 & \rho_1\\ \rho_1 & 1\end{pmatrix}\begin{pmatrix}a_1\\ a_2\end{pmatrix}=\begin{pmatrix}\rho_1\\ \rho_2\end{pmatrix}$，得 $\begin{cases}\rho_1=\dfrac{a_1}{1-a_2}\\ \rho_2=\dfrac{a_1^2+a_2(1-a_2)}{1-a_2}\end{cases}$。

12. Levinson 递推公式：若 Γ_{n+1} 正定，对 $1\leq k\leq n$ 有

$$\begin{cases}a_{1,1}=\gamma_1/\gamma_0\\ \sigma_0^2=\gamma_0\\ \sigma_k^2=\sigma_{k-1}^2(1-a_{k,k}^2)\\ a_{k+1,k+1}=\dfrac{\gamma_{k+1}-\gamma_k a_{k,1}-\gamma_{k-1}a_{k,2}-\cdots-\gamma_1 a_{k,k}}{\gamma_0-\gamma_1 a_{k,1}-\gamma_2 a_{k,2}-\cdots\gamma_k a_{k,k}}\\ a_{k+1,j}=a_{k,j}-a_{k+1,k+1}a_{k,k+1-j},\ 1\leq j\leq k.\end{cases}$$

其中 $\sigma_k^2\overset{\text{def}}{=}E(X_{k+1}-a_k^T X_k)^2$ 是用 X_k 预测 X_{k+1} 时的均方误差。

$a_{k,j}\quad a_{k,j+1}$

$a_{k+1,j}\quad a_{k+1-j+1}$

13. AR(1)和AR(2)模型：

对 $|a|<1$，AR(1)模型 $X_t=aX_{t-1}+\varepsilon_t,\ t\in\mathbb{Z}$，$\{\varepsilon_t\}\sim WN(0,\sigma^2)$ 有平稳解 $X_t=\sum_{j=0}^{\infty}a^j\varepsilon_{t-j}$。

自协方差函数 $\gamma_0=\sigma^2\sum_{j=0}^{\infty}a^{2j}=\dfrac{\sigma^2}{1-a^2}$，$\gamma_k=a\gamma_{k-1}=\cdots=a^k\gamma_0$。

自相关系数 $\rho_k=\gamma_k/\gamma_0=a^k$

谱密度 $f(\lambda)=\dfrac{\sigma^2}{2\pi|1-a\exp(i\lambda)|^2}=\dfrac{\sigma^2}{2\pi[1+a^2-2a\cos\lambda]}$，$\lambda\in[-\pi,\pi]$。

AR(2) 模型 $X_t=a_1X_{t-1}+a_2X_{t-2}+\varepsilon_t$，$t\in\mathbb{Z}$。其中 $\{\varepsilon_t\}$ 是 $WN(0,\sigma^2)$。

特征多项式 $A(z)=1-a_1z-a_2z^2$ 的根都在单位圆外的充分必要条件是 a_1,a_2 满足条件 $a_2\pm a_1<1$，$|a_2|<1$。

1101C-08 201412·2500

中国科学技术大学

设 z_1, z_2 是 $A(z)$ 的两个根，则 $|z_1|>1$，$|z_2|>1$。从 Yule-Walker 方程可得

自相关系数 $\rho_0=1$，$\rho_1=\frac{a_1}{1-a_2}$，$\rho_k=a_1\rho_{k-1}+a_2\rho_{k-2}$，$k\geq 1$.

(1) z_1, z_2 是实根 $\Leftrightarrow a_1^2+4a_2\geq 0$

(2) z_1, z_2 是复根 $\Leftrightarrow a_1^2+4a_2<0$

(3) 解 Yule-Walker 方程得到 $a_1=\frac{\rho_1(1-\rho_2)}{1-\rho_1^2}$，$a_2=\frac{\rho_2-\rho_1^2}{1-\rho_1^2}$.

从上式可以反解出 $\rho_1=\frac{a_1}{1-a_2}$，$\rho_2=a_2+\frac{a_1^2}{1-a_2}$。

当 (a_1, a_2) 在 $\mathscr{A}=\{(a_1,a_2)\,|\,a_2\pm a_1<1,\ |a_2|<1\}$ 内变化时，ρ_1, ρ_2 在 $\mathscr{C}=\{(\rho_1,\rho_2)\,|\,\rho_1^2<(1+\rho_2)/2,\ |\rho_1|<1, |\rho_2|<1\}$ 中取值。称 $\mathscr{C}$ 是 AR(2) 的允许域，$\mathscr{A}$ 是 AR(2) 的稳定域.

(4) AR(2) 序列的谱密度是 $f(\lambda)=\frac{\sigma^2}{2\pi}\cdot\frac{1}{|1-a_1e^{i\lambda}-a_2e^{2i\lambda}|^2}$.

(5) 由 Durbin-Levinson 方程组，

中国科学技术大学

14. MA(q)模型和MA(q)序列：设$\{\varepsilon_t\}$是$WN(0,\sigma^2)$，若实数$b_1,b_2,\cdots,b_q(b_q\neq0)$使得$B(z)=1+\sum_{j=1}^{q}b_jz^j\neq0$，$|z|<1$，就称$X_t=\varepsilon_t+\sum_{j=1}^{q}b_j\varepsilon_{t-j}$ (*)，$t\in\mathbb{Z}$是q阶滑动平均模型，简称为MA(q)模型；由$X_t=\varepsilon_t+\sum_{j=1}^{q}b_j\varepsilon_{t-j}$决定的平稳序列$\{X_t\}$是滑动平均序列，简称为MA(q)序列。

如果进一步要求多项式~~零点~~ $B(z)$在单位圆上也没有零点：$B(z)\neq0$当$|z|\le1$，则称(*)是可逆的MA(q)模型，称相应的平稳序列是可逆的MA(q)序列。

