

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ



دانشکده علوم ریاضی

رشته آمار، گرایش آمار ریاضی

پایان نامه کارشناسی ارشد

نسخه‌ای استوار از رگرسیون بریج

نگارنده: وحید گودرزی وانانی

استادان راهنما

دکتر محمدرضا ربیعی
دکتر محمد آرشی

بهمن ۱۳۹۷

تقدیم خالصانه و خاضعانه به:

پدرم به پاس سال‌های بی‌دریغش
مادرم به پاس سال‌های صبری بی‌انتہایش
برادرم به پاس صبری و همراهی فراوانش

سپاسگزاری

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

به مصداق لم یشکر المخلوق، لم یشکر الخالق، سپاس بی کران نثار پدر و مادرم، که ناتوان شدند تا من توانا شوم، مویشان سپید شد تا من روسفید شوم و عاشقانه سوختند تا گرمابخش وجودم و روشنگر راهم باشند. اگر پدرم و حمایت‌های بی دریغش نبود، این سطرها هیچ‌گاه نگاشته نمی‌شد. سپاسگزاری و قدردانی خالصانه و خاضعانه نثار اساتیدم، جناب آقای دکتر محمدرضا ربیعی و جناب آقای دکتر محمد آرشی که با حمایت‌های بی دریغشان تاکنون هیچ‌گاه تنهایی نگذاشته‌اند و با صبر و تلاش فراوان ایشان، توانستم پایان‌نامه کارشناسی ارشد را گردآوری نمایم.

سپاسگزاری و قدردانی خالصانه را تقدیم خانم دکتر مینا نوروزی راد می‌نمایم که همواره با صبر و تلاش فراوان، همانند خواهری دلسوز، در سختی‌ها راه‌گشا و راهنمایم بوده‌اند. از اساتید محترم، دکتر احمد نزاکتی، دکتر داود شاهسونی، دکتر حسین باغیشنی که در طول دوران تحصیلی‌ام صادقانه و عاشقانه تلاش نمودند تا برایم نقالی دانسته‌ها را به نقادی اندیشه‌ها تبدیل کنند، قدردانی می‌کنم. از مدیر گروه دلسوز و پرتلاش گروه آمار دانشگاه صنعتی شاهرود، جناب آقای دکتر ربیعی به دلیل خدمات، راهنمایی‌ها و زحماتشان سپاسگزارم.

از اساتید داور محترم، جناب آقای دکتر حسین باغیشنی و دکتر داود شاهسونی که با حضور دلگرمشان تصحیح و داوری این پایان‌نامه را برعهده گرفته‌اند، سپاسگزارم. و دوباره، بارها و بارها از همدلی و همراهی و همگامی **مهربان خانواده‌ام** که همواره با قلبی آکنده از عشق و معرفت، محیطی سرشار از سلامت، امنیت، آرامش و آسایش برایم فراهم آورده‌اند و مرا در رسیدن به اهدافم یاری رساندند بی‌نهایت سپاس گزارم.

وحید گودرزی وانانی

بهمن ۱۳۹۷

تعهد نامه

اینجانب وحید گودرزی وانانی دانشجوی کارشناسی ارشد رشته آمار ریاضی دانشکده علوم ریاضی دانشگاه صنعتی شاهرود، نویسنده پایان نامه با عنوان **نسخه‌ای استوار از رگرسیون بریج**، تحت راهنمایی جناب آقای دکتر **محمد رضا ربیعی** و جناب آقای دکتر **محمد آرشی** متعهد می‌شوم:

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

وحید گودرزی وانانی

بهمن ۱۳۹۷

مالکیت نتایج و حق نشر

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

چکیده

در روش‌های رگرسیونی، برآورد پارامترهای مدل یکی از مهم‌ترین ارکان تحلیل و پیشگویی می‌باشد. یکی از معمول‌ترین این روش‌ها برآوردگر کمترین توان‌های دوم است که وقتی بین متغیرهای تبیینی همخطی چندگانه وجود دارد، این برآوردگر کارایی لازم را ندارد. از این رو استفاده از برآوردگرهای انقباضی برای برآورد ضرایب مدل پیشنهاد شده‌اند. یکی از این روش‌ها برآوردگر بریج است.

علاوه بر مشکل همخطی، حضور داده ناهمگون در مجموعه داده‌ها نیز می‌تواند بر عملکرد نه‌تنها برآوردگر کمترین توان‌های دوم، بلکه بر عملکرد روش‌های انقباضی هم تاثیرگذار باشد. در خصوص حضور داده ناهمگون در مجموعه داده‌ها مطالعاتی انجام شده و روش‌های استوار متعددی پیشنهاد شده‌اند، اما این برآوردگرها با وجود همخطی بین متغیرهای تبیینی قادر به برازش خوب مدل نخواهند بود. از این رو روش‌هایی مورد نیاز است که بتواند حضور داده ناهمگون و همخطی بالا بین متغیرهای تبیینی را کنترل کند.

در این پایان‌نامه، علاوه بر معرفی برآوردگر بریج استوار به مطالعه رفتار مجانبی آن می‌پردازیم و با استفاده از روش‌های مکرر، الگوریتمی برای محاسبه برآوردگر بریج معرفی می‌کنیم. همچنین به‌عنوان حالت‌های خاص برآوردگر بریج، به مطالعه برآوردگرهای ریج و لاسو استوار خواهیم پرداخت و این برآوردگرها را در حالت فازی نیز بررسی و ارائه می‌کنیم.

کلمات کلیدی: روش‌های انقباضی، رگرسیون بریج، روش‌های استوار، رگرسیون بریج استوار، رگرسیون فازی، برآوردگرهای انقباضی استوار فازی.

فهرست مقالات مستخرج از پایان نامه

1. Rabiei, M. R., Arashi, A., Goodarzi, V. and Kouhsar, F. P. (2018), Fuzzy ridge linear regression analysis for fuzzy input output data. *Proceeding of Seminar on Fuzzy Statistics and Probability*. **8**, 156-164.

۲. گودرزی و.، ربیعی م. ر.، آرشی م. (۱۳۹۷). رگرسیون ریج استوار با متغیرهای پاسخ، توضیحی و ضرایب فازی. چهاردهمین کنفرانس آمار ایران، شاهرود، ایران.

فهرست مطالب

ش	فهرست تصاویر
ث	فهرست جداول
۱	۱ تعاریف و مقدمات
۱	۱.۱ مقدمه
۲	۲.۱ مدل رگرسیون چندگانه
۲	۳.۱ برآوردگر کمترین توان‌های دوم
۳	۴.۱ همخطی و برآوردگر جریمه‌شده
۶	۱.۴.۱ برآوردگر بریج
۸	۲.۴.۱ برآوردگر ریج
۹	۳.۴.۱ برآوردگر لاسو
۹	۵.۱ روش‌های استوار
۱۰	۱.۵.۱ بررسی مشاهدات ناهمگون
۱۳	۶.۱ برآوردگرهای استوار در مدل خطی
۱۴	۱.۶.۱ M-برآوردگرها
۱۶	۲.۶.۱ MM-برآوردگرها
۱۷	۷.۱ برآوردهای مبتنی بر باقی‌مانده‌های استوار مقیاس
۱۸	۱.۷.۱ S-برآوردها
۱۸	۲.۷.۱ L-برآوردها
۱۸	۸.۱ برآوردگرهای جریمه‌شده استوار
۱۹	۹.۱ مقدمه‌ای بر مجموعه‌های فازی
۲۰	۱.۹.۱ حساب اعداد فازی
۲۱	۲.۹.۱ فاصله بین دو عدد فازی
۲۴	۱۰.۱ رگرسیون فازی
۲۴	۱.۱۰.۱ انواع رگرسیون فازی

۲۴	رگرسیون کمترین توان‌های دوم فازی	۲۰۱۰.۱
۲۶	برآوردگر استوار فازی بر پایه فاصله قدرمطلق	۳۰۱۰.۱
۲۹	برآوردگر جریمه‌شده فازی	۴۰۱۰.۱
۳۱	۲ برآوردگر ریج استوار	
۳۱	۱.۲ مقدمه	
۳۲	۲.۲ MM-برآوردگر برای رگرسیون ریج استوار	
۳۴	۳.۲ معادلات نرمال موزون و الگوریتم مکرر	
۳۵	۱.۳.۲ انتخاب ثابت‌ها	
۳۵	۴.۲ شبیه‌سازی	
۵۲	۵.۲ تحلیل داده‌های چربی بدن	
۵۲	۱.۵.۲ معرفی داده‌های Vo2	
۵۴	۶.۲ برآوردگر ریج فازی	
۵۴	۱.۶.۲ برآوردگر ریج فازی معمولی	
۵۶	۲.۶.۲ شاخص‌های نیکویی برازش مدل	
۵۷	۳.۶.۲ تحلیل رگرسیون ریج فازی برای داده‌های سیمان پرتلند	
۵۸	۴.۶.۲ برآوردگر ریج استوار فازی	
۵۹	۵.۶.۲ تحلیل رگرسیون ریج استوار فازی برای داده‌های سیمان پرتلند	
۶۱	۳ برآوردگر لاسو استوار	
۶۱	۱.۳ مقدمه	
۶۲	۲.۳ برآوردگرهای LAD-LASSO	
۶۴	۳.۳ ویژگی‌های نظری	
۶۵	۴.۳ تعیین پارامتر تنظیم‌کننده	
۶۷	۵.۳ شبیه‌سازی	
۷۲	۶.۳ تحلیل داده‌های چربی بدن	
۸۲	۷.۳ برآوردگر استوار فازی با استفاده از جریمه لاسو	
۸۳	۱.۷.۳ مثال کاربردی	
۸۵	۴ برآوردگر بریج استوار	
۸۵	۱.۴ مقدمه	
۸۶	۲.۴ برآوردگر بریج استوار	
۸۸	۳.۴ ویژگی‌های مجانبی برآوردگر ریج استوار	
۹۶	۴.۴ الگوریتم محاسبه برآوردهای MM-Bridge	
۹۷	۱.۴.۴ الگوریتم LQA	

۹۸	۲.۴.۴ تقریب کمترین توان‌های دوم (LS): برآوردگر یک مرحله‌ای . . .
۹۹	۵.۴ شبیه‌سازی
۱۰۴	۶.۴ تحلیل داده‌های چربی بدن
۱۱۴	۷.۴ نتیجه‌گیری و آینده تحقیق
۱۱۷	۵ گزیده‌ای از کدها
۱۱۷	۱.۵ کدهای مربوط به فصل دوم
۱۱۹	۲.۵ کدهای مربوط به فصل سوم
۱۲۵	۳.۵ کدهای مربوط به فصل چهارم
۱۳۵	مراجع

فهرست تصاویر

۱.۱	نمایش مشاهدات ناهمگون (آ)، در راستای محور y ها و (ب)، در راستای محور x ها در یک مثال ساختگی	۱۰
۲.۱	خطوط رگرسیونی کمترین توان‌های دوم در حضور انواع داده ناهمگون . .	۱۲
۱.۲	ماتریس نمودار پراکنش ۸ متغیر توضیحی در برابر یکدیگر که حاکی از همخطی برای $n = ۲۰۰$ مشاهده و همخطی بالای $\rho = ۰/۹$ ، زمانی که ۲۰% درصد مشاهدات آلوده می‌شوند (یعنی $\tau = ۰/۲$).	۳۷
۲.۲	نمایش نقاط نافذ بین متغیرهای تبیینی و متغیر پاسخ برای حالت $n = ۲۰۰$ ، $\rho = ۰/۹$ و $\tau = ۰/۲$ بالا: نمایش داده‌پرت در راستای متغیر پاسخ، پایین: نمایش مجموعه نقاط نافذ در متغیرهای تبیینی که در آن اعداد $۱, ۲, \dots, ۸$ نشان‌دهنده متغیرهای تبیینی $x_1, \dots, x_8$ هستند.	۳۸
۳.۲	نمودار جعبه‌ای مربوط به خطاهای پیشگویی برای ریح استوار	۵۳
۱.۳	نمودار جعبه‌ای مربوط به خطاهای پیشگویی برای لاسو استوار	۸۱
۱.۴	نمودار جعبه‌ای مربوط به خطاهای پیشگویی برای بریج استوار	۱۱۳

فهرست جداول

۱.۲	نتایج شبیه‌سازی مربوط به ریج استوار، طرح ۱، $n = 50$ ، $\rho = 0/3$ و $\tau = 0$	۳۹
۲.۲	نتایج شبیه‌سازی مربوط به ریج استوار، طرح ۱، $n = 50$ ، $\rho = 0/99$ و $\tau = 0$	۴۰
۳.۲	نتایج شبیه‌سازی مربوط به ریج استوار، طرح ۱، $n = 200$ ، $\rho = 0/3$ و $\tau = 0$	۴۰
۴.۲	نتایج شبیه‌سازی مربوط به ریج استوار، طرح ۱، $n = 200$ ، $\rho = 0/99$ و $\tau = 0$	۴۱
۵.۲	نتایج شبیه‌سازی مربوط به ریج استوار، طرح ۱، $n = 50$ ، $\rho = 0/9$ و $\tau = 0/2$	۴۱
۶.۲	نتایج شبیه‌سازی مربوط به ریج استوار، طرح ۱، $n = 50$ ، $\rho = 0/99$ و $\tau = 0/2$	۴۲
۷.۲	نتایج شبیه‌سازی مربوط به ریج استوار، طرح ۱، $n = 200$ ، $\rho = 0/9$ و $\tau = 0/2$	۴۲
۸.۲	نتایج شبیه‌سازی مربوط به ریج استوار، طرح ۱، $n = 200$ ، $\rho = 0/99$ و $\tau = 0/2$	۴۳
۹.۲	نتایج شبیه‌سازی مربوط به ریج استوار، طرح ۲، $n = 100$ ، $\rho = 0/3$ و $\tau = 0$	۴۴
۱۰.۲	نتایج شبیه‌سازی مربوط به ریج استوار، طرح ۲، $n = 100$ ، $\rho = 0/99$ و $\tau = 0$	۴۵
۱۱.۲	نتایج شبیه‌سازی مربوط به ریج استوار، طرح ۲، $n = 300$ ، $\rho = 0/3$ و $\tau = 0$	۴۶
۱۲.۲	نتایج شبیه‌سازی مربوط به ریج استوار، طرح ۲، $n = 300$ ، $\rho = 0/99$ و $\tau = 0$	۴۷
۱۳.۲	نتایج شبیه‌سازی مربوط به ریج استوار، طرح ۲، $n = 100$ ، $\rho = 0/9$ و $\tau = 0/2$	۴۸
۱۴.۲	نتایج شبیه‌سازی مربوط به ریج استوار، طرح ۲، $n = 100$ ، $\rho = 0/99$ و $\tau = 0/2$	۴۹
۱۵.۲	نتایج شبیه‌سازی مربوط به ریج استوار، طرح ۲، $n = 300$ ، $\rho = 0/9$ و $\tau = 0/2$	۵۰
۱۶.۲	نتایج شبیه‌سازی مربوط به ریج استوار، طرح ۲، $n = 300$ ، $\rho = 0/99$ و $\tau = 0/2$	۵۱
۱۷.۲	نتایج مربوط به برآوردگر ریج استوار برای داده‌های Vo2	۵۳
۱۸.۲	داده‌های فازی شده سیمان پرتلند	۵۷
۱۹.۲	شاخص‌های ارزیابی مدل برای رگرسیون ریج فازی برای مجموعه داده	۵۹
۵۹	سیمان پرتلند	
۱.۳	نتایج شبیه‌سازی مربوط به لاسو استوار، طرح ۱، $n = 50$ ، $\rho = 0/3$ و $\tau = 0$	۶۸
۲.۳	نتایج شبیه‌سازی مربوط به لاسو استوار، طرح ۱، $n = 50$ ، $\rho = 0/99$ و $\tau = 0$	۶۸

۶۹	۳.۳	نتایج شبیه‌سازی مربوط به لاسو استوار، طرح ۱، $n = 200$ ، $\rho = 0/3$ و $\tau = 0$
	۴.۳	نتایج شبیه‌سازی مربوط به لاسو استوار، طرح ۱، $n = 200$ ، $\rho = 0/99$ و $\tau = 0$
۶۹		نتایج شبیه‌سازی مربوط به لاسو استوار، طرح ۱، $n = 200$ ، $\rho = 0/3$ و $\tau = 0/2$
۷۰	۵.۳	نتایج شبیه‌سازی مربوط به لاسو استوار، طرح ۱، $n = 50$ ، $\rho = 0/3$ و $\tau = 0/2$
	۶.۳	نتایج شبیه‌سازی مربوط به لاسو استوار، طرح ۱، $n = 50$ ، $\rho = 0/99$ و $\tau = 0/2$
۷۰		نتایج شبیه‌سازی مربوط به لاسو استوار، طرح ۱، $n = 200$ ، $\rho = 0/3$ و $\tau = 0/2$
۷۱	۷.۳	نتایج شبیه‌سازی مربوط به لاسو استوار، طرح ۱، $n = 200$ ، $\rho = 0/99$ و $\tau = 0/2$
	۸.۳	نتایج شبیه‌سازی مربوط به لاسو استوار، طرح ۱، $n = 200$ ، $\rho = 0/99$ و $\tau = 0/2$
۷۱		نتایج شبیه‌سازی مربوط به لاسو استوار، طرح ۲، $n = 100$ ، $\rho = 0/3$ و $\tau = 0$
۷۳	۹.۳	نتایج شبیه‌سازی مربوط به لاسو استوار، طرح ۲، $n = 100$ ، $\rho = 0/99$ و $\tau = 0$
	۱۰.۳	نتایج شبیه‌سازی مربوط به لاسو استوار، طرح ۲، $n = 100$ ، $\rho = 0/99$ و $\tau = 0$
۷۴		نتایج شبیه‌سازی مربوط به لاسو استوار، طرح ۲، $n = 300$ ، $\rho = 0/3$ و $\tau = 0$
۷۵	۱۱.۳	نتایج شبیه‌سازی مربوط به لاسو استوار، طرح ۲، $n = 300$ ، $\rho = 0/99$ و $\tau = 0$
	۱۲.۳	نتایج شبیه‌سازی مربوط به لاسو استوار، طرح ۲، $n = 100$ ، $\rho = 0/3$ و $\tau = 0/2$
۷۶		نتایج شبیه‌سازی مربوط به لاسو استوار، طرح ۲، $n = 100$ ، $\rho = 0/99$ و $\tau = 0/2$
۷۷	۱۳.۳	نتایج شبیه‌سازی مربوط به لاسو استوار، طرح ۲، $n = 100$ ، $\rho = 0/99$ و $\tau = 0/2$
	۱۴.۳	نتایج شبیه‌سازی مربوط به لاسو استوار، طرح ۲، $n = 300$ ، $\rho = 0/3$ و $\tau = 0/2$
۷۸		نتایج شبیه‌سازی مربوط به لاسو استوار، طرح ۲، $n = 300$ ، $\rho = 0/99$ و $\tau = 0/2$
۷۹	۱۵.۳	نتایج شبیه‌سازی مربوط به لاسو استوار، طرح ۲، $n = 300$ ، $\rho = 0/99$ و $\tau = 0/2$
	۱۶.۳	نتایج شبیه‌سازی مربوط به لاسو استوار، طرح ۲، $n = 300$ ، $\rho = 0/99$ و $\tau = 0/2$
۸۰		نتایج شبیه‌سازی مربوط به لاسو استوار، طرح ۲، $n = 300$ ، $\rho = 0/99$ و $\tau = 0/2$
۸۱	۱۷.۳	نتایج شبیه‌سازی مربوط به لاسو استوار، طرح ۲، $n = 300$ ، $\rho = 0/99$ و $\tau = 0/2$
	۱۸.۳	نتایج شبیه‌سازی مربوط به لاسو استوار، طرح ۲، $n = 300$ ، $\rho = 0/99$ و $\tau = 0/2$
۸۴		نتایج شبیه‌سازی مربوط به لاسو استوار، طرح ۲، $n = 300$ ، $\rho = 0/99$ و $\tau = 0/2$
۱۰۰	۱.۴	نتایج شبیه‌سازی مربوط به لاسو استوار، طرح ۱، $n = 50$ ، $\rho = 0/3$ و $\tau = 0$
۱۰۰	۲.۴	نتایج شبیه‌سازی مربوط به لاسو استوار، طرح ۱، $n = 50$ ، $\rho = 0/99$ و $\tau = 0$
۱۰۱	۳.۴	نتایج شبیه‌سازی مربوط به لاسو استوار، طرح ۱، $n = 200$ ، $\rho = 0/3$ و $\tau = 0$
۱۰۱	۴.۴	نتایج شبیه‌سازی مربوط به لاسو استوار، طرح ۱، $n = 200$ ، $\rho = 0/99$ و $\tau = 0$
۱۰۲	۵.۴	نتایج شبیه‌سازی مربوط به لاسو استوار، طرح ۱، $n = 50$ ، $\rho = 0/9$ و $\tau = 0/2$

۶.۴	نتایج شبیه‌سازی مربوط به بریج استوار، طرح ۱، $n = 50$ ، $\rho = 0.99$ و $\tau = 0.2$	۱۰۳
۷.۴	نتایج شبیه‌سازی مربوط به بریج استوار، طرح ۱، $n = 200$ ، $\rho = 0.9$ و $\tau = 0.2$	۱۰۳
۸.۴	نتایج شبیه‌سازی مربوط به بریج استوار، طرح ۱، $n = 200$ ، $\rho = 0.99$ و $\tau = 0.2$	۱۰۴
۹.۴	نتایج شبیه‌سازی مربوط به بریج استوار، طرح ۲، $n = 100$ ، $\rho = 0.3$ و $\tau = 0$	۱۰۵
۱۰.۴	نتایج شبیه‌سازی مربوط به بریج استوار، طرح ۲، $n = 100$ ، $\rho = 0.99$ و $\tau = 0$	۱۰۶
۱۱.۴	نتایج شبیه‌سازی مربوط به بریج استوار، طرح ۲، $n = 300$ ، $\rho = 0.3$ و $\tau = 0$	۱۰۷
۱۲.۴	نتایج شبیه‌سازی مربوط به بریج استوار، طرح ۲، $n = 300$ ، $\rho = 0.99$ و $\tau = 0$	۱۰۸
۱۳.۴	نتایج شبیه‌سازی مربوط به بریج استوار، طرح ۲، $n = 100$ ، $\rho = 0.9$ و $\tau = 0.2$	۱۰۹
۱۴.۴	نتایج شبیه‌سازی مربوط به بریج استوار، طرح ۲، $n = 100$ ، $\rho = 0.99$ و $\tau = 0.2$	۱۱۰
۱۵.۴	نتایج شبیه‌سازی مربوط به بریج استوار، طرح ۲، $n = 300$ ، $\rho = 0.9$ و $\tau = 0.2$	۱۱۱
۱۶.۴	نتایج شبیه‌سازی مربوط به بریج استوار، طرح ۲، $n = 300$ ، $\rho = 0.99$ و $\tau = 0.2$	۱۱۲
۱۷.۴	نتایج مربوط به برآوردگر بریج استوار برای داده‌های Vo22	۱۱۳

فصل ۱

تعاریف و مقدمات

۱.۱ مقدمه

تحلیل رگرسیونی یکی از روش‌های آماری برای مدل‌بندی داده‌های چند عاملی است و از دیدگاه مفهومی یک معادله است که ارتباط بین دسته‌ای از متغیرها را بیان می‌کند. موفقیت در تحلیل به درک و تیزبینی در مقوله نظری و مسائل علمی نیاز دارد و زمانی آشکار می‌شود که تکنیک در به‌کارگیری داده‌ها و اطلاعات جهان واقعی، قرار می‌گیرد. تحلیل رگرسیونی فن و تکنیکی آماری برای بررسی و مدل‌بندی ارتباط بین متغیرهاست.

کلمه رگرسیون در لغت به معنای «بازگشت» است، این نام‌گذاری براساس مطالعات گالتون [۳۲] در مورد رابطه قد پسر و متوسط قد والدین می‌باشد. گالتون نمودار قد پسر در مقابل متوسط قد والدین را برای حجم بزرگی از داده‌ها رسم کرد و به این نتیجه رسید که اغلب والدین بلند قد، پسرانی کوتاه قد و والدین کوتاه قد، پسرانی بلند قدتر از خود دارند و این پدیده را «بازگشت به میانگین» نامید، بر این اساس نام رگرسیون بر روش‌های بررسی روابط بین متغیرهای آماری ماندگار شد.

در این فصل، ضمن معرفی مدل رگرسیون خطی چندگانه و روش کمترین توان‌های دوم در برآورد پارامترهای رگرسیونی، مباحثی همچون برآوردگرهای جریمه‌شده، داده‌های ناهمگون و برآوردگرهای استوار را تعریف کرده و سپس به حالت کلی برآوردگرهای جریمه‌شده استوار می‌پردازیم و در آخر نیز به تعریف مفاهیمی از رگرسیون فازی خواهیم پرداخت. عمدتاً مطالب

این فصل از مارونا [۴۸]، ربیعی [۳]، آرست [۱] و فرخی [۶] آورده شده‌اند.

۲.۱ مدل رگرسیون چندگانه

در حالت کلی اگر متغیر پاسخ به p متغیر توضیحی $X_1, X_2, \dots, X_p$ وابسته باشد، مدل رگرسیونی به صورت زیر نوشته می‌شود

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p + \epsilon \quad (۱.۱)$$

که در آن X_j ، j امین متغیر توضیحی، Y متغیر پاسخ و ϵ مؤلفه خطای تصادفی مدل است. به دلیل وجود p متغیر در این مدل رگرسیونی، رگرسیون چندگانه نام دارد. پارامترهای β_j ، $j = 0, 1, \dots, p$ ضرایب رگرسیونی هستند. این مدل یک ابر صفحه در فضای p -بعدی از متغیرهای رگرسیونی X_j است و پارامتر β_j نشان‌دهنده تغییرات مورد انتظار متغیر پاسخ، زمانی که همه متغیرهای توضیحی باقی‌مانده دیگر $X_z (z \neq j)$ ثابت نگه‌داشته شوند، به ازای یک واحد تغییر در X_j است. به همین جهت پارامترهای β_j ، $(j = 1, 2, \dots, p)$ را ضرایب جزئی رگرسیون نیز می‌نامند. از آنجایی که ارتباط تابعی واقعی بین $X_1, X_2, \dots, X_p$ و Y شناخته شده نیست، مدل رگرسیونی خطی چندگانه روی دامنه تغییرات متغیرهای توضیحی، تقریب مناسبی از ارتباط تابعی ذکر شده را نشان می‌دهد.

با در نظر گرفتن بردارهای ستونی p -بعدی $\mathbf{x}_i = (x_{i1}, \dots, x_{ip})^\top$ و $\beta = (\beta_1, \dots, \beta_p)^\top$ که $i = 1, \dots, n$ ، مدل (۱.۱) را می‌توان به صورت زیر نوشت

$$y_i = \mathbf{x}_i^\top \beta + \epsilon_i, \quad i = 1, \dots, n \quad (۲.۱)$$

۳.۱ برآوردگر کمترین توان‌های دوم

فرض کنید یک نمونه تصادفی به حجم n ($n > p$) از مدل (۲.۱) به صورت $\{(y_i, \mathbf{x}_i)\}_{i=1}^n$ در دسترس است. در این صورت می‌توان مدل متناظر با (۱.۱) را با استفاده از نماد ماتریسی به صورت زیر نوشت

$$\mathbf{y} = \mathbf{X}\beta + \epsilon \quad (۳.۱)$$

که در آن $\mathbf{y} = (y_1, \dots, y_n)^\top$ یک بردار n مولفه‌ای شامل متغیر پاسخ و y_i متناظر با i امین پاسخ مشاهده شده برای مشاهده i ام، $\mathbf{X} = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n)^\top$ ماتریس طرح $n \times p$ ، $\mathbf{x}_i \in \mathbb{R}^p$ و x_{ij} متناظر با i امین مشاهده در j امین سطح متغیر توضیحی است. نیاز داریم که $\mathbf{X}$ دارای رتبه کامل ستونی باشد تا رتبه ماتریس $\mathbf{X}$ برابر با p شود (یعنی p ستون مستقل خطی داشته باشیم)، در این صورت ماتریس p -بعدی $\mathbf{X}^\top \mathbf{X}$ معکوس‌پذیر خواهد شد. همچنین فرض

می‌کنیم مشاهدات مرکزی شده‌اند، که در این صورت پارامتر عرض از مبدأ از مدل حذف شده (برای مشاهده جزئیات بخش ۱۰.۰ از رجبلو [۴] را ببینید)، بنابراین $\beta = (\beta_1, \beta_2, \dots, \beta_p)^\top$ برداری p -بعدی شامل ضرایب رگرسیونی و $\epsilon = (\epsilon_1, \epsilon_2, \dots, \epsilon_n)^\top$ بردار n -مولفه‌ای خطای تصادفی است که در آن $\mathbb{E}(\epsilon) = 0$ و $\mathbb{E}(\epsilon\epsilon^\top) = \sigma^2 \mathbf{I}_n$ (عبارت $\mathbb{E}(\cdot)$ به معنای امید ریاضی می‌باشد).

برآوردگر کمترین توان‌های دوم معمولی^۱ (OLS) برآوردگری است که مجموع توان‌های دوم مانده‌ها^۲ (RSS) در زیر را مینیمم می‌کند

$$\begin{aligned} \text{RSS}(\beta) &= \sum_{i=1}^n \epsilon_i^2 = \epsilon^\top \epsilon \\ &= (\mathbf{y} - \mathbf{X}\beta)^\top (\mathbf{y} - \mathbf{X}\beta) \end{aligned} \quad (4.1)$$

و نهایتاً برآوردگر ضرایب مدل به صورت زیر به دست می‌آید

$$\begin{aligned} \hat{\beta}_{\text{OLS}} &= \arg \min_{\beta} [(\mathbf{y} - \mathbf{X}\beta)^\top (\mathbf{y} - \mathbf{X}\beta)] \\ &= (\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top \mathbf{y} \end{aligned}$$

ارتباط مقادیر برازش شده $\hat{\mathbf{y}}$ با مقادیر مشاهده شده $\mathbf{y}$ به صورت زیر است

$$\hat{\mathbf{y}} = \mathbf{X}\hat{\beta} = \mathbf{H}\mathbf{y}$$

که در آن $\mathbf{H} = \mathbf{X}(\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top$ ماتریس هت^۳ نامیده می‌شود و همچنین بردار مقادیر مشاهده شده را به بردار مقادیر برازش شده تصویر می‌کند. ماتریس $\mathbf{H}$ ماتریسی $p \times p$ ، متقارن و خودتوان^۴ است. برای کسب اطلاعات بیش‌تر به آرست [۱] مراجعه کنید.

۴.۱ همخطی و برآوردگر جریمه‌شده

در مدل‌های رگرسیونی عدم همخطی چندگانه یا وابستگی خطی نزدیک بین متغیرهای رگرسیونی از اهمیت بسزایی برخوردار است. وجود این وابستگی خطی نزدیک، توانایی برآورد ضرایب مدل رگرسیونی را با مشکل مواجه می‌کند. از آنجایی که به کارگیری و تعبیر یک مدل رگرسیونی چندگانه به‌طور ضمنی به برآوردهای ضرایب مدل رگرسیونی بستگی دارد، برای به دست آوردن این ضرایب، این فرض که متغیرهای توضیحی وابستگی قوی با یکدیگر ندارند و مستقل خطی هستند، بسیار لازم و ضروری است. وقتی هیچ ارتباط خطی بین متغیرهای توضیحی وجود نداشته باشد، این متغیرها را متعامد گویند.

¹Ordinary least squares

²Residual sum of squares

³Hat matrix

⁴Idempotent

متعامد بودن متغیرهای توضیحی در اکثر کاربردهای رگرسیونی وجود ندارد و در بعضی وضعیت‌ها متغیرهای توضیحی تقریباً به‌طور ناقص ارتباط خطی دارند. اگر یکی از متغیرهای توضیحی تابع خطی دقیقی از یک یا چند متغیر توضیحی دیگر باشد، گوییم همخطی کامل^۵ اتفاق افتاده، همچنین همخطی ناقص زمانی اتفاق می‌افتد که یکی از متغیرهای توضیحی به‌طور تقریبی تابعی خطی از یک یا چند متغیر توضیحی دیگر باشد. در چنین مواردی نتایج رگرسیونی مبهم خواهند بود. برای اطلاع بیشتر از چگونگی ایجاد همخطی کامل و ناقص به روزبه [۵] مراجعه کنید.

بنابراین دانستن دلایل پیدایش همخطی و پی بردن به وجود همخطی در بین متغیرهای توضیحی، امری بسیار مهم و ضروری است و در زیر به بعضی از این روش‌ها اشاره می‌کنیم. چهار منبع اولیه پیدایش همخطی عبارتند از:

۱. شیوه جمع‌آوری داده‌ها

۲. قیدهای روی مدل یا درون جامعه

۳. نوع مدل

۴. مدلی با متغیرهای توضیحی بیش از حد

فهمیدن منبع همخطی چندگانه دارای اهمیت دوچندانی است، بدین صورت که توصیه برای تحلیل داده‌ها و تعبیر مدل نتیجه شده تا حدودی به علت مسئله همخطی بستگی دارد. برای اطلاعات بیش‌تر در مورد همخطی چندگانه به میسون و همکاران [۵۱] مراجعه کنید. همچنین از روش‌هایی که برای تشخیص همخطی چندگانه به کار می‌روند، می‌توان به موارد زیر اشاره کرد:

۱. محک ماتریس همبستگی^۶

۲. عامل تورم واریانس^۷ (VIF)

۳. عدد شرطی^۸ (CN)

که مختصراً به معرفی آن‌ها می‌پردازیم.

محک ماتریس همبستگی

مدل (۳.۱) را در نظر بگیرید. یک معیار بسیار ساده برای تشخیص وجود همخطی چندگانه، بررسی اعضای غیر قطر اصلی ماتریس $X^T X$ است. اگر متغیرهای توضیحی x_i و x_j تقریباً وابسته خطی باشند، در این صورت قدرمطلق این اعضا نزدیک یک خواهد بود.

⁵Exact collinearity

⁶Examination correlation matrix

⁷Variance inflation factor

⁸Conditional number

عامل تورم واریانس

برای آشکارسازی همخطی چندگانه، اعضای قطری ماتریس $S = (X^T X)^{-1}$ کاربرد چندانی ندارند. زامین عضو قطری ماتریس S (S_{jj}) را می‌توان به صورت $S_{jj} = (1 - R_j^2)^{-1}$ بازنویسی کرد که در آن R_j^2 ضریب تعیین مدل رگرسیون x_j بر روی $p - 1$ متغیر توضیحی باقی‌مانده دیگر است. اگر x_j بر متغیرهای توضیحی باقی‌مانده تقریباً عمود باشد، R_j^2 کوچک و S_{jj} به یک نزدیک می‌شود، درحالی‌که اگر مقداری وابستگی خطی به بعضی از متغیرهای رگرسیون داشته باشد، R_j^2 نزدیک به یک 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 دلیل وابستگی نزدیک بین متغیرهای توضیحی، افزایش یافته است. معیار

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

را عامل تورم واریانس می‌نامند. این معیار اثر مرکب از ارتباط بین متغیرهای توضیحی روی واریانس آن جمله را برای هر جمله مدل اندازه‌گیری می‌کند. مقدار بزرگ برای VIF به معنای همخطی بین متغیرهاست. بر اساس تحقیقات پیشین اگر هریک از VIFها از ۵ یا ۱۰ تجاوز کند، نشان‌دهنده این است که ضرایب رگرسیونی به دست آمده به علت وجود همخطی چندگانه، به صورت ضعیفی برآورد شده‌اند. برای اطلاعات بیشتر به مارکوارت [۵۰] مراجعه فرمایید.

عدد شرطی

مقادیر ویژه $X^T X$ می‌توانند برای اندازه‌گیری وسعت همخطی چندگانه در داده‌ها مورد استفاده قرار گیرند. عدد شرطی معیاری برای کشف وجود همخطی چندگانه در ماتریس مشاهدات است که با استفاده از مقادیر ویژه $X^T X$ محاسبه می‌شود. عدد شرطی ماتریس $X^T X$ به صورت زیر تعریف می‌شود

$$CN = \frac{\lambda_{max}}{\lambda_{min}}$$

که در آن $\lambda_{max} = \max(\lambda_i)$ و $\lambda_{min} = \min(\lambda_i)$ و λ_i مقدار ویژه i ام $X^T X$ است. عدد شرطی معیاری برای اندازه‌گیری پراکندگی مقادیر ویژه $X^T X$ است. اگر عدد شرطی کمتر از ۱۰۰ باشد، همخطی چندگانه مشکل‌ساز نخواهد بود، اعداد شرطی بین ۱۰۰ و ۱۰۰۰ نشان‌دهنده همخطی چندگانه شدید و اگر از ۱۰۰۰ تجاوز کند حاکی از همخطی چندگانه بسیار شدید است.

لازم به ذکر است در بعضی مقالات، جذر رابطه بالا را در نظر می‌گیرند. در این حالت اگر جذر CN بزرگتر از ۱۰ باشد، آن‌گاه در ماتریس طرح همخطی وجود دارد، اگر این عدد شرطی بین ۳۰ تا ۱۰۰ باشد همخطی شدید و اگر بالاتر از ۱۰۰ باشد، همخطی بسیار جدی و مشکل‌ساز است (بلسلی و همکاران [۱۷]).

روش‌های برخورد با همخطی چندگانه

روش‌های متعددی نیز برای برخورد با مشکلات ناشی از همخطی چندگانه پیشنهاد شده‌اند. روش‌هایی همانند جمع‌آوری داده‌های اضافی و به‌کارگیری روش‌های برآورد غیر از کمترین توان‌های دوم و روش‌هایی که به‌طور ویژه برای مقابله با مشکلات ناشی از وجود همخطی چندگانه در بین متغیرهای توضیحی طراحی شده‌اند.

یکی از رهیافت‌های مقابله با همخطی، روش‌های انقباضی^۹ است. یک روش معمول دستیابی به برآوردگرهای انقباضی، استفاده از رگرسیون کمترین توان‌های دوم جریمه‌شده^{۱۰} (PLS) می‌باشد. به عبارت دیگر حاصل جواب رگرسیون کمترین توان‌های دوم جریمه‌شده یک برآوردگر انقباضی است. تحت فرض‌های مدل (۳.۱) در مسئله رگرسیون PLS به دنبال دستیابی به برآوردگر $\hat{\beta}^S$ هستیم، که جواب مسئله بهینه‌سازی زیر می‌باشد

$$\hat{\beta}^S = \arg \min_{\beta \in \mathbb{R}^p} \left(\text{RSS}(\beta) + k \sum_j P(|\beta_j|) \right) \quad (5.1)$$

که در آن $\text{RSS}(\beta)$ مجموع توان‌های دوم باقی‌مانده‌ها و $\sum_j P(|\beta_j|)$ عبارت جریمه و k پارامتر کنترل‌کننده می‌باشد. از انواع برآوردگرهای انقباضی می‌توان به برآوردگر انقباضی ریج^{۱۱}، که حالت خاصی از برآوردگر انقباضی ریج^{۱۲} است، اشاره کرد، که در سال ۱۹۷۰ توسط هورل و کنارد [۳۶] معرفی شد. تیبشیرانی [۶۰] نیز با استفاده از عملگر علامت، برآوردگر انقباضی لاسو^{۱۳} (LASSO) را برای اولین بار ارائه کرد. در ادامه این بخش به معرفی این برآوردگرهای انقباضی خواهیم پرداخت.

۱.۴.۱ برآوردگر بریج

فرانک و فریدمن [۳۰] مجموع توان‌های دوم باقی‌مانده در رابطه (۴.۱) را با استفاده از قید کلی $\sum_{j=1}^p |\beta_j|^\gamma \leq t$ ، به ازای $\gamma \geq 0$ در نظر گرفتند و آن را به عنوان رگرسیون بریج ارائه نمودند. بنابراین طبق آنچه در (۵.۱) گفته شد، عبارت جریمه را به صورت $\sum_{j=1}^p P(|\beta_j|) = \sum_{j=1}^p |\beta_j|^\gamma$ در نظر گرفته و داریم

$$\hat{\beta}_{\text{Bridge}} = \arg \min_{\beta \in \mathbb{R}^p} \left(\text{RSS}(\beta) + k \sum_{j=1}^p |\beta_j|^\gamma \right) \quad (6.1)$$

که در آن $k, \gamma, t \geq 0$ پارامترهای تنظیم‌کننده^{۱۴} بوده و میزان انقباض تحمیل‌شده به ضرایب را کنترل می‌کنند. انتخاب بهینه‌ی γ منجر به افزایش کارایی برآوردگر خواهد شد. اگر $\gamma = 2$ ،

^۹Shrinkage method

^{۱۰}Penalized least square

^{۱۱}Ridge

^{۱۲}Bridge

^{۱۳}Least absolute shrinkage and selection operator

^{۱۴}Tunning parameters

مسئله فوق تبدیل به رگرسیون ريج (بخش ۲.۴.۱) و اگر $\gamma = 1$ آن گاه به رگرسیون لاسو (بخش ۳.۴.۱) تبدیل می شود. با توجه به ساختار برآوردگر ريج، مشخص می شود به ازای $\gamma \in (1, 2)$ این برآوردگر بین برآوردگرهای لاسو و ريج رفتار می کند و به نوعی بین این دو برآوردگر پل برقرار می کند، از این رو نام آن را برآوردگر ريج نهادند. برآوردگر ريج در حالت کلی (به ازای $\gamma \geq 0$) دارای صورت بسته ای نیست و در واقع فرانک و فریدمن [۳۰] فقط آن را پیشنهاد دادند. در نهایت فو [۳۱] این مساله را برای $\gamma \geq 1$ حل نمود و الگوریتم نیوتون-رافسون^{۱۵} را برای برآوردگر ريج زمانی که $\gamma > 1$ است، ارایه کرد و همچنین الگوریتمی پیش برنده^{۱۶} برای لاسو پیشنهاد داد.

محاسبه برآوردگر ريج. برای محاسبه برآوردگر انقباضی ريج زمانی که $\gamma \geq 1$ است، باید مسئله بهینه سازی زیر را حل کنیم

$$\hat{\beta}_{\text{Bridge}} = \arg \min_{\beta \in \mathbb{R}^p} \left((\mathbf{Y} - \mathbf{X}\beta)^\top (\mathbf{Y} - \mathbf{X}\beta) + k \sum_{j=1}^p |\beta_j|^\gamma \right), \quad k \geq 0 \quad (7.1)$$

بنابراین برای یافتن مینیمم از معادله (۷.۱) نسبت به β مشتق گرفته و برابر با صفر قرار می دهیم. به این دلیل که صورت بسته ای برای معادله (۷.۱) موجود نیست و از آنجایی که برای $1 \leq \gamma < 2$ تابع $\sum_{j=1}^p |\beta_j|^\gamma$ در $\beta_j = 0$ مشتق پذیر نیست، مشتقات جزئی (۷.۱) را نسبت به $\beta_j \neq 0$ برای $j = 1, \dots, p$ محاسبه می نماییم. برای اطلاعات بیشتر به آرست [۱] و فو [۳۱] مراجعه کنید.

واریانس برآوردگر ريج. فو [۳۱] واریانس برآوردگر ريج را در قالب قضیه زیر ارائه داد

قضیه ۱۰.۴.۱ (فو [۳۱]). تحت مفروضات مدل (۱.۱) و معادله (۶.۱)، واریانس برآوردگر ريج برای $\gamma > 1$ به صورت زیر است

$$\text{Var}(\hat{\beta}_{\text{Bridge}}) = \sigma^2 (\mathbf{X}^\top \mathbf{X} + kD^*(\hat{\beta})|_{\mathbf{y}_0})^{-1} \mathbf{X}^\top \mathbf{X} (\mathbf{X}^\top \mathbf{X} + kD^*(\hat{\beta})|_{\mathbf{y}_0})^{-1}$$

که در آن $D^*(\hat{\beta}) = \text{diag}(\frac{2(\gamma-1)}{\gamma} |\hat{\beta}_j|^{\gamma-2})$ و $\mathbf{y}_0$ نقطه ای ثابت در فضای نمونه است.

برهان. برای اثبات به آرست [۲] و فو [۳۱] مراجعه کنید. □

پارامتر انقباضی γ و پارامتر کنترل k . بنابر استدلال فرانک و فریدمن [۳۰]، انتخاب بهینه پارامتر انقباضی γ باعث افزایش دقت در پیشگویی می شود. بنابراین برای انتخاب بهینه پارامترهای انقباضی γ و تنظیم کننده k از آماره اعتبارسنجی متقابل تعمیم یافته^{۱۷} (GCV)

¹⁵Newton-Raphson algorithm

¹⁶Shooting algorithm

¹⁷Generalized cross-validation

استفاده می‌کنیم، که توسط کراون و وهبا [۲۲] و تیبشیرانی [۶۰] ارائه شد. برای $k \geq 0$ و $\gamma \geq 1$ ، این معیار به صورت زیر تعریف می‌شود

$$GCV = \frac{1}{n} \sum_{i=1}^n \left(\frac{e_{i,k}}{1 - \frac{p(k)}{n}} \right)^2 \quad (۸.۱)$$

که در آن به ازای $\gamma = 1$

$$p(k) = \text{tr}(\mathbf{X}(\mathbf{X}^\top \mathbf{X} + k\Omega^-)^{-1} \mathbf{X}^\top) - n_0$$

و برای سایر مقادیر γ

$$p(k) = \text{tr}(\mathbf{X}(\mathbf{X}^\top \mathbf{X} + k\Omega^{-1})^{-1} \mathbf{X}^\top)$$

که n_0 تعداد $\hat{\beta}_j$ های برابر با صفر و Ω^- معکوس تعمیم‌یافته ماتریس Ω است. علاوه بر این $e_{i,k} = y_i - \hat{y}_i^{\text{Bridge}}$ و $\Omega = \text{diag}(\frac{\gamma}{\gamma-2} |\hat{\beta}_j|^{\gamma-2})$ پارامترهای γ و k چنان انتخاب می‌شوند که آماره GCV کمترین مقدار خود را اختیار کند.

۲.۴.۱ برآوردگر ریج

یکی از موثرترین روش‌های مقابله با مشکل همخطی، رگرسیون ریج است که توسط هورل و کنارد [۳۶] معرفی شد. اگر در رابطه (۵.۱)، RSS را نسبت به قید $\sum_{j=1}^p \beta_j^2 \leq t$ مینیمم کنیم، برآوردگر ریج به دست می‌آید. برآوردگر ریج به پارامتر کنترل‌کننده k بستگی دارد. افزودن این پارامتر به ماتریس $\mathbf{X}^\top \mathbf{X}$ به صورت زیر باعث بهبود وضعیت آن می‌شود

$$\begin{aligned} \hat{\beta}_{\text{Ridge}} &= \arg \min_{\beta} \left(\text{RSS} + k \sum_{j=1}^p \beta_j^2 \right) \\ &= (\mathbf{X}^\top \mathbf{X} + k\mathbf{I}_p)^{-1} \mathbf{X}^\top \mathbf{Y} \\ &= (k(\mathbf{X}^\top \mathbf{X})^{-1} + \mathbf{I}_p)^{-1} \hat{\beta}_{\text{OLS}} \\ &= T(k) \hat{\beta}_{\text{OLS}} \end{aligned} \quad (۹.۱)$$

که در آن $T(k) = (k(\mathbf{X}^\top \mathbf{X})^{-1} + \mathbf{I}_p)^{-1}$ و $k, t \geq 0$ پارامترهای تنظیم‌کننده بوده و میزان انقباض ضرایب را کنترل می‌کنند. برآوردگر ریج یک تبدیل خطی از برآوردگر کمترین توان‌های دوم است. امید ریاضی برآوردگر $\hat{\beta}_{\text{Ridge}}$ عبارت است از

$$\mathbb{E}(\hat{\beta}_{\text{Ridge}}) = \mathbb{E}((\mathbf{X}^\top \mathbf{X} + k\mathbf{I}_p)^{-1} \mathbf{X}^\top \mathbf{Y}) = (\mathbf{X}^\top \mathbf{X} + k\mathbf{I}_p)^{-1} \mathbf{X}^\top \mathbf{X} \beta$$

بنابراین می‌توان گفت برآوردگر $\hat{\beta}_{\text{Ridge}}$ برآوردگری اریب برای β است. همچنین ماتریس کوواریانس برآوردگر ریج به صورت زیر محاسبه می‌شود

$$\text{Cov}(\hat{\beta}_{\text{Ridge}}) = \sigma^2 (\mathbf{X}^\top \mathbf{X} + k\mathbf{I})^{-1} \mathbf{X}^\top \mathbf{X} (\mathbf{X}^\top \mathbf{X} + k\mathbf{I})^{-1}$$

در بخش بعدی این فصل به معرفی برآوردگر لاسو خواهیم پرداخت که حالت خاص دیگری از برآوردگر بریج است.

۳.۴.۱ برآوردگر لاسو

از دیدگاه آماری، لاسو مخفف عملگر انتخاب‌کننده، منقبض‌کننده و مینیمم‌کننده از جنس قدرمطلق است. برآوردگر لاسو نوع جریمه‌شده برآوردگر OLS است و از حل مسئله بهینه‌سازی زیر به دست می‌آید

$$\hat{\beta}_{\text{LASSO}} = \min_{\beta \in \mathbb{R}^p} \left(\text{RSS} \right) \quad \text{s.t.} \quad \sum_{j=1}^p |\beta_j| \leq t \quad (۱۰.۱)$$

که در آن $t > 0$ یک مقدار ثابت است. اگر $t = 0$ ، آن‌گاه متغیری در مدل وجود نخواهد داشت، درحالی‌که اگر $t = \infty$ ، مدل کامل خواهد بود. فرض کنید $\hat{\beta}^* = (\hat{\beta}_1^*, \dots, \hat{\beta}_p^*)^\top$ یک برآوردگر اولیه برای β است که معمولاً این برآوردگر را $\hat{\beta}_{\text{OLS}}$ در نظر می‌گیرند. اگر $t > \sum_{j=1}^p |\hat{\beta}_j^*|$ ، آن‌گاه روش لاسو برآوردگر OLS را نتیجه می‌دهد. با انتخاب $0 < t < \sum_{j=1}^p |\hat{\beta}_j^{\text{OLS}}|$ ، مسئله بهینه‌سازی (۱۰.۱) معادل است با

$$\hat{\beta}_{\text{LASSO}} = \arg \min_{\beta \in \mathbb{R}^p} \left(\text{RSS} + k \sum_{j=1}^p |\beta_j| \right)$$

که در آن k پارامتر تنظیم‌کننده است و سطح تنکی (تعداد پارامترهای صفر) را در برآوردگر لاسو کنترل می‌کند. وقتی $\lambda \rightarrow 0$ برآوردگر لاسو به برآوردگر OLS میل می‌کند و وقتی $\lambda \rightarrow \infty$ ، برآوردگر لاسو به سمت صفر میل می‌کند. به عبارت دیگر، هرچه جریمه بزرگتری به کار برده شود، تعداد بیشتری از ضرایب به سمت صفر منقبض می‌شوند. در حقیقت λ و t رفتار یا رابطه‌ای معکوس با یکدیگر دارند. از این‌رو برآوردگر لاسو می‌تواند متغیرهای اضافی را از مدل حذف نماید. در واقع برآوردگر لاسو، هم‌زمان هم انتخاب متغیر انجام می‌دهد و هم ضرایب را منقبض می‌کند. همچنین به دلیل تابع جریمه قدرمطلق، جواب صریحی برای برآوردگر لاسو وجود ندارد. تیبشیرانی [۶۰] این برآوردگر را از طریق برنامه‌ریزی درجه دوم به دست آورده بود.

۵.۱ روش‌های استوار

داده‌های جمع‌آوری‌شده، علاوه بر مشکل همخطی بین متغیرها که در بخش قبلی به آن پرداخته شد، ممکن است شامل یک یا چند مشاهده ناهمگون^{۱۸} باشند، یعنی مشاهداتی که به‌خوبی از بیشتر یا اکثر^{۱۹} داده‌ها جدا شده‌اند یا بعضاً از طرح کلی داده‌ها خارج شده‌اند. برآوردگرهای کلاسیک همانند میانگین، واریانس، کوواریانس و همبستگی نمونه‌ای یا برازش کمترین توان‌های دوم برای یک مدل رگرسیونی، می‌توانند به‌وسیله مشاهدات ناهمگون به شدت

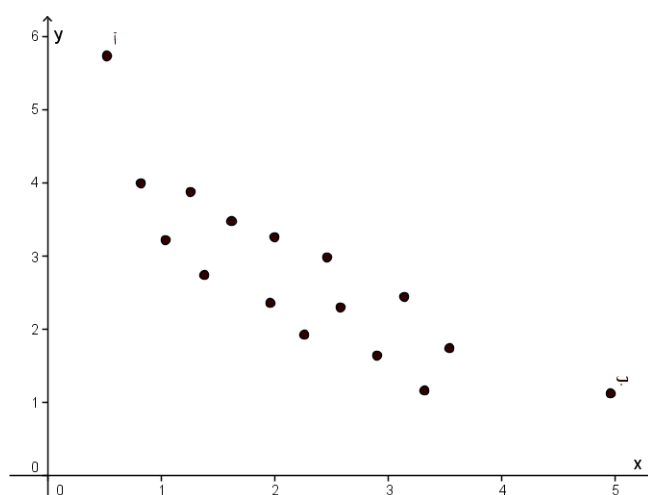
^{۱۸}Atypical

^{۱۹}Bulk

تحت تاثیر قرار گیرند، حتی اگر یک مشاهده ناهمگون موجود باشد. در این موارد اغلب قادر به ارائه برازش‌های خوبی به حجم داده نخواهیم بود. برای حل این مشکل، برآوردگرهای پارامتر استواری^{۲۰} وجود دارند که در حضور مشاهدات ناهمگون در داده برازشی به خوبی زمانی که مشاهده ناهمگون در داده وجود ندارد، برای حجم داده ارائه می‌کنند.

۱.۵.۱ بررسی مشاهدات ناهمگون

با توجه به اینکه همه نقاط داده در تعیین معادله خط وزن یکسانی می‌گیرند، وضعیت قرارگیری داده‌ها در برازش کمترین توان‌های دوم نقش مهمی را ایفا می‌کند و شیب این خط شدیداً تحت تاثیر داده‌های ناهمگون است. برای مثال شکل ۱.۱ را در نظر بگیرید، شیب خط در برازش کمترین توان‌های دوم تا حد زیادی به یکی از دو نقطه (آ) و (ب) وابسته است. همچنین حذف این دو داده از مجموعه داده‌ها برآورد کاملاً متفاوتی را از شیب به دست می‌دهد.



شکل ۱.۱: نمایش مشاهدات ناهمگون (آ)، در راستای محور y ها و (ب)، در راستای محور x ها در یک مثال ساختگی

در چنین مواردی نیاز به انجام روش‌هایی اصلاحی برای برآورد پارامترها داریم، از قبیل تحلیل بیشتر یا در صورت امکان حذف مشاهدات ناهمگون یا برآورد پارامترها با استفاده از روش‌هایی که در قیاس با روش کمترین توان‌های دوم از استواری بیشتری در حضور مشاهدات ناهمگون در مجموعه داده‌ها برخوردارند. همان‌طور که از شکل ۱.۱ مشاهده می‌شود، مشاهدات ناهمگون به داده‌های دورافتاده^{۲۱} و نقاط نافذ^{۲۲} تقسیم می‌شوند.

²⁰Robust

²¹Outliers

²²Leverage Points

داده‌های دورافتاده

یک داده دورافتاده در بین مانده‌ها، مانده‌ای است که از نظر قدرمطلق به‌طور قابل توجهی بزرگتر از بقیه باشد. حتی ممکن است به فاصله بیشتر از سه یا چهار برابر انحراف معیار میانگین داده‌ها قرار گیرد. یک داده دورافتاده جدا از بقیه داده‌ها نیست و باید تحت بررسی دقیق قرار گیرد تا علت وجودی آن معلوم شود.

قواعدی برای حذف داده‌های دورافتاده پیشنهاد شده‌اند. حذف داده‌های دورافتاده همیشه روش معقولی نیست، گاهی این داده‌ها حاوی اطلاعاتی هستند که نقاط دیگر قادر به بیان آن نیستند، زیرا ناشی از عوامل غیرمعمول بوده و حذف آن‌ها منجر به بررسی مجدد خواهد شد. به عنوان یک قاعده کلی، داده‌های دورافتاده را تنها زمانی می‌توان کنار گذاشت که احتمالاً در نتیجه خطا در ثبت مشاهدات یا تنظیم وسایل، ایجاد شده‌اند در غیر این صورت باید آن‌ها را به دقت بررسی کرد.

نقطه نافذ

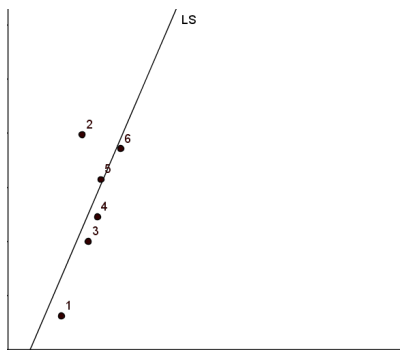
داده‌ای که در راستای یکی از متغیرهای تبیینی (توضیحی) از تفاوت قابل ملاحظه‌ای نسبت به دیگر نقاط داده برخوردار است را یک نقطه نافذ می‌نامند. نقاط نافذ در برآوردهای ضرایب رگرسیونی اثر نامناسبی دارند، زیرا این نقاط از بقیه داده‌ها طبعیت نمی‌کنند و اثر نامطلوبی در مدل می‌گذارند.

در شکل ۲.۱ در تصاویر (آ) و (ب) اثرات داده ناهمگون (مشاهده ۲) بر روی خط رگرسیونی روش کمترین توان‌های دوم نمایش داده شده است. همچنین در تصاویر (ج) و (د) اثرات نقطه نافذ خوب و نقطه نافذ بد بر روی خط رگرسیونی کمترین توان‌های دوم نمایش داده شده است. مطابق با تصویر (ج) نقطه نافذی که بر خط برازش شده تاثیر چندانی ندارد را نقطه نافذ خوب^{۲۳} و همچنین مطابق با تصویر (د)، نقطه نافذی که در روند برآورد ضرایب تاثیر به‌سزایی دارد را نقطه نافذ بد می‌نامیم.

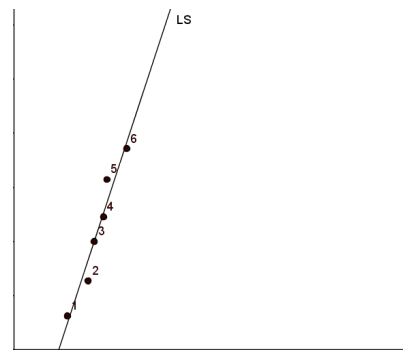
مشاهده‌ای که یک داده ناهمگون بزرگ تولید می‌کند به واقع خوب نیست، ولی لزوماً بدین معنی نیست که این مشاهده در برازش مدل انتخابی مؤثر باشد. در هر مجموعه‌ای از داده‌ها که در آن برآورد یک یا چند پارامتر قویا به تعداد اندکی از مشاهدات وابسته است، حضور نقاط نافذ مشکلاتی جدی پدید می‌آورد. یک راه فائق آمدن بر این مشکلات این است که ببینیم آیا حذف یک یا دو مشاهده ناهمگون در برازش مدل و نتایج حاصل از آن اثر چندانی دارد یا خیر. اگر اثر زیادی دارد، نتایج ضعیف‌اند و به داده‌های بیشتری احتیاج است. اما روش کارآمدتر این است که از روش‌های برآورد دیگری استفاده کنیم که نسبت به این داده‌ها استوار هستند. از معیارهای مهم اندازه‌گیری استواری می‌توان به نقطه فروریزش^{۲۴} اشاره کرد که تعریف آن

²³Good leverage point

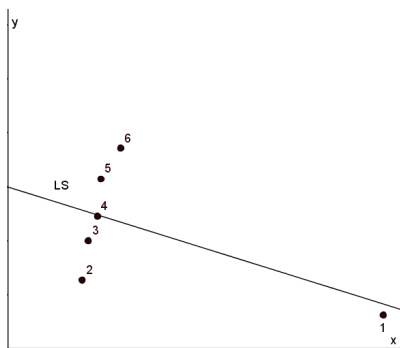
²⁴Breakdown point



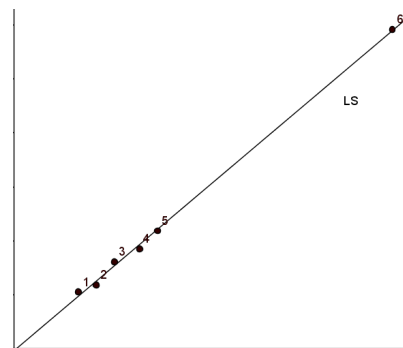
(ب) داده ناهمگون در راستای محور y



(آ) داده‌ها با رابطه خطی قوی بایکدیگر



(د) داده‌ای با داده ناهمگون از نوع متفاوت در راستای محور x جایگزین می‌شود.



(ج) یک داده با یک داده ناهمگون در راستای محور x جایگزین می‌شود.

شکل ۲.۱: خطوط رگرسیونی کمترین توان‌های دوم در حضور انواع داده ناهمگون

در زیر می‌آید.

تعریف ۱.۵.۱ (نقطه فروریزش (راسو و لروی [۵۵]). یک نمونه تصادفی از n نقطه داده را به صورت زیر در نظر می‌گیریم

$$Z = \{(x_{11}, \dots, x_{1p}, y_1), \dots, (x_{n1}, \dots, x_{np}, y_n)\}$$

همچنین فرض کنید T یک برآوردگر رگرسیونی باشد. به این معنی که به کار بردن T برای نمونه Z یک بردار از ضرایب رگرسیونی به دست می‌دهد. اکنون فرض کنید همه نمونه‌های آلوده Z' که به وسیله جایگزین کردن هر m تا از نقاط داده اصلی با نقاط دلخواه، به دست می‌آید (این اجازه می‌دهد داده‌های ناهمگون خیلی بد در مجموعه داده‌ها داشته باشیم). فرض کنیم بیش‌ترین آریبی که می‌تواند به وسیله چنین آلودگی ایجاد شود را $\text{bias}(m; T, Z)$ تعریف کنیم، در این صورت

$$\text{bias}(m; T, Z) = \sup_{Z'} \|T(Z') - T(Z)\| \quad (11.1)$$

که در آن سوپریمم روی همه مجموعه‌های ممکن Z' است. اگر $\text{bias}(m; T, Z)$ نامتناهی باشد، به این معنی که m داده ناهمگون می‌تواند تاثیر زیادی روی T داشته باشد، به طوری که بتوان گفت برآوردگر به نوعی فروریخت. از این رو، نقطه فروریزش برآوردگر T برای نمونه (متناهی) Z به صورت زیر تعریف می‌شود

$$\varepsilon_n^*(T, Z) = \min \left\{ \frac{m}{n}; \text{bias}(m; T, Z) \text{ نامتناهی است} \right\} \quad (12.1)$$

به عبارت دیگر، کمترین کسری از آلودگی است که می‌تواند باعث شود برآوردگر T مقادیری با فاصله زیاد از $T(Z)$ بگیرد. توجه کنید که این تعریف شامل هیچ توزیع احتمالی نمی‌شود. در روش OLS حتی یک مشاهده ناهمگون هم می‌تواند سبب فروریزش شود، لذا نقطه فروریزش در روش OLS برابر است با

$$\varepsilon^*(T, L) = \frac{1}{n}$$

که در نمونه‌های با حجم زیاد برابر صفر است، به عبارت دیگر روش OLS دارای نقطه فروریزش ۰٪ است، یعنی در مقابل حضور حتی یک داده ناهمگون در مجموعه داده بسیار حساس است و نتایج قابل اعتمادی ارائه نمی‌دهد.

۶.۱ برآوردگرهای استوار در مدل خطی

یک راه ساده برای کنترل داده‌های ناهمگون، شناسایی و حذف آن‌هاست. راه‌های زیادی برای شناسایی داده‌های ناهمگون وجود دارند، اما حذف آن‌ها اگرچه از انجام ندادن هیچ کاری بهتر است، در عین حال مشکلات زیادی را به وجود خواهد آورد. چندی از این مشکلات به صورت پرسش‌های زیر مطرح شده‌اند.

- چه موقع حذف کردن موجه است؟ حذف کردن به تصمیمات فردی نیاز دارد، چه موقع یک مشاهده "به حد کافی دورافتاده" حذف شود؟
- کاربر یا محقق ممکن است به این فکر کند که «یک مشاهده، یک مشاهده است» و احساس خوش‌آیندی درمورد حذف کردن آن‌ها نداشته باشد.
- از آن‌جا که در حالت کلی مقداری عدم اطمینان در مورد ناهمگونی یک مشاهده وجود دارد، خطای حذف کردن مشاهدات خوب به‌وجود می‌آید که می‌تواند نتیجه دست‌کم گرفتن تغییرپذیری داده‌ها باشد.

بنابر سوالات مطرح شده، حذف داده‌های ناهمگون خیلی منطقی به‌نظر نمی‌رسد. بنابراین به‌دنبال راه‌هایی برای مقابله با داده‌های ناهمگون هستیم. در ادامه به معرفی برخی از این روش‌های استوار نسبت به حضور داده‌های ناهمگون در مجموعه داده خواهیم پرداخت.

۱.۶.۱ M – برآوردگرها

این برآوردگرها از نوع برآوردگرهای ماکسیمم درست‌نمایی^{۲۵} (MLE) هستند که هم‌زمان کارایی و استواری نسبت به حضور داده ناهمگون در مجموعه داده را افزایش خواهند داد. مدل (۲.۱) را درنظر بگیرید، که در آن ماتریس X معین و ϵ_i ها دارای چگالی زیر باشند

$$\frac{1}{\sigma} f_{\circ} \left(\frac{\epsilon}{\sigma} \right)$$

که در آن σ پارامتر مقیاس است. برای مدل خطی (۲.۱) y_i ها مستقل هستند ولی هم‌توزیع نیستند و دارای چگالی زیر می‌باشند

$$\frac{1}{\sigma} f_{\circ} \left(\frac{y - \mathbf{x}_i^T \beta}{\sigma} \right)$$

و تابع درست‌نمایی برای β با فرض معلوم بودن مقدار σ به‌صورت زیر است

$$L(\beta) = \frac{1}{\sigma^n} \prod_{i=1}^n f_{\circ} \left(\frac{y_i - \mathbf{x}_i^T \beta}{\sigma} \right)$$

برآورد ماکسیمم درست‌نمایی به‌وسیله ماکسیمم کردن $L(\beta)$ معادل یافتن $\hat{\beta}$ است که به ازای $\beta = \hat{\beta}$ تابع

$$l_n(\beta) = \frac{1}{n} \sum_{i=1}^n \rho_{\circ} \left(\frac{r_i(\beta)}{\sigma} \right) + \log \sigma \quad (۱۳.۱)$$

مینیمم شود، که در آن $\rho = -\log f_{\circ}$ و $r_i(\beta) = y_i - \mathbf{x}_i^T \beta$. اکنون با فرض معلوم بودن σ و مشتق‌گیری نسبت به β برای به‌دست آوردن مینیمم، معادلات نرمال زیر را داریم

$$\sum_{i=1}^n \psi_{\circ} \left(\frac{r_i(\hat{\beta})}{\sigma} \right) \mathbf{x}_i = 0 \quad (۱۴.۱)$$

²⁵Maximum likelihood estimators

که در آن $\psi_0 = \rho'_0 = -f'_0/f_0$. اگر f_0 چگالی نرمال استاندارد باشد، آن گاه $\hat{\beta}$ برآوردگر کمترین توان‌های دوم خواهد بود و اگر f_0 چگالی نمایی دوگانه باشد آن گاه رابطه (۱۳.۱) به صورت زیر به دست می‌آید

$$l_n(\beta) = \sum_{i=1}^n |r_i(\beta)| \quad (15.1)$$

و برآورد β به دست آمده از آن، $\hat{\beta}$ ، را برآورد L_1 می‌نامیم که رگرسیون هم‌ارز با میانه است. قابل توجه است که این برآوردگر قبل از برآوردگر کمترین توان‌های دوم مورد مطالعه قرار گرفته است (توسط بوسکوویچ [۱۹]). با مشتق‌گیری از رابطه (۱۵.۱)، معادلات نرمال به صورت زیر است

$$\sum_{i=1}^n \text{sgn}(r_i(\hat{\beta})) \mathbf{x}_i = \mathbf{0} \quad (16.1)$$

که در آن "sgn" تابع علامت است. اگر مدل شامل عبارت عرض از مبدا نیز باشد، باقی‌مانده‌ها دارای میانه صفر خواهند بود. توجه داشته باشید که برخلاف برآورد LS، برآورد L_1 در کل دارای صورت بسته‌ای نیست اما الگوریتم‌های سریعی برای محاسبه آن وجود دارند (بارودال و رابرت [۱۵] و پورتنوی و کوینکر [۵۴]). بنابراین ممکن است $\hat{\beta}$ به دست آمده یکتا نباشد و دارای این ویژگی است که حداقل p باقی‌مانده صفر هستند (بلومفیلد و استیگر [۱۸]). رگرسیون M – برآوردگر را به عنوان جواب $\hat{\beta}$ وقتی معادله زیر مینیمم شود، در نظر می‌گیریم

$$l_n(\beta) = \sum_{i=1}^n \rho\left(\frac{r_i(\beta)}{\hat{\sigma}}\right) \quad (17.1)$$

که در آن $\hat{\sigma}$ برآورد مقیاسی خطاهاست. با مشتق‌گیری از عبارت (۱۷.۱)، معادلات نرمال زیر را داریم

$$\sum_{i=1}^n \psi\left(\frac{r_i(\hat{\beta})}{\hat{\sigma}}\right) \mathbf{x}_i = \mathbf{0} \quad (18.1)$$

که در آن $\psi = \rho'$. معادله (۱۸.۱) لزوماً معادله برآوردکننده یک MLE نیست. در این پایان‌نامه فرض می‌شود که $\hat{\sigma}$ از قبل به دست آمده باشد. همچنین از این به بعد فرض خواهد شد که ρ و ψ به ترتیب ρ – تابع و ψ – تابع باشند که در زیر به آن‌ها اشاره می‌شود. در این بخش نیز ماتریس $\mathbf{X}$ پرتبه فرض می‌شود.

تعریف ۱.۶.۱ (ρ – تابع). یک ρ – تابع توسط یک تابع ρ تعریف می‌شود به طوری که

R1 $\rho(x)$ تابعی نانزولی از $|x|$ است.

R2 $\rho(0) = 0$.

R3 $\rho(x)$ برای $x > \circ$ صعودی است، به طوری که $\rho(x) < \rho(\infty)$.

R4 اگر ρ کراندار باشد، آن گاه $\rho(\infty) = 1$.

تعریف ۲.۶.۱ $(\psi - \text{تابع})$. یک $\psi - \text{تابع}$ معرف تابعی از ψ است که مشتق $\rho - \text{تابع}$ است و به خصوص در شرایط زیر صدق کند

P1 ψ تابعی فرد است.

P2 برای $x \geq \circ$ ، $\psi(x) \geq \circ$.

برای اطلاعات بیشتر به مارونا [۴۸] رجوع شود.

۲.۶.۱ MM-برآوردگرها

روش برآورد رگرسیونی MM، برآوردگری مقاوم نسبت به داده‌های ناهمگون (داده‌های دورافتاده در متغیر پاسخ و همچنین نقاط نافذ در متغیرهای توضیحی) را به دست می‌دهد که تحت فرض نرمال بودن کارا هستند. روش MM مبتنی بر دو تابع ρ مختلف، $\rho_\circ$ و ρ_1 ، برای تعیین نقطه فروریزش و کارایی است. این توابع باید کراندار و دارای ویژگی‌های زیر باشند

A1 : (i) $\rho(\circ) = \circ$.

(ii) $\rho(-x) = \rho(x)$.

(iii) $\rho(x)$ پیوسته است.

(iv) $\sup \rho(x) = a < \infty$.

(v) اگر $\circ \leq x \leq y$ آن گاه $\rho(x) \leq \rho(y)$.

(vi) اگر $\rho(x) < a$ و $\circ \leq x < y$ ، آن گاه $\rho(x) < \rho(y)$.

A2 : $\rho_1(x) \leq \rho_\circ(x)$ و $\sup \rho_1(x) = \sup \rho_\circ(x) = a$.

برای به دست آوردن برآوردگر رگرسیونی MM ابتدا با یک برآوردگر اولیه، $\hat{\beta}_s$ ، برای بردار پارامتر β شروع می‌کنیم. این برآوردگر اولیه باید با نقطه فروریزش بالایی استوار باشد، اما لزومی ندارد کارا باشد. با استفاده از $\hat{\beta}_s$ باقی‌مانده‌های اولیه $r_i(\hat{\beta}_s) = y_i - \mathbf{x}_i^\top \hat{\beta}_s$ ، برای $i = 1, 2, \dots, n$ را محاسبه می‌کنیم. سپس M-برآوردگر مقیاس $\hat{\sigma}_s$ را برای باقی‌مانده‌های اولیه $r_i(\hat{\beta}_s)$ ، با حل معادله M-برآورد زیر برای پارامتر مقیاس، پیدا می‌کنیم

$$\frac{1}{n} \sum_{i=1}^n \rho_\circ \left(\frac{r_i(\hat{\beta}_s)}{\sigma} \right) = \circ / 5 \quad (19.1)$$

و در نهایت برای به‌دست آوردن MM-برآوردگر $(\hat{\beta}_{MM})$ برای بردار پارامتر β ، مینیمم مطلق عبارت زیر را پیدا می‌کنیم

$$l_n(\beta) = \sum_{i=1}^n \rho_1\left(\frac{r_i(\beta)}{\hat{\sigma}_s}\right) \quad (20.1)$$

که در آن $r_i(\beta) = y_i - \mathbf{x}_i^\top \beta$ است. اگر ρ_1 دارای مشتق اول $\rho_1'(t) = \psi_1(t)$ باشد، آن‌گاه برآوردگر رگرسیون MM از $\hat{\beta}_{MM}$ ، هر جوابی از معادله برآوردکننده زیر خواهد بود

$$\sum_{i=1}^n \psi_1\left(\frac{r_i(\beta)}{\hat{\sigma}_s}\right) \mathbf{x}_i = \mathbf{0} \quad (21.1)$$

که در نامساوی زیر صدق کند

$$l_n(\hat{\beta}_{MM}) \leq l_n(\hat{\beta}_s) \quad (22.1)$$

انتخاب تابع ρ برای داشتن کارایی و نقطه فروریزش بالا بسیار مهم است. یوهای [۶۵]، مارونا و همکاران [۴۸] و مارونا [۴۷] از ρ -تابع دو مربعی باضابطه زیر استفاده کردند

$$\rho(r) = \min\{1, 1 - (1 - r^2)^3\}, \quad r \in \mathbb{R} \quad (23.1)$$

و بنابراین تعریف می‌کنیم $\rho_0(r) = \rho(r/c_0)$ و $\rho_1(r) = \rho(r/c_1)$. لازم به ذکر است برای اینکه $\rho_1 \leq \rho_0$ باید $c_0 \leq c_1$ انتخاب شود. در این پایان‌نامه نیز ما از همین تابع استفاده خواهیم کرد. برای جزئیات و ویژگی‌های بیش‌تر از روش رگرسیون MM به یوهای [۶۵]، مارونا و همکاران [۴۸] و مارونا و یوهای [۴۹] مراجعه فرمایید.

