سیزدهمین کنفرانس آمار ایران
۲-۴ شهریور ۱۳۹۵

پیش‌بینی سری‌های زمانی یک متغیره فصلی بر اساس روش سری‌های زمانی تابعی

سیما شیخ ویسی[۱]، مجید عظیم محسنی[۱]، مهناز خلفی[۱]
[۱]دانشگاه گلستان

چکیدههدف در این مقاله، استفاده از تکنیک سری‌های زمانی تابعی برای پیش‌بینی سری‌های زمانی یک متغیره فصلی می‌باشد. همچنین در این مقاله، مقادیر پیش‌بینی به دست آمده را با مقادیر پیش‌بینی حاصل از طریق مدل‌سازی باکس-جنکینز مورد مقایسه قرار می‌دهیم. معیار ارزیابی در این مقایسه اریبی، میانگین مربعات خطای مدل و میانگین مربعات خطای پیش‌بینی می‌باشد.

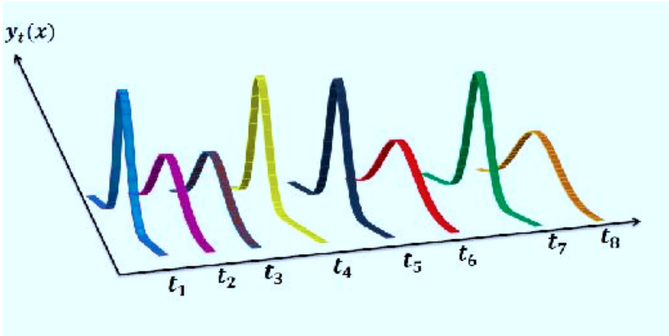
واژه‌های کلیدیسری زمانی تابعی، مولفه‌های اصلی تابعی، پیش‌بینی، دبی رودخانه
کد موضوع‌بندی ریاضی (۲۰۱۰): 62P12 ، 62M10 .

۱ مقدمه

سری زمانی تابعی $\{y_t(x_i), t = 1, ..., n, i = 1, ..., p\}$، دنباله‌ای از مشاهدات تابعی در طول زمان هستند. به عبارت دیگر، هنگامی که این منحنی‌ها، درطول یک دورهٔ زمانی مشاهده می‌شوند، یک سری زمانی تابعی را تشکیل می‌دهند (شکل۱) . بحث داده‌های تابعی اولین بار توسط رامسای و سیلورمن (۱۹۹۷) به کاربرده شده است. باسک (۲۰۰۰) ، سری‌های زمانی تابعی خطی را مورد مطالعه قرارداد و روی مدل اتو رگرسیو تابعی متمرکزشد. افرادی همچون هیندمن و شانگ (۲۰۰۹) ، دیدِرکسِن و همکاران (۲۰۱۱) و هورمن و کوکاسکا (۲۰۱۰) به مدل‌سازی و پیش‌بینی سری‌های زمانی تابعی پرداختند. یکی از روش‌های پیش‌بینی سری‌های زمانی تابعی، روش مولفه‌های اصلی تابعی است، که توسط هال و حسینی نسب (۲۰۰۶) مورد بررسی قرارگرفته است.

دراین مقاله به دنبال ارائه تکنیک سری‌های زمانی تابعی، جهت پیش‌بینی سری‌های زمانی فصلی یک متغیره هستیم. سپس این مقادیر پیش‌بینی شده را با مقادیر پیش‌بینی شده توسط روش باکس-جنکینز موردمقایسه قرارمی‌دهیم.

[۱]سیما شیخ‌ویسی : sima.shv.1992@gmail.com



شکل ۱: شکل نمادین از سری زمانی تابعی.

بخش دوم، چگونگی مدل‌سازی و پیش‌بینی سری‌های زمانی تابعی را با استفاده از روش مولفه‌های اصلی تابعی، شرح می‌دهد. بخش سوم، دیدگاه تابعی سری‌های زمانی فصلی یک متغیره را بیان می‌کند. بخش چهارم به مطالعات عددی و مقایسه مقادیر پیش‌بینی شده برای یک سری از داده‌های واقعی با استفاده از دو روش می‌پردازد. بخش پنجم، شامل بحث و نتیجه‌گیری می‌باشد.

