

مقایسه روش‌های خودگردان و شبکه عصبی خودگردان در سری‌های زمانی غیرخطی

مسعود عیسی‌پره*، نصراله ایران‌پناه*^۱، مرجان کائدی**

*گروه آمار دانشگاه اصفهان،

**گروه مهندسی فناوری اطلاعات، دانشگاه اصفهان

تاریخ پذیرش: ۱۳۹۳/۱۰/۱۳

تاریخ دریافت: ۱۳۹۳/۱/۵

چکیده: شبکه‌های عصبی مصنوعی، از جمله مدل‌های ریاضی جدیدی هستند که با دقت بالا به مدل‌بندی ساختار سری‌های زمانی غیرخطی می‌پردازند. مزیت این مدل‌ها در مقایسه با مدل‌های سری زمانی این است که نیاز به فرضیات محدودکننده نیست. دقت برآوردگرهای حاصل از شبکه عصبی به‌عنوان یک مدل ناپارامتری از مسائل مهم هست. برای این منظور با استفاده از روش‌های خودگردان، می‌توان دقت برآوردگرها را در ساختارهای پیچیده و غیرخطی سری‌های زمانی محاسبه کرد. این روش‌ها که در سال‌های اخیر معرفی شده‌اند، در محاسبه اریبی و واریانس برآوردگرها، در مقایسه با سایر روش‌ها از دقت بالاتری برخوردار هستند. در این مقاله ابتدا به معرفی الگوریتم‌های شبکه عصبی خودگردان، خودگردان اتورگرسیو، خودگردان بلوک متحرک و خودگردان مانده‌ها در سری‌های زمانی پرداخته می‌شود. سپس این چهار الگوریتم در یک مطالعه شبیه‌سازی برای مدل‌های اتورگرسیو، اتورگرسیو نمایی، اتورگرسیو آستانه‌ای خودمحرک و اتورگرسیو تغییر وضعیت هموار برای اندازه دقت برآورد واریانس استاندارد شده سه برآوردگر میانگین، اتوکوواریانس مرتبه صفر (واریانس) و اتوکوواریانس مرتبه یک، مورد مقایسه قرار می‌گیرند. درنهایت با استفاده از داده‌های قیمت نفت سنگین ایران به ارائه یک مثال کاربردی پرداخته می‌شود.

واژه‌های کلیدی: شبکه عصبی مصنوعی، خودگردان اتورگرسیو، خودگردان مانده‌ها، خودگردان بلوک متحرک، سری زمانی غیرخطی، شبیه‌سازی مونت کارلو.

رده‌بندی موضوعی: ۳۷۷۱۰، ۶۲۴۴۰

۱- مقدمه

سری‌های زمانی غیرخطی زمینه‌های رو به رشدی را در پژوهش‌های اقتصادسنجی و آمار کاربردی به‌خصوص در حیطه اقتصاد، به خود اختصاص داده‌اند. بر این اساس همواره برازش مدل‌های خطی یا غیرخطی مورد بحث محققین آمار و اقتصاد بوده است. تشخیص خطی یا غیرخطی بودن سری زمانی و نیز نوع مدل برازش داده‌شده در سری زمانی با محدودیت‌هایی از قبیل عدم برقراری فرض‌ها و دیگر مشکلات برازش همراه است. با این حال بیشتر مدل‌ها، دارای ساختارهایی غیرخطی می‌باشند.

از طرف دیگر، در تحلیل داده‌های سری زمانی، برازش مدلی با کمترین خطا، محاسبه اندازه دقت برآوردگرها و برآورد بازه‌های پیش‌بینی بر اساس یک نمونه مشاهده‌شده، از مسائل مهم است. اگر فرایند مورد نظر خطی و توزیع فرایند خطا معلوم باشد، ارائه بازه پیش‌بینی، برازش مدل و محاسبه اندازه دقت برآوردگرها، به‌خوبی مورد بحث قرار گرفته و توسعه یافته است. در غیر این صورت، برآورد اندازه دقت برآوردگرها و بازه‌های پیش‌بینی از دقت پایینی برخوردار هستند [۱]. برای حل این مشکل، چندین روش ناپارامتری بر اساس روش‌های خودگردان، پیشنهاد شده است. آلونسو و همکاران [۲] برای رفع این مشکل روش خودگردان اتورگرسیو، که در خانواده مدل‌های خطی آزاد توزیع است، پیشنهاد دادند. این روش برای فرایند خطی، کارا اما برای فرایند غیرخطی ضعیف عمل می‌کند. با این حال اگر فرایند تولید داده‌ها غیرخطی باشد، روش خودگردان اتورگرسیو روشی ناسازگار خواهد بود. بولمان [۳ و ۴] با استفاده از شبیه‌سازی نشان داد برای رده‌ای از مدل‌های اتورگرسیو نمایی^۱ (EXPAR) و مدل‌های اتورگرسیو آستانه‌ای خودمحرک^۲ (SETAR)، برآوردگرهای روش خودگردان اتورگرسیو، دارای اریبی است که این اریبی با افزایش حجم نمونه، کاهش نمی‌یابد.

جوردانو و همکاران [۵] روش جدیدی برای ساخت بازه‌های پیش‌بینی ناپارامتری برای داده‌های سری زمانی غیرخطی ارائه دادند. در این روش، خودگردان اتورگرسیو توسط بولمان [۳، ۴ و ۶] برای فرایند خطی پیشنهاد و برای تقریب فرایند غیرخطی، مدل‌های شبکه عصبی به کار گرفته شد. بولمان [۴ و ۶] به معرفی این روش در سری زمانی و مدل قدم زدن تصادفی پرداخت.

جوردانو و همکاران [۷] روش جدیدی مبتنی بر روش‌های غیرخطی ارائه دادند. آن‌ها به جای روش خودگردان اتورگرسیو، از روشی مبتنی بر مدل‌های غیرخطی شبکه عصبی پیش‌خورانه^۳

1- Exponential Autoregressive

2- Self-exiting Threshold Autoregressive

3- Feed-Forward Neural network

استفاده کردند. آن‌ها همچنین نشان دادند ایده اعمال روش خودگردان روی مانده‌های حاصل از شبکه عصبی پیش‌خورانه نتایج بهتری به دست می‌دهد.

جوردانو و همکاران [۵] کاربرد روش خودگردان را در پیش‌بینی سری‌های زمانی غیرخطی با استفاده از شبکه عصبی معرفی نمودند. آن‌ها به معرفی روش جدیدی برای ساختن بازه‌ی پیش‌بینی در سری زمانی غیرخطی پرداختند. این روش با بهره‌گیری از روش‌های شبکه عصبی و خودگردان، به بهبود پیش‌بینی می‌پردازد. جوردانو و همکاران [۷] خواص نظری و حدی شبکه عصبی خودگردان را ارائه کردند. آن‌ها ابتدا با برازش مدل شبکه عصبی پیش‌خورانه با یک لایه پنهان، مانده‌های مدل شبکه عصبی را برآورد کرده، سپس با بازنمونه‌گیری از مانده‌های شبکه عصبی و برآورد تابع توزیع تجربی آماره موردنظر، اندازه دقت این برآوردگرها را از روش شبکه عصبی خودگردان محاسبه کردند. پرز رودریگز و آندرادا فلیکس [۸] آزمونی برای استقلال و هم‌توزیعی مانده‌های مدل $GARCH^1$ معرفی نمودند. مقادیر بحرانی این آماره آزمون برای اندازه نمونه کوچک با روش‌های خودگردان به‌دست‌آمده و این ورودی‌ها برای شبکه عصبی به‌کاربرده شده‌اند. سیانی و دی پرتی [۹] برای آزمون ناهمسانی واریانس در مدل ARCH، با استفاده از خودگردان و شبکه عصبی موردبررسی قرار دادند.

