

مقایسه‌ی تجربی مدل‌های باکس - جنکینز، شبکه‌های عصبی مصنوعی و تحلیل مجموعه‌ی مقادیر تکین در پیش‌بینی سری‌های زمانی

*مسعود یارمحمدی^۱، *مهدی کلانتری و **رحیم محمودوند

**گروه آمار، دانشگاه پیام نور، صندوق پستی ۴۶۹۷-۱۹۳۹۵، تهران

**گروه آمار، دانشگاه بوعلی سینا

تاریخ دریافت: ۱۳۹۵/۷/۱۹ تاریخ پذیرش: ۱۳۹۵/۱۰/۱۷

چکیده: مدل باکس - جنکینز به‌عنوان یک روش پارامتری برای تحلیل سری‌های زمانی و برازش مدل‌های اتورگرسیو و میانگین متحرک فصلی و غیر فصلی به‌کار می‌رود؛ اما این روش برای سری‌های کوتاه مدت و نامانای مناسب نیست. در چنین شرایطی می‌توان از روش‌های ناپارامتری مانند شبکه‌های عصبی مصنوعی و تحلیل مجموعه‌ی مقادیر تکین استفاده کرد. هر دو روش از این امتیاز برخوردارند که نیازمند نرمال بودن توزیع خطاها و زیاد بودن تعداد داده‌ها نمی‌باشند. در این مقاله، پس از معرفی روش‌های فوق دقت آن‌ها در پیش‌بینی میزان فروش چهار نوع کالای غذایی، دارویی و بهداشتی یک شرکت پخش مواد غذایی و بهداشتی مقایسه می‌شود. علاوه بر آن در یک مطالعه شبیه‌سازی شده کارآمدی این روش‌ها برای پیش‌بینی‌های کوتاه مدت و بلند مدت ارزیابی شده است. نتایج حاصل برتری روش تحلیل مجموعه‌ی مقادیر تکین را در مقایسه با دو روش دیگر برحسب ریشه میانگین مربعات خطای پیش‌بینی نشان می‌دهد.

واژگان کلیدی: سری زمانی، مدل‌های باکس - جنکینز، روش تحلیل مجموعه‌ی مقادیر تکین، شبکه‌های عصبی مصنوعی.

رده‌بندی ریاضی: ۹۲B۲۰، ۳۷M۱۰.

۱- مقدمه

یکی از مهم‌ترین مدل‌های سری زمانی برای پیش‌بینی مقادیر آینده، مدل‌های میانگین متحرک جمع بسته اتورگرسیو فصلی^۱ است که معمولاً به مدل‌های باکس - جنکینز مشهور می‌باشند [۱]. نحوه‌ی برازش این مدل‌ها به داده‌های یک سری زمانی از اهمیت ویژه‌ای برخوردار بوده و مستلزم داشتن تجربه و مهارت کافی است.

از دیدگاه روش‌های ناپارامتری برای پیش‌بینی می‌توان به روش‌های شبکه‌های عصبی مصنوعی^۲ و تحلیل مجموعه‌ی مقادیر تکین^۳ اشاره کرد. شبکه‌های عصبی به دلیل دارا بودن ویژگی‌هایی نظیر پردازش موازی، هوشمندی و انعطاف‌پذیری جایگاه ویژه‌ای در حل مسائل پیچیده همچون تشخیص الگو، شناسایی و برآورد، خوشه‌بندی و مدل‌سازی دارد [۲-۵]. این روش با الگوبرداری مناسب از نرون‌های موجود در مغز انسان سعی می‌کند تا از طریق توابع تعریف شده ریاضی، رفتار درون سلولی نرون‌های مغز را شبیه‌سازی کرده و با اختصاص وزن‌های محاسباتی به خطوط ارتباطی نرون‌های مصنوعی، مدل مناسبی برای عملکرد سیناپسی نرون‌های طبیعی به دست آورد. یکی از کاربردهای مهم شبکه‌های عصبی در تحلیل سری‌های زمانی و پیش‌بینی مقادیر آینده‌ی سامانه‌های غیرخطی است. برای کسب اطلاعات بیشتر در مورد کاربردهای شبکه‌های عصبی در تحلیل سری‌های زمانی به [۶-۸] مراجعه کنید.

در سال‌های اخیر، روشی به نام تحلیل مجموعه‌ی مقادیر تکین که به اختصار آن را SSA می‌نامیم، در حوزه‌ی تحلیل سری‌های زمانی توسعه یافته است. این روش ناپارامتری به دلیل دارا بودن ویژگی‌هایی منحصر به فرد، کاربردهای وسیعی در بسیاری از شاخه‌های علوم از جمله هواشناسی، ژئوفیزیک، علوم دریایی، پزشکی، بیمه، مهندسی، پردازش تصویر، پردازش سیگنال، اقتصاد سنجی و ریاضیات مالی دارد. در روش SSA سری زمانی به اجزایی نظیر روند، مؤلفه‌های فصلی (با دوره تناوب‌های مختلف) و خطای تصادفی (نوفه) تجزیه می‌شود. به منظور انجام چنین تجزیه‌ی ای، از روش تجزیه‌ی مقدار تکین^۴ (SVD) که روشی شناخته شده و پرکاربرد در جبر خطی است، استفاده می‌شود. پس از آن، سری زمانی به کمک اجزای غیر تصادفی (روند و مؤلفه‌های فصلی) که سیگنال نامیده می‌شوند، بازسازی می‌شود. پس از بازسازی سری می‌توان به روش‌هایی نظیر روش بازگشتی و روش برداری، پیش‌بینی‌هایی در مورد آینده‌ی سری زمانی انجام داد.

1- Seasonal Autoregressive Integrated Moving Average (SARIMA)

2- Artificial Neural Networks

3- Singular Spectrum Analysis (SSA)

4- Singular Value Decomposition (SVD)

در این مقاله، نخست مرور مختصری بر مدل‌های باکس-جنکینز داشته و سپس روش شبکه عصبی مصنوعی و تحلیل مجموعه‌ی مقادیر تکین به تفصیل معرفی خواهند شد. در ادامه، یک مطالعه تجربی انجام شده و دقت پیش‌بینی هر سه مدل در داده‌های میزان فروش چهار نوع کالای غذایی، دارویی و بهداشتی یک شرکت پخش مواد غذایی و بهداشتی و داده‌های شبیه‌سازی شده مقایسه شده و بهترین روش پیش‌بینی مشخص خواهد شد.

۲- معرفی مدل‌های تحت بررسی

در این بخش مدل‌های باکس-جنکینز، شبکه‌های عصبی مصنوعی و تحلیل مجموعه‌ی مقادیر تکین به تفصیل معرفی خواهند شد.