۲ سری‌های زمانی تابعی

درعمل ، ساختار مشاهدات سری‌های زمانی تابعی به صورت زیر می‌باشد:

$$y_t(x_i) = f_t(x_i) + \sigma_t(x_i)\varepsilon_{t,i} \qquad t = 1, ..., n \qquad i = 1, ..., p$$

که $\{f_t(x_i)\}$ توابع معین زمانی، $\varepsilon_{t,i}$ یک نوفهٔ سفید است و $\sigma_t(x_i)$ تابع زمانی واریانس می‌باشد.

مدل‌سازی سری‌های زمانی تابعی، با استفاده از روش مولفه‌های اصلی تابعی به صورت زیر انجام می‌شود:

۱) پیداکردن پایه‌های متعامد $\varphi_1(x), \varphi_2(x), ..., \varphi_k(x)$

۲) تجزیه کردن $f_t(x)$ برحسب پایه‌ها و ضرایب مربوط به آن‌ها

$$f_t(x) = \bar{f}(x) + \sum_{k=1}^{K} \hat{\beta}_{t,k}\hat{\varphi}_k(x) + \hat{\varepsilon}_t(x)$$

که $\bar{f}(x) = \frac{1}{n}\sum_{t=1}^{n} f_t(x)$ برآورد تابع میانگین می‌باشد. $\hat{\varphi}_k(x)$ ، تابع ویژه عملگر کوواریانس زیر می‌باشد:

$$\hat{\Gamma}(x) = \frac{1}{n}\sum_{t=1}^{n}[f_t(x) - \bar{f}(x)][f_t(x) - \bar{f}(x)]$$

ضرایب $\hat{\beta}_{t,k}$ ، وزن‌های زمانی هستند، که میزان اثرگذاری هر تابع ویژه را در مدل کنترل می‌کنند، و به صورت زیر محاسبه می‌شوند:

$$\hat{\beta}_{t,k} =< f_t(x) - \bar{f}(x), \hat{\varphi}_k(x) >= \int_x [f_t(x) - \bar{f}(x)]\hat{\varphi}_k(x)dx$$

از آنجایی که k=1,...,K ،$\hat{\beta}_{t,k}$ سری‌های زمانی ناهمبسته می‌باشند، می‌توان به هر کدام یک مدل جداگانه برازش داد و بر اساس این

مدل‌ها، پیش‌بینی گام h ام $\hat{\beta}_{n+h|n,k}$ را محاسبه کرد. در نهایت پیش‌بینی گام h ام سری زمانی تابعی از فرمول زیر حاصل می‌شود:

$$\hat{y}_{n+h|n,k}(x) = \bar{f}(x) + \sum_{k=1}^{K} \hat{\beta}_{n+h|n,k}\hat{\varphi}_k(x)$$

۳ دیدگاه تابعی به سری زمانی فصلی یک متغیره

اگر $\{x_1, x_2, ..., x_n\}$ یک سری زمانی فصلی یک متغیره با دوره تناوب T باشد، کهn=kT ، برای مدل‌سازی و پیش‌بینی این سری زمانی معمولا از مدل‌های باکس-جنکینز فصلی استفاده می‌شود. در این بخش این داده‌ها را بر حسب یک سری زمانی تابعی بیان می‌کنیم، تا بتوانیم از تکنیک‌های پیش‌بینی سری‌های زمانی تابعی استفاده کنیم.