۷.۱ برآوردهای مبتنی بر باقی‌مانده‌های استوار مقیاس

در این بخش خانواده‌ای از برآوردهای رگرسیونی را معرفی می‌کنیم که به مقیاس اولیه بستگی ندارد و از این رو به‌عنوان برآوردگر اولیه برای MM-برآورد مفید هستند. اعضای به‌خصوصی از این خانواده، برآورد اولیه $\hat{\beta}_s$ و باقی‌مانده‌های متناظر خوبی برای تعریف برآورد اولیه مقیاس $\hat{\sigma}_s$ در روش MM-برآورد به‌دست می‌دهند.

اگر $\hat{\sigma} = \hat{\sigma}(\mathbf{r})$ یک برآورد مقیاس استوار هم‌وردای مقیاسی مبتنی بر بردار باقی‌مانده‌های زیر باشد

$$\mathbf{r}(\beta) = (r_1(\beta), \dots, r_n(\beta))^\top \quad (24.1)$$

آن‌گاه یک برآوردگر رگرسیونی می‌تواند به‌صورت زیر تعریف شود

$$\hat{\beta} = \arg \min_{\beta \in \mathbb{R}^p} \hat{\sigma}(\mathbf{r}(\beta)) \quad (25.1)$$

۱.۷.۱ S-برآوردها

یکی از انواع خیلی مهم از برآوردگرهای رگرسیونی (۲۵.۱) زمانی است که $\hat{\sigma}(\mathbf{r})$ برای هر $\mathbf{r}$ ، M -برآورد مقیاسی تعریف شود بدین صورت که

$$\frac{1}{n} \sum_{i=1}^n \rho\left(\frac{r_i}{\hat{\sigma}}\right) = \delta \quad (26.1)$$

که در آن $\delta \in [0, 0.5]$ و ρ یک ρ -تابع کراندار است. این برآورد نتیجه‌شده از (۲۵.۱) را یک S -برآورد می‌گویند (راسو و یوحای [۵۶]).

۲.۷.۱ L-برآوردها

یک راه دیگر برای استفاده از یک M -مقیاس استفاده از L -برآورد مقیاس است. مقادیر مطلق ترتیبی برای باقی‌مانده‌ها را به صورت $|r|_{(1)} \leq \dots \leq |r|_{(n)}$ تعریف می‌کنیم، سپس می‌توانیم برآوردهای مقیاسی را به‌عنوان ترکیباتی از $|r|_{(i)}$ در یکی از دو شکل زیر تعریف کنیم

$$\hat{\sigma} = \sum_{i=1}^n a_i |r|_{(i)} \quad \text{یا} \quad \hat{\sigma} = \left(\sum_{i=1}^n a_i |r|_{(i)}^p \right)^{\frac{1}{p}}$$

که در آن a_i ها ثابت‌هایی نامنفی هستند.

۸.۱ برآوردگرهای جریمه‌شده استوار

تاکنون به بررسی برآوردگرهای جریمه‌شده (برآوردگر بریج و حالات خاص آن) و همچنین برآوردگرهای استوار (MM -برآوردگرها) پرداخته شد. اکنون با ترکیب برآوردگرهای رگرسیونی MM و بریج، حالت خاصی از برآوردگرهای جریمه‌شده استوار را معرفی می‌کنیم (آرسلان [۱۳]).

تحت مفروضات مدل (۳.۱) در مسئله رگرسیونی جریمه‌شده استوار به‌دنبال دستیابی به برآوردگری استوار از β می‌باشیم، که جوابی از مسئله بهینه‌سازی زیر می‌باشد

$$\hat{\beta} = \arg \min_{\beta \in \mathbb{R}^p} \left[\hat{\sigma}_s^2 \sum_{i=1}^n \rho_1\left(\frac{r_i(\beta)}{\hat{\sigma}_s}\right) + k \sum_j P(|\beta_j|) \right] \quad (27.1)$$

که در آن $\rho, \rho(\cdot)$ -تابع برای باقی‌مانده‌ها و عبارت $\sum_j P(\beta_j)$ ، عبارت جریمه می‌باشد. همچنین عبارت $\hat{\sigma}_s^2$ تضمین‌کننده این است که برآوردگر حاصل در صورت انتخاب $\rho_1(r) = r^2$ به برآوردگر بریج معرفی‌شده در (۷.۱) تبدیل شود. در این باره به تفصیل در فصل ۴ بحث خواهد شد.

در ادامه با توجه به این که در فصل‌های بعد، صورت فازی برآوردگرهای مورد بررسی نیز ارائه می‌شود، لازم است که به‌طور مختصر مجموعه‌های فازی و مباحث مرتبط به آن تعریف شوند.

۹.۱ مقدمه‌ای بر مجموعه‌های فازی

در مسائل کاربردی که نادقیق بودن مولفه‌ها واضح است، استفاده از مجموعه‌های فازی در بررسی ویژگی‌ها و اندازه‌گیری میزان کارایی آن‌ها می‌تواند برخی از محدودیت‌های فعلی را برطرف نماید. در رگرسیون کلاسیک نیز فرض‌هایی هم‌چون دقت بودن متغیرها و مشاهدات مربوط به آن‌ها و عدم وجود هم‌خطی بین متغیرها در نظر گرفته می‌شود و در صورت برقرار نبودن این فرض‌ها، رگرسیون کلاسیک کارایی مطلوبی نخواهد داشت. از این رو با استفاده از مجموعه‌های فازی برآوردهایی را معرفی می‌کنیم که تا حدودی کارایی برآوردهای کلاسیک را بهبود می‌دهند. برای یادآوری مروری بر نظریه مجموعه‌های فازی خواهیم کرد اما برای اطلاعات بیشتر به طاهری [۷] و طاهری و ماشین‌چی [۸] مراجعه کنید.

مجموعه فازی $\tilde{A}$ از مجموعه مرجع $\mathcal{X}$ به صورت $\mathcal{F} = \{(x, A(x)) | x \in \mathcal{X}\}$ است که در آن $A(x) : \mathcal{X} \rightarrow [0, 1]$ تابع عضویت x در $\tilde{A}$ است. گوییم مجموعه فازی $\tilde{A}$ زیر مجموعه‌ی مجموعه فازی $\tilde{B}$ است، اگر برای هر $x \in \mathcal{X}$ ، $A(x) \leq B(x)$ یعنی

$$\tilde{A} \subseteq \tilde{B} \Leftrightarrow A(x) \leq B(x), \forall x \in \mathcal{X}$$

تعریف ۱.۹.۱. اگر عدد فازی $\tilde{A}$ دارای تابع عضویت زیر باشد

$$A(x) = \begin{cases} L\left(\frac{m-x}{l}\right) & x \leq m \\ R\left(\frac{x-m}{r}\right) & x > m \end{cases} \quad (28.1)$$

که در آن L و R توابعی غیرصعودی از $\mathbb{R}^+$ به $[0, 1]$ هستند و $L(0) = R(0) = 1$ ، آن‌گاه $\tilde{A}$ را یک عدد فازی LR نامیده و با نماد $\tilde{A} = (m, l, r)_{LR}$ نشان می‌دهیم. عدد m را مقدار نما (میانه) و اعداد مثبت l و r را به ترتیب پهنای چپ و پهنای راست $\tilde{A}$ می‌نامیم. L و R نیز توابع مرجع (یا توابع شکل) نامیده می‌شوند.

تعریف ۲.۹.۱. فرض کنید $\tilde{A} = (m, l, r)_{LR}$ ، در این صورت $\tilde{A}$ را یک عدد فازی مثلثی نامیده و با $\tilde{A} = (m, l, r)_T$ نشان می‌دهیم، اگر به صورت زیر باشد

$$A(x) = \begin{cases} \frac{x-(m-l)}{l} & m-l \leq x < m \\ \frac{(m+r)-x}{r} & m \leq x < m+r \\ 0 & \text{در غیر این صورت} \end{cases} \quad (29.1)$$

همچنین مجموعه تمام اعداد فازی مثلثی را با $T(\mathbb{R})$ نمایش می‌دهیم.

تعریف ۳.۹.۱. برای هر عدد فازی $\tilde{A} = (m_A, l_A, r_A)_T$ و $h \in [0, 1]$ ، $-h$ برش $\tilde{A}$ به صورت زیر تعریف می‌شود

$$[A^L(h), A^U(h)] = [m_A - (1-h)l_A, m_A + (1-h)r_A] \quad (30.1)$$

تعریف ۴.۹.۱. عدد فازی $\tilde{A}$ را مثبت (منفی) گوییم اگر برای هر $(x \geq 0)$ ، $A(x) = 0$ ، عدد فازی مثبت و منفی را به ترتیب با نمادهای $\tilde{A} \succ 0$ و $\tilde{A} \prec 0$ نمایش می‌دهیم.

تعریف ۵.۹.۱. اگر برای عدد فازی $\tilde{A}$ ، $L = R$ و $l = r$ ، آن‌گاه $\tilde{A}$ را یک عدد فازی متقارن نامیده و به صورت $\tilde{A} = (m, l)_L$ نشان می‌دهیم. همچنین یک عدد حقیقی می‌تواند به‌عنوان یک عدد مثلثی فازی در نظر گرفت که پهنای آن صفر هستند.

۱.۹.۱ حساب اعداد فازی

تعریف ۶.۹.۱. فرض کنید $\tilde{A}$ و $\tilde{B}$ دو عدد فازی و $X : \mathbb{R} \times \mathbb{R}$ یک عملگر دوتایی بر اعداد حقیقی باشد. اگر عملگر $*$ را برای اعداد فازی با $\otimes$ نشانه دهیم، حاصل عبارت $\tilde{A} \otimes \tilde{B}$ با استفاده از اصل توسیع به‌صورت یک مجموعه فازی با تابع عضویت زیر تعریف می‌شود

$$(\tilde{A} \otimes \tilde{B})(z) = \sup_{z=x*y} \inf[A(x), B(y)] \quad (۳۱.۱)$$

و این تعریف برای چهار عمل اصلی به‌صورت‌های زیر است

$$\begin{aligned} (\tilde{A} \oplus \tilde{B})(z) &= \sup_{z=x+y} \inf[A(x), B(y)] \\ (\tilde{A} \ominus \tilde{B})(z) &= \sup_{z=x-y} \inf[A(x), B(y)] \\ (\tilde{A} \otimes \tilde{B})(z) &= \sup_{z=x \times y} \inf[A(x), B(y)] \\ (\tilde{A} \oslash \tilde{B})(z) &= \sup_{z=x/y} \inf[A(x), B(y)] \\ (k \odot \tilde{B})(z) &= \sup_{z=k \cdot y} \inf[A(x), B(y)] \end{aligned}$$

برخی از اعمال جبری برای اعداد فازی LR بسیار ساده و دارای الگوی مشخصی است و برای برخی دیگر از اعمال جبری نیز الگوهای تقریبی موجود است که در زیر برخی روابط به‌صورت قضیه‌هایی بدون اثبات در حساب اعداد فازی LR بیان می‌شوند.

قضیه ۱.۹.۱. (دووا و پراد [۲۷]) اگر $\tilde{A} = (m_A, l_A, r_A)_{LR}$ و $\tilde{B} = (m_B, l_B, r_B)_{LR}$ آن‌گاه بر اساس اصل گسترش داریم

$$\tilde{A} \oplus \tilde{B} = (m_A + m_B, l_A + l_B, r_A + r_B)_{LR} \quad (۳۲.۱)$$

$$\tilde{A} \ominus \tilde{B} = (m_A - m_B, l_A + r_B, r_A + l_B)_{LR} \quad (۳۳.۱)$$

برای ضرب یک عدد فازی مثلثی با عدد حقیقی k داریم

$$k\tilde{A} = \begin{cases} (km_A, kl_A, kr_A)_T & k > 0 \\ (km_A, -kr_A, -kl_A)_T & k < 0 \end{cases} \quad (۳۴.۱)$$

اما در مورد ضرب یا تقسیم اعداد فازی نمی‌توان به این راحتی عمل کرد، همچنین در حالتی که $\tilde{A}$ و $\tilde{B}$ دو عدد فازی LR باشند، نتیجه ضرب آن‌ها یک عدد فازی LR نخواهد بود. قضیه‌های زیر روابطی تقریبی برای ضرب و تقسیم بیان می‌کند.

قضیه ۲.۹.۱. (دووا و پراد [۲۷]) اگر $\tilde{A} = (m_A, l_A, r_A)_{LR}$ و $\tilde{B} = (m_B, l_B, r_B)_{LR}$ آن‌گاه

$$(\tilde{A} \otimes \tilde{B}) \simeq \begin{cases} (m_A m_B, m_A l_B + m_B l_A, m_A r_B + m_B r_A)_{LR} & \tilde{A} \succ \circ, \tilde{B} \succ \circ \\ (m_A m_B, m_A l_B - m_B r_A, m_A r_B - m_B l_A)_{RL} & \tilde{A} \succ \circ, \tilde{B} \prec \circ \\ (m_A m_B, -m_B r_A - m_A r_B, -m_B l_A - m_A l_B)_{RL} & \tilde{A} \prec \circ, \tilde{B} \prec \circ \end{cases} \quad (35.1)$$

قضیه ۳.۹.۱. (دووا و پراد [۲۷]) اگر $\tilde{A} = (m_A, l_A, r_A)_{LR}$ یک عدد فازی مثبت باشد، آن‌گاه

$$\tilde{A}^{-1} \simeq (m_A^{-1}, r_A m_A^{-2}, l_A m_A^{-2})_{RL}$$

همچنین اگر $\tilde{A} = (m_A, l_A, r_A)_{LR}$ و $\tilde{B} = (m_B, l_B, r_B)_{LR}$ دو عدد فازی مثبت باشند، آن‌گاه

$$(\tilde{A} \oslash \tilde{B}) \simeq \left(\frac{m_A}{m_B}, \frac{r_B m_A + l_A m_B}{m_B^2}, \frac{l_B m_A + r_A m_B}{m_B^2} \right)_{LR}$$

هرچه پهناهای اعداد فازی $\tilde{A}$ و $\tilde{B}$ نسبت به مقادیر نمای آن‌ها کوچکتر باشد، تقریب‌های فوق دقیق‌تر هستند و در همسایگی نما، این روابط از دقت بیشتری برخوردارند.

قضیه ۴.۹.۱. (دووا و پراد [۲۷]) فرض کنید $\tilde{A}$ ، $\tilde{B}$ و $\tilde{C}$ اعداد فازی باشند، در این صورت

$$(\tilde{A} \oplus \tilde{B}) \otimes \tilde{C} \subseteq (\tilde{A} \otimes \tilde{C}) \oplus (\tilde{B} \otimes \tilde{C}) \quad (36.1)$$

و تساوی زمانی برقرار است اگر $\tilde{A}, \tilde{B}, \tilde{C} \succ \circ$.

۲.۹.۱ فاصله بین دو عدد فازی

در بحث رگرسیون فازی برای بدست آوردن پارامترهای مدل لازم است خطا یا به عبارت دقیق‌تر فاصله بین دو مقدار مشاهده‌شده و برآوردشده فازی محاسبه شود. بنابراین به معرفی برخی از انواع این فاصله‌ها می‌پردازیم.

فاصله موزون

ژو و لی [۶۴] فاصله موزون^{۲۶} بین دو عدد فازی را به صورت زیر تعریف می‌کند

²⁶Weighted distance

تعریف ۷.۹.۱. فرض کنید $\tilde{A}$ و $\tilde{B}$ دو عدد فازی باشند، فاصله موزون بین این دو عدد فازی به صورت زیر تعریف می شود

$$d(\tilde{A}, \tilde{B}) = \left[\int_0^1 f(h) d^{\vee}(A_h, B_h) dh \right]^{\frac{1}{\vee}}$$

که در آن A_h و B_h ، h -برش های دو عدد فازی هستند و

$$d^{\vee}(A_h, B_h) = [A_h^L - B_h^L]^{\vee} + [A_h^U - B_h^U]^{\vee}$$

و $A_h^L, B_h^L, A_h^U, B_h^U$ کران های بالا و پایین h -برش $\tilde{A}$ و $\tilde{B}$ هستند. همچنین $f(h)$ یک تابع صعودی روی بازه $[0, 1]$ است و $f(0) = 0$ و $\int_0^1 f(h) dh = \frac{1}{\vee}$. تابع $f(h)$ را می توان به عنوان وزن $d^{\vee}(A_h, B_h)$ در نظر گرفت. ویژگی صعودی بودن $f(h)$ تضمین کننده این است که در تعیین فاصله بین دو عدد فازی $\tilde{A}$ و $\tilde{B}$ ، درجات بالاتر اهمیت بیشتری داشته باشند. شرط های $f(0) = 0$ و $\int_0^1 f(h) dh = \frac{1}{\vee}$ نیز این اطمینان را می دهند که فاصله فوق، یک تعمیم معمولی از فاصله بین دو عدد حقیقی است.

قضیه ۵.۹.۱. (بالاساندرام [۱۴]) تحت مفروضات تعریف ۷.۹.۱ فرض کنید $\tilde{A}$ و $\tilde{B}$ دو عدد فازی مثلثی باشند، آن گاه بر پایه ی تابع وزنی $f(h) = h$ داریم

$$\begin{aligned} d^{\vee}(\tilde{A}, \tilde{B}) &= (m_A - m_B)^{\vee} \\ &+ \frac{1}{\vee} (m_A - m_B) [(r_A - r_B) - (l_A - l_B)] \\ &+ \frac{1}{1+\vee} [(l_A - l_B)^{\vee} + (r_A - r_B)^{\vee}] \end{aligned} \quad (37.1)$$

نتیجه ۱.۹.۱. اگر $\tilde{A} = (m_A, l_A)_T$ و $\tilde{B} = (m_B, l_B)_T$ دو عدد فازی مثلثی متقارن باشند، آن گاه رابطه (۳۷.۱) به صورت زیر به دست می آید

$$d^{\vee}(\tilde{A}, \tilde{B}) = (m_A - m_B)^{\vee} + \frac{1}{\vee} (l_A - l_B)^{\vee} \quad (38.1)$$

نتیجه ۲.۹.۱. در رابطه (۳۷.۱) فاصله هر عدد فازی مثلثی مانند $\tilde{A}$ از مبدأ، $\tilde{o} = (0, 0, 0)_T$ به صورت

$$\begin{aligned} d^{\vee}(\tilde{A}, \tilde{o}) &= \|\tilde{A}\|^{\vee} \\ &= m_A^{\vee} + \frac{1}{\vee} m_A (r_A - l_A) + \frac{1}{1+\vee} (l_A^{\vee} + r_A^{\vee}) \end{aligned} \quad (39.1)$$

محاسبه می شود.

تعریف ۸.۹.۱. (بالاساندرام [۱۴]) اگر $\tilde{W} = (\tilde{W}_1, \dots, \tilde{W}_n)^T$ یک بردار فازی باشد به طوری که به ازای $i = 1, \dots, n$ ها اعداد فازی هستند، آن گاه نرم بردار فازی $\tilde{W}$ به صورت زیر تعریف می شود

$$\|\tilde{W}\|^{\vee} = \sum_{i=1}^n \|\tilde{W}_i\|^{\vee} \quad (40.1)$$

فاصله دونسو

دونسو و همکاران [۲۶] فاصله بین دو عدد فازی را به صورت زیر تعریف می کنند.

تعریف ۹.۹.۱. فرض کنید $\tilde{A} = (m_A, l_A, r_A)_T$ و $\tilde{B} = (m_B, l_B, r_B)_T$ دو عدد فازی باشند، فاصله بین دو عدد فازی به صورت زیر تعریف می شود

$$d^2(\tilde{A}, \tilde{B}) = k_1(m_A - m_B)^2 + k_2((m_A - l_A - (m_B - l_B))^2 + (m_A + r_A - (m_B + r_B))^2) \quad (۴۱.۱)$$

که در آن k_1 و k_2 اعداد حقیقی بین صفر و یک هستند که توسط متخصص تعیین می شوند. در صورتی که $k_1 > k_2$ اهمیت مرکز بیش تر از پهنای پاسخ ها در نظر گرفته شده و اگر $k_1 < k_2$ ، پهنای پاسخ ها از اهمیت بیش تری برخوردار است. در صورت تساوی ($k_1 = k_2$) اهمیت پهنای و مراکز پاسخ ها یکسان می شوند.

فاصله دیاموند

تعریف ۱۰.۹.۱. (دیاموند [۲۴]) این فاصله حالت خاصی از فاصله دونسو است که در آن $k_1 = k_2 = ۱$. برای هر دو عدد فازی $\tilde{A} = (m_A, l_A, r_A)_T$ و $\tilde{B} = (m_B, l_B, r_B)_T$ فاصله دیاموند به صورت زیر تعریف می شود

$$d^2(\tilde{A}, \tilde{B}) = [(m_A - l_A) - (m_B - l_B)]^2 + [(m_A + r_A) - (m_B + r_B)]^2 + [m_A - m_B]^2 \quad (۴۲.۱)$$

فاصله همینگ

تعریف ۱۱.۹.۱. (ایزدی خواه [۳۹]) برای اعداد فازی $\tilde{A} = (m_A, l_A, r_A)_T$ و $\tilde{B} = (m_B, l_B, r_B)_T$ فاصله همینگ به صورت زیر تعریف می شود

$$d(\tilde{A}, \tilde{B}) = \int_{\mathbb{R}} |A(x) - B(x)| dx \quad (۴۳.۱)$$

ملاحظه ۱.۹.۱. توجه کنید که اگر $\tilde{A} = [A_h^L, A_h^U]$ ، با انتخاب $h = ۱$ می توان مقدار مرکز و با انتخاب $h = ۰$ ، مقدار پهنای چپ و راست $\tilde{A}$ را مشخص نمود.

فاصله قدرمطلق

حسن پور و همکاران [۳۴] فاصله بین دو عدد فازی را به صورت زیر بیان کردند.

تعریف ۱۲.۹.۱. فرض کنید $\tilde{A} = (m_A, l_A, r_A)_T$ و $\tilde{B} = (m_B, l_B, r_B)_T$ دو عدد فازی باشند. فاصله $d : T(\mathbb{R}) \times T(\mathbb{R}) \rightarrow \mathbb{R}$ را به صورت زیر تعریف می کنیم

$$d(\tilde{A}, \tilde{B}) = |m_A - m_B| + |l_A - l_B| + |r_A - r_B| \quad (۴۴.۱)$$

ملاحظه ۲.۹.۱. برای دو عدد غیرفازی (حقیقی مقدار)، فاصله (۴۴.۱) به قدرمطلق تفاوت بین آن‌ها تبدیل می‌شود.

۱۰.۱ رگرسیون فازی

در این بخش مروری بر رگرسیون در محیط فازی خواهیم کرد و رگرسیون کمترین توان‌های دوم فازی^{۲۷} (FLS) را با استفاده از فاصله‌های موزون و قدرمطلق بیان می‌کنیم.

۱.۱۰.۱ انواع رگرسیون فازی

در یک تقسیم‌بندی کلی، می‌توان رگرسیون فازی را به سه حالت زیر تقسیم کرد:

۱. حالتی که ارتباط بین متغیرها فازی در نظر گرفته شود. به عبارت دیگر ضرایب معادله رگرسیونی فازی در نظر گرفته شوند.

۲. حالتی که متغیرها (اعم از توضیحی یا پاسخ) نادقیق و فازی هستند.

۳. در حالتی که هم متغیرها و هم ضرایب مدل، فازی در نظر گرفته شوند.

لازم به ذکر است که تنوع در رگرسیون فازی به حالت‌های فوق محدود نمی‌شود و شیوه‌هایی برای هر حالت پیشنهاد شده‌اند. این شیوه‌ها تنوع بسیاری را در روش‌های رگرسیون فازی ایجاد کرده‌اند. همانند رگرسیون کلاسیک، رگرسیون فازی را نیز می‌توان تقسیم‌بندی نمود. از این انواع می‌توان به روش‌های رگرسیونی امکانی، رگرسیون کمترین توان‌های دوم خطای فازی و رگرسیون قدرمطلق خطای فازی اشاره کرد. در ادامه به معرفی برخی از این روش‌ها خواهیم پرداخت.

ملاحظه ۱.۱۰.۱. در سرتاسر این پایان‌نامه از اعداد فازی مثلثی استفاده شده است.

۲.۱۰.۱ رگرسیون کمترین توان‌های دوم فازی

روش کمترین توان‌های دوم فازی برای اولین بار توسط دیاموند [۲۳] و سلمینس [۲۱] ارائه شد. این روش تعمیمی از رگرسیون کمترین توان‌های دوم در حالت کلاسیک است با این تفاوت که بر پایه‌ی تعریف‌هایی برای فاصله بین اعداد فازی می‌باشد. دیاموند [۲۳] با تعریف فاصله‌ای برای اعداد فازی، مدل‌هایی را برای برازش کمترین توان‌های دوم برای مدل با متغیر مستقل غیرفازی و متغیر وابسته فازی و همچنین برای مدل با متغیر مستقل و وابسته فازی، پیشنهاد داد. دیاموند و کورنر [۲۵] رگرسیون فازی را براساس متغیرهای تصادفی فازی

²⁷Fuzzy Least Squares

ارائه کردند. همچنین ژو و لی [۶۷] روش کمترین توان‌های دوم را برای برآورد ضرایب فازی مدل با استفاده از متری بر اساس سطوح تراز ارائه کردند و با یک معیار نیکویی برازش، میزان دقت مدل را اندازه‌گیری کردند. در روش کمترین توان‌های دوم فازی، پارامترها به گونه‌ای برآورد می‌شوند که به ازای هر مقدار، فاصله‌ی بین متغیر وابسته مشاهده‌شده و متغیر وابسته برآوردشده، بر اساس فاصله d مینیمم شود. به عبارتی به دنبال مینیمم کردن مجموع توان‌های دوم خطا هستیم.

ابتدا یک مدل خطی با ضرایب فازی $\tilde{A}_j$ ، $j = 0, \dots, p$ ، به صورت

$$\tilde{Y}_i = \tilde{A}_0 + \tilde{A}_1 x_{i1} + \dots + \tilde{A}_p x_{ip}, \quad i = 1, \dots, n \quad (۴۵.۱)$$

بر پایه مشاهدات

$$(\tilde{Y}_i, x_{i1}, \dots, x_{ip}), i = 1, \dots, n$$

که در آن x_{ij} ها اعداد حقیقی و $\tilde{Y}_i$ ها اعداد فازی هستند، به داده‌ها برازش می‌دهیم. بدون کاستن از کلیت مساله فرض می‌کنیم $x_{ij} > 0$ (همچنین در صورتی که x_{ij} ها اعداد منفی باشند، می‌توان با یک انتقال آن‌ها را به اعداد مثبت تبدیل کرد). با استفاده از روش کمترین توان‌های دوم خطا و با توجه به فاصله (۳۸.۱)، مدل (۴۵.۱) را به داده‌ها برازش می‌دهیم، به این معنی که ضرایب فازی را به گونه‌ای می‌یابیم که مجموع توان‌های دوم خطای زیر، مینیمم شود

$$SSE(\tilde{A}_0, \tilde{A}_1, \dots, \tilde{A}_p) = \sum_{i=1}^n d^2(\tilde{Y}_i, \tilde{A}_0 + \tilde{A}_1 x_{i1} + \dots + \tilde{A}_p x_{ip}) \quad (۴۶.۱)$$

برای راحتی در انجام محاسبات، بحث را به حالتی که $\tilde{Y}_i = (m_{Y_i}, l_{Y_i})_T$ و $\tilde{A}_j = (m_{A_j}, l_{A_j})_T$ اعداد فازی مثلثی متقارن هستند، محدود می‌کنیم. براساس عملگرهای فازی بین اعداد فازی، داریم

$$\tilde{A}_0 + \tilde{A}_1 x_{i1} + \dots + \tilde{A}_p x_{ip} = (m_{A_0} + m_{A_1} x_{i1} + \dots + m_{A_p} x_{ip}, l_{A_0} + l_{A_1} x_{i1} + \dots + l_{A_p} x_{ip})_T$$

حال با استفاده از فاصله موزون گفته‌شده در بخش ۲.۹.۱ و نتیجه (۳۸.۱)، مجموعه توان‌های دوم خطا به صورت زیر حاصل می‌شود

$$\begin{aligned} SSE(\tilde{A}_0, \tilde{A}_1, \dots, \tilde{A}_p) &= \sum_{i=1}^n (m_{A_0} + m_{A_1} x_{i1} + \dots + m_{A_p} x_{ip} - m_{Y_i})^2 \\ &\quad + \frac{1}{6} \sum_{i=1}^n (l_{A_0} + l_{A_1} x_{i1} + \dots + l_{A_p} x_{ip} - l_{Y_i})^2 \end{aligned}$$

برای مینیمم کردن تابع $SSE(\tilde{A}_0, \tilde{A}_1, \dots, \tilde{A}_p)$ نسبت به m_{A_j} و l_{A_j} مشتق گرفته و برابر صفر

قرار می‌دهیم، معادلات زیر حاصل می‌شوند

$$\sum_{i=1}^n (m_{A_0} + m_{A_1}x_{i1} + \cdots + m_{A_p}x_{ip} - m_{Y_i})x_{ij} = 0$$

$$\sum_{i=1}^n (l_{A_0} + l_{A_1}x_{i1} + \cdots + l_{A_p}x_{ip} - l_{Y_i})x_{ij} = 0$$

که در آن $j = 1, \dots, p$ و به ازای $i = 1, \dots, n$ ، $x_{i0} = 1$ منظور شده است. می‌توان معادلات بالا را به صورت زیر نوشت

$$m_{A_0} \sum_{i=1}^n x_{i0}x_{ij} + m_{A_1} \sum_{i=1}^n x_{i1}x_{ij} + \cdots + m_{A_p} \sum_{i=1}^n x_{ip}x_{ij} = \sum_{i=1}^n m_{Y_i}x_{ij} \quad (47.1)$$

$$l_{A_0} \sum_{i=1}^n x_{i0}x_{ij} + l_{A_1} \sum_{i=1}^n x_{i1}x_{ij} + \cdots + l_{A_p} \sum_{i=1}^n x_{ip}x_{ij} = \sum_{i=1}^n l_{Y_i}x_{ij} \quad (48.1)$$

اکنون فرض کنید

$$\mathbf{X} = \begin{pmatrix} x_{10} & x_{11} & \cdots & x_{1p} \\ x_{20} & x_{21} & \cdots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n0} & x_{n1} & \cdots & x_{np} \end{pmatrix}$$

$$\mathbf{m}_A = (m_{A_0}, \dots, m_{A_p})^\top, \quad \mathbf{m}_Y = \left(\sum_{i=1}^n m_{Y_i}x_{i0}, \dots, \sum_{i=1}^n m_{Y_i}x_{ip} \right)^\top$$

$$\mathbf{l}_A = (l_{A_0}, \dots, l_{A_p})^\top, \quad \mathbf{l}_Y = \left(\sum_{i=1}^n l_{Y_i}x_{i0}, \dots, \sum_{i=1}^n l_{Y_i}x_{ip} \right)^\top$$

در این صورت معادلات (47.1) و (48.1) را می‌توان به صورت ماتریسی زیر نمایش داد

$$(\mathbf{X}^\top \mathbf{X})\mathbf{m}_A = \mathbf{m}_Y, \quad (\mathbf{X}^\top \mathbf{X})\mathbf{l}_A = \mathbf{l}_Y \quad (49.1)$$

نتیجه ۱.۱۰.۱. اگر $\text{Rank}(\mathbf{X}) = p+1$ و $(\mathbf{X}^\top \mathbf{X})^{-1}\mathbf{l}_Y \geq 0$ ، آن گاه دستگاه (49.1) دارای جواب یکتاست و از روابط زیر به دست می‌آید

$$\hat{\mathbf{m}}_A = (\mathbf{X}^\top \mathbf{X})^{-1}\mathbf{m}_Y, \quad \hat{\mathbf{l}}_A = (\mathbf{X}^\top \mathbf{X})^{-1}\mathbf{l}_Y \quad (50.1)$$

لازم به ذکر است، شرط $(\mathbf{X}^\top \mathbf{X})^{-1}\mathbf{l}_Y \geq 0$ تضمین کننده نامنفی بودن پهنای ضرایب مدل فازی، l_{A_j} ها، است.

۳.۱۰.۱ برآوردگر استوار فازی بر پایه فاصله قدرمطلق

اکنون حالتی را در نظر بگیرید که مجموعه داده‌های فازی به صورت زیر باشند

$$\{(\tilde{x}_{i1}, \tilde{x}_{i2}, \dots, \tilde{x}_{ip}, \tilde{y}_i) \mid i = 1, \dots, n\}$$

که در آن مقدار z امین متغیر ورودی نیز عدد فازی نامتقارن است، $\tilde{x}_{ij} = (m_{x_{ij}}, l_{x_{ij}}, r_{x_{ij}})_T$ ، $j = 1, \dots, p$ و $i = 1, \dots, n$ است که همه آن‌ها مثبت و اگر نباشند با استفاده از یک انتقال ساده داده‌ها مثبت می‌شوند (دیاموند [۲۴] را ببینید). همچنین مشاهدات پاسخ نیز اعداد فازی مثلثی نامتقارن به صورت $\tilde{y}_i = (m_{y_i}, l_{y_i}, r_{y_i})_T$ ، $i = 1, \dots, n$ هستند. در این صورت مدل رگرسیونی خطی فازی (۴۵.۱) به صورت زیر تبدیل می‌شود

$$\tilde{Y}_i = \tilde{A}_0 \oplus (\tilde{A}_1 \otimes \tilde{x}_{1i}) \oplus \dots \oplus (\tilde{A}_p \otimes \tilde{x}_{pi}), \quad i = 1, \dots, n \quad (۵۱.۱)$$

از آنجا که برای برآورد ضرایب رگرسیونی، محاسبه ضرب دو عدد فازی مثلثی $\tilde{x}_j$ و $\tilde{A}_j$ اجتناب‌ناپذیر است. بنابراین طبق قضیه ۲.۹.۱ این ضرب جواب دقیق نداشته و تقریب آن نیز به علامت اعداد فازی مثلثی وابسته است. از این رو شخص باید برای مدل‌های مختلف، علامت‌های مختلف ضرایب رگرسیونی بسازد. بنابراین از مدلی استفاده می‌کنیم که به علامت ضرایب رگرسیونی وابسته نباشد.

بیاد بیاورید که یک عدد حقیقی a نمایشی متناهی به صورت $a = a' - a''$ دارد، که در آن a' و a'' اعداد حقیقی نامنفی هستند. فرض کنید $\tilde{A} = (m_A, l_A, r_A)_T$ و $\tilde{x} = (m_x, l_x, r_x)_T$ ، به‌طوری که در آن $\tilde{A}$ نامقید و $\tilde{x} \succ \tilde{0}$. قرار می‌دهیم $\tilde{A} = \tilde{A}' + \tilde{A}''$ که در آن $\tilde{A}' = (m_{A'}, l_{A'}, r_{A'})_T$ و $\tilde{A}'' = (-m_{A''}, l_{A''}, r_{A''})_T$ همچنین داریم

$$\begin{aligned} m_A &= m_{A'} - m_{A''} \quad , \quad l_A = l_{A'} + l_{A''} \quad , \quad r_A = r_{A'} + r_{A''} \\ \text{s.t.} \quad m_{A'}, m_{A''}, l_{A'}, l_{A''}, r_{A'}, r_{A''} &\geq 0, \\ m_{A'} &\geq l_{A'}, m_{A''} &\geq r_{A''} \end{aligned}$$

حال با استفاده از قضیه‌های ۱.۹.۱، ۲.۹.۱ و ۴.۹.۱ خواهیم داشت

$$\begin{aligned} \tilde{A} \otimes \tilde{x} &= ((m_{A'}, l_{A'}, r_{A'}) \oplus (-m_{A''}, l_{A''}, r_{A''})) \otimes \tilde{x} \\ &\subseteq (m_{A'}, l_{A'}, r_{A'}) \otimes (m_x, l_x, r_x) \oplus (m_{A''}, l_{A''}, r_{A''}) \otimes (m_x, l_x, r_x) \\ &\simeq (m_{A'}m_x, m_{A'}l_x + l_{A'}m_x, m_{A'}r_x + r_{A'}m_x) \\ &\quad \oplus (-m_{A''}m_x, m_{A''}r_x + l_{A''}m_x, m_{A''}l_x + r_{A''}m_x) \\ &= (m_{A'}m_x - m_{A''}m_x, m_{A'}l_x + l_{A'}m_x + m_{A''}r_x + l_{A''}m_x, \\ &\quad m_{A'}r_x + r_{A'}m_x + m_{A''}l_x + r_{A''}m_x) \\ &= ((m_{A'} - m_{A''})m_x, m_{A'}l_x + l_{A'}m_x + m_{A''}r_x + l_{A''}m_x, \\ &\quad m_{A'}r_x + r_{A'}m_x + m_{A''}l_x + r_{A''}m_x) \end{aligned} \quad (۵۲.۱)$$

اکنون قرار می‌دهیم $\tilde{A}_j = (m_{A'_j}, l_{A'_j}, r_{A'_j})$ ، $\tilde{A}_j'' = (-m_{A''_j}, l_{A''_j}, r_{A''_j})$ و $\tilde{A}_j = \tilde{A}'_j \oplus \tilde{A}''_j$ برای $j = 1, \dots, p$. برای هر انتخاب $\tilde{A}'_j$ و $\tilde{A}''_j$ ، با استفاده از تقریب (۵۲.۱) و قضیه ۱.۹.۱ تقریب زیر

را برای i امین پاسخ برآورد شده $(\hat{Y}_i, i = 1, \dots, n)$ داریم

$$\begin{aligned}\hat{Y}_i &\simeq \tilde{A}_0 + \tilde{A}_1 \otimes \tilde{x}_{i1} + \dots + \tilde{A}_p \otimes \tilde{x}_{ip} \\ &= \left(m_{A_0} + \sum_{j=1}^p (m_{A'_j} - m_{A''_j}) m_{x_{ij}} \right. \\ &\quad , l_{A_0} + \sum_{j=1}^p (m_{A'_j} l_{x_{ij}} + l_{A'_j} m_{x_{ij}} + m_{A''_j} r_{x_{ij}} + l_{A''_j} m_{x_{ij}}) \\ &\quad \left. , r_{A_0} + \sum_{j=1}^p (m_{A'_j} r_{x_{ij}} + r_{A'_j} m_{x_{ij}} + m_{A''_j} l_{x_{ij}} + r_{A''_j} m_{x_{ij}}) \right) \quad (53.1)\end{aligned}$$

طرف راست معادله (۵۳.۱) را با $\tilde{Y}_i(\tilde{A}', \tilde{A}'')$ نشان می‌دهیم، که در آن $\tilde{A}'$ و $\tilde{A}''$ بردارهای p -بعدی هستند و در آن‌ها به ترتیب $\tilde{A}'_j$ و $\tilde{A}''_j$ ، j امین عضو می‌باشد. همانطور که می‌دانیم $\tilde{Y}_i$ لزوماً یک عدد فازی مثلثی نیست، اگرچه برای هر انتخاب $\tilde{A}'$ و $\tilde{A}''$ ، $\tilde{Y}_i(\tilde{A}', \tilde{A}'')$ یک برآورد مثلثی از $\tilde{Y}_i$ است که مستقل از علامت ضرایب رگرسیونی می‌باشد. برای دستیابی به بهترین انتخاب، تابع زیر را با روش‌های عددی مینیمم می‌کنیم

$$\sum_{i=1}^n d^{\vee}(\tilde{Y}_i(\tilde{A}', \tilde{A}''), \tilde{Y}_i) \quad (54.1)$$

به شرط آن که $m'_A \geq l'_A$ و $m''_A \geq r''_A$ و همچنین $m'_A, m''_A, l'_A, l''_A, r'_A, r''_A \geq 0$. با استفاده از فاصله قدرمطلق معرفی شده در (۴۴.۱) به دست می‌آوریم

$$\begin{aligned}\sum_{i=1}^n d(\tilde{Y}_i(\tilde{A}', \tilde{A}''), \tilde{Y}_i) &= \sum_{i=1}^n \left(\left| m_{Y_i} - m_{Y_i(\tilde{A}', \tilde{A}'')} \right| + \left| l_{Y_i} - l_{Y_i(\tilde{A}', \tilde{A}'')} \right| \right. \\ &\quad \left. + \left| r_{Y_i} - r_{Y_i(\tilde{A}', \tilde{A}'')} \right| \right) \quad (55.1)\end{aligned}$$

حسن‌پور و همکاران [۳۴] نشان دادند که این برآوردگر نسبت به حضور داده‌پرت در مجموعه داده‌ها استوار است.

زمانیکه همخطی بین متغیرهای توضیحی وجود دارد، معیاری برای اندازه‌گیری این همخطی، شاخص عامل تورم واریانس تعمیم‌یافته است که به صورت زیر تعریف می‌شود.

تعریف ۱.۱۰.۱. (فرخی [۶]) برای شناسایی همخطی بین متغیرهای ورودی فازی از شاخص VIF تعمیم‌یافته که به صورت زیر تعریف می‌شود، استفاده می‌کنیم

$$VIF_{G_j} = \frac{1}{1 - R_{G_j}^{\vee}}, \quad j = 1, 2, \dots, p-1 \quad (56.1)$$

که در آن

$$R_{G_j}^{\vee} = \frac{\sum_{i=1}^n d^{\vee}(\hat{X}_{ij}, \bar{\bar{X}}_j)}{\sum_{i=1}^n d^{\vee}(\hat{X}_{ij}, \bar{\bar{X}}_j)}$$

ضریب تعیین تعمیم یافته است و $\hat{X}_{ij}$ مقادیر رگرسیون شده متغیر $\tilde{X}_j$ روی دیگر متغیرهای توضیحی است. اگر $\tilde{X}_{ij} = (m_{X_{ij}}, l_{X_{ij}}, r_{X_{ij}})_T$ به ازای $i = 1, \dots, n$ و $j = 1, \dots, p$ ، آن گاه $\bar{\bar{X}}_j$ از رابطه زیر به دست می آید

$$\bar{\bar{X}}_j = \left(\frac{\sum_{i=1}^n m_{X_{ij}}}{n}, \frac{\sum_{i=1}^n l_{X_{ij}}}{n}, \frac{\sum_{i=1}^n r_{X_{ij}}}{n} \right)$$

۴.۱۰.۱ برآوردگر جریمه شده فازی

به این دلیل که عدم وجود همخطی در بین متغیرهای ورودی کمتر رخ می دهد، دیگر نمی توان تنها با مینیمم کردن معادله (۵۴.۱) به بهترین جواب دست یافت. یک رهیافت برای برطرف سازی این مشکل استفاده از برآوردگرهای انقباضی است، بنابراین همانند روش های انقباضی بیان شده در بخش ۴.۱ و با استفاده از معادله (۶.۱) به طور مشابه تابع هدف زیر را مینیمم می کنیم

$$\sum_{i=1}^n d(\tilde{Y}_i(\tilde{\mathbf{A}}', \tilde{\mathbf{A}}''), \tilde{Y}_i) + k \|\tilde{\mathbf{A}}\|^\gamma \quad (57.1)$$

که در آن k پارامتر کنترل کننده نامیده شده و $k \geq 0$ است. در این خصوص نیز در فصل های آتی بیشتر بحث خواهد شد.

فصل ۲

برآوردگر ريج استوار

در این فصل به معرفی حالت خاصی از برآوردگرهای انقباضی یعنی برآوردگر ريج استوار پرداخته، پس از توضیحی چند از برآوردگر ريج استوار، به معرفی نسخه فازی آن می‌پردازیم و کارایی این دو مدل را با مثالی کاربردی نشان می‌دهیم.

۱.۲ مقدمه

رگرسیون ريج معرفي شده توسط هورل و کنارد [۳۶]، یک ابزار مفید برای بهبود پیشگویی در مسائل رگرسیونی با همخطی بالای بین متغیرهاست. با وجود رهیافت‌های جدید مانند لاسو (LASSO)، رگرسیون کمترین زوایا^۱ (LARS) معرفي شده توسط افرون و همکاران [۲۸] و دیگر روش‌های رگرسیونی، رگرسیون ريج در مباحث زیادی به‌خصوص در شیمی‌سنجی کاربرد دارد. به‌ویژه زمانی مفید است که بتوانیم فرض کنیم همه ضرایب تقریباً دارای درجه یکسانی از بزرگی هستند (برخلاف مسائل تُنک که فرض می‌شود بیشتر ضرایب صفر باشند). فرانک و فریدمن [۳۰] عملکرد چندین برآوردگر رايج مورد استفاده در شیمی‌سنجی مانند رگرسیون کمترین توان‌های دوم جزئی^۲ (PLS) و مؤلفه‌های اصلی^۳ (PCR) را باهم مقایسه کردند. نهایتاً به این نتیجه رسیدند: «در تمام شرایطی که مطالعه شد، رگرسیون ريج بر دیگر روش‌ها غالب

^۱Least angle regression

^۲Partial least square

^۳Principle component regression

شده و به ترتيب به PLS و PCR، بسيار نزديک است».

از آن جا که رگرسيون ريج مبتنی بر کمترین توان های دوم است، نسبت به داده های ناهمگون موجود در داده ها بسيار حساس است. لارنس و مارش [۴۲] مدلی برای پیشگویی مرگ و مير در ایالات متحده آمریکا ارائه کردند. معادن زغال سنگ براساس یک مجموعه داده با ۶۴ مشاهده و ۶ ویژگی تعريف شده و نشان داده شد به وسیله داده های پرت تحت تاثیر قرار گرفته اند و همچنین دارای همخطی بالایی می باشند. آن ها نشان دادند که نسخه ی استوار رگرسيون ريج، نتایجی بهتر از رگرسيون کمترین توان های دوم معمولی (OLS) و رگرسيون ريج کلاسیک، به دست می دهد. اگرچه از دیدگاه امروزی این استوارسازی می تواند بهبود یابد. برای مجموعه داده (X, y) مدل رگرسيونی زیر را در نظر می گیریم

$$y = X\beta + \epsilon \quad (1.2)$$

که در آن $y = (y_1, \dots, y_n)^T$ یک بردار n مولفه ای شامل متغیرهای پاسخ، $X = (x_1, x_2, \dots, x_n)^T$ ماتریس طرح $n \times p$ با رتبه کامل ستونی، $\beta = (\beta_1, \beta_2, \dots, \beta_p)^T$ بردار پارامترهای مدل، و $\epsilon = (\epsilon_1, \epsilon_2, \dots, \epsilon_n)^T$ بردار n - مولفه ای خطای تصادفی است که در آن $\mathbb{E}(\epsilon) = 0$ و $\mathbb{E}(\epsilon\epsilon^T) = \sigma^2 I_n$.

۲.۲ MM - برآوردگر برای رگرسيون ريج استوار

حالت کلی برآوردگرهای ريج معمولی معرفی شده در بخش ۲.۴.۱ تحت مفروضات مدل (۱.۲) را که به صورت زیر است، بیاد آورید

$$\hat{\beta}_{\text{Ridge}} = \arg \min_{\beta} \left(\text{RSS} + k \sum_{j=1}^p \beta_j^2 \right) \quad (2.2)$$

که در آن $k \geq 0$ پارامتر تنظیم کننده بوده و میزان انقباض ضرایب را کنترل می کند. RSS همان مجموع توان های دوم مانده های معرفی شده در (۴.۱) می باشد.

برای اطمینان از استواری و کارایی تحت مدل نرمال (مدل رگرسيونی با خطای نرمال)، رهیافت MM - برآورد معرفی شده در بخش ۲.۶.۱ را بکار می گیریم (یوهای [۶۵] را ببینید). ابتدا با یک برآوردگر اولیه استوار ولی نه لزوماً کارا $\hat{\beta}_s$ شروع می کنیم، سپس با استفاده از باقی مانده های متناظر، یک برآوردگر استوار مقیاسی ($\hat{\sigma}_s$) به دست می آوریم. سپس M - برآوردگر را با استفاده از تابع زیانی که تضمین کننده کارایی مورد نظر باشد و $\hat{\sigma}_s$ مقیاسی ثابت شده، محاسبه می کنیم. لازم به ذکر است شروع تکرارها از $\hat{\beta}_s$ است. در اینجا «کارایی» به طور دلخواه به عنوان «شبهات با برآوردگر رگرسيون ريج کلاسیک در مدل نرمال» تعريف خواهد شد.

در ابتدا یادآور می شویم که یک M - برآوردگر مقیاسی برای بردار داده ی $r = (r_1, r_2, \dots, r_n)^T$

به عنوان $\hat{\sigma} = \hat{\sigma}(r)$ جواب معادله زیر است

$$\frac{1}{n} \sum_{i=1}^n \rho\left(\frac{r_i}{\sigma}\right) = \delta \quad (3.2)$$

که در آن ρ یک تابع کراندار (مارونا و همکاران [۴۸]) بوده و $\delta \in (0, 0.5)$ نشان‌دهنده نقطه فروریزش $\hat{\sigma}$ است. همانطور که در تعریف ۱.۵.۱ اشاره شد، نقطه فروریزش یک برآوردگر به بیشترین نسبت از مشاهدات گفته می‌شود که می‌تواند به‌طور دلخواه تغییر داده شود و برآوردگر همچنان با فاصله از کران مجموعه پارامتر (که می‌تواند در بی‌نهایت باشد)، کراندار باقی بماند. مارونا [۴۷] از ρ -تابع دومربعی استفاده کرد که به صورت زیر تعریف می‌شود

$$\rho_{\text{bis}}(r) = \min\{1, 1 - (1 - r^2)^3\} \quad (4.2)$$

همانطور که گفته شد، اگر $\hat{\beta}_s$ برآوردگر اولیه باشد، قرار می‌دهیم $r = r(\hat{\beta}_s)$ و اگر $\hat{\sigma}_s$ یک M-برآوردگر مقیاس برای r باشد، با جایگذاری داریم

$$\frac{1}{n} \sum_{i=1}^n \rho_{\circ}\left(\frac{r_i}{\hat{\sigma}_s}\right) = \delta \quad (5.2)$$

که در آن $\rho_{\circ}$ یک ρ -تابع کراندار و δ به دلخواه انتخاب شده است. از این رو MM-برآوردگر برای رگرسیون ريج (از این به بعد به اختصار MM-Ridge نامیده می‌شود) به وسیله معادله (۲.۲)، به صورت زیر تعریف می‌شود

$$L(\mathbf{X}, \mathbf{y}, \beta) = \hat{\sigma}_s^2 \sum_{i=1}^n \rho_1\left(\frac{r_i(\beta)}{\hat{\sigma}_s}\right) + k \|\beta\|^2 \quad (6.2)$$

که در آن $\|\cdot\|$ همان نرم اقلیدوسی و ρ_1 یک ρ -تابع کراندار دیگری است بطوری که $\rho_1 \leq \rho_{\circ}$. ضریب $\hat{\sigma}_s^2$ قبل از مجموع، برای منطبق ساختن برآوردگر با برآوردگر کلاسیک، وقتی $\rho(r) = r^2$ ، به کار گرفته می‌شود. از این رو تعریف می‌کنیم

$$\rho_{\circ}(r) = \rho_{\text{bis}}\left(\frac{r}{c_{\circ}}\right), \quad \rho_1(r) = \rho_{\text{bis}}\left(\frac{r}{c_1}\right) \quad (7.2)$$

که در آن بر اساس مارونا و همکاران [۴۸] برای کنترل استواری و کارایی، $c_{\circ} < c_1$ در نظر گرفته می‌شود.

ملاحظه ۱.۲.۲. مرکزی و مقیاسی کردن ستون‌های ماتریس $\mathbf{X}$ دو موضوع بسیار مهم در رگرسیون ريج است. از آن‌جا که برآوردگر شامل عبارت عرض از مبدا $\hat{\beta}_{\circ}$ نیز می‌باشد، مرکزی کردن تنها بر $\hat{\beta}_{\circ}$ اثر خواهد گذاشت و بر بردار ضرایب $\hat{\beta}$ اثری ندارد، از این رو مرکزی کردن ضروری نیست، اگرچه در الگوریتم MM-Ridge داده‌ها از قبل مرکزی شده است اما این تنها برای راحتی فرآیند عددی انجام شده است. همچنین ستون‌های ماتریس $\mathbf{X}$ با استفاده از میانگین یک M-برآوردگر مقیاس، مقیاسی شده‌اند.

۳.۲ معادلات نرمال موزون و الگوریتم مکرر

می‌دانیم که برآوردگر رگرسیون ريج کمترین توان‌های دوم کلاسیک در معادلات نرمال زیر صدق می‌کند

$$(\mathbf{X}^\top \mathbf{X} + k\mathbf{I}_p)\hat{\beta}_1 = \mathbf{X}^\top \mathbf{y} \quad (۸.۲)$$

که در آن $\mathbf{I}_p$ ماتریس همانی است. همچنین معادلات مشابهی برای MM-Ridge نیز وجود دارند. فرض کنید

$$\psi(r) = \rho'(r), \quad w(\beta) = \frac{\psi(r(\beta))}{r(\beta)} \quad (۹.۲)$$

برای حذف اثر واریانس، باقی‌مانده‌ها را بر واریانس اولیه ($\hat{\sigma}_s$) تقسیم می‌کنیم و بنابراین فرض می‌کنیم

$$w_i(\beta) = \left(\frac{\psi(r_i(\beta/\hat{\sigma}_s))}{r_i(\beta)/\hat{\sigma}_s} \right), \quad \mathbf{w}(\beta) = (w_1(\beta), \dots, w_n(\beta))^\top, \quad \mathbf{W}(\beta) = \text{diag}(\mathbf{w}(\beta))$$

با گرفتن مشتق از (۶.۲) نسبت به β و برابر صفر قرار دادن آن برای برآوردگر MM-Ridge معادلات نرمال به صورت زیر بدست می‌آید

$$(\mathbf{X}^\top \mathbf{W} \mathbf{X} + 2k\mathbf{I}_p)\hat{\beta} = \mathbf{X}^\top \mathbf{W} \mathbf{y} \quad (۱۰.۲)$$

با استفاده از معادله (۱۰.۲) برای برآوردگر β به جواب زیر خواهیم رسید

$$\hat{\beta} = (\mathbf{X}^\top \mathbf{W}(\beta) \mathbf{X} + 2k\mathbf{I}_n)^{-1} \mathbf{X}^\top \mathbf{W}(\beta) \mathbf{y} \quad (۱۱.۲)$$

با توجه به این که $\hat{\mathbf{y}} = \mathbf{X}\hat{\beta}$ در این صورت داریم

$$\begin{aligned} \hat{\mathbf{y}} &= \mathbf{X}(\mathbf{X}^\top \mathbf{W}(\beta) \mathbf{X} + 2k\mathbf{I}_n)^{-1} \mathbf{X}^\top \mathbf{W}(\beta) \mathbf{y} \\ &= \mathbf{H} \mathbf{y} \end{aligned} \quad (۱۲.۲)$$

که در آن $\mathbf{H}$ را ماتریس هت می‌گویند، اثر ماتریس $\mathbf{H}$ را می‌توان برای انتخاب پارامتر تنظیم‌کننده k با استفاده از یک معیار اطلاع‌بیزی^۴ (BIC) به کار برد.

از آنجایی که برای ρ -تابع انتخاب شده، $\mathbf{W}(\beta)$ یک تابع نزولی برای $|r|$ است، مشاهدات با باقی‌مانده بزرگتر وزن کمتری خواهند گرفت. همان‌طور که در روش‌های استوار معمول است، برای این معادلات نرمال موزون یک روند مکرر پیشنهاد می‌شود. در این روند با یک برآورد اولیه $\hat{\beta}_s$ شروع کرده و با استفاده از این برآوردگر بردار باقی‌مانده‌های r و وزن‌های w را به دست می‌آوریم. $\hat{\beta}_s$ را کنار می‌گذاریم و با w به دست آمده، $\hat{\beta}_1$ را از معادله (۱۰.۲) به دست می‌آوریم. مجدداً با جایگذاری $\hat{\beta}_1$ ، w را محاسبه و ادامه می‌دهیم تا جایی که تغییرات باقی‌مانده‌ها به قدر کفایت کوچک باشد.

^۴Bayesian information criteria

۱.۳.۲ انتخاب ثابت‌ها

اکنون با مسئله انتخاب δ در معادله (۵.۲) و انتخاب c_0 و c_1 در معادله (۷.۲) مواجه هستیم. برای مورد $k = 0$ یعنی M - برآوردگر استاندارد برای یک مسئله رگرسیونی با p پارامتر، انتخاب معمول $\delta = 0.5 \left(\frac{1-p}{n} \right)$ است (از این به بعد، یک برآوردگر استاندارد به معنی برآوردگری با $k = 0$ می‌باشد). این انتخاب δ می‌تواند نقطه فروریزش $\hat{\sigma}$ را ماکزیمم کند. همچنین δ می‌تواند به‌وسیله معادله زیر برای مقیاس نیز به‌دست آید

$$\frac{1}{n-p} \sum_{i=1}^n \rho \left(\frac{r_i}{\sigma} \right) = 0.5$$

یک انتخاب رایج برای $k > 0$ نیز به‌صورت زیر است

$$\delta = 0.5 \left(1 - \frac{\text{tr}(\mathbf{H})}{n} \right) \quad (13.2)$$

که در آن $\text{tr}(\cdot)$ معرف اثر ماتریس است که وابسته به k می‌باشد و زمانی که $k = 0$ ، چون $\mathbf{H}$ خودتوان است، $\text{tr}(\mathbf{H}) = \text{Rank}(\mathbf{X})$. از این رو می‌توان اثر ماتریس $\mathbf{H}$ را عدد هم‌ارزی پیشگویی^۵ نامید.

از آن‌جا که δ در مقدار ثابت شده 0.5 متوقف می‌شود، به‌منظور اطمینان از سازگاری $\hat{\sigma}$ برای داده‌های نرمال وقتی $k = 0$ است، c_0 باید به δ وابسته باشد. از این رو $c_0(\delta)$ به‌کار گرفته شده را به‌صورت زیر تعریف می‌کنیم

$$\mathbb{E} \rho \left(\frac{z}{c_0(\delta)} \right) = \delta, \quad z \sim N(0, 1)$$

که می‌تواند به‌صورت عددی محاسبه شود. به‌طور مشابه c_1 نیز با توجه به کارایی مطلوب برآوردگر استاندارد در مدل نرمال انتخاب خواهد شد.

۴.۲ شبه‌سازی

در این بخش برای ارزیابی عملکرد و مقایسه روش رگرسیونی ریدج استوار (MM-Ridge) بر حسب برآورد استوار پارامترهای رگرسیونی با نمونه متناهی، یک مطالعه شبه‌سازی را در دو طرح برای حالت $n > p$ و در دو حالت با حضور داده‌پرت در متغیر پاسخ و تبیینی و بدون داده‌پرت بررسی می‌کنیم. در این شبه‌سازی روش MM-Ridge را با روش‌های ریدج معمولی (RR) و یک MM - برآوردگر با کارایی و نقطه فروریزش بالا، مقایسه می‌کنیم. کلیه محاسبات در این مطالعه شبه‌سازی با استفاده از نرم‌افزار R انجام شده‌اند. برآوردهای ریدج معمولی با استفاده از بسته glmnet در R محاسبه شده‌اند. برای محاسبه MM - برآوردگر استوار نیز از پکیج robust در R

⁵Equivalent number of prediction

استفاده می‌کنیم. در این پکیج تابع ImRob به‌طور خودکار یک الگوریتم مناسب برای محاسبه برآورد استوار با کارایی و نقطه فروریزش بالا انتخاب می‌کند. سپس برآورد نهایی براساس یک برآورد اولیه با نقطه فروریزش بالا محاسبه می‌شود. برآوردگر اولیه از برآوردهای M و S انتخاب می‌شود. از آنجایی که روش‌های انقباضی برای مقابله با همخطی چندگانه است، لذا در این طرح داده‌ها را از یک مدل هم‌خط شبیه‌سازی می‌کنیم. مک‌دونالد و گلارینو [۴۶] و گیبونز [۳۳] روش زیر را برای ایجاد سطوح مختلف همخطی پیشنهاد دادند

$$x_{ij} = (1 - \rho^2)^{\frac{1}{2}} z_{ij} + \rho z_{ip}, \quad i = 1, \dots, n, \quad j = 1, \dots, p \quad (14.2)$$

که در آن اعداد تصادفی از توزیع نرمال و ρ^2 همبستگی بین هر دو متغیر توضیحی است. در ادامه در قالب دو طرح، ابتدا نحوه تولید داده‌های شبیه‌سازی شده را بیان و سپس نتایج کاربرد روش را ارائه می‌دهیم.

در این مطالعه شبیه‌سازی از ρ - تابع دو مربعی توکی استفاده می‌کنیم. پارامتر تنظیم‌کننده و پارامتر کنترل در ρ - تابع، $\hat{\sigma}$ ، می‌توانند براساس یک انتخاب‌گر پارامتر مبتنی بر BIC انتخاب شوند که توسط وانگ و همکاران [۶۲] پیشنهاد شده است، به این معنی که مقدار بهینه این پارامترها می‌توانند توسط مینیمم کردن معیار BIC به‌صورت زیر انتخاب شوند

$$\text{BIC} = n \log \left(\hat{\sigma}_s^2 \sum_{i=1}^n \rho_1 \left(\frac{r_i(\hat{\beta}_{MM-BR})}{\hat{\sigma}_s} \right) \right) + \text{tr}(\mathbf{H}) \log(n) \quad (15.2)$$

که در آن

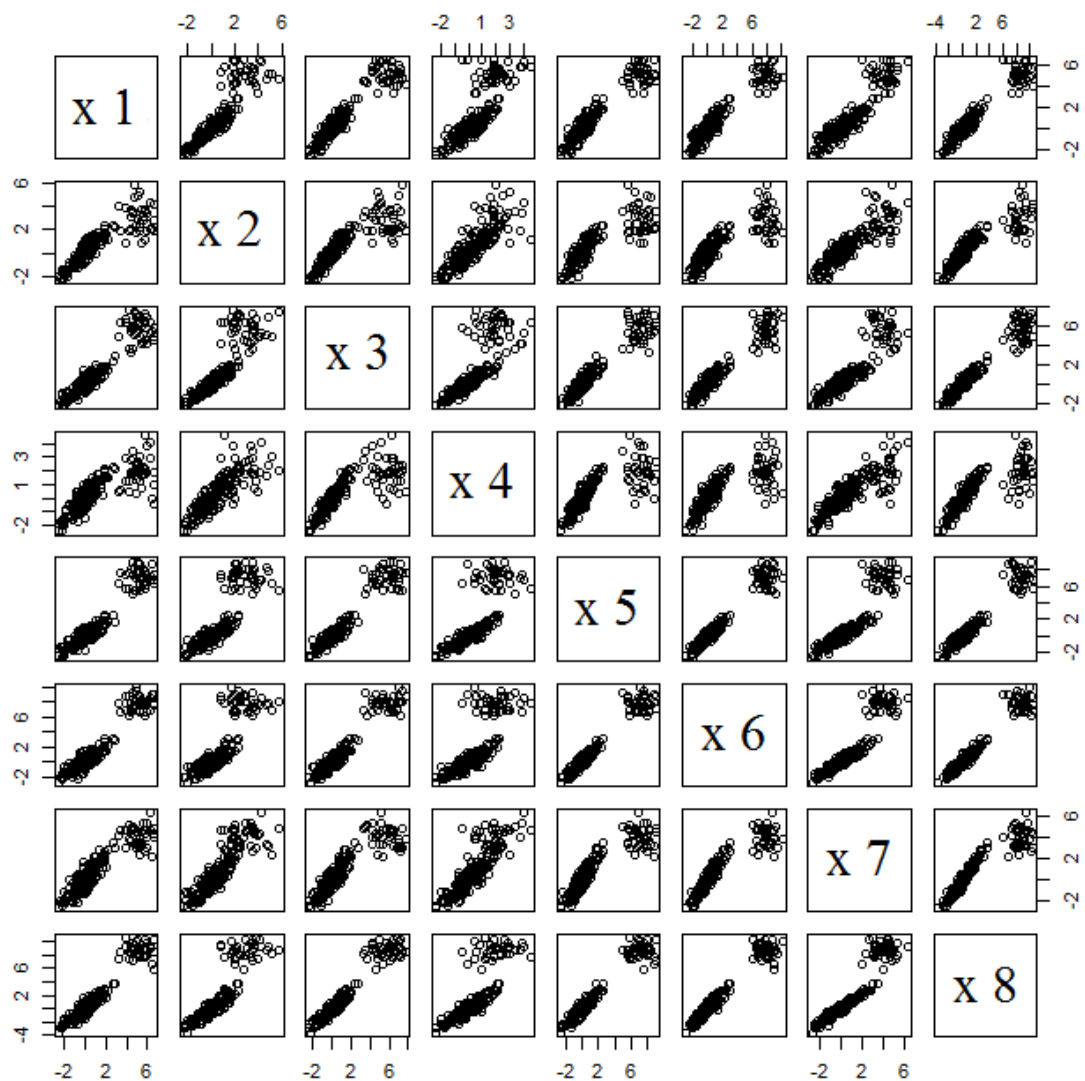
$$\mathbf{H} = \mathbf{X}(\mathbf{X}^T \mathbf{W}(\hat{\beta}_{MM-BR})\mathbf{X} + k\gamma \mathbf{W}_0(\hat{\beta}_{MM-BR}))^{-1} \mathbf{X}^T \mathbf{W}(\hat{\beta}_{MM-BR}) \quad (16.2)$$

برای ساده‌سازی محاسبات، پارامتر کنترل ρ - تابع را همانند مقدار داده شده در مارونا [۴۷] ثابت در نظر گرفته و تنها مقدار λ_n را با استفاده از معیار BIC انتخاب می‌کنیم.

طرح ۱. در این طرح نرخ آلودگی را با τ نمایش داده و طرح را در دو حالت بدون آلودگی و $\tau/2$ آلودگی را بررسی می‌کنیم. حجم نمونه با n ، تعداد پارامتر با p و تعداد داده‌های آلوده را با m نشان داده و به‌صورت $m = [\tau n]$ در نظر می‌گیریم که در آن $[\cdot]$ به مفهوم جزء صحیح است و $\beta_0 = (\beta_{01}, \beta_{02}, \dots, \beta_{0p})^T$ بردار پارامتر واقعی می‌باشد.

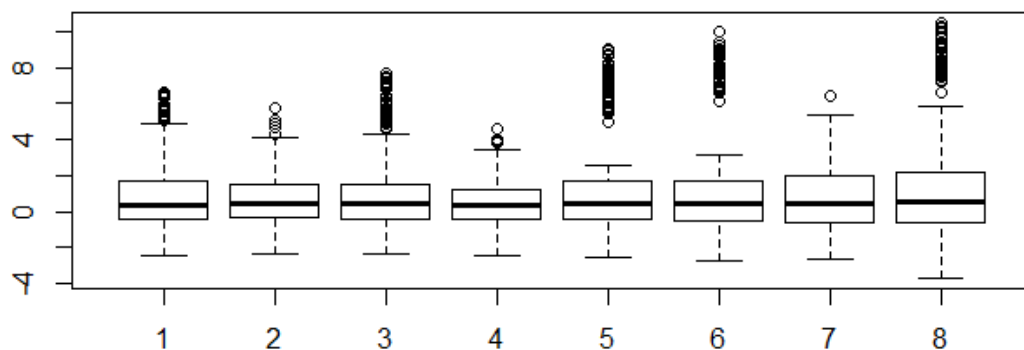
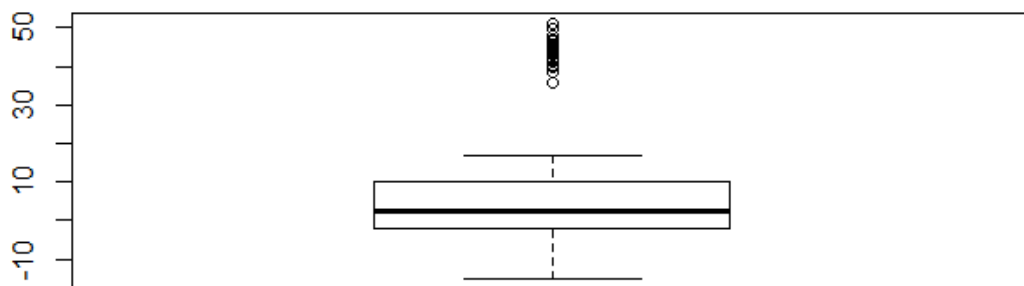
در اینجا $p = 8$ ، $n = 500$ و بردار پارامتر واقعی را، $\beta_0 = (3, 1/5, 0, 0, 2, 0, 0, 0)^T$ ، در نظر می‌گیریم. به‌طور تصادفی $n - m$ مشاهده را از $\mathbf{z}_q \sim N_p(0, \mathbf{V})$ تولید می‌کنیم که در آن $\mathbf{V}_{p \times p} = (v_{ij})$ و تعریف می‌کنیم $v_{ij} = 0.5^{|i-j|}$ ، سپس $\mathbf{x}_q$ با استفاده از رابطه (۱۴.۲) محاسبه و مقادیر متغیر پاسخ از مدل $y_q = \mathbf{x}_q^T \beta_0 + \sigma \varepsilon_q$ تولید می‌شوند که در آن $q = 1, 2, \dots, n - m$ و ε دارای توزیع t با سه درجه آزادی است. این توزیع دم‌پهن باعث ایجاد داده‌پرت در راستای محور عمودی می‌شود. مقدار σ را 0.5 در نظر می‌گیریم، این $n - m$ نقطه داده خوب تولید شده، حجم اصلی داده‌ها را شکل می‌دهند. اکنون m نقطه داده آلوده را به‌این صورت تولید

می‌کنیم: ابتدا $\mathbf{x}_z \sim N_p(\boldsymbol{\mu}, \mathbf{I})$ را تولید کرده که در آن $\boldsymbol{\mu} = (5, 3, 6, 2, 7, 8, 4, 9)^\top$. همچنین با در نظر گرفتن بردار $\boldsymbol{\beta}_1 = (1, 1, 1, 1, 1, 1, 1, 1)^\top$ ، مقادیر متغیر پاسخ برای مشاهدات آلوده را با استفاده از مدل $y_z = \mathbf{x}_z \boldsymbol{\beta}_1$ به دست می‌آوریم که در آن $z = 1, \dots, m$. اکنون با ترکیب این دو مجموعه داده، مجموعه داده اصلی $(\mathbf{x}_i, y_i)$ ، $i = 1, \dots, n$ ، را می‌سازیم. در آخر نیز مدل رگرسیونی $y = \mathbf{x}^\top \boldsymbol{\beta} + \varepsilon$ را به مجموعه داده‌ها برازش می‌دهیم و بردار پارامتر نامعلوم $\boldsymbol{\beta}$ را برآورد می‌کنیم. شکل ۱.۲ نمایشی از همخطی به وجود آمده در بین متغیرهای تبیینی را نشان می‌دهد. همچنین شکل ۲.۲ نمایشی از ایجاد داده پرت در هریک از متغیرها و متغیر پاسخ می‌باشد. در این حالت در ۲۰ درصد داده‌ها آلودگی وجود دارد.



شکل ۱.۲: ماتریس نمودار پراکنش ۸ متغیر توضیحی در برابر یکدیگر که حاکی از همخطی برای $n = 200$ مشاهده و همخطی بالای $\rho = 0.9$ ، زمانی که ۲۰ درصد مشاهدات آلوده می‌شوند (یعنی $\tau = 0.2$).

ملاحظه ۱.۴.۲. باید متذکر شویم، همان‌طور که در شکل ۱.۲ مشاهده می‌شود، همخطی اعمال‌شده بر مجموعه داده‌ها تنها بر $n - m$ داده خوب، اعمال خواهد شد. بنابراین برای اطمینان از وجود همخطی در هر حالت میانگین عدد شرطی را برای داده‌های شبیه‌سازی شده گزارش می‌کنیم.



شکل ۲.۲: نمایش نقاط نافذ بین متغیرهای تبیینی و متغیر پاسخ برای حالت $n = 200$ ، $\rho = 0.9$ و $\tau = 0.2$: بالا: نمایش داده‌پرت در راستای متغیر پاسخ، پایین: نمایش مجموعه نقاط نافذ در متغیرهای تبیینی که در آن اعداد ۱، ۲، ...، ۸ نشان‌دهنده متغیرهای تبیینی $x_1, \dots, x_p$ هستند.

برای حالت بدون داده‌پرت، n داده را از $\mathbf{z}_i \sim N_p(0, \mathbf{V})$ ، $i = 1, \dots, n$ ، تولید می‌کنیم و سپس با استفاده از رابطه (۱۴.۲) مقادیر x_i را محاسبه می‌کنیم. برای اجتناب از تولید داده‌پرت در متغیر پاسخ، ε از نرمال استاندارد تولید می‌شود.