در این مقاله در بخش ۲ الگوریتم‌های خودگردان اتورگرسیو، خودگردان بلوکی و خودگردان مانده‌ها در سری‌های زمانی ارائه می‌گردند. در بخش ۳ مدل‌های شبکه عصبی معرفی و در ادامه الگوریتم شبکه عصبی خودگردان ارائه می‌شود. در بخش ۴ در یک مطالعه شبیه‌سازی چهار روش مختلف مورد مقایسه قرار می‌گیرند. در بخش ۵ روش‌های ارائه‌شده برای تحلیل قیمت نفت سنگین ایران مورد استفاده قرار گرفته و در نهایت در بخش ۶ نتیجه‌گیری ارائه می‌شود.

۲- الگوریتم‌های خودگردان در سری‌های زمانی

افران [۱۰] روش خودگردان را بر اساس ایده‌ی بازنمونه‌گیری از داده‌ها برای برآورد اندازه‌های دقت برآوردگرها مانند اریبی، واریانس و توزیع برآوردگرها و همچنین بازه‌ی اطمینان و آزمون فرض برای پارامترهای نامعلوم ارائه نمود. برای حفظ ساختار همبستگی داده‌ها در نمونه‌های خودگردان می‌توان از روش‌های خودگردان نیم پارامتری و ناپارامتری استفاده نمود. روش خودگردان نیم پارامتری بر اساس بازنمونه‌گیری از مانده‌های مدل است که مدل موردنظر با استفاده از معیارهای انتخاب مدل، برازش می‌یابد. این معیارها توسط آکائیک [۱۱] و [۱۲] معرفی شد. در زیر روش خودگردان اتورگرسیو ارائه می‌شود.

۲-۱- روش خودگردان اتورگرسیو

سری زمانی $\{Y_t\}_{t \in Z}$ یک فرایند مانای اتورگرسیو از مرتبه p است هرگاه به صورت

$$Y_t = \phi_1 Y_{t-1} + \dots + \phi_p Y_{t-p} + \varepsilon_t$$

باشد، که در آن ε_t از توزیعی نامعلوم با میانگین صفر و واریانس ثابت σ_ε^2 می‌باشد. هدف محاسبه اندازه دقت برآوردگر $T = t(Y_1, \dots, Y_n)$ است. مراحل مختلف الگوریتم خودگردان اتورگرسیو بر اساس سری زمانی مشاهده شده Y_1, \dots, Y_n به صورت زیر است:

۱- با استفاده از معیار اطلاع تصحیح اریبی شده

$$AICC = -n \log(\sigma_\varepsilon^2) + 2(p+1)n / (n-p-2)$$

مرتبه سری زمانی اتورگرسیو (p) برآورد می‌گردد.

۲- با استفاده از معادلات یول والکر پارامترهای $\hat{\Phi}_p = (\hat{\phi}_1, \dots, \hat{\phi}_p)'$ برآورد می‌گردند.

۳- برآورد مانده‌های سری زمانی اتورگرسیو به صورت $\hat{\varepsilon}_t = \sum_{j=0}^p \hat{\phi}_j (Y_{t-j} - \bar{Y})$ به دست می‌آیند، که در آن $\hat{\phi}_0 = 1, \quad \hat{\phi}_1, \dots, \hat{\phi}_p$

۴- برآورد مانده‌های مدل به صورت $\tilde{\varepsilon}_t = \hat{\varepsilon}_t - \bar{\hat{\varepsilon}}_t$ مرکزی می‌گردند، که در آن $\bar{\hat{\varepsilon}}_t = (n-p)^{-1} \sum_{t=p+1}^n \hat{\varepsilon}_t$ میانگین مانده‌ها است.

۵- نمونه خودگردان مانده‌های $\varepsilon_{p+1}^*, \dots, \varepsilon_n^*$ از توزیع تجربی مانده‌های مرکزی شده $F_n(\varepsilon)$ به صورت مستقل و هم‌توزیع به روش نمونه‌گیری تصادفی ساده با جایگذاری از $\tilde{\varepsilon}_{p+1}, \dots, \tilde{\varepsilon}_n$ تولید می‌شوند.

۶- نمونه خودگردان Y_1^*, \dots, Y_n^* به ازای $t \leq p$ به صورت $Y_t^* = \bar{Y}$ و به ازای $t > p$ از رابطه $\sum_{j=0}^p \hat{\phi}_j (Y_{t-j}^* - \bar{Y}) = \varepsilon_t^*$ به دست می‌آیند. (برای به دست آوردن نمونه خودگردان مناسب بر اساس مدل مورد نظر، ابتدا $n+100$ نمونه تولید و سپس 100 نمونه اول حذف می‌گردد.)

۷- برآوردگر T بر اساس نمونه خودگردان به صورت $T^* = t(Y_1^*, \dots, Y_n^*)$ محاسبه می‌شود.

۸- مراحل ۵ تا ۷، را B بار تکرار و T_1^*, \dots, T_B^* به دست می‌آید.

۹- اندازه‌های دقت برآوردگر T مانند اریبی و واریانس به صورت

$$\hat{bias}^*(T) = \bar{T}^* - T \quad \text{و} \quad \hat{Var} = \frac{1}{B-1} \sum_{i=1}^B (T_i^* - \bar{T}^*)^2$$

محاسبه می‌گردند که در آن $\bar{T}^* = \frac{1}{B} \sum_{i=1}^B T_i^*$. برای اطلاع بیشتر در مورد برآورد اندازه‌های دقت به روش خودگردان به کتاب افرون و تیشیرانی [۱۳] مراجعه نمایید.

۲-۲- روش خودگردان بلوک متحرک

در روش خودگردان بلوک متحرک (MBB)^۱ به عنوان یک روش ناپارامتری برای سری زمانی Y_1, \dots, Y_n ، بلوک‌های متداخل به اندازه b به صورت $B_t = \{Y_t, \dots, Y_{t+b-1}\}$ به ازای $t=1, \dots, n-b+1$ در نظر گرفته می‌شوند. اگر $k = n/b$ عددی صحیح باشد، k بلوک خودگردان B_1^*, \dots, B_k^* با نمونه‌گیری تصادفی ساده با جایگذاری از $n-b+1$ بلوک مشاهدات به دست آورده و نمونه خودگردان Y_1^*, \dots, Y_n^* از به هم پیوستن k بلوک خودگردان حاصل می‌شود که در آن بلوک‌های خودگردان به صورت $B_i^* = \{Y_{(i-1)b+1}^*, \dots, Y_{ib}^*\}$ در نظر گرفته می‌شوند. در حالتی که n/b عدد صحیح نباشد آنگاه k برابر با کوچک‌ترین عدد صحیح است که $kb \geq n$ ، در این صورت k بلوک خودگردان B_1^*, \dots, B_k^* با نمونه‌گیری تصادفی ساده با جایگذاری از $n-b+1$ بلوک مشاهدات به دست آورده و نمونه‌ی خودگردان برابر با n مقدار اولیه‌ای است که از به هم پیوستن k بلوک خودگردان حاصل می‌شود. برای مثال بلوک‌های متحرک به طول $b=3$ برای سری زمانی به طول $n=6$ به صورت

$$X_1, X_2, X_3, X_4, X_5, X_6$$

است. ادامه مراحل مانند مراحل ۷ الی ۹ روش خودگردان اتورگرسیو است.