۲-۱- مدل‌های باکس-جنکینز

یک مدل میانگین متحرک جمع بسته اتورگرسیو فصلی که از سوی باکس و جنکینز [۱] معرفی شده است، به صورت زیر بیان می‌شود:

$$\Phi_p(B^S)\phi_p(B)(1-B)^d(1-B^S)^D y_t = \Theta_Q(B^S)\theta_q(B)\varepsilon_t \quad (1)$$

که در آن $\theta_q(B)$, $\phi_p(B)$, $\Theta_Q(B^S)$, $\Phi_p(B^S)$ میانگین متحرک فصلی، اتورگرسیو غیر فصلی و میانگین متحرک غیر فصلی بوده و به صورت زیر تعریف می‌شوند:

$$\begin{aligned} \Phi_p(B^S) &= (1 - \Phi_1 B^S - \Phi_2 B^{2S} - \dots - \Phi_p B^{pS}) \\ \Theta_Q(B^S) &= (1 - \Theta_1 B^S - \Theta_2 B^{2S} - \dots - \Theta_Q B^{QS}) \\ \phi_p(B) &= (1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p) \\ \theta_q(B) &= (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q) \end{aligned}$$

همچنین ε_t و y_t به ترتیب سری زمانی و فرآیند تصادفی اغتشاش محض با میانگین صفر هستند. مدل (۱) را اغلب با نماد $SARIMA(p, d, q) \times (P, D, Q)_S$ نشان می‌دهند که در آن زیرنویس S نشان دهنده‌ی طول دوره‌ی فصلی است.

پیدا کردن مدل‌های مناسب برای سری‌های زمانی کاری چندان آسان نیست. برای یافتن یک مدل مناسب SARIMA که بتواند رفتار سری زمانی مشاهده شده را تبیین کرده و مقادیر آینده‌ی سری زمانی را پیش‌بینی کند، مراحل زیر باید طی شوند:

۱- شناسایی مدل ۲- برازش مدل ۳- ارزیابی مدل برازانده شده

مراحل فوق را می‌توان چندین بار به کاربرد تا در انتها به یک مدل مناسب دست‌یافت، ولی باید به این نکته توجه داشت که طی کردن این مراحل به‌منظور دست‌یابی به یک مدل مناسب، مستلزم محاسبات زیاد بوده و زمان‌بر است. خوشبختانه روش‌ها و نرم‌افزارهایی وجود دارند که یافتن یک مدل SARIMA مناسب را بسیار آسان کرده‌اند. در این مقاله از بسته‌ی forecast در نرم‌افزار R برای یافتن بهترین مدل استفاده‌شده است. در این مقاله، بهترین مدل بدین معنی است که کمترین مقدار معیار اطلاع آکائیک^۱ (AIC) را داشته باشد [۹].

۲-۲- شبکه عصبی مصنوعی

در این بخش به معرفی شبکه عصبی در چارچوب تحلیل سری‌های زمانی خواهیم پرداخت. در این راستا از مرجع [۱۰] بیشترین بهره را برده‌ایم. شبکه‌های عصبی مصنوعی و یا به عبارت ساده‌تر، شبکه‌های عصبی^۲ (NN) روش‌های پیش‌بینی هستند که بر پایه‌ی الگوهای ساده‌ی ریاضی مغز انسان بنا نهاده شده‌اند. این روش‌ها، امکان بررسی روابط غیرخطی پیچیده بین متغیر پاسخ و پیش‌بینی‌کننده‌های آن را میسر می‌کنند.

یک شبکه‌ی عصبی را می‌توان شبکه‌ای از نرون‌ها^۳ که در لایه‌هایی سازمان یافته‌اند، تصور نمود. پیش‌بینی‌کننده‌ها (یا ورودی‌ها) لایه‌ی زیرین و پیش‌بینی شونده‌ها (یا خروجی‌ها) لایه‌ی بالایی را تشکیل می‌دهند. ممکن است لایه‌های میانی شامل نرون‌های پنهان هم وجود داشته باشد. شبکه‌های بسیار ساده، لایه‌های پنهان ندارند و معادل رگرسیون خطی هستند. ضرایبی که به این پیش‌بینی‌کننده‌ها اختصاص داده می‌شود، «وزن‌ها» نامیده می‌شوند. پیش‌بینی‌ها به‌وسیله‌ی یک ترکیب خطی از ورودی‌ها به دست می‌آیند. در چارچوب شبکه‌های عصبی، وزن‌ها با استفاده از یک الگوریتم یادگیری^۴ انتخاب می‌شوند به‌طوری‌که یک تابع زیان نظیر کمترین مربعات خطا را کمینه سازند.

با افزودن لایه‌های میانی با نرون‌های پنهان، یک شبکه‌ی عصبی غیرخطی به دست می‌آید. یک مثال ساده در شکل ۱ نشان داده‌شده است. این شبکه که هر لایه از نرون‌ها، ورودی‌هایی از لایه‌های قبلی دریافت می‌کند، به شبکه‌ی چندلایه‌ی پیش‌خور^۵ معروف است. خروجی‌های نرون‌ها در یک لایه، ورودی‌هایی برای لایه‌ی بعدی هستند. ورودی‌های هر نرون با استفاده از یک ترکیب خطی موزون باهم ادغام می‌شوند. سپس نتیجه قبل از اینکه به‌عنوان خروجی ظاهر شود،

1- Akaike's Information Criterion (AIC)

2- Neural Networks (NN)

3- Neurons

4- Learning Algorithm

5- Multilayer Feed-Forward Network

به‌وسیله‌ی یک تابع غیرخطی اصلاح می‌شود. مثلاً ورودی‌های نرون پنهان Z در شکل ۱ به‌صورت خطی ترکیب می‌شوند تا مقدار

$$Z_j = b_j + \sum_{i=1}^F w_{i,j} x_i$$

به دست آید. سپس به‌منظور به دست آوردن ورودی برای لایه‌ی بعدی، این نتیجه در لایه‌ی پنهان با استفاده از یک تابع غیرخطی نظیر تابع حلقوی^۱ (تابع لجستیک استاندارد) $s(z) = (1 + e^{-z})^{-1}$ اصلاح می‌شود. این تبدیل منجر به کاهش اثر ورودی‌های کرانگین و در نتیجه ساختن شبکه‌ای نسبتاً استوار در مقابل نقاط دورافتاده می‌شود. توجه کنید که تعداد لایه‌های پنهان و تعداد نرون‌ها در هر لایه‌ی پنهان باید پیشاپیش تعیین شوند.

پارامترهای b_1, b_2, b_3 و $w_{1,1}, \dots, w_{4,3}$ از داده‌ها «برآورد» می‌شوند. اغلب اوقات برای جلوگیری از بزرگی بیش‌ازاندازه‌ی وزن‌ها، آن‌ها را مقید می‌کنند. پارامتری که وزن‌ها را مقید می‌کند به پارامتر زوال^۲ معروف است و اغلب برابر با ۰/۱ قرار داده می‌شود. در آغاز، وزن‌ها مقادیر تصادفی می‌گیرند، سپس با استفاده از داده‌های مشاهده شده به‌روز می‌شوند. در نتیجه جنبه‌ای از تصادف در پیش‌بینی‌های تولیدشده به‌وسیله‌ی یک شبکه‌ی عصبی وجود دارد. از این‌رو، معمولاً شبکه چندین مرتبه با استفاده از نقاط شروع تصادفی مختلف اجرا شده و نتایج میانگین‌گیری می‌شوند.

