

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ



دانشکده ریاضی

گروه ریاضی کاربردی

رشته آمار

برآورد انقباضی در مدل‌های چندمتغیره

دانشجو: سیدعلی اصغر تجدد

استاد راهنما

دکتر محمد آرشی

استاد مشاور

دکتر حسین باغیشنی

پایان نامه ارشد جهت اخذ درجه کارشناسی ارشد

مهر ۱۳۹۱

ساها خواندم و خواندم کتاب دیگران

تا که امروز ز خود مطلبی آرام به زبان

همه حاصل عمرم همین است، همین

پیشکش به مهاد و پدر بادل و جان

پیشکش به یاری، همسرم و صبر و دعاش

تک تک نایه داد به من شوق و توان

تقدیم به استاد کراتقدر و صبور

که آموخت به من آنچه یا موخت به جان

قدردانی

پاس خدا را بر آنچه از شکرش به ما الهام فرموده و بر آن دهنی دانشی که به پروردگارش به ما نکرده است. الهی ادای شکر تو را بهیچ زبان یار نیست و دیهای فضل تو را بهیچ کلام نیست و سر حقیقت تو بهیچ کس عیان نیست. هدایت کن بر ما راهی که بهتر از آن نیست.

الکون که به یاری خداوند متعال این دوره را به پایان رسانده ام، به رسم ادب و احترام بوسه بردستان زده و بر خود واجب می دانم از زحمات استاد راهنمای عزیز و بزرگوارم جناب آقای دکتر محمد آرشی که علاوه بر تحقیق و پژوهش با نظرات ارزنده و مدبرانه خود به سرفکر کردن، بهتر زندگی کردن و بهتر بودن را به من آموختند، شکر کنم. همچنین از استاد مشاور مهربان و دلسوزم جناب آقای دکتر حسین باغینی که با صبر و بردباری من را راهنمایی و این پایان نامه را تصحیح نمودند، صمیمانه شکر کنم. و نیز از اساتید محترم آقایان دکتر داوود شایونی و دکتر مهدی روزبه که با حضور و لگزشان تصحیح و داوری این پایان نامه را بر عهده گرفته اند سپاسگزارم.

بر خود واجب می دانم که از زحمات اساتید کرامت و پرستش زحماتش دانشجو و شاگرد صمیمانه شکر کنم.

و در نهایت از کلیه دوستان عزیزم که در طی دوران تحصیل فضایی آرام و دوستانه را با ایشان تجربه کردم سپاسگزارم.

تعهد نامه

اینجانب سیدعلی اصغر تجدد دانشجوی دوره کارشناسی ارشد رشته آمارریاضی دانشکده ریاضی دانشگاه صنعتی شاهرود نویسنده پایان نامه برآورد انقباضی در مدل های چندمتغیره تحت راهنمایی دکتر محمد آرشی متعهد می شوم

- تحقیقات در این پایان نامه توسط اینجانب انجام شده است و از صحت و اصالت برخوردار است.
- در استفاده از نتایج پژوهشهای محققان دیگر به مرجع مورد استفاده استناد شده است.
- مطالب مندرج در پایان نامه تاکنون توسط خود یا فرد دیگری برای دریافت هیچ نوع مدرک یا امتیازی در هیچ جا ارائه نشده است.
- کلیه حقوق معنوی این اثر متعلق به دانشگاه صنعتی شاهرود می باشد و مقالات مستخرج با نام «دانشگاه صنعتی شاهرود» و یا «Shahrood University of Technology» به چاپ خواهد رسید.
- حقوق معنوی تمام افرادی که در به دست آمدن نتایج اصلی پایان نامه تأثیرگذار بوده اند در مقالات مستخرج از پایان نامه رعایت می گردد.
- در کلیه مراحل انجام این پایان نامه، در مواردی که از موجود زنده (یا بافتهای آنها) استفاده شده است ضوابط و اصول اخلاقی رعایت شده است.
- در کلیه مراحل انجام این پایان نامه، در مواردی که به حوزه اطلاعات شخصی افراد دسترسی یافته یا استفاده شده است اصل رازداری، ضوابط و اصول اخلاق انسانی رعایت شده است.

تاریخ ۱۳۹۱/۰۷/۰۵

امضای دانشجو

* متن این صفحه نیز باید در ابتدای نسخه های تکثیر شده پایان نامه وجود داشته باشد.

چکیده

پس از آنکه در سال ۱۹۵۶، چارلز استاین نشان داد در توزیع نرمال، هنگامی که بعد فضای پارامتر بزرگتر یا مساوی ۳ است، میانگین نمونه برآوردگر غیرمجاز برای میانگین توزیع نرمال است، تاکنون تلاش‌های زیادی در جهت بهبود برآوردگر میانگین در حالتی که بعد فضای پارامتر بزرگتر یا مساوی ۳ است انجام شده است. بهبود برآوردگرها تنها محدود به توزیع نرمال نشده است بلکه به عنوان تعمیمی از آن، در خانواده توزیع‌های بیضی‌گون که شامل توزیع‌های زیادی از جمله نرمال، t ، اسلش و غیره می‌باشد، نیز مورد بررسی قرار گرفته‌اند. در این پایان‌نامه برآوردگرهای بهبودیافته، که در بازه‌ای خاص از فضای پارامتر بر برآوردگرهای کمترین توان‌های دوم برتری دارند را در مدل‌های چندمتغیره بیضی‌گون ماتریسی، معرفی و رفتار آنها را مورد بررسی قرار می‌دهیم. در این راستا ابتدا برای برآورد پارامتر مکان در توزیع نرمال ماتریسی، برآوردگرهای درست‌نمایی ماکزیمم محدودنشده، درست‌نمایی ماکزیمم محدودشده، انقباضی نوع استاین و انقباضی نوع استاین مثبت را معرفی، و با محاسبه دقیق مخاطره به بررسی رفتار برآوردگرهای مذکور می‌پردازیم. در نهایت با بکارگیری تکنیک خاصی نتایج را برای برآورد پارامتر ماتریسی مکان به خانواده توزیع‌های بیضی‌گون ماتریسی تعمیم می‌دهیم.

واژه‌های کلیدی: مدل رگرسیون خطی چندمتغیره، تابع مخاطره، توزیع بیضی‌گون ماتریسی، برآوردگر انقباضی، برآوردگر نوع استاین

فهرست مقالات مستخرج از متن پایان نامه

1. Tajadod, S. A. A. and Arashi, M. (2012), "Shrinkage estimation in growth curve models" 11th Iran. Statist. Conf. 703-709.
2. Arashi, M., Kibria, B. M. G. and Tajadod, S. A. A. (2012), "On shrinkage estimators in matrix variate elliptical models" Revised.

نمادها

χ^2_q : توزیع کی-دو مرکزی با درجه آزادی q

$\chi^2_{q,\delta}$: توزیع کی-دو غیرمرکزی با درجه آزادی q و پارامتر غیرمرکزی δ

$N_n(\mu, \Sigma)$: توزیع نرمال n متغیره با میانگین μ و مقیاس Σ

$MN_{p \times q}(\mu, \Sigma \otimes \psi)$: توزیع نرمال ماتریسی در اندازه $p \times q$ با میانگین μ و مقیاس $\Sigma \otimes \psi$

$EC_n(\mu, \Sigma, \psi)$: توزیع بیضی‌گون با میانگین μ ، مقیاس Σ و تابع مولد مشخصه ψ

$EC_n(\mu, \Sigma, g)$: توزیع بیضی‌گون با میانگین μ ، مقیاس Σ و تابع مولد چگالی g

$LS_{p \times q}(\phi)$: توزیع چپ-کروی با تابع مشخصه $\phi(T'T)$

$RS_{p \times q}(\phi)$: توزیع راست-کروی با تابع مشخصه $\phi(T'T)$

$SS_{p \times q}(\phi)$: توزیع کروی ماتریسی با تابع مشخصه $\phi(T'T)$

$MEC_{p \times q}(\mu, \Sigma \otimes \psi, f)$: توزیع بیضی‌گون ماتریسی با پارامترهای μ ، Σ و ψ و تابع مولد چگالی f

$W_n(\psi, p, \Delta)$: توزیع ویشارت غیرمرکزی با مقیاس ψ ، p درجه آزادی و پارامتر غیرمرکزی Δ

$W^{-1}(\psi^{-1}, p)$: توزیع ویشارت معکوس با پارامتر مقیاس ψ و p درجه آزادی

$$\Gamma_n(.)\text{: تابع گامای چندمتغیره}$$

$$B(a,b)\text{: تابع بتای کامل}$$

$$\mathcal{R}^{p\times q}\text{: فضای }p\times q\text{ بعدی حقیقی}$$

$$F_p(a_1,...,a_p;b_1,...,b_q;z)\text{: تابع فوق هندسی تعمیم یافته}$$

$$X=Y^d\text{: هم توزیع بودن دو متغیر }X\text{ و }Y$$

$$\exp(i tr(X))=e^{i\times trace(X)}$$

$$\tilde{\theta}\text{: برآوردگر محدودنشده }\theta$$

$$\hat{\theta}\text{: برآوردگر محدودشده }\theta$$

$$\hat{\theta}_S\text{: برآوردگر انقباضی}$$

$$\hat{\theta}^s\text{: برآوردگر انقباضی نوع استاین}$$

$$\hat{\theta}^{s+}\text{: برآوردگر انقباضی نوع استاین مثبت}$$

$$R(X,\theta;W)\text{: تابع مخاطره وزنی برآوردگر }X\text{ در خصوص برآورد پارامتر }\theta\text{ با وزن }W$$

$$\delta_1 \succ \delta_2\text{: برتری برآوردگر }\delta_1\text{ بر }\delta_2$$

$$I(A\leq a)\text{: تابع نشانگر مجموعه }A$$

$$X\sim_{p\times q}(\mu,\Sigma\otimes\psi)\text{: ماتریس تصادفی }X\text{ در اندازه }p\times q\text{ با میانگین }\mu\text{ و ماتریس کوواریانس}$$

$$\Sigma\otimes\psi$$

علائم اختصاری

ECD: توزیع با منحنی تراز بیضی شکل (Elliptically contoured distribution)

BLUE: بهترین برآوردگر خطی ناریب (Best linear unbiased estimator)

MLE: برآوردگر درست‌نمایی ماکزیمم (maximum likelihood estimator)

GLS: کمترین توان‌های دوم تعمیم‌یافته (Generalized least square)

RGLS: کمترین توان‌های دوم تعمیم‌یافته محدودشده (Restricted generalized least square)

UE: برآوردگر محدودنشده (Unrestricted Estimator)

RE: برآوردگر محدودشده (Restricted estimator)

SSE: برآوردگر انقباضی نوع استاین (Stein-type shrinkage estimator)

PRSSE: برآوردگر انقباضی نوع استاین مثبت (Positive-rule Stein-type shrinkage estimator)

GCM: مدل منحنی رشد (Growth curve model)

فهرست مندرجات

۱	مقدمه و دورنما	۱
۲	۱-۱ مقدمه و تاریخچه.....	۲
۸	۲-۱ تعاریف و نتایج.....	۸
۸	۱-۲-۱ توابع ریاضی.....	۸
۱۰	۲-۲-۱ توزیع‌های آماری.....	۱۰
۱۵	۳-۲-۱ لم‌ها و قضایای اساسی.....	۱۵
۲۶	۲ مدل رگرسیون چندمتغیره محدودشده	۲۶
۲۷	۱-۲ مقدمه.....	۲۷
۲۸	۲-۲ مدل رگرسیون خطی چندگانه.....	۲۸
۳۸	۳-۲ مدل رگرسیون خطی چندمتغیره.....	۳۸
۳۹	۱-۳-۲ برآورد پارامترها وقتی Σ معلوم است.....	۳۹
۴۳	۲-۳-۲ برآورد پارامترها در حالت $\Sigma = \sigma^2 V$	۴۳

۳ برآورد انقباضی ۴۵

۱-۳	مقدمه.....	۴۶
۲-۳	برآوردگر انقباضی.....	۴۶
۳-۳	برآورد انقباضی در توزیع نرمال ماتریسی.....	۵۱
۱-۳-۳	روش ساخت برآوردگر انقباضی.....	۵۱
۲-۳-۳	لمها و قضایای اساسی.....	۵۵
۳-۳-۳	محاسبه توابع مخاطره برآوردگرها.....	۶۶
۴-۳-۳	مقایسه برآوردگرها.....	۷۱
۴-۳	برآورد انقباضی در مدل منحنی رشد.....	۷۵
۱-۴-۳	مدل منحنی رشد.....	۷۵
۲-۴-۳	برآوردگر انقباضی.....	۷۸
۳-۴-۳	محاسبه مخاطره.....	۷۹
۴-۴-۳	مقایسه برآوردگرها.....	۸۱

۴ برآورد انقباضی در مدل های بیضی گون ماتریسی ۸۵

۱-۴	مقدمه.....	۸۶
۲-۴	مختصری درباره توزیع های بیضی گون.....	۸۷

- ۱-۲-۴ توزیع‌های گروهی ماتریسی ۸۷
- ۲-۲-۴ ارتباط بین کلاس‌های توزیع گروهی ماتریسی ۹۱
- ۳-۲-۴ توزیع‌های بیض‌گون ماتریسی ۹۲
- ۴-۲-۴ مثال‌هایی از توزیع بیضی‌گون ماتریسی ۹۸
- ۳-۴ برآورد انقباضی در توزیع بیضی‌گون ماتریسی ۱۰۰
- ۱-۳-۴ محاسبه توابع مخاطره ۱۰۱

پیوست ۱۰۵

- ۱-A جبر ماتریس‌ها ۱۰۶
- ۲-A پیشنهادات برای پژوهش‌های آینده ۱۲۱

مراجع ۱۲۲

چکیده انگلیسی

فصل اول

مقدمه و دورنما

۱-۱ مقدمه و تاریخچه

در تحلیل‌های آماری، استفاده از اطلاعات پیشین بر روی تعدادی یا تمام پارامترهای مدل، معمولاً منجر به استفاده از روش‌های پیشرفته و تعمیم‌یافته‌ای می‌شود که با توجه به قطعی یا غیرقطعی بودن این اطلاعات، به کار برده می‌شوند. در حالت قطعی بودن اطلاعات پیشین، وارد کردن آن در مدل به صورت قیدی^۱ بر روی فضای پارامتر، باعث به وجود آمدن مدل محدودشده^۲ می‌شود. تحلیل چنین قیدهایی منتهی به روش‌های آماری پیشرفته، مانند مقایسه برآوردگرهای پارامترها در مدل محدودشده، در مقابل مدل محدودنشده^۳ که تمام فضای پارامتر را در بر می‌گیرد، می‌شود. زمانی که در مدل‌های آماری با اطلاعات پیشین غیرقطعی، به صورت یک قید بر روی پارامترها مواجه می‌شویم آن را به عنوان یک پارامتر مزاحم^۴ در تحلیل مدل مورد نظر می‌گیریم. یک روش مناسب برای از بین بردن عدم قطعیت اطلاعات پیشین در مدل، استفاده از رهیافت فیشر^۵ است. این روش با استفاده از برآوردگر آزمون مقدماتی^۶ (PTE)، پارامتر مزاحم را از مدل حذف می‌کند. به این ترتیب که بر اساس اعتبار و درستی اطلاعات پیشین غیرقطعی، یکی از دو مدل محدودشده یا نشده انتخاب می‌شوند. بنکرافت^۷ (۱۹۴۴، ۱۹۶۴، ۱۹۶۵) برای اولین بار ایده آزمون مقدماتی را برای برآورد واریانس در جدول آنالیز واریانس مورد استفاده قرار داد. این ایده از روش پیشنهادی سندگر^۸ (۱۹۳۸)، که به منظور آزمون برابری میانگین‌ها با فرض تساوی واریانس-ها (نامعلوم) طراحی شده بود، الهام گرفته شده است. در روش سندگر اگر آزمون برابری واریانس‌ها رد

¹ Constraint

² Restricted model

³ Unrestricted model

⁴ Nuisance parameter

⁵ Fisher approach

⁶ Preliminary test estimator

⁷ Bancroft

⁸ Snedecor

نشود، آنگاه در آزمون استیودنت، از واریانس آمیخته و در غیر این صورت از دسته‌بندی بهرنز- فیشر^۹ استفاده می‌شود. با تحقیقات و بررسی‌های بیشتر در این زمینه، آشکار شد که روش آزمون مقدماتی به عواملی همچون سطح معنی‌داری آزمون و مقادیر برآوردگر انتخابی، وابسته است. در این راستا هان^{۱۰} و بنکرافت (۱۹۶۸) ثابت کردند با استفاده از یک روش ماکزیمم-مینیمم بر روی سطح معنی‌داری آزمون، می‌توان به اندازه بهینه برآوردگر آزمون مقدماتی میانگین، در مسائل دونمونه‌ای رسید. ماستیلر^{۱۱} (۱۹۴۸) به بررسی این مدل‌ها از دیدگاه بیزی پرداخت. کیتاگاوا^{۱۲} (۱۹۶۳) ایده آزمون مقدماتی را بر روی پیشامدهای متوالی، با به‌دست آوردن توزیع و گشتاورهای برآوردگر آزمون مقدماتی در مسائل دونمونه‌ای، مورد بررسی قرار داد. بوزیویچ^{۱۳}، بنکرافت و هارتلی^{۱۴} (۱۹۵۶) ویژگی‌های اندازه و توان برآوردگر مقدماتی را مورد مطالعه قرار دادند. بنت^{۱۵} (۱۹۵۲، ۱۹۵۶)، هانس برگر^{۱۶} (۱۹۵۵)، آسانو و ساتو^{۱۷} (۱۹۶۲) کاربرد برآوردگر آزمون مقدماتی را در روش‌های مختلف مورد توجه قرار دادند. استاین^{۱۸} (۱۹۵۵، ۱۹۵۶، ۱۹۸۱) و جیمز^{۱۹} و استاین (۱۹۶۱) به یک تضاد آماری رسیدند که حل آن یک قرن به طول انجامید. نتیجه این بررسی‌ها نشان داد که برآوردگر میانگین، تحت تابع زیان مربع خطا، در توزیع نرمال چندمتغیره p بعدی ($p \geq 3$) غیر مجاز^{۲۰} است. این موضوع انگیزه‌ای برای بهبود برآوردگرها در مسایل برآوردیابی شد. بعد از آن، جیمز و استاین (۱۹۶۱) برآوردگری ارائه دادند که بر میانگین نمونه برتری داشت. این برآوردگر به برآوردگر جیمز-استاین معروف شد. این یافته، تحول بزرگی در شاخه‌های مختلف آمار ایجاد کرد.

⁹ Behrens – Fisher

¹⁰ Han

¹¹ Mosteller

¹² Kitagawa

¹³ Bozovich

¹⁴ Hartley

¹⁵ Bennett

¹⁶ Huntssberger

¹⁷ Asano and Sato

¹⁸ Charles Stein

¹⁹ James

²⁰ Inadmissible

تاکنون، مطالعات زیادی به منظور بهبود برآوردگرهای پارامترها در مدل‌های آماری با استفاده از یافته‌های استاین (۱۹۵۶) و جیمز و استاین (۱۹۶۱) انجام شده‌اند. ایده اصلی، استفاده مناسب از اطلاعات پیشین نمونه‌ای و غیرنمونه‌ای در توسعه هر چه بهتر برآوردگرهای نوع جیمز-استاین می‌باشد. در این خصوص می‌توان به جایلز^{۲۱} (۱۹۹۱)، جاج و باک^{۲۲} (۱۹۷۸) و طباطبایی (۱۹۹۵) در حالاتی که توزیع خطاها نرمال یا t است، مراجعه کرد.

برآوردگر جیمز-استاین، برآوردگری غیرخطی است که بر برآوردگر خطی میانگین نمونه برتری دارد. لذا تمامی خواص خوب برآوردگر میانگین نمونه‌ای که بر اساس نظریه کمترین توان‌های دوم و درست‌نمایی ماکزیمم به دست می‌آیند، مانند ناریبی و پایایی، تحت تاثیر برآوردگر جیمز-استاین قرار گرفتند. بر این اساس به افتخار پرفسور چارلز استاین، برآوردگری را که به طور یکنواخت بر برآوردگر استاندارد (برآوردگرهای کمترین توان‌های دوم، برآوردگرهای درست‌نمایی ماکزیمم و غیره) برتری دارد، را برآوردگر نوع استاین^{۲۳} (SE) می‌نامند.

رفتار برآوردگرهای نوع استاین وابسته به عامل انقباض^{۲۴} است. برآوردگر نوع استاین، با استفاده از حاصل-ضرب دو عبارت که یکی تفاضل برآوردگر محدودشده از برآوردگر محدودنشده و دیگری حاصل‌ضرب عامل انقباض در آماره آزمون مناسب قید اعمال شده بر روی فضای پارامتر است، برآوردگر محدودنشده را بهبود می‌بخشد. به طور کلی، آماره آزمون به کار رفته در برآوردگر نوع استاین، فاصله بین برآوردگر محدودشده و برآوردگر محدودنشده را اندازه می‌گیرد که معمولاً این آماره دارای توزیع کی-دو غیرمرکزی^{۲۵} یا فیشر

²¹ Giles

²² Judge and Bock

²³ Stein-type estimator

²⁴ Shrinkage factor

²⁵ Non-central chi-square

غیرمرکزی^{۲۶} می‌باشد. مقدار مخاطره^{۲۷} برآوردگر نوع استاین به پارامتر غیرمرکزی توزیع آماره آزمون بستگی دارد.

با توجه به آن که احتمال منفی شدن برآوردگر نوع استاین مثبت است، در تعمیمی از آن، تنها قسمت مثبت آن را مورد استفاده قرار می‌دهند که به برآوردگر نوع استاین مثبت^{۲۸} (PRSE) معروف است.

تاکنون در مدل‌های مختلف آماری، در حالاتی که ماتریس کوواریانس جامعه مورد بررسی به صورت $\sigma^2 I_n$ است، که در آن I_n ماتریس واحد با بعد n است، برآوردهای PTE، SE و PRSE مورد بررسی قرار گرفته‌اند. در این زمینه می‌توان به کلارک^{۲۹} و همکاران (۱۹۸۷)، علی و صالح^{۳۰} (۱۹۹۱)، سینگ^{۳۱} (۱۹۹۱)، جایلز (۱۹۹۱، ۱۹۹۲)، کبیریا^{۳۲} (۱۹۹۶)، خان^{۳۳} و صالح (۱۹۹۷)، طباطبایی و همکاران (۲۰۰۴a، ۲۰۰۴b) اشاره کرد. همچنین صالح (۲۰۰۶)، برای توزیع پاسخ نرمال، مدل‌های رگرسیون خطی چندگانه، ANOVA و موازی را تحت قیدهای قطعی و غیرقطعی مورد بررسی قرار داد. آرشی (۲۰۰۸) با تعمیم ماتریس کوواریانس به حالت‌های کلی‌تر ماتریس Σ (معلوم) و $\sigma^2 V$ ، که در آن σ^2 مجهول و ماتریس V معلوم است، با در نظر گرفتن توزیع بیضی‌گون برای بردار خطا، به بررسی و مقایسه برآوردهای PTE، SE و PRSE در مدل رگرسیون خطی چندگانه پرداخت.

اگرچه در رویارویی با مدل‌بندی داده‌های کاربردی و واقعی بیشتر با مدل‌های چندمتغیره سروکار داریم، با این حال برآورد انقباضی تحت مدل‌های چندمتغیره کمتر مورد توجه محققان این رشته قرار گرفته است. از این رو و با توجه به اهمیت موضوع، هدف از تالیف این پایان‌نامه، به‌دست آوردن

Non-central Fisher²⁶Risk²⁷Positive-rule Stein-type estimator²⁸Clarke²⁹Ali and Saleh³⁰Singh³¹Kibria³²Khan³³

برآوردگرهای بهبودیافته، به خصوص انواع برآوردگر انقباضی برای برآورد پارامترهای مکانی در مدل‌های چندمتغیره، که غالباً از جنس ماتریس می‌باشند، است. برای این منظور، در این پایان‌نامه فرض کرده‌ایم که مدل از خانواده بزرگ توزیع‌های بیضی‌گون ماتریسی تبعیت می‌کند. همچنین بهبودی برآوردگرها با معیار اندازه مخاطره‌شان تعیین می‌گردد. به عبارتی برآوردگرهایی بهتر هستند که دارای مخاطره کمتر باشند.

محاسبه دقیق مخاطره برآوردگرها با استفاده از لم‌ها و قضایای کاملاً جدید، مقایسه رفتار برآوردگرها از جنبه نظری و نیز با استفاده از رسم نمودار تابع مخاطره و بروز بودن منابع اصلی از ویژگی‌های این پایان‌نامه محسوب می‌شود. همچنین بررسی رفتار انواع برآوردگرهای پارامترهای مدل منحنی رشد^{۳۴} (GCM)، که یک مدل رگرسیونی چندمتغیره و کاربردی در زیست‌شناسی و تحقیقات پزشکی است، از نتایج فصل سوم این مجموعه به شمار می‌رود.

این مجموعه شامل ۴ فصل و ۱ ضمیمه می‌باشد. مطالب هر فصل به طور مختصر عبارتند از:

- در فصل اول، علاوه بر مقدمات، مروری بر تاریخچه موضوع مورد بررسی و ضرورت و اهداف پایان‌نامه که شرح آن گذشت، برخی توابع ریاضی، توزیع‌های آماری و لم‌های اساسی که در این پایان‌نامه استفاده شده‌اند نیز آمده است.
- در فصل دوم، در مدل رگرسیون خطی چندگانه، بدون اینکه فرض بخصوصی برای توزیع بردار خطای مدل در نظر بگیریم (آزاد توزیع)، برآوردگرهای کمترین توان‌های دوم تعمیم‌یافته^{۳۵} (GLS) و محدودشده^{۳۶} (RGLS) را مطرح کرده‌ایم و در پایان با معرفی مدل رگرسیون چند متغیره، برآوردگرهای مذکور را برای این مدل نیز به‌دست آورده‌ایم. در واقع در این فصل

³⁴ Growth curve model

³⁵ Generalized least square

³⁶ Restricted generalized least square

برآوردگرهای لازم در ساخت برآوردگر انقباضی که در فصل سوم به آن می‌پردازیم، آورده شده است. تکنیک دستیابی به این برآوردگرها (GLS و RGLS) در مدل‌های چندمتغیره موضوع مورد توجه نویسندگان در این فصل است.

- در فصل سوم، ابتدا برآوردگرهای محدودشده، انقباضی نوع استاین و نوع استاین مثبت برای ماتریس میانگین در مدل نرمال ماتریسی مطرح شده است. در ادامه این برآوردگرها برای ماتریس پارامتر مدل رگرسیونی منحنی رشد در حالتی که خطای مدل دارای توزیع نرمال ماتریسی است، مورد بررسی قرار گرفته‌اند.
- در فصل چهارم، موارد بدست آمده در فصل سوم، در حالتی که مدل بیضی‌گون ماتریسی است، بررسی شده‌اند.
- در پیوست، برخی تعاریف و قضایا از جبر ماتریس‌ها که در اثبات‌ها و قضایای این پایان‌نامه مورد استفاده قرار گرفته‌اند، آورده شده‌اند. همچنین پیشنهادها برای ادامه کار در آینده، مطرح شده است.

قسمت‌هایی که با * مشخص شده‌اند نشان دهنده قضیه یا لم موردنظر در مرجع بوده که اثبات آن‌ها توسط نویسندگان این پایان‌نامه انجام شده، و ** بدین معنی است که قضیه و یا لم توسط نویسندگان این پایان‌نامه، بیان و اثبات شده است.

۲-۱ تعاریف و نتایج

در این بخش ابتدا برخی توابع ریاضی و توزیع‌های آماری که در این پایان‌نامه از آن‌ها استفاده شده است را معرفی کرده و در پایان لم‌ها و قضایایی که در اثبات قضایای فصل‌های بعدی مورد استفاده قرار گرفته‌اند را بیان می‌کنیم. البته توجه کنید بعضی تعاریف، اصطلاحات و قضایایی که در این بخش به آن اشاره‌ای نشده ولی در متن پایان‌نامه وجود دارد را می‌توانید در پیوست (جبر ماتریس‌ها) ببینید.

۱-۲-۱ توابع ریاضی

تعریف ۱-۱ تابع گامای چندمتغیره

به ازای هر مقدار حقیقی t بزرگتر از $(n-1)/2$ ، تابع گامای چندمتغیره را با $\Gamma_n(t)$ نشان می‌دهیم و عبارتست از

$$\Gamma_n(t) = \pi^{\frac{n(n-1)}{4}} \prod_{i=1}^n \Gamma\left[t - \frac{1}{2}(i-1)\right],$$

که در آن $\Gamma(a) = \int_0^\infty y^{a-1} e^{-y} dy$ تابع گاما می‌باشد.

تعریف ۲-۱ تابع بتا

به ازای هر دو مقدار حقیقی a و b ، تابع بتا را با $B(a, b)$ نشان می‌دهیم و عبارتست از

$$\begin{aligned} B(a, b) &= \int_0^1 x^{a-1} (1-x)^{b-1} dx \\ &= \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}. \end{aligned}$$

تعریف ۳-۱ تابع فوق هندسی تعمیم یافته^{۳۷}

به ازای مقادیر حقیقی $a_1, a_2, \dots, a_p$ و $b_1, b_2, \dots, b_p$ تابع فوق هندسی تعمیم یافته به صورت زیر است

$${}_qF_p(a_1, \dots, a_p; b_1, \dots, b_p; z) = \sum_{k=0}^{\infty} \frac{(a_1)_k \dots (a_p)_k}{(b_1)_k \dots (b_p)_k} \times \frac{z^k}{k!}$$

که در آن $(a)_k = a(a+1) \dots (a+k-1)$. برای آگاهی بیشتر در خصوص این تابع، میرهد (۱۹۸۲) را ببینید.

تعریف ۴-۱ تابع زیان توان دوم موزون

اگر ماتریس B^* در اندازه $p \times q$ برآوردگری برای ماتریس پارامتر مکان B باشد، آن گاه زیان وزنی استفاده از B^* در خصوص برآورد B را با $L(B^*, B; W)$ نشان می دهیم و به صورت زیر تعریف می شود:

$$L(B^*, B; W) = (B^* - B)'W(B^* - B),$$

که در آن W ماتریس وزن معلوم و نامنفرد با بعد p است.

تعریف ۵-۱ تابع مخاطره توان دوم موزون^{۳۸}

با استفاده از تعریف ۴-۱، تابع مخاطره توان دوم موزون برآوردگر B^* را با $R(B^*, B; W)$ نشان می دهیم و عبارت است از:

$$R(B^*, B; W) = E[L(B^*, B; W)].$$

³⁷ Generalized hypergeometric function

³⁸ Weighted quadratic risk function

تعریف ۱-۶ تابع مشخصه^{۳۹}

تابع مشخصه بردار p مولفه‌ای X به صورت زیر تعریف می‌شود:

$$\varphi_X(t) = E[\exp(it'X)], \quad t \in \mathbb{R}^{p \times 1}$$

تابع مشخصه ماتریس تصادفی X در اندازه $p \times q$ به صورت زیر تعریف می‌شود:

$$\varphi_X(t) = E[\exp(itr(T'X))], \quad t \in \mathbb{R}^{p \times q}$$

با توجه به ویژگی (۴) در قضیه A-۲۵، در پیوست، تابع مشخصه ماتریس X را به صورت زیر نیز می‌توان بیان کرد:

$$\varphi_X(t) = E[\exp(ivec'TvecX)].$$

در رابطه با عملگر $vec(.)$ و ویژگی‌های آن پیوست را ببینید.

۲-۲-۱ توزیع‌های آماری

تعریف ۱-۷ توزیع نرمال چندمتغیره

بردار p مولفه‌ای X دارای توزیع نرمال چندمتغره p بعدی است و با $X \sim N_p(\mu, \Sigma)$ نمایش می‌دهیم،

اگر و فقط اگر به ازای هر بردار p مولفه‌ای ثابت a ، $a'X$ دارای توزیع نرمال یک متغیره باشد.

اگر Σ یک ماتریس معین مثبت باشد، آنگاه X دارای تابع چگالی احتمال به صورت زیر است:

$$f(x) = \frac{|\Sigma|^{-1/2}}{(2\pi)^{p/2}} \exp\left[-\frac{(x - \mu)' \Sigma^{-1}(x - \mu)}{2}\right],$$

که در آن $|\Sigma|^{1/2}$ 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 در نتیجه $X \sim N(\mu, \sigma^2)$

تعریف ۸-۱ توزیع کی-دو مرکزی^{۴۰}

اگر $X \sim N_p(0, I)$ باشد، آنگاه $Y = X'X$ دارای توزیع کی-دو مرکزی با n درجه آزادی است و با

$Y \sim \chi_p^2$ نمایش می دهند. تابع چگالی احتمال Y به صورت زیر تعریف می شود:

$$f(y) = \frac{e^{-y/2} y^{n/2-1}}{2^{n/2} \Gamma(n/2)}, \quad y > 0, \quad n \in \mathbb{N}.$$

تعریف ۹-۱ توزیع کی-دو غیرمرکزی

اگر $X \sim N_p(\mu, I)$ باشد، آنگاه $Y = X'X$ دارای توزیع کی-دو غیرمرکزی با n درجه آزادی و پارامتر

غیرمرکزی δ است و با $Y \sim \chi_{n,\delta}^2$ نشان می دهیم. تابع چگالی احتمال Y به صورت زیر تعریف می شود:

$$f(y) = e^{-\delta/2} \cdot F_1(n/2; \delta y/4) \times \frac{e^{-y/2} y^{n/2-1}}{2^{n/2} \Gamma(n/2)}.$$

لازم به ذکر است که تابع چگالی احتمال Y را می توان به صورت زیر نیز نوشت

$$f(y) = \sum_{k=0}^{\infty} P(K = k) g_{n+2k}(y),$$

که در آن K متغیر تصادفی دارای توزیع پواسن با میانگین $\delta/2$ و $g_{n+2k}(y)$ تابع چگالی احتمال یک

متغیر تصادفی کی-دو مرکزی با $n + 2k$ درجه آزادی در نقطه y است. (میرهد، ۱۹۸۲)

تعریف ۱۰-۱ توزیع t چندمتغیره^{۴۱}

بردار p مولفه‌ای X دارای توزیع t چندمتغیره p بعدی با پارامترهای μ ، Σ و درجه آزادی ν است و با $X \sim t_p(\mu, \Sigma, \nu)$ نشان می‌دهیم، اگر تابع چگالی احتمال X به صورت زیر باشد:

$$f(x) = \frac{\Gamma\left(\frac{\nu+p}{2}\right) |\Sigma|^{-1/2}}{(\nu p)^{p/2} \Gamma\left(\frac{\nu}{2}\right)} \left(1 + \frac{(x - \mu)' \Sigma^{-1} (x - \mu)}{\nu}\right)^{-\frac{(\nu+p)}{2}},$$

$$x, \mu \in \mathbb{R}^{p \times 1}, \Sigma \in \mathbb{R}^{p \times p}, \Sigma > 0, \nu \in \mathbb{N}.$$

برای آگاهی بیشتر در مورد این توزیع، ناداراجا و کاتز^{۴۲} (۲۰۰۵) را ببینید.

