دانشگاه صنعتی شاهرود

دانشکده ریاضی
گروه آمار

پایان‌نامه

برای دریافت درجه کارشناسی ارشد در رشته
آمار، گرایش آماررياضی

عنوان

برآورد انقباضی در مدل‌های محدود شده

استاد راهنما

دکتر محمد آرشی

پژوهشگر

حمید کرمی کبیر

اردیبهشت ۱۳۹۲

نام خانوادگی دانشجو: کرمی کبیر　　　　نام: حمید

عنوان: برآورد انقباضی در مدل‌های محدود شده

استاد راهنما: دکتر محمد آرشی

مقطع تحصیلی: کارشناسی ارشد　　　رشته: آمار　　　گرایش: آماررياضی

دانشگاه: صنعتی شاهرود　　　دانشکده ریاضی

تاریخ فارغ‌التحصیلی: اردیبهشت ۱۳۹۲　　　تعداد صفحات: ۱۲۱

واژگان کلیدی: برآوردگر انقباضی، توزیع متقارن کروی، توزیع نرمال چندمتغیره، فضای پارامتر محدود شده، مخاطره

چکیده

در این پایان‌نامه مسئله برآورد بردار پارامتر مکان p بعدی θ در کلاس توزیع‌های متقارن کروی با پارامتر ماتریس مقیاس معلوم و مجهول، تحت محدودیت را مورد بررسی قرار می‌دهیم. سه محدودیت بر روی بردار پارامتر θ در نظر می‌گیریم. ابتدا فرض می‌کنیم همه مؤلفه‌های بردار θ نامنفی باشند و سپس فرض می‌کنیم تنها زیر مجموعه‌ای از مؤلفه‌های بردار θ نامنفی باشند. به عنوان محدودیت سوم حالتی را در نظر می‌گیریم که در آن پارامتر θ متعلق به مخروط $\mathscr{C} \in \mathbb{R}^p$ است. هدف یافتن کلاسی از برآوردگرهای انقباضی برتر برای پارامتر مکان مجهول θ، در فضای پارامتر محدود شده، تحت توابع زیان مربعی و مربعی موزون است. ملاک برتری برآوردگرها کمتر بودن نسبی مخاطره آن‌ها از برآوردگر طبیعی است.

خاضعانه و خالصانه تقدیم می گردد به محضر نورانی

امام مهدی (عج)

سپاس گزاری

دَعْواهُمْ فیها سُبْحانَکَ اللَّهُمَّ وَ تَحِیَّتُهُمْ فیها سَلامٌ وَ آخِرُ دَعْواهُمْ أَنِ الْحَمْدُ لِلَّهِ رَبِّ الْعالَمینَ.
نیایش آنان در آنجا، این است که خدایا تو پاك و منزّهی و درودشان در آنجا سلام است و پایان نیایش آنان این است که ستایش ویژه پروردگار جهانیان است (سوره یونس آیه ۱۰).

سپاس خداوندی را که بخشنده، مهربان و شایسته پرستش است. خداوندی که با لطف بی‌کران خود، آدمی را زیور عقل آراست و وی را کیمیای محبت کرامت فرمود.

در آغاز وظیفه خود می‌دانم از زحمات بی‌دریغ استاد راهنمای خود، جناب آقای دکتر محمد آرشی، صمیمانه تشکر و قدردانی کنم. ایشان نه تنها راهنمای مسائل علمی بلکه راهنمای عملی راه و رسم زندگی نیز بودند. همچنین وظیفه خود می‌دانم از تمامی اساتید محترم گروه آمار دانشگاه شاهرود آقایان داود شاهسونی، احمد نزاکتی رضازاده و اساتید محترم داور آقایان حسین باغیشنی و مهدی روزبه قدردانی به عمل بیاورم.

هر چند این اوراق در بیان ارادت قلبی، به پدر و مادر عزیزم ناتوان‌اند، اما به رسم ادب، صمیمانه و خالصانه، از پدرم برای صبر و فداکاری‌اش و مادرم به خاطر گذشت و مهربانی‌اش، تشکر و قدردانی به عمل می‌آورم. همچنین از برادر و خواهرانم به ویژه سرکار خانم کرمی کبیر که از کودکی معلم ریاضی بنده بوده‌اند، کمال تشکر را دارم.

وظیفه خود می‌دانم تا از دوستانم آقایان احسان اسحقی و میعادولیپور بواسطه همراهی و هم‌فکری‌شان تشکر نمایم. همچنین از دوست عزیزم آقای پیمان برآبادی که آشنایی با ایشان تاثیری شگرف در اینجانب داشت، از صمیم قلب سپاسگذارم.

در پایان از تمامی دانشجویان آمار و ریاضی دانشکده ریاضی و همچنین از مسئولین محترم آموزش دانشکده جناب آقای حسین‌پور و سرکار خانم خداوردی به پاس لطف و محبت‌هایشان قدردانی می‌نمایم.

حمید کرمی کبیر
اردیبهشت ۱۳۹۲

فهرست مقالات مستخرج از متن پایان نامه

Karamikabir H. and Arashi M. (2012), " A Note On Shrinkage Estimator For Restricted Parameter Space" **Communications in Statistics - Theory and Methods**, Under Review.

کرمی کبیر، ح. و آرشی، م. (۱۳۹۱)، " رآوردگر انقباضی در توزیع نرمال چندمتغیره تحت فضای پارامتر محدود"ب، **مجله علوم آماری**، در حال بررسی.

پیش گفتار

از زمانی که استاین (۱۹۵۶) نشان داد در توزیع نرمال، وقتی که بعد فضای پارامتر بزرگتر از ۲ است، میانگین نمونه برآوردگری غیرمجاز برای میانگین توزیع نرمال است، تاکنون تلاش‌های فراوانی در جهت بهبود برآوردگر میانگین صورت گرفته است. با این حال هنوز برآوردگر مجاز به طور یکنواخت در تمام فضای پارامتر برای این حالت به دست نیامده است.

به دست آوردن برآوردگر بهبود یافته نه تنها در توزیع نرمال بلکه در خانواده توزیع‌های بیضی‌گون به خصوص توزیع متقارن کروی نیز مورد توجه قرار دارد. بدین منظور در این پایان‌نامه در جستجوی یافتن کلاسی از برآوردگرهای انقباضی بهبود یافته برای پارامتر مکان مجهول θ، که در فضای محدود شده قرار دارد، نسبت به برآوردگر طبیعی هستیم. در این راستا با در نظر گرفتن سه نوع محدودیت بر روی فضای پارامتر مکان توزیع متقارن کروی با پارامتر مقیاس معلوم و مجهول، برآوردگرهای انقباضی بهبود یافته را معرفی خواهیم کرد. این مجموعه شامل ۴ فصل و ۳ پیوست می‌باشد. مطالب هر فصل به طور خلاصه عبارتند از:

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

- در فصل دوم با قرار دادن سه نوع محدودیت روی بردار پارامتر θ، رده‌ای از برآوردگرهای انقباضی به صورت

$$\delta(X,U) = \delta_\circ(X) + g(X)U^TU,$$

که تحت تابع زیان مربعی بهتر از برآوردگر طبیعی $\delta_\circ(X)$ است را خواهیم یافت. مطالب این فصل عمدتاً از منبع فودینیر و همکاران (۲۰۰۳a) می‌باشد.

- در فصل سوم همانند فصل دوم با قرار دادن دو محدودیت بر روی بردار پارامتر θ، در توزیع متقارن کروی با پارمتر مقیاس $\sigma^2 I$ با σ^2 مجهول و همچنین در توزیع نرمال p – متغیره با پارامتر مکان θ و پارامتر مقیاس $\sigma^2 I$ با σ^2 مجهول ($N_p(\theta,\sigma^2 I_p)$) رده‌ای از برآوردگرهای

انقباضی به صورت

$$\delta(X,U)=\delta_\circ(X)+g(X,S)U^TU,$$

که تحت تابع زیان درجه دوم بهتر از برآوردگر طبیعی $\delta_\circ(X)$ است را خواهیم یافت.

- در فصل چهارم ضمن ارائه یک مثال واقعی برای پی بردن چگونگی کار با برآوردگر انقباضی، با استفاده از شبیه‌سازی برتری برآوردگر انقباضی را در توزیع نرمال چند متغیره تحت محدودیت کامل، به تصویر خواهیم کشید.

- در پیوست‌ها ابتدا به تعاریف جبر خطی مورد نیاز که در متن پایان‌نامه به آن‌ها اشاره شده است، می‌پردازیم. سپس در بخش دوم، به ارائه پیشنهادات پرداخته و در پایان دستورهای مورد نیاز برنامه‌نویسی بر اساس نرم‌افزار R، که در پایان‌نامه استفاده شده، را می‌آوریم.

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

فهرست مطالب

لیست تصاویر

فصل ۱
مقدمه و دورنما

۱.۱ مقدمه

در تحلیل آماری، استفاده از اطلاعات پیشین بر روی تعدادی یا تمام پارامترهای مدل معمولاً منجر به استفاده از روش‌های پیشرفته و تعمیم یافته ای می‌شود، که با توجه به قطعی یا غیرقطعی بودن این اطلاعات، به کار برده می‌شود. در حالت اول، اعمال آن بر روی مدل به صورت قید[۱]، باعث به وجود آمدن مدل محدودشده[۲] می‌شوند. برآوردگری که بر اساس مدل‌های محدودشده به دست می‌آید، برآوردگر محدودشده[۳] نامیده می‌شود. اعتبار و کارآیی برآوردگرهای محدودشده بر اساس قید اعمال شده بر روی فضای پارامتر مورد سنجش و بررسی قرار می‌گیرند، در حالی که برآوردگرهای محدودنشده در کل فضای پارامتر بررسی می‌شوند. بنابراین، نتایج بدست آمده بر پایه مدل‌های محدودشده و محدودنشده تنها با سنجش درست و اعتبار محدودیت‌ها و قیود اعمال شده، می‌تواند معقول و کاربردی باشد (برای آگاهی بیشتر در این زمینه آرشی، ۱۳۸۷ را ببینید).

در سال ۱۹۵۶، چارلز استاین نشان داد که در توزیع نرمال p - متغیره ($p \geq ۳$)، برآوردگر طبیعی[۴] میانگین نمونه تحت تابع زیان مربعی، غیرمجاز[۵] است. این موضوع سبب ایجاد انگیزه برای بهبود برآوردگرها در مسائل مختلف برآوردیابی شد. سپس جیمز و استاین (۱۹۶۱) برآوردگری ارائه دادند

[۱] Constraint
[۲] Restricted model
[۳] Restricted estimator
[۴] Natural estimator
[۵] Inadmissible

که بر میانگین نمونه برتری داشت. به عبارت دیگر این برآوردگر دارای مخاطره کمتری نسبت به میانگین نمونه بود. این یافته تحول بزرگی در شاخه‌های مختلف آمار ایجاد کرد. زیرا پس از آن افراد زیادی تلاش کردند برآوردگرهایی از نوع جیمز ـ استاین ارائه دهند که در شرایط مختلف فضای پارامتر دارای مخاطره کمتر باشد.

برآوردگر جیمز ـ استاین برآوردگری غیرخطی است، که بر برآوردگر خطی میانگین نمونه برتری دارد. در نتیجه تمامی خواص خوب برآوردگر میانگین نمونه‌ای، از قبیل نااریبی و پایایی، که بر اساس نظریه کمترین توان‌های دوم و درستنمایی ماکزیمم بدست می‌آیند، تحت تاثیر برآوردگر جیمز- استاین قرار گرفتند. بر این اساس به افتخار پرفسور چارلز استاین، برآوردگری که به طور یکنواخت بر برآوردگر استاندارد (برآوردگر کمترین توان‌های دوم[۶] (LSE) و برآوردگر درستنمایی ماکزیمم[۷] (MLE)) برتری دارد، را برآوردگر نوع استاین [۸] (SE) می‌نامند. فرض کنید $X \in \mathbb{R}^{۲}$ برآوردگر طبیعی پارامتر مکان باشد، ساختار کلی برآوردگرهای نوع استاین مکان به صورت زیر است.

$$\delta_a^{JS}(X) = \left(۱ - \frac{a}{\|X\|^{۲}}\right)X, \quad ۰ < a < ۲(p-۲).$$

که در آن p بعد فضای پارامتر است. بنابه قضیه گوس–مارکف[۹]، برآوردگر LS در کلاس برآوردگرهای نااریب خطی دارای کمترین واریانس است (شائو، ۲۰۰۳). اما ممکن است در ازای اریب شدن برآوردگرها به برآوردگری دست پیدا کنیم که دارای واریانسی (مخاطره) به مراتب کمتر از برآوردگر LS باشد. برآوردگر انقباضی [۱۰] یکی از برآوردگرهایی است که دارای این خاصیت است.

برآوردگر انقباضی، برآوردگری است که به طور واضح یا مجازی در اثر انقباض به دست می‌آید (تعریف کامل این برآوردگر در ادامه خواهد آمد). شکل کلی برآوردگر انقباضی به صورت $X+g(X)$ است که در آن g یک تابع اندازه‌پذیر می‌باشد. همانطور که مشخص است برآوردگر استاین نیز با فرض $g(X) = -\frac{a}{\|X\|^{۲}}X$، یک برآوردگر انقباضی است. لذا به برآوردگر نوع استاین، برآوردگر انقباضی نوع استاین [۱۱] (SSE) نیز گفته می‌شود. به طور کلی، آماره آزمون به کار رفته در برآوردگر

[۶] Least Squares Estimator
[۷] Maximum-Likelihood Estimator
[۸] Stein-type Estimator
[۹] Gauss–Markov Theorem
[۱۰] Shrinkage estimator
[۱۱] Stein-type Shrinkage Estimator

نوع استاین، فاصله بین برآوردگر محدودشده و برآوردگر محدودنشده را اندازه می‌گیرد.

در اغلب پدیده‌های طبیعی، با توجه به شرایط خاص طبیعت، ممکن است محدودیت‌هایی روی صفت‌ها و مشخصه‌های متغیر گذاشته شود که این مورد، موجب تخمین این مشخصه‌ها در مدل‌های محدودشده می‌شود. مثلاً در یک فرآیند تولید لامپ، متوسط طول عمر همواره عددی مثبت است، حال اگر علاقه‌مند به تخمین طول عمر باشیم، باید محدودیت مثبت بودن را به مدل مورد بررسی اضافه کنیم. لازم به ذکر است که چنانچه از توزیع گاما برای مدل‌سازی طول عمر لامپ‌ها استفاده کنیم، محدودیت مثبت بودن در ذات تکیه‌گاه و فضای پارامتر توزیع است. ولی چنانچه از توزیع نرمال (بنابه اطلاعات نمونه) استفاده کنیم، آن‌گاه باید حتماً محدودیت مثبت بودن را بر روی برآوردگر پارامتر مکان اعمال کنیم، چون طول عمر نمی‌تواند نامثبت باشد. بنابراین برآورد پارامتر محدودشده، رده‌ای وسیع از مثال‌های موجود در دنیای واقعی را شامل می‌شود. چرا که در اغلب این مثال‌ها محدودیت‌هایی وجود دارد که توسط آزمایشگر اتخاذ می‌شود. برای آگاهی بیشتر و مشاهده مثالی کاربردی به تجدد (۱۳۹۱)، در زمینه کاربرد برآوردگر انقباضی در بازی بیس‌بال، مراجعه کنید.

برآوردگر انقباضی از دو جهت در برآورد پارامتر محدودشده دارای اهمیت است. یکی این‌که این برآوردگر، مقدار مخاطره را در حالت کلی نسبت به برآوردگر LS کاهش می‌دهد و دیگر این‌که نقش تابع $g(X)$ در اضافه کردن اثر محدودیت به برآوردگر اولیه X غیر قابل چشم‌پوشی است. از این‌رو در این پایان‌نامه به دنبال یافتن برآوردگر انقباضی مناسب و بررسی رفتار آن در مدل‌های محدودشده هستیم.

مسئله برآورد پارامتر مکان محدودشده، در سال‌های اخیر پیشرفت های متعددی یافته است. در این زمینه، کسلا و استرادرمان[۱۲] (۱۹۸۱) و بیکل[۱۳] (۱۹۸۱) مسئله برآورد پارامتر مکان را تحت محدودیت $|\theta| < m$ در توزیع نرمال یک متغیره مورد مطالعه قرار دادند. در همین راستا، گاتسونیس[۱۴] و همکاران (۱۹۸۵) برآوردگر بیز[۱۵] را با محدودیت مشابه با کسلا و استرادرمان

[۱۲] Strawderman
[۱۳] Bickel
[۱۴] Gatsonis
[۱۵] Bayes estimator

برای میانگین توزیع نرمال ارائه کردند. همچنین کاریا[۱۶](۱۹۸۹)، پرون و جیری[۱۷](۱۹۸۹) و مارچاند[۱۸](۱۹۹۴)، توزیع نرمال p - متغیره $N_p(\mu, \Sigma)$ را با محدودیت $\mu^T\Sigma^{-1}\mu = \lambda_0$ (برای μ و Σ مجهول)، در نظر گرفتند. مارچاند (۱۹۹۳) توزیع‌های متقارن کروی برای θ (اساساً برای توزیع‌های نرمال آمیخته) را با محدودیت $\|\theta\| = (\theta^T\theta)^{1/2} = \lambda_0$ در نظر گرفت. وی بهترین برآوردگر هم‌پایا [۱۹](BEE) را یافت و ثابت کرد که BEE بهتر از برآوردگر درستنمایی ماکزیمم[۲۰] و بهترین برآوردگر خطی نااریب[۲۱] (BLUE) است. مارچاند و جیری (۱۹۹۳) کلاسی از برآوردگرهای از نوع جیمز ـ استاین را در نظر گرفته و نتایج را برای توزیع‌های نرمال آمیخته مقیاسی[۲۲] تعمیم دادند.

فودینیر و اوسو [۲۳](۲۰۰۰)، توزیع کروی[۲۴] معمولی وقتی که مشاهدات به شکل (X, U) و دارای توزیع متقارن کروی[۲۵] حول بردار $(\theta, 0)$ بودند را در حالت θ محدودشده مورد بررسی قرار دادند. در ادامه، وان[۲۶] و همکاران (۲۰۰۰) برآوردگرهای مینیماکس[۲۷] و گاما-مینیماکس[۲۸] در توزیع پواسن، وقتی که فضای پارامتر به فاصله $\theta \in [0, \beta]$ متعلق است را بدست آوردند. همچنین اوسو و استرادرمان (۲۰۰۲) توزیع متقارن کروی را با محدودیت روی مخروط‌ها مورد بررسی قرار دادند و در همین راستا فودینیر و همکاران (a۲۰۰۳) در توزیع متقارن کروی با در نظر گرفتن سه محدودیت بر روی فضای پارامتر مطالعات را ادامه دادند.

در زمینه مدل‌های محدودشده، مارچاند و استرادرمان (۲۰۰۵) برای خانواده مکان با در نظر گرفتن محدودیتی به صورت $\theta > a$ کار را ادامه دادند. همچنین مارچاند و پرون (۲۰۰۵) نتایج را برای

[۱۶] Kariya
[۱۷] Perron and Giri
[۱۸] Marchand
[۱۹] Best Equivariant Estimator
[۲۰] Maximum Likelihood Estimator
[۲۱] Best Linear Unbiased Estimator
[۲۲] Scale mixture of normal distribution
[۲۳] Fourdrinier and Ouassou
[۲۴] Spherical distribution
[۲۵] Spherically symmetri distribution
[۲۶] Wan
[۲۷] Minimax
[۲۸] Γ-minimax

توزیع‌های متقارن کروی تحت محدودیت

$$\Theta(m) = \{\theta \in \mathbb{R}^p : \|\theta\| \leq m : \quad m > 0\}.$$

بدست آوردند. مارچاند و پارسیان (۲۰۰۶)، محدودیت به شکل $\theta \in [0, m]$ را برای کلاس وسیعی از مدل‌های گسسته در نظر گرفتند. در ادامه فودینیر و همکاران (۲۰۰۶) و فودینیر و مارچاند (۲۰۱۰) برای توزیع‌های متقارن کروی، پارامتر مکان را با نوع دیگری از محدودیت‌ها برآورد نمودند.

در توزیع نرمال چندمتغیره[۲۹]، زین‌الدینی (۱۳۸۸) ابتدا به برآورد میانگین در حالت‌هایی که ماتریس کوواریانس معلوم یا مجهول می‌باشد از دیدگاه نظریه تصمیم پرداخت و سپس برآورد میانگین توزیع‌های غیرنرمال و نرمال آمیخته مقیاسی تحت تابع زیان درجه دوم را مورد بررسی قرار داد.

اخیراً مارچاند و استرادرمان (۲۰۱۲) رهیافت واحدی را برای برآورد مینیماکس در فضای پارامتر محدودشده ارائه داده‌اند. در نهایت کورتبی[۳۰] و مارچاند (۲۰۱۲) برآوردگر خطی بریده‌شده[۳۱] با محدودیت $\|\theta\| < m$ در توزیع نرمال چندمتغیره را ارائه کرده و رفتار آن را مورد بررسی قرار دادند.

در این پایان‌نامه، فرض می‌کنیم که جامعه مورد بررسی دارای توزیع متقارن کروی است. در این راستا بردار پارامتر p بعدی $\theta = (\theta_1, \ldots, \theta_p)$، وقتی که مشاهدات به صورت (X, U)، که در آن بعد X برابر با p و بعد U برابر با k بوده را در نظر می‌گیریم. به عبارتی (X, U) دارای توزیع متقارن کروی حول $(\theta, 0)$ است. سپس پارامتر θ را تحت محدودیت‌های مختلف برآورد کرده و برآوردگرهایی از نوع انقباضی ارائه می‌دهیم.

---

[۲۹] Multivariate normal distribution

[۳۰] Kortbi

[۳۱] Truncated linear estimtor

۲.۱ تعاریف

در این بخش تعاریف اساسی که در طول این مجموعه از آن‌ها استفاده می‌شود را می‌آوریم.

۱.۲.۱ تعاریف و توابع ریاضی

برای اطلاع دقیقتر از تعاریف و قضایای این بخش می‌توان به رودین[۳۲] (۱۳۷۳)، میرهد[۳۳] (۱۹۸۲)، لهمن و کسلا[۳۴] (۱۹۹۸)، مدقالچی (۱۳۸۴)، بوید و واندربرگ[۳۵] (۲۰۰۹)، گات [۳۶](۲۰۰۵)، آرشی (۱۳۸۷) و راس[۳۷](۲۰۱۰) مراجعه کرد.

تعریف ۱.۱.۲.۱. تابع گاما[۳۸]

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

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

تعریف ۲.۱.۲.۱. تابع بتا[۳۹]

تابع بتا را به ازای اعداد حقیقی $p, k > 0$ با نماد $\mathbf{B}(p,k)$ نشان داده و به صورت زیر تعریف می‌کنیم.

$$\mathbf{B}(p,k) = \int_0^1 x^{p-1}(1-x)^{k-1} dx.$$

لازم به ذکر است که بین تابع بتا و تابع گاما رابطه زیر برقرار است.

$$\mathbf{B}(p,k) = \frac{\Gamma(p)\Gamma(k)}{\Gamma(p+k)}.$$

---
[۳۲] Rudin
[۳۳] Muirhead
[۳۴] Lehmann and Casella
[۳۵] Boyd and Vandenberghe
[۳۶] Gut
[۳۷] Ross
[۳۸] Gamma function
[۳۹] Beta function

تعریف ۳.۲.۱. مجموعه محدب[۴۰]

مجموعه C محدب است، اگر پاره خط بین هر دو نقطه دلخواه در C، درون C قرار گیرد. به عبارتی به ازای هر $x_۱$ و $x_۲$ عضو C و به ازای هر θ ثابت، با شرط $۱ \leq \theta \leq ۰$، داشته باشیم

$$\theta x_۱ + (۱-\theta)x_۲ \in C.$$

تعریف ۴.۲.۱. تابع محدب[۴۱]

تابع $f : \mathbb{R}^n \to \mathbb{R}$ محدب است، اگر دامنه f یک مجموعه محدب باشد و به ازای هر X, Y عضو دامنه f و $۱ \leq \theta \leq ۰$ داشته باشیم

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

تابع $f(\cdot)$ را مقعر [۴۲] گوییم اگر $(-f(\cdot))$ محدب باشد.

تعریف ۵.۲.۱. مخروط[۴۳]

مجموعه C مخروط است، اگر به ازای هر x عضو مجموعه C و $\theta \geq ۰$، θx عضو C باشد.

تعریف ۶.۲.۱. مخروط محدب[۴۴]

مجموعه C را مخروط محدب گوییم، اگر هم محدب و هم مخروط باشد. به عبارت دیگر به ازای هر $x_۱$ و $x_۲$ عضو مجموعه C و $\theta_۱, \theta_۲ \geq ۰$، $\theta_۱ x_۱ + \theta_۲ x_۲$ عضو C باشد.

تعریف ۷.۲.۱. مخروط مثبت[۴۵]

مخروط C در $\mathbb{R}^p$ را مخروط مثبت گویند، اگر

۱. به ازای هر $X \in C$ و $\theta \geq ۰$. آن‌گاه $\theta X \in C$.

۲. اگر $X_۱, X_۲ \in C$، آن‌گاه $X_۱ + X_۲ \in C$،

۳. اگر $X \in C$ و $-X \in C$ ، آن‌گاه $X = 0$.

[۴۰] Convex set
[۴۱] Convex function
[۴۲] Concave function
[۴۳] Cone
[۴۴] Convex cone
[۴۵] Positive cone

تعریف ۱.۲.۸. نرم[۴۶]

تابع $f : \mathbb{R}^p \to \mathbb{R}$ را نرم گوییم، اگر

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

۲. $f(\cdot)$ معین[۴۷] باشد، یعنی $f(X) = ۰$ اگر و تنها اگر $X = 0$.

۳. $f(\cdot)$ همگن[۴۸] باشد، یعنی به ازای هر $X \in \mathbb{R}^p$ و $t \in \mathbb{R}$ داشته باشیم، $f(tX) = |t|f(X)$.

۴. $f(\cdot)$ در نامساوی مثلثی[۴۹] صدق کند، یعنی به ازای هر $X, Y \in \mathbb{R}^p$ داشته باشیم

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

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

نکته ۱.۲.۹. (بوید و واندربرگ ، ۲۰۰۹) فرض کنید $\|\cdot\|$ یک نرم دلخواه روی $\mathbb{R}^p$ باشد. مخروط مرتبط با نرم $\|\cdot\|$، مجموعه

$$C = \{(X,t) | \|X\| \leq t\} \subseteq \mathbb{R}^{p+۱},$$

است که محدب می‌باشد (شکل ۱.۱ را ببینید).

تعریف ۱.۲.۱۰. نرم اقلیدسی[۵۰]

$\|X\|$ را روی بردار $X = (X_۱, \ldots, X_p) \in \mathbb{R}^p$ نرم اقلیدسی گوییم هرگاه

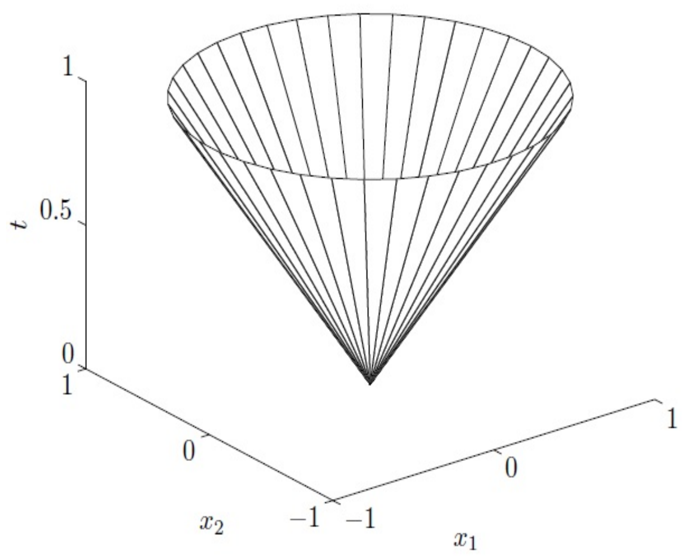
$$\|X\| = (X^T X)^{\frac{۱}{۲}} = (\sum_{i=۱}^{p} X_i^۲)^{۱/۲}.$$

نکته ۱.۲.۱۱. (بوید و واندربرگ، ۲۰۰۹) نرم اقلیدسی یک نرم اکیداً محدب است.

در این پایان‌نامه هر جا سخن از نرم در میان باشد، منظور نرم اقلیدسی است.

---

[۴۶] Norm
[۴۷] Definite
[۴۸] Homogeneous
[۴۹] Triangle inequality
[۵۰] Euclidean norm



شکل ۱.۱: کران مخروط مرتبه دوم در $\mathbb{R}^۳$، $\{(x_۱, x_۲, t)|(x_۱^۲ + x_۲^۲)^{۱/۲} \leq t\}$.

تعریف ۱۲.۲.۱. گوی[۵۱](گوی باز)

فرض کنید X و θ عضو $\mathbb{R}^p$ باشند. به ازای $\rho > ۰$ مجموعه

$$B_{\theta,\rho} = \{X| \|X - \theta\| < \rho\},$$

را یک گوی باز به مرکز θ و شعاع ρ می‌نامیم. به همین ترتیب

$$B_{[\theta,\rho]} = \{X| \|X - \theta\| \leq \rho\},$$

را یک گوی بسته[۵۲] به مرکز θ و شعاع ρ می‌نامیم.

نکته ۱۳.۲.۱. (رودین، ۱۳۷۳) فرض کنید $\|\cdot\|$ یک نرم دلخواه روی $\mathbb{R}^p$ باشد. بنابه خواص کلی نرم‌ها می‌توان نشان داد که یک گوی به شعاع ρ و مرکز θ ($\{X| \|X - \theta\| \leq \rho\}$)، مجموعه‌ای محدب است.

تعریف ۱۴.۲.۱. کره[۵۳]

فرض کنید X و θ عضو $\mathbb{R}^p$ باشند. به ازای $\rho > ۰$ مجموعه

$$S_{\theta,\rho} = \{X| \|X - \theta\| = \rho\},$$

[۵۱] Open ball
[۵۲] Closed ball
[۵۳] Sphere

را یک کره به مرکز θ و شعاع ρ می‌نامیم. به عبارت دیگر کره عبارت است از مجموعه نقاط در $\mathbb{R}^p$ به طوری که نرم آن‌ها از یک نقطه ثابت مانند $\theta \in \mathbb{R}^p$ برابر ρ باشد.

تعریف ۱.۲.۱۵. نقطه درونی مجموعه[۵۴]

فرض کنید $X \subseteq \mathbb{R}^p$. نقطه $\theta \in X$ را یک نقطه درونی مجموعه X می‌نامیم در صورتی که گوی بازی مانند $B_{\theta,\rho}$ وجود داشته باشد به طوری که جزء X باشد. به عبارت دیگر

$$\forall\, \theta \in X \;\; \exists \quad B_{\theta,\rho} \subseteq X.$$

تعریف ۱.۲.۱۶. نقطه انباشتگی مجموعه[۵۵](نقطه حدی مجموعه)

فرض کنید $X \subseteq \mathbb{R}^p$. نقطه $\theta \in \mathbb{R}^p$ را یک نقطه انباشتگی X می‌نامیم در صورتی که هر گوی باز به مرکز θ شامل نقطه‌ای از X غیر از θ باشد. به عبارت دیگر

$$\forall \;\; \theta \in \mathbb{R}^p, \;\; \forall \quad \Big(B_{\theta,\rho} - \{\theta\}\Big) \cap X \neq \varnothing.$$

تعریف ۱.۲.۱۷. مجموعه باز[۵۶]

مجموعه $X \subseteq \mathbb{R}^p$ را باز می‌نامیم در صورتی که هر نقطه آن درونی باشد و X را بسته[۵۷] می‌نامیم در صورتی که $\mathbb{R}^p - X$ باز باشد.

تعریف ۱.۲.۱۸. فاصله یک نقطه از مجموعه بسته[۵۸]

فاصله نقطه $X_\circ \in \mathbb{R}^p$ از مجموعه بسته $C \subseteq \mathbb{R}^p$، با نرم $\|\cdot\|$ را با $\operatorname{dist}(X_\circ, C)$ نمایش داده و به صورت زیر تعریف می‌کنیم.

$$\operatorname{dist}(X_\circ, C) = \inf\{\|X_\circ - X\| | X \in C\}.$$

این مقدار اینفیمم همواره موجود است (بوید و واندربرگ، ۲۰۰۹).

---

[۵۴] Interior point of a set
[۵۵] Accumulation point of a set
[۵۶] Open set
[۵۷] Closed set
[۵۸] Distance of a point from a closed set

تعریف ۱۹.۲.۱. تصویر روی یک مجموعه[۵۹]

تصویر بردار $X_\circ$ روی مجموعه C، را با نماد $P_C(X_\circ)$ نمایش داده و به ازای هر تابع $P_C:\mathbb{R}^p\to\mathbb{R}^p$، به صورت زیر تعریف می‌نماییم.

$$\forall X_\circ\in\mathbb{R}^p,\quad P_C(X_\circ)\in C,\quad \|X_\circ-P_C(X_\circ)\|=dist(X_\circ,C).$$

به عبارت دیگر داریم

$$P_C(X_\circ)=\text{argmin}\{\|X_\circ-X\|\,|X\in C\}.$$

که در آن argmin به صورت زیر تعریف می‌شود.

$$\underset{x}{\text{argmin}}f(x)=\{x|\forall y:f(y)\geq f(x)\}.$$

برای آگاهی بیشتر در این خصوص بوید و واندربرگ (۲۰۰۹) را ببینید.

نکته ۲۰.۲.۱. (بوید و واندربرگ، ۲۰۰۹) تصویر نقطه دلخواه $X_\circ\in\mathbb{R}^p$ روی مجموعه $C\subseteq\mathbb{R}^p$، یکتا است.

تعریف ۲۱.۲.۱. تابع زبرهمساز[۶۰]

اگر تابع $f:\mathbb{R}^p\to\mathbb{R}$ دارای مشتق دوم باشد، آن‌گاه f در $\mathbb{R}^p$ زبرهمساز است اگر و تنها اگر به ازای هر $X=(X_1,\ldots,X_p)\in\mathbb{R}^p$، داشته باشیم

$$\sum_{i=1}^{p}\frac{\partial^2}{\partial X_i^2}f(X)\leq \circ.$$

چنانچه علامت کوچکتری در این تعریف به تساوی تبدیل شود، تابع همساز [۶۱] و اگر به بزرگتری تبدیل شود تابع زیرهمساز [۶۲] را داریم.

[۵۹] Projection on a set
[۶۰] Superharmonic function
[۶۱] Harmonic function
[۶۲] Subrharmonic function

تعریف ۲۲.۲.۱. تابع مشتق‌پذیر ضعیف[۶۳]

تابع $p_{\theta+\Delta}(x)/p_\theta(x)$ در θ مشتق‌پذیر ضعیف است اگر تابع اندازه‌پذیر q، به ازای هر $h(\cdot)$ که در شرط

$$\int h^{۲}(x)p_\theta(x)d\mu(x) < \infty,$$

صدق می‌کند، وجود داشته باشد به طوری که

$$\lim_{\Delta\to ۰} \int h(x)\Big\{\Big[\Delta^{-۱}\Big(\frac{p_{\theta+\Delta}(x)}{p_\theta(x)} - ۱\Big)\Big] - q(x)\Big\}p_\theta(x)d\mu(x) = ۰.$$

در واقع شرط لازم و کافی برای مشتق‌پذیر ضعیف بودن، معادل است با وجود تابع $q_\theta(x)$ طوری که

$$\frac{\partial}{\partial\theta}E_\theta(\delta) = E\delta(q).$$

که در آن δq یک آماره دلخواه است طوری که $E_\theta(|\delta q|^{۲}) < \infty$.

تعریف ۲۳.۲.۱. تابع فوق هندسی تعمیم یافته [۶۴]

به ازای مقادیر حقیقی $a_۱, a_۲, \dots, a_p$ و $b_۱, b_۲, \dots, b_q$، تابع فوق هندسی تعمیم یافته به صورت زیر است

$${}_pF_q(a_۱, a_۲, \dots, a_p; b_۱, b_۲, \dots, b_q; z) = \sum_{k=۰}^{\infty} \frac{(a_۱)_k \cdots (a_p)_k}{(b_۱)_k \cdots (b_q)_k} \times \frac{z^k}{k!},$$

که در آن $(a)_k = a(a+۱)\cdots(a+k-۱)$. برای آگاهی بیشتر در خصوص این تابع، میرهد (۱۹۸۲) را ببینید.

تعریف ۲۴.۲.۱. تابع زیان مربعی[۶۵]

اگر بردار p مولفه‌ای $\delta = (\delta_۱, \dots, \delta_p)$ برآوردگری برای بردار پارامتر $\theta \in \mathbb{R}^p$ باشد، آن‌گاه زیان استفاده از δ در خصوص برآورد θ را با $L(\theta,\delta)$ نشان داده که برابر است با

$$L(\theta,\delta) = (\delta-\theta)^T(\delta-\theta) = \|\delta-\theta\|^{۲} = \sum_{i=۱}^{p}(\delta_i-\theta_i)^{۲},$$

[۶۳] Weakly differentiable function
[۶۴] Generalized hypergeometric function
[۶۵] Quadratic loss function

چنان‌چه پارامتر مقیاس مجهول $\sigma^2 \geq 0$ در مدل موجود باشد، آن‌گاه از تابع زیان پایا[۶۶] زیر استفاده می‌کنیم.

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

تعریف ۲۵.۲.۱. تابع مخاطره [۶۷]

با استفاده از تعریف تابع زیان، تابع مخاطره برآوردگر δ را با $\mathscr{R}(\theta, \delta)$ نشان داده که برابر است با

$$\mathscr{R}(\theta, \delta) = E_\theta[L(\theta, \delta)] = \int_\chi L(\theta, \delta) f_\theta(x) dx.$$

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

لم ۲۶.۲.۱. ***نامساوی هولدر*** [۶۸]

فرض کنید p و q دو عدد در $\mathbb{R}$ باشند به طوری که $p^{-1}+q^{-1}=1$ و X و Y دو متغیر تصادفی روی χ باشند. اگر $E|X|^p < \infty$ و $E|Y|^q < \infty$، آن‌گاه

$$|E[XY]| \leq E|XY| \leq (E|X|^p)^{1/p} \cdot (E|Y|^q)^{1/q}.$$

نتیجه ۲۷.۲.۱. ***نامساوی کوشی - شوارتز*** [۶۹]

به ازای $p = q = 2$ در نامساوی هولدر، نامساوی کوشی - شوارتز بدست می‌آید.

۲.۲.۱ توزیع‌های آماری

برای اطلاع دقیقتر از تعاریف و قضایای این بخش می‌توان به میرهد (۱۹۸۲)، لهمن و کسلا (۱۹۹۸)، گات (۲۰۰۵)، پارسیان (۱۳۸۶)، آرشی (۱۳۸۷)، زین الدینی (۱۳۸۸) و راس (۲۰۱۰) مراجعه کرد.

