بسم الله الرحمن الرحیم

دانشگاه صنعتی شاهرود

دانشکده علوم ریاضی

گروه آمار

پایان‌نامه

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

عنوان

بررسی رفتار برآوردگر جک‌نایف ریج

استاد راهنما

دکتر محمد آرشی

دانشجو

مریم برزوئی بیدگلی

۱۳۹۳

پیر ما گفت خطا بر قلم صنع نرفت

آفرین بر نظر پاک خطاپوشش باد

«حافظ»

به

پدر و مادرم

و روح پاک الهه

سپاس گزاری...

سپاس بی‌کران پروردگار یکتا را که هستی‌مان بخشید، و به طریق علم و دانش رهنمون ساخت، و به همنشینی رهروان علم و معرفت، مفتخرمان نمود، و خوشه‌چینی از علم و آگاهی را روزی‌مان نهاد.

با سپاس از جناب آقای دکتر محمد آرشی، استاد راهنمای ارجمند که با ارائه رهنمودها، انتقادها و پیشنهادهایشان، در تمامی مراحل اجرای پایان‌نامه مرا حمایت و تشویق نمودند. از استادان شایسته جناب آقایان دکتر احمد نزاکتی و دکتر محمدرضا ربیعی که زحمت داوری این رساله را متقبل شدند، نهایت تشکر و قدردانی را دارم.

از تنها سرمایه زندگی‌ام، پدر بزرگوار و مادر عزیز، که هرچه امروز دارم از دعای خیر ایشان است و همواره قوت قلبی برای ادامه راهم بوده‌اند، تشکری ویژه و صمیمانه دارم. سپاس از خواهران و برادرانم که همواره در طول تحصیل حامی من در مواجهه با مشکلات بودند و هستی‌شان مایه‌ی دلگرمی من است.

بوسه می‌زنم بر دستان اولین معلمم سرکار خانم «سنجر» عزیز، که با همه وجودش آموخت تا بیاموزم و سپاسگزارم از جناب آقای «دکتر علی دولتی» استادی که سپیدی را بر تخته سیاه زندگیم نگاشت و برایم زندگی، بودن و انسان بودن را معنا کرد.

از استادان محترم جناب آقایان دکتر داود شاهسونی، دکتر حسین باغیشنی، دکتر علی مس فروش و سرکار خانم دکتر نگار اقبال، که در طول دوران تحصیلی‌ام در دوره کارشناسی ارشد، جهت آموزش و ارتقای علمی بنده، زحمت کشیده‌اند و صادقانه و عاشقانه تلاش می‌کنند تا نقّالی دانسته‌ها را به نقّادی اندیشه‌ها تبدیل سازند، قدردانی می‌نمایم. همچنین از جناب آقای دکتر مهدی کبرائی (دانشگاه صنعتی امیر کبیر)، به خاطر زحماتی که برای توصیف داده‌ها کشیدند، صمیمانه قدردانی می‌نمایم.

بسی شایسته است تشکر کنم از دوست عزیزم، خانم «مینا نوروزی راد» به پاس محبت بی‌دریغش که بدون هیچ چشم‌داشتی مشکلات مسیر را برایم تسهیل نمود.

به پاس قدردانی از دوستان خوبم خانم‌ها، مریم خان احمدی، سوسن عشوره، مهشید اخوان نژاد، عطیه سابقی، الهام آشوری و کوثر السادات تفاخ که لذت لحظه‌های ناب با هم بودن، آرامش روحی و آسایش فکری‌ام مدیون حضور سبز آن‌هاست.

مریم برزوئی بیدگلی
۱۳۹۳

تعهدنامه

اینجانب **مریم برزوئی بیدگلی** دانشجوی کارشناسی ارشد رشته **آمار** **دانشکده علوم ریاضی** دانشگاه شاهرود، نویسنده پایان‌نامه با عنوان **بررسی رفتار برآوردگر جک‌نایف ریج** ، تحت راهنمایی **دکتر محمد آرشی** متعهد می‌شوم:

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

مریم برزوئی بیدگلی
۱۳۹۳

مالکیت نتایج و حق نشر

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

چکیده

هنگامی‌که همخطی بین ستون‌های ماتریس طرح در رگرسیون خطی وجود دارد، استفاده از روش کمترین توان‌های دوم در برآورد ضرایب مدل، معمولاً باعث می‌شود که برآوردهای خیلی ضعیفی به‌دست آید. ثابت شده است که واریانس برآوردگرهای کمترین توان‌های دوم ضرایب رگرسیون ممکن است در حد قابل توجهی افزایش یابند و طول بردار برآورد کمترین توان‌های دوم پارامتر به طور متوسط خیلی زیاد می‌شود. در این راستا، یکی از راه‌های رفع مشکلات مطرح شده، استفاده از برآوردگرهای اریب است.

مسئله همخطی و کاربردی نبودن برآوردگر کمترین توان‌های دوم، منجر به گستردگی استفاده از برآوردگرهای اریب شده است. روش‌های متعددی مانند رگرسیون ریج، لیو و ... برای به‌دست آوردن برآوردهای اریب ضرایب رگرسیون مطرح شده است. یکی از این روش‌ها استفاده از برآوردگر اصلاح‌شده جک‌نایف لیو است، که اساس کار این برآوردگر این است که دو برآوردگر نوع لیو و جک‌نایف لیو را با هم ترکیب می‌کند و سپس برآوردگری ارائه می‌دهد که عملکرد بهتری نسبت به برآوردگر کمترین توان‌های دوم دارد. این برآوردگر اریب بوده اما، دارای واریانس کمتری نسبت به برآوردگر کمترین توان‌های دوم است.

در این پایان‌نامه مدل رگرسیون ریج را مورد بررسی قرار داده ایم. بدین منظور از نتایج برخی محققین بهره برده ایم که در نهایت منجر به برآوردگر جک‌نایف لیو اصلاح شده گردیده است. با توجه به پیچیدگی برآوردگر ارائه شده، از روش‌های شبیه‌سازی مونت‌کارلویی و جک‌نایف به منظور ارائه اریبی و مخاطره آن بهره برده ایم.

کلمات کلیدی: برآوردگر جک‌نایف، برآوردگر ریج، برآوردگر لیو، رگرسیون چندگانه، مخاطره، همخطی

پیش گفتار

آماردان‌ها از قدیم با مشکل انتخاب بین دو برآوردگر اریب مواجه بودند. در راستای انتخاب بین برآوردگرهای اریب، سعی نموده ایم با بهره‌گیری از تکنیک جک‌نایف برآوردگری را انتخاب کنیم که اریبی کمتری دارد. بدین منظور این تکنیک را برای برآوردگر اریب لیو به‌کار برده و برآوردگر جک‌نایف لیو اصلاح شده را معرفی کرده ایم.

در این راستا با مبنا قرار دادن مقاله‌ی آکدنیز دوران و آکدنیز[۱](۲۰۱۲) به دنبال کاهش اریبی برآوردهای اریب ضرایب رگرسیون می‌باشیم. آکدنیز دوران و آکدنیز (۲۰۱۲) در مقاله‌ی خود از روش جک‌نایف ارائه شده توسط هینکلی [۲](۱۹۷۷)، سینگ[۳]و همکاران (۱۹۸۶)، نیکوئیست [۴](۱۹۸۸) و باتا[۵]و همکاران (۲۰۰۸) برای محاسبه‌ی برآوردگر جک‌نایف لیو و برآوردگر جک‌نایف لیو تعمیم یافته، بهره برده اند. بدین منظور با یاری از نتایج مقاله آکدنیز دوران و آکدنیز (۲۰۱۲) و همچنین لی و یانگ[۶] (۲۰۱۲)، برآوردگر جک‌نایف لیو اصلاح شده را معرفی نموده ایم. با توجه به پیچیدگی برآوردگر ارائه شده، از روش‌های شبیه‌سازی مونت‌کارلویی و جک‌نایف به منظور محاسبه اریبی و مخاطره آن بهره برده ایم. این مجموعه شامل ۴ فصل و ۵ ضمیمه می‌باشد. مطالب هر فصل به طور مختصر عبارتست از:

- در فصل۱، مقدمات، نگاهی بر تاریخچه موضوع مورد بررسی آورده شده است. بنا به ضرورت، مقدماتی از رگرسیون خطی چندگانه، رگرسیون ریج، و برآوردگر نوع لیو ذکر گردیده است. در ادامه روش جک‌نایف تعریف و کاربرد آن در برآورد اریبی و واریانس ذکر شده است.
- در فصل۲، نوع دیگری از برآوردگر لیو را با عنوان برآوردگر جک‌نایف اصلاح شده لیو معرفی و کارایی آن را مورد بررسی قرار داده ایم. در نهایت با یک مثال عددی و همچنین شبیه‌سازی تلاش شده است که به برآوردگری که عملکرد بهتری نسبت به برآوردگر لیو تعمیم یافته دارد، دست یابیم.
- در فصل۳، برآوردگر معرفی شده توسط لی و یانگ (۲۰۱۲)، با عنوان برآوردگر لیو اصلاح شده مورد

[۱] Akdeniz Duran and Akdeniz

[۲] Hinkley

[۳] Singh

[۴] Nyquist

[۵] Batah

[۶] Li and Yang

تحلیل قرار گرفته است. در این فصل نیز با یک مثال عددی و شبیه سازی عملکرد این برآوردگر نسبت به سایر برآوردگرهای مطرح شده، مورد بررسی و مقایسه قرار گرفته است.

• در فصل۴، برآوردگر لی و یانگ (۲۰۱۲) را با بهره‌گیری از تکنیک جک‌نایف به برآوردگری تبدیل کرده که دارای اریبی کمتری نسبت به برآوردگر قبلی است. برای مقایسه‌ی کارایی این برآوردگر با سایر برآوردگرها از روش‌های شبیه‌سازی مونت کارلویی استفاده کرده‌ایم.

• در ضمیمه۱، مقدمه‌ای بر جبر ماتریس‌ها به منظور سادگی درک اثبات قضایا و لم‌ها در متن پایان‌نامه مطرح گردیده است.

• در ضمیمه۲، تعاریف و لم‌های اساسی که در محاسبات به‌کار گرفته شده است، را بیان نموده ایم.

• در ضمیمه۳، توضیحاتی در خصوص آماره کاپا و عامل تورم واریانس به عنوان عاملی برای شناخت همخطی در مدل رگرسیون خطی داده شده است.

• در ضمیمه۴، برنامه‌های کامپیوتری با نرم افزارR ارائه شده است.

• در ضمیمه۵، خلاصه، بحث و نتیجه‌گیری و پیشنهادها برای ادامه‌ی کار در آینده مطرح گردیده است.

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

لیست مقالات مستخرج از پایان‌نامه

۱. برزوئی بیدگلی م. و آرشی م.، (۱۳۹۳). "کارایی برآوردگر جک‌نایف ریج"، دوازدهمین کنفرانس آمار ایران، کرمانشاه، ایران.

۲. برزوئی بیدگلی م. و آرشی م.، (۱۳۹۳). "کاربرد برآوردگر لیو اصلاح شده در صنعت پتروشیمی"، نخستین کنفرانس بین المللی تحقیق در عملیات برای تصمیم‌سازی بهینه در صنایع نفت، گاز، پتروشیمی و پالایش و پخش، تهران، ایران.

۳. برزوئی بیدگلی م. و آرشی م.، (۱۳۹۳). "برآوردگر جک‌نایف لیو اصلاح شده"، مجله علوم آماری، تحت داوری.

فهرست نشانه‌ها، نمادها و علائم اختصاری

$\mathbf{X}$	ماتریسX
X	بردار تصادفیX
$\mathbf{x}$	بردار تصادفی
x	یافته‌ی بردار تصادفی X
A'	ترانهاده‌ی ماتریس A
$X_{(i)}$	i- امین نمونه جک‌نایف
$\hat{\gamma}_{OLSE}$	برآوردگر کمترین توان‌های دوم γ
$\hat{\gamma}(k)$	برآوردگر ریج γ
$\hat{\gamma}(d)$	برآوردگر لیو γ
$\hat{\gamma}(\mathbf{D})$	برآوردگر لیو تعمیم یافته γ
$\hat{\gamma}_{JGLE}(d)$	برآوردگر جک‌نایف لیو γ
$\hat{\gamma}_{JGLE}(\mathbf{D})$	برآوردگر جک‌نایف لیو تعمیم یافته γ
$\hat{\gamma}_{MJGLE}(\mathbf{D})$	برآوردگر اصلاح شده جک‌نایف لیو تعمیم یافته γ
$\hat{\gamma}_{JMLE}(d,b)$	برآوردگر جک‌نایف لیو اصلاح شده γ
$\chi^{۲}(q)$	توزیع کی- دو مرکزی با درجه آزادی q
$N_p(\mu, \mathbf{\Sigma})$	نرمال p- متغیره با میانگین μ و واریانس Σ
$\kappa(\mathbf{X})$	عدد شرطی $\mathbf{X}$
VIF_j	عامل تورم واریانس برای مشاهده j- ام
$\hat{\sigma}^{۲}$	برآورد واریانس جامعه
ρ	ضریب همبستگی
$MSEM$	ماتریس میانگین توان دوم خطا
$SMSE$	اثر ماتریس میانگین توان دوم خطا
$ABIAS$	اثر ماتریس قدر مطلق اریبی
d_{mm}	برآورد نوع لیو تعمیم یافته مقدار d
$\hat{\beta}_{OLSE}$	برآوردگر کمترین توان‌های دوم β
$\hat{\beta}(k)$	برآوردگر ریج β
$\hat{\beta}(d)$	برآوردگر لیو β
$\hat{\beta}_{ME}$	برآوردگر آمیخته β
$\tilde{\beta}$	برآوردگر کمترین توان‌های دوم محدود شده β
$\hat{\beta}_{MRE}(k,b)$	برآوردگر ریج اصلاح شده β
$\hat{\beta}_{MLE}(k,b)$	برآوردگر لیو اصلاح شده β

فهرست نشانه‌ها، نمادها و علائم اختصاری

LSE	برآوردگر کمترین توان‌های دوم
LE	برآوردگر نوع لیو
GLE	برآوردگر لیو تعمیم یافته
JLE	برآوردگر جک‌نایف لیو
$JGLE$	برآوردگر جک‌نایف لیو تعمیم یافته
$MJLE$	برآوردگر جک‌نایف اصلاح شده لیو
$MJGLE$	برآوردگر جک‌نایف اصلاح شده لیو تعمیم یافته
RE	برآوردگر ریج
MRE	برآوردگر ریج اصلاح شده
MLE	برآوردگر لیو اصلاح شده
$JMLE$	برآوردگر جک‌نایف لیو اصلاح شده

فهرست مطالب

لیست تصاویر ش

لیست جداول ص

۱ مقدمه و دورنما ۱

۱.۱ مقدمه و تاریخچه ۱

۲.۱ رگرسیون خطی چندگانه ۴

۱.۲.۱ برآورد پارامترهای مدل ۴

۲.۲.۱ همخطی در رگرسیون چندگانه ۹

۳.۲.۱ روش‌هایی برای برخورد با همخطی چندگانه ۱۰

۳.۱ رگرسیون ریج ۱۱

۴.۱ برآوردگر نوع لیو ۱۵

۱.۴.۱ برآوردگر لیو تعمیم یافته ۱۶

۲.۴.۱ مثال عددی ۱۷

۵.۱ برآوردگر آمیخته ۲۱

۶.۱ روش جک‌نایف ۲۲

۲ برآوردگر جک‌نایف اصلاح شده لیو ۲۹

۱.۲ مقدمه ۲۹

۲.۲ معرفی برآوردگر جک‌نایف لیو تعمیم یافته ۲۹

۱.۲.۲ مثال عددی ۳۳

۳.۲ برآوردگر جک‌نایف اصلاح شده لیو تعمیم یافته ۳۴

۴.۲ کارایی برآوردگر جک‌نایف اصلاح شده لیو تعمیم یافته ۳۷

۱.۴.۲ مثال عددی ۳۸

۵.۲ مقایسه برآوردگر جک‌نایف اصلاح شده لیو تعمیم یافته و برآوردگر جک‌نایف لیو تعمیم یافته ۳۸

۶.۲ مثال عددی ۴۱

۷.۲ مطالعه شبیه‌سازی ۴۷

۳ برآوردگر لیو اصلاح شده لی و یانگ ۵۰
۱.۳ مقدمه ۵۰
۲.۳ برآوردگر لیو اصلاح شده ۵۴
۳.۳ مقایسه برآوردگرها ۵۵
۱.۳.۳ مقایسه بین برآوردگر لیو اصلاح شده و لیو ۵۶
۲.۳.۳ مقایسه بین برآوردگرهای لیو ، ریج و لیو اصلاح شده ۵۷
۳.۳.۳ مقایسه بین برآوردگرهای ریج اصلاح شده و لیو اصلاح شده ۶۰
۴.۳ مثال عددی ۶۲
۵.۳ مطالعه شبیه‌سازی ۶۵

۴ برآوردگر جک‌نایف اصلاح شده لی و یانگ ۶۸
۱.۴ مقدمه ۶۸
۲.۴ معرفی برآوردگر جک‌نایف لیو اصلاح شده ۶۹
۳.۴ اریبی و واریانس برآوردگر ۷۱
۴.۴ کارایی برآوردگر جک‌نایف لیو اصلاح شده ۷۳
۵.۴ مطالعه شبیه‌سازی ۷۴
۶.۴ نتیجه‌گیری ۷۵
۷.۴ پیشنهادات برای آینده تحقیق ۷۶

آ جبر ماتریس‌ها ۷۸

ب تعاریف، لم‌ها و قضایای اساسی ۸۲

پ شاخص‌های همخطی ۸۵
پ.۱ آماره کاپا ۸۵
پ.۲ عامل تورم واریانس ۸۵

ت برنامه‌های کامپیوتری با نرم افزار R ۸۷
ت.۱ برنامه مربوط به مثال ۲.۴.۱ ۸۷
ت.۲ برنامه مربوط به مثال ۱.۲.۲. ۹۱
ت.۳ برنامه مربوط به مثال ۱.۴.۲ ۹۴
ت.۴ برنامه مربوط به مثال ۶.۲ ۹۷
ت.۵ برنامه مربوط به بخش ۷.۲ ۱۰۳
ت.۶ برنامه مربوط به مثال ۴.۳ ۱۰۶

ت.۷ برنامه مربوط به بخش ۵.۳ . . . . ۱۱۱
ت.۸ برنامه مربوط به بخش ۵.۴ . . . . ۱۱۴

مراجع **۱۱۸**

واژه‌نامه فارسی به انگلیسی **۱۲۳**

واژه‌نامه انگلیسی به فارسی **۱۲۵**

نمایه **۱۲۷**

لیست تصاویر

۱.۱ عملکرد برآوردگر نوع لیو به ازای مقادیر مختلف d ، $\circ < d < ۱$ ۲۰

۱.۲ عملکرد برآوردگر لیو به ازای مقادیر مختلف d ، $\circ < d < ۱$ ۳۵
۲.۲ عملکرد برآوردگر لیو به ازای مقادیر مختلف d ، $\circ < d < ۱$ ۳۹
۳.۲ مقایسه برآوردگرها به ازای مقادیر مختلف d ، $\circ < d < ۱$ ۴۵

۱.۳ برآورد MSE برای برآوردگرهای RE و MRE ۶۴
۲.۳ برآورد MSE برای برآوردگرهای LE و MLE ۶۴

لیست جداول

۱.۱ داده‌های در صد تبدیل n- هپتان به استیلن ۱۸
۲.۱ برآورد $\hat{\beta}(d)$ برای مقادیر مختلف d ۱۹

۱.۲ برآورد $\hat{\beta}(d)$ برای مقادیر مختلف d ۳۴
۲.۲ برآورد $\hat{\beta}(d)$ برای مقادیر مختلف d ۴۰
۳.۲ همبستگی بین متغیرها ۴۲
۴.۲ برآورد β برای مقادیر مختلف d ۴۳
۵.۲ برآورد β برای مقادیر مختلف d ۴۴
۶.۲ برآورد β برای مقادیر مختلف d ۴۴
۷.۲ برآورد $MSEM$ برای $OLSE$ ، LE ، JLE و $MJLE$ به ازای (الف) $n=۱۵$ و $p=۳$، (ب) $n=۵۰$ و $p=۳$ ، (پ) $n=۱۰۰$ و $p=۳$ ۴۸
۸.۲ برآورد قدر مطلق اریبی برای $OLSE$ ، LE ، JLE و $MJLE$ به ازای (الف) $n=۱۵$ و $p=۳$ ، (ب) $n=۵۰$ و $p=۳$، (پ) $n=۱۰۰$ و $p=۳$ ۴۸

۱.۳ برآورد MSE برای LSE ، MLE ، LE ، RE و MRE ۶۳
۲.۳ برآورد MSE برای LSE ، MLE ، LE ، RE و MRE به ازای $\gamma=۰/۷۰$. ۶۶
۳.۳ برآورد MSE برای LSE ، MLE ، LE، RE و MRE به ازای $\gamma=۰/۹۰$ . . ۶۶
۴.۳ برآورد MSE برای LSE ، MLE ، LE ، RE و MRE به ازای $\gamma=۰/۹۹$. ۶۶

۱.۴ برآورد MSE برای LSE ، LE ، MLE و $JMLE$ به ازای $\gamma=۰/۷۰$ ۷۵
۲.۴ برآورد MSE برای LSE ، LE ، MLE و $JMLE$ به ازای $\gamma=۰/۹۰$ ۷۵
۳.۴ برآورد MSE برای LSE ، LE ، MLE و $JMLE$ به ازای $\gamma=۰/۹۹$ ۷۵

فصل ۱

مقدمه و دورنما

۱.۱ مقدمه و تاریخچه

در کتب آماری، رگرسیون به عنوان روش تعیین و تحلیل روابط نادقیق بین متغیرهای آماری تعریف می‌شود. بنابراین در رگرسیون دو مطلب تعیین روابط بین متغیرها و تحلیل روابط به‌دست آمده مورد توجه قرار می‌گیرد.

رگرسیون در لغت به معنای «برگشت» است. دلیل این نام‌گذاری به مطالعات گالتون[۱] (۱۸۸۶) در مورد رابطه بین قد پسر و قد پدر مربوط می‌شود. گالتون با رسم نمودار قد پسر در مقابل قد پدر (نمودار پراکنش) بر اساس تعداد زیادی از داده‌ها (مشاهدات) به این نتیجه رسید که اغلب افراد بلند قد، پسرانی کوتاه قدتر و اکثر افراد کوتاه قد، پسرانی بلند قدتر از خود دارند و این پدیده را «برگشت به میانگین[۲]» نامید و از آن پس نام رگرسیون بر روش‌های بررسی روابط بین متغیرهای آماری باقی ماند. تعیین روابط بین متغیرهای آماری را رگرسیون توصیفی و تحلیل این روابط را رگرسیون استنباطی می‌نامند.

به عنوان مثالی از یک مسئله که در آن تحلیل رگرسیونی می‌تواند مفید واقع شود، فرض کنید یک مهندس صنایع، توسط یک سازنده نوشابه استخدام شده که محصول تحویلی و عملیات سرویس و خدمت رسانی برای ماشین‌های فروش را تجزیه و تحلیل کند. او می‌خواهد بداند که زمان لازم برای این‌که یک فرد تحویل‌دهنده، این ماشین را سرویس دهد بستگی به تعداد موارد محصول تحویل شده دارد یا خیر. مهندس ۲۵ فروشنده جزئی را که دارای ماشین فروش می‌باشند، به تصادف بازدید کرده است و زمان تحویل (بر حسب دقیقه) و حجم محصول تحویل داده شده (بر حسب مورد) برای هر فقره، مشاهده می‌شود.

رابطه بین زمان تحویل با متغیرهای ذکر شده در بالا به شکل معادله یا الگویی است که متغیر وابسته (زمان تحویل) را به یک یا چند متغیر پیش‌بین[۳] مربوط می‌کند. در رگرسیون خطی متغیر پاسخ پیوسته

[۱] Francis Galton

[۲] Regression to the mean

[۳] Predictor

است ولی متغیرهای پیش‌بین می‌توانند پیوسته، گسسته یا رسته‌ای[۴] باشند. در این مثال زمان تحویل متغیر وابسته بر حسب دقیقه محاسبه می‌شود و متغیرهای پیش‌بین تعداد مختلف محصول تحویل شده می‌باشد. متغیر پاسخ را با y و متغیرهای پیش‌بین را با $X_۱, X_۲, \cdots, X_p$ نشان می‌دهیم.

متغیرهای پیش‌بین را «متغیرهای مستقل» نیز می‌نامند، اما بکار بردن این اصطلاح ممکن است تا حدی گمراه کننده باشد، زیرا همیشه $X_۱, X_۲, \cdots, X_p$ ، مستقل از یکدیگر نیستند و معمولاً از y مستقل نیستند. متغیر پاسخ نیز، گاهی اوقات «متغیر وابسته» نامیده می‌شود. برای توضیحات بیشتر، نیرومند (۱۳۸۷) را ببینید.

تحلیل رگرسیونی کاربردی‌ترین و مهم‌ترین روش آماری در تحلیل‌های آماری، اقتصادی، مالی، صنعتی و ... می‌باشد؛ به‌خصوص در ۴ دهه اخیر به طور ماهرانه‌ای در حل و بررسی مسائل کاربردی مورد استفاده قرار گرفته است. در بین تعداد زیاد شاخه‌های تحلیل رگرسیونی، روش برآورد کمترین توان‌های دوم تعمیم یافته[۵]، بر اساس نظریه معروف گوس- مارکف[۶] (کاریا و کوراتا[۷]، ۲۰۰۴ را ببینید) روشی پایه‌ای می‌باشد و هنوز نقشی اساسی در بسیاری از جنبه‌های نظری و کاربردی استنباط آماری به‌خصوص مدل رگرسیون خطی ایفا می‌کند.

اگر هیچ رابطه خطی بین متغیرهای پیش‌بین وجود نداشته باشد، این متغیرها متعامد[۸] نامیده می‌شوند. هنگامی‌که متغیرهای پیش‌بین متعامد باشند، برآوردگر کمترین توان‌های دوم پر کاربرد خواهند بود. اما متأسفانه، در بسیاری از موارد در عمل، متغیرهای پیش‌بین متعامد نیستند. گاهی اوقات، عدم وجود تعامد در متغیرهای پیش‌بین، مسأله بسیار حادی نیست. اما اگر در یک الگوی رگرسیون متغیرهای پیش‌بین به صورت کامل و یا تقریباً کامل، وابسته خطی باشند، در این صورت ممکن است استنباط بر اساس این الگوی رگرسیون تا حدی بسیار گمراه‌کننده و پر اشتباه باشد. در این هنگام می‌گوییم که میان متغیرها در الگوی رگرسیون همخطی وجود دارد.

مسئله همخطی و کاربردی نبودن برآوردگر کمترین توان‌های دوم، منجر به استفاده گسترده ای از برآوردگرهای اریب شده است. برخی روش‌های برآوردیابی اریب برای رفع مشکل همخطی، عبارتند از برآورد انقباضی استاین[۹]، برآورد ریج[۱۰] ، برآورد نوع لیو[۱۱] و ... (برای مطالعه بیشتر در زمینه برآوردگرهای بهبود یافته، به لیو[۱۲] ۲۰۰۳، راداکریشنا رائو و توتنبرگ[۱۳] ۱۹۹۵، سوئیندل[۱۴] ۱۹۷۶ و حسن زاده بشتیان ۱۳۸۸ مراجعه کنید).

[۴] Categorical
[۵] Generalized least squares
[۶] Gauss- Markov
[۷] Kariya and Kurata
[۸] Orthogonal
[۹] Stein shrinkage estimation
[۱۰] Ridge estimation
[۱۱] Liu-type estimation
[۱۲] Liu
[۱۳] Radhakrishna Rao and Tuotenburg
[۱۴] Swindel

یکی از مباحث مهم در استنباط آماری، بررسی ویژگی‌های اصلی برآوردها از جمله اریبی، واریانس و توزیع نمونه‌ای آن‌هاست. استفاده از برخی روش‌های معمول برای این منظور، عمدتاً پیچیده و حتی غیر ممکن هستند و در برخی از آنان نیاز به برقراری فرض‌های تئوری پیچیده وجود دارد. در چنین شرایطی، استفاده از روش‌های بازنمونه‌گیری می‌تواند راهگشا باشد، زیرا مهم‌ترین مزیت این روش‌ها عدم نیاز به محاسبات تئوری و برقرای فرض‌های پیچیده است. در سال‌های اخیر توجه زیادی به روش‌های بازنمونه‌گیری در رگرسیون شده است. روش‌های بوت استرپ[۱۵] و جک‌نایف[۱۶] عمده ترین روش بازنمونه‌گیری می‌باشند. روش جک‌نایف در سال ۱۹۴۹ توسط کوئینلی معرفی شد. روش بوت استرپ برای اولین بار توسط بردلی افرن[۱۷] در سال ۱۹۷۹ در مقاله‌ای با عنوان «بوت استرپ: نگاهی دیگر به روش جک‌نایف»، معرفی و در این مقاله از آن برای برآورد اریبی، واریانس و توزیع نمونه‌ای آماره‌ها استفاده شد. بوت استرپ معرفی شده توسط افرن، بوت استرپ ناپارامتری[۱۸] بود. ولی بعدها انواع مختلف بوت استرپ از جمله بوت استرپ پارامتری[۱۹]، بوت استرپ هموار[۲۰]، بوت استرپ بیزی[۲۱] و بوت استرپ دوگانه[۲۲] نیز معرفی شدند. برای آگاهی بیشتر به روبین[۲۳] (۱۹۸۱)، سیلورمن و یانگ[۲۴] (۱۹۸۷)، وایند[۲۵] (۱۹۹۵)، و دیویسن و هینکلی[۲۶] (۱۹۹۷) مراجعه کنید.

استفاده از روش‌های بوت استرپ در رگرسیون اولین بار توسط دیوید فریدمن[۲۷] در مقاله‌ای با عنوان «بوت استرپ کردن مدل‌های رگرسیون» در سال ۱۹۸۱ مطرح شد.

کاربرد روش‌های بوت استرپ در رگرسیون بر اساس ثابت بودن و یا تصادفی بودن متغیرهای پیش‌بین است. در صورت ثابت بودن متغیرهای پیش‌بین از روش بازنمونه‌گیری از مانده‌ها و در صورت تصادفی بودن، از روش بازنمونه‌گیری از مشاهدات استفاده می‌شود.

از زمان چاپ مقاله ای توسط افرن و گانگ[۲۸] در سال ۱۹۸۳ توجه بسیاری از آماردانان به روش جک‌نایف جلب شد و اهمیت و کاربرد آن روز به روز آشکارتر گردید، به طوری‌که از اواخر دهه ۷۰ تا اواسط دهه ۸۰، بیش از ده هزار مقاله در مورد بوت استرپ، ویژگی‌ها و کاربردهای آن در زمینه‌های مختلف نوشته شد. نیکوئیست در سال ۱۹۸۸ تکنیک جک‌نایف را برای رگرسیون ریج به کار برد و به این نتیجه رسید که برآوردگر ریج جک‌نایف دارای پارامتر اریبی کمتر و واریانس بیشتری نسبت به

[۱۵] Bootstrap
[۱۶] Jackknife
[۱۷] Bradley Efron
[۱۸] Non-parametric bootstrap
[۱۹] Parametric bootstrap
[۲۰] Smoothed bootstrap
[۲۱] Bayesian bootstrap
[۲۲] Double bootstrap
[۲۳] Rubin
[۲۴] Silverman and Yang
[۲۵] Vinod
[۲۶] Davison and Hinkley
[۲۷] Freedman
[۲۸] Efron and Gong

برآوردگر ریج است. در این پایان‌نامه از روش جک‌نایف برای انجام رگرسیون استفاده شده است.

در ادامه، برای روشن‌تر شدن اهداف اصلی و خط مشی پایان‌نامه، به معرفی مختصری از روش رگرسیون چندگانه، رگرسیون ریج، برآوردگر نوع لیو و روش جک‌نایف می‌پردازیم.

۲.۱ رگرسیون خطی چندگانه

رابطه‌ی p متغیر پیش‌بین و یک متغیر پاسخ در رگرسیون خطی چندگانه[۲۹] به صورت زیر فرمول‌بندی می‌شود:

$$y = \mathbf{X}\beta + \epsilon \qquad (۱.۱)$$

که درآن $y = (y_۱, \cdots, y_n)'$ یک بردار $n \times ۱$ از پاسخ‌ها، $\mathbf{X} = (X_۱, \cdots, X_n)'$ یک ماتریس $n \times p$ غیر تصادفی و معلوم از n بردار p -بعدی، $\beta = (\beta_۱, \cdots, \beta_p)'$ بردار ضرایب رگرسیون مجهول (پارامترهای مدل) و $\epsilon = (\epsilon_۱, \cdots, \epsilon_n)'$ بردار خطای تصادفی هستند. به عبارتی دیگر داریم

$$\begin{bmatrix} y_۱ \\ y_۲ \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} x_{۱۱} & x_{۱۲} & \cdots & x_{۱p} \\ x_{۲۱} & x_{۲۲} & \cdots & x_{۲p} \\ \vdots & & & \\ x_{n۱} & x_{n۲} & \cdots & x_{np} \end{bmatrix} \begin{bmatrix} \beta_۱ \\ \beta_۲ \\ \vdots \\ \beta_p \end{bmatrix} + \begin{bmatrix} \epsilon_۱ \\ \epsilon_۲ \\ \vdots \\ \epsilon_n \end{bmatrix}.$$

اغلب در مسائل کاربردی، ستون اول $\mathbf{X}$ بردار $n \times ۱$ از یک‌ها به منظور برآوردکردن مقدار ثابت در مدل رگرسیون می‌باشد.

رگرسیون خطی چندگانه تعمیمی از رگرسیون خطی ساده[۳۰] است. این روش پارامتری هنگامی مناسب است که فرض خطی بودن رابطه‌ی میان متغیر پاسخ و متغیرهای پیش‌بین دست کم به صورت تقریبی برقرار باشد. اما هنگامی‌که رابطه میان متغیر پاسخ و متغیرهای پیش‌بین نامعلوم است، استفاده از رگرسیون خطی ممکن است منجر به نتایج گمراه کننده‌ای شود.

۱.۲.۱ برآورد پارامترهای مدل

مدل خطی (۱.۱) را در نظر گرفته و فرض کنید ماتریس واریانس- کوواریانس مؤلفه خطا برابر $\sigma^۲\mathbf{I}$ باشد، به عبارتی $\mathrm{cov}(\epsilon) = \sigma^۲\mathbf{I}$ ، که در آن $\sigma^۲$ یک مقدار اسکالر نامعلوم است.

برآوردگر کمترین توان‌های دوم، برآوردگری است که مجموع توان‌های دوم مانده‌ها یعنی

$$\begin{aligned} S(\beta_۰, \beta_۱, \cdots, \beta_k) &= \sum_{i=۱}^{n} \epsilon_i^۲ \\ &= \sum_{i=۱}^{n} \left(y_i - \beta_۰ - \sum_{j=۱}^{p} \beta_j \mathbf{X}_{ij} \right)^۲ \end{aligned}$$

[۲۹] Multiple linear regression

[۳۰] Ordinary linear regression

$$= (y - \mathbf{X}\beta)'(y - \mathbf{X}\beta)$$
$$= y'y - 2y'\mathbf{X}\beta + \beta'\mathbf{X}'\mathbf{X}\beta \qquad (2.1)$$

را مینیمم کند. با مشتق‌گیری از S نسبت به $(\beta_0, \beta_1, \cdots, \beta_k)$ به شکل زیر

$$\frac{\partial S}{\partial \beta} = \mathbf{0} - 2\mathbf{X}'y + 2\mathbf{X}'\mathbf{X}\beta$$

و مساوی با صفر قرار دادن $\frac{\partial S}{\partial \beta}$، در صورت نامنفرد بودن[31] ماتریس $\mathbf{X}'\mathbf{X}$ (معکوس پذیر بودن)، برآوردگرهای کمترین توان‌های دوم ضرایب رگرسیونی به صورت زیر به‌دست می‌آید

$$\hat{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'y. \qquad (3.1)$$

بردار مقادیر برازش شده $\hat{y}$ با مقادیر مشاهده شده y چنین خواهد بود

$$\hat{y} = \mathbf{X}\hat{\beta} = \mathbf{H}y \qquad (4.1)$$

که در آن $\mathbf{H} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$ ماتریس هَت[32] یا برازش شده[33] نامیده می‌شود زیرا باعث می‌شود روی y یک علامت هَت گذاشته شود و از طرفی بردار مقادیر مشاهدات را به بردار مقادیر برازش شده تصویر می‌کند. ماتریس H ماتریسی $p \times p$، متقارن و خودتوان[34] است و بنا به قضیه آ.۰.۲۳، رابطه زیر بین رتبه و اثر ماتریس برقرار است

$$rank(\mathbf{H}) = tr(\mathbf{H}) = p.$$

ماتریس واریانس- کوواریانس $\hat{\beta}$ نیز به صورت زیر محاسبه می‌شود

$$\begin{aligned}
\mathrm{cov}(\hat{\beta}) &= \mathrm{cov}\left((\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'y\right) \\
&= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathrm{cov}(y)\left((\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\right)' \\
&= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathrm{cov}(y)\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1} \\
&= \sigma^2(\mathbf{X}'\mathbf{X})^{-1}. \qquad (5.1)
\end{aligned}$$

قضیه ۱.۲.۱. *(قضیه گوس- مارکف) اگر $E(y) = \boldsymbol{X}\beta$ و $\mathrm{cov}(y) = \sigma^2\boldsymbol{I}$، آن‌گاه برآوردهای کمترین توان‌های دوم $\hat{\beta}_j$، $j = 0, 1, \cdots, k$ در میان تمام برآوردگرهای نااریب خطی دارای می‌نیمم واریانس هستند.*

برهان. یک برآوردگر خطی مانند $\mathbf{A}y$ از β در نظر می‌گیریم و ماتریس $\mathbf{A}$ را به قسمی پیدا می‌کنیم که $\mathbf{A}y$ یک برآوردگر نااریب با می‌نیمم واریانس برای β شود. برای آن‌که $\mathbf{A}y$ یک برآوردگر نااریب β باشد، باید داشته باشیم $E(\mathbf{A}y) = \beta$. با استفاده از فرض $E(y) = \mathbf{X}\beta$، این را می‌توان به صورت زیر نوشت

$$E(\mathbf{A}y) = \mathbf{A}E(y) = \mathbf{A}\mathbf{X}\beta = \beta$$

[31] Non- singular
[32] Hat matrix
[33] Fitting
[34] Idempotent

که شرط نااریبی زیر را می‌دهد

$$\mathbf{AX} = \mathbf{I}$$

زیرا رابطه $\mathbf{AX}\beta = \beta$ باید برای تمام مقادیر ممکن β برقرار باشد.

ماتریس واریانس- کوواریانس برای برآوردگر $\mathbf{A}y$ عبارت است از

$$\mathrm{cov}(\mathbf{A}y) = \mathbf{A}(\sigma^{2}I)\mathbf{A}' = \sigma^{2}\mathbf{AA}'.$$

واریانس $\hat{\beta}_j$ روی قطر $\sigma^{2}\mathbf{AA}'$ قرار دارند، بنابراین باید $\mathbf{A}$ را به قسمی انتخاب کنیم که (با توجه به $\mathbf{AX} = \mathbf{I}$) اعضای قطری $\mathbf{AA}'$ می‌نیمم شوند. برای ارتباط $\mathbf{A}y$ به $\hat{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'y$، عبارت $(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$ را اضافه و کم می‌کنیم

$$\mathbf{AA}' = [\mathbf{A} - (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'][\mathbf{A} - (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}']'.$$

اگر این عبارت را نسبت به $\mathbf{A} - (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$ و $(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$ بسط دهیم، چهار جمله به‌دست می‌آوریم که دو جمله آن بنابر $\mathbf{AX} = \mathbf{I}$، صفر می‌شوند. در نتیجه داریم

$$\mathbf{AA}' = [\mathbf{A} - (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'][\mathbf{A} - (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}']' + (\mathbf{X}'\mathbf{X})^{-1} \qquad (۶.۱)$$

ماتریس $[\mathbf{A} - (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'][\mathbf{A} - (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}']'$ از سمت راست (۶.۱) نیمه معین مثبت است (قضیه آ.۰.۲۰ را ببینید)، بنابه قسمت اول قضیه آ.۰.۱۸، اعضای قطری بزرگتر یا مساوی صفرند. با انتخاب $\mathbf{A} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$، این اعضای قطری را می‌توان برابر صفر قرار داد. (این مقدار $\mathbf{A}$ در شرط نااریب بودن $\mathbf{AX} = \mathbf{I}$ نیز صدق می‌کند.) پس برآوردگرهای می‌نیمم واریانس حاصل برای β عبارت است از

$$\mathbf{A}y = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'y$$

که برابر همان برآورد کمترین توان‌های دوم $\hat{\beta}$ است. □

برآوردگر نااریب σ^{2} عبارتست از میانگین توان‌های دوم مانده‌ها

$$E(s^{2}) = \frac{E(SSE)}{n-p} = \frac{(n-p)\sigma^{2}}{n-p} = \sigma^{2} \qquad (۷.۱)$$

که در آن $SSE = (y - \mathbf{X}\beta)'(y - \mathbf{X}\beta)$، مجموع توان‌های دوم مانده‌ها بوده و

$$\frac{(n-p)S^{2}}{\sigma^{2}} \sim \chi^{2}(n-p).$$

قضیه ۱.۲.۲. * *(کاریا و کوراتا، ۲۰۰۴) اگر در مدل (۱.۱)، $\mathrm{cov}(y) = \mathrm{cov}(\epsilon) = \sigma^{2}\Sigma$، که در آن Σ ماتریسی معین مثبت است، آن‌گاه با قرار دادن*

$$(\boldsymbol{X}'\Sigma^{-1}\boldsymbol{X}) = \boldsymbol{C}.$$

(الف) برآوردگر کمترین توان‌های دوم ضرایب رگرسیونی به صورت زیر است

$$\hat{\beta} = \boldsymbol{C}^{-1}\boldsymbol{X}'\Sigma^{-1}y. \qquad (۸.۱)$$

(ب) ماتریس کوواریانس $\hat{\beta}$ عبارت است از

$$\mathrm{cov}(\hat{\beta}) = \sigma^{2}\boldsymbol{C}^{-1}. \qquad (۹.۱)$$

(پ) برآوردگر نااریب $\hat{\sigma}^{2}$ به شکل زیر است

$$\hat{\sigma}^{2} = \frac{(y - \boldsymbol{X}\hat{\beta})'\Sigma^{-1}(y - \boldsymbol{X}\hat{\beta})}{n-p}. \qquad (۱۰.۱)$$

برهان. (الف) از این‌که Σ ماتریس معین مثبت است، ماتریس ناویژه $\mathbf{P}_{n\times n}$ وجود دارد که در آن $\Sigma = \mathbf{P}\mathbf{P}'$ (قضیه آ.۰.۱۹). با ضرب نمودن $\mathbf{P}^{-1}$ از سمت چپ در $y = \mathbf{X}\beta + \epsilon$، داریم

$$\mathbf{P}^{-1}y = \mathbf{P}^{-1}\mathbf{X}\beta + \mathbf{P}^{-1}\epsilon$$

که در آن

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

همچنین بنا به قضیه آ.۰.۲۱، می‌توان نوشت

$$\begin{aligned} \mathrm{cov}(\mathbf{P}^{-1}\epsilon) &= \mathbf{P}^{-1}\mathrm{cov}(\epsilon)(\mathbf{P}^{-1})' \\ &= \mathbf{P}^{-1}\sigma^2\mathbf{\Sigma}(\mathbf{P}^{-1})' \\ &= \sigma^2\mathbf{P}^{-1}\mathbf{P}\mathbf{P}'(\mathbf{P}')^{-1} \\ &= \sigma^2\mathbf{I}. \end{aligned}$$

بنابراین

$$\begin{aligned} \hat{\beta} &= \left[\mathbf{X}'(\mathbf{P}^{-1})'\mathbf{P}^{-1}\mathbf{X}\right]^{-1}\mathbf{X}'(\mathbf{P}^{-1})'\mathbf{P}^{-1}y \\ &= \left[\mathbf{X}'(\mathbf{P}')^{-1}\mathbf{P}^{-1}\mathbf{X}\right]^{-1}\mathbf{X}'(\mathbf{P}')^{-1}\mathbf{P}^{-1}y \\ &= \left[\mathbf{X}'(\mathbf{P}\mathbf{P}')^{-1}\mathbf{X}\right]^{-1}\mathbf{X}'(\mathbf{P}\mathbf{P}')^{-1}y \\ &= \mathbf{C}^{-1}\mathbf{X}'\mathbf{\Sigma}^{-1}y. \end{aligned}$$

(ب)

$$\begin{aligned} \mathrm{cov}\left[\mathbf{C}^{-1}\mathbf{X}'y\right] &= \mathbf{C}^{-1}\mathbf{X}'\mathrm{cov}(y)\mathbf{X}\mathbf{C}^{-1} \\ &= \mathbf{C}^{-1}\mathbf{X}'\sigma^2\mathbf{\Sigma}\mathbf{X}\mathbf{C}^{-1} \\ &= \sigma^2\mathbf{C}^{-1}. \end{aligned}$$

(پ)
با توجه به این که

$$\begin{aligned} \hat{\beta}'\mathbf{X}'\Sigma^{-1}\mathbf{X}\hat{\beta} &= \hat{\beta}'\mathbf{X}'\Sigma^{-1}\mathbf{X}\mathbf{C}^{-1}\mathbf{X}'\mathbf{\Sigma}^{-1}y \\ &= \hat{\beta}'\mathbf{C}\mathbf{C}^{-1}\mathbf{X}'\mathbf{\Sigma}^{-1}y \\ &= \hat{\beta}'\mathbf{X}'\Sigma^{-1}y \end{aligned}$$

بنابراین

$$\begin{aligned}(y-\mathbf{X}\hat{\beta})'\mathbf{\Sigma}^{-1}(y-\mathbf{X}\hat{\beta}) &= y'\mathbf{\Sigma}^{-1}y - y'\mathbf{\Sigma}^{-1}\mathbf{X}\hat{\beta}\\ &- \hat{\beta}'\mathbf{X}'\Sigma^{-1}y + \hat{\beta}'\mathbf{X}'\Sigma^{-1}\mathbf{X}\hat{\beta}\\ &= y'\mathbf{\Sigma}^{-1}y - y'\mathbf{\Sigma}^{-1}\mathbf{X}\mathbf{C}^{-1}\mathbf{X}'\mathbf{\Sigma}^{-1}y\\ &= y'\left[\mathbf{\Sigma}^{-1} - \mathbf{\Sigma}^{-1}\mathbf{X}\mathbf{C}^{-1}\mathbf{X}'\mathbf{\Sigma}^{-1}\right]y\end{aligned}$$

و بنا به قضیه ب.۳۴.۰ و با در نظر گرفتن $\mathbf{A}=\left[\mathbf{\Sigma}^{-1}-\mathbf{\Sigma}^{-1}\mathbf{X}C^{-1}\mathbf{X}'\mathbf{\Sigma}^{-1}\right]$، $\mathbf{V}=\sigma^{2}\mathbf{\Sigma}$ و $\mu=\mathbf{X}\beta$ نتیجه حاصل می‌شود. □

قضیه ۳.۲.۱. * *(کاریا و کوراتا، ۲۰۰۴) اگر در مدل خطی (۱.۱)، y دارای توزیع $N_n(\boldsymbol{X}\beta,\sigma^{2}\Sigma)$ باشد، که در آن $\boldsymbol{X}$ ماتریس $n\times p$ با رتبه p و Σ ماتریس معین مثبت معلوم است، آن‌گاه برآوردگرهای درستنمایی ماکزیمم β و σ^{2} عبارت است از*

$$\hat{\beta}=\boldsymbol{C}^{-1}\boldsymbol{X}'\mathbf{\Sigma}^{-1}y \tag{۱۱.۱}$$

و

$$\hat{\sigma}^{2}=\frac{1}{n}(y-\boldsymbol{X}\hat{\beta})'\mathbf{\Sigma}^{-1}(y-\boldsymbol{X}\hat{\beta}). \tag{۱۲.۱}$$

برهان. با توجه به تابع چگالی نرمال n- متغیره، تابع درستنمایی برابر است با

$$\begin{aligned}\mathrm{Ln}f_Y(y) &= -\frac{n}{2}\mathrm{Ln}(2\pi)-\frac{n}{2}\mathrm{Ln}\sigma^{2}-\frac{n}{2}\mathrm{Ln}|\mathbf{\Sigma}|\\ &- \frac{(y-\mathbf{X}\beta)'\mathbf{\Sigma}^{-1}(y-\mathbf{X}\beta)}{2\sigma^{2}}\\ \frac{\partial\mathrm{Ln}f_Y(y)}{\partial\beta} &= \frac{-2\mathbf{X}'\mathbf{\Sigma}^{-1}(y-\mathbf{X}\beta)}{2\sigma^{2}}=\mathbf{0}\\ &\Longrightarrow\ \hat{\beta}=\mathbf{C}^{-1}\mathbf{X}'\mathbf{\Sigma}^{-1}y\end{aligned} \tag{۱۳.۱}$$

همچنین

$$\begin{aligned}\frac{\partial\mathrm{Ln}f_Y(y)}{\partial\sigma^{2}} &= -\frac{n}{2\sigma^{2}}+\frac{(y-\mathbf{X}\hat{\beta})'\mathbf{\Sigma}^{-1}(y-\mathbf{X}\hat{\beta})}{2\sigma^{4}}=\mathbf{0}\\ &\Longrightarrow\ \hat{\sigma}^{2}=\frac{(y-\mathbf{X}\hat{\beta})'\mathbf{\Sigma}^{-1}(y-\mathbf{X}\hat{\beta})}{n}.\end{aligned} \tag{۱۴.۱}$$

□

در این صورت مقادیر برازش شده به صورت زیر می‌باشند

$$\hat{y}=\mathbf{X}\hat{\beta}=\mathbf{H}\mathbf{\Sigma}^{-\frac{1}{2}}y \tag{۱۵.۱}$$

که در آن $\mathbf{H}=\mathbf{X}\mathbf{\Sigma}^{-\frac{1}{2}}(\mathbf{X}'\mathbf{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{\Sigma}^{-\frac{1}{2}}$. برای آگاهی بیشتر به کاریا و کوراتا (۲۰۰۴) مراجعه کنید.