تعریف ۱۱-۱ توزیع با منحنی تراز بیضی شکل^{۴۳} (توزیع بیضی‌گون)

بردار n مولفه‌ای X دارای توزیع بیضی‌گون با پارامترهای μ ، Σ و تابع مولد مشخصه^{۴۴} ψ است و با $X \sim EC_n(\mu, \Sigma, \psi)$ نشان می‌دهیم، اگر تابع مشخصه آن به صورت زیر باشد

$$\varphi_X(t) = \exp(it' \mu) \psi\left(\frac{t' \Sigma t}{2}\right),$$

که در آن $\psi(\cdot)$ متعلق به کلاس توابعی به صورت $\psi(t): [0, \infty) \rightarrow \mathbb{R}$ است، به طوری که $\psi\left(\sum_{i=1}^n t_i^2\right)$ یک تابع مشخصه n -بعدی می‌باشد.

Multivariate Student distribution⁴¹Nadarajah and Kotz⁴²Elliptically contoured distribution⁴³Characteristic generating function⁴⁴

توزیع‌های بیضی‌گون، توزیع‌هایی هستند که منحنی‌های تراز آنها بیضی شکل می‌باشند. از این که $X \sim EC_n(\mu, \Sigma, \psi)$ نمی‌توان نتیجه گرفت که متغیر تصادفی X دارای تابع چگالی احتمال است، ولی در صورتی که تابع چگالی وجود داشته باشد، به صورت زیر خواهد بود:

$$f_X(x) = c_n |\Sigma|^{-1/2} g \left[\frac{(x - \mu)' \Sigma^{-1} (x - \mu)}{2} \right],$$

که در آن $g(\cdot)$ تابع مولد چگالی^{۴۵} است و دارای شرط زیر می‌باشد

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

و همچنین c_n ضریب نرمال‌سازی می‌باشد که به صورت زیر تعیین می‌گردد:

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

در این صورت می‌نویسیم $X \sim EC_n(\mu, \Sigma, g)$. نکته حائز اهمیت آن است که اگر تابع چگالی وجود داشته باشد، تابع g ، تابع ψ را مشخص می‌کند و برعکس. برای آگاهی بیشتر در این زمینه، فنگ و همکاران^{۴۶} (۱۹۹۰) و میرهد^{۴۷} (۱۹۸۲) را ببینید.

با توجه به تعریف ۱-۱۱، $E(X) = \mu$ و $cov(X) = -2\psi'(0)$ وجود دارد در صورتی که $|\psi'(0)| < \infty$.

Density generator⁴⁵

Fang et al⁴⁶

Muirhead⁴⁷

تعریف ۱-۱۲ توزیع نرمال ماتریسی

ماتریس تصادفی X در اندازه $n \times p$ ، دارای توزیع نرمال ماتریسی با ماتریس میانگین μ و ماتریس کوواریانس $\Sigma \otimes \Omega$ است و با $X \sim MN_{n \times p}(\mu, \Sigma \otimes \Omega)$ نشان می‌دهیم، اگر تابع چگالی احتمال X به صورت زیر باشد:

$$f(X) = \frac{|\Omega|^{-\frac{n}{2}} |\Sigma|^{-\frac{p}{2}}}{(2\pi)^{\frac{np}{2}}} \exp \left\{ -\frac{1}{2} \text{tr}[\Omega^{-1}(X - \mu)' \Sigma^{-1}(X - \mu)] \right\},$$

که در آن Σ ماتریس کوواریانس ستونی و Ω ماتریس کوواریانس سطری می‌باشند.

تعریف ۱-۱۳ توزیع ویشارت

اگر $X \sim MN_{n \times p}(\mu, \psi \otimes I)$ و $\psi \geq 0$ باشد، آنگاه $U = XX'$ دارای توزیع ویشارت است. حال اگر $\mu = 0$ توزیع را ویشارت مرکزی با پارامتر مقیاس ψ و p درجه آزادی گوئیم و با $U \sim w_n(\psi, p)$ نمایش می‌دهیم و اگر $\mu \neq 0$ آنگاه U دارای توزیع ویشارت غیرمرکزی است و با $U \sim w_n(\psi, p, \Delta)$ نمایش می‌دهیم، که در آن $\Delta = \mu\mu'$ پارامتر غیرمرکزی است.

تعریف ۱-۱۴ توزیع ویشارت معکوس

با استفاده از تعریف ۱-۱، ماتریس مربعی U با بعد n ، دارای توزیع ویشارت معکوس با پارامتر مقیاس ψ^{-1} و m درجه آزادی است و با $U \sim W^{-1}(\psi^{-1}, m)$ نشان می‌دهیم، اگر تابع چگالی احتمال آن به صورت زیر باشد:

$$g(U) = \frac{|\psi|^{m/2} |U|^{-(m+n+1)/2} e^{-\text{trace}(\psi U^{-1})/2}}{2^{mn/2} \Gamma_n(m/2)}.$$

۳-۲-۱ لم‌ها و قضایای اساسی

قضیه ۱-۱ اگر Y بردار تصادفی، X ماتریس تصادفی، a و b بردارهای ثابت و A و B ماتریس‌های ثابت باشند، آنگاه با فرض قابل انجام بودن ضرب، داریم:

$$E(AY) = AE(Y)$$

$$E(a'Xb) = a'E(X)b$$

$$E(AXB) = AE(X)B$$

قضیه ۲-۱ فرض کنید ماتریس‌های $A_{k \times p}$ و $B_{m \times p}$ ثابت و $Y_{p \times 1}$ بردار تصادفی با ماتریس کوواریانس Σ باشد، همچنین $Z = AY$ و $W = BY$ ، آنگاه

$$\text{cov}(Z) = \text{cov}(AY) = A\Sigma A',$$

$$\text{cov}(Z, W) = \text{cov}(AY, BY) = A\Sigma B'.$$

قضیه ۳-۱ اگر A ماتریس متقارن ثابت، Y بردار تصادفی با بردار میانگین μ و ماتریس کوواریانس Σ باشند، آنگاه

$$E(Y'AY) = \text{tr}(A\Sigma) + \mu' A \mu.$$

قضیه ۴-۱ فرض کنید $A_{p \times p}$ ماتریس متقارن ثابت با رتبه r و $Y \sim N_p(\mu, \Sigma)$ باشد. آنگاه شرط لازم و کافی برای آن که $Y'AY \sim \chi^2(r, \delta)$ ، که در آن $\delta = \frac{\mu' A \mu}{2}$ ، باشد، آن است که $A\Sigma$ خودتوان باشد.

نتیجه ۱-۱ اگر $Y \sim N_p(\mu, \sigma^2 I)$ باشد، آنگاه $Y'AY \sim \chi^2\left(r, \frac{\mu' A \mu}{2\sigma^2}\right)$ است، اگر و فقط اگر A خودتوان با رتبه r باشد.

قضیه ۵-۱ اگر ماتریس B در اندازه $k \times p$ ثابت، ماتریس مربعی A با بعد p ماتریس متقارن ثابت و

$$B\Sigma A = 0 \text{ اگر و فقط اگر و مستقلند، از هم } Y'AY \text{ و } BY \text{ آنگاه } Y \sim N_p(\mu, \Sigma)$$

قضیه ۶-۱ اگر A و B ماتریس‌های متقارن ثابت و $Y \sim N_p(\mu, \Sigma)$ باشند، آنگاه $Y'AY$ و $Y'BY$ از

$$A\Sigma B = 0 \text{ اگر و فقط اگر و مستقلند، اگر و فقط اگر}$$

برای اثبات قضایای ۱-۱ تا ۶-۱ سیرل^{۴۸} (۱۹۷۱) را ببینید.

قضیه ۷-۱ (اندرسن^{۴۹}، ۲۰۰۳). اگر $X \sim N_p(\mu, \Sigma)$ باشد، در این صورت تابع مشخصه $\varphi_X(T)$ به

صورت زیر است:

$$\varphi_X(T) = \exp\left(i(T'\mu) - \frac{1}{2}(T'\Sigma T)\right).$$

قضیه ۸-۱ (کولو و وان‌رزن^{۵۰}، ۲۰۰۵). اگر $X \sim MN_{p \times n}(\mu, \Sigma \otimes \psi)$ در این صورت

(۱) تابع مشخصه $\varphi_X(T)$ به صورت زیر است:

$$\varphi_X(T) = \exp\left(i \operatorname{tr}(T'\mu) - \frac{1}{2} \operatorname{tr}(\Sigma T \psi T')\right).$$

$$E(XAX') = \operatorname{tr}(\psi A)\Sigma + \mu A \mu' \quad (۲)$$

اثبات: (۱) از تعریف ۶-۱ داریم:

$$\varphi_X(T) = E[\exp(i \operatorname{vec}' T \operatorname{vec} X)],$$

Searle⁴⁸

Anderson⁴⁹

Kollo and Von Rosen⁵⁰

که در آن $vecX \sim N_{pn}(vec\mu, \Sigma \otimes \psi)$ بنابراین با استفاده از قضیه ۷-۱ خواهیم داشت

$$\varphi_X(T) = \exp\left(i(vec'Tvec\mu) - \frac{1}{2}vec'T(\Sigma \otimes \psi)vecT\right).$$

و در نهایت از (A-۲۱) نتیجه حاصل می‌شود.

(۲) فرض کنید $Y \sim N_{p \times n}(0, \Sigma \otimes \psi)$ بنابراین X و $Y + \mu$ هم توزیع هستند. لذا

$$\begin{aligned} vecE(XAX') &= E[(Y + \mu) \otimes (Y + \mu)vecA] \\ &= E[(Y \otimes Y)vecA] + vec(\mu A \mu'). \end{aligned}$$

از طرفی $E(Y \otimes Y) = vec\Sigma vec'\psi$ در نتیجه

$$vecE(XAX') = vec\Sigma vec'\psi vecA + vec(\mu A \mu') = tr(\psi A)vec\Sigma + vec(\mu A \mu').$$

و اثبات کامل است. $\square$

لم ۱-۱ (اندرس، ۲۰۰۳). فرض کنید A, B, C_1 و C_2 ماتریس‌های به ترتیب با ابعاد $n \times q, m \times p$

$m \times n$ و $p \times q$ باشند. همچنین فرض کنید $X \sim MN_{p \times n}(\mu, \Omega \otimes \Phi)$ در این صورت

$$AX + C_1 \sim MN_{m \times n}(A\mu + C_1, (A\Omega A') \otimes \Phi),$$

$$XB + C_2 \sim MN_{p \times q}(\mu B + C_2, \Omega \otimes (B'\Phi B)),$$

$$AXB \sim MN_{p \times q}(A\mu B, (A\Omega A') \otimes (B'\Phi B)).$$

اثبات: فرض کنید $Z = AX + C_1$ در این صورت تابع مشخصه $\varphi_Z(T)$ به صورت زیر است:

$$\begin{aligned}
 \varphi_Z(T) &= E[\exp(\text{itr}(T'Z))] = E[\exp(\text{itr}(T'(AX + C_1)))] \\
 &= \exp(\text{itr}(T'C_1)) E[\exp(\text{itr}(A'T)'X)] \\
 &= \exp(\text{itr}(T'C_1)) \exp\left(\text{itr}(A'T)'\mu - \frac{1}{2} \text{trace}(\Omega A'T\Phi T'A)\right) \\
 &= \exp\left[\text{itr}T'(A\mu + C_1) - \frac{1}{2} \text{trace}(A\Omega A'T\Phi T')\right]
 \end{aligned}$$

که تابع مشخصه توزیع نرمال ماتریسی با میانگین $A\mu + C_1$ و ماتریس کوواریانس $(A\Omega A') \otimes \Phi$ است. با روشی مشابه، دیگر رابطه‌ها نیز اثبات می‌شود. $\square$

قضیه ۹-۱ (کولو و وان‌رز، ۲۰۰۵) اگر $X \sim MN_{p \times n}(\mu, \Sigma \otimes \psi)$ باشد، و $U = X\psi^{-1}X'$ در این صورت

$$U \sim W_p(\Sigma, r(\psi), \Delta),$$

که در آن $\Delta = \mu\psi^{-1}X'$

اثبات: با توجه به اینکه ψ ماتریس متقارن است، لذا از قضیه (A-۱۵) داریم $\psi = Q\Lambda Q'$ از طرفی با استفاده از لم ۱-۱ می‌توان نوشت

$$XQ\Lambda^{-\frac{1}{2}} \sim MN_{p,r(\psi)}\left(\mu Q\Lambda^{-\frac{1}{2}}, \Sigma \otimes I\right).$$

بنابراین طبق تعریف ۱-۱۳، $XQ\Lambda^{-\frac{1}{2}}\Lambda^{-\frac{1}{2}}Q'X' = X\psi^{-1}X'$ و اثبات کامل است. $\square$

لم ۲-۱ (اندرسن، ۲۰۰۳) فرض کنید متغیر تصادفی X دارای توزیع ویشارت معکوس با پارامترهای Σ و k است، $X \sim W_n^{-1}(\Sigma, k)$ در این صورت داریم:

$$E(X) = \frac{\Sigma^{-1}}{k-2}.$$

لم ۳-۱ (آرشی، ۲۰۰۸). گشتاور n متغیر تصادفی $Y = \frac{1}{\chi_{q,\theta}^2}$ برابر است با $E(Y^n) = \psi(q, \theta, n)$

که در آن

$$\psi(q, \theta, n) = \sum_{k=0}^{\infty} \frac{e^{-\theta/2} (\theta/2)^k}{\Gamma(k+1)} \times \frac{\Gamma(q/2 + k - n)}{2^n \Gamma(q/2 + k)}.$$

همچنین در حالت $\theta = 0$ داریم:

$$E \left[\left(\frac{1}{\chi_q^2} \right)^n \right] = \frac{\Gamma(q/2 - n)}{2^n \Gamma(q/2)}.$$

اثبات: با استفاده از تعریف ۱-۶، داریم:

$$\begin{aligned} E \left[\left(\frac{1}{\chi_{q,\theta}^2} \right)^n \right] &= \sum_{k=0}^{\infty} \frac{e^{-\theta/2} (\theta/2)^k}{\Gamma(k+1)} E \left[\left(\frac{1}{\chi_{q+2k}^2} \right)^n \right] \\ &= \sum_{k=0}^{\infty} \frac{e^{-\theta/2} (\theta/2)^k}{\Gamma(k+1)} \times \frac{\Gamma(q/2 + k - n)}{2^n \Gamma(q/2 + k)}. \end{aligned}$$

□

و اثبات کامل است.

نتیجه ۲-۱* فرض کنید $\chi_{q+s}^2(\Delta)$ یک متغیر تصادفی کی-دو غیرمرکزی باشد، در این صورت روابط

زیر از لم ۲-۱ نتیجه می‌شوند:

$$\begin{aligned} E \left(\chi_{p+2}^{-2}(\Delta) \right) - E \left(\chi_{p+2}^{-2}(\Delta) \right) &= 2E \left(\chi_{p+2}^{-4}(\Delta) \right), \quad \forall p > 2; \\ E \left(\chi_{p+2}^{-2}(\Delta) \right) - (p-2)E \left(\chi_{p+2}^{-4}(\Delta) \right) &= \Delta E \left(\chi_{p+4}^{-4}(\Delta) \right). \end{aligned}$$

اثبات: با استفاده از لم ۳-۱، برای $q = p + 2$ ، $q = p$ و به ازای $n = 1$ داریم:

$$\begin{aligned}
 E\left(\chi_p^{-2}(\Delta)\right) &= \sum_{k=0}^{\infty} \frac{e^{-\Delta/2}(\Delta/2)^k}{\Gamma(k+1)} \times \frac{\Gamma(p/2 + k - 1)}{2\Gamma(p/2 + k)} \\
 &= \sum_{k=0}^{\infty} \frac{e^{-\Delta/2}(\Delta/2)^k}{\Gamma(k+1)} \times \frac{1}{p + 2k - 2}, \\
 E\left(\chi_{p+2}^{-2}(\Delta)\right) &= \sum_{k=0}^{\infty} \frac{e^{-\Delta/2}(\Delta/2)^k}{\Gamma(k+1)} \times \frac{\Gamma(p + 2/2 + k - 1)}{2\Gamma(p + 2/2 + k)} \\
 &= \sum_{k=0}^{\infty} \frac{e^{-\Delta/2}(\Delta/2)^k}{\Gamma(k+1)} \times \frac{1}{p + 2k}.
 \end{aligned}$$

در نتیجه

$$E\left(\chi_p^{-2}(\Delta)\right) - E\left(\chi_{p+2}^{-2}(\Delta)\right) = \sum_{k=0}^{\infty} \frac{e^{-\Delta/2}(\Delta/2)^k}{\Gamma(k+1)} \times \frac{1}{p + 2k} \times \frac{1}{p + 2k - 2}.$$

از طرفی برای $q = p + 2$ و $n = 2$ داریم:

$$\begin{aligned}
 2E\left(\chi_{p+2}^{-4}(\Delta)\right) &= \sum_{k=0}^{\infty} \frac{e^{-\Delta/2}(\Delta/2)^k}{\Gamma(k+1)} \times \frac{\Gamma(p + 2/2 + k - 2)}{4\Gamma(p + 2/2 + k)} \\
 &= \sum_{k=0}^{\infty} \frac{e^{-\Delta/2}(\Delta/2)^k}{\Gamma(k+1)} \times \frac{1}{p + 2k} \times \frac{1}{p + 2k - 2}.
 \end{aligned}$$

و اثبات رابطه اول کامل است.

برای رابطه دوم بنا بر لم ۱-۳ داریم:

$$\begin{aligned}
 &E\left(\chi_{p+2}^{-2}(\Delta)\right) - (p - 2)E\left(\chi_{p+2}^{-4}(\Delta)\right) \\
 &= \sum_{k=0}^{\infty} \frac{e^{-\Delta/2}(\Delta/2)^k}{\Gamma(k+1)} \times \frac{2k}{(p + 2k)(p + 2k - 2)}.
 \end{aligned}$$

حال با استفاده از تغییر متغیر $k = j + 1$ خواهیم داشت:

$$\sum_{j=0}^{\infty} \frac{e^{-\Delta/2} (\Delta/2)^j}{(j+1)\Gamma(j+1)} \times \frac{2(j+1) \times (\Delta/2)}{(p+2j+2)(p+2j)} = \Delta E \left(\chi_{p+4}^{-4}(\Delta) \right).$$

□

و اثبات کامل است.

لم ۴-۱ (آرشی، ۲۰۰۸). اگر p عدد طبیعی بزرگتر از $2m$ باشد، آن گاه

$$E \left[\left(1 - \frac{d}{\chi_{q,\theta}^2} \right) I(\chi_{q,\theta}^2 \leq d) \right] = \chi_{q,\theta}^2(d) + \Xi_1(q, \theta, d),$$

$$E \left[\left(1 - \frac{d}{\chi_{q,\theta}^2} \right)^2 I(\chi_{q,\theta}^2 \leq d) \right] = \chi_{q,\theta}^2(d) + \Xi_2(q, \theta, d),$$

که در آن d عددی حقیقی، $I(A)$ تابع نشانگر مجموعه A

$$\Xi_1(q, \theta, d) = \sum_{r=0}^{\infty} \frac{-d e^{-\theta/2} (\theta/2)^r \chi_{q+2r}^2(d)}{r! (q+2r-2)},$$

$$\Xi_2(q, \theta, d) = \sum_{r=0}^{\infty} \frac{d[d-2q-4r+8] e^{-\theta/2} (\theta/2)^r \chi_{q+2r}^2(d)}{r! (q+2r-2)(q+2r-4)},$$

و $\chi_q^2(\alpha)$ چندک α درصد بالایی توزیع کی-دو با q درجه آزادی است.

اثبات: با استفاده از تعریف ۹-۱،

$$E \left[\left(\frac{1}{\chi_{q,\theta}^2} \right)^m I(\chi_{q,\theta}^2 \leq d) \right] = \sum_{r=0}^{\infty} \frac{e^{-\theta/2} (\theta/2)^r}{r!} E \left[\left(\frac{1}{\chi_{q+2r}^2} \right)^m I(\chi_{q+2r}^2 \leq d) \right]$$

$$= \sum_{r=0}^{\infty} \frac{e^{-\theta/2} (\theta/2)^r}{2^m r!} \times \frac{\Gamma(q/2 + r - m)}{\Gamma(q/2 + r)} \times \chi_{q+2r}^2(d)$$

بنابراین می توان نتیجه گرفت

$$\begin{aligned}
 E \left[\left(1 - \frac{d}{\chi_{q,\theta}^2} \right)^2 I(\chi_{q,\theta}^2 \leq d) \right] &= \chi_{q,\theta}^2(d) - 2dE \left[\left(\frac{1}{\chi_{q,\theta}^2} \right) I(\chi_{q,\theta}^2 \leq d) \right] \\
 &\quad + d^2E \left[\left(\frac{1}{\chi_{q,\theta}^2} \right)^2 I(\chi_{q,\theta}^2 \leq d) \right] \\
 &= \chi_{q,\theta}^2(d) + \sum_{r=0}^{\infty} \frac{e^{-\theta/2}(\theta/2)^r}{r! \Gamma(q/2 + r)} \times \chi_{q+2r}^2(d) \\
 &\quad \times \left[\frac{d^2 \Gamma(q/2 + r - 2)}{4} - d \Gamma(q/2 + r - 1) \right] \\
 &= \chi_{q,\theta}^2(d) \\
 &\quad + \sum_{r=0}^{\infty} \frac{d[d - 2q - 4r + 8]e^{-\theta/2}(\theta/2)^r \chi_{q+2r}^2(d)}{r! (q + 2r - 2)(q + 2r - 4)}.
 \end{aligned}$$

□

و اثبات کامل است.

لم ۵-۱ (جاج و باک، ۱۹۷۸) فرض کنید متغیر تصادفی ω ، دارای توزیع نرمال چندمتغیره n بعدی با میانگین μ و ماتریس کوواریانس I_n باشد. همچنین A ماتریسی معین مثبت و $\phi(\cdot)$ تابعی اندازه پذیر بورل^{۵۱} هستند. در این صورت داریم

$$\begin{aligned}
 E[\phi(\omega' \omega) \omega] &= \mu E[\phi(\chi_{n+2, \mu' \mu/2}^2)], \\
 E[\phi(\omega' \omega) \omega' A \omega] &= \text{trace}(A) E[\phi(\chi_{n+2, \mu' \mu/2}^2)] + \mu' A \mu E[\phi(\chi_{n+4, \mu' \mu/2}^2)].
 \end{aligned}$$

⁵¹ Borel measurable

لم ۱-۶ (چو^{۵۲}، ۱۹۷۳) فرض کنید متغیر تصادفی X با n مولفه، به صورت $X \sim EC_n(\mu, \Sigma, g)$ توزیع شده است که در آن g تابع مولد چگالی می‌باشد. در این صورت تابع چگالی احتمال X را می‌توان به صورت زیر بیان کرد:

$$f_X(x) = \int_0^\infty \omega(t) h(x) dt$$

که در آن $h(\cdot)$ نشان‌دهنده تابع چگالی احتمال $N_n(\mu, t^{-1}\Sigma)$ است. بنابراین

$$f_X(x) = \int_0^\infty \omega(t) \frac{1}{(2\pi)^{n/2}} |t^{-1}\Sigma|^{-1/2} \exp \left[-\frac{1}{2} (x - \mu)' (t^{-1}\Sigma)^{-1} (x - \mu) \right] dt,$$

که در آن

$$\omega(t) = (2\pi)^{n/2} |t^{-1}\Sigma|^{1/2} \mathcal{L}^{-1}[f(s)],$$

$$\mathcal{L}^{-1}[f(s)] \text{ نشان‌دهنده تبدیل معکوس لاپلاس } h(s) \text{ به ازای } s = \frac{(x-\mu)'\Sigma^{-1}(x-\mu)}{2} \text{ می‌باشد.}$$

برای آگاهی از خواص تبدیل معکوس لاپلاس به گرادشتاین و ری‌زیک^{۵۳} (۲۰۰۰) مراجعه کنید.

توجه به این نکته ضروری است که ممکن است تابعی تبدیل معکوس لاپلاس نداشته باشد. در این صورت لم فوق کاربردی ندارد. اما در بیشتر توزیع‌های آماری چندمتغیره معروف، تبدیل معکوس لاپلاس وجود دارد. در این زمینه، جدول (۱-۱) که از چنگ^{۵۴} (۱۹۹۹) برگرفته شده است را ببینید.

Chu⁵²
Gradshteyn and Ryzhik⁵³
Cheong⁵⁴

جدول (۱-۱) مقادیر $\omega(t)$ برای برخی توزیع‌های آماری به ازای $s = \frac{(x-\mu)'\Sigma^{-1}(x-\mu)}{2}$

توزیع	تابع چگالی احتمال	$\omega(t)$
نرمال چند متغیره	$\frac{ \Sigma ^{-1/2}}{2\pi^{n/2}} e^{-s}$	$\delta(t)$
اسلش تعمیم یافته	$\frac{\nu s^{-n/2-\nu} \Sigma ^{-1/2}}{(2\pi)^{n/2}} [\Gamma(n/2 + \nu) - \Gamma(n/2 + \nu - s)]$	$\nu \tau^{\nu-1}$
چندمتغیره نوع کاتز	$\frac{q^{m-1+n/2} \Gamma(n/2) \Sigma ^{-1/2}}{\pi^{n/2} \Gamma(m+1+n/2)} (2s)^{m-1} e^{-2qs}$	$\frac{(2q)^{m-1+n/2} \Gamma(n/2)}{\Gamma(m-1+n/2)} \tau^{-n/2} \delta^{m-1}(\tau - 2q)$
پیرسن چندمتغیره نوع VII	$\frac{\Gamma(m) \Sigma ^{-1/2}}{(q\pi)^{n/2} \Gamma(m-n/2)} (1 + 2s/q)^{-m}$	$\frac{t^{m-n/2-1} e^{-qt/2}}{(q/2)^{n/2-m} \Gamma(m-n/2)}$
لاپلاس چند متغیره	$\frac{\Gamma(n/2) \Sigma ^{-1/2} e^{-\sqrt{2}s}}{(2\pi)^{n/2} \Gamma(n)}$	$\delta(t - \sqrt{2})$

در جدول (۱-۱)، $\delta(t)$ تابع دلتا دیراک^{۵۵} می‌باشد و دارای خاصیت زیر است

$$\int_0^\infty \delta(t) dt = 1.$$

فصل دوم

مدل رگرسیون چندمتغیره محدود شده

۱-۲ مقدمه

تحلیل رگرسیونی کاربردی‌ترین و مهم‌ترین روش آماری در تحلیل‌های آماری، اقتصادی، مالی، صنعتی و غیره می‌باشد. به‌ویژه در چهار دهه اخیر، به طور ماهرانه ای در حل و بررسی مسایل کاربردی مورد استفاده قرار گرفته است. در بین تعداد زیادی از روش‌های تحلیل رگرسیونی، روش برآورد کمترین توان-های دوم تعمیم‌یافته، براساس نظریه معروف گوس-مارکوف، که در زیر به آن اشاره می‌شود، روشی پایه-ای است و هنوز نقشی اساسی در بسیاری از جنبه‌های نظری و کاربردی مدل‌های رگرسیونی به ویژه مدل رگرسیون خطی، ایفا می‌کند.

قضیه گوس مارکوف (کاریا و کوراتا^۱، ۲۰۰۴). یک برآوردگر خطی نااریب یکتا با کمترین واریانس برای تابع برآوردپذیر $c'\theta$ برابر است با $c'\hat{\theta}$ ، به طوری که $\hat{\theta}$ برآوردگر کمترین توان‌های دوم بردار θ است.

در شرایطی که ساختار ماتریس کوواریانس، درمدل رگرسیون خطی، کلی است (به صورت Σ یا $\sigma^2 V$ و نه $\sigma^2 I$) در مسایل برآورد ضرایب رگرسیونی، یک روش برآوردیابی کارا، روش گوس-مارکوف یا کمترین توان‌های دوم تعمیم‌یافته می‌باشد. با این روش بهترین برآوردگر خطی نااریب^۲ (BLUE) برای پارامترهای مدل رگرسیونی به‌دست می‌آید.

در این فصل، ابتدا با فرض این که مدل تحت بررسی یک مدل رگرسیون خطی چندگانه است، برآوردگرهایی از نوع کمترین توان‌های دوم تعمیم‌یافته برای بردار ضرایب رگرسیونی به‌دست آورده و خواص آن‌ها را در حالت خاصی مورد بررسی قرار می‌دهیم. در ادامه، با توجه به این که در مسائل کاربردی اغلب با مشاهدات ماتریسی روبه‌رو هستیم، فرض می‌کنیم مدل تحت بررسی یک مدل رگرسیون

¹ Kariya and Kurata
² Best linear unbiased estimator

خطی چندمتغیره است. از نتایج رگرسیون چندگانه استفاده کرده و برآوردگرهایی از نوع کمترین توان‌های دوم تعمیم‌یافته برای ماتریس ضرایب رگرسیونی را نیز به‌دست خواهیم آورد.

مدل رگرسیون خطی چندگانه، مدلی انعطاف‌پذیر است، به این معنی که بررسی برآوردگرهای ضرایب رگرسیونی در این مدل، امکان تعمیم آن را، به مدل‌های رگرسیون به ظاهر نامرتبط^۳، مدل همبستگی ترتیبی^۴، مدل ناهمپراش^۵ و مدل با همبستگی برابر^۶ فراهم می‌سازد. برای آگاهی بیشتر در خصوص مدل‌های فوق و بررسی برآوردگرهای از نوع کمترین توان‌های دوم تعمیم‌یافته به زلنر^۷ (۱۹۶۲)، اندرسن (۱۹۴۸)، تیلور^۸ (۱۹۷۸)، سریواستوا و جایلز^۹ (۱۹۸۷)، کوراتا^{۱۰} (۱۹۹۹) و آرشی (۲۰۱۲) مراجعه کنید.

۲-۲ مدل رگرسیون خطی چندگانه

در حقیقت مدل رگرسیون خطی چندگانه به صورت

$$Y_i = x_{i1}\beta_1 + x_{i2}\beta_2 + \dots + x_{ik}\beta_k + e_i \\ = x'_i\beta + e_i,$$

است، که در آن $x'_i = (x_{i1}, \dots, x_{ik})$ و $i = 1, \dots, n$ و $\beta' = (\beta_1, \dots, \beta_k)$ می‌باشد. نمایش

برداری مدل رگرسیون خطی چندگانه فوق به صورت

$$Y = X\beta + e,$$

³ Seemingly unrelated regression model

⁴ Serial correlation model

⁵ Heteroscedastic model

⁶ Equi-correlated model

⁷ Zellner

⁸ Taylor

⁹ Srivastava

¹⁰ Kurata

می‌باشد، که در آن Y یک بردار $n \times 1$ پاسخ یا متغیر وابسته، X یک ماتریس $n \times k$ غیرتصادفی و معلوم از مقادیر تبیینی یا پیش‌بین^{۱۱} با رتبه کامل k ، β بردار ضرایب رگرسیونی مجهول (پارامترهای مدل) و $e = (e_1, \dots, e_n)'$ بردار خطای تصادفی هستند. به عبارتی داریم

$$\begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} x_{11} & \cdots & x_{1p} \\ \vdots & \ddots & \vdots \\ x_{n1} & \cdots & x_{np} \end{bmatrix} \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_n \end{bmatrix} + \begin{bmatrix} e_1 \\ \vdots \\ e_n \end{bmatrix}.$$

اغلب در مسائل کاربردی به خصوص در طرح آزمایش‌ها، ستون اول X بردار $n \times 1$ از یک‌ها می‌باشد.

در این قسمت فرض کنید Y یک بردار تصادفی n بعدی با میانگین $X\beta$ و ماتریس کوواریانس Σ است که به صورت زیر نمایش می‌دهیم:

$$Y \sim_n (X\beta, \Sigma) \quad (۱-۲)$$

قضیه ۱-۲ (کاریا و کوراتا، ۲۰۰۴). در مدل (۱-۲) فرض کنید X یک ماتریس رتبه کامل k و $\Sigma =$

$\sigma^2 V$ که در آن V یک ماتریس معین مثبت و معلوم است. آنگاه با قرار دادن $C = (X'V^{-1}X)$

الف- برآوردگر کمترین توان‌های دوم محدودنشده، $\tilde{\beta}$ ، برای بردار β به صورت زیر است:

$$\tilde{\beta} = C^{-1}X'V^{-1}Y. \quad (۲-۲)$$

ب- ماتریس کوواریانس $\tilde{\beta}$ عبارتست از:

$$\text{cov}(\tilde{\beta}) = \sigma^2 C^{-1}. \quad (۳-۲)$$

ج- برآوردگر ناریب σ^2 به صورت زیر می باشد:

$$S^2 = \frac{(Y - X\tilde{\beta})' V^{-1} (Y - X\tilde{\beta})}{n-k}. \quad (4-2)$$

اثبات: الف- با توجه به قضیه A-17 از اینکه V ماتریس معین مثبت است، ماتریس نامفرد P $n \times n$

وجود دارد به طوری که $V = PP'$. با ضرب نمودن P^{-1} در طرفین رابطه $Y = X\beta + e$ داریم:

$$P^{-1}Y = P^{-1}X\beta + P^{-1}e$$

که در آن،

$$\begin{aligned} E(P^{-1}e) &= P^{-1}E(e) \\ &= P^{-1}0 = 0. \end{aligned}$$

همچنین

$$\begin{aligned} cov(P^{-1}e) &= P^{-1}cov(e)(P^{-1})' \\ &= P^{-1}\sigma^2 V(P^{-1})' \\ &= \sigma^2 P^{-1}PP'(P')^{-1} = \sigma^2 I. \end{aligned}$$

بنابراین،

$$\begin{aligned} \tilde{\beta} &= [X'(P^{-1})'P^{-1}X]^{-1}X'(P^{-1})'P^{-1}Y \\ &= [X'(P')^{-1}P^{-1}X]^{-1}X'(PP')^{-1}Y \\ &= [X'(PP')^{-1}X]^{-1}X'(PP')^{-1}Y \\ &= [X'V^{-1}X]^{-1}X'V^{-1}Y \\ &= C^{-1}X'V^{-1}Y. \end{aligned}$$

ب- با استفاده از قضیه ۱-۲ داریم:

$$Cov[C^{-1}X'V^{-1}Y] = C^{-1}X'V^{-1}cov(Y)V^{-1}XC^{-1}$$

$$= C^{-1}X'V^{-1}\sigma^2VV^{-1}XC^{-1} = \sigma^2C^{-1}$$

ج- طبق تعریف می توان نوشت

$$\begin{aligned}(Y - X\tilde{\beta})'V^{-1}(Y - X\tilde{\beta}) &= \\ Y'V^{-1}Y - 2Y'V^{-1}XC^{-1}X'V^{-1}Y + Y'V^{-1}XC^{-1}X'V^{-1}XC^{-1}X'V^{-1}Y \\ &= Y'(V^{-1} - V^{-1}XC^{-1}X'V^{-1})Y.\end{aligned}$$

اکنون بنا به قضیه ۳-۱ و با در نظر گرفتن $\mathcal{A} = [V^{-1} - V^{-1}XC^{-1}X'V^{-1}]$ و $\Sigma = \sigma^2V$ و $\mu = X\beta$ داریم:

$$\begin{aligned}E(Y - X\tilde{\beta})'V^{-1}(Y - X\tilde{\beta}) &= E[Y'(V^{-1} - V^{-1}XC^{-1}X'V^{-1})Y] \\ &= \text{trace}[(V^{-1} - V^{-1}XC^{-1}X'V^{-1})\sigma^2V] \\ &\quad + \beta'X'(V^{-1} - V^{-1}XC^{-1}X'V^{-1})\beta X \\ &= \sigma^2\text{trace}(I_n - I_k) = \sigma^2(n - k).\end{aligned}$$

و اثبات کامل است. $\square$

قضیه ۲-۲ (رنچر، ۲۰۰۰). اگر $Y \sim N_n(X\beta, \sigma^2V)$ که در آن X ماتریس $n \times k$ با رتبه k و V ماتریس معین مثبت معلوم است، آنگاه برآوردهای درستنمایی ماکزیمم β و σ^2 عبارتست از:

$$\hat{\beta} = C^{-1}X'V^{-1}Y \quad (۵-۲)$$

$$\hat{\sigma}^2 = \frac{1}{n}(Y - X\hat{\beta})'V^{-1}(Y - X\hat{\beta}) \quad (۶-۲)$$

اثبات: با استفاده از تابع درستنمایی مدل (۱-۲) داریم

$$L(\beta, \sigma^2) = \frac{1}{(2\pi)^{n/2}(\sigma^2)^{n/2}|V|^{n/2}} \exp \frac{-(y - X\beta)'V^{-1}(y - X\beta)}{2\sigma^2}.$$

بنابراین

$$\text{Ln}L(\beta, \sigma^2) = -\frac{n}{2}\text{Ln}(2\pi) - \frac{n}{2}\text{Ln}\sigma^2 - \frac{n}{2}\text{Ln}|V| - \frac{(y - X\beta)'V^{-1}(y - X\beta)}{2\sigma^2}.$$

حال با مشتق گیری از لگاریتم درستنمایی نسبت به پارامترها و مساوی صفر قرار دادن آنها، می توان

نتیجه گرفت

$$\frac{\partial \text{Ln}L(\beta, \sigma^2)}{\partial \beta} = \frac{-2X'V^{-1}(y - X\beta)}{2\sigma^2} = 0.$$

بنابراین

$$\hat{\beta} = C^{-1}X'V^{-1}Y.$$

همچنین

$$\frac{\partial \text{Ln}L(\beta, \sigma^2)}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{(y - X\hat{\beta})'V^{-1}(y - X\hat{\beta})}{2\sigma^4} = 0.$$

بنابراین

$$\hat{\sigma}^2 = \frac{(y - X\hat{\beta})'V^{-1}(y - X\hat{\beta})}{n}.$$

□

و اثبات کامل است.

