随机过程统计

（苏）P·Ш·里普切尔　A·H·史里亚耶夫著

张纬国译

SUIJIGUOCHENGTONGJI

宇航出版社

随机过程统计

非线性滤波及有关问题

〔苏〕 P. Ш. 里普切尔
A. H. 史里亚耶夫 著

张纬国 译
张志鸿 校

宇航出版社

内 容 简 介

本书针对离散时间和连续时间两种情况系统地阐述了最佳非线性滤波理论。关于序贯估计、线性滤波(卡尔曼-布西滤波器)以及随机过程的平滑和预测等应用问题在本书中占了很重要的地位，书中对鞅理论的主要方面进行了实质性的论述。

本书既可供从事概率及数学统计理论研究的专业人员，也可供在实际工作中使用概率统计方法解决信号从噪声中分离、统计假设的识别、斯笃客体按不完全数据的最佳控制等问题的读者阅读。

СТАТИСТИКА СЛУЧАЙНЫХ ПРОЦЕССОВ

Р. Ш. ЛИПЦЕР, А. Н. ШИРЯЕВ

ИЗДАТЕЛЬСТВО «НАУКА» 1970

随机过程统计

〔苏〕 Р. Ш. 里普切尔
A. H. 史里亚耶夫 著

张纬国 译

张志鸿 校

宇航出版社出版

新华书店北京发行所发行

各地新华书店经售

北京科技印刷厂印刷

开本：850×1168 1/32	字数：617 千字
1987 年 2 月第 1 版	印张：23.75
1987 年 2 月第 1 次印刷	印数：1—5,000 册
统一书号：15244·0096	定价：5.50 元

目　录

绪 论

1.随机过程的统计问题大部分是在以下概型范围内来描述的。

在某个概率空间 $(\Omega, \mathscr{F}, \mathbf{P})$ 上给出了部分可观测的随机过程 $(\theta, \xi) = (\theta_t, \xi_t)$，$t \geqslant 0$，在这一过程中，只可能观测第二个分量 $\xi = (\xi_t)$，$t \geqslant 0$。在每一瞬时 t 都要求基于对 $\xi_0^t = \{\xi_s, 0 \leqslant s \leqslant t\}$ 的观测给出（非观测）值 θ_t 的估计。θ_t 依 ξ_0^t 的这种估计问题(又称之为滤波问题)将在本书中进行研究。

众所周知，如果 $\mathbf{M}\theta_t^2 < \infty$，则验后平均值 $m_t = \mathbf{M}(\theta_t | \mathscr{F}_t^\xi)$ 就是均方意义内 θ_t 依 ξ_0^t 的最佳估计，其中，$\mathscr{F}_t^\xi = \sigma\{\omega: \xi_s, s \leqslant t\}$ 为由变量 ξ_0^t 产生的 σ-代数。因此，(在均方意义内)解最佳滤波问题归结为寻求条件数学期望 $m_t = \mathbf{M}(\theta_t | \mathscr{F}_t^\xi)$。

原则上，条件数学期望 $\mathbf{M}(\theta_t | \mathscr{F}_t^\xi)$ 可按贝叶斯 (Bayes) 公式计算。但是，即使在很多比较简单的情况下利用贝叶斯公式所得到的表达式也是十分繁杂的，即无论对实际应用，还是对这样获得的结构和性质的研究都是很困难的。

从计算的观点来看，希望确定"滤波器" m_t，$t \geqslant 0$ 的公式应具有递推的性质。粗略地说，这意味着值 $m_{t+\Delta}$，$\Delta > 0$，应按值 m_t 和观测 $\xi_t^{t+\Delta} = \{\xi_s, t \leqslant s \leqslant t + \Delta\}$ 来恢复。在 $t = 0, 1, 2, \cdots$ 离散时间的情况下，例如，方程

$$\Delta m_t = a(t, m_t) + b(t, m_t)(\xi_{t+1} - \xi_t), \qquad (1)$$

可以作为这种递推关系的最简单的形式，式中，$\Delta m_t = m_{t+1} - m_t$。

在 $t \geqslant 0$ 的连续时间的情况下，斯笃微分方程*)

$$dm_t = a(t, m_t)dt + b(t, m_t)d\xi_t \tag{2}$$

具有这种形式。

很清楚，若对过程 (θ, ξ) 的结构不作专门的假设，很难设想最佳估计 m_t 将满足(1)、(2)形式的递推关系。因此，在阐述我们所研究的过程 (θ, ξ) 的结构之前（对此过程而言，本书中研究的是滤波问题），我们先从一些实例入手。

