

روش‌های مونت کارلویی همیلتونی برای تحلیل مدل آمیخته خطی تعمیم‌یافته فضایی چوله

فاطمه حسینی^۱، امید کریمی

تاریخ دریافت: ۱۴۰۰/۰۵/۰۶

تاریخ پذیرش: ۱۴۰۰/۰۷/۱۵

چکیده:

برای مدل‌بندی پاسخ‌های فضایی گسسته، مدل‌های آمیخته خطی تعمیم‌یافته فضایی استفاده می‌شود. در این مدل‌ها همبستگی فضایی داده‌ها به صورت متغیرهای پنهان فضایی وارد مدل می‌شود. معمولاً برای سادگی فرض می‌شود که متغیرهای پنهان دارای توزیع نرمال هستند که نادرست بودن این فرض بر روی دقت نتایج تأثیرگذار است. در این مقاله متغیرهای پنهان با میدان تصادفی چوله گاوسی بسته مدل‌بندی می‌شوند که بزرگ‌تر و انعطاف‌پذیرتر از میدان تصادفی گاوسی می‌باشد. یک الگوریتم جدید برای به دست آوردن برآورد ماکسیمم درست‌نمایی پارامترها معرفی می‌شود. اساس الگوریتم معرفی شده بر مبنای الگوریتم بیشینه‌سازی امیدریاضی و نوعی الگوریتم مونت کارلویی همیلتونی است. کارایی و سرعت الگوریتم معرفی شده در یک مثال شبیه‌سازی بررسی می‌شود.

واژه‌های کلیدی: مدل‌های آمیخته خطی تعمیم‌یافته فضایی، الگوریتم مونت کارلویی همیلتونی، میدان تصادفی چوله گاوسی بسته.

۱ مقدمه

[۱۰] اشاره نمود. یکی دیگر از مشکلات مدل‌های آمیخته خطی تعمیم‌یافته فضایی وجود متغیرهای پنهان فضایی در مدل است که برای سادگی و خواص خوب توزیع گاوسی از میدان تصادفی گاوسی برای مدل‌بندی این متغیرها استفاده می‌شود. در عمل بررسی برقراری فرض نرمال بودن متغیرهای پنهان امکان‌پذیر نیست و در صورت عدم برقراری نرمال بودن برآوردها و پیشگویی‌ها از دقت کافی برخوردار نیست. یکی از توزیع‌هایی که انعطاف‌پذیر است و شامل توزیع نرمال، خانواده توزیع چوله نرمال است. اولین بار توزیع چوله نرمال را [۲] معرفی نمود. همچنین [۴] و [۳] به‌طور مفصل‌تر به مطالعه‌ی خانواده توزیع چوله نرمال چند متغیره و خواص آن پرداختند. با توجه به مشکلات و عدم برقراری برخی خواص توزیع گاوسی [۹] یک خانواده توزیع بزرگ‌تر به نام چوله نرمال بسته را معرفی نمودند که شامل خانواده توزیع چوله نرمال و نرمال می‌باشد و دارای خواص بسیار خوبی است، ([۱۱]، [۱۳]، [۹]، [۱۲]، [۲۶]). در مطالعات فضایی اولین بار [۲۰] میدان تصادفی چوله گاوسی را معرفی نمود. [۲۱]، [۲۲] و [۲۹] برای مدل‌های فضایی چوله میدان تصادفی چوله گاوسی بسته را معرفی کردند. همچنین در مدل‌های رگرسیونی غیرخطی [۱] و [۳۴] پیشنهاد دادند که خطاهای تصادفی در این مدل‌ها با توزیع چوله نرمال مدل‌بندی شود. [۲۳] در مدل‌های رگرسیونی فضایی خطاها را به صورت توزیع چوله نرمال بسته در نظر گرفتند. [۱۴]، [۱۵]، [۱۸] و [۱۹] در مدل‌های آمیخته خطی تعمیم‌یافته فضایی استفاده از توزیع چوله نرمال بسته

در مدل‌های آمیخته خطی تعمیم‌یافته فضایی معمولاً متغیرهای پنهان فضایی موجود در مدل را با میدان تصادفی گاوسی و متغیرهای پاسخ فضایی که رسته‌ای و ناگاوسی فرض می‌شوند با خانواده توزیع‌های نمایی مدل‌بندی می‌نمایند. نرمال بودن متغیرهای پنهان از پذیره‌های معمول بر روی این مدل‌ها است که به‌عنوان مثال می‌توان به کارهای [۷]، [۸]، [۳۶]، [۳۱]، [۳۵] اشاره نمود.

به دلیل ناگاوسی بودن متغیرهای پاسخ و وجود متغیرهای پنهان در مدل، به دست آوردن برآورد ماکسیمم درست‌نمایی در این مدل‌ها به دلیل عدم وجود شکل بسته برای تابع درست‌نمایی به‌راحتی امکان‌پذیر نمی‌باشد و به همین دلیل تاکنون کارهای تحقیقاتی زیادی در این زمینه انجام و برای حل این مشکل رهیافت‌هایی ارائه شده‌اند. [۳۶] یک الگوریتم بر اساس الگوریتم بیشینه‌سازی امیدریاضی و تولید نمونه‌های مونت کارلویی معرفی نمود. [۵] و [۶] بر مبنای الگوریتم همسانه‌سازی داده‌ها یک رهیافت جدید برای به دست آوردن برآورد ماکسیمم درست‌نمایی پیشنهاد کردند. [۱۶] و [۱۷] با ترکیب الگوریتم بیشینه‌سازی امیدریاضی، درست‌نمایی دوبه‌دویی و استفاده از همسایگی‌های الگوریتم‌هایی برای استنباط پیرامون مدل‌های آمیخته خطی تعمیم‌یافته فضایی ارائه نمودند. از کارهای تحقیقاتی دیگر که با در نظر گرفتن متغیرهای پنهان به صورت نرمال به تحلیل این مدل‌ها پرداختند می‌توان به [۳۲]، [۳۳] و

$$\times \Phi_q(\Gamma(x - \mu); \nu, \Delta), \quad (1)$$

باشد، که در آن $k = [\Phi_q(\cdot; \nu, \Delta + \Gamma \Sigma \Gamma^T)]^{-1}$ بردار پارامتر مکان، ماتریس $n \times n$ معین مثبت Σ ماتریس مقیاس، عناصر ماتریس $\Gamma_{q \times n}$ پارامترهای چولگی هستند. $\Phi_q(\Gamma(x - \mu); \nu, \Delta)$ تابع توزیع تجمعی q متغیره نرمال با بردار میانگین ν و ماتریس واریانس کوواریانس Δ است. به‌طور خلاصه این توزیع به‌صورت $CSN_{n,q}(\mu, \Sigma, \Gamma, \nu, \Delta)$ برای $q = 1$ ، $\nu = 0$ ، $\Delta = 1$ و $\Gamma = \gamma^T \Sigma^{-\frac{1}{2}}$ تابع چگالی چوله نرمال بسته همان تابع چگالی توزیع چوله نرمال خواهد شد. وقتی Γ ماتریس صفر تعریف شود، تابع چگالی نرمال حاصل می‌شود. گشتاور مرتبه‌ی اول توزیع چوله نرمال بسته به‌صورت

$$E(X) = \mu + \Sigma \Gamma^T \Psi, \quad (2)$$

به دست می‌آید که در آن $\Psi = \frac{\Phi_q^*(\cdot; \nu, \Delta + \Gamma \Sigma \Gamma^T)}{\Phi_q(\cdot; \nu, \Delta + \Gamma \Sigma \Gamma^T)}$ است. برای ماتریس معین مثبت Ω ،

$$\Phi_q^*(s; \nu, \Omega) = [\nabla_s \Phi_q(s; \nu, \Omega)]^T,$$

که در آن $\nabla_s = (\frac{\partial}{\partial s_1}, \dots, \frac{\partial}{\partial s_q})^T$ می‌باشد. واریانس توزیع چوله نرمال بسته به‌صورت

$$\begin{aligned} \text{Var}(X) &= E(XX^T) - (E(X))(E(X))^T, \\ &= \Sigma + \Sigma \Gamma^T \Lambda \Gamma \Sigma - \Sigma \Gamma^T \Psi \Psi^T \Gamma \Sigma \end{aligned} \quad (3)$$

است، که در آن $\Lambda = \frac{\Phi_q^{**}(\cdot; \nu, \Delta + \Gamma \Sigma \Gamma^T)}{\Phi_q(\cdot; \nu, \Delta + \Gamma \Sigma \Gamma^T)}$ و $\Phi_q^{**}(s; \nu, \Omega) = \nabla_s \nabla_s^T \Phi_q(s; \nu, \Omega)$ تعریف می‌شوند.

