مسئله پالایش در معادلات دیفرانسیل تصادفی و کاربرد آن در برآوردیابی

مهدی شمس،[۱] غلامرضا حسامیان[۲]

تاریخ دریافت: ۱۳۹۶/۱۱/۲۶
تاریخ پذیرش: ۱۳۹۸/۷/۵

چکیده:

در این مقاله پس از معرفی انتگرال تصادفی، روشی برای حل مسئله پالایش مطرح می‌شود که به حل دو معادله دیفرانسیل تصادفی (سیستم و مشاهده) از دیدگاه ایتو منجر می‌شود. همچنین با ذکر چند مثال، کاربردهایی از مسئله پالایش نظیر مشاهدات پراغتشاش ناشی از فرایندهای آمیخته ثابت و حرکت براونی، برآوردیابی پارامتر جامعه و حل یک معادله دیفرانسیل بر حسب جریان در یک مدار و نیروی الکترومغناطیس نوسانی بیان می‌شوند.

واژه‌های کلیدی: انتگرال تصادفی، انتگرال ایتو، معادله دیفرانسیل تصادفی، پالایش، برآورد.

۱ مقدمه

گسترهٔ چشمگیر مسائل طبیعی، اجتماعی و علوم زیستی ناشی از پیشرفت علم حساب در زمینه نظریهٔ توابع روی متغیرهای تصادفی حقیقی از زمان نیوتن و لیبنیتز است. از نخستین اجزای علم حساب مشتق‌گیری برای توصیف نرخ تغییرات، انتگرال‌گیری برای عبور حد در مجموع‌های تقریبی و قضیه‌های اساسی حساب است که تمام این موارد ناشی از مفهوم متداول معادلات دیفرانسیل و کاربردهای این معادلات در مدل‌بندی پدیده‌های طبیعی است که این موضوع نشان‌دهندهٔ توان علم حساب در کاربردهای عملی است.

حساب تصادفی، معادلات دیفرانسیل معمولی را با فرایندهای تصادفی پیوسته در می‌آمیزد و از این‌که بیشتر این فرایندها مانند حرکت براونی مشتق‌پذیر نیستند، حساب تصادفی از حساب معمولی متمایز می‌شود.

اصطلاح معادلات دیفرانسیل تصادفی[۳] برای اوّلین بار توسط برنشتاین [۴، ۱۵] در مطالعهٔ حدی یک دنباله از زنجیرهای مارکوف[۴] در یک طرح تفاضلی تصادفی مطرح شد. وی تنها به توزیع‌های فرایند حدی علاقه‌مند بود. گیهمن [۸، ۹، ۱۰] نظریهٔ معادلات دیفرانسیل تصادفی را، به‌طور مستقل از تحقیقات ایتو [۱۳، ۱۴] گسترش داد و شرایط وجود و یکتایی را کامل نمود. بعد از تحقیقات ایتو و گیهمن در پی توجه به نظریهٔ معادلات دیفرانسیل تصادفی، نتایج چشم‌گیری در کاربردهای عملی حاصل شد. ماهیت طبیعی و شهودی این نظریه در پژوهش‌های فیزیکی منجر به توجه ویژه پژوهشگران در بین علوم کاربردی شده است.

معادلات دیفرانسیل تصادفی، شامل معادله‌ای بر حسب بردار حالت X_t (که تابعی از پارامتر با مقادیر حقیقی زمان t است) و مشتق‌های آن هستند. به‌عنوان مثال مدل رشد ساده زیر

$$\frac{dX_t}{dt} = \lambda X_t,$$

بیان‌کنندهٔ این مفهوم است که متغیر زمان X_t در فاصله $[t, t+dt]$ متناسب با حالتی در این نقطه از زمان است که از مثال‌های واقعی این مدل می‌توان به رشد جوامع یا میرایی اشعه‌های

[۱] هیأت علمی گروه آمار، دانشگاه کاشان، کاشان، ایران
[۲] هیأت علمی گروه آمار، دانشگاه پیام نور شهرکرد، شهرکرد، ایران

[۳] Stochastic differential equations

[۴] Markov chain

رادیواکتیو اشاره کرد. جواب معادله بالا با شرط اوّلیه $X_{t_\circ}$، به‌صورت زیر است:

$$X_t = \exp\left[\lambda\left(t-t_\circ\right)\right] X_{t_\circ}$$

.

در علوم اجتماعی، معادلات دیفرانسیل تصادفی نظیر زیر

$$\frac{dX_t}{dt} = \lambda X_t + kW_t$$

کاربرد دارند که W_t یک فرایند نوفه سفید گاوسی با میانگین صفر و تابع اتوکوواریانس زیر است:

$$\begin{aligned}\gamma(t-s) &= E(W_tW_s)\\ &= \delta(t-s)\\ &= \begin{cases} ۱ & t=s\\ ۰ & t\neq s\end{cases}\end{aligned}$$

این حقیقت نشان می‌دهد فرایند W_t فقط در تمام فواصل کوتاه زمانی خودهمبسته‌اند. جواب معادله دیفرانسیل تصادفی اخیر برابر با عبارت زیر است:

$$X_t = \mathrm{e}^{\lambda(t-t_\circ)}X_{t_\circ} + k\int_{t_\circ}^{t}\mathrm{e}^{\lambda(t-s)}W_s ds,$$

در پی آن می‌توان جواب را بر حسب $W_s ds = dB(s)$ نوشت که در آن B_s فرایند پیوسته مشتق‌ناپذیر وینر (یا حرکت براونی) است. بنا بر این به‌منظور اجتناب از مشتق‌گیری از یک فرایند مشتق‌ناپذیر می‌نویسند،

$$dX_t = \lambda X_t dt + kdB_t$$

که جواب آن یعنی،

$$X_t = \mathrm{e}^{\lambda(t-t_\circ)}X_{t_\circ} + k\int_{t_\circ}^{t}\mathrm{e}^{\lambda(t-s)}dB_s$$

در حالتی که شرط اوّلیه گاوسی یا ثابت باشد، یک فرایند گاوسی است.