15. MA(q)序列$\{X_t\}$的自协方差函数q后截尾：$\gamma_q=\sigma^2b_q$，$\gamma_k=0$，$|k|>q$；（这是MA(q)序列的充要条件）

其谱密度$f(\lambda)=\frac{\sigma^2}{2\pi}|B(e^{i\lambda})|^2=\frac{1}{2\pi}\sum_{k=-q}^{q}\gamma_ke^{-ik\lambda}$，$\lambda\in[-\pi,\pi]$。

16. 可逆MA(1)和MA(2)模型：设$\{\varepsilon_t\}$是$WN(0,\sigma^2)$。

对可逆MA(1)序列$X_t=\varepsilon_t+b\varepsilon_{t-1}$，$\varepsilon_t\sim WN(0,\sigma^2)$，$|b|<1$。自协方差函数$\gamma_0=\sigma^2(1+b^2)$，$\gamma_1=b\sigma^2$和$\gamma_k=0$，$|k|\ge2$。自相关系数$\rho_k=\begin{cases}b/(1+b^2), & |k|=1\\ 0, & |k|>1\end{cases}$

谱密度$f(\lambda)=\frac{\sigma^2}{2\pi}|1+be^{i\lambda}|^2=\frac{\sigma^2}{2\pi}(1+b^2+2b\cos\lambda)$，$\lambda\in[-\pi,\pi]$。

对可逆MA(2)模型$X_t=\varepsilon_t+b_1\varepsilon_{t-1}+b_2\varepsilon_{t-2}$，$t\in\mathbb{Z}$的特征多项式是$B(z)=1+b_1z+b_2z^2\neq0$，$|z|<1$。

(1) 可逆域是$\{(b_1,b_2):B(z)\neq0,|z|\le1\}=\{(b_1,b_2):b_2\pm b_1>-1,|b_2|<1\}$，和AR(2)的平稳域相对应。

1101C-08 201412·2500

中国科学技术大学

(2) 自协方差函数是 $\gamma_0=\sigma^2(1+b_1^2+b_2^2)$, $\gamma_1=\sigma^2(b_1+b_1b_2)$, $\gamma_2=\sigma^2 b_2$, $\gamma_k=0$, $|k|>2$.

(3) 自相关系数是 $\rho_1=\frac{b_1+b_1b_2}{1+b_1^2+b_2^2}$, $\rho_2=\frac{b_2}{1+b_1^2+b_2^2}$, $\rho_k=0$, $|k|>2$.

(4) 谱密度是 $f(\lambda)=\frac{\sigma^2}{2\pi}|1+b_1e^{i\lambda}+b_2e^{i2\lambda}|^2$.

17. ARMA(p,q) 模型及其平稳解：设 $\{\varepsilon_t\}$ 是 $WN(0,\sigma^2)$，实系数多项式 $A(z)$ 和 $B(z)$ 没有公共根，满足 $b_0=1$, $a_p b_q\neq 0$, $A(z)=1-\sum_{j=1}^{p}a_jz^j\neq 0$, $|z|\le 1$; $B(z)=\sum_{j=0}^{q}b_jz^j\neq 0$, $|z|<1$. 就称差分方程 $X_t=\sum_{j=1}^{p}a_jX_{t-j}+\sum_{j=0}^{q}b_j\varepsilon_{t-j}$, $t\in Z$ (*) 是一个自回归滑动平均模型，简称为 ARMA(p,q) 模型. 称满足上式的平稳序列 $\{X_t\}$ 为平稳解或 ARMA(p,q) 序列. 利用推移算子可以改写为 $A(\mathscr{B})X_t=B(\mathscr{B})\varepsilon_t$, $t\in Z$.

18. ARMA(p,q) 序列的 Wold 系数：$\exists\rho>1$ s.t. $\{|z|\le\rho\}$ 内 $A^{-1}(z)B(z)$ 解析，从而有 Taylor 展开式 $\Phi(z)=A^{-1}(z)B(z)=\sum_{j=0}^{\infty}\psi_jz^j$, $|z|\le\rho$. 定义 $\Phi(\mathscr{B})=A^{-1}(\mathscr{B})B(\mathscr{B})=\sum_{j=0}^{\infty}\psi_j\mathscr{B}^j$. 因此 $A(\mathscr{B})X_t=B(\mathscr{B})\varepsilon_t\Rightarrow X_t=A^{-1}(\mathscr{B})B(\mathscr{B})\varepsilon_t=\sum_{j=0}^{\infty}\psi_j\varepsilon_{t-j}$, $t\in Z$. 称 $\{\psi_k\}$ 为 X_t 的 Wold 系数. 上述定义的平稳序列是 ARMA(p,q) 模型 (*) 的唯一平稳解.

19. ARMA(p,q) 序列的自协方差函数可以由 Wold 系数 $\{\psi_j\}$ 表示：$\gamma_k=\sigma^2\sum_{j=0}^{\infty}\psi_j\psi_{j+k}$, $k\ge 0$.
计算 Wold 系数的递推方法：规定 $b_j=0$, $j>q$; $\psi_j=0$, $j<0$, $\psi_j=\begin{cases}1, & j=0\\ b_j+\sum_{k=1}^{p}a_k\psi_{j-k}, & j=1,2,\dots\end{cases}$

把往年卷都掌握!

中国科学技术大学

20. ARMA序列的谱密度和可逆性

$\{\gamma_k\}$绝对可和$\Rightarrow f(\lambda)=\frac{1}{2\pi}\sum_{k=-\infty}^{\infty}\gamma_k e^{-ik\lambda}=\frac{\sigma^2}{2\pi}\left|\sum_{j=0}^{\infty}\psi_j e^{ij\lambda}\right|^2=\frac{\sigma^2}{2\pi}\left|\frac{B(e^{i\lambda})}{A(e^{i\lambda})}\right|^2$.