设 θ——为具 $\mathbf{M}\theta = m$, $D\theta = \gamma$ 的高斯随机变量，为简单起见，记为 $\theta \sim N(m, \gamma)$。假定序列

$$\xi_t = \theta + \varepsilon_t,\ t = 1, 2, \cdots, \tag{3}$$

应被观测，式中，$\varepsilon_1, \varepsilon_2, \cdots$——为彼此（且对 θ）独立的高斯随机变量，具零均值和单位方差。使用正态相关定理（定理 13.1）**) 则很容易求出估计 $m_t = \mathbf{M}(\theta|\xi_1, \cdots, \xi_t)$ 和"跟随"误差 $\gamma_t = \mathbf{M}(\theta - m_t)^2$ 的公式：

$$m_t = \frac{m + \sum_{i=1}^{t} \xi_i}{1 + \gamma t},\quad \gamma_t = \frac{\gamma}{1 + \gamma t} \tag{4}$$

从而就得到 m_t 和 γ_t 的以下递推方程：

$$\Delta m_t = \frac{\gamma_t}{1 + \gamma_t}[\xi_{t+1} - m_t], \tag{5}$$

$$\Delta \gamma_t = -\gamma_t^2/(1 + \gamma_t), \tag{6}$$

式中，$\Delta m_t = m_{t+1} - m_t$，$\Delta \gamma_t = \gamma_{t+1} - \gamma_t$。

使所研究的例子再复杂一点。设 θ 及 $\varepsilon_1, \varepsilon_2, \cdots$ 和前面例子中假定的一样，而可观测过程 ξ_t, $t = 1, 2, \cdots$，由下列关系式确

---

*) "斯笃"是俄文"стохастический"，英文"stochastic"（随机的）词头译音。它们都来源于希腊文"στοχαστιξ"，原字是"中的"（打中目标）的意思。"随机微分方程"、"随机过程"等和音译的"斯笃微分方程"、"斯笃过程"这两种称呼法在各国都已通行——译注。

**) 本书中，定理、引理和公式采用双重编号。第一个数字表示章号，而第二个数字则表示章内的序号。

定:

$$\xi_{t+1} = A_0(t, \xi) + A_1(t, \xi)\theta + \varepsilon_{t+1} \tag{7}$$

其中，函数 $A_0(t, \xi)$ 和 $A_1(t, \xi)$ 假定为 $\mathscr{F}_t^\xi = \{\omega: \xi_0, \cdots, \xi_t\}$ -可测(即 $A_0(t,\xi)$和 $A_1(t, \xi)$当每个 t 时只取决于值 $\xi_0, \cdots, \xi_t$)。

我们指出，研究系数 $A_0(t, \xi)$ 和 $A_1(t, \xi)$ (取决于全部过去值($\xi_0, \cdots, \xi_t$))的必要性是由下列一些问题中产生的，例如，在控制论中(14 章 §3 节) 这些系数起“控制”作用，在信息论中 (16 章 § 4 节) 其中对偶函数 ($A_0(t, \xi)$, $A_1(t, \xi)$) 被解释为使用无噪声反馈的“编码”。

还有，概型 (7) 的最佳估计 $m_t = M(\theta_t|\mathscr{F}_t^\xi)$ 和条件方差 $\gamma_t = \mathbf{M}[(\theta - m_t)^2|\mathscr{F}_t^\xi]$ 也服从于递推方程(见 13 章§5 节):

$$\Delta m_t = \frac{\gamma_t A_1(t, \xi)}{1 + A_1^2(t, \xi)\gamma_t}(\xi_{t+1} - A_0(t, \xi) - A_1(t, \xi)m_t),$$

$$m_0 = m, \tag{8}$$

$$\Delta\gamma_t = -\frac{A_1^2(t, \xi)\gamma_t^2}{1 + A_1^2(t, \xi)\gamma_t}, \quad \gamma_0 = \gamma。 \tag{9}$$

