随机过程B

刘 杰

Email: jiel@ustc.edu.cn

第四章 平稳过程

平稳过程是指概率性质在时间平移下不变的随机过程. 在通讯、天文学、生物学和经济学领域获得了广泛的应用。

§ 4.1 定义和例子

在本章中，我们总是假定参数集T对加法封闭. 通常T取如下几种集合之一:

(1) $T=\{0,1,2,\cdots\}$

(2) $T=\{0,\pm 1,\pm 2,\cdots\}$

(3) $T=\{t:t\geq 0\}$

(4) $T=\{t:-\infty<t<+\infty\}$

定义4.1 设$X=\{X(t), t\in T\}$为一随机过程，若对任意正整数k及T中任意k个时刻 $t_1<t_2<\cdots<t_k$ 及T中的h，有

$$\{X(t_1),\cdots,X(t_k)\}\overset{d.}{=}\{X(t_1+h),\cdots,X(t_k+h)\}$$

则随机过程X称为**严格平稳过程**。这里 "$d.$"表示等式两边k维随机向量的分布相同。

注： 如果T是离散时间，且$X=\{X(t), t\in T\}$是一个严格平稳过程，则X称为**严格平稳序列**。

设$X=\{X(t), t\in T\}$是一个严格平稳过程。若**期望函数**$E[X(t)]$存在, 则必为常数, 即:

$$m_X(t)=E[X(t)]=m, \quad t\in T.$$

若**方差函数**存在, 则也必为常数, 即:

$$Var(X(t))=E[X(t)-m]^2=\sigma^2, \quad t\in T.$$

协方差函数仅与时间差有关, 即:

$$R_X(h)=R_X(t+h,t)=E[(X(t+h)-m)(X(t)-m)], \quad t\in T.$$

注: $Var(X(t))=R_X(0).$

设$X=\{X(t), t\in T\}$是一个严格平稳过程。若$E[X(t)X(t+h)]$存在, 则起点t无关, 即:

$$r(h)=E[X(t)X(t+h)], \quad t\in T.$$

称为**自相关函数**;

$$\rho_X(h)=R_X(h)/\sigma^2, \quad t\in T.$$

称为**标准自相关函数**.

注: 严格平稳过程是一个相当严格的条件, 在实际中也很难验证. 而在实际物理系统来看, 过程的前二阶矩已能反映该过程的许多主要性质. 因此, 我们引入宽平稳过程.

定义4.2 设$X=\{X(t), t\in T\}$为一实值随机过程，若对任意的$t\in T$, $E[X^2(t)]$存在，$E[X(t)]=m$以及协方差函数$E[(X(t)-m)(X(s)-m)]$仅与$t-s$有关，则称随机过程X为**宽平稳过程**，简称**平稳过程**。

注: 严格平稳过程与宽平稳过程是互不包含的. 但是只要过程的二阶矩存在，则严格平稳过程一定是宽平稳过程.

例4.1 设Z为非零随机变量，考虑如下两个随机过程：$X=\{X(t)=tZ,\ t\in T\}$和$Y=\{Y(t)=Z,\ t\in T\}$．试考察这两个随机过程的平稳性.

解: 因$Y(t)$与时间t无关, 故Y必然是严格平稳过程. 如果还有$E(Z^2)<\infty$, Y也是宽平稳过程.

由于X的一维分布$X(t)$与时间t有关，故X必然不是严格平稳过程．又因为

$$E[X(t)X(s)]=tsE(Z^2)$$

与ts有关, 因此X也不是宽平稳过程.

例4.2 $X=\{X(t),\ t\in(-\infty,\infty)\}$是一个零均值的平稳过程，且不恒等于一个随机变量，问$\{X(t)+X(0),\ t\in(-\infty,\infty)\}$是否仍然是平稳过程.

解: 由题设知:

$$E[X(t)]=0, \qquad E[X(t+h)X(t)]=R_X(h).$$

设$Y(t)=X(t)+X(0)$, 则

$$E[Y(t)]=0,$$

并且

$$E[Y(t+h)Y(t)]=E[(X(t+h)+X(0))(X(t)+X(0))]$$

$$=E[X(t+h)X(t)]+E[X(t+h)X(0)]$$

$$+E[X(t)X(0)]+E[X^2(0))]$$

$$=R_X(h)+R_X(t+h)+R_X(t)+R_X(0)$$

假设 $Y(t)=X(t)+X(0)$是一个平稳过程，则存在$C(h)$使得，对任意的t和h均有

$$R_X(t+h)+R_X(t) = C(h)$$

取$t=(m\text{-}1)h$得

$$R_X(mh)+R_X((m\text{-}1)h)=C(h).$$

因此，

$$R_X(mh)=\begin{cases} C(h)-R_X(0), & m\text{为奇数} \\ R_X(0), & m\text{为偶数} \end{cases}$$

由h的任意性得，

$$C(h)=R_X(0),\quad h\geq 0.$$

因此，对任意的t，$R_X(t)=R_X(0)$，从而

$$Var(X(t)-X(0))\equiv 0,$$

即: $X(t)$几乎恒等于一个随机变量$X(0)$.

例4.3 设有随机过程$Z(t)=Xsint+Ycost$, 其中X和Y是相互独立的随机变量, 它们都分别以2/3和1/3的概率取值-1和2, 讨论$Z(t)$的平稳性.

