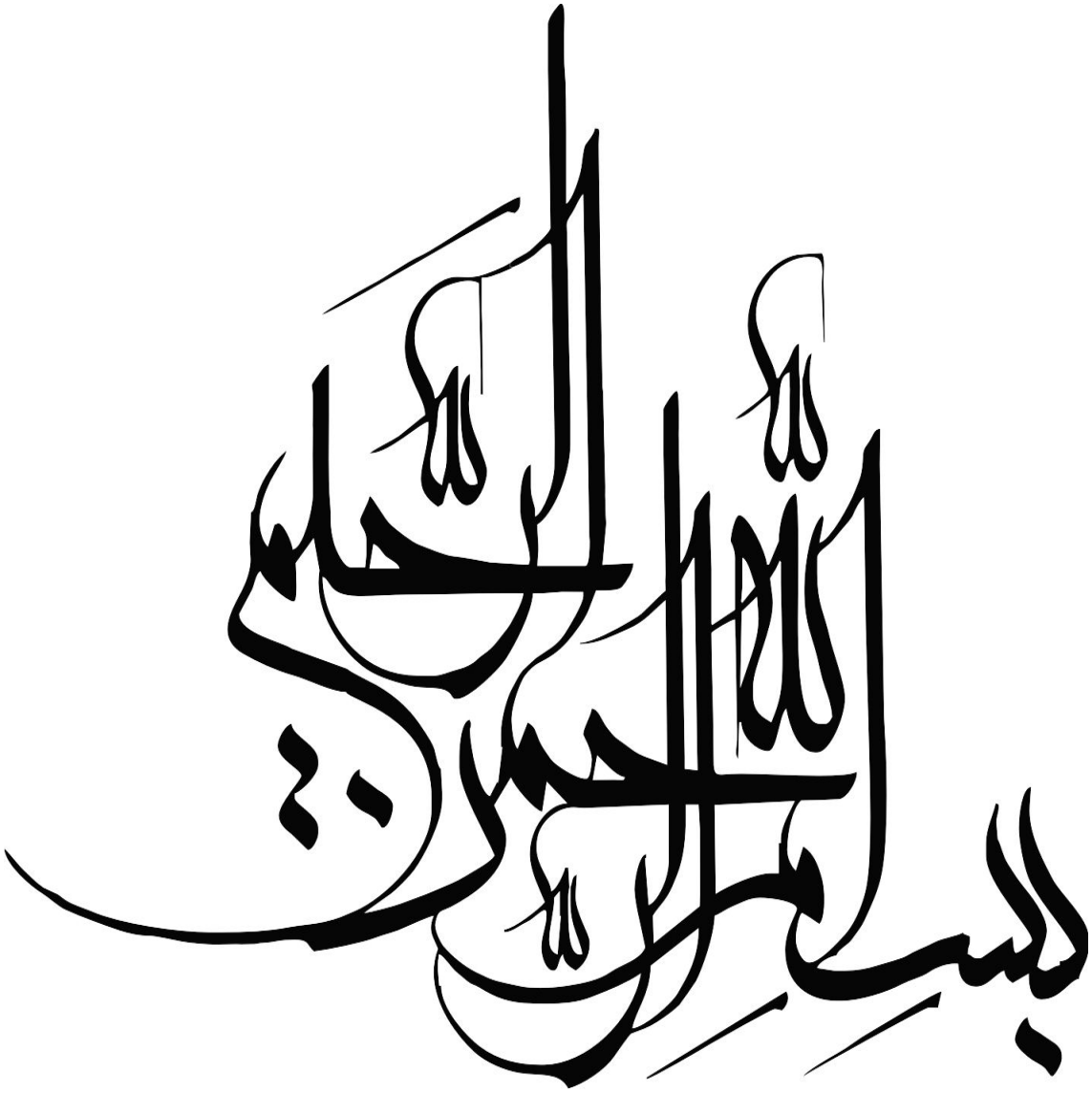
بسم الله الرحمن الرحيم

دانشکده علوم ریاضی

رشته آمار، گرایش آمار ریاضی

پایان‌نامه کارشناسی ارشد

برآوردگر انقباضی ماتریس کوواریانس در بعد بالا

زهرا حسین پور

استاد راهنما

دکتر محمد آرشی

تیر ۱۳۹۷

شماره:
تاریخ:

باسمه‌تعالی

دانشگاه صنعتی شاهرود

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 **خانم زهرا حسین پور** با شماره دانشجویی ۹۴۰۵۸۴۴ رشته **آمار** گرایش **آمار ریاضی** تحت عنوان: **برآوردگر انقباضی ماتریس کوواریانس در بعد بالا** که در تاریخ **۱۳۹۷/۰۴/۲۵** با حضور هیأت محترم داوران در دانشگاه صنعتی شاهرود برگزار گردید به شرح ذیل اعلام می‌گردد:

قبول (با درجه: خیلی خوب) ☑ مردود ☐

نوع تحقیق: نظری ☑ عملی ☐

عضو هیأت داوران	نام ونام خانوادگی	مرتبهٔ علمی	امضاء
۱ـ استادراهنمای اول	دکتر محمد آرشی	دانشیار	
۲- استادراهنمای دوم	------	------	------
۳- استاد مشاور	------	------	------
۴- نماینده تحصیلات تکمیلی	دکتر محمدرضا ربیعی	استادیار	
۵- استاد ممتحن اول	دکتر داود شاهسونی	دانشیار	
۶ـ استاد ممتحن دوم	دکتر حسین باغیشنی	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 به مقدس ترین واژه ها در لغت نامه دلم،

مادر مهربانم که زندگیم را مدیون مهر و عطوفت آن می دانم.

پدر، مهربانی مشفق، بردبار و حامی.

همسر و فرزندم که نشانه لطف الهی در زندگی من، هستند.

خواهر و برادرانم، همراهان همیشگی و پشتوانه های زندگیم.

سپاس‌گزاری

خدایا! شکر که تو را حدی برایت متصور نیست. سپاس تو را که بی نهایت مهربانی. خدایا! حمد و سپاس از آن توست که محیطی بر تمام عالم و تو تکیه گاهی بر تمام موجودات. معبودا! تو آن قدر گویایی که من از وصف تو ناتوانم. بند بند وجودم از شعاع نور توست که گرم و روشن می شود.

اکنون که با عنایت خداوند متعال این دوره‌ی علمی به پایان رسیده است وظیفه خود را می‌دانم که تشکر قلبی خود را نسبت به کسانی که به طریقی در انجام این پژوهش مرا یاری نموده‌اند ابراز دارم.

در آغاز وظیفه‌ی خود می‌دانم از زحمات بی‌دریغ استاد راهنمای خود، جناب آقای دکتر محمد آرشی صمیمانه تشکر و قدردانی کنم که از راهنمایی‌های ارزنده ایشان در راستای پیشبرد پژوهش حاصل فراوان بردم و همواره شاگرد مکتب علم و انسانیت و منش والای ایشان هستم.

همچنین لازم می‌دانم از اساتید فرهیخته جناب آقای دکتر داود شاهسونی، جناب آقای دکتر حسین باغیشنی که داوری این پایان‌نامه را به عهده گرفتند با تمام وجود تشکر و قدردانی نمایم.

در پایان از همه دوستان و عزیزانی که با اندیشه و قلبشان در این راه یاری‌ام نمودند، خالصانه تشکر و قدردانی می‌نمایم.

زهرا حسین پور

تیر ۱۳۹۷

تعهد نامه

اینجانب **زهرا حسین پور** دانشجوی کارشناسی ارشد رشته **آمار ریاضی** **دانشکده علوم ریاضی** دانشگاه صنعتی شاهرود، نویسنده پایان‌نامه با عنوان **برآوردگر انقباضی ماتریس کوواریانس در بعد بالا** ، تحت راهنمایی **دکتر محمد آرشی** متعهد می‌شوم:

- تحقیقات در این پایان‌نامه 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 پایان‌نامه، تا کنون توسط خود، یا فرد دیگری برای دریافت هیچ نوع مدرک یا امتیازی در هیچ‌جا ارایه نشده‌اند.
- حقوق معنوی این اثر، به دانشگاه صنعتی شاهرود تعلق دارد، و مقالات مستخرج با نام '' دانشگاه شاهرود '' یا " Shahrood University of Technology " به چاپ خواهند رسید.
- حقوق معنوی تمام افرادی که در به‌دست آوردن نتایج اصلی پایان‌نامه تاثیرگذار بوده‌اند، در مقالات مستخرج از پایان‌نامه رعایت می‌شود.
- در تمام مراحل انجام این پایان‌نامه، در مواردی که از موجود زنده (یا بافت‌های آن‌ها) استفاده شده است، ضوابط و اصول اخلاقی رعایت شده‌اند.
- در تمام مراحل انجام این پایان‌نامه، در مواردی که به حوزه اطلاعات شخصی افراد دسترسی یافته (یا استفاده) شده است، اصل رازداری و اصول اخلاق انسانی رعایت شده‌اند.

زهرا حسین پور

تیر ۱۳۹۷

مالکیت نتایج و حق نشر

- تمام حقوق معنوی این اثر و محصولات آن (مقالات مستخرج، کتاب، برنامه‌های رایانه‌ای، نرم‌افزارها و تجهیزات ساخته‌شده) 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 پایان‌نامه بدون ذکر منبع مجاز نیست.

چکیده

در تحلیل‌های چندمتغیره یکی از موارد چالش برانگیز برآورد ماتریس کوواریانس یک متغیر تصادفی، زمانی که تعداد متغیرها (در این‌جا بعد) بزرگ باشد، است. همچنین در بعضی بحث‌های چندمتغیره لازم است که تابعی از ماتریس کوواریانس برآورد شود. فرض کنید p نشان‌دهنده بعد باشد. در این پایان‌نامه برآوردگر ماتریس کوواریانس و تابعی از آن را برای حالت بعد بالا $(p \to \infty)$ مورد بررسی قرار می‌دهیم. در این راستا فرض می‌کنیم جامعه نمونه‌گیری دارای توزیع نرمال و غیرنرمال بوده و برآوردگرهای ناپارامتری انقباضی ماتریس کوواریانس در بعد بالا را معرفی کرده و برخی خواص آن‌ها را بررسی می‌کنیم. همچنین با استفاده از مثال‌های شبیه‌سازی، کارایی برآوردگرهای انقباضی معرفی شده برای ماتریس کوواریانس را در بعد بالا مورد ارزیابی قرار می‌دهیم.

کلمات کلیدی:
ماتریس کوواریانس، برآوردگر انقباضی، بعد بالا، سازگاری برآوردگر ناپارامتری.

پیش‌گفتار

مسئله‌ای را در نظر بگیرید که یک نمونه تصادفی به حجم n از یک بردار تصادفی p بعدی (p متغیره) در اختیار داریم و علاقه‌مندیم بر اساس این نمونه ماتریس کوواریانس Σ را که در اندازه $p \times p$ است برآورد کنیم. برآوردگر ماتریس کوواریانس نمونه S به راحتی قابل محاسبه است اما چنانچه بعد p بزرگ باشد خواص خوبی ندارد. در این حالت یافتن یک برآوردگر سازگار برای Σ در بعد بالا با توجه به کارایی آن از اهمیت بسزایی برخوردار است. یک راهکار مناسب و کارا استفاده از برآوردگر انقباضی Σ در بعد بالا است که تاکنون توسط محققین زیادی مورد مطالعه قرار گرفته است. در این راستا لدویت و ولف (۲۰۰۳)، سریواستاوا (۲۰۰۵) و فیشر و سان (۲۰۰۹) برآوردگرهای انقباضی ماتریس کوواریانس در بعد بالا را به خوبی بررسی کرده‌اند. در این پایان‌نامه، با توجه به اهمیت موضوع مورد بررسی، ساختار برآوردگرهای انقباضی خطی را برای ماتریس کوواریانس در بعد بالا بررسی کرده و تحقیقات جدید در خصوص برآوردگر انقباضی Σ و توابعی از آن را در جامعه نرمال و غیرنرمال ارائه می‌کنیم. با انجام مطالعات شبیه‌سازی برای مقادیر متفاوت حجم نمونه n و بعد متغیر p کارایی برآوردگرهای انقباضی را ارزیابی کرده و آن را از دیدگاه معیارهای متفاوت مقایسه می‌کنیم. این مجموعه شامل ۴ فصل و یک پیوست است که در ادامه به مطالب هر فصل به طور مختصر اشاره شده است.

در **فصل اول**، مقدمه‌ای بر برآوردگر ماتریس کوواریانس جامعه همرا با برآوردگرهای انقباضی نوع استاین Σ بیان می‌شود.

در **فصل دوم**، برآورد برخی توابع ماتریس کوواریانس را که در بحث آزمون فرضیه‌ها از آن‌ها استفاده می‌شود، در بعد بالا مورد بررسی قرار می‌دهیم.

فصل سوم به برآورد انقباضی ماتریس کوواریانس در بعد بالا در حالتی که توزیع جامعه نامعلوم است، اختصاص دارد.

فصل چهارم نیز در برگیرنده معرفی خانواده کلی برآوردگرهای انقباضی خطی Σ و مقایسه آن در بعد بالا است.

برنامه‌های کامپیوتری برای مثال‌های شبیه‌سازی به کار رفته، در پیوست آمده‌اند.

فهرست مطالب

فهرست جداول

فصل ۱

مقدمه

گاهی اوقات در آمار، ماتریس کوواریانس یک متغیر تصادفی چندمتغیره مجهول است و باید برآورد شود. برآورد ماتریس‌های کوواریانس با این سوال مواجه می‌شود که چگونه می‌توان ماتریس واقعی کوواریانس را بر اساس نمونه مشاهده شده از توزیع چندمتغیره تقریب نمود. در مواردی که مشاهده‌ها کامل هستند، می‌توان از ماتریس کوواریانس نمونه استفاده کرد. معمولا این ماتریس، یک برآوردگر ناریب و کارا برای ماتریس کوواریانس است.

تحلیل آماری داده‌های چندمتغیره، اغلب شامل مطالعات اکتشافی[۱] از متغیرهایی که با یکدیگر ارتباط دارند، می‌باشند و ممکن است با مدل‌های آماری صریح با ماتریس کوواریانس متغیرها بیان شود. بنابراین، برآورد ماتریس کوواریانس به‌طور مستقیم از داده‌های مشاهده‌شده، دو نقش مهم دارد:

(۱) برای ارایه برآوردهای اولیه که می‌تواند برای مطالعه روابط درون متغیرها مورد استفاده قرار گیرد.

(۲) برای ارایه برآوردهای نمونه‌ای که می‌تواند برای بررسی مدل به‌کار رود.

برآوردهای ماتریس کوواریانس در مراحل اولیه تحلیل مولفه‌های اصلی[۲] و تحلیل عاملی[۳] مورد نیاز است. همچنین در مدل‌های رگرسیونی که برخی از متغیرهای پیشگو به یکدیگر وابستگی دارند، به‌کار برده می‌شود.

فرض کنید $\boldsymbol{x}_1$، $\boldsymbol{x}_2$، ...، $\boldsymbol{x}_n$ مشاهده‌های مستقل از بردار تصادفی p بعدی $\boldsymbol{X}$ باشند. ماتریس کوواریانس به‌صورت زیر تعریف می‌شود

$$\boldsymbol{\Sigma} = E\left[(\boldsymbol{X} - E[\boldsymbol{X}])(\boldsymbol{X} - E[\boldsymbol{X}])^\top\right].$$

[۱] Exploratory

[۲] Principle component

[۳] Factor analysis

یک برآوردگر ناریب برای ماتریس Σ، ماتریس کوواریانس نمونه است که به‌صورت زیر تعریف می‌شود

$$\boldsymbol{S} = \frac{۱}{n-۱}\sum_{i=۱}^{n}(\boldsymbol{x}_i - \bar{\boldsymbol{x}})(\boldsymbol{x}_i - \bar{\boldsymbol{x}})^\top, \tag{۱.۱}$$

که در آن

$$\bar{\boldsymbol{x}} = \frac{۱}{n}\sum_{i=۱}^{n}\boldsymbol{x}_i.$$

در مواردی که بعد داده‌ها بالاست مانند مطالعات مربوط به بیماری سرطان، برآوردگر ماتریس کوواریانس نمونه، اغلب ضعیف عمل می‌کند. برای مثال به میرهد[۴] (۱۹۸۲)؛ جان استون[۵] (۲۰۰۱)؛ بیکل و لوینا[۶] (۲۰۰۸)؛ و فن و ال وی[۷] (۲۰۰۸) مراجعه کنید. روش‌هایی اخیرا برای برآورد بهبودیافته ماتریس کوواریانس ارایه شده‌اند. در این راستا می‌توان به روش‌هایی در لم و فن[۸] (۲۰۰۹)، رتمن[۹] و همکاران (۲۰۰۹)، بیکل و لوینا (۲۰۰۸)، وو و پوراحمدی[۱۰] (۲۰۰۳) و جان استون و لو[۱۱] (۲۰۰۹) اشاره کرد.

در برآورد بردار میانگین توزیع نرمال چندمتغیره، استاین[۱۲] (۱۹۵۶) نشان داد که می‌توان برآوردگر درستنمایی ماکسیمم[۱۳] (MLE) را زمانی که بعد p غیر قابل اغماض است، بهبود بخشید. استاین (۱۹۶۴) به این موضوع پی برد که بررسی این عقیده بهتر از MLE می‌باشد. نظرات مختلفی در این زمینه ارائه شده‌اند که به چندین مورد اشاره می‌کنیم. هاف[۱۴] (۱۹۸۱ و ۱۹۷۹) برآوردگر جدیدی برای ماتریس دقت[۱۵] که معکوس ماتریس کوواریانس، یعنی $\Sigma^{-۱}$، است، مبتنی بر دو تابع زیان مختلف برای توزیع ویشارت ارائه داد. افرون و موریس[۱۶] (۱۹۷۶) و یانگ و برگر[۱۷] (۱۹۹۴) با استفاده از رویکرد بیزی به ترتیب ماتریس دقت و ماتریس کوواریانس را برآورد کردند. دی و سرینیواسان[۱۸] (۱۹۸۵) به برآوردگری دست یافتند که تحت تابع زیان مشخص، مینیماکس است.

---

[۴] Muirhead
[۵] Johnstone
[۶] Bickel and Levina
[۷] Fan and Lv
[۸] Lam and Fan
[۹] Rothman
[۱۰] Wu and Pourahmadi
[۱۱] Lu
[۱۲] Stein
[۱۳] Maximum likelihood estimator
[۱۴] Haff
[۱۵] Precision matrix
[۱۶] Efron and Morris
[۱۷] Yang and Berger
[۱۸] Dey and Srinivasan

اساس کار همه محققین کاهش متوسط توابع زیان $L(\hat{\Sigma}, \Sigma)$ یا $L(\hat{\Sigma}^{-1}, \Sigma^{-1})$ در مقایسه با $L(S, \Sigma)$ یا $L(S^{-1}, \Sigma^{-1})$ و دسترسی به برآوردگر مناسبی برای کاهش این معیار بوده است. این کار برای توابع زیان مختلفی با روش‌های متفاوتی انجام شده است که در این قسمت تعدادی از این روش‌ها را بیان خواهیم نمود. به‌طور خلاصه تعدادی از کارهایی که به برآورد انقباضی منتهی نمی‌شود را بیان خواهیم نمود. بیشتر تکنیک‌ها، شرایطی را برای ماتریس کوواریانس، برای به‌دست آوردن برآوردگری مناسب، در نظر گرفته‌اند. رورتو[۱۹] (۲۰۰۰)، اسمیت و کهن[۲۰] (۲۰۰۲)، و چن و دانسون[۲۱] (۲۰۰۳)، از رویکرد تجزیه چولسکی[۲۲] استفاده کردند.

همل و هارتلی[۲۳] (۱۹۷۳) نشان دادند که چطور تابع هدف و مشتقات آن را برای برآورد درستنمایی ماکسیمم مولفه‌های واریانس می‌توان محاسبه کرد. این کار سپس توسط کربیل و سارل[۲۴] (۱۹۷۶) دنبال شد. هاف (۱۹۹۱) نظر کلی برای برآوردگر بیزی بردار میانگین و ماتریس کوواریانس را با کمینه کردن زیان بیزی فراهم کرد. فرالی و برنز[۲۵] (۱۹۹۵) نشان دادند که چگونه توابع احتمال و مشتقات آن‌ها توسط روش‌های ماتریس پراکنده (تنک[۲۶]) محاسبه می‌شود و نتایجی کلی را تعمیم دادند. پوراحمدی (۱۹۹۹) به بررسی مشترک مدل‌های میانگین‌ـ‌واریانس پرداخت. دنیلز و پوراحمدی[۲۷] (۲۰۰۲) و دنیل (۲۰۰۵) از یک رویکرد بیزی توسط کاوش‌های قبلی برای مشاهدات وابسته استفاده کردند. وو و پوراحمدی (۲۰۰۳) یک برآوردگر ناپارامتری را برای ماتریس کوواریانس پیشنهاد نمودند.

دنیلز (۲۰۰۶) چندین برآوردگر برای ماتریس کوواریانس مبتنی بر چندین مدل رایج رگرسیون بیزی پیشنهاد کرد. چادهوری و همکاران[۲۸] (۲۰۰۷) الگوریتمی را برای محاسبه برآورد درستنمایی ماکسیمم ماتریس کوواریانس تحت این محدودیت که کوواریانس صفر است، ارائه دادند. فن و ال وی (۲۰۰۸) در مدل کوواریانس به برآوردگر مناسبی با استفاده از تکنیک‌های تحلیل عاملی چندمتغیره دست یافتند.

کائو و بومن[۲۹] (۲۰۰۹)، ماتریس را مبتنی بر فرضیه‌ای که مقادیر ویژه با استفاده از تبدیل ماتریس پراکنده ارائه شده است، برآورد نمودند. فان و وو (۲۰۰۸) روش نیمه پارامتری را که در آن واریانس با استفاده از یک تابع ناپارامتری تخمین زده می‌شود و همبستگی با یک تابع پارامتری برآورد می‌شود، در برآورد در نظر گرفتند.

---

[۱۹] Roverato
[۲۰] Smith and Kohn
[۲۱] Chen and Dunson
[۲۲] Cholesky decomposition
[۲۳] Hemmerle and Hartley
[۲۴] Corbeil and Searle
[۲۵] Fraley and Burns
[۲۶] Sparse matrix
[۲۷] Daniels and Pourahmadi
[۲۸] Chaudhuri et al.
[۲۹] Cao and Buoman

مواردی که بیان شد تعدادی از پیشرفت‌های اخیر در برآورد ماتریس کوواریانس و ماتریس دقت می‌باشند. با این حال، در خصوص رده کلی از برآوردگرهای ماتریس کوواریانس و ماتریس دقت که به طور معمول به نام استاین یا انقباضی یا برآوردگر انقباضی[۳۰] نام گرفته است، مطالعات چندانی صورت نگرفته است. در این پایان‌نامه قصد داریم به مطالعه برآوردگرهای انقباضی ماتریس کوواریانس بپردازیم.

۱.۱ تعاریف و توابع ریاضی

برای اطلاع دقیق‌تر از تعاریف این بخش می‌توان به رودین[۳۱] (۱۳۷۳)، میرهد (۱۹۸۲) و کرمی‌کبیر (۱۳۹۲) مراجعه کرد.

تعریف ۱.۱.۱ (تابع گاما[۳۲]). تابع گاما را به ازای هر عدد حقیقی $a > ۰$ با نماد $\Gamma(a)$ نشان داده و به‌صورت زیر تعریف می‌کنیم

$$\Gamma(a) = \int_{۰}^{\infty} \mathrm{e}^{-x} x^{a-۱} \mathrm{d}x.$$

تعریف ۲.۱.۱ (توزیع ویشارت[۳۳]). فرض کنید $W = (w_{ij})$ ماتریس متقارن معین مثبت با احتمال یک باشد، و Σ نیز ماتریس $q \times q$ معین مثبت باشد. اگر $m \geq q$ (m عدد صحیح)، آنگاه W دارای توزیع ویشارت با m درجه آزادی است اگر تابع چگالی پیوسته‌ی $\frac{۱}{۲}q(q+۱)$ عناصر مجزا W (بالا مثلثی گفته می شود) به‌صورت زیر باشد

$$f(w_{۱۱}, w_{۱۲}, \cdots, w_{qq}) = c^{-۱} \det W^{\frac{m-q}{۲}-\frac{۱}{۲}} etr\Big(-\frac{۱}{۲}\Sigma^{-۱}W\Big)$$

که منظور از etr همان e^{trace} می باشد و

$$c = ۲^{\frac{mq}{۲}} \det \Sigma^{\frac{m}{۲}} \Gamma_q(\frac{m}{۲})$$

که $\Gamma_q(a)$ تابع گامای چندمتغیره به‌صورت

$$\Gamma_q(a) = \pi^{\frac{q(q-۱)}{۲}} \prod_{j=۱}^{q} \Gamma\Big[a + \frac{۱}{۲}(۱-j)\Big]$$

می‌باشد. همچنین می‌توان گفت که $W \sim W_q(m, \Sigma)$ است.

تعریف ۳.۱.۱. فرض کنید X_n و Y_n دو دنباله تصادفی حقیقی مقدار باشند. آن‌گاه وقتی $n \to \infty$،

[۳۰] Shrinkage estimator
[۳۱] Rudin
[۳۲] Gamma function
[۳۳] Wishart distribution

(۱) $X_n = O(Y_n)$، (O نماد اوی بزرگ) اگر و فقط اگر به ازای هر $\epsilon > ۰$، وجود داشته باشند $\delta > ۰$ و $N > ۰$ به‌طوری که برای همه $n > N$،

$$P\Big(\Big|\frac{X_n}{Y_n}\Big| > \delta\Big) < \epsilon.$$

(۲) $X_n = o(Y_n)$، (o نماد اوی کوچک) اگر و فقط اگر به ازای هر $\epsilon > ۰$،

$$\lim_{n\to\infty} P\Big(\Big|\frac{X_n}{Y_n}\Big| > \epsilon\Big) = ۰.$$

تعریف ۴.۱.۱ (مجموعه محدب[۳۴]). مجموعه C محدب است، اگر پاره‌خط بین هر دو نقطه دلخواه در C، درون C قرار گیرد. به عبارتی به ازای هر $x_۱$ و $x_۲$ عضو مجموعه C و به ازای هر θ ثابت، با شرط $۰ \le \theta \le ۱$، داشته باشیم

$$\theta x_۱ + (۱ - \theta)x_۲ \in C.$$

تعریف ۵.۱.۱ (تابع محدب). تابع $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ محدب است، اگر دامنه f یک مجموعه محدب باشد و به ازای هر x و y عضو دامنه f و $۰ \le \theta \le ۱$ داشته باشیم

$$f(\theta x + (۱ - \theta)y) \le \theta f(x) + (۱ - \theta)f(y).$$

تابع $f(.)$ را مقعر گوییم اگر $-f(.)$ محدب باشد.

تعریف ۶.۱.۱ (نرم[۳۵]). تابع $f : \mathbb{R}^p \longrightarrow \mathbb{R}$ را نرم گوییم، اگر

(۱) $f(.)$ تابعی نامنفی باشد، یعنی به ازای هر $\boldsymbol{x} \in \mathbb{R}^p$ داشته باشیم، $f(\boldsymbol{x}) \ge ۰$.

(۲) $f(.)$ معین[۳۶] باشد، یعنی $f(\boldsymbol{x}) = ۰$ اگر و تنها اگر $\boldsymbol{x} = ۰$.

(۳) $f(.)$ همگن[۳۷] باشد، یعنی به ازای هر $\boldsymbol{x} \in \mathbb{R}^p$ و $t \in \mathbb{R}$ داشته باشیم، $f(t\boldsymbol{x}) = |t|f(\boldsymbol{x})$.

(۴) $f(.)$ در نامساوی مثلثی[۳۸] صدق کند، یعنی به ازای هر $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^p$ داشته باشیم

$$f(\boldsymbol{x} + \boldsymbol{y}) \le f(\boldsymbol{x}) + f(\boldsymbol{y}).$$

در این‌صورت از نماد $f(\cdot) = \|\cdot\|$ برای نمایش نرم استفاده می‌کنیم.

تعریف ۷.۱.۱ (نرم ماتریسی). فرض کنید $\boldsymbol{A}$ یک ماتریس $m \times n$ باشد، در این‌صورت نرم ماتریسی $\|\boldsymbol{A}\|$ با خواص زیر تعریف می‌کنیم

---

[۳۴] Convex
[۳۵] Norm
[۳۶] Definite
[۳۷] Homogenous
[۳۸] Triangle inequality

(۱) $\|\boldsymbol{A}\| \geq ۰$، $\|\boldsymbol{A}\| = ۰$ اگر و فقط اگر $\boldsymbol{A}$ یک ماتریس صفر باشد.

(۲) به ازای هر اسکالر α، $\|\alpha\boldsymbol{A}\| = |\alpha|\|\boldsymbol{A}\|$.

(۳) $\|\boldsymbol{A}+\boldsymbol{B}\| \leq \|\boldsymbol{A}\| + \|\boldsymbol{B}\|$.

علاوه بر سه شرط فوق برای نرم ماتریسی، گاهی اوقات به شرایط زیر نیز نیاز دارند که لزوما برای همه نرم‌های ماتریس برقرار نیست

$$\|\boldsymbol{A}\boldsymbol{x}\| \leq \|\boldsymbol{A}\| \cdot \|\boldsymbol{x}\|,$$
$$\|\boldsymbol{A}\boldsymbol{B}\| \leq \|\boldsymbol{A}\| \cdot \|\boldsymbol{B}\|.$$

تعریف ۸.۱.۱ (نرم‌های مولفه‌ای)**.** اگر با همه‌ی $m \times n$ مولفه‌ی ماتریس $\boldsymbol{A}$ مشابه مولفه‌های یک بردار mn بعدی رفتار شود، آن‌گاه pنرم این بردار به‌صورت زیر تعریف می‌شود

$$\|A\|_p = \left\{ \sum_{i=۱}^{m} \sum_{j=۱}^{n} |\alpha_{ij}|^p \right\}^{\frac{۱}{p}}.$$

به ویژه سه حالت $p = ۱, ۲, \infty$ را در نظر می‌گیریم.

(۱) $p = ۱$، $\|\boldsymbol{A}\|_۱$ قدرمطلق مجموع همه عناصر ماتریس $\boldsymbol{A}$ است. به عبارتی

$$\|\boldsymbol{A}\|_۱ = \sum_{i=۱}^{m} \sum_{j=۱}^{n} |a_{ij}|.$$

(۲) $p = \infty$، $\|\boldsymbol{A}\|_\infty$، نرم بیشینه، ماکزیمم مقدار مطلق در بین همه mn عنصر ماتریس $\boldsymbol{A}$ است. یعنی،

$$\|\boldsymbol{A}\|_\infty = \max\{|a_{ij}| : ۱ \leq i \leq m,\ ۱ \leq j \leq n\}.$$

(۳) $p = ۲$، $\|\boldsymbol{A}\|_۲$ نرم فروبنیوس[۳۹] است که به‌صورت زیر می‌باشد

$$\|\boldsymbol{A}\|_۲ = \|\boldsymbol{A}\|_F = \sqrt{\sum_{i=۱}^{m} \sum_{j=۱}^{n} |a_{ij}|^۲} = \sqrt{\mathrm{tr}(\boldsymbol{A}^\top \boldsymbol{A})} = \sqrt{\sum_{i=۱}^{r} \lambda_i} = \sqrt{\sum_{i=۱}^{r} \sigma_i^۲},$$

که در آن $r \leq \min(m,n)$ رتبه ماتریس $\boldsymbol{A}$، $\lambda_i = \sigma_i^۲$، i امین مقدار ویژه نامنفی از $\boldsymbol{A}^\top\boldsymbol{A}$ است و $\sigma_i = \sqrt{\lambda_i}$، i امین مقدار تکین (ریشه مقدار ویژه) $\boldsymbol{A}$ $(i = ۱, ۲, \ldots, r)$ است.

نرم فروبینیوس ماتریس متعامد $\boldsymbol{R}$ که در شرط $\boldsymbol{R}^\top = \boldsymbol{R}^{-۱}$ یا $\boldsymbol{R}^\top\boldsymbol{R} = \boldsymbol{R}^{-۱}\boldsymbol{R} = \boldsymbol{I}$ صدق کند، برابر با $\|\boldsymbol{R}\|_F^۲ = \mathrm{tr}(\boldsymbol{R}^\top\boldsymbol{R}) = \mathrm{tr}(\boldsymbol{I}) = n$ می‌باشد.

[۳۹] Frobenius norm

تعریف ۹.۱.۱ (عدد شرطی)**.** نسبت بزرگترین مقدار ویژه ماتریس $\boldsymbol{A}$ به کوچکترین مقدار ویژه ماتریس $\boldsymbol{A}$ را عدد شرطی گوییم. به عبارتی

$$k(\lambda) = \Big|\frac{\lambda_{max}}{\lambda_{min}}\Big|.$$

تعریف ۱۰.۱.۱ (تابع زیان توان دوم)**.** اگر بردار p مولفه‌ای $\boldsymbol{\delta} = (\delta_1, \cdots, \delta_p)^\top$ برآوردگری برای بردار پارامتر $\boldsymbol{\theta} \in \mathbb{R}^p$ باشد، آن‌گاه زیان استفاده از $\boldsymbol{\delta}$ در خصوص برآورد $\boldsymbol{\theta}$ را با $L(\boldsymbol{\theta}, \boldsymbol{\delta})$ نشان داده که برابر است با

$$L(\boldsymbol{\theta}, \boldsymbol{\delta}) = (\boldsymbol{\delta} - \boldsymbol{\theta})^\top(\boldsymbol{\delta} - \boldsymbol{\theta}) = \|\boldsymbol{\delta} - \boldsymbol{\theta}\|^2 = \sum_{i=1}^{p}(\delta_i - \theta_i)^2.$$

چنان‌چه پارامتر مقیاس مجهول $\sigma^2 \geq 0$ در مدل موجود باشد، آن‌گاه از تابع زیان زیر استفاده می‌کنیم

$$L(\boldsymbol{\theta}, \boldsymbol{\delta}) = \frac{(\boldsymbol{\delta} - \boldsymbol{\theta})^\top(\boldsymbol{\delta} - \boldsymbol{\theta})}{\sigma^2} = \frac{\|\boldsymbol{\delta} - \boldsymbol{\theta}\|^2}{\sigma^2} = \frac{1}{\sigma^2}\sum_{i=1}^{p}(\delta_i - \theta_i)^2.$$

تعریف ۱۱.۱.۱ (تابع مخاطره)**.** با استفاده از تعریف تابع زیان، تابع مخاطره $\boldsymbol{\delta}$ را با $R(\boldsymbol{\theta}, \boldsymbol{\delta})$ نشان داده که برابر است با

$$R(\boldsymbol{\theta}, \boldsymbol{\delta}) = E_{\boldsymbol{\theta}}[L(\boldsymbol{\theta}, \boldsymbol{\delta})] = \int_{\chi} L(\boldsymbol{\theta}, \boldsymbol{\delta}) f_{\boldsymbol{\theta}}(x)\mathrm{d}x,$$

که در آن $f_{\boldsymbol{\theta}}(\cdot)$ تابع چگالی احتمال تعریف‌شده روی χ است.

تعریف ۱۲.۱.۱ (اثر یک ماتریس)**.** در جبرخطی، اثر ماتریس $n \times n$ دلخواه $\boldsymbol{A}$ به‌صورت مجموع عناصر قطر اصلی $\boldsymbol{A}$ است. به‌عبارت دیگر

$$\mathrm{tr}(\boldsymbol{A}) = \sum_{i=1}^{n} a_{ij} = a_{11} + a_{22} + \cdots + a_{nn},$$

که a_{ii}، iامین مولفه روی قطر اصلی ماتریس $\boldsymbol{A}$ است.

ویژگی‌های اثر ماتریس

(۱) به‌ازای ماتریس‌های متقارن مربعی $\boldsymbol{A}$ و $\boldsymbol{B}$ و مقدار ثابت $c \in \mathbb{R}$

$$\mathrm{tr}(\boldsymbol{A} + \boldsymbol{B}) = \mathrm{tr}(\boldsymbol{A}) + \mathrm{tr}(\boldsymbol{B}),$$
$$\mathrm{tr}(c\boldsymbol{A}) = c\ \mathrm{tr}(\boldsymbol{A}).$$

در حالت کلی، فرض کنید $k_1, \ldots, k_m \in \mathbb{R}$ و $\boldsymbol{A}_1, \cdots, \boldsymbol{A}_m$ ماتریس‌های مربعی باشند. در این‌صورت

$$\mathrm{tr}(\sum_{i=1}^{m} k_i \boldsymbol{A}_i) = \sum_{i=1}^{m} k_i\ \mathrm{tr}(\boldsymbol{A}_i).$$

(۲) به ازای هر ماتریس مربعی دلخواه $\boldsymbol{A}$، $\mathrm{tr}(\boldsymbol{A}) = \mathrm{tr}(\boldsymbol{A}^\top)$. زیرا ترانهاده کردن یک ماتریس، عناصر قطر اصلی را تغییر نمی‌دهد.

(۳) اثر حاصل‌ضرب ماتریس‌ها می‌تواند به‌صورت مجموع حاصل‌ضرب مولفه‌ها نوشته شود. یعنی

$$\mathrm{tr}(\boldsymbol{A}^\top\boldsymbol{B}) = \mathrm{tr}(\boldsymbol{A}\boldsymbol{B}^\top) = \mathrm{tr}(\boldsymbol{B}^\top\boldsymbol{A}) = \mathrm{tr}(\boldsymbol{B}\boldsymbol{A}^\top) = \sum_{i,j} a_{ij}b_{ij}.$$

بدین معنی است که حاصل‌ضرب توابع ماتریسی، مشابه حاصل‌ضرب داخلی بردارهاست. در حالت خاص اگر $\boldsymbol{A} = (a_{ij})$ ماتریس $m \times n$ باشد، آن‌گاه

$$\mathrm{tr}(\boldsymbol{A}^\top\boldsymbol{A}) = \sum_{i=۱}^{m}\sum_{j=۱}^{n} a_{ij}^۲.$$

(۴) برای ماتریس‌های حقیقی، اثر ماتریس حاصلضرب می‌تواند به‌صورت زیر نوشته شود

$$\mathrm{tr}(\boldsymbol{A}^\top\boldsymbol{B}) = \sum_{i,j}(\boldsymbol{A}\circ\boldsymbol{B})_{ij}.$$

که $\circ$ ضرب داخلی (ضرب مولفه به مولفه) را نشان می‌دهد.