۲.۲ میدان تصادفی چوله گاوسی بسته

[۲۹] میدان تصادفی چوله گاوسی بسته را به‌صورت زیر تعریف کرد:

تعریف ۱۰.۲. فرض کنید $U(s) = \{(U_1(s), U_2(s))^T, s \in D \subseteq R^d\}$ میدان تصادفی گاوسی دو متغیره باشد و $U_2 = [U_2(s'_1), \dots, U_2(s'_q)]$ که در آن q ثابت و متناهی است. آنگاه میدان تصادفی چوله گاوسی بسته به‌صورت $\{X(s) = [U_1(s)|U_2 \leq \cdot]\}$ تعریف می‌شود، اگر برای هر مجموعه متناهی (s_1, \dots, s_n) $X = (X(s_1), \dots, X(s_n))^T$ دارای توزیع چوله نرمال بسته باشد.

[۲۹] نشان دادند که وقتی موقعیت‌های (s'_1, \dots, s'_q) به‌اندازه‌ی کافی از مرزها دور، $n = q$ و $(s_1, \dots, s_n) = (s'_1, \dots, s'_n)$ باشند و با در نظر گرفتن شکل توزیع چوله نرمال بسته برای X به‌صورت

$$CSN_{n,n}(\mu, \sigma^2 C, \frac{\gamma}{\sigma} I_n, \nu | \nu, (1 - \gamma^2) I_n), \quad (4)$$

که در آن μ پارامتر مکان، σ^2 پارامتر مقیاس، $|\gamma| < 1$ پارامتر چولگی، C ماتریس همبستگی ایستا، I_n ماتریس واحد و $\nu \in R^n$ برداری با عناصر یک است، آنگاه $X(s)$ تقریباً ایستا است.

برای متغیرهای پنهان فضایی را پیشنهاد کردند. در اکثر منابع اشاره‌شده در بالا برای تولید نمونه از الگوریتم‌هایی مثل متروپلیس هاستینگس استفاده‌شده است. چون تابع درستمایی در مدل‌های آمیخته خطی تعمیم‌یافته فضایی بسیار پیچیده است، الگوریتم‌هایی مثل متروپلیس هاستینگس زمان‌بر و گاهی زنجیر به‌سختی همگرا می‌شود. [۲۷] و [۲۸] پیشنهاد کردند که در مدل‌های پیچیده و داده با بعد بالا بهتر است به‌جای استفاده از الگوریتم‌های متروپلیس هاستینگس از الگوریتم مونت‌کارلوی همیلتونی که بر اساس معادلات همیلتونی است، استفاده شود. در راستای کاهش زمان محاسبات در داده‌های بزرگ [۳۷] و [۳۸] استفاده از روش‌های مونت‌کارلوی همیلتونی را موردبررسی قرار دادند. هدف از ارائه این مقاله در راستای افزایش دقت برآوردها و کاهش زمان محاسبات، از میدان تصادفی چوله گاوسی بسته برای مدل‌بندی متغیرهای پنهان فضایی استفاده می‌شود و پس از معرفی مدل برای به دست آوردن برآورد ماکسیمم درستمایی پارامترها دو الگوریتم جدید به نام الگوریتم بیشینه‌سازی امیدریاضی مونت‌کارلویی همیلتونی و الگوریتم بیشینه‌سازی امیدریاضی مونت‌کارلویی معرفی می‌شوند و در یک مثال شبیه‌سازی مورد مقایسه قرار می‌گیرند. ساختار مقاله به این صورت است که در بخش دوم مدل‌های آمیخته خطی تعمیم‌یافته فضایی با متغیرهای پنهان چوله نرمال بسته و میدان تصادفی چوله گاوسی بسته که در این مقاله استفاده‌شده است، معرفی می‌شوند. در بخش چهارم دو الگوریتم بیشینه‌سازی امیدریاضی مونت‌کارلویی همیلتونی و الگوریتم بیشینه‌سازی امیدریاضی مونت‌کارلویی ارائه می‌گردد. در نهایت یک مثال شبیه‌سازی در بخش پنجم مورد مطالعه قرار گرفته است.

۲ مدل‌های آمیخته خطی تعمیم‌یافته فضایی چوله

در این بخش ابتدا توزیع چوله نرمال بسته و سپس میدان تصادفی چوله گاوسی بسته [۲۹] بیان و برخی خواص آن ارائه می‌گردد. در نهایت مدل‌های آمیخته خطی تعمیم‌یافته فضایی با متغیرهای پنهان چوله نرمال بسته معرفی می‌شوند.

۱۰.۲ توزیع چوله نرمال گاوسی بسته

بردار تصادفی n بعدی x دارای توزیع چوله نرمال بسته چند متغیره با پارامترهای μ, Σ, Γ, ν است اگر تابع چگالی آن به‌صورت

$$f_{n,q}(x|\mu, \Sigma, \Gamma, \nu, \Delta) = k \phi_n(x; \mu, \Sigma)$$

۳.۲ مدل

فرض کنید $x = (x_1, \dots, x_n)^T$ یک تحقق از میدان تصادفی گاوسی چوله بسته در n موقعیت $\{s_1, \dots, s_n\}$ با دامنه $D \subseteq \mathbb{R}^n$ و دارای توزیع چوله نرمال به صورت (۴) با پارامتر مکان $\mu = H\beta$ باشد، که در آن H ماتریسی با بعد $n \times (p+1)$ شامل متغیرهای کمکی و β بردار پارامترهای رگرسیونی است. برای ماتریس C ساختار فضایی نمایی همسانگرد

$$C(i, j) = \exp(-\|s_i - s_j\| / \varphi) \quad (5)$$

فرض می‌شود، که در آن $\|s_i - s_j\|$ فاصله ی اقلیدسی بین s_i و s_j ، φ پارامتر دامنه فضایی است. از رابطه ی (۱) تابع چگالی x به صورت

$$f(x|\eta) = \phi_n(x; H\beta, \sigma^2 C) \times \frac{\prod_{i=1}^n \Phi(\frac{\gamma}{\sigma}(x_i - \mu_i); \nu, 1 - \gamma^2)}{\phi_n(\sigma; \nu | n, (1 - \gamma^2)I_n + \gamma^2 C)} \quad (6)$$

