最佳估计原理

上海交通大学

袁天鑫 编

国防工业出版社

最佳估计原理

上海交通大学

袁天鑫 编

国防工业出版社

内 容 简 介

本书主要为高等学校工科电子类《自动控制》专业的高年级大学生选修和低年级研究生必修《最佳估计原理》课程而编写的。主要包括随机线性系统的数学描述，基本估计方法，离散线性系统的最佳估计（滤波、预测和平滑），连续线性系统的最佳估计以及最佳估计的稳定性、误差分析、发散现象和克服发散的方法等重要内容，最后还扼要地介绍了线性滤波方法对非线性系统滤波的一些推广应用。为了方便教学和一般工程技术人员阅读，在书后还编写了《矩阵和矢量》、《概率论和随机过程基础》两个附录。

本书还可作为有关工程技术人员和高等学校师生的参考书。

前　言

目前，我国科学技术的发展日新月异，电子数字计算机的应用也越来越广泛。在生产和科学技术领域中，遇到了许多需要借助于估计理论来解决的实际问题。因此，估计理论及其方法受到了各方面的重视。本书就是为了适应这方面的需要，在有关高等学校工科电子类《自动控制》专业开设《最佳估计原理》课程而编写的统编教材。

本书主要作为《自动控制》专业高年级大学生的选修教材，同时也兼顾到该专业低年级研究生使用。其主要任务是：通过本课程的学习，使读者初步掌握最佳估计，特别是线性状态估计的基本内容。材料的选择基本上包括了近二十年来国内外在离散和连续线性系统的状态估计方面的主要成果，此外，为了应用，还包括线性状态估计在复杂情况下的推广，最佳估计的稳定性和误差分析，非线性最佳估计的线性化等有关内容。

全书共分六章：第一章讨论随机线性系统的数学描述；第二章讨论几种基本的估计方法；第三章给出离散线性系统状态估计(滤波、预测和平滑)的基本方程；第四章给出连续线性系统状态估计的基本方程；第五章讨论最佳滤波的稳定性，误差分析，发散现象和克服发散的主要方法；第六章简单介绍非线性系统的最佳滤波。为了方便教学和其他读者阅读，在书后还编写了两个附录：A．矩阵和矢量；B．概率论和随机过程基础。

在使用本教材对高年级大学生讲授选修课时，建议用45学时讲授书中不带＊号部分的内容。其中，第一章10学时，第二章6学时，第三章12学时，第四章5学时，第五章7学时，第六章5学时。如果先学课程《线性代数》、《概率论和随机过程》不能满足该课程的要求，则建议用15到20学时先补充讲授附录的内容。即总教学时数建议为60到65学时。研究生使用该教材时，建议用60学时讲完全书的主要内容，即用15学时讲授书中带＊号的部分。其中，第三章增加5学时，第四章增加5学时，第五章增加5学时。

本书由北京工业学院张志方付教授担任主审。中国科技大学的涂其栵付教授，清华大学褚家晋同志和王其藩同志，北京工业学院徐和生同志和张惠同志，上海交通大学何焕熹同志以及中国科学院数学所张永光同志参加了本书初稿的审查，他们对本书提出了许多宝贵的意见。特别是张志方和涂其栵两位先生，他们逐字逐句地阅读了初稿的全部内容，并且提出了许多极为重要的、很有启发性的意见。在此表示衷心的感谢。编者尤要感谢上海交通大学张钟俊教授，因为在本书的编写过程中，自始至终得到了张先生的鼓励、帮助和指导。

由于编者水平不高，时间又比较紧，因此书中一定存在不少缺点和错误，希望同志们随时提出批评和指教。

编　者

1979.12.

目　录

前　言
第一章　随机线性系统的数学描述……………………………………（1）
§1-1　几个特殊的随机过程……………………………………（1）
1-1-1　高斯（正态）随机过程……………………………………（1）
1-1-2　马尔柯夫过程……………………………………（2）
1-1-3，高斯-马尔柯夫过程……………………………………（4）
1-1-4　二阶过程……………………………………（6）
1-1-5　独立增量过程和维纳过程……………………………………（10）
1-1-6　白噪声过程……………………………………（12）
§1-2　随机连续线性系统的数学描述……………………………………（15）
1-2-1　随机连续线性系统的一般描述……………………………………（15）
1-2-2　模型的基本假设……………………………………（16）
1-2-3　假设条件之下模型的特点……………………………………（18）
1-2-4　扩充状态模型……………………………………（23）
§1-3　随机离散线性系统的数学描述……………………………………（29）
1-3-1　随机离散线性系统的一般描述……………………………………（29）
1-3-2　模型的基本假设……………………………………（31）
1-3-3　假设条件之下模型的特点……………………………………（32）
1-3-4　扩充状态模型……………………………………（36）
第二章　基本估计方法……………………………………（41）
§2-1　估计问题的提出及其一般叙述……………………………………（41）
§2-2　最佳估计和最佳估计准则……………………………………（42）
§2-3　基本估计方法……………………………………（44）
2-3-1　最小二乘估计……………………………………（44）
2-3-2　线性最小方差估计……………………………………（47）
2-3-3　极大似然估计……………………………………（51）
2-3-4　贝叶斯估计……………………………………（52）
1．极大验后估计……………………………………（52）
2．最小方差估计……………………………………（54）
2-3-5　举例……………………………………（56）
§2-4　基本估计方法的比较及其相互关系……………………………………（64）
2-4-1　基本估计方法的比较……………………………………（64）
2-4-2　基本估计方法之间的关系……………………………………（65）

