arXiv:2601.01805v1 [math.ST] 5 Jan 2026

连续时间线性高斯模型中平滑分布的路径表示

Masahiro Kurisaki*†

2026年1月6日

摘要

我们研究连续时间线性高斯系统中的滤波与平滑问题。经典方法如 Kalman–Bucy 滤波器和 Rauch–Tung–Striebel（RTS）平滑器，为条件均值和协方差提供了递推公式；而我们提出了一种路径视角，将平滑误差动力学刻画为 Ornstein–Uhlenbeck过程。作为应用，我们证明，标准的滤波方程和平滑方程均可统一地作为主定理的推论导出。特别是，我们首次在连续时间情形下给出了 Bryson–Frazier 平滑器的数学严格推导。除有助于更透明地理解平滑分布外，我们的公式还能够对该分布进行路径采样，从而便于使用蒙特卡罗方法评估非线性泛函。

关键词— 线性滤波与平滑、卡尔曼-布西滤波器、Rauch–Tung–Striebel 平滑器、非线性滤波、渐近展开

1 引言

根据含噪观测估计随机动力系统的隐状态，是信号处理、控制理论和数理金融等诸多领域中的一个基本问题。本文考虑如下连续时间线性高斯模型：

$$dX_t = a(t)X_t dt + b(t)dV_t, \tag{1.1}$$
$$dY_t = c(t)X_t dt + \sigma(t)dW_t, \tag{1.2}$$

*东京大学大学院数理科学研究科，电子邮箱：makurisaki@g.ecc.u-tokyo.ac.jp

†日本科学技术振兴机构CREST 作者获得了日本科学技术振兴机构CREST JPMJCR2115和JSPS KAKENHI项目编号JP24KJ0667的资助。

其中， V 和 W 是相互独立的布朗运动， a、 b、 c和 σ 是已知的确定性函数。过程 X 不可观测，目标是根据观测到的过程 Y推断其行为。

估计条件期望的问题 $E[X_t \mid \mathcal{Y}_t]$，其中$\mathcal{Y}_t = \sigma(Y_s : 0 \le s \le t)$ 是由截至时间 σ的观测所生成的 t-域，被称为滤波问题。在线性高斯情形下，该问题可由经典的 Kalman–Bucy 滤波器解决；相关推导在离散时间模型（Kalman 1960年、S¨arkk¨a 2013年）和连续时间情形（Kalman 和布西 1961年、Kailath 1968年、Bain 和 Crisan 2009、Liptser 和 Shiryaev 2001a）中均有详尽论述。

另一个相关问题是估计 $E[X_s \mid \mathcal{Y}_t]$ 在 $0 \le s \le t$下的值，该问题被称为平滑问题。其解由 Rauch–Tung–Striebel（RTS）平滑器给出，离散时间（Rauch 等 1965年、S¨arkk¨a 2013年）和连续时间（Liptser 和 Shiryaev 2001b、Ander- son 等 2021、Rutkowski 1993年）中均有相关推导。传统的平滑方法侧重于推导条件均值和条件协方差的递推公式。

本文从不同角度探讨平滑问题，对给定 $\mathcal{Y}_t$ 时隐过程 $\{X_s\}_{0 \le s \le t}$ 的条件分布进行路径刻画。我们不只关注条件均值和协方差，还明确描述平滑误差过程的动态；结果表明，该过程满足一个线性随机微分方程。具体而言，我们证明，平滑误差演化为一个奥恩斯坦–乌伦贝克型过程，由在条件测度下独立的布朗运动驱动。

这种观点——将平滑误差视为随机过程——此前已见于工程学文献，尤以 Bello 等人（1986）的工作最为著名。然而，该研究缺乏数学严谨性，依赖缺乏坚实概率论基础的形式推理，并要求条件协方差满足非奇异性假设。相比之下，我们的推导完全严谨，基于概率学家熟悉的随机分析工具，例如吉尔萨诺夫定理和伊藤公式，并且不依赖任何额外的正则性假设。

我们的表述具有若干关键优势。首先，它通过揭示背后的随机机制，对平滑分布提供了更清晰、统一的理解。因此，Kalman–Bucy 滤波器和 RTS 平滑器等经典结果，都可以统一地作为我们主定理的推论导出。

特别是，我们的方法首次对连续时间设定下的 Bryson–Frazier（BF）平滑器进行了数学上严谨的推导。BF 平滑器最初是在离散时间背景下提出的（Bryson 和 Frazier 1963年，Bierman 1973），为 RTS 平滑器提供了另一种表述。BF表述的一个关键优势在于，它避免了计算条件协方差

逆矩阵，而 RTS 方法要求进行这项计算。尽管工程学文献中已经对 BF 平滑器的连续时间版本进行了形式化推导（Weinert 和 Desai（1981））， 但直到现在，这一推导仍缺乏严谨的概率论基础。

值得强调的是，尽管 BF 平滑器在数学上等价于 RTS 平滑器，但试图从 Kalman–Bucy 滤波器或 RTS 平滑器推导出 BF 平滑器，不可避免地会得到涉及条件协方差逆矩阵的表达式（见备注）。因此，必须施加额外的正则性假设，以确保定义良好。我们的方法通过直接构造 BF 平滑器，而不依赖那些已有的表述，避开了这一困难。

我们这种表述在实践中的另一项优势是，它能够通过显式的随机表示，从平滑分布中进行逐路径采样。这有助于对非线性泛函（例如 $E[\sup_{0\le s\le t} X_s \mid \mathcal{Y}_t]$）的期望，以及非线性模型 EM 算法中的 Q-函数进行蒙特卡洛近似；在此过程中，还可结合扩展卡尔曼滤波器等线性化技术。

本文其余部分安排如下。第 2 节通过一个简单的一维示例阐释主要思想。第 3 节给出线性高斯模型中平滑动力学的一般公式。第 4 节将经典滤波方程和平滑方程推导为我们结果的推论，并讨论其意义。最后，第 5 节概述该定理的进一步应用。

2 一维情形

在讨论一般问题之前，我们先通过以下一维模型来阐释主要思想：

$$Y = cX + \sigma Z,$$

其中 $X \sim N(\mu_x, \sigma_x^2)$ 和 $Z \sim N(0,1)$ 是相互独立的随机变量，且 $\sigma_x, \sigma > 0$。此外，我们假设 $c \neq 0$。在此模型中，给定 Y 时 X 的条件期望可表示为

$$\begin{aligned} E[X|Y=y] &= \frac{\displaystyle\int_{\mathbb{R}} x \exp\left(-\frac{1}{2\sigma^2}(cx-y)^2\right) p(x)\,dx}{\displaystyle\int_{\mathbb{R}} \exp\left(-\frac{1}{2\sigma^2}(cx-y)^2\right) p(x)\,dx} \\ &= \frac{E\left[X \exp\left(-\frac{c^2X^2}{2\sigma^2} + \frac{cXy}{\sigma^2}\right)\right]}{E\left[\exp\left(-\frac{c^2X^2}{2\sigma^2} + \frac{cXy}{\sigma^2}\right)\right]} \end{aligned} \tag{2.1}$$

其中 $p(x)$ 是 X 的概率密度函数。此公式类似于（3.6）。

这个最终表达式还可以进一步改写为 $\frac{\sigma^2}{c}\frac{d}{dy}\log E\left[\exp\left(-\frac{c^2X^2}{2\sigma^2}+\frac{cXy}{\sigma^2}\right)\right]$。由于 X 具有正态性，可以证明 log 部分是关于 y 的二次函数。因此，我们可以写成

$$E[X|Y=y]=\frac{E\left[X\exp\left(-\frac{c^2X^2}{2\sigma^2}+\frac{cXy}{\sigma^2}\right)\right]}{E\left[\exp\left(-\frac{c^2X^2}{2\sigma^2}+\frac{cXy}{\sigma^2}\right)\right]}=Ay+B.$$

对两边取期望（相对于 y），得到 $B=E[X]-AcE[X]$，因此

$$E[X|Y=y]=\frac{E\left[X\exp\left(-\frac{c^2X^2}{2\sigma^2}+\frac{cXy}{\sigma^2}\right)\right]}{E\left[\exp\left(-\frac{c^2X^2}{2\sigma^2}+\frac{cXy}{\sigma^2}\right)\right]}=E[X]+A(y-cE[X]). \tag{2.2}$$

此外，对 2.2 求导，A可以确定为

$$\begin{aligned}A=&\frac{d}{dy}\frac{E\left[X\exp\left(-\frac{c^2X^2}{2\sigma^2}+\frac{cXy}{\sigma^2}\right)\right]}{E\left[\exp\left(-\frac{c^2X^2}{2\sigma^2}+\frac{cXy}{\sigma^2}\right)\right]}\\=&\frac{c}{\sigma^2}\frac{E\left[X^2\exp\left(-\frac{c^2X^2}{2\sigma^2}+\frac{cXy}{\sigma^2}\right)\right]}{E\left[\exp\left(-\frac{c^2X^2}{2\sigma^2}+\frac{cXy}{\sigma^2}\right)\right]}-\frac{c}{\sigma^2}\left(\frac{E\left[X\exp\left(-\frac{c^2X^2}{2\sigma^2}+\frac{cXy}{\sigma^2}\right)\right]}{E\left[\exp\left(-\frac{c^2X^2}{2\sigma^2}+\frac{cXy}{\sigma^2}\right)\right]}\right)^2\end{aligned} \tag{2.3}$$

由于 A 不依赖于 y，我们可以在最终表达式中令 $y=0$，从而得到

$$A=\frac{c}{\sigma^2}\left\{\frac{E\left[X^2\exp\left(-\frac{c^2X^2}{2\sigma^2}\right)\right]}{E\left[\exp\left(-\frac{c^2X^2}{2\sigma^2}\right)\right]}-\left(\frac{E\left[X\exp\left(-\frac{c^2X^2}{2\sigma^2}\right)\right]}{E\left[\exp\left(-\frac{c^2X^2}{2\sigma^2}\right)\right]}\right)^2\right\}. \tag{2.4}$$

因此，（2.2）最终出现在条件均值公式中

$$E[X|Y=y] = E[X] + \frac{c}{\sigma^2}\left\{\frac{E\left[X^2\exp\left(-\frac{c^2X^2}{2\sigma^2}\right)\right]}{E\left[\exp\left(-\frac{c^2X^2}{2\sigma^2}\right)\right]} - \left(\frac{E\left[X\exp\left(-\frac{c^2X^2}{2\sigma^2}\right)\right]}{E\left[\exp\left(-\frac{c^2X^2}{2\sigma^2}\right)\right]}\right)^2\right\}(y - cE[X]). \tag{2.5}$$

此外，回顾贝叶斯公式（2.1），我们注意到（2.3）等价于

$$A = \frac{c}{\sigma^2}(E[X^2|Y=y] - E[X|Y=y]^2) = \frac{c}{\sigma^2}E[V|Y=y].$$

由于 A 由（2.4）给出，我们得到条件方差公式

$$V[X|Y=y] = \frac{E\left[X^2\exp\left(-\frac{c^2X^2}{2\sigma^2}\right)\right]}{E\left[\exp\left(-\frac{c^2X^2}{2\sigma^2}\right)\right]} - \left(\frac{E\left[X\exp\left(-\frac{c^2X^2}{2\sigma^2}\right)\right]}{E\left[\exp\left(-\frac{c^2X^2}{2\sigma^2}\right)\right]}\right)^2. \tag{2.6}$$

进一步对（2.3）求导，可得 X 在给定Y条件下的第三阶及更高阶累积量，但它们均为零，因为 A 是一个常数。由此可知， X 在给定 $Y=y$ 条件下的条件分布为正态分布，其均值和方差分别由（2.5）和（2.6）给出。

这里的关键贡献在于，我们通过对条件期望求导、在零点处取值，推导出（2.5）和（2.6）中的非平凡表达式，它们将条件分布表示为普通期望之商。在无限维情形下，这种表述使我们能够运用随机微积分中的经典工具（如伊藤公式和吉尔сан诺夫定理）来分析条件分布。