به دست می‌آید، که در آن $\mu_i = h_i^T \beta$ ، $h_i = (1, h_{i1}, \dots, h_{ip})^T$ ، $\eta = (\beta^T, \sigma, \varphi, \nu, \gamma)^T$ و H کمکی $n \times (p+1)$ بردار پارامترهای مدل است. فرض کنید $y^T = (y_1, \dots, y_k)$ بردار متغیرهای پاسخ فضایی در موقعیت‌های $\{s_1, \dots, s_k\}$ ، $k \leq n$ باشد. در عمل هدف پیش‌گویی متغیرهای پنهان و متغیرهای پاسخ در موقعیت‌های بدون مشاهده $\{s_{k+1}, \dots, s_n\}$ است. متغیرهای پنهان در موقعیت‌های دارای اندازه‌ی متغیر پاسخ را با $x^{obs} = Ax$ نشان می‌دهند، که در آن $A = [I_{k \times k} \mid 0_{k \times (n-k)}]$ ماتریس A بردار x را به دو بردار $x = (x^{obs}, x^{pred})^T$ تجزیه می‌کند. بردار متغیرهای پنهان در $(n-k)$ موقعیت فاقد مشاهده‌ی پاسخ است. فرض کنید متغیرهای پاسخ دارای تابع چگالی $f(y|x)$ و مطابق [۲۵] به‌طور شرطی مستقل از یک خانواده‌ی توزیع نمایی باشد. بنابراین در موقعیت s_i ، $i = 1, \dots, k$ می‌توان نوشت

$$f(y_i|x_i) = \exp\{y_i x_i - b(x_i) + c(y_i)\},$$

که در آن تابع کومولانت و مطابق مدل‌های خطی تعمیم‌یافته $E(y_i|x_i) = g^{-1}(x_i)$ که در آن $g(\cdot)$ یک تابع پیوند معلوم است. پس تابع درستنمایی

$$L(\eta|y) = \int \left\{ \prod_{i=1}^k f(y_i|x_i) \right\} f(x|\beta, \theta) dx \quad (7)$$

می‌شود. اکنون از رابطه‌ی (۶) تابع درستنمایی را می‌توان به صورت

$$L(\eta|y) = \int \exp\left\{ \sum_{i=1}^k (y_i x_i - b(x_i) + c(y_i)) - \frac{1}{2\sigma^2} (x - H\beta)^T C^{-1} (x - H\beta) \right\} \times \frac{\prod_{i=1}^n \Phi(\frac{\gamma}{\sigma}(x_i - \mu); \nu, 1 - \gamma^2)}{\phi_n(\sigma; \nu | n, (1 - \gamma^2)I_n + \gamma^2 C)} dx, \quad (8)$$

نوشت، که شکل بسته‌ای ندارد و برای به دست آوردن برآوردهای ماکسیمم درستنمایی نمی‌توان از روش‌های معمول استفاده کرد و نیاز به استفاده از الگوریتم‌های تکرارشونده است.

۳ برآورد پارامترها

در این بخش دو الگوریتم برای به دست آوردن برآورد ماکسیمم درستنمایی پارامترها معرفی می‌شوند. ابتدا الگوریتم معرفی شده در [۳۶] که بر مبنای نمونه‌های مونت‌کارلویی متروپلیس هاستینگس و الگوریتم پیشینه‌سازی امیدریاضی است را برای برآورد مدل معرفی شده در این مقاله تعمیم داده و سپس یک الگوریتم جدید با سرعت همگرایی بیشتر معرفی می‌شود.

الگوریتم MCEMG-CSN

گام ۱- مقدار اولیه برای پارامترها به صورت $\eta^{(0)}$ در نظر گرفته شود، به طوری که $L(\eta^{(0)}|y) > 0$ و قرار داده شود $m = 0$.

گام ۲- مرحله‌ی تولید نمونه:

نمونه‌های مونت‌کارلویی $x^{(1)}, \dots, x^{(N)}$ از $f_{x|y}(x|y, \eta^{(m)})$ با به کار بردن الگوریتم متروپلیس هاستینگس به صورت زیر استخراج شود.

- مقدار شروع $x^{(0)}$ در نظر گرفته شود.

- برای $t = 1 : N$

یک نامزد تولید شود $x^{(t-1)}$

تولید شود $u \sim U[0, 1]$

محاسبه شود $r(x^{(t)}, x^{(t-1)}) = \frac{q(x^{(t)}|x^{(t-1)})f(x^{(t)}|\eta^{(m)})f(y|x^{(t)})}{q(x^{(t-1)}|x^{(t)})f(x^{(t-1)}|\eta^{(m)})f(y|x^{(t-1)})}$

اگر $u < \min(r(x^{(t)}, x^{(t-1)}), 1)$ سپس

$$x^{(t)} = x^{(t-1)}$$

در غیر این صورت

$$x^{(t)} = x^{(t-1)}$$

پایان

- پایان

گام ۳- مرحله‌ی امیدریاضی:

با به کار بردن نمونه‌های مونت‌کارلویی $x^{(1)}, \dots, x^{(N)}$ محاسبه شود: $E\left\{\frac{\partial \ln f(x|\eta)}{\partial \eta}\right\}|y$ ، و $E\left\{\frac{\partial^2 \ln f(x|\eta)}{\partial \eta \partial \eta'}\right\}|y$

$$E\left\{\frac{\partial \ln f(x|\eta)}{\partial \eta \partial \eta'}\right\}|y = \frac{1}{N} \sum_{t=1}^N \left(\frac{\partial \ln f(x^{(t)}|\eta^{(m)})}{\partial \eta \partial \eta'} \right),$$

$$E\left\{\frac{\partial \ln f(x|\eta)}{\partial \eta}\right\}|y = \frac{1}{N} \sum_{t=1}^N \left(\frac{\partial \ln f(x^{(t)}|\eta^{(m)})}{\partial \eta} \right). \quad (9)$$

گام ۴- مرحله‌ی پیشینه‌سازی:

با به کار بردن الگوریتم نیوتن-رافسون و رابطه‌ی (۹) $\eta^{(m+1)}$

به صورت زیر محاسبه شود:

$r \sim N(0, M)$ تولید می‌شود. سپس توزیع توأم (θ, r) بر اساس مقدار کنونی پارامتر θ و مقدار تولید شده r ساخته می‌شود و معادلات همیلتونی زیر تکامل می‌یابد:

$$\frac{d\theta}{dt} = +\frac{\partial H}{\partial r} = +\frac{\partial K}{\partial r}$$

$$\frac{dr}{dt} = -\frac{\partial H}{\partial \theta} = -\frac{\partial K}{\partial \theta} - \frac{\partial U}{\partial \theta}$$

با در نظر گرفتن استقلال بین چگالی کمکی و چگالی هدف یعنی $p(r|\theta) = p(r)$ ، $\partial K/\partial \theta$ صفر می‌شود و معادلات به صورت زیر ساده می‌شوند:

$$\frac{d\theta}{dt} = +\frac{\partial K}{\partial r}$$

$$\frac{dr}{dt} = -\frac{\partial U}{\partial \theta}$$

مرحله بعدی الگوریتم مونت کارلوی همیلتونی به نام انتگرال‌گیری جهشی^۴ است که در این مرحله با استفاده از روش‌های عددی انتگرال‌گیری، معادلات سیستمی همیلتونی فراهم می‌شود. مشابه خیلی از روش‌های انتگرال‌گیری عددی برای این الگوریتم جهشی فاصله‌های زمانی کوچک به میزان δ در نظر گرفته می‌شود. این الگوریتم جهشی مقادیر θ و r را به طور مستقل از θ و مقدار قبلی r به صورت

$$r \leftarrow r - \frac{\delta}{\gamma} \frac{\partial U}{\partial \theta}$$

$$\theta \leftarrow \theta + \delta M^{-1} r$$

$$r \leftarrow r - \frac{\delta}{\gamma} \frac{\partial U}{\partial \theta} \quad (10)$$

به روزرسانی می‌کند. با L تکرار سه معادله فوق مقادیر (r^*, θ^*) حاصل می‌شوند. مقادیر (θ^*, r^*) در مرحله پذیرش متروپلیس به صورت

$$\min(1, \exp(H(r, \theta) - H(r^*, \theta^*)))$$

پذیرفته یا رد می‌شود. بر اساس الگوریتم مونت کارلوی همیلتونی که یک الگوریتم جدید برای به دست آوردن برآورد ماکسیمم درستمایی پارامترها معرفی می‌شود. در این الگوریتم معادلات همیلتونی برای تولید نامزد جدید در الگوریتم متروپلیس استفاده می‌شود و این امر باعث افزایش سرعت و رسیدن سریع‌تر به همگرایی نسبت به الگوریتم نمونه‌گیری گیبز و متروپلیس هاستینگس می‌شود. فرض کنید هدف تولید نمونه از توزیع پسین برای توضیحات بیشتر به [۲۷]، [۳۰]، [۲۸] و [۲۴] مراجعه شود.

