برآورد ماکسیمم درستنمایی پارامترهای مدل اتورگرسیو فضایی مرتبه دوم یکطرفه

آزاده مجیری[۱]، یداله واقعی[۲]، حمیدرضا نیلی ثانی[۳]، غلامرضا محتشمی برزادران[۴]

[۱] دانشگاه بیرجند، a.mojiri۳۲۳۰@yahoo.com؛ [۲] دانشگاه بیرجند، ywaghei@birjand.ac.ir

[۳] دانشگاه بیرجند، nilisani@yahoo.com؛ [۴] دانشگاه فردوسی مشهد، gmb۱۳۳۴@yahoo.com

چکیده

به داده‌های وابسته‌ای که از موقعیتهای فضایی مختلف به دست می‌آیند، داده‌های فضایی می‌گویند. در صورتی که موقعیت داده‌های فضایی مربوط به مکان‌های ناحیه‌ای باشند، داده‌های شبکه‌ای نامیده می‌شوند. داده‌های شبکه‌ای معمولا با استفاده از مدلهای اتورگرسیو فضایی مدل‌بندی می‌شوند. یک رده مدل‌های فضایی یکطرفه که برای داده‌های شبکه‌ای استفاده می‌گردد مدل‌های اتورگرسیو فضایی مرتبه دوم یکطرفه می‌باشد. در این مقاله مدل اتورگرسیو فضایی مرتبه دوم یکطرفه معرفی و پارامترهای آن به روش ماکسیمم درستنمایی برآورد گردیده است.

واژه‌های کلیدی: داده‌های شبکه‌ای، مدل اتورگرسیو فضایی مرتبه دوم یکطرفه، برآورد ماکسیمم درستنمایی
رده‌بندی موضوعی ریاضی (۲۰۱۰): ۶۲M۳۰، ۹۱B۷۲، ۶۲H۱۱.

۱ مقدمه

فرض کنید $s \in D \subset \mathbb{R}^d$ موقعیت یك مشاهده در فضای اقلیدسی dبعدی و Z_s یك متغیر تصادفی در موقعیت s باشد. گردایه‌ای از متغیرهای تصادفی به صورت $\{Z_s; s \in D\}$ را میدان تصادفی می‌نامند. مشاهدات حاصل از میدان تصادفی، داده‌های فضایی نامیده می‌شود. داده‌های شبکه‌ای نوع خاصی از داده‌های فضایی هستند که موقعیت مکانی داده‌ها به صورت ناحیه‌ای است، این مکان‌ها می‌تواند منظم یا نامنظم باشند.

در صورتی که داده‌ها در فضای دو بعدی بر روی یک شبکه منظم قرار بگیرند متغیرها به صورت $\{Z_{ij}, i=1,2,\ldots,m;\ j=1,2,\ldots n\}$ نمایش می‌دهیم.

مدلهای فضایی که برای تحلیل داده‌های فضایی شبکه‌ای به کار می‌رود شامل مدل‌های اتورگرسیو فضایی (SAR) (ویتل ۱۹۵۴)، اتورگرسیو شرطی (CAR) (بارتلت ۱۹۷۱ و بسج ۱۹۷۴) و میانگین متحرک (MA) (هاینینگ۱۹۷۸b) می‌باشند.

محققانی که قبلا بر روی حالتهای مدل اتورگرسیو فضایی SAR کارکرده‌اند میانگین داده‌ها را صفرفرض کرده‌اند. در این مقاله می‌خواهیم مدل SAR را به حالتی که میانگین غیر ثابت باشد تعمیم داده و برآورد مدل انجام دهیم.

۲ برآورد پارامترهای مدل اتورگرسیو فضایی

مدل اتورگرسیو فضایی یکطرفه که به $SAR(p_1,p_2)$ نشان می‌دهیم به صورت $\Phi(B_1,B_2)Z_{i,j}=\varepsilon_{i,j}$ تعریف می‌شود که در آن $\Phi(B_1,B_2)=\sum_{k=0}^{p_1}\sum_{l=0}^{p_2}\alpha_{kl}B_1^k B_2^l$ و $\varepsilon_{i,j}$ها متغیرهای تصادفی مستقل با توزیع دلخواه، میانگین صفر و واریانس ثابت σ^2 می‌باشند.