3 条件过程的动力学

现在，我们继续讨论原始的连续时间情形。设 $(\Omega\ \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, P)$ 为一个滤过概率空间，并假设过程 $\{X_t\}$ 和 $\{Y_t\}$ 是以下线性方程的解：

$$dX_t = a(t)X_t\,dt + b(t)\,dV_t, \tag{3.1}$$
$$dY_t = c(t)X_t\,dt + \sigma(t)\,dW_t, \quad Y_0 = 0 \tag{3.2}$$

其中， a、 b、 c和 σ 是分别取值于 $M_{d_1}(\mathbb{R})$、$M_{d_1,\ m_1}(\mathbb{R})$、$M_{d_2,\ d_1}(\mathbb{R})$和$M_{d_2,\ m_2}(\mathbb{R})$的可测函数。这里， $\{V_t\}$和 $\{W_t\}$

是相互独立的 m_1维和 m_2维 $\{\mathcal{F}_t\}$-布朗运动。我们还假设 X_0 服从正态分布，且与 $\{V_t\}$ 和 $\{W_t\}$相互独立，因此 $\{(X_t,\ Y_t)\}$ 构成一个 d_1+d_2维高斯过程。此外，我们假设对于每个 $t\geq 0$,

$$\begin{aligned}&\int_0^t |a(s)|\,ds<\infty, \quad \int_0^t |b(s)|^2\,ds<\infty,\\ &\int_0^t |c(s)|^2\,ds<\infty, \quad \int_0^t |\sigma(s)|^2\,ds<\infty,\end{aligned} \tag{3.3}$$

并且存在常数 $C>0$，使得对于每个 $t\geq 0$,

$$\lambda_{\min}(\sigma(t)\sigma(t)^\top)>C, \tag{3.4}$$

其中，$\lambda_{\min}(\sigma(t)\sigma(t)^\top)$ 表示 $\sigma(t)\sigma(t)^\top$ 的最小特征值。在本文中，$|A|$ 表示矩阵 A 的弗罗贝尼乌斯范数。

我们首先给出贝叶斯公式的连续时间类似物（2.1），在滤波领域中称为卡利安普尔–斯特里贝尔公式。

命题 3.1。设 $T>0$，并通过 $Q(A)=E\left[1_A Z_T^{-1}\right]$ 定义概率测度 Q，其中对于 $A\in\mathcal{F}$,

$$\begin{aligned}Z_t=&\exp\left(\int_0^t X_s^\top c(s)^\top(\sigma(s)\sigma(s)^\top)^{-1}\sigma(s)dW_s\right.\\ &\left.+\frac{1}{2}\int_0^t X_s^\top c(s)^\top(\sigma(s)\sigma(s)^\top)^{-1}c(s)X_s ds\right)\\ =&\exp\left(\int_0^t X_s^\top c(s)^\top(\sigma(s)\sigma(s)^\top)^{-1}dY_s\right.\\ &\left.-\frac{1}{2}\int_0^t X_s^\top c(s)^\top(\sigma(s)\sigma(s)^\top)^{-1}c(s)X_s ds\right).\end{aligned} \tag{3.5}$$

那么，在新的测度 Q 下，过程

$$\overline{Y}_t=\int_0^t(\sigma(t)\sigma(t)^\top)^{\frac{1}{2}}dY_s \quad (0\leq t\leq T)$$

是一个与 $\{X_t\}_{0\leq t\leq T}$ 独立的布朗运动，并且 $\{X_t\}_{0\leq t\leq T}$ 的分布与其在 P 下的分布相同。此外，对于 $0\leq t\leq T$ 以及满足 $E[|U|]<\infty$ 的随机变量 U，有 $E[Z_t|\mathcal{Y}_t]>0$ 几乎处处。并且

$$E[U|\mathcal{Y}_t]=\frac{E_Q\left[UZ_t|\mathcal{Y}_t\right]}{E_Q\left[Z_t|\mathcal{Y}_t\right]}. \tag{3.6}$$

证明。参见 Bain 和 Crisan(2009) 中的习题 3.11、3.14 和命题 3.16。

□

接下来，我们旨在确定 Z_t 表达式中的 Y。为此，将 Z_t 重写为

$$\begin{aligned} Z_t =& \exp\left(\int_0^t X_s^\top c(s)^\top(\sigma(s)\sigma(s)^\top)^{-1}dY_s \right. \\ & \left. -\frac{1}{2}\int_0^t X_s^\top c(s)^\top(\sigma(s)\sigma(s)^\top)^{-1}c(s)X_s ds\right) \\ =& \exp\left(X_t^\top \hat{Y}_t - \int_0^t \hat{Y}_s^\top dX_s - \frac{1}{2}\int_0^t X_s^\top c(s)^\top(\sigma(s)\sigma(s)^\top)^{-1}c(s)X_s ds\right), \end{aligned} \tag{3.7}$$

其中 $\hat{Y}_s = \int_0^s c(u)^\top(\sigma(u)\sigma(u)^\top)^{-1}dY_u$，并记

$$\begin{aligned} z_t(y) =& \exp\left(X_t^\top y(t) - \int_0^t y(s)^\top dX_s \right. \\ & \left. -\frac{1}{2}\int_0^t X_s^\top c(s)^\top(\sigma(s)\sigma(s)^\top)^{-1}c(s)X_s ds\right) \end{aligned} \tag{3.8}$$

对于有界可测函数 $y: [0,t] \to \mathbb{R}^{d_1}$。

于是，由于 X 和 Y 相互独立，且 X 在 P 和 Q下具有相同的分布（命题3.1），自然可以预期

$$E_Q[Z_t \mid \mathcal{Y}_t] = E[z_t(y)]\big|_{y=\hat{Y}}. \tag{3.9}$$

这个恒等式确实成立，不过论证需要谨慎，因为 Z_t 中的随机积分不是逐路径定义的。引理A.1给出了严格的证明。

接下来，我们证明（3.9）是 Y的“二次型”。

引理 3.2。对于 $t \in [0,T]$ 以及有界可测函数 $y: [0,t] \to \mathbb{R}^{d_1}$，可以写成

$$\begin{aligned} \log E_Q[z_t(y)] =& h_0(t) + h_1(t)y(t) + y(t)^\top h_2(t)y(t) + \int_0^t y(s)^\top k_1(s;t)ds \\ & + \int_0^t y(s)^\top k_2(s;t)y(s)ds + \int_0^t y(s)^\top k_3(s;t)y(t)ds \end{aligned}$$

其中 $\int_0^t |k_i(s;t)|ds < \infty\,(i=1,2,3)$。

这个引理的证明只需直接计算期望。详情请参见第 B 节。现在我们给出一个公式，将条件期望化为普通期望；该公式对应于一维情形中的方程（2.5）和（2.6）。

定理 3.3。 设 $t \geq 0$、 s、 $u \in [0,t]$，并记

$$\begin{aligned}
&\gamma(s,u;t) \\
&= \frac{E\left[X_s \otimes X_u \exp\left(-\frac{1}{2}\int_0^t X_s^\top c(s)^\top (\sigma(s)\sigma(s)^\top)^{-1} c(s) X_s ds\right)\right]}{E\left[\exp\left(-\frac{1}{2}\int_0^t X_s^\top c(s)^\top (\sigma(s)\sigma(s)^\top)^{-1} c(s) X_s ds\right)\right]} \\
&- \frac{E\left[X_s \exp\left(-\frac{1}{2}\int_0^t X_s^\top c(s)^\top (\sigma(s)\sigma(s)^\top)^{-1} c(s) X_s ds\right)\right]}{E\left[\exp\left(-\frac{1}{2}\int_0^t X_s^\top c(s)^\top (\sigma(s)\sigma(s)^\top)^{-1} c(s) X_s ds\right)\right]} \\
&\otimes \frac{E\left[X_u \exp\left(-\frac{1}{2}\int_0^t X_s^\top c(s)^\top (\sigma(s)\sigma(s)^\top)^{-1} c(s) X_s ds\right)\right]}{E\left[\exp\left(-\frac{1}{2}\int_0^t X_s^\top c(s)^\top (\sigma(s)\sigma(s)^\top)^{-1} c(s) X_s ds\right)\right]}.
\end{aligned} \tag{3.10}$$

那么， $\{X_s\}_{0\leq s\leq t}$ 在给定 $\mathcal{Y}_t$ 条件下的条件分布是一个均值为

$$E[X_s|\mathcal{Y}_t] = E[X_s] + \int_0^t \gamma(s,u;t) c(u)^\top (\sigma(u)\sigma(u)^\top)^{-1} (dY_u - c(u)E[X_u]\,du), \tag{3.11}$$

、协方差为

$$\mathrm{Cov}(X_s, X_u|\mathcal{Y}_t) = \gamma(s,u;t). \tag{3.12}$$

我们注意到，方程（3.11）和（3.12）与其一维对应项，即方程（2.5）和（2.6），十分相似。这也阐明了卡尔曼-布西滤波器与卡梅隆-马丁公式之间的联系；该公式在 Liptser 和 Shiryaev （2001a）的第7节中讨论过，两者都涉及黎卡提方程。

这些公式可以通过对（3.6）中的商进行求导直观地推导出来，其中 $U = X_s$，关于“dY_u”求导（或通过 Malliavin 导数），然后在 $Y = 0$处计算结果。然而，要以当前形式严格证明这些操作并不简单。

为克服这一数学困难，我们通过以下方式引入 Z_t 和路径 $\{X_s\}_{0\leq s\leq t}$ 的离散化版本：

$$\begin{aligned}
Z_t^{(n)} = \exp\Bigg(&X_t^\top \hat{Y}_t - \sum_{i=1}^n \hat{Y}_{s_{i-1}}^\top (X_{s_i} - X_{s_{i-1}}) \\
&-\frac{1}{2}\int_0^t X_s^\top c(s)^\top (\sigma(s)\sigma(s)^\top)^{-1} c(s) X_s ds\Bigg),
\end{aligned} \tag{3.13}$$

and

$$X_s^{(n)} = \sum_{i=1}^{n} X_{s_{i-1}} 1_{[s_{i-1},s_i)}(s) + X_t 1_{\{t\}}(s) = X_{\frac{t}{n}[\frac{s}{t}n]} \tag{3.14}$$

其中 $s_i = s_i^n = (i/n)t$。随后可得以下收敛性结论。

引理 3.4。对于任意固定的 u_1、$\cdots$、$u_k \in [0,t]$，有

$$\frac{E_Q[X_{u_1}^{(n)} \otimes \cdots \otimes X_{u_k}^{(n)} Z_t^{(n)} | \mathcal{Y}_t]}{E_Q[Z_t^{(n)} | \mathcal{Y}_t]} \xrightarrow{L^2} \frac{E_Q[X_{u_1} \otimes \cdots \otimes X_{u_k} Z_t | \mathcal{Y}_t]}{E_Q[Z_t | \mathcal{Y}_t]} \quad (n \to \infty). \tag{3.15}$$

证明。若定义 $\hat{Y}^{(n)}(s) = \sum_{i=1}^{n} \hat{Y}_{s_{i-1}} 1_{[s_{i-1},s_i)}(s) + \hat{Y}_t 1_{\{t\}}(s)$，则有 $Z_t^{(n)} = z_t(\hat{Y}^{(n)})$，因此

$$E_Q[Z_t^{(n)} | \mathcal{Y}_t] = E_Q[z_t(y)]\big|_{y=\hat{Y}^{(n)}}$$

and

$$E_Q[X_{u_1}^{(n)} \otimes \cdots \otimes X_{u_k}^{(n)} Z_t^{(n)} | \mathcal{Y}_t] = E_Q[X_{u_1}^{(n)} \otimes \cdots \otimes X_{u_k}^{(n)} z_t(y)]\big|_{y=\hat{Y}^{(n)}}.$$

