انقباض مدل‌های اتورگرسیو فضایی بیزی بعد بالا

سیده نرگس ملائکه، محسن محمدزاده
گروه آمار، دانشگاه تربیت مدرس

چکیده: در این مقاله مدل اتورگرسیو فضایی با ماتریس مشخصه‌سازی فضایی نمایی، مطرح می‌شود. سپس از رهیافت هموارسازی بیزی برای مدل‌بندی و برآورد پارامترها استفاده می‌شود. در حقیقت رهیافت بیز سلسله‌مراتبی، از پیشین‌های انقباضی برای انتخاب متغیر بیزی در داده‌های بعد بالا استفاده می‌کند. در این مقاله دو پیشین انقباضی معرفی می‌شوند سپس عملکرد هر یک از آن‌ها با استفاده از مطالعه شبیه‌سازی ارزیابی می‌شود.

واژه‌های کلیدی: بعد بالا، انتخاب متغیر بیزی، پیشین‌های انقباضی، اتورگرسیو فضایی، بیز سلسله مراتبی .
کد موضوع‌بندی ریاضی (۲۰۱۰): 62M11، 62M30.

۱ مقدمه

در یک مدل آماری، وقتی تعداد متغیرهای تبیینی خیلی بیشتر از تعداد مشاهدات باشند، نمی‌توان از روش‌های متداول آماری مانند ماکسیمم درستنمایی یا کمترین توان‌های دوم عادی برای برآورد پارامترهای مدل استفاده کرد و لازم است روش‌های آماری بعد بالا را به‌کار گرفت. چالش اصلی داده‌های بعد بالا، تعداد زیاد پارامترهای مدل نسبت به حجم نمونه است که منجر به پارامتری شدن بیش از حد[1] مدل می‌شود. این موضوع در داده‌های فضایی از اهمیت بیشتری برخوردار است، چرا که به‌طور کلی همبستگی داده‌های فضایی موجب پیچیدگی بیشتر مدل می‌شود. برای حل این مشکل در مدل‌های اتورگرسیو فضایی[2] (SAR)، رویکرد میانگین–مدل بیزی توسط لسج و پرنت (۲۰۰۷) پیشنهاد شده است. در این رویکرد بجای استنباط بر اساس یک مدل، از میانگین موزون برآوردهای پارامتر، بر اساس ترکیب‌های زیادی از متغیرهای تبیینی استفاده می‌شود (کوپ، ۲۰۰۳). رهیافت بیزی نیز در تحلیل داده‌های بعد بالا به کرات مورد استفاده قرار گرفته است، زیرا این

[1] Overparametrization
[2] Spatial Autoregressive

* نام وآدرس پست الکترونیکی ارائه‌دهنده مقاله: سیده نرگس ملائکه، n.malaeke@modares.ac.ir

رهیافت از طریق بکارگیری اطلاعات پیشین داده‌ها در مدل، پیچیدگی برآورد پارامترها را کاهش می‌دهد. بعلاوه قدرت محاسباتی رهیافت بیزی، کاربرد آن را در تحلیل داده‌های بعد بالا افزایش داده است. فنون انتخاب متغیر گیبز[۳] (GVS)، میانگین مدل بیزی[۴] (BMA) و جستجوی تصادفی انتخاب متغیر[۵] (SSVS)، از روش‌های مختلف انتخاب متغیر بیزی هستند. اغلب روش‌های انتخاب متغیر بیزی با تعیین پیشین‌های تیر و تخته[۶] ضرایب، به الگوریتم‌های مونت کارلو زنجیر مارکوفی[۷] (MCMC) وابسته هستند. در روش‌های هموارسازی[۸] بیزی با مشخص کردن پیشین‌های انقباضی، انتخاب متغیر و برآورد پارامتر به‌طور همزمان فراهم می‌شود. این روش‌ها به‌طور طبیعی منجر به استنباط بیزی می‌شوند و مزایایی نسبت به سایر روش‌ها دارند (مالیک، ۲۰۱۳). در این روش‌ها، با استفاده از الگوریتم MCMC مقدار خطای استاندارد از توزیع پسین حاصل می‌شود و بازه باورمندی پارامترها به سادگی قابل محاسبه است. از طرفی پارامتر هموارساز[۹] که به‌عنوان توان تابع پیشین در نظر گرفته می‌شود (روزت و ژو، ۲۰۰۶)، با دیگر پارامترها تواماً برآورد می‌شود. در حقیقت، درجه انقباض از طریق واریانس پیشین کنترل می‌شود، به‌طوری‌که با واریانس کمتر، انقباض ضرایب بیشتر می‌شود. از دیگر مزایای این روش، استفاده از نمونه‌گیری گیبز و الگوریتم‌های MCMC برای انتخاب مدل، بدون برازش همه مدل‌های ممکن است بنابراین از محاسبات وقت‌گیر جلوگیری می‌شود (فریدلی، ۲۰۰۹).