الگوریتم تبدیل یک سری زمانی فصلی یک متغیره با دوره تناوب T به سری زمانی تابعی، به صورت زیر می‌باشد:

۱) داده‌های سری زمانی $\{x_1, x_2, ..., x_n\}$ را برحسب یک سری زمانی تابعی $y_t(u)$ ، به فرم زیر می‌نویسیم:

$$y_t(u) = x_{(t-1)T+u} \qquad t = 1, ..., k \qquad u = 1, ..., T$$

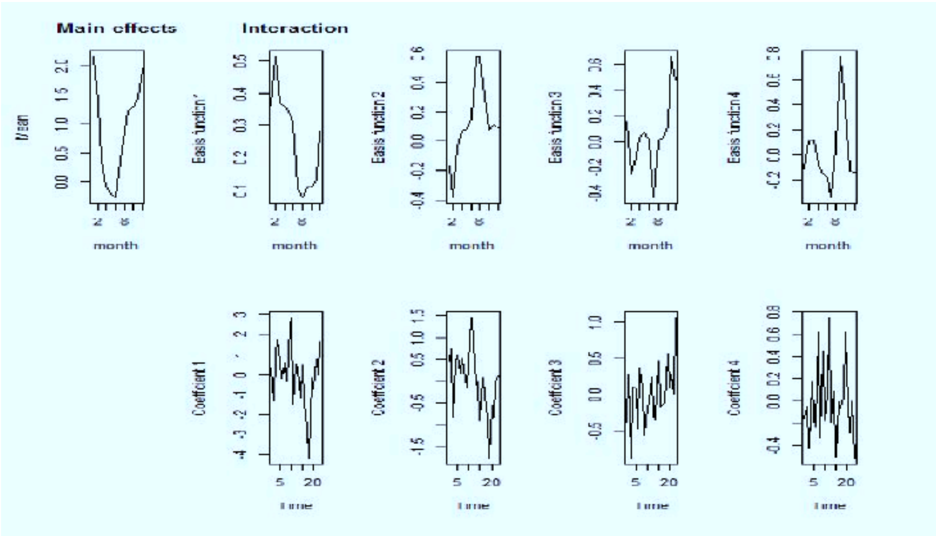
۲) سری زمانی تابعی $y_t(u)$ را هموارسازی می‌کنیم.

با الگوریتم بالا می‌توان مقادیر پیش‌بینی برای یک دوره تناوب آینده را، از تکنیک سری زمانی تابعی به دست آورد.

۴ داده‌های واقعی

برای ارزیابی روش این مقاله و ارائه راه‌کار، از داده‌های ماهانه دبی رودخانه، که دارای دوره تناوب ۱۲ می‌باشد، استفاده می‌شود. دراین بخش، ما از لگاریتم مقادیر ثبت شده ماهانهٔ دبی رودخانه کاسکان، که یکی از شاخه‌های رود کارون می‌باشد و از سال‌های [۲۰۱۲-۱۹۸۸] به ثبت رسیده‌اند، استفاده کرده‌ایم. برای مقایسه روش سری‌های زمانی تابعی با مدل باکس-جنکینز فصلی، هر دو مدل به داده‌ها برازش داده می‌شود و دو مدل از لحاظ اریبی و میانگین مربعات خطای مدل و پیش‌بینی با یکدیگر مقایسه می‌شوند.

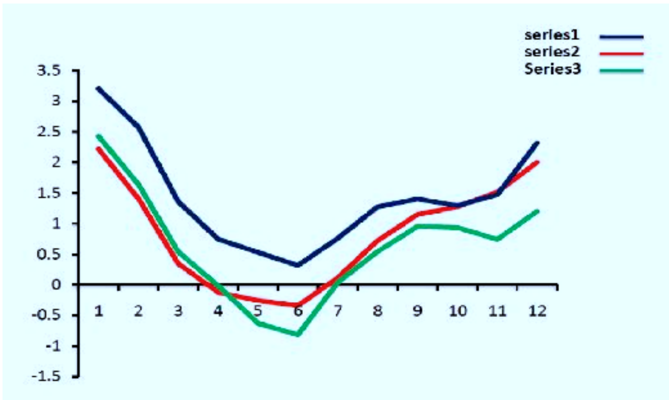
بااستفاده از روش سری‌های زمانی تابعی، یک مدل با K=۴ پایه انتخاب می‌شود. معیار انتخاب K=۴ کمترین میزان خطای پیش‌بینی می‌باشد، هیندمن و یولاه (۲۰۰۷) . با برازش دادن این مدل به داده‌ها ۹۶ درصد از تغییرات، توسط مولفه‌های اصلی تابعی، توجیه شده است. توابع ویژه و ضرایب مربوط به آن‌ها در شکل ۲ نشان داده شده است.