因此，由于当 $p \geq 1$ 时 $\sup_{0\leq s\leq t} E_Q\left[\left|\hat{Y}_s - \hat{Y}_{(}{}_s^n\right|^p\right] \to 0$ 收敛，且我们可以写成

$$\frac{E_Q[X_{u_1} \otimes \cdots \otimes X_{u_n} z_t(y)]}{E_Q[z_t(y)]} = \bigotimes_{j=1}^{n} \left\{ p_j(u_j;t) + q_j(u_j;t)y(t) + \int_0^t r_j(u_j,s;t)y(s)ds \right\} \tag{3.16}$$

其中 $\int_0^t |r_j(u_j, s; t)| < \infty$（见第 B 节）。

□

定理 *3.3* 的证明。首先，写作

$$z_t^{(n)}(y_0, \cdots, y_n) = \exp\left(X_t^\top y_n - \sum_{i=1}^{n} y_{i-1}^\top (X_{s_i} - X_{s_{i-1}}) \right.$$
$$\left. -\frac{1}{2}\int_0^t X_s^\top c(s)^\top (\sigma(s)\sigma(s)^\top)^{-1} c(s) X_s ds \right),$$

并考虑函数 $F(y_0, \cdots, y_n) = E_Q\left[z_t^{(n)}(y_0, \cdots, y_n)\right]$，其中 $y_0, \cdots, y_n \in \mathbb{R}^{d_1}$。那么很容易验证

$$\partial_{y_i} F(y_0, \cdots, y_n) = -E_Q\left[(X_{s_{i+1}} - X_{s_i}) z_t^{(n)}(y_0, \cdots, y_n)\right] \quad (i = 0, 1, \cdots, n),$$

其中 $X_{s_{n+1}}=0$，为方便起见，更高阶导数也可以用同样的方法计算。因此，我们有

$$\begin{aligned}
&\sum_{i,j=0}^{n}\partial_{y_i}\otimes\partial_{y_j}\log F(y_0,\cdots,y_n)1_{[0,s_{i+1})}(s)1_{[0,s_{j+1})}(u)\\
=&\sum_{i,j=0}^{n}\frac{E_Q[(X_{s_{i+1}}-X_{s_i})\otimes(X_{s_{j+1}}-X_{s_j})z_t^{(n)}(y_0,\cdots,y_n)]}{E_Q[z_t^{(n)}(y_0,\cdots,y_n)]}1_{[0,s_{i+1})}(s)1_{[0,s_{j+1})}(u)\\
&-\sum_{i,j=0}^{n}\frac{E_Q[(X_{s_{i+1}}-X_{s_i})z_t^{(n)}(y_0,\cdots,y_n)]}{E_Q[z_t^{(n)}(y_0,\cdots,y_n)]}1_{[0,s_{i+1})}(s)\\
&\otimes\frac{E_Q[(X_{s_{j+1}}-X_{s_j})z_t^{(n)}(y_0,\cdots,y_n)]}{E_Q[z_t^{(n)}(y_0,\cdots,y_n)]}1_{[0,s_{j+1})}(u)\\
=&\frac{E_Q[X_s^{(n)}\otimes X_u^{(n)}z_t^{(n)}(y_0,\cdots,y_n)]}{E_Q[z_t^{(n)}(y_0,\cdots,y_n)]}\\
&-\frac{E_Q[X_s^{(n)}z_t^{(n)}(y_0,\cdots,y_n)]\otimes E_Q[X_u^{(n)}z_t^{(n)}(y_0,\cdots,y_n)]}{E_Q[z_t^{(n)}(y_0,\cdots,y_n)]^2},
\end{aligned}\tag{3.17}$$

其中 $X^{(n)}$ 定义见（3.14）。

另一方面，由引理 3.2 可知，$\log F(y_1,\cdots,y_n)$ 是关于 $y_0,\cdots,y_n$ 的二次型，因此 $\partial_{y_i}\otimes\partial_{y_j}\log F(y_0,\cdots,y_n)$ 不依赖于 $y_0,\cdots,y_n$。因此，（3.17）不依赖于 $y_0,\cdots,y_n$，于是我们有

$$\begin{aligned}
&\frac{E_Q[X_s^{(n)}\otimes X_u^{(n)}z_t^{(n)}(y_0,\cdots,y_n)]}{E_Q[z_t^{(n)}(y_0,\cdots,y_n)]}\\
&-\frac{E_Q[X_s^{(n)}z_t^{(n)}(y_0,\cdots,y_n)]\otimes E_Q[X_u^{(n)}z_t^{(n)}(y_0,\cdots,y_n)]}{E_Q[z_t^{(n)}(y_0,\cdots,y_n)]^2}\\
=&\frac{E_Q[X_s^{(n)}\otimes X_u^{(n)}z_t^{(n)}(0,\cdots,0)]}{E_Q[z_t^{(n)}(0,\cdots,0)]}\\
&-\frac{E_Q[X_s^{(n)}z_t^{(n)}(0,\cdots,0)]\otimes E_Q[X_u^{(n)}z_t^{(n)}(0,\cdots,0)]}{E_Q[z_t^{(n)}(0,\cdots,0)]^2}.
\end{aligned}$$

^^ 将 $(Y_{s_0},\cdots,Y_{s_n})$ 代入 $(y_0,\cdots,y_n)$，得到

$$\begin{aligned}
&\frac{E_Q[X_s^{(n)}\otimes X_u^{(n)}Z_t^{(n)}|\mathcal{Y}_t]}{E_Q[Z_t^{(n)}|\mathcal{Y}_t]}-\frac{E_Q[X_s^{(n)}Z_t^{(n)}|\mathcal{Y}_t]\otimes E_Q[X_u^{(n)}Z_t^{(n)}|\mathcal{Y}_t]}{E_Q[Z_t^{(n)}|\mathcal{Y}_t]^2}\\
=&\frac{E\left[X_s^{(n)}\otimes X_u^{(n)}\exp\left(-\frac{1}{2}\int_0^t X_s^\top c(s)^\top(\sigma(s)\sigma(s)^\top)^{-1}c(s)X_s ds\right)\right]}{E\left[\exp\left(-\frac{1}{2}\int_0^t X_s^\top c(s)^\top(\sigma(s)\sigma(s)^\top)^{-1}c(s)X_s ds\right)\right]}
\end{aligned}$$

$$
- \frac{E\left[X_s^{(n)} \exp\left(-\frac{1}{2}\int_0^t X_s^\top c(s)^\top(\sigma(s)\sigma(s)^\top)^{-1}c(s)X_s ds\right)\right]}{E\left[\exp\left(-\frac{1}{2}\int_0^t X_s^\top c(s)^\top(\sigma(s)\sigma(s)^\top)^{-1}c(s)X_s ds\right)\right]}
$$

$$
\otimes \frac{E\left[X_u^{(n)} \exp\left(-\frac{1}{2}\int_0^t X_s^\top c(s)^\top(\sigma(s)\sigma(s)^\top)^{-1}c(s)X_s ds\right)\right]}{E\left[\exp\left(-\frac{1}{2}\int_0^t X_s^\top c(s)^\top(\sigma(s)\sigma(s)^\top)^{-1}c(s)X_s ds\right)\right]} \tag{3.18}
$$

因为 $z_t^{(n)}(\hat{Y}_{s_0}, \cdots, \hat{Y}_{s_n}) = Z_t^{(n)}$，如（3.13）中所引入。这里，我们利用了 $\{X_t\}_{0\le t\le T}$ 在 E 和 E_Q下的分布相同。根据引理 3.4和命题 3.1，该式左侧收敛于 L^2 到

$$
\frac{E_Q[X_s \otimes X_u Z_t|\mathcal{Y}_t]}{E_Q[Z_t|\mathcal{Y}_t]} - \frac{E_Q[X_s Z_t|\mathcal{Y}_t] \otimes E_Q[X_u Z_t|\mathcal{Y}_t]}{E_Q[Z_t|\mathcal{Y}_t]^2} = \mathrm{Cov}(X_s, X_u|\mathcal{Y}_t)
$$

当 $n \to \infty$。另一方面， L^2-右侧的收敛直接由 L^2-收敛 $X_s^{(n)}$得出。因此，我们得到（3.12）。

同理，由于 $\log F(y_0, \cdots, y_n)$ 的三阶及更高阶导数均为零，我们有

$$
0 = \sum_{i_1,\cdots,i_k=0}^{n} \partial_{y_{i_1}} \otimes \cdots \otimes \partial_{y_{i_k}} \log F(\hat{Y}_{s_0}, \cdots, \hat{Y}_{s_n}) 1_{[0,s_{i_1}]}(u_1) \cdots 1_{[0,s_{i_k}]}(u_k)
$$

对于 $k \ge 3$。右侧依概率收敛到 $X_{u_1}, \cdots, X_{u_k}$ 的第 k 阶累积量，因此它几乎必然为零，这意味着$\{X_s\}_{0\le s\le t}$ 是条件高斯的。

为证明（3.11），我们考虑

$$
\sum_{i=0}^{n} \partial_{y_i} \log F(\hat{Y}_{s_0}, \cdots, \hat{Y}_{s_n}) 1_{[0,s_{i+1})}(s) = \frac{E_Q[X_s^{(n)} Z_t^{(n)}|\mathcal{Y}_t]}{E_Q[Z_t^{(n)}|\mathcal{Y}_t]}, \tag{3.19}
$$

并令 D 为关于布朗运动 $Y_s = \int_0^s (\sigma(s)\sigma(s)^\top)^{-\frac{1}{2}} dY_s$ 在 Q 上的 Malliavin 导数。由于 $\log F(y_0, \cdots, y_n)$ 是一个二次型， $\partial_{y_i} \log F(y_0, \cdots, y_n)$ 是线性的。因此，根据光滑随机变量的 Malliavin 导数定义（Nualart，2006年）以及 $\hat{Y}_t = \int_0^t c(s)^\top(\sigma(s)\sigma(s)^\top)^{-\frac{1}{2}} dY_s$，(3.19) 的 Malliavin 导数为

$$
\begin{aligned}
&D_u \frac{E_Q[X_s^{(n)} Z_t^{(n)}|\mathcal{Y}_t]}{E_Q[Z_t^{(n)}|\mathcal{Y}_t]} \\
=&\sum_{i=0}^{n} (\sigma(u)\sigma(u)^\top)^{-\frac{1}{2}} c(u) \partial_{y_j}\partial_{y_i} \log F(\hat{Y}_{s_0}, \cdots, \hat{Y}_{s_n}) 1_{[0,s_i]}(s) 1_{[0,s_j]}(u) \\
=&(\sigma(u)\sigma(u)^\top)^{-\frac{1}{2}} c(u) \left\{ \frac{E[X_u^{(n)} \otimes X_s^{(n)} Z_t^{(n)}|\mathcal{Y}_t]}{E[Z_t^{(n)}|\mathcal{Y}_t]} - \frac{E[X_u^{(n)} Z_t^{(n)}|\mathcal{Y}_t] \otimes E[X_s^{(n)} Z_t^{(n)}|\mathcal{Y}_t]}{E[Z_t^{(n)}|\mathcal{Y}_t]^2} \right\}.
\end{aligned}
$$

根据（3.18），这是确定性的。因此，根据克拉克-奥康公式（Nualart, 2006年），我们可以写成

$$\frac{E_Q[X_s^{(n)}Z_t^{(n)}|\mathcal{Y}_t]}{E_Q[Z_t^{(n)}|\mathcal{Y}_t]} = g^{(n)}(s;t) + \int_0^t \left(\frac{E[X_u^{(n)} \otimes X_s^{(n)}Z_t^{(n)}|\mathcal{Y}_t]}{E[Z_t^{(n)}|\mathcal{Y}_t]} - \frac{E_Q[X_u^{(n)}Z_t^{(n)}|\mathcal{Y}_t] \otimes E_Q[X_s^{(n)}Z_t^{(n)}|\mathcal{Y}_t]}{E_Q[Z_t^{(n)}|\mathcal{Y}_t]^2} \right)^\top c(u)^\top (\sigma(u)\sigma(u)^\top)^{-1} dY_u,$$

