
بسم الله الرحمن الرحيم

دانشکده علوم ریاضی
گروه آمار

پایان‌نامه

برای دریافت درجه کارشناسی ارشد در رشته
آمار، گرایش آمار ریاضی

عنوان

برآوردگر انقباضي ماتريس كوواريانس توزيع نرمال چند متغيره با بعد بالا

استاد راهنما

دکتر محمد آرشی

دانشجو

الهام آشوری

شهریور ۱۳۹۳

(اللهم عجل لولیک الفرج)

خدایا...

« به علمای ما مسئولیت و به عوام ما علم و به مومنان ما روشنایی و به روشنفکران ما ایمان و به متعصبین ما فهم و به فهمیدگان ما تعصب و به زنان ما شعور و به مردان ما شرف و به پیران ما آگاهی و به جوانان ما اصالت و به اساتید ما عقیده و به دانشجویان ما نیز عقیده و به خفتگان ما بیداری و به دینداران ما دین و به نویسندگان ما تعهد و به هنرمندان ما درد و به شاعران ما شعور و به محققان ما هدف و به نومیدان ما امید و به ضعیفان ما نیرو و به محافظه‌کاران ما گستاخی و به نشستگان ما قیام و به راکدان ما تکان و به مردگان ما حیات و به کوران ما نگاه و به خاموشان ما فریاد و به مسلمانان ما قرآن و به شیعیان ما علی(ع) و به فرقه های ما وحدت و به حسودان ما شفا و به خودبینان ما انصاف و به فحاشان ما ادب و به مجاهدان ما صبر و به مردم ما خودآگاهی و به همه ملت ما همت، تصمیم و استعداد فداکاری و شایستگی نجات و عزت ببخش » [1].

[1] دکتر علی شریعتی

سپاس‌گزاری

بخوان به نام پروردگارت که جهان را آفرید، همان که انسان را از خون بسته‌ای خلق کرد. بخوان که پروردگارت از (همه) بزرگوارتر است همان کس که به‌وسیله قلم تعلیم نمود و به انسان آنچه نمی‌دانست یاد داد (سوره علق آیه ۱-۴).
سپاس خداوندی را که بخشنده، مهربان و شایسته پرستش است. خداوندی که با لطف بی‌کران خود، آدمی را زیور عقل آراست و وی را کیمیای محبت کرامت فرمود.
در آغاز وظیفه خود می‌دانم از زحمات بی‌دریغ استاد راهنمای خود، جناب آقای دکتر محمد آرشی، صمیمانه تشکر و قدردانی کنم که بدون راهنمایی‌های ارزنده ایشان، این مجموعه به انجام نمی‌رسید. از اساتید محترم داور جناب آقای دکتر حسین باغیشنی و جناب آقای دکتر محمد رضا ربیعی که زحمت داوری این مجموعه را به عهده گرفتند و با نقد و ارائه نظرها و پیشنهادهای سازنده خود باعث غنای هر چه بیشتر آن شده‌اند، صمیمانه تشکر و قدردانی می‌کنم.
همچنین از جناب آقای دکتر احمد رضازاده نزاکتی، جناب آقای دکتر داود شاهسونی و جناب آقای دکتر حسین باغیشنی که افتخار حضور در کلاس درسشان نصیبم شد و از آن‌ها سخت کوشی و پشتکار را آموختم، کمال تشکر را دارم.
بر خود لازم می‌دانم از اساتید گرامی و گرانقدر دانشگاه پیام‌نور گنبد کاووس جناب آقای مجتبی کاشانی، جناب آقای آنه گلدی محمیانی و جناب آقای آنه ارجنلی و دیگر اساتید گروه آمار و ریاضی این دانشگاه نهایت تشکر و قدردانی را داشته باشم. وظیفه خود می‌دانم تا از دوست خوبم سرکار خانم مینا نوروزی‌راد که همواره از مشورت و همفکری ایشان برخوردار بوده‌ام نیز تشکر نمایم.
هر چند این اوراق در بیان ارادت قلبی، به پدر و مادر عزیزم ناتوان‌اند، اما به رسم ادب، صمیمانه و خالصانه، از پدرم برای صبر و فداکاری‌اش و مادرم به خاطر گذشت و مهربانی‌اش، تشکر و قدردانی به عمل می‌آورم. همچنین از برادران و خواهران خوبم که مشوق بنده بوده‌اند، کمال تشکر را دارم.
همچنین از دوستان خوبم خانمها سمیرا سوخت‌سرایی، سیمین پویا، مریم برزوئی بیدگلی و زهره فرهادی به واسطه همراهی و همفکری‌شان از صمیم قلب سپاس‌گزارم. در پایان از تمامی دانشجویان آمار و ریاضی دانشکده ریاضی و همچنین از مسئول آموزش سرکار خانم خداوردی به پاس لطف و محبت‌هایش قدردانی نمایم.

الهام آشوری

شهریور ۱۳۹۳

تعهدنامه

اینجانب **الهام آشوری** دانشجوی کارشناسی ارشد رشته **آمار** **دانشکده علوم ریاضی** دانشگاه شاهرود، نویسنده پایان‌نامه با عنوان **برآوردگر انقباضي ماتريس كوواريانس توزيع نرمال چند متغيره با 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 نشده است.
- حقوق معنوی این اثر، به دانشگاه شاهرود متعلق دارد، و مقالات مستخرج با نام « دانشگاه شاهرود » یا « University Shahrood » به چاپ خواهد رسید.
- حقوق معنوی تمام افرادی که در به‌دست آوردن نتایج اصلی پایان‌نامه تاثیرگذار بوده‌اند، در مقالات مستخرج از پایان‌نامه رعایت می‌گردد.
- در تمام مراحل انجام این پایان‌نامه، در مواردی که از موجود زنده (یا بافت‌های آنها) استفاده شده است، ضوابط و اصول اخلاقی رعایت شده است.
- 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 شده است.

الهام آشوری

شهریور ۱۳۹۳

مالکیت نتایج و حق نشر

- تمام حقوق معنوی این اثر و محصولات آن (مقالات مستخرج، کتاب، برنامه‌های رایانه‌ای، نرم‌افزارها و تجهیزات ساخته شده) متعلق به دانشگاه 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 نمی‌باشد.

چکیده

در بسیاری از کاربردها نیاز به برآوردگری برای ماتریس واریانس-کوواریانس است که معکوس‌پذیر باشد. در ابعاد بالا ماتریس واریانس-کوواریانس معکوس‌پذیر نخواهد بود. در این حالت یک روش معمول برای برآورد ماتریس واریانس-کوواریانس در ابعاد بالاتر استفاده از روش انقباضی [۲] است.

در این پایان‌نامه برآوردگر، به‌صورت ترکیبی محدب از ماتریس کوواریانس نمونه و ماتریس هدف (معین مثبت و نامنفرد) در نظر گرفته می‌شود. مجموعه کارهایی که در این راستا انجام شده است، نشانگر این است که پارامتر انقباضی بهینه تحت میانگین توان‌های دوم خطا وجود دارد و باید توسط داده‌ها برآورده شود. در این مجموعه برآوردگرهای جدیدی با استفاده از ترکیب محدب ارائه شده و سه ماتریس هدف معرفی شده است. مطالعات شبیه‌سازی و مثال واقعی که در این پایان نامه انجام شده است، بهبود برآوردگرها را در حالتی که بعد بالاست، نشان می‌دهد.

[۲] Shrinkage

پیشگفتار

پس از کشف پدیده استاین [۳] (۱۹۵۶)، بسیاری از تلاشها در جهت بهبود برآورد ماتریس واریانس-کوواریانس زمانی که بعد آن بالاست، صورت گرفته است. در بسیاری از کاربردها بعد، p، بزرگتر از تعداد مشاهدات، N می‌باشد؛ که در این صورت برآوردگری که با استفاده از روش درستنمایی ماکسیمم حاصل می‌شود، مناسب نمی‌باشد. بدین منظور در این پایان‌نامه در جستجوی برآوردگری مناسب برای ماتریس واریانس-کوواریانس با استفاده از روش انقباضی هستیم.

در این مجموعه ابتدا به یافت برآوردگر انقباضی برای ماتریس واریانس-کوواریانس در توزیع نرمال چندمتغیره پرداخته و از آنجایی‌که بهبود برآوردگرها تنها محدود به توزیع نرمال نشده است بلکه در توزیع‌های متقارن زیادی از جمله توزیع t نیز مورد بررسی قرار گرفته است و توزیع t یکی از توزیع‌های پرکاربرد آماری است، و در حالت خاصی نیز به توزیع نرمال تبدیل می‌شود، به دنبال برآوردگر بهبود یافته برای پارامتر ماتریس واریانس-کوواریانس، در حالتی که بعد فضای پارامتر بزرگتر از مشاهدات است، نیز می‌باشیم. بدین منظور در این پایان نامه با در نظر گرفتن سه ماتریس هدف متفاوت، به دنبال دستیابی به برآوردگر انقباضی ماتریس واریانس-کوواریانس می‌باشیم. این مجموعه شامل ۳ فصل و دو پیوست می‌باشد. مطالب هر فصل بطور خلاصه عبارتند از:

- در فصل اول مقدمات کار را با ارائه تعاریف اولیه و لم‌های مورد نیاز در فصل‌های بعدی فراهم می‌کنیم.

- در فصل دوم به دنبال برآوردگر انقباضی برای ماتریس واریانس-کوواریانس در توزیع نرمال چندمتغیره [۴] در حالتی که بعد متغیر (p) بزرگتر از تعداد مشاهدات (N) است، می‌باشیم همچنین مقاله‌ای تحت عنوان « برآوردگر انقباضی ماتریس واریانس-کوواریانس » از آن مستخرج شده است.

- در فصل سوم، همانند فصل دوم به دنبال برآوردگر انقباضی برای ماتریس واریانس-کوواریانس هستیم با این تفاوت که توزیع مورد نظر، توزیع t چندمتغیره [۵] است. همچنین مقاله‌ای تحت

[۳] Stein
[۴] Multivariate normal distribution
[۵] Multivariate t distribution

عنوان « برآوردگر انقباضی ماتریس واریانس-کوواریانس با استفاده از ترکیب محدب » از آن مستخرج شده است.

- در پیوست آ نتیجه‌گیری به همراه ارائه پیشنهادات برای آینده تحقیق ارائه شده است و در پایان دستورهای مورد نیاز برنامه نویسی بر اساس نرم‌افزار R، که در این پایان‌نامه استفاده شده است، را می‌آوریم.
- در پیوست ب به تعاریف جبر خطی مورد نیاز که در متن پایان نامه به آن اشاره شده است، می‌پردازیم.

در این مجموعه، در مواردی که برهان از نویسنده بوده از علامت * و در صورتی که هم قضیه و هم برهان از نویسنده می‌باشد، از نماد ** استفاده شده است.

لیست مقالات مستخرج از پایان‌نامه

- آشوری، ا. و آرشی، م. (۱۳۹۳)، برآوردگر انقباضی ماتریس واریانس-کوواریانس، مجله اندیشه آماری، تحت داوری.

- Ashori E. and Arashi M. (2014), Shirinkage Estimation of Covariance Matrix Through Convex Optimization, The 7th International Conference of Iranian Operations Research Society, Semnan, Iran.

نمادها

X: متغیر تصادفی
$\mathbf{X}$: ماتریس
$\mathbf{x}$: بردار تصادفی
A^T: ترانهاده ماتریس A
$diag(A)$: ماتریس قطری حاصل از عناصر A
$det(A)$: دترمینان ماتریس مربع A
$tr(A)$: اثر ماتریس مربع A
$etr(A)$: معادل $\exp(tr(A))$ ، زمانی که A ماتریسی مربع باشد.
R^+: فضای حقیقی مثبت
R^n: فضای n بعدی حقیقی
$J(x \longrightarrow y)$: ژاکوپی تبدیل x به y
ave: میانگین
μ_λ: میانگین مقادیر ویژه ماتریس Σ
$\approx$: خیلی نزدیک
λ: پارامتر انقباضی
Λ: ماتریس مقادیر ویژه
Γ: ماتریس بردارهای ویژه
I: ماتریس همانی
Ω: ماتریسی که عناصر قطر اصلی آن صفر است.
Θ: فضای پارامتر
$\langle f, g \rangle$: حاصل ضرب داخلی f و g
$||f||$: نرم تابع f
$||f||_F$: نرم فروبنیوس
df: درجه آزادی
$N_n(\mu, \Sigma)$: توزیع نرمال n-متغیره با میانگین μ و ماتریس کواریانس Σ
$G(a, b)$: توزیع گاما با پارامترهای a و b
$InvG(a, b)$: توزیع گاما معکوس با پارامترهای a و b
$t_n(\mu, \Sigma, \nu)$: توزیع t-استودنت n-متغیره با میانگین μ، ماتریس کواریانس Σ و درجه آزادی ν
$SL_n(\mu, \Sigma, \nu)$: توزیع اسلش n-متغیره با میانگین μ، ماتریس کواریانس Σ و درجه آزادی ν
$CN_n(\mu, \Sigma, \nu, \gamma)$: توزیع نرمال آلایشی n-متغیره با میانگین μ، ماتریس کواریانس Σ، nu پارامتر نشاندهنده‌ی درصد داده‌های دور افتاده و γ می‌تواند تفسیری از فاکتور ماتریس کواریانس باشد.
$L(\theta, \delta)$: تابع زیان
$\Re(\theta, \delta)$: تابع مخاطره

MLE: برآوردگر درستنمایی ماکسیمم
MSE: مجموع توان‌های دوم خطا

فهرست مطالب

لیست تصاویر س

لیست جداول ش

۱ مقدمه و دورنما ۳

۱.۱ مقدمه . . . ۳

۲.۱ تعاریف و توابع ریاضی . . . ۵

۱.۲.۱ ژاکوبی تبدیل‌ها . . . ۷

۳.۱ توزیع‌های آماری . . . ۹

۱.۳.۱ توزیع‌های چندمتغیره . . . ۱۰

۴.۱ توزیع‌های آمیخته مقیاسی . . . ۱۷

۵.۱ مدل نرمال آمیخته مقیاسی . . . ۱۸

۶.۱ مثال‌هایی از توزیع نرمال آمیخته مقیاسی . . . ۱۹

۱.۶.۱ توزیع t–استیودنت . . . ۱۹

۲.۶.۱ توزیع اسلش . . . ۱۹

۳.۶.۱ توزیع نرمال آلایشی . . . ۲۰

۷.۱ قضایا و لم‌های اساسی . . . ۲۰

۲ برآوردگر انقباضی ماتریس کواریانس در توزیع نرمال چندمتغیره ۲۱

۱.۲ مقدمه . . . ۲۱

۲.۲ برآورد . . . ۲۲

۳.۲ برآوردگر انقباضی نوع استاین . . . ۲۴

۴.۲ انواع پارامترهای انقباضی . . . ۲۶

۱.۴.۲ برآورد انقباضی LW . . . ۲۶

۲.۴.۲ برآورد انقباضی $RBLW$ . . . ۲۶

۳.۴.۲ برآورد انقباضی SS . . . ۲۷

۵.۲ برآوردگر جدید تحت شرایط بهینه . . . ۲۸

۶.۲ برآوردگر انقباضی، برای ماتریس هدف همانی . . . ۳۴

۷.۲ برآوردگر انقباضی، برای ماتریس هدف غیر همانی ۳۶
۸.۲ مقایسه برآوردگرها ۴۲
۱.۸.۲ مطالعه شبیه‌سازی ۴۲
۹.۲ هزینه شبیه‌سازی ۴۷
۱۰.۲ مثال واقعی ۴۹

۳ برآوردگر انقباضی ماتریس کوواریانس در توزیع t چندمتغیره ۵۳
۱.۳ مقدمه ۵۳
۲.۳ برخی خواص توزیع t چندمتغیره ۵۴
۳.۳ برآوردگر انقباضی ۵۸
۴.۳ برآوردگر انقباضی، برای ماتریس هدف همانی ۶۱
۵.۳ برآوردگر انقباضی، برای ماتریس هدف غیر همانی ۶۲
۶.۳ مطالعه شبیه‌سازی ۶۶

آ خلاصه، نتیجه‌گیری، پیشنهادات برای آینده تحقیق و برنامه‌های نرم‌افزار R ۸۱
آ.۱ دستورهای نرم‌افزار R ۸۳

ب جبرخطی ۹۹
ب.۱ تعاریف و قضایا ۹۹

مراجع ۱۰۳

واژه‌نامه فارسی به انگلیسی ۱۰۹

واژه‌نامه انگلیسی به فارسی ۱۱۱

لیست تصاویر

۱.۱ نمودارهای تابع چگالی و منحنی‌های تراز توزیع نرمال دومتغیره ۱۲
۲.۱ نمودارهای تابع چگالی و منحنی‌های تراز توزیع نرمال دومتغیره ۱۳
۳.۱ نمودارهای تابع چگالی و منحنی‌های تراز توزیع کوشی دومتغیره ۱۳
۴.۱ نمودارهای تابع چگالی و منحنی‌های تراز توزیع لجستیک دومتغیره ۱۴
۵.۱ نمودارهای تابع چگالی و منحنی‌های تراز توزیع نمایی توانی دومتغیره ۱۴
۶.۱ نمودارهای تابع چگالی و منحنی‌های تراز توزیع کاتز دومتغیره ۱۴
۷.۱ نمودارهای تابع چگالی و منحنی‌های تراز توزیع لاپلاس دومتغیره ۱۵

۱.۲ پارامتر انقباضی LW ۲۷
۲.۲ پارامتر انقباضی $RBLW$ ۲۷
۳.۲ پارامتر انقباضی SS ۲۸
۴.۲ پارامتر انقباضی FS ۲۹
۵.۲ زمان مطالعه برای $p = ۲۰$ ۴۸
۶.۲ زمان مطالعه برای $n = ۳۰$ ۴۹
۷.۲ زمان مطالعه برای $n = p$ یکسان ۵۰

لیست جداول

۱.۲ برآورد λ برای $n = ۴۰$، $p = ۲۰$ و $T = \mu_\lambda I$ ۴۳
۲.۲ برآورد λ برای $n = ۳۰$، $p = ۳۰$ و $T = \mu_\lambda I$ ۴۳
۳.۲ برآورد λ برای $n = ۵$، $p = ۱۰۰$ و$T = \mu_\lambda I$ ۴۳
۴.۲ برآوردگر انقباضی برای $p = ۲۰$، $n = ۴۰$ و $T = \mu_\lambda I$ ۴۳
۵.۲ برآوردگر انقباضی برای $p = ۳۰$، $n = ۳۰$ و $T = \mu_\lambda I$ ۴۴
۶.۲ برآوردگر انقباضی برای $p = ۱۰۰$، $n = ۵$ و $T = \mu_\lambda I$ ۴۴
۷.۲ برآوردگر انقباضی برای $p = ۲۰$، $n = ۴۰$ و $T = I$ ۴۴
۸.۲ برآوردگر انقباضی برای $p = ۳۰$، $n = ۳۰$ و $T = I$ ۴۵
۹.۲ برآوردگر انقباضی برای $p = ۱۰۰$، $n = ۵$ و $T = I$ ۴۵
۱۰.۲ برآورد λ برای $n = ۴۰$، $p = ۲۰$ و $T = D_s$ ۴۵
۱۱.۲ برآورد λ برای $n = ۳۰$، $p = ۳۰$ و $T = D_s$ ۴۵
۱۲.۲ برآورد λ برای $n = ۵$، $p = ۱۰۰$ و $T = D_s$ ۴۶
۱۳.۲ برآوردگر انقباضی برای $p = ۲۰$ و $n = ۴۰$ و $T = D_s$ ۴۶
۱۴.۲ برآوردگر انقباضی برای $p = ۳۰$ و $n = ۳۰$ و $T = D_s$ ۴۶
۱۵.۲ برآوردگر انقباضی برای $p = ۱۰۰$ و $n = ۵$ و $T = D_s$ ۴۶
۱۶.۲ عدد شرطی برآوردگرها و مقدار عددی پارامتر انقباضی برای داده‌های $E.Coli$ برای $p = ۱۰۲$ و $n = ۸$ ۵۱

۱.۳ برآورد λ برای $n = ۴۰$، $p = ۲۰$، $T = \mu_\lambda I$ و$df = ۵$ ۶۷
۲.۳ برآورد λ برای $n = ۴۰$، $p = ۲۰$، $T = \mu_\lambda I$ و$df = ۱۰$ ۶۷
۳.۳ برآورد λ برای $n = ۴۰$، $p = ۲۰$، $T = \mu_\lambda I$ و$df = ۱۰۰$ ۶۷
۴.۳ برآوردگر انقباضی برای $p = ۲۰$، $n = ۴۰$، $T = \mu_\lambda I$ و $df = ۵$ ۶۸
۵.۳ برآوردگر انقباضی برای $p = ۲۰$، $n = ۴۰$، $T = \mu_\lambda I$ و $df = ۱۰$ ۶۸
۶.۳ برآوردگر انقباضی برای $p = ۲۰$، $n = ۴۰$، $T = \mu_\lambda I$ و $df = ۱۰۰$ ۶۸
۷.۳ برآورد λ برای $n = ۳۰$، $p = ۳۰$، $T = \mu_\lambda I$ و$df = ۵$ ۶۸
۸.۳ برآورد λ برای $n = ۳۰$، $p = ۳۰$، $T = \mu_\lambda I$ و$df = ۱۰$ ۶۹
۹.۳ برآورد λ برای $n = ۳۰$، $p = ۳۰$، $T = \mu_\lambda I$ و$df = ۱۰۰$ ۶۹

۱۰.۳ برآوردگر انقباضی برای $p = ۳۰$، $n = ۳۰$، $T = \mu_\lambda I$ و $df = ۵$ ۶۹
۱۱.۳ برآوردگر انقباضی برای $p = ۳۰$، $n = ۳۰$، $T = \mu_\lambda I$ و $df = ۱۰$ ۶۹
۱۲.۳ برآوردگر انقباضی برای $p = ۳۰$، $n = ۳۰$، $T = \mu_\lambda I$ و $df = ۱۰۰$ ۷۰
۱۳.۳ برآورد λ برای $n = ۵$، $p = ۱۰۰$، $T = \mu_\lambda I$ و$df = ۵$ ۷۱
۱۴.۳ برآورد λ برای $n = ۵$، $p = ۱۰۰$، $T = \mu_\lambda I$ و$df = ۱۰$ ۷۱
۱۵.۳ برآورد λ برای $n = ۵$، $p = ۱۰۰$، $T = \mu_\lambda I$ و$df = ۱۰۰$ ۷۱
۱۶.۳ برآوردگر انقباضی برای $p = ۱۰۰$، $n = ۵$، $T = \mu_\lambda I$ و $df = ۵$ ۷۱
۱۷.۳ برآوردگر انقباضی برای $p = ۱۰۰$، $n = ۵$، $T = \mu_\lambda I$ و $df = ۱۰$ ۷۲
۱۸.۳ برآوردگر انقباضی برای $p = ۱۰۰$، $n = ۵$، $T = \mu_\lambda I$ و $df = ۱۰۰$ ۷۲
۱۹.۳ برآورد λ برای $n = ۴۰$، $p = ۲۰$، $T = I$ و $df = ۱۰$ ۷۲
۲۰.۳ برآورد λ برای $n = ۴۰$، $p = ۲۰$، $T = I$ و$df = ۵$ ۷۲
۲۱.۳ برآورد λ برای $n = ۴۰$، $p = ۲۰$، $T = I$ و $df = ۱۰۰$ ۷۳
۲۲.۳ برآوردگر انقباضی برای $p = ۲۰$، $n = ۴۰$، $T = I$ و $df = ۵$ ۷۳
۲۳.۳ برآوردگر انقباضی برای $p = ۲۰$، $n = ۴۰$، $T = I$ و $df = ۱۰$ ۷۳
۲۴.۳ برآوردگر انقباضی برای $p = ۲۰$، $n = ۴۰$، $T = I$ و $df = ۱۰۰$ ۷۳
۲۵.۳ برآورد λ برای $n = ۳۰$، $p = ۳۰$، $T = I$ و $df = ۵$ ۷۳
۲۶.۳ برآورد λ برای $n = ۳۰$، $p = ۳۰$، $T = I$ و$df = ۱۰$ ۷۴
۲۷.۳ برآورد λ برای $n = ۳۰$، $p = ۳۰$، $T = I$ و$df = ۱۰۰$ ۷۴
۲۸.۳ برآوردگر انقباضی برای $p = ۳۰$، $n = ۳۰$، $T = I$ و $df = ۵$ ۷۴
۲۹.۳ برآوردگر انقباضی برای $p = ۳۰$، $n = ۳۰$، $T = I$ و $df = ۱۰$ ۷۴
۳۰.۳ برآوردگر انقباضی برای $p = ۳۰$، $n = ۳۰$، $T = I$ و $df = ۱۰۰$ ۷۴
۳۱.۳ برآورد λ برای $n = ۵$، $p = ۱۰۰$، $T = I$ و$df = ۵$ ۷۵
۳۲.۳ برآورد λ برای $n = ۵$، $p = ۱۰۰$، $T = I$ و$df = ۱۰$ ۷۵
۳۳.۳ برآورد λ برای $n = ۵$، $p = ۱۰۰$، $T = I$ و$df = ۱۰۰$ ۷۵
۳۴.۳ برآوردگر انقباضی برای $p = ۱۰۰$، $n = ۵$، $T = I$ و $df = ۵$ ۷۵
۳۵.۳ برآوردگر انقباضی برای $p = ۱۰۰$، $n = ۵$، $T = I$ و $df = ۱۰$ ۷۵
۳۶.۳ برآوردگر انقباضی برای $p = ۱۰۰$، $n = ۵$، $T = I$ و $df = ۱۰۰$ ۷۶
۳۷.۳ برآورد λ برای $n = ۴۰$، $p = ۲۰$، $T = D_s$ و$df = ۵$ ۷۶
۳۸.۳ برآورد λ برای $n = ۴۰$، $p = ۲۰$، $T = D_s$ و$df = ۱۰$ ۷۶
۳۹.۳ برآورد λ برای $n = ۴۰$، $p = ۲۰$، $T = D_s$ و$df = ۱۰۰$ ۷۶
۴۰.۳ برآوردگر انقباضی برای $p = ۲۰$، $n = ۴۰$، $T = D_s$ و $df = ۵$ ۷۶
۴۱.۳ برآوردگر انقباضی برای $p = ۲۰$، $n = ۴۰$، $T = D_s$ و $df = ۱۰$ ۷۷
۴۲.۳ برآوردگر انقباضی برای $p = ۲۰$، $n = ۴۰$، $T = D_s$ و $df = ۱۰۰$ ۷۷
۴۳.۳ برآورد λ برای $n = ۳۰$، $p = ۳۰$، $T = D_s$ و $df = ۵$ ۷۷

۴۴.۳ برآورد λ برای $n = ۳۰$، $p = ۳۰$، $T = D_s$ و $df = ۱۰$ ۷۷
۴۵.۳ برآورد λ برای $n = ۳۰$، $p = ۳۰$، $T = D_s$ و $df = ۱۰۰$ ۷۷
۴۶.۳ برآوردگر انقباضی برای $p = ۳۰$، $n = ۳۰$، $T = D_s$ و $df = ۵$ ۷۸
۴۷.۳ برآوردگر انقباضی برای $p = ۳۰$، $n = ۳۰$، $T = D_s$ و درجه آزادی $df = ۱۰$ . . ۷۸
۴۸.۳ برآوردگر انقباضی برای $p = ۳۰$، $n = ۳۰$، $T = D_s$ و $df = ۱۰۰$ ۷۸
۴۹.۳ برآورد λ برای $n = ۵$، $p = ۱۰۰$، $T = D_s$ و $df = ۵$ ۷۸
۵۰.۳ برآورد λ برای $n = ۵$، $p = ۱۰۰$، $T = D_s$ و $= ۱۰$ ۷۸
۵۱.۳ برآورد λ برای $n = ۵$، $p = ۱۰۰$، $T = D_s$ و $df = ۱۰۰$ ۷۸
۵۲.۳ برآوردگر انقباضی برای $p = ۱۰۰$، $n = ۵$، $T = D_s$ و $df = ۵$ ۷۹
۵۳.۳ برآوردگر انقباضی برای $p = ۱۰۰$، $n = ۵$، $T = D_s$ و $df = ۱۰$ ۷۹
۵۴.۳ برآوردگر انقباضی برای $p = ۱۰۰$، $n = ۵$، $T = D_s$ و $df = ۱۰۰$ ۷۹

فصل ۱
مقدمه و دورنما

۱.۱ مقدمه

فرض کنید $\mathbf{x}_1, \ldots, \mathbf{x}_n$ بردارهای تصادفی مستقل و هم توزیع (p-بعدی) با ماتریس واریانس-کوواریانس Σ با بعد $p \times p$ باشد. هدف برآورد ماتریس واریانس-کوواریانس Σ برای نمونه‌ی $\{\mathbf{x}_i : i = 1, \ldots, n\}$ می‌باشد. برآورد ماتریس واریانس-کوواریانس، مسئله‌ی اصلی و مهم در تحلیل‌های چندمتغیره است. برای مثال می‌توان به کاربرد آن در تحلیل‌های خوشه‌ای، تحلیل مولفه‌های اصلی، تحلیل ممیزی درجه‌ی دو و خطی و تحلیل‌های رگرسیونی اشاره کرد. در مواردی که بعد داده‌ها بالاست مانند مطالعات آب و هوایی، برآوردگر طبیعی ماتریس کوواریانس نمونه، اغلب ضعیف عمل می‌کند. برای مثال به میرهد [۱] (۱۹۸۷)، جان استون [۲] (۲۰۰۱)، بیکل و لوینا [۳] (۲۰۰۸) و فان و لو [۴] (۲۰۰۸) مراجعه کنید. روش هایی اخیرا برای برآورد بهبودیافته ماتریس واریانس-کوواریانس ارایه شده‌اند. در این راستا می‌توان به روش‌هایی در هانگ [۵] (۲۰۰۶)، زو و همکاران [۶] (۲۰۰۶)، لیم و فان [۷] (۲۰۰۷)، رتمن و همکاران[۸] (۲۰۰۸)، بیکل و لوینا (۲۰۰۸)، ال کاروی [۹] (۲۰۰۸)، وو و پوراحمدی [۱۰] (۲۰۰۹) 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 را زمانی که بعد (p) غیر قابل اغماض (بزرگ) است، بهبود بخشید. استاین (۱۹۶۴) به این موضوع پی برد که بررسی این عقیده بهتر از MLE می‌باشد. نظرات مختلفی در این

[۱] Muirhead
[۲] Johnstone
[۳] Bickel and Levina
[۴] Fan and Lv
[۵] Huang
[۶] Zou
[۷] Lam and Fan
[۸] Rothman
[۹] El Karoui
[۱۰] Wu and Pourahmadi
[۱۱] Lu
[۱۲] Stein

زمینه ارائه شده است که به چندین مورد اشاره می‌کنیم. هاف [۱۳] (۱۹۷۹ و ۱۹۸۱) برآوردگر جدیدی برای ماتریس دقت [۱۴] (Σ^{-1}) مبتنی بر دو تابع زیان مختلف برای توزیع ویشارت ارائه داد. افرون و موریس [۱۵] (۱۹۷۶) و یانگ و برگر [۱۶] (۱۹۹۴) با استفاده از رویکرد بیزی به ترتیب ماتریس دقت و ماتریس کوواریانس را برآورد کردند. دی و سرینیواسان [۱۷] (۱۹۸۵) به برآوردگری دست یافتند که تحت تابع زیان مشخص، مینیماکس است.

اساس کار همه محققین، کاهش متوسط توابع زیان های $L(\hat{\Sigma}, \Sigma)$ 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 واریانس-کوواریانس، برای به‌دست آوردن برآوردگری مناسب، در نظر گرفته‌اند. ایتون و الکین [۱۸] (۱۹۸۷)، رورتو [۱۹] (۲۰۰۰)، اسمیت و کهن [۲۰] (۲۰۰۲)، چن و دانسون [۲۱] (۲۰۰۳)، پوراحمدی و همکاران (۲۰۰۷) و پوراحمدی (۲۰۰۹) از رویکرد تجزیه چولسکی [۲۲] استفاده کردند.

همرل 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 سرال [۲۴] (۱۹۷۶) دنبال شد. هاف (۱۹۹۱) نظر کلی برای برآوردگر بیزی بردار میانگین و ماتریس واریانس-کوواریانس 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 میانگین-واریانس پرداخت. دنیل و پوراحمدی (۲۰۰۲) و دنیل [۲۷] (۲۰۰۵) از رویکرد بیزی توسط کاوش‌های قبلی برای مشاهدات وابسته استفاده کردند. وو [۲۸] و پوراحمدی (۲۰۰۳) برآوردگر ناپارامتری را برای ماتریس واریانس-کوواریانس پیشنهاد نمودند.

دنیل (۲۰۰۶) چندین برآوردگر برای ماتریس کوواریانس مبتنی بر چندین مدل رایج رگرسیون بیزی

---

[۱۳] Haff
[۱۴] Precision matrix
[۱۵] Morris and Efron
[۱۶] Yang and Berger
[۱۷] Dey and Srinivasan
[۱۸] Eaton and Olkin
[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
[۲۸] Wu

پیشنهاد کرد. چادهوری [۲۹] و همکاران (۲۰۰۷) الگوریتمی را برای محاسبه برآورد درستنمایی ماکسیمم ماتریس واریانس-کوواریانس تحت این محدودیت که کوواریانس صفر است، ارائه دادند. فان [۳۰]، فان و لو [۳۱] (۲۰۰۸) در مدل کوواریانس به برآوردگر مناسبی با استفاده از تکنیک‌های تحلیل عاملی چندمتغیره دست یافتند.

کائو و بومن [۳۲] (۲۰۰۹)، ماتریس Σ را مبتنی بر فرضیه‌ای که مقادیر ویژه‌ی Σ با استفاده از تبدیل ماتریس پراکنده ارائه شده است، برآورد نمودند. فان و وو (۲۰۰۸) روش نیمه‌پارامتری 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.

مواردی که بیان شد تعدادی از پیشرفت‌های اخیر در برآورد ماتریس واریانس-کوواریانس و ماتریس دقت می‌باشند. با این حال، درخصوص رده کلی از برآوردگرهای ماتریس واریانس-کوواریانس و ماتریس دقت که به‌طور معمول به نام استاین یا انقباضی یا برآوردگر انقباضی [۳۳] نام گرفته است، مطالعات چندانی صورت نگرفته است. در این پایان‌نامه قصد داریم به مطالعه‌ی برآوردگرهای انقباضی ماتریس واریانس-کوواریانس در توزیع‌های نرمال چندمتغیره و t چندمتغیره بپردازیم.

با توجه به این‌که در فصول مختلف از توزیع‌های آماری و توابع ریاضی متفاوتی استفاده شده است، در این فصل به معرفی این توزیع‌ها و توابع ریاضی می‌پردازیم. از آن‌جایی که توزیع نرمال چندمتغیره و توزیع tچندمتغیره متعلق به خانواده توزیع‌های بیضی‌گون می‌باشند، به معرفی توزیع بیضی‌گون و توزیع آمیخته مقیاسی و ارائه چند مثال از آن‌ها نیز می‌پردازیم.

۲.۱ تعاریف و توابع ریاضی

برای اطلاع دقیق‌تر از تعاریف این بخش می‌توان به رودین [۳۴] (۱۳۷۳)، میرهد (۱۹۸۲) و کرمی (۱۳۹۲) مراجعه کرد.

تعریف ۱.۲.۱. تابع گاما [۳۵]

تابع گاما را به ازای هر عدد حقیقی $a > 0$ با نماد $\Gamma(a)$ نشان داده و به‌صورت زیر تعریف می‌نماییم

$$\Gamma(a) = \int_0^\infty e^{-x} x^{a-1} dx$$

تعریف ۲.۲.۱. مجموعه محدب [۳۶]

مجموعه C محدب است، اگر پاره خط بین هر دو نقطه دلخواه در C، درون C قرار گیرد. به عبارتی به

[۲۹] Chaudhur
[۳۰] Fan
[۳۱] Lv
[۳۲] Cao and Buoman
[۳۳] Shrinkage estimator
[۳۴] Rudin
[۳۵] Gamma function
[۳۶] Convex

ازای هر $x_۱$ و $x_۲$ عضو C و به ازای هر θ ثابت، با شرط $۰ \leq \theta \leq ۱$، داشته باشیم

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

تعریف ۱.۲.۱. تابع محدب[۳۷]

تابع $f : R^n \longrightarrow R$ محدب است، اگر دامنه f یک مجموعه‌ی محدب باشد و به ازای هر X و Y عضو دامنه f و $۰ \leq \theta \leq ۱$ داشته باشیم

$$f(\theta X + (۱ - \theta)Y) \leq \theta f(X) + (۱ - \theta)f(Y)$$

تابع $f(.)$ را مقعر [۳۸] گوییم اگر $-f(.)$ محدب باشد.

تعریف ۲.۲.۱. نرم[۳۹]

تابع $f : R^p \longrightarrow R$ را نرم گوییم، اگر

i) $f(.)$ تابعی نامنفی باشد، یعنی به ازای هر $X \in R^p$ داشته باشیم، $f(X) \geq ۰$،

ii) $f(.)$ معین [۴۰] باشد، یعنی $f(X) = ۰$ اگر و تنها اگر $X = ۰$،

iii) $f(.)$ همگن [۴۱] باشد، یعنی به ازای هر $X \in R^p$ و $t \in R$ داشته باشیم، $f(tX) = |t|f(X)$،

iv) $f(.)$ در نامساوی مثلثی [۴۲] صدق کند، یعنی به ازای هر $X, Y \in R^p$ داشته باشیم

$$f(X + Y) \leq f(X) + f(Y)$$

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

تعریف ۳.۲.۱. نرم ماتریسی

فرض کنید A یک ماتریس $m \times n$ باشد، در این‌صورت نرم ماتریسی $||A||$ با خواص زیر تعریف می‌کنیم

i) $||A|| \geq ۰$، $||A|| = ۰$ اگر و فقط اگر A یک ماتریس صفر باشد،

ii) به ازای هر اسکالر α، $||\alpha A|| = |\alpha|||A||$،

iii) $||A + B|| \leq ||A|| + ||B||$.

تعریف ۴.۲.۱. نرم فروبنیوس [۴۳]

یک نرم ماتریسی مهم، نرم فروبنیوس می‌باشد که به‌صورت زیر تعریف می‌شود.

$$||A||_F = \left[\sum_{j=۱}^{n}\sum_{i=۱}^{m}|a_{ij}|^۲\right]^{\frac{۱}{۲}}$$

در این پایان نامه هر جا سخن از نرم در میان باشد، منظور نرم فروبنیوس است.

