



پایان نامه جهت اخذ درجه کارشناسی ارشد

آمار ریاضی

بررسی رفتار برآوردگرهای رگرسیونی ريج در مدل های خطی منفرد

نگارش

سمانه نجاریان

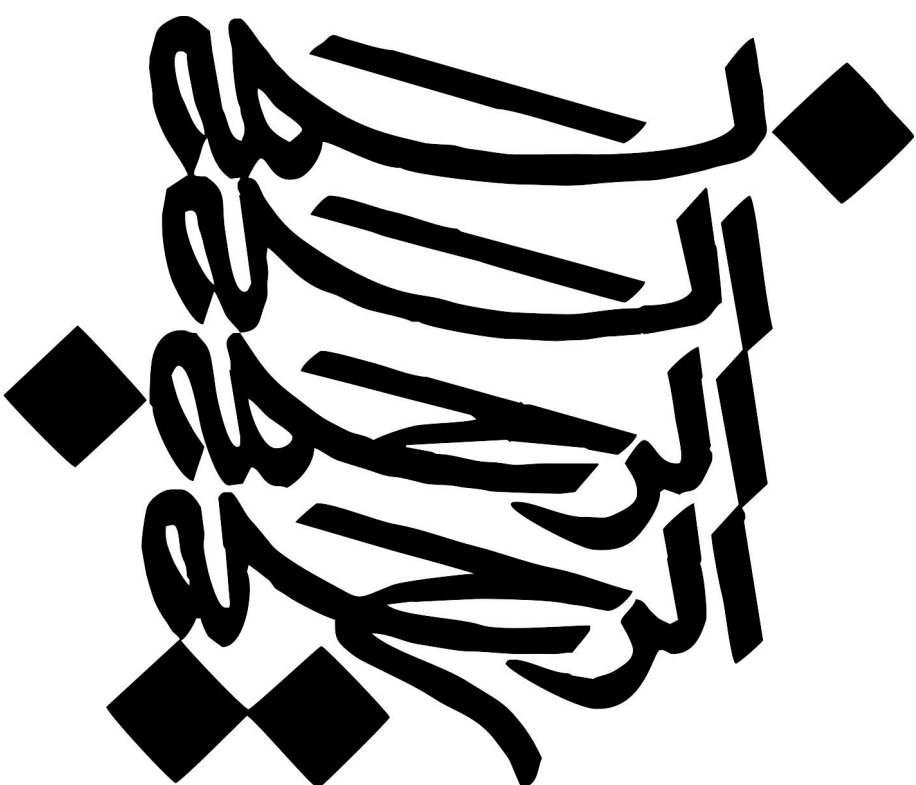
استاد راهنما

دکتر محمد آرشی

استاد مشاور

دکتر جعفر فتحعلی

شهریور ۱۳۹۰



تقدیم به:

همه آنانی که

به من اندیشیدن آموختند

تا امروز

کورسوی ظلمت را از روشنایی معرفت

باز شناسم.

فهرست مقالات مستخرج از متن پایان نامه

[۱] نجاریان. س. و آرشی. م. (۱۳۸۹)، شبیه سازی در مدل رگرسیون خطی ريج منفرد، ششمین همایش ملی

آمار – دانشگاه پیام نور اهواز.

[2] Najarian. S. and Arashi. M. (2010), Ridge Estimation In Singular Linear Model: New Approach, The 10th Iranian Statistical Conference.

[3] Najarian. S., Arashi. M. and Golam Kibria. B. M. (2011), A Simulation Study On Some Ridge Regression Estimators, Communication in statistics– simulation Restricted and Computation.

پیشگفتار

در دهه‌های اخیر، تلاش‌های زیادی در راستای یافتن برآوردهای بهبودیافته انجام گرفته است. یکی از مسائلی که زمینه‌ساز این تلاش بوده، همخطی چندگانه می‌باشد. وقتی که ستون‌های ماتریس طرح وابسته خطی باشند، برآوردهای کمترین توان دوم بدست آمده به روش گوس-مارکف به دلیل نزدیک به صفر بودن دترمینان ماتریس $X'X$ ، کارایی لازم را نخواهند داشت. این مطلب منجر به ارائه برآوردهای بهبودیافته از جمله برآوردهای نوع ريج شده است. هدف از این پایان‌نامه، ارائه برآوردهای ريج در مدل‌های خطی محدودشده و محدودنشده و تحلیل و تفسیر ویژگی‌های برآوردهای ارائه‌شده، است. این مجموعه شامل ۵ فصل و ۴ ضمیمه می‌باشد. مطالب هر فصل بطور مختصر عبارتند از:

در فصل ۱، مقدمه‌ای بر تحلیل رگرسیونی، ارائه و چگونگی برآورد ضرایب رگرسیونی و همچنین چگونگی بروز همخطی در مدل‌های رگرسیون چندگانه خطی و روش‌های تشخیص همخطی بررسی شده است.

در فصل ۲، برآوردهای رگرسیونی ريج در مدل خطی محدودشده مورد بررسی قرار گرفته و خطای برآوردهای محاسبه شده‌است و به کمک شبیه‌سازی و مثال عددی کارایی برآوردهای ريج و برتریش بر برآوردهای کمترین توان‌های دوم محدودشده نشان داده شده است.

در فصل ۳، برآوردهای رگرسیونی ريج محدودشده و محدودنشده در مدل خطی منفرد ارائه‌شده و ویژگی‌های برآوردهای ريج بررسی شده است.

در فصل ۴، برآوردهای رگرسیونی ريج در مدل خطی منفرد با در نظر گرفتن محدودیتی که دارای جمله خطای

تصادفی است، ارائه و ویژگی‌های برآوردگر بررسی و به کمک شبیه‌سازی کارایی برآوردگر تایید شده است. در فصل ۵، برآوردگر رگرسیونی ریج منفرد در مدل به ظاهر نامرتبط (SUR) ارائه و ویژگی‌های برآوردگر بطور خلاصه بیان شده است.

در ضمیمه A ، بحث و نتیجه‌گیری و پیشنهادها برای ادامه کار در آینده، مطرح شده‌اند. در ضمیمه B ، عملیات جبری ماتریس‌ها بمنظور سادگی درک اثبات قضایا و لم‌ها در متن پایان‌نامه بیان شده است.

در ضمیمه C ، توضیحاتی در خصوص عامل تورم واریانس به عنوان عاملی برای شناخت همخطی در مدل رگرسیونی خطی داده شده است.

در ضمیمه D ، توضیحات کاملتری درباره عدد شرطی که عاملی دیگر برای تشخیص مسئله همخطی است، ارائه شده است.

* در این مجموعه، قضایایی که برهان آن از نویسنده این پایان‌نامه می‌باشد با علامت ستاره مشخص شده است.

مدلهای ارائه شده در این مجموعه

$Y = X\beta + E$ $R\beta = r$ $E(E) = 0 \quad Cov(E) = \sigma^2 I$ $Y \sim n \times 1 \quad X \sim n \times p \quad \beta \sim p \times 1$ $E \sim n \times 1 \quad R \sim q \times p \quad r \sim q \times 1$	فصل دوم
$Y = X\beta + E$ $R\beta = r$ $E(E) = 0 \quad Cov(E) = \sigma^2 \Sigma$ $Y \sim n \times 1 \quad X \sim n \times p \quad \beta \sim p \times 1$ $E \sim n \times 1 \quad R \sim q \times p \quad r \sim q \times 1$	فصل سوم
$Y = X\beta + E$ $R\beta + v = r$ $E(E) = 0 \quad Cov(E) = \sigma^2 I$ $Y \sim n \times 1 \quad X \sim n \times p \quad \beta \sim p \times 1$ $E \sim n \times 1 \quad v \sim q \times 1$ $R \sim q \times p \quad r \sim q \times 1$	فصل چهارم
$Y = X\beta + E$ $R\beta = r$ $E(E) = 0 \quad Cov(E) = \Sigma \otimes I_q$ $Y \sim n \times 1 \quad X \sim n \times k \quad \beta \sim k \times 1$ $E \sim n \times 1 \quad R \sim q \times p \quad r \sim q \times 1$ $k = \sum_{i=1}^p k_i \quad n = pq$	فصل پنجم

قدردانی

منت خدای را عزوجل که طاعتش موجب قربت است و به شکر اندرش مزید نعمت

در پایان لازم می‌دانم از استاد گرامی، جناب آقای دکتر آرشی که راهنمایی و هدایت این پایان‌نامه را عهده‌دار بودند و تلاش زیادی در راستای جهت‌دهی به تفکرات و تلاش‌های اینجانب داشته‌اند، قدردانی کنم. همچنین

از جناب آقای دکتر فتحعلی جهت مشاوره و راهنمایی‌های بی‌دریغشان کمال تشکر را دارم.

از اساتید داور محترم، آقایان دکتر داوود شاهسونی (دانشگاه صنعتی شاهرود) و دکتر مهدی عمادی (دانشگاه

فردوسی مشهد) که با حضور گرمشان تصحیح و داوری پایان‌نامه را بر عهده گرفتند، سپاسگزارم.

از جناب آقای پروفسور کیبیریا (دانشگاه فلوریدای آمریکا) بخاطر راهنمایی‌هایشان در استخراج مقاله‌ای از فصل

یک این پایان‌نامه، تشکر می‌کنم.

از جناب آقای دکتر فکور و آقای دکتر روزبه بخاطر همکاری‌شان با بنده کمال تشکر را دارم.

همچنین از مسئولین کتابخانه دکتر فاطمی دانشگاه فردوسی مشهد بخصوص آقای مسئله‌گو بخاطر همکاری

صمیمانه‌شان متشکرم.

سمانه نجاریان

تابستان ۱۳۹۰

علائم اختصاری

MSE : میانگین توان دوم خطا *Mean Square Error*

MSEM : ماتریس میانگین توان دوم خطا *Mean Square Error Matrix*

LSE : برآوردگر کمترین توان‌های دوم *Least Squares Estimator*

RLSE : برآوردگر کمترین توان‌های دوم محدودشده *Restricted Least Squares Estimator*

GLSE : برآوردگر کمترین توان‌های دوم محدودشده *Generalized Least Squares Estimator*

RE : برآوردگر ریج *Ridge Estimator*

RRE : برآوردگر ریج محدودشده *Restricted Ridge Estimstor*

RLM : مدل خطی محدودشده *Restricted Linear Model*

SE : برآوردگر استاین *Stein Estimator*

LE : برآوردگر لویی *Lui Estimator*

PCE : برآوردگر مولفه‌های اصلی *Principal Components Estimator*

PLSE : برآوردگر کمترین توان‌های دوم جزئی *Partial Least Squares Estimator*

AUE : برآوردگر تقریباً ناریب *Almost Unbaised Estimator*

BLUE : بهترین برآوردگر ناریب خطی *Best Linear Unbaised Estimator*

SUR : مدل به ظاهر نامرتبط *Seemingly Unrelated Regression*

Asymptotic Normal *AN* : نرمال مجانبی

Variance Inflation Factor *VIF* : عامل تورم واریانس

Condition Number *CN* : عدد شرطی

فهرست مطالب

۱	مقدمه و دورنما	۱
۳	۱.۱ مدل رگرسیونی چندگانه	۳
۴	۲.۱ برآورد پارامترهای مدل	۴
۶	۳.۱ همخطی در رگرسیون خطی چندگانه	۶
۸	۴.۱ آثار همخطی چندگانه	۸
۹	۵.۱ شاخص‌های همخطی چندگانه	۹
۹	۱.۵.۱ محک ماتریس همبستگی	۹
۱۰	۲.۵.۱ عامل تورم واریانس	۱۰
۱۰	۳.۵.۱ عدد شرطی	۱۰
۱۲	۲ برآوردگر رگرسیونی ريج محدودشده	۱۲
۱۶	۱.۲ معرفی برآوردگر	۱۶
۲۰	۲.۲ محاسبه خطای برآوردگر	۲۰
۲۷	۳.۲ برآورد مقادیر k	۲۷
۲۹	۴.۲ شبیه‌سازی	۲۹
۳۵	۵.۲ کاربرد	۳۵

۳۶		مثال ۱	۱.۵.۲
۳۸		مثال ۲	۲.۵.۲
۳۹		مثال ۳	۳.۵.۲
۴۰		نتیجه‌گیری	۶.۲
۴۲		برآوردگر رگرسیونی رنج منفرد	۳
۴۴		معرفی برآوردگر	۱.۳
۴۶		ویژگی‌های برآوردگر	۲.۳
۵۸		شبیه‌سازی	۳.۳
۶۰		نتیجه‌گیری	۴.۳
۶۱		برآوردگر رنج منفرد در مدل با محدودیت تصادفی	۴
۶۲		معرفی برآوردگر	۱.۴
۶۶		ویژگی‌های برآوردگر	۲.۴
۶۹		شبیه‌سازی	۳.۴
۷۱		نتیجه‌گیری	۴.۴
۷۲		برآوردگر رنج منفرد در مدل <i>SUR</i>	۵
۷۶		معرفی برآوردگر	۱.۵
۷۹		ویژگی‌های برآوردگر	۲.۵
۸۰		نتیجه‌گیری	۳.۵
۸۱		الف خلاصه و پیشنهادات برای تحقیقات آینده	
۸۱		الف.۱ خلاصه	

الف. ۲. پیشنهادات برای پژوهش‌های آینده (ارائه زمینه‌های تحقیق)	۸۲
ب جبر ماتریس‌ها	۸۴
پ عامل تورم واریانس	۸۹
ت عدد شرطی	۹۱
مراجع	۹۴

فصل ۱

مقدمه و دورنما

مقدمه

تحلیل رگرسیونی یکی از روش‌های آماری برای تحلیل داده‌های چندعاملی است که حوزه کاربرد آن بیشترین وسعت را دارد، نتایج پرجاذبه آن از نظر مفهومی، فرآیند ساده بکارگیری یک معادله است که ارتباط بین دسته‌ای از متغیرها را بیان می‌کند. تحلیل رگرسیونی همچنین از جهت نظری به لحاظ ظرافت و زیبایی ریاضیات دارای جذابیت است. موفقیت در تحلیل، نیاز به درک و تیزبینی در دو مقوله تئوری و مسائل عملی دارد و زمانی بروز می‌کند و خود را نشان می‌دهد که تکنیک در بکارگیری داده‌ها و اطلاعات جهان واقع قرار می‌گیرد.

تحلیل رگرسیونی فن و تکنیکی آماری برای بررسی و مدل‌بندی ارتباط بین متغیرهاست. کاربردهای رگرسیون متعدد است و تقریباً در هر زمینه‌ای از جمله مهندسی، فیزیک، اقتصاد، مدیریت، علوم زیستی و بیولوژی و علوم اجتماعی صورت می‌پذیرد.

کلمه "رگرسیون" در لغت به معنای "برگشت" است. دلیل این نامگذاری به مطالعات "گالتون" در مورد رابطه قد پسر و قد پدر مربوط می‌شود. گالتون با رسم نمودار قد پسر در مقابل قد پدر (نمودار پراکنش) براساس حجم بزرگی از داده‌ها به این نتیجه رسید که اغلب افراد بلندقد، پسرانی کوتاه‌قد و پدارن کوتاه‌قد، پسرانی بلندقدتر از خود دارند و این پدیده را "برگشت به میانگین" نامید و از آن پس نام رگرسیون بر روش‌های بررسی روابط بین متغیرهای آماری باقی ماند.

به عنوان مثالی از یک مسئله که در آن تحلیل رگرسیونی می‌تواند مفید واقع شود، فرض می‌کنیم 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 گمان می‌کند که زمان لازم برای اینکه یک فرد تحویل دهنده، این ماشین را سرویس دهد بستگی به تعداد موارد محصول تحویل شده دارد. مهندس ۲۵ فروشنده جزئی را که دارای ماشین فروش می‌باشند، به تصادف بازدید کرده است و زمان تحویل (برحسب دقیقه) و حجم محصول تحویل داده شده (بر حسب مورد) برای هر یک مشاهده می‌شود. بوسیله نمودار پراکنش، ارتباط بین

زمان تحویل و حجم تحویل مشخص می‌شود. نقاطی که نشان‌دهنده داده‌ها هستند در امتداد یک خط مستقیم قرار می‌گیرند، نه دقیقاً روی خط. اگر Y نشان‌دهنده زمان تحویل و X را برای حجم تحویل شده قرار دهیم، در این صورت معادله یک خط مستقیم که این دو متغیر را به هم مربوط می‌سازد، چنین است

$$Y = \beta_0 + \beta_1 X \quad (1.1)$$

که β_0 عرض از مبدا و β_1 شیب خط است. از آنجایی که نقاط داده‌ها دقیقاً روی یک خط مستقیم قرار نمی‌گیرند بنابراین (۱.۱) بایستی به عنوان برآوردکننده آن اصلاح و تعدیل شود. اگر اختلاف بین مقدار مشاهده‌شده Y و خط مستقیم $(\beta_0 + \beta_1 X)$ را خطای E قرار می‌دهیم، مناسب است که خطای E به عنوان یک خطای آماری تلقی شود. بدین معنی که متغیری تصادفی است که عدم برازش را بیان و اندازه ناتوانی مدل از برازش دقیق داده‌ها را برآورد می‌کند. این خطا ممکن است به لحاظ اثرات دیگر متغیرها، خطاهای اندازه‌گیری و غیره روی زمان تحویل صورت پذیرد. بنابراین مدلی با قابلیت پذیرش بیشتر برای داده‌های زمان تحویل عبارتست از

$$Y = \beta_0 + \beta_1 X + E \quad (2.1)$$

معادله (۲.۱) یک مدل رگرسیون خطی نامیده می‌شود. در این مدل X متغیر مستقل و Y متغیر وابسته نامیده می‌شود. در هر حال این نامگذاری باعث می‌شود که این مفهوم با مفهوم استقلال آماری اشتباه شود. چون در (۲.۱) فقط یک متغیر رگرسیونی وجود دارد، این مدل را مدل رگرسیونی ساده نامیده‌اند.

۱.۱ مدل رگرسیونی چندگانه

در حالت کلی ممکن است متغیر پاسخ به k متغیر رگرسیونی $X_1, X_2, \dots, X_k$ مربوط باشد، بنابراین مدل رگرسیونی به صورت زیر نوشته می‌شود

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k + E \quad (3.1)$$

چون در این مدل رگرسیونی بیش از یک متغیر درگیر است، رگرسیونی چندمتغیره نام دارد. پارامترهای β_j ($j = 0, 1, \dots, k$) ضرایب رگرسیون نامیده می‌شود. بکارگیری صفت خطی بودن در مدل فوق اشاره بر این دارد که مدل، برحسب پارامترهای $\beta_0, \beta_1, \dots, \beta_k$ خطی است و نه به دلیل اینکه Y تابعی خطی از X هاست. این مدل یک ابرصفحه در فضای k بعدی از متغیرهای رگرسیونی X_j است. پارامتر β_j نشان‌دهنده تغییرات مورد انتظار متغیر پاسخ به ازای یک واحد تغییر در X_j است و قتیکه همه متغیرهای رگرسیونی باقیمانده دیگر X_i ($i \neq j$) ثابت نگهداشته شوند. به همین جهت پارامترهای β_j ($j = 1, 2, \dots, k$) ضرایب جزئی رگرسیون نامیده می‌شوند. مدل‌های رگرسیون چندگانه، اغلب به عنوان تقریب توابع بکار می‌روند. بدین معنی که ارتباط تابعی واقعی بین $X_1, X_2, \dots, X_k$ و Y شناخته شده نیست، اما مدل رگرسیون خطی روی دامنه تغییرات متغیرهای رگرسیونی، تقریب مناسبی از ارتباط تابعی مذکور می‌باشد.

۲.۱ برآورد پارامترهای مدل

برای برآورد ضرایب رگرسیون، روش کمترین توان‌های دوم بکار می‌رود. فرض می‌کنیم $n > k$ مشاهده در دسترس است و Y_i نمایش دهنده i امین پاسخ مشاهده شده و X_{ij} نمایش دهنده j امین مشاهده در i امین سطح متغیر رگرسیونی باشد. فرض می‌کنیم جمله خطای E در مدل دارای $E(E) = 0$ و $Var(E) = \sigma^2$ و خطاها ناهمبسته هستند. می‌توان مدل متناظر با (۳.۱) را بصورت زیر نوشت

$$\begin{cases} Y = X\beta + E \\ E(E) = 0, \quad Cov(E) = \sigma^2 I \end{cases} \quad (4.1)$$

تابع کمترین توان‌های دوم خطا بصورت زیر نوشته می‌شود

$$\begin{aligned} S(\beta_0, \beta_1, \dots, \beta_k) &= \sum_{i=1}^n E_i^2 \\ &= \sum_{i=1}^n \left(Y_i - \beta_0 - \sum_{j=1}^k \beta_j X_{ij} \right)^2 \\ &= (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})'(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}) \end{aligned}$$

با مشتقگیری از S نسبت به $\beta_0, \beta_1, \dots, \beta_k$ و مساوی با صفر قرار دادن مشتق حاصل، در صورت نامنفرد بودن

ماتریس $\mathbf{X}'\mathbf{X}$ ، برآوردگرهای ضرایب مدل رگرسیونی را بصورت زیر بدست می‌آوریم

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'\mathbf{Y} \quad (5.1)$$

بردار مقادیر برازش شده $\hat{Y}_i$ با مقادیر مشاهده شده Y_i چنین خواهد بود

$$\hat{\mathbf{Y}} = \mathbf{X}\hat{\boldsymbol{\beta}} = \mathbf{H}\mathbf{Y} \quad (6.1)$$

ماتریس $\mathbf{H} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'$ ، ماتریسی $p \times p$ ، متقارن و خودتوان است و بنابراین $\text{Rank}(\mathbf{H}) =$

$\text{tr}(\mathbf{H}) = p$ که ماتریس هت یا برازش نامیده می‌شود زیرا که بردار مقادیر مشاهدات را به بردار مقادیر

برازش شده تصویر می‌کند. ماتریس برازش و خواص آن نقشی مرکزی در تحلیل رگرسیونی بازی می‌کند.

ماتریس واریانس-کوواریانس $\hat{\boldsymbol{\beta}}$ عبارتست از

$$\text{Cov}(\hat{\boldsymbol{\beta}}) = \sigma^2 (\mathbf{X}'\mathbf{X})^{-1} \quad (7.1)$$

همچنین برآوردگر ناریب σ^2 عبارتست از میانگین توان‌های دوم باقیمانده‌ها

$$\hat{\sigma}^2 = \frac{SSE}{n - p} \quad (8.1)$$

که در آن $SSE = (Y - X\beta)'(Y - X\beta)$ ، مجموع توان‌های دوم باقیمانده‌هاست. حال اگر در مدل (۴.۱)، $Cov(E) = \sigma^2 \Sigma$ باشد، که در آن Σ ماتریسی معیت مثبت است. برآوردگر ضرایب رگرسیونی و ماتریس واریانس-کوواریانس برآوردگرها و همچنین برآوردگر ناریب σ^2 به ترتیب به صورت زیر خواهند بود

$$\hat{\beta} = (X' \Sigma^{-1} X)^{-1} X' \Sigma^{-1} Y \quad (9.1)$$

9

$$Cov(\hat{\beta}) = \sigma^2 (X' \Sigma^{-1} X)^{-1} \quad (10.1)$$

9

$$\hat{\sigma}^2 = \frac{(Y - X\hat{\beta})' \Sigma^{-1} (Y - X\hat{\beta})}{n - p} \quad (11.1)$$

در این صورت مقادیر برآزش شده به فرم زیر می‌باشند

$$\hat{Y} = X\hat{\beta} = H \Sigma^{-1/2} Y \quad (12.1)$$

که در آن $H = X \Sigma^{-1/2} (X' \Sigma^{-1} X)^{-1} X' \Sigma^{-1/2}$.

۳.۱ همخطی در رگرسیون خطی چندگانه

مدل‌های رگرسیونی در حوزه کاربردی وسیعی مورد استفاده قرار می‌گیرند. یک مسئله جدی که می‌تواند استفاده از مدل رگرسیونی را با اشکال مواجه کند، همخطی چندگانه یا وابستگی خطی نزدیک بین متغیرهای رگرسیونی است. وجود وابستگی خطی نزدیک، توانایی برآورد ضرایب مدل رگرسیون را با مشکل مواجه می‌کند.

بکارگیری و تعبیر یک مدل رگرسیون چندگانه بطور ضمنی یا بطور صریح به برآوردهای ضرایب مدل رگرسیون بستگی دارد. بعضی از استنباط‌هایی که بطور معمول در مدل رگرسیونی صورت می‌پذیرد، شامل

موارد زیر است:

۱- مشخص کردن اثرات نسبی متغیرهای رگرسیونی

۲- پیش‌بینی و یا برآورد

۳- انتخاب دسته‌ای مناسب از متغیرها برای مدل

اگر هیچ رابطه خطی بین متغیرهای رگرسیونی نباشد، گفته می‌شود که متغیرهای رگرسیونی مستقلند. اگر متغیرهای رگرسیونی مستقل باشند، استنباط‌هایی مانند مواردی که در بالا ذکر شد، نسبتاً بسادگی صورت می‌پذیرد. متأسفانه در بیشتر کاربردهای رگرسیون، متغیرهای رگرسیونی مستقل نیستند. هنگامی که ارتباط خطی نزدیکی بین متغیرهای رگرسیونی وجود دارد، گفته می‌شود مسئله همخطی چندگانه وجود دارد. چهار منبع اولیه همخطی وجود دارد:

۱- شیوه جمع‌آوری داده‌ها

۲- قیدهای روی مدل یا درون جامعه

۳- نوع مدل

۴- مدلی با متغیرهای رگرسیونی بیش از حد

درک اختلاف بین این منابع همخطی چندگانه اهمیت دارد، به‌طوری‌که توصیه برای تحلیل داده‌ها و تعبیر مدل نتیجه‌شده تا اندازه‌ای به علت مسئله همخطی بستگی دارد. (برای اطلاعات بیشتر در مورد همخطی

چندگانه رجوع کنید به میسون، گونست و وبستر^۱ (۱۹۷۵).

