

حاشا لله  
الرحمن الرحيم



دانشکده علوم ریاضی

پایان نامه کارشناسی ارشد آمار

# مروری بر روش‌های انتخاب متغیر در مدل‌های رگرسیون با خطاهای سری زمانی

نگارنده: سعید شیرزایی

استاد راهنما

دکتر محمد آرشی

استاد مشاور

دکتر مینا نوروزی راد

شهریور ۱۳۹۸

شماره:

تاریخ:

باسمه تعالی



دانشگاه گیلان

مدیریت تحصیلات تکمیلی

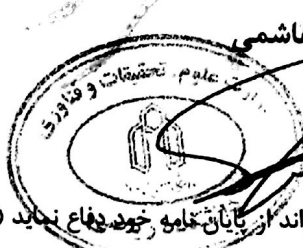
فرم شماره (۳) صورتجلسه نهایی دفاع از پایان نامه دوره کارشناسی ارشد

با نام و یاد خداوند متعال، ارزیابی جلسه دفاع از پایان نامه کارشناسی ارشد آقای سعید شیرزایی با شماره دانشجویی ۹۶۰۹۳۵۴ رشته آمار ریاضی گرایش - تحت عنوان: مروری بر روش های انتخاب متغیر در مدل های رگرسیونی با خطاهای سری زمانی که در تاریخ ۱۳۹۸/۰۶/۱۲ با حضور هیأت محترم داوران در دانشگاه صنعتی شاهرود برگزار گردید به شرح ذیل اعلام می گردد:

قبول (با درجه:  خلی خوب)  مردود

نوع تحقیق:  نظری  عملی

عضو هیأت داوران	نام و نام خانوادگی	مرتبه علمی	امضاء
۱- استاد راهنمای اول	دکتر محمد آرشی	دانشیار	
۲- استاد راهنمای دوم	-----	-----	-----
۳- استاد مشاور	دکتر مینا نوروزی راد	استادیار	
۴- نماینده تحصیلات تکمیلی	دکتر احمد نزاکتی	دانشیار	
۵- استاد ممتحن اول	دکتر داود شاهسونی	دانشیار	
۶- استاد ممتحن دوم	دکتر محمدرضا ربیعی	استادیار	



نام و نام خانوادگی رئیس دانشکده: دکتر ابراهیم هاشمی

تاریخ و امضاء و مهر دانشکده:

تبصره: در صورتی که کسی مردود شود حداکثر یکبار دیگر (در مدت مجاز تحصیل) می تواند از پایان نامه خود دفاع نماید (دفاع مجدد نباید زودتر از ۴ ماه برگزار شود).

تقدیم خالصانہ و خاضعانہ بہ:

پدرم بہ پاس سال ہا حمایت ہا بی درینغش  
مادرم بہ پاس سال ہا صبوری ہا بی انتہایش

# سپاس‌گزاری

منت‌خدای را عزوجل که طاعتش موجب قربت است و به شکر اندرش فرزند نعمت.

به مصداق لم یشکر المخلوق، لم یشکر الخالق،  
سپاس بی‌کران نثار پدر و مادرم، که ناتوان شدند تا من توانا شوم، مویشیان سپید شد تا من  
روسفید شوم و عاشقانه سوختند تا گرمابخش وجودم و روشن‌گر راهم باشند. اگر پدرم و  
حمایت‌های بی‌دریغش نبود، این سطرها هیچ‌گاه نگاشته نمی‌شد.  
سپاس‌گزاری و قدردانی خالصانه‌ام نثار اساتیدم جناب آقای دکتر محمد آرشی و سرکار خانم  
دکتر مینا نوروزی‌راد که با حمایت‌های بی‌دریغشان تاکنون هیچ‌گاه تنه‌ایم نگذاشته‌اند و با  
صبر و تلاش فراوان ایشان توانستم پایان‌نامه کارشناسی ارشد را گردآوری نمایم.  
از اساتید محترم، دکتر احمد نزاقتی، دکتر داود شاهسونی، دکتر حسین باغی‌شینی، دکتر  
محمد رضا ربیعی که در دوران تحصیلی‌ام صادقانه و عاشقانه تلاش نمودند تا برایم نقالی  
دانسته‌ها را به نقادی اندیشه‌ها تبدیل کنند، قدردانی می‌کنم.  
از مدیر گروه دلسوز و پرتلاش گروه آمار دانشگاه صنعتی شاهرود، جناب آقای دکتر حسین  
باغی‌شینی به دلیل خدمات، راهنمایی‌ها و زحماتشان سپاس‌گزارم.  
از اساتید داور محترم، جناب آقای دکتر داود شاهسونی و دکتر محمد رضا ربیعی که با حضور  
دلگرمشان تصحیح و داوری این پایان‌نامه را بر عهده گرفته‌اند سپاس‌گزارم و دوباره بارها و بارها  
از همدلی و همراهی و همگامی **خانواده مهربانم** که همواره با قلبی آکنده از عشق و معرفت  
محیطی سرشار از سلامت، امنیت، آرامش و آسایش برایم فراهم آورده‌اند و مرا در رسیدن به  
اهدافم یاری رسانده‌اند، بی‌نهایت سپاس‌گزارم.

سعید شیرزایی  
شهریور ۱۳۹۸

## تعهد نامه

اینجانب سعید شیرزایی دانشجوی کارشناسی ارشد رشته آمار دانشکده علوم ریاضی دانشگاه شاهرود، نویسنده پایان نامه با عنوان **مروری بر روش‌های انتخاب متغیر در مدل‌های رگرسیون با خطاهای سری زمانی**، تحت راهنمایی محمد آرشی متعهد می‌شوم:

- تحقیقات در این پایان نامه توسط اینجانب انجام شده است و از صحت و اصالت برخوردار است.
- در استفاده از نتایج پژوهش‌های دیگر پژوهش‌گران، به مرجع مورد استفاده استناد شده است.
- مطالب این پایان نامه، تا کنون توسط خود، یا فرد دیگری برای دریافت هیچ نوع مدرک یا امتیازی در هیچ‌جا ارایه نشده است.
- حقوق معنوی این اثر، به دانشگاه صنعتی شاهرود تعلق دارد، و مقالات مستخرج با نام ”دانشگاه صنعتی شاهرود“ یا ”Shahrood University of Technology“ به چاپ خواهد رسید.
- حقوق معنوی تمام افرادی که در به‌دست آوردن نتایج اصلی پایان نامه تاثیرگذار بوده‌اند، در مقالات مستخرج از پایان نامه رعایت می‌گردد.
- در تمام مراحل انجام این پایان نامه، در مواردی که از موجود زنده (یا بافت‌های آنها) استفاده شده است، ضوابط و اصول اخلاقی رعایت شده است.
- در تمام مراحل انجام این پایان نامه، در مواردی که به حوزه اطلاعات شخصی افراد دسترسی یافته (یا استفاده شده است)، اصل رازداری و اصول اخلاق انسانی رعایت شده است.

سعید شیرزایی

شهریور ۱۳۹۸

### مالکیت نتایج و حق نشر

- تمام حقوق معنوی این اثر و محصولات آن (مقالات مستخرج، کتاب، برنامه‌های رایانه‌ای، نرم‌افزارها و تجهیزات ساخته شده) متعلق به دانشگاه صنعتی شاهرود می‌باشد. این مطلب باید به نحو مقتضی، در تولیدات علمی مربوطه ذکر شود.
- استفاده از اطلاعات و نتایج موجود در این پایان نامه بدون ذکر منبع مجاز نمی‌باشد.

## چکیده

مدل‌بندی سری‌های زمانی مالی از اهمیت ویژه‌ای برخوردار است. در این راستا، مدل‌های آماری زیادی طراحی شده است که از آن جمله می‌توان به مدل‌های میانگین متحرک (MA)، اتورگرسیو (AR) و ترکیب این دو (ARMA) برای سری‌های زمانی ایستا اشاره کرد. اما این‌ها برای مدل‌بندی واریانس، تغییرات و ناپایداری‌های ناشی از نمونه‌های موجود در مدیریت ریسک، مدیریت اوراق بهادار، بازار بورس، تخصیص دارایی‌ها و موارد گوناگون دیگر از این دسته، کارا و مناسب نبوده و در عوض از سری‌های زمانی واریانس ناهمگن شرطی اتورگرسیو (ARCH) و نوع تعمیم‌یافته آن (GARCH) استفاده می‌شود. از طرفی، چنانچه در بحث مدل‌بندی با متغیرهای ورودی (توضیحی) و خروجی (پاسخ) سرو کار داشته باشیم شایسته است که از مدل رگرسیون با خطاهای سری زمانی استفاده کنیم. در این پایان‌نامه، تلفیق روش‌های جریمه‌شده با تاکید بر لاسوی سازوار، در انتخاب متغیرهای بااهمیت ورودی در مدل‌های رگرسیونی با خطای سری زمانی ARMA-GARCH را مورد بررسی قرار داده و با استفاده از یک سری مطالعات عددی کارایی این مدل‌ها را بررسی می‌کنیم.

**کلمات کلیدی:** انتخاب متغیر، برآوردگر لاسو سازوار، رگرسیون سری زمانی، مدل  
ARMA-GARCH.

# فهرست مطالب

۱	تعاریف و مقدمات	۱
۱	مقدمه	۱.۱
۲	سری زمانی	۲.۱
۲	روند	۱.۲.۱
۳	تغییرات فصلی	۲.۲.۱
۳	تغییرات دوره‌ای	۳.۲.۱
۴	تغییرات نامنظم	۴.۲.۱
۶	مانایی	۳.۱
۷	تابع خودهمبستگی (ACF)	۱.۳.۱
۹	مدل‌های سری‌های زمانی	۴.۱
۱۰	مدل میانگین متحرک	۱.۴.۱
۱۱	مدل اتورگرسیو	۲.۴.۱
۱۳	مدل اتورگرسیو میانگین متحرک	۳.۴.۱
۱۷	مدل‌های سری زمانی برای اندازه‌گیری تغییرات	۲
۱۷	مفاهیم مالی و اقتصادی	۱.۲
۲۳	مدل‌های واریانس ناهمگن شرطی اتورگرسیو	۱.۱.۲
۲۶	مدل‌سازی و برآورد پارامترها	۲.۱.۲
۲۷	مدل ARCH مرتبه اول	۳.۱.۲
۳۰	مدل GARCH	۴.۱.۲
۳۲	مدل GARCH(1, 1)	۵.۱.۲
۳۳	مدل‌های ARMA-GARCH	۲.۲
۳۴	انتخاب پارامترهای مدل‌های پیش‌بینی	۳.۲
۴۵	روش‌های انتخاب متغیر در رگرسیون سری زمانی	۳
۴۵	مدل‌های خطی	۱.۳



۴۷	.....	برآورد ضرایب مدل	۱.۱.۳
۴۹	.....	انتخاب متغیرها	۲.۳
۵۱	.....	لاسو	۱.۲.۳
۵۲	.....	لاسو سازوار	۲.۲.۳
۵۳	.....	الاستیکنت	۳.۲.۳
۵۴	.....	انتخاب پارامترهای جریمه	۴.۲.۳
۵۷	.....	پیاپی سازی های کامپیوتری	۳.۳
۵۷	.....	شبیه سازی	۴.۳
۶۹		<b>رگرسیون جریمه شده با خطاهای سری زمانی ARMA-GARCH</b>	<b>۴</b>
۶۹	.....	مقدمه	۱.۴
۷۰	.....	برآوردگر لاسو سازوار برای مدل های رگرسیون با خطاهای ARMA	۲.۴
		الگوریتم های ALASSO جدید برای خطای مدل های ARMA و ARMA	۱.۲.۴
۷۱	.....	GARCH	
۷۵	.....	مطالعات عددی	۳.۴
۷۵	.....	مطالعه شبیه سازی	۱.۳.۴
۷۹	.....	تحلیل مثال واقعی	۴.۴
۸۶	.....	نتیجه گیری و آینده تحقیق	۵.۴
۸۸		<b>آ</b>	<b>گزیده های از برنامه های رایانه ای</b>
۹۵			<b>مراجع</b>

# فصل ۱

## تعاریف و مقدمات

### ۱.۱ مقدمه

سری‌های زمانی نقش بسیار مهمی را در بسیاری از حوزه‌های علوم و مهندسی همچون اقتصاد، نیروی برق، محیط زیست و اکولوژی ایفا می‌کنند. به‌عنوان مثال، می‌توان به موارد زیر اشاره کرد:

- کشاورزی: مقدار محصول و قیمت‌های سالیانه غلات
- بازرگانی و اقتصاد: قیمت‌های موجود در پایان روز، نرخ‌های بهره هفتگی، شاخص قیمت‌های ماهیانه، فروش سه‌ماهه و درآمدهای سالیانه
- مهندسی: صدا، علایم الکتریکی و ولتاژ
- ژئوفیزیک: تلاطم‌هایی نظیر موج‌های اقیانوس و اغتشاش زمین در ناحیه‌ای خاص
- هواشناسی: سرعت باد در ساعت، درجه حرارت روزانه و میزان باران سالیانه
- علوم اجتماعی: میزان زاد و ولد سالیانه، نرخ‌های مرگ و میر، تصادفات و جنایت‌های گوناگون

## ۲.۱ سری زمانی

مجموعه‌ای از مشاهدات که بر حسب زمان مرتب شده باشند را سری زمانی می‌گویند که معمولاً به صورت  $\{x_t; t \in T\}$  نشان داده می‌شود که  $t = 1, 2, \dots, T$  بیان‌کننده زمانی است که مشاهده  $x$  در آن اندازه‌گیری شده و  $T$  طول سری زمانی یا تعداد کل مشاهدات است. سری زمانی معمولاً به صورت گسسته و پیوسته بررسی می‌شود، اگر مشاهدات به‌طور پیوسته بر حسب زمان در نظر گرفته شوند، سری زمانی را پیوسته و اگر مشاهدات به‌طور منظم در فاصله‌های مساوی ثبت شوند، یک سری گسسته به‌دست می‌آید. همان‌گونه که در تحلیل‌های آماری هر نمونه  $x_t$  به عنوان نماینده‌ای از جامعه آماری  $X_t$  فرض می‌شود، در داده‌های سری زمانی نیز هر مشاهده از سری  $x_t$  به عنوان مثالی از متغیر تصادفی  $X_t$  است که این رشته محدود  $X_t$  در طی زمان مرتب شده، قسمتی از یک رشته فرضی نامحدود است که آن را فرآیند تصادفی می‌نامیم (معروف‌خانی، ۱۳۷۷، ص. ۷۱).

معمولاً می‌توان الگوی مدل تغییرات یک سری زمانی را به چهار مولفه تفکیک کرد:

(۱) روند<sup>۱</sup>،

(۲) تغییرات فصلی<sup>۲</sup>،

(۳) تغییرات دوره‌ای<sup>۳</sup>،

(۴) تغییرات نامنظم<sup>۴</sup>.

به جز تغییرات نامنظم که قابل شناسایی نمی‌باشند، بقیه موارد قابل شناسایی‌اند. در ادامه، تعریف ساده‌ای از این مولفه‌ها از **مونت‌گمری و همکاران (۲۰۱۱)** آورده می‌شود.

### ۱.۲.۱ روند

تغییرات بلندمدت در میانگین سری زمانی، به عبارت دیگر، مسیر طبیعی سری زمانی در بلندمدت را روند می‌گویند. از مطالعه داده‌ها در یک دوره طولانی می‌توان یک ایده کلی نسبت به رفتار پدیده مورد بررسی به‌دست آورد که در پیش‌بینی آینده به کار گرفته می‌شود. مثلاً اگر سری زمانی، یک روند را در جهتی نشان دهد، با این فرض که این فرآیند در آینده نزدیک نیز به همین شکل ادامه پیدا می‌کند، مقادیر آینده را می‌توان پیش‌بینی کرد. روند می‌تواند خطی، کاهشی، افزایشی و یا نمایی باشد. تعداد کلاس‌هایی که در ده سال آینده در کشور لازم است، مثالی از مولفه روند است.

<sup>1</sup>Trend

<sup>2</sup>Seasonal

<sup>3</sup>Cyclic

<sup>4</sup>Irregular

اشکال تعریف روند در تعبیر «بلندمدت» است، برای مثال، گاهی اوقات، متغیرهای مربوط به آب و هوا، تغییرات دوره‌ای را در یک دوره زمانی بسیار طولانی مانند ۵ سال نشان می‌دهند. اگر دقیقاً، داده‌های مربوط به ۲۰ سال را در اختیار داشته باشیم، این نوسانات بلندمدت به صورت یک روند ظاهر می‌شود اما اگر چندین صد سال داده در دسترس باشند، نوسانات بلندمدت محو خواهند شد. با وجود این، در کوتاه‌مدت در نظر گرفتن این قبیل نوسانات بلندمدت به عنوان روند، هنوز دارای معنی می‌باشد. بنابراین، وقتی از یک روند صحبت به میان می‌آوریم، باید تعداد مشاهدات موجود را به حساب آورده و از «بلندمدت» ارزیابی ذهنی به عمل آوریم.

### ۲.۲.۱ تغییرات فصلی

تغییراتی که در دوره‌های تناوبی کوتاه پیش می‌آید را تغییرات فصلی می‌گویند. این تغییرات مربوط به عواملی است که به طریقی منظم و چرخه‌ای روی یک دوره کمتر از یک سال عمل می‌کند. اگر مشاهدات سری زمانی به صورت فصلی، ماهانه، هفتگی و یا روزانه ثبت شود، تغییرات فصلی در سری زمانی وجود دارد. مانند فروش بستنی در طول یک سال، که در تابستان بیشترین مقدار خود را دارد، که در بررسی این گونه (فروش بستنی) در طی سال‌های مختلف اگر اطلاعات به صورت ماهانه ثبت شود، مشاهده می‌کنیم که در تابستان بیشترین فروش و در زمستان کمترین فروش وجود داشته است. به عنوان مثالی دیگر، تامین میزان سوخت که در برخی از ماه‌ها، زیاد و در برخی ماه‌های دیگر، کم است.

تجزیه و تحلیل سری‌های زمانی که تغییرات را نشان می‌دهند به این بستگی دارد که بخواهیم (الف) مولفه‌های فصلی را اندازه بگیریم و یا (ب) آنها را حذف کنیم.

برای سری‌هایی که روند کمی را نشان می‌دهند، معمولاً مناسب است که برای هر ماه (یا یک فصل یا دوره چهار هفته‌ای) میانگین را محاسبه کرده و آن را به عنوان یک تفاضل یا یک نسبت، با رقم متوسط کلی مقایسه نماییم. یک اثر فصلی را با تفاضلی کردن می‌توان حذف کرد. (باکس و جنکینز، ۱۹۷۰)

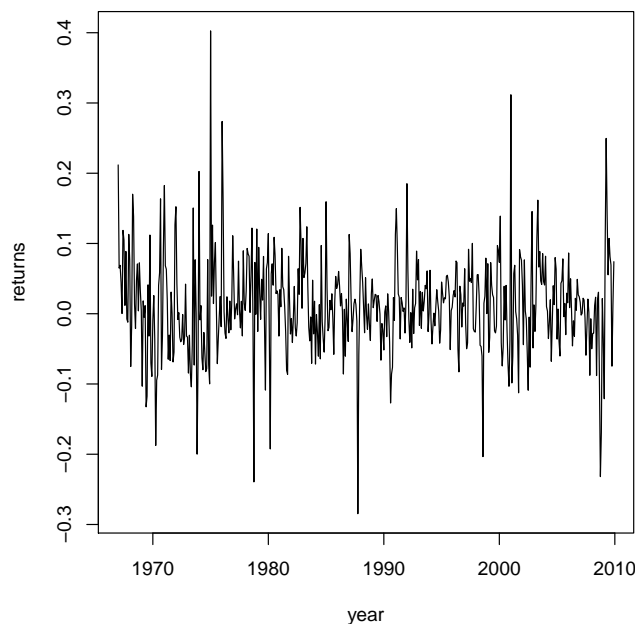
### ۳.۲.۱ تغییرات دوره‌ای

حرکت نوسانی در یک سری زمانی با دوره نوسان بیشتر از یک سال را تغییرات دوره‌ای می‌نامند. این تغییرات در سری‌های زمانی به واسطه نوساناتی است که بعد از یک دوره بیشتر از یک سال رخ می‌دهد. اگرچه شناسایی تغییرات دوره‌ای و تشخیص آن از روند مقداری مشکل است ولی این تغییرات نیز قابل شناسایی‌اند. مثلاً هر ۱۳ سال تولید گندم در کشور انگلستان به یک ترتیب است (یعنی چرخه دارد)، یعنی هر ۱۳ سال کم یا زیاد شدن تولید گندم تکرار می‌شود.

## ۴.۲.۱ تغییرات نامنظم

بعد از آن که روند و تغییرات فصلی از مجموعه داده‌ها حذف شوند، سری باقی مانده‌ها حاصل می‌شود که ممکن است تصادفی باشد. در واقع این عوامل تصادفی باعث تغییرات در سری زمانی می‌شوند مانند سیل، زلزله، جنگ، خشکسالی یا آتش‌سوزی که در تولید یک کارخانه تاثیر می‌گذارد. از الگوهای میانگین متحرک<sup>۵</sup> یا فرآیندهای اتورگرسیو<sup>۶</sup> برای تجزیه و تحلیل برخی از داده‌های نامنظم استفاده می‌شود.

**مثال ۱.۲.۱.** مجموعه داده‌های مرکز تحقیقات امنیت قیمت‌ها<sup>۷</sup> (CRSP)، نمونه‌ای از داده‌های بازده ماهانه Desile10 است که تعداد مشاهدات آن ۵۱۶ است. این مجموعه کوچک، شامل ۱۰ درصد از سهام در بازار سرمایه‌گذاری NYSE/AMEX/NASDAQ است. شکل ۱.۱ نمودار سری زمانی این داده‌ها را نمایش می‌دهد که تغییر کم نوسانات به خوبی در آن قابل مشاهده است.



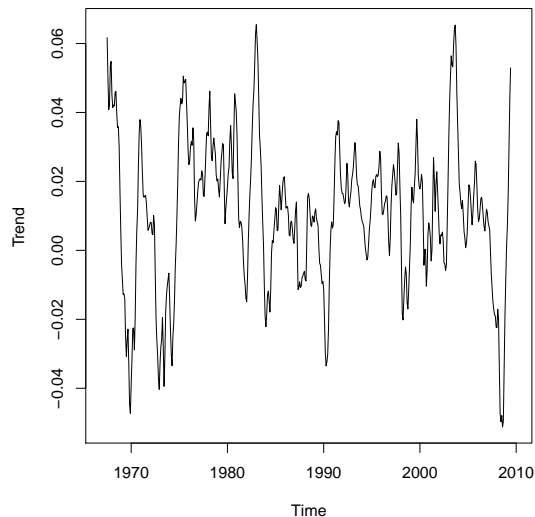
شکل ۱.۱: نمایش سری زمانی بازده ماهانه Desile10

<sup>5</sup>Moving Average

<sup>6</sup>Autoregressive

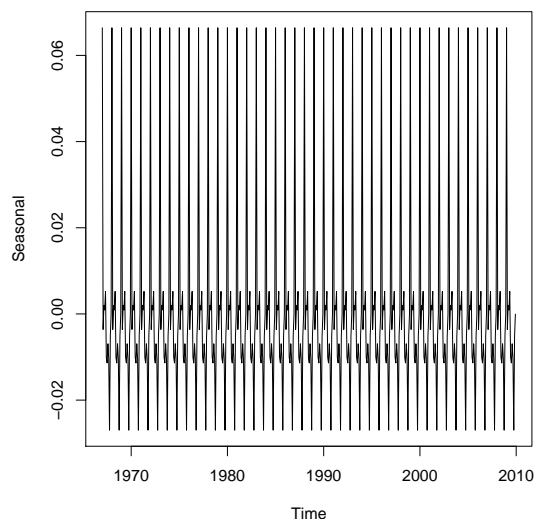
<sup>7</sup>The Center for Research in Security Prices

نمودار روند مربوط به این داده‌ها در شکل ۲.۱ مشاهده می‌شود.



شکل ۲.۱: نمایش روند داده‌های Desile10

تغییرات فصلی این داده‌ها در شکل ۳.۱ نمایش داده شده است.



شکل ۳.۱: نمایش تغییرات فصلی داده‌های Desile10

## ۳.۱ مانایی

مدل سازی اقتصادسنجی سری زمانی مبتنی بر فرض مانایی<sup>۸</sup> متغیرهای سری زمانی است. بر اساس این فرض، میانگین، واریانس و کواریانس داده‌ها در طول زمان ثابت بوده و مستقل از زمان است. اما بررسی‌هایی که از سال ۱۹۹۰ به بعد انجام شده، نشان داده که این فرض در مورد بسیاری از متغیرهای سری زمانی اقتصاد کلان نادرست بوده و بیشتر این متغیرها وابسته به زمان و نامانا<sup>۹</sup> است. مطالعات نشان داده که در صورت عدم تحقق فرض مانایی، یعنی نامانا بودن متغیرها در سری‌های زمانی استفاده از آماره‌های  $t$  و  $F$  گمراه‌کننده بوده و امکان این که نتایج به دست آمده تنها از یک رگرسیون دروغین بوده و هیچ‌گونه رابطه اقتصادی واقعی و تعادلی وجود نداشته باشد، موجود است. بنابراین بررسی مانایی متغیرها الزامی است (بیدرام، ۱۳۸۱).

قبل از برآورد تابع رگرسیون، نخستین گام، آزمون مانایی متغیرهاست. سری‌های زمانی از لحاظ خواص احتمالی به دو دسته، مانا و نامانا و سری‌های زمانی مانا نیز به دو دسته مانای اکید<sup>۱۰</sup> و مانای ضعیف<sup>۱۱</sup> تقسیم می‌شوند. اگر توزیع مشترک آن در طول زمان بدون تغییر باقی بماند یک سری زمانی اکیدا ماناست. اما در عمل چون تعیین توزیع مشترک متغیر تصادفی سری زمانی سخت است، مانایی ضعیف به کار برده می‌شود. مفهوم مانایی این است که آیا یک متغیر در طی زمان یک رفتار باثبات و پایدار از خود نشان می‌دهد یا نه.

اگر در سری زمانی  $\{x_t; t \in T\}$  برای هر زمان  $t = 1, 2, \dots, T$  و هر تاخیر  $k$ ، توزیع  $x_1, x_2, \dots, x_T, \dots$  همانند توزیع  $x_{1+k}, x_{2+k}, \dots, x_{T+k}$  باشد، سری مانا گفته می‌شود. اگر گشتاورهای توزیعی تا مرتبه معینی نسبت به زمان ثابت باشد، فرآیند را مانای ضعیف در آن مرتبه می‌نامند. مثلا اگر در یک سری زمانی، میانگین نسبت به زمان ثابت باشد، مانایی ضعیف از مرتبه اول برقرار است و اگر میانگین و واریانس نسبت به زمان ثابت باشد مانایی ضعیف از مرتبه دوم برقرار است. در سری‌های مانای اکید تمام گشتاورها نسبت به زمان ثابت است. مانایی ضعیف از مرتبه دوم در کنار فرض نرمال بودن سری، مانایی اکید را نتیجه می‌دهد. در مجموع، مفهوم مانایی این است که قوانین حاکم بر فرآیند با زمان تغییر نکند. سری زمانی  $X_t$ ، مانا گفته می‌شود اگر ویژگی‌های زیر را دارا باشد:

۱. میانگین آن برای همه نقاط زمانی، ثابت بماند. اگر میانگین را با  $\mu$  نشان دهیم، آن‌گاه

$$\forall t \quad E(X_t) = \mu$$

<sup>8</sup>Stationarity

<sup>9</sup>Non stationarity

<sup>10</sup>Strict stationary

<sup>11</sup>Weak stationary

۲. واریانس آن برای همه نقاط در طول زمان ثابت باشد. به عبارت دیگر،

$$\forall t \quad V(X_t) = E(X_t - \mu)^2 = \sigma^2$$

۳. کواریانس هر دو مقدار از سری زمانی، بستگی به فاصله آنها دارد نه به زمان آنها، پس اگر  $\gamma_k$ ، کواریانس بین دو مقدار از سری زمانی، با فاصله زمانی  $K$  باشد، در این صورت:

$$\begin{aligned} \text{Cov}(X_t, X_{t-k}) &= E[(X_t - \mu)(X_{t-k} - \mu)] = \gamma_k \\ \text{Cov}(X_{t+j}, X_{t+j-k}) &= \text{Cov}(X_t, X_{t-k}) = \gamma_k \end{aligned}$$

نقض هر یک از سه شرط بیان شده در یک سری زمانی بدین معناست که سری زمانی ناماناست. بیشتر سری‌های زمانی در اقتصاد کلان، نامانا هستند. استفاده از روش‌های رایج اقتصادسنجی مثل روش کمترین توان‌های دوم معمولی<sup>۱۲</sup> (OLS) برای سری‌های زمانی نامانا، بسیاری از اوقات به تفسیر نادرست منجر می‌شود.

یک روش متداول و آسان برای تبدیل یک سری نامانا به یک سری مانا استفاده از روش تفاضل مرتبه اول است. مثلاً در سری زمانی  $X_t$ ، تفاضل مرتبه اول چنین تعریف می‌شود.

$$\nabla X_t = X_t - X_{t-1}$$

معمولاً تفاضل‌گیری (یک بار یا بیشتر)، یک سری نامانا را به سری مانا تبدیل می‌کند. یک سری زمانی نامانا که با  $d$  بار تفاضل‌گیری به سری مانا تبدیل شود، دارای درجه مانایی  $d$  است و با  $I(d)$  نمایش داده می‌شود. البته ناگفته نماند که با تفاضل‌گیری اطلاعات بلندمدت از دست می‌رود و ممکن است توجیه نظری برای متغیر ساخته شده وجود نداشته باشد. معمولاً اگر نامانایی در میانگین باشد، با یک بار تفاضل‌گیری سری موردنظر مانا خواهد شد و اگر نامانایی در واریانس باشد با دو بار تفاضل‌گیری مانا خواهد شد (کارتر و همکاران، ۲۰۰۷).

روش‌های مختلفی برای تشخیص مانایی وجود دارد که یکی از آنها استفاده از نمودار خودهمبستگی است. یکی از روش‌های تشخیص سری زمانی مانا و نامانا، استفاده از توابع خودهمبستگی<sup>۱۳</sup> است.

### ۱.۳.۱ تابع خودهمبستگی (ACF)

این تابع، توصیف جزئی از فرآیند را با هدف الگوسازی فراهم می‌کند و از تقسیم کواریانس بر واریانس سری زمانی به صورت

$$\rho_k = \frac{\text{Cov}(X_t, X_{t-k})}{V(X_t)} = \frac{\lambda_k}{\lambda_0} \quad (1.1)$$

<sup>12</sup>Ordinary least squares

<sup>13</sup>Autocorrelation function



به دست می آید که  $k$  تعداد وقفه آزمون و  $\rho_k$  ضریب همبستگی با  $k$  وقفه سری زمانی است که در واقع همبستگی بین مقادیر سری زمانی با  $k$  وقفه را نشان می دهد. برای یک سری زمانی مانا، تابع خودهمبستگی، تابعی از پارامتر  $k$  و متغیر تصادفی و مستقل از  $t$  است. چنانچه مقدار تابع خودهمبستگی نمونه  $\hat{\rho}_k$  را در مقابل وقفه  $k$  در نموداری رسم کرده و نمودار همبستگی نگار را به دست آوریم، نوع خاصی از سری زمانی که فرآیند تولید داده از نظر جمعی به بار آورده نتیجه می شود. از این رو اگر متغیر مورد نظر مانا باشد، با بزرگ شدن  $k$ ، ACF کوچک شده و به سمت صفر گرایش پیدا می کند. در یک فرآیند آماری کاملاً تصادفی و در حالی که وقفه سری بیشتر از صفر باشد، اتوکواریانس صفر می شود. بنابراین در یک فرآیند تصادفی مانا، مقدار ضریب همبستگی در وقفه صفر، یک می شود. برای آزمون اینکه آیا فقط یک مقدار خاص تابع خودهمبستگی نمونه، صفر است یا اینکه تمام مقادیر تابع برای  $k$  های بزرگتر از صفر، صفر است، از آزمون های متفاوتی استفاده می شود.

اگر فرض شود که  $X_t$  دارای توزیع نرمال استاندارد باشد، در این صورت ضرایب خودهمبستگی نیز تقریباً توزیع نرمال خواهند داشت

$$\hat{X}_s \sim \mathcal{N}\left(0, \frac{1}{n}\right)$$

$n$ ، حجم نمونه و  $\hat{X}_s$  بیان گر ضرایب خودهمبستگی است که از نمونه به دست می آید. این نتایج برای آزمون معنی داری ضرایب خود همبستگی و تشکیل فاصله اطمینان نیز به کار برده می شود. به عنوان مثال فاصله اطمینان ۹۵ درصدی برای ضرایب خودهمبستگی به صورت  $\left(-\frac{1.96}{\sqrt{n}}, +\frac{1.96}{\sqrt{n}}\right)$  است. اگر ضریب خودهمبستگی در این فاصله قرار بگیرد، فرضیه صفر مبنی بر اینکه مقدار واقعی ضریب خودهمبستگی با وقفه  $s$  برابر صفر است، رد می شود. همچنین می توان هر فرضیه مشترکی در مورد ضرایب خودهمبستگی را آزمون کرد. اگر فرضیه مشترک به صورت  $\rho_1 = \rho_2 = \dots = \rho_m = 0$  باشد، آن گاه می توان با استفاده از تابع  $Q$ ، رابطه (۲.۱) که توسط **باکس و پییرس (۱۹۷۰)** مطرح شده، آزمون کرد

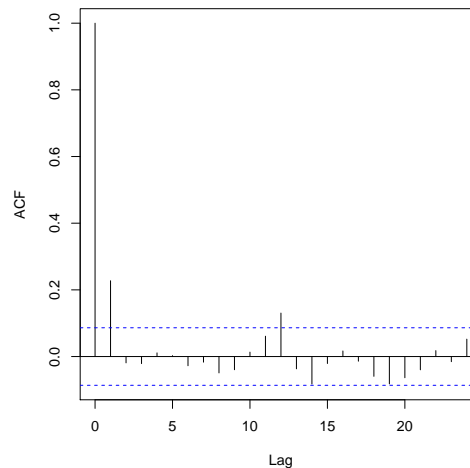
$$Q = T \sum_{k=1}^m \hat{\rho}_k^2 \quad (2.1)$$

که  $T$ ، حجم نمونه و  $m$ ، ضرایب خودهمبستگی است. ضرایب خودهمبستگی به این دلیل مجذور می شوند که علامت های مثبت و منفی یکدیگر را خنثی نکنند. از آنجا که مجذور متغیرهای نرمال استاندارد دارای توزیع  $\chi^2$  است، لذا  $Q$  دارای توزیع مجانبی  $\chi_m^2$  تحت فرضیه صفر (همه ضرایب خودهمبستگی، صفر هستند) است. به هر حال، آزمون باکس- پییرس برای نمونه های کوچک، ضعیف است و ممکن است منجر به نتایج نادرست شود. آماره تعدیل شده دیگری که معروف به تابع لیونگ- باکس (**لیونگ و باکس، ۱۹۷۸**) است، به صورت

$$Q = T(T+2) \sum_{k=1}^m \frac{\hat{\rho}_k^2}{T-k} \quad (3.1)$$

تعریف می‌شود. از رابطه (۳.۱) مشخص است که اگر حجم نمونه به سمت بی‌نهایت میل کند،  $T+2$  و  $T-K$  همدیگر را خنثی می‌کنند و  $Q^*$  معادل با آزمون باکس - پییرس می‌شود (سوری، ۱۳۸۹، ص. ۱۴۷).

مثال ۱.۳.۱. نمودار تابع خودهمبستگی مربوط به داده‌های Desile10 در مثال ۱.۲.۱ در شکل ۴.۱ نشان داده شده است.



شکل ۴.۱: نمودار تابع خودهمبستگی داده‌های Desile10

همان‌طور که شکل ۴.۱ نشان می‌دهد مجموعه داده Desile10 مانا است، زیرا نمودار تابع خودهمبستگی روند کاهشی دارد.

## ۴.۱ مدل‌های سری‌های زمانی

هدف اصلی در تجزیه و تحلیل سری زمانی در مورد یک پدیده، ایجاد یک مدل آماری برای داده‌های وابسته به زمان بر اساس اطلاعات گذشته آن پدیده است. با این کار امکان پیش‌بینی در مورد آینده پدیده مورد بحث میسر می‌شود. به بیان دیگر، تحلیل سری زمانی، ایجاد مدلی گذشته‌نگر است تا امکان تصمیمات آینده‌نگر را فراهم سازد. الگوهای سری زمانی بر خلاف الگوهای اقتصادسنجی که از اطلاعات مربوط به نظریه‌های اقتصادی و داده‌های آماری سود می‌جویند، تنها از اطلاعات مربوط به داده‌های آماری استفاده می‌کنند و توجهی به مبانی نظری تئوری‌های اقتصادی ندارند. این الگوها که تنها مقادیر فعلی یک متغیر را به مقادیر گذشته آن نسبت می‌دهند، الگوهای سری زمانی یک متغیره<sup>۱۴</sup> نامیده می‌شوند. این الگوها، عبارتند از: فرآیندهای اتورگرسیو، میانگین متحرک، اتورگرسیو میانگین متحرک<sup>۱۵</sup> (نوفرستی، ۱۳۷۸).