---

[۳۷] Convex function
[۳۸] Convave function
[۳۹] Norm
[۴۰] Definit
[۴۱] Homogeneous
[۴۲] Triangle inequality
[۴۳] Frobenius norm

تعریف ۷.۲.۱. عدد شرطی
نسبت بزرگترین مقدار ویژه ماتریس A (λ_{max}) به کوچکترین مقدار ویژه ماتریس A (λ_{min}) را عدد شرطی گوییم به عبارتی داریم

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

تعریف ۸.۲.۱. تابع زیان مربعی
اگر بردار p مولفه‌ای $\delta = (\delta_1, \ldots, \delta_p)$ برآوردگری برای بردار پارامتر $\theta \in R^p$ باشد، آنگاه زیان استفاده از δ در خصوص برآورد θ را با $L(\theta, \delta)$ نشان داده که برابر است با

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

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

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

تعریف ۹.۲.۱. تابع مخاطره
با استفاده از تعریف تابع زیان، تابع مخاطره δ را با $\Re(\theta, \delta)$ نشان داده که برابر است با

$$\Re(\theta, \delta) = E_\theta[L(\theta, \delta)] = \int_\chi L(\theta, \delta) f_\theta(x) dx$$

که در آن $f_\theta(.)$ چگالی تعریف شده روی χ است.

۱.۲.۱ ژاکوبی تبدیل‌ها

فرض کنید $\mathbf{x}$ و $\mathbf{y}$ دو بردار باشند که دارای تعداد مولفه‌های یکسان و مستقل به ترتیب $x_1, \ldots, x_p$ و $y_1, \ldots, y_p$ باشند. تبدیل $\mathbf{y} = F(\mathbf{x})$ را در نظر بگیرید. ژاکوبی تبدیل $\mathbf{x}$ به $\mathbf{y}$ به‌صورت زیر است

$$J(\mathbf{x} \to \mathbf{y}) = \det \begin{pmatrix} \frac{\partial x_1}{\partial y_1} & \cdots & \frac{\partial x_1}{\partial y_p} \\ \vdots & & \vdots \\ \frac{\partial x_p}{\partial y_1} & \cdots & \frac{\partial x_p}{\partial y_p} \end{pmatrix}$$

تعریف ۱۰.۲.۱. 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، بیز و غیره با اطلاعات غیرنمونه‌ای که معمولا در قالب یک یا چند محدودیت به مدل اضافه می‌شوند، به‌صورت برآوردگری جدید بهبود می‌یابد. این بهبود در جهت بهبود برآوردگر اولیه است. در حالت کلی بهبود برآوردگرها را می‌توان با معیار مخاطره یا توان دوم خطا [۴۴] (MSE) اندازه‌گیری کرد. در این‌صورت انقباض می‌تواند به سمت صفر یا یک مقدار ثابت باشد. اثر این انقباض معمولا به‌صورت تبدیل یک برآوردگر نااریب به اریب ولی با MSE کمتر می‌باشد.

[۴۴] Mean squared error

در عین حال، می‌توان برآوردگرهایی انقباضی از نوع استاین و نااریب را نیز یافت. برای آگاهی بیشتر در این خصوص به نوروزی‌راد، (۱۳۹۰) مراجعه کنید.
شکل کلی برآوردگر انقباضی به‌صورت $X + g(X)$ است که در آن g یک تابع اندازه‌پذیر می‌باشد. پس از حل نامعادله

$$\forall \theta \in \Theta : \quad \Re(\theta, X + g(X)) \leq \Re(\theta, X)$$

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

۱- برآوردگر انقباضی به سمت صفر $(۱ - c)X = X - cX$.،

۲- برآوردگر انقباضی به سمت یک عدد ثابت m

$$m + c(X - m) = (۱ - c)m + cX.$$

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

تعریف ۱.۲.۱۱. سازگاری
فرض کنید $X_۱, \ldots, X_n$ 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 مشاهده‌ی اول باشد. هدف از سازگاری T_n بررسی رفتار حدی دنباله $\{T_n\}_{n=۱}^{\infty}$ از برآوردگرهاست. گوییم دنباله‌ی 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 به طوری 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.

تعریف ۱.۲.۱۲. متغیر تصادفی تباهیده [۴۵]
X را متغیر تصادفی تباهیده می‌نامیم، در صورتی که $a \in R$ وجود داشته باشد به طوری‌که $P(X = a) = ۱$. چنانچه متغیر تصادفی تباهیده نباشد، آن را متغیر تصادفی ناتباهیده [۴۶] گوییم.

تعریف ۱.۲.۱۳. تابع توزیع تباهیده [۴۷]
تابع توزیع یک متغیر تصادفی تباهیده را تابع توزیع تباهیده می‌نامیم. اگر Fتابع توزیع تباهیده باشد، آنگاه $a \in R$ وجود دارد به‌طوری که

$$F(x) = \begin{cases} ۰ & x < a \\ ۱ & x \geq a \end{cases}$$

[۴۵] Degenerate random variable
[۴۶] Non-Degenerate random variable
[۴۷] Degenerate distribution function

چنانچه تابع توزیع تباهیده نباشد، آن را تابع توزیع ناتباهیده [۴۸] گوییم.

لم ۱۴.۲.۱. *(شفر و استریمر، ۲۰۰۵) فرض کنید $x_1, \ldots, x_k$ نمونه‌ای از متغیر تصادفی X_i باشد. هم چنین فرض کنید x_{ki}، k امین مشاهده‌ی متغیر X_i و $\bar{x}_i = \frac{1}{n}\sum_{k=1}^{n} x_{ki}$ میانگین نمونه باشد. همچنین فرض کنید $S_{ij} = \frac{1}{n-1}\sum_{i=1}^{N}(x_{ki} - \bar{x}_i)(x_{kj} - \bar{x}_j)^T$، S_{ii} عناصر قطری ماتریس کوواریانس نمونه، $w_{kij} = (x_{ki} - \bar{x}_i)(x_{kj} - \bar{x}_j)$ و $\bar{w}_{kij} = \frac{1}{n}\sum_{k=1}^{n} w_{kij}$ باشد. آن‌گاه کوواریانس تجربی به‌صورت زیر است*

$$\widehat{Cov}(X_i, X_j) = S_{ij} = \frac{n}{n-1}\bar{w}_{ij}$$

و واریانس نیز به‌صورت

$$\widehat{Var}(x_i) = S_{ii} = \frac{n}{n-1}\bar{w}_{ii}$$

می‌باشد. همچنین به‌صورت مشابه داریم

$$\begin{aligned}\widehat{Var}(S_{ij}) &= \frac{n^2}{(n-1)^2}\widehat{Var}(\bar{w}_{ij}) \\ &= \frac{n}{(n-1)^2}\widehat{Var}(w_{ij}) \\ &= \frac{n}{(n-1)^2}\sum_{k=1}^{n}(w_{kij} - w_{ij})^2.\end{aligned}$$

۳.۱ توزیع‌های آماری

در این بخش توزیع‌های آماری را که در این مجموعه استفاده شده‌اند، آورده‌ایم. برای اطلاع دقیق‌تر از تعاریف و قضایای این بخش می‌توان به ماردیا [۴۹] (۱۳۷۶)، میرهد (۱۹۸۲)، گوپتار و ناگار [۵۰] (۲۰۰۰) و اندرسن [۵۱](۲۰۰۳) مراجعه کرد.

تعریف ۱.۳.۱. توزیع نرمال یک متغیره

متغیر تصادفی X دارای توزیع نرمال یک‌متغیره با میانگین μ و واریانس σ^2 است و به‌صورت $X \sim N(\mu, \sigma^2)$ نشان داده می‌شود، اگر تابع چگالی احتمال آن به‌صورت زیر باشد

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} exp\{-\frac{1}{2\sigma^2}(x-\mu)^2\}, \quad -\infty < x < +\infty, \quad -\infty < \mu < +\infty, \quad \sigma > 0$$

تعریف ۲.۳.۱. توزیع گاما

متغیر تصادفی X دارای توزیع گامای یک‌متغیره با پارامترهای a و b است و با نماد $X \sim G(a, b)$ نشان داده می‌شود، اگر تابع چگالی احتمال آن به‌صورت زیر باشد

$$f(x) = \frac{1}{\Gamma(a)b^a} x^{a-1} e^{-\frac{x}{b}}, \quad x > 0, \quad a > 0, \quad b > 0$$

[۴۸] Non-Degenerate distribution function
[۴۹] Mardia
[۵۰] Gupta and Nagar
[۵۱] Anderson

که $E(X) = ab$، $var(X) = ab^۲$ و $\Gamma(a)$ تابع گاما است.
اگر $a = \frac{n}{۲}$ و $b = \frac{۱}{۲}$ آن گاه X دارای توزیع کای‌دو با n درجه آزادی است و می‌نویسیم

$$X \sim G(\frac{n}{۲}, \frac{۱}{۲}) \equiv \chi^۲_{(n)}$$

تعریف ۳.۳.۱. توزیع t استودنت[۵۲]
متغیر تصادفی X دارای توزیع t استودنت با n درجه آزادی می‌باشد و به‌صورت $X \sim t_{(n)}$ نشان داده می‌شود، اگر تابع چگالی احتمال آن به‌صورت زیر باشد

$$f(x) = \frac{۱}{\Gamma(\alpha)\sigma^\alpha} x^{-\alpha-۱} \exp\left(\frac{-۱}{\sigma x}\right), \quad x > ۰$$

که $E(X) = ۰$ و $var(X) = \frac{n}{n-۲}$ می‌باشد.

تعریف ۴.۳.۱. توزیع گامای معکوس
متغیر تصادفی X دارای توزیع گامای معکوس با پارامترهای α و σ است و به‌صورت $X \sim InVG(\alpha, \sigma)$ نشان داده می‌شود، اگر تابع چگالی احتمال آن به‌صورت زیر باشد

$$f(x) = \frac{۱}{\Gamma(\alpha)\sigma^\alpha} x^{-\alpha-۱} \exp\left(\frac{-۱}{\sigma x}\right), \quad x > ۰, \quad \alpha > ۰ \quad > \sigma > ۰$$

که $E(X) = \frac{۱}{(\alpha-۱)\sigma}$ و $var(X) = \frac{۱}{\sigma^۲(\alpha-۱)^۲(\alpha-۲)}$ می‌باشد.
اگر X دارای توزیع گاما باشد، $\frac{۱}{W}$ دارای توزیع گامای معکوس است.

تعریف ۵.۳.۱. توزیع بتا [۵۳]
متغیر تصادفی X دارای توزیع بتا با پارامترهای p و k است و به‌صورت $B(p,k)$ نمایش داده می‌شود، اگر تابع چگالی احتمال آن به‌صورت زیر باشد.

$$f(x) = \begin{cases} \frac{۱}{B(p,k)} x^{p-۱}(۱-x)^{k-۱}, & ۰ < x < ۱, \\ ۰, & \text{سایر مقادیر.} \end{cases}$$

که $E(X) = \frac{p}{p+k}$ و $var(X) = \frac{pk}{(p+k+۱)(p+k)^۲}$ می‌باشد و همچنین $B(p,k)$ تابع بتا است.

۱.۳.۱ توزیع‌های چندمتغیره

در این قسمت، توزیع‌های چندمتغیره مورد استفاده در این پایان‌نامه را به طور مختصر بیان خواهیم کرد.

تعریف ۶.۳.۱. توزیع نرمال چندمتغیره
بردار p مولفه‌ای $\mathbf{x}$ دارای توزیع نرمال p-متغیره با میانگین $\mu \in R^p$ و ماتریس کوواریانس Σ است و با نماد $\mathbf{x} \sim N_p(\mu, \Sigma)$ نمایش می‌دهند اگر و فقط اگر به ازای هر بردار p مولفه‌ای ثابت a، $a^T\mathbf{x}$ دارای نرمال یک متغیره باشد.
اگر Σ ماتریسی معین مثبت باشد [۵۴]، آن گاه $\mathbf{x}$ دارای تابع چگالی احتمال به‌صورت زیر است

$$f(x) = \frac{|\Sigma|^{\frac{۱}{۲}}}{(۲\pi)^{\frac{p}{۲}}} exp\left\{-\frac{۱}{۲}(x-\mu)^T\Sigma^{-۱}(x-\mu)\right\}$$

[۵۲] Student s t distribution
[۵۳] Beta distribution
[۵۴] Positive definite matrix

که در آن $|\Sigma|^{\frac{1}{2}}$ ریشه دوم دترمینان Σ است. توجه کنید اگر $p = 1$، آن گاه صورت درجه دوم در توان نمایی، به $(x-\mu)^2/\sigma^2$ تبدیل می شود و $\mathbf{x} \sim N_p(\mu, \Sigma)$. همچنین اگر Σ یک ماتریس نیمه معین مثبت باشد، آن گاه $\mathbf{x}$ دارای توزیع تباهیده است (ایرانمش، ۱۳۹۱). برای مثال توزیع یک متغیره $N(0, 0)$ تباهیده در صفر است.

تعریف ۷.۳.۱. توزیع بیضی‌گون [۵۵]

بردار n مولفه‌ای $\mathbf{x}$ دارای توزیع بیضی‌گون با پارامترهای μ و Σ و تابع مولد مشخصه ψ است و با $\mathbf{x} \sim EC_n(\mu, \Sigma, \psi)$ نشان می‌دهیم اگر تابع مشخصه آن به ازای بردار n مولفه‌ای t به‌صورت زیر باشد

$$\Phi_{\mathbf{x}}(t) = \exp(it\mu)\psi\left(\frac{t^T\Sigma t}{2}\right)$$

که در آن $\psi(.)$ متعلق به کلاس توابعی به‌صورت $\psi(.) : [0, \infty) \to R$ است به طوری که $\psi(\sum_{i=1}^{n} t_i^2)$ یک تابع مشخصه n-بعدی می‌باشد.

توزیع‌های بیضی گون، توزیع‌هایی هستند که منحنی‌های تراز آن‌ها بیضی شکل می‌باشند (آرشی، ۱۳۸۷).

در صورت وجود تابع چگالی آن، به‌صورت زیر می‌باشد

$$f_{\mathbf{x}}(x) = c_n|\Sigma|^{-\frac{1}{2}} g\left[\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\right]$$

که در آن تابع $g(.)$ تابع مولد چگالی [۵۶] بوده و دارای شرط همگرایی زیر است.

$$\int_0^\infty x^{\frac{n}{2}-1} g(x)dx < \infty$$

به علاوه c_n ضریب نرمال‌سازی می‌باشد که به‌صورت زیر تعیین می‌گردد

$$c_n = \frac{\gamma(\frac{n}{2})}{(2\pi)^{\frac{n}{2}}}\left[\int_0^\infty x^{\frac{n}{2}-1} g(x)dx\right]^{-1}$$

در این صورت، می‌نویسیم $\mathbf{x} \sim EC_n(\mu, \Sigma, \psi)$. شایان ذکر است در صورتی که تابع چگالی وجود داشته باشد، تابع g، تابع ψ را مشخص می‌کند و برعکس. برای آگاهی بیشتر در این زمینه فنگ [۵۷] و همکاران (۱۹۹۰) و آرشی (۱۳۸۷) را ببینید.

با توجه به تعریف ۷.۳.۱، $E(\mathbf{x}) = \mu$ و $Cov(\mathbf{x}) = -2\psi'(0)\Sigma$ که در آن $\psi'(0)$ مشتق اول ψ در نقطه صفر است و در صورتی وجود دارد که $|\psi'(0)| < \infty$.

توزیع‌های زیادی متعلق به خانواده توزیع‌های بیضی‌گون می‌باشند که از آن جمله می‌توان به توزیع‌های نرمال چندمتغیره، t چندمتغیره [۵۸] (Mt)، کوشی چندمتغیره [۵۹]، پیرسن چندمتغیره نوع II [۶۰]، پیرسن چندمتغیره نوع IIV [۶۱]، بسل چندمتغیره [۶۲]، نمایی توانی چندمتغیره [۶۳]، لاپلاس چندمتغیره [۶۴]، اسلش

[۵۵] Elliptically distribution
[۵۶] Density generator function
[۵۷] Fang
[۵۸] Multivatriate t distribution
[۵۹] Multivariate Cauchy distribution
[۶۰] Multivariate Pearson type II
[۶۱] Multivariate Pearson type IIV
[۶۲] Multivariate Bessel
[۶۳] Multivariate exponential power
[۶۴] Multivariate Laplase

تعمیم یافته [۶۵]، و کاتز چندمتغیره اشاره کرد.

حال به عنوان نمونه، تابع چگالی احتمال و تابع مولد چگالی را برای چند توزیع مهم بیضی‌گون، همراه با نمودار آن‌ها ارائه می‌دهیم. تمامی شکل‌های این مثال با نرم‌افزار R-۳.۰.۱ و با استفاده از بسته $lattice$ رسم شده‌اند. در پیوست آ دستورهای مورد نیاز برای رسم شکل‌ها در نرم‌افزار آمده‌اند.
دقت کنید در تمامی این نمودارها $n = ۲$، $\mu = \begin{bmatrix} ۰ \\ ۰ \end{bmatrix}$ و $\Sigma = \begin{bmatrix} ۱ & ۰ \\ ۰ & ۱۰ \end{bmatrix}$ فرض شده‌اند.

۱. توزیع نرمال چندمتغیره: توابع مولد چگالی و چگالی احتمال به ترتیب عبارتند از

$$g(u) = e^{-u}$$

$$f_X(x) = \frac{|\Sigma|^{-\frac{۱}{۲}}}{(\sqrt{۲\pi})^n} \exp\left\{-\frac{۱}{۲}(x-\mu)^T\Sigma^{-۱}(x-\mu)\right\}$$

نمودارهای تابع چگالی توزیع نرمال دومتغیره و منحنی‌های تراز آن در شکل ۱.۱ داده شده است.

(ب) (آ)

شکل ۱.۱: نمودارهای تابع چگالی و منحنی‌های تراز توزیع نرمال دومتغیره

۲. توزیع t چندمتغیره: توابع مولد چگالی و چگالی احتمال به ترتیب عبارتند از

$$g_n(u) = \left(۱ + \frac{۲u}{\nu}\right)^{-\frac{(n+\nu)}{۲}}$$

$$f_X(x) = \frac{\Gamma\left(\frac{\nu+n}{۲}\right)|\Sigma|^{-\frac{۱}{۲}}}{(\nu\pi)^{\frac{n}{۲}}\Gamma\left(\frac{\nu}{۲}\right)}\left[۱ + \frac{۱}{\nu}(x-\mu)^T\Sigma^{-۱}(x-\mu)\right]^{-\frac{(\nu+n)}{۲}} \qquad (۱.۱)$$

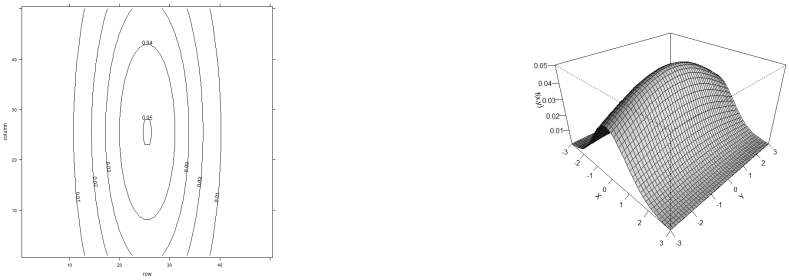
نمودارهای تابع چگالی توزیع t دومتغیره و منحنی تراز به ازای $\nu = ۳$ در شکل ۲.۱ داده شده است.

۳. توزیع کوشی چندمتغیره: به ازای $\nu = ۱$، توزیع t به توزیع کوشی چندمتغیره تبدیل می‌شود.
نمودارهای تابع چگالی توزیع کوشی دومتغیره و منحنی‌های تراز آن در شکل ۳.۱ داده شده است.

[۶۵] Generalized Slash

(ب) (آ)

شکل ۲.۱: نمودارهای تابع چگالی و منحنی‌های تراز توزیع نرمال دومتغیره



(ب) (آ)

شکل ۳.۱: نمودارهای تابع چگالی و منحنی‌های تراز توزیع کوشی دومتغیره

۴. توزیع لجستیک چندمتغیره: توابع مولد چگالی و چگالی احتمال به ترتیب عبارتند از

$$g(u) = \frac{e^{-u}}{(۱ + e^{-u})^۲}$$

$$f_X(x) = \frac{\left(\sum_{j=۱}^{\infty} (-۱)^{j-۱} j^{۱-\frac{n}{۲}}\right)^{-۱} |\Sigma|^{-\frac{۱}{۲}}}{(۲\pi)^{-\frac{n}{۲}}} \times \frac{\exp\left[-\frac{۱}{۲}(x-\mu)^T \Sigma^{-۱}(x-\mu)\right]}{\left\{۱ + \exp\left[-\frac{۱}{۲}(x-\mu)^T \Sigma^{-۱}(x-\mu)\right]\right\}^۲}$$

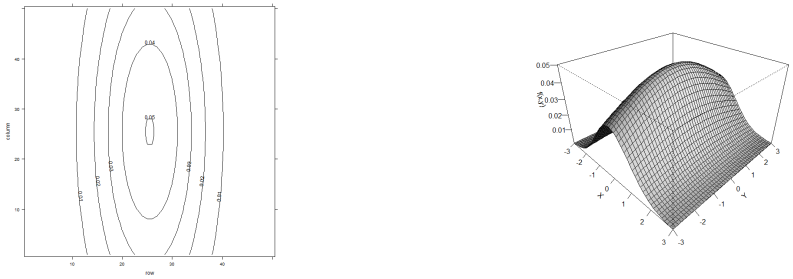
نمودارهای تابع چگالی توزیع لجستیک دومتغیره و منحنی‌های تراز آن در شکل ۴.۱ داده شده است.

۵. توزیع نمایی توانی چندمتغیره: توابع مولد چگالی و چگالی احتمال به ترتیب عبارتند از

$$g(u) = e^{-ru^s}, \quad r, s > ۰$$

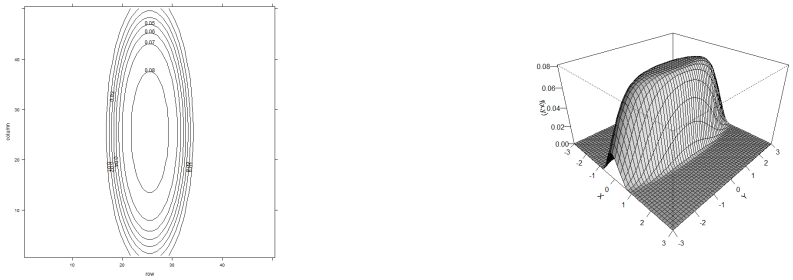
$$f_X(x) = \frac{s\Gamma\left(\frac{n}{۲}\right) |\Sigma|^{-\frac{۱}{۲}} r^{\frac{n}{۲s}}}{(۲\pi)^{\frac{n}{۲}} \Gamma\left(\frac{n}{۲s}\right)} \exp\left\{-\frac{r}{۲}\left[-\frac{۱}{۲}(x-\mu)^T \Sigma^{-۱}(x-\mu)\right]^s\right\}$$

نمودارهای تابع چگالی و منحنی‌های تراز به ازای $r = s = ۳$ در شکل ۵.۱ داده شده است.



(آ) (ب)

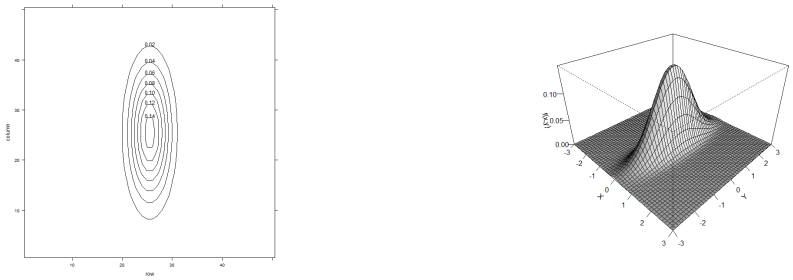
شکل ۴.۱: نمودارهای تابع چگالی و منحنی‌های تراز توزیع لجستیک دومتغیره



(آ) (ب)

شکل ۵.۱: نمودارهای تابع چگالی و منحنی‌های تراز توزیع نمایی توانی دومتغیره

۶. توزیع کاتز چندمتغیره: به ازای $s = ۱$، توزیع نمایی توانی به نوع کاتز تبدیل می‌شود. نمودارهای تابع چگالی و منحنی‌های تراز به ازای $r = ۳$ در شکل ۶.۱ داده شده است.



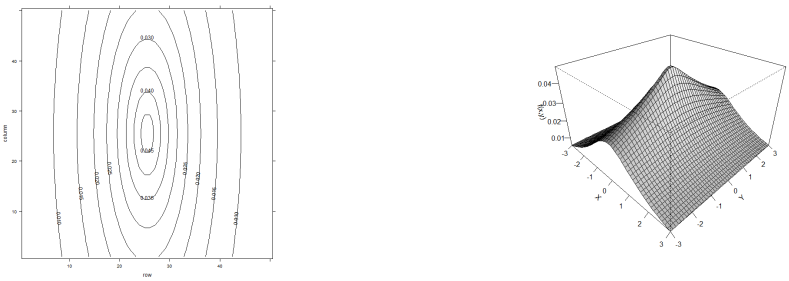
(آ) (ب)

شکل ۶.۱: نمودارهای تابع چگالی و منحنی‌های تراز توزیع کاتز دومتغیره

۷. توزیع لاپلاس (نمایی دوگانه) چندمتغیره: به ازای $s = \frac{۱}{۲}$ و $r = \sqrt{۲}$، توزیع نمایی توانی به توزیع لاپلاس تبدیل می‌شود. نمودارهای تابع چگالی و منحنی‌های تراز در شکل ۷.۱ داده شده است.

خواص توزیع t چندمتغیره

۱. اگر $\mu = ۰$ توزیع t مرکزی نامیده می‌شود و اگر $\mu \neq ۰$ توزیع t غیرمرکزی نامیده می‌شود.



(ب) (آ)

شکل ۷.۱: نمودارهای تابع چگالی و منحنی‌های تراز توزیع لاپلاس دومتغیره

۲. اگر $n = ۱$ و $\mu = ۰$ و $\Sigma = ۱$ آن‌گاه رابطه‌ی (۱.۱) تابع چگالی توزیع t یک متغیره با درجه آزادی ν می‌باشد. تابع چگالی توزیع t یک متغیره به‌صورت زیر است.

$$f(x) = \frac{\Gamma(\frac{\nu+۱}{۲})}{(\pi\nu)^{\frac{۱}{۲}}\Gamma(\frac{\nu}{۲})}\left(۱ + \frac{۱}{\nu}x^{۲}\right)^{-\frac{\nu+۱}{۲}}$$

۳. اگر $\nu = ۱$ آن گاه رابطه (۱.۱) توزیع کوشی n متغیره [۶۶]نامیده می‌شود و تابع چگالی آن به‌صورت زیر است

$$f(x) = \frac{\Gamma(\frac{۱+p}{۲})}{\pi^{\frac{p}{۲}}\Gamma(\frac{۱}{۲})\det(\Sigma)^{\frac{۱}{۲}}}\left[۱ + (x-\mu)'\Sigma^{-۱}(x-\mu)\right]^{-\frac{۱+p}{۲}} \qquad (۲.۱)$$

۴. حد رابطه (۱.۱) وقتی $\nu \to \infty$ برابر تابع چگالی توام توزیع نرمال n متغیره با بردار میانگین μ و ماتریس کوواریانس Σ است. بنابراین رابطه (۱.۱) به طور تقریبی نشان‌دهنده‌ی توزیع نرمال چندمتغیره می‌باشد.

قضیه ۸.۳.۱. *(آرمانفر، ۱۳۸۹) اگر $u \sim \chi^{۲}_{n}$ و $y \sim N_p(۰, \Sigma)$ باشند و با فرض اینکه u و y مستقل هستند، داریم*

$$x = \left(\frac{u}{n}\right)^{-\frac{۱}{۲}} y + \mu \sim t_p(n, \mu, \Sigma) \qquad (۳.۱)$$

برهان.

$$\begin{aligned} f(y,u) &= f(y)f(u) \\ &= \frac{۱}{(۲\pi)^{\frac{p}{۲}}\det(\Sigma)^{\frac{۱}{۲}}}\exp\left(-\frac{۱}{۲}y^T\Sigma^{-۱}y\right)\frac{۱}{۲^{\frac{n}{۲}}\Gamma(\frac{n}{۲})}u^{\frac{n}{۲}-۱}\exp\left(-\frac{u}{۲}\right) \\ &= \frac{۱}{(۲\pi)^{\frac{p}{۲}}۲^{\frac{n}{۲}}\Gamma(\frac{n}{۲})\det(\Sigma)^{\frac{۱}{۲}}}u^{\frac{n}{۲}-۱}\exp\left(-\frac{u}{۲}\right)\exp\left(-\frac{۱}{۲}y^T\Sigma^{-۱}y\right) \end{aligned}$$

حال با تغییر متغیر $x = (\frac{u}{n})^{-\frac{۱}{۲}}y + \mu$ خواهیم داشت

$$y = \left(\frac{u}{n}\right)^{-\frac{۱}{۲}}(x-\mu) \Longrightarrow \quad J(y \to x) = \left(\frac{u}{n}\right)^{-\frac{p}{۲}}$$

[۶۶] P-Variate Cauchy Distribution

$$
\begin{aligned}
f(x,u) &= \frac{۱}{(۲\pi)^{\frac{p}{۲}}۲^{\frac{n}{۲}}\Gamma(\frac{n}{۲})\det(\Sigma)^{\frac{۱}{۲}}}u^{\frac{n}{۲}-۱}\exp\left(-\frac{u}{۲}\right)\\
&\times\exp\left[-\frac{۱}{۲}(x-\mu)^T(\frac{u}{n})\Sigma^{-۱}(x-\mu)\right]\left(\frac{u}{n}\right)^{\frac{p}{۲}}\\
&= \frac{۱}{(۲\pi)^{\frac{p}{۲}}۲^{\frac{n}{۲}}\Gamma(\frac{n}{۲})\det(\Sigma)^{\frac{۱}{۲}}}u^{\frac{n+p}{۲}-۱}n^{-\frac{p}{۲}}\\
&\times\exp\left\{-\frac{۱}{۲}[(x-\mu)^T(\frac{u}{n})(\Sigma^{-۱})(x-\mu)+u]\right\}\\
&= \frac{۱}{(۲\pi)^{\frac{p}{۲}}۲^{\frac{n}{۲}}\Gamma(\frac{n}{۲})\det(\Sigma)^{\frac{۱}{۲}}}u^{\frac{n+p}{۲}-۱}n^{-\frac{p}{۲}}\\
&\times\exp\left\{-\frac{u}{۲}[۱+\frac{۱}{n}(x-\mu)^T\Sigma^{-۱}(x-\mu)]\right\}
\end{aligned}
$$

با انتگرال‌گیری از $f(x,u)$ نسبت به u داریم

$$
\begin{aligned}
f(x) &= \frac{۱}{(۲\pi)^{\frac{p}{۲}}۲^{\frac{n}{۲}}\Gamma(\frac{n}{۲})\det(\Sigma)^{\frac{۱}{۲}}}n^{-\frac{p}{۲}}\\
&\times\int_{۰}^{\infty}u^{\frac{n+p}{۲}-۱}\exp\left\{-\frac{u}{۲}\left[۱+\frac{۱}{n}(x-\mu)^T\Sigma^{-۱}(x-\mu)\right]\right\}du
\end{aligned}
$$

با تغییر متغیر $t=\frac{u}{۲}\left[۱+\frac{۱}{n}(x-\mu)^T\Sigma^{-۱}(x-\mu)\right]$ می‌توان نتیجه گرفت

$$
\begin{aligned}
u &= \frac{۲t}{\left[۱+\frac{۱}{n}(x-\mu)^T\Sigma^{-۱}(x-\mu)\right]}\\
du &= \frac{۲dt}{\left[۱+\frac{۱}{n}(x-\mu)^T\Sigma^{-۱}(x-\mu)\right]}
\end{aligned}
$$

$$
\begin{aligned}
f(x) &= \frac{۱}{(۲\pi)^{\frac{p}{۲}}۲^{\frac{n}{۲}}\Gamma(\frac{n}{۲})\det(\Sigma)^{\frac{۱}{۲}}}n^{-\frac{p}{۲}}\int_{۰}^{\infty}\left(\frac{۲t}{\left[۱+\frac{۱}{n}(x-\mu)^T\Sigma^{-۱}(x-\mu)\right]}\right)^{\frac{n+p}{۲}}\\
&\times e^{-t}\frac{۲dt}{\left[۱+\frac{۱}{n}(x-\mu)^T\Sigma^{-۱}(x-\mu)\right]}\\
&= \frac{\Gamma(\frac{n+p}{۲})}{(n\pi)^{\frac{p}{۲}}\Gamma(\frac{n}{۲})\det(\Sigma)^{\frac{۱}{۲}}}\left[۱+\frac{۱}{n}(x-\mu)^T\Sigma^{-۱}(x-\mu)\right]^{-\frac{n+p}{۲}}
\end{aligned}
$$

□

در ادامه تنها توزیع با متغیر ماتریسی را که در این پایان‌نامه استفاده شده را به طور مختصر معرفی می‌کنیم.

تعریف ۹.۳.۱. توزیع ویشارت [۶۷]

فرض کنید $W=(w_{ij})$ ماتریس متقارن معین مثبت با احتمال یک باشد، و Σ نیز ماتریس $q\times q$

[۶۷] Wishart distribution

معین مثبت باشد. اگر $m \geq q$ (m عدد صحیح)، آنگاه W دارای توزیع ویشارت با m درجه‌ی آزادی است اگر تابع چگالی پیوسته‌ی $\frac{۱}{۲}q(q+۱)$ عناصر مجزا W (بالا مثلثی گفته می‌شود) به‌صورت زیر باشد

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

که منظور از etr همان e^{trace} می‌باشد و

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

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

$$\Gamma_q(a) = \pi^{\frac{q(q-۱)}{۴}} \prod_{j=۱}^{q} \Gamma\left[a + \frac{۱}{۲}(۱-j)\right]$$

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

لم ۱۰.۳.۱. *(میرهد، ۱۹۸۲) فرض کنید $\mathbf{x}_۱, \ldots, \mathbf{x}_m$ بردارهای مستقل و دارای توزیع $N_q(۰, \Sigma)$ باشد آنگاه $W = \sum_{i=۱}^{m} \mathbf{x}_i\mathbf{x}_i^T$ دارای توزیع ویشارت با m درجه‌ی آزادی و ماتریس مقیاس Σ می‌باشد.*

قضیه ۱۱.۳.۱. *(گوپتار و ناگار، ۲۰۰۰) فرض کنید $\mathbf{x}_۱, \ldots, \mathbf{x}_N$ بردارهای تصادفی مستقل از توزیع $N_p(\mu, \Sigma)$ باشند به‌طوری‌که $\Sigma > ۰$ و $N > p$. اگر داشته باشیم $S = \sum_{i=۱}^{N}(\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})^T$ و $\bar{\mathbf{x}} = \frac{۱}{N}\sum_{i=۱}^{N} \mathbf{x}_i$ آنگاه*

$$S \sim W_p(n, \Sigma), \quad n = N - ۱$$

اکثر توزیع‌های بیضی‌گون را می‌توان به‌صورت یک توزیع نرمال آمیخته مقیاسی در نظر گرفت. در این راستا ابتدا توزیع‌های آمیخته مقیاسی را معرفی کرده و سپس به طور خاص‌تری به بررسی توزیع‌های نرمال آمیخته مقیاسی می‌پردازیم.

۴.۱ توزیع‌های آمیخته مقیاسی

معمول است که توزیع مربوط به خصوصیات مورد بررسی یک جامعه آماری را نرمال فرض کنیم، اما ممکن است همیشه چنین فرضی صحیح نباشد. از این رو محققان بر آن شدند که توزیع نرمال را تعمیم داده و توزیعی را در نظر بگیرند که دارای خصوصیاتی مشابه توزیع نرمال باشد که در آن توزیع نرمال به عنوان یک مولفه اصلی در تابع چگالی وجود دارد. در این راستا در این بخش به معرفی رده‌ای از توزیع‌های آمیخته مقیاسی [۶۸] می‌پردازیم و در ادامه توزیع نرمال آمیخته مقیاسی [۶۹] را که در سال‌های اخیر مورد توجه بسیار قرار گرفته، به عنوان یک حالت خاص از

[۶۸] Scale mixture distribution
[۶۹] Scale mixture of normal distribution

این خانواده، معرفی می‌کنیم. خانواده توزیع‌های آمیخته مقیاسی اولین بار توسط آندرو و مالو [۷۰] (۱۹۷۴) معرفی شد.
فرض کنید T یک متغیر مثبت تک نمایی با تابع توزیع $G(t,\alpha)$ (α می‌تواند یک عدد یا یک بردار باشد) و $f(x|\theta,\beta)$ یک تابع چگالی از خانواده مکانی-مقیاسی یا مقیاسی، با پارامتر مکان θ و پارامتر مقیاس β باشد که β و θ می‌توانند یک عدد یا یک بردار باشند. همچنین لازم به ذکر است که متغیر تصادفی T می‌تواند پیوسته، گسسته یا آمیخته باشد. بنابراین یک مدل آمیخته مقیاسی به‌صورت زیر تعریف می‌شود.

تعریف ۱.۴.۱. (آندرو و مالو، ۱۹۷۴) متغیر تصادفی X دارای توزیع آمیخته مقیاسی است اگر بتوان تابع چگالی آن را به‌صورت زیر نوشت
الف) اگر T یک متغیر تصادفی پیوسته باشد

$$\int_t f(x|\theta,t^{-1}\beta)dG(t,\alpha) \qquad (۴.۱)$$

ب) اگر T یک متغیر تصادفی گسسته باشد

$$\sum_{\{t:\Sigma G(t,\alpha)=1\}} f(x|\theta,t^{-1}\beta) \qquad (۵.۱)$$

که در آن $G(t,\alpha)$ تابع توزیع متغیر تصادفی مثبت T با پارامتر α است.

در ادامه به عنوان عضوی معروف از خانواده آمیخته مقیاسی، مدل نرمال آمیخته مقیاسی را به همراه خصوصیات مهمی از این خانواده معرفی خواهیم نمود.