اگر سری زمانی از یک مدل اتورگرسیو پیروی کند، مقادیر گذشته سری می‌توانند به‌عنوان ورودی‌های یک شبکه‌ی عصبی مورد استفاده قرار گیرند. در این صورت شبکه‌ی عصبی مربوطه را یک شبکه‌ی عصبی اتورگرسیو^۳ نامند [۱۰]. برای نشان دادن یک شبکه‌ی عصبی اتورگرسیو پیش‌خور با یک لایه‌ی پنهان، p ورودی گذشته و k نرون در لایه‌ی پنهان، از نماد $NNAR(p, k)$ استفاده می‌شود [۱۰]. مثلاً مدل $NNAR(5, 9)$ نشان دهنده‌ی یک شبکه‌ی عصبی اتورگرسیو پیش‌خور با یک لایه‌ی پنهان با ورودی‌های $y_{t-1}, y_{t-2}, \dots, y_{t-9}$ (برای پیش‌بینی خروجی y_t) و پنج نرون در لایه‌ی پنهان است. توجه کنید که مدل $NNAR(p, 0)$ هم‌ارز است با مدل $AR(p)$ ولی بدون قیود روی پارامترها به‌منظور اطمینان از مانایی سری.

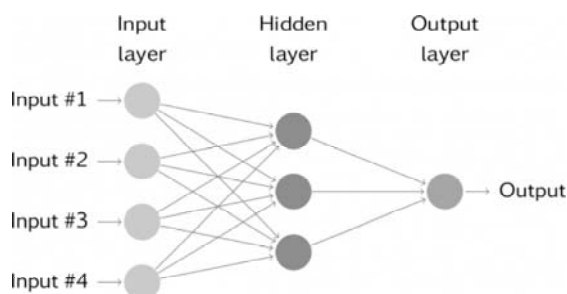
در یک سری زمانی فصلی با دوره‌ی تناوب m نیز نمادگذاری مشابهی وجود دارد. نماد $NNAR(p, P, k)_m$ نشان دهنده‌ی یک شبکه‌ی عصبی اتورگرسیو پیش‌خور با یک لایه‌ی پنهان با ورودی‌های $y_{t-1}, y_{t-2}, \dots, y_{t-p}, y_{t-m}, y_{t-2m}, \dots, y_{t-pm}$ و k نرون در لایه‌ی پنهان است. مثلاً مدل $NNAR(3, 1, 2)_{12}$ یک شبکه‌ی عصبی اتورگرسیو پیش‌خور با یک لایه‌ی

1- Sigmoid

2- Decay Parameter

3- Autoregressive Neural Network

پنهان با ورودی‌های $y_{t-1}, y_{t-2}, y_{t-3}$ و دو نرون در لایه‌ی پنهان، برای داده‌های فصلی با دوره‌ی تناوب ۱۲ را نشان می‌دهد. توجه کنید که مدل $NNAR(p, P, \circ)_m$ هم‌ارز است با مدل $SARIMA(p, \circ, \circ) \times (P, \circ, \circ)_m$ ولی بدون قیود روی پارامترها به‌منظور اطمینان از مانایی سری.



شکل (۱): یک شبکه‌ی عصبی با چهار ورودی و یک لایه‌ی پنهان متشکل از سه نرون پنهان.

در این مقاله، برای برازش یک مدل $NNAR(p, P, k)_m$ از نرم‌افزار R و تابع `nnetar` در بسته‌ی `forecast` استفاده شده است. در تابع `nnetar` این امکان فراهم شده است که مقادیر p ، P و k توسط کاربر نیز به‌صورت خودکار انتخاب شوند. در سری‌های زمانی غیرفصلی، مقدار p طوری انتخاب می‌شود که معیار `AIC` یک مدل $AR(p)$ کمترین مقدار را داشته باشد. ولی در سری‌های زمانی فصلی، مقدار پیش‌فرض P برابر با یک است و مقدار p نیز از یک مدل `AR` بهینه‌ی برازش داده‌شده به داده‌های تعدیل یافته‌ی فصلی به دست می‌آید. اگر مقدار k مشخص نشده باشد، برابر با $k = 2^{-1}(p + P + 1)$ قرار داده می‌شود (گرد شده به نزدیک‌ترین عدد صحیح). در این مقاله، برای پیش‌بینی میزان فروش کالاها از مقادیر پیش‌فرض فوق استفاده شده است. برای کسب اطلاعات بیشتر در مورد نحوه‌ی تعیین مقادیر p ، P و k به [۱۰] مراجعه کنید.

۳-۲- روش تحلیل مجموعه‌ی مقادیر تکین

اولین اثر چاپ شده که می‌توان به‌عنوان مبدأ پیدایش روش تحلیل مجموعه‌ی مقادیر تکین (و یا به‌طور کلی روش‌های مبتنی بر زیر فضا در پردازش سیگنال) در نظر گرفت، به قرن هجدهم برمی‌گردد [۱۱]. ولی ورود و ظهور اصلی `SSA`، با انتشار مقالات برومهید و کینگ در سال ۱۹۸۶ بود [۱۲-۱۳]. پس از آن رشد و توسعه‌ی `SSA` و به‌کارگیری آن در علوم مختلف شدت گرفت. اولین کتاب مقدماتی در مورد کاربرد `SSA` در تحلیل سری‌های زمانی توسط السنر و سونیس در سال ۱۹۹۶ به رشته‌ی تحریر در آمد [۱۴]. پس از آن در سال ۱۹۹۷ مجموعه‌ای از مقالات

آماردانان دانشگاه سن پترزبورگ در مورد SSA، در کتابی به زبان روسی گردآوری و چاپ شد [۱۵]. بعد از انتشار این کتاب، توجه بسیاری از محققین در حوزه‌ی تحلیل سری‌های زمانی به روش SSA جلب شد و تاکنون مقالات زیادی درباره به‌کارگیری این روش در تحلیل داده‌ها در علوم مختلف به چاپ رسیده است. از دیگر کتاب‌های مهم در این زمینه می‌توان به [۱۶-۱۸] اشاره کرد. برای کسب اطلاعات بیشتر در مورد SSA و آگاهی در مورد دیگر کاربردهای آن به [۱۹-۲۵] مراجعه کنید.

روش SSA از دو مرحله تشکیل شده است: تجزیه^۱ و بازسازی^۲. هر کدام از این مراحل نیز شامل دو گام هستند. برای تشریح مراحل فرض کنید $Y_N = \{y_1, y_2, \dots, y_N\}$ یک سری زمانی مشاهده شده به طول N بوده و L نیز عددی صحیح که طول پنجره^۳ نامیده می‌شود، باشد به طوری که $1 < L < N$. در این صورت مراحل SSA در بخش‌های ۱-۳-۲ و ۲-۳-۲ به صورت خلاصه ارائه شده است.

۱-۳-۲- تجزیه

این مرحله شامل دو گام است: نشانیدن^۴ و تجزیه‌ی مقدار تکین.