解: 由题设易求出:

$$E(X)=E(Y)=0, \quad E(X^2)=E(Y^2)=2, \quad E(XY)=0.$$

因此,

$$m_Z(t)=E(X)sint+E(Y)cost=0,$$

$$R_Z(t+h,t)$$

$$=E[(X\sin(t+h)+Y\cos(t+h))(X\sin t+Y\cos t)]$$

$$=E(X^2)\sin(t+h)\sin t+E(Y^2)\cos(t+h)\cos t$$

$$+E(XY)(\sin(t+h)\cos t+\cos(t+h)\sin t)$$

$$=2\cos(h)$$

所以, $Z(t)$是宽平稳过程. 又因为

$$E[Z^3(t)]=2(\sin^3 t+\cos^3 t)$$

所以, $Z(t)$不是严格平稳过程.

练习 设$\{A_n, 1\le n\le N\}$和$\{B_n, 1\le n\le N\}$是两个互不相关的实随机变量序列, 且

$$E(A_n)=E(B_n)=0, \qquad 1\le n\le N,$$

$$E(A_nB_m)=0, \qquad 1\le n,m\le N,$$

$$E(A_nA_m)=E(B_nB_m)=\delta_{nm}\sigma^2, \qquad 1\le n,m\le N.$$

又设$c_1, c_2, \ldots, c_N$是任意正数, 证明:

$$X(t)=\sum_{n=1}^{N}[A_n\cos(c_nt)+B_n\sin(c_nt)], \quad -\infty<t<\infty$$

是平稳过程.

证明: 因为

$$E[X(t)]=0$$

$$E[X(t)X(s)]=\sigma^2\sum_{n=1}^{N}[\cos(c_n t)\cos(c_n s)+\sin(c_n t)\sin(c_n s)]$$

$$=\sigma^2\sum_{n=1}^{N}\cos[c_n(t-s)]$$

所以$\{X(t),-\infty<t<\infty\}$是平稳过程.

定义4.3 设G={G(t), t∈(-∞,∞)}为一随机过程，若对任一正整数k以及k个时刻$t_1 < t_2 < \cdots < t_k$,

$$(G_1(t_1), G_2(t_2), \cdots, G_k(t_k))$$

的联合分布为k维正态分布, 则G为**Gauss过程**。

注: 对于Guess过程而言，严格平稳过程与宽平稳过程是等价的.

平稳过程中还有一类常见的过程是周期平稳过程.

定义4.4 设X={X(t), t∈T}为一平稳过程，若存在正常数τ使

$$X(t+\tau)=X(t),$$

则称X为**周期平稳过程,** τ为该过程的周期。

注: 若X是周期平稳过程, 则其协方差函数$R_X(t)$也是周期函数, 且周期相同.

例4.4 (周期振动) 设X(t)=f(t)Z, 其中Z是实随机变量, 满足E(Z)=0, E(Z^2)=σ^2, f(t)为一个非随机的复值函数. 考虑复值过程X={X(t), t∈(-∞,∞)}. 试证明: X为平稳过程的充要条件是

$$f(t)=Ce^{i(\lambda t+\theta)}$$

其中 $i=\sqrt{-1}, C, \lambda, \theta$为常数.

证: (充分性) 因为 $f(t)=Ce^{i(\lambda t+\theta)}$, 所以

$$E[X(t)]=0, \quad R_X(t+h,t)=E[X(t+h)\overline{X(t)}]=\sigma^2C^2e^{i\lambda h}$$

故X是平稳过程.

(必要性) 因为X为平稳过程, 则

$$R_X(h)=E[X(t+h)\overline{X(t)}]=E(X^2)f(t+h)\overline{f(t)}=\sigma^2 f(t+h)\overline{f(t)},$$

故 $f(t+h)\overline{f(t)}$ 与t无关. 取h=0, 我们有

$$R_X(0)=f(t)\overline{f(t)}=|f(t)|=C^2 \quad (\text{某个非负} C),$$

故 $f(t)=Ce^{i\psi(t)}$, 这里 $\psi(t)$ 为一实数. 由此,

$$f(t+h)\overline{f(t)}=C^2 e^{i(\psi(t+h)-\psi(t))}$$

与t无关, 因此

$$\frac{d(\psi(t+h)-\psi(t))}{dt}=0, \quad 即: \quad \frac{d\psi(t)}{dt}=\frac{d\psi(s)}{ds}$$

则 $\psi(t)=\lambda t+\theta$ 形式. 证毕.

课外作业:

Page 106 Ex 1, 3, 5

§ 4.2 遍历性定理

设$X_1,X_2,\ldots,X_n$是来自总体X的样本, 记EX=m.

由大数定律知:

$$\hat{m}_X = \frac{1}{n}\sum_{k=1}^{n} X_k$$

又设$Y_1,Y_2,\ldots,Y_n$是来自Y的样本,记R=Cov(X,Y),

由大数定律知:

$$\hat{R} = \frac{1}{n}\sum_{k=1}^{n}(X_k - \hat{m}_X)(Y_k - \hat{m}_Y)$$

对于随机过程$\{X(t), t\in T\}$, 我们记

$$m(t) = E[X(t)]$$

$$R(t,s) = E[(X(t)-m(t))(X(s)-m(s))]$$

由于对随机过程进行多次观察一般来说是很难做到的，而比较容易的是做一次观察，获得一条样本路径. 我们希望由这一条样本路径来估计m(t)和R(t,s). 对于一般的随机过程这是不太可能的，但是对于平稳过程，只要加上一些很一般的条件就可以了. 这就是著名的遍历性定理.