حال دوباره مدل خطی (۱-۲) را در نظر بگیرید و فرض کنید قیدی خطی به صورت زیر به مدل تحمیل

شده است.

$$B\beta + b = 0 \quad (7-2)$$

که در آن B در اندازه $q \times k$ ، ماتریس معلوم غیر صفر با رتبه q است که ماتریس محدودیت نامیده می-شود و b بردار معلوم با q مولفه است. مدل خطی معرفی شده در (۱-۲) تحت محدودیت (۷-۲)، مدل خطی محدودشده نامیده می-شود. برآوردگری که بر اساس مدل‌های محدودشده به دست می‌آید، برآوردگر محدودشده (RE) نامیده می‌شود.

به عنوان مثالی برای درک بهتری از ماتریس محدودیت، فرض کنید بر اساس اطلاعات پیشین از نمونه‌ای تصادفی، دریافتیم که روابط زیر بین بردار پارامترها برقرار است. (لازم به ذکر است که از ساختار زیر به عنوان ضرایب مقابله در طرح آزمایش‌ها استفاده می‌شود).

$$\beta_1 + \beta_2 + \beta_3 = 1,$$

$$\beta_3 - 2\beta_2 = 3,$$

$$2\beta_1 + 3\beta_2 = 0.$$

در این صورت

$$B = \begin{bmatrix} 1 & 1 & 1 \\ 0 & -2 & 1 \\ 2 & 3 & 3 \end{bmatrix}, \quad b = \begin{bmatrix} 1 \\ 3 \\ 0 \end{bmatrix}.$$

قضیه ۲-۳ (کوآن^{۱۲}، ۲۰۰۲). مدل خطی (۱-۲) را تحت محدودیت (۷-۲) در نظر بگیرید. در این صورت برآوردگر کمترین توان‌های دوم محدود شده (RLSE)^{۱۳} عبارتست از:

$$\hat{\beta} = \tilde{\beta} - C^{-1}B'(BC^{-1}B')^{-1}(B\tilde{\beta} + b) \quad (8-2)$$

¹² Kuan

¹³ Restricted least square estimator

که دارای خواص زیر است :

الف- تحت قید $B\beta + b = 0$ برآوردگری ناریب برای β است.

ب- کوواریانس آن برابر است با:

$$\sigma^2[C^{-1} - C^{-1}B'(BC^{-1}B')^{-1}BC^{-1}] \quad (9-2)$$

اثبات: برای بدست آوردن RE تحت فرضیه $b + B\beta = 0$ از روشی که توسط راثو در سال ۱۹۹۵ پیشنهاد داده شد، استفاده می‌کنیم. این روش تلفیقی از روش نظریه کمترین توان‌های دوم و لاگرانژ است. تابع زیر که تابعی خطی از مجموع توان‌های دوم خطا و محدودیت اعمال شده بر روی مدل است، مینیمم می‌گردد.

$$\begin{aligned} Q(\beta, \lambda) &= e'V^{-1}e + 2\lambda'(B\beta + b) \\ &= (Y - X\beta)'V^{-1}(Y - X\beta) + 2\lambda'(B\beta + b). \end{aligned}$$

با مشتق‌گیری نسبت به λ و β و مساوی صفر قرار دادن، داریم:

$$\begin{aligned} \frac{\partial Q(\beta, \lambda)}{\partial \lambda} &= 2(B\beta + b) = 0 \Rightarrow B\beta = -b, \\ \frac{\partial Q(\beta, \lambda)}{\partial \beta} &= -2X'V^{-1}(Y - X\beta) + 2B'\lambda = 0. \end{aligned}$$

بنابراین

$$2X'V^{-1}X\beta = 2X'V^{-1}Y - 2B'\lambda$$

و در نهایت داریم:

$$\beta = C^{-1}X'V^{-1}Y - C^{-1}B'\lambda$$

از طرفی

$$B\beta + b = B[C^{-1}X'V^{-1}Y - C^{-1}B'\lambda] + b = 0,$$

در نتیجه،

$$BC^{-1}B'\lambda = BC^{-1}X'V^{-1}Y + b,$$

و بنابراین

$$\lambda = (BC^{-1}B')^{-1}(BC^{-1}X'V^{-1}Y + b).$$

با جایگذاری رابطه فوق در جواب مشتقات بدست آمده، می توان نتیجه گرفت:

$$\beta = C^{-1}X'V^{-1}Y - C^{-1}B'(BC^{-1}B')^{-1}(BC^{-1}X'V^{-1}Y + b),$$

و بنابراین RLSE عبارتست از:

$$\hat{\beta} = \tilde{\beta} - C^{-1}B'(BC^{-1}B')^{-1}(B\tilde{\beta} + b).$$

الف- از آنجایی که $E(\tilde{\beta}) = \beta$ ، لذا برای اثبات نااریب بودن برآوردگر $\hat{\beta}$ می توان نوشت:

$$\begin{aligned} E(\hat{\beta}) &= E(\tilde{\beta}) - C^{-1}B'(BC^{-1}B')^{-1}(BE(\tilde{\beta}) + b) \\ &= \beta. \end{aligned}$$

ب- ماتریس کوواریانس $\hat{\beta}$ به صورت زیر به دست می آید :

$$\begin{aligned} Var(\hat{\beta}) &= Var\left(\tilde{\beta} - C^{-1}B'(BC^{-1}B')^{-1}(B\tilde{\beta} + b)\right) \\ &= Var(\tilde{\beta}) + C^{-1}B'(BC^{-1}B')^{-1}BVar(\tilde{\beta})B'(BC^{-1}B')^{-1}BC^{-1} \\ &\quad - 2Cov(\tilde{\beta}, C^{-1}B'(BC^{-1}B')^{-1}B\tilde{\beta}) \\ &= \sigma^2C^{-1} + \sigma^2C^{-1}B'(BC^{-1}B')^{-1}BC^{-1}B'(BC^{-1}B')^{-1}BC^{-1} \\ &\quad - 2\sigma^2C^{-1}B'(BC^{-1}B')^{-1}BC^{-1} \\ &= \sigma^2[C^{-1} - C^{-1}B'(BC^{-1}B')^{-1}BC^{-1}]. \end{aligned}$$

□

و اثبات کامل است.

قضیه ۴-۲ (میرهد، ۱۹۸۲). تحت قید $B\beta + b = 0$ برآوردگر $S^{2*} = \frac{(Y - X\hat{\beta})' V^{-1} (Y - X\hat{\beta})}{n - k + q}$

برآوردگری ناریب برای σ^2 است.

اثبات: ابتدا صورت کسر S^{2*} را به روش زیر بسط می‌دهیم:

$$\begin{aligned} & (Y - X\hat{\beta})' V^{-1} (Y - X\hat{\beta}) \\ &= \{Y - X[\tilde{\beta} - C^{-1}B'(BC^{-1}B')^{-1}(B\tilde{\beta} + b)]\}' V^{-1} \\ &\quad \times \{Y - X[\tilde{\beta} - C^{-1}B'(BC^{-1}B')^{-1}(B\tilde{\beta} + b)]\} \\ &= \{(Y - X\tilde{\beta}) + XC^{-1}B'(BC^{-1}B')^{-1}(B\tilde{\beta} + b)\}' V^{-1} \\ &\quad \times \{(Y - X\tilde{\beta}) + XC^{-1}B'(BC^{-1}B')^{-1}(B\tilde{\beta} + b)\} \\ &= (Y - X\tilde{\beta})' V^{-1} (Y - X\tilde{\beta}) \\ &\quad + (Y - X\tilde{\beta})' V^{-1} XC^{-1}B'(BC^{-1}B')^{-1}(B\tilde{\beta} + b) \\ &\quad + (B\tilde{\beta} + b)' (BC^{-1}B')^{-1} BC^{-1}X' V^{-1} (Y - X\tilde{\beta}) \\ &\quad + (B\tilde{\beta} + b)' (BC^{-1}B')^{-1} BC^{-1}CC^{-1}B'(BC^{-1}B')^{-1}(B\tilde{\beta} + b). \end{aligned}$$

حال با توجه به اینکه $(Y - X\tilde{\beta})' V^{-1} XC^{-1}B'(BC^{-1}B')^{-1}(B\tilde{\beta} + b) = 0$ داریم:

$$\begin{aligned} & (Y - X\hat{\beta})' V^{-1} (Y - X\hat{\beta}) \\ &= (Y - X\tilde{\beta})' V^{-1} (Y - X\tilde{\beta}) + (B\tilde{\beta} + b)' (BC^{-1}B')^{-1} (B\tilde{\beta} + b). \end{aligned}$$

و در نتیجه

$$\begin{aligned} & E(Y - X\hat{\beta})' V^{-1} (Y - X\hat{\beta}) \\ &= E(Y - X\tilde{\beta})' V^{-1} (Y - X\tilde{\beta}) + E(B\tilde{\beta} + b)' (BC^{-1}B')^{-1} (B\tilde{\beta} + b). \end{aligned}$$

با توجه به رابطه (۴-۲)، $E(Y - X\tilde{\beta})'V^{-1}(Y - X\tilde{\beta}) = \sigma^2(n - k)$ و برای رابطه دوم بنا به قضیه ۳-۱ و با در نظر گرفتن $A = (BC^{-1}B')$ و $\Sigma = B\sigma^2C^{-1}B'$ و $\mu = B\beta + b$ داریم:

$$E(B\tilde{\beta} + b)'(BC^{-1}B')^{-1}(B\tilde{\beta} + b) = \text{trace}[(BC^{-1}B')^{-1}(B\sigma^2C^{-1}B')] \\ = \sigma^2q.$$

لذا می توان نتیجه گرفت

$$E(Y - X\hat{\beta})'V^{-1}(Y - X\hat{\beta}) = \sigma^2(n - k) + \sigma^2q = \sigma^2(n - k + q)$$

و اثبات کامل است. $\square$

قضیه ۵-۲ (راویشانکر و دی، ۲۰۰۲). تحت مفروضات قضیه ۳-۲ و با فرض

$$\text{cov}(Y) = \text{cov}(e) = \Sigma,$$

الف- برآوردگر GLS به صورت زیر است:

$$\tilde{\beta}_{GLS} = (X'\Sigma^{-1}X)^{-1}X'\Sigma^{-1}Y. \quad (۹-۲)$$

ب- ماتریس کوواریانس $\tilde{\beta}_{GLS}$ عبارتست از:

$$\text{cov}(\tilde{\beta}_{GLS}) = (X'\Sigma^{-1}X)^{-1}. \quad (۱۰-۲)$$

ج- با این فرض که ماتریس X غیرتصادفی و $E(Y) = X\beta$ ، برآوردگر $\tilde{\beta}_{GLS}$ ، برای پارامتر β ، بهترین برآوردگر نااریب خطی است.

د- برآوردگر RGLS به صورت زیر است:

$$\hat{\beta}_{RGLS} = \tilde{\beta}_{GLS} - (X'\Sigma^{-1}X)^{-1}B'[B(X'\Sigma^{-1}X)^{-1}B']^{-1}(B\tilde{\beta}_{GLS} + b). \quad (۱۱-۲)$$

ه- ماتریس کوواریانس $\hat{\beta}_{RGLS}$ برابر است با:

$$(X'\Sigma^{-1}X)^{-1} - (X'\Sigma^{-1}X)^{-1}B'[B(X'\Sigma^{-1}X)^{-1}B']^{-1}B(X'\Sigma^{-1}X)^{-1} \quad (12-2)$$

اثبات: مشابه اثبات قضایای ۱-۲ و ۳-۲.

۳-۲ مدل رگرسیون خطی چندمتغیره

در این بخش، به عنوان تعمیمی بر نتایج بخش قبل مدل رگرسیون خطی چندمتغیره را معرفی کرده، و برخی از خواص آن را بررسی می‌کنیم.

مدل رگرسیون خطی چندمتغیره به صورت

$$Y = XB + E,$$

می‌باشد، که در آن Y یک ماتریس تصادفی $n \times m$ از پاسخ‌ها، X یک ماتریس $n \times k$ غیرتصادفی و معلوم با رتبه $k < n$ و $B = (\beta_1, \dots, \beta_m)$ ماتریس $k \times m$ پارامترهای مدل و

$$E \sim_{n \times m}(0, \Sigma \otimes I)$$

نوشت، که در آن Y_i ها، $i = (1, \dots, m)$ بردارهای تصادفی n بعدی، با میانگین $X\beta_i$ و

$$cov(Y_i, Y_j) = \sigma_{ij}I_n \text{ هستند.}$$

به طور خلاصه، مدل رگرسیون خطی چندمتغیره را می‌توان به صورت زیر نیز نوشت:

$$Y \sim_{n \times m} (XB, \Sigma \otimes I_m) \quad (13-2)$$

در ادامه با معلوم فرض کردن ماتریس Σ ، قصد داریم با اعمال قیدهای مختلف به فضای پارامتر مدل (13-2) برآوردگر RGLS پارامتر ماتریسی B را بیابیم. و سپس با فرض $\Sigma = \sigma^2 V$ برآوردگر σ^2 را به دست می آوریم.

۲-۳-۱ برآورد پارامترها وقتی ماتریس Σ معلوم است

قضیه ۲-۶* (کوباک^{۱۴}، ۲۰۰۷) مدل رگرسیونی (13-2) را در نظر بگیرید. همچنین فرض کنید قیدی خطی به صورت $G_{q \times k} B_{k \times m} H_{m \times r} + D_{q \times r} = 0_{q \times r}$ به فضای پارامتر مدل تحمیل شده است، که در آن $r(H_{m \times r}) = r < m$ و $r(G_{q \times k}) = q < k$ آنگاه

الف- برآوردگر RGLS ماتریس B عبارتست از:

$$\hat{\hat{B}} = \hat{B} - (X'X)^{-1} G' [G(X'X)^{-1} G']^{-1} (G \hat{B} H + D) (H' \Sigma H)^{-1} H' \Sigma, \quad (14-2)$$

که در آن $\hat{B} = (X'X)^{-1} X'Y$.

ب- ماتریس کوواریانس $\hat{\hat{B}}$ برابر است با

$$Var[vec(\hat{\hat{B}})] = \Sigma \otimes (X'X)^{-1}$$

$$-[\Sigma H(H'\Sigma H)^{-1}H'\Sigma] \otimes \{(X'X)^{-1}G'[G(X'X)^{-1}G']^{-1}G(X'X)^{-1}\}. \quad (۱۵-۲)$$

اثبات: در قضیه ۵-۲ دیدیم که برای $Y \sim_n (X\beta, \Sigma)$ با اعمال قید $B_{q,k}\beta_{k,1} + b_{q,1} = 0$ برآوردگر RGLS برای β به صورت زیر است:

$$\hat{\beta} = \hat{\beta} - (X'\Sigma^{-1}X)^{-1}B'[B(X'\Sigma^{-1}X)^{-1}B']^{-1}(B\hat{\beta} + b). \quad (۱۶-۲)$$

همچنین

$$Var(\hat{\beta}) = (X'\Sigma^{-1}X)^{-1} - (X'\Sigma^{-1}X)^{-1}B'[B(X'\Sigma^{-1}X)^{-1}B']^{-1}B(X'\Sigma^{-1}X)^{-1}, \quad (۱۷-۲)$$

$$\hat{\beta} = (X'\Sigma^{-1}X)^{-1}X'\Sigma^{-1}Y \text{ که در آن}$$

از طرفی با گرفتن vec از مدل (۱۳-۲) و قید اعمال شده، می‌توان نتیجه گرفت:

$$vec(Y) \sim_{nm} [(I \otimes X)vec(B), \Sigma \otimes I],$$

$$vec[G_{q \times k}B_{k \times m}H_{m \times r} + D_{q \times r} = 0_{q \times r}] = (H' \otimes G)vec(B) + vec(D) = 0.$$

بنابراین فقط کافیهست در رابطه (۱۶-۲) قرار دهیم:

$$X = (I \otimes X),$$

$$B = (H' \otimes G),$$

$$\Sigma = (\Sigma \otimes I),$$

$$Y = vec(Y),$$

$$\beta = vec(B),$$

$$b = vec(D)0$$

با اعمال این تغییرات در رابطه (۱۶-۲) خواهیم داشت:

$$\begin{aligned} vec(\hat{\hat{B}}) &= vec(\hat{B}) - [(I \otimes X)'(\Sigma \otimes I)^{-1}(I \otimes X)]^{-1} \\ &\quad \times (H' \otimes G)' \left[(H' \otimes G)((I \otimes X)'(\Sigma \otimes I)^{-1}(I \otimes X))^{-1}(H' \otimes G)' \right]^{-1} \\ &\quad \times [(H' \otimes G)vec(\hat{B}) + vec(D)] \end{aligned} \quad (۱۸-۲)$$

و از طرفی داریم:

$$\begin{aligned} [(I \otimes X)'(\Sigma \otimes I)^{-1}(I \otimes X)]^{-1} &= [(I \otimes X')(\Sigma^{-1} \otimes I)(I \otimes X)]^{-1} \\ &= [\Sigma^{-1} \otimes X'X]^{-1} = \Sigma \otimes (X'X)^{-1}, \end{aligned} \quad (۱۹-۲)$$

$$\begin{aligned} vec(\hat{B}) &= [(I \otimes X)'(\Sigma \otimes I)^{-1}(I \otimes X)]^{-1}(I \otimes X)'(\Sigma \otimes I)^{-1}vec(Y) \\ &= vec((X'X)^{-1}X'Y), \end{aligned} \quad (۲۰-۲)$$

$$\begin{aligned} (H' \otimes G)' \left[(H' \otimes G)((I \otimes X)'(\Sigma \otimes I)^{-1}(I \otimes X))^{-1}(H' \otimes G)' \right]^{-1} \\ = [H(H'\Sigma H)^{-1}] \otimes [G'(G(X'X)^{-1}G')^{-1}], \end{aligned} \quad (۲۱-۲)$$

$$\begin{aligned} [(H' \otimes G)vec(\hat{B}) + vec(D)] &= vec(G\hat{B}H) + vec(D) \\ &= vec(G\hat{B}H + D) \end{aligned} \quad (۲۲-۲)$$

در پایان با جایگذاری روابط (۱۹-۲) تا (۲۲-۲) در رابطه (۱۸-۲) و انجام محاسبات جبری، نتیجه مورد نظر حاصل می‌شود.

□

به طور مشابه قسمت ب اثبات می شود.

نتیجه ۱-۲ در مدل رگرسیونی (۱۳-۲)، اگر قید را به صورت $G_{q \times k} B_{k \times m} + D_{q \times m} = 0_{q \times m}$ داشته باشیم، آنگاه

$$\hat{\hat{B}} = \hat{B} - (X'X)^{-1}G'[G(X'X)^{-1}G']^{-1}(G\hat{B} + D),$$

$$\begin{aligned} Var[vec(\hat{\hat{B}})] &= \Sigma \otimes \{(X'X)^{-1} - (X'X)^{-1}G'[G(X'X)^{-1}G']^{-1}G(X'X)^{-1}\} \\ &= \Sigma \otimes [M_{G'}X'XM_{G'}]^+, \end{aligned}$$

که در آن $P_{G'} = G'(GG')^{-1}G$ و $M_{G'} = I - P_{G'}$

نتیجه ۲-۲ در مدل رگرسیونی (۱۳-۲)، اگر قید را به صورت $B_{k \times m}H_{m \times r} + D_{k \times r} = 0_{k \times r}$ داشته باشیم، در این صورت

$$\hat{\hat{B}} = \hat{B} - (\hat{B}H + D)(H'\Sigma H)^{-1}H'\Sigma,$$

$$\begin{aligned} Var[vec(\hat{\hat{B}})] &= [\Sigma - \Sigma H(H'\Sigma H)^{-1}H'\Sigma] \otimes (X'X)^{-1} \\ &= (M_H\Sigma^{-1}M_H)^+ \otimes (X'X)^{-1}. \end{aligned}$$

قضیه ۷-۲* (کوباک، ۲۰۰۷). در مدل (۱۳-۲) فرض کنید قیدها به صورت

$$trace(G_1 B) + g_1 = 0, \dots, trace(G_q B) + g_q = 0,$$

هستند، و ماتریس $\tilde{G} = (vec(G'_1), \dots, vec(G'_q))'$ در سطرها تمام رتبه است، یعنی $r(\tilde{G}) = q$. $q < km$ آنگاه

$$\begin{aligned} vec(\hat{\hat{B}}) &= vec(\hat{B}) \\ &\quad - [\Sigma \otimes (X'X)^{-1}] \tilde{G}' \{ \tilde{G} [\Sigma \otimes (X'X)^{-1}] \tilde{G}' \}^{-1} [\tilde{G} vec(\hat{B}) + g], \end{aligned}$$

$$Var \left[vec \left(\hat{\hat{B}} \right) \right] = \{ M_{\tilde{G}} [\Sigma^{-1} \otimes (X'X)] M_{\tilde{G}} \}^+.$$

اثبات: با توجه به اینکه $trace(AB) = vec'(A')vec(B)$ می‌توان قیدها را به صورت زیر نوشت:

$$trace(G_1 B) + g_1 = 0 \quad \Rightarrow \quad vec'(G'_1)vec(B) + g_1 = 0$$

$$trace(G_2 B) + g_2 = 0 \quad \Rightarrow \quad vec'(G'_2)vec(B) + g_2 = 0$$

$$\vdots \quad \quad \quad \vdots$$

$$trace(G_q B) + g_q = 0 \quad \Rightarrow \quad vec'(G'_q)vec(B) + g_q = 0$$

حال با توجه به فرض قضیه، می‌توان قید را به صورت زیر بازنویسی کرد:

$$\tilde{G}_{q \times kq} vec(B)_{kq \times 1} + g_{q \times 1} = 0_{q \times 1}.$$

بنابراین کافیت در اثبات قضیه ۲-۶ قرار دهیم،

$$b_{q \times 1} = g_{q \times 1}$$

$$B = \tilde{G}$$

$$\beta = vec(B)$$

$$X = (I \otimes X)$$

$$\Sigma = \Sigma \otimes I$$

□

۲-۳-۲ برآورد پارامترها در حالتی که $\Sigma = \sigma^2 V$

قضیه ۲-۸* (کوباک، ۲۰۰۷). تحت مفروضات قضیه (۲-۶) و با فرض $\Sigma = \sigma^2 V$ ، برآوردگر σ^2

عبارتست از:

$$\hat{\sigma}^2 = trace(v_I' v_I V^{-1}) / [m(n - k) + qr],$$

که در آن

$$v_I = Y - X\hat{\hat{B}} = v + k_I, \quad v = M_X Y,$$

$$k_I = X(X'X)^{-1}G'[G(X'X)^{-1}G']^{-1}(G\hat{B}H + D)(H'VH)^{-1}H'V,$$

$$\hat{B} = (X'X)^{-1}X'Y.$$

اثبات: در قضیه ۲-۴ دیدیم که در مدل $Y \sim_n (X\beta, \sigma^2 V)$ با اعمال قید $B_{q \times k}\beta_{k \times 1} + b_{q \times 1} = 0$ برآوردگر σ^2 به صورت زیر است

$$\hat{\sigma}^2 = v_I' V^{-1} v_I / (n + q - k),$$

که در آن

$$v_I = Y - X\hat{\beta} = Y - X\hat{\beta} + X(X'V^{-1}X)^{-1}B'[B(X'V^{-1}X)^{-1}B']^{-1}(B\hat{\beta} + b),$$

$$\hat{\beta} = (X'V^{-1}X)^{-1}X'V^{-1}Y.$$

□ لذا با جایگذاری $n = mn$ و $k = mk$ و $q = qr$ در اثبات قضیه ۲-۴، اثبات کامل می‌شود.

قضیه ۲-۹ (کوباکک، ۲۰۰۷) فرض کنید V, V_I و k_I ماتریس‌های تعریف شده در قضیه (۲-۸) باشند.

اگر Y دارای توزیع نرمال باشد، آنگاه

$$k_I' k_I \sim W_m(q, \sigma^2 V H (H' V H)^{-1} H' V)$$

اثبات: با استفاده از قضیه ۱-۹ و اینکه Y دارای توزیع نرمال است، می‌توان نوشت

$$(G\hat{B}H + D) \sim N_{q \times r}\{0, (H'\Sigma H) \otimes [G(X'X)^{-1}G']\}.$$

بنابراین از (۲-۲۳) خواهیم داشت

$$(G\hat{B}H + D)' [G(X'X)^{-1}G']^{-1} (G\hat{B}H + D) \sim W_r(q, (H'\Sigma H)).$$

از طرفی

$$k'_I k_I = VH(H'VH)^{-1}(G\hat{B}H + D)'[G(X'X)^{-1}G']^{-1}(G\hat{B}H + D)(H'VH)^{-1}H'V$$

در نتیجه

$$k'_I k_I \sim W_m(q, \sigma^2 VH(H'VH)^{-1}H'V)$$

□

و اثبات کامل است.

فصل سوم

بر آورد انقباضی

۳-۱ مقدمه

در این فصل، ابتدا به معرفی برآوردگرهای انقباضی پرداخته و سپس با فرض این که ماتریس تصادفی X دارای توزیع نرمال ماتریسی است، کلاسی از برآوردگرهای انقباضی را، برای برآورد پارامتر ماتریسی میانگین، با یک عامل انقباض کلی تشکیل داده و به عنوان معیاری برای مقایسه برآوردگرها، مخاطره آنها را به دست می آوریم. در ادامه، از این مخاطره استفاده کرده و با جایگزین کردن این عامل با عوامل انقباض استاین و استاین مثبت، مخاطره آنها را نیز محاسبه می کنیم. در پایان به عنوان کاربردی از مطالب ذکر شده مدل منحنی رشد را معرفی می کنیم و برآوردگرهای انقباضی را برای آن نوشته و مخاطره آنها را نیز به دست می آوریم.

۳-۲ برآوردگر انقباضی

در آمار، برآوردگر انقباضی، برآوردگری است که به طور واضح یا مجازی در اثر انقباض به دست آید. به عبارت دیگر برآوردگر انقباضی، برآوردگری است که از ترکیب اطلاعات نمونه ای یعنی برآوردگر خام یا اولیه مانند MLE، برآوردگر کمترین توان های دوم یا بیزی با اطلاعات غیرنمونه ای یعنی یک سری محدودیت، به صورت برآوردگری جدید بهبود می یابد. این بهبودی در جهت مقادیری است که از اطلاعات دیگر به دست آمده نه در جهت برآوردگر اولیه.

اطلاعات غیرنمونه ای (محدودیت) می توانند از روش های زیر به دست آیند:

۱. اطلاعاتی که با استفاده از تحقیق های نظری یا آزمایشی درستی و صحت آنها معلوم شده است.

۲. فرض هایی که قصد آزمون آنها را داریم.

۳. این اطلاعات می‌توانند یک شرط ساختگی باشند که به منظور حذف بخش‌های زائد (متغیرها یا

حتی پارامترها)، در تشریح مدل، بر مدل تحمیل می‌شوند.

لازم به ذکر است که اطلاعات غیرنمونه‌ای یا محدودیت‌های اعمال شده بر روی مدل در قالب یک

برآوردگر (برآوردگر محدودشده) به کار گرفته می‌شود.

در حالت کلی، بهبود برآوردگرها را می‌توان با معیار توان دوم خطا (MSE) اندازه‌گیری کرد. در

این صورت انقباض می‌تواند به سمت صفر یا یک مقدار ثابت باشد. اثر این انقباض معمولاً به صورت

تبدیل یک برآوردگر ناریب به اریب ولی با MSE کمتر است. در عین حال می‌توان برآوردگرهای انقباضی

از نوع استاین و ناریب را نیز یافت. برای آگاهی بیشتر در این خصوص به نوروزی‌راد (۱۳۹۰) مراجعه

کنید.

شکل کلی برآوردگر انقباضی به صورت $X + g(X)$ می‌باشد. پس از حل نامعادله

$$R(X + g(X), \theta) \leq R(X, \theta); \quad \forall \theta,$$

به کمک اتحاد استاین (استاین، ۱۹۸۱) تابع $g(X)$ را در هر یک از حالت‌های زیر به دست می‌آوریم:

(۱) انقباض به سمت صفر

$$(1 - c)X = X - cX$$

(۲) انقباض به سمت یک عدد ثابت m .

$$m + c(X - m) = (1 - c)m + cX$$

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

مفهوم برآوردگر انقباضی را با استفاده از، مثالی از صالح^۱ (۲۰۰۶) بیشتر روشن می‌کنیم.

مسابقه بیس‌بال را در نظر بگیرید. این ورزش گروهی با چوب و توپ توسط ۹ بازیکن انجام می‌شود.

به عنوان مثال در یک بازی بیس‌بال سه بازیکن یا بیشتر را در نظر بگیرید. علاقه‌مندیم میانگین بتینگ^۲ را برای هر یک از بازیکنان پیش‌بینی کنیم. این میانگین در آمار بیس‌بال برای شناخت عملکرد بازیکنان بسیار موثر است. آماردان‌هایی که از روش انقباضی استفاده می‌کنند، انتظار دارند تا ضربه‌های آینده بازیکنان را دقیق‌تر از مقادیر میانگین بتینگ جداگانه هر بازیکن، پیش‌بینی کنند. روش استاین را بر پایه‌ی داده‌های بیس‌بال گزارش‌شده در افرون و موریس^۳ در زیر نشان می‌دهیم:

جدول (۱-۳) مقادیر میانگین بتینگ از ۴۵ ضربه‌ی اول ۱۸ بازیکن را نشان می‌دهد:

جدول (۱-۳) میانگین بتینگ ۱۸ بازیکن بیس‌بال و پیش‌بینی به روش انقباضی

$\hat{q}_i$	میانگین بتینگ بعد از ۴۵ ضربه‌ی اول ($\bar{X}_i$)	پیش‌بینی میانگین به روش انقباضی ($\bar{X}_i^{JS}$)	$\hat{q}_i$	میانگین بتینگ بعد از ۴۵ ضربه‌ی اول ($\bar{X}_i$)	پیش‌بینی میانگین به روش انقباضی ($\bar{X}_i^{JS}$)
۱	۰/۴	۰/۲۹۴	۱۰	۰/۲۴۴	۰/۲۶۱
۲	۰/۴۷۸	۰/۲۸۹	۱۱	۰/۲۲۲	۰/۲۵۶
۳	۰/۳۵۶	۰/۲۴۵	۱۲	۰/۲۲۲	۰/۲۵۶
۴	۰/۳۳۳	۰/۲۸	۱۳	۰/۲۲۲	۰/۲۵۶
۵	۰/۳۱۱	۰/۲۷۵	۱۴	۰/۲۲۲	۰/۲۵۶
۶	۰/۳۱۱	۰/۲۷۵	۱۵	۰/۲۲۲	۰/۲۵۶
۷	۰/۲۸۹	۰/۲۷	۱۶	۰/۲	۰/۲۵۲
۸	۰/۲۶۷	۰/۲۶۶	۱۷	۰/۱۷۸	۰/۲۴۱
۹	۰/۲۴۴	۰/۲۶۱	۱۸	۰/۱۵۶	۰/۲۴۳

¹ Saleh

² Batting average

³ Efron and Morris

نحوه محاسبه پیش‌بینی در روش استاین به این صورت است که در ابتدا باید میانگین تمام میانگین‌ها، $\bar{X}_T$ ، محاسبه شود. در این مثال داریم:

$$\bar{X}_T = \frac{\bar{X}_1 + \bar{X}_2 + \dots + \bar{X}_{18}}{18} = 0/2654$$

هدف، انقباض ۱۸ میانگین به سمت میانگین کل است که با توجه به فرضیات (حالت دوم؛ انقباض به سمت عدد ثابت m که در این مثال $\bar{X}_T$ است)، منطقی به نظر می‌رسد.