در ابتدا برخی مفاهیم مقدماتی تعریف می‌شوند.

تعریف ۱.۱. *[۷] خانواده $\mathcal{F}$ از زیرمجموعه‌های Ω را یک σ-جبر[۵] روی Ω گویند، هرگاه:*

الف) $\varnothing \in \mathcal{F}$.

ب) اگر $A \in \mathcal{F}$، آن‌گاه $A^c \in \mathcal{F}$.

ج) برای هر $A_۱, A_۲, \ldots \in \mathcal{F}$، $\bigcup_{i=۱}^{\infty} A_i \in \mathcal{F}$.

دنباله‌ای از متغیرهای تصادفی $\{X_t\}_{t\geq ۰}$ را یک فرایند تصادفی[۶] گویند. توجه کنید که برای هر $t \geq ۰$، تابع $w \to X_t(w)$ که $w \in \Omega$ یک متغیر تصادفی است و برای هر $w \in \Omega$، تابع $t \to X_t(w)$ که $t \geq ۰$ را یک مسیر[۷] از X_t نامند. فرایند $\{X_t\}_{t\geq ۰}$ را نوفه سفید[۸] (اغتشاش خالص) گویند، هرگاه:

الف) برای $t_۱ \neq t_۲$، $X_{t_۱}$ و $X_{t_۲}$ مستقل باشند.

ب) $\{X_t\}_{t\geq ۰}$ مانا[۹] باشد، یعنی توزیع توأم $(X_{t_۱+t}, \ldots, X_{t_k+t})$ به t بستگی نداشته باشد.

ج) برای هر t، $E(X_t) = ۰$ [۲۰].

تعریف ۲.۱. *[۱۲] خانواده صعودی $\{\mathcal{F}_t\}_{t\geq ۰}$ از زیرσ-جبرهای $\mathcal{F}$ (یعنی برای هر $۰ \leq s < t$، $\mathcal{F}_s \subset \mathcal{F}_t$) $\mathcal{F}_s \subset \mathcal{F}$ را یک پالایش[۱۰] گویند.*

تعریف ۳.۱. *[۱۸] فرایند $f(t,w) : [۰,\infty)\times\Omega \to \mathbb{R}^n$ را نسبت به پالایش $\mathcal{F}_t$، $\mathcal{F}_t$- سازوارشده[۱۱] گویند، هرگاه برای هر $t \geq ۰$، تابع $w \to f(t,w)$، $\mathcal{F}_t$-اندازه‌پذیر باشد.*

---

[۵] σ-Algebra

[۶] Stochastic process

[۷] Path

[۸] White noise

[۹] Stationary

[۱۰] Filtering

[۱۱] F_t-Adapted

خانواده توابع $f(t,w):[0,\infty)\times\Omega\to\mathbb{R}$ که در شرایط زیر صدق کند را با $v=v(S,T)$ نمایش می‌دهند:

الف) $f(t,w)$، $(t,w)\to f(t,w)$، $\mathcal{B}\times\mathcal{F}$-اندازه‌پذیر باشد که در آن $\mathcal{B}$، σ-جبر بورل روی $[0,\infty)$ است.

ب) $f(t,w)$، $\mathcal{F}$-سازوارشده باشد.

ج) $E[\int_S^T (f(t,w))^2 dt]<\infty$ [۲۰].

تعریف ۴.۱. *[۱۲] فرایند تصادفی n-بعدی $\{X_t\}_{t\geq 0}$ را یک مارتینگل[۱۲] نسبت به پالایش $\{\mathcal{F}_t\}_{t\geq 0}$ است، هرگاه*

الف) X_t برای هر t، $\mathcal{F}_t$-اندازه‌پذیر باشد.

ب) برای هر t، $E|X_t|<\infty$.

ج) برای هر $s\geq t$، $E[X_s|\mathcal{F}_t]=X_t$.

مثال ۵.۱. *[۱۸] اگر $\{Y_i\}_{i\in\mathbb{N}}$ ها متغیرهای تصادفی مستقل با میانگین صفر باشند، $\{X_i\}_{i\in\mathbb{N}}$ یک مارتینگل است که در آن $X_i=\sum_{j=1}^{i} Y_j$.*

تعریف ۶.۱. *[۱۶] یک خانواده از فرایندهای تصادفی $\{B_t\}_{t\geq 0}$ را یک حرکت براونی[۱۳] (فرایند وینر[۱۴]) گویند، هرگاه*

الف) $B_0=0$.

ب) نموهای $B_{s_i+t_i}-B_{s_i}$ روی مجموعه‌های متناهی دلخواه از فواصل مجزای (s_i,s_i+t_i) متغیرهای تصادفی مستقل باشند.

ج) برای هر $t,s\geq 0$، $B_{s+t}-B_s$ دارای توزیع نرمال $N(0,t)$ باشد.