برای مقایسه روش‌های ذکرشده، همانند آرسلان [۱۳] این شبه‌سازی را $B = 100$ بار تکرار کرده و برای هر برآوردگر میانگین توان‌های دوم خطا را برای هر ضریب به صورت

$$\text{MSE}(\beta_j) = \frac{1}{B} \sum_{i=1}^B (\hat{\beta}_{ij} - \beta_{0j})^2, \quad j = 1, \dots, p \quad (17.2)$$

محاسبه می‌کنیم. همچنین میانگین توان دوم خطای بردار ضرایب را به صورت

$$\text{MSE}(\beta) = \frac{1}{B} \sum_{i=1}^B (\hat{\beta}_i - \beta_0)^\top (\hat{\beta}_i - \beta_0) \quad (18.2)$$

و جذر میانگین خطای پیشگویی (MPE) که به صورت

$$\sqrt{\text{MPE}} = \sqrt{\frac{1}{B} \sum_{i=1}^B \sum_{j=1}^B (\hat{y}_i - y)^\top (\hat{y}_i - y)} \quad (19.2)$$

به دست می‌آید، معیار دیگری است که برای نشان دادن قدرت پیش‌گویی برآوردگرها ارائه شده است. معیار $\text{REFF}(\text{MPE})$ نسبت خطای پیشگویی برآوردگر ريج به دیگر برآوردگرها می‌باشد. همچنین میانگین عدد شرطی (MCN) نیز در هر جدول گزارش شده است که شدت هم‌خطی داده‌ها را در B تکرار شبه‌سازی شده نشان می‌دهد.

جدول ۱.۲: نتایج شبه‌سازی مربوط به ريج استوار، طرح ۱، $n = 50$ ، $\rho = 0.3$ و $\tau = 0$

RR – MM		MM		Ridge		ضرایب
MSE	mean	MSE	mean	MSE	mean	
0/0428	3/0261	0/0327	3/0132	0/2616	2/5954	$\hat{\beta}_1$
0/0426	1/5186	0/0317	1/5297	0/0667	1/3354	$\hat{\beta}_2$
0/0329	0/0223	0/0288	0/0218	0/0213	0/0244	$\hat{\beta}_3$
0/0563	-0/0055	0/0458	0/0058	0/0278	0/0205	$\hat{\beta}_4$
0/0485	1/9584	0/0368	1/9865	0/1305	1/7172	$\hat{\beta}_5$
0/0397	-0/0539	0/0365	-0/0392	0/0276	-0/0271	$\hat{\beta}_6$
0/0355	0/0230	0/0324	0/0213	0/0210	-0/0044	$\hat{\beta}_7$
0/0394	-0/0086	0/0354	-0/0149	0/0531	0/1645	$\hat{\beta}_8$
0/3377	—	0/2800	—	0/6096	—	MSE
RR – MM		MM		Ridge		
8/5878		8/4467		9/1575		$\sqrt{\text{MPE}}$
1/0663		1/0842		1/0000		REff
11/6777		11/6777		11/6777		MCN

همان‌طور که در جدول‌های ۱.۲ و ۳.۲ مشاهده می‌شود، زمانی که داده‌پرت در مجموعه داده حضور ندارد و هم‌خطی کمی بین متغیرها وجود دارد، عملکرد برآوردگرهای استوار MM، ريج و ريج استوار تقریباً مشابه است، اما با افزایش هم‌خطی در بین متغیرها، جدول‌های ۲.۲ و ۴.۲، برآوردگر استوار MM برآورد خوبی از ضرایب نداشته ولی برآوردگرهای ريج و ريج استوار

جدول ۲.۲: نتایج شبیه‌سازی مربوط به ريج استوار، طرح ۱، $n = 50$ ، $\rho = 0.99$ و $\tau = 0$

RR – MM		MM		Ridge		ضرایب
MSE	mean	MSE	mean	MSE	mean	
۱/۱۳۸۰	۲/۵۸۲۳	۱/۴۹۴۳	۳/۰۸۹۳	۴/۳۷۶۲	۰/۹۲۳۳	$\hat{\beta}_1$
۱/۰۱۵۱	۱/۴۵۸۳	۱/۴۴۹۶	۱/۷۰۱۰	۰/۵۱۳۲	۰/۸۰۸۷	$\hat{\beta}_2$
۰/۸۴۵۹	۰/۱۶۵۹	۱/۳۱۶۱	۰/۱۴۷۱	۰/۴۴۰۷	۰/۶۵۳۹	$\hat{\beta}_3$
۰/۹۷۵۷	۰/۱۶۵۲	۲/۰۹۶۲	۰/۰۳۹۴	۰/۴۴۹۶	۰/۶۶۰۰	$\hat{\beta}_4$
۱/۲۲۰۷	۱/۵۴۵۳	۱/۶۸۱۰	۱/۹۰۸۵	۱/۴۳۴۹	۰/۸۱۶۰	$\hat{\beta}_5$
۰/۸۱۹۹	–۰/۰۸۴۰	۱/۶۶۹۱	–۰/۲۶۵۲	۰/۴۱۵۹	۰/۶۲۸۱	$\hat{\beta}_6$
۰/۸۸۶۸	۰/۲۱۷۹	۱/۴۷۹۸	۰/۱۴۴۴	۰/۴۴۷۶	۰/۶۵۳۴	$\hat{\beta}_7$
۱/۲۴۳۸	۰/۳۸۱۰	۸/۳۸۲۴	–۰/۲۳۷۵	۰/۴۲۸۷	۰/۶۴۸۸	$\hat{\beta}_8$
۸/۱۴۵۹	–	۱۹/۵۶۸۴	–	۸/۵۰۶۷	–	MSE
		RR – MM		MM	Ridge	
		۸/۱۹۵۶		۸/۴۴۶۷	۱۰/۸۱۵۷	$\sqrt{\text{MPE}}$
		۱/۳۱۹۷		۱/۲۸۰۵	۱/۰۰۰۰	REff
		۱۵۰۷/۳۵۰۵		۱۵۰۷/۳۵۰۵	۱۵۰۷/۳۵۰۵	MCN

جدول ۳.۲: نتایج شبیه‌سازی مربوط به ريج استوار، طرح ۱، $n = 200$ ، $\rho = 0.3$ و $\tau = 0$

RR – MM		MM		Ridge		ضرایب
MSE	mean	MSE	mean	MSE	mean	
۰/۰۰۶۹	۲/۹۷۴۴	۰/۰۰۵۷	۲/۹۹۲۰	۰/۰۷۵۶	۲/۷۳۲۹	$\hat{\beta}_1$
۰/۰۰۸۶	۱/۴۸۹۹	۰/۰۰۸۹	۱/۵۰۰۳	۰/۰۲۲۲	۱/۳۷۴۵	$\hat{\beta}_2$
۰/۰۰۶۷	۰/۰۰۷۶	۰/۰۰۶۱	۰/۰۰۸۴	۰/۰۰۵۵	۰/۰۱۰۷	$\hat{\beta}_3$
۰/۰۰۶۹	۰/۰۰۲۴	۰/۰۰۶۵	۰/۰۰۷۹	۰/۰۰۵۰	۰/۰۱۰۰	$\hat{\beta}_4$
۰/۰۰۷۹	۱/۹۷۶۲	۰/۰۰۶۰	۱/۹۹۳۷	۰/۰۳۶۴	۱/۸۲۲۹	$\hat{\beta}_5$
۰/۰۰۸۰	–۰/۰۲۳۷	۰/۰۰۸۰	–۰/۰۲۳۸	۰/۰۰۶۳	–۰/۰۱۶۰	$\hat{\beta}_6$
۰/۰۰۸۲	–۰/۰۱۰۷	۰/۰۰۷۶	–۰/۰۱۰۲	۰/۰۰۶۱	–۰/۰۰۰۷	$\hat{\beta}_7$
۰/۰۰۹۳	۰/۰۱۰۲	۰/۰۰۷۹	۰/۰۰۲۷	۰/۰۲۰۳	۰/۱۲۴۸	$\hat{\beta}_8$
۰/۰۶۲۵	–	۰/۰۵۶۸	–	۰/۱۷۷۵	–	MSE
		RR – MM		MM	Ridge	
		۱۳/۹۵۶۶		۱۳/۸۹۱۸	۱۵/۰۷۱۸	$\sqrt{\text{MPE}}$
		۱/۰۷۹۹		۱/۰۸۴۹	۱/۰۰۰۰	REff
		۸/۸۴۱۶		۸/۸۴۱۶	۸/۸۴۱۶	MCN

جدول ۴.۲: نتایج شبیه‌سازی مربوط به ريج استوار، طرح ۱، $n = 200$ ، $\rho = 0.99$ و $\tau = 0$

RR – MM		MM		Ridge		ضرایب
MSE	mean	MSE	mean	MSE	mean	
۳/۱۷۶۶	۱/۲۸۴۳	۲/۶۰۱۹	۲/۸۲۸۶	۴/۶۷۸۳	۰/۸۳۷۳	$\hat{\beta}_1$
۰/۶۶۸۹	۰/۹۵۴۳	۴/۰۶۲۲	۱/۵۰۵۷	۰/۴۷۴۰	۰/۸۱۲۳	$\hat{\beta}_2$
۰/۷۰۰۴	۰/۶۲۰۸	۲/۷۸۳۷	۰/۱۸۰۳	۰/۶۱۸۳	۰/۷۸۵۹	$\hat{\beta}_3$
۰/۶۰۷۹	۰/۶۰۶۲	۲/۹۷۰۲	۰/۱۶۷۶	۰/۶۱۹۴	۰/۷۸۶۵	$\hat{\beta}_4$
۱/۱۶۳۱	۱/۰۰۵۱	۲/۷۳۲۶	۱/۸۶۵۸	۱/۴۱۲۴	۰/۸۱۱۹	$\hat{\beta}_5$
۰/۴۶۰۴	۰/۴۶۵۱	۳/۶۲۵۵	–۰/۵۰۷۱	۰/۵۹۸۵	۰/۷۷۳۱	$\hat{\beta}_6$
۰/۵۵۳۷	۰/۵۳۵۸	۳/۴۷۴۸	–۰/۲۱۷۸	۰/۶۰۵۲	۰/۷۷۷۵	$\hat{\beta}_7$
۳/۶۱۱۲	۰/۹۷۰۶	۲۴/۱۶۶۲	۰/۶۴۲۲	۰/۵۷۶۵	۰/۷۵۹۰	$\hat{\beta}_8$
۱۰/۹۴۲۱	–	۴۶/۴۱۷۱	–	۹/۵۸۲۶	–	۱MSE
RR – MM		MM		Ridge		ضرایب
MSE	mean	MSE	mean	MSE	mean	
۱۳/۹۳۰۶	–	۱۳/۸۹۱۸	–	۱۴/۳۴۳۱	–	$\sqrt{\text{MPE}}$
۱/۰۲۹۶	–	۱/۰۳۲۵	–	۱/۰۰۰۰	–	REff
۱۳۵۷۷/۱۰۵۸	–	۱۳۵۷۷/۱۰۵۸	–	۱۳۵۷۷/۱۰۵۸	–	MCN

جدول ۵.۲: نتایج شبیه‌سازی مربوط به ريج استوار، طرح ۱، $n = 50$ ، $\rho = 0.9$ و $\tau = 0.2$

RR – MM		MM		Ridge		ضرایب
MSE	mean	MSE	mean	MSE	mean	
۰/۲۷۸۳	۲/۷۶۴۱	۰/۲۷۰۲	۲/۷۸۶۲	۵/۷۱۷۱	۰/۶۰۹۴	$\hat{\beta}_1$
۰/۲۹۱۷	۱/۱۵۱۲	۰/۲۶۱۶	۱/۱۹۲۸	۰/۵۴۱۳	۰/۷۶۶۹	$\hat{\beta}_2$
۰/۱۵۶۴	۰/۰۶۱۱	۰/۱۶۱۸	۰/۰۴۸۶	۰/۲۷۴۰	۰/۵۲۱۶	$\hat{\beta}_3$
۰/۳۸۴۵	–۰/۳۶۶۹	۰/۴۲۵۰	–۰/۳۸۶۰	۰/۶۳۴۶	۰/۷۹۲۵	$\hat{\beta}_4$
۰/۱۳۵۶	۲/۰۰۲۷	۰/۱۴۱۵	۲/۰۸۷۱	۲/۳۴۹۱	۰/۴۶۷۸	$\hat{\beta}_5$
۰/۳۳۴۷	۰/۳۱۰۷	۰/۳۲۳۶	۰/۲۹۱۳	۰/۱۷۴۶	۰/۴۱۶۳	$\hat{\beta}_6$
۰/۲۷۱۱	–۰/۱۸۸۶	۰/۲۳۸۷	–۰/۱۹۳۱	۰/۴۷۳۷	۰/۶۸۵۸	$\hat{\beta}_7$
۰/۹۷۵۲	۰/۶۰۲۸	۱/۰۵۴۲	۰/۵۴۲۰	۰/۱۳۴۵	۰/۳۶۵۷	$\hat{\beta}_8$
۲/۸۲۷۵	–	۲/۸۷۶۶	–	۱۰/۲۹۸۹	–	MSE
RR – MM		MM		Ridge		ضرایب
MSE	mean	MSE	mean	MSE	mean	
۲۷/۴۲۸۷	–	۲۷/۵۳۷۰	–	۶۹/۹۳۹۳	–	$\sqrt{\text{MPE}}$
۲/۵۴۹۹	–	۲/۵۳۹۸	–	۱/۰۰۰۰	–	REff
۳۳۱/۹۳۸۳	–	۳۳۱/۹۳۸۳	–	۳۳۱/۹۳۸۳	–	MCN

جدول ۶.۲: نتايج شبیه‌سازی مربوط به ريج استوار، طرح ۱، $n = 50$ ، $\rho = 0.99$ و $\tau = 0.2$

RR – MM		MM		Ridge		ضرایب
MSE	mean	MSE	mean	MSE	mean	
۳/۶۷۳۲	۱/۴۶۶۳	۳/۵۷۳۷	۱/۷۱۷۱	۵/۷۵۳۶	۰/۶۰۱۸	$\hat{\beta}_1$
۱/۶۳۷۹	۰/۵۸۲۹	۱/۹۵۷۴	۰/۶۳۲۰	۰/۵۴۲۶	۰/۷۶۵۷	$\hat{\beta}_2$
۱/۰۱۴۷	۰/۴۷۵۷	۱/۳۶۸۶	۰/۳۷۳۰	۰/۲۷۵۴	۰/۵۲۲۸	$\hat{\beta}_3$
۰/۴۸۳۹	–۰/۱۹۴۰	۰/۸۳۵۱	–۰/۲۶۰۴	۰/۶۷۳۳	۰/۸۱۷۱	$\hat{\beta}_4$
۱/۵۶۶۴	۱/۴۲۶۵	۲/۲۱۷۰	۱/۴۲۸۲	۲/۳۶۴۶	۰/۴۶۲۹	$\hat{\beta}_5$
۱/۶۰۸۱	۰/۹۰۴۳	۲/۲۳۹۴	۰/۷۶۹۵	۰/۱۷۳۵	۰/۴۱۴۹	$\hat{\beta}_6$
۰/۶۵۰۱	۰/۱۱۹۸	۰/۹۱۵۰	۰/۰۰۲۱	۰/۴۷۸۹	۰/۶۸۹۹	$\hat{\beta}_7$
۳/۴۰۲۷	۱/۵۹۰۲	۵/۱۰۲۹	۱/۶۹۱۳	۰/۱۳۹۰	۰/۳۷۱۴	$\hat{\beta}_8$
۱۴/۰۳۶۹	–	۱۸/۲۰۸۹	–	۱۰/۴۰۰۸	–	MSE
		RR – MM		MM	Ridge	
		۱۳/۹۷۰۲		۱۶/۱۴۶۷	۶۹/۶۴۹۵	$\sqrt{\text{MPE}}$
		۴/۹۸۵۶		۴/۳۱۳۶	۱/۰۰۰۰	REff
		۱۵۷۱/۶۸۲۴		۱۵۷۱/۶۸۲۴	۱۵۷۱/۶۸۲۴	MCN

جدول ۷.۲: نتايج شبیه‌سازی مربوط به ريج استوار، طرح ۱، $n = 200$ ، $\rho = 0.9$ و $\tau = 0.2$

RR – MM		MM		Ridge		ضرایب
MSE	mean	MSE	mean	MSE	mean	
۰/۰۳۱۵	۲/۹۳۰۴	۰/۰۲۸۴	۳/۰۰۱۵	۵/۶۰۴۲	۰/۶۳۲۹	$\hat{\beta}_1$
۰/۱۵۱۵	۱/۲۸۱۹	۰/۱۴۷۵	۱/۳۰۲۴	۰/۴۹۶۳	۰/۷۹۶۴	$\hat{\beta}_2$
۰/۱۰۸۱	۰/۱۸۱۸	۰/۰۹۹۹	۰/۱۴۹۰	۰/۲۹۸۸	۰/۵۴۶۱	$\hat{\beta}_3$
۰/۱۷۷۵	–۰/۲۴۲۹	۰/۱۸۴۲	–۰/۲۶۷۳	۰/۶۳۰۴	۰/۷۹۲۹	$\hat{\beta}_4$
۰/۰۴۶۶	۲/۰۲۸۳	۰/۰۵۹۱	۲/۱۰۲۹	۲/۲۹۶۸	۰/۴۸۴۷	$\hat{\beta}_5$
۰/۱۱۰۶	۰/۱۸۵۵	۰/۱۰۸۰	۰/۱۵۶۶	۰/۱۸۴۲	۰/۴۲۸۶	$\hat{\beta}_6$
۰/۱۷۴۶	–۰/۲۵۰۳	۰/۱۹۱۸	–۰/۲۵۷۷	۰/۴۲۹۸	۰/۶۵۴۹	$\hat{\beta}_7$
۰/۲۸۲۵	۰/۳۴۶۳	۰/۲۶۹۳	۰/۲۹۷۸	۰/۱۴۴۲	۰/۳۷۹۳	$\hat{\beta}_8$
۱/۰۸۳۰	–	۱/۰۸۸۱	–	۱۰/۰۸۴۹	–	MSE
		RR – MM		MM	Ridge	
		۵۰/۴۰۷۵		۵۰/۴۷۲۷	۱۳۱/۹۰۴۵	$\sqrt{\text{MPE}}$
		۲/۶۱۶۸		۲/۶۱۳۴	۱/۰۰۰۰	REff
		۲۰۳/۳۱۳۷		۲۰۳/۳۱۳۷	۲۰۳/۳۱۳۷	MCN

جدول ۸.۲: نتایج شبیه‌سازی مربوط به رنج استوار، طرح $n = 200$ ، $\rho = 0.99$ و $\tau = 0.2$

RR – MM		MM		Ridge		ضرایب
<i>MSE</i>	<i>mean</i>	<i>MSE</i>	<i>mean</i>	<i>MSE</i>	<i>mean</i>	
۲/۰۴۷۹	۱/۶۹۲۹	۱/۹۴۳۶	۱/۷۹۲۶	۵/۶۹۱۸	۰/۶۱۴۴	β_1
۰/۹۸۵۷	۰/۶۱۰۳	۰/۹۵۶۶	۰/۶۷۵۵	۰/۵۱۳۲	۰/۷۸۴۴	β_2
۰/۸۰۶۵	۰/۷۴۵۵	۰/۸۱۸۸	۰/۶۹۳۳	۰/۲۸۸۸	۰/۵۳۶۸	β_3
۰/۲۳۶۵	–۰/۲۹۴۰	۰/۳۲۲۷	–۰/۳۴۲۵	۰/۶۵۶۲	۰/۸۰۹۳	β_4
۰/۵۳۹۰	۱/۴۴۱۵	۰/۵۵۹۷	۱/۵۳۹۱	۲/۳۲۲۲	۰/۴۷۶۳	β_5
۰/۸۶۷۹	۰/۸۲۹۳	۰/۸۴۲۵	۰/۷۶۵۱	۰/۱۷۹۵	۰/۴۲۳۱	β_6
۰/۱۸۸۱	–۰/۰۹۳۸	۰/۳۱۰۹	–۰/۱۶۸۰	۰/۴۶۲۰	۰/۶۷۹۰	β_7
۲/۲۲۱۹	۱/۴۴۳۴	۲/۲۲۲۳	۱/۴۲۴۴	۰/۱۴۴۰	۰/۳۷۹۰	β_8
۷/۸۹۳۵	–	۷/۹۷۸۲	–	۱۰/۲۵۷۷	–	MSE
RR – MM				MM	Ridge	
۱۹/۰۳۴۱				۲۰/۴۱۷۱	۱۳۲/۷۷۹۰	$\sqrt{\text{MPE}}$
۶/۹۷۵۸				۶/۵۰۳۳	۱/۰۰۰۰	REff
۳۵۰/۶۶۶۹				۳۵۰/۶۶۶۹	۳۵۰/۶۶۶۹	MCN

عملکرد بهتری را نشان می‌دهند. در حجم نمونه بالاتر این نتیجه به‌خوبی نشان داده شده است.

اکنون با ایجاد داده‌پرت در متغیرهای تبیینی و پاسخ، در جدول‌های ۵.۲ و ۷.۲ در حضور همخطی بالا، مشاهده می‌شود که برآوردگر ریج معمولی ظاهراً عملکرد خوبی در برآورد پارامترها دارد اما با محاسبه خطای پیشگویی، عملکرد به‌شدت ضعیف این مدل در حضور داده‌های پرت نمایان می‌شود، بخصوص در برآورد ضرایب تاثیر گذار مدل (β_1) ، همچنین با افزایش همخطی، در جدول‌های ۶.۲ و ۸.۲ نیز شاهد این نتیجه هستیم.

طرح ۲. در این طرح نحوه تولید داده همانند طرح یک می باشد، در این طرح بردار ضرایب را به صورت $\beta_0 = (\underbrace{0, \dots, 0}_{1^0}, \underbrace{2, \dots, 2}_{1^0}, \underbrace{0, \dots, 0}_{1^0}, \underbrace{2, \dots, 2}_{1^0})^T$ در نظر می گیریم. بردار μ از توزیع یکنواخت $[0, 1^0]$ تولید شده و $\beta_1 = (\underbrace{1, \dots, 1}_{4^0})$ در نظر گرفته می شود. در این مدل گروه هایی از متغیرهای موثر و غیرموثر وجود دارند که همبستگی بین آن ها یکسان است، این طرح را برای حجم های $n = 100, 300$ بررسی می کنیم.

در جدول‌های ۹.۲، ۱۰.۲، ۱۱.۲ و ۱۲.۲ عملکرد مدل ریح استوار در قیاس با ریح معمولی و M- برآوردگر معرفی شده را مقایسه می‌کنیم، همان‌طور که مشاهده می‌شود در همخطی کم بین متغیرها، برآوردگر ریح استوار نیز عملکرد مشابهی با برآوردگر استوار معمولی نشان می‌دهد اما با افزایش همخطی بین متغیرها، عملکرد خوب برآوردگر ریح استوار نمایان می‌شود.

در جدول‌های ۱۳.۲، ۱۴.۲، ۱۵.۲ و ۱۶.۲ در متغیر پاسخ و به میزان ۲۰ درصد در متغیرهای تبیینی، داده‌پرت ایجاد می‌کنیم. مشاهده می‌شود که برآوردگر ریج معمولی در حضور داده‌های پرت موجود، به شدت عملکرد ضعیفی ارائه می‌کند، و همچنین در حضور همخطی بالا،

جدول ۹.۲: نتايج شبیه‌سازی مربوط به ريج استوار، طرح ۲، $n = ۱۰۰$ ، $\rho = ۰/۳$ و $\tau = ۰$

ضرایب	RR – MM		MM		Ridge	
	MSE	mean	MSE	mean	MSE	mean
$\hat{\beta}_1$	۰/۰۴۲۶	۰/۰۰۳۶	۰/۰۴۳۹	۰/۰۰۷۶	۰/۰۷۷۶	۰/۲۰۲۷
$\hat{\beta}_2$	۰/۰۵۳۲	۰/۰۱۴۷	۰/۰۵۵۹	۰/۰۱۱۵	۰/۰۶۹۵	۰/۲۰۰۰
$\hat{\beta}_3$	۰/۰۶۰۶	۰/۰۳۳۷	۰/۰۵۸۵	۰/۰۲۹۸	۰/۰۵۰۳	۰/۱۳۲۱
$\hat{\beta}_4$	۰/۰۵۳۹	-۰/۰۵۴۰	۰/۰۴۹۱	-۰/۰۴۹۸	۰/۰۴۹۶	۰/۱۳۵۰
$\hat{\beta}_5$	۰/۰۵۱۱	۰/۰۰۶۰	۰/۰۵۴۸	-۰/۰۰۱۴	۰/۰۴۳۳	۰/۱۱۵۷
$\hat{\beta}_6$	۰/۰۵۴۶	-۰/۰۱۴۳	۰/۰۵۸۲	-۰/۰۱۷۹	۰/۰۵۳۴	۰/۱۴۹۵
$\hat{\beta}_7$	۰/۰۶۵۷	۰/۰۱۴۶	۰/۰۷۰۲	۰/۰۱۷۱	۰/۰۵۶۴	۰/۱۴۴۱
$\hat{\beta}_8$	۰/۰۷۳۷	-۰/۰۱۵۵	۰/۰۷۴۹	-۰/۰۲۳۷	۰/۰۶۲۹	۰/۱۷۱۴
$\hat{\beta}_9$	۰/۰۶۴۹	۰/۰۲۲۷	۰/۰۶۲۸	۰/۰۲۷۸	۰/۰۸۰۱	۰/۲۲۶۳
$\hat{\beta}_{10}$	۰/۰۵۵۱	۰/۰۰۴۱	۰/۰۵۴۳	-۰/۰۱۶۹	۰/۲۳۲۱	۰/۴۳۸۲
$\hat{\beta}_{11}$	۰/۰۵۴۴	۲/۰۰۲۴	۰/۰۴۷۷	۲/۰۲۲۹	۰/۳۸۹۱	۱/۴۱۷۲
$\hat{\beta}_{12}$	۰/۰۵۰۲	۱/۹۷۴۱	۰/۰۴۷۴	۱/۹۷۲۴	۰/۲۲۱۶	۱/۶۰۲۱
$\hat{\beta}_{13}$	۰/۰۵۲۳	۱/۹۸۷۳	۰/۰۴۹۵	۱/۹۸۸۱	۰/۱۴۴۵	۱/۶۷۱۵
$\hat{\beta}_{14}$	۰/۰۶۳۰	۱/۹۹۹۸	۰/۰۶۶۹	۲/۰۱۱۳	۰/۱۱۳۳	۱/۷۴۰۱
$\hat{\beta}_{15}$	۰/۰۶۴۰	۱/۹۶۵۱	۰/۰۵۷۶	۱/۹۷۸۶	۰/۱۴۱۹	۱/۶۸۷۸
$\hat{\beta}_{16}$	۰/۰۷۳۱	۲/۰۱۶۱	۰/۰۶۸۶	۲/۰۱۰۴	۰/۱۲۴۸	۱/۶۹۹۲
$\hat{\beta}_{17}$	۰/۰۷۲۴	۱/۹۳۸۴	۰/۰۶۷۳	۱/۹۵۱۳	۰/۱۳۰۱	۱/۶۹۷۴
$\hat{\beta}_{18}$	۰/۰۶۰۷	۲/۰۱۱۳	۰/۰۵۹۸	۲/۰۱۲۳	۰/۱۶۸۴	۱/۶۵۶۷
$\hat{\beta}_{19}$	۰/۰۵۵۶	۱/۹۸۵۱	۰/۰۴۸۰	۱/۹۹۶۸	۰/۲۲۶۷	۱/۶۰۲۶
$\hat{\beta}_{20}$	۰/۰۵۰۹	۲/۰۰۴۳	۰/۰۵۰۴	۲/۰۱۳۲	۰/۴۱۷۷	۱/۳۹۹۲
$\hat{\beta}_{21}$	۰/۰۵۳۹	۰/۰۱۲۶	۰/۰۵۴۰	۰/۰۰۲۴	۰/۲۲۵۴	۰/۴۳۳۵
$\hat{\beta}_{22}$	۰/۰۵۲۴	-۰/۰۱۰۴	۰/۰۵۶۵	-۰/۰۱۴۴	۰/۰۸۰۳	۰/۲۱۵۸
$\hat{\beta}_{23}$	۰/۰۴۷۱	۰/۰۱۱۳	۰/۰۴۵۰	۰/۰۰۸۳	۰/۰۴۶۶	۰/۱۴۴۸
$\hat{\beta}_{24}$	۰/۰۵۵۶	۰/۰۰۶۹	۰/۰۵۹۷	۰/۰۰۸۵	۰/۰۴۰۵	۰/۱۱۷۵
$\hat{\beta}_{25}$	۰/۰۷۲۷	-۰/۰۰۹۲	۰/۰۷۳۵	-۰/۰۰۶۶	۰/۰۷۰۶	۰/۱۷۵۳
$\hat{\beta}_{26}$	۰/۰۵۳۷	۰/۰۰۴۷	۰/۰۵۴۷	۰/۰۰۰۸	۰/۰۶۶۴	۰/۱۶۲۹
$\hat{\beta}_{27}$	۰/۰۶۸۹	۰/۰۱۰۹	۰/۰۶۷۴	۰/۰۰۵۹	۰/۰۵۶۷	۰/۱۶۸۸
$\hat{\beta}_{28}$	۰/۰۴۳۱	۰/۰۱۶۳	۰/۰۴۱۵	۰/۰۲۲۲	۰/۰۵۸۱	۰/۱۲۷۱
$\hat{\beta}_{29}$	۰/۰۵۰۴	۰/۰۳۱۷	۰/۰۵۲۱	۰/۰۱۸۵	۰/۰۸۱۴	۰/۲۱۹۰
$\hat{\beta}_{30}$	۰/۰۵۶۶	-۰/۰۰۲۵	۰/۰۵۳۲	-۰/۰۱۷۵	۰/۲۱۷۰	۰/۴۲۶۱
$\hat{\beta}_{31}$	۰/۰۴۸۳	۲/۰۱۵۱	۰/۰۴۳۶	۲/۰۲۸۶	۰/۳۹۸۵	۱/۴۱۲۵
$\hat{\beta}_{32}$	۰/۰۴۷۷	۱/۹۹۵۵	۰/۰۴۷۱	۱/۹۹۷۷	۰/۱۹۳۸	۱/۶۲۷۲
$\hat{\beta}_{33}$	۰/۰۵۳۶	۱/۹۸۴۷	۰/۰۵۴۴	۱/۹۸۴۲	۰/۱۴۷۰	۱/۶۷۷۰
$\hat{\beta}_{34}$	۰/۰۴۶۱	۲/۰۱۳۹	۰/۰۴۸۳	۲/۰۱۸۶	۰/۱۳۰۷	۱/۶۸۵۲
$\hat{\beta}_{35}$	۰/۰۶۴۱	۱/۹۶۷۱	۰/۰۶۲۷	۱/۹۷۱۵	۰/۱۵۸۰	۱/۶۷۵۱
$\hat{\beta}_{36}$	۰/۰۷۲۶	۲/۰۴۰۸	۰/۰۷۲۵	۲/۰۴۱۹	۰/۱۳۰۲	۱/۶۹۵۹
$\hat{\beta}_{37}$	۰/۰۶۵۳	۱/۹۷۷۹	۰/۰۶۳۹	۱/۹۸۵۹	۰/۱۴۱۲	۱/۶۸۳۳
$\hat{\beta}_{38}$	۰/۰۶۸۶	۲/۰۰۸۳	۰/۰۶۷۱	۲/۰۱۱۳	۰/۱۳۶۲	۱/۶۸۶۶
$\hat{\beta}_{39}$	۰/۰۵۶۹	۲/۰۰۵۷	۰/۰۵۵۷	۲/۰۱۲۷	۰/۱۳۱۵	۱/۶۸۰۹
$\hat{\beta}_{40}$	۰/۰۶۲۰	۱/۹۹۷۴	۰/۰۶۴۵	۱/۹۹۷۱	۰/۱۵۵۷	۱/۶۴۷۲
MSE	۲/۳۱۱۷	—	۲/۲۸۲۹	—	۵/۵۱۹۰	—
	RR – MM		MM		Ridge	
	۱۵/۳۹۰۸		۱۵/۲۵۳۵		۳۵/۳۱۳۰	$\sqrt{MPE}$
	۲/۲۹۴۴		۲/۳۱۵۱		۱/۰۰۰۰	REff
	۱۵۷/۹۶۰۱		۱۵۷/۹۶۰۱		۱۵۷/۹۶۰۱	MCN

جدول ۱۰.۲: نتایج شبه‌سازی مربوط به رنج استوار، طرح ۲، $n = ۱۰۰$ ، $\rho = ۰/۹۹$ و $\tau = ۰$

ضرایب	RR – MM		MM		Ridge	
	MSE	mean	MSE	mean	MSE	mean
$\hat{\beta}_1$	۱/۴۱۵۶	-۰/۱۵۹۴	۲/۰۰۶۵	۰/۰۵۱۶	۰/۷۵۳۰	۰/۸۶۵۸
$\hat{\beta}_2$	۱/۶۰۱۳	۰/۳۳۹۱	۲/۵۵۵۴	۰/۰۷۸۰	۰/۷۴۶۷	۰/۸۶۲۳
$\hat{\beta}_3$	۱/۴۸۴۶	۰/۲۳۷۷	۲/۶۷۲۹	۰/۲۰۱۲	۰/۷۴۸۶	۰/۸۶۳۴
$\hat{\beta}_4$	۱/۴۴۸۲	-۰/۱۵۸۶	۲/۲۴۳۰	-۰/۳۳۶۸	۰/۷۴۹۸	۰/۸۶۴۱
$\hat{\beta}_5$	۱/۳۵۳۲	۰/۰۴۲۹	۲/۵۰۸۱	-۰/۰۰۹۷	۰/۷۵۱۶	۰/۸۶۵۳
$\hat{\beta}_6$	۱/۵۲۱۵	-۰/۰۶۰۰	۲/۶۶۳۵	-۰/۱۲۰۸	۰/۷۵۴۴	۰/۸۶۶۷
$\hat{\beta}_7$	۱/۵۳۶۸	۰/۱۶۷۰	۳/۲۱۰۲	۰/۱۱۵۴	۰/۷۵۴۹	۰/۸۶۷۰
$\hat{\beta}_8$	۱/۹۹۶۵	-۰/۰۰۷۱	۳/۴۲۶۹	-۰/۱۶۰۳	۰/۷۵۷۰	۰/۸۶۸۱
$\hat{\beta}_9$	۱/۶۶۸۲	۰/۲۴۳۷	۲/۸۷۲۵	۰/۱۸۷۹	۰/۷۶۱۶	۰/۸۷۰۶
$\hat{\beta}_{10}$	۱/۳۷۹۵	۰/۱۳۷۶	۲/۴۸۲۳	-۰/۱۱۴۲	۰/۷۷۰۰	۰/۸۷۵۳
$\hat{\beta}_{11}$	۱/۳۶۸۲	۱/۸۳۶۸	۲/۱۸۱۰	۲/۱۵۴۹	۱/۲۴۷۶	۰/۸۸۵۰
$\hat{\beta}_{12}$	۱/۴۳۴۰	۱/۸۴۰۶	۲/۱۶۵۵	۱/۸۱۳۷	۱/۲۳۵۱	۰/۸۹۰۸
$\hat{\beta}_{13}$	۱/۴۰۶۸	۱/۸۹۹۹	۲/۲۶۴۵	۱/۹۱۹۸	۱/۲۳۱۲	۰/۸۹۲۷
$\hat{\beta}_{14}$	۱/۷۳۵۰	۱/۹۹۷۹	۳/۰۵۸۳	۲/۰۷۶۴	۱/۲۲۶۵	۰/۸۹۴۷
$\hat{\beta}_{15}$	۱/۷۴۹۶	۱/۸۶۱۹	۲/۶۳۴۵	۱/۸۵۵۰	۱/۲۳۰۱	۰/۸۹۳۱
$\hat{\beta}_{16}$	۱/۹۱۹۳	۱/۹۲۵۱	۳/۱۳۶۵	۲/۰۷۰۶	۱/۲۳۰۸	۰/۸۹۲۸
$\hat{\beta}_{17}$	۲/۱۲۶۸	۱/۷۰۹۹	۳/۰۷۷۶	۱/۶۷۰۹	۱/۲۳۴۵	۰/۸۹۱۰
$\hat{\beta}_{18}$	۱/۷۸۳۳	۱/۹۶۳۵	۲/۷۳۶۳	۲/۰۸۳۴	۱/۲۳۶۱	۰/۸۹۰۲
$\hat{\beta}_{19}$	۱/۴۵۹۶	۱/۸۷۳۶	۲/۱۹۴۵	۱/۹۷۸۴	۱/۲۳۶۷	۰/۸۸۹۸
$\hat{\beta}_{20}$	۱/۴۰۷۶	۱/۹۴۳۲	۲/۳۰۳۲	۲/۰۸۹۵	۱/۲۴۲۱	۰/۸۸۷۵
$\hat{\beta}_{21}$	۱/۵۲۹۳	۰/۱۳۰۷	۲/۴۶۸۱	۰/۰۱۶۴	۰/۷۷۷۱	۰/۸۷۹۳
$\hat{\beta}_{22}$	۱/۶۵۳۸	۰/۰۷۱۱	۲/۵۸۱۵	-۰/۰۹۷۳	۰/۷۶۵۶	۰/۸۷۳۱
$\hat{\beta}_{23}$	۱/۵۲۴۱	۰/۰۷۸۴	۲/۰۵۷۲	۰/۰۵۶۴	۰/۷۶۰۲	۰/۸۷۰۱
$\hat{\beta}_{24}$	۱/۲۹۹۱	۰/۱۲۶۱	۲/۷۲۹۰	۰/۰۵۷۳	۰/۷۵۹۹	۰/۸۶۹۹
$\hat{\beta}_{25}$	۱/۶۶۵۱	۰/۱۴۴۰	۳/۳۶۰۴	-۰/۰۴۴۶	۰/۷۵۶۱	۰/۸۶۷۸
$\hat{\beta}_{26}$	۱/۴۸۰۲	-۰/۰۳۲۲	۲/۵۰۰۳	۰/۰۰۵۳	۰/۷۵۶۹	۰/۸۶۸۲
$\hat{\beta}_{27}$	۱/۷۰۲۲	۰/۱۱۹۲	۳/۰۸۲۴	۰/۰۳۹۹	۰/۷۵۸۸	۰/۸۶۹۴
$\hat{\beta}_{28}$	۱/۱۵۶۲	۰/۲۴۳۴	۱/۸۹۶۶	۰/۱۵۰۴	۰/۷۶۲۴	۰/۸۷۱۳
$\hat{\beta}_{29}$	۱/۲۳۷۶	۰/۱۲۴۶	۲/۳۸۳۶	۰/۱۲۴۹	۰/۷۶۲۵	۰/۸۷۱۴
$\hat{\beta}_{30}$	۱/۵۵۰۱	۰/۱۵۹۲	۲/۴۳۳۷	-۰/۱۱۸۵	۰/۷۶۷۳	۰/۸۷۳۹
$\hat{\beta}_{31}$	۱/۳۴۷۲	۱/۹۶۰۴	۱/۹۹۲۶	۲/۱۹۳۲	۱/۲۴۸۷	۰/۸۸۴۴
$\hat{\beta}_{32}$	۱/۳۴۳۳	۱/۹۶۱۱	۲/۱۵۳۱	۱/۹۸۴۷	۱/۲۴۰۵	۰/۸۸۸۲
$\hat{\beta}_{33}$	۱/۷۱۱۵	۱/۹۶۸۳	۲/۴۸۹۷	۱/۸۹۳۵	۱/۲۳۵۲	۰/۸۹۰۷
$\hat{\beta}_{34}$	۱/۴۷۸۹	۱/۸۸۱۰	۲/۲۰۸۲	۲/۱۲۵۷	۱/۲۳۴۱	۰/۸۹۱۱
$\hat{\beta}_{35}$	۲/۱۳۴۴	۱/۹۸۵۲	۲/۸۶۸۴	۱/۸۰۷۴	۱/۲۴۱۳	۰/۸۸۷۹
$\hat{\beta}_{36}$	۲/۱۱۰۸	۱/۹۸۵۸	۳/۳۱۴۰	۲/۲۸۳۶	۱/۲۵۱۳	۰/۸۸۳۳
$\hat{\beta}_{37}$	۱/۶۹۴۹	۱/۸۹۲۴	۲/۹۲۲۱	۱/۹۰۴۹	۱/۲۶۷۴	۰/۸۷۶۰
$\hat{\beta}_{38}$	۱/۸۷۰۹	۱/۹۲۹۹	۳/۰۶۶۲	۲/۰۷۶۲	۱/۳۰۳۱	۰/۸۶۰۱
$\hat{\beta}_{39}$	۱/۶۳۴۲	۱/۹۱۲۷	۲/۵۴۵۸	۲/۰۸۵۶	۱/۳۷۱۳	۰/۸۳۰۴
$\hat{\beta}_{40}$	۷/۵۱۴۷	۱/۷۶۲۲	۱۷/۰۷۷۵	۱/۸۷۱۵	۱/۴۸۵۱	۰/۷۸۲۵
MSE	۶۹/۴۳۴۱	-	۱۱۸/۵۲۳۴	-	۴۰/۴۰۳۱	-
		RR – MM		MM		Ridge
		۱۳/۹۲۷۳		۱۵/۲۵۳۵		۶۸/۶۲۱۷
		۴/۹۲۷۱		۴/۴۹۸۷		۱/۰۰۰۰
		۶۱۹۳/۵۷۴۸		۶۱۹۳/۵۷۴۸		۶۱۹۳/۵۷۴۸
						$\sqrt{\text{MPE}}$
						REff
						MCN

جدول ۱۱.۲: نتايج شبیه‌سازی مربوط به ريج استوار، طرح ۲، $n = 300$ ، $\rho = 0.3$ و $\tau = 0$

RR – MM		MM		Ridge		ضرایب
MSE	mean	MSE	mean	MSE	mean	
0/0069	0/0004	0/0072	0/0018	0/0212	0/1171	$\hat{\beta}_1$
0/0073	−0/0055	0/0073	−0/0067	0/0155	0/0829	$\hat{\beta}_2$
0/0102	0/0155	0/0102	0/0151	0/0120	0/0696	$\hat{\beta}_3$
0/0090	−0/0208	0/0090	−0/0196	0/0147	0/0775	$\hat{\beta}_4$
0/0096	0/0055	0/0094	0/0049	0/0172	0/0778	$\hat{\beta}_5$
0/0071	0/0107	0/0070	0/0101	0/0120	0/0695	$\hat{\beta}_6$
0/0085	−0/0029	0/0082	−0/0034	0/0133	0/0722	$\hat{\beta}_7$
0/0104	−0/0001	0/0106	−0/0007	0/0166	0/0804	$\hat{\beta}_8$
0/0111	−0/0037	0/0113	−0/0047	0/0245	0/1125	$\hat{\beta}_9$
0/0089	0/0100	0/0091	0/0080	0/1084	0/2937	$\hat{\beta}_{10}$
0/0084	1/9959	0/0081	1/9996	0/1628	1/6337	$\hat{\beta}_{11}$
0/0107	1/9953	0/0106	1/9951	0/0580	1/8107	$\hat{\beta}_{12}$
0/0076	1/9919	0/0072	1/9935	0/0379	1/8416	$\hat{\beta}_{13}$
0/0084	1/9854	0/0085	1/9841	0/0348	1/8535	$\hat{\beta}_{14}$
0/0091	2/0092	0/0089	2/0093	0/0336	1/8611	$\hat{\beta}_{15}$
0/0104	1/9836	0/0107	1/9856	0/0362	1/8457	$\hat{\beta}_{16}$
0/0084	1/9995	0/0081	2/0010	0/0331	1/8601	$\hat{\beta}_{17}$
0/0099	2/0083	0/0099	2/0100	0/0421	1/8375	$\hat{\beta}_{18}$
0/0114	1/9902	0/0117	1/9900	0/0520	1/8176	$\hat{\beta}_{19}$
0/0079	1/9968	0/0076	2/0002	0/1586	1/6413	$\hat{\beta}_{20}$
0/0068	0/0085	0/0068	0/0060	0/1018	0/2891	$\hat{\beta}_{21}$
0/0095	0/0178	0/0094	0/0175	0/0291	0/1239	$\hat{\beta}_{22}$
0/0104	−0/0111	0/0099	−0/0101	0/0143	0/0816	$\hat{\beta}_{23}$
0/0087	0/0005	0/0088	−0/0010	0/0144	0/0849	$\hat{\beta}_{24}$
0/0091	0/0013	0/0089	−0/0001	0/0149	0/0768	$\hat{\beta}_{25}$
0/0094	0/0190	0/0096	0/0186	0/0164	0/0886	$\hat{\beta}_{26}$
0/0080	−0/0116	0/0081	−0/0128	0/0128	0/0746	$\hat{\beta}_{27}$
0/0085	0/0166	0/0087	0/0185	0/0188	0/0858	$\hat{\beta}_{28}$
0/0072	−0/0151	0/0072	−0/0150	0/0243	0/1148	$\hat{\beta}_{29}$
0/0083	0/0202	0/0079	0/0170	0/1167	0/3106	$\hat{\beta}_{30}$
0/0089	1/9947	0/0087	1/9976	0/1774	1/6201	$\hat{\beta}_{31}$
0/0091	1/9889	0/0090	1/9890	0/0578	1/8059	$\hat{\beta}_{32}$
0/0082	1/9968	0/0080	1/9984	0/0411	1/8407	$\hat{\beta}_{33}$
0/0083	2/0016	0/0085	2/0014	0/0407	1/8402	$\hat{\beta}_{34}$
0/0093	2/0027	0/0095	2/0043	0/0348	1/8486	$\hat{\beta}_{35}$
0/0097	1/9987	0/0090	1/9999	0/0394	1/8551	$\hat{\beta}_{36}$
0/0084	1/9878	0/0080	1/9904	0/0403	1/8380	$\hat{\beta}_{37}$
0/0090	2/0000	0/0088	1/9980	0/0342	1/8502	$\hat{\beta}_{38}$
0/0084	1/9939	0/0084	1/9979	0/0384	1/8437	$\hat{\beta}_{39}$
0/0095	2/0010	0/0095	1/9968	0/0518	1/8019	$\hat{\beta}_{40}$
0/3564	—	0/3536	—	1/8238	—	MSE
RR – MM		MM		Ridge		$\sqrt{\text{MPE}}$
1/85887		1/85801		34/5865		
1/8606		1/8615		1/0000		
64/2164		64/2164		64/2164		
						REff
						MCN

جدول ۱۲.۲: نتایج شبه‌سازی مربوط به رنج استوار، طرح ۲، $n = 300$ ، $\rho = 0.99$ و $\tau = 0$

ضرایب	RR – MM		MM		Ridge	
	MSE	mean	MSE	mean	MSE	mean
$\hat{\beta}_1$	۰/۲۹۸۰	-۰/۰۱۲۵	۰/۳۲۹۳	۰/۰۱۲۴	۰/۸۲۷۱	۰/۹۰۸۸
$\hat{\beta}_2$	۰/۲۱۷۸	۰/۱۱۷۴	۰/۳۳۳۸	-۰/۰۴۵۶	۰/۸۲۲۲	۰/۹۰۶۱
$\hat{\beta}_3$	۰/۳۶۰۶	۰/۱۴۹۲	۰/۴۶۶۵	۰/۱۰۲۰	۰/۸۲۲۳	۰/۹۰۶۲
$\hat{\beta}_4$	۰/۲۹۴۲	-۰/۰۳۵۸	۰/۴۱۱۰	-۰/۱۳۲۹	۰/۸۲۳۰	۰/۹۰۶۶
$\hat{\beta}_5$	۰/۲۸۳۶	۰/۱۰۳۹	۰/۴۳۰۱	۰/۰۳۳۴	۰/۸۲۶۵	۰/۹۰۸۵
$\hat{\beta}_6$	۰/۲۴۴۱	۰/۱۱۸۳	۰/۳۱۸۴	۰/۰۶۸۳	۰/۸۲۹۴	۰/۹۱۰۰
$\hat{\beta}_7$	۰/۲۹۲۷	۰/۰۷۷۶	۰/۳۷۷۰	-۰/۰۲۳۱	۰/۸۳۴۱	۰/۹۱۲۷
$\hat{\beta}_8$	۰/۳۰۱۳	۰/۰۶۷۴	۰/۴۸۵۸	-۰/۰۰۴۶	۰/۸۳۸۷	۰/۹۱۵۱
$\hat{\beta}_9$	۰/۳۸۷۶	۰/۰۵۹۲	۰/۵۱۷۲	-۰/۰۳۱۵	۰/۸۴۹۸	۰/۹۲۱۱
$\hat{\beta}_{10}$	۰/۳۹۷۸	۰/۲۵۷۲	۰/۴۱۷۵	۰/۰۵۴۱	۰/۸۶۸۷	۰/۹۳۱۱
$\hat{\beta}_{11}$	۰/۳۸۳۲	۱/۸۰۱۳	۰/۳۶۸۵	۱/۹۹۷۳	۱/۱۰۴۱	۰/۹۵۰۸
$\hat{\beta}_{12}$	۰/۳۵۵۰	۱/۸۸۰۲	۰/۴۸۵۷	۱/۹۶۶۹	۱/۰۸۴۱	۰/۹۶۱۰
$\hat{\beta}_{13}$	۰/۲۴۵۲	۱/۸۹۷۳	۰/۳۲۹۴	۱/۹۵۶۰	۱/۰۷۶۷	۰/۹۶۴۶
$\hat{\beta}_{14}$	۰/۳۰۲۲	۱/۸۳۱۸	۰/۳۸۹۴	۱/۸۹۲۴	۱/۰۷۰۷	۰/۹۶۷۸
$\hat{\beta}_{15}$	۰/۳۴۱۵	۱/۹۹۱۵	۰/۴۰۸۳	۲/۰۶۲۹	۱/۰۷۰۰	۰/۹۶۷۹
$\hat{\beta}_{16}$	۰/۳۲۲۷	۱/۸۳۵۰	۰/۴۹۰۷	۱/۹۰۲۵	۱/۰۷۲۱	۰/۹۶۶۶
$\hat{\beta}_{17}$	۰/۲۵۳۳	۲/۰۰۰۶	۰/۳۷۱۳	۲/۰۰۷۰	۱/۰۷۴۲	۰/۹۶۵۶
$\hat{\beta}_{18}$	۰/۲۵۳۰	۱/۹۷۴۵	۰/۴۵۲۸	۲/۰۶۷۵	۱/۰۷۳۵	۰/۹۶۶۰
$\hat{\beta}_{19}$	۰/۳۵۰۳	۱/۸۸۳۶	۰/۵۳۵۳	۱/۹۳۲۲	۱/۰۸۶۵	۰/۹۵۹۲
$\hat{\beta}_{20}$	۰/۳۳۳۷	۱/۸۲۹۶	۰/۳۴۹۴	۲/۰۰۱۴	۱/۱۰۸۳	۰/۹۴۸۲
$\hat{\beta}_{21}$	۰/۲۸۰۷	۰/۱۸۹۸	۰/۳۱۱۴	۰/۰۴۰۵	۰/۸۶۳۷	۰/۹۲۸۶
$\hat{\beta}_{22}$	۰/۳۲۲۹	۰/۱۵۱۱	۰/۴۲۹۴	۰/۱۱۸۳	۰/۸۴۵۸	۰/۹۱۸۹
$\hat{\beta}_{23}$	۰/۳۱۲۳	۰/۰۴۵۴	۰/۴۵۴۵	-۰/۰۶۸۳	۰/۸۳۵۹	۰/۹۱۳۶
$\hat{\beta}_{24}$	۰/۳۰۳۷	۰/۰۹۱۹	۰/۴۰۰۴	-۰/۰۰۶۹	۰/۸۳۳۹	۰/۹۱۲۷
$\hat{\beta}_{25}$	۰/۳۰۹۹	۰/۰۸۹۰	۰/۴۰۷۲	-۰/۰۰۱۰	۰/۸۳۱۰	۰/۹۱۱۱
$\hat{\beta}_{26}$	۰/۳۳۸۸	۰/۱۷۵۶	۰/۴۳۸۱	۰/۱۲۶۰	۰/۸۳۱۴	۰/۹۱۱۲
$\hat{\beta}_{27}$	۰/۲۴۸۳	۰/۰۴۶۹	۰/۳۷۰۴	-۰/۰۸۶۵	۰/۸۳۲۴	۰/۹۱۱۷
$\hat{\beta}_{28}$	۰/۳۱۳۳	۰/۱۳۶۷	۰/۳۹۹۶	۰/۱۲۵۱	۰/۸۳۸۱	۰/۹۱۵۰
$\hat{\beta}_{29}$	۰/۲۳۰۰	۰/۰۳۳۰	۰/۳۲۹۶	-۰/۱۰۱۱	۰/۸۴۱۱	۰/۹۱۶۶
$\hat{\beta}_{30}$	۰/۴۰۰۸	۰/۲۹۳۵	۰/۳۶۳۲	۰/۱۱۵۲	۰/۸۵۷۱	۰/۹۲۵۱
$\hat{\beta}_{31}$	۰/۳۵۷۱	۱/۷۷۷۹	۰/۳۹۸۱	۱/۹۸۴۱	۱/۱۱۳۸	۰/۹۴۶۲
$\hat{\beta}_{32}$	۰/۳۰۶۱	۱/۸۵۱۴	۰/۴۱۱۵	۱/۹۲۵۳	۱/۰۹۶۹	۰/۹۵۴۷
$\hat{\beta}_{33}$	۰/۲۹۰۳	۱/۹۵۰۴	۰/۳۶۵۸	۱/۹۸۹۵	۱/۰۸۹۵	۰/۹۵۸۲
$\hat{\beta}_{34}$	۰/۳۳۱۲	۱/۹۲۳۲	۰/۳۹۰۰	۲/۰۰۹۵	۱/۰۸۷۵	۰/۹۵۹۲
$\hat{\beta}_{35}$	۰/۲۹۷۲	۱/۹۳۴۵	۰/۴۳۲۵	۲/۰۲۹۲	۱/۰۹۶۷	۰/۹۵۴۵
$\hat{\beta}_{36}$	۰/۲۸۰۵	۱/۹۴۳۳	۰/۴۱۳۵	۱/۹۹۹۱	۱/۱۰۵۵	۰/۹۵۰۲
$\hat{\beta}_{37}$	۰/۲۶۳۸	۱/۸۹۴۷	۰/۳۶۴۶	۱/۹۳۵۰	۱/۱۲۵۰	۰/۹۴۰۸
$\hat{\beta}_{38}$	۰/۳۱۳۰	۱/۹۲۸۲	۰/۴۰۴۲	۱/۹۸۶۴	۱/۱۶۲۵	۰/۹۲۳۲
$\hat{\beta}_{39}$	۰/۲۹۹۷	۱/۸۶۴۳	۰/۳۸۴۸	۱/۹۸۶۱	۱/۲۳۷۶	۰/۸۸۸۶
$\hat{\beta}_{40}$	۱/۶۰۷۶	۱/۸۷۰۱	۲/۶۲۴۰	۲/۰۶۰۱	۱/۳۷۲۰	۰/۸۲۹۱
MSE	۱۳/۶۲۵۰	—	۱۸/۳۶۰۱	—	۳۹/۰۵۹۵	—
	RR – MM		MM		Ridge	
	۱۸/۴۰۳۹		۱۸/۵۸۰۱		۶۵/۳۸۸۰	$\sqrt{\text{MPE}}$
	۳/۵۵۲۹		۳/۵۱۹۲		۱/۰۰۰۰	REF
	۴۰۵۰/۷۴۱۰		۴۰۵۰/۷۴۱۰		۴۰۵۰/۷۴۱۰	MCN

جدول ۱۳.۲: نتايج شبیه‌سازی مربوط به ريج استوار، طرح ۲، $n = ۱۰۰$ ، $\rho = ۰/۹$ و $\tau = ۰/۲$

RR – MM		MM		Ridge		ضرایب
MSE	mean	MSE	mean	MSE	mean	
۰/۱۵۷۰	−۰/۰۴۸۱	۰/۱۵۶۱	−۰/۰۴۴۷	۰/۱۵۲۸	۰/۳۹۰۱	$\hat{\beta}_1$
۰/۲۳۵۱	۰/۰۷۲۷	۰/۲۵۹۳	۰/۰۵۳۵	۰/۲۲۵۹	۰/۴۷۴۲	$\hat{\beta}_2$
۰/۲۷۹۷	−۰/۰۲۹۷	۰/۳۱۸۲	−۰/۰۳۴۰	۰/۱۵۸۰	۰/۳۹۶۵	$\hat{\beta}_3$
۰/۲۷۴۸	۰/۰۵۷۶	۰/۲۹۳۸	۰/۰۵۰۷	۰/۱۹۸۴	۰/۴۴۴۴	$\hat{\beta}_4$
۰/۲۲۷۹	−۰/۰۷۲۵	۰/۲۴۱۵	−۰/۰۸۹۶	۰/۴۴۴۲	۰/۶۶۴۹	$\hat{\beta}_5$
۰/۲۵۶۸	۰/۲۱۵۹	۰/۲۸۷۹	۰/۲۰۱۳	۰/۷۴۲۷	۰/۸۶۰۲	$\hat{\beta}_6$
۰/۲۶۰۸	−۰/۱۱۱۸	۰/۲۸۵۰	−۰/۱۲۹۹	۰/۱۸۲۰	۰/۴۲۵۷	$\hat{\beta}_7$
۰/۲۹۴۷	۰/۱۵۲۷	۰/۲۸۱۸	۰/۱۴۴۲	۰/۴۵۴۱	۰/۶۷۲۲	$\hat{\beta}_8$
۰/۱۹۶۸	−۰/۰۲۰۴	۰/۲۰۶۴	−۰/۰۶۷۲	۰/۱۸۲۶	۰/۴۲۶۳	$\hat{\beta}_9$
۰/۳۳۳۱	۰/۲۴۸۷	۰/۳۷۰۵	۰/۲۴۳۲	۰/۷۱۹۳	۰/۸۴۶۱	$\hat{\beta}_{10}$
۰/۲۷۲۱	۱/۸۰۴۰	۰/۲۷۶۴	۱/۸۲۵۷	۲/۴۸۷۵	۰/۴۲۳۱	$\hat{\beta}_{11}$
۰/۲۴۲۲	۱/۹۲۱۴	۰/۲۷۸۲	۱/۹۰۳۹	۲/۰۶۹۰	۰/۵۶۲۰	$\hat{\beta}_{12}$
۰/۲۱۱۶	۲/۰۰۰۴	۰/۲۳۷۱	۲/۰۳۴۵	۲/۱۰۸۱	۰/۵۴۸۵	$\hat{\beta}_{13}$
۰/۲۸۷۳	۱/۷۸۸۶	۰/۳۳۱۸	۱/۸۱۱۸	۲/۴۴۳۵	۰/۴۳۷۱	$\hat{\beta}_{14}$
۰/۲۶۱۴	۲/۰۷۹۱	۰/۲۷۷۱	۲/۰۹۸۶	۱/۴۰۱۶	۰/۸۱۷۴	$\hat{\beta}_{15}$
۰/۲۸۱۷	۲/۰۱۲۶	۰/۲۹۶۹	۲/۰۲۶۶	۱/۶۱۱۹	۰/۷۳۱۲	$\hat{\beta}_{16}$
۰/۲۴۶۳	۱/۸۱۹۷	۰/۲۵۳۹	۱/۸۲۲۹	۲/۳۹۴۸	۰/۴۵۲۸	$\hat{\beta}_{17}$
۰/۱۹۰۰	۲/۰۸۶۴	۰/۲۱۴۱	۲/۱۱۳۹	۱/۲۳۳۷	۰/۸۹۰۷	$\hat{\beta}_{18}$
۰/۱۸۰۸	۱/۹۶۶۳	۰/۲۱۱۵	۱/۹۸۰۰	۱/۵۸۳۳	۰/۷۴۲۸	$\hat{\beta}_{19}$
۰/۲۸۴۷	۱/۹۳۷۴	۰/۳۲۹۹	۱/۹۶۷۶	۱/۳۰۹۸	۰/۸۵۶۶	$\hat{\beta}_{20}$
۰/۲۵۵۲	۰/۱۵۸۴	۰/۲۷۸۷	۰/۰۹۸۶	۰/۷۶۷۱	۰/۸۷۳۱	$\hat{\beta}_{21}$
۰/۲۱۹۰	۰/۰۲۶۱	۰/۲۳۵۰	۰/۰۱۶۰	۰/۵۱۵۷	۰/۷۱۶۵	$\hat{\beta}_{22}$
۰/۲۶۴۶	۰/۱۷۰۷	۰/۲۷۳۶	۰/۱۸۱۴	۰/۷۹۱۷	۰/۸۸۶۹	$\hat{\beta}_{23}$
۰/۲۴۷۸	۰/۰۸۰۰	۰/۲۶۷۸	۰/۰۴۶۱	۰/۷۸۷۹	۰/۸۸۵۴	$\hat{\beta}_{24}$
۰/۲۰۹۹	−۰/۰۰۵۲	۰/۲۳۰۸	−۰/۰۲۰۰	۰/۴۰۹۲	۰/۶۳۸۴	$\hat{\beta}_{25}$
۰/۲۳۸۰	۰/۰۵۲۵	۰/۲۵۷۰	۰/۰۴۴۱	۰/۳۰۷۲	۰/۵۵۳۰	$\hat{\beta}_{26}$
۰/۲۲۹۱	−۰/۱۱۲۸	۰/۲۵۲۳	−۰/۱۳۰۲	۰/۱۴۰۴	۰/۳۷۳۸	$\hat{\beta}_{27}$
۰/۲۴۱۰	۰/۱۳۲۴	۰/۲۵۶۰	۰/۱۳۴۲	۰/۷۱۹۴	۰/۸۴۵۹	$\hat{\beta}_{28}$
۰/۲۴۵۵	۰/۰۳۱۵	۰/۲۳۵۵	۰/۰۰۳۳	۰/۴۳۹۸	۰/۶۶۱۷	$\hat{\beta}_{29}$
۰/۲۶۶۸	۰/۲۲۵۷	۰/۲۸۹۶	۰/۱۹۷۶	۰/۳۰۵۵	۰/۵۵۱۶	$\hat{\beta}_{30}$
۰/۳۱۵۸	۱/۷۳۶۷	۰/۳۳۴۱	۱/۷۶۰۹	۲/۷۰۴۵	۰/۳۵۵۶	$\hat{\beta}_{31}$
۰/۱۸۹۳	۱/۹۸۵۷	۰/۲۱۴۸	۱/۹۹۵۷	۲/۲۸۸۸	۰/۴۸۷۴	$\hat{\beta}_{32}$
۰/۲۱۷۴	۲/۰۳۳۱	۰/۲۵۰۹	۲/۰۷۶۲	۲/۰۳۵۷	۰/۵۷۳۷	$\hat{\beta}_{33}$
۰/۲۵۲۵	۱/۹۲۹۰	۰/۲۶۶۴	۱/۹۳۴۲	۲/۴۳۱۵	۰/۴۴۱۰	$\hat{\beta}_{34}$
۰/۲۳۱۵	۲/۰۶۹۴	۰/۲۷۱۶	۲/۱۰۰۴	۲/۱۵۰۵	۰/۵۳۷۸	$\hat{\beta}_{35}$
۰/۱۹۹۹	۲/۰۱۵۹	۰/۲۰۹۶	۲/۰۳۰۷	۲/۰۶۵۸	۰/۵۶۳۲	$\hat{\beta}_{36}$
۰/۲۱۳۷	۱/۸۹۸۰	۰/۲۲۱۱	۱/۹۰۵۰	۲/۳۹۹۴	۰/۴۵۱۳	$\hat{\beta}_{37}$
۰/۲۵۷۲	۱/۹۹۱۷	۰/۲۹۸۷	۱/۹۹۲۰	۱/۸۸۲۸	۰/۶۲۸۳	$\hat{\beta}_{38}$
۰/۲۰۹۴	۱/۹۵۸۳	۰/۲۸۸۹	۱/۹۸۵۹	۲/۰۴۰۰	۰/۵۷۲۱	$\hat{\beta}_{39}$
۱/۰۰۱۱	۱/۷۹۰۷	۱/۲۴۴۰	۱/۷۸۱۰	۲/۳۶۹۲	۰/۴۶۱۰	$\hat{\beta}_{40}$
۱۰/۴۷۹۵	—	۱۱/۵۸۳۹	—	۴۹/۶۵۵۰	—	MSE
RR – MM		MM		Ridge		$\sqrt{\text{MPE}}$
۳۹/۷۳۰۶		۴۰/۵۷۹۳		۴۸۳/۳۴۵۲		
۱۲/۱۶۵۶		۱۱/۹۱۱۱		۱/۰۰۰۰		
۳۶۸۹/۹۵۶۸		۳۶۸۹/۹۵۶۸		۳۶۸۹/۹۵۶۸		MCN

جدول ۱۴.۲: نتایج شبه‌سازی مربوط به رنج استوار، طرح ۲، $n = ۱۰۰$ ، $\rho = ۰/۹۹$ و $\tau = ۰/۲$

ضرایب	RR – MM		MM		Ridge	
	MSE	mean	MSE	mean	MSE	mean
$\hat{\beta}_1$	۰/۷۲۹۴	۰/۱۶۶۴	۰/۹۷۶۲	۰/۲۴۹۱	۰/۱۵۰۵	۰/۳۸۷۱
$\hat{\beta}_2$	۰/۹۱۶۴	۰/۵۴۵۵	۱/۱۰۱۲	۰/۳۴۸۸	۰/۲۲۳۱	۰/۴۷۱۳
$\hat{\beta}_3$	۰/۷۰۶۸	۰/۱۶۸۶	۱/۲۳۱۹	۰/۰۳۳۴	۰/۱۵۶۱	۰/۳۹۴۱
$\hat{\beta}_4$	۰/۸۹۲۱	۰/۱۸۶۱	۱/۳۷۹۲	۰/۰۸۹۹	۰/۱۹۶۳	۰/۴۴۲۰
$\hat{\beta}_5$	۰/۹۶۴۵	۰/۲۵۶۰	۱/۵۴۶۵	۰/۲۰۹۸	۰/۴۴۳۶	۰/۶۶۴۵
$\hat{\beta}_6$	۰/۸۱۵۹	۰/۳۴۹۰	۱/۳۲۲۳	۰/۲۶۱۴	۰/۷۵۴۸	۰/۸۶۷۵
$\hat{\beta}_7$	۰/۷۳۱۹	۰/۲۴۸۲	۱/۰۸۹۷	۰/۱۴۳۳	۰/۱۷۹۶	۰/۴۲۲۸
$\hat{\beta}_8$	۰/۸۷۰۲	۰/۲۱۷۶	۱/۲۰۱۴	۰/۱۲۷۷	۰/۴۵۵۰	۰/۶۷۳۰
$\hat{\beta}_9$	۰/۶۲۸۴	۰/۳۳۱۰	۰/۸۹۴۱	۰/۱۸۸۳	۰/۱۸۰۲	۰/۴۲۳۵
$\hat{\beta}_{10}$	۱/۱۳۶۹	۰/۵۸۶۳	۱/۳۵۵۷	۰/۳۸۴۸	۰/۷۲۷۰	۰/۸۵۰۹
$\hat{\beta}_{11}$	۰/۸۳۸۳	۱/۶۶۸۹	۱/۱۷۶۸	۱/۷۸۳۵	۲/۴۹۸۴	۰/۴۱۹۶
$\hat{\beta}_{12}$	۰/۹۲۳۶	۱/۵۳۹۴	۱/۲۷۱۷	۱/۷۱۰۴	۲/۰۸۱۳	۰/۵۵۷۷
$\hat{\beta}_{13}$	۰/۸۶۴۷	۱/۷۱۷۱	۱/۱۰۵۵	۱/۸۸۹۵	۲/۱۲۰۵	۰/۵۴۴۲
$\hat{\beta}_{14}$	۰/۹۹۴۸	۱/۶۳۷۶	۱/۲۸۰۶	۱/۷۰۱۰	۲/۴۵۶۶	۰/۴۳۲۹
$\hat{\beta}_{15}$	۰/۸۳۶۹	۱/۸۴۳۹	۱/۳۵۳۰	۱/۹۲۷۶	۱/۴۱۳۹	۰/۸۱۲۰
$\hat{\beta}_{16}$	۱/۱۹۰۸	۱/۷۴۶۲	۱/۵۸۱۲	۱/۹۲۶۷	۱/۶۲۲۹	۰/۷۲۶۹
$\hat{\beta}_{17}$	۰/۸۷۷۹	۱/۸۰۵۵	۱/۱۷۴۲	۱/۹۲۲۲	۲/۴۱۱۰	۰/۴۴۷۵
$\hat{\beta}_{18}$	۰/۹۴۲۳	۱/۶۲۴۹	۱/۳۸۹۱	۱/۶۸۹۳	۱/۲۱۸۸	۰/۸۹۷۳
$\hat{\beta}_{19}$	۰/۸۴۰۷	۱/۸۰۳۶	۱/۳۰۴۹	۱/۹۸۷۸	۱/۵۸۴۰	۰/۷۴۲۴
$\hat{\beta}_{20}$	۱/۰۵۲۹	۱/۴۷۶۵	۱/۶۳۶۶	۱/۵۵۴۶	۱/۳۰۴۰	۰/۸۵۸۹
$\hat{\beta}_{21}$	۱/۳۹۲۸	۰/۶۶۷۹	۱/۸۴۷۶	۰/۵۲۵۲	۰/۸۰۳۶	۰/۸۹۴۳
$\hat{\beta}_{22}$	۱/۰۶۹۱	۰/۲۵۲۱	۱/۷۳۵۱	۰/۱۱۶۲	۰/۵۱۸۸	۰/۷۱۸۸
$\hat{\beta}_{23}$	۰/۸۳۴۹	۰/۳۱۴۴	۱/۲۷۷۹	۰/۲۳۷۳	۰/۸۱۷۲	۰/۹۰۱۵
$\hat{\beta}_{24}$	۰/۶۰۱۵	۰/۲۸۱۷	۱/۰۵۷۸	۰/۱۷۸۱	۰/۸۱۸۵	۰/۹۰۲۹
$\hat{\beta}_{25}$	۰/۶۱۲۴	۰/۲۲۱۹	۰/۸۷۸۲	۰/۱۴۶۷	۰/۴۰۸۷	۰/۶۳۸۱
$\hat{\beta}_{26}$	۰/۹۹۱۱	۰/۳۶۵۹	۱/۱۹۸۳	۰/۲۳۶۲	۰/۳۰۵۴	۰/۵۵۱۴
$\hat{\beta}_{27}$	۰/۹۹۵۷	۰/۰۷۶۴	۱/۲۲۰۱	–۰/۰۸۱۳	۰/۱۳۸۶	۰/۳۷۱۳
$\hat{\beta}_{28}$	۰/۶۶۹۹	۰/۲۵۹۸	۰/۹۵۳۱	۰/۱۶۵۵	۰/۷۳۳۹	۰/۸۵۴۹
$\hat{\beta}_{29}$	۰/۸۳۴۵	۰/۳۶۱۸	۱/۲۴۳۳	۰/۳۰۵۷	۰/۴۴۰۵	۰/۶۶۲۴
$\hat{\beta}_{30}$	۱/۴۸۰۶	۰/۷۵۸۹	۱/۶۰۲۲	۰/۶۲۱۹	۰/۳۰۴۸	۰/۵۵۰۹
$\hat{\beta}_{31}$	۱/۲۱۳۹	۱/۲۷۴۳	۱/۵۰۵۸	۱/۳۴۴۲	۲/۷۱۵۳	۰/۳۵۲۴
$\hat{\beta}_{32}$	۰/۷۰۹۸	۱/۶۷۴۱	۰/۹۹۷۱	۱/۷۹۱۶	۲/۲۹۸۶	۰/۴۸۴۲
$\hat{\beta}_{33}$	۰/۸۹۵۴	۱/۷۵۹۲	۱/۳۰۶۱	۱/۹۶۰۹	۲/۰۴۸۹	۰/۵۶۹۱
$\hat{\beta}_{34}$	۰/۶۷۰۶	۱/۹۳۸۴	۱/۱۲۰۶	۱/۹۶۳۲	۲/۴۴۳۳	۰/۴۳۷۲
$\hat{\beta}_{35}$	۰/۶۸۴۱	۱/۷۹۹۷	۰/۸۸۶۲	۱/۹۶۸۹	۲/۰۸۵۱	۰/۵۶۰۵
$\hat{\beta}_{36}$	۰/۹۲۲۶	۱/۷۱۸۰	۱/۱۴۹۳	۱/۸۱۱۳	۲/۰۷۶۳	۰/۵۵۹۶
$\hat{\beta}_{37}$	۰/۷۴۵۷	۱/۸۹۴۸	۱/۰۱۸۵	۱/۹۷۹۰	۲/۴۱۰۶	۰/۴۴۷۷
$\hat{\beta}_{38}$	۰/۸۸۷۲	۱/۸۲۵۸	۱/۲۵۱۰	۲/۰۰۳۱	۱/۸۷۵۳	۰/۶۳۱۲
$\hat{\beta}_{39}$	۰/۹۳۹۱	۱/۶۷۶۶	۱/۳۹۸۳	۱/۷۹۸۴	۲/۰۲۰۸	۰/۵۷۸۹
$\hat{\beta}_{40}$	۳/۳۲۸۳	۱/۰۸۴۰	۵/۸۷۰۶	۰/۹۳۱۶	۲/۳۴۹۷	۰/۴۶۷۴
	۳۸/۲۳۴۷	–	۵۴/۸۸۸۷	–	۴۹/۷۹۱۶	–
	RR – MM		MM		Ridge	
	۳۷/۵۱۱۵		۴۴/۳۷۰۱		۴۸۳/۲۷۵۵	$\sqrt{\text{MPE}}$
	۱۲/۸۸۳۴		۱۰/۸۹۱۹		۱/۰۰۰۰	REff
	۳۶۳۸۴/۸۳۳۳		۳۶۳۸۴/۸۳۳۳		۳۶۳۸۴/۸۳۳۳	MCN