نشانیدن

در این گام، سری زمانی Y_N به K زیر سری X_1, \dots, X_K که $X_i = (y_i, \dots, y_{i+L-1})^T$ ، به ازای $i = 1, \dots, K$ و $K = N - L + 1$ تبدیل می‌شود. نتیجه این گام یک ماتریس X با ابعاد $L \times K$ است که ماتریس مسیره^۵ نامیده می‌شود و به صورت زیر تعریف می‌شود:

$$X = [X_1 : \dots : X_K] = \begin{bmatrix} y_1 & y_2 & y_3 & \dots & y_K \\ y_2 & y_3 & y_4 & \dots & y_{K+1} \\ y_3 & y_4 & y_5 & \dots & y_{K+2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ y_L & & y_{L+2} & \dots & y_N \end{bmatrix}$$

دقت کنید که عناصر روی قطرهای فرعی ماتریس X باهم برابرند. چنین ماتریسی را ماتریس هتکل^۶ می‌نامند. واضح است که با در اختیار داشتن ماتریس X ، می‌توان سری زمانی Y_N را به دست آورد. در واقع با کنار هم قرار دادن ستون اول و سطر آخر (یا سطر اول و ستون آخر)

- 1- Decomposition
- 2- Reconstruction
- 3- Window Length
- 4- Embedding
- 5- Trajectory Matrix
- 6- Hankel Matrix

ماتریس X ، سری زمانی Y_N حاصل می‌شود؛ بنابراین یک تناظر یک‌به‌یک بین ماتریس مسیر و سری زمانی وجود دارد.

تجزیه‌ی مقدار تکین

در این گام، تجزیه‌ی مقدار تکین (SVD) ماتریس مسیر را به دست می‌آوریم. فرض کنید $\lambda_1, \dots, \lambda_L$ که $\lambda_1 \geq \dots \geq \lambda_L \geq 0$ ، مقادیر ویژه ماتریس XX^T و U_1, \dots, U_L نیز بردارهای یک‌معامد ویژه‌ی متناظر با آن‌ها باشند. در این صورت SVD ماتریس X به صورت

$$X = X_1 + \dots + X_L$$

نوشته می‌شود که در آن برای $i = 1, \dots, L$ داریم $X_i = \sqrt{\lambda_i} U_i V_i^T$ و $V_i = X^T U_i / \sqrt{\lambda_i}$. سه‌تایی $(\sqrt{\lambda_i}, U_i, V_i)$ را سه‌تایی ویژه^۱ (ET) می‌نامند. گاهی اوقات U_i را بردار تکین چپ، V_i را بردار تکین راست و $\sqrt{\lambda_i}$ را مقدار تکین نامند.

۲-۳-۲- بازسازی

این مرحله شامل دو گام است: گروه‌بندی^۲ و میانگین‌گیری قطری^۳.

گروه‌بندی

هدف از اجرای این مرحله، تشخیص مؤلفه‌های سری زمانی نظیر روند، مؤلفه‌های فصلی، نوفه و غیره بر اساس گروه‌بندی سه‌تایی‌های ویژه‌ی ماتریس مسیر است. در این گام، پس‌ازاینکه SVD ماتریس مسیر به دست آمد، مجموعه‌ی اندیس‌های $\{1, \dots, L\}$ را به m زیرمجموعه‌ی I_1, \dots, I_m افراز می‌کنیم. فرض کنید $I = \{i_1, \dots, i_p\}$. در این صورت ماتریس X_I متناظر با گروه I به صورت $X_I = X_{i_1} + \dots + X_{i_p}$ تعریف می‌شود. مثلاً اگر $I = \{1, 4, 5\}$ آنگاه $X_I = X_1 + X_4 + X_5$. بدین ترتیب می‌توان X_I را به ازای $I = I_1, \dots, I_m$ به دست آورد. در این صورت از SVD ماتریس X نتیجه می‌گیریم که $X = X_{I_1} + \dots + X_{I_m}$. شیوه‌ی انتخاب مجموعه‌های I_1, \dots, I_m ، گروه‌بندی نامیده می‌شود. برای پیدا کردن سه‌تایی‌های ویژه‌ی مرتبط با اجزای سری زمانی، از اطلاعات نهفته در مقادیر تکین و بردارهای ویژه که در قالب نمودارهایی نظیر نمودار لگاریتم مقادیر تکین، نمودار بردارهای ویژه و نمودار پراکندگی بردارهای ویژه

1- Eigen Triple (ET)

2- Grouping

3- Diagonal Averaging

ارائه‌شده‌اند، می‌توان استفاده کرد. لازم به ذکر است که در عمل، تشخیص سه‌تایی‌های ویژه‌ی مرتبط با اجزای سری زمانی، همیشه کار آسانی نیست و در برخی مواقع از پیچیدگی‌هایی برخوردار است و موفقیت در این راه نیاز به تجربه و تبحر دارد. اطلاعات بیشتر در مورد نمودارهای فوق و راهنمایی‌های کلی برای گروه‌بندی در [۱۶-۱۷] ارائه شده است.

میانگین‌گیری قطری

هدف اصلی در این گام، تبدیل هر ماتریس \mathbf{X}_{I_j} ، $j=1, \dots, m$ ، محاسبه‌شده در گام گروه‌بندی به یک سری زمانی به طول N است. همان‌طور که در گام نشانیدن اشاره شد، با در اختیار داشتن یک ماتریس هنکل می‌توان سری زمانی متناظر با آن را به دست آورد؛ ولی ماتریس‌های \mathbf{X}_{I_j} که در مرحله‌ی گروه‌بندی به دست می‌آیند دارای خاصیت هنکلی نیستند. هنکل سازی ماتریس \mathbf{X}_{I_j} به وسیله‌ی میانگین‌گیری روی عناصر قطرهای فرعی انجام می‌شود. بدین ترتیب ماتریس \mathbf{X}_{I_j} به یک ماتریس هنکل تبدیل می‌شود. پس از این تبدیل، زیر سری بازسازی شده‌ی \mathbf{X} ، $Y_N^{(j)} = \{\tilde{y}_1^{(j)}, \dots, \tilde{y}_N^{(j)}\}$ به طول N به دست می‌آید؛ بنابراین با توجه به SVD ماتریس \mathbf{X} ، سری زمانی اصلی $Y_N = \{y_1, y_2, \dots, y_N\}$ به صورت مجموع m زیر سری بازسازی شده به شکل زیر بازسازی می‌شود:

$$y_t = \sum_{j=1}^m \tilde{y}_t^{(j)}, \quad t = 1, \dots, N$$

۳-۳-۲- پیش‌بینی در چارچوب SSA

هر چند که روش SSA یک روش ناپارامتری است و نیاز به هیچ‌گونه فرضی درباره‌ی مدل سری زمانی ندارد، ولی به‌منظور دست‌یابی به پیش‌بینی‌هایی در مورد آینده‌ی سری زمانی، فرض می‌کنیم که ساختار سری زمانی به‌وسیله‌ی مدل‌هایی به نام روابط بازگشتی خطی^۱ (LRR) قابل توصیف است. گوییم سری زمانی $Y_N = \{y_1, y_2, \dots, y_N\}$ به‌وسیله‌ی یک LRR تولیدشده است هرگاه ضرایب a_1, \dots, a_t وجود داشته باشند به‌طوری‌که

$$y_{i+t} = \sum_{k=1}^t a_k y_{i+t-k}, \quad 1 \leq i \leq N-t, a_t \neq 0, t < N$$

خوشبختانه فرض فوق محدودیت جدی ایجاد نمی‌کند؛ زیرا کلاس سری‌های زمانی تولیدشده به‌وسیله‌ی LRR به‌قدر کافی بزرگ است و بسیاری از سری‌های زمانی که در عمل با آن‌ها مواجه

1- Linear Recurrence Relation (LRR)

می‌شویم را در بر دارد. این کلاس سری‌هایی که به صورت ترکیب خطی حاصل ضرب سری‌های نمایی، چند جمله‌ای و همساز هستند را نیز شامل می‌شود [۱۷]. برای کسب اطلاعات بیشتر در مورد نظریه و کاربرد LRR در SSA به [۱۶] مراجعه کنید.