۲-۳- روش خودگردانی مانده‌ها

در روش خودگردان مانده‌ها با استفاده از معیار AICC، بهترین مدل برآورد می‌گردد. در این روش با برازش مدل مقادیر مانده‌ها برآورد و سپس به صورت $\hat{\varepsilon}_t = \bar{\varepsilon}_t - \varepsilon_t$ مرکزی می‌گردند،

که در آن $\bar{\hat{\epsilon}}_t = (n-p)^{-1} \sum_{t=p+1}^n \hat{\epsilon}_t$ است. در ادامه با استفاده از روابط بازگشتی و با بازنمونه‌گیری از مقادیر مانده‌های مرکزی شده، نمونه‌ی خودگردان Y_1^*, \dots, Y_n^* محاسبه و مقدار برآوردگر خودگردان $T^* = t(Y_1^*, \dots, Y_n^*)$ محاسبه می‌شود. درنهایت با تکرار B بار مراحل قبل مانند روش خودگردان نیم‌پارامتری اتورگرسیو، اندازه‌های دقت آماره موردنظر محاسبه می‌شوند. در این مقاله این روش برای مدل‌های اتورگرسیو (AR)، اتورگرسیو نمایی (EXPAR)، اتورگرسیو آستانه‌ای خودمحرک (SETAR) و اتورگرسیو تغییر وضعیت هموار (STAR) به کار می‌رود.

۳- مدل‌های شبکه عصبی

مدل‌های شبکه عصبی مصنوعی، الگوهایی ریاضی هستند که به‌طور مستقیم از روی مغز و کارکردهای مغز الگوبرداری شده‌اند و با مدل‌سازی و روش‌های مختلف، ابتدا آموزش می‌بینند و سپس با انعطاف‌پذیری بالا خطا را کاهش می‌دهند. شبکه عصبی مصنوعی از جمله مدل‌هایی است که قادر به برآورد انواع مختلف روابط خطی و غیرخطی در داده‌ها است. شبکه‌های عصبی مصنوعی دارای انعطاف‌پذیری بالایی در محاسبات برای مدل‌های غیرخطی می‌باشند و می‌توانند طیف وسیعی از توابع را با درجه دقت بالا تخمین بزنند. یک مدل شبکه عصبی همانند مدل‌های سری زمانی دارای پارامترهایی است. تعداد این پارامترها وابسته به تعداد لایه پنهان و تعداد نرون در لایه پنهان است. تعداد لایه پنهان برای برآورد بهتر روابط غیرخطی متغیرها تغییر می‌کند. تعداد وزن‌های شبکه تعداد این پارامترهاست. با افزایش تعداد لایه‌های پنهان و حجم آن، روابط غیرخطی بهتر مدل می‌شوند اما این افزایش یک مقدار بهینه دارد که با استفاده از مینی‌موم کردن میانگین توان دوم حاصل می‌شود.

در شبکه‌های عصبی چندلایه، یک‌لایه ورودی وجود دارد که اطلاعات را دریافت می‌کند، تعدادی لایه پنهان وجود دارد که اطلاعات را از لایه‌های قبلی می‌گیرد. در اصل وجود لایه‌ی پنهان زمانی مفید است که تابع تبدیل غیرخطی باشد. درنهایت یک‌لایه خروجی وجود دارد که نتیجه محاسبات به آن رفته و خروجی آن، خروجی نهایی شبکه است.

جوردانو و همکاران [۷] مهم‌ترین دلایل در استفاده از شبکه‌های عصبی پیش‌خورانه را به‌صورت زیر ارائه کرده‌اند:

- ۱) شبکه عصبی پیش‌خورانه یک رده از مدل‌های عمومی هستند که در سری‌های زمانی غیرخطی با دقت بالایی به پیش‌بینی می‌پردازد.
- ۲) مدل‌های شبکه عصبی هیچ حساسیتی روی بعد داده‌ها ندارند. بنابراین اگر با داده‌های با ابعاد بالا سروکار داشته باشیم، اولین مدل ناپارامتری، مدل شبکه عصبی است.

۳) مباحث نظری این مدل‌ها به خوبی بسط یافته است.

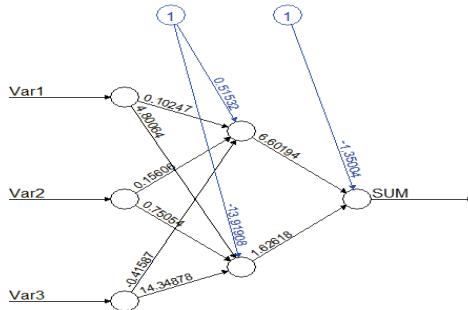
فرض کنید $\{Y_t ; t=1,2,3,\dots\}$ یک فرآیند تصادفی با مقادیر حقیقی است که به صورت

$$Y_t = g(Y_{t-1}) + \varepsilon_t$$

مدل‌سازی می‌شوند، که در آن ε_t دارای میانگین صفر و واریانس ثابت σ_ε^2 است. تابع نامعلوم $g(\cdot)$ می‌تواند توسط یک مدل شبکه عصبی پیش‌خورانه با یک لایه پنهان به صورت

$$h(y_{t-1}; \eta) = \sum_{k=1}^m c_k L(a_k y_{t-1} + b_k) + c_0$$

تقریب زده شود، که در آن m حجم لایه پنهان، a_k وزن ارتباطی بین لایه ورودی و لایه پنهان، $k=1, \dots, m$ و c_k وزن پیوند (ارتباط) بین نرون لایه پنهان و خروجی، b_k و c_0 به ترتیب آریبی لایه‌های پنهان و خروجی، و $L(\cdot)$ تابع فعال‌سازی لایه پنهان می‌باشند. بردار پارامترهای $\eta = (c_0, c_1, \dots, c_m, a_1, a_2, \dots, a_m, b_1, \dots, b_m)$ به صورت η تعریف می‌گردد. به‌طور معمول در برنامه‌های شبکه عصبی کاربردی، تابع سیگموئید، لوژستیک یا تانژانت هیپربولیک به‌عنوان تابع فعال‌سازی در نظر گرفته می‌شود که با استفاده از این توابع، شبکه‌های عصبی یک لایه می‌توانند به‌عنوان یک تقریب مناسب مورد استفاده قرار گیرند [۱۴].



شکل ۱: شبکه عصبی با یک لایه پنهان به حجم ۲ نرون و ۳ متغیر ورودی

در شکل ۱ شبکه عصبی با یک لایه پنهان که دارای دو نرون در لایه پنهان و سه ورودی است نشان داده شده است. بردارهایی که با عدد ۱ مشخص شده‌اند، آریبی لایه و سایر بردارها، وزن‌های شبکه می‌باشند. شبکه فوق دارای ۱۱ پارامتر است. الگوریتم اجرا شده برای داده‌های واقعی به این صورت است که با ورود اولین بردار از داده‌ها، بردار پارامترها را محاسبه کرده، سپس با ورود دومین بردار ورودی از داده‌ها، پارامترها تعدیل شده تا نهایتاً داده‌ها در یک دوره

تکراری مثلاً ۲۰ بار به شبکه آموزش می‌دهند و آن قدر بردار پارامترها تعدیل می‌یابد تا خطا حداقل گردد. شبکه‌های تک لایه توانایی پیاده‌سازی توابع غیرخطی را ندارند به همین دلیل از شبکه‌هایی که از چندلایه تشکیل شده‌اند استفاده می‌کنیم. این شبکه‌ها دارای توانایی‌های بیشتری هستند. از کان [۱۵] از مدل‌های شبکه عصبی برای مدل‌بندی مؤلفه خطی و غیرخطی در پیش‌بینی نرخ ارز کشور ترکیه استفاده کرد.