تعریف ۲۸.۲.۱. فرض کنید δ_1 و δ_2 دو برآوردگر $g(\theta)$ باشند. اگر به ازای هر $\theta \in \Theta$، داشته باشیم

$$\mathscr{R}(\theta, \delta_1) \leq \mathscr{R}(\theta, \delta_2),$$

---
۶۶ Invariant loss function
۶۷ Quadratic risk function
۶۸ Holder inequality
۶۹ Cauchy-Schwarz inequality

و به ازای حداقل یک $\theta \in \Theta$ نامساوی اکید باشد، آن‌گاه برآوردگر δ_1 را بهتر (برتر) از برآوردگر δ_2 گوییم. به عبارت دیگر گوییم δ_1 اکیداً بر δ_2 غلبه دارد.

تعریف ۲۹.۲.۱. برآوردگر غیرمجاز

برآوردگر δ را غیرمجاز گوییم هرگاه برآوردگر δ^* وجود داشته باشد به طوری که به ازای هر $\theta \in \Theta$ داشته باشیم

$$\mathscr{R}(\theta, \delta^*) \leq \mathscr{R}(\theta, \delta),$$

و حداقل یک مقدار $\theta \in \Theta$ وجود داشته باشد به طوری که نامساوی اکید باشد. هرگاه چنین برآوردگری وجود نداشته باشد، آن‌گاه δ را برآوردگر مجاز[۷۰] گویند.
به بیانی دیگر هرگاه هیچ برآوردگری موجود نباشد که از δ برتر باشد، گوییم δ یک برآوردگر مجاز است و اگر δ برآوردگری مجاز نباشد آن را غیرمجاز می‌خوانیم.

چون مجاز بودن می‌تواند خصوصیت مطلوبی باشد پس تعیین تمام برآوردگرهای مجاز مورد توجه است. البته باید این نکته را مورد توجه قرار داد که مجاز بودن به تنهایی یک معیار ضعیف برای بهینگی است و برآوردگرهای مجاز فراوانی می‌توان یافت. از جمله برآوردگر ثابت $\delta(X) = \theta_0$ که هیچ اطلاعی از X را به کار نمی‌برد، مجاز است. پس تنها بر اساس معیار مجاز بودن نمی‌توانیم این مسئله، که کدام برآوردگر باید انتخاب شود را حل کنیم (زین‌الدینی، ۱۳۸۸).

تعریف ۳۰.۲.۱. برآوردگر مینیماکس [۷۱]

به برآوردگر δ^M که ماکزیمم مخاطره را مینیمم می‌کند، را برآوردگر مینیماکس گویند. یعنی

$$\inf_{\delta} \sup_{\theta} \mathscr{R}(\theta, \delta) = \sup_{\theta} \mathscr{R}(\theta, \delta^M).$$

به عبارت دیگر برآوردگر را مینیماکس گویند هرگاه به ازای هر برآوردگر δ داشته باشیم

$$\sup_{\theta} \mathscr{R}(\theta, \delta^M) \leq \sup_{\theta} \mathscr{R}(\theta, \delta).$$

قضیه ۳۱.۲.۱. (*زین الدینی، ۱۳۸۸*) *هر برآوردگری که بهتر از برآوردگر مینیماکس δ باشد، مینیماکس است.*

[۷۰] Admissible
[۷۱] Minimax estimator

تعریف ۳۲.۲.۱. 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 نمونه‌ای یعنی برآوردگر خام یا اولیه (MLE، برآوردگر کمترین توان‌های دوم، بیز و غیره) با اطلاعات غیر نمونه‌ای که معمولاً در قالب یک یا چند محدودیت به مدل اضافه می‌شوند، به صورت برآوردگری جدید بهبود می‌یابد. این بهبود در جهت مقادیری است که از اطلاعات دیگر به دست آمده نه در جهت برآوردگر اولیه. در حالت کلی بهبود برآوردگرها را می‌توان با معیار مخاطره یا توان دوم خطا (MSE) [۷۲] اندازه‌گیری کرد. در این صورت انقباض می‌تواند به سمت صفر یا یک مقدار ثابت باشد. اثر این انقباض معمولا به صورت تبدیل یک برآوردگر نااریب به اریب ولی با MSE کمتر می‌باشد. در عین حال، می‌توان برآوردگرهایی انقباضی از نوع استاین و نااریب را نیز یافت (برای آگاهی بیشتر در این خصوص به نوروزی‌راد، ۱۳۹۰ مراجعه کنید).

شکل کلی برآوردگر انقباضی به صورت $X + g(X)$ است که در آن g یک تابع اندازه‌پذیر می‌باشد. پس از حل نامعادله‌ی

$$\forall\ \theta \in \Theta \ : \qquad \mathscr{R}(\theta, X + g(X)) \leq \mathscr{R}(\theta, X).$$

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

۱. برآوردگر انقباضی به سمت صفر $(1 - c)X = X - cX$.

۲. برآوردگر انقباضی به سمت یک عدد ثابت m

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

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

[۷۲] Mean Squared Error

تعریف ۱.۲.۳۳. متغیر تصادفی تباهیده[۷۳]

X را یک متغیر تصادفی تباهیده می‌نامیم، در صورتی که $a \in \mathbb{R}$ وجود داشته باشد به طوری که $P(X = a) = ۱$. چنانچه متغیر تصادفی تباهیده نباشد، آن را متغیر تصادفی ناتباهیده[۷۴] گوییم.

تعریف ۱.۲.۳۴. تابع توزیع تباهیده[۷۵]

تابع توزیع یک متغیر تصادفی تباهیده را تابع توزیع تباهیده نامیم. اگر F تابع توزیع تباهیده باشد، آن‌گاه $a \in \mathbb{R}$ وجود دارد به طوری که

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

چنانچه تابع توزیع تباهیده نباشد، آن را تابع توزیع ناتباهیده[۷۶] گوییم.

تعریف ۱.۲.۳۵. توزیع نرمال چندمتغیره

بردار p مولفه‌ای X دارای توزیع نرمال p - متغیره با میانگین $\mu \in \mathbb{R}^p$ و ماتریس کوواریانس Σ است و با $X \sim N_p(\mu, \Sigma)$ نمایش می‌دهند اگر و فقط اگر به ازای هر بردار p مولفه‌ای ثابت a، $a^T X$ دارای نرمال یک متغیره باشد.

اگر Σ یک ماتریس معین مثبت [۷۷] باشد، آن‌گاه X دارای تابع چگالی احتمال به صورت زیر است

$$f(x) = \frac{|\Sigma|^{-۱/۲}}{(۲\pi)^{p/۲}} exp\Big[\frac{-(x-\mu)^T \Sigma^{-۱}(x-\mu)}{۲}\Big],$$

که در آن $|\Sigma|^{۱/۲}$ ریشه دوم دترمینان Σ است. توجه کنید اگر $p = ۱$، آن‌گاه صورت درجه دوم در توان نمایی، به $(x-\mu)^۲/\sigma^۲$ تبدیل می‌شود و $X \sim N(\mu, \sigma^۲)$. همچنین اگر Σ یک ماتریس نیمه معین مثبت باشد، آن‌گاه X دارای توزیع تباهیده است (ایرانمنش، ۱۳۹۱). برای مثال توزیع یک متغیره $N(۰, ۰)$ تباهیده در صفر است.

[۷۳] Degenerate random variable
[۷۴] Non-Degenerate random variable
[۷۵] Degenerate distribution function
[۷۶] Non-Degenerate distribution function
[۷۷] Positive definite matrix

تعریف ۳۶.۲.۱. توزیع با منحنی تراز بیضی‌شکل[۷۸](توزیع بیضی‌گون)

بردار n مولفه‌ای X دارای توزیع بیضی‌گون با پارامترهای μ و Σ و تابع مولد مشخصه[۷۹] ψ است و با $X \sim EC_n(\mu, \Sigma, \psi)$ نشان می‌دهیم اگر تابع مشخصه آن به ازای بردار n مولفه‌ای t به صورت زیر باشد

$$\phi_X(t) = exp(it\mu)\psi\Big(\frac{t^T\Sigma t}{2}\Big),$$

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

توزیع‌های بیضی‌گون، توزیع‌هایی هستند که منحنی‌های تراز آنها بیضی‌شکل می‌باشند (آرشی، ۱۳۸۷). در صورت وجود تابع چگالی آن، به صورت زیر می‌باشد

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

که در آن تابع $g(\cdot)$ تابع مولد چگالی[۸۰] است و دارای شرط همگرایی زیر می‌باشد.

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

به علاوه c_n ضریب نرمال‌سازی می‌باشد که به صورت زیر تعیین می‌گردد.

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

در این صورت، می‌نویسیم $X \sim EC_n(\mu, \Sigma, \psi)$. شایان ذکر است در صورتی که تابع چگالی وجود داشته باشد، تابع g، تابع ψ را مشخص می‌کند و برعکس. برای آگاهی بیشتر در این زمینه فنگ[۸۱] و همکاران (۱۹۹۰) و میرهد (۱۹۸۲) را ببینید.

با توجه به تعریف ۳۶.۲.۱، $E(X) = \mu$ و $Cov(X) = -2\psi'(0)\Sigma$ که در آن $\psi'(0)$ مشتق اول ψ در نقطه صفر است و در صورتی وجود دارد که $|\psi'(0)| < \infty$.

[۷۸] Elliptically Contoured distribution
[۷۹] Characteristic generator function
[۸۰] Density generator
[۸۱] Fang

توزیع‌های زیادی متعلق به خانواده توزیع‌های بیضی‌گون می‌باشند که از آن جمله می‌توان به توزیع‌های نرمال چندمتغیره، t چندمتغیره[۸۲] (Mt)، کوشی چندمتغیره [۸۳]، پیرسن چندمتغیره نوع II[۸۴]، پیرسن چندمتغیره نوع IIV[۸۵]، بسل چندمتغیره[۸۶]، نمایی توانی چندمتغیره[۸۷]، لاپلاس چندمتغیره[۸۸]، اسلش تعمیم‌یافته[۸۹] و کاتز چندمتغیره[۹۰] اشاره کرد.

حال به عنوان نمونه، تابع چگالی احتمال و تابع مولد چگالی را برای چند توزیع مهم بیضی‌گون، همراه با نمودار آن‌ها ارائه می‌دهیم. تمامی شکل‌های این مثال با نرم‌افزار R-2.15.3 و با استفاده از بسته[۹۱] lattice رسم شده‌اند. در پیوست **پ** دستورهای مورد نیاز برای رسم شکل‌ها در نرم‌افزار آمده‌اند.

دقت کنید که در تمامی نمودارها $n = 2$، $\mu = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ و $\Sigma = \begin{bmatrix} 1 & 0 \\ 0 & 10 \end{bmatrix}$ فرض شده است.

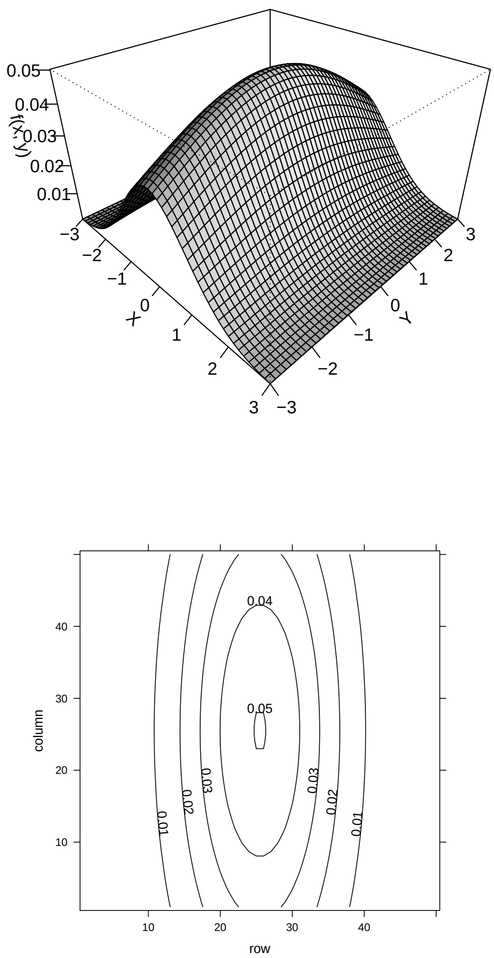
---
[۸۲] Multivariate t distribution
[۸۳] Multivariate Cauchy distribution
[۸۴] Multivariate Pearson type II
[۸۵] Multivariate Pearson type IIV
[۸۶] Multivariate Bessel
[۸۷] Multivariate exponential power
[۸۸] Multivariate Laplace
[۸۹] Generalized Slash
[۹۰] Multivariate Kots type
[۹۱] Package

۱. توزیع نرمال چندمتغیره ($X \sim N_n(\mu, \Sigma)$): توابع مولد چگالی و چگالی احتمال به ترتیب عبارتند از

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

$$f_X(x) = \frac{|\Sigma|^{-1/2}}{(\sqrt{2\pi})^n} exp\Big\{-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\Big\}.$$

نمودارهای تابع چگالی و منحنی‌های تراز در شکل ۲.۱ داده شده است.



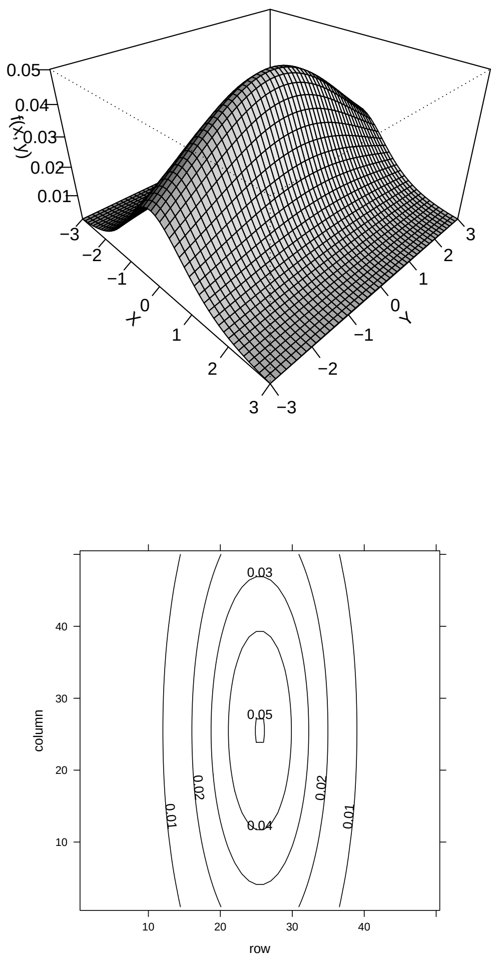
شکل ۲.۱: نمودارهای تابع چگالی و منحنی‌های تراز توزیع نرمال دومتغیره

۲. توزیع t چندمتغیره ($X \sim t_n(\mu, \Sigma, \nu)$): توابع مولد چگالی و چگالی احتمال به ترتیب عبارتند از

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

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

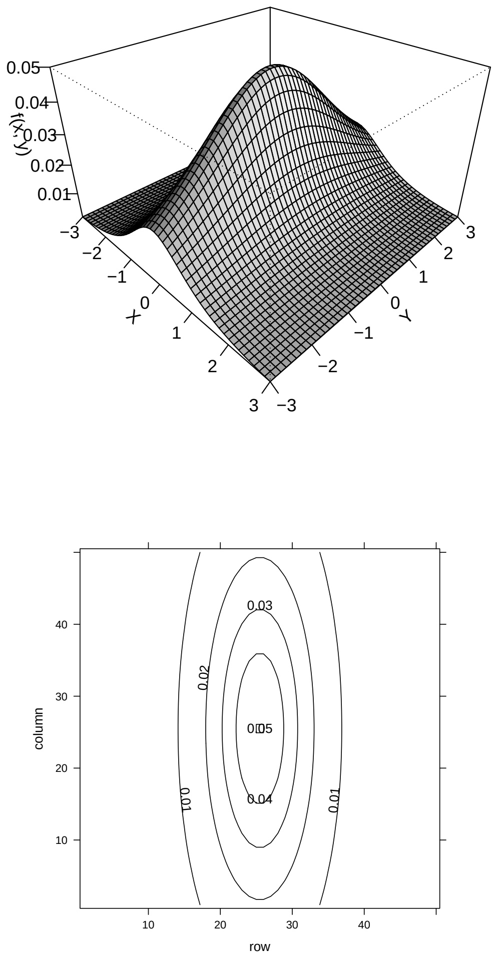
نمودارهای تابع چگالی و منحنی‌های تراز به ازای $\nu = ۳$ در شکل ۳.۱ داده شده است.



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

۳. توزیع کوشی چندمتغیره ($X \sim C_n(\mu, \Sigma, \nu)$): به ازای $\nu = ۱$، توزیع t به توزیع کوشی چندمتغیره تبدیل می‌شود.

نمودارهای تابع چگالی و منحنی‌های تراز در شکل ۴.۱ داده شده است.



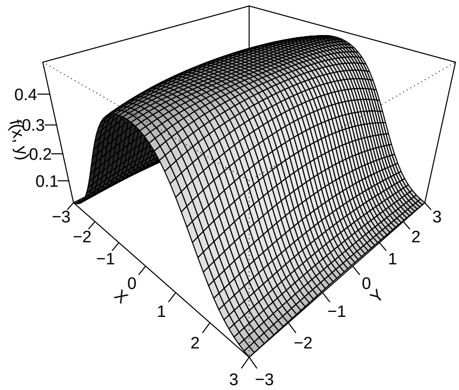
شکل ۴.۱: نمودارهای تابع چگالی و منحنی‌های تراز توزیع کوشی دومتغیره

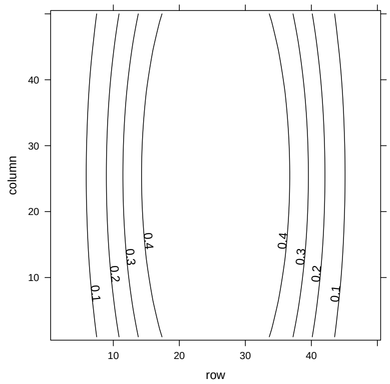
۴. توزیع لجستیک چندمتغیره ($X \sim ML_n(\mu, \Sigma)$): توابع مولد چگالی و چگالی احتمال به ترتیب عبارتند از

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

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

نمودارهای تابع چگالی و منحنی‌های تراز در شکل ۵.۱ داده شده است.



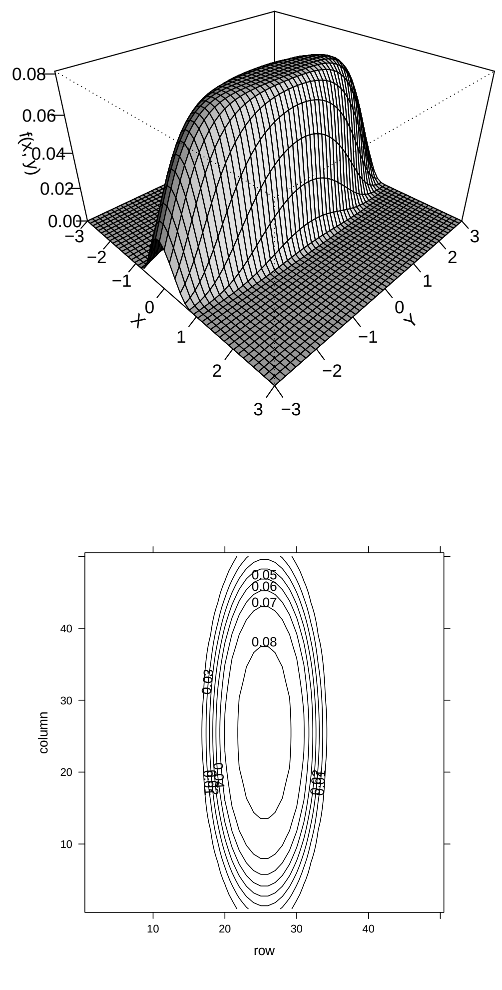


شکل ۵.۱: نمودارهای تابع چگالی و منحنی‌های تراز توزیع لجستیک دومتغیره

۵. توزیع نمایی توانی چندمتغیره ($X \sim EP_n(\mu, \Sigma, r, s)$): توابع مولد چگالی و چگالی احتمال به ترتیب عبارتند از

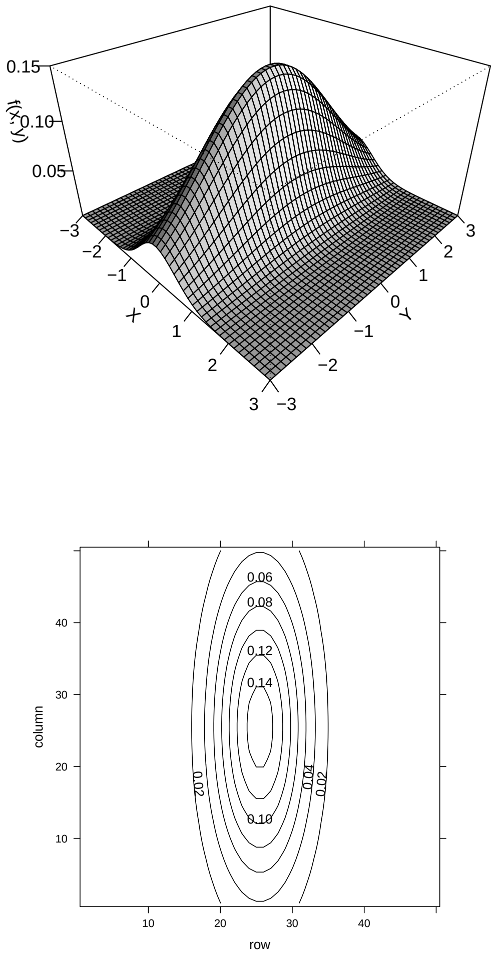
$$g(u) = e^{-ru^s}, \quad r, s > 0,$$
$$f_X(x) = \frac{s\Gamma\left(\frac{n}{2}\right)|\Sigma|^{-1/2} r^{\frac{n}{2s}}}{(2\pi)^{\frac{n}{2}}\Gamma\left(\frac{n}{2s}\right)} exp\Big\{-\frac{r}{2}\Big[-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\Big]^s\Big\}.$$

نمودارهای تابع چگالی و منحنی‌های تراز به ازای $r = s = 3$ در شکل ۶.۱ داده شده است.



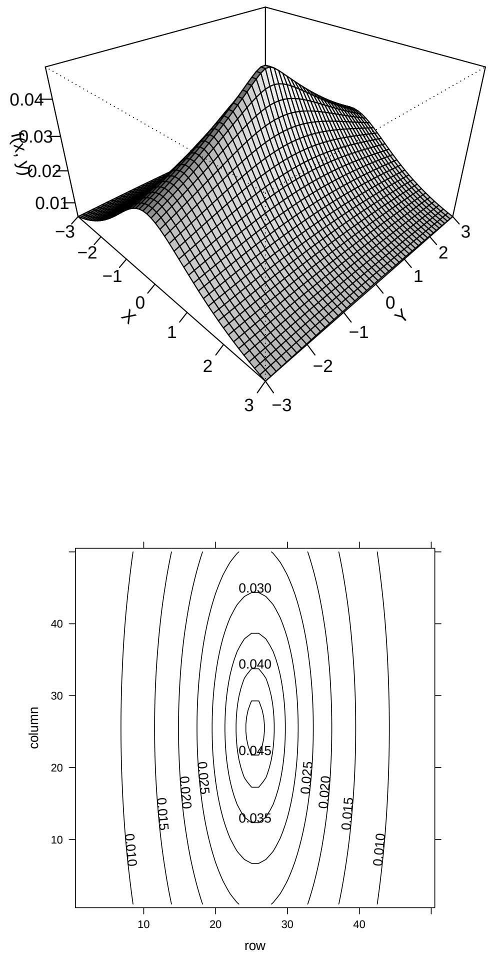
شکل ۶.۱: نمودارهای تابع چگالی و منحنی‌های تراز توزیع نمایی توانی دومتغیره

۶. توزیع کاتز چندمتغیره ($X \sim K_n(\mu, \Sigma, r)$): به ازای $s = ۱$، توزیع نمایی توانی به نوع کاتز تبدیل می‌شود. نمودارهای تابع چگالی و منحنی‌های تراز به ازای $r = ۳$ در شکل ۷.۱ داده شده است.



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

۷. توزیع لاپلاس (نمایی دوگانه) چندمتغیره ($(X \sim Lap_n(\mu, \Sigma)$): به ازای $s = ۱/۲$ و $r = \sqrt{۲}$، توزیع نمایی توانی به توزیع لاپلاس تبدیل می‌شود. نمودارهای تابع چگالی و منحنی‌های تراز در شکل ۸.۱ داده شده است.

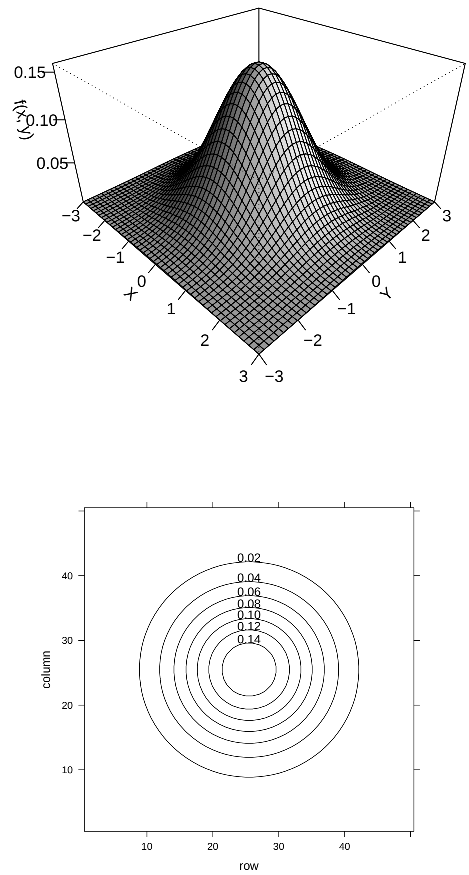


شکل ۸.۱: نمودارهای تابع چگالی و منحنی‌های تراز توزیع لاپلاس دومتغیره

تعریف ۱.۲.۳۷. توزیع کروی

بردار تصادفی p بعدی X دارای توزیع کروی است اگر X و HX به ازای هر ماتریس متعامد[۹۲] H با بعد p دارای توزیع یکسان باشند و به صورت $X \sim SS_n(\mu, \Sigma, \psi)$ نشان می‌دهیم.

در صورتی که تابع چگالی X موجود باشد، آن‌گاه واضح است که تابع چگالی آن تنها از طریق مقدار $X'X$ به X وابسته است. در این توزیع، منحنی‌های تراز به شکل دایره هستند. به عبارتی دقیق‌تر با فرض $\Sigma = I_p$ در توزیع‌های بیضی‌گون به توزیع کروی دست می‌یابیم. برای مثال نمودارهای تابع چگالی و منحنی‌های تراز توزیع نرمال استاندارد دومتغیره ($X \sim N_۲(0, I_۲)$) در شکل ۹.۱ نشان داده شده است.



شکل ۹.۱: نمودارهای تابع چگالی و منحنی‌های تراز توزیع نرمال استاندارد دومتغیره

[۹۲] Orthogonal matrix

تعریف ۳۸.۲.۱. توزیع متقارن کروی

متغیر تصادفی $X = (X_1, \ldots, X_p)$ دارای توزیع متقارن کروی حول پارامتر $\theta = (\theta_1, \ldots, \theta_p)$ است اگر دوران[۹۳] X حول θ، توزیع را تغییر ندهد. به عبارت دیگر

$$X - \theta \overset{d}{\sim} H(X - \theta),$$

که در آن H ماتریس متعامد با بعد p است و به ازای هر مقدار σ^2 مثبت به صورت $X \sim SS_p(\theta, \sigma^2 I_p)$ نشان داده می‌شود. در صورت وجود تابع چگالی، تابع چگالی این توزیع به صورت $g((X-\theta)^T(X-\theta))$ بوده که در آن $X \in \mathbb{R}^p$ و $g(\cdot) : [0, \infty) \to \mathbb{R}$.

تعریف ۳۹.۲.۱. توزیع کی - دو غیرمرکزی[۹۴]

با استفاده از تعریف ۲۳.۲.۱، فرض کنید متغیر تصادفی X دارای توزیع $N_p(\mu, I_p)$ باشد، در این صورت متغیر تصادفی $Z = X^T X$ دارای توزیع کی - دو غیرمرکزی با p درجه آزادی و پارامتر غیر مرکزی $\delta = \mu^T \mu$ است و با $X \sim \chi^2_{p,\delta}$ نمایش می‌دهند اگر تابع چگالی احتمال X به صورت زیر باشد.

$$f(z) = e^{-\frac{\delta}{2}} {}_0F_1\left(\frac{p}{2}; \frac{\delta z}{4}\right) \times \frac{e^{-\frac{z}{2}} z^{\frac{p}{2}-1}}{2^{\frac{p}{2}} \Gamma(\frac{p}{2})}.$$

قضیه ۴۰.۲.۱. (میرهد، ۱۹۸۲) *فرض کنید X دارای توزیع $N_p(\theta, \Sigma)$ باشد، که در آن Σ ماتریس $p \times p$ نامنفرد است، آن‌گاه*

$$(X-\theta)^T \Sigma^{-1} (X - \theta) \sim \chi^2_p.$$

نتیجه ۴۱.۲.۱. *فرض کنید X دارای توزیع $N_p(\theta, \sigma^2 I_p)$ باشد، که در آن σ^2 معلوم است، آن‌گاه*

$$\frac{(X-\theta)^T(X-\theta)}{\sigma^2} \sim \chi^2_p.$$

تعریف ۴۲.۲.۱. توزیع بتا[۹۵]

متغیر تصادفی X دارای توزیع بتا با پارامترهای p و k است و به صورت $B(p,k)$ نمایش داده

[۹۳] Rotation
[۹۴] Noncentral χ^2 distribution
[۹۵] Beta distribution

می‌شود، اگر تابع چگالی احتمال آن به صورت زیر باشد.

$$f(x) = \begin{cases} \frac{1}{\mathbf{B}(p,k)} x^{p-1}(1-x)^{k-1}, & 0 < x < 1, \\ 0, & \text{سایر مقادیر}. \end{cases}$$

که در آن $\mathbf{B}(p,k)$ تابع بتا است.

قضیه ۱.۲.۴۳. (پارسیان، ۱۳۸۶) *فرض کنید متغیر تصادفی X دارای توزیع χ^2_p و متغیر تصادفی Y دارای توزیع χ^2_k باشد. آنگاه رابطه زیر همواره برقرار است.*

$$\frac{X}{X+Y} \sim B\Big(\frac{p}{2}, \frac{k}{2}\Big).$$

لم ۱.۲.۴۴. *** *فرض کنید متغیر تصادفی X دارای توزیع $B(\frac{p}{2}, \frac{k}{2})$ باشد. در این صورت مقدار امید ریاضی $\frac{X}{1-X}$ برابر است با $\frac{p}{k-2}$.*

برهان. بنابه تعریف امید ریاضی داریم

$$\begin{aligned} E\left[\frac{X}{1-X}\right] &= \int_0^1 \frac{1}{\mathbf{B}(\frac{p}{2}, \frac{k}{2})} \frac{X}{1-X} X^{\frac{p}{2}-1}(1-X)^{\frac{k}{2}-1} dX \\ &= \frac{\mathbf{B}(\frac{p}{2}+1, \frac{k}{2}-1)}{\mathbf{B}(\frac{p}{2}, \frac{k}{2})} \int_0^1 \frac{1}{\mathbf{B}(\frac{p}{2}+1, \frac{k}{2}-1)} X^{\frac{p}{2}}(1-X)^{\frac{k}{2}-2} dX \\ &= \frac{\Gamma(\frac{p}{2}+1)\Gamma(\frac{k}{2}-1)/\Gamma(\frac{p}{2}+\frac{k}{2})}{\Gamma(\frac{p}{2})\Gamma(\frac{k}{2})/\Gamma(\frac{p}{2}+\frac{k}{2})} \\ &= \frac{(\frac{p}{2})!(\frac{k}{2}-2)!}{(\frac{p}{2}-1)!(\frac{k}{2}-1)!} \\ &= \frac{\frac{p}{2}}{\frac{k}{2}-1} = \frac{p}{k-2}. \end{aligned}$$

□

تعریف ۱.۲.۴۵. مخاطره بیز[۹۶]

امید ریاضی $\mathscr{R}(\theta,\delta)$ نسبت به تابع توزیع پیشین $\pi(\theta) \in \Pi$ که در آن Π کلاس تمام توابع توزیع پیشین است، را مخاطره بیز می‌نامند و با $r(\pi,\theta)$ نشان می‌دهند که برابر است با

$$\begin{aligned} r(\pi,\theta) &= E_\delta[\mathscr{R}(\theta,\delta)] \\ &= \begin{cases} \sum_{\theta\in\Theta} \mathscr{R}(\theta,\delta)\Pi(\theta) & \text{گسسته} \\ \int_\Theta \mathscr{R}(\theta,\delta)\pi(\theta)d\theta & \text{پیوسته} \end{cases} \end{aligned}$$

[۹۶] Bayes risk

تعریف ۴۶.۲.۱. برآوردگر بیز

برآوردگری که در کلاس تمام برآوردگرها مخاطره بیز را مینیمم کند، برآوردگر بیز نام دارد و با نماد δ^π نشان داده می‌شود. به عبارت دیگر داریم

$$r(\pi, \delta\pi) = \inf_{\delta \in D} r(\pi, \delta).$$

که در آن D کلاس تمام برآوردگرها است.

تعریف ۴۷.۲.۱. توزیع پیشین سره[۹۷]

توزیع پیشین $\pi(\theta) \in \Pi$ را سره گوییم اگر π یک تابع احتمال باشد، به عبارت دیگر داشته باشیم

$$\int_\Theta d\Pi(\theta) = \int_\Theta \pi(\theta) d\theta = ۱.$$

و ناسره[۹۸] گوییم اگر $\int_\Theta d\Pi(\theta) \neq ۱$ که در آن Π کلاس تمام توابع توزیع پیشین است.

قضیه ۴۸.۲.۱. (*لهمن و کسلا، ۱۹۹۸*) *فرض کنید $\{\pi_K(\theta), K \geq ۱\}$ دنباله‌ای از توزیع‌های پیشین سره روی Π و δ^{π_K} برآوردگر بیز متناظر با توزیع پیشین $\pi_K(\theta)$ با مخاطره بیز زیر باشد.*

$$r_K = r(\pi_K, \delta^{\pi_K}) = \int_\Theta \mathscr{R}(\theta, \delta^{\pi_K}) d\Pi_K(\theta).$$

اگر $\lim_{K\to\infty} r_K = r$ و برآوردگر $\delta_۰$ وجود داشته باشد به طوری که $\sup_\Theta \mathscr{R}(\theta, \delta_۰) \leq r$، آنگاه $\delta_۰$ یک برآوردگر مینیماکس است.

۳.۲.۱ لم‌های اساسی

لم ۴۹.۲.۱. ***لم استاین*** (*استاین، ۱۹۸۱*)

اگر بردار X دارای توزیع نرمال استاندارد باشد، آن‌گاه به ازای تمام توابع پیوسته $g : \mathbb{R} \to \mathbb{R}$ که $E|g'| < \infty$، داریم

$$E[g'(X)] = E[Xg(X)]. \qquad (۱.۱)$$

[۹۷] Proper prior distributions
[۹۸] Improper prior distributions

برهان. فرض کنید g تابعی مطلقاً پیوسته است که $E|g'(z)| < \infty$. اگر W دارای توزیع نرمال استاندارد باشد، آن‌گاه داریم

$$\begin{aligned} E[g'(W)] &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} g'(w) e^{-w^2/2} dw \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{0} g'(w) \left(\int_{-\infty}^{w} -x e^{-x^2/2} dx \right) dw \\ &+ \frac{1}{\sqrt{2\pi}} \int_{0}^{\infty} g'(w) \left(\int_{w}^{\infty} x e^{-x^2/2} dx \right) dw \end{aligned}$$

در این قسمت ترتیب انتگرال‌ها را عوض می‌کنیم.

$$\begin{aligned} E[g'(W)] &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{0} \left(\int_{x}^{0} g'(w) dw \right) (-x) e^{-x^2/2} dx \\ &+ \frac{1}{\sqrt{2\pi}} \int_{0}^{\infty} \left(\int_{0}^{x} g'(w) dw \right) x e^{-x^2/2} dx \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{0} \left(g(0) - g(x) \right) (-x) e^{-x^2/2} dx \\ &+ \frac{1}{\sqrt{2\pi}} \int_{0}^{\infty} \left(g(x) - g(0) \right) x e^{-x^2/2} dx \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \left(g(x) - g(0) \right) x e^{-x^2/2} dx \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x g(x) e^{-x^2/2} dx - \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x g(0) e^{-x^2/2} dx \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x g(x) e^{-x^2/2} dx - \frac{g(0)}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x e^{-x^2/2} dx \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} w g(w) e^{-w^2/2} dw - 0 \\ &= E[Wg(W)]. \end{aligned} \tag{2.1}$$

اثبات کامل است. □

لم ۱.۲.۰۵. ***تعمیم لم استاین*** *(استاین، ۱۹۸۱)*

فرض کنید $X \sim N_p(\theta, \sigma^2 I_p)$ و $g : \mathbb{R}^p \to \mathbb{R}^p$ باشد، که در آن σ^2 معلوم و منظور از I_p ماتریس همانی p بعدی است، آن‌گاه

$$E[(X-\theta)^T g(X)] = \sigma^2 E[\nabla \cdot g(X)].$$

لم ۵۱.۲.۱. (فودینیر و استرادرمان، ۱۹۹۶) *فرض کنید (X,U)، یک بردار تصادفی $p+k$ بعدی با توزیع متقارن کروی حول بردار $(\theta,0)$ با پارامتر θ مجهول باشد.که در آن بعد X برابر با بعد θ و مساوی با p و بعد U برابر با بعد بردار 0 و مساوی با k است. همچنین فرض کنید $h:\mathbb{R}\to\mathbb{R}$ و $g:\mathbb{R}^p\to\mathbb{R}^p$ توابعی اندازه‌پذیر باشند، در این صورت به ازای هر $\theta\in\mathbb{R}^p$، به شرط وجود امید ریاضی داریم*

$$E_\theta\left[h(U'U)(X-\theta)g(X)\right]=-E_\theta\left[\frac{H(U'U)}{(U'U)^{k/۲-۱}}\nabla\cdot g(X)\right]. \qquad (۳.۱)$$

که در آن

$$H(t)=\int -\frac{۱}{۲}h(t)t^{k/۲-۱}dt,$$

و همچنین ریشه آن صفر است.