在概型(3)和(7)中，实质上，这里所谈的问题是数理统计的传统问题——随机参数依观测值 ξ_0^t 的贝叶斯估计。使概型(7)更为复杂的以下步骤在于，用研究随机过程 θ_t 来代替随机变量 θ。

假定，随机过程 $(\theta, \xi) = (\theta_t, \xi_t)$, $t = 0, 1, \cdots$, 是用以下递推方程描述的:

$$\begin{aligned} \theta_{t+1} &= a_0(t, \xi) + a_1(t, \xi)\theta_t + b(t, \xi)\varepsilon_1(t + 1), \\ \xi_{t+1} &= A_0(t, \xi) + A_1(t, \xi)\theta_t + B(t, \xi)\varepsilon_2(t + 1), \end{aligned} \tag{10}$$

式中，$\varepsilon_1(t)$, $\varepsilon_2(t)$, $t = 1, 2, \cdots$, ——为有正态分布 $N(0, 1)$ 并且也不取决于 (θ_0, ξ_0) 的独立随机变量序列。系数 $a_0(t, \xi)$, $\cdots$, $B(t, \xi)$ 假定在每一个 $t = 0, 1, \cdots$时为 $\mathscr{F}_t^\xi$-可测的。

为了获得估计 $m_t = \mathbf{M}(\theta_t|\mathscr{F}_t^\xi)$ 和条件方差 $\gamma_t = \mathbf{M}\{[\theta_t - m_t]^2|\mathscr{F}_t^\xi\}$ 的递推方程，假定条件分布 $\mathbf{P}(\theta_0 \leqslant x|\xi_0)$ (对几乎所有的 ξ_0) 为正态分布，$N(m, \gamma)$。这一假定的实质在于，它可以

证明(见 13 章)由方程 (10) 制约的序列 (θ, ξ) 为条件高斯序列。这意味着,其中条件分布 $\mathbf{P}(\theta_t \leqslant x | \mathscr{F}_0^\xi)$ (几乎一定) 是高斯分布。但这一分布只是用本身的两个条件矩 m_t, γ_t 来描述的,这就给出获得条件矩的以下闭合方程组的可能性:

$$
\begin{aligned}
m_{t+1} &= a_0 + a_1 m_t + \frac{a_1 A_1 \gamma_t}{B^2 + A_1^2 \gamma_t}[\xi_{t+1} - A_0 - A_1 m_t], \quad m_0 = m, \\
\gamma_{t+1} &= [a_1^2 \gamma_t + b^2] - \frac{(a_1 A_1 \gamma_t)^2}{B^2 + A_1^2 \gamma_t}, \quad \gamma_0 = \gamma,
\end{aligned}
\tag{11}
$$

(为书写简单起见,系数 $a_0, \cdots, B$ 中已省略了引数 t 和 ξ)。

方程(11)(在更一般的情况下)将在第 13 章中推出。对于此方程的推导而言,实际上,除了与正态相关定理有关外,什么也不要求了。也是在第 13 章里,将导出外推最佳估计(即当 $\tau > t$ 时 θ_τ 依 ξ_0^t 的估计)的方程和内插最佳估计 (即当 $\tau < t$ 时 θ_τ 依 ξ_0^t 的估计)的方程。这些方程对于随机序列各种不同统计问题、控制问题及有关构造线性代数伪解问题的应用将在第十四章中描述。

这两章可以与书的其它材料分开来单独阅读,这就是说,使关心非线性滤波问题但随机过程的一般理论知识还不充分的读者可从这两章入手。

2. 本书的主要材料就是连续时间情况下的最佳滤波问题 (以及内插、外推、序贯估计、假设的识别等有关问题)。在连续时间情况下这些问题之所以令人感兴趣是因为 (除了它们本身的兴趣之外)对它们已成功地获得简洁表达与紧凑的公式。还必须补充说,有时方便的是先从研究与连续时间情况相类似的离散时间问题入手,然后再利用获得的结果来研究开始提出的问题。

(对连续时间情况)所作出的简洁表达,当然不是无代价的给出的——必须采用很复杂的随机过程理论工具。在本书中所使用的有关方法和工具的更具体的内容,下面就会谈到,而现在的目的是举例说明我们所研究滤波的一些情况。