در این مقاله رویکرد هموارسازی بیزی برای برآورد پارامترهای مدل رگرسیونی فضایی با ابعاد بالا به‌کارگرفته می‌شود. آن‌گاه در مطالعه شبیه‌سازی دقت و زمان محاسبه روش‌های مختلف مورد ارزیابی و مقایسه قرار می‌گیرند.

۲ مدل اتورگرسیو فضایی

در اغلب روش‌های آماری فرض بر این است که مشاهدات تحت شرایط یکسان و به‌صورت مستقل از هم جمع‌آوری شده‌اند. اما اگر مشاهدات مستقل نباشند و وابستگی آن‌ها تابعی از فاصله بین موقعیت‌های مشاهدات باشد، داده‌های فضایی نامیده می‌شوند. در یک مدل رگرسیون فضایی، رابطه بین متغیرهای تبیینی و متغیر پاسخ با در نظر گرفتن اثر موقعیت داده‌ها مشخص می‌شود. مدل SAR (کلیف و ارد، ۱۹۷۳) یک فرایند تصادفی به صورت

$$\mathbf{Y} = \alpha\mathbf{W}\mathbf{Y} + \mathbf{X}\beta + \epsilon, \quad \epsilon \sim N(\mathbf{0}_n, \sigma^2\mathbf{I}_n) \tag{۱.۲}$$

است، که در آن $\mathbf{Y}$ بردار $n \times 1$ متغیر پاسخ، $\mathbf{X}$ ماتریس $n \times p$ متغیرهای تبیینی، α ضریب اتورگرسیو فضایی مشاهدات، $|\alpha| < 1$، β بردار ضرایب رگرسیونی و $\mathbf{W}$ ماتریس وزن فضایی نامنفی و نشان‌دهنده ارتباط بین مشاهدات نمونه است. ارد (۱۹۷۵) برآوردگر سازگار کمترین توان‌های دوم عادی را معرفی کرد، اما با افزایش پارامتر α، کارایی این برآوردگر نسبت به برآوردگر ماکسیمم درستنمایی کاهش می‌یابد. چالش اساسی در برآورد پارامترهای این مدل با ماکسیمم تابع درستنمایی، محاسبه دترمینان ماتریس $\mathbf{S}(\alpha) = \mathbf{I} - \alpha\mathbf{W}$ است که یک چندجمله‌ای درجه n از α است. لسج و پیس (۲۰۰۷) به‌جای $\mathbf{S}(\alpha)$ استفاده از ماتریس مشخصه‌سازی فضایی نمایی[۱۰] معین مثبت

$$\mathbf{S}(\rho) = \exp(\rho\mathbf{W}) = \sum_{\ell=0}^{\infty} \rho^{\ell}\mathbf{W}^{\ell}(\ell!)^{-1} \tag{۲.۲}$$

[3] Gibbs Variable Selection
[4] Bayesian Model Averaging
[5] Stochastic Search Variable Selection
[6] Spike and Slab Priors
[7] Markov Chain Monte Carlo
[8] Regularization
[9] Regularization Parameter
[10] Matrix Exponential Spatial Specification

را پیشنهاد کردند، که در آن $\rho \in (-\infty, \infty)$ همبستگی فضایی را مشخص می‌کند و به دست آوردن برآوردهای ماکسیمم درستنمایی و بیزی را تسهیل می‌کند. اگر چه هر دو رویکرد مشخصه‌سازی، استنباط و برآوردهای مشابهی ارائه می‌کنند، اما ماتریس مشخصه‌سازی فضایی نمایی، مزایای محاسباتی بویژه در فضای داده‌های بعد بالا دارد. با در نظر گرفتن مدل اتورگرسیو فضایی به صورت

$$\mathbf{S}(\rho)\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}, \quad \boldsymbol{\epsilon} \sim N(\mathbf{0}_n, \boldsymbol{\Omega}) \tag{۳.۲}$$

تابع درستنمایی به‌صورت

$$\pi(\mathbf{Y}|\mathbf{X}, \boldsymbol{\beta}, \sigma^2, \rho) = (2\pi)^{-n/2} \det(\boldsymbol{\Omega})^{-1/2} \exp(-\mathbf{e}^{\mathbf{T}}\boldsymbol{\Omega}^{-\mathbf{1}}\mathbf{e}/\mathbf{2}) \tag{۴.۲}$$