شکل ۲: سطر بالا به ترتیب، میانگین تابعی و چهار تابع ویژه اول را نشان می‌دهد.سطر پایین،نشان دهندهٔ وزن‌های زمانی مربوط به هر تابع ویژه می‌باشد.

بااستفاده از روش باکس-جنکینز فصلی، مدل$(۲,۰,۳) \times (۰,۱,۳)$ را به داده‌های سری زمانی برازش دادیم. مقدار معنی‌داری آماره باکس-پیرس، برای نکویی برازش $p - value = ۰/۶۷۷$ را گزارش داد، که نشان دهندهٔ برازش مناسب مدل می‌باشد.

مقادیر پیش‌بینی از هر دو روش، به همراه مقادیر واقعی آن‌ها در شکل ۳ رسم شده است. سری ۱ نشان دهندهٔ مقادیر واقعی لگاریتم دبی رودخانه کاسکان در سال ۲۰۱۳ می‌باشد، سری۲ مربوط به مقادیر پیش‌بینی شده سال ۲۰۱۳ با استفاده از روش سری‌های زمانی تابعی و سری۳ مربوط به مقادیر پیش‌بینی شده سال ۲۰۱۳ به روش باکس-جنکینز می‌باشد.



شکل ۳: مقادیر واقعی لگاریتم دبی سال ۲۰۱۳ ومقادیر پیش‌بینی شده آن‌ها از دو روش سری‌های زمانی تابعی و باکس-جنکینز.

شاخص‌های ارزیابی و مقایسه دو مدل در جدول ۱ گزارش شده است.

جدول ۱: اریبی،متوسط مربعات خطای مدل و پیش‌بینی.

روش	متوسط مربعات خطای پیش‌بینی	متوسط مربعات خطای مدل	اریبی
باکس-جنکینز	۰/۷۱۰۱۴۱۷	۰/۲۷۱۴	۰/۸۰۷۵۲۱۷
سری‌های زمانی تابعی	۰/۵۰۳۶۳۳۴	۰/۰۲۸۳۱	۰/۶۰۱۳۹۳۳

با توجه به شکل۳ و نتایج جدول ۱، روش سری زمانی تابعی، پیش‌بینی هایی با اریبی و خطای کمتر ارائه می‌دهد.

۵ بحث و نتیجه گیری

معمولا تکنیک‌هایی که در استنباط آماری در ابعاد بالاتر شکل می‌گیرد، تعمیمی از روش‌های یک متغیره می‌باشد. در این مقاله سعی بر این شد که با رویکردی متفاوت، با سری‌های زمانی یک متغیره فصلی از دیدگاه تابعی برخورد گردد. همانطور که از نتایج عددی حاصل گردید، این تکنیک، یک روش کارا در پیش‌بینی سری‌های زمانی فصلی ارائه می‌دهد، که با انتخاب مناسب پایه‌ها و هموارسازی کارا حتی می‌توان دقت و صحت پیش‌بینی را افزایش داد.

مراجع

Bosq D. (2000), *Linear processes in function spaces*, Springer , New York.

Didericksen D., Kokoszka, p. , and Zhang, X. (2011), Empirical properties of forecast with the functional autoregressive model, *Springer* .

Hall P. and Hossininasab M. (2006), on properties of functional principal components analysis, *Jornal of the Royal Statistical Society.* 61, 109 - 126.

Hormann S., Kokoszka, p. (2010), Weakly dependent functional data, *The Annalys of statistics*. 38(3): 1845-1884 .

Hyndman R., Shang H. (2009), Forecasting functional time series , *Jornal of Korean Statistical Society* . 199-211

Hyndman R., Ullah M. (2007), Robust forecasting of mortality and fertility rate: a functional data approach. *Computational Statistics and Data Analysis*, 51, 4942-4956.

Ramsay J., Silverman B. (1997), *Functional Data Analysis*, Springer .