(۵) ماتریس‌ها در اثر حاصلضرب می‌توانند بدون تغییر جابه‌جا شوند. یعنی اگر $\boldsymbol{A}$ ماتریس $m \times n$ و $\boldsymbol{B}$ ماتریس $n \times m$ باشند، آن‌گاه

$$\mathrm{tr}(\boldsymbol{A}\boldsymbol{B}) = \mathrm{tr}(\boldsymbol{B}\boldsymbol{A}).$$

در حالت کلی، اثر تحت جایگشت‌های دایره‌ای[۴۰] پایا است یعنی

$$\mathrm{tr}(\boldsymbol{A}\boldsymbol{B}\boldsymbol{C}\boldsymbol{D}) = \mathrm{tr}(\boldsymbol{B}\boldsymbol{C}\boldsymbol{D}\boldsymbol{A}) = \mathrm{tr}(\boldsymbol{C}\boldsymbol{D}\boldsymbol{A}\boldsymbol{B}) = \mathrm{tr}(\boldsymbol{D}\boldsymbol{A}\boldsymbol{B}\boldsymbol{C}).$$

این ویژگی به خاصیت دایره‌ای معروف است. دقت کنید که برای هر جایگشت دلخواهی رابطه برقرار نیست. در حالت کلی

$$\mathrm{tr}(\boldsymbol{A}\boldsymbol{B}\boldsymbol{C}) \neq \mathrm{tr}(\boldsymbol{A}\boldsymbol{C}\boldsymbol{B}).$$

با این وجود اگر حاصل‌ضرب ماتریس‌های متقارن درنظر گرفته شود، هر جایگشتی مجاز است زیرا

$$\begin{aligned}\mathrm{tr}(\boldsymbol{A}\boldsymbol{B}\boldsymbol{C}) &= \mathrm{tr}(\boldsymbol{A}^\top\boldsymbol{B}^\top\boldsymbol{C}^\top) = \mathrm{tr}(\boldsymbol{A}^\top(\boldsymbol{C}\boldsymbol{B})^\top) = \mathrm{tr}((\boldsymbol{C}\boldsymbol{B})^\top\boldsymbol{A}^\top)\\ &= \mathrm{tr}((\boldsymbol{A}\boldsymbol{B}\boldsymbol{C})^\top)\end{aligned}$$

تساوی آخر از ویژگی (۲) نتیجه شده است.

[۴۰] Cyclic Permutations

(۶) حاصلضرب ماتریس‌ها با حاصلضرب اثر ماتریس‌ها برابر نیست به‌عبارت دیگر

$$\operatorname{tr}(\boldsymbol{AB}) \neq \operatorname{tr}(\boldsymbol{A})\operatorname{tr}(\boldsymbol{B}).$$

اما اثر حاصلضرب کرونکر [۴۱] دو ماتریس، حاصلضرب اثر آن‌هاست یعنی

$$\operatorname{tr}(\boldsymbol{A} \otimes \boldsymbol{B}) = \operatorname{tr}(\boldsymbol{A})\operatorname{tr}(\boldsymbol{B}).$$

(۷) اگر $\boldsymbol{A}$ متقارن و $\boldsymbol{B}$ نامتقارن باشد، آن‌گاه $\operatorname{tr}(\boldsymbol{AB}) = ۰$.

(۸) اثر ماتریس همانی، بعد فضای ماتریس است، این ویژگی منجر به تعمیم بعد با استفاده از اثر می‌شود.

(۹) اثر یک ماتریس خودتوان $\boldsymbol{A}$ (یعنی $\boldsymbol{A}^۲ = \boldsymbol{A}$) برابر رتبه ماتریس $\boldsymbol{A}$ است. برای جزئیات بیشتر، لانگ[۴۲] (۱۹۷۸) را ببینید.

تعریف ۱۳.۱.۱ (برآوردگر انقباضی). در آمار برآوردگر انقباضی، برآوردگری است که به‌طور واضح یا مجازی در اثر انقباض به‌دست می‌آید. به عبارت دیگر، برآوردگر انقباضی برآوردگری است که از ترکیب اطلاعات نمونه‌ای یعنی برآوردگر خام یا اولیه مثلا MLE، برآوردگر کمترین توان‌های دوم، بیزی و غیره با اطلاعات غیر نمونه‌ای که معمولا در قالب یک یا چند محدودیت به مدل اضافه می‌شوند، به‌صورت برآوردگر بهبود یافته جدید به‌دست می‌آید. این بهبود در جهت بهبود برآوردگر اولیه است. در حالت کلی بهبود برآوردگرها را می‌توان با معیار مخاطره یا توان دوم خطا[۴۳] (MSE) اندازه‌گیری کرد. در این‌صورت انقباض می‌تواند به سمت صفر یا یک مقدار ثابت باشد. اثر این انقباض معمولا به‌صورت تبدیل یک برآوردگر ناریب به اریب ولی با MSE کمتر می‌باشد.

در عین حال، می‌توان برآوردگرهایی انقباضی از نوع استاین و ناریب را نیز یافت. برای آگاهی بیشتر در این خصوص به نوروزی راد (۱۳۹۰) مراجعه کنید.

شکل کلی برآوردگر انقباضی به‌صورت $\boldsymbol{X} + g(\boldsymbol{X})$ است که در آن g یک تابع اندازه‌پذیر می‌باشد. پس از حل نامعادله

$$\forall\, \theta \in \Theta : \quad R(\boldsymbol{\theta}, \boldsymbol{X} + g(\boldsymbol{X})) \leq R(\boldsymbol{\theta}, \boldsymbol{X}),$$

به کمک لم استاین، تابع $g(\cdot)$ را در هر یک از حالت‌های زیر به‌دست می‌آوریم

(۱) برآوردگر انقباضی به سمت صفر، $(۱ - c)\boldsymbol{X} = \boldsymbol{X} - c\boldsymbol{X}$.

[۴۱] Kronecher product
[۴۲] Lang
[۴۳] Mean squared error

(۲) برآوردگر انقباضی به سمت بردار ثابت $\boldsymbol{m}$

$$\boldsymbol{m} + c(\boldsymbol{X} - \boldsymbol{m}) = (۱ - c)\boldsymbol{m} + c\boldsymbol{X},$$

که در آن c ثابت انقباضی است و بسته به نوع توزیع و تابع زیان، مقادیر متفاوتی می‌پذیرد.

تعریف ۱۴.۱.۱ (سازگاری)**.** فرض کنید $X_۱$، ...، و X_n دنباله‌ای از متغیرهای تصادفی مستقل با توزیعی از خانواده توزیع‌های $\{F_\theta : \theta \in \Theta\}$ باشد. همچنین فرض کنید T_n یک برآوردگر $\gamma(\theta)$ بر اساس n مشاهده اول باشد. هدف از سازگاری T_n بررسی رفتار حدی دنباله $\{T_n\}_{n=۱}^{\infty}$ از برآوردگرهاست. گوییم دنباله برآوردگرهای T_n برای $\gamma(\theta)$ سازگار است اگر چنان‌چه n به سمت ∞ میل کند، آن‌گاه

$$T_n \xrightarrow{p} \gamma(\theta),$$

یعنی برای هر $\epsilon > ۰$ و $\gamma > ۰$ عدد صحیح مثبتی مانند $n_۰ = n_۰(\epsilon, \delta)$ وجود دارد به‌طوری که برای هر $n > n_۰$،

$$P_\theta(|T_n - \gamma(\theta)| > \epsilon) < \delta,$$

یا برای هر $\epsilon > ۰$،

$$\lim_{n\to\infty} P_\theta(|T_n - \gamma(\theta)| > \epsilon) = ۰.$$

هر دنباله از برآوردگرها که در شرط فوق صدق نکند ناسازگار می‌نامیم.

اندازه‌گیری‌های حاصل روی یک مشاهده، به‌صورت برداری به شکل زیر است

$$\boldsymbol{x}_i = \begin{pmatrix} x_{i۱} \\ x_{i۲} \\ \vdots \\ x_{ip} \end{pmatrix},$$

که x_i $(i = ۱, \ldots, n)$، iامین مشاهده از نمونه تصادفی را ارائه می‌دهد. مجموعه تمام اندازه‌گیری‌های حاصل از مشاهدات، ماتریس زیر را تشکیل می‌دهد

$$\mathbf{X} = (\boldsymbol{x}_۱, \ldots, \boldsymbol{x}_n),$$

که $\mathbf{X}$ ماتریسی $n \times p$ است.

لم ۱۵.۱.۱ (شفر و استریمر، ۲۰۰۵)**.** *فرض کنید $x_{۱i}$، ...، x_{ni} نمونه‌ای تصادفی از متغیر تصادفی X_i باشد. همچنین فرض کنید x_{ki}، k امین مشاهده X_i و $\bar{x}_i = \frac{۱}{n}\sum_{k=۱}^{n} x_{ki}$ میانگین نمونه باشد. همچنین فرض کنید $S_{ij} = \frac{۱}{n-۱}\sum_{i=۱}^{n}(x_{ki} - \bar{x}_i)(x_{ki} - \bar{x}_j)$، S_{ii} عناصر قطری ماتریس کوواریانس نمونه، $w_{kij} = (x_{ki} - \bar{x}_i)(x_{ki} - \bar{x}_j)$ و $\bar{w}_{ij} = \frac{۱}{n}\sum_{k=۱}^{n} w_{kij}$ باشد. آن‌گاه کوواریانس تجربی به‌صورت زیر است*

$$\widehat{Cov}(X_i, X_j) = S_{ij} = \frac{n}{n-۱}\bar{w}_{ij},$$

و واریانس نیز به‌صورت

$$\widehat{Var}(X_i) = S_{ii} = \frac{n}{n-۱}\bar{w}_{ii},$$

می‌باشد. همچنین به‌صورت مشابه

$$\begin{aligned}\widehat{Var}(S_{ij}) &= \frac{n^۲}{(n-۱)^۲}\widehat{Var}(\bar{w}_{ij}) \\ &= \frac{n}{(n-۱)^۲}\sum_{k=۱}^{n}(w_{kij}-\bar{w}_{ij})^۲.\end{aligned}$$

۲.۱ برآورد

تحلیل چندمتغیره بخشی از آمار است که مربوط به تحلیل داده‌ها شامل یک یا بیشتر از یک اندازه‌گیری (متغیر، صفت) در افراد یا اشیا می‌باشد. نتایج حاصل از اندازه‌گیری را بعد نامیده و با p نشان می‌دهیم. نمونه n تایی را از جمعیتی با متغیرهای p بعدی در نظر بگیرید.

اگر فرض کنیم توزیع بردار $\boldsymbol{x}_i$ متعلق به خانواده توزیع‌های احتمال pبعدی زیر باشد

$$\mathcal{F} = \{f(x;\mu,\boldsymbol{\Sigma})|\ \ \mu \in R^p, \boldsymbol{\Sigma} > ۰\}, \tag{۲.۱}$$

آن‌گاه بردار میانگین $\boldsymbol{\mu}$ و ماتریس کوواریانس $\boldsymbol{\Sigma}$ به‌صورت زیر می‌باشند

$$\boldsymbol{\mu} = \begin{pmatrix} EX_۱ \\ EX_۲ \\ \vdots \\ EX_p, \end{pmatrix} = \begin{pmatrix} \mu_۱ \\ \mu_۲ \\ \vdots \\ \mu_p \end{pmatrix},$$

$$\boldsymbol{\Sigma} = \begin{pmatrix} cov(X_۱,X_۱) & cov(X_۱,X_۲) & \dots & cov(X_۱,X_p) \\ cov(X_۲,X_۱) & cov(X_۲,X_۲) & \dots & cov(X_۲,X_p) \\ \vdots & \vdots & \ddots & \vdots \\ cov(X_p,X_۱) & cov(X_p,X_۲) & \dots & cov(X_p,X_p) \end{pmatrix} = \begin{pmatrix} \sigma_{۱۱} & \sigma_{۱۲} & \dots & \sigma_{۱p} \\ \sigma_{۲۱} & \sigma_{۲۲} & \dots & \sigma_{۲p} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{p۱} & \sigma_{p۲} & \dots & \sigma_{pp} \end{pmatrix},$$

به‌طوری که $\mu_i = E[X_i]$ و $\boldsymbol{\Sigma}_{ij} = E[(X_i-\mu_i)(X_j-\mu_j)]$ برای $i \neq j$.

به‌طور کلی در خانواده توزیع‌های (۲.۱) علاقمند به برآورد دو پارامتر بردار میانگین $\boldsymbol{\mu}$ و ماتریس کوواریانس $\boldsymbol{\Sigma}$ می‌باشیم. برآورد معمول برای $\boldsymbol{\mu}$، میانگین نمونه $\bar{\boldsymbol{x}}$ و برای $\boldsymbol{\Sigma}$، ماتریس کوواریانس نمونه $\boldsymbol{S}$ می‌باشد که بر اساس یک اندازه نمونه n تایی به‌صورت زیر می‌باشند

$$\bar{\boldsymbol{x}} = \begin{pmatrix} \bar{x}_۱ \\ \vdots \\ \bar{x}_p \end{pmatrix},$$

که در آن

$$\bar{x}_i = \frac{۱}{n}\sum_{j=۱}^{n} x_{ij},$$

میانگین iامین نمونه است. ماتریس کوواریانس نمونه نیز به‌صورت زیر قابل محاسبه است

$$\boldsymbol{S} = \frac{۱}{n-۱}\sum_{j=۱}^{n}(\boldsymbol{x}_j - \bar{\boldsymbol{x}})(\boldsymbol{x}_j - \bar{\boldsymbol{x}})^\top, \tag{۳.۱}$$

که $\boldsymbol{S}$ را می‌توان به‌صورت زیر نوشت

$$\boldsymbol{S} = \frac{۱}{n-۱}(\boldsymbol{X} - \bar{\boldsymbol{X}})(\boldsymbol{X} - \bar{\boldsymbol{X}})^\top,$$

که $\bar{\boldsymbol{X}}$ ماتریس $p \times n$ با هر ستون متشکل از $\bar{\boldsymbol{x}}$ است.

براساس یافته‌های فیشر و سان [۴۴] (۲۰۱۱)، $\bar{\boldsymbol{x}}$ و $\boldsymbol{S}$ به دلیل ناااریبی و سازگاری برآوردگرهای مناسبی برای $\boldsymbol{\mu}$ و $\boldsymbol{\Sigma}$ هستند.

قضیه ۱.۲.۱ (میرهد، ۱۹۸۲). *اگر $\boldsymbol{x}_۱$، ...، $\boldsymbol{x}_n$ بردارهای تصادفی مستقل دارای توزیع $\mathcal{N}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ باشند، به‌طوری که $n > p$، آن‌گاه برآوردگر درستنمایی ماکزیمم $\boldsymbol{\mu}$ و $\boldsymbol{\Sigma}$ به‌صورت*

$$\bar{\boldsymbol{x}} = \frac{۱}{N}\sum_{i=۱}^{n}\boldsymbol{x}_i,$$

و $\hat{\boldsymbol{\Sigma}} = \frac{n}{n+۱}\boldsymbol{S}$ می‌باشند که $\boldsymbol{S}$ به‌صورت (۳.۱) تعریف می‌شود.

۳.۱ برآوردگر انقباضی نوع استاین

در این بخش به دنبال برآوردگری مناسب برای ماتریس کوواریانس زمانی که حجم نمونه کمتر از تعداد متغیرها $(p > n)$ است، می‌باشیم. رویکردی رایج در بهبود برآورد ماتریس کوواریانس در این حالت، «انقباضی» یا «برآورد اریب» است که به‌طور گسترده توسط افرون و موریس (۱۹۷۵) بررسی شده است.

فرض کنید $\boldsymbol{S}$ برآوردگر ارائه شده در (۳.۱) است که میانگین توان‌های دوم خطای آن به‌صورت زیر محاسبه می‌شود

$$MSE(\boldsymbol{S}) = (Bias(\boldsymbol{S}))^\top(Bias(\boldsymbol{S})) + Cov(\boldsymbol{S}),$$

که در آن $Bias(\boldsymbol{S}) = \boldsymbol{S} - E(\boldsymbol{S})$. برای ماتریس کوواریانس نمونه $\boldsymbol{S}$ اریبی صفر است (فیشر و سان، ۲۰۰۹)، از این رو بهبود دقت آن در گرو کاهش واریانس است. با تعریف برآوردگری اریب، می‌توان واریانس را کاهش داد، که ایده اصلی برآوردگر جیمز و استاین (۱۹۶۱) است.

[۴۴] Sun and Fisher

در واقع در این روش ترکیبی محدب از ماتریس کوواریانس نمونه با ماتریسی هدف به‌صورت زیر در نظر گرفته می‌شود

$$\boldsymbol{S}^* = (1-\lambda)\boldsymbol{S} + \lambda \boldsymbol{T}, \qquad (۴.۱)$$

که $\lambda \in (0,1)$ پارامتر (ضریب) انقباضی[۴۵] نام دارد و $\boldsymbol{T}$ ماتریس هدف[۴۶] است که باید معین مثبت باشد. این ترکیب را طوری در نظر می‌گیریم که $\boldsymbol{S}^*$ واریانس کمتری نسبت به $\boldsymbol{S}$ داشته باشد. با توجه به این که ماتریس هدف $\boldsymbol{T}$ از قبل تعیین شده است، مساله اصلی در ارتباط با برآوردگر انقباضی، تعیین پارامتر انقباضی λ است. پارامتر انقباضی λ طوری انتخاب می‌شود که برآوردگر $\boldsymbol{S}$ را بهبود بخشد. اگر n در مقایسه با p بزرگ باشد، $\boldsymbol{S}$ باید واریانس کمتری داشته باشد و برآوردگر خوبی باشد، از این رو $\lambda \longrightarrow 0$. اگر p بزرگ باشد، $\boldsymbol{S}$ واریانس بزرگتری دارد و باید وزنی را برای ماتریس هدف فراهم کرد به طوری که $\lambda \longrightarrow 1$. انتخاب پارامتر انقباضی (λ) را می‌توان با روش‌های بوت‌استرپ[۴۷]، اعتبارسنجی متقابل[۴۸] و سایر روش‌ها تعیین کرد.

زمانی که بعد بالاست، مقادیر ویژه ماتریس کوواریانس برآوردگرهای ضعیفی برای مقادیر ویژه واقعی هستند. طبق آن، کوچکترین مقادیر ویژه به سمت صفر میل می‌کند، همچنان که بزرگترین آن به سمت بی نهایت میل می‌کند. ایده نظری، انقباض مقادیر ویژه $\boldsymbol{S}$ به مقادیر ویژه $\boldsymbol{T}$ است.

ماتریس هدف انتخاب‌شده معین مثبت (معکوس‌پذیر) بوده و باید ماتریس کوواریانس واقعی را برای کاربرد ما ارائه دهد. لازم به ذکر است که این شرط دشوار بوده و باید داده‌ها در دسترس باشند. برآوردگر انقباضی چندین امتیاز نسبت به ماتریس کوواریانس نمونه دارد. برآوردگر انقباضی $\boldsymbol{S}^*$ برای هر بعدی، ماتریس معین مثبت و نامنفرد (معکوس‌پذیر) است. لدویت و ولف[۴۹] (۲۰۰۳) قضیه‌ای را ارائه دادند که بیانگر پارامتر انقباضی مطلوب از نظر واریانس‌ـ کوواریانس در ماتریس کوواریانس نمونه و ماتریس هدف است. تحت میانگین توان‌های دوم خطا و با در نظر گرفتن نرم فربنیوس، یک مقدار بهینه برای پارامتر انقباضی وجود دارد. لدویت و ولف (۲۰۰۴) جزئیات پارامتر بهینه را برای ماتریس هدف $\boldsymbol{T} = \mu_\lambda \boldsymbol{I}$، در نظر گرفتند که μ_λ میانگین مقادیر ویژه $\boldsymbol{\Sigma}$ و $\boldsymbol{I}$ ماتریس همانی است. پارامتر انقباضی بهینه در نظر گرفته‌شده به‌صورت زیر است

$$\lambda = \frac{\beta^2}{\delta^2}$$

که در آن

$$\beta^2 = E[\|\boldsymbol{S} - \boldsymbol{\Sigma}\|^2],$$

[۴۵] Shrinkage intensity
[۴۶] Target matrix
[۴۷] Bootstrap
[۴۸] Cross-validation
[۴۹] Ledoit and Wolf

و

$$\delta^۲ = E[\|\boldsymbol{S} - \mu_\lambda \boldsymbol{I}\|^۲].$$

متاسفانه μ_λ، $\beta^۲$ و $\delta^۲$ همه به اطلاعاتی در مورد $\boldsymbol{\Sigma}$ نیاز دارند و از آن جایی که مجهول هستند، باید برآورد شوند.

لدویت و ولف (۲۰۰۴) برآوررگرهای سازگاری برای μ_λ و λ فراهم کردند. چن و همکاران[۵۰] (۲۰۱۰) با در نظر گرفتن قضیه رائو–بلکول، برآوردگر لدویت و ولف (۲۰۰۴) را بهبود بخشیدند. شفر و استریمر (۲۰۰۵) شاخص ناارییبی را نسبت به سازگاری در برآوردگرهای μ_λ و λ در نظر گرفتند.

۴.۱ انواع پارامترهای انقباضی

فرض کنید $\{\boldsymbol{x}_۱, \ldots, \boldsymbol{x}_n\}$ نمونه‌ای تصادفی از توزیع pمتغیره نرمال با بردارهای میانگین صفر و کوواریانس $\boldsymbol{\Sigma}$ باشد. نکته‌ای که باید بدان توجه نمود این است که فرض $n \geq p$ را نداریم. به دنبال برآوردگری مانند $\hat{\boldsymbol{\Sigma}}$ هستیم که تابع زیان زیر را کمینه کند

$$E\left[\left\|\hat{\boldsymbol{\Sigma}}\left(\{\boldsymbol{x}_i\}_{i=۱}^n\right) - \boldsymbol{\Sigma}\right\|_F^۲\right], \tag{۵.۱}$$

که در آن $\hat{\boldsymbol{\Sigma}}$ برآوردگری برای $\boldsymbol{\Sigma}$ می‌باشد. در واقع باید پارامتر انقباضی λ را (از روی مشاهدات) طوری یافت که تابع زیان (۵.۱) را کمینه کند. حال به معرفی پارامترهای انقباضی می‌پردازیم.

۱.۴.۱ پارامتر انقباضی LW

لدویت و ولف (۲۰۰۴) تحت فرض این‌که $\boldsymbol{x}_i$ها دارای توزیع نرمال pمتغیره با میانگین صفر و ماتریس کوواریانس $\boldsymbol{\Sigma}$ هستند، پارامتر انقباضی λ را به‌صورت زیر معرفی کردند

$$\lambda_{\text{LW}} = \min\left\{\frac{\sum_{i=۱}^n \left\|\boldsymbol{x}_i\boldsymbol{x}_i^\top - \hat{\boldsymbol{S}}\right\|_F^۲}{n^۲\,\text{tr}(\hat{\boldsymbol{S}}^۲) - \frac{۱}{p}\,\text{tr}^۲(\hat{\boldsymbol{S}})}, ۱\right\}, \tag{۶.۱}$$

که در آن

$$\hat{\boldsymbol{S}} = \frac{۱}{n}\sum_{i=۱}^n \boldsymbol{x}_i\boldsymbol{x}_i^\top.$$

[۵۰] Chen et al.

۲.۴.۱ پارامتر انقباضی RBLW

چن و همکاران (۲۰۱۰) با در نظر گرفتن قضیه رائو بلکول، برآوردگر لدویت و ولف (۲۰۰۴) را بهبود دادند. آن‌ها تحت فرضیه‌ی نرمال بودن داده‌ها و همچنین با در نظر گرفتن $\boldsymbol{S}$ به عنوان آماره بسنده، پارامتر انقباضی را به‌صورت زیر تعریف کردند

$$\lambda_{\text{RBLW}} = \min\left\{\frac{\frac{n-۲}{n}\operatorname{tr}(\hat{\boldsymbol{S}}^۲) + \operatorname{tr}^۲(\hat{\boldsymbol{S}})}{(n+۲)[\operatorname{tr}(\hat{\boldsymbol{S}}^۲) - \frac{۱}{p}\operatorname{tr}^۲(\hat{\boldsymbol{S}})]}, ۱\right\}. \tag{۷.۱}$$

برآوردگر انقباضی با استفاده از این برآوردگر برای پارامتر بهینه بر برآوردگر لدویت و ولف ارجحیت دارد. آن‌ها برای داده‌های سری زمانی و شبیه‌سازی‌های خود از این رویکرد استفاده کردند و در مثال‌های خود تاثیر برآوردگرها را مشاهده کردند.

۳.۴.۱ پارامتر انقباضی SS

شفر و استریمر (۲۰۰۵) نیز تحت فرض نرمال بودن داده‌ها، پارامتر انقباضی را به‌صورت زیر ارائه دادند

$$\lambda_{\text{SS}} = \min\left\{\frac{\sum_{i=۱}^{n}\sum_{j=۱,j\neq i}^{n}\widehat{\operatorname{Var}}(S_{ij} + \sum_{i=۱}^{n}\widehat{\operatorname{Var}}(S_{ii}))}{\sum_{i=۱}^{n}\sum_{j=۱,j\neq i}^{n}S_{ij}^۲ + \sum_{i=۱}^{n}(S_{ii}-\mu_\lambda)^۲}, ۱\right\}, \tag{۸.۱}$$

که در آن S_{ii} عناصر قطری ماتریس کوواریانس نمونه $\boldsymbol{S}$ و μ_λ میانگین عناصر قطری S_{ii} است. برای جزئیات بیشتر به شفر و استریمر (۲۰۰۵) مراجعه کنید.

۴.۴.۱ پارامتر انقباضی FS

فیشر و سان (۲۰۱۱) پارامتر انقباضی زیر را پیشنهاد دادند

$$\lambda_{\text{FS}} = \min\left\{\frac{\frac{۱}{n}\hat{a}_۱ + \frac{p}{n}\hat{a}_۱^۲}{\frac{n+۱}{n}\hat{a}_۲ + \frac{p-a}{n}\hat{a}_۱}, ۱\right\},$$

که در آن $a_i = \frac{۱}{p}\operatorname{tr}(\boldsymbol{\Sigma}^i)$. در این‌صورت

$$\begin{aligned}\hat{a}_۱ &= \frac{\operatorname{tr}(\boldsymbol{S})}{p},\\ \hat{a}_۲ &= \frac{n^۲}{(n+۱)(n+۲)}\frac{۱}{p}\left[\operatorname{tr}(\boldsymbol{S}^۲) - \frac{\operatorname{tr}^۲(\boldsymbol{S})}{n}\right],\end{aligned}$$

که برآوردگرهایی سازگار برای $a_۱$ و $a_۲$ هستند.

فصل ۲

برآورد توابع ماتریس کوواریانس در بعد بالا

۱.۲ مقدمه

برآورد ماتریس کوواریانس بزرگ موضوعی مهم در بطن تحلیل داده‌های موضوعات مدرن و پیچیده مانند ژنومیک، تحقیقات سرطان، مطالعات بالینی، شناسایی الگو و هندسه محدب محاسباتی است. معمولا در این مسائل هدف برآورد ماتریس کوواریانس Σ بر پایه یک نمونه n تایی از بردارهای تصادفی p متغیره مستقل و هم‌توزیع $\boldsymbol{x}_1$، ...، $\boldsymbol{x}_n$ با بردار میانگین $\boldsymbol{\mu}$ در حالت n کوچک و p بزرگ است. در این موارد، معمولا تعداد مشاهدات n در مقایسه با تعداد پارامترها بسیار کوچکتر است.

فرض کنید $\boldsymbol{x}_1 \dots \boldsymbol{x}_n$، بردارهای تصادفی p متغیره مستقل باشند که هر یک از $\boldsymbol{x}_i$ها می‌تواند به‌صورت زیر بیان شود

$$\boldsymbol{x}_i = \boldsymbol{\mu} + \Sigma^{\frac{1}{2}} \boldsymbol{z}_i, \tag{۱.۲}$$

که در آن $\boldsymbol{z}_i$ دارای توزیع دلخواه $\mathcal{F}$ با بردار میانگین صفر و ماتریس کوواریانس Σ می‌باشد، که Σ ماتریس معین مثبت $p \times p$ است.

در آزمون‌های آماری مربوط به تحلیل‌های چندمتغیره برای داده‌های بعد بالا، اغلب به برآورد $\mathrm{tr}(\Sigma^2)$ نیاز داریم (اسکات[۱] (۲۰۰۷)؛ سریواستاوا و فوجیکوشی[۲] (۲۰۰۶)؛ چن و کین (۲۰۱۰)؛ فوجیکوشی و همکاران (۲۰۰۴)).

سریواستاوا (۲۰۰۵) برآوردگر ناااریب و سازگار $a_2 = \mathrm{tr}(\Sigma^2)/p$ را با در نظر گرفتن فرض نرمال در قالب قضیه زیر پیشنهاد کرد.

---

[۱] Schott

[۲] Srivastava and Fujikoshi

قضیه ۱.۱.۲ (سریواستاوا، ۲۰۰۵). *برآوردگر ناااریب و سازگار $a_2 = \mathrm{tr}(\Sigma^2)/p$، تحت فرض نرمال بودن $\mathcal{F}$، به‌صورت*

$$\tilde{a}_2 = \frac{n^2}{(n-1)(n+2)}\frac{1}{p}\left[\mathrm{tr}(\boldsymbol{S}^2) - \frac{1}{n}(\mathrm{tr}(\boldsymbol{S}))^2\right],$$

است.

پس از آن، سریواستاوا و همکاران (۲۰۱۱) سازگاری برآوردگر $\tilde{a}_2$ را حتی زمانی که $\mathcal{F}$ نرمال نباشد بررسی کردند. نکته‌ای که باید به آن توجه کرد این است که برآوردگر آنها ناااریب نیست. چن و کین (۲۰۱۰) و چن و همکاران (۲۰۱۰) نیز برآوردگر سازگار دیگری برای a_2 پیشنهاد کردند. برآوردگر پیشنهادی چن و کین (۲۰۱۰)، ناااریب نیست و همین‌طور برآوردگر پیشنهادی آن‌ها نیز صورت پیچیده‌ای دارد. علاوه بر این، سازگاری این برآوردگرها تحت شرایط گشتاوری قویتری نیز به‌دست آمده است. بنابراین برآوردگرهای سازگار ناااریب برای $(\mathrm{tr}(\Sigma))^2$ و κ_{11} علاوه بر $\mathrm{tr}(\Sigma^2)$ نیز پیشنهاد شده است که در آن κ_{ij} به‌صورت

$$\kappa_{ij} = E[z^\top \Sigma^i z z^\top \Sigma^i z] - 2\,\mathrm{tr}(\Sigma^{i+j}) - \mathrm{tr}(\Sigma^i)\,\mathrm{tr}(\Sigma^j), \qquad (۲.۲)$$

تعریف می‌شود که $z \sim \mathcal{F}$. برای توزیع نرمال، $\kappa_{ij} = 0$ (مگنوس و نیودکر، ۱۹۷۹). معیار κ_{11} یکی از مهمترین پارامترهایی است که تفاوت حالت غیرنرمال را نشان می‌دهد. مطالب این فصل برگرفته از مرجع هیمنو و یامادا[۳] (۲۰۱۴) است.

۲.۲ ناارییبی و سازگاری برآوردگرها

ابتدا شرایط نظم زیر را در نظر بگیرید. وقتی $p \to \infty$، شرایط زیر برقرار هستند:

$$A_1 : \frac{n}{p} \longrightarrow c_0 \in (0, \infty),$$
$$A_2 : a_i = \frac{\mathrm{tr}\,\Sigma^i}{p} \longrightarrow a_{i0} \in (0, \infty) \quad i = 1, \ldots, 4.$$

در برآورد $\mathrm{tr}(\Sigma^2)$ اغلب از $\mathrm{tr}(\boldsymbol{S}^2)$ و $(\mathrm{tr}(\boldsymbol{S}))^2$ استفاده می‌کنیم، که امید ریاضی این آماره‌ها در قضیه زیر بیان شده است.

قضیه ۱.۲.۲. *تحت مفروضات بخش ۱.۲، بر اساس نمونه تصادفی $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$، امید ریاضی $\mathrm{tr}(\boldsymbol{S}^2)$ و $(\mathrm{tr}(\boldsymbol{S}))^2$ به‌صورت زیر به‌دست می‌آیند*

$$E[\mathrm{tr}(\boldsymbol{S}^2)] = \frac{1}{n}\kappa_{11} + \frac{n}{n-1}\mathrm{tr}(\Sigma^2) + \frac{1}{n-1}(\mathrm{tr}(\Sigma))^2, \qquad (۳.۲)$$

$$E[(\mathrm{tr}(\boldsymbol{S}))^2] = \frac{1}{n}\kappa_{11} + \frac{2}{n-1}\mathrm{tr}(\Sigma^2) + (\mathrm{tr}(\Sigma))^2. \qquad (۴.۲)$$

[۳] Himeno and Yamada

برهان. می‌توان ماتریس $\boldsymbol{S}$ را به‌صورت زیر بازنویسی کرد

$$\begin{aligned}\boldsymbol{S} &= \frac{۱}{n-۱}\sum_{i=۱}^{n}(\boldsymbol{x}_i-\bar{\boldsymbol{x}})(\boldsymbol{x}_i-\bar{\boldsymbol{x}})^\top \\ &= \frac{۱}{n-۱}\sum_{i=۱}^{n}(\boldsymbol{x}_i-\boldsymbol{\mu}+\boldsymbol{\mu}-\bar{\boldsymbol{x}})(\boldsymbol{x}_i-\boldsymbol{\mu}+\boldsymbol{\mu}-\bar{\boldsymbol{x}})^\top. \qquad (۵.۲)\end{aligned}$$

فرض کنید به‌ازای هر $i=۱,\cdots,n$، $\boldsymbol{y}_i=\boldsymbol{x}_i-\boldsymbol{\mu}$ در این‌صورت

$$\begin{aligned}\boldsymbol{x}_i-\boldsymbol{\mu}+\boldsymbol{\mu}-\bar{\boldsymbol{x}} &= \boldsymbol{y}_i+(\boldsymbol{\mu}-\frac{۱}{n}\sum_{i=۱}^{n}\boldsymbol{x}_i) \\ &= \boldsymbol{y}_i+(\frac{۱}{n}\sum_{i=۱}^{n}\boldsymbol{\mu}-\frac{۱}{n}\sum_{i=۱}^{n}\boldsymbol{x}_i) \\ &= \boldsymbol{y}_i+\frac{۱}{n}\sum_{i=۱}^{n}(\boldsymbol{\mu}-\boldsymbol{x}_i) \\ &= \boldsymbol{y}_i-\frac{۱}{n}\sum_{j=۱}^{n}\boldsymbol{y}_j \\ &= \boldsymbol{y}_i-\frac{۱}{n}\boldsymbol{y}_i-\frac{۱}{n}\sum_{\substack{j=۱\\j\neq i}}^{n}\boldsymbol{y}_j \\ &= \frac{n-۱}{n}\boldsymbol{y}_i-\frac{۱}{n}\sum_{\substack{j=۱\\j\neq i}}^{n}\boldsymbol{y}_j \qquad (۶.۲)\end{aligned}$$

بنابراین با جای‌گذاری (۶.۲) در (۵.۲) نتیجه می‌شود

$$\begin{aligned}\boldsymbol{S} &= \frac{۱}{n-۱}\sum_{i=۱}^{n}(\frac{n-۱}{n}\boldsymbol{y}_i-\frac{۱}{n}\sum_{\substack{j=۱\\j\neq i}}^{n}\boldsymbol{y}_j)(\frac{n-۱}{n}\boldsymbol{y}_i-\frac{۱}{n}\sum_{\substack{j=۱\\j\neq i}}^{n}\boldsymbol{y}_j)^\top \\ &= \frac{۱}{n-۱}\sum_{i=۱}^{n}\frac{(n-۱)^۲}{n^۲}\boldsymbol{y}_i\boldsymbol{y}_i^\top-\frac{۱}{n-۱}\sum_{i=۱}^{n}\frac{(n-۱)}{n^۲}\boldsymbol{y}_i(\sum_{\substack{j=۱\\j\neq i}}^{n}\boldsymbol{y}_j^\top) \\ &\quad -\frac{۱}{n-۱}\sum_{i=۱}^{n}\frac{(n-۱)}{n^۲}(\sum_{\substack{j=۱\\j\neq i}}^{n}\boldsymbol{y}_j)\boldsymbol{y}_i^\top+\frac{۱}{n-۱}\frac{۱}{n^۲}\sum_{i=۱}^{n}(\sum_{\substack{j=۱\\j\neq i}}^{n}\boldsymbol{y}_j)(\sum_{\substack{j=۱\\j\neq i}}^{n}\boldsymbol{y}_j^\top) \\ &= \frac{n-۱}{n^۲}\sum_{i=۱}^{n}\boldsymbol{y}_i\boldsymbol{y}_i^\top-\frac{۱}{n^۲}\sum_{i=۱}^{n}\sum_{\substack{j=۱\\j\neq i}}^{n}\boldsymbol{y}_i\boldsymbol{y}_j^\top-\frac{۱}{n^۲}\sum_{i=۱}^{n}\sum_{\substack{j=۱\\j\neq i}}^{n}\boldsymbol{y}_j\boldsymbol{y}_i^\top \\ &\quad +\frac{۱}{n^۲(n-۱)}\sum_{i=۱}^{n}(\sum_{\substack{j=۱\\j\neq i}}^{n}\boldsymbol{y}_j)(\sum_{\substack{j=۱\\j\neq i}}^{n}\boldsymbol{y}_j^\top) \\ &= \frac{n-۱}{n^۲}\sum_{i=۱}^{n}\boldsymbol{y}_i\boldsymbol{y}_i^\top-\frac{۲}{n^۲}\sum_{i=۱}^{n}\sum_{\substack{j=۱\\j\neq i}}^{n}\boldsymbol{y}_i\boldsymbol{y}_j^\top \\ &\quad +\frac{۱}{n^۲(n-۱)}\sum_{i=۱}^{n}\sum_{\substack{j=۱\\j\neq i}}^{n}\boldsymbol{y}_j\boldsymbol{y}_j^\top+\frac{۱}{n^۲(n-۱)}\sum_{i=۱}^{n}\sum_{\substack{j=۱\\j\neq i}}^{n}\sum_{\substack{k=۱\\k\neq i\\k\neq j}}^{n}\boldsymbol{y}_j\boldsymbol{y}_k^\top\end{aligned}$$

$$
\begin{aligned}
&= \frac{۱}{n}\sum_{i=۱}^{n}\boldsymbol{y}_i\boldsymbol{y}_i^\top - \frac{۱}{n^۲}\sum_{i=۱}^{n}\boldsymbol{y}_i\boldsymbol{y}_i^\top - \frac{۲}{n^۲}\sum_{i=۱}^{n}\sum_{\substack{j=۱\\ j\neq i}}^{n}\boldsymbol{y}_i\boldsymbol{y}_j^\top \\
&\quad + \frac{۱}{n^۲(n-۱)}\sum_{i=۱}^{n}\sum_{\substack{j=۱\\ j\neq i}}^{n}\boldsymbol{y}_j\boldsymbol{y}_j^\top + \frac{۱}{n^۲(n-۱)}\sum_{i=۱}^{n}\sum_{\substack{j=۱\\ j\neq i}}^{n}\sum_{\substack{k=۱\\ k\neq i\\ k\neq j}}^{n}\boldsymbol{y}_j\boldsymbol{y}_k^\top .
\end{aligned}
$$