مطابق [۲۸] در الگوریتم مونت کارلویی همیلتونی تابع انرژی جنبشی

^۲Kinetic energy function

^۳Potential energy function

^۴Leapfrog integrator

$$\eta^{(m+1)} = \eta^{(m)} - \left[\frac{1}{N} \sum_{t=1}^N \left(\frac{\partial^r \ln f(\mathbf{x}^{(t)}|\eta^{(m)})}{\partial \eta \partial \eta'} \right) \right]^{-1} \left[\frac{1}{N} \sum_{t=1}^N \left(\frac{\partial \ln f(\mathbf{x}^{(t)}|\eta^{(m)})}{\partial \eta} \right) \right],$$

گام ۵- قرار داده شود $m = m + 1$. به مرحله‌ی (۲) برگرد و تا همگرایی حاصل شود.

۱۰۳ الگوریتم مونت کارلوی همیلتونی

الگوریتم مونت کارلوی همیلتونی یک روش مونت کارلوی زنجیر مارکوفی است که از مشتق‌های تابع چگالی مورد نظر برای تولید نمونه‌های کارا استفاده می‌کند. این روش از یک تقریب دینامیکی همیلتونی بر اساس انتگرال‌های عددی استفاده می‌کند که به وسیله مرحله پذیرش متروپلیس اصلاح می‌شود. فرض کنید پارامتر θ و مشاهدات y هستند و هدف تولید نمونه از توزیع پسین $p(\theta|y)$ است. الگوریتم مونت کارلوی همیلتونی از متغیر کمکی r به وسیله توزیع توأم

$$p(r, \theta) = p(r|\theta)p(\theta|y)$$

نمونه‌گیری از پسین را انجام می‌دهد، که اغلب از توزیع نرمال چند متغیره به عنوان چگالی کمکی استفاده می‌شود که به پارامتر θ وابسته نیست و آن را به صورت $r \sim N(0, M)$ در نظر می‌گیرند، که در آن M یک ماتریس معین مثبت است و بر اساس کوواریانس نمونه‌های تولید شده در مرحله داغیدن تعیین می‌شود. با استفاده از توزیع توأم $p(r, \theta)$ معادله دینامیکی همیلتونی به صورت

$$H(r, \theta) = -\log p(r, \theta)$$

$$= -\log p(r|\theta) - \log p(\theta|y).$$

$$= K(r|\theta) + U(\theta),$$

در نظر گرفته می‌شود که در آن

$$K(r|\theta) = -\log p(r|\theta)$$

و به آن انرژی جنبشی^۲ می‌گویند و منفی لگاریتم چگالی

$$U(\theta) = -\log p(\theta|y)$$

را انرژی نهفته^۳ می‌نامند. یک تغییر وضعیت به مرحله جدید با شروع از مقادیر کنونی پارامتر θ توسط دو مرحله زیر قبل از مرحله پذیرش متروپلیس صورت می‌گیرد. ابتدا مقادیری برای r از توزیع

$$E\left\{\frac{\partial \ln f(\mathbf{x}|\boldsymbol{\eta})}{\partial \boldsymbol{\eta}}|\mathbf{y}\right\} = \frac{1}{N} \sum_{t=1}^N \left(\frac{\partial \ln f(\mathbf{x}^{(t)}|\boldsymbol{\eta}^{(m)})}{\partial \boldsymbol{\eta}}\right). \quad (۱۲)$$

گام ۴- **مرحله‌ی بیشینه‌سازی:**

از الگوریتم نیوتن رافسون و رابطه‌ی (۱۲) را به صورت زیر به دست آورید:

$$\boldsymbol{\eta}^{(m+1)} = \boldsymbol{\eta}^{(m)} - \left[\frac{1}{N} \sum_{t=1}^N \left(\frac{\partial^2 \ln f(\mathbf{x}^{(t)}|\boldsymbol{\eta}^{(m)})}{\partial \boldsymbol{\eta} \partial \boldsymbol{\eta}'}\right)\right]^{-1} \left[\frac{1}{N} \sum_{t=1}^N \left(\frac{\partial \ln f(\mathbf{x}^{(t)}|\boldsymbol{\eta}^{(m)})}{\partial \boldsymbol{\eta}}\right)\right]$$

گام ۵- قرار دهید $m = m + 1$. تا رسیدن به همگرایی برگرد به مرحله‌ی ۲ و الگوریتم تکرار شود.

۴ پیش‌گویی فضایی

یکی از اهداف مهم در مدل‌های آمیخته خطی تعمیم‌یافته فضایی پیش‌گویی متغیرهای پنهان، x^{pred} در موقعیت‌های $\{s_{k+1}, \dots, s_n\}$ است. متغیرهای پنهان در $n - k$ موقعیت فاقد مشاهده‌ی پاسخ به صورت $x^{pred} = A_1 x$ در آن $A_1 = [I_{(n-k) \times k} | I_{(n-k) \times (n-k)}]^\circ$

قضیه ۱۰۴. فرض کنید $x = (x^{obs \top}, x^{pred \top})^\top$ دارای توزیع چوله نرمال بسته به صورت

$$CSN_{n,n}(H\beta, \sigma^2 C_\theta, \frac{\gamma}{\sigma} I_n, \nu, \nu, (1 - \gamma^2) I_n), \quad (۱۳)$$

و $x | y$ متعلق به خانواده‌ی توزیع نمایی باشد، آنگاه پیش‌گویی فضایی با حداقل میانگین توان دوم خطا برای x^{pred} در $n - k$ موقعیت $\{s_{k+1}, \dots, s_n\}$ از رابطه‌ی

$$E(x^{pred} | y, \boldsymbol{\eta}) = \boldsymbol{\mu}_p + C_{p \circ} C_{\circ \circ}^{-1} (E(x^{obs} | y, \boldsymbol{\eta}) - \boldsymbol{\mu}_\circ) + \frac{\gamma}{\sigma} \boldsymbol{\Sigma}_p \boldsymbol{\Gamma}_1 \boldsymbol{\Psi}_p, \quad (۱۴)$$

به دست می‌آید، که در آن

$$\begin{aligned} \boldsymbol{\Sigma}_p &= \sigma^2 (C_{pp} - C_{p \circ} C_{\circ \circ}^{-1} C_{\circ p}), \\ \boldsymbol{\Psi}_p &= \boldsymbol{\Psi}(\circ, \boldsymbol{\Sigma}_p, \frac{\gamma}{\sigma} \boldsymbol{\Gamma}_1, \nu_p, (1 - \gamma^2) I_n) \\ \nu_p &= \nu I_n - \frac{\gamma}{\sigma} (\boldsymbol{\Gamma}_1 + \boldsymbol{\Gamma}_2 C_{p \circ} C_{\circ \circ}^{-1}) (x^{obs} - \boldsymbol{\mu}_\circ) \end{aligned} \quad (۱۵)$$