مرحله بعد یافتن توان دوم فاصله‌ی وزنی بین ۱۸ میانگین جدول و میانگین کل است، که در این مثال 19/045 می‌شود. به دنبال این مرحله، ثابت انقباضی را که در صالح (۲۰۰۶) به صورت زیر معرفی شده، به دست می‌آوریم:

$$c = \left(1 - \frac{2 - \text{درجه آزادی}}{\text{توان دوم فاصله وزنی}} \right) = \left(1 - \frac{17 - 15}{19/045} \right) = 0/212$$

ثابت انقباضی را از آن جهت ثابت گویند که برای هر ۱۸ بازیکن ثابت است و در محاسبه آن باید این نکته را مد نظر داشت که تنها میانگین تعداد ضربات اصابت شده به توپ ۱۷ بازیکن دیگر بر بازیکن مد نظر دخالت دارد و لذا درجه‌ی آزادی برابر عدد ۱۷ خواهد بود. پیش‌بینی میانگین به روش انقباضی جیمز-استاین از رابطه زیر به دست می‌آید:

$$\bar{X}_i^{JS} = \bar{X}_T + c(\bar{X}_i - \bar{X}_T); \quad i = 1, 2, \dots, 18$$

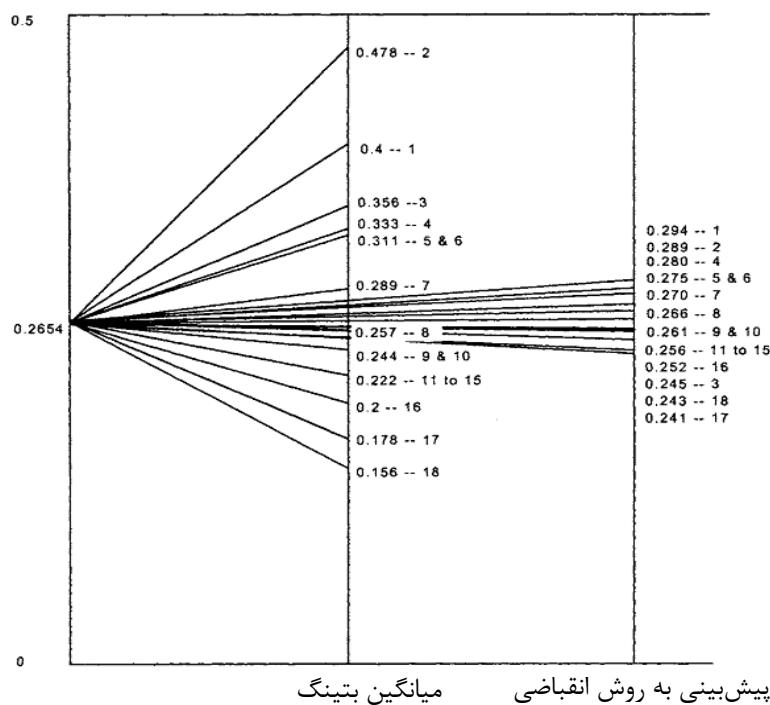
به عنوان مثال برای بازیکن‌های اول و هجدهم به ترتیب داریم:

$$\bar{X}_1^{JS} = \bar{X}_T + c(\bar{X}_1 - \bar{X}_T) = 0/265 + 0/212(0/4 - 0/265) = 0/294$$

$$\bar{X}_{18}^{JS} = \bar{X}_T + c(\bar{X}_{18} - \bar{X}_T) = 0/265 + 0/212(0/156 - 0/265) = 0/243$$

در حالی که میانگین واقعی این دو بازیکن به ترتیب ۰/۳۴۶ و ۰/۲ می باشد.

در شکل (۱-۳) مقادیر میانگین بتینگ و پیش‌بینی استاین (جدول ۱-۳) به صورت خطوط واصل به میانگین کل آمده است. همچنانکه مشاهده می‌شود میانگین‌های به دست آمده به روش انقباضی در مقایسه با میانگین بتینگ معمولی بیشتر در اطراف میانگین کل متمرکز شده‌اند. که نشان دهنده مفهوم انقباض به میانگین کل است. به عبارت دیگر روش استاین برآوردهای اولیه را به سمت میانگین کل هدایت می‌کند.



شکل (۱-۳) نمودار میانگین بتینگ و مقادیر پیش‌بینی شده بر اساس فرمول استاین

۳-۳ برآورد انقباضی در توزیع نرمال ماتریسی

فرض کنید ماتریس تصادفی X در اندازه $q \times k$ دارای توزیع نرمال ماتریسی به صورت زیر باشد

$$X \sim MN_{q \times k}(\theta, \Lambda \otimes I_k), \quad (1-3)$$

که در آن θ در اندازه $q \times k$ ماتریس میانگین توزیع و Λ ماتریس معین مثبت با بعد و رتبه q ، است.

۱-۳-۳ روش ساخت برآوردگر انقباضی

همانطور که گفته شد برای ساخت برآوردگر انقباضی به یک برآوردگر اولیه و یک برآوردگر جدید نیاز است. در این بخش از برآوردگر درستنمایی ماکزیمم پارامتر ماتریسی θ به عنوان برآوردگر اولیه، و از برآوردگر محدودشده آن تحت یک قید خطی، به عنوان یک برآوردگر جدید استفاده می‌کنیم. از فصل قبل به سهولت می‌توان دید که ماتریس X برآوردگر درستنمایی ماکزیمم محدودنشده برای پارامتر ماتریسی θ در (۱-۳) می‌باشد، یعنی:

$$\tilde{\theta} = X. \quad (2-3)$$

حال فرض کنید قیدی خطی به صورت

$$L_1 \theta L_2 = d \quad (3-3)$$

به فضای پارامتر توزیع نرمال ماتریسی (۱-۳)، تحمیل شده است، که در آن L_1 و L_2 ماتریس‌های محدودیت و معلوم به ترتیب با ابعاد $[p \times q, (p < q)]$ و $[k \times m, (m \leq k)]$ و رتبه کامل هستند، و d ماتریس معلوم با بعد $p \times m$ می‌باشد. در لم ۱-۳ نحوه ساخت برآوردگر درستنمایی ماکزیمم θ تحت قید (۳-۳) بررسی می‌شود.

لم ۳-۱* (کرونزیزا، ۲۰۱۲). در توزیع نرمال ماتریسی (۱-۳)، تحت درستی قید (۲-۳)، برآوردگر درست‌نمایی ماکزیمم محدودشده برای پارامتر θ عبارت است از:

الف-

$$\hat{\theta} = X - \Lambda L'_1 (L_1 \Lambda L'_1)^{-1} (L_1 X L_2 - d) (L'_2 L_2)^{-1} L'_2 0 \quad (۴-۳)$$

ب- این برآوردگر تحت فرض $L_1 \theta L_2 = d$ ، برآوردگری نااریب برای θ است.

اثبات: الف- برای به دست آوردن برآوردگر RE از روش ضریب لاگرانژ استفاده می‌کنیم. با توجه به تابع درست‌نمایی

$$L(\theta, \Lambda) = (2\pi)^{-\frac{1}{2}kq} |\Lambda|^{-\frac{k}{2}} \exp \left\{ -\frac{1}{2} \text{trace}[\Lambda^{-1}(X - \theta)(X - \theta)'] \right\},$$

داریم

$$\ln(L(\theta, \Lambda)) = \ell(\theta) \propto -\frac{1}{2} \text{trace}[\Lambda^{-1}(X - \theta)(X - \theta)'].$$

اکنون تابع لاگرانژ را به صورت زیر تشکیل می‌دهیم

$$\lambda(\theta) = \ell(\theta) + \lambda(L_1 \theta L_2 - d),$$

که در آن ضریب لاگرانژ λ برابر است با $vec \alpha$ -. حال کفایت از تابع لاگرانژ نسبت به θ مشتق بگیریم، یعنی:

$$\frac{\partial \lambda(\theta)}{\partial \theta} = \frac{\partial \ell(\theta)}{\partial \theta} + \frac{\partial \lambda(L_1 \theta L_2 - d)}{\partial \theta}$$

برای محاسبه این مشتق از قضایای (A-۲۸) و (A-۲۹) در پیوست A استفاده شده است.

ابتدا فرض می‌کنیم $A = \Lambda^{-1}(X - \theta)(X - \theta)'$ بنابراین

$$\frac{\partial A}{\partial \theta} = -2[(X' \otimes \Lambda^{-1}) - (\theta' \otimes \Lambda^{-1})].$$

در نتیجه

$$\begin{aligned} \frac{\partial \ell(\theta)}{\partial(\theta)} &= -\frac{1}{2} \frac{\partial A}{\partial \theta} \text{vec} I = (X' \otimes \Lambda^{-1}) \text{vec} I - (\theta' \otimes \Lambda^{-1}) \text{vec} I \\ &= \text{vec}(\Lambda^{-1} X) - \text{vec}(\Lambda^{-1} \theta). \end{aligned}$$

همچنین

$$\frac{\partial \lambda(L_1 \theta L_2 - d)}{\partial \theta} = \lambda(L_2 \otimes L_1') = -(L_2 \otimes L_1') \text{vec} \alpha.$$

اکنون با مساوی صفر قرار دادن مقدار مشتق خواهیم داشت:

$$\frac{\partial \lambda(\theta)}{\partial \theta} = \text{vec}(\Lambda^{-1} X) - \text{vec}(\Lambda^{-1} \theta) - \text{vec}(L_1' \alpha L_2') = \text{vec} o.$$

$$\Rightarrow \Lambda^{-1}(X - \theta) - (L_1' \alpha L_2') = o.$$

$$\Rightarrow \Lambda^{-1} \theta = \Lambda^{-1} X - L_1' \alpha L_2'.$$

$$\Rightarrow \theta = X - \Lambda L_1' \alpha L_2'. \quad (۵-۳)$$

با جایگزین کردن (۵-۳)، در رابطه $L_1 \theta L_2 = d$ خواهیم داشت:

$$L_1(X - \Lambda L_1' \alpha L_2')L_2 = d$$

$$\Rightarrow L_1 X L_2 - L_1 \Lambda L_1' \alpha L_2' L_2 = d$$

$$\Rightarrow \alpha = (L_1 \Lambda L_1')^{-1} (L_1 X L_2 - d) (L_2' L_2)^{-1} \quad (۶-۳)$$

و در نهایت با جایگذاری رابطه (۳-۶) در (۳-۵)، نتیجه حاصل می شود.

ب- با گرفتن امید ریاضی از $\hat{\theta}$ داریم

$$E(\hat{\theta}) = E(X) - \Lambda L_1' (L_1 \Lambda L_1')^{-1} (L_1 E X L_2 - d) (L_2' L_2)^{-1} L_2' = \theta,$$

و اثبات کامل است. $\square$

بنابراین می توان کلاسی از برآوردگرهای انقباضی را به صورت زیر تشکیل داد:

$$\hat{\theta}_S = \hat{\theta} + h(\| (X - \hat{\theta}) L_2 \|_{\Xi_1, \Xi_2}^2) (X - \hat{\theta}) \quad (۳-۷)$$

که در آن $\Xi_1 = L_1' (L_1 \Lambda L_1')^{-1} L_1$ و $\Xi_2 = (L_2' L_2)^{-1}$ و $\tilde{\theta}$ و نیز به ترتیب در رابطه های (۳-۲) و (۳-۴) داده شده اند.

در رابطه (۳-۷)، برآوردگر انقباضی $\hat{\theta}_S$ با یک عامل انقباض کلی $(h(\|\cdot\|))$ بیان شده است. با جایگزین کردن آن با عوامل انقباض دیگری می توان برآوردگرهای جدیدی ساخت. البته آن چه مهم است، داشتن برآوردگری با مخاطره کمتر می باشد. در زیر نمونه هایی خاص ذکر گردیده است.

مثال ۳-۱ (صورت های دقیقی از برآوردگر انقباضی)

$$(۱) \text{ با قراردادن } h(x) = 0 \text{ داریم } \hat{\theta}_S = \hat{\theta}$$

$$(۲) \text{ با قراردادن } h(x) = 1 \text{ نتیجه می گیریم } \hat{\theta}_S = \tilde{\theta} = X$$

$$(۳) \text{ با قراردادن } h(x) = 1 - \frac{pq-2}{x}, \text{ به برآوردگر انقباضی نوع استاین به صورت زیر می رسیم:}$$

$$\hat{\theta}_s = \hat{\theta}^s = \hat{\theta} + \left(1 - \frac{pq-2}{\psi}\right)(X - \hat{\theta}),$$

$$\psi = \text{trace} \left((X - \hat{\theta})' L_1' (L_1 \Lambda L_1')^{-1} L_1 (X - \hat{\theta}) \right)$$
 که در آن

$$h(x) = \left(1 - \frac{pq-2}{x}\right) I(x > pq-2)$$
 با قراردادن (۴)

صورت زیر نتیجه می گیریم:

$$\hat{\theta}_s = \hat{\theta}^{s+} = \hat{\theta} + \left(1 - \frac{pq-2}{\psi}\right) I(\psi > pq-2)(X - \hat{\theta})$$

در ادامه به محاسبه توابع مخاطره می پردازیم. برای این منظور ابتدا لم ها و قضایای اساسی را، که ابزاری لازم برای محاسبه توابع مخاطره می باشند، بیان و اثبات می کنیم.

۲-۳-۳ لم ها و قضایای اساسی

لم ۲-۳* (کرونزیزا، ۲۰۱۲). فرض کنید $\Sigma^{-1} = \Lambda$ ، $\Sigma^* = \Lambda L_1' (L_1 \Lambda L_1')^{-1} L_1 \Lambda$ و

$$\delta = \Lambda L_1' (L_1 \Lambda L_1')^{-1} (L_1 \theta L_2 - d)$$
 حال اگر $X \sim N_{q \times k}(\theta, \Lambda \otimes I_k)$ در این صورت

$$\xi = (X - \hat{\theta}) L_2 \sim N_{q \times m}(\delta, \Sigma^* \otimes L_2' L_2), \quad (۸-۳)$$

$$\rho = (X - \theta) L_2 \sim N_{q \times m}(0, \Sigma^{-1} \otimes L_2' L_2), \quad (۹-۳)$$

$$\eta = (\hat{\theta} - \theta) L_2 \sim N_{q \times m}(-\delta, (\Sigma^{-1} - \Sigma^*) \otimes L_2' L_2, \quad (۱۰-۳)$$

که در آن‌ها $\hat{\theta}$ در رابطه (۳-۴) داده شده است.

اثبات: با توجه به رابطه (۳-۴) داریم

$$\xi = \Lambda L_1' (L_1 \Lambda L_1')^{-1} (L_1 X L_2 - d), \quad (۱۱-۳)$$

$$\eta = X L_2 - \Lambda L_1' (L_1 \Lambda L_1')^{-1} (L_1 X L_2 - d) - \theta L_2, \quad (۱۲-۳)$$

$$\rho = X L_2 - \theta L_2, \quad (۱۳-۳)$$

حال با توجه به این که $EX = \theta$ ، می‌توان نتیجه گرفت

$$\begin{aligned} E(\xi) &= \Lambda L_1' (L_1 \Lambda L_1')^{-1} (L_1 EX L_2 - d) \\ &= \Lambda L_1' (L_1 \Lambda L_1')^{-1} (L_1 \theta L_2 - d), \\ E(\eta) &= EX L_2 - \Lambda L_1' (L_1 \Lambda L_1')^{-1} (L_1 EX L_2 - d) - \theta L_2 \\ &= -\Lambda L_1' (L_1 \Lambda L_1')^{-1} (L_1 \theta L_2 - d), \\ E(\rho) &= EX L_2 - \theta L_2 = 0, \end{aligned}$$

و بنابر لم ۱-۱ خواهیم داشت:

$$\begin{aligned} cov(\xi) &= (\Lambda L_1' (L_1 \Lambda L_1')^{-1} L_1 \Lambda L_1' (L_1 \Lambda L_1')^{-1} L_1 \Lambda) \otimes (L_2' L_2) \\ &= \Lambda L_1' (L_1 \Lambda L_1')^{-1} L_1 \Lambda \otimes L_2' L_2. \end{aligned}$$

از آنجایی که

$$\begin{aligned} \eta &= X L_2 - \Lambda L_1' (L_1 \Lambda L_1')^{-1} L_1 X L_2 + \Lambda L_1' (L_1 \Lambda L_1')^{-1} d - \theta L_2 \\ &= [I - \Lambda L_1' (L_1 \Lambda L_1')^{-1} L_1] X L_2 + \Lambda L_1' (L_1 \Lambda L_1')^{-1} d - \theta L_2, \end{aligned}$$

می توان نوشت

$$\begin{aligned} cov(\eta) &= ([I - \Lambda L'_1 (L_1 \Lambda L'_1)^{-1} L_1] \Lambda [I - L'_1 (L_1 \Lambda L'_1)^{-1} L_1 \Lambda]) \otimes L'_2 L_2 \\ &= (\Lambda - \Lambda L'_1 (L_1 \Lambda L'_1)^{-1} L_1 \Lambda - \Lambda L'_1 (L_1 \Lambda L'_1)^{-1} L_1 \Lambda \\ &\quad + \Lambda L'_1 (L_1 \Lambda L'_1)^{-1} L_1 \Lambda L'_1 (L_1 \Lambda L'_1)^{-1} L_1 \Lambda) \otimes L'_2 L_2 \\ &= (\Lambda - \Lambda L'_1 (L_1 \Lambda L'_1)^{-1} L_1 \Lambda) \otimes L'_2 L_2 \end{aligned}$$

در نهایت

$$cov(\rho) = \Lambda \otimes L'_2 L_2$$

□

و اثبات کامل است.

قضیه ۳-۱ (کرونزیزا، ۲۰۱۲). فرض کنید $X \sim MN_{k \times q}(M, Y \otimes \Lambda_1)$ که در آن Λ_1 ماتریس

معین مثبت و Y ماتریس معین نامنفی با رتبه $p \leq k$ است. همچنین Ξ ماتریس معین مثبت و متقارن،

با دو شرط زیر می باشد:

الف- $Y\Xi$ خودتوان است.

ب- $\Xi Y \Xi M = \Xi M$.

در این صورت،

$$E(h(\text{trace}(\Lambda_1^{-1} X' \Xi Y \Xi X)) X) = E\left(h\left(\chi_{pq+2}^2(\text{trace}(\Lambda_1^{-1} M' \Xi Y \Xi M))\right)\right) M.$$

(۱۴-۳)

اثبات: طبق فرض قضیه $P_{k \times k} = Y\Xi$ ماتریس خودتوان است. لذا

$$P^2 = P \implies Y\Xi Y \Xi = Y\Xi.$$

از آنجایی که Ξ معین مثبت است لذا نامنفرد می باشد. حال با ضرب طرفین رابطه بالا در Ξ^{-1} داریم:

$$\Upsilon \Xi \Upsilon \Xi \Xi^{-1} = \Upsilon \Xi \Xi^{-1},$$

$$\Upsilon \Xi \Upsilon = \Upsilon.$$

در این قسمت فرض کنید $R = \Xi^{1/2} \Upsilon \Xi^{1/2}$. ماتریس مربعی R با بعد k ، متقارن است و علاوه بر این

$$R^2 = \Xi^{1/2} \Upsilon \Xi^{1/2} \Xi^{1/2} \Upsilon \Xi^{1/2} = \Xi^{1/2} \Upsilon \Xi \Upsilon \Xi^{1/2} = \Xi^{1/2} \Upsilon \Xi^{1/2} = R,$$

که در نتیجه R خودتوان می‌باشد. لذا با استفاده از قضیه (A-۲۳) ماتریس متعامدی مانند Q وجود دارد به طوری که

$$QRQ' = \begin{bmatrix} I_p & 0 \\ 0 & 0_{k-p} \end{bmatrix}_{k \times k}. \quad (3-15)$$

همچنین فرض کنید $V = Q\Xi^{1/2}X\Lambda_1^{-1/2}$. بنابراین طبق لم ۱-۱

$$V \sim MN_{k \times q}(\mu, (QRQ') \otimes I_q),$$

که در آن $\mu = Q\Xi^{1/2}M\Lambda_1^{-1/2}$ و

$$vec(V) = ((Q\Xi^{1/2}) \otimes \Lambda_1^{-1/2}) vec(X). \quad (3-16)$$

در این قسمت با توجه به رابطه (3-۱۴) می‌توان نوشت:

$$V \sim MN_{k \times q}(\mu, \begin{bmatrix} I_p & 0 \\ 0 & 0 \end{bmatrix} \otimes I_q).$$

حال $vec(V)$ را به صورت زیر افراز می‌کنیم:

$$vec(V) = \begin{bmatrix} V_1 \\ V_2 \end{bmatrix} \sim MN_{kq \times 1} \left(\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{bmatrix} I_p & 0 \\ 0 & 0 \end{bmatrix} \otimes I_q \right),$$

که در آن بردارهای V_1 و V_2 به ترتیب دارای pq و $(k-p)q$ مولفه‌اند، و این افراز را طوری در نظر می‌گیریم که $V_2 = 0$ در نتیجه $\mu_2 = 0$ بنابراین،

$$\begin{aligned} E(V_1) &= \mu_1 = [I_{pq}, 0][Evec(V)] \\ &= ([I_p, 0] \otimes I_q)(Q\Xi^{1/2} \otimes \Lambda_1^{-1/2})vec(M). \end{aligned} \quad (۱۷-۳)$$

حال با استفاده از (۱۳-A) داریم:

$$\begin{aligned} trace(\Lambda_1^{-1}X'\Xi Y \Xi X) &= trace(\Lambda_1^{-1/2}X'\Xi^{1/2}Q'Q\Xi^{1/2}Y\Xi^{1/2}X\Lambda_1^{-1/2}) \\ &= trace(V'QRQ'V) \\ &= vec'(V)((QRQ') \otimes I_q)vec(V) \\ &= [V_1' \ V_2'] \left[\begin{pmatrix} I_p & 0 \\ 0 & 0 \end{pmatrix} \otimes I_q \right] \begin{bmatrix} V_1 \\ V_2 \end{bmatrix} = V_1'V_1. \end{aligned} \quad (۱۸-۳)$$

از رابطه های (۱۶-۳) تا (۱۸-۳)، می‌توانیم نتیجه بگیریم که

$$\begin{aligned} vec[h(trace(\Lambda_1^{-1}X'\Xi Y \Xi X))X] &= vec[h(V_1'V_1)X] \\ &= (h(V_1'V_1) \otimes I_q)vecX \\ &= (h(V_1'V_1) \otimes I_q)(\Xi^{-1/2}Q' \otimes \Lambda_1^{1/2})vec(V) \\ &= h(V_1'V_1)(\Xi^{-1/2}Q' \otimes \Lambda_1^{1/2}) \begin{pmatrix} V_1 \\ V_2 \end{pmatrix} \\ &= (\Xi^{-1/2}Q' \otimes \Lambda_1^{1/2})h(V_1'V_1) \begin{pmatrix} V_1 \\ V_2 \end{pmatrix} \\ &= (\Xi^{-1/2}Q' \otimes \Lambda_1^{1/2}) \begin{pmatrix} h(V_1'V_1)V_1 \\ h(V_1'V_1)V_2 \end{pmatrix}. \end{aligned}$$

بنابراین

$$\begin{aligned} \text{vec} \left[E \left(h(\text{trace}(\Lambda_1^{-1} X' \Xi Y \Xi X)) \right) X \right] &= E \left[\text{vec} \left(h(\text{trace}(\Lambda_1^{-1} X' \Xi Y \Xi X)) \right) X \right] \\ &= \left(\Xi^{-1/2} Q' \otimes \Lambda_1^{1/2} \right) \begin{pmatrix} E(h(V_1' V_1) V_1) \\ E(h(V_1' V_1) V_2) \end{pmatrix}. \end{aligned} \quad (۱۹-۳)$$

و از آنجایی که $V_2 = 0$ ، لذا

$$E[h(V_1' V_1) V_2] = 0. \quad (۲۰-۳)$$

و همچنین بنا بر لم ۵-۱

$$E[h(V_1' V_1) V_1] = \mu_1 E \left[h \left(\chi_{pq+2}^2 (\mu_1' \mu_1) \right) \right], \quad (۲۱-۳)$$

که در آن با توجه به رابطه (۱۷-۳)

$$\begin{aligned} \mu_1' \mu_1 &= \text{vec}'(M) \left(\Xi^{1/2} Q' \otimes \Lambda_1^{-1/2} \right) \left([I_p, 0]' \otimes I_q \right) \\ &\quad \times \left([I_p, 0] \otimes I_q \right) \left(Q \Xi^{1/2} \otimes \Lambda_1^{-1/2} \right) \text{vec}(M) \\ &= \text{vec}'(M) \left(\Xi^{1/2} Q' \otimes \Lambda_1^{-1/2} \right) \left(\begin{bmatrix} I_p & 0 \\ 0 & 0 \end{bmatrix} \otimes I_q \right) \\ &\quad \times \left(Q \Xi^{1/2} \otimes \Lambda_1^{-1/2} \right) \text{vec}(M) \\ &= \text{vec}'(M) \left(\Xi^{1/2} Q' \otimes \Lambda_1^{-1/2} \right) (Q \Xi^{1/2} Y \Xi^{1/2} Q' \otimes I_q) \\ &\quad \times \left(Q \Xi^{1/2} \otimes \Lambda_1^{-1/2} \right) \text{vec}(M) \\ &= \text{vec}'(M) \left(\Xi^{1/2} Q' Q \Xi^{1/2} Y \Xi^{1/2} Q' Q \Xi^{1/2} \otimes \Lambda_1^{-1/2} \Lambda_1^{-1/2} \right) \text{vec}(M) \\ &= \text{vec}'(M) (\Xi Y \Xi \otimes \Lambda_1^{-1}) \text{vec}(M) \end{aligned}$$

$$\begin{aligned}
 &= \text{trace}(M' \Xi Y \Xi M \Lambda_1^{-1}) \\
 &= \text{trace}(\Lambda_1^{-1} M' \Xi Y \Xi M). \tag{۲۲-۳}
 \end{aligned}$$

در نهایت با ترکیب رابطه‌های (۱۹-۳) تا (۲۱-۳)، داریم:

$$\begin{aligned}
 &\text{vec}(E[h(\text{trace}(\Lambda_1^{-1} X' \Xi Y \Xi X))X]) \\
 &= E\left(h\left(\chi_{pq+2}^2(\mu'_1 \mu_1)\right)\right) \left(\Xi^{-1/2} Q' \otimes \Lambda_1^{1/2}\right) \\
 &\quad \times \left(\begin{bmatrix} I_p & 0 \\ 0 & 0 \end{bmatrix} \otimes I_q\right) (Q \Xi^{1/2} \otimes \Lambda_1^{-1/2}) \text{vec}(M) \\
 &= E\left(h\left(\chi_{pq+2}^2(\mu'_1 \mu_1)\right)\right) \left(\Xi^{-1/2} Q' \otimes \Lambda_1^{1/2}\right) (Q \Xi^{1/2} Y \Xi^{1/2} Q' \otimes I_q) \\
 &\quad \times (Q \Xi^{1/2} \otimes \Lambda_1^{-1/2}) \text{vec}(M) \\
 &= E\left(h\left(\chi_{pq+2}^2(\mu'_1 \mu_1)\right)\right) \left(\Xi^{-1/2} Q' Q \Xi^{1/2} Y \Xi^{1/2} Q' Q \Xi^{1/2} \right. \\
 &\quad \left. \otimes \Lambda_1^{1/2} \Lambda_1^{-1/2}\right) \text{vec}(M) \\
 &= E\left(h\left(\chi_{pq+2}^2(\mu'_1 \mu_1)\right)\right) (Y \Xi \otimes I_q) \text{vec}(M) \\
 &= E\left(h\left(\chi_{pq+2}^2(\mu'_1 \mu_1)\right)\right) \text{vec}(Y \Xi M).
 \end{aligned}$$

بنا بر شرط ب در صورت قضیه، و اینکه Ξ یک ماتریس نامنفرد است، داریم

$$\Xi^{-1} \Xi Y \Xi M = \Xi^{-1} \Xi M \implies Y \Xi M = M.$$

در نتیجه خواهیم داشت:

$$\text{vec}(E[h(\text{trace}(\Lambda_1^{-1} X' \Xi Y \Xi X))X]) = \text{vec}\left(E\left(h\left(\chi_{pq+2}^2(\mu'_1 \mu_1)\right)\right) M\right),$$

□

و اثبات کامل است.

به منظور نشان دادن برقراری شرایط الف و ب از قضیه ۱-۳ به مثال زیر توجه کنید.

مثال ۱-۳: فرض کنید $\Xi = \Lambda^{-1}$ ، $M = \Lambda L_1' (L_1 \Lambda L_1')^{-1} (L_1 \theta L_2 - d)$ ، $\Lambda_1 = L_2' L_2$ و $Y = \Lambda L_1' (L_1 \Lambda L_1')^{-1} L_1 \Lambda$ که در آن ماتریس‌های L_1 ، L_2 ، Λ و θ در بخش ۱-۳-۳ معرفی شده‌اند. همچنین توجه کنید که $M = E[(\hat{\theta} - \theta)L_2]$ بنابراین

$$Y \Xi Y \Xi = \Lambda L_1' (L_1 \Lambda L_1')^{-1} L_1 = Y \Xi,$$

و شرط الف برقرار است. از طرفی

$$Y \Xi E M = L_1' (L_1 \Lambda L_1')^{-1} (L_1 \theta L_2 - d) = \Xi M.$$

□

بنابراین شرط ب نیز برقرار است.

قضیه ۲-۳ (کرونزیا، ۲۰۱۲) فرض کنید

$$(X', Y')' \sim MN_{2k \times q} \left((M_1', M_2'), \begin{pmatrix} Y_{11} \otimes \Lambda_{11} & 0 \\ 0 & Y_{22} \otimes \Lambda_{22} \end{pmatrix} \right),$$

که در آن Λ_{11} یک ماتریس معین مثبت، و Y_{11} ، Y_{22} و Λ_{22} ماتریس‌های معین نامنفی با رتبه $p \leq k$ هستند. همچنین فرض کنید Ξ یک ماتریس معین مثبت و متقارن با دو شرط زیر باشد:

الف- $Y_{11} \Xi$ یک ماتریس خودتوان است.

ب- $\Xi Y_{11} \Xi M_1 = \Xi M_1$.

در این صورت برای هر تابع اندازه‌پذیر و انتگرال‌پذیر بول h ، و هر ماتریس معین نامنفی A ، داریم

$$\begin{aligned} E[h(\text{trace}(\Lambda_{11}^{-1}X'\Xi Y_{11}\Xi X))Y'AX] \\ = E[h(\chi_{pq+2}^2(\text{trace}(\Lambda_{11}^{-1}M_1'\Xi Y_{11}\Xi M_1)))M_2'AM_1']. \end{aligned} \quad (23-3)$$

اثبات: از آن جایی که X و Y از هم مستقل هستند، پس:

$$\begin{aligned} E[h(\text{trace}(\Lambda_{11}^{-1}X'\Xi Y_{11}\Xi X))Y'AX] \\ = (E(Y))'AE[h(\text{trace}(\Lambda_{11}^{-1}X'\Xi Y_{11}\Xi X))X]. \end{aligned}$$

همچنین بنا به قضیه ۱-۳،

$$E[h(\text{trace}(\Lambda_{11}^{-1}X'\Xi Y_{11}\Xi X))X] = E\left[h\left(\chi_{pq+2}^2(\text{trace}(\Lambda_{11}^{-1}M_1'\Xi Y_{11}\Xi M_1))\right)\right]M_1$$

□ حال با توجه به اینکه $(E(Y))' = M_2'$ ، اثبات کامل می شود.

مثال ۲-۳: به منظور نشان دادن شرایط الف و ب از قضیه ۲-۳ فرض کنید Ξ ، Y ، Λ_1 و M همان

ماتریس های داده شده در مثال ۱-۳ باشند. همچنین $Y_{11} = Y$ ، $M_1 = M$ ، $M_2 = -M$ ، $\Lambda_{11} =$

$\Lambda_{22} = \Lambda - Y$ و $Y_{22} = Y$. به وضوح ماتریس های Ξ ، Y_{11} ، M_1 و Λ_{11} برقراری شرایط الف و ب را

نشان می دهند. همچنین دقت کنید که $(\hat{\theta} - \theta)L_2$ نقش Y و $(X - \hat{\theta})L_2$ نقش X در قضیه ۲-۳ را

ایفا می کنند.

قضیه ۳-۳ (کرونزیزا، ۲۰۱۲). فرض کنید $X \sim MN_{k \times q}(M, Y \otimes \Lambda)$ که در آن Λ ماتریس معین

مثبت و Y یک ماتریس معین نامنفی با رتبه $p \leq k$ است. همچنین فرض کنید A و Ξ ماتریس های

متقارن معین مثبت هستند که برای Ξ دو شرط زیر برقرار است:

الف- $Y\Xi$ ماتریس خودتوان است.

$$\mathbb{E}Y\mathbb{E}M = \mathbb{E}M \quad \text{ب-}$$

در این صورت برای هر تابع اندازه‌پذیر و انتگرال‌پذیر بول h ، داریم:

$$\begin{aligned} & E[h(\text{trace}(\Lambda^{-1}X'\mathbb{E}Y\mathbb{E}X))\text{trace}(X'AX)] \\ &= E\left[h\left(\chi_{pq+2}^2(\text{trace}(\Lambda^{-1}M'\mathbb{E}Y\mathbb{E}M))\right)\right]\text{trace}(AY)\text{trace}(\Lambda) \\ &+ E\left[h\left(\chi_{pq+4}^2(\text{trace}(\Lambda^{-1}M'\mathbb{E}Y\mathbb{E}M))\right)\right]\text{trace}(M'AM). \end{aligned} \quad (24-3)$$

اثبات: از رابطه (۱۸-۳) می‌توان نتیجه گرفت

$$h(\text{trace}(\Lambda^{-1}X'\mathbb{E}Y\mathbb{E}X)) = h(V_1'V_1).$$

همانند قضیه ۱-۳، فرض کنید $V = Q\mathbb{E}^{1/2}X\Lambda^{-1/2}$ لذا $X = \mathbb{E}^{-1/2}Q'V\Lambda^{1/2}$ و

$$X'AX = \Lambda^{1/2}V'Q\mathbb{E}^{-1/2}A\mathbb{E}^{-1/2}Q'V\Lambda^{1/2}. \quad (25-3)$$

در نتیجه

$$\begin{aligned} & E[h(\text{trace}(\Lambda^{-1}X'\mathbb{E}Y\mathbb{E}X))\text{trace}(X'AX)] \\ &= E[h(V_1'V_1)\text{trace}(\Lambda^{1/2}V'Q\mathbb{E}^{-1/2}A\mathbb{E}^{-1/2}Q'V\Lambda^{1/2})] \\ &= E[h(V_1'V_1)\text{trace}(V'Q\mathbb{E}^{-1/2}A\mathbb{E}^{-1/2}Q'V\Lambda)] \\ &= E[h(V_1'V_1)\text{vec}'(V)(Q\mathbb{E}^{-1/2}A\mathbb{E}^{-1/2}Q' \otimes \Lambda)\text{vec}(V)]. \end{aligned}$$

حال فرض کنید که

$$(Q\Xi^{-1/2}A\Xi^{-1/2}Q' \otimes \Lambda) = G = \begin{bmatrix} G_{11} & G_{12} \\ G_{21} & G_{22} \end{bmatrix}, \quad (26-3)$$

که زیرماتریس‌های مربع G_{11} و G_{22} به ترتیب با بعدهای pq و $(k-p)q$ هستند.

