پانزدهمین سمینار احتمال و فرآیندهای تصادفی
۸ و ۹ شهریور ۱۴۰۴
دانشگاه کردستان

15th Seminar On Probability and Stochastic Processes
August 30-31, 2025
University of Kurdistan

انجمن آمار ایران

روش استوار مونت‌کارلوی بهینه

امید کریمی [1] ، **فاطمه حسینی**

گروه آمار دانشگاه سمنان

چکیده: اغلب موارد در تحلیل فرایندهای سری‌زمانی و فضایی با مدل‌های آماری پیچیده‌ای مواجه می‌شویم که تابع درست‌نمایی آن‌ها براحتی قابل محاسبه نیستند. به همین دلیل، بسیاری از روش‌های استنباط بدون درست‌نمایی با محدودیت‌هایی در دقت و کارایی مواجه هستند. در این مقاله، یک روش استنباط بدون درست‌نمایی با استفاده از روش استوار مونت‌کارلوی بهینه را برای تحلیل داده‌های وابسته فضایی ارائه می‌کنیم. این رهیافت یک چارچوب نوین و کارآمد برای استنباط بدون درست‌نمایی است که نمونه‌های وزنی دقیقی از توزیع پسین ارائه می‌دهد. در نهایت، روش پیشنهادی روی داده‌های شبیه‌سازی شده پیاده‌سازی و ارزیابی می‌شود.

واژه‌های کلیدی: استنباط بیزی، داده‌های وابسته، نمونه‌های مونت کارلویی، استنباط بدون درست‌نمایی.

کد موضوع‌بندی ریاضی (۲۰۲۰): 62F15، 62H11، 62M05.

۱ مقدمه

مدل‌های مبتنی بر شبیه‌ساز به دلیل انعطاف‌پذیری بالایی که در مدل‌سازی ارائه می‌دهند، بسیار جذاب هستند. در واقع، هر مکانیزم تولید داده که بتوان آن را به صورت مجموعه‌ای محدود از گام‌های الگوریتمی تعریف کرد، می‌تواند به عنوان یک مدل مبتنی بر شبیه‌ساز برنامه‌نویسی شود. از این رو، این مدل‌ها اغلب برای شبیه‌سازی پدیده‌های فیزیکی در علوم طبیعی، مانند ژنتیک، اپیدمیولوژی، یا علوم اعصاب، به کار می‌روند (گاتمان و کورندر، ۲۰۱۶). در این مدل‌ها، تولید نمونه‌ها با استفاده از روش‌های مونت کارلویی امکان پذیر است، اما محاسبه تابع درست‌نمایی غیرممکن است. ناتوانی در محاسبه تابع درست‌نمایی، استنباط بدون درست‌نمایی (LFI)[1] ، یعنی تقریب توزیع پسین بدون استفاده از تابع درست‌نمایی، را به چالشی بزرگ تبدیل می‌کند. روش مونت‌کارلوی بهینه (OMC)[2] که توسط میدز و ولینگ (۲۰۱۵) پیشنهاد شد، یک روش نوین برای LFI است. ایده اصلی این روش، تبدیل مکانیزم تصادفی تولید داده به

[1] سخنران، omid.karimi@semnan.ac.ir

[1] Likelihood-Free Inference
[2] Optimization Monte Carlo

مجموعه‌ای از فرآیندهای بهینه‌سازی است. گولمیس و همکاران (۲۰۱۶) برخی محدودیت‌های کلیدی OMC را شناسایی و نسخه بهبودیافته‌ای به نام مونت‌کارلوی بهینه استوار (ROMC)[۳] را پیشنهاد کردند.

در این مقاله، نحوه پیاده‌سازی ROMC را برای تقریب بهینه توزیع پسین روی داده‌های وابسته ارائه می‌کنیم. ROMC یک چارچوب استنباط بدون درست‌نمایی است که مجموعه‌ای از گام‌های الگوریتمی را برای تقریب توزیع پسین تعریف می‌کند، بدون آنکه الگوریتم خاصی را برای هر گام الزامی کند. بنابراین، ROMC را می توان به عنوان یک روش پایه استفاده کرد. در نهایت با یک مطالعه شبیه‌سازی روش پیشنهادی را ارزیابی و تحلیل می‌کنیم.