定义4.5 设X={X(t), t∈T}为一平稳过程(序列)，若

$$\overline{X} = \lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T}X(t)dt \overset{L_2}{=} m \quad 或 \overline{X} = \lim_{N\to\infty}\frac{1}{2N+1}\sum_{k=-N}^{N}X(k) \overset{L_2}{=} m$$

则称X的**均值有遍历性**。如果

$$\hat{R}(\tau) = \lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T}[X(t)-m][X(t+\tau)-m]dt \overset{L_2}{=} R(\tau)$$

或
$$\hat{R}(\tau) = \lim_{N\to\infty}\frac{1}{2N+1}\sum_{k=-N}^{N}[X(k)-m][X(k+\tau)-m] \overset{L_2}{=} R(\tau)$$

则称X的**协方差有遍历性**。若随机过程(序列)的均值和协方差函数都有遍历性，则称此随机过程(序列)有**遍历性**.

注: 我们这里的L_2收敛是指均方收敛. 比如:

$$\lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T}X(t)dt\overset{L_2}{=}m$$

的含义是

$$\lim_{T\to\infty}E\left(\frac{1}{2T}\int_{-T}^{T}X(t)dt-m\right)^2=0.$$

定理4.1 (均值遍历性定理)

(i) 设X={X_n, n=0,±1,±2,…}为一平稳序列，其协方差函数为R(τ), 则X均值有遍历性的充要条件是

$$\lim_{N\to\infty}\frac{1}{N}\sum_{\tau=0}^{N-1}R(\tau)=0.$$

(ii) 设X={X(t), -∞<t<∞}为一平稳过程，其协方差函数为R(τ), 则X均值有遍历性的充要条件是

$$\lim_{T\to\infty}\frac{1}{T}\int_0^{2T}(1-\frac{\tau}{2T})R(\tau)d\tau=0.$$

证: 连续型过程给出证明(离散型序列类似).

$$\overline{X}_T = \frac{1}{2T}\int_{-T}^{T} X(t)dt$$

则

$$E(\overline{X}_T) = \frac{1}{2T}\int_{-T}^{T} E[X(t)]dt = m.$$

下面计算其方差.

$$Var(\overline{X}_T) = E\left(\frac{1}{2T}\int_{-T}^{T} X(t)dt - m\right)^2$$

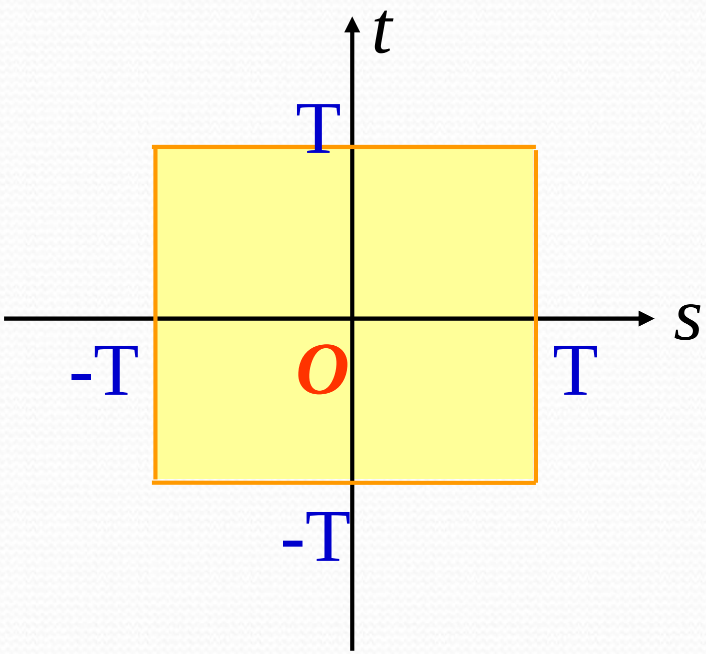
$$= \frac{1}{4T^2} E\left(\int_{-T}^{T}[X(t)-m]dt\int_{-T}^{T}[X(s)-m]ds\right)$$

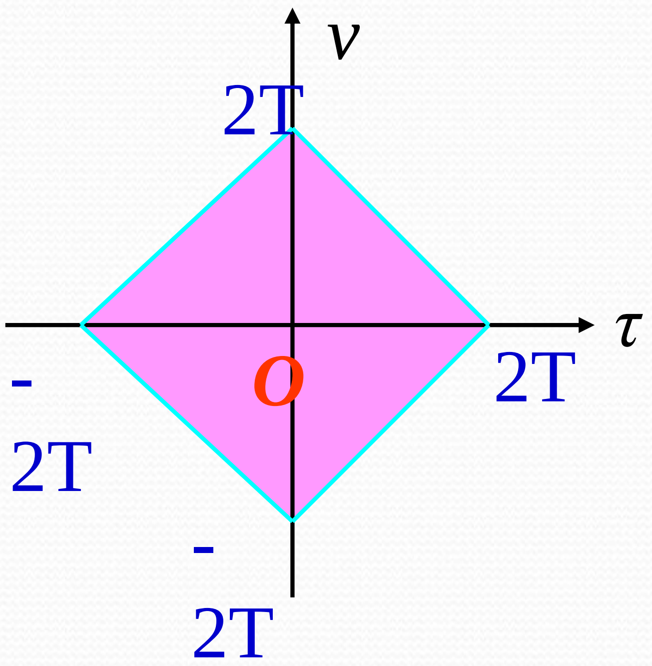
$$= \frac{1}{4T^2}\int_{-T}^{T}\int_{-T}^{T} E[(X(t)-m)(X(s)-m)]dtds$$

$$= \frac{1}{4T^2}\int_{-T}^{T}\int_{-T}^{T} R(t-s)dtds$$

作积分变换$\begin{cases}\tau = t - s \\ v = t + s\end{cases}$，则变换的Jacobi行列式为:

$$J = \begin{vmatrix} \frac{\partial t}{\partial \tau} & \frac{\partial t}{\partial v} \\ \frac{\partial s}{\partial \tau} & \frac{\partial s}{\partial v} \end{vmatrix} = \begin{vmatrix} \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{vmatrix} = \frac{1}{2}$$





积分区域变换为: D={-2T≤τ±v≤2T}. 由R(τ)是偶函数得:

$$Var(\overline{X}_T) = \frac{1}{4T^2} \cdot \frac{1}{2} \iint_D R(\tau) d\tau dv$$

$$= \frac{1}{4T^2} \cdot \int_{-2T}^{2T} R(\tau)(2T - |\tau|) d\tau$$

$$= \frac{1}{T} \cdot \int_0^{2T} R(\tau)(1 - \frac{\tau}{2T}) d\tau$$

若上式极限是0, 则由均方收敛定义得证.