۵.۱ مدل نرمال آمیخته مقیاسی

خانواده نرمال آمیخته مقیاسی در سال‌های اخیر بسیار مورد توجه قرار گرفته است. این خانواده شامل توزیع‌های مهمی از جمله t-استودنت، اسلش [۷۱]، نمایی-توانی [۷۲] و نرمال آلایشی است. لانگ و سینشمر [۷۳](۱۹۹۳) توزیع‌های نرمال آمیخته مقیاسی را، که شامل گروهی از توزیع‌های دم ضخیم [۷۴] است و اغلب برای مدل‌بندی‌های دقیق در مورد داده‌های متقارن استفاده می‌شوند، گسترش دادند. بنابراین، متقارن بودن و داشتن دم‌های ضخیم از خصوصیات اصلی این خانواده است. این خانواده نسبت به توزیع نرمال، به عنوان یک توزیع متقارن دارای دم های پهن‌تری است. در عمل ممکن است با داده‌هایی سروکار داشته باشیم که انباشتگی در دم‌های هیستوگرام این داده‌ها نسبت به سایر نقاط بیشتر باشد (داده‌های مربوط به درآمد خانوار از این دسته هستند)، در این صورت توزیع‌های خانواده نرمال آمیخته مقیاسی نسبت به توزیع نرمال، می‌توانند مدل مناسب‌تری برای برازش به این نوع داده‌ها باشند.

[۷۰] Andrews and Mallows
[۷۱] Slash distribution
[۷۲] Power-exponential distribution
[۷۳] Lange and Sinsheimer
[۷۴] Thick tailed

تعریف ۱.۵.۱. اگر تابع f معرفی شده در تعریف ۱.۴.۱ با تابع چگالی نرمال n-متغیره جایگزین شود، آن‌گاه تابع چگالی مدل نرمال آمیخته مقیاسی به‌صورت زیر به‌دست می‌آید.

الف) اگر متغیر تصادفی T پیوسته باشد در این صورت داریم

$$\int_t \Phi_n(x|\mu, t^{-1}\Sigma) dG(t, \alpha) \qquad (۶.۱)$$

ب) اگر متغیر تصادفی T گسسته باشد داریم

$$\sum_{t:\sum G(t,\alpha)=1} \phi_n(x|\mu, t^{-1}\Sigma) \qquad (۷.۱)$$

در این صورت می‌گوییم متغیر تصادفی X دارای توزیع نرمال آمیخته مقیاسی با پارامترهای مکان μ، مقیاس Σ و G است و با نماد زیر نمایش می‌دهیم

$$X \sim SMN_p(\mu, \Sigma, G) \qquad (۸.۱)$$

همچنین می‌توان نتیجه گرفت که اگر $X \sim SMN_n(\mu, \Sigma, G)$، آنگاه

$$X|T = t \sim N_n(\mu, t^{-1}\Sigma) \qquad (۹.۱)$$

برای آگاهی بیشتر تورەزاده (۱۳۹۲) را ببینید.

۶.۱ مثال‌هایی از توزیع نرمال آمیخته مقیاسی

حال به معرفی اجمالی سه توزیع مهم از خانواده توزیع‌های نرمال آمیخته مقیاسی، که در تحلیل‌های آماری نقش مهمی ایفا می‌کنند، می‌پردازیم.

۱.۶.۱ توزیع t-استیودنت

یکی از توزیع‌های متعلق به خانواده توزیع‌های نرمال آمیخته مقیاسی، توزیع t-استیودنت است. استفاده از این توزیع در بسیاری از موارد به جای توزیع نرمال پیشنهاد شده است. لانگ و همکاران [۷۵] (۱۹۸۹) برای مدل بندی‌های قوی، توزیع t-استیودنت را به عنوان یک پیشنهاد مناسب به کار برده‌اند. برای آگاهی بیشتر و مشاهده برخی تعمیم‌های این توزیع ایرانمنش (۱۳۹۱) را ببینید.

۲.۶.۱ توزیع اسلش

یکی دیگر از توزیع‌های نرمال آمیخته مقیاسی توزیع اسلش با پارامتر شکل ν است. این توزیع نسبت به توزیع نرمال دارای دم‌های ضخیم‌تری است. زمانی که $\nu \to \infty$، توزیع نرمال به عنوان توزیع حدی اسلش در نظر گرفته می‌شود.

[۷۵] Lange

روش ساخت

فرض کنید متغیر تصادفی T در تعریف ۱.۵.۱، دارای توزیع بتا $B(\nu, ۱)$ باشد، آنگاه تابع چگالی توزیع اسلش n–متغیره به‌صورت زیر معرفی می‌شود

$$f_X(x) = \nu \int_{۰}^{۱} t^{\nu-۱} \Phi_n(x|\mu, t^{-۱}\Sigma) dt$$

و از نماد زیر برای نمایش این توزیع استفاده می‌کنیم

$$X \sim SL_n(\mu, \Sigma, \nu)$$

۳.۶.۱ توزیع نرمال آلایشی

یکی دیگر از خانواده توزیع‌های نرمال آمیخته مقیاسی، توزیع نرمال آلایشی است. معمولا از این توزیع برای مدل‌بندی داده‌های متقارن با مشاهدات پرت نیز استفاده می‌شود.

روش ساخت

فرض کنید متغیر تصادفی T در تعریف ۱.۵.۱ دارای تابع چگالی زیر باشد

$$h(t, \nu) = \nu I_{(t=\gamma)} + (۱ - \nu) I_{(t=۱)}, \quad \nu = (\nu, \gamma)^T$$

آنگاه با استفاده از تعریف ۱.۴.۱ تابع چگالی نرمال آمیخته n–متغیره به‌صورت زیر تعریف می‌شود

$$f_X(x) = \nu \Phi\left(x|\mu, \frac{\Sigma}{\gamma}\right) + (۱ - \nu)\Phi_n(x|\mu, \Sigma)$$

و از نماد زیر برای نمایش این توزیع استفاده می‌کنیم

$$X \sim CN_n(\mu, \Sigma, \nu, \gamma), \quad ۰ \leq \nu \leq ۱, \quad ۰ < \gamma \leq ۱$$

که در آن ν پارامتر نشان‌دهنده درصد مشاهدات دورافتاده و γ پارامتر مقیاس است.

۷.۱ قضایا و لم‌های اساسی

لم ۱.۷.۱. *لم استاین (استاین،۱۹۸۱)*
اگر بردار تصادفی p– بعدی $\mathbf{x}$ دارای توزیع نرمال استاندارد باشد، آنگاه به ازای تمام توابع پیوسته $g : R^p \to R$ که $E|g'| < \infty$، داریم

$$E[g'(\mathbf{x})] = E[\mathbf{x}g(\mathbf{x})] \qquad (۱۰.۱)$$

برای آگاهی بیشتر در خصوص لم استاین و کاربردهای آن کرمی (۱۳۹۲) را ببینید.

فصل ۲

برآوردگر انقباضی ماتریس کواریانس در توزیع نرمال چندمتغیره

۱.۲ مقدمه

یکی از مسائلی که در آمار مورد توجه است مسئله برآورد و یافتن برآوردگر مناسب است. در دنیای پیرامون ما در علوم و زمینه‌های مختلف اعم از اجتماعی، اقتصادی، روانشناختی، سیاسی، زمین‌شناسی، پزشکی، نجوم و غیره با پارامترهای مجهولی روبه‌رو هستیم که همواره برآورد این پارامترها از اهداف اصلی می‌باشد. یافتن برآوردگرهای مناسب در عین سادگی در ظاهر، پیچیدگی‌های خاص خود را داراست و همواره در یافتن آن‌ها نیاز است جوانب امر سنجیده شود. در این فصل به دنبال برآوردگری مناسب برای ماتریس واریانس-کواریانس در توزیع نرمال چندمتغیره زمانی که بعد متغیرها (مشاهدات) بالاست، می‌باشیم. مطالب این فصل عمدتا از منبع سان و فیشر (۲۰۰۹) می‌باشد و مقاله‌ای تحت عنوان « برآوردگر انقباضی ماتریس واریانس-کوواریانس » از آن استخراج شده است.

آمار 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. نمونه‌ی N تایی را از جمعیتی با متغیرهای p-بعدی در نظر بگیرید. اندازه‌گیری‌های حاصل روی یک مشاهده به‌صورت برداری به شکل زیر است

$$\mathbf{x_i} = \begin{pmatrix} x_{i1} \\ x_{i2} \\ \vdots \\ x_{ip} \end{pmatrix}$$

که x_i $(i = 1, \ldots, N)$، i-امین مشاهده از نمونه تصادفی را ارائه می‌دهد. مجموعه تمام اندازه‌گیری‌های حاصل از مشاهدات ماتریس زیر را تشکیل می‌دهد

$$\mathbf{X} = (\mathbf{x_1}, \ldots, \mathbf{x_N})$$

که $\mathbf{X}$ ماتریسی $p \times N$ است.
اگر فرض کنیم توزیع بردار $\mathbf{x_i}$ متعلق به خانواده توزیع‌های احتمال p-بعدی زیر باشد

$$\mathcal{F} = \{f(x;\mu,\Sigma)| \quad \mu \in R^p, \Sigma > ۰\} \tag{۱.۲}$$

آن‌گاه بردار میانگین μ و ماتریس واریانس-کوواریانس Σ به‌صورت زیر می‌باشند

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

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

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

۲.۲ برآورد

به‌طور کلی در خانواده توزیع‌های (۱.۲) علاقمند به برآورد دو پارامتر بردار میانگین μ و ماتریس کوواریانس Σ می‌باشیم. برآورد معمول برای μ، میانگین نمونه $\bar{\mathbf{x}}$ و برای Σ، ماتریس کوواریانس نمونه S می‌باشد که براساس یک اندازه نمونه‌ی $N = n + ۱$ تایی به صورت زیر می‌باشند

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

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

$$S = \frac{۱}{N-۱}\sum_{j=۱}^{N}(\mathbf{x}_j - \bar{\mathbf{x}})(\mathbf{x}_j - \bar{\mathbf{x}})^T \tag{۲.۲}$$

که S را می‌توان به‌صورت زیر نوشت

$$S = \frac{۱}{N-۱}(\mathbf{X} - \bar{\mathbf{X}})(\mathbf{X} - \bar{\mathbf{X}})^T \tag{۳.۲}$$

که $\bar{\mathbf{X}}$ ماتریسی $p \times N$ با هر ستون متشکل از $\bar{\mathbf{x}}$ است.
براساس یافته‌های سان و فیشر [۱] (۲۰۰۹)، $\bar{\mathbf{x}}$ و S به دلیل ناااریبی، سازگاری و برآوردگرهای درستنمایی ماکزیمم بودن، برآوردگرهای مناسبی برای μ و Σ هستند.

[۱] Sun and Fisher

قضیه ۱.۲.۲. *(میرهد، ۱۹۸۲) اگر* $\mathbf{x}_1, \ldots, \mathbf{x}_N$ *بردارهای تصادفی مستقل دارای توزیع* $N_p(\mu, \Sigma)$ *باشند، به‌طوری‌که* $N > p$*، آن‌گاه برآوردگر درستنمایی ماکزیمم* (MLE) μ *و* Σ *به‌صورت* $\hat{\mu} = \bar{\mathbf{x}}$ *و* $\hat{\Sigma} = \frac{n}{N}S$ *می‌باشند که* $N = n + 1$*،*

$$\bar{\mathbf{x}} = \frac{1}{N}\sum_{i=1}^{N}\mathbf{x}_i$$

و S *هم به‌صورت* (۲.۲) *تعریف می‌شوند.*

برهان. تابع درستنمایی بدون در نظر گرفتن ثابت، به‌صورت

$$L(\mu, \Sigma) = \det \Sigma^{-\frac{N}{2}} etr(-\frac{1}{2}\Sigma^{-1}A) exp\{-\frac{1}{2}N(\bar{\mathbf{x}} - \mu)^T\Sigma^{-1}(\bar{\mathbf{x}} - \mu)\}$$

می‌باشد که $L(\mu, \Sigma) \leq \det \Sigma^{-\frac{N}{2}} etr(-\frac{1}{2}\Sigma^{-1}A)$ و همچنین $A = \sum_{i=1}^{N}(\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})^T$. تابع مورد نظر بیشترین مقدار خود را زمانی می‌گیرد که $(\bar{\mathbf{x}} - \mu)^T\Sigma^{-1}(\bar{\mathbf{x}} - \mu) = 0$ و این حالتی است که $\bar{\mathbf{x}} = \mu$ باشد. بنابراین $\bar{\mathbf{x}}$ برآوردگر درستنمایی ماکزیمم μ می‌باشد. از طرفی داریم

$$\begin{aligned} g(\Sigma) = logL(\bar{\mathbf{x}}, \Sigma) &= -\frac{1}{2}Nlog\det\Sigma - \frac{1}{2}tr(\Sigma^{-1}A) \\ &= -\frac{1}{2}Nlog\det\Sigma^{-1}A - \frac{1}{2}tr(\Sigma^{-1}A) - \frac{1}{2}Nlog\det A \\ &= \frac{1}{2}Nlog\det A^{\frac{1}{2}}\Sigma^{-1}A^{\frac{1}{2}} - \frac{1}{2}tr(A^{\frac{1}{2}}\Sigma^{-1}A^{\frac{1}{2}}) - \frac{1}{2}Nlog\det A \\ &= \frac{1}{2}\sum_{i=1}^{p}(Nlog\lambda_i - \lambda_i) - \frac{1}{2}Nlog\det A \end{aligned}$$

که $\lambda_1, \ldots, \lambda_p$ ریشه‌های $A^{\frac{1}{2}}\Sigma^{-1}A^{\frac{1}{2}}$ یعنی $\Sigma^{-1}A$ می‌باشند. چون تابع

$$f(\lambda) = Nlog\lambda - \lambda$$

ماکزیمم مقدار خود را در $\lambda = N$ دارد، نتیجه می‌شود که

$$g(\Sigma) \leq \frac{1}{2}pNlogN - \frac{1}{2}pN - \frac{1}{2}Nlog\det A$$

یا

$$L(\bar{\mathbf{x}}, \Sigma) \leq N^{\frac{Np}{2}}e^{-\frac{Np}{2}}(\det A)^{-\frac{N}{2}}.$$

با توجه به این‌که $\lambda_i = N \; (i = 1, \ldots, p)$، داریم: $A^{\frac{1}{2}}\Sigma^{-1}A^{\frac{1}{2}} = NI_p$، از این‌رو $\Sigma = \frac{1}{N}A$ می‌باشد. بنابراین می‌توان نتیجه گرفت

$$L(\mu, \Sigma) \leq N^{\frac{Np}{2}}e^{\frac{-pN}{2}}(\det A)^{-\frac{N}{2}}$$

اگر و تنها اگر $\mu = \bar{\mathbf{x}}$ و $\Sigma = \frac{1}{N}A$ باشد. □

در حالتی که $n > p$ می‌باشد، طبق روش درستنمایی ماکزیمم،

$$S = \frac{1}{n}(\mathbf{X} - \bar{\mathbf{X}})(\mathbf{X} - \bar{\mathbf{X}})^T$$

برآوردگری مناسب برای Σ است و در این حالت، معین مثبت و نامنفرد می‌باشد. همچنین براساس روش درستنمایی ماکزیمم، برآوردگری نااریب برای Σ است. خاصیت دیگری که S برای Σ دارد، سازگاری آن است. برای جزئیات بیشتر به سان و فیشر (۲۰۰۹) مراجعه شود.
اگرچه S دارای چندین ویژگی مطلوب است اما در عمل ماتریس کوواریانس نمونه (S) دارای تعدادی محدودیت نیز می‌باشد. در بسیاری از کاربردها نیاز به برآوردگری برای ماتریس واریانس-کوواریانس هست که معکوس‌پذیر باشد. ولی برای ابعاد (p) بالا یا زمانی که $n \approx p$ (خیلی بهم نزدیک باشند)، دیگر شرایط خوبی برای ماتریس واریانس-کوواریانس نداریم، زیرا در این حالت ماتریس واریانس-کوواریانس معکوس‌پذیر نخواهد بود و نمی‌توان برآوردگر مناسبی را یافت. برای جزئیات بیشتر به بای و یین [۲] (۱۹۹۳)، لدویت و ولف[۳] (۲۰۰۳) ، شفر و استریمر [۴] (۲۰۰۵) و بیکل و لوینا [۵] (۲۰۰۸) مراجعه کنید. در بسیاری از روش‌های آماری و همچنین کارهایی که توسط لدویت و ولف (۲۰۰۳ و ۲۰۰۴) انجام شده که در آن‌ها بعد بالاست، برآورد دقیقی از ماتریس واریانس-کوواریانس نیاز است.
در این فصل به دنبال آن هستیم تا برای پارامتر واریانس-کوواریانس Σ، (زمانی که بعد p بزرگ است، در حالتی که توزیع موردنظر نرمال چندمتغیره می‌باشد، برآوردگر مناسبی بیابیم.

۳.۲ برآوردگر انقباضی نوع استاین

در این بخش به دنبال برآوردگری مناسب برای ماتریس واریانس-کوواریانس زمانی که حجم نمونه کمتر از تعداد متغیرها $(p > n)$ است، می‌باشیم. رویکردی رایج در بهبود برآورد ماتریس واریانس-کوواریانس در این حالت، «انقباضی» یا «برآورد اریب» می‌باشد که به‌طور گسترده توسط افرون [۶] (۱۹۷۵) و افرون و موریس [۷] (۱۹۷۵) و بررسی شده است.
فرض کنید S برآوردگر ارائه شده در (۲.۲) باشد که میانگین توان‌های دوم خطای آن به‌صورت زیر محاسبه می‌شود

$$MSE(S) = (Bais(S))^T (Bais(S)) + Cov(S)$$

که در آن $Bias(S) = S - E(S)$. برای ماتریس کوواریانس نمونه S اریبی صفر می‌باشد (سان و فیشر، ۲۰۰۹)، از این‌رو بهبود دقت آن در گرو کاهش واریانس است. با تعریف برآوردگری اریب، می‌توان واریانس را کاهش داد، که ایده‌ی اصلی جیمز و استاین [۸] (۱۹۶۱) است. در واقع در این روش ترکیبی محدب از ماتریس کوواریانس نمونه با ماتریسی هدف به‌صورت زیر در نظر

[۲] Bai and Yin
[۳] Ledoit and Wolf
[۴] Schafer and Strimmer
[۵] Bickel and Levina
[۶] Efron
[۷] Moriss
[۸] James and Stien

گرفته می‌شود

$$S^* = (۱ - \lambda)S + \lambda T \qquad (۴.۲)$$

که $\lambda \in (۰, ۱)$ پارامتر (ضریب) انقباضی [۹]نام دارد و T ماتریس هدف [۱۰] است که باید معین مثبت باشد. این ترکیب را طوری در نظر می‌گیریم که S^* واریانس کمتری نسبت به S داشته باشد.

با توجه به این‌که ماتریس هدف T از قبل تعیین شده است، مساله اصلی در ارتباط با برآوردگر انقباضی، تعیین پارامتر انقباضی λ می‌باشد. پارامتر انقباضی λ طوری انتخاب می‌شود که برآوردگر S را بهبود بخشد. اگر n در مقایسه با p بزرگ باشد، S باید واریانس کمتری داشته باشد و باید برآوردگر خوبی باشد، از این‌رو $\lambda \longrightarrow ۰$. به‌طور مشابه اگر p بزرگ باشد، S واریانس بزرگتری دارد و باید وزنی را برای ماتریس هدف فراهم کرد به طوری که $\lambda \longrightarrow ۱$. انتخاب پارامتر انقباضی (λ) را می‌توان از روش بوت‌استرپ [۱۱]، روش اعتبارسنجی متقابل [۱۲] و روش‌های دیگری تعیین کرد. یادآوری می‌کنیم که در این فصل به‌طور عمده به دنبال برآوردگر انقباضی ماتریس کوواریانس با در نظر گرفتن مینیمم میانگین مربع خطا می‌باشیم.

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، همچنان‌که بزرگترین آن به سمت بی‌نهایت میل می‌کند. ایده نظری، انقباض مقادیر ویژه‌ی S به مقادیر ویژه T می‌باشد. ماتریس هدف انتخاب شده 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. S^* (برآوردگر انقباضی) برای هر بعدی، ماتریس معین مثبت و نامنفرد (معکوس‌پذیر) می‌باشد.

لدویت و ولف (۲۰۰۳) قضیه‌ای را ارائه دادند که بیانگر پارامتر انقباضی مطلوب از نظر واریانس-کوواریانس در ماتریس کواریانس نمونه و ماتریس هدف می‌باشد. تحت میانگین توان‌های دوم خطا و با در نظر گرفتن نرم فروبنیوس، یک مقدار بهینه برای پارامتر انقباضی وجود دارد. لدویت و ولف (۲۰۰۴b) جزئیات پارامتر بهینه را برای ماتریس هدف $T = \mu_\lambda I$، در نظر گرفتند که μ_λ میانگین مقادیر ویژه‌ی Σ و I ماتریس همانی می‌باشد. پارامتر انقباضی بهینه در نظر گرفته شده به‌صورت زیر می‌باشد

$$\lambda = \frac{\beta^۲}{\delta^۲} \qquad (۵.۲)$$

که در آن

$$\beta^۲ = E[||S - \Sigma||^۲] \qquad (۶.۲)$$

و

$$\delta^۲ = E[||S - \mu_\lambda I||^۲]. \qquad (۷.۲)$$

[۹] Shrinkage intensity
[۱۰] Target matrix
[۱۱] Bootstrap
[۱۲] Cross-validation

متاسفانه μ_λ، $\beta^۲$ و $\delta^۲$ همه به اطلاعاتی در مورد Σ نیاز دارند و از آن‌جایی که مجهول هستند باید برآورد شوند.
لدویت و ولف (۲۰۰۴b) برآوردگرهای سازگاری برای μ_λ و λ فراهم کردند. اخیرا چن و همکاران [۱۳] (۲۰۰۹) با در نظر گرفتن قضیه رائو-بلکول، برآوردگر لدویت و ولف (۲۰۰۴b) را بهبود بخشیدند. شفر و استریمر (۲۰۰۵) شاخص ناایبی را نسبت به سازگاری در برآوردگرهای μ_λ و λ در نظر گرفتند.

۴.۲ انواع پارامترهای انقباضی

فرض کنید $\{\mathbf{x}_i\}_{i=۱}^n$ نمونه‌ای تصادفی از توزیع $p-$متغیره نرمال با بردارهای میانگین صفر و کوواریانس Σ باشد. نکته‌ای که باید بدان توجه نمود این است که فرض $n \geq p$ را نداریم. به دنبال برآوردگری مانند $\hat{\Sigma}$ هستیم که تابع زیان زیر را مینیمم کند

$$E[||\hat{\Sigma}(\{\mathbf{x}_i\}_{i=۱}^n) - \Sigma||_F^۲] \qquad (۸.۲)$$

که در آن $\hat{\Sigma}(\{\mathbf{x}_i\})$ برآوردگری برای Σ می‌باشد و اندیس Fنشانگر نرم فروبنیوس است.
در واقع باید پارامتر انقباضی λ را (از روی مشاهدات) طوری یافت که تابع زیان (۸.۲) را مینیمم کند. حال به معرفی پارامترهای انقباضی می‌پردازیم.

۱.۴.۲ برآورد انقباضی LW

لدویت و ولف (۲۰۰۴) تحت فرض این‌که $\{\mathbf{x}_i\}_{i=۱}^n$ دارای توزیع نرمال $p-$متغیره با میانگین صفر و ماتریس کوواریانس Σ می‌باشند، پارامتر انقباضی λ را به‌صورت زیر معرفی نمودند

$$\lambda_{LW} = min\left\{\frac{\sum_{i=۱}^N ||\mathbf{x}_i\mathbf{x}_i^T - \hat{S}||_F^۲}{n^۲[tr(\hat{S})^۲] - \frac{۱}{p}tr^۲(\hat{S})}, ۱\right\} \qquad (۹.۲)$$

که $\hat{S}$ به‌صورت زیر می‌باشد

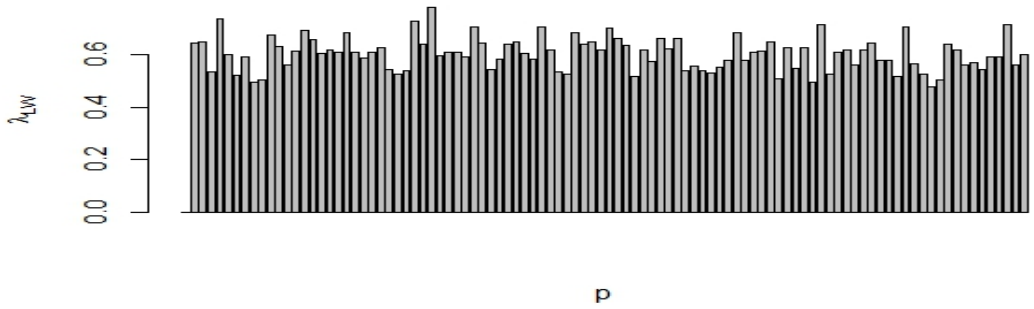
$$\hat{S} = \frac{۱}{N}\sum_{i=۱}^N \mathbf{x}_i\mathbf{x}_i^T$$

نمودار این پارامتر را بر اساس یک نمونه‌ی شبیه‌سازی شده Nتایی بر حسب p $(p = ۲, \ldots, ۱۰۰)$ رسم نموده‌ایم که شکل ۱.۲ حاصل شده است

۲.۴.۲ برآورد انقباضی $RBLW$

چن و همکاران (۲۰۰۹) با در نظر گرفتن قضیه رائوبلکول، برآوردگر لدویت و ولف (۲۰۰۴) را بهبود بخشیدند. آن‌ها تحت فرضیه‌ی نرمال بودن داده‌ها و همچنین با در نظر گرفتن S به عنوان

[۱۳] Chen

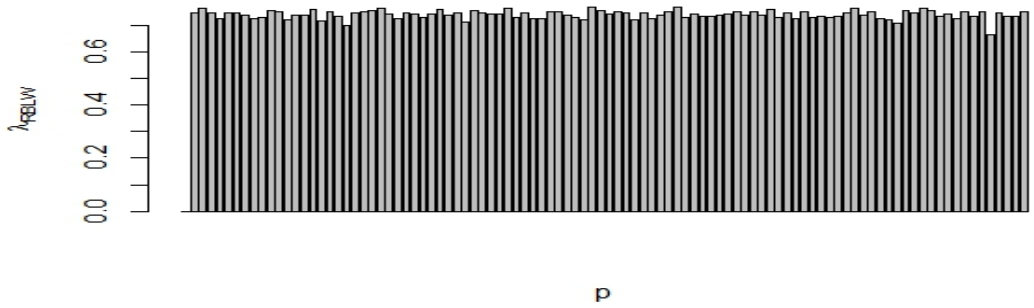


شکل ۱.۲: پارامتر انقباضی LW

آماره بسنده ضریب انقباضی را به‌صورت زیر تعریف نمودند

$$\lambda_{RBLW} = min\left\{\frac{\frac{(n-۲)}{n}tr(\hat{S}^۲) + tr^۲(\hat{S})}{(n+۲)[tr(\hat{S}^۲) - \frac{۱}{p}tr^۲(\hat{S})]}, ۱\right\} \qquad (۱۰.۲)$$

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 نمودند.
نمودار این پارامتر را بر اساس یک نمونه‌ی شبیه‌سازی شده Nتایی بر حسب p $(p = ۲, \ldots, ۱۰۰)$ رسم نموده‌ایم که شکل ۲.۲ به‌دست آمده است



شکل ۲.۲: پارامتر انقباضی $RBLW$

۳.۴.۲ برآورد انقباضی SS

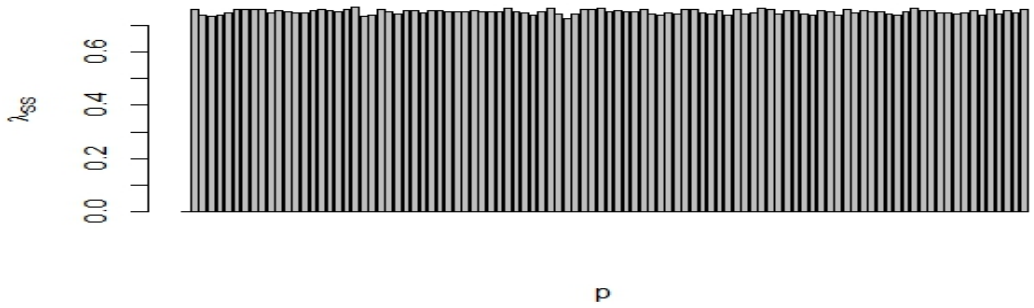
شفر و استریمر (۲۰۰۵) نیز تحت فرض نرمال بودن داده‌ها، پارامتر انقباضی را به‌صورت زیر ارائه دادند

$$\lambda_{SS} = min\left\{\frac{\sum_{i\neq j}\widehat{var}(S_{ij} + \sum_i \widehat{var}(S_{ii}))}{\sum_{i\neq j} S_{ij}^۲ + \sum_i (S_{ii} - \mu_\lambda)^۲}, ۱\right\} \qquad (۱۱.۲)$$

که در آن

$$S_{ij} = \frac{۱}{N-۱}\sum_{i=۱}^{N}(X_{ki} - \bar{X}_i)(X_{kj} - \bar{X}_j)^T$$

و S_{ii} عناصر قطری ماتریس کوواریانس نمونه S و $\mu_\lambda = ave(S_{ii})$ می‌باشد. برای جزئیات بیشتر به شفر و استریمر (۲۰۰۵) مراجعه شود. نمودار این پارامتر را بر اساس یک نمونه‌ی شبیه‌سازی شده Nتایی بر حسب p $(p = ۲, \ldots, ۱۰۰)$ رسم نموده‌ایم که شکل ۳.۲ به‌دست آمده است.



شکل ۳.۲: پارامتر انقباضی SS

برآورد انقباضی FS

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

$$\lambda_{FS} = min\left\{\frac{\frac{۱}{n}\hat{a}_۱ + \frac{p}{n}\hat{a}_۱^۲}{\frac{n+۱}{n}\hat{a}_۲ + \frac{p-n}{n}\hat{a}_۱}, ۱\right\} \qquad (۱۲.۲)$$

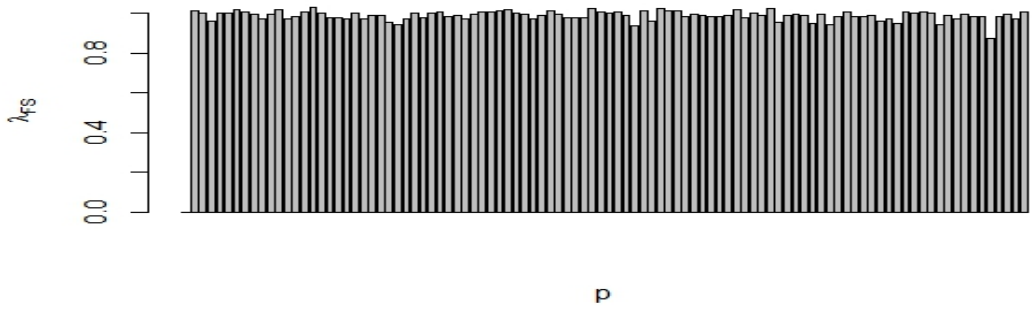
که در آن $a_i = \frac{۱}{p}tr(\Sigma^i)$. در این صورت

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

که برآوردگرهایی سازگار برای $a_۱$ و $a_۲$ هستند. نمودار این پارامتر را بر اساس یک نمونه‌ی شبیه‌سازی شده Nتایی بر حسب p $(p = ۲, \ldots, ۱۰۰)$ رسم نموده‌ایم که شکل ۴.۲ به‌دست آمده است

۵.۲ برآوردگر جدید تحت شرایط بهینه

لدویت و ولف (۲۰۰۴)، برآوردگری مجانبی اما با رویکردی متفاوت را برای کاربردهای متناهی در نظر گرفتند. همانطورکه در قسمت ۳.۲ بیان شد مقدار بهینه برای λ، $\beta^۲/\delta^۲$ می‌باشد. حال



شکل ۴.۲: پارامتر انقباضی FS

در این قسمت به برآورد ضرایب δ^2، β^2، α^2 و μ_λ می‌پردازیم؛ اما ابتدا فرض و لم‌های زیر را مطرح می‌کنیم.

فرض کنید

$$a_i = \frac{1}{p} tr\Sigma^i$$

که میانگین $i-$امین مقادیر ویژه Σ می‌باشد.

لم ۱.۵.۲. *(سان و فیشر، ۲۰۰۹) فرض کنید* $W^T = (w_1, \ldots, w_p)$ *که* w_i*ها دارای توزیع* $N_n(0, I)$ *و* $\nu = w_i^T w_i$ *متغیرهای تصادفی کای‌دو با درجه آزادی* n *باشند، آنگاه داریم*

الف. $E[\nu_{ii}^4] = n(n+2)(n+4)(n+6)$

ب. $E[\nu_{ii}^2] = n(n+2)$

پ. $E[\nu_{ii}\nu_{ij}^2] = n(n+2)$

ت. $E[\nu_{ij}^4] = 3n(n+2)$

ث. $E[\nu_{ij}\nu_{ik}\nu_{jk}] = n(n+2)^2$

ج. $E[\nu_{ij}\nu_{ik}\nu_{jk}] = n$

چ. $E[\nu_{ii}\nu_{ij}\nu_{ik}\nu_{jk}] = n(n+2)$

ح. $E[\nu_{ii}^2] = n(n+2)(n+4)$

خ. $E[\nu_{ii}] = n$

د. $E[\nu_{ii}^2\nu_{ij}^2] = n(n+2)(n+4)$

ذ. $E[\nu_{ij}^2] = n$

ر. $E[\nu_{ij}^2\nu_{jk}^2] = n(n+2)$

ز. $E[\nu_{ij}\nu_{il}\nu_{jk}\nu_{kl}] = n$

لم ۲.۵.۲. *(سان و فیشر، ۲۰۰۹) فرض کنید $X_1, \ldots, X_n$ نمونه‌ای تصادفی از توزیع نرمال p-متغیره با میانگین صفر و ماتریس کوواریانس Σ باشد، در این صورت خواهیم داشت*

$$E[tr(S^2)] = \frac{n+1}{n} tr\Sigma^2 + \frac{1}{n}(tr\Sigma)^2 \qquad (۱۳.۲)$$

که در آن $\Sigma > 0$ و S نیز به‌صورت

$$S = \frac{1}{N}\sum_{i=1}^{N}(X_i - \bar{X})(X_i - \bar{X})^T \qquad (۱۴.۲)$$

تعریف می‌شود.