<sup>14</sup>Univariate time-series models

<sup>15</sup>Autoregressive and moving average

## ۱.۴.۱ مدل میانگین متحرک

برای پیش‌بینی سری زمانی ایستا که روند از آن خارج شده، می‌توان از مدل میانگین متحرک استفاده کرد. در این مدل، سری زمانی  $X_t$  به صورت

$$X_t = C_0 + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q} \quad (۴.۱)$$

نوشته می‌شود که در آن  $a_t$  همان خطای تصادفی با میانگین صفر و واریانس ثابت،  $C_0$ ، امیدریاضی  $X_t$  (اغلب برابر صفر در نظر گرفته می‌شود) است. یک فرآیند میانگین متحرک از مرتبه  $q$  را با نماد اختصاری  $MA(q)$  نمایش می‌دهند که با استفاده از عملگر پسر و به صورت

$$X_t = C_0 + (1 - \theta_1 \beta - \theta_2 \beta^2 - \dots - \theta_q \beta^q) a_t = C_0 + \theta(\beta) a_t \quad (۵.۱)$$

نوشته می‌شود. این فرآیند همواره مانا است و برای اینکه معکوس‌پذیر باشد، با در نظر گرفتن بردارهای  $q$  بعدی  $\theta = (\theta_1, \dots, \theta_q)^\top$  و  $o = (o_1, \dots, o_q)^\top$  باید قدر مطلق ریشه‌های  $\theta\beta = o$  بزرگتر از یک باشد. از فرآیند میانگین متحرک در بسیاری از زمینه‌ها به‌ویژه در اقتصادسنجی استفاده می‌شود. این فرآیند در توصیف پدیده‌ای به کار می‌رود که در آن پیشامدها یک اثر آنی را به وجود می‌آورند که فقط برای مدتی کوتاه باقی می‌ماند. اصطلاح میانگین متحرک از این واقعیت ناشی می‌شود که  $X_t$  با استفاده از وزن‌های  $\theta_1, \theta_2, \dots, \theta_q$  برای متغیرهای  $a_{t-1}, a_{t-2}, \dots, a_{t-q}$  به دست می‌آید و اگر زمان را به اندازه یک واحد اضافه کرده و از این وزن‌ها برای تعیین  $a_{t+1}, a_{t+2}, \dots, a_{t+q}$  استفاده کرد،  $X_{t+1}$  حاصل می‌شود (ذوالفقاری و صادقی، ۱۳۹۰).

یکی از روش‌های شناسایی مرتبه مدل  $MA(q)$  استفاده از تابع خودهمبستگی است. برای یک سری زمانی با تابع خودهمبستگی  $\rho_l$  اگر  $\rho_q \neq 0$  اما  $\rho_l = 0$  برای  $l > q$  سپس مدل  $MA(q)$  را به دست می‌آوریم (تسای، ۲۰۱۳).

مثال ۱.۴.۱. با استفاده از داده‌های ۱.۲.۱ بازده ماهانه شاخص CRSP که از ژانویه ۱۹۲۶ تا دسامبر ۲۰۰۸ است. نمودار سری زمانی و تابع خودهمبستگی را در شکل ۵.۱ نشان می‌دهیم.

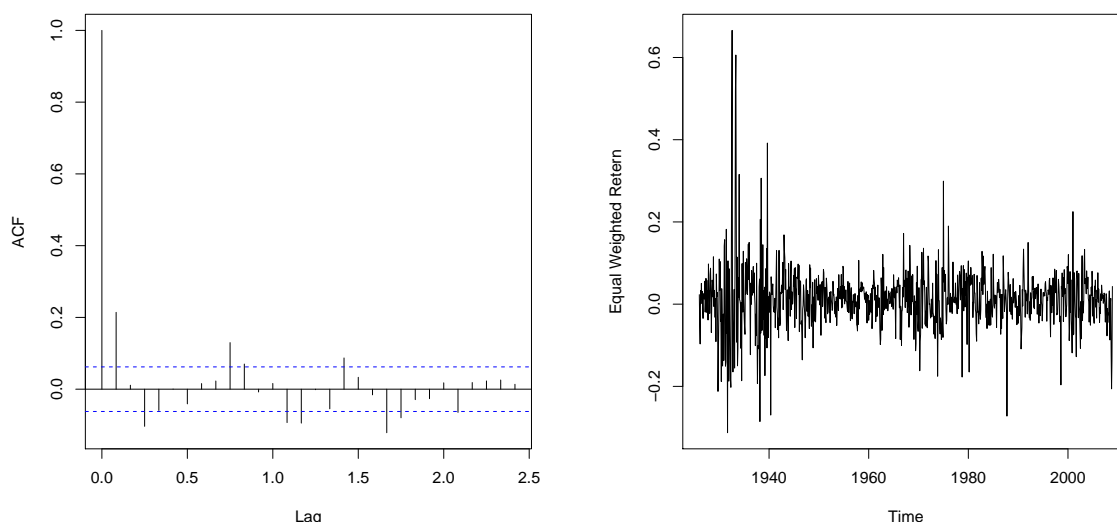
در شکل ۵.۱ چون تابع خودهمبستگی در لگ‌های اول و سوم و نهم معنی‌دار است بنابراین یک مدل  $MA(9)$  نتیجه می‌شود که فرمول آن به صورت

$$X_t = C_0 + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \theta_3 a_{t-3} - \dots - \theta_9 a_{t-9} \quad (۶.۱)$$

است که مقادیر  $\theta_1, \theta_2, \theta_3$  عبارت‌اند از:

$$X_t = 0.122 + a_t - 0.1909 a_{t-1} + 0.1199 a_{t-2} - 0.1227 a_{t-3}$$

که از آزمون لیونگ - باکس، با آماره آزمون  $17/604$  و درجه آزادی ۱۲ و  $p\text{-value} = 0.1283$ ،  $p\text{-value}$  تعدیل شده برابر است با  $0.4010828$  و چون لگ اول، سوم و نهم تابع خودهمبستگی از  $p\text{-value}$  تعدیل شده کمتر است پس معنی‌دار هستند.



شکل ۵.۱: نمودار سری زمانی و تابع خودهمبستگی مدل میانگین متحرک از مرتبه نهم بازده ماهانه شاخص CRSP همگن وزنی که از ژانویه ۱۹۲۶ تا دسامبر ۲۰۰۸ است.

### ۲.۴.۱ مدل اتورگرسیو

اگر مقدارهای سری زمانی ایستا به صورتی باشد که به مقدارهای قبلی خود بستگی داشته باشد، از مدل اتورگرسیو استفاده می‌شود. در این حالت،  $p$  تعداد مشاهدات گذشته در نظر گرفته می‌شود که برای پیش‌بینی یک مقدار به کار برده می‌شود. بنابراین مدل اتورگرسیو برای سری زمانی ایستا  $X_t$  به صورت

$$X_t = \phi_0 + \phi_1 x_{t-1} + \phi_2 x_{t-2} + \dots + \phi_p x_{t-p} + a_t \quad (7.1)$$

تعریف می‌شود. در این مدل،  $\phi_1, \phi_2, \dots, \phi_p$  پارامترهای مدل،  $\phi_0$ ، ثابت مدل و  $a_t$ ، خطای مدل است. چنین مدلی به صورت  $AR(p)$  نمایش داده می‌شود و  $p$ ، مرتبه مدل است. با توجه به مسئله رگرسیون می‌بینیم که رابطه (۷.۱) در واقع یک الگوی رگرسیون چندگانه است. با این تفاوت که در اینجا  $X_t$ ، روی متغیرهای مستقل رگرسیون نشده، بلکه روی مقادیر گذشته آن یعنی  $x_{t-1}, x_{t-2}, \dots, x_{t-p}$  رگرسیون شده است. و به همین دلیل، فرآیند  $X_t$  را اتورگرسیو نامیده‌اند. این فرآیند همواره معکوس‌پذیر است، یعنی همواره بدون گذاشتن شرط روی  $\phi_i$  ها می‌توان  $a_t$  را به صورت ترکیب خطی وزنی از مشاهدات حال و گذشته فرآیند نوشت. مدل به صورت

$$a_t = X_t - \phi_1 x_{t-1} - \phi_2 x_{t-2} - \dots - \phi_p x_{t-p} = \phi_p(\beta) X_t$$

نیز بیان می‌شود که در آن

$$\phi_p(\beta) = 1 - \phi_1 \beta - \phi_2 \beta^2 - \dots - \phi_p \beta^p$$

که  $\beta$  عملگر و پسر و است. به طور مثال،

$$X_{t-1} = \beta X_t, X_{t-2} = \beta^2 X_t, \dots, X_{t-p} = \beta^p X_t \quad (۸.۱)$$

برای مانایی فرآیند باید ریشه‌های معادله  $\phi(\beta) = 0$  خارج از دایره واحد باشند. برای توضیح بیشتر، فرآیند  $AR(1)$  به صورت

$$X_t = \phi_1 X_{t-1} + a_t \quad (۹.۱)$$

تولید می‌شود. بنابراین

$$a_t = X_t - \phi_1 X_{t-1} \quad (۱۰.۱)$$

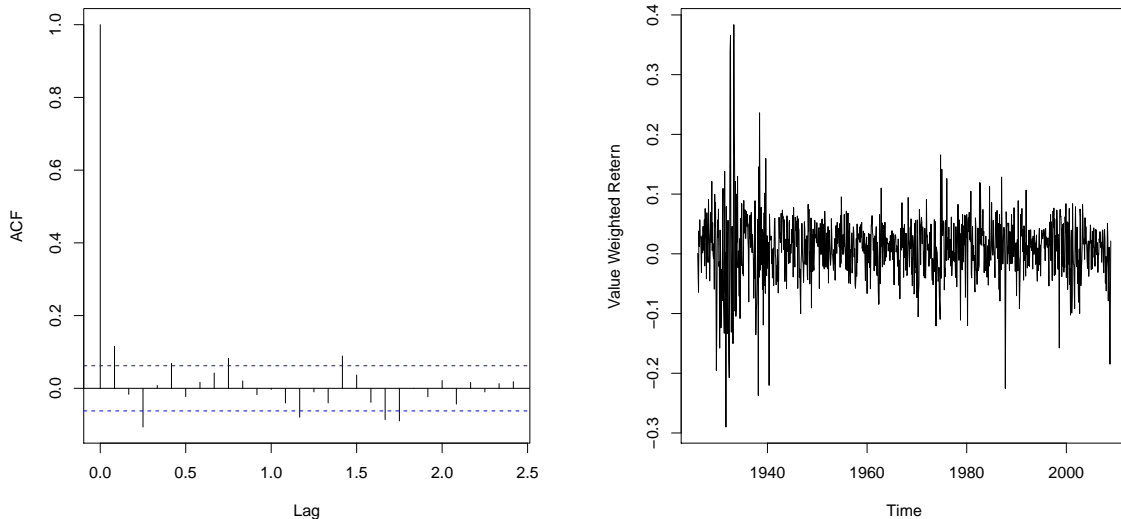
بر اساس تعریف عملگر پسر و در رابطه (۸.۱)، می‌توان رابطه (۱۰.۱) را به صورت

$$a_t = X_t - \phi_1(\beta X_t) = (1 - \phi_1 \beta) X_t \quad (۱۱.۱)$$

نوشت. بنابراین، با توجه به  $(1 - \phi_1 \beta) = 0$  نتیجه می‌شود

$$\beta = \frac{1}{\phi_1}.$$

مثال ۲.۴.۱. با توجه به داده‌های مثال ۱.۴.۱ برای یک مدل  $AR(3)$  برای سری بازده ماهانه همگن وزنی، نمودار سری زمانی و خودهمبستگی را در شکل ۶.۱ نشان داده شده است.



شکل ۶.۱: نمودار سری زمانی و تابع خودهمبستگی مدل اتورگرسیو از مرتبه ۳ بازده ماهانه شاخص CRSP همگن وزنی که از ژانویه ۱۹۲۶ تا دسامبر ۲۰۰۸ است.

با استفاده از شکل ۶.۱ معادله مدل میانگین متحرک از مرتبه ۳ به صورت

$$X_t = 0.088 + 0.114x_{t-1} - 0.106x_{t-3} + a_t$$

نوشته شده است که بر اساس آزمون لیونگ - باکس که در اینجا مقدار آماره آزمون برابر با  $۱۶/۸۳$  و  $p\text{-value} = ۰/۰۷۸$ ، از آن جایی که لگ‌های اول و دوم و سوم از مقدار  $p\text{-value} = ۰/۰۷۸$  کمتر هستند، پس معنی‌دار خواهند بود.

### ۳.۴.۱ مدل اتو رگرسیو میانگین متحرک

با توجه به خصوصیات هر یک از روش‌های اتو رگرسیو میانگین متحرک، امکان ترکیب این دو روش نیز وجود دارد. **باکس و جنکینز (۱۹۷۰)**، چنین مدلی را با نام اتو رگرسیو میانگین متحرک معرفی کردند که به صورت

$$X_t = \phi_0 + \sum_{i=1}^p \phi_i x_{t-i} + a_t - \sum_{i=1}^q \theta_i a_{t-i} \quad (۱۲.۱)$$

نوشته می‌شود. این مدل را با نماد  $ARMA(p, q)$  نشان می‌دهند که در آن  $p$ ، مرتبه مدل AR و  $q$ ، مرتبه مدل MA است. هم‌چنین این مدل را می‌توان به صورت

$$\phi_p(\beta)X_t = \phi_0 + \theta(\beta)a_t \quad (۱۳.۱)$$

که در آن

$$\begin{aligned} \phi_p(\beta) &= 1 - \phi_1\beta - \phi_2\beta^2 - \dots - \phi_p\beta^p \\ \theta(\beta) &= 1 - \theta_1\beta - \theta_2\beta^2 - \dots - \theta_q\beta^q \end{aligned}$$

نوشته. برای یافتن مقادیر مناسب  $p$  و  $q$  در مدل  $ARMA(p, q)$  می‌توان از رسم نمودار توابع خودهمبستگی نسبی<sup>۱۶</sup> برای  $p$  و رسم نمودار توابع خودهمبستگی<sup>۱۷</sup> برای تخمین  $q$  استفاده نمود. برخی از محققان نیز پیشنهاد کرده‌اند که از ملاک AIC برای تعیین مقادیر  $p$  و  $q$  استفاده شود. نکته حائز اهمیت در استفاده از مدل ARMA برای یک سری زمانی، ایستایی سری می‌باشد. این مدل تنها در صورتی که مشاهدات سری زمانی ایستا باشند، دارای کارایی مناسبی برای مدل نمودن سری است. به عبارت دیگر، چنان‌چه قوانین احتمالی فرآیند در طی انتقال آن در طول زمان تغییر کند، این تغییرات در مدل منعکس نمی‌شود و در نتیجه میزان کارایی سیستم در پیش‌بینی سری زمانی به میزان قابل توجهی کاهش می‌یابد. بنابراین قبل از برآزش یک مدل ARMA به داده‌های یک سری زمانی، همواره ایستایی مشاهدات مورد بررسی قرار می‌گیرند. از آن جایی که سری‌های زمانی مالی عمدتاً غیرایستا می‌باشند، معمولاً از تبدیل زیر به منظور ایستا نمودن مشاهدات استفاده می‌شود.

$$Y_t = \log X_t - \log X_{t-1} \quad (۱۴.۱)$$

<sup>۱۶</sup>Partial autocorrelation function

<sup>۱۷</sup>Autocorrelation function

در این تبدیل به سری جدید  $Y_t$ ، سری بازده درآمد یا سود اطلاق می‌شود. در سری  $Y_T$  میانگین مشاهدات نزدیک به صفر می‌باشد و همچنین واریانس آن‌ها ثابت است. بنابراین سری بازده، دارای خاصیت ایستایی ضعیف یا ایستایی مرتبه دوم می‌باشد.

### شناسایی مرتبه یک فرآیند اتورگرسیو میانگین متحرک

توابع خودهمبستگی و خودهمبستگی جزئی برای شناسایی دقیق فرآیند ARMA توانایی لازم را ندارند. برای این منظور رهیافتی جایگزین توسط **تسای و تیائو (۱۹۸۴)** به صورت تابع خودهمبستگی تعمیم‌یافته<sup>۱۸</sup> برآورد شده است که در بسته TSA در نرم‌افزار R در دسترس است. خروجی این تابع، یک تابلو از عنوان‌های O و X است. مقادیر AR در سطر و MA در ستون قرار گرفته است و مقدار بهینه فرآیند ARMA در این تابلو به صورت یک مثلث از مقادیر O در ضلع بالا و سمت چپ قابل شناسایی است.

**مثال ۳.۴.۱.** نمودار سری زمانی و مدل ARMA را با استفاده از داده‌های بازده ماهانه شاخص سهام شرکت  $M3$  از فوریه سال ۱۹۴۶ تا دسامبر ۲۰۰۸ که دارای ۷۵۵ مشاهده است در شکل ۷.۱ ترسیم شده است.

در ادامه، برای تشخیص مرتبه مدل ARMA بر روی داده‌ها از دستور eacf استفاده کرده که ملاحظه می‌شود مدل  $ARMA(0,0)$  است.

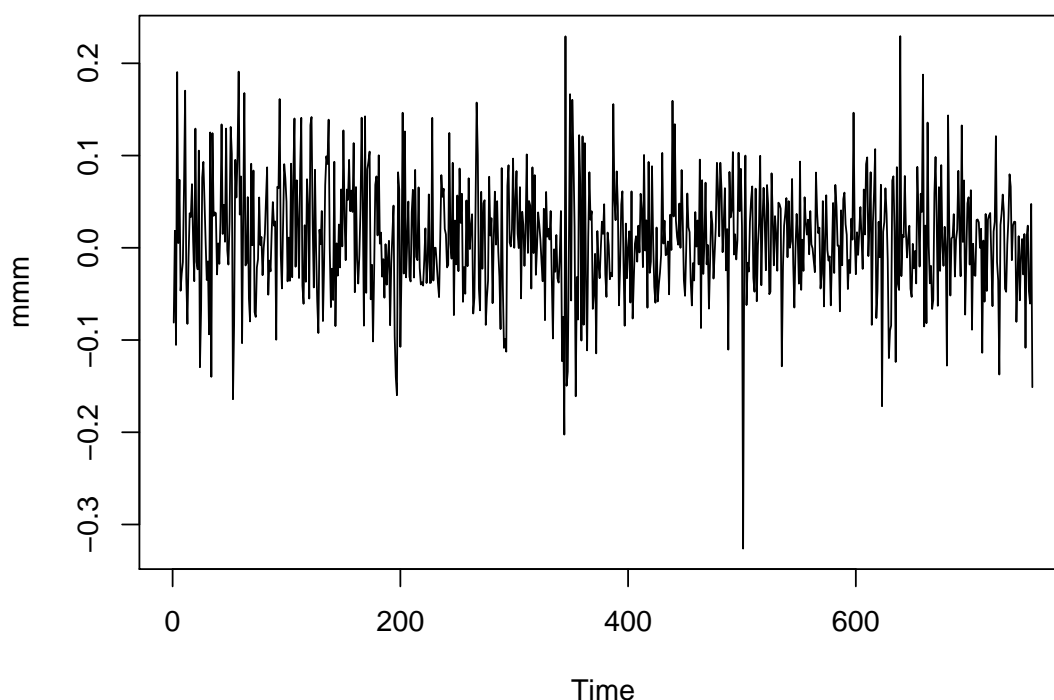
AR/MA

	0	1	2	3	4	5	6	7	8	9	10	11	12	13
0	o	o	x	o	o	x	o	o	o	x	o	x	o	o
1	x	o	x	o	o	x	o	o	o	o	o	x	o	o
2	x	x	x	o	o	x	o	o	o	o	o	o	o	o
3	x	x	x	o	o	o	o	o	o	o	o	o	o	o
4	x	o	x	o	o	o	o	o	o	o	o	o	o	o
5	x	x	x	o	x	o	o	o	o	o	o	o	o	o
6	x	x	x	x	x	o	o	o	o	o	o	o	o	o
7	x	x	x	x	o	x	o	o	o	o	o	o	o	o

### برآورد یک فرآیند اتورگرسیو میانگین متحرک

برای تخمین ARMA از روش‌های مختلفی مانند کمترین توان‌های دوم و ماکسیمم درست‌نمایی می‌توان استفاده کرد. که با نرم‌افزار R قابل انجام است. ولی باید توجه کرد که در یک مدل

<sup>18</sup>Extended autocorrelation function



شکل ۷.۱: نمایش سری زمانی بازده ماهانه شاخص سهام شرکت ۳M

$AR(p)$ ، مولفه خطا در پارامترها خطی و در یک فرایند  $MA(q)$ ، مولفه خطا در پارامترها غیرخطی است. بنابراین، برای آزمون مربوط به  $AR(p)$  می‌توان هم از روش برآورد کمترین توان‌های دوم و هم از روش برآورد ماکسیمم درست‌نمایی شرطی استفاده کرد، ولی برای آزمون مربوط به میانگین متحرک سری زمانی از روش برآورد کمترین توان‌های دوم به دلیل غیرخطی بودن پارامترها نمی‌توان استفاده کرد.

الگوریتم باکس-جنکینز (باکس و جنکینز، ۱۹۷۰) درصدد پاسخ به این سوال است که چگونه می‌توان تشخیص داد یک سری زمانی از یک فرایند  $AR$  خالص،  $MA$  خالص و یا از یک فرایند  $ARMA$  پیروی می‌کند، این روش شامل مراحل زیر است.

(۱) **تشخیص** در این مرحله، هدف یافتن مقادیر واقعی  $(p, d, q)$  است. بدین منظور از ابزار نمودار خودهمبستگی و خودهمبستگی جزئی می‌توان استفاده کرد.

(۲) **برآورد مدل** بعد از مرحله تشخیص به تخمین پارامترهای مدل پرداخته می‌شود. گاهی اوقات برای تخمین پارامترها از روش کمترین توان‌های دوم استفاده کرده اما زمانی که مدل نسبت به پارامترها غیرخطی باشد، روش‌های غیرخطی به کار برده می‌شود.

(۳) **پیش بینی** یکی از مهم‌ترین اهداف مدل‌سازی سری زمانی این است که مقادیر سری برای



زمان‌های آینده پیش‌بینی شود. مسئله‌ای که در اینجا مطرح می‌شود، کاربرد کلمه‌های پیش‌بینی یا پیش‌گویی به جای یکدیگر است. کلمه پیش‌بینی برای توصیف روش‌های غیرعینی و انجام سیستماتیک روند آن به کار می‌رود، در حالی که کلمه پیش‌گویی برای روش‌های عینی و نگاه به آینده به کار می‌رود. پیش‌بینی در بسیاری از کاربردها ارتباط بسیار نزدیکی با کنترل دارد. به‌عنوان مثال می‌توان پیش‌بینی انجام داد که آیا یک فرآیند تولیدی در حال توقف است یا خیر، و اگر چنین است، یک عمل اصلاحی انجام گیرد.

برای مطالعه بیشتر، **ماکریداکیس و همکاران (۱۳۹۴، فصل ۷)** را ببینید.

## فصل ۲

# مدل‌های سری زمانی برای اندازه‌گیری تغییرات

در فصل ۱، به اختصار مدل‌های ARMA برای نمایش سری‌های زمانی بررسی شده‌اند. اما گاهی اوقات، سری‌های زمانی‌ای که در واقعیت با آن‌ها روبرو می‌شویم ممکن است همیشه ذات یک فرآیند خطی را به نمایش نگذارند یا مدل ARMA، برای سری‌های زمانی، مدل خوبی نباشد. مدل‌های مربوط به سری‌های زمانی واریانس ناهمگن شرطی در ارتباط با مدل‌بندی فراتر از خانواده ARMA که برای سری‌های زمانی اقتصادی و مالی مناسب هستند، بررسی شده است. این سری‌ها، درجه‌های متفاوت بی‌ثباتی یا تغییرپذیری (شرطی) وابسته به پیشینه گذشته سری را به نمایش می‌گذارند. ابتدا برخی از مفاهیم مالی و اقتصادی معرفی می‌شود و سپس مدل‌های آماری مناسب برای آن‌ها بررسی می‌گردند.

### ۱.۲ مفاهیم مالی و اقتصادی

#### ریسک

برای اندازه‌گیری مخاطره (ریسک از دیدگاه اقتصاددانان) معیارهای متعددی طی دهه‌های اخیر معرفی شده است و هر کدام به نحوی به مقوله عدم اطمینان نگرینسته و تعدادی از آن‌ها

نیز مکمل همدیگر هستند. در سال ۱۹۳۸، مک کوئلی، معیار دوره<sup>۱</sup> را برای اندازه‌گیری ریسک اوراق بهادار با درآمد ثابت معرفی کرد. نخستین بار که مقوله ریسک کنار بازدهی مورد استفاده قرار گرفت در تحقیقات مارکوویتز (۱۹۵۲) هنگام ارزیابی مدلی کمی برای انتخاب سبد دارایی بود. وی انحراف معیار را ملاک سنجش ریسک قرار داد.

### ارزش در معرض ریسک

با پیشرفت سایر علوم از جمله علوم رایانه‌ای، ریاضی و آمار، ارزیابی ریسک‌های نامطلوب به‌طور گسترده مورد استفاده قرار گرفت، این پیشرفت‌ها در سال ۱۹۹۳ منجر به معرفی ارزش در معرض ریسک توسط موسسه جی. پی. مورگان<sup>۲</sup> گردید. به‌طوری که این معیار با در نظر گرفتن عوامل مختلف در ریسک یک عدد را به‌عنوان معیار ریسک معرفی می‌کند. جذاب بودن این معیار منجر به استفاده بیشتر از آن طی سال‌های اخیر شده است که با روش‌های مختلف آماری و ریاضی، دقت این مدل افزایش یافته و به این دلیل با استقبال سرمایه‌گذاران و تحلیل‌گران مالی روبه‌رو شده است (حیرانی و روشن‌ضمیر، ۱۳۹۷).

زیان از ترکیب دو عامل به‌وجود می‌آید: (۱) نوسان متغیرهای مالی بنیادی و (۲) قرار گرفتن در معرض این منابع ریسک. از آنجایی که بنگاه‌ها هیچ نوع کنترلی روی نوسان متغیرهای مالی ندارند، آن‌ها می‌توانند قرار گرفتن در معرض این ریسک‌ها را به‌عنوان مثال با استفاده از مشتق‌ها، تعدیل و تنظیم کنند. ارزش در معرض ریسک، اثر مشترک نوسان متغیرهای مالی بنیادی و قرارگیری در معرض ریسک‌های مالی را نشان می‌دهند.

ریسک می‌تواند به‌عنوان پراکندگی نتایج اندازه‌گیری شود. یک توزیع گسترده‌تر نشان‌دهنده ریسک بیشتری نسبت به یک توزیع متمرکزتر است. جنبه منفی ریسک معمولاً بر اساس چندک‌های توزیع سنجیده می‌شود. چندک، عبارت است از مقداری از متغیر ( $q$ ) که سطح سمت راست (یا چپ) آن نشان‌دهنده مقدار احتمال معین ( $\alpha$ ) است. در صورتی که توزیع ارزش سبد سرمایه موجود باشد ریسک را به راحتی می‌توان از رابطه

$$c = P(x \geq q) = \int_q^{+\infty} f(x)dx = 1 - F(q) \quad (1.2)$$

محاسبه کرد که در آن  $x$  ارزش دارایی،  $f(x)$  چگالی توزیع ارزش دارایی و  $F(x)$  توزیع تجمعی ارزش دارایی است. احتمال آنکه ارزش دارایی یا سبد دارایی کمتر از حد معینی باشد، برابر است با  $(1 - \alpha)$ .

اگر تابع توزیع، نرمال باشد چندک‌های آن از جداول آماری قابل استخراج خواهد بود. بدین ترتیب ارزش در معرض ریسک، حداکثر زیانی است که کاهش ارزش سبد دارایی برای دوره‌ی معینی در آینده با درصد اطمینانی معین، از آن بیشتر نمی‌شود. به‌عبارت دیگر، حداکثر زیانی که در یک سرمایه‌گذاری مالی به سبد دارایی‌های یک سرمایه‌گذار ممکن است در یک افق زمانی مشخص و سطح اطمینان معین وارد شود، برابر با ارزش در معرض ریسک<sup>۳</sup> (VaR)

<sup>1</sup>Duration

<sup>2</sup>J. P. Morgan

<sup>3</sup>Value at Risk (VaR)

است. بیان ریاضی ارزش در معرض ریسک به صورت

$$P(P_1 - P_0 \leq \text{VaR}) \leq \alpha \quad (2.2)$$

که  $P_0$  ارزش سبد دارایی در زمان صفر،  $P_1$  ارزش سبد در زمان یک و  $\alpha$  نیز سطح معناداری آماری است. رابطه (۲.۲) بیان می کند که احتمال اینکه کاهش ارزش سبد دارایی در دوره آتی، کمتر از ارزش در معرض ریسک باشد، حداکثر برابر با  $\alpha$  است. این مفهوم در اقتصاد مالی به ارزش در معرض ریسک در موقعیت خرید معروف است. بدین ترتیب در جهت دیگر احتمال این که زیان سبد دارایی در دوره آتی، بیش از ارزش در معرض ریسک باشد،  $1 - \alpha$  درصد است که به آن موقعیت فروش می گویند. حال اگر تابع توزیع تجمعی ارزش سبد دارایی در دوره آتی را به صورت  $F(p)$  نشان دهیم، معکوس تابع توزیع تجمعی نشان دهنده چندک توزیع بازدهی است. بنابراین، ارزش در معرض ریسک برای موقعیت های دوگانه را می توان به صورت

$$\text{VaR}(\alpha) = F^{-1}(\alpha)$$

$$\text{VaR}(\alpha) = F^{-1}(1 - \alpha)$$

به دست آورد. بدین ترتیب، برای آن که بتوان ارزش در معرض ریسک یک سبد دارایی را در بازار مالی اندازه گیری کرد، لازم است توزیع احتمال ارزش سبد دارایی ها برای دوره آتی برآورد شود. بنابراین، برآورد ارزش در معرض ریسک برای سبد دارایی، معادل فرآیند تخمین توزیع احتمال ارزش سبد دارایی در دوره آتی یا دوره نگهداری است.

### نوسان

نوسان<sup>۴</sup> برای کسانی که در بازارهای مالی شرکت می کنند، از اهمیت بالایی برخوردار است. در علم مالی، نوسان یک اندازه آماری برای بازده دارایی<sup>۵</sup> است. این اندازه آماری در طول زمان تغییر می کند و به صورت یک سری زمانی بیان می شود. در مسائل مالی، واریانس شرطی را نیز، نوسان نامیده و طبق تحقیقات انگل (۱۹۸۲) نوسان های بازار مالی قابل پیش بینی هستند. این بدان معنی است که سرمایه گذاران می توانند با استفاده نکردن از دارایی های پر نوسان، ریسک سبدها را کاهش دهند. در اینجا نوسان به معنای واریانس شرطی سود سرمایه اصلی است. در مسائل تجاری و نرخ ارز خارجی، نوسان نقش مهمی ایفا می کند. پیش بینی نوسان همانند پیش بینی واریانس، به بررسی اینکه میزان قیمت ها کم است یا زیاد می پردازد. به طور کلی پیش بینی نوسان یک عمل دقیق نیست. مانند پیش بینی باران، شما می توانید در پیش بینی باران درست عمل کنید اما هیچ بارانی ندارد.

نوسان، یک مفهوم مهم در امور مالی است که بروز آن دلایل متعددی دارد. در بیست سال اخیر مدل های نوسان و پیش بینی آن ها مورد توجه محققان دانشگاهی و فعالان بازار قرار گرفته

<sup>4</sup>Volatility

<sup>5</sup>Asset Returns

است. این توجه به خاطر این است که از نوسان به عنوان معیاری برای محاسبه ریسک استفاده شده است. برای شرکت‌های خصوصی، نوسان می‌تواند یک عامل مهم در بررسی احتمال ورشکستگی به حساب آید. هر چقدر نوسان در یک شرکت بیشتر باشد، احتمال پرداخت نکردن بدهی بیشتر است. نوسان در قیمت‌گذاری بسیاری از دارایی‌ها از جمله بیمه نقش بسزایی دارد. وقتی قیمت دارایی‌ها با تغییر زمان به شدت دچار نوسان می‌شود، مدل‌سازی نوسان و پیش‌بینی آن از بی‌اعتمادی سرمایه‌گذاران نسبت به بازار سرمایه جلوگیری می‌کند.

با پایه‌گذاری کمیته بازل<sup>۶</sup> در سال ۱۹۹۶، مدیریت ریسک مالی برای بسیاری از موسسات مالی در سرتاسر جهان جلوه مهمی پیدا کرد. در پیمان یاد شده، نوسان قیمت سهام به منظور نگهداری سرمایه ذخیره در قبال ارزش در معرض ریسک<sup>۷</sup> برای موسسات مالی به صورت اجبار درآمد. بنابراین، اندازه‌گیری نوسان برای تمامی موسسات مالی بسیار مهم می‌باشد (حیرانی و روشن‌ضمیر، ۱۳۹۷).

مدل‌سازی نوسان در سری‌ها زمانی مالی، از منظر پژوهشگران و فعالان بخش مالی، به لحاظ موارد استفاده آن در پیش‌بینی نوسان و ریسک سهام، موضوع جذابی برای مدل‌سازی است. زیرا این پیش‌بینی‌ها در مواردی چون مدیریت ریسک قیمت‌گذاری مشتقات مالی و پوشش ریسک ناشی از آن‌ها، بازاریابی، انتخاب سبدهای مالی و ... مورد استفاده قرار می‌گیرد. از این نظر، مدل‌سازی و پیش‌بینی نوسان در بازارهای مالی، اهمیت می‌یابد.

در ارتباط با مسائل اقتصادی و سری‌های زمانی مالی، مدل‌های متعددی برای بررسی نوسان ساخته شده است. یک فرض اولیه به هنگام مدل‌سازی نوسان این است که می‌توان نوسان را به دو بخش قابل پیش‌بینی و غیرقابل پیش‌بینی تقسیم کرد. با توجه به این حقیقت که در سری‌های زمانی، ارزش اضافه ریسک، تابعی از نوسان بازده است، تمرکز تحقیقات علمی بر جز قابل پیش‌بینی نوسان بازده می‌باشد. بنابر همین واقعیت، هر چند نوسان، همان ریسک محسوب نمی‌شود. ولی دانستن مقادیر نوسان به خاطر ارتباط آن با ریسک، مهم است.

به طور کلی مدل‌سازی ریسک در اقتصاد با سه رهیافت متفاوت پارامتریک<sup>۸</sup>، ناپارامتریک<sup>۹</sup> و نیمه‌پارامتریک<sup>۱۰</sup> انجام می‌شود. ویژگی مشترک هر سه رویکرد، استفاده از داده‌های تاریخی به منظور تعیین شکل توزیع احتمال است (حیرانی و روشن‌ضمیر، ۱۳۹۷).

در میان مدل‌های معرفی شده، رویکرد پارامتریک فرض خاصی را در مورد توابع احتمال بازدهی دارایی‌ها در نظر می‌گیرد و تمامی محاسبات براساس این فرض استوار خواهد بود. این توزیع‌ها می‌تواند نرمال، نرمال چوله،  $t$ -استیودنت،  $t$ -استیودنت چوله، توزیع خطای تعمیم‌یافته، توزیع خطای تعمیم‌یافته چوله یا دیگر توزیع‌های آماری در نظر گرفته شود. اگر  $\varepsilon_t$  دارای توزیع نرمال استاندارد  $\Phi(\varepsilon)$  باشد، آن‌گاه  $\text{VaR}(\alpha) = \mu + \sigma_t \Phi^{-1}(\alpha)$  است که در آن

<sup>6</sup>Basel committee

<sup>7</sup>Value at risk

<sup>8</sup>Parametric approaches

<sup>9</sup>NonParametric approaches

<sup>10</sup>Semi-Parametric approaches

$\Phi^{-1}(\alpha)$  چندک  $\alpha$  توزیع نرمال استاندارد با  $\Phi^{-1}(\alpha) < 0$  است. اگر فرض شود که  $\varepsilon_t$  دارای توزیع  $t$ -استیودنت با درجه آزادی  $k$  باشد، آن گاه  $\text{VaR}_t(\alpha) = \mu + \sigma_t T_k^{-1}(\alpha)$  که در آن  $T_k^{-1}(\alpha)$  چندک  $\alpha$  از توزیع  $t$ -استیودنت با درجه آزادی  $k$  است.

در رویکرد ناپارامتریک، هیچ فرضی بر تابع توزیع بازدهی‌ها تحمیل نمی‌شود و به داده‌ها اجازه داده می‌شود که در مورد خود، اظهار نظر کنند. در این رویکرد، فرض بر آن است که آینده نه چندان دور تا اندازه زیادی شبیه گذشته نزدیک بوده و می‌توان برای پیش‌بینی ریسک آینده از اطلاعات گذشته استفاده کرد، زیرا نحوه شکل‌گیری واریانس و ضریب همبستگی در نمونه تاریخی می‌تواند در آینده تکرار شود. به بیان دیگر از آخرین توزیع تجربی بازده و نه یک توزیع نظری برای برآورد سنج‌های ریسک بهره می‌گیرد. مهم‌ترین مدل‌های رویکرد ناپارامتریک که بر اساس شبیه‌سازی داده‌های تاریخی به برآورد ریسک آینده می‌پردازند، عبارت‌اند از: (۱) شبیه‌سازی تاریخی پایه‌ای، (۲) روش هیبریدی و (۳) روش شبیه‌سازی مونت کارلو.