و ماتریس‌های $C_{\circ \circ}, C_{\circ p}, C_{pp}, C_{\circ \circ}, \boldsymbol{\mu}_\circ, \boldsymbol{\mu}_p$ و $\boldsymbol{\Gamma}_1$ و $\boldsymbol{\Gamma}_2$ از افراز $H\beta$ به صورت زیر به دست می‌آیند:

$$H\beta = \begin{bmatrix} \boldsymbol{\mu}_\circ \\ \boldsymbol{\mu}_p \end{bmatrix}, C_\theta = \begin{bmatrix} C_{\circ \circ} & C_{\circ p} \\ C_{p \circ} & C_{pp} \end{bmatrix}, \frac{\gamma}{\sigma} I_n = \frac{\gamma}{\sigma} (\boldsymbol{\Gamma}_1, \boldsymbol{\Gamma}_2)$$

اثبات. با فرض این‌که $x = (x_{k \times 1}^{obs \top}, x_{(n-k) \times 1}^{pred \top})^\top$ دارای توزیع به صورت رابطه‌ی (۱۳) باشد، آنگاه طبق خاصیت بسته بودن توزیع

به صورت $K(r) = \frac{1}{\gamma} r^\top r$ در نظر می‌گیریم و مطابق [۳۷]، تابع انرژی نهفته به صورت منهای لگاریتم تابع چگالی احتمال هدف در نظر گرفته می‌شد. در این مقاله هدف نمونه‌گیری از تابع چگالی شرطی متغیرهای پنهان به شرط متغیرهای پاسخ و پارامترها است. بنابراین $U(x)$ به صورت

$$\begin{aligned} U(x) &= -\log f(x | y, \boldsymbol{\eta}) \\ &= \sum_{i=1}^k -\log f(y_i | x_i, \boldsymbol{\eta}) - \log f(x | \boldsymbol{\eta}), \quad (۱۱) \end{aligned}$$

تعریف می‌شود.

الگوریتم HMCEMG-CSN

گام ۱- مقدار اولیه به صورت $\boldsymbol{\eta}^{(0)}$ با شرط $L(\boldsymbol{\eta}^{(0)} | y) > 0$ در نظر گرفته و $m = 1$ قرار داده شود.

گام ۲- **مرحله‌ی تولید نمونه:**

نمونه‌های مونت‌کارلوی $x^{(1)}, \dots, x^{(N)}$ از $f_{x|y}(x|y, \boldsymbol{\eta}^{(m)})$ با به کار بردن الگوریتم مونت‌کارلوی همیلتون زیر تولید شود:

• مقدار اولیه $x^{(1)}$ ؛ اندازه‌ی گام δ تنظیم شود.

• **برای** $t = 1 : N$

تولید شود $r^{(t)} \sim N(0, I)$ قرار دهید $(x_\circ, r_\circ) = (x^{(t)}, r^{(t)})$

از معادلات پویای همیلتونی گسسته‌سازی انجام شود:

برای $l = 1 : L$

$$r_{l-1} \leftarrow r_{l-1} - \frac{\delta}{\gamma} \nabla U(x_{l-1})$$

$$x_l \leftarrow x_{l-1} + \delta r_{l-1}$$

$$r_l \leftarrow r_{l-1} - \frac{\delta}{\gamma} \nabla U(x_l)$$

پایان

$$(x^*, r^*) = (x_L, r_L)$$

تصحیح متروپولیس هاستینگس انجام شود:

استخراج کنید $\alpha \sim U[0, 1]$

قرار دهید $r = \exp[H(x^{(t)}, r^{(t)}) - H(x^*, r^*)]$

اگر $\alpha < \min\{1, r\}$ سپس

$$(x^{(t+1)}, r^{(t+1)}) = (x^*, r^*)$$

در غیر این صورت

$$(x^{(t+1)}, r^{(t+1)}) = (x^{(t)}, r^{(t)})$$

• **پایان**

گام ۳- **مرحله‌ی امیدریاضی:**

امیدریاضی‌های $E\left\{\frac{\partial \ln f(\mathbf{x}|\boldsymbol{\eta})}{\partial \boldsymbol{\eta}}|\mathbf{y}\right\}$ و $E\left\{\frac{\partial^2 \ln f(\mathbf{x}|\boldsymbol{\eta})}{\partial \boldsymbol{\eta} \partial \boldsymbol{\eta}'}|\mathbf{y}\right\}$ به صورت زیر برآورد شوند:

$$E\left\{\frac{\partial \ln f(\mathbf{x}|\boldsymbol{\eta})}{\partial \boldsymbol{\eta}}|\mathbf{y}\right\} = \frac{1}{N} \sum_{t=1}^N \left(\frac{\partial \ln f(\mathbf{x}^{(t)}|\boldsymbol{\eta}^{(m)})}{\partial \boldsymbol{\eta}}\right),$$

آمیخته خطی تعمیم‌یافته فضایی با متغیرهای پنهان چوله نرمال بسته و نرمال ارائه می‌شود.

چوله نرمال بسته تحت شرطی‌سازی $x^{pred} | x^{obs}, \eta$ دارای توزیع CSN به صورت

$$CSN \sim (\mu_p + C_p C_o^{-1} (x^{obs} - \mu_o), \Sigma_p, \frac{\gamma}{\sigma} \Gamma, \nu_p, (1 - \gamma^2) I_n), \quad (16)$$

۵ شبیه‌سازی

بر روی نقشه‌ی کشور ایران و به صورت یک مشبکه‌ی نامنظم $n = 200$ موقعیت تولید شده است. برای ماتریس ساختار C_φ همسانگرد نمایی با پارامتر دامنه همبستگی φ در نظر گرفته شده است. با در نظر گرفتن مقادیر زیر برای پارامترهای مدل

$$\beta_o = -2.5, \beta_1 = 0.8, \sigma^2 = 1, \varphi = 4, \gamma = 0.95,$$

متغیرهای پنهان از توزیع چوله نرمال بسته به صورت

$$(x|\eta) \sim CSN_{n,n}(\beta_o + \beta_1 h, \sigma^2 C_\varphi, \frac{\gamma}{\sigma}, \nu|n, (1 - \sigma^2) I_n),$$

استخراج شد، که در آن h بردار مقادیر استاندارد شده عرض جغرافیایی است که به عنوان متغیر کمکی وارد مدل می‌شود. پارامتر ν معلوم و صفر فرض شد.

متغیر پاسخ به شرط متغیرهای پنهان، $y_j, j = 1, \dots, 200$ ، بواسن به صورت

$$y_j \sim Poiss(p_j), p_j = \exp(x_j)$$

تولید می‌شود. این فرآیند شبیه‌سازی ۱۰۰۰ بار تکرار می‌شود. یک نمونه از موقعیت‌ها و داده‌های تولید شده در شکل ۱ رسم شده است. برنامه‌نویسی در نرم‌افزار متلب انجام شده است.

