

## بررسی آماری دو فرایند انتشار روی چنبره و کاربرد آن‌ها

میلاذ رحیمی، موسی گل‌علی‌زاده\*:

دانشگاه تربیت مدرس، دانشکده علوم ریاضی، گروه آمار

دریافت ۹۲/۴/۴

پذیرش ۹۳/۶/۱۰

### چکیده

محققان در زمینه‌های مختلف علوم از جمله علوم زیستی، به فرایندهای انتشار مثل حرکت براونی و فرایند اورنشتاین-اولن‌بک که کلاسی از فرایندهای تصادفی هستند توجه ویژه دارند. در بررسی چنین فرایندهایی معمولاً فرض می‌شود مشاهدات حاصل از آن‌ها در فضاهای اقلیدسی قرار دارند. اما در بعضی از پدیده‌های فیزیکی، شیمیایی و زیستی داده‌هایی یافت می‌شوند که به دلایلی مثل تناوبی بودن مقادیری از فضاهای اقلیدسی به حساب نمی‌آیند. در نتیجه آن‌ها با مدل‌بندی‌های معمول ریاضی که برای فضاهای اقلیدسی وجود دارند بررسی نمی‌شوند. علاوه بر این، از نقطه نظر آمار، بررسی و تحلیل آن‌ها با استفاده از روش‌های مرسوم آمار خطی ممکن نیست. زاویه‌های دوسطحی که برای شناسایی، مدل‌بندی و پیش‌بینی ساختار اصلی پروتئین‌ها استفاده می‌شوند مثالی از این دست داده‌هاست. چون این زوایا مقادیری را روی چنبره نمایش می‌دهند، در نتیجه انتظار می‌رود مدل‌بندی مناسب آماری فرایندهای انتشار روی چنبره بتواند کمک شایانی به فعالیت‌های معطوف به شبیه‌سازی پویای مولکولی در پیش‌بینی ساختار اصلی پروتئین‌ها کند. در این مقاله، با استفاده از فاصله‌های ریمانی روی چنبره، معادلات دیفرانسیل تصادفی برای نمایش حرکت براونی و فرایند اورنشتاین-اولن‌بک روی این شکل هندسی به‌دست آورده می‌شود. سپس با محاسبه توزیع مانای فرایندهای بررسی شده و ارزیابی تعدادی از توزیع‌های غیراقلیدسی موجود، ارتباط نتایج حاصل با مفاهیم موجود در آمار غیرخطی برجسته خواهد شد.

واژه‌های کلیدی: فرایندهای انتشار، معادله دیفرانسیل تصادفی، توزیع‌های مانا، آمار غیرخطی.

### مقدمه

بررسی‌های تاریخی نشان می‌دهند که علاقه‌مندی دانشمندان به مدل‌بندی پدیده‌های تصادفی آن‌ها را متوجه فرایندهای تصادفی به‌ویژه فرایندهای مارکف کرده است. از میان آن‌ها فرایندهای انتشار<sup>۱</sup> نقش بسیار برجسته‌ای در مدل‌بندی الگوهای تصادفی موجود در پدیده‌های طبیعی دارند. یکی از مهم‌ترین و پرکاربردترین فرایندهای انتشار فرایند وینر یا حرکت براونی است. این فرایند که رابرت براون گیاه شناس اسکاتلندی در سال ۱۸۲۸ معرفی کرد برای توجیه ریاضی نحوه برخورد گرده گیاهان با مولکول‌های آب استفاده شد. جالب است که با وجود پیوستگی مسیرهای حرکت براونی، آن هیچ‌جا مشتق پذیر نیست [۱۳]. از این رو، کاربرد مستقیم آن در توصیف بعضی از پدیده‌های تصادفی ممکن نیست. به‌عنوان مثال، برای بررسی سرعت حرکت یک ذره،

\*نویسنده مسنول golalizadeh@modares.ac.ir

1. Diffusion

نمی‌توان مستقیماً از حرکت براونی استفاده کرد. برای رفع این مشکل، اورنشتاین و اولن‌بک فرایند معروف اورنشتاین-اولن‌بک ( $OU$ )<sup>۲</sup> را برای مدل‌بندی سرعت ذره معلق در جسم سیال که وضعیت آن در زمانی خاص با فرایند براونی توصیف می‌شود، ارائه کردند [۲۶]. این فرایند نیز در مدل‌بندی بسیاری از پدیده‌های تصادفی استفاده می‌شود.

فرض معمول در بررسی اکثر فرایندهای مرسوم تصادفی این است که داده‌های بررسی شده رخداد‌های آزمایش تصادفی از فضای اقلیدسی هستند. اما گاهی اوقات ماهیت داده‌های آماری به‌گونه‌ای است که نمی‌توان آن‌ها را برآمدی از یک فضای اقلیدسی در نظر گرفت. یکی از این موارد داده‌های سودار<sup>۳</sup> است که به مشاهدات با اندازه واحد در فضای  $\mathbb{R}^n$ ،  $n \geq 2$  اطلاق می‌گردد. مثال‌هایی از داده‌های سودار عبارتند از زاویه پرواز پرندگان، زاویه قرار گرفتن قطب‌نما و زوایای دوسطحی حاصل از موقعیت اتم‌های یک مولکول. بحث جامع آماری و ریاضی راجع به این‌گونه داده‌ها در [۱۹] موجود است. مدل‌بندی فرایندهای تصادفی داده‌های سوئی حوزه جدیدی از تحقیق است که اخیراً کاربرد آن در بررسی پدیده‌های متنوع، نظر محققان زیادی را به خود جلب کرده است. جالب این‌جاست که بدون توجه به کاربرد حرکت براونی در فضای نااقلیدسی، پوشیدا در دهه پنجاه به بررسی این حرکت در کره سه بعدی پرداخت [۳].

یکی از ابزارهای مفید برای توصیف تغییرات فرایندهای تصادفی معادله دیفرانسیل تصادفی (SDE)<sup>۴</sup> است. با این حال، حل این معادلات به سادگی یافتن جواب برای معادلات دیفرانسیل معمولی نیست. برای محاسبه انتگرال‌های مرتبط با SDE ایتو [۸] فرمولی را ارائه کرد که در منابع مرتبط با فرایندهای تصادفی به انتگرال تصادفی یا انتگرال ایتو<sup>۵</sup> معروف شد. لازم به ذکر است که انگیزه ایتو برای حل انتگرال‌های تصادفی ساخت فرایند انتشار با SDE بود [۱۷]. توجه شود که اگرچه حل تحلیلی SDEها می‌تواند گام مناسبی را برای دسترسی به پاسخ یک فرایند مهیا کند ولی در بسیاری از مواقع انجام اینکار ممکن نیست. در این حالات حل تقریبی این‌گونه معادلات از طریق گسسته سازی به‌دست می‌آید که روش‌های متنوعی برای حل عددی آن‌ها در منابعی مثل [۱۵] وجود دارد.