推论4.1 若 $\int_{-\infty}^{\infty}|R(\tau)|d\tau<\infty$, 则均值遍历性成立.

证: 这是因为对任意的0≤τ≤2T, 我们有

$$\left|(1-\frac{\tau}{2T})R(\tau)\right|\le|R(\tau)|.$$

推论4.2 对平稳序列, 若 $R(\tau)\to 0\ (\tau\to\infty)$, 则均值遍历性成立.

证: 由Stolz定理,

$$\lim_{N\to\infty}\frac{1}{N}\sum_{\tau=0}^{N-1}R(\tau)=\lim_{N\to\infty}R(N-1)=0.$$

定理4.2 (协方差函数遍历性定理)

设X={X(t), -∞<t<∞}为一平稳过程，对于任意给定的τ, X的协方差函数为R(τ)有遍历性的充要条件是

$$\lim_{T\to\infty}\frac{1}{T}\int_0^{2T}(1-\frac{u}{2T})[B(u)-R^2(\tau)]du=0,$$

其中

$$B(u)=E[X(t+\tau+u)X(t+u)X(t+\tau)X(t)].$$

注: 定理4.2由于牵涉到过程的四阶矩, 一般难以验证.

定理4.3 (协方差函数遍历性定理)

设X={X_n, n=0,±1,±2,…}为一均值为0的Gauss平稳序列. 如果

$$\lim_{T\to\infty}\frac{1}{N}\sum_{k=0}^{N-1}R^2(k)=0,$$

则Gauss过程X的协方差函数有遍历性.

例4.5 设X(t)=acos(ωt+Y), ω≠0, Y~U(0,2π), 则

X={X(t), -∞<t<∞}的均值有遍历性.

证: 首先 $E[X(t)]=\int_{-\infty}^{+\infty} a\cos(\omega t+y) f_Y(y)dy$

$$=\frac{1}{2\pi}\int_0^{2\pi} a\cos(\omega t+y)dy=0$$

其协方差函数为:

$$R(\tau)=E[a\cos(\omega(t+\tau)+Y)a\cos(\omega t+Y)]$$

$$=\frac{a^2}{2\pi}\int_0^{2\pi}\cos(\omega(t+\tau)+y)\cos(\omega t+y)dy$$

$$=\frac{a^2}{4\pi}\int_0^{2\pi}[\cos(\omega(2t+\tau)+2y)+\cos(\omega\tau)]dy=\frac{a^2}{2}\cos(\tau\omega)$$

故X是平稳的. 由于

$$\frac{1}{T}\int_0^{2T}(1-\frac{\tau}{2T})R(\tau)d\tau=\frac{a^2}{2T}\int_0^{2T}(1-\frac{\tau}{2T})\cos(\omega\tau)d\tau$$

$$=\frac{a^2}{2T}\cdot\frac{\sin(2T\omega)}{\omega}-\frac{a^2}{4T^2}\int_0^{2T}\tau\cos(\omega\tau)d\tau$$

由分布积分知

$$\left|\int_0^{2T}\tau\cos(\omega\tau)d\tau\right|=\left|\frac{1}{\omega}[2T\sin(2T\omega)-\frac{1}{\omega}(1-\cos(2T\omega))]\right|\le\frac{2T}{\omega}$$

因此,

$$\frac{1}{T}\int_0^{2T}(1-\frac{\tau}{2T})R(\tau)d\tau\to 0\qquad (T\to\infty).$$

故由定理4.1知, X的均值有遍历性.

课外作业:

Page 108 Ex 16

§ 4.3 平稳序列的预报

设X表示某次“试验”的一个结果. 假定X是未知的, 在将来才会被观察到, 现在要求预报X的值. 设 $\hat{X}$ 是X的一个预报, 用 $X-\hat{X}$ 表示预报误差. 我们采用均方误差 $E(X-\hat{X})^2$ 来衡量预报误差, 即在均方意义下, 使预报误差 $E(X-\hat{X})^2$ 达到最小.

如果有估计$\hat{X}$, 使 $E(X-\hat{X})^2$ 达到最小, 则称 $\hat{X}$为X的**最小均方误差预报**.

我们的任务是找出X的预报 $\hat{X}$, 使 $E(X-\hat{X})^2$达到最小. 为此, 我们必须指明预报 $\hat{X}$ 所能取值的范围. 一般 $\hat{X}$ 的形式由试验的知识或可能的X的分布函数的知识来确定. 以下用H表示允许的预报类.

例4.6 设某次试验的一个结果为X，其均值为μ，方差为σ^2，此外对X，没有进一步的信息，则在均方误差最小的意义下对X的最好预报是$\hat{X}=\mu$.