با تغییر پارامترهای $\alpha_{10}=\alpha_1$، $\alpha_{01}=\alpha_2$، $\alpha_{11}=\alpha_3$، $\alpha_{20}=\alpha_4$ و $\alpha_{21}=\alpha_5$ مدل اتورگرسیو فضایی مرتبه دوم یکطرفه ($SAR(2,1)$) به صورت زیر تبدیل می‌شود:

$$Z_{i,j}=\alpha_1 Z_{i-1,j}+\alpha_2 Z_{i,j-1}+\alpha_3 Z_{i-1,j-1}+\alpha_4 Z_{i-2,j}+\alpha_5 Z_{i-2,j-1}+\varepsilon_{i,j} \qquad (۱.۲)$$

فرض کنید $\mathbf{Z}=(Z_{1,1},Z_{1,2},\ldots,Z_{1,n},\ldots,Z_{m,1},\ldots,Z_{m,n})'$ و $\varepsilon=(\varepsilon_{1,1},\varepsilon_{1,2},\ldots,\varepsilon_{1,n},\ldots,\varepsilon_{m,1},\ldots,\varepsilon_{m,n})'$ بردارهایی با $N=m\times n$ عنصر باشند با فرض صفر بودن نقاط مرزی مدل (۱.۲) را می‌توان به فرم ماتریسی زیر نوشت:

$$\mathbf{Z}=\mathbf{A}\mathbf{Z}+\varepsilon \rightarrow (\mathbf{I}_N-\mathbf{A})\,\mathbf{Z}=\varepsilon \qquad (۲.۲)$$

که در آن $\mathbf{A}$ ماتریس $N\times N$ شامل پارامترهای مدل که آن را می‌توان به پنج ماتریس مجزا تجزیه کرد و $\mathbf{I}_N$ ماتریس همانی $N\times N$ است.

در صورتی که $E(Z_{i,j})=\mu$ باشد با استفاده از μ به جای مقادیر مرزی فرم ماتریسی مدل فوق برابر است با:

$$(\mathbf{I}_N-\mathbf{A})(\mathbf{Z}-\mu\mathbf{1})=\varepsilon \qquad (۳.۲)$$

اگر میانگین $\mathbf{Z}$ ثابت نباشد و به صورت $E(\mathbf{Z})=X\beta$ باشد به عنوان تعمیم دیگری از مدل (۲.۲) داریم:

$$(\mathbf{I}_N-\mathbf{A})(\mathbf{Z}-X\beta)=\varepsilon \qquad (۴.۲)$$

تابع درستنمایی $\boldsymbol{\theta}'=(\alpha_1,\alpha_2,\alpha_3,\alpha_4,\alpha_5,\mu,\sigma^2)$ برای مدل (۳.۲) برابر است با:

$$L(\boldsymbol{\theta}|\mathbf{Z})=(2\pi)^{-N/2}(\sigma^2)^{-N/2}\exp\left\{-\frac{1}{2\sigma^2}(\mathbf{Z}-\mu\mathbf{1})'(\mathbf{I}_N-\mathbf{A}')(\mathbf{I}_N-\mathbf{A})(\mathbf{Z}-\mu\mathbf{1})\right\} \qquad (۵.۲)$$

به منظور برآورد $\boldsymbol{\alpha}=(\alpha_1,\alpha_2,\alpha_3,\alpha_4,\alpha_5)$، $E(Z_{i,j})=\mu$ و $Var(Z_{i,j})=\sigma^2$ مشتقات تابع لگاریتم درستنمایی را نسبت به α_i، μ و σ^2 صفر می‌گذاریم.

فرض کنید $\mathbf{X}^*=[\mathbf{S}_1|\mathbf{S}_2|\mathbf{S}_3|\mathbf{S}_4|\mathbf{S}_5]$، $\mathbf{S}_i=\mathbf{W}_i\mathbf{Y}$، $\mathbf{Y}=\mathbf{Z}-\mu\mathbf{1}$ و $\mathbf{B}=(\mathbf{I}_N-\mathbf{A})$ باشد برآورد پارامتر ها به ترتیب از روابط زیر به دست می‌آید:

$$\hat{\boldsymbol{\alpha}}=\left(\mathbf{X}^{*\prime}\mathbf{X}^*\right)^{-1}\left(\mathbf{X}^{*\prime}\mathbf{Y}\right) \qquad (۶.۲)$$

$$\hat{\mu}=\frac{\mathbf{1}'\hat{\mathbf{B}}'\hat{\mathbf{B}}}{\mathbf{1}'\hat{\mathbf{B}}'\hat{\mathbf{B}}\mathbf{1}}\mathbf{Z} \qquad (۷.۲)$$

$$\hat{\sigma^2}=\frac{\left\{(\mathbf{Z}-\hat{\mu}\mathbf{1})'\hat{\mathbf{B}}'\hat{\mathbf{B}}(\mathbf{Z}-\hat{\mu}\mathbf{1})\right\}}{N} \qquad (۸.۲)$$