برهان. ترکیب قطری بردارهای

$$y_k = \sum_{i=1}^{N} \Omega_{ki} X_i$$

را در نظر بگیرید که در آن $y_N = \sqrt{N}\bar{X}$ و y_k $(k = 1, \ldots, n)$ دارای توزیع نرمال $p-$متغیره با میانگین صفر و ماتریس کوواریانس Σ و مستقل از y_N بوده و Ω_{ki} مولفه kiام ماتریس $\Omega = (\Omega_{ik})$ می‌باشد (برای جزئیات بیشتر سان و فیشر ۲۰۰۹ را ببینید).
فرض کنید

$$S = \frac{1}{n}\sum_{i=1}^{n} y_i y_i^T,$$

و همچنین طبق لم ۱۰.۳.۱، $V = nS = YY^T \sim W_p(n, \Sigma)$ که $Y = (y_1, \ldots, y_n)$ و $y_i \sim N_p(0, \Sigma)$ و مستقل می‌باشند. تجزیه طیفی $\Sigma = \Gamma^T \Lambda \Gamma$ را در نظر بگیرید که $\Lambda = diag(\lambda_1, \ldots, \lambda_p)$ مقادیر ویژه Σ و Γ ماتریس بردارهای ویژه است. فرض کنید $U = (u_1, u_2, \ldots, u_n)$، که u_i دارای توزیع $N_p(0, I)$ هستند و می‌توان نوشت $Y = \Sigma^{\frac{1}{2}} U$ که $\Sigma^{\frac{1}{2}}\Sigma^{\frac{1}{2}} = \Sigma$ می‌باشد. تعریف می‌کنیم $W^T = (w_1, \ldots, w_p) = U^T\Gamma^T$ که w_iها دارای توزیع $N_n(0, I)$ هستند. همچنین تعریف می‌کنیم $\nu_{ii} = w_i^T w_i$ که متغیرهای تصادفی کای‌دو [۱۴] با n درجه آزادی می‌باشند. طبق لم ۶.۱ در سریواستاوا [۱۵] (۲۰۰۵) داریم

$$\begin{aligned} trS &= \frac{1}{n} trU^T\Sigma U \qquad (۱۵.۲)\\ &= \frac{1}{n} trU^T\Gamma^T\Lambda\Gamma U \\ &= \frac{1}{n} trW^T\Lambda W \\ &= \frac{1}{n}\sum_{i=1}^{p}\lambda_i w_i^T w_i \end{aligned}$$

[۱۴] Chi-squared
[۱۵] Srivastava

به‌طور مشابه

$$n^۲(trS)^۲ = \sum_{i=۱}^{p} \lambda_i^۲ \nu_{ii}^۲ + ۲ \sum_{i<j}^{p} \lambda_i \lambda_j \nu_{ii} \nu_{jj} \tag{۱۶.۲}$$

و

$$n^۲ trS^۲ = \sum_{i=۱}^{p} \lambda_i^۲ \nu_{ii}^۲ + ۲ \sum_{i<j}^{p} \lambda_i \lambda_j \nu_{ij}^۲ \tag{۱۷.۲}$$

حال داریم

$$trS^۲ = \frac{۱}{n^۲} \left[\sum_{i=۱}^{p} \lambda_i^۲ \nu_{ii}^۲ + ۲ \sum_{i<j}^{p} \lambda_i \lambda_j \nu_{ij}^۲ \right]$$

با گرفتن امید ریاضی از طرفین عبارت بالا و استفاده از لم ۱.۵.۲ داریم

$$\begin{aligned} E[trS^۲] &= \frac{۱}{n^۲} \left[\sum_{i=۱}^{p} \lambda_i^۲ E[\nu_{ii}^۲] + ۲ \sum_{i<j}^{p} \lambda_i \lambda_j E[\nu_{ij}^۲] \right] \\ &= \frac{n(n+۲)}{n^۲} \sum_{i=۱}^{p} \lambda_i^۲ + ۲ \frac{n}{n^۲} \sum_{i<j}^{p} \lambda_i \lambda_j \\ &= \frac{n+۲}{n} tr\Sigma^۲ + \frac{۱}{n} \left((tr\Sigma)^۲ - tr\Sigma^۲ \right) \\ &= \frac{n+۱}{n} tr\Sigma^۲ + \frac{۱}{n} (tr\Sigma)^۲ \end{aligned} \tag{۱۸.۲}$$

□

لم ۳.۵.۲. *(سان و فیشر، ۲۰۰۹) فرض کنید S ماتریس کوواریانس نمونه تعریف شده در (۲.۲) باشد. در برآورد Σ بر اساس یک نمونه تصادفی از توزیع $N_p(\mu, \Sigma)$ داریم*

$$E[tr(\Sigma^T S)] = tr\Sigma^T [E(S)] = tr\Sigma^T \Sigma = tr(\Sigma^۲)$$

حال به محاسبه‌ی ضرایب موردنظر با توجه به نرم فروبنیوس و لم ۲.۵.۲ می‌پردازیم. با توجه به این که

$$\mu_\lambda = \langle \Sigma, I \rangle = \frac{tr(\Sigma)}{p} = a_۱$$

می‌توان نتیجه گرفت

$$\begin{aligned} \beta^۲ = E[||S - \Sigma||^۲] &= E\left[\frac{tr[(S-\Sigma)(S-\Sigma)^T]}{p} \right] \\ &= E\left[\frac{tr(SS^T)}{p} - \frac{۲tr(S\Sigma)^T}{p} + \frac{tr(\Sigma\Sigma^T)}{p} \right] \\ &= E||S||^۲ - ۲E[\langle S, \Sigma \rangle] + E||\Sigma||^۲ \\ &= E\frac{tr(S^۲)}{p} - \frac{۲}{p} E\left[tr(\Sigma^T S) \right] + \frac{۱}{p} tr(\Sigma^۲) \end{aligned}$$

$$
\begin{aligned}
&= \frac{۱}{p}E[tr(S^۲)] - \frac{۲}{p}E\left[tr(\Sigma^T S)\right] + \frac{۱}{p}tr(\Sigma^۲) \\
&= \frac{n+۱}{np}tr(\Sigma^۲) + \frac{۱}{np}(tr\Sigma)^۲ - \frac{۲}{p}tr(\Sigma^۲) + \frac{۱}{p}tr\Sigma^۲ \\
&= \frac{n+۱}{n}a_۲ + \frac{p}{n}a_۱^۲ - ۲a_۲ + a_۲ \\
&= \left(\frac{n+۱}{n} - ۲ + ۱\right)a_۲ + \frac{p}{n}a_۱^۲ \\
&= \frac{۱}{n}a_۲ + \frac{p}{n}a_۱^۲
\end{aligned}
$$

حال به محاسبه $\delta^۲$ می‌پردازیم، برای این منظور داریم

$$
\begin{aligned}
\delta^۲ = E[||S-\mu_\lambda I||^۲] &= E\left[\frac{tr(S-\mu_\lambda I)(S-\mu_\lambda I)^T}{p}\right] \\
&= E\left[\frac{tr[(S-\mu_\lambda I)(S^T-\mu_\lambda^T I^T)]}{p}\right] \\
&= E\left[\frac{tr(SS^T - SI^T\mu_\lambda^T - \mu_\lambda IS^T + \mu_\lambda II^T\mu_\lambda^T)}{p}\right] \\
&= E\left[\frac{tr(SS^T)}{p} - ۲\mu_\lambda E\frac{tr(SI^T)}{p} + \mu_\lambda^۲ E\frac{tr(II^T)}{p}\right] \\
&= E||S||^۲ - ۲\mu_\lambda E\frac{tr(SI^T)}{p} + \mu_\lambda^۲ E||I||^۲ \\
&= \frac{۱}{p}Etr(S^۲) - \frac{۲\mu_\lambda}{p}Etr(I^T S) + \frac{۲}{p}Etr(I^۲) \\
&= \frac{n+۱}{np}tr\Sigma^۲ + \frac{۱}{np}(tr\Sigma)^۲ - \frac{۲\mu}{p}tr\Sigma + \frac{\mu^۲}{p}p \\
&= \frac{n+۱}{n}a_۲ + \frac{p}{n}a_۱^۲ - ۲a_۱a_۱ + a_۱^۲ \\
&= \frac{n+۱}{n}a_۲ + \left(\frac{p}{n} - ۲ + ۱\right)a_۱^۲ \\
&= \frac{n+۱}{n}a_۲ + \left(\frac{p-n}{n}\right)a_۱^۲
\end{aligned}
$$

سریواستاوا (۲۰۰۵) برآوردگرهای $a_۱$ و $a_۲$ را به ترتیب به صورت زیر پیشنهاد کرد

$$\hat{a}_۱ = \frac{trS}{p}$$

و

$$\hat{a}_۲ = \frac{n^۲}{(n-۱)(n+۲)}\frac{۱}{p}\left[trS^۲ - \frac{۱}{n}(trS)^۲\right]$$

او نشان داد که $\hat{a}_۱$ و $\hat{a}_۲$ نااریب هستند و به ترتیب برای $a_۱$ و $a_۲$ سازگار می‌باشند. در خصوص نااریبی با توجه به تعریف برآوردگرها داریم

$$E(\hat{a}_۱) = \frac{۱}{p}E(trS)$$

با توجه به (۱۵.۲) در ادامه خواهیم داشت

$$\begin{aligned} E(\hat{a}_۱) &= \frac{۱}{p}E(trS) \\ &= \frac{۱}{p}\left(\frac{۱}{n}\sum_{i=۱}^{p}\lambda_i w_i^T w_i\right) \\ &= \frac{۱}{p}\frac{۱}{n}\sum_{i=۱}^{p}\lambda_i E(w_i^T w_i) \\ &= \frac{trS}{p} = \hat{a}_۱ \end{aligned}$$

همچنین

$$E(\hat{a}_۲) = \frac{۱}{p}E\left[trS^۲ - \frac{۱}{n}(trS)^۲\right] \qquad (۱۹.۲)$$

$$= \frac{n-۱}{n^۳p}E\left[\sum_{i=۱}^{p}\lambda_i^۲\nu_i^۲\right] + \frac{۲}{n^۲p}E\left[\sum_{i<j}^{p}\lambda_i\lambda_j z_{ij}\right] \qquad (۲۰.۲)$$

$$= \frac{n-۱}{n^۳p}n(n+۲)\sum_{i=۱}^{p}\lambda_i^۲ + \frac{۲}{n^۲p}\sum_{i<j}^{p}\lambda_i\lambda_j E(z_{ij})$$

که

$$\begin{aligned} z_{ij} &= \nu_{ij}^۲ - \frac{۱}{n}\nu_{ii}\nu_{jj} \\ &= (w_i^T w_j)^۲ - \frac{۱}{n}(w_i^T w_i)(w_j^T w_j) \end{aligned}$$

از طرفی $E(z_{ij}) = ۰$ است زیرا با توجه به اینکه $w_i \sim N_n(۰, I_n)$، داریم

$$\begin{aligned} E(z_{ij}) &= E\left[w_i^T w_j w_j^T w_i - \frac{۱}{n}(w_i^T w_i)(w_j^T w_j)\right] \\ &= E(w_i^T w_i) - \frac{۱}{n}\times n \times n \\ &= ۰ \end{aligned}$$

در ادامه رابطه (۱۹.۲) خواهیم داشت

$$\frac{(n-۱)(n+۲)}{n^۲}\left(\frac{tr\Sigma^۲}{p}\right)$$

از این‌رو

$$\hat{a}_۲ = \frac{n^۲}{(n-۱)(n+۲)}\frac{۱}{p}\left[trS^۲ - \frac{۱}{n}(trS)^۲\right]$$

برآوردگری ناریب برای $\frac{tr\Sigma^۲}{p}$ می‌باشد. برای اثبات سازگاری به سریواستاوا (۲۰۰۵) مراجعه شود.

در یک جمع‌بندی، برآوردگرهای μ، $\beta^۲$ و $\delta^۲$ به ترتیب عبارتند از

$$\hat{\mu_\lambda} = \hat{a}_۱$$

$$\hat{\beta}^{۲} = \frac{۱}{n}\hat{a}_{۲} + \frac{p}{n}\hat{a}_{۱}^{۲}$$

و

$$\hat{\delta}^{۲} = \frac{n+۱}{n}\hat{a}_{۲}^{۲} + \frac{p-n}{n}\hat{a}_{۱}^{۲}$$

ضریب انقباضی (λ) هم توسط $\hat{\beta}^{۲}/\hat{\delta}^{۲}$ برآورد می‌شود.

۶.۲ برآوردگر انقباضی، برای ماتریس هدف همانی

در این قسمت به دنبال برآورد ضریب انقباضی، در حالتی که ماتریس هدف در نظر گرفته‌شده قطری بوده و واریانس آن واحد می‌باشد یا به عبارت دیگر ماتریس هدف انتخاب شده به‌صورت ماتریس همانی[۱۶] در نظر گرفته شود $(T = I)$، می‌باشیم. مقدار مناسب برای پارامتر انقباضی را می‌توان با استفاده از روش لدویت و ولف (۲۰۰۴) به‌دست آورد. برای این منظور ابتدا لم زیر و به دنبال آن قضیه‌ای را ارائه می‌دهیم.

لم ۱.۶.۲. *(لدویت و ولف، ۲۰۰۴) تحت مفروضات لم ۲.۵.۲، فرض کنید* $\delta^{۲} = E[||S-I||^{۲}]$، $\alpha^{۲} = E[||\Sigma - I||^{۲}]$ *و* $\beta^{۲} = E[||S - \Sigma||^{۲}]$ *که* S *ماتریس کوواریانس نمونه،* Σ *ماتریس واریانس-کوواریانس و* I *ماتریس همانی است؛ در آن صورت ترکیب زیر را خواهیم داشت*

$$\delta^{۲} = \alpha^{۲} + \beta^{۲}$$

برهان. می‌دانیم $E(S) = \Sigma$ ، بنابراین داریم

$$\begin{aligned}
\delta^{۲} &= E[||S - I||^{۲}] \\
&= E[||S - \Sigma||^{۲}] + E[||\Sigma - I||^{۲}] + ۲E[\langle S - \Sigma, \Sigma - I\rangle] \\
&= E[||S - \Sigma||^{۲}] + ||\Sigma - I||^{۲} + ۲E[\langle S - \Sigma, \Sigma - I\rangle] \\
&= E[||S - \Sigma||^{۲}] + ||\Sigma - I||^{۲} \\
&= \alpha^{۲} + \beta^{۲}
\end{aligned}$$

□

ترکیب ارائه شده در لم ۱.۶.۲ محاسبات را برای به‌دست آوردن مقادیر $\alpha^{۲}$ و $\beta^{۲}$ در به‌دست آوردن پارامتر انقباضی ساده می‌کند.
در این قسمت ما نیز به دنبال پارامتر انقباضی بهینه هستیم. در واقع باید λی را یافت که مخاطره را مینیمم کند. برای این منظور قضیه اساسی زیر را ارائه می‌کنیم.

[۱۶] Identity

قضیه ۲.۶.۲. *(لدویت و ولف، ۲۰۰۴) تحت مفروضات لم ۲.۵.۲ ، فرض کنید* $\Sigma^* = \lambda I + (1-\lambda)S$. *در این صورت پارامتر* λ*ای که مقدار* $E||\Sigma^* - \Sigma||^2$ *را مینیمم می‌کند عبارتست از* $\lambda = \frac{\beta^2}{\delta^2}$ *و داریم*

$$\Sigma^* = \frac{\beta^2}{\delta^2} I + \frac{\alpha^2}{\delta^2} S \qquad (۲۱.۲)$$

همچنین مقدار می‌نیمم برابر است با $\frac{\alpha^2\beta^2}{\delta^2}$.

برهان. با توجه به این‌که $E(S) = \Sigma$ و در نظر گرفتن لم ۱.۶.۲ می‌توان نوشت

$$\begin{aligned} E||\Sigma^* - \Sigma||^2 &= E[||\lambda I + (1-\lambda)S - \lambda\Sigma - (1-\lambda)\Sigma||^2] \\ &= E[||\lambda(\Sigma - I) + (1-\lambda)(S-\Sigma)||^2] \\ &= \lambda^2 E[||\Sigma - I||^2] + (1-\lambda)^2 E[||S-\Sigma||^2] \\ &+ \lambda(1-\lambda)E[\langle \Sigma - I, S - \Sigma\rangle] \\ &= \lambda^2 ||\Sigma - I||^2 + (1-\lambda)^2 E[||S-\Sigma||^2] \end{aligned}$$

که با مشتق گیری نسبت به λ و مساوی صفر قرار دادن، داریم

$$\frac{\partial E[||\Sigma^* - \Sigma||^2]}{\partial \lambda} = 2\lambda\alpha^2 - 2(1-\lambda)\beta^2 = 0$$

که می‌توان نتیجه گرفت

$$2\lambda\alpha^2 = 2(1-\lambda)\beta^2$$

که در نهایت خواهیم داشت

$$\lambda = \frac{\beta^2}{\alpha^2 + \beta^2} = \frac{\beta^2}{\delta^2} \qquad (۲۲.۲)$$

با توجه به (۲۲.۲)، $1 - \lambda = \alpha^2/\delta^2$ می‌باشد. بنابراین رابطه $E||\Sigma^* - \Sigma||^2$ را می‌توان به‌صورت زیر نوشت.

$$\left(\frac{\beta^2}{\delta^2}\right)\alpha^2 + \left(\frac{\alpha^2}{\delta^2}\right)\beta^2 = \frac{\alpha^2\beta^2}{\delta^2}$$

□

حال با توجه به لم ۲.۵.۲ و تحت فرض نرمال بودن داده‌ها، می‌توان ضرایب α^2، β^2 و δ^2 را به‌صورت زیر محاسبه نمود

$$\begin{aligned} \alpha^2 &= ||\Sigma - I||^2 \\ &= ||\Sigma||^2 - 2\langle \Sigma, I\rangle + ||I||^2 \\ &= \frac{1}{p} tr\Sigma^2 - 2\frac{1}{p} tr\Sigma + 1 \\ &= a_2 - 2a_1 + 1 \end{aligned}$$

همچنین برای محاسبه‌ی $\delta^۲$ داریم

$$\begin{aligned}\delta^۲ &= E[||S-I||^۲] = E\left[\frac{tr[(S-I)(S-I)^T]}{p}\right] \\ &= E\left[\frac{trSS^T}{p} - \frac{trSI^T}{p} - \frac{trSI^T}{p} + \frac{trII^T}{p}\right] \\ &= \frac{۱}{p}E[tr(S^۲)] - ۲E\frac{trSI}{p} + ۱ \\ &= \frac{n+۱}{np}tr\Sigma^۲ + \frac{۱}{n}(tr\Sigma)^۲ - \frac{۲}{p}tr\Sigma + ۱ \\ &= \frac{n+۱}{n}a_۲ + \frac{p}{n}a_۱{}^۲ - ۲a_۱^۲ + ۱\end{aligned}$$

از طرفی داریم

$$\beta^۲ = E[||S-\Sigma||^۲] = \frac{۱}{n}a_۲ + \frac{p}{n}a_۱{}^۲$$

در نتیجه پارامتر انقباضی به‌صورت زیر به‌دست می‌آید

$$\lambda = \frac{\frac{۱}{n}a_۲ + \frac{p}{n}a_۱{}^۲}{\frac{n+۱}{n}a_۲ + \frac{p}{n}a_۱{}^۲ - ۲a_۱^۲ + ۱}$$

$a_۱$ و $a_۲$ را می‌توان با برآوردگرهای سازگار و ناااریب $\hat{a}_۱$ و $\hat{a}_۲$ جایگذاری نمود. در این صورت پارامتر انقباضی به‌صورت

$$\hat{\lambda} = \frac{\hat{\beta}^۲}{\hat{\delta}^۲}$$

ارائه می‌شود که در آن

$$\hat{\beta}^۲ = \frac{۱}{n}\hat{a}_۲ + \frac{p}{n}\hat{a}_۱^۲$$

و

$$\hat{\delta}^۲ = \frac{n+۱}{n}\hat{a}_۲ + \frac{p}{n}\hat{a}_۱^۲ - ۲\hat{a}_۱^۲ + ۱$$

بنابراین می‌توان انتظار داشت که برآوردگر حاصل شده در ابعاد بالا با حجم نمونه‌ی کوچک برآوردگر خوبی باشد.

۷.۲ برآوردگر انقباضی، برای ماتریس هدف غیر همانی

در این بخش به بررسی حالتی برای برآورد پارامتر انقباضی می‌پردازیم که ماتریس هدف به‌صورت $T = D_s$ باشد که در آن $D_s = diag(S)$ ماتریس قطری شامل عناصر روی قطر اصلی ماتریس S می‌باشد. ویژگی این ماتریس نیز معین مثبت بودن آن است. پس در این حالت می‌توان برآوردگر ماتریس کوواریانس نمونه (۴.۲) را به‌صورت زیر نوشت

$$S^* = \lambda D_s + (۱-\lambda)S$$

با استفاده از روشی مشابه روش لدویت و ولف (۲۰۰۴) و محاسبات بخش‌های ۵.۲ و ۶.۲، پارامتر انقباضی را با استفاده از کمینه کردن زیان درجه دو و با در نظر گرفتن نرم فروبنیوس به‌دست می‌آوریم. مشابه بخش ۳.۲ که ماتریس هدف $T = \mu_\lambda I$ بیان گردید، در این قسمت با در نظر گرفتن $T = D_s$ ضرایب مورد نظر را محاسبه می‌نماییم. اما قبل از آن لم و قضیه‌ای در ادامه آورده شده است که برای محاسبه پارامتر انقباضی بهینه با استفاده از اسکالرها باید مورد توجه قرار گیرد.

لم ۱.۷.۲. *(لدویت و ولف، ۲۰۰۴) تحت مفروضات لم ۲.۵.۲، فرض کنید* $\delta_D^2 = E[||S - D_s||^2]$، $\alpha_D^2 = E[||S - \Sigma||^2]$ *و* $\beta_D^2 = E[||\Sigma - D_s||^2]$ *که در آن* $\gamma_D^2 = E[\langle S - \Sigma, \Sigma - D_s \rangle]$ *ماتریس کوواریانس نمونه معرفی‌شده در (۲.۲)،* Σ *ماتریس واریانس-کوواریانس و* D_s *ماتریس هدف (قطری)*[۱۷] *می‌باشد، آنگاه ترکیب زیر را خواهیم داشت*

$$\delta_D^2 = \alpha_D^2 + \beta_D^2 + 2\gamma_D^2 \tag{۲۳.۲}$$

برهان.

$$\begin{aligned} \delta_D^2 &= E[||S - D_s||^2] \\ &= E[||S - \Sigma||^2] + E[||\Sigma - D_s||^2] + 2E[\langle S - \Sigma, \Sigma - D_s \rangle] \\ &= \alpha_D^2 + \beta_D^2 + 2\gamma_D^2 \end{aligned} \tag{۲۴.۲}$$

□

ترکیب ارائه شده (۲۳.۲) به ما امکان این را خواهد داد که به راحتی بتوانیم پارامتر انقباضی بهینه را بیابیم. در این قسمت ما نیز به دنبال پارامتر انقباضی بهینه هستیم. در واقع باید λی را یافت که مخاطره را می‌نیمم کند. برای این منظور قضیه اساسی زیر را ارائه می‌دهیم.

قضیه ۲.۷.۲. *(لدویت و ولف، ۲۰۰۴) تحت مفروضات لم ۱.۷.۲، فرض کنید* $\Sigma^* = \lambda I + (1 - \lambda)S$. *در این صورت پارامتر* λ*ای که مقدار* $E||\Sigma^* - \Sigma||^2$ *را مینیمم می‌کند عبارتست از* $\lambda = \frac{\alpha_D^2 + \gamma_D^2}{\delta_D^2}$ *و داریم*

$$\Sigma^* = \frac{\alpha_D^2 + \gamma_D^2}{\delta_D^2} S + \frac{\beta_D^2 + \gamma_D^2}{\delta_D^2} D_s \tag{۲۵.۲}$$

برهان. با استفاده از لم ۱.۷.۲ می‌توان نوشت

$$\begin{aligned} E[||\Sigma^* - \Sigma||^2] &= E[||\lambda D_s - \lambda\Sigma + (1 - \lambda)S - (1 - \lambda)\Sigma||^2] \\ &= \lambda^2 E[||D_s - \Sigma||^2] + (1 - \lambda)^2 E[||S - \Sigma||^2] \\ &\quad + 2\lambda(1 - \lambda)E[\langle S - \Sigma, D_s - \Sigma \rangle] \\ &= \lambda^2 \beta_D^2 + (1 - \lambda)^2 \alpha_D^2 - 2\lambda(1 - \lambda)\gamma_D^2 \end{aligned}$$

[۱۷] Diagonal

که با مشتق گیری نسبت به λ و مساوی صفر قرار دادن، داریم

$$\frac{\partial E[||\Sigma^* - \Sigma||^۲]}{\partial \lambda} = ۲\lambda\alpha_D^۲ - ۲(۱-\lambda)\beta_D^۲ - ۲\lambda\gamma_D^۲ + ۴\lambda\gamma_D^۲ = ۰$$

که می‌توان نتیجه گرفت

$$\lambda(\alpha_D^۲ + \beta_D^۲ + ۲\gamma_D^۲) = \alpha_D^۲ + \gamma_D^۲$$

که در نهایت خواهیم داشت

$$\lambda = \frac{\alpha_D^۲ + \gamma_D^۲}{\delta_D^۲} \tag{۲۶.۲}$$

با توجه به (۲۶.۲)، $۱ - \lambda = \frac{\beta_D^۲ + \gamma_D^۲}{\delta_D^۲}$. با محاسبه‌ی مشتق دوم می‌توان نشان داد که پارامتر انقباضی مینیمم کننده‌ی عبارت $E||\Sigma^* - \Sigma||^۲$ است یعنی

$$\frac{\partial^۲ E[||\Sigma^* - \Sigma||^۲]}{\partial \lambda} = ۲\beta_D^۲ + ۲\alpha_D^۲ + ۴\gamma_D^۲ > ۰$$

□

حال فرض کنید

$$D_\Sigma = diag\Sigma = \begin{pmatrix} \sigma_{۱۱} & ۰ & \dots & ۰ \\ ۰ & \sigma_{۲۲} & \dots & ۰ \\ \vdots & \vdots & \ddots & \vdots \\ ۰ & ۰ & \dots & \sigma_{pp} \end{pmatrix} \tag{۲۷.۲}$$

و

$$a_i^* = \frac{۱}{p} tr D_\Sigma^i$$

مانند بخش‌های ۵.۲ و ۶.۲ تحت پذیره نرمال بودن داده‌ها، می‌توان $\delta_D^۲$، $\alpha_D^۲$، $\beta_D^۲$ و $\gamma_D^۲$ را یافت. قبل از برآورد ضرایب مورد نظر، لم‌های زیر را ارائه می دهیم.

لم ۳.۷.۲. *(سان و فیشر، ۲۰۰۹) فرض کنید $X_۱, \dots, X_n$ دارای توزیع نرمال p-متغیره باشد. در این‌صورت داریم*

$$E[tr(D_s^۲)] = \frac{n+۲}{n} tr D_\Sigma^۲ \tag{۲۸.۲}$$

که در آن D_s ماتریس قطری که عناصر آن، عناصر قطر اصلی ماتریس واریانس-کوواریانس (۴.۲) و D_Σ ماتریس تعریف‌شده در (۲۷.۲) می‌باشد.

برهان. مشابه تکنیک استفاده‌شده در مود (۱۳۷۷) داریم

$$n\frac{S_{ii}}{\sigma_{ii}} \sim \chi_n^۲$$

می‌توان نتیجه گرفت

$$tr(D_s) = \sum_{i=۱}^{p} S_{ii} = \frac{۱}{n}\sum_{i=۱}^{p} \sigma_{ii}\nu_{ii}$$

که ν_{ii} با توجه به برهان لم ۲.۵.۲ دارای توزیع $\chi^۲$ با n درجه آزادی است. چون D_s ماتریسی قطری است، داریم

$$\begin{aligned} tr(D_s^۲) &= \sum_{i=۱}^{p} S_{ii}^۲ \\ &= \frac{۱}{n^۲} \sum_{i=۱}^{p} \sigma_{ii}^۲ \nu_{ii}^۲ \end{aligned}$$

که با گرفتن امید ریاضی از طرفین تساوی فوق نتیجه می‌شود

$$\begin{aligned} E[tr(D_s^۲)] &= \frac{۱}{n^۲} \sum_{i=۱}^{p} \sigma_{ii}^۲ E(\nu_{ii})^۲ \\ &= \frac{n(n+۲)}{n^۲} \sigma_{ii}^۲ \\ &= \frac{n+۲}{n} tr D_\Sigma^۲ \end{aligned}$$

□

لم ۴.۷.۲. *(سان و فیشر، ۲۰۰۹) تحت فرضیات لم ۳.۷.۲ داریم*

$$E[tr(SD_s)] = E[tr(D_s S)] \tag{۲۹.۲}$$

$$= E[tr(D_s^۲)] \tag{۳۰.۲}$$

برهان. با توجه به این که $tr(D_s S) = tr(SD_s)$، D_s یک ماتریس قطری است که عناصر آن S_{ii} است بنابراین SD_s یا $D_s S$ دارای عناصر روی قطر اصلی $S_{۱۱}^۲, \ldots, S_{pp}^۲$ است و نتیجه موردنظر حاصل می‌شود. □

حال با در نظر گرفتن لم‌های ۲۸.۲، ۲۹.۲ و رابطه (۱۳.۲) داریم

$$\begin{aligned} \delta_D^۲ = E[||S - D_s||^۲] &= E\left[\frac{tr[(S-D_s)(S-D_s)^T]}{p}\right] \\ &= E\left[\frac{tr(SS^T - SD_S^T - SD_S^T + D_s D_S^T)}{p}\right] \\ &= E\left[\frac{trSS^T}{p} - \frac{۲trSD_S^T}{p} + \frac{trD_s D_S^T}{p}\right] \\ &= \frac{۱}{p}E[tr(S^۲)] - \frac{۲}{p}E[tr(SD_s)] + \frac{۱}{p}E(trD_s^۲) \\ &= \frac{n+۱}{np} tr\Sigma^۲ + \frac{۱}{np}(tr\Sigma)^۲ - \frac{۲}{p}E[tr(D_s^۲)] + \frac{۱}{p}E(trD_s^۲) \\ &= \frac{n+۱}{np} tr\Sigma^۲ + \frac{۱}{np}(tr\Sigma)^۲ - \frac{n+۲}{np} trD_\Sigma^۲ \\ &= \frac{n+۱}{n} a_۲ + \frac{p}{n} a_۱ - \frac{n+۲}{n} a_۲^* \end{aligned}$$

همچنین خواهیم داشت

$$\begin{aligned}\gamma_D^۲ &= E\left[\langle \Sigma - D_s, S - \Sigma\rangle\right] \\ &= E\left[\frac{tr(S\Sigma) - tr\Sigma^۲ - tr(SD_s) + tr(\Sigma D_s)}{p}\right] \\ &= \frac{۱}{p}E[tr(S\Sigma)] - \frac{۱}{p}E[tr(\Sigma^۲)] - \frac{۱}{p}E[tr(SD_s)] + \frac{۱}{p}E[tr(\Sigma D_s)] \\ &= \frac{۱}{p}tr\Sigma^۲ - \frac{۱}{p}tr\Sigma^۲ - \frac{n+۲}{np}trD_\Sigma^۲ + \frac{۱}{p}trD_\Sigma^۲ \\ &= -\frac{۲}{np}trD_\Sigma^۲ \\ &= -\frac{۲}{n}a_۲^*\end{aligned}$$

چون $E[tr(S\Sigma)] = tr[E(S)\Sigma] = tr(\Sigma^۲)$ و $E[tr(D_s\Sigma)] = tr(D_\Sigma^۲)$، بنابراین

$$\begin{aligned}\alpha_D^۲ = E[||S-\Sigma||^۲] &= E\left[\frac{tr[(S-\Sigma)(S-\Sigma)^T]}{p}\right] \\ &= E\left[\frac{tr(SS^T) - tr(S\Sigma^T) - tr(\Sigma^T S) + tr(\Sigma\Sigma^T)}{p}\right] \\ &= \frac{Etr(S^۲)}{p} - \frac{۲E(tr(S\Sigma^T))}{p} + \frac{Etr\Sigma^۲}{p} \\ &= \frac{۱}{p}\left(Etr(S^۲) - Etr\Sigma^۲\right) \\ &= \frac{n+۱}{np}tr\Sigma^۲ + \frac{۱}{np}(tr\Sigma)^۲ - \frac{۱}{p}tr\Sigma^۲ \\ &= \frac{n+۱}{n} + \frac{p}{n}a_۱^۲ - a_۲ \\ &= \frac{۱}{n}a_۲ + \frac{p}{n}a_۱^۲\end{aligned}$$

لذا

$$\hat{a}_۱^* = \frac{۱}{p}tr(D_s) \tag{۳۱.۲}$$

و

$$\hat{a}_۲^* = \frac{n}{(n+۲)p}tr(D_s^۲) \tag{۳۲.۲}$$

در ادامه نشان می‌دهیم $\hat{a}_۱^*$ و $\hat{a}_۲^*$ برای $a_۱$ و $a_۲$ ناااریب هستند. برای بررسی سازگاری به سان و فیشر (۲۰۰۹) مراجعه کنید.

نتیجه اول واضح است و نتیجه دوم از با استفاده از لم (۴.۷.۲) حاصل می‌شود. برای محاسبه واریانس $\hat{a}_۲^۲$ ابتدا لم زیر را بیان می‌کنیم.

لم ۵.۷.۲. *(سان و فیشر، ۲۰۰۹) تحت فرضیات لم ۳.۷.۲ تعریف می‌کنیم*

$$\hat{a}_۲^* = \frac{n}{(n+۲)p}tr(D_s^۲)$$

که در آن D_s ماتریس قطری شامل عناصر ماتریس S تعریف‌شده در (۲.۲)، p بعد و n مشاهدات می‌باشند. آن‌گاه $\hat{a}_2^*$ برای $a_2^* = \frac{tr\Sigma^2}{p}$، نااریب است.

برهان. با توجه به لم‌های ۳.۷.۲ و ۴.۷.۲ داریم

$$\begin{aligned} E(\hat{a}_2^*) &= \frac{n}{n+2} E(tr(D_s^2)/p) \\ &= \frac{n}{n+2}\frac{n+2}{n}\frac{1}{p} tr\Sigma^2 \\ &= \frac{tr\Sigma^2}{p} \end{aligned}$$

□

حال به محاسبه‌ی واریانس $\hat{a}_2^*$ می‌پردازیم.

لم ۶.۷.۲. *(سان و فیشر، ۲۰۰۹) واریانس $\hat{a}_2^*$ معرفی شده در رابطه (۳۲.۲) تقریبا برابر است با*

$$var(\hat{a}_2^*) \simeq \frac{8}{np} a_4^*$$

برهان. ابتدا به این نکته توجه شود که برای n بزرگ می‌توان گفت: $\frac{n}{n+2} \simeq 1$؛ لذا با توجه با لم ۱.۵.۲ داریم

$$\begin{aligned} var(\hat{a}_2^*) = \frac{1}{p} var(tr(D_s^2)) &= \frac{1}{n^4 p^2} \sum_{i=1}^{n} \sigma_{ii}^4 var(\nu_{ii}^2) \\ &= \frac{1}{n^4 p^2} \sum_{i=1}^{p} \sigma_{ii}^4 (E[\nu_{ii}^4] - E[\nu_{ii}^2]^2) \\ &= \frac{1}{n^4 p^2} \sum_{i=1}^{p} \sigma_{ii}^4 \left(n(n+2)(n+4)(n+6) - n^2(n+2)^2 \right) \\ &= \frac{n(n+2)}{n^4 p^2} \sum_{i=1}^{p} \sigma_{ii}^4 (n^2 + 10n + 24 - n^2 - 2n) \\ &= \frac{8n(n+2)(n+3)}{n^4 p} a_4^* \simeq \frac{8}{np} a_4^* \end{aligned}$$

□

حال با توجه به مطالبی که بیان شد، می‌توان گفت که پارامتر بهینه به‌صورت زیر برآورد می‌شود

$$\hat{\lambda} = \frac{\hat{\beta}_D^2 + \hat{\gamma}_D^2}{\hat{\delta}_D^2}$$

که در آن

$$\hat{\beta}_D^2 = \frac{1}{n}(\hat{a}_2 + p\hat{a}_1^2)$$

$$\hat{\gamma}_D^{۲} = -\frac{۲}{n}\hat{a}_۲^*$$

و

$$\hat{\delta}_D^{۲} = \frac{n+۱}{n}\hat{a}_۲ + \frac{p}{n}\hat{a}_۱ - \frac{n+۲}{n}\hat{a}_۲^*$$

۸.۲ مقایسه برآوردگرها

در این بخش برآوردگرهای مختلف Σ را که از جانشینی λ های معرفی‌شده به‌دست می‌آید، با کمک شبیه‌سازی با هم مقایسه می‌کنیم. بدین منظور علاوه بر محاسبه‌ی مخاطره، از معیار دیگری به‌نام درصد بهبود نسبی در میانگین خطا $(PRIAL)$ [۱۸] استفاده می‌کنیم که به صورت زیر می‌باشد

$$PRIAL(S^*) = \frac{E[||S-\Sigma||_F^{۲}] - E[||S^*-\Sigma||_F^{۲}]}{E[||S-\Sigma||_F^{۲}]} \times ۱۰۰ \qquad (۳۳.۲)$$

که در آن S و S^* دو برآوردگر Σ هستند. اگر مخاطره‌ی برآوردگر S^* کوچک و $PRIAL$ آن مثبت باشد، آنگاه برآوردگر S^* در مقایسه با برآوردگر S، ماتریس واریانس-کوواریانس Σ را بهتر برآورد می‌کند.

۱.۸.۲ مطالعه شبیه‌سازی

در این بخش کارایی برآوردگرها با یکدیگر مورد مقایسه قرار می‌گیرد. حالتی را در نظر می‌گیریم که در آن ماتریس هدف به‌صورت $T = \mu_\lambda I$ است. بدین منظور $n+۱$ داده از توزیع نرمال $p-$متغیره با بردار میانگین صفر و ماتریس معین مثبت Σ تولید می‌کنیم. ماتریس معین مثبت Σ را به گونه‌ای ایجاد می‌کنیم که مقادیر ویژه‌ی آن از توزیع یکنواخت روی بازه $(۰/۵, ۱۰/۵)$ تولید شده باشند (کد برنامه موردنظر در پیوست آ آورده شده است).
پارامتر انقباضی برآورد شده با استفاده از رابطه‌های (۹.۲)، (۱۰.۲)، (۱۱.۲) و (۱۲.۲) به ترتیب روش‌های لدویت و ولف (۲۰۰۴)، چن و همکاران (۲۰۰۹)، شفر و استریمر (۲۰۰۵) و سان و فیشر (۲۰۱۱) به ازای (الف) $n = ۴۰$ و $p = ۲۰$، (ب) $n = ۳۰$ و $p = ۳۰$ و (ج) $n = ۵$ و $p = ۱۰۰$ در جدول‌های ۱.۲، ۲.۲ و ۳.۲ ارائه شده است.

به تعداد $m = ۱۰۰۰$ بار نمونه‌ی $n+۱$ تایی از توزیع نرمال $p-$متغیره تولید کردیم و هر بار مقادیر λ_{SS}، λ_{RBWL}، λ_{LW} و λ_{FS} را محاسبه کردیم که λ_{FS} مربوط به پارامتر انقباضی سان و فیشر (۲۰۱۱)، λ_{LW} مربوط به پارامتر انقباضی لدویت و ولف (۲۰۰۴)، λ_{RBLW} مربوط به چن و همکاران (۲۰۰۹) و λ_{SS} مربوط به پارامتر شفر و استریمر (۲۰۰۵) می‌باشد. میانگین این برآوردها با پارامتر انقباضی واقعی، $\lambda = \frac{\beta^{۲}}{\delta^{۲}}$، که بر اساس ماتریس Σ قابل محاسبه می‌باشد، مقایسه شده‌اند.
جدول ۳.۲ مثالی از ابعاد بالا را نشان می‌دهد. با توجه به نتایج جداول ۱.۲، ۲.۲ و ۳.۲ مشخص

[۱۸] Percentage relative improvement in average loss

جدول ۱.۲: برآورد λ برای $n = ۴۰$، $p = ۲۰$ و $T = \mu_\lambda I$

λ	خطای استاندارد	میانگین شبیه‌سازی‌شده
بهینهλ	۰	۰/۶۱۷۴
λ_{FS}	۰/۰۶۰۴	۰/۶۱۶۹
λ_{LW}	۰/۰۶۱۶	۰/۵۸۵۹
λ_{RBLW}	۰/۰۵۴۳	۰/۶۰۰۴
λ_{SS}	۰/۰۵۹۵	۰/۶۱۲۹

جدول ۲.۲: برآورد λ برای $n = ۳۰$، $p = ۳۰$ و $T = \mu_\lambda I$

λ	خطای استاندارد	میانگین شبیه‌سازی‌شده
بهینهλ	۰	۰/۸۶۳۵
λ_{FS}	۰/۰۰۰۰۶	۰/۸۶۸۵
λ_{LW}	۰/۰۰۰۰۶	۰/۸۱۲۱
λ_{RBLW}	۰/۰۰۰۰۶	۰/۸۳۹۲
λ_{SS}	۰/۰۰۰۰۶	۰/۸۳۸۹

جدول ۳.۲: برآورد λ برای $n = ۵$، $p = ۱۰۰$و $T = \mu_\lambda I$

λ	خطای استاندارد	میانگین شبیه‌سازی‌شده
بهینهλ	۰	۰/۹۲۳۶
λ_{FS}	۰/۰۲۰۷	۰/۹۲۸۶
λ_{LW}	۰	۱
λ_{RBLW}	۰/۰۱۶۵	۰/۷۴۴۳
λ_{SS}	۰/۱۹۶۶	۰/۹۱۰۱

می‌شود که با افزایش بعد، مقدار عددی پارامتر انقباضی بیشتر می‌شود. هم‌چنین از این جدول‌ها ملاحظه می‌گردد که پارامتر انقباضی سان و فیشر (λ_{FS}) در مقایسه با سایر برآوردگرها بهتر عمل می‌کند. پس از این، پارامترهای λ_{SS}، λ_{RBLW} و λ_{LW} در جایگاه‌های بعدی قرار می‌گیرند. جدول‌های ۴.۲، ۵.۲ و ۶.۲ به ترتیب تحت جدول‌های ۲.۱، ۲.۲ و ۳.۲ ساخته شده‌اند.