^۱ Mason, Gunst and Webster

۴.۱ آثار همخطی چندگانه

وجود همخطی چندگانه تاثیرات جدی بر برآوردهای کمترین توان‌های دوم ضرایب رگرسیون دارد. برای درک بهتر این آثار، مدلی با دو متغیر رگرسیونی را در نظر بگیرید.

$$Y = \beta_1 X_1 + \beta_2 X_2 + E$$

دستگاه معادلات کمترین توان‌های دوم نرمال عبارتند از

$$(X'X)\hat{\beta} = X'Y$$

$$\begin{bmatrix} 1 & r_{12} \\ r_{12} & 1 \end{bmatrix} \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix} = \begin{bmatrix} r_{1y} \\ r_{2y} \end{bmatrix}$$

که در آن r_{12} همبستگی ساده بین X_1 و X_2 و r_{jy} همبستگی ساده بین X_j و Y به ازای $j = 1, 2$ می‌باشد. معکوس $(X'X)$ عبارتست از

$$(X'X)^{-1} = \begin{bmatrix} \frac{1}{(1-r_{12}^2)} & \frac{-r_{12}}{(1-r_{12}^2)} \\ \frac{-r_{12}}{(1-r_{12}^2)} & \frac{1}{(1-r_{12}^2)} \end{bmatrix} \quad (13.1)$$

و برآورد ضرایب رگرسیون عبارتند از

$$\hat{\beta} = \frac{r_{1y} - r_{12}r_{2y}}{(1 - r_{12}^2)}, \quad \hat{\beta}_2 = \frac{r_{2y} - r_{12}r_{1y}}{(1 - r_{12}^2)} \quad (14.1)$$

اگر بین X_1 و X_2 همخطی چندگانه شدید موجود باشد، در این صورت ضریب همبستگی r_{12} بزرگ خواهد بود.

از (۱۳.۱) خواهیم دید اگر $|r_{12}| \rightarrow 1$ آنگاه $Var(\hat{\beta}_j) = S_{jj}\hat{\sigma}^2 \rightarrow \infty$ و $Cov(\hat{\beta}_1, \hat{\beta}_2) = S_{12}\hat{\sigma}^2 \rightarrow \infty$.

$\pm\infty$ بنابراین همخطی زیاد بین X_1 و X_2 منجر به واریانس‌ها و کوواریانس‌ها بزرگ برای برآوردهای کمترین

توان‌های دوم ضرایب رگرسیون خواهد شد. در نتیجه اگر از سطوح یکسان X نمونه‌های مختلف گرفته شود،

برآوردهای بدست آمده برای هر پارامتر با مقدار واقعی پارامتر، تفاوت زیادی خواهد داشت.

وقتی که بیش از دو متغیر رگرسیونی وجود دارد، همخطی اثرات مشابهی ایجاد می‌کند. می‌توان نشان داد که

اعضای قطر اصلی ماتریس $S = (X'X)^{-1}$ عبارتند از

$$S_{jj} = \frac{1}{1 - R_j^2}, \quad j = 1, 2, \dots, p \quad (15.1)$$

که در آن R_j^2 ضریب تعیین چندگانه از رگرسیون X_j نسبت به $p - 1$ متغیر رگرسیونی باقیمانده است. اگر همخطی شدید بین X_j و هر زیرمجموعه‌ای از $p - 1$ متغیر رگرسیونی دیگر وجود داشته باشد، در این صورت مقدار R_j^2 نزدیک به واحد خواهد بود. چون واریانس $\hat{\beta}_j$ برابر است با $(1 - R_j^2)^{-1} \hat{\sigma}^2$ ، $Var(\hat{\beta}_j) = S_{jj} \hat{\sigma}^2$ ، همخطی چندگانه شدید موجب می‌شود که واریانس برآورد کمترین توان‌های دوم ضرایب رگرسیونی β_j بسیار زیاد شود. بطور کلی اگر X_i و X_j در یک رابطه خطی درگیر باشند، کوواریانس $\hat{\beta}_i$ و $\hat{\beta}_j$ نیز بزرگ خواهد بود و وجود همخطی در ماتریس طرح باعث می‌شود برآوردگرهای کمترین توان‌های دوم $\hat{\beta}_j$ از نظر قدرمطلق خیلی بزرگ باشند. (برای اطلاعات بیشتر در مورد آثار همخطی رجوع کنید به داگلاس مونتگومری و الیزابت پک^۱ (۱۹۹۲).)

۵.۱ شاخص‌های همخطی چندگانه

در این قسمت بعضی از روش‌هایی که برای تشخیص همخطی بکار می‌روند، ارائه می‌شود.

۱.۵.۱ محک ماتریس همبستگی

فرض کنید X یک ماتریس $n \times p$ باشد. یک اندازه بسیار ساده همخطی چندگانه بررسی اعضای غیرقطر r_{ij} در ماتریس $X'X$ است. اگر متغیرهای رگرسیونی X_i و X_j تقریباً وابسته خطی باشند، در این صورت $|r_{ij}|$ نزدیک واحد خواهد بود.

^۱Montgomery, Douglas, C. and Peck, Elizabeth, A.

۲.۵.۱ عامل تورم واریانس

اعضای قطر ماتریس $S = (X'X)^{-1}$ برای آشکارسازی همخطی چندگانه کاربرد زیادی دارند. S_{jj} ، زامین عضو قطر ماتریس S می‌تواند بصورت $S_{jj} = (1 - R_j^2)^{-1}$ نوشته شود، که در آن R_j^2 ضریب تعیین بدست آمده است هنگامی که X_j به $p - 1$ متغیر رگرسیونی باقیمانده برگشته است. اگر X_j تقریباً بر متغیرهای رگرسیونی باقیمانده عمود باشد، R_j^2 کوچک و S_{jj} به واحد نزدیک می‌شود، در حالی که اگر وابستگی تقریباً خطی به بعضی متغیرهای رگرسیونی داشته باشد R_j^2 نزدیک به واحد و S_{jj} بزرگ می‌شود. چون واریانس زامین ضریب رگرسیون برابر است با $S_{jj}\sigma^2$ می‌توان S_{jj} را به عنوان عاملی دید که به وسیله آن واریانس $\hat{\beta}_j$ به لحاظ بستگی نزدیک بین متغیرهای رگرسیونی، افزایش یافته است. عامل

$$VIF_j = S_{jj} = (1 - R_j^2)^{-1} \quad (۱۶.۱)$$

را عامل تورم واریانس می‌نامند. این نامگذاری مربوط به مارکوارت^۱ (۱۹۷۰) است. برای هر جمله مدل، VIF اثر مرکب از ارتباط بین متغیرهای رگرسیونی روی واریانس آن جمله را اندازه‌گیری می‌کند. یک VIF بزرگ اشاره بر همخطی چندگانه دارد. تجربیات عملی حاکی از اینست که اگر هر یک از VIF ها از ۵ یا ۱۰ تجاوز کنند، اختطاری مبنی بر این است که ضرایب رگرسیونی ذیربط به علت همخطی چندگانه بصورت ضعیفی برآورد شده‌اند (برای اطلاعات کاملتر درباره عامل تورم واریانس پیوست B را ببینید).

۳.۵.۱ عدد شرطی

ریشه‌های مشخصه با مقادیر ویژه $X'X$ یعنی $\lambda_1, \lambda_2, \dots, \lambda_p$ می‌توانند برای اندازه‌گیری وسعت همخطی چندگانه در داده‌ها مورد استفاده قرار گیرند. عدد شرطی $X'X$ به صورت زیر تعریف می‌شود

$$CN = \frac{\lambda_{Max}}{\lambda_{Min}} \quad (۱۷.۱)$$

^۱Marquardt

عدد شرطی یک اندازه گستردگی در طیف مقادیر ویژه $X'X$ است. در حالت کلی اگر عدد شرطی کمتر از ۱۰۰ باشد، مشکل جدی در مورد همخطی چندگانه وجود ندارد. اعداد شرطی بین ۱۰۰ و ۱۰۰۰ منتج به همخطی چندگانه شدید و اگر عدد شرطی از ۱۰۰۰ تجاوز کند، همخطی چندگانه بسیار شدید را حکایت می‌کند (اطلاعات کاملتر در خصوص عدد شرطی در پیوست C آمده است).

فصل ۲

برآوردگر رگرسیونی ریج محدود شده

مقدمه

در مسئله برآورد کردن پارامترها، همواره به دنبال برآوردگری هستیم که کمترین میزان خطا را داشته باشد. بدین منظور در بین کلاس‌های مختلف، برآوردگری را می‌یابیم که کمترین واریانس را دارا باشد. زیرا که مقدار خطای برآوردگر وابسته به دو معیار اریبی و واریانس است. به عبارتی میانگین توان دوم خطای (MSE) برآوردگر $\hat{\beta}$ در برآورد پارامتر β بصورت زیر تعریف می‌شود

$$MSE(\hat{\beta}) = E(\hat{\beta} - \beta)^2 = Var(\hat{\beta}) + [E(\hat{\beta} - \beta)]^2$$

همواره به دنبال برآوردگری می‌گردیم که کمترین واریانس را در بین برآوردگرهای نااریب داشته باشد.

حال مدل رگرسیونی زیر را در نظر بگیرید

$$Y = X\beta + E$$

که در آن $Y_{n \times 1}$ برداری از مقادیر پاسخ، $X_{n \times p}$ ماتریس طرح با رتبه کامل p ، $\beta_{p \times 1}$ برداری از ضرایب مجهول رگرسیونی و $E_{n \times 1}$ بردار مولفه خطای تصادفی مدل است. بنا به قضیه گوس-مارکف^۱، برآوردگر کمترین توان‌های دوم (LSE) پارامتر β عبارتست از

$$\hat{\beta}_{LS} = (X'X)^{-1} X'Y$$

بهترین برآوردگر خطی نااریب ($BLUE$) می‌باشد. لازم به ذکر است که این برآوردگر، به روش کمترین توان‌های دوم خطا (MSE) بدست می‌آید. اما همانطور که در مقدمه آمد، وقتی که در داده‌ها همخطی وجود داشته باشد، این برآوردگر کارایی لازم را ندارد.

^۱ Gauss- Markov

دو نوع همخطی وجود دارد: همخطی کامل و همخطی ناقص. در همخطی کامل $|X'X|$ صفر است و لذا $(X'X)^{-1}$ وجود ندارد و می‌توان بجای آن از $(X'X)^{-}$ استفاده کرد. در همخطی ناقص، $|X'X|$ نزدیک صفر است و در این حالت $(X'X)^{-1}$ وجود دارد، با این تفاوت که $\hat{\beta}$ رفتار پایداری ندارد. در طبیعت هیچگاه با داده‌هایی که در آن‌ها همخطی کامل وجود داشته باشد، مواجه نمی‌شویم. لذا اغلب مسائلی که با آن‌ها سروکار داریم، داده‌هایی هستند که برای آن‌ها $|X'X|$ نزدیک به صفر است. در این پایان‌نامه، همخطی ناقص مورد بررسی قرار گرفته است.

مسئله همخطی و کاربردی نبودن برآوردگر کمترین توان‌های دوم، منجر به توسعه برآوردگرهای اریب مانند برآوردگر انقباضی استاین (SE)، برآوردگر مولفه‌های اصلی (PCE)، برآوردگر کمترین توان‌های دوم جزئی ($PLSE$)، برآوردگر نوع لویی (LE)، برآوردگر تقریباً ناریب (AUE) و برآوردگر نوع ريج (RE) که برآوردگرهای بهبودیافته را ارائه می‌دهند، شده‌است. برای اطلاعات بیشتر در زمینه برآوردگرهای بهبودیافته، رجوع کنید به آکدنیز و کاکیرانلار^۱ (۱۹۹۵)، لیو^۲ (۲۰۰۳) راداکریشنا رائو و تونتبرگ^۳ (۱۹۹۵)، سوئیندل (۱۹۷۶)^۴. در میان این برآوردگرهای اریب، برآوردگر نوع ريج کاربرد فراوانی دارد و در سال‌های اخیر بسیار مورد توجه قرار گرفته‌است. سارکار^۵ (۱۹۹۲) برآوردگری جدید برای مدل خطی استاندارد که تحت تاثیر همخطی است پیشنهاد کرد و ویژگی این برآوردگر و همچنین شرایط کافی برای برتری این برآوردگر بر برآوردگر کمترین توان‌های دوم محدودشده ($RLSE$) را بررسی کرد.

^۱ Akdenize and Kaciranlar^۲ Liu^۳ Radhakrishna Rao and Tuotenburg^۴ Swindel^۵ Nityananda Sarkar

شای^۱ (۲۰۰۱) پارامترهای رگرسیونی ريج را در مدل رگرسیونی خطی برآورد کرد و برتریش را بر بهترین برآوردگر نااریب خطی ($BLUE$) محدودشده در MSE و ماتریس $MSE(MSEM)$ نشان داد.

گراب^۲ (۲۰۰۳) رگرسیون ريج را برای مدل رگرسیونی خطی محدودشده تعمیم داد و شرایط لازم و کافی برای برتری این برآوردگر بر $RLSE$ را بررسی کرد.

برآوردگر ريج در سال ۱۹۷۰ توسط هورل و کنارد با حل تغییر یافته معادلات نرمال به صورت زیر معرفی

شد

$$(X'X + kI)\hat{\beta} = X'Y$$

$$\hat{\beta} = (X'X + kI)^{-1} X'Y$$

در واقع این معادلات حاصل مینیمم سازی ترکیبی خطی از کمترین توان دوم خطا و محدودیت اعمال شده بر روی فضای پارامتر به فرم $\beta'\beta \leq \phi^2$ است. یعنی مینیمم کردن عبارت زیر

$$F(\beta, \phi) = (Y - X\beta)'(Y - X\beta) + k(\beta'\beta - \phi^2)$$

که در آن $k > 0$ مقداری ثابت و مثبت است که پارامتر ريج نامیده می شود. روش فوق را به دلیل اینکه پایه ریاضی آن شبیه تحلیل ريج که پیشتر توسط هورل و کنارد در سال ۱۹۵۹ برای توصیف رفتار پاسخ درجه ۲ استفاده شده است، می باشد، رگرسیون ريج می نامند. برآوردگر ريج در واقع یک تبدیل خطی $\mathbb{R}^p \rightarrow \mathbb{R}^p$ از برآوردگر کمترین توان های دوم خطاست. در مسائل کاربردی و مدل های اقتصادی، استفاده از اطلاع پیشین در خصوص بعضی از پارامترهای مورد علاقه در مدل آماری تحت بررسی، اغلب منجر به استفاده از روش های پیشرفته و تعمیم یافته ای می شود که با توجه به قطعی یا غیرقطعی بودن این اطلاع، به کار برده می شوند.

^۱ Shi

^۲ Grob

استفاده از اطلاع پیشین معلوم در مدل‌های آماری، معمولاً به صورت یک قید مطرح است که در نهایت به مدل‌های محدودشده یا مقید منتج می‌شود. برآوردگری که براساس مدل‌های محدودشده بدست می‌آید، برآوردگر محدودشده (RE) نامیده می‌شود. اعتبار و کارایی برآوردگرهای محدودشده براساس قید اعمال شده بر روی فضای پارامتر مورد سنجش و بررسی قرار می‌گیرند، در حالی که برآوردگرهای محدودنشده در کل فضای پارامتر بررسی می‌شوند. بنابراین نتایج بدست آمده بر پایه مدل‌های محدودشده و محدودنشده تنها با سنجش درستی و اعتبار محدودیت‌ها و قیود اعمال شده، می‌تواند معقول و کاربردی باشد. باید توجه کرد که انتخاب برآوردگرهای محدودشده در شرایطی که اطلاع پیشین درست نباشد، توجیه‌پذیر نیست.

در بخش اول این فصل، برای مدل رگرسیون چندگانه با محدودیت تساوی خطی در صورت وجود همخطی در ماتریس طرح، برآوردگری پیشنهاد داده شده و در بخش دوم به منظور مقایسه این برآوردگر با برآوردگر کمترین توان‌های دوم محدودشده، خطای آن محاسبه شده است. در بخش سوم، چند مقدار بهینه برای پارامتر رنج معرفی شده و در بخش چهارم، به کمک شبیه‌سازی مونت-کارلو صحت نتایج بدست آمده بررسی و در بالاخره در بخش پنجم، چند مثال عددی برای تشریح کارایی برآوردگر، ارائه شده است. از نتایج این فصل، مقاله نجاریان، آرشی و کیریا (۲۰۱۱) استخراج شده است.

۱.۲ معرفی برآوردگر

مدل خطی زیر را در نظر بگیرید

$$Y = X\beta + E \quad , \quad E \sim N(0, \sigma^2 I) \quad (1.2)$$

که $Y_{n \times 1}$ بردار مشاهدات، $X_{n \times p}$ ماتریس طرح با رتبه کامل p ، $\beta_{p \times 1}$ برداری از ضرایب مجهول رگرسیونی و $E_{n \times 1}$ بردار خطا با $E(E) = 0$ و $Cov(E) = \sigma^2 I$ است. اکنون محدودیت خطی زیر را نیز برای β در

نظر بگیرید

$$R\beta = r \quad (2.2)$$

که $R_{q \times p}$ ماتریس معلوم غیرصفری است که ماتریس محدودیت نامیده می‌شود با $Rank(R) = q < p$ و $r_{q \times 1}$ برداری معلوم است. مدل خطی معرفی شده در معادله (۱.۲) با محدودیت (۲.۲)، مدل خطی محدودشده (RLM) نامیده می‌شود.

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

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

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

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

در این صورت

$$R = \begin{bmatrix} 1 & 1 & 1 \\ 0 & -2 & 1 \\ 2 & 3 & 0 \end{bmatrix}, \quad r = \begin{bmatrix} 1 \\ 3 \\ 0 \end{bmatrix}$$

به منظور یافتن برآوردگر β ، از روشی که توسط رانو در سال ۱۹۹۵ پیشنهاد داده شد، استفاده می‌کنیم. در این روش که تلفیقی از روش نظریه کمترین توان‌های دوم و لاگرانژ است، تابع زیر که تابعی خطی از مجموع توان‌های دوم خطا و محدودیت اعمال شده بر روی مدل است، مینیمم می‌گردد.

$$F(\beta, \lambda) = (Y - X\beta)'(Y - X\beta) + 2\lambda'(R\beta - r)$$

با مشتگیری از $F(\beta, \lambda)$ نسبت به β داریم

$$\frac{\partial F}{\partial \beta} = -\mathbf{X}'\mathbf{Y} + \mathbf{X}'\mathbf{X}\beta + \mathbf{R}'\lambda$$

با حل معادله $\frac{\partial F}{\partial \beta} = 0$ داریم

$$\mathbf{X}'\mathbf{X}\beta = \mathbf{X}'\mathbf{Y} - \mathbf{R}'\lambda$$

$$\begin{aligned} \Rightarrow \hat{\beta}_R &= (\mathbf{X}'\mathbf{X})^{-1}(\mathbf{X}'\mathbf{Y} - \mathbf{R}'\lambda) \\ &= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y} - (\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'\lambda \\ &= \hat{\beta}_{LS} - (\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'\lambda \end{aligned} \quad (۳.۲)$$

از آنجایی که

$$\mathbf{R}\hat{\beta}_R = \mathbf{R}\hat{\beta}_{LS} - \mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'\hat{\lambda} = \mathbf{r}$$

مجدداً با عملیات خبری، رابطه‌های زیر را به دست می‌آوریم

$$\begin{aligned} \Rightarrow \mathbf{R}\hat{\beta}_{LS} &= \mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}'\hat{\lambda} + \mathbf{r} \\ \Rightarrow \hat{\lambda} &= [\mathbf{R}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{R}']^{-1}(\mathbf{R}\hat{\beta}_{LS} - \mathbf{r}) \end{aligned} \quad (۴.۲)$$

بنابراین برای مدل RLM برآوردگر کمترین توان‌های دوم محدودشده ($RLSE$) به صورت زیر است

$$\hat{\beta}_R = \hat{\beta}_{LS} + \mathbf{S}^{-1}\mathbf{R}'(\mathbf{R}\mathbf{S}^{-1}\mathbf{R}')^{-1}(\mathbf{r} - \mathbf{R}\hat{\beta}) \quad (۵.۲)$$

که در آن $S = X'X$ و $\hat{\beta}_{LS} = S^{-1}X'Y$ است. یکی از بهترین روش‌ها برای مقابله با مسئله همخطی، اعمال محدودیت اضافی روی بردار پارامترهاست. سوامی^۱ و همکاران (۱۹۸۷)، مهتا^۲ و سوامی (۱۹۷۷) و ژانگ و یانگ^۳ (۲۰۰۷) روشی را برای برآورد پارامترها ارائه دادند که در آن مجموع توان دوم باقیمانده‌ها تحت دو محدودیت خطی و کروی مینیمم می‌شود، به عبارت دیگر رگرسیون خطی محدودشده تبدیل به مسئله بهینه‌سازی زیر با دو قید می‌شود.

$$\text{Min}(Y - X\beta)'(Y - X\beta)$$

$$\text{s.t. } \beta'\beta = \phi^2 \quad \& \quad R\beta = r$$

حال با استفاده از تلفیق مینیمم توان‌های دوم خطا و روش لاگرانژ، "برآوردگر رنج محدودشده" را برای این مدل بدست می‌آوریم. بدین منظور تابع زیر را که تابعی خطی از مجموع توان‌های دوم خطا و دو محدودیت اعمال شده روی فضای پارامترهاست، مینیمم می‌کنیم.

$$F(\beta, \phi, \lambda) = Y'Y - 2\beta'X'Y + \beta'X'X\beta + k\beta'\beta - k\phi^2 + 2\lambda'R\beta - 2\lambda'r$$

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

$$\frac{\partial F}{\partial \beta} = -2X'Y + 2X'X\beta + 2k\beta + 2R'\lambda = r$$

بنابراین

$$\hat{\beta}_R(k) = (X'X + kI)^{-1}(X'Y - R'\lambda)$$

^۱Swamy

^۲Mehta

^۳Zhang and Yang

حال از آنجایی که $R\beta = r$ داریم

$$R\beta = R(X'X + kI)^{-1}(X'Y - R'\lambda) = r$$

$$\Rightarrow \hat{\lambda} = [R(X'X + kI)^{-1}R']^{-1}[R(X'X + kI)^{-1}X'Y - r]$$

در نهایت، برآوردگر ريج محدودشده را به صورت زیر بدست می‌آوریم

$$\begin{aligned} \hat{\beta}_R(k) &= (X'X + kI)^{-1}X'Y \\ &- R'[R(X'X + kI)^{-1}R']^{-1}[R(X'X + kI)^{-1}X'Y - r] \end{aligned} \quad (۶.۲)$$

حال اگر تعريف كنيم $S_k = X'X + kI$ و $\hat{\beta}(k) = S_k^{-1}X'Y$ می‌توانیم این برآوردگر را به صورت ساده‌تری

بنویسیم

$$\hat{\beta}_R(k) = \hat{\beta}(k) + S_k^{-1}R'(RS_k^{-1}R')^{-1}(r - R\hat{\beta}(k)) \quad (۷.۲)$$

که در آن $\hat{\beta}(k) = (X'X + kI)^{-1}X'Y$ ”برآوردگر رگرسیونی ريج β ” است.