解: 由于X没有进一步的信息可利用，故把预报空间H取为所有的实数集合. 我们可以选取任一实数a作为X的预报. 根据方差最小性，对任意$a\neq\mu$，我们有

$$E(X-\mu)^2 > E(X-a)^2$$

因此，当$\hat{X}=\mu$时$E(X-\hat{X})^2$最小. 证毕.

例4.7 设随机变量X和Y的联合分布已知，且有有限方差．假定Y是可观察的，求根据Y的观察值的X的预报．

解: 由于除了联合分布之外，X没有进一步的信息．我们允许有有限方差的任一Y的函数$\hat{X}=f(Y)$作为X的预报, 即预报空间

$$H=\{\hat{X}=f(Y): f\text{为任一函数且}Var(\hat{X})<\infty\}.$$

在观察到Y=y后，我们可以得到最小均方误差预报是在Y=y下X的条件均值 $\mu_{X|Y}=E(X\mid Y)$ ．这是因为

$$E(X-\hat{X})^2 = E(X-\mu_{X|Y})^2 + 2E[(X-\mu_{X|Y})(\mu_{X|Y}-\hat{X})]$$

$$+E(\mu_{X|Y}-\hat{X})^2$$

在上式右边第二项中用条件期望平滑公式得:

$$E[(X-\mu_{X|Y})(\mu_{X|Y}-\hat{X})] = E_Y\left\{E[(X-\mu_{X|Y})(\mu_{X|Y}-\hat{X})\big|Y]\right\}$$

$$= E_Y\left\{(\mu_{X|Y}-\hat{X})E[(X-\mu_{X|Y})\big|Y]\right\}$$

$$= E_Y\left\{(\mu_{X|Y}-\hat{X})[E(X\big|Y)-\mu_{X|Y}]\right\}$$

$$= E_Y\left\{(\mu_{X|Y}-\hat{X})[\mu_{X|Y}-\mu_{X|Y}]\right\} = 0$$

所以

$$E(X-\hat{X})^2 = E(X-\mu_{X|Y})^2 + E(\mu_{X|Y}-\hat{X})^2$$

$$\geq E(X-\mu_{X|Y})^2$$

当且仅当 $\hat{X} = \mu_{X|Y}$ 时取等号, 故

$$\hat{X} = E(X \mid Y)$$

在例4.7中得到 $\hat{X}=E(X\mid Y)$，但是 $E(X\mid Y)$ 的计算需要X和Y联合分布的全部信息. 并且即使联合分布已知，这个条件期望的计算也往往由于太繁杂而得不到实际结果. 另外，在平稳过程的研究中我们仅假定过程的二阶矩存在，没有进一步关于联合分布的知识，因此要求有一种预报理论，在仅知道二阶矩的情况下也能导出简单的预报公式，方法之一是线性预报.

定理4.4 设X, Y分别有均值μ_X, μ_Y 和方差σ_X^2, σ_Y^2, 它们的协方差为$\sigma_{X,Y} = E(X-\mu_X)(Y-\mu_Y)$. 我们限制X的预报只能是Y的线性函数, 即$\hat{X} = a + bY$

或等价地$\hat{X} = a + b(Y-\mu_Y)$这种形式, 这里a和b为任意实数. 则X根据Y的**最佳线性预报**为:

$$\hat{X}^* = \mu_X + \frac{\sigma_{X,Y}}{\sigma_Y^2}(Y-\mu_Y)$$

证: 令 $\hat{X}$ 为X的任一线性预报, 即

$$\hat{X} \in H = \{a + b(Y - \mu_Y), \quad a, b\text{为实数}\}$$

注意到$\hat{X}*-\hat{X}$仍是X的线性预报, 故$\hat{X}*-\hat{X} \in H$,且

$$\hat{X}*-\hat{X} = a' + b'(Y - \mu_Y),$$

其中

$$a' = \mu_X - a, \ \ b' = \frac{\sigma_{X,Y}}{\sigma_Y^2} - b$$

同例4.7一样计算得:

$$E(X - \hat{X})^2 = E(X - \hat{X}*)^2 + 2E(X - \hat{X})(\hat{X}*-\hat{X})$$

$$+ E(\hat{X}*-\hat{X})^2.$$

因为

$$E(X-\hat{X})(\hat{X}^*-\hat{X})$$

$$=E\left\{\left[(X-\mu_X)-\frac{\sigma_{X,Y}}{\sigma_Y^2}(Y-\mu_Y)\right][a'+b'(Y-\mu_Y)]\right\}$$

$$=a'E\left[(X-\mu_X)-\frac{\sigma_{X,Y}}{\sigma_Y^2}(Y-\mu_Y)\right]$$

$$+b'E\left[(X-\mu_X)(Y-\mu_Y)-\frac{\sigma_{X,Y}}{\sigma_Y^2}(Y-\mu_Y)^2\right]$$

$$=0+b'\left(\sigma_{X,Y}-\frac{\sigma_{X,Y}}{\sigma_Y^2}\sigma_Y^2\right)=0$$

因此

$$E(X-\hat{X})^2 = E(X-\hat{X}^*)^2 + E(\hat{X}^* - \hat{X})^2.$$

当 $\hat{X}=\hat{X}^*$ 时, 上式在一切线性预报中取最小值, 故

$$\hat{X}^* = \mu_X + \frac{\sigma_{X,Y}}{\sigma_Y^2}(Y-\mu_Y)$$

是所述的线性最小均方误差预报.

注: 如果(X,Y)为正态分布, 则

$$E(X \mid Y=y) = \mu_X + \frac{\sigma_{X,Y}}{\sigma_Y^2}(y-\mu_Y).$$

此时, 线性最小均方误差预报等于最小均方误差预报.