第三章　离散线性系统的最佳估计……………………………………………………………………(66)
§3-1　问题的叙述和正交投影………………………………………………………………………(66)
3-1-1　问题的叙述…………………………………………………………………………………(66)
3-1-2　正交投影……………………………………………………………………………………(67)
§3-2　离散线性系统的最佳滤波……………………………………………………………………(70)
3-2-1　概述…………………………………………………………………………………………(70)
3-2-2　卡尔曼滤波方程……………………………………………………………………………(72)
3-2-3　卡尔曼滤波的具体计算……………………………………………………………………(80)
3-2-4　卡尔曼滤波的特点与性质…………………………………………………………………(82)
3-2-5　卡尔曼滤波的推广…………………………………………………………………………(84)
1．白噪声作用下一般线性系统的卡尔曼滤波方程…………………………………………(84)
2．有色噪声情况下线性系统的卡尔曼滤波方程……………………………………………(88)
3-2-6　举例…………………………………………………………………………………………(95)
*§3-3　离散线性系统的最佳预测…………………………………………………………………(107)
3-3-1　概述…………………………………………………………………………………………(107)
3-3-2　最佳固定区间预测…………………………………………………………………………(108)
3-3-3　最佳固定点预测……………………………………………………………………………(108)
3-3-4　最佳固定超前预测…………………………………………………………………………(109)
3-3-5　最佳预测的误差特性………………………………………………………………………(109)
3-3-6　举例…………………………………………………………………………………………(111)
*§3-4　离散线性系统的最佳平滑…………………………………………………………………(113)
3-4-1　概述…………………………………………………………………………………………(113)
3-4-2　最佳固定区间平滑…………………………………………………………………………(114)
3-4-3　最佳固定点平滑……………………………………………………………………………(120)
3-4-4　最佳固定滞后平滑…………………………………………………………………………(125)
3-4-5　举例…………………………………………………………………………………………(130)
第四章　连续线性系统的最佳估计……………………………………………………………………(134)
§4-1　问题的叙述……………………………………………………………………………………(134)
§4-2　连续线性系统的最佳滤波……………………………………………………………………(134)
4-2-1　连续型卡尔曼滤波问题的提法……………………………………………………………(134)
4-2-2　等效的离散线性系统………………………………………………………………………(135)
4-2-3　连续型卡尔曼滤波方程……………………………………………………………………(137)
4-2-4　连续型卡尔曼滤波的推广…………………………………………………………………(141)
1．白噪声作用下一般连续线性系统的卡尔曼滤波方程……………………………………(141)
2．有色噪声情况下连续线性系统的卡尔曼滤波方程………………………………………(142)
4-2-5　举例…………………………………………………………………………………………(145)
*§4-3　连续线性系统的最佳预测…………………………………………………………………(154)
*§4-4　连续线性系统的最佳平滑…………………………………………………………………(157)

4-4-1 连续型最佳固定区间平滑……(157)
4-4-2 连续型最佳固定点平滑……(161)
4-4-3 连续型最佳固定滞后平滑……(162)
4-4-4 举例……(164)
第五章 最佳估计的稳定性和误差特性……(171)
§5-1 最佳估计的稳定性……(171)
5-1-1 最佳滤波的稳定性概念……(171)
5-1-2 最佳滤波的稳定性条件……(172)
§5-2 最佳估计的误差分析……(174)
5-2-1 最佳滤波误差方差阵之微分方程和差分方程的解……(174)
1. 误差方差阵之微分方程的解……(174)
2. 误差方差阵之差分方程的解……(177)
5-2-2 误差方差阵的上、下界……(181)
5-2-3 误差方差阵的渐近特性……(184)
5-2-4 模型误差分析……(197)
1. 存在模型误差时，最佳滤波的实际误差方差阵……(198)
2. 特殊情况的讨论……(202)
*§5-3 最佳估计的发散现象和克服发散的方法……(203)
5-3-1 最佳估计的发散现象……(203)
5-3-2 引起发散的原因……(205)
5-3-3 克服发散的主要方法……(206)
1. 限定下界法……(206)
2. 扩充状态法……(208)
3. 渐消记忆（衰减记忆）滤波……(208)
4. 限定记忆滤波……(212)
5. 自适应滤波……(214)
6. 最佳滤波的平方根算法……(214)
第六章 非线性最佳估计引论……(217)
§6-1 非线性最佳估计问题的叙述……(217)
6-1-1 随机非线性系统的数学描述……(217)
6-1-2 非线性最佳估计问题的叙述……(218)
§6-2 非线性滤波的线性化……(219)
6-2-1 线性化卡尔曼滤波……(219)
1. 连续型线性化卡尔曼滤波……(219)
2. 离散型线性化卡尔曼滤波……(221)
6-2-2 推广卡尔曼滤波……(225)
6-2-3 举例……(228)
附 录
附录 A 矩阵和矢量……(235)
附录 B 概率论和随机过程基础……(260)

第一章　随机线性系统的数学描述

本章主要讨论在随机信号作用下，物理系统的数学描述。由于在以后主要讨论随机线性系统的状态估计问题，因此，本章也就主要研究随机信号作用下线性系统的数学描述。

为了比较深刻地理解随机过程通过线性系统以后的一些特性，我们将首先在§1-1节讨论几个特殊的随机过程。然后在§1-2节讨论在随机信号作用下，连续线性系统的数学描述，而在§1-3节讨论在随机信号作用下，离散线性系统的数学描述。有关随机非线性系统的数学描述，则不作详细讨论，对此问题感兴趣的读者可参阅资料[2]的4.5节。

§1-1　几个特殊的随机过程

1-1-1　高斯（正态）随机过程

讨论矢量随机过程的情况。**如果 n 维矢量随机过程$\{X(t);\ t\in T\}$在 $t\in T$ 中的任意 m 个时间点 $t_1,\ t_2,\cdots\cdots,\ t_m$ 上（m 是任意整数）的 m 个 n 维随机矢量 $X(t_1),\ X(t_2),\cdots X(t_m)$的联合分布是高斯的，即如果其联合特征函数为**

$$\phi_{x^1x^2\cdots x^m}(s^1,t_1;\ s^2,t_2;\ \cdots;s^m,t_m)=e^{j\sum\limits_{K=1}^{m}(\mu^k)^T s^k-\frac{1}{2}\sum\limits_{K=1}^{m}\sum\limits_{j=1}^{m}(s^k)^T P_{xx}(k,j)s^j} \tag{1-1-1}$$

或其联合概率密度函数为

$$p(x^1,t_1;x^2,t_2;\cdots;\ x^m,t_m)=\frac{1}{(2\pi)^{mn/2}|P^*|^{1/2}}e^{-\frac{1}{2}(x^*-\mu^*)^T(P^*)^{-1}(x^*-\mu^*)} \tag{1-1-2}$$