می‌دانیم

$$
\sum_{i=۱}^{n}\sum_{\substack{j=۱\\ j\neq i}}^{n}\boldsymbol{y}_j\boldsymbol{y}_j^\top = \sum_{i=۱}^{n}(n-۱)\boldsymbol{y}_i\boldsymbol{y}_i^\top = (n-۱)\sum_{i=۱}^{n}\boldsymbol{y}_i\boldsymbol{y}_i^\top
$$

و

$$
\sum_{i=۱}^{n}\sum_{\substack{j=۱\\ j\neq i}}^{n}\sum_{\substack{k=۱\\ k\neq i\\ k\neq j}}^{n}\boldsymbol{y}_j\boldsymbol{y}_k^\top = \sum_{i=۱}^{n}\sum_{\substack{j=۱\\ j\neq i}}^{n}(n-۲)\boldsymbol{y}_i\boldsymbol{y}_j^\top = (n-۲)\sum_{i=۱}^{n}\sum_{\substack{j=۱\\ j\neq i}}^{n}\boldsymbol{y}_i\boldsymbol{y}_j^\top .
$$

در این‌صورت

$$
\begin{aligned}
\boldsymbol{S} &= \frac{۱}{n}\sum_{i=۱}^{n}\boldsymbol{y}_i\boldsymbol{y}_i^\top - \frac{۲}{n^۲}\sum_{i=۱}^{n}\sum_{\substack{j=۱\\ j\neq i}}^{n}\boldsymbol{y}_i\boldsymbol{y}_j^\top \\
&\quad + \frac{(n-۲)}{n^۲(n-۱)}\sum_{i=۱}^{n}\sum_{\substack{j=۱\\ j\neq i}}^{n}\boldsymbol{y}_i\boldsymbol{y}_j^\top \\
&= \frac{۱}{n}\sum_{i=۱}^{n}\boldsymbol{y}_i\boldsymbol{y}_i^\top + \left(\frac{n-۲}{n^۲(n-۱)} - \frac{۲}{n^۲}\right)\sum_{i=۱}^{n}\sum_{\substack{j=۱\\ j\neq i}}^{n}\boldsymbol{y}_i\boldsymbol{y}_j^\top \\
&= \frac{۱}{n}\sum_{i=۱}^{n}\boldsymbol{y}_i\boldsymbol{y}_i^\top - \frac{۱}{n(n-۱)}\sum_{i=۱}^{n}\sum_{\substack{j=۱\\ j\neq i}}^{n}\boldsymbol{y}_i\boldsymbol{y}_j^\top .
\end{aligned}
$$

از طرفی با توجه به مدل (۱.۲)

$$
\boldsymbol{y}_i = \boldsymbol{\Sigma}^{\frac{۱}{۲}}\boldsymbol{z}_i
$$

و می‌توان نوشت

$$
\boldsymbol{S} = \frac{۱}{n}\sum_{i=۱}^{n}\boldsymbol{\Sigma}^{\frac{۱}{۲}}\boldsymbol{z}_i\boldsymbol{z}_i^\top\boldsymbol{\Sigma}^{\frac{۱}{۲}} - \frac{۱}{n(n-۱)}\sum_{i=۱}^{n}\sum_{\substack{j=۱\\ j\neq i}}^{n}\boldsymbol{\Sigma}^{\frac{۱}{۲}}\boldsymbol{z}_i\boldsymbol{z}_j^\top\boldsymbol{\Sigma}^{\frac{۱}{۲}} .
$$

از عبارت بالا می‌توانیم $\mathrm{tr}(\boldsymbol{S}^۲)$ و $(\mathrm{tr}(\boldsymbol{S}))^۲$ را به‌صورت زیر بیان کنیم

$$
\begin{aligned}
\mathrm{tr}(\boldsymbol{S}^۲) &= \frac{۱}{n^۲}\sum_{i=۱}^{n}(\boldsymbol{z}_i^\top\boldsymbol{\Sigma}\boldsymbol{z}_i)^۲ + \frac{۱}{n^۲}\sum_{i=۱}^{n}\sum_{\substack{j=۱\\ j\neq i}}^{n}(\boldsymbol{z}_i^\top\boldsymbol{\Sigma}\boldsymbol{z}_j)^۲
\end{aligned}
$$

$$
\begin{aligned}
&+\frac{۱}{n^۲(n-۱)^۲}\sum_{i=۱}^{n}\sum_{j=۱}^{n}\sum_{\substack{k=۱\\k\neq j}}^{n}\sum_{\substack{l=۱\\l\neq i}}^{n}\boldsymbol{z}_i^\top\boldsymbol{\Sigma}\boldsymbol{z}_j\boldsymbol{z}_k^\top\boldsymbol{\Sigma}\boldsymbol{z}_l\\
&-\frac{۲}{n^۲(n-۱)}\sum_{i=۱}^{n}\sum_{j=۱}^{n}\sum_{\substack{k=۱\\k\neq i}}^{n}\boldsymbol{z}_i^\top\boldsymbol{\Sigma}\boldsymbol{z}_j\boldsymbol{z}_j^\top\boldsymbol{\Sigma}\boldsymbol{z}_k\\
=\;&\frac{۱}{n^۲}\sum_{i=۱}^{n}(\boldsymbol{z}_i^\top\boldsymbol{\Sigma}\boldsymbol{z}_i)^۲+\frac{n^۲-۲n+۲}{n^۲(n-۱)^۲}\sum_{i=۱}^{n}\sum_{\substack{j=۱\\j\neq i}}^{n}(\boldsymbol{z}_i^\top\boldsymbol{\Sigma}\boldsymbol{z}_j)^۲\\
&+\frac{۱}{n^۲(n-۱)^۲}\sum_{i=۱}^{n}\sum_{\substack{j=۱\\j\neq i}}^{n}\boldsymbol{z}_i^\top\boldsymbol{\Sigma}\boldsymbol{z}_i\boldsymbol{z}_j^\top\boldsymbol{\Sigma}\boldsymbol{z}_j\\
&-\frac{۴}{n^۲(n-۱)}\sum_{i=۱}^{n}\sum_{\substack{j=۱\\j\neq i}}^{n}\boldsymbol{z}_i^\top\boldsymbol{\Sigma}\boldsymbol{z}_i\boldsymbol{z}_i^\top\boldsymbol{\Sigma}\boldsymbol{z}_j\\
&+\frac{۲}{n^۲(n-۱)^۲}\sum_{i=۱}^{n}\sum_{\substack{j=۱\\j\neq i}}^{n}\sum_{\substack{k=۱\\k\neq i\\k\neq j}}^{n}\boldsymbol{z}_i^\top\boldsymbol{\Sigma}\boldsymbol{z}_i\boldsymbol{z}_j^\top\boldsymbol{\Sigma}\boldsymbol{z}_k\\
&-\frac{۲n-۴}{n^۲(n-۱)^۲}\sum_{i=۱}^{n}\sum_{\substack{j=۱\\j\neq i}}^{n}\sum_{\substack{k=۱\\k\neq i\\k\neq j}}^{n}\boldsymbol{z}_i^\top\boldsymbol{\Sigma}\boldsymbol{z}_j\boldsymbol{z}_i^\top\boldsymbol{\Sigma}\boldsymbol{z}_k\\
&+\frac{۱}{n^۲(n-۱)^۲}\sum_{i=۱}^{n}\sum_{\substack{j=۱\\j\neq i}}^{n}\sum_{\substack{k=۱\\k\neq i\\k\neq j}}^{n}\sum_{\substack{l=۱\\l\neq i\\l\neq j\\l\neq k}}^{n}\boldsymbol{z}_i^\top\boldsymbol{\Sigma}\boldsymbol{z}_j\boldsymbol{z}_k^\top\boldsymbol{\Sigma}\boldsymbol{z}_l
\end{aligned}
$$

$$
\begin{aligned}
(\mathrm{tr}(\boldsymbol{S}))^۲ =\;&\frac{۱}{n^۲}\sum_{i=۱}^{n}(\boldsymbol{z}_i^\top\boldsymbol{\Sigma}\boldsymbol{z}_i)^۲+\frac{۱}{n^۲}\sum_{i=۱}^{n}\sum_{\substack{j=۱\\j\neq i}}^{n}\boldsymbol{z}_i^\top\boldsymbol{\Sigma}\boldsymbol{z}_i\boldsymbol{z}_j^\top\boldsymbol{\Sigma}\boldsymbol{z}_j\\
&+\frac{۱}{n^۲(n-۱)^۲}\sum_{i=۱}^{n}\sum_{\substack{j=۱\\j\neq i}}^{n}\sum_{k=۱}^{n}\sum_{\substack{l=۱\\l\neq k}}^{n}\boldsymbol{z}_i^\top\boldsymbol{\Sigma}\boldsymbol{z}_j\boldsymbol{z}_k^\top\boldsymbol{\Sigma}\boldsymbol{z}_l\\
&-\frac{۲}{n^۲(n-۱)}\sum_{i=۱}^{n}\sum_{j=۱}^{n}\sum_{\substack{k=۱\\k\neq j}}^{n}\boldsymbol{z}_i^\top\boldsymbol{\Sigma}\boldsymbol{z}_i\boldsymbol{z}_j^\top\boldsymbol{\Sigma}\boldsymbol{z}_k\\
=\;&\frac{۱}{n^۲}\sum_{i=۱}^{n}(\boldsymbol{z}_i^\top\boldsymbol{\Sigma}\boldsymbol{z}_i)^۲+\frac{۲}{n^۲(n-۱)^۲}\sum_{i=۱}^{n}\sum_{\substack{j=۱\\j\neq i}}^{n}(\boldsymbol{z}_i^\top\boldsymbol{\Sigma}\boldsymbol{z}_j)^۲\\
&+\frac{۱}{n^۲}\sum_{i=۱}^{n}\sum_{\substack{j=۱\\j\neq i}}^{n}\boldsymbol{z}_i^\top\boldsymbol{\Sigma}\boldsymbol{z}_i\boldsymbol{z}_j^\top\boldsymbol{\Sigma}\boldsymbol{z}_j\\
&-\frac{۴}{n^۲(n-۱)}\sum_{i=۱}^{n}\sum_{\substack{j=۱\\j\neq i}}^{n}\boldsymbol{z}_i^\top\boldsymbol{\Sigma}\boldsymbol{z}_i\boldsymbol{z}_i^\top\boldsymbol{\Sigma}\boldsymbol{z}_j
\end{aligned}
$$

$$
\begin{aligned}
&-\frac{۲}{n^{۲}(n-۱)}\sum_{i=۱}^{n}\sum_{\substack{j=۱\\ j\neq i}}^{n}\sum_{\substack{k=۱\\ k\neq j\\ k\neq i}}^{n} z_i^\top \boldsymbol{\Sigma} z_i z_j^\top \boldsymbol{\Sigma} z_k \\
&+\frac{۴}{n^{۲}(n-۱)^{۲}}\sum_{i=۱}^{n}\sum_{\substack{j=۱\\ j\neq i}}^{n}\sum_{\substack{k=۱\\ k\neq j\\ k\neq i}}^{n} z_i^\top \boldsymbol{\Sigma} z_i z_i^\top \boldsymbol{\Sigma} z_k \\
&+\frac{۱}{n^{۲}(n-۱)^{۲}}\sum_{i=۱}^{n}\sum_{\substack{j=۱\\ j\neq i}}^{n}\sum_{\substack{k=۱\\ k\neq j\\ k\neq i}}^{n}\sum_{\substack{l=۱\\ l\neq k\\ l\neq j\\ l\neq i}}^{n} z_i^\top \boldsymbol{\Sigma} z_j z_k^\top \boldsymbol{\Sigma} z_l
\end{aligned}
$$

□

از آنجایی که سه پارامتر مجهول وجود دارند، برآورد $\operatorname{tr}(\boldsymbol{\Sigma}^{۲})$ تنها با استفاده از $\operatorname{tr}(\boldsymbol{S}^{۲})$ و $(\operatorname{tr}(\boldsymbol{S}))^{۲}$ غیرممکن است. بنابراین از آماره زیر استفاده می‌کنیم

$$
Q = \frac{۱}{n-۱}\sum_{i=۱}^{n}\Big((\boldsymbol{x}_i-\bar{\boldsymbol{x}})^\top(\boldsymbol{x}_i-\bar{\boldsymbol{x}})\Big)^{۲} \tag{۷.۲}
$$

قضیه ۲.۲.۲. *برای مدل* (۱.۲) *برآوردگرهای نااریب* $\operatorname{tr}(\boldsymbol{\Sigma}^{۲})$، $(\operatorname{tr}(\boldsymbol{\Sigma}))^{۲}$ و $\kappa_{۱۱}$ *به‌صورت زیر به‌دست می‌آیند*

$$
\begin{aligned}
\widehat{\operatorname{tr}(\boldsymbol{\Sigma}^{۲})} &= \frac{n-۱}{n(n-۲)(n-۳)}\{(n-۱)(n-۲)\operatorname{tr}(\boldsymbol{S}^{۲})+(\operatorname{tr}(\boldsymbol{S}))^{۲}-nQ\}\\
\widehat{(\operatorname{tr}(\boldsymbol{\Sigma}))^{۲}} &= \frac{n-۱}{n(n-۲)(n-۳)}\{۲\operatorname{tr}(\boldsymbol{S}^{۲})+(n^{۲}-۳n+۱)(\operatorname{tr}(\boldsymbol{S}))^{۲}-nQ\}\\
\hat{\kappa}_{۱۱} &= \frac{-۱}{(n-۲)(n-۳)}\{۲(n-۱)^{۲}\operatorname{tr}(\boldsymbol{S}^{۲})+(n-۱)^{۲}(\operatorname{tr}(\boldsymbol{S}))^{۲}-n(n+۱)Q\}
\end{aligned}
$$

برهان. می‌توان $\boldsymbol{Q}$ را به‌صورت زیر نوشت

$$
\begin{aligned}
\boldsymbol{Q} &= \frac{۱}{n-۱}\sum_{i=۱}^{n}\{(\boldsymbol{y}_i^\top\boldsymbol{y}_i)^{۲}+۴(\boldsymbol{y}_i^\top\bar{\boldsymbol{y}})^{۲}+(\bar{\boldsymbol{y}}^\top\bar{\boldsymbol{y}})^{۲}\\
&\quad +۲\boldsymbol{y}_i^\top\boldsymbol{y}_i\bar{\boldsymbol{y}}^\top\bar{\boldsymbol{y}}-۴\boldsymbol{y}_i^\top\boldsymbol{y}_i\boldsymbol{y}_i^\top\bar{\boldsymbol{y}}-۴\boldsymbol{y}_i^\top\bar{\boldsymbol{y}}\bar{\boldsymbol{y}}^\top\bar{\boldsymbol{y}}\}\\
&= \frac{۱}{n-۱}\sum_{i=۱}^{n}(\boldsymbol{y}_i^\top\boldsymbol{y}_i)^{۲}+\frac{۴}{n^{۲}(n-۱)}\sum_{i=۱}^{n}\left(\boldsymbol{y}_i^\top\sum_{j=۱}^{n}\boldsymbol{y}_j\right)^{۲}\\
&\quad +\frac{۲}{n^{۲}(n-۱)}\sum_{i=۱}^{n}\boldsymbol{y}_i^\top\boldsymbol{y}_i\left(\sum_{j=۱}^{n}\sum_{k=۱}^{n}\boldsymbol{y}_j^\top\boldsymbol{y}_k\right)\\
&\quad -\frac{۴}{n(n-۱)}\sum_{i=۱}^{n}\boldsymbol{y}_i^\top\boldsymbol{y}_i\boldsymbol{y}_i^\top\left(\sum_{j=۱}^{n}\boldsymbol{y}_j\right)-\frac{۳n}{n-۱}(\bar{\boldsymbol{y}}^\top\bar{\boldsymbol{y}})^{۲}\\
&= \frac{n^{۲}-۳n+۳}{n^{۳}}\sum_{i=۱}^{n}(z_i^\top\boldsymbol{\Sigma}z_i)^{۲}+\frac{۴n-۶}{n^{۳}(n-۱)}\sum_{i=۱}^{n}\sum_{\substack{j=۱\\ j\neq i}}^{n}(z_i^\top\boldsymbol{\Sigma}z_j)^{۲}\\
&\quad +\frac{۲n-۳}{n^{۳}(n-۱)}\sum_{i=۱}^{n}\sum_{\substack{j=۱\\ j\neq i}}^{n}z_i^\top\boldsymbol{\Sigma}z_iz_j^\top\boldsymbol{\Sigma}z_j
\end{aligned}
$$

$$
\begin{aligned}
&-\frac{۴(n^۲-۳n+۳)}{n^۳(n-۱)}\sum_{i=۱}^{n}\sum_{\substack{j=۱\\ j\neq i}}^{n} z_i^\top \boldsymbol{\Sigma} z_i z_i^\top \boldsymbol{\Sigma} z_j \\
&+\frac{۲n-۶}{n^۳(n-۱)}\sum_{i=۱}^{n}\sum_{\substack{j=۱\\ j\neq i}}^{n}\sum_{\substack{k=۱\\ k\neq j\\ k\neq i}}^{n} z_i^\top \boldsymbol{\Sigma} z_i z_j^\top \boldsymbol{\Sigma} z_k \\
&+\frac{۴n-۱۲}{n^۳(n-۱)}\sum_{i=۱}^{n}\sum_{\substack{j=۱\\ j\neq i}}^{n}\sum_{\substack{k=۱\\ k\neq j\\ k\neq i}}^{n} z_i^\top \boldsymbol{\Sigma} z_j z_i^\top \boldsymbol{\Sigma} z_k \\
&-\frac{۳}{n^۳(n-۱)}\sum_{i=۱}^{n}\sum_{\substack{j=۱\\ j\neq i}}^{n}\sum_{\substack{k=۱\\ k\neq j\\ k\neq i}}^{n}\sum_{\substack{l=۱\\ l\neq k\\ l\neq j\\ l\neq i}}^{n} z_i^\top \boldsymbol{\Sigma} z_j z_k^\top \boldsymbol{\Sigma} z_l. \qquad (۸.۲)
\end{aligned}
$$

مقدار امید ریاضی Q به‌صورت زیر بیان می‌شود

$$
E[Q] = \frac{n^۲-۳n+۳}{n^۲}\kappa_{۱۱} + \frac{۲(n-۱)}{n}\operatorname{tr}(\boldsymbol{\Sigma}^۲) + \frac{n-۱}{n}(\operatorname{tr}(\boldsymbol{\Sigma}))^۲. \qquad (۹.۲)
$$

از حل کردن همزمان معادله‌های (۳.۲)، (۴.۲) و (۹.۲) قضیه اثبات می‌شود. □

اکنون سازگاری $\hat{a}_۲ = \widehat{\operatorname{tr}(\boldsymbol{\Sigma}^۲)}/p$ و $\hat{a}_۱^۲ = \widehat{(\operatorname{tr}(\boldsymbol{\Sigma}))}^۲/p^۲$ را بررسی می‌کنیم. با استفاده از تعاریف $\widehat{\operatorname{tr}(\boldsymbol{\Sigma}^۲)}$ و $\widehat{(\operatorname{tr}(\boldsymbol{\Sigma}))}^۲$، این آماره‌ها به‌صورت زیر بیان می‌شوند

$$
\begin{aligned}
\widehat{\operatorname{tr}(\boldsymbol{\Sigma}^۲)} &= \frac{۱}{n(n-۱)}\sum_{i=۱}^{n}\sum_{\substack{j=۱\\ j\neq i}}^{n}(z_i^\top \boldsymbol{\Sigma} z_j)^۲ \\
&\quad -\frac{۲}{n(n-۱)(n-۲)}\sum_{i=۱}^{n}\sum_{\substack{j=۱\\ j\neq i}}^{n}\sum_{\substack{k=۱\\ k\neq i\\ k\neq j}}^{n} z_i^\top \boldsymbol{\Sigma} z_j z_i^\top \boldsymbol{\Sigma} z_k \\
&\quad +\frac{۱}{n(n-۱)(n-۲)(n-۳)}\sum_{i=۱}^{n}\sum_{\substack{j=۱\\ j\neq i}}^{n}\sum_{\substack{k=۱\\ k\neq i\\ k\neq j}}^{n}\sum_{\substack{l=۱\\ l\neq i\\ l\neq j\\ l\neq k}}^{n} z_i^\top \boldsymbol{\Sigma} z_j z_k^\top \boldsymbol{\Sigma} z_l \\
\widehat{(\operatorname{tr}(\boldsymbol{\Sigma}))}^۲ &= \frac{۱}{n(n-۱)}\sum_{i=۱}^{n}\sum_{\substack{j=۱\\ j\neq i}}^{n} z_i^\top \boldsymbol{\Sigma} z_i z_j^\top \boldsymbol{\Sigma} z_j \\
&\quad -\frac{۲}{n(n-۱)(n-۲)}\sum_{i=۱}^{n}\sum_{\substack{j=۱\\ j\neq i}}^{n}\sum_{\substack{k=۱\\ k\neq i\\ k\neq j}}^{n} z_i^\top \boldsymbol{\Sigma} z_i z_j^\top \boldsymbol{\Sigma} z_k \\
&\quad +\frac{۱}{n(n-۱)(n-۲)(n-۳)}\sum_{i=۱}^{n}\sum_{\substack{j=۱\\ j\neq i}}^{n}\sum_{\substack{k=۱\\ k\neq i\\ k\neq j}}^{n}\sum_{\substack{l=۱\\ l\neq i\\ l\neq j\\ l\neq k}}^{n} z_i^\top \boldsymbol{\Sigma} z_j z_k^\top \boldsymbol{\Sigma} z_l.
\end{aligned}
$$

لم ۳.۲.۲ (هیمنو و یامادا، ۲۰۰۹). *برای مدل* (۱.۲) *واریانس‌های* $\widehat{\mathrm{tr}(\boldsymbol{\Sigma}^2)}$ و $\widehat{(\mathrm{tr}(\boldsymbol{\Sigma}))^2}$ که در قضیه ۲.۲.۲ تعریف شد، به‌صورت زیر به‌دست می‌آیند

$$\begin{aligned}
\mathrm{Var}(\widehat{\mathrm{tr}(\boldsymbol{\Sigma}^2)}) &= \frac{2}{n(n-1)}E[(z_1^\top \boldsymbol{\Sigma} z_2)^4] + \frac{4(n-3)}{n(n-2)}\kappa_{22} \\
&\quad - \frac{8}{n(n-1)(n-2)}E[(z_1^\top \boldsymbol{\Sigma} z_2)^2 z_1^\top \boldsymbol{\Sigma}^2 z_2] \\
&\quad - \frac{2(n+1)(n-4)}{n(n-1)(n-2)(n-3)}(\mathrm{tr}(\boldsymbol{\Sigma}^2))^2 \\
&\quad + \frac{4(2n^2-13n+23)}{n(n-2)(n-3)}\mathrm{tr}(\boldsymbol{\Sigma}^4) \\
\mathrm{Var}((\widehat{\mathrm{tr}(\boldsymbol{\Sigma})})^2) &= \frac{2}{n(n-1)}\kappa_{11}^2 + \frac{4(2n-5)}{n(n-1)(n-2)}\kappa_{11}\mathrm{tr}(\boldsymbol{\Sigma}^2) + \frac{4}{n}\kappa_{11}(\mathrm{tr}(\boldsymbol{\Sigma}))^2 \\
&\quad - \frac{8}{n(n-1)(n-2)}E[z_1^\top \boldsymbol{\Sigma} z_1 z_1^\top \boldsymbol{\Sigma}^2 z_2 z_2^\top \boldsymbol{\Sigma} z_2] \\
&\quad + \frac{8(n^2-6n+10)}{n(n-1)(n-2)(n-3)}(\mathrm{tr}(\boldsymbol{\Sigma}^2))^2 \\
&\quad + \frac{4(2n-3)}{n(n-1)}\mathrm{tr}(\boldsymbol{\Sigma}^2)(\mathrm{tr}(\boldsymbol{\Sigma}))^2 \\
&\quad + \frac{16}{n(n-1)(n-2)(n-3)}\mathrm{tr}(\boldsymbol{\Sigma}^4).
\end{aligned}$$

از لم ۳.۲.۲ قضیه زیر نتیجه می‌شود. ابتدا شرایط زیر را در نظر بگیرید

$$\begin{aligned}
&C_1 : E[(z_1^\top \Sigma z_2)^4] = o(p^4),\ \kappa_{22} = o(p^3) \quad z_1, z_2 \sim \mathcal{F}, \\
&C_2 : \kappa_{11} = o(p^3).
\end{aligned}$$

قضیه ۴.۲.۲ (هیمنو و یامادا، ۲۰۰۹). *مدل* (۱.۲) *را در نظر بگیرید. با توجه به شرایط نظم* A_1 *و* A_2 *و همچنین شرط* C_1، $\hat{a}_2$ *برآوردگر سازگار* a_2 *می‌باشد. علاوه بر این تحت شرایط نظم* A_1، A_2 *و همچنین شرط* C_2، $\widehat{a_1^2}$ *برآوردگر سازگار* a_1^2 *می‌باشد.*

نتایج زیر با استفاده از نامساوی کوشی ـ شوارتز به‌دست آمده‌اند

تحت شرط C_1

$$\begin{aligned}
E[|(z_1^\top \boldsymbol{\Sigma} z_2)^2 z_1^\top \boldsymbol{\Sigma}^2 z_2|] &\le \sqrt{E[(z_1^\top \boldsymbol{\Sigma} z_2)^4]E[(z_1^\top \boldsymbol{\Sigma}^2 z_2)^2]} \\
&= \sqrt{E[(z_1^\top \boldsymbol{\Sigma} z_2)^4]\,\mathrm{tr}(\boldsymbol{\Sigma}^4)} = \sqrt{o(p^4)o(p)} \qquad (10.2)
\end{aligned}$$

در نتیجه، $E[(z_1^\top \boldsymbol{\Sigma} z_2)^2 z_1^\top \boldsymbol{\Sigma}^2 z_2] = o(p^{\frac{5}{2}})$. *تحت شرط* C_2

$$\begin{aligned}
E[|z_1^\top \boldsymbol{\Sigma} z_1)^2 z_1^\top \boldsymbol{\Sigma}^2 z_2 z_2^\top \boldsymbol{\Sigma} z_2|] &\le \sqrt{E[(z_1^\top \boldsymbol{\Sigma} z_1)^2 (z_2^\top \boldsymbol{\Sigma} z_2)^2]E[(z_1^\top \boldsymbol{\Sigma}^2 z_2)^2]} \\
&= \sqrt{\{\kappa_{11} + 2\,\mathrm{tr}(\boldsymbol{\Sigma}^2) + (\mathrm{tr}(\boldsymbol{\Sigma}))^2\}^2\,\mathrm{tr}(\boldsymbol{\Sigma}^4)} \\
&= o(p^{\frac{7}{2}}). \qquad (11.2)
\end{aligned}$$

۳.۲ شبیه‌سازی‌های عددی

در این بخش، کارایی برآوردگرهای پیشنهادی را با استفاده از شبیه‌سازی مونت‌کارلو ارزیابی می‌کنیم.

با توجه به این‌که $\hat{a}_۲$، $\widehat{a_۱^۲}$ و $\hat{\kappa}_{۱۱}$ برای μ پایا هستند، فرض کنیم $\boldsymbol{x}_i = \boldsymbol{\Sigma}^{\frac{۱}{۲}} \boldsymbol{z}_i$ (به ازای $i = ۱, \ldots, n$) که $\boldsymbol{z}_i$ ها نمونه تصادفی و مستقل از رده $\mathcal{F}$ هستند. ماتریس $\boldsymbol{\Sigma}$ را به‌صورت $\Sigma_{ij} = (۰/۲)^{|i-j|}$ و حجم نمونه n را برابر با ۴۰، ۸۰ یا ۱۲۰ و نیز بعد پارامتر p را با ۴۰، ۸۰ یا ۱۲۰ در نظر می‌گیریم. چهار حالت زیر را برای رده توزیع‌های $\mathcal{F}$ فرض می‌کنیم

$D_۱$ توزیع نرمال استاندارد، $N(۰, ۱)$.

$D_۲$ توزیع t با $\frac{p}{۴}$ درجه آزادی $t_{\frac{p}{۴}}$.

$D_۳$ توزیع $\chi^۲$ با ۱ درجه آزادی، $\chi_۱^۲$.

$D_۴$ توزیع p متغیره t با ۱۰ درجه آزادی، $Mt_{۱۰}$.

دقت کنید که توزیع $\mathcal{F}$ استاندارد شده است، بنابراین $\mathrm{Var}(\boldsymbol{z}) = \boldsymbol{I}_p$. تحت این شرط، برآوردهای $\hat{a}_۲$، $\tilde{a}_۲$، $\widehat{a_۱^۲}$ و $\tilde{a}_۱^۲ = (\frac{\mathrm{tr}(\boldsymbol{S})}{p})^۲$ و نیز $\frac{\hat{\kappa}_{۱۱}}{p^۲}$ (واریانس نمونه نااریب از این برآوردها) و مقادیر واقعی پارامترها را وقتی $i = ۱$ در حالت‌های $D_۱$، $D_۲$، $D_۳$ و به ازای $i = ۲$ در حالت $D_۴$ با تکرار ۱۰۰۰۰ به‌دست می‌آوریم.

جدول ۱.۲ مقادیر واقعی $a_۲$ و برآوردهای $\hat{a}_۲$ و $\tilde{a}_۲$ را نشان می‌دهد. مقادیر داخل پرانتز خطاهای استاندارد را بیان می‌کند. کارایی $\tilde{a}_۲$ برای همه توزیع‌ها خوب است به ویژه اگر توزیع نرمال باشد خطای استاندارد $\hat{a}_۲$ نزدیک خطای استاندارد $\tilde{a}_۲$ است. اگر توزیع نرمال نباشد، کارایی $\tilde{a}_۲$ چندان خوب نیست.

جدول ۲.۲ مقادیر واقعی $a_۱^۲$ و برآوردهای $\widehat{a_۱^۲}$ و $\tilde{a}_۱^۲$ را نشان می‌دهد. بر اساس این جدول به نظر می‌رسد $\tilde{a}_۱^۲$ بیش‌برازش شده است. اما بر اساس این جدول تفاوت بین $\hat{a}_۱^۲$ و $\tilde{a}_۱^۲$ به سختی تایید می‌شود.

جدول ۳.۲ مقادیر واقعی $\frac{\kappa_{۱۱}}{p^۲}$ و برآوردهای $\frac{\hat{\kappa}_{۱۱}}{p^۲}$ را به ازای $i = ۱$ برای حالت‌های $D_۱$، $D_۲$، $D_۳$ و $i = ۲$ در حالت $D_۴$ نشان می‌دهد. وقتی $\mathcal{F}$ متقارن باشد، به نظر می‌رسد تقریب مناسب باشد. با این وجود اگر $\mathcal{F}$ کی‌دو باشد، تقریب‌ها بد می‌شود. در این حالت به p و n بزرگتری نیاز داریم. همچنین به ازای n معین، با افزایش p، مقدار برآوردها افزایش می‌یابد.

جدول ۱.۲: مقدار واقعی a_2 و برآوردهای $\hat{a}_2$ و $\tilde{a}_2$ (و خطاهای استاندارد)

توزیع	n	p	a_2	$\hat{a}_2$	$\tilde{a}_2$
		۴۰	۱/۰۸۱	۱/۰۸۱(۰/۱۱۰)	۱/۰۸۱(۰/۱۰۸)
	۴۰	۸۰	۱/۰۸۲	۱/۰۸۱(۰/۰۹۷)	۱/۰۸۱(۰/۰۹۵)
		۱۲۰	۱/۰۸۳	۱/۰۸۲(۰/۰۸۴)	۱/۰۸۲(۰/۰۸۲)
		۴۰	۱/۰۸۱	۱/۰۸۱(۰/۰۷۶)	۱/۰۸۱(۰/۰۷۴)
D_1	۸۰	۸۰	۱/۰۸۲	۱/۰۸۲(۰/۰۶۳)	۱/۰۸۲(۰/۰۶۲)
		۱۲۰	۱/۰۸۳	۱/۰۸۳(۰/۰۵۸)	۱/۰۸۳(۰/۰۵۵)
		۴۰	۱/۰۸۱	۱/۰۸۱(۰/۰۶۵)	۱/۰۸۱(۰/۰۶۲)
	۱۲۰	۸۰	۱/۰۸۲	۱/۰۸۳(۰/۰۵۰)	۱/۰۸۳(۰/۰۴۸)
		۱۲۰	۱/۰۸۳	۱/۰۸۳(۰/۰۴۰)	۱/۰۸۳(۰/۰۴۹)
		۴۰	۱/۰۸۱	۱/۰۸۰(۰/۱۱۲)	۱/۱۰۴(۰/۱۵۶)
	۴۰	۸۰	۱/۰۸۲	۱/۰۸۴(۰/۰۹۵)	۱/۰۹۲(۰/۰۹۰)
		۱۲۰	۱/۰۸۳	۱/۰۸۲(۰/۰۸۷)	۱/۰۸۸(۰/۰۸۱)
		۴۰	۱/۰۸۱	۱/۰۸۲(۰/۰۷۴)	۱/۰۹۴(۰/۰۷۴)
D_2	۸۰	۸۰	۱/۰۸۲	۱/۰۸۲(۰/۰۵۸)	۱/۰۸۷(۰/۰۵۷)
		۱۲۰	۱/۰۸۳	۱/۰۸۳(۰/۰۴۳)	۱/۰۸۶(۰/۰۴۲)
		۴۰	۱/۰۸۱	۱/۰۸۲(۰/۰۶۹)	۱/۰۸۹(۰/۰۶۷)
	۱۲۰	۸۰	۱/۰۸۲	۱/۰۸۳(۰/۰۵۹)	۱/۰۸۵(۰/۰۵۷)
		۱۲۰	۱/۰۸۳	۱/۰۸۲(۰/۰۳۷)	۱/۰۸۴(۰/۰۳۸)
		۴۰	۱/۰۸۱	۱/۰۸۱(۰/۲۱۳)	۱/۳۶۷(۰/۳۴۲)
	۴۰	۸۰	۱/۰۸۲	۱/۰۸۱(۰/۱۷۱)	۱/۰۸۱(۰/۰۹۵)
		۱۲۰	۱/۰۸۳	۱/۰۸۴(۰/۱۴۹)	۱/۳۷۰(۰/۲۱۵)
		۴۰	۱/۰۸۱	۱/۰۸۲(۰/۱۵۶)	۱/۲۲۴(۰/۱۹۲)
D_3	۸۰	۸۰	۱/۰۸۲	۱/۰۸۴(۰/۱۱۲)	۱/۲۳۱(۰/۱۵۲)
		۱۲۰	۱/۰۸۳	۱/۰۸۳(۰/۰۹۱)	۱/۲۲۸(۰/۱۱۲)
		۴۰	۱/۰۸۱	۱/۰۸۱(۰/۱۲۴)	۱/۰۸۱(۰/۱۵۴)
	۱۲۰	۸۰	۱/۰۸۲	۱/۰۸۲(۰/۰۹۳)	۱/۰۸۱(۰/۰۹۹)
		۱۲۰	۱/۰۸۳	۱/۰۸۲(۰/۰۷۴)	۱/۰۸۰(۰/۰۶۹)
		۴۰	۱/۰۸۱	۱/۰۸۲(۰/۲۴۰)	۱/۴۲۱(۰/۵۹۸)
	۴۰	۸۰	۱/۰۸۲	۱/۰۸۰(۰/۲۳۱)	۱/۷۱۹(۰/۶۹۸)
		۱۲۰	۱/۰۸۳	۱/۰۸۴(۰/۲۲۷)	۲/۰۵۹(۰/۹۸۴)
		۴۰	۱/۰۸۱	۱/۰۷۸(۰/۱۶۱)	۱/۲۴۰(۰/۲۴۸)
D_4	۸۰	۸۰	۱/۰۸۲	۱/۰۸۲(۰/۱۵۲)	۱/۴۱۵(۰/۳۴۹)
		۱۲۰	۱/۰۸۳	۱/۰۸۵(۰/۱۴۸)	۱/۵۴۹(۰/۴۱۷)
		۴۰	۱/۰۸۱	۱/۰۸۱(۰/۱۲۷)	۱/۱۸۴(۰/۱۷۹)
	۱۲۰	۸۰	۱/۰۸۲	۱/۰۸۱(۰/۱۲۵)	۱/۳۰۷(۰/۲۱۳)
		۱۲۰	۱/۰۸۳	۱/۰۸۱(۰/۱۲۲)	۱/۴۱۸(۰/۲۷۹)

جدول ۲.۲: مقدار واقعی a_1^2 و برآوردهای $\widehat{a_1^2}$ و $\tilde{a}_1^2$ (و خطاهای استاندارد)