یکی از کاربردهای نوین آمار در پیش‌بینی ساختار پروتئین است. یکی از مرسوم‌ترین روش‌های موجود روشی به نام جمع‌آوری قطعه‌ای است [۲۳]. در مقابل آن، روشی وجود دارد که بر اساس مدل‌های احتمالی پایمیزی شده است [۶]. در این بین، استفاده از زاویه برای معرفی مکان هندسی اتم‌ها در زنجیره اصلی<sup>۶</sup> پروتئین مورد توجه خاصی قرار دارد. دلیل آن امکان پایمیزی مدل‌های احتمالی پیوسته و بهره‌برداری از توزیع‌های آماری غیراقلیدسی است. توجه کنید که ساختار بسیاری از پروتئین‌ها به‌گونه‌ای است که مشاهدات حاصل از آن‌ها می‌توانند به‌صورت داده‌های آماری روی سطح چنبره در نظر گرفته شوند. بررسی و مطالعه مقالات علوم زیستی که رفتارهای تصادفی بخشی از جفت زوایای پروتئین را مدنظر قرار دادند، نشان می‌دهد که مدل‌بندی آماری فرایندهای انتشار این پدیده زیستی از موضوعات نوین تحقیقاتی در حوزه علوم زیستی

2. Ornstein-Uhlenbeck

3. Directional

4. Stochastic Differential Equation (SDE)

5. Itô Integral

6. Backbone

به‌ویژه بیوانفورماتیک است. تا جایی که نویسندگان مقاله حاضر اطلاع دارند تنها منبع موجود که به این موضوع پرداخته [۲۵] است. برای دسته‌بندی بهتر، مطالب مقاله حاضر به‌صورت ذیل تدوین شده است.

در بخش بعد ابتدا مقدمه‌ای گذرا از فرایند انتشار و نحوه محاسبه توزیع‌های مانای آن بیان و سپس نحوه مدل‌بندی آن‌ها با SDE ارائه می‌شود. در ادامه با ارائه مقدمه‌ای از آمار در فضای ناقلیدسی و معرفی عملگر مشخصه فرایند انتشار، نحوه مدل‌بندی چنین فرایندهایی در فضاهای ناقلیدسی بررسی می‌شود. در پی آن SDE حرکت براونی روی چنبره به‌دست می‌آید. در انتها، به‌منظور مدل‌بندی زوایای دوسطحی پروتئین فرایند OU روی چنبره پیشنهاد و توزیع مانای آن محاسبه می‌شود.

### فرایند انتشار و توزیع مانای آن

از نقطه نظر ریاضی، فرایند مارکف زمان پیوسته که مسیرهای نمونه‌ای آن پیوسته باشد را فرایند انتشار می‌نامند [۱۲]. این فرایند در شکل‌گیری و ساخت بسیاری از فرایندهای دیگر به‌کار گرفته می‌شود. دلیل اصلی این موضوع را می‌توان به علت ویژگی‌های شناخته شده این فرایند شامل خاصیت‌های مارکفی، گاوسی، مارتینگل<sup>۷</sup> بودن و پیوستگی مسیرش دانست [۲۲]. معمولاً فرایندهای انتشار را با دو شاخصه معرفی می‌کنند که در اکثر منابع مرتبط با فرایندهای تصادفی مانند [۲۱] و [۲۴] به آن‌ها ضرایب رانش<sup>۸</sup> و انتشار می‌گویند.

چنان‌که اشاره شد یکی از مهم‌ترین فرایندهای انتشار، حرکت براونی (فرایند وینر) است. یکی از ویژگی‌های خاصش این است که مسیرهایش هیچ‌جا مشتق‌پذیر نیست. از این‌رو اگر حرکت براونی بیان‌گر وضعیت ذره معلق در آب در زمان  $t$  باشد، با مشتق‌گیری از آن نمی‌توان سرعت ذره را تعیین کرد. برای حل این مشکل از فرایند OU استفاده می‌شود. این فرایند را [۲۶] برای مدل‌بندی سرعت یک ذره معلق در جسم سیال<sup>۹</sup> که وضعیت آن در زمان  $t$  با فرایند براونی توصیف می‌شود، معرفی کرد. جزئیات ریاضی این فرایند در ادامه می‌آید.

اگر  $X(t) = X_t$  نشان‌دهنده فرایند OU و  $b(t, x) = -\beta x$  و  $\sigma^2(t, x) = \sigma^2$  به ترتیب ضرایب رانش و انتشار فرایند باشند آن‌گاه  $X_{t+s}$  به شرط  $X_s = x$ ، دارای توزیع نرمال با میانگین

$$E[X_{t+s} | X_s = x] = xe^{-\beta t}$$

و واریانس

$$\text{Var}[X_{t+s} | X_s = x] = \left(\frac{1 - e^{-2\beta t}}{2\beta}\right)\sigma^2$$

است. توجه شود که فرایند OU تنها فرایند انتشاری است که گاوسی و مانا است [۲].

یکی از مسائل مورد علاقه در بررسی فرایندهای تصادفی محاسبه توزیع مانای فرایند است که به‌صورت گذرا به مفهوم آن و همچنین روش محاسبه آن به این شرح است:

فرض کنید  $X_t$  فرایند انتشار همگن باشد و برای هر  $t > 0$  و هر  $x \in \mathcal{H}^n$ ، تابع چگالی انتقال  $p(x; t; y)$  موجود،  $b(t, x)$  و  $a = \sigma\sigma^T$  به ترتیب ضرایب رانش و انتشار فرایند باشند. اگر تابع  $\psi(x)$  با دامنه  $\mathcal{H}^n$  و

7. Martingale  
8. Drift  
9. Fluid

برد  $\mathcal{H}$  در معادله

$$0 = -\sum_{j=1}^n \frac{\partial}{\partial x_j} (b_j(t, x)\psi(x)) + \frac{1}{2} \sum_{j,k=1}^n \frac{\partial^2}{\partial x_j \partial x_k} (\sigma_{ij}(t, x)\psi(x)) \quad (1)$$

صدق کند آن‌گاه فرایند انتشار مربوط، مانا و توزیع مانای آن برابر  $\psi(x)$  خواهد بود [۲۴]. در حالت یک بعدی معادله صریحی برای توزیع مانای یک فرایند انتشار وجود دارد که تشریح روابط مربوط بدین صورت است:

فرض کنید  $X_t$  فرایند انتشار یک متغیره با فضای حالت به صورت بازه  $(r_1, r_2)$  باشد. توابع

$$s(y) = \exp\left(\int^y -\frac{2b(u)}{\sigma^2(u)} du\right)$$

$$S(x) = \int^x s(y) dy$$

$$m(x) = \frac{1}{\sigma^2(x)s(x)} \quad (2)$$

را در نظر بگیرید. بنا به [۱۳]، با حل رابطه (۱) توزیع مانای فرایند به صورت

$$\psi(x) = C_1 \frac{S(x)}{\sigma^2(x)s(x)} + C_2 \frac{1}{\sigma^2(x)s(x)} = m(x)[C_1 S(x) + C_2]$$

به دست می‌آید که در آن مقادیر ثابت  $C_1$  و  $C_2$  طوری تعیین می‌شوند که  $\psi(x)$  در شرایط تابع چگالی صدق کند. توجه شود که ممکن است تابع  $\psi(x)$  در نقاط انتهایی بازه  $(r_1, r_2)$  بی‌نهایت شود. در این حالت معمولاً ابتدا  $C_1$  برابر صفر قرار داده می‌شود و سپس مقدار  $C_2$  طوری محاسبه می‌شود که انتگرال  $\psi(x)$  روی بازه‌اش برابر

یک شود. به‌طور خلاصه، تابع چگالی مانای فرایند یکتا و برابر

$$\psi(x) = \frac{m(x)}{\int_{r_1}^x m(y) dy} = \frac{1}{\sigma^2(x)s(x) \int_{r_1}^x [\sigma^2(y)s(y)]^{-1} dy} \quad (3)$$

است. بحث جامع راجع به وضعیت نقاط بازه  $(r_1, r_2)$  و یکتایی تابع چگالی در [۱۳] موجود است.

یکی از نکات حائز اهمیت در مورد فرایندهای انتشار این است که می‌توان آن‌ها را با ضرایب رانش و انتشار به‌طور کامل تعیین کرد. علاوه بر این، تحت فرضیات مناسبی روی ضرایب رانش و انتشار، توزیع مانای فرایند ترکیب شده با ضرایب مربوطه در معادلات دیفرانسیل مشخصی صدق می‌کنند. این معادلات را معادلات پیش‌رو<sup>۱۰</sup> و پس‌روی کولموگوروف<sup>۱۱</sup> (به ترتیب KFE و KBE) می‌نامند. با این حال، حل مستقیم این معادلات و یافتن توزیع مانای فرایند بررسی شده جز در مواردی خاص ممکن نیست.

شناسایی و ساخت فرایندهای انتشار در فضاهای خاص مانند فضای ناقلیدسی اهمیت به‌سزایی دارد. یکی از روش‌های آسان برای ساخت فرایند انتشار استفاده از ماتریس تانسور و به‌کارگیری نظریه ایتو در معرفی فرایند با SDE است. با الهام از [۱۷]، جزئیات بیشتر این موضوع بدین صورت است:

### معادلات دیفرانسیل تصادفی و فرایند انتشار

از ریاضیات مقدماتی بیاد داریم که معادله دیفرانسیل، بیان‌گر رابطه بین یک تابع و مشتقات آن است. این معادلات برای مدل‌بندی بسیاری از پدیده‌های فیزیکی، مهندسی، زیست‌شناسی و غیره کاربرد دارند [۱۶]. به‌عنوان مثال معادله

10. Kolmogorov Forward Equation

11. Kolmogorov Backward Equation

$$\frac{dN}{dt} = a_t N_t \quad (۲)$$

یک معادله دیفرانسیل را نشان می‌دهد که در آن  $N_t$  معرف حجم نمونه در زمان  $t$  و  $a_t$  نرخ نسبی رشد در زمان  $t$  است. مانند بسیاری از مدل‌های آماری، مثل مدل‌های رگرسیونی و سری زمانی که در آنها برخی از متغیرهای بررسی شده تصادفی هستند، ضریب معادله (۴) را نیز می‌توان تصادفی در نظر گرفت. از این رو فرض کنید تغییرات  $a_t$  در معادله (۴) بدین صورت است:

$$a_t = r_t + \text{نوفه}$$

که رفتار "نوفه" در این تساوی کاملاً مشخص نیست و فقط توزیع احتمال آن معلوم است. با این فرض معادله (۴) به معادله دیفرانسیل تصادفی تبدیل می‌شود. بنا به [۲۰] معادله دیفرانسیلی که ضرایب آن متغیر تصادفی باشد را معادله دیفرانسیل تصادفی می‌گویند. اکنون می‌توان با لحاظ حرکت براونی به‌عنوان یک نوفه صورت کلی فرایندها را برحسب  $SDE$  نوشت.

فرض کنید  $B_t$  معرف یک حرکت براونی در زمان  $t$  است. در این صورت بنا بر مطالب ذکر شده معادله

$$(۵) dX_t = b(t, X_t)dt + \sigma(t, X_t)dB_t$$

معرف  $SDE$  فرایند  $X_t$  است. معادله (۵) را می‌توان به‌کمک عملگر انتگرال بدین صورت نیز نوشت:

$$(۶) X_t = X_0 + \int_0^t b(s, X_s)ds + \int_0^t \sigma(s, X_s)dB_s$$

قسمت دوم رابطه (۶) که انتگرال نسبت به حرکت براونی  $B_t$  است را نمی‌توان مانند انتگرال ریمان اشتبلیس بیان کرد. رویکردهای گوناگونی برای حل این انتگرال وجود دارد که از جمله آنها می‌توان به انتگرال ایتو و انتگرال استراتونویچ<sup>۱۲</sup> اشاره کرد که در این مقاله توجه‌مان را تنها معطوف به انتگرال ایتو کرده‌ایم. یکی از جنبه‌های مهم انتگرال ایتو امکان به‌کارگیری قاعده زنجیری<sup>۱۳</sup> برای حل آن است. به بیانی دقیق‌تر، در انتگرال ایتو می‌توان از فرمول ایتو که مانند قاعده زنجیری عمل می‌کند، استفاده کرد. شرط لازم و کافی برای استفاده از فرمول ایتو این است که فرایند مورد نظر فرایند ایتو باشد. بر اساس این ویژگی و با کاربرد فرمول ایتو می‌توان فرایند جدیدی را با  $SDE$  متناظرش به‌دست آورد. جزئیات بیشتر در این مورد در [۲۰] وجود دارد.