اوّلین تحقیقات صورت‌گرفته بر حرکت براونی به بچلیر [۹] در پی پژوهش بر نوسانات قیمت سهام نسبت داده شده است. بچیلر از واژه حرکت براونی استفاده نکرد، بلکه B_t را به‌صورت توصیفی برای نوسانات بازارکه روی قیمت دارایی X_t تأثیر می‌گذارد، معرفی کرد. او افزایش قیمت بی‌نهایت کوچک dX_t رامتناسب با نموهای dB_t در نظر گرفت، یعنی معادله دیفرانسیل تصادفی زیر را معرفی کرد:

$$dX_t=\sigma dB_t$$

که در آن σ یک ثابت مثبت است. از این رو دارایی با قیمت اوّلیه $X_0=x$ داری ارزش $X_t=x+\sigma B_t$ در زمان t هست. ایرادی که به این مدل وارد هست این است که برای هر $t>0$، قیمت X_t با احتمال غیر صفر می‌تواند منفی باشد. برای رفع این ایراد قیمت نسبی dX_t/X_t به‌عنوان عکس العمل نوسانات بازار در نظر گرفته می‌شود که باید متناسب با dB_t باشد، یعنی

$$dX_t=\sigma X_t dB_t.$$

این معادله به‌خاطر این‌که B_t هیچ‌جا مشتق‌پذیر نیست، غیر قابل حل بوده است تا بالاخره ایتو در سال ۱۹۴۰ این مشکل را رفع کرد و نظریهٔ انتگرال‌های تصادفی و معادلات دیفرانسیل تصادفی شکل گرفت. از این رو معادله بالا از دیدگاه ایتو به‌صورت زیر بازنویسی می‌شود:

$$X_t=x+\sigma\int_0^t X_t dB_t$$

که انتگرال سمت راست تساوی به انتگرال تصادفی ایتو معروف است که آن در ادامه معرفی خواهد شد. انیشتین [۵، ۶] چگالی انتقال برای حرکت براونی در نظریه مولکولی-جنبشی[۱۵] گرما را محاسبه کرد. مطالعه‌های عمیق ریاضی روی حرکت براونی توسط وینر [۲۲، ۲۳] انجام شده است. دیدگاه انتگرال تصادفی نسبت به حرکت براونی اوّلین بار توسط پالی و دیگران [۲۱] برای انتگرال‌های غیر تصادفی مطرح شد و پس از آن ایتو [۱۳، ۱۴] آن را تعمیم داد. وجه تسمیهٔ حرکت براونی از روی حرکت‌های نامنظم دانه‌های گرده معلق در آب الهام شده است که

[۱۲] Martingle

[۱۳] Brownian motion

[۱۴] Weiner process

[۱۵] Molecular-leinetic

[۱۶] Robert brown

این تحقیقات توسط رابرت براون[16] زیست شناس معروف در سال ۱۸۲۸ انجام شده است. این حرکت‌های تصادفی منسوب به نوسانات گرده است که توسط مولکول‌های آب تولید می‌شود و نتیجه پراکندگی[17] یا انتشار[18] گرده در آب است. حرکت براونی در مطالعه ذره‌های بسیار کوچک معلق، مدل‌های قیمت سهام، اغتشاشات حرارتی در مدارهای الکتریکی، رفتارهای حدی در سیستم‌های صف و موجودی کالا و به‌طور کلّی اختلاهای تصادفی در فیزیک، شیمی، زیست‌شناسی، اقتصاد و سیستم‌های مدیریت کاربرد دارد.

نظریه حسابان تصادفی یا انتگرال‌های تصادفی و معادلات دیفرانسیل تصادفی توسط ریاضی‌دان ژاپنی ایتو در سال ۱۹۴۰ مطرح شد که در ذیل معرفی شده است.

تعریف ۷.۱. *[۲۰] برای تابع اندازه‌پذیر $f(t\ ,\ \omega)\in v(S,T)$ و حرکت براونی یک‌بعدی $B_t(\omega)$ با نقطه شروع مبدأ، انتگرال ایتو[19] به‌صورت*

$$\int_S^T f(t\ ,\ \omega)\ dB_t(\omega)=\lim_{n\to\infty}\ \int_S^T \varphi_n(t\ ,\ \omega)\ dB_t(\omega)$$

تعریف می‌شود که حد در L^2 و $\{\varphi_n\}$ دنباله توابع مقدماتی[20] (توابع ساده[21]) است به‌طوری که

$$\lim_{n\to\infty} E\int_S^T [f(t\ ,\ \omega)-\varphi_n(t\ ,\ \omega)]^2\ dt\ =\ 0.$$

انتگرال ایتو را می‌توان بر اساس انتگرال ریمان نیز توصیف کرد. حد اقل دو تفاوت عمده بین انتگرال ریمان[22] و انتگرال ایتو وجود دارد. اوّلین تفاوت نوع همگرایی هست. تقریب‌های انتگرال ریمان در $\mathbb{R}$ همگرا هستند، در صورتی که انتگرال ایتو توسط دنباله‌ای از متغیرهای تصادفی که در L^2 همگرا هستند تقریب زده می‌شود. دومین تفاوت این است که مجموع‌های ریمان انتگرال یک تابع $f:[0,T]\to\mathbb{R}$ را به‌صورت

$$\sum_{j=0}^{n-1} f(s_j)(t_{j+1}-t_j)$$

تقریب می‌زنند که در آن $0=t_0<t_1<\dots<t_n=T$ و برای هر j، s_j یک نقطه دلخواه در $[t_j,t_{j+1})$ هست. قابل توجه است که مقدار انتگرال ریمان به انتخاب نقاط $s_j\in[t_j,t_{j+1})$ بستگی ندارد. در حالت تصادفی، مجموع‌های تقریب زده شده به‌صورت

$$\sum_{j=0}^{n-1} f(s_j)\left(B_{t_{j+1}}-B_{t_j}\right)$$

هستند. در این حالت حد چنین تقریب‌هایی به انتخاب نقاط $s_j\in[t_j,t_{j+1})$ بستگی دارد. به‌عنوان نمونه قضیهٔ زیر ساختار انتگرال ریمان برای یک انتگرال تصادفی را مشخص می‌کند.