توزیع	N	p	a_1^2	$\hat{a}_1^2$	$\tilde{a}_1^2$
D_1	۴۰	۴۰	۱/۰۰۰	۱/۰۰۰(۰/۰۷۵)	۱/۰۰۱(۰/۰۷۵)
		۸۰	۱/۰۰۰	۱/۰۰۱(۰/۰۵۳)	۱/۰۰۱(۰/۰۵۳)
		۱۲۰	۱/۰۰۰	۰/۹۹۹(۰/۰۴۲)	۱/۰۰۰(۰/۰۴۲)
	۸۰	۴۰	۱/۰۰۰	۱/۰۰۰(۰/۰۵۲)	۱/۰۰۰(۰/۰۵۲)
		۸۰	۱/۰۰۰	۱/۰۰۰(۰/۰۳۷)	۱/۰۰۱(۰/۰۳۷)
		۱۲۰	۱/۰۰۰	۱/۰۰۱(۰/۰۳۰)	۱/۰۰۱(۰/۰۳۰)
	۱۲۰	۴۰	۱/۰۰۰	۱/۰۰۰(۰/۰۴۳)	۱/۰۰۰(۰/۰۴۳)
		۸۰	۱/۰۰۰	۱/۰۰۰(۰/۰۳۰)	۱/۰۰۰(۰/۰۳۰)
		۱۲۰	۱/۰۰۰	۱/۰۰۰(۰/۰۲۵)	۱/۰۰۰(۰/۰۲۵)
D_2	۴۰	۴۰	۱/۰۰۰	۰/۹۹۹(۰/۰۹۰)	۱/۰۰۱(۰/۰۹۰)
		۸۰	۱/۰۰۰	۱/۰۰۱(۰/۰۵۷)	۱/۰۰۲(۰/۰۵۷)
		۱۲۰	۱/۰۰۰	۱/۰۰۰(۰/۰۴۶)	۱/۰۰۱(۰/۰۴۶)
	۸۰	۴۰	۱/۰۰۰	۱/۰۰۱(۰/۰۶۴)	۱/۰۰۲(۰/۰۶۴)
		۸۰	۱/۰۰۰	۱/۰۰۰(۰/۰۴۰)	۱/۰۰۲(۰/۰۴۰)
		۱۲۰	۱/۰۰۰	۱/۰۰۰(۰/۰۳۲)	۱/۰۰۱(۰/۰۳۲)
	۱۲۰	۴۰	۱/۰۰۰	۱/۰۰۰(۰/۰۵۲)	۱/۰۰۲(۰/۰۵۲)
		۸۰	۱/۰۰۰	۱/۰۰۰(۰/۰۳۲)	۱/۰۰۱(۰/۰۳۲)
		۱۲۰	۱/۰۰۰	۱/۰۰۰(۰/۰۲۶)	۱/۰۰۰(۰/۰۲۶)
D_3	۴۰	۴۰	۱/۰۰۰	۱/۰۰۲(۰/۱۹۰)	۱/۰۱۱(۰/۱۹۳)
		۸۰	۱/۰۰۰	۰/۹۹۹(۰/۱۳۳)	۱/۰۰۳(۰/۱۳۴)
		۱۲۰	۱/۰۰۰	۱/۰۰۱(۰/۱۰۹)	۱/۰۰۴(۰/۱۱۰)
	۸۰	۴۰	۱/۰۰۰	۱/۰۰۱(۰/۱۳۱)	۱/۰۰۶(۰/۱۳۲)
		۸۰	۱/۰۰۰	۱/۰۰۲(۰/۰۹۵)	۱/۰۰۴(۰/۰۹۵)
		۱۲۰	۱/۰۰۰	۱/۰۰۱(۰/۰۷۷)	۱/۰۰۲(۰/۰۷۷)
	۱۲۰	۴۰	۱/۰۰۰	۱/۰۰۰(۰/۱۱۰)	۱/۰۰۳(۰/۱۱۰)
		۸۰	۱/۰۰۰	۱/۰۰۰(۰/۰۷۷)	۱/۰۰۲(۰/۰۷۷)
		۱۲۰	۱/۰۰۰	۰/۹۹۹(۰/۰۶۳)	۱/۰۰۱(۰/۰۶۳)
D_4	۴۰	۴۰	۱/۰۰۰	۱/۰۰۰(۰/۲۰۵)	۱/۰۱۰(۰/۲۱۲)
		۸۰	۱/۰۰۰	۰/۹۹۶(۰/۱۹۱)	۱/۰۰۵(۰/۱۹۵)
		۱۲۰	۱/۰۰۰	۱/۰۰۲(۰/۱۹۰)	۱/۰۱۱(۰/۱۹۵)
	۸۰	۴۰	۱/۰۰۰	۰/۹۹۹(۰/۱۴۳)	۱/۰۰۴(۰/۱۴۵)
		۸۰	۱/۰۰۰	۰/۹۹۹(۰/۱۳۵)	۱/۰۰۴(۰/۱۳۷)
		۱۲۰	۱/۰۰۰	۰/۹۹۹(۰/۱۳۴)	۱/۰۰۳(۰/۱۳۶)
	۱۲۰	۴۰	۱/۰۰۰	۱/۰۰۰(۰/۱۱۶)	۱/۰۰۳(۰/۱۱۷)
		۸۰	۱/۰۰۰	۰/۹۹۹(۰/۱۱۱)	۱/۰۰۲(۰/۱۱۲)
		۱۲۰	۱/۰۰۰	۰/۹۹۹(۰/۱۰۹)	۱/۰۰۲(۰/۱۱۰)

جدول ۳.۲: مقدار واقعی κ_{11}/p و برآوردهای $\hat{\kappa_{11}}/p$ (و خطاهای استاندارد)

توزیع	n	p	κ_{11}/p	$\hat{\kappa}_{11}/p$
D_1	۴۰	۴۰	۰/۰۰۰	−۰/۰۰۴(۰/۵۶۴)
		۸۰	۰/۰۰۰	۰/۰۰۵(۰/۵۱۵)
		۱۲۰	۰/۰۰۰	۰/۰۰۱(۰/۵۲۷)
	۸۰	۴۰	۰/۰۰۰	−۰/۰۰۷(۰/۳۴۹)
		۸۰	۰/۰۰۰	۰/۰۰۰(۰/۳۱۷)
		۱۲۰	۰/۰۰۰	۰/۰۰۴(۰/۳۵۹)
	۱۲۰	۳۸	۰/۰۰۰	۰/۰۰۲(۰/۲۹۴)
		۸۰	۰/۰۰۰	۰/۰۰۱(۰/۲۷۴)
		۱۲۰	۰/۰۰۰	۰/۰۰۴(۰/۲۸۷)
D_2	۴۰	۴۰	۱/۰۰۰	۰/۹۹۷(۰/۹۴۴)
		۸۰	۰/۳۷۵	۰/۳۶۹(۰/۶۴۱)
		۱۲۰	۰/۲۳۱	۰/۲۳۶(۰/۵۴۸)
	۸۰	۴۰	۱/۰۰۰	۱/۰۱۱(۰/۶۸۸)
		۸۰	۰/۳۷۵	۰/۳۶۴(۰/۴۲۲)
		۱۲۰	۰/۲۳۱	۰/۲۲۸(۰/۳۹۴)
	۱۲۰	۴۰	۱/۰۰۰	۰/۹۹۶(۰/۵۷۸)
		۸۰	۰/۳۷۵	۰/۳۷۵(۰/۳۴۹)
		۱۲۰	۰/۲۳۱	۰/۲۲۴(۰/۳۱۶)
D_3	۴۰	۴۰	۱۲/۰۰۰	۱۲/۰۵۸(۶/۸۷۹)
		۸۰	۱۲/۰۰۰	۱۱/۹۶۰(۵/۵۹۸)
		۱۲۰	۱۲/۰۰۰	۱۱/۹۷۵(۴/۹۲۶)
	۸۰	۴۰	۱۲/۰۰۰	۱۱/۹۷۰(۴/۹۸۵)
		۸۰	۱۲/۰۰۰	۱۲/۰۴۷(۳/۹۳۵)
		۱۲۰	۱۲/۰۰۰	۱۲/۰۶۰(۳/۴۵۷)
	۱۲۰	۴۰	۱۲/۰۰۰	۱۲/۰۲۰(۴/۲۲۴)
		۸۰	۱۲/۰۰۰	۱۱/۹۹۰(۳/۲۲۷)
		۱۲۰	۱۲/۰۰۰	۱۱/۹۵۷(۲/۷۵۶)
D_4	۴۰	۴۰	۰/۳۵۱	۰/۳۵۶(۰/۵۸۷)
		۸۰	۰/۳۴۲	۰/۳۴۰(۰/۳۲۷)
		۱۲۰	۰/۳۳۹	۰/۳۴۱(۰/۳۲۹)
	۸۰	۴۰	۰/۳۵۱	۰/۳۵۱(۰/۲۳۴)
		۸۰	۰/۳۴۲	۰/۳۴۲(۰/۲۳۸)
		۱۲۰	۰/۳۳۹	۰/۳۳۹(۰/۲۴۸)
	۱۲۰	۴۰	۰/۳۵۱	۰/۳۵۴(۰/۲۱۷)
		۸۰	۰/۳۴۲	۰/۳۴۰(۰/۱۸۴)
		۱۲۰	۰/۳۳۹	۰/۳۴۱(۰/۱۹۱)

فصل ۳

برآوردگرهای ناپارامتری انقباضی نوع استاین ماتریس کوواریانس در بعد بالا

۱.۳ مقدمه

برآورد ماتریس کوواریانس، نقش مهمی در مسائل بعد بالا، وقتی تعداد متغیرها بیشتر از تعداد مشاهدات باشد، دارد. در مطالعات انجام‌شده، رهیافت‌های انقباضی برای برآورد ماتریس کوواریانس به کار گرفته شده است تا محدودیت‌های ماتریس کوواریانس نمونه (از جمله معکوس‌پذیر نبودن) را حل کند. یک خانواده جدید از برآودگرهای ناپارامتری انقباضی نوع استاین ماتریس کوواریانس که مولفه‌های آن به‌صورت ترکیب خطی محدب از ماتریس کوواریانس نمونه و ماتریس هدف معکوس‌پذیر از قبل تعیین‌شده، پیشنهاد شده است. این ترکیب خطی محدب شامل یک پارامتر انقباضی یا همان ضریب موازنه است که باید به درستی تعیین شود.

در این فصل، یک فرآیند ساده اما تاثیرگذار برای یافتن برآوردگرهای ناپارامتری سازگار برای پارامتر انقباضی بهینه برای یک ماتریس هدف رایج معرفی شده است.

در این فصل به روش انقباضی نوع استاین (استاین، ۱۹۵۶) که توسط لدویت و ولف (۲۰۰۴) برای برآورد ماتریس کوواریانس به کار برده می‌شود، می‌پردازیم. این روش برآوردگرهایی را نتیجه می‌دهد که ویژگی‌های زیر را دارند:

(۱) معکوس‌پذیر هستند.

(۲) به جایگشت‌های مختلف از p متغیر پایا است.

(۳) نسبت به فاصله از توزیع نرمال چندمتغیره استوار است.

(۴) به‌صورت یک فرم ساده بیان می‌شود.

(۵) صرف‌نظر از تعداد متغیرهای p، هزینه محاسباتی کمی دارد.

لدویت و ولف (۲۰۰۴) برآوردگر نوع استاین ماتریس کوواریانس Σ را به‌صورت

$$\boldsymbol{S}^* = (۱ - \lambda)\boldsymbol{S} + \lambda\nu\boldsymbol{I}_p, \tag{۱.۳}$$

پیشنهاد دادند که انتخاب بهینه λ تابع مخاطره $E[\|\boldsymbol{S}^* - \boldsymbol{\Sigma}\|_F^۲]$ را کمینه می‌کند و در قضیه زیر بیان شده است.

قضیه ۱.۱.۳. *مقدار بهینه λ در برآوردگر $\boldsymbol{S}^*$ در رابطه (۱.۳) برابر است با*

$$\begin{aligned}\lambda &= \frac{\frac{۱}{p}E[\mathrm{tr}((\boldsymbol{S}-\boldsymbol{\Sigma})^\top(\boldsymbol{S}-\nu\boldsymbol{I}_p))]}{E[\|\boldsymbol{S}-\nu\boldsymbol{I}_p\|_F^۲]}\\ \nu &= \frac{\mathrm{tr}(\boldsymbol{\Sigma})}{p}\end{aligned}$$

برهان. ابتدا مخاطره برآوردگر $\boldsymbol{S}^*$ را به‌دست می‌آوریم. داریم

$$\begin{aligned}R(\boldsymbol{\Sigma},\boldsymbol{S}^*) &= E[\|\boldsymbol{S}^*-\boldsymbol{\Sigma}\|_F^۲]\\ &= \frac{۱}{p}E[\mathrm{tr}(\boldsymbol{S}^*-\boldsymbol{\Sigma})^\top(\boldsymbol{S}^*-\boldsymbol{\Sigma})]\\ &= \frac{۱}{p}E[\mathrm{tr}((۱-\lambda)\boldsymbol{S}+\lambda\nu\boldsymbol{I}_p-\Sigma)^\top((۱-\lambda)\boldsymbol{S}+\lambda\nu\boldsymbol{I}_p-\Sigma)]\\ &= \frac{۱}{p}E[\mathrm{tr}(S-\lambda(\boldsymbol{S}-\nu\boldsymbol{I}_p)-\boldsymbol{\Sigma})^\top(\boldsymbol{S}-\lambda(\boldsymbol{S}-\nu\boldsymbol{I}_p)-\Sigma)]\\ &= \frac{۱}{p}E[\mathrm{tr}((\boldsymbol{S}-\boldsymbol{\Sigma})-\lambda(\boldsymbol{S}-\nu\boldsymbol{I}_p))^\top((\boldsymbol{S}-\boldsymbol{\Sigma})-\lambda(\boldsymbol{S}-\nu\boldsymbol{I}_p))]\\ &= \frac{۱}{p}E[\mathrm{tr}((\boldsymbol{S}-\boldsymbol{\Sigma})^\top(\boldsymbol{S}-\boldsymbol{\Sigma}))]-۲\lambda\frac{۱}{p}E[\mathrm{tr}((\boldsymbol{S}-\boldsymbol{\Sigma})^\top(\boldsymbol{S}-\nu\boldsymbol{I}_p))]\\ &\quad +\lambda^۲\frac{۱}{p}E[\mathrm{tr}((\boldsymbol{S}-\nu\boldsymbol{I}_p)^\top(\boldsymbol{S}-\nu\boldsymbol{I}_p))] \qquad (۲.۳)\\ &= E[\|\boldsymbol{S}-\boldsymbol{\Sigma}\|_F^۲]-۲\lambda\frac{۱}{p}E[\mathrm{tr}((\boldsymbol{S}-\boldsymbol{\Sigma})^\top(\boldsymbol{S}-\nu\boldsymbol{I}_p))]+\lambda^۲E[\|\boldsymbol{S}-\nu\boldsymbol{I}_p\|_F^۲].\\ \frac{\partial R(\boldsymbol{\Sigma},\boldsymbol{S}^*)}{\partial\lambda} &= -۲\frac{۱}{p}E[\mathrm{tr}((\boldsymbol{S}-\boldsymbol{\Sigma})^\top(\boldsymbol{S}-\nu\boldsymbol{I}_p))]+۲\lambda E[\|\boldsymbol{S}-\nu\boldsymbol{I}_p\|_F^۲]=۰\end{aligned}$$

اکنون با حل معادله نتیجه می‌شود

$$\lambda = \frac{\frac{۱}{p}E[\mathrm{tr}((\boldsymbol{S}-\boldsymbol{\Sigma})^\top(\boldsymbol{S}-\nu\boldsymbol{I}_p))]}{E[\|\boldsymbol{S}-\nu\boldsymbol{I}_p\|_F^۲]}.$$

برای به‌دست آوردن مقدار بهینه ν، مخاطره $R(\boldsymbol{S}^*,\boldsymbol{\Sigma})$ در رابطه (۲.۳) را برحسب ν به‌صورت زیر بازنویسی می‌کنیم

$$\begin{aligned}R(\boldsymbol{S}^*,\boldsymbol{\Sigma}) &= E[\|\boldsymbol{S}-\boldsymbol{\Sigma}\|_F^۲]-۲\lambda\frac{۱}{p}E[\mathrm{tr}(\boldsymbol{S}^\top\boldsymbol{S})-\mathrm{tr}(\boldsymbol{\Sigma}^\top\boldsymbol{S})-\nu\,\mathrm{tr}(\boldsymbol{S}^\top)\\ &\quad +\nu\,\mathrm{tr}(\boldsymbol{\Sigma}^\top)]+\lambda^۲E[\mathrm{tr}(\boldsymbol{S}^\top\boldsymbol{S})-۲\nu\,\mathrm{tr}(\boldsymbol{S})+\nu^۲\,\mathrm{tr}(\boldsymbol{I}_p)]\\ &= E[\|\boldsymbol{S}-\boldsymbol{\Sigma}\|_F^۲]-۲\lambda\frac{۱}{p}E[\boldsymbol{S}^۲-\mathrm{tr}(\boldsymbol{\Sigma}^\top\boldsymbol{\Sigma})-\nu\,\mathrm{tr}(\boldsymbol{\Sigma}^\top)+\nu\,\mathrm{tr}(\boldsymbol{\Sigma}^\top)]\end{aligned}$$

$$
\begin{aligned}
& +\lambda^{۲} E[\operatorname{tr}(\boldsymbol{S}^{۲}) - ۲\nu \operatorname{tr}(\boldsymbol{S}) + p\nu^{۲}] \\
= \; & E[\|\boldsymbol{S} - \boldsymbol{\Sigma}\|_F^{۲}] - ۲\lambda \frac{۱}{p}[E(\operatorname{tr}(\boldsymbol{S}^{۲})) - \operatorname{tr}(\boldsymbol{\Sigma}^{۲})] \\
& +\lambda^{۲} E[(\operatorname{tr}(\boldsymbol{S}^{۲}))] - ۲\nu \operatorname{tr}(\boldsymbol{\Sigma}) + p\nu^{۲} \qquad (۳.۳)
\end{aligned}
$$

بنابراین

$$
\frac{\partial R(\boldsymbol{\Sigma}, \boldsymbol{S}^*)}{\partial \nu} = \lambda^{۲}(-۲ \operatorname{tr}(\boldsymbol{\Sigma}) + ۲p\nu) = ۰
$$

و $\nu = \operatorname{tr}(\boldsymbol{\Sigma})/p$ نتیجه می‌شود و برهان کامل می‌شود. □

با توجه به تعریف $\boldsymbol{S}^*$، پارامتر انقباضی بهینه λ پیشنهاد می‌کند که چه اندازه باید مقادیر ویژه ماتریس کوواریانس نمونه به سمت مقدار ویژه ماتریس هدف $\nu \boldsymbol{I}_p$ منقبض شود. برای مثال $\lambda = ۰$ نتیجه می‌دهد که ماتریس $\nu \boldsymbol{I}_p$ هیچ سهمی در برآوردگر $\boldsymbol{S}^*$ ندارد در حالی که $\lambda = ۱$، بدین معنی است که ماتریس $\boldsymbol{S}$ سهمی در $\boldsymbol{S}^*$ ندارد. مقادیر بین ۰ و ۱ برای λ، به‌طور همزمان برای $\boldsymbol{S}$ و $\nu \boldsymbol{I}_p$ در $\boldsymbol{S}^*$ سهمی قائل می‌شود. با وجود این تفسیر جالب، ماتریس $\boldsymbol{S}^*$ برآوردگر ماتریس کوواریانس نیست، زیرا ν و λ به ماتریس مشاهده نشده Σ بستگی دارند. بدین منظور، لدویت و ولف (۲۰۰۴) برآوردگرهای ناپارامتری n سازگار را برای ν و λ در (۱.۳) پیشنهاد داده‌اند و برآوردگر نتیجه را به عنوان برآوردگر انقباضی ماتریس $\boldsymbol{\Sigma}$ معرفی کردند. اگرچه به نظر می‌رسد ν با $\hat{\nu} = \frac{۱}{p}\operatorname{tr}(S)$ به‌صورت مناسبی برآورد شده است. لدویت و ولف (۲۰۰۴) در قالب یک مطالعه شبیه‌سازی نشان دادند برآوردگر پیشنهادی λی آن‌ها، در حالت بعد بالا و نیز به ازای $\boldsymbol{\Sigma} = \boldsymbol{I}_p$، برآوردگر اریبی است. به عنوان یک اثبات ساده از این موضوع، فرض کنید $\lambda = ۱$. در این حالت انتظار داریم برآوردگر جای‌گذاری $\boldsymbol{S}^*$ تا حد امکان به ماتریس هدف $\nu \boldsymbol{I}_p$ نزدیک باشد که به سادگی اریب بودن قابل مشاهده است. در این راستا فیشر و سان (۲۰۱۱) برآوردگرهای ماتریس کوواریانس انقباضی نوع استاین را برای سایر ماتریس‌های هدف پیشنهاد کردند. اما با توجه به اینکه ساختار فرضی آن‌ها بر اساس مدل نرمال چندمتغیره است، برآوردگر پیشنهادی آن‌ها به‌صورت پارامتری معرفی شده است.

با توجه به مطالب بیان‌شده، برآوردگر پارامتر انقباضی بهینه با برآوردگر سازگار λ در حالت بعد بالا بهبود بخشیده می‌شود. برای ساختن برآوردگر λ، از سه مرحله ساده استفاده می‌شود.

(۱) با در نظر گرفتن فرضیه مدل نرمال چندمتغیره، عبارت‌های صورت و مخرج λ را بسط می‌دهیم.

(۲) اثبات می‌کنیم که این نسبت، λ^*، به‌طور مجانبی معادل λ است.

(۳) هر پارامتر مجهول در λ^* را با برآوردگرهای سازگار و نااریب جای‌گذاری می‌کنیم.

آخرین مرحله، برآوردگر ناپارامتری و سازگار λ را تعیین می‌کند. علاوه بر این، فرضیه نرمال بودن در مقاله فیشر و سان (۲۰۱۱) برای ماتریس هدف $\nu \boldsymbol{I}_p$ در (۱.۳) را نادیده می‌گیریم و نشان می‌دهیم که چگونه پارامترهای انقباضی بهینه در حالت بعد بالا به‌طور سازگار برآورد می‌شوند. به عبارت دیگر، خانواده جدیدی از برآوردگرهای ناپارامتری انقباضی نوع استاین

برای حالت بعد بالا مناسب است که ویژگی‌های جذاب بیان‌شده را تعیین می‌کند و می‌توانند برای ماتریس‌های هدف دلخواه دیگری نیز به کاربرده شوند، پیشنهاد می‌شود. نتایج این فصل عمدتا برگرفته از مرجع تولومیس (۲۰۱۵) است.

۲.۳ چارچوب بعد بالا

فرض کنید $\boldsymbol{x}_1$، ...، $\boldsymbol{x}_n$ یک نمونه تصادفی از بردارهای تصادفی p متغیره از مدل ناپارامتری

$$\boldsymbol{x}_i = \boldsymbol{\Sigma}^{\frac{1}{2}} \boldsymbol{z}_i + \boldsymbol{\mu}, \tag{۴.۳}$$

باشد، که در آن $\boldsymbol{\mu} = E[\boldsymbol{x}_i]$ بردار میانگین p متغیره، $\boldsymbol{\Sigma} = \mathrm{Cov}[\boldsymbol{x}_i] = \boldsymbol{\Sigma}^{\frac{1}{2}} \boldsymbol{\Sigma}^{\frac{1}{2}}$ ماتریس کوواریانس $p \times p$ و z_1، ...، $\boldsymbol{z}_n$ یک نمونه هم‌توزیع و مستقل از بردارهای تصادفی p متغیره هستند.

به جای مفروضات توزیعی، محدودیت‌های لحظه‌ای روی متغیرهای تصادفی در z_i اعمال می‌شود.

به‌طور خاص، فرض کنید z_{ia}، aامین متغیر تصادفی در $\boldsymbol{z}_i$ باشد و فرض کنید $E[z_{ia}] = 0$، $E[z_{ia}^2] = 1$، $E[z_{ia}^4] = 3 + B$ که در آن $-2 \leq B < \infty$ ، و برای اعداد صحیح نامنفی و دلخواه $l_1, \ldots, l_4$ به‌طوری که $\sum_{\nu=1}^{4} l_\nu \leq 4$،

$$E[z_{ia_1}^{l_1} z_{ia_2}^{l_2} z_{ia_3}^{l_3} z_{ia_4}^{l_4}] = E[z_{ia_1}^{l_1}] E[z_{ia_2}^{l_2}] E[z_{ia_3}^{l_3}] E[z_{ia_4}^{l_4}]$$

که در آن a_1، ...، a_4 از یکدیگر متمایز هستند.

کمیت λ را نیز به‌صورت زیر در نظر بگیرید

$$\lambda = \frac{E\left[\|\boldsymbol{S} - \boldsymbol{\Sigma}\|_F^2\right]}{E\left[\|\boldsymbol{S} - \nu \boldsymbol{I}_p\|_F^2\right]}$$

۳.۳ نتایج اصلی

تحت مدل ناپارامتری (۴.۳)، مقدار مورد انتظار در صورت و مخرج λ برابر با

$$\lambda = \frac{E[\|\boldsymbol{S} - \boldsymbol{\Sigma}\|_F^2]}{E[\|\boldsymbol{S} - \nu \boldsymbol{I}_p\|_F^2]} = \frac{\mathrm{tr}(\boldsymbol{\Sigma}^2) + \mathrm{tr}^2(\boldsymbol{\Sigma}) + B\,\mathrm{tr}(\boldsymbol{D}_{\boldsymbol{\Sigma}}^2)}{n\,\mathrm{tr}(\boldsymbol{\Sigma}^2) + \frac{p-n+1}{p}\,\mathrm{tr}^2(\boldsymbol{\Sigma}) + B\,\mathrm{tr}(\boldsymbol{D}_{\boldsymbol{\Sigma}}^2)}, \tag{۵.۳}$$

که در آن برای هر ماتریس معین مثبت $\boldsymbol{A}$،

$$\mathrm{tr}(\boldsymbol{D}_{\boldsymbol{A}}^2) = \mathrm{tr}(\boldsymbol{A} \circ \boldsymbol{A}) \leq \mathrm{tr}(\boldsymbol{A}^2) \leq \mathrm{tr}^2(\boldsymbol{A}),$$

اگر از سهم $B\,\mathrm{tr}(\boldsymbol{D}_{\boldsymbol{\Sigma}}^2)$ برای λ چشم‌پوشی کنیم، خواهیم داشت

$$\lambda^* = \frac{E[\|\boldsymbol{S} - \boldsymbol{\Sigma}\|_F^2]}{E[\|\boldsymbol{S} - \nu \boldsymbol{I}_p\|_F^2]} = \frac{\mathrm{tr}(\boldsymbol{\Sigma}^2) + \mathrm{tr}^2(\boldsymbol{\Sigma})}{n\,\mathrm{tr}(\boldsymbol{\Sigma}^2) + \frac{p-n+1}{p}\,\mathrm{tr}^2(\boldsymbol{\Sigma})},$$

که در آن λ^* میزان انقباض بهینه مدل نرمال چندمتغیره است. یا به‌طور کلی، اگر در مدل (۴.۳) $B=0$، آن‌گاه $\lambda=\lambda^*$. در حالت کلی

$$\begin{aligned}|\lambda-\lambda^*| &= \frac{(n-1)\Big(\operatorname{tr}(\boldsymbol{\Sigma}^2)-\frac{1}{p}\operatorname{tr}^2(\boldsymbol{\Sigma})\Big)}{n\Big(\operatorname{tr}(\boldsymbol{\Sigma}^2)-\frac{1}{p}\operatorname{tr}^2(\boldsymbol{\Sigma})\Big)+\frac{p+1}{p}tr^2(\boldsymbol{\Sigma})} \\ &\quad\times\frac{|B|\operatorname{tr}(\boldsymbol{D}_{\boldsymbol{\Sigma}}^2)}{n\Big(\operatorname{tr}(\boldsymbol{\Sigma}^2)-\frac{1}{p}\operatorname{tr}^2(\boldsymbol{\Sigma})\Big)+\frac{p+1}{p}\operatorname{tr}^2(\boldsymbol{\Sigma})+B\operatorname{tr}(\boldsymbol{D}_{\boldsymbol{\Sigma}}^2)} \\ &\leq \frac{|B|\frac{\operatorname{tr}(\boldsymbol{D}_{\boldsymbol{\Sigma}}^2)}{\operatorname{tr}^2(\boldsymbol{\Sigma})}}{n\Big(\frac{\operatorname{tr}(\boldsymbol{\Sigma}^2)}{\operatorname{tr}^2(\boldsymbol{\Sigma})}-\frac{1}{p}\Big)+\frac{p+1}{p}+B\frac{\operatorname{tr}(\boldsymbol{D}_{\boldsymbol{\Sigma}}^2)}{\operatorname{tr}^2(\boldsymbol{\Sigma})}}\longrightarrow 0.\end{aligned}$$

بنابراین میزان انقباض بهینه بدست آمده تحت حالت نرمال به‌طور مجانبی معادل میزان انقباض بهینه با توجه به مدل (۴.۳) است.

این بدین معناست که برای برآورد λ از برآوردگر ناپارامتری و سازگار λ^* استفاده شود. برای انجام این کار پارامترهای مجهول $\operatorname{tr}(\boldsymbol{\Sigma})$ و $\operatorname{tr}(\boldsymbol{\Sigma}^2)$ در λ^* را با برآوردگرهای ناااریب زیر جایگزین می‌کنیم

$$Y_{1N}=U_{1N}-U_{4N}=\frac{1}{n}\sum_{i=1}^{n}\boldsymbol{x}_i^\top\boldsymbol{x}_i-\frac{1}{P_2^n}\sum_{i=1}^{n}\sum_{\substack{j=1\\j\neq i}}^{n}\boldsymbol{x}_j^\top\boldsymbol{x}_i,$$

و

$$\begin{aligned}Y_{2N} &= U_{2N}-2U_{5N}+U_{6N} \\ &= \frac{1}{P_2^n}\sum_{i=1}^{n}\sum_{\substack{j=1\\j\neq i}}^{n}(\boldsymbol{x}_i^\top\boldsymbol{x}_j)^2-2\frac{1}{P_3^n}\sum_{i=1}^{n}\sum_{\substack{j=1\\j\neq i}}^{n}\sum_{\substack{k=1\\k\neq i\\k\neq j}}^{n}\boldsymbol{x}_i^\top\boldsymbol{x}_j\boldsymbol{x}_i^\top\boldsymbol{x}_k \\ &\quad+\frac{1}{P_4^n}\sum_{i=1}^{n}\sum_{\substack{j=1\\j\neq i}}^{n}\sum_{\substack{k=1\\k\neq i\\k\neq j}}^{n}\sum_{\substack{l=1\\l\neq i\\l\neq j\\l\neq k}}^{n}\boldsymbol{x}_i\boldsymbol{x}_j^\top\boldsymbol{x}_k\boldsymbol{x}_l^\top.\end{aligned}$$

که در آن $P_t^s=\frac{s!}{(s-t)!}$ و U_{1N} و U_{2N} به ترتیب برآوردگرهای ناااریب $\operatorname{tr}(\boldsymbol{\Sigma})$ و $\operatorname{tr}(\boldsymbol{\Sigma}^2)$ وقتی که داده‌ها حول بردار میانگین هستند (برای مثال $\boldsymbol{\mu}=0$)، می‌باشند.

Uـآماره‌های U_4، U_5 و U_6 به ازای $\mu\neq 0$ باعث تضمین ناااریبی Y_{1N} و Y_{2N} می‌شوند.

بنابراین برآوردگر سازگار λ به‌صورت زیر است:

$$\hat{\lambda}=\frac{Y_{2N}+Y_{1n}^2}{nY_{2N}+\frac{p-n+1}{p}Y_{1N}^2}.$$

برآوردگر انقباضی نوع استاین پیشنهادشده برای $\boldsymbol{\Sigma}$ به‌صورت زیر می‌باشد

$$\hat{\boldsymbol{S}}^*=(1-\hat{\lambda})\boldsymbol{S}+\hat{\lambda}\hat{\nu}\boldsymbol{I}_p,$$

که با جایگزینی $\hat{\nu}=\frac{Y_{1N}}{p}$ و $\hat{\lambda}$ در (۱.۳) بدست می‌آید.

۴.۳ شبیه‌سازی

در این بخش از یک شبیه‌سازی برای بررسی کارایی برآورگرهای پیشنهادی نوع استاین ماتریس کوورایانس استفاده کرده‌ایم. فرض کنید مشاهدات pبعدی $\boldsymbol{x}_1$، $\boldsymbol{x}_2$، ...، $\boldsymbol{x}_n$ از مدل $\boldsymbol{x}_i = \boldsymbol{\Sigma}^{\frac{1}{2}} z_i + \boldsymbol{\mu}$ تولید شده باشند که سه توزیع زیر را برای متغیرهای z_1، z_2،... و z_n در نظر گرفته‌ایم.

سناریوی ۱ یک سناریوی نرمال که $z_{ia} \sim \mathcal{N}(0, 1)$.

سناریوی ۲ یک سناریوی گاما با $z_{ia} = (z^*_{ia} - 8)/4$، که در آن $z^*_{ia} \sim \Gamma(4, 0.5)$ و بنابراین $B = 12$.

سناریوی ۳ ترکیبی از سناریوهای ۱ و ۲ که در این سناریو، $p/2$ مولفه اول بردار z_i از سناریوی اول و سایر مولفه‌ها از سناریوی دوم تولید شده‌اند.

سناریوی ۲ و ۳ به منظور بررسی ماهیت ناپارامتری روش پیشنهادی در این فصل در نظر گرفته شده‌اند. برای بررسی ساختار بعد بالا، فرض کنید n از ۱۰ تا ۱۰۰ با گام ۱۰ تغییر کند و همچنین $p = 100, 500, 1000, 1500, 2500$. خانواده پیشنهادی از برآوردگرهای کوورایانس با استفاده از $\hat{\boldsymbol{S}}^*$، که همان برآوردگر ماتریس کوواریانس پیشنهادی به ازای $\boldsymbol{T} = \nu \boldsymbol{I}_p$ است، محاسبه و سپس با برآوردگرهای $\hat{\boldsymbol{S}}^*_{\text{LW}}$ و $\hat{\boldsymbol{S}}^*_{\text{FS}}$، که به ترتیب ماتریس‌های متناظر با روش‌های لدویت و ولف (۲۰۰۴) و فیشر و سان (۲۰۱۱) هستند، مقایسه می‌شوند. مشابه لدویت و ولف (۲۰۰۴)، فرض می‌کنیم $\boldsymbol{\mu} = 0$ و تغییرات لازم را روی $\hat{\boldsymbol{S}}^*$ و $\hat{\boldsymbol{S}}^*_{\text{FS}}$ اعمال می‌کنیم. با توجه به اینکه $\hat{\boldsymbol{S}}^*_{\text{FS}}$ بر اساس فرض نرمال بودن ساخته شده است، کارایی آن فقط در سناریوی ۱ بررسی شده است. لازم به توضیح است که برآوردگر شفر و استریمر (۲۰۰۵) در مطالعات شبیه‌سازی در نظر گرفته نشده است زیرا روش آنها، برآوردگر سازگار λ را در مسایل بعد بالا تضمین نمی‌کند. به ازای $\boldsymbol{\Sigma}$، n و p از قبل تعیین‌شده در هر یک از سناریوهای تعیین‌شده، معیار درصد بهبودی نسبی در میانگین خطا[1] PRIAL شبیه‌سازی‌شده (SPRIAL) ماتریس $\hat{\boldsymbol{S}}^*$ و $\hat{\boldsymbol{S}}^*_{FS}$ را برای برآورد $\boldsymbol{\Sigma}$ محاسبه کرده‌ایم. معیار SPRIAL برای ماتریس $\hat{\boldsymbol{\Sigma}}$ به‌صورت زیر تعریف می‌شود:

$$\text{SPRIAL}(\hat{\boldsymbol{\Sigma}}) = \frac{\sum_{b=1}^{1000} \left\| \hat{\boldsymbol{S}}_{LW,b} - \boldsymbol{\Sigma} \right\|_F^2 - \sum_{b=1}^{1000} \left\| \hat{\boldsymbol{\Sigma}}_b - \boldsymbol{\Sigma} \right\|_F^2}{\sum_{b=1}^{1000} \left\| \hat{\boldsymbol{S}}_{LW,b} - \boldsymbol{\Sigma} \right\|_F^2} \times 100\%$$

که در آن $\hat{\boldsymbol{S}}^*_{\text{LW,b}}$ برآوردگر لدویت و ولف (۲۰۰۴) در تکرار bام ($b = 1, \ldots, 1000$) و $\hat{\boldsymbol{\Sigma}}_b$ برآوردگر متناظر از فرآیند برآورد رقیب در برآورد ماتریس کوواریانس را نشان می‌دهد. بر اساس تعریف، $\text{SPRIAL}(\hat{\boldsymbol{\Sigma}}) = 100\%$ و $\text{SPRIAL}(\hat{\boldsymbol{S}}^*_{LW}) = 0\%$. بنابراین مقادیر مثبت (منفی) از $\text{SPRIAL}(\hat{\boldsymbol{\Sigma}})$ نتیجه می‌دهد که برآوردگر $\hat{\boldsymbol{\Sigma}}$ در مقایسه با برآوردگر $\hat{\boldsymbol{S}}^*_{\text{LW}}$ کارایی بیشتری (کمتری) دارد در

[1] Percentage relative improvement in average loss

حالی‌که مقادیر نزدیک صفر، نتیجه می‌دهد که دو برآوردگر تقریبا کارا هستند. دقت کنید از آن‌جایی که لدویت و ولف (۲۰۰۴) کارایی برآوردگر پیشنهادی‌شان را برای ماتریس کووریانس نمونه نشان دادند، برآوردگر $\hat{\Sigma}^*_{\text{LW}}$ را به عنوان برآوردگر پایه‌ای در نظر گرفتیم.

برای بررسی حالتی که ماتریس هدف با ماتریس کوواریانس واقعی برابر باشد، $\Sigma = \boldsymbol{I}_p$ را در نظر می‌گیریم. در این حالت، برآورد دقیقی از λ حیاتی است زیرا صرف‌نظر از مقادیر n و p، مقدار λ برابر ۱ است.

جدول ۱.۳ نتایج شبیه‌سازی را برای سناریوی ۱ برای میانگین پارامتر انقباضی بهینه بر اساس روش پیشنهادی ($\hat{\lambda}$)، رهیافت لدویت و ولف (۲۰۰۴) ($\hat{\lambda}_{\text{LW}}$) و رهیافت فیشر و سان (۲۰۱۱) ($\hat{\lambda}_{\text{FS}}$) برای $n = ۱۰, ۵۰, ۱۰۰$ و $p = ۱۰۰, ۱۰۰۰, ۲۵۰۰$ نشان می‌دهد. برآوردگرهای $\hat{\lambda}$، $\hat{\lambda}_{\text{LW}}$ و $\hat{\lambda}_{\text{FS}}$ در بازه $(۰, ۱)$ محدود شده‌اند. در مقایسه با $\hat{\lambda}_{\text{LW}}$، برآوردگر $\hat{\lambda}$ در برآورد λ دقت بیشتری دارند. این برآوردگر ذاتا اریبی کمتری دارد و به ازای همه مقادیر n و p، خطای استاندارد کوچکتری دارد. اگرچه اریبی برآوردگر $\hat{\lambda}_{\text{LW}}$، با افزایش n کاهش می‌یابد، به نظر می‌رسد که برآوردگر $\hat{\lambda}_{\text{LW}}$، یک برآوردگر اریب حتی در حالت $n = ۱۰۰$ است. همان‌طور که انتظار می‌رفت، تفاوت معنی‌داری بین برآوردگرهای $\hat{\lambda}$ و $\hat{\lambda}_{\text{LW}}$ در سناریوی ۱ وجود ندارد.