### آمار در فضای نااقلیدسی

تئوری و بسط بی‌حد و وصف مطالب ریاضی و آماری راجع به مشاهدات و داده‌های آماری که در فضای اقلیدسی قرار دارند بر کسی پوشیده نیست. این پیشرفت‌ها محققان را با مسائل و ایده‌های جدیدی روبرو کرده است که نقش آمار اقلیدسی در آنها به چالش کشیده می‌شود. از این دست داده‌ها به آمار غیرخطی یا نااقلیدسی یاد می‌شود که در سال‌های اخیر محققان آمار کاربردی نقش آنها را در علوم زیستی، محیطی و کامپیوتری، برجسته‌تر کردند. بنا به [۱۰] آمار نااقلیدسی مثل آمار دایره‌ای، کروی و شکل به داده‌هایی اطلاق می‌گردند که مقادیرشان را در فضایی نااقلیدسی اختیار می‌کنند. یکی از حالات خاص آمار نااقلیدسی، آمار سوار است [۱۹]. در سال‌های اخیر دانشمندان بسیاری در حوزه‌های مختلف علوم، مثل زیست‌شناسی، هواشناسی و

12. Stratonovich integral

13. Chain rule

بیوانفورماتیک به این نوع خاص از داده‌ها توجه کرده‌اند که مثال مناسب برای هر کدام به ترتیب عبارتند از: بررسی زاویه پرواز پرندگان در مهاجرت به مناطق دیگر، بررسی جهت حرکت باد و پیش‌بینی ساختار اصلی پروتئین‌ها بر اساس زوایای درونی اتم‌های تشکیل‌دهنده آن‌ها. یکی از ساده‌ترین مثال‌ها برای توصیف داده‌های سودار، داده‌های دایره‌ای است و آن داده‌هایی است که مقادیرش روی دایره اتفاق می‌افتد. ویژگی اساسی داده‌های دایره‌ای این است که آغاز و پایان آن‌ها یکسان است (داده‌ها دارای دور هستند). به‌طور دقیق، اگر داده‌های روی دایره را با زاویه مشخص کنیم، آنگاه واضح است زوایای  $\theta$  و  $\theta + 2\pi$  معادل هم هستند از این‌رو، باید این موضوع در محاسبات آماری راجع به زوایا مدنظر قرار گیرد.

چنان‌که که بیان شد به دلیل ویژگی‌های خاص آمار سودار مثل تناوبی بودن، می‌توان انتظار داشت معیارهای مرسوم آمار خطی به‌طور مستقیم قابل استفاده برای داده‌های روی دایره یا هر داده ناقلیدسی دیگر نباشند. یکی از این موارد معیار مرکزی میانگین حسابی است که برای داده‌های سودار بر اساس قاعده جمع مشاهدات تقسیم بر تعداد مشاهدات عمل نمی‌کند. مورد دیگر مربوط به تعریف مدل احتمالاتی است. به‌طور دقیق، یکی از ابزارهای مفید برای بررسی رفتارهای تصادفی داده‌های سودار، بررسی توزیع‌های آماری در این فضا است. می‌توان انتظار داشت که توزیع نرمال نمی‌تواند مستقیماً برای داده‌های سودار به‌کار گرفته شود از این‌رو، توزیع‌های دیگری در این حوزه نقشی اساسی دارند. یکی از این توزیع‌ها، توزیع فون‌میسز<sup>۱۴</sup> ( $VM$ ) است. این توزیع مثل توزیع نرمال که دارای کاربردهای زیادی در آمار است، نقش مهمی در استنباط آماری روی دایره ایفا می‌کند.

تابع چگالی توزیع فون‌میسز برای متغیر تصادفی دایره‌ای  $\theta$  با پارامترهای  $K$  و  $\mu$  برابر

$$f(\theta; \mu, k) = \frac{1}{2\pi I_0(k)} \exp[k \cos(\theta - \mu)] \quad 0 \leq \theta \leq 2\pi, \quad 0 \leq \mu \leq 2\pi$$

است که پارامتر  $\mu$  را میانگین سودار و  $k$  را پارامتر تمرکز می‌گویند [۱۹]. به علاوه،  $I_0(k)$  تابع بسل تغییر یافته نوع اول و از مرتبه صفر است [۱].

### مدل‌بندی فرایندهای انتشار در فضاها ناقلیدسی

چنان‌که که قبلاً بیان شد فرایندهای تصادفی و به‌خصوص فرایندهای انتشار برای مدل‌بندی پدیده‌های گوناگون به‌کار گرفته می‌شوند. یکی از موضوعات مورد علاقه محققان بررسی فرایندهای تصادفی مربوط به رفتار تصادفی داده‌های ناقلیدسی است. انتظار می‌رود مدل‌بندی فرایندهای انتشار در فضای ناقلیدسی کمک شایانی به حل بسیاری از مسائل مربوط به این‌گونه داده‌ها باشد. برای نیل به این هدف به چند ابزار ریاضی نیاز داریم که بدین شرح است:

از نقطه نظر ریاضی، متناظر با فرایند انتشار، یک عملگر دیفرانسیلی با مشتقات جزئی مرتبه دوم وجود دارد که رابطه مستقیمی با ضرایب رانش و انتشار فرایند دارد. از این رو با داشتن این عملگر و با استفاده از مطالب ارائه شده در بخش‌های قبل، می‌توان به راحتی فرایند انتشار مورد نظر را به وسیله SDE متناظر مدل‌بندی کرد. یکی از کاربردهای این عملگر ساخت حرکت براونی روی خمینه ریمانی است [۲۲]. لازم به ذکر است

14. Von Mises distribution

خمینه ریمانی يك فضای نااقلیدسی و به عبارتی تعمیم ساده‌ای از سطوح ناهموار بر روی فضای اقلیدسی است. خواننده علاقه‌مند برای تحقیق بیشتر در مورد خمینه ریمانی می‌تواند به منابع مرتبط با هندسه دیفرانسیل مثل [۹] و [۱۱] مراجعه کند. نوع ارتباط بین عملگر دیفرانسیلی و حرکت براونی روی خمینه ریمانی در پی می‌آید.

فرض کنید  $X_t$  فرایند ایتوی انتشار باشد، آن‌گاه عملگر مشخصه  $A$  این فرایند به صورت

$$Af(x) = \lim_{u \downarrow x} \frac{E^x[f(X_{\tau_u})] - f(x)}{E^x[\tau_u]} \quad (۸)$$

تعریف می‌شود که در آن  $\tau_u = \inf \{t > 0; X_t \notin U\}$  اولین زمان خروج  $X_t$  از  $U$  است [۲۰]. عملگر مشخصه (۸) با ضرایب رانش و انتشار فرایند رابطه مستقیمی دارد طوری که با این ضرایب عملگر مشخصه فرایند به راحتی قابل محاسبه است. قضیه ۷ نشان‌دهنده این موضوع است.