نتیجه ۵۲.۲.۱. (فودینیر و همکاران، a۲۰۰۳) *به ازای هر تابع مشتق‌پذیر ضعیف $g:\mathbb{R}^p\to\mathbb{R}^p$، هر عدد صحیح q و به ازای هر $\theta\in\mathbb{R}^p$ به شرط وجود امید ریاضی، داریم*

$$E_\theta\left[(U^TU)^q(g(X))'(X-\theta)\right]=\frac{۱}{k+۲q}\left[(U^TU)^{q+۱}\nabla\cdot(g(X))\right].$$

قضیه ۵۳.۲.۱. (مدقالچی، ۱۳۸۴) *اگر تابع f بر (a,b) محدب (اکیداً محدب) باشد و $x_\circ\in(a,b)$، آن‌گاه $f'_+(x_\circ)$ و $f'_-(x_\circ)$ موجودند. به علاوه مشتقات چپ و راست بر (a,b) صعودی‌اند.*

نتیجه ۵۴.۲.۱. *تابع مشتق‌پذیر f بر (a,b) محدب است اگر و تنها اگر f' صعودی باشد و مقعر است اگر و فقط اگر f' نزولی باشد.*

لم ۵۵.۲.۱. (فودینیر و همکاران، a۲۰۰۳) *اگر r تابع نامنفی، مقعر و مشتق‌پذیر باشد، آن‌گاه r روی $\mathbb{R}_+$ نانزولی است و تابع $f(t)=\frac{r(t)}{t}$ روی $\mathbb{R}_+$ ناصعودی است. به علاوه اگر r دارای مشتق دوم باشد، آن‌گاه تابع $f(x)=\frac{r(\|x\|^۲)}{\|x\|^۲}$ برای $p\geq ۴$ زبرهمساز است.*

برهان. بنابه فرض r مشتق‌پذیر و مقعر است، در نتیجه با استفاده از نتیجه ۵۴.۲.۱، r' نزولی است. به عبارتی به ازای هر t و $t_\circ$، عضو $\mathbb{R}_+$ هرگاه $t_\circ\leq t$ آن‌گاه می‌توان نتیجه گرفت $r'(t)\leq r'(t_\circ)$.

از طرف دیگر بنابه تعریف مشتق داریم

$$r'(t) = \lim_{t \to t_\circ} \frac{r(t) - r(t_\circ)}{t - t_\circ} \leq r'(t_\circ),$$

در نتیجه می‌توان نوشت

$$r(t) - r(t_\circ) \leq r'(t_\circ)(t - t_\circ). \tag{۴.۱}$$

حال با استفاده از برهان خلف فرض می‌کنیم r ناصعودی باشد. اگر $t_\circ \geq \circ$ طوری وجود داشته باشد که $r'(t_\circ) < \circ$، بنابه فرض خلف با توجه به ناصعودی بودن r داریم $\lim_{t\to\infty} r(t) = -\infty$. که به تناقض می‌رسیم، چون بنابه فرض $r(t) \geq \circ$. بنابراین به ازای هر t، $r'(t) \geq \circ$ و در نتیجه r نانزولی است.

به طور مشابه به ازای هر $t > \circ$ و با استفاده از رابطه (۴.۱) با قرار دادن $t_\circ = t$ و $t = \circ$ داریم

$$r(\circ) - r(t) \leq r'(t)(\circ - t),$$

در نتیجه

$$tr'(t) - r(t) \leq -r(\circ) \leq \circ.$$

از طرف دیگر

$$t^{۲}\left(\frac{r(t)}{t}\right)' = tr'(t) - r(t) \leq \circ,$$

پس می‌توان نتیجه گرفت که تابع $f(t) = \frac{r(t)}{t}$ ناصعودی است.

سرانجام برای اثبات زبرهمسازی تابع $\frac{r(\|x\|^{۲})}{\|x\|^{۲}}$ با محاسبه مشتق دوم آن داریم

$$\begin{aligned}
\Delta\left[\frac{r(\|x\|^{۲})}{\|x\|^{۲}}\right] &= \frac{\partial}{\partial x_i}\sum_{i=۱}^{p}\Big[\frac{۲x_i r'(\|x\|^{۲})\|x\|^{۲} - ۲x_i r(\|x\|^{۲})}{\|x\|^{۴}}\Big] \\
&= ۲\frac{\partial}{\partial x_i}\sum_{i=۱}^{p}\Big[\frac{x_i r'(\|x\|^{۲})}{\|x\|^{۲}} - \frac{x_i r(\|x\|^{۲})}{\|x\|^{۴}}\Big] \\
&= ۲\sum_{i=۱}^{p}\Bigg\{\left[\frac{\Big(pr'(\|x\|^{۲}) + ۲x_i^{۲} r''(\|x\|^{۲})\Big)\|x\|^{۲} - ۲x_i^{۲} r'(\|x\|^{۲})}{\|x\|^{۴}}\right] \\
&\quad - \left[\frac{(pr(\|x\|^{۲}) + ۲x_i^{۲} r'(\|x\|^{۲}))\|x\|^{۴} - ۴x_i^{۲}\|x\|^{۲} r(\|x\|^{۲})}{\|x\|^{۸}}\right]\Bigg\}
\end{aligned}$$

$$
\begin{aligned}
&= 2\left[\frac{pr'(\|x\|^2)\|x\|^2 + 2\|x\|^4 r''(\|x\|^2) - 2\|x\|^2 r'(\|x\|^2)}{\|x\|^4}\right] \\
&- 2\left[\frac{pr(\|x\|^2)\|x\|^4 + 2\|x\|^6 r'(\|x\|^2) - 4r(\|x\|^2)\|x\|^4}{\|x\|^8}\right] \\
&= 2p\frac{r'(\|x\|^2)}{\|x\|^2} + 4r''(\|x\|^2) - 4\frac{r'(\|x\|^2)}{\|x\|^2} - 2p\frac{r(\|x\|^2)}{\|x\|^4} \\
&- 4\frac{r'(\|x\|^2)\|x\|^4}{\|x\|^2} + 8\frac{r(\|x\|^2)}{\|x\|^4} \\
&= \frac{2}{\|x\|^4}\Big[p\|x\|^2 r'(\|x\|^2) + 2\|x\|^4 r''(\|x\|^2) - 2\|x\|^2 r'(\|x\|^2) \\
&- pr(\|x\|^2) - 2\|x\|^2 r'(\|x\|^2) + 4r(\|x\|^2)\Big] \\
&= \frac{2}{\|x\|^4}\Big[2\|x\|^4 r''(\|x\|^2) + (p-4)(\|x\|^2 r'(\|x\|^2) - r(\|x\|)^2)\Big],
\end{aligned}
$$

که در آن $r''(\cdot)$ مشتق دوم $r(\cdot)$ است. حال با توجه به این‌که $r''(\cdot) \leq 0$ (بنابه فرض نانزولی و مقعر بودن r)، $tr'(t) - r(t) \leq 0$ و $p \geq 4$ داریم

$$
\Delta\left[\frac{r(\|x\|^2)}{\|x\|^2}\right] \leq 0.
$$

در نتیجه بنابه تعریف ۲۱.۲.۱ تابع فوق زبرهمساز است. □

لم ۵۶.۲.۱. *** *فرض کنید X یک متغیر تصادفی حقیقی مقدار با توزیع متقارن تک‌مدی برای $\theta \in \mathbb{R}_+$ و $\sigma^2 \geq 0$ باشد. اگر f تابع نامنفی روی $\mathbb{R}_+$ باشد، آن‌گاه*

$$
E_\theta\left[f(X^2)\frac{X^2}{\sigma^2} I_{[X<0]}\right] \leq \frac{1}{2} E_\theta\left[\frac{(X-\theta)^2}{\sigma^2} f(X^2)\right].
$$

برهان. چگالی X تنها در صورتی متقارن است که به شکل $g((X-\theta)^2)$ باشد. که در آن g یک تابع ناصعودی است. بنابراین می‌توان نتیجه گرفت

$$
\begin{aligned}
&E_\theta\left[\frac{f(X^2)}{\sigma^2}\Big\{X^2 I_{[X<0]} - \frac{1}{2}(X^2 - 2\theta X + \theta^2)\Big\}\right] \\
=&E_\theta\left[\frac{f(X^2)}{\sigma^2}\Big\{\Big(\frac{1}{2}X^2 + \theta X - \frac{1}{2}\theta^2\Big) I_{[X<0]} - \Big(\frac{1}{2}X^2 - \theta X + \frac{1}{2}\theta^2\Big) I_{[X\geq 0]}\Big\}\right] \\
=&E_\theta\left[\frac{f(X^2)}{\sigma^2}\Big(\frac{1}{2}X^2 + \theta X - \frac{1}{2}\theta^2\Big) I_{[X<0]} \Big/ |X|\right] \\
-&E_\theta\left[\frac{f(X^2)}{\sigma^2}\Big(\frac{1}{2}X^2 - \theta X + \frac{1}{2}\theta^2\Big) I_{[X\geq 0]} \Big/ |X|\right]
\end{aligned}
$$

$$=\int_{I_{[\frac{1}{2}X^2-\theta|X|-\frac{1}{2}\theta^2>0]}}\frac{f(X^2)}{\sigma^2}\Big(\frac{1}{2}X^2-\theta|X|-\frac{1}{2}\theta^2\Big)g((-|X|-\theta)^2)dx$$
$$-\int_{I_{[\frac{1}{2}X^2-\theta|X|-\frac{1}{2}\theta^2>0]}}\frac{f(X^2)}{\sigma^2}\Big(\frac{1}{2}X^2-\theta|X|+\frac{1}{2}\theta^2\Big)g(((|X|-\theta)^2)\Big]dx. \qquad (۵.۱)$$

از طرف دیگر همواره به ازای هر $\theta > 0$ داریم

$$(-|X|-\theta)^2 \geq (|X|-\theta)^2,$$

در نتیجه با توجه به آن که $g(\cdot)$ ناصعودی است، می‌توان نوشت

$$g((-|X|-\theta)^2) \leq g((|X|-\theta)^2),$$

و از اینرو کران بالای (۵.۱) به صورت زیر است.

$$\int_{I_{[\frac{1}{2}X^2-\theta|X|-\frac{1}{2}\theta^2>0]}}\frac{f(X^2)}{\sigma^2}\Big(\frac{1}{2}X^2-\theta|X|-\frac{1}{2}\theta^2\Big)g(((|X|-\theta)^2)dx$$
$$-\int_{I_{[\frac{1}{2}X^2-\theta|X|-\frac{1}{2}\theta^2>0]}}\frac{f(X^2)}{\sigma^2}\Big(\frac{1}{2}X^2-\theta|X|+\frac{1}{2}\theta^2\Big)g(((|X|-\theta)^2)dx$$
$$=\int_{I_{[\frac{1}{2}X^2-\theta|X|-\frac{1}{2}\theta^2>0]}}\frac{f(X^2)}{\sigma^2}\Big(\frac{1}{2}X^2-\theta|X|-\frac{1}{2}\theta^2-\frac{1}{2}X^2+\theta|X|-\frac{1}{2}\theta^2\Big)g(((|X|-\theta)^2)dx$$
$$=\int_{I_{[\frac{1}{2}X^2-\theta|X|-\frac{1}{2}\theta^2>0]}}\frac{f(X^2)}{\sigma^2}(-\theta^2)g(((|X|-\theta)^2)dx. \qquad (۶.۱)$$

در انتگرال (۶.۱) با توجه به این‌که بنابه فرض $f(\cdot)$ یک تابع نامنفی، $g(\cdot)$ تابعی ناصعودی و کران انتگرال مثبت است، به علت وجود علامت منفی در ضریب θ^2، داریم

$$\int_{I_{[\frac{1}{2}X^2-\theta|X|-\frac{1}{2}\theta^2>0]}}\frac{f(X^2)}{\sigma^2}(-\theta^2)g(|X|-\theta)^2)dx$$
$$=E_\theta\Big[\frac{f(X^2)}{\sigma^2}I_{[\frac{1}{2}X^2-\theta|X|-\frac{1}{2}\theta^2>0]}(-\theta^2)\Big]\leq 0.$$

اثبات کامل است. □

نکته ۵۷.۲.۱. (فودینیر و همکاران، ۲۰۰۳a) دقت کنید که ضریب $\frac{1}{2}$ در سمت راست رابطه موجود در لم ۵۶.۲.۱ کوچکترین ثابت مثبت موجود است و تساوی برای حالت $\theta = 0$ برقرار است.

لم ۱.۲.۵۸. (لهمن و کسلا، ۱۹۹۸) *فرض کنید Y یک متغیر تصادفی باشد و $g(y)$ و $h(y)$ دو تابع حقیقی روی مقادیر Y باشند طوری که $E[g(Y)]$، $E[h(Y)]$ و $E[g(Y)h(Y)]$ موجود باشند. آن‌گاه*

۱. *اگر یکی از توابع $h(\cdot)$ یا $g(\cdot)$ ناصعودی و دیگری نانزولی باشد، داریم*

$$E[g(Y)h(Y)] \leq E[g(Y)]E[h(Y)].$$

۲. *اگر هر دو تابع ناصعودی یا نانزولی باشند، داریم*

$$E[g(Y)h(Y)] \geq E[g(Y)]E[h(Y)].$$

فصل ۲
برآورد انقباضی پارامتر مکان با مقیاس معلوم

۱.۲ مقدمه

در این فصل، با استفاده از برآوردگرهای انقباضی سعی در برآورد بردار پارامتر در حالتی که این بردار دارای محدودیت است، داریم. فرض می‌کنیم توزیع جامعه مورد نظر متقارن کروی باشد. مسئله برآورد پارامتر محدودشده با استفاده از برآوردگرهای انقباضی را در سه حالت ۱. محدودیت کامل، ۲. محدودیت جزئی و ۳. حالتی که پارامتر θ محدود به مخروط‌های تودرتو و ماتریس‌های متعامد است، مورد بررسی قرار می‌دهیم.

در حالت اول، فرض می‌کنیم تمام مولفه‌های بردار پارامتر نامنفی باشند. به عبارت دیگر $\theta_i \geq 0$، $i = 1, \ldots, p$ و در حالت دوم تنها یک زیرمجموعه نامنفی از بردار پارامتر θ را در نظر می‌گیریم. با این دو پیش‌فرض، یک کلاس از برآوردگرهای انقباضی به صورت

$$\delta(X, U) = \delta_0(X) + g(X) U^T U,$$

را خواهیم یافت که تحت تابع زیان مربعی بهتر از برآوردگر طبیعی $\delta_0(X)$ است. در حالت سوم، مسئله برآورد وقتی که پارامتر θ متعلق به مخروط $\mathscr{C}$ از $\mathbb{R}^p$ است را مورد بررسی قرار داده و در بخش پایانی در توزیع نرمال p - متغیره با پارامتر مکان θ و پارامتر مقیاس $\sigma^2 I_p$ با σ^2 معلوم، تحت دو محدودیت کامل و جزئی، برآوردگر برتر را خواهیم یافت.

نتایج بخش‌های ۲.۲ تا ۴.۲ از منبع فودینیر (۲۰۰۳a) می‌باشد.

۲.۲ برآورد انقباضی در حالت با محدودیت کامل

فرض کنید (X, U) یک بردار تصادفی $p+k$ بعدی با توزیع متقارن کروی حول بردار $(\theta, 0)$ باشد که در آن $X = (X_1, \ldots, X_p)$ و $U = (U_1, \ldots, U_k)$ است و پارامتر θ مجهول است. واضح است که بعد X برابر با بعد $\theta = (\theta_1, \ldots, \theta_p)$ و مساوی با p و بعد U برابر با بعد بردار 0 و مساوی با k است.

علاقه‌مندیم پارامتر مجهول θ را تحت تابع زیان مربعی برآورد کنیم. در این بخش، منظور از محدودیت کامل، نامنفی بودن مؤلفه θ_i $(i = 1, \ldots, p)$ در بردار پارامتر θ است. به عبارت دیگر $(\theta_1 \geq 0, \ldots, \theta_p \geq 0)$. واضح است که برآوردگر طبیعی θ که آن را با $\delta_0(X)$ نشان می‌دهیم، عبارت است از

$$\delta_0(X) = (\delta_{01}, \ldots, \delta_{0p}), \qquad (۱.۲)$$

که در آن

$$\delta_{0i}(X) = \max(X_i, 0), \quad i = 1, \ldots, p. \qquad (۲.۲)$$

چون اثبات مینیماکس بودن $\delta_0(\cdot)$ ساده به نظر نمی‌رسد، یک اثبات کلی برای حالت بیان شده در قضیه زیر، ارائه می‌دهیم.

قضیه ۱.۲.۲. (فودینیر و همکاران، $2003a$) *فرض کنید (X, U) دارای توزیع متقارن کروی با واریانس معلوم $\sigma^2 < \infty$ باشد. همچنین فرض کنید $\mathscr{C}$ یک مخروط محدب بسته با درون ناتهی باشد و $\delta_0(X, U) = P_{\mathscr{C}}(X)$ (δ_0 تصویر بردار X روی مخروط $\mathscr{C}$ است). آن‌گاه δ_0 تحت تابع زیان مربعی برای θ مینیماکس است.*

برهان. چون δ_0 تصویر بردار X بر روی مخروط $\mathscr{C}$ است، وقتی که تابع زیان مربعی است، تصویر تبدیل به طول می‌گردد و در نتیجه طول این تصویر از طول X کمتر خواهد بود. به عبارتی داریم

$$\forall \ \epsilon > 0 \quad : \sup_{\theta \in \mathscr{C}} \mathscr{R}(\theta, \delta_0) - \epsilon \leq E_\theta[\|X - \theta\|^2] - \epsilon.$$

اکنون فرض کنیم $\delta_\circ$ مینیماکس نباشد، آن‌گاه δ وجود دارد به طوری که

$$\sup_{\theta\in\mathscr{C}} \mathscr{R}(\theta,\delta) \leq \sup_{\theta\in\mathscr{C}} \mathscr{R}(\theta,\delta_\circ) - \epsilon \leq E_\theta[\|X-\theta\|^۲] - \epsilon = R_\circ - \epsilon. \qquad (۳.۲)$$

از طرفی بنابه فرض، $\mathscr{C}$ حداقل دارای یک نقطه درونی نظیر $\theta_\circ$ است. بنابه تعریف نقطه درونی، عدد مثبت $\rho > \circ$ وجود دارد به طوری که $B_{\theta_\circ,\rho} \subset \mathscr{C}$ که در آن $B_{\theta_\circ,\rho}$ یک گوی باز به مرکز ρ و شعاع $\theta_\circ$ است.

ابتدا ادعا می‌کنیم به ازای هر $n > \circ$، $B_{n\theta_\circ,n\rho} \subset \mathscr{C}$. کافی است نشان دهیم به ازای هر $x \in B_{n\theta_\circ,n\rho}$، داریم $x \in \mathscr{C}$. فرض کنیم x عضو دلخواهی در $B_{n\theta_\circ,n\rho}$ باشد. بنابراین $\|x-n\theta_\circ\| < n\rho$، در نتیجه $\|\frac{x}{n}-\theta_\circ\| < \rho$، لذا

$$\frac{x}{n} \in B_{\theta_\circ,\rho} \subset \mathscr{C}.$$

چون $\mathscr{C}$ یک مخروط محدب است، کافی است در تعریف ۶.۲.۱ قرار دهیم $\theta_۱ = n$، $\theta_۲ = \circ$ و $x_۱ = \frac{x}{n}$. بنابراین $\frac{x}{n} \times n + \circ \times x_۲ \in \mathscr{C}$. یعنی $x \in \mathscr{C}$. که در نتیجه

$$B_{n\theta_\circ,n\rho} \subset \mathscr{C}.$$

در ادامه فرض کنید $\delta_n(X) = \delta(X+n\theta_\circ) - n\theta_\circ$، در نتیجه داریم

$$\begin{aligned}
\mathscr{R}(\theta,\delta_n) &= E(\|\delta_n - \theta\|^۲)\\
&= E(\|\delta(X+n\theta_\circ) - n\theta_\circ - \theta\|^۲)\\
&= E(\|\delta(X+n\theta_\circ) - (n\theta_\circ + \theta)\|^۲)\\
&= \mathscr{R}(n\theta_\circ + \theta, \delta^*). \qquad (۴.۲)
\end{aligned}$$

از آن جایی که $B_{n\theta_\circ,n\rho} \subset \mathscr{C}$ ، با استفاده از نامساوی (۳.۲) داریم

$$\sup_{\theta\in B_{\circ,n\rho}} \mathscr{R}(\theta,\delta_n) = \sup_{\theta\in B_{n\theta_\circ,n\rho}} \mathscr{R}(\theta,\delta) \leq \sup_{\theta\in\mathscr{C}} \mathscr{R}(\theta,\delta) \leq R_\circ - \epsilon.$$

همچنین مخاطره بیز r_n، متناظر با توزیع پیشین یکنواخت π_n روی $B_{\circ,n\rho}$، همگرا به $R_\circ$ است (فودینیر و همکاران، ۲۰۰۳a). از طرف دیگر وقتی که $\theta_\circ = \circ$، بنابه رابطه (۴.۲)، داریم

و لذا $\mathscr{R}(\theta,\delta_n)=\mathscr{R}(\theta,\delta)$

$$\begin{aligned} r_n = E(\mathscr{R}(\theta,\delta)) &= \int_{B_{\circ,n\rho}} \mathscr{R}(\theta,\delta)\pi(\theta)d\theta \\ &= \int_{B_{\circ,n\rho}} \mathscr{R}(\theta,\delta_n)\pi(\theta)d\theta \\ &\leq \underset{\theta\in B_{\circ,n\rho}}{sup}\ \mathscr{R}(\theta,\delta_n). \end{aligned} \tag{۵.۲}$$

از اینرو با توجه به تعریف حد، به ازای هر $\epsilon' \geq \circ$، $n_\circ$ وجود دارد به طوری که $r_n \geq R_\circ - \epsilon'$. اما بنابه نامساوی‌های (۳.۲) و (۵.۲) داریم

$$R_\circ - \epsilon \geq \underset{\theta\in\mathscr{C}}{sup}\ \mathscr{R}(\theta,\delta_\circ) - \epsilon \geq \underset{\theta\in\mathscr{C}}{sup}\ \mathscr{R}(\theta,\delta) \geq \underset{\theta\in B_{\circ,n\rho}}{sup}\ \mathscr{R}(\theta,\delta_n) \geq r_n \geq R_\circ - \epsilon'.$$

که این یک تناقض است، چون ϵ ثابت بود و ϵ' می‌تواند هر مقدار مثبت را داشته باشد. لذا $\delta_\circ$ یک برآوردگر مینیماکس است. به عبارت دیگر هیچ برآوردگر δ دیگری وجود ندارد که دارای ماکزیمم مخاطره کمتری نسبت به آن باشد. □

لازم به ذکر است که اگر $\sigma^۲$ ثابت نباشد با استدلالی مشابه می‌توان مینیماکس بودن را برای تابع زیان $\frac{\|\delta-\theta\|^۲}{\sigma^۲}$ ثابت کرد.

برآوردگر $\delta_\circ(X)$ در (۱.۲) را به صورت زیر نیز می‌توان نشان داد.

$$\delta_\circ(X) = X + \gamma(X). \tag{۶.۲}$$

که در آن $\gamma(X) = (\gamma_۱(X),\dots,\gamma_p(X))$ و هر مؤلفه $\gamma(X)$ به صورت زیر تعریف می‌شود.

$$\gamma_i(X) = \begin{cases} -X_i, & X_i < \circ, \\ \circ, & X_i \geq \circ. \end{cases}$$

حال رده‌ای از برآوردگرهای انقباضی به صورت زیر را در نظر بگیرید.

$$\begin{aligned} \delta(X,U) &= \delta_\circ(X) + g(X)U^TU \\ &= X + \gamma(X) + g(X)U^TU. \end{aligned} \tag{۷.۲}$$

که در آن $g:\mathbb{R}^p \to \mathbb{R}^p$ تابعی حقیقی با خواصی می‌باشد که در ادامه به آن اشاره می‌کنیم. همان طور که دیده می‌شود، این کلاس از برآوردگرها شامل بردار باقیمانده U است.

ابتدا دقت کنید برای این‌که مسئله برآورد منطقی به نظر برسد، نیاز داریم که مخاطره $\mathscr{R}(\theta,\delta_\circ)$ متناهی باشد. چراکه تا مادامی می‌توانیم درباره کوچکتر شدن یک مخاطره به طور نسبی صحبت کنیم که مخاطره‌ها متناهی باشند. حال می‌توان نشان داد که به محض این‌که تابع مخاطره برآوردگر طبیعی، متناهی باشد یا به طور معادل $E_\theta[\|X\|^{۲}] \leq \infty$، مخاطره متناهی است. زیرا

$$\begin{aligned}\mathscr{R}(\theta,\delta_\circ) &= E(\|\delta_\circ-\theta\|^{۲}) = E(\|X+\gamma(X)-\theta\|^{۲})\\ &= E(\|X\|^{۲}) + E(\|\gamma(X)-\theta\|^{۲}) + ۲E(X^T(\gamma(X)-\theta)).\end{aligned} \qquad (۸.۲)$$

در عبارت (۸.۲) دو امید ریاضی $E(\|\gamma(X)-\theta\|^{۲})$ و $E(X^T(\gamma(X)-\theta))$ متناهی می‌باشند و برای متناهی بودن مخاطره باید $E(\|X\|^{۲})$ نیز متناهی باشد. به عبارت دیگر $E_\theta[\|X\|^{۲}] \leq \infty$. تحت فرضیه اخیر به راحتی می‌توان بررسی کرد که مخاطره $\mathscr{R}(\theta,\delta)$ متناهی است، اگر و فقط اگر

$$\begin{aligned}\mathscr{R}(\theta,\delta) &= E(\|\delta-\theta\|^{۲}) = E(\|X+\gamma(X)+g(X)U^TU-\theta\|^{۲})\\ &= E(\|X+\gamma(X)-\theta\|^{۲}) + E(\|g(X)\|^{۲}(U^TU)^{۲})\\ &\quad + ۲E(U^TU(g(X))^T(X+\gamma(X)-\theta)),\end{aligned} \qquad (۹.۲)$$

متناهی باشد. در عبارت (۹.۲) امید ریاضی $E(\|X+\gamma(X)-\theta\|^{۲})$ مشابه با استدلال رابطه (۸.۲) متناهی است، اگر $E_\theta[\|X\|^{۲}] \leq \infty$ و با استفاده از نامساوی کوشی - شوارتز برای متناهی بودن امید ریاضی $E(U^TU(g(X))^T(X+\gamma(X)-\theta))$ داریم

$$E(U^TU(g(X))^T(X+\gamma(X)-\theta))^{۲} \leq E(\|g(X)\|^{۲}(U^TU)^{۲})E(\|X+\gamma(X)-\theta\|^{۲}). \qquad (۱۰.۲)$$

عبارت (۱۰.۲) وقتی متناهی است که $E(X+\gamma(X)-\theta)^{۲}$ و $E_\theta[\|g\|^{۲}(U^TU)^{۲}]$ متناهی باشند. برای اثبات متناهی بودن $E(X+\gamma(X)-\theta)^{۲}$ مجدداً بر اساس نامساوی کوشی - شوارتز داریم

$$E(X+\gamma(X)-\theta)^{۲} \leq E(\|X\|^{۲})E(\|\gamma(X)-\theta\|^{۲}).$$

که متناهی است اگر $E(\|X\|^{۲}) \leq \infty$ و این مهم در قسمت قبل بیان شد.

لم ۲.۲.۲. (میرهد، ۱۹۸۲) *فرض کنید (Z,U) دارای توزیع متقارن کروی حول بردار 0 باشد، که در آن بعد Z برابر با p و بعد U برابر با k است، در این صورت توزیع $Z^TZ/(Z^TZ+U^TU)$ برابر*

با $B(\frac{p}{۲},\frac{k}{۲})$ *و مستقل از* Z^TZ+U^TU *است.*

برهان. این قضیه باید به ازای هر توزیع دلخواه از خانواده توزیع‌های متقارن کروی درست باشد. لذا برای توزیع نرمال نیز باید درست باشد. فرض کنید $Z=HX$ که در آن H یک ماتریس متعامد $p+k$ بعدی است. همچنین فرض کنید $X\sim N_p(\circ,I_p)$، در این صورت

$$Z^TZ=X^TH^THX=X^TX\sim\chi^۲_p.$$

مجدداً فرض کنید $U=HX$ و $U\sim N_p(\circ,I_p)$ آن‌گاه با استدلال مشابه بالا داریم

$$U^TU\sim\chi^۲_p.$$

در مرحله بعد با تشکیل نسبت $Z^TZ/(Z^TZ+U^TU)$ بنابه قضیه ۴۳.۲.۱ داریم

$$\frac{Z^TZ}{Z^TZ+U^TU}=\frac{\chi^۲_p}{\chi^۲_p+\chi^۲_k}\sim B\Big(\frac{p}{۲},\frac{k}{۲}\Big).$$

□

لم ۳.۲.۲. (دوپلسیس[۱]، ۱۹۷۰) *اگر g یک تابع زبرهمساز روی $\mathbb{R}^p$ و $Z\in\mathbb{R}^p$ یک متغیر تصادفی با توزیع یکنواخت روی کره به مرکز مبدا و شعاع τ باشد، آن‌گاه به ازای هر $\theta\in\mathbb{R}^p$ امید ریاضی $E[g(\theta+Z)]$ تابعی ناصعودی از τ است.*

قضیه زیر یکی از کاربردهای نتیجه ۵۲.۲.۱ است. پیش از آن که به این قضیه بپردازیم، ابتدا برای تسهیل در روند اثبات قضیه، تفاضل مخاطره، $\Delta\mathscr{R}$، را در زیر محاسبه می‌نماییم.

$$\begin{aligned}
\Delta\mathscr{R}&=\mathscr{R}(\theta,\delta)-\mathscr{R}(\theta,\delta_\circ)\\
&=E_\theta[\|\delta-\theta\|^۲]-E_\theta[\|\delta_\circ-\theta\|^۲]\\
&=E_\theta[\|\delta-\theta\|^۲-\|\delta_\circ-\theta\|^۲]\\
&=E_\theta[\|X+\gamma(X)+g(X)U^TU-\theta\|^۲-\|X+\gamma(X)-\theta\|^۲]\\
&=E_\theta[\|X+\gamma(X)-\theta\|^۲+\|g(X)U^TU\|^۲\\
&+۲U^TU(g(X))^T(X+\gamma(X)-\theta)-\|X+\gamma(X)-\theta\|^۲]
\end{aligned}$$

[۱] Du Plessis

$$= E_\theta[\|g(X)\|^{۲}(U^TU)^{۲} + ۲U^TU(g(X))^T(X-\theta)$$
$$+ ۲U^TU(g(X))^T\gamma(X)]. \qquad (۱۱.۲)$$

همانطور که در رابطه (۱۱.۲) دیده می‌شود، دومین جمله از آخرین امید ریاضی به θ وابسته است. برای از بین بردن نقش θ، از لم ۵۱.۲.۱ استفاده خواهیم کرد. در این قسمت فرض می‌کنیم در نتیجه ۵۲.۲.۱، $q = ۱$ باشد، در این صورت اگر g مشتق‌پذیر ضعیف باشد، داریم

$$E_\theta[(U^TU)(g(X))^T(X-\theta)] = \frac{۱}{k+۲}[(U^TU)^{۲}\nabla \cdot (g(X))]. \qquad (۱۲.۲)$$

بنابراین تفاضل مخاطره (۱۱.۲) با استفاده از (۱۲.۲) برابر است با

$$\Delta\mathscr{R} = E_\theta[\|g(X)\|^{۲}(U^TU)^{۲} + ۲U^TU(g(X))^T(X-\theta) + ۲U^TU(g(X))^T\gamma(X)]$$
$$= E_\theta[\|g(X)\|^{۲}(U^TU)^{۲} + \frac{۲(U^TU)^{۲}}{k+۲}\nabla \cdot g(X) + ۲U^TU(g(X))^T\gamma(X)]. \qquad (۱۳.۲)$$

قضیه ۴.۲.۲. *(فودینیر و همکاران، ۲۰۰۳) تحت مفروضات این بخش، فرض کنید که $p > ۴$، $k > ۲ + \frac{۴p}{p-۴}$ و توزیع نیز تک‌مدی باشد. برآوردگر انقباضی $\delta(X,U) = X + \gamma(X) + g(X)U^TU$ را در نظر بگیرید که در آن تابع g به صورت $g(X) = -\frac{c\, r(\|X\|^{۲})}{\|X\|^{۲}}X$ است. اگر تابع r و ثابت c در سه شرط زیر صدق کنند.*

۱. $\circ \leq r(\|X\|^{۲}) \leq ۱$،

۲. $r(\|X\|^{۲})$ *مقعر و دارای مشتق دوم باشد،*

۳. $\circ < c \leq ۲(\frac{p-۲}{k+۲}) - \frac{p}{k-۲}$،

آن‌گاه به ازای هر $\theta \in \mathbb{R}^p$، برآوردگر $\delta(X,U)$ بر برآوردگر $\delta_\circ(X) = X + \gamma(X)$ برتری دارد.

برهان. با استفاده از (۱۳.۲) و لم ۵۵.۲.۱، از آن‌جایی که r نامنفی، مقعر و مشتق‌پذیر است،

$r'(\cdot) \geq ۰$، تفاضل مخاطره $\Delta\mathscr{R}$ برابر است با

$$\begin{aligned}
\Delta\mathscr{R} &= E_\theta\Big[\|g(X)\|^۲(U^TU)^۲ + \frac{۲(U^TU)^۲}{k+۲}\nabla\cdot g(X) + ۲U^TU(g(X))^T\gamma(X)\Big] \\
&= E_\theta\left[\|-\frac{c\,r(\|X\|^۲)}{\|X\|^۲}X\|^۲(U^TU)^۲ + ۲(U^TU)\left(-\frac{c\,r(\|X\|^۲)}{\|X\|^۲}X\right)^T\gamma(X)\right. \\
&\left.+ ۲\frac{(U^TU)^۲}{k+۲}\nabla\cdot\left(-\frac{c\,r(\|X\|^۲)}{\|X\|^۲}X\right)\right] \\
&= E_\theta\left[(U^TU)^۲c^۲\frac{r^۲(\|X\|^۲)}{\|X\|^۴}\|X\|^۲ + ۲c\frac{(U^TU)r(\|X\|^۲)}{\|X\|^۲}\sum_{i=۱}^{p}X_i^۲I_{[X_i\leq ۰]}\right. \\
&\left.- ۲c\frac{(U^TU)^۲}{k+۲}\left\{\sum_{i=۱}^{p}\left(\frac{(۲X_i^۲r'(\|X\|^۲)+p\,r(\|X\|^۲))\|X\|^۲-۲X_i^۲r(\|X\|^۲)}{\|X\|^۴}\right)\right\}\right] \\
&= E_\theta\left[(U^TU)^۲c^۲\frac{r^۲(\|X\|^۲)}{\|X\|^۲} + ۲c\frac{(U^TU)r(\|X\|^۲)}{\|X\|^۲}\sum_{i=۱}^{p}X_i^۲I_{[X_i\leq ۰]}\right. \\
&\left.- ۲c\frac{(U^TU)^۲}{k+۲}\left(\frac{r'(\|X\|^۲)\|X\|^۴+pr(\|X\|^۲)\|X\|^۲-۲r(\|X\|^۲)\|X\|^۲}{\|X\|^۴}\right)\right] \\
&= E_\theta\left[(U^TU)^۲c^۲\frac{r^۲(\|X\|^۲)}{\|X\|^۲} + ۲c\frac{(U^TU)r(\|X\|^۲)}{\|X\|^۲}\sum_{i=۱}^{p}X_i^۲I_{[X_i\leq ۰]}\right. \\
&\left.- ۲c\frac{(U^TU)^۲}{k+۲}\left((p-۲)\frac{r(\|X\|^۲)}{\|X\|^۲}+۲r'(\|X\|^۲)\right)\right] \\
&= E_\theta\left[(U^TU)^۲c^۲\frac{r^۲(\|X\|^۲)}{\|X\|^۲} + ۲c\frac{(U^TU)r(\|X\|^۲)}{\|X\|^۲}\sum_{i=۱}^{p}X_i^۲I_{[X_i\leq ۰]}\right. \\
&\left.- ۲c\frac{(p-۲)(U^TU)^۲r(\|X\|^۲)}{(k+۲)\|X\|^۲} - ۴c\frac{(U^TU)^۲r'(\|X\|^۲)}{k+۲}\right] \\
&= E_\theta\left[(U^TU)^۲\frac{r(\|X\|^۲)}{\|X\|^۲}c\left(cr(\|X\|^۲) - ۲\frac{p-۲}{k+۲} + ۲\frac{\sum_{i=۱}^{p}X_i^۲I_{[X_i\leq ۰]}}{U^TU}\right)\right. \\
&\left.- ۴c\frac{(U^TU)^۲r'(\|X\|^۲)}{k+۲}\right] \\
&\leq E_\theta\left[(U^TU)^۲\frac{r(\|X\|^۲)}{\|X\|^۲}c\left(cr(\|X\|^۲) - ۲\frac{p-۲}{k+۲} + ۲\frac{\sum_{i=۱}^{p}X_i^۲I_{[X_i\leq ۰]}}{U^TU}\right)\right].
\end{aligned}$$

حال با استفاده از $۰ \leq r(\|X\|^۲) \leq ۱$ و با قرار دادن مقدار ماکزیمم $r(\|X\|^۲)$ که بنابه فرض برابر

با ۱ است، داریم

$$\begin{aligned}\Delta\mathscr{R} &\leq E_\theta\left[(U^TU)^{۲}\frac{r(\|X\|^{۲})}{\|X\|^{۲}}c\left(cr(\|X\|^{۲}) - ۲\frac{p-۲}{k+۲} + ۲\frac{\sum_{i=۱}^p X_i^{۲} I_{[X_i\leq ۰]}}{U^TU}\right)\right]\\ &\leq E_\theta\left[(U^TU)^{۲}\frac{r(\|X\|^{۲})}{\|X\|^{۲}}c\left(c - ۲\frac{p-۲}{k+۲} + ۲\frac{\sum_{i=۱}^p X_i^{۲} I_{[X_i\leq ۰]}}{U^TU}\right)\right]. \qquad (۱۴.۲)\end{aligned}$$

در آخرین امید ریاضی ابتدا برای i معین جمله

$$\left((U^TU)^{۲}\frac{r(\|X\|^{۲})}{\|X\|^{۲}}\right)\left(\frac{\sum_{i=۱}^p X_i^{۲} I_{[X_i\leq ۰]}}{U^TU}\right),$$

را در نظر می‌گیریم و درباره امید ریاضی شرطی روی (U^TU) و $(X_j-\theta_j)^{۲}$ برای $j\neq i$ بحث می‌کنیم. دقت کنید که با توجه به این‌که بنابه فرض r نامنفی، مقعر و مشتق‌پذیر است، بنابه لم ۵۵.۲.۱ ، $r(\|X\|^{۲})/\|X\|^{۲}$ در $\|X\|^{۲}$ ناصعودی است. آن‌گاه چون توزیع شرطی X_i بر روی U^TU و $(X_j-\theta_j)^{۲}$ برای $j\neq i$، برای θ_i متقارن و تک‌مدی است، بنابه لم ۵۶.۲.۱ در حالت $\sigma^{۲}=۱$، داریم

$$\begin{aligned}&cE\left[U^TU\frac{r(\|X\|^{۲})}{\|X\|^{۲}}X_i^{۲}I_{[X_i\leq ۰]}\Big/ U^TU=\rho^{۲}, (X_j-\theta_j)^{۲}=Z_j^{۲}, j\neq i\right]\\ &\leq \frac{c}{۲}E\left[U^TU\frac{r(\|X\|^{۲})}{\|X\|^{۲}}Z_i^{۲}\Big/ U^TU=\rho^{۲}, (X_j-\theta_j)^{۲}=Z_j^{۲}, j\neq i\right]. \qquad (۱۵.۲)\end{aligned}$$

بنابراین تفاضل مخاطره در (۱۴.۲) با فرض $Z=X-\theta$ و U، دارای کران بالای به صورت زیر است.

$$\begin{aligned}&E_\theta\left[(U^TU)^{۲}\frac{r(\|X\|^{۲})}{\|X\|^{۲}}c\left(c - ۲\frac{p-۲}{k+۲} + ۲\frac{\sum_{i=۱}^p X_i^{۲} I_{[X_i\leq ۰]}}{U^TU}\right)\right]\\ &\leq cE_\theta\left[(U^TU)^{۲}\frac{r(\|Z+\theta\|^{۲})}{\|Z+\theta\|^{۲}}\left(c - ۲\frac{p-۲}{k+۲} + \frac{Z^TZ}{U^TU}\right)\right]. \qquad (۱۶.۲)\end{aligned}$$

حال فرض می‌کنیم که $U^TU+Z^TZ=R^{۲}$ و یک کران بالا برای امید ریاضی شرطی (۱۶.۲) با $R^{۲}$ مزبور خواهیم یافت. مجدداً دقت کنید که چون r نامنفی، مقعر و دارای مشتق دوم است، با استفاده از لم ۵۵.۲.۱ واضح است که $r(\|X\|^{۲})/\|X\|^{۲}$ برای $p\geq ۴$ زبرهمساز و در $\|X\|^{۲}$

ناصعودی است. حال با شرطی کردن نامساوی (۱۶.۲) روی $R^۲$ و Z^TZ، داریم

$$\begin{aligned} E\left[(R^۲-Z^TZ)E\left[\frac{r(\|Z+\theta\|^۲)}{\|Z+\theta\|^۲}\Big/ Z^TZ,R^۲\right]\left(c-۲\frac{p-۲}{k+۲}+\frac{Z^TZ}{R^۲-Z^TZ}\right)\Big/ R^۲\right] \\ \leq cE\left[(R^۲-Z^TZ)E\left[\frac{r(\|Z+\theta\|^۲)}{\|Z+\theta\|^۲}\Big/ Z^TZ,R^۲\right]\Big/ R^۲\right] \\ \times E\left[\left(c-۲\frac{p-۲}{k+۲}+\frac{Z^TZ}{R^۲-Z^TZ}\right)\Big/ R^۲\right]. \end{aligned} \quad (۱۷.۲)$$

چون توزیع شرطی Z به شرط $R^۲$ متقارن کروی است، $E[r(\|Z+\theta\|^۲)/\|Z+\theta\|^۲]$ با استفاده از لم ۳.۲.۲ در Z^TZ برای $Z^TZ < R^۲$ ناصعودی است. همچنین $(R^۲-Z^TZ)^۲$ در Z^TZ ناصعودی و $Z^TZ/(R^۲-Z^TZ)$ نیز نانزولی است. بنابراین نامساوی (۱۷.۲) را می‌توان از لم ۵۸.۲.۱ نتیجه گرفت.