۲.۲ محاسبه خطای برآوردگر

برای هر برآوردگر، میانگین توان دوم خطا (MSE) به صورت زیر بدست می‌آید

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

لم ۱.۲.۲. (ژانگ و یانگ (۲۰۰۷)) اگر β محدودیت خطی (۲.۲) را داشته باشد، آنگاه

$$MSEM(\hat{\beta}_R(k), \beta) = \sigma^2 tr(M_k S M_k) + k^2 tr(M_k \beta' \beta M_k) \quad (۸.۲)$$

که در آن $M_k = S_k^{-1} - S_k^{-1}R'(RS_k^{-1}R')^{-1}RS_k^{-1}$

برهان: ابتدا امید ریاضی برآوردگر ريج را به صورت زیر محاسبه می‌کنیم

$$\begin{aligned}
 E(\hat{\beta}(k)) &= E((X'X + kI)^{-1} X'Y) \\
 &= (X'X + kI)^{-1} X'X\beta \\
 &= (X'X + kI)^{-1} ((X'X + kI) - kI)\beta \\
 &= \beta - k(X'X + kI)^{-1} \beta \\
 &= (I - kS_k^{-1})\beta
 \end{aligned}$$

”برآوردگر ريج محدودشده“ $\hat{\beta}_R(k)$ را به صورت زیر می‌نویسیم

$$\begin{aligned}
 \hat{\beta}_R(k) &= \hat{\beta}(k) + S_k^{-1} R' [RS_k^{-1} R']^{-1} (r - R\hat{\beta}(k)) \\
 &= S_k^{-1} S_k \hat{\beta}(k) + S_k^{-1} R' [RS_k^{-1} R']^{-1} (r - RS_k^{-1} S_k \hat{\beta}(k))
 \end{aligned}$$

حال اگر تعریف کنیم $M_k = S_k^{-1} - S_k^{-1} R' [RS_k^{-1} R']^{-1} RS_k^{-1}$ می‌توانیم بنویسیم

$$\begin{aligned}
 \hat{\beta}_R(k) &= M_k S_k \hat{\beta}(k) + S_k^{-1} R' [RS_k^{-1} R']^{-1} r \\
 &= M_k S_k \hat{\beta}(k) + S_k^{-1} R' [RS_k^{-1} R']^{-1} R\beta, \\
 &= M_k S_k \hat{\beta}(k) + S_k^{-1} R' [RS_k^{-1} R']^{-1} R\beta, - S_k^{-1} S_k \beta, + \beta, \\
 &= M_k S_k \hat{\beta}(k) - M_k S_k \beta, + \beta, \\
 &= M_k S_k (\hat{\beta}(k) - \beta,) + \beta, \\
 &= M_k S_k \hat{\beta}(k) - M_k S_k \beta, + \beta,
 \end{aligned}$$

بنابراین می‌توان امید ریاضی برآوردگر ريج محدودشده را به صورت زیر نوشت

$$\begin{aligned} E(\hat{\beta}_R(k)) &= \mathbf{M}_k \mathbf{S}_k ((\mathbf{I} - k \mathbf{S}_k^{-1}) \beta - \beta_*) + \beta_* \\ &= \mathbf{M}_k \mathbf{S}_k (\beta - \beta_*) + \beta_* - k \mathbf{M}_k \beta \end{aligned}$$

از طرفی، با توجه به اینکه $R\beta = r$ و $R\beta_* = r$ و اینکه $\mathbf{M}_k \mathbf{S}_k (\beta - \beta_*) = \beta - \beta_*$ داریم

$$E(\hat{\beta}_R(k)) = \beta - k \mathbf{M}_k \beta$$

بنابراین می‌توانیم اریبی برآوردگر ريج محدودشده را به صورت زیر بنویسیم

$$E(\hat{\beta}_R(k) - \beta) = -k \mathbf{M}_k \beta$$

از طرفی با توجه به اینکه

$$\text{Cov}(\hat{\beta}_R(k)) = \sigma^2 \mathbf{M}_k \mathbf{X}' \mathbf{X} \mathbf{M}_k$$

لذا بنابه تعریف ماتریس مجموع توان‌های دوم خطا، داریم

$$\begin{aligned} \text{MSEM}(\hat{\beta}_R(k), \beta) &= \text{Cov}(\hat{\beta}_R(k)) + [E(\hat{\beta}_R(k) - \beta)][E(\hat{\beta}_R(k) - \beta)]' \\ &= \sigma^2 \text{tr}(\mathbf{M}_k \mathbf{S} \mathbf{M}_k) + k^2 \text{tr}(\mathbf{M}_k \beta' \beta \mathbf{M}_k) \end{aligned}$$

قضیه ۱.۲.۲. * ماتریس میانگین توان‌های دوم خطای "برآوردگر ريج محدودشده" به صورت زیر است

$$\text{MSEM}(\hat{\beta}_R(k), \beta) = \sigma^2 \sum_{i=1}^p \frac{\lambda_i (\lambda_i + k - r_{ii}^*)^2}{(\lambda_i + k)^4} + k^2 \sum_{i=1}^p \left[\frac{\alpha_i (\lambda_i + k - r_{ii}^*)}{(\lambda_i + k)^2} \right]^2 \quad (9.2)$$

برهان: برای ماتریس نیمه معین مثبت S ماتریس متعامد Γ وجود دارد، به طوری که $S = \Gamma \Lambda \Gamma'$ و

$$S_k = \Gamma(\Lambda + kI)^{-1} \Gamma'.$$

اکنون ملاحظه کنید که

$$\begin{aligned} M_k S M_k &= (S_k^{-1} - S_k^{-1} R' (R S_k^{-1} R')^{-1} R S_k^{-1}) S (S_k^{-1} - S_k^{-1} R' (R S_k^{-1} R')^{-1} R S_k^{-1}) \\ &= (S_k^{-1} S S_k^{-1}) - [(S_k^{-1} R' (R S_k^{-1} R')^{-1} R S_k^{-1} S S_k^{-1}) \\ &\quad (S_k^{-1} S S_k^{-1} R' (R S_k^{-1} R')^{-1} R S_k^{-1})] \\ &\quad + (S_k^{-1} R' (R S_k^{-1} R')^{-1} R S_k^{-1} S S_k^{-1} R' (R S_k^{-1} R')^{-1} R S_k^{-1}) \\ &= (\Lambda + kI)^{-1} \Lambda (\Lambda + kI)^{-1} - [(\Lambda + kI)^{-1} R^* (\Lambda + kI)^{-1} \Lambda (\Lambda + kI)^{-1}] \\ &\quad - [(\Lambda + kI)^{-1} \Lambda (\Lambda + kI)^{-1} R^* (\Lambda + kI)^{-1}] \\ &\quad + [(\Lambda + kI)^{-1} R^* (\Lambda + kI)^{-1} \Lambda (\Lambda + kI)^{-1} R^* (\Lambda + kI)^{-1}] \end{aligned}$$

حال با توجه به اینکه $S_k^{-1} = \Gamma(\Lambda + kI)^{-1} \Gamma'$ داریم

$$\begin{aligned} tr(S_k^{-1} S S_k^{-1}) &= tr[\Gamma(\Lambda + kI)^{-1} \Gamma' \Gamma \Lambda \Gamma' \Gamma(\Lambda + kI)^{-1} \Gamma'] \\ &= tr[(\Lambda + kI)^{-1} \Lambda] \\ &= \sum_{i=1}^p \frac{\lambda_i}{(\lambda_i + k)^2} \end{aligned} \tag{۱۰.۲}$$

و همچنین

$$\begin{aligned}
& tr[S_k^{-\backslash} S S_k^{-\backslash} R' (R S_k^{-\backslash} R')^{-\backslash} R S_k^{-\backslash}] \\
&= tr[\Gamma(R S_k^{-\backslash} R')^{-\backslash} \Gamma' \Gamma \Lambda \Gamma' \Gamma(R S_k^{-\backslash} R')^{-\backslash} \Gamma' (R S_k^{-\backslash} R')^{-\backslash} R \Gamma(R S_k^{-\backslash} R')^{-\backslash} \Gamma'] \\
&= tr[R^* (\Lambda + kI)^{-\backslash} \Lambda] \\
&= \sum_{i=\backslash}^p \frac{r_{ii}^* \lambda_i}{(\lambda_i + k)^{\backslash}} \tag{۱۱.۲}
\end{aligned}$$

به روش مشابه داریم

$$\begin{aligned}
& tr[S_k^{-\backslash} R' (R S_k^{-\backslash} R')^{-\backslash} R S_k^{-\backslash} S S_k^{-\backslash}] \\
&= tr[\Gamma \Lambda \Gamma' \Gamma (\Lambda + kI)^{-\backslash} \Gamma' (R S_k^{-\backslash} R')^{-\backslash} R \Gamma (\Lambda + kI)^{-\backslash} \Gamma' \Gamma \Lambda \Gamma' \Gamma (\Lambda + kI)^{-\backslash} \Gamma'] \\
&= tr[R^* (\Lambda + kI)^{-\backslash} \Lambda] \\
&= \sum_{i=\backslash}^p \frac{r_{ii}^* \lambda_i}{(\lambda_i + k)^{\backslash}} \tag{۱۲.۲}
\end{aligned}$$

$$\begin{aligned}
& tr[S_k^{-\backslash} R' (R S_k^{-\backslash} R')^{-\backslash} R S_k^{-\backslash} S S_k^{-\backslash}] \\
&= tr[\Gamma (\Lambda + kI)^{-\backslash} \Gamma' R' (R S_k^{-\backslash} R')^{-\backslash} R \Gamma (\Lambda + kI)^{-\backslash} \Gamma' \Gamma \Lambda \Gamma \\
&\quad \Gamma (\Lambda + kI)^{-\backslash} \Gamma' R' (R S_k^{-\backslash} R')^{-\backslash} R \Gamma (\Lambda + kI)^{-\backslash} \Gamma'] \\
&= tr[R^* \Lambda (\Lambda + kI)^{-\backslash}] \tag{۱۳.۲}
\end{aligned}$$

که $\mathbf{R}^* = \mathbf{\Gamma}' \mathbf{R}' (\mathbf{R} \mathbf{S}_k^{-1} \mathbf{R}')^{-1} \mathbf{R} \mathbf{\Gamma}$ و $\text{Diag}(\mathbf{R}^*) = (r_{11}^*, r_{22}^*, \dots, r_{pp}^*)$.

با گردآوری معادلات (۱۰.۲) تا (۱۳.۲) به دست می‌آوریم

$$\begin{aligned} \text{tr}(\mathbf{M}_k \mathbf{S} \mathbf{M}_k) &= \sum_{i=1}^p \frac{\lambda_i}{(\lambda_i + k)^2} - \sum_{i=1}^p \frac{r_{ii}^* \lambda_i}{(\lambda_i + k)^3} + \sum_{i=1}^p \frac{r_{ii}^* \lambda_i}{(\lambda_i + k)^3} \\ &= \sum_{i=1}^p \frac{\lambda_i (\lambda_i + k - r_{ii}^*)}{(\lambda_i + k)^3} \end{aligned} \quad (۱۴.۲)$$

حال

$$\begin{aligned} \mathbf{M}_k^* &= (\mathbf{S}_k^{-1} - \mathbf{S}_k^{-1} \mathbf{R}' (\mathbf{R} \mathbf{S}_k^{-1} \mathbf{R}')^{-1} \mathbf{R} \mathbf{S}_k^{-1}) (\mathbf{S}_k^{-1} - \mathbf{S}_k^{-1} \mathbf{R}' (\mathbf{R} \mathbf{S}_k^{-1} \mathbf{R}')^{-1} \mathbf{R} \mathbf{S}_k^{-1}) \\ &= \mathbf{S}_k^{-2} - \mathbf{S}_k^{-1} \mathbf{R}' (\mathbf{R} \mathbf{S}_k^{-1} \mathbf{R}')^{-1} \mathbf{R} \mathbf{S}_k^{-1} \\ &\quad - \mathbf{S}_k^{-1} \mathbf{R}' (\mathbf{R} \mathbf{S}_k^{-1} \mathbf{R}')^{-1} \mathbf{R} \mathbf{S}_k^{-1} + \mathbf{S}_k^{-1} \mathbf{R}' (\mathbf{R} \mathbf{S}_k^{-1} \mathbf{R}')^{-1} \mathbf{R} \mathbf{S}_k^{-1} \mathbf{R}' (\mathbf{R} \mathbf{S}_k^{-1} \mathbf{R}')^{-1} \mathbf{R} \mathbf{S}_k^{-1} \end{aligned}$$

سپس با انجام عملیات جبری داریم

$$\text{tr}(\mathbf{S}_k^{-2}) = \mathbf{\Gamma} (\mathbf{\Lambda} + k \mathbf{I})^{-1} \mathbf{\Gamma}' \mathbf{\Gamma} (\mathbf{\Lambda} + k \mathbf{I})^{-1} \mathbf{\Gamma}' \quad (۱۵.۲)$$

$$\begin{aligned} &\text{tr}[\mathbf{S}_k^{-1} \mathbf{R}' (\mathbf{R} \mathbf{S}_k^{-1} \mathbf{R}')^{-1} \mathbf{R} \mathbf{S}_k^{-1}] \\ &= \text{tr}[\mathbf{\Gamma} (\mathbf{\Lambda} + k \mathbf{I})^{-1} \mathbf{R}^* (\mathbf{\Lambda} + k \mathbf{I})^{-1} \mathbf{\Gamma}] \\ &= \sum_{i=1}^p \frac{r_{ii}^*}{(\lambda_i + k)^3} \end{aligned} \quad (۱۶.۲)$$

$$\begin{aligned}
& tr[S_k^{-\backslash} R'(RS_k^{-\backslash} R')^{-\backslash} RS_k^{-\backslash}] \\
&= [\Gamma(\Lambda + kI)^{-\backslash} \Gamma' R'(RS_k^{-\backslash} R')^{-\backslash} R\Gamma(\Lambda + kI)^{-\backslash} \Gamma' \Gamma(\Lambda + kI)^{-\backslash} \Gamma'] \\
&= \sum_{i=\backslash}^p \frac{r_{ii}^*}{(\lambda_i + k)^{\backslash}} \tag{۱۷.۲}
\end{aligned}$$

$$\begin{aligned}
& tr[S_k^{-\backslash} R'(RS_k^{-\backslash} R')^{-\backslash} RS_k^{-\backslash} R'(RS_k^{-\backslash} R')^{-\backslash} RS_k^{-\backslash}] \\
&= tr[\Gamma(\Lambda + kI)^{-\backslash} \Gamma' R'(RS_k^{-\backslash} R')^{-\backslash} R\Gamma(\Lambda + kI)^{-\backslash} \Gamma'] \\
&\quad \Gamma(\Lambda + kI)^{-\backslash} \Gamma' R'(RS_k^{-\backslash} R')^{-\backslash} R\Gamma(\Lambda + kI)^{-\backslash} \Gamma' \\
&= tr[R^*(\Lambda + kI)^{-\backslash}] \\
&= \sum_{i=\backslash}^p \frac{r_{ii}^*}{(\lambda_i + k)^{\backslash}} \tag{۱۸.۲}
\end{aligned}$$

با تلفیق معادلات (۱۵.۲) تا (۱۸.۲) بدست می‌آوریم

$$\begin{aligned}
tr(M_k^{\backslash}) &= \sum_{i=\backslash}^p \left[\frac{\backslash}{(\lambda_i + k)^{\backslash}} - \frac{\backslash r_{ii}^*}{(\lambda_i + k)^{\backslash}} + \frac{r_{ii}^*}{(\lambda_i + k)^{\backslash}} \right] \\
&= \sum_{i=\backslash}^p \frac{(\lambda_i + k - r_{ii}^*)^{\backslash}}{(\lambda_i + k)^{\backslash}} \tag{۱۹.۲}
\end{aligned}$$

نتیجه موردنظر از لم (۱.۲) و رابطه زیر بدست می‌آید.

$$tr(\alpha' M_k \alpha) = \sum_{i=\backslash}^p \left[\frac{\alpha_i^{\backslash} (\lambda_i + k - r_{ii}^*)^{\backslash}}{(\lambda_i + k)^{\backslash}} \right]$$

که در آن $\alpha = \Gamma'\beta = (\alpha_1, \alpha_2, \dots, \alpha_p)$ بنابراین

$$\begin{aligned} MSEM(\hat{\beta}_R(k), \beta) &= \sigma^2 tr(\mathbf{M}_k \mathbf{S} \mathbf{M}_k) + k^2 tr(\mathbf{M}_k \beta' \beta \mathbf{M}_k) \\ &= \sigma^2 \sum_{i=1}^p \frac{\lambda_i (\lambda_i + k - r_{ii}^*)^2}{(\lambda_i + k)^4} + k^2 \sum_{i=1}^p \left[\frac{\alpha_i (\lambda_i + k - r_{ii}^*)}{(\lambda_i + k)^2} \right]^2 \end{aligned} \quad (20.2)$$

اثبات کامل است. $\square$

۳.۲ برآورد مقادیر k

اکنون برای بررسی کارایی MSE از مقادیر مختلف پارامتر ریح k استفاده می‌کنیم. هاکنینگ^۱ و همکاران (۱۹۷۶) نشان دادند که به ازای k بهینه، برآوردگر رگرسیون ریح محدودشده بر همه برآوردگرهای دیگر در کلاس برآوردگرهای اریب برتری دارد، اگر چه مقدار بهینه k به σ^2 و α_i مجهول وابسته است که باید از داده‌های مشاهده شده، برآورد شوند. یعنی

$$\hat{k} = \frac{\hat{\sigma}^2}{\hat{\alpha}_i^2} \quad (21.2)$$

که $\hat{\sigma}^2 = \frac{\sum_{i=1}^p E_i^2}{n-p}$ برآوردگر میانگین توان دوم خطاست که برآوردگر نارایب σ^2 است و $\hat{\alpha}_i$ ، $\hat{\alpha}$ امین مولفه $\hat{\alpha}$ که برآوردگر نارایب α است. لازم به ذکر است که پارامتر k به روش‌های دیگری نیز برآوردشده که در زیر به صورت اجمالی آن‌ها را معرفی می‌کنیم.

۱. هورل و کنارد (۱۹۷۰) برآوردگر زیر را برای k پیشنهاد دادند

$$\hat{k}_{HK} = \frac{\hat{\sigma}^2}{\hat{\alpha}_{Max}^2} \quad (22.2)$$

^۱Hocking

که در آن $\hat{\alpha}_{Max}^2$ بزرگترین عضو $\hat{\alpha}$ است. اگر σ^2 و α معلوم باشند، برآوردگر ريج به ازای $\hat{k}_{HK}$ خطای کمتری از LSE دارد.

۲. هورل، کنار و بالدوین^۲ (۱۹۷۵) پیشنهاد دادند که از برآوردگر زیر برای محاسبه k استفاده شود

$$\hat{k}_{HKB} = \frac{p\hat{\sigma}^2}{\sum_{i=1}^p \hat{\alpha}_i^2} = \frac{p\hat{\sigma}^2}{\hat{\alpha}'\hat{\alpha}} = \frac{p\hat{\sigma}^2}{\hat{\beta}'\hat{\beta}} \quad (23.2)$$

که در آن $\hat{\beta}$ برآوردگر کمترین توان‌های دوم پارامتر β است.

۳. براساس برآوردگر نقطه‌ای بیز، لاولس و وانگ^۱ (۱۹۷۶) برآوردگر زیر را برای k پیشنهاد دادند

$$\hat{k}_{LW} = \frac{p\hat{\sigma}^2}{\sum_{i=1}^p \lambda_i \hat{\alpha}_i^2} = \frac{p\hat{\sigma}^2}{\hat{\alpha}'\mathbf{X}'\mathbf{X}\hat{\alpha}} = \frac{p\hat{\sigma}^2}{\hat{\beta}'\mathbf{\Lambda}\hat{\beta}} \quad (24.2)$$

۴. هوکینگ، اسپید و لین^۲ (۱۹۷۶) نیز برآوردگر زیر را معرفی نمودند

$$\hat{k}_{HSL} = \hat{\sigma}^2 \frac{\sum_{i=1}^p (\lambda_i \hat{\alpha}_i)^2}{(\sum_{i=1}^p \lambda_i \hat{\alpha}_i^2)^2} \quad (25.2)$$

^۲Baldwin

^۱Lawless and Wang

^۲Hocking, Speed and Lynn

کيبريا^۱ (۲۰۰۳) عملکرد سه برآوردگر ريج را مورد بررسی قرار داد:

الف) ابتدا با استفاده از میانگين حسابی k_i برآوردگر زیر را ارائه نمود

$$\hat{k}_{AM} = \frac{1}{p} \sum_{i=1}^p \frac{\hat{\sigma}^2}{\hat{\alpha}_i^2} \quad (26.2)$$

ب) سپس با استفاده از میانگين هندسی k_i ، برآوردگر جديد زیر را معرفی کرد

$$\hat{k}_{GM} = \frac{\hat{\sigma}^2}{(\prod_{i=1}^p \hat{\alpha}_i^2)^{1/p}} \quad (27.2)$$

ج) و سرانجام پیشنهاد داد که برای برآوردکردن k از میانه k_i استفاده کنیم

$$\hat{k}_{MED} = \text{Median} \left(\frac{\hat{\sigma}^2}{\hat{\alpha}_i^2} \right), \quad i = 1, 2, \dots, p \quad \forall p \geq 1 \quad (28.2)$$

برآوردگرهای رگرسیونی ريج محدودشده از جایگذاری $\hat{k}$ های فوق به دست می‌آیند

$$\hat{\beta}_R(k) = (X'X + kI)^{-1} X'Y - R'[R(X'X + kI)^{-1} R']^{-1} [R(X'X + kI)^{-1} X'Y - r]$$

که در آن $\hat{k}$ یکی از برآوردگرهای $\hat{k}_{MED}, \dots, \hat{k}_{HK}$ است.

۴.۲ شبیه‌سازی

همخطی در ماتریس‌های طرح، متفاوت است. برای داشتن داده‌هایی با میزان همخطی مختلف، مک‌دونالد و

گالارنیو^۲ (۱۹۷۵) و گیبونز^۳ (۱۹۸۱) پیشنهاد دادند که داده‌ها را به روش زیر تولید کنیم

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

^۱ Kibria

^۲ Mc Donald and Galarneau

^۳ Gibbons

که در آن z_{ij} اعداد تصادفی مستقل از توزیع نرمال استاندارد و γ^2 همبستگی بین هر دو متغیر توضیحی است. این متغیرها استاندارد شده‌اند، به طوری که $X'X$ و $X'Y$ همبسته هستند. حال با استفاده از رابطه زیر، n مشاهده برای متغیر وابسته تعیین می‌شود

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_p x_{ip} + e_i \quad , \quad i = 1, 2, \dots, n$$

که در آن $\beta_0 = 0$ و e_i اعداد تصادفی مستقل از توزیع نرمال با میانگین صفر و واریانس σ^2 هستند. نیوهوس و اومن^۱ (۱۹۷۱) بیان کردند: در صورتی که میانگین توان دوم خطا، تابعی از β ، σ^2 و k باشد و متغیرهای توضیحی ثابت باشند، اگر β بردار ویژه نرمال شده متناظر با بزرگترین مقدار ویژه ماتریس $X'X$ تحت قید $\beta'\beta = 1$ باشد، آنگاه MSE مینیمم می‌شود. بنابراین از بردار ویژه متناظر با بزرگترین مقدار ویژه ماتریس $X'X$ برای برآورد ضرایب $\beta_1, \beta_2, \dots, \beta_p$ در شبیه‌سازی مقادیر متغیر پاسخ استفاده می‌شود. لازم به ذکر است که استفاده از بردار ویژه متناظر با کوچکترین مقدار ویژه، نتایج آنچنان متفاوتی نسبت به حالت قبل ارائه نمی‌دهد. برای انجام شبیه‌سازی، مجموعه‌های مختلفی از داده‌ها را به ازای همبستگی‌های متفاوت $\gamma = 0/70, 0/80, 0/90, 0/99$ که در نتیجه همبستگی‌های متفاوتی دارند، تولید نموده و سپس میزان خطای آن‌ها را با برآوردگر کمترین توان‌های دوم معمولی مقایسه می‌کنیم. حجم نمونه را ۱۵ و ۲۵ و تعداد متغیرهای توضیحی را ۴ و ۱۰ در نظر گرفته و به ازای ۸ مقدار σ ، برآوردگرهای k ارائه شده در بخش قبل را محاسبه نمودیم. مقادیر ویژه و ضرایب رگرسیونی $X'X$ برای مقادیر مختلف n, p, γ و ρ^2 در جدول (۱.۲) درج شده‌اند. رابطه زیر بین σ^2 و نسبت اختلال یعنی ρ^2 برقرار است.

$$\rho^2 = \frac{\beta'\beta}{\sigma^2} \quad (29.2)$$

^۱Newhouse and Oman

جدول ۱.۲: مقادیر β و λ به ازای مقادیر مختلف $\square$ و $\square$ به کار رفته در شبیه‌سازی

γ		۰.۷۰	۰.۸۰	۰.۹۰	۰.۹۹	۰.۷۰	۰.۸۰	۰.۹۰	۰.۹۹
n		n=۱۵				n=۲۵			
λ_1		۱۱,۴۸۷۴	۱۲,۴۳۳۳	۱۵,۳۰۲۰	۲۱,۸۹۶۶	۱۲۴,۴۳۵۶	۱۴۳,۶۸۱۳	۱۶۲,۶۹۴۷	۱۷۶,۸۴۳۶
λ_2		۵,۶۶۷۹	۶,۲۸۱۰	۶,۷۱۴۰	۷,۷۹۹۸	۲۵,۵۱۴۷	۱۸,۲۷۷۸	۹,۸۳۹۲	۱,۰۵۵۱
λ_3		۳,۵۲۸۳	۳,۶۵۵۹	۲,۵۷۴۰	۰,۳۳۰۵	۱۲,۲۴۸۰	۹,۵۸۴۲	۵,۹۷۴۱	۰,۷۳۰۵
λ_4		۲,۴۱۷۹	۲,۲۸۳۳	۱,۷۰۱۶	۰,۲۵۲۰	۱۱,۲۰۰۲	۸,۵۲۱۲	۴,۷۰۲۹	۰,۵۲۰۱
λ_5						۹,۸۴۹۹	۷,۰۸۸۷	۴,۱۳۶۹	۰,۴۸۵۵
λ_6						۹,۱۰۸۸	۶,۵۰۱۹	۳,۳۶۵۴	۰,۳۴۹۲
λ_7						۶,۰۴۹۷	۴,۸۵۷۰	۲,۶۹۹۹	۰,۳۰۲۹
λ_8						۴,۳۴۸۴	۳,۸۴۹۵	۲,۴۶۶۰	۰,۲۷۲۸
λ_9						۳,۵۸۰۲	۲,۵۷۷۴	۱,۳۷۶۳	۰,۱۴۶۱
λ_{10}						۲,۶۲۱۴	۱,۹۲۴۴	۰,۹۳۷۱	۰,۰۷۱۴
β_1		۰,۷۹۹۰	۰,۹۱۶۴	۰,۹۹۰۶	۰,۹۹۷۶	۰,۳۲۳۰	۰,۳۳۷۹	۰,۲۵۳۴	۰,۳۷۱۹
β_2		۰,۴۹۲۹	۰,۳۱۹۳	۰,۰۹۳۵	-۰,۰۶۶۸	-۰,۰۵۹۴	-۰,۰۵۳۴	-۰,۰۴۱۱	-۰,۰۱۴۰
β_3		-۰,۳۴۲۴	-۰,۲۲۶۱	-۰,۰۶۸۵	۰,۰۰۳۵	۰,۰۷۱۳	۰,۰۶۶۳	۰,۰۵۳۱	۰,۰۱۹۲
β_4		۰,۰۳۸۳	۰,۰۸۴۱	۰,۰۷۲۷	۰,۰۱۸۶	۰,۱۲۴۲	۰,۱۰۷۴	۰,۰۷۸۶	۰,۰۲۴۶
β_5						۰,۲۶۸۲	۰,۲۹۱۴	۰,۳۱۸۸	۰,۳۵۹۹
β_6						۰,۴۱۶۱	۰,۴۰۶۹	۰,۳۹۷۸	۰,۳۸۴۵
β_7						۰,۲۷۷۷	۰,۳۰۱۴	۰,۳۲۷۷	۰,۳۶۳۸
β_8						۰,۲۹۶۷	۰,۳۱۶۱	۰,۳۳۷۹	۰,۳۶۶۹
β_9						۰,۳۰۱۵	۰,۳۱۸۸	۰,۳۳۸۹	۰,۳۶۶۹
β_{10}						۰,۶۰۸۹	۰,۵۶۹۳	۰,۵۱۷۹	۰,۴۲۵۹

مقادیر ρ^2 متناظر با σ^2 به ترتیب ۱۰۰۰۰، ۱۰۰، ۱۶، ۴، ۱، ۰/۰۶۲۵، ۰/۰۱۲۳، ۰/۰۰۲۵ هستند. برای

محاسبه خطای برآوردگرها ابتدا ۲۰۰۰ خطا را محاسبه نموده و میانگین خطاها را از رابطه زیر بدست آوردیم

$$MSE(\hat{\beta}(k)) = \frac{1}{2000} \sum_{i=1}^{2000} (\hat{\beta}(k) - \beta)'(\hat{\beta}(k) - \beta) \quad (30.2)$$

برای انجام شبیه‌سازی از نرم‌افزار R استفاده شده است. نتایج شبیه‌سازی در جداول (۱.۲) تا (۹.۲) ارائه شده‌اند.