جدول ۱۵.۲: نتايج شبیه‌سازی مربوط به ريج استوار، طرح ۲، $n = 300$ ، $\rho = 0.9$ و $\tau = 0.2$

RR – MM		MM		Ridge		ضرایب
MSE	mean	MSE	mean	MSE	mean	
0/0۳۶۵	−0/۱0۷۷	0/0۳۶۹	−0/۱0۶۲	0/۱۵۸۳	0/۳۹۷۵	$\hat{\beta}_1$
0/0۴۱۶	0/0۶۹۸	0/0۴۴۸	0/0۶۳۱	0/۲۳۴۳	0/۴۸۳۷	$\hat{\beta}_2$
0/0۴۵۸	−0/0۷۸۶	0/0۴۷0	−0/0۷۷۴	0/۱۶۴۳	0/۴0۵0	$\hat{\beta}_3$
0/0۳۹۹	−0/0۴۷۷	0/0۴۱۳	−0/0۵۱۶	0/۲0۷۹	0/۴۵۵۶	$\hat{\beta}_4$
0/0۳۷۶	0/0۴۲۹	0/0۳۹۸	0/0۴۳۹	0/۴۵۷۴	0/۶۷۵۹	$\hat{\beta}_5$
0/0۵۵۵	0/۱۱0۳	0/0۵۶۵	0/۱۱0۲	0/۷۶۳۷	0/۸۷۳۲	$\hat{\beta}_6$
0/0۵۷۶	−0/۱۴۹۳	0/0۵۹۳	−0/۱۵۴۵	0/۱۹0۹	0/۴۳۶۵	$\hat{\beta}_7$
0/0۴۶۷	0/۱۴0۹	0/0۴۷۷	0/۱۳۹۹	0/۴۷۲۵	0/۶۸۶۹	$\hat{\beta}_8$
0/0۵۷۴	−0/۱۲۴۵	0/0۵۹۳	−0/۱۲۴0	0/۱۹۱۲	0/۴۳۷0	$\hat{\beta}_9$
0/0۷۱۷	0/۱۵۵۷	0/0۶۹0	0/۱۴۹۴	0/۷۴۷۶	0/۸۶۳۸	$\hat{\beta}_{10}$
0/0۵۵۵	۱/۸۵۳۴	0/0۵۳۵	۱/۸۵۶۸	۲/۴۵۵۶	0/۴۳۳0	$\hat{\beta}_{11}$
0/0۴۸۶	۲/0۵۳۸	0/0۴۸۷	۲/0۵۷۸	۲/0۲۴۴	0/۵۷۷۳	$\hat{\beta}_{12}$
0/0۳۹۹	۱/۹۹۹۲	0/0۴0۵	۱/۹۹۹۶	۲/0۶۸۸	0/۵۶۱۸	$\hat{\beta}_{13}$
0/0۵0۹	۱/۸۷۵0	0/0۵۱۸	۱/۸۷۶۱	۲/۴۱۳۳	0/۴۴۶۶	$\hat{\beta}_{14}$
0/0۴۷۳	۲/0۸۷۷	0/0۴۷۲	۲/0۸۶۷	۱/۳۶۹۲	0/۸۳0۳	$\hat{\beta}_{15}$
0/0۳۳۷	۲/0۵۶۸	0/0۳۴۴	۲/0۵۷۹	۱/۵۸۲0	0/۷۴۲۶	$\hat{\beta}_{16}$
0/0۶0۷	۱/۸۱۵۹	0/0۶0۳	۱/۸۱۷۲	۲/۳۶۵۶	0/۴۶۲۱	$\hat{\beta}_{17}$
0/0۵۱۶	۲/۱۵۵۳	0/0۵۲۷	۲/۱۵۶۳	۱/۱۸۹0	0/۹۱0۱	$\hat{\beta}_{18}$
0/0۴۴۸	۱/۹۴۵۳	0/0۴۴۶	۱/۹۴۶۳	۱/۵۳۴۲	0/۷۶۱۶	$\hat{\beta}_{19}$
0/0۴۶۴	۱/۹۸۶۹	0/0۴۶0	۱/۹۹۱۵	۱/۲۵۴۶	0/۸۸0۶	$\hat{\beta}_{20}$
0/0۶0۲	0/۱۵۴۹	0/0۵۹۲	0/۱۵۱0	0/۷۸۹۶	0/۸۸۷۶	$\hat{\beta}_{21}$
0/0۳۹۳	−0/0۶۶۳	0/0۴0۵	−0/0۷0۴	0/۵۴۴۹	0/۷۳۷۶	$\hat{\beta}_{22}$
0/0۴۲۶	0/0۵۳0	0/0۴۳۹	0/0۵۳۴	0/۸0۹0	0/۸۹۸۵	$\hat{\beta}_{23}$
0/0۴۲۵	0/۱0۹۵	0/0۴۲۲	0/۱0۸۴	0/۸۱۱۳	0/۹000	$\hat{\beta}_{24}$
0/0۴0۶	−0/0۶۳۷	0/0۴۱۵	−0/0۶۸0	0/۴۲0۵	0/۶۴۸۱	$\hat{\beta}_{25}$
0/0۴۶۲	0/0۹۲۸	0/0۴۶۳	0/0۹۲۳	0/۳۲0۹	0/۵۶۶۱	$\hat{\beta}_{26}$
0/0۷۲۸	−0/۱۸۹۲	0/0۷۵۳	−0/۱۹۲۶	0/۱۴۵۹	0/۳۸۱۶	$\hat{\beta}_{27}$
0/0۶۲۳	0/۱۴۳۱	0/0۶۵0	0/۱۴۳۹	0/۷۵۳۷	0/۸۶۷۵	$\hat{\beta}_{28}$
0/0۴۴۸	−0/0۱۳۸	0/0۴۵۶	−0/0۱۵۹	0/۴۵۸۲	0/۶۷۶۴	$\hat{\beta}_{29}$
0/0۳۸۹	0/0۴۸۷	0/0۳۸۹	0/0۴۳۵	0/۳۲۱۸	0/۵۶۶۹	$\hat{\beta}_{30}$
0/0۶۳۴	۱/۸۵0۱	0/0۶۳۵	۱/۸۵۳۱	۲/۶۷۶۳	0/۳۶۴۱	$\hat{\beta}_{31}$
0/0۴۳۴	۲/0۱۱۷	0/0۴۵۲	۲/0۱۴۹	۲/۲۴۴۸	0/۵0۱۹	$\hat{\beta}_{32}$
0/0۴۲0	۲/0۲۵0	0/0۴۳۵	۲/0۲۶۷	۱/۹۹۳۷	0/۵۸۸۲	$\hat{\beta}_{33}$
0/0۶۶۷	۱/۸۲۴۵	0/0۶۸۲	۱/۸۲۲۹	۲/۳۹۵۶	0/۴۵۲۳	$\hat{\beta}_{34}$
0/0۹۹۵	۲/۲۳۲۹	0/۱000	۲/۲۳۵۸	۲/0۷۸۴	0/۵۵۹۸	$\hat{\beta}_{35}$
0/0۳۹۴	۱/۹۳۲0	0/0۳۹۴	۱/۹۳0۸	۲/0۲۸۶	0/۵۷۵۹	$\hat{\beta}_{36}$
0/0۳۹۵	۱/۹0۶۳	0/0۳۹۹	۱/۹0۸0	۲/۳۶۷۸	0/۴۶۱۳	$\hat{\beta}_{37}$
0/0۴۵۷	۲/0۳۳۲	0/0۴۶0	۲/0۳۳۱	۱/۸۴۴۵	0/۶۴۲۱	$\hat{\beta}_{38}$
0/0۳۴۳	۲/00۹۹	0/0۳۴0	۲/0۱۳۹	۲/00۶۴	0/۵۸۳۷	$\hat{\beta}_{39}$
0/0۲0۳۶	۲/0۲۴۴	0/۲0۶۷	۲/0۳۲۷	۲/۳۳0۲	0/۴۷۳۶	$\hat{\beta}_{40}$
۲/۱۳۷۵	—	۲/۱۶۶0	—	۴۹/۱۸۶۶	—	MSE
RR – MM		MM		Ridge		$\sqrt{\text{MPE}}$
۵۳/۲۶۹۱		۵۳/۳۹۵۱		۸۱۵/0۵۵۱		
۱۵/۳00۷		۱۵/۲۶۴۶		۱/0000		
۹۴۷/۱۴۲۹		۹۴۷/۱۴۲۹		۹۴۷/۱۴۲۹		MCN

جدول ۱۶.۲: نتایج شبیه‌سازی مربوط به ريج استوار، طرح ۲، $n = 300$ ، $\rho = 0.99$ و $\tau = 0.2$

RR – MM		MM		Ridge		ضرایب
MSE	mean	MSE	mean	MSE	mean	
0/۳۸۷۸	−0/۰۴۲۸	0/۳۳۳۱	−0/۰۹۷۴	0/۱۵۶۲	0/۳۹۴۸	$\hat{\beta}_1$
0/۲۸۷۴	0/۳۰۶۰	0/۲۸۶۷	0/۲۰۵۶	0/۲۳۲۳	0/۴۸۱۵	$\hat{\beta}_2$
0/۲۹۷۴	0/0۰۳۴	0/۳۳۱۳	−0/۰۷۸۳	0/۱۶۲۴	0/۴۰۲۶	$\hat{\beta}_3$
0/۳۴۲۴	0/۱۳۰۱	0/۳۶۴۷	0/۰۴۴۸	0/۲۰۵۸	0/۴۵۳۳	$\hat{\beta}_4$
0/۳۰۲۳	0/۱۰۶۹	0/۳۰۰۳	0/۰۵۱۱	0/۴۵۹۴	0/۶۷۷۴	$\hat{\beta}_5$
0/۳۲۴۰	0/۱۹۳۵	0/۳۳۹۵	0/۱۲۵۱	0/۷۷۹۸	0/۸۸۲۵	$\hat{\beta}_6$
0/۲۴۵۵	0/۰۸۷۷	0/۳۰۴۶	0/۰۳۲۲	0/۱۸۸۶	0/۴۳۳۹	$\hat{\beta}_7$
0/۳۱۴۷	0/۱۹۷۹	0/۳۲۵۶	0/۱۲۰۹	0/۴۷۲۶	0/۶۸۷۰	$\hat{\beta}_8$
0/۳۱۹۴	0/۱۱۰۷	0/۳۵۵۹	−0/۰۱۱۳	0/۱۸۸۹	0/۴۳۴۳	$\hat{\beta}_9$
0/۳۸۲۳	0/۳۷۰۷	0/۳۴۶۸	0/۲۷۷۱	0/۷۵۷۱	0/۸۶۹۴	$\hat{\beta}_{10}$
0/۴۱۱۶	۱/۶۰۱۸	0/۳۸۴۶	۱/۶۹۴۷	۲/۴۶۷۳	0/۴۲۹۳	$\hat{\beta}_{11}$
0/۳۲۸۱	۱/۹۸۰۷	0/۳۸۵۴	۲/۰۸۱۲	۲/۰۳۶۲	0/۵۷۳۲	$\hat{\beta}_{12}$
0/۳۰۸۰	۱/۸۳۷۹	0/۳۰۷۱	۱/۹۱۰۹	۲/۰۸۱۸	0/۵۵۷۳	$\hat{\beta}_{13}$
0/۳۶۷۰	۱/۸۱۳۹	0/۴۰۴۲	۱/۸۷۲۵	۲/۴۲۶۹	0/۴۴۲۳	$\hat{\beta}_{14}$
0/۲۴۲۷	۲/۰۰۸۱	0/۳۵۳۵	۲/۱۰۳۳	۱/۳۷۴۶	0/۸۲۷۹	$\hat{\beta}_{15}$
0/۲۹۹۰	۱/۹۱۷۵	0/۳۵۷۳	۲/۰۱۷۸	۱/۵۹۱۸	0/۷۳۸۸	$\hat{\beta}_{16}$
0/۳۰۶۲	۱/۸۱۸۲	0/۳۳۳۵	۱/۸۴۳۲	۲/۳۷۸۶	0/۴۵۷۹	$\hat{\beta}_{17}$
0/۲۶۴۶	۱/۹۷۱۵	0/۳۰۸۲	۲/۰۸۹۹	۱/۱۷۷۴	0/۹۱۵۵	$\hat{\beta}_{18}$
0/۲۹۲۰	۱/۸۷۷۹	0/۲۹۰۴	۱/۹۶۴۲	۱/۵۴۰۶	0/۷۵۹۱	$\hat{\beta}_{19}$
0/۳۱۰۶	۱/۸۰۳۴	0/۳۲۸۴	۱/۹۰۰۰	۱/۲۵۱۴	0/۸۸۲۰	$\hat{\beta}_{20}$
0/۴۶۰۰	0/۴۵۰۲	0/۴۴۱۱	0/۳۴۱۷	0/۸۲۳۸	0/۹۰۶۷	$\hat{\beta}_{21}$
0/۳۴۰۴	0/۰۹۵۴	0/۴۰۶۳	0/۰۱۵۲	0/۵۴۷۸	0/۷۳۹۵	$\hat{\beta}_{22}$
0/۳۴۳۸	0/۱۸۵۴	0/۳۵۲۷	0/۱۰۷۲	0/۸۴۵۳	0/۹۱۸۵	$\hat{\beta}_{23}$
0/۳۵۹۸	0/۱۹۴۰	0/۳۴۷۹	0/۱۲۴۵	0/۸۴۸۸	0/۹۲۰۶	$\hat{\beta}_{24}$
0/۳۹۹۱	0/۱۰۸۶	0/۴۰۹۹	0/۰۴۲۹	0/۴۲۱۸	0/۶۴۹۰	$\hat{\beta}_{25}$
0/۲۷۸۸	0/۱۳۲۶	0/۳۰۹۹	0/۰۵۷۸	0/۳۲۰۰	0/۵۶۵۳	$\hat{\beta}_{26}$
0/۲۵۰۲	−0/۰۱۴۰	0/۲۹۸۰	−0/۱۱۹۵	0/۱۴۴۰	0/۳۷۹۱	$\hat{\beta}_{27}$
0/۳۵۶۱	0/۲۵۱۸	0/۳۷۰۵	0/۱۷۷۲	0/۷۷۰۷	0/۸۷۷۲	$\hat{\beta}_{28}$
0/۳۴۳۰	0/۱۳۰۱	0/۳۴۶۶	0/۰۴۲۳	0/۴۵۸۴	0/۶۷۶۴	$\hat{\beta}_{29}$
0/۳۶۳۷	0/۳۱۹۲	0/۳۱۴۶	0/۲۱۶۸	0/۳۱۹۲	0/۵۶۴۶	$\hat{\beta}_{30}$
0/۴۲۲۸	۱/۶۹۵۱	0/۳۹۳۹	۱/۷۷۶۱	۲/۶۸۷۴	0/۳۶۰۷	$\hat{\beta}_{31}$
0/۳۲۲۰	۱/۸۷۴۹	0/۳۳۳۳	۱/۹۶۷۷	۲/۲۵۶۳	0/۴۹۸۱	$\hat{\beta}_{32}$
0/۲۴۹۱	۱/۹۱۵۷	0/۲۴۳۹	۱/۹۸۶۹	۲/۰۰۶۶	0/۵۸۳۶	$\hat{\beta}_{33}$
0/۳۳۹۴	۱/۷۵۰۴	0/۳۵۴۱	۱/۸۱۹۱	۲/۴۰۹۶	0/۴۴۷۸	$\hat{\beta}_{34}$
0/۲۷۴۴	۲/۱۳۲۸	0/۳۲۰۹	۲/۲۴۶۱	۲/۰۲۰۲	0/۵۸۰۲	$\hat{\beta}_{35}$
0/۳۲۴۹	۱/۸۲۹۴	0/۳۵۱۳	۱/۹۱۱۴	۲/۰۳۸۶	0/۵۷۲۴	$\hat{\beta}_{36}$
0/۳۱۱۰	۱/۸۱۵۲	0/۳۲۳۰	۱/۸۸۱۶	۲/۳۷۸۳	0/۴۵۷۹	$\hat{\beta}_{37}$
0/۲۸۰۶	۱/۸۸۴۳	0/۲۸۹۰	۱/۹۶۰۲	۱/۸۳۵۴	0/۶۴۵۴	$\hat{\beta}_{38}$
0/۳۰۹۰	۱/۸۴۴۵	0/۳۲۷۹	۱/۹۶۸۲	۱/۹۸۶۸	0/۵۹۰۶	$\hat{\beta}_{39}$
۱/۴۱۲۴	۱/۴۰۳۶	۱/۷۹۵۱	۱/۴۱۵۴	۲/۳۰۹۸	0/۴۸۰۳	$\hat{\beta}_{40}$
۱۴/۰۷۳۹	—	۱۵/۰۷۰۹	—	۴۹/۳۵۸۶	—	MSE
RR – MM		MM		Ridge		$\sqrt{MPE}$
۵۵/۶۳۶۹		۶۰/۱۰۷۳		۸۱۸/۱۵۶۸		
۱۴/۰۰۵۳		۱۳/۶۱۱۶		۱/۰۰۰۰		
۳۷۹۲/۷۷۰۳		۳۷۹۲/۷۷۰۳		۳۷۹۲/۷۷۰۳		MCN

برآوردگر ریج استوار از برآوردگر استوار معمولی، عملکرد بهتری ارائه می‌دهد. همچنین با بالا بردن میزان همخطی داده‌ها نیز این روند عملکرد به‌خوبی مشاهده می‌شود.

۵.۲ تحلیل داده‌های چربی بدن

در این بخش رفتار برآوردگر ریج استوار معرفی شده را در قالب یک مثال واقعی مورد ارزیابی قرار می‌دهیم. در این مثال مجموعه داده‌های Vo2^۶ را از رساله نوروزی‌راد [۱۰] مورد استفاده قرار می‌دهیم که همخطی بین متغیرهای تبیینی نیز بررسی شده است.

۱.۵.۲ معرفی داده‌های Vo2

مجموعه داده‌های Vo2 مربوط به مطالعات تعیین تناسب اندام فیزیکی است. این مجموعه داده شامل اطلاعات مربوط به ۲۳۳ مشاهده است. در این مطالعه، در مجموع ۷ متغیر وجود دارند که یکی از آن‌ها به‌عنوان متغیر وابسته و ۶ متغیر تبیینی دارای همخطی و همراه با داده‌های پرت در نظر گرفته شده است. در این مجموعه داده، متغیر وابسته (Vo2) ماکزیمم مصرف اکسیژنی است که با آزمون روی تردمیل اندازه‌گیری شده است (واحد محاسبه شده ml/kg/min است).

مجموعه متغیرهای توضیحی شامل سن (AGE) برحسب سال، ماکزیمم ضربان قلب (MHR) برحسب ضربه بر دقیقه، مدت زمان آزمون تردمیل (DUR) برحسب ثانیه، ماکزیمم فشار خون سیاستولیک روی آزمون تردمیل (MXS) برحسب میلی‌متر جیوه، ماکزیمم فشار خون دیاستولیک روی آزمون تردمیل (MXD) برحسب میلی‌متر جیوه و اختلال هوازی کاربردی (FAI) برحسب درصد نسبی به سن و جنسیت است.

برای برآورد مقادیر ضرایب رگرسیونی براساس معادله (۱۱.۲) نیاز به انتخاب پارامتر ریج (k) بهینه داریم. از این رو برای انتخاب k بهینه از معیار BIC معرفی شده در (۱۵.۲) استفاده می‌کنیم.

برای مقایسه برآوردگر ریج استوار با برآوردگرهای ریج معمولی و M-برآوردگر استوار در این داده‌ها از معیار میانگین خطای پیشگویی استفاده می‌کنیم. بدین منظور داده‌ها را به دو بخش آموزشی^۷ ((x^{Train}, y^{Train})) و آزمون^۸ ((x^{Test}, y^{Test})) تقسیم می‌کنیم. ۷۰ درصد داده‌ها را به‌عنوان داده‌های آموزشی و ۳۰ درصد بقیه را به‌عنوان داده‌های آزمون در نظر می‌گیریم. بر اساس داده‌های آموزشی، برآورد پارامترهای مدل را به‌صورت $\hat{\beta}^{Train}$ به‌دست می‌آوریم و از این

^۶این مجموعه داده در وب‌سایت <http://people.umass.edu/be640/yr2004/resources/data2002/VO2.txt> در دسترس است.

^۷Train

^۸Test

برآوردگر برای پیش‌گویی متغیر پاسخ در داده‌های آزمون بهره می‌گیریم. به عبارت دیگر،

$$\hat{\mathbf{y}}^{\text{Test}} = \mathbf{x}^{\text{Test}} \hat{\boldsymbol{\beta}}^{\text{Train}} \quad (20.2)$$

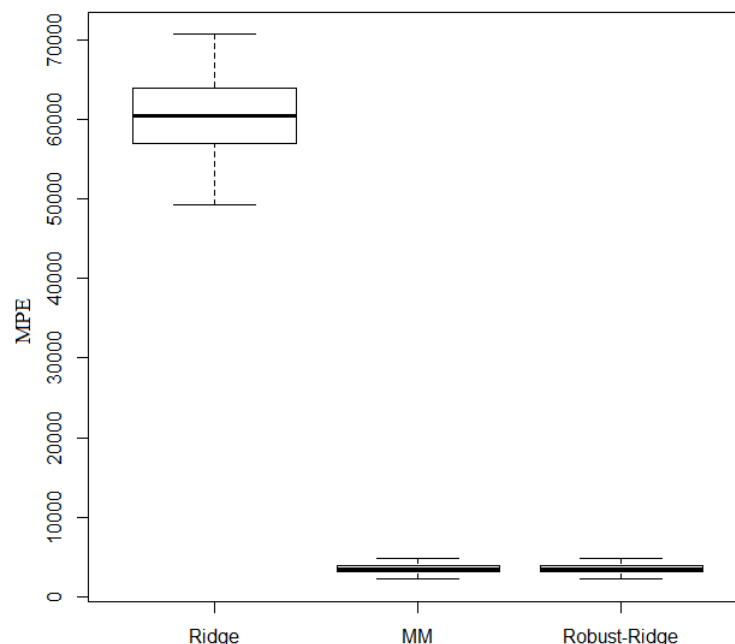
این رویه را ۱۰۰ بار تکرار نموده یعنی به تعداد ۱۰۰ بار، داده‌ها را به صورت تصادفی و با جایگذاری تقسیم می‌کنیم. اگر به‌ازای بردار دلخواه $\mathbf{a} = (a_1, \dots, a_p)^T$ ، تعریف کنیم $\|\mathbf{a}\| = \sum_{i=1}^n a_i^2$ ، آن‌گاه میانگین خطای پیش‌گویی را به صورت

$$\text{MPE} = \frac{1}{100} \|\hat{\mathbf{y}}^{\text{Test}} - \mathbf{y}^{\text{Test}}\| \quad (21.2)$$

به دست می‌آوریم.

جدول ۱۷.۲: نتایج مربوط به برآوردگر ريج استوار برای داده‌های Vo2

ريج استوار	M – برآودگر	ريج	
۵۹/۰۸۶۶	۵۸/۷۷۵	۲۴۵/۷۷۶	$\sqrt{\text{MPE}}$



شکل ۳.۲: نمودار جعبه‌ای مربوط به خطاهای پیش‌گویی برای ريج استوار

همان‌طور که در جدول ۱۷.۲ مشاهده می‌شود، با محاسبه معیار جذر (MPE) برای برآوردگرهای گفته شده، در حضور داده‌های پرت برآوردگر ريج معمولی عملکرد به شدت ضعیفی نشان می‌دهد

اما برآوردگرهای استوار قادر به کنترل این داده‌های پرت موجود در مجموعه داده هستند و عملکرد بهتری دارند. دلیل شباهت برآوردگرهای استوار می‌تواند به دلیل مکرر بودن روش M-برآوردگر استفاده شده در این محاسبات باشد.

۶.۲ برآوردگر ریج فازی

۱.۶.۲ برآوردگر ریج فازی معمولی

تاکنون به معرفی برآوردگر ریج پرداختیم، حال اگر شرط دقیق بودن داده‌ها برقرار نباشد، دیگر قادر به استفاده از برآوردگر ریج معمولی نیستیم و باید برآوردگر جدیدی را معرفی کنیم. مدل رگرسیون زیر را با ورودی و خروجی فازی و همچنین ضرایب رگرسیون فازی در نظر بگیرید

$$\tilde{Y}_i = \tilde{A}_0 \oplus (\tilde{A}_1 \otimes \tilde{X}_{i1}) \oplus \cdots \oplus (\tilde{A}_p \otimes \tilde{X}_{ip}), \quad i = 1, \dots, n \quad (22.2)$$

که در آن $\tilde{Y}_i$ متغیر پاسخ فازی i ام، $\tilde{X}_{ij}$ متغیر توضیحی فازی برای مشاهده i ام و متغیر j ام و $\tilde{A}_j$ ضریب رگرسیونی فازی برای متغیر j ام است. برای برآورد ضرایب مدل (۲۲.۲) با حاصلضرب عدد فازی در عدد فازی مواجه هستیم. همان‌طور که در بخش ۳.۱۰.۱ توضیح داده شد این حاصلضرب وابسته به علامت مقادیر فازی ضرایب است. این پارامترها مجهول هستند و باید برآورد شوند، اما اطلاعاتی درباره مثبت و منفی بودن این پارامترها نداریم که بسته به مثبت و منفی بودن این پارامترها نتیجه این حاصلضرب متفاوت است. برای دستیابی به بهترین انتخاب، تابع زیر را با روش‌های عددی به کمک نرم‌افزار Mathematica 11 مینیمم می‌کنیم

$$\sum_{i=1}^n d^2(\tilde{Y}_i(\tilde{A}', \tilde{A}''), \tilde{Y}_i) \quad (23.2)$$

به شرط آن که $m'_A \geq r''_A$ و $m''_A \geq l'_A$ و همچنین $m'_A, m''_A, l'_A, l''_A, r'_A, r''_A \geq 0$ ، با استفاده از فاصله موزون معرفی شده در (۳۷.۱) داریم

$$\begin{aligned} d^2(\tilde{Y}_i(\tilde{A}', \tilde{A}''), \tilde{Y}_i) = & [m_0 + \sum_{j=1}^p (m'_j - m''_j)x_{ij} - y_i]^2 + \frac{1}{3}(a_0 + \sum_{j=1}^p (m'_j - m''_j)x_{ij} - y_i) \\ & \times [(r_0 + \sum_{j=1}^p (m'_j r_{x_{ij}} + r'_j x_{ij} + m''_j l_{x_{ij}} + r''_j x_{ij}) - r_{y_i})] \\ & - (l_0 + \sum_{j=1}^p (m'_j l_{x_{ij}} + l'_j x_{ij} + m''_j r_{x_{ij}} + l''_j x_{ij}) - l_{y_i}) \\ & + \frac{1}{12} [(r_0 + \sum_{j=1}^p (m'_j r_{x_{ij}} + r'_j x_{ij} + m''_j l_{x_{ij}} + r''_j x_{ij}) - r_{y_i})^2 \\ & + (l_0 + \sum_{j=1}^p (m'_j l_{x_{ij}} + l'_j x_{ij} + m''_j r_{x_{ij}} + l''_j x_{ij}) - l_{y_i})^2] \quad (24.2) \end{aligned}$$

به این دلیل که عدم وجود همخطی در بین متغیرهای ورودی کمتر رخ می‌دهد، دیگر نمی‌توان تنها با مینیمم کردن معادله (۲۳.۲) به بهترین جواب دست یافت. بنابراین همانند مدل رگرسیون ریج بیان شده در بخش ۲.۴.۱ و با استفاده از معادله (۲.۲) به طور مشابه داریم

$$\sum_{i=1}^n d^2(\tilde{Y}_i(\tilde{A}', \tilde{A}''), \tilde{Y}_i) + k \|\tilde{A}\|^2 \quad (25.2)$$

که در آن k ضریب ریج نامیده شده و $k \geq 0$ است. فاصله استفاده شده فوق فاصله موزون معرفی شده در بخش ۲.۹.۱ می‌باشد، مسئله رگرسیون ریج (۲.۲) می‌تواند به عنوان مسئله بهینه‌سازی تعریف شده در زیر حل شود

$$\begin{aligned} \text{Minimize} \quad & Q(A_0, A_1, \dots, A_p) = k(\|\tilde{A}_0\|^2 + \|\tilde{A}\|^2) + \sum_{i=1}^n d^2(\tilde{Y}_i(\tilde{A}', \tilde{A}''), \tilde{Y}_i), \\ \text{s.t.} \quad & m'_j \geq l'_j, m''_j \geq r''_j \quad j = 1, \dots, p, \\ & m_0 \in \mathbb{R}, l_0, r_0 \geq 0, \text{ and } m'_j, m''_j, l'_j, l''_j, r'_j, r''_j \geq 0, \quad j = 1, \dots, p \end{aligned} \quad (26.2)$$

که در آن

$$\|A\|^2 = \sum_{j=1}^p \|A_j\|^2 = \sum_{j=1}^p \|(m'_j - m''_j, l'_j + l''_j, r'_j + r''_j)\|^2 \quad (27.2)$$

بنابراین با استفاده از (۲۴.۲) و تعریف $\tilde{A}_0 = (m_{A_0}, l_{A_0}, r_{A_0})_T$ به مسئله بهینه‌سازی محدب زیر برای مدل (۲۶.۲) می‌رسیم

$$\begin{aligned} \text{Minimize} \quad & Q(A_0, A_1, \dots, A_p) = k \left(m_0^2 + \frac{1}{\alpha} m_0 (r_0 - l_0) + \frac{1}{\alpha} (l_0^2 + r_0^2) \right) \\ & + \sum_{j=1}^p \left[(m'_j - m''_j)^2 + \frac{1}{\alpha} (m'_j - m''_j) [(r'_j + r''_j) - (l'_j + l''_j)] \right. \\ & \left. + \frac{1}{\alpha} [(l'_j + l''_j)^2 + (r'_j + r''_j)^2] \right] + \sum_{i=1}^n (m_0 + \sum_{j=1}^p (m'_j - m''_j) x_{ij} - y_i)^2 \\ & + \frac{1}{\alpha} \sum_{i=1}^n \left\{ \left(m_0 + \sum_{j=1}^p (m'_j - m''_j) x_{ij} - y_i \right) \right. \\ & \quad \times \left[\left(r_0 + \sum_{j=1}^p (m'_j r_{x_{ij}} + r'_j x_{ij} + m''_j l_{x_{ij}} + r''_j x_{ij}) - r_{y_i} \right) \right. \\ & \quad \left. \left. - \left(l_0 + \sum_{j=1}^p (m'_j l_{x_{ij}} + l'_j x_{ij} + m''_j r_{x_{ij}} + l''_j x_{ij}) - l_{y_i} \right) \right] \right\} \\ & + \frac{1}{\alpha} \sum_{i=1}^n \left[\left(r_0 + \sum_{j=1}^p (m'_j r_{x_{ij}} + r'_j x_{ij} + m''_j l_{x_{ij}} + r''_j x_{ij}) - r_{y_i} \right)^2 \right. \\ & \quad \left. + \left(l_0 + \sum_{j=1}^p (m'_j l_{x_{ij}} + l'_j x_{ij} + m''_j r_{x_{ij}} + l''_j x_{ij}) - l_{y_i} \right)^2 \right] \end{aligned} \quad (28.2)$$

مشروط به اینکه

$$\begin{aligned} m'_j &\geq l'_j, m''_j \geq r''_j, \quad j = 1, \dots, p, \\ m_o &\in \mathbb{R}, l_o, r_o \geq 0, \\ m'_j, m''_j, l'_j, l''_j, r'_j, r''_j &\geq 0, \quad j = 1, \dots, p \end{aligned} \quad (29.2)$$

با مینیمم کردن مسئله بهینه‌سازی (۲۸.۲) تحت قیود (۲۹.۲)، برآورد ضرایب مدل رگرسیونی محاسبه می‌گردد. برای مینیمم کردن این معادله از نرم افزار Mathematica 11 استفاده می‌کنیم.

۲.۶.۲ شاخص‌های نیکویی برازش مدل

خطای برآورد

برای ارزیابی عملکرد مدل رگرسیونی فازی، کیم و بیشا [۴۵] معیار زیر را به عنوان اختلاف بین دو تابع عضویت $Y_i(x)$ و $y_i(x)$ (امین برآورد $Y_i(x)$) معرفی کردند که در آن $S_{\tilde{Y}_i}$ و $S_{\tilde{y}_i}$ به ترتیب تکیه‌گاه‌های اعداد فازی مثلثی $\tilde{Y}_i$ و $\tilde{y}_i$ هستند.

$$E_i = \int_{S_{\tilde{Y}_i} \cup S_{\tilde{y}_i}} |Y_i(x) - y_i(x)| dx, \quad i = 1, 2, \dots, n \quad (30.2)$$

همچنین $E = \sum_{i=1}^n E_i$ مقدار کل خطای برآورد است.

پارامتر λ

برای ارزیابی عملکرد مدل رگرسیونی فازی، حسن‌پور و همکاران [۳۴] معیار دیگری بنام λ را معرفی کردند که اختلاف بین مراکز مشاهدات و پاسخ‌های برآورد شده را به صورت زیر محاسبه می‌کند

$$\lambda = \sum_{i=1}^n \lambda_i \quad (31.2)$$

که در آن $\lambda_i = \left| m_{y_i} - \sum_{j=0}^p m_{A_j} m_{X_{ij}} \right|$.

اعتبار سنجی متقابل

تعریف ۱.۶.۲. برای ارزیابی نیکویی برازش بین مقدار مشاهده شده $\tilde{Y}_i$ و مقدار برآورد شده $\tilde{y}_i$ ، میانگین توان‌های دوم خطای پیشگویی^۹ (MPE) را به صورت زیر تعریف می‌کنیم

$$\text{MPE} = \frac{1}{n} \sum_{i=1}^n d^2(\tilde{y}_i, \tilde{Y}_i) \quad (32.2)$$

^۹Mean squared prediction error

يکي ديگر از معيارهاي ارزيابي عملکرد براي مدل، اعتبار سنجي متقابل^{۱۰} (cv) است. به منظور اندازه گيري قدرت پيشگويي مدل، هربار i امين مشاهده ($i = 1, 2, \dots, n$) به عنوان مجموعه اعتبارسنج کنار گذاشته مي شود سپس $n - 1$ مشاهده باقي مانده براي برازش مدل استفاده مي شود و مدل برازش شده براي پيشگويي $\tilde{Y}_{(-i)}$ مشاهده کنار گذاشته شده، استفاده مي گردد. در نهايت براي به دست آوردن فاصله بين مشاهده $\tilde{Y}_i$ و پيشگويي $\tilde{Y}_{(-i)}$ (مقدار پيشگويي شده به وسيله $n - 1$ مشاهده باقي مانده)، فاصله موزون (۳۷.۱) را به صورت زير محاسبه مي کنيم

$$MPE_{cv} = \frac{1}{n} \sum_{i=1}^n d^w(\tilde{Y}_i, \tilde{Y}_{(-i)}) \quad (33.2)$$

۳.۶.۲ تحليل رگرسيون ريج فازي براي داده هاي سيمان پرتلند

داده هاي جدول ۱۸.۲ مربوط به سيمان پرتلند هستند که در وودز و همکاران [۶۳] آورده شده است. اين داده ها از مطالعه تاثير ترکيبات مختلف روي حرارت بيرون آمده در طول سخت شدن سيمان پورتلند به دست مي آيد. در اينجا داده هايي که از چهار ترکيب و ميزان گرمای بيرون آمده بعد از ۸۰ روز به دست آمده اند را بررسي مي کنيم. اين چهار ترکيب، تری سدیم آلومينات $4CaO.Al_2O_3.Fe_2O_3$ ، تری کلسیم سيليكات $3CaO.SiO_2$ ، تتراکلسيوم آلمينوفريت $2CaO.SiO_2$ هستند که به ترتيب X_1, X_2, X_3 و X_4 و گرمای بيرون آمده Y ، تعريف مي شوند.

جدول ۱۸.۲: داده هاي فازي شده سيمان پرتلند

$\tilde{Y}$	$\tilde{X}_4$	$\tilde{X}_3$	$\tilde{X}_2$	$\tilde{X}_1$	i
(۷۸/۵, ۸/۶۳۵, ۱۰/۲۰۵)	(۶۰, ۶/۶, ۷/۸)	(۶, ۰/۶۶, ۰/۷۸)	(۲۶, ۲/۸۶, ۳/۳۸)	(۷, ۰/۷۷, ۰/۹۱)	۱
(۷۴/۳, ۸/۹۱۶, ۱۴/۸۶)	(۵۲, ۶/۲۴, ۱۰/۴)	(۱۵, ۱/۸, ۳)	(۲۹, ۳/۴۸, ۵/۸)	(۱, ۰/۱۲, ۰/۲)	۲
(۱۰۴/۳, ۱۹/۸۱۷, ۱۱/۴۷)	(۲۰, ۳/۸, ۲/۲)	(۸, ۱/۵۲, ۰/۸۸)	(۵۶, ۱۰/۶۴, ۶/۱۶)	(۱۱, ۲/۰۹, ۱/۲۱)	۳
(۸۷/۶, ۱۳/۱۴, ۱۴/۸۹۲)	(۴۷, ۷/۰۵, ۷/۹۹)	(۸, ۱/۲, ۱/۳۶)	(۳۱, ۴/۶۵, ۵/۲۷)	(۱۱, ۱/۶۵, ۱/۸۷)	۴
(۹۵/۹, ۱۵/۳۴۴, ۱۶/۳۰۳)	(۳۳, ۵/۲۸, ۵/۶۱)	(۶, ۰/۹۶, ۱/۰۲)	(۵۲, ۸/۳۲, ۸/۸۴)	(۷, ۱/۱۲, ۱/۱۹)	۵
(۱۰۹/۲, ۱۲/۰۱۲, ۱۶/۳۰۳)	(۲۲, ۲/۴۲, ۳/۳)	(۹, ۰/۹۹, ۱/۳۵)	(۵۵, ۶/۰۵, ۸/۲۵)	(۱۱, ۱/۲۱, ۱/۶۵)	۶
(۱۰۲/۷, ۱۲/۳۲۴, ۱۱/۲۹۷)	(۶, ۰/۷۲, ۰/۶۶)	(۱۷, ۲/۰۴, ۱/۸۷)	(۷۱, ۸/۵۲, ۷/۸۱)	(۳, ۰/۳۶, ۰/۳۳)	۷
(۷۲/۵, ۱۳/۰۵, ۱۰/۱۵)	(۴۴, ۷/۹۲, ۶/۱۶)	(۲۲, ۳/۹۶, ۳/۰۸)	(۳۱, ۵/۵۸, ۴/۳۴)	(۱, ۰/۱۸, ۰/۱۴)	۸
(۹۳/۱, ۱۴/۸۹۶, ۱۴/۸۹۶)	(۲۲, ۳/۵۲, ۳/۵۲)	(۱۸, ۲/۸۸, ۲/۸۸)	(۵۴, ۸/۶۴, ۸/۶۴)	(۲, ۰/۳۲, ۰/۳۲)	۹
(۱۱۵/۹, ۱۱/۵۹, ۲۰/۸۹۶)	(۲۶, ۲/۶, ۴/۶۸)	(۴, ۰/۴, ۰/۷۲)	(۴۷, ۴/۷, ۸/۴۶)	(۲۱, ۲/۱, ۳/۷۸)	۱۰
(۸۳/۳, ۱۵/۸۲۷, ۹/۹۹۶)	(۳۴, ۶/۴۶, ۴/۰۸)	(۲۳, ۴/۳۷, ۲/۷۶)	(۴۰, ۷/۶, ۴/۸)	(۱, ۰/۱۹, ۰/۱۲)	۱۱
(۱۱۳/۳, ۱۵/۸۶۲, ۱۳/۵۹۶)	(۱۲, ۱/۶۸, ۱/۴۴)	(۹, ۱/۲۶, ۱/۰۸)	(۶۶, ۹/۲۴, ۷/۹۲)	(۱۱, ۱/۵۴, ۱/۳۲)	۱۲
(۱۰۹/۴, ۱۸/۵۹, ۱۲/۰۳)	(۱۲, ۲/۰۴, ۱/۳۲)	(۸, ۱/۳۶, ۰/۸۸)	(۶۸, ۱۱/۵۶, ۷/۴۸)	(۱۰, ۱/۷, ۱/۱)	۱۳

با توجه به برخي ابهامات در اندازه گيري هاي مربوط به متغيرهاي ورودی و مشاهدات

¹⁰Cross-validation

متغیر پاسخ، با مشورت با یک فرد متخصص با استفاده از یک عدد فازی مثلثی با پهنای راست و چپ متناسب برای مراکز $\tilde{X}_1, \tilde{X}_2, \tilde{X}_3$ و $\tilde{X}_4$ و همچنین $\tilde{Y}$ ، فازی سازی شده‌اند (فرخی [۶]). از آنجا که متغیرهای ورودی فازی هستند، از معیار VIF_G برای تشخیص همخطی بین متغیرها استفاده می‌کنیم. در این صورت

$$VIF_{G_1} = 3/705, \quad VIF_{G_2} = 41/4263, \quad VIF_{G_3} = 4/2225 \quad VIF_{G_4} = 15/847$$

که همخطی بین متغیرها را نشان می‌دهد ($\exists i = 1, 2, 3, 4, VIF_{G_i} \geq 5$). لذا باید از روش رگرسیون ریح فازی پیشنهادی استفاده کرد. برای محاسبه ضریب ریح بهینه به ازای $k = 1/3$ مقدار $MPE_{cv} = 7/2$ که در بین بقیه k ها و مخصوصا در قیاس با $k = 0$ (روش معمولی) که $MPE_{cv} = 7/71$ است، کمترین مقدار را داراست که نشان‌دهنده عملکرد خوب مدل در حضور همخطی چندگانه در داده‌ها است. بنابراین ضریب ریح را برابر $k = 1/3$ اختیار می‌کنیم که با مینیمم کردن رابطه (۲۸.۲) تحت قیود (۲۹.۲) با استفاده از نرم‌افزار 11 Mathematica مدل رگرسیونی ریح فازی به صورت زیر بدست می‌آید

$$\begin{aligned} \tilde{Y} = (&0/038, 0, 0/1) \oplus (2/167, 0/0003, 0) \otimes \tilde{X}_1 \oplus (1/16, 0/00003, 0) \otimes \tilde{X}_2 \\ &\oplus (0/729, 0/0004, 0) \otimes \tilde{X}_3 \oplus (0/49, 0/0002, 0) \otimes \tilde{X}_4 \end{aligned} \quad (34.2)$$

و همچنین برای $k = 1/3$ دیگر شاخص‌های نیکویی برازش را به دست آورده و داریم

$$E = 38/29, \quad \lambda = 20/09.$$

۴.۶.۲ برآوردگر ریح استوار فازی

همانطور که در تابع هدف (۲۸.۲) نیز مشاهده می‌شود، به دلیل وجود توان‌های دوم، این رهیافت در مقابل داده‌های دورافتاده موجود استوار نخواهد بود، بنابراین با استفاده از فاصله قدرمطلق معرفی شده در تعریف ۱۲.۹.۱ که حسن‌پور و همکاران [۳۴] نشان دادند این فاصله نسبت به حضور داده‌های دورافتاده موجود در مجموعه داده استوار است، تابع هدف جدیدی تعریف می‌کنیم که تحت تاثیر داده‌های دورافتاده نخواهد بود. از این رو معادله (۲۵.۲) به صورت زیر تبدیل خواهد شد

$$\sum_{i=1}^n d(\tilde{Y}_i(\tilde{A}', \tilde{A}''), \tilde{Y}_i) + k \|\tilde{A}\|^2 \quad (35.2)$$

که در آن k ضریب ریح نامیده شده و $k \geq 0$ است. می‌توان آن را به صورت زیر بازنویسی کرد

$$\begin{aligned} \sum_{i=1}^n d(\tilde{Y}_i(\tilde{A}', \tilde{A}''), \tilde{Y}_i) + k \|\tilde{A}\|^2 = \sum_{i=1}^n \left[\left| m_{Y_i} - \left(m_{A_0} + \sum_{j=1}^p (m'_{A_j} - m''_{A_j}) m_{X_{ij}} \right) \right| \right. \\ \left. + \left| l_{Y_i} - \left(l_{A_0} + \sum_{j=1}^p (m'_{A_j} l_{X_{ij}} + l'_{A_j} m_{X_{ij}} + m''_{A_j} r_{X_{ij}} + l''_{A_j} m_{X_{ij}}) \right) \right| \right] \end{aligned}$$

$$+ \left| r_{Y_i} - \left(r_{A_0} + \sum_{j=1}^p (m'_{A_j} r_{X_{ij}} + r'_{A_j} m_{X_{ij}} + m''_{A_j} l_{X_{ij}} + r''_{A_j} m_{X_{ij}}) \right) \right| \Bigg] \\ + k \left[\sum_{j=0}^p \left(m_{A_j}^2 + \frac{1}{3} m_{A_j} (r_{A_j} - l_{A_j}) + \frac{1}{12} (l_{A_j}^2 + r_{A_j}^2) \right) \right] \quad (36.2)$$

با مينيمم کردن رابطه (۳۶.۲) تحت قيود زیر

$$m'_{A_j} \geq l'_{A_j}, m''_{A_j} \geq r''_{A_j}, \quad l_{A_0}, r_{A_0}, m'_{A_j}, m''_{A_j}, l'_{A_j}, l''_{A_j}, r'_{A_j}, r''_{A_j} \geq 0, \\ j = 1, 2, \dots, p \quad (37.2)$$

برآورد ضرايب مدل رگرسيونی محاسبه می گردد. برای مينيمم کردن اين معادله از نرم افزار Mathematica 11 استفاده می کنیم.

۵.۶.۲ تحليل رگرسيون ريج استوار فازي برای داده های سيمان پرتلند

در اینجا نیز همانند مثال ۳.۶.۲ برای داده های جدول ۱۸.۲ با استفاده از برآوردگر ريج استوار فازي و انتخاب پارامتر $k = 2$ ، با مينيمم کردن رابطه (۳۶.۲) تحت قيود (۳۷.۲) با استفاده از نرم افزار Mathematica 11 مدل رگرسيون ريج فازي به صورت زیر به دست می آيد

$$\tilde{Y} = (0/2, 0/04, 0) + (2/17, 0, 0/0006) \tilde{X}_1 + (1/15, 0, 0) \tilde{X}_2 \\ + (0/81, 0/008, 0/0003) \tilde{X}_3 + (0/47, 0, 0/0001) \tilde{X}_4 \quad (38.2)$$

همچنين شاخص های ارزیابی مدل را برای اين مسئله به دست می آوريم و نتايج به دست آمده برای مدل رگرسيون کمترین توان های دوم معمولی فازي، رگرسيون ريج فازي معمولی و رگرسيون ريج استوار فازي را در جدول ۱۹.۲ خلاصه می کنیم.

جدول ۱۹.۲: شاخص های ارزیابی مدل برای رگرسيون ريج فازي برای مجموعه داده سيمان پرتلند

مدل رگرسيونی	k	شاخص E	شاخص λ
OLS	۰	۳۷/۷۳	۱۹/۸۲
ريج فازي معمولی	۱/۳	۳۸/۲۹	۲۰/۰۹
ريج استوار فازي	۲	۳۵/۵۲	۱۸/۷۵

همان طور که در جدول ۱۹.۲ نشان داده شد، مدل رگرسيون ريج استوار فازي در قياس با ديگر برآوردگرها از عملکرد بهتری برخوردار است و نتايج عدم کارایی رگرسيون فازي معمولی و رگرسيون ريج فازي معمولی را نشان می دهند.

فصل ۳

برآوردگر لاسو استوار

۱.۳ مقدمه

همان‌طور که در فصل ۲ دیدیم در برخی مسائل کاربردی با مجموعه مشاهداتی سروکار داریم که در آنها داده‌های پرت وجود دارد یا خطای مدل دم‌پهن است. در چنین موقعیت‌هایی به‌خوبی واضح است که روش کمترین توان‌های دوم معمولی در تولید برآوردگری مطمئن ناتوان است و برآوردگر کمترین قدرمطلق انحرافات^۱ (LAD) می‌تواند مفید واقع شود. مخصوصاً زمانی که بخواهیم خواص $\sqrt{n}$ -سازگاری و نرمال‌بودن مجانبی برای این برآوردگرها را بدون در نظر گرفتن هیچ‌گونه شرط گشتاوری برای باقی‌مانده‌ها، بررسی کنیم. به‌دلیل این حقیقت که تابع هدف (به‌طور مثال معیار LAD) تابعی ناهموار است، بسط تیلور معمولی نمی‌تواند برای مطالعه ویژگی‌های مجانبی برآوردگر LAD به‌طور مستقیم به‌کار برده شود. از این رو، در طول دهه‌های گذشته، تلاش‌های زیادی برای اثبات $\sqrt{n}$ -سازگاری و ویژگی‌های مجانبی انواع برآوردگرهای LAD انجام شده‌است. برای آگاهی بیشتر، بسیط و کوینکر [۱۶]، پلارد [۵۳]، بلوم‌فیلد و استیگر [۱۸]، نایت [۴۰]، پنگ و یاوو [۵۲] و لینگ [۴۴] را ببینید، که مسائل مهمی از انتخاب مدل استوار را مورد بررسی قرار داده‌اند.

در مدل‌های رگرسیونی به‌خوبی واضح است که حذف یک متغیر تبیینی مهم، ممکن است مقداری اریبی در برآوردهای پارامتر و نتایج پیشگویی تولید کند (هیستی و همکاران، [۳۵]).

^۱Least absolute deviation

از طرف دیگر، اضافه شدن متغیرهای غیرضروری به مدل می‌تواند کارایی برآورد نتیجه شده را کاهش دهد و پیشگویی‌های با دقت کمتری به دست دهد. بنابراین انتخاب بهترین مدل بر پایه یک نمونه متناهی، همیشه مسئله مورد علاقه از دیدگاه مباحث نظری و کاربردی بوده است. تحت شرط‌های گشتاوری مناسب برای باقی‌مانده‌ها، مسئله انتخاب مدل به‌طور گسترده‌ای در آثار علمی، از قبیل شائو [۵۸]، هارویچ و تسای [۳۸] و شی و تسای [۵۹]، مورد مطالعه قرار گرفته است و دو معیار انتخاب مهم، معیار اطلاع آکایک (AIC) (آکائیک [۱۲]) و معیار اطلاع بیزی (شوارز [۵۷])، به‌طور گسترده‌ای در عمل استفاده شده‌اند. به دلیل اینکه معروف‌ترین معیارهای انتخاب (مانند AIC و BIC) مبتنی بر برآوردهای OLS توسعه یافته‌اند، عملکرد نمونه متناهی آن‌ها تحت خطاهای دم‌پهن ضعیف است. بنابراین هارویچ و تسای [۳۸] مجموعه‌ای از معیارهای انتخاب مدل مفید (مانند AIC، AICc و BIC) را مبتنی بر برآوردهای LAD به دست آوردند. اما علی‌رغم مفید بودنشان، این معیارهای انتخاب متغیر مبتنی بر LAD، محدودیت‌هایی دارند که مهم‌ترین آن‌ها سنگینی حجم محاسبات است. لازم به ذکر است که با افزایش تعداد متغیرهای رگرسیونی، تعداد کل مدل‌های منتخب ممکن به صورت نمایی افزایش می‌یابد. از این رو اگر تعداد پیشگوها نسبتاً زیاد باشد، اجرای روش انتخاب بهترین زیرمجموعه با در نظر گرفتن همه مدل‌های منتخب ممکن، سخت خواهد بود. برای جبران کاستی‌های روش‌های انتخاب مدل معیارهای قدیمی (مانند AIC و BIC)، تیبشیرانی [۶۰] لاسو را ارائه کرد که به‌طور مؤثر و به‌صورت هم‌زمان می‌تواند متغیرهای تبیینی مهم را انتخاب کرده و پارامترهای رگرسیونی را برآورد کند.

در این فصل برآوردگر انقباضی استواری را بررسی می‌کنیم که هم رگرسیون انقباضی و هم مانند لاسو انتخاب متغیر را انجام می‌دهد و نسبت به داده‌های پرت با خطاهای دم‌پهن، مانند LAD، مقاوم است. ایده اصلی این برآوردگر، ترکیب ضابطه LAD معمولی و جریمه لاسو با یکدیگر، برای تولید روش LAD-LASSO است. در قیاس با LAD، برآوردگر LAD-LASSO می‌تواند برآورد پارامتر و انتخاب متغیر را هم‌زمان انجام دهد. در قیاس با لاسو، این برآوردگر نسبت به داده‌های پرت و خطاهای دم‌پهن مقاوم‌تر است.

۲.۳ برآوردگرهای LAD-LASSO

مدل رگرسیون خطی زیر را در نظر بگیرید

$$y_i = \mathbf{x}_i^\top \boldsymbol{\beta} + \epsilon_i, \quad i = 1, \dots, n \quad (1.3)$$

که در آن y_i متغیر پاسخ، $\mathbf{x}_i = (x_{i1}, \dots, x_{ip})^\top$ متغیرهای تبیینی p -بعدی، $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)^\top$ بردار ضرایب رگرسیونی و ϵ_i خطاهای تصادفی مستقل و هم‌توزیع با میانه صفر می‌باشند. علاوه بر این فرض کنید برای $j \leq p_0$ ، $\beta_j \neq 0$ و برای $j > p_0$ ، $\beta_j = 0$ که در آن $0 \leq p_0 \leq p$. بنابراین مدل واقعی، دارای p_0 ضریب غیرصفر و $(p - p_0)$ ضریب صفر خواهد بود. به‌طور معمول

پارامترهای نامعلوم مدل (۱.۳) می‌توانند به وسیله کمینه کردن تابع هدف کمترین توان‌های دوم معمولی، یعنی $\sum_{i=1}^n (y_i - \mathbf{x}_i^\top \beta)^2$ ، برآورد شوند. علی‌رغم این روش، تیبشیرانی [۶۰] برای منقبض کردن ضرایب غیرضروری به سمت صفر، تابع هدف لاسو زیر را پیشنهاد کرد (بخش ۳.۴.۱ را ببینید).

$$\sum_{i=1}^n (y_i - \mathbf{x}_i^\top \beta)^2 + k \sum_{j=1}^p |\beta_j|$$

که در آن $k > 0$ پارامتر تنظیم‌کننده می‌باشد. به دلیل اینکه برآوردگر لاسو برای همه ضرایب رگرسیونی، پارامتر تنظیم‌کننده یکسانی را به کار می‌برد، برآوردگرهای نتیجه شده میزان قابل توجهی اریبی خواهند داشت (فن و لی، [۲۹]). بنابراین از این به بعد مدل لاسوی تعدیل‌یافته^۲ زیر را استفاده می‌کنیم

$$\sum_{i=1}^n (y_i - \mathbf{x}_i^\top \beta)^2 + \sum_{j=1}^p k_j |\beta_j| \quad (2.3)$$

که به ما اجازه می‌دهد برای هر ضریب، پارامتر تنظیم‌کننده خاص خودش را داشته باشیم. به عنوان یک نتیجه، برآوردگر لاسو تعدیل‌یافته قادر به تولید جواب‌های تنگ کارتری نسبت به لاسو است. با این وجود، بسیار واضح است که تابع هدف OLS استفاده شده در لاسو تعدیل‌یافته نسبت به داده‌های پرت بسیار حساس است. برای به دست آوردن برآوردگر استوار نوع لاسو، وانگ و همکاران [۶۲] پیشنهاد دادند تابع هدف LAD را در تابع هدف لاسوی تعدیل‌یافته جایگذاری کرده و تابع هدف به دست آمده زیر را به عنوان LAD-LASSO ارائه کرد

$$Q(\beta) = \sum_{i=1}^n |y_i - \mathbf{x}_i^\top \beta| + \sum_{j=1}^p k_j |\beta_j| \quad (3.3)$$

همان‌طور که مشاهده می‌شود، ضابطه LAD-LASSO ضابطه LAD و جریمه لاسو را ترکیب می‌کند. بنابراین انتظار می‌رود نتایج این برآوردگر در مقابل داده‌پرت استوار باشد و نتایج تنگی را ارائه دهد. پیدا کردن برآوردگر LAD-LASSO بسیار راحت است. برای این منظور، مجموعه داده افزوده^۳ $\{(y_i^*, \mathbf{x}_i^*)\}$ ، به ازای $i = 1, \dots, n+p$ را در نظر بگیرید، که در آن به ازای $1 \leq j \leq n$ ، $(y_i^*, \mathbf{x}_i^*) = (y_i, \mathbf{x}_i)$ و به ازای $1 \leq j \leq p$ ، $(y_{n+j}^*, \mathbf{x}_{n+j}^*) = (0, n\mathbf{k}_j \mathbf{e}_j)$ است که در آن $\mathbf{e}_j$ یک بردار p -بعدی است که مؤلفه j ام آن برابر با یک و دیگر مؤلفه‌های آن برابر با صفر می‌باشند. در این صورت

$$\text{LAD-LASSO} = Q(\beta) = \sum_{i=1}^{n+p} |y_i^* - \mathbf{x}_i^* \beta| \quad (4.3)$$

این معادله، همان ضابطه LAD معمولی است که با در نظر گرفتن $(y_i^*, \mathbf{x}_i^*)$ به عنوان داده اصلی، به دست آمده است. در نتیجه، هر برنامه استاندارد LAD جریمه نشده می‌تواند برای یافتن برآوردگر LAD-LASSO، بدون اضافه کردن برنامه دیگری، استفاده شود.

²Adaptive lasso

³Augmented dataset

۳.۳ ویژگی‌های نظری

برای ساده‌سازی، بردار ضرایب رگرسیونی را به صورت $\beta = (\beta_a^\top, \beta_b^\top)^\top$ تجزیه می‌کنیم که در آن $\beta_a = (\beta_1, \dots, \beta_{p_0})^\top$ و $\beta_b = (\beta_{p_0+1}, \dots, \beta_p)^\top$ می‌باشند. برآوردگر LAD-LASSO متناظر با آن به صورت $\hat{\beta} = (\hat{\beta}_a^\top, \hat{\beta}_b^\top)^\top$ تعریف می‌شود و تابع هدف آن را با $Q(\beta) = Q(\beta_a, \beta_b)$ نشان می‌دهند. علاوه بر این، متغیرهای کمکی را نیز به صورت $\mathbf{x}_i = (\mathbf{x}_{ia}^\top, \mathbf{x}_{ib}^\top)^\top$ تجزیه می‌کنیم که در آن $\mathbf{x}_{ia} = (x_{i1}, \dots, x_{ip_0})^\top$ و $\mathbf{x}_{ib} = (x_{i(p_0+1)}, \dots, x_{ip})^\top$ می‌باشند. برای مطالعه ویژگی‌های نظری برآوردگر LAD-LASSO، پذیره‌های زیر ضروری هستند

A(1): خطای τ_i در نقطه شروع، دارای تابع چگالی مثبت و پیوسته می‌باشد.

A(2): ماتریس $\text{cov}(\mathbf{x}_1) = \Sigma$ موجود و معین مثبت می‌باشد.

توجه کنید که این پذیره‌ها، برای اثبات $\sqrt{n}$ -سازگاری و نرمال بودن مجانبی برآوردگر LAD جریمه‌شده مورد نیاز هستند. علاوه بر این تعریف می‌کنیم

$$a_n = \max\{k_j, 1 \leq j \leq p_0\}$$

و

$$b_n = \min\{k_j, p_0 < j \leq p\}$$

که در آن k_j تابعی از n است. ابتدا سازگاری برآوردگر LAD-LASSO را بیان می‌کنیم.

لم ۱.۳.۳ ($\sqrt{n}$ -سازگاری). فرض کنید $(\mathbf{x}_i, y_i)$ به ازای $i = 1, \dots, n$ مستقل و هم‌توزیع باشند و همچنین مدل رگرسیون خطی (۱.۳) در پذیره‌های A(1) و A(2) صدق کند. اگر $\sqrt{n}a_n \rightarrow 0$ ، آن‌گاه برآوردگر LAD-LASSO، $\sqrt{n}$ -سازگار است.

برهان. برای اثبات به وانگ و همکاران [۶۲] مراجعه کنید. $\square$

لم ۱.۳.۳ نتیجه می‌دهد که اگر پارامترهای تنظیم‌کننده وابسته به متغیرهای معنی‌دار سریع‌تر از $n^{-\frac{1}{4}}$ به صفر همگرا شوند، آن‌گاه برآوردگر LAD-LASSO متناظر با آن می‌تواند $\sqrt{n}$ -سازگار باشد.

اکنون می‌خواهیم نشان دهیم اگر پارامترهای تنظیم‌کننده وابسته به متغیرهای معنی‌دار آهسته‌تر از $n^{-\frac{1}{4}}$ به صفر منقبض شوند، آن‌گاه همان ضرایب رگرسیونی می‌توانند با احتمال تقریباً یک، دقیقاً صفر برآورد شوند.

لم ۲.۳.۳ (تُنکی). تحت پذیره‌های لم ۱.۳.۳ و این پذیره که با احتمال تقریباً یک $\sqrt{n}b_n \rightarrow \infty$ ، برآوردگر LAD-LASSO معرفی شده، $\hat{\beta} = (\hat{\beta}_a^\top, \hat{\beta}_b^\top)^\top$ در شرط $\hat{\beta}_b^\top = 0$ صدق می‌کند.

برهان. برای اثبات به وانگ و همکاران [۶۲] مراجعه کنید. $\square$

لم ۲.۳.۳ نشان می‌دهد که برآوردگر LAD-LASSO به‌طور سازگار قادر به تولید جواب‌های تُنگ برای ضرایب رگرسیونی ناچیز می‌باشد. بنابراین انتخاب متغیر و برآورد پارامتر می‌تواند به‌صورت هم‌زمان انجام شود. از هر دو لم گفته شده به این نتیجه می‌رسیم که سازگاری ریشه n ام برآوردگر $\hat{\beta}$ ، باید وقتی پارامتر تنظیم‌کننده شرایط مناسب را فراهم می‌کند در قید $1 \rightarrow P(\hat{\beta}_\tau = 0)$ صدق کند. در آخر، توزیع مجانبی برآوردگرهای LAD-LASSO را یادآور می‌شویم.

قضیه ۱.۳.۳ (ویژگی پیشگویی). فرض کنید (x_i, y_i) ، $i = 1, \dots, n$ ، مستقل و هم‌توزیع باشند و مدل رگرسیونی (۱.۳) در فرض‌های A(1) و A(2) صدق کند. علاوه بر این اگر $\sqrt{n}a_n \rightarrow 0$ و $\sqrt{n}b_n \rightarrow \infty$ ، آن‌گاه برای برآوردگر پارامترهای LAD-LASSO داریم $\hat{\beta}^\top = (\hat{\beta}_a^\top, \hat{\beta}_b^\top)^\top$

$$\begin{aligned} (1) \quad & P(\hat{\beta}_b = 0) \rightarrow 1 \\ (2) \quad & \sqrt{n}(\hat{\beta}_a - \beta_a) \rightarrow N_{p_0}\left(0, \frac{0.25\Sigma_0^{-1}}{f^2(0)}\right) \end{aligned} \quad (5.3)$$

که در آن $\Sigma_0 = \text{cov}(x_{ia})$ و $f(t)$ چگالی τ_i می‌باشد.

برهان. برای اثبات به وانگ و همکاران [۶۲] مراجعه کنید. $\square$

تحت شروط $\sqrt{n}a_n \rightarrow 0$ و $\sqrt{n}b_n \rightarrow \infty$ ، می‌توان نتیجه گرفت که برآوردگر LAD-LASSO در مقابل خطاهای دم‌پهن استوار است. همچنین این قضیه بیان می‌کند که برآوردگر LAD-LASSO نتیجه‌شده دارای توزیع مجانبی یکسانی با برآوردگر LAD به‌دست آمده تحت مدل واقعی، است. از این رو ویژگی اراکل برآوردگر LAD-LASSO اثبات شده می‌باشد. علاوه بر این، به سبب محدب بودن و تکه‌ای خطی بودن تابع هدف $Q(\beta)$ ، ویژگی‌های برآوردگر LAD-LASSO مطرح شده در این بخش، فراموشی می‌باشند.

۴.۳ تعیین پارامتر تنظیم کننده

در این بخش به برآورد پارامتر تنظیم‌کننده می‌پردازیم. برای این هدف، می‌توان از روش‌های اعتبار سنجی متقابل، اعتبار سنجی متقابل تعمیم‌یافته، AIC، AICs و BIC به‌طور مستقیم برای انتخاب بهینه پارامتر تنظیم‌کننده، k_j ، استفاده کرد (بعد از تغییرات مناسب مانند جایگزین کردن عبارت کمترین توان‌های دوم با تابع LAD). اگرچه استفاده از این‌ها در وضعیت فعلی، حداقل بنا به دلایل بیان شده در بخش ۱.۳، سخت خواهد بود. اول اینکه تابع هدف LAD-LASSO مجموعاً p پارامتر تنظیم‌کننده را شامل می‌شود و برآورد هم‌زمان این تعداد پارامتر تنظیم‌کننده از لحاظ محاسباتی بسیار پرهزینه است. دوم اینکه، ویژگی‌های آماری آن‌ها نیز تحت خطاهای دم‌پهن، به راحتی قابل فهم نیست. از این رو می‌توانیم گزینه زیر را مد نظر قرار دهیم.

با استفاده از ایده تیبشیرانی [۶۰]، می‌توانیم به برآوردگر LAD-LASSO به‌عنوان برآوردگری بیزی که هر ضریب رگرسیونی از یک پیشین‌نمایی دوگانه با پارامتر مکان صفر و پارامتر مقیاس nk_j پیروی می‌کند، نگاه کنیم، بنابراین k_j بهینه می‌تواند به‌وسیله کمینه کردن تابع لگاریتم درست‌نمایی پسین منفی زیر به‌دست آید

$$\sum_{i=1}^n |y_i - \mathbf{x}_i^\top \beta| + n \sum_{j=1}^p k_j |\beta_j| - \log(\circ / \Delta nk_j) \quad (۶.۳)$$

که در آن $k_j = \frac{1}{(n|\beta_j|)}$ را نتیجه می‌دهد. در مسائل کاربردی مقادیر β_j را نمی‌دانیم؛ اما می‌توانیم به راحتی به‌وسیله برآوردگر جریمه‌نشده LAD، $\tilde{\beta}_j$ ، آن‌ها را برآورد کنیم. در این صورت داریم

$$\tilde{k}_j = \frac{1}{(n|\tilde{\beta}_j|)}$$

توجه کنید که $\sqrt{n}$ – سازگار بودن $\tilde{\beta}_j$ ، نشان می‌دهد که برای $j \leq p$ ، $\sqrt{n}\tilde{k}_j \rightarrow \circ$ است. بنابراین یکی از شرایط لازم در قضیه برقرار شد. اگرچه شرط دیگر، $\sqrt{n}\tilde{k}_j \rightarrow \infty$ به ازای $j > p$ ، ضرورتاً برقرار نیست. بنابراین به کار بردن مستقیم این چنین برآورد پارامترهای تنظیم‌کننده، اگر مدل واقعی پایه در حقیقت از بعد متناهی باشد، می‌تواند منجر به تولید نتایج بیش‌برازش شده شود. این ویژگی بسیار شبیه به AIC است. در مقابل، توجه کنید که در تابع (۶.۳)، میزان پیچیدگی مدل نهایی به‌وسیله $-\log(k)$ کنترل می‌شود، که به‌سادگی می‌تواند به‌عنوان تعداد متغیرهای معنی‌دار در AIC در نظر گرفته شود. از این رو می‌توانیم تابع هدف نوع BIC زیر را در نظر می‌گیریم

$$\sum_{i=1}^n |y_i - \mathbf{x}_i^\top \beta| + n \sum_{j=1}^p k_j |\beta_j| - \log(\circ / \Delta nk_j) \log(n) \quad (۷.۳)$$

که با عامل $\log(n)$ ، در روشی مشابه با BIC، پیچیدگی مدل را جریمه می‌کند. این تابع برآورد منجر به پارامتر تنظیم‌کننده زیر می‌شود

$$\hat{k}_j = \frac{\log(n)}{n|\hat{\beta}_j|} \quad (۸.۳)$$

که در قیود $\sqrt{n}\hat{k}_j \rightarrow \circ$ برای $j \leq p$ و $\sqrt{n}\hat{k}_j \rightarrow \infty$ برای $j > p$ صدق می‌کند. بنابراین سازگاری انتخاب متغیر توسط برآوردگر نهایی، تضمین شده است.

تا اینجا روش LAD-LASSO را ارائه کردیم که ایده‌های LAD و لاسو را برای به‌دست آوردن روش رگرسیونی انقباضی استوار که هم‌زمان قادر به انتخاب متغیر نیز می‌باشد، ترکیب کردیم. علاوه بر این، ایده‌های مشابهی می‌تواند برای M – برآوردهای هوبر با استفاده از تابع هدف نوع ψ زیر برای لاسو پیشنهاد شود

$$Q(\beta) = \sum_{i=1}^n \psi(y_i - \mathbf{x}_i^\top \beta) + n \sum_{j=1}^p k_j |\beta_j| \quad (۹.۳)$$

که در آن $\psi(\cdot)$ ، ψ - تابع هوبر است. به دلیل این حقیقت که توابع ψ ، به عنوان یک نوع خاص، تابع LAD ($|\cdot|$) را شامل می‌شوند، بنابراین در صورتی که یک پارامتر تنظیم کننده بهینه استفاده شود انتظار می‌رود حتی سازگارتر از LAD-LASSO باشند. اگرچه برای مجموعه داده واقعی، چگونگی انتخاب یک پارامتر تنظیم کننده بهینه مسئله‌ای جالب و چالش برانگیز است.

۵.۳ شبيه‌سازی

در این بخش نیز همانند شبيه‌سازی بخش ۴.۲، برای ارزیابی مدل لاسو استوار (LAD-LASSO) دو طرح $n > p$ و دو حالت با داده پرت و بدون داده پرت را بررسی می‌کنیم و سپس برآوردگر LAD-LASSO را با یک برآوردگر لاسو معمولی و یک M-برآوردگر معمولی مقایسه خواهیم کرد. تمامی محاسبات انجام شده در این مثال شبيه‌سازی با استفاده از نرم افزار R اجرا شده‌اند. برای محاسبه برآوردگر LAD-LASSO از روش پیشنهادی وانگ و همکاران [۶۲] استفاده می‌کنیم. همچنین برآوردگر لاسو معمولی را با استفاده از پکیج glmnet از نرم افزار R و M-برآوردگر معمولی را با استفاده از پکیج robust موجود در نرم افزار R محاسبه و نتایج این برآوردگرها را با استفاده از معیارهای معرفی شده در بخش ۴.۲ مقایسه می‌کنیم. در اینجا، منظور از REff(MPE) نسبت خطای پیش گویی برآوردگر لاسو نسبت به سایر برآوردگرها است. همچنین از آنجایی که در برآوردگر لاسو انتخاب متغیر نیز امکان پذیر می‌باشد، دو معیار C و IC را هم محاسبه می‌کنیم. معیار C نشان دهنده میانگین تعداد ضرایب صفری است که به درستی صفر برآورد می‌شوند و IC نشان دهنده ضرایب غیر صفری است که به اشتباه صفر برآورد می‌شوند.

طرح ۱ در این طرح روش تولید داده‌ای مشابه با طرح ۱ در مثال شبيه‌سازی فصل ۲ داریم. حجم داده و نحوه ایجاد همخطی و داده پرت نیز به همان صورت می‌باشد. تعداد پارامتر $p = 8$ ، داده‌ها به حجم‌های $n = 50, 200$ و مقادیر واقعی بردار پارامتر را، $\beta_0 = (3, 1/5, 0, 0, 2, 0, 0, 0)^T$ در نظر می‌گیریم.

همان طور که در جدول‌های ۱.۳ و ۲.۳ مشاهده می‌کنیم، با افزایش همخطی بین متغیرهای تبیینی برآوردگر استوار معمولی در برآورد ضرایب مدل عملکرد خوبی ندارد. همچنین مشاهده می‌شود برآوردگر لاسو نیز در بحث پیشگویی مدل، عملکرد خوبی نداشته که می‌تواند به علت حضور داده‌های پرت ناخواسته در متغیر پاسخ باشد. در جدول‌های ۳.۳ و ۴.۳ با افزایش حجم داده، زمانی که همخطی کمی بین داده‌ها موجود است، برآوردگر لاسو از دیگر برآوردگرها عملکرد بهتری دارد. یکی از روش‌های برخورد با داده پرت نیز جمع‌آوری داده اضافی می‌باشد و می‌توان گفت این افزایش حجم باعث کاهش تاثیر داده پرت شده است.