است، که پارامترهای آن در رابطه‌ی (۱۴) تعریف شده‌اند. برای به دست آوردن پیش‌گو با حداقل میانگین توان دوم خطا x^{pred} به صورت $(x^{pred} | y, \eta)$ باید امیدریاضی شرطی $(x^{pred} | x^{obs}, \eta)$ به دست آورده شود. از رابطه‌ی (۲):

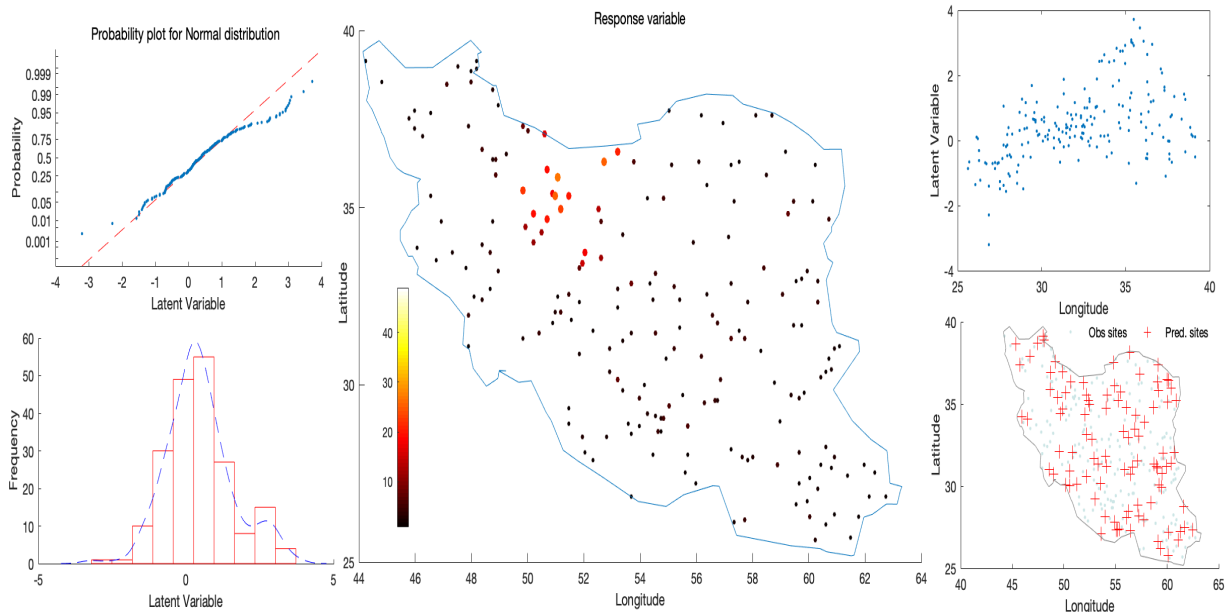
$$E(x^{pred} | x^{obs}, \eta) = \mu_p + C_p C_o^{-1} (x^{obs} - \mu_o) + \frac{\gamma}{\sigma} \Sigma_p \Gamma \Psi_p.$$

پس پیش‌گوی حداقل میانگین توان دوم خطا برای x^{pred} در $(n-k)$ موقعیت جدید به صورت

$$\begin{aligned} E(x^{pred} | y, \eta) &= E(E(x^{pred} | x^{obs}) | y, \eta) \\ &= E(\mu_p + C_p C_o^{-1} (x^{obs} - \mu_o) + \frac{\gamma}{\sigma} \Sigma_p \Gamma \Psi_p | y, \eta) \\ &= \mu_p + C_p C_o^{-1} (E(x^{obs} | y, \eta) - \mu_o) + \frac{\gamma}{\sigma} \Sigma_p \Gamma \Psi_p, \end{aligned} \quad (17)$$

به دست می‌آید. □

در قسمت بعد یک مطالعه‌ی شبیه‌سازی به منظور مقایسه‌ی کارایی دو الگوریتم HMCMEG و الگوریتم MCEMG در دو مدل



شکل ۱. یک تحقق از مدل بواسن فضایی شبیه‌سازی شده: راست - بالا) نمودار پراکنش y در مقابل عرض جغرافیایی، راست - پایین) موقعیت‌های انتخاب شده برای پیش‌گویی روی نقشه ایران وسط) داده‌های پاسخ بواسن تولید شده چپ - بالا) چندک - چندک متغیرهای پنهان تولید شده چپ - پایین) هیستوگرام متغیرهای پنهان.

تعریف می‌شوند. برای این منظور i امین نمونه حذف می‌شود و بر اساس سایر داده‌ها در ۱۹۹ موقعیت $\{s_1, \dots, s_{i-1}, s_{i+1}, \dots, s_{200}\}$ پیش‌گویی در i امین موقعیت به صورت $\hat{y}_{(i)}$ به دست می‌آید. پس از تکرار برای ۲۰۰ موقعیت مقادیر $CV-RRMSPE$ و $CV-MSPE$ برای دو مدل نرمال و چوله و دو الگوریتم $MCCEMG$ و $HMCEMG$ محاسبه و خلاصه‌ی نتایج در جدول ۲ ارائه شده است.

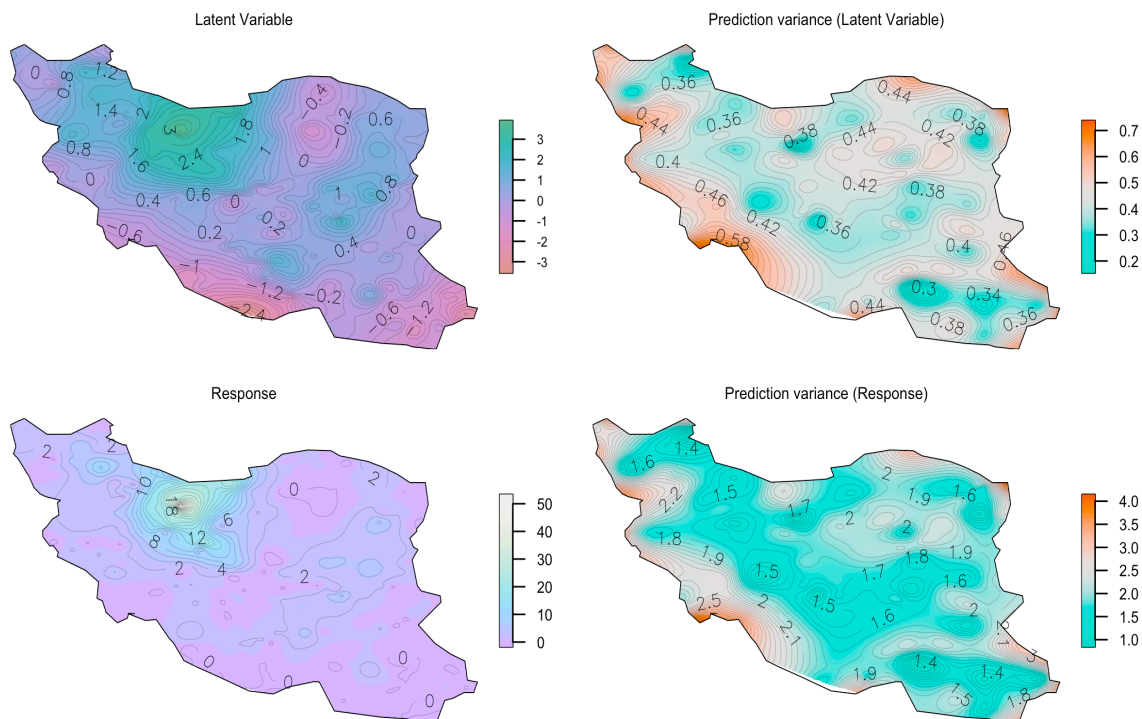
همان‌طور که ملاحظه می‌شود دقت پیش‌گویی مدل آمیخته خطی تعمیم یافته که متغیرهای پنهان به صورت چوله گاوسی در نظر گرفته شد بیشتر از مدل گاوسی است و همچنین دقت برآورد پارامترها و در نتیجه پیش‌گویی بر اساس الگوریتم همیلتونی مونت کارلویی بیشتر از الگوریتم متروپلیس هاستینگس و زمان اجرای الگوریتم همیلتونی کمتر از متروپلیس هاستینگس است.