若$B(z)=1+\sum_{j=1}^{q}b_j z^j\neq 0,\ |z|\leq 1$. 则称ARMA$(p,q)$模型(*)为可逆的ARMA模型,相应的平稳解为可逆的ARMA(p,q)序列.

对于上述可逆的ARMA(p,q)模型,由于$B^{-1}(z)A(z)$在$\{z:|z|\leq\rho\}$ $(\rho>1)$内解析,所以有Taylor展式$B^{-1}(z)A(z)=\sum_{j=0}^{\infty}\varphi_j z^j,\ |z|\leq\rho$. 其中$|\varphi_j|=o(\rho^{-j})$, 当$j\to\infty$.

因此定义$B^{-1}(\mathcal{B})A(\mathcal{B})=\sum_{j=0}^{\infty}\varphi_j\mathcal{B}^j$. 则$\varepsilon_t=B^{-1}(\mathcal{B})A(\mathcal{B})X_t=\sum_{j=0}^{\infty}\varphi_j X_{t-j},\ t\in\mathbb{Z}$.

此式为$X_t=\sum_{j=0}^{\infty}\psi_j\varepsilon_{t-j}$的逆转形式,表明可逆ARMA$(p,q)$序列及其噪声序列可以相互线性表示.

21. ARIMA(p,d,q)模型和SARIMA模型

ARIMA: 设d是一个正整数,如果$Y_t=(1-\mathcal{B})^d X_t=\sum_{k=0}^{d}C_d^k(-1)^k X_{t+k},\ t\in\mathbb{Z}$, 是一个ARMA$(p,q)$序列,且模型MA部分的特征多项式没有等于1的特征根,则称$\{X_t\}$是一个求和ARIMA(p,d,q)序列,简称为ARIMA(p,d,q)序列。于是ARIMA(p,d,q)序列满足的模型是$A(\mathcal{B})(1-\mathcal{B})^d X_t=B(\mathcal{B})\varepsilon_t,\ t\in\mathbb{Z}$. $A(z)=1-\sum_{j=1}^{p}a_j z^j\neq 0,\ |z|\leq 1$; $B(z)=\sum_{j=0}^{q}b_j z^j\neq 0,\ |z|<1$.

SARIMA: 设d和D是非负整数,若差分方程$Y_t\triangleq(1-B)^d(1-B^s)^D X_t$是因果ARMA过程$\phi(B)\Phi(B^s)Y_t=\theta(B)\Theta(B^s)\varepsilon_t$, $\{\varepsilon_t\}\sim WN(0,\sigma^2)$, 则称$\{X_t\}$是周期为$s$的季节ARIMA$(p,d,q)\times(P,D,Q)_s$过程. 其中$\phi(z)=1-\phi_1 z-\cdots-\phi_p z^p$, $\Phi(z)=1-\Phi_1 z-\cdots-\Phi_P z^P$, $\theta(z)=1+\theta_1 z+\cdots+\theta_q z^q$, $\Theta(z)=1+\Theta_1 z+\cdots+\Theta_Q z^Q$.

1101C-08 201412·2500

22. 白噪声检验：为了判断一个平稳序列中是否含有信息，是否可以继续分析，需要对该序列进行白噪声检验。该检验用到的统计量是χ^2分布，原假设H_0：$\{X_t\}$是独立白噪声。

样本自相关置信区间检验：若$\{X_t\}$为独立同分布白噪声时$\sqrt{n}(\hat{\rho}_1,\hat{\rho}_2,\cdots\hat{\rho}_m)$近似服从$N_m(0,I_m)$，则$Pr(\sqrt{n}\,|\hat{\rho}_k|>1.96)\approx 0.05$。若超过5%的$|\hat{\rho}_j|\geq\frac{1.96}{\sqrt{n}}$，否定$H_0$。

23. 最佳线性预测：设Y和X_j $(1\leq j\leq n)$是均值为0、方差有限的随机变量，记$X=(X_1,\cdots,X_n)^T$，若$a\in\mathbb{R}^n$（$a=(a_1,\cdots,a_n)\in\mathbb{R}^n$），使得对任何$b\in\mathbb{R}^n$都有$E(Y-a^TX)^2\leq E(Y-b^TX)^2$，则称$a^TX$是用$X_1,\cdots,X_n$对$Y$进行预测的最佳线性预测，记作$\hat{Y}$或$L(Y|X)$，即$L(Y|X)=a^TX$。当$\hat{Y}=a^TX$时，$Y-\hat{Y}=Y-a^TX$为预测误差，$E(Y-\hat{Y})^2=E(Y-a^TX)^2$是预测的均方误差。若$EY=b$，$EX=\mu$，则$L(Y|X)=L(Y-b|X-\mu)+b$。

以下用$L_n=\overline{sp}(X_1,\cdots X_n)$表示随机r.v. $X_1,X_2,\cdots,X_n$的线性组合全体。

24. 决定性与非决定性平稳序列：设$X_{n,m}$ $\{X_t\}$是零均值平稳序列，记$X_{n,m}=(X_n,X_{n-1},\cdots,X_{n-(m-1)})^T$，$m$表示向量维数。定义$\hat{X}_{n+1,m}=L(X_{n+1}|X_{n,m})$，定义$\sigma_1^2=\lim\limits_{m\to\infty}\sigma_{1,m}^2<\infty$（$\sigma_{1,m}^2=E(X_{n+1}-\hat{X}_{n+1,m})^2$）。若$\sigma_1^2=0$，称$\{X_t\}$是决定性平稳序列；若$\sigma_1^2>0$，称$\{X_t\}$是非决定性平稳序列，并且称$\sigma_1^2=\lim\limits_{m\to\infty}\sigma_{1,m}^2$为$\{X_t\}$的一步预测的均方误差。设$\{X_t\}$是非决定性的平稳序列，如果$\lim\limits_{k\to\infty}\sigma_k^2=\gamma_0$，就称$\{X_t\}$是纯非决定性的。