جدول ۲.۳، مقادیر $SPRIAL(\hat{\boldsymbol{S}}^*)$ را در سناریوی ۲ و جدول ۳.۳، این مقادیر را برای سناریوی ۳ نشان می‌دهد.

جدول ۱.۳: برآورد پارامتر انقباضی $\lambda = ۱$ برای سناریوی ۱ به ازای $\Sigma = \boldsymbol{I}_p$

n	p	$\hat{\lambda}$	$\hat{\lambda}_{\text{LW}}$	$\hat{\lambda}_{\text{FS}}$
۱۰	۱۰۰	۰/۹۹۱۵(۰/۰۱۲)	۰/۸۹۹۵(۰/۰۱۸)	۰/۹۹۱۵(۰/۰۱۰)
	۱۰۰۰	۰/۹۹۸۸(۰/۰۰۱)	۰/۸۹۹۹(۰/۰۰۱)	۰/۹۹۹۱(۰/۰۰۱)
	۲۵۰۰	۰/۹۹۹۵(۰/۰۰۰)	۰/۹۰۰۰(۰/۰۰۰)	۰/۹۹۹۶(۰/۰۰۰)
۵۰	۱۰۰	۰/۹۹۱۸(۰/۰۱۰)	۰/۹۸۷۸(۰/۰۱۴)	۰/۹۹۲۰(۰/۰۰۹)
	۱۰۰۰	۰/۹۹۹۰(۰/۰۰۱)	۰/۹۷۹۹(۰/۰۰۲)	۰/۹۹۹۳(۰/۰۰۱)
	۲۵۰۰	۰/۹۹۹۵(۰/۰۰۰)	۰/۹۷۹۹(۰/۰۰۰)	۰/۹۹۹۶(۰/۰۰۵)
۱۰۰	۱۰۰	۰/۹۹۱۹(۰/۰۱۱)	۰/۹۸۵۵(۰/۰۱۴)	۰/۹۹۲۰(۰/۰۱۱)
	۱۰۰۰	۰/۹۹۹۱(۰/۰۰۱)	۰/۹۸۹۸(۰/۰۰۲)	۰/۹۹۹۲(۰/۰۰۱)
	۲۵۰۰	۰/۹۹۹۸(۰/۰۰۰)	۰/۹۸۹۹(۰/۰۰۰)	۰/۹۹۹۶(۰/۰۰۰)

جدول ۲.۳: مقادیر $\mathrm{SPRIAL}(\hat{\boldsymbol{S}}^*)$ برای سناریوی ۲ به ازای $\boldsymbol{\Sigma} = \boldsymbol{I}_p$

	p				
n	۱۰۰	۵۰۰	۱۰۰۰	۱۵۰۰	۲۵۰۰
۱۰	۶۱/۱۵	۸۹/۲۴	۹۲/۰۴	۹۴/۶۰	۹۸/۶۵
۲۰	۴۲/۹۹	۹۰/۰۲	۹۳/۴۲	۹۴/۶۳	۹۸/۷۵
۳۰	۲۹/۸۰	۸۸/۸۷	۹۳/۹۴	۹۵/۰۶	۹۸/۹۰
۴۰	۲۲/۰۲	۸۳/۲۵	۹۴/۹۵	۹۵/۷۵	۹۸/۹۲
۵۰	۱۸/۹۲	۸۲/۶۳	۹۳/۲۱	۹۶/۰۲	۹۸/۹۷
۶۰	۱۷/۲۴	۷۹/۸۲	۹۲/۵۴	۹۵/۴۸	۹۸/۲۴
۷۰	۱۶/۷۶	۷۸/۲۴	۹۲/۴۳	۹۴/۸۷	۹۸/۵۴
۸۰	۱۵/۴۸	۷۳/۵۹	۹۰/۵۷	۹۳/۰۶	۹۸/۲۳
۹۰	۱۴/۲۶	۶۸/۳۶	۸۵/۶۹	۹۲/۸۵	۹۷/۸۶
۱۰۰	۱۳/۶۹	۶۱/۰۳	۸۴/۰۶	۹۱/۶۹	۹۸/۲۶

جدول ۳.۳: مقادیر $\mathrm{SPRIAL}(\hat{\boldsymbol{S}}^*)$ برای سناریوی ۳ به ازای $\boldsymbol{\Sigma}$ ماتریسی است که (a,b) امین عضو آن به‌صورت $\Sigma_{ab} = ۰/۱I(|a-b|)$ است

	p				
n	۱۰۰	۵۰۰	۱۰۰۰	۱۵۰۰	۲۵۰۰
۱۰	۷۸/۵۶	۹۸/۳۶	۹۸/۵۶	۹۸/۹۷	۹۹/۵۴
۲۰	۳۸/۷۵	۷۸/۶۹	۸۲/۳۶	۸۵/۶۹	۹۰/۶۳
۳۰	۱۸/۳۶	۴۲/۶۹	۶۳/۲۶	۷۵/۶۹	۸۳/۲۶
۴۰	۱۵/۲۴	۲۲/۳۹	۴۲/۲۸	۵۹/۳۰	۸۳/۲۷
۵۰	۱۴/۶۷	۱۹/۲۴	۳۵/۶۹	۴۲/۱۶	۴۷/۶۹
۶۰	۱۴/۳۱	۱۵/۳۶	۲۶/۷۸	۳۶/۴۷	۳۹/۸۴
۷۰	۱۴/۰۳	۱۳/۶۸	۱۵/۷۴	۲۰/۰۶	۲۴/۰۳
۸۰	۱۳/۷۵	۱۳/۰۶	۱۴/۶۹	۱۵/۰۳	۲۰/۱۶
۹۰	۱۳/۴۷	۱۲/۹۸	۱۴/۰۵	۱۴/۳۶	۱۹/۲۴
۱۰۰	۱۳/۰۲	۱۲/۲۳	۱۳/۵۷	۱۴/۱۰	۱۸/۳۲

فصل ۴

مقایسه برآوردگر انقباضی خطی ماتریس کوواریانس در توزیع نرمال و غیر نرمال

۱.۴ مقدمه

هنگامی‌که تعداد متغیرها یعنی p بزرگتر از اندازه نمونه n است، ماتریس کوواریانس نمونه معکوس‌پذیر نیست. وقتی p بزرگ و نزدیک به n باشد، معکوس ماتریس کوواریانس نمونه حتی اگر $n > p$، ممکن است سازگار نباشد. بنابراین به برآوردگری برای ماتریس کوواریانس نیاز داریم که هم معکوس‌پذیر و هم سازگار باشد.

تاکنون رهیافت‌های بسیاری برای رسیدن به این هدف بررسی شده‌اند. یک رهیافت ساده روش انقباضی خطی است که ترکیب محدب از ماتریس کوواریانس نمونه و ماتریس هدف معین مثبت است . این رهیافت توسط دنیلز و کاس (۲۰۰۱)، لدویت و ولف (۲۰۰۳ و ۲۰۰۴)، شافر و استریمر (۲۰۰۵)، سریواستارا و کوبوکاوا[۱] (۲۰۰۷)، کنو [۲] (۲۰۰۹)، چن وهمکاران (۲۰۱۰)، فیشر و سان (۲۰۱۱)، بای و شی (۲۰۱۱) و تولومیس (۲۰۱۵) مورد مطالعه قرار گرفته است. لدویت و ولف (۲۰۱۲ و۲۰۱۵) روش‌های انقباضی غیرخطی را بسط داده‌اند. از رهیافت‌های دیگر می‌توان به تکنیک‌های منظم‌سازی[۳] و آستانه‌سازی[۴] مانند بیکل و لوینا (۲۰۰۸)، راتمن[۵]

[۱] Serivastara and Kubokawa
[۲] Konno
[۳] Regularization
[۴] Thresholding
[۵] Rothm

و همکاران (۲۰۰۹)، کای و ژائو[۶] (۲۰۱۰)، کای و لیو[۷] (۲۰۱۱) و فن[۸] و همکاران (۲۰۱۱ و ۲۰۱۳) اشاره کرد.

در این فصل، برآوردگرهای انقباضی خطی ساده را بررسی می‌کنیم و یک روش منطقی برای برآورد وزن‌های بهینه در حالت‌های ناپارامتری ارائه می‌شود. لازم به ذکر است که برآوردگرهای انقباضی خطی، حتی در حالت‌های بعد پایین مینیماکس یا مجاز نیستند. مطالب این فصل برگرفته از مرجع ایکدا و همکاران[۹] (۲۰۱۶) است.

۲.۴ برآوردگر انقباضی خطی

فرض کنید $\boldsymbol{S}$ یک ماتریس کوواریانس نمونه با $E(\boldsymbol{S}) = \boldsymbol{\Sigma}$ باشد و همچنین $\hat{\boldsymbol{\Lambda}}$ ماتریس هدف معین مثبت بر پایه $\boldsymbol{S}$ باشد. برآوردگرهای انقباضی خطی، ترکیب‌های محدب $\boldsymbol{S}$ و $\hat{\boldsymbol{\Lambda}}$ هستند که به‌صورت

$$\begin{aligned}\hat{\boldsymbol{\Sigma}}_w(\hat{\boldsymbol{\Lambda}}) &= (1-w)\boldsymbol{S} + w\hat{\boldsymbol{\Lambda}} \\ &= \boldsymbol{S} - w(\boldsymbol{S} - \hat{\boldsymbol{\Lambda}})\end{aligned}$$

تعریف می‌شوند که در آن w یک ثابت دلخواه است که $0 \leq w \leq 1$. این نوع از برآوردگرها را، برآوردگرهای انقباضی تکی[۱۰] می‌نامند. می‌خواهیم وزن w را به‌صورت ناپارامتری بر اساس $\boldsymbol{S}$ با یک وزن بهینه برآورد کنیم. برآوردخطی ناپارامتری توسط هارتیگان[۱۱] (۱۹۶۹)، رابینز[۱۲] (۱۹۸۳) و گاش و لاهیری[۱۳] (۱۹۸۷) مطالعه شده است. مسئله برآورد $\boldsymbol{\Sigma}$ با برآوردگر $\hat{\Sigma}$ را نسبت به تابع زیان $L(\boldsymbol{\Sigma}, \hat{\boldsymbol{\Sigma}}) = \mathrm{tr}(\hat{\boldsymbol{\Sigma}} - \boldsymbol{\Sigma})^2$ و در نتیجه تابع مخاطره $R(\boldsymbol{\Sigma}, \hat{\boldsymbol{\Sigma}}) = E[\mathrm{tr}(\hat{\boldsymbol{\Sigma}} - \boldsymbol{\Sigma})^2]$ در نظر بگیرید. تابع مخاطره بر اساس $\hat{\boldsymbol{\Sigma}}_w(\boldsymbol{\Lambda})$ به‌صورت زیر خواهد بود

$$\begin{aligned}R\big(\boldsymbol{\Sigma}, \hat{\boldsymbol{\Sigma}}_w\big) &= E\big[\,\mathrm{tr}(\hat{\boldsymbol{\Sigma}}_w(\hat{\boldsymbol{\Lambda}}) - \boldsymbol{\Sigma})^2\big] \\ &= E\big[\,\mathrm{tr}(\boldsymbol{S} - w(\boldsymbol{S} - \hat{\boldsymbol{\Lambda}}) - \boldsymbol{\Sigma})^2\big] \\ &= E\big[\,\mathrm{tr}\left((\boldsymbol{S} - \boldsymbol{\Sigma}) - w(\boldsymbol{S} - \hat{\boldsymbol{\Lambda}})\right)^2\big] \\ &= E\big[\,\mathrm{tr}\left((\boldsymbol{S} - \boldsymbol{\Sigma})^2 - 2w(\boldsymbol{S} - \boldsymbol{\Sigma})(\boldsymbol{S} - \hat{\boldsymbol{\Lambda}}) + w^2(\boldsymbol{S} - \hat{\boldsymbol{\Lambda}})^2\right)\big] \\ &= E\big[\,\mathrm{tr}(\boldsymbol{S} - \boldsymbol{\Sigma})^2\big] - 2wE\big[\,\mathrm{tr}(\boldsymbol{S} - \boldsymbol{\Sigma})(\boldsymbol{S} - \hat{\boldsymbol{\Lambda}})\big] + w^2 E\big[\,\mathrm{tr}(\boldsymbol{S} - \hat{\boldsymbol{\Lambda}})^2\big].\end{aligned}$$

---

[۶] Cai and Zhou
[۷] Cai and Liu
[۸] Fan
[۹] Ikeda et al.
[۱۰] Single Shrinkage Estimate
[۱۱] Hartigan
[۱۲] Robins
[۱۳] Gosh and Lahiri

برای به‌دست آوردن وزن بهینه کافیست مخاطره کمینه شود. برای این منظور با مشتق‌گیری نسبت به w داریم

$$\frac{d}{dw}R(\boldsymbol{\Sigma},\hat{\boldsymbol{\Sigma}}_w(\boldsymbol{\Lambda}))=۲wE\big[\operatorname{tr}(\boldsymbol{S}-\hat{\boldsymbol{\Lambda}})^۲\big]-۲E\big[\operatorname{tr}(\boldsymbol{S}-\boldsymbol{\Sigma})(\boldsymbol{S}-\hat{\boldsymbol{\Lambda}})\big]=۰$$

و در نتیجه

$$w^*=\frac{E\big[\operatorname{tr}(\boldsymbol{S}-\boldsymbol{\Sigma})(\boldsymbol{S}-\hat{\boldsymbol{\Lambda}})\big]}{E\big[\operatorname{tr}(\boldsymbol{S}-\hat{\boldsymbol{\Lambda}})^۲\big]}. \tag{۱.۴}$$

در این فصل، یک روش نااریب برای برآورد صورت و مخرج w^* در حالت‌های غیرنرمال پیشنهاد می‌شود.

مثال‌های معمولی از ماتریس هدف $\hat{\boldsymbol{\Lambda}}$، ماتریس کروی $\hat{a}_۱\boldsymbol{I}_p$ به ازای $\hat{a}_۱=\operatorname{tr}(\boldsymbol{S})/p$ و ماتریس قطری $D_s=\operatorname{diag}(s_{۱۱},\ldots,s_{pp})$ برای $\boldsymbol{S}=(s_{ij})$ هستند. برای این ماتریس‌های هدف، برآوردگرهای ناپارامتری به‌دست آمده‌اند. یک برآوردگر انقباضی خطی دیگر، برآوردگر انقباضی دوگانه [۱۴] خواهد بود. وقتی چندین ماتریس هدف در نظر گرفته شوند، منطقی است که $\boldsymbol{S}$ را به سمت آن ماتریس‌های هدف منقبض کنیم. برای مثال، فرض کنید این ماتریس‌های هدف، ماتریس کروی $\hat{a}_۱\boldsymbol{I}_p$ و ماتریس قطری $\boldsymbol{D}_s$ به عنوان دو نامزد از ماتریس‌های هدف باشند. می‌توان ماتریس هدف را به‌صورت ترکیب محدب زیر در نظر گرفت

$$\boldsymbol{\Sigma}_D=(۱-\alpha)\hat{a}_۱\boldsymbol{I}_p+\alpha\boldsymbol{D}_S.$$

در این‌صورت رهیافت انقباضی خطی برآوردگر زیر پیشنهاد می‌شود

$$\begin{aligned}\hat{\boldsymbol{\Sigma}}(\hat{a}_۱\boldsymbol{I}_p,\boldsymbol{D}_S) &= (۱-\beta)\boldsymbol{S}+\beta\boldsymbol{\Sigma}_D\\ &= (۱-\beta)\boldsymbol{S}+\beta\Big\{(۱-\alpha)\hat{a}_۱\boldsymbol{I}_p+\alpha\boldsymbol{D}_S\Big\}\end{aligned}$$

که ثابت‌های α و β در بازه $[۰,۱]$ قرار دارند. اگر $w_۱=(۱-\alpha)\beta$ و $w_۲=\alpha\beta$، آن‌گاه برآوردگر $\hat{\boldsymbol{\Sigma}}(\hat{a}_۱\boldsymbol{I}_p,\boldsymbol{D}_S)$ به‌صورت زیر بیان می‌شود

$$\begin{aligned}\hat{\boldsymbol{\Sigma}}_{w_۱,w_۲}(\hat{a}_۱\boldsymbol{I}_p,\boldsymbol{D}_s) &= (۱-w_۱-w_۲)\boldsymbol{S}+\boldsymbol{w}_۱\hat{a}_۱\boldsymbol{I}_p+w_۲\boldsymbol{D}_S\\ &= \boldsymbol{S}-w_۱(\boldsymbol{S}-\hat{a}_۱\boldsymbol{I}_p)-w_۲(\boldsymbol{S}-\boldsymbol{D}_s)\end{aligned}$$

که در این‌صورت این برآوردگر را برآوردگر انقباضی دوگانه می‌نامند.

۳.۴ برآورد وزن بهینه

بردارهای تصادفی p متغیره مستقل و هم‌توزیع $\boldsymbol{x}_n,\ldots,\boldsymbol{x}_۱$ با بردار میانگین $\boldsymbol{\mu}$ و ماتریس کوواریانس $\boldsymbol{\Sigma}=\boldsymbol{\Sigma}^{۱/۲}\boldsymbol{\Sigma}^{۱/۲}$ را در نظر بگیرید که $\boldsymbol{\Sigma}^{۱/۲}$ تجزیه چولسکی[۱۵] با عناصر قطری مثبت

[۱۴] Double Shrinkage

[۱۵] Cholesky Decomposition

است. فرض کنید بردارهای مشاهدات $\boldsymbol{x}_i$ به‌صورت زیر تولید شده‌اند

$$\boldsymbol{x}_i = \boldsymbol{\mu} + \boldsymbol{\Sigma}^{\frac{1}{2}}\boldsymbol{u}_i \qquad i = 1, \ldots, n,$$

به‌طوری که

$$E(\boldsymbol{u}_i) = 0, \qquad \mathrm{Cov}(\boldsymbol{u}_i) = \boldsymbol{I}_p,$$

و به ازای γ_1، ... و γ_k که $0 \leq \sum_{k=1}^{p} \gamma_k \leq 4$

$$E\left[\prod_{k=1}^{p} u_{ki}^{\gamma_k}\right] = \prod_{k=1}^{p} E(u_{ki}^{\gamma_k}) \qquad i = 1, \ldots, n,$$

که در آن u_{ki}، k امین مولفه از بردار $\boldsymbol{u}_i = (u_{1i}, \ldots, u_{pi})$ است. این فرضیه برای وجود تابع مخاطره ضروری است. سومین وچهارمین گشتاور u_{ki} به‌صورت $E(u_{ki}^3) = \kappa_3$ و $E(u_{ki}^4) = \kappa_4 + 3$ نوشته می‌شوند. در حالت توزیع نرمال، $\kappa_3 = \kappa_4 = 0$. فرض کنید $\boldsymbol{\Sigma} = (\sigma_{ij})$. آن‌گاه از نمادهای زیر استفاده می‌شود:

$$\begin{aligned} a_1 &= \frac{1}{p}\mathrm{tr}(\boldsymbol{\Sigma}) \\ a_2 &= \frac{1}{p}\mathrm{tr}(\boldsymbol{\Sigma}^2) \\ a_{20} &= \frac{1}{p}\sum_{i=1}^{p}\sigma_{ii}^2. \end{aligned}$$

برآوردگر ناریب $\boldsymbol{\Sigma}$، ماتریس کواریانس $\boldsymbol{S}$ است اما در حالت $p > n$ و حتی زمانی که p نزدیک به n است، خوش‌ـوضع [16] نیست. بنابراین منطقی است که ترکیب‌های محدب $\boldsymbol{S}$ و ماتریس معین مثبت بر پایه $\boldsymbol{S}$ در نظر بگیریم. برای مثال، محدودیت کروی بودن $\boldsymbol{\Sigma} = \sigma^2\boldsymbol{I}_p$ را در نظر بگیرید. آن‌گاه σ^2 با $\hat{a}_1 = \mathrm{tr}(\boldsymbol{S})/p$ برآورد می‌شود و ترکیب‌های محدب $(1-w)\boldsymbol{S} + w\hat{a}_1\boldsymbol{I}_p$ به‌ازای $0 \leq w \leq 1$ پیشنهاد شده است. این برآوردگرهای انقباضی خطی، ماتریس $\boldsymbol{S}$ را به سمت ماتریس هدف معین مثبت $\hat{a}_1\boldsymbol{I}_p$ منقبض می‌کنند. ماتریس هدف دیگر، ماتریس قطری $\boldsymbol{D}_s = \mathrm{diag}(s_{11}, \ldots, s_{pp})$ برای $\boldsymbol{S} = (s_{ij})$ است. چنین برآوردگرهای انقباضی خطی توسط لدویت و ولف (۲۰۰۴)، فیشر وسان (۲۰۱۱)، تولومیس (۲۰۱۵) و سایرین مورد بررسی قرار گرفته‌اند.

در حالت کلی، ممکن است محدودیت $\boldsymbol{\Sigma} = \boldsymbol{\Lambda}(\theta)$ در نظر گرفته شود. دقت کنید که برای محدودیت کروی بودن، $\boldsymbol{\Lambda}(\theta) = \sigma^2\boldsymbol{I}_p$ و برای محدودیت قطری بودن، $\boldsymbol{\Lambda}(\theta) = \mathrm{diag}(\sigma_{11}, \ldots, \sigma_{pp})$ است. فرض کنید $\hat{\theta}$ برآوردگر θ بر حسب $\boldsymbol{S}$ تحت محدودیت $\boldsymbol{\Sigma} = \boldsymbol{\Lambda}(\theta)$ باشد. در این صورت، برآوردگرهای انقباضی خطی کلی به‌صورت

$$\hat{\boldsymbol{\Sigma}}_w(\Lambda) = (1-w)\boldsymbol{S} + w\hat{\boldsymbol{\Lambda}}$$

[16] Well-conditioned

است که $\hat{\boldsymbol{\Lambda}} = \boldsymbol{\Lambda}(\hat{\theta})$ و $۰ \leq w \leq ۱$.

وزن بهینه رابطه (۱.۴) را می‌توان به‌صورت زیر نوشت

$$w^* = \frac{E\big[\operatorname{tr}(\boldsymbol{S}-\hat{\boldsymbol{\Lambda}})(\boldsymbol{S}-\hat{\boldsymbol{\Lambda}})\big]}{E\big[\operatorname{tr}(\boldsymbol{S}-\hat{\boldsymbol{\Lambda}})^۲\big]} = \frac{E\big[\operatorname{tr}\left(\boldsymbol{S}(\boldsymbol{S}-\hat{\boldsymbol{\Lambda}})\right)\big] - E\big[\operatorname{tr}\left(\boldsymbol{\Sigma}(\boldsymbol{S}-\hat{\boldsymbol{\Lambda}})\right)\big]}{E\big[\operatorname{tr}(\boldsymbol{S}-\hat{\boldsymbol{\Lambda}})^۲\big]}.$$

همچنین

$$\begin{aligned} E\big[\operatorname{tr}\left(\boldsymbol{\Sigma}(\boldsymbol{S}-\hat{\boldsymbol{\Lambda}})\right)\big] &= \operatorname{tr}\left(E\big(\boldsymbol{\Sigma}(\boldsymbol{S}-\hat{\boldsymbol{\Lambda}})\big)\right) \\ &= \operatorname{tr}\left(\boldsymbol{\Sigma}E(\boldsymbol{S}) - E(\boldsymbol{\Sigma}\hat{\boldsymbol{\Lambda}})\right) \\ &= \operatorname{tr}(\boldsymbol{\Sigma}^۲) - E\Big[\operatorname{tr}(\boldsymbol{\Sigma}\hat{\boldsymbol{\Lambda}})\Big]. \end{aligned}$$

دقت کنید که $\operatorname{tr}[\boldsymbol{S}(\boldsymbol{S}-\hat{\boldsymbol{\Lambda}})]$ و $\operatorname{tr}[(\boldsymbol{S}-\hat{\boldsymbol{\Lambda}})^۲]$ برآوردگرهای ناریب دقیق $E\Big[\operatorname{tr}(\boldsymbol{S}(\boldsymbol{S}-\hat{\boldsymbol{\Lambda}}))\Big]$ و $E\Big[\operatorname{tr}(\boldsymbol{S}-\hat{\boldsymbol{\Lambda}})^۲\Big]$ هستند. فرض کنید

$$G(\boldsymbol{\Lambda}) = E\operatorname{tr}\Big[\boldsymbol{\Sigma}(\boldsymbol{S}-\hat{\boldsymbol{\Lambda}})\Big] = \operatorname{tr}(\boldsymbol{\Sigma}^۲) - E\Big[\operatorname{tr}(\boldsymbol{\Sigma}\hat{\boldsymbol{\Lambda}})\Big].$$

یک برآوردگر به‌طور مجانبی ناریب از $G(\boldsymbol{\Lambda})$ را با $\hat{G}$ نشان می‌دهیم، در این‌صورت

$$E\Big[\hat{G}\Big] \approx G(\boldsymbol{\Lambda}) = \operatorname{tr}\Big[\boldsymbol{\Sigma}E(\boldsymbol{S}-\hat{\boldsymbol{\Lambda}})\Big].$$

لذا برآوردگر ناپارامتری w^* به‌صورت

$$\hat{w^*} = \frac{\operatorname{tr}\big[\boldsymbol{S}(\boldsymbol{S}-\hat{\boldsymbol{\Lambda}})\big] - \hat{G}}{\operatorname{tr}[(\boldsymbol{S}-\hat{\boldsymbol{\Lambda}})^۲]}$$

خواهد بود. از آن جایی که w در بازه $(۰,۱)$ محدود شده است، از برآوردگر

$$\hat{w} = \max\{۰, \min\{۱, \hat{w}^*\}\}$$

استفاده می‌کنیم.

۴.۴ ماتریس هدف کروی

حالتی را در نظر بگیرید که $\boldsymbol{\Lambda}_۱ = \sigma^۲\boldsymbol{I}_p$ و به ازای $\hat{a}_۱ = \operatorname{tr}(\boldsymbol{S})/p$،

$$\hat{\boldsymbol{\Lambda}}_۱ = \hat{a}_۱\boldsymbol{I}_p.$$

در این حالت، برآوردگر انقباضی خطی به‌صورت زیر نوشته می‌شود

$$\hat{\boldsymbol{\Sigma}}_w(\hat{a}_۱\boldsymbol{I}_p) = (۱-w)\boldsymbol{S} + w\hat{a}_۱\boldsymbol{I}_p.$$

می‌دانیم

$$E[\hat{a}_۱] = E\left[\frac{۱}{p}\operatorname{tr}(\boldsymbol{S})\right] = \frac{۱}{p}\operatorname{tr}(E(\boldsymbol{S})) = \frac{۱}{p}\operatorname{tr}(\boldsymbol{\Sigma}) = a_۱$$

از آنجایی که $E\hat{a}_۱ = a_۱$ داریم

$$\begin{aligned} G(\boldsymbol{\Lambda}_۱) &= E\big[\operatorname{tr}(\boldsymbol{\Sigma}(\boldsymbol{S} - \hat{a}_۱\boldsymbol{I}_p))\big] \\ &= \operatorname{tr}\left(\boldsymbol{\Sigma}E(\boldsymbol{S} - \hat{a}_۱\boldsymbol{I}_p)\right) \\ &= \operatorname{tr}\left(\boldsymbol{\Sigma}(\boldsymbol{\Sigma} - a_۱\boldsymbol{I}_p)\right) \\ &= \operatorname{tr}(\boldsymbol{\Sigma}^۲) - a_۱\operatorname{tr}(\boldsymbol{\Sigma}) \\ &= p\left\{\frac{۱}{p}\operatorname{tr}(\boldsymbol{\Sigma}^۲) - a_۱\frac{۱}{p}\operatorname{tr}(\boldsymbol{\Sigma})\right\} \\ &= p\left\{a_۲ - a_۱^۲\right\}. \end{aligned}$$

بنابراین باید بدون فرض نرمال بودن، $a_۲$ و $a_۱^۲$ را به‌طور نااریب برآورد کنیم.
یک برآوردگر برای $a_۲$ را در حالت توزیع نرمال به‌صورت زیر در نظر بگیرید

$$\hat{a}_۲ = \frac{(n-۱)^۲}{(n-۲)(n+۱)}\left[\frac{۱}{p}\operatorname{tr}(\boldsymbol{S}^۲) - \frac{p}{n-۱}\hat{a}_۱^۲\right]. \qquad (۲.۴)$$

برای اثبات اینکه برآوردگر $\hat{a}_۲$ در رابطه (۲.۴) برای $a_۲$ نااریب است، ابتدا لم زیر را در نظر بگیرید.

لم ۱.۴.۴. *فرض کنید ν متغیر تصادفی از توزیع $\chi^۲$ با n درجه آزادی باشد. آن‌گاه*

$$\begin{aligned} E(\nu^۲) &= n(n+۲),\ldots,n(n+۲r-۲) \qquad r = ۱,۲,\ldots \\ Var(\nu) &= ۲n \\ Var(\nu^۲) &= ۸n(n+۲)(n+۳) \\ E(\nu-n)^۳ &= ۸n \\ E(\nu-n)^۴ &= ۱۲n(n+۴) \\ E(\nu^۲ - n(n+۲))^۴ &= ۳n(n+۲)(۲۷۲n^۴ + O(n^۳)) \end{aligned}$$

قضیه ۲.۴.۴. *در حالت نرمال، برآوردگر $\hat{a}_۲$ در رابطه (۲.۴) برای $a_۲$ نااریب است،*

برهان. عبارت‌های $\operatorname{tr}(\boldsymbol{S})$، $\operatorname{tr}(\boldsymbol{S}^۲)$ و $(\operatorname{tr}(\boldsymbol{S}))^۲$ را برحسب متغیرهای تصادفی کی‌دو می‌نویسیم. فرض کنید

$$\boldsymbol{V} = (n-۱)\boldsymbol{S} = \boldsymbol{Y}\boldsymbol{Y}^\top \sim \mathcal{W}_p(\boldsymbol{\Sigma}, n)$$

که در آن $\boldsymbol{Y} = (\boldsymbol{y}_۱, \boldsymbol{y}_۲, \ldots, \boldsymbol{y}_n)^\top$ و $\boldsymbol{y}_i \overset{iid}{\sim} \mathcal{N}_p(\circ, \boldsymbol{\Sigma})$. همچنین $\boldsymbol{\Gamma}$ یک ماتریس متعامد $p \times n$ باشد به‌طوری که $\boldsymbol{\Gamma}\boldsymbol{\Sigma}\boldsymbol{\Gamma}^\top = \boldsymbol{\Lambda}$ که در آن $\boldsymbol{\Lambda} = \operatorname{diag}(\lambda_۱, \ldots, \lambda_p)$ و λ_iها مقادیر ویژه ماتریس $\boldsymbol{\Sigma}$ هستند. فرض کنید $\boldsymbol{U} = (\boldsymbol{u}_۱, \ldots, \boldsymbol{u}_n)^\top$ به‌طوری که $\boldsymbol{u}_i \overset{iid}{\sim} \mathcal{N}_p(\circ, \boldsymbol{I}_p)$ و $\boldsymbol{Y} = \boldsymbol{\Sigma}^{\frac{۱}{۲}}\boldsymbol{U}$ و $\boldsymbol{\Sigma}^{\frac{۱}{۲}}\boldsymbol{\Sigma}^{\frac{۱}{۲}} = \boldsymbol{\Sigma}$. بنابراین

$$\operatorname{tr}(\boldsymbol{S}) = \operatorname{tr}\left(\frac{۱}{n-۱}\boldsymbol{Y}\boldsymbol{Y}^\top\right)$$

$$
\begin{aligned}
&= \operatorname{tr}\left(\frac{۱}{n-۱}\boldsymbol{Y}^\top\boldsymbol{Y}\right)\\
&= \operatorname{tr}\left(\frac{۱}{n-۱}\boldsymbol{U}^\top\boldsymbol{\Sigma}^{\frac{۱}{۲}}\boldsymbol{\Sigma}^{\frac{۱}{۲}}\boldsymbol{U}\right)\\
&= \operatorname{tr}\left(\frac{۱}{n-۱}\boldsymbol{U}^\top\boldsymbol{\Sigma}\boldsymbol{U}\right)\\
&= \operatorname{tr}\left(\frac{۱}{n-۱}\boldsymbol{U}(\boldsymbol{\Gamma}^\top\boldsymbol{\Lambda}\boldsymbol{\Gamma})\boldsymbol{U}\right)\\
&= \operatorname{tr}\left(\frac{۱}{n-۱}(\boldsymbol{\Gamma}^\top\boldsymbol{U})\boldsymbol{\Lambda}(\boldsymbol{\Gamma}\boldsymbol{U})\right)\\
&= \operatorname{tr}\left(\frac{۱}{n-۱}\boldsymbol{W}^\top\boldsymbol{\Lambda}\boldsymbol{W}\right)\qquad \boldsymbol{W}=\boldsymbol{\Gamma}\boldsymbol{U}\\
&= \frac{۱}{n-۱}\sum_{i=۱}^{p}\lambda_i W_i^\top W_i \qquad (۳.۴)
\end{aligned}
$$

به‌طوری که $\boldsymbol{W}=\boldsymbol{\Gamma}_{p\times n}\boldsymbol{U}_{n\times ۱}=(\boldsymbol{w}_۱,\ldots,\boldsymbol{w}_p)$ و $\boldsymbol{w}_i \overset{iid}{\sim} \mathcal{N}_n(\circ,\boldsymbol{I}_n)$. بنابراین، اگر $v_{ii}=\boldsymbol{w}_i^\top\boldsymbol{w}_i$ آن‌گاه $\nu_{ii}\sim\chi^۲_n$. پس

$$
\begin{aligned}
(n-۱)p\hat{a}_۱ &= (n-۱)(\operatorname{tr}(\boldsymbol{S}))=\sum_{i=۱}^{p}\lambda_i\nu_{ii}\\
(n-۱)^۲(\operatorname{tr}(\boldsymbol{S}))^۲ &= \left[\sum_{i=۱}^{p}\lambda_i^۲\nu_{ii}^۲ + ۲\sum_{i=۱}^{p}\sum_{\substack{j=۱\\ j\neq i}}^{p}\lambda_i\lambda_j\nu_{ii}\nu_{jj}\right]
\end{aligned}
$$

و

$$
\begin{aligned}
(n-۱)^۲\operatorname{tr}(\boldsymbol{S}^۲) &= \operatorname{tr}\Big(\big(\boldsymbol{W}^\top\boldsymbol{\Lambda}\boldsymbol{W}\big)\big(\boldsymbol{W}^\top\boldsymbol{\Lambda}\boldsymbol{W}\big)\Big)\\
&= \operatorname{tr}\left[\Big(\sum_{i=۱}^{p}\lambda_i\boldsymbol{W}_i\boldsymbol{W}_i^\top\Big)\Big(\sum_{i=۱}^{p}\lambda_i\boldsymbol{W}_i\boldsymbol{W}_i^\top\Big)\right]\\
&= \operatorname{tr}\left[\sum_{i=۱}^{p}\lambda_i^۲\boldsymbol{W}_i\boldsymbol{W}_i^\top\boldsymbol{W}_i\boldsymbol{W}_i^\top + ۲\sum_{i=۱}^{p}\sum_{\substack{j=۱\\ j\neq i}}^{p}\lambda_j\lambda_i\boldsymbol{W}_i\boldsymbol{W}_i^\top\boldsymbol{W}_j\boldsymbol{W}_j^\top\right]\\
&= \operatorname{tr}\left[\sum_{i=۱}^{p}\lambda_i^۲(\boldsymbol{W}_i\boldsymbol{W}_i^\top)^۲ + ۲\sum_{i=۱}^{p}\sum_{\substack{j=۱\\ j\neq i}}^{p}\lambda_i\lambda_j(\boldsymbol{W}_i\boldsymbol{W}_j)^۲\right]\\
&= \sum_{i=۱}^{p}\lambda_i^۲(\boldsymbol{W}_i\boldsymbol{W}_i^\top)^۲ + ۲\sum_{i=۱}^{p}\sum_{\substack{j=۱\\ j\neq i}}^{p}\lambda_i\lambda_j(\boldsymbol{W}_i\boldsymbol{W}_j)^۲\\
&= \sum_{i=۱}^{p}\lambda_i^۲\nu_{ii}^۲ + ۲\sum_{i=۱}^{p}\sum_{\substack{j=۱\\ j\neq i}}^{p}\lambda_i\lambda_j\nu_{ij}^۲
\end{aligned}
$$

که در آن

$$\nu_{ij}=\boldsymbol{W}_i^\top\boldsymbol{W}_j \qquad i\neq j$$

و

$$\nu_{ii}=\boldsymbol{W}_i^\top\boldsymbol{W}_i.$$

با توجه به اینکه $\frac{n^۲}{(n-۱)(n+۲)} \simeq ۱$، بنابراین

$$\begin{aligned}
a_۲ &\simeq \frac{۱}{p}\left[\operatorname{tr}(\boldsymbol{S}^۲) - \frac{۱}{(n-۱)}(\operatorname{tr}(\boldsymbol{S}))^۲\right] \\
&= \frac{۱}{(n-۱)^۲ p}\left[\sum_{i=۱}^{p}\lambda_i^۲\nu_{ii}^۲ + ۲\sum_{i=۱}^{p}\sum_{\substack{j=۱\\ j\neq i}}^{p}\lambda_i\lambda_j\nu_{ij}^۲ - \frac{۱}{(n-۱)}\left(\sum_{i=۱}^{p}\lambda_i^۲\nu_{ii}^۲ + ۲\sum_{i=۱}^{p}\sum_{\substack{j=۱\\ j\neq i}}^{p}\lambda_i\lambda_j\nu_{ij}^۲\right)\right] \\
&= \frac{۱}{(n-۱)^۲ p}\left[\left(۱ - \frac{۱}{(n-۱)}\right)\sum_{i=۱}^{p}\lambda_i^۲\nu_{ii}^۲ + ۲\sum_{i=۱}^{p}\sum_{\substack{j=۱\\ j\neq i}}^{p}\lambda_i\lambda_j\left(\nu_{ij}^۲ - \frac{۱}{(n-۱)}\nu_{ii}\nu_{jj}\right)\right] \\
&= \frac{۱}{(n-۱)^۲ p}\left[\left(\frac{n-۲}{n-۱}\right)\sum_{i=۱}^{p}\lambda_i^۲\nu_{ii}^۲ + ۲\sum_{i=۱}^{p}\sum_{\substack{j=۱\\ j\neq i}}^{p}\lambda_i\lambda_j\left(\nu_{ij}^۲ - \frac{۱}{(n-۱)}\nu_{ii}\nu_{jj}\right)\right] \\
&= \left[\frac{n-۲}{(n-۱)^۳ p}\sum_{i=۱}^{p}\lambda_i^۲\nu_{ii}^۲\right] + \left[\frac{۲}{(n-۱)^۲ p}\sum_{i=۱}^{p}\sum_{\substack{j=۱\\ j\neq i}}^{p}\lambda_i\lambda_j z_{ij}^۲\right] \\
&= q_۱ + q_۲
\end{aligned}$$

که در آن

$$\begin{aligned}
z_{ij} &= \nu_{ij}^۲ - \frac{۱}{(n-۱)}\nu_{ii}\nu_{jj} \\
&= (\boldsymbol{W}_i^\top \boldsymbol{W}_j) - \frac{۱}{(n-۱)}(\boldsymbol{W}_i^\top \boldsymbol{W}_i)(\boldsymbol{W}_j^\top \boldsymbol{W}_j).
\end{aligned}$$

با توجه به رابطه (۴.۴)،

$$\begin{aligned}
E(a_۲) &= \frac{۱}{p}E\left[\operatorname{tr}(\boldsymbol{S}^۲) - \frac{۱}{n}(\operatorname{tr}(\boldsymbol{S}))^۲\right] \\
&= \frac{n-۲}{(n-۱)^۳ p}E\left[\sum_{i=۱}^{p}\lambda_i^۲\nu_{ii}^۲\right] + \frac{۲}{(n-۱)^۲ p}E\left[\sum_{i=۱}^{p}\sum_{\substack{j=۱\\ j>۱}}^{p}\lambda_i\lambda_j z_{ij}\right] \\
&= \frac{n-۲}{(n-۱)^۳ p}(n-۱)(n+۱)\sum_{i=۱}^{p}\lambda_i^۲ + \frac{۲}{(n-۱)^۲ p}\sum_{i=۱}^{p}\sum_{\substack{j=۱\\ j>۱}}^{p}\lambda_i\lambda_j E(z_{ij}) \\
&= \frac{(n-۲)(n+۱)}{(n-۱)^۲}\left(\frac{\operatorname{tr}(\boldsymbol{\Sigma}^۲)}{p}\right)
\end{aligned}$$

بنابراین

$$\frac{(n-۱)^۲}{(n-۲)(n+۱)}\frac{۱}{p}E\left[\operatorname{tr}(\boldsymbol{S}^۲) - \frac{۱}{(n-۱)}(\operatorname{tr}(\boldsymbol{S}))^۲\right]$$

برآوردگر ناریب $\operatorname{tr}(\boldsymbol{\Sigma}^۲)/p$ است. □

با این وجود تحت فرض غیرنرمال بودن، سریواستاوا و همکاران (۲۰۱۴) نشان دادند که

$$E\left[\hat{a}_۲\right] = a_۲ + \frac{n-۱}{n(n+۱)}\kappa_۴ a_{۲۰}.$$

که نتیجه می‌دهد وقتی $a_{۲۰} = O(۱)$، برآوردگر $\hat{a}_۲$ اریبی مرتبه دوم دارد. سریواستاوا و همکاران (۲۰۱۴) و هیمنو و یامادا (۲۰۱۴) به‌طور مستقل از یکدیگر، برآوردگر دقیق ناریب $\hat{a}_{۲C}$ را به‌صورت زیر پیشنهاد دادند

$$\hat{a}_{۲C} = \frac{n-۱}{n(n-۲)(n-۳)} \left\{ (n-۱)(n-۲)\frac{۱}{p}\operatorname{tr}(\boldsymbol{S}^۲) + p\hat{a}_۱^۲ - \frac{۱}{p}n\boldsymbol{Q} \right\}$$

که در آن

$$\boldsymbol{Q} = \frac{۱}{n-۱}\sum_{i=۱}^{n} \left\{ (\boldsymbol{x}_i - \bar{\boldsymbol{x}})^\top (\boldsymbol{x}_i - \bar{\boldsymbol{x}}) \right\}^۲ \qquad (۴.۴)$$

همچنین نشان دادند که برآوردگر $\hat{a}_{۲C}$ با برآوردگر پیشنهادی چن و همکاران (۲۰۱۰) یکسان است. در مقایسه با شکل ساده چن و همکاران (۲۰۱۰) عبارت $\hat{a}_{۲C}$ ساده است و به آسانی به‌کار برده می‌شود.

