
بسم الله
الرحمن الرحیم

دانشکده علوم ریاضی

رشته آمار، گرایش آمار ریاضی

پایان‌نامه کارشناسی ارشد

استنباط مبتنی بر درستنمایی در رگرسیون ریج برای داده‌های طولی

نگارنده: محمد رحمانی

استاد راهنما

دکتر محمد آرشی

شهریور ۱۳۹۷

شماره:
تاریخ:

بسمه تعالی

دانشگاه صنعتی شاهرود

مدیریت تحصیلات تکمیلی

فرم شماره (۳) صورتجلسه نهایی دفاع از پایان نامه دوره کارشناسی ارشد

با نام و یاد خداوند متعال، ارزیابی جلسه دفاع از پایان نامه کارشناسی ارشد خانم / آقای محمد رحمانی با شماره دانشجویی ۹۴۰۷۷۷۴ رشته آمار گرایش آمار ریاضی تحت عنوان استنباط مبتنی بر درستنمایی در رگرسیون ریج برای داده‌های طولی که در تاریخ ۹۷/۶/۱۲ با حضور هیأت محترم داوران در دانشگاه صنعتی شاهرود برگزار گردید به شرح ذیل اعلام می‌گردد:

قبول (با درجه: عالی.....) ☑ مردود ☐

نوع تحقیق: نظری ☑ عملی ☐

عضو هیأت داوران	نام ونام خانوادگی	مرتبهٔ علمی	امضاء
۱ـ استادراهنمای اول	دکتر محمد آرشی	دانشیار	
۲- استادراهنمای دوم	------	------	
۳- استاد مشاور	------	------	
۴- نماینده تحصیلات تکمیلی	دکتر نگار اقبال	استادیار	
۵- استاد ممتحن اول	دکتر حسین باغیشنی	استادیار	
۶ـ استاد ممتحن دوم	دکتر محمدرضا ربیعی	استادیار	

نام و نام خانوادگی رئیس دانشکده:

دکتر ابراهیم هاشمی

تاریخ و امضاء و مهر دانشکده:

تبصره: در صورتی که کسی مردود شود حداکثر یکبار دیگر (در مدت مجاز تحصیل) می تواند از پایان نامه خود دفاع نماید (دفاع مجدد نباید زودتر از ۴ ماه برگزار شود).

تقدیم به
پدر بزرگوار و مادر مهربانم

آن دو فرشته‌ای که از برایم عاشقانه و خالصانه ز خویشتن گذشتند...

پس، هستی من ز هستی اوست تا هستم و هست دارمش دوست

ایرج میرزا

تقدیم به
استاد گرانقدرم دکتر آرشی

که از چشمه جوشان علم و اخلاق شایسته ایشان بهره بردم.
زیر سایه لطف الهی، پایدار باشند و برقرار...

سپاس گزاری...

منّت خدای را عزوجل که طاعتش موجب قربتست و به شکر اندرش مزید نعمت، هر نفسی که فرو می رود ممدّ حیاتست و چون برمی آید مفرّح ذات، پس در هر نفسی دو نعمت موجودست و بر هر نعمت شکری واجب...

از دست و زبان که برآید کز عهده شکرش به در آید

سعدی

پس از حمد خداوند باری تعالی بر خود لازم می‌دانم تا زحمات اساتید گرانقدرم را ارج نهم.
بسیار خرسندم که افتخار شاگردی استاد فرزانه جناب آقای دکتر محمد آرشی را دارم، بیش از همه سپاس‌گزارِ زحمات، محبت‌ها و راهنمایی‌های ایشان هستم و از خداوند رحمان بهترین‌ها را برای ایشان آرزو دارم.
از اساتید بزرگوار و دلسوزم آقایان دکتر حسین باغیشنی و دکتر محمدرضا ربیعی که خالصانه از علم خویش به من آموختند و زحمت داوری این پایان‌نامه را نیز متقبل شدند، کمال تقدیر و تشکر را دارم.
بر خود لازم می‌دانم از زحمات اساتید ارجمندم آقایان دکتر شاهسونی، دکتر نزاکتی و سرکارخانم دکتر اقبال که همواره راهنما و راه‌گشای اینجانب بودند و تمامی عزیزانی که به هر نحوی در طول دوران تحصیل و تهیه این مجموعه مرا یاری دادند صمیمانه تشکر و قدردانی نمایم و از خداوند متعال برای همه ایشان موفقیت و سربلندی را خواستارم.

سعی نابرده در این راه به جایی نرسی مزد اگر میطلبی طاعت استاد بیر

حافظ

محمد رحمانی
شهریور ۱۳۹۷

تعهد نامه

اینجانب **محمد رحمانی** دانشجوی کارشناسی ارشد رشته **آمار** دانشکده **علوم ریاضی** دانشگاه صنعتی شاهرود، نویسنده پایان‌نامه با عنوان **استنباط مبتنی بر درستنمایی در رگرسیون ریج برای داده‌های طولی** ، تحت راهنمایی **محمد آرشی** متعهد می‌شوم:

- تحقیقات در این پایان‌نامه توسط اینجانب انجام شده است و از صحت و اصالت برخوردار است.
- در استفاده از نتایج پژوهش‌های دیگر پژوهش‌گران، به مرجع مورد استفاده استناد شده است.
- مطالب این پایان‌نامه، 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 پایان‌نامه تاثیرگذار بوده‌اند، در مقالات مستخرج از پایان‌نامه رعایت می‌گردد.
- در تمام مراحل انجام این پایان‌نامه، در مواردی که از موجود زنده (یا بافت‌های آنها) استفاده شده است، ضوابط و اصول اخلاقی رعایت شده است.
- در تمام مراحل انجام این پایان‌نامه، در مواردی که به حوزه اطلاعات شخصی افراد دسترسی یافته (یا استفاده شده است)، اصل رازداری و اصول اخلاق انسانی رعایت شده است.

محمد رحمانی
شهریور 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 پایان‌نامه بدون ذکر منبع مجاز نمی‌باشد.

چکیده

داده‌های طولی به داده‌هایی اطلاق می‌شود که در جمع آوری آن‌ها، اندازه‌گیری ویژگی‌های مختلف برای هر کدام از واحدهای آزمایشی در طول زمان تکرار می‌شود. یکی از ابزارهای تحلیل داده‌های طولی استفاده از مدل‌های آمیخته خطی است، اما وجود مشکل همخطی در این نوع داده‌ها استفاده از روش‌های مختلفی همچون روش‌های انقباضی را موجب می‌شود. در این پایان‌نامه به تحلیل این نوع داده‌ها با استفاده از مدل‌های آمیخته خطی و ترکیب آن‌ها با برخی از روش‌هایی همچون رگرسیون ریج پرداخته شده و برآوردگرهای ریج و لیو با خطای بیضی‌گون ارائه شده‌اند. همچنین رفتار برآوردگرهای حاصل از این روش‌ها به کمک مطالعات شبیه‌سازی و تحلیل داده‌های واقعی مورد بررسی قرار گرفته‌است.

کلمات کلیدی: الگوریتم EM، برآوردگر انقباضی، برآوردگر لیو، توزیع بیضی‌گون، رگرسیون ریج، مدل‌های آمیخته خطی.

پیش‌گفتار

پیشینه روش‌های تحلیل داده‌های طولی به حدود نیم قرن پیش باز می‌گردد. پایه این روش‌ها توسط یک ستاره‌شناس انگلیسی بنا شد که زمینه‌ای برای فعالیت‌های پژوهشگران در این موضوع، همچون لیرد و ویر (۱۹۸۲) با معرفی مدل اثرات تصادفی برای داده‌های طولی فراهم آمد. با گسترش این روش‌ها تاکنون مدل‌های مختلفی مبنی بر تحلیل داده‌های طولی از جمله مدل‌های حاشیه‌ای، مدل‌های انتقال و مدل‌های آمیخته توسط پژوهشگران معرفی شده‌اند. در شرایطی که در داده‌های متغیرهای توضیحی همخطی وجود دارد، روش‌های معمول برآورد کارا نبوده و باید از روش‌های جایگزینی همچون ریج (هورل و کنارد، ۱۹۷۰) استفاده کرد. تمرکز این پایان‌نامه بر تحلیل داده‌های طولی با استفاده از مدل‌های آمیخته خطی است. الیوت و همکاران (۲۰۱۱) روش رگرسیون ریج را در مواجهه با داده‌های همخط در مدل آمیخته خطی، به منظور افزایش کارایی در تحلیل داده‌های طولی به کار بردند. در این مجموعه رهیافت آن‌ها را بررسی کرده و روشی جایگزین را ارائه می‌کنیم. همچنین تعمیم‌های دیگری از جمله تغییر توزیع جامعه هدف به بیضی‌گون و بررسی مدل‌های خطی تعمیم‌یافته را در نظر می‌گیریم. این مجموعه مشتمل بر ۳ فصل و یک پیوست است که محتویات آن‌ها به‌طور خلاصه در زیر آمده‌است.

- در فصل اول پس از معرفی برخی از مدل‌های ذکرشده مبنی بر تحلیل داده‌های طولی، مدل‌های آمیخته خطی را با نگاه دقیق‌تری مورد مطالعه قرار داده‌ایم.
- در فصل دوم با هدف دستیابی به نتایج مطمئن‌تر و دقیق‌تر در تحلیل داده‌های طولی، پس از معرفی و بررسی مدل آمیخته خطی ریج به منظور مقابله با وجود مشکل همخطی در داده‌ها، مدل آمیخته خطی ریج با خطای بیضی‌گون و برآوردگرهای انقباضی در مدل آمیخته خطی ریج به جهت بهبود برآوردگرها ارائه شده‌اند.
- در فصل سوم به سبب ویژگی خطی بودن برآوردگر لیو نسبت به پارامتر آن، مدل آمیخته خطی لیو، مدل آمیخته خطی لیو با خطای بیضی‌گون و برآوردگر انقباضی در مدل آمیخته خطی لیو ارائه شده و با بهره‌گیری از روش‌های شبیه‌سازی مورد بررسی قرار گرفته‌اند.
- همچنین گزیده‌ای از برنامه‌های کامپیوتری که تحت محیط R نوشته شده‌اند، در پیوست آمده‌است.

لازم به ذکر است محتویات بخش‌های ۳.۲، ۴.۲ و فصل ۳ کاملاً جدید بوده و از یافته‌های نویسنده این مجموعه است که حاصل آن تهیه ۴ مقاله است که در فهرست مقالات مستخرج از پایان‌نامه آمده‌اند.

لیست مقالات مستخرج از پایان‌نامه

1. Rahmani, M. and Arashi, M. (2017), Longitudinal Data Analysis Using Liu Regression, *11th Seminar on Probability and Stochastic Processes*, Imam Khomeini International University, Qazvin, Iran.

2. Rahmani, M. and Arashi, M. (2017), Ridge Regression for Longitudinal Data with Elliptically Contoured Errors, *48th Annual Iranian Mathematics Conference*, Bu-Ali Sina University, Hamedan, Iran.

3. Rahmani, M., Arashi, M. and Jenatabadi, H.S. (2018), Longitudinal data analysis using Liu regression, *Biostatistics and Biometric Open Access Journal*, 7(5).

4. Rahmani, M., Arashi, M., Mamode Khan, N. and Sunecher, Y. (2018), Improved mixed model for longitudinal data analysis using shrinkage method, Submitted.

فهرست مطالب

فهرست تصاویر

فهرست جداول

فصل ۱

مدل‌های خطی برای داده‌های طولی

۱.۱ مقدمه

از آنجایی‌که آشنایی با یک موضوع جدید، مستلزم شناخت مقدمات مربوط به آن موضوع می‌باشد، در این فصل به معرفی داده‌های طولی[1] و همچنین شرح چگونگی مدل‌بندی این داده‌ها می‌پردازیم. مطالب این فصل در قسمت تعریف داده‌های طولی و انواع مدل‌های آن عمدتاً برگرفته از تعاونی (۱۳۹۲) و در قسمت مدل‌های خطی (تعمیم‌یافته) از مک کولاک و سیرل (۲۰۰۱) است.

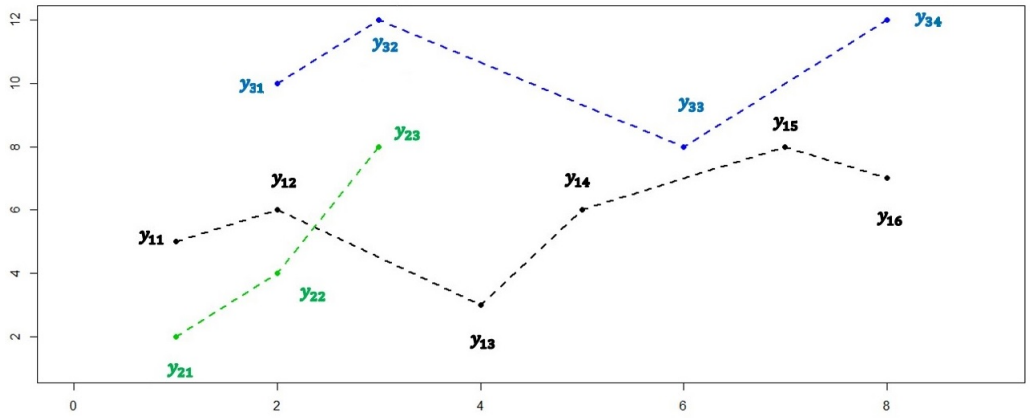
۲.۱ داده‌های طولی

نوع مطالعه و روش گردآوری داده‌ها یکی از مسائل مهم در یک تحلیل آماری می‌باشد. تحلیل‌های رگرسیونی و سری‌های زمانی دو روش متداول در تحلیل داده‌ها به شمار می‌آیند.

تحلیل داده‌های طولی، ترکیبی از تحلیل‌های رگرسیونی و سری‌های زمانی که در آن ارتباط بین چندین ویژگی در طول زمان مد نظر است را فراهم می‌آورد. بنابراین داده‌های طولی به مجموعه وسیعی از داده‌های گسسته و پیوسته اطلاق می‌شود که در آن ویژگی‌های چندین واحد آزمایشی (مانند انسان، حیوان یا نمونه‌های آزمایشگاهی) به‌عنوان نمونه‌ای از اعضای

[1] Longitudinal data

جامعه، در طول زمان اندازه‌گیری می‌شوند. لازم به ذکر است همان‌طور که در شکل ۱.۱ نشان داده شده‌است، در داده‌های طولی تعداد مشاهدات همه واحدها و همچنین فاصله نقاط زمانی آن‌ها لزوماً برابر نیستند. جهت آشنایی بیشتر با ساختار داده‌های طولی در ادامه به معرفی نمادگذاری متداول این داده‌ها می‌پردازیم.



شکل ۱.۱: نمودار ساختگی به منظور توجیه ساختار داده‌های طولی

۱.۲.۱ نمادگذاری داده‌های طولی

جدول ۱.۱ را در نظر بگیرید که در آن y_{ij}ها مقادیر متغیر پاسخ و x_{ijk} مقدار kامین متغیر پیش‌بین مربوط به واحد i ام در زمان j ام می‌باشند، به‌طوری‌که $i = ۱,\ldots,n$ ، $j = ۱,\ldots,t_i$ و $k = ۱,\ldots,p$، که در آن n بیان‌گر تعداد واحدهای تحت مطالعه، t_i تعداد دوره‌های زمانی برای واحد iام و p تعداد متغیرهای پیش‌بین است. لازم به ذکر است که هر دو متغیر پاسخ و پیش‌بین ممکن است گسسته یا پیوسته باشند. همچنین متغیرهای پیش‌بین می‌توانند در طول زمان ثابت (مانند نژاد) یا متغیر (مانند سن) باشند.

جدول ۱.۱: ساختار داده‌های طولی

واحد	زمان	y_{ij}	x_{ijk}
۱	۱	$y_{۱۱}$	$x_{۱۱۱}, \ldots, x_{۱۱p}$
⋮	⋮	⋮	⋮
۱	$t_۱$	$y_{۱t_۱}$	$x_{۱t_۱۱}, \ldots, x_{۱t_۱p}$
۲	۱	$y_{۲۱}$	$x_{۲۱۱}, \ldots, x_{۲۱p}$
⋮	⋮	⋮	⋮
۲	$t_۲$	$y_{۲t_۲}$	$x_{۱t_۲۱}, \ldots, x_{۱t_۲p}$
⋮	⋮	⋮	⋮
n	۱	$y_{n۱}$	$x_{n۱۱}, \ldots, x_{n۱p}$
⋮	⋮	⋮	⋮
n	t_n	y_{nt_n}	$x_{nt_nn}, \ldots, x_{nt_np}$

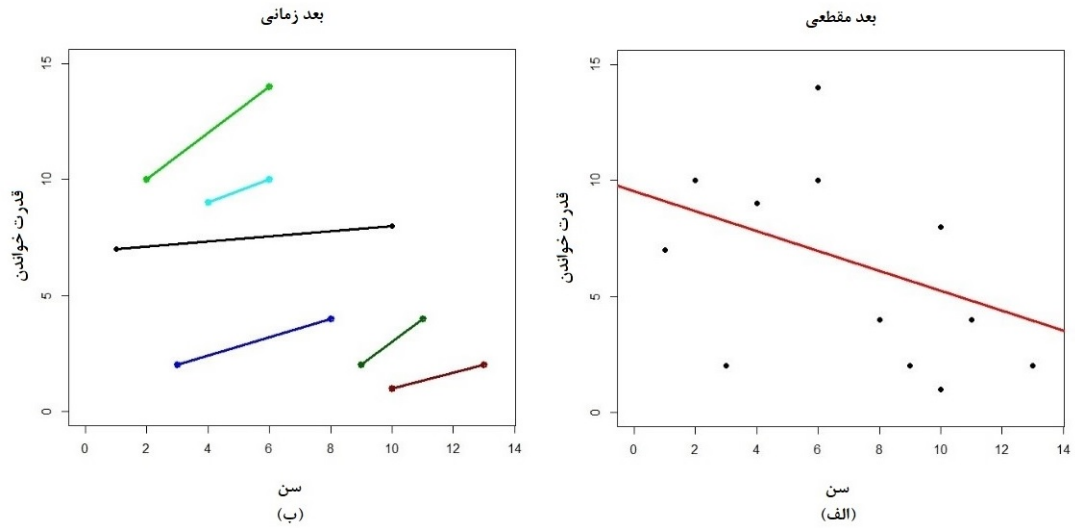
طبق آن‌چه که بیان شد، این داده‌ها را می‌توان از دو بعد زمانی و مقطعی[۲] مورد توجه قرار داد. این ماهیت دو بعدی، اهداف و لزوم ارائه ترکیبی از ابزارهای آماری را در تحلیل داده‌های طولی آشکار می‌سازد.

۲.۲.۱ اهداف و ضرورت تحلیل داده‌های طولی

با جمع‌آوری مشاهداتِ مربوط به واحدها در طول زمان، علاوه‌بر دستیابی به رابطه بین متغیرها می‌توان تغییراتِ واحدها در طول زمان را نیز مورد مطالعه قرار داد. به‌عنوان مثال اگر داده‌هایی که در زمان jام جمع‌آوری شده‌اند را در نظر بگیریم، در واقع واحدهای اول تا nام را تنها در یک مقطع از زمان مورد بررسی قرار داده‌ایم. بنابراین برای هر دوره زمانی، داده‌ها مقطعی هستند. اما اگر داده‌های مربوط به واحد iام را در نظر بگیریم، تنها یک واحد را در زمان‌های متوالی مورد بررسی قرار داده‌ایم که در این صورت، داده‌ها اثر زمان را نشان می‌دهند. لذا در تحلیل داده‌های طولی سه هدف اصلی می‌تواند مد نظر هستند: ۱) مطالعه تغییرات واحدها در طول زمان، ۲) مطالعه پیرامون رابطه بین متغیرهای پاسخ و پیش‌بین، ۳) مطالعه اثر زمان بر رابطه بین متغیرها.

برای درک بهتر این موضوع به مثالی از دیگل و همکاران (۲۰۰۲) اشاره می‌کنیم.