۳-۱- الگوریتم شبکه عصبی خودگردان

فرض کنید Y_1, \dots, Y_n مقادیر مشاهده‌شده سری زمانی، θ پارامتر موردنظر و همچنین $\hat{\theta} = t(Y_1, \dots, Y_n)$ یک برآوردگر بر اساس داده‌ها باشد، مراحل مختلف الگوریتم شبکه عصبی خودگردان به صورت زیر است:

(۱) با داشتن Y_1, \dots, Y_n یک شبکه عصبی پیش‌خورانه، با حجم لایه پنهان افزایشی بر روی n داده سری زمانی برازش می‌دهیم. حجم لایه پنهان \hat{m} ، با معیار AICC، به صورت

$$AICC = -n \log(\hat{\sigma}^2) + 2(l+1)n / (n-l-2)$$

انتخاب می‌شود، که در آن l تعداد وزن‌های شبکه (تعداد وزن‌های لایه‌ها و اربیبی‌های لایه) و $\hat{\sigma}^2$ برآوردگر سازگار σ_ε^2 است.

(۲) در این مرحله ابتدا بردار وزن‌های شبکه را به صورت $(\hat{b}_m, \dots, \hat{b}_m, \hat{a}_m, \dots, \hat{a}_m, \hat{c}_m, \dots, \hat{c}_m)$ ، با مینیمم کردن میانگین توان دوم خطا، برآورد و سپس مانده‌ها به صورت

$$\hat{\varepsilon}_t = Y_t - h(Y_{t-1}; \hat{\eta}) = Y_t - \sum_{k=1}^m \hat{c}_k L(\hat{a}_k Y_{t-1} + \hat{b}_k) - \hat{c}_0; \quad t = 2, \dots, n$$

برآورد و در نهایت مانده‌های مرکزی شده به صورت $\tilde{\varepsilon}_t = \hat{\varepsilon}_t - (n-1)^{-1} \sum_{t=2}^n \hat{\varepsilon}_t$ محاسبه می‌شوند.

(۳) تابع توزیع تجربی مانده‌های مرکزی شده $\tilde{\varepsilon}_t$ ، به صورت $\hat{F}_{\tilde{\varepsilon}}(x) = (n-1)^{-1} \sum_{t=2}^n I(\tilde{\varepsilon}_t \leq x)$ تعیین می‌گردد که در آن $I(\cdot)$ تابع نشانگر است.

(۴) مانده‌های خودگردان $\varepsilon_2^*, \dots, \varepsilon_n^*$ به صورت مستقل و هم‌توزیع از توزیع تجربی $\hat{F}_{\tilde{\varepsilon}}(x)$ به صورت نمونه‌گیری تصادفی ساده با جایگذاری از $\tilde{\varepsilon}_2, \dots, \tilde{\varepsilon}_n$ با نمونه‌گیری شده و نمونه خودگردان Y_2^*, \dots, Y_n^* به صورت بازگشتی

$$Y_t^* = h(Y_{t-1}^*; \hat{\eta}) + \varepsilon_t^*, \quad t=2, \dots, n.$$

به دست می‌آیند و اولین نمونه خودگردان برابر میانگین $Y_1^* = \bar{Y}$ در نظر گرفته می‌شود.

(۵) برآوردگر خودگردان $\hat{\theta}^*$ به صورت $\hat{\theta}^* = t(Y_1^*, \dots, Y_n^*)$ محاسبه می‌گردد.

(۶) به‌طور معمول توزیع خودگردان برآوردگر $\hat{\theta}$ را می‌توان با شبیه‌سازی مونت کارلو با تکرار B بار مراحل ۴ و ۵ و محاسبه $\hat{\theta}_1^*, \dots, \hat{\theta}_B^*$ بر اساس تابع توزیع تجربی به صورت

$$\hat{F}^*(x) = B^{-1} \sum_{b=1}^B I(\hat{\theta}_b^* \leq x)$$

ارائه و اندازه‌های دقت مانند اریبی و واریانس برآوردگر $\hat{\theta}$ را مشابه مرحله ۹ الگوریتم خودگردان اتورگرسیو برآورد نمود.

۴- مطالعه شبیه‌سازی

به‌منظور مقایسه دقت و عملکرد الگوریتم‌های شبکه عصبی خودگردان، خودگردان اتورگرسیو، خودگردان بلوک متحرک و خودگردان مانده‌ها در سری زمانی، شبیه‌سازی از ۴ مدل

$$M_1: Y_t = -0/8Y_{t-1} + \varepsilon_t$$

$$M_2: Y_t = -0/8Y_{t-1} + 6/5Y_{t-1}\exp(-Y_{t-1}^2) + \varepsilon_t$$

$$M_3: Y_t = -0/8Y_{t-1} - 1/4 \frac{1}{1 + \exp(-10Y_{t-1})} Y_{t-1} + \varepsilon_t$$

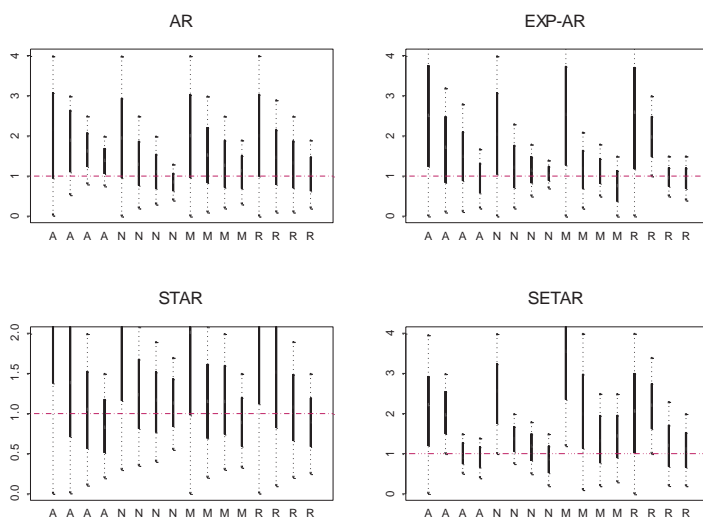
$$M_4: Y_t = (1/5 - 0/9Y_{t-1})I_t + (-0/4 - 0/6Y_{t-1})(1 - I_t) + \varepsilon_t$$

استفاده می‌شود که در آن I_t تابع نشانگر است و به صورت $I_t = \begin{cases} 1 & Y_{t-1} \leq 0 \\ 0 & Y_{t-1} > 0 \end{cases}$ تعریف می‌شود.

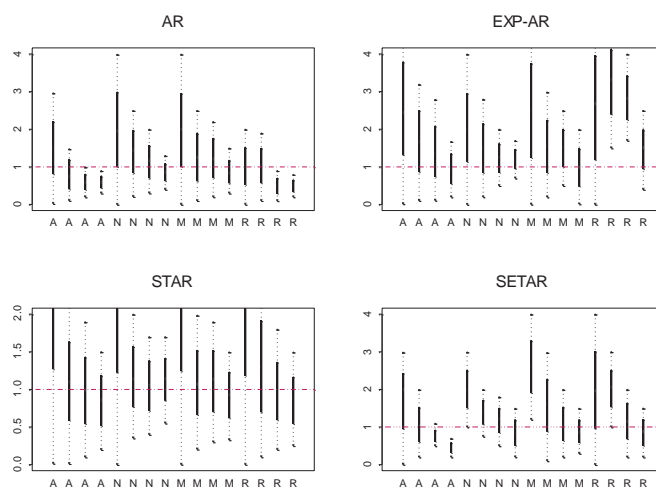
مدل M_1 اتورگرسیو خطی با مرتبه ۱ است. مدل M_2 اتورگرسیو نمایی مرتبه ۲ با یک چرخه محدود است. مدل M_3 یک مدل STAR است. مدل M_4 یک مدل SETAR با دو رژیم است. سه مدل M_2 ، M_3 و M_4 غیرخطی هستند. مانده‌های مدل‌ها دارای توزیع‌های نرمال استاندارد، چوله نرمال با پارامتر چولگی یک، تی استیودنت با پنج درجه آزادی و چوله-تی $St(0, \delta, 1)$ در نظر گرفته می‌شوند. پارامتر θ را به ترتیب میانگین (μ_T) ، واریانس (γ_0) و تابع اتوکوواریانس با تأخیر یک (γ_1) با برآوردگرهای \bar{Y} ، $\hat{\gamma}_0$ و $\hat{\gamma}_1$ که معیارهای مهم مرکزی، پراکندگی و خودهمبستگی داده‌هاست، در نظر می‌گیریم.