مشابه قضیه ۱-۳، $vec(V)$ را به صورت زیر افراز می‌کنیم:

$$vec(V) = (V_1', o)'$$

در این صورت با استفاده از رابطه (۲۶-۳) داریم:

$$vec'(V)(Q\Xi^{-1/2}A\Xi^{-1/2}Q' \otimes \Lambda)vec(V) = V_1'G_{11}V_1.$$

بنابراین

$$E[h(V_1'V_1)vec'(V)(Q\Xi^{-1/2}A\Xi^{-1/2}Q' \otimes \Lambda)vec(V)] = E[h(V_1'V_1)V_1'G_{11}V_1].$$

حال بنا بر لم ۱-۵ خواهیم داشت:

$$\begin{aligned} E[h(V_1'V_1)V_1'G_{11}V_1] &= E \left[h \left(\chi_{pq+2}^2(\mu_1'\mu_1) \right) \right] trace(G_{11}) \\ &+ E \left[h \left(\chi_{pq+2}^2(\mu_1'\mu_1) \right) \right] (\mu_1'G_{11}\mu_1), \end{aligned} \quad (27-3)$$

که در آن μ_1 در رابطه (۱۷-۳) داده شده است.

به‌علاوه توجه کنید که

$$G_{11} = ([I_p, 0] \otimes I_q) G \left(\begin{bmatrix} I_p \\ 0 \end{bmatrix} \otimes I_q \right). \quad (28-3)$$

اکنون با ترکیب رابطه (28-3) با (17-3) و با استفاده از رابطه (15-3) داریم:

$$\begin{aligned} \mu_1' G_{11} \mu_1 &= \text{vec}'(M) (\Xi^{1/2} Q' \otimes \Lambda^{-1/2}) \left(\begin{bmatrix} I_p & 0 \\ 0 & 0 \end{bmatrix} \otimes I_q \right) G \left(\begin{bmatrix} I_p & 0 \\ 0 & 0 \end{bmatrix} \otimes I_q \right) \\ &\quad \times (Q \Xi^{1/2} \otimes \Lambda^{-1/2}) \text{vec}(M) \\ &= \text{vec}'(M) (\Xi^{1/2} Q' \otimes \Lambda^{-1/2}) (Q \Xi^{1/2} \Upsilon \Xi^{1/2} Q' \otimes I_q) \\ &\quad \times (Q \Xi^{-1/2} A \Xi^{-1/2} Q' \otimes \Lambda) (Q \Xi^{1/2} \Upsilon \Xi^{1/2} Q' \otimes I_q) \\ &\quad \times (Q \Xi^{1/2} \otimes \Lambda^{-1/2}) \text{vec}(M) \\ &= \text{vec}'(M) (\Xi \Upsilon A \Upsilon \Xi \otimes I_q) \text{vec}(M) \\ &= \text{trace}(M' \Xi \Upsilon A \Upsilon M) \\ &= \text{trace}(M' \Xi \Upsilon \Xi \Upsilon A M). \end{aligned}$$

تحت فرض الف، $\Xi \Upsilon \Xi \Upsilon = \Xi \Upsilon$. بنابراین $\mu_1' G_{11} \mu_1 = \text{trace}(M' \Xi \Upsilon A M)$ همچنین تحت فرض

ب، $M' \Xi \Upsilon = M'$ در نتیجه

$$\mu_1' G_{11} \mu_1 = \text{trace}(M' A M). \quad (29-3)$$

افزون بر آن با استفاده از رابطه‌های (17-3)، (26-3) و (28-3) داریم

$$\begin{aligned} \text{trace}(G_{11}) &= \text{trace}[(Q \Xi^{-1/2} A \Xi^{-1/2} Q' \otimes \Lambda) (Q \Xi^{1/2} \Upsilon \Xi^{1/2} Q' \otimes I_q)] \\ &= \text{trace}(A \Upsilon \otimes \Lambda) \\ &= \text{trace}(A \Upsilon) \text{trace}(\Lambda). \end{aligned} \quad (30-3)$$

در نهایت، اثبات، با ترکیب رابطه‌های (۲۲-۳)، (۲۷-۳)، (۲۹-۳) و (۳۰-۳) به صورت زیر کامل می‌شود.

$$\begin{aligned} & E[h(\text{trace}(\Lambda^{-1}X'\Xi Y \Xi X))\text{trace}(X'AX)] \\ &= E\left[h\left(\chi_{pq+2}^2(\mu_1'\mu_1)\right)\right]\text{trace}(G_{11}) + E\left[h\left(\chi_{pq+4}^2(\mu_1'\mu_1)\right)\right](\mu_1'G_{11}\mu_1) \\ &= E\left[h\left(\chi_{pq+2}^2(\text{trace}(\Lambda^{-1}M'\Xi Y \Xi M))\right)\right]\text{trace}(AY)\text{trace}(\Lambda) \\ &\quad + E\left[h\left(\chi_{pq+4}^2(\text{trace}(\Lambda^{-1}M'\Xi Y \Xi M))\right)\right]\text{trace}(M'AM) \end{aligned}$$

□

و اثبات کامل است.

۳-۳-۳ محاسبه توابع مخاطره برآوردگرها

با توجه به اینکه برآوردگر انقباضی $\hat{\theta}_S$ را، با یک عامل انقباض کلی نوشتیم، لذا می‌توان با جایگزین کردن عوامل انقباض دیگری، به برآوردگرهای جدیدی رسید. برای محاسبه تابع مخاطره برآوردگرها، ابتدا مخاطره $\hat{\theta}_S$ را به دست آورده، آن‌گاه با استفاده از مثال ۱-۳ مخاطره دیگر برآوردگرها ($\hat{\theta}$ ، $\tilde{\theta}$ ، $\hat{\theta}^S$ و $\hat{\theta}^{sp}$) را محاسبه می‌کنیم.

در این قسمت، بر اساس تعریف ۱-۵، تابع مخاطره موزون را به صورت زیر در نظر می‌گیریم:

$$R(\hat{\theta}_S, \theta; W) = E\left[\text{trace}\left(L_2'(\hat{\theta}_S - \theta)'W(\hat{\theta}_S - \theta)L_2\right)\right]. \quad (۳۱-۳)$$

قضیه ۳-۴ (کرونزیا، ۲۰۱۲). تابع مخاطره برآوردگر $\hat{\theta}_S$ به صورت زیر است

$$\begin{aligned} R(\hat{\theta}_S, \theta; W) &= \text{trace}(W(\Sigma^{-1} - \Sigma^*))\text{trace}(L_2'L_2) + \text{trace}(\delta'W\delta) \\ &\quad - 2E\left[h\left(\chi_{pq+2}^2(\text{trace}(\Xi_2\delta'\Xi_1\delta))\right)\right]\text{trace}(\delta'W\delta) \\ &\quad + E\left[h^2\left(\chi_{pq+2}^2(\text{trace}(\Xi_2\delta'\Xi_1\delta))\right)\right]\text{trace}(W\Sigma^*)\text{trace}(L_2'L_2) \end{aligned}$$

$$+E \left[h^2 \left(\chi_{pq+4}^2 \left(\text{trace}(\Xi_2 \delta' \Xi_1 \delta) \right) \right) \right] \text{trace}(\delta' W \delta), \quad (32-3)$$

که در آن $\delta, \Sigma^*, \Sigma^{-1}$ در لم ۳-۲ داده شده است.

اثبات: برای سهولت در نوشتن، فرض کنید

$$\eta = (\hat{\theta} - \theta) L_2, \quad \xi = (X - \hat{\theta}) L_2$$

که توزیع η و ξ در لم ۳-۲ آمده است. بنابراین از رابطه (۷-۳)، داریم:

$$\hat{\theta}_S = \hat{\theta} + h \|\xi\|_{\Xi_1, \Xi_2}^2 \xi L_2^{-1}.$$

از طرفی،

$$\begin{aligned} L_2'(\hat{\theta}_S - \theta)' &= L_2' \left((\hat{\theta} - \theta) + h \|\xi\|_{\Xi_1, \Xi_2}^2 \xi L_2^{-1} \right)' \\ &= L_2' \left((\hat{\theta} - \theta)' + (L_2')^{-1} \xi' h \|\xi\|_{\Xi_1, \Xi_2}^2 \right) \\ &= L_2'(\hat{\theta} - \theta)' + L_2'(L_2')^{-1} \xi' h \|\xi\|_{\Xi_1, \Xi_2}^2 \\ &= \eta' + \xi' h \|\xi\|_{\Xi_1, \Xi_2}^2 = (\eta + h \|\xi\|_{\Xi_1, \Xi_2}^2 \xi)' . \end{aligned}$$

بنابراین رابطه (۳۱-۳) را می توان به صورت زیر نوشت:

$$\begin{aligned} R(\hat{\theta}_S, \theta; W) &= E \left[\text{trace} \left((\eta + h \|\xi\|_{\Xi_1, \Xi_2}^2 \xi)' W (\eta + h \|\xi\|_{\Xi_1, \Xi_2}^2 \xi) \right) \right] \\ &= E \left[\text{trace} (\eta' W \eta + \eta' W h \|\xi\|_{\Xi_1, \Xi_2}^2 \xi + \xi' h \|\xi\|_{\Xi_1, \Xi_2}^2 W \eta \right. \\ &\quad \left. + \xi' h \|\xi\|_{\Xi_1, \Xi_2}^2 W h \|\xi\|_{\Xi_1, \Xi_2}^2 \xi) \right] \\ &= E \left[\text{trace} (\eta' W \eta + 2h \|\xi\|_{\Xi_1, \Xi_2}^2 \eta' W \xi + h^2 \|\xi\|_{\Xi_1, \Xi_2}^2 \xi' W \xi) \right]. \end{aligned}$$

لذا وقتی h یک تابع حقیقی مقدار باشد، خواهیم داشت:

$$R(\hat{\theta}_S, \theta; W) = E\{trace(\eta'W\eta)\} + 2E\{(h\|\xi\|_{\Xi_1, \Xi_2}^2)trace(\eta'W\xi)\} \\ + E\{(h^2\|\xi\|_{\Xi_1, \Xi_2}^2)trace(\xi'W\xi)\}. \quad (33-3)$$

حال با استفاده از قضایای ۸-۱، ۲-۳ و ۳-۳ و لم‌های ۱-۱ و ۲-۳ داریم:

$$E\{trace(\eta'W\eta)\} = trace\{E(\eta'W\eta)\} \\ = trace\{trace[(\Sigma^{-1} - \Sigma^*)W']L'_2L_2 + \delta'W\delta\} \\ = trace[W(\Sigma^{-1} - \Sigma^*)L'_2L_2] + trace(\delta'W\delta) \\ = trace(W(\Sigma^{-1} - \Sigma^*))trace(L'_2L_2) + trace(\delta'W\delta),$$

$$E\{(h\|\xi\|_{\Xi_1, \Xi_2}^2)trace(\eta'W\xi)\} = E[h(trace(\xi'\Xi_1\xi\Xi_2))\eta'W\xi] \\ = E[h(trace(\Xi_2\xi'\Xi_1\xi))\eta'W\xi] \\ = -E[h(\chi_{pq+2}^2(trace(\Xi_2\delta'\Xi_1\delta)))]trace(\delta'W\delta), \\ E\{(h^2\|\xi\|_{\Xi_1, \Xi_2}^2)trace(\xi'W\xi)\} = E[h^2(trace(\Xi_2\xi'\Xi_1\xi))trace(\xi'W\xi)] \\ = E[h^2(\chi_{pq+2}^2(trace(\Xi_2\delta'\Xi_1\delta)))]trace(W\Sigma^*)trace(L'_2L_2) \\ + E[h^2(\chi_{pq+4}^2(trace(\Xi_2\delta'\Xi_1\delta)))]trace(\delta'W\delta),$$

□

و اثبات کامل است.

حال با استفاده از مخاطره $\hat{\theta}_S$ و بر اساس برآوردگرهای ارائه شده در مثال ۳-۱، مخاطره برآوردگرهای

UE، RE، SSE و PRSSE را به دست می آوریم. بنابراین:

$$R(X, \theta; W) = trace(W(\Sigma^{-1} - \Sigma^*))trace(L'_2L_2) + trace(W\Sigma^*)trace(L'_2L_2)$$

$$= \text{trace}(W\Sigma^{-1})\text{trace}(L'_2L_2), \quad (34-3)$$

$$R(\hat{\theta}, \theta; W) = R(X, \theta; W) - \text{trace}(W\Sigma^*)\text{trace}(L'_2L_2) + \text{trace}(\delta'W\delta). \quad (35-3)$$

برای محاسبه مخاطره برآوردگرهای استاین و استاین مثبت با فرض $\Delta = \text{trace}(\Xi_2\delta'\Xi_1\delta)$ می توان

نوشت:

$$\begin{aligned} R(\hat{\theta}^s, \theta; W) &= R(X, \theta; W) - \text{trace}(W\Sigma^*)\text{trace}(L'_2L_2) + \text{trace}(\delta'W\delta) \\ &\quad - 2E\left(1 - \frac{pq-2}{\chi^2_{pq+2}(\Delta)}\right)\text{trace}(\delta'W\delta) \\ &\quad + E\left(1 - \frac{pq-2}{\chi^2_{pq+2}(\Delta)}\right)^2 \text{trace}(W\Sigma^*)\text{trace}(L'_2L_2) \\ &\quad + E\left(1 - \frac{pq-2}{\chi^2_{pq+4}(\Delta)}\right)^2 \text{trace}(\delta'W\delta) \\ &= R(X, \theta; W) - \text{trace}(W\Sigma^*)\text{trace}(L'_2L_2) + \text{trace}(\delta'W\delta) \\ &\quad - 2\text{trace}(\delta'W\delta) + 2(pq-2)E\left(\chi^{-2}_{pq+2}(\Delta)\right)\text{trace}(\delta'W\delta) \\ &\quad + \text{trace}(W\Sigma^*)\text{trace}(L'_2L_2) - 2(pq-2)E\left(\chi^{-2}_{pq+2}(\Delta)\right) \\ &\quad \quad \times \text{trace}(W\Sigma^*)\text{trace}(L'_2L_2) \\ &\quad + (pq-2)^2E\left(\chi^{-4}_{pq+2}(\Delta)\right)\text{trace}(W\Sigma^*)\text{trace}(L'_2L_2) \\ &\quad + \text{trace}(\delta'W\delta) - 2(pq-2)E\left(\chi^{-2}_{pq+4}(\Delta)\right)\text{trace}(\delta'W\delta) \\ &\quad + (pq-2)E\left(\chi^{-4}_{pq+4}(\Delta)\right)\text{trace}(\delta'W\delta) \end{aligned}$$

$$\begin{aligned}
 &= R(X, \theta; W) - (pq - 2) \left[2E \left(\chi_{pq+2}^{-2}(\Delta) \right) - (pq - 2)E \left(\chi_{pq+2}^{-4}(\Delta) \right) \right] \\
 &\quad \times \text{trace}(W\Sigma^*)\text{trace}(L'_2L_2) \\
 &+ (pq - 2) \left[2E \left(\chi_{pq+2}^{-2}(\Delta) \right) - 2E \left(\chi_{pq+4}^{-2}(\Delta) \right) + (pq - 2)E \left(\chi_{pq+4}^{-4}(\Delta) \right) \right] \\
 &\quad \times \text{trace}(\delta'W\delta).
 \end{aligned}$$

حال با استفاده از نتیجه ۱-۲، برای عبارت داخل کروشه در سطر آخر داریم:

$$\begin{aligned}
 &2E \left(\chi_{pq+2}^{-2}(\Delta) \right) - 2E \left(\chi_{pq+4}^{-2}(\Delta) \right) + (pq - 2)E \left(\chi_{pq+4}^{-4}(\Delta) \right) \\
 &= (pq + 2)E \left(\chi_{pq+4}^{-4}(\Delta) \right),
 \end{aligned}$$

که در نتیجه

$$\begin{aligned}
 R(\hat{\theta}^s, \theta; W) &= R(X, \theta; W) - (pq - 2)\text{trace}(L'_2L_2)\text{trace}(W\Sigma^*) \\
 &\quad \times \left[2E \left(\chi_{pq+2}^{-2}(\Delta) \right) - (pq - 2)E \left(\chi_{pq+2}^{-4}(\Delta) \right) \right] \\
 &\quad + \text{trace}(\delta'W\delta)((pq)^2 - 4)E \left(\chi_{pq+4}^{-4}(\Delta) \right). \quad (۳۶-۳)
 \end{aligned}$$

به طریقی مشابه

$$\begin{aligned}
 R(\hat{\theta}^{s+}, \theta; W) &= R(X, \theta; W) - \text{trace}(W\Sigma^*)\text{trace}(L'_2L_2) + \text{trace}(\delta'W\delta) \\
 &\quad - 2E \left[\left(1 - \frac{pq - 2}{\chi_{pq+2}^2(\Delta)} \right) I_{(\chi_{pq+2}^2(\Delta) > (pq-2))} \right] \text{trace}(\delta'W\delta) \\
 &\quad + E \left[\left(1 - \frac{pq - 2}{\chi_{pq+2}^2(\Delta)} \right)^2 I_{(\chi_{pq+2}^2(\Delta) > (pq-2))} \right] \text{trace}(L'_2L_2)\text{trace}(W\Sigma^*) \\
 &\quad + E \left[\left(1 - \frac{pq - 2}{\chi_{pq+4}^2(\Delta)} \right)^2 I_{(\chi_{pq+4}^2(\Delta) > (pq-2))} \right] \text{trace}(\delta'W\delta).
 \end{aligned}$$

حال با توجه به اینکه $E \left[\left(1 - \frac{r}{\chi}\right) I_{(\chi > r)} \right] = E \left(1 - \frac{r}{\chi}\right) - E \left[\left(1 - \frac{r}{\chi}\right) I_{(\chi \leq r)} \right]$ خواهیم داشت:

$$\begin{aligned} R(\hat{\theta}^{s+}, \theta; W) &= R(\hat{\theta}^s, \theta; W) \\ &+ 2E \left[\left(1 - \frac{pq-2}{\chi_{pq+2}^2(\Delta)}\right) I_{(\chi_{pq+2}^2(\Delta) \leq (pq-2))} \right] \text{trace}(\delta' W \delta) \\ &- E \left[\left(1 - \frac{pq-2}{\chi_{pq+2}^2(\Delta)}\right)^2 I_{(\chi_{pq+2}^2(\Delta) \leq (pq-2))} \right] \text{trace}(L'_2 L_2) \text{trace}(W \Sigma^*) \\ &- E \left[\left(1 - \frac{pq-2}{\chi_{pq+4}^2(\Delta)}\right)^2 I_{(\chi_{pq+4}^2(\Delta) \leq (pq-2))} \right] \text{trace}(\delta' W \delta) \end{aligned} \quad (37-3)$$

و با فرض $d = pq - 2$ و $H_v(x; \Delta) = P\{\chi_v^2(\Delta) \leq x\}, x \geq 0$ می توان نتیجه گرفت:

$$\begin{aligned} R(\hat{\theta}^{s+}, \theta; W) &= R(\hat{\theta}^s, \theta; W) + d \text{trace}(L'_2 L_2) \text{trace}(W \Sigma^*) \\ &\times \left[2E \left\{ \chi_{pq+2}^{-2}(\Delta) I_{(\chi_{pq+2}^2(\Delta) \leq d)} \right\} - d E \left\{ \chi_{pq+2}^{-4}(\Delta) I_{(\chi_{pq+2}^2(\Delta) \leq d)} \right\} \right] \\ &- \text{trace}(L'_2 L_2) \text{trace}(W \Sigma^*) H_{pq+2}(d; \Delta) + \text{trace}(\delta' W \delta) \\ &\times [2H_{pq+2}(d; \Delta) - H_{pq+4}(d; \Delta)] - d \text{trace}(\delta' W \delta) \\ &\times \left[2E \left\{ \chi_{pq+2}^{-2}(\Delta) I_{(\chi_{pq+2}^2(\Delta) \leq d)} \right\} - 2E \left\{ \chi_{pq+4}^{-4}(\Delta) I_{(\chi_{pq+4}^2(\Delta) \leq d)} \right\} \right. \\ &\left. + d E \left\{ \chi_{pq+4}^{-4}(\Delta) I_{(\chi_{pq+4}^2(\Delta) \leq (pq-2))} \right\} \right]. \end{aligned} \quad (38-3)$$

۴-۳-۳ مقایسه برآوردگرها

در این زیربخش با استفاده از توابع مخاطره بدست آمده در ۳-۳-۳، رفتار برآوردگرهای مورد نظر را با یکدیگر مورد بررسی قرار می‌دهیم.

مقایسه برآوردگرهای $\hat{\theta}$ و $\tilde{\theta}$

به ازای وزن $W = \Lambda^{-1}$ داریم

$$\begin{aligned} \text{trace}(W\Sigma^*) &= \text{trace}(\Lambda^{-1}\Sigma^*) = \text{trace}(I_p) = p, \\ \text{trace}(\delta'W\delta) &= \text{trace}(\delta'\Lambda^{-1}\delta) \\ &= (L_1\theta L_2 - d)'(L_1\Lambda L_1')^{-1}L_1\Lambda\Lambda^{-1}\Lambda L_1'(L_1\Lambda L_1')^{-1}(L_1\theta L_2 - d) \\ &= (L_1\theta L_2 - d)'(L_1\Lambda L_1')^{-1}(L_1\theta L_2 - d). \end{aligned}$$

بنابراین از (۳۴-۳) و (۳۵-۳) خواهیم داشت:

$$\begin{aligned} \mathcal{D}_1 &= R(\tilde{\theta}, \theta; \Lambda^{-1}) - R(\hat{\theta}, \theta; \Lambda^{-1}) \\ &= \text{trace}(L_2' L_2) \text{trace}(\Lambda^{-1} \Sigma^*) - \text{trace}(\delta' \Lambda^{-1} \delta) \\ &= p \text{trace}(L_2' L_2) - \text{trace}\{(L_1 \theta L_2 - d)'(L_1 \Lambda L_1')^{-1}(L_1 \theta L_2 - d)\}. \end{aligned}$$

حال از آنجا که تحت محدودیت (۳-۳)،

0

لذا $\mathcal{D}_1 > 0$ یعنی برآوردگر $\tilde{\theta}$ در مقایسه با $\hat{\theta}$ دارای مخاطره بزرگتری است. در این صورت می‌گوییم برآوردگر $\hat{\theta}$ بهتر از برآوردگر $\tilde{\theta}$ است و با $\hat{\theta} > \tilde{\theta}$ نشان می‌دهیم.

از طرفی $\Delta = \text{trace}(\delta' \Xi_1 \delta (L_2' L_2)^{-1})$ که در آن

$$\delta' \Xi_1 \delta = (L_1 \theta L_2 - d)'(L_1 \Lambda L_1')^{-1} L_1 \Lambda \Xi_1 \Lambda L_1' (L_1 \Lambda L_1')^{-1} (L_1 \theta L_2 - d). \quad (۳۹-۳)$$

$\Lambda \Xi_1$ یک ماتریس خودتوان است، یعنی $\Lambda \Xi_1 \Lambda \Xi_1 = \Lambda \Xi_1$ و Ξ_1 ماتریس معین مثبت و در نتیجه نامنفرد است. بنابراین

$$\Lambda \Xi_1 \Lambda \Xi_1 \Xi_1^{-1} = \Lambda \Xi_1 \Xi_1^{-1} \Rightarrow \Lambda \Xi_1 \Lambda = \Lambda.$$

از این رو رابطه (۳-۳۹) را می توان به صورت زیر نیز نوشت:

$$\begin{aligned} \delta' \Xi_1 \delta &= (L_1 \theta L_2 - d)' (L_1 \Lambda L_1')^{-1} L_1 \Lambda L_1' (L_1 \Lambda L_1')^{-1} (L_1 \theta L_2 - d) \\ &= (L_1 \theta L_2 - d)' (L_1 \Lambda L_1')^{-1} (L_1 \theta L_2 - d). \end{aligned} \quad (۴۰-۳)$$

در نهایت

$$\begin{aligned} \Delta &= \text{trace}(\delta' \Xi_1 \delta (L_2' L_2)^{-1}) \\ &> \text{trace}(\delta' \Xi_1 \delta) \\ &= \text{trace}(L_1 \theta L_2 - d)' (L_1 \Lambda L_1')^{-1} (L_1 \theta L_2 - d). \end{aligned} \quad (۴۱-۳)$$

در نتیجه زمانی که اطلاعاتی از درستی محدودیت (۳-۳) نداریم، از (۴۰-۳) و (۴۱-۳) خواهیم داشت

$$\mathcal{D}_1 > 0 \text{ اگر } \Delta < p \text{trace}(L_2' L_2) \text{ باشد.}$$

همچنین اگر $p L_2' L_2 - (L_1 \theta L_2 - d)' (L_1 \Lambda L_1')^{-1} (L_1 \theta L_2 - d)$ یک ماتریس معین مثبت

$$\mathcal{D}_1 > 0 \text{ باشد با توجه به قضیه } A-۱۶ \text{ و } A-۱۴، \text{ قویا می توان گفت که}$$

و اگر $p L_2' L_2 = (L_1 \theta L_2 - d)' (L_1 \Lambda L_1')^{-1} (L_1 \theta L_2 - d)$ در این صورت $\Delta = pm$ و

$$\mathcal{D}_1 = 0$$

مقایسه برآوردگرهای $\hat{\theta}^s$ و $\tilde{\theta}$

به ازای وزن $W = \Lambda^{-1}$ ، و روابط (۳۴-۳)، (۳۶-۳) و (۴۰-۳) و نتیجه ۱-۲ می توان نوشت

$$\begin{aligned}
 \mathcal{D}_2 &= R(\tilde{\theta}, \theta; \Lambda^{-1}) - R(\hat{\theta}^s, \theta; \Lambda^{-1}) \\
 &= p(pq - 2)\text{trace}(L'_2 L_2) \left[2E\left(\chi_{pq+2}^{-2}(\Delta)\right) - (pq - 2)E\left(\chi_{pq+2}^{-4}(\Delta)\right) \right] \\
 &\quad - \text{trace}(\delta' \Lambda^{-1} \delta)((pq)^2 - 4)E\left(\chi_{pq+4}^{-4}(\Delta)\right) \\
 &> -\Delta((pq)^2 - 4)E\left(\chi_{pq+4}^{-4}(\Delta)\right) \\
 &\quad + p(pq - 2)\text{trace}(L'_2 L_2) \left[2E\left(\chi_{pq+2}^{-2}(\Delta)\right) - (pq - 2)E\left(\chi_{pq+2}^{-4}(\Delta)\right) \right] \\
 &= p\text{trace}(L'_2 L_2) \left\{ (pq - 2)E\left(\chi_{pq+2}^{-2}(\Delta)\right) \right. \\
 &\quad \left. + \left[1 - \frac{(pq)^2 - 4}{p(pq - 2)\text{trace}(L'_2 L_2)} \right] (pq - 2)\Delta E\left(\chi_{pq+4}^{-4}(\Delta)\right) \right\}
 \end{aligned}$$

لذا $\mathcal{D}_2 > 0$ به شرط آنکه

$$\text{trace}(L'_2 L_2) < \frac{(pq)^2 - 4}{p(pq - 2)} \quad (۴۱-۳)$$

بنابراین با برقراری شرط (۴۱-۳)، $\hat{\theta}^s > \tilde{\theta}$.

مقایسه برآوردگرهای $\hat{\theta}^{s+}$ و $\hat{\theta}^s$

این مقایسه را به ازای ماتریس وزن دلخواه W انجام می دهیم. بنابراین از (۳۷-۳) و (۳۸-۳) داریم

$$\mathcal{D}_3 = R(\hat{\theta}^s, \theta; W) - R(\hat{\theta}^{s+}, \theta; W)$$

$$\begin{aligned}
 &= +trace(L'_2 L_2) trace(W \Sigma^*) E \left[\left(1 - d \chi_{pq+2}^{-2}(\Delta) \right)^2 I(\chi_{pq+2}^2(\Delta) \leq d) \right] \\
 &\quad - 2trace(\delta' W \delta) E \left[\left(1 - d \chi_{pq+2}^{-2}(\Delta) \right) I(\chi_{pq+2}^2(\Delta) \leq d) \right] \\
 &\quad + trace(\delta' W \delta) E \left[\left(1 - d \chi_{pq+4}^{-2}(\Delta) \right)^2 I(\chi_{pq+4}^2(\Delta) \leq d) \right]. \quad (42-3)
 \end{aligned}$$

حال با توجه به این نکته که به ازای مقادیر $\chi_{pq+2}^2(\Delta) \leq d$ داریم

$$\left[\left(1 - d \chi_{pq+2}^{-2}(\Delta) \right) I(\chi_{pq+2}^2 \leq d) \right] < 0$$

بنابراین در رابطه (42-3)،

$$-2trace(\delta' W \delta) E \left[\left(1 - d \chi_{pq+2}^{-2}(\Delta) \right) I(\chi_{pq+2}^2(\Delta) < d) \right] > 0.$$

همچنین می‌دانیم که امید ریاضی متغیر تصادفی مثبت، مثبت است. بنابراین

$$E \left[\left(1 - d \chi_{pq+2}^{-2}(\Delta) \right)^2 I(\chi_{pq+2}^2(\Delta) < d) \right] > 0,$$

$$E \left[\left(1 - d \chi_{pq+4}^{-2}(\Delta) \right)^2 I(\chi_{pq+4}^2(\Delta) < d) \right] > 0.$$

در نتیجه تفاضل در (42-3) به ازای هر مقدار θ ، مثبت است. یعنی $\mathcal{D}_3 > 0$ و بنابراین $\hat{\theta}^{s+} > \hat{\theta}^s$.

از آنجا که مقایسه را بر اساس ماتریس وزن دلخواه انجام دادیم، می‌توان اینگونه بیان کرد که برآوردگر

$\hat{\theta}^{s+}$ به طور یکنواخت بر برآوردگر $\hat{\theta}^s$ برتری دارد.

۴-۳ برآورد انقباضی در مدل منحنی رشد

در فصل دوم، در مدل رگرسیون خطی چندمتغیره، بدون این که فرض ویژه‌ای برای توزیع ماتریس خطای مدل در نظر بگیریم، برآوردگرهای ناریب را به دست آوردیم. همچنین در بخش‌های قبلی، بدون در نظر گرفتن یک مدل رگرسیونی خاص، برآوردگرهای انقباضی برای میانگین در توزیع نرمال ماتریسی، مورد بحث قرار گرفت. در این بخش ابتدا مدل منحنی رشد، که تعمیمی از مدل رگرسیون چندمتغیره (۲-۱۳) می‌باشد را معرفی کرده، سپس با فرض این که ماتریس خطای مدل دارای توزیع نرمال ماتریسی است، برآوردگرهای انقباضی نوع استاین و نوع استاین مثبت را برای برآورد ماتریس پارامترهای مدل به دست می‌آوریم.

۳-۴-۱ مدل منحنی رشد

منحنی رشد، یک مدل تجربی از تکامل تدریجی یک متغیر کمی در طول زمان است که آن را با نماد GCM نمایش می‌دهند، این مدل آنالیز واریانس چندمتغیره تعمیم‌یافته^۵ (GMANOVA) است که برای آنالیز داده‌های طولی که در آن‌ها متغیر پاسخ در تعدادی از بازه‌های زمانی مشخص، مکرراً اندازه‌گیری می‌شود، مورد استفاده قرار می‌گیرد. بنابراین می‌توان یک الگوی کلی از تغییرات متغیر پاسخ را، پیدا کرد. به‌ویژه در مواقعی که با بازه‌های زمانی کوتاه‌مدت یا میان‌مدت سروکار داریم و روش سری زمانی کاربردی ندارد، مدل منحنی رشد بسیار مفید است. GCM با کاربردهای زیادی که دارد در زمینه‌های بسیاری راه یافته است. از آن جمله می‌توان به کاربرد آن در تجزیه و تحلیل منحنی‌های رشد یا روند در آمار زیستی، زیست‌شناسی و تحقیقات پزشکی اشاره کرد (پان و فانگ^۶، ۲۰۰۲ و وان رزن، ۱۹۸۹).

^۵ Generalized multivariate analysis of variance

^۶ Pan and Fang

اولین مقاله در ارتباط با منحنی‌های رشد توسط ویشارت^۷ (۱۹۳۸) منتشر شده است. اما GCM برای اولین بار توسط پات‌هاف و روی^۸ (۱۹۶۴) معرفی شد. برای آگاهی بیشتر در مورد منحنی‌های رشد به کتاب شیرساگار و اسمیت^۹ (۱۹۹۵) و برای تشخیص آماری با مدل‌های منحنی رشد می‌توان به پان و فانگ (۲۰۰۲) مراجعه نمود. همچنین می‌توان اطلاعات جامع‌تری را در کتاب کولو و وان‌رزن (۲۰۰۵) یافت.

تعریف ۱-۳ (GCM) فرض کنید X ماتریسی در اندازه $p \times n$ ، A ماتریسی در اندازه $p \times q$ ،

$q \leq p$ ، B در اندازه $q \times k$ و C در اندازه $k \times n$ باشند به طوری که $r(C) + p \leq n$ و Λ یک ماتریس معین مثبت با بعد p باشد. در این صورت مدل منحنی رشد را به صورت زیر تعریف می‌کنیم:

$$X = ABC + \Lambda^{1/2}E, \quad (43-3)$$

که در آن $E \sim MN_{p \times n}(0, I_p \otimes I_n)$. بنابراین با توجه به لم ۱-۱ می‌توان نوشت:

$$X \sim MN_{p \times n}(ABC, \Lambda \otimes I_n), \quad (44-3)$$

که در آن A ماتریس طرح درون گروهی و معلوم، C ماتریس طرح بین گروهی و معلوم، B و Λ ماتریس-های پارامتر نامعلوم هستند. ستون‌های ماتریس X مستقل از هم و دارای توزیع نرمال p -متغیره، با ماتریس کوواریانس نامعلوم Λ هستند و همچنین $E(X) = ABC$.

برای درک بهتر از مدل (۴۳-۳)، مثال زیر را ببینید.

⁷ Wishart
⁸ Potthoff and Roy
⁹ Kshirsagar and Smith

مثال ۳-۴: فرض کنید ۳ گروه تیمار از حیوانات داریم. بطوریکه در گروه j -ام، n_j حیوان وجود دارد. هر گروه تحت q روش درمانی مختلف قرار می‌گیرد. هدف، بررسی میزان افزایش وزن در هر گروه است. برای این منظور همه حیوانات پس از اعمال هر روش درمانی، در p نقطه زمانی مشابه اندازه‌گیری می‌شوند. (مثلاً $t_r, r = 1, 2, \dots, p$). که این موضوع (اندازه‌گیری در نقاط زمانی مشابه) شرط لازم برای کاربرد GCM است. بنابراین تغییرات وزن n حیوان در p نقطه را می‌توان در قالب یک ماتریس تصادفی، مثل X ، با بعد $p \times n$ نمایش داد، که در آن $n = n_1 + n_2 + n_3$. هر ستون X بیانگر مشاهدات اندازه‌گیری شده برای یک حیوان در p نقطه زمانی است. اولین n_1 ستون از ماتریس X متعلق به گروه یک، n_2 ستون بعدی متعلق به گروه دو و الی آخر.

$$X = \begin{pmatrix} X_{11} & X_{12} & \dots & X_{1n_1} & X_{1(n_1+1)} & \dots & X_{1(n_1+n_2)} & X_{1(n_1+n_2+1)} & \dots & X_{1(n_1+n_2+n_3)} \\ X_{21} & X_{22} & \dots & X_{2n_1} & X_{2(n_1+1)} & \dots & X_{2(n_1+n_2)} & X_{2(n_1+n_2+1)} & \dots & X_{2(n_1+n_2+n_3)} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ X_{p1} & X_{p2} & \dots & X_{pn_1} & X_{p(n_1+1)} & \dots & X_{p(n_1+n_2)} & X_{p(n_1+n_2+1)} & \dots & X_{p(n_1+n_2+n_3)} \end{pmatrix}_{p \times n}$$

بر اساس تجربه دریافته‌اند که وزن هر گروه از حیوانات دارای توزیع نرمال چندمتغیره با میانگین μ_j ، $j = 1, 2, 3$ و ماتریس کوواریانس Λ است. همچنین اندازه‌گیری حیوانات در گروه‌های مختلف، مستقل از هم می‌باشد و فرض می‌شود میانگین توزیع برای هر گروه تیمار، یک چند جمله‌ای از درجه $q - 1$ نسبت به زمان است. بنابراین میانگین μ_j از j -امین گروه تیمار در زمان t به صورت زیر است:

$$\mu_j = \beta_{1j} + \beta_{2j}t + \dots + \beta_{qj}t^{q-1}, \quad j = 1, 2, 3 \quad (۴۵-۳)$$

که در آن پارامترهای نامعلوم هستند.