[2] Sectional



شکل ۲.۱: نمودار قدرت خواندن در کودکان

در شکل ۲.۱، داده‌ها حاصل یک مطالعه طولی مربوط به قدرت خواندن کودکان هستند که در آن شاخصِ توانایی مطالعه هر کودک، دو بار در طول زمان اندازه‌گیری شده‌است. بُعد مقطعی داده‌ها در قسمت (الف)، کاهش قدرت خواندن کودکان با گذشت زمان را نشان می‌دهد، درحالی‌که بُعد زمانی داده‌ها در قسمت (ب)، افزایش توانایی هر کودک با گذشت زمان را نشان می‌دهد. به‌طور کلی در مورد این داده‌ها می‌توان گفت توانایی خواندن در کودکان کم سن‌تر سطح بالاتری دارد و همچنین تمام کودکانِ مورد مطالعه، توانایی خود را با گذشت زمان بهبود می‌بخشند. بنابراین تحلیل داده‌های طولی که با هدف مشاهده تغییر رفتار پدیده‌ها در طول زمان انجام می‌گیرد، نسبت به داده‌های صرفاً مقطعی یا صرفاً زمانی از اهمیت ویژه‌ای برخوردار است.

علاوه بر حجیم بودن اطلاعات حاصل از داده‌های طولی و ماهیت دو بعدی آن‌ها که پیچیدگی داده‌ها را سبب می‌شوند، دلیل دیگری نیز وجود دارد که تحلیل این داده‌ها را با چالش رو برو می‌سازد. بر خلاف روش‌های آماری متداول که بر پایه پذیره استقلال مشاهدات طراحی شده‌اند، در این داده‌ها مجموعه مشاهدات مربوط به هر واحد، دارای همبستگی درونی بوده و از مشاهدات سایر واحدها مستقل می‌باشند. واضح است که برای دستیابی به استنباط آماری و معتبر، درنظر گرفتن این همبستگی امری ضروری است.

دلایل فوق از جمله مهمترین مواردی است که روش‌های آماری کلاسیک برای تحلیل داده‌های طولی از قبیل رگرسیون و سری‌های زمانی به تنهایی قادر به تحلیل دقیق این داده‌ها نمی‌باشند. با هدف تحلیل معتبر این داده‌ها آماردانان روش‌های متعددی معرفی نموده‌اند که در ادامه به آن‌ها خواهیم پرداخت.

۳.۱ انواع مدل‌بندی داده‌های طولی

در این بخش به معرفی و بررسی مختصرِ برخی از مدل‌های رگرسیونی همچون مدل‌های حاشیه‌ای[۳] ، مدل‌های انتقال[۴] و مدل‌های خطی آمیخته [۵] که در تحلیل داده‌های طولی به کار می‌روند، می‌پردازیم.

۱.۳.۱ مدل‌های حاشیه‌ای

مدل‌های حاشیه‌ای، مدل‌هایی هستند که در آن‌ها اثر متغیرهای پیش‌بین بر روی متغیر پاسخ به طور مجزا از همبستگی بین پاسخ‌ها، مدل‌بندی می‌شوند. به عبارت دیگر، در این روش مدل‌های جداگانه‌ای برای میانگین حاشیه‌ای، واریانس حاشیه‌ای و همبستگی حاشیه‌ای در نظر گرفته می‌شود. در صورت پیوسته بودن متغیر پاسخ، مدل حاشیه‌ای به صورت یک مدل خطی کلی با در نظر گرفتن همبستگی داده‌ها ارائه می‌شود. مدل خطی زیر یک مدل حاشیه‌ای به شمار می‌آید:

$$\mathbf{y}_i = \mathbf{X}_i^T \boldsymbol{\beta} + \boldsymbol{\epsilon}_i, i = ۱, \ldots, n, \qquad (۱.۱)$$

که در آن $\mathbf{y}_i = (y_{i۱}, y_{i۲}, \cdots, y_{in_i})^T$ و $\mathbf{X_i} = (\mathbf{x}_۱, \ldots, \mathbf{x}_{t_i})^T$. $\mathbf{x}_{t_i} \in \mathbb{R}^p$ بردار متغیرهای پیش‌بین مربوط به پارامتر β_i از $\boldsymbol{\beta} = (\beta_۱, \ldots, \beta_p)^T$ و $\boldsymbol{\epsilon}_i = (\boldsymbol{\epsilon}_۱, \ldots, \boldsymbol{\epsilon}_{t_i})^T$ بردار خطای تصادفی بوده، به‌طوری‌که $\boldsymbol{\epsilon}_i$ها ازهم مستقل و هم‌توزیع (iid) با توزیع زیر هستند

$$\boldsymbol{\epsilon}_i \overset{iid}{\sim} \mathcal{N}_{t_i}(\mathbf{0}, \mathbf{V}_i).$$

همچنین $\mathbf{V_i} > ۰, i = ۱, \ldots, n$، ماتریس معین مثبت است. وجود پذیره استقلال بین واحدهای مختلف و همبستگی درون واحدها (همبستگی حاشیه‌ای) که در مدل به صورت ماتریس $\mathbf{V_i}$ نشان داده شده‌است، تمایز مدل‌های حاشیه‌ای را با مدل‌های خطی کلاسیک، آشکار می‌سازد.

برآورد پارامترهای مدل

روش‌های متعددی برای برآورد پارامترهای مدل خطیِ (۱.۱) وجود دارند، که از جمله متداول‌ترین آن‌ها می‌توان به روش ماکزیمم درستنمایی [۶] (ML) و کمترین توان‌های دوم تعمیم‌یافته [۷](GLS) اشاره کرد. در این بخش به برآورد پارامترهای مدل با استفاده از روش ML خواهیم پرداخت.

[3] Marginal
[4] Transition
[5] Linear mixed models
[6] Maximum likelihood
[7] Generalized least squares

با توجه به مدل (۱.۱)، می‌توان نوشت

$$E(\mathbf{y}|\mathbf{X}) = \mathbf{X}\boldsymbol{\beta} \qquad (۲.۱)$$

که در آن $\mathbf{y} = (\mathbf{y}_۱^T, \ldots, \mathbf{y}_n^T)^T$ و $\mathbf{X} = (\mathbf{X}_۱, \cdots, \mathbf{X}_n)^T$. تحت مفروضات مدل (۱.۱) داریم

$$\mathbf{y}_i \overset{iid}{\sim} \mathcal{N}_{t_i}(\mathbf{X}_i^T\boldsymbol{\beta}, \mathbf{V}_i).$$

فرض کنید ماتریس کواریانس مدل به یک پارامتر α وابسته باشد که آن را با $\mathbf{V}_i(\alpha)$ نشان می‌دهیم. دلیل این تخصیص بررسی انواع ساختارهای همبستگی است که جلوتر به آن اشاره می‌کنیم.

تابع درستنمایی مدل (۱.۱) و لگاریتم آن به صورت زیر حاصل می‌شوند

$$L(\boldsymbol{\beta}, \alpha) = \prod_{i=۱}^{n} (۲\pi)^{\frac{-t_i}{۲}} |\mathbf{V}_i(\alpha)|^{\frac{-۱}{۲}} \exp\left\{-\frac{۱}{۲}(\mathbf{y}_i - \mathbf{X}_i^T\boldsymbol{\beta})^T\mathbf{V}_i(\alpha)^{-۱}(\mathbf{y}_i - \mathbf{X}_i^T\boldsymbol{\beta})\right\} \qquad (۳.۱)$$

$$l(\boldsymbol{\beta}, \alpha) = \sum_{i=۱}^{n}\left\{-\frac{t_i}{۲}\log(۲\pi) - \frac{۱}{۲}\log|\mathbf{V}_i(\alpha)|\right\} - \frac{۱}{۲}\sum_{i=۱}^{n}\left\{(\mathbf{y}_i - \mathbf{X}_i^T\boldsymbol{\beta})^T\mathbf{V}_i(\alpha)^{-۱}(\mathbf{y}_i - \mathbf{X}_i^T\boldsymbol{\beta})\right\}.$$

فرض کنید $U_i(\alpha) = \frac{d}{d\alpha}\mathbf{V}_i(\alpha)$، در این صورت با استفاده از دو قانون مشتق‌گیری

$$\frac{d\log|\mathbf{A}|}{dx} = tr\left(\mathbf{A}^{-۱}\frac{d\mathbf{A}}{dx}\right), \qquad \frac{d\mathbf{A}^{-۱}}{dx} = -\mathbf{A}^{-۱}\frac{d\mathbf{A}}{dx}\mathbf{A}^{-۱}$$

معادلات نرمال به صورت زیر به‌دست می‌آیند

$$\frac{dl(\boldsymbol{\beta}, \alpha)}{d\boldsymbol{\beta}} = \sum_{i=۱}^{n}\mathbf{X}_i^T\mathbf{V}_i(\alpha)^{-۱}(\mathbf{y}_i - \mathbf{X}_i^T\boldsymbol{\beta}) = ۰ \qquad (۴.۱)$$

$$\frac{dl(\boldsymbol{\beta}, \alpha)}{d\alpha} = \sum_{i=۱}^{n} -\frac{۱}{۲}tr\left(\mathbf{V}_i^{-۱}(\alpha)U_i(\alpha)\right)$$
$$+\frac{۱}{۲}(\mathbf{y}_i - \mathbf{X}_i^T\boldsymbol{\beta})^T\left(\mathbf{V}_i^{-۱}(\alpha)U_i(\alpha)\mathbf{V}_i^{-۱}(\alpha)\right)(\mathbf{y}_i - \mathbf{X}_i^T\boldsymbol{\beta}) = ۰ \qquad (۵.۱)$$

با حل رابطه (۴.۱)، برآورد پارامتر $\boldsymbol{\beta}$ به صورت زیر به‌دست می‌آید

$$\hat{\boldsymbol{\beta}} = \left(\sum_{i=۱}^{n}\mathbf{X}_i^T\hat{\mathbf{V}}_i^{-۱}(\alpha)\mathbf{X}_i\right)^{-۱}\sum_{i=۱}^{n}\mathbf{X}_i^T\hat{\mathbf{V}}_i^{-۱}(\alpha)\mathbf{y}_i \qquad (۶.۱)$$

که در آن $\hat{\mathbf{V}}_i(\alpha) = \mathbf{V}_i(\hat{\alpha})$ که از رابطه (۵.۱) حاصل می‌شود.

به سبب وجود $\hat{\mathbf{V}}_i^{-۱}(\alpha)$ در ساختار برآوردگر $\hat{\boldsymbol{\beta}}$، یافتن برآورد آن از طریق روش‌های عددی همچون نیوتن–رافسون امکان‌پذیر است. در ادامه فرض می‌کنیم ماتریس کواریانس $\mathbf{V}_i(\alpha)$، به صورت زیر است

$$\mathbf{V}_i(\alpha) = \phi\mathbf{A}_i^{\frac{۱}{۲}}\mathbf{R}_i(\alpha)\mathbf{A}_i^{\frac{۱}{۲}}. \qquad (۷.۱)$$

که در آن، $\mathbf{A}_i$ یک ماتریس قطری در اندازه $t_i \times t_i$ با عناصر روی قطر $\phi = var(y_{ij})$ بوده و $\mathbf{R}_i(\alpha)$ به‌عنوان ماتریس همبستگی عملی شناخته می‌شود که برای آن ساختارهای متعددی در نظر گرفته شده‌است. نظر به این که برآورد پارامترهای رگرسیونی به انتخاب این ساختارها وابسته است، می‌توان در انواع ساختارها به موارد زیر اشاره نمود.

ساختار استقلال[8]: که در آن پاسخ‌های مکرر هر واحد آزمایشی مستقل از هم فرض می‌شود.

برای ساده نویسی از نماد α در آرگومان $\mathbf{R}_i(\alpha)$ صرف نظر می‌کنیم.

$$\mathbf{R}_i = \begin{pmatrix} ۱ & ۰ & \dots & ۰ \\ ۰ & ۱ & \dots & ۰ \\ \vdots & \vdots & \ddots & \vdots \\ ۰ & ۰ & \dots & ۱ \end{pmatrix}_{t_i \times t_i} . \qquad (۸.۱)$$

بدیهی است در این حالت $\mathbf{R}_i = \mathbf{R}_i(۰)$.

ساختار همبستگی تبادل‌پذیر[9]: در آن مشاهدات مربوط به هر واحد آزمایشی، دارای همبستگی یکسان می‌باشند.

$$\mathbf{R}_i = \begin{pmatrix} ۱ & \alpha & \dots & \alpha \\ \alpha & ۱ & \dots & \alpha \\ \vdots & \vdots & \ddots & \vdots \\ \alpha & \alpha & \dots & ۱ \end{pmatrix}_{t_i \times t_i} . \qquad (۹.۱)$$

همبستگی غیر ساختاری[10]: نوع دیگری از ساختارهای همبستگی که به صورت زیر است

$$\mathbf{R}_i = \begin{pmatrix} ۱ & \alpha_{۱۲} & \dots & \alpha_{۱t_i} \\ \alpha_{۲۱} & ۱ & \dots & \alpha_{۲t_i} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{t_i ۱} & \alpha_{t_i ۲} & \dots & ۱ \end{pmatrix}_{t_i \times t_i} . \qquad (۱۰.۱)$$

همبستگی m[11]: در آن تنها مشاهداتی دارای همبستگی می‌باشند که اختلاف طول زمانی مشاهده آن‌ها کمتر از m است. بنابراین ساختار همبستگی به صورت زیر است

$$Corr(y_{ij}, y_{ij'}) = \begin{cases} ۱ & s = ۰ \\ \alpha_s & s = ۱, ۲, \dots, m \\ ۰ & s > m \end{cases} , s = |i - j|.$$

[8] Independence
[9] Exchangeable
[10] Unstructured
[11] m-dependent

صورت ماتریسی آن برای حالت $t_i = ۴$ و $m = ۲$ دارای ساختار زیر است

$$\mathbf{R}_i = \begin{pmatrix} ۱ & \alpha_۱ & \alpha_۲ & ۰ \\ \alpha_۱ & ۱ & \alpha_۱ & \alpha_۲ \\ \alpha_۲ & \alpha_۱ & ۱ & \alpha_۱ \\ ۰ & \alpha_۲ & \alpha_۱ & ۱ \end{pmatrix}_{۴\times۴} . \qquad (۱۱.۱)$$

همبستگی اتو رگرسیو[۱۲]: متداول‌ترین آن اتو رگرسیو مرتبه اول $AR(۱)$ است و به صورت زیر تعیین می‌شود

$$Corr(y_{ij}, y_{ij'}) = \alpha^s, \quad s = ۰, ۱, \ldots, t_{i-j} \qquad (۱۲.۱)$$

صورت ماتریسی آن برای حالت $t_i = ۴$ دارای ساختار زیر است

$$\mathbf{R}_i = \begin{pmatrix} ۱ & \alpha & \alpha^۲ & \alpha^۳ \\ \alpha & ۱ & \alpha & \alpha^۲ \\ \alpha^۲ & \alpha & ۱ & \alpha \\ \alpha^۳ & \alpha^۲ & \alpha & ۱ \end{pmatrix}_{۴\times۴} . \qquad (۱۳.۱)$$

در صورتی که طبق اطلاعات قبلی ساختار ماتریس همبستگی معلوم باشد، در این صورت ماتریس همبستگی به صورت یک ماتریس ثابت و معلوم در روابط محاسبه برآورد پارامتر β، جایگزین می‌شود.

پس از آشنایی با ساختارهای مختلف ماتریس همبستگی، به روش‌های عددی در برآورد پارامترهای مدل می‌پردازیم. در رابطه (۴.۱) و (۷.۱) دیدیم که معادلات نرمال علاوه بر پارامتر β، به پارامترهای همبستگی α و مقیاس[۱۳] ϕ نیز وابسته است. لذا به جهت این که معادلات نرمال به صورت تابعی از β به‌دست آیند، لازم است که α و ϕ برآورد شوند. یکی از راه‌حل‌ها، محاسبه برآورد گشتاوری آن‌ها به صورت زیر است

$$\hat{\phi} = \frac{۱}{N-p} \sum_{i=۱}^{n} \sum_{j=۱}^{t_i} R_{ij}^۲ \qquad (۱۴.۱)$$

که در آن، $R_{ij} = \frac{y_{ij} - E(y_{ij})}{\sqrt{var(y_{ij})}}$ باقیمانده‌های پیرسون و $N = \sum_{i=۱}^{n} t_i$ تعداد کل مشاهدات است.

برآورد α وابسته به انتخاب ساختارهای همبستگی است. به‌عنوان مثال برای همبستگی‌های تبادل‌پذیر، غیر ساختاری و اتو رگرسیو مرتبه اول، این برآوردگرها به ترتیب به صورت زیر

[12] Auto regressive
[13] Scale parameter

می‌باشند

$$\begin{aligned}
\hat{\alpha} &= \hat{\phi}^{-1} \sum_{i=1}^{n} \frac{1}{n_i(n_i-1)} \sum_{j \neq j'} R_{ij} R_{ij'}, \\
\hat{\alpha} &= \hat{\phi}^{-1} \frac{1}{n} \sum_{i=1}^{n} R_{ij} R_{ij'}, \\
\hat{\alpha} &= \hat{\phi}^{-1} \sum_{i=1}^{n} \frac{1}{n_i - 1} \sum_{j \leq n_i - 1} R_{ij} R_{ij+1},
\end{aligned}$$

که در آن‌ها R_{ij} مؤلفه سطر iام و ستون jام ماتریس همبستگی می‌باشد.

با به‌دست آمدن برآوردهای فوق در نهایت معادلات نرمال تابعی برحسب $\boldsymbol{\beta}$ است. برای به‌دست آوردن $\hat{\boldsymbol{\beta}}$ از روش نیوتن–رافسون[14] استفاده می‌کنیم. در زیر الگوریتم محاسبه $\hat{\boldsymbol{\beta}}$ آورده شده‌است.

الگوریتم ۱ الگوریتم نیوتن–رافسون در برآورد پارامتر مدل حاشیه‌ای

گام ۱) قرار می‌دهیم $r = 0$ و مقدار اولیه $\boldsymbol{\beta}^0$ را برآورد LS یا ML مدل خطی کلاسیک تحت فرض استقلال محاسبه می‌کنیم.

گام ۲) پارامتر مقیاس ϕ و ماتریس همبستگی عملی را براساس مانده‌های پیرسن و انتخاب یکی از ساختارهای معرفی‌شده، به ازای $\boldsymbol{\beta}^0$ محاسبه می‌کنیم.

گام ۳) برآورد ماتریس کواریانس را از رابطه زیر محاسبه می‌کنیم

$$\hat{\mathbf{V}}_i(\alpha) = \hat{\mathbf{A}}_i^{\frac{1}{2}} \mathbf{R}_i(\hat{\alpha}) \hat{\mathbf{A}}_i^{\frac{1}{2}}.$$

که در آن

$$\hat{\mathbf{A}}_i = \mathbf{A}_i(\hat{\phi})$$

گام ۴) مقدار $\boldsymbol{\beta}^{r+1}$ را با استفاده از رابطه زیر محاسبه می‌کنیم

$$\boldsymbol{\beta}^{r+1} = \boldsymbol{\beta}^r + \left(\sum_{i=1}^{n} \mathbf{X}_i \hat{\mathbf{V}}_i(\alpha) \mathbf{X}_i^T \right)^{-1} \left(\sum_{i=1}^{n} \mathbf{X}_i \hat{\mathbf{V}}_i(\alpha) (\mathbf{y}_i - \mathbf{X}_i^T \boldsymbol{\beta}^r) \right)$$

گام ۵) گام‌های (۲) تا (۴) را تا رسیدن به همگرایی $\left(|\boldsymbol{\beta}^{r+1} - \boldsymbol{\beta}^r| < \gamma \right)$ به ازای یک مقدار γ کوچک و از پیش تعیین‌شده، تکرار می‌کنیم.