لگاریتم تابع درستنمایی در حالت میانگین غیر ثابت برای مدل (۴.۲) برابر است با:

$$l(\boldsymbol{\theta}|\mathbf{Z})=-\frac{N}{2}\ln(2\pi)-\frac{N}{2}\ln(\sigma^2)-\frac{1}{2\sigma^2}\left\{(\mathbf{Z}-X\beta)'(\mathbf{I}_N-\mathbf{A}')(\mathbf{I}_N-\mathbf{A})(\mathbf{Z}-X\beta)\right\} \qquad (۹.۲)$$

با تغیر متغییر $\mathbf{Y}=\mathbf{Z}-\mathbf{X}\beta$ برآورد α همانند رابطه (۶.۲) است. همچنین برای به دست آوردن برآورد β و σ^2 اگر $\mathbf{B}=(\mathbf{I}_N-\mathbf{A})$ قرار دهیم، داریم:

$$\hat{\beta}=\left(\mathbf{X}'\hat{\mathbf{B}}'\hat{\mathbf{B}}\mathbf{X}\right)^{-1}\mathbf{X}'\hat{\mathbf{B}}'\hat{\mathbf{B}}\mathbf{Z} \qquad (۱۰.۲)$$

$$\hat{\sigma}^2=\frac{\left\{\left(\mathbf{Z}-X\hat{\beta}\right)'\hat{\mathbf{B}}'\hat{\mathbf{B}}\left(\mathbf{Z}-X\hat{\beta}\right)\right\}}{N} \qquad (۱۱.۲)$$

توجه کنید که برآورد پارامترها به یکدیگر وابسته هستند لذا به صورت صریح و با فرمول بسته به گونه‌ای که بر حسب داده‌ها قابل محاسبه باشند، به دست نمی‌آیند. برای رفع این مشکل ابتدا برآورد اولیه ای برای $\hat{\mu}$ به دست آورده سپس الگوریتم زیر را به کار می‌بریم:

- گام ۱: برآورد α با استفاده از رابطه (۶.۲)
- گام ۲: برآورد μ (β) با استفاده از رابطه (۷.۲) ((۱۰.۲))
- گام ۳: برآورد σ^2 با استفاده از رابطه (۸.۲) یا (۱۱.۲)

گام‌های ۱ تا ۳ را تا رسیدن به همگرایی ادامه می‌دهیم.

۳ مثال کاربردی

به منظور ارزیابی فنون ارائه شده تحت پارامترهایی که در سطر دوم جدول (۱) مشخص است از هر یک از چهار مدل زیر ۲۵۰ داده ($m=25, n=10$) شبیه‌سازی کرده‌ایم:

$$\begin{cases}(1)\ E(Z_{i,j})=\mu,\\ (2)\ E(Z_{i,j})=\beta_0+\beta_1 i+\beta_2 j,\\ (3)\ E(Z_{i,j})=\beta_0+\beta_1 i+\beta_2 j+\beta_3 ij,\\ (4)\ E(Z_{i,j})=\beta_0+\beta_1 i+\beta_2 j+\beta_3 ij+\beta_4 i^2+\beta_5 j^2.\end{cases} \qquad (۱.۳)$$

سپس با استفاده از روش ماکسیمم درستنمایی برآورد پارامترهای هر مدل را با استفاده از داده‌های همان مدل به دست آورده‌ایم، این نتایج در جدول (۱) آمده است. تعداد کل پارامترهای مدل را به k نشان می‌دهیم و از میانگین قدر مطلق خطا ($MAE=\sum_{i=1}^{k}\left|\theta_i-\hat{\theta}_i\right|/k$) برای ارزیابی مدلها استفاده می‌کنیم. از آنجایی که مقدار MAE مدل (۲) کمتر از سه مدل دیگر است، در این مثال مدل (۲) بهتر می‌باشد.