۲.۲.۱ همخطی در رگرسیون چندگانه

تعبیر و تفسیر معادله‌ی رگرسیون چندگانه به طور ضمنی به این فرض وابسته است که متغیرهای پیش‌بین وابستگی قوی با یکدیگر ندارند و به عبارتی مستقل خطی هستند. زیرا تعبیر یک ضریب رگرسیون به صورت اندازه‌گیری تغییر در متغیر پاسخ وقتی متغیر پیش‌بین متناظر یک واحد افزایش یافته و سایر متغیرهای پیش‌بین ثابت نگاه داشته شوند امری متداول است؛ اما این تغییر در صورتی که روابط خطی قوی بین متغیرهای پیش‌بین وجود داشته باشد معتبر نخواهد بود.

وقتی هیچ ارتباط خطی میان متغیرهای پیش‌بین وجود نداشته باشد، این متغیرها را متعامد می‌گویند. در اکثر کاربردهای رگرسیون متغیرهای پیش‌بین متعامد نیستند. در بعضی وضعیت‌ها متغیرهای پیش‌بین تقریباً به طور کامل ارتباط خطی دارند و در این چنین مواردی نتایج رگرسیون مبهم می‌شود.

اگر یکی از متغیرهای پیش‌بین یک تابع دقیق خطی از یک یا چند متغیر پیش‌بین دیگر باشد گوییم رگرسیون دارای همخطی کامل[۳۵] است. همخطی ناقص[۳۶] وقتی اتفاق می‌افتد که یکی از متغیرهای پیش‌بین بطور تقریبی یک تابع خطی از یک یا چند متغیر پیش‌بین دیگر باشد. به عبارت دیگر همخطی کامل زمانی رخ می‌دهد که دست کم به ازای یک i ($i = ۱, \cdots, p$) داشته باشیم $R_i^۲ = ۱$ و همخطی ناقص زمانی پیش می‌آید که دست کم به ازای یک i داشته باشیم $R_i^۲ \cong ۱$، که در آن $R_i^۲$ ضریب تعیین رگرسیون خطی متغیر x_i بر سایر متغیرهای پیش‌بین است. باید دقت داشت که ضریب تعیین $(R^۲)$ حاصل تقسیم پراکندگی بیان‌شده به پراکندگی کل است. لذا در مدل رگرسیون (۱.۱)، ضریب تعیین y روی $\mathbf{X}$ برابر است با

$$R^۲ = \frac{\hat{\beta}'\mathbf{X}'y - n\bar{y}^۲}{y'y - n\bar{y}^۲}, \quad \bar{y} = \frac{۱}{n}\sum_{i=۱}^{n} y_i.$$

همخطی ناقص باعث می‌شود که دترمینان ماتریس $\mathbf{X}'\mathbf{X}$ بسیار کوچک شود و در نتیجه ماتریس $(\mathbf{X}'\mathbf{X})^{-۱}$، دارای اعضایی باشند که از لحاظ قدر مطلق بسیار بزرگ هستند و به عبارت دیگر، واریانس برآوردگر $\hat{\beta}$ بسیار بزرگ می‌شود. در این‌صورت، $\hat{\beta}$ برآورد دقیقی از β نمی‌باشد. به عبارت دیگر دقت برآورد ضرایب رگرسیون بسیار کم می‌شود. علاوه بر این، $\hat{\beta}$ ناپایدار است. یعنی با تغییر کوچکی در مقادیر متغیرهای پیش‌بین، مقادیر موجود در $\hat{\beta}$ تغییرات زیادی پیدا می‌کنند و در نتیجه محاسبه $\hat{\beta}$ توسط دو رایانه متفاوت (با دو روش متفاوت) ممکن است جواب‌های کاملاً متفاوتی به‌دست دهد. همچنین در صورت وجود همخطی کامل ماتریس $\mathbf{X}'\mathbf{X}$ وارون پذیر نبوده (زیرا این ماتریس تمام رتبه[۳۷] نمی‌شود) و در نتیجه برای برآورد ضرایب جواب یکتا وجود نخواهد داشت. بنابراین مهم است بدانیم که چه‌وقت همخطی وجود دارد تا مواظب پیامدهای ممکن آن باشیم. در مورد آثار و شاخص‌های همخطی به روزبه (۱۳۹۰) و نجاریان (۱۳۹۱) مراجعه کنید.

---

[۳۵] Exact collinearity

[۳۶] Partial collinearity

[۳۷] Full rank

۳.۲.۱ روش‌هایی برای برخورد با همخطی چندگانه

روش‌های متعددی برای برخورد با مشکلات ناشی از همخطی چندگانه پیشنهاد شده است. روش‌های کلی شامل جمع آوری داده‌های اضافی، تخصیص مدل مجدد و بکارگیری روش‌های برآورد غیر از کمترین توان‌های دوم که به طور ویژه طراحی شده اند برای از بین بردن مشکلات ایجاد شده به لحاظ همخطی چندگانه است.

وقتی که روش کمترین توان‌های دوم[۳۸] در مورد داده‌های غیر متعامد به‌کار گرفته می‌شود، معمولاً برای ضرایب رگرسیون برآوردهای خیلی ضعیف به‌دست می‌آید. گفته می‌شود که واریانس برآوردهای کمترین توان‌های دوم ضرایب رگرسیون ممکن است در حد قابل توجهی افزایش یابد و طول بردار برآورد کمترین توان‌های دوم پارامتر به طور متوسط خیلی زیاد است. این بدین معنی است که قدر مطلق برآوردهای کمترین توان‌های دوم خیلی بزرگ می‌باشد و لذا بسیار ناپایدار می‌باشند، بدین معنی که با ارائه نمونه‌های متفاوت، اندازه‌ها و علامت‌ها در حد قابل توجهی تغییر خواهند کرد.

قضیه ۱.۲.۱ این اطمینان را می‌دهد که برآوردگر کمترین توان‌های دوم در کلاس برآوردگرهای خطی نااریب[۳۹] دارای می‌نیمم واریانس است، اما تضمینی وجود ندارد که این واریانس کوچک‌ترین باشد. اگر واریانس $\hat{\beta}$ بزرگ باشد در نتیجه فاصله اطمینان برای β عریض می‌شود که در این‌صورت برآورد نقطه‌ای $\hat{\beta}$ بسیار ناپایدار است. یک راه کوچک جلوه دادن این مسئله، کنارگذاشتن این ضرورت است که بایستی برآوردگر β نااریب باشد.

فرض کنیم که بتوان یک برآوردگر اریب $\hat{\beta}^*$ برای β پیدا کرد که دارای واریانس کمتر از واریانس برآوردگر نااریب $\hat{\beta}$ باشد. میانگین توان‌های دوم خطای[۴۰] (MSE)، برآوردگر $\hat{\beta}^*$ بنا به تعریف ب.۰.۰.۳ به صورت زیر است

$$MSE\left(\hat{\beta}^*\right) = E\left(\hat{\beta}^* - \beta\right)^2 = Var\left(\hat{\beta}^*\right) + \left[E(\hat{\beta}^*) - \beta\right]^2 \qquad (۱۶.۱)$$

یا

$$MSE(\hat{\beta}^*) = Var(\hat{\beta}^*) + \left(Bias(\hat{\beta}^*)\right)^2 .$$

توجه داریم که MSE برابر با امید ریاضی توان دوم فاصله $\hat{\beta}^*$ تا β می‌باشد. با پذیرفتن مقدار کمی اریبی برای $\hat{\beta}^*$، واریانس $\hat{\beta}^*$ می‌تواند طوری کوچک شود که MSE برای $\hat{\beta}^*$ از واریانس برآوردگر نااریب $\hat{\beta}$ کمتر باشد. باید توجه داشت که نمی‌توان همزمان هم اریبی و هم واریانس را کاهش داد. با کاستن یکی، دیگری افزایش می‌یابد. به این خاصیت مبادله پایاپای اریبی و واریانس می‌گویند که در این موارد معمولاً مقدار MSE را کمینه می‌کنند.

با به‌کارگیری برآوردگر اریب برای β بایستی فاصله اطمینان β باریک‌تر شود. واریانس کوچک برای برآوردگر اریب نیز نتیجه می‌دهد که $\hat{\beta}^*$ برآوردگری پایدارتر از برآوردگر نااریب $\hat{\beta}$ برای β می‌باشد.

[۳۸] Least square

[۳۹] linear unbiased estimators

[۴۰] Mean square error

۳.۱ رگرسیون ریج

وقوع همخطی چندگانه بین متغیرهای پیش‌بین در تحلیل رگرسیون خطی ممکن است باعث ناپایداری شدید در برآوردهای کمترین توان‌های دوم پارامترهای رگرسیونی شود، به این معنی که بزرگی و علامت بردار پارامترها در نمونه‌های مختلف به‌طور قابل ملاحظه‌ای بی ثبات خواهد بود که در نتیجه‌ی آن، برآورد کمترین توان‌های دوم به‌دست آمده، قابل اعتماد نخواهند بود.
به طور کلی می‌توان برای بوجود آمدن همخطی در داده‌ها چهار عامل را در نظر گرفت: ۱) طریقه جمع‌آوری داده‌ها ۲) قیدهای در نظر گرفته شده بر روی مدل ۳) نوع مدل ۴) تعداد زیاد متغیرها در مدل.
تأثیر همخطی در داده‌ها را می‌توان به صورت زیر در نظر گرفت.
مدل $y = \beta_1 X_1 + \beta_2 X_2 + \epsilon$ را در نظر بگیرید. با استفاده از روابط

$$S_y = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \bar{y}_i)^2}{n-1}}, \qquad y'_i = \frac{1}{\sqrt{n-1}}\left(\frac{y_i - \bar{y}_k}{S_y}\right)$$

$$S_k = \sqrt{\frac{\sum_{i=1}^{n}(x_{ik} - \bar{x}_k)^2}{n-1}}, \qquad X'_{ik} = \frac{1}{\sqrt{n-1}}\left(\frac{x_{ik} - \bar{x}_k}{S_k}\right)$$

$$\sum_{i=1}^{n}(X'_{i1})^2 = \sum_{i=1}^{n}\left(\frac{x_{i1} - \bar{x}_1}{\sqrt{n-1}s_1}\right)^2, \qquad \sum_{i=1}^{n}X'_{i1}X'_{i2} = \sum_{i=1}^{n}\left(\frac{x_{i1} - \bar{x}_1}{\sqrt{n-1}s_1}\right)\left(\frac{x_{i2} - \bar{x}_2}{\sqrt{n-1}s_2}\right)$$

که در آن s_1 و s_2 به ترتیب انحراف معیار X_1 و X_2 هستند، داده‌ها را استاندارد می‌کنیم. بنابراین معادله نرمال به صورت زیر می‌باشد

$$(\mathbf{X}'\mathbf{X})\hat{\beta} = \mathbf{X}'y.$$

به عبارتی

$$\begin{bmatrix} 1 & r_{12} \\ r_{21} & 1 \end{bmatrix}\begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix} = \begin{bmatrix} r_1 y \\ r_2 y \end{bmatrix}$$

که در آن ضرایب همبستگی بین متغیرها به صورت زیر تعریف می‌شود

$$r_{21} = \mathrm{Corr}(x_1, x_2), \qquad r_{jy} = \mathrm{Corr}(x_j, y), \quad j = 1, 2.$$

در نتیجه، برآورد کمترین توان‌های دوم برابر است با

$$\hat{\beta} = \begin{bmatrix} \frac{1}{1-r_{12}^2} & \frac{-r_{12}}{1-r_{12}^2} \\ \frac{-r_{21}}{1-r_{21}^2} & \frac{1}{1-r_{12}^2} \end{bmatrix}\begin{bmatrix} r_1 y \\ r_2 y \end{bmatrix}$$

بنابراین

$$\begin{cases} \hat{\beta}_1 = \dfrac{r_1 y - r_{12} r_2 y}{1 - r_{12}^2} \\ \hat{\beta}_2 = \dfrac{r_2 y - r_{12} r_1 y}{1 - r_{12}^2}. \end{cases}$$

حال اگر X_1 و X_2 دارای همبستگی شدیدی باشند، آنگاه درایه‌های $\mathbf{C}^{-1} = (\mathbf{X}'\mathbf{X})^{-1}$ به سمت $\pm\infty$ میل می‌کند، و چون $\mathrm{Cov}(\hat{\beta}) = \sigma^2\mathbf{C}^{-1}$ یعنی واریانس‌ها و کوواریانس‌های $\hat{\beta}$ بسیار زیاد می‌شود و این

بدان معنی است که نمونه‌های متفاوت برآودهایی متفاوت برای $\beta_۱$ و $\beta_۲$ تولید می‌کنند. وجود همخطی علاوه بر مشکل فوق برآورد β_j را به روش فوق بیشتر از آن‌چه باید باشد، نشان می‌دهد. اگر توان دوم تفاضل β از مقدار واقعی آن را به صورت زیر در نظر بگیریم

$$\begin{aligned} \mathbf{L}_۱^۲ &= (\hat{\beta}-\beta)'(\hat{\beta}-\beta) \\ E(\mathbf{L}_۱^۲) &= E[(\hat{\beta}-\beta)'(\hat{\beta}-\beta)] = \sum_{j=۱}^{p} E(\hat{\beta}_j-\beta_j)^۲. \end{aligned} \tag{۱۷.۱}$$

از طرفی داریم

$$\begin{aligned} \sum_{j=۱}^{p} \mathrm{Var}(\hat{\beta}_j) &= \sigma^۲ \mathrm{tr}[(\mathbf{X}'\mathbf{X})^{-۱}] \\ &= \sigma^۲ \sum_{j=۱}^{p} \frac{۱}{\lambda_j} \end{aligned}$$

که در آن λ_j ها مقادیر ویژه ماتریس $\mathbf{C} = \mathbf{X}'\mathbf{X}$ می‌باشند.

در صورت وجود همخطی، حداقل یکی از مقادیر ویژه بسیار کوچک خواهد بود و در این‌صورت امید ریاضی توان دوم فاصله خیلی بزرگ می‌شود. بنابراین به‌طور متوسط برآورد کمترین توان‌های دوم از مقدار واقعی آن خیلی دور خواهد بود. با توجه به رابطه زیر

$$\begin{aligned} E(L_۱^۲) &= E[(\hat{\beta}-\beta)'(\hat{\beta}-\beta)] \\ &= E(\hat{\beta}'\hat{\beta}) - \beta'\beta \\ &= \sigma^۲ \sum_{j=۱}^{p} \frac{۱}{\lambda_j}. \end{aligned}$$

در صورت وجود همخطی برآورد کمترین توان‌های دوم برای β به‌طور متوسط دارای طولی بزرگ‌تر می‌باشد. یک روش برای رفع این مشکل یافتن برآوردگری است که اریب، اما میانگین خطایی کم‌تر نسبت به برآوردگر نااریب $\hat{\beta}$ دارد. این برآوردگر را $\hat{\beta}^*$ می‌نامیم. با کوچک در نظر گرفتن میزان اریبی $\hat{\beta}^*$، واریانس $\hat{\beta}^*$ را می‌توان کوچک کرد به‌طوری که MSE، $\hat{\beta}^*$ کمتر از واریانس برآوردگر نااریب $\hat{\beta}$ شود.

روش‌های مختلفی به منظور به‌دست آوردن برآوردگرهای اریب ضرایب رگرسیون، توسعه داده شده است، که یکی از این روش‌ها رگرسیون ریج[۴۱] است که توسط هورل و کنارد[۴۲] در سال ۱۹۷۰ پیشنهاد شد.

مدل رگرسیون خطی چندگانه (۱.۱) را در نظر بگیرید. فرض می‌کنیم که دو یا تعداد بیشتری از $\mathbf{X}$ ها وابسته خطی باشند. پس مدل دارای مشکل همخطی چندگانه است. اگر $\mathbf{\Lambda}$ یک ماتریس $p \times p$ قطری باشد که اعضای قطر اصلی آن مقادیر ویژه $(\lambda_۱, \lambda_۲, \cdots, \lambda_p)$ ماتریس $\mathbf{X}'\mathbf{X}$ می‌باشند و $\mathbf{T}$ ماتریس

[۴۱] Ridge regression

[۴۲] Hoerl and Kennard

متناظر متعامد از بردارهای ویژه باشد (بنا به تعریف آ.۰.۱۶)، آنگاه

$$\mathbf{T}'\mathbf{X}'\mathbf{X}\mathbf{T} = \mathbf{\Lambda}.$$

با فرض $\mathbf{Z} = \mathbf{X}\mathbf{T}$ و $\gamma = \mathbf{T}'\beta$ صورت کانونی[۴۳] مدل خطی (۱.۱) عبارتست از

$$\begin{aligned} y &= \mathbf{X}\beta + \epsilon \\ &= (\mathbf{Z}\mathbf{T}')(\mathbf{T}\gamma) + \epsilon \\ &= \mathbf{Z}\gamma + \epsilon. \end{aligned} \tag{۱۸.۱}$$

در این‌صورت

$$\begin{aligned} \mathbf{Z}'\mathbf{Z} &= (\mathbf{X}\mathbf{T})'(\mathbf{X}\mathbf{T}) \\ &= (\mathbf{T}'\mathbf{X}')(\mathbf{X}\mathbf{T}) \\ &= \mathbf{T}'(\mathbf{X}'\mathbf{X})\mathbf{T} \\ &= \mathbf{\Lambda}. \end{aligned} \tag{۱۹.۱}$$

برآوردگر کمترین توان‌های دوم معمولی به صورت زیر به‌دست می‌آید

$$\begin{aligned} \hat{\gamma}_{OLSE} &= (\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'y \\ &= \mathbf{\Lambda}^{-1}\mathbf{Z}'y. \end{aligned} \tag{۲۰.۱}$$

همچنین، بردار برآورد پارامترهای اولیه به صورت زیر است

$$\hat{\beta} = \mathbf{T}\hat{\gamma}_{OLSE}.$$

برآوردگر ریج، برآوردگری است که از جایگزینی $\mathbf{\Lambda}$ با $(\mathbf{\Lambda} + k\mathbf{I})$ در رابطه (۲۰.۱) به دست می آید که در آن $k > 0$ مقداری ثابت است. این برآوردگر که به پارامتر k بستگی دارد، با افزودن این پارامتر به ماتریس $\mathbf{Z}'\mathbf{Z}$ باعث بهبود وضعیت آن می‌شود

$$\hat{\gamma}(k) = (\mathbf{\Lambda} + k\mathbf{I})^{-1}\mathbf{Z}'y$$

که در آن $k > 0$ با تحلیل‌های مختلف عددی به‌دست می‌آید. در واقع برآوردگر ریج یک تبدیل خطی از برآوردگر کمترین توان‌های دوم است

$$\begin{aligned} \hat{\gamma}(k) &= (\mathbf{\Lambda} + k\mathbf{I})^{-1}\mathbf{Z}'y \\ &= (\mathbf{\Lambda} + k\mathbf{I})^{-1}\mathbf{\Lambda}\hat{\gamma}_{OLSE} \\ &= \mathbf{T}_k\hat{\gamma}_{OLSE} \end{aligned} \tag{۲۱.۱}$$

که در آن $\mathbf{T}_k = (k\mathbf{\Lambda}^{-1} + \mathbf{I})^{-1}$. باید توجه داشت که برآوردگر ریج نسبت به پارامتر k غیر خطی است.

[۴۳] Canonical form

در حقیقت برآوردگر ریج $\hat{\beta}(k)$ با حل تغییر یافته معادلات نرمال به صورت زیر محاسبه می شود.

$$(\mathbf{X}'\mathbf{X}+k\mathbf{I})\hat{\beta}(k)=\mathbf{X}'y.$$

در نتیجه

$$\hat{\beta}(k)=(\mathbf{X}'\mathbf{X}+k\mathbf{I})^{-۱}\mathbf{X}'y. \qquad (۲۲.۱)$$

در واقع این معادلات حاصل مینیمم‌سازی ترکیبی خطی از کمترین توان‌های دوم خطا و محدودیت اعمال‌شده بر روی فضای پارامتر به صورت $\beta\beta'\leq\phi^{۲}$ است. یعنی مینیمم کردن A در عبارت زیر

$$A(\beta,\phi)=(y-\mathbf{X}\beta)'(y-\mathbf{X}\beta)+k(\beta'\beta-\phi^{۲})$$

که در آن $k\geq ۰$ ثابتی است که توسط تحلیل‌گر انتخاب می‌شود. توجه داریم که اگر $k=۰$ برآوردگر ریج همان برآوردگر کمترین توان‌های دوم است. روش فوق به دلیل شباهت ریاضیات آن به روش تحلیل ریج که قبل از این توسط هورل و کنارد (۱۹۵۹) برای تشریح رفتار سطح پاسخ درجه دوم مورد استفاده قرار گرفت، رگرسیون ریج نامیده می‌شود.

امید ریاضی $\hat{\beta}(k)$ عبارت است از

$$\mathrm{E}\left(\hat{\beta}(k)\right)=E[\mathbf{T}_k\hat{\beta}]=\mathbf{T}_k\beta. \qquad (۲۳.۱)$$

بنابراین $\hat{\beta}(k)$ یک برآوردگر اریب β می‌باشد. همچنین ماتریس واریانس- کوواریانس به صورت زیر است

$$\mathrm{Var}\left(\hat{\beta}(k)\right)=\sigma^{۲}(\mathbf{X}'\mathbf{X}+k\mathbf{I})^{-۱}\mathbf{X}'\mathbf{X}(\mathbf{X}'\mathbf{X}+k\mathbf{I})^{-۱}. \qquad (۲۴.۱)$$

میانگین توان‌های دوم خطای برآوردگر ریج عبارت است از

$$\begin{aligned}\mathrm{MSE}\left(\hat{\beta}(k)\right)&=Var\left(\hat{\beta}(k)\right)+\left(Bias(\hat{\beta}(k))\right)^{۲} \qquad (۲۵.۱)\\&=\sigma^{۲}tr\left[(\mathbf{X}'\mathbf{X}+k\mathbf{I})^{-۱}\mathbf{X}'\mathbf{X}(\mathbf{X}'\mathbf{X}+k\mathbf{I})^{-۱}\right]+k^{۲}\beta'(\mathbf{X}'\mathbf{X}+k\mathbf{I})^{-۲}\beta\\&=\sigma^{۲}\sum_{j=۱}^{p}\frac{\lambda_j}{(\lambda_j+k)^{۲}}+k^{۲}\beta'(\mathbf{X}'\mathbf{X}+k\mathbf{I})^{-۲}\beta\end{aligned}$$

که در آن $\lambda_۱,\lambda_۲,\cdots,\lambda_p$ مقادیر ویژه $\mathbf{X}'\mathbf{X}$ می‌باشند. اولین جمله طرف راست (۲۵.۱) مجموع واریانس‌های پارامترهای درون $\hat{\beta}(k)$ و جمله دوم توان دوم اریبی می‌باشد. اگر $k>۰$، باید توجه داشت که اریبی درون $\hat{\beta}(k)$ با k افزایش می‌یابد؛ هرچند با افزایش k ، واریانس کاهش می‌یابد.

ایده‌ی رگرسیون ریج این است که مقداری از k را انتخاب کند که کاهش در واریانس بزرگتر از افزایش در توان دوم اریبی باشد. اگر چنین امری بتواند انجام شود، میانگین توان دوم خطای برآوردگر ریج $\hat{\beta}(k)$ کمتر از واریانس برآوردگر کمترین توان‌های دوم $\hat{\beta}$ خواهد بود. هورل و کنارد (۱۹۷۰) ثابت کردند، اگر $\beta'\beta$ کران‌دار باشد مقداری غیر صفر از k وجود دارد که برای آن $MSE(\hat{\beta}(k))$ از واریانس برآوردگر کمترین توان‌های دوم $\hat{\beta}$ کوچکتر است. مقدار مثبتی برای k وجود دارد که به ازای آن برآوردهای ریج نسبت به تغییرات کم در برآورد داده‌ها پایدار خواهند بود.

اثر ریج یک نمودار اعضای $\hat{\beta}(k)$ در مقابل k برای مقادیری از k می‌باشد که معمولاً در فاصله‌ی ۱–۰

قرار دارند. مارکوارت و سنی[۴۴](۱۹۷۵) به‌کارگیری تا ۲۵ مقدار k را پیشنهاد می‌کنند که تقریباً به صورت لگاریتمی روی فاصله‌ی (۱ , ۰) قرار دارند. اگر همخطی شدید باشد، ناپایداری در ضرایب رگرسیون به لحاظ اثر ریج بدیهی خواهد بود. با افزایش k بعضی از برآوردهای ریج به صورت نمایشی تغییر خواهند کرد. به ازای بعضی مقادیر k برآوردهای ریج پایدار خواهند بود. هدف انتخاب مقدار کوچک قابل قبول از k است که در آن برآوردهای ریج $\hat{\beta}(k)$ پایدار باشند. امید است که این کار مجموعه‌ای از برآوردها را با MSE کوچکتر از برآوردهای کمترین توان‌های دوم به‌دست دهد. برای آگاهی بیشتر به مونتگومری[۴۵] (۱۹۹۲) و چاترجی[۴۶] و همکاران (۱۹۳۸) مراجعه کنید.

۴.۱ برآوردگر نوع لیو

از آن‌جایی که برآوردگر ریج ارائه شده توسط هورل وکنارد (۱۹۷۰) نسبت به پارامتر ریج، برآوردگری غیر خطی است و همچنین محاسبه پارامتر ریج بهینه نیاز به محاسبات عددی دارد، لیو (۱۹۹۳) برآوردگری بر پایه برآوردگر کمترین توان‌های دوم برای حل مشکل همخطی ارائه کرد که نسبت به پارامتر انقباضی خطی باشد، که در این‌صورت بهینه سازی این برآوردگر بر اساس پارامتر آن راحت‌تر بود.
برای رفع مشکل همخطی ناقص، لیو (۱۹۹۳)، برآوردگر استاین (۱۹۵۶) را با برآوردگر رگرسیون معمولی ریج ترکیب نمود و برآوردگری به صورت زیر را پیشنهاد کرد که پس از آن به دلیل استفاده مکرر از آن به برآوردگر نوع لیو[۴۷](LE)، معروف شد

$$\hat{\gamma}(d) = (\mathbf{\Lambda} + \mathbf{I})^{-۱}(\mathbf{Z}'y + d\hat{\gamma}_{OLSE}) \qquad (۲۶.۱)$$

که در آن $۰ < d < ۱$ پارامتر اریبی[۴۸] نام دارد. این برآوردگر ترکیب خطی از برآوردگر کمترین توان‌های دوم است، به عبارتی داریم

$$\hat{\gamma}(d) = (\mathbf{\Lambda} + \mathbf{I})^{-۱}\mathbf{Z}'y + d(\mathbf{\Lambda} + \mathbf{I})^{-۱}\hat{\gamma}_{OLSE}.$$

در این حالت می‌توانیم بنویسیم

$$\begin{aligned}\hat{\gamma}(d) &= (\mathbf{\Lambda} + \mathbf{I})^{-۱}(\mathbf{Z}'y + d\hat{\gamma}_{OLSE}) \\ &= (\mathbf{\Lambda} + \mathbf{I})^{-۱}(\mathbf{\Lambda} + d\mathbf{I})\hat{\gamma}_{OLSE} \\ &= \mathbf{F}_d\hat{\gamma}_{OLSE} \qquad (۲۷.۱)\end{aligned}$$

که در آن $\mathbf{F}_d = (\mathbf{\Lambda} + \mathbf{I})^{-۱}(\mathbf{\Lambda} + d\mathbf{I})$.

[۴۴] Marquardt and Snee
[۴۵] Montgomery
[۴۶] Chatterjee
[۴۷] Liu-type estimator
[۴۸] Biasing parameter

۱.۴.۱ برآوردگر لیو تعمیم یافته

آکدنیز و کاچیرانلار[۴۹](۱۹۹۵) با جایگذاری پارامتر اریبی d به یک ماتریس اریبی مانند $\mathbf{D} > 0$ برآوردگر لیو (۱۹۹۳) را تعمیم دادند و آن را برآوردگر لیو تعمیم یافته[۵۰](GLE)، نامیدند. برآوردگر لیو تعمیم یافته دارای صورت فرمولی زیر است

$$\hat{\gamma}(\mathbf{D}) = (\mathbf{\Lambda} + \mathbf{I})^{-1}(\mathbf{Z}'y + \mathbf{D}\hat{\gamma}_{OLSE}). \tag{۲۸.۱}$$

در برآوردگر آکدنیز و کاچیرانلار (۱۹۹۵) ماتریس اریبی $\mathbf{D}$ یک ماتریس قطری به صورت

$$\mathbf{D} = \begin{bmatrix} d_1 & 0 & \cdots & 0 \\ 0 & \ddots & \cdots & 0 \\ \vdots & \cdots & \ddots & 0 \\ 0 & \cdots & \cdots & d_p \end{bmatrix}$$

بود که $0 < d_i < 1, i = 1, 2, \cdots, p$. می‌توان نشان داد برآوردگر لیو تعمیم یافته نیز یک ترکیب خطی از برآوردگر کمترین توان‌های دوم است و نمایش خلاصه‌تر آن عبارتست از

$$\begin{aligned}
\hat{\gamma}(\mathbf{D}) &= (\mathbf{\Lambda} + \mathbf{I})^{-1}(\mathbf{Z}'y + \mathbf{D}\hat{\gamma}_{OLSE}) \\
&= (\mathbf{\Lambda} + \mathbf{I})^{-1}(\mathbf{\Lambda}\hat{\gamma}_{OLSE} + \mathbf{D}\hat{\gamma}_{OLSE}) \\
&= (\mathbf{\Lambda} + \mathbf{I})^{-1}(\mathbf{\Lambda} + \mathbf{D})\hat{\gamma}_{OLSE} \\
&= \left[(\mathbf{\Lambda} + \mathbf{I})^{-1}\mathbf{\Lambda} + (\mathbf{\Lambda} + \mathbf{I})^{-1}\mathbf{D}\right]\hat{\gamma}_{OLSE} \\
&= \left[(\mathbf{\Lambda} + \mathbf{I})^{-1}\mathbf{\Lambda} + (\mathbf{\Lambda} + \mathbf{I})^{-1}(\mathbf{I} - \mathbf{I} + \mathbf{D})\right]\hat{\gamma}_{OLSE} \\
&= \left[(\mathbf{\Lambda} + \mathbf{I})^{-1}\mathbf{\Lambda} - (\mathbf{\Lambda} + \mathbf{I})^{-1}(\mathbf{I} - \mathbf{I} - \mathbf{D})\right]\hat{\gamma}_{OLSE} \\
&= \left[(\mathbf{\Lambda} + \mathbf{I})^{-1}\mathbf{\Lambda} - (\mathbf{\Lambda} + \mathbf{I})^{-1}(\mathbf{I} - \mathbf{D}) + (\mathbf{\Lambda} + \mathbf{I})^{-1}\right]\hat{\gamma}_{OLSE} \\
&= \left[(\mathbf{\Lambda} + \mathbf{I})^{-1}(\mathbf{\Lambda} + \mathbf{I}) - (\mathbf{\Lambda} + \mathbf{I})^{-1}(\mathbf{I} - \mathbf{D})\right]\hat{\gamma}_{OLSE} \\
&= \left[\mathbf{I} - (\mathbf{\Lambda} + \mathbf{I})^{-1}(\mathbf{I} - \mathbf{D})\right]\hat{\gamma}_{OLSE} \\
&= \mathbf{F}_{\mathbf{D}}\hat{\gamma}_{OLSE}
\end{aligned} \tag{۲۹.۱}$$

که در آن

$$\begin{aligned}
\mathbf{F}_{\mathbf{D}} &= (\mathbf{\Lambda} + \mathbf{I})^{-1}(\mathbf{\Lambda} + \mathbf{D}) \\
&= (\mathbf{\Lambda} + \mathbf{I})^{-1}\mathbf{\Lambda} + (\mathbf{\Lambda} + \mathbf{I})^{-1}\mathbf{D} \\
&= (\mathbf{\Lambda} + \mathbf{I})^{-1}\mathbf{\Lambda} - (\mathbf{\Lambda} + \mathbf{I})^{-1}(\mathbf{I} - \mathbf{D}) + (\mathbf{\Lambda} + \mathbf{I})^{-1} \\
&= (\mathbf{\Lambda} + \mathbf{I})^{-1}(\mathbf{\Lambda} + \mathbf{I}) - (\mathbf{\Lambda} + \mathbf{I})^{-1}(\mathbf{I} - \mathbf{D}) \\
&= \mathbf{I} - (\mathbf{\Lambda} + \mathbf{I})^{-1}(\mathbf{I} - \mathbf{D}).
\end{aligned}$$

[۴۹] Akdenize and Kaciranlar

[۵۰] Generalized liu-type estimator

مقدار اریبی را برای این برآوردگر به‌دست می‌آوریم. برای این منظور می‌توان نوشت

$$\begin{aligned} \text{bias}(\hat{\gamma}(\mathbf{D})) &= E(\hat{\gamma}(\mathbf{D})) - \gamma \\ &= [\mathbf{I} - (\mathbf{\Lambda} + \mathbf{I})^{-1}(\mathbf{I} - \mathbf{D})]E(\hat{\gamma}_{OLSE}) - \gamma \end{aligned} \tag{۳۰.۱}$$

که بنا به رابطه‌های (۲۰.۱) و (۲۹.۱) داریم

$$\begin{aligned} \text{bias}(\hat{\gamma}(\mathbf{D})) &= [\mathbf{I} - (\mathbf{\Lambda} + \mathbf{I})^{-1}(\mathbf{I} - \mathbf{D})]\gamma - \gamma \\ &= -(\mathbf{\Lambda} + \mathbf{I})^{-1}(\mathbf{I} - \mathbf{D})\gamma \\ &= (\mathbf{F_D} - \mathbf{I})\gamma. \end{aligned} \tag{۳۱.۱}$$

هرگاه $d_1 = d_2 = \ldots = d_p = d, 0 < d < 1$ ، برآوردگر لیو تعمیم یافته به برآوردگر نوع لیو ساده می‌شود.
با توجه به (۲۹.۱) و (۲۰.۱) ماتریس واریانس- کوواریانس برآوردگر لیو تعمیم‌یافته به صورت زیر محاسبه می‌شود

$$\begin{aligned} \text{cov}(\hat{\gamma}(\mathbf{D})) &= \text{cov}(\mathbf{F_D}\hat{\gamma}_{OLSE}) \\ &= \mathbf{F_D}\text{cov}(\hat{\gamma}_{OLSE})F'_D \\ &= \sigma^2\mathbf{F_D}\mathbf{\Lambda}^{-1}\mathbf{F'_D}. \end{aligned} \tag{۳۲.۱}$$

بنا به تعریف ب.۰.۰.۳۰ و با توجه به روابط (۳۲.۱) و (۳۱.۱) مقدار ماتریس میانگین توان دوم خطا[۵۱] ($MSEM$) برای برآوردگر لیو تعمیم یافته به صورت زیر قابل محاسبه است

$$\begin{aligned} MSEM(\hat{\gamma}(\mathbf{D})) &= \text{cov}(\hat{\gamma}(\mathbf{D})) + (\text{bias}(\hat{\gamma}(\mathbf{D})))^2 \\ &= \sigma^2\mathbf{F_D}\mathbf{\Lambda}^{-1}\mathbf{F'_D} + (\mathbf{F_D} - \mathbf{I})\gamma\gamma'(\mathbf{F_D} - \mathbf{I})'. \end{aligned} \tag{۳۳.۱}$$

۲.۴.۱ مثال عددی

استیلن در صنعت پتروشیمی به عنوان ماده اولیه و پیش ماده در سنتز و تهیه مواد شیمیایی مختلف به‌کار می‌رود. از این رو، روش‌های صنعتی زیادی برای تولید انبوه این ماده ابداع شده است. کوچک‌ترین عضو خانواده آلکین‌های هیدروکربنی است که از دو اتم هیدروژن و دو اتم کربن که با هم پیوند سه گانه دارند ساخته شده است. به خاطر وجود این پیوند سه گانه استیلن جزو مواد شیمیایی اشباع‌نشده شناخته می‌شود.
داده‌های جدول ۱.۱ نشان‌دهنده درصد تبدیل n- هپتان به استیلن با سه متغیر توضیحی شامل: درجه

[۵۱] Mean square error matrix

جدول ۱.۱: داده‌های در صد تبدیل n- هپتان به استیلن

در صد تبدیل n- هپتان به استیلن	درجه رآکتور	نسبت $H_۲$ به n- هپتان	زمان اتصال
۴۹	۱۳۰۰	۷/۵	۰/۰۱۲
۵۰/۲	۱۳۰۰	۹	۰/۰۱۲
۵۰/۵	۱۳۰۰	۱۱	۰/۰۱۱۵
۴۸/۵	۱۳۰۰	۱۳/۵	۰/۰۱۳
۴۷/۵	۱۳۰۰	۱۷	۰/۰۱۳۵
۲۸	۱۳۰۰	۲۳	۰/۰۱۲
۳۱/۵	۱۲۰۰	۵/۳	۰/۰۴
۳۱/۵	۱۲۰۰	۷/۵	۰/۰۳۸
۳۴/۵	۱۲۰۰	۱۱	۰/۰۳۲
۳۵	۱۲۰۰	۱۳/۵	۰/۰۲۶
۳۸	۱۲۰۰	۱۷	۰/۰۳۴
۳۸/۵	۱۲۰۰	۲۳	۰.۰۴۱
۱۵	۱۱۰۰	۵/۳	۰/۰۸۴
۱۷	۱۱۰۰	۷/۵	۰/۰۹۸
۲۰/۵	۱۱۰۰	۱۱	۰/۹۲
۲۹/۵	۱۱۰۰	۱۷	۰/۰۸۶

رآکتور، زمان اتصال، نسبت H به n- هپتان می‌باشد. (نجاریان۱۳۹۰ را ببینید.)
الگوی مدل پارامتری به صورت زیر است

$$y = \beta_۰ + \beta_۱ X_{۱i} + \beta_۲ X_{۲i} + \beta_۳ X_{۳i} + \epsilon_i$$

که در آن y نشان‌دهنده درصد تبدیل n- هپتان به استیلن، $X_۱$ درجه رآکتور، $X_۲$ زمان اتصال، $X_۳$ نسبت H به n- هپتان است.

VIF معیاری برای تشخیص همخطی است و از رابطه‌ی زیر محاسبه می‌گردد (پیوست پ را ببینید.)

$$VIF_j = (x'_j x_j)/(x'_j M_j x_j)$$

که در آن

$$\begin{cases} M_j = \mathbf{I}_n - X_j(X'_j X_j)^{-۱} X'_j \\ \mathbf{X} = (x_j, X_j). \end{cases}$$

و x_j مشاهدات مربوط به متغیر x_j و $\mathbf{X}_j$ ماتریس مشاهدات باقیمانده پس از حذف متغیر x_j است.
برای این متغیرها مقادیر VIF به ترتیب به صورت

$$۳۲۵/۷۳,\ ۳۲۹/۱۶۳,\ ۶/۴۸,\ ۱/۵۴$$

جدول ۲.۱: برآورد $\hat{\beta}(d)$ برای مقادیر مختلف d

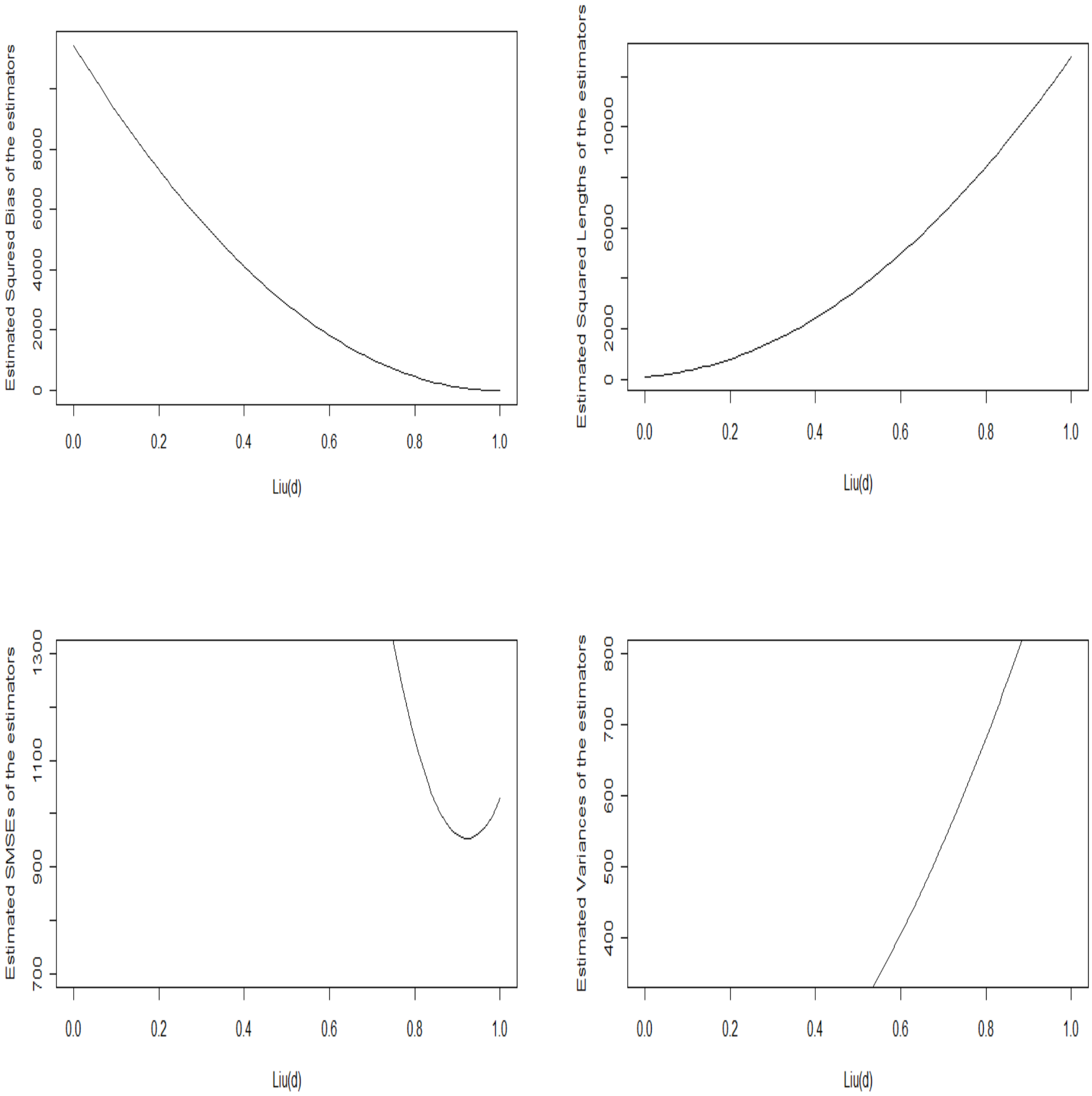
	$d=۰/۵۵$	$d=۰/۷۵$	$d=۱$
β_0	−۶۵/۰۶۱	−۸۶/۴۱۶	−۱۱۳/۱۱
ضریب درجه رآکتور	۰/۰۸۳	۰/۱۰۰۸	۰/۱۲۳
ضریب نسبت H_2 به n− هپتان	۰/۰۰۹	−۰/۰۱۱	−۰/۰۳۳
ضریب زمان اتصال	−۴/۱۲۱	−۲/۹۲۲	−۱/۴۲۴
طول توان‌دوم ضرایب	۴۲۴۹/۹۱۲	۷۴۷۶/۱۹۸	۱۲۷۹۵/۶۴
اثر میانگین توان‌دوم خطا	۲۶۶۴/۴۷۵	۱۳۲۰/۲۷۲	۱۰۲۷/۸۹۲

است. بعضی از VIF ها بزرگتر از ۳۰ هستند که دلالت بر همخطی شدید بین متغیرهاست. در این حالت، ضرایب رگرسیونی بی معنی‌اند بنابراین، استنباط و پیش‌گویی معتبر با استفاده از کمترین توان‌های دوم مشکل است. یک راه حل برای رفع این مشکل استفاده از برآوردگر لیو است. مقدار این برآوردگر را به ازای d های مختلف محاسبه نموده ایم. برآوردها و طول توان دوم ضرایب رگرسیون $(\beta'\beta)$ و مقادیر MSE ها که برابر $\mathrm{trace}(MSEM)$ است به ازای مقادیر مختلف d در جدول ۲.۱ آمده است. برنامه مربوط به این مثال در پیوست ت.۱ آمده است.

برآورد واریانس از رابطه زیر محاسبه می‌شود

$$\hat{\sigma}^2=(y-\mathbf{Z}\hat{\gamma}_{OLSE})'(y-\mathbf{Z}\hat{\gamma}_{OLSE})/(n-p).$$

شکل ۱.۱ عملکرد برآوردگر نوع لیو را در نواحی مختلف نشان می‌دهد که با توجه به جدول ۲.۱ و شکل ۱.۱ با افزایش مقدار d طول توان دوم برآوردگر افزایش می‌یابد و مقدار MSE کاهش پیدا می‌کند. همچنین هرچه مقدار d به سمت ۱ افزایش پیدا می‌کند مقدار توان دوم اریبی برآوردگر کاهش و برآورد واریانس برآوردگرها افزایش می‌یابد.



شکل ۱.۱: عملکرد برآوردگر نوع لیو به ازای مقادیر مختلف d ، $0 < d < 1$

۵.۱ برآوردگر آمیخته

تیل و گلدبرگر[۵۲] (۱۹۶۱) و تیل (۱۹۶۳) روش برآوردگر آمیخته[۵۳] را معرفی کرده‌اند. این روش اطلاع پیشین یا اضافی را به طور مستقیم به جای یک توزیع پیشین به داده‌ها می‌افزاید. برآوردگر آمیخته با مدل رگرسیون معمولی (۱.۱) شروع می‌شود و فرض می‌شود که محقق می‌تواند مجموعه‌ای از $r < p$ محدودیت‌های پیشین را بر روی β در نظر بگیرد، به این‌صورت که

$$a = \mathbf{T}\beta + \delta$$

که در آن $E(\delta) = \mathbf{0}$، $\mathrm{Var}(\delta) = \mathbf{V}$، $\mathbf{T}$ ماتریس $r \times p$ معلوم و ثابت با رتبه r می‌باشد و a یک بردار تصادفی $r \times ۱$ هستند. اگر y و $\mathbf{X}$ را به آن بیفزاییم خواهیم داشت

$$\begin{bmatrix} y \\ a \end{bmatrix} = \begin{bmatrix} \mathbf{X} \\ \mathbf{T} \end{bmatrix} \beta + \begin{bmatrix} \epsilon \\ \delta \end{bmatrix}$$

و با به کار بردن روش کمترین توان‌های دوم، برآوردگر آمیخته نااریب به‌دست می‌آید

$$\hat{\beta}_{ME} = \left(\frac{۱}{\sigma^۲}\mathbf{X}'\mathbf{X} + \mathbf{T}'\mathbf{V}^{-۱}\mathbf{T}\right)^{-۱}\left(\frac{۱}{\sigma^۲}\mathbf{X}'y + \mathbf{T}'\mathbf{V}^{-۱}a\right).$$

حال اگر $\mathbf{T} = \mathbf{A}$ (که در آن $\mathbf{A}'\mathbf{A} = \mathbf{I}$)، $a = \mathbf{0}$ و $\mathbf{V} = \sigma_۱^۲\mathbf{I}$، در نظر بگیریم، آن‌گاه

$$\hat{\beta}_{ME} = (\mathbf{X}'\mathbf{X} + k\mathbf{I})^{-۱}\mathbf{X}'y \equiv \hat{\beta}(k)$$

که در آن $k = \frac{\sigma^۲}{\sigma_۱^۲}$.

برآوردگر آمیخته و رگرسیون ریج را می‌توان از نظر عددی معادل یکدیگر در نظر گرفت. در برآورد آمیخته a یک متغیر تصادفی است در حالی‌که در رگرسیون ریج a ثابتی معین است، که موجب اریبی برآوردگر می‌شود.

خواص و تعمیم‌هایی بر برآوردگر لیو تعمیم یافته توسط چندین محقق دیگر نیز مورد مطالعه و بررسی قرار گرفته است. در این میان یانگ[۵۴] و همکاران (۲۰۰۹) برآوردگر لیو وزنی را معرفی کردند که ترکیبی از برآوردگر آمیخته وزنی و برآوردگر نوع لیو است.

هیبرت و وایجکوان[۵۵](۲۰۰۶) برآوردگر محدودکننده تصادفی را پیشنهاد کردند که ترکیبی از برآوردگر آمیخته با برآوردگر نوع لیو است. یانگ و سو[۵۶](۲۰۰۹) برآوردگر محدود شده لیو را برای بردار پارامترها وقتی محدودیت خطی روی بردار پارامترها فرض می‌شود، در یک مدل رگرسیون خطی معرفی و بهبود بخشیدند. آکدنیز و آکدنیز دوران[۵۷](۲۰۰۹) برآوردگر نوع لیو برای بردار پارامترهای β در یک مدل رگرسیون نیمه پارامتری $Y = \mathbf{X}\beta + f + \epsilon$ معرفی کردند که f تابعی هموار و نامعلوم است.