其中 $g^{(n)}(s;t)$ 是一个确定性函数。根据引理 3.4，对于每个s，t，$g^{(n)}(s;t)$ 收敛，并且

$$\left\{ \frac{E_Q[X_u^{(n)} \otimes X_s^{(n)}Z_t^{(n)}|\mathcal{Y}_t]}{E_Q[Z_t^{(n)}|\mathcal{Y}_t]} - \frac{E_Q[X_u^{(n)}Z_t^{(n)}|\mathcal{Y}_t] \otimes E_Q[X_s^{(n)}Z_t^{(n)}|\mathcal{Y}_t]}{E_Q[Z_t^{(n)}|\mathcal{Y}_t]^2} \right\}_{0\le u\le t}$$

在 $L^2(\Omega; H)$ 中收敛，其中 $H = L^2([0,t])$。另一方面，我们已在上文证明，对于每个 s，$u \in [0,t]$，它收敛到 $\gamma(s, u; t)$，且收敛于 L^2 中。因此，令 $n \to \infty$，可得

$$E[X_s|\mathcal{Y}_t] = \frac{E_Q[X_s Z_t|\mathcal{Y}_t]}{E_Q[Z_t|\mathcal{Y}_t]} = g(s;t) + \int_0^t \gamma(s,u;t)c(u)^\top(\sigma(u)\sigma(u)^\top)^{-1}dY_u,$$

其中 $g(s;t)$ 是一个确定性函数。对等式两边关于 P 取期望，可得

$$E[X_s] = E[E[X_s|\mathcal{Y}_t]] = g(s;t) + \int_0^t \gamma(s,u;t)c(u)^\top(\sigma(u)\sigma(u)^\top)^{-1}c(u)E[X_u]du.$$

这给出了 $g(s;t)$ 的表达式，从而得到所需结果。 □

现在，路径 $\{X_s\}_{0\le s\le t}$ 在给定 $\mathcal{Y}_t$ 的条件密度可归结为计算 $\gamma(s, u;t)$（见4.3）。由于这是无条件期望之比，我们可以运用随机分析中成熟的方法。

为计算 γ，对每个固定的 $t \ge 0$，我们引入一个取值于 $M_{d_1}(\mathbb{R})$ 的可微函数 $\phi(s;t)$，将其定义为矩阵黎卡提方程的负半定对称解

$$\frac{d\phi}{ds}(s;t) = -\phi(s;t)b(s)b(s)^\top\phi(s;t) - a(s)^\top\phi(s;t) - \phi(s;t)a(s) + c(s)^\top(\sigma(s)\sigma(s)^\top)^{-1}c(s) \tag{3.20}$$

满足边界条件 $\phi(t;t) = 0$。该解是良定义的。具体而言，Riccati 方程

$$\frac{d\tilde{\phi}}{ds}(s;t) = -\tilde{\phi}(s;t)b(t-s)b(t-s)^\top\tilde{\phi}(s;t) - a(t-s)^\top\tilde{\phi}(s;t) - \tilde{\phi}(s;t)a(t-s) + c(t-s)^\top(\sigma(t-s)\sigma(t-s)^\top)^{-1}c(t-s)$$

˜ ˜满足 $\phi(0;t) = 0$ 存在一个半正定解 $\phi(s;t)$，对于 $0 \le s \le t$，根据 Potter (1965年中的定理 2.1 和 2.2 可知。于是立即˜可得 $\phi(s;t) = -\phi(t-s;t)$ 满足 (3.31)，而唯一性可利用格朗沃尔引理证明。

˜此外，令 $\{V_s\}$ 为定义在 $(\Omega\ \mathcal{F}, \{\mathcal{F}_t\}, P)$ 上的 d_1 维布朗运动，且与 $\{Y_s\}$ 和 X_0 相互独立。令 $\{\xi^0_{s;t}\}_{0\le s\le t}$ 为该随机微分方程的解

$$d_s\xi^0_{s;t} = \Big\{a(s) + b(s)b(s)^\top\phi(s;t)\Big\}\,\xi^0_{s;t}\,ds + b(s)\,d\tilde{V}_s,\quad \xi^0_{0;t} = X_0. \tag{3.21}$$

由此得到以下关键引理，它是吉尔萨诺夫定理和伊藤公式的直接推论。

引理 3.5. 设 $X_\cdot$ 和 $\xi^0_\cdot$ 分别为 $\{X_s\}_{0\le s\le t}$ 和 $\{\xi^0_{s,t}\}_{0\le s\le t}$的路径，且$\mu_X$ 和 μ_{ξ^0} 为它们在 P 上的分布。

$$\begin{aligned}\frac{d\mu_{\xi^0}}{d\mu_X}(X_\cdot) = \exp\Big(&-\frac{1}{2}\int_0^t X_s^\top c(s)^\top(\sigma(s)\sigma(s)^\top)^{-1}c(s)X_s ds\\ &-\frac{1}{2}\int_0^t b(s)^\top\phi(s;t)b(s)ds - \frac{1}{2}X_0^\top\phi(0;t)X_0\Big).\end{aligned} \tag{3.22}$$

证明。根据 Liptser 和 Shiryaev (2001a) 中的 (7.138)，μ_X 和 μ_{ξ^0} 等价，且有

$$\begin{aligned}\frac{d\mu_{\xi^0}}{d\mu_X}(X_\cdot) = \exp\Big(&\int_0^t X_s^\top\phi(s;t)dX_s\\ &-\frac{1}{2}\int_0^t X_s^\top\phi(s;t)(2a(s) + b(s)b(s)^\top\phi(s;t))X_s ds\Big).\end{aligned} \tag{3.23}$$

另一方面，根据伊藤公式和（3.20），我们有

$$\begin{aligned}&X_t^\top\phi(t;t)X_t - X_0^\top\phi(0;t)X_0\\ =&2\int_0^t X_s^\top\phi(s;t)dX_s + \int_0^t X_s^\top\frac{\partial}{\partial s}\phi(s;t)X_s ds + \int_0^t b(s)^\top\phi(s;t)b(s)ds\\ =&2\int_0^t X_s^\top\phi(s;t)dX_s - \int_0^t X_s^\top\phi(s;t)b(s)b(s)^\top\phi(s;t)X_s ds\\ &-\int_0^t X_s^\top a(s)^\top\phi(s;t)X_s ds - \int_0^t X_s^\top\phi(s;t)a(s)X_s ds\\ &+\int_0^t X_s^\top c(s)^\top(\sigma(s)\sigma(s)^\top)^{-1}c(s)X_s ds + \int_0^t b(s)^\top\phi(s;t)b(s)ds\\ =&2\int_0^t X_s^\top\phi(s;t)dX_s - \int_0^t X_s^\top\phi(s;t)b(s)b(s)^\top\phi(s;t)X_s ds\end{aligned}$$

$$
\begin{aligned}
&-2\int_0^t X_s^\top \phi(s;t)a(s)X_s ds \\
&+\int_0^t X_s^\top c(s)^\top(\sigma(s)\sigma(s)^\top)^{-1}c(s)X_s ds+\int_0^t b(s)^\top\phi(s;t)b(s)ds.
\end{aligned}
$$

因此，注意到 $\phi(t;t)=0$，可得

$$
\begin{aligned}
&\int_0^t X_s^\top \phi(s;t)dX_s \\
=&\frac{1}{2}\int_0^t X_s^\top \phi(s;t)b(s)b(s)^\top\phi(s;t)X_s ds+\int_0^t X_s^\top\phi(s;t)a(s)X_s ds \\
&-\frac{1}{2}\int_0^t X_s^\top c(s)^\top(\sigma(s)\sigma(s)^\top)^{-1}c(s)X_s ds \\
&-\frac{1}{2}\int_0^t b(s)^\top\phi(s;t)b(s)ds-\frac{1}{2}X_0^\top\phi(0;t)X_0.
\end{aligned}
$$

结合（3.23），我们得到所需结果。 □

根据这一结果，（$\gamma(s,u;t)$ 在（4.3）中的表达式的指数项可以改写为

$$
\begin{aligned}
&\exp\left(-\frac{1}{2}\int_0^t X_s^\top c(s)^\top(\sigma(s)\sigma(s)^\top)^{-1}c(s)X_s\,ds\right) \\
=&\exp\left(\frac{1}{2}X_0^\top\phi(0;t)X_0+\frac{1}{2}\int_0^t b(s)^\top\phi(s;t)b(s)ds\right)\frac{d\mu_{\xi^0}}{d\mu_X}(X_\cdot).
\end{aligned}
\tag{3.24}
$$

现在，让我们修改 $\xi^0_{s;t}$ 的初值，并定义一个新过程 $\xi_{s;t}$，其定义为

$$
d_s\xi_{s;t}=\left\{a(s)+b(s)b(s)^\top\phi(s;t)\right\}\xi_{s;t}\,ds+b(s)\,d\tilde{V}_s, \tag{3.25}
$$

其中，$\xi_{0;t}$ 是一个高斯随机变量，与 $\{V_s\}$和 $\{Y_s\}$相互独立，其均值为零，协方差矩阵为

$$
\mathrm{Var}[\xi_{0;t}]=V[X_0]^{1/2}\left(I_{d_1}-V[X_0]^{1/2}\phi(0;t)V[X_0]^{1/2}\right)^{-1}V[X_0]^{1/2}. \tag{3.26}
$$

这里需要注意，矩阵 $I_{d_1}-V[X_0]^{1/2}\phi(0;t)V[X_0]^{1/2}$ 是正定的，因为 ϕ 被定义为负半定解。于是，$\gamma(s,u;t)$与过程 ξ的协方差矩阵一致。

命题 3.6。对于每个 s、$u\in[0,t]$，都有 $\gamma(s、u;t)$ =协方差（$\xi_{s;t}$、$\xi_{u;t}$）。

证明。 利用（3.24），我们得到

$$\begin{aligned}
\gamma(s,u;t) =& \frac{E\left[X_s \otimes X_u \exp\left(\frac{1}{2}X_0^\top \phi(0;t) X_0\right) \frac{d\mu_{\xi^0}}{d\mu_X}(X.)\middle| \mathcal{Y}_t\right]}{E\left[\exp\left(\frac{1}{2}X_0^\top \phi(0;t) X_0\right) \frac{d\mu_{\xi^0}}{d\mu_X}(X.)\middle| \mathcal{Y}_t\right]} \\
&- \frac{E\left[X_s \exp\left(\frac{1}{2}X_0^\top \phi(0;t) X_0\right) \frac{d\mu_{\xi^0}}{d\mu_X}(X.)\middle| \mathcal{Y}_t\right]}{E\left[\exp\left(\frac{1}{2}X_0^\top \phi(0;t) X_0\right) \frac{d\mu_{\xi^0}}{d\mu_X}(X.)\middle| \mathcal{Y}_t\right]} \\
&\otimes \frac{E\left[\exp\left(\frac{1}{2}X_0^\top \phi(0;t) X_0\right) \frac{d\mu_{\xi^0}}{d\mu_X}(X.)\middle| \mathcal{Y}_t\right]}{E\left[\exp\left(\frac{1}{2}X_0^\top \phi(0;t) X_0\right) \frac{d\mu_{\xi^0}}{d\mu_X}(X.)\middle| \mathcal{Y}_t\right]} \\
=& \frac{E_Q\left[\xi^0_{s;t} \otimes \xi^0_{u;t} \exp\left(\frac{1}{2}X_0^\top \phi(0;t) X_0\right)\right]}{E_Q\left[\exp\left(\frac{1}{2}X_0^\top \phi(0;t) X_0\right)\right]} \\
&- \frac{E_Q\left[\xi^0_{s;t} \exp\left(\frac{1}{2}X_0^\top \phi(0;t) X_0\right)\right]}{E_Q\left[\exp\left(\frac{1}{2}X_0^\top \phi(0;t) X_0\right)\right]} \otimes \frac{E_Q\left[\xi^0_{u;t} \exp\left(\frac{1}{2}X_0^\top \phi(0;t) X_0\right)\right]}{E_Q\left[\exp\left(\frac{1}{2}X_0^\top \phi(0;t) X_0\right)\right]}.
\end{aligned} \tag{3.27}$$