در جدول‌های ۵.۳ و ۶.۳ در حجم داده‌ها میزان ۲۰ درصد از داده‌ها را آلوده می‌کنیم. با محاسبه معیارهای سنجش، مشاهده می‌شود این آلودگی تاثیر به سزایی بر برآوردگر لاسو معمولی دارد و برآوردگر استوار معمولی حتی در همخطی بالا نیز از برآوردگر لاسو بهتر عمل

جدول ۱.۳: نتایج شبیه‌سازی مربوط به لاسو استوار، طرح ۱، $n = 50$ ، $\rho = 0.3$ و $\tau = 0$

LAD – LASSO		MM		LASSO		ضرایب
<i>MSE</i>	<i>mean</i>	<i>MSE</i>	<i>mean</i>	<i>MSE</i>	<i>mean</i>	
0.0581	3.0049	0.0433	3.0054	0.1382	2.7805	$\hat{\beta}_1$
0.0643	1.4986	0.0509	1.5265	0.1143	1.3273	$\hat{\beta}_2$
0.0485	0.0133	0.0478	0.0170	0.0150	0.0227	$\hat{\beta}_3$
0.0495	0.0114	0.0670	0.0049	0.0128	0.0234	$\hat{\beta}_4$
0.0668	1.9372	0.0613	1.9910	0.2296	1.6651	$\hat{\beta}_5$
0.0472	-0.0176	0.0640	-0.0506	0.0091	0.0013	$\hat{\beta}_6$
0.0451	0.0190	0.0581	0.0396	0.0079	0.0067	$\hat{\beta}_7$
0.0334	0.0158	0.0431	-0.0178	0.0079	0.0129	$\hat{\beta}_8$
0.4128	—	0.4354	—	0.5346	—	MSE
		MM – BR		MM	LASSO	
		8.4059		8.4466	9.6505	$\sqrt{\text{MPE}}$
		1.1481		1.1425	1.0000	REff
		24.9436		24.9436	24.9436	MCN
		2.2900		1.7500	4.1400	C
		0.4400		0.4000	0.7800	IC

جدول ۲.۳: نتایج شبیه‌سازی مربوط به لاسو استوار، طرح ۱، $n = 50$ ، $\rho = 0.99$ و $\tau = 0$

LAD – LASSO		MM		LASSO		ضرایب
<i>MSE</i>	<i>mean</i>	<i>MSE</i>	<i>mean</i>	<i>MSE</i>	<i>mean</i>	
1.7509	2.5347	1.9781	3.0367	1.4762	2.4938	$\hat{\beta}_1$
1.6518	1.3705	2.3260	1.6795	1.3062	1.1829	$\hat{\beta}_2$
0.6251	0.3519	2.1873	0.1152	0.4690	0.2828	$\hat{\beta}_3$
0.4287	0.3024	3.0640	0.0332	0.2842	0.2363	$\hat{\beta}_4$
1.9114	1.0196	2.8034	1.9393	1.9161	0.9271	$\hat{\beta}_5$
0.3142	0.2237	2.9255	-0.3419	0.2721	0.2081	$\hat{\beta}_6$
0.4457	0.3002	2.6566	0.2675	0.1539	0.1320	$\hat{\beta}_7$
0.6112	0.3351	4.3792	-0.2116	0.5023	0.2953	$\hat{\beta}_8$
7.7390	—	22.3200	—	6.3799	—	MSE
		MM – BR		MM	LASSO	
		7.9991		8.4466	10.8734	$\sqrt{\text{MPE}}$
		1.3593		1.2873	1.0000	REff
		860.5062		860.5062	860.5062	MCN
		3.4300		0.2000	3.7900	C
		1.0100		0.0900	1.0000	IC

جدول ۳.۳: نتایج شبه‌سازی مربوط به لاسو استوار، طرح ۱، $n = 200$ ، $\rho = 0.3$ و $\tau = 0$

LAD – LASSO		MM		LASSO		ضرایب
MSE	mean	MSE	mean	MSE	mean	
0.0093	2.9968	0.0080	2.9906	0.0083	2.9620	$\hat{\beta}_1$
0.0152	1.4825	0.0141	1.5001	0.0120	1.4670	$\hat{\beta}_2$
0.0137	0.0148	0.0098	0.0088	0.0032	0.0213	$\hat{\beta}_3$
0.0122	0.0166	0.0102	0.0100	0.0036	0.0170	$\hat{\beta}_4$
0.0149	1.9785	0.0093	1.9969	0.0122	1.9351	$\hat{\beta}_5$
0.0128	-0.0104	0.0121	-0.0247	0.0031	0.0014	$\hat{\beta}_6$
0.0144	-0.0012	0.0121	-0.0039	0.0023	0.0003	$\hat{\beta}_7$
0.0119	-0.0087	0.0090	0.0028	0.0026	0.0118	$\hat{\beta}_8$
0.1043	—	0.0846	—	0.0472	—	MSE
		MM – BR		MM	LASSO	
		13.9962		13.8918	13.8655	$\sqrt{\text{MPE}}$
		0.9907		0.9981	1.0000	REff
		18.9899		18.9899	18.9899	MCN
		3.2300		3.2200	4.4600	C
		0.5900		0.6400	0.8900	IC

جدول ۴.۳: نتایج شبه‌سازی مربوط به لاسو استوار، طرح ۱، $n = 200$ ، $\rho = 0.99$ و $\tau = 0$

LAD – LASSO		MM		LASSO		ضرایب
MSE	mean	MSE	mean	MSE	mean	
3.0555	2.3780	3.6624	2.7999	2.4317	2.8692	$\hat{\beta}_1$
2.0857	1.1132	6.4281	1.5031	1.9526	1.2591	$\hat{\beta}_2$
0.7888	0.4160	4.4475	0.1884	0.4792	0.3076	$\hat{\beta}_3$
1.0279	0.4740	6.6393	0.2135	0.8515	0.4162	$\hat{\beta}_4$
2.6429	0.8547	4.2245	1.9329	2.4799	0.7286	$\hat{\beta}_5$
0.7401	0.2953	5.5047	-0.5263	0.3659	0.1824	$\hat{\beta}_6$
0.7254	0.3687	5.4941	-0.0831	0.2298	0.1653	$\hat{\beta}_7$
1.7590	0.5462	12.3461	0.4484	1.0847	0.4639	$\hat{\beta}_8$
12.8253	—	46.7467	—	9.8753	—	MSE
		MM – BR		MM	LASSO	
		13.8088		13.8918	13.8981	$\sqrt{\text{MPE}}$
		1.0065		1.0005	1.0000	REff
		5734.6549		5734.6549	5734.6549	MCN
		3.6400		0.1800	3.8600	C
		1.2300		0.1100	1.1500	IC

جدول ۵.۳: نتایج شبیه‌سازی مربوط به لاسو استوار، طرح ۱، $n = 50$ ، $\rho = 0.3$ و $\tau = 0.2$

LAD – LASSO		MM		LASSO		ضرایب
<i>MSE</i>	<i>mean</i>	<i>MSE</i>	<i>mean</i>	<i>MSE</i>	<i>mean</i>	
0.1021	2.9842	0.0860	2.9352	3.2989	1.3008	$\hat{\beta}_1$
0.1433	1.3248	0.1215	1.3477	2.2398	0.0039	$\hat{\beta}_2$
0.4180	0.5408	0.4363	0.5608	0.3309	0.2919	$\hat{\beta}_3$
0.1848	-0.2532	0.1769	-0.2461	0.0000	0.0000	$\hat{\beta}_4$
0.1355	2.0868	0.1292	2.0491	1.4250	1.0382	$\hat{\beta}_5$
0.3611	0.5067	0.4329	0.5670	0.5098	0.4465	$\hat{\beta}_6$
0.1594	-0.2438	0.1783	-0.2526	0.0000	0.0000	$\hat{\beta}_7$
0.2232	0.3959	0.2491	0.4121	0.3274	0.3833	$\hat{\beta}_8$
1.7274	—	1.8101	—	8.1319	—	MSE
		MM – BR		MM	LASSO	
		15.5321		15.3084	74.2398	$\sqrt{\text{MPE}}$
		4.7798		4.8496	1.0000	REff
		134.3297		134.3297	134.3297	MCN
		0.8000		0.4700	3.4600	C
		0.0900		0.0400	1.6100	IC

جدول ۶.۳: نتایج شبیه‌سازی مربوط به لاسو استوار، طرح ۱، $n = 50$ ، $\rho = 0.99$ و $\tau = 0.2$

LAD – LASSO		MM		LASSO		ضرایب
<i>MSE</i>	<i>mean</i>	<i>MSE</i>	<i>mean</i>	<i>MSE</i>	<i>mean</i>	
4.3259	1.0075	4.2803	1.0512	8.8510	0.0290	$\hat{\beta}_1$
1.2222	0.4991	1.1852	0.6107	2.2500	0.0000	$\hat{\beta}_2$
1.2242	0.9392	1.4308	1.0031	0.1742	0.1549	$\hat{\beta}_3$
0.1503	-0.0869	0.3828	-0.2384	0.0000	0.0000	$\hat{\beta}_4$
1.2554	1.0559	1.3283	1.1310	2.6129	0.5471	$\hat{\beta}_5$
1.9107	1.2407	1.8054	1.1709	1.5537	0.8863	$\hat{\beta}_6$
0.2888	0.2399	0.4479	0.2335	0.0000	0.0000	$\hat{\beta}_7$
2.5112	1.4866	2.3699	1.4251	2.0242	1.1945	$\hat{\beta}_8$
12.8886	—	13.2306	—	17.4660	—	MSE
		MM – BR		MM	LASSO	
		8.1657		8.6110	73.5831	$\sqrt{\text{MPE}}$
		9.0113		8.5452	1.0000	REff
		1941.1845		1941.1845	1941.1845	MCN
		0.8000		0.3500	3.1300	C
		0.4200		0.2300	2.7000	IC

جدول ۷.۳: نتایج شبیه‌سازی مربوط به لاسو استوار، طرح ۱، $n = 200$ ، $\rho = 0.3$ و $\tau = 0.2$

LAD – LASSO		MM		LASSO		ضرایب
<i>MSE</i>	<i>mean</i>	<i>MSE</i>	<i>mean</i>	<i>MSE</i>	<i>mean</i>	
0.0193	3.0471	0.0159	2.9643	2.2560	1.5373	$\hat{\beta}_1$
0.0821	1.2570	0.0816	1.2507	2.2500	0.0000	$\hat{\beta}_2$
0.3143	0.5380	0.3639	0.5863	0.1423	0.2474	$\hat{\beta}_3$
0.0904	-0.2662	0.0701	-0.2271	0.0000	0.0000	$\hat{\beta}_4$
0.0343	2.0907	0.0218	2.0078	1.0394	1.0522	$\hat{\beta}_5$
0.2864	0.5134	0.3475	0.5730	0.2406	0.3568	$\hat{\beta}_6$
0.1174	-0.3025	0.1031	-0.2787	0.0000	0.0000	$\hat{\beta}_7$
0.2090	0.4422	0.2380	0.4758	0.2580	0.4371	$\hat{\beta}_8$
1.1532	—	1.2419	—	6.1863	—	MSE
		MM – BR		MM	LASSO	
		24.6157		24.3322	132.5499	$\sqrt{\text{MPE}}$
		5.3848		5.4475	1.0000	<i>REff</i>
		95.1834		95.1834	95.1834	MCN
		0.1900		0.2300	2.8800	C
		0.0000		0.0000	1.4300	IC

جدول ۸.۳: نتایج شبیه‌سازی مربوط به لاسو استوار، طرح ۱، $n = 200$ ، $\rho = 0.99$ و $\tau = 0.2$

LAD – LASSO		MM		LASSO		ضرایب
<i>MSE</i>	<i>mean</i>	<i>MSE</i>	<i>mean</i>	<i>MSE</i>	<i>mean</i>	
3.8014	1.0668	4.0991	0.9859	8.9489	0.0093	$\hat{\beta}_1$
1.1992	0.4381	1.1909	0.4371	2.2500	0.0000	$\hat{\beta}_2$
0.9675	0.9317	0.9612	0.9434	0.1136	0.1775	$\hat{\beta}_3$
0.0460	-0.0023	0.0405	0.0415	0.0000	0.0000	$\hat{\beta}_4$
0.7971	1.1472	0.7825	1.1434	2.3777	0.5123	$\hat{\beta}_5$
1.3625	1.1421	1.4287	1.1782	0.8186	0.7977	$\hat{\beta}_6$
0.1027	0.1941	0.1296	0.2669	0.0000	0.0000	$\hat{\beta}_7$
2.3417	1.5098	2.2292	1.4784	2.1521	1.3909	$\hat{\beta}_8$
10.6180	—	10.8617	—	16.6609	—	MSE
		MM – BR		MM	LASSO	
		15.4503		15.0943	138.2280	$\sqrt{\text{MPE}}$
		8.9466		9.1576	1.0000	<i>REff</i>
		35.06669		35.06669	35.06669	MCN
		0.6500		0.6300	2.6700	C
		0.1300		0.0500	2.6000	IC

می‌کند. اما برآوردگر LAD-LASSO برآورد بهتری نسبت به دیگر برآوردگرها به دست می‌دهد. در جدول‌های ۷.۳ و ۸.۳ مشاهده می‌شود با افزایش داده، مخصوصاً در همخطی بالا، برآوردگر لاسو عملکرد خوبی ندارد. این نتیجه نشان می‌دهد همیشه با افزایش حجم داده نمی‌توان تاثیر داده‌های پرت را کاهش داد، اما همان‌طور که مشاهده می‌شود برآوردگر LAD-LASSO عملکرد خوبی نسبت به برآوردگر لاسو نشان داده است، حتی نسبت به برآوردگر استوار معمولی نیز عملکرد بهتری در برآورد ضرایب دارد.

طرح ۲. در این طرح نیز تولید داده‌ای همانند طرح ۲ در شبیه‌سازی بخش ۴.۲ خواهیم داشت و بردار ضرایب را به صورت $\beta_0 = (\underbrace{0, \dots, 0}_{10}, \underbrace{2, \dots, 2}_{10}, \underbrace{0, \dots, 0}_{10}, \underbrace{2, \dots, 2}_{10})^T$ در نظر می‌گیریم. در این مدل گروه‌هایی از متغیرهای موثر و غیر موثر وجود دارند که همبستگی بین آن‌ها یکسان است. این طرح را برای حجم‌های $n = 100, 300$ بررسی می‌کنیم.

همان‌طور که در جدول‌های ۹.۳ و ۱۰.۳ مشاهده می‌شود در عدم حضور داده‌پرت، در کل به‌خصوص زمانی که میزان همخطی بین متغیرها افزایش می‌یابد برآوردگر LAD-LASSO عملکرد بهتری نشان می‌دهد، همچنین در جدول‌های ۱۱.۲ و ۱۲.۲، با افزایش حجم داده، در همخطی پایین، برآوردگر لاسو در برآورد ضرایب عملکرد بهتری نسبت به حجم داده کمتر در جدول ۹.۳ نشان می‌دهد، این نتیجه می‌تواند به دلیل حجم کم مجموعه داده و شناسایی تعدادی از نقاط داده به عنوان داده‌پرت باشد که با افزایش حجم برطرف می‌شود، اما در حضور همخطی بالا این عملکرد به شدت تحت تاثیر قرار می‌گیرد، یکی از دلایل این امر می‌تواند نحوه ایجاد همخطی در بین متغیرها باشد. در کل با افزایش همخطی، برآوردگر LAD-LASSO عملکرد بهتری را در برآورد ضرایب و پیشگویی نشان می‌دهد.

در جدول‌های ۱۳.۳، ۱۴.۳، ۱۵.۳ و ۱۶.۳ با ایجاد ۲۰ درصد آلودگی در مجموعه داده، برآوردگر لاسو معمولی ضعیف‌ترین عملکرد را در بین برآوردگرهای دیگر نشان می‌دهد و با افزایش همخطی برآوردگر LAD-LASSO نسبت به برآوردگر استوار معمولی عملکرد بهتری را در برآورد ضرایب نشان می‌دهد.

۶.۳ تحلیل داده‌های چربی بدن

در این بخش رفتار برآوردگر ریج استوار معرفی شده را در قالب یک مثال واقعی مورد ارزیابی قرار می‌دهیم. در این مثال مجموعه داده‌های Vo2 معرفی شده در بخش ۵.۲ استفاده می‌شود و برآوردگر LAD-LASSO را با برآوردگرهای لاسو معمولی و M-برآوردگر معرفی شده در شبیه‌سازی مقایسه می‌کنیم. در این جا نیز همانند بخش ۵.۲ برای مقایسه برآوردگرها از معیار جذر میانگین خطای پیشگویی (۱۹.۲) در بخش ۴.۲ استفاده می‌کنیم.

بر اساس جدول ۱۷.۳ و نمودار ۱.۳ مشاهده می‌شود در حضور داده‌پرت در مجموعه داده،

جدول ۹.۳: نتایج شبیه‌سازی مربوط به لاسو استوار، طرح ۲، $n = 100$ ، $\rho = 0.3$ و $\tau = 0$

LAD – LASSO		MM		LASSO		ضرایب
MSE	mean	MSE	mean	MSE	mean	
0.0323	0.0204	0.0439	0.0076	0.0000	0.0000	$\hat{\beta}_1$
0.0359	-0.0101	0.0559	0.0115	0.0000	0.0000	$\hat{\beta}_2$
0.0401	0.0350	0.0585	0.0298	0.0001	0.0015	$\hat{\beta}_3$
0.0250	-0.0218	0.0491	-0.0498	0.0000	0.0000	$\hat{\beta}_4$
0.0246	-0.0097	0.0548	-0.0014	0.0000	0.0000	$\hat{\beta}_5$
0.0269	0.0099	0.0582	-0.0179	0.0000	0.0000	$\hat{\beta}_6$
0.0281	0.0004	0.0702	0.0171	0.0000	0.0001	$\hat{\beta}_7$
0.0452	-0.0184	0.0749	-0.0237	0.0000	0.0000	$\hat{\beta}_8$
0.0347	0.0251	0.0628	0.0278	0.0000	0.0000	$\hat{\beta}_9$
0.0324	0.0083	0.0543	-0.0169	0.0000	0.0003	$\hat{\beta}_{10}$
0.0419	1.9865	0.0477	2.0229	0.7237	1.2917	$\hat{\beta}_{11}$
0.0394	1.9753	0.0474	1.9724	0.3888	1.6539	$\hat{\beta}_{12}$
0.0363	1.9951	0.0495	1.9881	0.3434	1.6661	$\hat{\beta}_{13}$
0.0477	2.0208	0.0669	2.0113	0.2899	1.6961	$\hat{\beta}_{14}$
0.0547	1.9477	0.0576	1.9786	0.3669	1.6299	$\hat{\beta}_{15}$
0.0644	2.0206	0.0686	2.0104	0.2933	1.6647	$\hat{\beta}_{16}$
0.0343	1.9629	0.0673	1.9513	0.3041	1.6914	$\hat{\beta}_{17}$
0.0433	2.0153	0.0598	2.0123	0.4138	1.6219	$\hat{\beta}_{18}$
0.0368	1.9875	0.0480	1.9968	0.4197	1.6636	$\hat{\beta}_{19}$
0.0402	1.9958	0.0504	2.0132	0.7118	1.3259	$\hat{\beta}_{20}$
0.0424	0.0155	0.0540	0.0024	0.0012	0.0079	$\hat{\beta}_{21}$
0.0295	-0.0043	0.0565	-0.0144	0.0000	0.0000	$\hat{\beta}_{22}$
0.0225	-0.0093	0.0450	0.0083	0.0001	0.0009	$\hat{\beta}_{23}$
0.0228	0.0092	0.0597	0.0085	0.0000	0.0002	$\hat{\beta}_{24}$
0.0290	0.0216	0.0735	-0.0066	0.0004	0.0031	$\hat{\beta}_{25}$
0.0251	-0.0220	0.0547	0.0008	0.0000	0.0000	$\hat{\beta}_{26}$
0.0362	0.0158	0.0674	0.0059	0.0000	0.0000	$\hat{\beta}_{27}$
0.0252	0.0116	0.0415	0.0222	0.0000	0.0000	$\hat{\beta}_{28}$
0.0299	0.0176	0.0521	0.0185	0.0007	0.0032	$\hat{\beta}_{29}$
0.0310	-0.0066	0.0532	-0.0175	0.0020	0.0068	$\hat{\beta}_{30}$
0.0416	2.0147	0.0436	2.0286	0.6894	1.3193	$\hat{\beta}_{31}$
0.0392	2.0001	0.0471	1.9977	0.3991	1.6663	$\hat{\beta}_{32}$
0.0463	1.9799	0.0544	1.9842	0.2805	1.7036	$\hat{\beta}_{33}$
0.0379	2.0031	0.0483	2.0186	0.2597	1.6748	$\hat{\beta}_{34}$
0.0415	1.9870	0.0627	1.9715	0.3781	1.6354	$\hat{\beta}_{35}$
0.0445	2.0223	0.0725	2.0419	0.2912	1.7247	$\hat{\beta}_{36}$
0.0457	1.9743	0.0639	1.9859	0.4251	1.5678	$\hat{\beta}_{37}$
0.0547	1.9906	0.0671	2.0113	0.3461	1.6489	$\hat{\beta}_{38}$
0.0439	2.0055	0.0557	2.0127	0.2557	1.6978	$\hat{\beta}_{39}$
0.0506	2.0179	0.0645	1.9971	1.4089	2.9920	$\hat{\beta}_{40}$
1.5034	—	2.2829	—	8.9937	—	MSE
		MM – BR		MM	LASSO	
		13.9634		15.2525	34.5908	$\sqrt{MPE}$
		2.4773		2.2677	1.0000	REff
		157.9601		157.9601	157.9601	MCN
		9.7500		6.5000	19.9400	C
		4.8500		3.1500	10.9000	IC

جدول ۱۰.۳: نتایج شبیه‌سازی مربوط به لاسو استوار، طرح ۲، $n = 100$ ، $\rho = 0.99$ و $\tau = 0$

LAD – LASSO		MM		LASSO		ضرایب
MSE	mean	MSE	mean	MSE	mean	
0.3634	0.2675	2.0065	0.0516	0.0005	0.0023	$\hat{\beta}_1$
0.2376	0.1414	2.5554	0.0780	0.0019	0.0043	$\hat{\beta}_2$
0.2283	0.1811	2.6729	0.2012	0.0010	0.0044	$\hat{\beta}_3$
0.1859	0.0683	2.2430	-0.3368	0.0000	0.0000	$\hat{\beta}_4$
0.2211	0.1325	2.5081	-0.0097	0.0000	0.0000	$\hat{\beta}_5$
0.2299	0.1038	2.6635	-0.1208	0.0082	0.0104	$\hat{\beta}_6$
0.1723	0.0776	3.2102	0.1154	0.0006	0.0025	$\hat{\beta}_7$
0.3158	0.1333	3.4269	-0.1603	0.0066	0.0081	$\hat{\beta}_8$
0.3839	0.1813	2.8725	0.1879	0.0021	0.0086	$\hat{\beta}_9$
0.2474	0.1754	2.4823	-0.1142	0.0162	0.0171	$\hat{\beta}_{10}$
1.2808	1.6325	2.1810	2.1549	3.2509	0.3139	$\hat{\beta}_{11}$
1.5357	1.9764	2.1655	1.8137	2.5273	0.9444	$\hat{\beta}_{12}$
1.1772	2.0000	2.2645	1.9198	2.5146	0.9050	$\hat{\beta}_{13}$
1.3149	2.0523	3.0583	2.0764	2.2370	0.9640	$\hat{\beta}_{14}$
1.4941	1.7418	2.6345	1.8550	2.4472	1.0069	$\hat{\beta}_{15}$
1.8622	1.9951	3.1365	2.0706	2.6241	0.9368	$\hat{\beta}_{16}$
1.4252	1.8326	3.0776	1.6709	2.4149	0.8971	$\hat{\beta}_{17}$
1.6193	1.9381	2.7363	2.0834	2.6853	0.8633	$\hat{\beta}_{18}$
1.2616	1.9054	2.1945	1.9784	2.8776	0.7836	$\hat{\beta}_{19}$
1.2643	1.7638	2.3032	2.0895	2.9897	0.4184	$\hat{\beta}_{20}$
0.4268	0.2903	2.4681	0.0164	0.0118	0.0155	$\hat{\beta}_{21}$
0.1876	0.1194	2.5815	-0.0973	0.0000	0.0000	$\hat{\beta}_{22}$
0.1738	0.0916	2.0572	0.0564	0.0048	0.0069	$\hat{\beta}_{23}$
0.1877	0.0875	2.7290	0.0573	0.0025	0.0067	$\hat{\beta}_{24}$
0.2832	0.2105	3.3604	-0.0446	0.0000	0.0000	$\hat{\beta}_{25}$
0.1788	0.1029	2.5003	0.0053	0.0000	0.0000	$\hat{\beta}_{26}$
0.2268	0.1415	3.0824	0.0399	0.0000	0.0000	$\hat{\beta}_{27}$
0.2218	0.1344	1.8966	0.1504	0.0033	0.0059	$\hat{\beta}_{28}$
0.2185	0.1468	2.3836	0.1249	0.0000	0.0000	$\hat{\beta}_{29}$
0.4437	0.2543	2.4337	-0.1185	0.0119	0.0111	$\hat{\beta}_{30}$
1.0848	1.8357	1.9926	2.1932	3.1215	0.3407	$\hat{\beta}_{31}$
1.4852	1.9690	2.1531	1.9847	2.4715	0.8300	$\hat{\beta}_{32}$
1.1820	1.8737	2.4897	1.8935	2.5468	0.8757	$\hat{\beta}_{33}$
1.3442	1.9952	2.2082	2.1257	2.4079	0.9767	$\hat{\beta}_{34}$
1.4492	1.9263	2.8684	1.8074	2.5406	0.9017	$\hat{\beta}_{35}$
1.4447	1.9991	3.3140	2.2836	2.4666	1.0304	$\hat{\beta}_{36}$
1.5078	1.7407	2.9221	1.9049	2.5942	0.7118	$\hat{\beta}_{37}$
1.5851	2.0089	3.0662	2.0762	2.6417	0.6914	$\hat{\beta}_{38}$
1.4247	1.8554	2.5458	2.0856	3.1498	0.4063	$\hat{\beta}_{39}$
3.2871	1.0640	17.0775	1.8715	255.5645	17.5718	$\hat{\beta}_{40}$
35.1645	—	118.5234	—	306.1451	—	MSE
MM – BR		MM		LASSO		
12.3930		15.2535		68.6044		$\sqrt{\text{MPE}}$
5.5357		4.4976		1.0000		REff
6193.5748		6193.5748		6193.5748		MCN
13.2500		0.8700		19.7800		C
7.6200		0.6700		14.8600		IC

جدول ۱۱.۳: نتایج شبیه‌سازی مربوط به لاسو استوار، طرح ۲، $n = 300$ ، $\rho = 0.3$ و $\tau = 0$

LAD – LASSO		MM		LASSO		ضرایب
MSE	mean	MSE	mean	MSE	mean	
0.0058	-0.0055	0.0072	0.0018	0.0003	0.0014	$\hat{\beta}_1$
0.0067	-0.0059	0.0073	-0.0067	0.0000	-0.0008	$\hat{\beta}_2$
0.0091	0.0148	0.0102	0.0151	0.0001	0.0008	$\hat{\beta}_3$
0.0092	-0.0218	0.0090	-0.0196	0.0001	-0.0004	$\hat{\beta}_4$
0.0085	0.0028	0.0094	0.0049	0.0000	0.0004	$\hat{\beta}_5$
0.0062	0.0101	0.0070	0.0101	0.0001	0.0001	$\hat{\beta}_6$
0.0070	-0.0058	0.0082	-0.0034	0.0000	0.0000	$\hat{\beta}_7$
0.0090	-0.0034	0.0106	-0.0007	0.0001	0.0002	$\hat{\beta}_8$
0.0096	0.0026	0.0113	-0.0047	0.0002	0.0027	$\hat{\beta}_9$
0.0096	0.0027	0.0091	0.0080	0.0002	0.0014	$\hat{\beta}_{10}$
0.0110	1.9941	0.0081	1.9996	0.1616	1.6617	$\hat{\beta}_{11}$
0.0130	1.9990	0.0106	1.9951	0.0642	1.8163	$\hat{\beta}_{12}$
0.0096	1.9951	0.0072	1.9935	0.0628	1.8187	$\hat{\beta}_{13}$
0.0111	1.9786	0.0085	1.9841	0.0509	1.8370	$\hat{\beta}_{14}$
0.0102	2.0093	0.0089	2.0093	0.0568	1.8361	$\hat{\beta}_{15}$
0.0123	1.9913	0.0107	1.9856	0.0489	1.8440	$\hat{\beta}_{16}$
0.0101	1.9937	0.0081	2.0010	0.0645	1.8260	$\hat{\beta}_{17}$
0.0113	2.0121	0.0099	2.0100	0.0585	1.8383	$\hat{\beta}_{18}$
0.0126	1.9806	0.0117	1.9900	0.0537	1.8392	$\hat{\beta}_{19}$
0.0097	2.0066	0.0076	2.0002	0.1598	1.6795	$\hat{\beta}_{20}$
0.0076	0.0072	0.0068	0.0060	0.0001	0.0010	$\hat{\beta}_{21}$
0.0082	0.0108	0.0094	0.0175	0.0000	0.0002	$\hat{\beta}_{22}$
0.0110	-0.0019	0.0099	-0.0101	0.0004	0.0029	$\hat{\beta}_{23}$
0.0085	-0.0022	0.0088	-0.0010	0.0000	0.0002	$\hat{\beta}_{24}$
0.0077	-0.0012	0.0089	-0.0001	0.0001	0.0008	$\hat{\beta}_{25}$
0.0100	0.0278	0.0096	0.0186	0.0000	0.0005	$\hat{\beta}_{26}$
0.0082	-0.0176	0.0081	-0.0128	0.0000	0.0008	$\hat{\beta}_{27}$
0.0078	0.0089	0.0087	0.0185	0.0000	-0.0001	$\hat{\beta}_{28}$
0.0068	-0.0102	0.0072	-0.0150	0.0002	0.0014	$\hat{\beta}_{29}$
0.0088	0.0206	0.0079	0.0170	0.0000	0.0000	$\hat{\beta}_{30}$
0.0115	1.9964	0.0087	1.9976	0.1692	1.6626	$\hat{\beta}_{31}$
0.0112	1.9958	0.0090	1.9890	0.0763	1.8094	$\hat{\beta}_{32}$
0.0083	1.9866	0.0080	1.9984	0.0598	1.8241	$\hat{\beta}_{33}$
0.0105	1.9971	0.0085	2.0014	0.0598	1.8272	$\hat{\beta}_{34}$
0.0107	1.9994	0.0095	2.0043	0.0502	1.8400	$\hat{\beta}_{35}$
0.0104	1.9950	0.0090	1.9999	0.0699	1.8301	$\hat{\beta}_{36}$
0.0093	1.9867	0.0080	1.9904	0.0747	1.8013	$\hat{\beta}_{37}$
0.0087	2.0002	0.0088	1.9980	0.0525	1.8402	$\hat{\beta}_{38}$
0.0083	2.0064	0.0084	1.9979	0.0602	1.8367	$\hat{\beta}_{39}$
0.0115	2.0032	0.0095	1.9968	0.3589	2.4964	$\hat{\beta}_{40}$
0.3767	—	0.3536	—	1.8150	—	MSE
		MM – BR		MM	LASSO	
		18.7125		18.5801	34.6854	$\sqrt{\text{MPE}}$
		1.8536		1.8668	1.0000	REF
		64.2164		64.2164	64.2164	MCN
		14.4600		14.5800	19.9400	C
		7.2800		7.2700	9.9700	IC

جدول ۱۲.۳: نتایج شبیه‌سازی مربوط به لاسو استوار، طرح ۲، $n = 300$ ، $\rho = 0.99$ و $\tau = 0$

LAD – LASSO		MM		LASSO		ضرایب
MSE	mean	MSE	mean	MSE	mean	
0/1167	0/0409	0/3293	0/0124	0/0000	0/0000	$\hat{\beta}_1$
0/1396	0/0080	0/3338	-0/0456	0/0000	0/0000	$\hat{\beta}_2$
0/1516	0/1014	0/4665	0/1020	0/0000	0/0000	$\hat{\beta}_3$
0/1210	-0/0085	0/4110	-0/1329	0/0000	0/0000	$\hat{\beta}_4$
0/1627	0/0849	0/4301	0/0334	0/0000	0/0000	$\hat{\beta}_5$
0/0929	0/0768	0/3184	0/0683	0/0000	0/0000	$\hat{\beta}_6$
0/1365	0/0281	0/3770	-0/0231	0/0000	0/0003	$\hat{\beta}_7$
0/1156	0/0678	0/4858	-0/0046	0/0000	0/0000	$\hat{\beta}_8$
0/1668	0/0787	0/5172	-0/0315	0/0000	0/0008	$\hat{\beta}_9$
0/2157	0/1449	0/4175	0/0541	0/0039	0/0088	$\hat{\beta}_{10}$
0/5190	1/8150	0/3685	1/9973	2/3268	0/6037	$\hat{\beta}_{11}$
0/5648	2/0098	0/4857	1/9669	1/4039	1/1213	$\hat{\beta}_{12}$
0/4314	1/9177	0/3294	1/9560	1/2602	1/1887	$\hat{\beta}_{13}$
0/4184	1/8700	0/3894	1/8924	1/2123	1/2444	$\hat{\beta}_{14}$
0/4252	2/0695	0/4083	2/0629	1/1712	1/3087	$\hat{\beta}_{15}$
0/5417	1/9515	0/4907	1/9025	1/4247	1/2314	$\hat{\beta}_{16}$
0/3907	1/9828	0/3713	2/0070	1/2934	1/3250	$\hat{\beta}_{17}$
0/5435	2/0681	0/4528	2/0675	1/2832	1/3607	$\hat{\beta}_{18}$
0/6158	1/8092	0/5353	1/9322	1/3732	1/1919	$\hat{\beta}_{19}$
0/3902	2/0031	0/3494	2/0014	2/4586	0/5778	$\hat{\beta}_{20}$
0/1421	0/1090	0/3114	0/0405	0/0002	0/0014	$\hat{\beta}_{21}$
0/1751	0/1069	0/4294	0/1183	0/0000	0/0000	$\hat{\beta}_{22}$
0/2365	0/0464	0/4545	-0/0683	0/0000	0/0000	$\hat{\beta}_{23}$
0/1031	0/0667	0/4004	-0/0069	0/0000	0/0000	$\hat{\beta}_{24}$
0/1395	0/0683	0/4072	-0/0010	0/0000	0/0000	$\hat{\beta}_{25}$
0/2009	0/1357	0/4381	0/1260	0/0000	0/0000	$\hat{\beta}_{26}$
0/1334	0/0121	0/3704	-0/0865	0/0000	0/0000	$\hat{\beta}_{27}$
0/1138	0/0659	0/3996	0/1251	0/0000	0/0000	$\hat{\beta}_{28}$
0/1318	0/0294	0/3296	-0/1011	0/0000	0/0000	$\hat{\beta}_{29}$
0/1937	0/2007	0/3632	0/1152	0/0006	0/0024	$\hat{\beta}_{30}$
0/4621	1/8707	0/3981	1/9841	2/6096	0/4889	$\hat{\beta}_{31}$
0/4484	1/9396	0/4115	1/9253	1/5787	1/1263	$\hat{\beta}_{32}$
0/3358	1/9626	0/3658	1/9895	1/5938	1/1221	$\hat{\beta}_{33}$
0/4754	1/9419	0/3900	2/0095	1/3482	1/3799	$\hat{\beta}_{34}$
0/4658	1/9671	0/4325	2/0292	1/3234	1/2501	$\hat{\beta}_{35}$
0/5043	1/9567	0/4135	1/9991	1/4812	1/2622	$\hat{\beta}_{36}$
0/4310	1/9448	0/3646	1/9350	1/5361	1/1412	$\hat{\beta}_{37}$
0/4303	1/9781	0/4042	1/9864	1/7587	0/9749	$\hat{\beta}_{38}$
0/3506	1/9984	0/3848	1/9861	2/5560	0/5215	$\hat{\beta}_{39}$
2/0709	1/5391	2/6240	2/0601	173/0411	14/5691	$\hat{\beta}_{40}$
13/8046	—	18/3601	—	204/0390	—	MSE
		MM – BR		MM	LASSO	
		18/2760		18/5801	66/9599	$\sqrt{\text{MPE}}$
		3/6638		3/6038	1/0000	REff
		4050/7410		4050/7410	4050/7410	MCN
		10/0100		2/6500	19/9600	C
		5/0700		1/4000	11/8300	IC

جدول ۱۳.۳: نتایج شبیه‌سازی مربوط به لاسو استوار، طرح ۲، $n = 100$ ، $\rho = 0.3$ و $\tau = 0.2$

LAD – LASSO		MM		LASSO		ضرایب
MSE	mean	MSE	mean	MSE	mean	
0/0۳9۸	−0/0۳۲۴	0/0۶۱۶	−0/0۲9۳	۱/۱۷9۶	0/۴9۷۳	$\hat{\beta}_1$
0/0۵۱9	0/0۳۱۸	0/0۸۷9	0/0۶۲۷	0/۱۱۰۸	0/0۶۳۴	$\hat{\beta}_2$
0/0۸۷۰	−0/0۱۵۲	0/۱۴۶۲	0/0۰۸۴	0/۶۳۲۵	0/۲۷۰۴	$\hat{\beta}_3$
0/0۶۸۴	−0/0۳۶۳	0/۱۱۶۷	−0/0۲۰۳	0/09۶۳	0/0۶۱۷	$\hat{\beta}_4$
0/0۸۱۰	0/0۳۵۰	0/۱۳۴۷	0/0۲۳۳	0/0۰۰۰	0/0۰۰۰	$\hat{\beta}_5$
0/0۶9۵	0/0۷۰۵	0/۱۲۷۲	0/۱۰۸۶	0/0۰۰۰	0/0۰۰۰	$\hat{\beta}_6$
0/0۴۶۳	−0/0۳۲۱	0/۰۸۲۶	−0/0۵۲۶	0/9۱۰۰	0/۳۶۱۴	$\hat{\beta}_7$
0/0۵۱۳	0/0۳۴۳	0/0۸۶۱	0/0۴۴۵	0/0۰۰۰	0/0۰۰۰	$\hat{\beta}_8$
0/0۵۱۵	0/0۰۰۳	0/0۷۶۸	0/0۰۷۵	۱/۱۷9۵	0/۳۵۴۲	$\hat{\beta}_9$
0/0۶۸۵	0/۱۲۳۶	0/۱۱۵۴	0/۱۵۳9	0/0۰۰۰	0/0۰۰۰	$\hat{\beta}_{10}$
0/09۷۰	۱/۸۷۵۷	0/۱۳۵9	۱/۸۷۱9	۳/0۷۲۸	۱/۱۵۳9	$\hat{\beta}_{11}$
0/0۵99	۱/9۵۰۰	0/0۸۶۸	۱/9۲۴۱	۳/۸۵۱9	0/0۸۲۱	$\hat{\beta}_{12}$
0/0۸۴۱	۱/9۷۶۳	0/۱۲۰۰	۱/9۷۸۵	۳/۷۴۳۶	0/۱۰۱۴	$\hat{\beta}_{13}$
0/۱۰۳۰	۱/۸۸۲۳	0/۱۴۱۴	۱/۸۵9۷	۳/۲۱۸۶	۱/۵۸۱9	$\hat{\beta}_{14}$
0/09۵۱	۲/0۳۳۲	0/۱۰9۱	۲/۰۱9۱	۴/۰۰۰۰	0/0۰۰۰	$\hat{\beta}_{15}$
0/09۸۶	۲/0۰۲۸	0/۱۱۱۶	۲/۰۱۸۵	۴/0۰۰۰	0/0۰۰۰	$\hat{\beta}_{16}$
0/۱۰۵۷	۱/۸۳۴۴	0/۱۲۱9	۱/۸۴۳۱	۴/۶۶۸۳	۱/9۴۳۰	$\hat{\beta}_{17}$
0/0۷۸۴	۲/0۵9۰	0/۱۱۷9	۲/0۴۴9	۴/0۰۰۰	0/0۰۰۰	$\hat{\beta}_{18}$
0/0۸۱۱	۱/9۶۰۶	0/۱۲۸۲	۱/9۴۲۱	۴/0۰۰۰	0/0۰۰۰	$\hat{\beta}_{19}$
0/۱۱۱۱	۱/۸۵۲9	0/۱۵۲۴	۱/۸۴۷۰	۴/0۰۰۰	0/0۰۰۰	$\hat{\beta}_{20}$
0/0۷۶۷	0/۱۵۱۳	0/۱۴۸۱	0/۲۱۳۰	0/0۰۰۰	0/0۰۰۰	$\hat{\beta}_{21}$
0/0۵۶۵	0/0۰۶9	0/۱۱۶۸	−0/0۱۳۱	0/0۰۰۰	0/0۰۰۰	$\hat{\beta}_{22}$
0/0۴۲۵	0/0۵۱۱	0/09۲۰	0/09۳۲	0/0۰۰۰	0/0۰۰۰	$\hat{\beta}_{23}$
0/0۴9۵	0/0۴۷۵	0/۱۰۱۶	0/0۲۳۴	0/0۰۰۰	0/0۰۰۰	$\hat{\beta}_{24}$
0/0۴۵۳	0/0۰۲۲	0/09۶۸	0/0۱۰۸	0/0۰۰۰	0/0۰۰۰	$\hat{\beta}_{25}$
0/0۴۲۵	−0/0۰۴۱	0/0۸9۸	0/0۲۴۰	0/۲9۵۱	0/0۷۶۷	$\hat{\beta}_{26}$
0/0۵۱۰	−0/0۵۸9	0/0۸۵9	−0/0۶۷۶	۱/۷۸9۷	0/۶۰۴9	$\hat{\beta}_{27}$
0/0۵۳۵	0/0۶۱۳	0/۱۱۱۱	0/09۴۴	0/0۰۰۰	0/0۰۰۰	$\hat{\beta}_{28}$
0/0۵۲9	0/0۲۷۱	0/099۷	0/0۱۳۵	0/0۰۰۰	0/0۰۰۰	$\hat{\beta}_{29}$
0/0۸۰۸	0/۱۳9۵	0/۱۳۶۱	0/۱۸۲۸	0/0۰۰۰	0/0۰۰۰	$\hat{\beta}_{30}$
0/۱۲۰۴	۱/۷۷9۰	0/۱۷۱۴	۱/۷۳۸۱	۳/۵۱۳۳	۲/۱۷۱۲	$\hat{\beta}_{31}$
0/0۷۳۶	۱/9۷۳۱	0/09۶۳	۱/9۴۶۰	۳/۶۷۱۳	0/۱۸۲9	$\hat{\beta}_{32}$
0/0۸۴۵	۲/0۲۰۴	0/۱۳۰۰	۲/0۲۸۳	۳/9۲۸۷	0/09۰۰	$\hat{\beta}_{33}$
0/09۱۱	۱/9۰۶۵	0/۱۱۲۲	۱/9۳۰۸	۲/9۷۸۳	۱/۴۰۰۸	$\hat{\beta}_{34}$
0/0۸۶۸	۲/0۶۶۴	0/۱۰۱۰	۲/0۶۳9	۴/0۰۰۰	0/0۰۰۰	$\hat{\beta}_{35}$
0/۱۰۲۷	۱/9۴۴۸	0/۱۱۶۰	۱/9۶۷۶	۳/۷9۸۷	0/۲۱۴۷	$\hat{\beta}_{36}$
0/0۸۶۲	۱/9۰۶9	0/۱۲۴۱	۱/۸۸99	۳/۶۲۷۱	۱/۶۰۲۷	$\hat{\beta}_{37}$
0/09۸۰	۲/0۰۷۳	0/۱۲۳۷	۲/0۰۷۴	۳/99۲۷	0/0۰۱9	$\hat{\beta}_{38}$
0/0۷9۷	۱/9۸۵۱	0/0۸۶۱	۱/99۲۰	۳/9۰۶۴	0/0۴۲۸	$\hat{\beta}_{39}$
0/0۸۶۲	۱/9۸9۱	0/۱۲۰۷	۱/9۲۶۵	۳/۱۱۴۸	۱/۴۰۸۴	$\hat{\beta}_{40}$
۲/9۸9۴	−	۴/۵۱9۶	−	۸۱/۲۸۰۱	−	MSE
LAD – LASSO		MM		LASSO		$\sqrt{\text{MPE}}$
۳۶/9۴۰۳		۳۷/۶۴۳۰		۵۰۲/۲۰۵۸		
۱۳/۵9۵۱		۱۳/۳۴۱۳		۱/۰۰۰۰		
9۲۱/۱۳۲۸		9۲۱/۱۳۲۸		9۲۱/۱۳۲۸		
۸/۱۳۰۰		۵/۰۰۰۰		۱۸/۵۷۰۰		
۴/۰۱۰۰		۲/۵۰۰۰		۱۶/۸۷۰۰		IC

جدول ۱۴.۳: نتایج شبیه‌سازی مربوط به لاسو استوار، طرح ۲، $n = ۱۰۰$ ، $\rho = ۰/۹۹$ و $\tau = ۰/۲$

LAD – LASSO		MM		LASSO		ضرایب
MSE	mean	MSE	mean	MSE	mean	
۰/۴۸۲۰	۰/۳۹۸۳	۰/۸۰۹۶	۰/۲۸۳۵	۰/۲۸۳۲	۰/۱۱۷۷	$\hat{\beta}_1$
۰/۷۷۳۴	۰/۴۷۲۵	۱/۰۰۰۱	۰/۳۶۴۹	۲/۵۶۶۰	۰/۷۴۸۴	$\hat{\beta}_2$
۰/۷۴۸۵	۰/۳۸۶۰	۱/۲۹۴۳	۰/۱۶۰۴	۰/۳۹۷۹	۰/۱۷۴۴	$\hat{\beta}_3$
۰/۶۰۴۰	۰/۲۵۷۱	۱/۲۵۴۳	۰/۰۹۲۲	۳/۳۳۶۷	۰/۷۶۳۲	$\hat{\beta}_4$
۰/۵۹۸۰	۰/۳۶۸۹	۱/۳۷۳۹	۰/۲۵۴۰	۰/۳۳۳۱	۰/۰۸۸۹	$\hat{\beta}_5$
۰/۴۹۳۴	۰/۳۳۸۵	۱/۰۳۷۱	۰/۲۹۴۷	۰/۰۰۰۰	۰/۰۰۰۰	$\hat{\beta}_6$
۰/۵۱۸۴	۰/۳۱۰۷	۰/۹۲۳۶	۰/۱۱۰۸	۱/۹۲۲۳	۰/۶۰۷۳	$\hat{\beta}_7$
۰/۷۵۸۶	۰/۳۵۷۹	۱/۱۶۳۹	۰/۲۱۰۶	۰/۱۴۱۳	۰/۰۳۷۶	$\hat{\beta}_8$
۰/۶۴۲۷	۰/۴۱۴۴	۰/۹۰۸۹	۰/۲۶۷۲	۲/۲۲۳۷	۰/۶۵۱۶	$\hat{\beta}_9$
۱/۰۴۲۳	۰/۶۶۳۱	۱/۲۸۱۴	۰/۴۶۲۸	۰/۰۰۰۰	۰/۰۰۰۰	$\hat{\beta}_{10}$
۱/۰۱۸۷	۱/۴۰۱۷	۱/۱۴۷۳	۱/۷۲۱۴	۴/۳۸۵۱	۰/۶۳۸۲	$\hat{\beta}_{11}$
۱/۱۶۱۸	۱/۵۱۹۳	۱/۳۳۵۸	۱/۶۱۵۷	۴/۱۳۴۰	۰/۴۳۵۱	$\hat{\beta}_{12}$
۱/۰۹۱۶	۱/۷۲۸۹	۱/۲۲۱۷	۱/۸۸۱۸	۴/۹۰۹۰	۰/۸۴۸۹	$\hat{\beta}_{13}$
۱/۲۰۹۶	۱/۵۳۰۹	۱/۵۵۴۹	۱/۶۵۵۴	۳/۹۸۰۱	۰/۵۷۲۶	$\hat{\beta}_{14}$
۱/۰۲۴۵	۱/۷۰۶۴	۱/۱۹۸۶	۱/۸۶۸۴	۴/۰۰۰۰	۰/۰۰۰۰	$\hat{\beta}_{15}$
۰/۸۹۴۰	۱/۷۴۱۷	۱/۱۸۵۳	۱/۸۳۹۲	۴/۰۰۰۰	۰/۰۰۰۰	$\hat{\beta}_{16}$
۰/۸۴۷۶	۱/۷۰۰۴	۱/۰۸۳۹	۱/۸۴۲۱	۴/۱۲۸۶	۰/۹۷۶۴	$\hat{\beta}_{17}$
۱/۲۰۱۷	۱/۵۴۵۰	۱/۳۷۶۱	۱/۷۲۱۸	۴/۰۰۰۰	۰/۰۰۰۰	$\hat{\beta}_{18}$
۱/۰۵۹۸	۱/۶۸۵۸	۱/۴۹۹۷	۱/۸۵۲۳	۴/۰۰۰۰	۰/۰۰۰۰	$\hat{\beta}_{19}$
۱/۲۳۱۸	۱/۳۵۷۵	۱/۳۱۵۱	۱/۵۳۶۸	۴/۰۰۰۰	۰/۰۰۰۰	$\hat{\beta}_{20}$
۱/۱۸۶۴	۰/۷۳۲۴	۱/۵۶۶۰	۰/۶۱۶۴	۰/۰۰۰۰	۰/۰۰۰۰	$\hat{\beta}_{21}$
۰/۷۸۸۴	۰/۳۹۲۶	۱/۴۴۴۴	۰/۱۱۹۶	۰/۰۰۰۲۱	۰/۰۰۰۴۶	$\hat{\beta}_{22}$
۰/۷۶۲۹	۰/۵۰۹۱	۱/۲۲۷۷	۰/۴۶۹۴	۰/۰۰۰۰	۰/۰۰۰۰	$\hat{\beta}_{23}$
۰/۴۱۱۴	۰/۲۹۷۵	۰/۷۴۹۸	۰/۱۲۰۴	۰/۰۰۰۰	۰/۰۰۰۰	$\hat{\beta}_{24}$
۰/۴۳۳۶	۰/۳۰۵۳	۱/۱۵۰۸	۰/۲۸۸۸	۰/۹۸۰۱	۰/۲۲۵۹	$\hat{\beta}_{25}$
۰/۵۷۸۹	۰/۲۸۱۷	۱/۰۰۶۴	–۰/۰۰۵۸	۳/۶۹۵۶	۰/۶۸۳۰	$\hat{\beta}_{26}$
۰/۴۱۳۲	۰/۲۷۲۴	۰/۸۵۸۱	۰/۱۹۱۵	۰/۱۰۰۰	۰/۰۶۹۹	$\hat{\beta}_{27}$
۰/۵۶۴۰	۰/۳۹۴۰	۰/۹۸۶۱	۰/۲۷۶۷	۰/۰۰۰۰	۰/۰۰۰۰	$\hat{\beta}_{28}$
۰/۷۶۷۷	۰/۴۴۱۲	۱/۱۱۰۸	۰/۳۲۳۷	۰/۰۸۸۰	۰/۰۴۱۷	$\hat{\beta}_{29}$
۱/۱۸۳۷	۰/۷۴۳۰	۱/۶۵۹۶	۰/۶۹۱۴	۶/۴۶۳۵	۰/۹۰۵۶	$\hat{\beta}_{30}$
۱/۳۵۹۸	۱/۱۶۱۲	۱/۸۰۵۸	۱/۲۱۳۷	۳/۹۰۲۷	۰/۰۴۴۳	$\hat{\beta}_{31}$
۱/۰۵۹۰	۱/۵۲۱۹	۱/۱۶۹۵	۱/۷۰۷۰	۵/۶۰۲۰	۱/۱۳۶۰	$\hat{\beta}_{32}$
۱/۰۸۶۵	۱/۷۴۱۲	۱/۳۱۰۲	۱/۹۷۱۴	۴/۴۱۷۹	۰/۶۳۸۱	$\hat{\beta}_{33}$
۰/۸۴۵۳	۱/۸۸۱۲	۱/۲۱۶۸	۱/۹۵۵۱	۴/۰۳۶۲	۰/۶۹۹۶	$\hat{\beta}_{34}$
۰/۸۱۱۲	۱/۶۲۹۵	۰/۸۸۰۱	۱/۷۱۹۴	۴/۰۰۰۰	۰/۰۰۰۰	$\hat{\beta}_{35}$
۱/۱۰۶۶	۱/۷۶۹۳	۱/۳۶۴۴	۱/۸۷۸۷	۶/۳۹۷۰	۱/۱۲۲۶	$\hat{\beta}_{36}$
۰/۸۳۱۴	۱/۶۹۸۵	۰/۹۴۴۳	۱/۸۳۷۵	۵/۰۵۲۹	۰/۸۴۱۵	$\hat{\beta}_{37}$
۰/۹۰۴۶	۱/۷۷۸۱	۱/۰۰۸۳	۱/۸۷۲۱	۴/۹۰۹۹	۰/۴۱۷۴	$\hat{\beta}_{38}$
۱/۱۰۸۴	۱/۵۵۵۰	۱/۲۲۱۳	۱/۸۰۸۵	۴/۳۷۸۱	۰/۵۱۸۶	$\hat{\beta}_{39}$
۲/۱۵۲۸	۱/۱۴۴۱	۳/۹۹۰۷	۱/۰۴۲۶	۱۴/۵۶۹۲	۳/۲۴۹۱	$\hat{\beta}_{40}$
۳۵/۷۵۸۱	–	۵۰/۶۳۶۷	–	۱۲۱/۳۳۶۲	–	MSE
			MM – BR	MM	LASSO	$\sqrt{\text{MPE}}$
			۳۳/۸۷۰۷	۴۱/۶۴۹۹	۴۹۹/۱۰۹۹	REF
			۱۴/۷۳۵۷	۱۱/۹۸۳۵	۱/۰۰۰۰	MCN
			۳۶۳۸۴/۸۳۳۳	۳۶۳۸۴/۸۳۳۳	۳۶۳۸۴/۸۳۳۳	C
			۶/۹۲۰۰	۱/۶۵۰۰	۱۷/۹۷۰۰	IC
			۴/۰۴۰۰	۱/۰۹۰۰	۱۷/۲۸۰۰	

جدول ۱۵.۳: نتایج شبیه‌سازی مربوط به لاسو استوار، طرح ۲، $n = 300$ ، $\rho = 0.3$ و $\tau = 0.2$

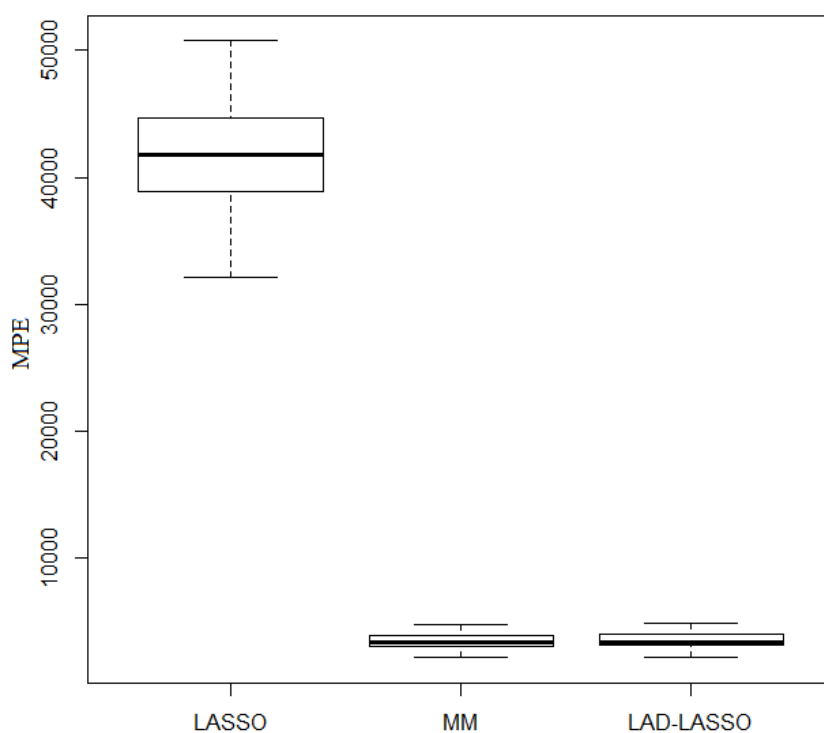
LAD – LASSO		MM		LASSO		ضرایب
MSE	mean	MSE	mean	MSE	mean	
0/0090	-0/0526	0/0124	-0/0368	0/5742	0/3323	$\hat{\beta}_1$
0/0090	0/0236	0/0153	0/0342	0/0066	0/0081	$\hat{\beta}_2$
0/0116	-0/0327	0/0145	-0/0264	0/1866	0/1387	$\hat{\beta}_3$
0/0109	-0/0162	0/0160	-0/0162	0/0090	0/0153	$\hat{\beta}_4$
0/0075	0/0013	0/0136	0/0082	0/0000	0/0000	$\hat{\beta}_5$
0/0152	0/0637	0/0211	0/0742	0/0000	0/0000	$\hat{\beta}_6$
0/0155	-0/0718	0/0193	-0/0661	0/0770	0/0645	$\hat{\beta}_7$
0/0107	0/0538	0/0186	0/0609	0/0000	0/0000	$\hat{\beta}_8$
0/0136	-0/0702	0/0181	-0/0604	0/0749	0/0894	$\hat{\beta}_9$
0/0222	0/1174	0/0339	0/1471	0/0000	0/0000	$\hat{\beta}_{10}$
0/0278	1/8670	0/0391	1/8415	1/8958	1/1952	$\hat{\beta}_{11}$
0/0100	2/0068	0/0127	1/9976	3/9838	0/0046	$\hat{\beta}_{12}$
0/0082	1/9701	0/0130	1/9615	3/8818	0/0396	$\hat{\beta}_{13}$
0/0187	1/9084	0/0219	1/9091	1/6507	1/8990	$\hat{\beta}_{14}$
0/0105	2/0340	0/0149	2/0299	4/0000	0/0000	$\hat{\beta}_{15}$
0/0097	1/9943	0/0127	1/9891	4/0000	0/0000	$\hat{\beta}_{16}$
0/0179	1/9071	0/0219	1/9025	2/2253	1/5798	$\hat{\beta}_{17}$
0/0113	2/0477	0/0139	2/0389	4/0000	0/0000	$\hat{\beta}_{18}$
0/0100	1/9499	0/0154	1/9505	4/0000	0/0000	$\hat{\beta}_{19}$
0/0113	1/9530	0/0192	1/9159	4/0000	0/0000	$\hat{\beta}_{20}$
0/0160	0/0862	0/0294	0/1227	0/0000	0/0000	$\hat{\beta}_{21}$
0/0099	-0/0138	0/0177	-0/0053	0/0000	0/0000	$\hat{\beta}_{22}$
0/0091	0/0152	0/0166	0/0310	0/0000	0/0000	$\hat{\beta}_{23}$
0/0120	0/0493	0/0191	0/0633	0/0000	0/0000	$\hat{\beta}_{24}$
0/0083	-0/0300	0/0150	-0/0337	0/0000	0/0000	$\hat{\beta}_{25}$
0/0096	0/0294	0/0154	0/0467	0/0000	0/0000	$\hat{\beta}_{26}$
0/0135	-0/0829	0/0167	-0/0864	0/8338	0/4679	$\hat{\beta}_{27}$
0/0114	0/0576	0/0196	0/0762	0/0000	0/0000	$\hat{\beta}_{28}$
0/0095	0/0162	0/0150	0/0156	0/0000	0/0000	$\hat{\beta}_{29}$
0/0095	0/0494	0/0186	0/0888	0/0000	0/0000	$\hat{\beta}_{30}$
0/0253	1/8734	0/0382	1/8472	2/7452	3/0367	$\hat{\beta}_{31}$
0/0097	1/9799	0/0133	1/9747	3/8200	0/0861	$\hat{\beta}_{32}$
0/0096	1/9826	0/0112	1/9749	4/0000	0/0000	$\hat{\beta}_{33}$
0/0201	1/9015	0/0251	1/9006	2/0759	1/9644	$\hat{\beta}_{34}$
0/0165	2/0909	0/0195	2/0912	4/0000	0/0000	$\hat{\beta}_{35}$
0/0167	1/9266	0/0223	1/9130	3/9517	0/0149	$\hat{\beta}_{36}$
0/0131	1/9536	0/0149	1/9500	1/5376	1/6631	$\hat{\beta}_{37}$
0/0113	1/9914	0/0145	1/9839	4/0000	0/0000	$\hat{\beta}_{38}$
0/0100	1/9803	0/0128	1/9741	3/9507	0/0163	$\hat{\beta}_{39}$
0/0139	2/0620	0/0144	2/0388	1/6922	1/4960	$\hat{\beta}_{40}$
0/5155	—	0/7367	—	67/1728	—	MSE
						MM – BR
						MM
						LASSO
						$\sqrt{MPE}$
						REF
						MCN
						C
						IC

جدول ۱۶.۳: نتایج شبیه‌سازی مربوط به لاسو استوار، طرح ۲، $n = 300$ ، $\rho = 0.99$ و $\tau = 0.2$

LAD – LASSO		MM		LASSO		ضرایب
MSE	mean	MSE	mean	MSE	mean	
0/۵۳۶۴	0/۵۹۹۲	0/۵۶۹۸	0/۶۷۳۹	0/0۹۱۹	0/0۷۹۳	$\hat{\beta}_1$
0/۵۵۹۲	0/۶۲۱۷	0/۶۳۱۷	0/۷0۴۲	۲/۲۳۸۶	0/۸۳۱0	$\hat{\beta}_2$
0/۴۹۲۵	0/۵۷۵۸	0/۵۴0۱	0/۶۴۴۴	0/۳۶۱0	0/۱۳۸0	$\hat{\beta}_3$
0/۴۱۶۲	0/۴۸۳0	0/۴۸۷0	0/۵۸۱۶	۲/۲۵0۷	0/۷۷۶۹	$\hat{\beta}_4$
0/۴۷۹۷	0/۵۵0۳	0/۴۹۶۲	0/۶0۵۵	0/۱۳۶۱	0/0۳۶۹	$\hat{\beta}_5$
0/۵0۶۴	0/۶0۲۳	0/۵۲۷۶	0/۶۶۲۷	0/0000	0/0000	$\hat{\beta}_6$
0/۴۵۱۳	0/۵۲۹۹	0/۵۱۸۵	0/۶۲۲۲	۱/0۶۹۵	0/۵0۶۸	$\hat{\beta}_7$
0/۵۴۶۸	0/۶0۲۲	0/۵۷0۵	0/۶۷۱۹	0/0000	0/0000	$\hat{\beta}_8$
0/۶0۶۸	0/۶۴۳0	0/۵۹۷۳	0/۶۸۳۱	0/۸۷۷۸	0/۴۵0۲	$\hat{\beta}_9$
0/۷۵0۶	0/۷۸۳۹	0/۷۹۴۵	0/۸۳۳۹	0/0000	0/0000	$\hat{\beta}_{10}$
0/۷۹۵۸	۱/۲۱۹۶	0/۷۹۶۵	۱/۱۷۹۲	۳/0۶۴0	0/۴۴0۹	$\hat{\beta}_{11}$
0/۵۵۲۷	۱/۴۲۱۵	0/۵۵۸۹	۱/۳۵۹۳	۳/۱۹۱۹	0/۷۶۸۱	$\hat{\beta}_{12}$
0/۵۳۴۲	۱/۳۷۷0	0/۵۶۲۲	۱/۳0۹۱	۲/۷۸۸۷	0/۹۹۹۶	$\hat{\beta}_{13}$
0/۴۶۴۵	۱/۴۵۳۷	0/۵۲۱۴	۱/۳۷۴۴	۲/۹۲۵۹	0/۶0۹۸	$\hat{\beta}_{14}$
0/۴0۳۹	۱/۵۳۵۷	0/۴۳0۷	۱/۴۵۲۹	۴/0000	0/0000	$\hat{\beta}_{15}$
0/۴۹0۱	۱/۵0۸۷	0/۴۶۹0	۱/۴۳۱۷	۴/0000	0/0000	$\hat{\beta}_{16}$
0/۴۶۴۱	۱/۴۵۹۱	0/۴۹۳۴	۱/۳۹0۵	۲/۹۳۵۱	۱/۱۵۳۶	$\hat{\beta}_{17}$
0/۵۱۹۶	۱/۴۱۴۸	0/۵۵۲۲	۱/۳۵۶۹	۴/0000	0/0000	$\hat{\beta}_{18}$
0/۴۸0۶	۱/۴۴۴۱	0/۵0۲۲	۱/۳۷۹۲	۴/0000	0/0000	$\hat{\beta}_{19}$
0/۷۴0۱	۱/۲۴۴۱	0/۷۴۴۸	۱/۲0۴۴	۴/0000	0/0000	$\hat{\beta}_{20}$
0/۹۲۳۸	0/۸۶۹۲	0/۸۹۸۲	0/۸۸۵۶	0/0000	0/0000	$\hat{\beta}_{21}$
0/۶۲۲۵	0/۶۶۲۸	0/۶۴۴۷	0/۷۱۸۱	0/0000	0/0000	$\hat{\beta}_{22}$
0/۵۴۵۱	0/۵۹۸0	0/۵۷۸۱	0/۶۷۶۹	0/0000	0/0000	$\hat{\beta}_{23}$
0/۵۷۷۷	0/۵۹۳۲	0/۵۸۷۴	0/۶۵۳۳	0/0000	0/0000	$\hat{\beta}_{24}$
0/۵0۵۹	0/۵۳۶0	0/۵0۷۳	0/۶00۹	0/0۲۲۲	0/0۱۴۹	$\hat{\beta}_{25}$
0/۵۳۳۴	0/۶۱0۶	0/۵۸۴۲	0/۶۷۷۷	۱/۶۹۷۷	0/۵۸۵۶	$\hat{\beta}_{26}$
0/۴۸۱۳	0/۵۷۵۹	0/۴۹۴۹	0/۶۲۴۹	0/0۲۸۴	0/0۲۸۵	$\hat{\beta}_{27}$
0/۵0۷۵	0/۶0۸۳	0/۵۹۴۹	0/۷0۹۶	0/0000	0/0000	$\hat{\beta}_{28}$
0/۵۶۱۹	0/۶۴0۱	0/۶۱۳۱	0/۶۹۹۲	0/00۴۹	0/00۹۹	$\hat{\beta}_{29}$
0/۸۵۲۱	0/۸۴0۶	0/۸۷۳۹	0/۸۷۸۳	۱/۸۶۲۶	0/۷۱۹۴	$\hat{\beta}_{30}$
0/۸۳0۱	۱/۱۷۶۳	0/۸۶۵۹	۱/۱۳0۱	۳/۹۴۶۳	0/0۱۸0	$\hat{\beta}_{31}$
0/۵۶۱۳	۱/۳۷۴۸	0/۵۸۹۲	۱/۳۱۸۴	۳/۶۹۴۲	۱/۲۲۷۳	$\hat{\beta}_{32}$
0/۵۲۲۶	۱/۴۲۲0	0/۵۳۲۹	۱/۳۶۲0	۳/۳۶۵۶	0/۵۲۵۸	$\hat{\beta}_{33}$
0/۴۷00	۱/۴۸۴0	0/۴۷۲0	۱/۴۱۴۳	۲/۸۷۱۹	۱/۲۱۹۷	$\hat{\beta}_{34}$
0/۴۲۶۶	۱/۴۸۳۸	0/۴۵۹۹	۱/۴0۵۵	۴/0000	0/0000	$\hat{\beta}_{35}$
0/۴۹۲۴	۱/۴۷۲۸	0/۵۲۵۹	۱/۳۷۸۸	۳/۳۴۴۷	0/۸۹۹۲	$\hat{\beta}_{36}$
0/۴0۸۹	۱/۵0۱۹	0/۴۵۶۲	۱/۴۱0۷	۲/۹۳۹۴	۱/۱۶۸۲	$\hat{\beta}_{37}$
0/۵۶۲۲	۱/۳۹۱۸	0/۵۶۷۸	۱/۳۳۲۸	۳/۸۷۵۲	0/0۴۶۸	$\hat{\beta}_{38}$
0/۶۹۶۵	۱/۲۹۴۸	0/۶۸۳۸	۱/۲۵۱۱	۳/۳۲۷۶	0/۴۵۶۷	$\hat{\beta}_{39}$
۱/۱۸۶۵	۱/00۴۴	۱/۱۶۴۶	0/۹۷۹۹	۸/۷۴۶۹	۳/۸۵۶۵	$\hat{\beta}_{40}$
۲۳/0۵۹۹	—	۲۴/0۵۹۳	—	۸۵/۶۶۸۶	—	MSE
			MM – BR	MM	LASSO	
			۲۳/۲۳۴۶	۲۹/۹۲۷۱	۸۳۷/۵۷۹۲	$\sqrt{MPE}$
			۲۵/۲0۲0	۲۷/۹۸۷۴	۱/0000	REF
			۳۷۹۲/۷۷0۳	۳۷۹۲/۷۷0۳	۳۷۹۲/۷۷0۳	MCN
			۲/0۸00	0/۶۹00	۱۷/۳۵00	C
			۱/۱۷00	0/۳۹00	۱۶/0۲00	IC

جدول ۱۷.۳: نتایج مربوط به برآوردگر لاسو استوار برای داده‌های V_{O2}

لاسو	MM – برآوردگر	LAD-LASSO	
۲۰۴/۲۵۳	۵۸/۷۷۵	۵۹/۶۱۳۳	$\sqrt{MPE}$



شکل ۱۰.۳: نمودار جعبه‌ای مربوط به خطاهای پیشگویی برای لاسو استوار

برآوردگر لاسو معمولی عملکرد بسیار ضعیفی نشان می‌دهد و تاثیر داده‌پرت بر این برآوردگر به‌وضوح نمایان است، اما برآوردگرهای استوار نسبت به این برآوردگر عملکرد خوبی نسبت به برآوردگر لاسو معمولی دارند، به این معنی که حضور داده‌های پرت در مجموعه داده تاثیر چندانی بر این برآوردگرها نمی‌گذارد.

۷.۳ برآوردگر استوار فازی با استفاده از جریمه لاسو

تاکنون برآوردگر لاسو استوار یا همان LAD-LASSO را معرفی کردیم اما اگر در مسائل کاربردی شرط دقیق بودن داده‌ها برقرار نباشد و داده‌هایی نادقیق (فازی) داشته باشیم، برآوردگرهای معمولی LAD-LASSO قادر به برازش مدلی خوب به مجموعه داده نخواهند بود. بنابراین در این بخش سعی در ارائه برآوردگری داریم که بتواند در حضور داده‌های پرت و وجود همخطی میان متغیرهای توضیحی (تبیینی)، زمانی که متغیر پاسخ، تبیینی و ضرایب فازی هستند، کارایی مطلوب را داشته باشد. در این برآوردگر استوار از جریمه لاسو برای مقابله با مشکل همخطی استفاده می‌نماییم. بنابراین برآوردگر استوار پیشنهاد شده قادر به انتخاب متغیر نیز خواهد بود.

اکنون مدل رگرسیون فازی زیر با ورودی، خروجی و ضرایب فازی زیر را در نظر بگیرید

$$\tilde{Y}_i = \tilde{A}_0 \oplus (\tilde{A}_1 \otimes \tilde{X}_{i1}) \oplus \cdots \oplus (\tilde{A}_p \otimes \tilde{X}_{ip}), \quad i = 1, \dots, n \quad (10.3)$$

که در آن $\tilde{Y}_i$ متغیر پاسخ فازی i ام، $\tilde{X}_{ij}$ متغیر توضیحی فازی برای مشاهده i ام و متغیر j ام و $\tilde{A}_j$ ضریب رگرسیونی فازی برای متغیر j ام است. برای برآورد ضرایب مدل (۱۰.۳) زمانی که همخطی بین متغیرهای توضیحی وجود دارد و در حضور داده‌های پرت، تابع هدف زیر را با استفاده از جریمه لاسو معرفی می‌کنیم

$$\sum_{i=1}^n d(\tilde{Y}_i(\tilde{A}', \tilde{A}''), \tilde{Y}_i) + k \|\tilde{A}\| \quad (11.3)$$

که در آن k پارامتر تنظیم‌کننده لاسو، $\tilde{Y}_i(\tilde{A}', \tilde{A}'')$ طرف راست معادله (۵۳.۱) خواهد بود و اگر $\tilde{\alpha} = (\alpha, \alpha, \alpha)_T$ باشد، $\|\tilde{A}\|$ را به صورت زیر تعریف می‌کنیم

$$\|\tilde{A}\| = d(\tilde{A}, \tilde{\alpha}) = |m_A - \alpha| + |l_A - \alpha| + |r_A - \alpha| \quad (12.3)$$

با استفاده از معادله (۱۲.۳) و توضیحات بخش ۳.۱۰.۱ تعریف می‌کنیم

$$\begin{aligned} \|\tilde{A}\| &= \sum_{j=1}^p \|\tilde{A}_j\| \\ &= \sum_{j=1}^p \|(m'_{A_j} - m''_{A_j}, l'_{A_j} + l''_{A_j}, r'_{A_j} + r''_{A_j})\| \\ &= \sum_{j=1}^p (|m'_{A_j} - m''_{A_j}| + |l'_{A_j} + l''_{A_j}| + |r'_{A_j} + r''_{A_j}|) \end{aligned} \quad (13.3)$$

اگر داشته باشیم $\tilde{A}_0 = (m_{A_0}, l_{A_0}, r_{A_0})^T$ با استفاده از تعاریف بالا می‌توان (۱۱.۳) را به صورت زیر بازنویسی کرد

$$\begin{aligned} & \sum_{i=1}^n d(\tilde{Y}_i(\tilde{A}', \tilde{A}''), \tilde{Y}_i) + k \|\tilde{A}\| \\ &= \sum_{i=1}^n \left[\left| m_{Y_i} - \left(m_{A_0} + \sum_{j=1}^p (m'_{A_j} - m''_{A_j}) m_{X_{ij}} \right) \right| \right. \\ & \quad + \left| l_{Y_i} - \left(l_{A_0} + \sum_{j=1}^p (m'_{A_j} l_{X_{ij}} + l'_{A_j} m_{X_{ij}} + m''_{A_j} r_{X_{ij}} + l''_{A_j} m_{X_{ij}}) \right) \right| \\ & \quad + \left. \left| r_{Y_i} - \left(r_{A_0} + \sum_{j=1}^p (m'_{A_j} r_{X_{ij}} + r'_{A_j} m_{X_{ij}} + m''_{A_j} l_{X_{ij}} + r''_{A_j} m_{X_{ij}}) \right) \right| \right] \\ & \quad + k \left[|m_{A_0}| + |l_{A_0}| + |r_{A_0}| + \sum_{j=1}^p (|m'_{A_j} - m''_{A_j}| + |l'_{A_j} + l''_{A_j}| + |r'_{A_j} + r''_{A_j}|) \right] \end{aligned}$$

با کمینه کردن رابطه (۱۴.۳) تحت قیود زیر

$$\begin{aligned} m'_{A_j} \geq l'_{A_j}, m''_{A_j} \geq r''_{A_j}, \quad l_{A_0}, r_{A_0}, m'_{A_j}, m''_{A_j}, l'_{A_j}, l''_{A_j}, r'_{A_j}, r''_{A_j} \geq 0, \\ j = 1, 2, \dots, p \end{aligned} \quad (15.3)$$

برآورد ضرایب مدل رگرسیونی محاسبه می‌گردد. برای کمینه کردن این معادله از نرم افزار Mathematica 11 استفاده می‌کنیم.