حاصل می‌شود، که در آن $\mathbf{e} = (\mathbf{S}(\rho)\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})$. با تعیین یک توزیع آمیخته معلوم، به عنوان مثال D، پیشین‌های پارامترهای رگرسیونی به‌صورت

$$\beta_j|\psi_j \sim N(0, \psi_j), \quad \psi_j \sim D \tag{۵.۲}$$

بیان می‌شوند. توزیع حاشیه‌ای β_j دم‌های سنگین‌تری از توزیع نرمال دارد، اما چگالی زیادی را در صفر قرار می‌دهد. بنابراین مدل‌های سلسه‌مراتبی، با پیشین‌های انقباضی، به‌طور همزمان متغیرهای بی‌اهمیت را از مدل حذف و ضرایب متغیرهای مهم را برآورد می‌کنند (گلمن و همکاران، ۲۰۱۳). شکل توزیع D وابسته به ابرپارامترهای پیشین است. با معرفی سطوح سلسله مراتبی بیشتر، انقباض سازوارتر و انعطاف‌پذیرتری حاصل می‌شود (فارهوفر و پیریباور، ۲۰۱۹). این رویکردها، پیشین‌های انقباضی فراموضعی‌ــ موضعی[11] نامیده می‌شوند و در شرایط بعد بالا خوب عمل می‌کنند (پولسون و اسکات، ۲۰۱۰).

پیشین انقباضی نرمال‌ــ گاما (گریفین و براون، ۲۰۱۰): این پیشین از طریق آمیخته مقیاسی توزیع‌های نرمال به صورت

$$\beta_j|\psi_j \sim N(0, 2\lambda^{-2}\psi_j), \quad \psi_j \sim G(\theta, \theta), \quad \lambda^2 \sim G(d_0, d_1) \tag{۶.۲}$$

معرفی شده است، که در آن ψ_j پارامتر مقیاس ویژه با توزیع گاما و دارای پارامتر خاص انقباض θ است. انقباض کلی توسط پارامتر فراموضعی λ کنترل می‌شود و دارای توزیع گاما است. این سطح مدل سلسله مراتبی، با انقباض همه میانگین‌های ضرایب به سمت صفر منجر به کاهش خطا از طریق پارامتر فراموضعی می‌شود. در حالی‌که برای ضرایب مهم مدل، با استفاده از پارامترهای مقیاسی موضعی، از این انقباض جلوگیری می‌شود. ابرپارامترها باید توسط محقق تعیین شوند، که در این مقاله به پیشنهاد فارهوفر و پیریباور (۲۰۱۹) مقادیر $d_0 = d_1 = 0/01$ و $\theta = 0/1$ به‌عنوان پیش‌فرض در نظر گرفته می‌شوند. بنابراین انقباض شدید روی هر پارامتر از مولفه فراموضعی ناشی می‌شود که در صورت لزوم، انعطاف کافی برای تشخیص ضرایب غیر صفر نشان می‌دهد. اگر $\theta = 1$، مدل لاسوی بیزی حاصل می‌شود (پارک و کسلا، ۲۰۰۸).

گریفین و براون (۲۰۱۰) توزیع پسینی ψ_j را وارون گاوسی تعمیم‌یافته به‌صورت

$$\psi_j|\lambda, \beta_j \sim \mathrm{GIG}(\theta - 1/2, \beta_j^2, \theta\lambda^2) \tag{۷.۲}$$

به‌دست آوردند. توزیع پسینی شرطی پارامتر فراموضعی نیز به‌صورت

$$\lambda^2|\boldsymbol{\psi} \sim G(d_0 + \theta K, d_1 + 2^{-1}\theta \sum_{j=1}^{p} \psi_j) \tag{۸.۲}$$

[11] Global-Local Shrinkage Priors

حاصل می‌شود. ماتریس کوواریانس پیشین $\boldsymbol{\beta}$ با $\underline{\boldsymbol{\Sigma}}$ نشان داده می‌شود که در حین برآورد به‌روزرسانی می‌شود. بردار پیشین ضرایب نیز $\underline{\boldsymbol{\beta}} = ۰$ است. در توزیع پیشین نرمال_گاما، ماتریس کوواریانس به صورت $\operatorname{diag}(\underline{\boldsymbol{\Sigma}}) = \boldsymbol{\psi}$ است. توزیع شرطی کامل ضرایب $\boldsymbol{\beta}$ به‌صورت

$$\boldsymbol{\beta}|\bullet \sim N(\overline{\boldsymbol{\beta}}, \overline{\boldsymbol{\Sigma}}), \quad \overline{\boldsymbol{\Sigma}} = (\underline{\boldsymbol{\Sigma}}^{-۱} + \mathbf{X}^T\boldsymbol{\Omega}^{-۱}\mathbf{X})^{-۱}, \quad \overline{\boldsymbol{\beta}} = \overline{\boldsymbol{\Sigma}}(\underline{\boldsymbol{\Sigma}}^{-۱}\underline{\boldsymbol{\beta}} + \mathbf{X}^T\boldsymbol{\Omega}^{-۱}\mathbf{S}(\rho)\mathbf{Y}) \qquad (۹.۲)$$

حاصل می‌شود، که در آن $\bullet$ نشان‌دهنده بردار مشاهدات و پارامترهای دیگر است (کوپ، ۲۰۰۳).