假定,部分可观测随机过程 $(\theta, \xi) = (\theta_t, \xi_t)$, $t \geqslant 0$, 为高

斯过程，可用以下斯笃微分方程来描述（与系 (10) 相比较）：

$$d\theta_t = a(t)\theta_t dt + b(t)dw_1(t), \quad d\xi_t = A(t)\theta_t dt + B(t)dw_2(t), \tag{12}$$

式中，$w_1(t)$ 和 $w_2(t)$——为彼此且与 (θ_0, ξ_0) 独立的标准维纳过程，而 $B(t) \geqslant C > 0$。我们认为，分量 $\theta = (\theta_t)$，$t \geqslant 0$，不可观测。所研究的滤波问题在于，在每一个瞬时 $t \geqslant 0$，使 θ_t 依 ξ_0^t 的估计（在均方意义内）为最佳估计。

根据假定，过程 (θ, ξ) 为高斯过程，因此，最佳估计 $m_t = \mathbf{M}(\theta_t | \mathscr{F}_t^\xi)$ 线性地取决于观察量 $\xi_0^t = \{\xi_s, s \leqslant t\}$。确切地说，存在有函数（引理 10.1）$G(t, s)$，具 $\int_0^t G^2(t, s)ds < \infty$，$t \geqslant 0$，（几乎一定）使

$$m_t = m_0 + \int_0^t G(t, s)d\xi_s。\tag{13}$$

如果形式上将此表达式微分，则我们就得到

$$dm_t = G(t, t)d\xi_t + \left(\int_0^t \frac{\partial G(t, s)}{\partial t} ds\right) dt。\tag{14}$$

如果函数 $G(t, s)$ 满足维纳-何甫 (Wiener-Hopf) 方程（见 10.25）则这一表达式右边可以进行变换，在我们所研究的情况下，维纳-何甫方程归结为方程

$$\frac{\partial G(t, s)}{\partial t} = \left[a(t) - \gamma_t \frac{A^2(t)}{B^2(t)}\right] G(t, s), \quad t > s, \tag{15}$$

$$G(s, s) = \frac{\gamma_s A(s)}{B^2(s)}, \quad \gamma_s = \mathbf{M}[\theta_s - m_s]^2。\tag{16}$$

考虑到 (15) 和 (16) 式，即得到最佳估计 m_t，$t \geqslant 0$，必满足线性斯笃微分方程

$$dm_t = a(t)m_t dt + \frac{\gamma_t A(t)}{B^2(t)} [d\xi_t - A(t)m_t dt]。\tag{17}$$

包含在这一方程中的跟随误差量 $\gamma_t = \mathbf{M}[\theta_t - m_t]^2$，是黎卡堤 (Riccati) 方程

$$\dot{\gamma}_t = 2a(t)\gamma_t - \frac{A^2(t)\gamma_t^2}{B^2(t)} + b^2(t) \tag{18}$$

的解（对过程 $[\theta_t - m_t]$ 的平方应用伊藤变量代换公式并最后取平均值，则很容易得到方程(18))。

我们要更详细地谈谈方程(17)，为简单起见，认为 $\xi_0 = 0$。我们表示为

$$\bar{w}_t = \int_0^t \frac{d\xi_s - A(s)m_s ds}{B(s)}。 \tag{19}$$

于是，方程 (17) 可写为以下形式:

$$dm_t = a(t)m_t dt + \frac{\gamma_t A(t)}{B(t)} d\bar{w}_t。 \tag{20}$$

所引入的过程 $(\bar{w}_t)$，$t \geqslant 0$，十分重要，在滤波问题中起重要作用。问题在于，第一，该过程可证明是维纳过程（相对于 σ-代数系 $(\mathscr{F}_t^\xi)$，$t \geqslant 0$)，第二，它本身包含有与过程 ξ 相同的"信息"。确切地说，这意味着对所有 $t \geqslant 0$，σ-代数 $\mathscr{F}_t^{\bar{w}} = \sigma\{\omega: \bar{w}_s, s \leqslant t\}$ 和 $\mathscr{F}_t^\xi = \sigma\{\omega: \xi_s, s \leqslant t\}$ 相合:

$$\mathscr{F}_t^{\bar{w}} = \mathscr{F}_t^\xi, \quad t \geqslant 0 \tag{21}$$

(见定理 7.16)。

正是过程 $\bar{w}$ 的这些性质就有基础将它称为更新过程 (innovation process, обновляющий процесс)。

由 σ-代数 $\mathscr{F}_t^\xi$ 和 $\mathscr{F}_t^{\bar{w}}$ 的相合性可得出，对 m_t 而言，不仅等式(13)是正确的，而且表达式

$$m_t = m_0 + \int_0^t F(t, s) d\bar{w}_s \tag{22}$$

也是正确的。式中，$\bar{w} = (\bar{w}_t)$，$t \geqslant 0$，——为更新过程，而函数 $F(t, s)$ 是这样的，它使 $\int_0^t F^2(t, s) ds < \infty$。在正文中（定理 7.16）将证明，表达式 (22) 实质上由（扩散型过程泛函结构的）一般结果即可得到。以表达式 (22) 为依据，使用较之以表达式 (13) 为根据的更为简单的方法即可导出方程 (20)。尽管如此，还必须指

出，表达式 (22) 的证明本身要求要比规定表达式 (13) 的正确性更困难一些。

在所研究的起源于卡尔曼和布西的例子中，最佳滤波乃是线性的，并且是在假定 (θ, ξ) 为高斯过程的前提下导出的。而现在我们再举些最佳滤波乃是非线性的例子。

设 (θ_t), $t \geqslant 0$——马尔科夫过程，它是由零开始且在随机瞬时 σ 具有 0 和 1 两个状态及有 $0 \to 1$ 单次转换的过程，而随机瞬时 σ 的分布（根据马尔科夫性的假定）为指数形式：$\mathbf{P}(\sigma > t) = e^{-\lambda t}$, $\lambda > 0$。假定可观测过程 $\xi = (\xi_t)$, $t \geqslant 0$, 有微分

$$d\xi_t = \theta_t dt + dw_t, \ \xi_0 = 0, \tag{23}$$

式中，$w = (w_t)$, $t \geqslant 0$——为不取决于过程 $\theta = (\theta_t)$, $t \geqslant 0$, 的维纳过程。

将过程 θ 在 σ 瞬时由 0 状态向 1 状态转换称为出现了《失效》。这就产生以下问题：在每个时刻 $t > 0$, 根据观测 ξ_0^t 要确定出在该瞬时之前是否出现了《失效》? 我们表示为 $\pi_t = \mathbf{P}(\theta_t = 1 | \mathscr{F}_t^\xi) = \mathbf{P}(\sigma \leqslant t | \mathscr{F}_t^\xi)$。很清楚，$\pi_t = m_t = \mathbf{M}(\theta_t | \mathscr{F}_t^\xi)$。因此，验后概率 π_t, $t \geqslant 0$（在均方意内）为不可观测过程 $\theta = (\theta_t)$, $t \geqslant 0$ 的最佳状态估计。

对于验后概率 π_t, $t \geqslant 0$ 而言，可导出以下斯笃微分方程(比如，使用贝叶斯公式及维纳测度相对于相应过程 ξ 测度的导数结果)：

$$d\pi_t = \lambda(1 - \pi_t)dt + \pi_t(1 - \pi_t)[d\xi_t - \pi_t dt], \pi_0 = 0。\tag{24}$$

需着重指出，如果在卡尔曼-布西概型中，最佳《滤波器》是线性的，则方程 (24) 实质上是非线性的。因此，可将方程 (24) 定义为最佳非线性滤波。

和在前面例子中一样，(更新)过程

$$\bar{w}_t = \int_0^t [d\xi_s - \pi_s ds], \ t \geqslant 0,$$

可证明是维纳过程，且 $\mathscr{F}_t^{\bar{w}} = \mathscr{F}_t^\xi$, $t \geqslant 0$。因而，方程 (24) 也