现在，将 $\xi^0_{s;t}$ 写为

$$\xi^0_{s;t} = \alpha(s;t)X_0 + \bar{\xi}_{s;t}, \tag{3.28}$$

其中

$$\begin{aligned}
\alpha(s;t) &= \exp\left(\int_s^t \left\{a(r) + b(r)b(r)^\top \phi(r;t)\right\}\right), \\
\bar{\xi}_{s;t} &= \int_0^s \exp\left(\int_u^s \left\{a(r) + b(r)b(r)^\top \phi(r;t)\right\}\right) b(s) d\tilde{V}_s
\end{aligned}$$

利用这一表达式，并由于 $\xi_{s;t} = \alpha(s;t)\xi_{0;t} + \bar{\xi}_{s;t}$ 以及 (3.26), (3.27) 可等价地改写为

$$\begin{aligned}
&\mathrm{Cov}(\bar{\xi}_{s;t}, \bar{\xi}_{u;t}) + \frac{E\left[\alpha(s;t)X_0 \otimes \alpha(u;t)X_0 \exp\left(\frac{1}{2}X_0^\top \phi(0;t) X_0\right)\right]}{E\left[\exp\left(\frac{1}{2}X_0^\top \phi(0;t) X_0\right)\right]} \\
&- \frac{E\left[\alpha(s;t)X_0 \exp\left(\frac{1}{2}X_0^\top \phi(0;t) X_0\right)\right]}{E\left[\exp\left(\frac{1}{2}X_0^\top \phi(0;t) X_0\right)\right]}
\end{aligned}$$

$$\otimes \frac{E\left[\alpha(u;t)X_0 \exp\left(\frac{1}{2}X_0^\top \phi(0;t)X_0\right)\right]}{E\left[\exp\left(\frac{1}{2}X_0^\top \phi(0;t)X_0\right)\right]}$$

=协方差($\xi_{s;t}^-$, $\xi_{u;t}$) + $\alpha(s;t)V[X_0]^{\frac{1}{2}}(I - V[X_0]^{\frac{1}{2}}\phi(0;t)V[X_0]^{\frac{1}{2}})^{-1}V[X_0]^{\frac{1}{2}}\alpha(u;t)^\top$=协方差($\xi_{s;t}$, $\xi_{u;t}$)。

□

综上所述，我们得到以下定理。

定理 3.7. $\{X_s\}_{0\le s\le t}$ 在给定 $\mathcal{Y}_t$ 的条件下，其条件分布是一个高斯过程，且

$$\begin{aligned} &E[X_s|\mathcal{Y}_t] = E[X_s] \\ &+ \int_0^t \mathrm{Cov}(\xi_{s;t},\xi_{u;t})c(u)^\top(\sigma(u)\sigma(u)^\top)^{-1}(dY_u - c(u)E[X_u]\,du), \end{aligned} \tag{3.29}$$

$$\mathrm{Cov}(X_s, X_u|\mathcal{Y}_t) = \mathrm{Cov}(\xi_{s;t},\xi_{u;t}), \tag{3.30}$$

其中

$$\begin{aligned} \frac{d\phi}{ds}(s;t) = &-\phi(s;t)b(s)b(s)^\top\phi(s;t) - a(s)^\top\phi(s;t) - \phi(s;t)a(s) \\ &+ c(s)^\top(\sigma(s)\sigma(s)^\top)^{-1}c(s), \end{aligned} \tag{3.31}$$

其中 $\phi(t;t) = 0$，并且

$$d_s\xi_{s;t} = \left\{a(s) + b(s)b(s)^\top\phi(s;t)\right\}\xi_{s;t}\,ds + b(s)\,d\tilde{V}_s \tag{3.32}$$

其中 $E[\xi_{0;t}] = 0$ 且

$$V[\xi_{0;t}] = V[X_0]^{\frac{1}{2}}\left(I_{d_1} - V[X_0]^{\frac{1}{2}}\phi(0;t)V[X_0]^{\frac{1}{2}}\right)^{-1}V[X_0]^{\frac{1}{2}}. \tag{3.33}$$

因此，平滑误差 $\{X_s - E[X_s|\mathcal{Y}_t]\}_{0\le s\le t}$ 与奥恩斯坦–乌伦贝克过程 $\{\xi_{s;t}\}_{0\le s\le t}$具有相同的分布。

该定理使得对条件误差进行路径采样成为可能。第5节介绍了其实际应用。

4 滤波与平滑方程

本节将说明，滤波方程和平滑方程可直接由主定理推导得出。我们记

$$\mu_{s;t} = E[X_s \mid \mathcal{Y}_t], \quad \gamma(s,u;t) = \mathrm{Cov}(X_s, X_u \mid \mathcal{Y}_t),$$

我们首先将 Bryson–Frazier 平滑器推导为定理3.7的直接推论。

命题 4.1。（Bryson–Frazier 平滑器）
条件均值 $\mu_{s;t}$ 满足方程

$$\begin{aligned} d_s\mu_{s;t} =& \left\{a(s)+b(s)b(s)^\top\phi(s;t)\right\}\mu_{s;t}ds - b(s)b(s)^\top\phi(s;t)E[X_s]ds \\ &+ b(s)b(s)^\top\rho_{s;t}ds, \\ \mu_{0;t} =& V[X_0]^{\frac{1}{2}}\left(I_{d_1} - V[X_0]^{\frac{1}{2}}\phi(0;t)V[X_0]^{\frac{1}{2}}\right)^{-1}V[X_0]^{\frac{1}{2}}\rho_{0;t} \end{aligned}$$

其中 $\rho_{s;t}$ 是方程的解

$$\begin{aligned} d_s\rho_{s;t} =& \left\{a(s)+b(s)b(s)^\top\phi(s;t)\right\}\rho_{s;t}ds \\ &+ c(s)^\top(\sigma(s)\sigma(s))^{-1}(dY_s - c(s)E[X_s]ds), \\ \rho_{t;t} =& 0. \end{aligned}$$

证明。 (3.32)的解表示为

$$\xi_{s;t} = \alpha(0,s;t)\xi_{0;t} + \int_0^s \alpha(r,s;t)b(r)d\tilde{V}_r,$$

其中 $\alpha(u,\ s;t)$ 是

$$\frac{\partial}{\partial s}\alpha(u,s;t) = \left\{a(s)+b(s)b(s)^\top\phi(s;t)\right\}\alpha(u,s;t),\quad \alpha(u,u;t) = I. \tag{4.1}$$

因此我们有

$$\begin{aligned} \gamma(s,u;t) = E[\xi_{s;t}\xi_{u;t}^\top] =& \alpha(0,s;t)V[\xi_{0;t}]\alpha(0,u;t)^\top \\ &+ \int_0^{s\wedge u}\alpha(r,s;t)b(r)b(r)^\top\alpha(r,u;t)^\top dr \end{aligned} \tag{4.2}$$

and

$$\frac{\partial}{\partial s}\gamma(s,u;t) = \begin{cases} \left\{a(s)+b(s)b(s)^\top\phi(s;t)\right\}\gamma(s,u;t) & (s<u) \\ \left\{a(s)+b(s)b(s)^\top\phi(s;t)\right\}\gamma(s,u;t) + b(s)b(s)^\top\alpha(s,u;t)^\top & (s\geq u) \end{cases}.$$

因此，对 (3.29) 求导得到

$$\begin{aligned} &d_s\mu_{s;t} = \left\{a(s)+b(s)b(s)^\top\phi(s;t)\right\}\mu_{s;t}ds - b(s)b(s)^\top\phi(s;t)E[X_s]ds \\ &+ b(s)b(s)^\top\int_s^t \alpha(s,u;t)^\top c(u)^\top(\sigma(u)\sigma(u)^\top)^{-1}(dY_u - c(u)E[X_u]du)ds. \end{aligned}$$

现在，通过写出下式即可得到所需结果

$$\rho_{s;t} = \int_s^t \alpha(s,u;t)^\top c(u)^\top(\sigma(u)\sigma(u)^\top)^{-1}(dY_u - c(u)E[X_u]du).$$

□

接下来，我们推导 $\gamma(t,\ t;t) = V[X_t|\mathcal{Y}_t]$ 所满足的微分方程。为此，我们将 $\gamma(t)$定义为黎卡提方程的半正定解

$$\frac{d}{dt}\gamma(t) = -\gamma(t)c(t)^\top(\sigma(t)\sigma(t)^\top)^{-1}c(t)\gamma(t) + a(t)\gamma(t) + \gamma(t)a(t)^\top + b(t)b(t)^\top \tag{4.3}$$

与 $\gamma(0) = V[X_0]$。解的存在性由 Potter (1965年) 中的定理 2.1 和 2.2 保证。

为避免 $\gamma(t)$ 可能出现奇异性的问题，我们引入 $\gamma^\epsilon(t)$ 和$\xi^\epsilon_{s;t}$，将它们作为相同方程的解

$$\begin{aligned}\frac{d}{dt}\gamma^\epsilon(t) = &-\gamma^\epsilon(t)c(t)^\top(\sigma(t)\sigma(t)^\top)^{-1}c(t)\gamma^\epsilon(t) + a(t)\gamma^\epsilon(t) + \gamma^\epsilon(t)a(t)^\top \\ &+ b(t)b(t)^\top\end{aligned} \tag{4.4}$$

and

$$d\xi^\epsilon_{s;t} = \left\{a(s) + b(s)b(s)^\top\phi(s;t)\right\}\xi^\epsilon_{s;t}\,ds + b(s)\,d\tilde{V}_s, \tag{4.5}$$

其中 $\gamma^\epsilon(0) = V[X_0] + \epsilon I_{d_1}$，且 $\xi^\epsilon_{0;t}$ 是均值为零、协方差为

$$V[\xi^\epsilon_{0;t}] = \left\{(V[X_0] + \epsilon I_{d_1})^{-1} - \phi(0;t)\right\}^{-1}. \tag{4.6}$$

根据 Potter (1965年) 中定理 2.1 的推论 1，$\gamma^\epsilon(t)$ 严格为正。

引理 4.2。*(1)* 对每个 $t \geq 0$，$\gamma^\epsilon(s)$ 有界，其中 $\epsilon \in [0,1]$ 和 $s \in[0,t]$。

(2) 矩阵 $\gamma^\epsilon(t)$ 在任意紧集上一致收敛到 $\gamma(t)$，当$\epsilon \to +0$时。

证明。（1）任取 $x \in \mathbb{R}^{d_1}$。则由（4.3）可得

$$\begin{aligned}\frac{d}{dt}x^\top\gamma^\epsilon(t)x =& -x^\top\gamma^\epsilon(t)c(t)^\top(\sigma(t)\sigma(t)^\top)^{-1}c(t)\gamma^\epsilon(t)x + 2x^\top a(t)\gamma^\epsilon(t)x \\ &+ x^\top b(t)b(t)^\top x \\ \leq& 2|x^\top a(t)\gamma^\epsilon(t)^{\frac{1}{2}}\gamma^\epsilon(t)^{\frac{1}{2}}x| + x^\top b(t)b(t)^\top x \\ \leq& 2\sqrt{x^\top a(t)\gamma^\epsilon(t)a(t)^\top x}\sqrt{x^\top\gamma^\epsilon(t)x} + x^\top b(t)b(t)^\top x \\ \leq& 2\|a(t)\|_2\|\gamma^\epsilon(t)\|_2|x|^2 + \|b(t)\|_2^2|x|^2,\end{aligned}$$

其中 $\|\cdot\|_2$ 是矩阵 2-范数，并且我们使用了

$$x^\top a(t)\gamma^\epsilon(t)a(t)^\top x = |\gamma^\epsilon(t)^{\frac{1}{2}}a(t)^\top x|^2 \leq \|\gamma^\epsilon(t)\|_2\|a(t)\|_2^2|x|^2.$$