در چارچوب SSA، چندین روش برای پیش‌بینی وجود دارد. نکته‌ی قابل توجه در پیش‌بینی‌های SSA این است که نه تنها می‌توان سیگنال را پیش‌بینی کرد، بلکه می‌توان تک تک مؤلفه‌های سری زمانی که در مرحله‌ی گروه‌بندی تشخیص داده شده‌اند را نیز پیش‌بینی نمود. فرض کنید I مجموعه‌ی سه‌تایی‌های ویژه‌ای باشد که در مرحله‌ی گروه‌بندی انتخاب شده‌اند. برای پیش‌بینی روند سری زمانی کافی است I را برابر مجموعه‌ی سه‌تایی‌های ویژه‌ی متناظر با روند قرار دهیم و به طریق مشابه، برای پیش‌بینی مؤلفه‌ی فصلی سری زمانی، I برابر مجموعه‌ی سه‌تایی‌های ویژه‌ی متناظر با مؤلفه‌ی فصلی خواهد بود. برای پیش‌بینی سیگنال نیز I را برابر مجموعه‌ی اولین r سه‌تایی‌های ویژه‌ی قرار می‌دهیم، به عبارت دیگر $I = \{1, \dots, r\}$.

یکی از مهم‌ترین روش‌های پیش‌بینی در SSA که در این مقاله نیز از آن استفاده خواهیم کرد، روش بازگشتی است که در ادامه به معرفی آن می‌پردازیم. فرض کنید I مجموعه‌ی سه‌تایی‌های ویژه‌ی انتخاب شده در مرحله‌ی گروه‌بندی بوده و $U_i \in R^L, i \in I$ بردارهای ویژه‌ی متناظر باشند. همچنین فرض کنید $\underline{U}_i \in R^{L-1}$ برداری شامل اولین $L-1$ مؤلفه‌ی بردار U_i بوده و π_i آخرین مؤلفه‌ی بردار U_i باشد و $v^T = \sum_{i \in I} \pi_i^T$. گیریم $\tilde{Y}_N = \{\tilde{y}_1, \dots, \tilde{y}_N\}$ نیز سری بازسازی شده به وسیله‌ی I باشد. بر این اساس پیش‌بینی به روش بازگشتی از رابطه‌ی بازگشتی زیر حاصل می‌شود:

$$\hat{y}_i = \begin{cases} \tilde{y}_i, & i = 1, \dots, N \\ \sum_{j=1}^{L-1} a_j \hat{y}_{i-j}, & i = N+1, \dots, N+h \end{cases}$$

که در آن بردار $R = (a_{L-1}, \dots, a_1)^T$ به صورت زیر تعریف می‌شود:

$$R = \frac{1}{1-v^T} \sum_{i \in I} \pi_i \underline{U}_i.$$

برای اطلاع از سایر روش‌های پیش‌بینی در SSA و نحوه‌ی ساختن فواصل پیش‌بینی به روش‌های مونت کارلو و بوت استرپ به [۱۶-۱۷] مراجعه کنید.

۳- مطالعه تجربی

در این بخش بر مبنای داده‌های شبیه‌سازی شده و داده‌های واقعی، کارایی هر سه مدل مطرح شده در بخش قبل را با یکدیگر مقایسه می‌کنیم. قبل از معرفی نتایج و مقایسه کارایی مدل‌ها با یکدیگر، روش انجام محاسبه‌ها شرح داده شده است. لازم به ذکر است که بسته‌های نرم‌افزاری کارآمدی برای هر سه روش در نرم‌افزار R تهیه شده است که تا حد زیادی از حجم محاسبات این کار کاسته است به‌طور مشخص در بسته forecast تابع auto.arima برای انتخاب مدل بهینه SARIMA و تابع nnetar برای انتخاب بهینه NNAR مورد استفاده قرار گرفته است؛ اما برای روش SSA بسته موجود در R یعنی Rssa تنها محاسبه‌های اصلی را انجام داده و یافتن مدل بهینه مستلزم کد نویسی است. برای این منظور کد مناسب در نرم‌افزار R توسط نویسندگان تهیه شده است که مبنای شبیه‌سازی و تحلیل داده‌های واقعی قرار گرفته است. معیار مورد استفاده برای مقایسه عملکرد پیش‌بینی‌ها و همچنین تعیین بهترین پیش‌بینی با استفاده از روش SSA مبتنی بر RMSE است که در ادامه به تشریح نحوه محاسبه آن می‌پردازیم. فرض کنید برای محاسبه مقدار RMSE مربوط به پیش‌بینی h گام جلوتر می‌خواهیم از b بار پیش‌بینی روش SSA استفاده کنیم. در این صورت الگوریتم زیر بهترین پیش‌بینی ممکن از روش SSA را فراهم می‌کند.

۱. قرار می‌دهیم $t=1$.

۲. بر مبنای مشاهده‌های $Y_{N-b-h+t} = \{y_1, y_2, \dots, y_{N-b-h+t}\}$ پارامترهای روش SSA را به‌گونه‌ای می‌یابیم که بهترین پیش‌بینی از مشاهده $y_{N-b-h+t}$ را فراهم کند.

۳. بر مبنای پارامترهای به دست آمده از مرحله ۲، پیش‌بینی مشاهده y_{N-b+t} را محاسبه کرده و خطای آن را با e_{N-b+t} نشان می‌دهیم.

۴. مراحل ۲ و ۳ را برای مقادیر $t = 2, 3, \dots, b$ محاسبه کرده که در نهایت منجر به محاسبه بردار خطاهای e_{N-b+1}, \dots, e_N می‌شود.

۵. مقدار RMSE مربوط به پیش‌بینی h گام جلوتر از فرمول زیر محاسبه می‌شود:

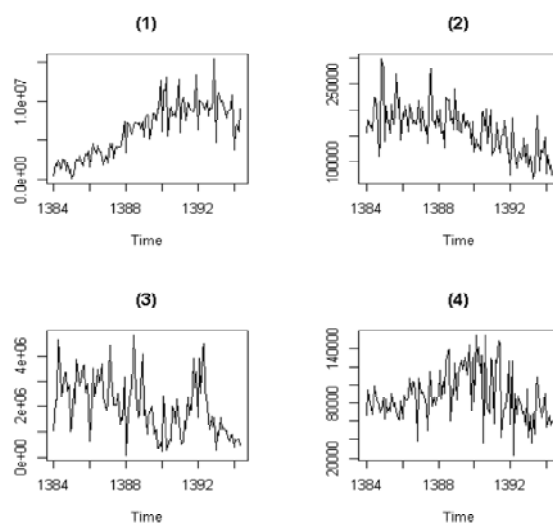
$$RMSE = \sqrt{\frac{1}{b} \sum_{t=1}^b e_{N-b+t}^2}$$

در استفاده از الگوریتم بالا باید توجه داشت که داده‌ها در هر بار اجرای الگوریتم به سه قسمت تقسیم می‌شود. بخش اول و دوم برای یافتن بهترین مدل پیش‌بین و بخش سوم برای بررسی کارایی واقعی مدل در مشاهده‌های خارج از نمونه. برای محاسبه مقدار RMSE در دو روش دیگر نیز از همین رویکرد (تقسیم مشاهده‌ها به سه قسمت) استفاده شده است. نکته دیگری که در

به کارگیری روش SSA باید به آن اشاره کرد، مربوط به پارامتر طول پنجره (L) است. مطابق با چند مطالعه انجام شده در این خصوص در نظر گرفتن مقدار این پارامتر نزدیک به نصف طول مشاهده‌ها ($N/2$) بهترین تقریب از سری را فراهم می‌کند ([۱۷-۱۸] و [۲۶-۲۸])؛ بنابراین در الگوریتم بالا از چنین جایگزینی استفاده شده است و در مرحله دوم مقدار پارامتر گروه‌بندی تعیین می‌شود. این کار باعث کاهش چشم‌گیر حجم محاسبات خواهد شد.