۱.۷.۳ مثال کاربردی

در اینجا نیز همانند مثال ۳.۶.۲ برای داده‌های جدول ۱۸.۲ با استفاده از برآوردگر لاسو استوار فازی و انتخاب پارامتر $k = 2/5$ ، با مینیمم کردن رابطه (۱۴.۳) تحت قیود (۱۵.۳) با استفاده از نرم‌افزار Mathematica 11 مدل رگرسیون لاسو فازی به صورت زیر بدست می‌آید

$$\begin{aligned} \tilde{Y} = (0/00004, 0, 0) + (2/20, 0, 0)\tilde{X}_1 + (1/15, 0, 0)\tilde{X}_2 \\ + (0/83, 0, 0)\tilde{X}_3 + (0/469, 0, 0)\tilde{X}_4 \end{aligned} \quad (16.3)$$

همچنین شاخص‌های ارزیابی مدل را برای این مسئله به دست می‌آوریم و نتایج به دست آمده برای مدل رگرسیون کمترین توان‌های دوم معمولی فازی، رگرسیون کمترین توان‌های دوم استوار فازی، رگرسیون ريج استوار فازی به دست آمده در بخش ۵.۶.۲ و رگرسیون لاسو فازی استوار را در جدول ۱۸.۳ خلاصه می‌کنیم.

همان‌طور که در جدول ۱۸.۳ نشان داده شد، مدل رگرسیون ريج استوار در قیاس با برآوردگر کمترین توان‌های دوم معمولی و کمترین توان‌های دوم استوار فازی از عملکرد بهتری برخوردار است و نتایج عدم کارایی رگرسیون فازی معمولی را نشان می‌دهند. اما برآوردگر لاسو استوار فازی، با بهبود دادن برآوردگر ريج استوار فازی عملکرد بهتری را نشان می‌دهد.

جدول ۱۸.۳: شاخص‌های ارزیابی مدل برای رگرسیون لاسو فازی برای مجموعه داده سیمان پرتلند

شاخص λ	شاخص E	k	مدل رگرسیونی
۱۹/۸۲	۳۷/۷۳	۰	OLS
۱۹/۱۴	۳۶/۳۷	۰	کمترین توان‌های دوم استوار
۱۸/۶۷۷	۳۵/۳۸	۱	ریج استوار فازی
۱۸/۶۲	۳۵/۱۲	۲/۵	لاسو استوار فازی

فصل ۴

برآوردگر بریج استوار

۱.۴ مقدمه

برآوردگر رگرسیونی بریج هردو برآوردگرهای رگرسیونی ریج و لاسو را بهبود می‌دهد، اما از آن‌جا که نسخه معمولی این برآوردگر مجموع توان‌های دوم باقی‌مانده‌ها را با یک جریمه L_γ مینیمم می‌کند، نسبت به داده‌های پرت موجود در مجموعه داده استوار نیست. تلاش‌هایی برای تعریف نسخه‌هایی استوار از این برآوردگر شده است. این روش‌های پیشنهادشده برآوردگرهای رگرسیونی بریج استوار نسبت به داده‌های پرت در متغیر پاسخ و خطاهای دم‌پهن استوار هستند، اما نسبت به نقاط نافذ استوار نمی‌باشند. در این فصل برآوردگر بریج استواری را بررسی می‌کنیم که روش‌های رگرسیونی بریج و MM-برآوردگرها را ادغام می‌کند و خواص جانبی آن را بررسی می‌کنیم. این برآوردگر در مقابل داده‌های پرت در متغیر پاسخ و نقاط نافذ استوار است. در ادامه به‌منظور ارزیابی نتایج نظری به‌دست آمده از یک مثال شبیه‌سازی و یک مثال داده واقعی استفاده خواهیم کرد. مطالب این فصل عمدتاً برگرفته از آرسلان [۱۳] و یوهای [۶۵] می‌باشند.

مدل رگرسیونی

$$\mathbf{Y} = \mathbf{X}\beta + \epsilon \quad (1.4)$$

را در نظر بگیرید، که در آن $X_{n \times p}$ ماتریس طرح، $Y_{n \times 1}$ بردار پاسخ و ϵ بردار خطاهای ϵ_i

$(i = 1, \dots, n)$ باشند. در این فصل فرض می‌کنیم $\beta_0 = (\beta_{01}, \beta_{02}, \dots, \beta_{0p})^T$ بردار پارامتر واقعی و $\Omega \in \mathbb{R}^p$ معرف فضای پارامتر باشد. همچنین فرض می‌کنیم $n > p$ ، Ω محدب و β_0 در مجموعه نقاط درونی Ω می‌باشد. یک روش برای برآورد بردار پارامتر نامعلوم β مینیمم کردن تابع هدف کمترین توان‌های دوم زیر است

$$Q_n(\beta) = \sum_{i=1}^n (y_i - \mathbf{x}_i^T \beta)^2 + k_n \sum_{j=1}^p |\beta_j|^\gamma \quad (2.4)$$

که در آن $\gamma > 0$ و $k = k_n$ پارامتر تنظیم‌کننده است و در اینجا اندیس n برای k نشان‌دهنده وابستگی پارامتر به حجم نمونه است که در اثبات قضایا از آن استفاده می‌شود. این روش کمترین توان‌های دوم جریمه شده، روش برآورد رگرسیونی بریج نامیده می‌شود که توسط فرانک و فریدمن [۳۰] معرفی شده است و تابع جریمه بریج به صورت زیر معرفی می‌شود

$$L_\gamma(\beta) = \sum_{j=1}^p |\beta_j|^\gamma \quad (3.4)$$

همان‌طور که مشاهده می‌شود با انتخاب $\gamma = 1$ و $\gamma = 2$ ، به ترتیب به برآوردگرهای ریج و لاسوی معرفی شده در فصل یک خواهیم رسید. ویژگی‌های برآوردگر بریج نیز در فصل یک بررسی شد، اما همان‌طور که مشاهده می‌شود، این برآوردگر نسبت به داده‌های پرت موجود در مجموعه داده استوار نمی‌باشد (به دلیل وجود تابع توان دوم خطاها). بنابراین برآوردگر جدیدی معرفی می‌کنیم که نسبت به داده‌های پرت متغیر پاسخ و نقاط نافذ موجود در مجموعه داده استوار خواهد بود.

۲.۴ برآوردگر بریج استوار

در این بخش با استفاده از روش‌های مکرر نسخه‌ای استوار از برآوردگر بریج ارائه می‌دهیم که از ادغام دو روش برآوردگر رگرسیونی بریج و MM-برآوردگرها به دست می‌آید. فرض می‌کنیم $\hat{\beta}_s$ برآوردگر استوار اولیه به دست آمده از MM-برآوردگرهای معرفی شده در بخش ۲.۶.۱ است و $\hat{\sigma}_s$ ، MM-برآوردگر مقیاسی برای باقی‌مانده‌های $r_i(\hat{\beta}_s)$ ، $i = 1, 2, \dots, n$ باشند. در این صورت برآوردگر MM-Bridge به صورت زیر تعریف می‌شود

$$\hat{\beta}_{MM-Bridge} = \arg \min_{\beta \in \mathbb{R}^p} L_n(\beta) \quad (4.4)$$

که در آن

$$L_n(\beta) = \hat{\sigma}_s^2 \sum_{i=1}^n \rho_1\left(\frac{r_i(\beta)}{\hat{\sigma}_s}\right) + k_n \sum_{j=1}^p |\beta_j|^\gamma \quad (5.4)$$

به طوری که k_n پارامتر تنظیم کننده و $r_i(\beta) = y_i - \mathbf{x}_i^\top \beta$ و $\hat{\sigma}_s$ همان برآوردگر مقیاسی برای باقی مانده ها است. عبارت $\hat{\sigma}_s^2$ برای اطمینان از این که با انتخاب $\rho_1(r) = r^2$ در نهایت تابع هدف (۵.۴) به تابع هدف (۲.۴) تبدیل شود، اضافه می گردد.

اکنون فرض کنید $\mathbf{1}_p = (1, \dots, 1)^\top$ یک بردار p مؤلفه ای از یک ها و $|\beta|^\gamma = (|\beta_1|^\gamma, \dots, |\beta_p|^\gamma)$ در این صورت رابطه (۵.۴) به صورت زیر بازنویسی می شود

$$L_n(\beta) = \hat{\sigma}_s^2 \sum_{i=1}^n \rho_1\left(\frac{r_i(\beta)}{\hat{\sigma}_s}\right) + k_n \mathbf{1}_p^\top |\beta|^\gamma \quad (6.4)$$

اگر ψ_1 مشتق تابع ρ_1 باشد، با توجه به این که $\frac{d|\beta_j|}{d\beta_j} = \frac{\beta_j}{|\beta_j|} = \text{sgn}(\beta_j)$ داریم

$$\frac{dL_n(\beta)}{d\beta} = \hat{\sigma}_s^2 \sum_{i=1}^n \psi_1\left(\frac{r_i(\beta)}{\hat{\sigma}_s}\right) \mathbf{x}_i - k_n \gamma \mathbf{V}(\beta) \quad (7.4)$$

که در آن $\mathbf{V}(\beta)$ یک بردار $p \times 1$ به صورت $(|\beta_1|^{\gamma-1} \text{sgn}(\beta_1), |\beta_2|^{\gamma-1} \text{sgn}(\beta_2), \dots, |\beta_p|^{\gamma-1} \text{sgn}(\beta_p))^\top$ است.

لم ۱.۲.۴. طبق پذیره های این بخش، از $\frac{dL_n(\beta)}{d\beta} = 0$ نتیجه می شود

$$(\mathbf{X}^\top \mathbf{W}(\beta) \mathbf{X} + k_n \gamma \mathbf{W}_\circ(\beta)) \beta = \mathbf{X}^\top \mathbf{W}(\beta) \mathbf{Y} \quad (8.4)$$

که در آن $\mathbf{W}_\circ(\beta) =$ و $W_i(\beta) = \frac{\psi_1(r_i(\beta)/\hat{\sigma}_s)}{r_i(\beta)/\hat{\sigma}_s}$ ، $\mathbf{W}(\beta) = \text{diag}(W_1(\beta), W_2(\beta), \dots, W_n(\beta))$ $\text{diag}(|\beta_1|^{\gamma-2}, \dots, |\beta_p|^{\gamma-2})$.

برهان. بدیهی است که

$$\frac{dL_n(\beta)}{d\beta} = (r_1(\beta)W_1(\beta)\mathbf{x}_1 + \dots + r_n(\beta)W_n(\beta)\mathbf{x}_n) - k_n \gamma \mathbf{V}(\beta) \quad (9.4)$$

از آن جایی که

$$\mathbf{V}(\beta) = \begin{bmatrix} |\beta_1|^{\gamma-1} \text{sgn}(\beta_1) \\ |\beta_2|^{\gamma-1} \text{sgn}(\beta_2) \\ \vdots \\ |\beta_p|^{\gamma-1} \text{sgn}(\beta_p) \end{bmatrix} = \begin{bmatrix} |\beta_1| \frac{\beta_1}{|\beta_1|} |\beta_1|^{\gamma-2} \\ |\beta_2| \frac{\beta_2}{|\beta_2|} |\beta_2|^{\gamma-2} \\ \vdots \\ |\beta_p| \frac{\beta_p}{|\beta_p|} |\beta_p|^{\gamma-2} \end{bmatrix}$$

نتیجه می شود

$$\mathbf{V}^\top(\beta) = [\beta_1, \beta_2, \dots, \beta_p]_{1 \times p} \times \begin{bmatrix} |\beta_1|^{\gamma-2} & & 0 \\ & \ddots & \\ 0 & & |\beta_p|^{\gamma-2} \end{bmatrix}_{p \times p}$$

با جای گذاری $r_i(\beta) = y_i - \mathbf{x}_i^\top \beta$ ، $i = 1, 2, \dots, n$ ، در رابطه (۹.۴) داریم

$$(y_1 - \mathbf{x}_1^\top \beta)W_1(\beta)\mathbf{x}_1 + \dots + (y_n - \mathbf{x}_n^\top \beta)W_n(\beta)\mathbf{x}_n - k_n \gamma \beta^\top \mathbf{W}_\circ(\beta) = 0$$

بنابراین

$$y_1 W_1(\beta) \mathbf{x}_1 - \mathbf{x}_1^\top \beta W_1 \mathbf{x}_1 + \cdots + y_n W_n(\beta) \mathbf{x}_n - \mathbf{x}_n^\top \beta W_n \mathbf{x}_n - k_n \gamma \beta^\top \mathbf{W}_o(\beta) = 0$$

با ترانهاده گرفتن از رابطه فوق، نتیجه می‌شود

$$\begin{aligned} \mathbf{x}_1^\top W_1(\beta) y_1 - \mathbf{x}_1^\top W_1(\beta) \mathbf{x}_1^\top \beta + \cdots + \mathbf{x}_n^\top W_n(\beta) y_n - \mathbf{x}_n^\top W_n(\beta) \mathbf{x}_n^\top \beta \\ - k_n \gamma \mathbf{W}_o(\beta) \beta = 0 \end{aligned} \quad (10.4)$$

چنانچه برابری (۱۰.۴) را با نماد ماتریسی بازنویسی کنیم رابطه زیر حاصل می‌شود

$$\mathbf{X}^\top \mathbf{W}(\beta) \mathbf{Y} - \mathbf{X}^\top \mathbf{W}(\beta) \mathbf{X} \beta - k_n \gamma \mathbf{W}_n(\beta) \beta = 0 \quad (11.4)$$

و اثبات کامل می‌شود. $\square$

با استفاده از لم ۱.۲.۴ برآوردگر $\hat{\beta}_{\text{MM-BR}}$ در رابطه (۴.۴) به صورت زیر حاصل می‌شود

$$\hat{\beta}_{\text{MM-BR}} = \left(\mathbf{X}^\top \mathbf{W}(\hat{\beta}_{\text{MM-BR}}) \mathbf{X} + k_n \gamma \mathbf{W}_o(\hat{\beta}_{\text{MM-BR}}) \right)^{-1} \mathbf{X}^\top \mathbf{W}(\hat{\beta}_{\text{MM-BR}}) \mathbf{Y} \quad (12.4)$$

مشروط به اینکه ماتریس $\left(\mathbf{X}^\top \mathbf{W}(\hat{\beta}_{\text{MM-BR}}) \mathbf{X} + k_n \gamma \mathbf{W}_o(\hat{\beta}_{\text{MM-BR}}) \right)$ پرتبه باشد. با استفاده از $\hat{\beta}_{\text{MM-BR}}$ به دست آمده در (۱۲.۴) رابطه پیشگویی y به صورت زیر به دست می‌آید

$$\hat{y} = \mathbf{H} y$$

که در آن

$$\mathbf{H} = \mathbf{X} \left(\mathbf{X}^\top \mathbf{W}(\hat{\beta}_{\text{MM-BR}}) \mathbf{X} + k_n \gamma \mathbf{W}_o(\hat{\beta}_{\text{MM-BR}}) \right)^{-1} \mathbf{X}^\top \mathbf{W}(\hat{\beta}_{\text{MM-BR}})$$

ماتریس $\mathbf{H}$ می‌تواند برای انتخاب پارامتر تنظیم‌کننده k_n و پارامتر γ ، با استفاده از یک معیار BIC، به کار برده شود. در حالت $k_n = 0$ ، جواب مسئله مینیمم‌سازی داده‌شده در معادله (۴.۴)، برآوردگر MM خواهد بود، در حالت $\gamma = 2$ ، برآوردگر رگرسیون ریج استوار معرفی شده در فصل دوم تعریف می‌شود و در حالت $\gamma = 1$ ، برآوردگر لاسو استوار مبتنی بر روش رگرسیونی MM حاصل می‌شود.

۳.۴ ویژگی‌های مجانبی برآوردگر ریج استوار

در این بخش ویژگی‌های سازگاری، نرمال بودن مجانبی و تنگی برآوردگر $\hat{\beta}_{\text{MM-BR}}$ را مورد بررسی قرار می‌دهیم. برای سادگی در نوشتار، در ادامه، برآوردگر $\hat{\beta}_{\text{MM-BR}}$ را با $\hat{\beta}_n$ نشان می‌دهیم. برای مطالعه ویژگی‌های مجانبی برآوردگر $\hat{\beta}_{\text{MM-BR}}$ به پذیره‌های زیر احتیاج داریم:

B1: فرض کنید

$$l(\beta) = \mathbb{E}_F \left(\rho_1 \left(\frac{r(\beta)}{\sigma_0} \right) \right)$$

که در آن $\mathbb{E}_F$ نشان‌دهنده امید ریاضی نسبت به توزیع خطاها (F) است، $r(\beta) = y - x^\top \beta$ و σ_0 با استفاده از $\sigma_0 = \mathbb{E}_F(\rho_0(r(\beta_0)/\sigma_0)) = 0.5$ تعیین می‌گردد. همچنین فرض می‌کنیم که $l(\beta)$ در β_0 مینیمم یکتا دارد. توجه کنید

$$\lim_{n \rightarrow \infty} \left(\frac{1}{n} \sum_{i=1}^n \rho_1 \left(\frac{r_i(\beta)}{\hat{\sigma}_0} \right) \right) = \mathbb{E}_F \left(\rho_1 \left(\frac{r(\beta)}{\sigma_0} \right) \right) = l(\beta)$$

(برای اطلاعات بیشتر یوهای [۶۵] را ببینید).

B2: فرض کنید که $\rho_1(x)$ به‌طور پیوسته دوبار مشتق‌پذیر بوده و m ای وجود دارد به‌طوری که برای $|x| \geq m$ داریم $\rho_1(x) = a$.

B3: فرض کنید وقتی $n \rightarrow \infty$

$$\frac{1}{n} \sum_{i=1}^n x_i x_i^\top \rightarrow C$$

که در آن C یک ماتریس نامنفرد است.**سازگاری.** برای بیان سازگاری برآوردگر $\hat{\beta}_{\text{MM-BR}}$ تعریف می‌کنیم

$$Z_n(\beta) = n^{-1} L_n(\beta) = \frac{\hat{\sigma}_s^2}{n} \sum_{i=1}^n \rho_1 \left(\frac{r_i(\beta)}{\hat{\sigma}_s} \right) + \frac{k_n}{n} \sum_{j=1}^p |\beta_j|^\gamma \quad (13.4)$$

که در نقطه $\arg \min_{\beta} L_n(\beta) = \hat{\beta}_n$ مینیمم خواهد شد. گزاره زیر نشان می‌دهد برآوردگر $\hat{\beta}_{\text{MM-BR}}$ یک برآوردگر سازگار برای β_0 است.

گزاره ۱.۳.۴. اگر به‌ازای $n \rightarrow \infty$ ، $\gamma > 0$ و $\frac{k_n}{n} \rightarrow \lambda_0 \geq 0$ تحت پذیره‌های $A1$ ، $A2$ از بخش ۲.۶.۱ و $B1$ ، آن‌گاه

$$\hat{\beta}_n \xrightarrow{P} \arg \min(Z(\beta))$$

که در آن

$$Z(\beta) = \sigma_0^2 l(\beta) + \lambda_0 \sum_{j=1}^p |\beta_j|^\gamma \quad (14.4)$$

برهان. توجه کنید وقتی $A1$ ، $A2$ و $B3$ برقرار باشند، اگر $\hat{\beta}_s$ برای β_0 سازگار باشد، در این صورت $\hat{\sigma}_s$ یک برآوردگر سازگار برای σ_0 است (توجه کنید که β_0 و σ_0 نشان‌دهنده مقدار واقعی‌شان

است). $Z_n(\beta)$ را مانند (۱۳.۴) تعریف می‌کنیم. برای اثبات سازگاری $\hat{\beta}_n$ ابتدا نیاز داریم نشان دهیم

$$\sup_{\beta \in \Omega} |Z_n(\beta) - Z(\beta)| \xrightarrow{P} 0$$

یعنی باید نشان دهیم که $Z_n(\beta)$ به‌طور یکنواخت در احتمال به $Z(\beta)$ ، تعریف‌شده در معادله (۱۴.۴)، همگراست. همچنین باید نشان دهیم که $\hat{\beta}_n = O_p(1)$ (یعنی $\hat{\beta}_n$ در احتمال، به‌طور یکنواخت کراندار است).

اثبات را برای دو حالت $\gamma \geq 1$ و $\gamma < 1$ بیان می‌کنیم. باید توجه داشته باشیم به ازای $\gamma \geq 1$ ، $Z_n(\beta)$ محدب بوده در صورتی که به ازای $\gamma < 1$ محدب نمی‌باشد. از آن‌جا که $\frac{k_n}{n} \rightarrow \lambda_0$ بنابراین

$$\frac{k_n}{n} \sum_{j=1}^p |\beta_j|^\gamma \rightarrow \lambda_0 \sum_{j=1}^p |\beta_j|^\gamma$$

همچنین طبق قانون ضعیف اعداد بزرگ و سازگاری $\hat{\sigma}_n$ برای σ_0 داریم

$$\frac{1}{n} \sum_{i=1}^n \rho_\gamma \left(\frac{r_i(\beta)}{\hat{\sigma}_s} \right) \xrightarrow{P} \mathbb{E}_F \left(\rho_\gamma \left(\frac{r(\beta)}{\sigma_0} \right) \right)$$

بنابراین

$$Z_n(\beta) \xrightarrow{P} \sigma_0^\gamma \mathbb{E}_F \left(\rho_\gamma \left(\frac{r(\beta)}{\sigma_0} \right) \right) + \lambda_0 \sum_{j=1}^p |\beta_j|^\gamma$$

که در آن طرف راست همگرایی همان $Z(\beta)$ است، لذا نتیجه می‌شود $Z_n(\beta) - Z(\beta) \xrightarrow{P} 0$ پس

$$\sup_{\beta \in \Omega} |Z_n(\beta) - Z(\beta)| \xrightarrow{P} 0$$

از آن‌جایی که Ω یک مجموعه فشرده می‌باشد (طبق پذیره‌های ابتدای بخش)، با استفاده از قاعده پیوستگی تابع احتمال داریم

$$\forall \epsilon > 0 \quad \exists n, n_0 \in \mathbb{N} \text{ s.t. } \forall n \geq n_0, \forall \beta \in \Omega; |Z_n(\beta) - Z(\beta)| \leq \epsilon$$

به این معنی که $Z_n(\beta) \xrightarrow{U} Z(\beta)$ ، (که در آن حرف U به معنی همگرایی یکنواخت می‌باشد). از آن‌جایی که هر دنباله به‌طور یکنواخت همگرا، نقطه به نقطه همگراست، مینیمم $Z_n(\beta)$ نیز به مینیمم $Z(\beta)$ همگراست.

از آن‌جا که $\gamma \geq 1$ و $Z_n(\beta)$ محدب و دارای همگرایی یکنواخت است و چون Ω فشرده است، پس β کراندار است و در نتیجه آن داریم

$$\arg \min_{\beta} (Z_n(\beta)) \xrightarrow{P} \arg \min_{\beta} (Z(\beta))$$

حال اگر $\gamma < 1$ ، آن‌گاه $Z_n(\beta)$ محدب نیست، اما کراندار است و $\hat{\beta}_n$ مینیمم‌کننده آن است. بنابراین داریم

$$Z_n(\beta) \geq \frac{\hat{\sigma}_s^2}{n} \sum_{i=1}^n \rho_1\left(\frac{r_i(\beta)}{\hat{\sigma}_s}\right)$$

طرف راست نامساوی بالا تابع هدف MM-برآوردگرها است. با توجه به کراندار بودن این برآوردگرها و نامساوی بالا داریم

$$\arg \min_{\beta} (Z_n(\beta)) = \hat{\beta}_n = O_p(1)$$

بنابراین

$$\arg \min_{\beta} (Z_n(\beta)) \xrightarrow{P} \arg \min_{\beta} (Z(\beta))$$

در نتیجه برای هر $\gamma > 0$ رابطه بالا برقرار خواهد بود و اثبات کامل می‌شود. $\square$

اکنون برای سازگاری $\hat{\beta}_n$ ، اگر $k_n = o(n)$ و $\frac{k_n}{n} \rightarrow 0$ ، آن‌گاه

$$Z(\beta) = \sigma_0^2 l(\beta)$$

و از آنجایی که در این حالت $Z(\beta)$ کمترین مقدار خود را دارد و این مقدار یکتاست و تابع هدف نیز به‌طور یکنواخت همگراست، پس اگر $\frac{k_n}{n} \rightarrow 0$ وقتی $n \rightarrow \infty$ ، آن‌گاه در احتمال $\hat{\beta}_n \rightarrow \beta_0$.

نرمال بودن مجانبی. گزاره زیر توزیع حدی را برای برآوردگر $\hat{\beta}_{\text{MM-BR}}$ وقتی $\gamma \geq 1$ و $\gamma < 1$ نتیجه می‌شود.

برای $\gamma > 1$ تعریف می‌کنیم

$$V(\mathbf{u}) = -\mathbf{u}^\top \mathbf{W} + \frac{1}{\gamma} \mathbf{u}^\top \left[B\left(\Psi_1\left(\frac{r(\beta_0)}{\sigma_0}\right)\right) C \right] \mathbf{u} + \lambda_0 \sum_{j=1}^p u_j \text{sgn}(\beta_{0j}) |\beta_{0j}|^{\gamma-1}$$

همچنین برای $\gamma = 1$

$$V(\mathbf{u}) = -\mathbf{u}^\top \mathbf{W} + \frac{1}{\gamma} \mathbf{u}^\top \left[B\left(\Psi_1\left(\frac{r(\beta_0)}{\sigma_0}\right)\right) C \right] \mathbf{u} + \lambda_0 \sum_{j=1}^p [u_j \text{sgn}(\beta_{0j}) \mathbf{I}(\beta_{0j} \neq 0) + |u_j| \mathbf{I}(\beta_{0j} = 0)]$$

که در آن

$$\mathbf{W} \sim N_p(0, \sigma_0^2 \mathbf{A}(\Psi_1(\frac{r(\beta_0)}{\sigma_0})) \mathbf{C})$$

و

$$\mathbf{A}\left(\Psi_1\left(\frac{r(\beta_0)}{\sigma_0}\right)\right) = \mathbb{E}_F\left(\Psi_1'\left(\frac{r(\beta_0)}{\sigma_0}\right)\right) \text{ و } B\left(\Psi_1\left(\frac{r(\beta_0)}{\sigma_0}\right)\right) = \mathbb{E}_F\left(\Psi_1\left(\frac{r(\beta_0)}{\sigma_0}\right)\right)$$

گزاره ۲.۳.۴. تحت درستی پذیره‌های A_1, A_2, B_1, B_2 و B_3 و این که برآوردگر اولیه $\hat{\beta}_s$ یک برآوردگر سازگار برای β_0 است، آن‌گاه به‌ازای $\gamma \geq 1$ ، برآوردگر $\hat{\beta}_n$ در موارد زیر صدق می‌کند

(i) اگر $\gamma \geq 1$ و $\frac{k_n}{\sqrt{n}} \rightarrow \lambda_0 \geq 0$ وقتی $n \rightarrow \infty$ ، آن گاه

$$\sqrt{n}(\hat{\beta}_n - \beta_0) \xrightarrow{d} \arg \min_u V(u)$$

که در آن $u = (u_1, \dots, u_p)^\top$ در فضای $\mathbb{R}^p$ و $\xrightarrow{d}$ به مفهوم همگرایی در توزیع است.

(ii) اگر به ازای $n \rightarrow \infty$ ، $\lambda < 1$ و $k_n/n^{\gamma/2} \rightarrow \lambda_0 \geq 0$ ، آن گاه به ازای $\gamma < 1$ ،

$$\sqrt{n}(\hat{\beta}_n - \beta_0) \xrightarrow{d} \arg \min_u V(u)$$

که در آن

$$V(u) = -u^\top W + \frac{1}{\gamma} u^\top \left[B \left(\Psi_\gamma \left(\frac{r(\beta_0)}{\sigma_0} \right) \right) C \right] u + \lambda_0 \sum_{j=1}^p |u_j|^\gamma I(\beta_{0j} = 0)$$

برهان. (i): برای به دست آوردن توزیع حدی $\hat{\beta}_n$ ، ابتدا تابع هدف موضعی شده $V_n(u)$ را تعریف می کنیم

$$V_n(u) = \hat{\sigma}_s^2 \sum_{i=1}^n \left(\rho_\gamma \left(\frac{r_i(\beta_0 + u/\sqrt{n})}{\hat{\sigma}_s} \right) - \rho_\gamma \left(\frac{r_i(\beta_0)}{\hat{\sigma}_0} \right) \right) + k_n \sum_{j=1}^p \left(|\beta_{0j} + \frac{u_j}{\sqrt{n}}|^\gamma - |\beta_{0j}|^\gamma \right)$$

که در آن $u = (u_1, \dots, u_p)^\top$ بردار پارامتر مکانی در $\mathbb{R}^p$ است، که هم علامت با بردار پارامتر اولیه (β_0) است.

باید توجه کنیم که $V_n(u)$ در $\hat{u}_n = \sqrt{n}(\hat{\beta}_n - \beta_0)$ مینیمم می شود. برای به دست آوردن توزیع $\hat{\beta}_n$ ابتدا باید نشان دهیم

$$V_n(u) \xrightarrow{d} V(u)$$

با استفاده از بسط سری تیلور $\hat{\sigma}_s^2 \sum_{i=1}^n \left(\rho_\gamma \left(\frac{r_i(\beta_0 + u/\sqrt{n})}{\hat{\sigma}_s} \right) - \rho_\gamma \left(\frac{r_i(\beta_0)}{\hat{\sigma}_0} \right) \right)$ حول $u = 0$ داریم

$$\begin{aligned} \hat{\sigma}_s^2 \sum_{i=1}^n \left(\rho_\gamma \left(\frac{r_i(\beta_0 + u/\sqrt{n})}{\hat{\sigma}_s} \right) - \rho_\gamma \left(\frac{r_i(\beta_0)}{\hat{\sigma}_0} \right) \right) &= -u^\top \left[\frac{\hat{\sigma}_s}{\sqrt{n}} \sum_{i=1}^n \psi_\gamma \left(\frac{r_i(\beta_0)}{\hat{\sigma}_s} \right) \mathbf{x}_i \right] \\ &\quad + \frac{1}{\gamma} u^\top \left[\frac{1}{n} \sum_{i=1}^n \psi'_\gamma \left(\frac{r_i(\beta_0)}{\hat{\sigma}_s} \right) \mathbf{x}_i \mathbf{x}_i^\top \right] u \end{aligned}$$

از آن جایی که

$$\begin{aligned} \frac{1}{n} \sum_{i=1}^n \psi'_\gamma \left(\frac{r_i(\beta_0)}{\hat{\sigma}_s} \right) \mathbf{x}_i \mathbf{x}_i^\top &\rightarrow \mathbf{B}(\psi_\gamma \left(\frac{r(\beta_0)}{\sigma_0} \right)) C \\ \frac{\hat{\sigma}_s}{\sqrt{n}} \sum_{i=1}^n \psi_\gamma \left(\frac{r_i(\beta_0)}{\hat{\sigma}_s} \right) \mathbf{x}_i &\xrightarrow{D} \mathbf{W} \sim N_p(0, \sigma_0^2 \mathbf{A}(\psi_\gamma \left(\frac{r(\beta_0)}{\sigma_0} \right)) C) \end{aligned}$$

به دست می‌آوریم

$$\hat{\sigma}_s^2 \sum_{i=1}^n \left(\rho_1 \left(\frac{r_i(\beta_\circ + \mathbf{u}/\sqrt{n})}{\hat{\sigma}_s} \right) - \rho_1 \left(\frac{r_i(\beta_\circ)}{\hat{\sigma}_\circ} \right) \right) \xrightarrow{d} -\mathbf{u}^\top \mathbf{W} + \frac{1}{\gamma} \mathbf{u}^\top \left[B(\psi_1 \left(\frac{r(\beta_\circ)}{\sigma_\circ} \right)) C \right] \mathbf{u}$$

با در نظر گرفتن قسمت دوم $V_n(\mathbf{u})$ ، همانند اثبات قضیه ۲ در نایت و فو [۴۱]، اگر $\gamma > 1$ وقتی $n \rightarrow \infty$ داریم

$$\begin{aligned} k_n \sum_{j=1}^p \left(|\beta_{\circ j} + \frac{u_j}{\sqrt{n}}|^\gamma - |\beta_{\circ j}|^\gamma \right) &= \frac{k_n}{\sqrt{n}} \sum_{j=1}^p u_j \operatorname{sgn}(\beta_{\circ j}) |\beta_{\circ j}|^{\gamma-1} \\ &\rightarrow \lambda_\circ \gamma \sum_{j=1}^p u_j \operatorname{sgn}(\beta_{\circ j}) |\beta_{\circ j}|^{\gamma-1} \end{aligned}$$

و اگر $\gamma = 1$ وقتی $n \rightarrow \infty$ داریم

$$\begin{aligned} k_n \sum_{j=1}^p \left(|\beta_{\circ j} + \frac{u_j}{\sqrt{n}}|^\gamma - |\beta_{\circ j}|^\gamma \right) &= \frac{k_n}{\sqrt{n}} \sum_{j=1}^p [u_j \operatorname{sgn}(\beta_{\circ j}) I(\beta_{\circ j} \neq \circ) + |u_j| I(\beta_{\circ j} = \circ)] \\ &\rightarrow \lambda_\circ \sum_{j=1}^p [u_j \operatorname{sgn}(\beta_{\circ j}) I(\beta_{\circ j} \neq \circ) + |u_j| I(\beta_{\circ j} = \circ)] \end{aligned}$$

از این رو با ترکیب این دو نتیجه برای هر $\gamma \geq 1$ به دست می‌آوریم

$$V_n(\mathbf{u}) \xrightarrow{d} V(\mathbf{u})$$

علاوه بر این $V_n(\mathbf{u})$ را می‌توان به صورت زیر بازنویسی کرد

$$\begin{aligned} V_n(\mathbf{u}) &= -\mathbf{u}^\top \left[\frac{\hat{\sigma}_s}{\sqrt{n}} \sum_{i=1}^n \psi_1 \left(\frac{r_i(\beta_\circ)}{\hat{\sigma}_s} \right) \mathbf{x}_i \right] + \frac{1}{\gamma} \mathbf{u}^\top \left[\frac{1}{n} \sum_{i=1}^n \psi_1' \left(\frac{r_i(\beta_\circ)}{\hat{\sigma}_s} \right) \mathbf{x}_i \mathbf{x}_i^\top \right] \mathbf{u} \\ &\quad + k_n \sum_{j=1}^p \left(|\beta_{\circ j} + \frac{u_j}{\sqrt{n}}|^\gamma - |\beta_{\circ j}|^\gamma \right) \end{aligned}$$

بنابراین، از آنجایی که $V(\mathbf{u})$ مینیمم یکتا دارد و $V_n(\mathbf{u})$ می‌تواند به وسیله یک تابع محدب تقریب زده شود، نتیجه زیر حاصل می‌شود

$$\arg \min_{\mathbf{u}} V_n(\mathbf{u}) = \sqrt{n}(\hat{\beta}_n - \beta_\circ) \xrightarrow{d} \arg \min_{\mathbf{u}} V(\mathbf{u})$$

(ii): اثبات این قسمت مشابه با اثبات قسمت اول می‌باشد. با کمی نکات اضافی که به سبب

نامحدب بودن تابع جریمه به وجود می‌آیند، همانند اثبات قضیه ۳ در نایت و فو [۴۱] و قضیه ۲ در کانر [۲۰]، اگر $\gamma < 1$ و $\frac{k_n}{n^{\gamma/\kappa}} \rightarrow \lambda_\circ \geq 0$ ، به دست می‌آوریم

$$\begin{aligned} k_n \sum_{j=1}^p \left(|\beta_{\circ j} + \frac{u_j}{\sqrt{n}}|^\gamma - |\beta_{\circ j}|^\gamma \right) &= \frac{k_n}{n^{\gamma/\kappa}} \sum_{j=1}^p |u_j|^\gamma I(\beta_{\circ j} = \circ) \\ &\rightarrow \lambda_\circ \sum_{j=1}^p |u_j|^\gamma I(\beta_{\circ j} = \circ) \end{aligned}$$

و همگرایی حول $\mathbf{u}$ در مجموعه‌های فشرده، یکنواخت است. با ترکیب نتایج مشاهده شده به این نتیجه می‌رسیم که

$$V_n(\mathbf{u}) \xrightarrow{d} V(\mathbf{u})$$

برای اثبات این که $\arg \min_{\mathbf{u}} V_n(\mathbf{u}) = \sqrt{n}(\hat{\beta}_n - \beta_0) \xrightarrow{d} \arg \min_{\mathbf{u}} V(\mathbf{u})$ باید نشان دهیم که $\arg \min_{\mathbf{u}} V_n(\mathbf{u}) = O_p(1)$ (به این معنی که $\arg \min_{\mathbf{u}} V_n(\mathbf{u})$ در احتمال به‌طور یکنواخت کراندار است). توجه کنید که برای هر $\mathbf{u}$ و n به قدر کافی بزرگ، داریم

$$\begin{aligned} V_n(\mathbf{u}) &\geq -\mathbf{u}^\top \left[\frac{\hat{\sigma}_s}{\sqrt{n}} \sum_{i=1}^n \psi_1 \left(\frac{r_i(\beta_0)}{\hat{\sigma}_s} \right) \mathbf{x}_i \right] + \frac{1}{\gamma} \mathbf{u}^\top \left[\frac{1}{n} \sum_{i=1}^n \psi'_1 \left(\frac{r_i(\beta_0)}{\hat{\sigma}_s} \right) \mathbf{x}_i \mathbf{x}_i^\top \right] \mathbf{u} \\ &\quad - k_n \sum_{j=1}^p (|u_j|/\sqrt{n})^\gamma \\ &\geq -\mathbf{u}^\top \left[\frac{\hat{\sigma}_s}{\sqrt{n}} \sum_{i=1}^n \psi_1 \left(\frac{r_i(\beta_0)}{\hat{\sigma}_s} \right) \mathbf{x}_i \right] + \frac{1}{\gamma} \mathbf{u}^\top \left[\frac{1}{n} \sum_{i=1}^n \psi'_1 \left(\frac{r_i(\beta_0)}{\hat{\sigma}_s} \right) \mathbf{x}_i \mathbf{x}_i^\top \right] \mathbf{u} \\ &\quad - (\lambda_0 + \delta) \sum_{j=1}^p (|u_j|)^\gamma \\ &= V_n^l(\mathbf{u}) \end{aligned}$$

که در آن δ ثابتی برای اطمینان از اینکه $\frac{k_n}{n^{\gamma/\gamma}} \leq (\lambda_0 + \delta)$ است، آورده شده (اثبات قضیه ۳ در نایت و فو [۴۱] و اثبات قضیه ۲ در کانر [۲۰] را ببینید). از آن جا که عبارت دوم (عبارت درجه دوم در u) از $V_n^l(\mathbf{u})$ سریع‌تر از عبارت‌های $|u_j|^\gamma$ افزایش می‌یابد (عبارت درجه دو بر عبارت با $|u_j|^\gamma$ غالب می‌شود)، نتیجه می‌شود $\arg \min_{\mathbf{u}} V_n^l(\mathbf{u}) = O_p(1)$. بنابراین با استفاده از نامساوی $V_n(\mathbf{u}) \geq V_n^l(\mathbf{u})$ ، به‌دست می‌آوریم

$$\sqrt{n}(\hat{\beta}_n - \beta_0) \xrightarrow{d} \arg \min_{\mathbf{u}} V(\mathbf{u})$$

□

باید توجه کنیم که وقتی $\lambda_0 = 0$

$$\sqrt{n}(\hat{\beta} - \beta_0) \xrightarrow{d} N_p \left(0, \sigma_0^2 \frac{A(\Psi_1(r(\beta_0)/\sigma_0))}{B(\Psi_1(r(\beta_0)/\sigma_0))^2} C^{-1} \right) \quad (15.4)$$

که عبارت فوق توزیع حدی برآوردگر رگرسیونی MM داده شده در یوهای [۶۵] می‌باشد. علاوه بر این وقتی $\gamma \geq 2$ ، توزیع $\arg \min V(\mathbf{u})$ به‌صورت است

$$\begin{aligned} \sqrt{n}(\hat{\beta}_n - \beta_0) &\xrightarrow{d} N_p \left(- \left(B \left(\Psi_1 \left(\frac{r(\beta_0)}{\sigma_0} \right) \right) C \right)^{-1} C^*(\beta_0) \beta_0, \right. \\ &\quad \left. \sigma_0^2 \frac{A(\Psi_1(r(\beta_0)/\sigma_0))}{B(\Psi_1(r(\beta_0)/\sigma_0))^2} C^{-1} \right) \end{aligned} \quad (16.4)$$

که در آن $C^*(\beta_0) = \text{diag}(|\beta_{01}|^{\gamma-2}, \dots, |\beta_{0p}|^{\gamma-2})$. در حالت خاص، توزیع حدی رگرسیون رنج استوار ($\gamma = 2$) مبتنی بر برآوردگر MM داده شده توسط مارونا [۴۷] به صورت زیر به دست می‌آید

$$\sqrt{n}(\hat{\beta}_n - \beta_0) \xrightarrow{d} N_p \left(- \left(B \left(\Psi_1 \left(\frac{r(\beta_0)}{\sigma_0} \right) \right) C \right)^{-1} \beta_0, \sigma_0^2 \frac{A(\Psi_1(r(\beta_0)/\sigma_0))}{B(\Psi_1(r(\beta_0)/\sigma_0))^2} C^{-1} \right) \quad (17.4)$$

تنکی. فرض کنید بتوانیم بردار پارامتر واقعی را به صورت $\beta_0 = (\beta_{01}^\top, \beta_{02}^\top)^\top$ تجزیه کنیم، که در آن β_{01} بردار $p_0 \times 1$ بعدی و β_{02} بردار $(p - p_0) \times 1$ بعدی، به شرط آن که $\beta_{01} \neq 0$ و $\beta_{02} = 0$ باشد. همچنین بردارهای $\beta = (\beta_1^\top, \beta_2^\top)^\top$ و $\hat{\beta}_n = (\hat{\beta}_{n1}, \hat{\beta}_{n2})^\top$ را همانند بردار پارامتر واقعی تجزیه کنیم. نمی‌دانیم که کدام یک از ضرایب صفر و کدام یک ناصفر هستند. بعلاوه $X_i = (X_{i1}^\top, X_{i2}^\top)^\top$ را نیز طبق تجزیه بردار پارامتر واقعی تجزیه می‌کنیم. اکنون گزاره زیر را که نشان‌دهنده ویژگی تنکی برآوردگر $\hat{\beta}_{\text{MM-BR}}$ است، تعریف می‌کنیم. ویژگی تنکی به این معنی است که ضرایب برآورد شده کوچک، به صورت خودکار توسط برآوردگر صفر برآورد شوند.

گزاره ۳.۳.۴ (آرسلان، [۱۳]). فرض کنید که $0 < \gamma \leq 1$. اگر وقتی $n \rightarrow \infty$ ، $\frac{k_n}{n} \rightarrow 0$ و $\frac{k_n}{\sqrt{n}} \rightarrow \infty$ ، آن‌گاه برآوردگر سازگار برنج استوار پارامتر β_{n2} در رابطه زیر صدق می‌کند

$$P(\hat{\beta}_{n2} = 0) \rightarrow 1$$

برهان. ادعای داده‌شده در گزاره ۳.۳.۴ را می‌توانیم به این صورت نیز بازنویسی کنیم. اگر وقتی $n \rightarrow \infty$ ، $\frac{k_n}{n} \rightarrow 0$ و $\frac{k_n}{\sqrt{n}} \rightarrow \infty$ ، آن‌گاه با احتمال یک برای هر β_1 داده‌شده حول β_{01} و هر ثابت a بر اساس رابطه (۱۳.۴)، می‌توانیم بنویسیم

$$Z_n(\beta_1, 0) = \min_{\|\beta_2\| \leq a/\sqrt{n}} Z_n(\beta_1, \beta_2)$$

برای اثبات این ادعا، کافی است نشان دهیم برای هر $e_n = \frac{a}{\sqrt{n}}$ کوچک و $j = p_0 + 1, \dots, p$ داریم

$$\begin{aligned} \frac{Z_n(\beta)}{\partial \beta_j} &< 0 \quad \text{for} \quad -e_n < \beta_j < 0 \\ \frac{Z_n(\beta)}{\partial \beta_j} &> 0 \quad \text{for} \quad 0 < \beta_j < e_n \end{aligned} \quad (18.4)$$

مشتق اول $Z_n(\beta)$ در هر نقطه‌ی مشتق‌پذیر $\beta = (\beta_1, \beta_2, \dots, \beta_p)^\top$ ، نسبت به β_j را در نظر بگیرد، $j = p_0 + 1, \dots, p$ ، داریم

$$\frac{\partial Z_n(\beta)}{\partial \beta_j} = -\left(\frac{\hat{\sigma}_s}{n}\right) \sum_{i=1}^n \psi_1\left(\frac{r_i(\beta)}{\hat{\sigma}_s}\right) x_{ij} + \left(\frac{k_n}{n}\right) \gamma \text{sgn}(\beta_j) |\beta_j|^{\gamma-1}$$

از آن جایی که

$$\left[\left(\frac{\hat{\sigma}_s}{\sqrt{n}} \right) \sum_{i=1}^n \psi_1 \left(\frac{r_i(\beta)}{\hat{\sigma}_s} \right) \mathbf{x}_i \right] \xrightarrow{d} N_p \left(\mathbf{0}, \sigma_0^2 \mathbf{A} \left(\psi_1 \left(\frac{r(\beta_0)}{\sigma_0} \right) \right) \mathbf{C} \mathbf{C} \right)$$

داریم

$$\left(\frac{\hat{\sigma}_s}{\sqrt{n}} \right) \sum_{i=1}^n \psi_1 \left(\frac{r_i(\beta)}{\hat{\sigma}_s} \right) \mathbf{x}_i = O_p(1)$$

و $(\hat{\sigma}_s/\sqrt{n}) \sum_{i=1}^n \psi_1(r_i(\beta)/\hat{\sigma}_s) \mathbf{x}_i$ در احتمال کراندار است، (شائو [۵۸] را ببینید) بنابراین

$$\frac{Z_n(\beta)}{\partial \beta_j} = \left(\frac{k_n}{n} \right) \left\{ -O_p \left(\frac{\sqrt{n}}{k_n} \right) + \gamma \operatorname{sgn}(\beta_j) |\beta_j|^{\gamma-1} \right\}$$

اگر وقتی $n \rightarrow \infty$ ، $\frac{k_n}{n} \rightarrow 0$ و $\frac{k_n}{\sqrt{n}} \rightarrow \infty$ ، آن گاه علامت مشتق توسط علامت β_j مشخص خواهد شد و این نشان دهنده این است که معادله (۱۸.۴) برقرار است. از این رو اثبات کامل است. $\square$

در قسمت دوم گزاره ۲.۳.۴، اگر $\gamma < 1$ ، بنابراین پارامترهای رگرسیونی غیر صفر به صورت استاندارد برآورد شده و اریبی جانبی نخواهند داشت، به عبارت دیگر، برآوردهای پارامترهای رگرسیونی صفر، با احتمالی مثبت کوچک خواهند شد. قسمت اول گزاره ۲.۳.۴ نشان می‌دهد که وقتی $\gamma \geq 1$ ، پارامترهای رگرسیونی غیر صفر، اگر $\lambda_0 > 0$ باشد، به طور جانبی مقدار کمی اریبی دارند. این نتایج با نتایج برآوردگر رگرسیونی بریج (نایت و فو [۴۱]) مشابه می‌باشد. علاوه بر این، گزاره ۲.۳.۴ نشان می‌دهد که برآوردگر ویژگی تنکی را برای حالت $0 < \gamma \leq 1$ دارا است. از این رو، وقتی $0 < \gamma < 1$ ، روش برآورد MM-Bridge، انتخاب متغیر را فراهم می‌کند و برآوردگر به طور جانبی نااریب خواهد بود. برای حالت $\gamma = 1$ ، روش برآورد MM-Bridge که یک برآوردگر لاسو استوار مبتنی بر برآورد MM است، انتخاب متغیر را فراهم می‌کند، اما برآوردگر نتیجه شده، به طور جانبی اریب است. در حالی که برای حالت $\gamma > 1$ ، روش برآورد MM-Bridge برآورد پارامترهای رگرسیونی را منقبض می‌کند اما انتخاب متغیر انجام نمی‌دهد.

۴.۴ الگوریتم محاسبه برآوردهای MM-Bridge

در این بخش، الگوریتمی برای مینیم کردن تابع هدف جریمه شده که در معادله (۵.۴) داده شده، ارائه می‌کنیم. برای حالت $\gamma > 1$ ، تابع هدف جریمه شده نسبت به β در همه جا مشتق پذیر است. از این رو، می‌توانیم با استفاده از الگوریتم‌های مبتنی بر مشتق استاندارد شده، مینیم کنیم. اگرچه برای $0 < \gamma \leq 1$ ، تابع هدف جریمه شده در معادله (۵.۴) مشتق ناپذیر می‌شود، همانند تابع جریمه که در $\beta = 0$ مشتق پذیر نیست. این مشکل تکینگی باعث سخت شدن مینیم کردن تابع هدف جریمه شده، وقتی $0 < \gamma \leq 1$ است، می‌شود. یک رهیافت برای حل مشکلات مینیم سازی مشابه، تقریب زدن تابع جریمه با یک تابع محدب

می باشد. همچنین می تواند با الگوریتم LQA ارائه شده توسط فن و لی [۲۹] یا الگوریتم LLA معرفی شده توسط ژو و لی [۶۷]، انجام شود. این تقریب ها به اجرای مسئله های بهینه سازی پیچیده، مانند معادله (۴.۴)، با استفاده از الگوریتم های استاندارد کمک می کند. این دو الگوریتم اغلب برای پیدا کردن برآوردهای LS جریمه شده و درست نمایی ماکزیمم جریمه شده، استفاده می شوند. برای محاسبه برآوردهای MM-Bridge از الگوریتم LQA استفاده می کنیم. همچنین به معرفی برآوردگر تقریب LS و برآوردگر یک گامی مبتنی بر تقریب LS و الگوریتم LQA می پردازیم.

۱.۴.۴ الگوریتم LQA

فرض کنید $\beta^{(0)}$ اولیه که تعریف می کنیم، به مقدار واقعی β نزدیک باشد. در صورتی که $\beta^{(0)}$ مولفه هایی نزدیک به صفر داشته باشد، آن مولفه ها را صفر قرار می دهیم. می توانیم برآورد MM جریمه نشده را به عنوان مقدار اولیه $\beta^{(0)}$ قرار دهیم. فن و لی [۲۹] تقریبی موضعی برای تابع جریمه با استفاده از تابع جریمه درجه دوم را در $\beta^{(0)} = (\beta_1^{(0)}, \beta_2^{(0)}, \dots, \beta_p^{(0)})^T$ به صورت زیر ارائه دادند

$$|\beta_j|^\gamma \approx |\beta_j^{(0)}|^\gamma + \frac{\gamma}{2} |\beta_j^{(0)}|^{\gamma-2} (\beta_j^2 - (\beta_j^{(0)})^2)$$

با استفاده از این تقریب، مسئله مینیمم سازی (۴.۴) می تواند به صورت زیر بازنویسی شود

$$\hat{\beta}_{\text{MM-BR}} = \arg \min_{\beta} L_n^*(\beta) \quad (۱۹.۴)$$

با تابع هدف جریمه شده زیر

$$\begin{aligned} L_n^*(\beta) &= \frac{\hat{\sigma}_s^2}{n} \sum_{i=1}^n \rho_1\left(\frac{r_i(\beta)}{\hat{\sigma}_s}\right) + \frac{k_n \gamma}{2n} \sum_{j=1}^p |\beta_j^{(0)}|^{\gamma-2} \beta_j^2 \\ &= \frac{\hat{\sigma}_s^2}{n} \sum_{i=1}^n \rho_1\left(\frac{r_i(\beta)}{\hat{\sigma}_s}\right) + \frac{k_n \gamma}{2n} \beta^T W_0(\beta^{(0)}) \beta \end{aligned}$$

تابع هدف $L_n^*(\beta)$ می تواند به عنوان تابع هدف برای یک مسئله رگرسیونی استوار موزون در نظر گرفته شود. در برآورد رگرسیونی ریح استوار ارائه شده توسط مارونا [۴۷] تابع جریمه دارای شکل $\sum_{j=1}^p \beta_j^2$ است، که در آن وزن های متفاوتی برای پارامترهای متفاوت تعریف می کنیم. با مساوی صفر قرار دادن مشتق $L_n^*(\beta)$ نسبت به β ، به دست می آید

$$(\mathbf{X}^T \mathbf{W}(\beta) \mathbf{X} + k_n \gamma \mathbf{W}_0(\beta^{(0)})) \beta = \mathbf{X}^T \mathbf{W}(\beta) \mathbf{Y}$$

از این رو به دست می آوریم

$$\hat{\beta} = (\mathbf{X}^T \mathbf{W}(\hat{\beta}) \mathbf{X} + k_n \gamma \mathbf{W}_0(\beta^{(0)}))^{-1} \mathbf{X}^T \mathbf{W}(\hat{\beta}) \mathbf{Y} \quad (۲۰.۴)$$

با این قید که $(\mathbf{X}^\top \mathbf{W}(\hat{\beta})\mathbf{X} + k_n \gamma \mathbf{W}_0(\beta^{(0)}))$ پرتبه است، با k_n, γ و $\beta^{(0)}$ مشخص، برآوردگر موزون برای $\hat{\beta}$ در (۲۰.۴)، الگوریتم موزون مکرر زیر را پیشنهاد می‌دهد

$$\beta^{k+1} = (\mathbf{X}^\top \mathbf{W}(\beta^{(k)})\mathbf{X} + k_n \gamma \mathbf{W}_0(\beta^{(k)}))^{-1} \mathbf{X}^\top \mathbf{W}(\beta^{(k)})\mathbf{Y} \quad (21.4)$$

برای $k = 0, 1, 2, 3, \dots$ ، نقطه همگرایی دنباله $\{\beta^{(k)}\}$ برای $k = 0, 1, 2, 3, \dots$ می‌تواند به‌عنوان برآورد MM-Bridge برای پارامتر رگرسیونی β باشد.

وقتی $\gamma < 2$ و تعدادی از مولفه‌های $\beta^{(k)}$ بسیار نزدیک به صفر هستند، گوییم برای z_j که $|\beta_j^{(k)}| < c$ (که در آن c قیدی کوچک و از پیش تعیین شده است) به‌سبب مولفه‌های قطری $\mathbf{W}_0(\beta^{(k)})$ ، ناپایداری عددی به‌وجود می‌آید. برای جلوگیری از این ناپایداری فن و لی [۲۹] پیشنهاد دادند که آن مولفه‌ها را برابر با صفر قرار داده و از x در تکرارهای بعدی حذف کرد. هانتر و لی [۳۷] نیز پیشنهاد کردند که برای جلوگیری از ناپایداری عددی $|\beta_j^k|$ را با $|\beta_j^k + c_0|$ جایگزین کرد، که در آن c_0 مقداری کوچک از اغتشاش و از پیش تعیین شده می‌باشد. توجه کنید که برای اجرای مینیمم‌سازی تابع هدف جریمه شده در معادله (۵.۴) نیز می‌توان از الگوریتم LLA معرفی شده توسط ژو و لی [۶۷] به‌جای الگوریتم LQA استفاده کنیم. اما از آن‌جا که الگوریتم LQA را برای محاسبه‌ی برآوردهای MM-Bridge استفاده می‌کنیم، این الگوریتم‌ها را شرح نمی‌دهیم.

۲.۴.۴ تقریب کمترین توان‌های دوم (LS): برآوردگر یک مرحله‌ای

تقریب LS را برای ساده کردن مسئله مینیمم‌سازی (۴.۴) به یک مسئله مینیمم‌سازی درجه دوم استفاده می‌کنیم. این تقریب توسط فن و لی [۲۹] و همچنین توسط وانگ و لنگ [۶۱] و لی و یو [۴۳] برای حل مسائل درست‌نمایی جریمه شده و LASSO استفاده شده است. این تقریب به ما اجازه می‌دهد یک برآوردگر یک مرحله‌ای مبتنی بر الگوریتم LQA تعریف کنیم. با استفاده از بسط سری تیلور در $\hat{\beta}_{MM}$ ، اولین عبارت معادله (۵.۴) را می‌توان به‌شکل زیر به‌صورت موضعی تقریب زد

$$\begin{aligned} \hat{\sigma}_s^2 \sum_{i=1}^n \rho_1 \left(\frac{r_i(\beta)}{\hat{\sigma}_s} \right) &\approx \hat{\sigma}_s^2 \sum_{i=1}^n \rho_1 \left(\frac{r_i(\hat{\beta}_{MM})}{\hat{\sigma}_s} \right) \\ &\quad - \left[\hat{\sigma}_s \sum_{i=1}^n \psi_1 \left(\frac{r_i(\hat{\beta}_{MM})}{\hat{\sigma}_s} \right) \mathbf{x}_i \right]^\top (\beta - \hat{\beta}_{MM}) \\ &\quad + \frac{1}{2} (\beta - \hat{\beta}_{MM})^\top \left[\sum_{i=1}^n \psi_1' \left(\frac{r_i(\hat{\beta}_{MM})}{\hat{\sigma}_s} \right) \mathbf{x}_i \mathbf{x}_i^\top \right] (\beta - \hat{\beta}_{MM}) \end{aligned} \quad (22.4)$$

با چشم‌پوشی از عبارت ثابت و توجه به این که $\sum_{i=1}^n \psi_1(r_i(\hat{\beta}_{MM})/\hat{\sigma}_s) \mathbf{x}_i = 0$ ، تقریب درجه دوم زیر را داریم

$$\hat{\sigma}_s^2 \sum_{i=1}^n \rho_1 \left(\frac{r_i(\beta)}{\hat{\sigma}_s} \right) \approx \frac{1}{2} (\beta - \hat{\beta}_{MM})^\top (\mathbf{X}^\top \mathbf{W}^*(\hat{\beta}_{MM}) \mathbf{X}) (\beta - \hat{\beta}_{MM}) \quad (23.4)$$

که در آن $W^*(\hat{\beta}_{MM}) = \text{diag}(\psi'_1(r_1(\hat{\beta}_{MM})/\hat{\sigma}_s), \dots, \psi'_p(r_p(\hat{\beta}_{MM})/\hat{\sigma}_s))$ با استفاده از این تقریب، تابع هدف جریمه شده (۵.۴) را می‌توان به صورت زیر بازنویسی کرد

$$L_n(\beta) \approx \frac{1}{\gamma} (\beta - \hat{\beta}_{MM})^\top (X^\top W^*(\hat{\beta}_{MM}) X) (\beta - \hat{\beta}_{MM}) + k_n \sum_{j=1}^p |\beta_j|^\gamma \quad (24.4)$$

تابع هدف تقریب زده شده (۲۴.۴) با تابع هدف رگرسیونی بریج داده شده در معادله (۲.۴) یکسان است. بنابراین با این تقریب، مسئله مینیمم‌سازی رگرسیونی MM-Bridge داده شده در (۴.۴)، می‌تواند به صورت هم‌ارز با مسئله مینیمم‌سازی LS جریمه‌شده نوشته شود

$$\tilde{\beta}_{MM-BR} = \arg \min_{\beta} \left[\frac{1}{\gamma} (\beta - \hat{\beta}_{MM})^\top (X^\top W^*(\hat{\beta}_{MM}) X) (\beta - \hat{\beta}_{MM}) + k_n \sum_{j=1}^p |\beta_j|^\gamma \right] \quad (25.4)$$

علاوه بر این، اگر از الگوریتم LQA استفاده کنیم و قرار دهیم $\beta^{(0)} = \hat{\beta}_{MM}$ ، به برآوردگر یک مرحله‌ای زیر خواهیم رسید

$$\tilde{\beta}_{MM-BR}^{(1)} = (D + k_n \gamma W_\circ(\hat{\beta}_{MM}))^{-1} D \hat{\beta}_{MM} \quad (26.4)$$

که در آن $D = X^\top W^*(\hat{\beta}_{MM}) X$ است. این یک برآوردگر نوع ریج مبتنی بر $\hat{\beta}_{MM}$ می‌باشد.

۵.۴ شبیه‌سازی

در این بخش برای ارزیابی عملکرد و مقایسه روش رگرسیونی بریج استوار (MM-BR) بر حسب برآورد استوار پارامترهای رگرسیونی با نمونه متناهی، یک مطالعه شبیه‌سازی در دو طرح برای حالت $n > p$ و دو حالت با حضور داده ناهگون در متغیر پاسخ و تبیینی و بدون داده ناهمگون بررسی می‌کنیم. برآوردگر رگرسیونی بریج استوار (MM-BR) را با برآوردگرهای رگرسیونی ریج و لاسو معمولی و همچنین M-برآوردگر استوار مقایسه خواهیم کرد. لازم به ذکر است تمام محاسبات انجام شده در این مثال با استفاده از نرم‌افزار R انجام می‌شود. همچنین همانند شبیه‌سازی ۵.۳ و ۴.۲، برآوردگرهای ریج و لاسو معمولی با استفاده از پکیج glmnet در نرم‌افزار R و همچنین M-برآوردگر استوار نیز با استفاده از پکیج robust در نرم‌افزار R استفاده خواهیم کرد. همچنین برای محاسبه برآوردگر MM-Bridge از الگوریتم موزون مکرر (۲۱.۴) استفاده می‌کنیم. برای سنجش مدل‌های معرفی شده، معیارهایی همانند معیارهای شبیه‌سازی ۴.۲ استفاده می‌کنیم، در اینجا منظور از REff(MPE) نسبت خطای پیش‌گویی برآوردگر ریج نسبت به سایر برآوردگرها است. در مدل MM-Bridge نیز برای راحتی انجام محاسبات $\gamma = 0.5$ انتخاب می‌شود. بنابراین برآوردگر معرفی شده قادر به انتخاب متغیر هم خواهد بود.

طرح ۱. در این طرح از روش تولید داده طرح ۱ شبیه‌سازی بخش ۴.۲ استفاده خواهیم کرد و همچنین تعداد پارامتر $p = 8$ ، $n = 50$ و بردار پارامتر واقعی را، $\beta_0 = (3, 1/5, 0, 0, 2, 0, 0, 0)^\top$ در نظر می‌گیریم.

جدول ۱.۴: نتایج شبیه‌سازی مربوط به بریج استوار، طرح ۱، $n = 50$ ، $\rho = 0.3$ و $\tau = 0$

MM – BR		MM		LASSO		Ridge		ضرایب
MSE	mean	MSE	mean	MSE	mean	MSE	mean	
0.1667	2.731	0.052	2.9831	0.1566	2.7562	0.54	2.4163	$\hat{\beta}_1$
0.2596	1.1148	0.0762	1.5071	0.1205	1.299	0.059	1.3586	$\hat{\beta}_2$
0.0047	0.0398	0.0563	0.0357	0.0073	0.0264	0.0777	0.1993	$\hat{\beta}_3$
0.0024	0.0269	0.0473	-0.0246	0.0067	0.0125	0.0576	0.1746	$\hat{\beta}_4$
0.249	1.5788	0.0548	2.0292	0.2561	1.6527	0.4054	1.4649	$\hat{\beta}_5$
0.0058	0.0393	0.0833	-0.0343	0.0062	0.0201	0.0533	0.1525	$\hat{\beta}_6$
0.005	0.032	0.0622	0.0064	0.0035	0.0024	0.0235	0.0497	$\hat{\beta}_7$
0.0036	0.0341	0.0467	0.0091	0.0071	0.0207	0.0415	0.131	$\hat{\beta}_8$
0.6968	—	0.4788	—	0.5639	—	1.258	—	MSE
				MM – BR	MM	LASSO	Ridge	
				8.5069	7.3313	8.686	9.4403	$\sqrt{\text{MPE}}$
				1.1097	1.2877	1.0868	1	REff
				28.083	28.083	28.083	28.083	MCN
				4.47	1.57	4.42	—	C
				0	0	0	—	IC

جدول ۲.۴: نتایج شبیه‌سازی مربوط به بریج استوار، طرح ۱، $n = 50$ ، $\rho = 0.99$ و $\tau = 0$

MM – BR		MM		LASSO		Ridge		ضرایب
MSE	mean	MSE	mean	MSE	mean	MSE	mean	
3.0392	1.7148	2.3801	2.8858	1.6758	2.3575	4.5256	0.8858	$\hat{\beta}_1$
1.2408	1.0594	2.4827	1.5478	1.1835	1.1072	0.5204	0.7994	$\hat{\beta}_2$
0.5058	0.4389	2.5746	0.2411	0.3675	0.2988	0.5114	0.7051	$\hat{\beta}_3$
0.3279	0.3226	2.1608	-0.166	0.2036	0.1877	0.4549	0.6673	$\hat{\beta}_4$
1.6229	1.1623	2.5056	2.1972	1.7915	1.0378	1.6358	0.7278	$\hat{\beta}_5$
0.6028	0.5229	2.8099	-0.232	0.3108	0.2313	0.4083	0.6299	$\hat{\beta}_6$
0.5557	0.3693	2.8453	0.0434	0.1685	0.1572	0.4047	0.6266	$\hat{\beta}_7$
1.0232	0.6672	4.2126	-0.0045	0.5455	0.3194	0.4184	0.6399	$\hat{\beta}_8$
8.9182	—	23.9717	—	6.2467	—	8.8795	—	MSE
				MM – BR	MM	LASSO	Ridge	
				6.9754	7.3313	1.00426	9.7896	$\sqrt{\text{MPR}}$
				1.4035	1.3353	0.9748	1	REff
				936.4978	936.4978	936.4978	936.4978	MCN
				1.06	0.18	3.68	—	C
				0.26	0.05	0.54	—	IC

جدول ۳.۴: نتایج شبیه‌سازی مربوط به بریج استوار، طرح ۱، $n = 200$ ، $\rho = 0/3$ و $\tau = 0$

MM – BR		MM		LASSO		Ridge		ضرایب
<i>MSE</i>	<i>mean</i>	<i>MSE</i>	<i>mean</i>	<i>MSE</i>	<i>mean</i>	<i>MSE</i>	<i>mean</i>	
0/2004	2/5881	0/0088	3/0029	0/0086	2/9672	0/0869	2/7175	$\hat{\beta}_1$
0/2717	0/9931	0/0083	1/4994	0/0085	1/4606	0/0093	1/4428	$\hat{\beta}_2$
0/0003	0/0102	0/0106	-0/011	0/0033	0/0078	0/0144	0/0862	$\hat{\beta}_3$
0/0003	0/0111	0/0113	0/0001	0/0034	0/0159	0/0198	0/1159	$\hat{\beta}_4$
0/351	1/4221	0/0089	1/9833	0/0121	1/9259	0/1066	1/6861	$\hat{\beta}_5$
0/0003	0/01	0/0103	0/0005	0/0032	0/0153	0/0205	0/1169	$\hat{\beta}_6$
0/0002	0/0107	0/0117	0/0089	0/0026	0/0106	0/0072	0/0181	$\hat{\beta}_7$
0/0002	0/0105	0/009	0/0018	0/0022	0/0069	0/0092	0/0633	$\hat{\beta}_8$
0/8243	—	0/0789	—	0/0438	—	0/274	—	MSE
				MM – BR	MM	LASSO	Ridge	
				20/5238	14/7669	14/6649	15/4259	$\sqrt{\text{MPE}}$
				0/7516	1/0446	1/0519	1	REff
				19/14	19/14	19/14	19/14	MCN
				5	3/32	4/55	—	C
				0	0	0	—	IC

جدول ۴.۴: نتایج شبیه‌سازی مربوط به بریج استوار، طرح ۱، $n = 200$ ، $\rho = 0/99$ و $\tau = 0$

MM – BR		MM		LASSO		Ridge		ضرایب
<i>MSE</i>	<i>mean</i>	<i>MSE</i>	<i>mean</i>	<i>MSE</i>	<i>mean</i>	<i>MSE</i>	<i>mean</i>	
0/8109	2/3487	0/4033	3/0194	0/3072	2/8912	3/6523	1/0929	$\hat{\beta}_1$
0/6614	0/8706	0/3786	1/4956	0/3552	1/2384	0/3439	0/9183	$\hat{\beta}_2$
0/0832	0/2051	0/486	-0/0741	0/0826	0/1182	0/5106	0/7118	$\hat{\beta}_3$
0/0906	0/2195	0/5149	0/0006	0/0856	0/1343	0/4803	0/6906	$\hat{\beta}_4$
0/9259	1/1918	0/4063	1/8871	0/5099	1/5008	1/4409	0/8014	$\hat{\beta}_5$
0/1138	0/2107	0/472	0/0032	0/1152	0/1477	0/4421	0/6621	$\hat{\beta}_6$
0/1082	0/248	0/5354	0/06	0/0678	0/1109	0/4164	0/6423	$\hat{\beta}_7$
0/2906	0/3778	0/9592	0/0925	0/1748	0/2138	0/4773	0/6896	$\hat{\beta}_8$
3/0846	—	4/1557	—	1/6983	—	7/7638	—	MSE
				MM – BR	MM	LASSO	Ridge	
				19/8448	14/7669	15/0154	15/6089	$\sqrt{\text{MPE}}$
				0/7865	1/057	1/0395	1	REff
				635/8134	635/8134	635/8134	635/8134	MCN
				1/88	0/47	3/67	—	C
				0/06	0/01	0/03	—	IC

همان طور که در جدول های ۱.۴ و ۲.۴ مشاهده می شود در زمانی که داده ناهمگون بین مجموعه داده وجود ندارد و همخطی کمی بین متغیرها تبیینی موجود است، M-برآوردگر استوار عملکرد بهتری در برآورد ضرایب و پیشگویی نشان می دهد اما با افزایش همخطی در بین متغیرها، روش های انقباضی عملکرد بهتری دارند و برآوردهای بهتری به دست می دهند، اما برآوردهای ریج و لاسو معمولی در پیشگویی مدل عملکردی به خوبی برآوردهای استوار نداشته اند که این امر می تواند نتیجه شناسایی بعضی از داده ها به عنوان داده پرت باشد که با افزایش حجم داده، جدول های ۳.۴ و ۴.۴، اثر این داده ها کم می شود.

جدول ۵.۴: نتایج شبیه سازی مربوط به بریج استوار، طرح ۱، $n = 50$ ، $\rho = 0.9$ و $\tau = 0.2$

MM - BR		MM		LASSO		Ridge		ضرایب
MSE	mean	MSE	mean	MSE	mean	MSE	mean	
0.2559	2.9929	0.2319	2.9633	8.0863	0.1986	5.7007	0.6131	$\hat{\beta}_1$
0.4493	1.0831	0.3805	1.1549	2.25	0	0.5223	0.7804	$\hat{\beta}_2$
0.4336	0.3699	0.466	0.3629	0.5336	0.3569	0.2889	0.535	$\hat{\beta}_3$
0.4192	-0.4598	0.4892	-0.5099	0	0	0.6168	0.7818	$\hat{\beta}_4$
0.2366	2.2427	0.2576	2.2692	2.1305	0.7629	2.3321	0.4735	$\hat{\beta}_5$
0.2954	0.2957	0.3338	0.2814	0.747	0.5584	0.179	0.4206	$\hat{\beta}_6$
0.2958	-0.3772	0.36	-0.4075	0	0	0.4167	0.6431	$\hat{\beta}_7$
0.3367	0.3373	0.3707	0.3535	1.7289	1.0923	0.1394	0.3716	$\hat{\beta}_8$
2.7225	-	2.8898	-	15.4762	-	10.1959	-	MSE
MM - BR		MM		LASSO		Ridge		
21.8177		22.2047		69.2137		67.5366		$\sqrt{MPE}$
3.0955		3.0415		0.9758		1		REff
434.9439		434.9439		434.9439		434.9439		MCN
1.25		0.58		3.21		-		C
0.05		0.02		2.04		-		IC

با ایجاد آلودگی در مجموعه داده، همان طور که مشاهده می شود، جدول های ۵.۴ و ۶.۴، با افزایش همخطی، برآوردگر MM-Bridge عملکرد بهتری در برآورد ضرایب و پیشگویی مدل دارد. همچنین در بحث انتخاب متغیر هم عملکرد بهتری نسبت به برآوردگر استوار معمولی دارد.

طرح ۲. در این طرح نیز همانند طرح ۲ داده ها را از روش شبیه سازی بخش ۴.۲ تولید می کنیم و بردار ضرایب را به صورت $\beta_0 = (\underbrace{0, \dots, 0}_{10}, \underbrace{2, \dots, 2}_{10}, \underbrace{0, \dots, 0}_{10}, \underbrace{2, \dots, 2}_{10})^T$ در نظر می گیریم، در این مدل گروه هایی از متغیرهای موثر و غیر موثر وجود دارند که همبستگی بین آن ها یکسان است، این طرح را برای حجم های $n = 100, 300$ بررسی می کنیم.