درباره برآورد $a_۱^۲$، سریواستاوا و همکاران (۲۰۱۱) نشان دادند که در حالت غیرنرمال،

$$\operatorname{Var}(\hat{a}_۱) = \frac{۲}{(n-۱)p}a_۲ + \frac{۱}{np}\kappa_۴ a_{۲۰}.$$

بنابراین

$$E\left[\hat{a}_۱^۲\right] = a_۱^۲ + \frac{۲}{(n-۱)p}a_۲ + \frac{۱}{np}\kappa_۴ a_{۲۰}$$

که بدین معنی است که وقتی $a_۲ = O(۱)$ و $a_{۲۰} = O(۱)$، در این صورت

$$E\left[\hat{a}_۱^۲\right] = a_۱^۲ + O\left(\frac{۱}{np}\right).$$

بنابراین از برآوردگر $\hat{a}_۱^۲$ برای $a_۱^۲$ استفاده می‌شود.

با استدلال‌های بالا، می‌توان $G(\boldsymbol{\Lambda}_۱)$ را با $\hat{G}_۱ = p(\hat{a}_{۲C} - \hat{a}_۱^۲)$ برآورد کرد و برآوردگر نتیجه‌شده با وزن بهینه

$$\hat{w}_{U۱} = \max\left\{ ۰, \min\left\{ \frac{\frac{۱}{p}\operatorname{tr}(\boldsymbol{S}^۲) - \hat{a}_{۲C}}{\frac{۱}{p}\operatorname{tr}(\boldsymbol{S}^۲) - \hat{a}_۱^۲} \right\}, ۱ \right\}$$

به‌صورت

$$\hat{\boldsymbol{\Sigma}}_{\boldsymbol{U}} = \hat{w}_U \boldsymbol{S} + (۱ - \hat{w}_U)\hat{a}_۱ \boldsymbol{I}_p$$

می‌باشد. مشابه لدویت و ولف (۲۰۰۴) معیار PRIAL روی ماتریس کوواریانس نمونه به‌صورت زیر تعریف می‌شود

$$\text{PRIAL} = \frac{E\left[\operatorname{tr}(\boldsymbol{S} - \boldsymbol{\Sigma})^۲\right] - E\left[\operatorname{tr}(\hat{\boldsymbol{\Sigma}}_{w^*}(\hat{a}_۱\boldsymbol{I}_p) - \boldsymbol{\Sigma})^۲\right]}{E\left[\operatorname{tr}(\boldsymbol{S} - \boldsymbol{\Sigma})^۲\right]}$$

قضیه ۳.۴.۴. *به ازای مقدار بزرگ p، فرض کنید $a_۱ = O(۱)$، $a_{۲۰} = O(۱)$ و $a_۲ = O(p^\delta)$. به‌ازای $\delta \geq ۰$،*

(حالت ۱) *اگر* $\frac{na_۲}{p} \to \infty$ *، آن‌گاه* $\text{PRIAL} \to ۰$.

(حالت ۲) *اگر* $\frac{na_۲}{p} \to ۰$ *، آن‌گاه* $\text{PRIAL} \to ۱$.

(حالت ۳) *اگر برای* C *ثابت مثبت،* $\frac{na_۲}{p} \to C$ *، در این‌صورت* PRIAL *به مقدار*

$$\frac{a_۱^۲ + \frac{۱}{n}C}{C - \frac{n}{p}a_۱^۲ + a_۱^۲ + \frac{۱}{n}C}$$

میل می‌کند.

برهان. می‌دانیم

$$\begin{aligned} E\left[\text{tr}(\hat{\boldsymbol{\Sigma}}_{w^*} - \boldsymbol{\Sigma})^۲\right] &= E\left[\text{tr}(\boldsymbol{S} - \boldsymbol{\Sigma})\right] - ۲E\left[\text{tr}(\boldsymbol{S} - \boldsymbol{\Sigma})(\boldsymbol{S} - \hat{a}_۱\boldsymbol{I}_p)\right] \\ &\quad + E\left[\text{tr}(\boldsymbol{S} - \hat{a}_۱\boldsymbol{I}_p)^۲\right] \end{aligned}$$

در این‌صورت PRIAL به‌صورت زیر به‌دست می‌آید.

$$\text{PRIAL} = w^* \frac{۲E\left[\text{tr}(\boldsymbol{S} - \boldsymbol{\Sigma})(\boldsymbol{S} - \hat{a}_۱\boldsymbol{I}_p)\right] - w^* E\left[\text{tr}(\boldsymbol{S} - \hat{a}_۱\boldsymbol{I}_p)^۲\right]}{E\left[\text{tr}(\boldsymbol{S} - \boldsymbol{\Sigma})^۲\right]}$$

با توجه به اینکه

$$w^* = \frac{E\left[\text{tr}(\boldsymbol{S} - \boldsymbol{\Sigma})(\boldsymbol{S} - \hat{a}_۱\boldsymbol{I}_p)\right]}{E\left[\text{tr}(\boldsymbol{S} - \hat{a}_۱\boldsymbol{I}_p)^۲\right]}$$

عبارت PRIAL برابر است با

$$\text{PRIAL} = \frac{E\left[\text{tr}(\boldsymbol{S} - \boldsymbol{\Sigma})(\boldsymbol{S} - \hat{a}_۱\boldsymbol{I}_p)\right]}{E\left[\text{tr}(\boldsymbol{S} - \hat{a}_۱\boldsymbol{I}_p)^۲\right]} \frac{A}{E\left[\text{tr}(\boldsymbol{S} - \boldsymbol{\Sigma})^۲\right]}$$

که

$$A = ۲E\left[\text{tr}(\boldsymbol{S} - \boldsymbol{\Sigma})(\boldsymbol{S} - \hat{a}_۱\boldsymbol{I}_p)\right] - \frac{E\left[\text{tr}(\boldsymbol{S} - \boldsymbol{\Sigma})(\boldsymbol{S} - \hat{a}_۱\boldsymbol{I}_p)\right]}{E\left[\text{tr}(\boldsymbol{S} - \hat{a}_۱\boldsymbol{I}_p)^۲\right]} E\left[\text{tr}(\boldsymbol{S} - \hat{a}_۱\boldsymbol{I}_p)^۲\right]$$

صورت و مخرج کسر فوق را می‌توان به‌ترتیب به‌صورت زیر نوشت:

$$\begin{aligned} &E\left[\text{tr}(\boldsymbol{S} - \boldsymbol{\Sigma})(\boldsymbol{S} - \hat{a}_۱\boldsymbol{I}_p)\right] \Big\{ ۲E\left[\text{tr}(\boldsymbol{S} - \boldsymbol{\Sigma})(\boldsymbol{S} - \hat{a}_۱\boldsymbol{I}_p)\right] E\left[\text{tr}(\boldsymbol{S} - \hat{a}_۱\boldsymbol{I}_p)^۲\right] \\ &\quad - E\left[\text{tr}(\boldsymbol{S} - \boldsymbol{\Sigma})(\boldsymbol{S} - \hat{a}_۱\boldsymbol{I}_p)\right] E\left[\text{tr}(\boldsymbol{S} - \hat{a}_۱\boldsymbol{I}_p)^۲\right] \Big\} \end{aligned}$$

و

$$E\left[\text{tr}(\boldsymbol{S} - \hat{a}_۱\boldsymbol{I}_p)^۲\right] E\left[\text{tr}(\boldsymbol{S} - \boldsymbol{\Sigma})^۲\right]$$

بنابراین PRIAL برابر است با

$$\text{PRIAL} = \frac{E\left[\text{tr}(\boldsymbol{S} - \boldsymbol{\Sigma})(\boldsymbol{S} - \hat{a}_۱\boldsymbol{I}_p)\right] E\left[\text{tr}(\boldsymbol{S} - \boldsymbol{\Sigma})(\boldsymbol{S} - \hat{a}_۱\boldsymbol{I}_p)\right] E\left[\text{tr}(\boldsymbol{S} - \hat{a}_۱\boldsymbol{I}_p)^۲\right]}{E\left[\text{tr}(\boldsymbol{S} - \hat{a}_۱\boldsymbol{I}_p)^۲\right] E\left[\text{tr}(\boldsymbol{S} - \boldsymbol{\Sigma})^۲\right]}$$

$$= \frac{\{E\left[\mathrm{tr}(\boldsymbol{S}-\boldsymbol{\Sigma})(\boldsymbol{S}-\hat{a}\boldsymbol{I}_p)\right]\}^2}{E\left[\mathrm{tr}(\boldsymbol{S}-\hat{a}_1\boldsymbol{I}_p)^2\right]E\left[\mathrm{tr}(\boldsymbol{S}-\boldsymbol{\Sigma})^2\right]}$$

با به‌کار بردن لم ۱.۴.۴، عبارت PRIAL به‌صورت زیر بیان می‌شود

$$\mathrm{PRIAL} = \frac{\left\{\frac{1}{n-1}a_2+\frac{p}{n-1}a_1^2+\frac{1}{n}\kappa_4 a_{20}\right\}\left\{a_2-a_1^2+\frac{p}{n-1}a_1^2+\frac{p-2}{(n-1)p}a_2+\frac{p-1}{np}\kappa_4 a_{20}\right\}}{\left\{\frac{p}{n-1}a_1^2+\frac{p-2}{(n-1)p}a_2+\frac{p-1}{np}\kappa_4 a_{20}\right\}^2}$$

$$= \frac{\left\{a_1^2+\frac{p-2}{p^2}a_2+\frac{(p-1)(n-1)}{np^2}\kappa_4 a_{20}\right\}^2}{\left\{a_1^2+\frac{1}{p}a_2+\frac{n-1}{n}\frac{1}{p}\kappa_4 a_{20}\right\}\left\{a_1^2+\frac{n-1}{p}a_2-\frac{n-1}{p}a_1^2+\frac{p-2}{p^2}a_2+\frac{(n-1)(p-1)}{np}\frac{1}{p}\kappa_4 a_{20}\right\}}$$

در این عبارت، می‌توان مقادیر PRIAL را در سه حالت بررسی کرد

(حالت اول) یعنی $\frac{n}{p}a_2 \to \infty$. با توجه به فرض‌های قضیه به راحتی نتیجه می‌شود $\mathrm{PRIAL} \to 0$

(حالت دوم) یعنی $\frac{n}{p}a_2 \to 0$، عبارت PRIAL به ازای $p \to \infty$ برابر است با

$$\frac{\left\{a_1^2+0+0\right\}^2}{\left\{a_1^2+0+0\right\}\left\{0+0+a_1^2+0\right\}}$$

(حالت سوم) یعنی $\frac{n}{p}a_2 \to 0$، عبارت PRIAL برابر است با

$$\lim_{p\to\infty}\frac{\left\{a_1^2+(1-\frac{p-2}{p})\frac{1}{n}C+\frac{(p-1)(n-1)}{np}\frac{1}{p}\kappa_4 a_{20}\right\}^2}{\left\{a_1^2+\frac{1}{n}C+\frac{n-1}{n}\frac{1}{p}\kappa_4 a_{20}\right\}\left\{a_1^2+C-\frac{n-1}{p}a_1^2+\left(1-\frac{2}{p}\right)\frac{1}{n}C+\frac{(n-1)(p-1)}{np}\frac{1}{p}\kappa_4 a_{20}\right\}}$$

$$= \frac{\left\{a_1^2+\frac{1}{n}C\right\}^2}{\left\{a_1^2+\frac{1}{n}C\right\}\left\{C-\frac{n}{p}a_1^2+a_1^2+\frac{1}{n}C\right\}}$$

و اثبات کامل می‌شود. □

در کاربرد قضیه ۳.۴.۴ می‌توان گفت چنانچه مقدار na_2 خیلی بزرگتر از p شود، مقادیر PRIAL نزدیک صفر بوده و چنانچه خیلی کوچک‌تر از P باشد به‌طوری که بتوان نتیجه گرفت $\frac{na_2}{P} \to 0$، آنگاه مقادیر PRIAL نزدیک ۱ است. در این حالت برآوردگر $\hat{\boldsymbol{\Sigma}}_{w^*}$ برآوردگری کارا بوده و بهتر از $\boldsymbol{S}$ است.

۵.۴ مطالعه شبیه‌سازی

در این بخش کارایی‌های عددی توابع مخاطره برآوردگرهای انقباضی خطی با به‌کارگیری شبیه‌سازی مونت کارلو مورد بررسی قرار می‌گیرد. در این راستا، فرض کنید $\Delta = \mathrm{diag}(d_1,\ldots,d_p)$ که در آن

$$d_i = 0/1 + 10 \times U_i \qquad (۵.۴)$$

و U_i عدد تصادفی است که از توزیع یکنواخت $(۰,۱)$ تولید شده است. همچنین σ_{ij}، (i,j) امین مولفه ماتریس کوواریانس Σ باشد که چهار ساختار زیر را برای آن در نظر می‌گیریم

(مدل الف)

$$\sigma_{ij} = d_i d_j \rho^{|i-j|}$$

(مدل ب)

$$\sigma_{ij} = \rho^{|i-j|}$$

(مدل ج)

$$\sigma_{ij} = d_i d_j \frac{|i-j+۱|^{۲h} - ۲|i-j|^{۲h} + |i-j-۱|^{۲h}}{۲}$$

(مدل د)

$$\sigma_{ij} = \frac{|i-j+۱|^{۲h} - ۲|i-j|^{۲h} + |i-j-۱|^{۲h}}{۲}.$$

مدل‌های (ب) و (د) به ترتیب متناظر با ساختار اتورگرسیو و حرکت براونی کسری [۱۷] (FBM) هستند، که هردو آن‌ها در چن و همکاران (۲۰۱۰) مورد بررسی قرار گرفته‌اند.

دقت کنید هر چهار مدل در نظر گرفته‌شده برای ماتریس کوواریانس در شرایط در نظر گرفته‌شده $a_i = O(۱)$ به ازای $i = ۱,۲$ و $a_{۲۰} = O(۱)$ صدق می‌کنند.

مشاهده‌های تصادفی $\boldsymbol{x}_۱$، ... و $\boldsymbol{x}_n$ به‌صورت $\boldsymbol{x}_i = \boldsymbol{\Sigma}^{\frac{۱}{۲}} \boldsymbol{u}_i$ تولید شده‌اند که $\boldsymbol{u}_i = (u_{i۱}, \ldots, u_{pi})^\top$ و $u_{۱i}$، ... و $\boldsymbol{u}_{pi}$ به‌طور مستقل به یکی از سه حالت زیر توزیع شده‌اند

(حالت اول)

$$u_{ji} \sim \mathcal{N}(۰,۱)$$

(حالت دوم)

$$u_{ji} \sim \sqrt{\frac{(m-۳)}{m}} z_{ji} \quad \text{که در آن} \quad z_{ji} \sim t_m$$

(حالت سوم)

$$u_{ji} \sim \frac{(z_{ji} - \nu)}{\sqrt{۲\nu}} \quad \text{که در آن} \quad z_{ji} \sim \chi^۲_\nu$$

به‌طوری که t_m توزیع t با m درجه آزادی و $\chi^۲_\nu$ توزیع کی دو با ν درجه آزادی را نشان می‌دهند. حالت دوم، توزیع‌های دم پهن و حالت سوم توزیع‌های چوله را در بر می‌گیرند.

کارایی برآوردگرهای پیشنهادی با سایر برآوردگرها با نماد $\hat{\Sigma}_U$، $\hat{\Sigma}_{LW}$، $\hat{\Sigma}_{RBLW}$، $\hat{\Sigma}_{OAS}$ و $\hat{\Sigma}_{FS}$ نشان داده شده و برای راحتی در جدول‌ها از U، LW، $RBLW$، OAS و FS استفاده شده است.

[۱۷] Fractional Brownian motion

مطالعه شبیه‌سازی برای $p = ۱۰۰$، $n = ۱۰, ۴۰, ۷۰, ۱۰۰, ۱۳۰$، $\rho = ۰/۲, ۰/۴$، $h = ۰/۷$، $m = ۵$ و $\nu = ۲$ انجام شده است. بر اساس ۱۰۰۰ تکرار، مخاطره تجربی این برآوردگرها محاسبه و معیار PRIAL برای هر برآوردگر $\hat{\Sigma}$ نسبت به ماتریس نمونه که به‌صورت زیر تعریف می‌شود، گزارش شده است

$$\text{PRIAL} = ۱۰۰ \times \frac{E[tr(\boldsymbol{S} - \boldsymbol{\Sigma})^۲] - E[tr(\hat{\boldsymbol{\Sigma}} - \boldsymbol{\Sigma})]}{E[tr(\boldsymbol{S} - \boldsymbol{\Sigma})]}$$

جدول ۱.۴ مقادیر PRIAL را برای مدل (الف) با $\rho = ۰/۲$ و $\rho = ۰/۹$ در توزیع‌های نرمال، $\sqrt{\frac{۳}{۵}}t_۵$ و $\frac{۱}{۴}(\chi^۲_۲ - ۲)$ که همان حالت‌های ۱، ۲و ۳ هستند، نشان می‌دهد.

جدول ۲.۴، مقادیر PRIAL را برای مدل (ب) با در نظر گرفتن همان فرض‌های جدول ۱.۴ گزارش می‌کند. مقادیر PRIAL برای مدل‌های (ج) و (د) در توزیع‌های $\sqrt{\frac{۳}{۵}}t_۵$ و $\frac{۱}{۴}(\chi^۲_۲ - ۲)$ در جدول ۳.۴ آورده شده‌است.

از این جداول نتایج زیر برداشت می‌شود:

۱) برآوردگر پیشنهادی بهتر از سایر برآوردگرها رفتار می‌کند اگرچه این تفاوت اندک است. این حالت‌ها شبیه ماتریس‌های قطری هستند که عناصر قطری آن تغییرپذیری زیادی دارند.

۲) کارایی برآوردگرهای $\hat{\Sigma}_{AOS}$ و $\hat{\Sigma}_{FS}$ بسیار نزدیک هم است.

۳) از آن‌جایی که $\hat{\Sigma}_{RBLW}$ برآوردگر رائو–بلکول شده‌ای از $\hat{\Sigma}_{LW}$ با در نظر گرفتن فرض نرمال است، این برآوردگر در حالت اول (توزیع نرمال) بهتر از $\hat{\Sigma}_{LW}$ رفتار می‌کند، اما این ویژگی لزوما در حالت‌های غیرنرمال (توزیع‌های $t_۵$ و $\chi^۲_۲$) برقرار نیست. دقت کنید $\hat{\Sigma}_U$ برآوردگر انقباضی خطی ناپارامتری است درحالی که $\hat{\Sigma}_{OAS}$ و $\hat{\Sigma}_{FS}$ از برآوردگرهای وزن بهینه با فرض نرمال بودن استفاده می‌کند.

نتایج جدول‌های ۱.۴–۳.۴، این حقیقت را تایید می‌کند که برآوردگرهای $\hat{\Sigma}_{OAS}$ و $\hat{\Sigma}_{FS}$ کمی بهتر از $\hat{\Sigma}_U$ در بیشتر موارد در نظر گرفته‌شده برای حالت اول (توزیع نرمال) است، اما برآوردگر پیشنهادی در بیشتر موارد گزارش‌شده برای توزیع‌های غیرنرمال بهتر است.

در جدول ۴.۴ ، کارایی برآوردگرهای انقباضی خطی برای مدل‌های (الف) و (ب) به ازای $\rho = ۰/۲$، $n = ۴۰$ و $p = ۱۰۰$ در حالت توزیع t با m درجه آزادی برای $m = ۳, ۴, ۵, ۶, ۱۰$ بررسی شده است. حالت‌های ۳ و ۴ را به منظور بررسی کارایی برآوردگر در حالت توزیع‌های دم پهن در نظر گرفتیم. مشاهده می‌شود که $\hat{\Sigma}_{LW}$ به $\hat{\Sigma}_{PBLW}$، $\hat{\Sigma}_{OAS}$ و $\hat{\Sigma}_{FS}$ برای مقادیر کوچک m برتری دارد و صرف نظر از مقدار m، برآوردگر پیشنهادی بهتر از $\hat{\Sigma}_{LW}$ است.

۶.۴ خلاصه و پیشنهادات

در این مجموعه برآورد ماتریس کوواریانس یک بردار p بعدی را وقتی p بزرگ است مورد مطالعه قرار دادیم. در این راستا برآوردگرهای انقباضی خطی ماتریس کوواریانس و برآوردگرهای سازگار توابعی از آن را بررسی کرده و با استفاده از مطالعات شبیه‌سازی مقایسه کردیم. از آنجایی که تعیین پارامتر انقباضی در برآوردگر انقباضی خطی ماتریس کوواریانس با بعد بالا از اهمیت بسزایی برخوردار است تحت سناریوهای مختلف شبیه‌سازی شده، مقدار بهینه آن را در هر حالت تعیین کردیم.

با توجه به مطالب ارائه‌شده می‌توان برای آینده این تحقیق موضوعات زیر را پیشنهاد داد:

۱) با توجه به اینکه در تحلیل ممیزی خطی به معکوس برآوردگر ماتریس کوواریانس نیاز است، می‌توان از برآوردگر انقباضی مطرح‌شده در این مجموعه در مسائل با بعد بالا استفاده کرد.

۲) در هیچ یک از حالت‌های مورد بحث دیدگاه بیزی برآوردگر کوواریانس بحث نشده که این موضوع می‌تواند در بعد بالا مورد توجه باشد.

۳) صالح (۲۰۰۶) یک دسته جدیدی از برآوردگرهای انقباضی واریانس را بر پایه برآوردگر انقباضی نوع استاین پیشنهاد کرد که خواص جالب توجهی دارد. می‌توان با الگو گرفتن از آن نتایج را برای حالت ماتریسی تعمیم داد.

جدول ۱.۴: مقدار PRIAL برآوردگرهای انقباضی خطی برای مدل (الف) به ازای $\rho =$ ۰/۲, ۰/۹ و $p =$ ۱۰۰ در حالت‌های ۱، ۲ و ۳

	n	U_1	U_2	LW	$RBLW$	OAS	FS
	۱۰	۸۳/۶۰	۹۳/۰۳	۰۰/۴۸	۸۲/۷۳	۸۳/۶۹	۸۳/۶۹
نرمال	۴۰	۵۴/۵۵	۹۰/۰۹	۵۴/۴۴	۵۴/۵۴	۵۴/۵۶	۵۴/۵۶
مدل الف	۷۰	۴۰/۵۴	۸۷/۳۸	۴۰/۵۲	۴۰/۵۵	۴۰/۵۵	۴۰/۵۵
$\rho =$ ۰/۲	۱۰۰	۳۲/۱۸	۸۴/۸۲	۳۲/۱۷	۳۲/۱۸	۳۲/۱۸	۳۲/۱۸
	۱۳۰	۲۶/۸۷	۸۲/۴۶	۲۶/۸۷	۲۶/۸۷	۲۶/۸۷	۲۶/۸۷
	۱۰	۸۴/۹۷	۸۳/۴۰	۸۱/۳۸	۸۱/۰۹	۸۳/۳۱	۸۳/۲۷
t_5	۴۰	۵۷/۹۳	۷۹/۳۷	۵۷/۷۹	۵۵/۱۶	۵۵/۴۴	۵۵/۴۴
مدل الف	۷۰	۴۴/۹۰	۷۵/۸۲	۴۴/۸۶	۴۱/۴۴	۴۱/۵۶	۴۱/۵۷
$\rho =$ ۰/۲	۱۰۰	۳۶/۶۷	۷۳/۳۹	۳۶/۶۴	۳۳/۵۷	۳۳/۶۴	۳۳/۶۴
	۱۳۰	۳۱/۴۳	۷۰/۷۲	۳۱/۴۲	۲۷/۷۸	۲۷/۸۲	۲۷/۸۲
	۱۰	۸۵/۱۴	۸۰/۹۲	۸۱/۶۰	۸۱/۱۴	۸۳/۷۴	۸۳/۵۶
χ^2_2	۴۰	۵۸/۱۲	۷۷/۶۰	۵۸/۰۰	۵۶/۲۸	۵۶/۶۴	۵۶/۶۴
مدل الف	۷۰	۴۴/۰۵	۷۵/۲۷	۴۴/۰۳	۴۲/۷۹	۴۲/۹۴	۴۲/۹۴
$\rho =$ ۰/۲	۱۰۰	۳۵/۵۸	۷۲/۹۶	۳۵/۵۷	۳۴/۵۳	۳۴/۶۱	۳۴/۶۲
	۱۳۰	۲۹/۸۳	۷۰/۸۵	۲۹/۸۳	۲۸/۹۸	۲۹/۰۳	۲۹/۰۳
	۱۰	۴۷/۷۶	۴۹/۰۴	۴۶/۹۰	۴۸/۳۰	۴۸/۵۲	۴۸/۵۶
χ^2_2	۴۰	۱۷/۶۵	۱۸/۶۹	۱۷/۷۲	۱۷/۷۹	۱۷/۷۴	۱۷/۷۴
مدل الف	۷۰	۱۰/۸۰	۱۱/۴۷	۱۰/۸۳	۱۰/۸۱	۱۰/۸۰	۱۰/۸۰
$\rho =$ ۰/۹	۱۰۰	۷/۸۴	۸/۳۹	۷/۸۵	۷/۸۷	۷/۸۶	۷/۸۶
	۱۳۰	۶/۵۵	۶/۹۶	۶/۵۶	۶/۵۶	۶/۵۶	۶/۵۶
	۱۰	۴۹/۸۷	۴۸/۹۶	۴۸/۷۲	۴۸/۶۶	۴۹/۴۹	۴۹/۴۹
χ^2_2	۴۰	۱۹/۳۹	۱۹/۱۱	۱۹/۴۳	۱۸/۵۵	۱۸/۵۷	۱۸/۵۷
مدل الف	۷۰	۱۲/۳۳	۱۱/۹۶	۱۲/۳۵	۱۱/۵۴	۱۱/۵۴	۱۱/۵۵
$\rho =$ ۰/۹	۱۰۰	۹/۱۹	۸/۷۶	۹/۱۹	۸/۴۳	۸/۴۳	۸/۴۳
	۱۳۰	۶/۶۷	۶/۴۶	۶/۶۸	۶/۱۸	۶/۱۸	۶/۱۸
	۱۰	۵۰/۰۸	۴۹/۴۸	۴۸/۸۹	۴۹/۱۸	۵۰/۱۰	۵۰/۰۹
χ^2_2	۴۰	۱۹/۲۵	۱۹/۳۸	۱۹/۲۹	۱۸/۸۵	۱۸/۸۸	۱۸/۸۹
مدل الف	۷۰	۱۱/۹۱	۱۲/۰۲	۱۱/۹۳	۱۱/۶۰	۱۱/۶۱	۱۱/۶۲
$\rho =$ ۰/۹	۱۰۰	۸/۷۱	۸/۸۰	۸/۷۲	۸/۴۶	۸/۴۷	۸/۴۷
	۱۳۰	۶/۸۱	۶/۹۲	۶/۸۲	۶/۶۳	۶/۶۳	۶/۶۳

جدول ۲.۴: مقدار PRIAL برآوردگرهای انقباضی خطی برای مدل (ب) به ازای $\rho =$ ۰/۲, ۰/۹ و $p =$ ۱۰۰ در حالت‌های ۱، ۲ و ۳

	n	U_1	U_2	LW	$RBLW$	OAS	FS
نرمال	۱۰	۹۹/۲۱	۹۷/۲۷	۹۵/۳۳	۹۵/۳۳	۹۹/۲۳	۹۹/۲۱
مدل ب	۷۰	۹۴/۶۰	۹۲/۶۵	۹۴/۵۳	۹۴/۵۳	۹۴/۶۰	۹۴/۶۰
$\rho =$ ۰/۲	۱۳۰	۹۰/۴۲	۸۸/۴۸	۹۰/۴۱	۹۰/۴۰	۹۰/۴۲	۹۰/۴۲
t_5	۱۰	۹۹/۲۱	۹۲/۷۴	۹۶/۵۷	۹۵/۱۳	۹۸/۴۶	۹۸/۴۳
مدل ب	۷۰	۹۴/۸۳	۰۸۸/۰۳	۹۳/۹۰	۹۴/۷۶	۹۴/۰۴	۹۴/۰۴
$\rho =$ ۰/۲	۱۳۰	۹۰/۷۹	۸۴/۳۳	۹۵/۱۳	۹۰/۷۷	۹۰/۲۹	۹۰/۳۰
χ^2_2	۱۰	۹۹/۲۰	۹۲/۷۰	۹۴/۷۵	۹۵/۱۳	۹۸/۹۸	۹۸/۹۵
مدل ب	۷۰	۹۴/۸۲	۸۷/۹۴	۹۰۸۱	۹۴/۷۵	۹۴/۵۵	۹۴/۵۴
$\rho =$ ۰/۲	۱۳۰	۹۰/۸۳	۸۳/۹۶	۸۵/۲۹	۹۰/۸۱	۹۰/۵۵	۹۰/۵۵
نرمال	۱۰	۵۷/۶۴	۵۶/۷۲	۶۸/۲۵	۵۶/۰۴	۵۸/۱۳	۵۸/۱۵
مدل ب	۷۰	۱۵۰۲۷	۱۴/۸۰	۸۰/۲۳	۱۵/۲۹	۱۵/۲۸	۱۵/۲۸
$\rho =$ ۰/۹	۱۳۰	۹/۰۶	۸/۷۷	۹/۲۵	۹/۰۶	۹/۰۶	۹/۰۶
t_5	۱۰	۵۹/۰۱	۵۶/۶۸	۵۶/۲۰	۵۷/۲۰	۵۸/۵۰	۵۸/۴۸
مدل ب	۷۰	۱۶/۸۴	۱۵/۶۷	۱۵/۲۴	۱۶/۸۴	۱۶/۳۱	۱۶/۳۲
$\rho =$ ۰/۹	۱۳۰	۹/۳۶	۸/۷۱	۹/۴۰	۹/۳۷	۹/۰۹	۹/۰۹
χ^2_2	۱۰	۵۹/۲۲	۵۷/۳۱	۵۸/۷۸	۵۷/۳۵	۵۹/۲۱	۵۹/۱۹
مدل ب	۷۰	۱۶/۴۵	۱۵/۶۰	۱۶/۲۴	۱۶/۴۶	۱۶/۲۵	۱۶/۲۶
$\rho =$ ۰/۹	۱۳۰	۹/۵۳	۹/۰۱	۹/۴۰	۹/۵۳	۹/۴۰	۹/۴۱

جدول ۳.۴: مقدار PRIAL برآوردگرهای انقباضی خطی برای مدل‌های (ج) و (د) به ازای $h = ۰/۷$ و $p = ۱۰۰$ در حالت‌های ۲ و ۳

	n	$U_۱$	$U_۲$	LW	$RBLW$	OAS	FS
$t_۵$	۱۰	۸۱/۸۷	۸۰/۸۸	۸۱/۷۱	۷۸/۴۷	۷۸/۴۸	۸۰/۳۵
مدل ج	۷۰	۹۴/۶۰	۹۲/۶۵	۹۴/۵۳	۹۴/۵۳	۹۴/۶۰	۹۴/۶۰
$h = ۰/۷$	۱۳۰	۹۰/۴۲	۸۸/۴۸	۹۰/۴۱	۹۰/۴۰	۹۰/۴۲	۹۰/۴۲
$\chi^۲_۲$	۱۰	۹۹/۲۱	۹۲/۷۴	۹۶/۵۷	۹۵/۱۳	۹۸/۴۶	۹۸/۴۳
مدل ج	۷۰	۹۴/۸۳	۰۸۸/۰۳	۹۳/۹۰	۹۴/۷۶	۹۴/۰۴	۹۴/۰۴
$h = ۰/۷$	۱۳۰	۹۰/۷۹	۸۴/۳۳	۹۵/۱۳	۹۰/۷۷	۹۰/۲۹	۹۰/۳۰
$t_۵$	۱۰	۹۹/۲۰	۹۲/۷۰	۹۴/۷۵	۹۵/۱۳	۹۸/۹۸	۹۸/۹۵
مدل د	۷۰	۹۴/۸۲	۸۷/۹۴	۹۰۸۱	۹۴/۷۵	۹۴.۵۵	۹۴.۵۴
$h = ۰/۷$	۱۳۰	۹۰.۸۳	۸۳.۹۶	۸۵.۲۹	۹۰.۸۱	۹۰.۵۵	۹۰.۵۵
$\chi^۲_۲$	۱۰	۵۷.۶۴	۵۶.۷۲	۶۸.۲۵	۵۶.۰۴	۵۸.۱۳	۵۸.۱۵
مدل د	۷۰	۱۵۰۲۷	۱۴/۸۰	۸۰/۲۳	۱۵/۲۹	۱۵/۲۸	۱۵/۲۸
$h = ۰/۷$	۱۳۰	۹/۰۶	۸/۷۷	۹/۲۵	۹/۰۶	۹/۰۶	۹/۰۶

جدول ۴.۴: مقدار PRIAL برآوردگرهای انقباضی خطی برای مدل‌های (الف) و (ب) به ازای $\rho = ۰/۲$، $p = ۱۰۰$ و $n = ۴۰$ در توزیع t_m برای $m = ۱۰, ۶, ۵, ۴, ۳$

	n	$U_۱$	$U_۲$	LW	$RBLW$	OAS	FS
	۱۰	۵۵/۲۹	۸۷/۷۶	۸۸/۶۰	۵۵/۱۷	۵۵/۱۱	۵۵/۲۱
	۶	۵۶/۳۶	۸۳/۷۰	۸۶/۱۲	۵۶/۲۴	۵۵/۳۳	۵۵/۵۲
مدل الف	۵	۵۸/۴۵	۷۸/۵۰	۸۴/۵۷	۹۰/۴۰	۵۴/۷۸	۵۵/۰۶
	۴	۶۴/۷۸	۵۰/۴۸	۵۰/۹۵	۵۰/۹۶	۵۰/۸۸	۵۰/۲۲
	۳	۸۸/۸۷	۲۲/۵۸	۲۳/۳۷	۲۳/۴۳	۲۳/۵۶	۲۳/۷۸
	۱۰	۹۰/۷۹	۸۴/۳۳	۹۵/۱۳	۹۰/۷۷	۹۰/۲۹	۹۰/۳۰
	۶	۹۹/۲۰	۹۲/۷۰	۹۴/۷۵	۹۵/۱۳	۹۸/۹۸	۹۸/۹۵
مدل ب	۵	۹۴/۸۲	۸۷/۹۴	۹۰۸۱	۹۴/۷۵	۹۴/۵۵	۹۴/۵۴
	۴	۹۰/۸۳	۸۳/۹۶	۸۵/۲۹	۹۰/۸۱	۹۰/۵۵	۹۰/۵۵
	۳	۵۷/۶۴	۵۶/۷۲	۶۸/۲۵	۵۶/۰۴	۵۸/۱۳	۵۸/۱۵