[14] Newton–Raphson method

۲.۳.۱ مدل‌های انتقال

مدل انتقال، بسط مدل‌های خطی تعمیم‌یافته[15] (که در بخش ۵.۱ در مورد آن‌ها صحبت خواهیم کرد) برای توصیف توزیع شرطی بر روی هر پاسخ y_{ij} به‌عنوان تابعی معلوم از متغیرهای پاسخ قبلی $y_{i1}, y_{i2}, \ldots, y_{i(j-1)}$ و متغیر پیش‌بین می‌باشد. مدل‌های انتقال یا مدل‌های مارکوف، مسئله همبستگی موجود هر واحد نمونه را با مدل‌بندی پاسخ در زمان حال به شرط پاسخ در زمان‌های گذشته مورد توجه قرار می‌دهند. به عبارت دیگر در مدل‌های خطی تعمیم‌یافته، اگر مشاهدات گذشته برای واحد iام را با

$$H_{ij} = \left\{y_{i1}, y_{i2}, \ldots, y_{i(j-1)}\right\}$$

و امید شرطی مشاهده iام را با

$$\mu_{ij} = E(y_{ij}|H_{ij})$$

و واریانس شرطی آن را با

$$\nu = var(y_{ij}|H_{ij})$$

نمایش دهیم، یک مدل انتقال با سه مشخصه مشخص می‌شود به‌طوری که $y_{ij}|H_{ij}$ از توزیعی از خانواده توزیع‌های نمایی پیروی می‌کند

$$\begin{aligned} f(y_{ij}|H_{ij}) &= \exp\left\{\frac{y_{ij}\theta_{ij} - b(\theta_{ij})}{\varphi} + c(y_{ij}, \varphi)\right\} \\ \mu_{ij} &= \frac{db(\theta_{ij})}{d\theta_{ij}} \\ h(\mu_{ij}) &= \mathbf{x}_{ij}^T\boldsymbol{\beta} + \sum_{r=1}^{s} f_r(\mathbf{H}_{ij}; \boldsymbol{\alpha}) \\ \nu_{ij} &= \varphi\nu(\mu_{ij}) \end{aligned}$$

که در آن h، ν و f_r توابعی مشخص هستند و $\boldsymbol{\alpha}$ برداری از پارامترها است. رایج‌ترین مدل مورد استفاده در مدل‌های انتقال استفاده از زنجیره مارکوف[16] است که در آن پاسخ فعلی y_{ij} به شرط H_{ij} تنها به q پاسخ پیش از خود وابسته است که از آن به‌عنوان مرتبه مدل یاد می‌شود. به‌عنوان مثال در یک مدل انتقال با مرتبه $q = 1$ داریم

$$h(\mu_{ij}) = h(E(y_{ij}|y_{i(j-1)})) = \mathbf{x}_{ij}^T\boldsymbol{\beta} + \alpha(\mathbf{y}_{i(j-1)} - \mathbf{x}_{i(j-1)}^T\boldsymbol{\beta}).$$

۳.۳.۱ مدل‌های آمیخته خطی

در بسیاری از تحقیقات و مطالعات پزشکی، پزشک بیماران مورد نظر را در طول زمان به دفعات مورد معاینه قرار می‌دهد. داده‌ها و اطلاعات حاصل از معاینه که مکرر ثبت خواهند شد، مثالی

[15] Generalized linear models

[16] Markov chain

از داده‌های طولی است. یکی از رویکردهای تحلیل این نوع داده‌ها استفاده از مدل‌های آمیخته است. مدل‌های آمیخته ترکیبی از مدل با اثرات تصادفی[۱۷] و مدل با اثرات ثابت[۱۸] هستند. برای درک بهتر این مفهوم به مثال زیر از مک کولاک و همکاران (۲۰۰۸) توجه کنید.

فرض کنید در یک آزمایش بالینی[۱۹] در یک کلینیک نوعی دارو در ۴ دُز مختلف تهیه شده‌است که اثر دُز مصرفی بیماران تشنجی بر تعداد تشنج‌های آن‌ها مورد مطالعه است. در مدل زیر y_{ij} تعداد تشنج‌های jامین شخصی که iامین دُز را دریافت کرده، μ میانگین کل و α_i اثر دز iام روی متوسط تعداد تشنج‌ها است.

$$E[y_{ij}] = \mu + \alpha_i, \quad i = ۱, \ldots, ۴. \tag{۱۵.۱}$$

واضح است که α_iها اثرات ثابت هستند و هدف تفاوت بین دزهای مختلف و اثر آن‌ها روی کاهش تعداد تشنج‌های بیمار است، اما آنچه حائز اهمیت است این است که اگر مدلی ترکیب خطی از عامل‌هایی با اثرات ثابت باشد، یک مدل خطی با اثرات ثابت است.

حال فرض کنید همین آزمایش را در ۲۰ مرکز کلینیک یک شهر انجام دهیم و بیمار تنها دُز شماره ۱ را دریافت کند. در این صورت y_{ij} نشان دهنده تعداد تشنج‌های jامین شخصی است که در iامین کلینیک دارو را دریافت کرده و α_i اثر کلینیک iام روی متوسط تعداد تشنج‌ها است، لذا مدل مورد بررسی به صورت زیر است

$$E[y_{ij}] = \mu + a_i, \quad i = ۱, \ldots, ۲۰. \tag{۱۶.۱}$$

کلینیک iام کلینیکی است که شخص بیمار به‌طور تصادفی از بین تمامی ۲۰ کلینیک موجود در شهر برای مراجعه به آن انتخاب می‌کند؛ لذا مراجعه هر شخص به هر مرکز کلینیک به‌طور تصادفی رخ خواهد داد، پس a_iها اثرات تصادفی هستند.

در حالتی دیگر فرض کنید در هر کلینیک برخلاف حالت قبل هر ۴ دُز دارو موجود باشد و شخصی که به‌طور تصادفی به یکی از ۲۰ کلینیک مراجعه می‌کند از بین ۴ دُز مختلف دارو آنچه را که مربوط به خود است دریافت کند. در این صورت خواهیم داشت

$$E[y_{ijk}] = \mu + \alpha_i + \beta_j + c_{ij} \tag{۱۷.۱}$$

به‌طوری که y_{ijk} تعداد تشنج‌های kامین شخصی است که دُز jام را در کلینیک iام دریافت کرده‌است و α_i اثر کلینیک iام روی متوسط تعداد تشنج‌های مربوط به کلینیک iام و β_j اثر دز jام روی متوسط تعداد تشنج‌های مربوط به دز jام و c_{ij} اثر متقابل[۲۰] بین کلینیک iام و دُز jام است. اکنون با توجه به توضیحات قبل واضح است که α_iها اثرات تصادفی و β_jها اثرات ثابت هستند و چون c_{ij} تقابلی بین اثر تصادفی و ثابت است، پس c_{ij} نیز یک اثر تصادفی است.

[17] Random effects
[18] Fixed effects
[19] Clinical trial
[20] Interaction

بنابراین مدل (۱۷.۱) آمیزه‌ای از اثرات تصادفی و ثابت است و به آن مدل آمیخته می‌گویند. همچنین چون ترکیبی خطی از این اثرات است به آن مدل آمیخته خطی گویند.

از آن جایی که تمرکز این پایان‌نامه بر روی مدل‌های آمیخته خطی است لذا در بخش بعدی به طور جداگانه به معرفی این مدل خواهیم پرداخت.

۴.۱ مدل آمیخته خطی

مدل خطی $E[\mathbf{y}|\mathbf{X}] = \mathbf{X}\boldsymbol{\beta}$ را در نظر بگیرید که در آن $\mathbf{y} = (\mathbf{y}_1, \ldots, \mathbf{y}_n)^T$ بردار مؤلفه‌های پاسخ، $\mathbf{X} = (\mathbf{X}_1^T, \ldots, \mathbf{X}_n^T)^T$ ماتریس طرح در اندازه $N \times p$ و $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_p)^T$ بردار اثرات ثابت است. تنها تفاوت مدل آمیخته خطی با مدل خطی افزودن جمله $\mathbf{Zb}$ است که در آن $\mathbf{Z}$ همانند $\mathbf{X}$ یک ماتریس معلوم از متغیرهای پیش‌بین و $\mathbf{b} = (\mathbf{b}_1^T, \ldots, \mathbf{b}_n^T)^T$ بردار اثرات تصادفی در رخ داد متغیر پاسخ $\mathbf{y}$ است که در آن $\mathbf{b}_i$ها بردارهای $q \times 1$ هستند، لذا مدل آمیخته خطی را می‌توان به‌صورت زیر نوشت:

$$E[\mathbf{y}|\mathbf{X}, \mathbf{Z}] = \mathbf{X}\boldsymbol{\beta} + \mathbf{Zb}. \qquad (۱۸.۱)$$

در مطالعات طولی که هر مشاهده برای هر فرد به دفعات اندازه‌گیری می‌شود، شکل ماتریسی مدل آمیخته خطی با در نظر گرفتن جمله خطاها به‌صورت زیر است

$$\underset{N\times 1}{\begin{pmatrix}\mathbf{y}_1\\ \mathbf{y}_2\\ \vdots\\ \mathbf{y}_n\end{pmatrix}} = \underset{N\times p}{\begin{pmatrix}\mathbf{X}_1\\ \mathbf{X}_2\\ \vdots\\ \mathbf{X}_n\end{pmatrix}} \underset{p\times 1}{\begin{pmatrix}\beta_1\\ \beta_2\\ \vdots\\ \beta_p\end{pmatrix}} + \underset{N\times nq}{\begin{pmatrix}\mathbf{Z}_1^T & \circ & \cdots & \circ\\ \circ & \mathbf{Z}_2^T & \cdots & \circ\\ \vdots & \vdots & \ddots & \vdots\\ \circ & \circ & \cdots & \mathbf{Z}_n^T\end{pmatrix}} \underset{nq\times 1}{\begin{pmatrix}\mathbf{b}_1\\ \mathbf{b}_2\\ \vdots\\ \mathbf{b}_n\end{pmatrix}} + \underset{N\times 1}{\begin{pmatrix}\boldsymbol{\epsilon}_1\\ \boldsymbol{\epsilon}_2\\ \vdots\\ \boldsymbol{\epsilon}_n\end{pmatrix}}$$

که می‌توان به صورت زیر نمایش داد

$$\mathbf{y}_i = \mathbf{X}_i\boldsymbol{\beta} + \mathbf{Z}_i{}^T\mathbf{b}_i + \boldsymbol{\epsilon}_i, \qquad i = 1, \ldots, n, \qquad (۱۹.۱)$$

که در آن $\mathbf{y}_i = (y_{i1}, y_{i2}, \ldots, y_{in_i})^T$ برداری از n_i مشاهده برای فرد iام، $\mathbf{X}_i$ ماتریس طرح $n_i \times p$ از اثرات ثابت و $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_p)^T$ بردار $p \times 1$ ضرایب ثابت و نامعلوم، و $\mathbf{Z}_i$ ماتریس طرح $q \times n_i$ از اثرات تصادفی، $\mathbf{b}_i$ بردار $q \times 1$ ضرایب تصادفی و $\boldsymbol{\epsilon}_i$ بردار $n \times 1$ خطاهای تصادفی است. همچنین فرض می‌شود به ازای هر $i = 1, \ldots, n$ ، $\mathbf{b}_i$ از $\boldsymbol{\epsilon}_i$ مستقل بوده و

$$\begin{cases} E(\mathbf{b}_i) = \circ \quad , \quad E(\mathbf{b}_i\mathbf{b}_i^T) = \mathbf{D} \\ E(\boldsymbol{\epsilon}_i) = \circ \quad , \quad E(\boldsymbol{\epsilon}_i\boldsymbol{\epsilon}_i^T) = \sigma^2\mathbf{I}_{n_i} \end{cases}.$$

با فرض نرمال بودن توزیع $\mathbf{b}_i$ها و $\boldsymbol{\epsilon}_i$ها، توزیع متغیر پاسخ به‌صورت زیر نتیجه می‌شود

$$\mathbf{y} \sim \mathcal{N}_N(\mathbf{X}\boldsymbol{\beta}, \mathbf{V}) \qquad (۲۰.۱)$$

که در آن $\mathbf{y} = (\mathbf{y}_1^T, \ldots, \mathbf{y}_n^T)^T$ و $\mathbf{X} = (\mathbf{X}_1^T, \ldots, \mathbf{X}_n^T)^T$ و

$$\begin{aligned} \mathbf{V} &= var(\mathbf{y}) = var(\mathbf{X}\boldsymbol{\beta} + \mathbf{Z}^T\mathbf{b} + \boldsymbol{\epsilon}) \\ &= var(\mathbf{Z}^T\mathbf{b}) + var(\boldsymbol{\epsilon}) \\ &= \mathbf{Z}\mathbf{D}\mathbf{Z}^T + \sigma^2\mathbf{I}. \end{aligned} \tag{۲۱.۱}$$

۱.۴.۱ برآورد کلاسیک پارامترهای مدل

در روش ML، برآورد پارامترها با ماکزیمم کردن تابع درستنمایی مدل محاسبه می‌شود. در این بخش برآورد اثرات ثابت مدل (۱۹.۱) را به این روش محاسبه می‌کنیم. لازم به ذکر است در طول این پایان‌نامه تمرکز بر روی برآورد اثرات ثابت $\boldsymbol{\beta}$ می‌باشد. برای برآورد اثرات تصادفی $\mathbf{b}$، مک کولاک و همکاران (۲۰۰۸) را ببینید.

تحت مفروضات مدل (۲۰.۱) داریم

$$f(\mathbf{y}; \boldsymbol{\beta}; \mathbf{V}) = \frac{|\mathbf{V}|^{-\frac{1}{2}}}{(2\pi)^{\frac{N}{2}}} \exp\left\{-\frac{1}{2}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})^T\mathbf{V}^{-1}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})\right\}.$$

بنابراین لگاریتم تابع درستنمایی عبارتست از

$$l(\mathbf{y}) = -\frac{N}{2}\log(2\pi) - \frac{1}{2}\sum_{i=1}^{n}\log|\mathbf{V}| - \frac{1}{2}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})^T\mathbf{V}^{-1}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}). \tag{۲۲.۱}$$

واضح است که ماکزیمم کردن تابع فوق برحسب $\boldsymbol{\beta}$ معادل با مینیمم کردن عبارت وابسته به $\boldsymbol{\beta}$ است، در نتیجه

$$\hat{\boldsymbol{\beta}} = \underset{\boldsymbol{\beta}\in\mathbb{R}^p}{argmin}\left\{(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})^T\mathbf{V}^{-1}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})\right\}. \tag{۲۳.۱}$$

با مساوی صفر قرار دادن مشتق رابطه (۲۳.۱) بر حسب $\boldsymbol{\beta}$ خواهیم داشت

$$\begin{aligned} \circ &= \frac{d}{d\boldsymbol{\beta}}\left[(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})^T\mathbf{V}^{-1}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})\right] \\ &= \frac{d}{d\boldsymbol{\beta}}\left[\mathbf{y}^T\mathbf{V}^{-1}\mathbf{y} - \mathbf{y}^T\mathbf{V}^{-1}\mathbf{X}\boldsymbol{\beta} - (\mathbf{X}\boldsymbol{\beta})^T\mathbf{V}^{-1}\mathbf{y} + (\mathbf{X}\boldsymbol{\beta})^T\mathbf{V}^{-1}(\mathbf{X}\boldsymbol{\beta})\right] \\ &= \frac{d}{d\boldsymbol{\beta}}\left[\mathbf{y}^T\mathbf{V}^{-1}\mathbf{y} - 2\mathbf{y}^T\mathbf{V}^{-1}\mathbf{X}\boldsymbol{\beta} + (\mathbf{X}\boldsymbol{\beta})^T\mathbf{V}^{-1}(\mathbf{X}\boldsymbol{\beta})\right] \\ &= -2\mathbf{y}^T\mathbf{V}^{-1}\mathbf{X} + 2\boldsymbol{\beta}^T\mathbf{X}^T\mathbf{V}^{-1}\mathbf{X}. \end{aligned}$$

با شرط معلوم بودن $\mathbf{V}$ معکوس پذیر بودن $(\mathbf{X}^T\mathbf{V}^{-1}\mathbf{X})$ داریم

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^T\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}^T\mathbf{V}^{-1}\mathbf{y} \tag{۲۴.۱}$$

به همین ترتیب نتیجه می‌شود

$$E(\hat{\boldsymbol{\beta}}) = \boldsymbol{\beta} \tag{۲۵.۱}$$

و

$$\begin{aligned} var(\hat{\boldsymbol{\beta}}) &= var((\mathbf{X}^T\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}^T\mathbf{V}^{-1}\mathbf{Y}) \\ &= (\mathbf{X}^T\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}^T\mathbf{V}^{-1}\mathbf{X}(\mathbf{X}^T\mathbf{V}^{-1}\mathbf{X})^{-1} \\ &= (\mathbf{X}^T\mathbf{V}^{-1}\mathbf{X})^{-1}. \end{aligned} \tag{۲۶.۱}$$

واضح است که در رابطه (۲۴.۱) محاسبه مقدار $\mathbf{V}$، مستلزم برآورد مؤلفه‌های واریانس یعنی $\boldsymbol{\theta} = (\sigma, \mathbf{D})$ است. بدین منظور لیرد و ویر (۱۹۸۲) الگوریتم EM را برای برآورد پارامتر $\boldsymbol{\beta}$ پیشنهاد داد که در ادامه بیان خواهیم کرد.

الگوریتم ۲ الگوریتم EM در برآورد پارامتر مدل آمیخته خطی

گام ۱) قرار می‌دهیم $\hat{\boldsymbol{\theta}}^{(t)} = \boldsymbol{\theta}_0$، سپس مقدار آماره‌های $\hat{t}_1^{(t)}$ و $\hat{t}_2^{(t)}$ را برای مرحله tام به‌صورت زیر محاسبه می‌کنیم:

$$\begin{aligned}\hat{t}_1^{(t)} &= E\left(\sum_{i=1}^{n}\boldsymbol{\epsilon}_i^T\boldsymbol{\epsilon}_i \mid \mathbf{y}_i, \hat{\boldsymbol{\beta}}^{(t)}, \hat{\boldsymbol{\theta}}^{(t)}\right) \\ &= \sum_{i=1}^{n}\left[\boldsymbol{\epsilon}_i^T(\hat{\boldsymbol{\theta}}^{(t)})\boldsymbol{\epsilon}_i(\hat{\boldsymbol{\theta}}^{(t)}) + trVar\left(\boldsymbol{\epsilon}_i \mid \mathbf{y}_i, \hat{\boldsymbol{\beta}}^{(t)}, \hat{\boldsymbol{\theta}}^{(t)}\right)\right], \qquad (۲۷.۱)\end{aligned}$$

$$\begin{aligned}\hat{t}_2^{(t)} &= E\left(\sum_{i=1}^{n}\mathbf{b}_i\mathbf{b}_i^T \mid \mathbf{y}_i, \hat{\boldsymbol{\beta}}^{(t)}, \hat{\boldsymbol{\theta}}^{(t)}\right) \\ &= \sum_{i=1}^{n}\left[\mathbf{b}_i(\hat{\boldsymbol{\theta}}^{(t)})\mathbf{b}_i^T(\hat{\boldsymbol{\theta}}^{(t)}) + Var\left(\mathbf{b}_i \mid \mathbf{y}_i, \hat{\boldsymbol{\beta}}^{(t)}, \hat{\boldsymbol{\theta}}^{(t)}\right)\right]. \qquad (۲۸.۱)\end{aligned}$$

گام ۲) برای به دست آوردن $\hat{\boldsymbol{\theta}}^{(t+1)}$، مقدار $\hat{\sigma}^{2(t+1)}$ و $\hat{\mathbf{D}}^{(t+1)}$ را از رابطه زیر محاسبه می‌کنیم:

$$\hat{\sigma}^{2(t+1)} = \hat{t}_1^{(t)}/N,$$

$$\hat{\mathbf{D}}^{(t+1)} = \hat{t}_2^{(t)}/n.$$

که در آن $N = \sum_{i=1}^{n} n_i$ و n برابر تعداد واحدهای آزمایشی است. حال برای محاسبه $\hat{\mathbf{V}}^{(t+1)}$ داریم

$$\hat{\mathbf{V}}^{(t+1)} = \mathbf{Z}\hat{\mathbf{D}}^{(t+1)}\mathbf{Z}^T + \hat{\sigma}^{2(t+1)}\mathbf{I}. \qquad (۲۹.۱)$$

گام ۳) در نهایت مقدار $\hat{\beta}$ از رابطه زیر بدست می‌آید

$$\hat{\boldsymbol{\beta}}^{(t+1)} = (\mathbf{X}^T\hat{\mathbf{V}}^{-1(t+1)}\mathbf{X})^{-1}\mathbf{X}^T\hat{\mathbf{V}}^{-1(t+1)}\mathbf{y}. \qquad (۳۰.۱)$$

گام ۴) گام‌های (۱) تا (۳) را تا رسیدن به همگرایی $\left(|\hat{\boldsymbol{\beta}}^{(t+1)} - \hat{\boldsymbol{\beta}}^{(t)}| < \gamma\right)$ به ازای یک مقدار γ کوچک و از پیش تعیین‌شده، تکرار می‌کنیم.