همچنین اندازه دقت را واریانس استاندارد شده برآوردگر $\hat{\theta}$ به صورت $\hat{R} = [\hat{\theta} - E(\hat{\theta})] / \sigma_{\hat{\theta}}$ در نظر می‌گیریم، که در آن $\sigma_{\hat{\theta}}$ خطای استاندارد $\hat{\theta}$ است که در شبیه‌سازی به‌طور مجزا با ۱۰۰۰۰۰ بار تکرار مونت‌کارلو تقریب می‌شود. در الگوریتم خودگردان اتورگرسیو، مرتبه مدل با استفاده از معیار $AICC$ محاسبه می‌شود. در این مطالعه برای هر مدل تعداد $N=2000$ سری زمانی به طول $n=100, 300, 500, 1000$ شبیه‌سازی می‌کنیم. همچنین تعداد تکرار خودگردان $B=1000$ در نظر گرفته می‌شود و $\hat{R}^* = [\hat{\theta}^* - E^*(\hat{\theta}^*)] / \sigma_{\hat{\theta}}$ محاسبه می‌شود. در هر شبیه‌سازی سری زمانی، برآوردگر $\hat{\theta}$ برای هر مدل، روش، توزیع مانده و اندازه نمونه تکرار می‌شود.

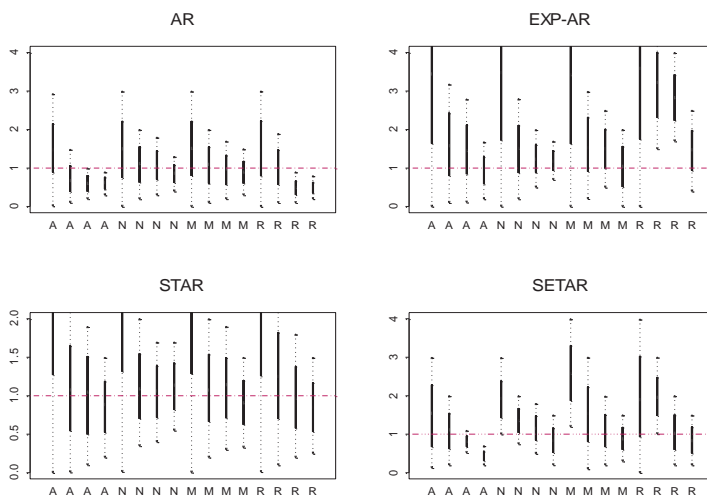
در شکل‌های ۲ تا ۱۳، نمودار جعبه‌ای $N=2000$ بار تکرار مونت‌کارلو برآورد خودگردان واریانس استاندارد شده برآوردگرهای میانگین، اتوکواریانس مرتبه صفر (واریانس)، اتوکواریانس نمونه‌ای مرتبه یک برای چهار مدل M_1 تا M_4 بر اساس ۴ توزیع نرمال استاندارد، نرمال چوله، تی استیودنت و چوله-تی برای مانده‌های مدل ارائه شده است. در هر نمودار چهار روش خودگردان اتورگرسیو (A)، شبکه عصبی خودگردان (N)، خودگردان بلوک متحرک (M) و خودگردان مانده‌ها (R) به ازای اندازه نمونه‌های به‌ترتیب $n=100, 300, 500, 1000$ رسم شده است.



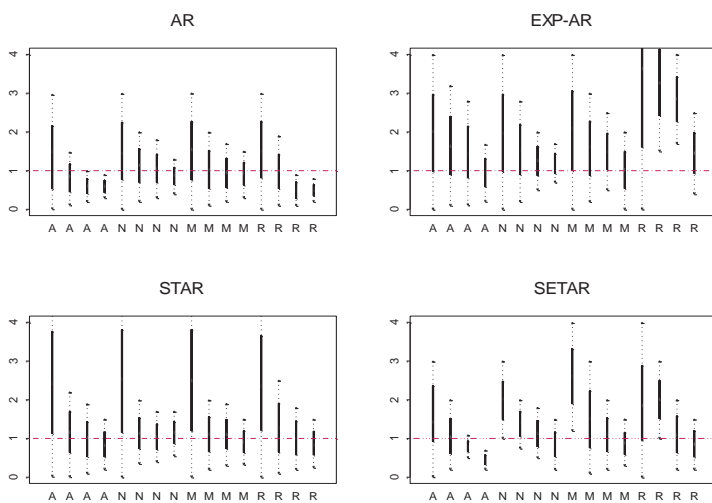
شکل ۲: نمودارهای جعبه‌ای برآورد خودگردان واریانس استاندارد شده \bar{Y} تحت توزیع نرمال استاندارد



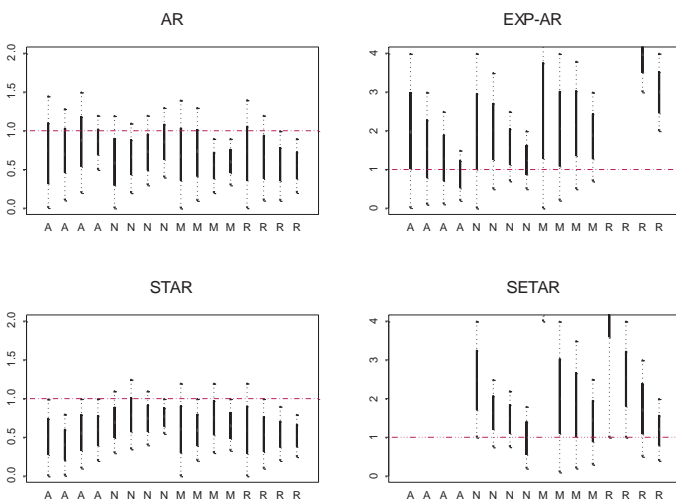
شکل ۳: نمودارهای جعبه‌ای برآورد خودگردان واریانس استاندارد شده \bar{Y} تحت توزیع چوله نرمال



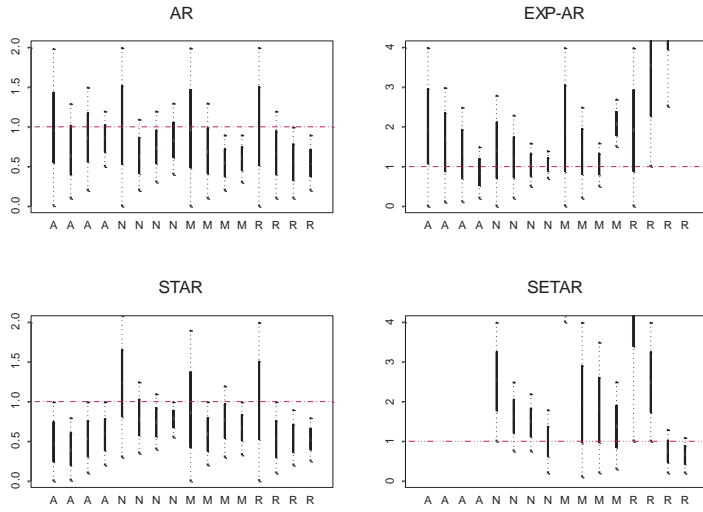
شکل ۴: نمودارهای جعبه‌ای برآورد واریانس خودگردان واریانس استاندارد شده \bar{Y} تحت توزیع تی