پیشین انقباضی دیریکله_لاپلاس (باتاچاریا و همکاران، ۲۰۱۵): این پیشین با پارامترهای انقباضی فراموضعی و موضعی سلسله مراتبی به‌صورت

$$\beta_j \sim N(۰, \varphi_j\phi_j^۲\tau^۲), \quad \varphi_j \sim \operatorname{Exp}(۱/۲), \quad \phi \sim \operatorname{Dir}(a, \ldots, a), \quad \tau \sim G(na, ۱/۲). \qquad (۱۰.۲)$$

معرفی شده است، که در آن φ_j پارامتر موضعی و دارای توزیع پیشین نمایی است. در مقابل پیشین نرمال_گاما که یک پارامتر فراموضعی λ دارد، رویکرد دیریکله_لاپلاس، برداری از مقیاس‌ها به صورت $(\phi_۱\tau, \ldots, \phi_p\tau)$ را برای انعطاف‌پذیری بیشتر در انقباض ضرایب خاص ارائه می‌کند، که در آن بردار $\boldsymbol{\phi} = (\phi_۱, \ldots, \phi_p)^T$ به‌گونه‌ای تعریف می‌شود که در مسئله بهینه‌سازی $(p-۱)$بعدی $\zeta^{p-۱} = \{x = (x_۱, \ldots, x_p)^T : x_j \geq ۰, \sum_{j=۱}^p x_j = ۱\}$ قرار گیرد و دارای توزیع پیشین دیریکله با ابرپارامتر a است که شدت و سختی پیشین را کنترل می‌کند. یک انتخاب استاندارد برای پارامتر توزیع دیریکله $a = ۱/p$ است. این امر باعث می‌شود که پیشین دیریکله_لاپلاس، بر خلاف پیشین نرمال_گاما به بعد مدل وابسته باشد. در پیشین دیریکله_لاپلاس φ_j را می‌توان از توزیع وارون گاوسی به طور مستقل تولید کرد. بنابراین با به‌دست آوردن $\tilde{\varphi_j}|\boldsymbol{\phi}, \boldsymbol{\beta}$ از دوباره پارامتری کردن توزیع وارون گاوسی تعمیم‌یافته با $\mu_j = \phi_j\tau/|\beta_j|$ به‌صورت

$$\tilde{\varphi_j}|\boldsymbol{\phi}, \boldsymbol{\beta} \sim \operatorname{GIG}(-۱/۲, ۱, \mu_j^{-۲}) \qquad (۱۱.۲)$$

حاصل می‌شود. در نتیجه قرار داده می‌شود $\varphi_j = ۱/\tilde{\varphi_j}$، تا توزیع پسینی شرطی φ_j حاصل شود. پارامتر انقباض فراموضعی τ از توزیع وارون گاوسی تعمیم‌یافته

$$\tau|\boldsymbol{\phi}, \boldsymbol{\beta} \sim \operatorname{GIG}(۱-P, ۲\sum_{j=۱}^{P} |\beta_j|/\phi_j, ۱) \qquad (۱۲.۲)$$

تولید می‌شود. برای تولید $\boldsymbol{\phi}$ ابتدا متغیرهای کمکی مستقل $T_۱, \ldots, T_P$ از توزیع وارون گاوسی تعمیم‌یافته $\operatorname{GIG}(a-۱, ۲|\beta_j|, ۱)$ تولید می‌شوند، در نتیجه $\phi_j = T_j/\sum_{j=۱}^P T_j$ به دست می‌آید. در نتیجه ساختار ماتریس کوواریانس پیشین $\underline{\boldsymbol{\Sigma}}$ به روزرسانی می‌شود، به‌طوری‌که $\operatorname{diag}(\underline{\boldsymbol{\Sigma}}) = (\varphi_۱\phi_۱^۲\tau^۲, \ldots, \varphi_p\phi_p^۲\tau^۲)^T$ است. در نهایت بردار ضرایب، براساس (۹.۲) تولید می‌شود.