则称此矢量随机过程为高斯分布矢量随机过程，并简称为高斯矢量随机过程。其中

$$\begin{cases}\mu^k=E\{X(t_k)\} \qquad (k=1,\ 2,\cdots,m);\\ P_{xx}(k,j)=\mathrm{cov}\{X(t_k),\ X(t_j)\}(k,\ j=1,2\cdots,m);\\ s^k\text{——}n\text{ 维非随机矢量}(k=1,2,\cdots n)\ ;\\ x^*\text{——由 }m\text{ 个 }n\text{ 维随机矢量 }X(t_k)(k=1,2,\cdots,m)\\ \qquad\text{组成的 }mn\text{ 维随机矢量 }X^*\text{ 的取值};\\ \mu^*=E\{X^*\};\\ P^*=\mathrm{var}\{X^*\}\end{cases}$$

显然，一个高斯矢量随机过程的统计特性完全由其平均值函数 $\mu_x(t)$和在 $t,\tau\in T$ 中的所有自协方差矩阵函数 $P_{xx}(t,\tau)$确定。

用类似的方法也可定义离散型高斯随机序列，读者可自行作出。

1-1-2 马尔柯夫过程

同样讨论矢量随机过程的情况。**n 维矢量随机过程$\{X(t);\ t\in T\}$在 $t\in T$ 中的任意 m 个时间点 $t_1<t_2<\cdots<t_m$(m 是任意整数) 上，如果 $X(t_m)$对于 $X(t_{m-1})$，$X(t_{m-2}),\cdots,X(t_1)$的条件概率具有如下特性：**

$$P\{X(t_m)\leqslant x^m \mid X(t_{m-1})=x^{m-1},X(t_{m-2})=x^{m-2},\cdots,X(t_1)=x^1\}$$
$$=P\{X(t_m)\leqslant x^m \mid X(t_{m-1})=x^{m-1}\} \qquad (1\text{-}1\text{-}3)$$

并对所有 m个 n 维随机矢量 $X(t_1)$，$X(t_2),\cdots,X(t_m)$ 都成立，则称此矢量随机过程为马尔柯夫过程。

由上述马尔柯夫过程的概念可知，对于所有任意 m 个 n 维随机矢量 $X^1=X(t_1),X^2=X(t_2),\cdots,X^m=X(t_m)$，其条件概率分布函数具有如下特性：

$$F(x^m,t_m \mid x^{m-1},t_{m-1};\ x^{m-2},\ t_{m-2};\cdots;x^1,t_1)=F(x^m,t_m \mid x^{m-1},t_{m-1}) \qquad (1\text{-}1\text{-}4)$$

而其条件概率密度函数具有如下特性：

$$p(x^m,t_m \mid x^{m-1},t_{m-1};\ x^{m-2},t_{m-2};\cdots;\ x^1,t_1)=p(x^m,t_m \mid x^{m-1},t_{m-1}) \qquad (1\text{-}1\text{-}5)$$

如果我们将时间 t_{m-1} 看作现在时刻，而将 $t_{m-2},t_{m-3},\cdots,t_1$ 看作过去时刻，那末，由式(1-1-4)或式(1-1-5)可知，过程在将来时刻 t_m 的统计特性仅仅取决于现在时刻的状态，而与过去时刻的状态无关，这是马尔柯夫过程的一个重要特性。常常称它为“马尔柯夫特性”。

如果矢量随机过程 $\{X(t);\ t\in T\}$ 是一个马尔柯夫过程，则对于 $t\in T$ 中的任意 m 个时间点 $t_1<t_2<\cdots<t_m$(m 是任意整数)上的所有 m 个随机矢量 $X^m=X(t_m)$；$X^{m-1}=X(t_{m-1})$；$\cdots$；$X^1=X(t_1)$ 的联合概率密度函数可以如下求得：

由条件概率密度函数公式可得

$$p(x^m,t_m;\ x^{m-1},t_{m-1};\cdots;x^1,t_1)=p(x^m,t_m \mid x^{m-1},\ t_{m-1};\ x^{m-2},t_{m-2};\cdots;\ x^1,t_1)$$
$$\cdot p(x^{m-1},t_{m-1};\ x^{m-2},t_{m-2};\cdots;\ x^1,t_1)$$

由于过程是马尔柯夫的，因此根据式 (1-1-5) 可将上式写成如下形式

$$p(x^m,t_m;\ x^{m-1},t_{m-1};\cdots;x^1,t_1)=p(x^m,t_m \mid x^{m-1},t_{m-1})\cdot p(x^{m-1},t_{m-1};x^{m-2},$$
$$t_{m-2};\cdots;x^1,t_1)$$

同理，由条件概率密度函数公式和考虑到过程是马尔柯夫的可得

$$p(x^{m-1},t_{m-1};\ x^{m-2},t_{m-2};\cdots;x^1,t_1)=p(x^{m-1},t_{m-1} \mid x^{m-2},t_{m-2})\cdot p(x^{m-2},t_{m-2};$$
$$x^{m-3},t_{m-3};\cdots;x^1,t_1)$$

将上式代入前一式得

$$p(x^m,t_m;x^{m-1},t_{m-1};\cdots;x^1,t_1)=p(x^m,t_m \mid x^{m-1},t_{m-1})\cdot p(x^{m-1},t_{m-1} \mid x^{m-2},t_{m-2})\cdot$$
$$p(x^{m-2},t_{m-2};x^{m-3},t_{m-3};\cdots;x^1,t_1)$$

重复应用条件概率密度函数公式和考虑到过程是马尔柯夫的，则最后得

$$p(x^m,t_m;x^{m-1},t_{m-1};\cdots;x^1,t_1)=p(x^m,t_m \mid x^{m-1},t_{m-1})\cdot p(x^{m-1},t_{m-1} \mid x^{m-2},t_{m-2})\cdots$$
$$p(x^2,t_2 \mid x^1,t_1)\cdot p(x^1,t_1) \qquad (1\text{-}1\text{-}6)$$

由上式可见，对于一个马尔柯夫过程$\{X(t);\ t\in T\}$只要知道了 $t=t_1$ 时刻过程的初始概率密度函数$p(x^1,t_1)$和在 $t,\tau\in T$中的所有 $t,\tau(t>\tau)$的转移概率密度函数 $p[x(t) \mid x(\tau)]$就完全确定了其联合概率密度函数。也就是说，如果 t_1 是 $t\in T$ 的一个初始时间，那末通过初始时