پیوست آ

برنامه‌های کامپیوتری

```
library(expm)
library(matrixcalc)
set.seed(8)
N.Sim = 100
N = 40
n = N - 1
p = 40


Sigma = matrix(0,p,p)
for(i in 1:p){
for(j in 1:p){
Sigma[i,j] = 0.2^(abs(i-j))
}
}

a2 <- function(Sigma){
matrix.trace(Sigma%*%Sigma)/p
}

a1.2 <- function(Sigma){
(matrix.trace(Sigma)/p)^2
}

a2.tilde <- function(S){
S2 <- S%*%S
((n^2)/((n-1)*(n+2)))*(1/p)*(matrix.trace(S2)-(1/n)*((matrix.trace(S))^2))
}

a2.hat <- function(S){
S2 <- S%*%S
```

```
tr.Sigma2 = ((N-1)/(N*(N-2)*(N-3)))*((N-1)*(N-2)*matrix.trace(S2)
+ ((matrix.trace(S))^2) - N*Q)
tr.Sigma2 / p
}

a1.2.hat <- function(S){
S2 <- S%*%S
tr.Sigma.2 = ((N-1)/(N*(N-2)*(N-3)))*(2*matrix.trace(S2)
 + (N^2 - 3*N + 1)*((matrix.trace(S))^2) - N*Q)
tr.Sigma.2 / (p^2)
}

a1.2.tilde <- function(S){
((matrix.trace(S))/p)^2
}

A2.tilde <- numeric(N.Sim)
A2.hat <- numeric(N.Sim)
A1.2.tilde <- numeric(N.Sim)
A1.2.hat <- numeric(N.Sim)

for(i.sim in 1:N.Sim){
X = matrix(0,nrow = N, ncol = p)
for(i in 1:N){
z = rnorm(p)
X[i,] = sqrtm(Sigma)%*%z
}
S = var(X)
bar.X = colMeans(X)

sum = 0
for(i in 1:N){
sum = sum + (t(X[i,]-bar.X)%*%(X[i,]-bar.X))^2}
Q = sum/(N-1)

A2.tilde[i.sim] = a2.tilde(S)
A2.hat[i.sim] = a2.hat(S)
A1.2.tilde[i.sim] = a1.2.tilde(S)
A1.2.hat[i.sim] = a1.2.hat(S)

}

a2.tilde.result = mean(A2.tilde)
a2.hat.result = mean(A2.hat)
a1.2.tilde.result = mean(A1.2.tilde)
a1.2.hat.result = mean(A1.2.hat)
```

```
a2(Sigma)
a2.tilde.result
a2.hat.result

sd(A2.tilde)
sd(A2.hat)


a1.2(Sigma)
a1.2.tilde.result
a1.2.hat.result
sd(A1.2.tilde)
sd(A1.2.hat)


rm(list = ls(all = TRUE))
set.seed(1234231)
####################
library(mnormt)
library(Matrix)
library(expm)
library(matrixcalc)
library(MASS)
p = 100;
N = 10 #10 - 40 - 70 - 100 - 130
rho = 0.2 #0.9
h = 0.7
m = 5
nu = 2
n = 5000 #number of simulations
#####################################
mat.tr.S.Sigma1 = matrix(NA,nr = n,nc = 3)
mat.tr.S.Sigma2 = matrix(NA,nr = n,nc = 3)
mat.tr.S.Sigma3 = matrix(NA,nr = n,nc = 3)
mat.tr.S.Sigma4 = matrix(NA,nr = n,nc = 3)
#####################################
mat.tr.U1.Sigma1 = matrix(NA,nr = n,nc = 3)
mat.tr.U1.Sigma2 = matrix(NA,nr = n,nc = 3)
mat.tr.U1.Sigma3 = matrix(NA,nr = n,nc = 3)
mat.tr.U1.Sigma4 = matrix(NA,nr = n,nc = 3)
#####################################
mat.tr.U2.Sigma1 = matrix(NA,nr = n,nc = 3)
mat.tr.U2.Sigma2 = matrix(NA,nr = n,nc = 3)
mat.tr.U2.Sigma3 = matrix(NA,nr = n,nc = 3)
mat.tr.U2.Sigma4 = matrix(NA,nr = n,nc = 3)
#####################################
mat.tr.DS.Sigma1 = matrix(NA,nr = n,nc = 3)
```

```
mat.tr.DS.Sigma2 = matrix(NA,nr = n,nc = 3)
mat.tr.DS.Sigma3 = matrix(NA,nr = n,nc = 3)
mat.tr.DS.Sigma4 = matrix(NA,nr = n,nc = 3)
####################################
mat.tr.LW.Sigma1 = matrix(NA,nr = n,nc = 3)
mat.tr.LW.Sigma2 = matrix(NA,nr = n,nc = 3)
mat.tr.LW.Sigma3 = matrix(NA,nr = n,nc = 3)
mat.tr.LW.Sigma4 = matrix(NA,nr = n,nc = 3)
####################################
mat.tr.RBLW.Sigma1 = matrix(NA,nr = n,nc = 3)
mat.tr.RBLW.Sigma2 = matrix(NA,nr = n,nc = 3)
mat.tr.RBLW.Sigma3 = matrix(NA,nr = n,nc = 3)
mat.tr.RBLW.Sigma4 = matrix(NA,nr = n,nc = 3)
####################################
mat.tr.OAS.Sigma1 = matrix(NA,nr = n,nc = 3)
mat.tr.OAS.Sigma2 = matrix(NA,nr = n,nc = 3)
mat.tr.OAS.Sigma3 = matrix(NA,nr = n,nc = 3)
mat.tr.OAS.Sigma4 = matrix(NA,nr = n,nc = 3)
####################################
mat.tr.FS.Sigma1 = matrix(NA,nr = n,nc = 3)
mat.tr.FS.Sigma2 = matrix(NA,nr = n,nc = 3)
mat.tr.FS.Sigma3 = matrix(NA,nr = n,nc = 3)
mat.tr.FS.Sigma4 = matrix(NA,nr = n,nc = 3)
####################################
#Sturctures of Sigma
U.rand <- runif(p,min = 0, max = 1)
D.vec <- 0.1 + 10*U.rand
D.mat <- diag(D.vec)
# Sigma matrix - structure 1 - Model A-D
# Sigma1 <- matrix(0,p,p)
# for(i in 1:p){
#   for(j in 1:p){
#     Sigma1[i,j] = D.vec[i]*D.vec[j]*rho^(abs(i-j))
#   }
# }
# write.matrix(Sigma1, file =  "/home/mina_norouzirad/results/Sigma1.txt", sep = " ")
Sigma1 <-as.matrix(read.table("/home/mina_norouzirad/results/Sigma1.txt"))
# Sigma matrix - structure 2 - Model A-I
Sigma2 <- matrix(0,p,p)
for(i in 1:p){
for(j in 1:p){
Sigma2[i,j] = rho^(abs(i-j))
}
}
# Sigma matrix - structure 3 - Model B-D
# Sigma3 <- matrix(0,p,p)
# for(i in 1:p){
#   for(j in 1:p){
```

```
#     Sigma3[i,j] = D.vec[i]*D.vec[j]*((abs(i - j +1)^(2*h)) - 2*(abs(i-j)^(2*h))
+ (abs(i - j - 1)^(2*h)))/2
#   }
# }
# write.matrix(Sigma3, file =  "/home/mina_norouzirad/results/Sigma3.txt", sep = " ")
Sigma3 <-as.matrix(read.table("/home/mina_norouzirad/results/Sigma3.txt"))
# Sigma matrix - structure 4 - Model B-I
Sigma4 <- matrix(0,p,p)
for(i in 1:p){
for(j in 1:p){
Sigma4[i,j] = ((abs(i - j +1)^(2*h)) - 2*(abs(i-j)^(2*h)) + (abs(i - j - 1)^(2*h)))/2
}
}
#Simulation loop
for(i.n in 2501:5000){
#Case 1 - Normal distribution
U1 <- matrix(rnorm(N*p,0,1),nr = N, nc = p)
X1 <- matrix(0, nr = N, nc = p)
for(i.1 in 1:N){
X1[i.1,] <- sqrtm(Sigma1)%*%U1[i.1,]
}
#Case 2 - t distribution with m degrees of freedom
Z2 <- matrix(rt(N*p,m),nr = N, nc = p)
X2 <- matrix(0, nr = N, nc = p)
U2 = sqrt((m-2)/m) * Z2
for(i.2 in 1:N){
X2[i.2,] <- sqrtm(Sigma1)%*%U2[i.2,]
}
#Case 3 - Chi square distribution with nu degrees of freedom
Z3 <- matrix(rchisq(N*p,m,df = nu),nr = N, nc = p)
X3 <- matrix(0, nr = N, nc = p)
U3 = ((Z3 - nu)/(sqrt(2*nu)))
for(i.3 in 1:N){
X3[i.3,] <- sqrtm(Sigma1)%*%U3[i.3,]
}
###################################
S1 = cov(X1)
S2 = cov(X2)
S3 = cov(X3)
###################################
###################################
S1.2 = S1%*%S1
S2.2 = S2%*%S2
S3.2 = S3%*%S3
###################################
###################################
a1.hat.1 <- matrix.trace(S1)/p
a1.hat.2 <- matrix.trace(S2)/p
```

```
a1.hat.3 <- matrix.trace(S3)/p
################################
################################
X1.bar = apply(X1,2,mean)
X2.bar = apply(X2,2,mean)
X3.bar = apply(X3,2,mean)
################################
sum.1 = sum.2 = sum.3 = 0
for(i.N in 1:N){
sum.1 = sum.1 + ((t(X1[i.N,]-X1.bar)%*%(X1[i.N,]-X1.bar))^2)
sum.2 = sum.2 + ((t(X2[i.N,]-X2.bar)%*%(X2[i.N,]-X2.bar))^2)
sum.3 = sum.3 + ((t(X3[i.N,]-X3.bar)%*%(X3[i.N,]-X3.bar))^2)
}
################################
################################
Q.1 = sum.1/(N-1)
Q.2 = sum.2/(N-1)
Q.3 = sum.3/(N-1)
################################
################################
a2C.hat.1 <- ((N-1)/(N*(N-2)*(N-3)))*((N-1)*(N-2)*matrix.trace(S1.2)/p + p*a1.hat.1^2 - N*Q.1/p)
a2C.hat.2 <- ((N-1)/(N*(N-2)*(N-3)))*((N-1)*(N-2)*matrix.trace(S2.2)/p + p*a1.hat.2^2 - N*Q.2/p)
a2C.hat.3 <- ((N-1)/(N*(N-2)*(N-3)))*((N-1)*(N-2)*matrix.trace(S3.2)/p + p*a1.hat.3^2 - N*Q.3/p)
################################
################################
#U1 method
################################
################################
w.U1.1 = max(0,min(1,((matrix.trace(S1.2)/p - a2C.hat.1)/(matrix.trace(S1.2)/p - (a1.hat.1^2)))))
w.U1.2 = max(0,min(1,((matrix.trace(S2.2)/p - a2C.hat.2)/(matrix.trace(S2.2)/p - (a1.hat.2^2)))))
w.U1.3 = max(0,min(1,((matrix.trace(S3.2)/p - a2C.hat.3)/(matrix.trace(S3.2)/p - (a1.hat.3^2)))))
################################
################################
Sigma.U1.1 = (1-w.U1.1)*S1 + w.U1.1*a1.hat.1*diag(1,p,p)
Sigma.U1.2 = (1-w.U1.2)*S2 + w.U1.2*a1.hat.2*diag(1,p,p)
Sigma.U1.3 = (1-w.U1.3)*S3 + w.U1.3*a1.hat.3*diag(1,p,p)
################################
a20.hat.K4.1 <- (-1/((N-2)*(N-3)*p))*(2*((N-1)^2)*matrix.trace(S1.2)
+ ((N-1)^2)*((matrix.trace(S1.2)^2)) - (N*(N+1)*Q.1))
a20.hat.K4.2 <- (-1/((N-2)*(N-3)*p))*(2*((N-1)^2)*matrix.trace(S2.2)
+ ((N-1)^2)*((matrix.trace(S2.2)^2)) - (N*(N+1)*Q.2))
a20.hat.K4.3 <- (-1/((N-2)*(N-3)*p))*(2*((N-1)^2)*matrix.trace(S3.2)
+ ((N-1)^2)*((matrix.trace(S3.2)^2)) - (N*(N+1)*Q.3))
################################
################################
D.S.1 = D.S.2 = D.S.3 = matrix(0, nr = p, nc = p)
diag(D.S.1) = diag(S1)
diag(D.S.2) = diag(S2)
```

```
diag(D.S.3) = diag(S3)
################################
################################
a20.hat.1 <- ((N-1)/((N+1)*p))*matrix.trace(D.S.1%^%2) - a20.hat.K4.1/(N+1)
a20.hat.2 <- ((N-1)/((N+1)*p))*matrix.trace(D.S.2%^%2) - a20.hat.K4.2/(N+1)
a20.hat.3 <- ((N-1)/((N+1)*p))*matrix.trace(D.S.3%^%2) - a20.hat.K4.3/(N+1)
################################
################################
#U2 method
################################
w.U2.1 = max(0,min(1,((matrix.trace(S1.2)/p - matrix.trace(S1%*%D.S.1)/p - a2C.hat.1 +
 a20.hat.1)/(matrix.trace(S1.2)/p - (a1.hat.1^2)))))
w.U2.2 = max(0,min(1,((matrix.trace(S2.2)/p - matrix.trace(S2%*%D.S.2)/p - a2C.hat.2 +
 a20.hat.2)/(matrix.trace(S2.2)/p - (a1.hat.2^2)))))
w.U2.3 = max(0,min(1,((matrix.trace(S3.2)/p - matrix.trace(S3%*%D.S.3)/p - a2C.hat.3 +
 a20.hat.3)/(matrix.trace(S3.2)/p - (a1.hat.3^2)))))
################################
################################
Sigma.U2.1 = (1-w.U2.1)*S1 + w.U2.1*D.S.1
Sigma.U2.2 = (1-w.U2.2)*S2 + w.U2.2*D.S.2
Sigma.U2.3 = (1-w.U2.3)*S3 + w.U2.3*D.S.3
################################
################################
mat.1 = matrix(c(matrix.trace((S1-a1.hat.1*diag(1,p,p))%^%2),
matrix.trace((S1-a1.hat.1*diag(1,p,p))%*%(S1-D.S.1)),matrix.trace((S1-a1.hat.1*diag(1,p,p))
%*%(S1-D.S.1)),matrix.trace((S1-D.S.1)%^%2)),nr = 2,nc = 2, byrow = TRUE)
mat.2 = matrix(c(matrix.trace((S2-a1.hat.2*diag(1,p,p))%^%2),
matrix.trace((S2-a1.hat.2*diag(1,p,p))%*%(S2-D.S.2)),matrix.trace((S2-a1.hat.2*diag(1,p,p))
%*%(S2-D.S.2)),matrix.trace((S2-D.S.2)%^%2)),nr = 2,nc = 2, byrow = TRUE)
mat.3 = matrix(c(matrix.trace((S3-a1.hat.3*diag(1,p,p))%^%2),
matrix.trace((S3-a1.hat.3*diag(1,p,p))%*%(S3-D.S.3)),matrix.trace((S3-a1.hat.3*diag(1,p,p))
%*%(S3-D.S.3)),matrix.trace((S3-D.S.3)%^%2)),nr = 2,nc = 2, byrow = TRUE)
################################
vec.1 = matrix(c(matrix.trace(S1.2) - p*a2C.hat.1, matrix.trace(S1.2)-matrix.trace(S1%*%D.S.1)
 - p*(a2C.hat.1 - a20.hat.1)))
vec.2 = matrix(c(matrix.trace(S2.2) - p*a2C.hat.2, matrix.trace(S2.2)-matrix.trace(S2%*%D.S.2)
- p*(a2C.hat.2 - a20.hat.2)))
vec.3 = matrix(c(matrix.trace(S3.2) - p*a2C.hat.3, matrix.trace(S3.2)-matrix.trace(S3%*%D.S.3)
- p*(a2C.hat.3 - a20.hat.3)))
################################
w.star.1 = solve(mat.1)%*%vec.1
w.star.2 = solve(mat.2)%*%vec.2
w.star.3 = solve(mat.3)%*%vec.3
################################
beta.hat.star.1 = w.star.1[1] + w.star.1[2]
beta.hat.star.2 = w.star.2[1] + w.star.2[2]
beta.hat.star.3 = w.star.3[1] + w.star.3[2]
################################
```

```
alpha.hat.star.1 = w.star.1[2]/(w.star.1[1]+w.star.1[2])
alpha.hat.star.2 = w.star.2[2]/(w.star.2[1]+w.star.2[2])
alpha.hat.star.3 = w.star.3[2]/(w.star.3[1]+w.star.3[2])
################################
alpha.hat.1 = max(0,min(alpha.hat.star.1,1))
alpha.hat.2 = max(0,min(alpha.hat.star.2,1))
alpha.hat.3 = max(0,min(alpha.hat.star.3,1))
################################
beta.hat.1 = max(0,min(beta.hat.star.1,1))
beta.hat.2 = max(0,min(beta.hat.star.2,1))
beta.hat.3 = max(0,min(beta.hat.star.3,1))
################################
wD1.hat.1 = (1-alpha.hat.1)*beta.hat.1
wD1.hat.2 = (1-alpha.hat.2)*beta.hat.2
wD1.hat.3 = (1-alpha.hat.3)*beta.hat.3
################################
wD2.hat.1 = alpha.hat.1 * beta.hat.1
wD2.hat.2 = alpha.hat.2 * beta.hat.2
wD2.hat.3 = alpha.hat.3 * beta.hat.3
################################
#Double Shrinkage method
################################
Sigma.DS.1 = (1-wD1.hat.1 - wD2.hat.1)*S1 + wD1.hat.1*a1.hat.1*diag(1,p,p) + wD2.hat.1*D.S.1
Sigma.DS.2 = (1-wD1.hat.2 - wD2.hat.2)*S2 + wD1.hat.2*a1.hat.2*diag(1,p,p) + wD2.hat.2*D.S.2
Sigma.DS.3 = (1-wD1.hat.3 - wD2.hat.3)*S3 + wD1.hat.3*a1.hat.3*diag(1,p,p) + wD2.hat.3*D.S.3
################################
sum.LW.1 = sum.LW.2 = sum.LW.3 = 0
for(i.LW in 1:N){
sum.LW.1 = sum.LW.1 + (matrix.trace(((X1[i.LW,]-X1.bar)%*%t(X1[i.LW,]-X1.bar) - S1)%^%2))
sum.LW.2 = sum.LW.2 + (matrix.trace(((X2[i.LW,]-X2.bar)%*%t(X2[i.LW,]-X2.bar) - S2)%^%2))
sum.LW.3 = sum.LW.3 + (matrix.trace(((X3[i.LW,]-X3.bar)%*%t(X3[i.LW,]-X3.bar) - S3)%^%2))
}
################################
W.LW.1 = max(0,min(1,sum.LW.1 / (((N-1)^2)*(matrix.trace(S1.2)-((matrix.trace(S1))^2)/p))))
W.LW.2 = max(0,min(1,sum.LW.2 / (((N-1)^2)*(matrix.trace(S2.2)-((matrix.trace(S2))^2)/p))))
W.LW.3 = max(0,min(1,sum.LW.3 / (((N-1)^2)*(matrix.trace(S3.2)-((matrix.trace(S3))^2)/p))))
################################
#Ledoit and Wolf method
################################
Sigma.LW.1 = (1-W.LW.1)*S1 + W.LW.1*a1.hat.1*diag(1,p,p)
Sigma.LW.2 = (1-W.LW.2)*S2 + W.LW.2*a1.hat.2*diag(1,p,p)
Sigma.LW.3 = (1-W.LW.3)*S3 + W.LW.3*a1.hat.3*diag(1,p,p)
################################
W.RBLW.1 = max(0,min(1,((N-3)*matrix.trace(S1.2)
+ (N-1)*(matrix.trace(S1))^2) / (((N-1)*(N+1))*(matrix.trace(S1.2)-((matrix.trace(S1))^2)/p))))
W.RBLW.2 = max(0,min(1,((N-3)*matrix.trace(S2.2)
+ (N-1)*(matrix.trace(S2))^2) / (((N-1)*(N+1))*(matrix.trace(S2.2)-((matrix.trace(S2))^2)/p))))
W.RBLW.3 = max(0,min(1,((N-3)*matrix.trace(S3.2)
```

```
+ (N-1)*(matrix.trace(S3))^2) / (((N-1)*(N+1))*(matrix.trace(S3.2)-((matrix.trace(S3))^2)/p))))
################################
#Rao-Blackwell-Ledoit-Wolf method
################################
Sigma.RBLW.1 = (1-W.RBLW.1)*S1 + W.RBLW.1*a1.hat.1*diag(1,p,p)
Sigma.RBLW.2 = (1-W.RBLW.2)*S2 + W.RBLW.2*a1.hat.2*diag(1,p,p)
Sigma.RBLW.3 = (1-W.RBLW.3)*S3 + W.RBLW.3*a1.hat.3*diag(1,p,p)
################################
W.OAS.1 = max(0,min(1,((p-2)*matrix.trace(S1.2) +
p*(matrix.trace(S1))^2) / (p*(N-2/p)*(matrix.trace(S1.2)-((matrix.trace(S1))^2)/p))))
W.OAS.2 = max(0,min(1,((p-2)*matrix.trace(S2.2)
+ p*(matrix.trace(S2))^2) / (p*(N-2/p)*(matrix.trace(S2.2)-((matrix.trace(S2))^2)/p))))
W.OAS.3 = max(0,min(1,((p-2)*matrix.trace(S3.2)
+ p*(matrix.trace(S3))^2) / (p*(N-2/p)*(matrix.trace(S3.2)-((matrix.trace(S3))^2)/p))))
################################
#Oracle-Approximating Shrinkage method
################################
Sigma.OAS.1 = (1-W.OAS.1)*S1 + W.OAS.1*a1.hat.1*diag(1,p,p)
Sigma.OAS.2 = (1-W.OAS.2)*S2 + W.OAS.2*a1.hat.2*diag(1,p,p)
Sigma.OAS.3 = (1-W.OAS.3)*S3 + W.OAS.3*a1.hat.3*diag(1,p,p)
################################
a2.hat.1 <- (((N-1)^2)/((N-1)*(N+1))) * (matrix.trace(S1.2)/p - (p/(N-1))*(a1.hat.1^2))
a2.hat.2 <- (((N-1)^2)/((N-1)*(N+1))) * (matrix.trace(S2.2)/p - (p/(N-1))*(a1.hat.2^2))
a2.hat.3 <- (((N-1)^2)/((N-1)*(N+1))) * (matrix.trace(S3.2)/p - (p/(N-1))*(a1.hat.3^2))
################################
W.FS.1 = max(0,min(1,(p*a1.hat.1^2 + a2.hat.1)/((N-1)*(a2.hat.1 - a1.hat.1^2)
+ p*a1.hat.1^2 + a2.hat.1)))
W.FS.2 = max(0,min(1,(p*a1.hat.2^2 + a2.hat.2)/((N-1)*(a2.hat.2 - a1.hat.2^2)
+ p*a1.hat.2^2 + a2.hat.2)))
W.FS.3 = max(0,min(1,(p*a1.hat.3^2 + a2.hat.3)/((N-1)*(a2.hat.3 - a1.hat.3^2)
+ p*a1.hat.3^2 + a2.hat.3)))
################################
#Fisher and Sun method
################################
Sigma.FS.1 = (1-W.FS.1)*S1 + W.FS.1*a1.hat.1*diag(1,p,p)
Sigma.FS.2 = (1-W.FS.2)*S2 + W.FS.2*a1.hat.2*diag(1,p,p)
Sigma.FS.3 = (1-W.FS.3)*S3 + W.FS.3*a1.hat.3*diag(1,p,p)
################################
mat.tr.S.Sigma1[i.n,1] <- matrix.trace((S1-Sigma1)%*%t((S1-Sigma1)))
mat.tr.S.Sigma2[i.n,1] <- matrix.trace((S1-Sigma2)%*%t((S1-Sigma2)))
mat.tr.S.Sigma3[i.n,1] <- matrix.trace((S1-Sigma3)%*%t((S1-Sigma3)))
mat.tr.S.Sigma4[i.n,1] <- matrix.trace((S1-Sigma4)%*%t((S1-Sigma4)))
################################
mat.tr.S.Sigma1[i.n,2] <- matrix.trace((S2-Sigma1)%*%t((S2-Sigma1)))
mat.tr.S.Sigma2[i.n,2] <- matrix.trace((S2-Sigma2)%*%t((S2-Sigma2)))
mat.tr.S.Sigma3[i.n,2] <- matrix.trace((S2-Sigma3)%*%t((S2-Sigma3)))
mat.tr.S.Sigma4[i.n,2] <- matrix.trace((S2-Sigma4)%*%t((S2-Sigma4)))
################################
```

```
mat.tr.S.Sigma1[i.n,3] <- matrix.trace((S3-Sigma1)%*%t((S3-Sigma1)))
mat.tr.S.Sigma2[i.n,3] <- matrix.trace((S3-Sigma2)%*%t((S3-Sigma2)))
mat.tr.S.Sigma3[i.n,3] <- matrix.trace((S3-Sigma3)%*%t((S3-Sigma3)))
mat.tr.S.Sigma4[i.n,3] <- matrix.trace((S3-Sigma4)%*%t((S3-Sigma4)))
################################
################################
mat.tr.U1.Sigma1[i.n,1] = matrix.trace((Sigma.U1.1-Sigma1)%*%t(Sigma.U1.1-Sigma1))
mat.tr.U1.Sigma2[i.n,1] = matrix.trace((Sigma.U1.1-Sigma2)%*%t(Sigma.U1.1-Sigma2))
mat.tr.U1.Sigma3[i.n,1] = matrix.trace((Sigma.U1.1-Sigma3)%*%t(Sigma.U1.1-Sigma3))
mat.tr.U1.Sigma4[i.n,1] = matrix.trace((Sigma.U1.1-Sigma4)%*%t(Sigma.U1.1-Sigma4))
################################
mat.tr.U1.Sigma1[i.n,2] = matrix.trace((Sigma.U1.2-Sigma1)%*%t(Sigma.U1.2-Sigma1))
mat.tr.U1.Sigma2[i.n,2] = matrix.trace((Sigma.U1.2-Sigma2)%*%t(Sigma.U1.2-Sigma2))
mat.tr.U1.Sigma3[i.n,2] = matrix.trace((Sigma.U1.2-Sigma3)%*%t(Sigma.U1.2-Sigma3))
mat.tr.U1.Sigma4[i.n,2] = matrix.trace((Sigma.U1.2-Sigma4)%*%t(Sigma.U1.2-Sigma4))
################################
mat.tr.U1.Sigma1[i.n,3] = matrix.trace((Sigma.U1.3-Sigma1)%*%t(Sigma.U1.3-Sigma1))
mat.tr.U1.Sigma2[i.n,3] = matrix.trace((Sigma.U1.3-Sigma2)%*%t(Sigma.U1.3-Sigma2))
mat.tr.U1.Sigma3[i.n,3] = matrix.trace((Sigma.U1.3-Sigma3)%*%t(Sigma.U1.3-Sigma3))
mat.tr.U1.Sigma4[i.n,3] = matrix.trace((Sigma.U1.3-Sigma4)%*%t(Sigma.U1.3-Sigma4))
################################
################################
mat.tr.U2.Sigma1[i.n,1] = matrix.trace((Sigma.U2.1-Sigma1)%*%t(Sigma.U2.1-Sigma1))
mat.tr.U2.Sigma2[i.n,1] = matrix.trace((Sigma.U2.1-Sigma2)%*%t(Sigma.U2.1-Sigma2))
mat.tr.U2.Sigma3[i.n,1] = matrix.trace((Sigma.U2.1-Sigma3)%*%t(Sigma.U2.1-Sigma3))
mat.tr.U2.Sigma4[i.n,1] = matrix.trace((Sigma.U2.1-Sigma4)%*%t(Sigma.U2.1-Sigma4))
################################
mat.tr.U2.Sigma1[i.n,2] = matrix.trace((Sigma.U2.2-Sigma1)%*%t(Sigma.U2.2-Sigma1))
mat.tr.U2.Sigma2[i.n,2] = matrix.trace((Sigma.U2.2-Sigma2)%*%t(Sigma.U2.2-Sigma2))
mat.tr.U2.Sigma3[i.n,2] = matrix.trace((Sigma.U2.2-Sigma3)%*%t(Sigma.U2.2-Sigma3))
mat.tr.U2.Sigma4[i.n,2] = matrix.trace((Sigma.U2.2-Sigma4)%*%t(Sigma.U2.2-Sigma4))
################################
mat.tr.U2.Sigma1[i.n,3] = matrix.trace((Sigma.U2.3-Sigma1)%*%t(Sigma.U2.3-Sigma1))
mat.tr.U2.Sigma2[i.n,3] = matrix.trace((Sigma.U2.3-Sigma2)%*%t(Sigma.U2.3-Sigma2))
mat.tr.U2.Sigma3[i.n,3] = matrix.trace((Sigma.U2.3-Sigma3)%*%t(Sigma.U2.3-Sigma3))
mat.tr.U2.Sigma4[i.n,3] = matrix.trace((Sigma.U2.3-Sigma4)%*%t(Sigma.U2.3-Sigma4))
################################
################################
mat.tr.DS.Sigma1[i.n,1] = matrix.trace((Sigma.DS.1-Sigma1)%*%t(Sigma.DS.1-Sigma1))
mat.tr.DS.Sigma2[i.n,1] = matrix.trace((Sigma.DS.1-Sigma2)%*%t(Sigma.DS.1-Sigma2))
mat.tr.DS.Sigma3[i.n,1] = matrix.trace((Sigma.DS.1-Sigma3)%*%t(Sigma.DS.1-Sigma3))
mat.tr.DS.Sigma4[i.n,1] = matrix.trace((Sigma.DS.1-Sigma4)%*%t(Sigma.DS.1-Sigma4))
################################
mat.tr.DS.Sigma1[i.n,2] = matrix.trace((Sigma.DS.2-Sigma1)%*%t(Sigma.DS.2-Sigma1))
mat.tr.DS.Sigma2[i.n,2] = matrix.trace((Sigma.DS.2-Sigma2)%*%t(Sigma.DS.2-Sigma2))
mat.tr.DS.Sigma3[i.n,2] = matrix.trace((Sigma.DS.2-Sigma3)%*%t(Sigma.DS.2-Sigma3))
mat.tr.DS.Sigma4[i.n,2] = matrix.trace((Sigma.DS.2-Sigma4)%*%t(Sigma.DS.2-Sigma4))
################################
```

```
mat.tr.DS.Sigma1[i.n,3] = matrix.trace((Sigma.DS.3-Sigma1)%*%t(Sigma.DS.3-Sigma1))
mat.tr.DS.Sigma2[i.n,3] = matrix.trace((Sigma.DS.3-Sigma2)%*%t(Sigma.DS.3-Sigma2))
mat.tr.DS.Sigma3[i.n,3] = matrix.trace((Sigma.DS.3-Sigma3)%*%t(Sigma.DS.3-Sigma3))
mat.tr.DS.Sigma4[i.n,3] = matrix.trace((Sigma.DS.3-Sigma4)%*%t(Sigma.DS.3-Sigma4))
################################
################################
mat.tr.LW.Sigma1[i.n,1] = matrix.trace((Sigma.LW.1-Sigma1)%*%t(Sigma.LW.1-Sigma1))
mat.tr.LW.Sigma2[i.n,1] = matrix.trace((Sigma.LW.1-Sigma2)%*%t(Sigma.LW.1-Sigma2))
mat.tr.LW.Sigma3[i.n,1] = matrix.trace((Sigma.LW.1-Sigma3)%*%t(Sigma.LW.1-Sigma3))
mat.tr.LW.Sigma4[i.n,1] = matrix.trace((Sigma.LW.1-Sigma4)%*%t(Sigma.LW.1-Sigma4))
################################
mat.tr.LW.Sigma1[i.n,2] = matrix.trace((Sigma.LW.2-Sigma1)%*%t(Sigma.LW.2-Sigma1))
mat.tr.LW.Sigma2[i.n,2] = matrix.trace((Sigma.LW.2-Sigma2)%*%t(Sigma.LW.2-Sigma2))
mat.tr.LW.Sigma3[i.n,2] = matrix.trace((Sigma.LW.2-Sigma3)%*%t(Sigma.LW.2-Sigma3))
mat.tr.LW.Sigma4[i.n,2] = matrix.trace((Sigma.LW.2-Sigma4)%*%t(Sigma.LW.2-Sigma4))
################################
mat.tr.LW.Sigma1[i.n,3] = matrix.trace((Sigma.LW.3-Sigma1)%*%t(Sigma.LW.3-Sigma1))
mat.tr.LW.Sigma2[i.n,3] = matrix.trace((Sigma.LW.3-Sigma2)%*%t(Sigma.LW.3-Sigma2))
mat.tr.LW.Sigma3[i.n,3] = matrix.trace((Sigma.LW.3-Sigma3)%*%t(Sigma.LW.3-Sigma3))
mat.tr.LW.Sigma4[i.n,3] = matrix.trace((Sigma.LW.3-Sigma4)%*%t(Sigma.LW.3-Sigma4))
################################
################################
mat.tr.RBLW.Sigma1[i.n,1] = matrix.trace((Sigma.RBLW.1-Sigma1)%*%t(Sigma.RBLW.1-Sigma1))
mat.tr.RBLW.Sigma2[i.n,1] = matrix.trace((Sigma.RBLW.1-Sigma2)%*%t(Sigma.RBLW.1-Sigma2))
mat.tr.RBLW.Sigma3[i.n,1] = matrix.trace((Sigma.RBLW.1-Sigma3)%*%t(Sigma.RBLW.1-Sigma3))
mat.tr.RBLW.Sigma4[i.n,1] = matrix.trace((Sigma.RBLW.1-Sigma4)%*%t(Sigma.RBLW.1-Sigma4))
################################
mat.tr.RBLW.Sigma1[i.n,2] = matrix.trace((Sigma.RBLW.2-Sigma1)%*%t(Sigma.RBLW.2-Sigma1))
mat.tr.RBLW.Sigma2[i.n,2] = matrix.trace((Sigma.RBLW.2-Sigma2)%*%t(Sigma.RBLW.2-Sigma2))
mat.tr.RBLW.Sigma3[i.n,2] = matrix.trace((Sigma.RBLW.2-Sigma3)%*%t(Sigma.RBLW.2-Sigma3))
mat.tr.RBLW.Sigma4[i.n,2] = matrix.trace((Sigma.RBLW.2-Sigma4)%*%t(Sigma.RBLW.2-Sigma4))
################################
mat.tr.RBLW.Sigma1[i.n,3] = matrix.trace((Sigma.RBLW.3-Sigma1)%*%t(Sigma.RBLW.3-Sigma1))
mat.tr.RBLW.Sigma2[i.n,3] = matrix.trace((Sigma.RBLW.3-Sigma2)%*%t(Sigma.RBLW.3-Sigma2))
mat.tr.RBLW.Sigma3[i.n,3] = matrix.trace((Sigma.RBLW.3-Sigma3)%*%t(Sigma.RBLW.3-Sigma3))
mat.tr.RBLW.Sigma4[i.n,3] = matrix.trace((Sigma.RBLW.3-Sigma4)%*%t(Sigma.RBLW.3-Sigma4))
################################
################################
mat.tr.OAS.Sigma1[i.n,1] = matrix.trace((Sigma.OAS.1-Sigma1)%*%t(Sigma.OAS.1-Sigma1))
mat.tr.OAS.Sigma2[i.n,1] = matrix.trace((Sigma.OAS.1-Sigma2)%*%t(Sigma.OAS.1-Sigma2))
mat.tr.OAS.Sigma3[i.n,1] = matrix.trace((Sigma.OAS.1-Sigma3)%*%t(Sigma.OAS.1-Sigma3))
mat.tr.OAS.Sigma4[i.n,1] = matrix.trace((Sigma.OAS.1-Sigma4)%*%t(Sigma.OAS.1-Sigma4))
################################
mat.tr.OAS.Sigma1[i.n,2] = matrix.trace((Sigma.OAS.2-Sigma1)%*%t(Sigma.OAS.2-Sigma1))
mat.tr.OAS.Sigma2[i.n,2] = matrix.trace((Sigma.OAS.2-Sigma2)%*%t(Sigma.OAS.2-Sigma2))
mat.tr.OAS.Sigma3[i.n,2] = matrix.trace((Sigma.OAS.2-Sigma3)%*%t(Sigma.OAS.2-Sigma3))
mat.tr.OAS.Sigma4[i.n,2] = matrix.trace((Sigma.OAS.2-Sigma4)%*%t(Sigma.OAS.2-Sigma4))
################################
```

```
mat.tr.OAS.Sigma1[i.n,3] = matrix.trace((Sigma.OAS.3-Sigma1)%*%t(Sigma.OAS.3-Sigma1))
mat.tr.OAS.Sigma2[i.n,3] = matrix.trace((Sigma.OAS.3-Sigma2)%*%t(Sigma.OAS.3-Sigma2))
mat.tr.OAS.Sigma3[i.n,3] = matrix.trace((Sigma.OAS.3-Sigma3)%*%t(Sigma.OAS.3-Sigma3))
mat.tr.OAS.Sigma4[i.n,3] = matrix.trace((Sigma.OAS.3-Sigma4)%*%t(Sigma.OAS.3-Sigma4))
################################
################################
mat.tr.FS.Sigma1[i.n,1] = matrix.trace((Sigma.FS.1-Sigma1)%*%t(Sigma.FS.1-Sigma1))
mat.tr.FS.Sigma2[i.n,1] = matrix.trace((Sigma.FS.1-Sigma2)%*%t(Sigma.FS.1-Sigma2))
mat.tr.FS.Sigma3[i.n,1] = matrix.trace((Sigma.FS.1-Sigma3)%*%t(Sigma.FS.1-Sigma3))
mat.tr.FS.Sigma4[i.n,1] = matrix.trace((Sigma.FS.1-Sigma4)%*%t(Sigma.FS.1-Sigma4))
################################
mat.tr.FS.Sigma1[i.n,2] = matrix.trace((Sigma.FS.2-Sigma1)%*%t(Sigma.FS.2-Sigma1))
mat.tr.FS.Sigma2[i.n,2] = matrix.trace((Sigma.FS.2-Sigma2)%*%t(Sigma.FS.2-Sigma2))
mat.tr.FS.Sigma3[i.n,2] = matrix.trace((Sigma.FS.2-Sigma3)%*%t(Sigma.FS.2-Sigma3))
mat.tr.FS.Sigma4[i.n,2] = matrix.trace((Sigma.FS.2-Sigma4)%*%t(Sigma.FS.2-Sigma4))
################################
mat.tr.FS.Sigma1[i.n,3] = matrix.trace((Sigma.FS.3-Sigma1)%*%t(Sigma.FS.3-Sigma1))
mat.tr.FS.Sigma2[i.n,3] = matrix.trace((Sigma.FS.3-Sigma2)%*%t(Sigma.FS.3-Sigma2))
mat.tr.FS.Sigma3[i.n,3] = matrix.trace((Sigma.FS.3-Sigma3)%*%t(Sigma.FS.3-Sigma3))
mat.tr.FS.Sigma4[i.n,3] = matrix.trace((Sigma.FS.3-Sigma4)%*%t(Sigma.FS.3-Sigma4))

}

#PRIAL
mean.tr.S.Sigma1 <- apply(mat.tr.S.Sigma1,2,mean)
mean.tr.S.Sigma2 <- apply(mat.tr.S.Sigma2,2,mean)
mean.tr.S.Sigma3 <- apply(mat.tr.S.Sigma3,2,mean)
mean.tr.S.Sigma4 <- apply(mat.tr.S.Sigma4,2,mean)
################################################
mean.tr.U1.Sigma1 <- apply(mat.tr.U1.Sigma1,2,mean)
mean.tr.U1.Sigma2 <- apply(mat.tr.U1.Sigma2,2,mean)
mean.tr.U1.Sigma3 <- apply(mat.tr.U1.Sigma3,2,mean)
mean.tr.U1.Sigma4 <- apply(mat.tr.U1.Sigma4,2,mean)
################################################
mean.tr.U2.Sigma1 <- apply(mat.tr.U2.Sigma1,2,mean)
mean.tr.U2.Sigma2 <- apply(mat.tr.U2.Sigma2,2,mean)
mean.tr.U2.Sigma3 <- apply(mat.tr.U2.Sigma3,2,mean)
mean.tr.U2.Sigma4 <- apply(mat.tr.U2.Sigma4,2,mean)
################################################
mean.tr.DS.Sigma1 <- apply(mat.tr.DS.Sigma1,2,mean)
mean.tr.DS.Sigma2 <- apply(mat.tr.DS.Sigma2,2,mean)
mean.tr.DS.Sigma3 <- apply(mat.tr.DS.Sigma3,2,mean)
mean.tr.DS.Sigma4 <- apply(mat.tr.DS.Sigma4,2,mean)
################################################
mean.tr.LW.Sigma1 <- apply(mat.tr.LW.Sigma1,2,mean)
mean.tr.LW.Sigma2 <- apply(mat.tr.LW.Sigma2,2,mean)
mean.tr.LW.Sigma3 <- apply(mat.tr.LW.Sigma3,2,mean)
mean.tr.LW.Sigma4 <- apply(mat.tr.LW.Sigma4,2,mean)
```

```
#################################################
mean.tr.RBLW.Sigma1 <- apply(mat.tr.RBLW.Sigma1,2,mean)
mean.tr.RBLW.Sigma2 <- apply(mat.tr.RBLW.Sigma2,2,mean)
mean.tr.RBLW.Sigma3 <- apply(mat.tr.RBLW.Sigma3,2,mean)
mean.tr.RBLW.Sigma4 <- apply(mat.tr.RBLW.Sigma4,2,mean)
#########################################################
mean.tr.OAS.Sigma1 <- apply(mat.tr.OAS.Sigma1,2,mean)
mean.tr.OAS.Sigma2 <- apply(mat.tr.OAS.Sigma2,2,mean)
mean.tr.OAS.Sigma3 <- apply(mat.tr.OAS.Sigma3,2,mean)
mean.tr.OAS.Sigma4 <- apply(mat.tr.OAS.Sigma4,2,mean)
#########################################################
mean.tr.FS.Sigma1 <- apply(mat.tr.FS.Sigma1,2,mean)
mean.tr.FS.Sigma2 <- apply(mat.tr.FS.Sigma2,2,mean)
mean.tr.FS.Sigma3 <- apply(mat.tr.FS.Sigma3,2,mean)
mean.tr.FS.Sigma4 <- apply(mat.tr.FS.Sigma4,2,mean)
#########################################################
PRIAL.Sigma1.case1.U1 <- 100*(mean.tr.S.Sigma1[1]-mean.tr.U1.Sigma1[1])/mean.tr.S.Sigma1[1]
PRIAL.Sigma1.case2.U1 <- 100*(mean.tr.S.Sigma1[2]-mean.tr.U1.Sigma1[2])/mean.tr.S.Sigma1[2]
PRIAL.Sigma1.case3.U1 <- 100*(mean.tr.S.Sigma1[3]-mean.tr.U1.Sigma1[3])/mean.tr.S.Sigma1[3]
#########################################################
PRIAL.Sigma2.case1.U1 <- 100*(mean.tr.S.Sigma2[1]-mean.tr.U1.Sigma2[1])/mean.tr.S.Sigma2[1]
PRIAL.Sigma2.case2.U1 <- 100*(mean.tr.S.Sigma2[2]-mean.tr.U1.Sigma2[2])/mean.tr.S.Sigma2[2]
PRIAL.Sigma2.case3.U1 <- 100*(mean.tr.S.Sigma2[3]-mean.tr.U1.Sigma2[3])/mean.tr.S.Sigma2[3]
#########################################################
PRIAL.Sigma3.case1.U1 <- 100*(mean.tr.S.Sigma3[1]-mean.tr.U1.Sigma3[1])/mean.tr.S.Sigma3[1]
PRIAL.Sigma3.case2.U1 <- 100*(mean.tr.S.Sigma3[2]-mean.tr.U1.Sigma3[2])/mean.tr.S.Sigma3[2]
PRIAL.Sigma3.case3.U1 <- 100*(mean.tr.S.Sigma3[3]-mean.tr.U1.Sigma3[3])/mean.tr.S.Sigma3[3]
#########################################################
PRIAL.Sigma4.case1.U1 <- 100*(mean.tr.S.Sigma4[1]-mean.tr.U1.Sigma4[1])/mean.tr.S.Sigma4[1]
PRIAL.Sigma4.case2.U1 <- 100*(mean.tr.S.Sigma4[2]-mean.tr.U1.Sigma4[2])/mean.tr.S.Sigma4[2]
PRIAL.Sigma4.case3.U1 <- 100*(mean.tr.S.Sigma4[3]-mean.tr.U1.Sigma4[3])/mean.tr.S.Sigma4[3]
#########################################################################
PRIAL.Sigma1.case1.U2 <- 100*(mean.tr.S.Sigma1[1]-mean.tr.U2.Sigma1[1])/mean.tr.S.Sigma1[1]
PRIAL.Sigma1.case2.U2 <- 100*(mean.tr.S.Sigma1[2]-mean.tr.U2.Sigma1[2])/mean.tr.S.Sigma1[2]
PRIAL.Sigma1.case3.U2 <- 100*(mean.tr.S.Sigma1[3]-mean.tr.U2.Sigma1[3])/mean.tr.S.Sigma1[3]
#########################################################
PRIAL.Sigma2.case1.U2 <- 100*(mean.tr.S.Sigma2[1]-mean.tr.U2.Sigma2[1])/mean.tr.S.Sigma2[1]
PRIAL.Sigma2.case2.U2 <- 100*(mean.tr.S.Sigma2[2]-mean.tr.U2.Sigma2[2])/mean.tr.S.Sigma2[2]
PRIAL.Sigma2.case3.U2 <- 100*(mean.tr.S.Sigma2[3]-mean.tr.U2.Sigma2[3])/mean.tr.S.Sigma2[3]
#########################################################
PRIAL.Sigma3.case1.U2 <- 100*(mean.tr.S.Sigma3[1]-mean.tr.U2.Sigma3[1])/mean.tr.S.Sigma3[1]
PRIAL.Sigma3.case2.U2 <- 100*(mean.tr.S.Sigma3[2]-mean.tr.U2.Sigma3[2])/mean.tr.S.Sigma3[2]
PRIAL.Sigma3.case3.U2 <- 100*(mean.tr.S.Sigma3[3]-mean.tr.U2.Sigma3[3])/mean.tr.S.Sigma3[3]
#########################################################
PRIAL.Sigma4.case1.U2 <- 100*(mean.tr.S.Sigma4[1]-mean.tr.U2.Sigma4[1])/mean.tr.S.Sigma4[1]
PRIAL.Sigma4.case2.U2 <- 100*(mean.tr.S.Sigma4[2]-mean.tr.U2.Sigma4[2])/mean.tr.S.Sigma4[2]
PRIAL.Sigma4.case3.U2 <- 100*(mean.tr.S.Sigma4[3]-mean.tr.U2.Sigma4[3])/mean.tr.S.Sigma4[3]
#########################################################################
```

```
PRIAL.Sigma1.case1.DS <- 100*(mean.tr.S.Sigma1[1]-mean.tr.DS.Sigma1[1])/mean.tr.S.Sigma1[1]
PRIAL.Sigma1.case2.DS <- 100*(mean.tr.S.Sigma1[2]-mean.tr.DS.Sigma1[2])/mean.tr.S.Sigma1[2]
PRIAL.Sigma1.case3.DS <- 100*(mean.tr.S.Sigma1[3]-mean.tr.DS.Sigma1[3])/mean.tr.S.Sigma1[3]
#########################################################
PRIAL.Sigma2.case1.DS <- 100*(mean.tr.S.Sigma2[1]-mean.tr.DS.Sigma2[1])/mean.tr.S.Sigma2[1]
PRIAL.Sigma2.case2.DS <- 100*(mean.tr.S.Sigma2[2]-mean.tr.DS.Sigma2[2])/mean.tr.S.Sigma2[2]
PRIAL.Sigma2.case3.DS <- 100*(mean.tr.S.Sigma2[3]-mean.tr.DS.Sigma2[3])/mean.tr.S.Sigma2[3]
#########################################################
PRIAL.Sigma3.case1.DS <- 100*(mean.tr.S.Sigma3[1]-mean.tr.DS.Sigma3[1])/mean.tr.S.Sigma3[1]
PRIAL.Sigma3.case2.DS <- 100*(mean.tr.S.Sigma3[2]-mean.tr.DS.Sigma3[2])/mean.tr.S.Sigma3[2]
PRIAL.Sigma3.case3.DS <- 100*(mean.tr.S.Sigma3[3]-mean.tr.DS.Sigma3[3])/mean.tr.S.Sigma3[3]
#########################################################
PRIAL.Sigma4.case1.DS <- 100*(mean.tr.S.Sigma4[1]-mean.tr.DS.Sigma4[1])/mean.tr.S.Sigma4[1]
PRIAL.Sigma4.case2.DS <- 100*(mean.tr.S.Sigma4[2]-mean.tr.DS.Sigma4[2])/mean.tr.S.Sigma4[2]
PRIAL.Sigma4.case3.DS <- 100*(mean.tr.S.Sigma4[3]-mean.tr.DS.Sigma4[3])/mean.tr.S.Sigma4[3]
#############################################################################
#############################################################################
PRIAL.Sigma1.case1.LW <- 100*(mean.tr.S.Sigma1[1]-mean.tr.LW.Sigma1[1])/mean.tr.S.Sigma1[1]
PRIAL.Sigma1.case2.LW <- 100*(mean.tr.S.Sigma1[2]-mean.tr.LW.Sigma1[2])/mean.tr.S.Sigma1[2]
PRIAL.Sigma1.case3.LW <- 100*(mean.tr.S.Sigma1[3]-mean.tr.LW.Sigma1[3])/mean.tr.S.Sigma1[3]
#########################################################
PRIAL.Sigma2.case1.LW <- 100*(mean.tr.S.Sigma2[1]-mean.tr.LW.Sigma2[1])/mean.tr.S.Sigma2[1]
PRIAL.Sigma2.case2.LW <- 100*(mean.tr.S.Sigma2[2]-mean.tr.LW.Sigma2[2])/mean.tr.S.Sigma2[2]
PRIAL.Sigma2.case3.LW <- 100*(mean.tr.S.Sigma2[3]-mean.tr.LW.Sigma2[3])/mean.tr.S.Sigma2[3]
#########################################################
PRIAL.Sigma3.case1.LW <- 100*(mean.tr.S.Sigma3[1]-mean.tr.LW.Sigma3[1])/mean.tr.S.Sigma3[1]
PRIAL.Sigma3.case2.LW <- 100*(mean.tr.S.Sigma3[2]-mean.tr.LW.Sigma3[2])/mean.tr.S.Sigma3[2]
PRIAL.Sigma3.case3.LW <- 100*(mean.tr.S.Sigma3[3]-mean.tr.LW.Sigma3[3])/mean.tr.S.Sigma3[3]
#########################################################
PRIAL.Sigma4.case1.LW <- 100*(mean.tr.S.Sigma4[1]-mean.tr.LW.Sigma4[1])/mean.tr.S.Sigma4[1]
PRIAL.Sigma4.case2.LW <- 100*(mean.tr.S.Sigma4[2]-mean.tr.LW.Sigma4[2])/mean.tr.S.Sigma4[2]
PRIAL.Sigma4.case3.LW <- 100*(mean.tr.S.Sigma4[3]-mean.tr.LW.Sigma4[3])/mean.tr.S.Sigma4[3]
#############################################################################
#############################################################################
PRIAL.Sigma1.case1.RBLW <- 100*(mean.tr.S.Sigma1[1]-mean.tr.RBLW.Sigma1[1])/mean.tr.S.Sigma1[1]
PRIAL.Sigma1.case2.RBLW <- 100*(mean.tr.S.Sigma1[2]-mean.tr.RBLW.Sigma1[2])/mean.tr.S.Sigma1[2]
PRIAL.Sigma1.case3.RBLW <- 100*(mean.tr.S.Sigma1[3]-mean.tr.RBLW.Sigma1[3])/mean.tr.S.Sigma1[3]
#########################################################
PRIAL.Sigma2.case1.RBLW <- 100*(mean.tr.S.Sigma2[1]-mean.tr.RBLW.Sigma2[1])/mean.tr.S.Sigma2[1]
PRIAL.Sigma2.case2.RBLW <- 100*(mean.tr.S.Sigma2[2]-mean.tr.RBLW.Sigma2[2])/mean.tr.S.Sigma2[2]
PRIAL.Sigma2.case3.RBLW <- 100*(mean.tr.S.Sigma2[3]-mean.tr.RBLW.Sigma2[3])/mean.tr.S.Sigma2[3]
#########################################################
PRIAL.Sigma3.case1.RBLW <- 100*(mean.tr.S.Sigma3[1]-mean.tr.RBLW.Sigma3[1])/mean.tr.S.Sigma3[1]
PRIAL.Sigma3.case2.RBLW <- 100*(mean.tr.S.Sigma3[2]-mean.tr.RBLW.Sigma3[2])/mean.tr.S.Sigma3[2]
PRIAL.Sigma3.case3.RBLW <- 100*(mean.tr.S.Sigma3[3]-mean.tr.RBLW.Sigma3[3])/mean.tr.S.Sigma3[3]
#########################################################
PRIAL.Sigma4.case1.RBLW <- 100*(mean.tr.S.Sigma4[1]-mean.tr.RBLW.Sigma4[1])/mean.tr.S.Sigma4[1]
PRIAL.Sigma4.case2.RBLW <- 100*(mean.tr.S.Sigma4[2]-mean.tr.RBLW.Sigma4[2])/mean.tr.S.Sigma4[2]
```

```
PRIAL.Sigma4.case3.RBLW <- 100*(mean.tr.S.Sigma4[3]-mean.tr.RBLW.Sigma4[3])/mean.tr.S.Sigma4[3]
##############################################################################
##############################################################################
PRIAL.Sigma1.case1.OAS <- 100*(mean.tr.S.Sigma1[1]-mean.tr.OAS.Sigma1[1])/mean.tr.S.Sigma1[1]
PRIAL.Sigma1.case2.OAS <- 100*(mean.tr.S.Sigma1[2]-mean.tr.OAS.Sigma1[2])/mean.tr.S.Sigma1[2]
PRIAL.Sigma1.case3.OAS <- 100*(mean.tr.S.Sigma1[3]-mean.tr.OAS.Sigma1[3])/mean.tr.S.Sigma1[3]
#########################################################
PRIAL.Sigma2.case1.OAS <- 100*(mean.tr.S.Sigma2[1]-mean.tr.OAS.Sigma2[1])/mean.tr.S.Sigma2[1]
PRIAL.Sigma2.case2.OAS <- 100*(mean.tr.S.Sigma2[2]-mean.tr.OAS.Sigma2[2])/mean.tr.S.Sigma2[2]
PRIAL.Sigma2.case3.OAS <- 100*(mean.tr.S.Sigma2[3]-mean.tr.OAS.Sigma2[3])/mean.tr.S.Sigma2[3]
#########################################################
PRIAL.Sigma3.case1.OAS <- 100*(mean.tr.S.Sigma3[1]-mean.tr.OAS.Sigma3[1])/mean.tr.S.Sigma3[1]
PRIAL.Sigma3.case2.OAS <- 100*(mean.tr.S.Sigma3[2]-mean.tr.OAS.Sigma3[2])/mean.tr.S.Sigma3[2]
PRIAL.Sigma3.case3.OAS <- 100*(mean.tr.S.Sigma3[3]-mean.tr.OAS.Sigma3[3])/mean.tr.S.Sigma3[3]
#########################################################
PRIAL.Sigma4.case1.OAS <- 100*(mean.tr.S.Sigma4[1]-mean.tr.OAS.Sigma4[1])/mean.tr.S.Sigma4[1]
PRIAL.Sigma4.case2.OAS <- 100*(mean.tr.S.Sigma4[2]-mean.tr.OAS.Sigma4[2])/mean.tr.S.Sigma4[2]
PRIAL.Sigma4.case3.OAS <- 100*(mean.tr.S.Sigma4[3]-mean.tr.OAS.Sigma4[3])/mean.tr.S.Sigma4[3]
##############################################################################################
##############################################################################
PRIAL.Sigma1.case1.FS <- 100*(mean.tr.S.Sigma1[1]-mean.tr.FS.Sigma1[1])/mean.tr.S.Sigma1[1]
PRIAL.Sigma1.case2.FS <- 100*(mean.tr.S.Sigma1[2]-mean.tr.FS.Sigma1[2])/mean.tr.S.Sigma1[2]
PRIAL.Sigma1.case3.FS <- 100*(mean.tr.S.Sigma1[3]-mean.tr.FS.Sigma1[3])/mean.tr.S.Sigma1[3]
#########################################################
PRIAL.Sigma2.case1.FS <- 100*(mean.tr.S.Sigma2[1]-mean.tr.FS.Sigma2[1])/mean.tr.S.Sigma2[1]
PRIAL.Sigma2.case2.FS <- 100*(mean.tr.S.Sigma2[2]-mean.tr.FS.Sigma2[2])/mean.tr.S.Sigma2[2]
PRIAL.Sigma2.case3.FS <- 100*(mean.tr.S.Sigma2[3]-mean.tr.FS.Sigma2[3])/mean.tr.S.Sigma2[3]
#########################################################
PRIAL.Sigma3.case1.FS <- 100*(mean.tr.S.Sigma3[1]-mean.tr.FS.Sigma3[1])/mean.tr.S.Sigma3[1]
PRIAL.Sigma3.case2.FS <- 100*(mean.tr.S.Sigma3[2]-mean.tr.FS.Sigma3[2])/mean.tr.S.Sigma3[2]
PRIAL.Sigma3.case3.FS <- 100*(mean.tr.S.Sigma3[3]-mean.tr.FS.Sigma3[3])/mean.tr.S.Sigma3[3]
#########################################################
PRIAL.Sigma4.case1.FS <- 100*(mean.tr.S.Sigma4[1]-mean.tr.FS.Sigma4[1])/mean.tr.S.Sigma4[1]
PRIAL.Sigma4.case2.FS <- 100*(mean.tr.S.Sigma4[2]-mean.tr.FS.Sigma4[2])/mean.tr.S.Sigma4[2]
PRIAL.Sigma4.case3.FS <- 100*(mean.tr.S.Sigma4[3]-mean.tr.FS.Sigma4[3])/mean.tr.S.Sigma4[3]

result.Sigma1.case1 =
 c(PRIAL.Sigma1.case1.U1,PRIAL.Sigma1.case1.U2,PRIAL.Sigma1.case1.DS,PRIAL.Sigma1.case1.LW,
 ...PRIAL.Sigma1.case1.RBLW,PRIAL.Sigma1.case1.OAS,PRIAL.Sigma1.case1.FS)
result.Sigma1.case2 =
 c(PRIAL.Sigma1.case2.U1,PRIAL.Sigma1.case2.U2,PRIAL.Sigma1.case2.DS,PRIAL.Sigma1.case2.LW,
 ...PRIAL.Sigma1.case2.RBLW,PRIAL.Sigma1.case2.OAS,PRIAL.Sigma1.case2.FS)
result.Sigma1.case3 =
 c(PRIAL.Sigma1.case3.U1,PRIAL.Sigma1.case3.U2,PRIAL.Sigma1.case3.DS,PRIAL.Sigma1.case3.LW,
 ...PRIAL.Sigma1.case3.RBLW,PRIAL.Sigma1.case3.OAS,PRIAL.Sigma1.case3.FS)

result.Sigma2.case1 =
 c(PRIAL.Sigma2.case1.U1,PRIAL.Sigma2.case1.U2,PRIAL.Sigma2.case1.DS,PRIAL.Sigma2.case1.LW,
```

```
 ...PRIAL.Sigma2.case1.RBLW,PRIAL.Sigma2.case1.OAS,PRIAL.Sigma2.case1.FS)
result.Sigma2.case2 =
 c(PRIAL.Sigma2.case2.U1,PRIAL.Sigma2.case2.U2,PRIAL.Sigma2.case2.DS,PRIAL.Sigma2.case2.LW,
 ...PRIAL.Sigma2.case2.RBLW,PRIAL.Sigma2.case2.OAS,PRIAL.Sigma2.case2.FS)
result.Sigma2.case3 =
 c(PRIAL.Sigma2.case3.U1,PRIAL.Sigma2.case3.U2,PRIAL.Sigma2.case3.DS,PRIAL.Sigma2.case3.LW,
 ...PRIAL.Sigma2.case3.RBLW,PRIAL.Sigma2.case3.OAS,PRIAL.Sigma2.case3.FS)

result.Sigma3.case1 =
 c(PRIAL.Sigma3.case1.U1,PRIAL.Sigma3.case1.U2,PRIAL.Sigma3.case1.DS,PRIAL.Sigma3.case1.LW,
 ...PRIAL.Sigma3.case1.RBLW,PRIAL.Sigma3.case1.OAS,PRIAL.Sigma3.case1.FS)
result.Sigma3.case2 =
 c(PRIAL.Sigma3.case2.U1,PRIAL.Sigma3.case2.U2,PRIAL.Sigma3.case2.DS,PRIAL.Sigma3.case2.LW,
 ...PRIAL.Sigma3.case2.RBLW,PRIAL.Sigma3.case2.OAS,PRIAL.Sigma3.case2.FS)
result.Sigma3.case3 =
 c(PRIAL.Sigma3.case3.U1,PRIAL.Sigma3.case3.U2,PRIAL.Sigma3.case3.DS,PRIAL.Sigma3.case3.LW,
 ...PRIAL.Sigma3.case3.RBLW,PRIAL.Sigma3.case3.OAS,PRIAL.Sigma3.case3.FS)

result.Sigma4.case1 =
 c(PRIAL.Sigma4.case1.U1,PRIAL.Sigma4.case1.U2,PRIAL.Sigma4.case1.DS,PRIAL.Sigma4.case1.LW,
 ...PRIAL.Sigma4.case1.RBLW,PRIAL.Sigma4.case1.OAS,PRIAL.Sigma4.case1.FS)
result.Sigma4.case2 =
 c(PRIAL.Sigma4.case2.U1,PRIAL.Sigma4.case2.U2,PRIAL.Sigma4.case2.DS,PRIAL.Sigma4.case2.LW,
 ...PRIAL.Sigma4.case2.RBLW,PRIAL.Sigma4.case2.OAS,PRIAL.Sigma4.case2.FS)
result.Sigma4.case3 =
 c(PRIAL.Sigma4.case3.U1,PRIAL.Sigma4.case3.U2,PRIAL.Sigma4.case3.DS,PRIAL.Sigma4.case3.LW,
 ...PRIAL.Sigma4.case3.RBLW,PRIAL.Sigma4.case3.OAS,PRIAL.Sigma4.case3.FS)

Result = rbind(result.Sigma1.case1,result.Sigma1.case2,result.Sigma1.case3,
result.Sigma2.case1,result.Sigma2.case2,result.Sigma2.case3,
result.Sigma3.case1,result.Sigma3.case2,result.Sigma3.case3,
result.Sigma4.case1,result.Sigma4.case2,result.Sigma4.case3)
colnames(Result) <- c("U1","U2","DS","LW","RBLW","OAS","FS")
Result
```