ρ^*	HK	HKB	LW	HSL	AM	GM	MED
1.000	0.0014	0.0014	0.0014	0.0014	0.0013	0.0014	0.0013
1.0	0.0014	0.0013	0.0014	0.0014	0.0006	0.0008	0.0006
1.6	0.0012	0.0009	0.0012	0.0013	0.0005	0.0006	0.0005
4	0.0009	0.0007	0.0009	0.0012	0.0005	0.0005	0.0005
1	0.0007	0.0006	0.0007	0.0009	0.0005	0.0005	0.0005
0.625	0.0005	0.0005	0.0005	0.0006	0.0005	0.0005	0.0005
0.123	0.0005	0.0005	0.0005	0.0006	0.0005	0.0005	0.0005
0.025	0.0005	0.0005	0.0005	0.0005	0.0005	0.0005	0.0005

[illegible]

ρ^*	HK	HKB	LW	HSL	AM	GM	MED
1.000	0.0015	0.0015	0.0015	0.0015	0.0014	0.0015	0.0014
1.0	0.0015	0.0014	0.0014	0.0015	0.0006	0.0008	0.0007
16	0.0013	0.0010	0.0013	0.0014	0.0006	0.0006	0.0006
4	0.0010	0.0007	0.0010	0.0011	0.0005	0.0006	0.0005
1	0.0007	0.0006	0.0007	0.0008	0.0005	0.0005	0.0005
0.625	0.0006	0.0005	0.0006	0.0006	0.0005	0.0005	0.0005
0.123	0.0005	0.0005	0.0005	0.0005	0.0005	0.0005	0.0005
0.025	0.0005	0.0005	0.0005	0.0005	0.0005	0.0005	0.0005

[illegible]

جدول ۶.۲: مقادیر MSE برآورد شده با $n = ۲۵$, $p = ۱۰$ و $\gamma = ۰/۷۰$

ρ^2	HK	HKB	LW	HSL	AM	GM	MED
۱۰۰۰۰	۰/۰۰۱۰	۰/۰۰۱۰	۰/۰۰۱۰	۰/۰۰۱۰	۰/۰۰۰۹	۰/۰۰۰۷	۰/۰۰۱۰
۱۰۰	۰/۰۰۱۰	۰/۰۰۱۰	۰/۰۰۱۰	۰/۰۰۱۰	۰/۰۰۰۱	۰/۰۰۰۱	۰/۰۰۱۰
۱۶	۰/۰۰۰۹	۰/۰۰۰۹	۰/۰۰۱۰	۰/۰۰۰۹	۰/۰۰۰۱	۰/۰۰۰۳	۰/۰۰۰۹
۴	۰/۰۰۰۷	۰/۰۰۰۸	۰/۰۰۱۰	۰/۰۰۰۷	۰/۰۰۰۳	۰/۰۰۰۴	۰/۰۰۰۶
۱	۰/۰۰۰۴	۰/۰۰۰۵	۰/۰۰۱۰	۰/۰۰۰۳	۰/۰۰۰۴	۰/۰۰۰۵	۰/۰۰۰۳
۰/۰۶۲۵	۰/۰۰۰۰	۰/۰۰۰۰	۰/۰۰۰۹	۰/۰۰۰۱	۰/۰۰۰۵	۰/۰۰۰۵	۰/۰۰۰۱
۰/۰۱۲۳	۰/۰۰۰۳	۰/۰۰۰۲	۰/۰۰۰۶	۰/۰۰۰۳	۰/۰۰۰۵	۰/۰۰۰۵	۰/۰۰۰۳
۰/۰۰۲۵	۰/۰۰۰۴	۰/۰۰۰۴	۰/۰۰۰۲	۰/۰۰۰۴	۰/۰۰۰۵	۰/۰۰۰۵	۰/۰۰۰۴

جدول ۷.۲: مقادیر MSE برآورد شده با $n = ۲۵$, $p = ۱۰$ و $\gamma = ۰/۸۰$

ρ^2	HK	HKB	LW	HSL	AM	GM	MED
۱۰۰۰۰	۰/۰۰۰۹	۰/۰۰۰۹	۰/۰۰۰۹	۰/۰۰۰۹	۰/۰۰۰۸	۰/۰۰۰۶	۰/۰۰۰۹
۱۰۰	۰/۰۰۰۹	۰/۰۰۰۹	۰/۰۰۰۹	۰/۰۰۰۹	۰/۰۰۰۰	۰/۰۰۰۰	۰/۰۰۰۹
۱۶	۰/۰۰۰۸	۰/۰۰۰۸	۰/۰۰۰۹	۰/۰۰۰۷	۰/۰۰۰۱	۰/۰۰۰۳	۰/۰۰۰۷
۴	۰/۰۰۰۵	۰/۰۰۰۶	۰/۰۰۰۹	۰/۰۰۰۵	۰/۰۰۰۳	۰/۰۰۰۴	۰/۰۰۰۴
۱	۰/۰۰۰۲	۰/۰۰۰۳	۰/۰۰۰۹	۰/۰۰۰۲	۰/۰۰۰۴	۰/۰۰۰۵	۰/۰۰۰۲
۰/۰۶۲۵	۰/۰۰۰۰	۰/۰۰۰۰	۰/۰۰۰۷	۰/۰۰۰۰	۰/۰۰۰۴	۰/۰۰۰۵	۰/۰۰۰۰
۰/۰۱۲۳	۰/۰۰۰۳	۰/۰۰۰۲	۰/۰۰۰۴	۰/۰۰۰۳	۰/۰۰۰۵	۰/۰۰۰۵	۰/۰۰۰۳
۰/۰۰۲۵	۰/۰۰۰۴	۰/۰۰۰۴	۰/۰۰۰۱	۰/۰۰۰۴	۰/۰۰۰۵	۰/۰۰۰۵	۰/۰۰۰۴

جدول ۸.۲: مقادیر MSE برآورد شده با $n = ۲۵$, $p = ۱۰$ و $\gamma = ۰/۹۰$

ρ^2	HK	HKB	LW	HSL	AM	GM	MED
۱۰۰۰۰	۰/۰۰۰۸	۰/۰۰۰۸	۰/۰۰۰۸	۰/۰۰۰۸	۰/۰۰۰۶	۰/۰۰۰۳	۰/۰۰۰۸
۱۰۰	۰/۰۰۰۷	۰/۰۰۰۷	۰/۰۰۰۸	۰/۰۰۰۷	۰/۰۰۰۰	۰/۰۰۰۰	۰/۰۰۰۷
۱۶	۰/۰۰۰۵	۰/۰۰۰۶	۰/۰۰۰۸	۰/۰۰۰۵	۰/۰۰۰۱	۰/۰۰۰۳	۰/۰۰۰۵
۴	۰/۰۰۰۲	۰/۰۰۰۳	۰/۰۰۰۸	۰/۰۰۰۲	۰/۰۰۰۳	۰/۰۰۰۴	۰/۰۰۰۲
۱	۰/۰۰۰۰	۰/۰۰۰۰	۰/۰۰۰۸	۰/۰۰۰۰	۰/۰۰۰۴	۰/۰۰۰۵	۰/۰۰۰۰
۰/۰۶۲۵	۰/۰۰۰۰	۰/۰۰۰۰	۰/۰۰۰۵	۰/۰۰۰۰	۰/۰۰۰۵	۰/۰۰۰۵	۰/۰۰۰۰
۰/۰۱۲۳	۰/۰۰۰۳	۰/۰۰۰۲	۰/۰۰۰۲	۰/۰۰۰۳	۰/۰۰۰۵	۰/۰۰۰۵	۰/۰۰۰۳
۰/۰۰۲۵	۰/۰۰۰۴	۰/۰۰۰۴	۰/۰۰۰۰	۰/۰۰۰۴	۰/۰۰۰۵	۰/۰۰۰۵	۰/۰۰۰۴

جدول ۹.۲: مقادیر MSE برآورد شده با $n = ۲۵$, $p = ۱۰$ و $\gamma = ۰/۹۹$

ρ^2	HK	HKB	LW	HSL	AM	GM	MED
۱۰۰۰۰	۰/۰۰۰۰	۰/۰۰۰۰	۰/۰۰۰۰	۰/۰۰۰۰	۰/۰۰۰۰	۰/۰۰۰۰	۰/۰۰۰۰
۱۰۰	۰/۰۰۰۰	۰/۰۰۰۰	۰/۰۰۰۰	۰/۰۰۰۰	۰/۰۰۰۰	۰/۰۰۰۰	۰/۰۰۰۰
۱۶	۰/۰۰۰۰	۰/۰۰۰۰	۰/۰۰۰۰	۰/۰۰۰۰	۰/۰۰۰۱	۰/۰۰۰۳	۰/۰۰۰۰
۴	۰/۰۰۰۰	۰/۰۰۰۰	۰/۰۰۰۰	۰/۰۰۰۰	۰/۰۰۰۳	۰/۰۰۰۴	۰/۰۰۰۰
۱	۰/۰۰۰۰	۰/۰۰۰۰	۰/۰۰۰۰	۰/۰۰۰۰	۰/۰۰۰۴	۰/۰۰۰۵	۰/۰۰۰۰
۰/۰۶۲۵	۰/۰۰۰۰	۰/۰۰۰۰	۰/۰۰۰۰	۰/۰۰۰۰	۰/۰۰۰۵	۰/۰۰۰۵	۰/۰۰۰۰
۰/۰۱۲۳	۰/۰۰۰۲	۰/۰۰۰۲	۰/۰۰۰۰	۰/۰۰۰۳	۰/۰۰۰۵	۰/۰۰۰۵	۰/۰۰۰۳
۰/۰۰۲۵	۰/۰۰۰۴	۰/۰۰۰۴	۰/۰۰۰۰	۰/۰۰۰۴	۰/۰۰۰۵	۰/۰۰۰۵	۰/۰۰۰۴

۵.۲ کاربرد

در این بخش مثالهایی واقعی، برای تشریح کارایی برآوردگر پیشنهاد داده شده، ارائه شده است.

۱.۵.۲ مثال ۱

جدول (۱۰.۲) داده‌هایی درباره درصد تبدیل n -هپتان به استیلن با سه متغیر توضیحی را نشان می‌دهد (هیمل بلا^۱ (۱۹۷۰)، کوناگی، تامورا و نایتو^۲ (۱۹۶۱) ماکارد و سنی^۳ (۱۹۷۵)). به‌طوری‌که مونت‌گومری و همکارانش (۲۰۰۱) نشان دادند، مدل درجه دوم کامل برای داده‌های استیلن به صورت زیر است

$$P = \beta_0 + \beta_1 T + \beta_2 H + \beta_3 C + \beta_{12} TH + \beta_{13} TC + \beta_{23} HC \\ + \beta_{11} T^2 + \beta_{22} H^2 + \beta_{33} C^2 + E$$

جدول ۱۰.۲: داده‌های استیلن

زمان اتصال	نسبت H_2 به n -هپتان	درجه رآکتور	تبدیل n -هپتان به استیلن	مشاهدات
۰.۱۲۰	۷.۵	۱۳۰۰	۴۹.۰	۱
۰.۱۲۰	۹.۰	۱۳۰۰	۵۰.۲	۲
۰.۱۱۵	۱۱.۰	۱۳۰۰	۵۰.۵	۳
۰.۱۳۰	۱۳.۵	۱۳۰۰	۴۸.۵	۴
۰.۱۳۵	۱۷.۰	۱۳۰۰	۴۷.۵	۵
۰.۱۲۰	۲۳.۰	۱۳۰۰	۲۸.۰	۶
۰.۴۰۰	۵.۳	۱۲۰۰	۳۱.۵	۷
۰.۳۸۰	۷.۵	۱۲۰۰	۳۱.۵	۸
۰.۳۲۰	۱۱.۰	۱۲۰۰	۳۴.۵	۹
۰.۲۶۰	۱۳.۵	۱۲۰۰	۳۵.۰	۱۰
۰.۳۴۰	۱۷	۱۲۰۰	۳۸.۰	۱۱
۰.۴۱۰	۲۳.۰	۱۲۰۰	۳۸.۵	۱۲
۰.۸۴۰	۵.۳	۱۱۰۰	۱۵.۰	۱۳
۰.۹۸۰	۷.۵	۱۱۰۰	۱۷.۰	۱۴
۰.۹۲۰	۱۱.۰	۱۱۰۰	۲۰.۵	۱۵
۰.۸۶۰	۱۷.۰	۱۱۰۰	۲۹.۵	۱۶

^۱ Himmelblau^۲ Kunugi, Tammura and Naito^۳ Marquadt and Snee

همانطور که در مقدمه و پیوست‌های A و B توضیح داده شده است، عدد شرطی و عامل تورم واریانس دو معیار برای سنجش میزان همخطی هستند، لذا برای بررسی وجود همخطی در این مثال‌ها، عدد شرطی و عامل تورم واریانس را محاسبه می‌کنیم. این مقادیر با استفاده از فرمول‌های ارائه‌شده در پیوست‌ها محاسبه شده‌اند و برای بدست آوردن عامل تورم واریانس از نرم افزار $SPSS$ استفاده شده است.

جدول ۱۱.۲: مقادیر عدد شرطی و عوامل تورم واریانس

CN	VIF _۱	VIF _۲	VIF _۳
۳۶۹۳۵/۹۱	۱۲/۱۹۵۱۲	۱/۰۶۱۵۷۱	۱۲/۳۴۵

همانطور که ملاحظه می‌شود، مقدار عدد شرطی بسیار بزرگ است و مقادیر عوامل تورم واریانس برای تک تک متغیرها نیز بزرگتر از یک است که نشان‌دهنده وجود همخطی در داده‌هاست.

فرض کنید ماتریس محدودیت بصورت زیر باشد

$$R = \begin{bmatrix} 2 & 4 & 5 \\ 1 & -2 & 3 \end{bmatrix}, \quad r = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

مقادیر برآوردگرها و میزان خطای آن‌ها به ازای $\hat{k}$ های بهینه مختلف در جدول (۱۲.۲) آورده شده است.

جدول ۱۲.۲: مقایسه برآوردگرها

برآوردگر	RLSE	$\hat{k}_{HK}$	$\hat{k}_{HKB}$	$\hat{k}_{LW}$
MSE	۱۵۳۸۰۵۵۷/۰۰۰۰	۳۰۸/۳۸۴۷۰	۷۴۲۶۵۱/۹۰۰۰	۷۳۴۷۸۹۱/۰۰۰۰
$\hat{\beta}_1$	۰/۰۰۹۹	۰/۰۰۹۹	۰/۸۳۲۶	۰/۸۳۲۶
$\hat{\beta}_2$	۱/۹۸۰۱	۱/۹۸۰۱	۰/۳۳۴۸	۰/۳۳۴۸
$\hat{\beta}_3$	-۰/۹۹۰۰	-۰/۹۹۰۰	-۰/۱۶۷۴	-۰/۱۶۷۴
برآوردگر	$\hat{k}_{HSL}$	$\hat{k}_{AM}$	$\hat{k}_{GM}$	$\hat{k}_{MED}$
MSE	۱۲۶۹۰۰۰/۰۰۰۰	۰/۰۰۰۰	۰/۰۰۰۰	۰/۰۰۰۰
$\hat{\beta}_1$	۰/۸۳۱۱	۰/۸۲۶۸	۰/۶۷۳۷	۰/۸۳۱۱
$\hat{\beta}_2$	۰/۳۳۷۷	۰/۳۴۶۴	۰/۶۵۲۶	۰/۳۳۷۷
$\hat{\beta}_3$	-۰/۱۶۸۸	-۰/۱۷۳۲	-۰/۳۲۶۳	-۰/۱۶۸۹

۲.۵.۲ مثال ۲

جدول زیر، داده‌های سیمان پورتلند را نشان می‌دهد (وود^۱ و همکاران (۱۹۳۲)).

$$X = \begin{bmatrix} 7 & 26 & 6 & 60 \\ 1 & 29 & 15 & 52 \\ 11 & 56 & 8 & 20 \\ 7 & 52 & 6 & 33 \\ 11 & 55 & 9 & 22 \\ 3 & 71 & 17 & 6 \\ 1 & 31 & 22 & 44 \\ 2 & 54 & 18 & 22 \\ 21 & 47 & 4 & 26 \\ 1 & 40 & 23 & 34 \\ 11 & 66 & 9 & 12 \\ 10 & 68 & 8 & 12 \end{bmatrix}, \quad Y = \begin{bmatrix} 78/5 \\ 74/3 \\ 104/3 \\ 87/6 \\ 95/9 \\ 109/2 \\ 102/7 \\ 72/5 \\ 93/1 \\ 115/9 \\ 83/8 \\ 113/3 \\ 109/4 \end{bmatrix}$$

فرض کنید ماتریس محدودیت به صورت زیر باشد

$$R = \begin{bmatrix} 1 & -1 & 1 & 0 \end{bmatrix}, \quad r = 0$$

عدد شرطی و عوامل تورم واریانس برای هر یک از متغیرها عبارتند از

جدول ۱۳.۲: مقادیر عدد شرطی و عوامل تورم واریانس

CN	VIF _۱	VIF _۲	VIF _۳	VIF _۴
۲۰/۵۸۴۶۳	۳۸/۴۶۱۵۴	۲۵۰	۴۷/۶۱۹۰۵	۲۵۰

اگر چه کوچک بودن CN همخطی جدی را نشان نمی‌دهد اما به دلیل بزرگ بودن مقادیر VIF می‌توان وجو

همخطی را در داده‌ها استنباط نمود. این داده‌ها را همانند مثال قبل بررسی کرده و نتایج (۱۴.۲) بدست آمده

است.

^۱Wood

جدول ۱۴.۲: مقایسه برآوردگرها

برآوردگر	RLSE	$\hat{k}_{HK}$	$\hat{k}_{HKB}$	$\hat{k}_{LW}$
MSE	۹۷۵,۸۲۲۴	۰,۶۶۸۰	۰,۶۵۷۶	۷۰,۳۰۲۸
$\hat{\beta}_1$	۰,۸۰۸۷	۰,۲۸۸۸	۰,۲۸۴۹	۰,۶۳۲۶
$\hat{\beta}_2$	۰,۲۸۰۹	۰,۲۸۰۶	۰,۲۸۰۶	۰,۳۰۰۳
$\hat{\beta}_3$	۰,۰۴۸۶	۰,۳۰۵۰	۰,۳۰۷۸	۰,۰۶۲۸
$\hat{\beta}_4$	۰,۰۸۷۶	۰,۲۱۵۸	۰,۲۱۷۲	۰,۰۹۴۷
برآوردگر	$\hat{k}_{HSL}$	$\hat{k}_{AM}$	$\hat{k}_{GM}$	$\hat{k}_{MED}$
MSE	۰,۶۵۹۰	۰,۶۵۶۷	۰,۶۵۷۱	۰,۶۵۶۶
$\hat{\beta}_1$	۰,۲۸۵۷	۰,۲۸۴۰	۰,۲۸۴۴	۰,۲۸۳۹
$\hat{\beta}_2$	۰,۲۸۰۶	۰,۲۸۰۵	۰,۲۸۰۶	۰,۲۸۰۵
$\hat{\beta}_3$	۰,۳۰۷۱	۰,۳۰۸۴	۰,۳۰۸۱	۰,۳۰۸۴
$\hat{\beta}_4$	۰,۲۱۶۹	۰,۲۱۷۵	۰,۲۱۷۳	۰,۲۱۷۵

۳.۵.۲ مثال ۳

در این مثال، داده‌های هزینه‌های تحقیق و تولید ملی خام به عنوان درصدی از تولید ناخالص ملی کشور آمریکا

را در سال‌های ۱۹۸۶-۱۹۷۲ بررسی می‌کنیم. (آکدنیز و ارول)^۱ (۲۰۰۳) و گروبر^۲ (۱۹۹۸).

مقادیر R و r به ترتیب در زیر آمده‌اند

$$R = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 3 & 1 \end{bmatrix}, r = \begin{bmatrix} 1/2170 \\ 1/0904 \end{bmatrix}$$

عدد شرطی و عوامل تورم واریانس به صورت زیر هستند

عدد شرطی نزدیک به ۱۰۰ است که می‌شود وجود همخطی را نتیجه گرفت و عوامل تورم واریانس نیز

حاکمی از همخطی در داده‌ها هستند. مقادیر خطا در جدول (۱۷.۲) درج شده‌اند.

^۱ Akdenize and Erol

^۲ Gruber

جدول ۱۵.۲: داده‌های هزینه‌های تحقیق و تولید ملی خام

سال	Y	X _۱	X _۲	X _۳	X _۴
۱۹۷۲	۲٫۳	۱٫۹	۲٫۲	۱٫۹	۳٫۷
۱۹۷۵	۲٫۲	۱٫۸	۲٫۲	۲	۳٫۸
۱۹۷۹	۲٫۲	۱٫۸	۲٫۴	۲٫۱	۳٫۶
۱۹۸۰	۲٫۳	۱٫۸	۲٫۴	۲٫۲	۳٫۸
۱۹۸۱	۲٫۴	۲	۲٫۵	۲٫۳	۳٫۸
۱۹۸۲	۲٫۵	۲٫۱	۲٫۶	۲٫۴	۳٫۷
۱۹۸۳	۲٫۶	۲٫۱	۲٫۶	۲٫۶	۳٫۸
۱۹۸۴	۲٫۶	۲٫۲	۲٫۶	۲٫۶	۴
۱۹۸۵	۲٫۷	۲٫۳	۲٫۸	۲٫۸	۳٫۷
۱۹۸۶	۲٫۷	۲٫۳	۲٫۷	۲٫۸	۳٫۸

جدول ۱۶.۲: مقادیر عدد شرطی و عوامل تورم واریانس

CN	VIF _۱	VIF _۲	VIF _۳	VIF _۴
۹۳٫۶۸۲۳۴	۶٫۸۹۶۵۵۲	۲۱٫۷۳۹۱۳	۲۹٫۴۱۱۷۶	۱٫۷۹۵۳۳۲

جدول ۱۷.۲: مقایسه برآوردگرها

برآوردگر	RLSE	$\hat{k}_{HK}$	$\hat{k}_{HKB}$	$\hat{k}_{LW}$
MSE	۵۵۳۸۷۷٫۵۰۰	۳۹۸۳٫۳۵۶۰	۴۱۰٫۴۶۳۷	۵۴۶۱۱۲٫۶۰۰
$\hat{\beta}_1$	۱٫۸۱۳۲	۰٫۶۹۰۵	۰٫۵۲۱۹	۱٫۸۱۱۵
$\hat{\beta}_2$	۱٫۲۵۱۰	۰٫۴۱۶۵	۰٫۰۷۰۹	۱٫۲۵۰۹
$\hat{\beta}_3$	۰٫۴۳۷۸	۰٫۷۲۶۰	۰٫۵۴۹۰	۰٫۴۳۹۴
$\hat{\beta}_4$	۰٫۵۴۸۸	۰٫۴۶۸۷	۰٫۲۶۷۵	۰٫۵۴۸۷
برآوردگر	$\hat{k}_{HSL}$	$\hat{k}_{AM}$	$\hat{k}_{GM}$	$\hat{k}_{MED}$
MSE	۱۳٫۵۹۷۴	۳۶٫۸۹۹۶	۵٫۲۳۷۹	۲۰۰٫۲۲۹۵۰
$\hat{\beta}_1$	۰٫۶۳۳۱	۰٫۳۷۶۰	۰٫۴۲۵۶	۰٫۳۹۷۵
$\hat{\beta}_2$	۰٫۳۰۱۶	-۰٫۲۴۰۹	-۰٫۱۳۳۹	-۰٫۱۹۴۵
$\hat{\beta}_3$	۰٫۶۶۸۵	۰٫۳۸۳۱	۰٫۴۴۰۴	۰٫۴۰۸۰
$\hat{\beta}_4$	۰٫۴۰۶۲	۰٫۰۶۳۰	۰٫۱۳۴۷	۰٫۰۹۴۲

۶.۲ نتیجه‌گیری

بر مبنای بررسی شبیه‌سازی دریافتیم که افزایش همبستگی بین متغیرهای مستقل روی MSE اثر منفی دارد.

وقتیکه حجم نمونه افزایش می‌یابد، MSE کاهش می‌یابد حتی اگر همبستگی بین متغیرهای مستقل زیاد

باشد.. در مثال‌های ارائه‌شده نیز خطای برآوردگر رنج به ازای همه k ‌های مختلف کمتر از خطای برآوردگر کمترین توان‌های دوم محدودشده است.

فصل ۳

برآوردگر رگرسیونی ریج منفرد

مقدمه

هنگامی که در مدل رگرسیونی خطی، $Cov(E) = \sigma^2 \Sigma$ ، به منظور یافتن برآوردگر کمترین توان‌های دوم خطا، بایستی معیار کمترین توان دوم خطای موزون را مینیمم کنیم، در عین حال یک روش خوب برای موزون کردن کمترین توان دوم خطا، استفاده از ماتریس $T = \Sigma + XUX'$ به جای Σ است که در آن $U \geq 0$ و

$$Rank(\Sigma; X) = Rank(T)$$

ماتریس T شامل اطلاعاتی از ماتریس واریانس-کوواریانس بردار خطا و همچنین ماتریس داده‌هاست. در این صورت می‌توان از اطلاعات ماتریس X در وزن‌دار کردن معیار کمترین توان‌های دوم استفاده کرد. ماتریس T منفرد است و بنابراین معکوس‌پذیر نمی‌باشد (برای آگاهی بیشتر درباره ماتریس منفرد و خواص آن پیوست B را ببینید)، بنابراین از معکوس تعمیم‌یافته این ماتریس استفاده می‌کنیم.

این فصل اختصاص به یافتن برآوردگر رنج برای مدل خطی منفرد دارد. در بخش اول ابتدا مدل خطی منفرد را معرفی کرده، سپس محدودیتی را روی مدل اعمال کرده و تحت این شرایط، برآوردگری را برای برآورد ضرایب رگرسیونی ارائه می‌دهیم. محدودیت اعمال شده روی مدل رگرسیونی با توجه به توزیع پیشین پارامترها و فرضیه تحت بررسی در مورد داده‌ها، در نظر گرفته می‌شود. در بخش دوم، ویژگی‌های برآوردگر را بررسی می‌کنیم و به منظور بررسی کارایی این برآوردگر، خطای آن را بدست می‌آوریم و در بخش سوم، به کمک شبیه‌سازی مونت-کارلو، برآوردگر رنج محدودشده را با برآوردگر کمترین توان‌های دوم محدودشده مقایسه می‌کنیم. و بالاخره در بخش چهارم، نتیجه‌ای که از این تحقیق گرفته می‌شود، ارائه شده است. لازم به ذکر است که نتایج این فصل برگرفته از منبع ژانگ و یانگ (۲۰۰۷) می‌باشد و نویسنده این مجموعه، تنها برهان ویژگی‌های برآوردگر و قضایای مربوطه را با جزئیات ارائه کرده است و برای فهم بیشتر، مطالب بخش (۴.۳) را به آن اضافه نموده است.