由此可得

$$\|\gamma^{\epsilon}(t)\|_2 = \max_{|x|=1} x^{\top}\gamma^{\epsilon}(t)x \leq \|\gamma^{\epsilon}(0)\|_2 + \int_0^t \{2\|a(s)\|_2\|\gamma^{\epsilon}(s)\|_2 + \|b(s)\|_2^2\}ds.$$

因此，根据格朗沃尔引理，我们得到

$$\begin{aligned}\|\gamma^{\epsilon}(t)\|_2 \leq & \|\gamma^{\epsilon}(0)\|_2 \exp\left(2\int_0^t \|a(r)\|_2 dr\right) \\ & + \int_0^t \exp\left(2\int_s^t \|a(r)\|_2 dr\right)\|b(s)\|_2^2 ds,\end{aligned}$$

由此可得所需结果。
利用（1）的结果和格朗沃尔引理，可以常规地证明陈述（2）。

□

命题 4.3。有 $\gamma(t) = \gamma(t,\ t;\ t)$。

证明。 首先，假设对于每个 $t \geq 0$， $\gamma(t)$ 都严格为正。那么我们有

$$\begin{aligned}\frac{d}{ds}\left(\gamma(s)^{-1}\right) = & -\gamma(s)^{-1}\frac{d}{ds}\gamma(s)\gamma(s)^{-1} \\ = & -\gamma(s)^{-1}b(s)b(s)^{\top}\gamma(s)^{-1} - a(s)^{\top}\gamma(s)^{-1} - \gamma(s)^{-1}a(s) \\ & + c(s)^{\top}(\sigma(s)\sigma(s)^{\top})^{-1}c(s).\end{aligned}$$

这意味着 $\gamma(s)^{-1}$ 满足与 3.31 相同的方程（ $\phi(s;\ t)$ 。因此，若令 $z(s;\ t) = \gamma(s)^{-1} - \phi(s;\ t)$，则可得

$$\begin{aligned}&\frac{d}{ds}z(s;t) = -\left\{\gamma(s)^{-1}b(s)b(s)^{\top}\gamma(s)^{-1} - \phi(s;t)^{-1}b(s)b(s)^{\top}\phi(s;t)^{-1}\right\} \\ &- a(s)^{\top}\left\{\gamma(s)^{-1} - \phi(s;t)\right\} - \left\{\gamma(s)^{-1} - \phi(s;t)\right\}a(s) \\ = & -z(s;t)b(s)b(s)^{\top}z(s;t) - \left\{a(s)^{\top} + \phi(s;t)b(s)b(s)^{\top}\right\}z(s;t) \\ &- z(s;t)\left\{a(s) + b(s)b(s)^{\top}\phi(s;t)\right\}.\end{aligned}$$

此外，如果我们令 $w(s;t) = z(s;t)^{-1}$（这是有良好定义的，因为 $\phi(s;\ t)$ 是负半定的），则有

$$\begin{aligned}\frac{d}{ds}w(s;t) = & -z(s;t)^{-1}\frac{d}{ds}z(s;t)z(s;t)^{-1} \\ = & b(s)b(s)^{\top} + w(s;t)\left\{a(s)^{\top} + \phi(s;t)b(s)b(s)^{\top}\right\} \\ & + \left\{a(s) + b(s)b(s)^{\top}\phi(s;t)\right\}w(s;t)\end{aligned} \tag{4.7}$$

因此， $w(s;t)$ 可以写成

$$w(s;t) = \alpha(0,s;t)V[\xi_{0;t}]\alpha(0,u;t) + \int_0^s \alpha(u,s;t)b(u)b(u)^\top \alpha(u,s;t)^\top du. \tag{4.8}$$

其中 α 在（4.1）中定义。这里，我们利用 $\gamma(0) = V[X_0]$ 和（3.33）来建立 $w(0;s) = (V[X_0]^{-1} - \phi(0;t))^{-1} = V[\xi_{0;t}]$.

因此，由（4.2）和（4.8）立即可得 $\gamma(s,s;t) = V[\xi_{s;t}] = w(s;t)$ 。特别地，由于

$$w(s;t) = \{\gamma(s)^{-1} - \phi(s;t)\}^{-1} \tag{4.9}$$

并且 $\phi(t;t) = 0$，我们有 $w(t;t) = \gamma(t)$，由此得到所需结果。

若对于某个 t， $\gamma(t)$ 变为奇异，则对 $\epsilon > 0$，我们用 $V[X_0] + \epsilon I_{d_1}$ 替换 $V[X_0]$，从而得到 $\gamma^\epsilon(t) = V[\xi^\epsilon_{t;t}]$，因为 $\gamma^\epsilon(t)$ 严格为正。然后令 $\epsilon \to 0$ ，即可得出结论。

□

注。 证明表明， $\psi(s;t)$ 可以表示为

$$\psi(s;t) = \left\{\gamma(s)^{-1} - \gamma(s,s;t)^{-1}\right\}^{-1},$$

只要 $\gamma(s)$、 $\gamma(s,\ s;\ t)$ 和 $\gamma(s)^{-1} - \gamma(s,\ s;\ t)^{-1}$ 非奇异即可。这一观察解释了为何在 Kalman–Bucy 滤波器和 RTS 平滑器的框架内推导 BF 平滑器时，逆矩阵 $\gamma(s)^{-1}$ 和 $\gamma(s,\ s;\ t)^{-1}$ 不可避免地会出现。

命题 4.4。 设 $0 \le u \le s \le t$。若对每个 $r \in [0,t]$， $\gamma(r)$ 均非奇异，则有

$$d_u\gamma(s,u;t) = \gamma(s,u;t)\left\{a(u)^\top + \gamma(u)^{-1}b(u)b(u)^\top\right\}du.$$

证明。 由（4.2）及 $w(s;\ t) = \gamma(s,\ s;\ t)$，我们可以表示为

$$\gamma(s,u;t) = \alpha(u,s;t)w(u;t), \tag{4.10}$$

其中 $w(u;t) = V[\xi_{u;t}]$ 由式（4.8）给出。因此，由式（4.7）可得

(4.9) 可知

$$
\begin{aligned}
&d_u\gamma(s,u;t)\\
=&\alpha(u,s;t)\left[-\left\{a(u)+b(u)b(u)^\top\phi(u;t)\right\}w(u;t)+\frac{\partial}{\partial u}w(u;t)\right]du\\
=&\alpha(u,s;t)\\
&\times\left[-\left\{a(u)+b(u)b(u)^\top\phi(u;t)\right\}w(u;t)+b(u)b(u)^\top\right.\\
&+w(u;t)\left\{a(u)^\top+\phi(u;t)b(u)b(u)^\top\right\}\\
&\left.+\left\{a(u)+b(u)b(u)^\top\phi(u;t)\right\}w(u;t)\right]du\\
=&\alpha(u,s;t)w(u,t)\\
&\times\left\{w(u;t)^{-1}b(u)b(u)^\top+a(u)^\top+\phi(u;t)b(u)b(u)^\top\right\}du\\
=&\mathrm{Cov}(\xi_{s;t},\xi_{u;t})\left[\left\{\gamma(u)^{-1}-\phi(u;t)\right\}b(u)b(u)^\top+a(u)^\top+\phi(u;t)b(u)b(u)^\top\right]du\\
=&\gamma(s,u;t)\left\{a(u)^\top+\gamma(u)^{-1}b(u)b(u)^\top\right\}du,
\end{aligned}
$$

由此得出结论。 □

命题 4.5. 对于每个 $s\geq 0$，均有

$$d_t\gamma(t,s;t)=\left\{a(t)-\gamma(t)c(t)^\top(\sigma(t)\sigma(t)^\top)^{-1}c(t)\right\}\gamma(t,s;t)dt.$$

证明。 首先，假设 $\gamma(t)$ 对每个 $t\geq 0$ 均为非奇异矩阵。在这种情况下，根据命题 4.3 和 4.4，我们可以写出

$$\gamma(t,s;t)=\mathrm{Cov}(\xi_{t;t},\xi_{s;t})=\gamma(t)\beta(s,t)^\top.$$

其中 $\beta(u,s)$ 是以下方程的解

$$\frac{\partial}{\partial s}\beta(u,s)=\left\{a(r)+b(r)b(r)^\top\gamma(r)^{-1}\right\}\beta(u,s),\quad \beta(u,u)=I. \tag{4.11}$$

因此，有

$$
\begin{aligned}
&d_t\gamma(t,s;t)\\
=&\left[\frac{d}{dt}\gamma(t)-\gamma(t)\{a(t)^\top+\gamma(t)^{-1}b(t)b(t)^\top\}\right]\beta(s,t)^\top dt\\
=&\left\{-\gamma(t)c(t)^\top(\sigma(t)\sigma(t)^\top)^{-1}c(t)\gamma(t)+a(t)\gamma(t)\right\}\beta(s,t)^\top dt\\
=&\left\{-\gamma(t)c(t)^\top(\sigma(t)\sigma(t)^\top)^{-1}c(t)+a(t)\right\}\gamma(t,s;t)dt,
\end{aligned}
$$

这就意味着该结论。对于 $\gamma(t)$ 奇异的情况，我们有

$$d_t\mathrm{Cov}(\xi^\epsilon_{t;t},\xi^\epsilon_{s;t})=\left\{-\gamma^\epsilon(t)c(t)^\top(\sigma(t)\sigma(t)^\top)^{-1}c(t)+a(t)\right\}\mathrm{Cov}(\xi^\epsilon_{t;t},\xi^\epsilon_{s;t})dt$$

并且 γ^ϵ 在任意紧集上一致收敛，从而得到所需结果。

□

前一个命题给出了卡尔曼-布西滤波器和劳赫-汤-斯特里贝尔平滑器的著名结果。

命题 4.6。（卡尔曼-布西滤波器）
过程 $\{\mu_{t,t}\}_{t\geq 0}$ 满足如下随机微分方程

$$d\mu_{t,t} = a(t)\mu_{t,t}dt + \gamma(t)c(t)^\top(\sigma(t)\sigma(t)^\top)^{-1}\{dY_t - c(t)\mu_{t,t}dt\}.$$

证明。令

$$\tilde{\mu}_{t;t} = \mu_{t;t} - E[X_t], \quad \tilde{Y}_t = Y_t - \int_0^t c(u)E[X_u]du. \tag{4.12}$$

那么，根据命题 4.5 和（3.29），我们有

$$\begin{aligned}
\tilde{\mu}_{t,t} &= \int_0^t \gamma(t,u;t)c(u)^\top(\sigma(u)\sigma(u)^\top)^{-1}d\tilde{Y}_u \\
&= \int_0^t \left[\gamma(u) + \int_u^t \left\{-\gamma(s)c(s)^\top(\sigma(s)\sigma(s)^\top)^{-1}c(s) + a(s)\right\}\gamma(s,u;s)ds\right] \\
&\quad \times c(u)^\top(\sigma(u)\sigma(u)^\top)^{-1}d\tilde{Y}_u \\
&= \int_0^t a(s)\int_0^s \gamma(s,u;s)c(u)^\top(\sigma(u)\sigma(u)^\top)^{-1}d\tilde{Y}_u ds \\
&\quad + \int_0^t \gamma(u)c(u)^\top(\sigma(u)\sigma(u)^\top)^{-1}d\tilde{Y}_u \\
&\quad - \int_0^t \gamma(s)c(s)^\top(\sigma(s)\sigma(s)^\top)^{-1}c(s)\int_0^s \gamma(s,u;s)c(u)^\top(\sigma(u)\sigma(u)^\top)^{-1}d\tilde{Y}_u ds \\
&= \int_0^t a(s)\tilde{\mu}_{s,s}ds + \int_0^t \gamma(u)c(u)^\top(\sigma(u)\sigma(u)^\top)^{-1}\{d\tilde{Y}_u - c(s)\tilde{\mu}_{s,s}ds\},
\end{aligned}$$

由此得到方程

$$d\tilde{\mu}_{t,t} = a(t)\tilde{\mu}_{t,t}dt + \gamma(t)c(t)^\top(\sigma(t)\sigma(t)^\top)^{-1}\{d\tilde{Y}_t - c(t)\tilde{\mu}_{t,t}dt\}.$$