جدول ۴.۲: برآوردگر انقباضی برای $p = ۲۰$، $n = ۴۰$ و $T = \mu_\lambda I$

برآوردگر	مخاطره	درصد بهبود نسبی
S	۹/۹۵۰۱	۰
S_{FS}	۳/۴۸۰۹	۶۵/۰۲۲۸
S_{LW}	۳/۵۱۲۸	۶۴/۶۹۶۱
S_{RBLW}	۳/۴۹۳۷	۶۴/۸۸۷۴
S_{SS}	۳/۴۸۲۷	۶۴/۹۹۷۷

جدول ۵.۲: برآوردگر انقباضی برای $p = ۳۰$، $n = ۳۰$ و $T = \mu_\lambda I$

برآوردگر	مخاطره	درصد بهبود نسبی
S	۲۳/۵۶۳۰	۰
S_{FS}	۶/۵۸۷۰	۷۲/۰۴۴۹
S_{LW}	۶/۶۴۵۷	۷۱/۷۹۵۹
S_{RBLW}	۶/۵۸۹۶	۷۲/۰۳۴۱
S_{SS}	۶/۵۹۷۵	۷۲/۰۰۰۴

نتایج جدول ۶.۲ نیز نتایج جدول ۴.۲ و ۵.۲ را تایید می‌کند. با توجه به مقدار درصد بهبود نسبی در میانگین خطا مشخص می‌شود که برآوردگر سان و فیشر به اندازه‌ی تقریبا ۹۳ درصد برآوردگر S را بهبود می‌بخشد. ضمنا از مقایسه‌ی جدول ۴.۲ و ۵.۲ ملاحظه می‌شود هنگامی که بعد p افزایش می‌یابد، مخاطره‌ی برآوردگر S افزایش می‌یابد.

حال به مقایسه برآوردگرها زمانی که ماتریس هدف به‌صورت $T = I$ می‌باشد، می‌پردازیم. طبق

جدول ۶.۲: برآوردگر انقباضی برای $p = ۱۰۰$، $n = ۵$ و $T = \mu_\lambda I$

برآوردگر	مخاطره	درصد بهبود نسبی
S	۹۱/۴۹۱۹	۰
S_{FS}	۶/۵۳۵۶	۹۲/۸۵۶۵
S_{LW}	۷/۰۷۵۹	۹۲/۲۶۶۰
S_{RBLW}	۹/۸۰۶۸	۸۹/۲۸۱۹
S_{SS}	۶/۵۶۰۴	۹۲/۸۲۹۵

قسمت قبل، سه حالت (الف) $n = ۴۰$ و $p = ۲۰$، (ب) $n = ۳۰$ و $p = ۳۰$ و (ج) $n = ۵$ و $p = ۱۰۰$ را برای مقایسه برآوردگرها مورد بررسی قرار می‌دهیم که نتایج در جداول ۷.۲، ۸.۲ و ۹.۲ آورده شده‌اند.

جدول ۷.۲: برآوردگر انقباضی برای $p = ۲۰$، $n = ۴۰$ و $T = I$

برآوردگر	مخاطره	درصد بهبود نسبی
S	۱۳/۵۷۹۷	۰
S_{FS}	۶/۹۲۲۸	۴۹/۰۲۱۱
S_{LW}	۷/۱۳۵۳	۴۷/۴۵۵۹
S_{SS}	۷/۰۴۲۰	۴۸/۱۴۳۳

با توجه به جداول ۷.۲ و ۸.۲ می‌توان نتیجه گرفت که برآوردگر سان و فیشر نسبت به برآوردگرهای دیگر بهتر عمل می‌کند و همچنین با افزایش بعد، مخاطره‌ی برآوردگر تجربی S افزایش می‌یابد. همچنین جدول ۹.۲، نتایج جدول ۷.۲ و ۸.۲ را تایید می‌کند. با توجه به درصد بهبود نسبی، مشخص می‌شود برآورد سان و فیشر به اندازه‌ی ۹۸ درصد برآوردگر S را بهبود می‌بخشد.

جدول ۸.۲: برآوردگر انقباضی برای $p = ۳۰$، $n = ۳۰$ و $T = I$

برآوردگر	مخاطره	درصد بهبود نسبی
S	۲۳/۵۶۳۰	۰
S_{FS}	۸/۰۲۷۳	۶۵/۹۳۲۵
S_{LW}	۸/۵۳۸۵	۶۳/۷۶۲۸
S_{SS}	۸/۲۸۶۳	۶۴/۸۳۳۴

جدول ۹.۲: برآوردگر انقباضی برای $p = ۱۰۰$، $n = ۵$ و $T = I$

برآوردگر	مخاطره	درصد بهبود نسبی
S	۵۲۹/۳۳۱۷	۰
S_{FS}	۹/۲۳۳۵	۹۸/۲۵۵۶
S_{LW}	۷۵/۷۸۴۰	۸۵/۶۸۳۰
S_{SS}	۳۴/۱۱۹۸	۹۳/۵۵۴۱

در نهایت حالتی را بررسی می‌کنیم که ماتریس هدف به‌صورت $T = D_s$، عناصر قطری S باشد. به‌طور مشابه مانند قبل به دنبال پارامتر انقباضی بهینه و به تبعیت از آن برآوردگر انقباضی مناسب هستیم. جداول ۱۰.۲، ۱۱.۲ و ۱۲.۲ به ترتیب مربوط به حالات (الف) $n = ۴۰$ و $p = ۲۰$، (ب) $n = ۳۰$ و $p = ۳۰$ و (ج) $n = ۵$ و $p = ۱۰۰$ می‌باشد. در این قسمت فقط λی مربوط به پارامتر λ_{FS} و λ_{SS} محاسبه شده و مقایسه انجام گرفته است.

در این قسمت نیز نتایجی مشابه مطالعات قبل داریم. با توجه به جدول ۱۰.۲ و ۱۱.۲ می‌توان گفت که با افزایش بعد مقدار عددی پارامتر انقباضی بیشتر می‌شود و پارامتر انقباضی سان و فیشر بهتر عمل می‌کند. جدول ۱۲.۲ هم نتایج جداول ۱۰.۲ و ۱۱.۲ را تایید می‌کند.

جدول ۱۰.۲: برآورد λ برای $n = ۴۰$، $p = ۲۰$ و $T = D_s$

λ	خطای استاندارد	میانگین شبیه‌سازی‌شده
بهینهλ	۰	۰/۶۷۸۰
λ_{FS}	۰/۰۰۰۷	۰/۶۸۶۵
λ_{SS}	۰/۰۰۰۰۶	۰/۶۸۶۷

جدول ۱۱.۲: برآورد λ برای $n = ۳۰$، $p = ۳۰$ و $T = D_s$

λ	خطای استاندارد	میانگین شبیه‌سازی‌شده
بهینهλ	۰	۰/۷۰۴۰
λ_{FS}	۰/۰۰۰۰۵	۰/۷۱۱۷
λ_{SS}	۰/۰۰۰۰۵	۰/۷۰۳۸

جدول ۱۲.۲: برآورد λ برای $n = ۵$، $p = ۱۰۰$ و $T = D_s$

λ	خطای استاندارد	میانگین شبیه‌سازی‌شده
بهینهλ	۰	۰/۹۸۷۸
λ_{FS}	۰/۰۰۰۰۲	۰/۹۸۷۶
λ_{SS}	۰/۰۰۰۰۱	۰/۷۹۴۵

جداول ۱۳.۲، ۱۴.۲ و ۱۵.۲ به ترتیب تحت جدول‌های ۱۰.۲، ۱۱.۲ و ۱۲.۲ حاصل شده‌اند. همانند مطالعات قبلی با توجه به جدول‌های ۱۳.۲ و ۱۴.۲ می‌توان نتیجه گرفت که برآوردگر سان و فیشر نسبت به برآوردگر شفر و استریمر بهتر عمل می‌کند. همچنین جدول ۱۵.۲، نتایج جدول ۱۳.۲ و ۱۴.۲ را تایید کرده و حاکی از این است که برآوردگر سان و فیشر ۹۶ درصد برآوردگر S را بهبود می‌بخشد.

یکی از معیارهایی که برای انتخاب برآوردگر مناسب ما بدان پرداخته‌ایم، هزینه (زمان) می‌باشد.

جدول ۱۳.۲: برآوردگر انقباضی برای $p = ۲۰$ و $n = ۴۰$ و $T = D_s$

برآوردگر	مخاطره	درصد بهبود نسبی
S	۱۳/۵۷۹۷	۰
S_{FS}	۶/۳۳۱۲	۵۳/۳۷۷۲
S_{SS}	۶/۳۰۷۴	۵۳/۵۵۲۶

جدول ۱۴.۲: برآوردگر انقباضی برای $p = ۳۰$ و $n = ۳۰$ و $T = D_s$

برآوردگر	مخاطره	درصد بهبود نسبی
S	۲۳/۵۶۳۰	۰
S_{FS}	۷/۶۵۰۱	۶۷/۵۳۳۴
S_{SS}	۷/۶۳۲۲	۶۷/۶۰۹۳

جدول ۱۵.۲: برآوردگر انقباضی برای $p = ۱۰۰$ و $n = ۵$ و $T = D_s$

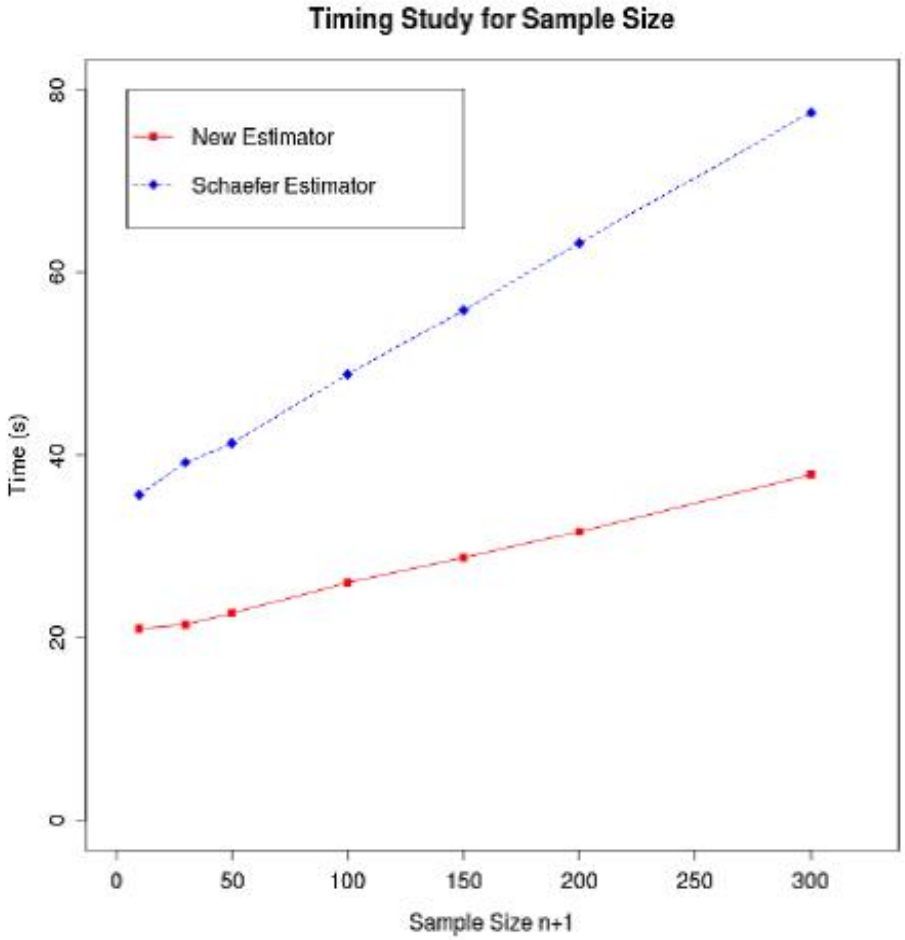
برآوردگر	مخاطره	درصد بهبود نسبی
S	۵۲۹/۳۳۱۷	۰
S_{FS}	۱۹/۰۳۱۹	۹۶/۴۰۴۵
S_{SS}	۳۸/۰۷۸۷	۹۲/۸۰۶۲

در قسمت بعدی به مطالعه هزینه شبیه‌سازی بهکار رفته برای دو برآوردگر انقباضی سان و فیشر (۲۰۰۱) و شفر و استریمر (۲۰۰۵) پرداخته‌ایم.

۹.۲ هزینه شبیه‌سازی

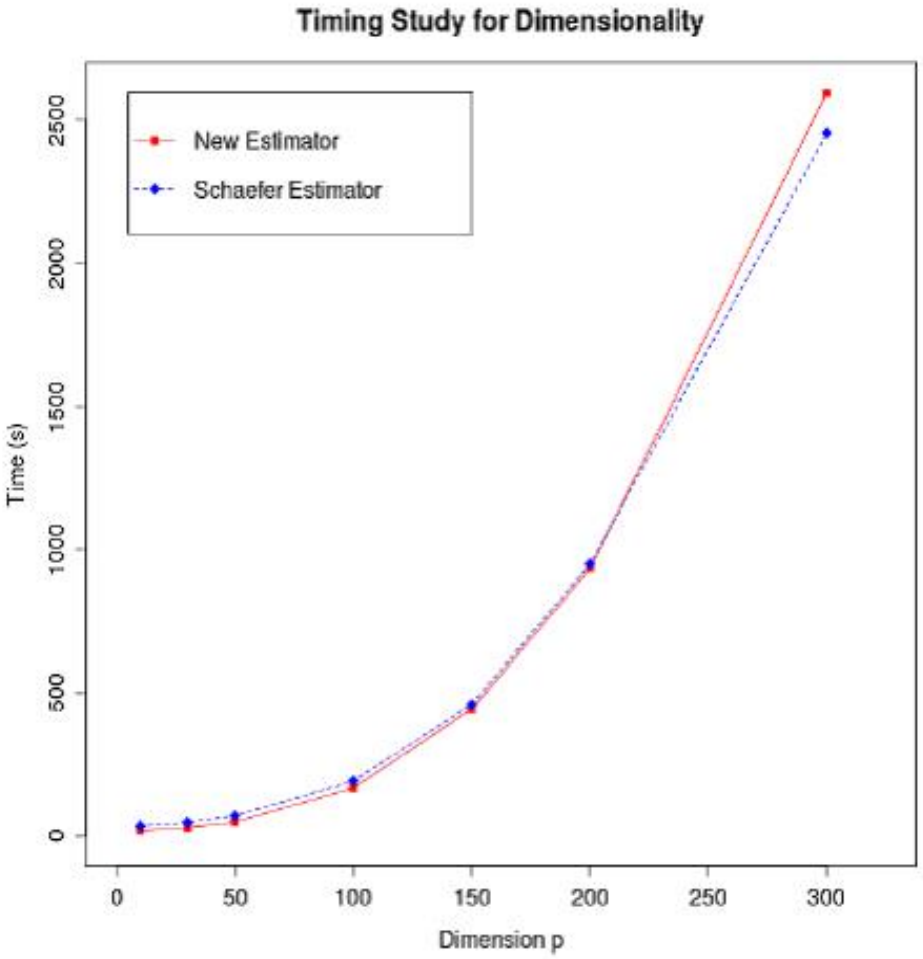
شفر و استریمر (۲۰۰۵) برای برآورد پارامتر انقباضی برای حالتی که ماتریس هدف قطری است، کارآمدی (کارایی) [۱۹] را در نظر گرفتند. در این قسمت قصد داریم نشان دهیم که کارایی برآوردگر سان و فیشر (۲۰۱۱) با برآوردگر شفر و استریمر (۲۰۰۵) تقریبا یکسان است. توجه کنید که هدف این مطالعه تشخیص برتری برآوردگر سان و فیشر (۲۰۱۱) بر برآوردگر شفر و استریمر (۲۰۰۵) نیست ولی با این وجود مقایسه‌ای بین این دو صورت گرفته است.
نمونه‌ای تصادفی با حجم n و بعد p با ماتریس کوواریانس Σ تولید کردیم. این محاسبات را برای $m = ۱۰۰۰۰$ نمونه تکرار کردیم. هر بار نمونه‌ی $N = n + ۱$ از توزیع نرمال با میانگین صفر و ماتریس کوواریانس Σ تولید کردیم. برآوردگر انقباضی سان و فیشر (۲۰۱۱) را به تعداد $m = ۱۰۰۰۰$ بار تولید کردیم؛ که هر بار زمان آن نیز ثبت شد. همین روند را برای برآوردگر شفر و استریمر (۲۰۰۵) نیز داریم.
به‌طور خلاصه سه مطالعه را انجام دادیم: ابتدا بعد را $p = ۲۰$ گرفته و حجم نمونه n را افزایش می‌دهیم. شکل ۵.۲ نتایج را برای حالتی که حجم نمونه افزایش می‌یابد و بعد ثابت است، نشان می‌دهد. با توجه به شکل می‌توان گفت که رفتار هر دو به هم نزدیک است و همچنین زمان یا به عبارتی هزینه شبیه‌سازی برآوردگر شفر و استریمر (۲۰۰۵) با افزایش n، صعود می‌کند. علاوه بر این، با توجه به شکل، برآوردگر شفر و استریمر (۲۰۰۵) در سطح بالاتری نسبت به برآوردگر سان و فیشر (۲۰۱۱) قرار دارد. همچنین در این حالت سرعت محاسبه برآوردگر انقباضی شفر و استریمر (۲۰۰۵) تقریبا دو برابر است. (منظور از $New\ Estimator$ در شکل، برآوردگر سان و فیشر (۲۰۱۱) می‌باشد.)
در مطالعه‌ی بعدی حالتی را بررسی کردیم که بعد در حال افزایش است و $n = ۳۰$ ثابت است. این مطالعه حالتی از بعد بالا می‌باشد. در شکل ۶.۲ نتایج را می‌بینیم. زمان محاسبه هر دو برآوردگر به‌صورت منحنی درجه‌ی دو می‌باشد و خیلی به یکدیگر شباهت دارند.

[۱۹] Efficiency



شکل ۵.۲: زمان مطالعه برای $p = ۲۰$

در نهایت حالتی را بررسی کردیم که هم حجم نمونه و هم بعد، افزایش می‌یابند. فرض کنید $n = p$ باشد. در شکل ۷.۲ دیده می‌شود که زمان محاسبه هر دو به‌صورت منحنی درجه دو می‌باشد. با این تفاوت که نمودار مربوط به برآوردگر سان و فیشر (۲۰۱۱) در سطح پایین‌تری از نمودار مربوط به برآوردگر شفر و استریمر (۲۰۰۵) قرار گرفته است و اندکی کارآمدتر از آن می‌باشد. روی هم رفته با توجه به مطالعات انجام شده می‌توان نتیجه گرفت که برآوردگر سان و فیشر (۲۰۱۱) با برآوردگر شفر و استریمر (۲۰۰۵) از نظر کارآمدی، تفاوت چندانی با هم ندارند.

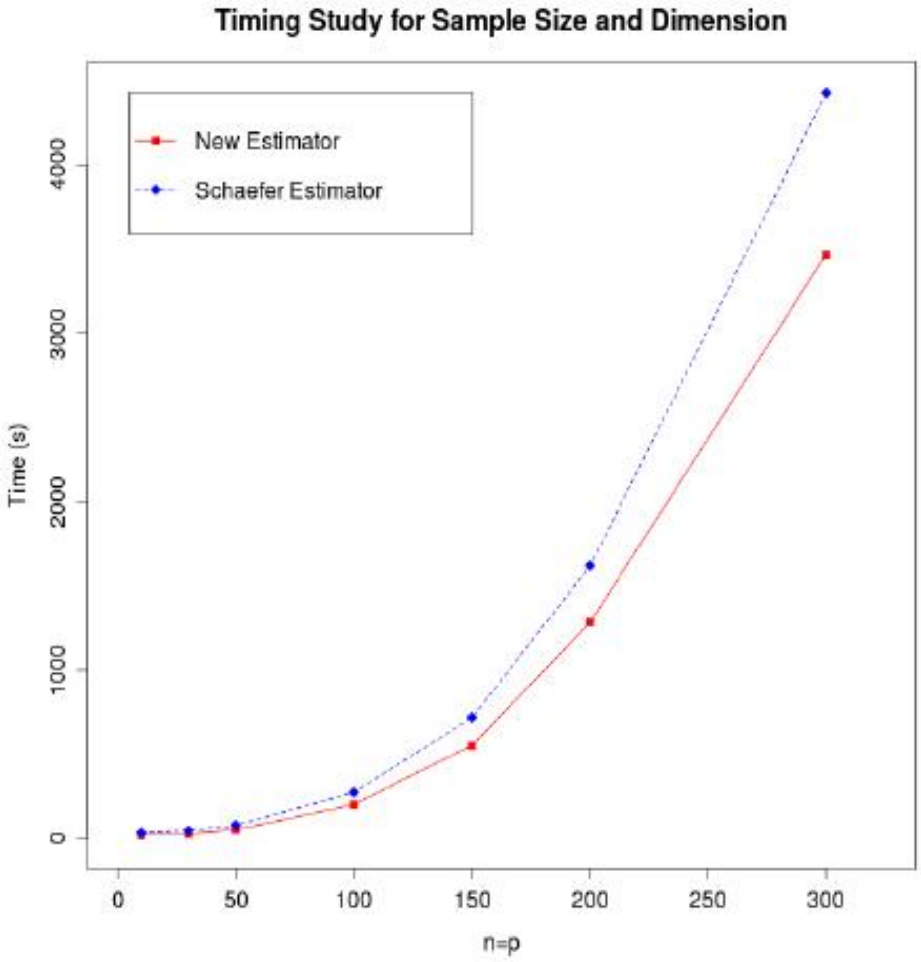


شکل ۶.۲: زمان مطالعه برای $n = ۳۰$

۱۰.۲ مثال واقعی

در این قسمت برآوردگرهای معرفی‌شده در این فصل را برای داده‌های واقعی بکار می‌بریم. داده‌های مورد بررسی برگرفته از موسسه‌ی میکروبیولوژی دانشگاه علوم کشاورزی وین است که مربوط به واکنش فشار $Microorganism\ Escherichia\ Coli$ (معمولا به عنوان $E.Coli$ شناخته می‌شود) می‌باشد. این داده‌ها همه‌ی ۴۲۸۹ ژن کدگذاری‌شده پروتئین را در ۸، ۱۵، ۲۲، ۴۵، ۶۸، ۹۰، ۱۵۰ و ۱۸۰ دقیقه بعد از القا پروتئین نوترکیب بررسی می‌کند. در مقایسه با نمونه‌های مختلف قبل از القا ۱۰۲ ژن توسط اسچمیدت-هک [۲۰] و همکاران (۲۰۰۴) مشخص شده‌اند. این منبع به‌طور متفاوت یک یا نمونه‌های بیشتری را بعد از القا توضیح می‌دهد. توجه داریم که ۱۰۲ متغیر با تنها ۸ مشاهده داریم. که این نمونه‌ای از یک بعد بالاست. در شبیه‌سازی که انجام شد، برآوردگر سان و فیشر (۲۰۱۱) نسبت به برآوردگرهای دیگر برتری داشت. حال در این قسمت برآوردگرهای انقباضی را با استفاده از سه ماتریس هدف معرفی‌شده در این فصل یعنی

[۲۰] Schmidt-Heck



شکل ۷.۲: زمان مطالعه برای $n = p$ یکسان

$T = \mu_\lambda I$، $T = I$ و $T = D_s$ (μ_λ میانگین مقادیر ویژه ماتریس مورد نظر) محاسبه نموده‌ایم. در جدول ۱۶.۲ برای حالتی که $T = \mu_\lambda I$ است، مقادیر عددی برآوردگر λ_{FS}، برآوردگر λ_{LW}، برآوردگر λ_{RBLW}، برآوردگر λ_{SS} و عدد شرطی [۲۱] (اعداد خارج از پرانتز) آن‌ها را آورده‌ایم. برای حالتی که $T = I$ است برآوردگرهای λ_{FS}، λ_{LW} و λ_{SS} و برای حالتی که $T = D_s$ می‌باشد تنها برآوردگر λ_{FS} و λ_{SS} را محاسبه نموده‌ایم. با توجه به این‌که ماتریس کوواریانس در دسترس نیست، تنها راه مقایسه برآوردگرها عدد شرطی برآوردگر است. با توجه به جدول ۱۶.۲ می‌توان نتیجه گرفت که برای هر ماتریس هدف، برآوردگر انقباضی سان و فیشر (۲۰۱۱) بهتر می‌کند.

[۲۱] Condition number

جدول ۱۶.۲: عدد شرطی برآوردگرها و مقدار عددی پارامتر انقباضی برای داده‌های $E.Coli$ برای $p = ۱۰۲$ و $n = ۸$

ماتریس هدف	$T = \mu_\lambda I$	$T = I$	$T = D_s$
FS	۱۵۶/۷۳(۰/۳۳)	۱۵۵/۹۵(۰/۳۳)	۴۶۸/۳۷(۰/۳۳)
LW	۳۸۴/۸۹(۰/۱۷)	۳۸۲/۹۷(۰/۱۷)	NA
$RBLW$	۲۱۲/۲۳ (۰/۲۷)	NA	NA
SS	۲۸۸/۷۹ (۰/۲۱)	۲۸۷/۳۵(۰/۲۱)	۷۱۵/۲۵(۰/۲۴)

فصل ۳

برآوردگر انقباضی ماتریس کوواریانس در توزیع t چندمتغیره

۱.۳ مقدمه

توزیع t چندمتغیره، اولین بار توسط کرنیش [۱] در سال ۱۹۵۴ مورد بررسی قرار گرفته است. محققین زیادی در نظریه توزیع‌های آماری خصوصیات متعددی از توزیع t را مورد مطالعه قرار دادند.

فرضیه‌ها و نظریه‌های زیادی در زمینه‌های اقتصادی، برای توضیح رفتارهای عوامل اقتصادی و روابط بین متغیرها بیان می‌شوند که ممکن است این گزاره‌ها به صورت نظری جذاب و از نظر منطقی صحیح باشند، ولی نمی‌توانند به صورت کاربردی ارائه شوند مگر اینکه از طرف داده‌های واقعی حمایت شوند. معمولا از مدل‌های رگرسیونی برای تحلیل روابط بین متغیرهای اقتصادی و بررسی مسائل اقتصادی استفاده‌های زیادی می‌شود. در این موارد این فرض پایه‌ای که توزیع خطاها t است، بسیار اهمیت دارد. امروزه در مطالعه رفتارها و مدل‌های پیچیده اقتصادی به دلیل وجود نقاط پرت [۲] و فرین [۳]، استفاده از توزیع نرمال به عنوان مدل خطای تصادفی باعث کاهش استواری و ناتوانی مدل در توجیه و تایید مسائل مطرح‌شده می‌شود. در این شرایط از توزیع t که احتمال در دنباله‌های آن بیشتر از توزیع نرمال است، استفاده می‌شود. برای آگاهی بیشتر در این زمینه می‌توان به تیکو[۴] و همکاران (۱۹۹۲) و برش [۵] و همکاران (۱۹۹۷) مراجعه کرد.

آرشی (۱۳۸۷) با در نظر گرفتن توزیع t چندمتغیره به عنوان مولفه خطای مدل رگرسیون چندگانه، برآوردگرهای انقباضی بردار پارامتر مدل را مورد بررسی قرار داد. جعفری (۱۳۸۹) با در نظرگرفتن این فرض که مدل مورد بررسی دارای توزیع t است، نظریه برآوردگرهای نوع استاین میانگین

---

[۱] Cornish
[۲] Outliers
[۳] Extrem values
[۴] Tiku
[۵] Breusch

جامعه را مورد بررسی قرار داد. همچنین می‌توان به مقالات بلتبرگ و ژونیدس [۶] (۱۹۷۴) که از مدل t چندمتغیره برای داده‌های بورس استفاده کردند و نیز زلنر [۷] (۱۹۷۶) که داده‌های بورس را با در نظر گرفتن یک مدل رگرسیون تحت فرض اینکه خطاها دارای مدل t چندمتغیره باشند، تحلیل کرده است، اشاره نمود. آرمانفر (۱۳۸۹) به بررسی خواص توزیع t چندمتغیره و توزیع t با متغیر ماتریسی و همچنین برآوردگرهای درستنمایی ماکزیمم برای این دو توزیع پرداخته است. تورەزاده (۱۳۹۲) به بررسی برآوردگر بیزی تعمیم یافته تحت تابع زیان نرمال منعکس‌شده پرداخته و برآوردگر انقباضی موجک را برای تعمیمی از توزیع نرمال (مدل نرمال آمیخته مقیاسی و t) به‌دست آورده است. همچنین اندرسن و فنگ [۸] (۱۹۸۶) برآورد درستنمایی ماکزیمم این توزیع را ارائه نمودند. سوترادار [۹] (۱۹۸۹) به بیان روابط بین توزیع‌های نرمال چندمتغیره، ویشارت و توزیع t چندمتغیره پرداخته است. گوپتا و ناگار [۱۰] (۲۰۰۰) با انتشار کتابی با عنوان توزیع‌هایی با متغیر ماتریسی به معرفی این توزیع‌ها از جمله توزیع t با متغیر ماتریسی پرداخته‌اند.
در این فصل به دنبال ارائه و بررسی خواص برآوردگر انقباضی ماتریس واریانس-کوواریانس در توزیع t چندمتغیره هستیم. لازم به ذکر است تمامی مطالب این فصل جدید می‌باشد. در ادامه برخی خواص توزیع t چندمتغیره از جمله نمایش تصادفی [۱۱] را ارائه می‌کنیم.

۲.۳ برخی خواص توزیع t چندمتغیره

قضیه ۱.۲.۳. *(حجتی، ۱۳۹۲) فرض کنید بردار تصادفی $\mathbf{z}$ دارای توزیع نرمال $n-$متغیره با میانگین صفر و ماتریس کوواریانس I_n و متغیر تصادفی s دارای توزیع کی‌دو با m درجه آزادی باشد، آن‌گاه*

$$y = \left(\frac{s}{m}\right)^{-\frac{1}{2}} \mathbf{z}$$

دارای توزیع t، $n-$متغیره با میانگین صفر، ماتریس مقیاس I_n و m درجه آزادی است.

برهان. با توجه به استقلال $\mathbf{z}$ و s، توزیع توام عبارت است از

$$\begin{aligned} f(\mathbf{z}, s) &= f(\mathbf{z}) f(s) \\ &= \frac{1}{(2\pi)^{\frac{p}{2}}} \exp\{-\frac{1}{2}\mathbf{z}^T\mathbf{z}\} \frac{1}{2^{\frac{m}{2}}\Gamma(\frac{m}{2})} s^{\frac{m}{2}-1} \exp\{-\frac{s}{2}\} \\ &= \frac{1}{(2\pi)^{\frac{p}{2}} 2^{\frac{m}{2}}\Gamma(\frac{m}{2})} s^{\frac{m}{2}-1} \exp\{-\frac{1}{2}\mathbf{z}^T\mathbf{z}\} \exp\{-\frac{s}{2}\} \end{aligned}$$

[۶] Blattberg and Gonedes
[۷] Zellnar
[۸] Anderson and Fang
[۹] Sutradhar
[۱۰] Gupta and Nagar
[۱۱] Stochastic representation

تغییر متغیر $\mathbf{z}=\left(\frac{s}{m}\right)^{-\frac{۱}{۲}}$ $y=\left(\frac{s}{m}\right)^{-\frac{۱}{۲}}$ را در نظر بگیرید. ژاکوپی تبدیل عبارت است از

$$\mathbf{z}=\left(\frac{s}{m}\right)^{-\frac{۱}{۲}}y \Longrightarrow J(\mathbf{z}\longrightarrow y)=\left(\frac{s}{m}\right)^{\frac{p}{۲}}$$

بنابرین داریم

$$\begin{aligned}f(y,s)&=\frac{۱}{(۲\pi)^{\frac{p}{۲}}۲^{\frac{m}{۲}}\Gamma(\frac{m}{۲})}s^{\frac{m}{۲}-۱}\exp\{-\frac{s}{۲}\}\exp\left\{-\frac{۱}{۲}y^T\left(\frac{s}{m}\right)y\}\left(\frac{s}{m}\right)^{\frac{p}{۲}}\right\}\\&=\frac{۱}{(۲\pi)^{\frac{p}{۲}}۲^{\frac{m}{۲}}\Gamma(\frac{m}{۲})}s^{\frac{m+p}{۲}-۱}m^{-\frac{p}{۲}}\exp\left\{-\frac{۱}{۲}\left[y^T\left(\frac{s}{m}\right)y+s\right]\right\}\\&=\frac{۱}{(۲\pi)^{\frac{p}{۲}}۲^{\frac{m}{۲}}\Gamma(\frac{m}{۲})}s^{\frac{m+p}{۲}-۱}m^{-\frac{p}{۲}}\exp\left\{-\frac{s}{۲}\left[۱+\left(\frac{۱}{m}\right)y^Ty\right]\right\}\end{aligned}$$

با انتگرال‌گیری نسبت به s داریم

$$f(y)=\frac{۱}{(۲\pi)^{\frac{p}{۲}}۲^{\frac{m}{۲}}\Gamma(\frac{m}{۲})}\int_{.}^{\infty}s^{\frac{m+p}{۲}-۱}\exp\left\{-\frac{s}{۲}\left[۱+\left(\frac{۱}{m}\right)y^Ty\right]\right\}ds$$

با تغییر متغیر $t=\frac{s}{۲}\left[۱+\left(\frac{m}{۲}\right)y^Ty\right]$ می‌توان نتیجه گرفت که

$$s=\frac{۲t}{\left[۱+\left(\frac{۱}{m}\right)y^Ty\right]}$$

و

$$ds=\frac{۲dt}{\left[۱+\left(\frac{۱}{m}\right)y^Ty\right]}$$

در نتیجه

$$\begin{aligned}f(y)&=\frac{۱}{(۲\pi)^{\frac{p}{۲}}۲^{\frac{m}{۲}}\Gamma(\frac{m}{۲})}m^{-\frac{p}{۲}}\int_{.}^{\infty}\left(\frac{۲t}{\left[۱+\left(\frac{۱}{m}\right)y^Ty\right]}\right)^{\frac{m+p}{۲}-۱}\exp\{-t\}\frac{۲dt}{\left[۱+\left(\frac{۱}{m}\right)y^Ty\right]}\\&=\frac{\Gamma\left(\frac{m+p}{۲}-۱\right)}{(m\pi)^{\frac{p}{۲}}\Gamma(\frac{m}{۲})}\left[۱+\left(\frac{۱}{m}\right)y^Ty\right]^{-\frac{m+p}{۲}}\end{aligned}$$

□

لم ۲.۲.۳. *(چو، ۱۹۷۳) فرض کنید متغیر تصادفی X با n مولفه، به‌صورت $X\sim EC_n(\mu,\Sigma,g)$ توزیع شده است که در آن g تابع مولد چگالی می‌باشد. در این صورت تابع چگالی احتمال X را می‌توان به‌صورت زیر بیان کرد.*

$$p_X(x)=\int_{.}^{\infty}W(t)f(x)dt$$

که در آن $f(.)$ نشان‌دهنده‌ی تابع چگالی احتمال $N_n(\mu,t^{-۱}\Sigma)$ و

$$W(t)=(۲\pi)^{\frac{n}{۲}}|\Sigma|^{\frac{۱}{۲}}t^{-\frac{n}{۲}}L^{-۱}(p_X(s))$$

که $L^{-۱}p_X(s)$ نشان‌دهنده‌ی تبدیل معکوس لاپلاس $p_X(s)$ به ازای $s=\frac{(x-\mu)^T\Sigma^{-۱}(x-\mu)}{۲}$ می‌باشد.

برهان. بدون کاسته شدن از کلیات فرض می‌کنیم $\mu = ۰$ و

$$p_X(x) = g(\frac{۱}{۲}x^T\Sigma^{-۱}x) = g(s), \qquad s = \frac{۱}{۲}x^T\Sigma^{-۱}x \geq ۰$$

از طرفی با توجه به اینکه

$$W(t) = (۲\pi)^{\frac{n}{۲}}|\Sigma|^{\frac{۱}{۲}}t^{-\frac{n}{۲}}L^{-۱}(p_X(s))$$

داریم

$$\begin{aligned} p_X(x) = g(s) &= L[W(t)(۲\pi)^{-\frac{n}{۲}}|\Sigma|^{-\frac{۱}{۲}}t^{\frac{n}{۲}}] \\ &= \int_{۰}^{\infty} W(t)(۲\pi)^{-\frac{n}{۲}}t^{-۱}\Sigma^{-\frac{۱}{۲}}e^{-ts}dt \\ &= \int_{۰}^{\infty} W(t)[(۲\pi)^{-\frac{n}{۲}}t^{-۱}\Sigma^{-\frac{۱}{۲}}e^{-\frac{۱}{۲}x^T(t^{-۱}\Sigma)^{-۱}x}dt \\ &= \int_{۰}^{\infty} W(t)f(x)dt \end{aligned}$$

□

لم ۳.۲.۳. *(چو، ۱۹۷۳) فرض کنید متغیر تصادفی X دارای توزیع t، pمتغیره با میانگین μ، ماتریس کوواریانس Σ و درجه آزادی ν است و $h_X(.)$ تابع چگالی احتمال آن باشد، در این صورت داریم*

$$h_X(x) = \int_{۰}^{\infty} W_.(t)f(x)dt$$

که در آن $f(.)$ نشاندهنده تابع چگالی احتمال یک متغیر تصادفی نرمال $p-$ متغیره با میانگین μ و ماتریس کوواریانس $t^{-۱}\Sigma$ بوده و

$$W_.(t) = \frac{۱}{\Gamma(\frac{\nu}{۲})}\left(\frac{\nu t}{۲}\right)^{\frac{\nu}{۲}}e^{-\frac{\nu t}{۲}}t^{-۱} \qquad (۱.۳)$$

لم ۴.۲.۳. **فرض کنید متغیر تصادفی t دارای تابع چگالی احتمال به‌صورت (۱.۳) باشد آن‌گاه $E(t^{-h})$ برابر است با*

$$E(t^{-h}) = \left(\frac{\nu}{۲}\right)^h\left(\frac{\Gamma(\frac{\nu}{۲}-h)}{\Gamma(\frac{\nu}{۲})}\right) \qquad (۲.۳)$$

برهان.

$$\begin{aligned}
E(t^{-h}) &= \int_{\cdot}^{\infty} t^{-h} W_{\cdot}(t)dt \\
&= \int_{\cdot}^{\infty} t^{-h} \frac{۱}{\Gamma(\frac{\nu}{۲})} \left(\frac{\nu t}{۲}\right)^{\frac{\nu}{۲}} e^{-\frac{\nu t}{۲}} t^{-۱} dt \\
&= \frac{۱}{\Gamma(\frac{\nu}{۲})} \left(\frac{\nu}{۲}\right)^{\frac{\nu}{۲}} \int_{\cdot}^{\infty} t^{\frac{\nu}{۲}-h-۱} e^{-\frac{\nu t}{۲}} dt \\
&= \frac{۱}{\Gamma(\frac{\nu}{۲})} \left(\frac{\nu}{۲}\right)^{\frac{\nu}{۲}} \frac{\Gamma\left(\frac{\nu}{۲}-h\right)}{\left(\frac{\nu}{۲}\right)^{\frac{\nu}{۲}-h}} \int_{\cdot}^{\infty} \frac{\left(\frac{\nu}{۲}\right)^{\frac{\nu}{۲}-h}}{\Gamma\left(\frac{\nu}{۲}-h\right)} t^{\frac{\nu}{۲}-h-۱} e^{-\frac{\nu t}{۲}} dt \\
&= \left(\frac{\nu}{۲}\right)^{h} \left(\frac{\Gamma(\frac{\nu}{۲}-h)}{\Gamma(\frac{\nu}{۲})}\right)
\end{aligned}$$

□

لم ۳.۲.۳، برای توزیع t چندمتغیره می‌باشد که کاربرد زیادی در به‌دست آوردن نتایج این فصل دارد.