۲ مدل‌های مبتنی بر شبیه‌ساز

در این بخش، ابتدا مقدمه‌ای کوتاه درباره مدل‌های مبتنی بر شبیه‌ساز ارائه می‌دهیم، سپس به روش OMC و نسخه استوار آن ROMC می‌پردازیم. مدل مبتنی بر شبیه‌ساز، مکانیزمی تصادفی و پارامتری برای تولید داده است که امکان نمونه‌گیری از داده‌ها را فراهم می‌کند، اما محاسبه تابع درست‌نمایی در آن امکان‌پذیر نیست. در واقع یک مدل مبتنی بر شبیه‌ساز، مجموعه‌ای پارامتری از توابع چگالی احتمال $\{p(\mathbf{y} \mid \boldsymbol{\theta})\}_{\boldsymbol{\theta}}$ است که شکل بسته ندارد یا از نظر محاسباتی غیرقابل محاسبه است. در چنین مواردی، تنها دسترسی به شبیه‌ساز $m_r(\boldsymbol{\theta})$ ممکن است، یعنی یک مکانیزم مبتنی بر الگوریتم برنامه‌نویسی که به صورت تصادفی داده‌های $\mathbf{y}$ را از مجموعه‌ای از پارامترها $\boldsymbol{\theta} \in \mathbb{R}^D$ تولید می‌کند. فرآیند نمونه‌گیری از شبیه‌ساز را با $\mathbf{y} \leftarrow m_r(\boldsymbol{\theta})$ نشان می‌دهیم. همان‌طور که میدز و ولینگ (۲۰۱۵) نشان داده‌اند، می‌توان تصادفی بودن شبیه‌ساز را با معرفی مجموعه‌ای از متغیرهای تصادفی $\mathbf{u} \sim p(\mathbf{u})$ جدا کرد. بنابراین، برای یک جفت مشخص $(\boldsymbol{\theta}, \mathbf{u})$، شبیه‌ساز به یک نگاشت قطعی g تبدیل می‌شود، به طوری که $\mathbf{y} = g(\boldsymbol{\theta}, \mathbf{u})$ است. فرض کنید $\mathbf{y}_\circ$ داده‌ی مشاهده‌شده باشد، مشکل اصلی ناتوانی در محاسبه تابع درست‌نمایی $L(\boldsymbol{\theta}) = p(\mathbf{y}_\circ \mid \boldsymbol{\theta})$ است. در واقع درست‌نمایی، احتمال تولید داده $\mathbf{y}$ مشابه داده مشاهده‌شده $\mathbf{y}_\circ$ با استفاده از پارامترهای $\boldsymbol{\theta}$ است. بنابراین، تابع درست‌نمایی $L(\boldsymbol{\theta})$ به‌صورت

$$L(\boldsymbol{\theta}) = \lim_{\epsilon \to \circ} c_\epsilon \int_{\mathbf{y} \in B_{d,\epsilon}(\mathbf{y}_\circ)} p(\mathbf{y} \mid \boldsymbol{\theta}) d\mathbf{y} = \lim_{\epsilon \to \circ} c_\epsilon \Pr(g(\boldsymbol{\theta}, \mathbf{u}) \in B_{d,\epsilon}(\mathbf{y}_\circ) \mid \boldsymbol{\theta})$$

می‌توان نوشت. که در آن c_ϵ یک ضریب تناسب وابسته به ϵ است و $B_{d,\epsilon}(\mathbf{y}_\circ)$ ناحیه‌ای به شعاع ϵ اطراف $\mathbf{y}_\circ$ است که با یک تابع فاصله d به‌صورت $B_{d,\epsilon}(\mathbf{y}_\circ) := \{\mathbf{y} : d(\mathbf{y}, \mathbf{y}_\circ) \leq \epsilon\}$ تعریف می‌شود.

در مواردی که $\mathbf{y}$ بعد بالا باشد از آماره خلاصه T استفاده و سپس فاصله d اعمال می‌شود. در این حالت، $B_{d,\epsilon}(\mathbf{y}_\circ) := \{\mathbf{y} : d(T(\mathbf{y}), T(\mathbf{y}_\circ)) \leq \epsilon\}$ خواهد شد. محاسبه $\Pr\left(g(\boldsymbol{\theta}, \mathbf{u}) \in B_{d,\epsilon}(\mathbf{y}_\circ) \mid \boldsymbol{\theta}\right)$ به عنوان کسری از نمونه‌هایی که در ناحیه ϵ اطراف $\mathbf{y}_\circ$ قرار دارند، در حالت حدی $\epsilon \to \circ$ از نظر محاسباتی غیرممکن است. بنابراین، محدودیت به $\epsilon > \circ$ کاهش می‌یابد که منجر به درست‌نمایی تقریبی $L_{d,\epsilon}(\boldsymbol{\theta}) = \Pr(\mathbf{y} \in B_{d,\epsilon}(\mathbf{y}_\circ) \mid \boldsymbol{\theta})$ می‌شود. در نتیجه، توزیع پسین تقریبی را می‌توان به‌صورت $p_{d,\epsilon}(\boldsymbol{\theta} \mid \mathbf{y}_\circ) \propto L_{d,\epsilon}(\boldsymbol{\theta}) p(\boldsymbol{\theta})$ نوشت. این روش تنها راه‌حل برای مواجهه با مشکل محاسباتی تابع درست‌نمایی نیست، بلکه رویکردهای دیگر شامل مدل‌سازی رابطه (تصادفی) بین $\boldsymbol{\theta}$ و $\mathbf{y}$، معکوس آن، یا تعریف استنباط بدون درست‌نمایی به عنوان یک مسئله براورد نسبت هستند (توماس و همکاران، ۲۰۲۲؛ هرمانس و همکاران، ۲۰۲۰). با این حال، روش‌های OMC و ROMC که در ادامه بیان می‌شوند، بر اساس تقریب ارائه‌شده در معادله درست‌نمایی تقریبی استوار هستند.