برخی از رویکردها همچون رویکرد نیمه‌پارامتریک در ادبیات ریسک مطرح شده است که رویکردی بینابین را اختیار کرده‌اند. رویکرد نیمه پارامتریک، ویژگی‌هایی از هر دو رویکرد را شامل می‌شود. در این رویکرد سعی شده است که مشکلات رویکردهای قبلی را کاهش دهد. از این حیث می‌توان به شبیه‌سازی تاریخی موزون، شبیه‌سازی تاریخی فیلترشده، شبیه‌سازی تاریخی با استفاده از تخمین چگالی ناپارامتریک و شبکه عصبی اشاره کرد. اغلب این روش‌ها بهبودیافته رویکردهای پارامتریک و ناپارامتریک است. در راستای برآورد VaR، واعظ (۱۳۹۰) و مجرد (۱۳۹۵) را ببینید.

برای برآورد واریانس شرطی یا نوسان بازدهی‌های مالی،  $\sigma_t^2$  می‌توان از تکنیک‌های مختلفی استفاده کرد (کشاورز حداد، ۱۳۹۴). این مدل‌ها به سه دسته اصلی تقسیم می‌شوند: (۱) مدل‌های واریانس کوواریانس، (۲) مدل‌های نوسان چندگانه، (۳) مدل‌های سری زمانی تعمیم‌یافته.

مناسب‌ترین روش برای بهترین پیش‌بینی نوسان، استفاده از مدل‌سازی نوسان سری‌های زمانی آماری می‌باشد. با استفاده از سری‌های زمانی آماری می‌توان تعیین کرد که آیا اطلاعات اخیر نسبت به اطلاعات قدیم مهمتر هستند یا خیر. همچنین می‌توان تعیین کرد که آیا نوسان نسبت به حرکت‌های بالا و پایین بازار به یک میزان، متغیر است یا خیر (ک.دالی، ۲۰۰۶).

مبنای روش‌های مدل‌سازی نوسان بر اساس انحراف معیارهای تاریخی<sup>۱۱</sup>، این است که برای تمامی  $\tau > 0$  در زمان  $t-1$  مشخص یا قابل تخمین است، ساده‌ترین مدل‌سازی تاریخی، استفاده از  $\sigma_{t-1}$  جهت پیش‌بینی  $\sigma_t$  است. با بسط این ایده، مدل میانگین تاریخی<sup>۱۲</sup>، روش میانگین متحرک ساده<sup>۱۳</sup>، روش هموارسازی نمایی<sup>۱۴</sup> و میانگین متحرک موزون نمایی<sup>۱۵</sup>

<sup>11</sup> Historical standard deviation

<sup>12</sup> Historical average

<sup>13</sup> Simple moving average

<sup>14</sup> Exponentially smoothing moving average

<sup>15</sup> Exponentially Weighted moving average

حاصل شده است. روش میانگین تاریخی از تمامی انحراف معیارهای گذشته استفاده می‌کند. در حالی که روش میانگین متحرک، مقادیر خیلی قدیمی را رها می‌کند، روش هموارسازی نمایی از تمامی تخمین‌های تاریخی استفاده می‌کند. لیکن، روش میانگین متحرک موزون نمایی فقط مقادیر اخیر را به کار می‌برد. ولی برخلاف دو روش اول، دو روش نمایی به تخمین‌های نوسان اخیر، وزن بیشتری می‌دهند (حیرانی و روشن‌ضمیر، ۱۳۹۷).

اصلی‌ترین تفاوت در میان مدل‌ها، اغلب (۱) انتخاب متغیرهایی است که باید شرطی شوند و (۲) انتخاب صورت تابع برای مدل‌بندی است. مسائل مدل‌سازی سری با نوسان متغیر، اغلب در اقتصادسنجی تحت مسئله ناهمگنی مشهور هستند اما بیشترین علاقه در مدل‌های نوسان متغیر با زمان است که از دارایی نشات می‌گیرد. به‌عنوان مثال، یکی از اساسی‌ترین ایده‌های قیمت‌گذاری دارایی این است که تغییرپذیری هر دارایی باید در قیمت آن منعکس شود. باید انتظار داشت که یک دارایی با واریانس (ریسک) بالا، سود بیشتری دهد (برای سرمایه‌گذارانی که هدفشان سرمایه‌گذاری بلندمدت است). اما استدلال اقتصادسنجی بیان می‌کند که ریسک نباید به‌عنوان واریانس غیرشرطی، بلکه به‌عنوان واریانس شرطی، واریانسی که به اطلاعات گذشته خودش وابسته است، اندازه‌گیری شود.

یکی از ابزارهای سنتی مدل‌سازی سری‌های زمانی، مدل‌های ARMA هستند که از ترکیب مدل‌های MA و AR تشکیل شده‌اند. به‌هنگام استفاده از مدل‌های MA و AR به‌تنهایی، یکی از مشکلات این است که به‌منظور توضیح ساختار پویای داده‌ها باید از تعداد پارامترهای زیادی استفاده کنیم. برای غلبه بر این مشکل، مدل‌های ARMA شکل گرفتند. (تسای، ۲۰۰۲) اگرچه مفهوم مدل‌های ARMA بسیار مرتبط به مدل‌سازی نوسان می‌باشد، ولی برای سری‌های بازده مالی، شانس استفاده از مدل‌های ARMA اندک است. (بولرسلو و همکاران، ۱۹۹۸) مدل‌های ARMA در مدل‌سازی سری‌های زمانی به‌طور کلی مفید هستند که دارای فرض واریانس ثابت خطا می‌باشند. این مساله در بسیاری از زمینه‌های مالی و اقتصادی واقع‌بینانه نیست.

مدل‌های واریانس ناهمگن شرطی اجازه می‌دهند که واریانس در طی زمان تغییر کند، به‌طور خاص در مدل‌سازی متغیرهای بازار مالی، بسیاری از سری‌های زمانی مالی خوشه‌بندی نوسان<sup>۱۶</sup> را نشان می‌دهند به این معنی که سری زمانی مربوطه، در بعضی دوره‌ها نوسان پایین و در بعضی دوره‌های دیگر نوسان بالایی دارد. در اکثر منابع، از مدل ارزشیابی نوسان در طی این زمان به‌عنوان مدل‌های آماری و اقتصادسنجی بحث می‌شود، عمومی‌ترین مدل توسعه داده‌شده در مطالعات گذشته، موضوع تثبیت و تحلیل نوسان خوشه‌ای، مدل‌های خانواده واریانس ناهمگن شرطی اتورگرسیو<sup>۱۷</sup> (ARCH) (انگل، ۱۹۸۲) می‌باشند.

درباره اهمیت مدل‌های ARCH، همان‌طور که اشاره شد در مدل‌سازی سری‌های زمانی داده‌های اقتصادی، تغییرپذیری نیازمند توجه به تغییر واریانس در طول دوره مورد بررسی

<sup>16</sup>Volatility clustering

<sup>17</sup>Autoregressive Conditional Heteroscedastic Models (ARCH)

است. واریانس در تمام دوره‌ها لزوماً ثابت نیست و متناسب با زمان تغییر می‌کند. در غالب پدیده‌های انسانی نیز ناهمسانی واریانس قابل مشاهده است. برای مثال، اگر فرآیند یادگیری در نظر گرفته شود، افراد براساس تجربه‌های گذشته خود رفتارهای آینده را اصلاح می‌کنند که این ویژگی می‌تواند ایجاد واریانس ناهمسانی نماید. یک مثال متداول از این موضوع بررسی رابطه بین میزان خطا و ساعات تمرین است. طبیعی است در ساعات اولیه تمرین یک کار، افراد بنا بر استعدادها متفاوتی که دارند میزان خطای کاملاً متفاوتی دارند، ولی با افزایش ساعات تمرین از دامنه خطا کاسته می‌شود. به‌عنوان مثالی دیگر می‌توان به تفاوت در واریانس سطح مصرف براساس سطوح مختلف درآمدی اشاره کرد. معمولاً در درآمدهای پایین افراد انتخاب‌های مصرفی چندان متنوعی ندارند و دامنه نوسان در مصرف، محدود است اما به تدریج که به مشاهدات با درآمدهای بالا می‌رسیم دامنه نوسان در مصرف افزایش می‌یابد. روش معمول برای رفع ناهمسانی واریانس، تبدیل داده‌هاست. ولی روش معقول استفاده از مدل‌هایی است که شرط ناهمسانی واریانس را در برازش مدل‌ها در نظر بگیرند. مدل‌های ARCH و تعمیم‌های آن از جمله متداول‌ترین مدل‌ها برای نوسانات سری‌های زمانی داده‌های اقتصادی با فرض ناهمسانی واریانس می‌باشند.

مدل‌های ARCH پس از مقاله مهم انگل (۱۹۸۲) توسعه بسیاری یافته‌اند. هم‌اکنون مدل‌ها ARCH به‌طور وسیعی در مدل‌سازی نوسان سری‌های زمانی مالی به کار می‌روند. این‌گونه مدل‌ها بر خلاف مدل‌های قبلی، از انحراف معیار نمونه استفاده نمی‌کنند. بلکه واریانس شرطی را مدل‌سازی می‌کنند. مشکلی که در عمل به‌هنگام استفاده از این مدل‌ها به‌وجود می‌آید، این است که هنگامی که مرتبه مدل، عدد بزرگی باشد، به‌طور معمول باعث نقض شدن فرض غیرمنفی بودن و شرایط مانایی معادله واریانس می‌شود. مدل تعمیم‌یافته واریانس ناهمگن شرطی اتورگرسیو<sup>۱۸</sup> (GARCH) (بولرسلو، ۱۹۸۶) راهکاری برای حل این مشکل و نیز مدلی توانمند در تعداد پارامترهاست. مدل‌های واریانس ناهمگن شرطی اجازه می‌دهند که واریانس در طی زمان تغییر کند، مخصوصاً در مدل‌سازی متغیرهای بازار مالی که در مسایل مالی، واریانس شرطی بی‌ثباتی نامیده می‌شود. نحوه اندازه‌گیری نوسان و بی‌ثباتی متنوع است. به‌طور مثال، در مطالعات نرخ بهره، دامنه (فاصله) بین نرخ‌های بلندمدت و کوتاه‌مدت برای اندازه‌گیری نوسان فروش، ممکن است استفاده شود. چنین عمومیتی در تعریف به این حقیقت بستگی دارد که هیچ اندازه مستقیمی از نوسان در دسترس نیست. همانند خبرگان اقتصادسنجی، آماردانان نیز نوسان را به‌معنی واریانس شرطی در نظر می‌گیرند.

## ۱.۱.۲ مدل‌های واریانس ناهمگن شرطی اتورگرسیو

در مدل  $ARMA(p, q)$ ، رابطه (۱۳.۱)، برای سری زمانی  $X_t$  وقتی خطاهای  $a_t$  متغیرهای تصادفی مستقل (با فرضیه‌های معمول میانگین صفر و واریانس ثابت  $\sigma_a^2$ ) هستند، یک پیامد آن است

<sup>18</sup>Generalized Autoregressive Conditional Heteroscedastic Models (GARCH)



که واریانس شرطی  $a_t$  به شرط اطلاعات گذشته مقداری ثابت بوده و وابسته به گذشته نیست. از مهم‌ترین پیشرفت در مدل‌سازی تغییرات نوسان، مدل‌های اتورگرسیو مشروط به ناهمگنی واریانس یا مدل ARCH است که توسط انگل (۱۹۸۲) معرفی شده است. مدل ARCH در سری داده‌های مالی و اقتصادی بسیاری از کشورها به کار رفته است. این کاربرد عمدتاً از تنوع مدل‌های ARCH در به دست آوردن بعضی از خصوصیات داده‌های مالی و اقتصادی ناشی می‌شود که شامل توزیع‌های غیرشرطی با دنباله پهن و تغییرات واریانس در طول زمان هستند. کاربرد مدل‌های ARCH برای آزمایش مدل‌های مختلف ارزش‌گذاری دارایی مانند مدل ارزش‌گذاری دارایی<sup>۱۹</sup> (CAPM) برای گرفتن فرآیند ریسک سیستماتیک کاملاً مشهود است. در اقتصاد کلان از مدل‌های ARCH استفاده شده است تا رابطه بین واریانس شرطی زمان متغیر و پاداش ریسک<sup>۲۰</sup> (سودی که در مقابل قبول ریسک به دست می‌آوریم) مشخص شود. مدل‌های ARCH هم‌چنین برای اندازه‌گیری تورم، در مطالعه عوامل موثر در بانک مرکزی و مشخص کردن رابطه اقتصاد کلان و بازار سهام استفاده می‌شوند.

به عبارت دیگر، از دلایل استفاده از مدل‌های ARCH می‌توان به وجود خطاهای پیش‌بینی کوچک و بزرگ در خوشه‌های اقتصادی (مانند نرخ ارز، تورم، سهام و ...) اشاره کرد. به طوری که ممکن است سری طی سال‌های مختلف رفتارهای متفاوتی را از خود به نمایش بگذارد. به مفهوم دیگر، در برخی سال‌ها دارای نوسان کم و در برخی از سال‌های دیگر دارای نوسان زیاد باشد. در چنین شرایطی انتظار بر این است که واریانس در طول روند تصادفی سری موردنظر ثابت نبوده و تابعی از رفتار خطا باشد. در واقع مزیت مدل‌های ARCH در این است که می‌توان با استفاده از آن، روند واریانس شرطی را بر اساس اطلاعات گذشته آن توضیح داد. هم‌چنین در مدل‌های ARCH، ساختار خودهمبستگی برای واریانس شرطی در نظر گرفته می‌شود.

قبل از آنکه مدل ARCH اصلی بررسی و بسط داده شود، ابتدا به طور خلاصه گام‌هایی را که در اندازه‌گیری نوسان بازدهی دارایی مفیدند را مورد بررسی قرار می‌دهیم مدل‌های ARCH همگی دارای یک مشخصه رایج آماری در ارتباط با واریانس شرطی هستند که عبارت است از میانگین شرطی و در حقیقت یک متغیر تصادفی است که به صورت

$$m_t = E[Y_t | F_{t-1}] = E_{t-1}[Y_t] \quad (۳.۲)$$

تعریف می‌شود. که  $Y_t$  نرخ بازدهی سهام یا سبد سرمایه در زمان  $t-1$  تا  $t$  می‌باشد.  $F_{t-1}$  مجموعه اطلاعات گذشته است که شامل مقادیر دریافت شده از همه متغیرهای مربوطه تا زمان  $t-1$  است و  $E$  عملگر امید ریاضی است. به عبارت دیگر یک سرمایه‌گذار اطلاعات را در زمان  $F_{t-1}$  می‌داند و وقتی در زمان  $t$  تصمیم به سرمایه‌گذاری می‌گیرد، بازده مورد انتظار مربوط به سرمایه‌گذاری بر اساس مقدار امید شرطی  $Y_t$  طبق معادله (۳.۲) به دست می‌آید.

<sup>19</sup>Capital asset pricing model

<sup>20</sup>Risk premium

واریانس شرطی  $Y_t$  از رابطه

$$\sigma_t^2 = E_{t-1}[Y_t - m_t] \quad (4.2)$$

به دست می آید. از آن جایی که نوسان، تغییرات بازده را اندازه می گیرد و به اطلاعات گذشته  $F_{t-1}$  وابسته است، سرمایه گذاران با استفاده از واریانس شرطی  $\sigma_t^2$  و اطلاعات گذشته  $F_{t-1}$  پیش بینی دقیق تری برای نوسانات انجام می دهند.

برای تجزیه بازدهی دارایی در زمان  $t$  سه گام زیر باید اعمال شود: (۱) معلوم کردن  $m_t$ ، (۲) معلوم کردن  $\sigma_t^2$  و (۳) معلوم کردن تابع چگالی خطای مدل.

به طور کلی در بازارهای مالی  $m_t$  به عنوان پاداش ریسک تعیین شده است که برای داده های با نوسان بالا، مقدار آن صفر است. بدین معنی که در زمان نوسان های شدید، سود قابل توجهی در برابر پذیرفتن ریسک سرمایه گذاری به دست نمی آید. معمولاً بیشتر روی گام ۲ تمرکز داشته و به ویژه محققان چگونگی وابستگی واریانس شرطی به اطلاعات گذشته را بررسی می کنند.

فرض کنید  $x_1, \dots, x_T$  مشاهدات یک سری زمانی باشد و  $F_{t-1}$  مجموعه ای از  $X_t$  ها تا زمان  $t$  باشد که  $t \geq 0$ . مدل های ARCH برای مجاز دانستن وابستگی واریانس شرطی خطا به مقدار توان دوم خطاهای گذشته در یک فرآیند کلی ARMA، به جای ثابت بودن آن مثل مدل های خانواده ARMA با خطاهای مستقل پیشنهاد شدند. به عنوان مثال، فرض کنید یک مدل  $AR(p)$  با خطاهای مدل  $ARCH(r)$  به صورت

$$X_t = \sum_{i=1}^p \phi_i x_{t-i} + \phi_0 + a_t \quad (5.2)$$

داده می شود که در آن  $E[a_t | a_{t-1}, a_{t-2}, \dots, a_0] = 0$  و

$$\sigma_t^2 \equiv V(a_t | a_{t-1}, a_{t-2}, \dots) = E[a_t^2 | a_{t-1}, a_{t-2}, \dots] = \alpha_0 + \sum_{i=1}^r \alpha_i a_{t-i}^2 \quad (6.2)$$

که  $r > 0$  و برای تضمین اینکه واریانس شرطی  $\sigma_t^2$  مثبت باشد،  $\alpha_0 > 0$  و به ازای  $i = 1, \dots, r$ ،  $\alpha_i \geq 0$ .

یک فرمول بندی هم ارز مدل ARCH به صورت

$$a_t = \sigma_t e_t^2$$

است که در آن دنباله ای از متغیرهای تصادفی مستقل و هم توزیع با میانگین صفر و واریانس ۱ است. چون  $e_t$  مستقل از گذشته است، در نتیجه

$$V(a_t | a_{t-1}, a_{t-2}, \dots) = E[\sigma_t e_t^2 | a_{t-1}, a_{t-2}, \dots] = \sigma_t E[e_t^2] = \sigma_t$$

خطاهای  $a_t$  دارای میانگین صفر هستند زیرا

$$E[a_t] = E[E[a_t | a_{t-1}, \dots]] = 0$$

و به‌طور پیاپی ناهمبسته هستند. به عبارت دیگر، به‌ازای  $j \neq 0$ ،  $E[a_t a_{t-j}] = 0$  زیرا مثلاً به‌ازای  $j > 0$ ،

$$E[a_t a_{t-j}] = E[E[a_t a_{t-j} | a_{t-1}, \dots]] = E[a_{t-j} E[a_t | a_{t-1}, \dots]] = 0$$

اما  $a_t$  ها متغیرهای تصادفی متقابلاً مستقل نیستند زیرا از طریق واریانس‌های شرطی‌شان (به عبارت دیگر، گشتاور مرتبه دوم خود) با هم ارتباط دارند. همچنین فرض می‌شود که  $a_t$  ها واریانس‌های غیرشرطی برابر دارند یعنی به‌ازای همه مقادیر  $t$ ،  $V(a_t) = \sigma_a^2$  به طوری که مانای ضعیف هستند.

## ۲.۱.۲ مدل‌سازی و برآورد پارامترها

واقعیت‌های پیشین، و تعمیم آن‌ها به مدل‌های مراتب بالا، انگیزه به‌کارگیری ACF و PACF توان‌های دوم  $a_t$  را برای مشخص‌سازی مدل و بازرسی مقدماتی مربوط به وجود اثرهای ARCH در خطاهای  $a_t$  فراهم می‌سازند. در عمل، نقطه شروع بررسی ACF و PACF نمونه‌ای توان دوم مانده‌های  $\hat{a}_t^2$  خواهد بود که  $\hat{a}_t$  مانده‌های به‌دست آمده از برازش یک مدل  $ARMA(p, q)$  معمول‌اند (مک‌لئود و لی، ۱۹۸۳).

به‌ویژه، فرض کنید  $r_k(\hat{a}^2)$  خودهمبستگی‌های توان دوم مانده‌های  $\hat{a}_t^2$  حاصل از یک مدل ARMA برازش شده را نشان می‌دهد، یعنی،

$$r_k(\hat{a}^2) = \frac{\sum_{t=1}^{n-k} (\hat{a}_t^2 - \hat{\sigma}_a^2)(\hat{a}_{t+k}^2 - \hat{\sigma}_a^2)}{\sum_{t=1}^n (\hat{a}_t^2)^2}$$

که در آن  $\hat{\sigma}_a^2 = n^{-1} \sum_{t=1}^n \hat{a}_t^2$  برآورد واریانس مانده‌هاست. مشابه با آماره کلی نگر باکس و پییرس (۱۹۷۰)، مک‌لئود و لی (۱۹۸۳) آماره کلی نگر

$$\tilde{Q}(\hat{a}^2) = n(n+2) \sum_{k=1}^n \frac{r_k^2 \hat{a}^2}{n-k} \quad (7.2)$$

را به‌عنوان یک بازرسی تشخیصی برای کشف نارسایی مدل پیشنهاد کردند. ماهیت نارسایی مدل که  $\tilde{Q}(\hat{a}^2)$  می‌تواند برای آن سودمند باشد، احتمالاً در جهت واریانس ناهمگن شرطی (یعنی، اثرهای ARCH و GARCH، و غیرنرمال بودن توزیع حاشیه‌ای  $a_t$  ها) است. مک‌لئود و لی (۱۹۸۳) ثابت کردند که آماره  $\tilde{Q}(\hat{a}^2)$  تقریباً دارای توزیع  $\chi^2$  با  $k$  درجه آزادی تحت فرض نارسایی مدل (یعنی،  $a_t$  ها مستقل بوده و بنابراین فاقد واریانس ناهمگن شرطی هستند) است.

برای برآورد ماکسیمم درست‌نمایی مدل‌های ARMA ی

$$\phi(\beta)X_t = \theta_0 + \theta(\beta)a_t$$

با خطاهای  $a_t$  ی ARCH، فرض کنید که  $a_t$  ها به‌طور شرطی دارای توزیع نرمال  $N(0, \sigma_t^2)$  هستند، آن‌گاه  $X$  ها، به شرط  $x_{t-1}, \dots$  نرمال شرطی با واریانس شرطی  $\sigma_t^2$  هستند و از روی

چگالی توام که به صورت

$$P(\mathbf{x}) = \prod_{t=1}^n P(x_t | x_{t-1}, \dots, x_1)$$

بیان می شود، تابع لگاریتم درست‌نمایی (صرف نظر از برخی شرایط آغازین) به صورت

$$l = \log(L) = -\frac{T}{2} \log(2\pi) - \frac{1}{2} \sum_{t=1}^T \log(\sigma_t^2) - \frac{1}{2} \sum_{t=1}^T \frac{a_t^2}{\sigma_t^2}$$

است که در آن  $a_t = X_t - \sum_{i=1}^p \phi_i x_{t-i} - \theta_0 + \sum_{i=1}^q \theta_i a_{t-i}$  و  $\sigma_t^2$  در رابطه (۶.۲) تعریف شده است. روش‌های برآورد درست‌نمایی ماکسیمم، خواص برآوردگرها، و شیوه‌های امتیازی یا ضرایب لاگرانژ (LM) و نسبت درست‌نمایی (LR) برای آزمون حضور اثرهای ARCH در واریانس‌های شرطی  $a_t$  توسط انگل (۱۹۸۲) بررسی شده‌اند. به‌ویژه، آزمون امتیازی برای آزمون  $\alpha_i = 0$  به‌ازای  $i = 1, \dots, r$  در مدل ARCH(r) صورتی نسبتاً ساده اختیار می‌کند. چون در مدل ARCH(r) و بنابراین  $\partial \sigma_t^2 / \partial \alpha_i = a_{t-i}^2$ ،

$$\frac{\partial}{\partial \alpha_i} \left( \frac{-a_t^2}{\sigma_t^2} \right) = \left( \frac{a_t^2}{\sigma_t^4} \right) a_{t-i}^2.$$

بردار امتیاز مربوط  $\partial l / \partial \alpha$ ، که در آن  $\alpha = (\alpha_1, \dots, \alpha_r)^\top$ ، پس از ارزیابی تحت مدل صفر یا  $\alpha_i = 0$  به‌صورت مقداری ثابت در عنصرهای  $\sum_{t=1}^T a_t^2 a_{t-i}^2$ ، به‌ازای  $i = 1, \dots, r$  خواهد بود. برای جزئیات بیشتر باکس و همکاران (۱۳۹۴) را ببینید.

### ۳.۱.۲ مدل ARCH مرتبه اول

حالت ساده مدل ARCH مرتبه اول یا ARCH(1) را به‌صورت

$$\sigma_t^2 = E[a_t^2 | a_{t-1}, \dots] = \alpha_0 + \alpha_1 a_{t-1}^2$$

در نظر بگیرید. فرض کنید  $0 \leq \alpha_1 < 1$  (واریانس باید مثبت باشد)، واریانس غیرشرطی برابر است با

$$\sigma_a^2 = V(a_t) = \frac{\alpha_0}{1 - \alpha_1} \quad (۸.۲)$$

زیرا

$$\sigma_a^2 = E[a_t^2] = E[E[a_t^2 | a_{t-1}, \dots]] = E[\alpha_0 + \alpha_1 a_{t-1}^2] = \alpha_0 + \alpha_1 \sigma_a^2$$

بنابراین، واریانس شرطی  $a_t$ ، خطای پیش‌بینی یک گام به جلو، می‌تواند به‌صورت

$$\sigma_t^2 = \alpha_0 + \alpha_1 a_{t-1}^2 = \sigma_a^2 (1 - \alpha_1) + \alpha_1 a_{t-1}^2 = \sigma_a^2 + \alpha_1 (a_{t-1}^2 - \sigma_a^2) \quad (۹.۲)$$

یا به شکل انحراف از واریانس غیرشرطی به صورت  $\sigma_t^2 - \sigma_a^2 = \alpha_1(a_{t-1}^2 - \sigma_a^2)$  نوشته می‌شود. در این صورت، واریانس شرطی  $a_t$  بزرگتر از واریانس غیرشرطی خواهد بود هرگاه  $a_{t-1}^2$  از واریانس غیرشرطی‌اش بزرگتر باشد.

واریانس‌های شرطی خطاهای پیش‌بینی چندگام به جلو  $e_t(l) = \sum_{j=0}^{l-1} \phi_j a_{t+1-j}$  را نیز می‌توان اثبات کرد که به‌ازای  $h > 0$  بر پایه استدلال (۱۰.۲) وابسته به توان دوم خطاهای گذشته هستند.

$$E[e_t^2(l)|a_t, \dots] = \sum_{j=0}^{l-1} \phi_j^2 E[a_{t+l-j}^2|a_t, \dots] \quad (10.2)$$

که در آن

$$\begin{aligned} E\{a_{t+h}^2|a_t, \dots\} &= E\left[E\left[a_{t+h}^2|a_{t+h-1}, \dots\right] | a_t, \dots\right] \\ &= \alpha_0 + \alpha_1 E\left[a_{t+h-1}^2|a_t, \dots\right] = \dots \\ &= \alpha_0(1 + \alpha_1 + \dots + \alpha_1^{h-1}) + \alpha_1^h a_t^2 \end{aligned}$$

از این نتیجه و با استفاده از (۸.۲) می‌توان نوشت

$$E\left[e_t^2(l)|a_t, \dots\right] = \sigma_a^2 \sum_{j=0}^{l-1} \phi_j^2 + \sum_{j=0}^{l-1} \phi_j^2 \alpha_1^{l-j} (a_t^2 - \sigma_a^2)$$

که نتیجه (۹.۲) را به واریانس‌های شرطی خطاهای پیش‌بینی چندگام به جلو تعمیم می‌دهد. علاوه بر آن، اگر  $a_t$ ، به شرط گذشته، به‌طور شرطی به‌صورت نرمال توزیع شده باشد، آن‌گاه

$$E\left[a_t^4|a_{t-1}, \dots\right] = 3\sigma_t^4 = 3(\alpha_0 + \alpha_1 a_{t-1}^2)^2$$

بنابراین گشتاور چهارم غیرشرطی  $a_t$  در رابطه

$$\begin{aligned} E\left[a_t^4\right] &= E\left[E\left[a_t^4|a_{t-1}, \dots\right]\right] \\ &= 3\left[\alpha_0^2 + 2\alpha_0\alpha_1 E\left[a_{t-1}^2\right] + \alpha_1^2 E\left[a_{t-1}^4\right]\right] \end{aligned} \quad (11.2)$$

صدق می‌کند. پس اگر فرض شود که  $\{a_t\}$  مانای مرتبه چهارم با گشتاور چهارم غیر شرطی  $\mu_4 = E\left[a_t^4\right]$  است. از رابطه (۱۱.۲) می‌توان نوشت:

$$\begin{aligned} \mu_4 &= \frac{3(\alpha_0^2 + 2\alpha_0\alpha_1\sigma_a^2)}{1 - 3\alpha_1^2} \equiv \frac{3\alpha_0^2(1 + \alpha_1)/(1 - \alpha_1)}{1 - 3\alpha_1^2} \\ &> 3 \frac{\alpha_0^2}{(1 - 3\alpha_1)^2} = 3\sigma_a^4 \end{aligned} \quad (12.2)$$

این نتیجه دو کاربرد مهم دارد:

(۱) از آن جایی که گشتاور چهارم مثبت است،  $\alpha_1$  باید در شرط  $1 - 3\alpha_1^2 > 0$  صدق کند یعنی  $0 \leq \alpha_1 < \frac{1}{\sqrt{3}}$ .

(۲) با توجه به اینکه  $E[a_t^4]$  از  $3\sigma_a^4$  (مقدار مربوط به توزیع نرمال) بزرگتر است، توزیع حاشیه‌ای  $a_t$  دم‌های کلفت‌تری از توزیع نرمال را به نمایش خواهد گذاشت زیرا  $\mu_4 = \mu_4 > 0$ ، به عبارت دیگر، خطای  $a_t$  تحت مدل  $ARCH(1)$  در تولید داده‌های پرت بیشتر از حالت نرمال نقش ایفا می‌کند.

این مسئله با نتایج تجربی مطابقت دارد که بیان می‌کند داده‌های پرت اغلب در سود سرمایه ظاهر می‌شوند تا اینکه توسط دنباله متغیرهای تصادفی نرمال تولید شوند. ویژگی‌های بیان شده برای مدل‌های ARCH کلی برقرارند اما فرمول‌ها برای مراتب بالاتر مدل‌های ARCH، پیچیده‌تر می‌شوند.

### برآورد پارامترها در مدل $ARCH(1)$

پارامترهای  $\alpha_0$  و  $\alpha_1$  را می‌توان با استفاده از روش درست‌نمایی ماکزیمم برآورد کرد. تابع چگالی احتمال توام  $X_1, X_2, \dots, X_T$  به صورت

$$f_{x_1, \dots, x_T}(x_1, \dots, x_T) = \left\{ \prod_{j=1}^T f_{x_j | x_1, \dots, x_{j-1}}(x_j | x_1, \dots, x_{j-1}) \right\} \times f_{x_1}(x_1) \quad (13.2)$$

و چگالی شرطی آن برای  $k = 2, \dots, T$  به صورت

$$f_{x_k | x_1, \dots, x_{k-1}}(x_k | x_1, \dots, x_{k-1}) = \frac{1}{\sqrt{2\pi(\alpha_0 + \alpha_1 x_{k-1}^2)}} \exp \left\{ \frac{-x_k^2}{2(\alpha_0 + \alpha_1 x_{k-1}^2)} \right\} \quad (14.2)$$

است. چگالی حاشیه‌ای  $X_1$  کنار گذاشته می‌شود و تابع درست‌نمایی حاصل عبارت است از:

$$L(\alpha_0, \alpha_1) = \prod_{j=2}^T \frac{1}{\sqrt{2\pi(\alpha_0 + \alpha_1 x_{j-1}^2)}} \exp \left\{ \frac{-x_j^2}{2(\alpha_0 + \alpha_1 x_{j-1}^2)} \right\}. \quad (15.2)$$

لگاریتم تابع درست‌نمایی، بدون در نظر گرفتن جمله ثابت، به صورت

$$l(\alpha_0, \alpha_1) = \frac{-1}{2} \sum_{j=2}^T \left\{ \log(\alpha_0 + \alpha_1 x_{j-1}^2) + \frac{x_j^2}{\alpha_0 + \alpha_1 x_{j-1}^2} \right\} \quad (16.2)$$

به دست می‌آید که می‌توان برآورد  $\hat{\alpha}_0$  و  $\hat{\alpha}_1$  را با حل معادلات

$$\frac{dl}{d\alpha_1} = 0, \quad \frac{dl}{d\alpha_0} = 0$$

به دست آورد که با مشتق‌گیری ساده به دست نمی‌آیند و باید از روش‌های عددی برای محاسبه آن‌ها استفاده نمود. (انگل، ۱۹۸۲)

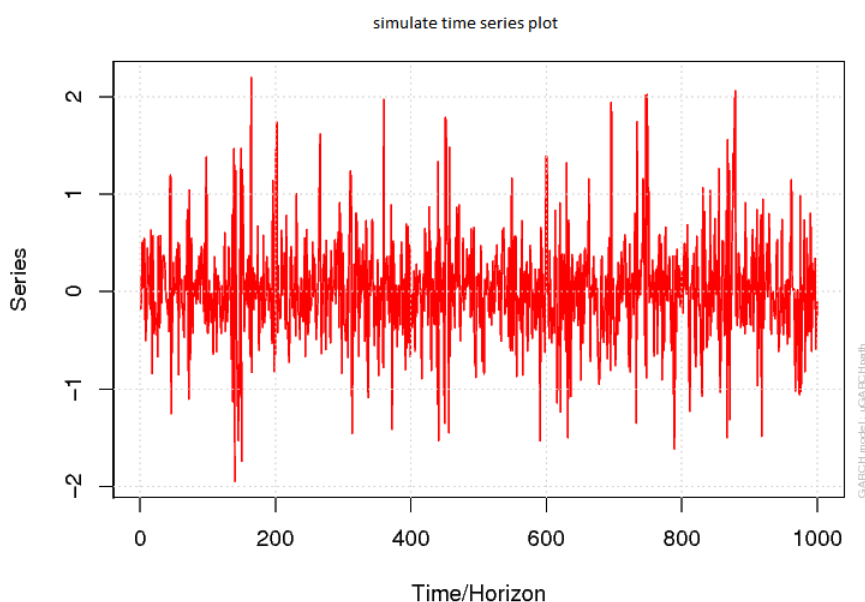
**مثال ۱.۱.۲.** در نرم‌افزار R، بسته‌ای به نام rugarch (گلانوس، ۲۰۱۴) وجود دارد که به کمک آن، یک مدل  $ARCH(1)$  شبیه‌سازی شده است. فرض کنید در رابطه (۱۳.۱)،  $\alpha_0 = 0.1$  و  $\alpha_1 = 0.8$ ، در این صورت، ابتدا با استفاده از دستورهای

```
arch1.spec = ugarchspec(variance.model=list(garchOrder= c(1,0)),
mean.model = list(armaOrder=c(0,0)),fixed.pars = list(mu = 0, omega=0.1,
alpha1=0.8))
```

مدل تعریف می‌شود و سپس با دستور

```
arch1.sim = ugarchpath(arch1.spec, n.sim = 1000)
```

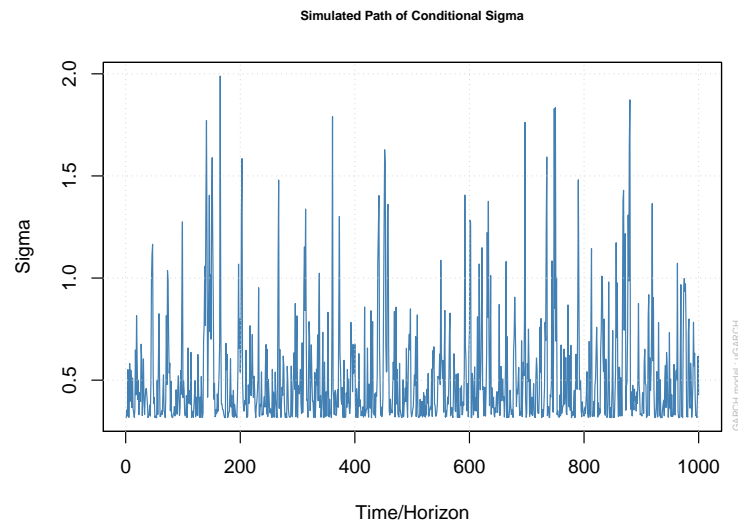
تعداد ۱۰۰۰ نمونه تصادفی از آن تولید شده است. شکل ۱.۲ نمودار سری زمانی و شکل ۲.۲ نمودار نوسان را برای مدل ARCH(۱) نشان می‌دهد.