قضیهٔ ۸.۱. *[۱۹] (نمایش ریمان). برای هر تابع پیوسته $f:\mathbb{R}\to\mathbb{R}$ اگر بازه $[0,T]$ را برای $0\le i\le n$ با $t_i=\frac{iT}{n}$ افراز کنیم، در این صورت داریم:*

$$\lim_{n\to\infty}\sum_{i=1}^{n} f\left(B_{t_{i-1}}\right)\left(B_{t_i}-B_{t_{i-1}}\right)=\int_0^T f\left(B_s\right)dB_s$$

که همگرایی در L^2 است.

گزارهٔ ۹.۱. *فرض کنید $f,g\in v(0,T)$، $0\le S<U<T$ و c یک ثابت باشد. در این‌صورت*

الف) به‌صورت تقریباً همه جا داریم:

$$\int_S^T fdB_t=\int_S^U fdB_t+\int_U^T fdB_t.$$

ب) به‌صورت تقریباً همه جا داریم:

$$\int_S^T (cf+g)dB_t=c\int_S^T fdB_t+\int_S^T gdB_t.$$

ج) $E[\int_S^T fdB_t]=0$.

د) $\int_S^T fdB_t$، $\mathcal{F}_T$– اندازه‌پذیر است.

ه) $\int_0^t f(s,w)dB_s$ یک مارتینگل نسبت به $\mathcal{F}_T$ است.

[17] Dispersal

[18] Diffusion

[19] Ito integral

[20] Elementary functions

[21] Simple functions

[22] Riemann integral

قضیهٔ ۱۰.۱. *[۱۱] فرض کنید X_t در معادله دیفرانسیل تصادفی زیر صدق کند:*

$$dX_t = udt + vdB_t$$

که در آن v طوری انتخاب می‌شود که تقریباً همه جا برای هر $t \geq ۰$،

$$\int_{۰}^{t} v(s,\omega)^{۲} ds < \infty$$

و B_t یک مارتینگل نسبت به $\mathcal{F}_t$ است و u به‌گونه‌ای در نظر گرفته می‌شود که $\mathcal{F}_t$-سازوارشده باشد و تقریباً همه جا برای هر $t \geq ۰$، داریم:

$$\int_{۰}^{t} |u(s,\omega)|\, ds < \infty,$$

همچنین

$$g(t,x) \in C^{۲}([۰,\infty) \times \mathbb{R}).$$

(یعنی g روی $[۰,\infty) \times \mathbb{R}$ مشتقات مرتبه دوم پیوسته داشته باشد)، آن‌گاه فرایند $Y_t = g(t,X_t)$ در معادله دیفرانسیل تصادفی زیر صدق می‌کند:

$$dY_t = \frac{\partial g}{\partial t}(t,X_t)dt + \frac{\partial g}{\partial x}(t,X_t)dX_t + \frac{۱}{۲}\frac{\partial^{۲} g}{\partial x^{۲}}(t,X_t).(dX_t)^{۲}$$

که در آن

$$(dX_t)^{۲} = (dX_t).(dX_t).$$

با توجه به قوانین $dt.d.t = dt.dB_t = dB_t.dt = ۰$ و $dB_t.dB_t = dt$ محاسبه می‌شود.

مثال ۱۱.۱. *[۲۰] برای محاسبه انتگرال ایتوی زیر:*

$$\int_{۰}^{t} B_s dB_s$$

با انتخاب $X_t = B_t$ و $g(t,x) = \frac{۱}{۲}x^{۲}$، داریم:

$$Y_t = g(t,B_t) = \frac{۱}{۲}B_t^{۲}$$

و در پی آن با توجه به قضیهٔ ۱۰.۱ داریم:

$$\begin{aligned} dY_t &= \frac{\partial g}{\partial t}\partial t + \frac{\partial g}{\partial x}dB_t + \frac{۱}{۲}\frac{\partial^{۲} g}{\partial x^{۲}}(dB_t)^{۲} \\ &= B_t dB_t + \frac{۱}{۲}(dB_t)^{۲} \\ &= B_t dB_t + \frac{۱}{۲}dt \end{aligned}$$

و از این رو

$$d(\frac{۱}{۲}B_t)^{۲} = B_t dB_t + \frac{۱}{۲}dt$$

و در پی آن

$$\int_{۰}^{t} B_s dB_s = \frac{۱}{۲}B_t^{۲} - \frac{۱}{۲}t.$$

۲ مسئله پالایش

در مسئله پالایش دو فرایند تصادفی X_t (غیر قابل مشاهده) و Z_t (قابل مشاهده) وجود دارند که X_t با استفاده از $\{Z_s : ۰ \leq s \leq t\}$ برآورد می‌شود. لذا

$$\widehat{X}_t = P_{k_t}(X_t) = E(X_t \,|\, G_t)$$

که در آن

$$G_t = \sigma\{Z_s : ۰ \leq s \leq t\}.$$

همچنین $P_{k_t}(X_t)$ تصویر متعامد X_t روی

$$K_t = K(Z,t) = \{Y : \Omega \to \mathbb{R}^n \,|Y \in L^{۲}, Y \in G_t\}$$

است که در آن $\widehat{X}_t$ میانگین توان دوم خطاهای زیر را کمینه می‌کند:

$$E(X_t - \widehat{X}_t)^{۲} = \inf_{Y \in K_t} E(X_t - Y)^{۲}.$$

مثال ۱.۲. *[۲۰] فرض کنید X، $W_۱$، $W_۲$، ... متغیرهای تصادفی مستقل‌اند که برای هر j:*

$$EX = EW_j = ۰;\ EX^{۲} = a^{۲};\ EW_j{}^{۲} = m^{۲}.$$

هدف یافتن بهترین برآورد خطی X یعنی

$$\widehat{X}_t = P_{K_t}(X) = E(X_t|G_t)$$

بر اساس $\{Z_j : j \leq k\}$ است که در آن $Z_j = X + W_j$. در این راستا، قرار دهید:

$$L = L(Z,k) = \{c_۱Z_۱ + ... + c_kZ_k \,|c_۱, ..., c_k \in R\}.$$

با استفاده از فرایند متعامد سازی گرام اشمیت[۲۳] متغیرهای تصادفی $A_۱$ و $A_۲$ و... طوری تولید می‌شوند که:

[۲۳] Gram- schmidt orthogonalisation process

الف) $EA_iA_j = 0\,;\ i \neq j$.

ب) برای هر k: $L(A,\ k) = L(Z,\ k)$.

در این صورت برای هر k:

$$\widehat{X}_k = \sum_{j=1}^{k} \frac{EXA_j}{EA_j^{2}} A_j \qquad (1)$$

که تساوی در رابطهٔ (۱۰) می‌تواند به‌صورت رابطهٔ بازگشتی زیر نوشته شود، بنا بر این داریم:

$$\widehat{X}_k = \widehat{X}_{k-1} + \frac{EXA_k}{EA_k^{2}} A_k.$$

برای هر $j < k$، داریم:

$$EW_jZ_k = EW_j\,.\,EX + EW_j\,.\,EW_k = 0,$$

از این رو با توجه به اینکه $P_{j-1}(W_j) = 0$ می‌توان نوشت:

$$\begin{aligned} A_j &= Z_j - P_{j-1}(Z_j) \\ &= Z_j - P_{j-1}(X) - P_{j-1}(W_j) \\ &= Z_j - \widehat{X}_{j-1} \end{aligned}$$

و در پی آن:

$$\begin{aligned} EXA_j &= E\left(X(Z_j - \widehat{X}_{j-1})\right) \\ &= E\left(X(X - \widehat{X}_{j-1})\right) + E\left(XW_j\right) \\ &= E\left((X - \widehat{X}_{j-1} + \widehat{X}_{j-1})(X - \widehat{X}_{j-1})\right) + E\left(X\right)E\left(W_j\right) \\ &= E\left(X - \widehat{X}_{j-1}\right)^{2} + E\left(\widehat{X}_{j-1}(X - \widehat{X}_{j-1})\right) \\ &= E\left(X - \widehat{X}_{j-1}\right)^{2}. \end{aligned}$$

همچنین

$$\begin{aligned} EA_j^{2} &= E\left(X - \widehat{X}_{j-1} + W_j\right)^{2} \\ &= E\left(X - \widehat{X}_{j-1}\right)^{2} + EW_j^{2} + 2E\left((X - \widehat{X}_{j-1})W_j\right) \\ &= E\left(X - \widehat{X}_{j-1}\right)^{2} + m^{2} \end{aligned}$$

که تساوی آخر به‌خاطر استقلال X و W_j است. حال با جایگذاری در رابطهٔ (۱۰) معادله بازگشتی زیر به دست می‌آید:

$$\widehat{X}_k = \widehat{X}_{k-1} + \frac{E\left(X - \widehat{X}_{j-1}\right)^{2}}{E\left(X - \widehat{X}_{j-1}\right)^{2} + m^{2}} \left(Z_k - \widehat{X}_{j-1}\right)$$

که برای حل آن از تساوی زیر

$$U_k = \alpha_k \overline{Z}_k \in L(Z,\ k)$$

که در آن

$$\overline{Z}_k = \tfrac{1}{k}\sum_{i=1}^{k} Z_i$$

و

$$\alpha_k = \frac{a^{2}}{a^{2} + \frac{1}{k}m^{2}}$$

و همچنین

$$\begin{aligned} &E\left((X - U_k)Z_i\right) = E\left(X(X + W_i)\right) - \alpha_k E\left(\overline{Z}_k Z_i\right) \\ &= EX^{2} + EX\,.\,EW_i - \tfrac{\alpha_k}{k}\sum_{j=1}^{k} E[(X + W_j)(X + W_i)] \\ &= a^{2} - \tfrac{\alpha_k}{k}\sum_{j=1}^{k}\{EX^{2} + E(W_j + W_i)\,.\,EX + EW_jW_i\} \\ &= a^{2} - \tfrac{\alpha_k}{k}\{ka^{2} + m^{2}\} \\ &= 0 \end{aligned}$$

می‌توان نتیجه گرفت که

$$(X - U_k) \perp L(Z,\ k)$$

و از این رو U_k (و همچنین $\widehat{X}_k$) تصویر X_t روی $L(Z, k)$ است. لذا بنا بر یکتایی تصویر $\widehat{X}_k = U_k$ عبارت زیر نتیجه می‌شود:

$$\widehat{X}_k = \frac{a^{2}}{a^{2} + \frac{1}{k}m^{2}}\overline{Z}_k.$$

به‌عنوان دو حالت خاص برای k به اندازهٔ کافی بزرگ، $\widehat{X}_k \approx \overline{Z}_k$ و برای $m^{2} >> a^{2}$، $\widehat{X}_k \approx 0$.