[3] Robust OMC

۱.۲ مونت‌کارلوی بهینه (OMC)

روش OMC برای اولین بار توسط میدز و ولینگ (۲۰۱۵) ارایه شده است. برای این منظور ابتدا تابع نشانگر $\mu_{B_{d,\epsilon}(\mathbf{y})}(\mathbf{x})$ را به‌صورت

$$\mu_{B_{d,\epsilon}(\mathbf{y})}(\mathbf{x}) = \begin{cases} ۱ & \text{اگر } \mathbf{x} \in B_{d,\epsilon}(\mathbf{y}) \\ ۰ & \text{در غیر این صورت} \end{cases} \tag{۱.۲}$$

تعریف می‌کنیم، پس، تابع درست‌نمایی تقریبی $L_{d,\epsilon}(\boldsymbol{\theta})$ را می‌توان به‌صورت زیر نوشت:

$$L_{d,\epsilon}(\boldsymbol{\theta}) = \Pr(\mathbf{y} \in B_{d,\epsilon}(\mathbf{y}_۰) \mid \boldsymbol{\theta}) = \int_{\mathbf{u}} \mu_{B_{d,\epsilon}(\mathbf{y}_۰)}(g(\boldsymbol{\theta}, \mathbf{u}))d\mathbf{u}.$$

در آن صورت با استفاده از نمونه‌های مونت‌کارلویی برگرفته از شبیه‌ساز به‌صورت زیر تقریب زده می‌شود:

$$L_{d,\epsilon}(\boldsymbol{\theta}) \approx \frac{۱}{N}\sum_{i=۱}^{N} \mu_{B_{d,\epsilon}(\mathbf{y}_۰)}(g(\boldsymbol{\theta}, \mathbf{u}_i)), \quad \mathbf{u}_i \sim p(\mathbf{u}). \tag{۲.۲}$$

در معادله (۲.۲)، برای هر $\mathbf{u}_i$، ناحیه‌ای به‌صورت $C_\epsilon^i = \{\boldsymbol{\theta} : g(\boldsymbol{\theta}, \mathbf{u}_i) \in B_{d,\epsilon}(\mathbf{y}_۰)\}$ در فضای پارامتر $\boldsymbol{\theta}$ وجود دارد که تابع نشانگر مقدار یک را برمی‌گرداند. بنابراین، می‌توان درست‌نمایی تقریبی و توزیع پسین را به صورت زیر نوشت:

$$L_{d,\epsilon}(\boldsymbol{\theta}) \approx \frac{۱}{N}\sum_{i=۱}^{N} \mu_{C_\epsilon^i}(\boldsymbol{\theta}), \qquad p_{d,\epsilon}(\boldsymbol{\theta} \mid \mathbf{y}_۰) \propto p(\boldsymbol{\theta}) \sum_{i=۱}^{N} \mu_{C_\epsilon^i}(\boldsymbol{\theta}). \tag{۳.۲}$$

در واقع، تقریب درست‌نمایی و توزیع پسین به توصیف مجموعه‌های C_ϵ^i تبدیل می‌شود. در روش OMC ، فرض می‌کند که فاصله d، فاصله اقلیدسی $\|\cdot\|_۲$ بین آمار خلاصه T داده‌های مشاهده‌شده و تولیدشده است و نواحی C_ϵ^i می‌توانند به خوبی با بیضی‌های بسیار کوچک تقریبی شوند. این فرضیات منجر به تقریب توزیع پسین با استفاده از نمونه‌های وزنی $\boldsymbol{\theta}_i^*$ می‌شوند که کمترین فاصله بین داده‌های مشاهده‌شده و شبیه‌سازی‌شده را برای هر تحقق $\mathbf{u}_i \sim p(\mathbf{u})$ به‌صورت $\boldsymbol{\theta}_i^* = \underset{\boldsymbol{\theta}}{\operatorname{argmin}} \|\Phi(\mathbf{y}_۰) - \Phi(g(\boldsymbol{\theta}, \mathbf{u}_i))\|_۲$ به‌دست می‌آورند.