۱.۳ معرفی برآوردگر

مدل خطی منفرد با محدودیت تساوی خطی زیر را در نظر بگیرید

$$\begin{cases} Y = X\beta + E \\ E(E) = 0, \quad Cov(E) = \sigma^2 \Sigma \\ R\beta = 0 \end{cases} \quad (۱.۳)$$

که $Y_{n \times 1}$ برداری از مشاهدات، $X_{n \times p}$ ماتریس طرح با رتبه p ، $\beta_{p \times 1}$ برداری از پارامترهای تحت بررسی، $E_{n \times 1}$ بردار خطای تصادفی و $\Sigma \geq 0$ ماتریس واریانس-کوواریانس E و $R_{q \times p}$ ماتریس محدودیت با رتبه q است.

حال، برآوردگر β را با مینیمم کردن معیار زیر بدست می‌آوریم. در حقیقت می‌توان این‌طور بیان کرد که برآوردگر β از کمینه کردن کمترین توان‌های دوم خطا در یک مدل خطی منفرد بدست می‌آید.

$$F(\beta, \lambda) = (Y - X\beta)'(Y - X\beta) + \lambda' R\beta$$

حال از $F(\beta, \lambda)$ نسبت به β مشتق می‌گیریم

$$\frac{\partial F}{\partial \beta} = -2X'T^{-}Y + 2X'T^{-}X\beta + 2R'\lambda$$

با انجام یک سری عملیات جبری، داریم

$$\begin{aligned} \hat{\beta}_R &= (X'T^{-}X)^{-1}(X'T^{-}Y - R'\lambda) \\ &= (X'T^{-}X)^{-1}X'T^{-}Y - (X'T^{-}X)^{-1}R'\lambda \\ &= \hat{\beta}_{LS} - (X'T^{-}X)^{-1}R'\lambda \end{aligned} \quad (۲.۳)$$

از آنجایی که

$$R\hat{\beta}_R = R\beta - R(X'T^{-1}X)^{-1}R'\lambda = 0$$

مجدداً با انجام عملیات جبری، رابطه‌های زیر را به دست می‌آوریم

$$\begin{aligned} R\hat{\beta}_{LS} &= R(X'T^{-1}X)^{-1}R'\lambda \\ \Rightarrow \lambda &= [R(X'T^{-1}X)^{-1}R']^{-1}R\hat{\beta}_{LS} \end{aligned} \quad (۳.۳)$$

سرانجام با قرار دادن معادله (۳.۳) در (۲.۳)، برآوردگر مقید در مدل رگرسیونی خطی منفرد را به صورت زیر بدست می‌آوریم

$$\hat{\beta}_R = \hat{\beta}_{LS} - (X'T^{-1}X)^{-1}R'[R(X'T^{-1}X)^{-1}R']^{-1}R\hat{\beta}_{LS} \quad (۴.۳)$$

حال اگر قرار دهیم $S = XT^{-1}X'$ و $W = S^{-1} - S^{-1}R'(RS^{-1}R')^{-1}RS^{-1}$ می‌توانیم عبارت ساده‌تری را برای $\hat{\beta}_R$ به صورت زیر بنویسیم.

$$\hat{\beta}_R = WX'T^{-1}Y \quad (۵.۳)$$

که $\hat{\beta}_R$ را "برآوردگر خطی مقید" برای β می‌نامند.

اگر در مدل، همخطی وجود داشته باشد، معادله (۵.۳) اصلاح و "برآوردگر ریج مقید" به صورت زیر پیشنهاد داده می‌شود.

$$\hat{\beta}_R(k) = (kW + I)^{-1}\hat{\beta}_R \quad (۶.۳)$$

که $k \geq 0$ ، مقداری ثابت است (k پارامتر ریج مقید نامیده می‌شود). به ازای مقادیر مختلف k ، می‌توانیم برآوردهای ریج متفاوتی را برای پارامترهای مدل رگرسیونی به دست آوریم. به خصوص به ازای $k = 0$ ، برآوردگر ریج مقید β به بهترین برآوردگر ناریب خطی مقید β تبدیل می‌شود. یعنی $\hat{\beta}_R(0) = \hat{\beta}_R$.

۲.۳ ویژگی‌های برآوردگر

به منظور ساده کردن محاسبه MSE ، ابتدا فرم کانونی مدل رگرسیون خطی منفرد با محدودیت تساوی خطی زیر را ارائه می‌دهیم

$$\begin{cases} Y = Z\alpha + E \\ E(E) = 0, \quad Cov(E) = \sigma^2 \Sigma \\ L\alpha = 0 \end{cases} \quad (7.3)$$

که $\alpha = Q'\beta$ ، $L = RQ$ ، $Z = XQ$ پارامتر کانونی شرطی نامیده می‌شود، Q ماتریسی متعامد است به‌طوری‌که $\Lambda = Q'WQ = Diag(\lambda_1, \lambda_2, \dots, \lambda_p)$ و $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p$ مقادیر ویژه ترتیبی W هستند.

در این صورت، داریم

$$\begin{aligned} \hat{\alpha} &= Q'\hat{\beta} \\ &= (Z'T^-Z)^{-1} - (Z'T^-Z)^{-1}L'[L(Z'T^-Z)^{-1}L']^{-1}L(Z'T^-Z)^{-1}Z'T^-Y \\ &= \Lambda Z'T^-Y \end{aligned}$$

که در آن $\Lambda = (Z'T^-Z)^{-1} - (Z'T^-Z)^{-1}L'[L(Z'T^-Z)^{-1}L']^{-1}L(Z'T^-Z)^{-1}$

توجه کنید که

$$\hat{\alpha}(k) = (k\Lambda + I)^{-1}\hat{\alpha} \quad (8.3)$$

واضح است که رابطه‌های زیر بین برآوردگرها و فرم کانونیشان برقرار است

$$\hat{\beta}_R = Q\hat{\alpha} \quad , \quad \hat{\beta}_R(k) = Q\hat{\alpha}(k) \quad (۹.۳)$$

همچنین روابط زیر بین MSE برآوردگرها و MSE شکل کانونی آنها برقرار است

$$MSE(\hat{\beta}_R) = MSE(\hat{\alpha}) \quad , \quad MSE(\hat{\beta}_R(k)) = MSE(\hat{\alpha}(k)) \quad (۱۰.۳)$$

لم ۱.۲.۳. $Rank(W) = p - q$ و $W = S^{-1} - S^{-1}R'(RS^{-1}R')^{-1}RS^{-1}$

برهان: از تعریف W داریم

$$\begin{aligned} WSW' &= [S^{-1} - S^{-1}R'(RS^{-1}R')^{-1}RS^{-1}]S[S^{-1} - S^{-1}R'(RS^{-1}R')^{-1}RS^{-1}] \\ &= S^{-1}SS^{-1} - S^{-1}SS^{-1} - S^{-1}SS^{-1}R'(RS^{-1}R')^{-1}RS^{-1} \\ &\quad - S^{-1}R'(RS^{-1}R')^{-1}RS^{-1}SS^{-1} \\ &\quad + S^{-1}R'(RS^{-1}R')^{-1}RS^{-1} - S^{-1}R'(RS^{-1}R')^{-1}RS^{-1} \\ &= S^{-1} - S^{-1}R'(RS^{-1}R')^{-1}RS^{-1} - S^{-1}R'(RS^{-1}R')^{-1}RS^{-1} \\ &\quad + S^{-1}R'(RS^{-1}R')^{-1}RS^{-1} \\ &= S^{-1} - S^{-1}R'(RS^{-1}R')^{-1}RS^{-1} \\ &= W \end{aligned}$$

از آنجایی که $S = X'T^{-1}X \geq 0$ و با توجه به لم (۱.۲) نتیجه می‌شود که $W \geq 0$.

برای اثبات $Rank(W) = p - q$ ، با توجه به این که $R_{q \times p}$ ماتریسی با $Rank(R) = q$ است. بنابراین

ماتریس وارنپذیر $A_{p \times p}$ وجود دارد به طوری که $A = (0, \dots, 0, 1, \dots, 1)_{q \times q}$ و $R =$

حال $AS^{-1}A'$ را به صورت زیر افراز می‌کنیم

$$AS^{-1}A' = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$$

که در آن A_{11} ماتریس $q \times q$ است. بنابراین داریم

$$\begin{aligned} AS^{-1}A' &= AS^{-1}A' - AS^{-1}R'(RS^{-1}R')^{-1}RS^{-1}R' \\ &= \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} - \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} I_{q \times q} \\ 0 \end{bmatrix} \\ &\times \left(\begin{bmatrix} I_{q \times q} & 0 \end{bmatrix} \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} I_{q \times q} \\ 0 \end{bmatrix}^{-1} \right) \times \begin{bmatrix} I_{q \times q} & 0 \end{bmatrix} \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \\ &= \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} - \begin{bmatrix} A_{11} \\ A_{21} \end{bmatrix} A_{11}^{-1} \begin{bmatrix} A_{11} & A_{12} \end{bmatrix} \\ &= \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} - \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{21}A_{11}^{-1}A_{12} \end{bmatrix} \\ &= \begin{bmatrix} 0 & 0 \\ 0 & A_{22} - A_{21}A_{11}^{-1}A_{12} \end{bmatrix} \end{aligned}$$

از آنجایی که $A_{21}A_{11}^{-1}A_{12}$ معکوس‌پذیر است، داریم

$$Rank(W) = R(A_{22} - A_{21}A_{11}^{-1}A_{12}) = p - q$$

لم ۲.۲.۳. در معادله (۷.۳)، داریم $\Lambda = Diag(\lambda_1, \lambda_2, \dots, \lambda_p)$ ، بنابراین

$$\lambda_{p-q+1} = \lambda_{p-q+2} = \dots = \lambda_p = 0$$

برهان: سارکار (۱۹۹۶) را ببینید.

لم ۳.۲.۳. در معادله (۷.۳)، داریم $\alpha = Q'\beta = (\alpha_1, \alpha_2, \dots, \alpha_p)'$ ، بنابراین

$$\alpha_{p-q+1} = \alpha_{p-q+2} = \dots = \alpha_p$$

برهان: سارکار (۱۹۹۶) را ببینید.

لم ۴.۲.۳. در معادله (۷.۳)، داریم $\hat{\alpha} = Q'\hat{\beta}_R = (\hat{\alpha}_1, \hat{\alpha}_2, \dots, \hat{\alpha}_p)'$ بنابراین

$$\hat{\alpha}_{p-q+1} = \hat{\alpha}_{p-q+2} = \dots = \hat{\alpha}_p$$

برهان: سارکار (۱۹۹۶) را ببینید.

ویژگی ۱.۲.۳. $\hat{\beta}_R(k)$ برآوردگر اریب و انقباضی β است.

برهان: برای اثبات انقباضی بودن برآوردگر رنج مقید، باید نشان دهیم که مربع نرم این برآوردگر از مربع

نرم برآوردگر نااریب خطی مقید کمتر است. با استفاده از فرمول برآوردگر رنج مقید و رابطه بین Λ و W

داریم

$$\begin{aligned} \|\hat{\beta}_R(k)\|^2 &= \|(k\mathbf{W} + \mathbf{I})^{-1}\hat{\beta}_R\|^2 \\ &= \|Q(k\Lambda + \mathbf{I})^{-1}Q'\hat{\beta}_R\|^2 \\ &= \|(k\Lambda + \mathbf{I})^{-1}Q'\hat{\beta}_R\|^2 \\ &< \|Q'\hat{\beta}_R\|^2 = \|\hat{\beta}_R\|^2 \end{aligned}$$

بنابراین $\hat{\beta}_R(k)$ برآوردگری انقباضی برای β است. حال برای اثبات اریب بودن برآوردگر رنج مقید، امید ریاضی

این برآوردگر را محاسبه می‌کنیم. داریم

$$E(\hat{\beta}_R(k)) = E((k\mathbf{W} + \mathbf{I})^{-1}\hat{\beta}_R) = (k\mathbf{W} + \mathbf{I})^{-1}\beta \neq \beta$$

واضح است که $\hat{\beta}_R(k)$ برآوردگری اریب است.

برآوردگر ريج مقید، داده‌ها را بیش‌برآورد یا کم‌برآورد می‌کند ولی همان‌طور که دیدیم، این برآوردگر انقباضی است که این اثر اریب بودن این برآوردگر را تعدیل می‌کند.

ویژگی ۲.۲.۳. برآوردگر ريج مقید در شرط $R\beta = 0$ صدق می‌کند. به عبارتی

$$\hat{\beta}_R(k) \in B = \{\beta : R\beta = 0\}$$

برهان: از آنجایی که $(kW + I)^{-1}W = W(kW + I)^{-1}$ و $RW = 0$ می‌توان نوشت

$$R\hat{\beta}_R(k) = R(kW + I)^{-1}WX'T^{-1}Y = RW(kW + I)^{-1}X'T^{-1}Y = 0$$

یعنی $\hat{\beta}_R(k) \in B$

ویژگی ۳.۲.۳. در معادله (۷.۳) داریم $\hat{\alpha}(k) = Q'\hat{\beta}_R(k) = (\hat{\alpha}_1(k), \hat{\alpha}_2(k), \dots, \hat{\alpha}_p(k))'$ بنابراین

$$\hat{\alpha}_{p-q+1}(k) = \hat{\alpha}_{p-q+2}(k) = \dots = \hat{\alpha}_p(k) = 0$$

برهان: از آنجایی که

$$\begin{aligned} \hat{\alpha}(k) &= Q'\hat{\beta}_R(k) \\ &= Q'WQ'(kW + I)^{-1}X'T^{-1}Y \\ &= \Lambda Q'(kW + I)^{-1}X'T^{-1}Y \end{aligned} \quad (۱۱.۳)$$

با استفاده از لم (۲.۳) می‌توان به راحتی نتیجه گرفته که

$$\hat{\alpha}_{p-q+1}(k) = \hat{\alpha}_{p-q+2}(k) = \dots = \hat{\alpha}_p(k) = 0$$

ویژگی ۴.۲.۳. کوواریانس برآوردگر ریج مقید $\hat{\beta}_R(k)$ همواره از کوواریانس برآوردگر مقید $\hat{\beta}_R$ کمتر است. یعنی

$$Cov(\hat{\beta}_R(k)) \leq Cov(\hat{\beta}_R)$$

برهان: با استفاده از (۲۱.B) می توان نوشت

$$\begin{aligned} Cov(\hat{\beta}_R) &= Cov(\mathbf{W}\mathbf{X}'\mathbf{T}^{-}\mathbf{Y}) \\ &= \sigma^2 \mathbf{W}\mathbf{X}'\mathbf{T}^{-}\Sigma\mathbf{T}^{-}\mathbf{X}\mathbf{W}' \\ Cov(\hat{\beta}_R(k)) &= Cov((k\mathbf{W} + \mathbf{I})^{-1}\hat{\beta}_R) \\ &= (k\mathbf{W} + \mathbf{I})^{-1}Cov(\hat{\beta}_R)(k\mathbf{W} + \mathbf{I})^{-1} \\ &= \sigma^2 (k\mathbf{W} + \mathbf{I})^{-1}\mathbf{W}\mathbf{X}'\mathbf{T}^{-}\Sigma\mathbf{T}^{-}\mathbf{X}\mathbf{W}'(k\mathbf{W} + \mathbf{I})^{-1} \end{aligned} \quad (۱۲.۳)$$

بنابراین

$$\begin{aligned} Cov(\hat{\beta}_R) - Cov(\hat{\beta}_R(k)) &= \sigma^2 \mathbf{W}\mathbf{X}'\mathbf{T}^{-}\Sigma\mathbf{T}^{-}\mathbf{X}\mathbf{W}' \\ &\quad - \sigma^2 (k\mathbf{W} + \mathbf{I})^{-1}\mathbf{W}\mathbf{X}'\mathbf{T}^{-}\Sigma\mathbf{T}^{-}\mathbf{X}\mathbf{W}'(k\mathbf{W} + \mathbf{I})^{-1} \\ &= \sigma^2 (k\mathbf{W} + \mathbf{I})^{-1}[(k\mathbf{W} + \mathbf{I})\mathbf{W}\mathbf{X}'\mathbf{T}^{-}\Sigma\mathbf{T}^{-}\mathbf{X}\mathbf{W}'(k\mathbf{W} + \mathbf{I}) \\ &\quad - \mathbf{W}\mathbf{X}'\mathbf{T}^{-}\Sigma\mathbf{T}^{-}\mathbf{X}\mathbf{W}'](k\mathbf{W} + \mathbf{I})^{-1} \\ &= (k\mathbf{W} + \mathbf{I})^{-1}[k^2\mathbf{W}^2\mathbf{X}'\mathbf{T}^{-}\Sigma\mathbf{T}^{-}\mathbf{X}\mathbf{W}' + k\mathbf{W}^2\mathbf{X}'\mathbf{T}^{-}\Sigma\mathbf{T}^{-}\mathbf{X}\mathbf{W}' \\ &\quad + k\mathbf{W}\mathbf{X}'\mathbf{T}^{-}\Sigma\mathbf{T}^{-}\mathbf{X}\mathbf{W}'](k\mathbf{W} + \mathbf{I})^{-1} \\ &= \sigma^2 (k\mathbf{W} + \mathbf{I})^{-1}\mathbf{M}(k\mathbf{W} + \mathbf{I})^{-1} \end{aligned}$$

توجه داشته باشید که

$$M = k^{\vee} W^{\vee} X' T^{-} \Sigma T^{-} X W^{\vee} + k W^{\vee} X' T^{-} \Sigma T^{-} X W' + k W X' T^{-} \Sigma T^{-} X W^{\vee}$$

حال، از آنجایی که $\Sigma \geq 0$ و $k \geq 0$ و از لم (۱.B) داریم

$$k^{\vee} W^{\vee} X' T^{-} \Sigma T^{-} X W^{\vee} \geq 0, \quad W^{\vee} X' T^{-} \Sigma T^{-} X W' \geq 0,$$

با توجه به این که ثابت کردیم $W \geq 0$ و لم (۲.B) نتیجه می‌دهد

$$k W^{\vee} X' T^{-} \Sigma T^{-} X W' \geq 0, \quad k W X' T^{-} \Sigma T^{-} X W^{\vee} \geq 0,$$

و در نهایت از لم (۳.B) نتیجه می‌گیریم که $M \geq 0$. لذا $Cov(\hat{\beta}_R(k)) \leq Cov(\hat{\beta}_R)$.

ویژگی ۵.۲.۳. $tr(Cov(\hat{\beta}_R(k)))$ تابعی یکنوا و کاهشی از k است.

برهان: برای اثبات این ویژگی، کفایت نشان دهیم که $\frac{tr(Cov(\hat{\beta}_R(k)))}{dk} \leq 0$.

فرض می‌کنیم

$$B_k = k W + I \quad (۱۳.۳)$$

با استفاده از قضیه (۱۸.B) مشتق B_k^{-1} را نسبت به k به صورت زیر می‌نویسیم.

$$\frac{dB_k^{-1}}{dk} = -B_k^{-1} \frac{dB_k}{dk} B_k^{-1} = -B_k^{-1} W B_k^{-1} \quad (۱۴.۳)$$

معادله (۱۳.۳) را در معادله (۱۲.۳) قرار می‌دهیم و $Cov(\hat{\beta}_R(k))$ را به فرم زیر می‌نویسیم.

$$Cov(\hat{\beta}_R(k)) = \sigma^2 B_k^{-1} W X' T^{-1} \Sigma T^{-1} X W' B_k^{-1}$$

سپس از $Cov(\hat{\beta}_R(k))$ نسبت به k مشتق می‌گیریم. داریم

$$\begin{aligned} \frac{dCov(\hat{\beta}_R(k))}{dk} &= -\sigma^2 B_k^{-1} (W B_k^{-1} W X' T^{-1} \Sigma T^{-1} X W' \\ &\quad + W X' T^{-1} \Sigma T^{-1} X W' B_k^{-1} W) B_k^{-1} \end{aligned}$$

بنابراین

$$\frac{dCov(\hat{\beta}_R(k))}{dk} \leq 0 \Leftrightarrow W B_k^{-1} W X' T^{-1} \Sigma T^{-1} X W' + W X' T^{-1} \Sigma T^{-1} X W' B_k^{-1} W \geq 0$$

حال اگر قرار دهیم

$$\begin{aligned} N &= W(kW + I)^{-1} [W X' T^{-1} \Sigma T^{-1} X W' (kW + I) \\ &\quad + (kW + I) W X' T^{-1} \Sigma T^{-1} X W'] (kW + I)^{-1} W \\ &= W(kW + I)^{-1} [kW X' T^{-1} \Sigma T^{-1} X W' + kW X' T^{-1} \Sigma T^{-1} X W' \\ &\quad + X W X' T^{-1} \Sigma T^{-1} X W] (kW + I)^{-1} W \end{aligned}$$

در این قسمت باید ثابت کنیم که عبارت بالا مثبت است. با کمک لم‌های (۱.B) و (۲.B) می‌توانیم نشان دهیم

که $N \geq 0$ یعنی $\frac{dCov(\hat{\beta}_R(k))}{dk}$ و اثبات کامل است.

قضیه ۱.۲.۳. اگر پارامتر ریج در شرط زیر صدق کند

$$0 < k < \frac{\sigma^2 \lambda_{p-q} \eta_{p-q}}{\alpha^2} \quad (۱۵.۳)$$

آنگاه میانگین توان دوم خطای ”برآوردگر ريج مقید“ از میانگین توان دوم خطای ”برآوردگر مقید“ کمتر است.

$$MSE(\hat{\beta}_R) - MSE(\hat{\beta}_R(k)) > 0$$

برهان: با توجه به تعريف میانگین توان دوم خطا، یعنی

$$MSE(\hat{\beta}) = E\|\hat{\beta} - \beta\|^2 = E(\hat{\beta} - \beta)'(\hat{\beta} - \beta)$$

داریم

$$\begin{aligned} MSE(\hat{\beta}_R(k)) &= MSE(\hat{\alpha}(k)) = tr[Cov(\hat{\alpha}(k))] + \|E(\hat{\alpha}(k) - \alpha)\|^2 \\ &= \sigma^2 tr[(k\Lambda + I)^{-1} \Lambda Z' T^{-1} \Sigma T^{-1} Z \Lambda (k\Lambda + I)^{-1}] \\ &\quad + \| (k\Lambda + I)^{-1} E(\hat{\alpha} - \alpha) \|^2 \\ &= \sigma^2 tr[(k\Lambda + I)^{-1} \Lambda Z' T^{-1} \Sigma T^{-1} Z \Lambda (k\Lambda + I)^{-1}] + \| (k\Lambda + I)^{-1} \alpha - \alpha \|^2 \end{aligned}$$

حال فرض کنید $\eta = Diag(\eta_1, \eta_2, \dots, \eta_p) = Z' T^{-1} \Sigma T^{-1} Z = Q' X' T^{-1} \Sigma T^{-1} X Q$ که در آن $\eta_1 \geq \eta_2 \geq \dots \geq \eta_p$

$\eta_2 \geq \dots \geq \eta_p$ مقادير ویژه ترتیبی ماتریس $X' T^{-1} \Sigma T^{-1} X$ هستند.

از طرفی

$$\begin{aligned} \|(k\Lambda + I)^{-1} \alpha - \alpha\|^2 &= (k\Lambda + I)^{-2} \alpha^2 + \alpha^2 - 2\alpha^2 (k\Lambda + I)^{-1} \\ &= \alpha^2 (k\Lambda + I)^{-1} [(k\Lambda + I)^{-1} + (k\Lambda + I) - 2I] \\ &= \alpha^2 (k\Lambda + I)^{-1} [(k\Lambda + I)^{-1} + (k\Lambda + I)] \\ &= \alpha^2 (k\Lambda + I)^{-1} [k^2 \Lambda^2 (k\Lambda + I)^{-1}] \\ &= k^2 \Lambda^2 \alpha^2 (k\Lambda + I)^{-2} \end{aligned}$$

بنابراین میانگین توان دوم خطای برآوردگر ریح بر حسب مقادیر ویژه عبارتست از

$$MSE(\hat{\beta}_R(k)) = \sigma^2 \sum_{i=1}^{p-q} \frac{\lambda_i^2 \eta_i}{(k\lambda_i + 1)^2} + k^2 \sum_{i=1}^{p-q} \frac{\lambda_i^2 \alpha_i^2}{(k\lambda_i + 1)^2}$$

در نهایت از میانگین توان دوم خطای برآوردگر ریح نسبت به k مشتق می‌گیریم

$$\begin{aligned} \frac{dMSE(\hat{\beta}_R(k))}{dk} &= \sigma^2 \sum_{i=1}^{p-q} \frac{-2\lambda_i^3 \eta_i}{(k\lambda_i + 1)^3} + 2k \sum_{i=1}^{p-q} \frac{\lambda_i \alpha_i^2}{(k\lambda_i + 1)^3} \\ &= 2 \sum_{i=1}^{p-q} \frac{\lambda_i^2 (k\alpha_i^2 - \sigma^2 \lambda_i \eta_i)}{(k\lambda_i + 1)^3} \end{aligned}$$

واضح است که اگر حدود پارامتر ریح مطابق با شرط (۱۵.۳) باشد، آنگاه $\frac{dMSE(\hat{\beta}_R(k))}{dk} \leq 0$. اگر $MSE(\hat{\beta}_R(k))$

پیوسته باشد، $MSE(\hat{\beta}_R(k))$ تحت شرط (۱۵.۳) تابعی یکنوازی نزولی از k است.

بنابراین $MSE(\hat{\beta}_R(k)) < MSE(\hat{\beta}_R(0))$.

قضیه ۲.۲.۳. اگر k و U را طوری انتخاب کنیم که داشته باشیم

$$kW + 2I - kWSUSW - WSUS - SUSW - \frac{k}{\sigma^2} \beta\beta' \geq 0$$

آنگاه ماتریس توان دوم خطای برآوردگر ریح مقید کمتر از ماتریس میانگین توان دوم خطای برآوردگر ناریب

مقید است. به عبارتی $MSEM(\hat{\beta}_R) - MSEM(\hat{\beta}_R(k)) \geq 0$.