[۵۲] Theil and Goldberger
[۵۳] Mixed estimation
[۵۴] Yang
[۵۵] Hubert and Wijekoon
[۵۶] Xu
[۵۷] Akdeniz Duran

۶.۱ روش جک‌نایف

روش جک‌نایف در سال ۱۹۴۹ توسط کوئینلی به عنوان روشی برای کاهش اریبی یک برآوردگر معرفی شد. در سال ۱۹۵۸ توکی[۵۸] نام این تکنیک را جک‌نایف نهاد و از آن برای برآورد خطای استاندارد بهره گرفت.

جک‌نایف، یک چاقوی جیبی سوئیسی است که به آسانی و در همه جا قابل استفاده است. مقصود از به‌کارگیری نام جک‌نایف برای این ابزار، ارائه گستره‌ی مفیدی از کاربرد این وسیله برای جایگزینی ابزارهای تخصصی است که ممکن است در دسترس نباشند. توکی این واژه را در علم آمار به عنوان یک رویکرد و روش نیرومندی معرفی نمود که تحت انواع وسیعی از شرایط به‌کار می‌رود. به عنوان مثال می‌توان به کاربرد آن در محاسبه‌ی بازه‌های اطمینان اشاره نمود (نوروزی راد ۱۳۹۰ به نقل از افرون ۱۹۷۹ را ببینید).

مهم‌ترین مزیت این روش عدم نیاز به محاسبات تئوری و برقراری فرض‌های پیچیده است. جک‌نایف روشی ساده برای برآورد اریبی و واریانس برآوردگرها می‌باشد. در این بخش جزئیات این روش را مورد بررسی قرار می‌دهیم.

فرض کنید $\mathbf{X} = (X_۱, X_۲, \cdots, X_n)$ یک نمونه تصادفی ازخانواده توزیع‌های نامعلوم $\mathcal{F}$ به صورت زیر باشد

$$\mathcal{F} = \{F_\theta : \theta \in \Theta \subseteq \mathbb{R}\}.$$

روش جک‌نایف بر اساس نمونه‌هایی است که در هر بار، یکی از مشاهدات، از نمونه اصلی حذف می‌شود. نمونه‌های جک‌نایف را به صورت زیر نشان می‌دهیم

$$X_{(i)} = (X_۱, X_۲, \cdots, X_{i-۱}, X_{i+۱}, \cdots, X_n) \qquad (۳۴.۱)$$

i-امین نمونه جک‌نایف شامل تمام مشاهدات نمونه اصلی به جز i-امین مشاهده است.

اگر $X_{(i)}$، i-امین نمونه جک‌نایف و $\hat{\theta} = S(X)$ برآوردگر معمول θ باشد، آن‌گاه متناظر با هر نمونه جک‌نایف، پاسخ جک‌نایف $\hat{\theta}$ به صورت زیر تعریف می‌شود

$$\hat{\theta}_{(i)} = S_{n-۱,i} = S(X_{(i)}). \qquad (۳۵.۱)$$

اگر $S_n = S_n(X_۱, \cdots, X_n)$ برآوردگری هموار از پارامتر θ باشد، آن‌گاه اریبی S_n به صورت زیر تعریف می‌شود

$$\text{bias}(S_n) = E(S_n) - \theta. \qquad (۳۶.۱)$$

برآورد اریبی جک‌نایف کوئینلی عبارتست از

$$\text{bias}_{jack}(S_n) = (n - ۱)(\bar{S}_n - S_n) \qquad (۳۷.۱)$$

که در آن

$$\bar{S}_n = \frac{۱}{n} \sum_{i=۱}^{n} S_{n-۱,i}.$$

[۵۸] Tukey

قضیه ۱.۶.۱. *اگر برآوردگر جک‌نایف θ را به صورت زیر تعریف کنیم*

$$\begin{aligned} S_{jack} &= S_n - \text{bias}_{jack}(S_n) \\ &= S_n - (n-۱)(\bar{S}_n - S_n) \\ &= nS_n - (n-۱)\bar{S}_n. \end{aligned} \tag{۳۸.۱}$$

آن‌گاه S_{jack} برآوردگری با اریبی کاهش یافته[۵۹] *است. به عبارت دیگر اریبی S_{jack} از اریبی $\hat{\theta} = S_n$ کمتر است.*

برهان. معمولاً برای برآوردگرها داریم (افرن، ۱۹۸۰ و شوکنی[۶۰] و همکاران، ۱۹۷۱)

$$E(S_n) = \theta + \frac{a}{n} + \frac{b}{n^۲} + O(n^{-۳})$$

و

$$\text{bias}(S_n) = \frac{a}{n} + \frac{b}{n^۲} + O(n^{-۳}) \tag{۳۹.۱}$$

که در آن a و b مجهولند و به n بستگی ندارند. (تعاریف O و o در پیوست ب.۲۶.۰ آمده است.)
چون $S_{n-۱,i}$ بر اساس n-۱ داده است، پس می‌توان نوشت

$$\text{bias}(S_{n-۱,i}) = \frac{a}{n-۱} + \frac{b}{(n-۱)^۲} + O((n-۱)^{-۳}).$$

با توجه به تعریف اریبی نتیجه می‌شود

$$\begin{aligned} \text{bias}(\frac{۱}{n}\sum_{i=۱}^{n} S_{n-۱,i}) &= E(\frac{۱}{n}\sum_{i=۱}^{n} S_{n-۱,i}) - \theta \\ &= \frac{۱}{n}E(\sum_{i=۱}^{n} S_{n-۱,i}) - \theta \\ &= \frac{۱}{n}\sum_{i=۱}^{n} E(S_{n-۱,i}) - \theta \\ &= \frac{۱}{n}\left[\sum_{i=۱}^{n} E(S_{n-۱,i}) - n\theta\right] \\ &= \frac{۱}{n}\sum_{i=۱}^{n} [E(S_{n-۱,i}) - \theta] \\ &= \frac{۱}{n}\sum_{i=۱}^{n} \text{bias}(S_{n-۱,i}). \end{aligned} \tag{۴۰.۱}$$

[۵۹] Bias-reduced estimator

[۶۰] Schucany

از هم‌توزیعی $S_{n-1,i}$، $i=1,2,\cdots,n$، و با توجه به رابطه (۴۰.۱)، داریم

$$\begin{aligned}\text{bias}(\bar{S}_n) &= \text{bias}(\frac{1}{n}\sum_{i=1}^{n} S_{n-1,i})\\ &= \frac{1}{n}\sum_{i=1}^{n}\text{bias}(S_{n-1,i})\\ &= \frac{1}{n}.n\text{bias}(S_{n-1,i})\\ &= \text{bias}(S_{n-1,i}).\end{aligned} \tag{۴۱.۱}$$

یعنی اریبی $\bar{S}_n$ با اریبی $S_{n-1,i}$ برابر است. بنابراین

$$\begin{aligned}\hat{b}_{jack} &= E(\text{bias}_{jack}(S_n))\\ &= E\left[(n-1)(\bar{S}_n - S_n)\right]\\ &= (n-1)\left[E(\bar{S}_n) - E(S_n)\right]\\ &= (n-1)\left[E(\bar{S}_n) - \text{bias}(S_n) - \theta\right]\\ &= (n-1)\left[\text{bias}(\bar{S}_n) - \text{bias}(S_n)\right]\\ &= (n-1)\left[(\frac{1}{n-1} - \frac{1}{n})a + (\frac{1}{(n-1)^2} - \frac{1}{n^2})b + O(\frac{1}{n^3})\right]\\ &= \frac{a}{n} + \frac{(2n-1)b}{n^2(n-1)} + O(\frac{1}{n^2}).\end{aligned} \tag{۴۲.۱}$$

لذا

$$\begin{aligned}\text{bias}(S_{jack}) &= E(S_{jack}) - \theta\\ &= E(S_n - \hat{b}_{jack}) - \theta\\ &= E(S_n) - E(\hat{b}_{jack}) - \theta\\ &= \text{bias}(S_n) - E(\hat{b}_{jack})\\ &= -\frac{b}{n(n-1)} + O(\frac{1}{n^2}).\end{aligned} \tag{۴۳.۱}$$

□

بنابراین با مقایسه (۴۳.۱) و (۳۹.۱) واضح است که روش جک‌نایف با حذف جمله درجه اول $\frac{a}{n}$ در $\text{bias}(S_n)$، برآوردگری با اریبی کمتر تولید می‌کند.

تعریف S_{jack} در (۳۸.۱) را می‌توان به صورت زیر نوشت

$$S_{jack} = \frac{1}{n}\sum_{i=1}^{n}\left[nS_n - (n-1)S_{n-1,i}\right]. \tag{۴۴.۱}$$

زیرا

$$\begin{aligned} S_{jack} &= nS_n - (n-1)\bar{S}_n \\ &= nS_n - (n-1)\frac{1}{n}\sum_{i=1}^{n} S_{n-1,i} \\ &= \frac{1}{n}n^2 S_n - (n-1)\frac{1}{n}\sum_{i=1}^{n} S_{n-1,i} \\ &= \frac{1}{n}\sum_{i=1}^{n} nS_n - (n-1)\frac{1}{n}\sum_{i=1}^{n} S_{n-1,i} \\ &= \frac{1}{n}\sum_{i=1}^{n}\left[nS_n - (n-1)S_{n-1,i}\right]. \end{aligned} \tag{۴۵.۱}$$

توکی (۱۹۵۸)، شبه مقادیر جک‌نایف را با $\tilde{S}_{n,i}$ نشان داده و به صورت زیر در نظر بگیرید
$$\tilde{S}_{n,i} = nS_n - (n-1)S_{n-1,i} \quad i = 1, 2, \cdots, n.$$

در این صورت داریم

(الف) شبه مقادیر $\tilde{S}_{n,i}$، $i = 1, 2, \cdots, n$، iid (نمونه تصادفی) هستند.

(ب) واریانس $\tilde{S}_{n,i}$ ها تقریباً برابر است با واریانس $\sqrt{n}S_n$.

با توجه به (الف) و (ب) طبیعی است که می‌توان از واریانس نمونه‌ای $\tilde{S}_{n,1}, \cdots, \tilde{S}_{n,n}$ برای برآورد $Var(\sqrt{n}S_n)$ استفاده کرد. پس می‌توان نوشت

$$\begin{aligned} Var(\sqrt{n}S_n) &= nVar(S_n) \\ &= \frac{1}{n-1}\sum_{i=1}^{n}(\tilde{S}_{n,i} - \frac{1}{n}\sum_{j=1}^{n}\tilde{S}_{n,j})^2 \\ \Rightarrow Var(S_n) &= \frac{1}{n(n-1)}\sum_{i=1}^{n}(\tilde{S}_{n,i} - \frac{1}{n}\sum_{j=1}^{n}\tilde{S}_{n,j})^2. \end{aligned} \tag{۴۶.۱}$$

با استفاده از (۴۴.۱)، $\frac{1}{n}\sum_{i=1}^{n}\tilde{S}_{n,i}$ را برحسب $S_{n-1,i}$ به‌دست می‌آوریم

$$\begin{aligned} \frac{1}{n}\sum_{i=1}^{n}\tilde{S}_{n,i} &= \frac{1}{n}\sum_{i=1}^{n}\left[nS_n - (n-1)S_{n-1,i}\right] \\ &= \frac{1}{n}\left[(\sum_{i=1}^{n} nS_n) - (n-1)\sum_{i=1}^{n} S_{n-1,i}\right] \\ &= \frac{n^2}{n}S_n - \frac{(n-1)}{n}\sum_{i=1}^{n} S_{n-1,i} \\ &= nS_n - \frac{(n-1)}{n}\sum_{i=1}^{n} S_{n-1,i}. \end{aligned} \tag{۴۷.۱}$$

پس داریم

$$
\begin{aligned}
Var_{jack}(S_n) &= \frac{۱}{n(n-۱)} \sum_{i=۱}^{n} \left[\tilde{S}_{n,i} - \frac{۱}{n} \sum_{j=۱}^{n} \tilde{S}_{n,j} \right]^{۲} \\
&= \frac{۱}{n(n-۱)} \sum_{i=۱}^{n} \left[\frac{۱}{n} \sum_{j=۱}^{n} \tilde{S}_{n,j} - \tilde{S}_{n,i} \right]^{۲} \\
&= \frac{۱}{n(n-۱)} \sum_{i=۱}^{n} \left[nS_n - \frac{(n-۱)}{n} \sum_{j=۱}^{n} S_{n-۱,j} - nS_n + (n-۱)S_{n-۱,i} \right]^{۲} \\
&= \frac{۱}{n(n-۱)} \sum_{i=۱}^{n} \left[(n-۱)S_{n-۱,i} - \frac{(n-۱)}{n} \sum_{j=۱}^{n} S_{n-۱,j} \right]^{۲} \\
&= \frac{۱}{n(n-۱)} \sum_{i=۱}^{n} [(n-۱)^{۲} S^{۲}_{n-۱,i} + \frac{(n-۱)^{۲}}{n^{۲}} (\sum_{j=۱}^{n} S_{n-۱,j})^{۲} \\
&- \frac{۲(n-۱)^{۲}}{n} S_{n-۱,j} (\sum_{j=۱}^{n} S_{n-۱,j})] \\
&= \frac{۱}{n(n-۱)} [(n-۱)^{۲} \sum_{i=۱}^{n} S^{۲}_{n-۱,i} + \frac{n(n-۱)^{۲}}{n^{۲}} (\sum_{j=۱}^{n} S_{n-۱,j})^{۲} \\
&- \frac{۲(n-۱)^{۲}}{n} (\sum_{j=۱}^{n} S_{n-۱,j}) \sum_{i=۱}^{n} S_{n-۱,i}] \\
&= \frac{(n-۱)^{۲}}{n(n-۱)} \sum_{i=۱}^{n} \left[S^{۲}_{n-۱,i} + \frac{۱}{n^{۲}} \left(\sum_{j=۱}^{n} S_{n-۱,j} \right)^{۲} - \frac{۲}{n} (\sum_{j=۱}^{n} S_{n-۱,j}) S_{n-۱,i} \right] \\
&= \frac{(n-۱)^{۲}}{n(n-۱)} \sum_{i=۱}^{n} (S_{n-۱,i} - \frac{۱}{n} \sum_{j=۱}^{n} S_{n-۱,j})^{۲} \\
&= \frac{n-۱}{n} \sum_{i=۱}^{n} (S_{n-۱,i} - \frac{۱}{n} \sum_{j=۱}^{n} S_{n-۱,j})^{۲}. \qquad (۴۸.۱)
\end{aligned}
$$

بنابراین برآورد جک‌نایف واریانس برابر است با

$$
Var_{jack}(S_n) = \frac{n-۱}{n} \sum_{i=۱}^{n} (S_{n-۱,i} - \frac{۱}{n} \sum_{j=۱}^{n} S_{n-۱,j})^{۲}. \qquad (۴۹.۱)
$$

در نتیجه برآورد جک‌نایف خطای استاندارد آماره S_n برابر خواهد شد با

$$
\begin{aligned}
se_{jack}(S_n) &= [Var_{jack}(S_n)]^{\frac{۱}{۲}} \\
&= \left[\frac{n-۱}{n} \sum_{i=۱}^{n} (S_{n-۱,i} - \frac{۱}{n} \sum_{j=۱}^{n} S_{n-۱,j})^{۲} \right]^{\frac{۱}{۲}}. \qquad (۵۰.۱)
\end{aligned}
$$

در حالت خاص اگر در نظر بگیریم $\hat{\theta} = S(X) = \bar{X}$، آن‌گاه می‌توان به آسانی نشان داد که

$$se_{jack}(\bar{X}) = \left[\frac{1}{n(n-1)}\sum_{i=1}^{n}(X_i - \bar{X})^2\right]^{\frac{1}{2}}. \tag{۵۱.۱}$$

یعنی اگر $\hat{\theta} = \bar{X}$، آن‌گاه برآورد جک‌نایف خطای استاندارد، با برآورد نااریب آن برابر است. این مطلب را در قضیه ۲.۶.۱ ثابت می‌کنیم.

قضیه ۲.۶.۱. *برآورد جک‌نایف خطای استاندارد $\bar{X}$ برابر برآورد خطای استاندارد واقعی آن است.*

برهان. می‌دانیم که برآورد واریانس $\bar{X}$ برابر است با

$$\begin{aligned}\frac{S^2}{n} &= \frac{\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2}{n}\\ &= \frac{1}{n(n-1)}\sum_{i=1}^{n}(X_i - \bar{X})^2.\end{aligned} \tag{۵۲.۱}$$

پس باید ثابت کنیم برآورد جک‌نایف واریانس $\bar{X}$ نیز با (۵۲.۱) برابر است.
اگر $S(X) = \bar{X}$، آن‌گاه داریم

$$\begin{aligned}S_{n-1,i} &= \frac{\sum_{j=1}^{n} X_j - X_i}{n-1}\\ &= \frac{n\bar{X} - X_i}{n-1}\\ &= \frac{n}{n-1}\bar{X} - \frac{1}{n-1}X_i\\ &\Rightarrow \frac{1}{n}\sum_{i=1}^{n} S_{n-1,i}\\ &= \frac{1}{n}\sum_{i=1}^{n}(\frac{n}{n-1}\bar{X} - \frac{1}{n-1}X_i)\\ &= \frac{n}{n-1}\bar{X} - \frac{1}{n-1}\bar{X}_i.\end{aligned} \tag{۵۳.۱}$$

پس داریم

$$\begin{aligned}S_{n-1,i} - \frac{1}{n}\sum_{i=1}^{n} S_{n-1,i} &= (\frac{n}{n-1}\bar{X} - \frac{1}{n-1}X_i) - (\frac{n}{n-1}\bar{X} - \frac{1}{n-1}\bar{X})\\ &= \frac{1}{n-1}(\bar{X} - X_i).\end{aligned} \tag{۵۴.۱}$$

با جایگذاری (۵۴.۱) در (۵۰.۱)، برآورد جک‌نایف خطای استاندارد $\bar{X}$ برابر می‌شود با

$$\begin{aligned} se_{jack}(\bar{X}) &= \left[\frac{n-1}{n}\sum_{i=1}^{n}\left(\frac{1}{n-1}(\bar{X}-X_i)\right))^2\right]^{\frac{1}{2}} \\ &= \left[\frac{n-1}{n}\sum_{i=1}^{n}\left(\frac{1}{(n-1)^2}(\bar{X}-X_i)^2\right))\right]^{\frac{1}{2}} \\ &= \left[\frac{1}{n(n-1)}\sum_{i=1}^{n}(X_i-\bar{X})^2\right]^{\frac{1}{2}} \end{aligned} \tag{۵۵.۱}$$

که برابر است با (۵۱.۱) . □

مثال ۳.۶.۱. فرض کنید $X_1, X_2, \cdots, X_n \overset{\text{iid}}{\sim} U(0, \theta)$. می‌دانیم آماره بسنده θ عبارتست از $S_n = X_{(n)} = \max_i X_i$ و همچنین به ازای هر $\theta > 0$، $\theta \left(1+\frac{1}{n}\right) = E[S_n]$.
برآوردگر جک‌نایف در این حالت عبارتست از

$$S_{jack} = \frac{1}{n}\sum_{i=1}^{n}\tilde{S}_{n,i}, \quad \tilde{S}_{n,i} = nS_n - (n-1)S_{n-1,i}$$

که در آن $S_{n-1,i} = \max\{X_1, X_2, \cdots, X_{i-1}, X_{i+1}, \cdots, X_n\}$. در نتیجه

$$\begin{aligned} E(S_{jack}) &= n(1+\frac{1}{n})\theta - (n-1)[1+\frac{1}{n-1}]\theta \\ &= \theta. \end{aligned}$$

دیده می شود که برآوردگر جک‌نایف عبارت از درجه $\frac{1}{n}$ را از اریبی حذف می‌کند. (خاصیت عمومی برآوردگرهای جک‌نایف.)
فرض کنید $\tilde{S}_{n,i}$، $i = 1, 2, \cdots, n$ متغیرهای تصادفی تقریباً مستقل و هم توزیع با میانگین θ باشد.
برآورد جک‌نایف واریانس نمونه عبارتست از

$$\widehat{Var}_{jack} = \frac{1}{n-1}\sum_{i=1}^{n}(\tilde{S}_{n,i} - S_{jack})^2.$$

و بنابراین آماره‌ی

$$\frac{S_{jack}-\theta}{(\frac{\widehat{Var}_{jack}}{n})^{\frac{1}{2}}}$$

تقریباً دارای توزیع t با $n-1$ درجه آزادی است و یک فاصله اطمینان تقریبی $100(1-\alpha)\%$ پارامتر θ عبارتست از

$$\left(S_{jack} - t_{(n-1),\frac{\alpha}{2}}\frac{\widehat{Var}_{jack}}{\sqrt{n}}, S_{jack} + t_{(n-1),\frac{\alpha}{2}}\frac{\widehat{Var}_{jack}}{\sqrt{n}}\right).$$

برای دستیابی به جزئیات بیشتر در این زمینه به غیورمرادی (۱۳۸۷) و هینکلی (۱۹۷۷) مراجعه کنید.

فصل ۲

برآوردگر جک‌نایف اصلاح شده لیو

۱.۲ مقدمه

در این فصل، با استفاده از برخی مهارت‌های جبری که هینکلی (۱۹۷۷)، سینگ و همکاران (۱۹۸۶)، نیکوئیست (۱۹۸۸) و باتا و همکاران (۲۰۰۸) برای به‌دست آوردن برآوردگر جک نایف در مدل‌های رگرسیون ریج پیشنهاد دادند، روش جک‌نایف را برای برآوردگر نوع لیو تعمیم یافته به‌کار گرفته و نشان می‌دهیم برآوردگری را که آکدنیز و آکدنیز دوران (۲۰۱۲) پیشنهاد کردند، $MSEM$ کمتری نسبت به برآوردگرهای لیو تعمیم یافته و جک‌نایف لیو تعمیم یافته دارد.

در بخش اول، مدل‌های لیو تعمیم یافته وجک‌نایف لیو تعمیم یافته توصیف شده‌اند و مقدار اریبی و واریانس آن محاسبه شده و در بخش دوم برآوردگر جک‌نایف اصلاح شده نوع لیو معرفی شده است. در بخش سوم، به منظور مقایسه این برآوردگر با برآوردگر جک‌نایف لیو تعمیم یافته، کارایی برآوردگر را مورد بررسی قرار داده‌ایم و در بخش چهارم با استفاده از معیار ماتریس میانگین توان دوم خطا برآوردگر جک‌نایف اصلاح شده لیو با برآوردگر جک‌نایف لیو تعمیم یافته مقایسه شده است. در بخش پنجم، یک مثال عددی برای تشریح کارایی برآوردگر آورده‌ایم و در نهایت در بخش ششم، با استفاده از شبیه‌سازی مونت‌کارلویی صحت نتایج مورد بررسی قرار گرفته شده است.
نتایج این فصل عمدتاً از مرجع آکدنیز و آکدنیز دوران (۲۰۱۲) می‌باشد. که از متن آن مقاله برزوئی بیدگلی و آرشی (۱۳۹۳) استخراج شده است.

۲.۲ معرفی برآوردگر جک‌نایف لیو تعمیم یافته

برآوردگر نوع لیو تعمیم‌یافته در قسمت ۴.۱ معرفی شد. در این بخش برآوردگر جک‌نایف لیو تعمیم یافته را به طور کامل‌تری شرح داده، اریبی و ماتریس واریانس- کوواریانس آن‌را محاسبه می‌کنیم. در نهایت برای نشان دادن برتری برآوردگر جک‌نایف لیو تعمیم یافته نسبت به برآوردگر نوع لیو تعمیم یافته مقادیر اریبی و معیار $MSEM$ (طبق تعریف ب.۰.۰.۳۰) هر دو برآوردگر را با هم مقایسه می‌کنیم.

صورت کانونی مدل رگرسیون خطی در رابطه (۱۸.۱) در نظر بگیرید. با توجه به رابطه (۳۵.۱) و با استفاده از قضیه آ.۱۲.۰، متناظر با هر نمونه جک‌نایف، پاسخ جک‌نایف $\hat{\gamma}(\mathbf{D})$ به صورت زیر تعریف می‌شود

$$\begin{aligned}
\hat{\gamma}(\mathbf{D})_{(i)} &= (\mathbf{A} - z_i z_i')^{-1}(\mathbf{Z}'y - z_i y_i) \\
&= \left(\mathbf{A}^{-1} + \frac{\mathbf{A}^{-1} z_i z_i' \mathbf{A}^{-1}}{1 - z_i' \mathbf{A}^{-1} z_i}\right)(\mathbf{Z}'y - z_i y_i) \\
&= \mathbf{A}^{-1}\mathbf{Z}'y - \mathbf{A}^{-1} z_i y_i + \frac{\mathbf{A}^{-1} z_i z_i' \mathbf{A}^{-1}}{1 - z_i' \mathbf{A}^{-1} z_i}\mathbf{Z}'y - \frac{\mathbf{A}^{-1} z_i z_i' \mathbf{A}^{-1}}{1 - z_i' \mathbf{A}^{-1} z_i} z_i z_i' \\
&= \hat{\gamma}(\mathbf{D}) - \mathbf{A}^{-1} z_i y_i \left(1 + \frac{z_i' \mathbf{A}^{-1} z_i}{1 - z_i' A^{-1} z_i}\right) + \frac{A^{-1} z_i z_i'}{1 - z_i' \mathbf{A}^{-1} z_i}\hat{\gamma}(\mathbf{D}) \\
&= \hat{\gamma}(\mathbf{D}) - \frac{(\mathbf{A}^{-1} z_i y_i)(1 - z_i' \mathbf{A}^{-1} z_i) + (z_i' \mathbf{A}^{-1} z_i)(A^{-1} z_i y_i) + A^{-1} z_i z_i' \hat{\gamma}(\mathbf{D})}{1 - z_i' \mathbf{A}^{-1} z_i} \\
&= \hat{\gamma}(\mathbf{D}) - \frac{\mathbf{A}^{-1} z_i y_i + \mathbf{A}^{-1} z_i z_i' \hat{\gamma}(D)}{1 - z_i' \mathbf{A}^{-1} z_i} \\
&= \hat{\gamma}(\mathbf{D}) - \frac{\mathbf{A}^{-1} z_i (y_i - z_i \hat{\gamma}(\mathbf{D}))}{1 - z_i' \mathbf{A}^{-1} z_i} \\
&= \hat{\gamma}(\mathbf{D}) - \frac{\mathbf{A}^{-1} z_i e_i}{1 - w_i} \qquad (1.2)
\end{aligned}$$

که z_i'، i–امین سطر از ماتریس $\mathbf{Z}$، $e_i = y_i - z_i' \hat{\gamma}(\mathbf{D})$، i–امین باقیمانده لیو، $w_i = z_i' A^{-1} z_i$، i–امین فاکتور فاصله و

$$\begin{aligned}
\mathbf{A}^{-1} &= (\mathbf{\Lambda} + \mathbf{I})^{-1}(\mathbf{I} + \mathbf{D}\mathbf{\Lambda}^{-1}) \\
&= (\mathbf{\Lambda} + \mathbf{I})^{-1}(\mathbf{\Lambda} + \mathbf{D})\mathbf{\Lambda}^{-1} \\
&= \mathbf{F}_{\mathbf{D}}\mathbf{\Lambda}^{-1} \qquad (2.2)
\end{aligned}$$

است. با توجه به این نکته که مقدار غیر صفر w_i، مدلِ با وزن‌های نابرابر را بازتاب می‌کند، از روش شبه مقادیر جک‌نایف وزنی که توسط هینکلی (۱۹۷۷) و نیکوئیست (۱۹۸۸) پیشنهاد شدند، استفاده می‌کنیم. در این خصوص شبه مقادیر وزنی جک‌نایف به صورت زیر تعریف می‌شود

$$Q_i = \hat{\gamma}(\mathbf{D}) + n(1 - w_i)(\hat{\gamma}(\mathbf{D}) - \hat{\gamma}(\mathbf{D})_{(i)}) \qquad (3.2)$$

برآوردگر جک‌نایف لیو تعمیم یافته[1](JGLE)، با اریبی کاهش یافته به صورت زیر به‌دست می‌آید

$$\begin{aligned}\hat{\gamma}_{JGLE}(\mathbf{D}) = \bar{Q} &= \frac{1}{n}\sum_{i=1}^{n} Q_i \\ &= \hat{\gamma}(\mathbf{D}) + \frac{1}{n}\sum_{i=1}^{n} n(1-w_i)(\hat{\gamma}(\mathbf{D}) - \hat{\gamma}(\mathbf{D})_{-i}) \\ &= \hat{\gamma}(\mathbf{D}) + \sum_{i=1}^{n}(1-w_i)\left(\hat{\gamma}(\mathbf{D}) - \hat{\gamma}(\mathbf{D}) + \frac{\mathbf{A}^{-1} z_i e_i}{1-w_i}\right) \\ &= \hat{\gamma}(\mathbf{D}) + \sum_{i=1}^{n} \mathbf{A}^{-1} z_i e_i \\ &= \hat{\gamma}(\mathbf{D}) + \mathbf{A}^{-1}\sum_{i=1}^{n} z_i e_i \end{aligned} \tag{۴.۲}$$

و همچنین

$$\begin{aligned}\sum_{i=1}^{n} z_i e_i &= \sum_{i=1}^{n}(y_i - z_i'\hat{\gamma}(\mathbf{D})) \\ &= \sum_{i=1}^{n} z_i y_i - \sum_{i=1}^{n} z_i z_i'(\mathbf{\Lambda}+\mathbf{I})^{-1}(z'y + \mathbf{D}\hat{\gamma}(\mathbf{D})) \\ &= \sum_{i=1}^{n} z_i y_i - \sum_{i=1}^{n} z_i z_i'(\mathbf{\Lambda}+\mathbf{I})^{-1}(\mathbf{I}+\mathbf{D}\mathbf{\Lambda}^{-1})\mathbf{Z}'y \\ &= \mathbf{Z}'y - \mathbf{Z}'\mathbf{Z}(\mathbf{\Lambda}+\mathbf{I})^{-1}(\mathbf{I}+\mathbf{D}\mathbf{\Lambda}^{-1})\mathbf{Z}'y \\ &= (\mathbf{I} - \mathbf{\Lambda}\mathbf{A}^{-1})\mathbf{Z}'y. \end{aligned} \tag{۵.۲}$$

با جایگذاری (۵.۲) در (۴.۲) داریم

$$\begin{aligned}\hat{\gamma}_{JGLE}(\mathbf{D}) &= \hat{\gamma}(\mathbf{D}) + \mathbf{A}^{-1}(\mathbf{I} - \mathbf{\Lambda}\mathbf{A}^{-1})\mathbf{Z}'y \\ &= \hat{\gamma}(\mathbf{D}) + \mathbf{A}^{-1}\mathbf{Z}'y - \mathbf{A}^{-1}\mathbf{\Lambda}\mathbf{A}^{-1}\mathbf{Z}'y \\ &= \hat{\gamma}(\mathbf{D}) + \hat{\gamma}(\mathbf{D}) - \mathbf{A}^{-1}\mathbf{\Lambda}\hat{\gamma}(\mathbf{D}) \\ &= 2\hat{\gamma}(\mathbf{D}) - \mathbf{A}^{-1}\mathbf{\Lambda}\hat{\gamma}(\mathbf{D}) \\ &= [2\mathbf{I} - \mathbf{A}^{-1}\mathbf{\Lambda}]\hat{\gamma}(\mathbf{D}) \end{aligned} \tag{۶.۲}$$

[1] Jackknifed generalized liu-type estimator

و با توجه به عبارت (۲۹.۱) می‌توان نتیجه گرفت

$$\begin{aligned}\hat{\gamma}_{JGLE}(\mathbf{D}) &= (2\mathbf{I}-\mathbf{A}^{-1}\mathbf{\Lambda})\hat{\gamma}(\mathbf{D})\\ &= \left[\mathbf{I}+(\mathbf{I}-\mathbf{A}^{-1}\mathbf{\Lambda})\right]\hat{\gamma}(\mathbf{D})\\ &= \left[\mathbf{I}+(\mathbf{\Lambda}+\mathbf{I})^{-1}(\mathbf{I}-\mathbf{D})\right]\hat{\gamma}(\mathbf{D})\\ &= (2\mathbf{I}-\mathbf{F_D})\hat{\gamma}(\mathbf{D}). \end{aligned} \tag{۷.۲}$$

با استفاده از رابطه (۲۹.۱) نتیجه می‌شود

$$\hat{\gamma}_{JGLE}(\mathbf{D}) = (2\mathbf{I}-\mathbf{F_D})\mathbf{F_D}\hat{\gamma}_{OLSE}. \tag{۸.۲}$$

i–امین مؤلفه برآوردگر جک‌نایف لیو تعمیم یافته به صورت زیر است

$$\begin{aligned}\hat{\gamma}_{JGLE}(\mathbf{D})_i &= \frac{\lambda_i^2+d_i^2+2(\lambda_i+d_i)}{(1+\lambda_i)^2}\hat{\gamma}_i\\ &= [g(d_i)]\hat{\gamma}_i \end{aligned} \tag{۹.۲}$$

که در آن $1 \leq g(d_i) \leq 0$ و $\hat{\gamma}_i$ ها، i –امین مؤلفه از بردار $\hat{\gamma}$ هستند.

معمولاً شبه مقادیر جک‌نایف برآوردگری را فراهم می‌کنند که دارای اریبی کمتری است. کاهش اریبی در مقایسه مقادیر $|\text{bias}(\hat{\gamma}(\mathbf{D})_i)|$ با $|\text{bias}(\hat{\gamma}_{JGLE}(\mathbf{D})_i)|$ مشخص می‌شود؛ که در آن $|\cdots|_i$ نشان‌دهنده قدر مطلق i–امین مؤلفه است.

با توجه به روابط (۲۹.۱) و (۸.۲)، برآوردگر نوع لیو تعمیم یافته و برآوردگر جک‌نایف لیو تعمیم یافته برآوردگرهایی اریب هستند. بنابراین با توجه به رابطه (۸.۲)، اریبی برآوردگر جک‌نایف لیو تعمیم یافته عبارتست از

$$\begin{aligned}\text{bias}(\hat{\gamma}_{JGLE}(\mathbf{D})) &= E(\hat{\gamma}_{JGLE}(\mathbf{D}))-\gamma\\ &= (2\mathbf{I}-\mathbf{F_D})\mathbf{F_D}\gamma-\gamma\\ &= [(2\mathbf{I}-\mathbf{F_D})\mathbf{F_D}-\mathbf{I}]\gamma\\ &= ([\mathbf{I}-(\mathbf{\Lambda}+\mathbf{I})^{-2}(\mathbf{I}-\mathbf{D})^2]-\mathbf{I})\gamma\\ &= -(\mathbf{\Lambda}+\mathbf{I})^{-2}(\mathbf{I}-\mathbf{D})^2\gamma\\ &= (\mathbf{I}-\mathbf{F_D})^2\gamma. \end{aligned} \tag{۱۰.۲}$$

تفاضل i-امین مؤلفه اریبی‌های برآوردگرجک‌نایف لیو تعمیم یافته و برآوردگر نوع لیو برابر است با

$$\begin{aligned}
&|\mathrm{bias}(\hat{\gamma}(\mathbf{D})_i)| - |\mathrm{bias}(\hat{\gamma}(\mathbf{D})_{JGLE}(\mathbf{D})_i)| \\
&= |\frac{(d_i - ۱)}{۱ + \lambda_i}|.|\gamma_i| - \frac{(۱ - d_i)^۲}{(۱ + \lambda_i)^۲}|\gamma_i| \\
&= |\frac{-(۱ - d_i)}{۱ + \lambda_i}|.|\gamma_i| - \frac{(۱ - d_i)^۲}{(۱ + \lambda_i)^۲}|\gamma_i| \\
&= \frac{(۱ - d_i)(۱ + \lambda_i + ۱ + d_i)}{(۱ + \lambda_i)^۲}|\gamma_i| \\
&= \frac{(۱ - d_i)(\lambda_i + d_i)}{(۱ + \lambda_i)^۲}|\gamma_i| \qquad (۱۱.۲)
\end{aligned}$$

که مقداری مثبت است و نشان می‌دهد برآوردگر نوع لیو دارای قدر مطلق اریبی بیشتری نسبت به برآوردگرجک‌نایف لیو تعمیم یافته است.
هرگاه $۰ < d < ۱$، $d_۱ = d_۲ = d_۳ = \ldots = d_p = d$، برآوردگر لیو جک‌نایف را با توجه به روابط (۲۹.۱) و (۷.۲) می‌توان به صورت زیر نوشت

$$\begin{aligned}
\hat{\gamma}_{JLE}(d) &= (۲\mathbf{I} - \mathbf{F}_d)\mathbf{F}_d \\
&= ۲\mathbf{F}_d - \mathbf{F}_d^۲ \\
&= (\mathbf{I} - (۱ - d)^۲(\mathbf{\Lambda} + \mathbf{I})^{-۲})\hat{\gamma}_{OLSE}. \qquad (۱۲.۲)
\end{aligned}$$

اگر در رابطه (۱۲.۲)، $d = ۱$، آنگاه برآوردکننده‌ی کمترین توان‌های دوم به‌دست می‌آید.
با استفاده از (۸.۲) و (۲۰.۱) کوواریانس برآوردگر جک‌نایف لیو تعمیم یافته به صورت زیر است

$$\begin{aligned}
\mathrm{cov}(\hat{\gamma}_{JGLE}(\mathbf{D})) &= \mathrm{cov}((۲\mathbf{I} - \mathbf{F_D})\mathbf{F_D}\hat{\gamma}_{OLSE}) \\
&= (۲\mathbf{I} - \mathbf{F_D})\mathbf{F_D}\mathrm{cov}(\hat{\gamma}_{OLSE})\mathbf{F}'_\mathbf{D}(۲\mathbf{I} - \mathbf{F_D})' \\
&= \sigma^۲(۲\mathbf{I} - \mathbf{F_D})\mathbf{F_D}\mathbf{\Lambda}^{-۱}\mathbf{F}'_\mathbf{D}(۲\mathbf{I} - \mathbf{F_D})'. \qquad (۱۳.۲)
\end{aligned}$$

مقدار $MSEM$ برای برآوردگرجک‌نایف لیو تعمیم یافته طبق روابط (۱۳.۲) و (۱۰.۲) به صورت زیر محاسبه می‌شود

$$\begin{aligned}
MSEM(\hat{\gamma}_{JGLE}(\mathbf{D})) &= \mathrm{cov}(\hat{\gamma}_{JGLE}(\mathbf{D})) + [\mathrm{bias}(\hat{\gamma}_{JGLE}(\mathbf{D}))][\mathrm{bias}(\hat{\gamma}_{JGLE}(\mathbf{D}))]' \\
&= \sigma^۲(۲\mathbf{I} - \mathbf{F_D})\mathbf{F_D}\mathbf{\Lambda}^{-۱}\mathbf{F}'_\mathbf{D}(۲\mathbf{I} - \mathbf{F_D})' + (\mathbf{I} - \mathbf{F_D})^۲\gamma\gamma'(\mathbf{I} - \mathbf{F_D})'^۲. \qquad (۱۴.۲)
\end{aligned}$$

۱.۲.۲ مثال عددی

مثال ۲.۴.۱ مربوط به درصد تبدیل n- هپتان به استیلن را در نظر بگیرید. در این‌جا از برآوردگر جک‌نایف لیو تعمیم یافته برای رفع مشکل همخطی استفاده می‌کنیم. مقدار این برآوردگر، طول توان دوم

جدول ۱.۲: برآورد $\hat{\beta}(d)$ برای مقادیر مختلف d

	$d=۰/۵۵$	$d=۰/۷۵$	$d=۱$
$\beta_۰$	-۹۲/۶۰۹	-۸۱۰۶/۷۸۱	-۱۱۳/۱۱
ضریب درجه رآکتور	۰/۱۰۶	۰/۱۱۸	۰/۱۲۳
ضریب نسبت $H_۲$ به n- هپتان	-۰/۰۱۶	-۰/۰۲۸	-۰/۰۳۳
ضریب زمان اتصال	-۳/۴۱۸	-۲/۰۴۱	-۱/۴۲۴
طول توان‌دوم ضرایب	۸۵۸۸/۱۴۸	۱۱۴۰۶/۵۱	۱۰۵۶/۱۵۷
اثر میانگین توان‌دوم خطا	۱۱۲۰/۶۲۶	۹۵۸/۹۲۰	۱۰۲۷/۸۹۲

ضرایب رگرسیون و مقادیر میانگین توان دوم خطا به ازای مقادیر مختلف d در جدول ۱.۲ آمده است. با توجه به جدول ۱.۲ و شکل ۱.۲ با افزایش مقدار d توان دوم ضرایب رگرسیون بیشتر می‌شود و مقدار MSE کمتر می‌شود. و هرچه مقدار d افزایش می‌یابد، مقدار توان دوم اریبی کمتر و برآورد واریانس برآوردگرها بیشتر می‌شود. برنامه مربوط به این مثال در پیوست ت.۲ آورده شده است.

۳.۲ برآوردگر جک‌نایف اصلاح شده لیو تعمیم یافته

برآوردگر جک‌نایف اصلاح شده لیو تعمیم یافته[۲](MJGLE)، توسط آکدنیز دوران و آکدنیز[۳] در سال ۲۰۱۲ پیشنهاد شده است. این برآوردگر به عنوان برآوردگر جک‌نایف اصلاح شده لیو تعمیم یافته طراحی شده است که عبارتست از

$$\begin{aligned}\hat{\gamma}_{MJGLE}(\mathbf{D}) &= [\mathbf{I}-(\mathbf{\Lambda}+\mathbf{I})^{-۲}(\mathbf{I}-\mathbf{D})^{۲}]\hat{\gamma}(\mathbf{D})\\ &= [\mathbf{I}-(\mathbf{\Lambda}+\mathbf{I})^{-۲}(\mathbf{I}-\mathbf{D})^{۲}]\mathbf{F_D}\hat{\gamma}_{OLSE}\\ &= [\mathbf{I}-(\mathbf{\Lambda}+\mathbf{I})^{-۲}(\mathbf{I}-\mathbf{D})^{۲}][\mathbf{I}-(\mathbf{\Lambda}+\mathbf{I})^{-۱}(\mathbf{I}-\mathbf{D})]\hat{\gamma}_{OLSE} \qquad (۱۵.۲)\end{aligned}$$

هرگاه $d_i=۱, i=۱,۲,\dots,p$، چون $\mathbf{I}-\mathbf{D}=\mathbf{0}$؛ آنگاه، با توجه به رابطه (۱۵.۲) i-امین مؤلفه این برآوردگر به صورت زیر خواهد بود

$$\hat{\gamma}_{MJGLE}(\mathbf{D})_i=\hat{\gamma}_i.$$

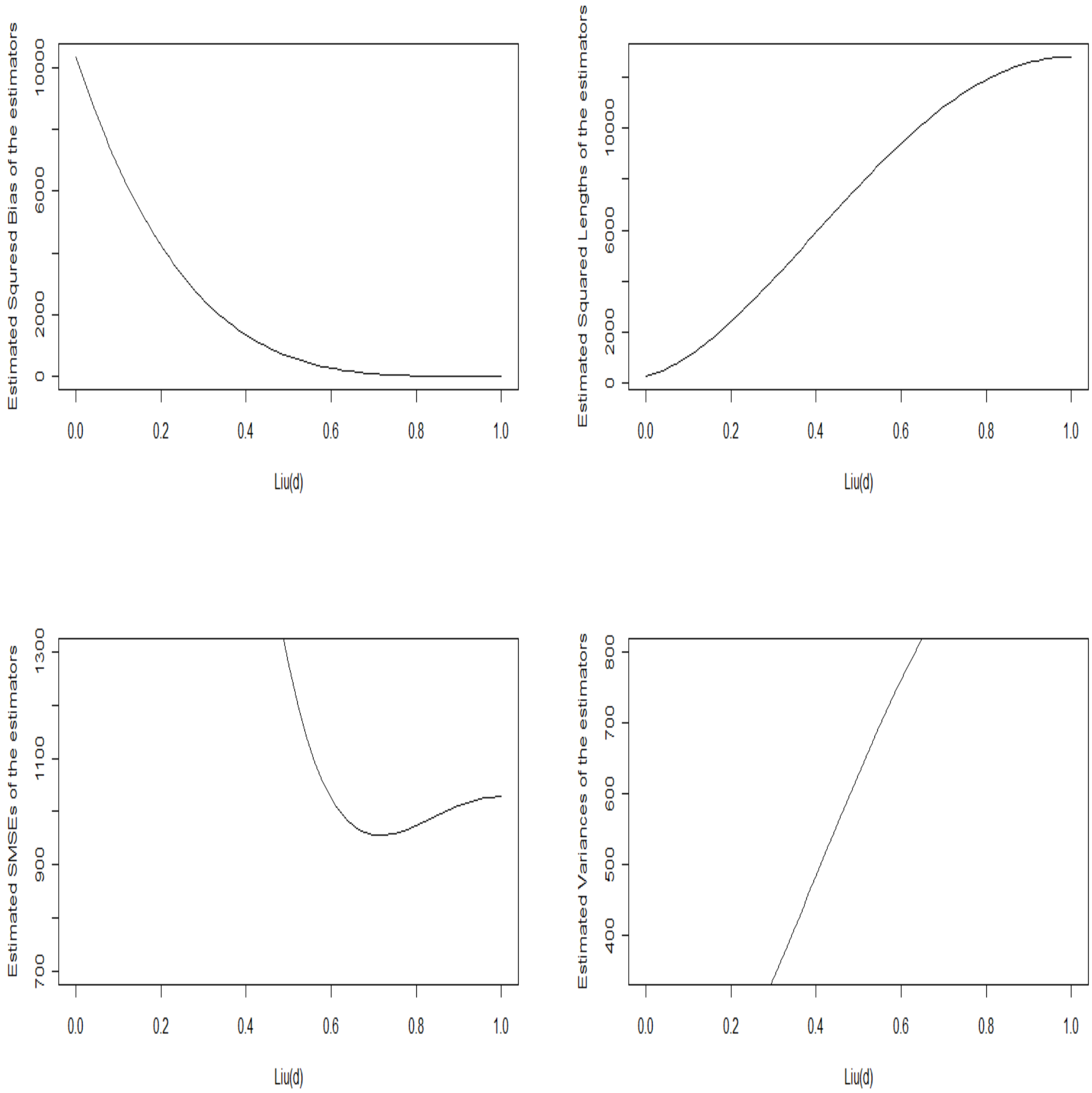
هرگاه $d_۱=d_۲=\cdots=d_p=d$ و $۰<d<۱$، آنگاه $\hat{\gamma}_{MJGLE}$ در رابطه (۱۵.۲) به صورت زیر نوشته می‌شود که آن را برآوردگر جک‌نایف اصلاح شده لیو[۴](MJLE)، می‌نامند

$$\hat{\gamma}_{MJLE}(d)=\Big[\mathbf{I}-(۱-d)^{۲}(\mathbf{\Lambda}+\mathbf{I})^{-۲}][\mathbf{I}-(۱-d)(\mathbf{\Lambda}+\mathbf{I})^{-۱}\Big]\hat{\gamma}_{OLSE}. \qquad (۱۶.۲)$$

[۲] Modified jackknifed generalized liu-type estimator

[۳] Akdeniz Duran and Akdeniz

[۴] Modified jackknifed liu-type estimator



شکل ۱.۲: عملکرد برآوردگر لیو به ازای مقادیر مختلف d ، $0 < d < 1$

با استفاده از روابط (۹.۲) و (۱۶.۲) برای i–امین مؤلفه داریم

$$\begin{aligned}
\hat{\gamma}_{MJGLE}(\mathbf{D})_i &= \left[1-\frac{(1-d_i)^2}{(1+\lambda_i)^2}\right]\left[1-\frac{1-d_i}{1+\lambda_i}\right]\hat{\gamma}_{OLSE} \\
&= \left(\frac{(1+\lambda_i)^2-(1-d_i)^2}{(1+\lambda_i)^2}\right)\left(\frac{1+\lambda_i-1+d_i}{1+\lambda_i}\right) \\
&= \frac{(\lambda_i+d_i)(1+\lambda_i)^2}{(1+\lambda_i)^3}-\frac{(\lambda_i+d_i)(1-d_i)^2}{(1+\lambda_i)^3} \\
&= \frac{(\lambda_i+d_i)}{(1+\lambda_i)}\left[1-\frac{(1-d_i)^2}{(1+\lambda_i)^2}\right] \\
&= [g(d_i)]\hat{\gamma}_i \qquad (17.2)
\end{aligned}$$

که در آن $g(d_i) \leq 1$ است.
اریبی این برآوردگر به صورت زیر قابل محاسبه است

$$\begin{aligned}
\text{bias}(\hat{\gamma}_{MJGLE}(\mathbf{D})) &= E(\hat{\gamma}_{MJGLE}(\mathbf{D}))-\gamma \\
&= [\mathbf{I}-(\mathbf{\Lambda}+\mathbf{I})^{-2}(\mathbf{I}-\mathbf{D})^2]F_{\mathbf{D}}\gamma-\gamma \\
&= [(\mathbf{I}-(\mathbf{\Lambda}+\mathbf{I})^{-2}(\mathbf{I}-\mathbf{D})^2)F_{\mathbf{D}}-\mathbf{I}]\gamma \\
&= -(\mathbf{\Lambda}+\mathbf{I})^{-1}[\mathbf{I}+(\mathbf{\Lambda}+\mathbf{I})^{-1}(\mathbf{I}-\mathbf{D})-(\mathbf{\Lambda}+\mathbf{I})^{-2}(\mathbf{I}-\mathbf{D})^2](\mathbf{I}-\mathbf{D})\gamma \\
&= -(\mathbf{\Lambda}+\mathbf{I})^{-1}[\mathbf{I}+\mathbf{F_D}-\mathbf{F_D^2}](\mathbf{I}-\mathbf{D})\gamma \\
&= -(\mathbf{\Lambda}+\mathbf{I})^{-1}\mathbf{W}(\mathbf{I}-\mathbf{D})\gamma \qquad (18.2)
\end{aligned}$$

که در آن $\mathbf{W}=\mathbf{I}+\mathbf{F_D}-\mathbf{F_D^2}$.
ماتریس واریانس- کوواریانس برآوردگر جک‌نایف اصلاح شده لیو تعمیم یافته نیز برابر است با

$$\begin{aligned}
\text{cov}(\hat{\gamma}_{MJGLE}(\mathbf{D})) =& [\mathbf{I}-(\mathbf{\Lambda}+\mathbf{I})^{-2}(\mathbf{I}-\mathbf{D})^2][\mathbf{I}-(\mathbf{\Lambda}+\mathbf{I})^{-1}(\mathbf{I}-\mathbf{D})]\text{cov}(\hat{\gamma}_{OLSE}) \\
&([\mathbf{I}-(\mathbf{\Lambda}+\mathbf{I})^{-2}(\mathbf{I}-\mathbf{D})^2][\mathbf{I}-(\mathbf{\Lambda}+\mathbf{I})^{-2}(\mathbf{I}-\mathbf{D})^2])' \\
&= \sigma^2\mathbf{\Phi}\mathbf{\Lambda}^{-1}\mathbf{\Phi}'
\end{aligned}$$

(۱۹.۲)

که در آن

$$\begin{aligned}
\mathbf{\Phi} &= \left[\mathbf{I}-(\mathbf{\Lambda}+\mathbf{I})^{-2}(\mathbf{I}-\mathbf{D})^2\right]\left[\mathbf{I}-(\mathbf{\Lambda}+\mathbf{I})^{-1}(\mathbf{I}-\mathbf{D})\right] \\
&= (2\mathbf{I}-\mathbf{F_D})\mathbf{F_D^2}.
\end{aligned}$$

و در نهایت مقدار $MSEM$ برابر ست با

$$\begin{aligned}
MSEM(\hat{\gamma}_{MJGLE}(\mathbf{D})) &= \text{cov}(\hat{\gamma}_{MJGLE}(\mathbf{D})) + [\text{bias}(\hat{\gamma}_{MJGLE}(\mathbf{D}))] \\
&= \sigma^{۲}\mathbf{\Phi}\mathbf{\Lambda}^{-۱}\mathbf{\Phi}' \\
&+ ((\mathbf{\Lambda}+\mathbf{I})^{-۱}\mathbf{W}(\mathbf{I}-\mathbf{D})\gamma)((\mathbf{\Lambda}+\mathbf{I})^{-۱}\mathbf{W}(\mathbf{I}-\mathbf{D})\gamma)' \\
&= \sigma^{۲}\mathbf{\Phi}\mathbf{\Lambda}^{-۱}\mathbf{\Phi}' \\
&+ (\mathbf{\Lambda}+\mathbf{I})^{-۱}\mathbf{W}(\mathbf{I}-\mathbf{D})\gamma\gamma'(\mathbf{I}-\mathbf{D})\mathbf{W}'(\mathbf{\Lambda}+\mathbf{I})^{-۱}. \quad (۲۰.۲)
\end{aligned}$$

۴.۲ کارایی برآوردگر جک‌نایف اصلاح شده لیو تعمیم یافته

در بخش‌های قبلی دیدیم که برآوردگر جک‌نایف اصلاح شده لیو تعمیم یافته اریب بوده و معیاری مناسب برای اندازه‌گیری کارایی این برآوردگر، $MSEM$ است. ابتدا برای درک این مطلب و سهولت در محاسبات بعدی، چند قضیه رابیان می‌کنیم و در ادامه کارایی برآوردگر جک‌نایف اصلاح شده لیو تعمیم یافته را با برآوردگر لیو تعمیم یافته و برآوردگر جک‌نایف لیو تعمیم یافته، با استفاده از معیار $MSEM$ مقایسه می‌کنیم.