定理4.5 设随机变量X有有限二阶矩. H为允许的预报 $\hat{X}$ 所组成的空间, 即 $H=\{\hat{X}:E(\hat{X}^2)<\infty\}$. 假设H关于线性运算封闭. 则预报 $\hat{X}*$有最小均方误差当且仅当对每个预报 $u\in H$ 有

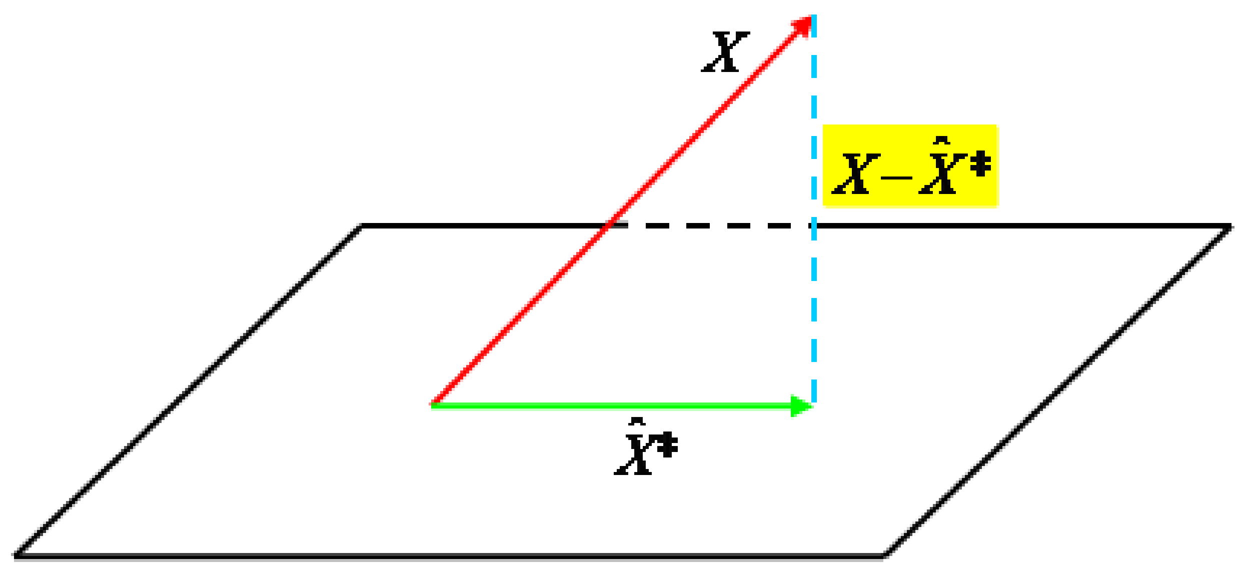
$$E[(X-\hat{X}*)u]=0.$$

此时我们称 $\hat{X}*$ 为**最小均方误差预报**, 简称最佳预报.

证明见附录A.

定理4.5直观的几何解释是:

如果 $\hat{X}*$ 是X在预报空间H上的最小均方误差预报, 则 $X-\hat{X}*$ 与空间H垂直. 又由于 $\hat{X}*$ 在H中, 故 $\hat{X}*$ 是X在空间H上的投影. 因此定理4.5又称为**投影定理**.



在最小均方误差预报中，如果把预报类仅限于H中随机变量的线性组合，则所得到的预报称为线性最小均方误差预报，简称为**线性最佳预报**．一般记为$\hat{E}(X\mid H)$. 如果H仅由有限个随机变量$X_1,X_2,\dots,X_n$生成, 即

$$H=\{a_1X_1+\cdots+a_nX_n:a_1,\cdots,a_n\text{为任意实数}\}.$$

记随机变量$Y=(X_1,\cdots,X_n)$, 则预报可写成$\hat{E}(X\mid Y)$, 则最佳预报的表达式为:

$$\hat{E}(X\mid Y)=EX+Cov(X,Y)(Var(Y))^{-1}(Y^T-E(Y^T)).$$

证明见附录B.

- **平稳序列的预报**

设$\{X_n, n=\ldots, -1, 0, 1, \ldots\}$为一平稳序列. 在时刻$n$我们要预报$X_{n+m}$的值, 这称为**m步预报**. 不妨设$X_1$, X_2, …, X_n为到时刻n为止所能获得的数据, 我们考虑X_{n+m}的线性最佳预报, 记为 $\hat{X}_{n+m|n,1}$, 则

$$\hat{X}_{n+m|n,1} = EX_{n+m} + Cov(X_{n+m}, Y_{n,1})(Var(Y_{n,1}))^{-1}(Y_{n,1}^T - E(Y_{n,1}^T)),$$

其中$Y_{n,1} = (X_1, X_2, \cdots, X_n)$.

注: $\hat{X}_{n+m|n,1}$随n的变化而变化, 是非平稳预报, 只有在数据量不大时才有使用价值, 这里暂不讨论.

为了克服非平稳预报的缺点，我们考虑$\hat{X}_{n+m|n,-\infty}$，简记为$\hat{X}_{n+m|n}$，这种预报的形式固定.

在实际中，我们不可能得到从-∞到n的全部数据，此时只要作为预报的时刻n充分大，则在预报公式中，0时刻以前的数据实际上对预报值已不起什么作用了，因而可以较随意地取值，比如取为0.

1. 自回归模型的线性最佳预报

设$\{\varepsilon_n, n = \ldots, -1, 0, 1, \ldots\}$是均值为0, 方差为$\sigma^2$的噪声序列. H_n表示由$\{\varepsilon_k, k \le n\}$所生成的线性空间.