برهان: با توجه به تعریف ماتریس میانگین توان دوم خطا، یعنی $MSEM(\hat{\beta}) = E(\hat{\beta} - \beta)(\hat{\beta} - \beta)'$

داریم

$$MSEM(\hat{\beta}_R) = Cov(\hat{\beta}_R) = \sigma^2 W X' T^{-1} \Sigma T^{-1} X W' \quad (۱۶.۳)$$

حال با جایگذاری $\Sigma = T - XUX'$ می‌توان نتیجه گرفت

$$\begin{aligned}
 MSEM(\hat{\beta}_R) &= \sigma^2 W X' T^{-1} (T - XUX') T^{-1} X W' \\
 &= \sigma^2 W X' T^{-1} X W' - \sigma^2 W X' T^{-1} X U X' T^{-1} X W' \\
 &= \sigma^2 W S W' - \sigma^2 W S U S W' \\
 &= \sigma^2 W - \sigma^2 W S U S W'
 \end{aligned}$$

و بطور مشابه بدست می‌آوریم

$$\begin{aligned}
 MSEM(\hat{\beta}_R(k)) &= Cov(\hat{\beta}_R(k)) + (E(\hat{\beta}_R(k)) - \beta)(E(\hat{\beta}_R(k)) - \beta)' \\
 &= \sigma^2 (kW + I)^{-1} W X' T^{-1} \Sigma T^{-1} X W' (kW + I)^{-1} \\
 &\quad + [(kW + I)^{-1} \beta - \beta][(kW + I)^{-1} \beta - \beta]' \\
 &= \sigma^2 (kW + I)^{-1} W (kW + I)^{-1} \\
 &\quad - \sigma^2 (kW + I)^{-1} W S U S W (kW + I)^{-1} \\
 &\quad + k^2 (kW + I)^{-1} W \beta \beta' W (kW + I)^{-1}
 \end{aligned}$$

از طرفی

$$\begin{aligned}
 &[(kW + I)^{-1} \beta - \beta][(kW + I)^{-1} \beta - \beta]' \\
 &= (kW + I)^{-1} \beta' \beta (kW + I)^{-1} - \beta \beta' (kW + I)^{-1} \\
 &\quad - (kW + I)^{-1} \beta \beta' + \beta \beta' \\
 &= (kW + I)^{-1} [\beta \beta' - kW \beta \beta' - \beta \beta' - \beta \beta' kW - \beta \beta' + kW \beta \beta' kW]
 \end{aligned}$$

در نهایت می‌توانیم $MSEM(\hat{\beta}_R(k))$ را به صورت زیر بنویسیم

$$\begin{aligned} MSEM(\hat{\beta}_R(k)) &= \sigma^2(k\mathbf{W} + \mathbf{I})^{-1}\mathbf{W}(k\mathbf{W} + \mathbf{I})^{-1} \\ &- \sigma^2(k\mathbf{W} + \mathbf{I})^{-1}\mathbf{W}\mathbf{S}\mathbf{U}\mathbf{S}\mathbf{W}(k\mathbf{W} + \mathbf{I})^{-1} \\ &+ k^2(k\mathbf{W} + \mathbf{I})^{-1}\mathbf{W}\beta\beta'\mathbf{W}(k\mathbf{W} + \mathbf{I})^{-1} \end{aligned}$$

با محاسبه تفاضل در ماتریس میانگین توان دوم خطا، داریم

$$\begin{aligned} &MSEM(\hat{\beta}_R) - MSEM(\hat{\beta}_R(k)) \\ &= \sigma^2\mathbf{W} - \sigma^2\mathbf{W}\mathbf{S}\mathbf{U}\mathbf{S}\mathbf{W} - \sigma^2(k\mathbf{W} + \mathbf{I})^{-1}\mathbf{W}(k\mathbf{W} + \mathbf{I})^{-1} \\ &+ \sigma^2(k\mathbf{W} + \mathbf{I})^{-1}\mathbf{W}\mathbf{S}\mathbf{U}\mathbf{S}\mathbf{W}(k\mathbf{W} + \mathbf{I})^{-1} - k^2(k\mathbf{W} + \mathbf{I})^{-1}\mathbf{W}\beta\beta'\mathbf{W}(k\mathbf{W} + \mathbf{I})^{-1} \\ &= \sigma^2(k\mathbf{W} + \mathbf{I})^{-1}[(k\mathbf{W} + \mathbf{I})\mathbf{W}(k\mathbf{W} + \mathbf{I}) - \mathbf{W}\mathbf{S}\mathbf{U}\mathbf{S}\mathbf{W} \\ &- (k\mathbf{W} + \mathbf{I})\mathbf{W}\mathbf{S}\mathbf{U}\mathbf{S}\mathbf{W}(k\mathbf{W} + \mathbf{I}) - \frac{k^2}{\sigma^2}\mathbf{W}\beta\beta'\mathbf{W}](k\mathbf{W} + \mathbf{I})^{-1} \end{aligned}$$

بنابراین نتیجه می‌شود

$$\begin{aligned}
 & MSEM(\hat{\beta}_R) - MSEM(\hat{\beta}_R(k)) \geq 0 \\
 \Leftrightarrow & (kW + I)W(kW + I) - W + WSUSW \\
 & - (kW + I)WSUSW(kW + I) - \frac{k^2}{\sigma^2}W\beta\beta'W \geq 0 \\
 \Leftrightarrow & k^2W' + 2kW' + W - W + WSUSW - k^2W'SUSW' - kW'SUSW \\
 & - kW'SUSW' - WSUSW - \frac{k^2}{\sigma^2}W\beta\beta'W \geq 0 \\
 \Leftrightarrow & kW' + 2W' - kW'SUSW' - W'SUW - WSUSW' - \frac{k}{\sigma^2}W\beta\beta'W \geq 0 \\
 \Leftrightarrow & kW + 2I - kW'SUSW - WSUS - SUSW - \frac{k}{\sigma^2}\beta\beta' \geq 0
 \end{aligned}$$

۳.۳ شبیه‌سازی

حال برای بررسی صحت نتایج ارائه شده، با استفاده از شبیه‌سازی مونت-کارلو که شرح آن در فصل دوم گذشت. "برآوردگر مقید" (RE) و "برآوردگر ريج مقید" (RRE) را به ازای k های مختلف مقایسه می‌کنیم. نتایج این قسمت نیز با کمک نرم‌افزار R بدست آمده است. همانند فصل دوم، داده‌ها را به روش زیر تولید می‌کنیم

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

که در آن z_{ij} اعداد تصادفی مستقل از توزیع نرمال استاندارد و γ^2 همبستگی بین هر دو متغیر توضیحی است. این متغیرها استاندارد شده‌اند، به طوری که $X'X$ و $X'Y$ همبسته هستند. در این قسمت فرض می‌کنیم

۰/۷۰. γ . حال با استفاده از رابطه زیر n مشاهده برای متغیر وابسته تعیین می‌شود

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_p x_{ip} + E_i, \quad i = 1, 2, \dots, n$$

که $\beta_0 = 0$ و E_i اعداد تصادفی مستقل از توزیع نرمال با میانگین صفر و واریانس σ^2 است.

در اینجا $n = 15$ و $p = 4$ در نظر گرفته شده‌اند.

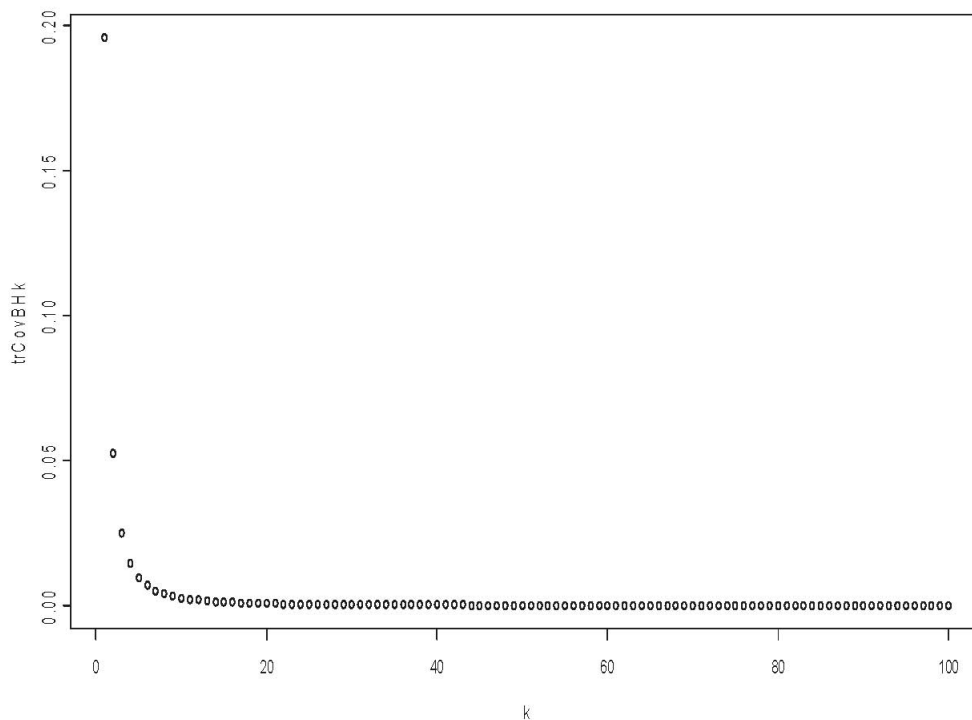
پس از انجام محاسبات، نتایج زیر حاصل شدند

$$\hat{\beta}_R = \begin{bmatrix} -9/2651 \\ -1/6460 \end{bmatrix} \Rightarrow \|\hat{\beta}_R\|^2 = 88/5523$$

همچنین نمودار اثر ماتریس کوواریانس برآوردگر RRE به صورت زیر می‌باشد.

k	۱	۱۱	۲۱	۳۱	۴۱
$\hat{\beta}_R(k)$	$\begin{bmatrix} 0/7141 \\ 0/3806 \end{bmatrix}$	$\begin{bmatrix} 0/0551 \\ 0/0561 \end{bmatrix}$	$\begin{bmatrix} 0/0284 \\ 0/0306 \end{bmatrix}$	$\begin{bmatrix} 0/0191 \\ 0/0210 \end{bmatrix}$	$\begin{bmatrix} 0/0144 \\ 0/0160 \end{bmatrix}$
$\ \hat{\beta}_R(k)\ ^2$	0/6548	0/0061	0/0017	0/00081	0/00046
k	۵۱	۶۱	۷۱	۸۱	۹۱
$\hat{\beta}_R(k)$	$\begin{bmatrix} 0/0116 \\ 0/0129 \end{bmatrix}$	$\begin{bmatrix} 0/0097 \\ 0/0090 \end{bmatrix}$	$\begin{bmatrix} 0/0083 \\ 0/0093 \end{bmatrix}$	$\begin{bmatrix} 0/0072 \\ 0/0082 \end{bmatrix}$	$\begin{bmatrix} 0/0064 \\ 0/0073 \end{bmatrix}$
$\ \hat{\beta}_R(k)\ ^2$	0/00030	0/00021	0/00015	0/00012	9/6171

شکل ۱.۳: نمودار اثر ماتریس کواریانس در مقابل پارامتر ریح



۴.۳ نتیجه‌گیری

در این فصل، مسئله برآورد پارامتر در مدل خطی ریح منفرد با محدودیت تساوی خطی بیان و با فرض وجود همخطی بین ستون‌های ماتریس طرح، برآوردگر ریح محدود شده ارائه شد. با توجه به قضایای ارائه شده، می‌توان نتیجه گرفت که RRE بر RE برتری دارد. با توجه به نتایج شبیه‌سازی، ویژگی‌های برآوردگر بررسی شدند و دیدیم که اندازه نرم اقلیدسی RRE به ازای مقادیر مختلف پارامتر ریح شرطی، کمتر از نرم اقلیدسی RE می‌باشد و همچنین شکل (۱.۳) درستی ویژگی (۵.۳) را به خوبی نشان می‌دهد، یعنی با افزایش پارامتر ریح، اثر ماتریس کواریانس بردار خطا کاهش می‌یابد.

فصل ۴

برآوردگر ریج منفرد در مدل با محدودیت تصادفی

مقدمه

در تحقیقات آماری، ممکن است علاوه بر مدل رگرسیونی، اطلاعات پیشین تجربی نیز دستخوش خطا شوند. بنابراین لازم می‌آید که برای محدودیت اعمال شده روی مدل نیز یک جمله خطای تصادفی در نظر گرفته شود و بطور کلی ما اطلاعات پیشین دقیقی مانند $R/\beta = r$ در روابط اقتصادی، ساختارهای صنعتی، برنامه‌ریزی و ... نداریم، بنابراین می‌توانیم رویدادهای تصادفی نامعلوم را اثر محدودیت‌های خطی اعمال کنیم. برای اطلاعات بیشتر در خصوص اینکه چرا ما با محدودیت‌های تصادفی کار می‌کنیم و کاربردهای آن، رجوع شود به جاج و باک^۱ (۱۹۸۷) و تیل^۲ (۱۹۶۳).

در بخش اول این فصل، پارامترهای مدل رگرسیونی محدود شده را با این فرض که محدودیت شامل خطای تصادفی است، برآورد می‌کنیم. در بخش دوم، ویژگی‌های آن را بررسی کرده و خطای برآوردگر را محاسبه می‌کنیم. در بخش سوم، به کمک شبیه‌سازی مونت-کارلو، خطای این برآوردگر را با خطای برآوردگر کمترین توان‌های دوم مقایسه می‌کنیم و در بخش چهارم، نتیجه‌گیری را ارائه می‌دهیم.

۱.۴ معرفی برآوردگر

مدل خطی زیر را در نظر بگیرید

$$\begin{cases} Y = X\beta + E \\ E(E) = 0, \quad Cov(E) = \sigma^2 I \end{cases} \quad (1.4)$$

که در آن Y بردار تصادفی n بعدی از پاسخ‌ها، $X_{n \times p}$ ماتریس طرح غیرتصادفی ثابت معلوم پر رتبه سطری، β برداری مجهول از ضرایب رگرسیونی p بعدی و E بردار خطای تصادفی n بعدی است. علاوه بر مدل (۱.۴)،

^۱Jock and Bock

^۲Theil

فرض کنید اطلاعات پیشین به شکل محدودیت تصادفی زیر نیز وجود دارند.

$$r = R\beta + v \quad (۲.۴)$$

که در آن R ماتریس معلوم $q \times p$ از ثابت‌های غیرتصادفی پررتبه سطری، برداری q بعدی از مقادیر معین و جمله خطای تصادفی است.

فرض کنید $E \sim N_n(\cdot, \sigma^2 I)$ و $v \sim N_q(\mu, \sigma^2 I)$ که μ بردار q بعدی از مقادیر معلوم و σ^2 مجهول است. به علاوه فرض کنید E و v مستقل باشند. تیل و گولدرگر^۱ (۱۹۶۱) پیشنهاد دادند که برای بدست آوردن رابطه‌ای خطی، اطلاعات پیشین تصادفی و مدل خطی به صورت زیر ترکیب شوند.

$$\begin{bmatrix} Y \\ r \end{bmatrix} = \begin{bmatrix} X \\ R \end{bmatrix} \beta + \begin{bmatrix} E \\ v \end{bmatrix} \quad (۳.۴)$$

در این حالت، محدودیت مورد نظر عبارتست از

$$H, : \begin{bmatrix} -R & I_q \end{bmatrix} \times \begin{bmatrix} \beta \\ R\beta + \mu \end{bmatrix} = \mu = \cdot \quad (۴.۴)$$

مدل (۳.۴) را به صورت زیر بازنویسی می‌کنیم

$$Y^* = Z\Phi + \Upsilon \quad (۵.۴)$$

که در آن

$$Y^* = \begin{bmatrix} Y \\ r \end{bmatrix}, Z = \begin{bmatrix} X & \cdot_{n \times q} \\ \cdot_{q \times p} & I_q \end{bmatrix}, \Phi = \begin{bmatrix} \beta \\ R\beta + \mu \end{bmatrix}$$

^۱Goldberger

که Y^* برداری $1 \times (n+q)$ بعدی از مشاهدات، Z ماتریس طرح معلوم $(p+q) \times (n+q)$ بعدی با رتبه $p+q$ ، Φ برداری $1 \times (p+q)$ بعدی از پارامترهای مورد بررسی، Υ بردار خطای تصادفی $1 \times (n+q)$ است. بنابراین محدودیت روی مدل جدید به صورت زیر بدست می‌آید.

$$H\Phi = \mu \quad (۶.۴)$$

که در آن H ماتریس محدودیت $(p+q) \times q$ بعدی با رتبه q است. حال، حالتی را در نظر می‌گیریم که ماتریس واریانس-کوواریانس بردار خطا منفرد است، بنابراین قرار می‌دهیم $T = \Sigma + ZUZ'$ که در آن $\Sigma = Cov(\Upsilon)$ و $U \geq 0$. اکنون تعریق می‌کنیم

$$F(\Phi, \lambda) = (Y^* - Z\Phi)'T^{-1}(Y^* - Z\Phi) + \lambda'H\Phi$$

با مشتقگیری از $F(\Phi, \lambda)$ نسبت به Φ و تساوی با صفر قرار دادن آن داریم

$$\begin{aligned} \frac{\partial F}{\partial \Phi} &= -\lambda'Z'T^{-1}Y^* + \lambda'Z'T^{-1}Z\Phi + \lambda'H' \\ \Rightarrow Z'T^{-1}Z\Phi &= Z'T^{-1}Y^* - H'\lambda \end{aligned}$$

بنابراین برآوردگر محدودشده عبارتست از

$$\begin{aligned} \hat{\Phi}_H &= (Z'T^{-1}Z)^{-1}(Z'T^{-1}Y^* - H'\lambda) \\ &= (Z'T^{-1}Z)^{-1}Z'T^{-1}Y^* - (Z'T^{-1}Z)^{-1}H'\lambda \\ &= \hat{\Phi}_{LS} - (Z'T^{-1}Z)^{-1}H'\lambda \end{aligned} \quad (۷.۴)$$

که در آن $\hat{\Phi}_{LS} = (Z'T^-Z)^{-1} Z'T^-Y^*$ برآوردگر محدود نشده پارامترهای رگرسیونی است.

از آنجایی که

$$\begin{aligned} H\hat{\Phi}_H &= H\hat{\Phi}_{LS} - H(Z'T^-Z)^{-1}H'\lambda = 0 \\ \Rightarrow H\hat{\Phi}_{LS} &= H(Z'T^-Z)^{-1}H'\lambda \\ \Rightarrow \lambda &= [H(Z'T^-Z)^{-1}H']^{-1}H\hat{\Phi}_{LS} \end{aligned} \quad (۸.۴)$$

با جایگذاری معادله (۸.۴) در معادله (۷.۴) بدست می‌آوریم

$$\hat{\Phi}_H = \hat{\Phi}_{LS} - (Z'T^-Z)^{-1}H'[H(Z'T^-Z)^{-1}H]H\hat{\Phi}_{LS} \quad (۹.۴)$$

با فرض اینکه $S = Z'T^-Z$ و $W = S^{-1} - S^{-1}H'(HS^{-1}H')^{-1}HS^{-1}$ ، بنابراین $\hat{\Phi}_H$ را می‌توان به

صورت زیر بازنویسی نمود

$$\hat{\Phi}_H = WZT^-Y^* \quad (۱۰.۴)$$

$\hat{\Phi}_H$ را بهترین برآوردگر نااریب خطی Φ می‌نامند.

در صورت وجود همخطی، معادله (۱۰.۴) را اصلاح می‌کنیم و برآوردگر ریج شرطی را به صورت زیر پیشنهاد

می‌دهیم.

$$\hat{\Phi}_H(k) = (kW + I)^{-1}\hat{\Phi}_H \quad (۱۱.۴)$$

۲.۴ ویژگی‌های برآوردگر

برای سادگی محاسبات MSE ، شکل کانونی مدل رگرسیون خطی منفرد با محدودیت تساوی خطی را معرفی می‌کنیم.

$$\begin{cases} Y^* = Z^* \alpha + \Upsilon \\ L\alpha = 0 \end{cases} \quad (۱۲.۴)$$

که در آن $Z^* = ZQ$ ، $L = HQ$ و $\alpha = Q'\Phi$ پارامتر کانونی شرطی نامیده می‌شود. Q ماتریسی است متعامد به‌طوری که $Q'WQ = \text{Diag}(\lambda_1, \lambda_2, \dots, \lambda_p)$ و $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p$ مقادیر ویژه ترتیبی ماتریس W هستند.

واضح است که

$$\hat{\alpha} = (Z^{*'}T^-Z^*)^{-1} - (Z^{*'}T^-Z^*)^{-1}L'[L(Z^{*'}T^-Z^*)^{-1}L']^{-1}L(Z^{*'}T^-Z^*)^{-1}Z^{*'}T^-Y^*$$

توجه کنید که

$$\hat{\alpha}(k) = (k\Lambda + I)^{-1}\hat{\alpha} \quad (۱۳.۴)$$

و می‌توانیم بدست آوریم

$$\hat{\Phi}_H = Q\hat{\alpha} \quad , \quad \hat{\Phi}_H(k) = Q\hat{\alpha}(k) \quad (۱۴.۴)$$

و همچنین

$$MSE(\hat{\Phi}_H) = MSE(\hat{\alpha}) \quad , \quad MSE(\hat{\Phi}_H(k)) = MSE(\hat{\alpha}(k)) \quad (۱۵.۴)$$

لم ۱.۲.۴. $Rank(W) = p$ و $W = S^{-1} - S^{-1}H'(HS^{-1}H')^{-1}HS^{-1}$

برهان: مشابه برهان لم (۱.۳) است.

ویژگی ۱.۲.۴. $\hat{\Phi}_H(k)$ برآوردگر انقباضی و اریب Φ است.

برهان: از معادله (۱۱.۴) و تعریف Λ در معادله (۱۲.۴) داریم

$$\begin{aligned}\|\hat{\Phi}_H(k)\|^2 &= \|(kW + I)^{-1} \hat{\Phi}_H\|^2 = \|Q(k\Lambda + I)^{-1} Q' \hat{\Phi}_H\|^2 \\ &= \|(k\Lambda + I)^{-1} Q' \hat{\Phi}_H\|^2 < \|Q' \hat{\Phi}_H\|^2 = \|\hat{\Phi}_H\|^2\end{aligned}$$

بنابراین $\hat{\Phi}_H(k)$ برآوردگر انقباضی Φ است.

با توجه به این که $E(\hat{\Phi}_H(k)) = E((kW + I)^{-1} \hat{\Phi}_H) = (kW + I)^{-1} \Phi \neq \Phi$ می بینیم که

$\hat{\Phi}_H(k)$ برآوردگری اریب برای Φ است.

ویژگی ۲.۲.۴. $\hat{\Phi}_H(k) \in B = \{\Phi : H\Phi = \bullet\}$ در شرط $H\Phi = \bullet$ صدق می کند. به عبارتی

برهان: از آنجایی که $(kW + I)^{-1} W \neq W(kW + I)^{-1}$ و $HW = \bullet$ بنابراین

$$H\hat{\Phi}_H(k) = H(kW + I)^{-1} Z'T^{-1}Y^* = HW(kW + I)^{-1} Z'T^{-1}Y^* = \bullet$$

یعنی $\hat{\Phi}_H(k) \in B$.

ویژگی ۳.۲.۴. کوواریانس برآوردگر ریج شرطی $\hat{\Phi}_H(k)$ همواره بر کوواریانس برآوردگر شرطی $\hat{\Phi}_H$ برتری دارد. یعنی

$$Cov(\hat{\Phi}_H(k)) \leq Cov(\hat{\Phi}_H)$$

برهان: مشابه برهان ویژگی (۴.۳) می باشد.

ویژگی ۴.۲.۴. $tr(Cov(\hat{\Phi}_H(k)))$ همواره تابعی یکنوازی نزولی از k است.

برهان: مشابه برهان ویژگی (۵.۳) است.

قضیه ۱.۲.۴. * اگر

$$0 < k < \frac{\sigma^2 \lambda_{p-q} \eta_{p-q}}{\alpha^2} \quad (۱۶.۴)$$

که در آن η_i مقادیر ویژه ترتیبی ماتریس $X'T^{-1}\Sigma T^{-1}X$ و λ_i مقادیر ویژه ترتیبی ماتریس $X'X$ هستند. آنگاه $MSE(\hat{\Phi}_H) - MS(\hat{\Phi}_H(k)) > 0$

برهان: چون $MSE(\hat{\Phi}) = E\|\hat{\Phi} - \Phi\|^2 = E(\hat{\Phi} - \Phi)'(\hat{\Phi} - \Phi)$ داریم

$$\begin{aligned} MSE(\hat{\Phi}_H(k)) &= MSE(\hat{\alpha}(k)) = tr[Cov(\hat{\alpha}(k))] + \|E(\hat{\alpha}(k) - \alpha)\|^2 \\ &= \sigma^2 tr[(k\Lambda + I)^{-1} \Lambda Z'T^{-1}\Sigma T^{-1}Z \Lambda (k\Lambda + I)^{-1}] \\ &\quad + \|(k\Lambda + I)^{-1} E(\hat{\alpha} - \alpha)\|^2 \\ &= \sigma^2 tr[(k\Lambda + I)^{-1} \Lambda Z'T^{-1}\Sigma T^{-1}Z \Lambda (k\Lambda + I)^{-1}] + \|(k\Lambda + I)^{-1} \alpha - \alpha\|^2 \end{aligned}$$

فرض کنید $\Phi = Diag(\eta_1, \eta_2, \dots, \eta_p) = Z'T^{-1}\Sigma T^{-1}Z = Q'Z'T^{-1}\Sigma T^{-1}ZQ$ که در آن $\eta_1 \geq \eta_2 \geq \dots \geq \eta_p$

مقادیر ویژه ترتیبی $Z'T^{-1}\Sigma T^{-1}Z$ هستند. در این صورت داریم

$$MSE(\hat{\Phi}_H(k)) = \sigma^2 \sum_{i=1}^p \frac{\lambda_i^2 \eta_i}{(k\lambda_i + 1)^2} + k^2 \sum_{i=1}^p \frac{\lambda_i^2 \alpha_i^2}{(k\lambda_i + 1)^2} \quad (۱۷.۴)$$

با مشتگیری از $MSE(\hat{\Phi}_H(k))$ نسبت به k نتیجه می‌شود

$$\begin{aligned} \frac{dMSE(\hat{\Phi}_H(k))}{dk} &= \sigma^2 \sum_{i=1}^p \frac{-2\lambda_i^2 \eta_i}{(k\lambda_i + 1)^3} + 2k \sum_{i=1}^p \frac{\lambda_i^2 \alpha_i^2}{(k\lambda_i + 1)^3} \\ &= 2 \sum_{i=1}^p \frac{\lambda_i^2 (k\alpha_i^2 - \sigma^2 \eta_i)}{(k\lambda_i + 1)^3} \end{aligned} \quad (۱۸.۴)$$

با توجه به رابطه بالا می‌بینیم که اگر شرط (۱۶.۴) درست باشد، آنگاه $\frac{dMSE(\hat{\Phi}_H(k))}{dk} < 0$ هنگامی که

$MSE(\hat{\Phi}_H(k))$ پیوسته باشد، می‌توانیم نتیجه بگیریم که $MSE(\hat{\Phi}_H(k))$ تابعی یکنوا نزولی از k است.

$$(MSE(\hat{\Phi}_H) - MSE(\hat{\Phi}_H(k))) > 0$$

قضیه ۲.۲.۴. اگر k و U را طوری داشته باشیم که

$$kW + 2I - kWSUSW - WSUS - SUSW - \frac{k}{\sigma^2} \Phi \Phi' \geq 0$$

آنگاه

$$MSEM(\hat{\Phi}_H) - MSEM(\hat{\Phi}_H(k)) \geq 0$$

برهان: مشابه برهان قضیه (۲.۳) می‌باشد.