قضیه [۷]: فرض کنید  $D_A$  مجموعه‌ای از همه توابع  $f$  باشد که برای هر  $x \in \mathbb{R}^n$  حد (۸) موجود است. اگر

$$f(x) \in C^2, \text{ یعنی } f \text{ دوبار مشتق پذیر و پیوسته باشد, آن‌گاه } f \in D_A \text{ و} \quad (۹)$$

$$Af(x) = \sum_i b_i \frac{\partial f}{\partial x_i} + \frac{1}{2} \sum_{i,j} (\sigma \sigma^T)_{i,j} \frac{\partial^2 f}{\partial x_i \partial x_j}$$

چنان‌که قبلاً ذکر شد یکی از کاربردهای عملگر  $A$  ساخت فرایند براونی روی خمینه‌های ریمانی است. نحوه ساخت این فرایند بدین صورت است:

عملگر مشخصه حرکت براونی روی خمینه ریمانی  $M$  با تانسور متریک  $g$  و مختصات موضعی (کارت<sup>۱۵</sup>)

به صورت  $x = (x_1, \dots, x_n)$  برابر  $\frac{\Delta_M}{2}$  است، که در آن عملگر لاپلاس بلترامی<sup>۱۶</sup> بوده و به صورت

$$\Delta_M = \frac{1}{\sqrt{\det(g)}} \sum_i \frac{\partial}{\partial x_i} \left[ \sqrt{\det(g)} \sum_j g^{ij} \frac{\partial}{\partial x_j} \right] \quad (۱۰)$$

تعریف می‌شود، که در آن به‌ازای  $\{i, j = 1, \dots, n\}$  عنصر  $g^{ij}$ ، ام معکوس ماتریس تانسور متریک  $g$

است [7]. با کمی محاسبات جبری تساوی (10) را می‌توان به صورت

$$\begin{aligned} \Delta_M &= \frac{1}{\sqrt{\det(g)}} \sum_i \left[ \sum_j \frac{\partial}{\partial x_i} (\sqrt{\det(g)} g^{ij} \frac{\partial}{\partial x_j}) \right] \\ &= \frac{1}{\sqrt{\det(g)}} \sum_i \left[ \sum_j \left( \frac{\partial}{\partial x_i} (\sqrt{\det(g)} g^{ij}) \frac{\partial}{\partial x_j} + \sqrt{\det(g)} g^{ij} \frac{\partial}{\partial x_i} \left( \frac{\partial}{\partial x_j} \right) \right) \right] \\ &= \sum_{i=1}^n \left[ \frac{1}{\sqrt{\det(g)}} \sum_{j=1}^n \frac{\partial}{\partial x_j} (\sqrt{\det(g)} g^{ij}) \right] \frac{\partial}{\partial x_i} + \sum_{i=1}^n \sum_{j=1}^n g^{ij} \frac{\partial^2}{\partial x_i \partial x_j}. \end{aligned}$$

برابر  $M$  ساده کرد. بنا بر این عملگر مشخصه حرکت براونی روی خمینه ریمانی

$$Af(x) = \frac{1}{2} \Delta_M = \sum_{i=1}^n \left[ \frac{1}{2\sqrt{\det(g)}} \sum_{j=1}^n \frac{\partial}{\partial x_j} (\sqrt{\det(g)} g^{ij}) \right] \frac{\partial}{\partial x_i} + \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n g^{ij} \frac{\partial^2}{\partial x_i \partial x_j}. \quad (۱۱)$$

است. با مقایسه روابط (۹) و (۱۱) می‌توان دریافت که ضرایب رانش و انتشار حرکت براونی روی خمینه ریمانی  $M$  به ترتیب عبارتند از:

$$(\sigma \sigma^T)_{ij} = g^{ij} \quad (۱۲)$$

15 Chart  
16 Laplace-Beltrami

$$b_i(x) = \frac{1}{2\sqrt{\det(g)}} \sum_{j=1}^n \frac{\partial}{\partial x_j} (\sqrt{\det(g)} g^{ij}) \quad (13)$$

چنان‌که ملاحظه می‌شود عبارت (۱۱) بیان‌گر رابطه بین ضرایب رانش و انتشار حرکت براونی با عمل‌گر مشخصه آن است. به زبانی دیگر، می‌توان گفت که اگر ضرایب رانش و انتشار حرکت براونی بررسی شده روی یک خمینه معلوم باشد به راحتی می‌توان عمل‌گر مشخصه آن را به دست آورد. بر عکس، اگر عمل‌گر مشخصه حرکت براونی روی خمینه‌ی بررسی شده معلوم باشد، آن‌گاه ضرایب رانش و انتشار آن مشخص و در نتیجه می‌توان SDE آن را به دست آورد [۲۲]. این ارتباط دو طرفه در بخش پیش‌رو و بخش‌های بعد استفاده می‌شود.

### حرکت براونی روی چنبره

موضوع بررسی پروتئین‌ها و خواص آن‌ها از مسائل بسیار مهم علوم زیستی است. به‌طور کلی پروتئین‌ها با اتصال تعدادی اسید آمینه در زوایای مختلف شکل می‌گیرند. به‌علاوه زنجیره اصلی پروتئین می‌تواند با دنباله‌ای از جفت زاویه دو سطحی  $\psi$  و  $\phi$  که هر دوی آن‌ها در بازه  $[-\pi, \pi]$  قرار دارند، مشخص شود [۴]. اخیراً، [۵] نشان دادند، زوایای فوق را می‌توان معرف نقاطی روی چنبره در نظر گرفت و آن‌گاه برای مدل‌بندی زوایای دوسطحی یک مدل احتمالی معرفی کردند. از آن‌جا که ساختار بسیاری از پلیمرها مانند DNA مینی‌پلاسمیدها می‌توانند با چنین زاویه‌های دوسطحی نمایش داده شوند [۲۵]، یافتن خواص آماری این دو زاویه وقتی به‌عنوان متغیرهای تصادفی در نظر گرفته می‌شوند، می‌تواند کمک شایانی به محققان حوزه علوم زیستی کند. به‌علاوه، نحوه مدل‌بندی فرایندهای تصادفی روی چنبره می‌تواند به پیش‌بینی دقیق‌تر ساختار پروتئین و DNA منجر شود. اضافه بر آن بررسی توزیع مانای فرایندهای تصادفی روی چنبره و ارائه الگوریتم‌های آماری مناسب برای شبیه‌سازی از SDE‌های متناظر می‌تواند اطلاعاتی مشابه آن‌چه که در علوم زیستی تحت عنوان شبیه‌سازی مولکولی ساختاری معروف است و برای پیش‌گویی ساختار پروتئین‌ها به‌کار می‌رود [۴]، فراهم کند.