قضیه ۱.۴.۲. * *(آکدنیز و آکدنیز دوران، ۲۰۱۲) فرض کنید $\boldsymbol{D}$ یک ماتریس قطری معین مثبت $p\times p$ باشد. برآوردگر جک‌نایف اصلاح شده لیو تعمیم یافته واریانس کمتری نسبت به برآوردگر نوع لیو تعمیم یافته دارد.*

برهان. با توجه به روابط (۳۲.۱) و (۱۹.۲) داریم

$$\text{cov}(\hat{\gamma}(\mathbf{D})) - \text{cov}(\hat{\gamma}_{MJGLE}(\mathbf{D})) = \sigma^{۲}\mathbf{H} \quad (۲۱.۲)$$

که در آن

$$\begin{aligned}
\mathbf{H} &= [\mathbf{I}-(\mathbf{\Lambda}+\mathbf{I})^{-۱}(\mathbf{I}-\mathbf{D})]\mathbf{\Lambda}^{-۱}[\mathbf{I}-(\mathbf{\Lambda}+\mathbf{I})^{-۱}(\mathbf{I}-\mathbf{D})]' - \sigma^{۲}\mathbf{\Phi}\mathbf{\Lambda}^{-۱}\mathbf{\Phi}' \\
&= \sigma^{۲}(\mathbf{I}-\mathbf{F_D})[\mathbf{\Lambda}^{-۱} - (\mathbf{I}-\mathbf{F}_{\mathbf{D}}^{۲})\mathbf{\Lambda}^{-۱}(\mathbf{I}-\mathbf{F}_{\mathbf{D}}^{۲})'](\mathbf{I}-\mathbf{F_D})'.
\end{aligned}$$

بنابراین ماتریس $\mathbf{H}$ یک ماتریس قطری است. i–امین مؤلفه این ماتریس مثبت و به صورت زیر است

$$h_{ii} = \frac{\lambda_i(\lambda_i+d_i)(۱-d_i)^{۲}}{(۱+\lambda_i)^{۵}}[(۱+\lambda_i)^{۲} + (\lambda_i+d_i)(۲+\lambda_i-d_i)].$$

بنابراین طبق تعریف آ.۱۷.۰ ماتریس $\mathbf{H}$ یک ماتریس معین مثبت بوده و برهان کامل است. □

قضیه ۲.۴.۲. * *(آکدنیز و آکدنیز دوران، ۲۰۱۲) فرض کنید $\boldsymbol{D}$ یک ماتریس قطری معین مثبت $p\times p$ باشد، آن‌گاه اختلاف*

$$\mathbf{\Delta}_{۱} = MSEM(\hat{\gamma}_{OLSE}, \hat{\gamma}(\boldsymbol{D})) - MSEM(\hat{\gamma}_{OLSE}, \hat{\gamma}_{MJGLE}(\boldsymbol{D})) \quad (۲۲.۲)$$

یک ماتریس معین غیر منفی است، اگر و تنها اگر در نامساوی زیر صدق کند

$$\gamma'[\mathbf{L}^{-1}(\sigma^2\mathbf{H}+\mathbf{F}_{\boldsymbol{D}}\gamma\gamma'\mathbf{F}'_{\boldsymbol{D}})\mathbf{L}'^{-1}]^{-1}\gamma \leqslant 1 \qquad (۲۳.۲)$$

که در آن $\mathbf{L}=\mathbf{F}_{\boldsymbol{D}}\mathbf{W}$ *و* $\mathbf{W}=\mathbf{I}+\mathbf{F}_{\boldsymbol{D}}-\mathbf{F}^2_{\boldsymbol{D}}$.

برهان. با استفاده از رابطه‌های (۳۳.۱) و (۲۰.۲) داریم

$$\begin{aligned}
\boldsymbol{\Delta}_1 &= MSEM(\hat{\gamma},\hat{\gamma}(\mathbf{D}))-MSEM(\hat{\gamma},\hat{\gamma}_{MJGLE}(\mathbf{D}))\\
&=\sigma^2\mathbf{F_D}\boldsymbol{\Lambda}^{-1}\mathbf{F}'_{\mathbf{D}}+(\mathbf{F_D}-\mathbf{I})\gamma\gamma'(\mathbf{F_D}-\mathbf{I})'-\sigma^2\boldsymbol{\Phi}\Lambda^{-1}\boldsymbol{\Phi}'\\
&+(\boldsymbol{\Lambda}+\mathbf{I})^{-1}\mathbf{W}(\mathbf{I}-\mathbf{D})\gamma\gamma'(\mathbf{I}-\mathbf{D})\mathbf{W}'(\boldsymbol{\Lambda}+\mathbf{I})^{-1}\\
&=\sigma^2\mathbf{F_D}\boldsymbol{\Lambda}^{-1}\mathbf{F}'_{\mathbf{D}}-\sigma^2\boldsymbol{\Phi}\boldsymbol{\Lambda}^{-1}\boldsymbol{\Phi}'+\mathbf{F_D}\gamma\gamma'\mathbf{F}'_{\mathbf{D}}-\mathbf{F_D}\mathbf{W}\gamma\gamma'\mathbf{W}'\mathbf{F}'_{\mathbf{D}}\\
&=\sigma^2\mathbf{H}+\mathbf{F_D}\gamma\gamma'\mathbf{F}'_{\mathbf{D}}-\mathbf{F_D}\mathbf{W}\gamma\gamma'\mathbf{W}'\mathbf{F}'_{\mathbf{D}} \qquad (۲۴.۲)
\end{aligned}$$

که در آن $\mathbf{W}=\mathbf{I}+\mathbf{F_D}-\mathbf{F}^2_{\mathbf{D}}$ یک ماتریس قطری با عناصر مثبت است که طبق تعریف آ.۱۷.۰، معین مثبت می‌باشد. در قضیه ۱.۴.۲ دیدیم که ماتریس $\mathbf{H}$ نیز معین مثبت می‌باشد. بنابراین اختلاف $\boldsymbol{\Delta}_1$ معین غیر منفی است اگر و تنها اگر $\mathbf{L}^{-1}\boldsymbol{\Delta}_1\mathbf{L}'^{-1}$ معین نامنفی باشد. با توجه به این‌که

$$\mathbf{L}^{-1}\boldsymbol{\Delta}_1\mathbf{L}'^{-1}=\mathbf{L}^{-1}\left(\sigma^2\mathbf{H}+\mathbf{F_D}\gamma\gamma'\mathbf{F}'_{\mathbf{D}}\right)\mathbf{L}'^{-1}-\gamma\gamma' \qquad (۲۵.۲)$$

و این مطلب که ماتریس $\sigma^2\mathbf{H}+\mathbf{F_D}\gamma\gamma'\mathbf{F}'_{\mathbf{D}}$ متقارن و معین مثبت است، با استفاده از لم ب.۳۲.۰، ماتریس $\mathbf{L}^{-1}\boldsymbol{\Delta}_1\mathbf{L}'^{-1}$ معین نامنفی است اگر و تنها اگر

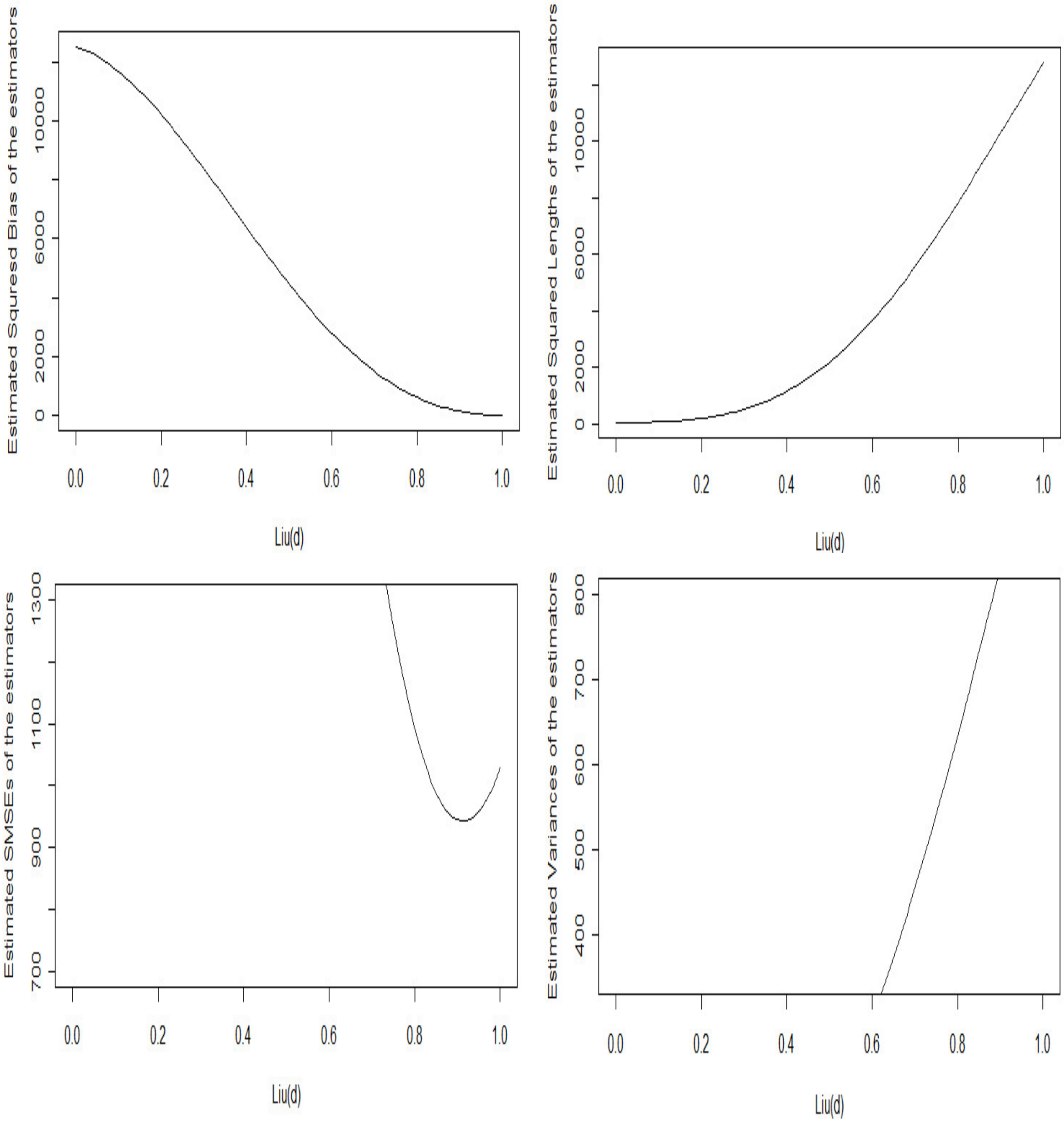
$$\gamma'\left[\mathbf{L}^{-1}(\sigma^2\mathbf{H}+\mathbf{F_D}\gamma\gamma'\mathbf{F}'_{\mathbf{D}})\mathbf{L}'^{-1}\right]^{-1}\gamma\leq 1 \qquad (۲۶.۲)$$

و برهان کامل است. □

۱.۴.۲ مثال عددی

مثال ۲.۴.۱ مربوط به در صد تبدیل n- هپتان به استیلن را در نظر بگیرید. در این‌جا از برآوردگر جک‌نایف اصلاح شده لیو تعمیم یافته برای رفع مشکل همخطی استفاده می‌کنیم. مقدار این برآوردگر، طول توان دوم ضرایب رگرسیون و مقادیر میانگین توان دوم خطا به ازای مقادیر مختلف d در جدول ۲.۲ آمده است.

با توجه به جدول ۲.۲ و شکل ۲.۲ با افزایش مقدار d توان دوم ضرایب رگرسیون بیش‌تر می‌شود و مقدارMSE کم‌تر می‌شود. و هرچه مقدار d افزایش می‌یابد، مقدار توان دوم اریبی کم‌تر و برآورد واریانس برآوردگرها بیش‌تر می‌شود. برنامه مربوط به این مثال در پیوست ت.۳ آورده شده است.



شکل ۲.۲: عملکرد برآوردگر لیو به ازای مقادیر مختلف d ، $0 < d < 1$

جدول ۲.۲: برآورد $\hat{\beta}(d)$ برای مقادیر مختلف d

	$d = ۰/۵۵$	$d = ۰/۷۵$	$d = ۱$
$\beta_۰$	−۵۳/۳۳	−۸۱/۵۹	−۱۱۳/۱۱
ضریب درجه راکتور	۰/۰۷۴	۰/۰۹۷	۰/۱۲۳
ضریب نسبت $H_۲$ به n− هپتان	۰/۰۱۹	−۰/۰۱۱	−۰/۰۳۳
ضریب زمان اتصال	−۵/۰۴۶	−۳/۳۵۴	−۱/۴۲۴
طول توان‌دوم ضرایب	۲۸۶۹/۵۹۳	۶۶۶۸/۵۳۳	۱۲۷۹۵/۶۴
اثر میانگین توان دوم خطا	۲۵۵۴/۸۲۹	۱۲۵۶/۴۸۱	۱۰۲۷/۸۹۲

۵.۲ مقایسه برآوردگر جک‌نایف اصلاح شده لیو تعمیم یافته و برآوردگر جک‌نایف لیو تعمیم یافته

در این بخش نشان می‌دهیم که برآوردگر جک‌نایف اصلاح شده لیو تعمیم یافته عملکرد بهتری نسبت به برآوردگر جک‌نایف لیو تعمیم یافته دارد. در ابتدا چند قضیه را بیان می‌کنیم.

قضیه ۱.۵.۲. * *(آکدنیز و آکدنیز دوران، ۲۰۱۲) فرض کنید $\boldsymbol{D}$ یک ماتریس قطری معین مثبت باشد. برآوردگر جک‌نایف اصلاح شده لیو تعمیم یافته ماتریس واریانس- کوواریانس کمتری نسبت به برآوردگر جک‌نایف لیو تعمیم یافته دارد.*

برهان. با استفاده از رابطه‌های (۱۳.۲) و (۱۹.۲) داریم

$$\begin{aligned}\operatorname{cov}(\hat{\gamma}_{JGLE}(\mathbf{D})) &= \sigma^۲\left[\mathbf{I}-(\boldsymbol{\Lambda}+\mathbf{I})^{-۲}(\mathbf{I}-\mathbf{D})^{-۲}\right]\boldsymbol{\Lambda}^{-۱}\left[\mathbf{I}-(\boldsymbol{\Lambda}+\mathbf{I})^{-۲}(\mathbf{I}-\mathbf{D})^{۲}\right]' \\ &= \sigma^۲\mathbf{V}\mathbf{U}\boldsymbol{\Lambda}^{-۱}\mathbf{U}'\mathbf{V}'\end{aligned} \tag{۲۷.۲}$$

و همچنین

$$\begin{aligned}\operatorname{cov}(\hat{\gamma}_{MJGLE}(\mathbf{D})) &= \sigma^۲\boldsymbol{\Phi}\boldsymbol{\Lambda}^{-۱}\boldsymbol{\Phi}' \\ &= \sigma^۲\mathbf{V}\mathbf{U}\mathbf{V}\boldsymbol{\Lambda}^{-۱}\mathbf{V}'\mathbf{U}'\mathbf{V}'\end{aligned} \tag{۲۸.۲}$$

که در آن $\mathbf{V}=\mathbf{I}-\mathbf{F}_{\mathbf{D}}, \mathbf{U}=\mathbf{I}+\mathbf{F}_{\mathbf{D}}$.

با استفاده از روابط (۲۷.۲) و (۲۸.۲) می‌توان نتیجه گرفت

$$\begin{aligned}\operatorname{cov}(\hat{\gamma}_{JGLE}(\mathbf{D}))-\operatorname{cov}(\hat{\gamma}_{MJGLE}(\mathbf{D})) &= \sigma^۲\mathbf{V}\mathbf{U}\boldsymbol{\Lambda}^{-۱}\mathbf{U}'\mathbf{V}'-\sigma^۲\mathbf{V}\mathbf{U}\mathbf{V}\boldsymbol{\Lambda}^{-۱}\mathbf{V}'\mathbf{U}'\mathbf{V}' \\ &= \sigma^۲\mathbf{V}\mathbf{U}(\boldsymbol{\Lambda}^{-۱}-\mathbf{V}\boldsymbol{\Lambda}^{-۱}\mathbf{V}')\mathbf{U}'\mathbf{V}' \\ &= \sigma^۲\boldsymbol{\Sigma}\end{aligned} \tag{۲۹.۲}$$

که در آن Σ یک ماتریس قطری بوده و i–امین عنصر قطر این تفاضل برابر $\frac{\sigma^{2}(\lambda_i+d_i)^{3}(2+\lambda_i-d_i)^{3}}{\lambda_i(1+\lambda_i)^{5}}$ است. با توجه به این‌که این مقدار مثبت می‌باشد، با استفاده از تعریف آ.۱۷.۰، برهان کامل است. □

در ادامه قضیه‌ای می‌آوریم که شرط لازم و کافی را برای برتری برآوردگر جک‌نایف اصلاح شده لیو تعمیم یافته نسبت به برآوردگر جک‌نایف لیو تعمیم یافته به‌دست می‌دهد.

قضیه ۲.۵.۲. * *(آکدنیز و آکدنیز دوران، ۲۰۱۲) فرض کنید $\boldsymbol{D}$ یک ماتریس معین مثبت قطری $p\times p$ باشد. آن‌گاه اختلاف*

$$\boldsymbol{\Delta}_{2} = MSEM(\gamma, \hat{\gamma}_{JGLE}(\boldsymbol{D})) - MSEM(\gamma, \hat{\gamma}_{MJGLE}(\boldsymbol{D}))$$

یک ماتریس نامنفی است اگر و تنها اگر در نامساوی زیر صدق کند

$$\gamma' \left[\mathbf{L}^{-1}(\sigma^{2}\boldsymbol{\Sigma}+\mathbf{F}_{\boldsymbol{D}}^{2}\gamma\gamma'\mathbf{F}_{\boldsymbol{D}}'^{2})\mathbf{L}'^{-1}\right]^{-1}\gamma \leqslant 1$$

که در آن $\mathbf{L}$ ماتریس معین مثبت با بعد $p\times p$ می باشد.

برهان. با توجه به روابط (۱۴.۲) و (۲۰.۲) داریم

$$\begin{aligned}\boldsymbol{\Delta}_{2} &= \operatorname{cov}(\hat{\gamma}_{JGLE}(\mathbf{D}) - \operatorname{cov}(\hat{\gamma}_{MJGLE}(\mathbf{D})) + (\mathbf{I}-\mathbf{F_D})^{2}\gamma\gamma'(\mathbf{I}-\mathbf{F_D})'^{2}\\ &\quad - (\boldsymbol{\Lambda}+\mathbf{I})^{-1}\mathbf{W}(\mathbf{I}-\mathbf{D})\gamma\gamma'(\mathbf{I}-\mathbf{D})\mathbf{W}'(\boldsymbol{\Lambda}+\mathbf{I})^{-1}\\ &= \sigma\boldsymbol{\Sigma}+\mathbf{F_D^{2}}\gamma\gamma'\mathbf{F_D'^{2}} - \mathbf{F_D}\mathbf{W}\gamma\gamma'W'\mathbf{F_D'}.\end{aligned} \tag{۳۰.۲}$$

با توجه به قضیه ۱.۵.۲، ماتریس Σ معین مثبت است. بنابراین، اختلاف $\boldsymbol{\Delta}_2$ معین نامنفی است اگر و تنها اگر $\mathbf{L}^{-1}\boldsymbol{\Delta}_2\mathbf{L}'^{-1}$ معین نامنفی باشد. دقت داشته باشید که ماتریس $\mathbf{L}^{-1}\boldsymbol{\Delta}_2\mathbf{L}'^{-1}$ را می‌توان به صورت زیر نوشت

$$\mathbf{L}^{-1}\boldsymbol{\Delta}_{2}\mathbf{L}'^{-1} = \mathbf{L}^{-1}(\sigma^{2}\boldsymbol{\Sigma}+\mathbf{F_D^{2}}\gamma\gamma'\mathbf{F_D'^{2}})\mathbf{L}'^{-1} - \gamma\gamma'.$$

اختلاف $\boldsymbol{\Delta}_2$ ماتریس معین نامنفی است اگر و تنها اگر $\mathbf{L}^{-1}\boldsymbol{\Delta}_2\mathbf{L}'^{-1}$ ماتریس معین نامنفی باشد. در حالی‌که ماتریس $(\sigma^{2}\boldsymbol{\Sigma}+\mathbf{F_D^{2}}\gamma\gamma'\mathbf{F_D'^{2}})$ متقارن و معین مثبت است، با استفاده از لم ب.۳۲.۰، عبارت $\mathbf{L}^{-1}\Delta_2\mathbf{L}'^{-1}$ معین نامنفی است اگر و تنها اگر رابطه‌ی زیر برقرار باشد

$$\gamma' \left[\mathbf{L}^{-1}(\sigma^{2}\boldsymbol{\Sigma}+\mathbf{F_D^{2}}\gamma\gamma'\mathbf{F_D'^{2}})\mathbf{L}'^{-1}\right]^{-1}\gamma \leqslant 1.$$

که در این صورت، برهان کامل است. □

۶.۲ مثال عددی

به منظور بررسی رفتار برآوردگرهای معرفی شده در این فصل در یک مثال واقعی، داده‌های مربوط به ارزش خانه‌ها را در نظر می‌گیریم. این داده‌ها شامل ۹۲ خانه مجزا در منطقه اتاوا[5] کشور کانادا، که در

[5] Ottowa

طول سال ۱۹۸۷ فروخته شده است، می‌باشد (به یاتچیو[۶]، ۲۰۰۳ مراجعه کنید).
الگوی مدل پارامتری به صورت زیر است

$$\begin{aligned} y_i = \beta_0 + \beta_1(\mathrm{frplc})_i + \beta_2(\mathrm{grge})_i + \beta_3(\mathrm{lux})_i + \beta_4(\mathrm{anginc})_i + \beta_5(\mathrm{dhwy})_i \\ + \beta_6(\mathrm{lot\ area})_i + \beta_7(\mathrm{nrbed})_i + \beta_8(\mathrm{usrspc})_i + \epsilon_i \end{aligned} \tag{۳۱.۲}$$

که در آن y قیمت فروش خانه (متغیر پاسخ) است، $frplc = 1$ به این معناست که خانه دارای شومینه (بخاری) است، اگر خانه دارای گاراژ باشد $grge = 1$ خواهد بود. $lux = 1$ دارا بودن خانه از فضای مجلل است، $avginc$ متوسط درآمد محلی را نشان می‌دهد. فاصله خانه از بزرگراه به عنوان یک متغیر مستقل به صورت $dhwy$ نشان داده می‌شود. $arealot$ متغیر پیوسته‌ای است که اندازه زمین در فوت مربع را نشان می‌دهد. $nrbed$ تعداد اتاق خواب است و در نهایت $usespcis$ متغیری است که زیر بنا بر حسب فوت را بیان می‌کند.
ماتریس ضریب همبستگی بین متغیرها در جدول ۳.۲ نشان داده شده است. برنامه مربوط به این مثال

جدول ۳.۲: همبستگی بین متغیرها

usespc	nrbed	lot area	dhwy	avginc	lux	grge	frplc	sp	
۰/۴۷۷	۰/۳۷۱	۰/۱۴۶	−۰/۱	۰/۳۴۲	۰/۵۰۶	۰/۲۹۹	۰/۳۳۷	۱	sp
۰/۴۶۳	۰/۳۱۴	۰/۱۵۹	۰/۱۰۸	۰/۳۸۱	۰/۰۵۴	۰/۱۴۶	۱	۰/۳۳۷	frplc
۰/۲۲۳	۰/۱۳۵	۰/۱۷۰	۰/۰۵۸	۰/۰۲۷	۰/۰۷۹	۱	۱۴۶/	۰/۲۹۹	grge
۰/۱۴۸	۰/۱۹۷	۰/۰۰۲	۰/۰۹۴	۰/۰۰۱	۱	۰/۰۷۹	۰/۰۵۴	۰/۵۰۶	lux
۰/۲۷۸	۰/۱۱۶	۰/۱۳۳	−۰/۱۱۰	۱	۰/۰۰۱	۰/۰۲۷	۰/۳۸۱	۰/۳۴۲	avginc
۰/۰۲۳	۰/۰۲۹	۰/۰۸۱	۱	−۰/۱۱۰	۰/۰۹۴	۰/۰۵۸	۰/۱۰۸	−۰/۱	dhwy
۰/۱۵۴	۰/۳۲۴	۱	۰/۰۸۱	۰/۱۳۳	۰/۰۰۲	۰/۱۷۰	۰/۱۵۹	۰/۱۴۶	lot area
۰/۵۷۷	۱	۰/۳۲۴	۰/۰۲۹	۰/۱۱۶	۰/۱۹۷	۰/۱۳۵	۰/۳۱۴	۰/۳۷۱	nrbed
۱	۰/۵۷۷	۰/۱۵۴	۰/۰۲۳	۰/۲۷۸	۰/۱۴۸	۰/۲۲۳	۰/۴۶۳	۰/۴۷۷	usespc

در پیوست ت.۴ آمده است.
بلسلی[۷] (۱۹۹۱) متذکر شد که همبستگی و همخطی معنی یکسانی ندارند. ممکن است داده‌هایی داشته باشیم که دارای همخطی باشند در حالی‌که همبستگی بین جفت متغیرهای توصیفی کم‌تر از ۰/۷ است. از این‌رو، همبستگی‌های زوجی برای بررسی همخطی چندگانه کافی نیستند؛ به همین علت آماره کاپا (عدد شرطی) و عامل تورم واریانس (VIF) را برای این داده‌ها بررسی می‌کنیم. (پیوست‌های پ.۱ و پ.۲ را ببینید.)
فرض کنید $\mathbf{X}_{92\times 9}$ ماتریس مشاهدات شامل متغیرهای: $frplc$ ، $grge$ ، lux ، $avginc$ ، $dhwy$ ،

[۶] Yatchew
[۷] Belsley

lot area ، *nrbed* ، *usespc* و یک ستون اضافی از یک‌ها باشد. مقادیر ویژه ماتریس $\mathbf{X'X}$ عبارتند از

$$\lambda_۱ = ۱/۴۷,\ \lambda_۲ = ۳/۷۷,\ \lambda_۳ = ۴/۵۲,\ \lambda_۴ = ۱۵/۳۳,\ \lambda_۵ = ۱۸/۵۷$$

$$\lambda_۶ = ۲۰/۹۷,\ \lambda_۷ = ۴۱/۷۹,\ \lambda_۸ = ۲۷۱/۱۵,\ \lambda_۹ = ۲۳۹۵۱۳/۶۸.$$

همچنین آماره کاپا برای این مثال برابر است با

$$\kappa = ۴۰۳/۲۷ > ۱۰$$

که نشان می‌دهد $\mathbf{X}$ دارای همخطی شدید است.

VIF معیار دیگری برای تشخیص همخطی است و برای این داده‌ها مقادیر VIF به ترتیب عبارتند از

$$۵۹/۵۶۱۰,\ ۴/۷۵۷۶,\ ۳/۰۲۶۴,\ ۱/۱۲۱۳,\ ۲۳/۹۰۸۵,$$

$$۴/۸۲۱۷,\ ۲۶/۰۸۱۸,\ ۵۰/۱۵۴۲,\ ۳۱/۳۵۶۴.$$

بعضی از VIF ها بزرگتر از ۳۰ هستند که دلالت بر همخطی شدید بین متغیرهاست. در این حالت، ضرایب رگرسیونی بی معنی‌اند بنابراین، استنباط و پیش‌گویی معتبر با استفاده از کمترین توان‌های دوم مشکل است. برای رفع این مشکل به جای استفاده از روش کمترین توان‌های دوم از برآوردگرهای رگرسیون ریج یا برآوردگر نوع لیو استفاده می‌کنیم. برآوردکردن d نگرانی اصلی در این مثال نیست، به هر حال تکنیک‌های مختلف بسیاری برای برآوردکردن d توسط لیو در سال ۱۹۹۳ پیشنهاد شده است. مدل (۳۱.۲) یک مدل رگرسیون خطی چندگانه به صورت (۷.۲) می‌باشد. در جداول ۴.۲ ، ۵.۲ و ۶.۲ برآوردگرهای $\hat{\gamma}(d)$، $\hat{\gamma}_{JLE}(d)$ و $\hat{\gamma}_{MJLE}(d)$ برای این داده‌ها به ازای $d = ۰/۵۵, ۰/۷۵, ۱$ به‌دست آمدند. برآوردها و طول توان دوم ضرایب رگرسیون $(\beta'\beta)$ و مقادیر MSE ها که برابر $trace(MSEM)$ است به ازای مقادیر مختلف d در جداول ۴.۲، ۵.۲ و ۶.۲ آمده است.

$\hat{\gamma}$ برآورد $\hat{\gamma}_{OLSE}$ در مدل کانونی است و برآورد واریانس از رابطه‌ی زیر محاسبه می‌گردد

$$\hat{\sigma}^۲ = (y - Z\hat{\gamma}_{OLSE})'(y - Z\hat{\gamma}_{OLSE})/(n-p).$$

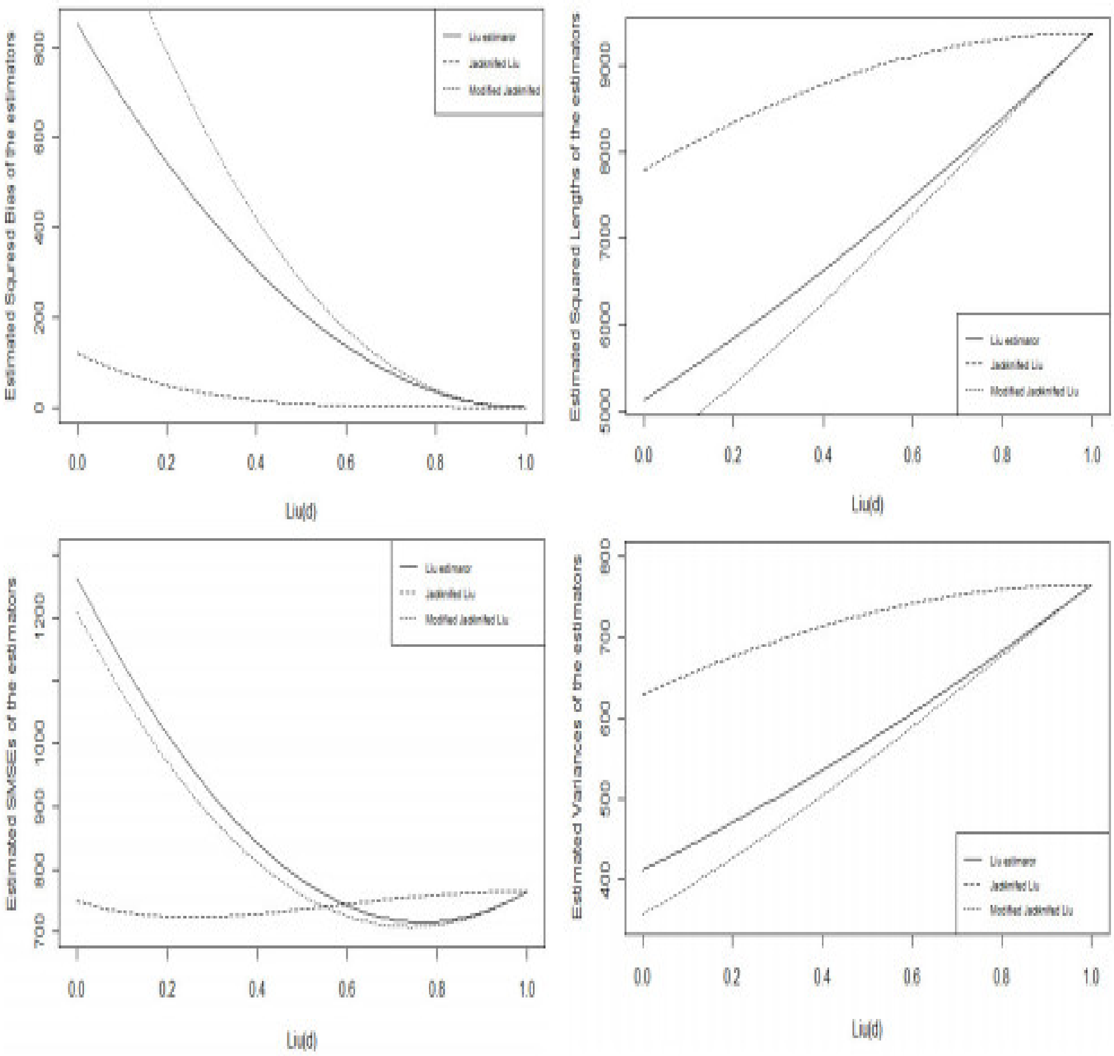
شکل ۳.۲ نواحی که هر برآوردگر از لحاظ طول توان دوم، توان دوم اریبی، واریانس و $MSEM$ نسبت به برآوردگرهای دیگر برتری دارد را نشان می‌دهد. در شکل ۳.۲ سه تابع $MSEM$ با سه برآوردگر داریم. طبق شکل $MSEM$ در مقابل d واضح است که برآوردگر جک‌نایف اصلاح شده لیو عملکرد ضعیف‌تری در حالت $۰ < d < ۰/۵۳$ نسبت به برآوردگر جک‌نایف لیو دارد، در حالی‌که وقتی $d > ۰/۵۳$ است برآوردگر جک‌نایف اصلاح‌شده لیو اندکی بهتر از برآوردگر جک‌نایف لیو در این مثال است.

جدول ۴.۲: برآورد β برای مقادیر مختلف d

$d=۱$ $\hat{\beta}_{OLSE}$	$d=۰/۷۵$ $\hat{\beta}(d)$	$d=۰/۵۵$ $\hat{\beta}(d)$	
۶۲/۵۱۹۷	۵۶/۲۱۳۶	۵۱/۱۶۸۷	$\beta_۰$
۶/۴۲۷۶	۵/۹۲۹۳	۵/۵۳۰۶	frplc
۱۳/۱۱۱۸	۱۳/۱۸۰۰	۱۳/۲۳۴۵	grge
۶۶/۴۲۵۸	۶۳/۱۳۲۰	۶۰/۴۹۷۱	lux
۰/۶۰۶۶	۰/۶۵۰۲	۰/۶۸۵۱	avginc
-۱۱/۱۷۰۷	-۱۰/۳۷۰۷	-۹/۷۳۰۷	dhwy
۰/۷۲۹۹	۰/۹۰۶۵	۱/۰۴۷۸	lot area
۳/۶۸۸۰	۴/۸۱۵۶	۵/۷۱۷۶	nrbed
۲۶/۶۹۱۳	۲۶/۰۰۹۱	۲۵/۴۶۳۴	usespc
۹۳۸۶	۸۱۶۳	۷۲۶۱/۲	$\beta'\beta$
۷۶۴/۷	۷۱۵/۸	۷۵۹/۹	SMSE

جدول ۵.۲: برآورد β برای مقادیر مختلف d

$d=۱$ $\hat{\beta}_{OLSE}$	$d=۰/۷۵$ $\hat{\beta}_{JLE}(d)$	$d=۰/۵۵$ $\hat{\beta}_{JLE}(d)$	
۶۲/۵۱۹۷	۶۱/۸۷۳۵	۶۰/۴۲۶۰	$\beta_۰$
۶/۴۲۷۶	۶/۳۶۵۳	۶/۲۲۵۹	frplc
۱۳/۱۱۱۸	۱۳/۱۲۴۶	۱۳/۱۵۳۳	grge
۶۶/۴۲۵۸	۶۶/۲۴۰۹	۶۵/۸۲۶۸	lux
۰/۶۰۶۶	۰/۶۱۰۷	۰/۶۱۹۹	avginc
-۱۱/۱۷۰۷	-۱۱/۱۰۲۰	-۱۰/۹۴۷۹	dhwy
۰/۷۲۹۹	۰/۷۴۹۱	۰/۷۹۲۱	lot area
۳/۶۸۸۰	۳/۷۸۵۰	۴/۰۰۲۲	nrbed
۲۶/۶۹۱۳	۲۶/۶۹۲۴	۲۶/۶۹۴۸	usespc
۹۳۸۶	۹۲۸۰	۹۰۴۵/۷	$\beta'\beta$
۷۶۴/۷	۷۵۶/۱	۷۴۰/۴	SMSE



شکل ۳.۲: مقایسه برآوردگرها به ازای مقادیر مختلف d ، $0 < d < 1$

جدول ۶.۲: برآورد β برای مقادیر مختلف d

	$d=\circ/۵۵$ $\hat{\beta}_{MJLE}(d)$	$d=\circ/۷۵$ $\hat{\beta}_{MJLE}(d)$	$d=۱$ $\hat{\beta}_{OLSE}$
$\beta_\circ$	۴۹/۴۵۸۱	۵۵/۶۳۳۱	۶۲/۵۱۹۷
frplc	۵/۳۶۹۶	۵/۸۷۴۰	۶/۴۲۷۶
grge	۱۳/۲۷۰۴	۱۳/۱۹۱۸	۱۳/۱۱۱۸
lux	۵۹/۹۶۸۲	۶۲/۹۵۹۲	۶۶/۴۲۵۸
avginc	۰/۶۹۵۹	۰/۶۵۳۹	۰/۶۰۶۶
dhwy	۹/۵۴۶۴-	۱۰/۳۰۸۶-	۱۱/۱۷۰۷-
lot area	۱/۰۹۷۳	۰/۹۲۳۵	۰/۷۲۹۹
nrbed	۵/۹۸۲۱	۴/۹۰۴۰	۳/۶۸۸۰
usespc	۲۵/۴۴۶۰	۲۶/۰۰۶۶	۲۶/۶۹۱۳
$\beta'\beta$	۷۰۲۳/۳	۸۰۷۵/۴	۹۳۸۶
SMSE	۷۳۷/۹	۷۰۸/۲	۷۶۴/۷

۷.۲ مطالعه شبیه‌سازی

اکنون رفتار برآوردگرهای پیشنهادی را با استفاده از تکنیک شبیه‌سازی مونت کارلو بررسی می‌کنیم. یک مطالعه شبیه‌سازی، مقایسه عملکرد برآوردگرهای گوناگون را تحت درجه‌های مختلف هم‌خطی ممکن می‌سازد. برای شبیه‌سازی فرض می‌کنیم مدل رگرسیون خطی (۱.۱) مفروض است. متغیرهای توضیحی طبق رابطه‌ی زیر تولید می‌شوند (ام سی دونالد و گالارنیو[۸]، ۱۹۷۵)

$$x_{ij}=\left(۱-\rho^۲\right)^{۱/۲}w_{ij}+\rho w_{ip} \quad i=۱,\cdots,n;\ j=۱,\cdots,p \qquad (۳۲.۲)$$

به طوری‌که w_{ij} اعداد شبه‌تصادفی نرمال استاندارد مستقل و همبستگی بین هر دو متغیر پیش‌بین به شکل $\rho^۲$ مشخص می‌شود. در این مطالعه، اثرهای درجه‌های هم‌خطی مختلف روی برآوردگرها را با مقادیر $\rho=\circ/۸,\circ/۹۵,\circ/۹۹$ بررسی می‌کنیم. اعداد شرطی متناظر اشاره ضعیفی به هم‌خطی شدید دارد. داده‌ها را با اندازه نمونه $n=۱۵,۵\circ,۱\circ\circ$ شبیه‌سازی می‌کنیم. انحراف استاندارد در این مطالعه به صورت $\sigma=\circ/۱,۱,۵$ در نظر گرفته شده است.

متغیر پاسخ y_i، $(i=۱,\cdots,n)$ با توجه به فرمول زیر تولید می‌شود

$$y_i=\beta_\circ+\beta_۱x_{i۱}+\beta_۲x_{i۲}+\cdots+\beta_px_{ip}+\epsilon_i, \quad i=۱,\cdots,n \qquad (۳۳.۲)$$

که در آن ϵ_i اعداد شبه‌تصادفی نرمال و مستقل هستند و دارای میانگین صفر و واریانس $\sigma^۲$ می‌باشد. وقتی برآوردگرها را محاسبه می‌کنیم مدل اصلی را به مدل کانونی تبدیل می‌کنیم و برآوردگر γ را به‌دست می‌آوریم، سپس γ به β تبدیل می‌شود. برای هر انتخاب $\rho,n,\sigma^۲$ آزمایش ۱۰۰۰ بار تکرار می‌شود و

[۸] McDonald and Galarnaeu

ماتریس میانگین توان‌دوم خطا[۹] و قدر مطلق اریبی[۱۰] را بر اساس روابط زیر محاسبه می‌کنیم

$$MSEM(\hat{\beta}) = \frac{۱}{۱۰۰۰}\sum_{l=۱}^{۱۰۰۰} \parallel \hat{\beta}^{(l)} - \beta \parallel^{۲}, \tag{۳۴.۲}$$

$$ABIAS(\hat{\beta}) = \frac{۱}{۱۰۰۰}\sum_{l=۱}^{۱۰۰۰}\sum_{i=۱}^{p} |\hat{\beta}_{il} - \beta_i| \tag{۳۵.۲}$$

که در آن $\hat{\beta}^{(l)}$ برآورد پارامترها در l–امین شبیه‌سازی است و $\hat{\beta}_{il}$ ، برآورد i–امین پارامتر در l شبیه‌سازی است. $MSEM$ دلالت بر این دارد که بردار پارامتر تا چه اندازه درست برآورد شده است. برای هر تکرار، مقدار d نیز به صورت زیر برآورد می‌شود

$$\hat{d}_{mm} = ۱ - \hat{\sigma}^{۲}\left[\sum_{i=۱}^{p}\frac{۱}{\lambda_i(۱+\lambda_i)}\Big/\sum_{i=۱}^{p}\frac{\hat{\gamma}^{۲}_{OLSE}}{(۱+\lambda_i)^{۲}}\right] \tag{۳۶.۲}$$

دقت کنید که $\hat{d}_{mm}$ طوری انتخاب شده است که MSE برآوردگر لیو را می‌نیمم می‌کند (لیو، ۱۹۹۳). در اینجا $\hat{\gamma} = \mathbf{T}'\hat{\beta}$ و $\hat{\sigma}$ برآوردگر کمترین توان‌های دوم γ و $\sigma^{۲}$ هستند و $\mathbf{T}$ ماتریس بردارهای ویژه متناظر با مقادیر ویژه $\mathbf{X}'\mathbf{X}$ است.

جدول ۷.۲: برآورد $MSEM$ برای $OLSE$ ، LE ، JLE و $MJLE$ به ازای (الف) $n = ۱۵$ و $p = ۳$، (ب) $n = ۵۰$ و $p = ۳$ ، (پ) $n = ۱۰۰$ و $p = ۳$

$\sigma = ۵$			$\sigma = ۱$			$\sigma = ۰/۱$			
$\rho = ۰/۹۹$	$\rho = ۰/۹۵$	$\rho = ۰/۸$	$\rho = ۰/۹۹$	$\rho = ۰/۹۵$	$\rho = ۰/۸$	$\rho = ۰/۹۹$	$\rho = ۰/۹۵$	$\rho = ۰/۸$	
									(الف)
۱/۵۰۰۱۳	۱/۰۶۵۶۴	۰/۹۶۹۸۷	۱/۱۱۰۷۰	۰/۸۵۳۸۶	۰/۸۲۷۸۹	۰/۷۷۱۸۹	۰/۷۹۴۸۰	۰/۸۱۹۳۹	$\hat{\beta}$
۱/۱۸۸۰۷	۱/۰۰۵۳۸	۰/۹۵۷۱۰	۰/۹۳۹۶۸	۰/۸۲۲۴۰	۰/۸۲۴۳۳	۰/۷۶۱۲۴	۰/۷۹۲۵۶	۰/۸۱۸۵۷	$\hat{\beta}(d)$
۱/۳۴۴۲۸	۱/۰۳۷۰۷	۰/۹۶۱۳۳	۱/۰۲۵۳۸	۰/۸۳۶۸۲	۰/۸۲۳۸۰	۰/۷۷۰۴۴	۰/۷۹۴۴۷	۰/۸۱۹۳۸	$\hat{\beta}_{JLE}(d)$
۱/۱۴۶۶۹	۰/۹۹۷۲۱	۰/۹۵۵۲۲	۰/۹۱۶۴۱	۰/۸۱۷۴۳	۰/۸۲۳۱۴	۰/۷۶۰۱۷	۰/۷۹۲۵۱	۰/۸۱۸۵۷	$\hat{\beta}_{MJ}(d)$
۲۴۷/۳۹	۴۹	۷/۷۴	۱۸۶/۹۷	۲۹/۰۹	۱۲/۰۳	۲۳۱/۸۶	۴۶/۱۹	۸/۳۳	κ
									(ب)
۱/۰۰۶۷۶	۰/۹۷۷۸۳	۰/۹۶۹۳۷	۰/۹۱۳۹۰	۰/۸۹۷۹۸	۰/۸۸۴۹۸	۰/۸۵۵۵۱	۰/۸۵۴۴۴	۰/۸۴۸۹۳	$\hat{\beta}$
۰/۹۸۴۶۷	۰.۹۷۳۶۷	۰.۹۷۰۶۱	۰.۹۰۰۱۷	۰.۸۹۴۵۱	۰.۸۸۶۷۶	۰.۸۵۴۸۲	۰.۸۵۴۳۰	۰.۸۴۸۹۳	$\hat{\beta}(d)$
۰/۹۹۵۶۱	۰/۹۷۵۳۱	۰/۹۶۸۹۵	۰/۹۰۷۲۵	۰/۸۹۶۲۰	۰/۸۸۴۸۳	۰/۸۵۵۴۹	۰/۸۵۴۴۴	۰/۸۴۸۹۳	$\hat{\beta}_{JLE}(d)$
۰/۹۸۱۶۵	۰/۹۷۳۰۴	۰/۹۷۰۶۴	۰/۸۹۸۱۷	۰/۸۹۳۸۱	۰/۸۸۶۷۱	۰/۸۵۴۸۰	۰/۸۵۴۳۰	۰/۸۴۹۸۳	$\hat{\beta}_{MJ}(d)$
۱۳۳/۱۵	۲۸/۶۰	۷/۹۴	۱۷۸/۶۴	۳۳/۹۷	۷/۱۸	۱۸۰/۲۲	۳۲/۹۱	۶/۸۱	κ
									(پ)
۰/۹۸۷۵۶	۰/۹۷۹۳۵	۰/۹۷۸۱۳	۰/۹۲۸۸۷	۰/۹۲۵۶۹	۰/۹۲۵۹۸	۰/۸۹۵۵۵	۰/۸۹۱۴۴	۰/۸۹۲۱۴	$\hat{\beta}$
۰/۹۸۱۶۷	۰/۹۷۸۵۲	۰/۹۷۹۴۹	۰/۹۲۵۱۷	۰/۹۲۴۸۸	۰/۹۲۶۴۰	۰/۸۹۵۳۷	۰/۸۹۱۴۱	۰/۸۹۲۱۵	$\hat{\beta}(d)$
۰/۹۸۴۴۷	۰/۹۷۸۷۲	۰/۹۷۸۱۱	۰/۹۲۷۲۰	۰/۹۲۵۳۱	۰/۹۲۵۹۴	۰/۸۹۵۵۵	۰/۸۹۱۴۴	۰/۸۹۲۱۴	$\hat{\beta}_{JLE}(d)$
۰/۹۸۰۸۵	۰/۹۷۸۳۵	۰/۹۷۹۵۸	۰/۹۲۴۵۹	۰/۹۲۴۶۹	۰/۹۲۶۳۷	۰/۸۹۵۳۶	۰/۸۹۱۴۱	۰/۸۹۲۱۵	$\hat{\beta}_{MJ}(d)$
۱۷۲/۹۵	۳۵/۹۴	۶/۹۲	۱۴۰/۷۸	۳۹/۴۹	۷/۳۵	۱۵۷/۶۹	۳۰/۰۴	۶/۴۱	κ

نتایج عددی شبیه‌سازی‌ها در جداول ۷.۲ و ۸.۲ خلاصه شده‌اند. برنامه مربوط به این مثال در پیوست ت.۵ آورده شده است.