间上的概率密度函数和转移概率密度函数，可完全确定一个马尔柯夫过程。这是马尔柯夫过程的又一重要特性。

根据边际概率密度公式可得

$$\int_{-\infty}^{\infty} p(x^k, t_k; x^{k-1}, t_{k-1}; x^{k-2}, t_{k-2}) dx^{k-1} = p(x^k, t_k; x^{k-2}, t_{k-2})$$

而对于马尔柯夫过程有

$$p(x^k, t_k; x^{k-1}, t_{k-1}; x^{k-2}, t_{k-2}) = p(x^k, t_k | x^{k-1}, t_{k-1}) \cdot p(x^{k-1}, t_{k-1} | x^{k-2}, t_{k-2}) \cdot p(x^{k-2}, t_{k-2})$$

又由条件概率密度函数公式得

$$p(x^k, t_k; x^{k-2}, t_{k-2}) = p(x^k, t_k | x^{k-2}, t_{k-2}) \cdot p(x^{k-2}, t_{k-2})$$

于是综合以上三式的结果得

$$\int_{-\infty}^{\infty} p(x^k, t_k | x^{k-1}, t_{k-1}) \cdot p(x^{k-1}, t_{k-1} | x^{k-2}, t_{k-2}) dx^{k-1} = p(x^k, t_k | x^{k-2}, t_{k-2}) \tag{1-1-7}$$

这是马尔柯夫过程转移概率密度函数乘积的一个很重要的关系式。

上述马尔柯夫过程的概念还可以推广，从而得到所谓高阶马尔柯夫过程。例如，**一个 n 维矢量随机过程 $\{X(t);\ t\in T\}$ 在 $t\in T$ 中的任意 m 个时间点 $t_1<t_2<\cdots<t_m$（m 是任意整数）上，如果 $X(t_m)$ 对于 $X(t_{m-1})$，$X(t_{m-2}),\cdots,X(t_1)$ 的条件概率具有如下特性**

$$P\{X(t_m)\leqslant x^m | X(t_{m-1}) = x^{m-1},\ X(t_{m-2}) = x^{m-2}, \cdots, X(t_1) = x^1\} = P\{X(t_m)\leqslant x^m | X(t_{m-1}) = x^{m-1},\ X(t_{m-2}) = x^{m-2}\} \tag{1-1-8}$$

并且对所有 m 个 n 维随机矢量 $X(t_1)$，$X(t_2),\cdots,X(t_m)$ 都成立，则称此矢量随机过程为马尔柯夫 -2 过程。同样，还可定义马尔柯夫 -K 过程。

一个马尔柯夫-2 过程 $\{X(t);t\in T\}$，其任意 m 个 n 维随机矢量 $X^1=X(t_1),X^2=X(t_2),\cdots,X^m=X(t_m)$ 的条件概率分布函数和条件概率密度函数具有如下特点：

$$F(x^m, t_m | x^{m-1}, t_{m-1}; x^{m-2}, t_{m-2}; \cdots; x^1, t_1) = F(x^m, t_m | x^{m-1}, t_{m-1}; x^{m-2}, t_{m-2}) \tag{1-1-9}$$

和

$$p(x^m, t_m | x^{m-1}, t_{m-1}; x^{m-2}, t_{m-2}; \cdots; x^1, t_1) = p(x^m, t_m | x^{m-1}, t_{m-1}; x^{m-2}, t_{m-2}) \tag{1-1-10}$$

马尔柯夫过程还具有如下特点：

(1) 如果矢量随机过程 $\{X(t);\ t\in T\}$ 是马尔柯夫过程，而 T_1 是 T 的一个子集，则随机过程 $\{X(t); t\in T_1\}$ 也是马尔柯夫过程。这个特性的正确性可从马尔柯夫过程的定义立即得出。

(2) 马尔柯夫过程也是一个逆时间马尔柯夫过程。也就是说，如果 $\{X(t); t\in T\}$ 是马尔柯夫过程，则对于 $t_m<t_{m+1}<\cdots<t_{m+k}$（$m$ 和 k 都是任意整数）有

$$p(x^m, t_m | x^{m+1}, t_{m+1}; x^{m+2}, t_{m+2}; \cdots; x^{m+k}, t_{m+k}) = p(x^m, t_m | x^{m+1}, t_{m+1}) \tag{1-1-11}$$

这可证明如下：

由条件概率密度函数公式得

$$p(x^m, t_m | x^{m+1}, t_{m+1}; x^{m+2}, t_{m+2}; \cdots; x^{m+k}, t_{m+k}) = \frac{p(x^m, t_m; x^{m+1}, t_{m+1}; \cdots; x^{m+k} t_{m+k})}{p(x^{m+1}, t_{m+1}; x^{m+2}, t_{m+2}; \cdots; x^{m+k}, t_{m+k})}$$

由于过程是马尔柯夫的，因此上式可改写成

$$p(x^m,t_m|x^{m+1},t_{m+1};x^{m+2},t_{m+2};\cdots;x^{m+k},t_{m+k})$$

$$=\frac{p(x^{m+k},t_{m+k}|x^{m+k-1},t_{m+k-1})\cdot p(x^{m+k-1},t_{m+k-1}|x^{m+k-2},t_{m+k-2})\cdots}{p(x^{m+k},t_{m+k}|x^{m+k-1},t_{m+k-1})p(x^{m+k-1},t_{m+k-1}|x^{m+k-2},t_{m+k-2})\cdots}$$

$$\frac{p(x^{m+1},t_{m+1}|x^m,t_m)\cdot p(x^m,t_m)}{p(x^{m+2},t_{m+2}|x^{m+1},t_{m+1})p(x^{m+1},t_{m+1})}$$

$$=\frac{p(x^{m+1},t_{m+1};x^m,t_m)}{p(x^{m+1},t_{m+1})}=p(x^m,t_m|x^{m+1},t_{m+1})$$

从而特性(2)证得。

对于马尔柯夫 -2 过程同样具有以上两个特性。

类似地也可定义马尔柯夫序列，并且可以得到同以上类似的特性。

1-1-3 高斯-马尔柯夫过程

也讨论矢量随机过程。**如果一个 n 维矢量随机过程 $\{X(t);\ t\in T\}$ 既是高斯过程又是马尔柯夫过程，则称它为高斯-马尔柯夫过程；而如果一个 n 维矢量随机过程 $\{X(t);t\in T\}$ 既是高斯过程，又是马尔柯夫 -2 过程，则称它为高斯-马尔柯夫 -2 过程。**