۳ پالایش در معادله دیفرانسیل تصادفی

معادلهٔ دیفرانسیل تصادفی زیر را در نظر بگیرید:

$$\frac{dX_t}{dt} = b(t,\, X_t) + \sigma(t,\, X_t)\, dW_t \qquad (2)$$

که در آن دو تابع $b : [0\ ,\ T] \times \mathbb{R} \to \mathbb{R}$ و $\sigma : [0\ ,\ T] \times \mathbb{R} \to \mathbb{R}$ به‌ترتیب ضریب رانش[۲۴] و ضریب انتشار[۲۵] نامیده می‌شوند و در شرط زیر (که منجر به وجود و یکتایی جواب معادله فوق می‌شود) صدق می‌کنند:

[۲۴] Drift coefficient

[۲۵] Diffusion coefficient

برای هر $x, y \in \mathbb{R}$ و $t \in [\circ\ ,\ T]$، ثابت‌های c و d وجود دارند که:

$$|b(t\ ,\ x)| + |\sigma(t\ ,\ x)| \leq c(۱ + |x|)$$

$$|b(t\ ,\ x) - b(t\ ,\ y)| + |\sigma(t\ ,\ x) - \sigma(t\ ,\ y)| \leq d|x - y|$$

و W_t یک فرایند نوفه سفید است [۱۴]. با استفاده از تعبیر ایتو، معادله (۱۹) را می‌توان به‌صورت زیر نوشت:

$$dX_t = b(t\ ,\ X_t)\ dt + \sigma(t\ ,\ X_t)\ dU_t$$

که U_t حرکت براونی مستقل از $X_\circ$ با توزیع معلوم است [۲].

در حالت پیوسته مسئله پالایش فرض می‌شود مشاهدات Z_t به‌صورت معادله زیر هستند:

$$dZ_t = c(t\ ,\ X_t)\, dt + \gamma(t\ ,\ X_t)\, dV_t$$

$$Z_\circ = \circ$$

که در آن توابع c و γ در شرایط یکتایی صدق می‌کنند و V_t حرکت براونی مستقل از U_t و $X_\circ$ است.

برای سادگی حالتی در نظر گرفته می‌شود که معادلات سیستم و مشاهدات به‌صورت خطی هستند:

$$dX_t = F(t) X_t\, dt + C(t)\, dU_t,\ \ F(t), C(t) \in R \qquad (۳)$$

$$dZ_t = G(t)\, X_t dt + D(t)\, dV_t\,,\ \ G(t), D(t) \in R, Z_\circ = \circ \qquad (۴)$$

که در آن $X_\circ$ دارای توزیع نرمال و از $\{U_t\}$ و $\{V_t\}$ مستقل است.

قضیهٔ زیر رهیافتی برای حل مسئله پالایش رابطه‌های (۱) و (۵) است.

قضیهٔ ۱.۳ *[۱۵] جواب $\widehat{X}_t = E(X_t | G_t)$ برای مسئله پالایش خطی یک‌بعدی (۱) و (۵) در معادله دیفرانسیل تصادفی زیر صدق می‌کند، بنا بر این:*

$$d\widehat{X}_t = \left(F(t) - \frac{G^۲(t)S(t)}{D^۲(t)}\right)\widehat{X}_t dt + \frac{G(t)S(t)}{D^۲(t)}\, dZ_t;$$

$$X_\circ = EX_\circ \qquad (۵)$$

که

$$S(t) = E\left(X_t - \widehat{X}_t\right)^۲$$

در معادلهٔ غیر تصادفی زیر صدق می‌کند:

$$\frac{dS}{dt} = ۲F(t)S(t) - \frac{G^۲(t)}{D^۲(t)}S^۲(t) + C^۲(t)$$

$$S(\circ) = Var(X_\circ)$$

که این معادله به معادله ریکاتی[۲۶] مشهور است.

۴ کاربردها

در این بخش با ذکر چند مثال، کابردهایی از مسئله پالایش معرفی می‌شوند.

مثال ۱.۴. *(مشاهدات پراغتشاش از فرایند ثابت). فرض کنید*

$$dX_t = \circ\ (X_t = X_\circ\ \text{یا})\ ;\ \widehat{X}_\circ = EX_\circ = \circ\ ;\ Var(X_\circ) = a^۲$$

$$dZ_t = X_t dt + m dV_t\ ;\ Z_\circ = \circ$$

که در آن

$$F(t) = C(t) = \circ\ \ ;\ \ G(t) = ۱\ \ ;\ \ D(t) = m.$$

لذا معادلهٔ ریکاتی به‌صورت زیر است:

$$\frac{dS}{dt} = -\frac{۱}{m^۲}S^۲$$

که با $S(\circ) = a^۲$ حاصل می‌شود و برای حل آن با انتگرال‌گیری از طرفین داریم:

$$\int_\circ^t -\frac{۱}{S^۲}dS = \int_\circ^t \frac{۱}{m^۲}\, dt$$

و لذا

$$(S(t))^{-۱} - (S(\circ))^{-۱} = \frac{t}{m^۲}.$$

اکنون با جایگذاری رابطهٔ زیر

$$S(t) = \frac{a^۲m^۲}{m^۲ + a^۲t}$$

[۲۶] Riccati equation

در رابطهٔ (۷) داریم:

$$d\widehat{X}_t = -\frac{a^2}{m^2 + a^2 t}\widehat{X}_t\, dt \,+\, \frac{a^2}{m^2 + a^2 t}\, dZ_t$$
$$\widehat{X}_0 = EX_0 = 0.$$

که با ضرب دو طرف معادله در عبارت

$$exp\{\int_0^t \tfrac{a^2}{m^2+a^2 s}ds\}$$

به راحتی می‌توان نتیجه گرفت که:

$$d\left(\widehat{X}_t \exp\left\{\int_0^t \frac{a^2}{m^2+a^2 s}ds\right\} = \right.$$
$$\exp\left\{\int_0^t \frac{a^2}{m^2+a^2 s}ds\right\} \frac{a^2}{m^2+a^2 t}\, dZ_t.$$