با توجه به ساختار X ، ماتریس طرح بین گروهی، C ، در اندازه $3 \times n$ به صورت

$$C = \begin{pmatrix} \overbrace{1 \ 1 \ \dots \ 1}^{n_1} & \overbrace{0 \ 0 \ \dots \ 0}^{n_2} & \overbrace{0 \ 0 \ \dots \ 0}^{n_3} \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{pmatrix}_{3 \times n}$$

و ماتریس طرح درون گروهی، A ، در اندازه $p \times q$ به صورت زیر نتیجه می‌شود:

$$A = \begin{pmatrix} 1 & t_1 & \dots & t_1^{q-1} \\ 1 & t_2 & \dots & t_2^{q-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & t_p & \dots & t_p^{q-1} \end{pmatrix}$$

از رابطه (۳-۴۵) می‌توان امید ریاضی ماتریس X را به صورت زیر نوشت:

$$E(X) = ABC$$

به‌طوری‌که B یک ماتریس $q \times 3$ از پارامترهای β_{ij} می‌باشد. که در آن $i = 1, 2, \dots, q$ و $j = 1, 2, 3$.

۳-۴-۲ برآوردگر انقباضی

همانند بخش قبل، از برآوردگر درست‌نمایی ماکزیمم پارامتر ماتریسی B به عنوان برآوردگر اولیه، و از برآوردگر محدودشده آن تحت یک قید خطی، به عنوان یک برآوردگر جدید استفاده می‌کنیم.

قضیه ۳-۵ (کولو و وان‌رزن، ۲۰۰۵). اگر A و C تمام رتبه باشند، یعنی $r(A) = q$ و $r(C) = k$ در

این صورت برآوردگر درست‌نمایی ماکزیمم B در مدل منحنی رشد (۳-۳۸)، به صورت زیر است

$$\tilde{B}_{q \times k} = (A'S^{-1}A)^{-1}A'S^{-1}XC'(CC')^{-1} \quad (46-3)$$

$$S_{p \times p} = X(I - C'(CC')^{-1}C)X'$$

که در آن

حال فرض کنید قیدی خطی به صورت

$$B = B_0 \quad (47-3)$$

به فضای پارامتر مدل (43-3)، تحمیل شده است. در این صورت به سهولت می‌توان دید که برآوردگر درست‌نمایی ماکزیمم محدود شده به صورت زیر است:

$$\hat{B}_{q \times k} = B_0. \quad (48-3)$$

حال مشابه بخش قبل، کلاسی از برآوردهای انقباضی را با یک عامل انقباض کلی به صورت زیر تشکیل می‌دهیم:

$$\hat{B}_S = \hat{B} + h(\|\tilde{B} - \hat{B}\|)(\tilde{B} - \hat{B}), \quad (49-3)$$

که در آن $\tilde{B}$ و $\hat{B}$ به ترتیب در روابط (46-3) و (47-3) داده شده‌اند.

۳-۴-۳ محاسبه مخاطره

در این زیربخش، برای محاسبه توابع مخاطره از تعاریف ۴-۱ و ۵-۱ استفاده شده است. همچنین فرض

$$\delta = B - B_0 \text{ و } \Sigma = (A'S^{-1}A)^{-1}A'S^{-1}$$

شده است

قضیه ۳-۶** تابع مخاطره برآوردگر $\hat{B}_S$ به صورت زیر است

$$\begin{aligned}
 R(\hat{B}_S, B; W) &= \text{trace}(\delta' W \delta) - 2E \left[h \left(\chi_{pq+2}^2 (\text{trace}(\delta' \delta)) \right) \right] \text{trace}(\delta' W \delta) \\
 &+ E \left[h^2 \left(\chi_{pq+2}^2 (\text{trace}(\delta' \delta)) \right) \right] \text{trace}(W \Sigma \Lambda \Sigma') \text{trace}(C C')^{-1} \\
 &+ E \left[h^2 \left(\chi_{pq+4}^2 (\text{trace}(\delta' \delta)) \right) \right] \text{trace}(\delta' W \delta). \quad (50-3)
 \end{aligned}$$

اثبات: برای سهولت در نوشتن، فرض کنید

$$\eta = \hat{B} - B, \quad \xi = \tilde{B} - B$$

$$\xi \sim MN_{q \times k}(\delta, (\Sigma \Lambda \Sigma') \otimes (C C')^{-1}) \text{ و } \eta \sim_{q \times k}(-\delta, 0)$$

بنابراین با روشی مشابه قضیه ۳-۴ می توان نوشت:

$$\begin{aligned}
 R(\hat{B}_S, B; W) &= E\{\text{trace}(\eta' W \eta)\} + 2E\{(h \|\xi\|_{\Xi_1, \Xi_2}^2) \text{trace}(\xi' W \eta)\} \\
 &+ E\{(h^2 \|\xi\|_{\Xi_1, \Xi_2}^2) \text{trace}(\xi' W \xi)\}.
 \end{aligned}$$

حال با استفاده از قضایای ۱-۳، ۳-۳ و ۳-۳ و ۱-۱ داریم:

$$\begin{aligned}
 E\{\text{trace}(\eta' W \eta)\} &= \text{trace}(\delta' W \delta) \\
 E\{h(\|\xi\|^2) \text{trace}(\xi' W \eta)\} &= -E \left[h \left(\chi_{pq+2}^2 (\text{trace}(\delta' \delta)) \right) \right] \text{trace}(\delta' W \delta) \\
 E\{h^2 \|\xi\|^2 \text{trace}(\xi' W \xi)\} &= E \left[h^2 \left(\chi_{pq+2}^2 (\text{tr}(\delta' \delta)) \right) \right] \text{tr}(W \Sigma \Lambda \Sigma') \text{tr}(C C')^{-1} \\
 &+ E \left[h^2 \left(\chi_{pq+4}^2 (\text{trace}(\delta' \delta)) \right) \right] \text{trace}(\delta' W \delta).
 \end{aligned}$$

□

و اثبات کامل است.

حال بنا بر مثال ۱-۳، مخاطره برآوردگرهای دیگر را با کمک $R(\hat{B}_S, B; W)$ به دست می آوریم.

بنابراین داریم:

$$R(\hat{B}, B; W) = \text{trace}(\delta' W \delta). \quad (51-3)$$

$$R(\tilde{B}, B; W) = \text{trace}(W \Sigma \Lambda \Sigma') \text{trace}(C C')^{-1}. \quad (52-3)$$

برای محاسبه مخاطره برآوردگرهای استاین و استاین مثبت با فرض $\Delta = \text{trace}(\delta' \delta)$ می توان نوشت

$$\begin{aligned} R(\hat{B}^s, B; W) &= R(\tilde{B}, B; W) \\ &\quad - (pq - 2) \text{trace}(W \Sigma \Lambda \Sigma') \text{trace}(C C')^{-1} \\ &\quad \times \left[2E \left(\chi_{pq+2}^{-2}(\Delta) \right) - (pq - 2)E \left(\chi_{pq+2}^{-4}(\Delta) \right) \right] \\ &\quad + \text{trace}(\delta' W \delta) ((pq)^2 - 4)E \left(\chi_{pq+4}^{-4}(\Delta) \right). \end{aligned} \quad (53-3)$$

و با فرض $d = pq - 2$ داریم:

$$\begin{aligned} R(\hat{B}^{s+}, B; W) &= R(\hat{B}^s, B; W) \\ &\quad + 2E \left[\left(1 - d \chi_{pq+2}^{-2}(\Delta) \right) I_{(\chi_{pq+2}^2(\Delta))} < d \right] \text{trace}(\delta' W \delta) \\ &\quad - E \left[\left(1 - d \chi_{pq+2}^{-2}(\Delta) \right)^2 I_{(\chi_{pq+2}^2(\Delta))} \right. \\ &\quad \left. < d \right] \text{trace}(W \Sigma \Lambda \Sigma') \text{trace}(C C')^{-1} \\ &\quad - E \left[\left(1 - d \chi_{pq+4}^{-2}(\Delta) \right)^2 I_{(\chi_{pq+4}^2(\Delta))} < d \right] \text{trace}(\delta' W \delta) \end{aligned} \quad (54-3)$$

۴-۴-۳ مقایسه برآوردگرها

در این قسمت با استفاده از توابع مخاطره بدست آمده در ۳-۴-۳، رفتار برآوردگرهای مورد نظر را با یکدیگر مورد بررسی قرار می‌دهیم. در این رابطه تابع مخاطره برآوردگرها را با رسم نمودار نیز نمایش می‌دهیم. برای رسم نمودارها فرض شده که $p = 3, q = 1, k = 2, m = 3, A = (0, 1, -1)'$

$$C = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, S = \text{Diag}(0/3, 2, 5), \Lambda = \text{Diag}(0/2, 2, 10)$$

مقایسه برآوردگرهای $\hat{B}$ و $\tilde{B}$

به ازای وزن $W = I$ ، از رابطه‌های (۵۱-۳) و (۵۲-۳) داریم:

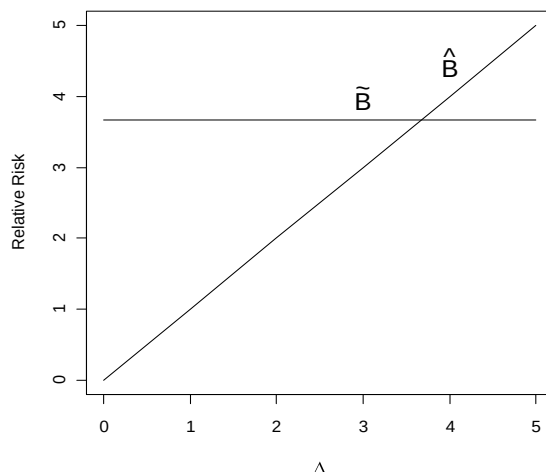
$$\begin{aligned} \mathcal{D}_1 &= R(\tilde{B}, B; I) - R(\hat{B}, B; I) \\ &= \text{trace}(\Sigma \Lambda \Sigma') \text{trace}(CC')^{-1} - \text{trace}(\delta' \delta) \\ &= \text{trace}(\Sigma \Lambda \Sigma') \text{trace}(CC')^{-1} - \Delta. \end{aligned} \quad (۵۵-۳)$$

تحت محدودیت (۴۷-۳)، $\delta = 0$ ، لذا $\mathcal{D}_1 > 0$ و بنابراین $\hat{B} > \tilde{B}$

همچنین از رابطه (۵۵-۳) داریم

$$\mathcal{D}_1 > 0 \iff \Delta < \text{trace}(\Sigma \Lambda \Sigma') \text{trace}(CC')^{-1}.$$

در شکل (۲-۳) به منظور مقایسه برآوردگرهای $\hat{B}$ و $\tilde{B}$ نمودار توابع مخاطره آنها را بر حسب Δ رسم کرده‌ایم. محل تلاقی دو نمودار نقطه $\text{trace}(\Sigma \Lambda \Sigma') \text{trace}(CC')^{-1}$ می‌باشد. مشاهده می‌شود که مخاطره برآوردگر درست‌نمایی ماکزیمم محدود نشده ثابت است، در حالیکه مخاطره برآوردگر درست‌نمایی ماکزیمم محدود شده تحت درستی قید (۴۷-۳) صفر، و با افزایش Δ یعنی با دور شدن از درستی فرض افزایش یافته و از کنترل خارج می‌شود.



شکل (۲-۳) نمودار توابع مخاطره برآوردهای محدودنشده و محدودشده

مقایسه برآوردهای $\hat{B}^s$ و $\tilde{B}$

به ازای وزن $W = I$ ، و روابط (۵۲-۳)، (۵۳-۳) می توان نوشت

$$\begin{aligned} \mathcal{D}_2 &= R(\tilde{B}, B; I) - R(\hat{B}^s, B; I) \\ &= (pq - 2) \text{trace}(\Sigma \Lambda \Sigma') \text{trace}(CC')^{-1} \\ &\quad \times \left[2E\left(\chi_{pq+2}^{-2}(\Delta)\right) - (pq - 2)E\left(\chi_{pq+2}^{-4}(\Delta)\right) \right] \\ &\quad - \Delta((pq)^2 - 4)E\left(\chi_{pq+4}^{-4}(\Delta)\right). \end{aligned} \quad (57-3)$$

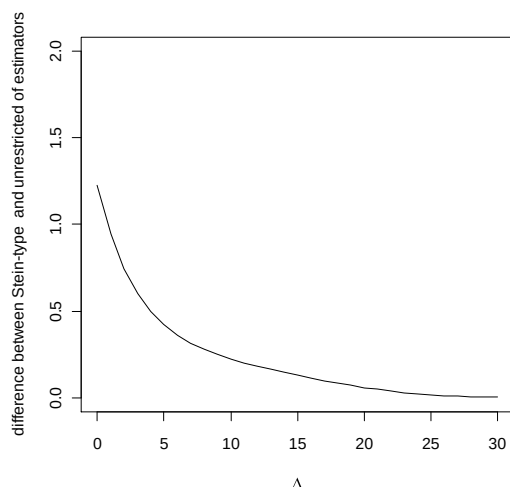
حال از آنجا که رابطه (۵۷-۳) یک تابع نافزایشی بر حسب Δ می باشد و در نقطه $\Delta = 0$ مقدار ماکزیمم

$$\hat{B}^s > \tilde{B} \text{ و } \mathcal{D}_2 > 0 \text{ بنابراین}$$

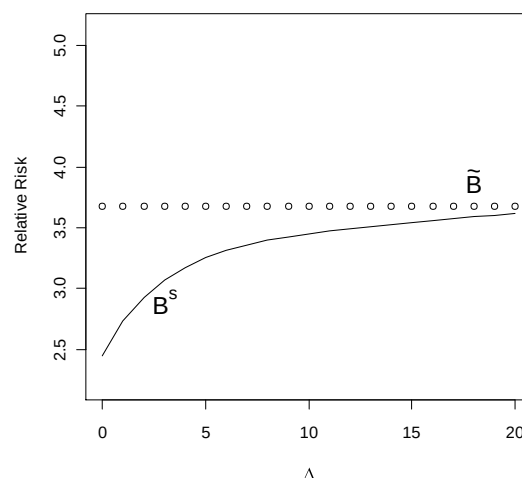
در شکل (۳-۳) که اختلاف اندازه مخاطره برآوردهای $\hat{B}^s$ و $\tilde{B}$ را نشان می دهد، مشاهده می شود که

این اختلاف بر حسب Δ همواره مثبت است. و در شکل (۴-۳) که نمودار تابع مخاطره آنها را نمایش می -

دهد، ملاحظه می‌شود که مخاطره $\hat{B}^S$ بر حسب Δ همواره کمتر از مخاطره $\tilde{B}$ است. در واقع برآوردگر نوع استاین همواره بر برآوردگر محدودنشده برتری دارد.



شکل (۳-۳) نمودار اختلاف اندازه مخاطره برآوردگرهای محدودنشده و نوع استاین



شکل (۴-۳) نمودار تابع مخاطره برآوردگرهای محدودنشده و نوع استاین

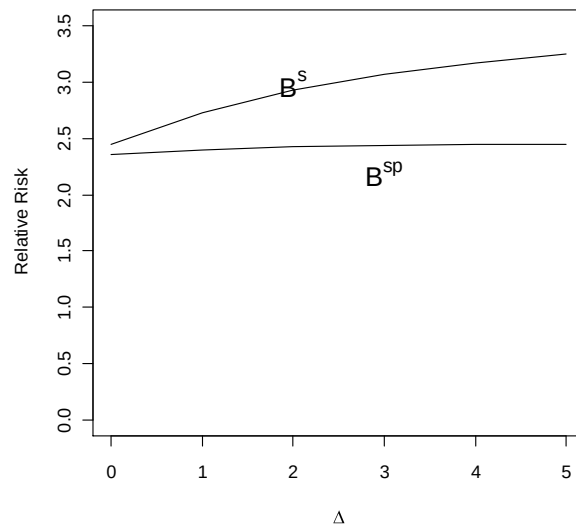
مقایسه برآوردگرهای $\hat{B}^{S+}$ و $\hat{B}^S$

این مقایسه را به ازای ماتریس وزن دلخواه W انجام می‌دهیم. بنابراین از (۵۳-۳) و (۵۴-۳) داریم

$$\begin{aligned} \mathcal{D}_3 &= R(\hat{B}^S, B; W) - R(\hat{B}^{S+}, B; W) \\ &= + E \left[\left(1 - d\chi_{pq+2}^{-2}(\Delta) \right)^2 I_{(\chi_{pq+2}^2(\Delta) < d)} \right] \text{trace}(W\Sigma\Lambda\Sigma') \text{trace}(CC')^{-1} \\ &\quad - 2E \left[\left(1 - d\chi_{pq+2}^{-2}(\Delta) \right) I_{(\chi_{pq+2}^2(\Delta) < d)} \right] \text{trace}(\delta'W\delta) \\ &\quad + E \left[\left(1 - d\chi_{pq+4}^{-2}(\Delta) \right)^2 I_{(\chi_{pq+4}^2(\Delta) < d)} \right] \text{trace}(\delta'W\delta) \end{aligned} \quad (58-3)$$

حال با استدلالی همانند بخش قبل خواهیم داشت $\mathcal{D}_3 > 0$ و بنابراین $\hat{B}^{S+} > \hat{B}^S$

شکل (۵-۳) نمودار تابع مخاطره برآوردهای $\hat{B}^s$ و $\hat{B}^{s+}$ را نشان می‌دهد. نمودار حاکی از این مطلب است که برآوردهای نوع استاین مثبت همواره از برآوردهای نوع استاین بهتر است.



شکل (۵-۳) نمودار تابع مخاطره برآوردهای نوع استاین و استاین مثبت

فصل چهارم

برآورد انقباضی در مدل‌های بیضی‌گون ماتریسی

۴-۱ مقدمه

اگرچه توزیع‌های نرمال تک‌متغیره و چندمتغیره نقش مهمی در آمار کلاسیک و بیزی ایفا می‌کنند، بسیاری از پدیده‌های طبیعی از مدل نرمال پیروی نمی‌کنند و از مدل‌های یک یا چندمتغیره غیرنرمال برای مدل‌بندی آنها استفاده می‌شود. در این میان به خانواده توزیع‌های کروی^۱ و بیضی‌گون^۲ (توزیع‌های با منحنی تراز بیضی شکل)، در دو دهه اخیر، توجه بیشتری شده است. توزیع‌های بیضی‌گون، بسیاری از خواص توزیع نرمال چندمتغیره را دارا هستند. بنابراین بسیاری از روش‌های مرسوم تحلیل داده‌ها را می‌توان به طور مستقیم در خانواده توزیع‌های بیضی‌گون به کار برد.

مقاله‌هایی در این خصوص توسط ماکسول^۳ (۱۸۶۰)، بارتلت^۴ (۱۹۳۴) و هارتمن و ویتنر^۵ (۱۹۴۰) ارائه شده‌اند. تحقیقات تازه در خصوص کاربردهای این توزیع‌ها در مهندسی، توسط بلیک و توماس^۶ (۱۹۶۸) و مک‌گرو و واگنر^۷ (۱۹۶۸) شروع و توسط چو (۱۹۷۳) دنبال شد. اولین مجموعه کامل درباره این توزیع‌ها، توسط کلکر^۸ (۱۹۷۰) ارائه شده و پس از آن بررسی رفتار منظم این خانواده از توزیع‌ها، توسط کمبانیس^۹ و همکاران (۱۹۸۱) انجام گرفت. چندین مقاله، با محتوای مروری بر توزیع‌های بیضی‌گون و استنباط آماری در این زمینه، از چیلسکی^{۱۰} (۱۹۸۱)، اندرسن و فنگ^{۱۱} (۱۹۹۱) و اندرسن (۱۹۹۳) وجود دارد. جانسون^{۱۲} و کاتز (۱۹۷۲) بحث کاملی بر روی این خانواده از توزیع‌ها، به خصوص

¹ Spherical distribution

² Elliptically contoured distribution

³ Maxwell

⁴ Bartlett

⁵ Hartman and Wintner

⁶ Blake and Thomas

⁷ McGraw and Wagner

⁸ Kelker

⁹ Cambanis

¹⁰ Chmielewski

¹¹ Anderson and Fang

¹² Johnson

توزیع استیودنت چندمتغیره انجام دادند. فنگ و اندرسن (۱۹۹۰)، چهل مقاله تحقیقاتی بر روی این توزیع‌ها گردآوری کردند. مطالعه جامعی بر روی این توزیع‌ها و استنباط آماری آن‌ها را می‌توان در مجموعه کتاب‌های میرهد^{۱۳} (۱۹۸۲)، فنگ و ژانگ^{۱۴} (۱۹۹۰)، فنگ و همکاران (۱۹۹۰) و گوپتا و وارگا^{۱۵} (۱۹۹۵) یافت. کاربردهای جدید توزیع‌های بیضی‌گون در رگرسیون و برآورد انقباضی را می‌توان در آرشی (۲۰۱۰) یافت.

در این فصل توجه خود را به توزیع‌های بیضی‌گون ماتریسی معطوف کرده و نتایج فصل ۳ را برای این خانواده از توزیع‌ها تعمیم می‌دهیم.

۲-۴ مختصری درباره توزیع‌های بیضی‌گون ماتریسی

از آنجا که توزیع‌های بیضی‌گون، صورت تعمیم‌یافته توزیع‌های کروی می‌باشند، لذا ابتدا توزیع‌های کروی ماتریسی را معرفی می‌کنیم.

۱-۲-۴ توزیع‌های کروی ماتریسی

در این زیربخش چهار کلاس از توزیع‌های کروی ماتریسی یعنی توزیع‌های چپ-کروی^{۱۶}، کروی چندمتغیره^{۱۷}، کروی برداری^{۱۸} و کروی‌گون^{۱۹} را مطابق با فنگ و ژانگ (۱۹۹۰) معرفی کرده و برخی از ویژگی‌های آن‌ها را ارائه می‌کنیم.

¹³ Muirhead

¹⁴ Zhang

¹⁵ Gupta and Varga

¹⁶ Left-spherical distribution

¹⁷ Multivariate spherical

¹⁸ Vector-spherical

¹⁹ Spherical distribution

تعریف ۱-۴ ماتریس تصادفی X در اندازه $n \times p$ دارای توزیع چپ-کروی است اگر X و ΛX هم‌توزیع باشند به عبارتی داریم

$$X \stackrel{d}{=} \Lambda X$$

که در آن $\Lambda \in O(n)$

لم ۱-۴ (فنگ و ژانگ، ۱۹۹۰) فرض کنید X دارای توزیع چپ کروی و $\varphi_X(T)$ تابع مشخصه آن باشد؛

که در آن T ماتریسی در اندازه $n \times p$ می‌باشد. در این صورت

$$\varphi_X(T) = \phi(T'T).$$

به عبارتی تابع مشخصه، تابعی از $T'T$ است.

اثبات: با استفاده از تعریف تابع مشخصه داریم:

$$\begin{aligned} \varphi_X(T) &= E[\exp(\text{itr}(T'X))] \\ &= E[\exp(\text{itr}(T'\Lambda' \Lambda X))], \quad \Lambda \in O(n) \\ &= E[\exp(\text{itr}(\Lambda T)' \Lambda X)]. \end{aligned}$$

بنا بر تعریف (۱-۴)، $X \stackrel{d}{=} \Lambda X$ لذا

$$\begin{aligned} \varphi_X(T) &= E[\exp(\text{itr}(\Lambda T)' X)] \\ &= \varphi_X(\Lambda T) \end{aligned}$$

بنابراین $\varphi(T)$ دارای خاصیت پایایی ماکزیمال بر روی تبدیلات متعامد می‌باشد و اثبات کامل است. $\square$

اگر X در اندازه $n \times p$ دارای توزیع چپ-کروی، با تابع مشخصه $\phi(T'T)$ باشد، آن را با نماد

$$X \sim LS_{n \times p}(\phi) \text{ و یا به طور ساده‌تر با نماد } X \sim LS(\phi) \text{ نمایش می‌دهیم.}$$

تعریف ۲-۴ اگر X در اندازه $n \times p$ دارای توزیع چپ-کروی، و $X'X = I_p$ ، آن گاه X دارای توزیع یکنواخت است و با نماد $X \sim \mathcal{U}_{n \times p}$ نمایش داده می‌شود.

تعریف ۳-۴ اگر X' دارای توزیع چپ-کروی باشد، در این صورت X دارای توزیع راست-کروی^{۲۰} بوده و با نماد $X \sim RS_{n \times p}(\phi)$ نمایش داده می‌شود. همچنین X دارای توزیع کروی گون است، اگر هم چپ کروی و هم راست کروی باشد و با نماد $X \sim SS_{n \times p}(\phi)$ نمایش داده می‌شود.

تعریف‌های دیگری برای توزیع‌های راست کروی و کروی گون توسط گوپتا و ناگار^{۲۱} (۱۹۹۹) به صورت زیر ارائه شده‌اند:

تعریف ۴-۴ ماتریس تصادفی X در اندازه $n \times p$ دارای توزیع

$$1. \text{ راست-کروی است اگر } X \stackrel{d}{=} X\Lambda \text{ که در آن } \Lambda \in O(p)$$

$$2. \text{ کروی گون است اگر } X \stackrel{d}{=} \Gamma X \Lambda \text{ که در آن } \Lambda \in O(p), \Gamma \in O(n)$$

قضیه ۱-۴ (فنگ و ژانگ، ۱۹۹۰). فرض کنید Y در اندازه $n \times p$ چپ-کروی، و $Y'Y \stackrel{d}{=} X'X$ در این

$$\text{صورت } X \stackrel{d}{=} Y$$

قضیه ۲-۴ (گوپتا و ناگار، ۱۹۹۹). فرض کنید $X \sim LS_{n \times p}(\phi)$ در این صورت ماتریس تصادفی A در اندازه $p \times p$ وجود دارد، به طوری که

$$X \stackrel{d}{=} UA,$$

که در آن $U \sim \mathcal{U}_{n \times p}$ مستقل از A می‌باشد.

²⁰ Right-spherical distribution
²¹ Nagar

اثبات: برای X در اندازه $n \times p$ ، ماتریس تصادفی A در اندازه $p \times p$ وجود دارد به طوری که $A'A \stackrel{d}{=} I_p$ و همچنین A مستقل از $U \sim \mathcal{U}_{n \times p}$ است.

حال فرض کنید $Y = UA$. در این صورت Y دارای توزیع چپ-کروی و

$$Y'Y = A'U'UA = A'A \stackrel{d}{=} I_p.$$

بنابراین با استفاده از قضیه ۴-۱، داریم $Y \stackrel{d}{=} X$ و اثبات کامل است. $\square$

در این قسمت فرض کنید A یک ماتریس معین نامنفی با بعد p ، و $\lambda_1 \geq \dots \geq \lambda_p \geq 0$ مقادیر ویژه آن باشند. همچنین فرض کنید $\lambda(A) = \text{diag}(\lambda_1, \dots, \lambda_p)$

قضیه ۴-۳ (گوپتا و ناگار، ۱۹۹۹). اگر X در اندازه $n \times p$ و $n \leq p$ دارای توزیع کروی‌گون باشد، آنگاه

$$X \stackrel{d}{=} U\Lambda V,$$

که در آن $U \sim \mathcal{U}_{n \times p}$ ، $V \sim \mathcal{U}_{p \times p}$ و $\Lambda = \lambda \left((X'X)^{\frac{1}{2}} \right)$ مستقل از یکدیگرند.

قضیه ۴-۴ (گوپتا و ناگار، ۱۹۹۹). اگر X دارای توزیع کروی باشد، در این صورت تابع مشخصه آن به فرم $\phi(\lambda(T'T))$ می‌باشد.

تعریف ۴-۵ ماتریس تصادفی X در اندازه $n \times p$ دارای توزیع کروی چندمتغیره است، اگر تابع مشخصه آن به فرم $\phi(t_1't_1, \dots, t_p't_p)$ باشد، که در آن $T = (t_1, \dots, t_p)$. در این صورت آن را با نماد $X \sim MS_{n \times p}(\phi)$ نمایش می‌دهند.

تعریف ۴-۶ فرض کنید $X = (x_1, \dots, x_p)$ یک ماتریس $n \times p$ و $vec(X) = (x'_1, \dots, x'_p)$.

اگر $vec(X) \sim S_{np}(\phi)$ در این صورت X دارای توزیع کروی برداری است و با نماد $X \sim VS_{n \times p}(\phi)$ نمایش می‌دهند.

۴-۲-۲ ارتباط بین کلاس‌های توزیع‌های کروی ماتریسی

فرض کنید

$$\mathcal{F}_1 = \{F(X): X \sim LS(\phi)\},$$

$$\mathcal{F}_2 = \{F(X): X \sim MS(\phi)\},$$

$$\mathcal{F}_3 = \{F(X): X \sim VS(\phi)\},$$

$$\mathcal{F}_s = \{F(X): X \sim S(\phi)\},$$

که در آن X در اندازه $n \times p$ است و $F(X)$ تابع توزیع X را نشان می‌دهد.

لم ۴-۲ (فنگ و ژانگ، ۱۹۹۰). رابطه زیر بین کلاس توزیع‌های کروی ماتریسی برقرار است:

$$\mathcal{F}_1 \supset \mathcal{F}_2 \supset \mathcal{F}_3 \text{ و } \mathcal{F}_1 \supset \mathcal{F}_s \supset \mathcal{F}_3.$$

بنابراین می‌توان نتیجه گرفت که کلاس $\mathcal{F}_1$ بزرگترین کلاس توزیع‌های کروی است.

فرض کنید $X \in \mathcal{F}_i$ که در آن s یا $i = 1, 2, 3$. به طور کلی X الزاماً دارای تابع چگالی نیست، ولی

اگر X دارای تابع چگالی باشد، در این صورت:

۱. $X \in \mathcal{F}_1$ اگر و فقط اگر تابع چگالی X به صورت $f(X'X)$ باشد.

۲. $X \in \mathcal{F}_2$ اگر و فقط اگر تابع چگالی X به صورت $f(x'_1 x_1, \dots, x'_p x_p)$ باشد.

۳. $X \in \mathcal{F}_3$ اگر و فقط اگر تابع چگالی X به صورت $f(tr(X'X))$ باشد.

۴. $X \in \mathcal{F}_s$ اگر و فقط اگر تابع چگالی X به صورت $f(\lambda(X'X))$ باشد.

۳-۲-۴ توزیع‌های بیضی‌گون ماتریسی

فرض کنید X در اندازه $n \times p$ یک ماتریس تصادفی باشد، و $X \in \mathcal{F}_i$ که در آن s یا $i = 1, 2, 3$. همچنین فرض کنید M و B ماتریس‌های ثابت، به ترتیب با ابعاد $m \times p$ و $n \times m$ باشند. اگر

$$Y \stackrel{d}{=} M + B'X, \quad \Sigma = B'B$$

در این صورت، Y دارای توزیع بیضی‌گون ماتریسی است و با نماد

$$1. \quad Y \sim ELS_{m \times p}(M, \Sigma, \phi) \text{ نشان می‌دهیم، اگر } X \sim LS(\phi)$$

$$2. \quad Y \sim EMS_{m \times p}(M, \Sigma, \phi) \text{ نشان می‌دهیم، اگر } X \sim MS(\phi)$$

$$3. \quad Y \sim EVS_{m \times p}(M, \Sigma, \phi) \text{ نشان می‌دهیم، اگر } X \sim VS(\phi)$$

$$4. \quad Y \sim ESS_{m \times p}(M, \Sigma, \phi) \text{ نشان می‌دهیم، اگر } X \sim SS(\phi)$$

لم ۳-۴ (فنگ و ژانگ، ۱۹۹۰).

۱. تابع مشخصه $Y \sim ELS_{m \times p}(M, \Sigma, \phi)$ به صورت $\exp(iT'M) \phi(T'\Sigma T)$ و تابع چگالی

$$\text{آن به صورت } \left| \Sigma \right|^{-\frac{p}{2}} f\left((Y - M)' \Sigma^{-1} (Y - M)\right) \text{ می‌باشد.}$$

۲. تابع مشخصه $Y \sim EMS_{m \times p}(M, \Sigma, \phi)$ به صورت $\exp(iT'M) \phi(t'_1 \Sigma t_1, \dots, t'_p \Sigma t_p)$ و تابع چگالی آن به صورت

$$\left| \Sigma \right|^{-\frac{p}{2}} f(y'_1 \Sigma^{-1} y_1, \dots, y'_p \Sigma^{-1} y_p) \text{ می‌باشد.}$$

۳. تابع مشخصه $Y \sim EVS_{m \times p}(M, \Sigma, \phi)$ به صورت $\exp(iT'M) \phi(\text{tr}(T'\Sigma T))$ و تابع چگالی آن به صورت

$$\left| \Sigma \right|^{-\frac{p}{2}} f(\text{tr}(\Sigma^{-1} (Y - M)(Y - M)')) \text{ می‌باشد.}$$

۴. تابع مشخصه $Y \sim ESS_{m \times p}(M, \Sigma, \phi)$ به صورت $\exp(it'TM) \phi(\lambda(T'\Sigma T))$ و تابع

$$\text{چگالی آن به صورت } |\Sigma|^{-\frac{p}{2}} f\left(\lambda((Y-M)'\Sigma^{-1}(Y-M))\right) \text{ می‌باشد.}$$

که در آن‌ها $T = (t_1, \dots, t_p)$ در اندازه $m \times p$ می‌باشد.

قضیه ۴-۵ (گوپتا و ناگار، ۱۹۹۹). فرض کنید $Y \sim ELS_{m \times p}(M, \Sigma, \phi)$ و C و D ماتریس‌های ثابت

و در اندازه‌های $m \times l$ و $l \times p$ باشند. در این صورت:

$$Z = C'Y + D \sim ELS_{l \times p}(C'M + D, C'\Sigma C, \phi).$$

اثبات: از تعریف تابع مشخصه داریم:

$$\begin{aligned} \varphi_Z(T) &= E[\exp(itr(T'Z))] = E[\exp(itr(T'(C'Y + D)))] \\ &= \exp(itr(T'D)) E[\exp(itr(T'C'Y))] \\ &= \exp(itr(T'D)) \varphi_Y(TC) \\ &= \exp(itr(T'D)) \exp(itr(TC)'M) \phi((TC)'\Sigma(TC)) \\ &= \exp(itr(T'(C'M + D))) \phi(T'(C'\Sigma C)T) \end{aligned}$$

□

و اثبات کامل است.