（怎么写？见试卷）

中国科学技术大学

25. 平稳序列的递推序列：设$\{X_t\}$是零均值平稳序列，对任何$n\in\mathbb{N}_+$，自协方差矩阵Γ_n正定，则最佳线性预测$\hat{X}_{n+1}\overset{\text{def}}{=}L(X_{n+1}|\vec{X}_n)=\sum_{j=1}^{n}\theta_{n,j}Z_{n+1-j}$，$n=1,2,\cdots$。其中$\vec{X}_n=(X_1,X_2,\cdots,X_n)^T$，$Z_n=X_n-L(X_n|\vec{X}_{n-1})$，系数$\{\theta_{n,j}\}$和预测的均方误差$\nu_n=EZ_{n+1}^2$满足如下的递推公式

$$\begin{cases}\nu_0=\gamma_0,\\ \theta_{n,n-k}=[\gamma_{n-k}-\sum_{j=0}^{k-1}\theta_{k,k-j}\theta_{n,n-j}\nu_j]/\nu_k,\quad 0\le k\le n-1,\\ \nu_n=\gamma_0-\sum_{j=0}^{n-1}\theta_{n,n-j}^2\nu_j,\end{cases}$$

其中$\sum_{j=0}^{-1}(\cdot)\overset{\text{def}}{=}0$。递推的顺序是

$\nu_0\to\theta_{1,1}\to\nu_1\to\theta_{2,2},\theta_{2,1}\to\nu_2\to\theta_{3,3},\theta_{3,2},\theta_{3,1}\to\nu_3\to\cdots$

$Z_n=X_n-L(X_n|\vec{X}_{n-1})$和$\vec{X}_{n-1}$正交，是不能被$\vec{X}_{n-1}$包含的信息。因此$Z_n$称为样本新息。

26. AR(p)序列的预测：设$\{X_t\}$满足AR(p)序列$X_t=a_1X_{t-1}+\cdots+a_pX_{t-p}+\varepsilon_t=\sum_{j=1}^{p}a_jX_{t-j}+\varepsilon_t$，$t\in\mathbb{Z}$。其中$\{\varepsilon_t\}$是白噪声。特征多项式$A(z)=1-\sum_{j=1}^{p}a_jz^j\neq0$，$|z|\le1$。

(i) 用$\vec{X}_n$预测X_{n+1}：1° 对于$1\le n\le p-1$，$\hat{X}_{n+1}=L(X_{n+1}|\vec{X}_n)=\vec{\gamma}_n^T\Gamma_n^{-1}\vec{X}_n$，$\Gamma_n$是$\{X_t\}$的n阶自协方差矩阵，$\vec{\gamma}_n=(\gamma_n,\gamma_{n-1},\cdots,\gamma_1)^T$。预测的均方误差为$E(X_{n+1}-\hat{X}_{n+1})^2=\gamma_0-\vec{\gamma}_n^T\Gamma_n^{-1}\vec{\gamma}_n$。

2° 对于$n\ge p$，$\hat{X}_{n+1}=L(X_{n+1}|X_n,X_{n-1},\cdots,X_{n-p+1})=\sum_{j=1}^{p}a_jX_{n+1-j}$。

(ii) 用$\vec{X}_n$预测X_{n+k}：$L(X_{n+k}|\vec{X}_n)=\sum_{j=1}^{p}a_jL(X_{n+k-j}|X_n,X_{n-1},\cdots,X_{n-p+1})$。

27. MA(q)序列的预测：$\{\varepsilon_t\}$是WN$(0,\sigma^2)$. 实系数多项式 $B(z)=1+b_1z+\cdots+b_qz^q$ 在单位圆内无根：$B(z)\neq 0$, $|z|<1$. $\{X_t\}$满足MA(q)模型 $X_t=B(\mathcal{B})\varepsilon_t$, $t\in\mathbb{Z}$. $L_n=\overline{sp}\{X_1,X_2,\cdots,X_n\}=\overline{sp}\{\hat{\varepsilon}_1,\hat{\varepsilon}_2,\cdots,\hat{\varepsilon}_n\}$. $\hat{\varepsilon}_n=X_n-L(X_n|\vec{X}_{n-1})$ 是$\{X_t\}$的逐步预测误差，$\vec{X}_n=(X_1,\cdots,X_n)^T$. 以下 $n\geq q$.

对 X_{n+1} 的预测：$L(X_{n+1}|\vec{X}_n)=\sum_{j=1}^{q}\theta_{n,j}\hat{\varepsilon}_{n+1-j}$，预测的均方误差为 $\nu_n=E\hat{\varepsilon}_{n+1}^2=\gamma_0-\sum_{j=1}^{q}\theta_{n,j}^2\nu_{n-j}$，$\theta_{n,j}$, ν_n 由25中表达式递推计算.

对 X_{n+k+1} 的预测：
$$L(X_{n+k+1}|X_n)=\begin{cases}\sum_{j=k+1}^{q}\theta_{n+k,j}\hat{\varepsilon}_{n+k+1-j}, & 1\leq k\leq q-1,\\ 0, & k\geq q\end{cases}$$
预测的均方误差 $E[X_{n+k+1}-L(X_{n+k+1}|\vec{X}_n)]^2=\gamma_0-\sum_{j=k+1}^{q}\theta_{n+k,j}^2\nu_{n+k-j}$，$1\leq k\leq q-1$，其中$\theta_{n+k,j}$，$\nu_{n+k-j}$ 由25中表达式递推得到.