۲.۲ روش استوار مونت‌کارلوی بهینه (ROMC)

گولمیس و همکاران (۲۰۱۶) نشان دادند که فرض بیضی‌های بسیار کوچک می‌تواند به توزیع‌های پسین بیش از حد مطمئن منجر شود. به طور شهودی، این مشکل به این دلیل رخ می‌دهد که وزن‌ها در OMC تنها بر اساس اطلاعات محلی در $\boldsymbol{\theta}_i^*$ محاسبه می‌شوند و استفاده صرف از اطلاعات محلی می‌تواند گمراه‌کننده باشد. برای مثال، اگر انحنای $\|T(\mathbf{y}_۰) - T(g(\boldsymbol{\theta}, \mathbf{u}_i))\|_۲$ در $\boldsymbol{\theta}_i^*$ تقریباً صاف باشد، ممکن است به اشتباه نشان دهد که C_ϵ^i بسیار بزرگ‌تر از اندازه واقعی آن است. روش استوار OMC پیشنهادی این مشکل را برطرف می‌کند. ROMC نواحی پذیرش C_ϵ^i را با جعبه‌های مرزی D-بعدی $\hat{C}_\epsilon^i$ تقریبی می‌کند. یک توزیع یکنواخت $q_i(\boldsymbol{\theta})$ روی هر جعبه مرزی تعریف می‌شود و به عنوان توزیع پیشنهادی برای تولید نمونه‌های پسین $\boldsymbol{\theta}_{ij} \sim q_i$ عمل می‌کند. نمونه‌ها وزن w_{ij} دریافت می‌کنند که استفاده از توزیع‌های پیشنهادی $q_i(\boldsymbol{\theta})$ به جای پیشین $p(\boldsymbol{\theta})$ را تعدیل می‌کند:

$$w_{ij} = \mu_{C_\epsilon^i}(\boldsymbol{\theta}_{ij}) \frac{p(\boldsymbol{\theta}_{ij})}{q_i(\boldsymbol{\theta}_{ij})}. \tag{۴.۲}$$

با توجه به نمونه‌های وزنی، هر امید $\mathbb{E}_{p(\boldsymbol{\theta}|\mathbf{y}_\circ)}[h(\boldsymbol{\theta})]$ از یک تابع $h(\boldsymbol{\theta})$ می‌تواند به صورت زیر تقریبی شود:

$$\mathbb{E}_{p(\boldsymbol{\theta}|\mathbf{y}_\circ)}[h(\boldsymbol{\theta})] \approx \frac{\sum_{ij} w_{ij} h(\boldsymbol{\theta}_{ij})}{\sum_{ij} w_{ij}}. \tag{۵.۲}$$

برای محاسبه وزن w_{ij} در معادله ۴.۲، باید بررسی کنیم که آیا نمونه‌های $\boldsymbol{\theta}_{ij}$، که از جعبه‌های مرزی گرفته شده‌اند، در ناحیه پذیرش C_ϵ^i قرار دارند. این بررسی به عنوان یک مکانیزم ایمنی عمل می‌کند که نادرستی‌های احتمالی در ساخت $\hat{C}_\epsilon^i$ را اصلاح می‌کند. اما این بررسی نیازمند ارزیابی تابع فاصله $d_i(\boldsymbol{\theta}_{ij})$ است که در مدل‌های پیچیده ممکن است پرهزینه باشد.

به‌طور کلی، ROMC به دو بخش آموزش و استنباط تقسیم می‌شود. بخش آموزش شامل گام‌هایی برای تخمین نواحی پیشنهادی است و بخش استنباط نمونه‌های وزنی را محاسبه می‌کند. در واقع ابتدا، نواحی پیشنهادی $\hat{C}_\epsilon^i$ تعیین و برآورد می‌شوند و سپس توزیع پسین تقریب زده می‌شود. در ادامه یک مطالعه شبیه‌سازی برای پیاده‌سازی روش پیشنهادی روی داده‌های وابسته ارائه می‌شود.

۳ مطالعه شبیه‌سازی

در این بخش، روش پیشنهادی را ابتدا روی مدل میانگین متحرک مرتبه دوم (MA2)، که یکی از مدل‌های فرایندهای سری زمانی است، پیاده‌سازی و ارزیابی می‌کنیم. سپس روی یک مدل رگرسیون فضایی ساده اجرا و تحلیل می‌شود. استنباط با سه رویکرد مختلف روش ROMC انجام می‌شود: (۱) با بهینه‌ساز مبتنی بر گرادیان، (۲) با بهینه‌سازی بیزی و (۳) با برازش یک شبکه عصبی به عنوان مدل جایگزین. مورد آخر نشان‌دهنده قابلیت توسعه روش پیشنهادی است، که در آن بخشی از ROMC با یک روش مشخص توسط کاربر جایگزین می‌شود.