نتیجه ۴-۱ افراز زیر را در نظر بگیرید

$$Y = \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix}, \quad M = \begin{bmatrix} M_1 \\ M_2 \end{bmatrix}, \quad \Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}$$

که در آن Y_1 و M_1 در اندازه $q \times p$ و Σ_{11} ماتریس متقارن با بعد q است. در این صورت

$$Y_1 \sim ELS_{q \times p}(M_1, \Sigma_{11}, \phi).$$

□

اثبات: با قرار دادن $D = 0$ و $C' = (I_q \quad 0)$ در قضیه ۴-۵ اثبات کامل می‌شود.

قضیه ۴-۶ (کولو و وان‌رزن، ۲۰۰۵). فرض کنید X یک ماتریس تصادفی در اندازه $p \times n$ باشد. در این صورت سه جمله زیر هم‌ارز هستند:

$$۱. \text{تابع مشخصه } vecX \text{ به صورت } \phi(vec'TvecT) = \phi(tr(T'T)) \text{ است.}$$

۲. X دارای نمایش احتمالی به صورت $X \stackrel{d}{=} RU$ است، که در آن $R \geq 0$ مستقل از U ، و $vecU$ دارای توزیع یکنواخت است.

$$۳. \Gamma \in O(pn) \text{ که در آن } vecX \stackrel{d}{=} \Gamma' vecX.$$

در ادامه تعریف دیگری از توزیع بیضی‌گون ماتریسی ارائه می‌دهیم که مبنای استنتاج نتایج بعدی در این فصل می‌باشد. تعریف زیر تعمیم تعریف ۴-۴ به بیضی‌گون است.

تعریف ۴-۷ فرض کنید ماتریس تصادفی X در اندازه $r \times s$ دارای توزیع کروی‌گون و $\Sigma = \tau\tau'$ ، $\Omega = \gamma\gamma'$ ماتریس‌های معین نامنفی، به ترتیب با ابعاد $n \times n$ و $p \times p$ باشند، که در آن τ در اندازه $n \times r$ و γ در اندازه $p \times s$ است. در این صورت، ماتریس Y در اندازه $n \times p$ دارای توزیع بیضی‌گون با پارامترهای M ، Σ و Ω است و با نماد $Y \sim MEC_{n \times p}(M, \Sigma \otimes \Omega, \phi)$ یا $Y \sim MEC_{n \times p}(M, \Sigma, \Omega, \phi)$ نمایش می‌دهند، اگر

$$Y \stackrel{d}{=} M + \tau X \gamma', \quad (۱-۴)$$

که در آن M در اندازه $m \times p$ می‌باشد.

قضیه ۴-۷ (کولو و وان‌رزن، ۲۰۰۵). فرض کنید $Y \sim MEC_{n \times p}(M, \Sigma, \Omega, \phi)$ که در آن M در اندازه $n \times p$ ، $\Sigma = \tau\tau'$ و $\Omega = \gamma\gamma'$ در این صورت:

تابع مشخصه Y به صورت زیر است:

$$\begin{aligned}\varphi_Y(T) &= \exp(\text{ivec}'T \text{vec}M) \phi(\text{vec}'T(\Omega \otimes \Sigma)\text{vec}T) \\ &= \exp(\text{itr}(T'M)\phi(\text{tr}(T'\Sigma T\Omega)))\end{aligned}\quad (2-4)$$

همچنین، اگر Y دارای تابع چگالی باشد، تابع چگالی آن به صورت زیر است:

$$g(Y) = d_{n,p} |\Omega|^{-\frac{n}{2}} |\Sigma|^{-\frac{p}{2}} f(\Omega^{-1}(Y - M)'\Sigma^{-1}(Y - M)) \quad (3-4)$$

که در آن $f(\cdot)$ تابع مولد چگالی است. در این صورت می‌نویسیم $Y \sim MEC_{n \times p}(M, \Sigma \otimes \Omega, f)$.

اثبات: با استفاده از تعریف ۶-۱ برای تابع مشخصه،

$$\varphi_Y(T) = E[\exp(\text{itr}(T'Y))] = E[\exp(\text{ivec}'T\text{vec}Y)].$$

حال از تعریف ۷-۴ خواهیم داشت:

$$\begin{aligned}\varphi_Y(T) &= E[\exp(\text{ivec}'T\text{vec}(M + \tau X\gamma))] \\ &= \exp(\text{ivec}'T\text{vec}M) E[\exp(\text{itr}(T'\tau X\gamma'))] \\ &= \exp(\text{ivec}'T\text{vec}M) E[\exp(\text{itr}(\tau'T\gamma)'X)] \\ &= \exp(\text{ivec}'T\text{vec}M) \varphi_X(\tau'T\gamma).\end{aligned}$$

در این قسمت از قضیه ۶-۴ نتیجه می‌گیریم

$$\begin{aligned}\varphi_Y(T) &= \exp(\text{ivec}'T\text{vec}M) \phi(\text{vec}'(\tau'T\gamma)\text{vec}(\tau'T\gamma)) \\ &= \exp(\text{ivec}'T\text{vec}M) \phi(\text{vec}'T(\gamma \otimes \tau)(\gamma' \otimes \tau')\text{vec}T) \\ &= \exp(\text{ivec}'T\text{vec}M) \phi(\text{vec}'T(\Omega \otimes \Sigma)\text{vec}T) \\ &= \exp(\text{itr}(T'\mu)) \phi(\text{tr}(T'\Sigma T\Omega)).\end{aligned}$$

□

و اثبات کامل است.

قضیه ۴-۸ (کولو و وان‌رزن، ۲۰۰۵)

فرض کنید $Y \sim MEC_{n \times p}(M, \Sigma \otimes \Omega, f)$ و A, B و C ماتریس‌هایی در اندازه $p \times m, q \times n$ و $q \times m$ باشند. در این صورت

$$AYB + C \sim MEC_{q \times m}(AMB + C, (A\Sigma A') \otimes (B'\Omega B), f). \quad (۴-۴)$$

اثبات: از تعریف ۴-۷، $Y \stackrel{d}{=} M + \tau X \gamma'$ که در آن $\Sigma = \tau \tau'$ و $\Omega = \gamma \gamma'$. بنابراین

$$AYB + C \stackrel{d}{=} AMB + C + A\tau X \gamma' B.$$

حال فرض می‌کنیم $Z = AMB + C + A\tau X \gamma' B$ و با استفاده از تابع مشخصه توزیع Z را به‌دست می‌آوریم. لذا

$$\begin{aligned} \varphi_Z(T) &= E[\exp(\text{itr}(T'Z))] \\ &= \exp\left(\text{itr}(T'(AMB + C))\right) E[\exp(\text{itr}(T'A\tau X \gamma' B))] \\ &= \exp\left(\text{itr}(T'(AMB + C))\right) E[\exp(\text{itr}(\gamma' B T' A \tau X))] \\ &= \exp\left(\text{itr}(T'(AMB + C))\right) \varphi_X(\tau' A' T B' \gamma) \\ &= \exp\left(\text{itr}(T'(AMB + C))\right) \phi(\text{vec}'(\tau' A' T B' \gamma) \text{vec}(\tau' A' T B' \gamma)) \\ &= \exp\left(\text{itr}(T'(AMB + C))\right) \phi(\text{vec}' T (B' \gamma \otimes A \tau) (\gamma' B \otimes \tau' A') \text{vec} T) \\ &= \exp\left(\text{itr}(T'(AMB + C))\right) \phi(\text{vec}' T ((B' \Omega B) \otimes (A \Sigma A')) \text{vec} T). \end{aligned}$$

□

و در نهایت اثبات با استفاده از قضیه ۴-۷ کامل است.

قضیه ۹-۴ (گوپتا و وارگا، ۱۹۹۵). فرض کنید $Y \sim MEC_{n \times p}(M, \Sigma \otimes \Omega, f)$ که در آن تابع چگالی احتمال $g(Y)$ ، به صورت زیر می‌باشد

$$g(Y) = |\Sigma|^{-\frac{p}{2}} |\Omega|^{-\frac{n}{2}} h\{tr[\Omega^{-1}(Y - M)' \Sigma^{-1}(Y - M)]\}.$$

اگر $h(t), t \in [0, \infty)$ تبدیل معکوس لاپلاس داشته باشد (که آن را با نماد $\mathcal{L}^{-1}[h(t)]$ نشان می‌دهند)، در این صورت

$$g(Y) = \int_0^\infty f_{MN_{n \times p}(M, Z^{-1} \Sigma \otimes \Omega)}(Y) \omega(z) dz, \quad (۴-۴)$$

که در آن $f_{MN_{n \times p}(M, Z^{-1} \Sigma \otimes \Omega)}(Y)$ نشان‌دهنده تابع چگالی احتمال

$$Y \sim MN_{n \times p}(M, Z^{-1} \Sigma \otimes \Omega) \text{ است و}$$

$$\omega(z) = (2\pi)^{\frac{np}{2}} z^{-\frac{np}{2}} \mathcal{L}^{-1}[h(2t)] \quad (۵-۴)$$

حال اگر $\omega(z)$ یک تابع، و تعریف کنیم: $u(z) = \frac{1}{z^2} \omega\left(\frac{1}{z}\right)$ ، در این صورت

$$g(Y) = \int_0^\infty f_{MN_{n \times p}(M, Z \Sigma \otimes \Omega)}(Y) u(z) dz. \quad (۶-۴)$$

قضیه ۱۰-۴* (گوپتا و وارگا، ۱۹۹۵). فرض کنید $\lambda \in \mathbb{R}^{p \times m}$ یک تابع اندازه‌پذیر بورل ماتریسی است.

اگر $Y \sim MN_{n \times p}(M, \Sigma \otimes \Omega)$ ، در این صورت $E(\lambda(Y))$ وجود دارد و آن را با نماد

$$E_{MN_{n \times p}(M, \Sigma \otimes \Omega)}(\lambda(Y))$$
 نشان می‌دهیم. حال اگر $Y \sim MEC_{n \times p}(M, \Sigma \otimes \Omega, f)$ با تابع

چگالی احتمال $g(Y) = \int_0^\infty f_{MN_{n \times p}(M, Z^{-1} \Sigma \otimes \Omega)}(Y) \omega(z) dz$ باشد، و در صورتی که

$$E(\lambda(Y)) \text{ وجود داشته باشد، آن را با نماد } E_{MEC_{n \times p}(M, Z^{-1} \Sigma \otimes \Omega, f)}(\lambda(Y)) \text{ نمایش داده و داریم}$$

$$E_{MEC_{n \times p}(M, \Sigma \otimes \Omega, f)}(\lambda(Y)) = \int_0^\infty E_{MN_{n \times p}(M, z^{-1} \Sigma \otimes \Omega)}(\lambda(Y)) \omega(z) d(z). \quad (7-4)$$

و اگر $g(Y) = \int_0^\infty f_{MN_{n \times p}(M, z \Sigma \otimes \Omega)}(Y) u(z) d(z)$ و $E(\lambda(Y))$ وجود داشته باشد در این صورت

$$E_{MEC_{n \times p}(M, \Sigma \otimes \Omega, f)}(\lambda(Y)) = \int_0^\infty E_{MN_{n \times p}(M, z \Sigma \otimes \Omega)}(\lambda(Y)) \omega(z) d(z). \quad (8-4)$$

اثبات: از تعریف امید ریاضی و قضیه ۹-۴ خواهیم داشت:

$$\begin{aligned} E_{MEC_{n \times p}(M, \Sigma \otimes \Omega, f)}(\lambda(Y)) &= \int \lambda(Y) f_{MEC_{n \times p}(M, \Sigma \otimes \Omega, f)}(Y) d(Y) \\ &= \int \lambda(Y) \int_0^\infty f_{MN_{n \times p}(M, z^{-1} \Sigma \otimes \Omega)}(Y) \omega(z) d(z) d(Y) \\ &= \int_0^\infty \omega(z) \int \lambda(Y) f_{MN_{n \times p}(M, z^{-1} \Sigma \otimes \Omega)}(Y) d(Y) d(z) \\ &= \int_0^\infty \omega(z) E_{MN_{n \times p}(M, z^{-1} \Sigma \otimes \Omega)}(\lambda(Y)) d(z). \end{aligned}$$

□

و اثبات کامل است.

۴-۲-۴ مثال‌هایی از توزیع بیضی‌گون ماتریسی

در این قسمت به عنوان نمونه، تابع چگالی احتمال را برای چند توزیع مهم بیضی‌گون ماتریسی ارائه می‌دهیم.

(۱) توزیع نرمال ماتریسی $(X \sim MN_{n,p}(M, \Sigma, \Omega))$: مطابق تعریف ۱۲-۱

(۲) توزیع t ماتریسی $(X \sim MT_{n,p}(M, \Sigma, \Omega, \nu))$:^{۲۲}

$$f(X) = \frac{|\Omega|^{-\frac{n}{2}} |\Sigma|^{-\frac{p}{2}}}{g_{n,p}} |I_n + \Omega^{-1}(X - M)' \Sigma^{-1}(X - M)|^{-\frac{n+p+v-1}{2}},$$

که در آن M میانگین توزیع، Ω و Σ به ترتیب ماتریس‌های مقیاس سطری و ستونی، v درجه آزادی و

$$g_{n,p} = \frac{(\nu\pi)^{\frac{np}{2}} \Gamma_p\left(\frac{\nu+p-1}{2}\right)}{\Gamma_p\left(\frac{\nu+n+p-1}{2}\right)}.$$

(۳) توزیع پیرسن ماتریسی نوع VII^{۲۳} $(X \sim MPVII_{n,p}(M, \Sigma, \Omega, m, q))$:

$$f(X) = \frac{|\Omega|^{-\frac{n}{2}} |\Sigma|^{-\frac{p}{2}}}{h_{n,p,m,q}} \left| I_n + \frac{1}{q} \Omega^{-1}(X - M)' \Sigma^{-1}(X - M) \right|^{-m},$$

که در آن M میانگین توزیع، Ω و Σ به ترتیب ماتریس‌های مقیاس سطری و ستونی، m و q پارامترهای

توزیع و

$$h_{n,p,m,q} = \frac{(q\pi)^{\frac{np}{2}} \Gamma_p\left(m - \frac{n}{2}\right)}{\Gamma_p(m)}.$$

(۴) توزیع نمایی توانی ماتریسی^{۲۴} $(X \sim MPE_{n,p}(M, \Sigma, \Omega, r, s))$:

$$f(X) = \frac{|\Omega|^{-\frac{n}{2}} |\Sigma|^{-\frac{p}{2}}}{j_{n,p,r,s}} \exp\left(-\frac{r}{2} (tr[\Omega^{-1}(X - M)' \Sigma^{-1}(X - M)])^s\right),$$

که در آن M میانگین توزیع، Ω و Σ به ترتیب ماتریس‌های مقیاس سطری و ستونی، r و s پارامترهای

توزیع و

$$j_{n,p,r,s} = \frac{(2\pi)^{\frac{np}{2}} \Gamma\left(\frac{np}{2s}\right)}{s \Gamma\left(\frac{np}{2}\right) r^{\frac{np}{2s}}}.$$

Matrix variate pearson type-VII²³
Matrix variate power exponential²⁴

(۵) توزیع لاپلاس ماتریسی $(X \sim ML_{n,p}(M, \Sigma, \Omega))$ ^{۲۵}:

$$f(X) = \frac{|\Omega|^{-\frac{n}{2}} |\Sigma|^{-\frac{p}{2}}}{l_{n,p}} \exp\left(-\frac{\sqrt{2}}{2} (tr[\Omega^{-1}(X - M)' \Sigma^{-1}(X - M)])^{\frac{1}{2}}\right),$$

که در آن M میانگین توزیع، Ω و Σ به ترتیب ماتریس‌های مقیاس سطری و ستونی و

$$l_{n,p} = \frac{2\pi^{\frac{np}{2}} \Gamma(np)}{\Gamma\left(\frac{np}{2}\right)}.$$

(۶) توزیع ماتریسی نوع کاتز $(X \sim MK_{n,p}(M, \Sigma, \Omega))$ ^{۲۶}:

$$f(X) = \frac{|\Omega|^{-\frac{n}{2}} |\Sigma|^{-\frac{p}{2}}}{t_{n,p,r}} \exp\left(-\frac{r}{2} (tr[\Omega^{-1}(X - M)' \Sigma^{-1}(X - M)])\right),$$

که در آن M میانگین توزیع، Ω و Σ به ترتیب ماتریس‌های مقیاس سطری و ستونی، S پارامتر توزیع و

$$t_{n,p,r} = \frac{(2\pi)^{\frac{np}{2}}}{r^{\frac{np}{2}}}.$$

۳-۴ برآورد انقباضی در توزیع بیضی‌گون ماتریسی

فرض کنید ماتریس تصادفی X در اندازه $q \times k$ دارای توزیع بیضی‌گون ماتریسی به صورت

$$X \sim MEC_{q \times k}(\theta, \Lambda \otimes I_k, f), \quad (۹-۴)$$

باشد، که در آن θ در اندازه $q \times k$ میانگین توزیع و Λ در اندازه $q \times q$ یک ماتریس معین مثبت با

رتبه q است. مشابه بخش ۳-۳، کلاسی از برآوردهای انقباضی را به صورت زیر تشکیل می‌دهیم:

$$\hat{\theta}_S = \hat{\theta} + h(\| (X - \hat{\theta}) L_2 \|_{\Xi_1, \Xi_2}^2)(X - \hat{\theta}),$$

²⁵ Matrix variate laplace

²⁶ Matrix variate Kotz-type

که در آن $\hat{\theta}$ برآوردگر درست‌نمایی ماکزیمم محدودشده تحت قید $L_1\theta L_2 = d$ بوده و مقدار آن در رابطه (۴-۳) داده شده است.

در واقع این بخش، تعمیم بخش ۳-۳ برای توزیع‌های بیضی‌گون ماتریسی است، و مثال ۱-۳ در اینجا نیز برقرار است.

۴-۳-۱ محاسبه توابع مخاطره

در این قسمت ابتدا لم‌ها و قضایای مطرح شده در ۲-۳-۳ را برای توزیع بیضی‌گون ماتریسی نوشته و سپس به کمک آن‌ها و رابطه (۳۱-۳)، توابع مخاطره را برای $\hat{\theta}_S, \tilde{\theta}, \hat{\theta}, \hat{\theta}^S$ و $\hat{\theta}^{S+}$ به دست می‌آوریم.

لم ۴-۴** تحت مفروضات لم ۲-۳، اگر $X \sim MEC_{q \times k}(\theta, \Lambda \otimes I_k, f)$ ، آنگاه داریم:

$$(\xi, \eta) \sim MEC_{q \times 2m} \left((\delta, -\delta), \begin{pmatrix} \Sigma^* \otimes L'_2 L_2 & 0 \\ 0 & (\Sigma^{-1} - \Sigma^*) \otimes L'_2 L_2 \end{pmatrix}, f \right).$$

اثبات: از لم ۲-۳ داریم $\xi = \Lambda L'_1 (L_1 \Lambda L'_1)^{-1} (L_1 X L_2 - d)$ اکنون بنا به قضیه ۸-۴ و با در نظر

گرفتن $A = \Lambda L'_1 (L_1 \Lambda L'_1)^{-1} L_1$ ، $B = L_2$ و $C = -\Lambda L'_1 (L_1 \Lambda L'_1)^{-1} d$ ، نتیجه مطلوب حاصل

می‌شود. همچنین $\eta = [I - \Lambda L'_1 (L_1 \Lambda L'_1)^{-1} L_1] X L_2 + \Lambda L'_1 (L_1 \Lambda L'_1)^{-1} d - \theta L_2$

بنابراین با فرض $A = I - \Lambda L'_1 (L_1 \Lambda L'_1)^{-1} L_1$ ، $B = L_2$ و $C = \Lambda L'_1 (L_1 \Lambda L'_1)^{-1} d - \theta L_2$

در قضیه ۸-۴ اثبات کامل می‌شود.

□

لم ۴-۵** فرض کنید $X \sim MEC_{k \times q}(M, Y \otimes \Lambda_1, f)$ که در آن Λ_1 ماتریس معین مثبت و Y ماتریس معین نامنفی با رتبه $p \leq k$ است. همچنین فرض کنید Ξ ماتریس معین مثبت و متقارن، با دو شرط زیر باشد:

الف- $Y\Xi$ خودتوان است.

ب- $\Xi Y \Xi M = \Xi M$.

در این صورت، برای هر تابع اندازه‌پذیر و انتگرال‌پذیر بورل h داریم:

$$E(h(\text{trace}(\Lambda_1^{-1} X' \Xi Y \Xi X)) X) = \kappa_2^1(\text{trace}(\Lambda_1^{-1} M' \Xi Y \Xi M) M,$$

که در آن

$$\kappa_j^i(x) = \int_0^\infty E \left[h^i \left(\chi_{pq+j}^2(zx) \right) \right] u(z) d(z). \quad (۱۰-۴)$$

اثبات: از قضیه ۱-۳ برای $X \sim MN_{k \times q}(M, Y \otimes \Lambda_1)$ داریم

$$E(h(\text{trace}(\Lambda_1^{-1} X' \Xi Y \Xi X)) X) = E \left(h \left(\chi_{pq+2}^2((z \text{trace}(\Lambda_1^{-1} M' \Xi Y \Xi M)) \right) \right) M.$$

حال با استفاده از قضیه ۱۰-۴ خواهیم داشت:

$$\begin{aligned} & E(h(\text{trace}(\Lambda_1^{-1} X' \Xi Y \Xi X)) X) \\ &= \int_0^\infty E \left(h \left(\chi_{pq+2}^2((z \text{trace}(\Lambda_1^{-1} M' \Xi Y \Xi M)) \right) \right) M u(z) d(z). \end{aligned}$$

□

و اثبات کامل است.

لم ۴-۶** فرض کنید

$$(X', Y')' \sim MEC_{2k \times q} \left((M'_1, M'_2), \begin{pmatrix} Y_{11} \otimes \Lambda_{11} & 0 \\ 0 & Y_{22} \otimes \Lambda_{22} \end{pmatrix}, f \right),$$

که در آن Λ_{11} یک ماتریس معین مثبت؛ Υ_{11} ، Υ_{22} و Λ_{22} ماتریس‌های معین نامنفی با رتبه $p \leq k$ هستند. همچنین فرض کنید Ξ یک ماتریس معین مثبت و متقارن با دو شرط زیر باشد:

الف- $\Upsilon_{11}\Xi$ ماتریس خودتوان است.

$$\Xi\Upsilon_{11}\Xi M_1 = \Xi M_1 \quad \text{ب-}$$

در این صورت برای هر تابع اندازه‌پذیر و انتگرال‌پذیر بول h ، و هر ماتریس معین نامنفی A ، داریم:

$$E[h(\text{trace}(\Lambda_{11}^{-1}X'\Xi\Upsilon_{11}\Xi X))Y'AX] = \kappa_2^1(\text{trace}(\Lambda_{11}^{-1}M_1'\Xi M_1))M_2'AM_1,$$

که در آن $\kappa_2^1(x)$ در رابطه (۴-۱۰) داده شده است.

□ اثبات: با استفاده از قضیه ۳-۲ و ۴-۱۰ و روشی مشابه اثبات لم ۴-۵ نتیجه حاصل می‌شود.

لم ۴-۷** فرض کنید $X \sim MEC_{k \times q}(M, Y \otimes \Lambda, f)$ که در آن Λ ماتریس معین مثبت و Y

یک ماتریس معین نامنفی با رتبه $p \leq k$ است. همچنین فرض کنید A و Ξ ماتریس‌های متقارن معین مثبت هستند که برای Ξ دو شرط زیر برقرار است:

الف- $Y\Xi$ ماتریس خودتوان است.

$$\Xi Y \Xi M = \Xi M \quad \text{ب-}$$

در این صورت برای هر تابع اندازه‌پذیر و انتگرال‌پذیر بول h ، داریم:

$$\begin{aligned} & E[h(\text{trace}(\Lambda^{-1}X'\Xi Y \Xi X))\text{trace}(X'AX)] \\ &= \kappa_2^1(\text{trace}(\Lambda^{-1}M'\Xi Y \Xi M))\text{trace}(AY)\text{trace}(\Lambda) \\ &+ \kappa_4^1(\text{trace}(\Lambda^{-1}M'\Xi Y \Xi M))\text{trace}(M'AM). \end{aligned}$$

□ اثبات: از قضایای ۳-۳ و ۴-۱۰ نتیجه حاصل می‌شود.

قضیه ۴-۱۱** تابع مخاطره برآوردگر $\hat{\theta}_S$ به صورت زیر است

$$\begin{aligned} R(\hat{\theta}_S, \theta; W) &= \text{trace}(W(\Sigma^{-1} - \Sigma^*))\text{trace}(L'_2 L_2) + \text{trace}(\delta' W \delta) \\ &\quad - 2\kappa_2^1(\text{trace}(\Xi_2 \delta' \Xi_1 \delta))\text{trace}(\delta' W \delta) \\ &\quad + \kappa_2^2(\text{trace}(\Xi_2 \delta' \Xi_1 \delta))\text{trace}(W \Sigma^*)\text{trace}(L'_2 L_2) \\ &\quad + \kappa_4^2(\text{trace}(\Xi_2 \delta' \Xi_1 \delta))\text{trace}(\delta' W \delta). \end{aligned}$$

□ اثبات: از قضیه ۳-۴ و لم‌های ۴-۶ و ۷-۴ به نتیجه مطلوب می‌رسیم.

بنا بر مثال ۳-۱، مخاطره دیگر برآوردگرها را با کمک $R(\hat{\theta}_S, \theta; W)$ به دست می‌آوریم. بنابراین داریم:

$$R(X, \theta; W) = \text{trace}(W \Sigma^{-1})\text{trace}(L'_2 L_2), \quad (۱۱-۴)$$

$$R(\hat{\theta}, \theta; W) = R(X, \theta; W) - \text{trace}(W \Sigma^*)\text{trace}(L'_2 L_2) + \text{trace}(\delta' W \delta). \quad (۱۲-۴)$$

برای محاسبه مخاطره برآوردگرهای استاین و استاین مثبت با فرض $\Delta = \text{trace}(\Xi_2 \delta' \Xi_1 \delta)$ می‌توان نوشت:

$$\begin{aligned} (\hat{\theta}^s, \theta; W) &= R(X, \theta; W) + \text{trace}(\delta' W \delta)((pq)^2 - 4)(K_4^4) \\ &\quad - (pq - 2)\text{trace}(L'_2 L_2)\text{trace}(W \Sigma^*)[2(K_2^2) - (pq - 2)(K_2^4)], \end{aligned} \quad (۱۳-۴)$$

$$K_{(j)}^{(i)} = \int_0^\infty E \left(\chi_{pq+j}^{-i}(z\Delta) \right) u(z) d(z) \quad \text{که در آن}$$

به طریقی مشابه و با فرض $d = pq - 2$.

$$\begin{aligned} R(\hat{\theta}^{s+}, \theta; W) &= R(\hat{\theta}^s, \theta; W) \\ &\quad - \text{trace}(L'_2 L_2)\text{trace}(W \Sigma^*)K_2^{*2} \\ &\quad + 2\text{trace}(\delta' W \delta)K_1^{*2} \end{aligned}$$

$$- \operatorname{trace}(\delta' W \delta) K_{2,}^{*4}, \quad (۱۴-۴)$$

$$.K_{(j)}^{*(i)} = \int_0^\infty E \left[\left(1 - d \chi_{pq+i}^{-2}(z\Delta) \right)^j I_{(\chi_{pq+j}^2(z\Delta) \leq d)} \right] u(z) d(z) \quad \text{که در آن}$$

پیوست

A-۱ جبر ماتریس‌ها

در این قسمت تعاریف و قضایایی از جبر ماتریس‌ها، که در این پایان‌نامه از آنها استفاده شده است را بدون اثبات آورده‌ایم. البته خواننده برای اثبات قضایای این بخش می‌تواند به رنچر^۱ (۲۰۰۰)، مونتگومری^۲ (۲۰۰۱)، کولو و وان‌رزن (۲۰۰۵)، جیری^۳ (۲۰۰۴)، مولر و استیوارت^۴ (۲۰۰۶)، اندرسن (۲۰۰۳) و سیرل (۱۹۸۲) مراجعه کند.

تعریف A-۱ اگر جای سطرها یا ستون‌های ماتریس A را با ستون و یا سطرهای آن تعویض نماییم، ماتریس حاصل ترانپوز^۵ ماتریس A نامیده می‌شود و با نماد A' نمایش داده می‌شود.

$$A' = (a_{ij})' = (a_{ji})$$

قضیه A-۱ برای هر ماتریس A داریم

$$(A')' = A$$

تعریف A-۲ اگر $A' = A$ و یا بطور معادل $(a_{ji}) = (a_{ij})$ ، آنگاه ماتریس A متقارن^۶ است.

تعریف A-۳ اگر A و B ماتریس‌های $n \times p$ باشند، آنگاه $C = A + B$ نیز یک ماتریس $n \times p$

است، که در آن $C = (c_{ij}) = (a_{ij} + b_{ij})$

قضیه A-۲ اگر A و B ماتریس‌های $n \times p$ باشند آنگاه

$$\begin{aligned} A + B &= B + A \\ (A + B)' &= A' + B' \end{aligned}$$

¹ Rencher

² montgomery

³ Giri

⁴ Muller and Stewart

⁵ transpose

⁶ symmetric

تعریف A-۴ اگر A یک ماتریس $n \times m$ و B یک ماتریس $m \times p$ باشد، آنگاه (ij) امین عنصر

$$C = AB, \text{ به صورت } c_{ij} = \sum_k a_{ik} b_{kj} \text{ تعریف می‌شود.}$$

قضیه A-۳ اگر A یک ماتریس $n \times p$ و B یک ماتریس $p \times m$ باشد، آنگاه

$$(AB)' = B'A'$$

قضیه A-۴ اگر A یک ماتریس $n \times p$ باشد، آنگاه

$$(۱) \quad AA' \text{ و } A'A \text{ متقارن هستند.}$$

$$(۲) \quad \text{اگر } A'A = 0 \text{ باشد، آنگاه } A = 0$$

تعریف A-۵ بردارهای $a_1, a_2, \dots, a_n$ وابسته خطی‌اند، هرگاه اعداد حقیقی $c_1, c_2, \dots, c_n$ (همگی

صفر نباشند) وجود داشته باشند، به طوری که

$$c_1 a_1 + c_2 a_2 + \dots + c_n a_n = 0.$$

تعریف A-۶ بردارهای $a_1, a_2, \dots, a_n$ مستقل خطی‌اند، اگر $Ac = 0$ نتیجه دهد $c = 0$.

تعریف A-۷ اگر A یک ماتریس مربع باشد، آنگاه تعداد ستون‌ها یا سطرهاى مستقل خطی A رتبه^۷ آن

نامیده می‌شود.

اگر A یک ماتریس $n \times p$ با رتبه p و $p < n$ باشد، آنگاه A دارای ماکزیمم رتبه ممکن است و گفته

می‌شود A تمام رتبه^۸ است. به طور کلی رتبه هر ماتریس $A_{n \times p}$ به صورت $\min(n, p)$ است.

قضیه A-۵ فرض کنید A یک ماتریس در اندازه $p \times q$ باشد، آنگاه

Rank⁷
Full rank⁸

$$0 \leq \text{rank}(A) \leq \min(p, q) \quad (۱)$$

$$\text{rank}(A) = 0 \quad \text{اگر } A = 0 \text{ در این صورت} \quad (۲)$$

$$\text{rank}(A) = \text{rank}(A') = \text{rank}(AA') = \text{rank}(A'A) \quad (۳)$$

$$\text{rank}(A + B) \leq \text{rank}(A) + \text{rank}(B) \quad (۴)$$

$$(۵) \quad \text{اگر ضرب ماتریس‌های } A \text{ و } B \text{ تعریف شده باشد، آنگاه}$$

$$\text{rank}(AB) \leq \min(\text{rank}(A), \text{rank}(B)).$$

$$(۶) \quad \text{اگر } B \text{ و } C \text{ تمام رتبه باشند، آنگاه}$$

$$\text{rank}(AB) = \text{rank}(CA) = \text{rank}(A).$$

تعریف A-۸ اگر A یک ماتریس مربع و رتبه آن با بعدش یکسان باشد، آنگاه ماتریس A نامنفرد^۹ نامیده می‌شود، دارای معکوس منحصر بفرد است و

$$AA^{-1} = A^{-1}A = I.$$

اگر A یک ماتریس مربع باشد و رتبه آن کمتر از بعدش باشد، آنگاه معکوس ندارد و ماتریس منفرد^{۱۰} نامیده می‌شود.

قضیه A-۶ اگر A یک ماتریس نامنفرد باشد، A' نیز نامنفرد است و معکوس آن را می‌توان به صورت زیر نوشت:

$$(A')^{-1} = (A^{-1})'$$

قضیه A-۷ اگر A و B ماتریس‌های نامنفرد، مربع و هم‌بعد باشد، آنگاه AB نامنفرد است و

$$(AB)^{-1} = B^{-1}A^{-1}.$$

nonsingular^۹
singular^{۱۰}

تعریف A-9 فرض کنید A یک ماتریس $n \times k$ باشد. اگر B ماتریسی $k \times n$ باشد که در رابطه

$$ABA = A,$$

صدق می‌کند، B را وارون تعمیم‌یافته^{۱۱} A نامیم و با A^- نشان می‌دهیم.

بدیهی است که اگر A مربع و وارون‌پذیر باشد $A^- = A^{-1}$. همچنین وارون تعمیم‌یافته ممکن است یکتا نباشد.

تعریف A-۱۰ فرض کنید A یک ماتریس $n \times k$ باشد. اگر B ماتریسی $k \times n$ باشد که در رابطه‌های

$$ABA = A,$$

$$BAB = B,$$

$$(AB)' = AB,$$

$$(BA)' = BA,$$

صدق می‌کند، B را وارون تعمیم‌یافته مور-پنروز^{۱۲} A نامیم و با A^+ نشان می‌دهیم.

وارون تعمیم‌یافته مور-پنروز ماتریس A ، یکتا است.

تعریف A-۱۱ دترمینان ماتریس A در اندازه $n \times n$ (با نماد $|A|$ نشان می‌دهیم) تابعی حقیقی مقدار

است و به صورت زیر تعریف می‌شود:

$$|A| = a_{11} \quad n = 1,$$

$$|A| = \sum_{j=1}^n a_{1j} |A_{1j}| (-1)^{1+j} \quad n > 1,$$

¹¹ Generalized inverse

¹² Moore-Penrose generalized inverse

که در آن A_{1j} ماتریسی در اندازه $(n-1) \times (n-1)$ است و با حذف اولین سطر و ستون ماتریس A به دست می‌آید.

قضیه A-۸

(۱) اگر $D = \text{diag}(d_1, d_2, \dots, d_n)$ ، آنگاه $|D| = \prod_{i=1}^n d_i$.

(۲) اگر A منفرد باشد، آنگاه $|A| \neq 0$.

(۳) $|A'| = |A|$.

(۴) اگر A نا منفرد باشد، آنگاه $|A^{-1}| = \frac{1}{|A|}$.