مراجع

[۱] آشوری، ا. (۱۳۹۳). *برآوردگر انقباضی ماتریس کوورایانس توزیع نرمال چندمتغیره با بعد بالا*، پایان‌نامه کارشناسی ارشد، دانشگاه صنعتی شاهرود.

[۲] رودین، و. (۱۳۷۳) *اصول آنالیز ریاضی*، ترجمه عالم‌زاده، ع.ا.، علمی و فنی، تهران.

[۳] کرمی‌کبیر، ح. (۱۳۹۲). *برآورد انقباضی در مدل‌های محدودشده*، پایان‌نامه کارشناسی ارشد، دانشگاه صنعتی شاهرود.

[4] Abramsky, S., Jung, A., (1994). Domain theory, in: Handbook of Logic in Computer Science. *Clarendon Press, Oxford. Vol. 3.* pp. 1-68.

[5] Bai, J., Shi, S., (2011). Estimating high dimensional covariance matrices and its applications. *Ann. Econ. Finance* **12**, 199-215.

[6] Bickel, P., Levina, E., (2008). Covariance regularization by thresholding. *Ann. Statist.*, **36**, 2577-2604.

[7] Birkhoff, G., (1967). Lattice Theory. *Amer. Math. Soc. 3rd Edition, AMS Colloq. Publication, Vol. 25.*

[8] Cai, T., Liu, W., (2011). Adaptive threshold for sparse covariance matrix estimation. *J. Amer. Statist. Assoc.* **106,** 672-684.

[9] Cai, T., Zhou, H., (2010). Optimal rates of convergence for sparse covariance matrix estimation. *Manuscript. University of Pennsylvania.*

[10] Cao, G., and Bouman, C., (2009). Covariance estimation for high dimensional data vectors using the sparse matrix transform, in *Advances in Neural Information Processing Systems 21*, editors: Koller, D., Schuurmans, D., Bengio, L., and Bottou, L., 225-232.

[11] Chaudhuri, S., Drton, M., Richardson, T. S. (2007). Estimation of covariance matrix with Zeros, *Biometrika*, **94**(1), 199-216.

[12] Chen, Z. and Dunson, D., (2003). Random effects selection in linear mixed models. *Biometrics*, **59**, 762-769.

[13] Chen, S.X., Qin, Y.L., (2010). A two-sample test for high-dimensional data with applications to gene-set testing, *Ann. Statist.* **38**, 808-835.

[14] Chen, Y., Wiesel, A., Eldar, C.Y., Hero, A.O., (2010). Shrinkage algorithms for MMSE covariance estimation. *IEEE Trans. Signal Process.* **58**, 5016-5029.

[15] Chen, S.X., Zhang, L.X., Zhong, P.S., (2010). Tests for high-dimensional covariance matrices. *J. Amer. Statist. Assoc.* **105**, 810-819.

[16] Corbeil, R.R. and Searle, S.R., (1976). A comparison of variance component estimators, *Biometric.* **32**(4), 779-791.

[17] Daniels, M.J., (2005) Shrinkage priors for the dependence structure in longitudinal data. *J. Statist. Plan. Inference.* **127**, 119-130.

[18] Daniels, M.J., (2006). Bayesian modeling of several covariance matrices and some results on the propriety of the posterior for linear regression with correlated and/or heterogeneous errors. *J. Mult. Anal.* **97**, 1185-1207.

[19] Daniels, M.J. and Kass, R., (2001). Shrinkage estimators for a covariance matrix. *Biometrics.* **57**, 1173-1184.

[20] Daniels, M. J., and Pourahmadi, M., (2002). Bayesian analysis of covariance matrices and dynamic models of longitudinal data, *Biometrika.* **89**(3), 553-566.

[21] Dey, D.K. and Srinivasan, C., (1985). Estimation of a covariance matrix under Stein' loss, *Ann. Statist.* **4**, 1581-1591.

[22] Edwards, D.A., (1978). On the existence of probability measures with given marginals. *Ann. Inst. Fourier, Grenoble.* **28**, 53-78.

[23] Efron, B., and Morris, C., (1975). Data analysis using Stein's estimator and its generalizations, *J. Am. Statist. Assoc.* **70** (350), 311-319.

[24] Efron, B., and Morris, C., (1976). Multivariate empirical Bayes and estimation of covariance matrices, *Ann. Statist.* **4** (1), 22-32.

[25] Fan, J. and Lv, J. (2008). Sure independence screening for ultrahigh dimensional feature space, *J. Royal. Statist. Soc. Series B.* **70**(5), 849-911.

[26] Fan, J., and Wu, Y., (2009). Semiparametric estimation of covariance matrices for longitudinal data, *J. Am. Statist. Assoc.*, **103**(484), 1520-1533.

[27] Fraley, C. and Burns, P. J., (1995). Large-Scale Estimation of Variance and Covariance Components. *SIAM J. Sci. Comput.*, **16**(1), 192-209.

[28] Fisher, T. j and Sun, X., (2011), Improved Stein-type shrinkage estimators for the high-dimonsional multivariate normal covariance matrix. *Comp. Statist. Data Anal.*, **55**, 1909-1918.

[29] Fujikoshi, Y., Himeno, T., Wakaki, H., (2004). Asymptotic results of a high dimensional MANOVA test and power comparison when the dimension is large compared to the sample size. *J. Japan Statist. Soc.* **34**, 19-26.

[30] Ghosh, M., Lahiri, P., (1987). Robust empirical Bayes estimation of means from stratified samples. *J. Amer. Statist. Assoc.* **82**, 1153-1162.

[31] Haff, L.R., (1979). Estimation of the inverse covariance matrix: random mixtures of the inverse Wishart matrix and the identity, *Ann. Statist.*, **7**(6), 1264-1276.

[32] Haff, L.R., (1981). Corrections to Estimation of the inverse covariance matrix: random mixtures of the inverse Wishart matrix and the identity, *Ann. Statist.* **9**(5), 11-32.

[33] Haff, L.R., (1991). The variational form of certain Bayes estimators. *Ann. Statist.*, **19** 1163-1190.

[34] Hannart, A., Naveau, P., (2014). Estimating high dimensional covariance matrices: A new look at the Gaussian conjugate framework. *J. Mult. Anal.* **131**, 149-162.

[35] Hartigan, J.A., (1969). Linear Bayes methods. *J. Royal. Statist. Soc. Ser. B* **31**, 446-454.

[36] Hemmerle, W. J. and Hartley, H. O. (1973). Computing maximum likelihood estimates for the mixed A.O.V. model using the W transformation, *Technometric*, **15**(4), 819-831.

[37] Himeno, T., Yamada, T., (2014). Estimation for some functions of covariance matrix in high dimension under non-normality and its applications. *J. Mult. Anal.* **130**, 27-44.

[38] Horn, A., Tarski, A., (1948). Measures on Boolean algebras. *Trans. Amer. Math. Soc.*, **64**, 467-497.

[39] Ikeda, Y., Kubokowa, T., Srivastava, M. S., (2016). Comparison of linear shrinkage estimators of a large covariance matrix in normal and non-normal distributions, *Comp. Statatist. and Data Anal.*, **95**, 95-108.

[40] James, W. and Stein, C., (1961). Estimation with quadratic loss. *In the Fourth Berkeley Symposium on Mathematical Statistics and Probability, CA: University of California Press.* pp. 361-379.

[41] Johnstone, I.M., (2001). On the distribution of the largest eigenvalue on principle component analysis, *Ann. Statist.*, **29**(2), 295-357.

[42] Johnstone, I. M., and Lu, A. Y., (2009). On consistency and sparsity for principle components analysis in high dimensions, *J. Amer. Statist. Assoc.*, **104**, 682-693.

[43] Konno, Y., (2009). Shrinkage estimators for large covariance matrices in multivariate real and complex normal distributions under an invariant quadratic loss. *J. Mult. Anal.* **100**, 2237-2253.

[44] Lam, C., Fan, J., (2009). Sparsistency and rates of convergence in large covariance matrix estimation. *Ann. Statist.*, **37**, 4254-4278.

[45] Lang, S., (1978). *Linear algebra*, Springer.

[46] Ledoit, O., Wolf, M., (2003), Improved estimation of the covariance matrix of stock returns with an application to portfolio selection. *J. Empir. Finance.*, **10**, 603-621.

[47] Ledoit, O., Wolf, M., (2004). A well-conditioned estimator for large-dimensional covariance matrices. *J. Mult. Anal.* **88**, 365-411.

[48] Ledoit, O., Wolf, M., (2012). Nonlinear shrinkage estimation of large-dimensional covariance matrices. *Ann. Statist.* **40**, 1024-1060.

[49] Ledoit, O., Wolf, M., (2015). Spectrum estimation: A unified framework for covariance matrix estimation and PCA in large dimensions. *J. Mult. Anal.* **139**, 360-384.

[50] Magnus, J. R., Neudecker, H., (1979). The commutation matrix: some properties and applications. *Ann. Statist.* **7**, 381-894.

[51] Muirhead, R.J., (1982), *Aspects Stat. Theory.* John Wiley, New York.

[52] Pourahmadi, M., (1999). Joint mean-covariance models with applications to longitudinal data: unconstrained parameterisation, *Biometrika*, **86**(3), 677-690.

[53] Roeverto, A., (2000) Cholesky decomposition of hyper invers Wishart matri, *Biometrika*, **87**(1), 99-112.

[54] Robbins, H., (1983) Some thoughts on empirical Bayes estimation. *Ann. Statist.* **11**, 713-723.

[55] Rothman, A., Levina, E., Zhu, J., (2009). Generalized thresholding of large covariance matrices. *J. Amer. Statist. Assoc.*, **104**, 177-186.

[56] Saleh, A. K. M. E., (2006). Theory of Preliminary Test and Stein-Type Estimation with Applications, *John Wiley and Sons.*

[57] Schafer, J., Strimmer, K., (2005). A shrinkage approach to large-scale covariance matrix estimation and implications for functional genomics. *Statist. Appl. Genet. Molec. Biol.*

[58] Schott, J. R., (2007). Some high-dimensional tests for a one-way MANOVA. *J. Mult. Anal.*, **98**, 1825-1839.

[59] Smith, M., and Kohn, R., (2002), Parsimonious Covariance Matrix Estimation for Longitudinal Data, *J. Amer. Statist. Assoc.*, **97**, 1141-1153.

[60] Srivastava, M. S., (2005). Some tests concerning the covariance matrix in high dimensional data. *J. Japan Statist. Soc.* **35**, 251-272.

[61] Srivastava, M. S., Fujikoshi, Y., (2006). Multivariate analysis of variance with fewer observations than the dimension. *J. Mult. Anal.* **97**, 1927-1940.

[62] Srivastava, M. S., Kollo, T., Rosen, D. V., (2011). Some tests for the covariance matrix with fewer observations than the dimension under non-normality. *J. Mult. Anal.* **102**, 1090-1103.

[63] Srivastava, M.S., Kubokawa, T., (2007). Comparison of discrimination methods for high dimensional data. *J. Japan Statist. Soc.* **37**, 123-134.

[64] Srivastava, M.S., Yanagihara, H., Kubokawa, T., (2014). Tests for covariance matrices in high dimension with less sample size. *J. Mult. Anal.* **130**, 289-309.

[65] Stein, C., (1956), Inadmissibility of the usual estimator for the mean of a multivariate normal distribution. *In Proceedings of the Third Berkeley Symposium*.,**1**, 197-206.

[66] Touloumis, A., (2015). Nonparametric Stein-type shrinkage covariance matrix estimators in high-dimensional setting. *Comp. Statist. Data Anal.* **83**, 251-261.

[67] Wu, W. B., and Pourahmadi, M., (2003). Nonparametric estimation of large covariance matrices of longitudinal data, *Biometrika*, **90**(4), 831-844.

[68] Yang, R. and Berger, J. O., (1994). Estimation of a covariance matrix using the reference prior, *Ann. Statist.*, **22** (3), 1195-1211.

Abstract

In multivariate analysis, one of the most challenging cases of estimating a covariance matrix of a random variable is when the number of variables (dimension) is large. Also, in some multivariate discussions, it is necessary to estimate the function of a covariance matrix. Suppose p represents the dimension. In this thesis, we examine the covariance matrix estimator and some functions of it for the high dimension $(p \to \infty)$. In this regard, we assume that the population of sampling has normal and non-normal distribution and introduce non-parametric shrinkage estimators of the covariance matrix in the high dimension and examine some of their properties. Also, using simulation examples, we evaluate the efficiency of the proposed shrinkage estimators for the covariance matrix in the high dimension.

Keywords

Covariance Matrix, Shrinkage Estimator, High Dimension, Consistent of Nonparametric Estimator.

Shahrood University of Technology

Faculty of Mathematical Sciences
MSc Thesis in: Mathematical Statistics

Shrinkage Covariance Matrix Estimators in High-Dimensional Settings

By: Zahra Hosseinpour

Supervisor

Dr. Mohammad Arashi

July 2018