پیشین برای دیگر پارامترها: ماتریس کوواریانس خطا به‌صورت $\boldsymbol{\Omega} = \sigma^۲\mathbf{I}_n$ فرض می‌شود، که در آن $\mathbf{I}_n$ ماتریس همانی است. برای تکمیل پیشین‌ها، باید توزیع پیشین $\sigma^۲$ و پارامتر همبستگی فضایی ρ نیز تعیین شوند. در اینجا از مشخصه‌سازی استاندارد مطرح شده در لسج و پیس (۲۰۰۹) استفاده می‌شود. پیشین گامای وارون برای واریانس جمله خطا به‌صورت $\sigma^۲ \sim \operatorname{G}^{-۱}(\underline{a}, \underline{b})$ فرض می‌شود، که در آن $\underline{a}$ و $\underline{b}$ ابرپارامترها هستند که با قرار دادن مقدار ۰/۰۱ برای آن‌ها، پیشین‌های ناآگاهی‌بخش منظور می‌شوند. با شرطی کردن روی دیگر پارامترهای مدل و داده‌ها، چگالی شرطی کامل واریانس خطا به‌صورت

$$\sigma^۲|\bullet \sim \operatorname{G}^{-۱}(\overline{a}, \overline{b}), \quad \overline{a} = \underline{a} + n/۲, \quad \overline{b} = \underline{b} + \mathbf{e}^T\mathbf{e}/۲,$$

به‌دست می‌آید. مقادیر $\boldsymbol{\beta}$ و $\sigma^{۲}$ استاندارد هستند و می‌توان با استفاده از الگوریتم گیبز از آن‌ها نمونه تولید کرد (کوپ، ۲۰۰۳). برای پارامتر همبستگی فضایی پیشین $\rho \sim N(۰, c)$ در نظر گرفته می‌شود، که در آن c با توجه به اطلاعات پیشین محقق از میزان رابطه فضایی داده‌ها، انتخاب می‌شود. برای مطالعه شبیه‌سازی، با انتخاب $c = ۱۰$ از مشخصه‌سازی ناآگاهی‌بخش استفاده می‌شود. در این‌صورت توزیع شرطی کامل ρ به‌صورت

$$\pi(\rho|\bullet) \propto \exp(-\mathbf{e}^T \boldsymbol{\Omega}^{-۱} \mathbf{e}/۲)\pi(\rho) \qquad (۱۳.۲)$$

حاصل می‌شود، که با استفاده از الگوریتم متروپلیس[12] درون الگوریتم گیبز می‌توان از آن مقادیر تصادفی تولید نمود. برای این‌کار، ابتدا یک مقدار از توزیع نرمال $\rho^* \sim N(\rho_{t-۱}, \tau)$ تولید می‌شود که زیروند $t - ۱$ نشان‌دهنده مقدار پارامتر از تکرار قبلی الگوریتم نمونه‌گیری و τ پارامتر تنظیم است. احتمال پذیرش مقدار پیشنهاد شده از $\min\{۱, \frac{g(\rho^*|\bullet)}{g(\rho_{t-۱}|\bullet)}\}$ محاسبه می‌شود. اگر مقدار پیشنهادی پذیرفته شود، $\rho_t = \rho^*$ تعیین و در غیر این‌صورت مقدار بدست آمده از تکرار قبلی باقی می‌ماند. در تولید نمونه برای ρ با داغیدن نیمی از تکرارهای الگوریتم، واریانس توزیع τ، به‌طوری تنظیم می‌شود که نرخ پذیرش بین ۰/۲ تا ۰/۴ باشد.

۳ مطالعه شبیه‌سازی

ویژگی‌های تجربی و مزایای پیشین‌های انقباضی در مدل اتورگرسیو فضایی، در یک مطالعه شبیه‌سازی بررسی می‌شود. داده‌ها با استفاده از پیشین‌های انقباضی نرمال‌ـ گاما و دیریکله‌ـ لاپلاس مدل‌بندی می‌شوند. به عنوان مشخصه معیار، در شرایط یکسان الگوریتم MCMC دیگری بدون در نظر گرفتن انقباض روی $\boldsymbol{\beta}$ اجرا و نتایج آن با دو روش انقباضی مقایسه می‌شود. برای هر رویکرد، یک زنجیر تولید می‌شود که با داغیدن نیمه اول نمونه، با رسم نمودار ACF[13] و انتخاب فاصله[14]، استنباط بر اساس ۱۰۰۰ نمونه آخر انجام می‌شود. برآورد نقطه‌ای پارامترها نیز از طریق میانگین نمونه‌های مستقل حاصل می‌شود. در این مطالعه حجم نمونه $n = ۱۰۰$ و $p = ۲۰۰$ در نظر گرفته می‌شود. علاوه‌براین باید مقادیر پارامترهای ضرایب $\boldsymbol{\beta}$، واریانس جمله خطا $\sigma^{۲}$ و پارامتر اتورگرسیو فضایی ρ تعیین شود. مدل اتورگرسیو فضایی با ماتریس مشخصه‌سازی فضایی نمایی به‌صورت

$$\mathbf{Y} = \exp(-\rho \mathbf{W})(X\boldsymbol{\beta} + \boldsymbol{\epsilon}), \quad \boldsymbol{\epsilon} \sim N(۰, \sigma^{۲}\mathbf{I}_n)$$