در ادامه چگونگی محاسبه برآوردگر β را در قالب یک مثال شهودی با استفاده از نرم‌افزار R بررسی می‌کنیم.

۲.۴.۱ مثال کاربردی

در این بخش یک نمونه از داده‌های طولی موجود در بسته nlme در نرم‌افزار R به‌نام Orthodont را مورد بررسی قرار می‌دهیم. این داده‌ها توسط محققان دانشکده دندان پزشکی دانشگاه

کارولینای شمالی[۲۱] جمع‌آوری شده‌اند و شامل اطلاعات رشد ۲۷ کودک است (۱۶ پسر و ۱۱ دختر) که از سن ۸ تا ۱۴ سالگی هر دو سال یک‌بار فاصله بین مرکز غده هیپوفیز[۲۲] تا شکاف پریگوماکسیلار[۲۳] واقع در جمجمه انسان اندازه‌گیری شده‌است. ۸ سطر ابتدایی داده‌ها به‌صورت است

```
  distance age Subject  Sex
1     26.0   8     M01 Male
2     25.0  10     M01 Male
3     29.0  12     M01 Male
4     31.0  14     M01 Male
5     21.5   8     M02 Male
6     22.5  10     M02 Male
7     23.0  12     M02 Male
8     26.5  14     M02 Male
```

متغیر Subject در این مجموعه داده یک عامل برای شناسایی افراد است (مثلاً M01 یعنی پسر شماره ۱ و F01 یعنی دختر شماره ۱). لذا به‌عنوان اثر تصادفی در مدل در نظر گرفته می‌شود، زیرا هر کودک به‌طور تصادفی برای انجام آزمایش انتخاب شده‌است. لذا مدل آمیخته خطی را به داده‌ها برازش می‌دهیم. اما قبل از هر چیز ممکن است یک سوال مطرح شود که چرا باید اثرات تصادفی را در مدل وارد کنیم؟

در ادامه پاسخ این سوال و دلیل این امر را با برازش یک مدل خطی به داده‌ها بدون در نظر گرفتن اثر تصادفی و مقایسه نتایج آن با مدل آمیخته خطی بیان می‌کنیم. برای این منظور ابتدا مدل (۲.۱) و سپس مدل (۱۸.۱) را با استفاده از بسته nlme در R به داده‌ها برازش می‌دهیم. کدهای مربوطه در ضمیمه آمده‌است.

نتایج خروجی این مدل در جدول ۲.۱ خلاصه شده‌است.

[21] North Carolina

[22] Pituitary

[23] Pterygomaxillary fissure

جدول ۲.۱: برآورد ضرایب با استفاده از مدل خطی و آمیخته خطی

مدل خطی			
متغیر	برآورد	انحراف استاندارد	P-value
عرض از مبدأ	۱۶/۳۴۰۶	۱/۴۱۶۲	< ۰/۰۰۰۱
سن	۰/۷۸۴۴	۰/۱۲۶۲	< ۰/۰۰۰۱
جنسیت	۱/۰۳۲۱	۲/۲۱۸۸	۰/۶۴۳۰
اثر متقابل سن و جنسیت	−۰/۳۰۴۸	۰/۱۹۷۷	۰/۱۲۶۰
مدل آمیخته خطی			
متغیر	برآورد	انحراف استاندارد	P-value
عرض از مبدأ	۱۶/۳۴۰۶	۰/۹۸۱۳	< ۰/۰۰۰۱
سن	۰/۷۸۴۳	۰/۰۷۷۵	< ۰/۰۰۰۱
جنسیت	۱/۰۳۲۱	۱/۵۳۷۴	۰/۵۰۸۲
اثر متقابل سن و جنسیت	−۰/۳۰۴۸	۰/۱۲۱۴	۰/۰۱۴۱

با توجه به نتایج جدول ۲.۱ در خروجی دو مدل دیده می‌شود که برآورد پارامترهای هر دو مدل بسیار بهم نزدیک هستند اما آنچه حائز اهمیت است میزان خطای استاندارد برآورد پارامترها می‌باشد که در مدل آمیخته خطی کاهش قابل ملاحظه‌ای داشته‌است و چنانچه اثر تصادفی در مدل نادیده گرفته شود، اثر متقابل دو متغیر سن و جنسیت معنا دار نخواهد بود.

۵.۱ مدل آمیخته خطی تعمیم‌یافته

در بخش قبل دیدیم که در استنباط مبتنی بر درستنمایی در مدل خطی فرض بر این است که متغیر پاسخ از توزیع نرمال پیروی می‌کند، اما در برخی موارد که متغیر پاسخ دو سطحی[۲۴] یا شمارشی[۲۵] است استفاده از مدل‌های خطی امکان‌پذیر نیست چرا که متغیر پاسخ از توزیع نرمال تبعیت نمی‌کند. در این صورت اگر متغیر پاسخ دارای توزیعی غیر از توزیع نرمال باشد از توزیع‌های متعلق به خانواده توزیع‌های نمایی در مدل‌های خطی تعمیم‌یافته، استفاده می‌شود. همچنین در مواردی از جمله توزیع پواسن که اغلب برای داده‌های شمارشی به کار می‌رود، میانگین توزیع پاسخ باید نامنفی باشد. در این حالت استفاده از مدل‌های خطی ممکن است برآورد منفی را برای میانگین پاسخ در پی داشته باشد. لذا در چنین مواردی با استفاده از

[24] Binary
[25] Count

مدل‌های خطی تعمیم‌یافته تابعی از میانگین پاسخ در مدل‌سازی استفاده می‌شود. بنابراین به سبب تحلیل این دسته از داده‌ها، مدل‌های خطی تعمیم‌یافته به عنوان تعمیمی از مدل‌های خطی به کار گرفته شده که با استفاده از یک تابع پیوند، بین میانگین پاسخ‌ها و متغیرهای بیش‌بین ارتباط برقرار می‌شود.

۱.۵.۱ مؤلفه‌های مدل خطی تعمیم‌یافته

یک مدل خطی تعمیم‌یافته دارای سه مؤلفه به‌صورت زیر است:

مؤلفه تصادفی

این مؤلفه توزیع متغیر پاسخ را مشخص می‌کند که باید از خانواده توزیع‌های نمایی باشد. خانواده توزیع‌های نمایی یک مؤلفه مهم در مدل‌های خطی تعمیم‌یافته است که تابع چگالی این خانواده را می‌توان به‌صورت زیر بیان کرد

$$f(y;\gamma,\tau)=\exp\left\{(y\gamma-b(\gamma))/\tau^{۲}-c(y,\tau)\right\} \qquad (۳۱.۱)$$

که در آن $b(\cdot)$ و $c(\cdot)$ توابعی معلوم هستند، $b(\cdot)$ یک تابع انباشته[۲۶]، پارامتر γ یک پارامتر مکانی[۲۷] و τ را اغلب پارامتر پراکندگی[۲۸] می‌نامند. همچنین میانگین و واریانس خانواده توزیع‌های نمایی به‌صورت زیر محاسبه می‌شوند:

$$E(y)=\mu=\frac{db(\gamma)}{d\gamma} \qquad (۳۲.۱)$$

$$Var(y)=\frac{d^{۲}b(\gamma)}{d\gamma^{۲}}\tau^{۲}. \qquad (۳۳.۱)$$

از جمله توزیع‌های مربوط به خانواده توزیع‌های نمایی می‌توان به نرمال، پواسن، هندسی، نمایی، گاما، نرمال وارون[۲۹] و دوجمله‌ای منفی اشاره کرد.

به عنوان مثال، تابع احتمال توزیع پواسن را در نظر بگیرید

$$f(y;\mu)=\frac{e^{-\mu}\mu^{y}}{y!}.$$

این تابع احتمال را می‌توان به‌صورت خانواده توزیع‌های نمایی نوشت:

$$f(y;\mu)=\exp\{y\ln\mu-\mu-\ln y!\}$$

[26] Cumulant
[27] Location parameter
[28] Dispersion parameter
[29] Inverse normal

که در آن $\gamma = \ln \mu$، $b(\gamma) = e^{\gamma}$، $\tau^{۲} = ۱$ و $c(y;\tau) = \exp\{-\ln(y!)\}$ است. برای محاسبه میانگین و واریانس آن نیز با استفاده از روابط (۳۲.۱) و (۳۳.۱) داریم

$$E(y) = \mu = \frac{db(\gamma)}{d\gamma} = \frac{de^{\gamma}}{d\gamma} = e^{\gamma} = e^{\ln \mu} = \mu$$

$$Var(y) = \frac{d^{۲} b(\gamma)}{d\gamma^{۲}} \tau^{۲} = \frac{d^{۲} e^{\gamma}}{d\gamma^{۲}} = e^{\gamma} = e^{\ln \mu} = \mu.$$

مؤلفه اصلی

این مؤلفه دارای یک پیش‌گوی خطی[۳۰] است که متغیرهای پیش‌بین را شامل شده و به‌صورت زیر بیان می‌شود

$$\eta_i = \mathbf{x}_i^T \boldsymbol{\beta} = \beta_{۰} + \sum_{i=۱}^{p} \beta_i x_i.$$

که در آن $\mathbf{x}_i = (x_{۱}, \ldots, x_p)^T$.

تابع پیوند

این مؤلفه پیش‌گوی خطی را به میانگین متغیر پاسخ مربوط می‌کند. به‌طوری که

$$\eta_i = g\left(E(\mathbf{y}_i|\mathbf{X}_i)\right) = \mathbf{x}_i^T \boldsymbol{\beta}$$

که در آن $g(\cdot)$ را تابع پیوند[۳۱] گویند که یک تابع یکنوای مشتق‌پذیر و همچنین معکوس‌پذیر است. لذا می‌توان آن را به‌صورت زیر بیان کرد:

$$\mu_i = E(\mathbf{y}_i|\mathbf{x}_i) = g^{-۱}(\eta_i) = g^{-۱}(\mathbf{x}_i^T \boldsymbol{\beta}). \qquad (۳۴.۱)$$

تابع پیوند انواع گوناگونی دارد که از جمله آن‌ها می‌توان به تابع پیوند همانی[۳۲]، $g(p) = p$، تابع پیوند لگاریتمی، $g(p) = \log(p)$، تابع پیوند لُجیت[۳۳]، $g(p) = \log\left(\frac{p}{۱-p}\right)$، تابع پیوند پرابیت[۳۴]، $g(p) = \Phi^{-۱}(p)$ و تابع پیوند Cloglog[۳۵]، $g(p) = \log\left(-\log\left(۱-p\right)\right)$، اشاره کرد.

با انتخاب تابع پیوند g، الگوهای رگرسیون متفاوتی همچون رگرسیون لُجستیک، رگرسیون پرابیت و رگرسیون پواسن به عنوان یکی از اعضای رده مدل‌های خطی تعمیم‌یافته در مدل‌بندی متغیرهای پاسخ استفاده می‌شود. واضح است که با در نظر گرفتن تابع پیوند همانی، الگوی رگرسیون خطی استاندارد نیز یک مدل خطی تعمیم‌یافته است.

[30] Linear predictor
[31] Link function
[32] Identity
[33] Logit
[34] Probit
[35] Complementary Log-Log

یکی از فرضیات مدل‌های خطی تعمیم‌یافته استقلال بردار پاسخ است. چنان‌چه این فرض برقرار نباشد از مدل‌های خطی آمیخته تعمیم‌یافته استفاده می‌شود. به عبارت دقیق‌تر، می‌توان گفت که در مدل (۱۸.۱)، رفتار متغیر پاسخ، استفاده از مدل‌های تعمیم‌یافته را به جهت موارد گفته‌شده، ایجاب می‌کند. لذا یکی از روش‌های تحلیل آن دسته از مجموعه داده‌های طولی که بردار پاسخ آن‌ها دارای چنین رفتاری است، مدل آمیخته خطی تعمیم‌یافته خواهد بود. زیرا همبستگی بین مشاهدات هر واحد در بردار پاسخ داده‌های طولی وجود دارد. با این توضیحات واضح است که تفاوت مدل‌های خطی تعمیم‌یافته با مدل‌های آمیخته خطی تعمیم‌یافته در مؤلفه اصلی خواهد بود، به طوری که داریم

$$E(\mathbf{y}_i|\mathbf{b}_i) = \mu_i$$

$$\eta_i = g(\mu_i) = \mathbf{X}_i^T\boldsymbol{\beta} + \mathbf{Z}_i{}^T\mathbf{b}_i \qquad (۳۵.۱)$$

$$\begin{aligned} E(\mathbf{y}_i) &= E(E(\mathbf{y}_i|\mathbf{b}_i)) \\ &= E(\mu_i) \\ &= E(g^{-1}(\mathbf{X}_i^T\boldsymbol{\beta} + \mathbf{Z}_i{}^T\mathbf{b}_i)). \end{aligned}$$

۲.۵.۱ برآورد پارامترهای مدل

مدل خطی تعمیم‌یافته

مدل خطی زیر را در نظر بگیرید

$$E[\mathbf{y}|\mathbf{X}] = \mathbf{X}\boldsymbol{\beta}$$

که در آن $\mathbf{y} = (\mathbf{y}_1^T, \ldots, \mathbf{y}_n^T)^T$ بردار مؤلفه‌های پاسخ، $\mathbf{X} = (\mathbf{X}_1^T, \ldots, \mathbf{X}_n^T)^T$ ماتریس طرح در اندازه $N \times p$ و $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_p)^T$ بردار اثرات ثابت است.

با فرض این که توزیع متغیر پاسخ از خانواده توزیع‌های نمایی باشد، داریم

$$\mathbf{y}_i \overset{iid}{\sim} f_{\mathbf{y}_i}(\mathbf{y}_i)$$

که در آن

$$f_{\mathbf{y}_i}(\mathbf{y}_i) = \exp\left\{(\mathbf{y}_i\gamma_i - b(\gamma_i))/\tau^2 - c(\mathbf{y}_i, \tau)\right\} \qquad (۳۶.۱)$$

لذا لگاریتم تابع درستنمایی به صورت زیر است

$$\ell(\boldsymbol{\beta}) = \sum_{i=1}^{n} \left[\mathbf{y}_i\gamma_i - b(\gamma_i)\right]/\tau^2 - \sum_{i=1}^{n} c(\mathbf{y}_i, \tau). \qquad (۳۷.۱)$$

قبل از ماکزیمم کردن تابع (۳۷.۱)، توجه به موارد زیر که در ادامه از آن‌ها بهره می‌گیریم ضروری است.

$$E\left[\frac{\partial \log f_{\mathbf{y}_i}(\mathbf{y}_i)}{\partial \gamma_i}\right] = ۰ \tag{۳۸.۱}$$

و

$$Var\left(\frac{\partial \log f_{\mathbf{y}_i}(\mathbf{y}_i)}{\partial \gamma_i}\right) = -E\left[\frac{\partial^۲ \log f_{\mathbf{y}_i}(\mathbf{y}_i)}{\partial \gamma_i^۲}\right]. \tag{۳۹.۱}$$

تحت شرایط نظم (کسلا و برگر (۱۹۹۰))، با استفاده از رابطه‌های (۳۶.۱) و (۳۸.۱) داریم

$$E\left[\left\{\mathbf{y}_i - \frac{\partial b(\gamma_i)}{\partial \gamma_i}\right\}/\tau^۲\right] = ۰, \tag{۴۰.۱}$$

$$E\left[\mathbf{y}_i\right] = \mu_i = \frac{\partial b(\gamma_i)}{\partial \gamma_i}. \tag{۴۱.۱}$$

همچنین با استفاده از رابطه‌های (۳۶.۱) و (۳۹.۱) داریم

$$Var\left(\left\{\mathbf{y}_i - \frac{\partial b(\gamma_i)}{\partial \gamma_i}\right\}/\tau^۲\right) = -E\left[-\frac{۱}{\tau^۲}\frac{\partial^۲ b(\gamma_i)}{\partial \gamma_i^۲}\right], \tag{۴۲.۱}$$

حال با استفاده از رابطه (۴۱.۱) می‌توان نوشت

$$\begin{aligned} Var\left(\frac{\mathbf{y}_i - \boldsymbol{\mu}_i}{\tau^۲}\right) &= \frac{۱}{\tau^۲}\frac{\partial^۲ b(\gamma_i)}{\partial \gamma_i^۲} \\ Var(\mathbf{y}_i) &= \tau^۲ \frac{\partial^۲ b(\gamma_i)}{\partial \gamma_i^۲} \end{aligned} \tag{۴۳.۱}$$

با قرار دادن $\nu(\mu_i) = \partial^۲ b(\gamma_i)/\partial \gamma_i^۲$ داریم

$$Var(\mathbf{y}_i) = \tau^۲ \nu(\mu_i).$$

همچنین

$$\frac{\partial \gamma_i}{\partial \mu_i} = \left(\frac{\partial \mu_i}{\partial \gamma_i}\right)^{-۱} = \left(\frac{\partial^۲ b(\gamma_i)}{\partial \gamma_i^۲}\right)^{-۱} = \frac{۱}{\nu(\mu_i)} \tag{۴۴.۱}$$

و طبق رابطه (۳۴.۱) داریم

$$\begin{aligned} \frac{\partial \mu_i}{\partial \boldsymbol{\beta}} &= \frac{\partial \mu_i}{\partial g(\mu_i)}\frac{\partial g(\mu_i)}{\partial \boldsymbol{\beta}} = \left(\frac{\partial g(\mu_i)}{\partial \mu_i}\right)^{-۱}\frac{\partial \mathbf{X}_i^T \boldsymbol{\beta}}{\partial \boldsymbol{\beta}} \\ &= \left(\frac{\partial g(\mu_i)}{\partial \mu_i}\right)^{-۱} \mathbf{X}_i^T. \end{aligned} \tag{۴۵.۱}$$

حال برای ماکزیمم کردن تابع (۳۷.۱) برحست β داریم

$$\begin{aligned}\frac{\partial \ell(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}} &= \frac{1}{\tau^2}\sum\left[\mathbf{y}_i\frac{\partial\gamma_i}{\partial\boldsymbol{\beta}} - \frac{\partial b(\gamma_i)}{\partial\gamma_i}\frac{\partial\gamma_i}{\partial\boldsymbol{\beta}}\right]\\ &= \frac{1}{\tau^2}\sum(\mathbf{y}_i-\mu_i)\frac{\partial\gamma_i}{\partial\boldsymbol{\beta}}\\ &= \frac{1}{\tau^2}\sum(\mathbf{y}_i-\mu_i)\frac{\partial\gamma_i}{\partial\mu_i}\frac{\partial\mu_i}{\partial\boldsymbol{\beta}}\\ &= \frac{1}{\tau^2}\sum\frac{(\mathbf{y}_i-\mu_i)}{\nu(\mu_i)g_\mu(\mu_i)}\mathbf{X}_i^T\\ &= \sum(\mathbf{y}_i-\mu_i)w_i g_\mu(\mu_i)\mathbf{X}_i^T, \qquad (۴۶.۱)\end{aligned}$$

که در آن $w_i = \left[\tau^2\nu(\mu_i)g_\mu^2(\mu_i)\right]^{-1}$.