28. ARMA(p,q)序列的预测：对ARMA(p,q)序列$\{X_t\}$，X_t满足$A(\mathcal{B})X_t=B(\mathcal{B})\varepsilon_t$，$A(z)=1-a_1z-\cdots-a_pz^p$，$B(z)=1+b_1z+\cdots+b_qz^q$. 定义$\{X_t\}$的逐步预测误差 $Z_t=X_t-L(X_t|\vec{X}_{t-1})$，$Z_1=X_1$，则
$$\hat{X}_{n+1}=\begin{cases}\sum_{j=1}^{n}\theta_{n,j}Z_{n+1-j}, & 1\leq n\leq m\\ \sum_{j=1}^{p}a_jX_{n+1-j}+\sum_{j=1}^{q}\theta_{n,j}Z_{n+1-j}, & n\geq m\end{cases}$$
，预测的均方误差 $EZ_{n+1}^2=\sigma^2\nu_n$.

$\sqrt{N}(\hat{a}_1-a_1,\cdots,\hat{a}_p-a_p)^T \xrightarrow{P} N(0,\sigma^2\Gamma_p^{-1})$

J 29. AR(p)模型的Yule-Walker估计: AR(p)模型 $X_t=\sum_{j=1}^{p}a_jX_{t-j}+\varepsilon_t,\ t\geq 0$; $\sigma^2=E\varepsilon_t^2$

p已知，考虑回归系数 $\vec{a}=(a_1,a_2,\cdots,a_p)^T$ 和零均值白噪声 $\{\varepsilon_t\}$ 的方差 σ^2 估计问题.

由观测样本 $x_1,x_2,\cdots,x_N$ 可以构造出样本自协方差函数的估计: $\hat{\gamma}_k=\frac{1}{N}\sum_{j=1}^{N-k}y_jy_{j+k}$, $k=0,1,2,\cdots,p$. 其中 $y_t=x_t-\bar{x}_N,\ t=1,2,\cdots,N$, $\bar{x}_N=\frac{1}{N}\sum_{j=1}^{N}x_j$. 因此AR(p)的自回归系数和白噪声方差的矩估计 $(\hat{a}_1,\hat{a}_2,\cdots,\hat{a}_p)^T$, $\hat{\sigma}^2$ 由样本Yule-Walker方程

$$\begin{bmatrix}\hat{\gamma}_1\\ \hat{\gamma}_2\\ \vdots\\ \hat{\gamma}_p\end{bmatrix}=\begin{bmatrix}\hat{\gamma}_0 & \hat{\gamma}_1 & \cdots & \hat{\gamma}_{p-1}\\ \hat{\gamma}_1 & \hat{\gamma}_0 & \cdots & \hat{\gamma}_{p-2}\\ \vdots & \vdots & & \vdots\\ \hat{\gamma}_{p-1} & \hat{\gamma}_{p-2} & \cdots & \hat{\gamma}_0\end{bmatrix}\begin{bmatrix}\hat{a}_1\\ \hat{a}_2\\ \vdots\\ \hat{a}_p\end{bmatrix}$$

和 $\hat{\sigma}^2=\hat{\gamma}_0-(\hat{a}_1\hat{\gamma}_1+\cdots+\hat{a}_p\hat{\gamma}_p)$ 决定.

另外，还可以采用 Levinson递推方法:

$$\begin{cases}\hat{\sigma}_0^2=\hat{\gamma}_0\\ \hat{a}_{1,1}=\hat{\gamma}_1/\hat{\sigma}_0^2\\ \hat{\sigma}_k^2=\hat{\sigma}_{k-1}^2(1-\hat{a}_{k,k}^2)\\ \hat{a}_{k+1,k+1}=\dfrac{\hat{\gamma}_{k+1}-\hat{\gamma}_k\hat{a}_{k,1}-\hat{\gamma}_{k-1}\hat{a}_{k,2}-\cdots-\hat{\gamma}_1\hat{a}_{k,k}}{\hat{\gamma}_0-\hat{\gamma}_1\hat{a}_{k,1}-\hat{\gamma}_2\hat{a}_{k,2}-\cdots-\hat{\gamma}_k\hat{a}_{k,k}}\\ \hat{a}_{k+1,j}=\hat{a}_{k,j}-\hat{a}_{k+1,k+1}\hat{a}_{k,k+1-j},\ 1\leq j\leq k,\ k\leq p.\end{cases}$$

递推最后得到矩估计 $(\hat{a}_1,\hat{a}_2,\cdots,\hat{a}_p)=(\hat{a}_{p,1},\hat{a}_{p,2},\cdots,\hat{a}_{p,p})$, $\hat{\sigma}^2=\hat{\sigma}_p^2$.

上述矩估计又被称为Yule-Walker估计.

以下写出带"J"标号部分的简化形式(即n,p取1,2等较小数值的情形)

1101C-08 201412·2500

12. 解: Levinson 递推公式: 当 n 取2时, 对 $1\le k\le 2$ 有

$$\begin{cases} a_{1,1}=\gamma_1/\gamma_0 \\ \sigma_0^2=\gamma_0 \\ \sigma_1^2=\sigma_0^2(1-a_{1,1}^2)=\dfrac{(\gamma_0-\gamma_1)^2}{\gamma_0}. \end{cases}$$