برای به دست آوردن یک نقشه پیش‌گویی، در موقعیت‌های در نظر گرفته شده برای پیش‌گویی که در شکل ۱ نمودار راست-پایین مشخص شده است مقادیر پیش‌گویی محاسبه و در نهایت نقشه پیش‌گویی و واریانس پیش‌گویی برای متغیر پاسخ و متغیر پنهان در شکل ۲ رسم شده است.

برای ۱۰۰۰ مجموعه داده‌ی تولید شده برای مدل چوله گاوسی بسته تعریف شده و مدل گاوسی دو الگوریتم $MCCEMG$ و $HMCEMG$ اجرا می‌شوند. نتایج به طور خلاصه در جدول ۱ ارائه شده است. نتایج نشان می‌دهد که برای همه پارامترها متوسط میانگین خطا برآورد همه‌ی پارامترها برای مدل چوله گاوسی بسته کوچک‌تر از مدل گاوسی است. همچنین با مقایسه برآورد، انحراف معیار و مقادیر MSE محاسبه شده به نظر می‌رسد دقت نتایج الگوریتم $HMCEMG$ از $MCCEMG$ بیشتر است و زمان اجرای این الگوریتم به طور متوسط ۳۲ درصد کمتر از الگوریتم $MCCEMG$ بود. برای مقایسه دقت پیش‌گویی دو مدل چوله و نرمال و دو الگوریتم $MCCEMG$ و $HMCEMG$ از معیارهای اعتبارسنجی متقابل کم‌ترین توان‌های دوم خطای پیش‌گویی^۵ ($CV-MSPE$) و جذر اعتبارسنجی متقابل کم‌ترین توان‌های دوم خطای پیش‌گویی نسبی^۶ ($CV-RRMSPE$) استفاده شد که به صورت

$$CV-MSPE = n^{-1} \sum (y_i - \hat{y}_{(i)})^2,$$

$$CV-RRMSPE = \sqrt{n^{-1} \sum \{(y_i - \hat{y}_{(i)}) \hat{y}_{(i)}^{-1}\}^2}.$$



شکل ۲. نقشه پیش‌گویی و واریانس پیش‌گویی: بالا) متغیر پنهان، پایین) متغیر پاسخ

⁵Cross validation mean squared prediction error

⁶Relative root MSPE

جدول ۱. نتایج شبیه‌سازی بر اساس $200 \times 1000 \times 1000$ مجموعه داده‌ی ۲۰۰ تایی تولیدشده.

HMC-CSN			MC-CSN			R. val.	Par.
Ave. MSE	Ave. Std.	Ave. Es.	Ave. MSE	Ave. Std.	Ave. Es.		
۰/۶۹۰۷	۰/۱۰۰۱	-۲۳۱۷۸	۰/۸۱۲۵	۰/۱۳۲۱	-۲۲۲۳۱	-۲/۵	β_0
۰/۴۵۸۷	۰/۰۹۸۸	۰/۹۱۰۹	۰/۶۸۷۳	۰/۱۰۱۲	۱/۳۰۱۱	۰/۸	β_1
۰/۲۶۰۸	۰/۰۴۱۱	۱/۳۱۰۴	۰/۴۸۰۹	۰/۰۷۳۸	۱/۵۳۸۱	۱	σ^2
۳/۱۶۳۸	۰/۴۸۲۰	۴/۰۹۱۳	۳/۳۱۱۱	۰/۵۳۳۰	۴/۱۴۹۸	۴	φ
۰/۰۵۷۷	۰/۰۱۶۵	۰/۸۲۱۱	۰/۰۶۶۱	۰/۰۲۰۰	۰/۷۲۲۱	۰/۹۵	γ

HMC-Normal			MC-Normal			R. val.	Par.
Ave. MSE	Ave. Std.	Ave. Es.	Ave. MSE	Ave. Std.	Ave. Es.		
۰/۸۸۹۶	۰/۱۵۳۴	-۲/۱۲۰۷	۱/۰۹۶۶	۰/۱۶۳۶	-۲/۱۰۲۱	-۲/۵	β_0
۰/۷۷۸۱	۰/۱۰۱۳	۱/۴۶۸۴۵	۰/۹۲۱۱	۰/۱۴۸۸	۱/۵۱۱۱	۰/۸	β_1
۰/۶۶۵۵	۰/۰۹۲۳	۱/۵۶۶۳	۰/۷۷۱۱	۰/۱۱۳۹	۱/۸۶۶۷	۱	σ^2
۳/۷۹۷۲	۱/۱۵۱۱۰	۴/۸۳۰۹	۴/۲۴۵۸	۱/۳۵۱۹	۵/۲۳۷۷	۴	φ

جدول ۲. مقادیر اعتبارسنجی متقابل حداقل مربعات خطای پیش‌گویی و جذر اعتبارسنجی متقابل حداقل مربعات خطای پیش‌گویی

متغیر پاسخ		متغیر پنهان		مدل-الگوریتم
$CV - RRMSPE$	$CV - MSPE$	$CV - RRMSPE$	$CV - MSPE$	
۰/۰۱۱۹	۱/۷۳۰۳	۰/۰۴۷۳۰	۰/۳۱۰۰	MC - N
۰/۰۰۴۱	۱/۰۹۸۶	۰/۰۳۳۳	۰/۲۳۷۲	MC - CSN
۰/۰۰۹۲	۱/۳۹۶۳	۰/۰۴۰۰	۰/۲۹۸۶	HMC - N
۰/۰۰۲۰	۰/۹۷۷۱	۰/۰۱۹۹	۰/۲۰۰۵	HMC - CSN

از رهیافت درستمایی برای برآورد پارامترهای مدل آمیخته خطی تعمیم‌یافته فضایی با متغیرهای پنهان گاوسی و چوله گاوسی بسته استفاده شد. به دلیل عدم وجود شکل بسته برای تابع درستمایی یک الگوریتم جدید تکرار شونده با ترکیب الگوریتم مونت کارلوی همیلتونی و الگوریتم بیشینه‌سازی امیدریاضی معرفی شد. با توجه به نتایج شبیه‌سازی الگوریتم معرفی شده علاوه بر اینکه دارای سرعت اجرای بالایی است از دقت خوبی برخوردار است.

۶ بحث و نتیجه‌گیری

معمولاً برای داده‌های بزرگ و مدل‌ها با تعداد پارامتر زیاد، الگوریتم‌های معمول مثل گیز و متروپلیس هاستینگس به راحتی همگرا نمی‌شوند و استفاده از الگوریتم‌های مثل مونت کارلوی همیلتونی پیشنهاد شده است. مدل آمیخته خطی تعمیم‌یافته فضایی به دلیل وجود متغیرهای پنهان فضایی و عدم گاوسی بودن متغیرهای پاسخ یکی از مدل‌های پیچیده آماری می‌باشد. در این مقاله

مراجع

- [1] Alodat, M. T., Shakhathreh, M. K. (2020). Gaussian process regression with skewed errors *Journal of Computational and Applied Mathematics*, **370**, 112665.
- [2] Azzalini, A. (1985). A class of distributions which includes the normal ones. *Scandinavian journal of statistics* **12 (2)** 171-178.