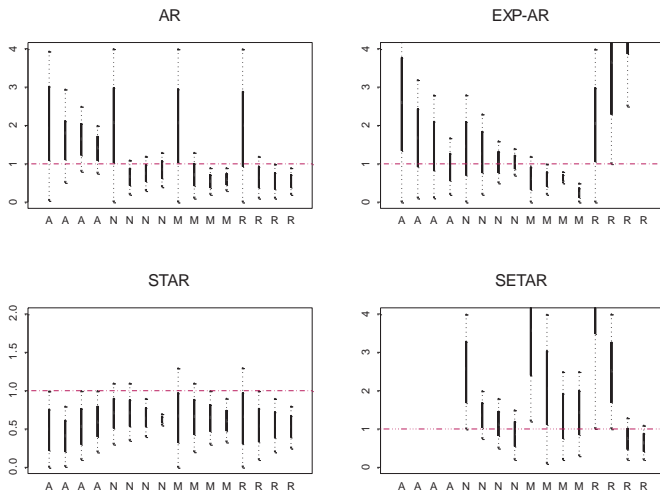
شکل ۵: نمودارهای جعبه‌ای برآورد خودگردان واریانس استاندارد شده \bar{Y} تحت توزیع چوله-تی



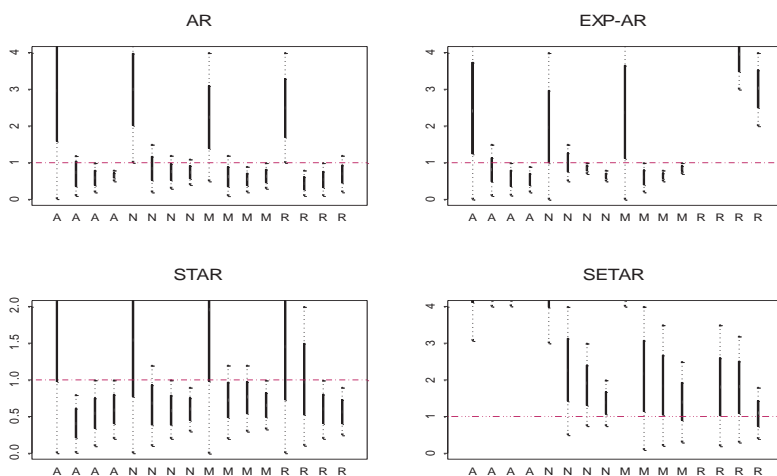
شکل ۶: نمودارهای جعبه‌ای برآورد خودگردان واریانس استاندارد شده $\hat{\gamma}_0$ تحت توزیع نرمال



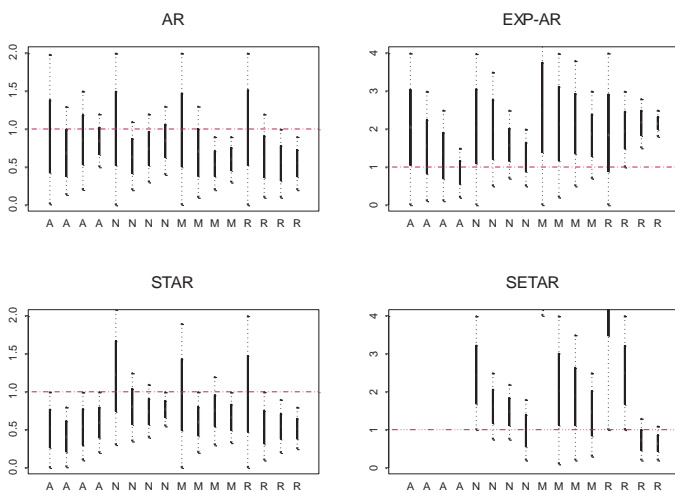
شکل ۷: نمودارهای جعبه‌ای برآورد خودگردان واریانس استاندارد شده $\hat{\gamma}_0$ تحت توزیع چوله نرمال



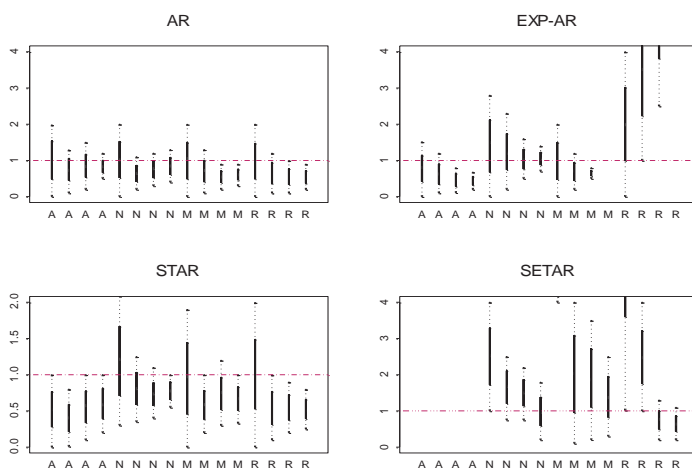
شکل ۸: نمودارهای جعبه‌ای برآورد خودگردان واریانس استاندارد شده $\hat{\gamma}_0$ تحت توزیع تی



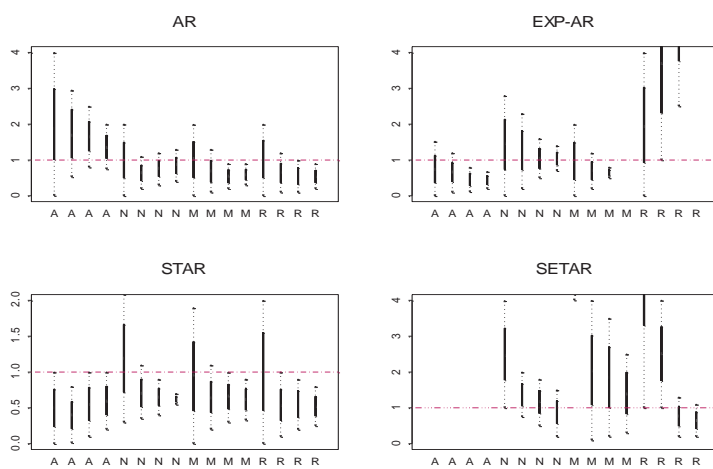
شکل ۹: نمودارهای جعبه‌ای برآورد خودگردان واریانس استاندارد شده $\hat{\gamma}_0$ تحت توزیع چوله-تی



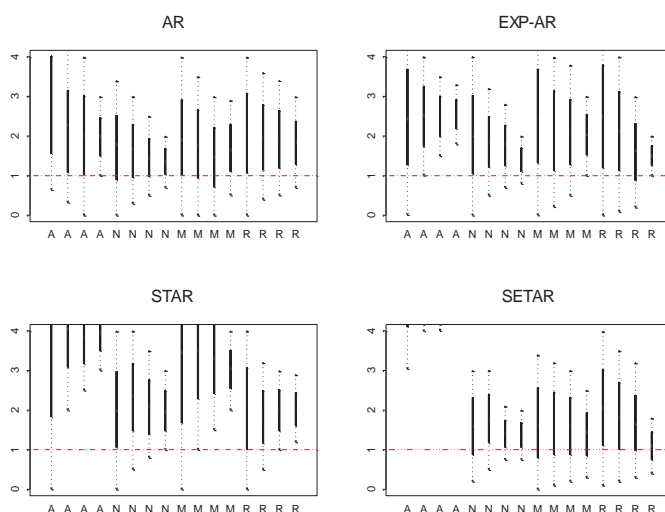
شکل ۱۰: نمودارهای جعبه‌ای برآورد خودگردان واریانس استاندارد شده $\hat{\gamma}_1$ تحت توزیع نرمال



شکل ۱۱: نمودارهای جعبه‌ای برآورد خودگردان واریانس استاندارد شده $\hat{\gamma}_1$ تحت توزیع چوله نرمال



شکل ۱۲: نمودارهای جعبه‌ای برآورد خودگردان واریانس استاندارد شده $\hat{\gamma}_1$ تحت توزیع تی