۳.۴ شبیه‌سازی

با استفاده از تکنیک شبیه‌سازی ارائه‌شده در فصل دوم (شبیه‌سازی مونت-کارلو)، ابتدا داده‌ها را به صورت زیر

تولید کرده

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

و سپس n مشاهده را برای متغیر مستقل از رابطه زیر بدست می‌آوریم

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_p x_{ip} + E_i, \quad i = 1, 2, \dots, n$$

به طریق مشابه مقادیر بردار r را نیز تولید می‌کنیم

$$r_i = \beta_0 + \beta_1 R_{i1} + \beta_2 R_{i2} + \dots + \beta_p R_{ip} + v_i, \quad i = 1, 2, \dots, q$$

سپس با روی هم قرار دادن مدل رگرسیونی و اطلاعات پیشین تصادفی بردارهای Y^* و Φ و ماتریس Z را به

صورت زیر را به صورت زیر بدست می‌آوریم

$$Y^* = \begin{bmatrix} Y \\ r \end{bmatrix}, Z = \begin{bmatrix} X & \cdot_{n \times q} \\ \cdot_{q \times q} & I_q \end{bmatrix}, \Phi = \begin{bmatrix} \beta \\ R\beta + \mu \end{bmatrix}$$

با در نظر گرفتن $\gamma = 0.70$ و حل مسئله به ازای $n = 15$ و $p = 4$ ، نتایج زیر بدست آمده است.

جدول ۱.۴: مقادیر λ و β به کار رفته در شبیه سازی

γ	۰.۷۰	۰.۸۰	۰.۹۰	۰.۹۹
n	$n = 15$		$p = 4$	
λ_1	۹,۰۱۴۰.۳۵	۱۹,۹۴۵۷۸۲۱	۰,۹۳۰.۸۴۹۸	۰,۹۴۷۰.۸۰۱
λ_2	۵,۱۷۶۴۸۱	۱۴,۹۹۶۷۱۷۵	۸,۰۲۰.۷۲۰۵	۰,۴۵۴۰.۲۲۰
λ_3	۲,۳۲۸۷۳۵	۲,۷۲۹۳۱۲۷	۲,۴۲۷۸۲۲۲	۰,۱۳۰.۲۲۹۴
λ_4	۱,۲۷۰.۴۲۷	۰,۹۰۸۵۱۸	۰,۳۷۳۴۰.۱۴	۰,۱۱۰.۲۰۱۸
β_1	۰,۸۸۷۹۸۹۷۴	-۰,۴۴۸۹۸۱۱۰	-۰,۰۶۴۳۰.۵۸۶	۰,۰۷۵۸۶۱۴۲
β_2	-۰,۴۸۹۸۴۲۴۷	-۰,۰۹۲۹۸۰۷۴	-۰,۲۵۹۵۵۸۵۶	۰,۸۲۷۰.۶۶۰۷
β_3	۰,۸۰۰.۹۸۴۸۵	۰,۱۸۸۳۲۶۸۷	-۰,۰۱۲۷۴۲۷۷	-۰,۵۶۸۱۴۹۵۳
β_4	۰,۹۹۸۰۰.۷۷۳	۰,۰۱۱۰.۲۴۴۴	-۰,۰۴۲۶۷۳۰.۵	-۰,۰۴۵۱۴۴۶۴

جدول ۲.۴: مقادیر MSE برآورد شده با $n = ۱۵$ ، $p = ۴$ و $\gamma = ۰/۷۰$

$RLSE$	$k = ۰/۰۰۰۱$	$k = ۰/۰۰۰۳$
۹۹۴,۷۶۳۸	$۵,۳۰۶۸۵۶ \text{ e-}۰۵$	۰/۰۱۴۶۴۵۸۱
$k = ۰/۰۰۱$	$k = ۰/۰۰۲$	$k = ۰/۰۱$
۰/۰۱۰۶۲۲۸۳۱	۰/۵۰۳۴۹۵	۷۹/۰۰۹۳۳

۴.۴ نتیجه‌گیری

همان‌طور که ملاحظه شد، برآوردگر ریج پیشنهاد داده شده در این مدل نسبت به برآوردگر کمترین توان‌های دوم، کوواریانس کمتری دارد و برآوردگری اریب و انقباضی است و نتایج شبیه‌سازی نیز نشان‌دهنده برتری قابل توجه این برآوردگر بر برآوردگر کمترین توان‌های دوم محدود شده است.

فصل ۵

برآوردگر ریج منفرد در مدل SUR

مقدمه

معمولا در معاملات کلان اقتصادی، تجاری و بازار بورس، میزان سرمایه و نقدینگی یک شرکت، تاثیر بسزایی در تصمیمات اقتصادی آینده‌نگر و خرید و فروش‌های سهام آن شرکت دارد. در این راستا، شرکت‌های تجاری با توجه به میزان نقدینگی‌شان در اول هر سال، در خرید و فروش‌های (فرصت‌های کلان) اقتصادی و معاملات درازمدت تصمیم‌گیری می‌کنند. بدین منظور، تعیین مدلی برای پیش‌بینی و تخمین (برآورد) میزان نقدینگی سال‌های آینده یک شرکت، اهمیت زیادی دارد. توجه به این نکته، الزامی است که نه تنها مدل سرمایه یا میزان نقدینگی شرکت‌ها از یکدیگر مستقل نیستند، بلکه در شرایط کنونی رقابتی، به دلیل وجود تاثیر بعضی عوامل از قبیل میزان سود سهام بورس و سهام واریز نشده هر سال، تعاملی متقابل بین مدل‌های سرمایه شرکت‌ها وجود دارد. یعنی میزان سرمایه یک شرکت بر میزان مبادلات و معاملات شرکت دیگر (شرکت رقیب) تاثیر دارد. بوت و دویت^۱ (۱۹۶۰) را ببینید. از طرفی برای تعیین مدلی برای میزان نقدینگی یا سرمایه، نیاز به تخمین درست ضرایب تاثیر (پارامترهای مدل) هر کدام از عوامل موثر فوق می‌باشد. لذا در چنین شرایطی، به جای مدل‌بندی جداگانه برای هر شرکت، مدل‌های توأم با خطای متقابل را در نظر می‌گیریم. در چنین حالتی، فرض پایه‌ای این است که اگر مثلا سرمایه سهام یک شرکت تغییر کند، بر روی سرمایه سهام شرکت دیگر (رقیب) نیز تاثیر دارد. برای این منظور مدل‌هایی که که تاکنون برای مدل سرمایه به کار می‌رفته‌اند از قبیل رگرسیون چندگانه یا رگرسیون چندمتغیره به تنهایی جوابگوی بستگی روابط بین شرکت‌ها نمی‌باشند. لذا از مدل رگرسیون به ظاهر نامرتبط استفاده می‌کنیم.

مدل رگرسیون به ظاهر نامرتبط (*SUR*)، اولین بار توسط زلنر^۲ (۱۹۶۲) به صورت تکنیکی برای تحلیل یک مجموعه از معادلات چندگانه شامل معادلات متقابل پارامترها و عبارات خطای وابسته مورد استفاده قرار گرفت. در تجزیه و تحلیل مسائل آماری و اقتصادی بخصوص مطالعه رفتار توابع تولید و سرمایه ممکن است با

^۱ Boot and de Witt^۲ Zellner

شرایطی مواجه شویم که در آن معادلات رگرسیونی چندگانه طاهرا از یکدیگر مستقل باشند ولی عبارات خطای آنها از طریق معادلات به یکدیگر مربوط باشند. مدل SUR در حالاتی که علاقه‌مند به بررسی و مقایسه توابع سرمایه چند شرکت و کارخانه هستیم، کاربرد بسطایی دارد. با در نظر گرفتن مدل SUR امکان بررسی و مدل‌بندی توام در چنین شرایطی بوجود می‌آید. مدل SUR تعمیمی از مدل رگرسیون خطی است که در آن عبارات خطای معادلات رگرسیونی به یکدیگر وابسته‌اند. این وابستگی خطاها تاثیر مدل شرکت‌ها را بر یکدیگر به‌خوبی توجیه می‌کند.

صورت کلی مدل SUR عبارتست از

$$\begin{aligned} y_{it} &= \beta_{11}x_{it,1} + \beta_{12}x_{it,2} + \dots + \beta_{1k}x_{it,k} + E_{it} \\ &\vdots \\ t &= 1, 2, \dots, q \end{aligned} \quad (1.5)$$

$$y_{nt} = \beta_{p1}x_{pt,1} + \beta_{p2}x_{pt,2} + \dots + \beta_{pk}x_{pk,k} + E_{nt}$$

مدل (۱.۵) را می‌توان به صورت زیر نیز نشان داد

$$Y_i = X_i\beta_i + E_i \quad \forall i = 1, 2, \dots, p \quad (2.5)$$

که در آن برداری Y_i از $q \times 1$ بعدی از مقادیر وابسته، X_i یک ماتریس طرح $q \times k_i$ بعدی از مقادیر نمونه k_i متغیر مستقل و β_i برداری $k_i \times 1$ بعدی از ضرایب رگرسیونی می‌باشند. در مدل SUR فرض می‌شود.

$$E(E_i) = 0, \quad E(E_i E_i') = \sigma_{ii} I_q \quad (3.5)$$

مدل (۲.۵) را می‌توان به صورت ماتریسی زیر نیز نشان داد.

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_p \end{bmatrix} = \begin{bmatrix} X_1 & \cdot & \dots & \cdot \\ \cdot & X_2 & \dots & \cdot \\ \vdots & \vdots & \ddots & \vdots \\ \cdot & \cdot & \dots & X_p \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_p \end{bmatrix} + \begin{bmatrix} E_1 \\ E_2 \\ \vdots \\ E_p \end{bmatrix} \quad (۴.۵)$$

به طور کلی‌تر به ازای $k = \sum_{i=1}^p k_i$ و $n = pq$ داریم

$$Y = X\beta + E \quad (۵.۵)$$

که در آن $Y_{n \times 1}$ برداری از متادیر پاسخ، $X_{n \times k}$ یک ماتریس غیرتصادفی بلوکی-قطری، $\beta_{k \times 1}$ برداری از پارامترهای مجهول و $E_{n \times 1}$ برداری از خطاهای تصادفی است. در این صورت با استفاده از (۲.۵) داریم

$$E[EE'] = \begin{bmatrix} \sigma_{11}I_q & \sigma_{12}I_q & \dots & \sigma_{1p}I_q \\ \sigma_{21}I_q & \sigma_{22}I_q & \dots & \sigma_{2p}I_q \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{p1}I_q & \sigma_{p2}I_q & \dots & \sigma_{pp}I_q \end{bmatrix} = \Sigma \otimes I_q \quad (۶.۵)$$

که در آن $\otimes$ عملگر ضرب کرونکر است. (پیوست B را ببینید). با توجه به صورت پیچیده مدل *SUR* در عین کاربردی بودن آن، در حالتی که ماتریس Σ معلوم است، می‌توان نتایجی نیرومند بدست آورد ولی این حالت مسائل کاربردی و واقعی را پوشش نمی‌دهد، لذا به منظور کاربردی بودن و استفاده از مشاهدات واقعی، تنها می‌توان رفتار مجانبی مدل *SUR* را مورد تجزیه و تحلیل قرار داد. برای مشاهده توجیه و دلیل این مطلب می‌توان به زلنر (۱۹۶۲ و ۱۹۶۳) و روانکار^۱ (۱۹۷۴ و ۱۹۷۶) مراجعه کرد. همچنین برای آگاهی بیشتر در زمینه بررسی خواص مدل‌های ساده *SUR* می‌توان به سریواستاوا و جایلز^۲ (۱۹۸۷) مراجعه کرد. علاوه بر این منبع اخیر، کاریا و کوراتا^۳ (۲۰۰۴) در قسمت‌هایی محدود، به بحث در مورد برآوردهای کمترین توان‌های

^۱Revankar

^۲Srivastava and Giles

^۳Kariya and Kurata

دوم تعمیم‌یافته در مدل کلی SUR (۳.۵) پرداخته است.

در این فصل، فرض می‌کنیم که بردار خطای تصادفی مدل (۳.۵) دارای توزیع نرمال چندمتغیره است و بر این اساس در بخش اول، برآوردگر کمترین توان‌های دوم تعمیم‌یافته و برآوردگر کمترین توان‌های دوم تعمیم‌یافته محدودشده را بدست آورده، سپس آن را تحت مدل خطی منفرد بررسی کرده و در آخر برآوردگر ريج را برای مدل SUR ارائه می‌دهیم و در بخش دوم ویژگی‌های برآوردگر مورد بررسی را بطور خلاصه بیان می‌کنیم.

۱.۵ معرفی برآوردگر

همانطور که در مقدمه به آن اشاره شده، به دلیل صورت پیچیده مدل SUR در (۳.۵)، مطالعه برآوردگرهای بردار پارامتر β به صورت دقیق، ممکن نیست. در عین حال بدست آوردن برآوردگرهای مجانبی، با استفاده از برآوردگرهای دقیق امکانپذیر می‌باشد. بدین منظور، ابتدا فرض می‌کنیم ماتریس Σ معلوم است و برآوردگرهای موردنظر را می‌یابیم و سپس در حالت مجانبی خواص برآوردگرها را بررسی می‌کنیم. با استفاده از نتایج بدست آمده در کاریا و کوراتا (۲۰۰۴)، برآوردگر کمترین توان‌های دوم بردار ضرایب رگرسیونی مدل SUR (۳.۵) برابر است با

$$\hat{\beta}_{LS} = (X' \Phi^{-1} X)^{-1} X' \Phi^{-1} Y, \quad \Phi = \Sigma \otimes I_q \quad (۷.۵)$$

با توجه به اینکه $\Phi^{-1} = \Sigma^{-1} \otimes I_q$ ، با استفاده از (۲.۵) برآوردگر کمترین توان‌های دوم β با فرض معلوم بودن Σ به صورت زیر است

$$\hat{\beta}_{LS} = \begin{bmatrix} \sigma_{11} X_1' X_1 & \sigma_{12} X_1' X_2 & \dots & \sigma_{1p} X_1' X_p \\ \sigma_{21} X_2' X_1 & \sigma_{22} X_2' X_2 & \dots & \sigma_{2p} X_2' X_p \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{p1} X_p' X_1 & \sigma_{p2} X_p' X_2 & \dots & \sigma_{pp} X_p' X_p \end{bmatrix} \times \begin{bmatrix} \sum_{i=1}^p \sigma_{1i} X_1' Y_i \\ \vdots \\ \sum_{i=1}^p \sigma_{pi} X_p' Y_i \end{bmatrix} \quad (۸.۵)$$

فرض کنید اطلاعاتی بیشتر در مورد بردار پارامتر β به صورت یک مجموعه قید خطی $R\beta = r$ داریم که در

آن $R_{k \times k}$ ماتریسی معلوم از پیش تعیین شده و r برداری با k مولفه است. بنابراین به تعداد $s = (p - 1)k$

محدودیت داریم که بر فضای پارامتر $\beta \in \mathbb{R}^k$ اعمال شده است. این گونه قید خطی معمولاً به صورت ضرایب

مقابل در مسائل مربوط به طرح آزمایش‌ها و آزمون‌های فرضیه در جداول آنالیز- واریانس بکار می‌روند.

قبل از اینکه بخواهیم آماره آزمون برای فرضیه $H_0: R\beta = r$ را بیابیم، باید توجه کرد که برآوردگر

کمترین توان‌های دوم محدودشده بردار پارامتر β تحت قید $R\beta = r$ برابر است با

$$\hat{\beta}_R = \hat{\beta}_{LS} - (X' \Phi^{-1} X)^{-1} R' [R(X' \Phi^{-1} X)^{-1} R']^{-1} R(\hat{\beta}_{LS} - r) \quad (9.5)$$

حال شرایطی را در نظر بگیرید که Σ مجهول است. در این حالت، برآوردگر Σ برابر است با

$$\hat{\Sigma} = \frac{(Y - X\hat{b})'(Y - X\hat{b})}{q - k} \quad (10.5)$$

که در آن

$$\hat{b} = (X'X)^{-1} X'Y \quad (11.5)$$

زلنر (۱۹۶۲) نشان داد، با جایگذاری (۱۱.۵) در (۹.۵) برآوردگر کمترین توان‌های دوم تعمیم‌یافته ($GLSE$)

پارامتر β برابر است با

$$\begin{aligned} \hat{\beta} &= (X' \hat{\Phi}^{-1} X)^{-1} X' \hat{\Phi}^{-1} Y + O(q^{-1}) \\ &= \left[\begin{array}{cccc} s_{11} X_1' X_1 & s_{12} X_1' X_2 & \dots & s_{1p} X_1' X_p \\ s_{21} X_2' X_1 & s_{22} X_2' X_2 & \dots & s_{2p} X_2' X_p \\ \vdots & \vdots & \ddots & \vdots \\ s_{p1} X_p' X_1 & s_{p2} X_p' X_2 & \dots & s_{pp} X_p' X_p \end{array} \right]^{-1} \left[\begin{array}{c} \sum_{i=1}^p s_{1i} X_1' Y_i \\ \vdots \\ \sum_{i=1}^p s_{pi} X_p' Y_i \end{array} \right] \quad (12.5) \end{aligned}$$

که در آن $\hat{\Phi} = \hat{\Sigma} \otimes I_q$ و $O(q^{-1})$ نشان‌دهنده مقداری در رتبه q^{-1} در احتمال است و s_{ii} مقادیر برآوردشده

σ_{ii} هستند. یعنی توزیع مجانبی $q^{1/2}(\hat{\beta} - \beta)$ و $q^{1/2}(\hat{\beta} - \beta)$ هر دو یکسان و نرمال می‌باشند. به عبارتی

$$q^{1/2}(\hat{\beta} - \beta) \sim AN_k[\bullet, (X' \hat{\Phi}^{-1} X)^{-1}] \quad (13.5)$$

حال اگر قرار دهیم $S = X' \Phi^{-1} X$ ، می‌توانیم برآوردگر کمترین توان‌های دوم تعمیم‌یافته را به صورت زیر بنویسیم

$$\hat{\beta} = S^{-1} X' \Phi^{-1} Y + O(q^{-1}) \quad (14.5)$$

همچنین با جایگذاری (۱۱.۵) در (۱۰.۵) برآوردگر کمترین توان‌های دوم تعمیم‌یافته محدودشده ($RGLSE$) β با فرض $r = 0$ عبارتست از

$$\hat{\beta}_R = \hat{\beta}_{LS} - S^{-1} R' (RS^{-1} R')^{-1} R \hat{\beta} + O(q^{-1}) \quad (15.5)$$

و بطور مشابه

$$q^{1/2}(\hat{\beta}_R - \beta) \sim AN_k[-\delta, W] \quad (16.5)$$

که در آن $\delta = S^{-1} R' (RS^{-1} R')^{-1} R \beta - r$ و $W = S^{-1} - S^{-1} R' (RS^{-1} R')^{-1} R S^{-1}$. واضح است که تحت فرضیه $H_0: R\beta = 0$ ، داریم $\delta = 0$ و لذا آریبی مجانبی $RGLSE$ صفر می‌باشد که می‌توان نتیجه گرفت برآوردگر $\hat{\beta}_R$ بطور مجانبی برای β ناریب است.

در این مدل نیز می‌توان برای وزن‌دار کردن میانگین توان‌های دوم خطا، از عبارت $T = \Phi - XUX'$ استفاده کرد و برآوردگر منفرد را بصورت زیر بدست آورد

$$\beta_R^* = \beta_{LS}^* - S^{-1} R' (RS^{-1} R')^{-1} R \beta_{LS}^* + O(q^{-1}) \quad (17.5)$$

که در آن $C = X' T^{-1} X$ و $\beta_{LS}^* = C^{-1} X' T^{-1} Y + O(q^{-1})$.

حال فرض کنید بین ستون‌های ماتریس طرح، همخطی وجود داشته باشد، در این صورت برآوردگر ریج

محدودشده را به صورت زیر ارائه می‌دهیم

$$\hat{\beta}_R(k) = (kW + I)^{-1} \beta_R^* + O(q^{-1}) \quad (18.5)$$

با توجه به اینکه یکی از اهداف اصلی این بخش، بررسی بهترین انتخاب برآوردگر، از نظر بهینگی در مدل سرمایه است، لذا از این به بعد فرض می‌کنیم ماتریس R و بردار r در مجموعه قید خطی اعمال شده بر روی فضای پارامتر به صورت زیر باشند

$$R = \begin{bmatrix} I_{k_1} & -I_{k_2} & \cdot & \dots & \cdot & \cdot & \cdot \\ \cdot & I_{k_2} & -I_{k_3} & \dots & \cdot & \cdot & \cdot \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ \cdot & \cdot & \cdot & \dots & I_{k_{p-2}} & -I_{k_{p-1}} & \cdot \\ \cdot & \cdot & \cdot & \dots & \cdot & I_{k_{p-1}} & -I_{k_p} \end{bmatrix}, \quad r = \begin{bmatrix} \cdot \\ \cdot \\ \vdots \\ \vdots \\ \vdots \\ \cdot \end{bmatrix}$$

عبارات در (۱۳.۵) را می‌توان به صورت ساده‌تر فرضیه صفر زیر نشان داد که در اکثر مطالعات رگرسیونی مورد استفاده قرار می‌گیرد.

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_p \quad (19.5)$$

فرضیه صفر در (۱۴.۵) بیان می‌کند که واحدهای جزئی یا ضرایب رگرسیونی جزئی همسان هستند یا به عبارتی اریبی تجمعی در مدل وجود ندارد. برای آگاهی بیشتر در مورد اهمیت این فرضیه و اریبی تجمعی در مدل‌های خطی تجمعی ساده، تیل^۱ و زلنر (۱۹۶۲) را ببینید.

۲.۵ ویژگی‌های برآوردگر

در این بخش، به دلیل اینکه اثبات ویژگی‌های برآوردگر مشابه برآوردگر ارائه‌شده در فصل (۲) است، از ذکر اثبات‌ها صرف‌نظر شده است.

ویژگی ۱.۵ $\hat{\beta}_R(k)$ برآوردگری اریب و انقباضی برای β است.

^۱Theil

ویژگی ۲.۵ کوواریانس $\hat{\beta}_R(k)$ کمتر از کوواریانس $\hat{\beta}_R$ است.

$$Cov(\hat{\beta}_R(k)) \leq Cov(\hat{\beta}_R)$$

۳.۵ نتیجه‌گیری

ویژگی‌های ارائه‌شده برای برآوردگر ريج در مدل SUR منفرد، نشان‌دهنده کارایی این برآوردگر می‌باشد.

پیوست الف

خلاصه و پیشنهادات برای تحقیقات آینده

الف. ۱. خلاصه

بهبود دادن برآوردگرهای معادلات رگرسیونی، زمینه کاری بسیاری از آماردانان دهه اخیر است. یکی از عواملی که منجر به تلاش برای یافتن برآوردگرهای بهبودیافته می‌شود، وجود همخطی در ماتریس طرح است. همانطور که در این پایان‌نامه شرح آن گذشت، یکی از این برآوردگرهای بهبودیافته که در سال‌های اخیر از سوی آماردانان مورد توجه زیادی قرار گرفته است، برآوردگر ریج می‌باشد. وقتی که با مسئله همخطی مواجه هستیم، برآوردگر ریج در کلاس برآوردگرهای اریب، بهترین برآوردگر است و همانطور که نشان دادیم، خطای آن از خطای برآوردگر کمترین توان‌های دوم خطا 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 چندمتغیره با در نظر گرفتن مسئله همخطی و استفاده از داده‌هایی که در آن‌ها این مشکل به چشم می‌خورد، برآورد کنیم.

ابتدا برآوردگری را برای مدل رگرسیونی چندمتغیره خطی وقتی که ماتریس واریانس-کوواریانس خطا، $I\sigma^2$ است و اطلاعات پیشین تجربی به صورت محدودیتی روی فضای پارامترها اعمال شده است، یافتیم و به منظور نشان دادن کارایی این برآوردگر، خطای آن را محاسبه و به کمک شبیه‌سازی و مثال عددی برتری آن را بر برآوردگر کمترین توان‌های دوم خطای محدودشده، نشان دادیم.

سپس همان معادله رگرسیونی را در حالت کلیتری که ماتریس واریانس-کوواریانس بردار خطا، ثابت نبوده و برابر با Σ^2 باشد، بررسی کردیم و در حالتی که ماتریس Σ منفرد است، ضرایب مدل رگرسیونی را برآورد کردیم. در این مدل خطی منفرد نیز، برآوردگر ریج محدودشده بر برآوردگر کمترین توان‌های دوم محدودشده برتری دارد.

از آنجایی که محققان معمولاً اطلاعات پیشین دقیق ندارند و این اطلاعات نیز می‌توانند دستخوش خطا شوند، در این پایان‌نامه سعی بر آن شده است که حالی کلی‌تر را برای مدل رگرسیونی محدودشده در نظر گرفته، به عبارتی ضرایب مدل رگرسیونی را در حالتی که محدودیت اعمال شده روی فضای پارامتر دارای جمله خطای تصادفی می‌باشد، برآوردگر و ویژگی‌های آن بررسی شد و علاوه بر آن به کمک شبیه‌سازی کارایی برآوردگر تایید شد.