از نقطه نظر هندسی چنبره سطحی است که از چرخش یک دایره حول خطی صاف متعلق به صفحه دایره تشکیل می‌شود. واضح است هنگامی که محور چرخش قطر دایره باشد، شکل حاصل کره است. از منظر ریاضی چنبره را می‌توان به‌صورت پارامتری

$$x = (R + r \cos v) \cos u, \quad y = (R + r \cos v) \sin u, \quad z = r \sin v$$

نوشت که در آن  $0 \leq u \leq 2\pi$ ،  $0 \leq v \leq 2\pi$ ،  $R$  فاصله مرکز دایره تا مرکز چنبره (فاصله محور از مرکز دایره که شعاع دایره بزرگ هم گفته می‌شود) و  $r$  شعاع دایره دوار (شعاع دایره کوچک) است [۹]. به‌طور کلی، بنا به مقادیر متفاوت  $R$  و  $r$  (هر دو مثبت) سه نوع چنبره متفاوت وجود دارد که در این مقاله چنبره‌ای مدنظر قرار می‌گیرد که  $R > r$ . چنبره یک خمینه ریمانی است و بنا بر [۱۱] می‌توان یک مختصات موضعی برای چنبره به‌صورت

$$f(x, y, z) = \left( \tan^{-1} \frac{y}{x}, \sin^{-1} \frac{z}{r} \right) = (u, v)$$

تعریف کرد. همچنین با محاسباتی ساده و براساس [۹] ماتریس تانسور متریک  $(g)$  و معکوس آن  $(g^{-1})$  برای این چنبره به‌ترتیب عبارتند از:



$$g = \begin{pmatrix} (R + r \cos v)^2 & 0 \\ 0 & r^2 \end{pmatrix}, \quad g^{-1} = \begin{pmatrix} 1/(R + r \cos v)^2 & 0 \\ 0 & 1/r^2 \end{pmatrix}$$

اکنون با استفاده از روابط (۱۲) و (۱۳) می‌توان ضرایب رانش و انتشار حرکت براونی روی چنبره را محاسبه کرد. با محاسباتی ساده می‌توان نوشت:

$$\begin{aligned} \sum_{j=1}^2 \frac{\partial}{\partial x_j} (\sqrt{\det(g)} g^{1j}) &= \frac{\partial}{\partial u} \left( \frac{r(R + r \cos v)}{(R + r \cos v)^2} \right) + \frac{\partial}{\partial v} (r(R + \cos v) \times 0) \\ &= \frac{\partial}{\partial u} \left( \frac{r}{R + r \cos v} \right) = 0 \end{aligned}$$

و

$$\begin{aligned} \sum_{j=1}^2 \frac{\partial}{\partial x_j} (\sqrt{\det(g)} g^{2j}) &= \frac{\partial}{\partial u} (r(R + r \cos v) \times 0) + \frac{\partial}{\partial v} \left( \frac{r(R + r \cos v)}{r^2} \right) \\ &= \frac{\partial}{\partial v} \left( \frac{R + r \cos v}{r} \right) = -\sin v \end{aligned}$$

از این رو ضرایب رانش چنین فرایندی عبارتند از:

$$b_1(x) = \frac{1}{2\sqrt{\det(g)}} \times 0 = 0, \quad b_2(x) = \frac{-\sin v}{2\sqrt{\det(g)}} = \frac{-\sin v}{2r(R + r \cos v)}$$

همچنین، می‌توان ملاحظه کرد که بنا به معکوس ماتریس تانسور متریک عناصر ماتریس انتشار به صورت

$$\begin{aligned} (\sigma\sigma^T)_{11} = g^{11} &= \frac{1}{(R + r \cos v)^2}, \quad (\sigma\sigma^T)_{22} = g^{22} = \frac{1}{r^2}, \\ (\sigma\sigma^T)_{12} = (\sigma\sigma^T)_{21} &= 0 \end{aligned}$$

هستند. در نتیجه SDE حرکت براونی  $X_t = (U_t, V_t)^T$  روی چنبره را می‌توان به صورت

$$\begin{pmatrix} dU_t \\ dV_t \end{pmatrix} = \begin{pmatrix} 0 \\ \sin V_t \\ -\frac{1}{2r(R + r \cos V_t)} \end{pmatrix} dt + \begin{pmatrix} \frac{1}{R + r \cos V_t} & 0 \\ 0 & \frac{1}{r} \end{pmatrix} \begin{pmatrix} dB_{1t} \\ dB_{2t} \end{pmatrix} \quad (14)$$

نوشت که در آن  $B_t = (B_{1t}, B_{2t})$  بیانگر حرکت براونی دو بعدی (در صفحه مختصات دکارتی) است. اکنون می‌توان از این حرکت برای ساخت فرایند اورنشستاین-اولن‌بک روی چنبره استفاده کرد.

### فرایند اورنشستاین-اولن‌بک روی چنبره

چنان‌که در بخش‌های قبل بیان شد، از فرایندهای تصادفی روی چنبره می‌توان برای مدل‌بندی زوایای دوسطحی پروتئین بهره جست. یکی از فرایندهای تصادفی روی خط که برای مدل‌بندی پدیده‌هایی با هدف حرکت معین (اما تحت تاثیر نوفه) استفاده می‌شود، فرایند  $OU$  روی خط است. از این رو، چون جهت‌گیری حرکتی پروتئین‌ها به سمت خاصی است، به نظر می‌رسد معرفی، شناسایی و تحلیل فرایند  $OU$  روی چنبره با زوایای دوسطحی می‌تواند به پیش‌بینی بهتر ساختارهای گوناگون پروتئین‌ها کمک کند. به همین منظور در ذیل با الهام از حرکت براونی روی چنبره فرایند  $OU$  روی آن معرفی و ویژگی‌های آن بررسی می‌شوند.

برای ساختن فرایند  $OU$  روی چنبره به تبعیت از [۱۴]، می‌توان تابع  $-\lambda \sin(V_t - \mu)$  را به ضریب رانش فرایند تصادفی  $V_t$  در معادله (۱۴) اضافه کرد و سپس توزیع مانای فرایند حاصل را بررسی کرد. اما، به جای این کار ما ابتدا یک تابع کلی مثل  $\frac{\partial g}{\partial v}(V_t, \theta)$  را به ضریب رانش فرایند تصادفی  $V_t$  اضافه کرده و سپس توزیع مانای فرایند حاصل را محاسبه و ارزیابی می‌کنیم. واضح است که حالت خاص این تابع کلی را می‌توان برابر  $-\lambda \sin(V_t - \mu)$  اختیار کرد و توزیع مانای فرایند متناظر را به دست آورد. بنا به مطالب بالا  $SDE$  یک خانواده از فرایندهای  $OU$  روی چنبره را به صورت

$$\begin{pmatrix} dU_t \\ dV_t \end{pmatrix} = \begin{pmatrix} 0 \\ \frac{\partial g}{\partial v}(V_t, \theta) - \frac{\sin V_t}{2r(R + r \cos V_t)} \end{pmatrix} dt + \begin{pmatrix} \frac{1}{R + r \cos V_t} & 0 \\ 0 & \frac{1}{r} \end{pmatrix} \begin{pmatrix} dB_{1t} \\ dB_{2t} \end{pmatrix} \quad (15)$$

پیشنهاد می‌کنیم. این فرایند ویژگی‌های جالب ریاضی و زیستی دارد. اما، آن جنبه‌ای از این فرایند که برای محققان آمار و ریاضی ممکن است حائز اهمیت باشد محاسبه توزیع مانا و چگونگی استنباط آمار راجع به پارامترهای فرایند است که در بخش بعد به آن پرداخته می‌شود.