类似地；**如果一个 n 维矢量随机序列 $\{X(k);\ k=1,\ \cdots,\ N\}$ 既是高斯序列，又是马尔柯夫序列，则称它为高斯-马尔柯夫序列；而如果一个 n 维矢量随机序列 $\{X(k);\ k=1,\ \cdots,\ N\}$ 既是高斯序列，又是马尔柯夫 -2 序列，则称它为高斯-马尔柯夫 -2 序列。**

显然，高斯-马尔柯夫过程（或序列）的高斯特性决定了它的幅度的概率分布特点，而马尔柯夫特性决定了过程（或序列）的概率分布特点在时间上的传播。

［例 1-1-1］ 设标量随机过程 $\{X(t);\ t\geqslant t_0\}$ 由如下方程确定

$$\dot{X}(t)=-\frac{1}{t+1}X(t)$$

其中 $X(t_0)$ 是平均值为零，方差为 $\sigma_0^2>0$ 的高斯分布随机变量。试分析其特性。

［解］ 首先由 $X(t)$ 所满足的方程，对一切 $t\geqslant t_0$ 可解得

$$X(t)=\frac{t_0+1}{t+1}X(t_0)$$

并且由于 $X(t_0)$ 是高斯分布的，因此根据高斯分布的特点：高斯分布随机变量的线性变换是高斯的，可以证明，$\{X(t);\ t\geqslant t_0\}$ 在 $t\geqslant t_0$ 的任意 m 个时间点 $t_1,\ t_2,\cdots,\ t_m$（m 是任意整数）上的联合分布也是高斯的，因此它是高斯随机过程。

其次，对于 $t_2>t_1\geqslant t_0$，容易求得

$$X(t_2)=\frac{t_1+1}{t_2+1}X(t_1)$$

可见，过程 $\{X(t);\ t\geqslant t_0\}$ 在将来时刻（$t=t_2$ 时）的统计特性完全可由现在时刻（$t=t_1$ 时）的状态决定，而与过去时刻（$t<t_1$ 时）的状态无关，因此它又是一个马尔柯夫过程。

综上所述，过程 $\{X(t);\ t\geqslant t_0\}$ 是一个高斯-马尔柯夫过程。由于高斯分布随机变量的条件分布是高斯的，因此其转移概率密度函数 $p(x(t)|x(\tau))(t>\tau)$ 是高斯分布的，并且它完全

由如下的条件均值和条件方差决定：

$$\begin{cases} E\{X(t)|x(\tau)\}=\dfrac{\tau+1}{t+1}x(\tau) \\ \mathrm{var}\{X(t)|x(\tau)\}=E\{[X(t)-E\{X(t)|x(\tau)\}]^2|x(\tau)\}=0 \end{cases}$$

其中 $x(\tau)$ 是随机变量 $X(\tau)$ 的取值。因此 $p(x(t)|x(\tau))$ 是一个 $\mathcal{N}\left(\dfrac{\tau+1}{t+1}x(\tau),0\right)$ 分布，亦即 $p(x(t)|x(\tau))$ 是一个 Dirac-δ 函数。这说明，$X(\tau)$ 的取值 $x(\tau)$ 对于一切 $t>\tau\geqslant t_0$ 能唯一地确定 $X(t)$ 的值。

由于 $\{X(t);\ t\geqslant t_0\}$ 是高斯分布的，并且

$$\begin{cases} E\{X(t)\}=E\left\{\dfrac{t_0+1}{t+1}X(t_0)\right\}=0; \\ \mathrm{var}\{X(t)\}=E\left\{\left[\dfrac{t_0+1}{t+1}X(t_0)\right]^2\right\}=\left(\dfrac{t_0+1}{t+1}\right)^2\cdot\sigma_0^2 \end{cases}$$

因此可得

$$p(x,t)=\frac{(t+1)}{\sqrt{2\pi}(t_0+1)\sigma_0}e^{-\frac{(t+1)^2}{2(t_0+1)^2\sigma_0^2}x^2}$$

［例 1-1-2］ 设标量随机过程 $\{X_1(t);\ t\geqslant t_0\}$ 由如下方程决定

$$\ddot{X}_1(t)=0$$

其中 $X_1(t_0)$ 与 $\dot{X}_1(t_0)$ 是具有联合高斯分布的随机变量。试分析此过程的特点。

［解］ 首先由 $X_1(t)$ 所满足的微分方程可解得

$$X_1(t)=X_1(t_0)+\dot{X}_1(t_0)(t-t_0)$$

由于 $X_1(t_0)$ 与 $\dot{X}_1(t_0)$ 是联合高斯分布的，因此同例 1-1-1 的理由，过程 $\{X(t);\ t\geqslant t_0\}$ 是一个高斯随机过程。

其次，设 $t_3>t_2>t_1\geqslant t_0$，则有

$$\begin{cases} X_1(t_3)=X_1(t_0)+\dot{X}_1(t_0)\cdot(t_3-t_0) \\ X_1(t_2)=X_1(t_0)+\dot{X}_1(t_0)\cdot(t_2-t_0) \\ X_1(t_1)=X_1(t_0)+\dot{X}_1(t_0)\cdot(t_1-t_0) \end{cases}$$

由此得

$$\begin{cases} X_1(t_3)=X_1(t_2)+\dot{X}_1(t_0)(t_3-t_2) \\ X_1(t_2)=X_1(t_1)+\dot{X}_1(t_0)(t_2-t_1) \end{cases}$$

于是

$$X_1(t_3)=X_1(t_2)+\frac{t_3-t_2}{t_2-t_1}[X_1(t_2)-X_1(t_1)]$$

由此可见，过程 $\{X_1(t),t\geqslant t_0\}$ 的将来（$t=t_3$ 时）统计特性完全由现在（$t=t_2$ 时）和前一时刻（$t=t_1$ 时）的统计特性确定，而与过去时刻（$t<t_1$）的状态无关，因此它是一个马尔柯夫-2 过程。实际上，由于描述过程 $\{X_1(t);\ t\geqslant t_0\}$ 的微分方程是二阶的，因此必需有两个积分常数才能够定其解，这就要求有上述两个时间点来确定过程的统计特性。