قضیه A-۹ اگر A یک ماتریس متقارن، مربع و به شکل زیر افراز شده باشد:

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix},$$

و ماتریس‌های A_{11} و A_{22} مربع و نامنفرد باشند، آنگاه

$$\begin{aligned} |A| &= |A_{11}| |A_{22} - A_{21} A_{11}^{-1} A_{12}| \\ &= |A_{22}| |A_{11} - A_{12} A_{22}^{-1} A_{21}|. \end{aligned}$$

قضیه A-۱۰ اگر A و B دو ماتریس مربع و هم‌بعد باشند، آنگاه $|AB| = |A||B|$.

تعریف A-۱۲ دو بردار $a_{n \times 1}$ و $b_{n \times 1}$ را متعامد^{۱۳} می‌نامند، هرگاه

$$a'b = a_1 b_1 + a_2 b_2 + \dots + a_n b_n = 0$$

تعریف A-۱۳ ماتریس $C_{p \times p} = (c_1, \dots, c_p)$ متعامد است، هرگاه ستون‌های آن، بردارهای متعامد

باشند. و مجموعه ماتریس‌های متعامد را با $O(p)$ نشان می‌دهیم. یعنی

$$O(p) = \{C(p \times p): CC' = I_p\}.$$

قضیه A-۱۱ اگر C یک ماتریس متعامد باشد، آنگاه $C'C = CC' = I$.

قضیه A-۱۲ اگر $C_{p \times p}$ یک ماتریس متعامد و $A_{p \times p}$ هر ماتریس دیگری باشد، آنگاه

$$|C| = \pm 1 \quad (۱)$$

$$|C'AC| = |A| \quad (۲)$$

$$-1 \leq c_{ij} \leq 1, \text{ که در آن } c_{ij} \text{ عنصر ماتریس } C \text{ است.} \quad (۳)$$

$$C^{-1} = C' \quad (۴)$$

تعریف A-۱۴ اگر A یک ماتریس مربع با بعد n باشد، آنگاه مجموع عناصر روی قطر اصلی، اثر^{۱۴} ماتریس A نام دارد و به صورت زیر تعریف می‌شود:

$$tr(A) = \sum_{i=1}^n a_{ii}.$$

قضیه A-۱۳

$$(۱) \text{ اگر } A \text{ ماتریس } n \times n \text{ باشد، آنگاه } tr(A) = tr(A').$$

$$(۲) \text{ اگر } A \text{ ماتریس } n \times n, \text{ و } k \neq 0 \text{ یک عدد حقیقی باشد، آنگاه } tr(kA) = ktr(A).$$

$$(۳) \text{ اگر } A \text{ و } B \text{ دو ماتریس } n \times n \text{ باشند، آنگاه } tr(A \pm B) = tr(A) \pm tr(B).$$

$$(۴) \text{ اگر } A \text{ ماتریس } n \times p \text{ و } B \text{ ماتریس } p \times n \text{ باشند، آنگاه } tr(AB) = tr(BA).$$

$$(۵) \text{ اگر } A \text{ ماتریس } n \times p \text{ باشد، آنگاه } tr(A'A) = \sum_{j=1}^n a'_j a_j = \sum_{i=1}^n \sum_{j=1}^p a_{ij}^2.$$

$$(۶) \text{ اگر } A \text{ ماتریس } n \times n \text{ و } P \text{ یک ماتریس } n \times n \text{ نامنفرد باشد، آنگاه}$$

$$tr(P^{-1}AP) = tr(A)$$

(۷) اگر A ماتریس $n \times n$ و C یک ماتریس $n \times n$ متعامد باشد، آنگاه $tr(C'AC) = tr(A)$.

تعریف A-۱۵ برای هر ماتریس مربع A ، عدد حقیقی λ و بردار غیرصفر X را می‌توان یافت به طوری که

$$AX = \lambda X$$

که در آن λ مقدار ویژه^{۱۵} و X بردار ویژه^{۱۶} ماتریس A نامیده می‌شوند و با حل $|A - \lambda I| = 0$ نسبت به λ به دست می‌آیند.

قضیه A-۱۴

(۱) اگر P یک ماتریس نامنفرد باشد، آنگاه A و $P^{-1}AP$ دارای مقادیر ویژه یکسان می‌باشند.

(۲) اگر C یک ماتریس متعامد باشد، آنگاه A و $C'AC$ دارای مقادیر ویژه یکسان می‌باشند.

(۳) اگر $A_{n \times n}$ متقارن باشد، آنگاه $\lambda_1, \lambda_2, \dots, \lambda_n$ مقادیری حقیقی و بردارهای ویژه $x_1, \dots, x_n$

ماتریس A دو به دو متعامد می‌باشند، یعنی به ازای $i \neq j$ داریم $x_i x_j = 0$.

(۴) اگر $A_{n \times n}$ با مقادیر ویژه $\lambda_1, \lambda_2, \dots, \lambda_n$ باشد، آنگاه

$$|A| = \prod_{i=1}^n \lambda_i,$$

$$tr(A) = \sum_{i=1}^n \lambda_i.$$

(۵) رتبه ماتریس A با تعداد مقادیر ویژه غیرصفر آن برابر است.

(۶) ماتریس A یک ماتریس نامنفرد است، اگر و فقط اگر تمام مقادیر ویژه آن غیرصفر باشند.

¹⁵Eigenvalue
¹⁶Eigenvector

قضیه A-۱۵ اگر A ماتریس متقارن با بعد n باشد، در این صورت ماتریس متعامد Q وجود دارد به طوری که

$$A = Q\Lambda Q',$$

که در آن $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_{n_1}, 0, \dots, 0)$ و $n_1 < n$.

تعریف A-۱۶ اگر A یک ماتریس متقارن و y یک بردار باشد، آنگاه

$$y' Ay = \sum_i a_{ii} y_i^2 + \sum_{i \neq j} a_{ij} y_i y_j$$

فرم درجه دوم^{۱۷} نامیده می‌شود. اگر $x_{n \times 1}$ یک بردار و $y_{p \times 1}$ یک بردار ویژه و $A_{n \times p}$ یک ماتریس باشد، آنگاه

$$x' Ay = \sum_{ij} a_{ij} x_i y_j$$

فرم دوخطی^{۱۸} نامیده می‌شود.

تعریف A-۱۷ اگر A یک ماتریس متقارن با ویژگی $y' Ay > 0$ ($y' Ay \geq 0$) به ازای هر $y \neq 0$ باشد، آنگاه فرم درجه دوم $y' Ay$ ، معین مثبت^{۱۹} (نیمه معین مثبت^{۲۰}) و A ماتریس معین مثبت (نیمه معین مثبت) نامیده می‌شود. و آن را با نماد $A > 0$ ($A \geq 0$) نشان می‌دهیم.

اگر $A > 0$ یا $A \geq 0$ گوئیم ماتریس A معین نامنفی^{۲۱} است.

¹⁷ Quadratic Form

¹⁸ Bilinear Form

¹⁹ Positive Definite

²⁰ Positive Semidefinite

قضیه A-۱۶ اگر $A_{n \times n}$ معین مثبت (نیمه معین مثبت) با مقادیر ویژه $\lambda_1, \lambda_2, \dots, \lambda_n$ باشد، آنگاه

$$(۱) \text{ عناصر قطر اصلی } a_{ii} > 0 \text{ (} a_{ii} \geq 0 \text{) می‌باشند.}$$

$$(۲) \text{ } A \text{ نامنفرد است.}$$

$$(۳) \text{ } A^{-1} \text{ معین مثبت است.}$$

$$(۴) \text{ } P'AP \text{ معین مثبت (نیمه معین مثبت) است، که در آن } P \text{ یک ماتریس نامنفرد است.}$$

$$(۵) |A| > 0$$

$$(۶) i = 1, 2, \dots, n, \lambda_i > 0$$

قضیه A-۱۷ ماتریس متقارن A معین مثبت است اگر و فقط اگر ماتریس نامنفرد P وجود داشته باشد

$$\text{به طوری که } A = P'P$$

قضیه A-۱۸ اگر ماتریس A معین مثبت باشد، در این صورت ماتریس نامنفرد C وجود دارد به طوریکه

$$C'AC = I.$$

قضیه A-۱۹ فرض کنید x یک بردار ستونی با p مولفه و A یک ماتریس نیمه معین مثبت است. در

این صورت داریم:

$$\lambda_p \leq \frac{x'Ax}{x'x} \leq \lambda_1,$$

که در آن λ_1 و λ_p به ترتیب بزرگترین و کوچکترین مقادیر ویژه ماتریس A هستند.

تعریف A-۱۸ اگر A یک ماتریس $p \times q$ ، در این صورت نرم ماتریس A به صورت زیر تعریف می‌شود

$$\|A\|_{\Xi_1, \Xi_2}^2 = \text{trace}(A'\Xi_1 A \Xi_2),$$

که در آن E_1 یک ماتریس معین نامنفی معلوم و E_2 ماتریس معین مثبت معلوم می‌باشد

تعریف A-۱۹ ماتریس مربع A در اندازه $n \times n$ خودتوان نامیده می‌شود، هرگاه $A^2 = A$.

قضیه A-۲۰ تنها ماتریس خودتوان نامنفرد، I (ماتریس واحد^{۲۲}) است.

قضیه A-۲۱ اگر ماتریس $A_{n \times n}$ خودتوان، $P_{n \times n}$ نامنفرد و $C_{n \times n}$ متعامد باشد، آنگاه

$$tr(A) = rank(A) \quad (۱)$$

$$I_n - A \text{ خودتوان است.} \quad (۲)$$

$$(I - A)A = 0 \text{ و } A(I - A) = 0 \quad (۳)$$

$$P^{-1}AP \text{ خودتوان است.} \quad (۴)$$

$$C'AC \text{ خودتوان است (اگر } A \text{ متقارن باشد، } C'AC \text{ متقارن و خودتوان است).} \quad (۵)$$

قضیه A-۲۲ اگر $A_{n \times n}$ متقارن و برای یک k ، تساوی $A = \sum_{i=1}^k A_i$ برقرار باشد، که در آن A_i در

اندازه $n \times n$ متقارن است، آنگاه

$$A \text{ خودتوان است.} \quad (۱)$$

$$A_1, A_2, \dots, A_k \text{ خودتوان می‌باشند.} \quad (۲)$$

$$\forall i \neq j, A_i A_j = 0 \quad (۳)$$

قضیه A-۲۳ اگر A یک ماتریس مربع متقارن و خودتوان در اندازه n باشد، آنگاه یک ماتریس متعامد مانند Q وجود دارد، به طوری که:

$$Q'AQ = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix},$$

$$r = trace(A) \text{ که در آن}$$

تعریف A-۲۰ فرض کنید A یک ماتریس $n \times k$ و B یک ماتریس $m \times r$ باشد. در این صورت حاصلضرب کرونیگر^{۲۳} A و B ، با نماد $A \otimes B$ ، عبارت است از یک ماتریس $nm \times kr$ که به صورت افراز شده زیر است:

$$A \otimes B = \begin{bmatrix} a_{11}B & \cdots & a_{1k}B \\ \vdots & \ddots & \vdots \\ a_{n1}B & \cdots & a_{nk}B \end{bmatrix} = (a_{ij}B).$$

قضیه A-۲۴

(۱) اگر $\alpha, \beta \neq 0$ عدد حقیقی باشند، آنگاه $(\alpha A) \otimes (\beta B) = \alpha\beta(A \otimes B)$.

(۲) اگر A و B ماتریس‌های $m \times n$ و C هر ماتریس دیگری باشد، آنگاه

$$(A + B) \otimes C = (A \otimes C) + (B \otimes C),$$

$$C \otimes (A + B) = (C \otimes A) + (C \otimes B).$$

$$(A \otimes B) \otimes C = A \otimes (B \otimes C) \quad (۳)$$

$$(A \otimes B)' = A' \otimes B' \quad (۴)$$

$$(I \otimes A)B(I \otimes A') = B \otimes AA' \quad (۵)$$

(۶) اگر A و B ماتریس‌های $m \times m$ باشند، آنگاه $tr(A \otimes B) = tr(A)tr(B)$.

(۷) اگر ماتریس‌های A, B, C و D به ترتیب در اندازه $m \times n, p \times q, n \times r$ و $q \times s$ باشند،

$$(A \otimes B)(C \otimes D) = (AC) \otimes (BD) \quad \text{آنگاه}$$

(۸) اگر ماتریس‌های A و B نامنفرد باشند، آنگاه $(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}$.

(۹) اگر P و Q ماتریس‌های متعامد باشند، آنگاه $P \otimes Q$ نیز یک ماتریس متعامد است.

(۱۰) اگر P و Q ماتریس‌های معین مثبت باشند، آنگاه $P \otimes Q$ نیز یک ماتریس معین مثبت است.

(۱۱) اگر A یک ماتریس $m \times m$ و B ماتریس $n \times n$ باشد، آنگاه $|A \otimes B| = |A|^n |B|^m$.

(۱۲) اگر ماتریس A به صورت $A = [A_1, A_2]$ افراز شده باشد، آنگاه

$$[A_1, A_2] \otimes B = [A_1 \otimes B, A_2 \otimes B].$$

تعریف A-۲۱ شکل برداری ماتریس A که آن را با $vecA$ نشان می‌دهیم، عبارت است از یک ماتریس ستونی که از زیر هم قرار دادن ستون‌های A حاصل می‌شود. یا هم ارز آن $vecA$ عبارت است از یک ماتریس ستونی که از زیر هم قرار دادن سطرهای ماتریس A' حاصل می‌شود. یعنی اگر A یک ماتریس $n \times k$ باشد، آنگاه

$$vecA = \begin{pmatrix} a_1 \\ \vdots \\ a_k \end{pmatrix}_{nk \times 1}$$

که در آن $a_i, i = 1, \dots, k$ i -امین ستون ماتریس A یا i -امین سطر ماتریس A' می‌باشد.

قضیه A-۲۵ اگر ماتریس‌های A, B, C و D به ترتیب در اندازه $p \times q, q \times r, r \times s$ و $s \times p$ باشند، آنگاه

$$(۱) \quad vecA = vecA' \quad \text{با توجه به تعریف A-۲۱}$$

$$(۲) \quad vec(ABC) = (C' \otimes A)vecB$$

$$(۳) \quad vec(AB) = (I \otimes A)vecB = (B' \otimes I)vecA$$

$$(۴) \quad tr(AB) = vec'A'vecB$$

$$(۵) \quad tr(ABC) = vec'A'(I \otimes B)vecC$$

$$(۶) \quad tr(ABCD) = vec'A'(D' \otimes B)vecC$$

تعریف A-۲۲ فرض کنید $u = f(x)$ تابعی از $x = (x_1, \dots, x_p)'$ و $\partial u / \partial x_1, \dots, \partial u / \partial x_p$

مشتق‌های جزئی باشند. $\partial u / \partial x$ را به صورت زیر تعریف می‌کنیم:

$$\frac{\partial u}{\partial x} = \begin{pmatrix} \frac{\partial u}{\partial x_1} \\ \vdots \\ \frac{\partial u}{\partial x_p} \end{pmatrix}.$$

قضیه A-۲۶ فرض کنید $u = a'x = x'a$ که در آن $a' = (a_1, a_2, \dots, a_p)$ برداری ثابت است.

آنگاه

$$\frac{\partial u}{\partial x} = \frac{\partial(a'x)}{\partial x} = \frac{\partial(x'a)}{\partial x} = a.$$

قضیه A-۲۷ فرض کنید $u = x'Ax$ که در آن A ماتریس متقارن و ثابت است. آنگاه

$$\frac{\partial u}{\partial x} = \frac{\partial(x'Ax)}{\partial x} = 2Ax.$$

تعریف A-۲۳ فرض کنید درایه‌های $Y \in \mathbb{R}^{r \times s}$ توابعی از $X \in \mathbb{R}^{p \times q}$ و مشتق‌های جزئی $\frac{\partial y_{kl}}{\partial x_{ij}}$

وجود داشته باشند. ماتریس $\frac{dY}{dX} \in \mathbb{R}^{pq \times rs}$ ، مشتق ماتریس Y نسبت به X نامیده شده و به صورت زیر

تعریف می‌شود:

$$\frac{dY}{dX} = \frac{\partial}{\partial \text{vec} X} \text{vec}' Y,$$

که در آن، $\frac{\partial}{\partial \text{vec} X} = \left(\frac{\partial}{\partial x_{11}}, \dots, \frac{\partial}{\partial x_{p1}}, \frac{\partial}{\partial x_{12}}, \dots, \frac{\partial}{\partial x_{p2}}, \dots, \frac{\partial}{\partial x_{1q}}, \dots, \frac{\partial}{\partial x_{pq}} \right)'$

قضیه A-۲۸ فرض کنید همه ماتریس‌ها در اندازه‌های مناسب باشند و درایه‌های X مستقل از هم و

متغیر^{۲۴} باشند؛ آنگاه

$$\frac{dX}{dX} = I_{pq} \quad (۱)$$

$$(۲) \quad \frac{d(cX)}{dX} = cI_{pq} \text{ با فرض این که } c \text{ یک ثابت است،}$$

$$(۳) \quad \frac{dA'X}{dX} = I \otimes A$$

$$(۴) \quad \frac{d(A' \text{vec} X)}{dX} = A$$

$$(۵) \quad \frac{d(Y+Z)}{dX} = \frac{dY}{dX} + \frac{dZ}{dX}$$

$$(۶) \quad \frac{d \text{tr} Y}{dX} = \frac{dY}{dX} \text{vec} I$$

$$(۷) \quad \frac{d \text{tr}(A'X)}{dX} = \text{vec} A$$

قضیه A-۲۹ فرض کنید A و B ماتریس‌های ثابت در اندازه‌های مناسب باشند. آنگاه

$$(۱) \quad \frac{d(AXB)}{dX} = B \otimes A'$$

$$(۲) \quad \frac{d(AYB)}{dX} = \frac{dY}{dX} (B \otimes A')$$

تعریف A-۲۴ برداری که عناصر آن متغیرهای تصادفی است، بردار تصادفی نامیده می‌شود. و ماتریسی

که درایه‌های آن متغیرهای تصادفی است، ماتریس تصادفی نامیده می‌شود.

تعریف A-۲۵ امید ریاضی ماتریس تصادفی X در اندازه $p \times q$ به صورت زیر تعریف می‌شود:

²⁴Mathematically independent and variable

$$E(X) = E \begin{bmatrix} X_{11} & \cdots & X_{1q} \\ \vdots & \ddots & \vdots \\ X_{p1} & \cdots & X_{pq} \end{bmatrix} = \begin{bmatrix} E(X_{11}) & \cdots & E(X_{1q}) \\ \vdots & \ddots & \vdots \\ E(X_{p1}) & \cdots & E(X_{pq}) \end{bmatrix} = \begin{bmatrix} \mu_{11} & \cdots & \mu_{1q} \\ \vdots & \ddots & \vdots \\ \mu_{p1} & \cdots & \mu_{pq} \end{bmatrix} \\ = \mu,$$

$$E(X_{ij}) = \mu_{ij} \text{ که در آن}$$

تعریف A-۲۶ فرض کنید ماتریس تصادفی Y در اندازه $p \times q$ به صورت $Y = (Y_1, Y_2, \dots, Y_q)'$ افزایش شده است، که در آن $Y_1, Y_2, \dots, Y_q$ بردارهای ستونی با p مولفه، و دارای واریانس‌های $\sigma_1^2, \sigma_2^2, \dots, \sigma_p^2$ و کوواریانس‌های σ_{ij} ، $\forall i \neq j$ هستند. در این صورت ماتریس کوواریانس بردار Y_k ، $k = 1, \dots, q$ به صورت زیر تعریف می‌شود:

$$\begin{aligned} \text{var}(Y_k) &= E[(Y_k - \mu_k)(Y_k - \mu_k)'] = E(Y_k Y_k') - \mu_k \mu_k' \\ &= \begin{bmatrix} \sigma_{11} & \cdots & \sigma_{1p} \\ \vdots & \ddots & \vdots \\ \sigma_{p1} & \cdots & \sigma_{pp} \end{bmatrix} = \Sigma, \end{aligned}$$

که در آن $i = 1, 2, \dots, p$ ، $\sigma_{ii} = \sigma_i^2$ همچنین ماتریس Σ متقارن است.

همچنین ماتریس کوواریانس ماتریس Y به صورت زیر است:

$$\begin{aligned} \text{cov}(Y) &= \text{cov}(\text{vec} Y) = E\{\text{vec} Y \text{vec}' Y\} - E\{\text{vec} Y\} E\{\text{vec}' Y\} \\ &= \begin{bmatrix} \text{cov}(Y_1, Y_1) & \cdots & \text{cov}(Y_1, Y_q) \\ \vdots & \ddots & \vdots \\ \text{cov}(Y_q, Y_1) & \cdots & \text{cov}(Y_q, Y_q) \end{bmatrix}_{pq \times pq}. \end{aligned}$$

A-2 پیشنهادات برای پژوهش‌های آینده (ارائه زمینه‌های تحقیق)

با توجه به نتایج بدست آمده در این پایان‌نامه و موضوعات مطرح شده می‌توان در آینده بر روی موارد زیر تحقیق کرد.

- تابع زیان به کار رفته در تمام فصل‌های این پایان‌نامه از نوع تابع زیان توان دوم موزون است. می‌توان بررسی‌های انجام شده در فصل‌های سوم و چهارم را با در نظر گرفتن تابع زیان نمایی خطی چندپارامتری^۱ (MLINEX)، متوازن^۲ و نرمال منعکس شده^۳ (RN) ادامه داد.
- محدودیت اعمال شده بر روی فضای پارامتر در فصل‌های سوم و چهارم از نوع قیود قطعی فرض شده است. که می‌توان قیودی غیرقطعی برای این محدودیت‌ها به کار برد. مثلاً می‌توان از قید $L_1\theta L_2 = d + \epsilon$ استفاده کرد؛ که در آن ϵ یک متغیر تصادفی و دارای توزیع خاصی است.
- در خصوص مدل منحنی رشد، فرض شده که ستون‌ها از هم مستقلند، و در نتیجه ماتریس کوواریانس به صورت $I \otimes \Sigma$ می‌باشد. می‌توان فرض کرد که بین ستون‌ها رابطه وجود دارد و ماتریس کوواریانس به صورت $\psi \otimes \Sigma$ است و بر این اساس برآوردها را پیدا کرد.
- در خصوص فصل چهارم، بدون اینکه مدل رگرسیونی خاصی را فرض کنیم، برآوردها را برای ماتریس میانگین در توزیع بیضی‌گون ماتریسی بدست آوردیم. می‌توان با فرض اینکه در مدل رگرسیونی چندمتغیره ارائه شده در فصل دوم، ماتریس خطا دارای توزیع بیضی‌گون ماتریسی است، به بررسی و مقایسه برآوردهای محدود نشده، محدود شده، استاین و استاین نوع مثبت پرداخت.

¹ Multiparameter linear exponential loss function

² Balance Loss

³ Reflected normal loss function

مراجع

- [1] آرشی، م. (۱۳۸۷)، رساله دکتری: "برآوردگرهای بهبودیافته در مدل رگرسیون خطی چندگانه با خطاهای دارای توزیع بیضی‌گون"، دانشکده علوم ریاضی، دانشگاه فردوسی مشهد.
- [2] Ali, A. M. and Saleh, A. K. Md. E. (1991), "Preliminary test and empirical Bayes approach to shrinkage estimation of regression parameters" *J. Japan. Statist. Soc*, 21, 401-416.
- [3] Anderson, T. W. (2003), "An introduction to multivariate Statistical analysis" 3rd ed., JohnWiley, New York.
- [4] Anderson, T. W. and Fang, K. T. (1991), "Theory and applications of elliptically contoured and related distributions, in the development of statistics" Recent contributions from China, X.R. Chen, K. Fang and C.C.Yang. Eds. Longman scientific and Technical. Landon.
- [5] Arashi, M. (2012), "Preliminary test and Stein estimators in simultaneous linear equations" *Linear Algebra and its Applications.*, 436(5), 1195-1211.
- [6] Arashi, M. and Tabatabaey, S. M. M. (2010), "A note on Stein-type estimators in elliptically contoured models" *journal of Statistical Planning and Inference.*, 140, 1206-1213.
- [7] Arashi, M., Saleh, A. K. Md. E. and Tabatabaey, S. M. M. (2011), "On mathematical characteristics of some improved estimators of the mean and variance components in elliptically contoured models" *journal of Iranian Statistical Society (JIRSS).*, 10(2), 237-266.

- [8] Arashi, M., Tabatabaey, S. M. M. and Khan, Shahjahan. (2008), "Estimation in multiple regression model with elliptically contoured errors under MLINEX loss" *Journal, of Applied Probability and Statistics.*, 3(1), 23-35.
- [9] Asano, C. and Sato, S. (1962), "A bivariate analogue of pooling data bull" *J. Math. Statist.*, 10, 39-59.
- [10] Bancroft, T. A. (1944) "On biases in estimation due to the use of preliminary tests of significance" *Annal. Math. Statist.*, 15, 195-204.
- [11] Bancroft, T. A. (1964) "Analysis and inference for incompletely specified models involving the use of preliminary test(s) of significance" *J. Biometrics.*, 20, 427-442.
- [12] Bancroft, T. A. (1965) "Inference for incompletely specified models in the physical sciences (with discussion)" *Bull. ISI, Proc. 35 th Session.* 41(1), 497-515.
- [13] Bartlett, M. S. (1934), "The vector representation of a sample. *Proc. Cambridge Phil, Soc.*, 30, 327-340.
- [14] Bennet, B. M. (1952) "Estimation of means on the biases of preliminary tests of significance. *J. An. Inst. Stat. Math.* 4, 31-43.
- [15] Bennet, B. M. (1956) "On the use of preliminary tests in certain statistical procedures" *J. An. Inst. Stat. Math.* 8, 45-57.
- [16] Blake, I. F. and Thomas, J. B. (1968), "On a class of processes arising in linear estimation theory" *J. IEEE Trans. Inf. Theory.*, 20, 1-20.
- [17] Bozivich, H., Bancroft, T. A. and Hartley, H. o. (1956) "Power of analysis of variance test procedure for certain incompletely specified models" *J. An. Math. Statist.* 27, 1037-1043.
- [18] Cambanis, S., Huang, S. and Simons, G. (1981), "On the theory of elliptically contoured distributions" *J. Multivariate Anal.*, 11, 368-385.

- [19] Chemielewski, M. A. (1981), "Elliptically symmetric distributions: a review and bibliography" *J. Int. Statist.* 4 Rev., 49, 67-74.
- [20] Cheong, Y. H. (1999), Ph.D. thesis "The distribution of quadratic forms in elliptically contoured random vectors" The university of Western Ontario, London.
- [21] Chu, K. C. (1973) "Estimation and decision for linear systems with elliptically random process" *J. IEEE. Transactions on Automatic Control.* 18,499-505.
- [22] Clarke, J. A., Giles,D. E. A. and Wallace, D. (1987) "Preliminary test estimation of the error variance in linear regression" *J. Econometric. Theo.* 3, 299-304.
- [23] Fang, K. T. and Anderson, T. W. (1990), "*Statistical inference in elliptically contoured and related distributions*" Allerton press, New York.
- [24] Fang, k. T. and Zhang, Y. (1990) "*Generalized multivariate analysis*" Springerverlag nd science press, Berlin and Beijing.
- [25] Giles, J. A. (1991) "Pre-testing for linear restriction model with spherically symmetric distributions. *J. Econometrics.* 50, 337-398.
- [26] Giles, J. A. (1992) "Estimation of the error variance after a preliminary-test of homogeneity in a regression model with spherically symmetric distributions. *J. Econometrics.* 53, 345-361.
- [27] Giri, N. C. (2004), "*multivariate statistical analysis*" 2nd , Marcel Dekker,Inc. New York.
- [28] Gradshteyn, I. S. and Ryzhik, I. M. (2000) "*Tables of integral, series and products, sixthed*" Academic press. USA.
- [29] Gupta, A. K. and Nagar, D. K. (1999) "*Matrix variate distributions*" Chapman and Hall/CRC
- [30] Gupta, A. K. and Varga, T. (1995) "Normal mixture representations of matrix variate elliptically contoured distributions" *Sankhya*, 57, 68-78.

- [31] Han, C. P. and Bancroft, T. A. (1968) "On pooling means when variance is unknown, *J. Amer. Statist. Assoc.*, 63, 1333-1342.
- [32] Hartman, P. and Winter, A. (1940), "On the spherical approach to the normal distribution law" *J. Amer. Math.*, 62, 759-779.
- [33] Huntsberger, D. V. (1955) "A generalization of a preliminary test procedure of pooling data" *J. An. Stat. Math.* 26, 734-743.
- [34] James, W. and Stein, C. (1961) "Estimation with quadratic loss. Proc. fourth berkeley symp" *J. math. Statist and probability.* 1, 361-379.
- [35] Johnson, N. L. and Kotz, S. (1972), "*Distributions in statistics, continuous multivariate distributions*" Wiley, New York.
- [36] Judge, G. G. and Bock, M. E. (1978) "*The statistical implication of pre-test and stein-rule estimators in econometrics*" North-Holland, New York.
- [37] Kariya, T. and Kurata, H. (2004) "*Generalized least squares*" John Wiley and sons, England.
- [38] Kelker, D. (1970), "Distribution theory of spherical distributions and a location scale parameter generalization" *J. Sankhya A.*, 32, 419-430.
- [39] Khan, S. and Saleh, A. K. Md. E. (1997) "Shrinkage pre-test estimator of the intercept parameter for a regression model with multivariate-t errors" *J. Biometrical.* 39, 1-17.
- [40] Kibria, B. M. G. (1996) "On shrinkage ridge regression estimators for restricted linear models with multivariate t distribution" *J. Student.* 1, 177-188.
- [41] Kitagawa, T. (1963) "Estimation after preliminary tests of significance" *J. Uc. Publ. Stat.* 3, 147-186.
- [42] Kollo, T. and von Rosen, D. (2005) "*Advanced Multivariate Statistics with Matrices*" Springer, New York.

- [43] Kshirsagar, A. M. and Smith, W. B. (1995), "*Growth Curves*" Marcel Dekker, New York.
- [44] Kuan, Ch. M. (2002), "*Statistics: Concepts and methods*, 2nd Edition, Hua-Tai, Taipei.
- [45] Kurata, H. (1999), "On the efficiencies of the several generalized least squares estimators in a seemingly unrelated regression model and a heteroscedastic model" *J. Mult. Annal.*, 70, 86-94.
- [46] Maxwell, J. C. (1860), "Illustration of the dynamical theory of gases" *Part I. J. Phil. Mag.*, 19, 19-32.
- [47] McGraw, D. K. and Wagner, J. F. (1968), "Elliptically symmetric distributions" *J. IEEE Trans. Inf. Theory. IT.*, 14, 110-120.
- [48] Mosteller, F. (1948) "On pooling data" *J. Am. Stat. Assoc.* 43, 231-242.
- [49] Muirhead, R. J. (1982) "*Aspects multivariate statistical theory*" Wiley. New York.
- [50] Nadarajah, S. and Kotz, S. (2005) "Mathematical properties of the multivariate t distribution" *J. Acta. Appl. Math.* 89, 53-84.
- [51] Nkurunziza, S. (2012) "The risk of pre-test and shrinkage estimators" *J. of theoretical and applied statistics*, 46(3), 305-312.
- [52] Pan, J. X. and Fang, K. T. (2002), "*Growth Curve Model and Statistical Diagnostics*" Springer, New York.
- [53] Potthoff, R. F. and Roy, S. N. (1964), "A generalized multivariate analysis of variance model useful specially for growth curves" *J. Biometrika*, 51, 313-326.
- [54] Ravishanker, N. and Dey, D. K. (2002), "*A first course in linear model Theory*" Chapman and Hall/CRC.
- [55] Rencher, C. A. (2000), "*Linear models in statistics*" John Wiley and Sons.
- [56] Saleh, A. K. Md. E. (2006) "*Theory of preliminary test and Stein-type estimation with applications*" John Wiley, New York.

- [57] Searle, S. R. (1971) “*Linear models*” John Wiley and Sons. New York.
- [58] Singh, R.S. (1991) “James- Stein rule estimators in linear regression models with multivariate-t distribution error. *Austral. J. Statist.* 33, 145-158.
- [59] Snedecore, G. W. (1938) “*Statistical Methods*” Collegiate. Press, Iowa.
- [60] Srivastava, V. K. and Giles, D. E. A. (1987), “*Seemingly unrelated regression equation model*” Marcel Decker.
- [61] Stein, C. (1955) “A necessary and sufficient condition for admissibility” *J. An. Math. Statist.* 26, 518-522.
- [62] Stein, C. (1956) “Inadmissibility of the usual estimator of the mean of a multivariate normal distribution. Proc.3rd Berkeley Symp” *J. Math. Statist. Probability.* 1, 197-206.
- [63] Stein, C. (1981) “Estimation of the mean of a multivariate normal distribution” *J. An. Math. Statist.* 9,1135-1151.
- [64] Tabatabaey, S. M. M. (1995), Ph.D. thesis, “*Preliminary test approach estimation regression model with spherically symmetric errors*”. Carleton University, Ottawa.
- [65] Tabatabaey, S. M. M., Saleh, A. K. Md. E. and Kibria, B. M. G. (2004a) “Estimation strategies for parameters of the linear regression model with spherically symmetric distributions” *J. Statist. Res.* 38, 13-31.
- [66] Tabatabaey, S. M. M., Saleh, A. K. Md. E. and Kibria, B. M. G. (2004b) “Simultaneous estimation of regression parameters with spherically symmetric error under possible stochastic constraints” *Int. J. Statist. Sci.* 3, 1-20.
- [67] Taylor, W. E. (1978), “heteroscedastic linear model: exact finite sample result” *J. Econometrica*, 46, 663-675.
- [68] Von Rosen, D. (1989), “Maximum likelihood estimators in multivariate linear normal models” *J. Multivariate Analysis* 31, 187-200.

- [69] Wishart, J. (1938) "Growth rate determinations in nutrition studies with bacon pig, and thier analysis" *J. Biometrika*, 30, 16-28.
- [70] Zellner, A. (1962), "An efficient method of estimating seemingly unrelated regression and tests for aggregation bias" *J. Amer. Statist. Assoc.*, 58, 977-992.

Abstract

After magnificent work of Prof Charles Stein in 1956, regarding the inadmissibility of the minimax estimator of mean of multivariate normal, when the dimension is greater than 2, there are many studies to develop improved estimators in the sense of having smaller risk or MSE than the minimax estimator of mean under different settings. These improved estimators are shrinkage in some sense. In this approach, we consider multivariate models with matrix structure for errors. We study the performance of least squares as well as shrinkage estimators. The underlying models are matrix variate normal and matrix variate elliptical distributions. The risk functions of the proposed estimators are exactly derived and the performance of the estimators is compared by their risk measures.

Keywords: Multivariate regression model, Risk function, Elliptically contoured distribution, Shrinkage estimator, Stein-type estimator



Shahrood University of Technology

Faculty of Mathematical Sciences

Shrinkage estimation in multivariate models

Seyyed aliasghar Tajaddod

Supervisor:

Dr. Mohammad Arashi

Date: September 2012