در نظر گرفته شده است، که در آن ستون اول ماتریس X ثابت عرض از مبدا است. پارامترهای مورد نیاز نیز به صورت زیر تعیین می‌شوند:

۱ـ مقادیر هر یک از متغیرهای تبیینی، برای هر مشاهده $i = ۱, \ldots, n$ از توزیع نرمال استاندارد تولید شده، یعنی $x_{ij} \sim N(۰, ۱)$ و ماتریس $\mathbf{X} = (\iota_n, x_۱, \ldots, x_{p-۱})$ حاصل می‌شود که در آن ι_n بردار nـ بعدی با درایه‌های ۱ است.

۲ـ برای تولید مختصات مشاهدات، نقاط تصادفی از درون یک مربع واحد تولید می‌شوند و ماتریس سطرـ تصادفی $\mathbf{W}$ بر اساس پنج‌ـ نزدیک‌ترین همسایه ساخته می‌شود.

۳ـ بردار ضرایب تنک با تعیین تعداد پارامترهای غیرصفر به‌دست می‌آید. ۱۰ مولفه اول بردار ضرایب غیرصفر و دیگر مولفه‌ها صفر در نظر گرفته می‌شوند. بنابراین ۵ مولفه اول بردار ضرایب از طریق نمونه‌گیری از $N(۰, ۵)$ و ۵ مولفه دوم آن از $N(۰, ۱)$ تولید می‌شوند. بقیه ضرایب نیز صفر هستند.

[12] Metropolis
[13] AutoCorrelation Function
[14] Thin

۴ـ واریانس جمله خطا برابر یک است، یعنی $\sigma^{۲} = ۱$.

۵ـ پارامتر همبستگی فضایی ρ، با نمونه‌گیری از توزیع $N(۰, ۳)$ تولید می‌شود.

با اجرای الگوریتم MCMC برآورد نقطه‌ای پارامترها به دست می‌آید. ابتدا مقادیر اولیه در فضای پارامتر مطلوب تولید می‌شوند. در هر مدل مقادیر اولیه یکسان، وارد الگوریتم گیبز می‌شوند. $\sigma^{۲}$ ثابت و معلوم در نظر گرفته می‌شود. در مدل ساده، (بدون انقباض)، ماتریس کوواریانس پیشین $\boldsymbol{\beta}$ به‌صورت $\underline{\Sigma} = ۱۰۰۰\mathbf{I}_p$ تعیین می‌شود. در پیشین انقباضی نرمال‌ـگاما، پارامترهای موضعی ψ_j به‌صورت مستقل از توزیع وارون گاوسی تعمیم‌یافته در (۷.۲) تولید می‌شوند. پارامتر انقباض فراموضعی نیز از توزیع گاما (۸.۲) تولید می‌شود. ماتریس کوواریانس $\boldsymbol{\psi} = \operatorname{diag}(\underline{\boldsymbol{\Sigma}})$ با مقادیر جدید بدست می‌آید و پارامترهای رگرسیونی با استفاده از (۹.۲) تولید می‌شوند. در پیشین انقباضی دیریکله‌ـ لاپلاس، پارامترهای مقیاسی موضعی و فراموضعی به ترتیب از توزیع وارون گاوسی تعمیم‌یافته (۱۱.۲) و (۱۲.۲) تولید می‌شوند. متغیرهای کمکی T_j نیز از توزیع وارون گاوسی تعمیم‌یافته با $\phi_j = T_j / \sum_{j=۱}^{p} T_j$ تولید می‌شوند. مانند پیشین نرمال‌ـگاما، ماتریس کوواریانس باید مجددا محاسبه شود. بنابراین در هر تکرار $\operatorname{diag}(\underline{\boldsymbol{\Sigma}}) = (\varphi_۱\phi_۱^۲\tau^۲, \ldots, \varphi_p\phi_p^۲\tau^۲)^T$ محاسبه و ضرایب رگرسیونی از (۹.۲) تولید می‌شوند.