شکل ۱.۲: نمایش سری زمانی ARCH(۱)

## ۴.۱.۲ مدل GARCH

از آنجایی که (۱) مدل‌های ARCH، در مورد سودهای مثبت و منفی با روش مشابه عمل می‌کند. زیرا به توان دوم سودهای قبلی بستگی دارند. در عمل، برای سری‌های زمانی مالی، قیمت‌ها به‌طور متفاوتی، متأثر از سودهای مثبت و یا منفی هستند، (۲) مدل ARCH نسبتاً محدود است. به‌طور مثال برای مدل ARCH(۱)، اگر سری زمانی دارای گشتاور چهارم متناهی باشد، در این صورت  $\alpha_2$  باید در فاصله  $(\frac{1}{4}, \frac{5}{4}]$  باشد. که این شرط برای مدل‌های ARCH با مراتب بالاتر حتی قوی‌تر است؛ (۳) مدل ARCH بینش جدیدی برای درک سری‌های زمانی مالی ارائه نمی‌کند. این مدل فقط یک روش مکانیکی را برای توصیف رفتار واریانس شرطی ارائه می‌دهد. همچنین بیان نمی‌کند که به چه دلیل این چنین رفتاری اتفاق می‌افتد؛ **بولرسلو**



شکل ۲.۲: نمایش نوسان مدل ARCH(۱)

(۱۹۸۶) مدل ARCH تعمیم‌یافته (GARCH) را بیان کرد که در این مدل تعمیم‌یافته، معادله واریانس شرطی علاوه بر  $r$  توان دوم خطای گذشته، شامل  $s$  وقفه واریانس‌های شرطی گذشته نیز است.

فرآیند  $\{a_t\}$  یک مدل GARCH از مرتبه  $r$  و  $s$  است (مرتبه تاخیر  $s$ ، طول زمان را برای واریانس شرطی دنباله خطاها مشخص می‌کند) اگر  $\sigma_t^2$  در مدل‌های ARCH با

$$\sigma_t^2 = E \left[ a_t^2 | a_{t-1}, a_{t-2}, \dots \right] = \alpha_0 + \alpha_1 a_{t-1}^2 + \alpha_r a_{t-r}^2 + \beta_1 \sigma_{t-1}^2 + \dots + \beta_s \sigma_{t-s}^2$$

یا

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^r \alpha_i a_{t-i}^2 + \sum_{j=1}^s \beta_j \sigma_{t-j}^2$$

جایگزین شود که  $r > 0$ ،  $s > 0$ ،  $\alpha_0 > 0$ ، به ازای  $i = 1, \dots, r$ ،  $\alpha_i \geq 0$ ، هم‌چنین، به‌ازای  $j = 1, \dots, s$ ،  $\beta_j \geq 0$ ، شرط‌های  $\alpha_0 > 0$ ،  $\alpha_i \geq 0$  و  $\beta_j \geq 0$  برای تضمین مثبت بودن واریانس شرطی  $\sigma_t^2$ ، ضروری است.

نتایجی که برای مدل ARCH(1) بیان شدند، برای مدل‌های مراتب بالاتر ARCH(r) و GARCH نیز برقرار هستند. تاثیر بنیادی خطاهای ARCH و GARCH موجود در فرآیند در ارزیابی درستی پیش‌بینی‌هاست زیرا خطاهای پیش‌بینی دارای واریانس‌های شرطی خواهند بود که به گذشته وابسته‌اند. این امر تشکیل بازه‌های پیشگویی (شرطی) درست و آگاهی بخشتر از آنچه را تحت مفروضات معمولی مبنی بر اینکه واریانس‌های خطای شرطی ثابت و مستقل از گذشته‌اند به دست خواهند آمد، میسر می‌سازد.



## ۵.۱.۲ مدل GARCH(1, 1)

ساده‌ترین فرآیند GARCH، GARCH(1,1) است، که در آن

$$\sigma_t^2 = E[a_t^2 | a_{t-1}, \dots] = \alpha_0 + \alpha_1 a_{t-1}^2 + \beta_1 \sigma_{t-1}^2. \quad (17.2)$$

مشابه مدل ARCH(1)، با فرض  $\alpha_1 + \beta_1 < 1$ ، ثابت می‌شود که واریانس غیرشرطی  $a_t$  برابر است با

$$\sigma_a^2 = V(a_t) = \frac{\alpha_0}{1 - (\alpha_1 + \beta_1)}$$

اغلب این‌طور استنباط می‌شود که زمانی که مدل‌های ARCH به داده‌های مالی برازش داده می‌شوند، مرتبه بالاتری جهت یک برازش رضایت‌بخش مورد نیاز است (بولرسلو و همکاران، ۱۹۹۲). برای داده‌هایی که از فرآیند GARCH(1, 1) پیروی می‌کنند، با جای‌گذاری  $\sigma_{t-1}^2$  به‌طور بازگشتی در معادله (17.2) به‌صورت

$$\begin{aligned} \sigma_t^2 &= \alpha_0 + \alpha_1 a_{t-1}^2 + \beta_1 (\alpha_0 + \alpha_1 a_{t-2}^2 + \beta_1 \sigma_{t-2}^2) \\ &= \alpha_0 (1 + \beta_1) + \alpha_1 a_{t-1}^2 + \alpha_1 \beta_1 a_{t-2}^2 + \beta_1^2 \sigma_{t-2}^2 \\ &= \alpha_0 (1 + \beta_1) + \alpha_1 a_{t-1}^2 + \alpha_1 \beta_1 a_{t-2}^2 + \beta_1^2 (\alpha_0 + \alpha_1 a_{t-3}^2 + \beta_1 \sigma_{t-3}^2) \\ &= \alpha_0 (1 + \beta_1 + \beta_1^2) + \alpha_1 a_{t-1}^2 + \alpha_1 \beta_1 a_{t-2}^2 + \alpha_1 \beta_1^2 a_{t-3}^2 + \beta_1^3 \sigma_{t-3}^2 \\ &= \dots = \alpha_0 \sum_{j=1}^k \beta_1^{j-1} + \alpha_1 \sum_{j=1}^{k-1} \beta_1^{j-1} a_{t-j}^2 + \beta_1^k \sigma_{t-k}^2 \end{aligned}$$

به‌دست می‌آید. شرط  $\alpha_1 + \beta_1 < 1$  برای متناهی بودن واریانس  $a_t$ ، لازم است به این معنی که  $\beta_1$  اکیدا کوچکتر از یک است. بنابراین، اگر  $k \rightarrow \infty$ ،  $\sigma_t^2$  به‌صورت

$$\sigma_t^2 = \frac{\alpha_0}{1 - \beta_1} + \alpha_1 \sum_{j=1}^{\infty} \beta_1^{j-1} a_{t-j}^2 \quad (18.2)$$

است که اگر

$$\alpha_0^* = \frac{\alpha_0}{1 - \beta_1}, \quad \alpha_j^* = \alpha_1 \beta_1^{j-1}; \quad j = 1, \dots, \infty,$$

آن‌گاه رابطه (18.2) را می‌توان به صورت

$$\sigma_t^2 = \alpha_0^* + \sum_{j=1}^{\infty} \alpha_j^* a_{t-j}^2$$

بازنویسی کرد که متناظر با ARCH( $\infty$ ) است. این نتیجه، پیشنهاد می‌کند که اگر یک مدل GARCH(1,1) به جای یک مدل ARCH(r) با مرتبه بالا جایگزین شود، مدل مناسب‌تری به‌دست می‌آید.

از طرف دیگر، فرض کنید  $\nu_t = a_t^2 - \sigma_t^2$ ، به‌طوری که متغیرهای تصادفی  $\nu_t$  دارای میانگین صفر بوده و به‌طور پیایی ناهمبسته باشند، زیرا

$$E[(a_t^2 - \sigma_t^2)(a_{t-j}^2 - \sigma_{t-j}^2)] = E[E(a_t^2 - \sigma_t^2)(a_{t-j}^2 - \sigma_{t-j}^2) | a_{t-1}, \dots]$$

$$= E \left[ (a_{t-j}^2 - \sigma_{t-j}^2) E \left[ (a_t^2 - \sigma_t^2) | a_{t-1}, \dots \right] \right] = 0$$

در این صورت، با توجه به این که  $\sigma_t^2 = a_t^2 - \nu_t$ ، دقت کنید که مدل GARCH(1,1) می‌تواند به صورت

$$a_t^2 - \nu_t = \alpha_0 + \alpha_1 a_{t-1}^2 + \beta_1 (a_{t-1}^2 - \nu_{t-1})$$

یا

$$a_t^2 = \alpha_0 + (\alpha_1 + \beta_1) a_{t-1}^2 + \nu_t - \beta_1 \nu_{t-1} \quad (19.2)$$

بازآرایی شود. این صورت، ملاحظه می‌شود فرآیند توان دوم خطاها،  $a_t^2$ ، از مدل ARMA(1,1) با نوسازی‌های ناهمبسته  $\nu_t$  پیروی می‌کند (اما،  $\nu_t$ ها می‌توانند واریانس ناهمگن باشند). در حالت خاص، مدل ARCH(1)،  $\beta_1 = 0$  و (19.2) به صورت مدل AR(1) ساده برای  $a_t^2$ ، یعنی به  $a_t^2 = \alpha_0 + \alpha_1 a_{t-1}^2 + \nu_t$  کاهش پیدا می‌کند.

## ۲.۲ مدل‌های ARMA-GARCH

معمولاً خطای مدل به صورت نوفه سفید<sup>۲۱</sup> در نظر گرفته می‌شود که می‌تواند واریانس برابر ۱ یا مقدار ثابت  $\sigma^2$  داشته باشد. اما چنانچه بخواهیم نوسانات و تغییرات واریانس را نیز در مدل بندی سری زمانی وارد کنیم، لازم است از مدل‌های توسعه یافته تری در تکمیل مدل ARMA بهره گرفته و مدل تغییرات را نیز فرمول بندی کنیم. برای مثال، در مسائل مالی و تجاری بررسی و ارزیابی تغییرات و نوسانات بازده سود بانکی یا سهام از اهمیت بسزایی برخوردار است. مدلی که معمولاً برای نوسانات به کار می‌رود، مدل ARCH است، بنابراین می‌توان مدل بازده را در طول زمان ARMA در نظر گرفت و نوسانات آن را با ARCH یا نسخه تعمیم یافته آن، GARCH، مدل بندی کرد. در این صورت، مدل سری زمانی، ARMA-GARCH خواهد بود. مدل ترکیبی ARMA-GARCH تغییرات میانگین و ناپایداری واریانس شرطی داده‌ها را توأم می‌کند. این مدل به تنهایی دارای کارایی مناسبی برای مدل نمودن تغییرات ناگهانی در سری و ناپایداری واریانس شرطی مشاهدات می‌باشد.

مدل ARMA-GARCH تلفیقی از مدل‌های ARMA و GARCH می‌باشد. در این مدل، میانگین شرطی مشاهدات دارای مدل ARMA و واریانس شرطی مشاهدات به شرط مشاهدات قبلی دارای مدل GARCH می‌باشند. به عبارت دیگر، هرگاه فرآیند  $X_t$  در روابط زیر صدق کند، این فرآیند دارای مدل ARMA-GARCH خواهد بود:

$$X_t = \phi_0 + \phi_1 x_{t-1} + \dots + \phi_{t-p} x_{t-p} + \theta_1 a_{t-1} + \dots + \theta_q a_{t-q} + a_t \quad (20.2)$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^r \alpha_i a_{t-i}^2 + \sum_{j=1}^s \beta_j \sigma_{t-j}^2 \quad (21.2)$$

<sup>21</sup> White noise

که در این مدل،  $\alpha_0 > 0$  و  $\alpha_i \geq 0$  و  $i = 1, \dots, r$  و  $\beta_j \geq 0$  و  $j = 1, \dots, s$  پارامترهای مدل هستند.

فرض کنید  $\{\epsilon_t, t \in T\}$  یک دنباله مستقل و هم‌توزیع<sup>۲۲</sup> (i.i.d.) با میانگین صفر و واریانس یک است. می‌توان یک مدل سری زمانی  $ARMA(p, q)$  با خطای  $GARCH(r, s)$  را با استفاده از رابطه‌های (۲۲.۲) تا (۲۴.۲) نیز مشخص نمود.

$$X_t = c + a_t + \sum_{i=1}^p \phi_i x_{t-i} + \sum_{i=1}^q \theta_i a_{t-i} \quad (22.2)$$

$$a_t = \sigma_t \epsilon_t \quad (23.2)$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^r \alpha_i a_{t-i}^2 + \sum_{j=1}^s \beta_j \sigma_{t-j}^2 \quad (24.2)$$

بدیهی است با تغییر مقادیر  $r, s, p, q$  می‌توان گستره بزرگی از مدل‌های ساده‌تر را برای اهداف مدل‌بندی به دست آورد.

## ۳.۲ انتخاب پارامترهای مدل‌های پیش‌بینی

تشخیص اینکه آیا یک مدل برای سری زمانی مناسب است یا خیر، مشکل است. در عمل برای ارزیابی مدل‌ها، بعد از برازش آن مدل‌ها به مشاهدات سری زمانی و تعیین مرتبه مدل و ضرایب آن، از معیارهای AIC و BIC استفاده می‌شود. این دو معیار به شکل زیر تعیین می‌گردند.

$$AIC = -2LLF + 2NumParams \quad (25.2)$$

$$BIC = -2LLF + \log(NumObs)NumParams \quad (26.2)$$

که در آن NumParams تعداد پارامترهای مدل است، همچنین NumObs برابر با تعداد مشاهدات سری زمانی می‌باشد و LLF تابع درست‌نمایی است. بعد از محاسبه این معیارها برای مدل‌های مختلف، مدلی انتخاب می‌شود که دارای مقدار AIC و BIC کمینه باشد. نکته قابل توجه این است که این دو معیار همواره منجر به انتخاب یک مدل منحصر بفرد نمی‌شوند. همان‌طور که در فرمول‌های دقیق مشاهده می‌شود، میزان حساسیت معیار BIC به تعداد پارامترهای مدل بیشتر از معیار AIC می‌باشد و در نتیجه تعداد پارامترهای مدل انتخابی معیار BIC همواره کوچکتر یا مساوی مدل انتخابی AIC خواهد بود. برای انتخاب مدل ترکیبی ARMA-GARCH از معیارهای AIC و BIC استفاده می‌شود. از آنجایی که در اکثر سری‌های بازده، مرتبه‌های پایین مدل GARCH ناپایداری سری را به خوبی مدل می‌کنند.

**مثال ۱.۳.۲.** در این مثال، VaR بر اساس ARMA-GARCH برازش و پیش‌بینی می‌شود. بدین منظور، ابتدا مجموعه داده‌ای از مدل  $ARMA(1, 1) - GARCH(1, 1)$  با پارامترهای  $\phi_0 = 0$

<sup>22</sup>Independent and identical distribution

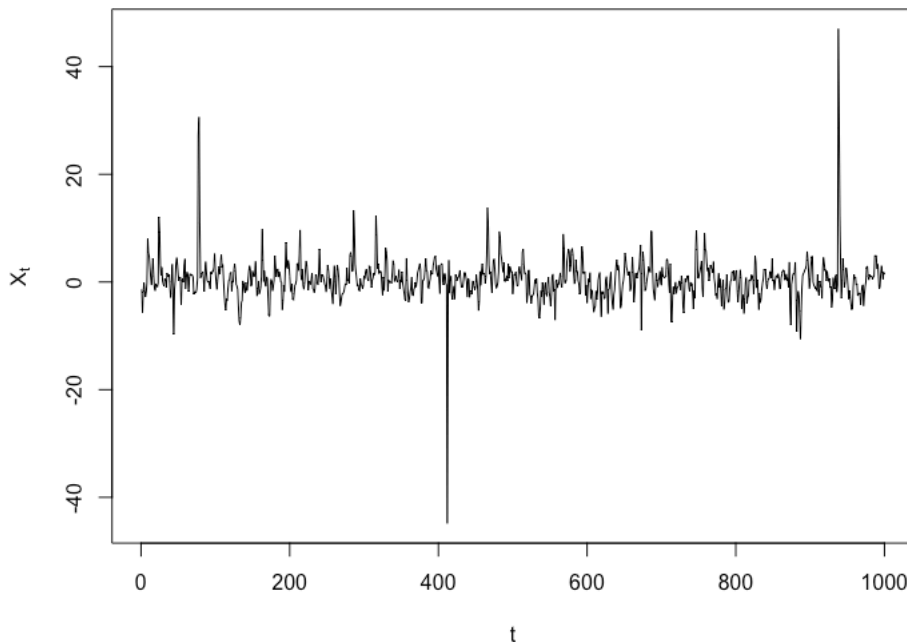
خطاها با استفاده از کد دستوری زیر در نرم‌افزار R تولید می‌شود.  
 $\beta_1 = 0.2$  و  $\alpha_1 = 0.4$ ،  $\alpha_0 = 4$ ،  $\theta_1 = 0.3$ ،  $\phi_1 = 0.5$  استودنت  $t$  - توزیع

```
library(rugarch)
library(qrmtools)
## Model specification (for simulation)
nu <- 3 # d.o.f. of the standardized distribution of Z_t
fixed.p <- list(mu = 0, # our mu (intercept)
ar1 = 0.5, # our phi_1 (AR(1) parameter of mu_t)
ma1 = 0.3, # our theta_1 (MA(1) parameter of mu_t)
omega = 4, # our alpha_0 (intercept)
alpha1 = 0.4, # our alpha_1 (GARCH(1) parameter of sigma_t^2)
beta1 = 0.2, # our beta_1 (GARCH(1) parameter of sigma_t^2)
shape = nu) # d.o.f. nu for standardized t_nu innovations
armaOrder <- c(1,1) # ARMA order
garchOrder <- c(1,1) # GARCH order
varModel <- list(model = "sGARCH", garchOrder = garchOrder)
spec <- ugarchspec(varModel, mean.model = list(armaOrder = armaOrder),
fixed.pars = fixed.p, distribution.model = "std") # t standardized residuals
## Simulate (X_t)
n <- 1000 # sample size (= length of simulated paths)
x <- ugarchpath(spec, n.sim = n, m.sim = 1, rseed = 271) # n.sim length of
#simulated path; m.sim = number of paths
## Note the difference:
## - ugarchpath(): simulate from a specified model
## - ugarchsim(): simulate from a fitted object
## Extract the resulting series
X <- fitted(x) # simulated process X_t = mu_t + epsilon_t for
#epsilon_t = sigma_t * Z_t
sig <- sigma(x) # volatilities sigma_t (conditional standard deviations)
eps <- x@path$residSim # unstandardized residuals epsilon_t = sigma_t * Z_t
## Note: There are no extraction methods for the unstandardized residuals
## epsilon_t
## Sanity checks (=> fitted() and sigma() grab out the right quantities)
```

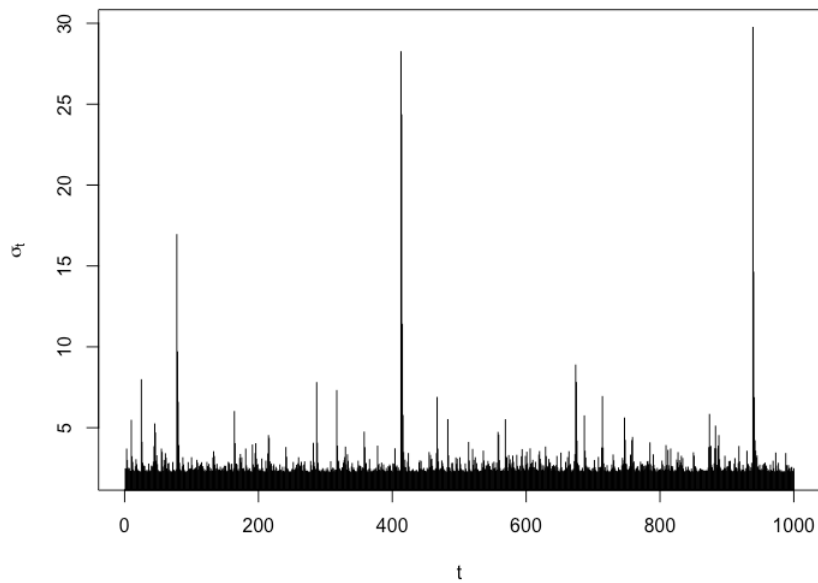
```
stopifnot(all.equal(X, x@path$seriesSim, check.attributes = FALSE),
all.equal(sig, x@path$sigmaSim, check.attributes = FALSE))
```

به منظور یک بررسی منطقی و دقیق، داده‌های شبیه‌سازی شده، انحراف استاندارد شرطی و باقی‌مانده‌ها با استفاده از کدهای زیر در ادامه رسم شده است. شکل‌های ۳.۲، ۴.۲ و ۵.۲ به ترتیب نمودارهای سری زمانی، انحراف معیار شرطی و باقی‌مانده‌های داده‌های شبیه‌سازی شده از مدل  $ARMA(1,1) - GARCH(1,1)$  را نشان می‌دهد.

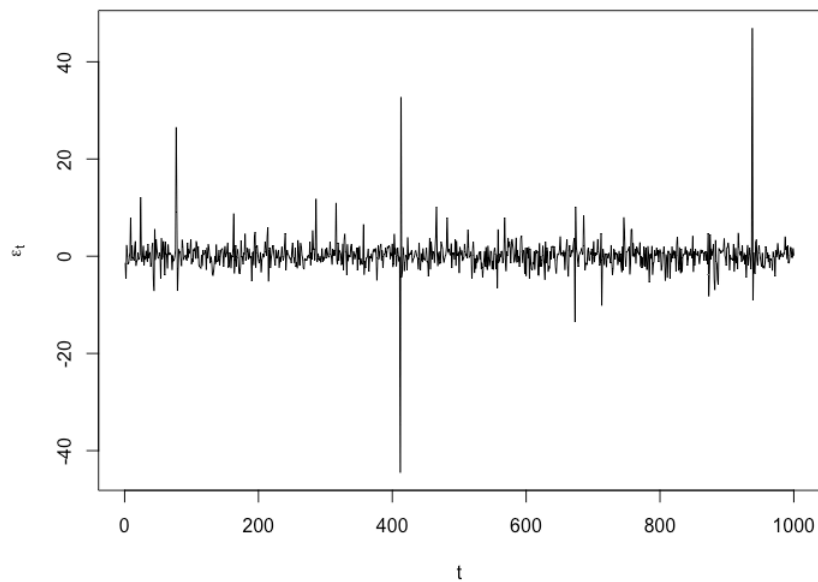
```
## Plots
#Time Series Plot
plot(X, type = "l", xlab = "t", ylab = expression(X[t]))
#Conditional Standard Deviations
plot(sig, type = "h", xlab = "t", ylab = expression(sigma[t]))
#Residuals
plot(eps, type = "l", xlab = "t", ylab = expression(epsilon[t]))
```



شکل ۳.۲: نمودار سری زمانی داده‌های شبیه‌سازی شده از مدل  $ARMA(1,1) - GARCH(1,1)$



شکل ۴.۲: نمودار انحراف معیار شرطی داده‌های شبیه‌سازی شده از مدل  $ARMA(1, 1)$  -  $GARCH(1, 1)$



شکل ۵.۲: نمودار باقی‌مانده‌های داده‌های شبیه‌سازی شده از مدل  $ARMA(1, 1)$  -  $GARCH(1, 1)$

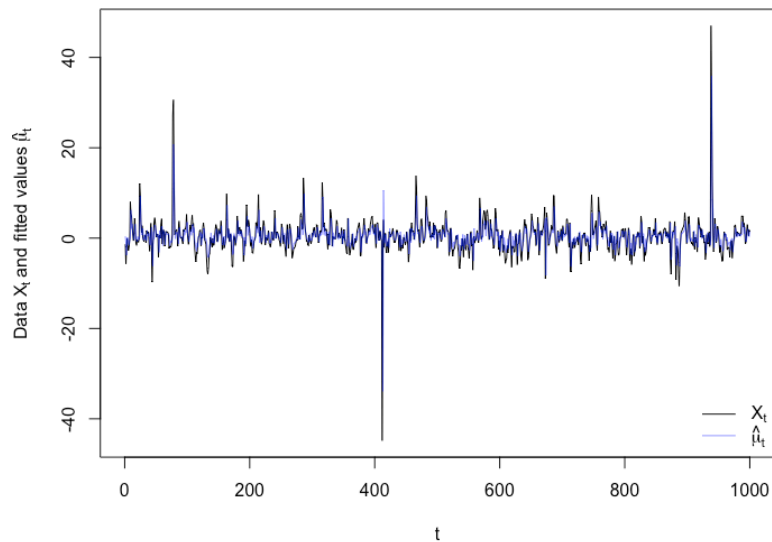
یک فرآیند ARMA-GARCH با مرتبه‌های صحیح و مشخص بر اساس کد زیر روی داده‌ها برازش شده است اما در حالت کلی، باید مدل با مرتبه‌های مختلف به داده‌ها برازش داده شود و بر اساس معیارهایی مانند AIC، BIC تصمیم گرفته شود از چه مرتبه‌ای استفاده شود.

```
## Fit an ARMA(1,1)-GARCH(1,1) model
spec <- ugarchspec(varModel, mean.model = list(armaOrder = armaOrder),
distribution.model = "std") # without fixed parameters here
fit <- ugarchfit(spec, data = X) # fit
## Extract the resulting series
mu. <- fitted(fit) # fitted  $\hat{\mu}_t (= \hat{X}_t)$ 
sig. <- sigma(fit) # fitted  $\hat{\sigma}_t$ 
## Sanity checks (=> fitted() and sigma() grab out the right quantities)
stopifnot(all.equal(as.numeric(mu.), fit@fit$fitted.values),
all.equal(as.numeric(sig.), fit@fit$sigma))
```

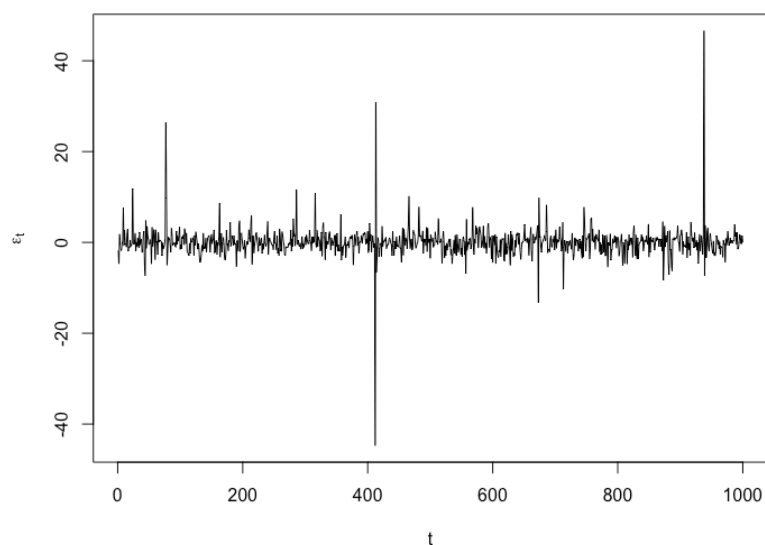
برای بررسی دقیق‌تر نتایج، نمودارهای داده‌ها و مقادیر پیش‌بینی شده (شکل ۶.۲)، باقی‌مانده‌های استاندارد نشده (شکل ۷.۲) و نمودار چندک-چندک باقی‌مانده‌های استاندارد شده در برابر توزیع  $t$ -استودنت مشخص (شکل ۸.۲) به کمک کدهای زیر رسم شده است.

```
## Plot data  $X_t$  and fitted  $\hat{\mu}_t$ 
plot(X, type = "l", xlab = "t",
ylab = expression("Data"~X[t]~"and fitted values"~hat(mu)[t]))
lines(as.numeric(mu.), col = adjustcolor("blue", alpha.f = 0.5))
legend("bottomright", bty = "n", lty = c(1,1),
col = c("black", adjustcolor("blue", alpha.f = 0.5)),
legend = c(expression(X[t]), expression(hat(mu)[t])))
## Plot the unstandardized residuals  $\epsilon_t$ 
resi <- as.numeric(residuals(fit))
stopifnot(all.equal(fit@fit$residuals, resi))
# check residuals  $\epsilon_t$ 
plot(resi, type = "l", xlab = "t", ylab = expression(epsilon[t]))
## Q-Q plot of the standardized residuals  $Z_t$  against their specified  $t$ 
## ( $t_{\nu}$  with variance 1)
Z <- fit@fit$z
stopifnot(all.equal(Z, as.numeric(resi/sig.)))
```

```
qq_plot(Z, FUN = function(p) sqrt((nu-2)/nu) * qt(p, df = nu),
main = substitute("Q-Q plot of (*Z[t]*) against a standardized~t[nu.],
list(nu. = nu)))
```

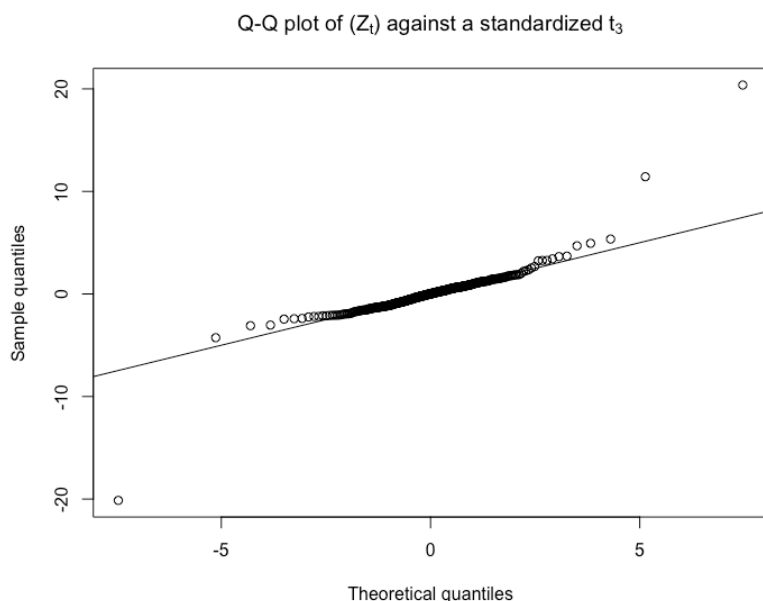


شکل ۶.۲: نمودار داده‌های شبیه‌سازی شده از مدل  $ARMA(1, 1) - GARCH(1, 1)$  و مقادیر برازش شده آن‌ها



شکل ۷.۲: نمودار باقی‌مانده‌های استاندارد نشده از مدل  $ARMA(1, 1) - GARCH(1, 1)$





شکل ۸.۲: نمودار چندک-چندک باقی‌مانده‌های استاندارد شده مدل  $ARMA(1,1)$  -  $GARCH(1,1)$  در برابر توزیع  $t$  با ۳ درجه آزادی.

در ادامه، سری‌های زمانی VaR محاسبه می‌شود.

```
## VaR confidence level we consider here
alpha <- 0.99
## Extract fitted VaR_alpha
VaR. <- as.numeric(quantile(fit, probs = alpha))
## Build manually and compare the two
nu. <- fit@fit$coef["shape"] # extract (fitted) d.o.f. nu
# VaR_alpha computed manually
VaR.. <- as.numeric(mu. + sig. * sqrt((nu.-2)/nu.) * qt(alpha, df = nu.))
stopifnot(all.equal(VaR., VaR..))
## => quantile(<rugarch object>, probs = alpha) provides VaR_alpha =
##  $\hat{\mu}_t + \hat{\sigma}_t * q_Z(\alpha)$ 
```

اکنون VaR بر اساس مدل برازش شده به کمک دستورهایی زیر می‌توان پیش‌بینی کرد.

```
## Predict from the fitted process
fspec <- getspec(fit) # specification of the fitted process
setfixed(fspect) <- as.list(coef(fit)) # set the parameters to the fitted ones
m <- ceiling(n / 10) # number of steps to forecast (roll/iterate m-1 times)
```

```
forward with frequency 1)
#predict from the fitted process
pred <- ugarchforecast(fspect, data = X, n.ahead = 1, n.roll = m-1, out.sample = m)
## Extract the resulting series
# extract predicted X_t (= conditional mean mu_t; note: E[Z] = 0)
mu.predict <- fitted(pred)
sig.predict <- sigma(pred) # extract predicted sigma_t
# corresponding predicted VaR_alpha
VaR.predict <- as.numeric(quantile(pred, probs = alpha))
## Checks
## Sanity checks
stopifnot(all.equal(mu.predict, pred@forecast$seriesFor, check.attributes = FALSE),
# sanity check
all.equal(sig.predict, pred@forecast$sigmaFor, check.attributes = FALSE))
## Build predicted VaR_alpha manually and compare the two
VaR.predict. <- as.numeric(mu.predict + sig.predict * sqrt((nu.-2)/nu.) *
qt(alpha, df = nu.)) # VaR_alpha computed manually
stopifnot(all.equal(VaR.predict., VaR.predict))
```

در انتها، جمله‌های بعدی از فرآیند با استفاده از کدهای زیر شبیه‌سازی شده‌اند و برای هر یک از آن‌ها، VaR برآورد می‌شود. از دستور `quantile()` نمی‌توان در این بسته استفاده کرد بنابراین باید از روش‌های باز نمونه‌گیری مانند بوت‌استرپ برای ساختن بازه‌های اطمینان VaR استفاده کرد.

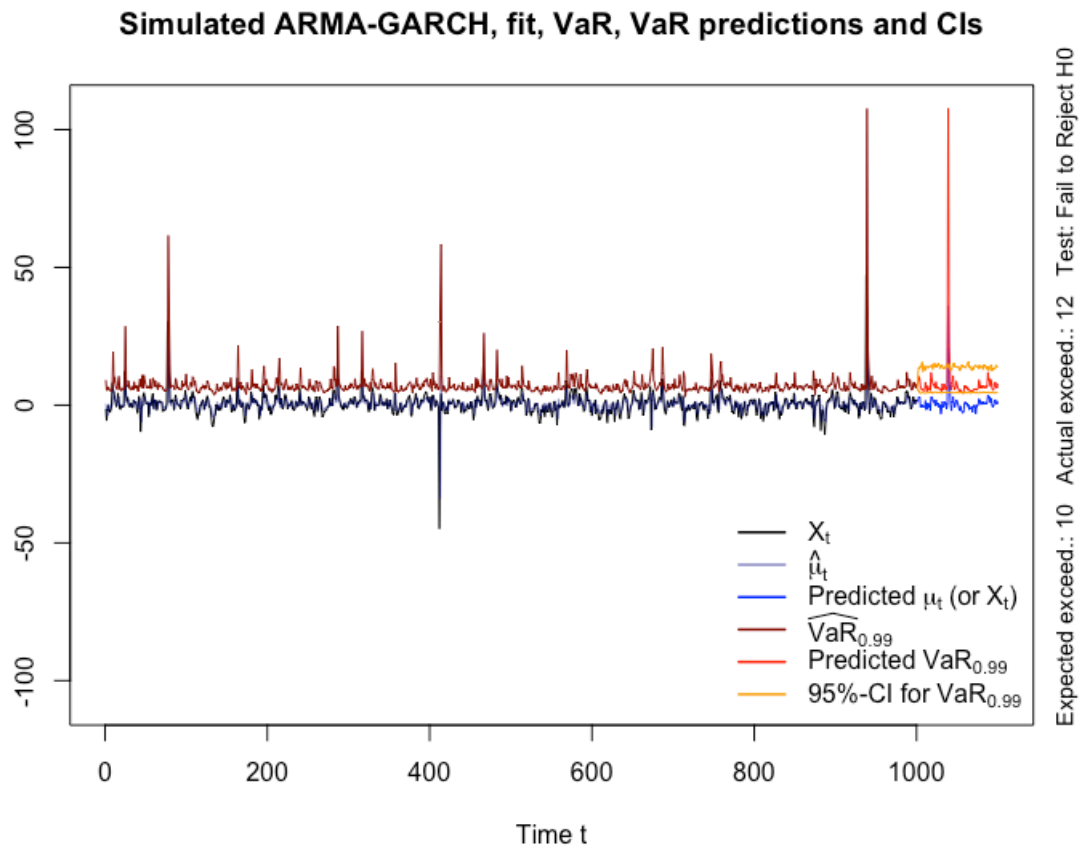
```
## Simulate B paths
B <- 1000
set.seed(271)
X.sim.obj <- ugarchpath(fspect, n.sim = m, m.sim = B) # simulate future paths
## Compute simulated VaR_alpha and corresponding (simulated) confidence
## intervals
## Note: Each series is now an (m, B) matrix
##(each column is one path of length m)
X.sim <- fitted(X.sim.obj) # extract simulated X_t
sig.sim <- sigma(X.sim.obj) # extract sigma_t
eps.sim <- X.sim.obj@path$residSim # extract epsilon_t
```

```
# (m, B) matrix
VaR.sim <- (X.sim - eps.sim) + sig.sim * sqrt((nu.-2)/nu.) * qt(alpha, df = nu.)
VaR.CI <- apply(VaR.sim, 1, function(x) quantile(x, probs = c(0.025, 0.975)))
```

سرانجام، همه نتایج به‌دست آمده با استفاده از کدهای زیر در شکل ۹.۲ نمایش داده شده است.

```
## Setup
yran <- range(X, # simulated path
mu., VaR., # fitted conditional mean and VaR_alpha
mu.predict, VaR.predict, VaR.CI) # predicted mean, VaR and CIs
myran <- max(abs(yran))
yran <- c(-myran, myran) # y-range for the plot
xran <- c(1, length(X) + m) # x-range for the plot
## Simulated (original) data (X_t), fitted conditional mean mu_t
## and VaR_alpha
plot(X, type = "l", xlim = xran, ylim = yran, xlab = "Time t", ylab = "",
main = "Simulated ARMA-GARCH, fit, VaR, VaR predictions and CIs")
# hat{\mu}_t
lines(as.numeric(mu.), col = adjustcolor("darkblue", alpha.f = 0.5))
lines(VaR., col = "darkred") # estimated VaR_alpha
mtext(paste0("Expected exceed.: ", btest$expected.exceed, " ",
"Actual exceed.: ", btest$actual.exceed, " ",
"Test: ", btest$cc.Decision),
side = 4, adj = 0, line = 0.5, cex = 0.9) # label
## Predictions
t. <- length(X) + seq_len(m) # future time points
lines(t., mu.predict, col = "blue") # predicted process X_t (or mu_t)
lines(t., VaR.predict, col = "red") # predicted VaR_alpha
lines(t., VaR.CI[1,], col = "orange") # lower 95%-CI for VaR_alpha
lines(t., VaR.CI[2,], col = "orange") # upper 95%-CI for VaR_alpha
legend("bottomright", bty = "n", lty = rep(1, 6), lwd = 1.6,
col = c("black", adjustcolor("darkblue", alpha.f = 0.5), "blue",
"darkred", "red", "orange"),
legend = c(expression(X[t]), expression(hat(mu)[t]),
```

```
expression("Predicted"~mu[t]~"(or"~X[t]*")"),
substitute(widehat(VaR)[a], list(a = alpha)),
substitute("Predicted"~VaR[a], list(a = alpha)),
substitute("95%-CI for"~VaR[a], list(a = alpha)))
```



شکل ۹.۲: نمودار داده‌های شبیه‌سازی شده از مدل  $ARMA(1,1) - GARCH(1,1)$ ، مقادیر برازش شده، VaR، VaR پیش‌بینی شده و فاصله‌های اطمینان



## فصل ۳

# روش‌های انتخاب متغیر در رگرسیون سری زمانی

در این فصل، برخی از روش‌های انتخاب متغیر در رگرسیون سری زمانی بررسی می‌شوند. بدین منظور، ابتدا به‌طور مختصر، رگرسیون سری زمانی را معرفی و سپس به مرور برخی از روش‌های انتخاب متغیر مدرن در خانواده رگرسیون جریمه‌شده مانند روش عملگر انتخاب‌کننده کمترین قدر مطلق و انقباضی<sup>۱</sup> (LASSO) پرداخته می‌شود. (یون و همکاران، ۲۰۱۷؛ ژانگ، ۲۰۱۷)

### ۱.۳ مدل‌های خطی

تحلیل رگرسیون، روش آماری برای بررسی و مدل‌سازی ارتباط بین متغیرهاست. کاربردهای رگرسیون متعدد است و تقریباً در هر زمینه‌ای از جمله مهندسی، فیزیک، اقتصاد، مدیریت، علوم زیستی، بیولوژی و علوم اجتماعی به کار می‌رود. در حقیقت تحلیل رگرسیونی ممکن است روش آماری با بیشترین و وسیع‌ترین کاربرد بین روش‌های آماری باشد. مدل‌های رگرسیونی برای اهدافی مانند: (۱) توصیف داده‌ها، (۲) برآورد پارامترها و (۳) پیش‌گویی و برآورد مورد استفاده قرار می‌گیرند.