شکل ۱۳: نمودارهای جعبه‌ای برآورد خودگردان واریانس استاندارد شده $\hat{\gamma}_1$ تحت توزیع چوله-تی

شکل‌های ۲ تا ۵، نمودارهای جعبه‌ای برآورد خودگردان واریانس استاندارد شده \bar{Y} تحت توزیع‌های نرمال، چوله نرمال، تی و چوله-تی است. نمودارهای ۶ تا ۹، نمودارهای جعبه‌ای برآورد خودگردان واریانس استاندارد شده $\hat{\gamma}_0$ تحت چهار توزیع موردنظر است. نمودارهای ۱۰ تا ۱۳، نمودارهای جعبه‌ای برآورد خودگردان واریانس استاندارد شده $\hat{\gamma}_1$ تحت چهار توزیع فوق است. در تمام نمودارها با افزایش حجم نمونه برآورد واریانس برآوردگرهای موردنظر کاهش می‌یابد. در توزیع‌های نامتقارن میزان تغییرات واریانس $\hat{\gamma}_0$ و $\hat{\gamma}_1$ وقتی حجم نمونه از ۱۰۰ به ۳۰۰ تغییر می‌کند برای M_1 تا M_4 با شدت زیادی کاهش پیدا می‌کند اما برای سایر حجم نمونه این تغییر کوچک است.

رفتار برآورد واریانس برآوردگرها در مدل‌های M_1 تا M_4 نیز متفاوت است. در مدل M_1 تفاوت چندانی بین چهار روش برآورد واریانس وجود ندارد. پراکندگی واریانس هر سه برآوردگر تحت توزیع‌های مختلف با پارامترهای مختلف، تفاوت زیادی ندارد. علت این امر را می‌توان در لحاظ کردن ساختار خطی در مدل AR(1) دانست.

در مدل‌های غیرخطی M_2 ، M_3 و M_4 پراکندگی واریانس برآوردگرها دارای رفتار متفاوتی است. در این مدل‌ها پراکندگی واریانس $\hat{\gamma}_0$ و $\hat{\gamma}_1$ زیاد است. در توزیع‌های نامتقارن پراکندگی واریانس برآوردگرها افزایش می‌یابد. در این سه مدل روش خودگردان اتورگرسو نسبت به روش شبکه عصبی خودگردان از دقت بسیار کمتری برخوردار است. روش خودگردان

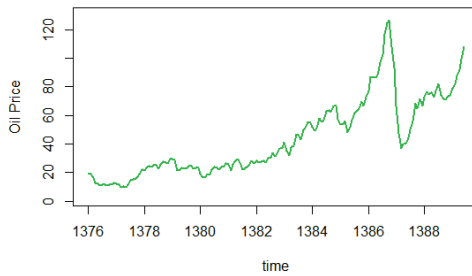
اتورگرسیو در مدل غیرخطی M_4 از کارایی کمتری برای محاسبه واریانس $\hat{\gamma}_0$ و $\hat{\gamma}_1$ برخوردار است. از این جهت روش‌های شبکه عصبی خودگردان، MBB و خودگردان مانده‌ها دارای عملکرد بهتری نسبت به روش خودگردان اتورگرسیو در مدل M_2 و M_4 است. وقتی توزیع چوله نرمال در تولید مانده‌ها مورد استفاده قرار می‌گیرد، روش خودگردان اتورگرسیو نسبت به روش شبکه عصبی خودگردان از دقت بسیار کمتری برخوردار است. در توزیع‌های نامتقارن اندازه دقت روش شبکه عصبی خودگردان نسبت به روش خودگردان اتورگرسیو افزایش می‌یابد. این موضوع در مورد γ_0 و γ_1 نیز برقرار است. اما برای \bar{Y} تفاوت چندانی بین روش‌های مختلف برآورد خودگردان واریانس وجود ندارد. در مدل خطی M_1 روش‌های مختلف برای برآورد خودگردان واریانس \bar{Y} دارای دقتی مشابه هستند. اما برای مدل‌های M_2 ، M_3 و M_4 روش شبکه عصبی خودگردان دارای دقت بهتری است.

به‌طور کلی در توزیع‌های نامتقارن تفاوت چندانی در برآورد خودگردان واریانس $\hat{\gamma}_1$ با استفاده از روش‌های شبکه عصبی خودگردان، خودگردان اتورگرسیو، MBB و خودگردان مانده‌ها مشاهده نمی‌گردد. واریانس سه برآوردگر در چهار روش در مدل‌های خطی مشابه است. در توزیع‌های نامتقارن چوله نرمال و چوله‌تی میزان تغییرپذیری برآورد خودگردان واریانس $\hat{\gamma}_1$ افزایش می‌یابد. تحت توزیع‌های نامتقارن روش شبکه عصبی خودگردان دقیق‌تر عمل می‌کند. این عملکرد تفاوت چندانی با سایر روش‌های برآورد واریانس سه برآوردگر ندارد. به‌طور خلاصه در مورد چهار روش برآورد واریانس، می‌توان گفت که روش خودگردان اتورگرسیو و شبکه عصبی خودگردان در مدل‌های خطی دارای دقت مشابهی هستند و دارای عملکرد بهتری نسبت به روش‌های خودگردان مانده‌ها و MBB هستند. اما در مدل‌های غیرخطی استفاده از روش خودگردان اتورگرسیو پیشنهاد نمی‌شود. در مدل‌های غیرخطی روش شبکه عصبی خودگردان از دقت بیشتری نسبت به سایر روش‌ها به‌خصوص روش خودگردان اتورگرسیو برخوردار است.

۵- مثال کاربردی

داده‌های مورد استفاده در این بخش مربوط به داده‌های قیمت ماهانه نفت سنگین ایران برحسب دلار طی سال‌های ۱۳۷۶ تا ۱۳۹۰ می‌باشد. نمودار سری زمانی فوق در شکل ۱۴ نشان داده شده است. روند صعودی در داده‌های قیمت ماهانه نفت سنگین ایران مشاهده می‌گردد. با توجه به نتایج شبیه‌سازی، روش شبکه عصبی خودگردان که دارای دقت بیشتری در برآورد اندازه دقت برآوردگرهای سری زمانی غیرخطی است، مورد استفاده قرار می‌گیرد. بهترین مدل بر اساس ملاک حداقل میانگین توان دوم خطا شبکه عصبی با یک‌لایه پنهان و ۵ نرون در لایه پنهان است. برای محاسبه اربیبی و خطای استاندارد سه برآوردگر \bar{Y} ، $\hat{\gamma}_0$ و $\hat{\gamma}_1$ ابتدا مدل شبکه عصبی مناسب روی داده‌ها برازش یافته سپس مانده‌های مدل برآورد شده و درنهایت

الگوریتم شبکه عصبی خودگردان با استفاده از بازنمونه‌گیری مانده‌ها بر اساس الگوریتم بخش ۳-۱ اندازه‌های دقت اریبی و خطای استاندارد سه برآوردگر محاسبه می‌گردد. تعداد تکرار خودگردان $B=1000$ در نظر گرفته شده است.