根据（4.12）和 $dE[X_t] = a$（t）$E[X_t]dt$，这等价于所需结果。

□

命题 4.7.（劳赫-汤-斯特里贝尔平滑器）
如果对于每个 $r \in [0,t]$，$\gamma(r)$ 都严格为正，那么对于每个 $0 \leq s \leq t$，该结论都成立

$$\mu_{s;t} = \mu_{t;t} - \int_s^t \left\{a(r)\mu_{r;t} + b(r)b(r)^\top\gamma(r)^{-1}(\mu_{r;t} - \mu_{r;r})\right\}dr, \tag{4.13}$$

and

$$\begin{aligned}
\frac{d}{ds}\gamma(s,s;t) = &- b(s)b(s)^\top + \gamma(s,s;t)\left\{a(s)^\top + \gamma(s)^{-1}b(s)b(s)^\top\right\} \\
&+ \left\{a(s) + b(s)b(s)^\top\gamma(s)^{-1}\right\}\gamma(s,s;t).
\end{aligned} \tag{4.14}$$

证明。令

$$\tilde{\mu}_{t;t} = \mu_{t;t} - E[X_t], \quad \hat{Y}_t = \int_0^t c(u)^\top (\sigma(u)\sigma(u)^\top)^{-1}(dY_u - c(u)E[X_u]\,du).$$

根据命题 4.4，对于 $s \le u$ 我们有

$$d_s\gamma(s,u;t) = \left\{a(s) + b(s)b(s)^\top\gamma(s)^{-1}\right\}\gamma(s,u;t)ds$$

另一方面，由（4.10）可知，对于 $s \ge u$

$$d_s\gamma(s,u;t) = \left\{a(s) + b(s)b(s)^\top\phi(s;t)\right\}\gamma(s,u;t)ds. \tag{4.15}$$

因此，有

$$\begin{aligned}
\tilde{\mu}_{s;t} =& \int_0^t \left\{\gamma(t,u;t) - \int_s^t \frac{\partial}{\partial r}\gamma(r,u;t)dr\right\} d\hat{Y}_u \\
=& \tilde{\mu}_{t;t} - \int_s^t \int_0^t \frac{\partial}{\partial r}\gamma(r,u;t)d\hat{Y}_u dr \\
=& \tilde{\mu}_{t;t} - \int_s^t \int_0^r \left\{a(r) + b(r)b(r)^\top\phi(r;t)\right\}\gamma(r,u;t)d\hat{Y}_u dr \\
& - \int_s^t \int_r^t \left\{a(r) + b(r)b(r)^\top\gamma(r)^{-1}\right\}\gamma(r,u;t)d\hat{Y}_u dr \\
=& \tilde{\mu}_{t;t} - \int_s^t \left\{a(r) + b(r)b(r)^\top\gamma(r)^{-1}\right\}\int_0^t \gamma(r,u;t)d\hat{Y}_u dr \\
& + \int_s^t \int_0^r b(r)b(r)^\top\{\gamma(r)^{-1} - \phi(r;t)\}\gamma(r,u;t)d\hat{Y}_u dr \\
=& \tilde{\mu}_{t;t} - \int_s^t \left\{a(r) + b(r)b(r)^\top\gamma(r)^{-1}\right\}\tilde{\mu}_{r;t}dr \\
& + \int_s^t b(r)b(r)^\top \int_0^r \gamma(r,r;t)^{-1}\gamma(r,u;t)d\hat{Y}_u dr.
\end{aligned}$$

在最后一个等式中，我们使用了关系（4.9）。

此外，根据命题 4.4，$\gamma(r,\ r;\ t)^{-1}\gamma(r,\ u;\ t)$ 不依赖于 t。因此，我们有

$$\int_0^r V[\xi_{r;t}]^{-1}\gamma(r,u;t)d\hat{Y}_u = \int_0^r \gamma(r,r;r)^{-1}\gamma(r,u;r)d\hat{Y}_u = \gamma(r)^{-1}\tilde{\mu}_{r,r}.$$

因此，我们得到

$$\tilde{\mu}_{s;t} = \tilde{\mu}_{t;t} - \int_s^t \left\{a(r) + b(r)b(r)^\top\gamma(r)^{-1}\right\}\tilde{\mu}_{r;t}dr + \int_s^t b(r)b(r)^\top\gamma(r)^{-1}\tilde{\mu}_{r,r}dr.$$

最后，我们通过使用（4.13）得到 $\tilde{\mu}_{s;t} = \mu_{s;t} - E[X_s]$ 和 $dE[X_s] = a(s)\ E[X_s]\ dt$. 为证明（4.14），在 $\gamma(s)^{-1}$ 和 $\phi(s;\ t)$ 中交换 4.7，并回忆 $w(s;\ t) = V[\xi_{s;t}] = \gamma(s,\ s;\ t)$。 □

命题 4.8。 对于 s、$u \geq 0$ 和 $t \geq s \vee u$，有

$$d_t\gamma(s,u;t) = -\gamma(t,s;t)^\top c(t)^\top(\sigma(t)\sigma(t)^\top)^{-1}c(t)\gamma(t,u;t)dt.$$

证明。 首先假设 $\gamma(t)$ 对每个 $t \geq 0$ 都为正。于是，由于 (4.14)，有

$$\begin{aligned}\frac{d}{ds}\gamma(s,s;t) =& -b(s)b(s)^\top + \gamma(s,s;t)\left\{a(s)^\top + \gamma(s)^{-1}b(s)b(s)^\top\right\}\\ &+\left\{a(s)+b(s)b(s)^\top\gamma(s)^{-1}\right\}\gamma(s,s;t).\end{aligned}$$

因此，对于 $s \leq t$，我们得到

$$\gamma(s,s;t) = \beta(t,s)\gamma(t)\beta(t,s)^\top + \int_s^t \beta(u,s)b(u)b(u)^\top\beta(u,s)^\top du,$$

其中 β 定义见 (4.11)，且我们使用了 $V[\xi_{t;t}] = \gamma(t)$。对其求导可得

$$\begin{aligned}&d_t\gamma(s,s;t) = \beta(t,s)\\ &\times\left[-\left\{a(t)+b(t)b(t)^\top\gamma(t)^{-1}\right\}\gamma(t) - \gamma(t)\left\{a(t)^\top+\gamma(t)^{-1}b(t)b(t)^\top\right\}\right.\\ &\left.+\frac{d}{dt}\gamma(t) + b(t)b(t)^\top\right]\beta(t,s)\\ =& -\beta(t,s)\gamma(t)c(t)^\top(\sigma(t)\sigma(t)^\top)^{-1}c(t)\gamma(t)\beta(t,s)t\end{aligned}$$

对于 $t \geq s$。因此，注意到由命题 4.4 可知 $\gamma(t)\beta(t,s)^\top = \gamma(t,s;t)$，我们得到

$$\begin{aligned}&d_t\gamma(s,u;t) = d_t\gamma(s,s;t)\beta(s,u)\\ =& -\beta(t,s)\gamma(t)c(t)^\top(\sigma(t)\sigma(t)^\top)^{-1}c(t)\gamma(t)\beta(t,u)dt\\ =& -\gamma(t,s;t)^\top c(t)^\top(\sigma(t)\sigma(t)^\top)^{-1}c(t)\gamma(t,u;t)dt\end{aligned} \tag{4.16}$$

对于 $0 \leq u \leq s \leq t$。当 $s \leq u$ 时，注意到 $\frac{\partial}{\partial t}\mathrm{Cov}(\xi_{s;t},\xi_{u;t}) = \frac{\partial}{\partial t}\mathrm{Cov}(\xi_{u;t},\xi_{s;t})^\top$，我们可以推导出相同的公式。最后，假设 $\gamma(t)$ 可能是奇异的。那么，我们有

$$d_t\mathrm{Cov}(\xi^\epsilon_{s;t},\xi^\epsilon_{u;t}) = -\mathrm{Cov}(\xi^\epsilon_{t;t},\xi^\epsilon_{s;t})^\top c(t)^\top(\sigma(t)\sigma(t)^\top)^{-1}c(t)\mathrm{Cov}(\xi^\epsilon_{t;t},\xi^\epsilon_{u;t})\,dt,$$

其中 $\gamma(s)^\epsilon$ 以及 $\xi^\epsilon_{s;t}$ 在（4.4）至（4.6）中定义。现在，注意到 $\mathrm{Cov}(\xi^\epsilon_{t;t},\xi^\epsilon_{s;t})$ 在任意紧集上均匀收敛于 $\mathrm{Cov}(\xi_{t;t},\xi_{s;t}) = \gamma(t,s;t)$，当 $\epsilon \to 0$ 时，即得所需结果。

□

命题 4.9。（不动点平滑器）
对于任意 $s \geq 0$，过程 $\{\mu_{s;t}\}_{t \geq s}$ 满足随机微分方程

$$d_t \mu_{s;t} = \gamma(s,t;t) c(t)^\top (\sigma(t)\sigma(t)^\top)^{-1} (dY_t - c(t)\mu_{t;t} dt).$$

证明。 根据 $\mu_{s;t}$ 的定义以及命题 4.8，可得

$$\begin{aligned}
& d_t \mu_{s;t} \\
=& \gamma(s,t;t) c(t)^\top (\sigma(t)\sigma(t)^\top)^{-1} (dY_t - c(t)E[X_t]dt) \\
& - \int_0^t \gamma(t,s;t)^\top c(t)^\top (\sigma(t)\sigma(t)^\top)^{-1} c(t) \gamma(t,u;t) c(u)^\top (\sigma(u)\sigma(u)^\top)^{-1} \\
& \times (dY_u - c(u)E[X_u]du) dt \\
=& \gamma(s,t;t) c(t)^\top (\sigma(t)\sigma(t)^\top)^{-1} (dY_t - c(t)E[X_t]dt) \\
& - \gamma(t,s;t)^\top c(t)^\top (\sigma(t)\sigma(t)^\top)^{-1} c(t)(\mu_{t;t} - E[X_t]) dt \\
=& \gamma(s,t;t) c(t)^\top (\sigma(t)\sigma(t)^\top)^{-1} (dY_t - c(t)\mu_{t;t} dt).
\end{aligned}$$

□

5 进一步应用

如上所述，定理 3.7支持从隐状态过程的条件分布中进行路径采样。这带来了一系列潜在应用，包括：

- **非线性泛函的期望：** 从完整平滑分布中进行路径采样的能力，使得可以近似涉及非线性和路径依赖泛函的期望，例如

$$E\left[\max_{0 \leq s \leq t} X_s \,\middle|\, \mathcal{Y}_t\right],$$

  而在标准平滑方程下，这些期望是无法解析求得的。这类期望与金融期权定价和工程阈值检测相关。

- **通过 EM 算法进行参数估计。** 当部分数据为潜变量时，EM 算法是一种广泛用于最大似然估计的方法。在 E 步中，必须根据隐状态的平滑分布，计算完整数据对数似然（Q 函数）的期望。我们的主定理允许通过蒙特卡罗模拟近似这一期望，从而即使在复杂环境中也能实现高效估计。

它还可以与线性化技术结合应用于非线性模型，例如扩展卡尔曼滤波器（EKF）。例如，在对系统动力学进行线性化并生成隐状态的样本路径后，无需进一步对似然中的观测（传递）函数进行线性化，即可近似 Q 函数。

- **同时置信带的构造**。不同于由平滑协方差矩阵导出的点态置信区间，我们的逐路径抽样方法能够为隐过程构造同时置信带，从而确保在整个时间区间内的覆盖率。这对于检测随时间变化的结构偏差或阈值超越尤其有用。