۱-۳- داده‌های واقعی

در این بخش داده‌های فروش ماهیانه‌ی چهار نوع کالا از محصولات یک شرکت پخش مواد غذایی و بهداشتی که از فروردین ۱۳۸۴ لغایت مرداد ۱۳۹۴ جمع‌آوری شده‌اند، استفاده شده است. کالای اول غذایی، کالای دوم دارویی و کالاهای سوم و چهارم هر دو بهداشتی می‌باشند. در شکل ۲ نمودار سری زمانی این داده‌ها نشان داده شده است. بر اساس این نمودارها، فروش کالای اول دارای یک روند نسبتاً صعودی است در حالی که روند فروش کالاهای ۲ و ۳ در دوره تحت بررسی نزولی است. روند تغییرات فروش کالای چهارم نشان دهنده یک روند افزایشی در بازه ۱۳۸۴ تا ۱۳۸۸ و پس از آن کاهش و برگشت فروش به میزانی کمتر از ابتدای دوره است.



شکل (۲): نمودار سری زمانی چهار نوع کالا

با به کارگیری الگوریتم ارائه شده در این بخش و بر اساس $b=5$ برای پیش‌بینی $h=1, 3, 6$ گام جلوتر نسبت RMSE مدل SSA به دو روش دیگر در جدول ۱ داده شده است. همان‌گونه که در این جدول ملاحظه می‌شود، روش SSA در پیش‌بینی فروش کالای ۱ موفق‌تر از دو روش

دیگر بوده است. در کالای دوم، مدل SARIMA پیش‌بینی‌های دقیق‌تری فراهم کرده است هر چند روش SSA نیز در پیش‌بینی یک گام جلوتر، بهتر از دو مدل دیگر است. در پیش‌بینی فروش کالای سوم، به‌جز برای افق پیش‌بینی ۱، مدل SSA بهتر عمل کرده است. در پیش‌بینی سری چهارم مدل SARIMA عملکرد بهتری نسبت به دو روش دیگر دارد. قابل‌ذکر است که مطابق بارانمای تابع (nnetar) تفاضل‌گیری برای ایستادن نمودن داده‌ها و تعیین مرتبه اتورگرسیو فصلی به‌صورت خودکار انجام نمی‌شود، بنابراین باید به‌صورت دستی این کار را انجام داد. با توجه به مرتبه اتورگرسیو فصلی در مدل SARIMA و بر اساس آزمایش و خطا مقادیر مختلفی از این پارامترها مورد آزمون قرار گرفت و در نهایت بهترین نتیجه ممکن برای این داده‌ها در نظر گرفته شد. توجه کنید که برای هر کالا با توجه به $b=5$ ، تعداد ۵ مدل SARIMA، ۵ مدل NNAR و ۵ مدل SSA به دست می‌آید؛ یعنی در مجموع تعداد ۶۰ مدل مختلف در مقایسه‌ها مورد استفاده قرار گرفته است. هر چند مرتبه برخی از این مدل‌ها یکسان است، به‌منظور جلوگیری از اطاله کلام از گزارش آن‌ها صرف‌نظر شد.

جدول (۱): مقادیر نسبی RMSE

نوع کالا	افق پیش‌بینی (h)	$\frac{SSA}{SARIMA} \times 100$	$\frac{SSA}{NNAR} \times 100$
کالای ۱	۱	۶۵	۴۴
	۳	۹۹	۷۶
	۶	۵۱	۲۸
کالای ۲	۱	۹۷	۷۴
	۳	۱۱۲	۷۲
	۶	۱۰۹	۵۱
کالای ۳	۱	۱۰۹	۸۲
	۳	۸۰	۷۰
	۶	۶۸	۵۹
کالای ۴	۱	۱۴۴	۱۰۲
	۳	۱۷۸	۱۲۹
	۶	۱۸۶	۱۰۵

۲-۳- مطالعه شبیه‌سازی

چنانکه در بخش قبل در تحلیل داده‌های واقعی دیدیم، روش SSA کارایی خوبی در پیش‌بینی کالای ۴ نداشت (جدول ۱). از طرفی نمودار این سری (شکل ۲) نشان‌دهنده مؤلفه‌های

هارمونیک در این سری است؛ بنابراین در این بخش از سری‌های هارمونیک شبیه‌سازی می‌کنیم. به‌علاوه مدل‌های شامل توابع مثلثاتی در بررسی و تقریب مدل‌های هارمونیک در حوزه‌های مختلفی چون داده‌های مالی و مهندسی کاربرد وسیعی دارند و در مطالعه‌های شبیه‌سازی معمولاً مورد استفاده قرار می‌گیرند. در این بخش داده‌های شبیه‌سازی شده از سری زمانی زیر تولید شده است.

$$y_t = \sin(w_1 \pi t) + \cos(w_2 \pi t) + z_t, \quad t = 1, \dots, T$$

که در آن $z_t \sim N(0, \sigma^2)$ ، فرآیند نوفه سفید است و w_1 و w_2 پارامترهای ثابت و مشخصی هستند. قابل ذکر است که مقادیر مختلفی از پارامترهای این سری توسط نویسندگان آزمون شده و با توجه به عدم تفاوت قابل توجه بین نتایج حاصل از مقایسه روش‌ها تصمیم گرفته شد تا به ذکر چند مورد پرداخته شود. در اینجا به بررسی موارد قید شده در جدول ۲ پرداخته شده است. علاوه بر این برای بررسی تأثیر میزان نوفه در مقایسه‌ها، دو سطح مختلف از واریانس نوفه در نظر گرفته شده است. برای این منظور نسبت واریانس سیگنال به واریانس نوفه (SNR) در دو سطح $\frac{1}{16}$ و $\frac{9}{16}$ در نظر گرفته شد. همچنین در تولید نمونه‌ها از اندازه نمونه $T = 100$ استفاده شده است.