پس با توجه به قضیه‌ی ۳.۲.۳ می‌توان توزیع t چندمتغیره را به‌صورت آمیزه‌ای از توزیع نرمال $N_p(\mu, t^{-۱}\Sigma)$ و توزیع گاما $t = \Omega \sim \Gamma(\frac{\nu}{۲}, \frac{۱}{۲})$ نوشت به عبارتی

$$h_X(x) = \int_{\cdot}^{\infty} \frac{|t^{-۱}\Sigma|^{-\frac{۱}{۲}}}{(۲\pi)^{\frac{p}{۲}}} exp\left(-\frac{۱}{۲}(x-\mu)^T(t^{-۱}\Sigma)^{-۱}(x-\mu)\right) W_{\cdot}(t)dt$$

که در آن

$$X|\Omega = t \sim N_p(\mu, t^{-۱}\Sigma), \qquad \Omega \sim \Gamma(\frac{\nu}{۲}, \frac{۱}{۲})$$

با استفاده از لم ۳.۲.۳ گشتاورهای بردار تصادفی $\mathbf{x} \sim t_p(\mu, \Sigma, \nu)$ را می‌توان به صورت زیر محاسبه کرد. اگر $\mathbf{x} \sim t_p(\nu, \Sigma, \mu)$ و $\mathbf{y} \sim N(\mu, t^{-۱}\Sigma)$ باشد، در این صورت داریم

$$\begin{aligned}
E(\mathbf{x}) &= E_{\Omega}\{E(\mathbf{y})|\Omega\} \\
&= E_{\Omega}\{\mu\} \\
&= \int \mu W_{\cdot}(t)dt \\
&= \mu
\end{aligned}$$

زیرا $\int W_{\cdot}(t)dt = ۱$. همچنین

$$E(\mathbf{x}\mathbf{x}^T) = E_{\Omega}[E(\mathbf{y}\mathbf{y}^T)|\Omega]$$

با توجه به اینکه $E(\mathbf{y}\mathbf{y}^T|\Omega = t) = \mu\mu^T + t^{-۱}\Sigma$ داریم

$$\begin{aligned}
E(\mathbf{x}\mathbf{x}^T) &= E_{\Omega}(\mu\mu^T + \Omega^{-۱}\Sigma) \\
&= \mu\mu^T + \left(\int t^{-۱} W_{\cdot}(t)\right)\Sigma dt
\end{aligned}$$

با توجه به (۱.۳)، داریم

$$\mu\mu^T + \left(\int t^{-1} W_\cdot(t) \right) \Sigma dt = \mu\mu^T + \frac{\nu}{\nu - 2}\Sigma$$

همچنین

$$\begin{aligned} cov(\mathbf{x}) &= \frac{\nu}{\nu - 2}\Sigma + \mu\mu^T - \mu\mu^T \\ &= \frac{\nu}{\nu - 2}\Sigma \end{aligned}$$

۳.۳ برآوردگر انقباضی

در این بخش، مشابه فصل قبل برآوردگر انقباضی را برای ماتریس کوواریانس توزیع t چندمتغیره ارائه کرده و خواص مربوط به آن را بررسی می‌کنیم.
ترکیب محدب زیر را در نظر بگیرید

$$S^* = (1 - \lambda)\dot{S} + \lambda T$$

که $\lambda \in (0, 1)$ پارامتر انقباضی نام دارد و $\dot{S}$ به صورت

$$\dot{S} = \sum_{i=1}^{N} (\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})^T \qquad (۳.۳)$$

می‌باشد که $\mathbf{x}_i \sim t_p(\mu, \Sigma, \nu)$ و T ماتریس هدف که باید معین مثبت باشد. این ترکیب را طوری در نظر می‌گیریم که S^* واریانس کمتری نسبت به $\dot{S}$ داشته باشد.
با توجه به اینکه ماتریس هدف T از قبل تعیین شده است، مساله اصلی در ارتباط با برآوردگر انقباضی، تعیین پارامتر λ می‌باشد. پارامتر انقباضی λ طوری انتخاب می‌شود که برآوردگر $\dot{S}$ را بهبود بخشد. همانند آنچه در فصل قبل بیان شد، اگر n در مقایسه با p بزرگ باشد، $\dot{S}$ باید واریانس کمتری داشته باشد و باید برآوردگر خوبی باشد، از این‌رو $\lambda \to 0$. بطور مشابه اگر p بزرگ باشد، $\dot{S}$ باید واریانس بزرگتری داشته باشد و باید وزنی را برای ماتریس هدف فراهم کرد به طوری‌که $\lambda \to 1$. تحت میانگین مربعات خطا و با در نظر گرفتن نرم فروبنیوس، مقدار بهینه برای پارامتر انقباضی را می‌یابیم.
ابتدا حالتی را بررسی می‌کنیم که ماتریس هدف به صورت $T = \mu_\lambda I$ تعریف شده باشد که μ_λ مقادیر ویژه ماتریس Σ و I ماتریس همانی می‌باشد. پارامتر انقباضی را به صورت زیر تعریف می‌کنیم

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

که در آن

$$\beta^2 = E[||\dot{S} - \Sigma||^2] = E_\Omega\{E[||\dot{S} - \Sigma||^2]|\Omega\}$$

و

$$\delta^2 = E[||\dot{S} - \mu_\lambda I||^2] = E_\Omega\{E[||\dot{S} - \mu_\lambda I||^2]|\Omega\}$$

متاسفانه μ_λ، β^2 و δ^2 به اطلاعاتی در مورد ماتریس Σ نیاز دارند که در دسترس نیست و باید برآورد شوند. حال به محاسبه‌ی ضرایب مورد نیاز با استفاده از نرم فروبنیوس و لم ۲.۵.۲ می‌پردازیم.
با توجه به اینکه

$$a_i = \frac{1}{p} tr(\Sigma^i)$$

و

$$\mu_\lambda = \langle \Sigma, I \rangle = \frac{tr(\Sigma)}{p} = a_1$$

داریم

$$\beta^2 = E[||\dot{S} - \Sigma||^2] = E_\Omega\{E[||\dot{S} - \Sigma||^2]|\Omega\}$$

که در آن

$$\begin{aligned}
E_t[||\dot{S} - \Sigma||^2] &= E[||\dot{S} - \Sigma||^2 | \Omega = t] \\
&= E_t\left[\frac{tr(\dot{S} - \Sigma)(\dot{S} - \Sigma)^T}{p}\right] \\
&= E_t\left[\frac{tr\dot{S}\dot{S}^T}{p} - \frac{tr\dot{S}\Sigma}{p} - \frac{tr\Sigma\dot{S}^T}{p} + \frac{tr\Sigma\Sigma^T}{p}\right] \\
&= \frac{1}{p}E_t[tr(\dot{S}^2)] - \frac{2}{p}E_t[tr\dot{S}\Sigma] + \frac{1}{p}tr\Sigma^2 \\
&= \frac{n+1}{np}tr(t^{-1}\Sigma)^2 - \frac{1}{np}(trt^{-1}\Sigma)^2 - \frac{2}{p}tr\Sigma E_t(\dot{S}) + \frac{1}{p}E_t tr\Sigma^2 \\
&= \frac{n+1}{np}trt^{-2}\Sigma^2 - \frac{1}{np}t^{-2}(tr\Sigma)^2 - \frac{2tr\Sigma}{p}(N-1)t^{-1}\Sigma + \frac{1}{p}E_t tr\Sigma^2 \\
&= \frac{n+1}{np}t^{-2}tr\Sigma^2 - \frac{1}{np}t^{-2}(tr\Sigma)^2 - \frac{2}{p}(N-1)t^{-1}\Sigma^2 + \frac{1}{p}tr\Sigma^2
\end{aligned}$$

حال با توجه به لم ۳.۲.۳ خواهیم داشت

$$\begin{aligned}
\beta^2 &= \int_0^\infty E_t[||\dot{S} - \Sigma||^2] W_0(t)dt \\
&= \int_0^\infty \frac{n+1}{np}t^{-2}tr\Sigma^2 W_0(t)dt - \int_0^\infty \frac{1}{np}t^{-2}(tr\Sigma)^2 W_0(t)dt \\
&- \int_0^\infty \frac{2}{p}(N-1)t^{-1}\Sigma^2 W_0(t)dt + \int_0^\infty \frac{1}{p}tr\Sigma^2 W_0(t)dt \\
&= \frac{n+1}{np}tr\Sigma^2 \int_0^\infty t^{-2}W_0(t)dt - \frac{1}{np}(tr\Sigma)^2 \int_0^\infty t^{-2}W_0(t)dt \\
&- \frac{2(N-1)\nu}{p(\nu-2)}tr\Sigma^2 + \frac{1}{p}tr\Sigma^2 \int_0^\infty W_0(t)dt
\end{aligned}$$

با در نظر گرفتن لم ۴.۲.۳ داریم

$$\beta^{۲} = (\frac{\nu}{۲})\left(\frac{\Gamma(\frac{\nu}{۲}-۲)}{\Gamma(\frac{\nu}{۲})}\right)\left[\frac{n+۱}{n}a_{۲}+\frac{p}{n}a_{۱}^{۲}\right]-\frac{۲(N-۱)\nu}{(\nu-۲)}a_{۲}+a_{۲} \qquad (۴.۳)$$
$$= \frac{۱}{(\frac{\nu}{۲}-۱)}\left[\frac{n+۱}{n}a_{۲}+\frac{p}{n}a_{۱}^{۲}\right]-a_{۲}\left(\frac{۲(N-۱)\nu}{(\nu-۲)}-۱\right)$$

بطور مشابه به محاسبه‌ی $\delta^{۲}$ می‌پردازیم، برای این منظور داریم

$$\delta^{۲} = E[||\dot{S}-\mu_{\lambda}I||^{۲}] = \int_{۰}^{\infty} E_t[||\dot{S}-\mu_{\lambda}I||^{۲}]W_{۰}(t)dt$$

که در آن

$$\begin{aligned}
E_t[||\dot{S}-\mu_{\lambda}I||^{۲}] &= E_t\left[\frac{tr(\dot{S}-\mu_{\lambda}I)(\dot{S}-\mu_{\lambda}I)^T}{p}\right]\\
&= E_t\left[\frac{tr\dot{S}\dot{S}^T}{p}-\frac{tr\dot{S}\mu_{\lambda}^T I^T}{p}-\frac{tr\mu_{\lambda}I\dot{S}^T}{p}+\frac{tr\mu_{\lambda}I\mu_{\lambda}^T I^T}{p}\right]\\
&= \frac{۱}{p}E_t(tr(\dot{S}^{۲}))-\frac{۲}{p}E_t[tr\dot{S}\mu_{\lambda}^T I^T]+\frac{۱}{p}E_t[tr\mu_{\lambda}^{۲}I^{۲}]\\
&= \frac{n+۱}{np}tr(t^{-۱}\Sigma)^{۲}+\frac{۱}{np}(trt^{-۱}\Sigma)^{۲}-\frac{۲}{p}\mu_{\lambda}trE_t(\dot{S})+\frac{۱}{p}\mu_{\lambda}^{۲}p\\
&= \frac{n+۱}{np}trt^{-۲}\Sigma^{۲}+\frac{۱}{np}t^{-۲}(tr\Sigma)^{۲}-\frac{۲}{p}\mu_{\lambda}(N-۱)t^{-۱}tr\Sigma+\mu_{\lambda}^{۲}
\end{aligned}$$

حال با توجه به قضیه ۳.۲.۳ خواهیم داشت

$$\begin{aligned}
\delta^{۲} &= \int_{۰}^{\infty} E[||\dot{S}-\mu_{\lambda}I||^{۲}]W_{۰}(t)dt\\
&= \int_{۰}^{\infty}\frac{n+۱}{np}trt^{-۲}\Sigma^{۲}W_{۰}(t)dt+\int_{۰}^{\infty}\frac{۱}{np}t^{-۲}(tr\Sigma)^{۲}W_{۰}(t)dt\\
&- \int_{۰}^{\infty}\frac{۲}{p}\mu_{\lambda}(N-۱)t^{-۱}tr\Sigma W_{۰}(t)dt+\int_{۰}^{\infty}\mu_{\lambda}^{۲}W_{۰}(t)dt\\
&= \frac{n+۱}{np}tr\Sigma^{۲}\int_{۰}^{\infty}t^{-۲}W_{۰}(t)dt+\frac{۱}{np}(tr\Sigma)^{۲}\int_{۰}^{\infty}t^{-۲}W_{۰}(t)dt\\
&- \frac{۲}{p}\mu_{\lambda}(N-۱)tr\Sigma\int_{۰}^{\infty}t^{-۱}W_{۰}(t)dt+\int_{۰}^{\infty}\mu_{\lambda}^{۲}W_{۰}(t)dt
\end{aligned}$$

با در نظر گرفتن لم ۴.۲.۳ داریم

$$\begin{aligned}
\delta^{۲} &= \frac{۱}{(\frac{\nu}{۲}-۱)}\left[\frac{n+۱}{n}a_{۲}+\frac{p}{n}a_{۱}^{۲}\right]-\frac{۲(N-۱)\nu}{(\nu-۲)}a_{۲}+\mu_{\lambda}{}_{۱}^{۲}\\
&= \frac{۱}{(\frac{\nu}{۲}-۱)}\left[\frac{n+۱}{n}a_{۲}+\frac{p}{n}a_{۱}^{۲}\right]-\frac{۲(N-۱)\nu}{(\nu-۲)}a_{۲}+a_{۱}^{۲}
\end{aligned}$$

در نهایت ضریب انقباضی در حالتی که $T=\mu_{\lambda}I$ می‌باشد، به‌صورت زیر است

$$\lambda = \frac{\frac{۱}{(\frac{\nu}{۲}-۱)}\left[\frac{n+۱}{n}a_{۲}+\frac{p}{n}a_{۱}^{۲}\right]-a_{۲}\left(\frac{۲(N-۱)\nu}{(\nu-۲)}-۱\right)}{\frac{۱}{(\frac{\nu}{۲}-۱)}\left[\frac{n+۱}{n}a_{۲}+\frac{p}{n}a_{۱}^{۲}\right]-\frac{۲(N-۱)\nu}{(\nu-۲)}a_{۲}+a_{۱}^{۲}}$$

۴.۳ برآوردگر انقباضی، برای ماتریس هدف همانی

در این قسمت به دنبال برآورد ضریب انقباضی، برای حالتی که ماتریس هدف یک ماتریس همانی $(T = I)$ در نظر گرفته شود، می‌باشیم.

قضیه ۱.۴.۳. *** فرض کنید $X_1, \ldots, X_n$ دارای توزیع t چندمتغیره باشد، همچنین $\beta^2 = E[||\dot{S} - \Sigma||^2]$ و $\delta^2 = E[||\dot{S} - I||^2]$ باشد که $\dot{S}$ به‌صورت (۳.۳)، Σ ماتریس واریانس-کوواریانس نمونه و $T = I$ ماتریس هدف باشد، آن‌گاه پارامتر انقباضی را به‌صورت زیر خواهیم داشت*

$$\lambda = \frac{\frac{1}{(\frac{\nu}{2}-1)}\left[\frac{n+1}{n}a_2 + \frac{p}{n}a_1^2\right] - a_2\left(\frac{2(N-1)\nu}{(\nu-2)} - 1\right)}{\frac{1}{(\frac{\nu}{2}-1)}\left[\frac{n+1}{n}a_2 + \frac{p}{n}a_1^2\right] - \frac{2(N-1)\nu}{(\nu-2)}a_1 + 1} \tag{۵.۳}$$

برهان. ابتدا به محاسبه δ^2 می‌پردازیم، برای این منظور داریم

$$\delta^2 = E[||\dot{S} - I||^2] = \int_0^\infty E_t[||\dot{S} - I||^2]W_0(t)dt$$

که در آن

$$\begin{aligned}
E_t[||\dot{S} - I||^2] &= E_t\left[\frac{tr(\dot{S} - I)(\dot{S} - I)^T}{p}\right] \\
&= E_t\left[\frac{tr\dot{S}\dot{S}^T}{p} - \frac{\dot{S}I^T}{p} - \frac{trI\dot{S}^T}{p} + \frac{trII^T}{p}\right] \\
&= \frac{1}{p}E_t[tr(\dot{S}^2)] - \frac{2}{p}E_t[trI\dot{S}] + \frac{p}{p} \\
&= \frac{n+1}{np}tr(t^{-1}\Sigma)^2 + \frac{1}{np}(trt^{-1}\Sigma)^2 - \frac{2}{p}ItrE_t(\dot{S}) + 1 \\
&= \frac{n+1}{np}t^{-2}tr\Sigma^2 + \frac{1}{np}t^{-2}(tr\Sigma)^2 - \frac{2}{p}(N-1)t^{-1}tr\Sigma + 1
\end{aligned}$$

حال با توجه به لم ۳.۲.۳ خواهیم داشت

$$\begin{aligned}
\delta^2 &= \int_0^\infty \frac{n+1}{np}t^{-2}tr\Sigma^2 W_0(t)dt + \int_0^\infty \frac{1}{np}t^{-2}(tr\Sigma)^2 W_0(t)dt \\
&- \frac{2}{p}(N-1)t^{-1}tr\Sigma W_0(t)dt + \int_0^\infty W_0(t)dt \\
&= \frac{n+1}{np}tr\Sigma^2\int_0^\infty t^{-2}W_0(t)dt + \frac{1}{np}(tr\Sigma)^2\int_0^\infty t^{-2}W_0(t)dt \\
&- \frac{2}{p}(N-1)tr\Sigma\int_0^\infty t^{-1}W_0(t)dt + \int_0^\infty W_0(t)dt
\end{aligned}$$

همچنین با در نظر گرفتن لم ۴.۲.۳ داریم

$$\delta^2 = \frac{1}{(\frac{\nu}{2}-1)}\left[\frac{n+1}{n}a_2 + \frac{p}{n}a_1^2\right] - \frac{2(N-1)\nu}{(\nu-2)}a_1 + 1$$

از طرفی $\beta^{۲}$ به‌صورت (۴.۳) می‌باشد، در نهایت پارامتر انقباضی به‌صورت زیر حاصل می‌شود

$$\lambda = \frac{\frac{۱}{(\frac{\nu}{۲}-۱)}\left[\frac{n+۱}{n}a_{۲} + \frac{p}{n}a_{۱}^{۲}\right] - a_{۲}\left(\frac{۲(N-۱)\nu}{(\nu-۲)} - ۱\right)}{\frac{۱}{(\frac{\nu}{۲}-۱)}\left[\frac{n+۱}{n}a_{۲} + \frac{p}{n}a_{۱}^{۲}\right] - \frac{۲(N-۱)\nu}{(\nu-۲)}a_{۱} + ۱}$$

□

۵.۳ برآوردگر انقباضی، برای ماتریس هدف غیر همانی

در این بخش به بررسی حالتی برای برآورد ضریب انقباضی می‌پردازیم که ماتریس هدف به‌صورت $T = D_{\dot{S}}$ باشد که در آن $D_{\dot{S}} = diag(\dot{S})$ ماتریس قطری شامل عناصر روی قطر اصلی ماتریس $\dot{S}$ می‌باشد. ویژگی این ماتریس نیز معین مثبت بودن آن است. پس در این حالت می‌توان برآوردگر ماتریس واریانس-کوواریانس (۳.۳) را به‌صورت

$$S^{*} = \lambda D_{\dot{S}} + (۱ - \lambda)\dot{S}$$

نوشت. در این قسمت نیز پارامتر انقباضی را با استفاده از کمینه کردن زیان درجه دو و با در نظر گرفتن نرم فروبنیوس به‌دست می‌آوریم. مشابه بخش ۳.۳ که ماتریس هدف $T = \mu_{\lambda}I$ بیان گردید، در این قسمت با در نظر گرفتن $T = D_{\dot{S}}$ ضرایب مورد نظر را محاسبه می‌نماییم. اما قبل از آن لمی در ادامه آورده شده است که برای محاسبه پارامتر انقباضی بهینه با استفاده از اسکالرها، باید مورد توجه قرار گیرد.

لم ۱.۵.۳. *** فرض کنید*

$$\alpha_{D_{\dot{s}}}^{۲} = E[||\Sigma - D_{\dot{S}}||^{۲}] = \int_{.}^{\infty} E_{t}[||\Sigma - D_{\dot{S}}||^{۲}]W_{.}(t)dt \qquad (۶.۳)$$

$$\beta_{D_{\dot{s}}}^{۲} = E[||\dot{S} - \Sigma||^{۲}] = \int_{.}^{\infty} E_{t}[||\dot{S} - \Sigma||^{۲}]W_{.}(t)dt \qquad (۷.۳)$$

و

$$\gamma_{D_{\dot{s}}}^{۲} = E[\langle \Sigma - D_{\dot{S}}, \dot{S} - \Sigma \rangle] = \int_{.}^{\infty} E_{t}[\langle \Sigma - D_{\dot{S}}, \dot{S} - \Sigma \rangle]W_{.}(t)dt \qquad (۸.۳)$$

که با ترکیب آنها خواهیم داشت

$$\begin{aligned}\delta_{D_{\dot{s}}}^{۲} = E[||\dot{S} - D_{\dot{S}}||^{۲}] &= \int_{.}^{\infty} E_{t}[||\dot{S} - D_{\dot{S}}||^{۲}]W_{.}(t)dt \\ &= \alpha_{D_{\dot{s}}}^{۲} + \beta_{D_{\dot{s}}}^{۲} + ۲\gamma_{D_{\dot{s}}}^{۲} \qquad (۹.۳)\end{aligned}$$

که در آن $\dot{S}$ ماتریس کواریانس نمونه معرفی‌شده در (۳.۳) می‌باشد، Σ ماتریس واریانس-کوواریانس است و $D_{\dot{S}}$ ماتریس هدف (قطری) می‌باشد.

برهان.

$$\begin{aligned}\delta^{۲}_{D_{\dot{s}}} &= \int_{.}^{\infty} E_t[||\dot{S} - D_{\dot{S}}||^{۲}]W_{.}(t)dt \\ &= \int_{.}^{\infty} E_t[||\dot{S} - \Sigma||^{۲}]W_{.}(t)dt + \int_{.}^{\infty} E_t[||\Sigma - D_{\dot{S}}||^{۲}]W_{.}(t)dt \\ &+ \int_{.}^{\infty} ۲E_t[\langle \dot{S} - \Sigma, \Sigma - D_{\dot{S}} \rangle]W_{.}(t)dt \\ &= \alpha^{۲}_{D_{\dot{s}}} + \beta^{۲}_{D_{\dot{s}}} + ۲\gamma^{۲}_{D_{\dot{s}}}\end{aligned}$$

ترکیب ارائه شده به ما امکان این را خواهد داد که به راحتی بتوانیم پارامتر انقباضی بهینه را بیابیم. برای برآورد ضرایب ابتدا موارد زیر را تعریف می‌نماییم

$$D_{\Sigma} = diag(\Sigma), \qquad a^{*}_{i} = \frac{۱}{p} tr D^{i}_{\Sigma}$$

$$E[tr(D^{۲}_{\dot{S}})] = \frac{n + ۲}{n} tr D^{۲}_{\Sigma} \tag{۱۰.۳}$$

و

$$E[tr(\dot{S} D_{\dot{S}})] = E[tr(D_{\dot{S}} \dot{S})] = E[tr(D^{۲}_{\dot{S}})] \tag{۱۱.۳}$$

قضیه ۲.۵.۳. *تحت مفروضات لم ۱.۵.۳، فرض کنید $\Sigma^{*} = \lambda D_{\dot{s}} + (۱ - \lambda)\dot{S}$، پارامتر λای که مقدار $E||\Sigma^{*} - \Sigma||^{۲}$ را مینیمم می‌کند عبارتست از $\lambda = \frac{\alpha^{۲}_{D_{\dot{s}}} + \beta^{۲}_{D_{\dot{s}}}}{\delta^{۲}_{D_{\dot{s}}}}$.*

برهان. اثبات مشابه قضیه ۲.۷.۲ است. □

قضیه ۳.۵.۳. *** اگر $\alpha^{۲}_{D_{\dot{s}}}$، $\beta^{۲}_{D_{\dot{s}}}$، $\gamma^{۲}_{D_{\dot{s}}}$ و $\delta^{۲}_{D_{\dot{s}}}$ به ترتیب روابط معرفی‌شده در (۶.۳)، (۷.۳)، (۸.۳)، (۹.۳) و ماتریس هدف به‌صورت $T = D_s$ باشد، در آن صورت پارامتر انقباضی را به‌صورت زیر خواهیم داشت*

$$\lambda = \frac{\alpha^{۲}_{D_{\dot{s}}} + \beta^{۲}_{D_{\dot{s}}}}{\delta^{۲}_{D_{\dot{s}}}}$$

برهان. با در نظر گرفتن روابط (۱۰.۳)، (۱۱.۳)،

$$\begin{aligned}E[tr(\dot{S}\Sigma)] &= tr\Sigma E(\dot{S}) \\ &= tr\Sigma \frac{(N - ۱)\nu}{\nu - ۲}\Sigma \\ &= \frac{(N - ۱)\nu}{\nu - ۲} tr\Sigma^{۲}\end{aligned}$$

و

$$E[tr(D_{\dot{S}}\Sigma)] = tr(D^{۲}_{\Sigma}) \tag{۱۲.۳}$$

ابتدا به محاسبه‌ی ضرایب موردنظر پرداخته و سپس پارامتر انقباضی را محاسبه می‌نماییم. برای این منظور داریم

$$\delta^{۲}_{D_{\dot{s}}} = E[||\dot{S} - D_{\dot{S}}||^{۲}] = \int_{۰}^{\infty} E_t[||\dot{S} - D_{\dot{S}}||^{۲}]W_{۰}(t)dt$$

که در آن

$$\begin{aligned}
E_t[||\dot{S} - D_{\dot{S}}||^{۲}] &= E_t\left[\frac{tr(\dot{S} - D_{\dot{S}})(\dot{S} - D_{\dot{S}})^T}{p}\right] \\
&= E_t\left[\frac{tr\dot{S}\dot{S}^T}{p} - \frac{۲tr\dot{S}D^T_{\dot{S}}}{p} + \frac{trD_{\dot{S}}D^T_{\dot{S}}}{p}\right] \\
&= \frac{۱}{p}E_t[tr(\dot{S}^{۲})] - \frac{۲}{p}E_t[tr(\dot{S}D_{\dot{S}})] + \frac{۱}{p}E_t[trD^{۲}_{\dot{S}}] \\
&= \frac{n+۱}{np}tr(t^{-۱}\Sigma)^{۲} + \frac{۱}{np}(trt^{-۱}\Sigma)^{۲} \\
&- \frac{۲}{p}E_t[tr(D^{۲}_{\dot{S}})] + \frac{۱}{p}E_t[tr(D^{۲}_{\dot{S}})] + \frac{۱}{p}E_t[tr(D^{۲}_{\dot{S}})] \\
&= \frac{n+۱}{np}t^{-۲}tr\Sigma^{۲} + \frac{۱}{np}t^{-۲}(tr\Sigma)^{۲} \\
&- \frac{۲(n+۲)}{np}trD^{۲}_{\Sigma} + \frac{n+۲}{np}trD^{۲}_{\Sigma}
\end{aligned}$$

حال با توجه به لم ۳.۲.۳ خواهیم داشت

$$\begin{aligned}
\delta^{۲}_{D_{\dot{s}}} = E[||\dot{S} - D_{\dot{S}}||^{۲}] &= \int_{۰}^{\infty} E_t[||\dot{S} - D_{\dot{S}}||^{۲}]W_{۰}(t)dt \\
&= \int_{۰}^{\infty} \frac{n+۱}{np}t^{-۲}tr\Sigma^{۲}W_{۰}(t)dt + \int_{۰}^{\infty} \frac{۱}{np}(trt^{-۱}\Sigma)^{۲}W_{۰}(t)dt \\
&- \int_{۰}^{\infty} \frac{۲(n+۲)}{np}trD^{۲}_{\Sigma}W_{۰}(t)dt + \int_{۰}^{\infty} \frac{n+۲}{np}trD^{۲}_{\Sigma}W_{۰}(t)dt \\
&= \frac{n+۱}{np}tr\Sigma^{۲}\int_{۰}^{\infty} t^{-۲}W_{۰}(t)dt + \frac{۱}{np}(tr\Sigma)^{۲}\int_{۰}^{\infty} t^{-۲}W_{۰}(t)dt \\
&- \frac{۲(n+۲)}{np}trD^{۲}_{\Sigma}\int_{۰}^{\infty} W_{۰}(t)dt + \frac{n+۲}{np}trD^{۲}_{\Sigma}\int_{۰}^{\infty} W_{۰}(t)dt
\end{aligned}$$

همچنین با در نظر گرفتن لم ۴.۲.۳ داریم

$$\delta^{۲}_{D_{\dot{s}}} = \frac{۱}{(\frac{\nu}{۲} - ۱)}\left[\frac{n+۱}{n}a_{۲} + \frac{p}{n}a^{۲}_{۱}\right] - \frac{n+۲}{n}a^{*}_{۲}$$

بطور مشابه داریم

$$\gamma^{۲}_{D_{\dot{s}}} = E[\langle\Sigma - D_{\dot{S}}, \dot{S} - \Sigma\rangle] = \int_{۰}^{\infty} E_t[\langle\Sigma - D_{\dot{S}}, \dot{S} - \Sigma\rangle]W_{۰}(t)dt$$

که در آن

$$E_t[\langle \Sigma - D_{\dot{S}}, \dot{S} - \Sigma \rangle] = E_t\left[\frac{tr(\Sigma - D_{\dot{S}})(\dot{S} - \Sigma)^T}{p}\right]$$
$$= E_t\left[\frac{tr\Sigma\dot{S}}{p} - \frac{tr\Sigma\Sigma^T}{p} - \frac{trD_{\dot{S}}\dot{S}^T}{p} + \frac{trD_{\dot{S}}\Sigma^T}{p}\right]$$
$$= \frac{۱}{p}trE_t(\dot{S}) - \frac{۱}{p}tr\Sigma^۲ - \frac{۱}{p}E_t(tr(D_{\dot{S}}^۲)) + \frac{۱}{p}\Sigma D_\Sigma$$
$$= \frac{(N-۱)\nu}{p(\nu-۲)}tr\Sigma - \frac{۱}{p}tr\Sigma^۲ - \frac{n+۲}{n}trD_\Sigma^۲ + \frac{۱}{p}tr\Sigma D_\Sigma$$

حال با توجه به لم ۳.۲.۳ خواهیم داشت

$$\gamma_{D_{\dot{s}}}^۲ = E[\langle \Sigma - D_{\dot{S}}, \dot{S} - \Sigma \rangle] = \int_۰^\infty E_t[\langle \Sigma - D_{\dot{S}}, \dot{S} - \Sigma \rangle]W_۰(t)dt$$
$$= \int_۰^\infty \frac{(N-۱)\nu}{p(\nu-۲)}tr\Sigma W_۰(t)dt - \int_۰^\infty \frac{۱}{p}tr\Sigma^۲ W_۰(t)dt$$
$$- \int_۰^\infty \frac{n+۲}{n}trD_\Sigma^۲ W_۰(t)dt + \int_۰^\infty \frac{۱}{p}tr\Sigma D_\Sigma W_۰(t)dt$$

همچنین با در نظر گرفتن لم ۴.۲.۳ داریم

$$\gamma_{D_{\dot{s}}}^۲ = \frac{(N-۱)\nu}{\nu-۲}a_۱ - a_۲ - \frac{n+۲}{n}trD_\Sigma^۲ + a_۱D_\Sigma$$

بطور مشابه برای $\alpha_{D_{\dot{s}}}^۲$ داریم

$$\alpha_{D_{\dot{s}}}^۲ = E[||\dot{S} - \Sigma||^۲] = \int_۰^\infty E_t[||\dot{S} - \Sigma||^۲]W_۰(t)dt$$

که در آن

$$E_t[||\dot{S} - \Sigma||^۲] = E_t\left[\frac{tr(\dot{S} - \Sigma)(\dot{S} - \Sigma)^T}{p}\right]$$
$$= \frac{E_t tr(\dot{S}^۲)}{p} - \frac{۲E_t(tr\dot{S}\Sigma)}{p} + \frac{E_t(tr\Sigma^۲)}{p}$$
$$= \frac{n+۱}{np}tr(t^{-۱}\Sigma)^۲ + \frac{۱}{np}t^{-۲}(tr\Sigma)^۲ - \frac{۲tr\Sigma(N-۱)}{p}\Sigma + a_۲$$

حال با توجه به لم ۳.۲.۳ خواهیم داشت

$$\begin{aligned}\alpha^{۲}_{D_{\dot{s}}} = E[||\dot{S}-\Sigma||^{۲}] &= \int_{۰}^{\infty} E_t[||\dot{S}-\Sigma||^{۲}]W_{۰}(t)dt \\ &= \int_{۰}^{\infty} \frac{n+۱}{np} tr(t^{-۱}\Sigma)^{۲} W_{۰}(t)dt + \int_{۰}^{\infty} \frac{۱}{np} t^{-۲}(tr\Sigma)^{۲} W_{۰}(t)dt \\ &- \int_{۰}^{\infty} \frac{۲tr\Sigma(N-۱)}{p} t^{-۱}\Sigma W_{۰}(t)dt + \int_{۰}^{\infty} a_{۲} W_{۰}(t)dt \\ &= \frac{n+۱}{np} tr\Sigma^{۲} \int_{۰}^{\infty} t^{-۲} W_{۰}(t)dt + \frac{۱}{np}(tr\Sigma)^{۲} \int_{۰}^{\infty} t^{-۲} W_{۰}(t)dt \\ &- \frac{۲tr\Sigma(N-۱)}{p} t^{-۱}\Sigma \int_{۰}^{\infty} W_{۰}(t)dt + \int_{۰}^{\infty} a_{۲} W_{۰}(t)dt\end{aligned}$$

همچنین با در نظر گرفتن لم ۴.۲.۳ داریم

$$\alpha^{۲}_{D_{\dot{s}}} = \frac{۱}{(\frac{\nu}{۲}-۱)}\left(\frac{n+۱}{n}a_{۲} + \frac{p}{n}a^{۲}_{۱}\right) - a_{۲}\left(\frac{۲(N-۱)\nu}{\nu-۲} - ۱\right)$$

برای محاسبه‌ی $\beta^{۲}_{D}$ با توجه به ترکیب (۹.۳)، خواهیم داشت

$$\begin{aligned}\beta^{۲}_{D_{\dot{s}}} &= \delta^{۲}_{D} - \alpha^{۲}_{D} - ۲\gamma^{۲}_{D} \\ &= \frac{۱}{(\frac{\nu}{۲}-۱)}\left[\frac{n+۱}{n}a_{۲} + \frac{p}{n}a^{۲}_{۱}\right] - \frac{n+۲}{n}a^{*}_{۲} \\ &- \frac{۱}{(\frac{\nu}{۲}-۱)}\left(\frac{n+۱}{n}a_{۲} + \frac{p}{n}a^{۲}_{۱}\right) - a_{۲}\left(\frac{۲(N-۱)\nu}{\nu-۲} - ۱\right) \\ &- ۲\left(\frac{(N-۱)\nu}{\nu-۲}a_{۱} - a_{۲} - \frac{n+۲}{n}trD^{۲}_{\Sigma} + a_{۱}D_{\Sigma}\right) \\ &= -\frac{n+۲}{n}a^{*}_{۲} - \frac{۲(N-۱)\nu}{\nu-۲}a_{۲} + a_{۲} - \frac{۲(N-۱)\nu}{\nu-۲}a_{۱} + ۲a_{۲} \\ &+ \frac{۲(n+۲)}{n}trD^{۲}_{\Sigma} - ۲a_{۱}D_{\Sigma}\end{aligned}$$

در نهایت پارامتر انقباضی به‌صورت زیر حاصل خواهد شد

$$\lambda = \frac{\alpha^{۲}_{D_{\dot{s}}} + \beta^{۲}_{D_{\dot{s}}}}{\delta^{۲}_{D_{\dot{s}}}}$$

□

۶.۳ مطالعه شبیه‌سازی

در این بخش کارایی برآوردگرها با یکدیگر مورد مقایسه قرار می‌گیرد. برای هر یک از سه ماتریس معرفی‌شده در این فصل یعنی $T = \mu I$، $T = I$ و $T = D_s$ مراحل زیر را انجام داده و مقایسه را انجام می‌دهیم.

بدین منظور $n+۱$ داده از توزیع t، $p-$متغیره با بردار میانگین صفر و ماتریس معین مثبت Σ تولید می‌کنیم. ماتریس معین مثبت Σ را به گونه‌ای ایجاد می‌کنیم که مقادیر ویژه‌ی آن از توزیع یکنواخت روی بازه $(۰/۵, ۱۰/۵)$ تولید شده باشند. همچنین مقدار μ برابر است با $\frac{tr(\Sigma)}{p}$ (کد برنامه موردنظر در پیوست آ آورده شده است).
پارامتر انقباضی برآورد شده با استفاده از رابطه‌های (۹.۲)، (۱۰.۲)، (۱۱.۲) و (۱۲.۲) به ازای n، p و درجه آزادی‌های مختلف و به دنبال آن برآوردگرهای انقباضی در جداول زیر ارائه شده است.