其他系数具体推导如下. $a_{2,2}=\dfrac{\gamma_2-\gamma_1 a_{1,1}}{\gamma_0-\gamma_1 a_{1,1}}=\dfrac{\gamma_2\gamma_0-\gamma_1^2}{\gamma_0^2-\gamma_1^2}$, $a_{2,1}=a_{1,1}-a_{2,2}\cdot a_{1,1}=$
$\gamma_1/\gamma_0-\dfrac{\gamma_2\gamma_0-\gamma_1^2}{\gamma_0^2-\gamma_1^2}\cdot\dfrac{\gamma_1}{\gamma_0}=\dfrac{\gamma_1}{\gamma_0}\cdot\dfrac{\gamma_0^2-\gamma_2\gamma_0}{\gamma_0^2-\gamma_1^2}=\dfrac{\gamma_1(\gamma_0-\gamma_2)}{\gamma_0-\gamma_1}$, $\sigma_2^2=\sigma_1^2(1-a_{2,2}^2)=\dfrac{(\gamma_0-\gamma_1)^2}{\gamma_0}\cdot$

$\left[1-\dfrac{(\gamma_1^2-\gamma_2\gamma_0)^2}{(\gamma_0^2-\gamma_1^2)^2}\right]=\dfrac{\gamma_0-\gamma_2}{\gamma_0-\gamma_1}\cdot\dfrac{\gamma_0^2-\gamma_1^2+\gamma_2\gamma_0-\gamma_1^2}{\gamma_0+\gamma_1}$, $a_{3,3}=\dfrac{\gamma_3-\gamma_2 a_{2,1}-\gamma_1 a_{2,2}}{\gamma_0-\gamma_1 a_{2,1}-\gamma_2 a_{2,2}}$

$$=\frac{\gamma_3-\gamma_2\cdot\dfrac{\gamma_1(\gamma_0-\gamma_2)}{\gamma_0-\gamma_1}-\gamma_1\cdot\dfrac{\gamma_2\gamma_0-\gamma_1^2}{\gamma_0-\gamma_1}}{\gamma_0-\gamma_1\cdot\dfrac{\gamma_1(\gamma_0-\gamma_2)}{\gamma_0-\gamma_1}-\gamma_2\cdot\dfrac{\gamma_2\gamma_0-\gamma_1^2}{\gamma_0-\gamma_1}}=\frac{\gamma_0(\gamma_3-2\gamma_1\gamma_2)+\gamma_1(\gamma_1^2+\gamma_2^2-\gamma_3)}{\gamma_0(\gamma_0-\gamma_1-\gamma_2^2-\gamma_1^2)+2\gamma_1^2\gamma_2}$$

$$a_{3,2}=a_{2,2}-a_{3,3}\cdot a_{2,1}=\frac{\gamma_2\gamma_0-\gamma_1^2}{\gamma_0^2-\gamma_1^2}-\frac{\gamma_0(\gamma_3-2\gamma_1\gamma_2)+\gamma_1(\gamma_1^2+\gamma_2^2-\gamma_3)}{\gamma_0(\gamma_0-\gamma_1-\gamma_2^2-\gamma_1^2)+2\gamma_1^2\gamma_2}\cdot\frac{\gamma_1(\gamma_0-\gamma_2)}{\gamma_0-\gamma_1},$$

$a_{3,1}=a_{2,1}-a_{3,3}\cdot a_{3,2}$ (下略).

19. ARMA(p,q) 序列: $X_t-a_1X_{t-1}-a_2X_{t-2}-\cdots-a_pX_{t-p}=\varepsilon_t+b_1\varepsilon_{t-1}+b_2\varepsilon_{t-2}+\cdots+b_q\varepsilon_{t-q}$
是一个滑动自回归平均model, 简称为ARMA(p,q)模型, 称满足上式的平稳序列 $\{X_t\}$ 为ARMA(p,q)序列.

$\exists\rho>1$ s.t. $A^{-1}(z)B(z)$ 在 $|z|\le\rho$ 内解析, 从而有Taylor展式 $\Phi(z)=\sum_{j=0}^{\infty}\psi_j z^j$, $|z|\le\rho$.
ψ_j 称为 $\{X_t\}$ 的Wold系数.

1101C-08 201412·2500

中国科学技术大学

$\to X_t = a_1X_{t-1}+a_2X_{t-2}+b_0\varepsilon_t+b_1\varepsilon_{t-1}+b_2\varepsilon_{t-2}$

19. 计算Wold系数的递推方法：对ARMA(2,2)模型，即$p=q=2$.

则$\psi_j=1$，$j=0$. $j=1$时，$\psi_j = b_j + \sum_{k=1}^{p} a_k\psi_{j-k} = b_1 + \sum_{k=1}^{2} a_k\psi_{1-k} = b_1 + a_1\cdot\psi_0 = b_1+a_1$.

$j=2$时，$\psi_j = b_2+\sum_{k=1}^{2} a_k\psi_{2-k} = b_2+a_1\psi_1+a_2\psi_0 = b_2+a_1(b_1+a_1)+a_2 = a_1(a_1+b_1)+a_2+b_2$.

若$j\geq 3$，则$\psi_j = a_1\psi_{j-1}+a_2\psi_{j-2}$.

对更简单的ARMA(2,1)模型，$p=2$，$q=1$. $\psi_j=1$，$j=0$. $\psi_1=b_1+a_1\psi_0=a_1+b_1$，

对$j=2$，$\psi_2=b_2+\sum_{k=1}^{2} a_k\psi_{2-k}=a_1\psi_1+a_2\psi_0$. 而对$j\geq 3$. $\psi_j=a_1\psi_{j-1}+a_2\psi_{j-2}$.