可写为以下等价形式：

$$d\pi_t = \lambda(1-\pi_t)dt + \pi_t(1-\pi_t)d\bar{w}_t,\ \pi_0 = 0。 \tag{25}$$

3. 可以发现，所有这些例子属于本书中所采用的以下一般概型范围。

设 $(\Omega, \mathscr{F}, \mathbf{P})$——某一概率空间，在该空间上能分离出非减的 σ-代数族 $(\mathscr{F}_t)$，$t \geqslant 0 (\mathscr{F}_s \subseteq \mathscr{F}_t \subseteq \mathscr{F},\ s \leqslant t)$。在此概率空间上，假定已给出部分可观测过程 (θ_t, ξ_t)，$t \geqslant 0$，和被估计过程 (h_t)，$t \geqslant 0$ 一般说来后者不仅取决于不可观测过程 θ_t，$t \geqslant 0$，而且也取决于可观测分量 (ξ_t)，$t \geqslant 0$。

假定可观测过程[*] $\xi = (\xi_t, \mathscr{F}_t)$ 有斯笃微分

$$d\xi_t = A_t(\omega)dt + dw_t,\ \xi_0 = 0, \tag{26}$$

式中，$w = (w_t, \mathscr{F}_t)$，$t \geqslant 0$，——标准的维纳过程（即具连续轨迹的且当 $t \geqslant s$ 时具 $\mathbf{M}[(w_t - w_s)^2 | \mathscr{F}_s] = t - s$ 及 $w_0 = 0$ 的平方可积鞅（квадратично интегрируемый мартингал）），而当 $A = (A_t(\omega), \mathscr{F}_t)$，$t \geqslant 0$，——某个可积的随机过程[**]。

不可观测过程 $\theta = (\theta_t, \mathscr{F}_t)$，$t \geqslant 0$ 的结构不能直接具体化，而假定被估计过程 $h = (h_t, \mathscr{F}_t)$，$t \geqslant 0$，允许有以下表达式：

$$h_t = h_0 + \int_0^t a_s(\omega)ds + x_t,\ t \geqslant 0, \tag{27}$$

式中，$a = (a_t(\omega), \mathscr{F}_t)$，$t \geqslant 0$，——某一可积过程，$x = (x_t, \mathscr{F}_t)$，$t \geqslant 0$，——平方可积鞅。

对所有可积过程 $g = (g_t, \mathscr{F}_t)$，$t \geqslant 0$，我们用 $\pi_t(g) = \mathbf{M}[g_t | \mathscr{F}_t^\xi]$ 表示。于是，如果 $\mathbf{M}g_t^2 < \infty$，则 $\pi_t(g)$（在均方意义内）将是 g_t 依 $\xi_0^t = \{\xi_s, s \leqslant t\}$ 的最佳估计。

本书的主要结果之一（定理 8.1）断言，$\pi_t(h)$ 的以下表达式是正确的：

---

[*] 记号 $\xi = (\xi_t, \mathscr{F}_t)$ 表示量 ξ_t 当每个 $t \geqslant 0$ 时为 $\mathscr{F}_t$-可测。

[**] 实际上在本书中研究的是更一般类型的过程 ξ（见 8 章）。

$$\pi_t(h)=\pi_0(h)+\int_0^t \pi_s(a)ds+\int_0^t \pi_s(D)d\bar{w}_s$$

$$+\int_0^t [\pi_s(h_A)-\pi_s(h)\pi_s(A)]d\bar{w}_s。 \qquad (28)$$

这里，$\bar{w}=(\bar{w}_t, \mathscr{F}_t^\xi)$，$t\geqslant 0$,——维纳过程(与上述两个例子中的更新过程相比较)，而过程 $D=(D_t, \mathscr{F}_t)$，$t\geqslant 0$，表示维纳过程 $w=(w_t, \mathscr{F}_t)$，$t\geqslant 0$，和鞅 $x=(x_t, \mathscr{F}_t)$，$t\geqslant 0$ 之间的《相关性》。确切地说，过程

$$D_t=\frac{d\langle x, w\rangle_t}{dt}, \quad t\geqslant 0, \qquad (29)$$

式中，$\langle x, w\rangle_t$——为杜勒-迈耶分解式中加入的鞅 x 和 w 乘积的随机过程：

$$\mathbf{M}[x_tw_t-x_sw_s|\mathscr{F}_s]=\mathbf{M}[\langle x, w\rangle_t-\langle x, w\rangle_s|\mathscr{F}_s]。 \qquad (30)$$