综上所述，过程 $\{X_1(t);\ t\geqslant t_0\}$ 是一个高斯-马尔柯夫 -2 过程。但是，如果设 $\dot{X}_1(t)=X_2(t)$ 并且令

$$X(t)=\begin{bmatrix}X_1(t)\\X_2(t)\end{bmatrix}=\begin{bmatrix}X_1(t)\\\dot{X}_1(t)\end{bmatrix}$$

则原微分方程可写成

$$\dot{X}(t)=\begin{bmatrix}0&1\\0&0\end{bmatrix}X(t)$$

并且可求得

$$\begin{cases}X(t_1)=\begin{bmatrix}1&t_1-t_0\\0&1\end{bmatrix}X(t_0)\\[2ex]X(t_2)=\begin{bmatrix}1&t_2-t_0\\0&1\end{bmatrix}X(t_0)\end{cases}$$

和

$$X(t_2)=\begin{bmatrix}1&t_2-t_1\\0&1\end{bmatrix}X(t_1)$$

可见，这时的二维矢量随机过程 $\{X(t);\ t\geqslant t_0\}$ 的将来时刻（$t=t_2$ 时）的统计特性就完全由其现在时刻（$t=t_1$ 时）的统计特性确定，而与其以前时刻（$t<t_1$ 时）的状态无关，因此它是一个马尔柯夫过程。由上面所得的结果还容易看出，由于初始状态是高斯分布的，因此过程 $\{X(t);\ t\geqslant t_0\}$ 也是高斯分布的。从而它是一个高斯-马尔柯夫过程。由此可见，一个高斯-马尔柯夫-2 过程，可以通过扩充状态（在本例中原是标量，扩充后变为二维矢量了）的办法，化为一个高斯-马尔柯夫过程。

1-1-4 二阶过程

二阶过程的基本概念 讨论标量随机过程的情况。**一个随机过程** $\{X(t);\ t\epsilon T\}$，**如果对于一切** $t\epsilon T$ **总有**

$$E\{X^2(t)\}<\infty \tag{1-1-12}$$

则称此过程为二阶过程。显然，对于一切二阶过程 $\{X(t);\ t\epsilon T\}$ 其平均值函数

$$\mu_X(t)=E\{X(t)\}$$

和自协方差函数

$$P_{XX}(t,\tau)=\mathrm{cov}\{X(t),X(\tau)\}=E\{[X(t)-\mu_X(t)][X(\tau)-\mu_X(\tau)]\}$$

总是存在的。

二阶过程的连续性　一个二阶过程 $\{X(t);\ t\epsilon T\}$，**如果对于一切** $t\epsilon T$ **有**

$$\lim_{h\to 0}E\{[X(t+h)-X(t)]^2\}=0 \tag{1-1-13}$$

则称它在 $t\epsilon T$ **内是处处均方连续的。**

一个二阶过程 $\{X(t);\ t\epsilon T\}$ 在 $t\epsilon T$ 内是处处均方连续的充要条件是：其平均值函数 $\mu_X(t)$ 在 t 处连续和其自协方差函数 $P_{XX}(t,\tau)$ 在 (t,t) 处连续。（证明略。）

由于自协方差函数，平均值函数和自相关函数之间具有如下关系（见附录B式(b-90)：

$$P_{XX}(t,\tau)=R_{XX}(t,\tau)-\mu_X(t)\mu_X(\tau)$$

因此 $\mu_X(t)$ 和 $P_{XX}(t,\tau)$ 在 t 和 (t,t) 处连续等效于 $R_{XX}(t,\tau)$ 在 (t,t) 处连续。因此上述均方连续的充要条件又可改述为：一个二阶过程 $\{X(t);\ t\in T\}$ 在 $t\in T$ 内处处均方连续的充要条件，是其自相关函数 $R_{XX}(t,\tau)$ 在 (t,t) 处连续。

最后，由附录B指出的均方收敛的性质可知，均方连续随机过程之和是均方连续的。还有，我们称一个二阶矢量随机过程是均方连续的，通常是指其每一个分量是均方连续的。

二阶过程的可微性　一个二阶过程 $\{X(t);\ t\in T\}$，如果对于一切 $t\in T$ 存在如下极限

$$\underset{h\to 0}{\mathrm{l\cdot i\cdot m}}\left[\frac{X(t+h)-X(t)}{h}\right]=\dot{X}(t) \tag{1-1-14}$$

即有

$$\lim_{h\to 0} E\left\{\left[\frac{X(t+h)-X(t)}{h}-\dot{X}(t)\right]^2\right\}=0 \tag{1-1-15}$$

则称它是均方可微的，并且称 $\{\dot{X}(t);\ t\in T\}$ 是 $\{X(t);\ t\in T\}$ 的均方导数。其中 $\mathrm{l\cdot i\cdot m}(\cdot)$ 是 $(\cdot)$ 的均方极限。

一个二阶过程 $\{X(t);\ t\in T\}$ 在 $t\in T$ 内为均方可微的充要条件，是其平均值函数 $\mu_X(t)$ 的导数 $\dot{\mu}_X(t)$ 在 t 存在和其自协方差函数的二阶混合偏导数 $\frac{\partial^2 P_{XX}(t,\tau)}{\partial t\partial\tau}$ 在 (t,t) 存在；或者等效地，其自相关函数 $R_{XX}(t,\tau)$ 的二阶混合偏导数 $\frac{\partial^2 R_{XX}(t,\tau)}{\partial t\partial\tau}$ 在 (t,t) 存在。(证明略)。