در سمت راست (۱۷.۲) اولین امید ریاضی شرطی نامنفی است. از اینرو تفاضل مخاطره شرطی ($\Delta\mathscr{R}$) به شرط آنکه دومین مقدار امید ریاضی شرطی در سمت راست (۱۷.۲) نامثبت باشد، نامثبت خواهد بود، اما چون با استفاده از لم ۲.۲.۲ توزیع $\beta = Z^TZ/R^۲$ به شرط $R^۲$، $B(\frac{p}{۲},\frac{k}{۲})$ است، با استفاده از لم ۴۴.۲.۱ داریم

$$E\left[\frac{Z^TZ}{R^۲-Z^TZ}\Big/ R^۲\right]=E\left[\frac{\beta}{۱-\beta}\right]=\frac{p}{k-۲}. \quad (۱۸.۲)$$

از اینرو تفاضل مخاطره شرطی تنها در صورتی نامثبت خواهد بود که داشته باشیم

$$c-۲\frac{p-۲}{k+۲}+E\left[\frac{Z^TZ}{R^۲-Z^TZ}\Big/ R^۲\right]\leq ۰$$

سپس

$$c-۲\frac{p-۲}{k+۲}-\frac{p}{k-۲}\leq ۰$$

آن‌گاه

$$c\leq ۲\frac{p-۲}{k+۲}-\frac{p}{k-۲},$$

و در نتیجه

$$۰<c\leq ۲\frac{p-۲}{k+۲}-\frac{p}{k-۲}.$$

که همان شرط سوم قضیه است.

برای این‌که کران بالای c مثبت باشد، داریم

$$۲\frac{p-۲}{k+۲}-\frac{p}{k-۲}>۰,$$

پس از مخرج مشترک گرفتن

$$\frac{۲(k-۲)(p-۲)-p(k+۲)}{(k+۲)(k-۲)}>۰,$$

سپس

$$k(p-۴)>۶p-۸,$$

بنابراین

$$k>\frac{۲(p-۴)+۴p}{p-۴},$$

در نتیجه

$$k>۲+\frac{۴p}{p-۴}.$$

اثبات کامل است. □

نکته ۵.۲.۲. به نظر می‌رسد که برای ادامه کار باید روی شرط $۰<c<۲\frac{p-۲}{k+۲}-\frac{p}{k-۲}$ برای ثابت انقباض متمرکز شویم. این امر در حالتی که $p>۴$ است، اهمیت بیشتری دارد. حال چون در صورت قضیه ۴.۲.۲ داشتیم $k>۲+\frac{۴p}{p-۴}$، با توجه به این‌که $p>۴$، همواره داریم $\frac{۴p}{p-۴}>۴$، پس لازم است که برای همه p ها $k\geq ۷$ باشد. برای $p=۵$، k باید حداقل ۲۳ و برای $p=۱۰$، k باید حداقل ۹ باشد.

۳.۲ برآورد انقباضی در حالت با محدودیت جزئی

در این بخش هدف دستیابی به کلاسی از برآوردگرهای برتر برای پارامتر مکان، در حالت با محدودیت جزئی، به این معنی که تنها زیرمجموعه‌ای از θ_i ها دارای محدودیت نامنفی بودن باشند، می‌باشد.

فرض کنید که s تا از θ_i ها نامنفی باشند. به عبارتی $\theta_۱\geq ۰,\theta_۲\geq ۰,\dots,\theta_s\geq ۰$ و

$\theta_{s+1}, \theta_{s+2}, \ldots, \theta_p$ هیچ محدودیتی نداشته باشند. آن‌گاه، مشابه با بخش دوم، برآوردگر طبیعی را به صورت

$$\delta_{\circ}^{s} = X + \gamma_s(X), \tag{۱۹.۲}$$

در نظر می‌گیریم. که در آن

$$\gamma_{s,j}(X) = \begin{cases} -X_j, & X_j < \circ, \\ \circ, & X_j \geq \circ. \end{cases} \tag{۲۰.۲}$$

و

$$\gamma_{s,j}(X) = \circ, \quad j > s.$$

همچنین در (۱۹.۲)، $\delta_{\circ}^{s}(X)$ در حالت نرمال برابر با MLE و در حالت متقارن کروی، مینیماکس است(فودینیر و همکاران، ۲۰۰۳a). در این صورت برآوردگر انقباضی را به صورت زیر در نظر می‌گیریم.

$$\delta^{s}(X, U) = X + \gamma_s(X) + U^T U g(X). \tag{۲۱.۲}$$

مشابه با قضیه ۴.۲.۲، قضیه زیر را در خصوص برتری برآوردگر انقباضی بر برآوردگر طبیعی بیان و اثبات می‌نماییم. مجدداً مانند بخش قبل پیش از آن که اثبات را آغاز نماییم، برای تسهیل در روند اثبات تفاضل مخاطره $\Delta\mathscr{R}$ را در زیر محاسبه می‌نماییم. محاسبات زیر با تفاوتی جزئی کاملاً مشابه محاسبات بخش قبل است.

$$\begin{aligned} \Delta\mathscr{R} &= \mathscr{R}(\theta, \delta^{s}) - \mathscr{R}(\theta, \delta_{\circ}^{s}) \\ &= E_\theta[\|\delta^{s} - \theta\|^{۲} - \|\delta_{\circ}^{s} - \theta\|^{۲}] \\ &= E_\theta[\|g(X)\|^{۲}(U^T U)^{۲} + ۲U^T U (g(X))^T (X - \theta) \\ &+ ۲U^T U (g(X))^T \gamma_s(X)]. \end{aligned} \tag{۲۲.۲}$$

قضیه ۱.۳.۲. (فودینیر و همکاران،۲۰۰۳a) *تحت مفروضات این بخش، فرض کنید* $p > ۴$، $p \geq s$، $k > ۲ + \frac{۴s}{۲p-s-۴}$ *و توزیع نیز تک‌مدی باشد. برآوردگر انقباضی به شکل*

$$\delta^{s}(X, U) = X + \gamma_s(X) + U^T U g(X)$$

را در نظر بگیرید که در آن تابع g به صورت $g(X) = -\frac{c\, r(\|X\|^۲)}{\|X\|^۲}X$ است. اگر تابع r و ثابت c در شرایط زیر صدق کنند.

۱. $۰ < r(\|X\|^۲) \leq ۱$،

۲. $r(\|X\|^۲)$ *دارای مشتق دوم و مقعر باشد،*

۳. $۰ < c \leq ۲\frac{p-۲}{k+۲} - \frac{s}{k-۲}$،

آن‌گاه به ازای $\theta_۱ \geq ۰, \theta_۲ \geq ۰, \ldots, \theta_s \geq ۰$، برآوردگر $\delta^s\,(X, U)$ بر برآوردگر $\delta^s\,(X, U) = X + \gamma_s(X)$ برتری دارد.

برهان. با استفاده از اثبات قضیه ۴.۲.۲، بدون آن که تغییر اساسی در آن ایجاد نماییم اثبات را آغاز می‌کنیم. مشابه با بخش قبل و با استفاده از نتیجه ۵۲.۲.۱ و لم ۵۵.۲.۱ از آن جایی که r نامنفی، مقعر و مشتق‌پذیر است، داریم $r'(\cdot) \geq ۰$، تفاضل مخاطره (۲۲.۲) برابر است با

$$\begin{aligned}
\Delta\mathscr{R} &= E_\theta\Big[\|g(X)\|^۲(U^TU)^۲ + \frac{۲(U^TU)^۲}{k+۲}\nabla\cdot g(X) + ۲U^TU(g(X))^T\gamma_s(X)\Big]\\
&= E_\theta\Bigg[\|-\frac{c\,r(\|X\|^۲)}{\|X\|^۲}X\|^۲(U^TU)^۲ + ۲(U^TU)\left(-\frac{c\,r(\|X\|^۲)}{\|X\|^۲}X\right)^T\gamma_s(X)\\
&+ ۲\frac{(U^TU)^۲}{k+۲}\nabla\cdot\left(-\frac{c\,r(\|X\|^۲)}{\|X\|^۲}X\right)\Bigg]\\
&= E_\theta\Bigg[(U^TU)^۲c^۲\frac{r^۲(\|X\|^۲)}{\|X\|^۲} + ۲c\frac{(U^TU)r(\|X\|^۲)}{\|X\|^۲}\sum_{i=۱}^{s}X_i^۲I_{[X_i\leq ۰]}\\
&- ۲c\frac{(p-۲)(U^TU)^۲r(\|X\|^۲)}{(k+۲)\|X\|^۲} - ۴c\frac{(U^TU)^۲r'(\|X\|^۲)}{k+۲}\Bigg]\\
&\leq E_\theta\Bigg[(U^TU)^۲\frac{r(\|X\|^۲)}{\|X\|^۲}c\left(cr(\|X\|^۲) - ۲\frac{p-۲}{k+۲} + ۲\frac{\sum_{i=۱}^{s}X_i^۲I_{[X_i\leq ۰]}}{U^TU}\right)\Bigg]. \quad (۲۳.۲)
\end{aligned}$$

حال با استفاده از $۰ \leq r(\|X\|^۲) \leq ۱$، با قرار دادن مقدار ماکزیمم $r(\|X\|^۲)$ که برابر با ۱ است

و مشابه نامساوی (۱۴.۲) برای رابطه (۲۳.۲) داریم

$$\begin{aligned} &E_\theta\left[(U^TU)^۲\frac{r(\|X\|^۲)}{\|X\|^۲}c\left(cr(\|X\|^۲)-۲\frac{p-۲}{k+۲}+۲\frac{\sum_{i=۱}^s X_i^۲ I_{[X_i\leq ۰]}}{U^TU}\right)\right]\\ &\leq E_\theta\left[(U^TU)^۲\frac{r(\|X\|^۲)}{\|X\|^۲}c\left(c-۲\frac{p-۲}{k+۲}+۲\frac{\sum_{i=۱}^s X_i^۲ I_{[X_i\leq ۰]}}{U^TU}\right)\right]. \end{aligned} \tag{۲۴.۲}$$

مشابه با قضیه ۴.۲.۲ در آخرین امید ریاضی ابتدا برای i معین که، $۰ < i < s$، جمله

$$\left((U^TU)^۲\frac{r(\|X\|^۲)}{\|X\|^۲}\right)\left(\frac{\sum_{i=۱}^s X_i^۲ I_{[X_i\leq ۰]}}{U^TU}\right),$$

را در نظر می‌گیریم و درباره امید ریاضی شرطی روی (U^TU) و $(X_j-\theta_j)^۲$ برای $j\neq i$ بحث می‌کنیم. دقت کنید که با توجه به این‌که بنابه فرض r نامنفی، مقعر و مشتق‌پذیر است، بنابه لم ۵۵.۲.۱ ، $r(\|X\|^۲)/\|X\|^۲$ در $\|X\|^۲$ ناصعودی است. آن‌گاه چون توزیع شرطی X_i بر روی U^TU و $(X_j-\theta_j)^۲$ برای $j\neq i$، برای θ_i متقارن و تک‌مدی است، بنابه لم ۵۶.۲.۱ داریم

$$\begin{aligned} &cE\left[U^TU\frac{r(\|X\|^۲)}{\|X\|^۲}X_i^۲I_{[X_i\leq ۰]}\Big/ U^TU=\rho^۲,(X_j-\theta_j)^۲=Z_j^۲, j\neq i\right]\\ &\leq \frac{c}{۲}E\left[U^TU\frac{r(\|X\|^۲)}{\|X\|^۲}Z_i^۲\Big/ U^TU=\rho^۲,(X_j-\theta_j)^۲=Z_j^۲, j\neq i\right]. \end{aligned} \tag{۲۵.۲}$$

دقت کنید که تمایز این اثبات با اثبات قضیه ۴.۲.۲ در کران مجموع است. در ادامه فرض می‌کنیم

$$\begin{cases} V=(Z_۱,\dots,Z_s),\\ \eta=(\theta_۱,\dots,\theta_s),\\ T=(Z_{s+۱},\dots,Z_p),\\ \mu=(\theta_{s+۱},\dots,\theta_p). \end{cases}$$

که در آن $Z=X-\theta$ و با فرض این‌که

$$\begin{cases} X_۱^s=(X_۱,\dots,X_s),\\ X_{s+۱}^p=(X_{s+۱},\dots,X_p). \end{cases}$$

داریم

$$V=(V_۱,\dots,V_s)=((X_۱-\theta_۱),\dots,(X_s-\theta_s))=X_۱^s-\eta,$$

و

$$T=(T_{s+۱},\dots,T_p)=((X_{s+۱}-\theta_{s+۱}),\dots,(X_p-\theta_p))=X_{s+۱}^p-\mu.$$

بنابراین می‌توان $\|X\|^2$ را به صورت نیز نوشت

$$\begin{aligned}\|X\|^2 &= X^TX = \sum_{i=1}^{p} X_i^2\\ &= \sum_{i=1}^{s} X_i^2 + \sum_{i=s+1}^{p} X_i^2\\ &= \|X_1^s\|^2 + \|X_{s+1}^p\|^2\\ &= \|(V+\eta)\|^2 + \|(T+\mu)\|^2.\end{aligned}$$

همچنین فرض کنید داشته باشیم

$$W^2 = V^TV + U^TU.$$

پس با در نظر گرفتن نامساوی (۲۵.۲) و لم ۵۶.۲.۱ تفاضل مخاطره در (۲۴.۲)، دارای کران بالا به صورت زیر است.

$$\begin{aligned}&E_\theta\left[(U^TU)^2\frac{r(\|X\|^2)}{\|X\|^2}c\left(c-2\frac{p-2}{k+2}+2\frac{\sum_{i=1}^{s}X_i^2I_{[X_i\leq 0]}}{U^TU}\right)\right]\\ &\leq cE_\theta\left[(W^2-V^TV)^2\frac{r(\|V+\eta\|^2+\|T+\mu\|^2)}{\|V+\eta\|^2+\|T+\mu\|^2}\right.\\ &\left.\times\left(c-2\frac{p-2}{k+2}+\frac{V^TV}{W^2-V^TV}\right)\right]. \qquad (۲۶.۲)\end{aligned}$$

دقت کنید که با استفاده از لم ۵۵.۲.۱ و همچنین شرایط (۱) و (۲) $r(\|X\|^2)/\|X\|^2$ برای $p\geq 4$ زبرهمساز و در $\|X\|^2$ ناصعودی است. با شرطی کردن (۲۶.۲) روی W^2 و T معین در تفاضل مخاطره شرطی (۲۶.۲) داریم

$$\begin{aligned}&E\left[(W^2-V^TV)\frac{r(\|V+\eta\|^2+\|T+\mu\|^2)}{\|V+\eta\|^2+\|T+\mu\|^2}c\left(c-2\frac{p-2}{k+2}+\frac{V^TV}{W^2-V^TV}\right)\Big/W^2,T\right]\\ &\leq cE\left[(W^2-V^TV)\frac{r(\|V+\eta\|^2+\|T+\mu\|^2)}{\|V+\eta\|^2+\|T+\mu\|^2}\Big/W^2,T\right]\\ &\times E\left[\left(c-2\frac{p-2}{k+2}+\frac{V^TV}{W^2-V^TV}\right)\Big/W^2,T\right]. \qquad (۲۷.۲)\end{aligned}$$

چون توزیع شرطی V به شرط W^2 متقارن کروی است، پس

$$E\left[\frac{r(\|V+\eta\|^2+\|T+\mu\|^2)}{\|V+\eta\|^2+\|T+\mu\|^2}\right],$$

با استفاده از لم ۳.۲.۲ در V^TV برای $V^TV < W^۲$ ناصعودی است. همچنین $(W^۲ - V^TV)$ در V^TV ناصعودی و $V^TV/(W^۲ - V^TV)$ نانزولی است. بنابراین نامساوی (۲۷.۲) را می‌توان از لم ۵۸.۲.۱ نتیجه گرفت. از اینرو، تنها کافیست تا نشان دهیم دومین امید ریاضی شرطی در سمت راست نامساوی (۲۷.۲) نامثبت است. در اینجا با در نظر گرفتن $W^۲$ و T، کسر $V^TV/W^۲$، بنابه لم ۲.۲.۲ دارای توزیع $B(\frac{s}{۲}, \frac{k}{۲})$ است و از اینرو با استفاده از لم ۴۴.۲.۱ داریم

$$E\left[\frac{V^TV}{(W^۲ - V^TV)}\Big/ W^۲\right] = \frac{s}{k-۲}.$$

بنابراین تفاضل مخاطره (۲۷.۲) هنگامی نامثبت است که

$$-۲\frac{p-۲}{k+۲} + E\left[\frac{V^TV}{W^۲ - V^TV}\Big/ W^۲\right] \leq ۰,$$

سپس

$$c - ۲\frac{p-۲}{k+۲} + \frac{s}{k-۲} \leq ۰,$$

بنابراین

$$c \leq ۲\frac{p-۲}{k+۲} - \frac{s}{k-۲},$$

در نتیجه

$$۰ < c \leq ۲\frac{p-۲}{k+۲} - \frac{s}{k-۲}.$$

از طرف دیگر کران بالای c تنها هنگامی مثبت است که

$$۲\frac{p-۲}{k+۲} - \frac{s}{k-۲} > ۰,$$

با مخرج مشترک گرفتن داریم

$$\frac{۲(p-۲)(k-۲) - s(k+۲)}{(k-۲)(k+۲)} > ۰,$$

سپس

$$۲pk - ۴p - ۴k + ۸ - sk - ۲s > ۰,$$

و

$$k(۲p - s - ۴) > ۴p + ۲s - ۸,$$

و در نتیجه

$$k > ۲ + \frac{۴s}{۲p - s - ۴}.$$

اثبات کامل است. □

فرع ۲.۳.۲. کران $۲\frac{p-۲}{k+۲} - \frac{s}{k-۲}$ برای $۰ < c \leq ۲\frac{p-۲}{k+۲} - \frac{s}{k-۲}$ برای $p > s$ بزرگتر از $۰ < c \leq ۲\frac{p-۲}{k+۲} - \frac{p}{k-۲}$ است و از اینرو قضیه ۱.۳.۲ قوی تر از قضیه ۴.۲.۲ است. به این معنی که در آن c دامنه‌ی وسیعتری از اعداد را شامل می‌شود. در حالتی که $p = s$، است قضیه ۱.۳.۲ به قضیه ۴.۲.۲ تبدیل می‌شود.

۴.۲ برآوردگر انقباضی برای مخروط‌های تودرتو و ماتریس‌های متعامد

در این بخش، فرض می‌کنیم که دو مخروط محدب بسته مثبت $\mathscr{C}_۱$ و $\mathscr{C}_۲$ طوری وجود دارند که

(الف) $\mathscr{C}_۱ \subset Q\mathscr{C}_۲$،

(ب) $Q\theta \in \mathscr{C}_۱$.

که در آن Q یک ماتریس متعامد است. همچنین فرض می‌کنیم جامعه دارای توزیع متقارن کروی باشد. وقتی که θ متعلق به یکی از مخروط‌های $\mathscr{C}_i(i = ۱, ۲)$ است، برآوردگر طبیعی به صورت $\delta^i_۰(X) = P_{\mathscr{C}_i}(X)$ می‌باشد که در آن $P_{\mathscr{C}_i}(X)$ تصویر X بر روی مخروط $\mathscr{C}_i$ است. صورت دیگر نمایش این برآوردگر به شکل

$$\delta^i_۰(X) = X + \gamma_i(X),$$

است، که در آن $\gamma_i(X) = P_{\mathscr{C}_i}(X) - X$. در این صورت برآوردگر انقباضی را به صورت زیر در نظر می‌گیریم.

$$\begin{aligned}\delta^i(X) &= \delta^i_۰(X) - g(\|X\|^۲)U^TUX \\ &= X + \gamma_i(X) - g(\|X\|^۲)U^TUX, \qquad i = ۱, ۲.\end{aligned}$$

که در آن g یک تابع نامنفی است.

در این بخش، قصد داریم تحت شرایط (الف) و (ب) نشان دهیم، وقتی θ در $\mathscr{C}_۱$ محدودشده است و برآوردگر $\delta^۱(X,U)$ بر برآوردگر طبیعی $\delta^۱_۰(X)$ برتری دارد، آن‌گاه وقتی θ در $\mathscr{C}_۲$ محدودشده است، برآوردگر انقباضی $\delta^۲(X,U)$ بر برآوردگر طبیعی $\delta^۲_۰(X)$ برتری دارد.

لم ۱.۴.۲. *(فودینیر و همکاران، ۲۰۰۳a) اگر $\mathscr{C}$ یک مخروط محدب بسته مثبت [۲] باشد، آن‌گاه به ازای هر X داریم*

$$(X - P_{\mathscr{C}}(X))^T P_{\mathscr{C}}(X) = ۰.$$

برهان. روشن است که لم در حالت $P_{\mathscr{C}}(X) = 0$ درست می‌باشد. فرض کنید $P_{\mathscr{C}}(X) \neq 0$ باشد. چون $P_{\mathscr{C}}(X)$ تصویر بردار X بر روی مخروط $\mathscr{C}$ است، وقتی که تابع زیان مربعی است، تصویر به طول تبدیل می‌گردد و در نتیجه طول این تصویر از طول X کمتر خواهد بود، از طرف دیگر چون $\mathscr{C}$ یک مخروط مثبت است، بنابه تعریف ۷.۲.۱ داریم

$$\forall X \in \mathscr{C}, a \geq ۰; aX \in \mathscr{C},$$

در نتیجه چون طول $P_{\mathscr{C}}(X)$ از طول X کمتر است،

$$\forall a \geq ۰, \quad aP_{\mathscr{C}}(X) \in \mathscr{C}.$$

همچنین بنابه نکته ۲۰.۲.۱، تصویر هر نقطه دلخواه $X \in \mathbb{R}^p$، یکتاست. از اینرو

$$\begin{aligned} Q(a) &= \|X - aP_{\mathscr{C}}(X)\|^۲ \\ &= \|X - P_{\mathscr{C}}(X) + (۱ - a)P_{\mathscr{C}}(X)\|^۲ \\ &= \|X - P_{\mathscr{C}}(X)\|^۲ + (۱ - a)^۲\|P_{\mathscr{C}}(X)\|^۲ + ۲(۱ - a)(P_{\mathscr{C}}(X))^T(X - P_{\mathscr{C}}(X)). \end{aligned}$$

$Q(a)$ برای $a > ۰$ در $a = ۱$ دارای مینیمم یکتا است. به عبارت دیگر $Q'(۱) = ۰$ یا

$$\begin{aligned} Q'(۱) &= [-۲(۱ - a)\|P_{\mathscr{C}}(X)\|^۲ - ۲(P_{\mathscr{C}}(X))^T(X - P_{\mathscr{C}}(X))]\Big|_{a=۱} \\ &= -۲P_{\mathscr{C}}(X))^T(X - P_{\mathscr{C}}(X)) = ۰. \end{aligned}$$

[۲] Positive closed convex cone

و در نتیجه

$$(X - P_{\mathscr{C}}(X))^T P_{\mathscr{C}}(X) = ۰.$$

□

لم ۲.۴.۲. (فودینیر و همکاران، ۲۰۰۳a) *اگر $\mathscr{C}_۱$ و $\mathscr{C}_۲$ دو مخروط محدب بسته مثبت باشند طوری‌که $\mathscr{C}_۱ \subset \mathscr{C}_۲$، آن‌گاه به ازای هر $X \in \mathbb{R}^p$*

$$X^T P_{\mathscr{C}_۱}(X) \leq X^T P_{\mathscr{C}_۲}(X).$$

همچنین به طور مشابه به ازای $\gamma_i(X) = P_{\mathscr{C}_i}(X) - X$ برای $i = ۱, ۲$، داریم

$$X^T \gamma_۱(X) \leq X^T \gamma_۲(X).$$

برهان. بنابه فرض، چون $\mathscr{C}_۱ \subset \mathscr{C}_۲$ و با استفاده از تعریف ۱۹.۲.۱، به ازای هر X

$$(dist(X, \mathscr{C}_۱))^۲ = \|X - P_{\mathscr{C}_۱}(X)\|^۲ \geq \|X - P_{\mathscr{C}_۲}(X)\|^۲ = (dist(X, \mathscr{C}_۲))^۲. \quad (۲۸.۲)$$

از طرف دیگر

$$\begin{aligned} \|X - P_{\mathscr{C}_i}(X)\|^۲ &= (X - P_{\mathscr{C}_i}(X))^T (X - P_{\mathscr{C}_i}(X)) \\ &= (X - P_{\mathscr{C}_i}(X))^T X - (X - P_{\mathscr{C}_i}(X))^T P_{\mathscr{C}_i}(X). \quad (۲۹.۲) \end{aligned}$$

حال بنابه لم ۱.۴.۲، $(X - P_{\mathscr{C}_۱}(X))^T P_{\mathscr{C}_۱}(X) = ۰$ و $(X - P_{\mathscr{C}_۲}(X))^T P_{\mathscr{C}_۲}(X) = ۰$. با استفاده از رابطه‌های (۲۸.۲) و (۲۹.۲)، می‌توان نتیجه گرفت

$$\begin{aligned} \|X - P_{\mathscr{C}_۱}(X)\|^۲ &= (X - P_{\mathscr{C}_۱}(X))^T X - (X - P_{\mathscr{C}_۱}(X))^T P_{\mathscr{C}_۱}(X) \\ &= (X - P_{\mathscr{C}_۱}(X))^T X - ۰ \\ &\geq (X - P_{\mathscr{C}_۲}(X))^T X - ۰ \\ &= (X - P_{\mathscr{C}_۲}(X))^T X - (X - P_{\mathscr{C}_۲}(X))^T P_{\mathscr{C}_۲}(X) \\ &= \|X - P_{\mathscr{C}_۲}(X)\|^۲ \end{aligned}$$

درنتیجه

$$X^T X - P_{\mathscr{C}_1}(X)^T X \geq X^T X - P_{\mathscr{C}_2}(X)^T X,$$

و بنابراین

$$X^T P_{\mathscr{C}_1}(X) \leq X^T P_{\mathscr{C}_2}(X). \tag{۳۰.۲}$$

برای اثبات قسمت دوم با کم کردن $X^T X$ از طرفین (۳۰.۲) داریم

$$X^T P_{\mathscr{C}_1}(X) - X^T X \leq X^T P_{\mathscr{C}_2}(X) - X^T X,$$

و

$$X^T (P_{\mathscr{C}_1}(X) - X) \leq X^T (P_{\mathscr{C}_2}(X) - X).$$

از طرفی بنابه فرض $\gamma_i(X) = P_{\mathscr{C}_i}(X) - X$، پس می‌توان نتیجه گرفت

$$X^T \gamma_1(X) \leq X^T \gamma_2(X).$$

$\square$

لم ۳.۴.۲. *(فودینیر و همکاران، $۲۰۰۳a$) اگر $\mathscr{C}$ یک مخروط محدب بسته و Q ماتریس متعامد متناظر با آن باشد، آن‌گاه*

$$P_{Q\mathscr{C}}(QX) = QP_{\mathscr{C}}(X),$$

و از اینرو

$$P_{Q\mathscr{C}}(QX) - QX = Q(P_{\mathscr{C}}(X) - X). \tag{۳۱.۲}$$

قضیه ۴.۴.۲. *(فودینیر و همکاران، $۲۰۰۳a$) تحت مفروضات این بخش فرض کنید $\mathscr{C}_1 \subset Q\mathscr{C}_2$ و $Q\theta \in \mathscr{C}_1$. اگر وقتی که θ روی $\mathscr{C}_1$ محدودشده است، برآوردگر انقباضی $\delta^1(X,U)$ بهتر از برآوردگر طبیعی $\delta_0^1(X)$ باشد، آن‌گاه وقتی که θ روی $\mathscr{C}_2$ محدودشده است، برآوردگر انقباضی $\delta^2(X,U)$ بهتر از برآوردگر طبیعی $\delta_0^2(X)$ است.*

برهان. چون $\mathscr{C}_1 \subset \mathscr{C}_2$ بین دو حالت ۱. $\theta \in \mathscr{C}_1$ و ۲. $\theta \in \mathscr{C}_2 - \mathscr{C}_1$، تمایز قائل می‌شویم. ابتدا فرض کنید که $\theta \in \mathscr{C}_1$ است. دقت کنید که تفاضل بین مخاطره دو برآوردگر $\delta^2(X,U)$ و

$\delta_\circ^{\text{۲}}(X)$ برابر است با

$$\begin{aligned}\Delta\mathscr{R}_{\text{۲}}(\theta) &= E_\theta[\|X+\gamma_{\text{۲}}(X)-g(\|X\|^{\text{۲}})U^TUX-\theta\|^{\text{۲}}-\|X+\gamma_{\text{۲}}(X)-\theta\|^{\text{۲}}]\\ &= E_\theta[\|X+\gamma_{\text{۲}}(X)-\theta\|^{\text{۲}}+\|g(\|X\|^{\text{۲}})U^TUX\|^{\text{۲}}\\ &-\text{۲}(X+\gamma_{\text{۲}}(X)-\theta)g(\|X\|^{\text{۲}})U^TUX-\|X+\gamma_{\text{۲}}(X)-\theta\|^{\text{۲}}]\\ &= E_\theta[(U^TU)^{\text{۲}}g^{\text{۲}}(\|X\|^{\text{۲}})\|X\|^{\text{۲}}-\text{۲}U^TUg(\|X\|^{\text{۲}})X^T(X-\theta)\\ &-\text{۲}U^TUg(\|X\|^{\text{۲}})X^T\gamma_{\text{۲}}(X)].\end{aligned}$$

با استفاده از لم ۲.۴.۲ به ازای هر $X\in\mathbb{R}^p$، $X^T\gamma_{\text{۲}}(X)\geq X^T\gamma_{\text{۱}}(X)$. از اینرو چون $g(\|X\|^{\text{۲}})\geq \circ$ و از طرف دیگر چون بنابه فرض $\delta^{\text{۱}}(X,U)$ برای $\theta\in\mathscr{C}$ بهتر از $\delta_\circ^{\text{۱}}(X)$ است، داریم

$$\begin{aligned}\Delta\mathscr{R}_{\text{۲}}(\theta) &= E_\theta[(U^TU)^{\text{۲}}g^{\text{۲}}(\|X\|^{\text{۲}})\|X\|^{\text{۲}}-\text{۲}U^TUg(\|X\|^{\text{۲}})X^T(X-\theta)\\ &-\text{۲}U^TUg(\|X\|^{\text{۲}})X^T\gamma_{\text{۲}}(X)]\\ &\leq E_\theta[(U^TU)^{\text{۲}}g^{\text{۲}}(\|X\|^{\text{۲}})\|X\|^{\text{۲}}-\text{۲}U^TUg(\|X\|^{\text{۲}})X^T(X-\theta)\\ &-\text{۲}U^TUg(\|X\|^{\text{۲}})X^T\gamma_{\text{۱}}(X)]\\ &= \Delta\mathscr{R}_{\text{۱}}(\theta)\leq\circ.\end{aligned}$$

در حالت دوم $\theta\in\mathscr{C}_{\text{۲}}-\mathscr{C}_{\text{۱}}$، فرض کنید Q ماتریس متعامد مفروض باشد. این‌بار فرض کنید که $Y=QX$ و این نکته را در نظر بگیرید که (Y,V) دارای توزیع متقارن کروی حول $(\nu,0)$ که در آن $\nu=Q\theta$ می‌باشد. همچنین تابع γ متناظر با Y را به صورت $\gamma^*(Y)$ فرض می‌کنیم. حال مسئله برآورد ν وقتی که ν به $Q\mathscr{C}_{\text{۲}}$ محدودشده است را در نظر بگیرید. چون $\mathscr{C}_{\text{۱}}\subset Q\mathscr{C}_{\text{۲}}$ با استفاده از نتیجه حالت اول که برای $Q\theta=\nu\in\mathscr{C}_{\text{۱}}$ بود، داریم

$$E_\nu[\|Y+\gamma_{\text{۲}}^*(Y)-V^TVg(\|Y\|^{\text{۲}})Y-\nu\|^{\text{۲}}]\leq E_\nu[\|Y+\gamma_{\text{۲}}^*(Y)-\nu\|^{\text{۲}}], \qquad (\text{۳۲.۲})$$

که در آن $\gamma_{\text{۲}}^*(Y)=\gamma_{\text{۲}}^*(QX)=P_{Q\mathscr{C}_{\text{۲}}}(QX)-QX$. در نتیجه، به طور معادل، می‌توان رابطه

(۳۲.۲) را به صورت زیر نوشت.

$$\begin{aligned} &E_\theta[\|QX + \gamma_2^*(QX) - \|QV\|^2 g(\|QX\|^2)QX - Q\theta\|^2] \\ &- E_\theta[\|QX + \gamma_2^*(QX) - Q\theta\|^2] \leq 0. \end{aligned} \tag{۳۳.۲}$$

از طرف دیگر با استفاده از لم ۳.۴.۲، داریم

$$\gamma_2^*(QX) = P_{Q\mathscr{C}_2}(QX) - QX = Q(P_{\mathscr{C}_2}(X) - X) = Q\gamma_2(X). \tag{۳۴.۲}$$

از اینرو با استفاده از رابطه (۳۴.۲) برای رابطه (۳۳.۲) داریم

$$\begin{aligned} &E_\theta[\|QX + \gamma_2^*(QX) - \|QV\|^2 g(\|QX\|^2)QX - Q\theta\|^2] \\ &\quad - E_\theta[\|QX + \gamma_2^*(QX) - Q\theta\|^2] \\ &\quad = E_\theta[\|QX + Q\gamma_2(X) - (QV)^T QV g((QX)^T QX)QX - Q\theta\|^2] \\ &\quad - E_\theta[\|QX + Q\gamma_2(X) - Q\theta\|^2] \\ &\quad = E_\theta[\|QX + Q\gamma_2(X) - QU^T U g(X^T Q^T QX)X - Q\theta\|^2] \\ &\quad - E_\theta[\|QX + Q\gamma_2(X) - Q\theta\|^2] \\ &\quad = E_\theta[\|Q(X + \gamma_2(X) - U^T U g(\|X\|^2)X - \theta)\|^2] \\ &\quad - E_\theta[\|Q(X + \gamma_2(X) - \theta)\|^2] \\ &\quad = E_\theta[\|X + \gamma_2(X) - U^T U g(\|X\|^2)X - \theta\|^2] - E_\theta[\|X + \gamma_2(X) - \theta\|^2] \\ &\quad = \Delta\mathscr{R}_2(\theta). \end{aligned}$$

حال چون در حالت اول قضیه برای $\theta \in \mathscr{C}_1$ نشان دادیم $\Delta\mathscr{R}_2(\theta) \leq \Delta\mathscr{R}_1(\theta) \leq 0$، با استفاده از این نامساوی در این قسمت نیز داریم

$$\Delta\mathscr{R}_2(\theta) \leq 0.$$

در نتیجه وقتی θ روی $\mathscr{C}_2$ محدودشده است، برآوردگر انقباضی $\delta^2(X,U)$ بهتر از برآوردگر طبیعی $\delta_0^2(X)$ است. $\square$

در این قسمت به بررسی حالت کلی که تنها به یک مخروط محدود است، می‌پردازیم. برای این حالت برآوردگر انقباضی را به صورت

$$\delta(X) = X + \gamma_{\mathscr{C}}(X) - g(\|X\|^{۲})X,$$

و برآوردگر طبیعی را به صورت

$$\delta_{\circ}(X) = X + \gamma_{\mathscr{C}}(X),$$

و همچنین برآوردگر انقباضی متناظر برای ماتریس متعامد Q را به صورت

$$\delta^{*}(X) = X + \gamma_{Q\mathscr{C}}(X) - g(\|X\|^{۲})X,$$

در نظر می‌گیریم.