- [3] Azzalini, A., Dalla-Valle, A. (1996). The multivariate skew-normal distribution. *Biometrika*, **83**, 715–726
- [4] Azzalini, A., Capitanio, A. (1999). Statistical applications of the multivariate skew normal distribution. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **61 (3)**, 579–602.
- [5] Baghishani, H., Mohammadzadeh, M. (2011). A data cloning algorithm for computing maximum likelihood estimates in spatial generalized linear mixed models, *Computational Statistics and Data Analysis*, **55**, 1748-1759.
- [6] Baghishani, H., Rue, H., Mohammadzadeh, M. (2012). On a Hybrid data cloning method and its application in generalized linear mixed models, *Statistics and Computing*, **22**, 597-613.
- [7] Breslow, N. E., Clayton, D. G. (1993). Approximate inference in generalized linear mixed models, *Journal of the American Statistical Association*, **88**, 9-25.
- [8] Diggle, P., Tawn, J.A., Moyeed, R. A. (1998). Model-based geostatistic, *Journal of the Royal Statistical Society, Series C. Applied Statistics*, **47**, 299-350.
- [9] Dominguez-Molina, J., Gonzalez-Farias, G., Gupta, A. (2003). The multivariate closed skew normal distribution. *Technical Report 03-12*, Department of mathematics and statistics, Bowling Green State University.
- [10] Evangelou, E., Zhu, Z., Smith, R. L. (2011). Estimation and prediction for spatial generalized linear mixed models using high order Laplace approximation. *Journal of Statistical Planning and Inference*, **141 (11)**, 3564-3577.
- [11] Gonzalez-Farias, G., Dominguez-Molina, A. and Gupta, A. K. (2004). The closed skew normal distribution. In: *Genton M. G., ed. Skew-elliptical distributions and their applications: A journey beyond normality. Boca Raton, FL: Chapman and Hall CRC*, 2542.
- [12] Gonzalez-Farias, G., Dominguez-Molina, A., Gupta, A. K. (2004). Additive properties of skew normal random vectors, *Journal of Statistical Planning and Inference*, **126: 2**, 521-534.
- [13] Gupta, A.K., Gonzales-Farias, G., Dominguez-Molina, J.A. (2003). A multivariate skew normal distribution. *Journal of Multivariate Analysis*, **89**, 181-190.
- [14] Hosseini, F., Eidsvik, J., Mohammadzadeh, M. (2011). Approximate Bayesian inference in Spatial GLMM with Skew Normal Latent Variables, *Computational Statistics and Data Analysis*, **55**, 1791-1806.
- [15] Hosseini, F., Mohammadzadeh, M. (2012). Bayesian prediction for spatial GLMM's with Closed Skew Normal latent variables, *Australian & New Zealand Journal of Statistics*, **54**, 43-62.
- [16] Hosseini, F. (2016). A new algorithm for estimating the parameters of the spatial generalized linear mixed models, *Environmental and Ecological Statistics*, **23**, 205-217.
- [17] Hosseini, F., Karimi, O. (2019). Approximate composite marginal likelihood inference in spatial generalized linear mixed models, *Journal of Applied Statistics* **46(3)**, 542-558.
- [18] Hosseini, F., Karimi, O. (2020). Approximate likelihood Inference in Spatial Generalized Linear Mixed Models with Closed Skew Normal latent Variables, *Communication in Statistics- Simulation and Computation* **49**, 121-134.
- [19] Hosseini, F., and Karimi, O. (2021) Approximate pairwise likelihood inference in SGLM models with skew normal latent variables, *Journal of Computational and Applied Mathematics*, **398**, 113692.
- [20] Kim, H.M., Mallick, B.K. (2004). A Bayesian prediction using the skew Gaussian distribution. *Journal of Statistical Planning and Inference*, **120**, 85–101.

- [21] Karimi, O., Omre, H., Mohammadzadeh, M. (2010). Bayesian Closed-skew Gaussian Inversion of Seismic AVO Data for Elastic Material Properties, *Geophysics*, **75**, R1-R11.
- [22] Karimi, O., Mohammadzadeh, M. (2011). Bayesian Spatial Prediction for Discrete Closed Skew Gaussian Random Field, *Mathematical Geosciences* **43**, 565–582.
- [23] Karimi, O., Mohammadzadeh, M. (2012). Bayesian spatial regression models with closed skew normal correlated errors and missing, *Statistical Papers* **53(1)**, 205-218.
- [24] Betancourt, M. (2018). A conceptual introduction to hamiltonian monte carlo. arXiv preprint arXiv:1701.02434, 2018. <https://arxiv.org/abs/1701.02434>.
- [25] McCullagh, P., Nelder, J. A. (1989). *Generalized l Linear Models*, Chapman and Hall, London.
- [26] Nadarajah, S., Li, R. (2017). The exact density of the sum of independent skew normal random variables, *Journal of Computational and Applied Mathematics*, **311**, 1–10.
- [27] Neal, R. M. (1992). Bayesian training of backpropagation networks by the hybrid monte carlo method. *University of Toronto Technical Report CRG-TR-92-1*.
- [28] Neal, R.M. (2010). MCMC using hamiltonian dynamics. *Handbook of Markov Chain Monte Carlo*, **2**, 113–162.
- [29] Rimstad, K., Omre, H. (2014). Skew-Gaussian Random Fields, *Spatial Statistics* **10**, 43-62.
- [30] Duane, S., Kennedy, A.D., Pendleton, B. J., and Roweth, D. (1987). Hybrid monte carlo. *Physics letters B*, **195(2)**, 216–222.
- [31] Smith, L. M., Stroup, W. W., Marx, D. B. (2020). Poisson cokriging as a generalized linear mixed model, *Spatial Statistics*, **35**, 100399.
- [32] Torabi, M. (2013). Likelihood inference in generalized linear mixed measurement error models, *Computational Statistics and Data Analysis*, **57**, 549-557.
- [33] Torabi, M. (2015). Likelihood inference for spatial generalized linear mixed models. *Communications in Statistics-Simulation and Computation*, **44**, 1692-1701.
- [34] Vicente G. C., Victor H. L., Edwin M. M. O. (2010). A nonlinear regression model with skew-normal errors, *Statistical Papers*, **51**, 547-558.
- [35] Walder, A., Hanks, E. M. (2020), Bayesian analysis of spatial generalized linear mixed models with Laplace moving average random fields, *Computational Statistics & Data Analysis*, **144**, 106861.
- [36] Zhang, H. (2002). On estimation and prediction for spatial generalized linear mixed models, *Biometrics*, **58**, 129-136.
- [37] Zhang, C., Shahbaba, B., Zhao, H. (2018). Variational Hamiltonian Monte Carlo via score matching *Bayesian Analysis* **13 (2)**, 485-506.
- [38] Zhang, C., Shahbaba, B., Zhao, H. (2017). Hamiltonian Monte Carlo acceleration using surrogate functions with random bases. *Statistics and Computing*, **27**, 1473–1490.

Hamiltonian Monte Carlo Methods for Analysing Skew GLM models

Fatemeh Hosseini and Omid Karimi

Semnan University, Department of Statistics

Abstract:

Spatial generalized linear mixed models are used commonly for modeling discrete spatial responses. In these models the spatial correlation of the data is considered as spatial latent variables. For simplicity, it is usually assumed in these models that spatial latent variables are normally distributed. An incorrect normality assumption may lead to inaccurate results and is therefore erroneous. In this paper we model the spatial latent variables in a general random field, namely the closed skew Gaussian random field which is more flexible and includes the Gaussian random field. We propose a new algorithm for maximum likelihood estimates of the parameters. A key ingredient in our algorithm is using a Hamiltonian Monte Carlo version of the EM algorithm. The performance of the proposed model and algorithm is presented through a simulation study.

Keywords: Spatial Generalized Linear Mixed Model, Hamiltonian Monte Carlo Algorithm, Closed Skew Gaussian Random Field.

Mathematics Subject Classification (2010): 91B72, 62J12, 60G60.