جدول ۱: برآورد پارامترهای مدلهای شبیه‌سازی شده به روش ماکسیمم درستنمایی

مدل	$\hat{\theta}_i$	$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\alpha}_3$	$\hat{\alpha}_4$	$\hat{\alpha}_5$	$\hat{\sigma^2}$	$\hat{\beta}_0=\hat{\mu}$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	$\hat{\beta}_5$	MAE
-	θ_i	۰/۷	۰/۹	۰/۲۵	۰/۱	-۰/۴	۰/۲۵	۱	۰/۲	۰/۵	-۰/۱۵	۰/۰۸	۰/۰۳	-
(۱)	-	۰/۶۷	۰/۹	۰/۲۷۵	۰/۱۲۳	-۰/۳۹۸	۰/۲۴۷	۱/۰۰۶	-	-	-	-	-	۰/۰۱۲۷
(۲)	-	۰/۶۷	۰/۹	۰/۲۷۵	۰/۱۲۲	-۰/۳۹۸	۰/۲۴۷	۱/۰۲۶	۰/۱۹۹	۰/۴۹۸	-	-	-	۰/۰۱۲۳*
(۳)	-	۰/۶۶۷	۰/۹۰۲	۰/۲۷۴	۰/۱۲۴	-۰/۳۹۶	۰/۲۴۵	۱/۲۱۷	۰/۱۸۱	۰/۴۵۱	-۰/۱۴۵	-	-	۰/۰۳۸۲
(۴)	-	۰/۶۶۵	۰/۹۰۱	۰/۲۷۷	۰/۱۲۶	-۰/۳۹۷	۰/۲۴۵	۱/۲۵۹	۰/۱۸۱	۰/۴۲۴	-۰/۱۴۵	۰/۰۷۹	۰/۰۳۳	۰/۰۳۸۳

در ادامه به منظور تولید یک مجموعه داده که ازمدل SAR پیروی نکند، یک مجموعه داده از نوع زمین آمار با میدان تصادفی گوسی و میانگین $\mu=0/7+x^3+\sqrt{y}$ با استفاده از دستور **grf** در پکیج **geoR** نرم افزار **R** ، تولید نموده‌ایم. تابع کواریانس داده‌ها دارای مدل نمایی با پارامترهای $\sigma^2=0/75$ ، دامنه ۰/۳ و آستانه صفر می‌باشد.

این مجموعه داده تقریبا شبیه داده‌های واقعی است که از چهار مدل فوق پیروی نمی‌کند. بر روی این مجوعه داده مدلهای (۱) تا (۴) را برازش داده‌ایم و پارامترهای آن را برآورد کرده ایم.

با توجه به اینکه معیار آکاییک ($AIC=-2\ln(L)+2k$) مدل (۴) کمتر از سه مدل دیگر است، در این مثال مدل (۴) بهتر می‌باشد، نتایج در جدول (۲) آمده است.

جدول ۲: برآورد پارامترها به روش ماکسیمم درستنمایی با استفاده از داده‌های زمین آماری

مدل	$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\alpha}_3$	$\hat{\alpha}_4$	$\hat{\alpha}_5$	$\hat{\sigma^2}$	$\hat{\beta}_0=\hat{\mu}$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	$\hat{\beta}_5$	AIC
(۱)	۰/۳۳۷	۰/۶۱۷	-۰/۰۳۲	۰/۰۴	-۰/۰۲۹	۰/۳۶	۲/۰۲۴	-	-	-	-	-	۴۲۲/۵۱۲
(۲)	۰/۳۶	۰/۵۸۲	۰/۰۱۵	۰/۰۴۲	-۰/۰۲۱	۰/۳۴۱	۰/۹۸	۰/۱۱۴	۰/۰۸۱	-	-	-	۴۱۴/۵۶۳
(۳)	۰/۳۷۱	۰/۵۴۲	۰/۱۱	۰/۰۵۱	-۰/۰۵۵	۰/۳۳۸	۰/۴۹۲	۰/۱۶	۰/۱۵۸	-۰/۰۱۱	-	-	۴۱۴/۲۴۳
(۴)	۰/۳۲۳	۰/۴۴۷	۰/۰۳۹	-۰/۰۴	-۰/۰۹۹	۰/۳۰۲	۰/۴۵۹	۰/۴۰۳	-۰/۱۴۷	-۰/۰۱۶	-۰/۰۱۶	۰/۰۲۴	۳۹۳/۴۷۲*

مراجع

[1] Awang, N., and Shitan, M. (2006), Estimating the Parameters of the Second Order Spatial Unilateral Autoregressive Model, *International Journal of Statistical Sciences*, **5**, 37-57.

[2] Bartlett, M. S. (1971), Physical Nearest-Neighbour Models and non-Linear Time Series, *J. Appl. Prob*, **8**, 222-232.

[3] Basu, S. and Reinsel G. C. (1993), Properties of the Spatial Unilateral First-Order ARMA Model, *Adv. Appl. Prob*, **25**, 631-648.

[4] Besag, J. (1974), Spatial Interaction and the Statistical Analysis of Lattice Systems, *J. R. Statist. Soc. B*, **36**, 192-236.

[5] Haining, R.P. (1978b), The Moving Average Model for Spatial Interaction, *Trans. Inst. Br. Geog*, **3**, 202-225.

[6] Whittle, P. (1954), On Stationary Processes in the Plane, *Biometrika*, **41**, 434-449.