将表达式 (28) 叫做(最佳非线性)滤波的基本方程。(在假定 (26)、(27) 范围内)大部分已知结果都可以由此方程导出。

例如，我们来证明，由 (28) 怎样导出卡尔曼-布西概型中的滤波方程 (17)、(18)，为简单起见，认为 $b(t)\equiv B(t)\equiv 1$ 。

将 (12) 与 (26) 及 (27) 相比较可看出，$A_t(\omega)=A(t)\theta_t$，$w_t=w_2(t)$。令 $h_t=\theta_t$，于是，根据 (12)

$$h_t=h_0+\int_0^t a(s)\theta_s ds+w_1(t)。 \qquad (31)$$

过程 $w_1=(w_1(t))$ 和 $w_2=(w_2(t))$，$t\geqslant 0$，为独立的平方可积鞅，因此，对于这两个过程而言，$D_t\equiv 0$($\mathbf{P}$-п.н.)。于是，根据(28)，$\pi_t(\theta)$ 有微分

$$d\pi_t(\theta)=a(t)\pi_t(\theta)dt+A(t)[\pi_t(\theta^2)-\pi_t^2(\theta)]d\bar{w}_t, \qquad (32)$$

即

$$dm_t=a(t)m_tdt+A(t)\gamma_td\bar{w}_t, \qquad (33)$$

其中我们使用了，根据过程 (θ, ξ) 的高斯性，$\mathbf{P}$-п.н.

$$\pi_t(\theta^2) - \pi_t^2(\theta) = \mathbf{M}[(\theta_t - m_t)^2 | \mathscr{F}_t^\xi]$$
$$= \mathbf{M}[\theta_t - m_t]^2 = \gamma_t。$$

为了由 (28) 导出 γ_t 的方程，我们取 $h_t = \theta_t^2$。于是，由系 (12) 的第一个方程按伊藤变量代换公式(定理 4.4)就可得到

$$\theta_t^2 = \theta_0^2 + \int_0^t a_s(\omega)ds + x_t, \tag{34}$$

式中

$$a_s(\omega) = 2a(s)\theta_s^2 + b^2(s) \text{ 及 } x_t = \int_0^t 2b(s)\theta_s dw_1(s)。$$

因此，根据 (28)

$$d\pi_t(\theta^2) = [2a(t)\pi_t(\theta^2) + b^2(t)]dt + A(t)[\pi_t(\theta^3) - \pi_t(\theta)\pi_t(\theta^2)]d\bar{w}_t。\tag{35}$$

由 (32) 和 (35) 可见，在使用基本滤波方程 (28) 时遇到了一点困难，即为了求得低阶条件矩，则要求知道高阶矩。因而，在寻求 $\pi_t(\theta^2)$的方程时，则要求知道三阶验后矩 $\pi_t(\theta^3)=\mathbf{M}(\theta_t^3|\mathscr{F}_t^\xi)$。在所研究的情况下，这一困难很容易解决，因为根据过程 (θ, ξ)的高斯性，对所有 $n \geqslant 3$ 可言，矩 $\pi_t(\theta^n) = \mathbf{M}(\theta_t^n|\mathscr{F}_t^\xi)$ 是用 $\pi_t(\theta)$ 和 $\pi_t(\theta^2)$ 来表示的。特别是，$\pi_t(\theta^3) - \pi_t(\theta)\pi_t(\theta^2) = \mathbf{M}[\theta_t^2(\theta_t - m_t)|\mathscr{F}_t^\xi] = 2m_t\gamma_t$，那么

$$d\pi_t(\theta^2) = [2a(t)\pi_t(\theta^2) + b^2(t)]dt + 2A(t)m_t\gamma_t d\bar{w}_t。\tag{36}$$

根据伊藤变量代换公式，由 (33) 即求得

$$dm_t^2 = 2m_t[a(t)m_t dt + A(t)\gamma_t m_t d\bar{w}_t] + A^2(t)\gamma_t^2 dt。$$

同方程(36)一起，这一关系对于 $\gamma_t = \pi_t(\theta^2) - m_t^2$ 而言，给出的是所求的方程 (18)。

方程(17)、(18)的上述推导在某种意义下是有可借鉴的地方，即由上述结果可见，为了获得确定最佳滤波的闭合方程组，应当引出有关高阶条件矩之间关系的补充知识。

本书中，特别注意到所谓条件高斯过程 (θ, ξ)，对该过程而