شکل ماتریسی رابطه (۴۶.۱) را می‌توان به صورت زیر نوشت

$$\frac{\partial\ell(\boldsymbol{\beta})}{\partial\boldsymbol{\beta}} = \mathbf{X}^T\mathbf{W}\boldsymbol{\Delta}(\mathbf{y}-\boldsymbol{\mu}). \qquad (۴۷.۱)$$

با مساوی صفر قرار دادن رابطه (۴۷.۱) نتیجه می‌شود

$$\mathbf{X}^T\mathbf{W}\boldsymbol{\Delta}\mathbf{y} = \mathbf{X}^T\mathbf{W}\boldsymbol{\Delta}\boldsymbol{\mu}, \qquad (۴۸.۱)$$

که در آن $\mathbf{W} = \{{}_n w_i\}$ و $\boldsymbol{\Delta} = \{{}_n g_\mu(\mu_i)\}$ و μ به پارامتر β وابسته می‌باشند.

برای حل معادله (۴۸.۱) بر حسب β، از روش کمترین توان‌های دوم موزون تکراری[۳۶] بهره می‌گیریم. در این روش برای بردار پارامتر مورد نظر $\boldsymbol{\theta}$ داریم

$$\boldsymbol{\theta}^{(m+1)} = \boldsymbol{\theta}^{(m)} + \mathbf{I}\left(\boldsymbol{\theta}^{(m)}\right)^{-1}\frac{\partial\ell}{\partial\boldsymbol{\theta}}\Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}^{(m)}}, \qquad (۴۹.۱)$$

که در آن (m) نشان دهنده تکرار mام و $\mathbf{I}(\boldsymbol{\theta})$ ماتریس اطلاع فیشر[۳۷] است که به صورت زیر محاسبه می‌شود

$$\mathbf{I}(\boldsymbol{\theta}) = -E\left[\frac{\partial^2\ell(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}^T}\right].$$

لذا با استفاده از رابطه (۴۷.۱) داریم

$$\frac{\partial^2\ell(\boldsymbol{\beta})}{\partial\boldsymbol{\beta}\partial\boldsymbol{\beta}^T} = -\mathbf{X}^T\mathbf{W}\boldsymbol{\Delta}\frac{\partial\mu}{\partial\boldsymbol{\beta}}^T + \mathbf{X}^T\frac{\partial\mathbf{W}\boldsymbol{\Delta}}{\partial\boldsymbol{\beta}}^T(\mathbf{y}-\boldsymbol{\mu}) \qquad (۵۰.۱)$$

در نتیجه

$$\begin{aligned}-E\left[\frac{\partial^2\ell(\boldsymbol{\beta})}{\partial\boldsymbol{\beta}\partial\boldsymbol{\beta}^T}\right] &= \mathbf{X}^T\mathbf{W}\boldsymbol{\Delta}\frac{\partial\mu}{\partial\boldsymbol{\beta}}^T + 0\\ &= \mathbf{X}^T\mathbf{W}\boldsymbol{\Delta}\boldsymbol{\Delta}^{-1}\mathbf{X}\\ &= \mathbf{X}^T\mathbf{W}\mathbf{X}. \qquad (۵۱.۱)\end{aligned}$$

[36] Iterrative weighted least squares method

[37] Fisher information matrix

اکنون طبق رابطه (۴۹.۱) و (۵۱.۱) برآوردگر خطی تعمیم‌یافته (GL) پارامتر $\boldsymbol{\beta}$ به صورت زیر حاصل می‌شود

$$\hat{\boldsymbol{\beta}}_{GL}^{(m+1)} = \hat{\boldsymbol{\beta}}_{GL}^{(m)} + (\mathbf{X}^T\mathbf{W}\mathbf{X})^{-1}\mathbf{X}^T\mathbf{W}\boldsymbol{\Delta}(\mathbf{y}-\boldsymbol{\mu}). \qquad (۵۲.۱)$$

مدل آمیخته خطی تعمیم‌یافته

مدل آمیخته خطی زیر را در نظر بگیرید

$$E[\mathbf{y}|\mathbf{X},\mathbf{Z}] = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{b}$$

که در آن $\mathbf{y} = (\mathbf{y}_1^T,\ldots,\mathbf{y}_n^T)^T$ بردار مؤلفه‌های پاسخ، $\mathbf{X} = (\mathbf{X}_1^T,\ldots,\mathbf{X}_n^T)^T$ ماتریس طرح در اندازه $N \times p$ و $\boldsymbol{\beta} = (\beta_1,\ldots,\beta_p)^T$ بردار اثرات ثابت و $\mathbf{Z}$ همانند $\mathbf{X}$ یک ماتریس معلوم از متغیرهای پیش‌بین و $\mathbf{b} = (\mathbf{b}_1^T,\ldots,\mathbf{b}_n^T)^T$ بردار اثرات تصادفی در رخ داد متغیر پاسخ $\mathbf{y}$ است.

با فرض این که توزیع متغیر پاسخ از خانواده توزیع‌های نمایی باشد، داریم

$$\mathbf{y}_i|\mathbf{b}_i \overset{iid}{\sim} f_{\mathbf{y}_i|\mathbf{b}}(\mathbf{y}_i|\mathbf{b}_i), \qquad (۵۳.۱)$$

که در آن

$$f_{\mathbf{y}_i|\mathbf{b}}(\mathbf{y}_i|\mathbf{b}_i) = \exp\left\{(\mathbf{y}_i\gamma_i - b(\gamma_i))/\tau^2 - c(\mathbf{y}_i,\tau)\right\}, \qquad (۵۴.۱)$$

به‌طوری که

$$\begin{aligned} E[\mathbf{y}_i|\mathbf{b}] &= \mu_i \\ g(\mu_i) &= \mathbf{X}_i^T\boldsymbol{\beta} + \mathbf{Z}_i^T\mathbf{b}_i \\ \mathbf{b} &\sim f_{\mathbf{b}}(\mathbf{b}|\mathbf{D}) \end{aligned}$$

لذا تابع درستنمایی و لگاریتم آن، به ترتیب به‌صورت زیر است

$$L(\boldsymbol{\beta}) = \int\left(\prod_i f_{\mathbf{y}_i|\mathbf{b}}(\mathbf{y}_i|\mathbf{b}_i)f_{\mathbf{b}}(\mathbf{b}_i)\right)d\mathbf{b}, \qquad (۵۵.۱)$$

$$\ell(\boldsymbol{\beta}) = \log\int f_{\mathbf{y}|\mathbf{b}}(\mathbf{y}|\mathbf{b})f_{\mathbf{b}}(\mathbf{b})d\mathbf{b} = \log f_{\mathbf{y}}(\mathbf{y}). \qquad (۵۶.۱)$$

با مشتق گیری از رابطه (۵۶.۱) نسبت به بردار اثرات ثابت $(\boldsymbol{\beta})$ داریم

$$\begin{aligned} \frac{\partial l(\boldsymbol{\beta})}{\partial\boldsymbol{\beta}} &= \frac{\partial}{\partial\boldsymbol{\beta}}\int f_{\mathbf{y}|\mathbf{b}}(\mathbf{y}|\mathbf{b})f_{\mathbf{b}}(\mathbf{b})\mathrm{d}\mathbf{b}\Big/f_{\mathbf{y}}(\mathbf{y}) \\ &= \int\left[\frac{\partial}{\partial\boldsymbol{\beta}}f_{\mathbf{y}|\mathbf{b}}(\mathbf{y}|\mathbf{b})\right]f_{\mathbf{b}}(\mathbf{b})\mathrm{d}\mathbf{b}\Big/f_{\mathbf{y}}(\mathbf{y}). \qquad (۵۷.۱) \end{aligned}$$

از آن‌جایی که $f_{\mathbf{b}}(\mathbf{b})$ به پارامتر $\boldsymbol{\beta}$ وابسته نیست، می‌توان به‌صورت زیر نوشت

$$\begin{aligned}\frac{\partial}{\partial\boldsymbol{\beta}}f_{\mathbf{y}|\mathbf{b}}(\mathbf{y}|\mathbf{b}) &= \left(\frac{1}{f_{\mathbf{y}|\mathbf{b}}(\mathbf{y}|\mathbf{b})}\frac{\partial f_{\mathbf{y}|\mathbf{b}}(\mathbf{y}|\mathbf{b})}{\partial\boldsymbol{\beta}}\right) f_{\mathbf{Y}|\mathbf{b}}(\mathbf{y}|\mathbf{b}) \\ &= \frac{\partial \log f_{\mathbf{y}|\mathbf{b}}(\mathbf{y}|\mathbf{b})}{\partial\boldsymbol{\beta}} f_{\mathbf{y}|\mathbf{b}}(\mathbf{y}|\mathbf{b}), \qquad (۵۸.۱)\end{aligned}$$

حال رابطه (۵۷.۱) را به‌صورت زیر بازنویسی می‌کنیم

$$\begin{aligned}\frac{\partial l(\boldsymbol{\beta})}{\partial\boldsymbol{\beta}} &= \int \frac{\partial \log f_{\mathbf{y}|\mathbf{b}}(\mathbf{y}|\mathbf{b})}{\partial\boldsymbol{\beta}} f_{\mathbf{y}|\mathbf{b}}(\mathbf{y}|\mathbf{b}) f_{\mathbf{b}}(\mathbf{b})\mathrm{d}\mathbf{b} \Big/ f_{\mathbf{y}}(\mathbf{y}) \\ &= \int \frac{\partial \log f_{\mathbf{y}|\mathbf{b}}(\mathbf{y}|\mathbf{b})}{\partial\boldsymbol{\beta}} f_{\mathbf{b}|\mathbf{y}}(\mathbf{b}|\mathbf{y})\mathrm{d}\mathbf{b}. \qquad (۵۹.۱)\end{aligned}$$

با استفاده از رابطه (۴۷.۱) داریم

$$\begin{aligned}\frac{\partial \ell(\boldsymbol{\beta})}{\partial\boldsymbol{\beta}} &= \int \mathbf{X}^T\mathbf{W}\boldsymbol{\Delta}(\mathbf{y}-\boldsymbol{\mu}) f_{\mathbf{b}|\mathbf{y}}(\mathbf{b}|\mathbf{y})\mathrm{d}\mathbf{b} \\ &= \mathbf{X}^T E_{\mathbf{b}|\mathbf{y}}\left[\mathbf{W}\boldsymbol{\Delta}|\mathbf{y}\right]\mathbf{y} - \mathbf{X}^T E_{\mathbf{b}|\mathbf{y}}\left[\mathbf{W}\boldsymbol{\Delta}\boldsymbol{\mu}|\mathbf{y}\right] \qquad (۶۰.۱)\end{aligned}$$

با مساوی صفر قرار دادن رابطه (۶۰.۱) داریم

$$\mathbf{X}^T E_{\mathbf{b}|\mathbf{y}}\left[\mathbf{W}\boldsymbol{\Delta}|\mathbf{y}\right]\mathbf{y} = \mathbf{X}^T E_{\mathbf{b}|\mathbf{y}}\left[\mathbf{W}\boldsymbol{\Delta}\boldsymbol{\mu}|\mathbf{y}\right] \qquad (۶۱.۱)$$

در نتیجه طبق رابطه (۴۹.۱) برآوردگر آمیخته خطی تعمیم‌یافته (GLM) پارامتر $\boldsymbol{\beta}$ به صورت زیر حاصل می‌شود

$$\hat{\boldsymbol{\beta}}_{GLM}^{(m+1)} = \hat{\boldsymbol{\beta}}_{GLM}^{(m)} + (\mathbf{X}^T E_{\mathbf{b}|\mathbf{y}}[\mathbf{W}|\mathbf{y}]\mathbf{X})^{-1}\mathbf{X}^T E_{\mathbf{b}|\mathbf{y}}[\mathbf{W}\boldsymbol{\Delta}(\mathbf{y}-\boldsymbol{\mu})|\mathbf{y}]. \qquad (۶۲.۱)$$

برای آگاهی بیشتر در مورد برآورد پارامترهای مدل آمیخته خطی تعمیم‌یافته به مک کولاک و همکاران (۲۰۰۸) مراجعه کنید.

فصل ۲

رگرسیون ریج برای داده‌های طولی

۱.۲ مقدمه

همان‌طور که در فصل قبل گفته شد، یکی از روش‌های مدل‌بندی داده‌های طولی استفاده از مدل‌های آمیخته خطی است. الیوت و همکاران (۲۰۱۱) برای بهبود بخشیدن به برآورد پارامترهای مدل آمیخته خطی، رگرسیون آمیخته ریج را مطرح کردند.
در این فصل علاوه‌بر شرح مدل آن‌ها، برآورد ضرایب رگرسیونی این مدل را با خطاهای بیضی‌گون بررسی کرده‌ایم. همچنین به جهت بهبود خطای برآوردگر مدل رگرسیون آمیخته ریج، برآوردگرهای انقباضی نوع استاین را در این مدل ارائه کرده و نتایج شبیه‌سازی و تحلیل مثال واقعی را گزارش می‌کنیم.
انتهای این فصل در بر گیرنده تعمیم نتایج در مدل‌های خطی تعمیم‌یافته است.

۲.۲ مدل آمیخته خطی ریج

مدل رگرسیون خطی زیر را در نظر بگیرید

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon} \qquad (۱.۲)$$

که در آن $\mathbf{y} = (y_1, \ldots, y_n)^T$ یک بردار n مؤلفه‌ای شامل متغیر پاسخ، y_i نشان دهنده iامین پاسخ مشاهده شده، $\mathbf{X} = (\mathbf{x}_1, \ldots, \mathbf{x}_n)^T$ ماتریس غیر تصادفی در اندازه $n \times p$ از متغیرهای پیش‌بین که در آن $\mathbf{x}_i = (x_{i1}, \ldots, x_{ip})^T \in \mathbb{R}^p$ و x_{ij} نشان دهنده iامین مشاهده در jامین ستون متغیر پیش‌بین است. فرض می‌کنیم ماتریس $\mathbf{X}$ دارای رتبه کامل ستونی است تا ماتریس $\mathbf{X}^T\mathbf{X}$ که بعد آن p بوده، معکوس پذیر باشد. لازم به ذکر است در این حالت $\mathbf{X}^T\mathbf{X}$ دارای p ستون مستقل خطی است و لذا معکوس پذیر است. بردار ضرایب رگرسیونی $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_p)^T$ پارامترهای مدل و $\boldsymbol{\epsilon} = (\epsilon_1, \ldots, \epsilon_n)^T$ بردار n مؤلفه‌ای خطای تصادفی است. در ابتدا فرض می‌کنیم

$$E(\boldsymbol{\epsilon}) = 0, \qquad E(\boldsymbol{\epsilon}\boldsymbol{\epsilon}^T) = \sigma^2\mathbf{I}$$

که $\sigma^2 \in \mathbb{R}^+$ مجهول است.

یکی از مسائلی که می‌تواند استفاده از برآوردگر کمترین توان‌های دوم در مدل رگرسیونی (۱.۲) را با مشکل مواجه کند، همبستگی چندگانه بین متغیرهای رگرسیونی است. چراکه این امر سبب پیدایش مشکل هم‌خطی در داده‌ها می‌شود.

اگر یکی از متغیرهای پیش‌بین یک تابع دقیق خطی از یک یا چند متغیر پیش‌بین دیگر باشد، گوییم رگرسیون دارای همخطی کامل[1] است و اگر این رابطه‌ی خطی، به‌طور تقریبی باشد همخطی ناقص[2] اتفاق می‌افتد. در صورت وجود همخطی کامل دترمینان ماتریس $\mathbf{X}^T\mathbf{X}$ صفر است و لذا معکوس آن وجود ندارد. اگرچه می‌توان از معکوس تعمیم‌یافته $\mathbf{X}^T\mathbf{X}$ استفاده کرد. لازم به ذکر است که معکوس تعمیم‌یافته یکتا نیست و لذا کارایی ندارد. از طرفی اگر معکوس تعمیم‌یافته مور–پنروز[3] که یکتا است نیز استفاده کنیم استنباط در خصوص برآوردگرهای حاصل به راحتی امکان پذیر نخواهد بود. وجود همخطی ناقص باعث می‌شود دترمینان ماتریس $\mathbf{X}^T\mathbf{X}$ بسیار کوچک شود و در نتیجه $(\mathbf{X}^T\mathbf{X})^{-1}$ دارای مولفه‌هایی با قدر مطلق بسیار بزرگ است که در این صورت دقت برآورد ضرایب رگرسیونی کم می‌شود و رفتار ناپایداری دارد. جهت برخورد با این مشکل استفاده از رگرسیون ریج[4] که توسط هورل و کنارد (۱۹۷۰) معرفی شد، می‌تواند راهکار مناسبی باشد. برای آگاهی بیشتر در خصوص مشکلات همخطی و چگونگی مواجه با آن به روزبه (۱۳۹۰) مراجعه کنید.

در رگرسیون ریج، پارامتر جریمه[5] λ در یک تابع جریمه درجه دوم از $\boldsymbol{\beta}$ ضرب شده و برآوردگر ریج با مینیمم کردن مجموع توان‌های دوم جریمه شده به صورت زیر محاسبه می‌شود

$$\widehat{\boldsymbol{\beta}}_{Ridge} = \arg\min_{\boldsymbol{\beta}\in\mathbb{R}^p} \{(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})^T(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) + \lambda\boldsymbol{\beta}^T\boldsymbol{\beta}\} \qquad (۲.۲)$$

[1] Exact collinearity
[2] Partial collinearity
[3] Moore–Penrose
[4] Ridge regression
[5] Penalty

با مشتق‌گیری از مجموع توان‌های دوم جریمه شده نسبت به پارامتر $\boldsymbol{\beta}$ داریم

$$\begin{aligned}
&\frac{d}{d\boldsymbol{\beta}}\left[(\mathbf{y}-\mathbf{X}\boldsymbol{\beta})^T(\mathbf{y}-\mathbf{X}\boldsymbol{\beta})+\lambda\boldsymbol{\beta}^T\boldsymbol{\beta}\right]\\
&=\frac{d}{d\boldsymbol{\beta}}\left[\mathbf{y}^T\mathbf{y}-\mathbf{y}^T\mathbf{X}\boldsymbol{\beta}-(\mathbf{X}\boldsymbol{\beta})^T\mathbf{y}+(\mathbf{X}\boldsymbol{\beta})^T(\mathbf{X}\boldsymbol{\beta})+\lambda\boldsymbol{\beta}^T\boldsymbol{\beta}\right]\\
&=\frac{d}{d\boldsymbol{\beta}}\left[\mathbf{y}^T\mathbf{y}-۲\mathbf{y}^T\mathbf{X}\boldsymbol{\beta}+(\boldsymbol{\beta}^T\mathbf{X}^T\mathbf{X}\boldsymbol{\beta})+\lambda\boldsymbol{\beta}^T\boldsymbol{\beta}\right]\\
&=-۲\mathbf{y}^T\mathbf{X}+۲\boldsymbol{\beta}^T\mathbf{X}^T\mathbf{X}+۲\lambda\boldsymbol{\beta}^T,
\end{aligned}$$

که با مساوی صفر قرار دادن رابطه فوق، برآوردگر ریج به صورت زیر حاصل می‌شود

$$\widehat{\boldsymbol{\beta}}_{Ridge}=(\mathbf{X}^T\mathbf{X}+\lambda\mathbf{I}_p)^{-۱}\cdot\mathbf{X}^T\mathbf{y}$$

امید ریاضی برآوردگر ریج به صورت زیر است

$$\begin{aligned}
E\left(\widehat{\boldsymbol{\beta}}_{Ridge}\right) &= E\left((\mathbf{X}^T\mathbf{X}+\lambda\mathbf{I}_p)^{-۱}\mathbf{X}^T\mathbf{y}\right)\\
&= (\mathbf{X}^T\mathbf{X}+\lambda\mathbf{I}_p)^{-۱}\mathbf{X}^T\mathbf{X}\boldsymbol{\beta}. \qquad (۳.۲)
\end{aligned}$$

لذا $\widehat{\boldsymbol{\beta}}_{Ridge}$ برآوردگری اریب برای $\boldsymbol{\beta}$ است. به همین ترتیب نتیجه می‌شود

$$\mathrm{Var}\left(\widehat{\boldsymbol{\beta}}_{Ridge}\right)=\sigma^۲(\mathbf{X}^T\mathbf{X}+\lambda\mathbf{I}_p)^{-۱}\mathbf{X}^T\mathbf{X}(\mathbf{X}^T\mathbf{X}+\lambda\mathbf{I}_p)^{-۱}.$$