<sup>۱</sup>Least Absolute Shrinkage and Selection Operator (LASSO)

ایده اصلی رگرسیون، تشخیص رابطه بین چند متغیر است. ساختار آن بدین صورت است که متغیر وابسته یا پاسخ ( $Y$ ) با یک یا چند متغیر مستقل یا توضیحی یا پیش‌گو ( $X_1, \dots, X_p$ ) مرتبط می‌باشد. لذا داده‌ها به صورت مجموعه  $\{(x_1, y_1), \dots, (x_n, y_n)\}$  در دسترس هستند که  $x_i = (x_{i1}, \dots, x_{ip})^T$  رگرسیون خطی یکی از گسترده‌ترین تکنیک‌های مدل‌سازی است. ماهیت مدل رگرسیون، به صورت خطی است و برای مشاهده  $i$ ام ( $i = 1, \dots, n$ ) می‌توان رابطه

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_p x_{ip} = \beta_0 + \mathbf{x}_i^T \boldsymbol{\beta} + \varepsilon_i \quad (1.3)$$

را نوشت. در حالت ماتریسی، مدل خطی

$$\mathbf{y}_n = \beta_0 \mathbf{1}_n + \mathbf{X}_n \boldsymbol{\beta} + \boldsymbol{\varepsilon}_n$$

نتیجه می‌شود که در آن

$$\mathbf{y}_n = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \mathbf{X}_n = \begin{bmatrix} \mathbf{x}_1^T \\ \mathbf{x}_2^T \\ \vdots \\ \mathbf{x}_n^T \end{bmatrix} = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1p} \\ x_{21} & x_{22} & \dots & x_{2p} \\ \vdots & \vdots & \vdots & \vdots \\ x_{n1} & x_{n2} & \dots & x_{np} \end{bmatrix}, \boldsymbol{\beta} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_p \end{bmatrix}, \boldsymbol{\varepsilon}_n = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}.$$

عدد نامعلوم  $\beta_0$  عرض از مبدا،  $\mathbf{1}_n = (1, \dots, 1)^T$  یک بردار  $n$  بعدی، ماتریس طرح،  $\boldsymbol{\beta}$ ، بردار پارامترها و  $\boldsymbol{\varepsilon}_n$ ، بردار خطاست.

رگرسیون سری زمانی شبیه رگرسیون معمولی است و تفاوت آن در این است که در رگرسیون سری‌های زمانی، مشاهدات در یک دنباله از زمان‌های مشخصی ثبت شده‌اند و بنابراین، اندیس  $i$  در رگرسیون معمولی با اندیس  $t$  در رگرسیون سری زمانی جانشین می‌شود و طبیعتاً تعداد مشاهدات در داده‌های سری زمانی،  $T$  خواهد بود که جایگزین  $n$  تعداد مشاهدات در رگرسیون معمولی است. بنابراین، رابطه (۱.۳) در رگرسیون سری زمانی به صورت

$$y_t = \beta_0 + \beta_1 x_{1t} + \beta_2 x_{2t} + \dots + \beta_p x_{pt} = \beta_0 + \mathbf{x}_t^T \boldsymbol{\beta} + \varepsilon_t, \quad t = 1, \dots, T \quad (2.3)$$

تبدیل می‌شود. در حالت ماتریسی، مدل خطی

$$\mathbf{y}_t = \beta_0 \mathbf{1}_t + \mathbf{X}_t \boldsymbol{\beta} + \boldsymbol{\varepsilon}_t$$

نتیجه می‌شود که در آن

$$\mathbf{y}_t = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_T \end{bmatrix}, \mathbf{X}_t = \begin{bmatrix} \mathbf{x}_1^T \\ \mathbf{x}_2^T \\ \vdots \\ \mathbf{x}_T^T \end{bmatrix} = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1p} \\ x_{21} & x_{22} & \dots & x_{2p} \\ \vdots & \vdots & \vdots & \vdots \\ x_{1T} & x_{2T} & \dots & x_{pT} \end{bmatrix}, \boldsymbol{\beta} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_p \end{bmatrix}, \boldsymbol{\varepsilon}_T = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_T \end{bmatrix}.$$

عدد نامعلوم  $\beta_0$  عرض از مبدا،  $\mathbf{1}_n = (1, \dots, 1)^\top$  یک بردار  $T$  بعدی، ماتریس  $X_t$ ، ماتریس طرح،  $\beta$ ، بردار پارامترها و  $\varepsilon_t$ ، بردار خطاست که میانگین و واریانس آن به ترتیب بردار  $p$  بعدی  $o$  و ماتریس  $\sigma^2 I_p$  است. دقت کنید که ماتریس طرح  $X_t$ ،  $T$  سطر، تعداد مشاهده‌ها و  $p$  ستون، تعداد متغیرها است.

مفهوم پایه‌ای در رگرسیون سری زمانی این است که متغیر سری زمانی  $Y_t$  با سایر متغیرهای سری زمانی مانند  $X_t$  ارتباط دارد. به عنوان مثال، می‌توان به پیش‌بینی فروش ماهانه ( $Y_t$ ) بر اساس هزینه‌های صرف‌شده در تبلیغات ( $x_t$ ) یا مصرف برق روزانه ( $Y_t$ ) بر اساس دما ( $X_{1t}$ ) و روز هفته ( $X_{2t}$ ) اشاره نمود.

مدل رگرسیون معمولی، به چهار پذیره پایه‌ای نیاز دارد: (۱) صورت مدل، (۲) استقلال باقی‌مانده‌ها، (۳) ثابت بودن واریانس و (۴) نرمال بودن باقی‌مانده‌ها. فرض کنید هر یک از پذیره‌ها برای مجموعه داده‌های مورد بررسی، معتبر باشد، در صورتی مدل رگرسیون مدل خوبی است که پیش‌بینی‌های مفیدی را ارائه دهد. هنگامی که خطاها، هر نوع الگویی را نشان دهند، آن مدل همه اطلاعات مفید در مجموعه داده را در بر ندارد. درک هر پذیره، نیازمند فهم دقیق نظریه آماری رگرسیون است.

در رگرسیون سری زمانی، باید پذیره‌های دیگری نیز مورد توجه قرار گیرند: (۱) ممکن است باقی‌مانده‌ها مستقل نباشند، (۲) ممکن است به تاثیرهای وابسته به زمان مانند روند، فصلی بودن یا تغییرات روزهای کاری نیاز باشد و (۳) متغیر  $Y_t$  باید پس از پیش‌بینی هر یک از متغیرهای پیشگوی  $X_t$ ، پیش‌بینی شود. برای بررسی بیشتر، خواننده علاقه‌مند می‌تواند **ماکریداکیس و همکاران (۱۳۹۴)** را ببیند.

مدل‌های رگرسیون خطی با این فرض که داده‌ها مستقل و هم‌توزیع‌اند، به عنوان ابزاری برای تحلیل رابطه بین متغیرهای پاسخ و پیشگو استفاده می‌شود. اما چنین ادعایی زمانی که داده‌ها در طول زمان به دست می‌آیند، نقض و بین مشاهدات همبستگی پیاپی ظاهر می‌شود. به منظور بازتاب آن بسیاری از نویسندگان از جمله **تسای (۱۹۸۴)**، **س.ا. گلاسی (۱۹۸۸)**، **شی و تسای (۲۰۰۴)**، **وانگ و همکاران (۲۰۰۷)**، مدل‌های رگرسیون خطی را با خطاهای AR و ARMA در نظر گرفتند.

### ۱.۱.۳ برآورد ضرایب مدل

در مسائل کاربردی، فقط مجموعه‌ای از مشاهده‌ها در دسترس است و هیچ اطلاعی از ضرایب نامعلوم  $\beta_0, \beta_1, \dots, \beta_p$  وجود ندارد که باید بر اساس داده‌ها برآورد شوند. یکی از معروف‌ترین روش‌های برآورد ضرایب نامعلوم (پارامترها)، روش کمترین توان‌های دوم معمولی<sup>۲</sup> است.

اصل کمترین توان‌های دوم، روش موثری برای انتخاب ضرایب را بر اساس کمینه کردن مجموع توان‌های دوم خطا ارائه می‌کند. بدین منظور، ضرایب  $\beta_0, \beta_1, \dots, \beta_p$  به گونه‌ای

<sup>۲</sup>Ordinary Least Square



انتخاب می‌شوند که عبارت

$$\begin{aligned}\varepsilon_t^\top \varepsilon_t &= \sum_{t=1}^T \varepsilon_t^2 \\ &= \sum_{i=1}^p (y_t - \beta_0 - \beta_1 x_{1t} - \beta_2 x_{2t} - \dots - \beta_p x_{pt})^2 \\ &= (\mathbf{y}_t - \mathbf{X}_t \boldsymbol{\beta})^\top (\mathbf{y}_t - \mathbf{X}_t \boldsymbol{\beta})\end{aligned}$$

را کمینه می‌کند. برای یافتن کمینه‌کننده  $\varepsilon_t^\top \varepsilon_t$ ، باید مشتق آن نسبت به  $\boldsymbol{\beta}$  را برابر صفر قرار داد و معادله را حل نمود. به عبارت دیگر

$$\frac{\partial}{\partial \boldsymbol{\beta}} \varepsilon_t^\top \varepsilon_t = \frac{\partial}{\partial \boldsymbol{\beta}} (\mathbf{y}_t - \mathbf{X}_t \boldsymbol{\beta})^\top (\mathbf{y}_t - \mathbf{X}_t \boldsymbol{\beta}) = -2 \mathbf{X}_t^\top \mathbf{y} + 2 \mathbf{X}_t^\top \mathbf{X}_t \boldsymbol{\beta} = 0$$

که نتیجه می‌دهد

$$\mathbf{X}_t^\top \mathbf{y} = (\mathbf{X}_t^\top \mathbf{X}_t) \boldsymbol{\beta}$$

بنابراین

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}_t^\top \mathbf{X}_t)^{-1} \mathbf{X}_t^\top \mathbf{y}. \quad (3.3)$$

با توجه به اینکه،  $\hat{\boldsymbol{\beta}}$  کمترین مقدار مجموع توان‌های دوم خطا را نتیجه می‌دهد، آن را برآوردگر کمترین توان‌های دوم می‌نامند. از رابطه (۳.۳) مشخص است که در صورتی برآوردگر OLS وجود دارد که ماتریس  $\mathbf{X}_t^\top \mathbf{X}_t$  معکوس‌پذیر باشد. اگر ماتریس طرح  $\mathbf{X}_t$ ، ماتریسی با رتبه کامل ستونی نباشد، معکوس ماتریس  $\mathbf{X}_t^\top \mathbf{X}_t$  وجود ندارد و نمی‌توان برآوردگر را به دست آورد. واریانس باقی‌مانده‌ها به صورت رابطه (۴.۳) برآورد می‌شود.

$$\hat{\sigma}^2 = \frac{1}{T - p - 1} (\mathbf{y}_t - \mathbf{X}_t \hat{\boldsymbol{\beta}})^\top (\mathbf{y}_t - \mathbf{X}_t \hat{\boldsymbol{\beta}}). \quad (4.3)$$

در نرم‌افزار R، به کمک تابع `tslm` می‌توان رگرسیون سری زمانی را به داده‌های سری زمانی برازش داد. این تابع، مشابه تابع `lm` در مدل‌های رگرسیون معمولی است.

از ویژگی‌های برآوردگر کمترین توان‌های دوم می‌توان به موارد زیر اشاره کرد: (۱)  $\hat{\boldsymbol{\beta}}$ ، برآوردگری نااریب با کمترین واریانس است؛ (۲) واریانس  $\hat{\boldsymbol{\beta}}$  به صورت  $V(\hat{\boldsymbol{\beta}}) = \hat{\sigma}^2 (\mathbf{X}^\top \mathbf{X})^{-1}$  است؛ (۳) میانگین توان دوم خطا<sup>۳</sup> (MSE) نیز به صورت  $\text{MSE}(\hat{\boldsymbol{\beta}}) = \hat{\sigma}^2 \text{tr}(\mathbf{X}^\top \mathbf{X})^{-1}$  که در آن `tr`، اثر ماتریس است؛ (۴) این برآوردگر در حالتی که توزیع خطاها نرمال باشد، مشابه برآوردگری است که به روش ماکسیمم درست‌نمایی به دست می‌آید و بنابراین، برآوردگری سازگار است؛ (۵) برآوردگر OLS، برآوردگر کارا است.

هنگامی که یک مدل رگرسیونی خطی به مجموعه‌ای از داده‌ها برازش داده می‌شود، ممکن است مسائلی بروز پیدا کند که برآوردگر OLS به شدت تحت تاثیر آن‌ها قرار می‌گیرد که می‌توان

<sup>3</sup>Mean Square Error (MSE)

به موارد (۱) بعد بالا<sup>۴</sup> (تعداد متغیرها از تعداد مشاهده‌ها بیشتر باشد که ماتریس  $X_t^T X_t$  معکوس پذیر نیست)؛ (۲) هم‌خطی<sup>۵</sup>؛ (دو یا بیش از دو متغیر مستقل به یک متغیر مستقل دیگر وابسته هستند) (۳) نقاط پرت<sup>۶</sup> (داده‌های پرت، داده‌هایی هستند که متفاوت از سایر داده‌ها است) و (۴) نقاط اهرمی<sup>۷</sup> (داده پرت در فضای  $X_t$  که آنها را نقاط اهرمی می‌نامند) اشاره کرد.

## ۲.۳ انتخاب متغیرها

در واقعیت، ساختن یک مدل رگرسیونی برای داده‌ها هیچ‌گاه فرآیند ساده‌ای نیست. مسئله مورد بررسی هرچه باشد، باید از متخصصین در آن رشته برای مشورت استفاده کرد. برای متغیر پاسخ  $Y_t$ ، لازم است بر اساس (۱) نظر متخصصان و افراد خبره، (۲) دردسترس بودن داده‌ها و (۳) محدودیت‌های زمانی و مکانی، لیستی از تمام متغیرهایی که بر آن موثر هستند، تهیه شود. لزوماً استفاده از تمام متغیرهای لیست شده ضروری نیست و معمولاً به منظورهای مختلفی با روش‌های متعددی می‌تواند به تعداد کمتری کاهش پیدا کند.

انتخاب متغیر، یکی از مهم‌ترین موضوعات در مدل‌سازی آماری محسوب می‌شود که کاربرد وسیعی در تحلیل‌های آماری دارد. بسیاری از روش‌های انتخاب متغیر با استفاده از مدل رگرسیون خطی انجام می‌شود. محققان، پیشنهاد‌های بسیاری را در مورد انتخاب متغیرهای مناسب برای مدل نهایی مطرح کرده‌اند. برخی اشاره کرده‌اند که نمودار متغیر پاسخ در برابر هر یک از متغیرهای پیش‌گو رسم شود و اگر رابطه قابل ملاحظه‌ای مشاهده نشد، آن را حذف کنند یا یک رگرسیون خطی چندگانه را روی همه متغیرهای توضیحی انجام داده و سپس آن متغیرهایی که مقدار آماره آزمون آن‌ها خیلی کوچک است، از مدل حذف شوند. با وجودی که استفاده از این رهیافت‌ها ساده است اما استفاده از آن‌ها منجر به یافتن یک مدل رگرسیون خوب و قابل‌اعتماد نمی‌شود.

یک روش ساده و بدیهی برای یافتن متغیرهای موثر، انتخاب یک زیرمجموعه از متغیرهای توضیحی با اهمیت و تاثیرگذار در مدل است. انتخاب این متغیرها می‌تواند بر اساس معنی‌داری آنها در مدل یا برخی از معیارهای اطلاع‌مانند AIC (آکاییک، ۱۹۷۴)، BIC (شوارتز، ۱۹۷۸) و ... انجام شود. اما مشکل بزرگی که در هنگام استفاده از این رهیافت مطرح می‌شود تعداد کل زیرمجموعه‌هایی که از متغیرها باید در نظر گرفته شوند به‌طور نمایی با افزایش  $p$  رشد می‌کند و از نظر محاسباتی برای تعداد متغیرهای بیشتر از ۴۰ یا ۵۰، تقریباً نشدنی است. مشکل اساسی دیگر این روش‌ها، ویژگی گسسته بودن آنهاست. به عبارت دیگر، تغییر بسیار

<sup>4</sup>High-dimensional

<sup>5</sup>Collinearity

<sup>6</sup>Outliers

<sup>7</sup>High-leverage points

کوچکی در داده‌ها، می‌تواند برآوردهای کاملاً متفاوتی را نتیجه دهد. در مجموع، روش‌های انتخاب بهترین زیرمجموعه، اغلب ناپایدار و به شدت تغییرپذیر هستند. از دیگر روش‌های انتخاب متغیر می‌توان به روش رگرسیون گام به گام اشاره کرد که به تفصیل می‌توان در **ماکریداکیس و همکاران (۱۳۹۴)**، فصل ۷ مطالعه کرد.

روش دیگر، برای رویارویی با این مسئله که در سال‌های اخیر کاربردی شده است، استفاده از رگرسیون جریمه شده<sup>۸</sup> است. هم‌چنین در بعضی از مواقع که با مسائلی همچون بعد بالا و یا هم‌خطی مواجه می‌شویم استفاده از رگرسیون معمولی، ناکارآمد بوده و برآورد حاصل از این روش مناسب نخواهد بود. در چنین شرایطی نیز از رگرسیون جریمه شده استفاده می‌شود که به نوعی به مسئله انتخاب متغیر اشاره می‌کند. در حالت کلی در رگرسیون خطی جریمه شده، هدف کمینه کردن تابع زیر است

$$\phi(\beta) = (\mathbf{y}_n - \mathbf{X}_n\beta)^\top (\mathbf{y}_n - \mathbf{X}_n\beta) + \lambda P(\beta). \quad (۵.۳)$$

با دقت در ساختار تابع (۵.۳)، استنباط می‌شود که به تابع توان دوم خطا در رگرسیون معمولی، عبارت جریمه  $P(\beta)$  اضافه شده است. در واقع با جریمه کردن پارامتر، سعی شده است مقدار خطا کمینه شود. این عبارت جریمه، تابعی از پارامترهاست و بنا به روش رگرسیون جریمه شده، شکل‌های متفاوتی دارد. در این حالت،  $\lambda > 0$  نیز پارامتر تنظیم‌کننده از قبل تعیین شده است.  $\lambda$  در اینجا، موازنه‌ای بین خطای پیشگویی و پیچیدگی مدل انجام می‌دهد و بنابراین، پاسخ این مسئله، به انتخاب  $\lambda$  به شدت وابسته است. تعیین مقدار مناسب برای پارامتر تنظیم‌کننده در آمار نیز، یک مسئله چالش برانگیز است.

جریمه‌های متفاوتی می‌تواند در نظر گرفته شود که در هر حالت، برآوردگر به دست آمده، نام خاصی به خود می‌گیرد. جریمه‌های  $\sum_{j=1}^p |\beta_j|$  و  $\sum_{j=1}^p \beta_j^2$  از معروفترین توابع جریمه هستند که به ترتیب بر اساس نرم‌های مستطیلی ( $L_1$ ) و اقلیدسی ( $L_2$ ) می‌باشند.

یکی از این روش‌های جریمه شده، رگرسیون جریمه شده ریج است که برای اولین بار در سال ۱۹۷۰ توسط **هورل و کنارد** معرفی شد که عملکرد آن انقباض پارامترها به سمت صفر است.

رگرسیون ریج توسط **هورل و کنارد (۱۹۷۰)** از کمینه کردن رابطه (۵.۳) با جای‌گذاری تابع جریمه

$$P(\beta) = \sum_{i=1}^p \beta_i^2 \quad (۶.۳)$$

به صورت

$$\hat{\beta}^{\text{ridge}}(\lambda) = (\mathbf{X}^\top \mathbf{X} + \lambda \mathbf{I}_p)^{-1} \mathbf{X}^\top \mathbf{y}$$

به دست می‌آید.

<sup>۸</sup>Penalized regression

رگرسیون ريج زمانی مناسب است که  $X^T X$  معکوس پذیر نباشد، چون برآوردگر کمترین توان های دوم (۱) حتی زمانی که  $X^T X$  معکوس پذیر نیست، یکتا نیست؛ (۲) هنگامی که واریانس بالایی داشته باشد،  $X^T X$  معکوس پذیر نیست. بنابراین علی رغم وجود نااریبی<sup>۹</sup> برآوردگر در مدل، واریانس بالا باعث افزایش خطای MSE شده است. بنابراین به ازای  $\lambda > 0$ ، برآوردگر ريج اریب است و همیشه به ازای هر  $\lambda > 0$ ، MSE برآوردگر ريج کمتر از MSE برآوردگر OLS می باشد. این نتایج در زمینه داده های مقطعی اثبات شده و استفاده از رگرسیون ريج را برای داده های سری زمانی مورد استفاده قرار می دهد. توجه کنید که اگرچه منقبض کردن، برآورد پارامتر (ضرایب) رگرسیون ريج به سمت صفر را نسبت به OLS کاهش می دهد، به عبارت دیگر، به ازای  $\lambda > 0$ ،  $\|\hat{\beta}\|^2 < \|\hat{\beta}^{\text{ridge}}(\lambda)\|^2$  اما برآوردگر  $\hat{\beta}^{\text{ridge}}(\lambda)$  دقیقاً برابر صفر نمی شود، بنابراین برآوردگر ريج مستقیماً برای انتخاب متغیر مناسب نیست.

### ۱.۲.۳ لاسو

اگر چه رگرسیون ريج، بهبود پیشگویی را نتیجه می دهد ولی با توجه به حضور همه متغیرها در این برآوردگر، تفسیرپذیری آن به سادگی امکان پذیر نیست. از طرف دیگر، حضور متغیرهایی که با متغیر پاسخ در یک مدل ارتباطی ندارند، در فرآیند برآورد، منجر به برآوردگری با واریانس زیاد می شود.

لاسو توسط **تیبشیرانی** (۱۹۹۶) از کمینه کردن رابطه (۵.۳) با جای گذاری تابع جریمه

$$P(\beta) = \sum_{i=1}^p |\beta_i| \quad (7.3)$$

به صورت

$$\hat{\beta}^{\text{LASSO}} = \operatorname{argmin}_{\beta} \left\{ (\mathbf{y}_n - \mathbf{X}_n \beta)^T (\mathbf{y}_n - \mathbf{X}_n \beta) + \lambda \sum_{i=1}^p |\beta_i| \right\} \quad (8.3)$$

پیشنهاد شد که در آن  $\lambda$  پارامتر حد آستانه است که سطح تنکی (تعداد پارامترهای صفر) را در برآوردگر لاسو کنترل می کند. وقتی  $\lambda \rightarrow 0$ ،  $\hat{\beta}^{\text{LASSO}} \rightarrow \hat{\beta}$  و وقتی  $\lambda \rightarrow \infty$ ،  $\hat{\beta}^{\text{LASSO}} \rightarrow 0$ . به عبارت دیگر، هرچه جریمه بزرگتری به کار برده شود، ضرایب بیشتری به سمت صفر منقبض می شوند.

انتخاب تابع جریمه (۷.۳) باعث می شود که برآوردگر لاسو، ضرایب را به سمت صفر منقبض کرده و متغیرهای اضافی را دقیقاً صفر نماید. در واقع، برآوردگر لاسو، همزمان هم انتخاب متغیر انجام می دهد و هم ضرایب را منقبض می کند. هم چنین، تابع جریمه (۷.۳) باعث می شود که جواب صریحی برای برآوردگر لاسو وجود نداشته باشد و **تیبشیرانی** (۱۹۹۶) این برآوردگر را از طریق برنامه ریزی درجه دوم به دست آورده بود. **افرون و همکاران** (۲۰۰۴)

<sup>9</sup>Unbiased

رگرسیون کمترین زاویه (LAR) نوعی از رگرسیون گام به گام را معرفی نمودند که برآوردهای لاسویی را نتیجه می‌دهد که هزینه محاسباتی آن با برآوردهای کمترین توان‌های دوم برابر است.

کاربرد جالب از برآوردهای لاسو در مدل‌های تنک (مدل‌هایی که تعداد پارامترهای صفر آن زیاد باشد) است. کاربرد دیگر برآوردهای لاسو، زمانی است که بعد فضای پارامتر،  $p$  بسیار بزرگتر از تعداد مشاهده‌ها،  $n$  باشد.

### ۲.۲.۳ لاسو سازوار

**فن و لی (۲۰۰۱a)** نشان دادند که یک تابع جریمه خوب باید برآوردگری با سه ویژگی مطلوب نااریبی، تنگی و پیوستگی را نتیجه دهد. جریمه  $L_1$  برآوردهایی را نتیجه می‌دهد که تنگی، نسبت به داده‌ها پیوسته و اریب می‌باشد زیرا اندازه انقباضی که به ضرایب رگرسیونی کوچک و بزرگ اختصاص می‌دهد، یکسان است. همچنین، لاسو برای یک مدل رگرسیونی خطی با  $p$  متغیر توضیحی و  $n$  مشاهده، حداکثر  $n$  متغیر انتخاب می‌کند و در بین یک گروه از متغیرهای به شدت همبسته، فقط یک متغیر را انتخاب می‌نماید. اگر همبستگی بالایی نیز بین متغیرهای توضیحی وجود داشته باشد، کارایی پیشگویی لاسو به خوبی برآوردهای ریبج نیست.

**تعریف ۱.۲.۳ (خاصیت پیشگویی).** یک متغیر یا رویه انتخاب متغیر دارای خاصیت پیشگویی گویند اگر بتواند زیرمجموعه درست و واقعی متغیرها را شناسایی کند. بنابراین مطلوب است که یک برآوردهای خاصیت پیشگویی باشد.

**ژو (۲۰۰۶)** نشان داد که برآوردهای لاسو، خاصیت پیشگویی ندارد و برآوردهای لاسو سازوار<sup>۱۰</sup> را معرفی کرد که از کمینه کردن رابطه (۵.۳) با جای‌گذاری تابع جریمه

$$P(\beta) = \sum_{i=1}^p \frac{1}{|\hat{\beta}_i|} |\beta_i| \quad (9.3)$$

به دست می‌آید که در آن  $\hat{\beta}_i$ ، برآوردهای کمترین توان‌های دوم است و آن را با نماد  $\hat{\beta}^{\text{ALASSO}}$  نشان می‌دهند.

توجه کنید که در مقایسه با برآوردهای ریبج و لاسو، جریمه برآوردهای لاسو سازوارها به واسطه برآوردهای OLS به داده‌ها وابسته است. ایده اصلی برآوردهای جریمه لاسو سازوار استفاده از وزن بیشتر متناسب با پارامترها در مقایسه با برآوردهای لاسو است که همه پارامترها را با همان پارامتر تنظیم‌کننده  $\lambda$  جریمه می‌کند. تا زمانی که برآوردهای OLS سازگار باشد، منجر به این می‌شود که ضرایب صفر، بیشتر از ضرایب غیرصفر جریمه شود. اگر پارامتر واقعی و البته نامعلوم  $\beta_i = 0$  در این صورت  $\hat{\beta}_i$  برای نمونه‌های بزرگ، به صفر نزدیک خواهد شد. بنابراین  $\sum_{i=1}^p \frac{1}{|\hat{\beta}_i|}$  به سمت بی‌نهایت میل می‌کند و هزینه‌های بالایی برای انتخاب متغیر مقادیر غیرصفر  $\beta_i$  در

<sup>10</sup> Adaptive LASSO

مساله کمینه‌سازی تحمیلی می‌کند. بنابراین ضرایب صفر، خیلی بیشتر از ضرایب غیرصفر جریمه می‌شوند.

ذکر این نکته ضروری است که در ابتدا ژو (۲۰۰۶) تابع جریمه

$$P(\beta) = \sum_{i=1}^p \frac{1}{|\hat{\beta}_i|^\alpha} |\beta_i| \quad \forall \alpha > 0 \quad (10.3)$$

را پیشنهاد داد. انتخاب  $\alpha = 1$  در عمل رایج است و از وجود پارامتر تنظیم‌کننده اضافی دیگر جلوگیری می‌شود. همچنین برآوردگرهای دیگری به جز برآوردگر کمترین توان‌های دوم می‌توانند به‌عنوان برآوردگر اولیه برای برآوردگر لاسو سازوار استفاده شوند. برآوردگر OLS زمانی که  $p > n$ ، یکتا نیست. در این حالت می‌توان از برآوردگر اولیه لاسو به‌جای برآوردگر OLS در تابع جریمه ۹.۳ استفاده کرد. به‌عبارت دیگر،

$$P(\beta) = \sum_{i=1}^p \frac{1}{|\hat{\beta}_i^{\text{LASSO}}|} |\beta_i|$$

بنابراین، برآوردگر لاسو سازوار در مدل‌های با بعد بالا، نیز کاربرد دارد.

### ۳.۲.۳ الاستیکنت

برآوردگر الاستیکنت<sup>۱۱</sup> توسط ژو و هیستی (۲۰۰۵) از کمینه کردن رابطه (۵.۳) با جای‌گذاری تابع جریمه

$$P(\beta) = \alpha \sum_{i=1}^p \beta_i^2 + (1 - \alpha) \sum_{i=1}^p |\beta_i| \quad \forall \alpha \in [0, 1] \quad (11.3)$$

به‌دست می‌آید و آن را با نماد  $\hat{\beta}^{\text{EN}}$  نشان می‌دهند. جریمه الاستیکنت، ترکیب محدبی از جریمه ریج ( $\alpha = 1$ ) و جریمه لاسو ( $\alpha = 0$ ) است. ایده اصلی برآوردگر الاستیکنت، ترکیبی از نقاط قوت این برآوردگرهاست.

همان‌طور که در ژو و هیستی (۲۰۰۵) مطرح شده است، در صورتی که یک گروه از متغیرهای توضیحی دارای همبستگی بالا باشد، لاسو تنها یکی از متغیرهای موجود در گروه را انتخاب می‌کند. در صورتی که بتوان یکی از متغیرهای با همبستگی بالا را به‌عنوان نماینده گروه در نظر گرفت، می‌توان تمام متغیرهای مربوطه را فقط به جای یکی از اعضای گروه انتخاب کرد. از این‌رو در چنین شرایطی ممکن است استفاده از برآوردگر الاستیکنت به برآوردگر لاسو ترجیح داده شود.

پس از معرفی برآوردگر لاسو، فرانک و فریدمن (۱۹۹۳)، تیبشیرانی و همکاران (۲۰۰۵)، ژو (۲۰۰۶)، ژانگ و و.لو (۲۰۰۷) و پارک و گ.کسلا (۲۰۰۸) نسخه‌های تعمیم‌یافته این برآوردگر را ارائه دادند. اخیراً، وانگ و همکاران (۲۰۰۷) روش‌های لاسو و لاسوی سازوار را برای

<sup>۱۱</sup>Elastic Net

مدل‌های رگرسیون خطی با خطاهای AR به کار برده‌اند. همان‌طور که در **وانگ و همکاران (۲۰۰۷)** اشاره شده است، روش لاسوی سازوار، پارامترهای تنظیم‌کننده متفاوتی را برای هر یک از متغیرهای رگرسیونی در نظر می‌گیرد و در نتیجه، برآورد دقیقی از ضرایب رگرسیونی با به کار بردن مقدار انقباض بیشتر برای متغیرهای بی‌اهمیت و مقدار انقباض کمتر برای متغیرهای بااهمیت نتیجه می‌دهد. هم‌چنین **پ.الکواپر و پ.دوخان (۲۰۱۱)** از روش‌های جریمه‌شده در سری‌های زمانی استفاده کرده‌اند.

### ۴.۲.۳ انتخاب پارامترهای جریمه

انتخاب پارامترهای جریمه بخش مهمی از برازش مدل در رگرسیون جریمه‌شده است و در کاربردهای آماری نقش مهمی ایفا می‌کند. پارامتر جریمه  $\lambda$  مستقیماً به تعداد متغیرهای موجود در مدل مربوط می‌شود، بنابراین مقداری کلیدی برای انتخاب درست مدل است. روش‌های بسیاری در این مورد وجود دارد اما مهم‌ترین روش‌های کاربردی بر اساس این است که برآوردگر جریمه‌شده چه قدر خوب می‌تواند مقدار واقعی متغیر پاسخ را پیشگویی کند.

#### اعتبارسنجی متقابل

منصفانه نیست از داده‌ها دوبار استفاده شود، یک بار برای برازش مدل و دوباره از همان داده‌ها برای برآورد دقت پیشگویی. این کار، منجر به بیش‌برازشی می‌شود. ایده‌ای که به ذهن می‌رسد، تقسیم‌بندی داده‌ها به دو مجموعه است. از یک مجموعه برای برازش برآوردگر جریمه‌شده  $\hat{\beta}$  استفاده شود و از مجموعه دیگر برای محاسبه اینکه کمیت  $X\hat{\beta}$  چه قدر خوب توانسته است مشاهدات مجموعه دیگر را پیشگویی کند. اما به ندرت تعداد داده‌ها آن قدر زیاد است که بتوان آنها را به دو قسمت تقسیم کرد به طوری که هر یک از مجموعه‌ها به تنهایی برای انتخاب پارامتر جریمه کافی باشند. از این رو، روش اعتبارسنجی متقابل<sup>۱۲</sup> (CV) پیشنهاد می‌شود.

ایده روش‌های CV این است که نمونه به دو زیرمجموعه جدا از هم<sup>۱۳</sup> تقسیم شود: (۱) مجموعه آموزشی<sup>۱۴</sup> (یا درون-نمونه<sup>۱۵</sup>)، (۲) مجموعه اعتبارسنجی<sup>۱۶</sup> (یا برون-نمونه<sup>۱۷</sup>). پارامترهای مدل فقط با استفاده از داده‌های مجموعه آموزشی، برآورد می‌شوند و کارایی مدل پس از بهره‌گیری از پارامترهای برآوردشده بر روی مجموعه اعتبارسنجی آزمون می‌شود. فرض کنید  $A$  و  $\Lambda$  به ترتیب مجموعه مقادیر اولیه‌ای باشند که برای  $\lambda$  و  $\alpha$  در نظر گرفته شده است.