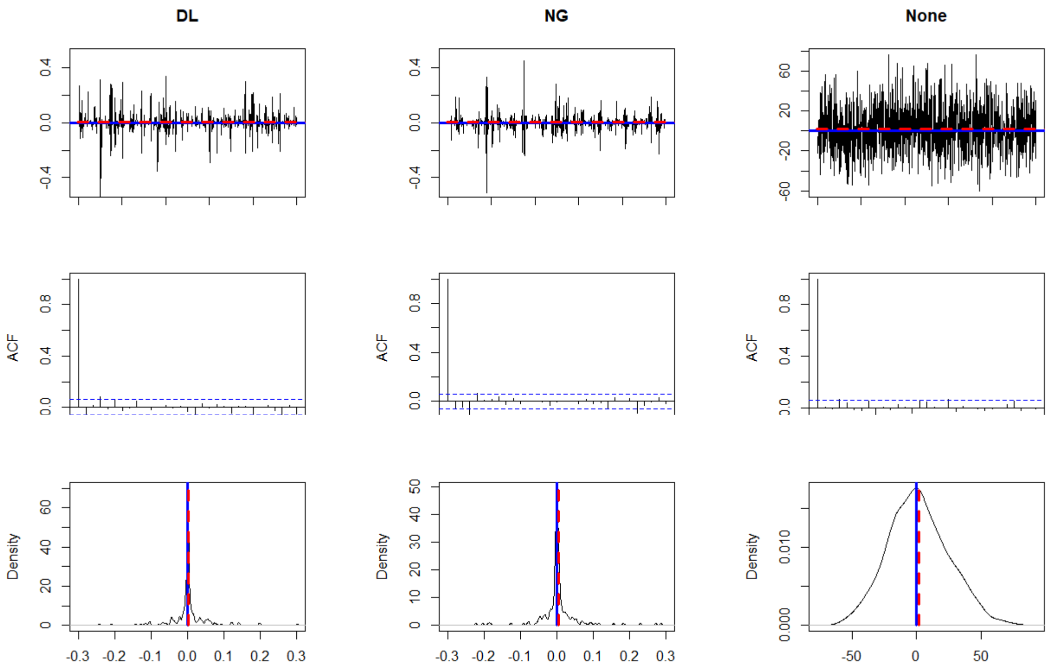
در مرحله آخر الگوریتم، پارامتر همبستگی ρ با استفاده از الگوریتم متروپلیس درون گیبز با توزیع پسینی شرطی (۱۳.۲) تولید می‌شود. برای تولید مقدار پیشنهادی ρ در الگوریتم متروپلیس، باید پارامتر تنظیم τ، واریانس توزیع کاندید، به طور تجربی برآورد شود. با اجرای ۵۰۰۰ تکرار الگوریتم متروپلیس و با نرخ پذیرش $r = ۰/۳۶$، واریانس توزیع کاندید $\hat{\tau} = ۳۰$ برآورد شد. نمودار اثر مدل‌های مختلف، در شکل ۳ برای پارامتر pام ترسیم شده است. خط آبی نشان‌دهنده مقدار واقعی پارامتر و خط چین قرمز نشان‌دهنده مقدار برآورد شده همان پارامتر با استفاده از الگوریتم گیبز و میانگین پسینی است. با توجه به دامنه محور عمودی یعنی مقادیر پارامتر در زنجیر، انقباض پارامتر در مدل نرمال‌ـگاما و دیریکله‌ـلاپلاس مشهود است. نمودار چگالی مقادیر پارامتر pام نیز به خوبی گواه این موضوع است. در داده‌های بعد بالا، دامنه تغییرات در مدل ساده بسیار زیاد است. در نمودار چگالی نیز قابل مشاهده است که دامنه مقادیر بزرگ‌تر از $(-۵۰, ۵۰)$ است. به نظر می‌رسد پیشین دیریکله‌ـلاپلاس انقباض بیشتری نسبت به پیشین نرمال‌ـگاما، روی پارامتر اعمال می‌کند و چگالی زیادی را در صفر قرار می‌دهد. بنابراین، انتظار می‌رود عملکرد بهتری از پیشین نرمال‌ـگاما داشته باشد. مقادیر محاسبه شده RMSE برای مدل‌های مختلف در جدول ۱ نشان داده شده است. همان‌طور که ملاحظه می‌شود مقادیر RMSE در دو پیشین انقباضی بسیار نزدیک به هم هستند و به طور متوسط پیشین دیریکله‌ـلاپلاس عملکرد بهتری نسبت به پیشین نرمال‌ـگاما در برآورد پارامترها دارد. زمان محاسبات (به دقیقه) در مدل ساده اختلاف بسیار کمی با دو مدل دیگر دارد، با این حال اختلاف مقادیر RMSE در مدل ساده نسبت به دو مدل دیگر زیاد است.