جدول ۱.۳: برآورد λ برای $n = ۴۰$، $p = ۲۰$، $T = \mu_\lambda I$ و $df = ۵$

λ	خطای استاندارد	میانگین شبیه‌سازی‌شده
λ بهینه	۰	۰/۸۳۵۷
λ_{FS}	۰/۰۶۷۰	۰/۷۲۰۸
λ_{LW}	۰	۱
λ_{RBLW}	۰/۰۶۸۱	۰/۷۳۳۷
λ_{SS}	۰/۰۶۴۷	۰/۶۹۶۴

جدول ۲.۳: برآورد λ برای $n = ۴۰$، $p = ۲۰$، $T = \mu_\lambda I$ و $df = ۱۰$

λ	خطای استاندارد	میانگین شبیه‌سازی‌شده
λبهینه	۰	۰/۷۸۷۶
λ_{FS}	۰/۰۶۷۰	۰/۷۴۱۴
λ_{LW}	۰	۱
λ_{RBLW}	۰/۰۶۷۸	۰/۷۳۷۳
λ_{SS}	۰/۰۶۹۱	۰/۷۲۶۹

جدول ۳.۳: برآورد λ برای $n = ۴۰$، $p = ۲۰$، $T = \mu_\lambda I$ و $df = ۱۰۰$

λ	خطای استاندارد	میانگین شبیه‌سازی‌شده
λبهینه	۰	۰/۶۹۲۶
λ_{FS}	۰/۰۶۴۷	۰/۶۶۵۲
λ_{LW}	۰	۱
λ_{RBLW}	۰/۰۶۸۷	۰/۷۲۷۳
λ_{SS}	۰/۰۶۲۴	۰/۶۴۷۳

جدول ۴.۳: برآوردگر انقباضی برای $p = ۲۰$، $n = ۴۰$، $T = \mu_\lambda I$ و $df = ۵$

برآوردگر	مخاطره	درصد بهبود نسبی
S	۶۳/۷۸۵۱	۰
S_{FS}	۸/۱۷۴۷	۸۷/۱۸۳۹
S_{LW}	۷/۹۴۵۳	۸۷/۵۴۳۵
S_{RBLW}	۷/۹۰۲۱	۸۷/۶۱۱۲
S_{SS}	۸/۷۶۰۲	۸۶/۲۶۶

جدول ۵.۳: برآوردگر انقباضی برای $p = ۲۰$، $n = ۴۰$، $T = \mu_\lambda I$ و $df = ۱۰$

برآوردگر	مخاطره	درصد بهبود نسبی
S	۴۵/۲۹۸۲	۰
S_{FS}	۶/۱۴۸۰	۸۶/۴۲۷۶
S_{LW}	۷/۳۶۴۲	۸۳/۷۴۲۶
S_{RBLW}	۶/۱۹۰۷	۸۶/۳۳۳۴
S_{SS}	۶/۳۰۷۵	۸۶/۰۷۵۴

جدول ۶.۳: برآوردگر انقباضی برای $p = ۲۰$، $n = ۴۰$، $T = \mu_\lambda I$ و $df = ۱۰۰$

برآوردگر	مخاطره	درصد بهبود نسبی
S	۱۵/۴۲۸۴	۰
S_{FS}	۴/۴۴۱۲	۷۱/۲۱۳۹
S_{LW}	۶/۳۶۹۴	۵۸/۷۱۶۴
S_{RBLW}	۴/۴۲۱۸	۷۱/۳۳۹۸
S_{SS}	۴/۴۷۸۹	۷۰/۹۶۹۶

جدول ۷.۳: برآورد λ برای $n = ۳۰$، $p = ۳۰$، $T = \mu_\lambda I$ و$df = ۵$

λ	خطای استاندارد	میانگین شبیه‌سازی‌شده
بهینهλ	۰	۰/۹۴۶۳
λ_{FS}	۰/۰۶۰۸	۰/۸۹۹۵
λ_{LW}	۰	۱
λ_{RBLW}	۰/۰۴۸۱	۰/۷۴۳۷
λ_{SS}	۰/۰۵۸۱	۰/۸۷۲۸

جدول ۸.۳: برآورد λ برای $n = ۳۰$، $p = ۳۰$، $T = \mu_\lambda I$ و $df = ۱۰$

λ	خطای استاندارد	میانگین شبیه‌سازی‌شده
بهینهλ	۰	۰/۸۳۶۸
λ_{FS}	۰/۰۵۶۷	۰/۷۸۳۲
λ_{LW}	۰	۱
λ_{RBLW}	۰/۰۴۹۸	۰/۷۳۱۵
λ_{SS}	۰/۰۵۳۶	۰/۷۶۵۲

جدول ۹.۳: برآورد λ برای $n = ۳۰$، $p = ۳۰$، $T = \mu_\lambda I$ و $df = ۱۰۰$

λ	خطای استاندارد	میانگین شبیه‌سازی‌شده
بهینهλ	۰	۰/۸۲۱۱
λ_{FS}	۰/۰۵۶۳	۰/۸۱۵۰
λ_{LW}	۰	۱
λ_{RBLW}	۰/۰۵۱۶	۰/۷۳۳۳
λ_{SS}	۰/۰۵۵۶	۰/۷۹۶۱

جدول ۱۰.۳: برآوردگر انقباضی برای $p = ۳۰$، $n = ۳۰$، $T = \mu_\lambda I$ و $df = ۵$

برآوردگر	مخاطره	درصد بهبود نسبی
S	۸۵/۲۴۳۸	۰
S_{FS}	۵/۵۲۶۹	۹۳/۵۱۶۲
S_{LW}	۵/۶۷۳۱	۹۳/۳۴۴۸
S_{RBLW}	۸/۸۹۷۱	۸۹/۵۶۲۶
S_{SS}	۵/۷۹۳۵	۹۳/۲۰۳۵

جدول ۱۱.۳: برآوردگر انقباضی برای $p = ۳۰$، $n = ۳۰$، $T = \mu_\lambda I$ و $df = ۱۰$

برآوردگر	مخاطره	درصد بهبود نسبی
S	۸۰/۴۹۴۲	۰
S_{FS}	۷/۷۴۷۳	۹۰/۳۷۵۲
S_{LW}	۸/۴۶۴۱	۸۹/۹۴۵۵
S_{RBLW}	۸/۹۱۰۸	۸۸/۹۲۹۸
S_{SS}	۸/۰۹۳۲	۸۹/۹۴۵۵

جدول ۱۲.۳: برآوردگر انقباضی برای $p = ۳۰$، $n = ۳۰$، $T = \mu_\lambda I$ و $df = ۱۰۰$

برآوردگر	مخاطره	درصد بهبود نسبی
S	۳۴/۹۶۶۰	۰
S_{FS}	۶/۸۵۸۱	۸۰/۳۸۶۳
S_{LW}	۸/۳۵۵۸	۷۶/۱۰۲۹
S_{RBLW}	۷/۱۲۴۳	۷۶/۶۲۴۹
S_{SS}	۶/۸۶۹۱	۸۰/۳۵۴۸

جدول ۱۳.۳: برآورد λ برای $n = ۵$، $p = ۱۰۰$، $T = \mu_\lambda I$ و $df = ۵$

λ	خطای استاندارد	میانگین شبیه‌سازی‌شده
بهینهλ	۰	۰/۹۸۹۹
λ_{FS}	۰/۰۲۱۲	۰/۹۸۹۱
λ_{LW}	۰/۰۳۹۹	۰/۹۶۷۵
λ_{RBLW}	۰/۰۱۶۴	۰/۷۴۳۷
λ_{SS}	۰/۰۱۳۷	۰/۷۹۵۱

جدول ۱۴.۳: برآورد λ برای $n = ۵$، $p = ۱۰۰$، $T = \mu_\lambda I$ و $df = ۱۰$

λ	خطای استاندارد	میانگین شبیه‌سازی‌شده
بهینهλ	۰	۰/۹۸۶۸
λ_{FS}	۰/۰۲۳۱	۰/۹۸۶۴
λ_{LW}	۰/۰۴۵۰	۰/۹۶۵۱
λ_{RBLW}	۰/۰۱۸۰	۰/۷۴۰۸
λ_{SS}	۰/۰۱۴۶	۰/۷۹۴۲

جدول ۱۵.۳: برآورد λ برای $n = ۵$، $p = ۱۰۰$، $T = \mu_\lambda I$ و $df = ۱۰۰$

λ	خطای استاندارد	میانگین شبیه‌سازی‌شده
بهینهλ	۰	۰/۹۷۵۹
λ_{FS}	۰/۰۲۴۷	۰/۹۸۳۸
λ_{LW}	۰/۰۴۵۹	۰/۹۶۲۷
λ_{RBLW}	۰/۰۱۷۵	۰/۷۵۰۲
λ_{SS}	۰/۰۱۵۰	۰/۷۹۳۰

جدول ۱۶.۳: برآوردگر انقباضی برای $p = ۱۰۰$، $n = ۵$، $T = \mu_\lambda I$ و $df = ۵$

برآوردگر	مخاطره	درصد بهبود نسبی
S	۱۹۵۲/۱۴	۰
S_{FS}	۷/۶۹۴۶	۹۹/۶۰۰۷
S_{LW}	۹/۰۴۳۴	۹۹/۵۳۶۷
S_{RBLW}	۱۳۰/۲۸۲۹	۹۳/۳۲۶۱
S_{SS}	۸۴/۹۶۱۹	۹۵/۶۴۷۷

جدول ۱۷.۳: برآوردگر انقباضی برای $p = ۱۰۰$، $n = ۵$، $T = \mu_\lambda I$ و $df = ۱۰$

برآوردگر	مخاطره	درصد بهبود نسبی
S	۶۴۴/۵۴۵۲	۰
S_{FS}	۹۸/۶۰۴۴	۸/۹۹۴۷
S_{LW}	۹/۲۸۹۹	۹۸/۵۵۸۶
S_{RBLW}	۴۸/۳۹۰۲	۹۲/۴۹۲۳
S_{SS}	۳۳/۱۲۹۲	۹۴/۸۶۰۰

جدول ۱۸.۳: برآوردگر انقباضی برای $p = ۱۰۰$، $n = ۵$، $T = \mu_\lambda I$ و $df = ۱۰۰$

برآوردگر	مخاطره	درصد بهبود نسبی
S	۴۹۶/۳۳۱۳	۰
S_{FS}	۹/۰۴۵۳	۹۸/۱۷۷۵
S_{LW}	۹/۱۶۵۲	۹۸/۱۵۳۴
S_{RBLW}	۳۷/۹۸۸۲	۹۲/۳۴۶۲
S_{SS}	۲۶/۵۸۶۱	۹۴/۶۴۳۴

جدول ۱۹.۳: برآورد λ برای $n = ۴۰$، $p = ۲۰$، $T = I$ و $df = ۱۰$

λ	خطای استاندارد	میانگین شبیه‌سازی‌شده
λ بهینه	۰	۰/۸۳۵۶
λ_{FS}	۰/۰۷۳۷	۰/۷۸۶۴
λ_{LW}	۰	۱
λ_{SS}	۰/۰۷۲۶	۰/۷۷۲۲

جدول ۲۰.۳: برآورد λ برای $n = ۴۰$، $p = ۲۰$، $T = I$ و$df = ۵$

λ	خطای استاندارد	میانگین شبیه‌سازی‌شده
بهینهλ	۰	۰/۸۳۱۶
λ_{FS}	۰/۰۵۷۳	۰/۶۵۳۹
λ_{LW}	۰	۱
λ_{SS}	۰/۰۵۷۵	۰/۶۳۷۹

جدول ۲۱.۳: برآورد λ برای $n = ۴۰$، $p = ۲۰$، $T = I$ و $df = ۱۰۰$

λ	خطای استاندارد	میانگین شبیه‌سازی‌شده
بهینهλ	۰	۰/۶۷۴۲
λ_{FS}	۰/۰۶۴۷	۰/۶۷۳۴
λ_{LW}	۰	۱
λ_{SS}	۰/۰۶۲۷	۰/۶۵۵۴

جدول ۲۲.۳: برآوردگر انقباضی برای $p = ۲۰$، $n = ۴۰$، $T = I$ و $df = ۵$

برآوردگر	مخاطره	درصد بهبود نسبی
S	۱۵۰/۶۱۳۱	۰
S_{FS}	۱۷/۷۲۶۴	۸۸/۲۳۰۴
S_{LW}	۳۳/۸۹۵۰	۷۷/۴۹۵۲
S_{RBLW}	۳۷/۹۸۸۲	۹۲/۳۴۶۲
S_{SS}	۱۸/۴۲۷۶	۸۷/۷۶۴۹

جدول ۲۳.۳: برآوردگر انقباضی برای $p = ۲۰$، $n = ۴۰$، $T = I$ و $df = ۱۰$

برآوردگر	مخاطره	درصد بهبود نسبی
S	۳۳/۵۴۸۴	۰
S_{FS}	۱۹/۸۳۰۰	۴۰/۸۹۱۳
S_{LW}	۳۳/۲۱۰۹	۱/۰۰۵
S_{SS}	۱۹/۱۹۵۲	۴۲/۷۸۳۳

جدول ۲۴.۳: برآوردگر انقباضی برای $p = ۲۰$، $n = ۴۰$، $T = I$ و $df = ۱۰۰$

برآوردگر	مخاطره	درصد بهبود نسبی
S	۱۸/۰۸۹۴	۰
S_{FS}	۱۴/۳۳۷۵	۲۰/۷۴۰۹
S_{LW}	۲۹/۲۷۶۰	۶۱/۸۴۰۰
S_{SS}	۱۳/۸۳۴۱	۲۳/۵۲۳۹

جدول ۲۵.۳: برآورد λ برای $n = ۳۰$، $p = ۳۰$، $T = I$ و $df = ۵$

λ	خطای استاندارد	میانگین شبیه‌سازی‌شده
بهینهλ	۰	۰/۹۳۴۰
λ_{FS}	۰/۰۶۲۸	۰/۸۷۴۸
λ_{LW}	۰	۱
λ_{SS}	۰/۰۶۰۵	۰/۸۵۱۶

جدول ۲۶.۳: برآورد λ برای $n = ۳۰$، $p = ۳۰$، $T = I$ و $df = ۱۰$

λ	خطای استاندارد	میانگین شبیه‌سازی‌شده
بهینهλ	۰	۰/۸۰۳۲
λ_{FS}	۰/۰۵۵۱	۰/۷۷۸۹
λ_{LW}	۰	۱
λ_{SS}	۰/۰۵۲۹	۰/۷۶۳۴

جدول ۲۷.۳: برآورد λ برای $n = ۳۰$، $p = ۳۰$، $T = I$ و $df = ۱۰۰$

λ	خطای استاندارد	میانگین شبیه‌سازی‌شده
بهینهλ	۰	۰/۸۳۶۰
λ_{FS}	۰/۰۵۷۶	۰/۸۰۴۸
λ_{LW}	۰	۱
λ_{SS}	۰/۰۵۵۸	۰/۷۸۸۶

جدول ۲۸.۳: برآوردگر انقباضی برای $p = ۳۰$، $n = ۳۰$، $T = I$ و $df = ۵$

برآوردگر	مخاطره	درصد بهبود نسبی
S	۴۴۴/۱۴۹۸	۰
S_{FS}	۲۲/۴۴۰۹	۹۴/۹۴۷۴
S_{LW}	۳۲/۸۷۳۰	۹۲/۵۹۶۸
S_{SS}	۲۲/۴۵۳۷	۹۴/۹۴۴۵

جدول ۲۹.۳: برآوردگر انقباضی برای $p = ۳۰$، $n = ۳۰$، $T = I$ و $df = ۱۰$

برآوردگر	مخاطره	درصد بهبود نسبی
S	۳۷/۵۶۶۷	۰
S_{FS}	۱۷/۳۱۷۰	۵۳/۹۰۳۳
S_{LW}	۲۵/۷۷۹۹	۳۱/۳۷۵۵
S_{SS}	۱۶/۹۵۹۵	۵۴/۸۵۴۹

جدول ۳۰.۳: برآوردگر انقباضی برای $p = ۳۰$، $n = ۳۰$، $T = I$ و $df = ۱۰۰$

برآوردگر	مخاطره	درصد بهبود نسبی
S	۲۹/۸۳۹۶	۰
S_{FS}	۱۵/۹۸۰۸	۴۶/۴۴۴۱
S_{LW}	۲۳/۴۳۱۷	۲۱/۴۷۴۵
S_{SS}	۱۵/۵۵۰۹	۴۷/۸۸۴۹

جدول ۳۱.۳: برآورد λ برای $n = ۵$، $p = ۱۰۰$، $T = I$ و $df = ۵$

λ	خطای استاندارد	میانگین شبیه‌سازی‌شده
بهینهλ	۰	۰/۹۹۵۸
λ_{FS}	۰/۰۲۱۷	۰/۹۸۹۴
λ_{LW}	۰/۰۴۲۰	۰/۹۶۷۸
λ_{SS}	۰/۰۱۴۴	۰/۶۹۵۷

جدول ۳۲.۳: برآورد λ برای $n = ۵$، $p = ۱۰۰$، $T = I$ و $df = ۱۰$

λ	خطای استاندارد	میانگین شبیه‌سازی‌شده
بهینهλ	۰	۰/۹۹۰۲
λ_{FS}	۰/۰۲۱۴	۰/۹۸۷۷
λ_{LW}	۰/۰۴۳۲	۰/۹۶۶۵
λ_{SS}	۰/۰۱۴۱	۰/۶۹۵۰

جدول ۳۳.۳: برآورد λ برای $n = ۵$، $p = ۱۰۰$، $T = I$ و $df = ۱۰۰$

λ	خطای استاندارد	میانگین شبیه‌سازی‌شده
بهینهλ	۰	۰/۹۸۵۲
λ_{FS}	۰/۰۲۲۲	۰/۹۸۸۸
λ_{LW}	۰/۰۴۳۱	۰/۹۶۶۹
λ_{SS}	۰/۰۱۳۹	۰/۶۹۵۰

جدول ۳۴.۳: برآوردگر انقباضی برای $p = ۱۰۰$، $n = ۵$، $T = I$ و $df = ۵$

برآوردگر	مخاطره	درصد بهبود نسبی
S	۴۸۳۰/۶۶۶	۰
S_{FS}	۲۳/۹۵۰۷	۹۹/۵۰۴۱
S_{LW}	۲۵/۸۶۴۹	۹۹/۴۶۴۵
S_{SS}	۲۰۵/۳۷۸۴	۹۵/۷۴۸۴

جدول ۳۵.۳: برآوردگر انقباضی برای $p = ۱۰۰$، $n = ۵$، $T = I$ و $df = ۱۰$

برآوردگر	مخاطره	درصد بهبود نسبی
S	۱۳۶۳/۷۸۶	۰
S_{FS}	۳۱/۲۲۸۰	۹۷/۶۱۰۲
S_{LW}	۳۰/۷۷۲۲	۹۷/۷۴۳۶
S_{SS}	۷۳/۹۶۵۷	۹۴/۵۷۶۴

جدول ۳۶.۳: برآوردگر انقباضی برای $p = ۱۰۰$، $n = ۵$، $T = I$ و $df = ۱۰۰$

برآوردگر	مخاطره	درصد بهبود نسبی
S	۸۸۳/۹۸۷۳	۰
S_{FS}	۳۳/۴۳۳۵	۹۶/۲۱۷۸
S_{LW}	۳۲/۶۰۷۴	۹۶/۳۱۱۳
S_{SS}	۵۷/۰۳۶۳	۹۳/۵۴۷۸

جدول ۳۷.۳: برآورد λ برای $n = ۴۰$، $p = ۲۰$، $T = D_s$ و$df = ۵$

λ	خطای استاندارد	میانگین شبیه‌سازی‌شده
بهینهλ	۰	۰/۹۵۶۲
λ_{FS}	۰/۰۷۰۰	۰/۷۱۳۹
λ_{SS}	۰/۰۶۸۸	۰/۷۱۳۴

جدول ۳۸.۳: برآورد λ برای $n = ۴۰$، $p = ۲۰$، $T = D_s$ و$df = ۱۰$

λ	خطای استاندارد	میانگین شبیه‌سازی‌شده
بهینهλ	۰	۰/۷۴۰۰
λ_{FS}	۰/۰۵۴۵	۰/۶۰۸۹
λ_{SS}	۰/۰۵۳۴	۰/۵۸۲۰

جدول ۳۹.۳: برآورد λ برای $n = ۴۰$، $p = ۲۰$، $T = D_s$ و$df = ۱۰۰$

λ	خطای استاندارد	میانگین شبیه‌سازی‌شده
بهینهλ	۰	۰/۵۶۳۲
λ_{FS}	۰/۰۵۷۹	۰/۵۸۴۶
λ_{SS}	۰/۰۵۶۴	۰/۵۸۵۶

جدول ۴۰.۳: برآوردگر انقباضی برای $p = ۲۰$، $n = ۴۰$، $T = D_s$ و $df = ۵$

برآوردگر	مخاطره	درصد بهبود نسبی
S	۳۰۲/۷۱۶۳	۰
S_{FS}	۱۲۴۷/۳۳۴	−۳۲۱/۶۲۷۲
S_{SS}	۱۲۷۴/۵۱۲	−۳۲۱/۰۲۵۱

جدول ۴۱.۳: برآوردگر انقباضی برای $p = ۲۰$، $n = ۴۰$، $T = D_s$ و $df = ۱۰$

برآوردگر	مخاطره	درصد بهبود نسبی
S	۳۸/۶۰۵۲	۰
S_{FS}	۳۶۹/۲۵۸۱	−۸۵۴/۰۲۷۱
S_{SS}	۳۳۸/۸۳۸۲	−۷۷۵/۴۳۳۲

جدول ۴۲.۳: برآوردگر انقباضی برای $p = ۲۰$، $n = ۴۰$، $T = D_s$ و $df = ۱۰۰$

برآوردگر	مخاطره	درصد بهبود نسبی
S	۱۰/۹۱۰۰	۰
S_{FS}	۱۵۳/۳۶۱۳	−۱۳۰۵/۶۸۸
S_{SS}	۱۵۳/۸۹۸۷	−۱۳۱۰/۶۱۳

جدول ۴۳.۳: برآورد λ برای $n = ۳۰$، $p = ۳۰$، $T = D_s$ و $df = ۵$

λ	خطای استاندارد	میانگین شبیه‌سازی‌شده
λبهینه	۰	۰/۸۸۶۰
λ_{FS}	۰/۰۵۶۲	۰/۸۰۳۷
λ_{SS}	۰/۰۵۳۸	۰/۷۸۵۵

جدول ۴۴.۳: برآورد λ برای $n = ۳۰$، $p = ۳۰$، $T = D_s$ و$df = ۱۰$

λ	خطای استاندارد	میانگین شبیه‌سازی‌شده
λبهینه	۰	۰/۸۴۷۹
λ_{FS}	۰/۰۵۵۱	۰/۷۸۴۰
λ_{SS}	۰/۰۵۲۹	۰/۷۶۵۳

جدول ۴۵.۳: برآورد λ برای $n = ۳۰$، $p = ۳۰$، $T = D_s$ و$df = ۱۰۰$

λ	خطای استاندارد	میانگین شبیه‌سازی‌شده
λبهینه	۰	۰/۸۶۸۱
λ_{FS}	۰/۰۵۶۵	۰/۸۴۰۸
λ_{SS}	۰/۰۵۵۲	۰/۸۱۸۶

جدول ۴۶.۳: برآوردگر انقباضی برای $p = ۳۰$، $n = ۳۰$، $T = D_s$ و $df = ۵$

برآوردگر	مخاطره	درصد بهبود نسبی
S	۱۱۹/۶۰۷۴	۰
S_{FS}	۱۱۵۲/۴۵۶	−۸۶۸/۵۳۲۸
S_{SS}	۱۱۰۱/۵۰۷	−۸۲۰/۹۳۵۸

جدول ۴۷.۳: برآوردگر انقباضی برای $p = ۳۰$، $n = ۳۰$، $T = D_s$ و درجه آزادی $df = ۱۰$

برآوردگر	مخاطره	درصد بهبود نسبی
S	۳۸/۲۹۱۰	۰
S_{FS}	۷۷۵/۹۳۴۹	−۱۹۲۶/۴۱۳
S_{SS}	۷۴۰/۰۹۲۷	−۱۸۳۲/۸۰۸

جدول ۴۸.۳: برآوردگر انقباضی برای $p = ۳۰$، $n = ۳۰$، $T = D_s$ و $df = ۱۰۰$

برآوردگر	مخاطره	درصد بهبود نسبی
S	۳۴/۸۷۱۴	۰
S_{FS}	۶۸۰/۶۱۲۳	−۱۸۵۱/۷۷۴
S_{SS}	۶۴۵/۰۷۳۳	−۱۷۴۹/۸۶

جدول ۴۹.۳: برآورد λ برای $n = ۵$، $p = ۱۰۰$، $T = D_s$ و$df = ۵$

λ	خطای استاندارد	میانگین شبیه‌سازی‌شده
بهینهλ	۰	۰/۹۸۳۶
λ_{FS}	۰/۰۲۲۱	۰/۹۸۷۴
λ_{SS}	۰/۰۱۴۴	۰/۷۹۴۲

جدول ۵۰.۳: برآورد λ برای $n = ۵$، $p = ۱۰۰$، $T = D_s$ و$= ۱۰$

λ	خطای استاندارد	میانگین شبیه‌سازی‌شده
بهینهλ	۰	۰/۹۸۵۶
λ_{FS}	۰/۰۲۲۷	۰/۹۸۷۸
λ_{SS}	۰/۰۱۴۳	۰/۷۹۴۶

جدول ۵۱.۳: برآورد λ برای $n = ۵$، $p = ۱۰۰$، $T = D_s$ و$df = ۱۰۰$

λ	خطای استاندارد	میانگین شبیه‌سازی‌شده
بهینهλ	۰	۰/۹۸۵۱
λ_{FS}	۰/۰۲۲۶	۰/۹۸۶۶
λ_{SS}	۰/۰۱۴۲	۰/۶۹۴۶

جدول ۵۲.۳: برآوردگر انقباضی برای $p = ۱۰۰$، $n = ۵$، $T = D_s$ و $df = ۵$

برآوردگر	مخاطره	درصد بهبود نسبی
S	۶۳۹/۹۵۱۲	۰
S_{FS}	۴۳۳۸/۵۸	−۵۷۷/۹۵۴۸
S_{SS}	۲۸۳۶/۵۱۳	−۳۴۳/۲۳۸۹

جدول ۵۳.۳: برآوردگر انقباضی برای $p = ۱۰۰$، $n = ۵$، $T = D_s$ و $df = ۱۰$

برآوردگر	مخاطره	درصد بهبود نسبی
S	۵۸۰/۶۴۴	۰
S_{FS}	۳۷۷۶/۹۳۶	−۵۵۰/۴۷۳۶
S_{SS}	۲۵۰۸/۷۰۴	−۳۳۲/۰۵۵۵

جدول ۵۴.۳: برآوردگر انقباضی برای $p = ۱۰۰$، $n = ۵$، $T = D_s$ و $df = ۱۰۰$

برآوردگر	مخاطره	درصد بهبود نسبی
S	۵۳۱/۲۵۳۱	۰
S_{FS}	۳۵۳۳/۰۹۸	−۵۶۵/۰۴۹۸
S_{SS}	۲۳۱۵/۱۰۹	−۳۳۵/۷۸۲۷

با توجه به نتایج ارائه شده می‌توان فهمید که با افزایش درجه آزادی توزیع t، p متغیره مخاطره و مقدار عددی پارامتر بهینه کاهش می‌یابد. همچنین با کاهش n و افزایش p مقدار عددی پارامتر انقباضی بزرگتر می‌شود.

پیوست آ

خلاصه، نتیجه‌گیری، پیشنهادات برای آینده تحقیق و برنامه‌های نرم‌افزار R

برآوردگر مناسب برای ماتریس واریانس-کوواریانس، زمانی که $n > p$ می‌باشد، برآوردگر حاصل از روش درستنمایی ماکزیمم است. اما با افزایش بعد دیگر شرایط خوب را برای ماتریس واریانس-کوواریانس نداشته و باید به دنبال برآوردگری بود که با توجه به تابع زیان در نظرگرفته شده دارای مخاطره کمتری باشد. در این پایان‌نامه چهار مجموعه از برآوردگرهای انقباضی در برآورد ماتریس واریانس-کوواریانس یک توزیع نرمال $p-$متغیره معرفی و مورد مقایسه قرار گرفت. مطالعات شبیه‌سازی و مثال واقعی نشان داد که برآوردگر سان و فیشر (۲۰۱۱) در ابعاد بالاتر، تحت فرض نرمال بودن، کارایی بیشتری در مقایسه با سایر برآوردگرها دارد. همچنین با استفاده از نتایج شبیه‌سازی دیده می‌شود با افزایش بعد برآوردگر ماتریس کوواریانس نمونه (S) نمی‌تواند به خوبی ماتریس واریانس-کوواریانس را برآورد نماید.
همچنین توانستیم با استفاده از تکنیک قضیه ۳.۲.۳ برای سه ماتریس هدف معرفی شده، پارامتر انقباضی مناسب و به دنبال آن برآوردگر مطلوبی برای ماتریس واریانس-کوواریانس توزیع t چندمتغیره بیابیم. با توجه به نتایج به‌دست آمده در این پایان‌نامه و موضوعات مطرح‌شده می‌توان در آینده بر موارد زیر تحقیق کرد.

۱- برآوردگر انقباضی مطرح‌شده در این پایان‌نامه ترکیبی خطی محدب بود؛ می‌توان برآوردگری تعریف کرد که حالت خطی نداشته باشد. یا از ماتریس‌های هدف غیرقطری در ترکیب مورد نظر استفاده کرد و برآوردگرهای حاصل را با حالتی که ماتریس هدف قطری در نظر گرفته می‌شود، مقایسه نمود (واریانس، اریبی و میزان خطا).

۲- تابع زیان مورد استفاده در این پایان‌نامه، تابع زیان درجه دو بود. می‌توان از توابع زیان

دیگر مانند تابع زیان لینکس [۱]، تابع زیان آنتروپی [۲] و ... استفاده نمود و رفتار برآوردگر را مورد بررسی قرار داد.

۳- فرضیه توزیع نرمال یا t بودن داده‌ها را حذف نمود و توزیع‌های دیگر بیضی‌گون را که در فصل اول به آن اشاره شد به عنوان مدل جامعه در نظر گرفت.

۴- تعمیم برای حالت بیزی.

[۱] Linex
[۲] Antropy

آ.۱ دستورهای نرم‌افزار R

در این قسمت دستورهای مورد نیاز برای رسم نمودارهای موجود در فصل اول و فصل دوم با استفاده از نرم‌افزار R آمده است.

دستور ۱. برای رسم نمودارهای تابع چگالی و منحنی‌های تراز در توزیع نرمال دومتغیره در شکل ۱.۱ از دستورهای زیر در نرم‌افزار R استفاده می‌کنیم.

```
f<-function(x,y){
 z<-((det(Sigma)^(-0.5))/(2*pi))*exp(-(0.5)*(x^2+0.1*y^2))
}
z<-outer(x,y,f)
persp(x,y,z,theta=45,phi=30,expand=0.6,ltheta=120,shade=0.75,
ticktype="detailed",xlab="X",ylab="Y",zlab="f(x,y)")
library(lattice)
contourplot(z)
```

دستور ۲. برای رسم نمودارهای تابع چگالی و منحنی‌های تراز در توزیع t دومتغیره در شکل ۲.۱ از دستورهای زیر در نرم‌افزار R استفاده می‌کنیم.

```
f<-function(x,y){
 z<-(((det(Sigma)^(-0.5))*gamma(3/2))/((pi)*gamma(1/2)))
*((1+(x^2+0.1*y^2))^(-3/2))
}
z<-outer(x,y,f)
persp(x,y,z,theta=45,phi=30,expand=0.6,ltheta=120,shade=0.75,
 ticktype="detailed",xlab="X",ylab="Y",zlab="f(x,y)")
library(lattice)
contourplot(z)
```

دستور ۳. برای رسم نمودارهای تابع چگالی و منحنی‌های تراز در توزیع کوشی دومتغیره در شکل ۳.۱ از دستورهای زیر در نرم‌افزار R استفاده می‌کنیم.

```
Sigma<-matrix(c(1,0,0,10),nrow=2)
y<-x<-seq(-3,3,length=50)
f<-function(x,y){
z<-(((det(Sigma)^(-0.5))*gamma(3/2))/((pi)*gamma(1/2)))
*((1+(x^2+0.1*y^2))^(-3/2))}
```

```
z<-outer(x,y,f)
persp(x,y,z,theta=45,phi=30,expand=0.6,ltheta=120,shade=0.75,
ticktype="detailed",xlab="X",ylab="Y",zlab="f(x,y)")
library(lattice)
contourplot(z)
```

دستور ۴. برای رسم نمودارهای تابع چگالی و منحنی‌های تراز در توزیع لجستیک دومتغیره در شکل ۴.۱ از دستورهای زیر در نرم‌افزار R استفاده می‌کنیم.

```
Sigma<-matrix(c(1,0,0,10),nrow=2)
y<-x<-seq(-3,3,length=50)
f<-function(x,y){
z<-det((det(Sigma)^(-0.5))*1/(2*pi)^(-1))*(exp(-0.5*(x^2
+0.1*y^2))/((1+exp(-0.5*(x^2+0.1*y^2)))^2))}
z<-outer(x,y,f)
persp(x,y,z,theta=45,phi=30,expand=0.6,ltheta=120,shade=0.75,
ticktype="detailed",xlab="X",ylab="Y",zlab="f(x,y)")
library(lattice)
contourplot(z)
```

دستور ۵. برای رسم نمودارهای تابع چگالی و منحنی‌های تراز در توزیع نمایی توانی دومتغیره در شکل ۵.۱ از دستورهای زیر در نرم‌افزار R استفاده می‌کنیم.

```
Sigma<-matrix(c(1,0,0,10),nrow=2)
y<-x<-seq(-3,3,length=50)
f<-function(x,y){
z<-(((det(Sigma)^(-0.5))*3*gamma(1)*(3^(1/3)))/((2*pi)*gamma(1/3)))*
exp(-(3/2)*((x^2+0.1*y^2)^3))}
z<-outer(x,y,f)
persp(x,y,z,theta=45,phi=30,expand=0.6,ltheta=120,shade=0.75,
ticktype="detailed",xlab="X",ylab="Y",zlab="f(x,y)")
library(lattice)
contourplot(z)
```

دستور ۶. برای رسم نمودارهای تابع چگالی و منحنی‌های تراز در توزیع کاتز دومتغیره در شکل ۶.۱ از دستورهای زیر در نرم‌افزار R استفاده می‌کنیم.

```
Sigma<-matrix(c(1,0,0,10),nrow=2)
y<-x<-seq(-3,3,length=50)
f<-function(x,y){
z<-(((det(Sigma)^(-0.5))*gamma(1)*3)/((2*pi)*gamma(1)))*
exp(-(3/2)*(x^2+0.1*y^2))^3}
z<-outer(x,y,f)
persp(x,y,z,theta=45,phi=30,expand=0.6,ltheta=120,shade=0.75,
ticktype="detailed",xlab="X",ylab="Y",zlab="f(x,y)")
library(lattice)
contourplot(z)
```

دستور ۷. برای رسم نمودارهای تابع چگالی و منحنی‌های تراز در توزیع لاپلاس دومتغیره در شکل ۷.۱ از دستورهای زیر در نرم‌افزار R استفاده می‌کنیم.

```
Sigma<-matrix(c(1,0,0,10),nrow=2)
y<-x<-seq(-3,3,length=50)
f<-function(x,y){
z<-(((det(Sigma)^(-0.5))*(0.5)*gamma(1)*2)/((2*pi)*gamma(2)))*
exp(-((sqrt(2))/2)*((x^2+0.1*y^2)^(1/2)))}
z<-outer(x,y,f)
persp(x,y,z,theta=45,phi=30,expand=0.6,ltheta=120,shade=0.75,
ticktype="detailed",xlab="X",ylab="Y",zlab="f(x,y)")
library(lattice)
contourplot(z)
```

دستور ۸. برای رسم نمودارهای ۱.۲، ۲.۲، ۳.۲ و ۴.۲ مربوط به پارامترهای انقباضی معرفی شده در فصل ۲ از دستورهای زیر در نرم‌افزار R استفاده می‌کنیم.

```
library("Matrix")
library("mnormt")
library("lattice")
library("matrixcalc")
#library("corpcor")
```

```
N=5
n=N-1
m=0.5
M=10.5
p=100
mean=as.vector(c(rep(0,p)))
Norm2=function(A,p){
  matrix.trace(A%*%t(A))/p
}
PD=function(p,m,M){
  d=as.vector(runif(p,m,M))
  lam=diag(d)
  H=diag(1,p,p)
  sigma=H%*%lam%*%t(H)
  return(sigma)
}
sigma=PD(p,m,M)
X=rmvnorm(n+1,mean,sigma)
S=cov(X)
FS<-function(X){
  N=dim(X)[1]
  p=dim(X)[2]
  n=N-1
  S<-cov(X)
  S.2<-S%*%S
  a1.hat=matrix.trace(S)/p
 a2.hat=(n^2/(n-1)/(n+2))*(sum(diag(S.2))-sum(diag(S))^2/n)/p;
  mu.hat=a1.hat
  beta2.hat=a2.hat/n+(p/n)*(a1.hat^2)
  delta2.hat=((n+1)/n)*a2.hat+((p-n)/n)*a1.hat^2
  optimal.landa=beta2.hat/delta2.hat
  return(optimal.landa)
}
```

```
LW<-function(X) {
  n=N-1
  S <- cov(X);
  mu <- matrix.trace(S)/p;
  Iden <- diag(rep(1,p));
  tmp <- S - mu*Iden;
  delta <- matrix.trace(tmp%*%tmp)/p;
  X <- t(X);
 Beta.tmp <- (sum(((X^2)%*%t(X^2)/(n+1))) - sum(S^2) )/((n+1)*p);
  Beta <- min(delta, Beta.tmp);
  alpha <- delta - Beta;
  c(alpha, Beta, delta);
  return(Beta/delta)
}

RBLW<-function(X){
  n=N-1
  Sum=matrix(0,p,p)
  for(i in 1:n)
    Sum=Sum+X[i,]%*%t(X[i,])
  S.hat=(1/n)*Sum
  sum=0
  for(i in 1:n){
   sum=sum+matrix.trace(((X[i,]%*%t(X[i,]))-S.hat)%*%t((X[i,]
    %*%t(X[i,]))-S.hat))/p
  }
 (((n-2)/n)*matrix.trace(S.hat%*%S.hat)+(matrix.trace(S.hat))^2)/((n+2)*
  (matrix.trace(S.hat%*%S.hat)-((matrix.trace(S.hat))^2)/p))
}

FS<-function(X){
  N=dim(X)[1]
  p=dim(X)[2]
  n=N-1
```

```
  S<-cov(X)
  S.2<-S%*%S
  a1.hat=matrix.trace(S)/p
 a2.hat=(n^2/(n-1)/(n+2))*(sum(diag(S.2))-sum(diag(S))^2/n)/p;
  mu.hat=a1.hat
  beta2.hat=a2.hat/n+(p/n)*(a1.hat^2)
  delta2.hat=((n+1)/n)*a2.hat+((p-n)/n)*a1.hat^2
  landa=beta2.hat/delta2.hat
  return(landa)
}

SS<-function(X){
  estimate.lambda(X,verbose=FALSE)
}

l.RBLW=as.vector(c(rep(0,p)))
l.FS=as.vector(c(rep(0,p)))
l.SS=as.vector(c(rep(0,p)))
l.LW=as.vector(c(rep(0,p)))
for(i in 2:p){
  X=rmvnorm(n+1,mean,sigma)
  l.LW[i]=LW(X)
  l.RBLW[i]=RBLW(X)
  l.FS[i]=FS(X)
l.SS[i]=SS(X)
            }
barplot(l.LW,xlab="p",ylab=expression(paste(lambda[LW])))
barplot(l.RBLW,xlab="p",ylab=expression(paste(lambda[RBLW])))
barplot(l.FS,xlab="p",ylab=expression(paste(lambda[FS])))
barplot(l.SS,xlab="p",ylab=expression(paste(lambda[SS])))
```