سرانجام با قرار دادن رابهٔ زیر

$$\int_0^t \tfrac{a^2}{m^2+a^2 s}ds = \ln \tfrac{m^2+a^2 s}{m^2}$$

در معادلهٔ بالا نتیجه زیر حاصل می‌شود:

$$\widehat{X}_t = \tfrac{a^2}{m^2 + a^2 t} Z_t \ ; \ t \geq 0.$$

مثال ۲.۴. *(مشاهدات پراغتشاش از حرکت براونی).* فرض کنید

$$dX_t = dU_t \quad ; \quad EX_0 = Var(X_0) = 0$$

$$dZ_t = X_t dt + dV_t \ \ ; \ \ Z_0 = 0$$

که در آن

$$F(t) = 0 \,; \quad C(t) = G(t) = D(t) = 1$$

و لذا معادلهٔ ریکاتی به‌صورت زیر است:

$$\frac{dS}{dt} = -S^2 + 1$$

که با $S(0) = 0$ تبدیل می‌شود. حال با جایگذاری رابطهٔ زیر

$$S(t) = \frac{e^{2t} - 1}{e^{2t} + 1} = \tanh(t)$$

در رابطهٔ (۷) داریم:

$$d\widehat{X}_t = -\tanh(t)\widehat{X}_t\, dt \,+\, \tanh(t)\, dZ_t$$
$$\widehat{X_0} = 0$$

و یا

$$d\left(\cosh(t)\widehat{X}_t\right) = \sinh(t) dZ_t$$

و بالاخره:

$$\widehat{X}_t = \frac{1}{\cosh(t)} \int_0^t \sinh(s)\, dZ_s.$$

مثال ۳.۴. *(برآورد پارامتر).* هدف برآورد پارامتر θ بر اساس مشاهدات Z_t بر اساس مدل زیر است:

$$dZ_t = \theta\, G(t)\, dt + D(t)\, dB_t$$

که $G(t)$ و $D(t)$ توابع معلوم هستند. بنا بر این رابطهٔ (۱) به‌صورت $d\theta = 0$ خواهد بود ($X_t = \theta$ و $V_t = B_t$) و معادلهٔ ریکاتی برای عبارت زیر:

$$S(t) = E(\theta - \widehat{\theta}_t)^2$$

به‌صورت زیر است:

$$\frac{dS}{dt} = -\left(\frac{G(t)S(t)}{D(t)}\right)^2$$

و جواب این معادله دیفرانسیل برابر با

$$S(t) = \left(S_0{}^{-1} + \int_0^t G^2(s) D^{-2}(s) ds\right)^{-1}$$

است. از این رو داریم:

$$d\widehat{\theta}_t = \frac{G(t)S(t)}{D^2(t)} \left(dZ_t - G(t)\widehat{\theta}_t dt\right)$$

یا این‌که

$$d\left[\left(S_0{}^{-1} + \int_0^t G^2(s)D^{-2}(s)ds\right)\widehat{\theta}_t\right] =$$
$$\left(S_0{}^{-1} + \int_0^t G^2(s)D^{-2}(s)ds\right) d\widehat{\theta}_t + G^2(t)D^{-2}(t)\widehat{\theta}_t dt$$
$$= G(t)D^{-2}(t)dZ_t.$$

[۲۷] Maximum likelihood estimator

بنا بر این برآورد پارامتر θ به‌صورت زیر به دست می‌آید:

$$\widehat{\theta}_t = \frac{\widehat{\theta}_\circ S_\circ^{-1} + \int_\circ^t G(s)D^{-2}(s)dZ_s}{S_\circ^{-1} + \int_\circ^t G^2(s)D^{-2}(s)ds}.$$

توجه شود که این برآورد برای حالتی که $S_\circ^{-1} = \circ$ باشد بر برآوردگر ماکسیمم درست‌نمایی (MLE[۲۷]) پارامترθ منطبق می‌شود [۱۵].

مثال ۴.۴. معادلهٔ دیفرانسیل تصادفی

$$L\frac{dI(t)}{dt} + RI(t) = \sigma\xi(t)$$

که $I(t)$ جریان [۲۸] در یک مدار مقاومت القایی[۲۹] و $\xi(t)$ نیروی الکترومغناطیس نوسانی[۳۰] (نوفه سفید) است به‌صورت زیر

$$dI(t) = -\frac{R}{L}I(t)\,dt + \frac{\sigma}{L}dV(t)$$
$$I(\circ) = I_\circ$$

تبدیل می‌شود که در آن $V(t)$ فرایند وینر، R/L ضریب اصطکاک [۳۱] و σ/L ضریب انتشار است [۱۷].

توجه شود که جریان آغازی $I_\circ$ یک متغیر تصادفی با $E(I_\circ) < \infty$ و $Var(I_\circ) < \infty$ هست. برای حل این معادله با استفاده از فرمول ایتو داریم:

$$I(t) = e^{-\frac{R}{L}I_\circ} + \frac{\sigma}{L}\int_\circ^t e^{-\frac{R}{L}(t-s)}dV(s)$$

لذا میانگین و واریانس جریان به‌صورت زیر حاصل می‌شوند:

$$E\left[I(t)\right] = e^{-\frac{R}{L}I_\circ}\,E\left(I_\circ\right)$$
$$Var\left[I(t)\right] = e^{-\frac{2Rt}{L}}\,Var\left(I_\circ\right) + \frac{\sigma^2}{2LR}\left(1 - e^{-\frac{2Rt}{L}}\right).$$