[۹] Mean square error matrix

[۱۰] Absolute bias

جدول ۸.۲: برآورد قدر مطلق اریبی برای $OLSE$ ، LE ، JLE و $MJLE$ به ازای (الف) $n = ۱۵$ و $p = ۳$ ، (ب) $n = ۵۰$ و $p = ۳$، (پ) $n = ۱۰۰$ و $p = ۳$

$\sigma = ۵$			$\sigma = ۱$			$\sigma = ۰/۱$			
$\rho = ۰/۹۹$	$\rho = ۰/۹۵$	$\rho = ۰/۸$	$\rho = ۰/۹۹$	$\rho = ۰/۹۵$	$\rho = ۰/۸$	$\rho = ۰/۹۹$	$\rho = ۰/۹۵$	$\rho = ۰/۸$	
									(الف)
۱/۷۱۹۷۸	۱/۶۷۹۷۵	۱/۶۷۰۹۹	۱/۵۶۱۳۲	۱/۵۲۲۵۸	۱/۵۳۹۴۶	۱/۴۶۶۴۳	۰۱/۵۰۹۲۱	۱/۳۳۰۶۷	$\hat{\beta}(d)$
۱/۷۶۴۳۸	۱/۶۷۹۶۶	۱/۶۶۴۸۹	۱/۵۷۶۵۵	۱/۵۱۴۱۸	۱/۵۳۸۸۲	۱/۴۶۶۴۸	۱/۵۰۹۷۶	۱/۳۲۹۸۰	$\hat{\beta}_{JLE}(d)$
۱/۷۱۳۸۳	۱/۶۷۹۵۷	۱/۶۷۱۹۲	۱/۵۵۹۵۰	۱/۵۲۳۶۳	۱/۵۳۷۳۰	۱/۴۶۶۲۹	۱/۵۰۹۱۹	۱/۳۳۰۶۸	$\hat{\beta}_{MJ}(d)$
۲۴۷/۳۹	۴۹	۷/۷۴	۱۸۶/۹۷	۲۹/۰۹	۱۲/۰۳	۲۳۱/۸۶	۴۶/۱۹	۸/۳۳	κ
									(ب)
۱/۷۰۴۱۰	۱/۷۰۵۶۳	۱/۷۰۵۱۳	۱/۶۳۱۱۵	۱/۶۳۱۴۹	۱/۹۱۰۹۸	۱/۵۸۸۲۱	۱/۵۸۴۲۷	۱/۵۸۳۰۷	$\hat{\beta}(d)$
۱/۷۰۳۸۲	۱/۷۰۴۸۹	۱/۷۰۳۲۱	۱/۶۳۰۲۸	۱/۶۳۰۳۷	۱/۶۰۶۹۵	۱/۵۸۸۱۲	۱/۵۸۴۲۱	۱/۵۸۳۰۳	$\hat{\beta}_{JLE}(d)$
۱/۷۰۴۱۲	۱/۷۰۵۶۷	۱/۷۰۵۲۴	۱/۶۳۱۲۰	۱/۶۳۱۳۸	۱/۶۱۱۲۸	۱/۵۸۸۲۱	۱/۵۸۴۲۷	۱/۵۸۳۰۷	$\hat{\beta}_{MJ}(d)$
۱۳۳/۱۵	۲۸/۶۰	۷/۹۴	۱۷۸/۶۴	۳۳/۹۷	۷/۱۸	۱۸۰/۲۲	۳۲/۹۱	۶/۸۱۶۸۱	κ
									(پ)
۱/۷۱۲۷۰	۱/۷۰۹۷۱	۱/۷۱۱۸۶	۱/۶۶۲۶۰	۱/۶۶۳۱۶	۱/۶۶۴۹۱	۱/۶۳۳۴۸	۱/۶۲۹۷۰	۱/۶۲۷۵۰	$\hat{\beta}(d)$
۱/۷۱۲۶۲	۱/۷۰۹۱۳۷	۱/۷۱۰۲۳	۱/۶۶۲۴۸	۱/۶۶۲۶۵	۱/۶۶۴۷۱	۱/۶۳۳۴۹	۱/۶۲۹۶۹	۱/۶۲۷۴۸	$\hat{\beta}_{GLE}(d)$
۱/۷۱۲۶۹	۱/۷۰۹۸۰	۱/۷۱۲۰۷	۱/۶۶۲۵۲	۱/۶۶۳۱۳	۱/۶۶۹۹۱	۱/۶۳۳۴۸	۱/۶۲۹۷۰	۱/۶۲۷۵۰	$\hat{\beta}_{MJ}(d)$
۱۷۲/۹۵	۳۵/۹۴	۶/۹۲	۱۴۰/۷۸	۳۹/۴۹	۷/۳۵	۱۵۷/۶۹	۳۰/۰۴	۶/۴۱	κ

در جدول ۷.۲ قسمت الف، برای اندازه نمونه‌های کوچک $n = ۱۵$ برآوردگر جک‌نایف اصلاح‌شده لیو عملکرد بهتری نسبت به دیگر برآوردگرها در همه‌ی حالت‌های ضریب همبستگی بین پیش‌بین‌ها و انحراف استانداردها دارد. در قسمت ب و پ جدول ۷.۲ ، برای اندازه نمونه‌های متوسط $n = ۵۰$ و اندازه نمونه‌های بزرگ $n = ۱۰۰$، برآوردگر جک‌نایف اصلاح‌شده لیو همانند برآوردگر جک‌نایف لیو عمل می‌کند؛ به هر حال برآوردگر جک‌نایف اصلاح‌شده لیو تحت حالت همخطی قوی و خیلی قوی بهتر از برآوردگر جک‌نایف لیو عمل می‌کند. به عبارت دیگر؛ برآوردگر جک‌نایف اصلاح‌شده لیو نسبت به دیگر برآوردها در معیار میانگین توان دوم خطا در مواقعی که درجه همخطی بین متغیرهای پیش‌بین افزایش می‌یابد کارایی بیشتری دارد.

در جدول ۸.۲ قسمت الف، برای اندازه نمونه‌های کوچک $n = ۱۵$ برآوردگر جک‌نایف اصلاح‌شده لیو تحت همخطی چندگانه قوی و خیلی قوی کمترین میانگین قدر مطلق اریبی را در بین دیگر برآوردگرها برای واریانس‌های مختلف دارد. در قسمت الف و پ جدول ۸.۲، برای اندازه نمونه‌های متوسط $n = ۵۰$ و اندازه نمونه‌های بزرگ $n = ۱۰۰$، برآوردگر جک‌نایف لیو تعمیم یافته کمترین میانگین قدر مطلق اریبی را در بیشتر حالت‌ها دارد. در تمام حالت‌های اندازه نمونه، همخطی و واریانس خطا، برآوردگر جک‌نایف لیو بهتر از برآوردگر جک‌نایف لیو در معیار میانگین قدر مطلق اریبی دارد.

به طور خلاصه، بیشتر حالات در نظر گرفته شده در جدول ۸.۲، برآوردگر جک‌نایف لیو در اندازه نمونه‌های کوچک خیلی بهتر از دیگر برآوردگرها است به طوری‌که برآوردگر جک‌نایف اصلاح‌شده لیو نسبت به برآوردگرهای دیگر در حالت همخطی‌های شدید است.

از همه‌ی نتایج شبیه‌سازی می‌توانیم بگوییم که برآوردگر جک‌نایف اصلاح‌شده لیو نسبت به برآوردگرهای نوع لیو و جک‌نایف لیو در برخی حالت‌ها عملکرد قابل توجهی داشته است.

فصل ۳

برآوردگر لیو اصلاح شده لی و یانگ

۱.۳ مقدمه

بر اساس اطلاع پیشین[۱]، سوئیندل (۱۹۷۶) برآوردگر ریج اصلاح شده[۲] (MRE)، را به صورت زیر پیشنهاد کرد

$$\hat{\beta}_{MRE}(k, \mathbf{b}_\circ) = (\mathbf{X}'\mathbf{X} + k\mathbf{I})^{-1}(\mathbf{X}'y + k\mathbf{b}_\circ) \qquad (۱.۳)$$

که در آن $\mathbf{b}_\circ$ بردار غیر تصادفی است به عنوان مقداری برای β که نشان‌دهنده اطلاعی پیشین در خصوص پارامتر β می‌باشد. وقتی‌که k به بی‌نهایت میل می‌کند، برآوردگر ریج اصلاح شده به مقدار $\mathbf{b}_\circ$ میل می‌کند. همچنین به ازای $k = \circ$ همان برآوردگر کمترین توان‌های دوم به‌دست می‌آید. این‌که چگونه برآوردگر محاسبه می‌شود، نیاز به بیان قضیه‌ای دارد که در ابتدا آن را مطرح می‌کنیم.

قضیه ۱.۱.۳. * *(رائو[۳]، ۱۹۹۵) مدل خطی (۱.۱) را که در آن* $\mathrm{cov}(y) = \mathrm{cov}(\epsilon) = \sigma^{۲}\boldsymbol{\Sigma}$ *و* $\boldsymbol{\Sigma}$ *ماتریسی معین مثبت است و* $\boldsymbol{C} = (\boldsymbol{X}'\boldsymbol{\Sigma}^{-۱}\boldsymbol{X})$ *در نظر بگیرید. همچنین فرض کنید* $\mathbf{H}\beta = \boldsymbol{h}$*، که در آن* $\mathbf{H}$ *ماتریس* $q \times p$ *معلوم با رتبه* q *و* h *یک بردار با* q *مؤلفه و از پیش تعیین شده باشد، در این‌صورت برآوردگر کمترین توان‌های دوم تحت درستی قید* $\mathbf{H}\beta = \boldsymbol{h}$*، (که به آن برآوردگر کمترین توان‌های دوم محدود شده[۴]* $\tilde{\beta}$ *می‌گویند) عبارتست از*

الف-

$$\tilde{\beta} = \hat{\beta} - \boldsymbol{C}^{-۱}\mathbf{H}'(\mathbf{H}\boldsymbol{C}^{-۱}\mathbf{H}')^{-۱}(\mathbf{H}\hat{\beta} - \boldsymbol{h})$$

که دارای خواص زیر است

ب- تحت فرض $\mathbf{H}\beta = \boldsymbol{h}$*، برآوردگری نااریب برای* β *است.*

---

[۱] Prior information

[۲] Modified ridge estimator

[۳] Rao

[۴] Restricted estimator

ج-ماتریس واریانس- کوواریانس آن برابر است با
$\sigma^{۲}[\boldsymbol{C}^{-۱} - \boldsymbol{C}^{-۱}\mathbf{H}'(\mathbf{H}\boldsymbol{C}^{-۱}\mathbf{H}')\mathbf{H}\boldsymbol{C}^{-۱}].$

برهان. الف- برای به‌دست آوردن برآوردگر محدود شده تحت فرضیه $\mathbf{H}\beta = \mathbf{h}$ از روش کمینه سازی لاگرانژ استفاده میکنیم. فرض کنید ۲λ یک بردار با q مؤلفه ضریب لاگرانژ است. حال تابع زیر را نسبت به β و λ می‌نیمم می‌کنیم.

$$\begin{aligned} Q(\beta,\lambda) &= \epsilon'\mathbf{\Sigma}^{-۱}\epsilon + ۲\lambda'(\mathbf{H}\beta - \mathbf{h}) \\ &= (y - \mathbf{X}\beta)'\mathbf{\Sigma}^{-۱}(y - \mathbf{X}\beta) + ۲\lambda'(\mathbf{H}\beta - \mathbf{h}). \end{aligned} \tag{۲.۳}$$

با مشتق‌گیری $Q(\beta,\lambda)$ نسبت به λ و β و مساوی با صفر قرار دادن این مشتق داریم

$$\begin{aligned} \frac{\partial Q(\beta,\lambda)}{\partial\lambda} &= ۲(\mathbf{H}\beta - \mathbf{h}) \\ &= \mathbf{0} \Longrightarrow \mathbf{H}\beta = \mathbf{h} \\ \frac{\partial Q(\beta,\lambda)}{\partial\beta} &= -۲\mathbf{X}'\mathbf{\Sigma}^{-۱}(y - \mathbf{X}\beta) + ۲\mathbf{H}'\lambda = \mathbf{0} \end{aligned} \tag{۳.۳}$$

که در این‌صورت

$$\begin{aligned} &\Longrightarrow ۲\mathbf{X}'\mathbf{\Sigma}^{-۱}y - ۲\mathbf{H}'\lambda = ۲\mathbf{X}'\mathbf{\Sigma}^{-۱}\mathbf{X}\beta \\ \beta &= \frac{۱}{۲}\mathbf{C}^{-۱}(۲\mathbf{X}'\mathbf{\Sigma}^{-۱}y - ۲\mathbf{H}'\lambda) \\ &= \mathbf{C}^{-۱}\mathbf{X}'\mathbf{\Sigma}^{-۱}y - \mathbf{C}^{-۱}\mathbf{H}'\lambda \end{aligned} \tag{۴.۳}$$

و

$$\begin{aligned} \Longrightarrow \mathbf{H}\beta &= \mathbf{H}\mathbf{C}^{-۱}\mathbf{X}'\mathbf{\Sigma}^{-۱}y - \mathbf{H}\mathbf{C}^{-۱}\mathbf{H}'\lambda \\ &= \mathbf{h} \\ \Rightarrow \mathbf{H}\mathbf{C}^{-۱}\mathbf{H}'\lambda &= \mathbf{H}\mathbf{C}^{-۱}\mathbf{X}'\mathbf{\Sigma}^{-۱}y - h \\ \lambda &= [\mathbf{H}\mathbf{C}^{-۱}\mathbf{H}']^{-۱}\{\mathbf{H}\mathbf{C}^{-۱}\mathbf{X}'\mathbf{\Sigma}^{-۱}y - \mathbf{h}\} \end{aligned} \tag{۵.۳}$$

در نتیجه β به صورت زیر به‌دست می‌آید

$$\begin{aligned} \beta = {} & \mathbf{C}^{-۱}\mathbf{X}'\mathbf{\Sigma}^{-۱}y \\ & - \mathbf{C}^{-۱}\mathbf{H}'[\mathbf{H}\mathbf{C}^{-۱}\mathbf{H}']^{-۱}\{\mathbf{H}\mathbf{C}^{-۱}\mathbf{X}'\mathbf{\Sigma}^{-۱}y - \mathbf{h}\}. \end{aligned}$$

بنابراین، برآوردگر کمترین توان‌های دوم محدود شده عبارتست از
$\tilde{\beta} = \hat{\beta} - \mathbf{C}^{-۱}\mathbf{H}'(\mathbf{H}\mathbf{C}^{-۱}\mathbf{H}')^{-۱}(\mathbf{H}\hat{\beta} - \mathbf{h}).$

ب- برآوردگر حاصل تحت فرضیه $\mathbf{H}\beta = \mathbf{h}$، برآوردگری نااریب برای β می‌باشد زیرا

$$\begin{aligned} E(\tilde{\beta}) &= E(\hat{\beta}) - \mathbf{C}^{-1}\mathbf{H}'(\mathbf{H}\mathbf{C}^{-1}\mathbf{H}')^{-1}(\mathbf{H}E(\hat{\beta}) - \mathbf{h}) \\ &= \beta. \end{aligned}$$

ج- همچنین ماتریس کوواریانس $\hat{\beta}$ به صورت زیر به‌دست می‌آید

$$\begin{aligned} Var(\tilde{\beta}) &= Var\left(\hat{\beta} - \mathbf{C}^{-1}\mathbf{H}'(\mathbf{H}\mathbf{C}^{-1}\mathbf{H}')^{-1}(\mathbf{H}\hat{\beta} - \mathbf{h})\right) \\ &= Var(\hat{\beta}) + \mathbf{C}^{-1}\mathbf{H}'(\mathbf{H}\mathbf{C}^{-1}\mathbf{H}')^{-1}\mathbf{H}Var(\hat{\beta})\mathbf{H}'(\mathbf{H}\mathbf{C}^{-1}\mathbf{H}')^{-1}\mathbf{C}^{-1}\mathbf{H} \\ &- 2Cov\left(\hat{\beta}, \mathbf{C}^{-1}\mathbf{H}'(\mathbf{H}\mathbf{C}^{-1}\mathbf{H}')^{-1}\mathbf{H}\hat{\beta}\right) \\ &= \sigma^2\mathbf{C}^{-1} + \sigma^2\mathbf{C}^{-1}\mathbf{H}'(\mathbf{H}\mathbf{C}^{-1}\mathbf{H}')^{-1}\mathbf{H}\mathbf{C}^{-1}\mathbf{H}'(\mathbf{H}\mathbf{C}^{-1}\mathbf{H}')^{-1}\mathbf{C}^{-1}\mathbf{H} \\ &- 2\sigma^2\mathbf{C}^{-1} - \mathbf{C}^{-1}\mathbf{H}'(\mathbf{H}\mathbf{C}^{-1}\mathbf{H}')^{-1}\mathbf{H}\mathbf{C}^{-1} \\ &= \sigma^2[\mathbf{C}^{-1} - \mathbf{C}^{-1}\mathbf{H}'(\mathbf{H}\mathbf{C}^{-1}\mathbf{H}')^{-1}\mathbf{H}\mathbf{C}^{-1}]. \end{aligned}$$

□

یکی از بهترین روش‌ها برای مقابله با همخطی، اعمال محدودیت اضافی روی بردار پارامترهاست. سوامی[۵] و همکاران (۱۹۷۸) و ژانگ و یانگ[۶] (۲۰۰۷) روشی را برای برآورد پارامترها ارائه دادند که در آن مجموع توان دوم مانده‌ها تحت دو محدودیت خطی و کروی می‌نیمم می‌شود، به عبارت دیگر رگرسیون خطی محدود شده تبدیل به مسئله بهینه سازی زیر

$$Min_{\beta}(y - \mathbf{X}\beta)'(y - \mathbf{X}\beta)$$

با دو قید $\beta'\beta = d^2$ و $\beta = \mathbf{b}_0$ می‌شود. حال با استفاده از تلفیق می‌نیمم توان‌های دوم خطا و روش لاگرانژ، برآوردگر ریج اصلاح شده را برای این مدل به‌دست می‌آوریم. برای این منظور تابع زیر را که تابعی خطی از مجموع توان‌های دوم خطا و دو محدودیت اعمال شده روی فضای پارامترهاست، می‌نیمم می‌کنیم

$$L(\beta, d, \lambda) = y'y - 2\beta'\mathbf{X}'y + \beta'\mathbf{X}'\mathbf{X}\beta + k\beta'\beta - kd^2 + 2\lambda'\beta - 2\lambda'\mathbf{b}_0.$$

با مشتق گرفتن از $L(\beta, d, \lambda)$ نسبت به β و λ و مساوی با صفر قرار دادن این مشتق داریم

$$\begin{aligned} \frac{\partial L}{\partial \beta} &= -2\mathbf{X}'y + 2\mathbf{X}'\mathbf{X}\beta + 2k\beta + 2\lambda \\ &= \mathbf{0} \Rightarrow (\mathbf{X}'\mathbf{X} + k\mathbf{I})\beta = \mathbf{X}'y - \lambda \\ \beta &= (\mathbf{X}'\mathbf{X} + k\mathbf{I})^{-1}(\mathbf{X}'y - \lambda) \end{aligned} \tag{۶.۳}$$

[۵] Swamy

[۶] Zhang and Yang

و

$$\begin{aligned} \frac{\partial L}{\partial \lambda} &= \mathbf{0} \\ \Rightarrow \beta &= \mathbf{b}_0 \\ (\mathbf{X}'\mathbf{X}+k\mathbf{I})^{-1}(\mathbf{X}'y-\lambda) &= \mathbf{b}_0 \\ (\mathbf{X}'\mathbf{X}+k\mathbf{I})^{-1}\mathbf{X}'y-(\mathbf{X}'\mathbf{X}+k\mathbf{I})^{-1}\lambda &= \mathbf{b}_0 \\ (\mathbf{X}'\mathbf{X}+k\mathbf{I})^{-1}\mathbf{X}'y-\mathbf{b}_0 &= (\mathbf{X}'\mathbf{X}+k\mathbf{I})^{-1}\lambda \\ \lambda &= \mathbf{X}'y-(\mathbf{X}'\mathbf{X}+k\mathbf{I})\mathbf{b}_0 \end{aligned} \tag{۳.۷}$$

که در این صورت با جایگذاری مقدار λ در فرمول β داریم

$$\hat{\beta}_{MRE}(k,b_0)=(\mathbf{X}'\mathbf{X}+k\mathbf{I})^{-1}(\mathbf{X}'y+k\mathbf{b}_0).$$

همان‌طور که توسط سوئیندل (۱۹۷۶) اشاره شد، به نظر می‌رسد در نظر گرفتن اطلاع پیشین در کاربردها معقول و مفید است، به همین دلیل برآورد ضرایب رگرسیون را براساس آن ارزیابی می‌کنیم.

می‌دانیم که آماردان‌ها تنها با مشکل انتخاب بین برآوردگر کمترین توان‌های دوم و یک برآوردگر اریب رو به رو نیستند، بلکه آن‌ها با مشکل انتخاب بین دو برآوردگر اریب مواجه هستند. لذا؛ برخی از آماردان‌ها تعدادی از برآوردگرهای اریب را باهم نیز مقایسه کرده اند. ساکالو[۷] و همکاران (۲۰۰۱) برآوردگر ریج را با برآوردگر نوع لیو مقایسه کردند و برآوردگر جدیدی به جای برآوردگر کمترین توان‌های دوم، در مواقعی که همخطی وجود دارد پیشنهاد دادند. آکدنیز و ارول (۲۰۰۳) برآوردگرهای تقریباً نااریب ریج تعمیم یافته و لیو تعمیم یافته را با استفاده از معیار ماتریس میانگین توان‌های دوم مقایسه کردند. ساکالو و کاچیرانلار[۸](۲۰۰۸) برآوردگر جدیدی را محاسبه کرده و آن را با برآوردگر کمترین توان‌های دوم، برآوردگرهای نوع لیو مقایسه کردند.

در این فصل، برآوردگر جدیدی از نوع لیو را معرفی می‌کنیم. این برآوردگر که بر اساس اطلاع پیشین از بردار پارامترها در مدل رگرسیون خطی است، برآوردگر لیو اصلاح شده نام دارد. با استفاده از معیار ماتریس میانگین توان دوم خطا، این برآوردگر بهتر از برآوردگر کمترین توان‌های دوم، برآوردگر نوع لیو، برآوردگر ریج و برآوردگر ریج اصلاح شده عمل می‌کند. در بخش اولِ این فصل، برآوردگر لیو اصلاح شده به طورکامل معرفی شده است. در بخش دوم، کارایی برآوردگر لیو اصلاح شده با دیگر برآوردگرها مقایسه شده‌است. در بخش سوم، برای شرح‌دادن نتایج نظری یک مثال عددی و یک شبیه‌سازی مونت‌کارلو انجام شده است.

نتایج این فصل عمدتاً از مرجع یالیان لی و هو یانگ (۲۰۱۲) می‌باشد. که از متن آن مقاله برزوئی بیدگلی و آرشی (۱۳۹۳) استخراج شده است.

[۷] Sakalhoglu

[۸] Sakahoglu and Kaciranlar

۲.۳ برآوردگر لیو اصلاح شده

فرض کنید اطلاعی پیشین در خصوص β به صورت $\beta = \mathbf{b}_\circ$ در اختیار داریم، همچنین در برآورد β مشکل همخطی وجود دارد. بنابراین در این حالت، می‌توانیم از برآوردگر ریج استفاده کرده و اطلاع پیشین را به منظور اصلاح آن به کار بریم. به همین منظور در این بخش برآوردگر لیو اصلاح شده که برآوردگر اریبی از نوع برآوردگر ریج است و از اطلاع پیشین بهره می‌گیرد، را معرفی می‌نماییم.
فرض کنید

$$\begin{aligned}
\mathbf{T}_k &= (\mathbf{X}'\mathbf{X}+k\mathbf{I})^{-1}\mathbf{X}'\mathbf{X} \\
&= (\mathbf{X}'\mathbf{X}+k\mathbf{I})^{-1}(\mathbf{X}'\mathbf{X}+k\mathbf{I}-k\mathbf{I}) \\
&= (\mathbf{X}'\mathbf{X}+k\mathbf{I})^{-1}(\mathbf{X}'\mathbf{X}+k\mathbf{I})-k\mathbf{I}(\mathbf{X}'\mathbf{X}+k\mathbf{I})^{-1} \\
&= \mathbf{I}-k(\mathbf{X}'\mathbf{X}+k\mathbf{I})^{-1}.
\end{aligned}$$

برآوردگر ریج اصلاح شده در (۱.۳) را می‌توان به صورت زیر نوشت

$$\begin{aligned}
\hat{\beta}_{MRE}(k,\mathbf{b}_\circ) &= (\mathbf{X}'\mathbf{X}+k\mathbf{I})^{-1}\mathbf{X}'y+k(\mathbf{X}'\mathbf{X}+k\mathbf{I})^{-1}\mathbf{b}_\circ \\
&= (\mathbf{X}'\mathbf{X}+k\mathbf{I})^{-1}\mathbf{X}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'y+(\mathbf{I}-\mathbf{T}_k)\mathbf{b}_\circ \\
&= \mathbf{T}_k\hat{\beta}_{OLSE}+(\mathbf{I}-\mathbf{T}_k)\mathbf{b}_\circ. \qquad (۸.۳)
\end{aligned}$$

بنابراین برآوردگر ریج اصلاح شده در حقیقت ترکیب محدبی از اطلاع پیشین $\mathbf{b}_\circ$ و برآوردگر کمترین توان‌های دوم است.
حال فرض کنید

$$\begin{aligned}
\mathbf{F}_d &= (\mathbf{X}'\mathbf{X}+\mathbf{I})^{-1}(\mathbf{X}'\mathbf{X}+d\mathbf{I}) \\
&= (\mathbf{\Lambda}+\mathbf{I})^{-1}(\mathbf{\Lambda}+d\mathbf{I}) \\
&= (\mathbf{\Lambda}+\mathbf{I})^{-1}\mathbf{\Lambda}+(\mathbf{\Lambda}+\mathbf{I})^{-1}d \\
&= (\mathbf{\Lambda}+\mathbf{I})^{-1}\mathbf{\Lambda}-(1-d)(\mathbf{\Lambda}+\mathbf{I})^{-1}+(\mathbf{\Lambda}+\mathbf{I})^{-1} \\
&= (\mathbf{\Lambda}+\mathbf{I})^{-1}(\mathbf{\Lambda}+\mathbf{I})-(1-d)(\mathbf{\Lambda}+\mathbf{I})^{-1} \\
&= \mathbf{I}-(1-d)(\mathbf{\Lambda}+\mathbf{I})^{-1}. \qquad (۹.۳)
\end{aligned}$$

به طور مشابه، برآوردگر لیو اصلاح شده[۹] (MLE)، بر اساس اطلاع پیشین و برآوردگر نوع لیو را به صورت زیر ارائه می‌کنیم

$$\begin{aligned}
\hat{\beta}_{MLE}(d,\mathbf{b}_\circ) &= \mathbf{F}_d\hat{\beta}_{OLSE}+(\mathbf{I}-\mathbf{F}_d)\mathbf{b}_\circ \\
&= (\mathbf{X}'\mathbf{X}+\mathbf{I})^{-1}(\mathbf{X}'\mathbf{X}+d\mathbf{I})\hat{\beta}_{OLSE}+(1-d)(\mathbf{X}'\mathbf{X}+\mathbf{I})^{-1}\mathbf{b}_\circ \\
&= (\mathbf{X}'\mathbf{X}+\mathbf{I})^{-1}\left[(\mathbf{X}'\mathbf{X}+d\mathbf{I})\hat{\beta}_{OLSE}+(1-d)\mathbf{b}_\circ\right]. \qquad (۱۰.۳)
\end{aligned}$$

[۹] Modified liu estimation

همان‌طور که ملاحظه می‌شود به ازای مقادیر مختلف d و $\mathbf{b}_\circ$ تساوی‌های زیر برقرار می‌باشند

$$\begin{aligned}\hat{\beta}_{MLE}(۱, \mathbf{b}_\circ) &= \hat{\beta}_{OLSE} \\ \hat{\beta}_{MLE}(\circ, \mathbf{b}_\circ) &= (\mathbf{X}'\mathbf{X}+\mathbf{I})^{-۱}(\mathbf{X}'\mathbf{X}+\mathbf{b}_\circ) \\ \hat{\beta}_{MLE}(d, \circ) &= \hat{\beta}(d).\end{aligned} \tag{۱۱.۳}$$

در حقیقت، برآوردگر لیو اصلاح شده که ترکیب محدبی از برآوردگر کمترین توان‌های دوم و اطلاع پیشین است، می‌تواند به عنوان تعمیم برآوردگر کمترین توان‌های دوم و برآوردگر لیو باشد.

صورت کانونی مدل رگرسیون خطی (۱.۱) در نظر بگیرید

$$y = \mathbf{Z}\gamma + \epsilon \tag{۱۲.۳}$$

که در آن $\mathbf{Z} = \mathbf{XT}$، $\alpha = \mathbf{T}'\beta$ و $\mathbf{T}$ ماتریس متعامد است که ستون‌های آن بردارهای ویژه $\mathbf{X}'\mathbf{X}$ است. بنابراین

$$\begin{aligned}\mathbf{Z}'\mathbf{Z} &= \mathbf{T}'\mathbf{X}'\mathbf{XT} \\ &= \mathbf{\Lambda} \\ &= \operatorname{diag}(\lambda_۱, \cdots, \lambda_p)\end{aligned}$$

که در آن $\lambda_۱ \geq \lambda_۲ \geq \cdots \geq \lambda_p > \circ$، مقادیر بردارهای ویژه ماتریس $\mathbf{X}'\mathbf{X}$ می‌باشند. برای مدل خطی (۱۲.۳) برآوردگرها به صورت زیر بیان می‌شوند

$$\hat{\gamma}_{OLSE} = \mathbf{\Lambda}^{-۱}\mathbf{Z}y \tag{۱۳.۳}$$

$$\hat{\gamma}(d) = (\mathbf{\Lambda}+\mathbf{I})^{-۱}(\mathbf{\Lambda}+d\mathbf{I})^{-۱}\Lambda^{-۱}\mathbf{Z}'y \tag{۱۴.۳}$$

$$\hat{\gamma}(k) = (\mathbf{\Lambda}+k\mathbf{I})^{-۱}\mathbf{Z}'y \tag{۱۵.۳}$$

$$\hat{\gamma}_{MRE}(k, \mathbf{b}) = (\mathbf{\Lambda}+k\mathbf{I})^{-۱}(\mathbf{Z}'y+k\mathbf{b}) \tag{۱۶.۳}$$

$$\hat{\gamma}_{MLE}(d, \mathbf{b}) = (\mathbf{\Lambda}+\mathbf{I})^{-۱}\left[(\mathbf{\Lambda}+d)\mathbf{\Lambda}^{-۱}\mathbf{Z}'y+(۱-d)\mathbf{b}\right] \tag{۱۷.۳}$$

که در آن $\mathbf{b} = \mathbf{T}'\mathbf{b}_\circ$.

۳.۳ مقایسه برآوردگرها

در این بخش برآوردگر لیو اصلاح شده را با برآوردگرهای لیو، ریج و ریج اصلاح شده مقایسه می‌کنیم. در زیر بخش اول نشان می‌دهیم که برآوردگر لیو اصلاح شده بر اساس معیار $MSEM$ بهتر از برآوردگر لیو عمل می‌کند. در زیر بخش دوم برآوردگر لیو اصلاح شده را با برآوردگر ریج مقایسه می‌کنیم و در نهایت مقایسه بین برآوردگر لیو اصلاح شده و برآوردگر ریج اصلاح شده در بخش سوم صورت پذیرفته است.

۱.۳.۳ مقایسه بین برآوردگر لیو اصلاح شده و لیو

برآوردگر لیو اصلاح شده بر اساس تبدیلات کانونی به صورت زیر قابل بازنویسی است

$$\hat{\gamma}_{MLE}(d,\mathbf{b}) = (\mathbf{\Lambda}+\mathbf{I})^{-1}\left[(\mathbf{\Lambda}+d)\mathbf{\Lambda}^{-1}\mathbf{Z}'y + (1-d)\mathbf{b}\right]$$

و بر اساس رابطه (۹.۳)

$$\begin{aligned}\mathbf{I}-(\mathbf{\Lambda}+\mathbf{I})^{-1}(\mathbf{\Lambda}+d\mathbf{I})\mathbf{b} &= (\mathbf{I}-\mathbf{I}+(1-d)(\mathbf{\Lambda}+\mathbf{I})^{-1}\mathbf{b})\\ &= (1-d)(\mathbf{\Lambda}+\mathbf{I})^{-1}\mathbf{b}.\end{aligned}$$

نتیجه می‌شود

$$\begin{aligned}\hat{\gamma}_{MLE}(d,\mathbf{b}) &= (\mathbf{\Lambda}+\mathbf{I})^{-1}(\mathbf{\Lambda}+d\mathbf{I})\hat{\gamma}_{OLSE} + (\mathbf{\Lambda}+\mathbf{I})^{-1}(1-d)\mathbf{b}\\ &= \mathbf{A}_d\hat{\gamma}_{OLSE} + (\mathbf{I}-\mathbf{A}_d)\mathbf{b}\end{aligned}$$

که در آن $\mathbf{A}_d = (\mathbf{\Lambda}+\mathbf{I})^{-1}(\mathbf{\Lambda}+d\mathbf{I})$.

قضیه ۱.۳.۳. *** تحت مدل خطی ۱۲.۳ برآوردگر لیو اصلاح شده نسبت به برآوردگر لیو برتری دارد اگر و تنها اگر*

$$\gamma\gamma' - (\gamma-\boldsymbol{b})(\gamma-\boldsymbol{b})' \geqslant 0.$$

برهان. بردار اریبی و ماتریس واریانس- کوواریانس برآوردگر لیو اصلاح شده به ترتیب عبارتند از

$$\begin{aligned}\text{bias}(\hat{\gamma}_{MLE}(d,\mathbf{b})) &= \mathbf{A}_d\gamma\\ &= (\mathbf{I}-\mathbf{A}_d)\mathbf{b}-\gamma\\ &= (\mathbf{A}_d-\mathbf{I})(\gamma-\mathbf{b})\end{aligned} \tag{۱۸.۳}$$

$$\begin{aligned}\text{cov}(\hat{\gamma}_{MLE}(d,\mathbf{b})) &= \text{cov}(\mathbf{A}_d\hat{\gamma}_{OLSE} + (\mathbf{I}-\mathbf{A}_d)\mathbf{b})\\ &= \text{cov}(\mathbf{A}_d\Lambda^{-1}\mathbf{Z}'y)\\ &= \mathbf{A}_d\mathbf{\Lambda}^{-1}\mathbf{Z}'\sigma^2\mathbf{Z}\mathbf{\Lambda}^{-1}\mathbf{A}_d'\\ &= \sigma^2(\mathbf{A}_d\mathbf{\Lambda}^{-1}\mathbf{Z}'\mathbf{Z}\mathbf{\Lambda}^{-1}\mathbf{A}_d')\\ &= \sigma^2\mathbf{A}_d\mathbf{\Lambda}^{-1}\mathbf{A}_d'.\end{aligned} \tag{۱۹.۳}$$

با استفاده از روابط (۱۸.۳) و (۱۹.۳) داریم

$$\begin{aligned}MSEM(\hat{\gamma}_{MLE}(d,\mathbf{b})) &= \sigma^2\mathbf{A}_d\mathbf{\Lambda}^{-1}\mathbf{A}_d'\\ &+ (\mathbf{A}_d-\mathbf{I})(\gamma-\mathbf{b})(\gamma-\mathbf{b})'(\mathbf{A}_d-\mathbf{I}))'.\end{aligned} \tag{۲۰.۳}$$

با استفاده از رابطه (۱۳.۳) اریبی، ماتریس واریانس-کوواریانس و $MSEM$ برآوردگر لیو اصلاح شده به صورت زیر قابل محاسبه است

$$\begin{aligned} bias(\hat{\gamma}(d)) &= (\mathbf{\Lambda}+\mathbf{I})^{-1}(\mathbf{\Lambda}+d\mathbf{I})\gamma-\gamma \\ &= (\mathbf{A}_d-\mathbf{I}))\gamma \end{aligned} \tag{۲۱.۳}$$

$$\begin{aligned} cov(\hat{\gamma}(d)) &= cov(\mathbf{A}_d\hat{\gamma}_{OLSE}) \\ &= \sigma^{2}\mathbf{A}_d\mathbf{\Lambda}^{-1}\mathbf{A}'_d \end{aligned} \tag{۲۲.۳}$$

$$\begin{aligned} MSEM(\hat{\gamma}(d)) &= \sigma^{2}\mathbf{A}_d\mathbf{\Lambda}^{-1}\mathbf{A}'_d \\ &+ (\mathbf{A}_d-\mathbf{I}))\gamma\gamma'(\mathbf{A}_d-\mathbf{I})'. \end{aligned} \tag{۲۳.۳}$$

با در نظر گرفتن روابط (۲۱.۳) تا (۲۳.۳)، اختلاف رابطه (۲۰.۳) از (۲۳.۳) برابر است با

$$\begin{aligned} \mathbf{\Delta}_1 &= MSEM(\hat{\gamma}(d)) - MSEM(\hat{\gamma}_{MLE}(d,\mathbf{b})) \\ &= (\mathbf{A}_d-\mathbf{I})\gamma\gamma'(\mathbf{A}_d-\mathbf{I})' - (\mathbf{A}_d-\mathbf{I})(\gamma-\mathbf{b})(\gamma-\mathbf{b})'(\mathbf{A}_d-\mathbf{I})' \\ &= (\mathbf{A}_d-\mathbf{I})\left[\gamma\gamma'-(\gamma-\mathbf{b})(\gamma-\mathbf{b})'\right](\mathbf{A}_d-\mathbf{I})'. \end{aligned} \tag{۲۴.۳}$$

در این صورت برهان کامل است. □

۲.۳.۳ مقایسه بین برآوردگرهای لیو ، ریج و لیو اصلاح شده

قضیه ۲.۳.۳. * *(یالیان لی و هو یانگ[۱۰]، ۲۰۱۲) فرض کنید $0<k<1$، مقداری ثابت باشد.*
(الف) اگر $0<d<d_{*i}<1$، آن‌گاه

$$MSEM(\hat{\gamma}(k)) - MSEM(\hat{\gamma}_{MLE}(d,\boldsymbol{b})) > 0$$

اگر و تنها اگر

$$b'_2\left[\sigma^2(C_1C'_1-C_2C'_2)+\boldsymbol{b}_1\boldsymbol{b}'_1\right]b_2 \leqslant 1.$$

(ب) اگر $0<d_{*i}<d<1$، آن‌گاه

$$MSEM(\hat{\gamma}_{MLE}(d,\boldsymbol{b})) - MSEM(\hat{\gamma}(k)) > 0$$

اگر و تنها اگر

$$\boldsymbol{b}'_1\left[\sigma^2(\mathbf{C}_2\mathbf{C}'_2-\mathbf{C}_1\mathbf{C}'_1)+b_2b'_2\right]\boldsymbol{b}_2 \leqslant 1$$

که در آن

$$d_{*i}=\frac{\lambda_i-k\lambda_i}{k+\lambda_i},\qquad i=1,2,\cdots,p.$$

[۱۰] Yalian Li and Hu Yang

برهان. اریبی و ماتریس واریانس- کوواریانس برآوردگر ریج به ترتیب برابر

$$\begin{aligned}\text{bias}(\hat{\gamma}(k)) &= (\mathbf{\Lambda}+k\mathbf{I})^{-1}\mathbf{Z}'y-\mathbf{\Lambda}^{-1}\mathbf{Z}'y\\ &= \mathbf{T}_k\gamma-\gamma\\ &= (\mathbf{T}_k-\mathbf{I})\gamma\\ &= -k(\mathbf{\Lambda}+k\mathbf{I})^{-1}\gamma \end{aligned} \tag{۲۵.۳}$$

و

$$\begin{aligned}\text{cov}(\hat{\gamma}(k)) &= Cov((\mathbf{\Lambda}+k\mathbf{I})^{-1}\mathbf{Z}'y)\\ &= (\mathbf{\Lambda}+k\mathbf{I})^{-1}\mathbf{Z}'Cov(y)\mathbf{Z}(\mathbf{\Lambda}+k\mathbf{I})^{-1}\\ &= \sigma^{2}(\mathbf{\Lambda}+k\mathbf{I})^{-1}\mathbf{\Lambda}(\mathbf{\Lambda}+k\mathbf{I})^{-1}. \end{aligned} \tag{۲۶.۳}$$

از رابطه (۲۵.۳) و (۲۶.۳) داریم

$$MSEM(\hat{\gamma}(k)) = \sigma^{2}\mathbf{A}_k\mathbf{\Lambda}\mathbf{A}_k'+k^{2}\mathbf{A}_k\gamma\gamma'\mathbf{A}_k' \tag{۲۷.۳}$$

که در آن $\mathbf{A}_k=(\Lambda+k\mathbf{I})^{-1}$.

اختلاف $MSEM$ برآوردگر لیو اصلاح شده از برآوردگر ریج برابر است با

$$\begin{aligned}\mathbf{\Delta}_2 &= MSEM(\hat{\gamma}(k))-MSEM(\hat{\gamma}_{MLE}(d,\mathbf{b}))\\ &= \sigma^{2}(\mathbf{A}_k\mathbf{\Lambda}\mathbf{A}_k'-\mathbf{A}_d\mathbf{\Lambda}^{-1}\mathbf{A}_d')+k^{2}\mathbf{A}_k\gamma\gamma'\mathbf{A}_k'-(\mathbf{A}_d-\mathbf{I})(\gamma-\mathbf{b})(\mathbf{A}_d-\mathbf{I})'\\ &= \sigma^{2}\mathbf{D}_1+\mathbf{b}_1\mathbf{b}_1'-\mathbf{b}_2\mathbf{b}_2' \end{aligned} \tag{۲۸.۳}$$

و اختلاف $MSEM$ برآوردگر ریج از برآوردگر لیو اصلاح شده به صورت زیر قابل محاسبه است

$$\begin{aligned}\mathbf{\Delta}_3 &= MSEM(\hat{\gamma}_{MLE}(d,\mathbf{b}))-MSEM(\hat{\gamma}(k))\\ &= \sigma^{2}(\mathbf{A}_d\mathbf{\Lambda}^{-1}\mathbf{A}_d'-\mathbf{A}_k\mathbf{\Lambda}\mathbf{A}_k')+(\mathbf{A}_d-\mathbf{I})(\gamma-\mathbf{b})(\mathbf{A}_d-\mathbf{I})'-k^{2}\mathbf{A}_k\gamma\gamma'\mathbf{A}_k'\\ &= \sigma^{2}\mathbf{D}_2+\mathbf{b}_2\mathbf{b}_2'-\mathbf{b}_1\mathbf{b}_1' \end{aligned} \tag{۲۹.۳}$$

که در آن‌ها

$$\begin{cases}\mathbf{D}_1=\mathbf{C}_1\mathbf{C}_1'-\mathbf{C}_2\mathbf{C}_2'\\ \mathbf{D}_2=\mathbf{C}_2\mathbf{C}_2'-\mathbf{C}_1\mathbf{C}_1'\\ \mathbf{C}_1=\mathbf{A}_k\mathbf{\Lambda}^{1/2}\\ \mathbf{C}_2=\mathbf{A}_d\mathbf{\Lambda}^{-1/2}\end{cases}$$

و

$$\begin{cases}\mathbf{b}_1=\text{bias}(\hat{\gamma}(k))=-k\mathbf{A}_k\gamma\\ \mathbf{b}_2=\text{bias}(\hat{\gamma}_{MLE}(d,\mathbf{b}))=(\mathbf{A}_d-\mathbf{I})(\gamma-\mathbf{b}).\end{cases}$$

ماتریس $\mathbf{D}_1$ را می‌توان به صورت زیر بازنویسی کرد

$$\begin{aligned}\mathbf{D}_1 &= \mathbf{C}_1\mathbf{C}_1' - \mathbf{C}_2\mathbf{C}_2' \\ &= \mathbf{A}_k\boldsymbol{\Lambda}\mathbf{A}_k' - \mathbf{A}_d\boldsymbol{\Lambda}^{-1}\mathbf{A}_d' \\ &= \mathbf{T}\mathrm{diag}(\tau_1, \cdots, \tau_p)\mathbf{T}'\end{aligned}$$

که در آن

$$\tau_i = \frac{(\lambda_i+1)^2\lambda_i^2 - (\lambda_i+d)^2(\lambda_i+k)^2}{\lambda_i(\lambda_i+1)^2(\lambda_i+k)^2}$$

بدیهی است که اگر $1 > \frac{\lambda_i - k\lambda_i}{k+\lambda_i} > d > 0$ آن‌گاه $\tau_i > 0$، چراکه

$$\lambda_i > 0 \Rightarrow (\lambda_i+1)^2\lambda_i^2 > (\lambda_i+d)^2(\lambda_i+k)^2$$

$$\longrightarrow (\lambda_i+1)\lambda_i > (\lambda_i+d)(\lambda_i+k) \Longrightarrow \frac{\lambda_i - \lambda_i k}{\lambda_i + k} > d > 0$$

که نتیجه می‌دهد $\mathbf{D}_1 > 0$.

به طور مشابه اگر $1 > d > \frac{\lambda_i - k\lambda_i}{k+\lambda_i} > 0$، آن‌گاه $\tau_i > 0$، که نتیجه می‌دهد $\mathbf{D}_2 > 0$.

برای مقدار ثابت d و $\frac{\lambda_i - d\lambda_i}{d+\lambda_i} > k > 0$، نتیجه می‌گیریم $\tau_i > 0$، و $\mathbf{D}_1 > 0$.

برای $1 > k > \frac{\lambda_i - d\lambda_i}{d+\lambda_i} > 0$، نتیجه می‌گیریم $\mathbf{D}_2 > 0$.

با استفاده از لم ب.۰.۳۳ برهان کامل است. □

قضیه ۳.۳.۳. * *(یالیان لی و هو یانگ، ۲۰۱۲) فرض کنید $1 > d > 0$، مقداری ثابت باشد.*

(الف) اگر $1 > k_{*i} > k > 0$، آن‌گاه

$$MSEM(\hat{\gamma}(k)) - MSEM(\hat{\gamma}_{MLE}(d, \boldsymbol{b})) > 0$$

اگر و تنها اگر

$$\boldsymbol{b}_2'\left[\sigma^2(\mathbf{C}_1\mathbf{C}_1' - \mathbf{C}_2\mathbf{C}_2') + \boldsymbol{b}_1\boldsymbol{b}_1'\right]\boldsymbol{b}_2 \leqslant 1.$$

(ب) اگر $1 > k > k_{*i} > 0$، آن‌گاه

$$MSEM(\hat{\gamma}_{MLE}(d, \boldsymbol{b})) - MSEM(\hat{\gamma}(k)) > 0$$

اگر و تنها اگر

$$\boldsymbol{b}_1'\left[\sigma^2(\mathbf{C}_2\mathbf{C}_2' - \mathbf{C}_1\mathbf{C}_1') + \boldsymbol{b}_2\boldsymbol{b}_2'\right]\boldsymbol{b}_2 \leqslant 1$$

که در آن

$$k_{*i} = \frac{\lambda_i - d\lambda_i}{d+\lambda_i}, \quad i = 1, 2, \cdots, p.$$

برهان. با استفاده از لم ب.۰.۳۳ و روابط قضیه ۲.۳.۳ برهان واضح است. □

قضیه ۴.۳.۳. * *(یالیان لی و هو یانگ، ۲۰۱۲) فرض کنید $\circ < d < ۱$، مقداری ثابت باشد. برآوردگر لیو اصلاح شده میانگین توان‌های دوم کمتری نسبت به برآوردگر توان‌های دوم معمولی دارد؛ به عبارتی*

$$MSEM(\hat{\gamma}_{OLSE}) - MSEM(\hat{\gamma}_{MLE}(d,\boldsymbol{b})) = \sigma^۲\mathbf{D}_۳ - b_۲b'_۲ \geqslant \circ$$

اگر و تنها اگر $\boldsymbol{b}'_۲\mathbf{D}_۳^{-۱}\boldsymbol{b}_۲ \leqslant \sigma^۲$.