قضیه زیر از بحث بالا نتیجه گرفته می‌شود.

قضیه ۵.۴.۲. *(فودینیر و همکاران، a۲۰۰۳) *فرض کنید $\mathscr{C}$ یک مخروط محدب بسته باشد، به طوری که به ازای هر $\theta \in \mathscr{C}$ برآوردگر $\delta(X)$ برتر از برآوردگر $\delta_{\circ}(X)$ باشد. در این صورت به ازای هر ماتریس متعامد Q و به ازای هر $\theta \in Q\mathscr{C}$، برآوردگر $\delta^{*}(X)$ برتر از برآوردگر $\delta_{\circ}(X)$ است.*

برهان. فرض کنید که $Y = QX$ و این نکته را در نظر بگیرید که (Y, V) دارای توزیع متقارن کروی حول $(\nu, 0)$ بوده که در آن $\nu = Q\theta$. همچنین تابع γ متناظر با Y را به صورت $\gamma^{*}(Y)$ فرض می‌کنیم. حال مسئله برآورد ν وقتی که ν به $Q\mathscr{C}$ محدودشده است را در نظر بگیرید. برای اثبات برتری برآوردگر $\delta^{*}(X)$ از برآوردگر $\delta_{\circ}(X)$ باید رابطه زیر برقرار باشد.

$$E_{\nu}[\|Y + \gamma^{*}(Y) - V^{T}Vg(\|Y\|^{۲})Y - \nu\|^{۲}] \leq E_{\nu}[\|Y + \gamma^{*}(Y) - \nu\|^{۲}], \qquad (۳۵.۲)$$

که در آن $\gamma^{*}(Y) = \gamma^{*}(QX) = P_{Q\mathscr{C}}(QX) - QX$. در نتیجه به طور معادل برای رابطه (۳۵.۲) داریم

$$E_{\theta}[\|QX + \gamma^{*}(QX) - \|QV\|^{۲}g(\|QX\|^{۲})QX - Q\theta\|^{۲}]$$
$$- E_{\theta}[\|QX + \gamma^{*}(QX) - Q\theta\|^{۲}] \leq \circ. \qquad (۳۶.۲)$$

حال با استفاده از لم ۳.۴.۲ می‌توان نتیجه گرفت

$$\gamma^*(QX) = P_{Q\mathscr{C}}(QX) - QX = Q(P_{\mathscr{C}}(X) - X) = Q\gamma(X). \tag{۳۷.۲}$$

از اینرو رابطه (۳۶.۲) را با استفاده از رابطه (۳۷.۲)، می‌توان به صورت زیر نوشت.

$$\begin{aligned}
&E_\theta[\|QX + \gamma^*(QX) - \|QV\|^2 g(\|QX\|^2)QX - Q\theta\|^2] \\
&\quad - E_\theta[\|QX + \gamma^*(QX) - Q\theta\|^2] \\
&\quad = E_\theta[\|QX + Q\gamma(X) - (QV)^T QV g((QX)^T QX)QX - Q\theta\|^2]
\end{aligned}$$

$$\begin{aligned}
&- E_\theta[\|QX + Q\gamma(X) - Q\theta\|^2] \\
&= E_\theta[\|QX + Q\gamma(X) - QU^T U g(X^T Q^T QX)X - Q\theta\|^2] \\
&- E_\theta[\|QX + Q\gamma(X) - Q\theta\|^2] \\
&= E_\theta[\|Q(X + \gamma(X) - U^T U g(\|X\|^2)X - \theta)\|^2] - E_\theta[\|Q(X + \gamma(X) - \theta)\|^2] \\
&= E_\theta[\|X + \gamma(X) - U^T U g(\|X\|^2)X - \theta\|^2] - E_\theta[\|X + \gamma(X) - \theta\|^2] \\
&= \Delta\mathscr{R}(\theta).
\end{aligned}$$

حال چون بنابه فرض برای $\theta \in \mathscr{C}$ برآوردگر $\delta(X)$ برتر از برآوردگر $\delta_0(X)$ است، می‌توان نتیجه گرفت $\Delta\mathscr{R}(\theta) \leq 0$. به عبارت دیگر به ازای هر ماتریس متعامد Q، برآوردگر $\delta^*(X)$ برتر از برآوردگر $\delta_0(X)$ برای $\theta \in Q\mathscr{C}$ است. □

۵.۲ برآوردگر انقباضی محدودشده در توزیع نرمال

در این بخش، مشابه با بخش دوم در حالت محدودیت کامل و بخش سوم در حالت محدودیت جزئی، به کلاسی از برآوردگرهای برتر برای پارامتر مجهول تحت دو محدودیت ذکر شده در توزیع نرمال می‌پردازیم.

فرض کنید $X = (X_1, \ldots, X_p)$ یک بردار تصادفی p بعدی با توزیع نرمال p - متغیره با پارامتر مکان مجهول θ و پارامتر مقیاس $\sigma^2 I_p$ با σ^2 معلوم و بدون بردار باقیمانده U است. علاقه‌مندیم

پارامتر مجهول θ را تحت تابع زیان مربعی برآورد کنیم. همانند بخش دوم، منظور از محدودیت کامل، نامنفی بودن مؤلفه $\theta_i(i = ۱, \ldots, p)$ در بردار پارامتر θ است.
برآوردگر طبیعی $\delta_\circ(X)$ را به صورت (۶.۲) در نظر گرفته و برآوردگر انقباضی را به صورت زیر در نظر می‌گیریم.

$$\begin{aligned} \delta(X) &= \delta_\circ(X) + g(X) \\ &= X + \gamma(X) + g(X). \end{aligned} \tag{۳۸.۲}$$

که در آن $g : \mathbb{R}^p \to \mathbb{R}^p$ تابعی حقیقی با خواصی می‌باشد که در ادامه به آن اشاره می‌کنیم. همان‌طور که دیده می‌شود این کلاس از برآوردگرها فاقد بردار باقیمانده U است.
تفاضل مخاطره $\Delta\mathscr{R}$ را در زیر محاسبه می‌نماییم.

$$\begin{aligned} \Delta\mathscr{R} &= \mathscr{R}(\theta, \delta) - \mathscr{R}(\theta, \delta_\circ) \\ &= E_\theta[\|\delta - \theta\|^۲ - \|\delta_\circ - \theta\|^۲] \\ &= E_\theta[\|X + \gamma(X) + g(X) - \theta\|^۲ - \|X + \gamma(X) - \theta\|^۲] \\ &= E_\theta[\|g(X)\|^۲ + ۲(g(X))^T(X + \gamma(X) - \theta)] \\ &= E_\theta[\|g(X)\|^۲ + (g(X))^T(X - \theta) + ۲(g(X))^T\gamma(X)]. \end{aligned} \tag{۳۹.۲}$$

در تفاضل مخاطره (۳۹.۲) ، دومین جمله از آخرین امید ریاضی به θ وابسته است. در این حالت برای از بین بردن نقش θ دیگر نیازی به استفاده از نتیجه ۵۲.۲.۱ نیست، زیرا در این حالت مشکل وجود θ در امیدریاضی را به کمک لم استاین حل خواهیم نمود. مشابه با قضیه ۴.۲.۲، قضیه زیر را در خصوص برتری برآوردگر انقباضی بر برآوردگر طبیعی بیان و اثبات می‌نماییم.

قضیه ۱.۵.۲. *** تحت مفروضات این بخش، فرض کنید $p > ۴$ و X دارای توزیع $N_p(\theta, \sigma^۲ I_p)$ با $\sigma^۲$ معلوم است. برآوردگر انقباضی به شکل $\delta(X) = X + \gamma(X) + g(X)$ را در نظر بگیرید، که در آن تابع g به صورت $g(X) = \frac{-c\sigma^۲ r(\|X\|^۲)}{\|X\|^۲} X$ می‌باشد. اگر تابع r و ثابت c در شرایط زیر صدق کنند.*

۱. $\circ \leq r(\|X\|^{۲}) \leq ۱$،

۲. $r(\|X\|^{۲})$ *مقعر و دارای مشتق دوم باشد،*

۳. $\circ < c \leq p - ۴$،

آن‌گاه برآوردگر انقباضی $\delta(X)$ *بر برآوردگر طبیعی* $\delta_\circ(X) = X + \gamma(X)$ *برتری دارد.*

برهان. با استفاده از لم ۵۵.۲.۱ از آن جایی که r نامنفی، مقعر و مشتق‌پذیر است، $r'(\cdot) \geq \circ$ و لم (تعمیم لم استاین) ۵۰.۲.۱، پس از جاگذاری مقدار $g(X)$ در تفاضل مخاطره (۳۹.۲)، داریم

$$\begin{aligned}
\Delta\mathscr{R} &= E_\theta[\|g(X)\|^{۲} + (g(X))^T(X-\theta) + ۲(g(X))^T\gamma(X)] \\
&= E_\theta[\|g(X)\|^{۲} + ۲\sigma^{۲}\nabla\cdot g(X) + ۲(g(X))^T\gamma(X)] \\
&= E_\theta\Bigg[\|-\frac{c\sigma^{۲}r(\|X\|^{۲})}{\|X\|^{۲}}X\|^{۲} + ۲\left(-\frac{c\sigma^{۲}r(\|X\|^{۲})}{\|X\|^{۲}}X\right)^T\gamma(X) \\
&+ ۲\sigma^{۲}\nabla\cdot\left(-\frac{c\sigma^{۲}r(\|X\|^{۲})}{\|X\|^{۲}}X\right)\Bigg] \\
&= E_\theta\Bigg[c^{۲}\sigma^{۴}\frac{r^{۲}(\|X\|^{۲})}{\|X\|^{۴}}\|X\|^{۲} + ۲c\sigma^{۲}\frac{r(\|X\|^{۲})}{\|X\|^{۲}}\sum_{i=۱}^{p}X_i^{۲}I_{[X_i\leq\circ]} \\
&- ۲c\sigma^{۴}\left(\frac{\Big(۲r'(\|X\|^{۲})\|X\|^{۲} + pr(\|X\|^{۲})\Big)\|X\|^{۲} - ۲r(\|X\|^{۲})\|X\|^{۲}}{\|X\|^{۴}}\right)\Bigg] \\
&= E_\theta\Bigg[c^{۲}\sigma^{۴}\frac{r^{۲}(\|X\|^{۲})}{\|X\|^{۲}} + ۲c\sigma^{۲}\frac{r(\|X\|^{۲})}{\|X\|^{۲}}\sum_{i=۱}^{p}X_i^{۲}I_{[X_i\leq\circ]} \\
&- ۲c\sigma^{۴}\left(\frac{۲r'(\|X\|^{۲})\|X\|^{۴} + pr(\|X\|^{۲})\|X\|^{۲} - ۲r(\|X\|^{۲})\|X\|^{۲}}{\|X\|^{۴}}\right)\Bigg] \\
&= E_\theta\Bigg[c^{۲}\sigma^{۴}\frac{r^{۲}(\|X\|^{۲})}{\|X\|^{۲}} + ۲c\sigma^{۲}\frac{r(\|X\|^{۲})}{\|X\|^{۲}}\sum_{i=۱}^{p}X_i^{۲}I_{[X_i\leq\circ]} \\
&- ۲c\sigma^{۴}\left((p-۲)\frac{r(\|X\|^{۲})}{\|X\|^{۲}} + ۲r'(\|X\|^{۲})\right)\Bigg] \\
&= E_\theta\Bigg[c^{۲}\sigma^{۴}\frac{r^{۲}(\|X\|^{۲})}{\|X\|^{۲}} + ۲c\sigma^{۲}\frac{r(\|X\|^{۲})}{\|X\|^{۲}}\sum_{i=۱}^{p}X_i^{۲}I_{[X_i\leq\circ]} \\
&- ۲c\sigma^{۴}(p-۲)\frac{r(\|X\|^{۲})}{\|X\|^{۲}} - ۴c\sigma^{۴}r'(\|X\|^{۲})\Bigg]
\end{aligned}$$

$$\leq E_\theta\left[c\sigma^۲\frac{r(\|X\|^۲)}{\|X\|^۲}\left(c\sigma^۲ r(\|X\|^۲)-۲\sigma^۲(p-۲)+۲\sum_{i=۱}^{p}X_i^۲ I_{[X_i\leq ۰]}\right)\right]. \qquad (۴۰.۲)$$

حال با استفاده از $۰ \leq r(\|X\|^۲) \leq ۱$ و با قرار دادن مقدار ماکزیمم $r(\|X\|^۲)$ که بنابه فرض برابر با ۱ است، داریم

$$\begin{aligned}\Delta\mathscr{R} &= E_\theta\left[c\sigma^۲\frac{r(\|X\|^۲)}{\|X\|^۲}\left(c\sigma^۲ r(\|X\|^۲)-۲\sigma^۲(p-۲)+۲\sum_{i=۱}^{p}X_i^۲ I_{[X_i\leq ۰]}\right)\right]\\ &\leq E_\theta\left[c\sigma^۲\frac{r(\|X\|^۲)}{\|X\|^۲}\left(c\sigma^۲-۲\sigma^۲(p-۲)+۲\sum_{i=۱}^{p}X_i^۲ I_{[X_i\leq ۰]}\right)\right]. \qquad (۴۱.۲)\end{aligned}$$

مشابه با استدلال قضیه ۴.۲.۲، بنابه لم ۵۶.۲.۱ برای حالت $\sigma^۲ = ۱$، داریم

$$\begin{aligned}&cE\left[\frac{r(\|X\|^۲)}{\|X\|^۲}X_i^۲ I_{[X_i\leq ۰]}\Big/ (X_j-\theta_j)^۲ = Z_j^۲, j\neq i\right]\\ &\leq \frac{c}{۲}E\left[\frac{r(\|X\|^۲)}{\|X\|^۲}Z_i^۲\Big/ (X_j-\theta_j)^۲ = Z_j^۲, j\neq i\right]. \qquad (۴۲.۲)\end{aligned}$$

بنابراین تفاضل مخاطره در (۴۱.۲) با فرض $Z = X-\theta$، دارای کران بالای به صورت زیر است.

$$\begin{aligned}&E_\theta\left[c\sigma^۲\frac{r(\|X\|^۲)}{\|X\|^۲}\left(c\sigma^۲-۲\sigma^۲(p-۲)+۲\sum_{i=۱}^{p}X_i^۲ I_{[X_i\leq ۰]}\right)\right]\\ &\leq cE_\theta\left[\frac{r(\|Z+\theta\|^۲)}{\|Z+\theta\|^۲}\sigma^۴\left(c-۲(p-۲)+\frac{Z^TZ}{\sigma^۲}\right)\right]. \qquad (۴۳.۲)\end{aligned}$$

دقت کنید چون r بنابه فرض، نامنفی، مقعر و دارای مشتق دوم است، با استفاده از لم ۵۵.۲.۱، واضح است که $r(\|X\|^۲)/\|X\|^۲$ برای $p \geq ۴$ زبرهمساز و در نتیجه در $\|X\|^۲$ ناصعودی است. اینک با شرطی کردن روی Z^TZ، برای عبارت شرطی (۴۳.۲) داریم

$$\begin{aligned}&E\left[E\left[\frac{r(\|Z+\theta\|^۲)}{\|Z+\theta\|^۲}\Big/ Z^TZ\right]\sigma^۴(c-۲(p-۲)+\frac{Z^TZ}{\sigma^۲})\Big/ Z^TZ\right]\\ &\leq cE\left[E\left[\frac{r(\|Z+\theta\|^۲)}{\|Z+\theta\|^۲}\Big/ Z^TZ\right]\right]\\ &\times E\left[\sigma^۴\left(c-۲(p-۲)+\frac{Z^TZ}{\sigma^۲}\right)\Big/ Z^TZ\right]. \qquad (۴۴.۲)\end{aligned}$$

از آن‌جایی که توزیع Z نرمال بوده و لذا متقارن کروی است، $E[(\|Z+\theta\|^۲/\|Z+\theta\|^۲]$ با استفاده از لم ۳.۲.۲، در Z^TZ ناصعودی می‌باشد و نامساوی (۴۴.۲) را می‌توان از لم ۵۸.۲.۱ نتیجه گرفت.

در سمت راست نامساوی (۴۴.۲)، اولین امید ریاضی شرطی نامنفی است. از اینرو، تفاضل مخاطره شرطی ($\Delta\mathscr{R}$) به شرط آنکه دومین مقدار امید ریاضی شرطی نامثبت باشد، نامثبت خواهد بود. اما بنابه نتیجه ۴۱.۲.۱، $Z^TZ/\sigma^۲$ دارای توزیع $\chi_p^۲$ است و از اینرو $E[Z^TZ/\sigma^۲] = p$. در نتیجه تفاضل مخاطره شرطی تنها در صورتی نامثبت خواهد بود که داشته باشیم

$$c - ۲(p-۲) + E\Big[\frac{Z^TZ}{\sigma^۲}\Big] \leq \circ,$$

سپس

$$c - ۲(p-۲) - p \leq \circ,$$

بنابراین

$$c \leq p - ۴,$$

در نتیجه

$$\circ < c \leq p - ۴$$

که همان شرط سوم قضیه است. □

در این قسمت، همانند بخش سوم، هدف دست‌یابی به کلاسی از برآوردگرهای برتر تحت محدودیت جزئی است. مانند قبل برآوردگر طبیعی را به صورت (۱۹.۲) در نظر گرفته و برآوردگر انقباضی را به صورت

$$\delta_s(X) = X + \gamma_s(X) + g(X), \qquad (۴۵.۲)$$

در نظر می‌گیریم.

پیش از آن که اثبات را آغاز نماییم، برای تسهیل در روند اثبات تفاضل مخاطره $\Delta\mathscr{R}$ را در زیر محاسبه می‌نماییم. محاسبات زیر، با تفاوتی جزئی، کاملاً مشابه محاسبات بخش سوم است.

$$\begin{aligned}
\Delta\mathscr{R} &= \mathscr{R}(\theta,\delta_s) - \mathscr{R}(\theta,\delta_\circ^s) \\
&= E_\theta[\|\delta_s-\theta\|^۲ - \|\delta_\circ^s-\theta\|^۲] \\
&= E_\theta[\|X+\gamma_s(X)+g(X)-\theta\|^۲ - \|X+\gamma_s(X)-\theta\|^۲]
\end{aligned}$$

$$= E_\theta[\|g(X)\|^۲ + ۲(g(X))^T(X-\theta) + ۲(g(X))^T\gamma_s(X)]. \qquad (۴۶.۲)$$

در ادامه، مشابه با قضیه ۱.۳.۲، قضیه زیر را در خصوص برتری برآوردگر انقباضی بر برآوردگر طبیعی بیان می‌نماییم.

قضیه ۲.۵.۲. *(فودینیر و همکاران، $۲۰۰۳a$) تحت مفروضات این بخش، فرض کنید $۴ \le s \le p$ و X دارای توزیع $N_p(\theta, \sigma^۲ I_p)$ با $\sigma^۲$ معلوم است. برآوردگر انقباضی به شکل $\delta_s(X) = X + \gamma_s(X) + g(X)$ را در نظر بگیرید، که در آن تابع g به صورت $g(X) = \frac{-c\sigma^۲ r(\|X\|^۲)}{\|X\|^۲}X$ می‌باشد. اگر تابع r و ثابت c در شرایط زیر صدق کنند.*

۱. $۰ < r(\|X\|^۲) \le ۱$،

۲. $r(\|X\|^۲)$ *مقعر و دارای مشتق دوم باشد،*

۳. $۰ < c \le ۲(p-۲) - s$،

آن‌گاه برآوردگر انقباضی $\delta_s(X)$ بر برآوردگر طبیعی $\delta_۰^s = X + \gamma_s(X)$ برتری دارد.

برهان. با استفاده از لم (تعمیم لم استاین) ۵۰.۲.۱ برای تفاضل مخاطره (۴۶.۲)، مشابه با تفاضل مخاطره (۴۰.۲) و با جاگذاری مقدار $g(X)$ داریم

$$\begin{aligned}
\Delta\mathscr{R} &= E_\theta[\|g(X)\|^۲ + ۲(g(X))^T(X-\theta) + ۲(g(X))^T\gamma_s(X)] \\
&= E_\theta[\|g(X)\|^۲ + ۲\sigma^۲\nabla\cdot g(X) + ۲(g(X))^T\gamma_s(X)] \\
&= E_\theta\Bigg[c^۲\sigma^۴\frac{r^۲(\|X\|^۲)}{\|X\|^۲} + ۲c\sigma^۲\frac{r(\|X\|^۲)}{\|X\|^۲}\sum_{i=۱}^{s}X_i^۲ I_{[X_i\le ۰]} \\
&\quad - ۲c\sigma^۴(p-۲)\frac{r(\|X\|^۲)}{\|X\|^۲} - ۴c\sigma^۴ r'(\|X\|^۲)\Bigg] \\
&\le E_\theta\left[c\sigma^۲\frac{r(\|X\|^۲)}{\|X\|^۲}\left(c\sigma^۲ r(\|X\|^۲) - ۲\sigma^۲(p-۲) + ۲\sum_{i=۱}^{s}X_i^۲ I_{[X_i\le ۰]}\right)\right].
\end{aligned}$$

حال بنابه لم ۵۵.۲.۱ که در آن $r'(\cdot) \ge ۰$ است و بنابه شرط ۱ که در آن $۰ \le r(\|X\|^۲) \le ۱$، با

قرار دادن مقدار ماکزیمم $r(\|X\|^{۲})$ که برابر با ۱ است و مشابه نامساوی (۱۴.۲) داریم

$$\begin{aligned} E_{\theta}\left[c\sigma^{۲}\frac{r(\|X\|^{۲})}{\|X\|^{۲}}\left(c\sigma^{۲}r(\|X\|^{۲})-۲\sigma^{۲}(p-۲)+۲\sum_{i=۱}^{s}X_i^{۲}I_{[X_i\leq ۰]}\right)\right] \\ \leq E_{\theta}\left[c\sigma^{۲}\frac{r(\|X\|^{۲})}{\|X\|^{۲}}\left(c\sigma^{۲}-۲\sigma^{۲}(p-۲)+۲\sum_{i=۱}^{s}X_i^{۲}I_{[X_i\leq ۰]}\right)\right]. \end{aligned} \tag{۴۷.۲}$$

در ادامه مانند قضیه ۱.۳.۲، برای اثبات فرض می‌کنیم

$$\begin{cases} V=(Z_۱,\ldots,Z_s), \\ \eta=(\theta_۱,\ldots,\theta_s), \\ T=(Z_{s+۱},\ldots,Z_p), \\ \mu=(\theta_{s+۱},\ldots,\theta_p). \end{cases}$$

از طرف دیگر چون $Z=X-\theta$ و با فرض این‌که

$$\begin{cases} X_۱^s=(X_۱,\ldots,X_s), \\ X_{s+۱}^p=(X_{s+۱},\ldots,X_p). \end{cases}$$

داریم

$$V=(V_۱,\ldots,V_s)=((X_۱-\theta_۱),\ldots,(X_s-\theta_s))=X_۱^s-\eta,$$

و

$$T=(T_{s+۱},\ldots,T_p)=((X_{s+۱}-\theta_{s+۱}),\ldots,(X_p-\theta_p))=X_{s+۱}^p-\mu.$$

در نتیجه می‌توان $\|X\|^{۲}$ را به صورت زیر نوشت.

$$\|X\|^{۲}=\|(V+\eta)\|^{۲}+\|(T+\mu)\|^{۲}.$$

پس مشابه با استدلال قضیه ۱.۳.۲ و با استفاده از عبارت (۱۶.۲) و لم ۵۶.۲.۱ می‌توان تفاضل مخاطره در (۴۷.۲) را به صورت زیر کراندار نمود. به عبارتی

$$\begin{aligned} E_{\theta}\left[c\sigma^{۲}\frac{r(\|X\|^{۲})}{\|X\|^{۲}}\left(c\sigma^{۲}-۲\sigma^{۲}(p-۲)+۲\sum_{i=۱}^{s}X_i^{۲}I_{[X_i\leq ۰]}\right)\right] \\ \leq cE_{\theta}\left[\frac{r(\|V+\eta\|^{۲}+\|T+\mu\|^{۲})}{\|V+\eta\|^{۲}+\|T+\mu\|^{۲}}\sigma^{۴}\left(c-۲(p-۲)+\frac{V^TV}{\sigma^{۲}}\right)\right]. \end{aligned} \tag{۴۸.۲}$$

ابتدا دقت کنید که با استفاده از لم ۵۵.۲.۱ و همچنین شرایط ۱ و ۲، $r(\|X\|^{۲})/\|X\|^{۲}$ برای $p\geq ۴$ زبرهمساز و در $\|X\|^{۲}$ ناصعودی و بنابه تعریف ۲۱.۲.۱ نامثبت است. حال با در نظر

گرفتن V^TV و T و شرطی کردن (۴۸.۲) روی V^TV و T معین، داریم

$$
\begin{aligned}
&cE_\theta\left[\frac{r(\|V+\eta\|^۲+\|T+\mu\|^۲)}{\|V+\eta\|^۲+\|T+\mu\|^۲}\sigma^۴\Big(c-۲(p-۲)\Big)+\frac{V^TV}{\sigma^۲}\Big)\right]\\
&\leq cE\left[\frac{r(\|V+\eta\|^۲+\|T+\mu\|^۲)}{\|V+\eta\|^۲+\|T+\mu\|^۲}\Big/V^TV,T\right]\\
&\times E\left[\sigma^۴\Big(c-۲(p-۲)+\frac{V^TV}{\sigma^۲}\Big)\Big/V^TV,T\right]. \qquad (۴۹.۲)
\end{aligned}
$$

از آن جایی که توزیع V نرمال بوده و لذا دارای توزیع متقارن کروی است،

$$E\left[\frac{r(\|V+\eta\|^۲+\|T+\mu\|^۲}{\|V+\eta\|^۲+\|T+\mu\|^۲}\right],$$

با استفاده از لم ۳.۲.۲ در V^TV ناصعودی است، نامساوی (۴۹.۲) را می‌توان از لم ۵۸.۲.۱ نتیجه گرفت. از اینرو، تنها کافیست تا نشان دهیم دومین امید ریاضی شرطی در (۴۹.۲) نامثبت است. در این جا با در نظر گرفتن V^TV و T، کسر $V^TV/\sigma^۲$، بنابه نتیجه ۴۱.۲.۱ دارای توزیع $\chi^۲_s$ است و از اینرو $E[V^TV/\sigma^۲]=s$. بنابراین

$$\sigma^۴\left(c-۲(p-۲)+E\Big[\frac{V^TV}{\sigma^۲}\Big]\right)\leq ۰,$$

سپس

$$\sigma^۴\Big(c-۲(p-۲)+s\Big)\leq ۰,$$

بنابراین

$$c\leq ۲(p-۲)-s,$$

در نتیجه

$$۰<c\leq ۲(p-۲)-s.$$

اثبات کامل است. □

فصل ۳
برآورد انقباضی پارامتر مکان با مقیاس مجهول

۱.۳ مقدمه

در این فصل، مشابه با فصل دوم، با استفاده از برآوردگرهای انقباضی سعی در برآورد بردار پارامتر در حالتی که این بردار دارای محدودیت است، داریم. مجدداً فرض می‌کنیم توزیع جامعه مورد نظر متقارن کروی باشد با این تفاوت که این بار پارامتر مقیاس ما مجهول است. به این صورت که توزیع جامعه را، توزیع متقارن کروی با پارامتر مکان θ و پارامتر مقیاس $\sigma^۲ I_p$، با $\sigma^۲$ مجهول در نظر می‌گیریم.

این بار مسئله برآورد پارامتر مکان محدودشده با استفاده از برآوردگرهای انقباضی را تنها در دو حالت ۱. محدودیت کامل و ۲. محدودیت جزئی، مورد بررسی قرار می‌دهیم.

نتایج این فصل از دو منظر قابل اهمیت می‌باشد. اول این‌که کلاس برآوردگرهای نوع بارانچیک[۱] (۱۹۷۰) به حالت فضای پارامتر محدود تعمیم داده شده و دوم این‌که کلاس برآوردگرهای انقباضی ارائه‌شده توسط فودینیر و همکاران (۲۰۰۳a) در توزیع‌های متقارن کروی و نرمال چندمتغیره، به حالت با واریانس مجهول گسترش داده شده است.

مشابه فصل دوم، در حالت اول فرض می‌کنیم تمام مولفه‌های بردار پارامتر نامنفی باشند، به عبارت دیگر $\theta_i \geq ۰$، $i = ۱, \ldots, p$ و در حالت دوم تنها یک زیرمجموعه نامنفی از بردار پارامتر θ را در

[۱] Baranchik

نظر می‌گیریم. با این دو پیش‌فرض یک کلاس از برآوردگرهای

$$\delta(X,U)=\delta_\circ(X)+g(X,S)U'U$$

که تحت تابع زیان مربعی (پایا) برای بردار پارامتر مکان θ بهتر از برآوردگر طبیعی $\delta_\circ(X)$ است را خواهیم یافت. در پایان هر دو محدودیت را در توزیع نرمال p - متغیره با پارامتر مکان θ و پارامتر مقیاس $\sigma^۲ I_p$ که در آن $\sigma^۲$ مجهول است را مورد بررسی قرار می‌دهیم.

۲.۳ برآورد انقباضی در حالت با محدودیت کامل

مشابه مفروضات بخش ۲.۲، فرض کنید (X,U)، که در آن $X=(X_۱,\dots,X_p)$ و $U=(U_۱,\dots,U_k)$، یک بردار تصادفی $p+k$ بعدی با توزیع متقارن کروی حول بردار $p+k$ بعدی $(\theta,0)$ باشد که در آن پارامتر θ مجهول است. واضح است که بعد X برابر با بعد $\theta=(\theta_۱,\dots,\theta_p)$ و مساوی با p و بعد U برابر با بعد بردار 0 و مساوی با k است. همچنین پارامتر مقیاس را $\sigma^۲ I_p$ با $\sigma^۲$ مجهول در نظر می‌گیریم.

علاقه‌مندیم پارامتر مجهول θ را تحت تابع زیان مربعی موزون ؟؟، برآورد کنیم. مانند فصل دوم، منظور از محدودیت کامل، نامنفی بودن مؤلفه $\theta_i (i=۱,\dots,p)$ در بردار پارامتر θ است. در این بخش، هدف برآورد بردار پارامتر تحت محدودیت کامل می‌باشد.

برآوردگر طبیعی θ که آن را با $\delta_\circ(X)$ نشان می‌دهیم، عبارت است از

$$\delta_\circ(X)=(\delta_{\circ۱},\dots,\delta_{\circ p}), \qquad (۱.۳)$$

که در آن

$$\delta_{\circ i}(X)=\max(X_i,\circ), \quad i=۱,\dots,p.$$

همانند فودینیر و همکاران (۲۰۰۳a) برای راحتی محاسبات شکل برآوردگر طبیعی را به صورت زیر نیز می‌توان نشان داد.

$$\delta_\circ(X)=X+\gamma(X), \qquad (۲.۳)$$

که در آن $\gamma(X) = (\gamma_۱(X), \ldots, \gamma_p(X))$ و هر مؤلفه $\gamma(X)$ به صورت زیر تعریف می‌شود.

$$\gamma_i(X) = \begin{cases} -X_i, & X_i < ۰, \\ ۰, & X_i \geq ۰. \end{cases} \tag{۳.۳}$$

حال رده‌ای از برآوردگرهای انقباضی به صورت زیر را در نظر بگیرید.

$$\begin{aligned} \delta(X,U) &= \delta_۰(X) + g(X,S)U^TU \\ &= X + \gamma(X) + g(X,S)U^TU, \end{aligned} \tag{۴.۳}$$

که در آن $g : \mathbb{R}^p \to \mathbb{R}^p$ تابعی حقیقی است. در طول این فصل تابع g به صورت زیر تعریف می‌شود. ابتدا فرض کنید $S^۲$ برآوردگر نا اریب $\sigma^۲$ است که از X مستقل می‌باشد. به ازای مقدار ثابت c، تابع g را به صورت زیر مشابه با بارانچیک (۱۹۷۰) در نظر می‌گیریم.

$$g(X,S) = -\frac{c\, r(F)}{F}X. \tag{۵.۳}$$

که در آن

$$F = \frac{\|X\|^۲}{S^۲}.$$

تابع r در شرایط هر قضیه به تناسب تعریف خواهد شد.

مشابه با (۱۱.۲) و با استفاده از تابع زیان مربعی (پایا)، تفاضل مخاطره $\Delta\mathscr{R}$ را در زیر محاسبه می‌نماییم.

$$\begin{aligned} \Delta\mathscr{R} &= \mathscr{R}(\theta,\delta) - \mathscr{R}(\theta,\delta_۰) \\ &= E_\theta\left[\frac{\|\delta-\theta\|^۲}{\sigma^۲} - \frac{\|\delta_۰-\theta\|^۲}{\sigma^۲}\right] \\ &= \frac{۱}{\sigma^۲}E_\theta[\|X+\gamma(X)+g(X,S)U^TU-\theta\|^۲ - \|X+\gamma(X)-\theta\|^۲] \\ &= \frac{۱}{\sigma^۲}E_\theta[\|g(X,S)U^TU\|^۲ + ۲U^TU(g(X,S))^T(X+\gamma(X)-\theta)] \\ &= \frac{۱}{\sigma^۲}E_\theta[\|g(X,S)\|^۲(U^TU)^۲ + ۲U^TU(g(X,S))^T(X-\theta) \\ &\quad + ۲U^TU(g(X,S))^T\gamma(X)]. \end{aligned} \tag{۶.۳}$$

همانطور که در رابطه (۶.۳) دیده می‌شود، دومین جمله از آخرین امید ریاضی به θ وابسته است. در این فصل برای از بین بردن نقش θ، بر خلاف فصل دوم نمی‌توان از نتیجه ۵۲.۲.۱ استفاده کرد.

چرا که تفاضل مخاطره به مقدار $\sigma^۲$ بستگی دارد. برای از بین بردن این مشکل مشابه با آرشی و طباطبایی (۲۰۱۰) از راهکار دیگری با استفاده از امید ریاضی دوگانه استفاده کرده و پس از شرطی کردن از نتیجه ۵۲.۲.۱ استفاده می‌کنیم.

همچنین در این‌جا نیز مشابه استدلال فودینیر و همکاران (۲۰۰۳a) برای آن‌که تفاضل مخاطره (۲۴.۳) متناهی باشد، باید $E_\theta[\|X\|^۲] < \infty$ و $E_\theta[\|g\|^۲] < \infty$ برقرار باشند.

قضیه زیر در خصوص برتری برآوردگر انقباضی بر برآوردگر طبیعی را بیان و اثبات می‌نماییم.

قضیه ۱.۲.۳. *** تحت مفروضات این بخش فرض کنید $p > ۴$، $k > ۲ + \frac{۴p}{p-۴}$ و توزیع نیز تک‌مدی باشد. برآوردگر انقباضی $\delta(X,U) = X + \gamma(X) + g(X,S)U^TU$ را در نظر بگیرید که در آن تابع g به صورت (۵.۳) است. اگر تابع r و ثابت c در سه شرط زیر صدق کنند.*

۱. $۰ \le r(\frac{\|X\|^۲}{S^۲}) \le ۱$،

۲. $r(\frac{\|X\|^۲}{S^۲})$ *مقعر و دارای مشتق دوم باشد،*

۳. $۰ < c \le ۲\frac{(p-۲)}{k+۲}\frac{E_{\sigma=۱}(S^۲)}{E_{\sigma=۱}(S^۴)} - \frac{p}{k-۲}\frac{E_{\sigma=۱}(S^۲)}{E_{\sigma=۱}(S^۴)}$،

آن‌گاه به ازای هر $\theta \in \mathbb{R}^p$، برآوردگر انقباضی $\delta(X,U)$ بر برآوردگر $\delta_۰(X) = X + \gamma(X)$ برتری دارد.