### توزیع مانای فرایند اورنشتاین-اولن‌بک روی چنبره

فرض کنید  $X_t = (U_t, V_t)^T$  معرف فرایند  $OU$  روی چنبره با  $SDE$  به صورت رابطه (۱۵) است. چون این فرایند دارای ضرایب رانش و انتشار مستقل از  $t$  است از این رو بنا به [۱۸] یک فرایند انتشار همگن است. بنا بر این توزیع مانای این فرایند در معادله پیش‌روی کولموگوروف (معادله (۱)) صدق می‌کند. از آنجا که روشی برای حل این معادله برای توزیع‌های دو متغیره موجود نیست محاسبه توزیع مانای فرایند  $X_t$  نیز تا حدودی مشکل است. اما چون دو فرایند  $U_t$  و  $V_t$  از هم مستقل هستند ما به جای حل معادله مذکور، ابتدا توزیع مانای این دو فرایند را جداگانه به دست آورده و سپس حاصل ضرب آن‌ها را به عنوان توزیع مانای فرایند  $X_t$  در نظر می‌گیریم. در انتها به منظور اثبات این‌که توزیع حاصل واقعاً توزیع مانای فرایند است، نشان می‌دهیم که آن در معادله پیش‌روی کولموگوروف (۱) صدق می‌کند.

به منظور سهولت در به دست آوردن توزیع مانا، شعاع کوچک برای چنبره ( $r$ ) را یک و شعاع دایره بزرگ ( $R$ ) را ۲ در نظر می‌گیریم. با توجه به این مقادیر ثابت و  $SDE$  رابطه (۱۵) ضرایب رانش و انتشار فرایندهای  $U_t$  و  $V_t$  به عنوان مؤلفه‌هایی از فرایند  $OU$  روی چنبره بدین صورت هستند:

$$b(V_t) = \frac{\partial g}{\partial v}(V_t, \theta) - \frac{\sin V_t}{2(2 + r \cos V_t)}, \quad \sigma(V_t) = 1$$

$$b(U_t) = 0, \quad \sigma(U_t) = \frac{1}{2 + \cos V_t}$$

اکنون، می‌توان توزیع مانای فرایند  $V_t$  را به دست آورد. مطابق مطالبی که در بخش دوم ارائه شد می‌توان توابع  $S(v)$ ،  $s(v)$  و  $m(v)$  را برای این فرایند براحتی محاسبه کرد. ابتدا، بنا به (۲) داریم:

$$s(v) = \exp\left(\int^v -\frac{2b(y)}{\sigma^2(y)} dy\right) \\ = \exp\left(\int^v \left[\frac{\sin y}{2 + \cos y} - 2\frac{\partial g(y, \theta)}{\partial y}\right] dy\right) = \frac{1}{2 + \cos v} \exp[-2g(v, \theta)]$$

از این‌رو توابع  $m(v)$  و  $S(v)$  به‌ترتیب برابر

$$m(v) = \frac{1}{\sigma^2(v)s(v)} = (2 + \cos v) \exp[2g(v, \theta)]$$

و

$$S(v) = \int^v s(y) dy = \int^v \frac{1}{2 + \cos y} \exp[-2g(v, \theta)]$$

خواهند بود. بنا بر این توزیع مانای فرایند  $V_t$  برابر

$$\psi(v) = (2 + \cos v) e^{2g(v, \theta)} [C_1 S(v) + C_2] \quad (۱۶)$$

است. برای حالت خاص فرض کنید  $\frac{\partial g(V_t, \theta)}{\partial v} = -\lambda \sin(V_t - \mu)$ . در آن صورت  $SDE$  فرایند  $V_t$  (معادله

۱۵) به‌صورت

$$dV_t = -\lambda \sin(V_t - \mu) - \frac{\sin V_t}{2(2 + \cos V_t)} dt + dB_2$$

ساده می‌شود. با توجه به این‌که  $g(V_t, \theta) = \lambda \cos(V_t - \mu)$  و بنا بر (۱۶) توزیع مانای فرایند  $V_t$  برابر

$$\psi(v) = (2 + \cos v) e^{2\lambda \cos(v - \theta)} [C_1 S(v) + C_2]$$

خواهد شد. حال با فرض  $C_1 = 0$ ، می‌توان یکی از توزیع‌های مانای فرایند  $V_t$  را بدین‌صورت محاسبه کرد:

$$\psi(v) = C_2 m(v) = \frac{m(v)}{\int_0^{2\pi} m(v)} = \frac{(2 + \cos v) e^{2\lambda \cos(v - \theta)}}{\int_0^{2\pi} m(v)} \quad (۱۷)$$

واضح است که مخرج کسر در (۱۷) ثابت نرمال‌ساز توزیع  $\psi(v)$  است. از این رو، با کمی محاسبات جبری

می‌توان نشان داد که توزیع مانای فرایند در این حالت به‌صورت

$$\psi(v) = \frac{2 \exp(2\lambda \cos(v - \mu)) + \cos v \exp(2\lambda \cos(v - \mu))}{2\pi(2I_0(2\lambda) + \cos \mu I_1(2\lambda))} \quad (۱۸)$$

است. که  $I_q(x)$  تابع بسل اصلاح شده از نوع یک و از مرتبه  $q$  است.

برای محاسبه توزیع مانای فرایند  $U_t$  ابتدا توابع  $s(v)$  و  $m(v)$  را محاسبه می‌کنیم. بنا بر (۲) این توابع به‌ترتیب

برابر  $s(u) = 1$  و

$$m(u) = \frac{1}{s(u)\sigma^2(u)} = (2 + \cos v)^2$$

هستند. در نتیجه توزیع مانای فرایند  $U_t$  برابر

$$\psi(u) = (2 + \cos v)^2 [C_1 S(u) + C_2], \quad 0 \leq u \leq 2\pi$$

است. مشابه روندی که برای فرایند  $V_t$  انجام شد توزیع مانای فرایند  $U_t$  برای  $0 \leq u \leq 2\pi$  به‌صورت

$$\psi(u) = C_2 m(u) = \frac{m(u)}{\int_0^{2\pi} m(u) du} = \frac{(2 + \cos v)^2}{\int_0^{2\pi} (2 + \cos v)^2 du} = \frac{1}{2\pi} \quad (۱۹)$$

است. با توجه به این‌که توزیع مانای فرایند  $U_t$  یکنواخت در بازه  $[0, 2\pi]$  است، امید ریاضی و واریانس آن

به‌ترتیب برابر  $\pi$  و  $\frac{\pi^2}{3}$  هستند.