در عدم حضور داده ناهمگون و همخطی ضعیف بین متغیرهای تبیینی، برآوردهای معمولی نتایج خوبی به دست می دهند، در جدول های ۹.۴ و ۱۰.۴، مشاهده می شود که بر اثر افزایش همخطی، این برآوردها نیز کارایی خود را از دست می دهند و برآوردهای جریمه شده

جدول ۶.۴: نتایج شبیه‌سازی مربوط به بریج استوار، طرح ۱، $n = 50$ ، $\rho = 0.99$ و $\tau = 0.2$

MM – BR		MM		LASSO		Ridge		ضرایب
<i>MSE</i>	<i>mean</i>	<i>MSE</i>	<i>mean</i>	<i>MSE</i>	<i>mean</i>	<i>MSE</i>	<i>mean</i>	
۲/۱۷۶۲	۱/۷۳۲۱	۱/۱۴۲۱	۲/۴۵۶۶	۸/۶۲۱۵	۰/۰۹۰۶	۵/۷۶۹۵	۰/۵۹۸۸	$\hat{\beta}_1$
۱/۴۳۳۶	۰/۳۷۳۸	۱/۶۸۱۶	۰/۸۱۳۹	۲/۲۵	۰	۰/۵۴۱	۰/۷۶۷۴	$\hat{\beta}_2$
۰/۶۲۸۵	۰/۵۲۱۳	۱/۳۸۲۳	۰/۴۳۷۱	۰/۳۸۹۱	۰/۲۸۱۱	۰/۲۷۹۷	۰/۵۲۶۴	$\hat{\beta}_3$
۰/۰۸۱۹	−۰/۰۳۹۷	۱/۱۴۲۷	−۰/۵۱۵۸	۰	۰	۰/۶۳۹۳	۰/۷۹۶۵	$\hat{\beta}_4$
۱/۱۲۵۵	۱/۷۹۰۱	۱/۱۶۶۱	۲/۰۲۰۷	۲/۳۶۷۸	۰/۷۰۳۹	۲/۳۵۶۹	۰/۴۶۵۴	$\hat{\beta}_5$
۱/۱۱۳۲	۰/۷۵۳۴	۱/۴۴۲۸	۰/۵۳۲۸	۰/۹۴۴۶	۰/۶۶۰۳	۰/۱۷۵۲	۰/۴۱۵۷	$\hat{\beta}_6$
۰/۰۵۵۱	۰/۰۸۷۹	۱/۱۱۱۵	−۰/۵۵۴۳	۰	۰	۰/۴۵۴۷	۰/۶۷۱۵	$\hat{\beta}_7$
۲/۵۱۸۸	۱/۲۵۳۳	۳/۱۴۱۴	۱/۱۸۰۱	۱/۹۷۴۱	۱/۱۶۸۱	۰/۱۴۰۸	۰/۳۷۳	$\hat{\beta}_8$
۹/۱۳۲۶	−	۱۲/۲۱۰۶	−	۱۶/۵۴۷	−	۱۰/۳۵۶۸	−	MSE
				MM – BR	MM	LASSO	Ridge	
				۱۰/۵۲۳۶	۱۵/۱۷۰۴	۶۹/۵۱۳۲	۶۷/۸۲۶۱	$\sqrt{\text{MPE}}$
				۶/۴۴۵۲	۴/۴۷۰۹	۰/۹۷۵۷	۱	REff
				۱۹۵۲/۱۸۵۵	۱۹۵۲/۱۸۵۵	۱۹۵۲/۱۸۵۵	۱۹۵۲/۱۸۵۵	MCN
				۱/۷۲	۰/۴	۳/۱۸	−	C
				۰/۴۲	۰/۰۸	۲/۳۱	−	IC

جدول ۷.۴: نتایج شبیه‌سازی مربوط به بریج استوار، طرح ۱، $n = 200$ ، $\rho = 0.9$ و $\tau = 0.2$

MM – BR		MM		LASSO		Ridge		ضرایب
MSE	mean	MSE	mean	MSE	mean	MSE	mean	
۰/۱۲۰۲	۲/۸۹۷۱	۰/۰۵۷۳	۲/۹۸۳۹	۸/۳۴	۰/۱۱۹۵	۵/۵۸۵۵	۰/۶۳۶۸	$\hat{\beta}_1$
۰/۹۱۶	۰/۶۲۴۸	۰/۲۰۵۳	۱/۲۰۹۷	۲/۲۵	۰	۰/۴۸۶۳	۰/۸۰۳۳	$\hat{\beta}_2$
۰/۰۷۰۱	۰/۱۴۸	۰/۲۰۹	۰/۲۷۷۶	۰/۰۸۳	۰/۱۴۶۶	۰/۳۰۳۲	۰/۵۵	$\hat{\beta}_3$
۰/۰۰۳۱	۰/۰۱۹	۰/۲۵۹	−۰/۳۸۲۱	۰	۰	۰/۶۳۹۶	۰/۷۹۸۸	$\hat{\beta}_4$
۰/۱۷۰۳	۱/۹۰۱۷	۰/۰۷۸	۲/۱۱۳۴	۲/۰۵۸۴	۰/۶۳۲۲	۲/۲۸۹۳	۰/۴۸۷۱	$\hat{\beta}_5$
۰/۰۴۴	۰/۱۱۹۷	۰/۱۲۵۲	۰/۲۰۷	۰/۵۵۳۱	۰/۶۲۶۴	۰/۱۸۸۵	۰/۴۳۳۸	$\hat{\beta}_6$
۰/۰۰۷	۰/۰۶۱۶	۰/۲۳۰۷	−۰/۳۳۳۲	۰	۰	۰/۴۳۸	۰/۶۶۱۳	$\hat{\beta}_7$
۰/۲۰۶۳	۰/۲۹۳۸	۰/۳۰۶۳	۰/۳۸۳۲	۲/۱۹۷۶	۱/۴۳۳۴	۰/۱۴۷۲	۰/۳۸۳۴	$\hat{\beta}_8$
۱/۵۳۷	−	۱/۴۷۰۶	−	۱۵/۴۸۲	−	۱۰/۰۷۷۶	−	MSE
				MM – BR	MM	LASSO	Ridge	
				۷۱/۷۰۰۳	۵۰/۶۸۲۲	۱۳۲/۰۰۵۸	۱۳۰/۸۷۹۴	$\sqrt{\text{MPE}}$
				۱/۸۲۵۴	۲/۵۸۲۴	۰/۹۹۱۵	۱	REff
				۲۰۷/۹۷۷۱	۲۰۷/۹۷۷۱	۲۰۷/۹۷۷۱	۲۰۷/۹۷۷۱	MCN
				۳/۴۸	۱/۰۷	۲/۷۴	−	C
				۰/۰۴	۰	۱/۸	−	IC

جدول ۸.۴: نتایج شبیه‌سازی مربوط به بریج استوار، طرح ۱، $n = 200$ ، $\rho = 0.99$ و $\tau = 0.2$

MM – BR		MM		LASSO		Ridge		ضرایب
MSE	mean	MSE	mean	MSE	mean	MSE	mean	
۲/۵۶۵۹	۱/۵۸۰۵	۱/۶۷۴۶	۱/۹۵۳	۸/۹۷۴۵	۰/۰۰۴۴	۵/۶۷۱۵	۰/۶۱۸۶	$\hat{\beta}_1$
۱/۶۰۲۸	۰/۲۵۵۶	۰/۹۸۹۶	۰/۶۶۱۴	۲/۲۵	۰	۰/۵۰۲۲	۰/۷۹۱۸	$\hat{\beta}_2$
۰/۵۳۹۱	۰/۵۸۸۳	۰/۸۸۹۸	۰/۷۴۳۱	۰/۰۶۲۳	۰/۱۱۲۳	۰/۲۹۲۹	۰/۵۴۰۷	$\hat{\beta}_3$
۰/۰۰۳۹	۰/۰۳۹۱	۰/۲۹۶۳	−۰/۳۱۳۸	۰	۰	۰/۶۷۰۱	۰/۸۱۷۸	$\hat{\beta}_4$
۰/۶۸۸۹	۱/۵۰۶۶	۰/۵۰۳۳	۱/۶۱۰۴	۲/۱۶۶۸	۰/۵۸۰۶	۲/۳۱۴۷	۰/۴۷۸۷	$\hat{\beta}_5$
۰/۸۵۲۱	۰/۷۴۵۷	۰/۹۱۸۹	۰/۷۸۱۴	۰/۹۰۶۲	۰/۸۴۸۵	۰/۱۸۳۷	۰/۴۲۸۲	$\hat{\beta}_6$
۰/۰۳۴۳	۰/۱۳۵	۰/۳۸۹۳	−۰/۲۷۵۴	۰	۰	۰/۴۷۳۵	۰/۶۸۷۶	$\hat{\beta}_7$
۲/۴۱۷۷	۱/۴۰۳۳	۱/۸۶۵۱	۱/۲۴۰۵	۲/۰۲۷۸	۱/۳۷۱۶	۰/۱۴۷۲	۰/۳۸۳۲	$\hat{\beta}_8$
۸/۷۰۴۶	—	۷/۵۲۶۹	—	۱۶/۳۸۷۶	—	۱۰/۲۵۵۷	—	MSE
				MM – BR	MM	LASSO	Ridge	
				۲۷/۳۸۰۴	۲۳/۶۷۹۸	۱۳۲/۳۷۶۱	۱۳۱/۵۵۱۷	$\sqrt{MPE}$
				۴/۸۰۴۶	۵/۵۵۵۴	۰/۹۹۳۸	۱	REff
				۳۵۵/۲۱۱۷	۳۵۵/۲۱۱۷	۳۵۵/۲۱۱۷	۳۵۵/۲۱۱۷	MCN
				۱/۷۸	۰/۴۳	۲/۷۵	—	C
				۰/۳۲	۰/۰۶	۲/۱	—	IC

با غلبه بر این همخطی نتایج بهتری در برآورد ضرایب به دست می‌دهند. اما در حجم‌های بالاتر برآوردگر بریج استوار از برآوردگرهای ریدج و لاسو عملکرد بهتری دارد و تقریباً مشابه با M - برآوردگر استوار عمل می‌کند.

جدول‌های ۱۳.۴ و ۱۴.۴ نتایج برآوردها در حضور داده دورافتاده در متغیر پاسخ و ۲۰ درصد داده آلوده در متغیرهای تبیینی (نقاط نافذ) هستند، همان‌طور که مشاهده می‌شود، وقتی بین متغیرهای تبیینی همخطی شدید موجود است، برآوردگر $MM-BR$ عملکرد بهتری در برآورد ضرایب و پیشگویی، نسبت به بقیه برآوردگرها دارد. همچنین در افزایش حجم داده، جدول‌های ۱۵.۴ و ۱۶.۴، با افزایش همخطی بین متغیرهای تبیینی، عملکرد بهتری را از برآوردگر $MM-BR$ مشاهده می‌کنیم.

۶.۴ تحلیل داده‌های چربی بدن

در این بخش رفتار برآوردگر بریج استوار معرفی شده را در قالب یک مثال واقعی مورد ارزیابی قرار می‌دهیم. در این مثال مجموعه داده‌های $Vo2$ معرفی شده در مثال ۵.۲ استفاده می‌شود و برآوردگر $MM-Bridge$ را با برآوردگرهای ریدج و لاسو معمولی و M - برآوردگر معرفی شده در مثال شبیه‌سازی مقایسه می‌کنیم.

در این مثال نیز همانند بخش ۵.۲ برای مقایسه برآوردگرها از معیار جذر میانگین خطای پیشگویی (۱۹.۲) در بخش ۴.۲ استفاده می‌کنیم. در اینجا برآوردگر بریج را با برآوردگرهای ریدج و لاسو معمولی و M - برآوردگر استوار مقایسه می‌کنیم. مشاهده می‌شود برآوردگر بریج

جدول ۹.۴: نتایج شبیه‌سازی مربوط به بریج استوار، طرح ۲، $n = ۱۰۰$ ، $\rho = ۰/۳$ و $\tau = ۰$

ضرایب	Ridge		LASSO		MM		MM - BR	
	MSE	mean	MSE	mean	MSE	mean	MSE	mean
$\hat{\beta}_1$	۰/۷۹۳	۰/۲۱۳۶	۰	۰	۰/۴۳۹	۰/۰۲۳۵	۰/۰۷۱۶	۰/۰۲۹
$\hat{\beta}_2$	۰/۰۶۱۸	۰/۱۷۹۸	۰	۰/۰۰۰۳	۰/۰۵۶۹	۰/۰۰۶۶	۰/۰۳۶۸	۰/۰۴۴۴
$\hat{\beta}_3$	۰/۰۴۱۸	۰/۱۲۳۹	۰	۰	۰/۰۵۶۸	۰/۰۰۳۳	۰/۰۰۰۳	۰/۰۴۱۹
$\hat{\beta}_4$	۰/۰۴۲۴	۰/۱۱۵۸	۰	۰/۰۰۰۷	۰/۰۷۴۴	۰/۰۳۸۲	۰/۰۳۳۲	۰/۰۵۸
$\hat{\beta}_5$	۰/۰۵۳۹	۰/۱۳۳۸	۰	۰/۰۰۰۴	۰/۰۵۲۵	۰/۰۱۲۹	۰/۰۱۸۷	۰/۰۳۶۵
$\hat{\beta}_6$	۰/۰۵۳۵	۰/۱۲۵۷	۰	۰/۰۰۰۴	۰/۰۶۱۳	۰/۰۰۲۰۷	۰/۰۰۰۰۷	۰/۰۴۴۱
$\hat{\beta}_7$	۰/۰۵۱۲	۰/۱۳۲۱	۰	۰	۰/۰۵۰۳	۰/۰۰۴۸	۰/۰۰۰۱۷	۰/۰۲۹۶
$\hat{\beta}_8$	۰/۰۴۸۷	۰/۱۴۵۶	۰	۰	۰/۰۵۸۶	۰/۰۰۰۱	۰/۰۰۰۷	۰/۰۴۱۱
$\hat{\beta}_9$	۰/۰۸۲	۰/۲۳۰۷	۰/۰۰۰۲	۰/۰۰۱۴	۰/۰۵۵۴	۰/۰۰۳۹۷	۰/۰۰۲۷۳	۰/۰۴۱۵
$\hat{\beta}_{10}$	۰/۲۳۵۵	۰/۴۴۶۲	۰/۰۰۸	۰/۰۱۸۳	۰/۰۵۵۸	۰/۰۴۳۹	۰/۰۵۰۴	۰/۰۴۱۹
$\hat{\beta}_{11}$	۰/۴۳۹۳	۱/۳۷۰۱	۱/۲۸۹۵	۰/۷۰۶۷	۱/۹۷۷۱	۰/۹۷۷۱	۱/۸۹۴	۰/۰۹۲۲
$\hat{\beta}_{12}$	۰/۲۰۲۷	۱/۶۰۴۶	۱/۶۵۹۹	۰/۳۸۸۵	۲/۰۴۱۶	۲/۰۴۱۶	۱/۹۹۷	۰/۰۹۲۷
$\hat{\beta}_{13}$	۰/۱۹۷۱	۱/۶۱۹۱	۱/۵۷۱	۰/۴۷۵۳	۱/۹۵۸	۱/۹۵۸	۱/۸۸۳۳	۰/۱۰۴۸
$\hat{\beta}_{14}$	۰/۱۴۴۵	۱/۶۹۷۷	۱/۷۰۵۴	۰/۳۰۹	۲/۰۱۷۴	۲/۰۱۷۴	۱/۹۸۴۹	۰/۰۶۹۱
$\hat{\beta}_{15}$	۰/۱۳۵	۱/۶۸۷۹	۱/۶۳۷۵	۰/۳۶۹۹	۱/۹۷۳۵	۱/۹۷۳۵	۱/۹۳۷۱	۰/۰۵۷۷
$\hat{\beta}_{16}$	۰/۱۳۳۴	۱/۶۹۳۹	۱/۶۹۹۸	۰/۳۵۲۸	۲/۰۲۲۵	۲/۰۲۲۵	۱/۹۹۰۶	۰/۰۶۵
$\hat{\beta}_{17}$	۰/۱۵۶۸	۱/۶۶۵۸	۱/۶۲۶۱	۰/۳۸۸۸	۱/۹۷۹۶	۱/۹۷۹۶	۱/۹۰۷۹	۰/۰۸۶۷
$\hat{\beta}_{18}$	۰/۱۵۴۳	۱/۶۸۱۶	۱/۶۲۴۵	۰/۳۹۳۵	۱/۹۹۷	۱/۹۹۷	۱/۹۶۲۲	۰/۰۸۱۴
$\hat{\beta}_{19}$	۰/۱۷۶۳	۱/۶۴۷۴	۱/۷۵۷۳	۰/۲۵۶۷	۱/۹۹۹۴	۱/۹۹۹۴	۱/۹۳۱۵	۰/۱۰۰۹
$\hat{\beta}_{20}$	۰/۴۴۹۴	۱/۳۷۰۴	۱/۲۲۵۷	۰/۸۷۹۱	۱/۹۸۱۹	۱/۹۸۱۹	۱/۹۰۱۴	۰/۰۶۶۶
$\hat{\beta}_{21}$	۰/۲۱۳۴	۰/۴۱۴۷	۰/۰۰۰۴	۰	۰/۰۵۴۲	۰/۰۰۶۹	۰/۰۲۴۶	۰/۰۳۴۷
$\hat{\beta}_{22}$	۰/۰۷۶۲	۰/۲۰۷۵	۰/۰۰۱۲	۰/۰۰۰۱	۰/۰۵۳	۰/۰۰۰۳	۰/۰۰۰۶۷	۰/۰۳۳
$\hat{\beta}_{23}$	۰/۰۵۱۴	۰/۱۵۲۳	۰	۰	۰/۰۵۱۱	۰/۰۲۶۲	۰/۰۲۳۶	۰/۰۳۶۳
$\hat{\beta}_{24}$	۰/۰۶۶۳	۰/۱۷۸۸	۰/۰۰۳۸	۰/۰۰۱۵	۰/۰۵۴۲	۰/۰۴۰۴	۰/۰۱۲۸	۰/۰۳۷۹
$\hat{\beta}_{25}$	۰/۰۵۲	۰/۱۵۱۳	۰	۰	۰/۰۵۰۱	۰/۰۰۰۷	۰/۰۰۰۷۴	۰/۰۳۹
$\hat{\beta}_{26}$	۰/۰۵۵۵	۰/۱۳۴	۰	۰	۰/۰۴۹۴	۰/۰۲۰۸	۰/۰۲۶۷	۰/۰۴۰۳
$\hat{\beta}_{27}$	۰/۰۵۹۸	۰/۱۵۴۳	۰	۰	۰/۰۵۳۳	۰/۰۰۶۲	۰/۰۰۸	۰/۰۴۲
$\hat{\beta}_{28}$	۰/۰۵۹۴	۰/۱۶۷۳	۰/۰۰۰۴	۰	۰/۰۶۱۹	۰/۰۰۳۳۲	۰/۰۱۷۷	۰/۰۵۱۶
$\hat{\beta}_{29}$	۰/۰۶۶۹	۰/۱۷۹۴	۰/۰۰۲۸	۰/۰۰۰۴	۰/۰۵۲۳	۰/۰۲۰۹	۰/۰۳۸۸	۰/۰۴۱۳
$\hat{\beta}_{30}$	۰/۲۱۲۹	۰/۴۱۹۹	۰/۰۰۴۴	۰/۰۰۰۸	۰/۰۵۳۷	۰/۰۰۲۸۹	۰/۰۰۲۰۷	۰/۰۳۹۸
$\hat{\beta}_{31}$	۰/۳۸۴۵	۱/۴۱۸۸	۱/۳۷۳۲	۰/۶۴۴۲	۲/۰۲۶۸	۲/۰۲۶۸	۱/۹۵۲۲	۰/۱۰۴۹
$\hat{\beta}_{32}$	۰/۱۹۵۳	۱/۶۱۵۸	۱/۶۴۰۹	۰/۳۸۹۵	۲/۰۲۳۵	۲/۰۲۳۵	۱/۹۷۹۳	۰/۰۷۴۲
$\hat{\beta}_{33}$	۰/۱۶۴۶	۱/۶۴۸۹	۱/۶۵۳۵	۰/۳۵۸	۱/۹۹۲۶	۱/۹۹۲۶	۱/۹۴۵	۰/۰۶۸
$\hat{\beta}_{34}$	۰/۱۶۱۱	۱/۶۵۴۶	۱/۵۹۷۲	۰/۳۹۹۷	۱/۹۸۹۸	۱/۹۸۹۸	۱/۹۴۵	۰/۰۶۶۴
$\hat{\beta}_{35}$	۰/۱۵۴۹	۱/۶۶۳۵	۱/۶۳۵۷	۰/۳۲۶۴	۲/۰۲۵۵	۲/۰۲۵۵	۱/۹۶۷۳	۰/۰۹۰۶
$\hat{\beta}_{36}$	۰/۱۳۲۴	۱/۷۰۱۱	۱/۶۵۱۱	۰/۳۱۹۹	۱/۹۶۴۲	۱/۹۶۴۲	۱/۹۰۶۶	۰/۰۷۳۴
$\hat{\beta}_{37}$	۰/۱۴۶۳	۱/۶۷۷۳	۱/۶۸۹۴	۰/۳۲۶۹	۲/۰۳۷۸	۲/۰۳۷۸	۱/۹۸۵۸	۰/۰۷۹
$\hat{\beta}_{38}$	۰/۱۳۷	۱/۶۹۰۴	۱/۶۶۴۹	۰/۳۲۱۲	۱/۹۸۶۴	۱/۹۸۶۴	۱/۹۵۶۱	۰/۰۵۷۶
$\hat{\beta}_{39}$	۰/۱۲۴۶	۱/۷۰۰۷	۱/۶۸۳۶	۰/۳۲۱۵	۱/۹۹۵۸	۱/۹۹۵۸	۱/۹۶۷۱	۰/۰۶۶۷
$\hat{\beta}_{40}$	۰/۱۵۰۷	۱/۶۵۵۴	۳/۰۰۲۷	۱/۳۸۶۷	۲/۰۳۱۲	۲/۰۳۱۲	۲/۰۹۸۳	۰/۰۹۰۴
MSE	۵/۶۰۴۳	—	۹/۳۲۵۴	—	۲/۲۵۹۳	—	—	۲/۳۹۲۴
		MM - BR	MM	LASSO	Ridge			
		۱۷/۳۶۹	۱۵/۳۰۷۴	۳۸/۵۷۲۴	۳۷/۱۳۹۷	$\sqrt{MPE}$		
		۲/۱۳۸۳	۲/۴۲۶۳	۰/۹۶۲۹	۱	REff		
		۱۴۸/۵۷۸۴	۱۴۸/۵۷۸۴	۱۴۸/۵۷۸۴	۱۴۸/۵۷۸۴	MCN		
		۱۰/۳۷	۶/۹۲	۱۹/۸۹	—	C		
		۰	۰	۰/۲	—	IC		

جدول ۱۰.۴: نتایج شبیه‌سازی مربوط به بریج استوار، طرح ۲، $n = ۱۰۰$ ، $\rho = ۰/۹۹$ و $\tau = ۰$

MM – BR		MM		LASSO		Ridge		ضرایب
MSE	mean	MSE	mean	MSE	mean	MSE	mean	
۱/۴۱۹۸	–۰/۴۴۹۱	۲/۰۰۶۶	–۰/۱۵۹۲	۰/۰۱۴۱	۰/۰۱۴۸	۰/۷۵۵۷	۰/۸۶۷۵	$\hat{\beta}_1$
۱/۶۳۲۴	۰/۴۶۶۶	۲/۶۰۲۴	۰/۰۴۴۷	۰/۰۰۵۹	۰/۰۱۴۴	۰/۷۵۵۲	۰/۸۶۷۴	$\hat{\beta}_2$
۱/۵۷۴۵	۰/۳۰۲۱	۲/۵۹۸۹	–۰/۲۲۳۵	۰	۰	۰/۷۵۶۲	۰/۸۶۷۹	$\hat{\beta}_3$
۲/۲۶۹۴	۰/۲۸۸۶	۳/۴۰۴	۰/۲۵۸۱	۰/۰۰۵۳	۰/۰۰۷۳	۰/۷۵۸۹	۰/۸۶۹۳	$\hat{\beta}_4$
۱/۴۱۷۱	۰/۳۷۸۷	۲/۳۹۹۶	۰/۰۸۷۵	۰/۰۰۱۸	۰/۰۰۵۳	۰/۷۵۵۷	۰/۸۶۷۴	$\hat{\beta}_5$
۱/۶۶۶۲	۰/۱۶۳۹	۲/۸۰۲۹	–۰/۱۴	۰/۰۰۰۲	۰/۰۰۵۲	۰/۷۵۴۱	۰/۸۶۶۵	$\hat{\beta}_6$
۱/۲۴۹۹	۰/۲۰۵۲	۲/۲۹۹۱	–۰/۰۳۲۴	۰/۰۰۰۹	۰/۰۰۰۳	۰/۷۵۸۵	۰/۸۶۹	$\hat{\beta}_7$
۱/۴۳۱۴	۰/۳۰۲۴	۲/۶۸۱۵	۰/۰۰۰۷	۰	۰	۰/۷۶۲۸	۰/۸۷۱۵	$\hat{\beta}_8$
۱/۴۷۹۹	۰/۱۲۰۹	۲/۵۳۲۹	–۰/۲۶۸۷	۰/۰۰۱۲	۰/۰۰۰۶۱	۰/۷۶۷۲	۰/۸۷۳۸	$\hat{\beta}_9$
۱/۵۷۵	۰/۴۴۷۴	۲/۵۵۱۶	۰/۲۹۶۹	۰/۰۳۷۳	۰/۰۴۴۴	۰/۷۷۶۲	۰/۸۷۸۹	$\hat{\beta}_{10}$
۲/۰۴۶۶	۱/۵۳۸۶	۲/۱۸۵۸	۱/۸۴۴۹	۳/۱۷۰۶	۰/۴۱۴۳	۱/۲۳۲۴	۰/۸۹۲	$\hat{\beta}_{11}$
۲/۵۶۲۴	۱/۸۶۷۹	۳/۲۹۳۶	۲/۲۸۱۴	۲/۸۱۲۲	۰/۷۲۱۴	۱/۲۱۹۶	۰/۸۹۷۹	$\hat{\beta}_{12}$
۲/۲۹۷۴	۱/۵۶۴۹	۲/۶۱۰۸	۱/۷۱۵۹	۲/۶۶۴۹	۰/۶۸۲۱	۱/۲۲۰۵	۰/۸۹۷۶	$\hat{\beta}_{13}$
۲/۰۸۰۴	۱/۷۵۲	۲/۶۱۸	۲/۱۱۷۹	۲/۰۸۷۴	۱/۱۳۰۸	۱/۲۱۱۵	۰/۹۰۲۲	$\hat{\beta}_{14}$
۲/۰۸۰۷	۱/۴۷۸۶	۲/۱۸۷۶	۱/۸۲۱	۲/۳۶۳۱	۰/۸۸۰۵	۱/۲۱۴۶	۰/۹۰۰۵	$\hat{\beta}_{15}$
۲/۰۰۴۲	۱/۸۷۱۹	۲/۲۸۶۹	۲/۱۵۲۲	۲/۴۰۱۸	۰/۹۹۰۴	۱/۲۱۲۵	۰/۹۰۱۲	$\hat{\beta}_{16}$
۱/۹۰۰۹	۱/۵۶۴۷	۲/۳۶۰۸	۱/۸۶۲۳	۲/۶۱۰۷	۰/۸۲	۱/۲۱	۰/۹۰۲۵	$\hat{\beta}_{17}$
۲/۴۰۵۶	۱/۷۶۲	۳/۰۴۹۷	۱/۹۷۹۹	۲/۴۷۱۹	۰/۹۸۳۲	۱/۲۱۱۴	۰/۹۰۱۷	$\hat{\beta}_{18}$
۲/۹۲۶۲	۱/۷۹۰۶	۳/۵۱۸۶	۱/۹۹۵۷	۲/۱۱۱	۱/۰۳۰۲	۱/۲۱۶۵	۰/۸۹۹۴	$\hat{\beta}_{19}$
۱/۶۳۴۵	۱/۴۹۳۱	۱/۸۴۳۱	۱/۸۷۷۴	۳/۳۶۰۸	۰/۲۶۰۸	۱/۲۳۰۵	۰/۸۹۲۶	$\hat{\beta}_{20}$
۱/۲۶۸۱	۰/۳۳۹	۲/۴۷۷۸	۰/۰۴۶۸	۰/۰۲۶۴	۰/۰۲۸۵	۰/۷۸۴۱	۰/۸۸۳۴	$\hat{\beta}_{21}$
۱/۱۰۰۸	۰/۱۹۷۸	۲/۴۲۳۲	–۰/۰۲۰۲	۰	۰	۰/۷۷۸۱	۰/۸۸۰۱	$\hat{\beta}_{22}$
۱/۳۸۳۱	۰/۲۹۲۲	۲/۳۳۷۲	۰/۱۷۷۱	۰/۰۰۰۶	۰/۰۰۰۲۴	۰/۷۷۲۶	۰/۸۷۷۲	$\hat{\beta}_{23}$
۱/۴۴۲۳	۰/۱۷۹۴	۲/۴۷۸۳	–۰/۲۷۳۲	۰	۰	۰/۷۶۷۲	۰/۸۷۴	$\hat{\beta}_{24}$
۱/۵۹۹۳	۰/۲۳۶۱	۲/۲۹۱۸	۰/۰۰۰۵	۰	۰	۰/۷۶۶۲	۰/۸۷۳۶	$\hat{\beta}_{25}$
۱/۴۵۲۴	۰/۲۴۵۱	۲/۲۶۰۷	۰/۱۴۰۹	۰/۰۰۰۲	۰/۰۰۱۳	۰/۷۷۰۷	۰/۸۷۶۱	$\hat{\beta}_{26}$
۱/۴۹۵۵	۰/۲۱۴	۲/۴۳۸۱	۰/۰۴۲	۰	۰/۰۰۰۶	۰/۷۷۰۳	۰/۸۷۶	$\hat{\beta}_{27}$
۲/۱۹۵۴	۰/۰۷۳۹	۲/۸۳۲۶	–۰/۲۲۴۸	۰	۰	۰/۷۶۵۶	۰/۸۷۳۳	$\hat{\beta}_{28}$
۱/۷۴۸۳	۰/۴۲۲۷	۲/۳۸۹۳	۰/۱۴۱۲	۰/۰۰۲۱	۰/۰۰۴۶	۰/۷۷۰۹	۰/۸۷۶۳	$\hat{\beta}_{29}$
۱/۴۹۸۴	۰/۰۷۵۳	۲/۴۵۴۹	–۰/۱۹۵۲	۰/۰۰۲۲	۰/۰۰۰۶۶	۰/۷۷۵۵	۰/۸۷۸۷	$\hat{\beta}_{30}$
۲/۶۲۱۲	۱/۹۶۵۶	۳/۷۶۴۲	۲/۱۸۱۵	۲/۸۵۵	۰/۵۱۴۷	۱/۲۳۱۲	۰/۸۹۲۴	$\hat{\beta}_{31}$
۲/۲۵۸۴	۱/۸۴۱	۲/۴۸۶۱	۲/۱۵۸۷	۲/۵۵۰۲	۰/۷۸۶۸	۱/۲۲۲۵	۰/۸۹۶۲	$\hat{\beta}_{32}$
۲/۱۸۴۴	۱/۶۳۹۷	۲/۲۱۹۷	۱/۹۵۰۱	۲/۳۲۵۵	۰/۹۹۲۸	۱/۲۱۸۱	۰/۸۹۸۶	$\hat{\beta}_{33}$
۲/۳۱۶۹	۱/۶۰۲۷	۲/۶۷۶۱	۱/۹۳۱۳	۲/۶۵۶۴	۰/۷۵۱۲	۱/۲۱۷۸	۰/۸۹۸۸	$\hat{\beta}_{34}$
۲/۴۴۵۳	۱/۸۴۷۱	۲/۸۷۱۲	۲/۱۷۲۳	۲/۳۰۶	۰/۹۸۴	۱/۲۲۱۱	۰/۸۹۷۵	$\hat{\beta}_{35}$
۲/۳۷۷۸	۱/۵۲۱۱	۲/۷۷۷۸	۱/۷۵۷۸	۲/۳۷۹۱	۰/۹۴۵۵	۱/۲۳۱	۰/۸۹۲۹	$\hat{\beta}_{36}$
۲/۳۳۱۴	۱/۹۲۴۴	۲/۷۲۳	۲/۲۵۵۶	۲/۵۲۵۱	۰/۸۴۶۵	۱/۲۴۹۶	۰/۸۸۴۳	$\hat{\beta}_{37}$
۲/۱۴۶۷	۱/۶۱۱۴	۲/۰۹۰۹	۱/۹۰۸۱	۲/۵۰۱	۰/۶۹۱۴	۱/۲۸۵۷	۰/۸۶۸	$\hat{\beta}_{38}$
۱/۹۷۷۳	۱/۶۸۶۶	۲/۴۱۳۱	۱/۹۷۱۶	۲/۸۹۷۷	۰/۴۶۸	۱/۳۵۴۲	۰/۸۳۷۸	$\hat{\beta}_{39}$
۱۰/۱۸۵۹	۲/۹۴۲۳	۱۴/۴۷۲۴	۲/۳۳۰۹	۲۵۵/۰۰۸۷	۱۷/۵۸۲۴	۱/۴۷۰۸	۰/۷۸۸۴	$\hat{\beta}_{40}$
۸۳/۶۸۳۴	–	۱۱۴/۷۱۲۶	–	۳۰۴/۱۵۸۸	–	۴۰/۲۰۳۱	–	MSE
MM – BR		MM		LASSO		Ridge		
۱۵/۴۶۳		۱۵/۳۰۷۴		۵۹/۵۰۳۳		۵۶/۹۳۵۸		$\sqrt{MPE}$
۳/۶۸۲۱		۳/۷۱۹۵		۰/۹۵۶۹		۱		REF
۶۱۵۰/۴۱۰۲		۶۱۵۰/۴۱۰۲		۶۱۵۰/۴۱۰۲		۶۱۵۰/۴۱۰۲		MCN
۲/۸۲		۱/۱۱		۱۹/۶۹		–		C
۱/۰۲		۰/۴۸		۹/۱۴		–		IC

جدول ۱۱.۴: نتایج شبیه‌سازی مربوط به بریج استوار، طرح ۲، $n = 300$ ، $\rho = 0.3$ و $\tau = 0$

MM – BR		MM		LASSO		Ridge		ضرایب
MSE	mean	MSE	mean	MSE	mean	MSE	mean	
0/0039	-0/0124	0/0063	0/0045	0/0001	0	0/0228	0/1144	$\hat{\beta}_1$
0/0082	-0/0036	0/0101	-0/0101	0	-0/0005	0/0122	0/0738	$\hat{\beta}_2$
0/0092	0/0046	0/0098	0/0033	0/0002	0/0002	0/0143	0/0786	$\hat{\beta}_3$
0/0092	-0/0096	0/0103	-0/0144	0	0/0003	0/0115	0/0558	$\hat{\beta}_4$
0/0006	0/0116	0/0071	0/0108	0/0001	0/0013	0/0132	0/0723	$\hat{\beta}_5$
0/0071	0/0036	0/0083	0/0054	0/0001	0/0014	0/0137	0/0712	$\hat{\beta}_6$
0/0075	0/0074	0/0092	0/0063	0/0001	0/0015	0/0137	0/072	$\hat{\beta}_7$
0/0095	0/0103	0/011	0/011	0	0/0007	0/0135	0/0712	$\hat{\beta}_8$
0/0125	-0/0122	0/0138	-0/0167	0/0002	0/0001	0/0217	0/1071	$\hat{\beta}_9$
0/0093	-0/0006	0/0108	-0/0049	0/0002	0/0023	0/0985	0/2818	$\hat{\beta}_{10}$
0/0377	1/9398	0/0094	2/0004	0/1744	1/6602	0/1661	1/6345	$\hat{\beta}_{11}$
0/03	1/9573	0/0099	2/0013	0/0512	1/8533	0/047	1/8309	$\hat{\beta}_{12}$
0/0224	1/9528	0/0075	1/9937	0/0747	1/8136	0/0428	1/8363	$\hat{\beta}_{13}$
0/0203	1/9616	0/0081	1/9984	0/0446	1/8606	0/0335	1/8648	$\hat{\beta}_{14}$
0/0154	1/9852	0/008	2/0125	0/0521	1/856	0/031	1/8658	$\hat{\beta}_{15}$
0/0252	1/9403	0/0095	1/9787	0/0628	1/8244	0/0369	1/8444	$\hat{\beta}_{16}$
0/0241	1/9678	0/0098	2/0005	0/0635	1/8333	0/0369	1/8555	$\hat{\beta}_{17}$
0/0256	1/9573	0/0141	1/9997	0/0476	1/8533	0/0311	1/8654	$\hat{\beta}_{18}$
0/016	1/9831	0/0084	2/0124	0/0454	1/8557	0/0428	1/8366	$\hat{\beta}_{19}$
0/0399	1/9385	0/0081	2/0058	0/1784	1/6709	0/1588	1/6414	$\hat{\beta}_{20}$
0/0079	0/0099	0/0093	0/0049	0/0003	0/0035	0/0978	0/2855	$\hat{\beta}_{21}$
0/0067	-0/0065	0/0077	-0/0032	0	0/0001	0/022	0/1073	$\hat{\beta}_{22}$
0/0086	0/0129	0/0097	0/0047	0	0/0007	0/0149	0/0796	$\hat{\beta}_{23}$
0/0081	-0/0093	0/0085	-0/0078	0/0001	-0/0015	0/0167	0/0804	$\hat{\beta}_{24}$
0/0076	-0/0052	0/0089	-0/0026	0/0001	0/0008	0/0136	0/0632	$\hat{\beta}_{25}$
0/0082	-0/0014	0/0101	-0/0067	0	0/0003	0/0109	0/0661	$\hat{\beta}_{26}$
0/0077	0/0004	0/0086	0/0006	0	0/0002	0/0139	0/0706	$\hat{\beta}_{27}$
0/0077	0/0014	0/009	0/0032	0/0001	-0/0004	0/0181	0/0851	$\hat{\beta}_{28}$
0/0074	-0/0101	0/0086	-0/0185	0	0/0002	0/0207	0/0946	$\hat{\beta}_{29}$
0/0088	0/0006	0/0103	0/0137	0/0003	0/0034	0/1003	0/2893	$\hat{\beta}_{30}$
0/0368	1/9452	0/0089	1/9979	0/1545	1/6902	0/152	1/6478	$\hat{\beta}_{31}$
0/0276	1/9585	0/0085	2/0046	0/0464	1/8562	0/0481	1/8298	$\hat{\beta}_{32}$
0/019	1/9422	0/0113	1/9746	0/066	1/8266	0/0382	1/8474	$\hat{\beta}_{33}$
0/0321	1/9819	0/0109	2/0252	0/0592	1/8603	0/0329	1/8684	$\hat{\beta}_{34}$
0/0232	1/9456	0/0085	1/9815	0/0755	1/7979	0/041	1/8335	$\hat{\beta}_{35}$
0/0147	1/9959	0/0094	2/0124	0/0452	1/8584	0/0336	1/8631	$\hat{\beta}_{36}$
0/0209	1/9514	0/0084	1/9863	0/0654	1/8331	0/0365	1/8542	$\hat{\beta}_{37}$
0/0204	1/9683	0/01	2/0022	0/0484	1/8519	0/0308	1/8599	$\hat{\beta}_{38}$
0/0136	1/9887	0/0084	1/9932	0/0549	1/8395	0/0326	1/8553	$\hat{\beta}_{39}$
0/0431	2/0738	0/0098	2/0154	0/3465	2/4788	0/0451	1/8117	$\hat{\beta}_{40}$
0/6686	-	0/3746	-	1/7585	-	1/6815	-	MSE
				MM – BR	MM	LASSO	Ridge	
				23/348	19/1383	32/7698	34/399	$\sqrt{\text{MPE}}$
				1/4733	1/7974	1/0497	1	REff
				64/9991	64/9991	64/9991	64/9991	MCN
				15/02	14/1	19/95	-	C
				0	0	0	-	IC

جدول ۱۲.۴: نتایج شبیه‌سازی مربوط به بریج استوار، طرح ۲، $n = 300$ ، $\rho = 0.99$ و $\tau = 0$

MM – BR		MM		LASSO		Ridge		ضرایب
MSE	mean	MSE	mean	MSE	mean	MSE	mean	
0.1689	-0.0969	0.2885	0.0302	0	0	0.8368	0.9143	$\hat{\beta}_1$
0.2679	0.2615	0.4638	-0.068	0	0.0001	0.8279	0.9094	$\hat{\beta}_2$
0.3279	0.0685	0.4496	0.0223	0	0	0.8263	0.9085	$\hat{\beta}_3$
0.3435	0.1017	0.47	-0.0974	0	0	0.8263	0.9085	$\hat{\beta}_4$
0.2145	0.1011	0.324	0.0733	0	0	0.8282	0.9095	$\hat{\beta}_5$
0.2575	0.1201	0.3778	0.0368	0	0	0.8323	0.9118	$\hat{\beta}_6$
0.3006	0.1113	0.4208	0.0426	0	0	0.8418	0.917	$\hat{\beta}_7$
0.3635	0.2198	0.5046	0.0742	0	0	0.8461	0.9192	$\hat{\beta}_8$
0.4586	0.04	0.633	-0.1127	0	0	0.8564	0.9247	$\hat{\beta}_9$
0.3384	0.082	0.4957	-0.0332	0	0.0008	0.8729	0.9334	$\hat{\beta}_{10}$
0.5147	1.8448	0.4286	2.0026	2.4231	0.5739	1.0963	0.9543	$\hat{\beta}_{11}$
0.5931	1.8679	0.451	2.0091	1.1626	1.2475	1.0769	0.9637	$\hat{\beta}_{12}$
0.4587	1.8126	0.3435	1.9573	1.356	1.1352	1.0693	0.9676	$\hat{\beta}_{13}$
0.4616	1.8469	0.3687	1.9894	1.1841	1.3332	1.065	0.9698	$\hat{\beta}_{14}$
0.4607	1.9365	0.365	2.0847	1.0779	1.4437	1.0643	0.9701	$\hat{\beta}_{15}$
0.6126	1.712	0.4357	1.8562	1.5194	1.1368	1.0656	0.9695	$\hat{\beta}_{16}$
0.5827	1.8937	0.4467	2.0341	1.3282	1.2944	1.0641	0.9701	$\hat{\beta}_{17}$
0.7746	1.8633	0.6426	1.9977	1.218	1.3916	1.0663	0.9691	$\hat{\beta}_{18}$
0.4904	1.9521	0.382	2.0841	1.2865	1.3085	1.0784	0.963	$\hat{\beta}_{19}$
0.4434	1.8862	0.3684	2.0392	2.7094	0.4606	1.1008	0.9519	$\hat{\beta}_{20}$
0.2891	0.1485	0.4258	0.0329	0.0001	0.0009	0.8707	0.9325	$\hat{\beta}_{21}$
0.2325	0.0803	0.3501	-0.0214	0	0	0.8506	0.9217	$\hat{\beta}_{22}$
0.2881	0.1702	0.4453	0.0318	0	0	0.8419	0.9171	$\hat{\beta}_{23}$
0.2455	0.1099	0.3899	-0.0528	0	0	0.8343	0.9129	$\hat{\beta}_{24}$
0.2777	0.0719	0.4091	-0.0179	0	0	0.8326	0.912	$\hat{\beta}_{25}$
0.3052	0.0943	0.4627	-0.0455	0	0	0.8347	0.9132	$\hat{\beta}_{26}$
0.2689	0.0486	0.3944	0.0043	0	0	0.8367	0.9143	$\hat{\beta}_{27}$
0.2687	0.1634	0.4135	0.0218	0	0	0.8389	0.9154	$\hat{\beta}_{28}$
0.2895	0.0028	0.3955	-0.1251	0	0.0001	0.846	0.9193	$\hat{\beta}_{29}$
0.337	0.1243	0.4721	0.0926	0.0001	0.0009	0.8645	0.9292	$\hat{\beta}_{30}$
0.5789	1.8368	0.4076	1.9859	2.4867	0.5547	1.1071	0.9488	$\hat{\beta}_{31}$
0.5272	1.8983	0.3876	2.0314	1.2862	1.3295	1.087	0.959	$\hat{\beta}_{32}$
0.6822	1.6881	0.5156	1.8283	1.7609	1.1484	1.082	0.9616	$\hat{\beta}_{33}$
0.608	2.0516	0.498	2.1705	1.3237	1.418	1.0798	0.9627	$\hat{\beta}_{34}$
0.5227	1.7222	0.3906	1.8752	1.621	1.0662	1.0875	0.9587	$\hat{\beta}_{35}$
0.5637	1.9589	0.4313	2.0841	1.1144	1.3728	1.0966	0.9543	$\hat{\beta}_{36}$
0.539	1.7797	0.3859	1.9074	1.396	1.2198	1.1178	0.9441	$\hat{\beta}_{37}$
0.6278	1.8866	0.4594	2.0147	1.5867	1.0323	1.1563	0.9259	$\hat{\beta}_{38}$
0.5204	1.854	0.3843	1.9541	2.7099	0.4417	1.2308	0.8914	$\hat{\beta}_{39}$
3.2085	2.2926	2.8008	2.1887	164.8995	14.2594	1.3625	0.8332	$\hat{\beta}_{40}$
19.6142	—	19.4797	—	195.4482	—	39.0003	—	MSE
MM – BR		MM		LASSO		Ridge		$\sqrt{MPE}$
21.5906	19.1383	62.0661	61.976					REff
2.8705	3.2383	0.9985	1					MCN
4146.2958	4146.2958	4146.2958	4146.2958					C
4.68	2.33	20	—					IC
0.03	0.01	3.88	—					

جدول ۱۳.۴: نتایج شبیه‌سازی مربوط به بریج استوار، طرح ۲، $n = ۱۰۰$ ، $\rho = ۰/۹$ و $\tau = ۰/۲$

ضرایب	Ridge		LASSO		MM		MM – BR	
	mean	MSE	mean	MSE	mean	MSE	mean	MSE
$\hat{\beta}_1$	۰/۳۸۸۱	۰/۱۵۱۴	۰/۳۴۸۸	۱/۳۶۴۴	۰/۰۷۱۸	۰/۲۰۲۴	۰/۱۴۲۹	۰/۱۴۱
$\hat{\beta}_2$	۰/۴۷۳۸	۰/۲۲۵۵	۰/۲۴۰۳	۰/۶۲۴۱	۰/۰۰۷۶	۰/۲۰۸۱	۰/۱۰۹۶	۰/۱۴۴۷
$\hat{\beta}_3$	۰/۳۹۴۴	۰/۱۵۶۳	۰/۲۶۳۵	۰/۵۹۷۴	۰/۰۵۰۷	۰/۲۶۶۱	۰/۱۰۸۶	۰/۱۸۸۵
$\hat{\beta}_4$	۰/۴۴۴۵	۰/۱۹۸۵	۰/۳۵۲۵	۱/۳۰۲۱	۰/۰۶۹۴	۰/۲۷۳۲	۰/۰۳۸۳	۰/۲۳
$\hat{\beta}_5$	۰/۶۶۶۱	۰/۴۴۵۸	۰/۰۱۸	۰/۰۳۲۳	۰/۰۶۶۶	۰/۲۶۷۲	۰/۰۸۲۷	۰/۲۰۱۴
$\hat{\beta}_6$	۰/۸۶۴	۰/۷۴۹۶	۰	۰	۰/۲۲۴۴	۰/۳۱۸۳	۰/۲۲۱۳	۰/۲۵۴۱
$\hat{\beta}_7$	۰/۴۲۴۷	۰/۱۸۱۲	۰/۰۶۵	۲/۰۵۶۹	۰/۱۵۴۶	۰/۳۱۰۴	۰/۰۰۹۴	۰/۲۰۴۶
$\hat{\beta}_8$	۰/۶۷۵	۰/۴۵۷۹	۰/۰۱۵۴	۰/۰۲۳۶	۰/۱۴۹۶	۰/۳۰۸۲	۰/۱۱۳۵	۰/۱۹۵۱
$\hat{\beta}_9$	۰/۴۲۴۷	۰/۱۸۱۳	۰/۲۸۵۹	۰/۶۴۲۹	۰/۱۰۱۱	۰/۲۰۰۹	۰/۰۱۱۵	۰/۱۳۲۱
$\hat{\beta}_{10}$	۰/۸۵۰۵	۰/۷۲۷۲	۰	۰	۰/۱۷۸۷	۰/۲۵۱۳	۰/۱۵۶۳	۰/۱۸۴۸
$\hat{\beta}_{11}$	۰/۴۲۲۷	۲/۴۸۸۹	۰/۰۱۲	۳/۲۰۵	۱/۸۶۰۴	۰/۲۴۳۵	۱/۷۸۷۵	۰/۳۱۰۳
$\hat{\beta}_{12}$	۰/۵۶۳۵	۲/۰۶۵	۰/۴۹۸۵	۴/۴۸۰۱	۲/۰۰۱۴	۰/۲۷۲۱	۱/۹۲۶۹	۰/۳۴۳
$\hat{\beta}_{13}$	۰/۵۵۰۱	۲/۱۰۳۸	۰/۴۴۵۲	۳/۷۶۵۴	۲/۰۲۵۶	۰/۲۳۸۷	۱/۹۹۴۲	۰/۲۵۹۴
$\hat{\beta}_{14}$	۰/۴۳۶۵	۲/۴۴۵۵	۰/۷۲۳۱	۴/۴۶۰۸	۱/۸۶۷۶	۰/۲۰۹۴	۱/۸۱۵۳	۰/۲۶۰۸
$\hat{\beta}_{15}$	۰/۸۱۳۶	۱/۴۱۱۱	۰	۴	۲/۰۹۳۹	۰/۱۶۸۸	۲/۰۴۵۶	۰/۱۷۸۲
$\hat{\beta}_{16}$	۰/۷۳۲۸	۱/۶۰۸۲	۰	۴	۲/۰۰۴۶	۰/۲۷۵۲	۱/۹۴۶۲	۰/۳۰۸۸
$\hat{\beta}_{17}$	۰/۴۵۲۳	۲/۳۹۶۳	۱/۰۵۳۹	۴/۴۶۸۸	۱/۸۶۸۵	۰/۲۳۱۱	۱/۸۳۱۲	۰/۲۸۳۷
$\hat{\beta}_{18}$	۰/۸۹۵۱	۱/۲۲۴۱	۰	۴	۲/۰۸۶۱	۰/۲۵۲۱	۲/۰۱۹۱	۰/۲۸۶۸
$\hat{\beta}_{19}$	۰/۷۵۰۷	۱/۵۶۳۴	۰	۴	۲/۰۰۲۷	۰/۲۶۶	۱/۹۳۳	۰/۳۳۱۶
$\hat{\beta}_{20}$	۰/۸۵۷۹	۱/۳۰۷۲	۰	۴	۱/۹۳۳	۰/۲۶۰۶	۱/۸۵۰۴	۰/۳۲۳۲
$\hat{\beta}_{21}$	۰/۸۷۵۴	۰/۷۷۲۸	۰	۰	۰/۱۸۰۱	۰/۲۵۹۵	۰/۱۶۷۱	۰/۱۸۹۲
$\hat{\beta}_{22}$	۰/۷۲۵۵	۰/۵۲۹	۰	۰	۰/۰۳۳۶	۰/۲۵۲۳	۰/۱۰۴۷	۰/۱۹۴
$\hat{\beta}_{23}$	۰/۸۹۰۸	۰/۷۹۷	۰	۰	۰/۱۰۹	۰/۳۰۹۴	۰/۱۲۸۱	۰/۲۲۹۳
$\hat{\beta}_{24}$	۰/۸۸۸۸	۰/۷۹۴۸	۰	۰	۰/۰۷۲۵	۰/۲۸۱۵	۰/۱۰۳۹	۰/۲۰۶۷
$\hat{\beta}_{25}$	۰/۶۳۸	۰/۴۰۸۹	۰	۰	۰/۰۱۹۵	۰/۳۱۱	۰/۰۷۱۴	۰/۲۵۴۸
$\hat{\beta}_{26}$	۰/۵۵۵۱	۰/۳۰۹۷	۰/۱۴۳۷	۰/۶۷۱۱	۰/۰۳۲۸	۰/۲۴۴	۰/۰۸۹۱	۰/۱۷۰۹
$\hat{\beta}_{27}$	۰/۳۷۱۵	۰/۱۳۸۶	۰/۳۳۲۱	۰/۸۹۴۶	۰/۰۹۵۱	۰/۲۰۷۴	۰/۰۱۵	۰/۱۵۰۷
$\hat{\beta}_{28}$	۰/۸۵۳	۰/۷۳۰۶	۰	۰	۰/۱۵۶۲	۰/۲۸۳	۰/۱۶۴۶	۰/۲۲۵۸
$\hat{\beta}_{29}$	۰/۶۶۰۲	۰/۴۳۷۸	۰/۰۰۹۹	۰/۰۰۹۸	۰/۰۸۴۹	۰/۳۰۶۵	۰/۱۰۳۴	۰/۲۳۳۹
$\hat{\beta}_{30}$	۰/۵۵۵۵	۰/۳۱۰۱	۰/۱۷۸۹	۰/۵۰۷۳	۰/۰۷۹۸	۰/۲۹۷۹	۰/۰۷۸۲	۰/۲۳۹۵
$\hat{\beta}_{31}$	۰/۳۵۵۹	۲/۷۰۳۸	۰/۱۹۳۴	۳/۵۱۸۵	۱/۸۰۰۷	۰/۳۱۷۶	۱/۷۴۴۶	۰/۳۶۹۴
$\hat{\beta}_{32}$	۰/۴۹۲۴	۲/۲۷۴	۱/۳۶	۴/۸۱۰۷	۱/۹۴۶۳	۰/۲۵۱۵	۱/۸۹۸۹	۰/۳۲۲۹
$\hat{\beta}_{33}$	۰/۵۷۲۸	۲/۰۳۸۴	۰/۴۷۷۲	۴/۰۹۰۹	۲/۰۵۶۱	۰/۲۵۶۴	۲/۰۱۶۱	۰/۳۰۳۸
$\hat{\beta}_{34}$	۰/۴۴۱۹	۲/۴۲۸۶	۱/۲۶۱۹	۴/۵۴۹۴	۱/۸۳۴۱	۰/۳۰۶	۱/۸۰۸۳	۰/۳۲۶۷
$\hat{\beta}_{35}$	۰/۵۲۹۴	۲/۱۷۸۶	۰	۴	۲/۱۷۷۲	۰/۲۶۰۴	۲/۱۰۶۵	۰/۲۸۸۸
$\hat{\beta}_{36}$	۰/۵۶۴۹	۲/۰۶۱	۰/۳۲۳۴	۳/۸۰۷۱	۱/۹۳۳۸	۰/۳۰۶۷	۱/۸۹۲۹	۰/۳۸۳۷
$\hat{\beta}_{37}$	۰/۴۵۲۸	۲/۳۹۴۹	۱/۱۲۴۸	۳/۵۶۷۱	۲/۰۰۱۴	۰/۲۷۳۱	۱/۹۶۴۴	۰/۳۲۱۲
$\hat{\beta}_{38}$	۰/۶۲۶۵	۱/۸۸۸۱	۰/۰۵۹۵	۳/۹۶۷۵	۱/۹۷۱۲	۰/۳۱۰۳	۱/۹۳۵۵	۰/۳۵۱۴
$\hat{\beta}_{39}$	۰/۵۶۷۴	۲/۰۵۳۵	۰/۱۸۲۹	۴/۱۰۸۱	۲/۰۵۱۲	۰/۲۳۴	۲/۰۴۴۹	۰/۲۶۵۹
$\hat{\beta}_{40}$	۰/۴۶۳۱	۲/۳۶۲۸	۴/۹۱۰۱	۱۵/۹۲۷۲	۱/۶۶۱	۱/۰۷۴	۱/۷۱۴۲	۱/۱۰۱۵
MSE	—	۴۹/۷۰۱۲	—	۱۰۱/۴۵۲۹	—	۱۱/۳۹۹۵	—	۱۰/۸۹۲۴
	Ridge		LASSO		MM		MM – BR	
$\sqrt{MPE}$	۴۸۲/۰۰۹۶		۴۹۷/۵۹۵۹		۴۱/۱۵۲۸		۴۰/۱۳۸۶	
REff	۱		۰/۹۶۸۷		۱۱/۷۱۲۷		۱۲/۰۰۸۶	
MCN	۳۵۸۶/۴۸۶۵		۳۵۸۶/۴۸۶۵		۳۵۸۶/۴۸۶۵		۳۵۸۶/۴۸۶۵	
C	—		۱۸/۶۴		۳/۳۵		۵/۴۷	
IC	—		۱۵/۷۳		۰/۰۳		۰/۰۴	

جدول ۱۴.۴: نتایج شبیه‌سازی مربوط به بریج استوار، طرح ۲، $n = ۱۰۰$ ، $\rho = ۰/۹۹$ و $\tau = ۰/۲$

MM – BR		MM		LASSO		Ridge		ضرایب
MSE	mean	MSE	mean	MSE	mean	MSE	mean	
۰/۹۰۴۴	۰/۰۷۴۲	۱/۲۰۸۵	۰/۲۲۹۷	۰/۳۰۸۷	۰/۱۲۶۵	۰/۱۴۸۴	۰/۳۸۴	$\hat{\beta}_1$
۰/۶۶۰۶	۰/۳۷۵۴	۱/۰۴۸۷	۰/۱۱۲۷	۴/۳۵۱	۰/۹۸۲۱	۰/۲۲۱۸	۰/۴۶۹۷	$\hat{\beta}_2$
۰/۷۵۲۵	۰/۵۰۲۴	۱/۰۸۲۳	۰/۲۶۴۳	۰/۵۰۸۸	۰/۱۸۶۳	۰/۱۵۳۵	۰/۳۹۰۶	$\hat{\beta}_3$
۰/۸۵۰۲	۰/۱۹۷۹	۱/۲۹۹	۰/۱۱۰۲	۲/۲۱۱۱	۰/۵۸۴۶	۰/۱۹۵۵	۰/۴۴۰۹	$\hat{\beta}_4$
۰/۷۶۱	۰/۴۱۷۱	۱/۱۷۸۸	۰/۲۱۳۴	۰/۰۳۷۵	۰/۰۳۰۱	۰/۴۴۳۷	۰/۶۶۴۵	$\hat{\beta}_5$
۰/۸۸۶۷	۰/۳۹۱	۱/۲۸۳۶	۰/۲۷۷۶	۰	۰	۰/۷۶۰۴	۰/۸۷۰۴	$\hat{\beta}_6$
۰/۷۴۷۸	۰/۲۴۲۷	۱/۱۴۵۷	۰/۰۰۸۸	۳/۱۱۷۲	۰/۶۷۹۶	۰/۱۷۸۵	۰/۴۲۱۳	$\hat{\beta}_7$
۰/۶۰۴۵	۰/۳۰۸۶	۱/۰۶۸	۰/۱۷۷۵	۰/۰۸۴۲	۰/۰۳۴۷	۰/۴۵۸۲	۰/۶۷۵۲	$\hat{\beta}_8$
۰/۴۵۹۱	۰/۱۵۴۵	۰/۷۳۸	–۰/۱۰۲۵	۰/۹۲۴۲	۰/۳۲۳۵	۰/۱۷۸۴	۰/۴۲۱۲	$\hat{\beta}_9$
۱/۰۵۶۴	۰/۶۵۰۷	۱/۷۵۳۱	۰/۶۸۹۴	۰	۰	۰/۷۳۰۱	۰/۸۵۲۶	$\hat{\beta}_{10}$
۱/۳۶۳۳	۱/۴۲۶۱	۱/۴۸۵۱	۱/۵۹۴۷	۳/۳۹۷۸	۰/۳۴۸۶	۲/۵۰۴۹	۰/۴۱۷۶	$\hat{\beta}_{11}$
۱/۵۱۱۳	۱/۶۱۷۳	۱/۵۳۵۶	۱/۷۸۶۶	۶/۹۷۴۸	۱/۱۶۰۱	۲/۰۸۱۴	۰/۵۵۷۹	$\hat{\beta}_{12}$
۱/۱۲۵۸	۱/۶۷۲۷	۱/۲۲۴۷	۱/۸۴	۶/۵۲۶۲	۱/۰۷۸۹	۲/۱۲۲۵	۰/۵۴۳۷	$\hat{\beta}_{13}$
۱/۲۱۶۴	۱/۵۵۳	۱/۱۴۵۶	۱/۷۴۹۸	۴/۲۵۶۵	۰/۴۳۲۸	۲/۴۶۲۹	۰/۴۳۱	$\hat{\beta}_{14}$
۰/۹۸۲۲	۱/۹۰۳۹	۱/۱۵۵۹	۲/۱۲۷۹	۴	۰	۱/۴۱۷۷	۰/۸۱۰۷	$\hat{\beta}_{15}$
۱/۱۲۸۳	۱/۷۳۸۷	۱/۵۰۶۸	۱/۸۸۴۷	۴	۰	۱/۶۲۰۵	۰/۷۲۸	$\hat{\beta}_{16}$
۱/۱۰۶۵	۱/۵۹۸۵	۱/۱۶۴۷	۱/۷۶۳	۴/۵۹۴	۰/۸۷۲۳	۲/۴۱۲۱	۰/۴۴۷۳	$\hat{\beta}_{17}$
۱/۰۶۳۷	۱/۷۵۵۷	۱/۲۲۳۳	۱/۹۷۱۷	۴	۰	۱/۲۱۷۶	۰/۸۹۸	$\hat{\beta}_{18}$
۱/۵۶۱۱	۱/۷۱۹۶	۱/۷۸۴۴	۱/۹۲۴۸	۴	۰	۱/۵۷۲	۰/۷۴۷۳	$\hat{\beta}_{19}$
۱/۳۹۶۴	۱/۴۷۸۶	۱/۳۹۵۳	۱/۶۷۲۷	۴	۰	۱/۳۰۶۷	۰/۸۵۸۱	$\hat{\beta}_{20}$
۰/۹۶۸	۰/۲۶۸۱	۱/۳۴۰۲	۰/۲۶۶۷	۰	۰	۰/۸۰۲	۰/۸۹۲۳	$\hat{\beta}_{21}$
۰/۷۲۸۴	۰/۳۴۴۵	۱/۱۲۳۷	۰/۱۵۹	۰	۰	۰/۵۲۶۵	۰/۷۲۳۸	$\hat{\beta}_{22}$
۰/۶۸۷۴	۰/۳۰۴۶	۱/۱۵۲	۰/۱۸۹۱	۰	۰	۰/۸۲۵۸	۰/۹۰۶۹	$\hat{\beta}_{23}$
۰/۷۰۳	۰/۲۷۸۲	۱/۲۹۷۸	۰/۱۵۲۷	۰	۰	۰/۸۲۴۵	۰/۹۰۵۷	$\hat{\beta}_{24}$
۰/۹۹۵۹	۰/۳۹۲۵	۱/۳۳۱۸	۰/۲۷۴۲	۰/۱۲۴۲	۰/۰۷	۰/۴۰۶۵	۰/۶۳۶۱	$\hat{\beta}_{25}$
۱/۱۸۲۹	۰/۲۵۴۳	۱/۵۰۲۵	۰/۰۵۸۳	۳/۱۸۲۴	۰/۶۸۶۳	۰/۳۰۸	۰/۵۵۳۳	$\hat{\beta}_{26}$
۰/۶۴۷۹	۰/۲۷۱۹	۱/۰۴۸۶	۰/۱۲۲	۰/۰۲۶	۰/۰۳۵	۰/۱۳۶۴	۰/۳۶۸۳	$\hat{\beta}_{27}$
۰/۷۵۸۱	۰/۳۵۵۸	۱/۲۸۰۵	۰/۲۵۱	۰	۰	۰/۷۴۴۷	۰/۸۶۱۴	$\hat{\beta}_{28}$
۰/۸۰۷۵	۰/۴۶	۱/۲۷۷۲	۰/۲۰۷۴	۰/۰۰۰۱	۰/۰۰۱۱	۰/۴۳۶۲	۰/۶۵۸۹	$\hat{\beta}_{29}$
۰/۸۶۲۲	۰/۴۹۱۹	۱/۴۲۵۶	۰/۳۵۹۶	۳/۴۵۱۶	۰/۷۴۴۸	۰/۳۰۶۷	۰/۵۵۲۲	$\hat{\beta}_{30}$
۲/۶۹۱۸	۱/۶۱۷۵	۱/۲۳۶۶	۱/۶۴۹۲	۳/۹۹۹۵	۰/۰۰۰۱	۲/۷۱۷۷	۰/۳۵۱۷	$\hat{\beta}_{31}$
۱/۱۸۷۲	۱/۶۶۸۷	۱/۱۸۲۵	۱/۸۸۴۵	۸/۸۸۳۵	۱/۸۳۵۵	۲/۲۹۳۱	۰/۴۸۶۲	$\hat{\beta}_{32}$
۱/۰۴۰۱	۱/۷۲۱۷	۰/۹۵	۱/۹۲۵۳	۵/۲۵۹۱	۰/۸۰۷۲	۲/۰۵۹۳	۰/۵۶۵۵	$\hat{\beta}_{33}$
۱/۰۹۱۳	۱/۷۴۰۴	۱/۰۹۰۸	۱/۹۰۶	۴/۱۲۷۷	۰/۹۱۱۴	۲/۴۴۶۸	۰/۴۳۶۱	$\hat{\beta}_{34}$
۱/۱۷۰۸	۱/۵۵۵۹	۱/۱۶۰۱	۱/۸۸۸۸	۴	۰	۲/۱۲۷۶	۰/۵۴۶۴	$\hat{\beta}_{35}$
۱/۰۰۴۴	۱/۷۴۱۵	۱/۰۹۸۷	۱/۹۰۹۳	۴/۷۲۹۹	۰/۷۷۵	۲/۰۷۴۷	۰/۵۶۰۱	$\hat{\beta}_{36}$
۱/۱۸۶	۱/۷۹۲۸	۱/۳۰۹۲	۱/۸۶۴۱	۳/۲۴۲	۰/۷۶۸۴	۲/۴۰۹۴	۰/۴۴۸۱	$\hat{\beta}_{37}$
۱/۲۰۲۹	۱/۶۰۶۳	۱/۰۳۰۸	۱/۷۸۵۵	۴/۳۱۸۸	۰/۱۵۹	۱/۸۸۲	۰/۶۲۸۸	$\hat{\beta}_{38}$
۱/۱۴۹۸	۱/۷۰۵۳	۱/۰۹۰۱	۱/۸۳۲۶	۳/۹۸۴۳	۰/۳۱۳۵	۲/۰۳۶۳	۰/۵۷۳۵	$\hat{\beta}_{39}$
۲/۳۱۰۶	۱/۳۴۳۸	۳/۵۰۳۱	۱/۱۶۲۵	۱۰/۵۲۶۳	۳/۲۹۸۱	۲/۳۴۷۳	۰/۴۶۸۳	$\hat{\beta}_{40}$
۴۲/۴۴۷	–	۵۱/۸۵۹۳	–	۱۱۷/۱۴۷۴	–	۴۹/۸۹۸۲	–	MSE
MM – BR		MM		LASSO		Ridge		
۴۱/۵۳۷۱		۴۴/۷۴۷۲		۴۹۴/۷۴۷۳		۴۸۶/۰۸۲۵		$\sqrt{\text{MPE}}$
۱۱/۷۰۲۴		۱۰/۸۶۲۹		۰/۹۸۲۵		۱		REF
۳۳۴۳۷/۰۸۲۷		۳۳۴۳۷/۰۸۲۷		۳۳۴۳۷/۰۸۲۷		۳۳۴۳۷/۰۸۲۷		MCN
۳/۵۷		۱/۵۶		۱۸/۱۹		–		C
۰/۸۷		۰/۳۴		۱۶/۱		–		IC

جدول ۱۵.۴: نتایج شبیه‌سازی مربوط به بریج استوار، طرح ۲، $n = 300$ ، $\rho = 0.9$ و $\tau = 0.2$

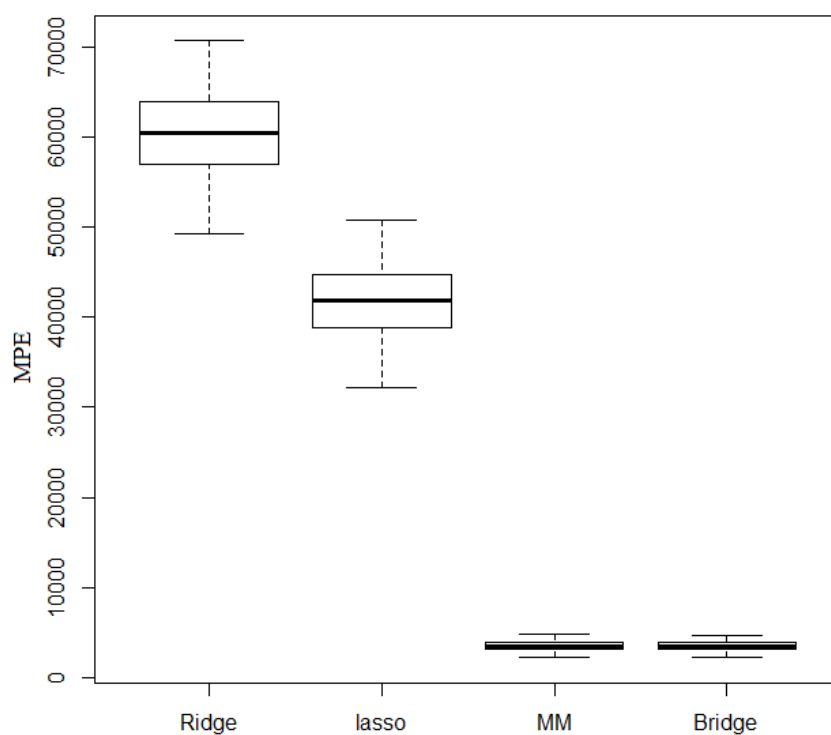
MM – BR		MM		LASSO		Ridge		ضرایب
MSE	mean	MSE	mean	MSE	mean	MSE	mean	
0/0۲۳۱	−0/0۹۲۶	0/0۲۷	−0/0۷۸۷	0/۲۸۲۷	0/۱۹۴۹	0/۱۵۵۷	0/۳۹۴۱	$\hat{\beta}_1$
0/0۴۲۱	0/0۸۷۶	0/0۴۵	0/0۸۱۷	0/۱۴۶	0/0۸۹۵	0/۲۳۲۳	0/۴۸۱۵	$\hat{\beta}_2$
0/0۳۶۱	−0/0۹۵۴	0/0۳۸۶	−0/۱0۵۶	0/۳0۹۷	0/۲۱۲۴	0/۱۶۳۶	0/۴0۰۴	$\hat{\beta}_3$
0/0۳۲۷	−0/0۱۷۹	0/0۳۹۱	−0/0۳۴۵	0/۵۴۲۶	0/۲۲۹۹	0/۲0۵۷	0/۴۵۳۱	$\hat{\beta}_4$
0/0۵	0/0۳۶۱	0/0۵۴۵	0/0۴۱۸	0	0	0/۴۵۵۳	0/۶۷۴۱	$\hat{\beta}_5$
0/0۵۷۹	0/۱۲۹۷	0/0۶۶۵	0/۱۴0۴	0	0	0/۷۶۳۷	0/۸۷۳۱	$\hat{\beta}_6$
0/0۵۳۷	−0/۱۲۸۹	0/0۶۲۸	−0/۱۵۶۹	0/۵۷۹۱	0/۲۹۵۳	0/۱۸۹۵	0/۴۳۴۸	$\hat{\beta}_7$
0/0۴۲۴	0/0۸۷۹	0/0۵0۵	0/0۸۲۷	0	0	0/۴۶۷۵	0/۶۸۳۱	$\hat{\beta}_8$
0/0۴۵۸	−0/۱0۶۶	0/0۴۹۶	−0/۱۱۴۲	0/۲0۹۲	0/۱۸۸۷	0/۱۸۸۶	0/۴۳۳۸	$\hat{\beta}_9$
0/0۶۶	0/۱۸0۸	0/0۶۹۲	0/۱۸۷۸	0	0	0/۷۳۴۶	0/۸۵۶۴	$\hat{\beta}_{10}$
0/0۶۵۳	۱/۸۲0۳	0/0۵۹۸	۱/۸۳۳	۲/۶00۳	0/۷۵۲۱	۲/۴۶۱	0/۴۳۱۴	$\hat{\beta}_{11}$
0/0۳۶۵	۱/۹۹۳	0/0۳۲۲	۲/00۳۲	۳/۶۶۵۱	0/۱۹۸۸	۲/0۲۹۶	0/۵۷۵۶	$\hat{\beta}_{12}$
0/0۵0۱	۲/0۳۷۷	0/0۴۷۴	۲/0۴۶۵	۳/۱۵۴۴	0/۴۴۴۶	۲/0۶۸۲	0/۵۶۲۱	$\hat{\beta}_{13}$
0/0۸۱	۱/۸۴۶۸	0/0۷۴	۱/۸۵۵۷	۳/۱۲۸	۱/۳0۶۴	۲/۴۱۶۴	0/۴۴۵۷	$\hat{\beta}_{14}$
0/0۶0۶	۲/0۶۳	0/0۶	۲/0۷۸۳	۴	0	۱/۳۹۹	0/۸۱۷۷	$\hat{\beta}_{15}$
0/0۷۶۲	۲/0۷۶۶	0/0۶۶۱	۲/0۹۳۹	۴	0	۱/۵۹0۳	0/۷۳۹۴	$\hat{\beta}_{16}$
0/0۷۱۴	۱/۸۲۱۱	0/0۶۴۸	۱/۸۳0۲	۲/۴۸۱۱	۱/۱۷۸۵	۲/۳۶۸۹	0/۴۶۱	$\hat{\beta}_{17}$
0/0۴۶۹	۲/0۹0۷	0/0۴۶۶	۲/۱۱۹	۴	0	۱/۱۸۴	0/۹۱۲۶	$\hat{\beta}_{18}$
0/0۴۵۹	۱/۹۸۲۱	0/0۴۱۱	۱/۹۹۸۲	۴	0	۱/۵۴۳۶	0/۷۵۸	$\hat{\beta}_{19}$
0/0۵۳۶	۱/۹۵۶۹	0/0۴۸۲	۱/۹۸۱۷	۴	0	۱/۲۸۸۵	0/۸۶۵۵	$\hat{\beta}_{20}$
0/0۴۲۶	0/۱۲۱	0/0۵0۳	0/۱۲۴۱	0	0	0/۷۶۷۳	0/۸۷۴۷	$\hat{\beta}_{21}$
0/0۳۴۲	−0/0۵۵	0/0۳۸۵	−0/0۶۵۴	0	0	0/۵۴۲۷	0/۷۳۶۱	$\hat{\beta}_{22}$
0/0۳۸۵	0/0۵۶	0/0۴۱۷	0/0۵۵۷	0	0	0/۸۱0۶	0/۸۹۹۳	$\hat{\beta}_{23}$
0/0۴۷۲	0/0۸۵۱	0/0۵۲۳	0/0۹۶۳	0	0	0/۸۱۹۶	0/۹0۴۳	$\hat{\beta}_{24}$
0/0۳۷۳	−0/0۲0۸	0/0۴0۸	−0/0۳۳۵	0	0	0/۴۱۴۳	0/۶۴۳۱	$\hat{\beta}_{25}$
0/0۳۲۷	0/0۳۵۳	0/0۳۶۸	0/0۲۹۵	0/0۲۱۱	0/0۱۸۴	0/۳۱۷۶	0/۵۶۳	$\hat{\beta}_{26}$
0/0۵۶۶	−0/۱۳۶۹	0/0۶۶۲	−0/۱۶۲۳	0/۴۳۱۴	0/۲۱۹۵	0/۱۴۴۶	0/۳۷۹۸	$\hat{\beta}_{27}$
0/0۵۶۸	0/۱۵۲۲	0/0۶۵۱	0/۱۵۵۲	0	0	0/۷۵۱۲	0/۸۶۵۹	$\hat{\beta}_{28}$
0/0۴0۳	0/0۲۶۳	0/0۴۸	0/0۲۸۳	0	0	0/۴۵۳۷	0/۶۷۲۹	$\hat{\beta}_{29}$
0/0۴0۷	0/0۶۵۳	0/0۴۳۴	0/0۶۶۳	0/0۴۷۹	0/0۴۱۶	0/۳۲۱۲	0/۵۶۶۱	$\hat{\beta}_{30}$
0/0۷۷۳	۱/۸۱۵۳	0/0۷۴	۱/۸۱۸۵	۳/۳۳0۷	0/۳0۸۷	۲/۶۸۱۱	0/۳۶۲۷	$\hat{\beta}_{31}$
0/0۴۸۲	۱/۹۹۷۳	0/0۴۴۷	۲/00۷۴	۲/۸۲۸۳	0/۸0۴۴	۲/۲۴۹۹	0/۵00۲	$\hat{\beta}_{32}$
0/0۴۹۷	۱/۹۹۲۱	0/0۴۵۱	۲/00۶۵	۳/۶۳0۶	0/۱۹۵۸	۲/00۷۳	0/۵۸۳۵	$\hat{\beta}_{33}$
0/0۸۲۴	۱/۸۲۶۴	0/0۷۵۲	۱/۸۳۴۶	۲/۴۸۱۹	۱/۳۴۹۲	۲/۴0۲۳	0/۴۵0۲	$\hat{\beta}_{34}$
0/0۸	۲/۱۹۲۲	0/0۸۳۷	۲/۲۱۸۶	۴	0	۲/0۹۵۷	0/۵۵۴۲	$\hat{\beta}_{35}$
0/0۴۶۷	۱/۹۴۶	0/0۴۲۳	۱/۹۵۴۳	۳/۵0۵۹	0/۲۴0۹	۲/0۳۶۵	0/۵۷۳۲	$\hat{\beta}_{36}$
0/0۶۸	۱/۸۷۳۵	0/0۶۱۴	۱/۸۸0۴	۲/۷۸۹۹	۱/۳۳۸۴	۲/۳۷۵۵	0/۴۵۸۹	$\hat{\beta}_{37}$
0/0۴۷۸	۲/0۵۹۴	0/0۴۴۸	۲/0۶۵۷	۳/۹۸۹۶	0/00۲۸	۱/۸۴۹۲	0/۶۴0۴	$\hat{\beta}_{38}$
0/0۵۱۳	۱/۹۹۹۸	0/0۴۷۳	۱/۹۸۹۸	۳/۸۳۹	0/0۸0۸	۲/0۱۸۷	0/۵۷۹۴	$\hat{\beta}_{39}$
0/۲۸۲۴	۲/0۵۷۱	0/۲۲۴۳	۲/0۱۷۱	۲۳/۸0۷۲	۶/۵۷۴۵	۲/۳۳۹۴	0/۴۷0۶	$\hat{\beta}_{40}$
۲/۲۹۸	−	۲/۲۶۹۲	−	۹۱/۸0۱۶	−	۴۹/۳0۴۴	−	MSE
				MM – BR	MM	LASSO	Ridge	
				۵۸/۹۳۳	۵۸/۸۶۱۳	۸۴۵/۷۳۷۵	۸۱۸/۴۷۱۲	$\sqrt{\text{MPE}}$
				۱۳/۸۸۸۲	۱۳/۹0۵۱	0/۹۶۷۸	۱	REF
				۹۶۵/۸۸۲۵	۹۶۵/۸۸۲۵	۹۶۵/۸۸۲۵	۹۶۵/۸۸۲۵	MCN
				۷/۹۳	۶/۷۲	۱۸/۶۹	−	C
				0	0	۱۴/۴۹	−	IC

جدول ۱۶.۴: نتایج شبیه‌سازی مربوط به بریج استوار، طرح ۲، $n = 300$ ، $\rho = 0.99$ و $\tau = 0.2$

MM – BR		MM		LASSO		Ridge		ضرایب
MSE	mean	MSE	mean	MSE	mean	MSE	mean	
0.1113	-0.0136	0.1903	0.0769	0.0864	0.0468	0.1537	0.3916	β_1
0.23	0.2339	0.3346	0.1091	2.0722	0.7777	0.2302	0.4793	β_2
0.2583	0.0733	0.3537	-0.0879	0.1166	0.0931	0.1617	0.4016	β_3
0.2882	0.1328	0.3769	0.0916	1.3027	0.5725	0.2038	0.451	β_4
0.2516	0.1524	0.3932	0.0268	0	0	0.4556	0.6744	β_5
0.285	0.2143	0.4221	0.2025	0	0	0.7763	0.8806	β_6
0.182	0.1017	0.2841	-0.0454	1.2593	0.4923	0.1876	0.4326	β_7
0.196	0.1217	0.3055	0.0092	0	0	0.4678	0.6834	β_8
0.1909	0.1611	0.2674	0.1142	0.7028	0.3763	0.1864	0.4312	β_9
0.2395	0.2488	0.3913	0.2898	0	0	0.7406	0.86	β_{10}
0.508	1.7194	0.3998	1.778	3.1123	0.4236	2.4721	0.4278	β_{11}
0.2944	1.812	0.2253	1.8904	3.1994	0.777	2.0417	0.5714	β_{12}
0.4452	2.0093	0.3458	2.0712	3.4268	1.2095	2.082	0.5573	β_{13}
0.4966	1.7972	0.3799	1.8442	3.3968	0.9291	2.4291	0.4416	β_{14}
0.4146	1.9537	0.3502	2.0559	4	0	1.4018	0.8164	β_{15}
0.3234	1.9928	0.264	2.0771	4	0	1.6007	0.7352	β_{16}
0.4859	1.6977	0.3614	1.76	2.8563	1.0048	2.3814	0.457	β_{17}
0.3568	2.0529	0.3211	2.1584	4	0	1.1763	0.916	β_{18}
0.4868	1.847	0.395	1.9609	4	0	1.5497	0.7555	β_{19}
0.4252	1.7675	0.3092	1.8908	4	0	1.2805	0.8689	β_{20}
0.2587	0.2388	0.3791	0.2307	0	0	0.7989	0.8929	β_{21}
0.2155	0.1705	0.3129	0.1046	0	0	0.5451	0.7378	β_{22}
0.2099	0.1187	0.2957	0.094	0	0	0.843	0.9174	β_{23}
0.2379	0.1729	0.3459	0.1776	0	0	0.8481	0.9202	β_{24}
0.1821	0.1041	0.2751	0.051	0.1963	0.0753	0.4158	0.6443	β_{25}
0.2255	0.1662	0.3084	0.1006	1.9443	0.6236	0.3166	0.5621	β_{26}
0.196	0.0079	0.2725	-0.0764	0.0369	0.0335	0.1428	0.3775	β_{27}
0.3077	0.3271	0.4062	0.2704	0	0	0.7621	0.8724	β_{28}
0.286	0.0507	0.396	-0.016	0	0	0.4527	0.6722	β_{29}
0.2774	0.254	0.3743	0.2608	2.2563	0.6897	0.319	0.5642	β_{30}
0.6445	1.5607	0.5229	1.6264	4	0	2.6913	0.3596	β_{31}
0.5974	2.057	0.4079	2.0692	3.4223	1.2451	2.2633	0.4957	β_{32}
0.3915	1.9137	0.2928	1.9987	3.2901	0.5078	2.0187	0.5794	β_{33}
0.4468	1.7771	0.3593	1.8187	2.5567	0.8399	2.416	0.4458	β_{34}
0.3837	2.0246	0.3258	2.175	4	0	2.0379	0.5743	β_{35}
0.3598	1.8836	0.2993	1.9627	3.5696	0.7741	2.0471	0.5694	β_{36}
0.5884	1.8216	0.4501	1.8441	3.1838	1.1318	2.3848	0.4559	β_{37}
0.4419	1.9624	0.3511	2.0252	3.9526	0.0138	1.8387	0.6443	β_{38}
0.4319	1.8619	0.3442	1.9089	3.3287	0.5001	1.9958	0.5875	β_{39}
1.8504	1.2449	1.9781	1.2232	1.05905	4.3269	2.3191	0.4773	β_{40}
15.0028	—	15.3682	—	87.8597	—	49.436	—	MSE
MM – BR		MM		LASSO		Ridge		$\sqrt{\text{MPE}}$
60.3616	63.0048	843.4525	818.1724					REF
13.5545	12.9859	0.97	1					MCN
3737.6087	3737.6087	3737.6087	3737.6087					C
5.28	3.14	17.66	—					IC
0.12	0.04	14.13	—					

جدول ۱۷.۴: نتایج مربوط به برآوردگر بریج استوار برای داده‌های Vo22

MM-Bridge	MM	LASSO	Ridge	
۵۷/۶۶۶	۵۸/۷۷۵	۲۰۴/۲۵۳	۲۴۵/۷۷۶	$\sqrt{\text{MPE}}$



شکل ۱.۴: نمودار جعبه‌ای مربوط به خطاهای پیشگویی برای بریج استوار

استوار از دیگر برآوردگرها عملکرد بهتری دارد و خطای پیشگویی کمتری را به دست می‌دهد و به خوبی اثرات همخطی بین متغیرهای تبیینی و حضور داده پرت در مجموعه داده را کنترل می‌کند.