<sup>12</sup>Cross-validation

<sup>13</sup>Disjoint

<sup>14</sup>Training Set

<sup>15</sup>In-Sample

<sup>16</sup>Validation Set

<sup>17</sup>Out-of-Sample

همچنین فرض کنید  $V \subseteq \{1, \dots, T\}$  اندیس مشاهدات در مجموعه اعتبارسنجی و مجموعه  $T \subseteq \{1, \dots, T\}$  نیز اندیس‌هایی از مشاهدات در مجموعه آموزشی باشند که برای برآورد پارامترها استفاده می‌شوند. در اغلب موارد اما نه همیشه،  $T := V^c$ . فرض کنید  $\hat{\beta}_T(\lambda, \alpha)$  برآوردگر بر اساس داده‌های آموزشی  $T$  با استفاده از پارامترهای تنظیم‌کننده  $(\lambda, \alpha) \in \Lambda \times \mathcal{A}$  باشد. همچنین، برای هر  $(\lambda, \alpha)$ ،

$$CV(\lambda, \alpha, V) = \sum_{t \in V} (y_t - \mathbf{x}_t^T \hat{\beta}_T(\lambda, \alpha))^2$$

خطای پیش‌گویی متناظر با مجموعه اعتبارسنجی  $V$  است. علاوه بر این، فرض کنید  $V = \{V_1, \dots, V_B\}$  خانواده معینی از مجموعه‌های اعتبارسنجی (متناظر با مجموعه‌های آموزشی  $\{T_1, \dots, T_B\}$ ) باشد. خطای اعتبارسنجی متقابل برای زوج مرتب  $(\lambda, \alpha)$  به صورت

$$CV(\lambda, \alpha) = \sum_{i=1}^B CV(\lambda, \alpha, V_i)$$

محاسبه می‌شود و مجموعه‌ای از پارامترهای انتخاب می‌شود که

$$(\hat{\lambda}, \hat{\alpha}) \in \underset{(\lambda, \alpha) \in \Lambda \times \mathcal{A}}{\operatorname{argmin}} CV(\lambda, \alpha).$$

پس برآورد پارامتر نهایی  $\hat{\beta}(\hat{\lambda}, \hat{\alpha})$  بر اساس همه مشاهدات  $1, \dots, T$  پیدا می‌شود. اکنون به انتخاب  $B$  و مجموعه متناظر  $\{V_1, \dots, V_B\}$  نگاه دقیق‌تری می‌شود. دو روش کلاسیک برای تقسیم کردن نمونه وجود دارد (۱)  $CV$  کامل<sup>۱۸</sup> و (۲)  $CV$  ناکامل<sup>۱۹</sup>. (برای مطالعاتی که اخیراً در مورد روش‌های  $CV$  برای انتخاب متغیر انجام شده است به **آرلوت و کلیسه**، ۲۰۱۰ مراجعه کنید).

مهم‌ترین عضو خانواده روش‌های  $CV$  کامل، روش  $CV_{\nu}$  حذفی است. ایده این است که از  $\nu$  مشاهده به‌عنوان مجموعه اعتبارسنجی و از سایر مشاهده‌ها یعنی  $T - \nu$  برای برآورد پارامترهای مدل استفاده می‌شود. این روش برای تمام حالت‌های ممکن انتخاب  $\nu$  مشاهده از بین  $T$  مشاهده انجام می‌شود. بنابراین، هر  $V_i$ ، عضو  $\nu$  به‌ازای  $i = 1, \dots, B$ ،  $T_i = V_i^c$  یک انتخاب مناسب این است که  $\nu = 1$  قرار داده شود به‌طوری‌که  $B = T$  و  $V_i = \{i\}$  که  $CV$  تک حذفی<sup>۲۰</sup> نامیده می‌شود.

روش‌های  $CV$  کامل، هزینه محاسباتی بالایی دارند زیرا باید روش  $CV$  بر روی بسیاری از تقسیمات نمونه انجام داده شود. یک راه‌حل برای غلبه بر هزینه بالای محاسباتی این روش‌ها، استفاده از روش‌های  $CV$  ناکامل است که در بین آن‌ها، روش اعتبارسنجی متقابل  $B$ -تایی<sup>۲۱</sup> محبوب‌ترین روش است. در روش  $CV_{B}$ -تایی، نمونه به‌طور تصادفی به  $B$  زیرمجموعه با

<sup>18</sup>Exhaustive CV

<sup>19</sup>Non-Exhaustive CV

<sup>20</sup>Leave-One-Out

<sup>21</sup>B-fold CV



(تقریباً) تعداد مشاهدات یکسان تقسیم می‌شود. بنابراین  $v_1, \dots, v_B$  تعداد اعضای تقریباً یکسان دارند. در حقیقت، هر  $v_i$  به گونه‌ای انتخاب می‌شود که  $T/B$  مشاهده دارد به طوری که  $\max_{1 \leq i \leq j \leq B} ||v_i| - |v_j|| \leq 1$ . در موارد بسیاری از اجراهای مختلف، هیچ یک از  $B$  زیرمجموعه اعتبارسنجی با یکدیگر اشتراکی ندارند و هر مشاهده فقط به یک زیرمجموعه تعلق دارد. دوباره می‌توان بیان کرد که  $T_i = v_i^0, i = 1, \dots, B$ .

روش‌های CV به طور گسترده در اقتصادسنجی و آمار برای انتخاب متغیر در مدل‌های خطی مورد استفاده قرار گرفته است. در چارچوب رگرسیون‌های جریمه‌شده، متغیرهای انتخاب‌شده به انتخاب پارامترهای جریمه بستگی دارد. با این وجود، برای داده‌های سری زمانی روش‌های CV که تاکنون مطرح شده‌اند، مشابه پذیره مشاهده‌های مستقل، انتخاب متغیر سازگار را تضمین نمی‌کند. برای جزئیات بیشتر، شائو (۱۹۹۳) و ریسین (۲۰۰۰) را ببینید.

یک راه‌حل کلیدی استفاده از روش اعتبارسنجی متقابل  $h$ -بلوکی<sup>۲۲</sup> است. ایده این روش به این صورت است که  $h$  مشاهده قبلی و بعدی از  $i$  امین مشاهده حذف شود و برآورد فقط بر اساس سایر مشاهده‌ها انجام گیرد. بنابراین، به ازای  $i = 1, \dots, T$

$$T_i = \{1, \dots, i-h-1\} \cup \{i+h+1, \dots, T\}$$

که اگر  $a > b$ ، آن گاه  $\{a, \dots, b\} = \emptyset$ . ممکن است  $V_i = \{i\}$  انتخاب شود.

با این وجود، ریسین (۲۰۰۰) نشان داد که روش اعتبارسنجی متقابل  $h$ -بلوکی نیز یک روش انتخاب مدل ناسازگار برای داده‌های سری زمانی است و به جای آن، روش اعتبارسنجی متقابل  $h\nu$ -بلوکی را پیشنهاد کرد. ایده این روش، به این صورت است که ابتدا  $\nu$  مشاهده به ازای  $i = 1, \dots, T-\nu$  از داده‌ها حذف و از  $v_i = \{i-\nu, i+\nu\}$  استفاده شود. سپس،  $h$  مشاهده قبلی و بعدی از  $i$  امین مشاهده حذف شود. به عبارت دیگر،  $T_i = \{1, \dots, i-\nu-h-1\} \cup \{i+\nu+h+1, \dots, T\}$  که به ازای  $a > b$ ،  $\{a, \dots, b\} = \emptyset$ . برای انتخاب  $h$  و  $\nu$  به ریسین (۲۰۰۰) مراجعه کنید. دقت کنید که نتایج ارائه شده در ریسین (۲۰۰۰) بر اساس برآوردگر OLS به دست آمده است.

## معیارهای اطلاع

پارامترهای جریمه مستقیماً به درجه آزادی مدل‌ها مربوط می‌شوند. معیارهای اطلاع<sup>۲۳</sup> کلاسیک مانند AIC (آکاییک، ۱۹۷۴)، BIC (شوارتز، ۱۹۷۸)، و نیز HQ (هنان و کویین، ۱۹۷۹) برای انتخاب  $\lambda$  می‌توانند استفاده شوند. فرض کنید

$$\hat{\sigma}^2(\lambda, \alpha) = \frac{1}{T} \sum_{t=1}^T (y_t - \mathbf{x}_t^\top \hat{\beta}(\lambda, \alpha))^2.$$

<sup>22</sup>  $h$ -block CV

<sup>23</sup> Information criteria

برای هر برآوردگر  $\hat{\beta}(\lambda, \alpha)$  مربوط به مقادیر مشخص  $(\lambda, \alpha)$  باشد. در این صورت، AIC و BIC و HQ به ترتیب به صورت

$$\text{AIC}(\lambda, \alpha) = \log(\hat{\sigma}^2(\lambda, \alpha)) + df(\lambda, \alpha) \frac{2}{T}$$

$$\text{BIC}(\lambda, \alpha) = \log(\hat{\sigma}^2(\lambda, \alpha)) + df(\lambda, \alpha) \frac{\log(T)}{T}$$

$$\text{HQ}(\lambda, \alpha) = \log(\hat{\sigma}^2(\lambda, \alpha)) + df(\lambda, \alpha) \frac{\log \log(T)}{T}$$

تعریف می‌شود که  $df(\lambda, \alpha)$ ، درجه آزادی روش برآوردی است که از آن استفاده شده است. برای روش‌های جریمه‌شده مانند لاسو، لاسو سازوار و الاستیکنت معمولاً  $df(\lambda, \alpha)$  تعداد پارامترهای غیرصفر برآورد شده برای  $(\lambda, \alpha)$  مشخص است. برای هر یک از معیارهای اطلاع، زوج مرتب  $(\hat{\lambda}, \hat{\alpha})$  که آن را کمینه کند که منجر به برآوردگر  $\hat{\alpha}(\hat{\lambda}, \hat{\alpha})$  می‌شود.

### ۳.۳ پیاده‌سازی‌های کامپیوتری

تمام مدل‌های ارائه‌شده در این فصل با استفاده از نرم‌افزار R اجرا شده است. بسته glmnet (فریدمن و همکاران، ۲۰۱۰) می‌تواند برآوردگرهای لاسو، الاستیکنت و نسخه‌های سازوار آن را اجرا کند. برای انتخاب  $\lambda$ ، بسته glmnet فقط از روش اعتبارسنجی متقابل استفاده می‌کند و برای بهره‌گیری از معیارهای اطلاعات، باید برنامه آن‌ها نوشته شود. نتیجه برآورد مدل‌های رگرسیونی جریمه‌شده با استفاده از glmnet، ارائه مسیر از بزرگ‌ترین مدل ممکن به مدلی که فقط عرض از مبدا دارد، می‌باشد. اگرچه glmnet محبوب‌ترین بسته برای اجرای برآوردگرهای انقباضی است، بسته‌های دیگری نیز در R وجود دارد.

### ۴.۳ شبیه‌سازی

فرآیند تولید داده را به صورت زیر در نظر بگیرید. (مدیروز و مندز، ۲۰۱۶)

$$y_t = \phi y_{t-1} + \mathbf{x}_{1,t-1}^\top \beta^\circ + \varepsilon_t, \quad \phi = 0.6$$

$$\varepsilon_t = \sigma_t u_t, u_t \sim t_{\delta}^*$$

$$\sigma_t^2 = 5 \times 10^{-4} + 0.9 \sigma_{t-1}^2 + 0.05 \varepsilon_{t-1}^2$$

$$\mathbf{x}_t = \begin{bmatrix} \mathbf{x}_{1,t} \\ \mathbf{x}_{2,t} \end{bmatrix} = \mathbf{A}_1 \begin{bmatrix} \mathbf{x}_{1,t-1} \\ \mathbf{x}_{2,t-1} \end{bmatrix} + \mathbf{A}_\varphi \begin{bmatrix} \mathbf{x}_{1,t-\varphi} \\ \mathbf{x}_{2,t-\varphi} \end{bmatrix} + \nu_t, \nu_t \sim t_\delta^*$$

پارامتر  $\beta$  در حالت کلی نامعلوم است اما در مدل‌های شبیه‌سازی،  $\beta$  معلوم در نظر گرفته می‌شود که برای سهولت آن را با نماد  $\beta^\circ$  نشان می‌دهند. داده‌ها بر اساس این مقدار معلوم ( $\beta^\circ$ ) تولید می‌شوند. سپس، فرض می‌شود مقدار پارامتر نامعلوم است و تلاش می‌شود با روش‌های مختلف برآورد شود که در نهایت با مقایسه با مقدار واقعی، کارایی روش‌ها سنجیده می‌شود.

فرض کنید  $S = \{i : \beta_i^\circ \neq 0\}$  اندیس‌های مربوط به ضرایب غیرصفر را نشان دهد و همچنین  $s_o = |S|$ ، تعداد اعضای مجموعه  $S$  باشد. اگر  $s_o \ll p$ ، آن‌گاه  $\beta^\circ$ ، تنک<sup>۲۴</sup> نامیده می‌شود. **مدیروز و مندز (۲۰۱۶)** پیشنهاد کردند که هر مولفه بردار  $\beta^\circ$  به‌ازای  $i = 1, \dots, s_o - 1$ ، به‌صورت  $\beta_i^\circ = \frac{1}{\sqrt{s_o}}(-1)^i$  در نظر گرفته شود. همچنین،  $\mathbf{x}_{1,t}$  یک بردار  $1 \times (s_o - 1)$  از متغیرهای غیرصفر (معنی‌دار) است و  $\{u_t\}$  و دنباله‌های مستقل و هم‌توزیع از توزیع  $t$  با  $\delta$  درجه آزادی است که استاندارد شده است تا میانگین صفر و واریانس یک داشته باشند که توزیع  $t$  استاندارد با نماد  $t^*$  نشان داده شده است. بردار  $\mathbf{x}_t = (\mathbf{x}_{1,t}^\top, \mathbf{x}_{2,t}^\top)^\top \in \mathbb{R}^{p-1}$  دارای  $p - s_o$  متغیر نامرتبب  $\mathbf{x}_{2,t}$  است که ضرایب آن‌ها صفر است و از مدل

$$\mathbf{x}_t = \psi_0 + \psi_1 \mathbf{x}_{t-1} + \psi_2 \mathbf{x}_{t-2} + \psi_3 \mathbf{x}_{t-3} + \psi_4 \mathbf{x}_{t-4} + \varepsilon_t, \quad t = 1, \dots, T \quad (12.3)$$

پیروی می‌کند. ماتریس‌های  $\mathbf{A}_1$  و  $\mathbf{A}_\varphi$  ماتریس‌های بلوکی هستند که بعد هر بلوک،  $5 \times 5$  است و به‌ترتیب، مقادیر قطر اصلی آن‌ها برابر با  $0.15$  و  $-0.1$  است. توزیع خطاها، نرمال نیستند و از توزیع  $t$  که توزیع دم-پهن است تولید شده است و فرض شده است واریانس آن‌ها با هم برابر نیست. واریانس غیرشرطی  $\varepsilon_t$  به‌صورت  $0.01 = 0.05 / (1 - 0.90 - 0.05) = 5 \times 10^{-4}$  محاسبه می‌شود.

به‌منظور ارزیابی اثرات نرمال نبودن توزیع و ناهمگن بودن واریانس‌ها، هم‌چنین در شبیه‌سازی مواردی که توزیع خطاها، نرمال باشد و هم‌چنین برابری واریانس‌ها نیز بررسی شده است. واریانس  $\nu_t$  برابر با  $1$  و واریانس  $\varepsilon_t$  برابر با  $0.01$  فرض شده است. شبیه‌سازی به‌ازای  $T = 100, 500, 1000$  و  $p = 100, 500, 1000$  و  $s_o = 5, 10, 20$  انجام شده است. مدل‌ها با استفاده از روش‌های لاسو، لاسو سازوار، الاستیکنت برآورد شده است. وزن‌های به‌کاررفته در محاسبه برآوردگر لاسو سازوار، معکوس مقادیر مطلق ضرایب برآوردشده از برآوردگر لاسو است. پارامتر  $\lambda$  با استفاده از معیار BIC انتخاب می‌شود و پارامتر  $\alpha$  در برآوردگر الاستیکنت برابر با  $0.5$  و در برآوردگر لاسو سازوار برابر با  $1$  و هم‌چنین تعداد تکرارهای شبیه‌سازی،  $1000$  در نظر گرفته شده است.

جدول‌های ۱.۳ و ۲.۳ به‌ترتیب متوسط اریبی (BIAS) و متوسط MSE برای برآوردگرها در

<sup>24</sup>Sparse

تکرارهای مونت کارلو و متغیرهای داوطلب را نشان می‌دهد. به عبارت دیگر،

$$\text{Bias} = \frac{1}{1000p} \sum_{j=1}^{1000} \left[ \hat{\phi}^j - 0.6 + \sum_{i=1}^p (\hat{\beta}_i^j - \beta_i^0) \right]$$

و

$$\text{MSE} = \frac{1}{1000p} \sum_{j=1}^{1000} \left[ (\hat{\phi}^j - 0.6)^2 + \sum_{i=1}^p (\hat{\beta}_i^j - \beta_i^0)^2 \right]$$

که  $\hat{\phi}^j$  و  $\hat{\beta}_i^j$  برآوردگرهای  $\phi$  و  $\beta_i$  را در زامین تکرار مونت کارلو نشان می‌دهند.

جدول ۱.۳: متوسط اریبی برآوردگرهای لاسو، لاسو سازوار و الاستیک‌نت

p/s <sub>0</sub>	T = 500			T = 100		
	1000	500	100	1000	500	100
لاسو						
5	-0.019	-0.025	-0.085	-0.059	-0.063	-0.186
10	-0.007	-0.019	-0.067	-0.074	-0.076	-0.121
20	-0.008	-0.015	-0.048	-0.164	-0.187	-0.101
لاسو سازوار						
5	-0.012	-0.009	-0.029	-0.014	-0.018	-0.059
10	-0.009	-0.007	-0.017	-0.034	-0.021	-0.038
20	-0.005	-0.002	-0.009	-0.084	-0.057	-0.043
الاستیک‌نت						
5	-0.062	-0.081	-0.105	-0.099	-0.112	-0.342
10	-0.023	-0.052	-0.097	-0.108	-0.129	-0.195
20	-0.018	-0.039	-0.041	-0.163	-0.254	-0.163

چندین واقعیت از جدول‌های ۱.۳ و ۲.۳ نمایان می‌شود: (۱) اریبی و MSE همه روش‌های مختلف با افزایش حجم نمونه، کاهش پیدا می‌کند. (۲) وقتی تعداد متغیرهای معنی‌دار، s<sub>0</sub>، افزایش می‌یابد، MSE نیز افزایش می‌یابد درحالی‌که اریبی، الگوی مشخصی را نشان نمی‌دهد. (۳) نه اریبی و نه MSE الگویی بر حسب p نشان نمی‌دهد. (۴) برآوردگر الاستیک‌نت نسبت به برآوردگر لاسو برتری ندارد. (۵) برآوردگر لاسو سازوار به برآوردگر لاسو برتری دارد.

جدول ۳.۳ نتایج انتخاب مدل را بر اساس درصد تکرارهایی که الگوی تنکی درست انتخاب شده است را ارائه می‌دهد. بر اساس جدول ۳.۳، برآوردگرهای لاسو و الاستیک‌نت عملکرد ضعیفی را از نظر تشخیص مدل درست نشان می‌دهند که ممکن است به این دلیل باشد که

جدول ۲.۳: متوسط MSE برآوردگرهای لاسو، لاسو سازوار و الاستیک‌نت

$T = 500$			$T = 100$			$p/s_0$
۱۰۰۰	۵۰۰	۱۰۰	۱۰۰۰	۵۰۰	۱۰۰	
<u>لاسو</u>						
۰/۰۰۱	۰/۰۰۸	۰/۰۱۲	۰/۰۱۹	۰/۰۲۱	۰/۰۵۴	۵
۰/۰۰۲	۰/۰۱۱	۰/۰۱۵	۰/۱۲۴	۰/۰۵۸	۰/۰۷۳	۱۰
۰/۰۰۶	۰/۰۱۴	۰/۰۱۹	۰/۷۶۱	۰/۸۶۷	۰/۱۴۵	۲۰
<u>لاسو سازوار</u>						
۰/۰۰۲	۰/۰۰۱	۰/۰۰۳	۰/۰۰۱	۰/۰۰۷	۰/۰۱۱	۵
۰/۰۰۷	۰/۰۰۲	۰/۰۰۵	۰/۰۵۲	۰/۰۳۱	۰/۰۱۹	۱۰
۰/۰۲۶	۰/۰۰۲	۰/۰۰۸	۰/۸۰۴	۰/۷۸۴	۰/۰۵۱	۲۰
<u>الاستیک‌نت</u>						
۰/۰۰۵	۰/۰۰۹	۰/۰۲۱	۰/۰۶۴	۰/۰۷۸	۰/۰۸۹	۵
۰/۰۰۶	۰/۰۱۱	۰/۰۲۲	۰/۲۸۹	۰/۲۳۵	۰/۱۰۹	۱۰
۰/۰۱۳	۰/۰۱۳	۰/۰۳۱	۰/۹۶۸	۱/۱۱۰	۰/۱۹۵	۲۰

متغیرهای پیش‌گو همبسته هستند که در این راستا، می‌توان به **ژائو و یو (۲۰۰۶)** مراجعه کرد. از طرفی، برآوردگر لاسو سازوار، نتایج خوبی را برای نمونه‌های بزرگ در عمل نشان می‌دهد. جدول ۴.۳ نتایج انتخاب مدل را بر اساس درصد تکرارهایی که متغیرهای معنی‌دار را شامل می‌شوند، ارائه می‌دهد. بر اساس جدول ۴.۳، یک نتیجه جالب این است که روش‌های لاسو و لاسوی سازوار تقریباً نتایج یکسانی دارند هر چند در برخی موارد نتایج حاکی از آن است که برآوردگر لاسوی سازگار درصد بیشتری از متغیرهای معنی‌دار را نسبت به لاسو شامل می‌شود. همچنین، هر چه تعداد متغیرها در مدل افزایش یابد، نتایج به‌دست آمده بدتر می‌شود.

جدول ۳.۳: جدول درصد تکرارهای شبه‌سازی از انتخاب الگوی تنکی درست

$T = 500$			$T = 100$			$p/s_0$
۱۰۰۰	۵۰۰	۱۰۰	۱۰۰۰	۵۰۰	۱۰۰	
لاسو						
۰/۰۲۱	۰/۰۳۹	۰/۰۶۱	۰/۰۲۵	۰/۰۰۳	۰/۰۱۰	۵
۰/۰۰۱	۰/۰۰۰	۰/۰۰۵	۰/۰۰۱	۰/۰۰۰	۰/۰۰۰	۱۰
۰/۰۰۰	۰/۰۰۰	۰/۰۰۱	۰/۰۰۰	۰/۰۰۰	۰/۰۰۰	۲۰
لاسو سازوار						
۱/۰۰۰	۱/۰۰۰	۱/۰۰۰	۰/۹۶۸	۰/۹۷۲	۰/۸۰۳	۵
۱/۰۰۰	۰/۹۹۴	۰/۹۹۷	۰/۸۸۴	۰/۹۶۷	۰/۶۱۸	۱۰
۱/۰۰۰	۰/۸۱۱	۰/۸۴۲	۰/۰۱۹	۰/۱۵۳	۰/۱۸۹	۲۰
الاستیک‌نت						
۰/۰۰۲	۰/۰۰۰	۰/۰۰۳	۰/۰۰۰	۰/۰۰۰	۰/۰۰۰	۵
۰/۰۰۰	۰/۰۰۰	۰/۰۰۰	۰/۰۰۰	۰/۰۰۰	۰/۰۰۰	۱۰
۰/۰۰۰	۰/۰۰۰	۰/۰۰۰	۰/۰۰۰	۰/۰۰۰	۰/۰۰۰	۲۰

جدول ۵.۳، MSE برای پیش‌بینی‌های مجموعه اعتبار در روش CV نشان می‌دهد.  $100$  مجموعه اعتبار در نظر گرفته شده است. همان‌طور که انتظار می‌رود نتایج برای  $T = 500$  بهتر از نتایج برای  $T = 100$  است. علاوه بر این، تمام روش‌ها با افزایش تعداد متغیرهای معنی‌دار یا آن‌هایی که معنی‌دار نیستند، برای مقادیر  $s_0$  و  $p$  بزرگ، ضعیف‌تر عمل می‌کنند. سرانجام، برآوردگر لاسو سازوار بر برآوردگر لاسو برتری دارد. هم‌چنین حالتی که فرض تنکی نادیده گرفته شود یعنی  $s_0 = p$  نیز در نظر گرفته شده است که در این حالت، نتایج طبیعتاً رضایت‌بخش نیست.

نتایج برای مدل نرمال در جدول‌های ۶.۳ تا ۱۰.۳ نشان داده شده است که مشابه نتایج قبلی است.

جدول ۴.۳: جدول درصد تکرارهای شبیه‌سازی از انتخاب مدل‌هایی که متغیرهای معنی‌دار را شامل می‌شوند.

$T = 500$			$T = 100$			$p/s_0$
۱۰۰۰	۵۰۰	۱۰۰	۱۰۰۰	۵۰۰	۱۰۰	
لاسو						
۱/۰۰۰	۱/۰۰۰	۱/۰۰۰	۱/۰۰۰	۱/۰۰۰	۱/۰۰۰	۵
۱/۰۰۰	۱/۰۰۰	۱/۰۰۰	۰/۹۴۷	۰/۹۹۱	۱/۰۰۰	۱۰
۱/۰۰۰	۱/۰۰۰	۱/۰۰۰	۰/۰۷۴	۰/۵۶۷	۱/۰۰۰	۲۰
لاسو سازوار						
۱/۰۰۰	۱/۰۰۰	۱/۰۰۰	۱/۰۰۰	۱/۰۰۰	۱/۰۰۰	۵
۱/۰۰۰	۱/۰۰۰	۱/۰۰۰	۰/۹۵۷	۰/۹۹۱	۱/۰۰۰	۱۰
۱/۰۰۰	۱/۰۰۰	۱/۰۰۰	۰/۰۵۳	۰/۴۲۱	۱/۰۰۰	۲۰
الاستیک‌نت						
۱/۰۰۰	۱/۰۰۰	۱/۰۰۰	۰/۹۹۸	۱/۰۰۰	۱/۰۰۰	۵
۱/۰۰۰	۱/۰۰۰	۱/۰۰۰	۰/۸۷۴	۰/۹۹۰	۱/۰۰۰	۱۰
۱/۰۰۰	۱/۰۰۰	۱/۰۰۰	۰/۰۳۸	۰/۴۰۲	۱/۰۰۰	۲۰

جدول ۵.۳: جدول مقادیر MSE پیش‌بینی خطای مجموعه‌های اعتبار. (مقادیر در ۱۰ ضرب شده است)

$T = 500$			$T = 100$			$p/s_0$
۱۰۰۰	۵۰۰	۱۰۰	۱۰۰۰	۵۰۰	۱۰۰	
لاسو						
۰/۱۱۸	۰/۱۲۵	۰/۱۱۲	۰/۲۵۹	۰/۱۷۲	۰/۱۶۲	۵
۰/۱۴۱	۰/۱۷۰	۰/۱۱۷	۱/۴۷۸	۰/۴۳۷	۰/۱۸۹	۱۰
۰/۱۸۱	۰/۱۴۵	۰/۱۲۱	۱۳/۰۰۱	۲/۵۰۱	۰/۳۴۵	۲۰
۱۲/۹۸۷	۶/۷۶۳	۰/۱۳۹	۲۳/۵۰۱	۲۲/۰۱۴	۶/۸۴۲	$p$
لاسو سازوار						
۰/۱۰۸	۰/۱۰۷	۰/۱۰۴	۰/۱۱۹	۰/۱۱۶	۰/۱۱۰	۵
۰/۱۵۹	۰/۱۳۴	۰/۱۱۰	۰/۷۳۷	۰/۲۰۳	۰/۱۲۴	۱۰
۰/۵۰۳	۰/۱۰۷	۰/۱۱۱	۱۱/۱۳۴	۱/۶۰۱	۰/۱۷۱	۲۰
۱۲/۹۹۹	۱۱/۰۰۱	۰/۱۴۱	۲۳/۱۱۰	۲۰/۹۹۷	۹/۰۰۱	$p$
الاستیک نت						
۰/۱۳۱	۰/۱۳۹	۰/۱۱۹	۱/۱۲۷	۰/۳۱۵	۰/۱۹۷	۵
۰/۲۰۹	۰/۲۱۱	۰/۲۰۱	۴/۱۰۳	۲/۰۰۶	۰/۲۱۵	۱۰
۰/۲۳۸	۰/۱۷۱	۰/۱۳۴	۱۴/۰۳۱	۴/۹۸۷	۰/۳۹۸	۲۰
۱۴/۳۵۹	۵/۶۹۸	۰/۱۳۵	۲۱/۳۵۹	۱۹/۷۴۵	۵/۸۹۷	$p$



جدول ۶.۳: متوسط اریبی برآوردگرهای لاسو، لاسو سازوار و الاستیکنت در مدل نرمال

$T = 500$			$T = 100$			$p/s_0$
۱۰۰۰	۵۰۰	۱۰۰	۱۰۰۰	۵۰۰	۱۰۰	
لاسو						
-۰/۰۲۱	-۰/۰۳۲	-۰/۰۹۴	-۰/۰۴۲	-۰/۰/۷۴	-۰/۲۰۱	۵
-۰/۰۱۳	-۰/۰۲۴	-۰/۰۴۵	-۰/۰۴۹	-۰/۰۶۱	-۰/۱۵۲	۱۰
-۰/۰۰۹	-۰/۰۱۹	-۰/۰۴۲	-۰/۱۶۷	-۰/۲۱۵	-۰/۰۸۹	۲۰
لاسو سازوار						
-۰/۰۰۷	-۰/۰۰۹	-۰/۰۲۳	-۰/۰۱۰	-۰/۰۲۹	-۰/۰۵۱	۵
-۰/۰۰۹	-۰/۰۰۸	-۰/۰۱۱	-۰/۰۲۹	-۰/۰۲۷	-۰/۰۴۳	۱۰
-۰/۰۰۵	-۰/۰۰۲	-۰/۰۰۰	-۰/۱۰۹	-۰/۱۹۰	-۰/۰۱۵	۲۰
الاستیکنت						
-۰/۰۲۱	-۰/۰۳۵	-۰/۰۹۹	-۰/۱۰۳	-۰/۱۲۳	-۰/۲۶۱	۵
-۰/۰۲۳	-۰/۰/۳۱	-۰/۰۷۲	-۰/۱۰۳	-۰/۱۳۳	-۰/۱۸۲	۱۰
-۰/۰۱۹	-۰/۲۳	-۰/۰۵۲	-۰/۱۹۱	-۰/۲۷۴	-۰/۱۲۳	۲۰



جدول ۹.۳: جدول درصد تکرارهای شبیه‌سازی از انتخاب مدل‌هایی که متغیرهای معنی‌دار را شامل می‌شوند در مدل‌های نرمال

$T = 500$			$T = 100$			$p/s_0$
۱۰۰۰	۵۰۰	۱۰۰	۱۰۰۰	۵۰۰	۱۰۰	
لاسو						
۱/۰۰۰	۱/۰۰۰	۱/۰۰۰	۱/۰۰۰	۱/۰۰۰	۱/۰۰۰	۵
۱/۰۰۰	۱/۰۰۰	۱/۰۰۰	۰/۹۴۲	۰/۹۹۸	۱/۰۰۰	۱۰
۱/۰۰۰	۱/۰۰۰	۱/۰۰۰	۰/۰۶۳	۰/۵۲۱	۱/۰۰۰	۲۰
لاسو سازوار						
۱/۰۰۰	۱/۰۰۰	۱/۰۰۰	۱/۰۰۰	۱/۰۰۰	۱/۰۰۰	۵
۱/۰۰۰	۱/۰۰۰	۱/۰۰۰	۰/۹۳۶	۰/۹۹۱	۱/۰۰۰	۱۰
۱/۰۰۰	۱/۰۰۰	۱/۰۰۰	۰/۰۵۲	۰/۴۰۳	۱/۰۰۰	۲۰
الاستیک‌نت						
۱/۰۰۰	۱/۰۰۰	۱/۰۰۰	۰/۹۹۹	۱/۰۰۰	۱/۰۰۰	۵
۱/۰۰۰	۱/۰۰۰	۱/۰۰۰	۰/۸۷۳	۰/۹۹۳	۱/۰۰۰	۱۰
۱/۰۰۰	۱/۰۰۰	۱/۰۰۰	۰/۰۲۹	۰/۳۱۲	۱/۰۰۰	۲۰

جدول ۱۰.۳: جدول مقادیر MSE پیش‌بینی خطای مجموعه‌های اعتبار در مدل‌های نرمال (مقادیر در ۱۰ ضرب شده است)

	$T = 500$			$T = 100$			$p/s_0$
	۱۰۰۰	۵۰۰	۱۰۰	۱۰۰۰	۵۰۰	۱۰۰	
لاسو							
	۰/۱۱۹	۰/۱۲۱	۰/۱۱۱	۰/۲۵۸	۰/۲۱۷	۰/۱۶۳	۵
	۰/۱۳۹	۰/۱۲۷	۰/۱۲۳	۱/۷۱۶	۰/۵۲۶	۰/۲۰۲	۱۰
	۰/۱۷۹	۰/۱۴۹	۰/۱۳۶	۱۲/۸۱۰	۶/۹۰۶	۰/۳۳۱	۲۰
	۱۲/۵۹۸	۶/۴۸۵	۰/۱۳۹	۲۲/۵۲۶	۲۱/۸۶۹	۶/۲۳۹	$p$
لاسو سازوار							
	۰/۱۰۳	۰/۱۰۵	۰/۱۰۹	۰/۱۱۰	۰/۱۱۱	۰/۱۱۹	۵
	۰/۱۵۹	۰/۱۰۷	۰/۱۰۹	۰/۸۲۳	۰/۲۳۵	۰/۱۲۷	۱۰
	۰/۳۶۹	۰/۸۲۳	۰/۶۰۱	۱۱/۲۰۳	۵/۲۱۹	۰/۱۶۱	۲۰
	۱۳/۳۳۰	۱۰/۹۹۸	۰/۱۴۲	۲۳/۰۰۶	۲۰/۹۸۱	۸/۶۹۸	$p$
الاستیک نت							
	۰/۱۳۹	۰/۱۲۷	۰/۱۱۹	۱/۲۰۱	۰/۶۱۲	۰/۳۰۱	۵
	۰/۱۷۵	۰/۱۵۲	۰/۱۱۹	۴/۳۸۷	۱/۸۹۶	۰/۲۵۸	۱۰
	۰/۲۶۱	۰/۱۳۸	۰/۴۲۶	۱۳/۰۳۹	۹/۰۹۳	۰/۲۰۳	۲۰
	۱۴/۹۲۰	۶/۲۸۱	۰/۱۲۹	۲۱/۰۹۹	۱۹/۶۱۳	۶/۳۵۹	$p$



## فصل ۴

# رگرسیون جریمه‌شده با خطاهای سری زمانی ARMA-GARCH

### ۱.۴ مقدمه

مدل‌های رگرسیونی با خطاهای سری زمانی ARMA هنگامی که همبستگی سریالی<sup>۱</sup> بین مشاهدات وجود دارد، کاربرد فراوانی دارد. در این فصل، روش لاسو سازوار به‌عنوان روشی که جواب‌های سازگار در انتخاب متغیر نتیجه می‌دهد، مورد بررسی قرار گرفته است. در این راستا، الگوریتم پیشنهادی **یون و همکاران (۲۰۱۷)** به منظور انتخاب متغیر در رگرسیون سری‌های زمانی با خطاهای ARMA و ARMA-GARCH بررسی شده است.

**یون و همکاران (۲۰۱۷)** روش ALASSO را برای مدل‌های رگرسیون خطی با خطاهای ARMA به‌کار برده‌اند. در واقع، تلاش آن‌ها در ادامه الگوریتم پیشنهادی **وانگ و همکاران (۲۰۰۷)** است که تنها مدل AR را برای خطاهای رگرسیونی در نظر گرفته بودند. باید به این نکته اشاره کرد که روش ALASSO پیشنهادی **وانگ و همکاران (۲۰۰۷)** را نمی‌توان به‌طور مستقیم برای مدل‌های رگرسیونی با خطای ARMA به‌کار برد. **یون و همکاران (۲۰۱۷)** ابتدا یک الگوریتم گام به گام ALASSO را پیشنهاد کردند که متغیرهای پیش‌گوی معنی‌دار را در مدل‌های رگرسیونی با خطای ARMA به کمک روش جریمه ALASSO تعیین می‌کند. به‌عبارت

<sup>۱</sup>serial correlation

دیگر، در تقابل با وانگ و لنگ (۲۰۰۸)، پارامترهای جریمه برای انتخاب مرتبه و همین‌طور پارامترهای مدل ARMA خطاها اتخاذ نشده است و از معیار BIC در الگوریتم پیشنهادی خود استفاده کرده‌اند. در بسیاری از مسائل کاربردی نیز، هدف اصلی محققان انتخاب متغیرهای معنی‌دار در مدل رگرسیونی است که از هزینه‌ها و تلاش‌های اضافی جلوگیری می‌کند. سپس، تلاش کرده‌اند که الگوریتم پیشنهادی خود را به مدل‌های رگرسیونی با خطای ARMA-GARCH تعمیم دهند.