همچنین میانگین توان‌های دوم خطای برآوردگر ریج عبارت است از

$$\begin{aligned}
\mathrm{MSE}\left(\widehat{\boldsymbol{\beta}}_{Ridge}\right) &= E\left(\left(\widehat{\boldsymbol{\beta}}_{Ridge}-\boldsymbol{\beta}\right)^T\left(\widehat{\boldsymbol{\beta}}_{Ridge}-\boldsymbol{\beta}\right)\right)\\
&= \mathrm{tr}\left(\mathrm{Var}\left(\widehat{\boldsymbol{\beta}}_{Ridge}\right)\right)+Bias^T\left(\widehat{\boldsymbol{\beta}}_{Ridge}\right)Bias\left(\widehat{\boldsymbol{\beta}}_{Ridge}\right)\\
&= \sigma^۲\,\mathrm{tr}\left((\mathbf{X}^T\mathbf{X}+\lambda\mathbf{I}_p)^{-۱}\mathbf{X}^T\mathbf{X}(\mathbf{X}^T\mathbf{X}+\lambda\mathbf{I}_p)^{-۱}\right)+\lambda^۲\boldsymbol{\beta}^T(\mathbf{X}^T\mathbf{X}+\lambda\mathbf{I})^{-۲}\boldsymbol{\beta}.
\end{aligned}$$
(۴.۲)

اولین جمله سمت راست (۴.۲) مجموع واریانس‌های مولفه‌های پارامتر $\widehat{\boldsymbol{\beta}}_{Ridge}$ و دومین عبارت، توان دوم اریبی (ترانهاده اریبی در خودش) است. بنابراین اریبی $\widehat{\boldsymbol{\beta}}_{Ridge}$ با افزایش λ، افزایش و واریانس آن با افزایش λ، کاهش می‌یابد و برعکس. در رگرسیون ریج، پارامتر λ طوری انتخاب می‌شود که کاهش در واریانس بیشتر از افزایش توان دوم اریبی باشد.
مقدار بهینه‌ی λ به روش اعتبارسنجی متقابل[۶] توسط کراون و واهبا (۱۹۷۹) از رابطه زیر به‌دست می‌آید

$$\widehat{\lambda}=\arg\min_{\lambda}\left\{n^{-۱}(\mathbf{y}-\widehat{\mathbf{y}})^T(\mathbf{y}-\widehat{\mathbf{y}})/(۱-\mathrm{tr}(\mathbf{S})/n)^۲\right\}, \qquad (۵.۲)$$

[6] Cross validation

که در آن $\mathbf{S} = \mathbf{X}(\mathbf{X}^T\mathbf{X} + \lambda\mathbf{I})^{-1}\mathbf{X}^T$ و $\widehat{\mathbf{y}} = \mathbf{X}\widehat{\boldsymbol{\beta}}_{Ridge}$.
همچنین جنتل و همکاران (۲۰۰۴) برای مقادیر کوچک $\mathrm{tr}(\mathbf{S})/n$ محاسبه مقدار تقریبی λی بهینه را از رابطه زیر پیشنهاد کردند

$$\widehat{\lambda} \approx \arg\min_{\lambda} \left\{ n^{-1}(\mathbf{y}-\widehat{\mathbf{y}})^T(\mathbf{y}-\widehat{\mathbf{y}}) + 2n^{-2}\,\mathrm{tr}(\mathbf{S})(\mathbf{y}-\widehat{\mathbf{y}})^T(\mathbf{y}-\widehat{\mathbf{y}}) \right\}.$$

لازم به ذکر است که برآوردگرهای λ محدود به دو حالت فوق نشده و می‌توان انواع دیگر آن را در روزبه و همکاران (۲۰۱۵) دید. از طرفی توجه به این نکته لازم است که برآورد ریج یک تابع غیر خطی نسبت به λ است و این خود موجب می‌شود که نتوان به راحتی آن را برآورد کرد.

روش رگرسیون ریج در برخورد با داده‌های دارای مشکل همخطی در مطالعات طولی نیز می‌تواند کارساز باشد. الیوت و همکاران (۲۰۱۱) روش رگرسیون آمیخته ریج[7] (MR) را برای مرتفع کردن این مشکل در داده‌های طولی مطرح کردند.

مدل آمیخته خطی زیر را در نظر بگیرید

$$\mathbf{y}_i = \mathbf{X}_i\boldsymbol{\beta} + \mathbf{Z}_i^T\mathbf{b}_i + \boldsymbol{\epsilon}_i, \tag{۶.۲}$$

که در آن $\mathbf{y}_i = (y_{i1}, y_{i2}, \ldots, y_{in_i})^T$ برداری از n_i مشاهده، $\mathbf{X}_i$ ماتریس طرح $n_i \times p$ از اثرات ثابت و $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_p)^T$ بردار $p\times 1$ ضرایب ثابت و نامعلوم، و $\mathbf{Z}_i$ ماتریس طرح $q\times n_i$ از اثرات تصادفی، $\mathbf{b}_i \sim \mathcal{N}_q(0, \mathbf{D})$ بردار $q \times 1$ ضرایب تصادفی و $\boldsymbol{\epsilon}_i \sim \mathcal{N}_n(0, \sigma^2\mathbf{I})$ بردار $n \times 1$ خطاهای تصادفی است. برای اعمال رگرسیون آمیخته ریج، تابع درستنمایی جریمه شده را می‌توان به صورت زیر نوشت

$$f(\boldsymbol{\beta}, \mathbf{V}; \mathbf{y}) = \frac{|\mathbf{V}|^{-\frac{1}{2}}}{(2\pi)^{\frac{N}{2}}} \exp\left\{ -\frac{1}{2}(\mathbf{y}-\mathbf{X}\boldsymbol{\beta})^T\mathbf{V}^{-1}(\mathbf{y}-\mathbf{X}\boldsymbol{\beta}) - \lambda\boldsymbol{\beta}^T\boldsymbol{\beta} \right\}.$$

در نتیجه لگاریتم تابع درستنمایی عبارتست از

$$\ell^{pen}(\boldsymbol{\beta}, \mathbf{V}; \mathbf{y}) = -\frac{N}{2}\log(2\pi) - \frac{1}{2}\sum_{i=1}^{n}\log|\mathbf{V}| - \frac{1}{2}(\mathbf{y}-\mathbf{X}\boldsymbol{\beta})^T\mathbf{V}^{-1}(\mathbf{y}-\mathbf{X}\boldsymbol{\beta}) - \lambda\boldsymbol{\beta}^T\boldsymbol{\beta}.$$

برای دستیابی به برآوردگر ماکزیمم درستنمایی عبارت وابسته به $\boldsymbol{\beta}$ را در تابع $\ell^{pen}(\boldsymbol{\beta}, \mathbf{V}; \mathbf{y})$ می‌نیمم می‌کنیم

$$\widehat{\boldsymbol{\beta}}_{MR} = \arg\min_{\boldsymbol{\beta}\in\mathbb{R}^p}\{(\mathbf{y}-\mathbf{X}\boldsymbol{\beta})^T\mathbf{V}^{-1}(\mathbf{y}-\mathbf{X}\boldsymbol{\beta}) + \lambda\boldsymbol{\beta}^T\boldsymbol{\beta}\}. \tag{۷.۲}$$

با مساوی صفر قرار دادن مشتق عبارت داخل آکولاد در رابطه‌ی (۷.۲) بر حسب $\boldsymbol{\beta}$ خواهیم

[7] Mixed ridge

داشت

$$\begin{aligned}\circ &= \frac{d}{d\boldsymbol{\beta}}\left[(\mathbf{y}-\mathbf{X}\boldsymbol{\beta})^T\mathbf{V}^{-1}(\mathbf{y}-\mathbf{X}\boldsymbol{\beta})+\lambda\boldsymbol{\beta}^T\boldsymbol{\beta}\right]\\ &= \frac{d}{d\boldsymbol{\beta}}\left[\mathbf{y}^T\mathbf{V}^{-1}\mathbf{y}-\mathbf{y}^T\mathbf{V}^{-1}\mathbf{X}\boldsymbol{\beta}-(\mathbf{X}\boldsymbol{\beta})^T\mathbf{V}^{-1}\mathbf{y}+(\mathbf{X}\boldsymbol{\beta})^T\mathbf{V}^{-1}(\mathbf{X}\boldsymbol{\beta})+\lambda\boldsymbol{\beta}^T\boldsymbol{\beta}\right]\\ &= \frac{d}{d\boldsymbol{\beta}}\left[\mathbf{y}^T\mathbf{V}^{-1}\mathbf{y}-2\boldsymbol{\beta}^T\mathbf{X}^T\mathbf{V}^{-1}\mathbf{y}+(\mathbf{X}\boldsymbol{\beta})^T\mathbf{V}^{-1}(\mathbf{X}\boldsymbol{\beta})+\lambda\boldsymbol{\beta}^T\boldsymbol{\beta}\right]\\ &= -2\mathbf{X}^T\mathbf{V}^{-1}\mathbf{y}+2\mathbf{X}^T\mathbf{V}^{-1}\mathbf{X}\boldsymbol{\beta}+2\lambda\boldsymbol{\beta}. \qquad (۸.۲)\end{aligned}$$

در نتیجه برآوردگر MR به شرط معلوم بودن $\mathbf{V}$ به صورت

$$\widehat{\boldsymbol{\beta}}_{MR}=(\mathbf{X}^T\mathbf{V}^{-1}\mathbf{X}+\lambda\mathbf{I}_p)^{-1}\mathbf{X}^T\mathbf{V}^{-1}\mathbf{y} \qquad (۹.۲)$$

حاصل می‌شود.

امیدریاضی برآوردگر MR عبارت است از

$$\begin{aligned}E\left(\widehat{\boldsymbol{\beta}}_{MR}\right) &= E\left((\mathbf{X}^T\mathbf{V}^{-1}\mathbf{X}+\lambda\mathbf{I}_p)^{-1}\mathbf{X}^T\mathbf{V}^{-1}\mathbf{y}\right)\\ &= (\mathbf{X}^T\mathbf{V}^{-1}\mathbf{X}+\lambda\mathbf{I}_p)^{-1}\mathbf{X}^T\mathbf{V}^{-1}\mathbf{X}\boldsymbol{\beta}\end{aligned}$$

همچنین واریانس آن برابر

$$\mathrm{Var}\left(\widehat{\boldsymbol{\beta}}_{MR}\right)=(\mathbf{X}^T\mathbf{V}^{-1}\mathbf{X}+\lambda\mathbf{I})^{-1}\mathbf{X}^T\mathbf{V}^{-1}\mathbf{X}(\mathbf{X}^T\mathbf{V}^{-1}\mathbf{X}+\lambda\mathbf{I})^{-1}$$

است. مقدار بهینه λ در برآوردگر MR نیز از معادله (۵.۲) با تعریف

$$\mathbf{S}=\mathbf{X}(\mathbf{X}^T\mathbf{V}^{-1}\mathbf{X}+\lambda\mathbf{I})^{-1}\mathbf{X}^T\mathbf{V}^{-1}+\mathbf{Z}\mathbf{D}\mathbf{Z}^T\left(\mathbf{I}-\mathbf{X}\left[(\mathbf{X}^T\mathbf{V}^{-1}\mathbf{X}+\lambda\mathbf{I})^{-1}\mathbf{X}^T\mathbf{V}^{-1}\right]\right)$$

محاسبه می‌شود.

همان‌طور که در فصل اول گفته شد محاسبه مقدار $\mathbf{V}$ در رابطه‌ی (۹.۲) مستلزم برآورد پارامتر $\boldsymbol{\theta}=(\sigma,\mathbf{D})$ است. بدین منظور الیوت و همکاران (۲۰۱۱) الگوریتم مطرح شده توسط لیرد و ویر (۱۹۸۲) را برای $\widehat{\beta}_{MR}$ تعمیم دادند که در ادامه بیان خواهیم کرد.

الگوریتم ۳ الگوریتم EM در برآورد پارامتر مدل آمیخته خطی ریج

گام ۱) قرار می‌دهیم $\hat{\boldsymbol{\theta}}^{(t)} = \boldsymbol{\theta}_0$ و $\hat{\lambda}^{(t)} = \lambda_0$ سپس مقدار آماره‌های $\hat{t}_1^{(t)}$ و $\hat{t}_2^{(t)}$ را برای مرحله tام به‌صورت زیر محاسبه می‌کنیم:

$$\begin{aligned}\hat{t}_1^{(t)} &= E\left(\sum_{i=1}^{n}\boldsymbol{\epsilon}_i^T\boldsymbol{\epsilon}_i \mid \mathbf{y}_i, \hat{\boldsymbol{\beta}}^{(t)}, \hat{\boldsymbol{\theta}}^{(t)}\right)\\ &= \sum_{i=1}^{n}\left[\boldsymbol{\epsilon}_i^T(\hat{\boldsymbol{\theta}}^{(t)})\boldsymbol{\epsilon}_i(\hat{\boldsymbol{\theta}}^{(t)}) + trVar\left(\boldsymbol{\epsilon}_i \mid \mathbf{y}_i, \hat{\boldsymbol{\beta}}^{(t)}, \hat{\boldsymbol{\theta}}^{(t)}\right)\right], \qquad (۱۰.۲)\end{aligned}$$

$$\begin{aligned}\hat{t}_2^{(t)} &= E\left(\sum_{i=1}^{n}\mathbf{b}_i\mathbf{b}_i^T \mid \mathbf{y}_i, \hat{\boldsymbol{\beta}}^{(t)}, \hat{\boldsymbol{\theta}}^{(t)}\right)\\ &= \sum_{i=1}^{n}\left[\mathbf{b}_i(\hat{\boldsymbol{\theta}}^{(t)})\mathbf{b}_i^T(\hat{\boldsymbol{\theta}}^{(t)}) + Var\left(\mathbf{b}_i \mid \mathbf{y}_i, \hat{\boldsymbol{\beta}}^{(t)}, \hat{\boldsymbol{\theta}}^{(t)}\right)\right]. \qquad (۱۱.۲)\end{aligned}$$

گام ۲) برای به دست آوردن $\hat{\boldsymbol{\theta}}^{(t+1)}$، مقدار $\hat{\sigma}^{2(t+1)}$ و $\hat{\mathbf{D}}^{(t+1)}$ را از رابطه زیر محاسبه می‌کنیم:

$$\hat{\sigma}^{2(t+1)} = \hat{t}_1^{(t)}/N,$$

$$\hat{\mathbf{D}}^{(t+1)} = \hat{t}_2^{(t)}/n.$$

که در آن $N = \sum_{i=1}^{n} n_i$ و n برابر تعداد واحدهای آزمایشی است. حال برای محاسبه $\hat{\mathbf{V}}^{(t+1)}$ داریم

$$\hat{\mathbf{V}}^{(t+1)} = \mathbf{Z}\hat{\mathbf{D}}^{(t+1)}\mathbf{Z}^T + \hat{\sigma}^{2(t+1)}\mathbf{I}. \qquad (۱۲.۲)$$

گام ۳) ماتریس $\hat{\mathbf{S}}^{(t+1)}$ را از رابطه زیر محاسبه می‌کنیم

$$\begin{aligned}\hat{\mathbf{S}}^{(t+1)} &= \mathbf{X}(\mathbf{X}^T\hat{\mathbf{V}}^{-1(t+1)}\mathbf{X} + \hat{\lambda}^{(t)}\mathbf{I})^{-1}\mathbf{X}^T\hat{\mathbf{V}}^{-1(t+1)}\\ &+ \mathbf{Z}\hat{\mathbf{D}}^{(t+1)}\mathbf{Z}^T\left(\mathbf{I} - \mathbf{X}\left[(\mathbf{X}^T\hat{\mathbf{V}}^{-1(t+1)}\mathbf{X} + \hat{\lambda}^{(t)}\mathbf{I})^{-1}\mathbf{X}^T\hat{\mathbf{V}}^{-1(t+1)}\right]\right), \qquad (۱۳.۲)\end{aligned}$$

سپس برای محاسبه $\hat{\lambda}^{(t+1)}$ داریم

$$\hat{\lambda}^{(t+1)} = \arg\min_{\lambda}\left\{n^{-1}(\mathbf{y} - \widehat{\mathbf{y}})^T(\mathbf{y} - \widehat{\mathbf{y}})/(1 - \mathrm{tr}(\hat{\mathbf{S}}^{(t+1)})/n)^2\right\} \qquad (۱۴.۲)$$

گام ۴) در نهایت مقدار $\hat{\boldsymbol{\beta}}_{MR}^{(t+1)}$ از رابطه زیر بدست می‌آید

$$\hat{\boldsymbol{\beta}}_{MR}^{(t+1)} = (\mathbf{X}^T\hat{\mathbf{V}}^{-1(t+1)}\mathbf{X} + \hat{\lambda}^{(t+1)})^{-1}\mathbf{X}^T\hat{\mathbf{V}}^{-1(t+1)}\mathbf{y}. \qquad (۱۵.۲)$$

گام ۵) گام‌های (۱) تا (۴) را تا رسیدن به همگرایی $\left(|\hat{\boldsymbol{\beta}}_{MR}^{(t+1)} - \hat{\boldsymbol{\beta}}_{MR}^{(t)}| < \gamma\right)$ به ازای یک مقدار γ کوچک و از پیش تعیین شده، تکرار می‌کنیم.

در ادامه چگونگی محاسبه $\widehat{\beta}_{MR}$ را در قالب یک مطالعه شبیه‌سازی و مثال واقعی با استفاده از نرم‌افزار R بررسی می‌کنیم.

۱.۲.۲ مطالعه شبیه‌سازی

در این بخش به بررسی عملکرد برآوردگر آمیخته ریج و مقایسه رفتار آن با برآوردگر آمیخته خطی در مدل آمیخته خطی خواهیم پرداخت. از آنجایی که روش ریج یکی از ابزارهای مقابله با همخطی چندگانه است، لذا مطالعه شبیه‌سازی معطوف به این امر است. برای داشتن داده‌هایی با میزان همخطی دلخواه، مک دونالد و همکاران (۱۹۷۵) و گیبونز (۱۹۸۱) پیشنهاد دادند که داده‌ها براساس رابطه‌ی

$$x_{ij} = (۱ - \rho^۲)^{\frac{۱}{۲}} \cdot \mathbf{Z}_{ij} + \rho \mathbf{Z}_{ip} \qquad i = ۱, ۲, \ldots, n_i,\ j = ۱, ۲, \ldots, p$$

تولید شوند که در آن $\mathbf{Z}_{ij}$ اعداد شبه تصادفی مستقل از توزیع نرمال استاندارد و $\rho^۲$ همبستگی بین هر دو متغیر توضیحی است. در ادامه ابتدا ویژگی داده‌های شبیه‌سازی شده را بیان و سپس نتایج را ارائه می‌دهیم.

در این مطالعه برای شبیه‌سازی مقادیر متغیر پاسخ در رابطه‌ی (۶.۲) فرض می‌کنیم

$$\boldsymbol{\beta} = (۰, ۰/۴, ۱/۰, ۱/۶, ۲/۰)$$

$b_{ijk} \sim \mathcal{N}(۰, ۰/۶)$ و $\varepsilon_{ijk} \sim \mathcal{N}(۰, ۱)$ به‌طوری که برای هر واحد آزمایشی iام ۴ $= n_i$ و همچنین هر متغیر پیش‌بین از توزیع نرمال با میانگین ۵ و واریانس ۱ در نظر گرفته شده است.

برای ایجاد همبستگی بین متغیرهای پیش‌بین در چند حالت مختلف، مقدار ρ از ۰ تا ۰/۹ فرض شده است. همچنین برای محاسبه‌ی خطای استاندارد برآوردگر روند شبیه‌سازی مونت کارلویی به تعداد $B = ۲۰۰$ تکرار شده است. لازم به ذکر است که در این مطالعه تعداد نمونه‌ی آزمایشی $n = ۴۰$ در نظر گرفته شده است. نتایج شبیه‌سازی در جدول ۱.۲ آمده است.