جدول (۲): پارامترهای سری زمانی شبیه‌سازی شده

پارامترها		نام سری زمانی
w_2	w_1	
$\frac{2}{9}$	۰	سری زمانی ۱
۰	$\frac{2}{7}$	سری زمانی ۲
$\frac{2}{9}$	$\frac{2}{7}$	سری زمانی ۳

۱-۲-۳- رویه محاسبه نتایج

ابتدا باید متذکر شویم که به‌منظور مقایسه جامع‌تر، مانند آنچه در بخش داده‌های واقعی گفته شد، در اینجا نیز افق‌های پیش‌بینی مختلفی به کار گرفته شده است. برای انجام محاسبه‌ها و مقایسه روش‌ها از معیار $RMSE$ استفاده شده است؛ اما رویه محاسبه این معیار در شبیه‌سازی با داده‌های واقعی تفاوت دارد. برای توضیح یادآوری می‌کنیم که اگر \hat{S}_{t+h} برآوردگر S_{t+h} باشد آنگاه مقدار واقعی این معیار بر اساس فرمول زیر محاسبه می‌شود:

$$RMSE = \sqrt{E \left(\hat{S}_{T+h} - S_{T+h} \right)^2}$$

بنابراین در رویکرد شبیه‌سازی می‌توانیم با توجه به امکان تکرار پیش‌بینی، برآورد واقعی‌تری از این معیار به دست آوریم. برای این منظور به صورت زیر عمل شده است.

$$1- \text{قرار می‌دهیم } S_t = \sin(w, \pi t) + \cos(w, \pi t) \quad , \quad t = 1, \dots, T$$

$$2- \text{قرار می‌دهیم } j = 1$$

۳- نمونه تصادفی $z_t^{(j)}$, $t = 1, \dots, T - h$ از توزیع نرمال با میانگین صفر و واریانس برابر با ضریبی (سیگنال به نوفه) از واریانس نمونه‌ای S_1, \dots, S_{T-h} تولید می‌کنیم.

۴- سری زمانی شبیه‌سازی شده $y_t^{(j)} = S_t + z_t^{(j)}$, $t = 1, \dots, T - h$ را محاسبه می‌کنیم.

۵- بر مبنای مشاهده‌های $\{y_1^{(j)}, \dots, y_{N-h}^{(j)}\}$ پیش‌بینی مشاهده S_N را به دست می‌آوریم و خطای آن را با $e_N^{(j)}$ نشان می‌دهیم.

۶- مراحل ۲ تا ۴ را برای $j = 2, \dots, b$ محاسبه کرده که در نهایت منجر به محاسبه بردار خطاهای $e_N^{(1)}, \dots, e_N^{(b)}$ می‌شود.

۷- مقدار برآورد $RMSE$ مربوط به پیش‌بینی h گام جلوتر از فرمول زیر محاسبه می‌شود:

$$RMSE = \sqrt{\frac{1}{b} \sum_{t=1}^b e_N^{(j)^2}}$$

جدول (۳): پارامتر گروه‌بندی در روش SSA

نام سری زمانی	سری زمانی ۱	سری زمانی ۲	سری زمانی ۳
پارامتر گروه‌بندی	۲	۳	۴

برای شبیه‌سازی در هر سه روش از بسته‌های موجود در نرم‌افزار R که پیش‌تر اشاره شد، استفاده شده است. برای پارامترهای روش SSA مقدار $L=24$ و برای پارامتر گروه‌بندی با توجه به رتبه ماتریس مسیر از جدول ۳ استفاده شده است. شایان ذکر است که در اینجا سیگنال کاملاً مشخص است و به‌سادگی می‌توان دید که چه تعداد از مؤلفه‌ها برای ساخت سیگنال و چه تعداد برای ساخت نوفه استفاده می‌شود.

با در نظر گرفتن مواردی که در بالا گفته شد، نسبت $RMSE$ مربوط به روش SSA در جدول‌های ۴ تا ۶ ارائه شده است. به‌طور خلاصه نتایج نشان می‌دهند که روش SSA نسبت به دو روش دیگر از کارایی بالاتری در پیش‌بینی‌های کوتاه مدت و بلند مدت سری‌های زمانی ۱، ۲ و ۳ برخوردار است. همچنین در مقایسه نتایج مشخص است که روش $SARIMA$ از روش $NNAR$ توانایی بیشتری در پیش‌بینی این سری‌های زمانی داشته است. از نتایج دیگر این جداول می‌توان به کاهش دقت پیش‌بینی روش $SARIMA$ با افزایش افق پیش‌بینی اشاره کرد. با این‌وجود هر سه سری زمانی نشان می‌دهند که با افزایش میزان نوفه دقت روش $SARIMA$ در مقایسه با روش SSA رو به افزایش است.

جدول (۴): مقادیر نسبی $RMSE$ برای پیش‌بینی در سری زمانی ۱

$\frac{SSA}{NNAR} \times 100$	$\frac{SSA}{SARIMA} \times 100$	افق پیش‌بینی (h)	SNR
۴	۲۷	۱	$\frac{۱}{۱۶}$
۷	۱۱	۶	$\frac{۱}{۱۶}$
۱۵	۴۷	۱	$\frac{۹}{۱۶}$
۱۸	۳۷	۶	$\frac{۹}{۱۶}$

جدول (۵): مقادیر نسبی $RMSE$ برای پیش‌بینی در سری زمانی ۲

$\frac{SSA}{NNAR} \times 100$	$\frac{SSA}{SARIMA} \times 100$	افق پیش‌بینی (h)	SNR
۵	۲۱	۱	$\frac{۱}{۱۶}$
۸	۸	۶	$\frac{۱}{۱۶}$
۱۷	۳۵	۱	$\frac{۹}{۱۶}$
۱۸	۲۴	۶	$\frac{۹}{۱۶}$

جدول (۶): مقادیر نسبی $RMSE$ برای پیش‌بینی در سری زمانی ۳

$\frac{SSA}{NNAR} \times 100$	$\frac{SSA}{SARIMA} \times 100$	افق پیش‌بینی (h)	SNR
۶	۳۰	۱	$\frac{۱}{۱۶}$
۸	۱۶	۶	$\frac{۱}{۱۶}$
۱۴	۵۱	۱	$\frac{۹}{۱۶}$
۱۹	۳۰	۶	$\frac{۹}{۱۶}$

۴- نتیجه‌گیری

در این مقاله، مدل‌های باکس-جنکینز و شبکه‌های عصبی مصنوعی به‌اختصار معرفی شده و سپس به‌تفصیل به معرفی روشی نسبتاً جدید و ناپارامتری در حوزه تحلیل سری‌های زمانی به نام تحلیل مجموعه‌ی مقادیر تکین (SSA) پرداختیم. ضمن تشریح مراحل چهارگانه‌ی SSA، توضیحاتی در مورد پارامترهای این روش و یک روش پیش‌بینی به نام روش بازگشتی ارائه شد. برای مقایسه‌ی میزان دقت پیش‌بینی روش‌های باکس-جنکینز، شبکه‌های عصبی و SSA، از داده‌های ماهیانه‌ی فروش چهار نوع کالای غذایی، دارویی و بهداشتی یک شرکت پخش مواد غذایی و بهداشتی و مشاهده‌های شبیه‌سازی‌شده استفاده شد. نتایج حاصل از شبیه‌سازی نشان داد که استفاده از روش ناپارامتری SSA به‌جای روش‌های باکس-جنکینز و شبکه‌های عصبی، مقدار ریشه میانگین مربع خطای پیش‌بینی (RMSE) را به میزان قابل‌توجهی کاهش می‌دهد؛ اما نتایج مربوط به داده‌های واقعی برتری SSA نسبت به دو روش دیگر را در همه موارد نشان نداد. از این‌رو، روش SSA می‌تواند در برخی موارد به‌عنوان یک انتخاب در تحلیل سری‌های زمانی مورد استفاده قرار بگیرد.