حال اگر مشاهدات به‌صورت زیر

$$dZ(t) = I(t)\,dt + dU(t)$$

باشد، هدف پیدا کردن بهترین برآورد جریان یعنی $\hat{I}(t)$ تحت مشاهدات $\{Z(s)\}_{s \leq t}$ است. ملاحظه می‌شود که عبارت زیر

$$S(t) = E\Big(I(t) - \hat{I}(t)\Big)^2$$

در معادلهٔ ریکاتی زیر

$$\frac{dS(t)}{dt} = -\frac{2R}{L}S(t) - S^2(t) + \frac{\sigma^2}{L^2}$$
$$S(\circ) = EI_\circ^{\,2} = A^2$$

صدق می‌کند. بنا بر این با جایگذاری جواب معادله بالا در معادله زیر

$$d\hat{I}(t) = \left(-\frac{R}{L} - S(t)\right)\hat{I}(t)dt + S(t)\,dZ(t)$$
$$\hat{I}(\circ) = EI_\circ = \circ$$

به راحتی معادله دیفرانسیل زیر

$$d\hat{I}(t) = \left(-\frac{R}{L} - \gamma\frac{1 + e^{C-2\gamma t}}{1 - e^{C-2\gamma t}}\right)\hat{I}(t)\,dt + \gamma\frac{1 + e^{C-2\gamma t}}{1 - e^{C-2\gamma t}}\,dZ(t)$$
$$\hat{I}(\circ) = EI_\circ = \circ$$

حاصل می‌شود که در آن

$$\gamma = \frac{\sqrt{R^2 + \sigma^2}}{L} \quad ; \quad C = \ln\left|\frac{A^2 - \frac{R}{L} - \gamma}{A^2 - \frac{R}{L} + \gamma}\right|$$

و از این رو:

$$\hat{I}(t) \approx \hat{I}(\circ)\,e^{-\gamma t} + (-\tfrac{R}{L} + \gamma)\,e^{-\gamma t}\int_\circ^t e^{\gamma s}\,dZ(s) = (-\tfrac{R}{L} + \gamma)\,e^{-\gamma t}\int_\circ^t e^{\gamma s}\,dZ(s).$$

---

[۲۸] Current

[۲۹] Inductance-resistance circuit

[۳۰] Electromagnetic force fluctuating

[۳۱] Coefficient of friction

مراجع

[1] Bachelier, L. (1900). Théorie de la spéculation. *In Annales Scientifiques de l'École Normale Supérieure*, **17**, 21-86.

[2] Bain, A., and Crisan, D. (2007). *Fundamentals of Stochastic Filtering*, Springer, New York.

[3] Bernstein, S. (1934). Principes de la théorie des équations différentielles stochastiques, *Trudy Fiziqeskogo Matematiqeskogo Instituta Steklova*, **5**, 95-124.

[4] Bernstein, S. (1938). Équations différentielles stochastiques. *Actualités Scientifiques et Industrielles*, **738**, 5-31.

[5] Einstein, A. (1905). On the movement of small particles suspended in a stationary liquid demanded by the molecular-kinetic theory of head, *Annalen der Physik*, **17**, 549-560.

[6] Einstein, A. (1956). *Investigations on the Theory of the Brownian Movement*, Courier Corporation.

[7] Feller, W. (1971). *An Introduction to Probability Theory and Its Applications*, **II**, Second Edition, John Wiley and Sons, New York.

[8] Gihman, I. I. (1947). A method of constructing random processes (in Russian), *Doklady Akademii Nauk SSSR* **58**, 961-964.

[9] Gihman, I. I. (1950a). Certain differential equations with random functions (in Russian), *Mathematische Zeitschrift*, **2**(**3**), 45-69.

[10] Gihman, I. I. (1950b). On the theory of differential equations of random processes (in Russian), *Mathematische Zeitschrift*, **2**(**4**), 37-63.

[11] Grimmett, G., and Stirzaker, D. (2001). *Probability and Random Processes*, Oxford.

[12] Gut, A. (2005). *Probability: A Graduate Course*, Springer, New York.

[13] Ito, K. (1942). Differential equations determining Markov processes (in Japanese), *Zenkoku Shijo Sugaku Danwakai*, **244**(**1077**), 1352-1400.

[14] Ito, K. (1944). Stochastic integral. *Proceedings of the Imperial Academy Tokyo*, **20**(**8**), 519-554.

[15] Jaswinski, A. H. (1970). *Stochastic Processes and Filtering Theory*, Academic Press, New York.

[16] Karatzas, I., and Shreve, S. E. (1991). *Brownian Motion and Stochastic Calculus*, Second Edition, Springer-Verlag, Berlin.

[17] Kolářová, E. (2007). An Application of the Kalman-Bucy filter to electrical Circuits. *Proceedings of the 6th mathematical workshop*, 71-72.

[18] Koralov, L. B., and Sinal, Y. G. (2007). *Theory of Probability and Random Processes*, Princeton.

[19] Liptser, R. S., and Shiryaev, A. N. (1978). *Statistics of Random Processes*, **II**, Springer-Verlag, New York.

[20] Oksendal, B. (2000). *Stochastic Differential Equations, An Introduction with Applications*, Fifth Edition, Springer-Verlag, New York.

[21] Paley, R. E. A. C., Weiner, N., and Zygmund, A. (1933). Notes on random functions. *Mathematische Zeitschrift*, **37**(**1**), 647-668.

[22] Weiner, N. (1923). Differential space. *Journal of Mathematics and Physics*, **2**, 131-174.

[23] Weiner, N. (1924). Un problème de probabilité dénombrables. *Bulletin de la Société Mathématique de France*, **52**, 569-578.