برهان. با جایگذاری مقدار g از رابطه (۵.۳) در تفاضل مخاطره (۶.۳) داریم

$$\begin{aligned}\Delta\mathscr{R} &= \frac{۱}{\sigma^۲}E_\theta[\|g(X,S)\|^۲(U^TU)^۲ + ۲U^TU(g(X,S))^T(X-\theta) + ۲U^TU(g(X,S))^T\gamma(X)]\\ &= \frac{۱}{\sigma^۲}E_\theta\Bigg[\| - \frac{c\, r(\frac{\|X\|^۲}{S^۲})S^۲}{\|X\|^۲}X\|^۲(U^TU)^۲ + ۲(U^TU)\left(-\frac{c\, r(\frac{\|X\|^۲}{S^۲})S^۲}{\|X\|^۲}X\right)^T\gamma(X)\\ &+ ۲\frac{(U^TU)^۲}{k+۲}\nabla\cdot\left(-\frac{c\, r(\frac{\|X\|^۲}{S^۲})S^۲}{\|X\|^۲}X\right)^T\Bigg]. \qquad (۷.۳)\end{aligned}$$

حال با شرطی کردن روی ($S^۲ = s^۲$) و با استفاده از نتیجه ۵۲.۲.۱ و با توجه به این‌که r نامنفی، مقعر و مشتق‌پذیر است، بنابه لم ۵۵.۲.۱، $r'(\cdot) \ge ۰$ در نتیجه تفاضل مخاطره (۷.۳) برابر است

با

$$\begin{aligned}
\Delta\mathscr{R} &= \frac{1}{\sigma^2}E_{S^2}\Bigg(E_\theta\Bigg[\|g(X,S)\|^2(U^TU)^2+\frac{2(U^TU)^2}{k+2}\nabla\cdot g(X,S)\\
&\quad+2U^TU(g(X,S))^T\gamma(X)\Bigg]\Bigg/S=s\Bigg)\\
&=\frac{1}{\sigma^2}E_{S^2}\Bigg(E_\theta\Bigg[(U^TU)^2c^2\frac{r^2(\frac{\|X\|^2}{S^2})S^4}{\|X\|^4}\|X\|^2+2c\frac{(U^TU)r(\frac{\|X\|^2}{S^2})S^2}{\|X\|^2}\sum_{i=1}^{p}X_i^2I_{[X_i\le 0]}\\
&\quad-2c\frac{(U^TU)^2}{k+2}\Bigg\{\sum_{i=1}^{p}\Bigg(\frac{\Big(2S^2\frac{X_i^2}{S^2}r'(\frac{\|X\|^2}{S^2})+pS^2r(\frac{\|X\|^2}{S^2})\Big)\|X\|^2}{\|X\|^4}\\
&\quad-\frac{2S^2X_i^2r(\frac{\|X\|^2}{S^2})}{\|X\|^4}\Bigg)\Bigg\}\Bigg]\Bigg/S=s\Bigg)\\
&=\frac{1}{\sigma^2}E_{S^2}\Bigg(E_\theta\Bigg[(U^TU)^2c^2\frac{r^2(\frac{\|X\|^2}{S^2})S^4}{\|X\|^2}+2c\frac{(U^TU)r(\frac{\|X\|^2}{S^2})S^2}{\|X\|^2}\sum_{i=1}^{p}X_i^2I_{[X_i\le 0]}\\
&\quad-2c\frac{(U^TU)^2}{k+2}\Bigg(\frac{2r'(\frac{\|X\|^2}{S^2})\|X\|^4+pS^2r(\frac{\|X\|^2}{S^2})\|X\|^2-2S^2r(\frac{\|X\|^2}{S^2})\|X\|^2}{\|X\|^4}\Bigg)\Bigg]\Bigg/S=s\Bigg)\\
&=\frac{1}{\sigma^2}E_{S^2}\Bigg(E_\theta\Bigg[(U^TU)^2c^2\frac{r^2(\frac{\|X\|^2}{S^2})S^4}{\|X\|^2}+2c\frac{(U^TU)r(\frac{\|X\|^2}{S^2})S^2}{\|X\|^2}\sum_{i=1}^{p}X_i^2I_{[X_i\le 0]}\\
&\quad-2c\frac{(U^TU)^2}{k+2}\Bigg((p-2)S^2\frac{r(\frac{\|X\|^2}{S^2})}{\|X\|^2}+2r'(\frac{\|X\|^2}{S^2})\Bigg)\Bigg]\Bigg/S=s\Bigg)\\
&=\frac{1}{\sigma^2}E_{S^2}\Bigg(E_\theta\Bigg[(U^TU)^2c^2\frac{r^2(\frac{\|X\|^2}{S^2})S^4}{\|X\|^2}+2c\frac{(U^TU)r(\frac{\|X\|^2}{S^2})S^2}{\|X\|^2}\sum_{i=1}^{p}X_i^2I_{[X_i\le 0]}\\
&\quad-2cS^2\frac{(p-2)(U^TU)^2r(\frac{\|X\|^2}{S^2})}{(k+2)\|X\|^2}-4c\frac{(U^TU)^2r'(\frac{\|X\|^2}{S^2})}{k+2}\Bigg]\Bigg/S=s\Bigg)\\
&\le\frac{1}{\sigma^2}E_{S^2}\Bigg(E_\theta\Bigg[(U^TU)^2\frac{r(\frac{\|X\|^2}{S^2})}{\|X\|^2}c\Bigg(cS^4r(\frac{\|X\|^2}{S^2})-2S^2\frac{p-2}{k+2}\\
&\quad+2S^2\frac{\sum_{i=1}^{p}X_i^2I_{[X_i\le 0]}}{U^TU}\Bigg)\Bigg]\Bigg/S=s\Bigg). \qquad (۸.۳)
\end{aligned}$$

حال بنابه شرایط ۱ و ۲، با استفاده از $1\le r(\frac{\|X\|^2}{S^2})\le 0$ و با قرار دادن مقدار ماکزیمم

$r(\|X\|^۲/S^۲)$ که بنابه فرض برابر با ۱ است، یک کران بالا برای نامساوی (۸.۳) عبارت است از

$$\frac{۱}{\sigma^۲}E_{S^۲}\left(E_\theta\left[(U^TU)^۲\frac{r(\frac{\|X\|^۲}{S^۲})}{\|X\|^۲}c\left(cS^۴-۲S^۲\frac{p-۲}{k+۲}\right.\right.\right.$$
$$\left.\left.\left.+۲S^۲\frac{\sum_{i=۱}^p X_i^۲ I_{[X_i\le ۰]}}{U^TU}\right)\right]\Big/ S=s\right). \qquad (۹.۳)$$

در آخرین امید ریاضی ابتدا برای i معین جمله

$$\left((U^TU)^۲\frac{r(\frac{\|X\|^۲}{S^۲})S^۲}{\|X\|^۲}\right)\left(\frac{\sum_{i=۱}^p X_i^۲ I_{[X_i\le ۰]}}{U^TU}\right),$$

را در نظر می‌گیریم و درباره امید ریاضی شرطی روی (U^TU) و $(X_j-\theta_j)^۲/\sigma^۲$ برای $j\neq i$ بحث می‌کنیم. دقت کنید که با توجه به این‌که r نامنفی، مقعر و مشتق‌پذیر است، بنابه لم ۵۵.۲.۱ ، $r(\frac{\|X\|^۲}{S^۲})S^۲/\|X\|^۲$ در $\|X\|^۲/S^۲$ ناصعودی است. آن‌گاه چون توزیع شرطی X_i به شرط U^TU و $(X_j-\theta_j)^۲/\sigma^۲$ برای $j\neq i$، برای θ_i متقارن و تک‌مدی است، بنابه لم ۵۶.۲.۱ داریم

$$cE\left[U^TU\frac{r(\frac{\|X\|^۲}{S^۲})}{\frac{\|X\|^۲}{S^۲}}X_i^۲ I_{[X_i\le ۰]}\Big/ U^TU=\rho^۲,\frac{(X_j-\theta_j)^۲}{\sigma^۲}=Z_j^۲, j\neq i\right]$$
$$\le \frac{c}{۲}E\left[U^TU\frac{r(\frac{\|X\|^۲}{S^۲})}{\frac{\|X\|^۲}{S^۲}}Z_i^۲\Big/ U^TU=\rho^۲,\frac{(X_j-\theta_j)^۲}{\sigma^۲}=Z_j^۲, j\neq i\right]. \qquad (۱۰.۳)$$

بنابراین تفاضل مخاطره در (۹.۳) با فرض $Z=\sigma^{-۱}(X-\theta)$ و U دارای کران بالای به صورت زیر است.

$$\frac{۱}{\sigma^۲}E_{S^۲}\left(E_\theta\left[(U^TU)^۲\frac{r(\frac{\|X\|^۲}{S^۲})}{\|X\|^۲}c\left(cS^۴-۲S^۲\frac{p-۲}{k+۲}+۲S^۲\frac{\sum_{i=۱}^p X_i^۲ I_{[X_i\le ۰]}}{U^TU}\right)\right]\Big/ S=s\right)$$
$$\le \frac{۱}{\sigma^۲}E_{S^۲}\left(E_\theta\left[(U'U)^۲\frac{r(\frac{\|\sigma Z+\theta\|^۲}{S^۲})}{\|\sigma Z+\theta\|^۲}c\left(cS^۴-۲S^۲\frac{p-۲}{k+۲}+S^۲\frac{Z'Z}{U'U}\right)\right]\Big/ S=s\right). \qquad (۱۱.۳)$$

حال فرض می‌کنیم که $U^TU+Z^TZ=R^۲$ و یک کران بالا برای امید ریاضی شرطی (۱۱.۳) با $R^۲$ مزبور خواهیم یافت. ابتدا دقت کنید که چون r نامنفی، مقعر و دارای مشتق دوم است، با استفاده از لم ۵۵.۲.۱ واضح است که $r(\frac{\|X\|^۲}{S^۲})/\frac{\|X\|^۲}{S^۲}$ برای $p\ge ۴$ زبرهمساز و بنابه تعریف ۲۱.۲.۱ نامثبت است. اینک با شرطی کردن روی $R^۲$، و با قرار دادن $\sigma=۱$ عبارت شرطی (۱۱.۳)

دارای کران بالای زیر است.

$$\begin{aligned}
&cE\left[(R^۲ - Z'Z)E\left[\frac{r(\frac{\|Z+\theta\|^۲}{s^۲})}{\|Z+\theta\|^۲}\Big/ Z'Z, R^۲\right]\left(cs^۴ - ۲s^۲\frac{p-۲}{k+۲} + s^۲\frac{Z'Z}{R^۲ - Z'Z}\right)\Big/ R^۲\right] \\
&\quad \leq cE\left[(R^۲ - Z'Z)E\left[\frac{r(\frac{\|Z+\theta\|^۲}{s^۲})}{\|Z+\theta\|^۲}\Big/ Z'Z, R^۲\right]\Big/ R^۲\right] \\
&\quad \times E\left[\left(cs^۴ - ۲s^۲\frac{p-۲}{k+۲} + s^۲\frac{Z'Z}{R^۲ - Z'Z}\right)\Big/ R^۲\right]. \qquad (۱۲.۳)
\end{aligned}$$

از آن‌جایی که توزیع شرطی Z به شرط $R^۲$ متقارن کروی است، $E\left[r(\frac{\|\sigma Z+\theta\|^۲}{S^۲})/\|\sigma Z+\theta\|^۲\right]$ با استفاده از لم ۳.۲.۲ در Z^TZ برای $Z^TZ < R^۲$ ناصعودی و همچنین $(R^۲ - Z^TZ)^۲$ در Z^TZ ناصعودی و $Z^TZ/(R^۲ - Z^TZ)$ نانزولی است و می‌توان نامساوی (۱۷.۲) را از لم ۵۸.۲.۱ نتیجه گرفت.

در سمت راست (۱۲.۳)، اولین امید ریاضی شرطی نامنفی است. از اینرو تفاضل مخاطره شرطی ($\Delta\mathscr{R}$) به شرط آنکه دومین مقدار امید ریاضی شرطی در سمت راست (۱۲.۳) نامثبت باشد، نامثبت خواهد بود، اما چون با استفاده از لم ۲.۲.۲ توزیع $\xi = Z^TZ/R^۲$ به شرط $R^۲$ دارای توزیع $B(\frac{p}{۲}, \frac{k}{۲})$ است و با استفاده از لم ۴۴.۲.۱، داریم

$$E\left[\frac{Z^TZ}{R^۲ - Z^TZ}\Big/ R^۲\right] = E\left[\frac{\xi}{۱-\xi}\right] = \frac{p}{k-۲}, \qquad (۱۳.۳)$$

از اینرو، تفاضل مخاطره شرطی تنها در صورتی نامثبت خواهد بود که داشته باشیم

$$cE_{\sigma=۱}(S^۴) - ۲E_{\sigma=۱}(S^۲)\frac{(p-۲)}{(k+۲)} + E_{\sigma=۱}(S^۲)E\left[\frac{Z^TZ}{R^۲ - Z^TZ}\right] \leq ۰,$$

سپس

$$cE_{\sigma=۱}(S^۴) - ۲E_{\sigma=۱}(S^۲)\frac{(p-۲)}{(k+۲)} - E_{\sigma=۱}(S^۲)\frac{p}{k-۲} \leq ۰,$$

و

$$c \leq ۲\frac{E_{\sigma=۱}(S^۲)}{E_{\sigma=۱}(S^۴)}\frac{(p-۲)}{(k+۲)} - \frac{E_{\sigma=۱}(S^۲)}{E_{\sigma=۱}(S^۴)}\frac{p}{k-۲},$$

در نتیجه

$$۰ < c \leq ۲\frac{E_{\sigma=۱}(S^۲)}{E_{\sigma=۱}(S^۴)}\frac{(p-۲)}{(k+۲)} - \frac{E_{\sigma=۱}(S^۲)}{E_{\sigma=۱}(S^۴)}\frac{p}{k-۲}.$$

که همان شرط سوم قضیه است. برای این‌که کران بالای c مثبت باشد، داریم

$$۲\frac{E_{\sigma=۱}(S^۲)}{E_{\sigma=۱}(S^۴)}\frac{(p-۲)}{(k+۲)} - \frac{E_{\sigma=۱}(S^۲)}{E_{\sigma=۱}(S^۴)}\frac{p}{k-۲} > ۰,$$

پس از گرفتن مخرج مشترک، داریم

$$\frac{۲(k-۲)(p-۲)E_{\sigma=۱}(S^۲) - p(k+۲)E_{\sigma=۱}(S^۲)}{E_{\sigma=۱}(S^۴)(k+۲)(k-۲)} > ۰,$$

سپس

$$k(p-۴) > ۶p-۸,$$

و

$$k > \frac{۲(p-۴)+۴p}{p-۴},$$

در نتیجه

$$k > ۲ + \frac{۴p}{p-۴}.$$

اثبات کامل است. □

۳.۳ برآورد انقباضی در حالت با محدودیت جزئی

در این بخش علاقه‌مندیم پارامتر مجهول θ را تحت محدودیت جزئی برآورد کنیم. همانند فصل دوم منظور از محدودیت جزئی این است که تنها زیر مجموعه‌ای از θ_i ها دارای محدودیت نامنفی بودن باشند. به عبارت دیگر $\theta_۱ \geq ۰, \theta_۲ \geq ۰, \ldots, \theta_q \geq ۰$ و $\theta_{q+۱}, \theta_{q+۲}, \ldots, \theta_p$ هیچ محدودیتی نداشته باشند. آن‌گاه مشابه فودینیر و همکاران (۲۰۰۳a) برآوردگر طبیعی را به صورت

$$\delta_۰^q(X) = X + \gamma_q(X), \tag{۱۴.۳}$$

در نظر می‌گیریم، که در آن برای $۱ \leq j \leq q$،

$$\gamma_{q,j}(X) = \begin{cases} -X_j, & X_j < ۰, \\ ۰, & X_j \geq ۰. \end{cases} \tag{۱۵.۳}$$

و برای $j > q$،

$$\gamma_{q,j}(X) = ۰.$$

در این صورت برآوردگر انقباضی را به صورت زیر در نظر می‌گیریم.

$$\delta^q(X,U) = X + \gamma_q(X) + U^TUg(X,S). \tag{۱۶.۳}$$

در این بخش هدف اثبات برتری برآوردگر انقباضی بر برآوردگر طبیعی است.

قضیه زیر را در خصوص برتری برآوردگر انقباضی بر برآوردگر طبیعی بیان و اثبات می‌نماییم. مجدداً مانند بخش قبل پیش از آن که اثبات را آغاز نماییم برای تسهیل در روند اثبات تفاضل مخاطره $\Delta\mathscr{R}$ را در زیر محاسبه می‌نماییم. محاسبات زیر با تفاوتی جزئی کاملاً مشابه محاسبات بخش قبل است.

$$\begin{aligned}\Delta\mathscr{R} &= \frac{1}{\sigma^2}\mathscr{R}(\theta,\delta^q) - \mathscr{R}(\theta,\delta^q_0)\\ &= \frac{1}{\sigma^2}E_\theta[\|\delta^s-\theta\|^2 - \|\delta^s_0-\theta\|^2]\\ &= \frac{1}{\sigma^2}E_\theta[\|g(X,S)\|^2(U^TU)^2 + 2U^TU(g(X,S))^T(X-\theta)\\ &+ 2U^TU(g(X,S))^T\gamma_q(X)].\end{aligned} \tag{۱۷.۳}$$

قضیه ۱.۳.۳. *** تحت مفروضات این بخش، فرض کنید $p>4$ ، $p \geq q$ ، $k > 2 + \frac{4q}{2p-q-4}$ و توزیع نیز تک‌مدی باشد. برآوردگر انقباضی به شکل*

$$\delta^q(X,U) = X + \gamma_q(X) + U^TUg(X,S)$$

را در نظر بگیرید که در آن تابع g به صورت (۵.۳) است. اگر تابع r و ثابت c در شرایط زیر صدق کنند.

۱. $0 < r(\frac{\|X\|^2}{S^2}) \leq 1$،

۲. $r(\frac{\|X\|^2}{S^2})$ *دارای مشتق دوم و مقعر باشد،*

۳. $0 < c \leq 2\frac{(p-2)}{k+2}\frac{E_{\sigma=1}(s^2)}{E_{\sigma=1}(s^4)} - \frac{q}{k-2}\frac{E_{\sigma=1}(s^2)}{E_{\sigma=1}(s^4)}$،

آن‌گاه به ازای $\theta_1 \geq 0,\ \theta_2 \geq 0,\ \ldots,\ \theta_q \geq 0$، برآوردگر انقباضی $\delta^q(X,U)$ بر برآوردگر طبیعی $\delta^q_0(X) = X + \gamma_q(X)$ برتری دارد.

برهان. با استفاده از اثبات قضیه ۱.۲.۳، بدون آن که تغییر اساسی در آن ایجاد نماییم اثبات را آغاز می‌کنیم. با توجه به این‌که r نامنفی، مقعر و مشتق‌پذیر است، بنابه لم ۵۵.۲.۱، $r'(\cdot) \geq 0$. در نتیجه تفاضل مخاطره (۱۷.۳)، پس از جاگذاری مقدار g و با شرطی کردن روی $(S^2 = s^2)$ و با استفاده از نتیجه ۵۲.۲.۱، مشابه با تفاضل مخاطره (۸.۳) برابر است با

$$
\begin{aligned}
\Delta\mathscr{R} &= \frac{1}{\sigma^2} E_{S^2}\Bigg(E_\theta\Bigg[\|g(X,S)\|^2 (U^TU)^2 + \frac{2(U^TU)^2}{k+2} \nabla\cdot g(X,S) \\
&\quad + 2U^TU(g(X,S))^T\gamma(X) \Bigg] \Bigg/ S=s \Bigg) \\
&= \frac{1}{\sigma^2} E_{S^2}\Bigg(E_\theta\Bigg[\Big\| -\frac{c\, r(\frac{\|X\|^2}{S^2})S^2}{\|X\|^2} X \Big\|^2 (U^TU)^2 + \frac{2(U^TU)^2}{k+2} \nabla\cdot \Bigg(-\frac{c\, r(\frac{\|X\|^2}{S^2})S^2}{\|X\|^2} X \Bigg) \\
&\quad + 2U^TU\Bigg(-\frac{c\, r(\frac{\|X\|^2}{S^2})S^2}{\|X\|^2} X \Bigg)^T \gamma(X) \Bigg] \Bigg/ S=s \Bigg) \\
&= \frac{1}{\sigma^2} E_{S^2}\Bigg(E_\theta\Bigg[(U^TU)^2 c^2 \frac{r^2(\frac{\|X\|^2}{S^2})S^4}{\|X\|^4}\|X\|^2 + 2c\frac{(U^TU)r(\frac{\|X\|^2}{S^2})S^2}{\|X\|^2} \sum_{i=1}^{q} X_i^2 I_{[X_i\leq 0]} \\
&\quad - 2S^2 c\frac{(p-2)(U^TU)^2 r(\frac{\|X\|^2}{S^2})}{(k+2)\|X\|^2} - 4c\frac{(U^TU)^2 r'(\frac{\|X\|^2}{S^2})}{k+2} \Bigg] \Bigg/ S=s \Bigg) \\
&\leq \frac{1}{\sigma^2} E_{S^2}\Bigg(E_\theta\Bigg[(U^TU)^2 \frac{r(\frac{\|X\|^2}{S^2})}{\|X\|^2} c\Bigg(cS^4 r(\frac{\|X\|^2}{S^2}) - 2S^2\frac{p-2}{k+2} \\
&\quad + 2S^2 \frac{\sum_{i=1}^{q} X_i^2 I_{[X_i\leq 0]}}{U^TU} \Bigg) \Bigg] \Bigg/ S=s \Bigg).
\end{aligned}
\tag{18.3}
$$

اینک با استفاده از $1 \leq r(\frac{\|X\|^2}{S^2}) < 0$، با قرار دادن مقدار ماکزیمم $r(\|X\|^2/S^2)$ که برابر با ۱ است و مشابه نامساوی (۱۴.۲) یک کران بالا برای عبارت (۱۸.۳) عبارت است از

$$
\begin{aligned}
&\frac{1}{\sigma^2} E_{S^2}\Bigg(E_\theta\Bigg[(U'U)^2 \frac{r(\frac{\|X\|^2}{S^2})}{\|X\|^2} c\Bigg(cS^4 - 2S^2\frac{p-2}{k+2} \\
&\qquad + 2S^2 \frac{\sum_{i=1}^{q} X_i^2 I_{[X_i\leq 0]}}{U'U} \Bigg) \Bigg] \Bigg/ S=s \Bigg).
\end{aligned}
\tag{19.3}
$$

در ادامه مانند فصل دوم فرض می‌کنیم

$$
\begin{cases}
V = (Z_1, \ldots, Z_q), \\
\eta = (\theta_1, \ldots, \theta_q), \\
T = (Z_{q+1}, \ldots, Z_p), \\
\mu = (\theta_{q+1}, \ldots, \theta_p),
\end{cases}
$$

که در آن، $Z=\sigma^{-1}(X-\theta)$ و با فرض این‌که

$$\begin{cases} X_1^q=(X_1,\ldots,X_q),\\ X_{q+1}^p=(X_{q+1},\ldots,X_p),\end{cases}$$

داریم

$$V=(V_1,\ldots,V_q)=(\sigma^{-1}(X_1-\theta_1),\ldots,\sigma^{-1}(X_q-\theta_q))=\sigma^{-1}(X_1^q-\eta),$$

و

$$T=(T_{q+1},\ldots,T_p)=(\sigma^{-1}(X_{q+1}-\theta_{q+1}),\ldots,\sigma^{-1}(X_p-\theta_p))=\sigma^{-1}(X_{q+1}^p-\mu).$$

بنابراین می‌توان نوشت.

$$\begin{cases} X_1^q=\sigma V+\eta,\\ X_{q+1}^p=\sigma T+\mu.\end{cases}$$

در نتیجه می‌توان $\|X\|^2$ را به صورت زیر نوشت.

$$\begin{aligned}\|X\|^2&=\|X_1^q\|^2+\|X_{q+1}^p\|^2\\ &=\|\sigma V+\eta\|^2+\|\sigma T+\mu\|^2.\end{aligned}$$

همچنین $W^2=V^TV+U^TU$ در نظر می‌گیریم. در نتیجه با استفاده از لم ۵۶.۲.۱ و با فرض $\sigma=1$ و با استدلال مشابه قضیه قبل، تفاضل مخاطره در (۱۹.۳) دارای کران بالا به صورت زیر است.

$$\begin{aligned}&\frac{1}{\sigma^2}E_{S^2}\left(E_\theta\left[(U'U)^2\frac{r(\frac{\|X\|^2}{S^2})}{\|X\|^2}c\left(cS^4-2S^2\frac{p-2}{k+2}+2S^2\frac{\sum_{i=1}^q X_i^2 I_{[X_i\le 0]}}{U'U}\right)\right]\Big/ S=s\right)\\ &\le\frac{1}{\sigma^2}E_{S^2}\left(E_\theta\left[(W^2-V'V)^2\frac{r(\frac{\|X_1^q\|^2+\|X_{q+1}^p\|^2}{S^2})}{\|X_1^q\|^2+\|X_{q+1}^p\|^2}\right.\right.\\ &\left.\left.\times\, c\left(cS^4-2S^2\frac{p-2}{k+2}+S^2\frac{V'V}{W^2-V'V}\right)\right]\Big/ S=s\right).\end{aligned}\tag{۲۰.۳}$$

ابتدا دقت کنید که با استفاده از لم ۵۵.۲.۱ و شرایط ۱ و ۲،

$$\frac{r\left(\frac{\|X_1^q\|^2+\|X_{q+1}^p\|^2}{S^2}\right)S^2}{\|X_1^q\|^2+\|X_{q+1}^p\|^2}$$

برای $p \geq ۴$ زبرهمساز و در نتیجه در $\|X\|^۲/S^۲$ ناصعودی است. بنابراین با شرطی کردن تفاضل مخاطره (۲۰.۳) بر روی $W^۲$ و T داریم

$$cE\left[(W^۲ - V'V)^۲ \frac{r\left(\frac{\|X_۱^q\|^۲+\|X_{q+۱}^p\|^۲}{s^۲}\right)}{\|X_۱^q\|^۲ + \|X_{q+۱}^p\|^۲}\left(cs^۴ - ۲s^۲\frac{p-۲}{k+۲} + s^۲\frac{V'V}{W^۲ - V'V}\right)\Big/ W^۲, T\right]$$
$$\leq cE\left[(W^۲ - V'V)^۲ \frac{r\left(\frac{\|X_۱^q\|^۲+\|X_{q+۱}^p\|^۲}{s^۲}\right)}{\|X_۱^q\|^۲ + \|X_{q+۱}^p\|^۲}\Big/ W^۲, T\right]$$
$$\times E\left[\left(cs^۴ - ۲s^۲\frac{p-۲}{k+۲} + s^۲\frac{V'V}{W^۲ - V'V}\right)\Big/ W^۲, T\right]. \qquad (۲۱.۳)$$

از آن جایی که توزیع V شرطی شده روی $W^۲$ متقارن کروی است،

$$E\left[\frac{r\left(\frac{\|X_۱^q\|^۲+\|X_{q+۱}^p\|^۲}{s^۲}\right)}{\|X_۱^q\|^۲ + \|X_{q+۱}^p\|^۲}\Big/ W^۲, T\right],$$

با استفاده از لم ۳.۲.۲ در V^TV برای $V^TV < W^۲$ ناصعودی است. همچنین $(W^۲ - V^TV)$ در V^TV ناصعودی و $V^TV/(W^۲ - V^TV)$ نانزولی است، در نتیجه نامساوی (۲۰.۳) با استفاده از لم ۵۸.۲.۱ حاصل می‌شود. از اینرو تنها کافیست تا نشان دهیم دومین امید ریاضی شرطی در (۲۱.۳) نامثبت است. در اینجا با در نظر گرفتن $W^۲$ و T، کسر $V^TV/W^۲$، بنابه لم ۲.۲.۲ دارای توزیع $B(\frac{s}{۲}, \frac{k}{۲})$ است و با استفاده از لم ۴۴.۲.۱ داریم

$$E\Big[\frac{V^TV}{(W^۲ - V^TV)}\Big] = \frac{q}{k-۲}.$$

آن‌گاه تفاضل مخاطره (۲۱.۳) وقتی نامثبت است که

$$cE_{\sigma=۱}(S^۴) - ۲E_{\sigma=۱}(S^۲)\frac{(p-۲)}{(k+۲)} + E_{\sigma=۱}(S^۲)E\left[\frac{V^TV}{(W^۲ - V^TV)}\right] \leq ۰$$

سپس

$$cE_{\sigma=۱}(S^۴) - ۲E_{\sigma=۱}(S^۲)\frac{(p-۲)}{(k+۲)} - E_{\sigma=۱}(S^۲)\frac{q}{k-۲} \leq ۰$$

و

$$c \leq ۲\frac{E_{\sigma=۱}(S^۲)}{E_{\sigma=۱}(S^۴)}\frac{(p-۲)}{(k+۲)} - \frac{E_{\sigma=۱}(S^۲)}{E_{\sigma=۱}(S^۴)}\frac{q}{k-۲}.$$

در نتیجه

$$0 < c \leq 2\frac{E_{\sigma=1}(S^2)}{E_{\sigma=1}(S^4)}\frac{(p-2)}{(k+2)} - \frac{E_{\sigma=1}(S^2)}{E_{\sigma=1}(S^4)}\frac{q}{k-2}.$$

که همان شرط سوم قضیه است. از طرف دیگر کران بالای c وقتی مثبت است که

$$2\frac{E_{\sigma=1}(S^2)}{E_{\sigma=1}(S^4)}\frac{(p-2)}{(k+2)} - \frac{E_{\sigma=1}(S^2)}{E_{\sigma=1}(S^4)}\frac{q}{k-2} > 0$$

پس از مخرج مشترک گرفتن، داریم

$$\frac{2E_{\sigma=1}(S^2)(p-2)(k-2) - E_{\sigma=1}(S^2)q(k+2)}{E_{\sigma=1}(S^4)(k-2)(k+2)} > 0$$

سپس

$$E_{\sigma=1}(S^2)\Big(2pk - 4p - 4k + 8 - qk - 2q\Big) > 0$$

بنابراین

$$k(2p - q - 4) > 4p + 2q - 8$$

در نتیجه

$$k > 2 + \frac{4q}{2p - q - 4}.$$

اثبات کامل است. □

۴.۳ برآوردگر انقباضی محدودشده در توزیع نرمال

فرض کنید بردار تصادفی p بعدی X دارای توزیع $N_p(\theta, \sigma^2 I_p)$ باشد. همچنین فرض کنید پارامتر σ^2 مجهول بوده و S^2 برآوردگر نااریب σ^2 و مستقل از X است.

در این بخش، هدف برآورد بردار $\theta = (\theta_1, \dots, \theta_p)$ تحت تابع زیان مربعی موزون ؟؟، با در نظر گرفتن دو محدودیت کامل و جزئی است.

مانند فصل دوم برآوردگر طبیعی را به صورت

$$\delta_0(X) = X + \gamma(X), \qquad (۲۲.۳)$$

و برآوردگر انقباضی را به شکل

$$\begin{aligned}\delta(X) &= \delta_\circ(X) + g(X,S) \\ &= X + \gamma(X) + g(X,S), \end{aligned} \tag{۲۳.۳}$$

در نظر می‌گیریم که در آن $g:\mathbb{R}^p \to \mathbb{R}^p$ تابعی حقیقی به صورت (۵.۳) و هر مؤلفه $\gamma(X)$ به صورت (۳.۳) است.

تفاضل مخاطره $\Delta\mathscr{R}$ عبارت است از

$$\begin{aligned}\Delta\mathscr{R}(\theta) &= \mathscr{R}(\theta,\delta) - \mathscr{R}(\theta,\delta_\circ) \\ &= E_\theta\left[\frac{\|\delta-\theta\|^۲}{\sigma^۲} - \frac{\|\delta_\circ-\theta\|^۲}{\sigma^۲}\right] \\ &= \frac{۱}{\sigma^۲}E_\theta[\|X+\gamma(X)+g(X,S)-\theta\|^۲ - \|X+\gamma(X)-\theta\|^۲] \\ &= \frac{۱}{\sigma^۲}E_\theta[\|X+\gamma(X)-\theta\|^۲ + \|g(X,S)\|^۲ \\ &+ ۲(g(X,S))^T(X+\gamma(X)-\theta) - \|X+\gamma(X)-\theta\|^۲] \\ &= \frac{۱}{\sigma^۲}E_\theta\Big[\|g(X,S)\|^۲ + ۲(g(X,S))^T(X-\theta) + ۲(g(X,S))^T\gamma(X)\Big]. \end{aligned} \tag{۲۴.۳}$$

لازم به ذکر است برای از بین بردن نقش θ مانند فصل دوم دیگر نمی‌توان مستقیماً از لم (تعمیم لم استاین) ۵۰.۲.۱ استفاده کرد. برای رفع این مشکل مشابه با آرشی و طباطبایی (۲۰۱۰) پس از استفاده از تکنیک امید ریاضی دوگانه و شرطی کردن، از تعمیم لم استاین استفاده خواهیم کرد.

در این قسمت با در نظر گرفتن محدودیت کامل (وقتی که هر مولفه θ_i نامنفی است)، در خصوص برتری برآوردگر انقباضی بر برآوردگر طبیعی، قضیه زیر را داریم.

قضیه ۱.۴.۳. ** *تحت مفروضات این بخش، فرض کنید* $p > ۴$ *و* X *دارای توزیع* $N_p(\theta,\sigma^۲ I_p)$ *با* $\sigma^۲$ *مجهول است. همچنین فرض کنید* $S^۲$ *برآوردگر ناااریب* $\sigma^۲$ *و مستقل از* X *باشد. برآوردگر انقباضی* $\delta(X)$ *به شکل (۲۳.۳) را در نظر بگیرید که در آن تابع* g *به صورت (۵.۳) است. اگر تابع* r *و ثابت* c *در شرایط زیر صدق کنند.*

۱. $\circ \le r(\frac{\|X\|^۲}{S^۲}) \le ۱$ ، ۲. $r(\frac{\|X\|^۲}{S^۲})$ *دارای مشتق دوم و مقعر باشد،*

۳. $0 < c \leq \frac{۲E_{\sigma=۱}(S^۲)(p-۲)}{E_{\sigma=۱}(S^۴)} - \frac{pE_{\sigma=۱}(S^۲)}{E_{\sigma=۱}(S^۴)}$،

آن‌گاه به ازای هر $\theta \in \mathbb{R}_+^p$، *برآوردگر انقباضی* $\delta(X)$ *بر برآوردگر طبیعی* $\delta_\circ(X) = X + \gamma(X)$ *برتری دارد.*

برهان. با استفاده از تفاضل مخاطره (۲۴.۳) با جایگذاری مقدار $g(X,S)$ داریم

$$\begin{aligned}\Delta\mathscr{R} &= E_\theta[\|g(X,S)\|^۲ + ۲(g(X,S))^T(X-\theta) + ۲(g(X,S))^T\gamma(X)]\\ &= E_\theta\Bigg[\| - \frac{cr(\frac{\|X\|^۲}{S^۲})S^۲}{\|X\|^۲}X\|^۲ + ۲\sigma^۲\left(-\frac{cr(\frac{\|X\|^۲}{S^۲})S^۲}{\|X\|^۲}X\right)^T(X-\theta)\\ &+ ۲\left(-\frac{cr(\frac{\|X\|^۲}{S^۲})S^۲}{\|X\|^۲}X\right)^T\gamma(X)\Bigg]. \end{aligned} \tag{۲۵.۳}$$

با توجه به این‌که r نامنفی، مقعر و مشتق‌پذیر است، بنابه لم ۵۵.۲.۱، $r'(\cdot) \geq 0$. در این صورت با شرطی کردن روی $(S^۲ = s^۲)$ و با استفاده از تعمیم لم استاین ۵۰.۲.۱، تفاضل مخاطره (۲۵.۳) برابر است با

$$\begin{aligned}\Delta\mathscr{R} &= \frac{۱}{\sigma^۲}E_\theta[\|g(X,S)\|^۲ + ۲\sigma^۲\nabla\cdot g(X,S) + ۲g(X,S)^T\gamma(X)]\\ &= \frac{۱}{\sigma^۲}E_{S^۲}\Bigg(E_\theta\Bigg[c^۲\frac{r^۲(\frac{\|X\|^۲}{S^۲})S^۴}{\|X\|^۴}\|X\|^۲ + ۲c\frac{r(\frac{\|X\|^۲}{S^۲})S^۲}{\|X\|^۲}\sum_{i=۱}^p X_i^۲ I_{[X_i\leq 0]}\\ &- ۲c\left(\frac{۲r'(\frac{\|X\|^۲}{S^۲})\frac{\|X\|^۴}{S^۴} + pr(\frac{\|X\|^۲}{S^۲})\frac{\|X\|^۲}{S^۲} - ۲r(\frac{\|X\|^۲}{S^۲})\frac{\|X\|^۲}{S^۲}}{\frac{\|X\|^۴}{S^۴}}\right)\Bigg]\Bigg/ S=s\Bigg)\\ &= \frac{۱}{\sigma^۲}E_{S^۲}\Bigg(E_\theta\Bigg[c^۲\frac{r^۲(\frac{\|X\|^۲}{S^۲})S^۴}{\|X\|^۲} + ۲c\frac{r(\frac{\|X\|^۲}{S^۲})S^۲}{\|X\|^۲}\sum_{i=۱}^p X_i^۲ I_{[X_i\leq 0]}\\ &- ۲c\left((p-۲)\frac{r(\frac{\|X\|^۲}{S^۲})S^۲}{\|X\|^۲} + ۲r'(\frac{\|X\|^۲}{S^۲})\right)\Bigg]\Bigg/ S=s\Bigg)\\ &= \frac{۱}{\sigma^۲}E_{S^۲}\Bigg(E_\theta\Bigg[c^۲\frac{r^۲(\frac{\|X\|^۲}{S^۲})S^۴}{\|X\|^۲} + ۲c\frac{r(\frac{\|X\|^۲}{S^۲})S^۲}{\|X\|^۲}\sum_{i=۱}^p X_i^۲ I_{[X_i\leq 0]}\\ &- ۲c(p-۲)\frac{r(\frac{\|X\|^۲}{S^۲})S^۲}{\|X\|^۲} - ۴cr'(\frac{\|X\|^۲}{S^۲})\Bigg]\Bigg/ S=s\Bigg)\end{aligned}$$

$$\leq \frac{1}{\sigma^{2}}E_{S^{2}}\left(E_{\theta}\left[c\frac{r(\frac{\|X\|^{2}}{S^{2}})}{\|X\|^{2}}\left(cr(\frac{\|X\|^{2}}{S^{2}})S^{4}-2S^{2}(p-2)\right.\right.\right.$$
$$\left.\left.\left.+2S^{2}\sum_{i=1}^{p}X_i^{2}I_{[X_i\leq 0]}\right)\right]\Big/ S=s\right). \qquad (۲۶.۳)$$

در نتیجه یک کران بالا برای تفاضل مخاطره (۲۶.۳) عبارت است از

$$\frac{1}{\sigma^{2}}E_{S^{2}}\left(E_{\theta}\left[c\frac{r(\frac{\|X\|^{2}}{S^{2}})}{\|X\|^{2}}\left(cS^{4}-2S^{2}(p-2)+2S^{2}\sum_{i=1}^{p}X_i^{2}I_{[X_i\leq 0]}\right)\right]\Big/ S=s\right). \qquad (۲۷.۳)$$

لازم به ذکر است از این حقیقت که مقدار ماکزیمم $r(\|X\|^{2}/S^{2})$ که بنابه فرض برابر با ۱ است در رابطه (۲۷.۳) استفاده کرده‌ایم. در این قسمت مشابه با استدلال فودینیر و همکاران (۲۰۰۳a)، برای i معین جمله

$$\left(\frac{r(\frac{\|X\|^{2}}{S^{2}})S^{2}}{\|X\|^{2}}\right)\left(\sum_{i=1}^{p}X_i^{2}I_{[X_i\leq 0]}\right),$$

را در نظر می‌گیریم و درباره امید ریاضی شرطی روی $(X_j-\theta_j)^{2}$ برای $j\neq i$ بحث می‌کنیم. دقت کنید با توجه به این‌که بنابه فرض r نامنفی، مقعر و مشتق‌پذیر و استفاده از لم ۵۵.۲.۱، $r(\frac{\|X\|^{2}}{S^{2}})S^{2}/\|X\|^{2}$ در $\|X\|^{2}/S^{2}$ ناصعودی است. آن‌گاه چون توزیع شرطی X_i بر روی $(X_j-\theta_j)^{2}/\sigma^{2}$ برای $j\neq i$، متقارن و تک‌مدی برای θ_i است، با در نظر گرفتن $Z=\sigma^{-1}(X-\theta)$ و با استفاده از لم ۵۶.۲.۱ داریم

$$cE\left[\frac{r(\frac{\|X\|^{2}}{s^{2}})}{\frac{\|X\|^{2}}{s^{2}}}\frac{X_i^{2}}{\sigma^{2}}I_{[X_i\leq 0]}\Big/\frac{(X_j-\theta_j)^{2}}{\sigma^{2}}=Z_j^{2}, j\neq i\right]$$
$$\leq \frac{c}{2}E\left[\frac{r(\frac{\|X\|^{2}}{s^{2}})}{\frac{\|X\|^{2}}{s^{2}}}\frac{Z_i^{2}}{\sigma^{2}}\Big/\frac{(X_j-\theta_j)^{2}}{\sigma^{2}}=Z_j^{2}, j\neq i\right].$$

بنابراین تفاضل مخاطره در (۲۷.۳) با فرض $Z=\sigma^{-1}(X-\theta)$ دارای کران بالای به صورت زیر می‌باشد.

$$\frac{1}{\sigma^{2}}E_{S^{2}}\left(E_{\theta}\left[c\frac{r(\frac{\|X\|^{2}}{S^{2}})}{\|X\|^{2}}\left(cS^{4}-2S^{2}(p-2)+2S^{2}\sum_{i=1}^{p}X_i^{2}I_{[X_i\leq 0]}\right)\right]\Big/ S=s\right)$$
$$\leq \frac{1}{\sigma^{2}}E_{S^{2}}\left(E_{\theta}\left[\frac{r(\frac{\|\sigma Z+\theta\|^{2}}{S^{2}})}{\|\sigma Z+\theta\|^{2}}c\left(cS^{4}-2S^{2}\frac{p-2}{k+2}+S^{2}Z^{T}Z\right)\right]\Big/ S=s\right). \qquad (۲۸.۳)$$

با قرار دادن $\sigma = 1$ و شرطی کردن روی Z^TZ، عبارت شرطی (۲۸.۳) دارای کران بالای زیر است.

$$\begin{aligned} cE&\left[E\left[\frac{r(\frac{\|Z+\theta\|^2}{s^2})}{\|Z+\theta\|^2}\Big/ Z^TZ\right]\left(cs^4 - 2s^2\frac{p-2}{k+2} + s^2Z^TZ\right)\Big/ Z^TZ\right] \\ &\leq cE\left[E\left(\frac{r(\frac{\|Z+\theta\|^2}{s^2})}{\|Z+\theta\|^2}\Big/ Z^TZ\right)\right] \\ &\times E\left[cs^4 - 2s^2(p-2) + s^2Z^TZ\Big/ Z^TZ\right]. \end{aligned} \tag{۲۹.۳}$$

دقت کنید که با استفاده از لم ۵۵.۲.۱ و فرضیات داریم که $r(\|X\|^2/s^2)s^2/\|X\|^2$ برای $p \geq 4$ زبرهمساز و در $\|X\|^2/s^2$ ناصعودی است. از آن جایی که توزیع Z نرمال بوده و لذا توزیع متقارن کروی است، $E[r(\frac{\|Z+\theta\|^2}{s^2})/\|Z+\theta\|^2]$ با استفاده از لم ۳.۲.۲ در Z^TZ ناصعودی می‌باشد. از اینرو تنها کافیست نشان دهیم دومین امید ریاضی شرطی در (۲۹.۳) نامثبت است. از طرف دیگر Z^TZ بنابه نتیجه ۴۱.۲.۱ دارای توزیع χ^2_p است و از اینرو $E[Z^TZ] = p$. در این صورت تفاضل مخاطره شرطی $\Delta\mathscr{R}$ نامثبت است اگر

$$cE_{\sigma=1}(S^4) - 2E_{\sigma=1}(S^2)(p-2) + E_{\sigma=1}(S^2)E\Big[Z^TZ\Big] \leq 0$$

سپس

$$cE_{\sigma=1}(S^4) - 2E_{\sigma=1}(S^2)(p-2) + pE_{\sigma=1}(S^2) \leq 0$$

و

$$c \leq \frac{2E_{\sigma=1}(S^2)(p-2)}{E_{\sigma=1}(S^4)} - \frac{pE_{\sigma=1}(S^2)}{E_{\sigma=1}(S^4)},$$

در نتیجه

$$0 < c \leq \frac{2E_{\sigma=1}(S^2)(p-2)}{E_{\sigma=1}(S^4)} - \frac{pE_{\sigma=1}(S^2)}{E_{\sigma=1}(S^4)}.$$

اثبات کامل است. $\square$

در این قسمت همانند بخش دوم، هدف دستیابی به کلاسی از برآوردگرهای برتر تحت محدودیت جزئی است.