۱.۳ مدل میانگین متحرک مرتبه دوم

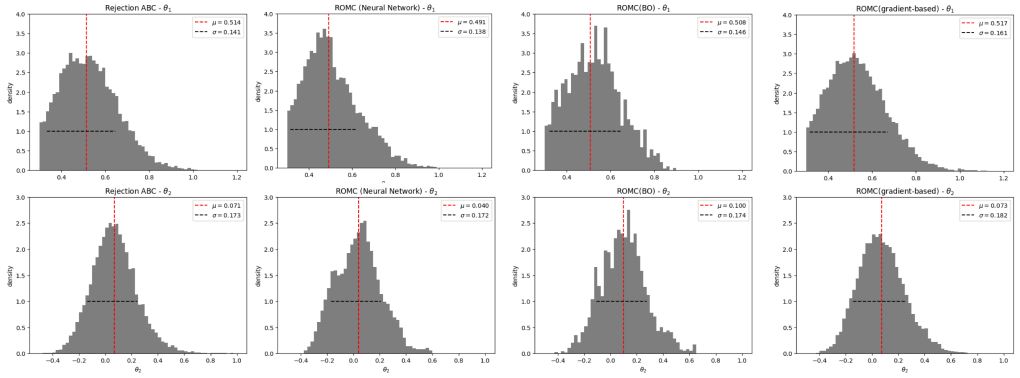
مدل MA2 به صورت زیر تعریف می‌شود:

$$y_t = w_t + \theta_1 w_{t-1} + \theta_2 w_{t-2}, \quad t = 1, \ldots, K \tag{۱.۳}$$

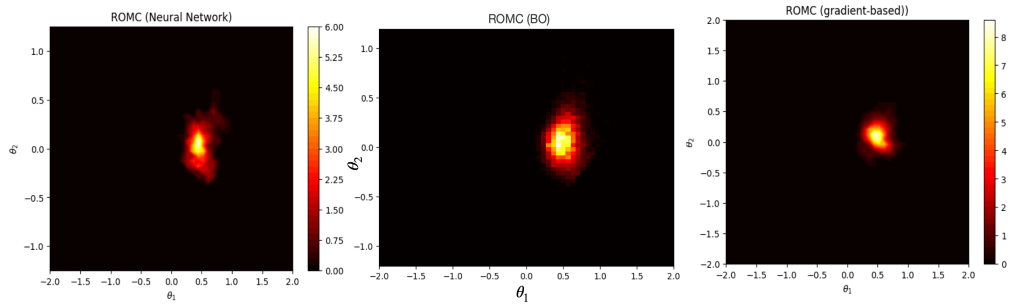
که در آن $\theta_1, \theta_2 \in \mathbb{R}$ و $w_t \sim \mathcal{N}(0, 1)$، است. متغیر تصادفی w_k نوفه سفید است و دو پارامتر مورد نظر، θ_1 و θ_2، وابستگی به مشاهدات قبلی را مدل می‌کنند. پارامتر K تعداد مشاهدات متوالی است که $K = 100$ در نظر گرفته شده است. برای اطمینان از قابل‌شناسایی بودن از پیشین پیشنهادی مارین و همکاران (۲۰۱۲) به‌صورت $p(\boldsymbol{\theta}) = p(\theta_1)p(\theta_2 \mid \theta_1)$ $\mathcal{U}(\theta_1; -2, 2)\mathcal{U}(\theta_2; \theta_1 - 1, \theta_1 + 1)$ استفاده می‌کنیم. بردار مشاهده $\mathbf{y}_\circ = (y_1, \ldots, y_{100})$ با $\boldsymbol{\theta}^* = (0.6, 0.2)$ تولید شده است. به دلیل بالا بودن بُعد خروجی $\mathbf{y}$، از آماره خلاصه استفاده می‌کنیم. با توجه به اینکه بردار خروجی یک تحقق سری‌زمانی را نشان می‌دهد، از اتوکواریانس‌ها با تأخیر ۱ و ۲ به‌صورت زیر استفاده شده است:

$$\begin{aligned} s_1(\mathbf{y}) &= \frac{1}{T-1}\sum_{t=2}^{T} y_t y_{t-1}, \quad s_2(\mathbf{y}) = \frac{1}{T-2}\sum_{t=3}^{T} y_t y_{t-2}, \\ s(\mathbf{y}) &= (s_1(\mathbf{y}), s_2(\mathbf{y})), \quad d = \|s(\mathbf{y}) - s(\mathbf{y}_\circ)\|_2^2, \end{aligned} \tag{۲.۳}$$