جدول ۱: مقادیر RMSE تجربی و زمان محاسبات(به دقیقه) برای مدل‌های فضایی.

مدل	RMSE_{β}	RMSE_{ρ}	Time(Min.)
ساده	۰/۳۹	۰/۹۴	۳۰/۵۹
نرمال‌ـگاما	۰/۲۹	۰/۶۶	۳۱/۷۸
دیریکله‌ـلاپلاس	۰/۲۲	۰/۶۶	۳۱/۲۹

بحث و نتیجه‌گیری

در این مقاله، روشی برای مدل‌بندی داده‌های بعد بالای فضایی، با استفاده از انقباض ضرایب، از طریق رهیافت بیزی مطرح شد. در هر مطالعه آماری، اطلاعاتی درباره پارامتر وجود دارد که منجر به تعریف توزیع پیشین می‌شود. مزیت مهم رهیافت



شکل ۱:   نمودارهای اثر، توابع خودهمبستگی و توزیع پسینی پارامتر در مدل‌های فضایی $p = ۲۰۰$ ستون‌ها به ترتیب از سمت راست: مدل ساده، نرمال_ گاما و دیریکله_ لاپلاس.

بیزی به‌دست آوردن بازه باورمندی همزمان با انتخاب متغیر و برآورد پارامترهای مدل است و نتایج معتبری در برآورد پارامترهای مدل اتورگرسیو فضایی با استفاده از پیشین‌های انقباضی فراهم شد. استفاده از پیشین‌های تیر و تخته دیگر می‌تواند در بهبود برآورد مفید باشد و مقدار قابل توجهی از چگالی پسینی را در صفر متمرکز کند که می‌تواند در مطالعات بعدی مورد توجه باشد.

تقدیر و تشکر

نویسندگان این مقاله از برگزارکنندگان و داوران محترم چهارمین سمینار آمار فضایی و کاربردهای آن و از حمایت قطب علمی تحلیل داده‌های وابسته فضایی و فضایی_ زمانی دانشگاه تربیت مدرس کمال تشکر و قدردانی را دارند.

مراجع

Bhattacharya, A., Pati, D., Pillai, N. S., and Dunson, D. B. (2015), Dirichlet–Laplace Priors for Optimal Shrinkage, *Journal of the American Statistical Association*, **110**, 1479-1490.

Cliff, A., and Ord, J. (1973), Spatial Autocorrelation, Pion, London.

Fridley, B. L. (2009), Bayesian Variable and Model Selection Methods for Genetic Association Studies, *Genetic Epidemiology: The Official Publication of the International Genetic Epidemiology Society*, **33**, 27-37.

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., and Rubin, D. B. (2013), Bayesian data analysis. CRC press.

Koop, G. M. (2003), Bayesian Econometrics, John Wiley and Sons Inc.

LeSage, J. P., and Parent, O. (2007), Bayesian Model Averaging for Spatial Econometric Models, *Geographical Analysis*, **39**, 241-267.

LeSage, J. P., and Pace, R. K. (2007), A Matrix Exponential Spatial Specification, *Journal of Econometrics*, **140**, 190-214.

LeSage, J., and Pace, R.K. (2009), *Introduction to Spatial Econometrics*, Chapman and Hall/CRC, New Yprk.

Griffin, J. E., and Brown, P. J. (2010), Inference with Normal-Gamma Prior Distributions in Regression Problems,*Bayesian Analysis*, **5**, 171-188.

Himel Mallick, N. Y. (2013), Bayesian Methods for High Dimensional Linear Models, *Biostatistics*, **1**, 005.

Ord, K. (1975), Estimation Methods for Models of Spatial Interaction, *Journal of the American Statistical Association*, **70**, 120-126.

Park, T., and Casella, G. (2008), The Bayesian Lasso, *Journal of the American Statistical Association*, **103**, 681-686.

Pfarrhofer, M., and Piribauer, P. (2019), Flexible Shrinkage in High-Dimensional Bayesian Spatial Autoregressive Models, *Spatial Statistics*, **29**, 109-128.

Piribauer, P. (2016), Heterogeneity in Spatial Growth Clusters, *Empirical Economics*, **51**, 659–680.

Polson, N. G., and Scott, J. G. (2010), Shrink Globally, Act Locally: Sparse Bayesian Regularization and Prediction, *Bayesian Statistics*, **9**, 501-538.

Rosset, S., and Zhu, J. (2006), Sparse, Flexible and Efficient Modeling using L1 Regularization, In Feature Extraction, 375-394.