۷.۴ نتیجه‌گیری و آینده تحقیق

هرگاه در بین متغیرهای توضیحی همخطی شدید وجود داشته باشد و مجموعه داده‌ها دارای مشاهده یا مشاهدات ناهمگون باشد، برآوردگرهای معمولی قادر به برازش خوب مجموعه داده نخواهند بود. در این صورت باید از روش‌های جایگزینی که در حضور مشکلات همخطی و داده‌های ناهمگون تحت تاثیر قرار نمی‌گیرند، استفاده کرد. در این پایان‌نامه برای مقابله با مشکلات همخطی و داده‌های ناهمگون موجود در مجموعه داده‌ها، یک برآوردگر انقباضی استوار به نام برآوردگر بریج استوار معرفی شد که از یک الگوریتم مکرر برای محاسبه آن استفاده شده است. در این الگوریتم با استفاده از یک برآوردگر اولیه استوار با نقطه فروریزش بالا به محاسبه برآوردگر بریج استوار پرداخته شده است و تا رسیدن به یک دقت مطلوب برای پارامترها ادامه می‌یابد.

در این پایان‌نامه، برآوردگر بریج را با یک MM – برآوردگر استوار ترکیب کرده و نسخه‌ای استوار از برآوردگر انقباضی بریج معرفی کردیم. برآوردگر بریج استوار معرفی شده، نسبت به داده‌های دورافتاده و نقاط نافذ موجود در مجموعه داده، استوار خواهد بود و در حضور همخطی بالای بین متغیرهای توضیحی نیز برآورد خوبی از ضرایب مدل ارائه می‌دهد. همچنین با استفاده از دو طرح شبیه‌سازی و تحلیل داده واقعی، کارایی برآوردگر معرفی شده را بررسی کردیم و در مجموع شاهد عملکرد خوبی از این برآوردگر بودیم، به خصوص در حضور داده‌ی ناهمگون و همخطی بالا در بین متغیرهای توضیحی، عملکرد حالت‌های خاص این برآوردگر، یعنی ریج و لاسوی استوار، را نیز مورد بررسی قرار دادیم که نسبتاً عملکرد خوبی از خود نشان دادند. علاوه بر این برای زمانی که متغیرها یا مشاهدات مربوط به آن‌ها دقیق نباشند، برآوردگر استواری برای حالت‌های ریج و لاسو ارائه کردیم و در مثال‌هایی شاهد عملکرد خوب این برآوردگرهای فازی استوار نیز بودیم که در حضور همخطی بین مجموعه داده‌ها، قادر به برازش خوب مدل هستند. در کل می‌توان گفت در حضور داده دورافتاده و نقاط نافذ و همچنین همخطی بالای بین متغیرهای توضیحی، روش بریج استوار ارائه شده، نسبت به دیگر روش‌های معرفی شده، برتری دارد.

می‌توان نتایج این پایان‌نامه را به برآوردگر بریج تطبیقی استوار نیز تعمیم داد. این برآوردگر به صورت زیر تعریف می‌شود.

$$\operatorname{argmin}_{\beta} \left\{ \hat{\sigma}_s^2 \sum_{i=1}^n \rho_1 \left(\frac{r_i(\beta)}{\hat{\sigma}_s} \right) + \sum_{j=1}^p \lambda_i |\beta_j|^\gamma \right\}$$

که در آن λ_i پارامتر تنظیم‌کننده برای هر پارامتر β_j به‌ازای $j = 1, \dots, p$ می‌باشد و مشابه مقاله ژو [۶۶] می‌توان برای محاسبه آن‌ها از برآوردگر اولیه بریج استوار استفاده نمود. همچنین می‌توان برآوردگر بریج استوار را در حالت فازی تعمیم داد. بررسی برآوردگر بریج استوار از دیدگاه فازی نیز می‌تواند موضوع پژوهش دیگری باشد.

فصل ۵

گزیده‌ای از کدها

۱.۵ کدهای مربوط به فصل دوم

کدهای زیر مربوط به رگرسیون ريج استوار فازی بخش ۳.۶.۲، با استفاده از نرم‌افزار متمتیکا، است:

```
X = Import["E:\\ridge-data.xlsx", {"Data", 2}];
n = Length[Transpose[X][[1]]];
a = 0;
b = 2;
c = 0.1;
p = 4;
mA = Array[h, {p, 1}];
lA = Array[h, {p, 1}];
rA = Array[h, {p, 1}];
Ao = {mA0, lA0, rA0};
mAp = {mp1, mp2, mp3, mp4};
lAp = {lp1, lp2, lp2, lp4};
rAp = {rp1, rp2, rp3, rp4};
mAz = {mz1, mz2, mz3, mz4};
lAz = {lz1, lz2, lz3, lz4};
rAz = {rz1, rz2, rz3, rz4};
Yhat = Array[h, {n, 3}];
Y = Array[w, {n, 3}];
```

```

d = {};
For[j = 1, j <= p, j++,
  d = Join[d, {Ao[[2]] >= 0, Ao[[3]] >= 0, mAp[[j]] >= 0, mAz[[j]] >= 0, lAp[[j]] >= 0, lAz[[j]] >= 0,
    rAp[[j]] >= 0, rAz[[j]] >= 0, mAp[[j]] - lAp[[j]] >= 0, mAz[[j]] - rAz[[j]] >= 0}]];
For[i = 1, i <= n, i++,
  Y[[i, 1]] = X[[i, 13]];
  Y[[i, 2]] = X[[i, 14]];
  Y[[i, 3]] = X[[i, 15]]; ];
For[i = 1, i <= n, i++,
  Yhat[[i, 1]] = Ao[[1]] + Sum[(mAp[[j]] - mAz[[j]])*X[[i, j]], {j, 1, p}];
  Yhat[[i, 2]] = Ao[[2]] + Sum[mAp[[j]]*X[[i, j + 4]] + lAp[[j]]*X[[i, j]] + mAz[[j]]*X[[i, j + 8]]
    + lAz[[j]]*X[[i, j]], {j, 1, p}];
  Yhat[[i, 3]] = Ao[[3]] + Sum[mAp[[j]]*X[[i, j + 8]] + rAp[[j]]*X[[i, j]] + mAz[[j]]*X[[i, j + 4]]
    + rAz[[j]]*X[[i, j]], {j, 1, p}]; ];
W = {"k", "Q1", "mA0", "lA0", "rA0", "mp1", "mp2", "mp3", "mp4", "lp1", "lp2", "lp3", "lp4", "rp1",
  "rp2", "rp3", "rp4", "mz1", "mz2", "mz3", "mz4", "lz1", "lz2", "lz3", "lz4", "rz1", "rz2", "rz3", "rz4"};
For[k = a, k <= b, k += c,
  Q1 = NMinimize[{Sum[Abs[Yhat[[i, 1]] - Y[[i, 1]]], {i, 1, n}] + Sum[Abs[Yhat[[i, 2]] - Y[[i, 2]]], {i, 1, n}]
    + Sum[Abs[Yhat[[i, 3]] - Y[[i, 3]]], {i, 1, n}] + k*(Sum[(mAp[[j]] - mAz[[j]])^2 + (1/3)*(mAp[[j]]
    - mAz[[j]])*(rAp[[j]] + rAz[[j]]) - (lAp[[j]] + lAz[[j]]) + (1/12)*((lAp[[j]] + lAz[[j]])^2
    + (rAp[[j]] + rAz[[j]])^2), {j, 1, p}] + (Ao[[1]]^2 + (1/3)*Ao[[1]]*(Ao[[3]] - Ao[[2]])
    + (1/12)*(Ao[[2]]^2 + Ao[[3]]^2))), d], {mA0, lA0, rA0, mp1, mp2, mp3, mp4, lp1, lp2, lp3, lp4,
    rp1, rp2, rp3, rp4, mz1, mz2, mz3, mz4, lz1, lz2, lz3, lz4, rz1, rz2, rz3, rz4}];
AppendTo[W, {k, Q1[[1]], Q1[[2, 1, 2]], Q1[[2, 2, 2]], Q1[[2, 3, 2]], Q1[[2, 4, 2]], Q1[[2, 5, 2]], Q1[[2, 6, 2]]
  , Q1[[2, 7, 2]], Q1[[2, 8, 2]], Q1[[2, 9, 2]], Q1[[2, 10, 2]], Q1[[2, 11, 2]], Q1[[2, 12, 2]]
  , Q1[[2, 13, 2]], Q1[[2, 14, 2]], Q1[[2, 15, 2]], Q1[[2, 16, 2]], Q1[[2, 17, 2]], Q1[[2, 18, 2]]
  , Q1[[2, 19, 2]], Q1[[2, 20, 2]], Q1[[2, 21, 2]], Q1[[2, 22, 2]], Q1[[2, 23, 2]], Q1[[2, 24, 2]]
  , Q1[[2, 25, 2]], Q1[[2, 26, 2]], Q1[[2, 27, 2]]}]; Print[k]]

W
B = {"k", "mA1", "mA2", "mA3", "mA4"}
lenk = Floor[(b - a)/c + 1];
mA = Array[h, {lenk, p}];
lA = Array[h, {lenk, p}];
rA = Array[h, {lenk, p}];
For[k = 1, k < Length[Transpose[W][[1]]], k++,
  For[j = 1, j <= p, j++,
    mA[[k, j]] = W[[k + 1, j + 5]] - W[[k + 1, j + 17]];
    lA[[k, j]] = W[[k + 1, j + 9]] + W[[k + 1, j + 21]];
    rA[[k, j]] = W[[k + 1, j + 13]] + W[[k + 1, j + 25]]; ];
AppendTo[B, {k, mA[[k, 1]], mA[[k, 2]], mA[[k, 3]], mA[[k, 4]]}]];
Yhat1 = Array[h, {lenk, n}];
Yhat2 = Array[h, {lenk, n}];
Yhat3 = Array[h, {lenk, n}];
For[k = 1, k <= lenk, k++, For[i = 1, i <= n, i++,
  Yhat1[[k, i]] = W[[k+1, 3]] + Sum[W[[k+1, j+5]]*X[[i, j]], {j, 1, p}] - Sum[W[[k+1, j+17]]*X[[i, j]], {j, 1, p}];
  Yhat2[[k, i]] = W[[k+1, 4]] + Sum[W[[k+1, j+5]]*X[[i, j+4]], {j, 1, p}] + Sum[W[[k+1, j+9]]*X[[i, j]]
    , {j, 1, p}] + Sum[W[[k+1, j+17]]*X[[i, j+8]], {j, 1, p}] + Sum[W[[k+1, j+21]]*X[[i, j]], {j, 1, p}];
  Yhat3[[k, i]] = W[[k+1, 5]] + Sum[W[[k+1, j+5]]*X[[i, j+8]], {j, 1, p}] + Sum[W[[k+1, j+13]]*X[[i, j]], {j, 1, p}]]

```

```

+Sum[W[[k+1,j+17]]*X[[i,j+4]],{j,1,p}]+Sum[W[[k+1,j+25]]*X[[i,j]],{j,1,p}]]];
min = Array[h, {lenk, n}];
max = Array[h, {lenk, n}];
yt = Array[h, {n, 1}];
Yt = Array[h, {lenk, n}];
For[k = 1, k <= lenk, k++, For[i = 1, i <= n, i++,
yt[[i]] = Piecewise[{{(x - (X[[i, 13]] - X[[i, 14]])) / X[[i, 14]], X[[i, 13]] - X[[i, 14]] <= x <= X[[i, 13]]}, {(X[[i, 13]]
+ X[[i, 15]] - x) / X[[i, 15]], X[[i, 13]] <= x <= X[[i, 13]] + X[[i, 15]]}}];
Yt[[k, i]] = Piecewise[{{(x - (Yhat1[[k, i]] - Yhat2[[k, i]])) / Yhat2[[k, i]], Yhat1[[k, i]] - Yhat2[[k, i]] <= x <=
Yhat1[[k, i]]}, {(Yhat1[[k, i]] + Yhat3[[k, i]] - x) / Yhat3[[k, i]], Yhat1[[k, i]] <= x <= Yhat1[[k, i]] + Yhat3[[k, i]]}}];
min[[k, i]] = Min[Yhat1[[k, i]] - Yhat2[[k, i]], X[[i, 13]] - X[[i, 14]]];
max[[k, i]] = Max[Yhat1[[k, i]] + Yhat3[[k, i]], X[[i, 13]] + X[[i, 15]]]; ];
V2 = {"k", "TE1", "T\[Lambda]"};
Ers = Array[h, {lenk, n}];
\[Lambda] = Array[h, {lenk, n}];
For[k = 1, k <= lenk, k++, For[i = 1, i <= n, i++,
Ers[[k, i]] = Integrate[Abs[Yt[[k, i]] - yt[[i]]], {x, min[[k, i]], max[[k, i]]}];
\[Lambda] [[k, i]] = Abs[Y[[i, 1]] - Yhat1[[k, i]]]; ];
T\[Lambda] = Array[h, {lenk, 1}];
TErs = Array[h, {lenk, 1}];
For[k = 1, k <= lenk, k++, TErs[[k]] = Sum[Ers[[k, i]], {i, 1, n}];
T\[Lambda] [[k]] = Sum[\[Lambda] [[k, i]], {i, 1, n}];
AppendTo[V2, {k, TErs[[k]], T\[Lambda] [[k]]}]];

```

۲.۵ کدهای مربوط به فصل سوم

کدهای R مربوط به جدول ۶.۳:

```

rm(list = ls())
#-----
set.seed(2234)
#----- packages -----
library(lasso2)
library(fit.models)
library(MASS)
library(mvtnorm)
library(Matrix)
library(foreach)
library(glmnet)
library(SparseM)
library(quantreg)
#-----
n <- 50                                #number of observation
B <- 100                                #number of iteration
epsilon <- 0.2                          #contamination level
beta0 <- c(3,1.5,0,0,2,0,0,0)          #real parametears
p <- length(beta0)                      #number of variables

```

```

Rho = 0.99                                #multiple colinarity inetensy
mu.xcont <- c(5,3,6,2,7,8,4,9)            #contaminated location parameter
sigma.cont <- diag(p)                     #contaminated scale parameter
beta1 <- c(rep(1,p))                      #contaminated coefficients
m <- floor(epsilon*n)                     #number of contaminated observation
sigmas <- 0.5                             #controls signal-to-noise ratio
msigma <- matrix(NA,nrow = p ,ncol = p) #matrix V
CN <- c(rep(NA,B))                        #for colinearity
#----- vij -----
for (i in 1:p) {
  for (j in 1:p) {
    msigma[i,j] <- 0.5^(abs(i-j))
  }
}
#----- iterater -----
betah_ladlasso <- matrix(NA,nrow = p ,ncol = B)
betah_mm <- matrix(NA,nrow = p ,ncol = B)
betah_olasso <- matrix(NA,nrow = p ,ncol = B)
for(b in 1:B){
#----- data generation -----
e <- rt(n-m, df =3)                       # error term(vertical observation)
z.new <- rmvnorm(n-m, mean = c(rep(0,p)), sigma = msigma)
xg= matrix(NA,nr = nrow(z.new),nc = ncol(z.new))
for(i.Rho in 1:nrow(z.new)){
  for(j.Rho in 1:ncol(z.new)){
    xg[i.Rho,j.Rho] <- (sqrt(1-Rho^2))*z.new[i.Rho,j.Rho] + Rho*z.new[i.Rho,p]
  }
}
yg <- c(xg %*% beta0 + sigmas*e)           #good responses
x.cont <- rmvnorm(m , mean=mu.xcont , sigma=sigma.cont) #conteminated predictors
y.cont <- c(x.cont %*% beta1)               #contaminated responses
x.total <- rbind(xg , x.cont)               #combine explanatory
y.total <- c(yg , y.cont)                  #combine responses
#----- initial beta with lmrob -----
mm.e <- rlm(y.total ~ -1 + x.total , method = "M")
betah_mm[,b] <- as.numeric(mm.e$coefficients)
lasso.e <- cv.glmnet(x.total , y.total , alpha = 1 , intercept = FALSE)
olasso_beta <- as.numeric(coef(lasso.e, s = "lambda.min"))
betah_olasso[,b] <- olasso_beta[-1]
#-----
LAD.LASSO<-function(X,Y){
  n=nrow(X)
  p=ncol(X)
  Y <- as.matrix(Y)
  lam<-numeric(p)
  e.mat<-matrix(0, nr=p, nc=p)
  Y1<-Y-median(Y)
  m.X<-apply(X,2,median)

```

```

X1<-X-m.X
data=as.data.frame(cbind(X1,Y1))
lad<-coef(rq(Y1 ~ X1-1,0.5))
for(j in 1:p){
  lam[j]= log(n)/(n*abs(lad[j]))
  e.mat[j,j]=1
}
X.new<-rbind(X,e.mat)
Y.new<-rbind(Y,matrix(rep(0,p),nr=p))
lad.lasso<-coef(rq(Y.new ~ X.new-1,0.5))
return(round(lad.lasso,9))
}

#-----
betah_ladlasso[,b] <- LAD.LASSO(x.total , y.total)
CN[b]= kappa(t(x.total)%*%x.total)
cat("CN=",CN[b], "b=",b)
}

#----- MSE for each one of beta -----

#
vSE_mmb <- matrix(NA,nrow = B ,ncol = p)
vSE_rrb <- matrix(NA,nrow = B ,ncol = p)
vSE_mrb <- matrix(NA,nrow = B ,ncol = p)
MSE_mmb <- matrix(NA,nrow = p ,ncol = 1)
MSE_rrb <- matrix(NA,nrow = p ,ncol = 1)
MSE_mrb <- matrix(NA,nrow = p ,ncol = 1)
for (j in 1:p){
  for(i in 1:B){
    vSE_mrb[i,j] <- betah_ladlasso[j,i] - beta0[j]
    vSE_mmb[i,j] <- betah_mm[j,i] - beta0[j]
    vSE_rrb[i,j] <- betah_olasso[j,i] - beta0[j]
  }
}
for (j in 1:p){
  MSE_mmb[j] <- (1/B)*(sum((vSE_mmb[,j])^2))
  MSE_rrb[j] <- (1/B)*(sum((vSE_rrb[,j])^2))
  MSE_mrb[j] <- (1/B)*(sum((vSE_mrb[,j])^2))
}

#----- MSE for vector beta -----
#
se_mrba <- betah_ladlasso - beta0
se_mmba <- betah_mm - beta0
se_rrba <- betah_olasso - beta0
semrbeta <- c(rep(NA,B))
semmbeta <- c(rep(NA,B))
serrbeta <- c(rep(NA,B))
for(i in 1:B){

```

```

semrbeta[i] <- t(se_mrba[,i])%*%se_mrba[,i]
semmbeta[i] <- t(se_mmba[,i])%*%se_mmba[,i]
serrbeta[i] <- t(se_rrba[,i])%*%se_rrba[,i]
}
#semrbeta
#semmbeta
#serrbeta
MSEbmr <- mean(semrbeta)
MSEbmm <- mean(semmbeta)
MSEbrr <- mean(serrbeta)
#----- mean prediction error -----
yhat_mrb <- matrix(NA,nrow = n ,ncol = B)
yhat_mmb <- matrix(NA,nrow = n ,ncol = B)
yhat_rrb <- matrix(NA,nrow = n ,ncol = B)
semry <- c(rep(NA,B))
semmy <- c(rep(NA,B))
serry <- c(rep(NA,B))
for(i in 1:B){
  yhat_mrb[,i] <- x.total%*%betah_ladlasso[,i]
  yhat_mmb[,i] <- x.total%*%betah_mm[,i]
  yhat_rrb[,i] <- x.total%*%betah_olasso[,i]
se_mry <- yhat_mrb - y.total
se_mmy <- yhat_mmb - y.total
se_rry <- yhat_rrb - y.total
}
for(i in 1:B){
  semry[i] <- t(se_mry[,i])%*%se_mry[,i]
  semmy[i] <- t(se_mmy[,i])%*%se_mmy[,i]
  serry[i] <- t(se_rry[,i])%*%se_rry[,i]
}
#----- count the zeros and non-zeros -----
zero_ladlasso <- 0
zero_olasso <- 0
zero_mm <- 0
true_nonzero <- 0
for (j in 1:p){
  for(i in 1:B){
    if (abs(betah_ladlasso[j,i]) < 10^(-1)){
      zero_ladlasso <- zero_ladlasso + 1
    }
  }
  if (abs(betah_olasso[j,i]) < 10^(-1)){
    zero_olasso <- zero_olasso + 1
  }
  if (abs(betah_mm[j,i]) < 10^(-1)){
    zero_mm <- zero_mm + 1
  }
}
}

```



```
C_ladlasso <- zero_ladlasso/B
C_olasso <- zero_olasso/B
C_mm <- zero_mm/B
number_zero <- matrix(c(C_mm,C_ladlasso,C_olasso), nrow = 1 , ncol = 3, byrow = F,
dimnames = list(1 , c("C_mm","C_ladlasso","C_olasso")))
#----- results -----
MSE_allbeta <- matrix(c(MSE_mmb,MSE_mrb,MSE_rrb), nrow = p , ncol = 3, byrow = F,
dimnames = list(c(seq(1,p)) , c("MSE_allmmb","MSE_alladb","MSE_allolassob")))
MSE_allbeta
MSE_vecbeta <- matrix(c(MSEbmm,MSEbmr,MSEbrr), nrow = 1 , ncol = 3, byrow = F,
dimnames = list(1 , c("MSE_vecmmb","MSE_vecadb","MSE_vecplassob")))
MSE_vecbeta
MSE_predicty <- matrix(c(mean(semmy),mean(semry),mean(serry),sqrt(mean(semmy))
,sqrt(mean(semry)),sqrt(mean(serry))), nrow = 2 , ncol = 3, byrow = T,
dimnames = list(c(1,2) , c("MSE_ymm","MSE_ylad","MSE_yolasso")))
MSE_predicty
mean(CN)
number_zero
```

کدهای زیر مربوط به رگرسیون لاسو استوار سازی بخش ۱.۷.۳، با استفاده از نرم افزار
متمتیکا، است:

```
X = Import["E:\\ridge-data.xlsx", {"Data", 2}];
n = Length[Transpose[X][[1]]];
a = 0;
b = 5;
c = 0.5;
p = 4;
mA = Array[h, {p, 1}];
lA = Array[h, {p, 1}];
rA = Array[h, {p, 1}];
Ao = {mA0, lA0, rA0};
mAp = {mp1, mp2, mp3, mp4};
lAp = {lp1, lp2, lp3, lp4};
rAp = {rp1, rp2, rp3, rp4};
mAz = {mz1, mz2, mz3, mz4};
lAz = {lz1, lz2, lz3, lz4};
rAz = {rz1, rz2, rz3, rz4};
Yhat = Array[h, {n, 3}];
Y = Array[w, {n, 3}];
d = {};
For[j = 1, j <= p, j++,
d=Join[d,{Ao[[2]]>=0,Ao[[3]]>=0,mAp[[j]]>=0,mAz[[j]] >=0,lAp[[j]]>=0,lAz[[j]]>= 0,rAp[[j]]>=0,
rAz[[j]] >= 0, mAp[[j]] - lAp[[j]] >= 0, mAz[[j]] - rAz[[j]] >= 0}]];
For[i = 1, i <= n, i++,
Y[[i, 1]] = X[[i, 13]];
Y[[i, 2]] = X[[i, 14]];
Y[[i, 3]] = X[[i, 15]]; ];
```

```

For[i = 1, i <= n, i++,
  Yhat[[i, 1]] = Ao[[1]] + Sum[(mAp[[j]] - mAz[[j]])*X[[i, j]], {j, 1, p}];
  Yhat[[i, 2]] = Ao[[2]] + Sum[mAp[[j]]*X[[i, j+4]] + lAp[[j]]*X[[i, j]] + mAz[[j]]*X[[i, j+8]] + lAz[[j]]*X[[i, j]], {j, 1, p}];
  Yhat[[i, 3]] = Ao[[3]] + Sum[mAp[[j]]*X[[i, j+8]] + rAp[[j]]*X[[i, j]] + mAz[[j]]*X[[i, j+4]] + rAz[[j]]*X[[i, j]], {j, 1, p}];];
W = {"k", "Q1", "mA0", "lA0", "rA0", "mp1", "mp2", "mp3", "mp4", "lp1", "lp2", "lp3", "lp4", "rp1", "rp2", "rp3", "rp4",
  "mz1", "mz2", "mz3", "mz4", "lz1", "lz2", "lz3", "lz4", "rz1", "rz2", "rz3", "rz4"};
For[k = a, k <= b, k += c,
  Q1 = NMinimize[{Sum[Abs[Yhat[[i, 1]] - Y[[i, 1]]], {i, 1, n}] + Sum[Abs[Yhat[[i, 2]] - Y[[i, 2]]], {i, 1, n}]
    + Sum[Abs[Yhat[[i, 3]] - Y[[i, 3]]], {i, 1, n}] + k*(Sum[Abs[mAp[[j]] - mAz[[j]]] + Abs[lAp[[j]] + lAz[[j]]]
    + Abs[rAp[[j]] + rAz[[j]]], {j, 1, p}] + Abs[Ao[[1]]] + Abs[Ao[[2]]] + Abs[Ao[[3]]]), d], {mA0, lA0, rA0, mp1, mp2,
    mp3, mp4, lp1, lp2, lp3, lp4, rp1, rp2, rp3, rp4, mz1, mz2, mz3, mz4, lz1, lz2, lz3, lz4, rz1, rz2, rz3, rz4}];
AppendTo[W, {k, Q1[[1]], Q1[[2, 1, 2]], Q1[[2, 2, 2]], Q1[[2, 3, 2]], Q1[[2, 4, 2]], Q1[[2, 5, 2]], Q1[[2, 6, 2]], Q1[[2, 7, 2]],
  Q1[[2, 8, 2]], Q1[[2, 9, 2]], Q1[[2, 10, 2]], Q1[[2, 11, 2]], Q1[[2, 12, 2]], Q1[[2, 13, 2]], Q1[[2, 14, 2]],
  Q1[[2, 15, 2]], Q1[[2, 16, 2]], Q1[[2, 17, 2]], Q1[[2, 18, 2]], Q1[[2, 19, 2]], Q1[[2, 20, 2]], Q1[[2, 21, 2]],
  Q1[[2, 22, 2]], Q1[[2, 23, 2]], Q1[[2, 24, 2]], Q1[[2, 25, 2]], Q1[[2, 26, 2]], Q1[[2, 27, 2]]}]; Print[k]]

W

B = {"k", "mA1", "lA1", "rA1", "mA2", "lA2", "rA2", "mA3", "lA3", "rA3", "mA4", "lA4", "rA4"};
lenk = Floor[(b - a)/c + 1];
mA = Array[h, {lenk, p}];
lA = Array[h, {lenk, p}];
rA = Array[h, {lenk, p}];
For[k = 1, k < Length[Transpose[W][[1]]], k++,
  For[j = 1, j <= p, j++,
    mA[[k, j]] = W[[k + 1, j + 5]] - W[[k + 1, j + 17]];
    lA[[k, j]] = W[[k + 1, j + 9]] + W[[k + 1, j + 21]];
    rA[[k, j]] = W[[k + 1, j + 13]] + W[[k + 1, j + 25]]; ];
AppendTo[B, {k, mA[[k, 1]], lA[[k, 1]], rA[[k, 1]], mA[[k, 2]], lA[[k, 2]], rA[[k, 2]], mA[[k, 3]], lA[[k, 3]], rA[[k, 3]],
  mA[[k, 4]], lA[[k, 4]], rA[[k, 4]]}]];

B

Yhat1 = Array[h, {lenk, n}];
Yhat2 = Array[h, {lenk, n}];
Yhat3 = Array[h, {lenk, n}];
For[k = 1, k <= lenk, k++, For[i = 1, i <= n, i++,
  Yhat1[[k, i]] = W[[k+1, 3]] + Sum[W[[k+1, j+5]]*X[[i, j]], {j, 1, p}] - Sum[W[[k+1, j+17]]*X[[i, j]], {j, 1, p}];
  Yhat2[[k, i]] = W[[k+1, 4]] + Sum[W[[k+1, j+5]]*X[[i, j+4]], {j, 1, p}] + Sum[W[[k+1, j+9]]*X[[i, j]], {j, 1, p}]
    + Sum[W[[k+1, j+17]]*X[[i, j+8]], {j, 1, p}] + Sum[W[[k+1, j+21]]*X[[i, j]], {j, 1, p}];
  Yhat3[[k, i]] = W[[k+1, 5]] + Sum[W[[k+1, j+5]]*X[[i, j+8]], {j, 1, p}] + Sum[W[[k+1, j+13]]*X[[i, j]], {j, 1, p}]
    + Sum[W[[k+1, j+17]]*X[[i, j+4]], {j, 1, p}] + Sum[W[[k+1, j+25]]*X[[i, j]], {j, 1, p}];];
min = Array[h, {lenk, n}];
max = Array[h, {lenk, n}];
yt = Array[h, {n, 1}];
Yt = Array[h, {lenk, n}];
For[k = 1, k <= lenk, k++, For[i = 1, i <= n, i++,
  yt[[i]] = Piecewise[{{(x - (X[[i, 13]] - X[[i, 14]]))/X[[i, 14]], X[[i, 13]] - X[[i, 14]] <= x <= X[[i, 13]]}, {X[[i, 13]]
    + X[[i, 15]] - x)/X[[i, 15]], X[[i, 13]] <= x <= X[[i, 13]] + X[[i, 15]]}]];
  Yt[[k, i]] = Piecewise[{{(x - (Yhat1[[k, i]] - Yhat2[[k, i]]))/Yhat2[[k, i]], Yhat1[[k, i]] - Yhat2[[k, i]] <= x <=
    Yhat1[[k, i]]}, {(Yhat1[[k, i]] + Yhat3[[k, i]] - x)/Yhat3[[k, i]], Yhat1[[k, i]] <= x <= Yhat1[[k, i]] + Yhat3[[k, i]]}]];
  min[[k, i]] = Min[Yhat1[[k, i]] - Yhat2[[k, i]], X[[i, 13]] - X[[i, 14]]];

```

```

max[[k, i]] = Max[Yhat1[[k, i]] + Yhat3[[k, i]], X[[i, 13]] + X[[i, 15]]]; ];
V2 = {"k", "TE1", "T\[Lambda]"};
Ers = Array[h, {lenk, n}];
\[Lambda] = Array[h, {lenk, n}];
For[k = 1, k <= lenk, k++, For[i = 1, i <= n, i++,
    Ers[[k, i]] = Integrate[Abs[Yt[[k, i]] - yt[[i]]], {x, min[[k, i]], max[[k, i]]}];
    \[Lambda] [[k, i]] = Abs[Y[[i, 1]] - Yhat1[[k, i]]]; ];
T\[Lambda] = Array[h, {lenk, 1}];
TErs = Array[h, {lenk, 1}];
For[k = 1, k <= lenk, k++,
    TErs[[k]] = Sum[Ers[[k, i]], {i, 1, n}];
    T\[Lambda] [[k]] = Sum\[Lambda] [[k, i]], {i, 1, n}];
AppendTo[V2, {k, TErs[[k]], T\[Lambda] [[k]]}];
V2

```

۳.۵ کدهای مربوط به فصل چهارم

کدهای R مربوط به جدول ۶.۴:

```

rm(list = ls())
#-----
set.seed(2234)
#----- packages -----
library(lasso2)
library(mvtnorm)
library(fit.models)
library(robust)
library(Matrix)
library(foreach)
library(glmnet)
#-----
#----- findelbow function -----
#-----
findElbow <- function(y, plot = FALSE, returnIndex = TRUE) {
  # The following helper functions were found at
  # paulbourke.net/geometry/pointlineplane/pointline.r
  # via the SO reference below. The short segment check
  # was modified to permit vectorization.
  ##=====
  ## Credits:
  ## Theory by Paul Bourke http://local.wasp.uwa.edu.au/~pbourke/geometry/pointline/
  ## Based in part on C code by Damian Coventry Tuesday, 16 July 2002
  ## Based on VBA code by Brandon Crosby 9-6-05 (2 dimensions)
  ## With grateful thanks for answering our needs!
  ## This is an R (http://www.r-project.org) implementation by Gregoire Thomas 7/11/08
  ##=====
  distancePointLine <- function(x, y, slope, intercept) {

```

```

## x, y is the point to test.
## slope, intercept is the line to check distance.
##
## Returns distance from the line.
##
## Returns 9999 on 0 denominator conditions.
x1 <- x-10
x2 <- x+10
y1 <- x1*slope+intercept
y2 <- x2*slope+intercept
distancePointSegment(x,y, x1,y1, x2,y2)
}

distancePointSegment <- function(px, py, x1, y1, x2, y2) {
  ## px,py is the point to test.
  ## x1,y1,x2,y2 is the line to check distance.
  ## Returns distance from the line, or if the intersecting
  ## point on the line nearest
  ## the point tested is outside the endpoints of the line, the distance to the
  ## nearest endpoint.
  ##
  ## Returns 9999 on 0 denominator conditions.
  lineMagnitude <- function(x1, y1, x2, y2) sqrt((x2-x1)^2+(y2-y1)^2)
  ans <- NULL
  ix <- iy <- 0 # intersecting point
  lineMag <- lineMagnitude(x1, y1, x2, y2)
  if(any(lineMag < 0.00000001)) { # modified for vectorization by BAH
    #warning("short segment")
    #return(9999)
    warning("At least one line segment given by x1, y1, x2, y2 is very short.")
  }
  u <- (((px - x1) * (x2 - x1)) + ((py - y1) * (y2 - y1)))
  u <- u / (lineMag * lineMag)
  if(any((u < 0.00001) || (u > 1))) { # BAH added any to vectorize
    ## closest point does not fall within the line segment
    ##, take the shorter distance
    ## to an endpoint
    ix <- lineMagnitude(px, py, x1, y1)
    iy <- lineMagnitude(px, py, x2, y2)
    if(ix > iy) ans <- iy
    else ans <- ix
  } else {
    ## Intersecting point is on the line, use the formula
    ix <- x1 + u * (x2 - x1)
    iy <- y1 + u * (y2 - y1)
    ans <- lineMagnitude(px, py, ix, iy)
  }
  ans
}

```

```
# End of helper functions by PB
### Now for the actual findElbow function!
# Find the elbow using the method described in
# stackoverflow.com/a/2022348/633251
# but translated to R (see above).
# Add an index to argument values for easy plotting
DF <- data.frame(x = 1:length(y), y = y)
fit <- lm(y ~ x, DF[c(1,nrow(DF)),]) # 2 point 'fit'
m <- coef(fit)[2]
b <- coef(fit)[1]
# Check to make sure the data is concave as described
# in the documentation, as arbitrary trends could give
# misleading answers. The following approach simply
# checks to make sure all values are either above or
# below the reference line. This allows the values
# to vary quite a bit and still return an answer.
concave <- FALSE
use <- 2:(nrow(DF)-1)
refpts <- m*DF$x[use] + b
if (all(refpts > DF$y[use]) | all(refpts < DF$y[use])) concave <- TRUE
if (!concave) stop("Your curve doesn't appear to be concave")
# Calculate the orthogonal distances
use <- 2:(nrow(DF)-1)
elbowd <- distancePointLine(DF$x[use], DF$y[use], coef(fit)[2], coef(fit)[1])
DF$dist <- c(NA, elbowd, NA) # first & last points don't have a distance
if (plot) {
  edm <- which.max(DF$dist)
  plot(DF[,1:2], type = "b", xlab = "index", ylab = "y values",
       main = "Looking for the Elbow")
  segments(DF$x[1], DF$y[1],
           DF$x[nrow(DF)], DF$y[nrow(DF)], col = "red")
  points(DF$x[edm], DF$y[edm], cex = 1.5, col = "red")
  points(DF$x[edm], DF$y[edm], pch = 20)
}
if (returnIndex) return(which.max(DF$dist))
if (!returnIndex) return(DF)
} # end of findElbow

#=====
#-----
n <- 50 #number of observation
gamma <- 0.5 #control parameter
B <- 100 #number of iteration
epsilon <- 0.2 #contamination level
beta0 <- c(3,1.5,0,0,2,0,0,0) #real parametears
p <- length(beta0) #number of variables
Rho = 0.99 #multiple colinarity inetensy
mu.xcont <- c(5,3,6,2,7,8,4,9) #contaminated location parameter
sigma.cont <- diag(p) #contaminated scale parameter
```

```

beta1 <- c(rep(1,p))                #contamination-beta-parameter
m <- floor(epsilon*n)                #number of contaminated observation
sigmas <- 0.5                        #controls signal-to-noise ratio
msigma <- matrix(NA,nrow = p ,ncol = p) #matrix V
CN <- c(rep(NA,B))                  #for colinearity
#_____ vij _____
for (i.vm in 1:p) {
  for (j.vm in 1:p) {
    msigma[i.vm,j.vm] <- 0.5^(abs(i.vm-j.vm))
  }
}
#_____ iterater _____
betah_update <- matrix(NA,nrow = p ,ncol = B)
betah_mm <- matrix(NA,nrow = p ,ncol = B)
betah_rr <- matrix(NA,nrow = p ,ncol = B)
betah_lasso <- matrix(NA,nrow = p ,ncol = B)
  for(b in 1:B){
#_____ data generation _____
e <- rt(n-m, df =3)
z.new <- rmvnorm(n-m, mean = c(rep(0,p)), sigma = msigma)
xg= matrix(NA,nr = nrow(z.new),nc = ncol(z.new))
for(i.Rho in 1:nrow(z.new)){
  for(j.Rho in 1:ncol(z.new)){
    xg[i.Rho,j.Rho] <- (sqrt(1-Rho^2))*z.new[i.Rho,j.Rho]
      + Rho*z.new[i.Rho,p]
  }
}
yg <- c(xg %%% beta0 + sigmas*e)
x.cont <- rmvnorm(m , mean=mu.xcont , sigma=sigma.cont)
y.cont <- c(x.cont %%% beta1)
x.total <- rbind(xg , x.cont)
y.total <- c(yg , y.cont)
#boxplot(x.total)
#_____ initial beta with lmrob _____
mm.e <- lmRob(y.total ~ -1 + x.total) #calculate MM-Estimation
betah_mm[,b] <- as.numeric(mm.e$coefficients)
#
rr.e <- cv.glmnet(x.total , y.total , alpha = 0 , intercept = FALSE)
oridge_beta <- as.numeric(coef(rr.e, s = "lambda.min"))
#
olasso.e <- cv.glmnet(x.total , y.total , alpha = 1 , intercept = FALSE)
olasso_beta <- as.numeric(coef(olasso.e, s = "lambda.min"))
#
betah_rr[,b] <- oridge_beta[-1]
betah_lasso[,b] <- olasso_beta[-1]
#
#
#_____ y_hat _____

```

```

a_re <- 0
beta_mm2 <- betah_mm[,b]
beta_iter <- matrix(0,nrow = p ,ncol = 1)
repeat{
  if(beta_mm2 - beta_iter <= 10^-1 ) break
y_hat <- x.total%*% beta_mm2
#----- residuals -----
#
r_beta <- y.total - y_hat
#
#----- r.beta/sigma_s -----
rt_beta <- r_beta/sigmas
#----- Hubor function -----
rho_1 <- array(NA , c(n))
  for(i.rho1 in 1:n){
    rho_1[i.rho1] <- min(1 , 1-(1-(rt_beta[i.rho1])^2)^3)
  }
#----- derivation Hubor function -----
psi_1 <- array(NA , c(n))
  for(i.psi1 in 1:n){
    psi_1[i.psi1] <- (rt_beta[i.psi1]>(-1) & rt_beta[i.psi1]<1)*
      ((6*rt_beta[i.psi1])-(12*(rt_beta[i.psi1]^3))+(6*(rt_beta[i.psi1]^5)))
      +(rt_beta[i.psi1]>=1)*0+(rt_beta[i.psi1]<=(-1))*0
  }
#----- Matrix W(beta) -----
w_s <- array(NA , c(n))
  for(i.w.s in 1:n){
    w_s[i.w.s] <- psi_1[i.w.s]/rt_beta[i.w.s]
  }
#
W_beta <- diag(w_s)
#
#----- Matrix W0(beta) -----
w0_s <- array(NA , c(p))
  for(j.w0.s in 1:p){
    w0_s[j.w0.s] <- abs(beta_mm2[j.w0.s])^(gamma - 2)
  }
#
W0_beta <- diag(w0_s)
#
#----- inverse matrix & estimata robust bridge _
lamb <- seq(0 , 5000 , by=1)
lambda <- lamb[-1]
l_1 <- length(lambda)
l.b <- array(NA , c(p,l_1))
#
tBIC_b <- array(NA , c(l_1))

```

```

#-----
  for(l in 1:l_l){
    l.b[,l] <- solve(t(x.total)%*%W_beta)%*%x.total + lambda[l]
    *gamma*W0_beta)%*%t(x.total)%*%W_beta)%*%y.total
#-----
y_hl <- x.total)%*% l.b[,l]
#-----
r_bl <- y.total - y_hl
#-----
rt_bl <- r_bl/sigmas
#-----
rho <- array(NA , c(n))
for(i.rho.t in 1:n){
  rho[i.rho.t] <- min(1 , 1-(1-(rt_bl[i.rho.t])^2)^3)
}
#-----
psi <- array(NA , c(n))
for(i.psi.t in 1:n){
  psi[i.psi.t] <- (rt_bl[i.psi.t]>(-1) & rt_bl[i.psi.t]<1)*
  ((6*rt_bl[i.psi.t])-(12*(rt_bl[i.psi.t]^3))+(6*(rt_bl[i.psi.t]^5)))
  +(rt_bl[i.psi.t]>=1)*0+(rt_bl[i.psi.t]<=(-1))*0
}
#-----
w_sl <- array(NA , c(n))
for(i.w1.s1 in 1:n){
  w_sl[i.w1.s1] <- psi[i.w1.s1]/rt_bl[i.w1.s1]
}
#-----
W_bl <- diag(w_sl)
#-----
w0_sl <- array(NA , c(p))
for(j.w0.s1 in 1:p){
  w0_sl[j.w0.s1] <- abs(l.b[j.w0.s1,l])^(gamma - 2)
}
#-----
W0_bl <- diag(w0_sl)
#-----
Hat <- (x.total)%*%(solve(t(x.total)%*%W_bl)%*%x.total
+ lambda[l]*gamma*W0_bl))%*%(t(x.total))%*%(W_beta)
#-----
tBIC_b[l] <- n*log((sigmas^2)*(sum(rho)))+tr(Hat)*log(n)
#-----
#
}
#tBIC_b
#length(tBIC_b)
#plot(tBIC_b)
#plot(y.total)

```



```
#pairs(x.total)
#par(mfcol = c(2, 1))
#boxplot(x.total)
#boxplot(y.total)
##par(mfcol = c(2, 1))
index = findElbow(tBIC_b)
index.b = tBIC_b[index]
betah.n <- l.b[,index.b]
beta_iter <- beta_mm2
beta_mm2 <- betah.n
a_re <- a_re+1
print(a_re)
    }

#
#betah.n
betah_update[,b] <- beta_mm2
CN[b]= kappa(t(x.total)%*%x.total)
cat("CN=",CN[b], "b=",b)
    # print(b)
    }

#----- MSE for each one of beta -----

#
vSE_mmb <- matrix(NA,nrow = B ,ncol = p)
vSE_rrb <- matrix(NA,nrow = B ,ncol = p)
vSE_mrb <- matrix(NA,nrow = B ,ncol = p)
vSE_lassob <- matrix(NA,nrow = B ,ncol = p)
#
MSE_mmb <- matrix(NA,nrow = p ,ncol = 1)
MSE_rrb <- matrix(NA,nrow = p ,ncol = 1)
MSE_mrb <- matrix(NA,nrow = p ,ncol = 1)
MSE_lassob <- matrix(NA,nrow = p ,ncol = 1)
#
for (j.MSEeb in 1:p){
  for(i.MSEeb in 1:B){
    vSE_mrb[i.MSEeb,j.MSEeb] <- betah_update[j.MSEeb,i.MSEeb] - beta0[j.MSEeb]
    vSE_mmb[i.MSEeb,j.MSEeb] <- betah_mm[j.MSEeb,i.MSEeb] - beta0[j.MSEeb]
    vSE_rrb[i.MSEeb,j.MSEeb] <- betah_rr[j.MSEeb,i.MSEeb] - beta0[j.MSEeb]
    vSE_lassob[i.MSEeb,j.MSEeb] <- betah_lasso[j.MSEeb,i.MSEeb] - beta0[j.MSEeb]
  }
}
#
for (j.meb in 1:p){
  MSE_mmb[j.meb] <- (1/B)*(sum((vSE_mmb[,j.meb])^2))
  MSE_rrb[j.meb] <- (1/B)*(sum((vSE_rrb[,j.meb])^2))
  MSE_mrb[j.meb] <- (1/B)*(sum((vSE_mrb[,j.meb])^2))
  MSE_lassob[j.meb] <- (1/B)*(sum((vSE_lassob[,j.meb])^2))
}
```

```

}
#----- MSE for vector beta -----
se_mrba <- betah_update - beta0
se_mmba <- betah_mm - beta0
se_rrba <- betah_rr - beta0
se_lassoba <- betah_lasso - beta0
semrbeta <- c(rep(NA,B))
semmbeta <- c(rep(NA,B))
serrbeta <- c(rep(NA,B))
selassobeta <- c(rep(NA,B))
for(i.se.be in 1:B){
  semrbeta[i.se.be] <- t(se_mrba[,i.se.be])%*%se_mrba[,i.se.be]
  semmbeta[i.se.be] <- t(se_mmba[,i.se.be])%*%se_mmba[,i.se.be]
  serrbeta[i.se.be] <- t(se_rrba[,i.se.be])%*%se_rrba[,i.se.be]
  selassobeta[i.se.be] <- t(se_lassoba[,i.se.be])%*%se_lassoba[,i.se.be]
}
MSEbmr <- mean(semrbeta)
MSEbmm <- mean(semmbeta)
MSEbrr <- mean(serrbeta)
MSEblasso <- mean(selassobeta)
#----- mean prediction error -----
yhat_mrb <- matrix(NA,nrow = n ,ncol = B)
yhat_mmb <- matrix(NA,nrow = n ,ncol = B)
yhat_rrb <- matrix(NA,nrow = n ,ncol = B)
yhat_lassob <- matrix(NA,nrow = n ,ncol = B)
semry <- c(rep(NA,B))
semmry <- c(rep(NA,B))
serrry <- c(rep(NA,B))
selassory <- c(rep(NA,B))
for(i.se.yh in 1:B){
  yhat_mrb[,i.se.yh] <- x.total%*%betah_update[,i.se.yh]
  yhat_mmb[,i.se.yh] <- x.total%*%betah_mm[,i.se.yh]
  yhat_rrb[,i.se.yh] <- x.total%*%betah_rr[,i.se.yh]
  yhat_lassob[,i.se.yh] <- x.total%*%betah_lasso[,i.se.yh]

  se_mry <- yhat_mrb - y.total
  se_mmy <- yhat_mmb - y.total
  se_rry <- yhat_rrb - y.total
  se_lassoy <- yhat_lassob - y.total
}
for(i.se.yh in 1:B){
  semry[i.se.yh] <- t(se_mry[,i.se.yh])%*%se_mry[,i.se.yh]
  semmy[i.se.yh] <- t(se_mmy[,i.se.yh])%*%se_mmy[,i.se.yh]
  serrry[i.se.yh] <- t(se_rry[,i.se.yh])%*%se_rry[,i.se.yh]
  selassoy[i.se.yh] <- t(se_lassoy[,i.se.yh])%*%se_lassoy[,i.se.yh]
}
#----- count the zeros and non-zeros -----
vec_zeros <- c(which(beta0 == 0))

```

```

zero_bridge <- 0
zero_olasso <- 0
zero_mm <- 0
for (j.zero in 1:length(vec_zeros)){
  for(i.zero in 1:B){
    if (abs(betah_update[vec_zeros[j.zero],i.zero]) < 10^(-1)){
      zero_bridge <- zero_bridge + 1
    }
    if (abs(betah_lasso[vec_zeros[j.zero],i.zero]) < 10^(-1)){
      zero_olasso <- zero_olasso + 1
    }
    if (abs(betah_mm[vec_zeros[j.zero],i.zero]) < 10^(-1)){
      zero_mm <- zero_mm + 1
    }
  }
}
C_bridge <- zero_bridge/B
C_olasso <- zero_olasso/B
C_mm <- zero_mm/B
#IC <- s_n/B
#IC
C_zero <- matrix(c(C_mm,C_bridge,C_olasso), nrow = 1 , ncol = 3, byrow = F,
dimnames = list(1 , c("C_mm","C_bridge","C_olasso")))
#----- number of nonzeros that cont as zero wrongly _
vec_nzero <- c(which(beta0 != 0))
nzw_bridge <- 0
nzw_olasso <- 0
nzw_mm <- 0
for(j.nonz in 1:length(vec_nzero)){
  for(i.nonz in 1:B){
    if (abs(betah_update[vec_nzero[j.nonz],i.nonz]) < 10^(-1)){
      nzw_bridge <- nzw_bridge + 1
    }
    if (abs(betah_lasso[vec_nzero[j.nonz],i.nonz]) < 10^(-1)){
      nzw_olasso <- nzw_olasso + 1
    }
    if (abs(betah_mm[vec_nzero[j.nonz],i.nonz]) < 10^(-1)){
      nzw_mm <- nzw_mm + 1
    }
  }
}
IC_bridge <- nzw_bridge/B
IC_olasso <- nzw_olasso/B
IC_mm <- nzw_mm/B
IC_nzero <- matrix(c(IC_mm,IC_bridge,IC_olasso), nrow = 1 , ncol = 3, byrow = F,
dimnames = list(1 , c("IC_mm","IC_br","IC_olasso")))
#----- results -----
#

```

```

MSE_allbeta <- matrix(c(MSE_mmb,MSE_mrb,MSE_rrb,MSE_lassob), nrow = p
,ncol = 4, byrow = F,
dimnames = list(c(seq(1,p)) , c("beta_MM","beta_MM-BR","beta_RR","beta_LASSO")))
MSE_allbeta
#
MSE_vecbeta <- matrix(c(MSEbmm,MSEbmr,MSEbrr,MSEblasso), nrow = 1
, ncol = 4, byrow = F,
dimnames = list(1 , c("beta_MM","beta_MM-BR","beta_RR","beta_LASSO")))
MSE_vecbeta
#
MSE_predicty <- matrix(c(mean(semmy),mean(semry),mean(serry)
,mean(selassoy),sqrt(mean(semmy)),sqrt(mean(semry)),sqrt(mean(serry))
,sqrt(mean(selassoy))), nrow = 2 , ncol = 4, byrow = T,dimnames =
list(c("MSE_p","sqrt") , c("beta_MM","beta_MM-BR","beta_RR","beta_LASSO")))
#
MSE_predicty
#
mean(CN)
C_zero
IC_nzero

```

مراجع

- [۱] آرست م. (۱۳۹۵)، مقایسه رفتار برخی برآوردگرهای انقباضی بریج در مدل رگرسیون چندگانه، پایان‌نامه ارشد، دانشگاه صنعتی شاهرود.
- [۲] آرست م.، آرشی م. و ربیعی م. (۱۳۹۸)، مطالعه رفتار برآوردگر انقباضی تحت یک قید خطی در مدل رگرسیون جریمه شده، مجله علوم آماری، ۱۳(۱)، ۱-۱۶.
- [۳] ربیعی، م. ر. (۱۳۸۰)، رگرسیون استوار و مقایسه آن با روش کمترین توان‌های دوم، پایان‌نامه کارشناسی ارشد، دانشگاه شهید بهشتی تهران.
- [۴] رجبلو، س. (۱۳۹۷)، رگرسیون جریمه‌شده در بعد بالا با تابع جریمه لاسوی مشتق‌پذیر، پایان‌نامه کارشناسی ارشد، دانشگاه صنعتی شاهرود.
- [۵] روزبه، م. (۱۳۹۰)، برآورد در مدل‌های خطی جزئی، رساله دکتری، دانشگاه فردوسی مشهد.
- [۶] فرخی، م. (۱۳۹۵)، مطالعه چندین روش رگرسیون ریج در محیط فازی، پایان‌نامه ارشد، دانشگاه صنعتی شاهرود.
- [۷] طاهری، م.، (۱۳۷۵)، آشنایی با نظریه مجموعه‌های فازی، انتشارات جهاد دانشگاهی مشهد.
- [۸] طاهری، م. و ماشین‌چی، م.، (۱۳۹۲)، مقدمه‌ای بر احتمال و آمار فازی، چاپ دوم، انتشارات دانشگاه شهید باهنر کرمان.
- [۹] مونت گمری د. و پک، ا.، (۲۰۱۲)، (ترجمه ابراهیم رضوی پاریزی) مقدمه‌ای بر تحلیل رگرسیون خطی، جلد اول، چاپ چهارم، انتشارات دانشگاه شهید باهنر کرمان.
- [۱۰] نوروزی‌راد، م. (۱۳۹۶)، برآوردگرهای بهبودیافته در برخی مدل‌های رگرسیونی خطی جریمه‌شده، رساله دکتری، دانشگاه صنعتی شاهرود.
- [۱۱] نیرومند ح. ع.، (۲۰۱۲)، تحلیل بر رگرسیون خطی، جلد اول، چاپ اول، انتشارات دانشگاه فردوسی مشهد.

-
- [12] Akaike, H. (1973), Information theory and an extension of the maximum likelihood principle, In 2nd International Symposium on Information Theory, eds. B. N. Petrov and F. Csaki, Budapest: Akademia Kiado, 267–281.
 - [13] Arslan, O. (2016), Penalized MM regression estimation with L_γ penalty: a robust version of bridge regression, *Statistics*, **50**(6), 1236-1260.
 - [14] Balasundaram, S. K. (2011), Weighted fuzzy ridge regression analysis with crisp inputs and triangular fuzzy outputs, *International Journal of Advanced Intelligence Paradigms*, **3**(1), 67-81.
 - [15] Barrodale, I. and Roberts, F. D. K. (1973), An improved algorithm for discrete l1 linear approximation, *SIAM Journal of Numerical Analysis*, **10**, 839–848.
 - [16] Bassett, G. and Koenker, R. (1978), Asymptotic theory of least absolute error regression, *Journal of the American Statistical Association*, **73**, 618–621.
 - [17] Belsely D. A., Kuh E. and Welsch R. E. (1980), *Regression Diagnostics*, John Wiley, New York.
 - [18] Bloomfield, P. and Steiger, W. L. (1983), *Least Absolute Deviation: Theory, Applications and Algorithms*, Boston: Birkhauser.
 - [19] Boscovich, R.J. (1757), De literaria expeditione per pontificiam ditionemet synopsis amplioris operis..., Bononiensi Scientiarum et Artum Instituto atque Academia Commentarii 4: 353-396. Reprinted with a Croatian translation by N. Cubranić, Institute of Higher Geodesy, Zagreb, 1961.
 - [20] Caner, M. (2009). LASSO-type GMM estimator, *Econometric Theory*, **25**, 270–290.
 - [21] Celmins, A. (1987). Least squares model fitting to fuzzy vector data, *Fuzzy Sets and Systems*, **22**, 260-269.
 - [22] Craven, P. and Wahba, G. (1979), Smoothing noise data with spline function: estimating the correct degree of smoothing by the method of generalized cross validation, *Numerische Mathematik*, **31**, 337–403.
 - [23] Diamond. P. (1987), Least squares fitting of several fuzzy variables, *In 2nd Int. Fuzzy Syst. Assoc. IFSA World Congress*, 329-331.
 - [24] Diamond, P. (1988), Fuzzy least squares, *Information Sciences*, **46**, 141-157.

- [25] Diamond, P. and Korner, R. (1997). Extended fuzzy linear models and least squares estimates, *Computers and Mathematics with Applications*, **33**, 15-32
- [26] Donoso, S., Marín, N. and Vila, M. A. (2007), Fuzzy ridge regression with non symmetric membership functions and quadratic models, *In International Conference on Intelligent Data Engineering and Automated Learning*, 135-144.
- [27] Dubois, D. and Prade, H. (1980), *Fuzzy Sets and Systems: theory and applications*, Academic press, Toulouse, France.
- [28] Efron, B., Hastie, T., Johnstone, I., and Tibshirani, R. (2004), Least angle regression, *The Annals of statistics*, **32(2)**, 407-499.
- [29] Fan, J. and Li, R. (2001), Variable selection via nonconcave penalized likelihood and its oracle properties, *Journal of the American Statistical Association*, **96(456)**, 1348-1360.
- [30] Frank, I. E. and Friedman, J. H. (1993), A statistical view of some chemometrics regression tools, *Technometrics*, **35(2)**, 109-135.
- [31] Fu, W. J. (1998), Penalized regression: the bridge versus lasso, *Journal of Computational and Graphical Statistics*, **7**, 397–416.
- [32] Galton F. (1886), Regression towards mediocrity in hereditary stature, *The Journal of the Anthropological Institute of Great Britain and Ireland*, **15**, 246-263.
- [33] Gibbons, D. G. A. (1981), Simulation study of some ridge estimators, *Journal of the American Statistical Association*, **76**, 131–139.
- [34] Hassanpour, H., Maleki, H. R. and Yaghoobi, M. A. (2010), Fuzzy linear regression model with crisp coefficients: a goal programming approach, *Iranian Journal of Fuzzy Systems*, **7(2)**, 1–153.
- [35] Hastie, T., Tibshirani, R. and Friedman (2009), *The Elements of Statistical Learning: Data Mining, Inference and Prediction* (2nd ed), Springer.
- [36] Hoerl, A.E. and Kennard, R.W. (1970), Ridge regression: biased estimation for non-orthogonal problems, *Technometrics*, **12**, 55-67.
- [37] Hunter, D. R. and Li, R. (2005), Variable selection using MM algorithms, *The Annals of Statistics*, **33**, 1617–1642.

- [38] Hurvich, C. M. and Tsai, C. L. (1990), Model selection for least absolute deviation regression in small samples, *Statistics and Probability Letters*, **9**, 259–265.
- [39] Izadikhah, M. (2009), Using the Hamming distance to extend TOPSIS in a fuzzy environment, *Journal of Computational and Applied Mathematics*, **231**(1), 200-207.
- [40] Knight, K. (1998), Limiting distributions for L1 regression estimators under general conditions, *The Annals of Statistics*, **26**, 755–770.
- [41] Knight, K. and Fu, W. (2000), Asymptotic for LASSO-type estimators, *Annals of Statistics*, **28**(5), 1356 - 1378.
- [42] Lawrence, K. D. and Marsh, L. C. (1984), Robust ridge estimation methods for predicting U.S. coal mining fatalities, *Communications in Statistics, Theory and Methods*, **13**, 139–149.
- [43] Li, B. and Yu, Q. (2009), Robust and sparse bridge regression, *Statistics and Its Interface*, **2**, 481–491.
- [44] Ling, S. (2005), Self-weighted least absolute deviation estimation for infinite variance autoregressive models, *Journal of the Royal Statistical Society Ser. B*, **67**, 1–13.
- [45] Kim, B. and Bishu, R. R. (1998), Evaluation of fuzzy linear regression models by comparing membership functions, *Fuzzy sets and systems*, **100**(1), 343-352.
- [46] McDonald, G. C. and Galarneau, D. I. (1975), A monte carlo evaluation of some ridge-type estimators, *Journal American Statistics Association*, **70**, 407–416.
- [47] Maronna, R. A. (2011), Robust ridge regression for high-dimensional data, *Technometrics*, **53**, 44–53.
- [48] Maronna, R. A., Martin, D. and Yohai, R. S. (2006), *Robust Statistics: Theory and Methods*, Wiley Series in Probability and Statistics.
- [49] Maronna, R. A. and Yohai, V. J. (2010), Correcting MM estimates for fat data sets, *Computational Statistics Data Analysis*, **54**, 3168–3173.
- [50] Marquart D. W. (1970), Generalized inverse ridge regression biased linear estimation and non-linear estimation, *Technometrics*, **12**, 591-612.
- [51] Mason R. L., Gunst. R. F. and Webster. J. T. (1975), Ridge analysis and problems of multicollinearity, *Communications in Statistics*, **4**(3), 277-292.

- [52] Peng, L. and Yao, Q. (2003), Least absolute deviation estimation for ARCH and GARCH models, *Biometrika*, **90**, 967–975.
- [53] Pollard, D. (1991), Asymptotic for least absolute deviation regression estimators, *Econometric Theory*, **7**, 186–199.
- [54] Portnoy, S. and Koenker, R. (1997), The Gaussian hare and the Laplacian tortoise: computability of squared-error versus absolute-error estimators, *Statistical Science*, **12**, 299–300.
- [55] Rousseeuw, P. J. and Leroy, A. M. (1987), *Robust regression and outlier detection*, John Wiley and Sons, New York.
- [56] Rousseeuw, P. J. and Yohai, V. J. (1984), Robust regression by means of S-estimators, *Robust and Nonlinear Time Series*, J. Franke, W. Härdle and R.D. Martin (eds.), *Lectures Notes in Statistics*, New York: Springer, **26**, 256–272.
- [57] Schwarz, G. (1978), Estimating the dimension of a model, *The Annals of Statistics*, **6**, 461–464.
- [58] Shao, J. (1997), An asymptotic theory for linear model selection, *Statistica Sinica*, **7**, 221–264.
- [59] Shi, P. and Tsai, C. L. (2002), Regression model selection: a residual likelihood approach, *Journal of the Royal Statistical Society, Ser. B*, **64**, 237–252.
- [60] Tibshirani, R. (1996), Regression shrinkage and selection via the lasso, *Journal of the Royal Statistical Society, Series B*, **58**, 267–288.
- [61] Wang, H. and Leng, C. (2007), Unified LASSO estimation by least squares approximation, *Journal of the American Statistical Association*, **102**, 1039–1048.
- [62] Wang, H., Li, R. and Jiang, J. (2007), Robust regression shrinkage and consistent variable selection through LAD-LASSO, *Journal of Business and Economic Statistics*, **25**, 347–355.
- [63] Woods, H., Steinour, H. H. and Starke, H. R. (1932), Effect of composition of Portland cement on heat evolved during hardening, *Industrial & Engineering Chemistry*, **24**(11), 1207–1214.
- [64] Xu, R. and Li, C. (2001), Multidimensional least squares fitting with a fuzzy model, *Fuzzy Sets and Systems*, **119**, 215–223.

- [65] Yohai, V. J. (1987), High breakdown point and high efficiency estimates for regression, *The Annals of Statistics*, **15**, 642–656.
- [66] Zou, H. (2006), The adaptive LASSO and its oracle properties, *Journal of the American Statistical Association*, **101**(476), 1418-1429.
- [67] Zou, H. and Li, R. (2008), One-step sparse estimates in non-concave penalized likelihood models, *The Annals of Statistics*, **36**, 1509–1533.

Abstract

The estimation of regression parameters is very important in analysing and predicting data. A common method is the least squares estimator that it does not have good performance in case of multicollinearity, so that the shrinkage estimators were born. One of these most popular techniques is bridge estimator. Also, the least squares and shrinkage estimators are influenced by outliers. Although, there are many papers about these estimators, the suggested methods are not good when the data suffers multicollinearity problem. In this dissertation, we propose a new method called robust bridge that controls the existence outliers and high multicollinearity. Also, the asymptotic behavior of this estimator is studied and the computational algorithm is introduced. As a special case, the robust LASSO and ridge and their fuzzied versions are verified.

Keywords: Shrinkage Methods, Bridge Regression, Robust Methods, Robust Bridge Regression, Fuzzy Regression, Robust Fuzzy Shrinkage Estimator.



Shahrood University of Technology

Faculty of Mathematical Sciences

MSc Thesis in: Statistics

A Robust Version of Bridge Regression

By: Vahid Goodarzi Vanani

Supervisors

Dr. Mohammad Reza Rabiei

Dr. Mohammad Arashi

January 2019