فاصله بین مشاهده و خروجی شبیه‌ساز با فاصله اقلیدسی، طبق معادله ۲.۳، اندازه‌گیری می‌شود. همچنین برای شبکه عصبی، از کلاس `MLPRegressor` از بسته `scikit-learn` استفاده می‌کنیم. که در آن $\hat{d}_i$ جایگزین تابع فاصله واقعی d_i در نواحی پیشنهادی



شکل ۱: هیستوگرام‌های توزیع‌های پسین حاشیه‌ای برای هر روش



شکل ۲: توزیع پسین برای سه نسخه مختلف روش ROMC

می‌شود. بنابراین، تمام اقدامات استنباط، یعنی نمونه‌گیری، محاسبه امید، و ارزیابی پسین، بر اساس $\hat{d}_i$ انجام می‌شود. در این‌جا از یک شبکه عصبی با دو لایه مخفی، هر یک با ۱۰ نورون، استفاده می‌کنیم و آن را با ۵۰۰ نمونه از هر ناحیه پیشنهادی آموزش می‌دهیم. برای مقایسه نتایج استنباط ROMC با الگوریتم سنتی بیز تقریبی ABC، نیز در تحلیل آورده شده است. در جدول ۱، میانگین و انحراف معیار نمونه‌های پسینی پارمترها برای هر روش ارائه شده است. مشاهده می‌شود که نمونه‌های به‌دست‌آمده از توزیع پسین تقریبی در تمام روش‌ها معادل و نزدیک به مقدار واقعی هستند. در شکل ۱، هیستوگرام‌های توزیع‌های پسین حاشیه‌ای پارامترها برای همه روش‌ها

جدول ۱: مقایسه نمونه‌های به‌دست‌آمده از توزیع پسین تقریبی با نمونه‌گیری رد و نسخه‌های مختلف ROMC

	$\hat{\mu}_{\theta_1}$	$\hat{\sigma}_{\theta_1}$	$\hat{\mu}_{\theta_2}$	$\hat{\sigma}_{\theta_2}$
نمونه‌گیری رد ABC	۰/۵۱۶	۰/۱۴۲	۰/۰۷۰	۰/۱۷۲
ROMC (مبتنی بر گرادیان)	۰/۵۰۱	۰/۱۴۲	۰/۰۳۳	۰/۱۶۹
ROMC (بهینه‌سازی بیزی)	۰/۵۱۳	۰/۱۶۹	۰/۰۹۰	۰/۱۷۴
ROMC (شبکه عصبی)	۰/۴۹۱	۰/۱۳۸	۰/۰۴۰	۰/۱۷۲

رسم شده‌اند که نشان می‌دهد توزیع پارامترها در همه روش‌ها نسبتا مشابه به هم هستند. در شکل ۲، توزیع پسین برای سه نسخه مختلف روش ROMC ارائه شده است. نتایج نشان می‌دهند که تمام نسخه‌های ROMC نتایج نسبتا مناسب و یکسانی ارائه می‌دهند که با الگوریتم نمونه‌گیری رد ABC هم‌خوانی دارند. با موازی کردن این روش می‌توان سرعت محاسبات را تا ۵ برابر افزایش داد.

۲.۳ مدل رگرسیون فضایی

در این بخش، روش پیشنهادی را روی مدل رگرسیون فضایی در یک مشبکه منظم ۱۰ × ۱۰ با مختصات $s_i = (i, j)$ و $i, j = ۱, \ldots, ۱۰$ پیاده‌سازی و ارزیابی می‌کنیم. مشابه بخش قبل استنباط با سه رویکرد مختلف روش ROMC انجام می‌شود. مدل رگرسیون فضایی به صورت زیر تعریف شده است:

$$y(s_i) = \beta_۰ + \beta_۱ x(s_i) + w(s_i) + \varepsilon_i, \quad i = ۱, \ldots, ۱۰۰,$$