بنا بر این با توجه به مطالب ذکر شده در ابتدای این بخش حاصل ضرب تابع چگالی با معادله (۱۸) و (۱۹) را به عنوان توزیع مانای فرایند  $OU$  با مؤلفه  $U_t$  و  $V_t$  در نظر می‌گیریم. یعنی، به ازای  $u$  و  $v$  در  $[0, 2\pi]$ ، توزیع مانای فرایند  $OU$  پیشنهادی در این مقاله به صورت

$$f(u, v) = \psi(u)\psi(v) = \frac{1}{2\pi} \frac{2 \exp(2\lambda \cos(v - \mu)) + \cos v \exp(2\lambda \cos(v - \mu))}{2\pi(2I_0(2\lambda) + \cos \mu I_1(2\lambda))} \quad (20)$$

است. اکنون بنا بر آنچه که بیان شد، برای ارزیابی درستی این ادعا که توزیع پیشنهادی (۲۰) توزیع مانای فرایند  $OU$  روی چنبره با  $SDE$  به صورت (۱۵) است، کافی است نشان داده شود که این توزیع در معادله پیش‌روی کولموگوروف متناظر با آن صدق می‌کند. اثبات این مطالب به صورت دستی کمی پر حجم اما شدنی است ولی با ابزارهای محاسباتی کامپیوتری جدید این موضوع بسیار آسان خواهد بود. این امر با برنامه‌نویسی کامپیوتری انجام و صحت توزیع مانای تأیید شد.

### بحث و نتیجه‌گیری

در این مقاله مدل‌بندی فرایندهای انتشار در فضاهاى نااقلیدسی بررسی شده. به‌ویژه، با استفاده از تانسور متریک، معادلات دیفرانسیل تصادفی حرکت‌های براونی و فرایندهای اورنشتاین-اولن‌بک روی چنبره به دست آمد. به علاوه نحوه شبیه‌سازی از معادلات دیفرانسیل تصادفی توصیف شد. همچنین توزیع‌های مانای فرایند اورنشتاین-اولن‌بک روی چنبره محاسبه شد. از آن‌جاکه زنجیره اصلی پروتئین سنگ‌بنای اصلی مواد شیمیایی متفاوت است انتظار می‌رود مدل‌بندی مناسب فرایندهای انتشار تغییرات تصادفی زنجیر، نیازمند اعمال محدودیت‌های دیگری نیز باشد. از این رو، به عنوان تحقیقات آتی می‌توان استفاده از فرایندهای انتشار چند متغیره همراه با هم‌بستگی فضایی بین اتم‌ها برای پیش‌بینی بهتر زنجیره اصلی پروتئین را پیشنهاد کرد. به علاوه به‌کارگیری تقریب‌های دیگر که مستقیماً در فضای نااقلیدسی به‌کار گرفته می‌شوند نیز می‌تواند موضوع تحقیقات آینده در این حوزه باشد. علاوه بر این موضوعات، نباید از استنباط آماری راجع به پارامترهای این مدل و مدل‌های دیگر غافل شد.

### منابع

1. Abramowitz M., Stegun I.A., "Handbook of Mathematical Functions", Dover, New York (1995).
2. Armitage P., Colten T., "Encyclopedia of Biostatistics", John Wiley and Sons, Chichester (2005).
3. Yoshida K., "Brownian Motion on the Surface of the 3-sphere", Annals of Mathematical Statistics, 20 (1949) 292-296.
4. Bourne P.E., Weissig H., "Structural Bioinformatics", John Wiley and Sons, New Jersey (2003).

5. Boomsma W., Mardia K.V., Taylor C.C., Ferkinghoff-Borg J., Kroghand A., Hamelryck T., "A Generative, Probabilistic Model of Local Protein Structure", *Proceedings of the National Academy of Sciences*, 105 (2008) 8932-8937.
6. Bystroff C., Thorsson V., Baker D., "HMMSTR: a Hidden Markov Model for Local Sequence-Structure Correlations in Proteins", *Journal of Molecular Biology*, 301 (2000) 173-190.
7. Hsu E.P., "Stochastic Analysis on Manifolds", American Mathematical Society (2002).
8. Ito K., "Stochastic Integral", *Proceedings of the Imperial Academy of Japan*, 20 (1944) 519-524.
9. Ivancevic V.G., "Applied Differential Geometry: a Modern Introduction", World Scientific Publishing, Singapore (2007).
10. Jammalamadaka S.R., SenGupta A., "Topic in Circular Statistics", World Scientific Publishing, Singapore (2001).
11. Jost J., "Riemannian Geometry and Geometric Analysis", Springer-Verlag, Berlin, (2005).
12. Karatzas L., Steven E.S., "Brownian Motion and Stochastic Calculus", Springer-Verlag, New York (1991).
13. Karlin S., Taylor H.M., "A Second Course in Stochastic Processes", Academic press, New York (1981).
14. Kent J.T., "Discussion of Professor Mardia's Paper", *Journal of the Royal Statistical Society, Series B*, 37 (1975) 377-378.
15. Kloden P.E., Platen E., "Numerical Solution of Stochastic Differential Equations", Springer-Verlag, Berlin (1992).
16. Krantz S.G., "Differential Equations", Demystified, McGraw-Hill (2004).
17. Kuo H.H., "Introduction to Stochastic Integration", Springer, New York (2006).
18. Mao X., "Stochastic Differential Equations and Their Applications", Horwood, Chichester (1997).
19. Mardia K.V., Jupp P.E., "Directional Statistics", John Wiley and Sons, Chichester (2000).
20. Oksendal B., "Stochastic Differential Equations: An Introduction with Applications", Springer-Verlag, Berlin (2005).
21. Platen E., Heath D., "A Benchmark Approach to Quantitative Finance", Springer-Verlag, Berlin (2006).

22. Rogers L., Williams D., "Diffusion, Markov Processes and Martingales1", Cambridge University Press, Cambridge (2000).
23. Simons K.T., Kooperberg C., Huang E., Baker D., "Assembly of Protein Tertiary Structures from Fragments with Similar Local Sequences Using Simulated Annealing and Bayesian Scoring Functions", Journal of Molecular Biology, 268 (1997) 209-225.
24. Soize C., "The Fokker-Planck Equation for Stochastic Dynamical Systems and Its Explicit Steady State Solutions", World Scientific, Singapore, (1994).
25. Thaokar R.M., "Brownian Motion of a Torus", Colloids and Surfaces A: Physicochem, Engineering Aspects, 317 (2008) 650-657.
26. Uhlenbeck G.E., Ornstein L.S., "On the Theory of Brownian Motion", Physics Review, 36 (1930) 823-841.