设随机序列$\{X_n, n = \ldots, -1, 0, 1, \ldots\}$满足

$$X_n = aX_{n-1} + \varepsilon_n, \qquad n = \cdots, -1, 0, 1, \cdots, |a| < 1,$$

称为**一阶自回归模型**，记为$AR(1)$.

容易求得：

$$X_n = \sum_{k=0}^{+\infty} a^k \varepsilon_{n-k},$$

$$EX_n = 0, \qquad \sigma_X^2 = E(X_n^2) = \frac{\sigma^2}{1-a^2}$$

$$Cov(X_n, X_{n+h}) = a^h \sigma_X^2 .$$

从而$\{X_n, n = \ldots, -1, 0, 1, \ldots\}$为一平稳序列.

下面我们讨论$AR(1)$的m步线性最佳预报.

注意到$\{X_k, k\le n\}$生产的线性空间与H_n相同, 根据定理4.5, X_{n+m}可以分解为两部分, 第一部分与H_n正交, 第二部分是H_n中的一个元素. 这种分解叫Wold分解. 由定理4.5知, X_{n+m}的线性最佳预报应该是X_{n+m}在H_n上的投影, 即:

$$\hat{X}_{n+m|n} = a^m \sum_{k=0}^{+\infty} a^k \varepsilon_{n-k} = a^m X_n$$

由定理4.5知，X_{n+m}的线性最佳预报应该是X_{n+m}在H_n上的投影, 即:

$$\hat{X}_{n+m|n} = a^m \sum_{k=0}^{+\infty} a^k \varepsilon_{n-k} = a^m X_n$$

预报误差为:

$$X_{n+m} - \hat{X}_{n+m|n} = \varepsilon_{n+m} + a\varepsilon_{n+m-1} + \cdots + a^{m-1}\varepsilon_{n+1}$$

$$E(X_{n+m} - \hat{X}_{n+m|n})^2 = \sum_{k=0}^{m-1} a^{2k}\sigma^2 = \frac{1-a^{2m}}{1-a^2}\sigma^2$$

设随机序列$\{X_n, n = \ldots, -1, 0, 1, \ldots\}$满足

$$X_n = \sum_{k=1}^{p} a_k X_{n-k} + \varepsilon_n , \qquad n = \cdots, -1, 0, 1, \cdots,$$

称为**p阶自回归模型**，记为AR(p). 可以改写为:

$$(1 - a_1 z - a_2 z^2 - \cdots - a_p z^p) X_n = \varepsilon_n ,$$

其中z表示后移算子, 即: $zX_n = X_{n-1}$. 因此, 形式上可以写为:

$$X_n = (1 - a_1 z - a_2 z^2 - \cdots - a_p z^p)^{-1} \varepsilon_n ,$$

如果多项式$1-a_1z-a_2z^2-\cdots-a_pz^p$的根都在单位圆外, 则

$$(1-a_1z-a_2z^2-\cdots-a_pz^p)^{-1}$$

可以展成z的收敛幂级数, 记为

$$(1-\sum_{k=1}^{p}a_kz^k)^{-1}=\psi(z)=\sum_{k=0}^{\infty}b_kz^k$$

$$(1-\sum_{k=1}^{p} a_k z^k)^{-1} = \sum_{k=0}^{\infty} b_k z^k$$

其中

$$b_1 = a_1$$

$$b_2 = a_1^2 + a_2$$

$$b_3 = a_1^3 + 2a_1 a_2 + a_3$$

$$b_4 = a_1^4 + 3a_1^2 a_2 + 2a_1 a_3 + a_2^2 + a_4$$

$$\cdots\cdots$$

通过严格地数学证明可知:

$$X_n = \sum_{k=0}^{\infty} b_k z^k \varepsilon_n = \sum_{k=0}^{\infty} b_k \varepsilon_{n-k}$$

因此,

$$EX_n = 0, \qquad E(X_n^2) = \sum_{k=0}^{\infty} b_k^2 \sigma^2,$$

$$Cov(X_n, X_{n+h}) = \sigma^2 \sum_{k=0}^{\infty} b_k b_{k+h},$$

从而$AR(p)$也是平稳过程.

同p=1的场合，由$\{X_k\ ,\ k\le n\}$生产的线性空间与H_n相同. 而X_{n+1}=a$_1X_n$+…+$a_pX_{n+1\text{-}p}$+ε_{n+1}, 其中ε_{n+1}与H_n垂直, a$_1X_n$+…+$a_pX_{n+1\text{-}p}$是H_n中的 元素，因此由定理4.5知 一步预报为:

$$\hat{X}_{n+1|n} = a_1 X_n + a_2 X_{n-1} + \cdots + a_p X_{n+1-p},$$

多步预报为:

$$\hat{X}_{n+m|n} = a_1 \hat{X}_{n+m-1|n} + a_2 \hat{X}_{n+m-2|n} + \cdots + a_p \hat{X}_{n+m-p|n},$$

另一方面，由 $X_{n+m}=\sum_{k=0}^{\infty}b_k\varepsilon_{n+m-k}$ 得:

$$\hat{X}_{n+m|n}=\sum_{k=0}^{\infty}b_{m+k}\varepsilon_{n-k}$$

并且

$$X_{n+m}-\hat{X}_{n+m|n}=\sum_{k=0}^{m-1}b_k\varepsilon_{n+m-k}$$

$$E(X_{n+m}-\hat{X}_{n+m|n})^2=\sigma^2\sum_{k=0}^{m-1}b_k^2 .$$

在n时刻, X_{n+m}的1-α置信区间为:

$$\hat{X}_{n+m|n}\pm u_{1-\frac{\alpha}{2}}\cdot\sigma\cdot\sqrt{\sum_{k=0}^{m-1}b_k^2}$$

例4.8 设$AR(2)$模型为

$$X_n = 1.2X_{n-1} - 0.55X_{n-2} + \varepsilon_n,$$

和

$$X_k = 0.761,\ X_{k-1} = 6.02,\ \sigma_\varepsilon^2 = 0.08^2$$

求模型的前三步预报值及X_{k+3}的0.95置信区间.