最简单的ARMA(1,1)模型，$p=1$，$q=1$. $\psi_0=1$，$\psi_1=b_1+a_1\psi_{1-1}=b_1+a_1\psi_0=b_1+a_1$，

$\psi_j=a_1\psi_{j-1}=\cdots=a_1^{j-1}\psi_1=a_1^{j-1}(a_1+b_1)$，$j=2,3,4,\cdots$

25. 对$n=2$的情形：$\hat{X}_3 \overset{def}{=} L(X_3|\vec{X}_2) = L(X_3|X_1,X_2) = \sum_{j=1}^{2} \theta_{n,j}Z_{3-j}$，$Z_2=X_2-L(X_2|X_1)$. $\{\theta_{n,j}\}$，$\nu_n=EZ_{n+1}^2$满足如下公式：

$\nu_0=\gamma_0$，$\theta_{1,1}=\gamma_1/\nu_0=\gamma_1/\gamma_0$，$\nu_1=\gamma_0-\theta_{1,1}^2\nu_0=\frac{\gamma_0^2-\gamma_1^2}{\gamma_0}$，$\theta_{2,2}=\gamma_2/\nu_0=\gamma_2/\gamma_0$，$\theta_{2,1}=(\gamma_1-\theta_{1,1}\theta_{2,2}\nu_0)/\nu_1=\gamma_1(\gamma_0-\gamma_2)/(\gamma_0^2-\gamma_1^2)$. $\nu_2=\gamma_0-\sum_{j=0}^{1}\theta_{2,2-j}^2\nu_j=\gamma_0-(\theta_{2,2}^2\nu_0+\theta_{2,1}^2\nu_1)$

$=\gamma_0-\frac{\gamma_2^2}{\gamma_0}+\frac{\gamma_1^2(\gamma_0-\gamma_2)^2}{(\gamma_0^2-\gamma_1^2)\gamma_0}$.

26. AR(p)模型的预测：以下讨论$p=1$和$p=2$两种特殊情况.

(1) AR(1)模型：$X_t-a_1X_{t-1}=\varepsilon_t$，$|a_1|<1$，$\varepsilon_t\sim WN(0,\sigma^2)$，$\gamma_0=\frac{\sigma^2}{1-a_1^2}$，$\gamma_k=a_1^k\gamma_0$.

$\rho_k=\frac{\gamma_k}{\gamma_0}=a_1^k$.

1101C-08 201412·2500

模型　　　　　　　　　　　　　　　　　　　　A.4

中国科学技术大学

(何书元书方法废弃，改用李东风《金融时序讲义》)

AR(p)模型：$x_t=\phi_0+\phi_1 x_{t-1}+\cdots+\phi_p x_{t-p}+\varepsilon_t$，在时刻h，用截止到h时刻的信息(即$x_1,\cdots,x_h$已知)预测$x_{h+l}$，得到的最佳预测记为$\hat{x}_h(l)$。

对$l=1$，因为$x_{h+1}=\phi_0+\phi_1 x_h+\cdots+\phi_p x_{h+1-p}+\varepsilon_{h+1}$；以及$E(\varepsilon_{h+1}|x_1,\cdots,x_h)=0$。因此有$\hat{x}_h(1)=E(x_{h+1}|x_1,\cdots,x_h)=\phi_0+\phi_1 x_h+\cdots+\phi_p x_{h+1-p}$。一步预测误差为$e_h(1)=x_{h+1}-\hat{x}_h(1)=\varepsilon_{h+1}$，$Var(e_h(1))=\sigma^2$。若$\varepsilon_t$服从正态分布，则$x_{h+1}$的超前一步预测的95%预测区间为$\hat{x}_h(1)\pm1.96\sigma$。

对$l=2$，注意到$x_{h+2}=\phi_0+\phi_1 x_{h+1}+\cdots+\phi_p x_{h+2-p}+\varepsilon_{h+2}$，因此$\hat{x}_h(2)=E(x_{h+2}|x_1,\cdots,x_h)$
$=\phi_0+\phi_1 E(x_{h+1}|x_1,\cdots,x_h)+\phi_2 x_h+\cdots+\phi_p x_{h+2-p}=\phi_0+\phi_1\hat{x}_h(1)+\phi_2 x_h+\cdots+\phi_p x_{h+2-p}$。预测误差为$e_h(2)=x_{h+2}-\hat{x}_h(2)=\phi_1 e_h(1)+\varepsilon_{h+2}$。预测均方误差为$E(e_h(2))^2=Var(e_h(2))=\sigma^2(1+\phi_1^2)$ ($\varepsilon_{h+2}\perp e_h(1)=x_{h+1}-\hat{x}_h(1)$)。

对一般的$l\in N^+$，有$\hat{x}_h(l)=\phi_0+\sum_{j=1}^{p}\phi_j\hat{x}_h(l-j)$，其中$\hat{x}_h(i)=\begin{cases}x_{h+i}, & i\le 0\\ \hat{x}_h(i), & i>0\end{cases}$

为计算$\hat{x}_h(l)$，只要递推计算$\hat{x}_h(1),\cdots,\hat{x}_h(l-1)$，再按上面的公式得到$\hat{x}_h(l)$。超前$l$步的预测误差为$e_h(l)=x_{h+l}-\hat{x}_h(l)$。对平稳AR(p)模型，当超前步数$l\to+\infty$时，$\hat{x}_h(l)\to\mu$。这种~~形式~~性质称为均值回转。

对AR(1)模型的零均值平稳解$\{x_t\}$，$\hat{x}_h(l)=\phi_1^l x_h$。

MA(q)模型：$X_t=\theta_0+\varepsilon_t+\sum_{k=1}^{q}\theta_k\varepsilon_{t-k}$。由于MA(q)序列在间隔超过q步后独立，因此超前多步预测只能预测到q步，从q+1步开始就只能用均值θ_0预测。

例1：对MA(1)模型$X_t=\theta_0+\varepsilon_t+\theta_1\varepsilon_{t-1}$，$\{\varepsilon_t\}\sim WN(0,\sigma^2)$.

对$l=1$，$\hat{X}_h(1)=E(X_{h+1}|X_1,\cdots,X_h)=\theta_0+\theta_1\varepsilon_h$，这里利用$E(\varepsilon_{h+1}|X_1,\cdots,X_h)=0$.