显然，如果一个二阶过程 $\{X(t);\ t\in T\}$ 对一切 $t\in T$ 是均方可微的，则对一切 $t\in T$ 它必均方连续。另外，直接由附录B所指出的均方收敛的性质和上述均方可微的定义可知，如果一个二阶过程 $\{X(t);\ t\in T\}$ 对一切 $t\in T$ 均方可微，则有

$$\mu_{\dot{X}}(t)=E\{\dot{X}(t)\}=\dot{\mu}_X(t) \tag{1-1-16}$$

$$P_{\dot{X}X}(t,\tau)=E\{[\dot{X}(t)-\mu_{\dot{X}}(t)][X(\tau)-\mu_X(\tau)]\}=\frac{\partial P_{XX}(t,\tau)}{\partial t} \tag{1-1-17}$$

$$P_{\dot{X}\dot{X}}(t,\tau)=E\{[\dot{X}(t)-\mu_{\dot{X}}(t)][\dot{X}(\tau)-\mu_{\dot{X}}(\tau)]\}=\frac{\partial^2 P_{XX}(t,\tau)}{\partial t\partial\tau} \tag{1-1-18}$$

$$R_{\dot{X}X}(t,\tau)=E\{\dot{X}(t)X(\tau)\}=\frac{\partial R_{XX}(t,\tau)}{\partial t} \tag{1-1-19}$$

$$R_{\dot{X}\dot{X}}(t,\tau)=E\{\dot{X}(t)\dot{X}(\tau)\}=\frac{\partial^2 R_{XX}(t,\tau)}{\partial t\partial\tau} \tag{1-1-20}$$

这些结果说明：对于一个二阶过程，均方求导运算与期望运算是可交换的。

应该指出，由附录B所指出的均方收敛的性质可知，均方可微二阶过程之和是均方可微的。并且其和的均方导数等于均方导数之和。另外还要指出，我们称一个二阶矢量随机过程是均方可微的，是指其每个分量是均方可微的。

［例 1-1-3］　设有随机过程 $\{X(t)=at;\ t\in T\}$ 和 $\{Y(t)=X^2(t)=a^2t^2;\ t\in T\}$，其中 a 是随机变量，并且 $E\{a^4\}<\infty$。

试证明：$\{\dot{X}(t)=a;\ t\in T\}$；$\{\dot{Y}(t)=2a^2t;\ t\in T\}$

［解］　首先由假定可知 $\{X(t);\ t\in T\}$ 和 $\{Y(t);\ t\in T\}$ 都是二阶过程，并且由于

$$E\left\{\left[\frac{a(t+h)-at}{h}-a\right]^2\right\}\equiv 0$$

因此

$$\lim_{h\to 0} E\left\{\left[\frac{a(t+h)-at}{h}-a\right]^2\right\}\equiv 0$$

亦即

$$\{\dot{X}(t)=a; t\in T\}$$

另一方面，由于

$$\left[\frac{a^2(t+h)^2-a^2t^2}{h}-2a^2t\right]^2=a^4h^2$$

因此

$$\lim_{h\to 0} E\left\{\left[\frac{a^2(t+h)^2-a^2t^2}{h}-2a^2t\right]^2\right\}=\lim_{h\to 0} E\{a^4\}h^2=0$$

亦即

$$\{\dot{Y}(t)=2a^2t;\ t\in T\}$$

［例 1-1-4］ 设有随机过程 $\{X(t)=\sin at;\ t\in T\}$，其中 a 是一个随机变量，并且 $E\{a^4\}<\infty$。试证明：$\{\dot{X}(t)=a\cos at;\ t\in T\}$。

［解］ 根据假设，$\{X(t)=\sin at;\ t\in T\}$ 是一个二阶过程。由于

$$\left[\frac{\sin a(t+h)-\sin at}{h}-a\cos at\right]^2$$

$$=\left[\frac{\sin at(\cos ah-1)+\cos at(\sin ah-ah)}{h}\right]^2$$

$$\leqslant 2\left[\frac{\sin^2 at(\cos ah-1)^2}{h^2}\right]+2\left[\frac{\cos^2 at(\sin ah-ah)^2}{h^2}\right]$$

又由于 $\cos\alpha\leqslant 1+\alpha^2$，$\sin\alpha\leqslant\alpha+\alpha^2$ 因此由上式得

$$\left[\frac{\sin a(t+h)-\sin at}{h}-a\cos at\right]^2\leqslant 2a^4h^2\sin^2 at+2a^4h^2\cos^2 at$$

$$=2a^4h^2$$

于是

$$\lim_{h\to 0} E\left\{\left[\frac{\sin a(t+h)-\sin at}{h}-a\cos at\right]^2\right\}=\lim_{h\to 0} 2E\{a^4\}h^2=0$$

亦即

$$\{\dot{X}(t)=a\cos at;\ t\in T\}$$

二阶过程的可积性 **设 $\{X(t);\ t\in T\}$ 是一个二阶过程，如果 $a,b\in T$，$a\leqslant b$，并将区间 $[a,b]$ 划分为 $a=t_0<t_1<\cdots<t_{n-1}<t_n=b$ 和 $\rho=\max\limits_{0\leqslant i\leqslant n-1}(t_{i+1}-t_i)$，$t_i\leqslant t'_i\leqslant t_{i+1}$，则当如下极限存在①**

$$\underset{\rho\to 0}{\mathrm{l\cdot i\cdot m}}\left(\sum_{i=0}^{n}X(t'_i)(t_{i+1}-t_i)\right)=\int_a^b X(t)\,dt \qquad (1\text{-}1\text{-}21)$$

即有

$$\lim_{\rho\to 0} E\left\{\left[\sum_{i=0}^{n-1}X(t'_i)(t_{i+1}-t_i)-\int_a^b X(t)\,dt\right]^2\right\}=0 \qquad (1\text{-}1\text{-}22)$$

① 在二阶过程的可积性中积分号上的 a 即为 a。

时，就称此二阶过程是均方黎曼（Riemann）可积的，并且称$\int_a^b X(t)dt$是$\{X(t);\ t\in T\}$在区间$[a,b]$上的均方黎曼积分。

一个二阶过程$\{X(t);\ t\in T\}$，在区间$[a,b]$上（其中$a,b\in T$，$a\leqslant b$）为均方黎曼可积的充要条件，是其平均值函数$\mu_X(t)$在区间$[a,b]$黎曼可积和其自协方差函数$P_{XX}(t,\tau)$在区域$[a,b]\times[a,b]$上黎曼可积；或等效地，其自相关函数$R_{XX}(t,\tau)$在区域$[a,b]\times[a,b]$上黎曼可积。（证明略。）

另外，如果二阶过程$\{X(t);\ t\in T\}$，对一切$t\in T$在区间$[a,b]$上是均方连续的，则它在区间$[a,b]$上必是均方黎曼可积的。