جدول ۱.۲: نتایج مطالعه شبیه‌سازی مدل آمیخته خطی ریج

		مدل آمیخته خطی		مدل آمیخته خطی ریج	
ρ	β	برآورد	انحراف استاندارد	برآورد	انحراف استاندارد
۰/۰	۰/۰	۰/۰۰۴۰۱۸۶۳	۰/۰۳۴۴۱۳۴۲	۰/۰۰۴۷۱۶۳۷	۰/۰۳۰۷۸۲۸۸
	۰/۴	۰/۴۰۸۸۷۲۳۱	۰/۰۳۰۹۵۰۹۹	۰/۴۲۸۲۲۴۶۲	۰/۰۲۹۷۴۹۷۷
	۱/۰	۱/۰۱۷۶۴۹۹۵	۰/۰۲۹۳۳۴۶۹	۱/۰۱۹۱۸۷۶۶	۰/۰۲۷۸۱۳۴۲
	۱/۶	۱/۶۱۸۰۹۶۸۱	۰/۰۳۳۴۰۱۲۴	۱/۶۳۲۱۳۹۲۱	۰/۰۳۲۷۵۳۹۲
	۲/۰	۱/۹۶۲۷۴۴۶۸	۰/۰۳۵۹۴۵۲۲	۱/۹۴۴۲۸۷۳۰	۰/۰۳۳۹۲۸۱۷
MSE		۰/۸۶۷۷۵۸		۰/۷۷۸۱۵	
۰/۳	۰/۰	−۰/۰۰۲۱۳۷۳۸	۰/۰۴۳۸۶۴۱۷	۰/۰۰۵۵۲۴۲۹	۰/۰۳۸۶۳۲۳۲
	۰/۴	۰/۴۰۹۰۴۹۹۵	۰/۰۳۶۰۳۶۹۱	۰/۳۹۱۰۳۴۴۶	۰/۰۳۵۲۱۹۰۰
	۱/۰	۰/۹۹۹۸۰۲۰۱	۰/۰۳۳۲۳۹۸۱	۰/۹۸۵۹۰۳۸۹	۰/۰۲۹۹۸۸۶۸
	۱/۶	۱/۶۱۴۲۷۰۳۲	۰/۰۴۰۷۷۵۳۵	۱/۶۶۸۸۲۷۱۴	۰/۰۳۹۸۸۶۱۱
	۲/۰	۱/۹۸۹۴۲۵۷۱	۰/۰۴۲۹۸۱۱۴	۱/۹۳۴۷۷۱۳۷	۰/۰۳۹۸۹۱۲۳
MSE		۱/۲۵۴۴۲۱		۱/۰۹۹۶۰	
۰/۶	۰/۰	۰/۰۰۴۱۶۶۲۰	۰/۰۵۴۳۴۴۱۰	۰/۴۱۰۸۰۶۵۳	۰/۰۵۰۹۷۳۹۳
	۰/۴	۰/۴۰۴۱۶۶۲۰	۰/۰۵۲۲۱۳۵۵	۰/۴۱۳۲۱۲۸	۰/۰۴۸۲۳۰۷۷
	۱/۰	۰/۹۸۶۴۴۱۴۲	۰/۰۴۲۴۹۵۰۵	۱/۰۱۷۸۴۲۸۲	۰/۰۳۷۵۸۸۴۸
	۱/۶	۱/۶۰۱۸۴۵۹۴	۰/۰۴۷۰۶۸۸۱	۱/۶۷۸۰۲۹۹۸	۰/۰۴۶۵۶۰۵۰
	۲/۰	۲/۰۷۰۲۹۹۸۲	۰/۰۵۲۲۲۵۸۷	۲/۱۹۵۵۰۶۱۶	۰/۰۵۰۳۷۹۷۶
MSE		۱/۹۹۳۷۰۵		۱/۳۰۳۸۶۵	
۰/۹	۰/۰	۰/۰۲۷۶۱۴۱۱	۰/۰۹۹۳۷۳۴۷	۰/۰۴۷۴۴۹۶۱	۰/۰۸۵۰۸۶۳۳
	۰/۴	۰/۴۰۵۴۹۸۱۰	۰/۰۸۴۸۷۲۵۲	۰/۴۳۰۵۷۳۶۲	۰/۰۷۰۲۲۱۶۲
	۱/۰	۱/۰۵۱۷۲۱۷۰	۰/۰۷۸۲۲۹۶۰	۱/۱۴۶۳۲۹۹۷	۰/۰۶۰۹۸۹۳۶
	۱/۶	۱/۶۰۱۳۶۷۸۷	۰/۰۸۶۸۶۹۳۴	۱/۵۳۲۶۷۱۳۸	۰/۰۷۸۴۸۶۳۶
	۲/۰	۲/۰۶۵۲۷۳۱۳	۰/۰۷۷۱۵۰۲۳	۲/۱۷۱۸۱۹۸۷	۰/۰۷۰۶۸۱۶۶
MSE		۵/۸۷۹۲۰۹		۴/۳۸۶۰۸۹	

همان‌طور که در جدول ۱.۲ ملاحظه می‌شود اگرچه برآوردهای حاصل از روش آمیخته ریج دارای اریبی $\left(\left|\widehat{\beta}-\beta\right|\right)$ بیشتری نسبت به برآوردهای حاصل از آمیخته خطی است اما به مراتب دارای انحراف استاندارد (sd) کمتری است. این امر را مقادیر میانگین توان‌های دوم خطای برآورد (MSE) نیز تایید می‌کنند. آنچه که در تحلیل نتایج خروجی این شبیه‌سازی حائز اهمیت است این است که هر چقدر مقدار ρ به سمت یک نزدیک‌تر باشد یعنی میزان همخطی قوی‌تری در داده حاکم باشد برآورد حاصل از مدل آمیخته خطی ریج به‌طور قابل توجهی دارای انحراف استاندارد کمتری نسبت به برآورد حاصل از مدل آمیخته خطی دارا است. همین امر نتیجه استفاده از رگرسیون ریج در مقابله با مشکل همخطی است. اگرچه که این میزان همبستگی در حالت $\rho = 0$ نیز سبب برتری برآوردگر حاصل از مدل آمیخته ریج نسبت به برآورد بدست آمده از روش مدل آمیخته خطی شده است.

در ادامه، یک مثال واقعی به‌کار گرفته شده توسط الیوت و همکاران (۲۰۱۱) را شرح می‌دهیم.

۲.۲.۲ مثال واقعی

اندوتوکسین[8] نوعی سم موجود در باکتری است که توسط باکتری تولید یا آزاد می‌شود و ارتباط مستقیم با سیستم ایمنی بدن دارد، کمبود این ماده در خون عملکرد سلول‌ها را مختل و همچنین باعث انعقاد خون، اختلال بدن در دریافت اکسیژن، افسردگی، تب و بسیاری از اثرات نامطلوب بر بدن می‌شود. از طرفی وجود بیش از حد مجاز آن نیز خطر آفرین است که از جمله عوارض آن می‌توان به شوک بدن به همراه خونریزی‌های داخلی اشاره کرد. از این رو مطالعه بر روی واکنش بدن در برابر میزان اندوتوکسین از اهمیت ویژه‌ای برخوردار است.

مجموعه داده مورد بررسی در این تحقیق به مطالعه GENE[9] معروف است که حاصل یک آزمایش مداوم برای تشخیص اثرات عوامل ژنتیکی به میزان نیاسین[10] و اندوتوکسین می‌باشد.

در این مطالعه ۱۸۹ نفر به جهت تزریق انسداد اندوتوکسین با دُز کم (LPS)، مورد آزمایش قرار گرفته‌اند. نتیجه این تزریق یک التهاب خفیف پس از حدود ۶ الی ۸ ساعت است که طی آن آپولیپوپروتئین A1[11] (Apo-A1)، آپولیپوپروتئین B (Apo-B)، کلسترول کل[12] (Chol)، کلسترول لیپوپروتئین با چگالی بالا[13] (HDL)، کلسترول لیپوپروتئین با چگالی کم[14] (LDL)، فسفولیپید[15]

[8] Endotoxin
[9] Genetics of Evoked-Responses to Niacin and Endotoxemia
[10] Niacin
[11] Apolipoprotein
[12] Total cholesterol
[13] High-density lipoprotein cholesterol
[14] Low-density lipoprotein cholesterol
[15] Phospholipids

Phos، تری گلیسیرید[۱۶] Trig و فاکتور نکروز تومور آلفا[۱۷] TNF در ۶، ۱۲ و ۲۴ ساعت پس از التهاب ناشی‌شده، اندازه‌گیری شده‌اند. لازم به ذکر است که در این مطالعه متغیر LPS به عنوان متغیر پاسخ و متغیر جنسیت به عنوان اثر تصادفی در مدل در نظر گرفته شده است.

توجه به یک نکته اساسی مهم است. همبستگی با همخطی تفاوت دارد و باید وجود همخطی با استفاده از ابزارهای مختلف از جمله محاسبه عدد شرطی $\mathbf{X}^T\mathbf{X}$ و عامل تورم واریانس[۱۸]، بررسی شود. اما همان‌طور که مونت گومری و همکاران (۲۰۱۲) به آن اشاره کردند، می‌توان از روی مقادیر همبستگی بالا نیز تا حدی وجود همخطی را تشخیص داد. برای آگاهی بیشتر در خصوص معیارهای تشخیص همخطی پایان‌نامه آرست (۱۳۹۵) را ببینید.

همبستگی بین تمام جفت متغیرهای پیش‌بین در جدول ۲.۲، استفاده از مدل آمیخته خطی ریج را به مدل آمیخته خطی ترجیح می‌دهد.
نتایج استفاده از این دو مدل در جدول ۳.۲ خلاصه شده‌است.

جدول ۲.۲: همبستگی بین متغیرهای پیش‌بین در داده‌های GENE

	TNF	Apo-A1	Apo-B	Chol	HDL	Trig	Phos	LDL
TNF	۱/۰۰	۰/۱۳	۰/۰۴	۰/۰۴	۰/۰۷	−۰/۰۵	۰/۰۲	۰/۰۳
Apo-A1	۰/۱۳	۱/۰۰	۰/۰۷	۰/۴۰	۰/۸۷	۰/۰۱	۰/۶۶	۰/۰۱
Apo-B	۰/۰۴	۰/۰۷	۱/۰۰	۰/۸۸	−۰/۱۲	۰/۴۰	۰/۵۶	۰/۹۵
Chol	۰/۰۴	۰/۴۰	۰/۸۸	۱/۰۰	۰/۲۹	۰/۳۱	۰/۷۹	۰/۹۰
HDL	۰/۰۷	۰/۸۷	−۰/۱۲	۰/۲۹	۱/۰۰	−۰/۳۲	۰/۴۶	−۰/۱۱
Trig	−۰/۰۵	۰/۰۱	۰/۴۰	۰/۳۱	−۰/۳۲	۱/۰۰	۰/۴۲	۰/۲۷
Phos	۰/۰۲	۰/۶۶	۰/۵۶	۰/۷۹	۰/۴۶	۰/۴۲	۱/۰۰	۰/۵۵
LDL	۰/۰۳	۰/۰۱	۰/۹۵	۰/۹۰	−۰/۱۱	۰/۲۷	۰/۵۵	۱/۰۰

[16] Triglycerides
[17] Tumor necrosis factoralpha
[18] Variance inflation factor

جدول ۳.۲: برآورد ضرایب در داده‌های GENE

	مدل آمیخته خطی ریج		مدل آمیخته خطی	
	انحراف استاندارد	برآورد	انحراف استاندارد	برآورد
Time	۰/۱۲	۰/۴۷	۰/۱۶	۰/۴۶
Time1	۰/۲۰	−۰/۴۰	۰/۲۶	−۰/۴۰
Time2	۰/۱۵	−۰/۱۸	۰/۱۸	−۰/۱۷
Apo-A1	۰/۰۴	۰/۰۷	۰/۰۵	۰/۰۸
Apo-B	۰/۰۷	−۰/۲۴	۰/۰۹	−۰/۱۸
Chol	۰/۹۴	۱/۱۶	۱/۲۰	۰/۴۹
HDL	۰/۹۴	−۱/۱۶	۱/۲۰	−۰/۵۵
LDL	۰/۹۴	−۰/۹۸	۱/۲۰	−۰/۳۵
Phos	۰/۰۲	−۰/۰۶	۰/۰۳	−۰/۰۵
Trig	۰/۱۹	−۰/۲۱	۰/۲۴	−۰/۰۸
TNF	۰/۱۳	۰/۱۱	۰/۲۰	۰/۱۵

همان‌طور که در جدول ۳.۲ ملاحظه می‌شود مقادیر مربوط به انحراف استاندارد برآورد ضرایب در مدل آمیخته خطی ریج به مراتب کمتر از انحراف استاندارد برآورد ضرایب در مدل آمیخته خطی است. لذا می‌توان از این حیث مدل آمیخته خطی ریج را بر مدل آمیخته خطی برتر دانست.

۳.۲ مدل آمیخته خطی ریج با خطای بیضی‌گون

در این بخش مدل آمیخته خطی ریج را با فرض پایه‌ای این‌که بردار خطای تصادفی دارای توزیع بیضی‌گون[19] (توزیع چندمتغیره با منحنی‌های تراز بیضی شکل) می‌باشند، مورد بررسی قرار می‌دهیم و نتایج مطالعه شبیه‌سازی را در انتهای بخش در مقایسه با مدل آمیخته خطی ارائه می‌دهیم.

توضیحات مربوط به توزیع بیضی‌گون عمدتاً برگرفته از آرشی (۱۳۸۷) و حسن‌زاده بشتیان (۱۳۸۸) است و سایر نتایج آن جدید هستند.

[19] Elliptically contoured distribution

۱.۳.۲ توزیع بیضی‌گون

بردار تصادفی n مولفه‌ای $\mathbf{X}$ دارای توزیع بیضی‌گون با پارامترهای μ، Σ و تابع مولد مشخصه[۲۰] ψ است و با $\mathbf{X} \sim \mathcal{E}_n(\mu, \Sigma, \psi)$ نشان می‌دهیم اگر تابع مشخصه آن به صورت زیر باشد

$$\phi_{\mathbf{X}}(t) = \exp(it^T\mu)\psi\left(\frac{t^T\Sigma t}{۲}\right)$$

که در آن $\psi(\cdot)$ متعلق به رده توابعی به صورت $\psi(t) : [۰, \infty] \to \mathbb{R}$ است به‌طوری که $\psi\left(\sum_{i=۱}^{n} t_i^۲\right)$ یک تابع مشخصه nـ بعدی باشد.

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

$$f_{\mathbf{X}}(x) = c_n\,|\Sigma|^{۱/۲}\,g\left[\frac{(x-\mu)^T\Sigma^{-۱}(x-\mu)}{۲}\right]$$

که در آن تابع $g(\cdot)$ تابع مولد چگالی[۲۱] است و دارای شرط زیر می‌باشد

$$\int_{۰}^{\infty} x^{n/۲-۱}g(x)\,\mathrm{d}x < \infty$$

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

$$c_n = \frac{\Gamma(n/۲)}{(۲\pi)^{n/۲}}\left[\int_{۰}^{\infty} x^{n/۲-۱}g(x)\,\mathrm{d}x\right]^{-۱}$$

در این صورت می‌نویسیم $\mathbf{X} \sim \mathcal{E}_n(\mu, \Sigma, g)$. این نکته حائز اهمیت است، در صورتی که تابع چگالی وجود داشته باشد، تابع g، تابع ψ را مشخص می‌کند و برعکس. برای آگاهی بیشتر در این زمینه فنگ و همکاران (۱۹۹۰)، میرهد (۱۹۸۲) و کوپتا و وارگا (۱۹۹۳) را ببینید.

در این توزیع، $E(\mathbf{X}) = \mu$ و $\mathrm{Cov}(\mathbf{X}) = -۲\psi^T(۰)\Sigma$ وجود ندارد در صورتی که $|\psi^T(۰)| < \infty$.

توزیع‌های زیادی متعلق به خانواده توزیع‌های بیضی‌گون می‌باشند که از آن جمله می‌توان به توزیع‌های نرمال چندمتغیره، t چندمتغیره، کوشی چندمتغیره[۲۲]، پیرسن چندمتغیره نوع II[۲۳]، پیرسن چندمتغیره نوع VII[۲۴]، بسل چندمتغیره[۲۵]، نمایی توانی چندمتغیره[۲۶]، لاپلاس چندمتغیره[۲۷]، اِسلش تعمیم‌یافته[۲۸] و کاتز چندمتغیره[۲۹] اشاره کرد.

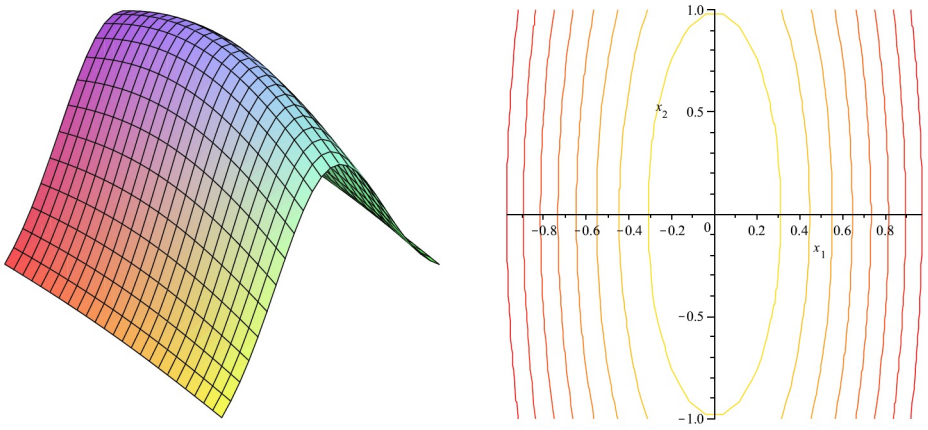
[20] Characteristic generator
[21] Density generator
[22] Multivariate Cauchy
[23] Multivariate Pearson type II
[24] Multivariate Pearson type VII
[25] Multivariate Bessel
[26] Multivariate Exponential power
[27] Multivariate Laplace
[28] Generalized Slash
[29] Multivariate Kotz type

حال به عنوان نمونه، تابع چگالی احتمال و تابع مولد چگالی را برای چند توزیع مهم بیضی‌گون، همراه با نمودار آن‌ها ارائه می‌دهیم. در نمودارها $n = ۲$ و $\boldsymbol{\Sigma} = \begin{bmatrix} ۱ & ۰ \\ ۰ & ۱ \end{bmatrix}$ فرض شده است.

۱. توزیع نرمال n متغیره $(\mathbf{X} \sim \mathcal{N}_n(\mu, \boldsymbol{\Sigma}))$: توابع مولد چگالی و چگالی احتمال به ترتیب عبارتند از

$$g(u) = e^{-u},$$
$$f_{\mathbf{X}}(x) = \frac{|\boldsymbol{\Sigma}|^{-۱/۲}}{(\sqrt{۲\pi})^n} \exp\left\{-\frac{۱}{۲}(x-\mu)^T \boldsymbol{\Sigma}^{-۱}(x-\mu)\right\}.$$

نمودارهای منحنی‌های تراز و تابع چگالی در حالت $n = ۲$ در شکل ۱.۲ داده شده است.