## ۲.۴ برآوردگر لاسو سازوار برای مدل‌های رگرسیون با خطاهای ARMA

مدل‌های رگرسیون خطی با خطاهای ARMA به صورت

$$y_t = x_t^\top \beta + e_t, \quad t = 1, \dots, T \quad (1.4)$$

تعریف می‌شود که  $x_t = (x_{t1}, \dots, x_{td})^\top$  متغیر پیش‌گوی  $d$  بعدی،  $\beta = (\beta_1, \dots, \beta_d)^\top$  ضرایب رگرسیون است که باید برآورد شوند و همچنین، خطاهای  $e_t$  از فرآیند  $(ARMA)(p, q)$  به صورت

$$e_t = \sum_{i=1}^p a_i e_{t-1} + \epsilon_t - \sum_{j=1}^q b_j \epsilon_{t-j} \quad (2.4)$$

پیروی می‌کنند که  $\epsilon_t$  مجموعه‌ای از متغیرهای تصادفی مستقل و هم‌توزیع با میانگین صفر و واریانس  $\sigma^2$  است.

پارامترهای مدل ARMA در اینجا با نماد  $\theta = (a_1, \dots, a_p, b_1, \dots, b_q)^\top$  نشان داده می‌شود و فرض شده است که مرتبه‌های  $p$  و  $q$  نامعلوم هستند. علاوه بر این فرض کنید  $y_t$  ها مرکزی و  $x_t$  ها استاندارد شده‌اند. به عبارت دیگر،

$$\sum_{t=1}^T y_t = 0, \quad \sum_{t=1}^T x_{tj} = 0, \quad \frac{1}{T} \sum_{t=1}^T x_{tj}^2 = 1, \quad j = 1, \dots, p.$$

یون و همکاران (۲۰۱۷) روش جریمه‌ای وانگ و همکاران (۲۰۰۷) را برای انتخاب پیش‌گوهای مهم و برآورد ضرایب  $\beta$  و  $\theta$  با کمینه کردن لاسو سازوار به صورت

$$Q_T(\beta, \theta) = \sum_{t=p+1}^T \left\{ y_t - x_t^\top \beta - \sum_{j=1}^p a_j (y_{t-j} - x_{t-j}^\top \beta) + \sum_{j=1}^q b_j \epsilon_{t-j} \right\}^2 + T \sum_{j=1}^p \lambda_j |\beta_j| + T \sum_{j=1}^p \delta_j |a_j| + T \sum_{j=1}^q v_j |b_j| \quad (3.4)$$

تغییر دادند که  $T_0 = T - p$  اندازه نمونه موثر است و  $\lambda_j, \delta_j, v_j$  پارامترهای جریمه هستند. مقادیر  $(\hat{\beta}, \hat{\theta})$  در (۳.۴) که  $Q_T(\beta, \theta)$  را کمینه می‌کند، برآوردگر لاسو سازوار نامیده می‌شود. توجه کنید که روش لاسو سازوار، مقادیر مختلفی را به ضرایب متفاوت اختصاص می‌دهد تا از آریبی برآوردگر LASSO جلوگیری کند. متأسفانه در اجرای مستقیم برآورد جریمه‌شده با تابع هدف (۳.۴) دو مشکل اساسی وجود دارد: (۱) معادله (۳.۴) شامل سه پارامتر جریمه متفاوت برای هر پارامترهای رگرسیون خطی و مدل ARMA است و بنابراین، مسئله بهینه‌سازی به‌سادگی نمی‌تواند حل شود. (۲) یک دنباله از متغیرهای تصادفی غیرقابل مشاهده  $\epsilon_t$  را شامل می‌شود و هم‌چنین مرتبه‌های واقعی  $p$  و  $q$  خطاهای ARMA در رابطه (۳.۴) مشخص نیست. پیش از به‌کار بردن تابع هدف (۳.۴)،  $\epsilon_t$  و مرتبه‌های نامعلوم باید به‌طور مناسبی با باقی‌مانده‌ها و مرتبه‌های برآورد شده جایگزین شود. بدین منظور، **یون و همکاران (۲۰۱۷)** الگوریتم ALASSO گام به گام پیشنهاد کردند.

## ۱.۲.۴ الگوریتم‌های ALASSO جدید برای خطای مدل‌های ARMA و ARMA-GARCH

همان‌طور که اشاره شد ابتدا یک الگوریتم ALASSO جدید برای مدل‌های رگرسیون خطی با خطاهای ARMA (الگوریتم ALASSO-ARMA) معرفی می‌شود که می‌توان آن را به مدل‌های رگرسیون ARMA-GARCH (الگوریتم ALASSO-ARMA-GARCH) گسترش داد. در این الگوریتم پیشنهادی، پارامترهای  $\beta$  و  $\theta$  در ۳.۴ به ترتیب به‌عنوان پارامتر ساختاری و پارامتر مزاحم شناخته می‌شوند. بنابراین تمرکز بر روی انتخاب متغیرهای معنی‌دار رگرسیونی است و پارامترهای جریمه برای انتخاب مرتبه‌ها و برآورد پارامترهای خطاها با مدل ARMA پذیرفته نیست. برای محاسبه ساده‌تر، از آن جایی که خطاهای ARMA اهمیت کمتری دارند، تابع هدف (۴.۴) که فقط شامل پارامترهای رگرسیون است، کمینه می‌شود.

$$Q_T(\beta, \theta) = \sum_{t=p+1}^T \left\{ y_t - x_t^\top \beta - \sum_{j=1}^p a_j (y_{t-j} - x_{t-j}^\top \beta) + \sum_{j=1}^Q b_j \epsilon_{t-j} \right\}^2 + T_0 \sum_{j=1}^p \lambda_j |\beta_j| \quad (4.4)$$

برای الگوریتم جدید ALASSO، مسئله کمینه‌سازی در (۴.۴) در نقطه صفر مشتق پذیر نیست یا تابعی مقعر است. بنابراین از تقریب موضعی درجه دوم پیشنهادی **فن و لی (۲۰۰۱b)** به‌عنوان راه‌حلی برای این مسئله استفاده می‌شود. با استفاده از تقریب موضعی درجه دوم در  $\beta^*$   $(\beta_1^*, \dots, \beta_d^*)^\top$  عبارت جریمه‌شده می‌تواند به صورت تقریب موضعی در  $\beta_0^*$  با یک تابع درجه دوم به‌صورت

$$|\beta_j| \approx |\beta_j^*| + \frac{1}{2} \frac{(\beta_j^2 - \beta_j^{*2})}{|\beta_j^*|}$$



باشد. سپس مسئله کمینه‌سازی (۴.۴) در نزدیکی  $\beta^*$  را می‌توان به معادله درجه دوم (۵.۴)

$$Q_T^*(\beta, \theta) = \sum_{t=p+1}^T \left\{ y_t - x_t^\top \beta - \sum_{j=1}^p a_j (y_{t-j} - x_{t-j}^\top \beta) + \sum_{j=1}^Q b_j \epsilon_{t-j} \right\}^2 + T \sum_{j=1}^p \lambda_j \frac{\beta_j^2}{|\beta_j^*|}$$

کاهش داد.

اکنون می‌توان، الگوریتم ALASSO گام به گام جدید را با استفاده از برآوردگر معمولی کمترین توان‌های دوم بدون در نظر گرفتن ساختار خودهمبستگی به‌عنوان یک برآوردگر اولیه برای بردار ضریب رگرسیون  $\beta$  توسعه داد. به عبارت دیگر،

$$\hat{\beta}^{(0)} = (\mathbf{X}_t^\top \mathbf{X}_t)^{-1} \mathbf{X}_t^\top \mathbf{y}_t \quad (6.4)$$

که  $\mathbf{X}_t^\top = (x_1, \dots, x_T)$  و  $\mathbf{y} = (y_1, \dots, y_T)$ . با استفاده از  $\hat{\beta}^{(0)}$  به‌عنوان برآوردگر اولیه، الگوریتم ALASSO-ARMA از  $\ell = 1$  به‌صورت زیر شروع می‌شود:

**الگوریتم ۱.۲.۴ گام ۱:** با استفاده از  $\hat{\beta}^{(\ell-1)}$ ، باقی‌مانده رگرسیون به صورت

$$\tilde{\epsilon}_t^{(\ell-1)} = y_t - x_t^\top \hat{\beta}^{(\ell-1)}, \quad t = 1, \dots, T \quad (7.4)$$

تعریف می‌شود.

**گام ۲:** با استفاده از باقی‌مانده‌های رگرسیون  $\tilde{\epsilon}_t^{(\ell-1)}$  و معیار BIC مرتبه  $p^{(\ell-1)}$  و  $q^{(\ell-1)}$  به ترتیب برای  $p$  و  $q$  برآورد می‌شود. سپس برآوردگر ماکسیمم درست‌نمایی شرطی  $\hat{\theta}^{(\ell-1)}$  (CMLE) برای پارامتر  $\theta$  به‌صورت

$$\hat{\theta}^{(\ell-1)} = \left( \hat{a}_1^{(\ell-1)}, \dots, \hat{a}_{p^{(\ell-1)}}^{(\ell-1)}, \hat{b}_1^{(\ell-1)}, \dots, \hat{b}_{q^{(\ell-1)}}^{(\ell-1)} \right)^\top$$

به‌دست می‌آید. با استفاده از باقی‌مانده‌های  $\{\tilde{\epsilon}_t^{(\ell-1)}\}$  و برآوردگر ماکسیمم درست‌نمایی شرطی پارامتر  $\hat{\theta}^{(\ell-1)}$ ، باقی‌مانده‌های ARMA به‌صورت

$$\tilde{\epsilon}_t(\hat{\theta}^{(\ell-1)}) = \tilde{\epsilon}_t^{(\ell-1)} - \hat{a}_0^{(\ell-1)} - \sum_{i=1}^{P^{(\ell-1)}} \hat{a}_i^{(\ell-1)} \tilde{\epsilon}_{t-i}^{(\ell-1)} + \sum_{j=1}^{Q^{(\ell-1)}} \hat{b}_j^{(\ell-1)} \hat{\epsilon}_{t-j}(\hat{\theta}^{(\ell-1)}), \quad t = 1, \dots, T$$

تعریف می‌شود.

**گام ۳:** از کمینه کردن تابع هدف (۸.۴) به‌دست می‌آید.

$$Q_T^{(\ell-1)}(\beta) = \sum_{t=P^{(\ell-1)}+1}^T \left( y_t - x_t^\top \beta - \sum_{i=1}^{P^{(\ell-1)}} \hat{a}_i^{(\ell-1)} (y_{t-i} - x_{t-i}^\top \beta) + \sum_{j=1}^{Q^{(\ell-1)}} \hat{b}_j^{(\ell-1)} \hat{\epsilon}_{t-j}(\hat{\theta}^{(\ell-1)}) \right)^2$$

<sup>2</sup>Conditional Maximum Likelihood Estimator (CMLE)

$$+T_0^{(\ell-1)} \sum_{j=1}^P \lambda_j \frac{\beta_j^2}{|\hat{\beta}_j^{(i-1)}|}$$

که  $T_0^{(\ell-1)} = T - p^{(\ell-1)}$ . قرار دهید  $\ell = \ell + 1$  و به گام اول بازگردید تا زمانی که برای برخی از اعداد صحیح از پیش تعیین شده  $L > 0$ ،  $\ell < L + 1$  یا پارامترهای برآورد شده  $\hat{\beta}^{(\ell)}$  همگرا شوند.

**ملاحظه ۱.۲.۴.** در گام ۲ و گام ۳، مقادیر اولیه برای  $\tilde{e}_0, \dots, \tilde{e}_{p^{(\ell-1)}}$  و  $\tilde{e}_0, \dots, \tilde{e}_{1-q^{(\ell-1)}}$  باید ثابت در نظر گرفته شوند و نه می‌توانند تصادفی و نه تابعی از پارامترها باشند.

اکنون، با این مقدمه الگوریتم ALASSO را می‌توان به مدل‌های رگرسیون خطی با خطاهای ARMA-GARCH تعمیم داد که در (۱.۴) دنباله متغیرها به صورت  $ARMA(p, q)$ - $GARCH(r, s)$  یعنی

$$e_t = \sum_{i=1}^p a_i e_{t-i} + \epsilon_t + \sum_{j=1}^q b_j \epsilon_{t-j},$$

$$\epsilon_t = z_t \sigma_t,$$

$$\sigma_t^2 = c_0 + \sum_{i=1}^r c_i \epsilon_{t-i}^2 + \sum_{j=1}^s d_j \sigma_{t-j}^2.$$

هستند. پارامتر مدل ARMA-GARCH با نماد  $\varphi = (\theta_1^\top, \theta_2^\top)^\top$  نشان داده شده است که  $\theta_1 = (a_1, \dots, a_p, b_1, \dots, b_q)^\top$  و  $\theta_2 = (c_0, \dots, c_r, d_1, \dots, d_s)^\top$  و مقادیر واقعی آن با  $\varphi_0 = (\theta_{01}^\top, \theta_{02}^\top)^\top$  مشخص شده است که  $\theta_{01} = (a_{01}, \dots, a_{0p}, b_{01}, \dots, b_{0q})^\top$  و  $\theta_{02} = (c_{00}, \dots, c_{0r}, d_{01}, \dots, d_{0s})^\top$ . فضای پارامتر  $\varphi$  به صورت

$$\varphi \subseteq \mathbb{R}^{p+q+1} \times (0, +\infty) \times [0, +\infty)^{r+s}$$

است. در اینجا،  $z_t$ ،  $t = 1, \dots, T$  دنباله‌ای از متغیرهای تصادفی مستقل با میانگین صفر و واریانس یک هستند و مرتبه‌های  $p, q, r, s$  نامعلوم فرض شده است که باید تعیین شوند. با در نظر گرفتن یک اصلاح جزئی در گام دوم و گام سوم الگوریتم ALASSO، به منظور بازتاب ساختار خطی ARMA-GARCH، الگوریتم ALASSO، در برخی از برنامه‌های مالی، مناسب می‌باشد. با استفاده از برآوردگر کمترین توان‌های دوم  $\hat{\beta}^{(0)}$  به عنوان برآوردگر اولیه در (۶.۴)، الگوریتم ALASSO-ARMA-GARCH به‌روزشده از  $\ell = 1$  دارای گام‌های زیر است:

**الگوریتم ۲.۲.۴.** گام ۱: با استفاده از  $\hat{\beta}^{(\ell-1)}$ ، باقی‌مانده‌ها

$$\tilde{e}_t^{(\ell-1)} = y_t - \mathbf{x}_t^\top \hat{\beta}^{(\ell-1)}, \quad t = 1, \dots, T$$

را تعریف کنید.

گام ۲: با استفاده از باقی مانده‌ها  $\{\tilde{e}_t^{(\ell-1)}\}$  و معیار BIC مرتبه‌های  $p^{(\ell-1)}$  و  $q^{(\ell-1)}$  و  $r^{(\ell-1)}$  و  $s^{(\ell-1)}$  به ترتیب برای  $p, q, r$  و  $s$  به دست آورید. سپس برآوردگر شبه ماکسیمم درست‌نمایی<sup>۳</sup> (QMLE) برای پارامتر  $\varphi$  توسط  $\hat{\varphi}^{(\ell-1)} = (\hat{\theta}_1^{(\ell-1)}, \hat{\theta}_\varphi^{(\ell-1)})^\top$  پیشنهاد شد که با نماد  $\hat{\theta}_1^{(\ell-1)} = (\hat{a}_1^{(\ell-1)}, \dots, \hat{a}_{p^{(\ell-1)}}^{(\ell-1)}, \hat{b}_1^{(\ell-1)}, \dots, \hat{b}_{q^{(\ell-1)}}^{(\ell-1)})^\top$  و  $\hat{\theta}_\varphi^{(\ell-1)} = (\hat{c}_0^{(\ell-1)}, \dots, \hat{c}_{r^{(\ell-1)}}^{(\ell-1)}, \hat{d}_1^{(\ell-1)}, \dots, \hat{d}_{s^{(\ell-1)}}^{(\ell-1)})^\top$  نشان داده می‌شود که با استفاده از باقی مانده‌های QMLE و  $\tilde{e}_t^{(\ell-1)}$  برای پارامتر  $\hat{\theta}_1^{(\ell-1)}$ ، باقی مانده‌های ARMA به صورت بازگشتی به فرم

$$\tilde{e}_t(\hat{\theta}_1^{(\ell-1)}) = \tilde{e}_t^{(\ell-1)} - \sum_{i=1}^{p^{(\ell-1)}} \hat{a}_i^{(\ell-1)} \tilde{e}_{t-i}^{(\ell-1)} - \sum_{j=1}^{q^{(\ell-1)}} \hat{b}_j^{(\ell-1)} \tilde{e}_{t-j}(\hat{\theta}_1^{(\ell-1)}) \quad , \quad t = 1, \dots, T$$

تعریف می‌شوند.

گام ۳: مشاهدات  $\hat{\beta}^{(\ell-1)}$  تابع هدف

$$Q_T^{(\ell-1)}(\beta) = \sum_{t=p^{(\ell-1)}+1}^T \left( \frac{y_t - x_t^\top \beta - \sum_{i=1}^{p^{(\ell-1)}} \hat{a}_i^{(\ell-1)} (y_{t-i} - x_{t-i}^\top \beta) - \sum_{j=1}^{q^{(\ell-1)}} \hat{b}_j^{(\ell-1)} \tilde{e}_{t-j}(\hat{\theta}_1^{(\ell-1)})}{\tilde{\delta}_t(\hat{\varphi}^{(\ell-1)})} \right)^2 + T_0^{(\ell-1)} \sum_{j=1}^p \lambda_j \frac{\beta_j^2}{|\hat{\beta}_j^{(\ell-1)}|}$$

را کمینه می‌کند که  $T_0^{(\ell-1)} = T - p^{(\ell-1)}$  و  $\tilde{\sigma}_t^2(\hat{\varphi}^{(\ell-1)})$  و  $t = 1, \dots, T$  به صورت مجزا با استفاده از  $\tilde{e}_t(\hat{\theta}_1^{(\ell-1)})$  و

$$\tilde{\sigma}_t^2(\hat{\varphi}^{(\ell-1)}) = \hat{c}_0^{(\ell-1)} + \sum_{i=1}^{r^{(\ell-1)}} \hat{c}_i^{(\ell-1)} \tilde{e}_{t-i}^2(\hat{\theta}_1^{(\ell-1)}) + \sum_{j=1}^{s^{(\ell-1)}} \hat{d}_j^{(\ell-1)} \tilde{\delta}_{t-j}^2(\hat{\varphi}^{(\ell-1)})$$

تعریف شده‌اند. قرار دهید  $\ell = \ell + 1$  و به گام اول بازگردید تا زمانی که برای یک  $L \geq 0$  از قبل تعیین شده،  $\ell < L + 1$ ، یا پارامتر برآورد شده  $\hat{\beta}^\ell$  همگرا شود.

**ملاحظه ۲.۲.۴.** در گام ۲ و گام ۳، مقادیر اولیه برای  $\tilde{e}_0, \dots, \tilde{e}_{1-(r^{(\ell-1)}-q^{(\ell-1)})-p^{(\ell-1)}}$  و  $\tilde{e}_{1-\max\{r^{(\ell-1)}, q^{(\ell-1)}\}}, \dots, \tilde{e}_{\min\{0, -r^{(\ell-1)}+q^{(\ell-1)}\}}$  باید ثابت در نظر گرفته شوند و نه می‌توانند تصادفی و نه تابعی از پارامترها باشند.

مشابه **وانگ و همکاران (۲۰۰۷)**، برای هر دو مدل رگرسیونی با خطاهای ARMA و ARMA-GARCH، می‌توان از

$$\lambda_j = \lambda \frac{\log(T)}{T |\hat{\beta}_j^{(\circ)}|}$$

استفاده کرده که  $\hat{\beta}^{(\circ)} = (\hat{\beta}_1^{(\circ)}, \dots, \hat{\beta}_d^{(\circ)})$  برآوردگر کمترین توان‌های دوم تعریف شده در رابطه (۶.۴) است. به منظور انتخاب پارامتر تنظیم‌کننده در الگوریتم ALASSO، از معیار BIC به صورت

$$\text{BIC} = \log(\hat{\sigma}^2) + df \frac{\log(T)}{T}$$

<sup>3</sup>Quasi-Maximum Likelihood Estimator (QMLE)

استفاده می‌شود که  $\hat{\sigma}^2$  مجموع توان دوم خطاهای مدل تقسیم بر  $T$  است و  $\hat{df}$  تعداد ضرایب غیرصفر  $\hat{\beta}^{(\ell)}$  است. بنابراین، پارامتر تنظیم‌کننده به‌گونه‌ای انتخاب می‌شود که معیار BIC کمینه شود.

## ۳.۴ مطالعات عددی

### ۱.۳.۴ مطالعه شبیه‌سازی

در این بخش، برای بررسی عملکرد الگوریتم‌های پیشنهادی ALASSO-ARMA و ALASSO-ARMAGARCH بر نمونه‌های محدود از روش مونت کارلو استفاده می‌شود. الگوریتم پیشنهادی با الگوریتم ALASSO (ژو، ۲۰۰۶) مقایسه شده است. مدل رگرسیون خطی در (۱.۴) را با چهار خطای مختلف به‌صورت زیر در نظر گرفته شده است:

(۱) خطاهای i.i.d.

$$e_t = \varepsilon_t \quad (۸.۴)$$

(۲) خطاهای AR(۲):

$$e_t = 0.5e_{t-1} - 0.7e_{t-2} + \varepsilon_t \quad (۹.۴)$$

(۳) خطاهای ARMA(۱, ۱):

$$e_t = 0.6e_{t-1} + \varepsilon_t - 0.4\varepsilon_t \quad (۱۰.۴)$$

(۳) خطاهای ARMA(۱, ۱) – GARCH(۱, ۱):

$$\begin{aligned} e_t &= 0.6e_{t-1} + \varepsilon_t - 0.4\varepsilon_t \\ \varepsilon_t &= z_t \sigma_t \\ \sigma_t^2 &= c_0 + 0.1\varepsilon_{t-1}^2 + 0.8\sigma_{t-1}^2 \end{aligned} \quad (۱۱.۴)$$

اکنون، چهار مجموعه داده‌های مختلف از مدل‌های رگرسیونی با خطاهای مستقل و هم‌توزیع در (۸.۴)، خطاها AR(۲)، رابطه (۹.۴) که در شبیه‌سازی با ALASSO-AR نشان داده شده است، خطاها ARMA(۱, ۱)، رابطه (۱۰.۴) که از ALASSO-ARMA برای نمایش آن بهره گرفته می‌شود و خطاهای ARMA(۱, ۱) – GARCH(۱, ۱)، رابطه (۱۱.۴) که با ALASSO-ARMA-GARCH نشان داده شده است، تولید گردیده است که  $\beta = (3, 1/5, 0, 0, 2, 0, 0, 0)^T$  و متغیرهای  $x_t = (x_{t1}, \dots, x_{t8})^T$  به‌طور مستقل از توزیع نرمال چندمتغیره با میانگین  $0_{8 \times 1}$  و همبستگی دوجه‌دو بین  $x_{t,j_1}$  و  $x_{t,j_2}$  برابر با  $0.5^{|j_1 - j_2|}$ ، تولید می‌شود. مقادیر ۰/۵ و ۳ برای  $\sigma$  در (۸.۴) تا (۱۰.۴) و ۰/۵ و ۳ برای  $c_0$  در (۱۱.۴) و  $T = 100, 300$  به‌عنوان اندازه نمونه در

نظر گرفته شده است. مشابه در **وانگ و همکاران (۲۰۰۷)**، پارامترهای تنظیم‌کننده بهینه از نقاط  $\{0, 0.01, 0.05, 0.1, 1, 10, 50\}$  با استفاده از معیار BIC انتخاب شده است. برای انتخاب مرتبه‌های AR و ARMA از  $p, q = 1, 2, 3, 4, 5$  و برای انتخاب مرتبه مدل GARCH از  $r, s = 1, 2$  استفاده شده است. هر برآورد پارامتری که قدرمطلق آن از  $10^{-6}$  کمتر باشد، به سمت صفر منقبض شده است یعنی متغیرهای رگرسیون متناظر حذف می‌شوند. شبیه‌سازی  $1000$  بار تکرار شده است و میانگین خطای مدل  $^4$  (ME)  $ME = (\hat{\beta} - \beta)^T E[(x_t^T x_t)(\hat{\beta} - \beta)]$  و درصد تعداد برآوردهای صحیح ضرایب رگرسیونی مقایسه شده است.

جدول‌های ۱.۴ تا ۸.۴ نتایج شبیه‌سازی را نشان می‌دهند.

بر اساس جدول‌های ۱.۴ و ۲.۴، وقتی خطاهای i.i.d. در نظر گرفته شده است، برآوردگر ALASSO بهترین کارایی را در بین همه برآوردگرها دارد. با این وجود، برآوردگر پیشنهادی ALASSO **یون و همکاران (۲۰۱۷)** با خطاهای AR و ARMA اندکی بدتر از برآوردگر ALASSO عمل کرده‌اند. دقت کنید که در الگوریتم ALASSO-AR، فقط جمله AR در نظر گرفته شده و همه ضرایب مدل MA در الگوریتم ALASSO-ARMA صفر فرض می‌شود.

جدول ۱.۴: جدول میانگین ME و انحراف معیار آن‌ها برای مدل رگرسیون با خطاهای i.i.d.

$\sigma = 3$	$\sigma = 0.5$	
		$T = 100$
0.5971(0.5113)	0.0089(0.0072)	ALASSO
0.0092(0.0079)	0.0091(0.0081)	ALASSO-AR
0.6001(0.4998)	0.0094(0.0080)	ALASSO-ARMA
		$T = 300$
0.1401(0.1310)	0.0031(0.0021)	ALASSO
0.1408(0.1318)	0.0030(0.0020)	ALASSO-AR
0.1419(0.1317)	0.0030(0.0021)	ALASSO-ARMA

از جدول‌های ۳.۴ تا ۸.۴ چنین برمی‌آید که:

(۱) برآوردگر ALASSO به‌خوبی عمل نمی‌کند زیرا وابستگی در خطاها در ساختار آن به‌طور مناسبی بازتابی ندارد.

(۲) الگوریتم پیشنهادی **یون و همکاران (۲۰۱۷)** به‌درستی متغیرهای معنی‌دار و بااهمیت با ضرایب غیرصفر در مدل و مرتبه‌های ARMA-GARCH را به‌طور همزمان انتخاب می‌کند، به‌ویژه، برای واریانس کمتر یا حجم نمونه بزرگ‌تر، تعداد ضرایب صفر بیشتری به‌طور موثری به صفر منقبض می‌شود.

<sup>4</sup>Model Error (ME)

جدول ۲.۴: جدول درصد تعداد ضرایب مدل رگرسیونی با خطاهای i.i.d که به طور صحیح برآورد شده‌اند

$\sigma = 3$		$\sigma = 0.5$		
i.i.d.	REG	i.i.d.	REG	
				$T = 100$
-	۷۳/۸	-	۸۸/۰	ALASSO
۹۰/۴	۷۳/۵	۹۰/۱	۸۶/۳	ALASSO-AR
۹۰/۱	۷۳/۱	۹۰/۰	۸۵/۹	ALASSO-ARMA
				$T = 300$
-	۸۷/۹	-	۹۱/۳	ALASSO
۹۲/۱	۸۸/۳	۹۲/۳	۹۱/۰	ALASSO-AR
۹۰/۲	۹۱/۰	۹۰/۱	۹۱/۲	ALASSO-ARMA

- (۳) کارایی روش پیشنهادی **یون و همکاران (۲۰۱۷)** به طور معنی داری انتخاب متغیر را در رگرسیون با افزایش نمونه از ۱۰۰ به ۳۰۰ نشان می‌دهند.
- (۴) در حالت واریانس بالا، روش پیشنهادی **یون و همکاران (۲۰۱۷)** منجر به برآوردهای دقیق و انتخاب متغیرهای بااهمیت و مرتبه‌های ARMA-GARCH می‌شود.
- (۵) تمام نتایج، از مزیت الگوریتم‌های ALASSO پیشنهادی **یون و همکاران (۲۰۱۷)** در انتخاب متغیر مدل‌های رگرسیونی با خطای ARMA یا ARMA-GARCH حمایت می‌کنند.

جدول ۳.۴: جدول میانگین ME و انحراف معیار آن‌ها برای مدل رگرسیون با خطاهای AR

$\sigma = 3$		$\sigma = 0.5$		
				$T = 100$
۱/۴۱۰۲(۱/۴۲۰۳)	۰/۰۵۳۸(۰/۰۴۱۲)			ALASSO
۰/۴۱۱۹(۰/۴۰۰۳)	۰/۰۰۷۱(۰/۰۰۴۷)			ALASSO-AR
				$T = 300$
۰/۴۰۰۹(۰/۳۳۳۷)	۰/۰۰۳۹(۰/۰۰۴۱)			ALASSO
۰/۰۸۱۶(۰/۰۵۱۶)	۰/۰۰۲۱(۰/۰۰۱۷)			ALASSO-AR

جدول ۴.۴: جدول میانگین ME و انحراف معیار آن‌ها برای مدل رگرسیون با خطاهای ARMA

$\sigma = 3$		$\sigma = 0.5$		
				$T = 100$
1/7003 (1/3021)		0/0463 (0/0401)		ALASSO
0/5103 (0/6101)		0/0081 (0/0093)		ALASSO-ARMA
				$T = 300$
0/4219 (0/3875)		0/0129 (0/0132)		ALASSO
0/1313 (0/1118)		0/0031 (0/0029)		ALASSO-ARMA

جدول ۵.۴: جدول میانگین خطای مدل (ME) و انحراف معیار آن‌ها برای مدل رگرسیون با خطاهای ARMA-GARCH

$\sigma = 3$		$\sigma = 0.5$		
				$T = 100$
8/015 (5/362)		0/8514 (0/6156)		ALASSO
3/3698 (2/1236)		0/3162 (0/3666)		ALASSO-ARMA-GARCH
				$T = 300$
2/2199 (1/9899)		0/2501 (0/1999)		ALASSO
0/6239 (0/5001)		0/0651 (0/0642)		ALASSO-ARMA-GARCH

جدول ۶.۴: جدول درصد تعداد ضرایب مدل رگرسیونی با خطاهای AR که به‌طور صحیح برآورد شده‌اند

$\sigma = 3$		$\sigma = 0.5$		
AR	REG	AR	REG	
				$T = 100$
-	0/57	-	50/1	ALASSO
95/9	54/3	96/1	91/8	ALASSO-AR
				$T = 300$
-	12/8	-	56/9	ALASSO
96/9	74/9	95/9	98/1	ALASSO-AR

جدول ۷.۴: جدول درصد تعداد ضرایب مدل رگرسیونی با خطاهای ARMA که به طور صحیح برآورد شده اند

$\sigma = 3$		$\sigma = 0.5$		
ARMA	REG	ARMA	REG	
				$T = 100$
-	0.7	-	50.1	ALASSO
31.9	47.1	34.9	90.7	ALASSO-ARMA
				$T = 300$
-	4.9	-	57.0	ALASSO
86.7	71.8	82.7	96.9	ALASSO-ARMA

جدول ۸.۴: جدول درصد تعداد ضرایب مدل رگرسیونی با خطاهای ARMA-GARCH که به طور صحیح برآورد شده اند

$\sigma = 3$		$\sigma = 0.5$		
ARMA-GARCH	REG	ARMA-GARCH	REG	
				$T = 100$
-	3.2	-	2.8	ALASSO
72.3	15.1	72.8	31.9	ALASSO-ARMA-GARCH
				$T = 300$
-	10.7	-	24.1	ALASSO
87.5	37.9	87.6	49.0	ALASSO-ARMA-GARCH

## ۴.۴ تحلیل مثال واقعی

در این بخش، یک مجموعه داده مربوط به مصرف برق متقاضیان شرکت برق و گاز سن دیگو (رامانازان، ۱۹۹۸) در نظر گرفته شده است. این مجموعه شامل ۸۷ داده فصلی، از فصل دوم سال ۱۹۷۲ تا فصل آخر سال ۱۹۹۳ است. متغیر پاسخ، لگاریتم مصرف برق هر مشترک بر حسب میلیون کیلووات ساعت (LKWH) است و متغیرهای توضیحی عبارتند از: سرانه درآمد (LY)، لگاریتم قیمت برق (LPRICE)، روزهای سرد (CDD) و روزهای گرم (HDD). مدل رگرسیون خطی مشابه رامانازان (۱۹۹۸) و وانگ و همکاران (۲۰۰۷) به صورت

$$LWKH = \beta_0 + \beta_1 LY + \beta_2 LPRICE + \beta_3 CDD + \beta_4 HDD + e_t \quad (12.4)$$



در نظر گرفته شده است که علامت‌های مورد انتظار برای  $\beta$ ها بر اساس **رامانازان (۱۹۹۸)** عبارت‌اند از:

$$\beta_1 > 0, \quad \beta_2 < 0, \quad \beta_3 > 0, \quad \beta_4 > 0.$$

برای مطالعه مقایسه‌ای، سه ساختار متفاوت خطا در مدل‌های رگرسیون خطی را در نظر گرفته شده است. ابتدا متغیرهای تصادفی مستقل و هم‌توزیع با میانگین صفر و واریانس  $\sigma^2$  برای  $\{e_t\}$  در نظر گرفته و سپس روش کمترین توان‌های دوم خطا به کار برده شده است. سپس مدل‌های رگرسیونی با خطای AR (REGAR)، ARMA (REGARMA) و خطای ARMA-GARCH (REGARMA-GARCH) فرض شده است که از الگوریتم پیشنهادی برای برآورد پارامترها استفاده شده است. برای انتخاب مرتبه‌های ARMA و GARCH،  $p, q = 1, 2, 3, 4$  و  $r, s = 1, 2$  استفاده شده است. دقت کنید که مرتبه  $q$  برای مدل‌های REGAR در الگوریتم، صفر در نظر گرفته می‌شود. مدل‌های REGAR فقط برای مقایسه استفاده شده است که نشان داده شود که ممکن است مدل‌های REGAR برای بعضی از داده‌های رگرسیون با خطاهای سری زمانی مناسب نباشند.

نتایج تحلیل داده‌ها با روش OLS برای مدل‌های رگرسیون خطی با خطاهای i.i.d و الگوریتم لاسو سازوار برای مدل‌های REGAR، REGARMA، REGARMA-GARCH در جدول ۹.۴ گزارش شده است.

جدول ۹.۴: جدول برآورد ضرایب، مرتبه‌های مدل‌های ARMA و GARCH و مقادیر BIC بر اساس باقی‌مانده‌ها

متغیر	ALASSO ARMA-GARCH	ALASSO ARMA	ALASSO AR	ALASSO
LY	۰	۰	۰/۰۸۴۰۴	-۰/۰۳۶۲۷
LPRICE	۰	۰	-۰/۰۹۷۸۷	-۰/۰۹۴۲۶
CCD	۰/۰۰۰۲۵	۰/۰۰۰۱۸	۰/۰۰۰۲۸	۰/۰۰۰۲۷
HDD	۰/۰۰۰۳۶	۰/۰۰۰۱۳	۰/۰۰۰۲۲	۰/۰۰۰۳۶
مرتبه ARMA	(۳, ۳)	(۴, ۳)	(۴, ۰)	-
مرتبه GARCH	(۱, ۰)	-	-	-
BIC	-۷/۱۵۲۲۴	-۷/۳۱۷۳۰	-۷/۱۲۲۲۹	

شکل‌های ۱.۴ تا ۴.۴، ACF و نمونه<sup>۵</sup> (SACF) و PACF نمونه<sup>۶</sup> از باقی‌مانده‌های حاصل از OLS و لاسو سازوار برای مدل‌های REGARMA، REGARMA-GARCH، REGARMA را نشان می‌دهند.

بر اساس جدول ۹.۴ و شکل‌های ۱.۴ تا ۴.۴ نتایج زیر به دست می‌آید:

(۱) برآوردهای OLS رضایت‌بخش نیست زیرا علامت ضریب برآورد شده متغیر LY منفی است در صورتی که علامت مورد انتظار در **رامانازان** (۱۹۹۸)، مثبت بود.

(۲) فرضیه استقلال خطاها بر اساس شکل ۱.۴ به نظر می‌رسد که تحت تاثیر همبستگی سریالی قوی در بین باقی‌مانده‌ها شدیداً نقض شده است.

(۳) بدیهی است که برآوردهای ALASSO ضرایب رگرسیونی با خطاهای AR نسبت به برآوردگر OLS کاملاً بهبود یافته است.

(۴) در شکل ۲.۴، همبستگی سریالی کمی در نمودارهای ACF و PACF دیده می‌شود. بنابراین، عبارت خطای AR ممکن است برای نشان داده وابستگی کامل در بین مشاهده‌ها کافی نباشد.

(۵) علامت‌های ضرایب برآورد شده برای مدل‌های رگرسیونی ALASSO با خطاهای ARMA و ARMA-GARCH علامت‌های یکسانی دارند و ضرایب متغیرهای LY و LPRICE دقیقاً برابر صفر برآورد شده‌اند.