关于(3.9)的引理

令 C_t 为连续函数 g 的空间：$[0,t] \to \mathbb{R}^{d_2}$，并令 $\mathcal{C}_t$ 为柱状 σ-代数。

引理 A.1。设 $\{H_t\}_{t\geq 0}$ 为一个 d_1维连续可测过程，并记 $\tilde{V}_s = \int_0^s b(s)\, dV_s$. 则存在一个可测函数$\Phi_t : (C_t \times C_t, \mathcal{C}_t \otimes \mathcal{C}_t) \to \mathbb{R}$ 使得

$$\Phi_t(H_\cdot, \tilde{V}_\cdot) = \int_0^t H_s d\tilde{V}_s \quad a.s., \tag{A.1}$$

并且对于 μ_H-几乎处处的 $h \in C_t$

$$\Phi_t(h, \tilde{V}_\cdot) = \int_0^t h(s) d\tilde{V}_s \quad a.s. \tag{A.2}$$

这里，$H_\cdot$ 和 $V_\cdot$ 分别表示 $\{H_s\}_{0\leq s\leq t}$ 和 $\{V_s\}_{0\leq s\leq t}$的路径，而 μ_H 是 $H_\cdot$ 在 C_t上的分布。

证明。首先，我们可以取一个序列 $\{n(k)\}_{k\in\mathbb{N}}$ 使得

$$\sum_{i=1}^{n(k)} H_{t/2^{n(k)-1}}^\top (\tilde{V}_{t/2^{n(k)}} - \tilde{V}_{t/2^{n(k)-1}}) \xrightarrow{\text{a.s.}} \int_0^t H_s d\tilde{V}_s. \tag{A.3}$$

现在我们引入一列可测函数 Φ_t^k： $C_t \times C_t \to \mathbb{R}$，其中$k \in \mathbb{N}$，定义如下：

$$\Phi_t^k(h,w) = \sum_{i=1}^{n(k)} h\left(\frac{t}{2^{n(k)-1}}\right)^\top \left(w\left(\frac{t}{2^{n(k)}}\right) - w\left(\frac{t}{2^{n(k)-1}}\right)\right).$$

接着我们定义 Φ_t： $C_t \times C_t \to \mathbb{R}$，定义如下:

$$\Phi_t(h,w) = \begin{cases} \lim\limits_{k\to\infty} \Phi_t^k(h,w) & \text{if } \Phi_t^k(h,w) \text{ converges,} \\ 0 & \text{otherwise.} \end{cases}$$

由于每个 Φ_t^k都是可测的， Φ_t 也可测，因此（A.1）可直接由 Φ_t的构造得出。

我们来验证（A.2）。若令 $A = \{(h,w) \in C_t \times C_t; \Phi_t^k(h,w)$ 收敛到 $\Phi_t(h,w)\}$，那么（A.3）意味着

$$\int_{C_t \times C_t} 1_{A^c}(h,w) d\mu_{(H,\tilde{V})}(h,w) = 0,$$

~其中 $\mu_{(H,\tilde{V})}$ 是 $(H_\cdot, V_\cdot)$ 的分布。由于 $H_\cdot$~与 $V_\cdot$ 相互独立，这可以改写为

$$\int_{C_t} \int_{C_t} 1_{A^c}(h,w) d\mu_{\tilde{V}}(w) d\mu_H(h) = 0,$$

因此，对于 μ_H-几乎处处的 $h \in C_t$，有

$$\int_{C_t} 1_{A^c}(h,w) d\mu_{\tilde{V}}(w) = 0.$$

由此可知，对于 μ_H-几乎必然的 $h \in C_t$，随机变量序列

$$\Phi_t^k(h, W_\cdot) = \sum_{i=1}^{n(k)} h(t/2^{n(k)-1})^\top (\tilde{V}_{t/2^{n(k)}} - \tilde{V}_{t/2^{n(k)-1}}) \tag{A.4}$$

~几乎必然收敛到 $\Phi_t(h, V_\cdot)$。另一方面，由于 $\{H_s\}_{0\le s\le t}$是连续的，我们可以假设 h 是连续的。对于这样的 h， （A.4）右端依概率收敛到 $\int_0^t h(s)\, d\tilde{V}_s$，从而得到（A.2）。

□

B 引理 3.2和（3.16）的证明

利用（3.24），可得

$$E_Q[z_t(y)] = E\left[\exp\left((\xi_{t;t}^0)^\top y(t) - \int_0^t y(s)^\top d\xi_{s;t}^0 + \frac{1}{2}\int_0^t b(s)^\top \phi(s;t) b(s) ds + \frac{1}{2} X_0^\top \phi(0;t) X_0\right)\right]. \tag{B.1}$$

此外，将分解式（3.28）代入（B.1），可写为

$$
\begin{aligned}
&E_Q[z_t(y)] = \exp\left(\frac{1}{2}\int_0^t b(s)^\top \phi(s;t) b(s) ds\right) \\
&\times E\Bigg[\exp\Bigg(X_0^\top \alpha(t;t)^\top y(t) \\
&- \int_0^t y(s)^\top \left\{a(s) + b(s)b(s)^\top \phi(s;t)\right\} \alpha(s;t) ds X_0 + \frac{1}{2} X_0^\top \phi(0;t) X_0\Bigg)\Bigg] \\
&\times E\left[\exp\left(\bar{\xi}_{t;t}^\top y(t) - \int_0^t y(s)^\top d\bar{\xi}_{s;t}\right)\right].
\end{aligned} \tag{B.2}
$$

对于右侧的第一个期望，令 $X_0 = \mu + \Sigma^{\frac{1}{2}} U$，其中 U 是标准正态随机变量，$\mu \in \mathbb{R}^{d_1}$，且 $\Sigma \in M_{d_1}(\mathbb{R})$。于是，通过直接计算可得

$$
\begin{aligned}
&E\Bigg[\exp\Bigg(X_0^\top \alpha(t;t)^\top y(t) \\
&- \int_0^t y(s)^\top \left\{a(s) + b(s)b(s)^\top \phi(s;t)\right\} \alpha(s;t) ds X_0 + \frac{1}{2} X_0^\top \phi(0;t) X_0\Bigg)\Bigg] \\
=&\sqrt{\det|I_{d_1} - \Sigma^{\frac{1}{2}} \phi(0;t) \Sigma^{\frac{1}{2}}|} \exp\left(\frac{1}{2} \beta(t;y)^\top \Sigma^{\frac{1}{2}} (I_{d_1} - \Sigma^{\frac{1}{2}} \phi(0;t) \Sigma^{\frac{1}{2}})^{-1} \Sigma^{\frac{1}{2}} \beta(t;y)\right) \\
&\times \exp\Bigg(\mu^\top \alpha(t;t)^\top y(t) \\
&- \int_0^t y(s)^\top \left\{a(s) + b(s)b(s)^\top \phi(s;t)\right\} \alpha(s;t) ds \mu + \frac{1}{2} \mu^\top \phi(0;t) \mu\Bigg).
\end{aligned} \tag{B.3}
$$

$$
\tag{B.4}
$$

这里使用了 $\phi(0;t)$ 的负半定性。

对于第二个期望，由于 $\xi_t^\top {}_{;t} y(t) = \int_0^t y(s)^\top d\xi_{s;t}$ 是均值为 0 的高斯随机变量，因此可得

$$
\begin{aligned}
&E\left[\exp\left(\bar{\xi}_{t;t}^\top y(t) - \int_0^t y(s)^\top d\bar{\xi}_{s;t}\right)\right] = \exp\left(\frac{1}{2} E\left[\left(\bar{\xi}_{t;t}^\top y(t) - \int_0^t y(s)^\top d\bar{\xi}_{s;t}\right)^2\right]\right) \\
=&\exp\Bigg(\frac{1}{2} y(t)^\top E[\bar{\xi}_{t;t} \bar{\xi}_{t;t}^\top] y(t) \\
&+ \frac{1}{2}\int_0^t y(s)^\top b(s) b(s)^\top y(s) ds - \int_0^t y(s)^\top b(s) b(s)^\top y(t) ds\Bigg).
\end{aligned} \tag{B.5}
$$

将（B.2）、（B.4）和（B.5）结合起来，我们得到引理 3.2。同样地

，（3.16）可由下式证明

$$
\begin{aligned}
&E_Q[X_{u_1}\otimes\cdots\otimes X_{u_n}z_t(y)]\\
=&E\Bigg[\xi^0_{u_1;t}\otimes\cdots\otimes\xi^0_{u_n;t}\exp\Bigg(\xi^{0\,\top}_{t;t}y(t)-\int_0^t y(s)^\top d\xi^0_{s;t}\\
&+\frac{1}{2}\int_0^t b(s)^\top\phi(s;t)b(s)ds+\frac{1}{2}X_0^\top\phi(0;t)X_0\Bigg)\Bigg]\\
=&\exp\left(\frac{1}{2}\int_0^t b(s)^\top\phi(s;t)b(s)ds\right)E\Big[\{\alpha(u_1;t)X_0+\bar{\xi}_{u_1;t}\}\otimes\cdots\otimes\{\alpha(u_1;t)X_0+\bar{\xi}_{u_1;t}\}\\
&\times\exp\Bigg(X_0^\top\alpha(t;t)^\top y(t)\\
&-\int_0^t y(s)^\top\left\{a(s)+b(s)b(s)^\top\phi(s;t)\right\}\alpha(s;t)dsX_0+\frac{1}{2}X_0^\top\phi(0;t)X_0\Bigg)\\
&\times\exp\left(\bar{\xi}^\top_{t;t}y(t)-\int_0^t y(s)^\top d\bar{\xi}_{s;t}\right)\Big].
\end{aligned}
$$

致谢

作者谨向 N. Yoshida 致以诚挚谢意，感谢他给予的宝贵建议和富有见地的讨论。

资金支持

本研究得到日本科学技术振兴机构 CREST（JP-MJCR2115）、JSPS KAKENHI（项目编号 JP24KJ0667）以及理化学研究所特别博士后研究员计划的资助。

参考文献

B. D. O. Anderson、A. N. Bishop、P. D. Moral 和 C. Palmier。反向非线性平滑扩散。《概率论及其应用》，66(2):245–262，2021年。A. Bain 和 D. Crisan。随机滤波基础，第 3 卷。Springer，2009年。M. Bello、A. Willsky、B. Levy 和 D. Castanon。平滑误差动力学及其在平滑与制图问题求解中的应用。*IEEE* 信息论汇刊，32(4):483–495，1986年。G. H. Bierman。离散测量下的固定区间平滑。《国际控制杂志》，8(1):65–75，1973年。数字对象标识符：10.1080/ 00207177308932487。

A. E. Bryson 和 M. C. Frazier。线性与非线性系统的平滑。技术报告 TDR-63-119，航空系统部，赖特-帕特森空军基地，1963年。T. Kailath。最小二乘估计的新息方法——第一部分：加性白噪声下的线性滤波。《*IEEE*自动控制汇刊》，13(6):646–655，1968年。数字对象标识符：10.1109/TAC.1968.1099025。R. E. Kalman。线性滤波与预测问题的一种新方法。《基础工程学杂志》，82(1):35–45，1960年。R. E. Kalman 和 R. S. Bucy。线性滤波与预测理论的新成果。《基础工程学杂志》，83(1):95–108，1961年。R. S. Liptser 和 A. N. Shiryaev。随机过程统计学：*I.* 一般理论。Springer，2001a。R. S. Liptser 和 A. N. Shiryaev。随机过程统计学 *II*：应用。Springer，2001b。D. Nualart。Malliavin演算及相关主题，2006年。J. E. Potter。统计滤波理论中的一个矩阵方程。技术报告，1965年。H. E. Rauch、F. Tung 和 C. T. Striebel。线性动态系统的最大似然估计。《*AIAA*期刊》，3(8):1445–1450，1965年。M. Rutkowski。卡尔曼-布西平滑估计公式的一个简单证明。《统计与概率快报》，17(5):377–385，1993年。国际标准连续出版物号 0167- 7152。S. S¨arkk¨a。贝叶斯滤波与平滑。数学统计学会教材丛书。剑桥大学出版社，2013年。H. Weinert 和 U. Desai。互补模型与固定区间平滑。《*IEEE*自动控制汇刊》，26(4):863–867，1981年。