که در آن $x(s_i)$ یک متغیر توضیحی در مکان s_i است که از توزیع یکنواخت $\mathcal{U}(۰, ۱۰)$ تولید شده است، $\beta = (\beta_۰, \beta_۱)$ بردار ضرایب رگرسیون، $w(s_i)$ فرآیند گاوسی با کوواریانس نمایی $\mathrm{Cov}(w(s_i), w(s_j)) = \sigma^۲ \exp(-\phi \|s_i - s_j\|)$ با $\sigma^۲ = ۱$ و $\varepsilon_i \sim \mathcal{N}(۰, \tau^۲)$ نوفه سفید است. فاصله $\|s_i - s_j\|$ فاصله اقلیدسی بین مکان‌های s_i و s_j روی مشبکه منظم است. پارامترهای مورد نظر برای استنباط $\boldsymbol{\theta} = (\beta_۰, \beta_۱, \phi, \tau^۲)$ هستند. از پیشین‌های معمول زیر استفاده می‌کنیم:

$$p(\boldsymbol{\theta}) = p(\beta_۰)p(\beta_۱)p(\phi)p(\tau^۲) = \mathcal{N}(\beta_۰; ۰, ۱۰)\mathcal{N}(\beta_۱; ۰, ۱۰)\mathcal{U}(\phi; ۰/۱, ۵)\mathcal{IG}(\tau^۲; ۲, ۱). \qquad (۳.۳)$$

که در آن $\mathcal{IG}$ توزیع گامای معکوس است. بردار مشاهده $\mathbf{y}_۰ = (y(s_۱), \ldots, y(s_{۱۰۰}))$ با $\boldsymbol{\theta}^* = (۱, ۰/۵, ۱, ۰/۱)$ تولید شده است. به دلیل بالا بودن بعد مشاهدات $\mathbf{y}$، از آماره‌های خلاصه استفاده می‌کنیم. با توجه به ماهیت داده‌های فضایی، آماره‌های خلاصه شامل میانگین و واریوگرام تجربی به‌صورت زیر استفاده شده است:

$$T_۱(\mathbf{y}) = \frac{۱}{n}\sum_{i=۱}^{n} y(s_i), \quad T_۲(\mathbf{y}) = \frac{۱}{|\mathcal{N}|}\sum_{(i,j)\in\mathcal{N}} (y(s_i) - y(s_j))^۲,$$
$$T(\mathbf{y}) = (T_۱(\mathbf{y}), T_۲(\mathbf{y})), \quad d = \|T(\mathbf{y}) - T(\mathbf{y}_۰)\|_۲^۲,$$

که در آن $\mathcal{N}$ مجموعه جفت‌های مکان‌هایی است که فاصله اقلیدسی آن‌ها کمتر از ۲ واحد است. در واقع از همسایگی‌های نزدیک مشابه استفاده کرده‌ایم. در این‌جا از یک شبکه عصبی با دو لایه مخفی، هر یک با ۱۲ نورون، استفاده می‌کنیم و آن را با ۵۰۰ نمونه از هر ناحیه پیشنهادی آموزش می‌دهیم. برای مقایسه نتایج استنباط ROMC با الگوریتم سنتی نمونه‌گیری رد ABC، نیز در این تحلیل آورده شده است. در جدول ۲، میانگین و انحراف معیار نمونه‌های پسینی پارامترهای $\beta_۰$، $\beta_۱$، ϕ، و $\tau^۲$ برای همه روش‌ها ارائه شده است. مشاهده می‌شود که نمونه‌های به‌دست‌آمده از توزیع پسین تقریبی در تمام روش‌ها معادل و تقریبا به مقدار واقعی نزدیک هستند. نتایج

جدول ۲: مقایسه نمونه‌های به‌دست‌آمده از توزیع پسین تقریبی با نمونه‌گیری رد و نسخه‌های مختلف ROMC

	$\hat{\mu}_{\beta_۰}$	$\hat{\sigma}_{\beta_۰}$	$\hat{\mu}_{\beta_۱}$	$\hat{\sigma}_{\beta_۱}$	$\hat{\mu}_{\phi}$	$\hat{\sigma}_{\phi}$	$\hat{\mu}_{\tau^۲}$	$\hat{\sigma}_{\tau^۲}$
نمونه‌گیری رد ABC	۰/۹۹۲	۰/۱۴۸	۰/۴۹۵	۰/۰۸۵	۱/۰۱۰	۰/۲۱۲	۰/۰۹۸	۰/۰۲۵
ROMC (مبتنی بر گرادیان)	۰/۹۹۸	۰/۱۴۵	۰/۴۹۹	۰/۰۸۳	۱/۰۰۷	۰/۲۰۸	۰/۰۹۹	۰/۰۲۴
ROMC (بهینه‌سازی بیزی)	۰/۹۸۹	۰/۱۵۰	۰/۴۹۴	۰/۰۸۷	۱/۰۱۳	۰/۲۱۵	۰/۰۹۷	۰/۰۲۶
ROMC (شبکه عصبی)	۰/۹۸۷	۰/۱۴۳	۰/۴۹۰	۰/۰۸۲	۱/۰۰۵	۰/۲۱۰	۰/۰۹۶	۰/۰۲۳

نشان می‌دهند که تمام نسخه‌های ROMC نتایج منسجمی ارائه می‌دهند که با الگوریتم نمونه‌گیری رد ABC هم‌خوانی دارند.