برهان. از آن‌جایی که به ازای $k = \circ$ برآوردگر ریج به برآوردگر کمترین توان‌های دوم ساده می‌شود، بنابراین

$$\begin{aligned}\Delta_۴ &= MSEM(\hat{\gamma}_{OLSE}) - MSEM(\hat{\gamma}_{MLE}(d,\mathbf{b}))\\ &= \sigma^۲ D_۳ - \mathbf{b}_۲\mathbf{b}'_۲\end{aligned}$$

که در آن

$$\begin{aligned}\mathbf{D}_۳ &= \mathbf{\Lambda}^{-۱} - \mathbf{A}_d\mathbf{\Lambda}^{-۱}\mathbf{A}'_d\\ &= \mathbf{T}\mathrm{diag}(\gamma_۱,\cdots,\gamma_p)\mathbf{T}'\end{aligned}$$

و

$$\gamma_i = \frac{(\lambda_i+۱)^۲-(\lambda_i+d)^۲}{\lambda_i(\lambda_i+۱)^۲}.$$

به ازای $\circ < d < ۱$، $\gamma_i > \circ$ که نتیجه می‌گیریم $\mathbf{D}_۳ > \circ$. لذا با استفاده از لم ب.۳۲.۰ برهان کامل است. □

۳.۳.۳ مقایسه بین برآوردگرهای ریج اصلاح شده و لیو اصلاح شده

برآوردگر ریج اصلاح شده بر پایه اطلاع پیشین را در نظر بگیرید. سوئیندل (۱۹۷۶) این برآوردگر اریب را در مواقعی که همخطی وجود دارد پیشنهاد نمود و نشان داد که این برآوردگر عملکرد بهتری نسبت به برآوردگر ریج دارد. در این زیر بخش برآوردگر ریج اصلاح شده را مورد بررسی قرار می‌دهیم و با استفاده از قضیه‌هایی که بیان و اثبات می‌کنیم، این برآوردگر را با برآوردگر لیو اصلاح شده مقایسه می‌نماییم.
برآوردگر ریج اصلاح شده به صورت زیر قابل بازنویسی است

$$\begin{aligned}\hat{\gamma}_{MRE}(k,\mathbf{b}) &= (\mathbf{\Lambda}+k\mathbf{I})^{-۱}(\mathbf{Z}'y+k\mathbf{b})\\ &= (\mathbf{\Lambda}+k\mathbf{I})^{-۱}\mathbf{Z}'y+k(\mathbf{\Lambda}+k\mathbf{I})^{-۱}\mathbf{b}\\ &= (\mathbf{\Lambda}+k\mathbf{I})^{-۱}\mathbf{\Lambda}\gamma+k(\mathbf{\Lambda}+k\mathbf{I})^{-۱}\mathbf{b}\\ &= \mathbf{T}_k\gamma+(\mathbf{I}-\mathbf{T}_k)\mathbf{b} \qquad (۳۰.۳)\end{aligned}$$

که در آن $\tilde{\mathbf{A}}_k = (\mathbf{\Lambda}+k\mathbf{I})^{-۱}\mathbf{\Lambda}$.

اریبی و ماتریس واریانس- کوواریانس برآوردگر ریج اصلاح شده در رابطه (۳۰.۳) به صورت زیر

می‌باشند

$$\text{bias}(\hat{\gamma}_{MRE}(k,\mathbf{b})) = \tilde{\mathbf{A}}_k\gamma + (\mathbf{I}-\tilde{\mathbf{A}}_k)\mathbf{b}-\gamma = (\tilde{\mathbf{A}}_k-\mathbf{I})(\gamma-\mathbf{b}) \tag{۳۱.۳}$$

$$\text{cov}(\hat{\gamma}_{MRE}(k,\mathbf{b})) = \sigma^{۲}\tilde{\mathbf{A}}_k\mathbf{\Lambda}^{-۱}\tilde{\mathbf{A}}'_k. \tag{۳۲.۳}$$

با استفاده از روابط (۳۱.۳) و (۳۲.۳) نتیجه می‌شود

$$MSEM(\hat{\gamma}_{MRE}(k,\mathbf{b})) = \sigma^{۲}\tilde{\mathbf{A}}_k\mathbf{\Lambda}^{-۱}\tilde{\mathbf{A}}'_k + (\tilde{\mathbf{A}}_k-\mathbf{I})(\gamma-\mathbf{b})(\gamma-\mathbf{b})'(\tilde{\mathbf{A}}_k-\mathbf{I})'. \tag{۳۳.۳}$$

با استفاده از قضیه‌های زیر برآوردگر لیو اصلاح شده را با برآوردگر ریج اصلاح شده مقایسه می‌کنیم.

قضیه ۵.۳.۳. * *(یالیان لی و هو یانگ، ۲۰۱۲) فرض کنید $۰ < k < ۱$، مقداری ثابت باشد.*
(الف) اگر $۰ < d < d_{*i} < ۱$، آن‌گاه برآوردگر لیو اصلاح شده نسبت به برآوردگر ریج اصلاح شده برتری دارد؛ اگر و تنها اگر

$$\boldsymbol{b}'_۲\left(\sigma^{۲}\mathbf{D}_۱+\boldsymbol{b}_۳\boldsymbol{b}'_۳\right)^{-۱}\boldsymbol{b}_۲ \leqslant ۱.$$

(ب) اگر $۰ < d_{*i} < d < ۱$، آن‌گاه برآوردگر ریج اصلاح شده نسبت به برآوردگر لیو اصلاح شده برتری دارد؛ اگر و تنها اگر

$$\boldsymbol{b}'_۳\left(\sigma^{۲}\mathbf{D}_۲+\boldsymbol{b}_۲\boldsymbol{b}'_۲\right)^{-۱}\boldsymbol{b}_۳ \leqslant ۱$$

که در آن

$$d_{*i} = \frac{\lambda_i - k\lambda_i}{k+\lambda_i}, \quad i = ۱,۲,\cdots,p.$$

برهان. با استفاده از رابطه‌های (۳.۳) و (۳۳.۳)، اختلاف مخاطره برآوردگر لیو اصلاح شده از مخاطره برآوردگر ریج اصلاح شده به صورت زیر است

$$\begin{aligned}\mathbf{\Delta}_۵ &= MSEM(\hat{\gamma}_{MRE}(k,\mathbf{b})) - MSEM(\hat{\gamma}_{MLE}(d,\mathbf{b}))\\ &= \sigma^{۲}(\tilde{\mathbf{A}}_k\mathbf{\Lambda}^{-۱}\tilde{\mathbf{A}}'_k - \mathbf{A}_d\mathbf{\Lambda}^{-۱}\mathbf{A}'_d)\\ &+ (\tilde{\mathbf{A}}_k-\mathbf{I})(\gamma-\mathbf{b})(\gamma-\mathbf{b})'(\tilde{\mathbf{A}}_k-\mathbf{I})'\\ &- (\mathbf{A}_d-\mathbf{I})(\gamma-\mathbf{b})(\gamma-\mathbf{b})'(\mathbf{A}_d-\mathbf{I})'\end{aligned} \tag{۳۴.۳}$$

و همچنین اختلاف مخاطره برآوردگر ریج اصلاح شده از مخاطره برآوردگر لیو اصلاح شده به صورت زیر قابل محاسبه می‌باشد

$$\begin{aligned}\mathbf{\Delta}_۶ &= MSEM(\hat{\gamma}_{MLE}(d,\mathbf{b})) - MSEM(\hat{\gamma}_{MRE}(k,\mathbf{b}))\\ &= \sigma^{۲}\mathbf{D}_۲ + \mathbf{b}_۲\mathbf{b}'_۲ - \mathbf{b}_۳\mathbf{b}'_۳\end{aligned} \tag{۳۵.۳}$$

که در آن

$$\mathbf{b}_3 = \text{bias}(\hat{\gamma}_{MRE}) = (\mathbf{T}_k - \mathbf{I})(\gamma - \mathbf{b}).$$

با استفاده از قضیه ۲.۳.۳ اثبات کامل است. □

قضیه ۶.۳.۳. * *(یالیان لی و هو یانگ، ۲۰۱۲) فرض کنید $0 < d < 1$، مقداری ثابت باشد*
(الف) اگر $0 < k < k_{*i} < 1$، آن‌گاه برآوردگر لیو اصلاح شده نسبت به برآوردگر ریج اصلاح شده برتری دارد؛ اگر و تنها اگر

$$\boldsymbol{b}_2' \left(\sigma^2 \mathbf{D}_1 + \boldsymbol{b}_3 \boldsymbol{b}_3'\right)^{-1} \boldsymbol{b}_2 \leqslant 1.$$

(ب) اگر $0 < k_{*i} < k < 1$، آن‌گاه برآوردگر ریج اصلاح شده نسبت به برآوردگر لیو اصلاح شده برتری دارد؛ اگر و تنها اگر

$$\boldsymbol{b}_3' \left(\sigma^2 \mathbf{D}_2 + \boldsymbol{b}_2 \boldsymbol{b}_2'\right)^{-1} \boldsymbol{b}_3 \leqslant 1$$

که در آن

$$k_{*i} = \frac{\lambda_i - d\lambda_i}{d + \lambda_i}, \quad i = 1, 2, \cdots, p.$$

برهان. اثبات همانند قضیه ۵.۳.۳ است. □

۴.۳ مثال عددی

در صنعت نفت برای سیمان در قسمت حفاری چاه نفت کاربردهای زیادی وجود دارد، مثلاً بعد از این‌که چاه نفت حفاری شد، برای این‌که کیفیت چاه حفظ شود، لوله‌های جداری داخل چاه می‌گذاریم و برای این‌که بین سنگ و لوله جداری فاصله‌ای نداشته باشیم، از سیمان استفاده می‌کنیم. به طور کلی، سیمان‌ها از چهار بخش اصلی: سیلیکات تری کلسیم، سیلیکات دی کلسیم، آلومینت تری کلسیم و آلومینو فرایت کلسیم تشکیل می‌شوند. تغییرات نسبت استفاده شده هر قسمت در میزان استحکام و زمان سخت شدن سیمان تأثیر گذار خواهد بود. مثلاً هرچه سیلیکات تری کلسیم بیشتر شود سیمان زودتر سخت می‌شود. یا میزان بیشتر آهن در سیمان باعث می‌شود، آب دهی کمتری داشته باشد. داده‌های زیر مربوط به سیمان پورتلند است که شامل ۱۳ مشاهده و ۵ متغیر می‌باشد. (وود[۱۱] و همکاران، ۱۹۳۲).
متغیر X_1 درصد آلومینت تری کلسیم در نمونه را نشان می‌دهد. درصد سیلیکات تری کلسیم در نمونه را با X_2 مشخص می‌کنیم. X_3 درصد آلمینو فرایت کلسیم و X_4 درصد سیلیکات دی کلسیم در نمونه را نشان می‌دهد. و در نهایت متغیر پاسخ Y، میزان گرمای حاصل از این واکنش بر حسب کالری بر گرم است.

[11] Woods

$$\mathbf{X}=\begin{pmatrix} ۱ & ۷ & ۲۶ & ۶ & ۶۰ \\ ۱ & ۱ & ۲۹ & ۱۵ & ۵۲ \\ ۱ & ۱۱ & ۵۶ & ۸ & ۲۰ \\ ۱ & ۱۱ & ۳۱ & ۸ & ۴۷ \\ ۱ & ۷ & ۵۲ & ۶ & ۳۳ \\ ۱ & ۱۱ & ۵۵ & ۹ & ۲۲ \\ ۱ & ۳ & ۷۱ & ۱۷ & ۶ \\ ۱ & ۱ & ۳۱ & ۲۲ & ۴۴ \\ ۱ & ۲ & ۵۴ & ۱۸ & ۲۲ \\ ۱ & ۲۱ & ۴۷ & ۴ & ۲۶ \\ ۱ & ۱ & ۴۰ & ۲۳ & ۳۴ \\ ۱ & ۱۱ & ۶۶ & ۹ & ۱۲ \\ ۱ & ۱۰ & ۶۸ & ۸ & ۱۲ \end{pmatrix}, \quad Y=\begin{pmatrix} ۷۸/۵ \\ ۷۴/۳ \\ ۱۰۴/۰۳ \\ ۸۷/۶ \\ ۹۵/۹ \\ ۱۰۹/۲ \\ ۱۰۲/۷ \\ ۷۲/۵ \\ ۹۳/۱ \\ ۱۱۵/۹ \\ ۸۳/۸ \\ ۱۱۳/۳ \\ ۱۰۹/۴ \end{pmatrix}$$

مقادیر ویژه برای این مثال به صورت زیر است

$$\lambda_۱=۴۴۶۷۶/۲۰۶,\ \lambda_۲=۵۹۶۵/۴۲۲,\ \lambda_۳=۸۰۹/۹۵۲,\ \lambda_۴=۱۰۵/۴۱۹,\ \lambda_۵=۰/۰۰۱۲۲.$$

برنامه مربوط به این مثال در پیوست ت.۶ آمده است. مقدار آماره کاپا برابر ۶۰۵۶/۳۴۴ است که دلالت بر همخطی شدید می‌باشد.

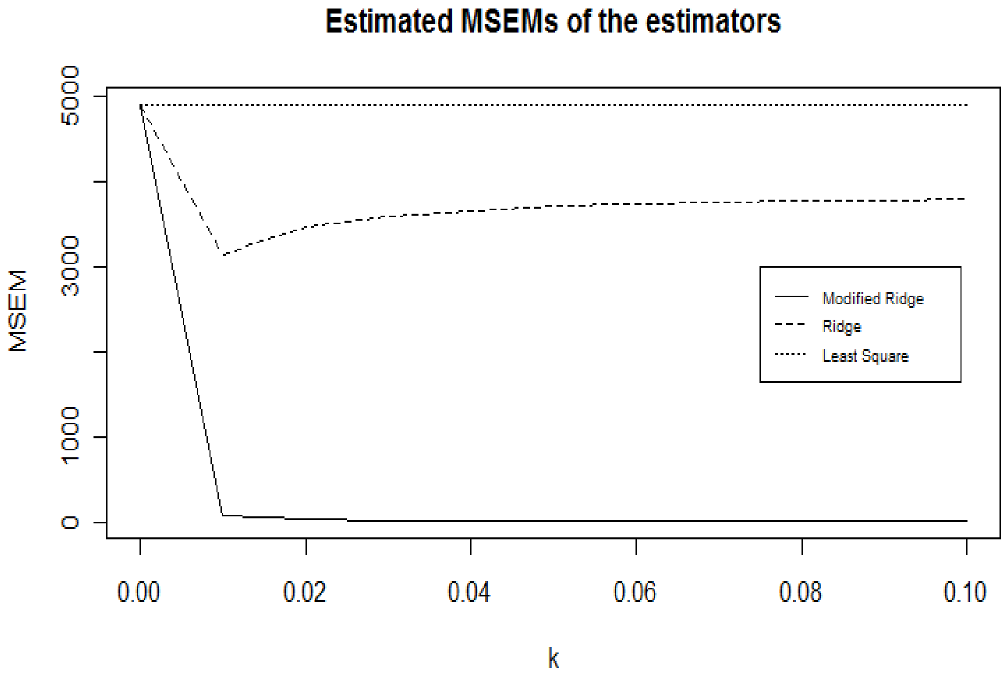
مقادیر میانگین توان‌های دوم خطا را برای برآوردگرهای لیو اصلاح شده، ریج اصلاح شده، ریج و لیو

جدول ۱.۳: برآورد MSE برای LSE، MLE، LE، RE و MRE

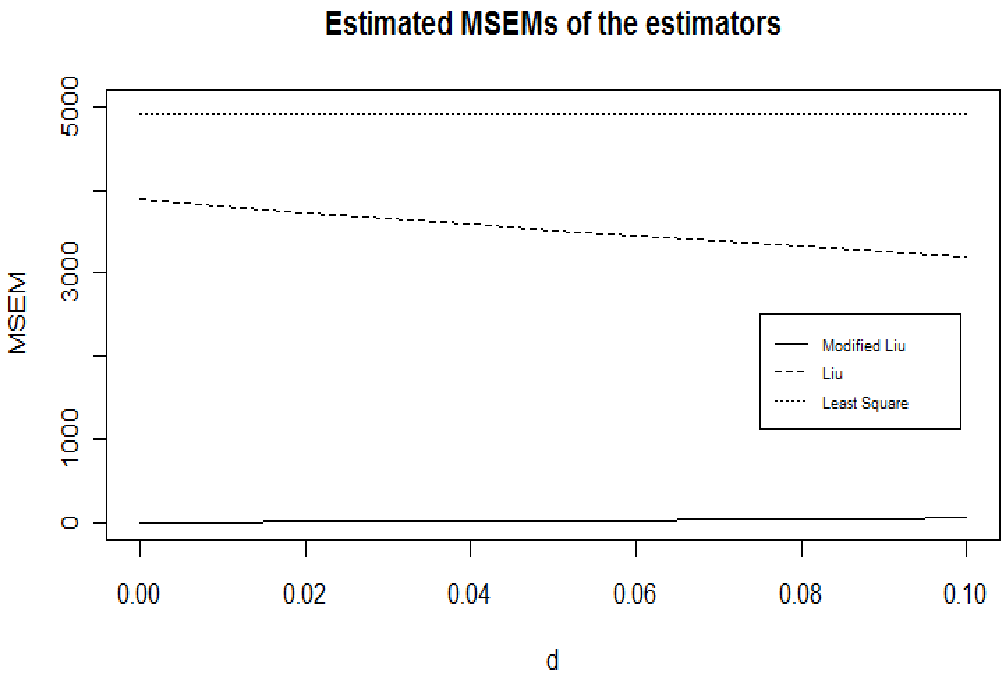
	k=۰	k=۰/۰۱	k=۰/۰۲	k=۰/۰۳	k=۰/۰۴	k=۰/۰۵	k=۰/۰۹	k=۰/۱
LSE	۴۹۱۲/۱	۴۹۱۲/۱	۴۹۱۲/۱	۴۹۱۲/۱	۴۹۱۲/۱	۴۹۱۲/۱	۴۹۱۲/۱	۴۹۱۲/۱
RE	۴۹۱۲/۱	۳۱۴۹/۲	۳۴۷۲/۶	۳۶۰۰/۱	۳۶۶۸/۰	۳۷۱۰/۲	۳۷۸۷/۹	۳۷۹۷/۹
MRE	۴۹۱۲/۱	۶۵/۷	۲۴/۹	۱۶/۵	۱۳/۵	۱۲/۱	۱۰/۴	۱۰/۳
	d=۰	d=۰/۰۱	d=۰/۰۲	d=۰/۰۳	d=۰/۰۴	d=۰/۰۵	d=۰/۰۹	d=۰/۱
LSE	۴۹۱۲/۱	۴۹۱۲/۱	۴۹۱۲/۱	۴۹۱۲/۱	۴۹۱۲/۱	۴۹۱۲/۱	۴۹۱۲/۱	۴۹۱۲/۱
LE	۳۸۸۰/۸	۳۸۰۴/۲	۳۷۲۹/۳	۳۶۵۶/۲	۳۵۸۴/۹	۳۵۱۵/۳	۳۲۵۴/۵	۳۱۹۳/۶
MLE	۹/۷۷۳۳	۱۰/۱۹	۱۱/۵۸۸	۱۳/۹۶۸	۱۷/۳۳۰	۲۱/۶۷۴	۴۸/۸۷۱	۵۸/۱۲۵

و کمترین توان‌های دوم، در جدول ۱.۳ آمده است، در تمام حالت‌ها برآوردگرهای لیو اصلاح شده، ریج اصلاح شده، ریج و نوع لیو عملکرد بهتری نسبت به برآوردگر کمترین توان‌های دوم دارند. با انتخاب مناسب اطلاع پیشین برآوردگر لیو اصلاح شده برتری قابل توجهی نسبت به برآوردگرهای کمترین توان‌های دوم، ریج، نوع لیو و لیو اصلاح شده دارد.

آن‌چه در شکل ۱.۳ قابل توجه می‌باشد، این است که، مقدار برآورد MSE برای برآوردگرهای ریج و



شکل ۱.۳: برآورد MSE برای برآوردگرهای RE و MRE



شکل ۲.۳: برآورد MSE برای برآوردگرهای LE و MLE

ریج اصلاح شده با افزایش k، کاهش می‌یابد. اما، کاهش برای برآوردگر ریج اصلاح شده چشم‌گیرتر از برآوردگر ریج است. به این مفهوم که در این مثال، برآوردگر ریج اصلاح شده عملکرد بهتری نسبت به برآوردگر ریج داشته است. این در حالی است که دو برآوردگر ریج و ریج اصلاح شده مقدار برآورد MSE کمتری نسبت به برآوردگر کمترین توان‌های دوم دارند.
شکل ۲.۳ نشان می‌دهد که برآورد MSEبرای برآوردگر لیو و برآوردگر لیو اصلاح شده به آرامی تغییر می‌کند، با این تفاوت که برای برآوردگر لیو مقدار این برآورد به آرامی کاهش و برای برآودگر لیو اصلاح شده، به آرامی افزایش می‌یابد، با این وجود، در نهایت مقدار این برآورد برای برآوردگر لیو اصلاح شده خیلی کم‌تر از برآوردگر لیو می‌باشد. و این در حالی است که مقدار برآورد MSE برای هر دو برآوردگر کمتر از مقدار برآورد MSE برای برآوردگر کمترین توان‌های دوم است.

۵.۳ مطالعه شبیه‌سازی

در این بخش برای مقایسه رفتار برآوردگرهای ارائه شده، روش شبیه‌سازی مونت کارلو را به‌کار می‌گیریم. برای داشتن داده‌هایی با میزان همخطی مختلف، لیو (۲۰۰۳) پیشنهاد کرد که داده‌ها به روش زیر تولید شوند

$$x_{ij} = (۱ - \gamma^۲)^{۱/۲}\omega_{ij} + \gamma\omega_{ip}, \quad i = ۱, ۲, \cdots, n, \; j = ۱, ۲, \cdots, p - ۱$$

که در آن اعداد ω_{ij} تصادفی مستقل از توزیع نرمال استاندارد و $\gamma^۲$ همبستگی بین دو متغیر پیش‌بین است.
این متغیرها استاندارد شده‌اند، به طوری‌که $\mathbf{X'X}$ و $\mathbf{X'}y$ همبسته هستند. حال با استفاده از رابطه زیر، n مشاهده برای متغیر پاسخ مشخص می‌شود

$$y_i = \beta_۱ x_۱ + \beta_۲ x_۲ + \beta_۳ x_۳ + e_i$$

که در آن e_i اعداد تصافی مستقل از توزیع نرمال با میانگین صفر و واریانس $\sigma^۲$ هستند.
این آزمایش را با مقادیر مختلف γ : ۰٫۷, ۰٫۹۰, ۰٫۹۹ و σ : ۰٫۰۱, ۰٫۱, ۱ برای ۱۰۰ بار تکرار می‌کنیم. مقدار میانگین توان دوم خطا را برای ۵ برآوردگر کمترین توان‌های دوم، برآوردگر لیو، برآوردگر ریج، برآوردگر ریج اصلاح شده، برآوردگر لیو اصلاح شده با استفاده از رابطه‌ی زیر برآورد می‌کنیم

$$MSE(\tilde{\beta}) = \frac{۱}{۱۰۰}\sum_{j=۱}^{۱۰۰}\sum_{i=۱}^{۳}(\tilde{\beta}_{ij} - \beta_i)^۲$$

که در آن $\tilde{\beta}_{ij}$ برآورد i–امین پارامتر در j–امین تکرار و β_i مقدار پارامتر است. برنامه مربوط به این مثال در پیوست ت.۷ آمده است.

نتایج شبیه‌سازی در جداول ۲.۳، ۳.۳ و ۴.۳ آمده است. مقادیر برآورد میانگین توان‌های دوم خطا در بیشتر حالت‌ها برای برآوردگر لیو، برآوردگر ریج، برآوردگر ریج اصلاح شده، برآوردگر لیو اصلاح شده نسبت به برآوردگر کمترین توان‌های دوم کمتر است. همچنین برآوردگر لیو اصلاح شده نسبت به برآوردگر لیو و برآوردگر ریج اصلاح شده نسبت به برآوردگر ریج، برآورد میانگین توان‌های دوم خطای کمتری دارد.

جدول ۲.۳: برآورد MSE برای LSE ، MLE ، LE ، RE و MRE به ازای $\gamma = ۰/۷۰$

$\sigma = ۱$			$\sigma = ۰/۱$			$\sigma = ۰/۰۱$			
k=۰/۱	k=۰/۰۵	k=۰	k=۰/۱	k=۰/۰۵	k=۰	k=۰/۱	k=۰/۰۵	k=۰	
۰/۰۹۳۵۹	۰/۰۹۳۵۹	۰/۰۹۳۵۹	۰/۸۹۹۳	۰/۸۹۹۳	۰/۸۹۹۳	۹/۲۳۴۶	۹/۲۳۴۶	۹/۲۳۴۶	LSE
۰/۰۹۳۰۵	۰/۰۹۳۴۳	۰/۰۹۳۵۹	۰/۸۵۹۷	۰/۸۷۸۵	۰/۸۹۹۳	۷/۵۷۱۷	۷/۵۸۸۰۶	۹/۲۳۴۶	RE
۰/۰۹۳۰۵	۰/۰۹۳۴۳	۰/۰۹۳۵۹	۰/۸۵۹۷	۰/۸۷۸۵	۰/۸۹۹۳	۷/۵۳۱۸	۷/۵۸۸۰۶	۹/۲۳۴۶	MRE
d=۱	d=۰/۰۵	d=۰	d=۱	d=۰/۰۵	d=۰	d=۱	d=۰/۰۵	d=۰	
۰/۰۹۳۵۹	۰/۰۹۳۵۹	۰/۰۹۳۵۹	۰/۸۹۹۳	۰/۸۹۹۳	۰/۸۹۹۳	۹/۲۳۴۶	۹/۲۳۴۶	۹/۲۳۴۶	LSE
۰/۰۸۷۲	۰/۰۷۸۱	۰/۰۸۲۸	۰/۸۷۹۸	۰/۸۱۶۱	۰/۸۴۱۲	۰/۸۶۹۳	۰/۹۱۹۵	۰/۸۳۶۴	LE
۰/۰۸۷۰	۰/۰۸۷۱	۰/۰۸۲۸	۰/۸۷۷۹	۰/۸۱۵۹	۰/۸۴۱۲	۰/۸۶۳۸	۰/۹۱۷۲	۰/۸۲۲۰	MLE

جدول ۳.۳: برآوردMSE برای LSE ، MLE ، LE، RE و MRE به ازای $\gamma = ۰/۹۰$

$\sigma = ۱$			$\sigma = ۰/۱$			$\sigma = ۰/۰۱$			
k=۰/۱	k=۰/۰۵	k=۰	k=۰/۱	k=۰/۰۵	k=۰	k=۰/۱	k=۰/۰۵	k=۰	
۲/۵۹۷۳	۲/۵۹۷۳	۲/۵۹۷۳	۶/۳۲۵۷	۶/۳۲۵۷	۶/۳۲۵۷	۲۸/۶۶۱۲	۲۸/۶۶۱۲	۲۸/۶۶۱۲	LSE
۰/۷۱۲۳	۰/۷۲۳۳	۲/۵۹۷۳	۰/۷۳۱۱	۰/۷۳۲۱	۶/۳۲۵۷	۲/۷۵۵۶	۲/۷۱۶۴	۲۸/۶۶۱۲	RE
۰/۷۱۰۲	۰/۷۱۵۳	۲/۵۹۷۳	۰/۷۳۱۱	۰/۷۳۲۱	۶/۳۲۵۷	۲/۳۲۲۰	۲/۶۱۵۵	۲۸/۶۶۱۲	MRE
d=۱	d=۰/۰۵	d=۰	d=۱	d=۰/۰۵	d=۰	d=۱	d=۰/۰۵	d=۰	
۲/۵۹۷۳	۲/۵۹۷۳	۲/۵۹۷۳	۶/۳۲۵۷	۶/۳۲۵۷	۶/۳۲۵۷	۸/۶۶۱۲/	۲۸/۶۶۱۲	۲۸/۶۶۱۲	LSE
۲/۵۵۳۰	۲/۰۱۶۶	۱/۸۱۴۳	۲/۵۵۲۰	۲/۰۸۹۱	۱/۸۹۶۳	۲/۷۹۴۳	۲/۱۹۷۸	۲/۰۵۲۲	LE
۲/۵۵۳۰	۲/۰۱۶۶	۱/۸۱۳۱	۲/۵۰۲۰	۲/۰۸۹۱	۱/۸۹۶۳	۲/۷۹۴۳	۲/۹۷۸	۲/۰۴۹۸	MLE

جدول ۴.۳: برآورد MSE برای LSE ، MLE ، LE ، RE و MRE به ازای $\gamma = ۰/۹۹$

$\sigma = ۱$			$\sigma = ۰/۱$			$\sigma = ۰/۰۱$			
k=۰/۱	k=۰/۰۵	k=۰	k=۰/۱	k=۰/۰۵	k=۰	k=۰/۱	k=۰/۰۵	k=۰	
۲/۶۳۸۰	۲/۶۳۸۰	۲/۶۳۸۰	۲۷/۸۸۱۲	۲۷/۸۸۲۱	۲۷/۸۸۲۱	۴۳/۵۳۲۲	۴۳/۵۳۲۲	۴۳/۵۳۲۲	LSE
۰/۸۹۷۱	۰/۹۵۶۷	۲/۶۳۸۰	۰/۳۴۲۱	۰/۸۹۳۴	۲۷/۸۸۲۱	۰/۷۵۴۱	۰/۷۳۸۹	۴۳/۵۳۲۲	RE
۰/۸۶۶۳	۰/۹۰۸۷	۲/۶۳۸۰	۰/۳۲۷۸	۰/۸۷۲۸	۲۷/۸۸۲۱	۰/۱۹۵۶	۰/۴۷۸۲	۴۳/۵۳۲۲	MRE
d=۱	d=۰/۰۵	d=۰	d=۱	d=۰/۰۵	d=۰	d=۱	d=۰/۰۵	d=۰	
۲/۶۳۸۰	۲/۶۳۸۰	۲/۶۳۸۰	۲۷/۸۸۱۲	۲۷/۸۸۲۱	۲۷/۸۸۲۱	۴۳/۵۳۲۲	۴۳/۵۳۲۲	۴۳/۵۳۲۲	LSE
۰.۰۷۲۵	۰/۰۸۵۹	۰/۰۳۷۲	۰/۰۶۰۲	۰/۰۷۱۵	۰/۰۳۷۹	۰/۶۹۶۳۱	۰/۲۹۸۵	۰/۶۶۷۷	LE
۰/۰۷۲۵	۰/۰۸۵۸	۰/۰۳۱۲	۰/۰۶۰۲	۰/۰۷۱۵	۰/۰۲۶۳	۰/۶۹۳۶۸	۰/۲۹۸۵	۰/۶۶۷۲	MLE

فصل ۴

برآوردگر جک‌نایف اصلاح شده لی و یانگ

۱.۴ مقدمه

برآوردگر لیو اصلاح شده بر اساس اطلاع پیشین ($\beta = b_\circ$)، توسط لی و یانگ در سال ۲۰۱۲ معرفی شد. این برآوردگر را می‌توان تعمیمی از برآوردگر کمترین توان‌های دوم و برآوردگر لیو دانست. این برآوردگر به طور کامل در بخش ۲.۳ معرفی شده است. لی و یانگ (۲۰۱۲) سعی داشته اند که نشان دهند زمانی که همخطی وجود دارد این برآوردگر عملکرد بهتری نسبت به برآوردگر کمترین توان‌های دوم و برآوردگر لیو دارد، که نتایج تلاش آن‌ها در بخش ۳.۳ آمده است.

در این فصل تلاش شده است با استفاده از روش جک‌نایف آکدنیز و آکدنیز دوران (۲۰۱۲)، به برآوردگری از نوع لیو اصلاح شده دست یابیم که اریبی را کاهش دهد. در این راستا، در بخش اول این فصل، با استفاده از این تکنیک به دنبال معرفی برآوردگری جدید هستیم. در بخش دوم اریبی و ماتریس واریانس- کوواریانس را برای برآوردگر پیشنهادی محاسبه نمودیم و سعی داریم، نشان دهیم که در مواقعی که همخطی وجود دارد برآوردگر معرفی شده عملکرد بهتری نسبت به برآوردگر کمترین توان‌های دوم، برآوردگر لیو و برآوردگر لیو اصلاح شده دارد. به همین منظور در بخش سوم، با استفاده از روش شبیه‌سازی مونت کارلویی و روش جک‌نایف برآوردگر جدید را با دیگر برآوردگرها مقایسه نمودیم و صحت نتایج را مورد بررسی قرار داده‌ایم. مطالب این فصل کاملاً جدید می‌باشد و از محتوای آن، مقاله برزوئی بیدگلی و آرشی (۱۳۹۳) استخراج شده است.

۲.۴ معرفی برآوردگر جک‌نایف لیو اصلاح شده

مدل رگرسیون خطی (۱۲.۳) را در نظر بگیرید. برای این مدل برآوردگر لیو تعمیم یافته را می‌توان به صورت زیر نوشت

$$\begin{aligned}\hat{\gamma}(d) &= (\mathbf{Z}'\mathbf{Z}+\mathbf{I})^{-۱}(\mathbf{Z}'\mathbf{Z}+d\mathbf{I})(\mathbf{Z}'\mathbf{Z})^{-۱}\mathbf{Z}'y\\ &= (\mathbf{\Lambda}+\mathbf{I})^{-۱}(\mathbf{\Lambda}+d\mathbf{I})\mathbf{\Lambda}^{-۱}\mathbf{Z}'y\\ &= \mathbf{A}^{-۱}\mathbf{Z}'y \qquad (۱.۴)\end{aligned}$$

که در آن $\mathbf{A}^{-۱} = (\mathbf{\Lambda}+\mathbf{I})^{-۱}(\mathbf{\Lambda}+d\mathbf{I})\mathbf{\Lambda}^{-۱}$.

همچنین برآوردگر لیو اصلاح شده با استفاده از رابطه (۱۳.۳) به صورت زیر است

$$\begin{aligned}\hat{\gamma}_{MLE}(d,\mathbf{b}) &= (\mathbf{Z}'\mathbf{Z}+\mathbf{I})^{-۱}\left[(\mathbf{Z}'\mathbf{Z}+d\mathbf{I})(\mathbf{Z}'\mathbf{Z})^{-۱}\mathbf{Z}'y+(۱-d)\mathbf{b}\right]\\ &= (\mathbf{\Lambda}+\mathbf{I})^{-۱}\left[(\mathbf{\Lambda}+d\mathbf{I})\mathbf{\Lambda}^{-۱}\mathbf{Z}'y+(۱-d)\mathbf{b}\right]\\ &= \mathbf{A}^{-۱}\mathbf{X}'y+(\mathbf{\Lambda}+\mathbf{I})^{-۱}(۱-d)\mathbf{b}\\ &= \hat{\gamma}(\mathbf{d})+\mathbf{B}\mathbf{b} \qquad (۲.۴)\end{aligned}$$

که در آن ماتریس $\mathbf{A}$ و $\mathbf{B}$ به صورت زیر می‌باشند

$$\begin{cases}\mathbf{A}^{-۱} = (\mathbf{\Lambda}+\mathbf{I})^{-۱}(\mathbf{\Lambda}+d\mathbf{I})\mathbf{\Lambda}^{-۱}\\ \mathbf{B} = (۱-d)(\mathbf{\Lambda}+\mathbf{I})^{-۱}.\end{cases}$$

با استفاده از تکنیک جک‌نایف آکدنیز و آکدنیز دوران (۲۰۱۲) که در بخش ۲.۲ آمده است، متناظر با هر نمونه جک‌نایف پاسخ جک‌نایف $\hat{\gamma}_{MLE}(d,b)$ به صورت زیر تعریف می‌شود

$$\begin{aligned}\hat{\gamma}_{MLE}(d,\mathbf{b})_{(i)} &= (\mathbf{A}-z_iz_i')^{-۱}(\mathbf{Z}'y-Z_iy_i)+(\mathbf{B}-z_iz_i')^{-۱}\mathbf{b}\\ &= \hat{\gamma}(d)+\mathbf{B}^{-۱}b-\frac{\mathbf{A}^{-۱}z_ie_i}{۱-w_i}+\frac{\mathbf{B}^{-۱}z_iz_i'\mathbf{B}^{-۱}}{۱-u_i}\mathbf{b}\\ &= \hat{\gamma}_{MLE}(d,\mathbf{b})-\frac{\mathbf{A}^{-۱}z_ie_i(۱-u_i)-(۱-w_i)\mathbf{B}^{-۱}z_iz_i'\mathbf{B}^{-۱}\mathbf{b}}{(۱-w_i)(۱-u_i)} \qquad (۳.۴)\end{aligned}$$

که در آن z_i'، i–امین سطر از ماتریس $\mathbf{Z}$، $e_i = y_i - z_i'\hat{\gamma}_{MLE}(d,\mathbf{b})$، i–امین باقیمانده لیو، و

$$\begin{cases}u_i = z_i'\mathbf{B}^{-۱}z_i\\ w_i = z_i'\mathbf{A}^{-۱}z_i.\end{cases}$$

بنابراین، شبه مقادیر وزنی جک‌نایف برآوردگر اصلاح شده لیو با تعریف رابطه‌ای که توسط هینکلی (۱۹۷۷) و نیکوئیست (۱۹۸۸) بیان شد، به صورت زیر است

$$\begin{aligned}Q_i &= \hat{\gamma}_{MLE}(d,\mathbf{b})\\ &\quad+\sum_{i=۱}^{n}(۱-h_i)\left(\frac{\mathbf{A}^{-۱}z_ie_i(۱-u_i)-(۱-w_i)\mathbf{B}^{-۱}z_iz_i'\mathbf{B}^{-۱}\mathbf{b}}{۱-h_i}\right)\end{aligned}$$

که در آن
$$h_i = ۱ - (u_i + w_i - u_i w_i).$$
با میانگین گرفتن از مقادیر جک‌نایف، برآوردگر جک‌نایف با اریبی کاهش یافته را به صورت زیر ارائه می‌کنیم

$$\begin{aligned}\hat{\gamma}_{JMLE}(d, \mathbf{b}) &= \frac{۱}{n}\sum_{i=۱}^{n} Q_i \\ &= \hat{\gamma}_{MLE}(d, \mathbf{b}) + \mathbf{A}^{-۱}\sum_{i=۱}^{n} z_i e_i(۱ - u_i) \\ &- \sum_{i=۱}^{n}(۱ - w_i)\mathbf{B}^{-۱} z_i z_i' \mathbf{B}^{-۱}\mathbf{b} \end{aligned} \tag{۴.۴}$$

که در آن

$$\begin{aligned}\sum_{i=۱}^{n} z_i e_i(۱ - u_i) &= \sum_{i=۱}^{n} z_i e_i - \sum_{i=۱}^{n} z_i e_i u_i \\ &= \sum_{i=۱}^{n} z_i(y_i - z_i'\hat{\gamma}_{MLE}(d, \mathbf{b})) - \sum_{i=۱}^{n} z_i(y_i - z_i'\hat{\gamma}_{MLE}(d, \mathbf{b}))(z_i'\mathbf{B}^{-۱} z_i) \\ &= \sum_{i=۱}^{n} z_i y_i - \sum_{i=۱}^{n} z_i z_i'\hat{\gamma}_{MLE}(d, \mathbf{b}) - \sum_{i=۱}^{n} z_i y_i + \sum_{i=۱}^{n} z_i z_i'\hat{\gamma}_{MLE}(d) z_i'\mathbf{B}^{-۱} z_i \\ &= -\mathbf{\Lambda}\hat{\gamma}_{MLE}(d, \mathbf{b}) + \sum_{i=۱}^{n}\frac{z_i z_i'}{(z_i'\mathbf{B}^{-۱} z_i)^{-۱}}\hat{\gamma}_{MLE}(d, \mathbf{b}) \end{aligned} \tag{۵.۴}$$

و

$$\begin{aligned}\sum_{i=۱}^{n}(۱ - w_i)\mathbf{B}^{-۱} z_i z_i'\mathbf{B}^{-۱} &= \sum_{i=۱}^{n}\mathbf{B}^{-۱} z_i z_i'\mathbf{B}^{-۱} - \sum_{i=۱}^{n}\mathbf{B}^{-۱} z_i z_i'\mathbf{B}^{-۱} z_i'\mathbf{A}^{-۱} z_i \\ &= \mathbf{B}^{-۱}\sum_{i=۱}^{n} z_i z_i'\mathbf{B}^{-۱} - \mathbf{B}^{-۱}\sum_{i=۱}^{n}\frac{z_i z_i'}{(z_i\mathbf{A}^{-۱} z_i)}\mathbf{B}^{-۱}. \end{aligned} \tag{۶.۴}$$

در نهایت برآوردگر جک‌نایف لیو اصلاح شده[۱] (JMLE)، با استفاده از روابط (۵.۴) و (۶.۴) به صورت زیر حاصل می‌شود

$$\begin{aligned}\hat{\gamma}_{JMLE}(d, \mathbf{b}) &= \hat{\gamma}_{MLE}(d, \mathbf{b}) + \mathbf{A}^{-۱}\left(\sum_{i=۱}^{n}\frac{z_i z_i'}{(z_i\mathbf{A}^{-۱} z_i)} - \mathbf{\Lambda}\right)\hat{\gamma}_{MLE}(d, \mathbf{b}) \\ &+ \mathbf{B}^{-۱}\left(\sum_{i=۱}^{n}\frac{z_i z_i'}{(z_i\mathbf{A}^{-۱} z_i)} - \mathbf{\Lambda}\right)\mathbf{B}^{-۱}\mathbf{b} \\ &= \hat{\gamma}_{MLE}(d, \mathbf{b}) + \mathbf{A}^{-۱}(\mathbf{T}'\mathbf{T} - \mathbf{\Lambda})\hat{\gamma}_{MLE}(d, \mathbf{b}) \\ &+ \mathbf{B}^{-۱}(\mathbf{C}'\mathbf{C} - \mathbf{\Lambda})\mathbf{B}^{-۱}\mathbf{b} \end{aligned} \tag{۷.۴}$$

[۱] Jackknifed modified liu estimator

که در آن سطر i–ام ماتریس‌های $\mathbf{T}$ و $\mathbf{C}$ عبارتند از

$$\begin{cases} T_i = \dfrac{z_i'}{(z_i'\mathbf{B}^{-1}z_i)^{\frac{-1}{2}}} \\ C_i = \dfrac{z_i'}{(z_i'\mathbf{A}^{-1}z_i)^{\frac{-1}{2}}}. \end{cases}$$

با جایگذاری مقدار ماتریس $\mathbf{B}$ در رابطه (۷.۴) به عبارت زیر دست می‌یابیم

$$\begin{aligned} \hat{\gamma}_{JMLE}(d,\mathbf{b}) &= \hat{\gamma}_{MLE}(d,\mathbf{b}) + \mathbf{A}^{-1}\mathbf{T}'\mathbf{T}\hat{\gamma}_{MLE}(d,\mathbf{b}) \\ &\quad - \mathbf{F}_d\hat{\gamma}_{MLE}(d,\mathbf{b}) + (\mathbf{I}-\mathbf{F}_d)(\mathbf{C}'\mathbf{C}-\mathbf{\Lambda})(\mathbf{I}-\mathbf{F}_d)\mathbf{b} \qquad (۸.۴) \end{aligned}$$

که در آن

$$\begin{cases} \mathbf{B}^{-1} = \mathbf{I} - \mathbf{F}_d \\ \mathbf{A}^{-1}\mathbf{\Lambda} = \mathbf{F}_d. \end{cases}$$

با فاکتورگیری از عبارت‌های مشترک در رابطه (۸.۴) و انجام محاسبات ماتریسی، برآوردگر جک‌نایف لیو اصلاح شده در نهایت به صورت زیر حاصل می‌شود

$$\begin{aligned} \hat{\gamma}_{JMLE}(d,\mathbf{b}) &= (\mathbf{A}^{-1}\mathbf{T}'\mathbf{T} + \mathbf{I} - \mathbf{F}_d)\hat{\gamma}_{MLE}(d,\mathbf{b}) + \mathbf{B}^{-1}(\mathbf{C}'\mathbf{C}-\mathbf{\Lambda})\mathbf{B}^{-1}\mathbf{b} \\ &= (\mathbf{A}^{-1}\mathbf{T}'\mathbf{T} + \mathbf{B}^{-1})\hat{\gamma}_{MLE}(d,\mathbf{b}) + \mathbf{B}^{-1}(\mathbf{C}'\mathbf{C}-\mathbf{\Lambda})\mathbf{B}^{-1}\mathbf{b}. \qquad (۹.۴) \end{aligned}$$

با استفاده از این تکنیک، امید است که برآوردگر معرفی شده دارای اریبی کمتری نسبت به برآوردگر لیو اصلاح شده داشته باشد. چرا که با توجه به ۱.۶.۱ هدف از جک‌نایف، کاهش اریبی برآوردگرها می‌باشد. در بخش بعد مقدار اریبی این برآوردگر را محاسبه می‌کنیم.