由附录B所指出的均方收敛的性质和上述均方黎曼可积的定义直接可推得：如果二阶过程$\{X(t);\ t\in T\}$和$\{Y(t);\ t\in T\}$在$t\in T$中是均方黎曼可积的，则对于$a,b\in T$和$c,d\in T$有

$$E\left\{\int_a^b X(t)dt\right\}=\int_a^b \mu_X(t)dt \tag{1-1-23}$$

$$E\left\{\int_a^b X(t)dt\int_c^d X(\tau)d\tau\right\}=\int_a^b\int_c^d R_{XX}(t,\tau)dtd\tau \tag{1-1-24}$$

$$\mathrm{cov}\left\{\int_a^b X(t)dt\ \int_c^d X(\tau)d\tau\right\}=\int_a^b\int_c^d P_{XX}(t,\tau)dtd\tau \tag{1-1-25}$$

上述结果说明，均方黎曼积分运算与期望运算是可交换的。利用式（1-1-24）和许瓦茨（Schwars）不等式还可以说明

$$\left[\int_a^b\int_c^d R_{XY}(t,\tau)\ dtd\tau\right]^2\leqslant\int_a^b\int_a^b R_{XX}(t,\tau)dtd\tau\ \cdot\int_c^d\int_c^d R_{YY}(t,\tau)dtd\tau \tag{1-1-26}$$

应该指出，根据附录B所指出的均方收敛的性质可知，在积分区间内的均方黎曼积分是线性的和可加的。而所谓一个二阶矢量随机过程是均方黎曼可积的，是指其每个分量是均方黎曼可积的。

由均方黎曼可积的定义知道，如果一个二阶过程$\{X(t);\ t\in T\}$对每一个$t\in[a,b]$在区间$[a,t]$上是均方黎曼可积的，则

$$Y(t)=\int_a^t X(\tau)d\tau \tag{1-1-27}$$

是一个定义在区间$[a,b]$上的随机过程$\{Y(t);\ t\in[a,b]\}$。显然，对于一切$t\in[a,b]$有

$$E\{Y(t)^2\}<\infty \tag{1-1-28}$$

因此$\{Y(t);\ t\in[a,b]\}$也是二阶过程；并且它具有如下特性：

（i） $\{Y(t);\ t\in[a,b]\}$在区间$[a,b]$均方连续；

(ii) 如果二阶过程$\{X(t);t\in T\}$在区间$[a,b]$上是均方连续的，则二阶过程$\{Y(t);\ t\in[a,b]\}$在区间$[a,b]$上是均方可微的，并且

$$\{\dot{Y}(t)=X(t);\ t\in[a,b]\} \tag{1-1-29}$$

最后，如果二阶过程$\{\dot{X}(t);\ t\in T\}$在区间$[a,b]$上均方黎曼可积，则对于$t\in[a,b]$，以概率1有

$$\left\{X(t)-X(a)=\int_a^t X(\tau)d\tau;\ t\in[a,b]\ \right\} \tag{1-1-30}$$

以上所有结果的证明，感兴趣的读者请见参考资料[8]的 3.4 节。

[例 1-1-5] 设有随机过程 $\{Z(t)=2a^2t;\ t\in T\}$，其中 a 是随机变量，并且 $E\{a^4\}<\infty$。试证明：

$$\left\{\int_0^t Z(\tau)d\tau = a^2t^2;\ t\in T\right\}$$

[解] 由假设条件知道，$\{Z(t)=2a^2t;\ t\in T\}$ 是一个二阶过程，并且由均方黎曼积分的定义得

$$\int_0^t Z(\tau)d\tau = \underset{\rho\to 0}{\text{l·i·m}}\left(\sum_{i=0}^{n-1}2a^2\tau'_i(\tau_{i+1}-\tau_i)\right)$$

由附录B所指出的均方收敛的性质（3）得

$$\int_0^t Z(\tau)d\tau = \underset{\rho\to 0}{\text{l·i·m}}\left(2a^2\sum_{i=0}^{n-1}\tau'_i[\tau_{i+1}-\tau_i]\right)$$

$$=2a^2\lim_{\rho\to 0}\left(\sum_{i=0}^{n-1}\tau'_i(\tau_{i+1}-\tau_i)\right)=a^2t^2$$

于是

$$\left\{\int_0^t Z(\tau)d\tau = a^2t^2;\ t\in T\right\}$$

证得。

1-1-5 独立增量过程和维纳过程

也讨论标量随机过程的情况。

独立增量过程 **一个随机过程** $\{X(t);\ t\in T\}$，**对一切** $t_i\in T$ （$t_{i+1}>t_i$，$i=0,1\cdots,N$），**如果恒有随机变量**

$X(t_0)$；$X(t_1)-X(t_0)$；$X(t_2)-X(t_1)$；…；$X(t_N)-X(t_{N-1})$

是相互独立的，则称此随机过程是一个独立增量过程。

一个独立增量过程 $\{X(t);\ t\in T\}$，如果对一切 $t,\tau\in T$ 和 $t+h$，$\tau+h\in T$，恒有增量 $X(t+h)-X(\tau+h)$ 的概率分布与增量 $X(t)-X(\tau)$ 的概率分布相同，则称此独立增量过程是平稳独立增量过程；而如果增量 $X(t)-X(\tau)$ 的概率分布是高斯的，则称此独立增量过程是高斯独立增量过程。

如果设 $X(t_0)=X_0$，$X(t_1)=X_1$，…，$X(t_N)=X_N$，则可将 X_N 表示成下式

$$X_N = X_0+\sum_{i=1}^{N}(X_i-X_{i-1}) \tag{1-1-31}$$

这就是说，X_N 是 $N+1$ 个独立随机变量之和。因此，对于独立增量过程来说，只要知道随机变量 X_0 的概率分布和增量 $X(t)-X(\tau)$ 的概率分布，那末，这个过程的概率分布也就完全确定了。

一个随机过程 $\{X(t);\ t\in T\}$，对于一切 $t_i\in T$ （其中 $t_{i+1}>t_i$，$i=0$，1，2，…，N）所相应的随机变量 $X(t_0)$，$X(t_1)-X(t_0)$，…，$X(t_N)-X(t_{N-1})$，如果恒有