۴ بحث و نتیجه‌گیری

در این مقاله، روش استنباط بدون درست‌نمایی ROMC برای تحلیل داده‌های وابسته، از جمله سری‌های زمانی و داده‌های فضایی، ارائه شد. این روش با استفاده از سه رویکرد مختلف پیاده‌سازی و ارزیابی شد. نتایج مطالعه شبیه‌سازی روی مدل میانگین متحرک مرتبه دوم و مدل رگرسیون فضایی نشان داد که هر سه نسخه ROMC نمونه‌های پسینی با میانگین و انحراف معیار نزدیک به مقادیر واقعی تولید می‌کنند و با الگوریتم نمونه‌گیری رد ABC نتایج مشابهی دارند. همچنین، موازی‌سازی روش ROMC می‌تواند سرعت محاسبات را تا ۵ برابر افزایش دهد، که نشان‌دهنده کارایی بالای این روش در مسائل پیچیده است.

جنبه اصلی این روش این است که ROMC یک چارچوب انعطاف‌پذیر و قابل گسترش ارائه می‌دهد که به کاربران اجازه می‌دهد تا بخش‌های مختلف روش، مانند تابع فاصله یا الگوریتم بهینه‌سازی، را با رویکردهای دلخواه جایگزین کنند و این امکان استفاده گسترده‌تر و توسعه بیشتر توسط پژوهشگران را فراهم می‌کند. با این حال، چالش‌هایی برای بهبود کارایی ،ROMC به‌ویژه در مسائل با ابعاد بالا، همچنان وجود دارد. مهم‌ترین چالش، امکان اجرای ROMC در محیط‌های توزیع‌شده، مانند خوشه‌های محاسباتی، است که می‌تواند زمان محاسبات را به‌طور قابل توجهی کاهش دهد. علاوه بر این، تنظیم تعداد نورون‌ها و لایه‌های شبکه عصبی برای مسائل پیچیده‌تر یا بهینه‌سازی انتخاب آماره‌های خلاصه برای داده‌های فضایی می‌تواند دقت تقریب توزیع پسین را بهبود بخشد. در آینده، تمرکز بر توسعه الگوریتم‌های موازی و بهینه‌سازی‌های عددی می‌تواند قابلیت‌های ROMC را برای کاربردهای گسترده‌تر تقویت کند.

مراجع

Gutmann, M. U. and Corander, J. (2016), *Bayesian Optimization for Likelihood-Free Inference of Simulator-Based Statistical Models*, *Journal of Machine Learning Research*, **17**, 1–47

Gkolemis, V., Gutmann, M. U. and Pesonen, H. (2024), *An Extendable Python Implementation of Robust Optimization Monte Carlo*, *Journal of Statistical Software*, **110**, 1–26.

Hermans, J., Begy, V. and Louppe, G. (2020), *Likelihood-Free MCMC with Amortized Approximate Ratio Estimators*, *International Conference on Machine Learning*, 4239–4248, PMLR.

Marin, J. M., Pudlo, P., Robert, C. P. and Ryder, R. J. (2012), *Approximate Bayesian Computational Methods*, *Statistics and Computing*, **22**, 1167–1180.

Meeds, T. and Welling, M. (2015), *Optimization Monte Carlo: Efficient and Embarrassingly Parallel Likelihood-Free Inference*, *Advances in Neural Information Processing Systems*, **28**.

Thomas, O., Dutta, R., Corander, J., Kaski, S. and Gutmann, M. U. (2022), *Likelihood-Free Inference by Ratio Estimation*, *Bayesian Analysis*, **17**, 1–31.

Robust Optimal Monte Carlo Method

Omid Karimi, Fatemeh Hosseini

Department of Statistics, Semnan University, Semnan, Iran

Abstract: In most cases in the analysis of time-series and spatial processes, we encounter complex statistical models whose likelihood functions are not easily computable. For this reason, many likelihood-free inference methods face limitations in terms of accuracy and efficiency. In this paper, we present a likelihood-free inference method using an optimal robust Monte Carlo approach for analyzing spatially dependent data. This approach provides a novel and efficient framework for likelihood-free inference, offering accurate weighted samples from the posterior distribution. Finally, the proposed method is implemented and evaluated on simulated data.

Keywords: Bayesian inference, dependent data, Monte Carlo methods, likelihood-free inference.

Mathematics Subject Classification (2020): 65C05, 90C26, 62H11.