۳.۴ اریبی و واریانس برآوردگر

همان‌طور که پیش‌تر بیان شد، یکی از مباحث مهم در استنباط آماری، بررسی ویژگی‌های اصلی برآوردها از جمله اریبی، واریانس آن‌هاست. استفاده از برخی روش‌های معمول برای این منظور، عمدتاً پیچیده و حتی غیر ممکن هستند و در برخی از آنان نیاز به برقراری فرض‌های تئوری پیچیده وجود دارد. برآوردگری که در این فصل معرفی می‌شود از جمله همین برآوردگرها می‌باشد که یافتن اریبی و واریانس آن نیازمند محاسبات و فرض‌های پیچیده است. برای پی بردن به این واقعیت، در این بخش طبق تعریفِ محاسبه اریبی برآوردگرها، اریبی برآوردگر جک‌نایف لیو اصلاح شده را با استفاده از رابطه (۱۸.۳) به شکل زیر

محاسبه می‌کنیم

$$\begin{aligned}\text{bias}(\hat{\gamma}_{JMLE}(d,\mathbf{b})) &= (\mathbf{A}^{-1}\mathbf{T}'\mathbf{T}+\mathbf{B}^{-1})E(\hat{\gamma}_{MLE}(d,\mathbf{b}))+\mathbf{B}^{-1}(\mathbf{C}'\mathbf{C}-\mathbf{\Lambda})\mathbf{B}^{-1}\mathbf{b}-\gamma\\ &= (\mathbf{A}^{-1}\mathbf{T}'\mathbf{T}+\mathbf{B}^{-1})\left[\mathbf{A}_d\gamma+(\mathbf{I}-\mathbf{A}_d)\mathbf{b}\right]+\mathbf{B}^{-1}(\mathbf{C}'\mathbf{C}-\mathbf{\Lambda})\mathbf{B}^{-1}\mathbf{b}-\gamma\\ &= (\mathbf{A}^{-1}\mathbf{T}'\mathbf{T}+\mathbf{B}^{-1})\mathbf{A}_d\gamma+(\mathbf{A}^{-1}\mathbf{T}'\mathbf{T}+\mathbf{B}^{-1})(\mathbf{I}-\mathbf{A}_d)\mathbf{b}\\ &+\mathbf{B}^{-1}(\mathbf{C}'\mathbf{C}\mathbf{b}-\mathbf{\Lambda})\mathbf{B}^{-1}-\gamma\\ &= \mathbf{W}\mathbf{A}_d\gamma+\mathbf{W}(\mathbf{I}-\mathbf{A}_d)\mathbf{b}+\mathbf{B}^{-1}(\mathbf{C}'\mathbf{C}-\mathbf{\Lambda})\mathbf{B}^{-1}\mathbf{b}-\gamma\end{aligned} \tag{۱۰.۴}$$

که در آن

$$\mathbf{W}=(\mathbf{A}^{-1}\mathbf{T}'\mathbf{T}+\mathbf{B}^{-1}).$$

با فاکتورگیری از عبارت‌های مشترک به روش زیر

$$\begin{aligned}\mathbf{W}\mathbf{A}_d\gamma-\gamma+\mathbf{W}(\mathbf{I}-\mathbf{A}_d)\mathbf{b} &= (\mathbf{W}\mathbf{A}_d-\mathbf{I})\gamma+\mathbf{W}(\mathbf{I}-\mathbf{A}_d)\mathbf{b}\\ &= (\mathbf{W}\mathbf{A}_d-\mathbf{I})\gamma-(\mathbf{W}\mathbf{A}_d-\mathbf{I})b+(\mathbf{W}-\mathbf{I})\mathbf{b}\\ &= (\mathbf{W}\mathbf{A}_d-\mathbf{I})(\gamma-\mathbf{b})+(\mathbf{W}-\mathbf{I})\mathbf{b}\end{aligned} \tag{۱۱.۴}$$

اریبی برآوردگر به صورت زیر به‌دست می‌آید

$$\begin{aligned}\text{bias}(\hat{\gamma}_{JMLE}(d,\mathbf{b})) &= (\mathbf{W}\mathbf{A}_d-\mathbf{I})(\gamma-\mathbf{b})+(\mathbf{W}-\mathbf{I})\mathbf{b}\\ &+\mathbf{B}^{-1}(\mathbf{C}'\mathbf{C}-\mathbf{\Lambda})\mathbf{B}^{-1}\mathbf{b}.\end{aligned} \tag{۱۲.۴}$$

که به صورت ضریبی از برآوردگر کمترین توان‌های دوم و مقدار اریبی به صورت اطلاع پیشین می‌باشد. ماتریس واریانس- کوواریانس به صورت زیر قابل محاسبه است

$$\begin{aligned}\text{cov}(\hat{\gamma}_{JMLE}(d,\mathbf{b})) &= \text{cov}\left[(\mathbf{A}^{-1}\mathbf{T}'\mathbf{T}+\mathbf{B}^{-1})\hat{\gamma}_{MLE}(d,\mathbf{b})+\mathbf{B}^{-1}(\mathbf{C}'\mathbf{C}-\mathbf{\Lambda})\mathbf{B}^{-1}\right]\mathbf{b}\\ &= \text{cov}\left[(\mathbf{A}^{-1}\mathbf{T}'\mathbf{T}+\mathbf{B}^{-1})\hat{\gamma}_{MLE}(d,\mathbf{b})\right]\\ &= \mathbf{W}\text{cov}(\hat{\gamma}_{MLE}(d,\mathbf{b}))\mathbf{W}'\\ &= \sigma^{2}\left(\mathbf{W}\mathbf{A}_d\mathbf{\Lambda}^{-1}\mathbf{A}_d'\mathbf{W}'\right)\end{aligned} \tag{۱۳.۴}$$

که در آن

$$\begin{cases}\mathbf{A}_d=\mathbf{F}_d\\ \mathbf{W}=\mathbf{A}^{-1}\mathbf{T}'\mathbf{T}+\mathbf{B}^{-1}.\end{cases}$$

۴.۴ کارایی برآوردگر جک‌نایف لیو اصلاح شده

در بخش ۳.۴ دیدیم که برآوردگر جک‌نایف لیو اصلاح شده اریب است. با توجه به این‌که معیاری مناسب برای اندازه‌گیری کارایی این برآوردگر، $MSEM$ است، ماتریس میانگین توان دوم خطا برای برآوردگر پیشنهادی را با استفاده از تعریف ب.۰.۰.۳ و رابطه‌های (۱۲.۴) و (۱۳.۴) به صورت زیر محاسبه می‌کنیم

$$\begin{aligned}\mathrm{MSEM}(\hat{\gamma}_{JMLE}(d,\mathbf{b})) &= \sigma^{2}\left(\mathbf{W}\mathbf{A}_d\mathbf{\Lambda}^{-1}\mathbf{A}_d'\mathbf{W}'\right)\\ &+[(\mathbf{W}\mathbf{A}_d-\mathbf{I})(\gamma-\mathbf{b})+(\mathbf{W}-\mathbf{I})\mathbf{b}+\mathbf{R}]\\ &\times[(\mathbf{W}\mathbf{A}_d-\mathbf{I})(\gamma-\mathbf{b})+(\mathbf{W}-\mathbf{I})\mathbf{b}+\mathbf{R}]' \qquad (۱۴.۴)\end{aligned}$$

که در آن

$$\mathbf{R}=\mathbf{B}^{-1}(\mathbf{C}'\mathbf{C}-\Lambda)\mathbf{B}^{-1}\mathbf{b}.$$

همچنین می‌دانیم که اریبی برآوردگر لیو اصلاح شده به صورت

$$\mathrm{bias}(\hat{\gamma}_{MLE}(d,\mathbf{b}))=(\mathbf{A}_d-\mathbf{I})(\gamma-\mathbf{b}) \qquad (۱۵.۴)$$

و کوواریانس و مخاطره آن نیز به صورت زیر است

$$\mathrm{cov}(\hat{\gamma}_{MLE}(d,\mathbf{b}))=\sigma^{2}\mathbf{A}_d\mathbf{\Lambda}^{-1}\mathbf{A}_d' \qquad (۱۶.۴)$$

$$\begin{aligned}MSEM(\hat{\gamma}_{MLE}(d,\mathbf{b})) &= \sigma^{2}\mathbf{A}_d\mathbf{\Lambda}^{-1}\mathbf{A}_d'\\ &+(\mathbf{A}_d-\mathbf{I})(\gamma-\mathbf{b})(\gamma-\mathbf{b})'(\mathbf{A}_d-\mathbf{I}))'. \qquad (۱۷.۴)\end{aligned}$$

بنابراین اختلاف مخاطره برآوردگر جک‌نایف لیو اصلاح شده از برآوردگر لیو اصلاح شده، با توجه به روابط (۱۴.۴) و (۱۷.۴) به صورت زیر می‌باشد

$$\begin{aligned}\Delta_{1} &= MSEM(\hat{\gamma}_{MLE}(d,\mathbf{b}))-\mathrm{MSEM}(\hat{\gamma}_{JMLE}(d,\mathbf{b}))\\ &=-\sigma^{2}\left(\mathbf{W}\mathbf{A}_d\Lambda^{-1}\mathbf{A}_d'\mathbf{W}'\right)+\sigma^{2}\mathbf{A}_d\mathbf{\Lambda}^{-1}\mathbf{A}_d'\\ &+[(\mathbf{W}\mathbf{A}_d-\mathbf{I})(\gamma-\mathbf{b})+(\mathbf{W}-\mathbf{I})\mathbf{b}+\mathbf{R}]\\ &\times[(\mathbf{W}\mathbf{A}_d-\mathbf{I})(\gamma-\mathbf{b})+(\mathbf{W}-\mathbf{I})\mathbf{b}+\mathbf{R}]'\\ &-(\mathbf{A}_d-\mathbf{I})(\gamma-\mathbf{b})(\gamma-\mathbf{b})'(\mathbf{A}_d-\mathbf{I})'. \qquad (۱۸.۴)\end{aligned}$$

در چنین شرایطی، که مقایسه برآوردگرها نیاز به محاسبات پیچیده و طاقت فرسا دارد، استفاده از روش‌های بازنمونه‌گیری می‌تواند راهگشا باشد، زیرا مهم‌ترین مزیت این روش عدم نیاز به محاسبات تئوری و برقراری فرض‌های پیچیده است. به همین علت در بخش بعد برای مقایسه کارایی دو برآوردگر جک‌نایف لیو اصلاح شده و برآوردگر لیو اصلاح شده از روش شبیه‌سازی استفاده شده است.

۵.۴ مطالعه شبیه‌سازی

در این بخش برای بررسی عملکرد برآوردگر جک‌نایف لیو اصلاح شده و مقایسه رفتار آن با سایر برآوردگرهای ارائه شده، روش شبیه‌سازی مونت کارلو را به‌کار می‌گیریم. همانند بخش ۵.۳ برای داشتن داده‌هایی با میزان همخطی مختلف، روش لیو (۲۰۰۳) را به‌کار می‌بریم

$$x_{ij} = (1-\gamma^2)^{1/2}\omega_{ij} + \gamma\omega_{ip}, \quad i = 1, 2, \cdots, n, \; j = 1, 2, \cdots, p-1$$

که در آن اعداد ω_{ij} تصادفی مستقل از توزیع نرمال استاندارد و γ^2 همبستگی بین دو متغیر پیش‌بین است.

این متغیرها استاندارد شده‌اند، به طوری‌که $\mathbf{X}'\mathbf{X}$ و $\mathbf{X}'y$ همبسته هستند. حال با استفاده از رابطه زیر، n مشاهده برای متغیر پاسخ مشخص می‌شود

$$y_i = \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + e_i$$

که در آن e_i اعداد تصافی مستقل از توزیع نرمال با میانگین صفر و واریانس σ^2 هستند.

این آزمایش را با مقادیر مختلف γ : ۰/۷, ۰/۹۰, ۰/۹۹ و σ : ۰/۰۱, ۰/۱, ۱ برای ۱۰۰ بار تکرار می‌کنیم. مقدار میانگین توان دوم خطا را برای برآوردگر کمترین توان‌های دوم، برآوردگر لیو، برآوردگر لیو اصلاح شده و برآوردگر جک‌نایف لیو اصلاح شده، با استفاده از رابطه‌ی زیر برآورد می‌کنیم

$$MSE(\tilde{\beta}) = \frac{1}{100}\sum_{j=1}^{30}\sum_{i=1}^{3}(\tilde{\beta}_{ij} - \beta_i)^2$$

که در آن $\tilde{\beta}_{ij}$ برآورد i–امین پارامتر در j–امین تکرار و β_i مقدار پارامتر است. برنامه مربوط به این مثال در پیوست ت.۸ آمده است.

نتایج شبیه‌سازی در جداول ۱.۴، ۲.۴ و ۳.۴ آمده است. مقادیر برآورد میانگین توان‌های دوم خطا در بیشتر حالت‌ها برای برآوردگر لیو، برآوردگر لیو اصلاح شده و برآوردگر جک‌نایف لیو اصلاح شده نسبت به برآوردگر کمترین توان‌های دوم کمتر است. بعلاوه می‌توان گفت با افزایش همخطی برآوردگر جک‌نایف اصلاح شده دارای خطای کمتر و در نتیجه عملکرد بهتری نسبت به دیگر برآوردگرها می‌باشد.

جدول ۱.۴: برآورد MSE برای LSE ، LE ، MLE و $JMLE$ به ازای $\gamma = ۰/۷۰$

$\sigma = ۱$			$\sigma = ۰/۱$			$\sigma = ۰/۰۱$			
d=۱	d=۰/۰۵	d=۰	d=۱	d=۰/۰۵	d=۰	d=۱	d=۰/۰۵	d=۰	
۰/۰۷۰۷	۰/۰۷۰۷	۰/۰۷۰۷	۲/۱۳۸۵	۲/۱۳۸۵	۲/۱۳۸۵	۷/۱۲۵۸	۷/۱۲۵۸	۷/۱۲۵۸	LSE
۰/۰۱۷۰	۰/۰۶۵۹	۰/۰۶۷۵	۰/۱۳۵۲	۰/۱۶۶۴	۱/۰۹۷۶	۴/۰۳۶۷	۶/۴۶۷۶	۶/۹۰۹۵	LE
۰/۰۱۶۱	۰/۰۶۶۱	۰/۰۶۷۵	۰/۰۴۷۵	۰/۱۶۶۹	۱/۰۹۷۶	۴/۰۳۳۵	۶/۴۵۹۴	۶/۹۰۹۵	MLE
۰/۰۰۸۳	۰/۰۰۱۷	۰/۰۰۵۱	۰/۰۰۰۱	۰/۰۰۹۹	۰/۰۳۲۱	۰/۰۰۷۳	۵/۰۹۶۱	۱/۰۲۷۸	JMLE

جدول ۲.۴: برآوردMSE برای LSE ، LE ، MLE و $JMLE$ به ازای $\gamma = ۰/۹۰$

$\sigma = ۱$			$\sigma = ۰/۱$			$\sigma = ۰/۰۱$			
d=۱	d=۰/۰۵	d=۰	d=۱	d=۰/۰۵	d=۰	d=۱	d=۰/۰۵	d=۰	
۰/۱۷۳۳	۰/۱۷۳۳	۰/۱۷۳۳	۰/۸۸۰۴	۰/۸۸۰۴	۰/۸۸۰۴	۳/۷۲۷۵	۳/۷۲۷۵	۳/۷۲۷۵	LSE
۰/۱۴۲۸	۰/۱۲۴۰	۰/۰۷۸۲	۰/۶۸۸۸	۰/۵۴۱۵	۰/۶۸۰۸	۳/۰۴۶۶	۱/۰۵۸۱	۰/۲۷۶۶	LE
۰/۱۴۲۸	۰/۱۲۴۰	۰/۰۷۸۱	۰/۶۸۸۸	۰/۵۳۸۹	۰/۶۳۴۷	۳/۰۴۶۶	۱/۰۵۴۹	۰/۲۶۶۵	MLE
۰/۰۰۳۰	۰/۰۰۳۱	۰/۰۰۰۷	۰/۰۰۹۹	۰/۰۰۶۹	۰/۰۰۰۱	۰/۰۸۰۲	۰/۰۱۵۲	۰/۰۰۹۷	JMLE

جدول ۳.۴: برآورد MSE برای LSE ، LE ، MLE و $JMLE$ به ازای $\gamma = ۰/۹۹$

$\sigma = ۱$			$\sigma = ۰/۱$			$\sigma = ۰/۰۱$			
d=۱	d=۰/۰۵	d=۰	d=۱	d=۰/۰۵	d=۰	d=۱	d=۰/۰۵	d=۰	
۰/۲۳۴۲	۰/۲۳۴۲	۰/۲۳۴۲	۰/۷۵۰۶	۰/۷۵۰۶	۰/۷۵۰۶	۱/۷۵۷۲	۱/۷۵۷۲	۱/۷۵۷۲	LSE
۰/۰۹۷۶	۰/۰۴۸۸	۰/۶۰۱	۰/۲۳۳۹	۰/۱۴۸۸	۰/۰۵۰۸	۰/۲۶۹۶	۰/۴۰۶۴	۰/۰۲۶۹	LE
۰/۰۹۷۶	۰/۴۶۶۱	۰/۰۶۰۱	۰/۲۳۳۹	۰/۱۴۷۱	۰/۰۵۰۸	۰/۲۶۹۶	۰/۳۹۹۲	۰/۰۲۶۹	MLE
۰/۰۰۰۶	۰/۰۰۰۸	۰/۰۰۱	۰/۰۰۱۸	۰/۰۰۹۲	۰/۰۰۰۳	۰/۰۳۰۱	۰/۰۰۴۵	۰/۰۰۱۵	JMLE

۶.۴ نتیجه‌گیری

بهبود دادن برآوردگرهای معادلات رگرسیونی، زمینه کاری بسیاری از آماردانان دهه اخیر است. یکی از عوامل که منجر به تلاش برای یافتن برآوردگرهای بهبود یافته می‌شود، وجود همخطی در ماتریس طرح است. همان‌طورکه در این پایان‌نامه شرح آن گذشت، یکی از این برآوردگرهای بهبود یافته که در سال‌های اخیر از سوی آماردانان مورد توجه زیادی قرار گرفته است، برآوردگر ریج می‌باشد.

هنگامی‌که با مسئله همخطی مواجه هستیم، برآوردگر ریج در کلاس برآوردگرهای اریب، بهترین برآوردگر است و همان‌طورکه نشان دادیم، خطای آن از خطای برآوردگر کمترین توان‌های دوم کمتر است و همچنین دارای ویژگی‌های مفید و خوبی است که استفاده از این برآوردگر را توجیه می‌کنند.

در این مجموعه سعی شده است، ضرایب معادلات رگرسیونی چندگانه با در نظرگرفتن مسئله همخطی و استفاده از داده‌هایی که در آن‌ها این مشکل به چشم می‌خورد، برآورد کنیم. در این راستا در فصل اول

برآوردگر نوع لیو به عنوان برآوردگری از نوع برآوردگرهای ریج، معرفی شده و خواص آن مورد بررسی قرار گرفت. در فصل دوم، ابتدا با استفاده از تکنیک جک‌نایف، برآوردگر جک‌نایف لیو محاسبه شد و به منظور نشان دادن کارایی این برآوردگر، خطای آن محاسبه گردید و به کمک شبیه‌سازی و مثال عددی برتری آن را بر برآوردگر کمترین توان‌های دوم، نشان دادیم.

به عنوان یک نوآوری در این پایان‌نامه، تکنیک جک‌نایف را برای برآوردگر لیو اصلاح شده به‌کار بردیم و ضرایب مدل رگرسیونی را برآورد کردیم. برآوردگر پیشنهادی با استفاده از روش‌های جک‌نایف معرفی شده توسط هینکلی (۱۹۷۷) و نیکوئیست (۱۹۸۸) محاسبه شده است و با استفاده از روش‌های شبیه‌سازی مونت‌کارلویی و جک‌نایف نشان داده‌ایم که در این مدل خطی نیز، برآوردگر جک‌نایف لیو اصلاح شده نسبت به برآوردگر کمترین توان‌های دوم و حتی برآوردگر لیو اصلاح شده برتری دارد.

۷.۴ پیشنهادات برای آینده تحقیق

با توجه به نتایج به‌دست آمده در این مجموعه و موضوعات مطرح شده، می‌توان در آینده بر روی موارد زیر تحقیق کرد که بعضی از آن‌ها در دست مطالعه و بررسی نویسنده این پایان‌نامه است.

۱. روش جک‌نایف مورد استفاده در این پایان‌نامه را می‌توان، برای برآوردگر لیو وزنی یانگ و همکاران (۲۰۰۹)، که ترکیبی از برآوردگر آمیخته وزنی و برآوردگر نوع لیو است، و برآوردگر محدودکننده تصادفی هیبرت و وایجکوان (۲۰۰۶)، به‌کار برد و عملکرد برآوردگر حاصل را مورد سنجش قرار داد.

۲. در مسائل کاربردی و مدل‌های اقتصادی، استفاده از اطلاع پیشین در خصوص بعضی از پارامترهای مورد علاقه در مدل آماری تحت بررسی، اغلب منجر به استفاده از روش‌های پیشرفته و تعمیم یافته‌ای می‌شود که با توجه به قطعی یا غیر قطعی بودن این اطلاع، به‌کار برده می‌شوند. استفاده از اطلاع پیشین معلوم در مدل‌های آماری، معمولاً به صورت یک قید مطرح است که در نهایت به مدل‌های محدود شده یا مقید منتج می‌شود. برآوردگری که بر اساس مدل‌های محدود شده به‌دست می‌آید، برآوردگر محدود شده نامیده می‌شود. به عنوان یک پیشنهاد می‌توان تکنیک‌های جک‌نایف معرفی شده در این مجموعه را بر روی این‌گونه مدل‌ها انجام و عملکرد برآوردگر جدید را مورد بررسی قرار داد، که در دست بررسی نویسنده این مجموعه می‌باشد و نتایجی در این خصوص نیز به‌دست آمده است.

۳. با استفاده از روش‌های مورد استفاده در این پایان‌نامه، می‌توان برآوردگر جدید ارائه شده در فصل۴ را در مدل‌های نیمه پارامتری با تلفیق نتایج مقاله آکدنیز و آکدنیز دوران (۲۰۰۹) به‌کار برد.

۴. تمامی برآوردگرهای مطرح شده در این پایان‌نامه را می‌توان در رویارویی با مدل‌های رگرسیونی استوار[۲] به‌کار برد.

[۲] Robust regression

۵. در نهایت تمامی برآوردگرهای مطرح شده در این پایان‌نامه را می‌توان در رویارویی با مدل‌های رگرسیونی فازی بررسی کرده و به‌کار برد.

پیوست آ

جبر ماتریس‌ها

داده‌های چند متغیری را به طور مناسبی می‌توان به صورت آرایه‌ای از اعداد نمایش داد. به طور کلی یک آرایه‌ی مستطیلی از اعداد با مثلاً p سطر و n ستون را یک ماتریس $p \times n$ بعدی می‌نامند. با استفاده از جبر ماتریس‌ها مطالعه‌ی روش‌های چندمتغیری بسیار آسان می‌شود.
نتایج جبر ماتریس‌های ارائه شده در این بخش به ما توان اجمالی روش‌های آماری را می‌دهد. علاوه بر این روابط متعارف بیان شده برحسب ماتریس‌ها به سهولت در رایانه‌ها برنامه‌ریزی شده، تا بتوانیم محاسبات عادی و مهم کمیت‌های آماری را انجام دهیم.
موضوع را با ارائه‌ی بعضی مفاهیم بنیادی که برای تعبیر و تفسیرهای هندسی و بیان جبری روش‌های آماری لازم هستند، شروع می‌کنیم. مطالب این پیوست عمدتاً برگرفته از رنچر (۲۰۰۲) و حسن زاده بشتیان (۱۳۸۸) می‌باشند.

تست

تعریف آ.۰.۱ . مجموعه بردارهای $\mathbf{x}_1, \cdots, \mathbf{x}_k$ را نامستقل خطی[۱] گوئیم هرگاه k عدد $(a_1, \cdots, a_k)$ که همگی صفر نباشند وجود داشته باشد که $a_1\mathbf{x}_1 + a_2\mathbf{x}_2 + \cdots + a_k\mathbf{x}_k = 0$ در غیر این‌صورت مجموعه‌ی بردارها را مستقل خطی می‌گویند.

تعریف آ.۰.۲ . اگر سطرها و ستون‌های ماتریس $\mathbf{A}$ را عوض کنیم، ماتریس حاصل را ترانهاده[۲] $\mathbf{A}$ گوئیم. به طور کلی اگر $\mathbf{A} = (a_{ij})$

$$\mathbf{A}' = (a_{ij})' = (a_{ji}).$$

قضیه آ.۰.۳. *اگر $\mathbf{A}$ یک ماتریس دلخواه باشد، آن‌گاه*

$$(\mathbf{A}')' = \mathbf{A}.$$

[۱] Linear dependence

[۲] Transpose

تعریف آ.۴.۰. فرض کنید $\mathbf{A}$ یک ماتریس مربع[۳] باشد. ماتریس $\mathbf{A}$ را متقارن[۴] گوئیم هرگاه $\mathbf{A} = \mathbf{A}'$.

تعریف آ.۵.۰. اگر $\mathbf{A}_{n\times m}$ و $\mathbf{B}_{m\times p}$، در این‌صورت $\mathbf{A}$ و $\mathbf{B}$ را سازگار[۵] گوئیم و حاصل‌ضرب $\mathbf{C} = \mathbf{AB}$ عبارت است از

$$c_{ij} = \sum_k a_{ik}b_{kj}$$

قضیه آ.۶.۰. *برای تمام ماتریس‌های $\mathbf{A}$، $\mathbf{B}$ و $\mathbf{C}$ (با بُعدهایی که حاصل‌ضرب‌ها تعریف‌شده باشند)*

$$(\mathbf{B}+\mathbf{C})\mathbf{A} = \mathbf{BA} + \mathbf{CA}$$

$$(\mathbf{AB})' = \mathbf{B}'\mathbf{A}'$$

تعریف آ.۷.۰. رتبه[۶] سطری یک ماتریس بیشترین تعداد سطرهای مستقل خطی[۷] هستند که به صورت بردار (یعنی بردارهای سطری) در نظر گرفته می‌شوند. رتبه‌ی ستونی یک ماتریس رتبه‌ی مجموعه ستون‌های آن است که به صورت بردار درنظر گرفته می‌شوند.
می‌توان نشان داد که رتبه‌ی سطری و رتبه‌ی ستونی یک ماتریس باهم برابرند.

تعریف آ.۸.۰. اگر $\mathbf{A}_{n\times p}$ دارای رتبه‌ی p باشد که در آن $p < n$، در این‌صورت $\mathbf{A}$ دارای رتبه‌ی ماکزیمم است و آن را با رتبه‌ی کامل گوئیم. به طور کلی، رتبه‌ی ماکزیمم ممکن یک ماتریس $n \times p$ برابر $\min(n,p)$ است.

تعریف آ.۹.۰. یک ماتریس مربع ناویژه است هرگاه رتبه اش برابر تعداد سطرها (یا ستون‌ها) باشد. یک ماتریس ناویژه $\mathbf{A}$ دارای معکوس منحصر بفرد است و

$$\mathbf{AA}^{-۱} = \mathbf{A}^{-۱}\mathbf{A} = \mathbf{I}$$

اگر $\mathbf{A}$ مربعی باشد ولی با رتبه‌ی کامل نباشد، آنگاه دارای معکوس نیست و آن را ویژه نامند.

تعریف آ.۱۰.۰. ماتریس $\mathbf{B}$ که در معادله‌ی $\mathbf{AB} = \mathbf{BA} = \mathbf{I}$ صدق کند را وارون[۸] $\mathbf{A}$ گوئیم و آن را با $\mathbf{A}^{-۱}$ نشان می‌دهند.

قضیه آ.۱۱.۰. *اگر $\mathbf{A}$ و $\mathbf{B}$ ماتریس‌های مربع با بُعد مساوی و دارای معکوس باشند، آن‌گاه موارد زیر برقرار است:*

$$(\mathbf{A}')^{-۱} = (\mathbf{A}^{-۱})'$$

$$(\mathbf{AB})^{-۱} = \mathbf{B}^{-۱}\mathbf{A}^{-۱}$$

---

[۳] Square matrix
[۴] Symmetric
[۵] Consistence
[۶] Rank
[۷] Linear independence
[۸] Inverse

قضیه آ.۱۲.۰. *اگر یک ماتریس مربع به صورت* $\mathbf{B}+\mathbf{cc}'$ *ناویژه باشد، که در آن* $\mathbf{c}$ *یک بردار و* $\mathbf{B}$ *یک ماتریس ناویژه است، آن‌گاه:*

$$(\mathbf{B}+\mathbf{cc}')^{-1}=\mathbf{B}^{-1}-\frac{\mathbf{B}^{-1}\mathbf{cc}'\mathbf{B}^{-1}}{1+\mathbf{cB}^{-1}\mathbf{c}}$$

تعریف آ.۱۳.۰. فرض کنید $\mathbf{A}$ یک ماتریس مربعی و λ یک مقدار ویژه[۹] $\mathbf{A}$ باشد. اگر $\mathbf{x}$ یک بردار مخالف صفر ($\mathbf{x}\neq 0$) باشد، به طوری که

$$\mathbf{Ax}=\lambda\mathbf{x}$$

آن‌گاه λ را مقدار ویژه ماتریس $\mathbf{A}$ که $\mathbf{x}$ بردار ویژه[۱۰] آن است، می‌نامند و با حل $|\mathbf{A}-\lambda\mathbf{I}|=0$ نسبت به λ بدست می‌آیند.

تعریف آ.۱۴.۰. دو بردار $\mathbf{a}_{n\times 1}$ و $\mathbf{b}_{n\times 1}$ را متعامد[۱۱] می‌نامند، هرگاه

$$\mathbf{a}'\mathbf{b}=a_1b_1+a_2b_2+\cdots+a_nb_n=0.$$

تعریف آ.۱۵.۰. ماتریس $\mathbf{A}=(\mathbf{c}_1,\mathbf{c}_2,\cdots,\mathbf{c}_p)$ متعامد است، هرگاه ستون‌های آن بردارهای متعامد باشند.

تعریف آ.۱۶.۰. تجزیه یک ماتریس به مقادیر ویژه و بردارهای ویژه را، تجزیه طیفی[۱۲] گویند. مثلاً فرض کنید ماتریس $\mathbf{A}$ دارای تجزیه زیر است

$$\mathbf{A}=\mathbf{P\Lambda P}^{-1}$$

در این‌صورت $\mathbf{\Lambda}$ یک ماتریس قطری با مقادیر ویژه $\mathbf{A}$ است و $\mathbf{P}$ ماتریس متعامدی است که ستون‌های آن بردارهای ویژه $\mathbf{A}$ متناظر با مقادیر ویژه آن است.

تعریف آ.۱۷.۰. ماتریس $\mathbf{A}$ را معین مثبت[۱۳] می‌گویند، اگر تمام مقادیر ویژه $\mathbf{A}$ مثبت باشد و اگر مقادیر ویژه $\mathbf{A}$ نامنفی باشند، گفته می‌شود ماتریس $\mathbf{A}$ نیمه معین مثبت[۱۴] است.

قضیه آ.۱۸.۰. *اگر* $\mathbf{A}_{n\times n}$ *معین مثبت (نیمه معین مثبت) با مقادیر ویژه* $\lambda_1,\lambda_2,\cdots,\lambda_n$ *باشد، آن‌گاه:*

(۱) عناصر قطر اصلی $a_{ii}>0$ $(a_{ii}\geqslant 0)$ *می‌باشند.*

(۲) $\mathbf{A}$ *ناویژه است.*

(۳) $\mathbf{A}^{-1}$ *معین مثبت است.*

(۴) $\mathbf{P}'\mathbf{AP}$ *معین مثبت (نیمه معین مثبت) است، که در آن* $\mathbf{P}$ *یک ماتریس ناویژه است.*

(۵) $\lambda_i>0,(\lambda_i\geqslant 0)$ *برای* $i=1,2,\cdots,n$.

[۹] Eigen value

[۱۰] Eigen vector

[۱۱] Orthogonal

[۱۲] Spectral factorization

[۱۳] Positive definite

[۱۴] Positive semidefinite

قضیه آ.۰.۱۹. *ماتریس متقارن* $\mathbf{A}$ *مثبت معین است اگر و فقط اگر ماتریس ناویژه* $\mathbf{P}$ *وجود داشته باشد به‌طوری که* $\mathbf{A} = \mathbf{P}'\mathbf{P}$.

قضیه آ.۰.۲۰. *فرض کنید* $\mathbf{A}$ *یک ماتریس* $n \times p$ *است.*
(۱) اگر $\operatorname{rank}(\mathbf{A}) = p$*، آن‌گاه* $\mathbf{A}'\mathbf{A}$ *معین مثبت است.*
(۲) اگر $\operatorname{rank}(\mathbf{A}) < p$*، آن‌گاه* $\mathbf{A}'\mathbf{A}$ *نیمه معین مثبت است.*

قضیه آ.۰.۲۱. *اگر* $\mathbf{A}$ *یک ماتریس ناویژه باشد،* $\mathbf{A}'$ *نیز ناویژه است و معکوس آن را می‌توان به صورت زیر نوشت*

$$(\mathbf{A}')^{-1} = (\mathbf{A}^{-1})'.$$

تعریف آ.۰.۲۲. یک ماتریس $\mathbf{A}$ را خودتوان گوییم اگر

$$\mathbf{A}^2 = \mathbf{A}.$$

قضیه آ.۰.۲۳. *اگر* $\mathbf{A}$ *یک ماتریس* $n \times n$ *متقارن و خودتوان با رتبه* r *باشد، آن‌گاه*

$$\operatorname{rank}(\mathbf{A}) = \operatorname{tr}(\mathbf{A}) = r.$$

قضیه آ.۰.۲۴. *فرض کنید* $u = \mathbf{a}'\mathbf{x} = \mathbf{x}'\mathbf{a}$*، که در آن* $\mathbf{a}' = (a_1, a_2, \cdots, a_p)$ *یک بردار از ثابت‌ها است. در این‌صورت:*

$$\frac{\partial u}{\partial \mathbf{x}} = \frac{\partial(\mathbf{a}'\mathbf{x})}{\partial \mathbf{x}} = \frac{\partial(\mathbf{x}'\mathbf{a})}{\partial \mathbf{x}} = \mathbf{a}$$

قضیه آ.۰.۲۵. *فرض کنید* $u = \mathbf{x}'\mathbf{A}\mathbf{x}$*، که در آن* $\mathbf{A}$ *یک ماتریس متقارن از ثابت‌ها است. در این‌صورت:*

$$\frac{\partial u}{\partial \mathbf{x}} = \frac{\partial(\mathbf{x}'\mathbf{A}\mathbf{x})}{\partial \mathbf{x}} = 2\mathbf{A}\mathbf{x}$$

برای اثبات قضایای (آ.۰.۳) تا (آ.۰.۲۵) به رنچر[15] (۲۰۰۰) و مونتگومری[16] (۱۹۹۲) مراجعه کنید.

---

[15] Rencher

[16] Montgomery

پیوست ب

تعاریف، لم‌ها و قضایای اساسی

مفاهیمی که در این پایان‌نامه مورد استفاده قرار گرفته‌اند، به صورت تعریف و به شرح زیر بیان کرده‌ایم.

تعریف ب.۰.۲۶. (ون در وارت[۱]، ۱۹۹۸) دو دنباله از اعداد ثابت $\{a_n\}$ و $\{b_n\}$ را در نظر بگیرید.
(۱) گوییم $b_n = o(a_n)$، هرگاه

$$\lim_{n\to\infty} |\frac{b_n}{a_n}| = \circ.$$

(۲) گوییم $b_n = O(a_n)$ ، اگر $\varepsilon > \circ$، $K(\varepsilon) > \circ$ و عدد صحیح مثبت $N(\varepsilon)$ وجود داشته باشند، آن‌گاه برای $n \geqslant N(\varepsilon)$،

$$|b_n| < K(\varepsilon)|a_n|.$$

تعریف ب.۰.۲۷. برای هر برآوردگر، میانگین توان دوم خطا به صورت زیر به‌دست می‌آید

$$MSE(\hat{\beta}, \beta) = E(\hat{\beta} - \beta)'(\hat{\beta} - \beta)$$

تعریف ب.۰.۲۸. (توزیع کی -دو مرکزی) اگر $\mathbf{X} \sim N_p(\mathbf{0}, \mathbf{I})$ باشد، آن‌گاه $\mathbf{U} = \mathbf{X}'\mathbf{X}$ دارای توزیع کی- دو مرکزی با p درجه آزادی (آن را با نماد $\mathbf{U} \sim \chi^۲(p)$ نمایش می‌دهیم) که تابع چگالی احتمال آن به صورت زیر است

$$f(u) = \frac{u^{\frac{p}{۲}} - ۱}{۲^{\frac{p}{۲}}\Gamma(\frac{p}{۲})} e^{\frac{-u}{۲}}, \quad u > \circ, p \in N.$$

تعریف ب.۰.۲۹. (توزیع نرمال چند متغیره[۲]) بردار تصادفی $\mathbf{X}$ دارای توزیع نرمال p- متغیره با بردار میانگین μ و ماتریس واریانس- کوواریانس $\mathbf{\Sigma}$ است (آن را با نماد $\mathbf{X} \sim N_p(\mu, \mathbf{\Sigma})$ نشان می‌دهیم) و تابع چگالی احتمال آن به صورت زیر است

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{۱}{(۲\pi)^{\frac{p}{۲}} |\mathbf{\Sigma}|^{\frac{۱}{۲}}} \exp \frac{-(\mathbf{x} - \mu)'\mathbf{\Sigma}^{-۱}(\mathbf{x} - \mu)}{۲}$$

که در آن $\mathbf{x}, \mu \in \mathbb{R}^{p\times p}, \mathbf{\Sigma} \geqslant \circ$.

[۱] Van der vaart

[۲] Multivariate normal

تعریف ب.۳۰.۰. ماتریس میانگین توان‌های دوم خطا، برای هر برآودگری به صورت زیر تعریف می‌شود

$$\mathrm{MSEM}(\hat{\beta}, \beta) = \mathrm{Cov}(\hat{\beta}) + [E(\hat{\beta} - \beta)][E(\hat{\beta} - \beta)']$$

تعریف ب.۳۱.۰. (بسط تیلور) برای تعریف بسط تیلور ابتدا نیاز داریم تا سری توانی را تعریف کنیم. یک سری به صورت زیر است

$$a_۰ + a_۱ x + a_۲ x^۲ + \cdots = \sum_{n=۰}^{\infty} a_n x^n,$$

که در آن $a_۰, a_۱, a_۲, \cdots$ اعدادی ثابت هستند، یک سری توانی از x می‌باشد. در حالت کلی، سری توانی برای $|x| < R$ همگرا و برای $|x| > R$ واگراست. در اینجا R عدد ثابتی است که آن را شعاع همگرایی سری می‌نامند. برای $|x| = R$، سری می‌تواند همگرا یا واگرا باشد.
بسط تیلور، نمایش یک تابع به صورت مجموع بی‌نهایت جمله است که از مشتق‌های تابع در یک نقطه به‌دست می‌آیند. با استفاده از این بسط، توابع را می‌توان حول یک نقطه با تعداد متناهی از جملات تقریب زد. بسط تیلور یک تابع $f(x)$ با مقادیر حقیقی یا مختلط که در همسایگی نقطه حقیقی یا مختلط $x_۰$ بی‌نهایت بار مشتق‌پذیر است، سری توانی زیر است

$$f(x) = f(x_۰) + \frac{f'(x_۰)(x - x_۰)}{۱} + \frac{f''(x_۰)(x - x_۰)^۲}{۲} + \frac{f'''(x_۰)(x - x_۰)^۳}{۳} + \cdots,$$

که می‌توانیم به صورت خلاصه‌تر زیر بنویسیم

$$f(x) = \sum_{n=۰}^{\infty} \frac{f^{(n)}(x_۰)}{n} (x - x_۰)^n.$$

لم ب.۳۲.۰. *(فاربرادر، ۱۹۷۰) فرض کنید ماتریس $\boldsymbol{M} > ۰$ و α یک بردار باشد، آن‌گاه $\boldsymbol{M} - \alpha\alpha' > ۰$ اگر و تنها اگر*

$$\alpha' \boldsymbol{M}^{-۱} \alpha < ۱.$$

برهان. فاربرادر[۳] (۱۹۷۰) را ببینید. □

لم ب.۳۳.۰. *(ترنکلر و توتنبرگ، ۱۹۹۰) فرض کنید $\hat{\beta}_۱$ و $\hat{\beta}_۲$ دو برآوردگر خطی مشابه از $\hat{\beta}$ به صورت زیر باشند*

$$\begin{cases} \hat{\beta}_۱ = \boldsymbol{A}_۱ y \\ \hat{\beta}_۲ = \boldsymbol{A}_۲ y \\ \boldsymbol{D} = (\boldsymbol{A}_۱ \boldsymbol{A}_۱' - \boldsymbol{A}_۲ \boldsymbol{A}_۲') > ۰. \end{cases}$$

آن‌گاه

$$\Delta(\hat{\beta}_۱, \hat{\beta}_۲) = MSEM(\hat{\beta}_۱) - MSEM(\hat{\beta}_۲) = \sigma^۲ \boldsymbol{D} + \boldsymbol{b}_۱ \boldsymbol{b}_۱' - \boldsymbol{b}_۲ \boldsymbol{b}_۲' > ۰$$

اگر و فقط اگر

$$\boldsymbol{b}_۲' \left[\sigma^۲ \boldsymbol{D} + \boldsymbol{b}_۱ \boldsymbol{b}_۱' \right] < ۱$$

[۳] Farebrother

که در آن، ماتریس میانگین توان‌های دوم و اریبی برآوردگرها به صورت زیر تعریف می‌شوند

$$\begin{cases} \boldsymbol{b}_i = Bias(\hat{\beta}_i) = (A_i\boldsymbol{X} - \boldsymbol{I})\beta \\ MSEM(\hat{\beta}_i) = cov(\hat{\beta}_i) + \boldsymbol{b}_i\boldsymbol{b}_i' \quad i = ۱, ۲. \end{cases}$$

برهان. ترنکلر و توتنبرگ[۴](۱۹۹۰) را ببینید. □

قضیه ب.۳۴.۰. *اگر* $\boldsymbol{A}$ *ماتریس متقارن ثابت،* y *بردار تصادفی با بردار میانگین* μ *و ماتریس واریانس–کوواریانس* $\boldsymbol{V}$ *باشند، آن‌گاه*

$$E(y'\boldsymbol{A}y) = tr(\boldsymbol{A}\boldsymbol{V}) + \mu'\boldsymbol{A}\mu$$

برهان. سیرل[۵] (۱۹۷۱) را ببینید. □

[۴] Trenkler and Toutenburg

[۵] Searle

پیوست پ

شاخص‌های همخطی

در این پیوست بعضی از روش‌هایی که برای تشخیص همخطی به کار می‌روند، ارائه می‌شود. مطالب این پیوست عمدتاً برگرفته از روزبه (۱۳۹۰) و نجاریان (۱۳۹۰) می‌باشند.

پ.۱ آماره کاپا

ساده‌ترین راه برای پی بردن به وجود همخطی در ماتریس طرح، محاسبه آماره کاپا[۱] برای ماتریس $\mathbf{X'X}$ است. این آماره که به آن عدد شرطی[۲] می‌گویند، به صورت زیر با استفاده از مقادیر ویژه ماتریس $\mathbf{X'X}$ محاسبه می‌شود:

$$\kappa(\mathbf{X'X}) = \sqrt{\frac{max_{1\leq i\leq p}\lambda_i}{min_{1\leq i\leq p}\lambda_i}} \qquad (\text{پ.۱})$$

که در آن $\lambda_p , \ldots ,\lambda_1$ مقادیر ویژه ماتریس $\mathbf{X'X}$ هستند. بلسلی و همکارانش در سال ۱۹۸۰ پیشنهاد دادند که اگر آماره کاپا برای ماتریس $\mathbf{X'X}$ بزرگتر از ۱۰ باشد، آنگاه در ماتریس طرح همخطی وجود دارد. آنگاه گفتند اگر $100 < \kappa(\mathbf{X'X}) < 30$ باشد همخطی شدید و چنان‌چه $100 < \kappa(\mathbf{X'X})$ باشد، همخطی خیلی جدی است.

پ.۲ عامل تورم واریانس

یکی از شیوه‌های تشخیص وجود همخطی که کاربرد زیادی دارد، عامل تورم واریانس[۳] می‌باشد. این عامل نشان می‌دهد که واریانس ضریب برآورد شده تا چه حد نسبت به حالتی که متغیرهای برآورد شده

[۱] Kappa

[۲] Condition number

[۳] Variance inflation factor

همبستگی خطی ندارند، متورم شده است. عامل تورم واریانس از رابطه زیر به‌دست می‌آید:

$$VIF_i = \frac{۱}{۱ - R_i^۲} \quad i = ۱, ۲, \cdots, p - ۱$$

که در آن $R_i^۲$ ضریب تعیین مدلی است که در آن متغیر پیش‌بین X_i روی سایر متغیرهای پیش‌بین رگرس شده است.

در حالتی که $R_i^۲ = ۰$ یا به عبارتی دیگر X_i به صورت خطی رابطه‌ای با سایر متغیرهای پیش‌بین در مدل رگرسیونی نداشته باشد، VIF_i برابر با یک خواهد بود. هنگامی‌که $R_i^۲ \neq ۰$، VIF_i بزرگتر از یک خواهد بود که نشان می‌دهد، واریانس X_i به دلیل وجود همبستگی بین متغیرهای پیش‌بین چقدر متورم شده است و مقدار آن افزایش پیدا کرده است. گراس[۴] (۲۰۰۳) عامل تورم واریانس را به صورت زیر بیان نمود:

$$VIF_j = (x_j' x_j)/(x_j' M_j x_j)$$

که در آن

$$\begin{cases} M_j = I_n - X_j(X_j'X_j)^{-۱}X_j' \\ \mathbf{X} = (x_j, X_j). \end{cases}$$

و x_j مشاهدات مربوط به متغیر x_j و X_j ماتریس مشاهدات باقیمانده پس از حذف متغیر x_j است.

هنگامی‌که یک متغیر پیش‌بین، وابستگی خطی کامل با سایر متغیرهای پیش‌بین مدل داشته باشد، ضریب تعیین مرتبط با آن برابر با یک خواهد بود و در نتیجه عامل تورم واریانس به سمت بی‌نهایت میل خواهد کرد و از آن می‌توان نتیجه گرفت که واریانس X_i نیز بی‌نهایت خواهد بود.

مقدار عامل تورم واریانس برای متغیرهای پیش‌بین، اغلب به عنوان شاخصی برای اندازه‌گیری شدت همخطی بودن در مدل استفاده می‌شود. اگر بیشترین مقدار عامل تورم واریانس در مدل بزرگتر از ۱۰ باشد، معمولاً به نشانه این مسئله در نظر گرفته می‌شود که همخطی موجود در مدل، اثر نامطلوبی بر برآوردگرهای روش کمترین توان‌های دوم دارد.

[۴] Gross

پیوست ت

برنامه‌های کامپیوتری با نرم افزارR

در این پیوست تمام دستورهای مورد نیاز برای رسم شکل‌ها و جداول با نرم‌افزار R-۲.۱۵.۳ آمده است. دستور زیر مربوط به فراخوانی بسته‌های مورد نیاز برای انجام برخی دستورالعمل‌های ماتریسی است که برای تمام برنامه‌ها به طور یکسان تکرار می‌شود.

```
require(MASS)
require(MASS)
require(stats)
require(graphics)
require(Matrix)
library(matrixcalc)
require(psych)
require(expm)
```

ت.۱ برنامه مربوط به مثال ۲.۴.۱

برای رسم نمودار عملکرد برآوردگر نوع لیو به ازای مقادیر مختلف d در شکل ۱.۱ از دستورهای زیر در نرم‌افزار R استفاده می‌کنیم.

• ابتدا داده‌های مربوط به درصد تبدیلn - هپتان به استیلن را با نام $Heptan$ و پسوند csv به صورت زیر فراخوانی می‌کنیم. و برای آنها ماتریس ضریب همبستگی، آماره کاپا و عامل تورم واریانس را محاسبه می‌کنیم.

```
##########
#########Variables######
##########
Data<-read.table("G:/Heptan.csv",sep = ",",head=TRUE)
```

```
Y<-Data[,1]
X=as.matrix(Data[2:5])
cor(X,Y)
S=t(X)%*%X
lam=as.vector(eigen(S)$values)
LAM=diag(eigen(S)$values)
Q=eigen(S)$vectors
kappa=sqrt(max(lam)/min(lam))
n=16
VIF=numeric(4)
for(j in 1:4){
  M=diag(1,n,n)-X[,-j]%*%solve(t(X[,-j])%*%X[,-j])%*%t(X[,-j])
  VIF[j]=t(X[,j])%*%X[,j]/(t(X[,j])%*% M%*%X[,j])}
```

• مقادیر برآوردگر نوع لیو را به ازای $d = ۱$، $d = ۰/۷۵$، $d = ۰/۵۵$ با تعریف تابع زیر محاسبه می‌کنیم.

```
###########
###############function###############
###############
I=diag(1,4,4)
Z=X%*%Q
G.hat=solve(LAM)%*%t(Z)%*%Y
beta.hat=solve(t(Q))%*%G.hat
D=function(d) diag(d,4,4)
Fun.D<-function(d) solve(LAM+I)%*%(LAM+D(d))
G.hat.d<-function(d)
  Fun.D(d)%*%G.hat
beta.hat.d<-function(d)
  solve(t(Q))%*%G.hat.d(d)
beta.hat.d(0.55)
beta.hat.d(0.75)
beta.hat.d(1)
```

• برای محاسبه طول توان دوم، MSE، توان دوم اریبی، برآورد واریانس برآوردگرها از دستورهای زیر در نرم‌افزار R استفاده شده است.

```
#####################BIAS############
bias.d<-function(d)
  (I-Fun.D(d))%*%G.hat
########
#####################variance##########
###########
p=4
s2.hat=t(Y-(Z%*%G.hat))%*%(Y-(Z%*%G.hat))/(n-p)
as.numeric(s2.hat)->s2.hat
cov.d<-function(d)
  s2.hat*(Fun.D(d)%*%solve(LAM)%*%t(Fun.D(d)))
############
################MSEMs################
############
MSEM.d<-function(d){
  b1=bias.d(d)%*%t(bias.d(d))
  b1+cov.d(d)
}
####################SMSE#############
SMSE.d=function(d){
  return(matrix.trace(MSEM.d(d)))
}
SMSE.d(0.55)
SMSE.d(0.75)
SMSE.d(1)
```

• در نهایت شکل ۱.۱ را با دستورهای زیر در نرم‌افزار R رسم نموده‌ایم.

```
##############
#####################graph:Squared Bias#############
BpB=function(B)
  t(B)%*%B
BpB(beta.hat.d(0.55))
BpB(beta.hat.d(0.75))
BpB(beta.hat.d(1))
d=seq(0,1,0.01)
```

```
B1=numeric(length(d))
for(i in 1:length(d)){
  B1[i]=BpB(bias.d(d[i]))
}
plot(d,B1,type="l",lty=1,xlab="Liu(d)",
ylab="Estimated Squresd Bias of the estimators",lwd=1.5)
###################graph:Squared Lengths#############
d=seq(0,1,0.01)
R1=numeric(length(d))
for(i in 1:length(d)){
  R1[i]=BpB(beta.hat.d(d[i]))
}
plot(d,R1,type="l",lty=1,,xlab="Liu(d)",
ylab="Estimated Squared Lengths of the estimators",lwd=1.5)
#######################graph:Estimated SMSEs#########
d=seq(0,1,0.01)
S1=numeric(length(d))
for(i in 1:length(d)){
  S1[i]=SMSE.d(d[i])
}
plot(d,S1,type="l",lty=1,
ylab="Estimated SMSEs of the estimators",
xlab="Liu(d)",lwd=1.5,ylim=c(700,1300))
######################Variance##########
cov1=function(d){
  matrix.trace(s2.hat*(Fun.D(d)%*%solve(LAM)%*%t(Fun.D(d))))
}
##############graph:Estimated Variances###############
d=seq(0,1,0.01)
C1=numeric(length(d))
for(i in 1:length(d)){
  C1[i]=cov1(d[i])
}
plot(d,C1,type="l",lty=1,
```

```
ylab="Estimated Variances of the estimators",
xlab="Liu(d)",lwd=1.5,ylim=c(350,800))
```

ت.۲ برنامه مربوط به مثال ۱.۲.۲.