و در آخر به منظور استفاده بیشتر و مفیدتر از معادلات رگرسیونی در اقتصاد و شرکت‌های تجاری و...، مدل رگرسیونی را روی مدل به ظاهر نامرتبط تعمیم دادیم.

الف. ۲. پیشنهادات برای پژوهش‌های آینده (ارائه زمینه‌های تحقیق)

با توجه به نتایج بدست 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.

- بدست آوردن خطای برآوردگر ارائه‌شده در فصل سوم بر حسب مقادیر ویژه ماتریس‌ها و مقایسه این خطا با خطای برآوردگر کمترین توان‌های دوم خطای محدودشده، به منظور بررسی کارایی این برآوردگر.
- در مدل ارائه‌شده در فصل چهارم (مدل با محدودیت تصادفی)، می‌توان وقتی که خطای مدل و همچنین خطای محدودیت، توزیع t چندمتغیره یا در حالت کلی‌تر توزیع بیضی‌گون دارند، ضرایب مدل رگرسیونی را برآورد کرد و ویژگی‌های برآوردگر را بررسی کرد.
- به منظور نشان دادن کارایی برآوردگر ارائه‌شده در فصل پنجم (برآوردگر مدل SUR)، شبیه‌سازی مونت-کarlo

کارلو را انجام دهیم.

• تلفیق مدل با محدودیت تصادفی و مدل، SUR و ارائه برآوردگر برای ضرایب مدل رگرسیونی در ایم مدل جدید.

• استفاده از معیار آکایی به جای معیار توان دوم خطا برای مقایسه برآوردگرها و مقایسه نتایج آن با نتایج بدست آمده در این مجموعه.

پیوست ب

جبر ماتریس‌ها

تعریف ب.۱.۰. ماتریس متعامد: ماتریس‌های مربعی که هر جفت از ستون‌ها شامل دو بردار متعامد است.

قضیه ب.۱.۰. اگر A ماتریسی متعامد باشد، آنگاه $A' = A^{-1}$.

قضیه ب.۲.۰. ماتریس A متعامد است اگر و فقط اگر $A'A = I$.

تعریف ب.۲.۰. ماتریس وارون: ماتریس A^{-1} را وارون ماتریس A می‌نامیم اگر و تنها اگر $AA^{-1} = I$.

قضیه ب.۳.۰. وارون وارون ماتریس A ، برابر خود ماتریس A است. یعنی $(A^{-1})^{-1} = A$.

قضیه ب.۴.۰.

$$(A')^{-1} = (A^{-1})' \quad (\text{ب.۱})$$

قضیه ب.۵.۰.

$$(AB)^{-1} = B^{-1}A^{-1} \quad (\text{ب.۲})$$

قضیه ب.۶.۰.

$$|A^{-1}| = \frac{1}{|A|} \quad (\text{ب.۳})$$

قضیه ب.۷.۰. وارون یک ماتریس متقارن، متقارن است.

تعریف ب.۳.۰. اثر ماتریس: اگر A ماتریسی مربعی باشد، آنگاه اثر A ، مجموع درایه‌های قطر اصلی A است.

اثر ماتریس A را با نماد $tr(A)$ نشان می‌دهیم.

قضیه ب.۸.۰. اگر A, B و C ماتریس‌های مربعی باشند، آنگاه $tr(AB) = tr(BA)$ و $tr(ABC) = tr(BCA) = tr(CAB)$ در صورتی که ضرب بین آن‌ها تعریف شده باشد.

تعریف ب.۴.۰. ماتریس قطری: ماتریس مربع متقارنی است که همه درایه‌های آن، به جز درایه‌های روی قطر اصلی، برابر صفر است. به عبارتی اگر A ماتریس قطری p بعدی باشد، داریم:

$$\begin{bmatrix} a_{11} & & 0 \\ & \ddots & \\ 0 & & a_{pp} \end{bmatrix}$$

تعریف ب.۵.۰. ماتریس اسکالر: اگر درایه‌های قطر اصلی ماتریس قطری با هم برابر باشند، آنگاه این ماتریس را اسکالر می‌نامند.

تعریف ب.۶.۰. ماتریس واحد: ماتریس اسکالری که درایه‌های روی قطر اصلی آن برابر با یک باشند، ماتریس واحد (یکه) نامیده می‌شود.

تعریف ب.۷.۰. ماتریس معین و نیمه‌معین مثبت: ماتریس A را مثبت معین می‌گویند اگر تمام مقادیر ویژه A مثبت باشد و اگر مقادیر ویژه A نامنفی باشند، گفته می‌شود ماتریس A نیمه‌معین مثبت است.

قضیه ب.۹.۰. ماتریس A نیمه‌معین مثبت است اگر و فقط اگر ماتریس نامنفرد P وجود داشته باشد بطوریکه $A = P'P$.

برهان: رنچر^۱ ۲۰۰۱ را ببینید.

تعریف ب.۸.۰. دترمینان یک ماتریس: دترمینان ماتریس مربعی A با بعد $p \times p$ ، یک مقدار عددی است که شامل مجموع تعدادی جمله است که هر جمله حاصلضرب p درایه ماتریس A می‌باشد.

قضیه ب.۱۰.۰. اگر A و B دو ماتریس مربعی باشند، آنگاه داریم $|AB| = |BA| = |A| \times |B|$

قضیه ب.۱۱.۰. اگر A ماتریسی $p \times p$ باشد، آنگاه $|kA| = k^p |A|$

تعریف ب.۹.۰. ماتریس منفرد: اگر $|A| \neq 0$ ، آنگاه ماتریس A را ماتریس منفرد (تکین) می‌نامیم. ماتریسی که دترمینان آن غیر صفر باشد، ماتریس نامنفرد (ناتکین) نامیده می‌شود.

قضیه ب.۱۲.۰. فرض کنید A یک ماتریس قطری نامنفرد و بصورت زیر باشد

$$A = \text{Diag}(a_1, a_2, \dots, a_p) \quad (\text{ب.۴})$$

در این صورت وارون آن، ماتریس قطری B است با $B = \text{Diag}(\frac{1}{a_1}, \frac{1}{a_2}, \dots, \frac{1}{a_p})$

^۱ Rencher

قضیه ب.۱۳.۰. ماتریس منفرد، وارنپذیر است.

تعریف ب.۱۰.۰. وابسته خطی: می‌گوییم بردارهای $x_1, x_2, \dots, x_p$ وابسته خطی هستند اگر و فقط اگر رابطه خطی به صورت زیر بین آن‌ها وجود داشته باشد.

$$c_1 x_1 + c_2 x_2 + \dots + c_p x_p = 0$$

که در آن همه ضرایب عددی $c_1, c_2, \dots, c_p$ صفر نیستند. در صورتی که یک مجموعه بردار، وابسته خطی نباشند، مستقل خطی هستند.

تعریف ب.۱۱.۰. رتبه ماتریس: رتبه ماتریس A که به صورت $r(A)$ یا $\text{Rank}(A)$ نوشته می‌شود، بعد بزرگترین زیر ماتریس مربعی و نامنفرد A است. همچنین تعداد ستون‌ها یا سطرهای مستقل خطی یک ماتریس، نشان‌دهنده رتبه آن ماتریس است.

قضیه ب.۱۴.۰. اگر A ماتریسی $p \times p$ و نامنفرد باشد، آنگاه رتبه آن برابر p است.

قضیه ب.۱۵.۰. اگر $A, m \times n$ باشد، آنگاه $r(A)$ نمی‌تواند از مینیمم m و n بیشتر باشد.

تعریف ب.۱۲.۰. وارون تعمیم‌یافته یک ماتریس: اگر B ماتریسی باشد بطوریکه Bh جوابی برای معادلات سازگار $Ax = h$ باشد، گوییم B وارون تعمیم‌یافته (یا بطور اختصار g -وارون) A است و با نماد A^- نشان می‌دهیم.

قضیه ب.۱۶.۰. وارون تعمیم‌یافته یک ماتریس، یکتا نیست.

قضیه ب.۱۷.۰. $AA^-A = A$

قضیه ب.۱۸.۰. فرض کنید A ماتریسی مربعی باشد، آنگاه مشتق ماتریس معکوس A نسبت به کمیت k از رابطه زیر بدست می‌آید

$$\frac{dA^{-1}}{dk} = -A^{-1} \frac{dA}{dk} A^{-1}$$

تعریف ب.۱۳.۰. عملگر ضرب کرونکر: فرض کنید $A = [a_{ij}]$ و B دو ماتریس در اندازه‌های $(m \times n)$ و $(p \times q)$ هستند. در این صورت ضرب کرونکر A و B ($A \otimes B$) به صورت زیر تعریف می‌شود

$$A \otimes B = \begin{bmatrix} a_{11}B & a_{12}B & \dots & a_{1n}B \\ a_{21}B & a_{22}B & \dots & a_{2n}B \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}B & a_{m2}B & \dots & a_{mn}B \end{bmatrix}$$

قضیه ب.۱۹۰.

$$(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}$$

قضیه ب.۲۰۰.

$$(A \otimes B)' = A' \otimes B'$$

اثبات قضایای (الف.۱) تا (الف.۲۰) را می‌توان در سرل^۱ (۱۹۸۲) دید.

قضیه ب.۲۱۰. فرض کنید $V = AY$ و $W = BY$ که در آن A ماتریس ثابت $n \times p$ و B ماتریس ثابت $p \times n$ و Y بردار تصادفی p بعدی با ماتریس واریانس-کواریانس Σ باشد، آنگاه

$$Var(V) = Var(AY) = A\Sigma A'$$

$$Cov(V, W) = Cov(AY, BY) = A\Sigma B'$$

برای اثبات اندرسن^۲ (۲۰۰۳) را ببینید.

لم ب.۱۰. فرض کنید A یک ماتریس متقارن حقیقی است، بنابراین مقادیر ویژه A نامنفی باشند $\Leftrightarrow p'Ap \geq 0, \forall p \Leftrightarrow A \geq 0$.

برهان: راثو (۱۹۹۵) را ببینید.

لم ب.۲۰. فرض کنید $A \geq 0$ و $B \geq 0$ بنابراین

$$\xi_n(A)\zeta_i(B) \leq \tau_i(AB) \leq \xi_1(A)\zeta_i(B) \quad \forall i = 1, 2, \dots, n$$

که در آن $\xi_1 \geq \xi_2 \geq \dots \geq \xi_n$ و $\zeta_1 \geq \zeta_2 \geq \dots \geq \zeta_n$ و به ترتیب مقادیر ویژه A ، B و $A+B$ هستند.

برهان: راثو (۱۹۹۵) را ببینید.

^۱Searle

^۲Anderson

لم ب.۳.۰. فرض کنید $A \geq 0$ و $B \geq 0$ بنابراین

$$\xi_n(A)\zeta_i(B) \leq \nu_i(AB) \leq \xi_1(A)\zeta_i(B) \quad \forall i = 1, 2, \dots, n$$

که در آن $\nu_1 \geq \nu_2 \geq \dots \geq \nu_n$ و $\zeta_1 \geq \zeta_2 \geq \dots \geq \zeta_n$ و $\xi_1 \geq \xi_2 \geq \dots \geq \xi_n$ به ترتیب مقادیر ویژه A ، B و AB هستند.

برهان: راثو (۱۹۹۵) را ببینید.

پیوست پ

عامل تورم واریانس

یکی از شیوه‌های تشخیص وجود همخطی که کاربرد زیادی دارد، عامل تورم واریانس می‌باشد. این عامل نشان می‌دهد که واریانس ضرایب برآوردشده تا چه حد نسبت به حالتی که متغیرهای برآوردشده همبستگی خطی ندارند، متورم شده است. عامل تورم واریانس از رابطه زیر بدست می‌آید

$$VIF_i = \frac{1}{1 - R_i^2}, \quad i = 1, 2, \dots, p - 1$$

که در آن R_i^2 ضریب تعیین مدلی است که در آن متغیر مستقل X_i روی سایر متغیرهای مستقل رگرس شده است.

در حالتی که $R_i^2 = 0$ یا به عبارتی دیگر X_i به صورت خطی رابطه‌ای با سایر متغیرهای مستقل در مدل رگرسیونی نداشته باشد، VIF_i برابر با یک خواهد بود. هنگامی که $VIF_i, R_i^2 \neq 0$ بزرگتر از یک خواهد بود که نشان می‌دهد، واریانس X_i به دلیل وجود همبستگی بین متغیرهای مستقل چقدر متورم شده است و مقدار آن افزایش پیدا کرده است.

هنگامی که یک متغیر مستقل، وابستگی خطی کامل با سایر متغیرهای مستقل 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 مستقل، 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 برآوردهای روش کمترین توان‌های دوم دارد.

پیوست ت

عدد شرطی

در آنالیز عددی، عدد شرطی یک تابع، بدترین وضعیتی که مقدار آن تابع نسبت به تغییرات کوچک می‌تواند تغییر کند را اندازه‌گیری می‌کند. مسئله‌ای با عدد شرطی پایین را خوب شرطی شده و معادله‌ای با عدد شرطی بالا را ناقص شرطی شده می‌گویند. به عنوان مثال، عدد شرطی در ارتباط با معادله $AX = B$ ، حدودی را برای نادرست برآورد شدن X از راه‌حل‌های تقریبی ارائه می‌دهد. دقت کنید که عدد شرطی نسبت به تغییرات در B نیز حساس است، یعنی اگر عدد شرطی زیاد باشد، حتی خطای کوچکی در B ممکن است باعث ایجاد خطای بزرگ در برآورد X شود. از سوی دیگر، اگر عدد شرطی کوچک باشد، خطا در X خیلی بزرگتر از خطا در B نخواهد بود.

عدد شرطی به صورت نسبت خطای نسبی در X به خطای نسبی در B تعریف شده است.

فرض کنید که E خطای نسبی در B و A ماتریس مربعی باشد، بنابراین $A^{-1}E$ خطا در معادله $A^{-1}B$

است. نسبت خطای نسبی در معادله به خطای نسبی در B به صورت زیر است

$$\frac{\|A^{-1}E\|/\|A^{-1}B\|}{\|E\|/\|B\|} \quad (\text{ت. ۱})$$

که به سادگی به صورت زیر تبدیل می شود

$$(\|A^{-1}E\|/\|E\|) \cdot (\|B\|/\|A^{-1}B\|) \quad (۲.ت)$$

بزرگترین مقدار (برای B و E غیر صفر) حاصلضرب دو نرم عملگر به صورت زیر است

$$k(A) = \|A\| \cdot \|A^{-1}\|$$

تعریف مشابهی هست که برای نرم های سازگار بکار می رود، غالباً این عدد در یک عبارت جبری خطی عددی

بوجود می آید که عدد شرطی ماتریس A نامیده می شود. البته این تعریف وابسته با انتخاب نرم است.

۱. اگر $\|\cdot\|$ نرم σ^2 باشد، بنابراین

$$k(A) = \frac{\sigma_{Max}(A)}{\sigma_{Min}(A)} \quad (۳.ت)$$

که σ_{Max} و σ_{Min} به ترتیب، بزرگترین و کوچکترین مقادیر منفرد ماتریس A هستند.

۲. اگر A نرمال باشد، آنگاه

$$k(A) = \left| \frac{\lambda_{Max}(A)}{\lambda_{Min}(A)} \right| \quad (۴.ت)$$

که در آن $\lambda_{Max}(A)$ و $\lambda_{Min}(A)$ به ترتیب، بزرگترین و کوچکترین مقادیر ویژه ماتریس A هستند.

۳. اگر A ماتریسی واحد باشد، آنگاه

$$k(A) = ۱ \quad (۵.ت)$$

۴. اگر $\|\cdot\|$ نرم بینهایت باشد و A ماتریسی منفرد و پایین مثلثی باشد (یعنی $a_{ii} \neq ۰, \forall i$)، بنابراین

$$k(A) = \frac{Max_i(|a_{ii}|)}{Min_i(|a_{ii}|)} \quad (۶.ت)$$

واژه‌نامه‌ی فارسی به انگلیسی

Aggregation bias	اریبی تجمعی
Over estimation	اریبی مثبت
Under estimation	اریبی منفی
Prior information	اطلاع پیشین
Trace	اثر
Regrassion to the mean	بازگشت به میانگین
Restricted estimator	برآوردگر محدودشده
Optimality	بهینگی
Ridge parameter	پارامتر ریج
Rank	رتبه
Multiple linear regression	رگرسیون چند متغیره خطی
Beggining-of-year capital stock	سهام بورس
Outstandy shaves	سهام واریزنشده
Linear constrain	قید خطی
Scaler matrix	ماتریس اسکالر

Diagonal matrix	ماتریس قطری
Block-diagonal matrix	ماتریس بلوکی - قطری
Design matrix	ماتریس طرح
Restricted matrix	ماتریس محدودیت
Singular matrix	ماتریس منفرد
Non-singular matrix	ماتریس نامنفرد
Unit matrix	ماتریس واحد
Orthogonal	متعامد
Response variable	متغیر پاسخ
Regressor variable	متغیر رگرسیونی
Symetric	متقارن
Singular linear model	مدل خطی منفرد
Inverse	معکوس
Generalized inverse	معکوس تعمیم یافته
Positive definite	معین مثبت
Constrast	مقابلہ
Linear dependence	وابسته خطی
Hat matrix	هت ماتریس - ماتریس برازش
Correlation	همبستگی
Multicollinearity	همخطی چندگانه

مراجع و مأخذ

- [1] Adkenize. F, Kacilanlar. S, (1995). On the almost unbiased generalized Lui estimator and unbiased estimation of the bias and MSE. *Commun. Statist. Theor. Mehta.* **24**, 1789–1797.
- [2] Adkeniz. F, Erol. H, (2003). Mean Squared error matrix comparisons of some biased estimator in linear regression. *Commun. Statist. Theor. Mehta.* **32**, 2389–2413.
- [3] Alkhamisi. M. A, Shukur. G, (2008). Developing ridge parameters for *SUR* model. *Commun. Statist. Theor. Mehta.* **37(4)**, 544–564.
- [4] Alkhamisi. M. A (2009). Simulation study of new estimators combining the *SUR* ridge regression and the restricted least squares methodologies, *Statist. Papers* in press.
- [5] Anderson. T. W. (2003). *An Introduction to multivariate Statistical Analysis* 3rd Ed; John, Wiley, New York.
- [6] Baaltagi. B (1980). On Seemingly unrelated regressions with error components, *Econometrica.* **48**, 1547–1552.
- [7] Creel. M, Forell. M (1996). *SUR* estimation of multiple time– series models with Heteroskedasticity and serial of unknown form. *Econ. Lett.* **53**, 239–245.
- [8] Dimov. I. T, Dimov. T. T, Grov (1998). A new iterative Monte Carlo approach for inverse matrix problem. *Journal of Computational and Applied Mathematics.* **92**, 15–35.

- [9] Gibbons. D. G (1981). A simulation study of some ridge estimators. *J. Amer. Statist. Assoc.* **76**, 1131-19.
- [10] Gisela. M, Kibria. B. M. G (2009). One some ridge regression estimators: An emprical comparisons. *Comm Statist. Sim. Comp.* **38(3)**, 621-630.
- [11] Grob, J. (2003). Restricted ridge estimation. *Statistic and probability Letters.* **65**, 57-64.
- [12] Gross. E. (2003). Restricted ridge estimation. *Statistic and pribability Letters.* **65**, 57-64.
- [13] Gruber. M. H. J. (1998). *Improving Efficiency by Shrinkage* NewYork: Marcel Dekker.
- [14] Himelblau. D. M. (1970). *Pross analysis by statistical methods*, Jhon Wiley, New York.
- [15] Hocking R. R, Speed. F. M, Lynn. M. J. (1976). A class of baised estimator in linear estimator in linear restricted model with non- normal disturbance. *Comm. Statist.* **A25**, 2349-2369.
- [16] Horel. A. E, Kennard. R. W. (1970). Ridge regression baised estimation for non-orthogonal problems. *Thechnometrics.* **12**, 69-89.
- [17] Horel. P. G, Port. S. C (1972). *Introcuction to stochastic procces*. Waveland Press Press, New York, 1994.
- [18] Horel. A. E, Kennard. R. W. (1975). Ridge regression some simulation. *Com-munication in Statistics.* 4105-4123.
- [19] Kaciranlar. S, Sakallioglu. S, Adkenize. F, Styne. G. P. H, Werner. H. J. (1999). A new biased estimator in linear regression and a detailed analysis of the widly-analysed data seton portland cement. *Sankhya Ind, J. Statist.* **61**, 443-456.
- [20] Kibria. B. M. G (2003). Performanceof some new ridge estimation. *Comm. Statist. Sim. Comp.* **32**, 419-435.

- [21] Kunugi. T, Tamura. T, Natio. T (1961). New acetylene process uses hydrogen dilution. *Chen. Eng. Prog.* **57**, (1965) 43-49.
- [22] Lawless. J. F, Wang. P (1976). A dimulation study of ridge and other regression estimators. *Comm. Statist. Theo. Mehta.* **5**, 307-323.
- [23] Lawless. J. F. (1978). Ridge and related estimation procedure. *Comm. Statist.* **A7**, 139-164.
- [24] Lui. K. (1993). A new class of biased estimate in linear regression. *Comm. Statist. Theo. Mehta.* **22**, 393-402.
- [25] Lui. K. (2003). Using Lui- type estimator to combat collinearity. *Comm. Statist. Theo. Meth.* **32**, 1009-1020.
- [26] Mansson. K, Shuker. G, Kibria. B. M. G. (2010). A simulation study of some ridge regression estimators under diferent distrinutional assumptions. *Comm. Statist. Sim. Comp.* **39(8)**, 1639-1670.
- [27] Marquart. D. W. (1970), Generalized inverse ridge regression biased linear estimation and non-linear estimation, *Technometrics.* **12**, 591-612.
- [28] Marquart. D. W, Snee. R. D. (1975). Ridge regression practice. *J. Amer. Statist.* **29(1)**, 3-20.
- [29] Mason. R. L, Gunst. R. F and Webster. J. T. (1975). Ridge analysis and problems of multicolliearity. *Commun. Statist.* **4(3)**, 277-292.
- [30] Mc. Donald. G. C, Galarneau. D. I. (1975). A monte carlo evalution of some ridge type estimators. *Amer. Statist. Assoc.* **70** , 407-416.
- [31] Montgomery. D. G. and Peck. A. E. (1992). Introduction to linear regression analysis, 2nd. ed.
- [32] Montgomery. D. G. and Peck. A. E, and Vining. G. G. (2001). Introduction to linear regression analysis, Jhon Wiley, 3rth Ed, New York.
- [33] Moon. H. R, Perron. P. (2004). Efficient estimation of SUR cointegration regression model and testing for purchasing power parity. *Econ. Rev.* **23**, 293-323.

- [34] Newhouse. J. P. and Oman. S. D. (1971). An evalution of ridge estimators. Rand Corporation. P-716-PR.
- [35] Rahakrishna Rao. C. and Toutenburg. H. (1995). Linear Models: Least Squares and Alternatives, Springer- Verlag.
- [36] Rao. C. R. (1995). Linear models- Least Squares and alternatives, Springer.
- [37] Revankar. N. S. (1976). Use of restricted residuals in SUR systems: Some finite sample results. *J. Amer. Statist. Assoc.* **71(353)**, 183-188.
- [38] Saleh. A. K. Md. E. and Kibria. B. M. G. (1993). Performance of some new preliminary test ridge regression estimators and their properties. *Comm. Statist.* **A22**, 2747-2764.
- [39] Sarkar. N. (1992). A new estimator combining the ridge regression and the restricted least squares methods of estimation. *Communication in Statistics- Theory and Method.* **21**, 1987-2000.
- [40] Saleh. A. K. Md. Ehsanes. (2006). Theory of preliminary test of Stein- type estimation with applications Jhon Wiley, New York.
- [41] Searl, S. R. (1982). Matrix Algebra Useful for Statistics, Jhon Wiley, New York.
- [42] Shi. J. (2001). The conditional ridge- type estimation of regression coefficient in restricted linear regression model. Journal of Shanxi Teacher's University Natural Science Edition. **15**, 10-16.
- [43] Swamy. P. A. V. B. and Mehta. J. S. (1977). A note of minimum avrage risk estimator for cofficients in linear models. *Commun. Statist. Theo. Mehta.* **6** 1161-1186.
- [44] Swamy. P. A. V. B., Mehta. J. S. and Robert. P. N. (1978). Two methods of evaluting Horel and Kennard's ridge regression. *Commun. Statist. Theo. Meth.* **12**, 1133-1155.
- [45] Swindel. B. F. (1976). Good estimators based on prior information. *Commun. Statist. Theor. J. Roy. Statist. Soc. B.* **36**, 103-106.

- [46] Theil. H. (1963). On the use of incomplete prior information in regression analysis. *J. Amer. Statist. Assoc.* **58**, 401–414.
- [47] Theobald. C. M. (1974). Generalizations of mean square error applied to ridge regression. *J. Roy. Statist. Soc.* **36**, 103–106.
- [48] Wang. S. (1994). *The Inequalities of Matrices* (Hefei: The Science Press).
- [49] Wood. H., Steinnour. H. H. and Strake. H. R. (1932). Effect of composition of portland cement on heat envolved during hardening. *Indust. Eng. Chem.* **24**, 1207–1242.
- [50] Zhang. Ch. and Yang. Hu. (2007). 'The conditional ridge– type estimation in singular linear model with linear equality restriction' *Statistics.* **41:6**, 485–494.

Abstract

In this project considers several estimators for estimating ridge regression estimators in singular linear model. In order to tackle the multicollinearity or ill-conditioning design matrices in singular linear model with linear equality restrictions, the conditional ridge-type estimation is proposed and some properties of the estimator are also studied. Furthermore, the conditional ridge-type estimation is compared with the restricted least squares estimator, and separate sufficient conditions are derived for the superiority of the new estimator over the restricted least squares estimator in terms of mean squares error and mean squares error matrix . A simulation study has been conducted to compare the performance of the estimators.



**Shahrood University
of Technology-Iran**

Faculty of Science

M.Sc thesis of Mathematics Statistics

Title:

Studying the performance of ridge regression estimators in singular linear models

By:

Samaneh Najarian

Supervisor:

Dr. Mohammad Arashi

Consultant:

Dr. Jafar Fahali

September 2011