解:

$$\hat{X}_{k+1|k} = 1.2X_k - 0.55X_{k-1} = 5.821$$

$$\hat{X}_{k+2|k} = 1.2\hat{X}_{k+1|k} - 0.55X_k = 2.7995$$

故

$$\hat{X}_{k+3|k} = 1.2\hat{X}_{k+2|k} - 0.55\hat{X}_{k+1|k} = 0.1579$$

对于给定的α=0.05, 查表得:

$$u_{1-\frac{\alpha}{2}} = u_{0.975} = 1.96$$

故

$$\hat{X}_{k+3|k} \pm u_{1-\frac{\alpha}{2}} \sigma \sqrt{\sum_{k=0}^{2} b_k^2}$$

$$= 0.158 \pm 1.96 \times 0.08 \times \sqrt{1 + 1.2^2 + 0.89^2}$$

因此, X_{k+3}的0.95置信区间为:

$$(0.124, 0.440)$$

课堂练习： 对于AR(1)

$$X_n = aX_{n-1} + \varepsilon_n\,, \qquad n = \cdots, -1, 0, 1, \cdots, |a| < 1,$$

证明: $\hat{X}_{n+m|n} = a^m X_n$

证：

$$\hat{X}_{n+1|n} = aX_n$$

$$\hat{X}_{n+k|n} = a\hat{X}_{n+k-1|n}$$

用数学归纳法易得结论.

2. 滑动平均模型的预报

设$\{\varepsilon_n, n=\ldots, -1, 0, 1, \ldots\}$是均值为0, 方差为$\sigma^2$的白噪声. H_n表示由$\{\varepsilon_k, k\le n\}$所生成的线性空间.

设随机序列$\{X_n, n = \ldots, -1, 0, 1, \ldots\}$满足

$$X_n = \varepsilon_n - \beta_1\varepsilon_{n-1} - \cdots - \beta_q\varepsilon_{n-q}, \quad n = \cdots, -1, 0, 1, \cdots$$

称$\{X_n\}$为**q阶滑动平均模型**，记为$MA(q)$.

利用后移算子, 滑动平均模型可改写成:

$$X_n = (1-\beta_1 z-\beta_2 z^2-\cdots-\beta_q z^q)\varepsilon_n\ ,$$

其中z表示后移算子, 即: $z\varepsilon_n=\varepsilon_{n-1}$. 通过严格地数学

证明可知:

$$EX_n=0, \qquad E(X_n^2)=(1+\sum_{k=1}^{q}\beta_k^2)\sigma^2,$$

$$Cov(X_n, X_{n+h})=\begin{cases}\sigma^2\sum_{k=h}^{q}\beta_k\beta_{k-h} & 0\le h\le q\\ 0 & h\ge q+1\end{cases}$$

从而$MA(q)$也是平稳过程.

由定理4.5知，X_{n+m}的线性最佳预报应该是X_{n+m}在H_n上的投影，即:

$$\hat{X}_{n+m|n}=\begin{cases}0, & m>q\\ -\sum_{k=m}^{q}\beta_k\varepsilon_{n+m-k}, & 1\le m\le q\end{cases}$$

由于$\{\varepsilon_n\}$不是数据，因此不便于直接使用上式预报. 为此，我们转化成利用$\{X_n\}$历史数据进行预报的形式.

由后移算子表示可知:

$$\varepsilon_n = (1-\beta_1 z - \beta_2 z^2 - \cdots - \beta_q z^q)^{-1} X_n,$$

$$= \sum_{k=0}^{\infty} a_k X_{n-k},$$

其中 $(1-\sum_{k=1}^{q}\beta_k z^k)^{-1} = \sum_{k=0}^{\infty} a_k z^k$, 并且 $a_0 = 1$, 因此

$$\varepsilon_n = X_n + \sum_{k=1}^{\infty} a_k X_{n-k},$$

故

$$X_n = \varepsilon_n - \sum_{k=1}^{\infty} a_k X_{n-k}.$$

因此，滑动平均模型可以看成无穷阶自回归模型．下面我们将利用自回归模型预报的方法给出滑动平均模型的一步预报. 首先注意到,

$$X_{n+1} = \varepsilon_{n+1} - \sum_{k=1}^{\infty} a_k X_{n+1-k} .$$

由定理4.5知,

$$\hat{X}_{n+1|n} = -\sum_{k=1}^{\infty} a_k X_{n+1-k} ,$$

$$E(X_{n+1} - \hat{X}_{n+1|n})^2 = E\varepsilon_{n+1}^2 = \sigma^2 .$$

关于滑动平均模型的多步预报. 注意到,

$$X_{n+m} = \varepsilon_{n+m} - \sum_{k=1}^{m-1} a_k X_{n+m-k} - \sum_{k=m}^{\infty} a_k X_{n+m-k}.$$

由定理4.5知,

$$\hat{X}_{n+m|n} = -\sum_{k=1}^{m-1} a_k \hat{X}_{n+m-k|n} - \sum_{k=m}^{\infty} a_k X_{n+m-k}.$$

上式可以作为递推公式, 递推得出多步预报, 略.

课外作业:

Page 109	Ex 31(1), (2), (3)
Page 110	Ex 34, 41, 42