• مرحله فراخوانی داده‌های این مثال همانند بخش اول ت.۱ است. تنها تفاوت در مقدار برآوردگر است که با تعریف تابع زیر در نرم‌افزار R به ازای مقادیر $d = ۰/۵۵,\ d = ۰/۷۵,\ d = ۱$ محاسبه شده است.

```
###########
###############function##############
###############
D=function(d) diag(d,4,4)
Fun.D<-function(d) solve(LAM+I)%*%(LAM+D(d))
G.hat.d<-function(d)
  Fun.D(d)%*%G.hat
beta.hat.d<-function(d)
  solve(t(Q))%*%G.hat.d(d)
G.hat.J.d<-function(d){
  ((2*I)-Fun.D(d))%*%Fun.D(d)%*%G.hat
}
beta.hat.J.d<-function(d)
  solve(t(Q))%*%G.hat.J.d(d)
beta.hat.J.d(0.55)
beta.hat.J.d(0.75)
beta.hat.J.d(1)
#####################BIAS###########
bias.J.d<-function(d){
  Inv=solve(LAM+I)
  Inv2=Inv%*%Inv
  id2=(I-D(d))%*%(I-D(d))
  (-Inv2%*%id2)%*%G.hat
}
########
####################variance#########
###########
```

```
p=4
s2.hat=t(Y-(Z%*%G.hat))%*%(Y-(Z%*%G.hat))/(n-p)
as.numeric(s2.hat)->s2.hat
cov.J.d<-function(d)
  s2.hat*(2*I-Fun.D(d))%*%Fun.D(d)%*%
  solve(LAM)%*%t(Fun.D(d))%*%t(2*I-Fun.D(d))
#############
###############MSEMs################
############
MSEM.J.d<-function(d){
  b2=bias.J.d(d)%*%t(bias.J.d(d))
  b2+cov.J.d(d)
}
################SMSE################
SMSE.J.d=function(d){
  return(matrix.trace(MSEM.J.d(d)))
}
SMSE.J.d(0.55)
SMSE.J.d(0.75)
SMSE.J.d(1)
##############
######################graph:Squared Bias###############
BpB=function(B)
  t(B)%*%B
BpB(beta.hat.J.d(0.55))
BpB(beta.hat.J.d(0.75))
BpB(beta.hat.J.d(0.1))
#############
d=seq(0,1,0.01)
B1=numeric(length(d))
for(i in 1:length(d)){
  B1[i]=BpB(bias.J.d(d[i]))
}
plot(d,B1,type="l",lty=1,xlab="Liu(d)",
```

```
ylab="Estimated Squresd Bias of the estimators",lwd=1.5)
####################graph:Squared Lengths#############
d=seq(0,1,0.01)
R1=numeric(length(d))
for(i in 1:length(d)){
  R1[i]=BpB(beta.hat.J.d(d[i]))
}
plot(d,R1,type="l",lty=1,,xlab="Liu(d)",
ylab="Estimated Squared Lengths of the estimators",lwd=1.5)
#################graph:Estimated SMSEs################
d=seq(0,1,0.01)
S1=numeric(length(d))
for(i in 1:length(d)){
  S1[i]=SMSE.J.d(d[i])
}
plot(d,S1,type="l",lty=1,
ylab="Estimated SMSEs of the estimators",
xlab="Liu(d)",lwd=1.5,ylim=c(700,1300))
################Variance###############
cov1=function(d){
  matrix.trace(s2.hat*(2*diag(1,4,4)-Fun.D(d))%*%Fun.D(d)
  %*%solve(LAM)%*%t(Fun.D(d))%*%(2*diag(1,4,4)-Fun.D(d)))
}
###############graph:Estimated Variances################
d=seq(0,1,0.01)
C1=numeric(length(d))
for(i in 1:length(d)){
  C1[i]=cov1(d[i])
}
plot(d,C1,type="l",lty=1,
ylab="Estimated Variances of the estimators",
xlab="Liu(d)",lwd=1.5,ylim=c(350,800))
```

ت.۳ برنامه مربوط به مثال ۱.۴.۲

• برای رسم نمودار عملکرد برآوردگر جک‌نایف اصلاح شده لیو به ازای مقادیر مختلف d در شکل ۲.۲ از دستورهای زیر در نرم‌افزار R استفاده می‌کنیم.

```
###########
##############function##############
###############
D=function(d) diag(d,4,4)
Fun.D<-function(d) solve(LAM+I)%*%(LAM+D(d))
G.hat.d<-function(d)
  Fun.D(d)%*%G.hat
beta.hat.d<-function(d)
  solve(t(Q))%*%G.hat.d(d)
G.hat.M.J.d<-function(d){
  Inv=solve(LAM+I)
  Inv2=Inv%*%Inv
  id2=(I-D(d))%*%(I-D(d))
  (I-Inv2%*%id2)%*%G.hat.d(d)
}
beta.hat.M.J.d<-function(d)
  solve(t(Q))%*%G.hat.M.J.d(d)
beta.hat.M.J.d(0.55)
beta.hat.M.J.d(0.75)
beta.hat.M.J.d(1)
####################BIAS############
bias.M.J.d<-function(d){
  W=I+Fun.D(d)-(Fun.D(d)%*%Fun.D(d))
  (-1)*solve(LAM+I)%*%W%*%(I-D(d))%*%G.hat
}
########
####################variance##########
###########
p=4
```

```
s2.hat=t(Y-(Z%*%G.hat))%*%(Y-(Z%*%G.hat))/(n-p)
as.numeric(s2.hat)->s2.hat
cov.M.J.d<-function(d){
  phi=(2*I-Fun.D(d))%*%(Fun.D(d)%*%Fun.D(d))
  s2.hat*(phi%*%solve(LAM)%*%t(phi))
}
#############
################MSEMs################
############
bias.d<-function(d)
  (I-Fun.D(d))%*%G.hat
MSEM.M.J.d<-function(d){
  b3=bias.d(d)%*%t(bias.d(d))
  b3+cov.M.J.d(d)
}
###################SMSE################
SMSE.M.J.d=function(d){
  return(matrix.trace(MSEM.M.J.d(d)))
}
SMSE.M.J.d(0.55)
SMSE.M.J.d(0.75)
SMSE.M.J.d(1)
##############
#####################graph:Squared Bias###############
BpB=function(B)
  t(B)%*%B
BpB(beta.hat.M.J.d(0.55))
BpB(beta.hat.M.J.d(0.75))
BpB(beta.hat.M.J.d(1))
#############
d=seq(0,1,0.01)
B1=numeric(length(d))
for(i in 1:length(d)){
  B1[i]=BpB(bias.M.J.d(d[i]))
```

```
}
plot(d,B1,type="l",lty=1,xlab="Liu(d)",
ylab="Estimated Squresd Bias of the estimators",lwd=1.5)
###################graph:Squared Lengths#############
d=seq(0,1,0.01)
R1=numeric(length(d))
for(i in 1:length(d)){
  R1[i]=BpB(beta.hat.M.J.d(d[i]))
}
plot(d,R1,type="l",lty=1,,xlab="Liu(d)",
ylab="Estimated Squared Lengths of the estimators",lwd=1.5)
################graph:Estimated SMSEs################
d=seq(0,1,0.01)
S1=numeric(length(d))
for(i in 1:length(d)){
  S1[i]=SMSE.M.J.d(d[i])
}
plot(d,S1,type="l",lty=1,
ylab="Estimated SMSEs of the estimators",
xlab="Liu(d)",lwd=1.5,ylim=c(700,1300))
#####################Variance#############
cov1=function(d){
  phi=(2*diag(1,4,4)-Fun.D(d))%*%Fun.D(d)%*%Fun.D(d)
  matrix.trace(s2.hat*phi%*%solve(LAM)%*%t(phi))
}
#####################graph:Estimated Variances#########
d=seq(0,1,0.01)
C1=numeric(length(d))

for(i in 1:length(d)){
  C1[i]=cov1(d[i])
}
plot(d,C1,type="l",lty=1,
ylab="Estimated Variances of the estimators",
```

```
xlab="Liu(d)",lwd=1.5,ylim=c(350,800))
```

ت.۴ برنامه مربوط به مثال ۶.۲

برای محاسبه مقادیر به‌دست آمده در جداول ۳.۲ و ۴.۲ از دستورهای زیر در نرم افزار R استفاده شده است.

```
##########
##########Variables#############
##########
Data<-read.table("G:/Yatchew.csv",
sep = ",",head=TRUE)
Y<-Data[,1]
X=as.matrix(Data[2:10])
cor(X,Y)
S=t(X)%*%X
lam=as.vector(eigen(S)$values)
LAM=diag(eigen(S)$values)
Q=eigen(S)$vectors
kappa=sqrt(max(lam)/min(lam))
n=92
VIF=numeric(9)
for(j in 1:9){
  M=diag(1,n,n)-X[,-j]%*%solve(t(X[,-j])%*%X[,-j])%*%t(X[,-j])
  VIF[j]=t(X[,j])%*%X[,j]/(t(X[,j])%*% M%*%X[,j])}
I=diag(1,9,9)
Z=X%*%Q
G.hat=solve(LAM)%*%t(Z)%*%Y
beta.hat=solve(t(Q))%*%G.hat
D=function(d) diag(d,9,9)
Fun.D<-function(d) solve(LAM+I)%*%(LAM+D(d))
G.hat.d<-function(d)
  Fun.D(d)%*%G.hat
beta.hat.d<-function(d)
  solve(t(Q))%*%G.hat.d(d)
```

```
beta.hat.d(0.75)
G.hat.J.d<-function(d){
  ((2*I)-Fun.D(d))%*%Fun.D(d)%*%G.hat
}
beta.hat.J.d<-function(d)
  solve(t(Q))%*%G.hat.J.d(d)
G.hat.M.J.d<-function(d){
  Inv=solve(LAM+I)
  Inv2=Inv%*%Inv
  id2=(I-D(d))%*%(I-D(d))
  (I-Inv2%*%id2)%*%G.hat.d(d)
}
beta.hat.M.J.d<-function(d)
  solve(t(Q))%*%G.hat.M.J.d(d)
beta.hat.d(0.55)
beta.hat.J.d(0.55)
beta.hat.M.J.d(0.55)

beta.hat.d(0.75)
beta.hat.J.d(0.75)
beta.hat.M.J.d(0.75)

beta.hat.d(1)
beta.hat.J.d(1)
beta.hat.M.J.d(1)
##############Norm###############
BpB=function(B)
  t(B)%*%B
BpB(beta.hat.d(0.55))
BpB(beta.hat.J.d(0.55))
BpB(beta.hat.M.J.d(0.55))

BpB(beta.hat.d(0.75))
BpB(beta.hat.J.d(0.75))
```

```
BpB(beta.hat.M.J.d(0.75))

BpB(beta.hat)
############BIAS#################
bias.d<-function(d)
  (I-Fun.D(d))%*%G.hat
bias.J.d<-function(d){
  Inv=solve(LAM+I)
  Inv2=Inv%*%Inv
  id2=(I-D(d))%*%(I-D(d))
  (-Inv2%*%id2)%*%G.hat
  }
bias.J.d(0.55)

bias.M.J.d<-function(d){
  W=I+Fun.D(d)-(Fun.D(d)%*%Fun.D(d))
  (-1)*solve(LAM+I)%*%W%*%(I-D(d))%*%G.hat
}
bias.M.J.d(0.55)
############MSEM###############
p=9
s2.hat=t(Y-(Z%*%G.hat))%*%(Y-(Z%*%G.hat))/(n-p)
as.numeric(s2.hat)->s2.hat
cov.d<-function(d)
  s2.hat*(Fun.D(d)%*%solve(LAM)%*%t(Fun.D(d)))
cov.J.d<-function(d)
  s2.hat*(2*I-Fun.D(d))%*%Fun.D(d)
 %*%solve(LAM)%*%t(Fun.D(d))%*%t(2*I-Fun.D(d))
cov.M.J.d<-function(d){
  phi=(2*I-Fun.D(d))%*%(Fun.D(d)%*%Fun.D(d))
  s2.hat*(phi%*%solve(LAM)%*%t(phi))
}
MSEM.d<-function(d){
  b1=bias.d(d)%*%t(bias.d(d))
```

```
  b1+cov.d(d)
  }
MSEM.J.d<-function(d){
  b2=bias.J.d(d)%*%t(bias.J.d(d))
  b2+cov.J.d(d)
}
MSEM.M.J.d<-function(d){
  b3=bias.d(d)%*%t(bias.d(d))
  b3+cov.M.J.d(d)
}
############SMSE################
SMSE.d=function(d){
  return(matrix.trace(MSEM.d(d)))
}
SMSE.J.d=function(d){
  return(matrix.trace(MSEM.J.d(d)))
}
SMSE.M.J.d=function(d){
  return(matrix.trace(MSEM.M.J.d(d)))
}
SMSE.d(0.55)
SMSE.J.d(0.55)
SMSE.M.J.d(0.55)

SMSE.d(0.75)
SMSE.J.d(0.75)
SMSE.M.J.d(0.75)
##########
##########################graphics############
##########Estimated Squresd Bias####
d=seq(0,1,0.01)
B1=numeric(length(d))
B2=numeric(length(d))
B3=numeric(length(d))
```

```
for(i in 1:length(d)){
  B1[i]=BpB(bias.d(d[i]))
  B2[i]=BpB(bias.J.d(d[i]))
  B3[i]=BpB(bias.M.J.d(d[i]))
}
plot(d,B1,type="l",lty=1,xlab="Liu(d)",
ylab="Estimated Squresd Bias of the estimators",lwd=1.5)
lines(d,B2,lty=2,lwd=1.5)
lines(d,B3,type="l",lty=3,lwd=1.5)
Bx=c("Liu estimaror","Jackknifed Liu","Modified Jackknifed Liu")
legend("topright",Bx,lty=1:3,cex=0.6,
text.width = strwidth("1,000,000"))
#############Squared Lengths#####
d=seq(0,1,0.01)
R1=numeric(length(d))
R2=numeric(length(d))
R3=numeric(length(d))
for(i in 1:length(d)){
  R1[i]=BpB(beta.hat.d(d[i]))
  R2[i]=BpB(beta.hat.J.d(d[i]))
  R3[i]=BpB(beta.hat.M.J.d(d[i]))
}
plot(d,R1,type="l",lty=1,,xlab="Liu(d)",
ylab="Estimated Squared Lengths of the estimators",lwd=1.5)
lines(d,R2,lty=2,lwd=1.5)
lines(d,R3,lty=3,lwd=1.5)
Rx=c("Liu estimaror","Jackknifed Liu","Modified Jackknifed Liu")
legend("bottomright",Rx,lty=1:3,cex=0.6,
 text.width = strwidth("100000,000,000"))
#############Estimated SMSEs####
d=seq(0,1,0.01)
S1=numeric(length(d))
S2=numeric(length(d))
S3=numeric(length(d))
```

```
for(i in 1:length(d)){
  S1[i]=SMSE.d(d[i])
  S2[i]=SMSE.J.d(d[i])
  S3[i]=SMSE.M.J.d(d[i])
}
plot(d,S1,type="l",lty=1,ylab="Estimated SMSEs of the estimators",
xlab="Liu(d)",lwd=1.5,ylim=c(700,1300))
lines(d,S2,lty=2,lwd=1.5)
lines(d,S3,lty=3,lwd=1.5)
Cx=c("Liu estimaror","Jackknifed Liu","Modified Jackknifed Liu")
legend("topright",Cx,lty=1:3,cex=0.6,
 text.width = strwidth("100000,000,000"))
#################Variance####
cov1=function(d){
  matrix.trace(s2.hat*(Fun.D(d)%*%solve(LAM)%*%t(Fun.D(d))))
}
cov2=function(d){
  matrix.trace(s2.hat*(2*diag(1,9,9)-Fun.D(d))
  %*%Fun.D(d)%*%solve(LAM)%*% t(Fun.D(d))
  %*%(2*diag(1,9,9)-Fun.D(d)))
}
cov3=function(d){
  phi=(2*diag(1,9,9)-Fun.D(d))%*%Fun.D(d)%*%Fun.D(d)
  matrix.trace(s2.hat*phi%*%solve(LAM)%*%t(phi))
}
###########Estimated Variances####
d=seq(0,1,0.01)
C1=numeric(length(d))
C2=numeric(length(d))
C3=numeric(length(d))
for(i in 1:length(d))
{
  C1[i]=cov1(d[i])
  C2[i]=cov2(d[i])
```

```
  C3[i]=cov3(d[i])
}
plot(d,C1,type="l",lty=1,ylab="Estimated Variances of
 the estimators",xlab="Liu(d)",lwd=1.5,ylim=c(350,800))
lines(d,C2,lty=2,lwd=1.5)
lines(d,C3,type="l",lty=3,lwd=1.5)
Cx=c("Liu estimaror","Jackknifed Liu","Modified Jackknifed Liu")
legend("bottomright",Cx,lty=1:3,cex=0.6,
text.width = strwidth("100000,000,000"))
```

ت.۵ برنامه مربوط به بخش ۷.۲

برای رسم نمودار ۳.۲ و محاسبه مقادیر به‌دست آمده در جداول ۷.۲ و ۸.۲ از دستورهای زیر در نرم افزار R استفاده شده است.

```
Main<-function(n,s,ro){
  Ip=diag(1,p,p)
  In=diag(1,n,n)
    W<-matrix(rnorm(n*p,0,1),nrow=n,ncol=p)
    X<-matrix(0,n,p)
      for(i in 1:n){
        for(j in 1:p)
          X[i,j]=sqrt(1-ro^2)*W[i,j]+ro*W[i,p]   }

  beta<-matrix(c(1,0,1),nc=1,nr=p)
  e<-matrix(rnorm(n,0,s^2),nr=n,nc=1)
  Y<-X%*%beta+e

  S=t(X)%*%X
  lam=as.vector(eigen(S)$values)
  LAM=diag(eigen(S)$values)
  Q=eigen(S)$vectors
  kappa=sqrt(max(lam)/min(lam))
  I=diag(1,p,p)
  D=function(d) diag(d,p,p)
```

```
  Fun.D<-function(d) solve(LAM+I)%*%(LAM+D(d))
###################OLS
#########
  Z=X%*%Q
  G.hat=solve(LAM)%*%t(Z)%*%Y
  beta.hat=solve(t(Q))%*%G.hat
##############Liu
###########
  G.hat.d<-function(d)
    Fun.D(d)%*%G.hat
  beta.hat.d<-function(d)
    solve(t(Q))%*%G.hat.d(d)
##############Jack.Liu
###########
  G.hat.J.d<-function(d)
    ((2*I)-Fun.D(d))%*%Fun.D(d)%*%G.hat
  beta.hat.J.d<-function(d)
     solve(t(Q))%*%G.hat.J.d(d)
##############Modified.Jackk.Liu
###########
  G.hat.M.J.d<-function(d){
    Inv=solve(LAM+I)
    Inv2=Inv%*%Inv
    id2=(I-D(d))%*%(I-D(d))
    (I-Inv2%*%id2)%*%G.hat.d(d)
  }
  beta.hat.M.J.d<-function(d)
     solve(t(Q))%*%G.hat.M.J.d(d)
##############d_min
########
  sum1=sum2=0

  for(i in 1:p){
     sum1=sum1+(1/(lam[i]*(1+lam[i])))
```

```
    sum2=sum2+((G.hat[i])^2/(1+lam[i])^2)}

  d.min=1-((s^2)*(sum1/sum2))

  result=list(beta.ols=beta.hat,beta.hat.d=beta.hat.d(d.min),
  beta.hat.J.d=beta.hat.J.d(d.min),
  beta.hat.M.J.d=beta.hat.M.J.d(d.min),kappa=kappa)
  return(result)
}

beta.N=matrix(0,nr=p,nc=N)
beta.d.N=matrix(0,nr=p,nc=N)
beta.J.d.N=matrix(0,nr=p,nc=N)
beta.M.J.d.N=matrix(0,nr=p,nc=N)
kappa=matrix(0,nr=1,nc=N)

SMSE<-function(b.h){
  sum=0
  for(i in 1:N)
    sum=sum+t(b.h[,i]-beta)%*%(b.h[,i]-beta)
  (1/N)*sum
}

ABIAS<-function(b.h){
  sum=0
  for(i in 1:N)
    sum=sum+sum(abs(b.h[,i]-beta))
  (1/N)*sum
}

for(i in 1:N){
  M=Main(15,0.1,0.99)
  beta.N[,i]<-M$beta.ols
  beta.d.N[,i]<-M$beta.hat.d
```

```
  beta.J.d.N[,i]<-M$beta.hat.J.d
  beta.M.J.d.N[,i]<-M$beta.hat.M.J.d
  kappa[,i]=M$kappa
  }

SMSE(beta.N)
SMSE(beta.d.N)
SMSE(beta.J.d.N)
SMSE(beta.M.J.d.N)

ABIAS(beta.N)
ABIAS(beta.d.N)
ABIAS(beta.J.d.N)
ABIAS(beta.M.J.d.N)
mean(kappa)
```

ت.۶ برنامه مربوط به مثال ۴.۳

برای محاسبه مقادیر به‌دست آمده در جدول ۱.۳ از دستورهای زیر در نرم افزار R استفاده شده است.

```
##########
##########Variables######
##########
Data<-read.table("G:/yang.csv",sep = ",",head=TRUE)
Y<-Data[,1]
X=as.matrix(Data[2:6])
S=t(X)%*%X
(lam=as.vector(eigen(S)$values))
LAM=diag(eigen(S)$values)
T=eigen(S)$vectors
kappa=sqrt(max(lam)/min(lam))
n=13
VIF=numeric(5)
for(j in 1:5){
  M=diag(1,n,n)-X[,-j]%*%solve(t(X[,-j])%*%X[,-j])%*%t(X[,-j])
```

```
  VIF[j]=t(X[,j])%*%X[,j]/(t(X[,j])%*% M%*%X[,j])}
Z=X%*%T
Alfa.hat=solve(LAM)%*%t(Z)%*%Y
(beta.hat=solve(t(T))%*%Alfa.hat)
###########
###############function###############
###############
Alfa.d<-function(d)
  (diag(1,5,5)-(1-d)*solve(LAM+diag(1,5,5)))%*%Alfa.hat
beta.hat.d=function(d)
  solve(t(Q))%*%Alfa.d(d)
Alfa.k<-function(k)
  solve(LAM+k*diag(1,5,5))%*%t(Z)%*%Y
beta.hat.k=function(k)
  solve(t(Q))%*%Alfa.k(k)
Alfa.k.b<-function(k,b)
  solve(LAM+k*diag(1,5,5))%*%(t(Z)%*%Y+k*b)
beta.hat.k.b=function(k,b)
  solve(t(Q))%*%Alfa.k.b(k,b)
Alfa.d.b<-function(d,b)
  Alfa.d(d)+solve(LAM+diag(1,5,5))%*%((1-d)*b)
beta.hat.d.b=function(d,b)
  solve(t(Q))%*%Alfa.d.b(d,b)
########
#####################variance############
###########
p=5
(s2.hat=t(Y-(Z%*%Alfa.hat))%*%(Y-(Z%*%Alfa.hat))/(n-p))
(as.numeric(s2.hat)->s2.hat)
S2=s2.hat*solve(t(Z)%*%Z)
m2=matrix.trace(S2)
Ad=function(d)
  solve(LAM+diag(1,5,5))%*%(LAM+d*diag(1,5,5))
Ak=function(k)
```

```
  solve(LAM+k*diag(1,5,5))
Aktild=function(k)
  solve(LAM+k*diag(1,5,5))%*%LAM
#############
###############MSEMs###############
############
MSEM.d=function(d){
  cov=s2.hat*(Ad(d)%*%solve(LAM)%*%t(Ad(d)))
  bias=beta.hat.d(d)-beta.hat
  return(t(bias%*%t(bias)+cov))
}
MSEM.k=function(k){
  cov=s2.hat*(Ak(k)%*%LAM%*%t(Ak(k)))
  bias=beta.hat.k(k)-beta.hat
  return(t(bias%*%t(bias)+cov))
}
MSEM.k.b=function(k,b){
  cov=s2.hat*(Aktild(k)%*%solve(LAM)%*%t(Aktild(k)))
  bias=beta.hat.k.b(k,b)-beta.hat
  return(t(bias%*%t(bias)+cov))
}
MSEM.d.b=function(d,b){
  cov=s2.hat*(Ad(d)%*%solve(LAM)%*%t(Ad(d)))
  bias=beta.hat.d.b(d,b)-beta.hat
  return(t(bias%*%t(bias)+cov))
}
##############
b=0.95*Alfa.hat
SMSE.d=function(d){
  return(matrix.trace(MSEM.d(d)))
}
SMSE.d.b=function(d,b){
  return(matrix.trace(MSEM.d.b(d,b)))
}
```

```
SMSE.k=function(k){
  return(matrix.trace(MSEM.k(k)))
}
SMSE.k.b=function(k,b){
  return(matrix.trace(MSEM.k.b(k,b)))
}
SMSE.d(0)
SMSE.d(0.01)
SMSE.d(0.02)
SMSE.d(0.03)
SMSE.d(0.04)
SMSE.d(0.05)
SMSE.d(0.09)
SMSE.d(0.1)

SMSE.d.b(0,b)
SMSE.d.b(0.01,b)
SMSE.d.b(0.02,b)
SMSE.d.b(0.03,b)
SMSE.d.b(0.04,b)
SMSE.d.b(0.05,b)
SMSE.d.b(0.09,b)
SMSE.d.b(0.1,b)

SMSE.k(0)
SMSE.k(0.01)
SMSE.k(0.02)
SMSE.k(0.03)
SMSE.k(0.04)
SMSE.k(0.05)
SMSE.k(0.09)
SMSE.k(0.1)

SMSE.k.b(0,b)
```

```
SMSE.k.b(0.01,b)
SMSE.k.b(0.02,b)
SMSE.k.b(0.03,b)
SMSE.k.b(0.04,b)
SMSE.k.b(0.05,b)
SMSE.k.b(0.09,b)
SMSE.k.b(0.1,b)
```

برای رسم شکل‌های ۱.۳ و ۲.۳ از دستور زیر در نرم افزار R استفاده شده است.

```
s2=m2
b=0.95*Alfa.hat
k=seq(0,0.1,0.01)
R1=numeric(length(k))
R2=numeric(length(k))
R3=numeric(length(k))
for(i in 1:length(k)){
  R1[i]=SMSE.k.b(k[i],b)
  R2[i]=SMSE.k(k[i])
  R3[i]=s2
}
plot(k,R1,type="l",lty=1,main="Estimated MSEMs of the estimators"
,xlab="k",ylab="MSEM",lwd=1.5)
lines(k,R2,lty=2,lwd=1.5)
lines(k,R3,lty=3,lwd=1.5)
Rx=c("Modified Ridge","Ridge","Least Square")
legend(0.075,3000,
       Rx,lty=1:3,cex=0.6, text.width = strwidth("1,000,000"))
$$$$$$$$$$$$$$$$$
s2=m2
b=0.95*Alfa.hat
d=seq(0,0.1,0.01)
C1=numeric(length(d))
C2=numeric(length(d))
C3=numeric(length(d))
for(i in 1:length(d)){
```

```
  C1[i]=SMSE.d.b(d[i],b)
  C2[i]=SMSE.d(d[i])
  C3[i]=s2
}
plot(d,C1,type="l",lty=1,main="Estimated MSEMs of the estimators",
xlab="d",ylab="MSEM",lwd=1.5,ylim=c(0,5000))
lines(d,C2,lty=2,lwd=1.5)
lines(d,C3,lty=3,lwd=1.5)
Cx=c("Modified Liu","Liu","Least Square")
legend(0.075,2500,
       Cx,lty=1:3,cex=0.6, text.width = strwidth("1,000,000"))
}
```

ت.۷ برنامه مربوط به بخش ۵.۳

برای محاسبه مقادیر به‌دست آمده در جداول ۲.۳ ، ۳.۳ و ۴.۳ از دستورهای زیر در نرم افزار R استفاده شده است.

```
##########Variables######
##########
N=100
p=4
n=100
beta.N=matrix(0,nr=p,nc=N)
beta.LE.N=matrix(0,nr=p,nc=N)
beta.RE.N=matrix(0,nr=p,nc=N)
beta.MRE.N=matrix(0,nr=p,nc=N)
beta.MLE.N=matrix(0,nr=p,nc=N)
Beta<-matrix(c(0,0,0,0),nc=1,nr=p)
gamma=0.7
s=0.01
DATA=list()
Y=matrix(0,nr=n,nc=N)
for(k in 1:N){
  Ip=diag(1,p,p)
  In=diag(1,n,n)
```

```
  W<-matrix(rnorm(n*p,0,1),nrow=n,ncol=p)
  X<-matrix(0,n,p)
  for(i in 1:n){
    for(j in 1:p)
      X[i,j]=sqrt(1-gamma^2)*W[i,j]+gamma*W[i,p]   }
  DATA[[k]]=X
  e<-matrix(rnorm(n,0,s^2),nr=n,nc=1)
  Y[,k]<-DATA[[k]]%*%Beta+e
}
Main<-function(s,gamma,X,Y,k,d){
  S=t(X)%*%X
  lam=as.vector(eigen(S)$values)
  LAM=diag(lam)
  Q=eigen(S)$vectors
  Z=X%*%Q
  #################LSE
  ####
  g.hat<-solve(t(Z)%*%Z)%*%t(Z)%*%Y
  b.h<- solve(t(Q))%*%g.hat
  b=0.000001*g.hat
  I=function(d) diag(d,p,p)
  ##################LE
  ####
  g.hat.LE<-function(d)
    solve(LAM+Ip)%*%(LAM+d*Ip)%*%g.hat
  b.h.LE<-function(d)
    solve(t(Q))%*%g.hat.LE(d)
  ##################RE
  ####
  g.hat.RE<-function(k)
    solve(LAM+k*Ip)%*%t(Z)%*%Y
  b.h.RE<-function(k)
    solve(t(Q))%*%g.hat.RE(k)
  ##################MRE
```

```
  ####
  g.hat.MRE<-function(b,k)
    solve(LAM+k*Ip)%*%((t(Z)%*%Y)+k*b)
  b.h.MRE<-function(b,k)
    solve(t(Q))%*%g.hat.MRE(b,k)
  ##################MLE
  #####
  g.hat.MLE<-function(d,b)
    solve(LAM+Ip)%*%((LAM+d)%*%solve(LAM)%*%t(Z)%*%Y+(1-d)*b)
  b.h.MLE<-function(d,b)
    solve(t(Q))%*% g.hat.MLE(d,b)

  result=list(beta.ols=b.h,beta.hat.LE=b.h.LE(d),
              beta.hat.RE=b.h.RE(k),beta.hat.MRE=b.h.MRE(b,k),
              beta.hat.MLE=b.h.MLE(d,b))
  return(result)
}
beta.N=matrix(0,nr=p,nc=N)
beta.LE.N=matrix(0,nr=p,nc=N)
beta.RE.N=matrix(0,nr=p,nc=N)
beta.MRE.N=matrix(0,nr=p,nc=N)
beta.MLE.N=matrix(0,nr=p,nc=N)
d=0
k=0.05
for(i in 1:N){
  M=Main(s,gamma,DATA[[i]],Y[,i],k,d)
  beta.N[,i]<-M$beta.ols
  beta.LE.N[,i]<-M$beta.hat.LE
  beta.RE.N[,i]<-M$beta.hat.RE
  beta.MRE.N[,i]<-M$beta.hat.MRE
  beta.MLE.N[,i]<-M$beta.hat.MLE
}
mse.N=numeric(N)
mse.LE.N=numeric(N)
```

```
mse.RE.N=numeric(N)
mse.MRE.N=numeric(N)
mse.MLE.N=numeric(N)
for(i in 1:N){
  mse.N[i]=t(beta.N[,i]-Beta)%*%(beta.N[,i]-Beta)
  mse.LE.N[i]=t(beta.LE.N[,i]-Beta)%*%(beta.LE.N[,i]-Beta)
  mse.RE.N[i]=t(beta.RE.N[,i]-Beta)%*%(beta.RE.N[,i]-Beta)
  mse.MRE.N[i]=t(beta.MRE.N[,i]-Beta)%*%(beta.MRE.N[,i]-Beta)
  mse.MLE.N[i]=t(beta.MLE.N[,i]-Beta)%*%(beta.MLE.N[,i]-Beta)
}
(MSE.beta.N=mean(mse.N))
(MSE.beta.RE.N=mean(mse.RE.N))
(MSE.beta.MRE.N=mean(mse.MRE.N))
(MSE.beta.N=mean(mse.N))
(MSE.beta.LE.N=mean(mse.LE.N))
(MSE.beta.MLE.N=mean(mse.MLE.N))
```

ت.۸ برنامه مربوط به بخش ۵.۴

```
##########Variables######
##########
N=100
p=4
n=30
gamma=0.99
s=1
########## Estimation of parameter for N=100
beta.N=matrix(0,nr=p,nc=N)
beta.LE.N=matrix(0,nr=p,nc=N)
beta.MLE.N=matrix(0,nr=p,nc=N)
beta.JMLE.N=matrix(0,nr=p,nc=N)
Beta<-matrix(c(0,0,0,0),nc=1,nr=p)
########### Matrix X & Y for N simulation
DATA=list()
```

```
Y=matrix(0,nr=n,nc=N)
for(k in 1:N){
  Ip=diag(1,p,p)
  In=diag(1,n,n)
  W<-matrix(rnorm(n*p,0,1),nrow=n,ncol=p)
  X<-matrix(0,n,p)
  for(i in 1:n){
    for(j in 1:p)
      X[i,j]=sqrt(1-gamma^2)*W[i,j]+gamma*W[i,p]   }
  DATA[[k]]=X
  e<-matrix(rnorm(n,0,s^2),nr=n,nc=1)
  Y[,k]<-DATA[[k]]%*%Beta+e
}
#### Function for different value of sigma=s,rho=gamma,X,Y,biasing parameters k,d
Main<-function(s,gamma,X,Y,k,d){
S=t(X)%*%X
lam=as.vector(eigen(S)$values)
LAM=diag(lam)
Q=eigen(S)$vectors
Z=X%*%Q
################## Least Square Error
g.hat=solve(LAM)%*%t(Z)%*%Y
b.h=Q%*%g.hat
b=0.000001*g.hat
################## Liu Etimator
####
g.hat.LE<-function(d)
  solve(LAM+Ip)%*%(LAM+d*Ip)%*%g.hat
b.h.LE<-function(d)
 solve(t(Q))%*%g.hat.LE(d)
################# Modified Liu Estimator
g.hat.MLE<-function(d,b)
  solve(LAM+Ip)%*%((LAM+d)%*%solve(LAM)%*%t(Z)%*%Y+(1-d)*b)
b.h.MLE<-function(d,b)
```

```
  solve(t(Q))%*% g.hat.MLE(d,b)
################ Jackknifed Modified Liu Estimator
(A=solve(LAM+Ip)%*%(LAM+d*Ip)%*%solve(LAM))
(B=(1-d)*solve(LAM+Ip))
Fun.d<-function(d) solve(LAM+Ip)%*%(LAM+d*Ip)
d1<-function(i)
  as.numeric(solve(t(Z[i,])%*%solve(B)%*%Z[i,])^(1/2))
T.mat=matrix(0,nr=n,nc=p)
for(i in 1:n)
  T.mat[i,]=t(Z[i,])*d1(i)
d2=function(i) as.numeric(solve(t(Z[i,])%*%solve(A)%*%Z[i,])^(1/2))
C.mat=matrix(0,nr=n,nc=p)
for(i in 1:n)
  C.mat[i,]=t(Z[i,])*d2(i)
g.hat.JMLE<-function(d,b)
  (A%*%t(T.mat)%*%T.mat+Ip-Fun.d(d))%*%g.hat.MLE(d,b)+
  solve(B)%*%(t(C.mat)%*%C.mat-LAM)%*%solve(B)%*%b
b.h.JMLE<-function(d,b)
  solve(t(Q))%*% g.hat.JMLE(d,b)
b.h.JMLE(d,b)
result=list(beta.ols=b.h,beta.hat.LE=b.h.LE(d),
            beta.hat.MLE=b.h.MLE(d,b),beta.hat.JMLE=b.h.JMLE(d,b))
return(result)
}
################## N simulation
beta.N=matrix(0,nr=p,nc=N)
beta.LE.N=matrix(0,nr=p,nc=N)
beta.MLE.N=matrix(0,nr=p,nc=N)
beta.JMLE.N=matrix(0,nr=p,nc=N)
k=0.1
d=0.0001
for(i in 1:N){
  M=Main(s,gamma,DATA[[i]],Y[,i],k,d)
  beta.N[,i]<-M$beta.ols
```

```
  beta.LE.N[,i]<-M$beta.hat.LE
  beta.MLE.N[,i]<-M$beta.hat.MLE
  beta.JMLE.N[,i]<-M$beta.hat.JMLE
}
############ Mean Square Error for N simulation
mse.N=numeric(N)
mse.LE.N=numeric(N)
mse.MLE.N=numeric(N)
mse.JMLE.N=numeric(N)
for(i in 1:N){
  mse.N[i]=t(beta.N[,i]-Beta)%*%(beta.N[,i]-Beta)
  mse.LE.N[i]=t(beta.LE.N[,i]-Beta)%*%(beta.LE.N[,i]-Beta)
  mse.MLE.N[i]=t(beta.MLE.N[,i]-Beta)%*%(beta.MLE.N[,i]-Beta)
  mse.JMLE.N[i]=t(beta.JMLE.N[,i]-Beta)%*%(beta.JMLE.N[,i]-Beta)
}
(MSE.beta.N=mean(mse.N))
(MSE.beta.LE.N=mean(mse.LE.N))
(MSE.beta.MLE.N=mean(mse.MLE.N))
(MSE.beta.JMLE.N=mean(mse.JMLE.N))
```

مراجع

[۱] حسن زاده بشتیان م، (۱۳۸۸). "برآوردگرهای انقباضی در مدل رگرسیون ریج با خطاهای دارای توزیع بیضی گون"، پایان‌نامه کارشناسی ارشد، دانشگاه فردوسی مشهد، مشهد.

[۲] روزبه م، (۱۳۹۰). "برآورد در مدل‌های خطی جزئی"، پایان‌نامه دکتری، دانشگاه فردوسی مشهد، مشهد.

[۳] غیورمرادی ز، (۱۳۸۷). "روش‌های بوت استرپ و جک‌نایف در رگرسیون خطی چندگانه"، پایان‌نامه کارشناسی ارشد، دانشگاه آزاد اسلامی واحد مشهد، مشهد.

[۴] نجاریان س، (۱۳۹۰). "بررسی رفتار برآوردگر ریج در مدل‌های خطی منفرد"، پایان‌نامه کارشناسی ارشد، دانشگاه صنعتی شاهرود، شاهرود.

[۵] نوروزی راد م، (۱۳۹۰). "مباحثی در برآورد تابع زیان"، پایان‌نامه کارشناسی ارشد، دانشگاه آزاد اسلامی واحد مشهد، مشهد.

[۶] نیرومند ح. ع، (۱۳۸۷). "تحلیل رگرسیون خطی ابزاری برای تحقیق"، انتشارات دانشگاه فردوسی مشهد، مشهد.

[7] Akdeniz F. and Kaciranlar S. (1995), "On the olmost unbiased genrelized Liu estimator and unbiased estimation of the Bias and MSE", *Comm. Statist. Theo Meth.*, **24**(7), 1789–1797.

[8] Akdeniz F. and Erol H. (2003), "Mean squared error comparisonsof some biased estimators in linear regression",*Comm. Statist. Comp. Simul.*, **32**, 12, 2389-2413.

[9] Akdeniz F. (2004), "New biased estimators under the LINEX loss function", *Stat. Papers*, **45**, 175-190.

[10] Akdeniz F. and Akdeniz Duran E. (2009), "Liu-type estimator in semiparametric regression Model", *J. Statist. Theo. Meth.*, **32**, 12, 2391-2415.

[11] Akdeniz Duran E. and Akdeniz F. (2012), "Efficiency of the modified jackknifed Liu-type estomator", *Stat. Papers*, **53**, 265-280.

[12] Batah F., Ramanathan T. K. and Gore S. D. (2008), “The efficiency of modified jackknife and ridge type regression estimators: a comparison”, *Surv. Math. Appl.*, **3**, 111-122.

[13] Belsley D. A., Kuh E. and Welsch R. E. (1980), *“Regression Diagnostics”*, John Wiley, New York.

[14] Belsley D. A. (1991), *“Conditioning diagnostics: collinearity and weak data in regression”*, John Wiley, New York.

[15] Chatterjee S., Hadi A. S. and Price B. (1938), *“Regression Anaalysis by Example”*, Third Edition, John Wiley, New York.

[16] Davison A. C. and Hinkley D. V (1997), *“Bootstrap Methods and their application”*,Cambridge university press, Cambridge.

[17] Efron B. (1979), “Bootstrap Methods: Another Look At The Jackknife”, *Ann. Statist.*, **7**, 1, 1-26.

[18] Efron B. (1980), *“The jackknife. the bootstrap, and other resampling plans”*, Stanford university press, California.

[19] Efron B. and Gong G. (1983), “A Leisurely Look at the Bootstrap, the Jackknife, and Cross Validation”, *J. Amer. Statist. Assoc.*, **37**, 1, 36-48.

[20] Farebrother R. W. (1976), “Further results on the mean square error of ridge regression”, *J. Royal. Statist. Soc.*, **38**, 245-250.

[21] Freedman D. A. (1981), “Bootstrapping regression models”, *Ann. Statist.*, **9**, 6, 1218-1228.

[22] Galton F. (1886), “Regression towards mediocrity in hereditary stature”, *The Journal of the Anthropological In stitute of Great Britain and Ireland*, **15**, 246-263.

[23] Hinkley D. V. (1977), “Jackknifing in unbalanced situations”, *The Chemometrics*, **19**, 3, 285-292.

[24] Hocking R. R. (1976), “The analysis and selection of variables in linear regression”, *Biometrika*, **32**, 1-49 .

[25] Hoerl A. E. and Kennard R. W. (1970), “Ridge regression: biased estimation for non-orthogonal problems”, *The Chemometrics*, **12**, 69-82.

[26] hubert M. H. and Wijekoon P. (2006), “Improvement of the Liu estimation in linear regression model”, *Stat. Papers*, **47**, 3, 471-479.

[27] Kariya T. and Kurata H. (2004), *"Generalized least squares"*, John Wiley and Sons, England .

[28] Li Y. and Yang H. (2012), "A new Liu-type estimator in linear regression model", *Stat. Papers*, **53**, 427-437.

[29] Liu K. J. (1993), "A new class of biased estimate in linear regression", *Comm. Statist. Theo. Meth.*, **22**, 2, 393-402.

[30] Liu K. J. (2003), "Using Liu type estimator to combat multicollinearity", *Comm. Statist. Theo. Meth.*, **32**, 5, 1009-1020.

[31] Marquardt D. W. and Snee R. D. (1975), "Ridge regression in practice", *J. Amer. Statist. Assoc.*, **29**, 1, 3-20.

[32] McDonald G. C., Galarnaeu D. I. and Keimel K. (1975), "A Monte Carlo evaiuation of some ridge-type estimators", *J. Amer. Statist. Assoc.*, **20**, 407-416.

[33] Montgomery D. C. and Peck A. E. (1992), "Introduction to linear regression analysis", *Scand J Statist*, 2nd Ed., John Wiley, New York.

[34] Nyquist H. (1988), "Applications of the jackknifed procedure in ridge regression", *Comp. Statist. Data. Anal.*, **6**, 177-183.

[35] Rao C. R. and Toutenberg H. (1995), *"Linear Models: Least Squares and Alternatives"*,2nd Ed., Springer Heidelberg.

[36] Rencher C. A. (2000), *"Linear Models in statistics"*, John Wiley, New York.

[37] Rubin D. B. (1981), "The bayesian bootstrap", *Ann. Statist*, **9**, 130-134.

[38] Sakalhooglu S., Kaciranlar S. and Akdeniz F. (2001), "Mean squared error comparisons of some biased estimators", *Comm. Statist. Theo. Meth.*, **30**, 2, 347-361.

[39] Sakalhooglu S. and Kaciranlar S. (2008), "A new biased estimator based on the ridge estimation", *Stat. Papers*, **49**, 669-689.

[40] Schucany W. R., Gray H. L. and Owen D. B. (1971), "On bias reduction in estimation", *J. Amer. Statist. Assoc.*, **66**, 335, 524-533.

[41] Searle S. R. (1971), *"Linear models"*, John Wiley, New York.

[42] Silverman B. W. and Yang G. A. (1987), "The bootstrap: to smooth or not smooth?", *Biometrika*, **74**, 469-479.

[43] Singh B., Chaubey Y. P. and Dwivedi T. D. (1986), “An almost unbiased ridge estimator”, *Sankhya*, **48**, 342-346.

[44] Stein M. C. (1956), “Inadmissibility of the usual estimator fort he mean of a multivariate normal distribution”, *Proc 3rd Berkeley Symp*, **1**, 197-206.

[45] Swamy P. A. V. B., Mehta J. S. and Robert P. N (1978), “Two methods of evaluting Horel and Kennard’s ridge regression”, *Comm. Statist. Theo. Meth.*, **12**, 1133-1155.

[46] Swindel B. F. (1976), ”Good ridge estimators based on prior information”, *Comm. Statist. Theo. Meth.*, **5**, 2, 1065-1075.

[47] Theil H. (1963), “On the use of incomplete prior information in regression analysis”, *J. Amer. Statist. Assoc.*, **58**, 401-414.

[48] Theil H. and Goldberger A. S. (1961), “On pure and mixed Statistical estimation in economics”, *Int. Econ. Rev.*, **2**, 65-78.

[49] Trenkler G. and Toutenburg H. (1990), “Mean squared error matrix comparison between biased estimators an overview of recent results”, *Stat Papers*, **31**, 165-179.

[50] Tukey J. W. (1958), “Bias an confidence in not- quitelarge samples”, *Ann. Statist. Math.*, **29**, 416.

[51] Van der vaart A. W. (1998), *Asymptotic Statistics*, 1st Ed., Cambridge University Press, New York.

[52] Vinod H. D. (1995), ”Double bootstrap for shrinkage estimators”, *J. Econometrics*, **68**, 287-302.

[53] Woods H., Steinour H. H. and Starke H. R. (1939), “Effect of composition of Portland cement on heat evolved during hardening”, *Industrial Eng. Chem.*, **24**, 1207-1241.

[54] Yang H. and Xu J. (2009), “An alternative stochastic restricted Liu estimator in linear regression”, *Stat. Papers*, **50**, 3, 639-647.

[55] Yang H., Chang X. F. and Liu D. (2009), “Improvement of the Liu estimator in weighted mixed regression”, *Comm. Statist. Theo. Meth.*, **38**, 285-292

[56] Yatchew A. (2003), *“Semiparametric regression for the applied econometrican”*,Cambridge University Press, Cambridge.

[57] Zhang Ch. and Yang Hu. (2007), "The conditional ridge- type estimation in singular linear model with linear equality restriction", *Ann. Statist.*, **41**, 6, 485-494.

واژه‌نامه فارسی به انگلیسی

اثر ریج	Ridge trace
اطلاع پیشین	Prior information
برآوردگر آمیخته	Mixed estimation
برآوردگر با اریبی کاهش یافته	Bias-reduced estimator
برآوردگر محدودشده	Restricted estimator
برازش	Fiting
بردار ویژه	Eigen vector
برگشت به میانگین	Regression to the mean
بوت استرپ بیزی	Bayesian bootstrap
بوت استرپ پارامتری	Parametric bootstrap
بوت استرپ دوگانه	Double bootstrap
بوت استرپ هموار	Smoothed bootstrap
پیش‌بین	Predictor
تجزیه طیفی	Spectral factorization
جک‌نایف	Jackknife
رگرسیون خطی چندگانه	Multiple linear regression
رگرسیون خطی ساده	Ordinary linear regression
عامل تورم واریانس	Variance inflation factor
عدد شرطی	Condition number
گوس- مارکف	Gauss-Markov
کاپا	Kappa
کمترین توان‌های دوم	Least square
متعامد	Orthogonal
مستقل خطی	Linear independence
معین مثبت	Positive definite

مقدار ویژه .... Eigen value
میانگین توان دوم خطا .... Mean square error square
نامستقل خطی .... Linear dependence
ناویژه .... Non-singular
نیمه معین مثبت .... Positive semidefinite
همخطی چندگانه .... Multiple collinearity

واژه‌نامه انگلیسی به فارسی

Bayesian bootstrap	بوت استرپ پارامتری
Bias-reduced estimator	برآوردگر با اریبی کاهش یافته
Condition number	عدد شرطی
Double bootstrap	بوت استرپ دوگانه
Eigen value	مقدار ویژه
Eigen vector	بردار ویژه
Fiting	برازش
Gauss-Markov	گوس- مارکف
Jackknife	جک‌نایف
Kappa	کاپا
Least square	کمترین توان‌های دوم
Linear dependence	نامستقل خطی
Linear independence	مستقل خطی
Linear unbiased estimators	برآوردگرهای خطی نااریب
Mean square error square	میانگین توان دوم خطا
Mixed estimation	برآوردگر آمیخته
Multiple collinearity	همخطی چندگانه
Multiple linear regression	رگرسیون خطی چندگانه
Non-singular	ناویژه
Ordinary linear regression	رگرسیون خطی ساده
Orthogonal	متعامد
Parametric bootstrap	بوت استرپ پارامتری
Positive definite	معین مثبت
Positive semidefinite	نیمه معین مثبت
Predictor	پیش‌بین

Prior information .... اطلاع پیشین
Regression to the mean .... برگشت به میانگین
Restricted estimator .... برآوردگر محدودشده
Ridge trace .... اثر ریج
Smoothed bootstrap .... بوت استرپ هموار
Spectral factorization .... تجزیه طیفی
Symmetric .... متقارن
Variance inflation factor .... عامل تورم واریانس

Aabstract

When multicollinearity among the columns of the design matrix of the linear regression exists, using the least squares method, is usually too weak to obtain efficient estimates of regression coefficients. It has been proved that the variance of least squares estimates may significantly increase and the length of parameter's vector get too big. In the presence of multicollineatity in this context, one way to overcome the problems, is using the biased estimators.

Several methods such as ridge regression, Liu and etc proposed for obtaining biased estimates of regression coefficients. One of these methods is using the modified Jackknife Liu estimator that is the combination of the Liu and Jackknife Liu estimators. However, this estimator is biased but has lower variance and also has a better performance comparing to the least squares estimator.

In this thesis, we have studied the ridge regression model. Using the results of Akdeniz Duran and Akdeniz (2012) and Li and Yang (2012), we construct a new estimator, namely the modified jackknife Liu estimator. Given the complexity of the proposed estimator, the Monte Carlo numerical methods are provided to illustarte the bias and risk of the namely proposed estimator.

Keywords: Jackknifed estimator, Ridge estimator, Liu estimator, Multiple regression, Risk, Multicollinearity

Shahrood University
Faculty Of Mathematical Sciences

Dissertation Submitted in Partial
Fulfillment of The Requirements For The
Degree of Master of Science in
Statistics

On the performance of Jackknifed ridge estimator

Supervisor
Dr. Mohammad Arashi

by
Maryam Borzoei Bidgoli

September 2013