برای حالتی که تنها زیر مجموعه‌ای از $\theta_i \geq 0$، برآوردگر طبیعی را به صورت

$$\delta^q_0(X) = X + \gamma_q(X). \tag{۳۰.۳}$$

در نظر می‌گیریم. که در آن هر مولفه $\gamma(X)$ به شکل (۱۵.۳) می‌باشد. همچنین برآوردگر انقباضی را به صورت زیر در نظر می‌گیریم.

$$\delta^q(X) = X + \gamma_q(X) + g(X, S). \tag{۳۱.۳}$$

مجدداً هدف اثبات برتری برآوردگر انقباضی بر برآوردگر طبیعی تحت تابع زیان مربعی (پایا) است. در این صورت مشابه با (۲۴.۳)، تفاضل مخاطره $\Delta\mathscr{R}$ برابر است با

$$\begin{aligned}\Delta\mathscr{R}(\theta) &= \mathscr{R}(\theta, \delta^q) - \mathscr{R}(\theta, \delta^q_\circ)\\ &= \frac{۱}{\sigma^۲} E_\theta[\|X + \gamma_q(X) + g(X, S) - \theta\|^۲ - \|X + \gamma_q(X) - \theta\|^۲]\\ &= \frac{۱}{\sigma^۲} E_\theta\Big[\|g(X, S)\|^۲ + ۲(g(X, S))^T(X - \theta) + ۲(g(X, S))^T\gamma_q(X)\Big]. \end{aligned} \tag{۳۲.۳}$$

در این حالت قضیه زیر را در خصوص برتری برآوردگر انقباضی بر برآوردگر طبیعی داریم.

قضیه ۲.۴.۳. *** فرض کنید که $۴ \leq q \leq p$ و X دارای توزیع $N_p(\theta, \sigma^۲ I_p)$ با $\sigma^۲$ مجهول است. همچنین فرض کنید $S^۲$ برآوردگر نااریب $\sigma^۲$ و مستقل از X باشد. برآوردگر انقباضی $\delta^q(X)$ به شکل (۳۱.۳) را در نظر بگیرید که در آن تابع g به صورت (۵.۳) است. اگر تابع r و ثابت c در شرایط زیر صدق کنند.*

۱. $\circ < r(\frac{\|X\|^۲}{S^۲}) \leq ۱$ ،

۲. $r(\frac{\|X\|^۲}{S^۲})$ *دارای مشتق دوم و مقعر باشد،*

۳. $\circ < c \leq \frac{۲E_{\sigma=۱}(S^۲)(p-۲)}{E_{\sigma=۱}(S^۴)} - \frac{qE_{\sigma=۱}(S^۲)}{E_{\sigma=۱}(S^۴)}$ ،

آن‌گاه به ازای هر $\theta \in \mathbb{R}^p_+$، برآوردگر انقباضی $\delta^q(X)$ بر برآوردگر طبیعی $\delta^q_\circ(X) = X + \gamma_q(X)$ برتری دارد.

برهان. با توجه به این‌که r نامنفی، مقعر و مشتق‌پذیر است، بنابه لم ۵۵.۲.۱، $r'(\cdot) \geq \circ$. در این صورت با شرطی کردن تفاضل مخاطره (۳۲.۳) روی $(S^۲ = s^۲)$ و با استفاده از تعمیم لم استاین،

مشابه با قضیه ۱.۴.۳ و رابطه (۲۵.۳)، تفاضل مخاطره (۳۲.۳) برابر است با

$$\begin{aligned}\Delta\mathscr{R} &= \frac{1}{\sigma^2}E_\theta\Big[\|g(X,S)\|^2 + 2(g(X,S))^T(X-\theta) + 2(g(X,S))^T\gamma_q(X)\Big]\\ &= \frac{1}{\sigma^2}E_\theta[\|g(X,S)\|^2 + 2\sigma^2\nabla\cdot g(X,S) + 2g(X,S)^T\gamma(X)]\\ &= \frac{1}{\sigma^2}E_{S^2}\Bigg(E_\theta\Bigg[c^2\frac{r^2(\frac{\|X\|^2}{S^2})S^4}{\|X\|^2} + 2c\frac{r(\frac{\|X\|^2}{S^2})S^2}{\|X\|^2}\sum_{i=1}^{q}X_i^2 I_{[X_i\le 0]}\\ &\quad - 2c(p-2)\frac{r(\frac{\|X\|^2}{S^2})S^2}{\|X\|^2} - 4cr'\Big(\frac{\|X\|^2}{S^2}\Big)\Bigg]\Bigg/ S=s\Bigg)\\ &\le \frac{1}{\sigma^2}E_{S^2}\Bigg(E_\theta\Bigg[c\frac{r(\frac{\|X\|^2}{S^2})}{\|X\|^2}\Bigg(cr\Big(\frac{\|X\|^2}{S^2}\Big)S^4 - 2S^2(p-2)\\ &\quad + 2S^2\sum_{i=1}^{q}X_i^2 I_{[X_i\le 0]}\Bigg)\Bigg]\Bigg/ S=s\Bigg). \qquad (۳۳.۳)\end{aligned}$$

در نتیجه با استفاده از این حقیقت که مقدار ماکزیمم $r(\|X\|^2/S^2)$ که بنابه فرض برابر با ۱ است، یک کران بالا برای عبارت (۳۳.۳) عبارت است از

$$\begin{aligned}&\frac{1}{\sigma^2}E_{S^2}\Bigg(E_\theta\Bigg[c\frac{r(\frac{\|X\|^2}{S^2})}{\|X\|^2}\Bigg(cr(\frac{\|X\|^2}{S^2})S^4 - 2S^2(p-2) + 2S^2\sum_{i=1}^{q}X_i^2 I_{[X_i\le 0]}\Bigg)\Bigg]\Bigg/ S=s\Bigg)\\ &\le \frac{1}{\sigma^2}E_{S^2}\Bigg(E_\theta\Big[c\frac{r(\frac{\|X\|^2}{S^2})}{\|X\|^2}\Big(cS^4 - 2S^2(p-2) + 2S^2\sum_{i=1}^{q}X_i^2 I_{[X_i\le 0]}\Big)\Big]\Bigg/ S=s\Bigg). \qquad (۳۴.۳)\end{aligned}$$

در ادامه مانند قضیه ۱.۳.۳، برای اثبات فرض می‌کنیم

$$\begin{cases} V = (Z_1,\dots,Z_q),\\ \eta = (\theta_1,\dots,\theta_q),\\ T = (Z_{q+1},\dots,Z_p),\\ \mu = (\theta_{q+1},\dots,\theta_p).\end{cases}$$

که در آن، $Z = \sigma^{-1}(X-\theta)$ و با فرض این‌که

$$\begin{cases} X_1^q = (X_1,\dots,X_q),\\ X_{q+1}^p = (X_{q+1},\dots,X_p).\end{cases}$$

داریم

$$V = \sigma^{-1}(X_1^q - \eta), T = \sigma^{-1}(X_{q+1}^p - \mu).$$

بنابراین می‌توان نوشت.

$$\begin{cases} X_1^q = \sigma V + \eta,\\ X_{q+1}^p = \sigma T + \mu.\end{cases}$$

در نتیجه می‌توان $\|X\|^{۲}$ را به صورت زیر نیز نوشت.

$$\|X\|^{۲}=\|X_{۱}^{q}\|^{۲}+\|X_{q+۱}^{p}\|^{۲}.$$

حال با استفاده از لم ۵۶.۲.۱ و با فرض $\sigma=۱$، می‌توان تفاضل مخاطره (۳۴.۳) را به صورت زیر کراندار نمود.

$$\begin{aligned}\frac{۱}{\sigma^{۲}}E_{S^{۲}}&\left(E_{\theta}\Big[c\frac{r(\frac{\|X\|^{۲}}{S^{۲}})}{\|X\|^{۲}}\Big(cS^{۴}-۲S^{۲}(p-۲)+۲S^{۲}\sum_{i=۱}^{q}X_i^{۲}I_{[X_i\leq ۰]}\Big)\Big]\Big/S=s\right)\\ &\leq\frac{۱}{\sigma^{۲}}E_{S^{۲}}\left(E_{\theta}\left[\frac{r(\frac{\|X_{۱}^{q}\|^{۲}+\|X_{q+۱}^{p}\|^{۲}}{S^{۲}})}{\|X_{۱}^{q}\|^{۲}+\|X_{q+۱}^{p}\|^{۲}}c\right.\right.\\ &\qquad\left.\left.\times\left(cS^{۴}-۲S^{۲}(p-۲)+S^{۲}V^TV\right)\right]\Big/S=s\right).\end{aligned}\qquad(۳۵.۳)$$

با شرطی کردن رابطه (۳۵.۳) بر حسب V^TV و T داریم

$$\begin{aligned}cE_{\theta}&\left[\frac{r(\frac{\|X_{۱}^{q}\|^{۲}+\|X_{q+۱}^{p}\|^{۲}}{s^{۲}})}{\|X_{۱}^{q}\|^{۲}+\|X_{q+۱}^{p}\|^{۲}}\left(cs^{۴}-۲s^{۲}(p-۲)+s^{۲}V^TV\right)\Big/V^TV,T\right]\\ &\leq cE_{\theta}\left[\frac{r(\frac{\|X_{۱}^{q}\|^{۲}+\|X_{q+۱}^{p}\|^{۲}}{s^{۲}})}{\|X_{۱}^{q}\|^{۲}+\|X_{q+۱}^{p}\|^{۲}}\Big/V^TV,T\right]\\ &\quad\times E_{\theta}\left[\left(cs^{۴}-۲s^{۲}(p-۲)+s^{۲}V^TV\right)\Big/V^TV,T\right].\end{aligned}\qquad(۳۶.۳)$$

دقت کنید که با استفاده از لم ۵۵.۲.۱ و فرضیات داریم که $r(\|X_{۱}^{q}\|^{۲}+\|X_{q+۱}^{p}\|^{۲}/s^{۲})$ برای $p\geq ۴$ زبرهمساز و در نتیجه در $\|X\|^{۲}/s^{۲}$ ناصعودی است. حال با در نظر گرفتن V^TV، چون توزیع V حالتی از توزیع متقارن کروی است، امید ریاضی

$$E\left[\frac{r\left(\frac{\|X_{۱}^{q}\|^{۲}+\|X_{q+۱}^{p}\|^{۲}}{s^{۲}}\right)}{\|X_{۱}^{q}\|^{۲}+\|X_{q+۱}^{p}\|^{۲}}\Big/V^TV,T\right].$$

با استفاده از لم ۳.۲.۲ در V^TV ناصعودی است. از اینرو تنها کافیست نشان دهیم دومین امید ریاضی شرطی در (۳۵.۳) نامثبت است. V^TV بنابه نتیجه ۴۱.۲.۱ دارای توزیع $\chi_p^{۲}$ است و از اینرو $E[V^TV]=q$. در این صورت تفاضل مخاطره شرطی $\Delta\mathscr{R}$ نامثبت است اگر

$$cE_{\sigma=۱}(S^{۴})-۲E_{\sigma=۱}(S^{۲})(p-۲)+E_{\sigma=۱}(S^{۲})E\Big[V^TV\Big]\leq ۰$$

سپس

$$cE_{\sigma=1}(S^4) - 2E_{\sigma=1}(S^2)(p-2) + qE_{\sigma=1}(S^2) \leq 0$$

و

$$c \leq \frac{2E_{\sigma=1}(S^2)(p-2)}{E_{\sigma=1}(S^4)} - \frac{qE_{\sigma=1}(S^2)}{E_{\sigma=1}(S^4)},$$

در نتیجه

$$0 < c \leq \frac{2E_{\sigma=1}(S^2)(p-2)}{E_{\sigma=1}(S^4)} - \frac{qE_{\sigma=1}(S^2)}{E_{\sigma=1}(S^4)}.$$

اثبات کامل است. $\square$

فصل ۴
مطالعه شبیه‌سازی

در این فصل، با استفاده از شبیه‌سازی مونت کارلو[۱]، توابع مخاطره برآوردگرهای ارائه شده در فصل‌های قبل را به ازای مقادیر مختلف $\|\theta\|$ محاسبه نموده و مورد بررسی قرار می‌دهیم. همچنین با استفاده از افزایش بعد فضای پارامتر و تغییر مقدار c و $\sigma^۲$ به بررسی چگونگی برتری برآوردگر انقباضی نسبت به برآوردگر طبیعی خواهیم پرداخت. مطالعه شبیه‌سازی در این فصل مربوط به حالت محدودیت کلی در فصل دوم برای توزیع نرمال p - متغیره است. برای آگاهی بیشتر در خصوص شبیه‌سازی مونت کارلو می‌توان به نوروزی‌راد (۱۳۹۰) مراجعه کرد.
در بخش دوم با ارائه یک مثال واقعی به چگونگی برتری برآوردگر انقباضی بر برآوردگر طبیعی خواهیم پرداخت. تمام دستورهای مورد نیاز برای رسم شکل‌ها و جداول با نرم‌افزار R-2.15.3 انجام شده که در پیوست **پ** آمده است.

۱.۴ شبیه‌سازی

توزیع مورد بررسی در این فصل، توزیع $N_p(\theta, \sigma^۲ I)$ با $\sigma^۲$ معلوم است. مقدار معلوم $\sigma^۲$ را در محاسبات به ازای مقادیر مختلف جایگزین نموده و پس از محاسبه مخاطره‌ها، با محاسبه تفاضل میانگین آن‌ها در جداول، برآوردگر برتر را می‌یابیم. در پایان، با رسم نمودار مخاطره‌ها به شناسایی برآوردگر برتر خواهیم پرداخت.

[۱] Monte Carlo

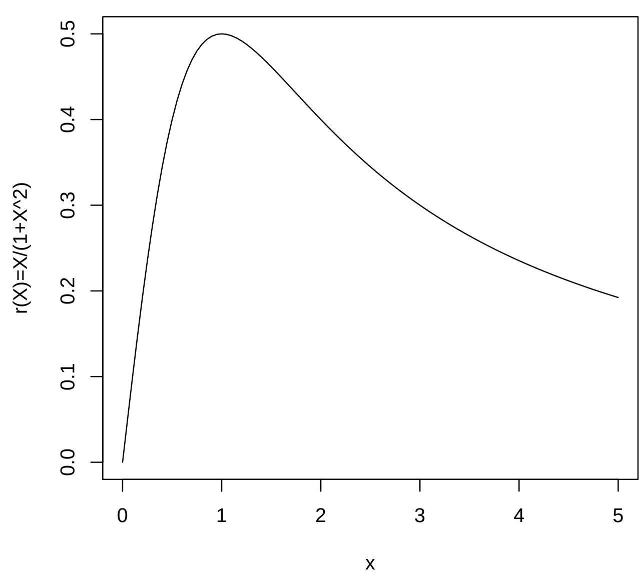
مقدار تابع g را مشابه با قضیه ۱.۵.۲ برابر با $g(X) = \frac{-c\sigma^2 r(\|X\|^2)}{\|X\|^2}X$ و تابع r را با توجه به شرط‌های

۱. $0 \leq r(\|X\|^2) \leq 1$،

۲. $r(\|X\|^2)$ مقعر و دارای مشتق دوم باشد،

برابر با $r(X) = \frac{X}{1+X^2}$ در نظر می‌گیریم. همانطور که در شکل (۱.۴) مشخص است، برد این تابع بین 0 و $\frac{1}{2}$ قرار داشته و در بازه $(0, \sqrt{3})$ مقعر است. از آن‌جایی که دامنه تابع مذکور همواره مثبت است (چون مقدار $\|X\|^2$ مثبت است)، قسمت‌هایی از تابع که در ناحیه منفی محور Xها قرار دارند، هرگز شرایط مسئله را دچار مشکل نخواهند کرد.

در این قسمت با استفاده از برآوردگر طبیعی $\delta_0(X) = X + \gamma(X)$ و برآوردگر انقباضی



شکل ۱.۴: نمودار تابع $r(X) = \frac{X}{1+X^2}$

$\delta(X) = \delta_0(X) + g(X)$، تفاضل مخاطره را به منظور مقایسه، به صورت زیر در نظر می‌گیریم.

$$\Delta\mathscr{R} = \mathscr{R}(\theta, \delta) - \mathscr{R}(\theta, \delta_0). \qquad (1.4)$$

در صورتی که تفاضل مخاطره (۱.۴) مقدار منفی داشته باشد، نتیجه می‌گیریم برآوردگر انقباضی بر برآوردگر طبیعی برتری دارد. این مطلب را تحت تغییر شرایط مختلف از قبیل تغییر مقدار σ^2، p و همچنین بردار θ در جدول‌هایی که در ادامه خواهند آمد، نمایش خواهیم داد.

برای سادگی $\theta = (\theta_1, \ldots, \theta_p)$ را به دو شکل $\theta_1 = (i, 0 \ldots, 0)$ و $\theta_2 = (\frac{i}{\sqrt{p}}, \ldots, \frac{i}{\sqrt{p}})$ در نظر می‌گیریم. در این صورت $\|\theta\|^2 = \theta'\theta = i^2$ و در نتیجه $\|\theta\| = i$. مقادیر مختلف $\|\theta\|$ در ازای i های بین ۰ تا ۵ با گام‌های ۰/۰۰۵ توسط یک دنباله در خطوط برنامه تعریف شده است. در جدول ۱.۴ با تغییر مقادیر p و σ^2 و با ثابت در نظر گرفتن مقدار $c = p - 5$، بر حسب بردار $\theta_1 = (i, 0 \ldots, 0)$ به بررسی تفاضل مخاطره دو برآوردگر می‌پردازیم. همان‌طور که در جدول مشخص است، با افزایش مقادیر p و σ^2 مقدار فاصله تفاضل مخاطرات افزایش می‌یابد.

مشابه با جدول ۱.۴، در جدول ۲.۴ نیز با تغییر مقادیر p و σ^2 و با مقدار ثابت $c = p - 5$، با

جدول ۱.۴: مقادیر تفاضل مخاطره به ازای بردار θ_1 و $c = p - 5$

σ^2	۱	۴	۱۰۰
$p = 15$	۰/۴۰۷۰۵۰	−۰/۶۴۴۶۰۰	−۰/۷۸۰۰۵۱
$p = 50$	−۰/۰۷۴۲۲۷	−۰/۸۸۴۷۱۱	−۰/۹۳۶۶۵۰

در نظر گرفتن بردار $\theta_2 = (\frac{i}{\sqrt{p}}, \ldots, \frac{i}{\sqrt{p}})$ به بررسی تفاضل مخاطره دو برآوردگر می‌پردازیم. در این جدول با افزایش مقادیر p و σ^2 نمی‌توان الگوی خاصی برای افزایش مقدار تفاضل مخاطرات یافت. تنها در ازای $\sigma^2 = 4$ مقادیر بیشینه تفاضل مخاطرات قابل مشاهده هستند. البته در این فصل هدف ما تنها نشان دادن مقدار منفی تفاضل مخاطره، که به معنی برتری برآوردگر انقباضی از برآوردگر طبیعی می‌باشد، بوده است. نمودارهایی که در ادامه می‌آیند برتری دو برآوردگر انقباضی و طبیعی را

جدول ۲.۴: مقادیر تفاضل مخاطره به ازای بردار θ_2 و $c = p - 5$

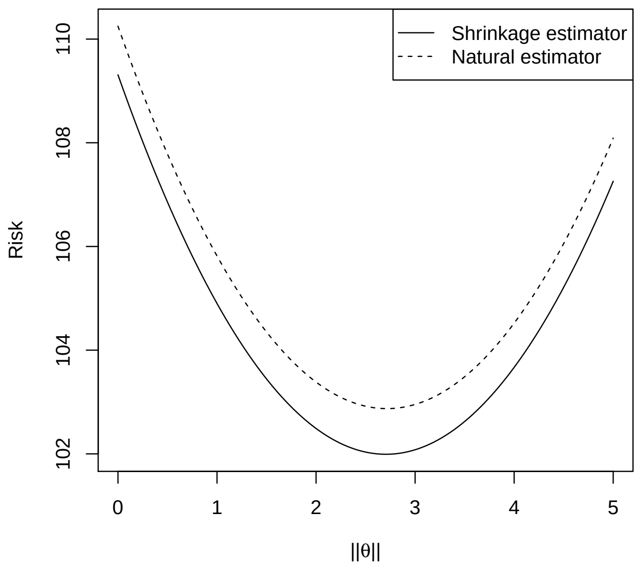
σ^2	۱	۴	۱۰۰
$p = 15$	−۰/۵۵۳۵۵۴	−۰/۸۲۱۵۰۱	−۰/۸۳۵۳۴۹
$p = 50$	−۰/۹۹۳۷۷۹	−۱/۰۶۷۵۸۱	−۰/۹۷۹۵۳۶

مورد مطالعه قرار می‌دهند. این نمودارها، حالات خاصی از اعداد مورد بررسی در جداول هستند

که تنها برای درک نوع برتری به گونه‌ای ملموس، آورده شده‌اند. محور افقی نمودارها مربوط به پارامتر $\|\theta\|$ و محور عمودی مربوط به مخاطره برآوردگر می‌باشد. در حاشیه نمودارها نوع برآوردگر مشخص شده است. در صورتی که یک منحنی در زیر منحنی دیگر قرار بگیرد به این معنی است که مخاطره برآوردگر مربوط به آن منحنی از مخاطره برآوردگر مربوط به منحنی دیگر کمتر بوده و بنابه تعریف آن برآوردگر بر برآوردگر دیگر برتری دارد. مجدداً مقدار ثابت c را برابر با $p-۵$ در نظر می‌گیریم.

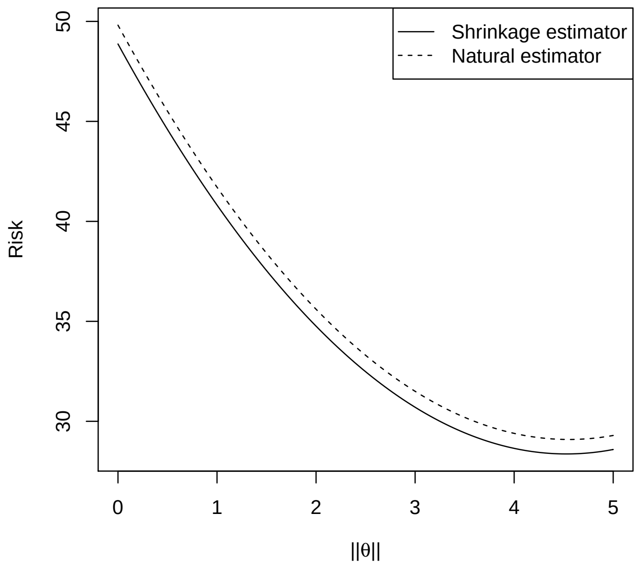
شکل ۲.۴، نمودار مخاطره در برابر نرم بردار $\theta_۱$ تحت حالتی است که در آن $p=۵۰$ و $\sigma^۲=۴$ است. همانطور که در شکل مشخص است منحنی مخاطره برآوردگر انقباضی پایین‌تر از منحنی برآوردگر طبیعی است. این بدین معنی است که برآوردگر انقباضی دارای برتری نسبت به برآوردگر طبیعی است.

در ادامه، نمودارهای چند حالت دیگر در شکل‌های ۳.۴، ۴.۴ و ۵.۴ نمایش داده شده‌اند. در ذیل

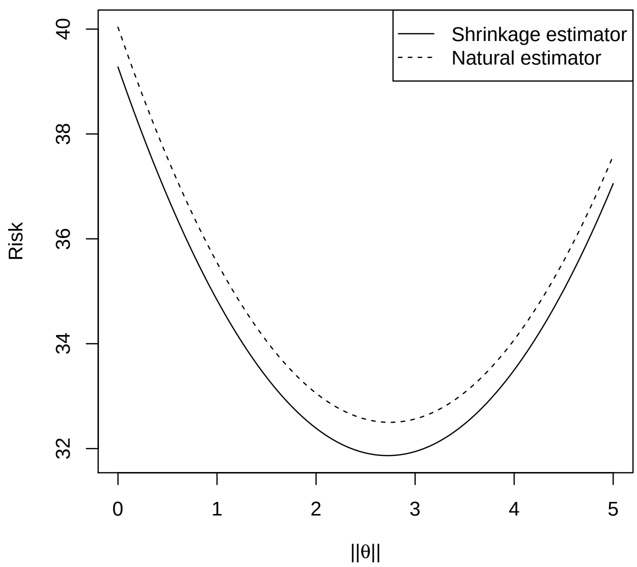


شکل ۲.۴: نمودار مخاطره برای بردار $\theta_۱$ به ازای $p=۵۰$ و $\sigma^۲=۴$

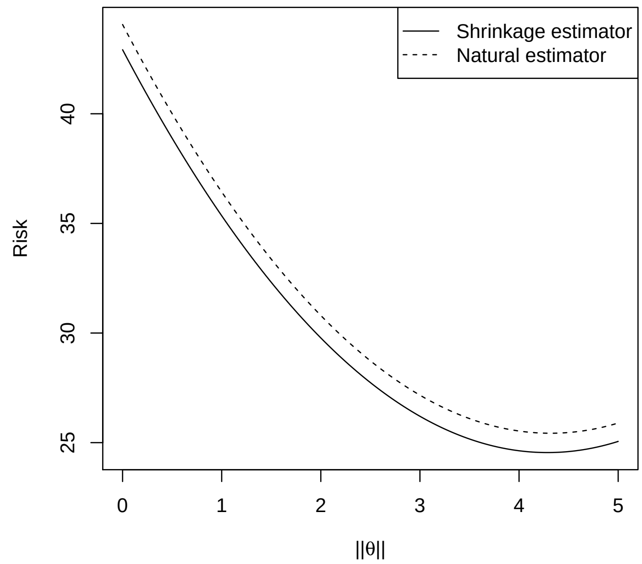
هر نمودار شرایطی را که نمودار تحت آن‌ها رسم شده، آمده است. تمامی نمودارها مؤید برتری برآوردگر انقباضی بر برآوردگر طبیعی هستند.



شکل ۳.۴: نمودار مخاطره برای بردار θ_2 به ازای ۱۵ $= p$ و ۴ $= \sigma^2$



شکل ۴.۴: نمودار مخاطره برای بردار θ_1 به ازای ۱۵ $= p$ و ۱۰۰ $= \sigma^2$



شکل ۵.۴: نمودار مخاطره برای بردار $\theta_۲$ به ازای $p = ۵۰$ و $\sigma^۲ = ۱$

۲.۴ مثال

در این بخش، با ذکر یک مثال واقعی به بررسی مقادیر برآوردشده از طریق برآوردگرهای انقباضی و طبیعی می‌پردازیم. به این صورت که در ازای مقادیر مختلف $\sigma^۲$ و ثابت c، مقادیر برآوردگر انقباضی و برآوردگر طبیعی را بدست می‌آوریم.

این مثال شامل مجموعه داده مربوط به آلودگی هوا در شهرهای ایالات متحده آمریکا در سال ۱۹۸۱ برگرفته از اوریت و هوتورن[۲] (۲۰۱۱) است. متغیرهایی که برای آن‌ها بررسی صورت گرفته است، شامل موارد زیر است.

- **SO2**: گاز SO2 موجود در هوا در مقیاس میکروگرم بر متر مکعب.
- **درجه حرارت**: متوسط درجه حرارت سالانه بر حسب درجه فارنهایت.
- **تولید**: تعداد شرکت‌های تولیدی با تعداد کارگر ۲۰ تن یا بیشتر.
- **جمعیت**: اندازه جمعیت بر حسب هزار بر اساس سرشماری سال ۱۹۷۰.

[۲] Everitt and Hothorn

- **باد**: متوسط سرعت باد سالانه بر حسب مایل بر ساعت.

- **بارش باران**: متوسط بارش سالانه بر حسب اینچ.

- **روز بارانی**: متوسط تعداد روزهای بارانی در هر سال است.

در جدول ۳.۴ مقادیر متغیرهای فوق بر اساس نام شهرها قابل مشاهده است.

برای آن‌که بتوانیم از شرایط حاکم بر مسئله استفاده نماییم، از بردار میانگین نمونه متغیرها به عنوان برآوردگر میانگین جامعه آن‌ها استفاده می‌نماییم. در این صورت بردار حاصل را به عنوان بردار مشاهدات در نظر می‌گیریم. بنابراین بردار $(\bar{X}_۱, \bar{X}_۲, \dots, \bar{X}_۷)$ مربوط به هفت متغیر آلودگی هوا را در محاسبات مورد نظر قرار می‌دهیم.

بردار میانگین نمونه، بر حسب میانگین هر یک از عوامل موثر بر آلودگی هوا بدست آمده است. این مقدار در زیر قابل مشاهده است.

$$\bar{X} = (۲۵٫۷۶, ۵۶٫۲۶, ۳۲۱٫۰۸, ۴۶۹٫۴۶, ۹٫۳۶, ۳۶٫۹۵, ۱۱۲٫۱۴).$$

این بردار را به عنوان بردار مشاهدات در نظر گرفته ($\bar{X} = X$) و با یافتن مقدار $\|X\|^۲ = X'X = ۳۴۱۳۴۰٫۷$ به جستجوی تابع r مناسب می‌پردازیم. فرض کنید

$$r(X) = -\frac{X^۲}{۱۰^{۱۵}}. \qquad (۲.۴)$$

از آن‌جایی که مقدار صورت (۲.۴) با قرار دادن $\|X\|^۲$ به جای X همواره کوچکتر از مخرج خواهد بود، در نتیجه $-۱ < r(X) < ۱$. برای آن‌که شرط $۰ \leq r(\|X\|^۲) \leq ۱$ برقرار باشد، تابع زیر را تعریف می‌کنیم.

$$r^+(X) = \begin{cases} -X, & X < ۰, \\ X, & X \geq ۰. \end{cases}$$

در نتیجه با قرار دادن مقادیر تابع $r(X)$ در تابع $r^+(X)$ داریم

$$۰ < r^+(r(X)) < ۱.$$

جدول ۳.۴: مجموعه داده آلودگی هوا شهرهای ایالات متحده آمریکا برای سال ۱۹۸۱ برگرفته از اوریت و هوتورن (۲۰۱۱)

	SO2	درجه حرارت	تولید	جمعیت	باد	بارش باران	روز بارانی
Albany	۴۶	۴۷٫۶	۴۴	۱۱۶	۸٫۸	۳۳٫۳۶	۱۳۵
Albuquerque	۱۱	۵۶٫۸	۴۶	۲۴۴	۸٫۹	۷٫۷۷	۵۸
Atlanta	۲۴	۶۱٫۵	۳۶۸	۴۹۷	۹٫۱	۴۸٫۳۴	۱۱۵
Baltimore	۴۷	۵۵٫۰	۶۲۵	۹۰۵	۹٫۶	۴۱٫۳۱	۱۱۱
Buffalo	۱۱	۴۷٫۱	۳۹۱	۴۶۳	۱۲٫۴	۳۶٫۱۱	۱۶۶
Charleston	۳۱	۵۵٫۲	۳۵	۷۱	۶٫۵	۴۰٫۷۵	۱۴۸
Chicago	۱۱۰	۵۰٫۶	۳۳۴۴	۳۳۶۹	۱۰٫۴	۳۴٫۴۴	۱۲۲
Cincinnati	۲۳	۵۴٫۰	۴۶۲	۴۵۳	۷٫۱	۳۹٫۰۴	۱۳۲
Cleveland	۶۵	۴۹٫۷	۱۰۰۷	۷۵۱	۱۰٫۹	۳۴٫۹۹	۱۵۵
Columbus	۲۶	۵۱٫۵	۲۶۶	۵۴۰	۸٫۶	۳۷٫۰۱	۱۳۴
Dallas	۹	۶۶٫۲	۶۴۱	۸۴۴	۱۰٫۹	۳۵٫۹۴	۷۸
Denver	۱۷	۵۱٫۹	۴۵۴	۵۱۵	۹٫۰	۱۲٫۹۵	۸۶
Des Moines	۱۷	۴۹٫۰	۱۰۴	۲۰۱	۱۱٫۲	۳۰٫۸۵	۱۰۳
Detroitv	۳۵	۴۹٫۹	۱۰۶۴	۱۵۱۳	۱۰٫۱	۳۰٫۹۶	۱۲۹
Hartford	۵۶	۴۹٫۱	۴۱۲	۱۵۸	۹٫۰	۴۳٫۳۷	۱۲۷
Houston	۱۰	۶۸٫۹	۷۲۱	۱۲۳۳	۱۰٫۸	۴۸٫۱۹	۱۰۳
Indianapolis	۲۸	۵۲٫۳	۳۶۱	۷۴۶	۹٫۷	۳۸٫۷۴	۱۲۱
Jacksonville	۱۴	۶۸٫۴	۱۳۶	۵۲۹	۸٫۸	۵۴٫۴۷	۱۱۶
Kansas City	۱۴	۵۴٫۵	۳۸۱	۵۰۷	۱۰	۳۷	۹۹
Little Rock	۱۳	۶۱٫۰	۹۱	۱۳۲	۸٫۲	۴۸٫۵۲	۱۰۰
Louisville	۳۰	۵۵٫۶	۲۹۱	۵۹۳	۸٫۳	۴۳٫۱۱	۱۲۳
Memphis	۱۰	۶۱٫۶	۳۳۷	۶۲۴	۹٫۲	۴۹٫۱۰	۱۰۵
Miami	۱۰	۷۵٫۵	۲۰۷	۳۳۵	۹	۵۹٫۸۰	۱۲۸
Milwaukee	۱۶	۴۵٫۷	۵۶۹	۷۱۷	۱۱٫۸	۲۹٫۰۷	۱۲۳
Minneapolis	۲۹	۴۳٫۵	۶۹۹	۷۴۴	۱۰٫۶	۲۵٫۹۴	۱۳۷
Nashville	۱۸	۵۹٫۴	۲۷۵	۴۴۸	۷٫۹	۴۶٫۰۰	۱۱۹
New Orleans	۹	۶۸٫۳	۲۰۴	۳۶۱	۸٫۴	۵۶٫۷۷	۱۱۳
Norfolk	۳۱	۵۹٫۳	۹۶	۳۰۸	۱۰٫۶	۴۴٫۶۸	۱۱۶
Omaha	۱۴	۵۱٫۵	۱۸۱	۳۴۷	۱۰٫۹	۳۰٫۱۸	۹۸
Philadelphia	۶۹	۵۴.۶	۱۶۹۲	۱۹۵۰	۹.۶	۳۹.۹۳	۱۱۵
Phoenix	۱۰	۷۰٫۳	۲۱۳	۵۸۲	۶٫۰	۷٫۰۵	۳۶
Pittsburgh	۶۱	۵۰٫۴	۳۴۷	۵۲۰	۹٫۴	۳۶٫۲۲	۱۴۷
Providence	۹۴	۵۰	۳۴۳	۱۷۹	۱۰٫۶	۴۲٫۷۵	۱۲۵
Richmond	۲۶	۵۷٫۸	۱۹۷	۲۹۹	۷٫۶	۴۲٫۵۹	۱۱۵
Salt Lake City	۲۸	۵۱	۱۳۷	۱۷۶	۸٫۷	۱۵٫۱۷	۸۹
San Francisco	۱۲	۵۶٫۷	۴۵۳	۷۱۶	۸٫۷	۲۰٫۶۶	۶۷
Seattle	۲۹	۵۱٫۱	۳۷۹	۵۳۱	۹٫۴	۳۸٫۷۹	۱۶۴
St. Louis	۵۶	۵۵٫۹	۷۷۵	۶۲۲	۹٫۵	۳۵٫۸۹	۱۰۵
Washington	۲۹	۵۷٫۳	۴۳۴	۷۵۷	۹٫۳	۳۸٫۸۹	۱۱۱
Wichita	۸	۵۶٫۶	۱۲۵	۲۷۷	۱۲٫۷	۳۰٫۵۸	۸۲
Wilmington	۳۶	۵۴٫۰	۸۰	۸۰	۹٫۰	۴۰٫۲۵	۱۱۴

جدول ۴.۴: مقادیر برآوردگر طبیعی به ازای $c = ۱$

	برآوردگر طبیعی
$SO۲$	۲۵/۷۵۶۷۵۶۷۵۶۷۵۶۸
درجه حرارت	۵۶/۲۵۶۷۵۶۷۵۶۷۵۶۸
تولید	۳۲۱/۰۸۱۰۸۱۰۸۱۰۸۱
جمعیت	۴۶۹/۴۵۹۴۵۹۴۵۹۴۵۹
باد	۹/۳۵۶۷۵۶۷۵۶۷۵۶۷۶
بارش باران	۳۶/۹۵۱۶۲۱۶۲۱۶۲۱۶
روزبارانی	۱۱۲/۱۳۵۱۳۵۱۳۵۱۳۵

در جدول ۴.۴ مقادیر برآوردگر طبیعی که دارای مقادیر یکسان در ازای تغییر مقادیر مختلف c و $\sigma^۲$ است، محاسبه شده است. علت ثابت بودن مقادیر این برآوردگر، اضافه نشدن مقدار $g(X)$ به آن، است.