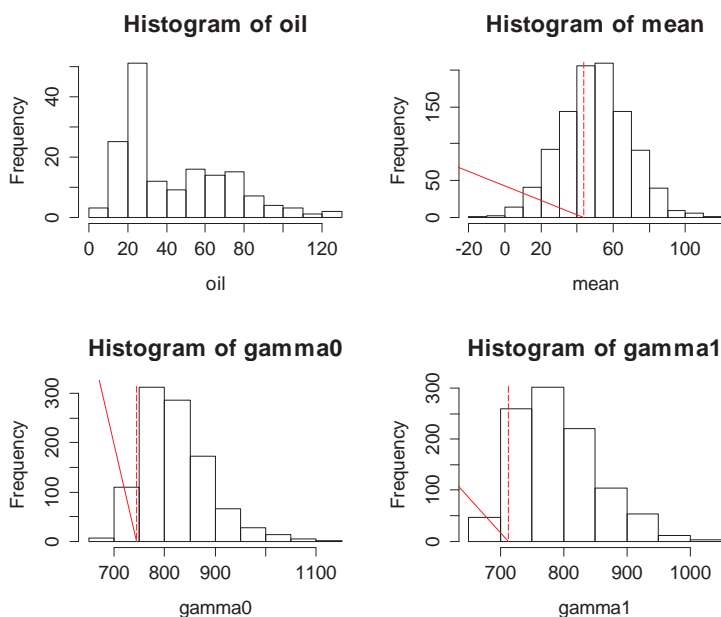
شکل ۱۴: نمودار داده‌های قیمت ماهانه نفت سنگین ایران.

در جدول ۱ مقادیر سه برآوردگر میانگین و اتوکواریانس‌های مرتبه صفر و یک به همراه برآورد اندازه‌های دقت اریبی و خطای استاندارد به روش شبکه عصبی خودگردان ارائه شده است. با استفاده از جدول، میزان اریبی \bar{Y} مشخص شده است. آماره‌های \hat{Y}_0 و \hat{Y}_1 اریب هستند. زیرا اریبی برابر صفر نیست. چولگی توزیع \hat{Y}_0 و \hat{Y}_1 نیز در اریبی این دو آماره تأثیرگذار است.

جدول ۱: مقادیر برآورد، اریبی و خطای استاندارد سه برآوردگر به روش شبکه عصبی خودگردان

خطای استاندارد	اریبی	برآورد	
۶/۲۸	۱/۱۲	۴۳/۷۴	\bar{Y}
۱۸۸/۱۲	۱۸/۱۴	۷۴۳/۹۴	\hat{Y}_0
۱۵۶/۹۸	-۱۵/۱۰	۷۱۲/۲۳	\hat{Y}_1

در شکل ۱۵ نمودار بافت‌نگار داده‌های قیمت ماهانه نفت سنگین ایران برای سه برآوردگر \bar{Y} ، \hat{Y}_0 و \hat{Y}_1 نشان داده شده است. با توجه به شکل ۱۵ توزیع \hat{Y}_0 و \hat{Y}_1 چوله به راست است، درحالی‌که بافت‌نگار \bar{Y} ، تقریباً متقارن و نرمال است. خطوط عمودی مقدار سه برآوردگر را نشان می‌دهند. مقدار اریبی دو برآوردگر اتوکواریانس‌های مرتبه صفر و یک نیز در شکل ۱۵ مشاهده می‌گردند که مخالف صفر است.



شکل ۱۵: بافت نگار داده‌ها و بافت‌نگار خودگردان میانگین و اتوکواریانس‌های مرتبه‌های صفر و یک

۶- نتیجه‌گیری

در این مقاله بعد از معرفی الگوریتم‌های خودگردان در سری‌های زمانی، الگوریتم شبکه عصبی خودگردان برای برآورد اندازه دقت برآوردگرهای سری زمانی خطی و غیرخطی معرفی شد. با فرض نرمال بودن مانده‌ها برآورد واریانس برآوردگرها در مدل خطی در روش خودگردان اتورگرسیون و شبکه عصبی خودگردان با همدیگر تقریباً معادل شد. اما در مدل‌های غیرخطی برآورد واریانس برآوردگرها در روش شبکه عصبی خودگردان مبتنی بر مانده‌های مدل شبکه عصبی بهتر برآورد می‌شود. همچنین در سری‌های زمانی غیرخطی حتی زمانی که مانده‌ها دارای توزیع نرمال هم باشند روش شبکه عصبی خودگردان دارای دقت بالاتری است. در توزیع‌های چوله نیز نتایج مشابه است. در یک نتیجه‌گیری کلی برای برآورد اندازه دقت برآوردگرهای سری‌های زمانی خطی و غیرخطی تحت توزیع‌های متقارن و چوله برای مانده‌های مدل استفاده از روش شبکه عصبی خودگردان پیشنهاد می‌شود.

مراجع

- [1] Thombs, L.A. and Schucany, W.R. (1990). Bootstrap prediction intervals for autoregression, *Journal of American Statistic Associat*, **85**, 486–492.
- [2] Alonso, A.M., Pena, D. and Romo, J. (2002). Forecasting time series with sieve bootstrap, *Journal of Statistical Planning and Inference*, **100**, 1–11.
- [3] Bühlmann, P. (1997). Sieve bootstrap for time series, *Bernoulli*, **3**, 123–148.
- [4] Bühlmann, P. (2002a). Bootstraps for time series, *Statistical Science*, **17**, 52–72.
- [5] Giordano, F., La Rocca, M. and Perna, C. (2007). Forecasting nonlinear time series with neural network sieve bootstrap. *Computational Statistics and Data Analysis*, **51**, 3871–3884.
- [6] Bühlmann, P. (2002b). Sieve bootstrap with variable -length Markov Chains for stationary categorical time series, *Journal of the American Statistical Association*, **97**, 443–471.
- [7] Giordano, F., La Rocca, M. and Perna, C. (2011). Properties of the Neural network sieve bootstrap. *Jorunal of Nonparametric Statistics*, **23**, 803-817.
- [8] Pérez-Rodríguez J.V. and Andrada-Félix J. (2013). Estimating critical values for testing the i.i.d in standardized residuals from GARCH models in finite sampeles. *Computational Statistics*, **28**, 701–734.
- [9] Siani, C. and de Peretti, C. (2007). Analysing the performance of bootstrap neural tests for conditional heteroskedasticity in ARCH-M models, *Computational Statistics and Data Analysis*, **51**, 2442–2460.
- [10] Efron, B. (1979). Another look at the jackknife. *Annals of Statistics*. **7**, 1–26.
- [11] Akaike, H. (1969). Fitting autoregressive models for prediction. *Ann. Inst.Statist. Math.* **21**, 243–247.
- [12] Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. *Proceedings of the Second International Symposium on Information Theory*. Akademiai Kiado, Budapest.
- [13] Efron, B. and Tibshirani, R. (1993). *An introduction to the Bootstrap*. Chapman and Hall.

- [14] Hagan, M.T., Demuth, H.B. and Beale, M. (1996). *Neural network design, Chap. 9*, PWS Publishing Company, Boston.
- [15] Özkan, F. (2013). Comparing the forecasting performance of neural network and purchasing power parity: The case of Turkey, *Economic Modelling*, **31**,752–758.

Comparison of the Bootstrap and Bootstrap Neural Network Methods in non Linear Time Series

Masoud Isapareh^{*}, Nasrollah Iranpanah^{*}, Marjan Kaedi^{**}

^{*}Department of Statistics, University of Isfahan, Isfahan, Iran

^{**}Department of Information Technology Engineering , University of Isfahan, Isfahan, Iran

Abstract

Neural networks are among those mathematical models which are used to model non-linear time series with high accuracy. The advantage with these linear times series as opposed to topical ones is that they don't require restrictive assumptions. The accuracy of neural network based estimators as nonparametric models is of high importance. In that light, we can use bootstrapping to calculate the accuracy of estimators in the time series' complex nonlinear structures. Though introduced in recent years these methods yield more accurate results in the bias calculation of estimators compared to the other ones. This paper introduces neural network bootstrap, bootstrap autoregressive, moving block bootstrap method and residual bootstrap methods in time . Then these four algorithms are compared with each other in a simulation study. Finally an example related to Iran's kerosene price monthly data is worked out.

Keywords: Artificial neural network, Auto regressive bootstrap, Residual bootstrap, Moving block bootstrap, Nonlinear time series, Monte Carlo simulation.

Mathematics Subject Classification (2010): 62F40, 37M10.