تقدیر و تشکر

نویسندگان از زحمات سردبیر و همچنین داوران محترم مجله که نظرات ارزنده آن‌ها موجب بهبود مقاله گردید، تشکر و قدردانی می‌نمایند.

مراجع

- [1] Box, G.E.P., Jenkins, G.M. and Reinsel, G.C. (2008). *Time Series Analysis: Forecasting and Control*, 4th ed., Hoboken, New Jersey: John Wiley & Sons.
- [2] Krose, B. and Smagt, P.V.D. (1996). *An Introduction to Neural Networks*, 8th ed. University of Amsterdam.
- [3] Dunne, R.A. (2007). *A Statistical Approach to Neural Networks for Pattern Recognition*, Hoboken, New Jersey: John Wiley & Sons.
- [4] Rao, A.R. and Cecchi, G.A. (eds), (2012). *The Relevance of the Time Domain to Neural Network Models*, Springer.
- [5] Dreyfus, G. (2005). *Neural Networks: Methodology and Applications*, Springer.
- [6] Tirozzi, B., Puca, S., Pittalis, S., Bruschi, A., Morucci, S., Ferraro, E. and Corsini, S. (2005). *Neural Networks and Sea Time Series: Reconstruction and Extreme-Event Analysis*, Boston, Birkhauser.

- [7] Petridis, V. and Kehagias, A. (1998). *Predictive Modular Neural Networks: Applications to Time Series*, Springer, New York.
- [8] Chatfield, C. (2000). *Time Series Forecasting*, Chapman & Hall/CRC.
- [9] Hyndman, R.J. and Khandakar, Y. (2008). Automatic Time Series Forecasting: The forecast Package for R. *Journal of Statistical Software*, **27** (3), 1-22.
- [10] Hyndman, R.J. and Athanasopoulos, G. (2013). *Forecasting: principles and practice*. Section 9/3. <http://otexts.org/fpp/9/3>. Accessed on 2016-02-02.
- [11] De Prony, G. (1795). Essai expérimental et analytique sur les lois de la dilatabilité des fluides élastiques et sur celles de la force expansive de la vapeur de l'eau et la vapeur de l'alkool à différentes températures. *J de l'Ecole Polytechnique*. **1**(2), 24–76.
- [12] Broomhead, D. and King, G. (1986). Extracting qualitative dynamics from experimental data. *Physica D*, **20**, 217–236.
- [13] Broomhead, D., King, G. (1986b). On the qualitative analysis of experimental dynamical systems. In: *Sarkar S (ed) Nonlinear Phenomena and Chaos*. Adam Hilger, Bristol, 113–144.
- [14] Elsner, J.B. and Tsonis, A.A. (1996). *Singular spectrum analysis: a new tool in time series analysis*. New York: Springer.
- [15] Danilov, D. and Zhigljavsky, A. (Eds), (1997). *Principal components of time series: the "Caterpillar" method*. St.Petersburg Press, St. Petersburg (in Russian).
- [16] Golyandina, N., Nekrutkin V. and Zhigljavsky, A. (2001). *Analysis of time series structure: SSA and related techniques*, Boca Raton: Chapman and Hall/CRC.
- [17] Golyandina, N. and Zhigljavsky, A. (2013). *Singular Spectrum Analysis for Time Series*, London: Springer.
- [18] Sanei, S. and Hassani, H. (2016). *Singular Spectrum Analysis of Biomedical Signals*, Taylor & Francis/CRC.
- [19] Zhigljavsky, A. (2010). Singular Spectrum Analysis for Time Series: Introduction to this Special Issue, *Statistics and Its Interface*, **3**, 255-258.
- [20] Hassani, H. (2007). Singular Spectrum Analysis: Methodology and Comparison. *Journal of Data Science*, **5**(2), 239–257.
- [21] Hassani, H. and Thomakos, D. (2010). A review on Singular Spectrum Analysis for Economic and Financial Time Series, *Statistics and Its Interface*, **3**, 377-397.

- [22] Mahmoudvand, R.; Alehosseini, F.; Rodrigues, C.P. (2015). Forecasting Mortality Rate by Singular Spectrum Analysis. *RevStat-Statistical Journal*, **13**, 193-206.
- [23] Mahmoudvand, R. and Rodrigues, P.C. (2016). Missing value imputation in time series using Singular Spectrum Analysis. *International Journal of Energy and Statistics*, **4** (1), 1650005.
- [24] Rodrigues, C.P. and Mahmoudvand, R. (2016). Correlation Analysis in Contaminated Data by Singular Spectrum Analysis. *Quality and Reliability Engineering International*, **32**(6), 2127-2137.
- [25] Ghanati, R; Kazemhafizi, M; Mahmoudvand, R. and Fallahsafari, M. (2016). Filtering and parameter estimation of surface-NMR data using singular spectrum analysis. *Journal of Applied Geophysics*, **130**, 118-130.
- [26] Hassani, H., Mahmoudvand, R. and Zokaei, M. (2011). Separability and window length in singular spectrum analysis. *Comptes Rendus Mathematique*, **349** (17-18), 987-990.
- [27] Mahmoudvand, R. and Zokaei, M. (2012). On the Singular Values of the Hankel Matrix with Application in Singular Spectrum Analysis. *Chilean Journal of Statistics*, **3**, 43-56.
- [28] Hassani, H., Webster, A., Silva, E. S. and Heravi, S., (2015). Forecasting U.S. Tourist arrivals using optimal Singular Spectrum Analysis, *Tourism Management*, **46**, 322-335.
- [29] Golyandina, N. and Korobeynikov, A. (2014). Basic Singular Spectrum Analysis and Forecasting with R, *Computational statistics and Data Analysis*, **71**, 934-954.

Empirical Comparison of Box-Jenkins Models, Artificial Neural Network and Singular Spectrum Analysis in Forecasting Time Series

Masoud Yarmohammadi*, Mahdi Kalantari* and Rahim Mahmoudvand**

*Department of Statistics, Payame Noor University, P. O. Box 19395-4697, Tehran, Iran

**Department of Statistics, Bu-Ali Sina University, Hamedan, Iran

Abstract

The Box-Jenkins model is applied as a parametric method for time series analysis and fitting seasonal and non-seasonal autoregressive moving average models. But this procedure is not useful for short length and non stationary time series data. To overcome these problems, two nonparametric methods i.e. Artificial Neural Network and Singular Spectrum Analysis are introduced. These procedures do not require any statistical assumptions about normality of errors and could be used for short time series data. In this article, after introducing the above methods, their accuracy in forecasting sales of four types of food products, pharmaceutical and health care of a distribution Corporation are compared. Then using simulation studies, the effectiveness of these methods for short-term and long-term predictions are evaluated. The results show the superiority of Singular Spectrum Analysis compared to the other two methods in terms of the root mean square error of forecasting.

Keywords: Time Series, Box-Jenkins Models, Singular Spectrum Analysis, Neural Networks.

Mathematics Subject Classification (2010): 92B20, 37M10.