در جداول ۵.۴، ۶.۴ و ۷.۴ مقادیر برآوردگرهای انقباضی، به ازای مقادیر مختلف c و $\sigma^۲$ محاسبه شده است. همانطور که در جداول مشخص است در ازای افزایش مقادیر c و $\sigma^۲$، مقادیر برآورد شده برآوردگر انقباضی، کاهش می‌یابد. شایان ذکر است در صورتی که به ازای بردار θ دلخواه مقادیر تفاضل مخاطره را محاسبه می‌نمودیم، این مقادیر بر برتری برآوردگر انقباضی بر برآوردگر طبیعی تاکید داشتند.

در پیوست **پ** دستورهای مورد نیاز برای رسم جداول با نرم‌افزار R-2.15.3 آمده است.

جدول ۵.۴: مقادیر برآوردگر انقباضی به ازای $c = ۱$

	$\sigma^۲ = ۱$	$\sigma^۲ = ۴$	$\sigma^۲ = ۱۰۰$
$SO۲$	۲۵/۷۵۶۷۵۶۷۴۷۹۶۴۹	۲۵/۷۵۶۷۵۶۷۲۱۵۸۹۴	۲۵/۷۵۶۷۵۵۸۷۷۵۷۳۷
درجه حرارت	۵۶/۲۵۶۷۵۶۷۳۷۵۵۴	۵۶/۲۵۶۷۵۶۶۷۹۹۴۵۹	۵۶/۲۵۶۷۵۴۸۳۶۴۸۴۵
تولید	۳۲۱/۰۸۱۰۸۰۹۷۱۴۸۳	۳۲۱/۰۸۱۰۸۰۶۴۲۶۸۹	۳۲۱/۰۸۱۰۷۰۱۲۱۲۷۶
جمعیت	۴۶۹/۴۵۹۴۵۹۲۹۹۲۱۴	۴۶۹/۴۵۹۴۵۸۸۱۸۴۷۷	۴۶۹/۴۵۹۴۴۳۴۳۴۸۹۶
باد	۹/۳۵۶۷۵۶۷۵۳۵۶۲۹۲	۹/۳۵۶۷۵۶۷۴۳۹۸۱۳۹	۹/۳۵۶۷۵۶۴۳۷۳۷۲۵۳
بارش باران	۳۶/۹۵۱۶۲۱۶۰۹۰۰۸۵	۳۶/۹۵۱۶۲۱۵۷۱۱۶۹۲	۳۶/۹۵۱۶۲۰۳۶۰۳۱۲۲
روزبارانی	۱۱۲/۱۳۵۱۳۵۰۹۶۸۵۹	۱۱۲/۱۳۵۱۳۴۹۸۲۰۳	۱۱۲/۱۳۵۱۳۱۳۰۷۵۰۶

جدول ۶.۴: مقادیر برآوردگر انقباضی به ازای $c = 2$

	$\sigma^2 = 1$	$\sigma^2 = 4$	$\sigma^2 = 100$
$SO2$	۲۵٫۷۵۶۷۵۶۷۳۹۱۷۳۱	۲۵٫۷۵۶۷۵۶۶۸۶۴۲۲۱	۲۵٫۷۵۶۷۵۴۹۹۸۳۹۰۷
درجه حرارت	۵۶٫۲۵۶۷۵۶۷۱۸۳۵۱۳	۵۶٫۲۵۶۷۵۶۶۰۳۱۳۵	۵۶٫۲۵۶۷۵۲۹۱۶۲۱۲۲
تولید	۳۲۱٫۰۸۱۰۸۰۸۶۱۸۸۵	۳۲۱٫۰۸۱۰۸۰۲۰۴۲۹۷	۳۲۱٫۰۸۱۰۵۹۱۶۱۴۷
جمعیت	۴۶۹٫۴۵۹۴۵۹۱۳۸۹۶۸	۴۶۹٫۴۵۹۴۵۸۱۷۷۴۹۴	۴۶۹٫۴۵۹۴۲۷۴۱۰۳۳۲
باد	۹٫۳۵۶۷۵۶۷۵۰۳۶۹۰۷	۹٫۳۵۶۷۵۶۷۳۱۲۰۶۰۲	۹٫۳۵۶۷۵۶۱۱۷۹۸۸۳۱
بارش باران	۳۶٫۹۵۱۶۲۱۵۹۶۳۹۵۴	۳۶٫۹۵۱۶۲۱۵۲۰۷۱۶۹	۳۶٫۹۵۱۶۱۹۰۹۹۰۰۲۹
روزبارانی	۱۱۲٫۱۳۵۱۳۵۰۵۸۵۸۳	۱۱۲٫۱۳۵۱۳۴۸۲۸۹۲۵	۱۱۲٫۱۳۵۱۲۷۴۷۹۸۷۷

جدول ۷.۴: مقادیر برآوردگر انقباضی به ازای $c = 3$

	$\sigma^2 = 1$	$\sigma^2 = 4$	$\sigma^2 = 100$
$SO2$	۲۵٫۷۵۶۷۵۶۷۳۰۳۸۱۳	۲۵٫۷۵۶۷۵۶۶۵۱۲۵۴۸	۲۵٫۷۵۶۷۵۴۱۱۹۲۰۷۷
درجه حرارت	۵۶٫۲۵۶۷۵۶۶۹۹۱۴۸۶	۵۶٫۲۵۶۷۵۶۵۲۶۳۲۴۱	۵۶٫۲۵۶۷۵۰۹۹۵۹۳۹۹
تولید	۳۲۱٫۰۸۱۰۸۰۷۵۲۲۸۷	۳۲۱٫۰۸۱۰۷۹۷۶۵۹۰۴	۳۲۱٫۰۸۱۰۴۸۲۰۱۶۶۵
جمعیت	۴۶۹٫۴۵۹۴۵۸۹۷۸۷۲۳	۴۶۹٫۴۵۹۴۵۷۵۳۶۵۱۲	۴۶۹٫۴۵۹۴۱۱۳۸۵۷۶۸
باد	۹٫۳۵۶۷۵۶۷۴۷۱۷۵۲۳	۹٫۳۵۶۷۵۶۷۱۸۴۳۰۶۵	۹٫۳۵۶۷۵۵۷۹۸۶۰۴۰۸
بارش باران	۳۶٫۹۵۱۶۲۱۵۸۳۷۸۲۳	۳۶٫۹۵۱۶۲۱۴۷۰۲۶۴۵	۳۶٫۹۵۱۶۱۷۸۳۷۶۹۳۵
روزبارانی	۱۱۲٫۱۳۵۱۳۵۰۲۰۳۰۶	۱۱۲٫۱۳۵۱۳۴۶۷۵۸۲	۱۱۲٫۱۳۵۱۲۳۶۵۲۲۴۸

پیوست آ
جبر خطی

برای اطلاع دقیقتر از تعاریف و قضایای این فصل می‌توان به میرهد (۱۹۸۲)، رنچر[۱](۲۰۰۲)، ارقامی (۱۳۸۳) و حسن‌زاده (۱۳۸۸) مراجعه کرد.

آ.۱ تعاریف و قضایا

تعریف آ.۱.۱. ماتریس ترانهاده [۲]

اگر جای سطرها یا ستون‌های ماتریس Σ را با ستون یا سطرهای آن تعویض نماییم، ماتریس حاصل ترانهاده ماتریس Σ نامیده می‌شود و با نماد Σ^T نمایش داده می‌شود.

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

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

قضیه آ.۱.۲. *به ازای هر ماتریس Σ داریم،* $\Sigma = {(\Sigma^T)}^T$.

تعریف آ.۱.۳. ماتریس متقارن[۳]

اگر $\Sigma = \Sigma^T$ یا به طور معادل $(\sigma_{ij}) = (\sigma_{ji})$، آن‌گاه ماتریس Σ متقارن است. همچنین ماتریس Σ را پاد - متقارن[۴] گوییم اگر $-\Sigma = \Sigma^T$.

[۱] Rencher
[۲] Transpose matrix
[۳] Symmetric matrix
[۴] Skew-symmetric

قضیه آ.۱.۴. *اگر Σ یک ماتریس $n \times p$ و Λ یک ماتریس $p \times m$ باشند، آن‌گاه* $(\Sigma\Lambda)^T = \Lambda^T\Sigma^T$.

قضیه آ.۱.۵. *اگر Σ یک ماتریس $n \times p$ باشد، آن‌گاه*

۱. *$\Sigma^T\Sigma$ و $\Sigma\Sigma^T$ متقارن هستند.*

۲. *اگر $\Sigma^T\Sigma = 0$ باشد، آن‌گاه $\Sigma = 0$.*

تعریف آ.۱.۶. دترمینان[۵]

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

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

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

قضیه آ.۱.۷. *اگر Σ و Λ دو ماتریس مربعی و هم‌بعد باشند، آن‌گاه* $|\Sigma\Lambda| = |\Sigma||\Lambda|$.

تعریف آ.۱.۸. بردار متعامد[۶]

دو بردار $X_{p\times ۱} = (X_۱, \ldots, X_p)$ و $Y_{p\times ۱} = (Y_۱, \ldots, Y_p)$ را متعامد می‌نامند، هرگاه

$$X^TY = X_۱Y_۱ + X_۲Y_۲ + \cdots + X_pY_p = ۰.$$

تعریف آ.۱.۹. ماتریس متعامد

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

قضیه آ.۱.۱۰. *اگر Σ یک ماتریس متعامد باشد، آن‌گاه* $\Sigma^T\Sigma = \Sigma\Sigma^T = I$.

قضیه آ.۱.۱۱. *اگر $\Sigma_{p\times p}$ یک ماتریس متعامد و $\Lambda_{p\times p}$ هر ماتریس دیگری باشد، آن‌گاه*

۱. $|\Sigma| = \pm ۱$

[۵] Determinant
[۶] Orthogonal vector

۲. $|\Sigma^T \Lambda \Sigma| = |\Lambda|$

۳. $-۱ \leq \sigma_{ij} \leq ۱$ ، *که در آن* σ_{ij} *عنصر ماتریس* Σ *است.*

تعریف آ.۱.۱۲. اگر ماتریس $\Sigma_{p\times p}$ موجود باشد، آن‌گاه مجموع عناصر روی قطر اصلی، اثر[۷] ماتریس Σ نام دارد و به صورت $tr(\Sigma) = \sum_{i=۱}^{p} \sigma_{ii}$ تعریف می‌شود.

تعریف آ.۱.۱۳. ماتریس معین مثبت[۸]

اگر Σ یک ماتریس متقارن با ویژگی $X^T\Sigma X > ۰$ ($X^T\Sigma X \geq ۰$) به ازای هر $X \neq ۰$ باشد، آن‌گاه فرم درجه دوم $X^T\Sigma X > ۰$ معین مثبت[۹] (نیمه معین مثبت[۱۰]) و ماتریس Σ را ماتریس معین مثبت (نیمه معین مثبت[۱۱]) می‌نامیم.

تعریف آ.۱.۱۴. ماتریس نامنفرد (وارون‌پذیر) [۱۲]

ماتریس p بعدی $\Sigma = (\sigma_{ij})$ را نامنفرد می‌نامیم هرگاه $|\Sigma| \neq ۰$ و منفرد[۱۳] (وارون‌ناپذیر) گوییم در صورتی که $|\Sigma| = ۰$.

تعریف آ.۱.۱۵. وارون ماتریس[۱۴]

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

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

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

قضیه آ.۱.۱۶. *فرض کنید* Σ *و* Λ *هم بعد و وارون‌پذیر باشند، در این صورت* $(\Sigma\Lambda)^{-۱} = \Lambda^{-۱}\Sigma^{-۱}$.

[۷] Trace
[۸] Positive definite matrix
[۹] Positive definite
[۱۰] Positive semi definite
[۱۱] Positive semi definite matrix
[۱۲] Nonsingular matrix
[۱۳] singular
[۱۴] Inverse of a matrix

تعریف آ.۱.۱۷. گرادیان[۱۵]

تابع $g : \mathbb{R}^p \to \mathbb{R}$ با ضابطه $\omega = g(X_1, X_2, \ldots, X_p)$ را در نظر بگیرید. هرگاه g دارای مشتقات جزئی باشد، بردار گرادیان آن را به صورت زیر تعریف می‌شود.

$$\nabla\omega = (\frac{\partial g}{\partial X_1}, \frac{\partial g}{\partial X_2}, \ldots, \frac{\partial g}{\partial X_p}).$$

تعریف آ.۱.۱۸. واگرایی[۱۶] (دیورژانس)

بردار $g(X) = g(X_1, X_2, \ldots, X_p)$ را در نظر بگیرید که در آن تابع g مشتق‌پذیر است، در این صورت واگرایی $g(X)$ را به صورت زیر تعریف می‌نماییم.

$$div(X) = \nabla \cdot g(X) = (\frac{\partial g}{\partial X_1} + \frac{\partial g}{\partial X_2} + \cdots + \frac{\partial g}{\partial X_p}) = \sum_{i=1}^{p} \frac{\partial g(X)}{\partial X_i}.$$

تعریف آ.۱.۱۹. عملگر لاپلاس[۱۷] (لاپلاسین)

بردار $g(X) = g(X_1, X_2, \ldots, X_p)$ را در نظر بگیرید که در آن تابع g دارای مشتق دوم است، در این صورت لاپلاسین $g(X)$ را به صورت زیر تعریف می‌نماییم.

$$\Delta g(X) = (\frac{\partial^2 g}{\partial X_1^2} + \frac{\partial^2 g}{\partial X_2^2} + \cdots + \frac{\partial^2 g}{\partial X_p^2}) = \sum_{i=1}^{p} \frac{\partial^2 g(X)}{\partial X_i^2}.$$

[۱۵] Gradient
[۱۶] Divergence
[۱۷] Laplace operator

پیوست ب

پیشنهادات برای آینده تحقیق

با توجه به نتایج بدست آمده در این پایان‌نامه و موضوعات مطرح‌شده می‌توان در آینده بر روی موارد زیر تحقیق کرد.

۱. در صورتی که در توزیع متقارن کروی پارامتر مقیاس را از حالت قطری $\sigma^۲ I_p$ با $\sigma^۲$ مجهول به ماتریس مجهول Σ تغییر دهیم، می‌توان به نتایج کلی‌تری در مسائل کاربردی برای این خانواده توزیع تحت محدودیت‌های مورد مطالعه دست یافت.

۲. با تغییر خانواده توزیع از کروی به بیضی‌گون، تحت محدودیت‌های ارائه شده در این پایان‌نامه می‌توان به نتایج کلی‌تری برای یافتن برآوردگر برتر دست یافت. در این زمینه فودینیر و همکاران (۲۰۰۳b) برای خانواده توزیع‌های بیضی‌گون با ماتریس کوواریانس مجهول در حالت بدون محدودیت موفق به یافتن برآوردگر مینیماکس شدند. در صورت امکان با اعمال محدودیت بر روی فضای پارامتر فودینیر و همکاران (۲۰۰۳b) می‌توان نتایج آن مقاله را برای حالت محدود شده ارتقا داد. همچنین در ادامه می‌توان نتایج آرشی و طباطبایی (۲۰۱۰) را به حالت فضای پارامتر محدود نیز تعمیم داد.

۳. در صورتی که در مسئله برآوردیابی نقطه‌ای ملاک برتری تابع مخاطره باشد، با توجه به وابسته بودن این ملاک به تابع زیان، می‌توان با تعویض این تابع نوع ملاک برای برتری برآوردگر را بر حسب تابع زیان و مخاطره متحول کرد. در این زمینه جوزانی و همکاران (۲۰۰۶)

با در نظر گرفتن تابع زیان وزنی متوازن[1] به جستجوی برآوردگر مینیماکس پرداختند. در این پایان‌نامه نیز می‌توان با استفاده از این تابع زیان در فضای پارامتر محدود شده به دنبال برآوردگر مینیماکس بود.

۴. همانطور که در مقدمه انواع محدودیت‌های در نظر گرفته شده بر روی فضای پارامتر ذکر شده‌اند، در این پایان‌نامه نیز می‌توان انواع محدودیت‌ها را بررسی نمود. چرا که شکل خاص برآوردگر طبیعی و برآوردگر انقباضی در فصل‌های دوم و سوم در هیچ یک از مقالات ارجاع داده شده مورد بررسی قرار نگرفته‌اند.

[1] Weighted balanced loss function

پیوست پ

دستورهای نرم‌افزار R

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

دستور زیر شامل تعریف پارامتر مقیاس و کران محورهای x و y می‌باشد که برای تمام دستورها یکسان است.

```
Sigma <- matrix(c(1,0,0,10),nrow=2)
y <- x <- seq(-3, 3, length= 50)
```

دستور ۱. برای رسم نمودارهای تابع چگالی و منحنی‌های تراز در توزیع نرمال دومتغیره در شکل ۲.۱ از دستورهای زیر در نرم‌افزار R استفاده می‌کنیم.

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

دستور ۲. برای رسم نمودارهای تابع چگالی و منحنی‌های تراز در توزیع t دومتغیره در شکل ۳.۱ از دستورهای زیر در نرم‌افزار R استفاده می‌کنیم.

```
f <- function(x,y) {
z <-  (((det(Sigma)^(-0.5))* gamma(5/2))/((3*pi)*gamma(3/2)))
 *((1+ ((x^2 + 0.1*y^2)/3))^(-5/2))}
z <- outer(x, y, f)
persp(x,y,z,theta=45,phi=30,expand=0.6,ltheta=120,shade=0.75,
ticktype = "detailed", xlab = "X", ylab = "Y", zlab = "f(x, y)")
library(lattice)
contourplot(z)
```

دستور ۳. برای رسم نمودارهای تابع چگالی و منحنی‌های تراز در توزیع کوشی دومتغیره در شکل ۴.۱ از دستورهای زیر در نرم‌افزار R استفاده می‌کنیم.

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

دستور ۴. برای رسم نمودارهای تابع چگالی و منحنی‌های تراز در توزیع لجستیک دومتغیره در شکل ۵.۱ از دستورهای زیر در نرم‌افزار R استفاده می‌کنیم.

```
f <- function(x,y) {
z <- ((det(Sigma)^(-0.5))*1/(2*pi)^(-1)) * (exp(-.5 * (x^2
+ 0.1*y^2))/((1+exp(-.5 * (x^2 + 0.1*y^2)))^2))}
z <- outer(x, y, f)
persp(x,y,z,theta=45,phi=30,expand=0.6,ltheta=120,shade=0.75,
ticktype = "detailed", xlab = "X", ylab = "Y", zlab = "f(x, y)")
library(lattice)
contourplot(z)
```

دستور ۵. برای رسم نمودارهای تابع چگالی و منحنی‌های تراز در توزیع نمایی توانی دومتغیره در شکل ۶.۱ از دستورهای زیر در نرم‌افزار R استفاده می‌کنیم.

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

دستور ۶. برای رسم نمودارهای تابع چگالی و منحنی‌های تراز در توزیع کاتز دومتغیره در شکل ۷.۱ از دستورهای زیر در نرم‌افزار R استفاده می‌کنیم.

```
f <- function(x,y) {
z <- (((det(Sigma)^(-0.5))*gamma(1)*3)/((2*pi)*gamma(1)))
```

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

دستور ۷. برای رسم نمودارهای تابع چگالی و منحنی‌های تراز در توزیع لاپلاس دومتغیره در شکل ۸.۱ از دستورهای زیر در نرم‌افزار R استفاده می‌کنیم.

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

دستور ۸. برای رسم نمودارهای تابع چگالی و منحنی‌های تراز در توزیع نرمال استاندارد دومتغیره در شکل ۹.۱ از دستورهای زیر در نرم‌افزار R استفاده می‌کنیم.

```
f <- function(x,y) {
z <- (1/(2*pi)) * exp(-.5 * (x^2 + y^2))}
z <- outer(x, y, f)
persp(x,y,z,theta=45,phi=30,expand=0.6,ltheta=120,shade=0.75,
```

```
ticktype = "detailed", xlab = "X", ylab = "Y", zlab = "f(x, y)")
library(lattice)
contourplot(z)
```

دستور ۹. برای محاسبه مقادیر بدست آمده در جداول ۱.۴ و ۲.۴ از دستورهای زیر در نرم‌افزار R استفاده می‌کنیم.

```
set.seed(1234)
library(mvtnorm)
p =15 # or 50
bound <- seq(0,5,0.005)
lt <-length(bound)
c=p-5 # and 0<c<=p-4
sigma2 <- 1  # or 4 or 100
sigma <- sigma2*diag(p)
gamma <- function(X){
for (i in 1:p){
if(X[i]<0){
X[i]= -X[i]}
else{X[i]=0}}
return(X)}
r <- function(X){X/(1+X^2)}
curve(r,0,5)
g <- function(X){
H <- as.vector((t(X)%*%X))
```

```
L <- (-c*sigma2*((r(H)/H)*X))
return(L)
}
f=numeric(lt*p)
H<-theta <-matrix(f,lt,p,byrow=T)
a=numeric(lt*lt)
Loss.S <-Loss.N <-matrix(a,lt,lt,byrow=T)
for(i in 1:lt){
thetai <-  c(bound[i],rep(0,p-1))# or thetai <- c(rep((bound[i]/(sqrt(p))),p))
x <- as.vector(rmvnorm(1, mean=thetai, sigma=sigma))
H[i,]<-x
theta[i,] <-thetai}
loss.S=loss.N=numeric(lt)
for(j in 1:lt){
 for(i in 1:lt){
   A  <- as.vector(H[i,] + gamma(H[i,])+ g(H[i,])-theta[j,])
   A1 <- as.vector(H[i,] + gamma(H[i,])-theta[j,])
   loss.S[i] <- as.vector((t(A))%*%(A))
   loss.N[i]<- as.vector((t(A1))%*%(A1))
  }
 Loss.S[j,] <-loss.S
 Loss.N[j,] <-loss.N
}
```

```
Risk.S <- rowMeans(Loss.S)
Risk.N <- rowMeans(Loss.N)
mean(Risk.S)
mean(Risk.N)
DeltaR <- mean(Risk.S)-mean(Risk.N)
```

دستور ۱۰. برای رسم نمودار مخاطره در شکل‌های ۲.۴ تا ۵.۴ از دستورهای زیر در نرم‌افزار R استفاده می‌کنیم.

```
matplot(bound, cbind(Risk.S,Risk.N),type="l",col = c("1","1")
,xlab=expression(paste("||",theta, "||")),ylab="Risk")
#matplot(bound, cbind(Risk.S,Risk.N),type="l", pch=1)
legend("topright", c("Shrinkage estimator", "Natural estimator"),
 col = c("1","1"),lty=c(1,2), merge = TRUE)
plot(bound,Risk.S,col="blue",type="l",ylim=range(c(Risk.S,Risk.N))
,xlab=expression(paste("||",theta, "||")),ylab="Risk")
par(new=T)
plot(bound,Risk.N,col="red",type="l",ylim=range(c(Risk.S,Risk.N)),
 axes = FALSE, xlab = "", ylab = "")
legend("topright", c("Shrinkage estimator", "Natural estimator"),
 col = c("4","2"),lty=c(1,1), merge = TRUE)
```

دستور ۱۱. برای وارد کردن داده‌ها در جدول ۳.۴ از دستورهای زیر در نرم‌افزار R استفاده می‌کنیم.

```
USairpollution = matrix(c(46, 47.6, 44, 116, 8.8, 33.36, 135,
```

```
 11, 56.8, 46, 244, 8.9, 7.77, 58, 24, 61.5, 368, 497, 9.1, 48.34,
   115, 47, 55.0, 625, 905, 9.6, 41.31, 111, 11, 47.1, 391, 463, 12.4,
 36.11, 166, 31, 55.2, 35, 71, 6.5, 40.75, 148,  23, 54.0, 462, 453,
 7.1, 39.04, 132, 26, 51.5, 266, 540, 8.6, 37.01, 134, 9, 66.2, 641,
 844, 10.9, 35.94, 78, 17, 51.9, 454, 515, 9.0, 12.95, 86, 17, 49.0,
 104, 201, 11.2, 30.85, 103, 56, 49.1, 412, 158, 9.0, 43.37, 127, 10,
 68.9, 721, 1233, 10.8, 48.19, 103, 28, 52.3, 361, 746, 9.7, 38.74, 121,
 14, 68.4, 136,529, 8.8, 54.47, 116, 14, 54.5, 381, 507, 10.0, 37.00,
 99, 13, 61.0, 91, 132, 8.2, 48.52, 100, 30, 55.6, 291, 593, 8.3,
 43.11, 123, 10, 61.6, 337, 624, 9.2, 49.10, 105, 10, 75.5, 207, 335,
 9.0, 59.80, 128, 16, 45.7, 569, 717, 11.8, 29.07, 123, 29, 43.5, 699,
 744, 10.6, 25.94, 137, 18, 59.4, 275, 448, 7.9, 46.00, 119, 9, 68.3,
 204, 361, 8.4, 56.77, 113, 31, 59.3, 96, 308, 10.6, 44.68, 116, 14,
 51.5, 181, 347, 10.9, 30.18, 98, 10, 70.3, 213, 582, 6.0, 7.05, 36,
 61, 50.4, 347, 520, 9.4, 36.22, 147, 94, 50.0, 343, 179, 10.6, 42.75,
  125, 26, 57.8, 197, 299, 7.6, 42.59, 115, 28, 51.0, 137, 176, 8.7,
15.17, 89, 12, 56.7, 453, 716, 8.7, 20.66, 67, 29, 51.1, 379, 531,
 9.4, 38.79, 164, 56, 55.9, 775, 622, 9.5, 35.89, 105, 29, 57.3, 434,
 757, 9.3, 38.89, 111, 8, 56.6, 125, 277, 12.7, 30.58, 82, 36, 54.0,
 80, 80, 9.0,  40.25,114),37,byrow=T)
cities = c("Albany", "Albuquerque", "Atlanta", "Baltimore", "Buffalo",
 "Charleston", "Cincinnati", "Columbus", "Dallas", "Denver", "Des Moines",
 "Hartford", "Houston", "Indianapolis", "Jacksonville", "Kansas City",
 "Little Rock", "Louisville","Memphis",
```

```
 "Miami","Milwaukee","Minneapolis","Nashville","New Orleans","Norfolk",
 "Omaha","Phoenix","Pittsburgh","Providence","Richmond",
 "Salt Lake City","San Francisco","Seattle","St. Louis","Washington",
 "Wichita","Wilmington")
variables = c("SO2","temp","manu","popul","wind","precip","predays")
colnames(USairpollution) = variables
rownames(USairpollution)= cities
```

دستور ۱۲. برای محاسبه مقادیر بدست آمده در جداول ۵.۴، ۶.۴ و ۷.۴ از دستورهای زیر در نرم‌افزار R استفاده می‌کنیم.

```
set.seed(1234)
xbar <- colMeans(x)
X=xbar
p=length(X)
c=p-6 # or p-5 or p-4 and 0<c<=p-4
sigma2 <- 1 # or 4 or 100 or 1000
gamma <- function(X){
for (i in 1:p){
if(X[i]<0){
X[i]= -X[i]}
else{X[i]=0}}
return(X)}
r <- function(X){-((X^2)/((10)^15))}
rplus <- function(X){
```

```
if(X<0){
X=-X}
else{X}
return(X)}
g <- function(X){
H <- as.vector((t(X)%*%X))
L <- (-c*sigma2*((rplus(r(H))/H)*X))
return(L)}
delta.S <- X + gamma(X)+ g(X)
delta.N <- X + gamma(X)
as.character(delta.S)
as.character(delta.N)
```

مراجع

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

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

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

[۴] پارسیان ا، (۱۳۸۶)، "**مبانی آمار ریاضی**"، چاپ چهارم، انتشارات دانشگاه اصفهان، اصفهان.

[۵] تجدد س. ع، (۱۳۹۱)، پایان‌نامه کارشناسی‌ارشد، "برآورد انقباضی در مدل‌های چندمتغیره"، دانشکده ریاضی، دانشگاه صنعتی شاهرود.

[۶] حسن‌زاده بشتیان م، (۱۳۸۸)، پایان‌نامه کارشناسی‌ارشد، "برآوردگرهای انقباضی در مدل رگرسیون ریج با خطاهای دارای توزیع بیضی‌گون "،دانشکده علوم ریاضی، دانشگاه فردوسی مشهد.

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

[۸] رویدن اچ. ال، (۱۳۷۴)، "**آنالیز حقیقی**"، ایزد دوستدار ن، چاپ اول، انتشارات دانشگاه تهران، تهران.

[۹] زین‌الدینی ش، (۱۳۸۸)، پایان‌نامه کارشناسی‌ارشد، "مروری بر برآورد بیزی مینیماکس میانگین در توزیع نرمال چندمتغیره آمیخته در مقیاس"، دانشکده ریاضی، آمار و کامپیوتر، دانشگاه تهران.

[۱۰] مدقالچی ع، (۱۳۸۴)، "**آنالیز ریاضی ۱**"، چاپ پنجم، انتشارات دانشگاه پیام نور، تهران.

[۱۱] نوروزی‌راد م، (۱۳۹۰)، پایان‌نامه کارشناسی‌ارشد، " مباحثی در برآورد تابع زیان"، دانشکده علوم پایه، دانشگاه آزاد اسلامی واحد مشهد.

[12] Arashi M. and Tabatabaey S. M. M. (2010), "A note on classical Stein-type estimators in elliptically contoured models", **Journal of Statistical Planning and Inference**, 140, 1206-1213.

[13] Baranchik A. J. (1970), "A Family of Minimax Estimators of the Mean of a Multivariate Normal Distribution", **The Annals of Mathematical Statistics**, 41, 2, 642-645.

[14] Bickel P. J. (1981), "Minimax estimation of the mean of a normal distribution when the parameter space is restricted", **The Annals of Statistics**, 9, 6, 1301-1309.

[15] Boyd S. and Vandenberghe L. (2009), **"Convex Optimization"**, Cambridge University Press, New York.

[16] Casella G. and Strawderman W. E. (1981), "Estimating a bounded normal mean", **The Annals of Statistics**, 9, 870-878.

[17] Du Plessis N. (1970), **"An Introduction to Potential Theory"**, Oliver and Boyd, Edinburgh.

[18] Everitt B. and Hothorn T. (2011), **"An Introduction to Applied Multivariate Analysis with R"**, Springer, New York.

[19] Fang K. T., Kotz S. and Ng K. W. (1990), **"Symmetric Multivariate and Related Distribution"**, Chapman and Hall, London, New York.

[20] Fourdrinier D. and Strawderman W. E. (1996), "A paradox concerning shrinkage estimators: should a known scale parameter be replaced by an estimated value in the shrinkage factor", **Journal of Multivariate Analysis**, 59, 2, 109-140.

[21] Fourdrinier D. and Ouassou I. (2000), "Estimation of the mean of a spherically symmetric distribution with constraints on the norm", **Canadian Journal of Statistics**, 28, 2, 399-415.

[22] Fourdrinier D, Ouassou I. and Strawderman W. E. (2003a), "Estimation of a parameter vector when some components are restricted", **Journal of Multivariate Analysis**, 86, 14-27.

[23] Fourdrinier D, Strawderman W. E. and Wells, M. T. (2003b), "Robust shrinkage estimation for elliptically symmetric distributions with unknown covariance matrix", **Journal of Multivariate Analysis**, 85, 24-39.

[24] Fourdrinier D, Strawderman W. E. and Wells M. T. (2006), "Estimation of a location parameter with restrictions or "vague information" for spherically symmetric distributions", **Annals of the Institute of Statistical Mathematics**, 58, 73-92.

[25] Fourdrinier D. and Marchand E. (2010), "On Bayes estimators with uniform priors on spheres and their comparative performance with maximum likelihood estimators for estimating bounded multivariate normal means", **Journal of Multivariate Analysis**, 101, 1390-1399.

[26] Gatsonis C, MacGibbon B. and Strawderman W. E. (1987), "On the estimation of a restricted normal mean", **Statistics and Probability Letters**, 6, 21-30.

[27] Gut A. (2005), **"Probability: A Graduate Course"**, Springer, New York.

[28] Jafari Jozani M. Marchand E. and Parsian A. (2006), "On estimation with weighted balanced-type loss function", **Statistics and Probability Letters**, 76, 773-780.

[29] Kariya T. (1989), "Equivariant estimation in a model with an ancillary statistics", **The Annals of Statistics**, 17, 920-928.

[30] Kortbi O. and Marchand E. (2012), "Truncated linear estimation of a bounded multivariate normal mean", **Journal of Statistical Planning and Inference**, 142, 2607-2618.

[31] Lehmann E. L. and Casella G. (1998), **"Theory of Point Estimation"**, 2nd Ed., Springer, New York.

[32] Marchand E. (1993), "Estimation of a multivariate normal mean with constraints on the norm", **Canadian Journal of Statistics**, 21, 4, 359-366.

[33] Marchand E. and Giri N. (1993), "James–Stein estimation with constraints on the norm", **Communications in Statistics - Theory and Methods**, 22, 10, 2903-2924.

[34] Marchand E. (1994), "On the estimation of the mean of a $N_p(\mu, \Sigma)$ population with $\mu'\mu$ known", **Statistics and Probability Letters**, 21, 69-75.

[35] Marchand E. and Perron F. (2005), "Improving on the mle of a bounded location parameter for spherical distributions", **Journal of Multivariate Analysis**, 92, 227-238.

[36] Marchand E. and Strawderman W. E. (2005), "Improving on the minimum risk equivariant estimator of a location parameter which is constrained to an interval or a half - interval", **Annals of the Institute of Statistical Mathematics**, 57, 1, 129-143.

[37] Marchand E. and Parsian A. (2006), "Minimax estimation of a bounded parameter of a discrete distribution", **Statistics and Probability Letters**, 76, 547-554.

[38] Marchand E. and Strqwderman W. E. (2012), "A unified minimax result for restricted parameter spaces", **Bernoulli**, 18, 2, 635-643.

[39] Muirhead J. R. (1982), **"Aspects of Multivariate Statistical Theory"**, John Wiley, New York.

[40] Ouassou I. and Strqwderman W. E. (2002), "Estimation of a parameter vector when some components are restricted to a cone", **Statistics and Probability Letters**, 56, 121-129.

[41] Perron F. and Giri N. (1989), "On the best equivariant estimator of the mean of multivariate normal population", **Journal of Multivariate Analysis**, 32, 1-16.

[42] Rencher C. A. (2002), **"Methods of Multivariate Analysis"**, 2nd Ed., John Wiley, New York.

[43] Ross S. (2010), **"A First Course in Probability"**, 8nd Ed., Upper Saddle River, New Jersey.

[44] Serfling R. J. (2006), "Multivariate symmetry and asymmetry", **Encyclopedia of Statistical Sciences**, 2nd Ed., , John Wiley, 8, 5338-5345.

[45] Shao J. (2007), **"Mathematical Statistics"**, 2nd Ed., Springer, New York.

[46] Stein M. C. (1956), "Inadmissibility of the usual estimator for the mean of the multivariate normal distribution", **Proceeding of the Third Berkeley Symposium** , 1, 197-206.

[47] Stein M. C. (1981), "Estimation of the Mean of a Multivariate Normal Distribution" **The Annals of Statistics**, 9, 6, 1135-1151.

[48] Wan A. T. K, Zou G. and Lee A. H. (2000), "Minimax and Γ-minimax estimation for the Poisson distribution under LINEX loss when the parameter space is restricted" **Statistics and Probability Letters**, 50, 23-32.

Surname: Karami Kabir Name: Hamid

Title: Shrinkage Estimation in Restricted Models.

Supervisor: Dr. Mohammad Arashi

Degree: Master of Science Subject: Statistics
Field: Srinkage Estimator

University of Shahrood Faculty Of Mathematical Sciences
Date: May 2013 Number of pages: 121

Keywords: Multivariate normal distribution, Restricted parameter space, Risk, Shrinkage estimator, Spherically symmetric distribution

Abstract

In this thesis we consider the problem of estimating the p-dimensional location parameter in spherically symmetric models. The scale component is assumed to be known and unknown. Three restrictions are imposed on the vector parameter θ. First we asaume that all elements of θ are nonnegative, at the second hand it is assumed that only a subset of elements is nonnegative. Lastly, we deal with the situation where the parameter θ belongs to a cone $\mathscr{C}$ of $\mathbb{R}^p$. We propose a class of shrinkage estimators wich dominates the natural estimator of θ under the quadratic and weighted quadratic loss functions.

University of Shahrood

Faculty Of Mathematical Sciences

Dissertation Submitted in Partial Fulfillment of The Requirements For The Degree of Master of Science in Statistics

Shrinkage Estimation in Restricted Models.

Supervisor

Dr. Mohammad Arashi

by

Hamid Karami Kabir

May 2013