(۶) به نظر می‌رسد مدل  $ARMA(4,3)$  مناسب‌ترین مدل باشد، زیرا کمترین مقدار BIC را دارد و بر اساس نمودار تابع‌های SACF و SPACF در شکل ۳.۴، باقی‌مانده‌های مدل ALASSO با خطاهای ARMA همبستگی‌های سریالی ندارند.

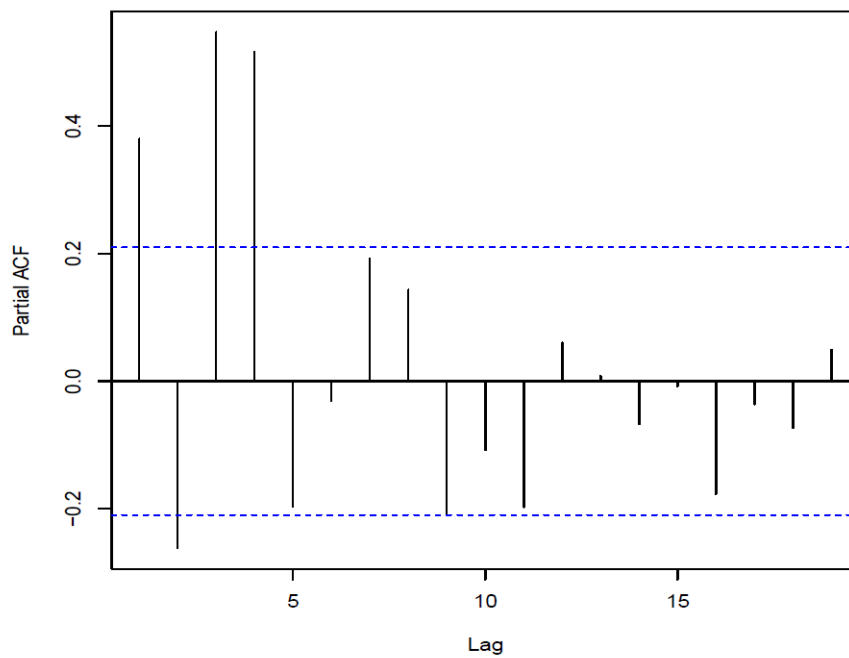
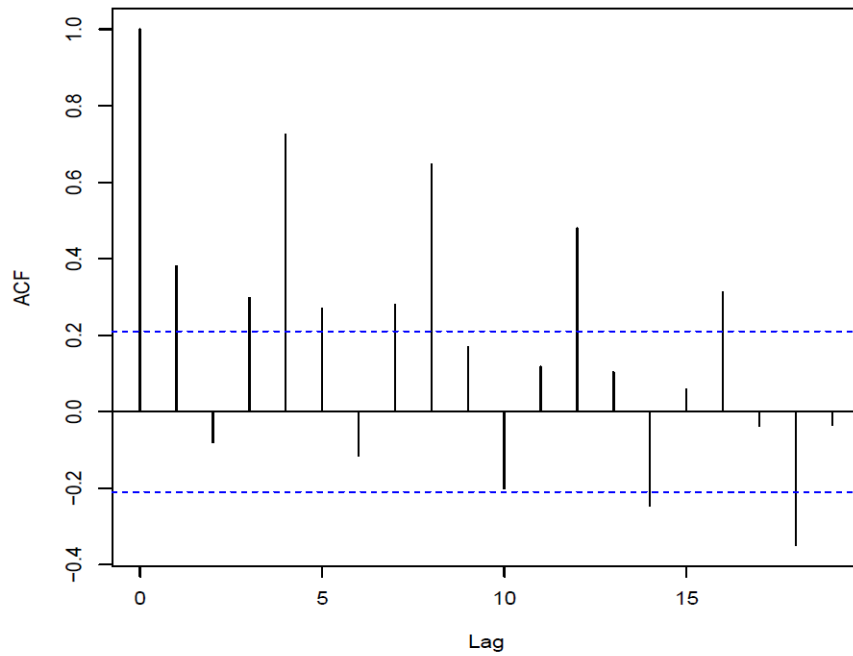
جدول ۱۰.۴، زمان محاسبات بر حسب ثانیه برای اجرای روش‌های برآورد در جدول ۹.۴ نشان می‌دهد. الگوریتم پیشنهادی **یون و همکاران** (۲۰۱۷) نسبت به برآوردگر OLS از هزینه محاسباتی خیلی بالایی برخوردار است. با این وجود نتایج جدول ۹.۴ و شکل‌های ۱.۴ تا ۴.۴ اهمیت و مزیت الگوریتم ALASSO را برای مدل‌های رگرسیونی با خطاهای ARMA و ARMA-GARCH در حضور همبستگی سریالی شدید در بین مشاهده‌ها نشان می‌دهند.

جدول ۱۰.۴: زمان محاسبات (ثانیه) برای اجرای روش‌های جدول ۹.۴

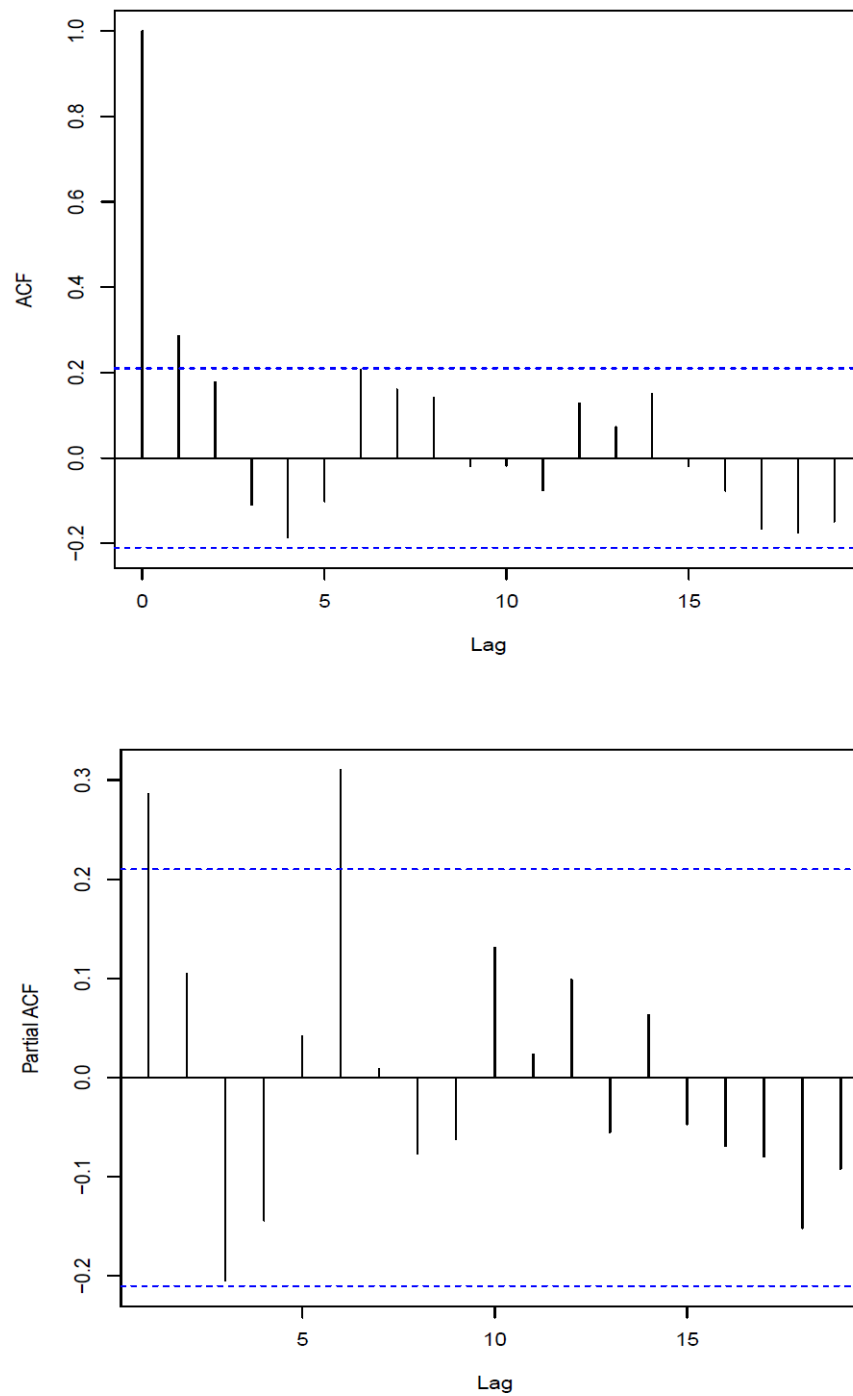
ALASSO	ALASSO	ALASSO	ALASSO
ARMA-GARCH	ARMA	AR	
۱۸۶۹/۵۱	۱۳۶/۵۲	۶/۹۸	۰/۲۵

<sup>5</sup>Sample Auto-Correlation Function (SACF)

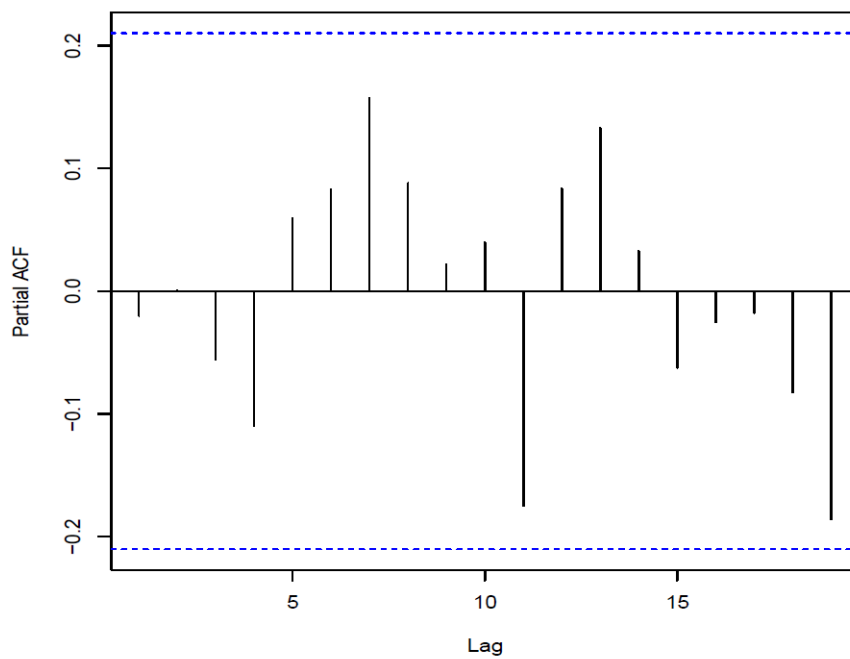
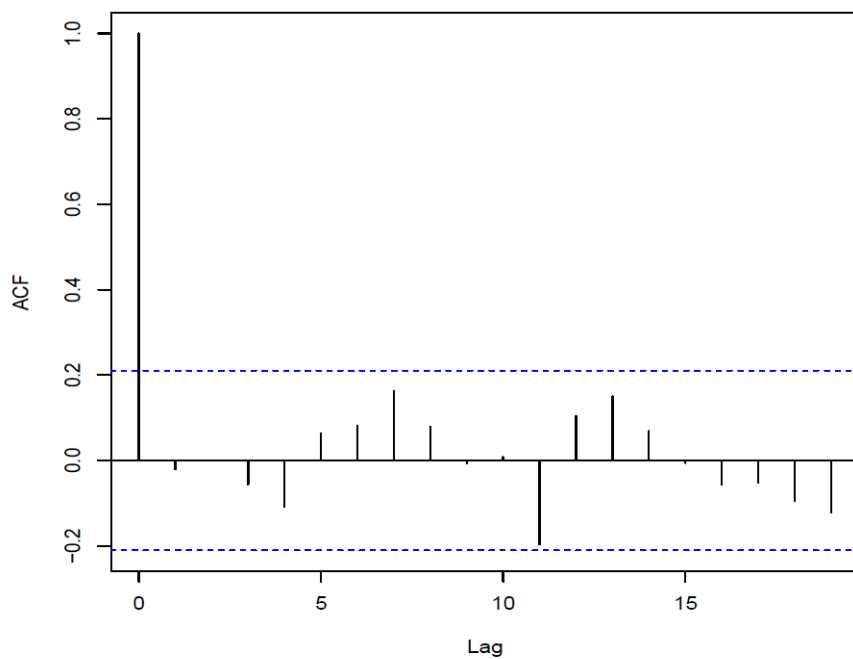
<sup>6</sup>Sample Partial Auto-Correlation Function (SACF)



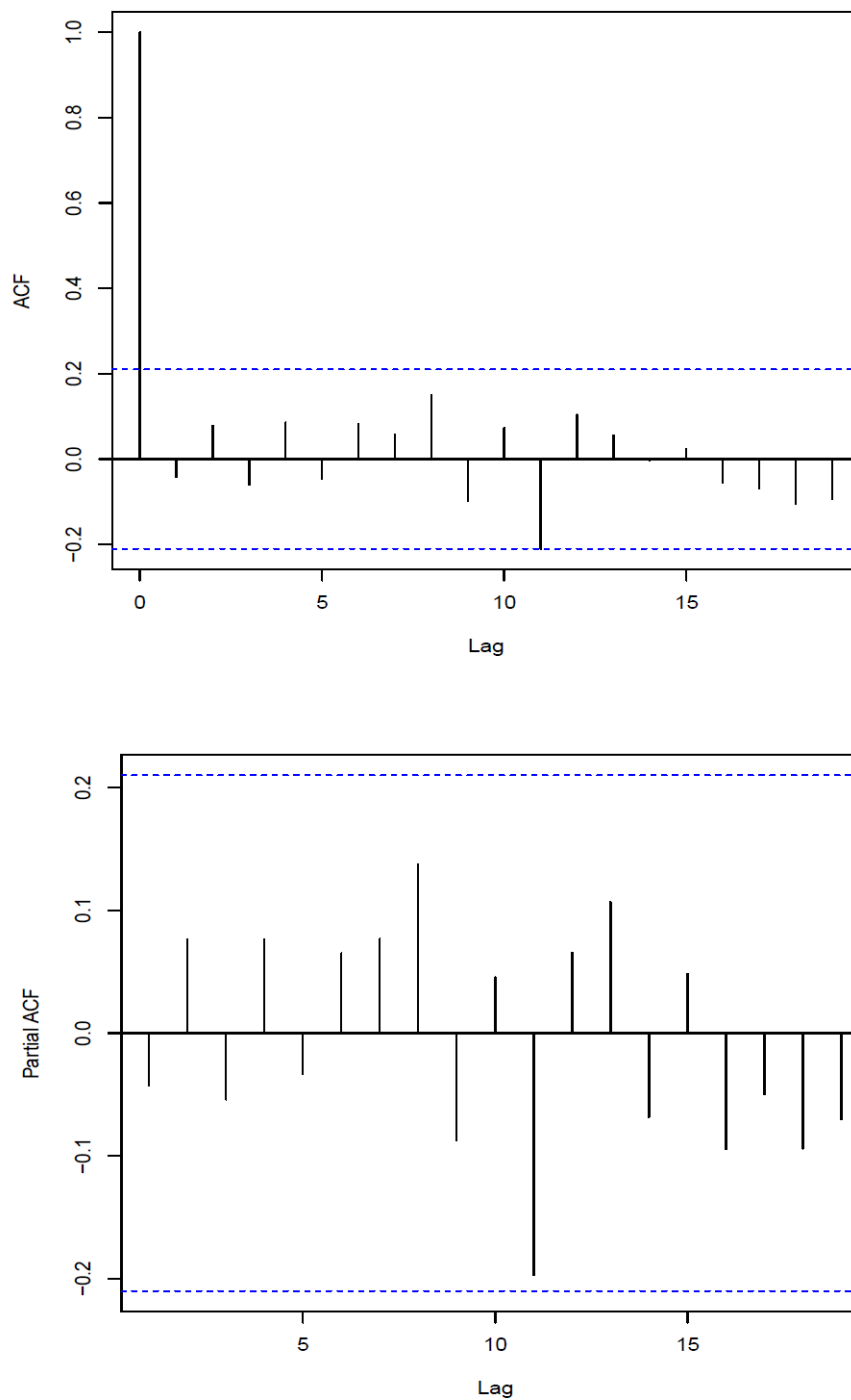
شکل ۱.۴: نمودارهای SACF و SPACF باقی‌مانده از OLS



شکل ۲.۴: نمودارهای SACF و SPACF باقی مانده از رگرسیون با خطای AR



شکل ۳.۴: نمودارهای SACF و SPACF باقی‌مانده از رگرسیون با خطای ARMA



شکل ۴.۴: نمودارهای SACF و SPACF باقی مانده از رگرسیون با خطای ARMA

## ۵.۴ نتیجه‌گیری و آینده تحقیق

در مدل‌سازی سری‌های زمانی داده‌های اقتصادی، تغییرپذیری نیازمند توجه به تغییر واریانس در طول دوره مورد بررسی است. واریانس در تمام دوره‌ها لزوماً ثابت نیست و متناسب با زمان تغییر می‌کند. مدل‌های ARCH و تعمیم‌های آن از جمله متداول‌ترین مدل‌ها برای نوسانات سری‌های زمانی داده‌های اقتصادی با فرض ناهمسانی واریانس می‌باشند. از این رو، مدل‌های رگرسیون خطی با خطاهای ARMA اغلب به‌منظور بازتاب همبستگی سریالی در مشاهده‌ها در نظر گرفته می‌شود. هدف اصلی این پایان‌نامه، بررسی روش لاسو سازوار برای رگرسیون با خطای ARMA-GARCH است که توسط **یون و همکاران (۲۰۱۷)** پیشنهاد شده است که نتایج نشان داد این الگوریتم به‌خوبی می‌تواند متغیرهای بااهمیت و معنی‌دار را در مدل‌های رگرسیونی با خطاهای ARMA و ARMA-GARCH تعیین کند.

می‌توان نتایج این پایان‌نامه را

- (۱) به مدل‌های سری‌زمانی دیگر مانند ARIMA و ARFIMA تعمیم داد.
- (۲) می‌توان به جای مدل GARCH یکی از اعضای دیگر از خانواده‌های مدل GARCH مانند EGARCH (نلسون، ۱۹۹۱)، APGARCH (دیانگ و همکاران، ۱۹۹۶)، GARCH-M (ک.دالی، ۲۰۰۶) و IGARCH (انگل و بولرسلو، ۱۹۸۶) جایگزین نمود.
- (۳) به جای روش جریمه‌شده ALASSO می‌توان از جریمه‌های دیگری مانند الاستیک‌نت (ژو و هیستی، ۲۰۰۵)، الاستیک‌نت سازوار (ژو و ژانگ، ۲۰۰۹)، لاسو گروهی (یوان و لین، ۲۰۰۶)، SCAD (فن و لی، ۲۰۰۱b)، بریج (فرانک و فریدمن، ۱۹۹۳) و ... بنا به ویژگی‌های داده‌ها و هدف مسئله مورد بررسی استفاده نمود.





# پیوست آ

## گزیده‌ای از برنامه‌های رایانه‌ای

### نمودار ۱.۱

```
da <- read.table("m-dec12910.txt", quote="\\"", comment.char="", header = T)
d10=da$dec10 # select the Decile 10 returns
dec10=ts(d10,frequency=12,start=c(1967,1))
plot(dec10,xlab='year',ylab='returns')
title(main="")
```

### نمودار ۲.۱

```
library(TTR)
decomp <- decompose(dec10)
plot(decomp$trend,ylab="Trend")
```

### نمودار ۳.۱

```
plot(decomp$seasonal,ylab="Seasonal")
```

#### نمودار ۴.۱

```
acf(d10,lag=24,main="")
```

#### نمودار ۵.۱

```
da=read.table("m-ibm3dx2608.txt",header=T)
ew=da$ewrtn
ew<-ts(da$ewrtn, frequency=12, start=c(1926, 1))
plot.ts(ew, ylab = "Equal Weighted Return")
acf(ew,main="")
m1=arima(ew,order=c(0,0,9),fixed=c(NA,0,NA,0,0,0,0,0,NA,NA))
m1
Box.test(m1$residuals,lag=12,type='Ljung')
pv=1-pchisq(17.6,9)
pv
```

#### نمودار ۶.۱

```
ibm <- read.table('m-ibm3dx2608.txt',header=T)
vw<-ts(ibm$vwrtn, frequency=12, start=c(1926, 1))
plot.ts(vw, ylab = "Value Weighted Return")
acf(vw,main="")
pacf(vw,main="")
m3=arima(vw,order=c(3,0,0),fixed=c(NA,0,NA,NA))
m3
Box.test(m3$residuals,lag=12,type='Ljung')
pv=1-pchisq(16.83,10)
pv
```

#### نمودار ۷.۱

```
da <- read.csv("J:/Thesis/ch2data/m-3m4608.txt", sep="")
head(da)
dim(da)
mmm=log(da$rtn+1)
mmm=ts(log(da$rtn+1))
plot(mmm)
library(TSA)
eacf(mmm)
```

## نمودار ۱.۲

```
rm(list=ls())
library(rugarch)
library(parallel)
set.seed(123)
arch1.spec = ugarchspec(variance.model=list(garchOrder= c(1,0)),
mean.model = list(armaOrder=c(0,0)),fixed.pars = list(mu = 0, omega=0.1,
alpha1=0.8))
class(arch1.spec)
show(arch1.spec)
set.seed(123)
arch1.sim = ugarchpath(arch1.spec, n.sim = 1000)
class(arch1.sim)
slotNames(arch1.sim)
plot(arch1.sim)
```

## نمودار ۲.۲

```
rm(list=ls())
library(rugarch)
library(parallel)
set.seed(123)
arch1.spec = ugarchspec(variance.model=list(garchOrder= c(1,0)),
mean.model = list(armaOrder=c(0,0)),fixed.pars = list(mu = 0, omega=0.1,
```

```

alpha1=0.8))
class(arch1.spec)
show(arch1.spec)
set.seed(123)
arch1.sim = ugarchpath(arch1.spec, n.sim = 1000)
class(arch1.sim)
slotNames(arch1.sim)
plot(arch1.sim)

```

## الگوریتم ARMA-GARCH

```

rm(list=ls())
library(timeDate)
library(timeSeries)
library(fBasics)
library(fGarch)
library(MASS)
spec = garchSpec(model = list(ar =0, ma = 0))
x1 <- garchSim(spec, n <- 30)
x2 <- garchSim(spec, n )
x3 <- garchSim(spec, n )
x4 <- garchSim(spec, n )
x5 <- garchSim(spec, n )
x6 <- garchSim(spec, n )
x7 <- garchSim(spec, n )
x8 <- garchSim(spec, n )
xtotal <- cbind(x1,x2,x3,x4,x5,x6,x7,x8)
beta0 <-c(3,1.5,0,0,2,0,0,0)
P=length(beta0)
yttotal= xtotal %*% beta0
yttotal=scale(yttotal,center = TRUE)
xttotal=scale(xtotal,center = T,scale=T)
beta0=solve(t(xtotal)%*%xttotal)%*%(t(xtotal)%*%yttotal)
bet=t(beta0)%*%beta0

```

```

bet
lambda=0.1

calcNum=0
#####

sigma2=1
#####
while(bet>0.1){
calcNum=calcNum+1
etotal=ytotal-xtotal%*%beta0

##### AIC #####

aicMat=matrix(numeric(4),ncol=2,nrow=2)
sapply(0:1, function(i){
sapply(0:1,function(j){
aicMat[i+1,j+1]<<-AIC(arima(etotal,order = c(i,0,j),method = "ML"))
})
})

current_order=which(aicMat==min(aicMat),arr.ind=T)-c(1,1)
current_theta=coef(arima(etotal,order = c(current_order[1],0,current_order[2]),
method = "ML"))
a0=current_theta[length(current_theta)]
if(current_theta>1){
current_theta=current_theta[-length(current_theta)];
}else{
current_theta=0
}

if(any(substr(names(current_theta),start = 1,stop = 2)=="ar")){
ar_index<-which(substr(names(current_theta),start = 1,stop = 2)=="ar")
}else{
ar_index<-0

```

---

```

}
if(any(substr(names(current_theta),start = 1,stop = 2)=="ma")){
ma_index<-which(substr(names(current_theta),start = 1,stop = 2)=="ma")
}else{
ma_index<-0
}

# ma_index=which(substr(names(current_theta),start = 1,stop = 2)=="ma")
# ##### epsilon #####
# if(firstshow==T){
#   epsilon=rnorm(length(ma_index)+length(ytotal),mean=0,sd=sigma2)
#   firstshow=F
# }

epsilon=numeric(length = n+length(ma_index))
etotal_new=c(rep(0,length(ar_index)),etotal)
for(tt in (length(ma_index)+1):n){
epsilon[tt]=etotal[tt]-a0-t(current_theta[ar_index+1])%*%
  etotal_new[(tt-length(ma_index)):(tt-length(ma_index)+length(ar_index)-1)]+
t(current_theta[ma_index+1])%*% epsilon[(tt-length(ma_index)):(tt-1)]
}
Lambda=lambda*log(n)/(n*abs(beta0))

z=rnorm(length(epsilon)+1)
epsilon_0=z[1]*sigma2
arma_model=arima(epsilon,order = c(1,0,1))
# arma_model$coef
ar_index_4sigma= arma_model$coef[which(substr(names(arma_model$coef),start = 1,
stop = 2)=="ar")]
ma_index_4sigma=arma_model$coef[which(substr(names(arma_model$coef),
start = 1,stop = 2)=="ma")]
intercept_index_4sigma=arma_model$coef[which(substr(names(arma_model$coef),
start = 1,stop = 3)=="int")]

```

```

Epsilon=c(epsilon_0,epsilon)

f_sigma2=function(ts){
sigma2=intercept_index_4sigma+sum(ma_index_4sigma*Epsilon[1:(ts-1)])
+sum(ar_index_4sigma*(Epsilon[1:(ts-1)]/z[2:ts]))
sqrt(sigma2)
}
#####
Qfun=function(beta){
f=function(ttt){
((ytotal[ttt]-xtotal[ttt,]*%*%beta-sum(ytotal[length(ar_index):1]
-xtotal[length(ar_index):1,]*%*%beta)+
t(current_theta[ma_index])*%*% epsilon[(ttt-length(ma_index)):(ttt-1)])/
(f_sigma2(ttt)))^2
}
f=Vectorize(f)
sum(f((length(ar_index)+1):n))+(n-length(ar_index))*sum(Lambda*beta/abs(beta0))
}
###
beta0 = nlm(Qfun,beta0)
beta0=beta0$estimate
bet=t(beta0)*%*%beta0
}

bet
calcNum

```

# مراجع

- باکس، ج. ای. بی.، گ. ام. جنکینز، و گ. س. رینسال (۱۳۹۴). تحلیل سری‌های زمانی: پیش‌بینی و کنترل. ترجمه‌ی مشکانی، م. ر. تهران: دانشگاه صنعتی شریف.
- بیدرام، رسول (۱۳۸۱). *Eviews* همگام با اقتصادسنجی. منشور بهره‌وری.
- حیرانی، مهرداد. و نسیم. روشن ضمیر (۱۳۹۷). مدل سازی سری‌های زمانی مالی با *R*. انتشارات بورس وابسته به شرکت اطلاع رسانی و خدمات بورس.
- ذوالفقاری، م. و ح. صادقی (۱۳۹۰). مبانی مدل‌های پیش‌بینی با در علوم اقتصادی. نور علم.
- سوری، علی (۱۳۸۹). اقتصادسنجی همراه با کاربرد نرم‌افزار *Eviews*. انتشارات فرهنگ‌شناسی.
- کشاوری حداد، غلامرضا (۱۳۹۴). اقتصادسنجی سری‌های زمانی مالی. نشر نی.
- ماکریداکیس، ر.، س. ویل‌رایت، و ر. هایندمن (۱۳۹۴). پیش‌بینی روش‌ها و کاربردها. ترجمه‌ی شاهسونی، د. و م. آرشی. شاهرود: دانشگاه صنعتی شاهرود.
- مجرد، محمد حسین (۱۳۹۵). اندازه‌گیری ریسک نقدینگی در موسسات مالی ایران با استفاده از معیار نقدینگی در معرض خطر (مقایسه عملکرد مدل‌های گارچ). پایان‌نامه کارشناسی ارشد، پایان‌نامه کارشناسی ارشد، دانشکده فنی و مهندسی دانشگاه علم و فرهنگ. فرهنگ.
- معروف‌خانی، علی (۱۳۷۷). بررسی ارتباط کسری بودجه دولت و تورم در اقتصاد ایران. پایان‌نامه کارشناسی ارشد، دانشگاه شهید بهشتی تهران.
- نوفرستی، محمد (۱۳۷۸). ریشه واحد و همجمعی در اقتصادسنجی. موسسه خدمات فرهنگی رسا.
- واعظ، آمنه (۱۳۹۰). مدل‌های *arch* و *garch* در اندازه‌گیری تلاطم. پایان‌نامه کارشناسی ارشد، پایان‌نامه کارشناسی ارشد، دانشکده علوم ریاضی و کامپیوتر دانشگاه شیخ بهایی.



- Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics* 31(3), 307–327.
- Bollerslev, T., R. Chou, and K.F.Kroner (1992). Arch modelling in finance: A review of the theory and empirical evidence. *Journal of Econometrics* 52, 5–29.
- Bollerslev, T., R. Engle, and J.M.Wooldridge (1998). A capital asset pricing model with time-varying covariances. *political Economy* 96, 116–131.
- Box, G. and G. Jenkins (1970). *Time Series Analysis: Forecasting and Control*. Holden-Day.
- Box, G. E. P. and D. A. Pierce (1970). Distribution of residual correlations in autoregressive-integrated moving average time series models. *Journal of the American Statistical Association* 65, 1509–1526.
- C.A.Glasbey (1988). Examples of regression with serially correlated errors. *The Statistician* 37, 277–291.
- D.B.Nelson (1991). Conditional heteroskedasticity in asset returns: a new approach, econometrica. *Journal of Econometric Society* 59, 347–370.
- Efron, B., T. Hastie, I. Johnstone, and R. Tibshirani (2004). Least angle regression. *Annals of Statistics* 32, 407–499.
- Engle, R. and T. Bollerslev (1986). Modelling the persistence of conditional variance. *Econometric Reviews* 5, 1–50.
- Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica* 50, 987–1007.
- Fan, J. and R. Li (2001a). Variable selection via nonconcave penalized likelihood and its oracle properties. *Journal of the American Statistical Association* 96, 1348–1360.
- Fan, J. and R. Li (2001b). Variable selection via nonconcave penalized likelihood and its oracle properties. *Journal of the American statistical Association* 96(456), 1348–1360.
- Frank, L. . E. . and J. . H. . Friedman (1993). A statistical view of some chemometrics regression tools. *Technometrics* 35(2), 109–135.
- Ghalanos, A. (2014). *rugarch: Univariate GARCH models*. R package version 1.4-0.

- G.Schwarz (1978). Estimating the dimension of a model. *The Annals of Statistics* 6(2), 461–464.
- H.Akaike (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control* 19(6), 716–723.
- Hoerl, A. E. and R. W. Kennard (1970). Ridge regression: biased estimation for non-orthogonal problems. *Technometrics* 12(1), 55–67.
- j. Racine (2000). Consistent cross-validators model-selection for dependent data: hv-block cross-validation. *Journal of Econometrics* 99, 39–61.
- J.Friedman, T. Hastie, and R.Tibshirani (2010). Regularization paths for generalized linear models via coordinate descent. *Journal of statistical software* 33(1), 1.
- J.Hannan, E. and B. G. Quinn (1979). The determination of the order of an autoregression. *Journal of the Royal Statistical Society. Series B (Methodological)*, 190–195.
- K.Daly (2006). volatility: Issues and measuring techniques. *Physica A* 387, 2377–2393.
- Ljung, G. M. and G. E. P. Box (1978). On a measure of a lack of fit in time series models. *Biometrika* 65(2), 297–303.
- Macaulay, F. (1938). *The Movements of Interest Rates, Bond Yields and Stock Prices in the United States Since 1856*. National Bureau of Economic Research.
- Markowitz, H. (1952). Portfolio selection. *Journal of Finance* 7, 77–91.
- McLeod, A. I. and W. K. Li (1983). Diagnostic checking arma time series models using squared residual autocorrelations. *Journal of Time Series Analysis* 4, 269–273.
- .Medeiros, M. C. and E. F. Mendes (2016). ‘1-regularization of high-dimensional time-series models with non-gaussian and heteroskedastic errors. *Journal of Econometrics* 191(1), 255–271.
- Montgomery, D. C., E. A. Peck, and G. G. Vinig (2011). *Introduction to Linear Regression Analysis (4th ed)*. John Wiley & Sons.
- P.Alquier and P.Doukhan (2011). Sparsity considerations for dependent variables. *Electronic Journal Statistics* 5, 750–774.

- 
- Park, T. and G. Casella (2008). The bayesian lasso. *Journal of the American Statistical Association* 103, 681–686.
- Ramanathan, R. (1998). *Introductory Econometrics with Applications*. TX: Dryden, Harcourt Brace College Publishers.
- R. Carter, W. Griffiths, and G. Lim (2007). *Principles of Econometrics*. Wiley, Third Edition.
- S., A. and A. Celisse (2010). A survey of cross-validation procedures for model selection. *Statistics Surveys* 4, 40–79.
- Shao, J. (1993). Linear model selection by cross-validation. *Journal of the American Statistical Association* 88, 486–495.
- Shi, P. and C. L. Tsay (2004). A joint regression variable and autoregressive order selection criterion. *Journal of Time Series Analysis* 25, 923–941.
- S. Tasy (2002). Analysis of financial time series. *Financial Economics*.
- Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society, Series B (Methodological)* 58, 267–288.
- Tibshirani, R., M. Saunders, S. Rosset, and K. Knight (2005). Sparsity and smoothness via the fused lasso. *Journal of the Royal Statistical Society* 67, 91–108.
- Tsay, R. (2013). *An Introduction to Analysis of Financial Data With R*. John Wiley & Sons.
- Tsay, R. S. (1984). Regression models with time series errors. *Journal of the American Statistical Association* 79, 118–124.
- Tsay, R. S. and G. C. Tiao (1984). Consistent estimates of autoregressive parameters and extended sample autocorrelation function for stationary and nonstationary arma models. *Journal of the American Statistical Association* 79, 84–96.
- Wang, H. and C. Leng (2008). A note on adaptive group lasso. *Computational statistics & data analysis* 52(12), 5277–5286.
- Wang, H., G. Li, and C. Tsai (2007). Regression coefficient and autoregressive order shrinkage and selection via the lasso. *Journal of the Royal Statistical Society* 69, 63–78.

- Wang, H., G. . Li, and C. . Tsai (2007). Regression coefficient and autoregressive order shrinkage and selection via the lasso. *Journal of the Royal Statistical Society* 69, 63–78.
- Yoon, Y. J., S. Lee, and T.Lee (2017). Adaptive lasso for linear regression models with arma-garch errors. *Communications in Statistics - Simulation and Computation* 46, 3479–3490.
- Yuan, M. and Y. Lin (2006). Model selection and estimation in regression with grouped variables. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 68(1), 49–67.
- Z.Ding, C. Granger, and R.F.Engle (1996). A long memory property of stock market returns and a new model. *Empirical Finance* 1, 83–106.
- Zhang, H. and W. .Lu (2007). Adaptive-lasso for cox’s proportional hazard model. *Biometrika* 94, 691–703.
- Zhang, L. (2017). *Penalized Regression Methods in Time Series and Functional Data Analysis*. Phd Thesis.
- Zhao, P. and B. Yu (2006). On model selection consistency of lasso. *Journal of Machine Learning Research* 7, 2541–2563.
- Zou, H. (2006). The adaptive lasso and its oracle properties. *Journal of the American Statistical Association* 101(476), 1418–1429.
- Zou, H. . and H. . H. . Zhang (2009). On the adaptive elastic-net with a diverging number of parameters. *Annals of Statistics* 37(4), 1733.
- Zou, H. and T. Hastie (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society, Series B (Statistical Methodology)* 67(2), 301–320.

## **Abstract**

Forecasting financial time series is especially important. There are many statistical models to predict time series, e.g.  $l_1$  AR,  $l_1$  MA, ARMA models. They are not good for modeling variance and instabilities in risk management, securities management, asset specialization and more. Instead, Autoregressive Conditional Heteroscedastic (or  $l_1$  ARCH) and Generalized Autoregressive Conditional Heteroscedastic (or GARCH) models have been used. If there are explanatory and response variables, it is better to use times series regression for modeling the data. In this dissertation, integrating penalized techniques like adaptive LASSO method in order to select important variables in time series regression models with errors  $l_1$  ARMA and  $l_1$  ARMA-GARCH is considered and verified efficiency of the proposed algorithm via some numerical studies.

**Keywords:** Adaptive LASSO estimator, ARMA model, ARMA-GARCH model, Time series regression, Variable selection.



**Shahrood University of Technology**

**Faculty of Mathematical Sciences**

**MSc Thesis in Statistics**

**On Variable Selection Methods in  
Regression Models with Time Series Error**

**By: Saeed Shirzayee**

**Supervisor**

**Dr. Mohammad Arashi**

**Advisor**

**Dr. Mina Norouzirad**

**August 2019**