دستور ۹. برای محاسبه مقادیر بدست آمده در جداول ۱۵.۲ - ۱.۲ از دستورهای زیر در نرم‌افزار R استفاده می‌کنیم.

```
library("base");library("colorspace");library("mnormt");library("lattice");
```

```
library("Matrix")
library("Rcpp");library("minqa");library("lme4");library("fastGHQuad");
library("mixAK")
library("mvtnorm");library("matrixcalc");library("corpcor")

PD <- function(p,m,M) {
  d <- as.vector(runif(p, m, M))
  Lam <- diag(d)
  H<- rRotationMatrix(1, p)
  Sigma <- H%*%Lam%*%t(H)
  return(Sigma)
}

Norm2=function(A,p){
  matrix.trace(A%*%t(A))/p
}

FS<-function(X){
  N=dim(X)[1]
  p=dim(X)[2]
  n=N-1
  S<-cov(X)
  S.2<-S%*%S
  a1.hat=matrix.trace(S)/p
 a2.hat=(n^2/(n-1)/(n+2))*(sum(diag(S.2))-sum(diag(S))^2/n)/p;
  mu.hat=a1.hat
  beta2.hat=a2.hat/n+(p/n)*(a1.hat^2)
  delta2.hat=((n+1)/n)*a2.hat+((p-n)/n)*a1.hat^2
  optimal.landa=beta2.hat/delta2.hat
  return(optimal.landa)
}

LW<-function(X) {
  n=N-1
```

```
  S <- cov(X);
  mu <- matrix.trace(S)/p;
  Iden <- diag(rep(1,p));
  tmp <- S - mu*Iden;
  delta <- matrix.trace(tmp%*%tmp)/p;
  X <- t(X);
 Beta.tmp <- (sum(((X^2)%*%t(X^2)/(n+1))) - sum(S^2) )/((n+1)*p);
  Beta <- min(delta, Beta.tmp);
  alpha <- delta - Beta;
  c(alpha, Beta, delta);
  return(Beta/delta)
}

RBLW<-function(X){
  n=N-1
  Sum=matrix(0,p,p)
  for(i in 1:n)
    Sum=Sum+X[i,]%*%t(X[i,])
  S.hat=(1/n)*Sum
  sum=0
  for(i in 1:n){
   sum=sum+matrix.trace(((X[i,]%*%t(X[i,]))-S.hat)%*%t((X[i,]%*%
t(X[i,]))-S.hat))/p
  }
 (((n-2)/n)*matrix.trace(S.hat%*%S.hat)+(matrix.trace(S.hat))^2)/((n+2)*
(matrix.trace(S.hat%*%S.hat)-((matrix.trace(S.hat))^2)/p))
}

SS<-function(X){
  estimate.lambda(X,verbose=FALSE)
 }

N=5
m=0.5
```

```
M=10.5
p=100
mean=rep(0,p)
Sigma=PD(p,m,M)
X=rmvnorm(n+1,mean,Sigma)
S=cov(X)
mu=matrix.trace(Sigma)/p
optimal.lambda=(Norm2(S-Sigma,p))/(Norm2(S-mu*diag(rep(1,p)),p))
optimal.lambda
m=1000
l.LW=l.RBLW=l.SS=l.FS=numeric(m)
for(i in 1:m){
  X=rmvnorm(n+1,mean,Sigma)
  l.LW[i]=LW(X)
  l.RBLW[i]=RBLW(X)
  l.FS[i]=FS(X)
  l.SS[i]=SS(X)}
mean(l.LW)
mean(l.RBLW)
mean(l.FS)
mean(l.SS)
sd(l.LW)
sd(l.RBLW)
sd(l.FS)
sd(l.SS)
T=mu*diag(1,p,p)
S.Star=function(l)
  (1-l)*S+(l*T)
Norm2(S-Sigma,p)
Norm2(S.Star(mean(l.LW))-Sigma,p)
Norm2(S.Star(mean(l.RBLW))-Sigma,p)
Norm2(S.Star(mean(l.FS))-Sigma,p)
Norm2(S.Star(mean(l.SS))-Sigma,p)
PRIAL<-function(S.star)
```

```
  ((Norm2(S-Sigma,p)-Norm2(S.star-Sigma,p))/Norm2(S-Sigma,p))*100
PRIAL(S.Star(mean(l.LW)))
PRIAL(S.Star(mean(l.RBLW)))
PRIAL(S.Star(mean(l.FS)))
PRIAL(S.Star(mean(l.SS)))
PRIAL(S)
```

دستور ۱۰. برای رسم نمودارهای ۷.۲، ۶.۲ و ۸.۲ مربوط به زمان مطالعه برآوردگرها در فصل ۲ از دستورهای زیر در نرم‌افزار R استفاده می‌کنیم.

```
# Here we calculate the respective shrinkage estimators.
# We note that the sames from the MVNormal will be the
# same since the seeds for the RNG are reset.
runDiagShrink <- function(n, Sigma) {
p <- dim(Sigma)[1];
X <- rmvnorm(n+1, rep(0, p), Sigma);
S.new <- cov.Diag.Shrink(X);
0; #Doesn't matter what we return, we are interested in time.
}
runSchaefShrink <- function(n, Sigma) {
p <- dim(Sigma)[1];
X <- rmvnorm(n+1, rep(0, p), Sigma);
S.schaef <- cov.shrink(X, lambda.var=0, verbose=FALSE);
0;
}
wrapperRuns <- function(n, p, m, seed) {
# We find a random Positive Definite Matrix Sigma
ev <- runif(p, 0.5, 10.5);
Sigma <- Posdef(p, ev);
# Set the seed to control the samples, calculate the
# time necessary for our estimator.
if(!(seed==Inf)) set.seed(seed,kind=NULL);
N <- rep(n,m);
ptm <- proc.time();
res1 <- sapply(N, runDiagShrink, Sigma=Sigma);
```

```
time.new <- (proc.time()-ptm)[3];
# Reset the seed, calculate the time for Schaefer
# and Strimmers estimator.
if(!(seed==Inf)) set.seed(seed,kind=NULL);
ptm <- proc.time();
res1 <- sapply(N, runSchaefShrink, Sigma=Sigma);
time.sch <- (proc.time()-ptm)[3];
# Return the two run-times for m-runs
c(time.new, time.sch);
}
timingDriverN <- function(p=20, m=1000, seed=Inf) {
if(!(seed==Inf)) set.seed(seed,kind=NULL);
x.axis<-c(10, 30, 50, 100, 150, 200, 300);
y.axis<-sapply(x.axis, wrapperRuns, p=p, m=m, seed=seed);
x.limit <- max(x.axis)+25;
y.limit <- ceiling(max(y.axis)+2);
plot(x.axis, y.axis[1,], type='p', pch=22, bg="red", col="red",
xlim=c(0,x.limit), ylim=c(0,y.limit), xlab="Sample Size n+1",
ylab="Time (s)", main="Timing Study for Sample Size" );
lines(x.axis, y.axis[1,], lty="solid", col="red");
points(x.axis, y.axis[2,], pch=23, bg="blue", col="blue");
lines(x.axis, y.axis[2,], lty="dashed", col="blue");
legend(5, (y.limit), legend = c("New Estimator", "Schaefer Estimator"),
col = c("red", "blue"), pch=c(22, 23),
pt.cex=1, pt.bg=c("red","blue"), lty = c(1, 2))
}
```

دستور ۱۱. برای محاسبه مقادیر بدست آمده در جداول ۱.۳-۵۴.۳ از دستورهای زیر در نرم‌افزار R استفاده می‌کنیم.

```
library("base");library("colorspace");library("mnormt");library("lattice");
library("Matrix")
library("Rcpp");library("minqa");library("lme4");library("fastGHQuad");
library("mixAK")
library("mvtnorm");library("matrixcalc");library("corpcor")
```

```
PD <- function(p,m,M) {
  d <- as.vector(runif(p, m, M))
  Lam <- diag(d)
  H<- rRotationMatrix(1, p)
  Sigma <- H%*%Lam%*%t(H)
  return(Sigma)
}

Norm2=function(A,p){
  matrix.trace(A%*%t(A))/p
}

FS<-function(X){
  N=dim(X)[1]
  p=dim(X)[2]
  n=N-1
  S<-cov(X)
  S.2<-S%*%S
  a1.hat=matrix.trace(S)/p
 a2.hat=(n^2/(n-1)/(n+2))*(sum(diag(S.2))-sum(diag(S))^2/n)/p;
  mu.hat=a1.hat
  beta2.hat=a2.hat/n+(p/n)*(a1.hat^2)
  delta2.hat=((n+1)/n)*a2.hat+((p-n)/n)*a1.hat^2
  optimal.landa=beta2.hat/delta2.hat
  return(optimal.landa)
}

LW<-function(X) {
  n=N-1
  S <- cov(X);
  mu <- matrix.trace(S)/p;
  Iden <- diag(rep(1,p));
  tmp <- S - mu*Iden;
  delta <- matrix.trace(tmp%*%tmp)/p;
```

```
  X <- t(X);
 Beta.tmp <- (sum(((X^2)%*%t(X^2)/(n+1))) - sum(S^2) )/((n+1)*p);
  Beta <- min(delta, Beta.tmp);
  alpha <- delta - Beta;
  c(alpha, Beta, delta);
  return(Beta/delta)
}

RBLW<-function(X){
  n=N-1
  Sum=matrix(0,p,p)
  for(i in 1:n)
    Sum=Sum+X[i,]%*%t(X[i,])
  S.hat=(1/n)*Sum
  sum=0
  for(i in 1:n){
   sum=sum+matrix.trace(((X[i,]%*%t(X[i,]))-S.hat)%*%t((X[i,]%*%
t(X[i,]))-S.hat))/p
  }
 (((n-2)/n)*matrix.trace(S.hat%*%S.hat)+(matrix.trace(S.hat))^2)/((n+2)*
(matrix.trace(S.hat%*%S.hat)-((matrix.trace(S.hat))^2)/p))
}

SS<-function(X){
  estimate.lambda(X,verbose=FALSE)
 }

N=5
m=0.5
M=10.5
p=100
mean=rep(0,p)
Sigma=PD(p,m,M)
X=rt(n+1,mean,Sigma,df)
```

```
S=cov(X)
mu=matrix.trace(Sigma)/p
optimal.lambda=(Norm2(S-Sigma,p))/(Norm2(S-mu*diag(rep(1,p)),p))
optimal.lambda
m=1000
l.LW=l.RBLW=l.SS=l.FS=numeric(m)
for(i in 1:m){
  X=rt(n+1,mean,Sigma,df=5)
  l.LW[i]=LW(X)
  l.RBLW[i]=RBLW(X)
  l.FS[i]=FS(X)
  l.SS[i]=SS(X)}
mean(l.LW)
mean(l.RBLW)
mean(l.FS)
mean(l.SS)
sd(l.LW)
sd(l.RBLW)
sd(l.FS)
sd(l.SS)
T=mu*diag(1,p,p)
S.Star=function(l)
  (1-l)*S+(l*T)
Norm2(S-Sigma,p)
Norm2(S.Star(mean(l.LW))-Sigma,p)
Norm2(S.Star(mean(l.RBLW))-Sigma,p)
Norm2(S.Star(mean(l.FS))-Sigma,p)
Norm2(S.Star(mean(l.SS))-Sigma,p)
PRIAL<-function(S.star)
 ((Norm2(S-Sigma,p)-Norm2(S.star-Sigma,p))/Norm2(S-Sigma,p))*100
PRIAL(S.Star(mean(l.LW)))
PRIAL(S.Star(mean(l.RBLW)))
PRIAL(S.Star(mean(l.FS)))
PRIAL(S.Star(mean(l.SS)))
```

```
PRIAL(S)
```

پیوست ب

جبرخطی

برای اطلاع دقیق‌تر از تعاریف و مطالب این فصل می‌توان به میرهد (۱۹۸۲)، رنچر [۱] (۲۰۰۲)، ارقامی (۱۳۸۳) و کرمی (۱۳۹۲) مراجعه کرد.

ب.۱ تعاریف و قضایا

تعریف ب.۱.۱. تعریف ماتریس
یک ماتریس، آرایه‌ای $m \times n$ تایی از اعداد است که دارای m سطر و n ستون است. اعضای ماتریسی مانند A بصورت (a_{ij}) می‌باشد که در آن a_{ij} عنصر سطر iام از ستون jام است. برای $n = ۱$، A را یک بردار می‌نامیم.

تعریف ب.۲.۱. اثر ماتریس
ماتریس $A = (a_{ij})$، $n \times n$ مفروض است. مجموع درایه‌های روی قطر اصلی A را اثر ماتریس A می‌نامیم و با نماد $tr(A)$ نشان می‌دهیم.

$$tr(A) = \sum_{i=۱}^{n} a_{ii}$$

تعریف ب.۳.۱. ترانهاده‌ی ماتریس [۲]
اگر جای سطرها و ستون‌های ماتریس Σ را با ستون یا سطرهای آن تعویض نماییم، ماتریس حاصل ترانهاده ماتریس Σ نامیده می‌شود و با نماد Σ^T نمایش داده می‌شود.

$$\Sigma^T = (\sigma_{ij})^T = (\sigma_{ji})$$

که در آن σ_{ij} عنصر موجود در سطر i ام و ستون j ام است.

قضیه ب.۴.۱. *به ازای هر ماتریس Σ داریم،* $(\Sigma^T)^T = \Sigma$.

[۱] Rencher
[۲] Transpose matrix

تعریف ب.۱.۵. تعریف مقادیر ویژه ماتریس A
فرض کنید ماتریس Σ از مرتبه‌ی $p \times p$ باشد آن‌گاه ریشه‌های معادله $det(\Sigma - \lambda I_p) = ۰$ مقادیر ویژه ماتریس A نامیده می‌شوند. که در این قسمت λ مقدار ویژه می باشد.

تعریف ب.۱.۶. ماتریس مربعی
ماتریس Σ مربعی است اگر تعداد سطرها و ستون‌های آن با هم برابر باشد و بصورت $\Sigma_{p\times p}$ نشان داده می شود.

تعریف ب.۱.۷. فرض کنید $\Sigma = (\sigma_{ij})$ یک ماتریس مربع از مرتبه‌ی p باشد در این صورت:

i) Σ را ماتریسی نامنفرد [۳] گوییم اگر $det(\Sigma) \neq ۰$ باشد.

ii) Σ را ماتریسی قطری [۴] گوییم اگر به ازای هر $i \neq j$، $\sigma_{ij} = ۰$ باشد و با نماد $Diag(\Sigma)$ یا $diag(\sigma_{۱۱}, \ldots, \sigma_{pp})$ نمایش می‌دهیم.

iii) Σ را ماتریسی همانی [۵] گوییم اگر قطری باشد و به ازای $i = ۱, \ldots, p$، $\sigma_{ii} = ۱$ و آن را با نماد I_p نمایش می‌دهیم.

iv) Σ را ماتریسی متقارن [۶] گوییم اگر $\Sigma = \Sigma^T$ یا به ازای هر $i \neq j$ داشته باشیم $\sigma_{ij} = \sigma_{ji}$. همچنین ماتریس Σ را پاد-متقارن [۷] گوییم اگر $\Sigma^T = -\Sigma$

v) Σ را ماتریسی بالا مثلثی گوییم اگر به ازای هر $i > j$ داشته باشیم $\Sigma_{ij} = ۰$.

vi) Σ را یک ماتریس پایین مثلثی گوییم اگر به ازای هر $i < j$ داشته باشیم $\Sigma_{ij} = ۰$.

vii) Σ را ماتریسی معین مثبت گوییم اگر Σ متقارن باشد و برای هر بردار غیر صفر p بعدی $\mathbf{x}$، $\mathbf{x}^T\Sigma\mathbf{x} > ۰$ که بصورت $\Sigma > ۰$ نمایش می‌دهیم.

تعریف ب.۱.۸. دترمینان
دترمینان ماتریس $\Sigma_{p\times p}$ را با نماد $|\ |$ نشان داده که تابعی اسکالر است و به صورت زیر تعریف می‌شود.

$$\begin{cases} |\Sigma| = \sigma_{۱۱} & p = ۱ \\ |\Sigma| = \sum_{j=۱}^{p} \sigma_{۱j}|\Sigma_{۱j}|(-۱)^{۱+j}, & p > ۱ \end{cases}$$

که در آن $\Sigma_{۱j}$ ماتریسی $(n-۱) \times (n-۱)$ است و با حذف jامین سطر و ستون ماتریس Σ بدست می‌آید.

---
[۳] Non-singular
[۴] Diagonal matrix
[۵] Identity matrix
[۶] Symmetric matrix
[۷] Skew-symmetric

تعریف ب.۱.۹. ماتریس نامنفرد (وارون پذیر) [۸]
ماتریس $p-$بعدی $\Sigma = (\sigma_{ij})$ را نامنفرد می‌نامیم هرگاه $|\Sigma| \neq ۰$ و منفرد [۹] (وارون ناپذیر) گوییم در صورتی که $|\Sigma| = ۰$.

تعریف ب.۱.۱۰. وارون ماتریس [۱۰]
ماتریس $p-$بعدی نامنفرد $\Sigma = (\sigma_{ij})$ را در نظر بگیرید. در این صورت ماتریس $p-$بعدی $\Lambda = (\lambda_{ij})$ وجود دارد به طوری که $\Sigma\Lambda = I_p$، که در آن $\Lambda = \Sigma^{-۱}$ وارون ماتریس Σ است و به صورت زیر محاسبه می‌شود.

$$\lambda_{ij} = \frac{|\Sigma_{ij}|(-i)^{i+j}}{|\Sigma|}.$$

که در آن $|\Sigma_{ij}|(-i)^{i+j}$ همسازه متناظر با عنصر σ_{ij} است.

تعریف ب.۱.۱۱. بردارمتعامد [۱۱]
دو بردار $\mathbf{x} = (X_۱, \ldots, X_p)$ و $\mathbf{y} = (Y_۱, \ldots, Y_p)$ را متعامد می‌نامند، هرگاه

$$\mathbf{x}^T\mathbf{y} = X_۱Y_۱ + X_۲Y_۲ + \ldots + X_pY_p = ۰$$

تعریف ب.۱.۱۲. ماتریس $\Sigma_{p\times p} = (\sigma_۱, \sigma_۲, \ldots, \sigma_p)$ متعامد است، هرگاه ستون‌های آن بردارهای متعامد باشند.

تعریف ب.۱.۱۳. تجزیه طیفی [۱۲]
فرض کنید Σ ماتریسی متقارن $p \times p$ با مقادیر ویژه $\lambda_۱ \leq, \ldots, \leq \lambda_p$ و Γ، ماتریس متعامد با بعد p باشد، در این صورت تجزیه طیفی ماتریس Σ، به صورت زیر است:

$$\Sigma = \Gamma D \Gamma^T$$

که در آن D، ماتریسی قطری به صورت زیر می‌باشد:

$$D = \begin{pmatrix} \lambda_۱ & ۰ & \ldots & ۰ \\ ۰ & \lambda_۲ & \ldots & ۰ \\ \vdots & \vdots & \ddots & \vdots \\ ۰ & ۰ & \ldots & \lambda_p \end{pmatrix}$$

تعریف ب.۱.۱۴. تجزیه چولسکی [۱۳]
تجزیه چولسکی یک ماتریس متقارن معین مثبت بصورت $\Sigma = L^TL$ است که در آن، L یک ماتریس بالا مثلثی است.

---

[۸] Nonsingular matrix
[۹] Singular
[۱۰] Inverse of a matrix
[۱۱] Orthogonal vector
[۱۲] Spectral decomposition
[۱۳] Cholesky decomposition

مراجع

[۱] آرشی م، (۱۳۸۷)، *برآوردگرهای بهبودیافته در مدل رگرسیون خطی چندگانه با خطاهای دارای توزیع بیضی‌گون*، رساله دکتری، دانشکده علوم ریاضی، دانشگاه فردوسی مشهد.

[۲] آرمانفر س، (۱۳۸۹)، *مباحثی بر توزیع t ماتریسی*، پایان نامه کارشناسی ارشد، دانشکده علوم ریاضی، دانشگاه آزاد اسلامی مشهد.

[۳] ارقامی ن. ر، (۱۳۸۳)، *جبر خطی برای آمار*، چاپ ششم، انتشارات دانشگاه پیام نور، تهران.

[۴] ایرانمنش ا، (۱۳۹۱)، *مباحثی در توزیع بیضی‌گون با تاکید بر توزیع t-استیودنت چندمتغیره*، رساله دکتری، دانشکده علوم ریاضی، دانشگاه فردوسی مشهد.

[۵] پارسیان ا، (۱۳۸۶)، *مبانی آمار ریاضی*، چاپ پنجم، انتشارات دانشگاه اصفهان، اصفهان.

[۶] توره زاده س، (۱۳۹۲)، *برآورد انقباضی موجک در حالت بیزی تحت تابع زیان لینکس*، پایان نامه کارشناسی ارشد، دانشکده علوم ریاضی، دانشگاه شاهرود.

[۷] جعفری ز، (۱۳۸۹)، *برآورد نوع استاین میانگین جامعه*، پایان نامه کارشناسی ارشد، دانشکده علوم ریاضی، دانشگاه آزاد اسلامی مشهد.

[۸] حجتی م ج، (۱۳۹۲)، *مباحثی در توزیع‌های بیضیگون با بررسی چند توزیع خاص*، پایان نامه کارشناسی ارشد، دانشکده علوم ریاضی، دانشگاه آزاد اسلامی مشهد.

[۹] رودین و، (۱۳۷۳)، *اصول آنالیز ریاضی*، عالم‌زاده ع. ا، چاپ هشتم، انتشارات علمی و فنی، تهران.

[۱۰] کرمی کبیر ح، (۱۳۹۲)، *برآورد انقباضی در مدل‌های محدود شده*، پایان نامه کارشناسی ارشد، دانشکده علوم ریاضی، دانشگاه شاهرود.

[۱۱] ماردیا، ک و کنت، ج و بی‌بی، ج. (۱۳۷۶) *تحلیل چندمتغیره*، ترجمه محمد مهدی طباطبایی، مرکز نشر دانشگاهی.

[۱۲] مود، ا، م و گریبل، ف، آ و بوز، د، س. (۱۳۷۷) *مقدمه‌ای بر نظریه آمار*، ترجمه دکتر علی مشکانی، انتشارات دانشگاه فردوسی.

[۱۳] نوروزی‌راد م، (۱۳۹۰)، *مباحثی در برآورد تابع زیان*، پایان‌نامه کارشناسی ارشد، دانشکده علوم پایه، دانشگاه آزاد اسلامی واحد مشهد.

[14] Anderson, T. W., Fang, K. T. and Hsu, H. (1986), Maximum-likelihood estimates and likelihood-ratio criteria for multivariate elliptically contoured distribution, *The Canad. J. Statist.*, **14**, 55-59.

[15] Anderson, T. W. (2003), *An introduction to multivariate statistical analysis*, Wiley Series in Probability and Statistics, Wiley-Interscience [John Wiley & Sons], Hoboken, NJ, third edition.

[16] Andrews, D. F. and Mallows, C. L. (1974), Scale mixtures of normal distribution, *J. Royal. Statist. Soc.*, B, **36**(1), 99-102.

[17] Bai, Z. D., Yin, Y. Q. (1993), Limit of the smallest eigenvalue of a large-dimensional sample covariance matrix, *Ann. Statist*, **21**(3), 1275–1294.

[18] Bickel, P. J., Levina, E. (2008), Regularized estimation of large covariance matrices, *Ann. Statist.*, **36**(1), 199–227.

[19] Blatberg, R. C. and Gonedes, N. J. (1974), A comparison of the stable and student distributions as statistical models for stock prices, *J. Business.*, **47**(2), 224–280.

[20] Breusch, T. S., Robertson, J. C. and Welsh, A. H. (1997), The emperor's new clothes: a critique of the multivariate t regression model, Statistica Neerlandica, **51**(3), 269-286.

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

[22] Chaudhuri, S., Drton, M. and Thomas, S. R. (2007), Estimation of a covariance matrix with zeros, *Biometrika.*, **94**(1), 199-216.

[23] Chen, Y., Wiesel, A., Hero, A. O. (2009), Shrinkage estimation of high dimensional covariance matrices. In: Acoustics, Speech, and Signal Processing, *IEEE International Conference* ., 2937–2940.

[24] Chen, Z. and Dunson, D. B. (2003), Random effects selection in linear mixed models, *Biometrics.*, **59**(4), 762-769.

[25] Chu, K. C. (1973), Estimation and decision for linear systems with elliptically random process, *IEEE Transactions on Automatic Control.*, **18**(486), 499-505.

[26] Corbeil, R. R. and Searle, S. R. (1976), Restricted maximum likelihood (REML) estimation of variance components in the mixed model, *Technometrics.*, **18**(1), 31-38.

[27] Cornish, E.A. (1954), The multivariate t-distribution associated with a set of normal sample deviates, *Austr. J. Physics.*, **7**, 531–542.

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

[29] Daniels, J. M. (2005), A class of shrinkage priors for the dependence structure in longitudinal data, *J. Statist. Plann. Inf.*, **127**(12), 119-130.

[30] Daniels, J. M. (2006), Bayesian modeling of several covariance matrices and some results on propriety of the posterior for linear regression with correlated and/or heterogeneous errors, *J. Mult. Ann. Statist.*, **97**(5), 1185-1207.

[31] Dey, D. K. and Srinivasan, C. (1985), Estimation of a covariance matrix under Stein's loss, *Ann. Statist.*, **13**(4), 1581-1591.

[32] Eaton, M. L and Olkin, I. (1987), Best equivariant estimators of a Cholesky decomposition, *Ann. Statist.*, **15**(4), 1639-1650.

[33] Efron, B. (1975), *Biased versus unbiased estimation.*, Advances in Math. **16**(3), 259–277.

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

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

[36] El Karoui, N. (2008), Operator norm consistent estimation of large dimensional sparse covariance matrices, *Ann. Statist.*, **36**(6), 2717–2756.

[37] Fan, J., Fan, Y., Lv, J. (2008), High dimensional covariance matrix estimation using a factor model, *J. Econometrics.*, **147**(1), 186–197.

[38] Fan, J. and Wu, Y. (2008), Semiparametric estimation of covariance matrixes for longitudinal data, *J. Amer. Statist. Assoc.*, **103**(484), 1520-1533.

[39] Fang, K. T., Katz S. and Ng K. W. (1990), Symmetric Multivariate and Related Distribution, Chapman and Hall, London, New York.

[40] Fisher, T. j (2009), On the testing and estimation of high-dimensional covariance Matrices, Clemson University, Doctor of Philosophy Mathematical Sciences.

[41] 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.

[42] Fraley, C. and Burns, P. J. (1995), Large-scale estimation of variance and covariance components, *SIAM J. Sci. Comput.*, **16**(1), 192-209.

[43] Haff, L. R. (1979), Estimation of the inverse covariance matrix: random mixtures of the inverse Wishart matrix and the identity, *Ann. Statist.*, **9**(5), 1132.

[44] Haff, L. R. (1981), Identity for the wishart distribution with applications, *J. Mult. Ann. Statist.*, **11**(4), 608-609.

[45] Haff, L. R. (1991), The variational form of certain Bayes estimators, *Ann. Statist.*, 19(3), 1163-1190.

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

[47] Huang, J., Liu, N., Pourahmadi, M. and Liu, L. (2006), Covariance matrix selection and estimation via penalised normal likelihood, *Biometrika.*, **93**, 85–98.

[48] Gupta, A. K. and Nagar, D. K. (2000), *Matrix variate distributions*, Chapman Hall.

[49] Guthke, R., Toepfer, S., Reischer, H., Duerrschmid K., Schmidt-Heck, W., and Bayer, K. (2004), Reverse engineering of the stress response during expression of a recombinant protein. In *Proceedings of the EUNITE symposium.*, 407-412.

[50] James, W. and Stein, C. (1961), Estimation with quadratic loss, In *Proceeding of the 4th Berkeley Symposium.*, **1**, 361-379.

[51] Johnstone, I. M. (2001), On the distribution of the largest eigenvalue in principal components analysis, *Ann. Statist.*, **29**, 295–327.

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

[53] Lam, C. and Fan, J. (2007), Sparsistency and rates of convergence in large covariance matrices estimation. *Technical report*, Princeton University.

[54] Lange, K. L., Little, J. A and Taylor, J. M. G. (1989), Robust Statistical modeling using the t distribution, *J. Amer. Statist. Assoc.*, **84**, 881-896.

[55] Lange, K. and Sinsheimer, J. S. (1993), Normal independent distributions and their applications in robust regression, *J. Comp. Graph. Statist.*, **2**, 175-198.

[56] Ledoit, O., Wolf, M. (2003), Improved estimation of the covariance matrix of stock returns with an application to portfolio selection, *J. Empir. Finance.*, **10** (5), 603–621.

[57] Ledoit, O., Wolf, M. (2004a), Honey, I shrunk the sample covariance matrix, *J. Portfolio Manage.*, **31**(1), 110–119.

[58] Ledoit, O., Wolf, M. (2004b), A well-conditioned estimator for large-dimensional covariance matrices, *J. Mult. Ann. Statist.*, **88**(2), 365–411.

[59] Muirhead, R. J. (1982), *Aspects of Multivariate Statistical Theory*, John Wiley, New York.

[60] Muirhead, R. J. (1987), Developments in eigenvalue estimation. In *Advan. Mult. Statist. Anal* (A. K. Gupta, ed.) 277–288. Reidel, Dordrecht.

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

[62] Pourahmadi, M. (2007), Cholesky decompositions and estimation of a covariance matrix: orthogonality of variance-correlation parameters, *Biometrika.*, **94**(4), 1006–1013.

[63] Pourahmadi, M., Daniels, M. J., Park, T. (2007), Simultaneous modelling of the Cholesky decomposition of several covariance matrices, *J. Mult. Ann. Statist.*, **98**(3), 568–587.

[64] Rencher C. A. (2002), Methods of mulltivariate analysis, 2nd Ed., John Wiley, New York.

[65] Rothman, A. J., Bickel, P. J., Levina, E. and Zhu, J. (2008), Sparse permutation invariant covariance estimation, *Electron. J. Stat.*, **2**, 494–515.

[66] Roverato, A. (2000), Cholesky decomposition of a hyper inverse Wishart matrix, *Biometrika.*, **87**, 99-112.

[67] Schafer, J and Strimmer, K. (2005), A shrinkage approach to large-scale covariance matrix estimation and implications for functional genomics, *Stat. Appl. Genet. Mol. Biol.*, 4, 28 pp.. Art. 32.

[68] Srivastava, M. S. (2005), Some tests concerning the covariance matrix in high dimensional data, *J. Japan Statist. Soc.*, **35**(2), 251–272.

[69] Smith, S. and Kohn, R. (2002), Parsimonious covariance matrix estimation for longitudinal data, *J. Amer. Statist. Assoc.*, **97**(460), 1141-1153.

[70] 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.

[71] Stein, C. (1981), Estimation of the mean of a Multivariate Normal Distribution, *Ann. Statist.*, 9, 6, 1135-1151.

[72] Sutradhar, B. C. and Ali, M. M. (1989), Estimation of the parameter of a regression model with a multivariate t error variable, *Communications in Statistics.*, A **15**(2), 429-450.

[73] Tiku, M. L., Tan, W. Y. and Balakrishnan, N. (1992), *Robust Inference.*, Marcel Dekker, New york.

[74] Wu, W. B. and Pourahmadi, M. (2009), Banding sample covariance matrices of stationary processes, *Statist. Sinica.*, **19**, 1755–1768.

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

[76] Zellner, A. (1976), Bayesian and non-Bayesian analysis of the regression model with multivariate Student t error terms, *J. Amer. Statist. Assoc.*, **71**, 400-405.

[77] Zou, H., Hastie, T. and Tibshirani, R. (2006), Sparse principal components analysis, *J. Comput. Graph. Statist.*, **15**, 265–286.

واژه‌نامه فارسی به انگلیسی

آنتروپی ..... Antropy

بسته ..... package

برآوردگر انقباضی ..... Shrinkage estimator

برآوردگر انقباضی ماتریس کواریانس از طریق بهینه سازی محدب ..... Shrinkage estimation of covariance matrix through convex optimization

برآورد همزمان ..... Simultaneous estimation

بردار متعامد ..... Orthogonal vector

پرت ..... Outliers

تابع توزیع تباهیده ..... Degenerate distribution function

تابع توزیع ناتباهیده ..... Non-Degenerate distribution function

تابع محدب ..... Convex function

تابع مولد چگالی ..... Density generator

تجزیه چولسکی ..... Cholesky decomposition

توزیع آمیخته مقیاسی ..... Scale mixture distribution

توزیع اسلش ..... Slash distribution

توزیع t استیودنت ..... Student t distribution

توزیع کوشی چندمتغیره ..... Multivariate cauchy distribution

توزیع نرمال آمیخته مقیاسی ..... Scale mixture of normal distribution

توزیع نمایی-توانی ..... Power-exponential distribution

توزیع ویشارت ..... Wishaart distribution

پارامتر انقباضی ..... Shrinkage intensity

توزیع بیضوی ..... Elliptical distribution

توزیع نرمال چندمتغیره ..... Multivariate normal distribution

توزیع t چندمتغیره ..... Multivariate t distribution

دترمینان ..... Determinant

درستنمایی ماکزیمم ..... Maximum likelihood
درصد بهبود نسبی ..... Precentage relative improvement in average loss
رویکرد درستنمایی جریمه شده ..... Penalized likelihood
عدد شرطی ..... Condition number
کارایی ..... Efficiency
کمینه کردن تابع زیان ..... Minimizing of a risk function
فرایند مکانی چندمتغیره ..... Multivariate spatial process
فرین ..... Extrem values
لینکس ..... Linex
ماتریس پادمتقارن ..... Skew-symmetric matrix
ماتریس ترانهاده ..... Transpose matrix
ماتریس دقت ..... Matrix precision
ماتریس قطری ..... Diagonal matrix
ماتریس نامنفرد ..... Nonsingular matrix
ماتریس همانی ..... Identity matrix
ماتریس هدف ..... Target matrix
ماتریس متقارن مثبت ..... Positive definite matrix
ماتریس متقارن ..... Symmetric matrix
متغیر تصادفی تباهیده ..... Degenerate random variable
متغیر تصادفی ناتباهیده ..... Non-Degenerate random variable
میانگین توان دوم خطا ..... Mean squared error
معین ..... Definit
مقعر ..... Convex
منفرد ..... Singular
نامساوی مثلثی ..... Triangle inequality
نرم ..... Norm
نرم فروبنیوس ..... Frobenius norm
وارون ماتریس ..... Inverse of a matrix
همگن ..... Homogeneous

واژه‌نامه انگلیسی به فارسی

آنتروپی ..... Antropy

تجزیه چولسکی ..... Cholesky decomposition

عدد شرطی ..... Condition number

مقعر ..... Convex

تابع محدب ..... Convex function

معین ..... Definit

تابع توزیع تباهیده ..... Degenerate distribution function

تابع مولد چگالی ..... Density generator

ماتریس قطری ..... Diagonal matrix

دترمینان ..... Determinant

کارایی ..... Efficiency

توزیع بیضوی ..... Elliptical distribution

فرین ..... Extrem values

همگن ..... Homogeneous

ماتریس همانی ..... Identity matrix

وارون ماتریس ..... Inverse of a matrix

لینکس ..... Linex

ماتریس دقت ..... Matrix precision

درستنمایی ماکزیمم ..... Maximum likelihood

میانگین توان دوم خطا ..... Mean squared error

کمینه کردن تابع زیان ..... Minimizing of a risk function

توزیع نرمال چندمتغیره ..... Multivariate normal distribution

توزیع t چندمتغیره ..... Multivariate t distribution

توزیع کوشی چندمتغیره ..... Multivariate cauchy distribution

فرایند مکانی چندمتغیره ..... Multivariate spatial process

تابع توزیع ناتباهیده ..... Non-Degenerate distribution function
ماتریس نامنفرد ..... Nonsingular matrix
نرم ..... Norm
بردار متعامد ..... Orthogonal vector
پرت ..... Outliers
بسته ..... package
رویکرد درستنمایی جریمه شده ..... Penalized likelihood
ماتریس متقارن مثبت ..... Positive definite matrix
توزیع نمایی-توانی ..... Power-exponential distribution
درصد بهبود نسبی ..... Precentage relative improvement in average loss
ماتریس هدف ..... Target matrix
ماتریس ترانهاده ..... Transpose matrix
نامساوی مثلثی ..... Triangle inequality
توزیع آمیخته مقیاسی ..... Scale mixture distribution
توزیع نرمال آمیخته مقیاسی ..... Scale mixture of normal distribution
ماتریس پادمتقارن ..... Skew-symmetric matrix
برآورد همزمان ..... Simultaneous estimation
منفرد ..... Singular
برآوردگر انقباضی ..... Shrinkage estimator
پارامتر انقباضی ..... Shrinkage intensity
برآوردگر انقباضی ماتریس کواریانس از طریق بهینه سازی محدب ..... Shrinkage estimation of covariance matrix through convex optimization
ماتریس متقارن ..... Symmetric matrix
توزیع اسلش ..... Slash distribution
توزیع t استیودنت ..... Student t distribution
توزیع ویشارت ..... Wishaart distribution

Aabstract

Many applications require an estimate for the covariance matrix that is non-singular and well-conditioned. As the dimensionality increases, the sample covariance matrix becomes ill-conditioned or even singular. A common approach to estimating the covariance matrix when the dimensionality is large is that of Stein-type shrinkage estimation. A convex combination of the sample covariance matrix and a well-conditioned target matrix is used to estimate the covariance matrix. Recent work in the literature has shown that an optimal combination exists under mean-squared loss, however it must be estimated from the data. we introduce a new set of estimators for the optimal convex combination for three commonly used target matrices. A simulation study shows an improvement over those in the literature in cases of extreme high-dimensionality of the data. A data analysis shows the estimators are effective in a discriminant and classification analysis.

Technology University of Shahrood
Faculty Of Mathematical Sciences

Dissertation Submitted in Partial
Fulfillment of The Requirements For The
Degree of Master of Science in
Statistic

Shrinkage estimator for the highdimensional multivariate normal covarianc matrix

Supervisor
Dr. M. Arashi

by
Elham Ashoori

SEPTEMBER 2014