$e_h(1)=X_{h+1}-\hat{X}_h(1)=\varepsilon_{h+1}$，$var\, e_h(1)=\sigma^2$.

对$l=2$，$\hat{X}_h(2)=E(\theta_0+\varepsilon_{h+2}+\theta_1\varepsilon_{h+1}|X_1,\cdots,X_h)=\theta_0$. 从两步开始的超前多步预测就变成$EX_t=\theta_0$. 因此对$k\geq 2$，$e_h(k)=X_{h+k}-\hat{X}_h(k)=\varepsilon_{h+k}+\theta_1\varepsilon_{h+k-1}$，$var(e_h(k))=(1+\theta_1^2)\sigma^2$。

e.g.2：对MA(2)模型$X_t=\theta_0+\varepsilon_t+\theta_1\varepsilon_{t-1}+\theta_2\varepsilon_{t-2}$，$\hat{X}_h(1)=\theta_0+\theta_1\varepsilon_h+\theta_2\varepsilon_{h-1}$ ($E(\varepsilon_{h+1}|X_h,\cdots,X_1)=0$).

$\hat{X}_h(2)=\theta_0+\theta_2\varepsilon_h$ ($E(\varepsilon_{h+1}|X_h,\cdots,X_1)=E(\varepsilon_{h+2}|X_h,\cdots,X_1)=0$). 对$l\geq 3$，$\hat{X}_h(l)=\theta_0=EX_t$. 因此$e_h(1)=X_{h+1}-\hat{X}_h(1)=\varepsilon_{h+1}$，$E(X_{h+1}-\hat{X}_h(1))^2=var(e_h(1))=\sigma^2$；

$e_h(2)=X_{h+2}-\hat{X}_h(2)=\varepsilon_{h+2}+\theta_1\varepsilon_{h+1}$，$E[X_{h+2}-\hat{X}_h(2)]^2=var(e_h(2))=(1+\theta_1^2)\sigma^2$.

对$l\geq 3$，$e_h(l)=X_{h+l}-\hat{X}_h(l)=\varepsilon_{h+l}+\theta_1\varepsilon_{h+l-1}+\theta_2\varepsilon_{h+l-2}$，$E(X_{h+l}-\hat{X}_h(l))^2=var(e_h(l))=(1+\theta_1^2+\theta_2^2)\sigma^2$.

对MA(q)模型$X_t=\theta_0+\varepsilon_t+\sum_{k=1}^{q}\theta_k\varepsilon_{t-k}$，可由归纳法得到如下结论：

(i) 对$1\leq l\leq q$，$\hat{X}_h(l)=\theta_0+\sum_{k=l}^{q}\theta_k\varepsilon_{h+l-k}$；对$l>q$，$\hat{X}_h(l)=\theta_0$；

(ii) $e_h(l)=X_{h+l}-\hat{X}_{h+l}=X_{h+l}-\hat{X}_h(l)$. 因此对$1\leq l\leq q$，$e_h(l)=\varepsilon_{h+l}+\sum_{k=1}^{l-1}\theta_k\varepsilon_{h+l-k}$（定义$\sum_{k=1}^{0}(\cdot)=0$）；对$l>q$，$e_h(l)=\varepsilon_{h+l}+\sum_{k=1}^{q}\theta_k\varepsilon_{h+l-k}$；

中国科学技术大学

(iii) $E(X_{n+l}-\hat{X}_{n+l})^2=var(e_n(l))$。因此对 $1\le l\le q$，$var(e_n(l))=var(\varepsilon_{n+l}+\sum_{k=1}^{l-1}\theta_k\varepsilon_{n+l-k})$
$=\sigma^2(1+\sum_{k=1}^{l-1}\theta_k^2)$（定义 $\sum_{k=1}^{l-1}(\cdot)=0$，当 $l<1$ 时）；对 $l>q$，$var(e_n(l))=var(\varepsilon_{n+l}+\sum_{k=1}^{q}\theta_k\varepsilon_{n+l-k})$
$=\sigma^2(1+\sum_{k=1}^{q}\theta_k^2)$.　（2018~2019，第一学期，第四题）.

ARMA(p,q) 模型：ARMA(p,q) 的预测都有关系，对 AR 部分仍是递推地向前预测，对 MA 部分需估计新息 ε_t 的值。

对 ARMA 模型 $X_t-\sum_{k=1}^{p}a_kX_{t-k}=\varepsilon_t+\sum_{j=1}^{q}b_j\varepsilon_{t-j}$，$\{\varepsilon_j\}\sim WN(0,\sigma^2)$. （$A(B)X_t=B(B)\varepsilon_t$，$A(z)=1-\sum_{j=1}^{p}a_jz^j$，$B(z)=1+\sum_{j=1}^{q}b_jz^j$）

(i) 超前一步预测：$\hat{x}_n(1)=\phi_0+\sum_{j=1}^{p}\phi_jx_{n+1-j}+\sum_{j=1}^{q}\theta_j\hat{\varepsilon}_{n+1-j}$，其中 $\hat{\varepsilon}_i=E(\varepsilon_i|x_1,\cdots,x_n)$

(ii) 对超前 k 步预测：$\hat{x}_n(k)=\phi_0+\sum_{j=1}^{p}\phi_j\hat{x}_{n+k-j}+\sum_{j=1}^{q}\theta_j\hat{\varepsilon}_{n+k-j}$，其中 $n+k-j\le n$ 时 $\hat{x}_{n+k-j}=x_{n+k-j}$，$n+k-j>n$ 时 $\hat{\varepsilon}_{n+k-j}=0$。预测误差为 $e_n(k)=x_{n+k}-\hat{x}_{n+k}$.