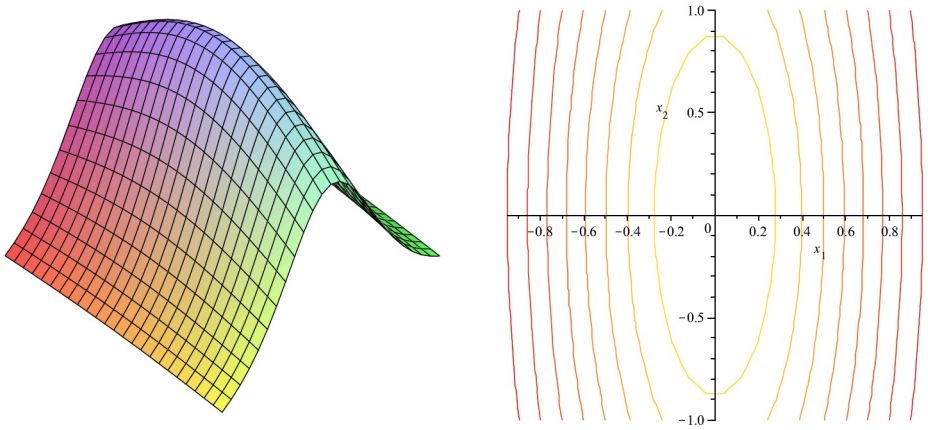


شکل ۱.۲: نمودارهای منحنی‌های تراز و تابع چگالی توزیع نرمال دومتغیره

۲. توزیع t چندمتغیره $(\mathbf{X} \sim \mathcal{T}_n(\mu, \boldsymbol{\Sigma}, \mathbf{V}))$: توابع مولد چگالی و چگالی احتمال به ترتیب عبارتند از

$$g_n(u) = \left(۱ + \frac{۲u}{\nu}\right)^{-(n+\nu)/۲},$$
$$f_{\mathbf{X}}(x) = \frac{\Gamma\left(\frac{n+\nu}{۲}\right)|\boldsymbol{\Sigma}|^{-۱/۲}}{(\nu\pi)^{n/۲}\Gamma\left(\frac{\nu}{۲}\right)} \left\{۱ + \frac{(x-\mu)^T\boldsymbol{\Sigma}^{-۱}(x-\mu)}{\nu}\right\}^{\frac{-(n+\nu)}{۲}}.$$

نمودارهای منحنی‌های تراز و تابع چگالی به ازای $n = ۲$ و $\nu = ۳$ در شکل ۲.۲ داده شده است.

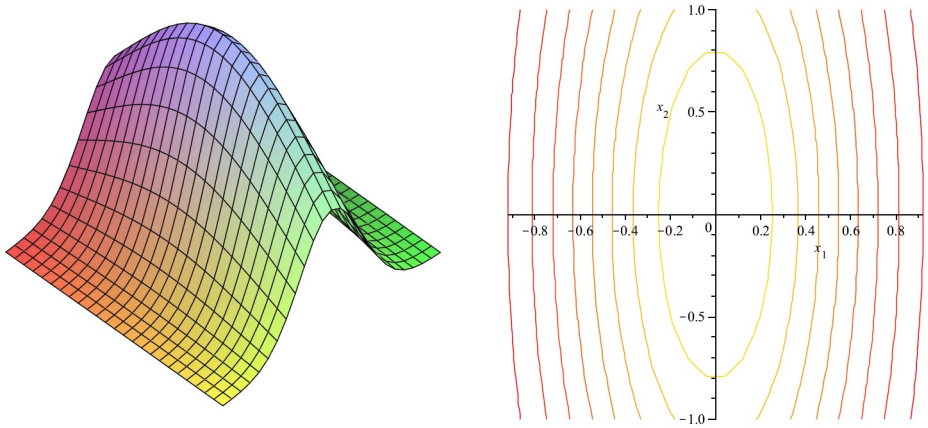


شکل ۲.۲: نمودارهای منحنی‌های تراز و تابع چگالی توزیع t دومتغیره

۳. توزیع کاتز n متغیره $(\mathbf{X} \sim \mathcal{K}_n(\mu, \mathbf{\Sigma}, r))$: توابع مولد چگالی و چگالی احتمال به ترتیب عبارتند از

$$g_n(u) = \mathrm{e}^{-ru}, r > 0,$$
$$f_{\mathbf{X}}(x) = \frac{\Gamma\left(\frac{n}{2}\right) |\mathbf{\Sigma}|^{-1/2} r^{n/2}}{(2\pi)^{n/2}\Gamma\left(\frac{n}{2}\right)} \exp\left\{-\frac{r}{2}\left[(x-\mu)^T \mathbf{\Sigma}^{-1}(x-\mu)\right]\right\}.$$

نمودارهای منحنی‌های تراز و تابع چگالی به ازای $r = 3$ در شکل ۳.۲ داده شده است.

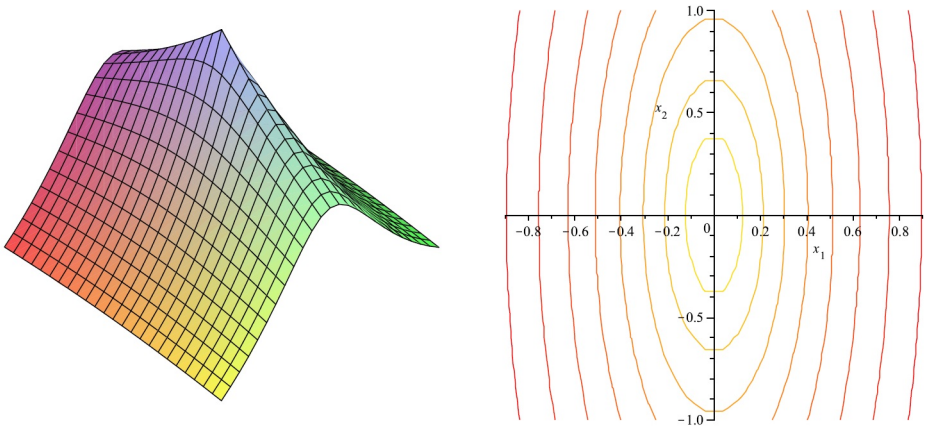


شکل ۳.۲: نمودارهای منحنی‌های تراز و تابع چگالی توزیع کاتز دومتغیره

۴. توزیع لاپلاس چندمتغیره $(\mathbf{X} \sim \mathcal{L}_n(\mu, \mathbf{\Sigma}, r))$: توابع مولد چگالی و چگالی احتمال به

ترتیب عبارتند از

$$g_n(u) = \mathrm{e}^{-\sqrt{۲}u^{۱/۲}},$$
$$f_{\mathbf{X}}(x) = \frac{\left(\frac{۱}{۲}\right)\Gamma\left(\frac{n}{۲}\right)|\mathbf{\Sigma}|^{-۱/۲}\sqrt{۲}^n}{(۲\pi)^{n/۲}\Gamma(n)} \exp\left\{-\frac{\sqrt{۲}}{۲}\left[(x-\mu)^T\mathbf{\Sigma}^{-۱}(x-\mu)\right]^{۱/۲}\right\}.$$



شکل ۴.۲: نمودارهای منحنی‌های تراز و تابع چگالی توزیع لاپلاس دومتغیره

در ادامه چند قضیه مهم را در خصوص توزیع بیضی‌گون ارائه می‌کنیم.

لم ۱.۳.۲. (چو، ۱۹۷۳). فرض کنید $\mathbf{X} \sim \mathcal{E}_n(\boldsymbol{\mu}, \mathbf{\Sigma}, \psi)$، در این صورت تابع چگالی احتمال $\mathbf{X}$ را می‌توان به صورت زیر نوشت

$$f_{\mathbf{X}}(\mathbf{x}) = \int_{۰}^{+\infty} w(t)h(x)\mathrm{d}t$$

که در آن $h(\cdot)$ نشان‌دهنده تابع چگالی احتمال $\mathcal{N}_n(\boldsymbol{\mu}, t^{-۱}\mathbf{\Sigma})$ است. بنابراین

$$f_{\mathbf{X}}(\mathbf{x}) = \int_{۰}^{+\infty} w(t)\frac{۱}{(۲\pi)^{n/۲}}|t^{-۱}\mathbf{\Sigma}|^{-۱/۲}\exp\left\{-\frac{۱}{۲}(\mathbf{x}-\boldsymbol{\mu})^T(t^{-۱}\mathbf{\Sigma})^{-۱}(\mathbf{x}-\boldsymbol{\mu})\right\}\mathrm{d}t$$

که در آن

$$w(t) = (۲\pi)^{n/۲}|t^{-۱}\mathbf{\Sigma}|^{۱/۲}L^{-۱}\left[f(s)\right]$$

و $L^{-۱}\left[f(s)\right]$ نشان‌دهنده تبدیل معکوس لاپلاس $h(s)$ به ازای

$$s = (\mathbf{x}-\boldsymbol{\mu})^T\mathbf{\Sigma}^{-۱}(\mathbf{x}-\boldsymbol{\mu})$$

می‌باشد.

قضیه ۱.۳.۲. اگر $\mathbf{X} \sim \mathcal{E}_p(\boldsymbol{\mu}, \mathbf{\Sigma}, \psi)$، $\mathbf{A}_{q\times p}$ و $\mathbf{C}_{q\times ۱}$ ماتریس‌های ثابت باشند، آنگاه

$$\mathbf{AX} + \mathbf{C} \sim \mathcal{E}_q(\mathbf{A}\boldsymbol{\mu} + \mathbf{C}, \mathbf{A\Sigma A}^T, \psi).$$

قضیه ۲.۳.۲. فرض کنید $\mathbf{X} \sim \mathcal{E}_n(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \psi)$ و Ω مجموعه‌ای در فضای $(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ باشد، به‌طوری که اگر $(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \in \Omega$ آنگاه $(\boldsymbol{\mu}, c\boldsymbol{\Sigma}) \in \Omega, \forall c > 0$. همچنین فرض کنید g به گونه‌ای باشد که $g(\mathbf{x}^T\mathbf{x})$ یک تابع چگالی در $\mathbb{R}^n$ و $y^{N/2}g(y)$ دارای یک مقدار ماکزیمم مثبت متناهی مانند y_g باشد و براساس تک مشاهده $\mathbf{x}$ از $|\boldsymbol{\Sigma}|^{-1/2}g\left[(\mathbf{x}-\boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right]$، برآوردگر درستنمایی ماکزیمم $(\tilde{\boldsymbol{\mu}}, \tilde{\boldsymbol{\Sigma}}) \in \Omega$ تحت نرمال بودن، موجود و منحصر بفرد باشد، آنگاه MLE برای g عبارت است از

$$\hat{\boldsymbol{\mu}} = \tilde{\boldsymbol{\mu}} \quad \hat{\boldsymbol{\Sigma}} = \frac{N}{y_g}\tilde{\boldsymbol{\Sigma}} \tag{۱۶.۲}$$

و تابع درستنمایی ماکزیمم به صورت زیر است

$$|\hat{\boldsymbol{\Sigma}}|^{-1/2}g(y_g). \tag{۱۷.۲}$$

اثبات: فرض کنید $\mathbf{B} = |\boldsymbol{\Sigma}|^{-1/N}\boldsymbol{\Sigma}$ و

$$d = (\mathbf{x}-\boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu}) = \frac{(\mathbf{x}-\boldsymbol{\mu})^T\mathbf{B}^{-1}(\mathbf{x}-\boldsymbol{\mu})}{|\boldsymbol{\Sigma}|^{1/N}} \tag{۱۸.۲}$$

آنگاه $(\tilde{\boldsymbol{\mu}}, \tilde{\boldsymbol{\Sigma}}) \in \Omega$ و $|\mathbf{B}| = 1$.
تابع درستنمایی عبارت است از

$$\left[(\mathbf{x}-\boldsymbol{\mu})^T\mathbf{B}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right]^{-N/2} d^{N/2}g(d) \tag{۱۹.۲}$$

تحت نرمال بودن $g(d) = (2\pi)^{-n/2}\mathrm{e}^{-d/2}$، ماکزیمم رابطه (۱۹.۲) در $\boldsymbol{\mu} = \tilde{\boldsymbol{\mu}}$،

$$\mathbf{B} = \tilde{\mathbf{B}} = |\tilde{\boldsymbol{\Sigma}}|^{-1/N}\tilde{\boldsymbol{\Sigma}}$$

و $d = N$ رخ می‌دهد. به‌طور کلی ماکزیمم رابطه (۱۹.۲) به ازای $\hat{\boldsymbol{\mu}} = \tilde{\boldsymbol{\mu}}$، $\hat{\mathbf{B}} = \tilde{\mathbf{B}}$ و $\hat{d} = y_g$ حاصل می‌شود. بنابراین MLE پارامتر $\boldsymbol{\Sigma}$ به صورت

$$\hat{\boldsymbol{\Sigma}} = |\hat{\boldsymbol{\Sigma}}|^{1/N}\hat{\mathbf{B}} = \frac{|\hat{\boldsymbol{\Sigma}}|^{1/N}}{|\tilde{\boldsymbol{\Sigma}}|^{1/N}}\tilde{\boldsymbol{\Sigma}}$$

است. با جایگذاری مقادیر $|\hat{\boldsymbol{\Sigma}}|^{1/N}$ و $|\tilde{\boldsymbol{\Sigma}}|^{1/N}$ از رابطه (۱۸.۲) به (۱۶.۲) می‌رسیم. همچنین با جایگذاری مقادیر فوق در رابطه (۱۹.۲)، رابطه (۱۷.۲) نتیجه می‌شود.

۲.۳.۲ مدل MR بیضی‌گون

مدل آمیخته خطی را در نظر بگیرید

$$\mathbf{y}_i = \mathbf{X}_i\boldsymbol{\beta} + \mathbf{Z}_i^T\mathbf{b}_i + \boldsymbol{\epsilon}_i \tag{۲۰.۲}$$

که در آن $\mathbf{y}_i = (y_{i1}, y_{i2}, \ldots, y_{in_i})^T$ برداری از n_i مشاهده، $\mathbf{X}_i$ ماتریس طرح $n_i \times p$ از اثرات ثابت و $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_p)^T$ بردار $p \times 1$ ضرایب ثابت و نامعلوم، و $\mathbf{Z}_i$ ماتریس طرح $q \times n_i$ از

اثرات تصادفی است. همچنین $\mathbf{b}_i$ بردار $q\times ۱$ ضرایب تصادفی و $\boldsymbol{\epsilon}_i$ بردار $n\times ۱$ خطاهای تصادفی هستند که دارای توزیع بیضی‌گون به صورت زیر می‌باشند

$$\mathbf{b}_i \sim \mathcal{E}_q(۰, \mathbf{D}, g)$$
$$\boldsymbol{\epsilon}_i \sim \mathcal{E}_n\left(۰, \sigma^۲\mathbf{I}, g\right),$$

و تابع مشخصه $\phi_{\boldsymbol{\epsilon}}(t)$ به صورت

$$\phi_{\boldsymbol{\epsilon}}(t) = \psi\left(\sigma^۲ t^T t\right)$$

است.

اگر $\boldsymbol{\epsilon}_i$ دارای یک تابع چگالی احتمال باشد آن‌گاه طبق آنچه گفته شد به صورت

$$f(\boldsymbol{\epsilon}_i)\alpha\left|\sigma^۲\mathbf{I}\right|^{-\frac{۱}{۲}} g\left(\frac{۱}{\sigma^۲}\boldsymbol{\epsilon}_i^T\boldsymbol{\epsilon}_i\right)$$

می‌باشد که در آن $g(\cdot)$ یک تابع نامنفی روی $\mathbb{R}^+$ بوده به‌طوری‌که $f(\cdot)$ یک تابع چگالی است. با توجه به قضیه ۱.۳.۲، هر ترکیب خطی از توزیع بیضی‌گون، دارای توزیع بیضی‌گون بوده و در این صورت داریم

$$\mathbf{y}_i \sim \mathcal{E}_n(\mathbf{X}_i\boldsymbol{\beta}, \mathbf{V}, g)$$

بنابراین لگاریتم تابع درستنمایی متغیر $\mathbf{y}$ به شرط معلوم بودن تابع $g(.)$ به صورت زیر است

$$\ell(\boldsymbol{\beta}, \mathbf{V}, \mathbf{y}) = -\frac{N}{۲}\log(۲\pi) - \frac{۱}{۲}\sum_{i=۱}^{n}\log|\mathbf{V}| + \log g\left(-\frac{۱}{۲}(\mathbf{y}-\mathbf{X}\boldsymbol{\beta})^T\mathbf{V}^{-۱}(\mathbf{y}-\mathbf{X}\boldsymbol{\beta})\right).$$

ماکزیمم کردن تابع فوق بر حسب $\boldsymbol{\beta}$، معادل با می‌نیمم کردن عبارت وابسته به $\boldsymbol{\beta}$ است. مشابه قضیه ۲.۳.۲ و به طور دقیق‌تر طبق آرشی و همکاران (۲۰۱۳) برآوردگر ماکزیمم درستنمایی (MLE) پارامتر $\boldsymbol{\beta}$ به صورت

$$\widehat{\boldsymbol{\beta}} = (\mathbf{X}^T\mathbf{V}^{-۱}\mathbf{X})^{-۱}\mathbf{X}^T\mathbf{V}^{-۱}\mathbf{y}. \qquad (۲۱.۲)$$

است. همانطور که مشاهده می‌شود، تابع g که نوع توزیع بیضی‌گون را مشخص می‌کند در ساختار برآوردگر $\boldsymbol{\beta}$ هیچ تأثیری ندارد. لذا اولین سوالی که در ذهن خواننده پیش می‌آید این است که دلیل استفاده از توزیع بیضی‌گون به عنوان مدل مؤلفه‌های خطا و اثر تصادفی چیست؟ در پاسخ می‌توان گفت گاهی اوقات تعیین توزیع دقیق متغیر پاسخ در مثال‌های واقعی کاری بس دشوار است در حالی که می‌توان متقارن بودن آن را به راحتی بر اساس داده‌ها بررسی نمود. در این حالت چنانچه بر اساس ضرایب چولگی پی به متقارن بودن توزیع داده‌ها ببریم، دیگر لازم نیست در تحلیل داده‌های طولی نگران توزیع دقیق مؤلفه پاسخ برای یافتن MLE باشیم و می‌توان از رابطه (۲۱.۲) به‌عنوان برآوردگر صحیح بردار ضرایب ثابت $\boldsymbol{\beta}$ استفاده کرد.

حال لگاریتم تابع درستنمایی متغیر $\mathbf{y}$ را بر پایه‌ی مدل آمیخته خطی با اعمال تابع جریمه در رگرسیون ریج، به صورت زیر داریم

$$\ell^{pen}(\boldsymbol{\beta}, \mathbf{V}, \mathbf{y}) = -\frac{N}{۲}\log(۲\pi) - \frac{۱}{۲}\sum_{i=۱}^{n}\log|\mathbf{V}| + \log g\left(-\frac{۱}{۲}(\mathbf{y}-\mathbf{X}\boldsymbol{\beta})^T\mathbf{V}^{-۱}(\mathbf{y}-\mathbf{X}\boldsymbol{\beta})\right) - \lambda\boldsymbol{\beta}^T\boldsymbol{\beta},$$

بنابراین با انجام محاسباتی ساده مشابه بخش ۲.۲ برآوردگر MR پارامتر $\boldsymbol{\beta}$ به صورت

$$\widehat{\boldsymbol{\beta}}_{MR} = (\mathbf{X}^T\mathbf{V}^{-۱}\mathbf{X} + \lambda\mathbf{I})^{-۱}\mathbf{X}^T\mathbf{V}^{-۱}\mathbf{y}.$$

حاصل می‌شود.

۳.۳.۲ مطالعه شبیه‌سازی

در این قسمت فرض کنید که مؤلفه‌های تصادفی مدل (۲۰.۲) برای شبیه‌سازی مقادیر متغیر پاسخ، بردار ضرایب رگرسیونی به صورت

$$\boldsymbol{\beta} = (۰, ۰/۴, ۱, ۱/۶, ۲)$$

، $b_{ijk} \sim \mathcal{T}(۰, ۰/۳, ۵)$ و $\varepsilon_{ijk} \sim \mathcal{T}(۰, ۱, ۵)$ فرض شده‌اند به‌طوری که برای هر واحد آزمایشی iام $n_i = ۴$ و هر متغیر پیش‌بین از توزیع نرمال با میانگین ۵ و واریانس ۱ فرض شده‌اند. همچنین برای ایجاد داده‌های هم‌خط از بسته‌ی EnvStat در نرم‌افزار R با ضریب همبستگی‌های متفاوت موجود در جدول ۴.۲

$$\rho \in \{۰/۰, ۰/۲, ۰/۵, ۰/۷, ۰/۹\}$$

استفاده شده است.

لازم به ذکر است که برای محاسبه‌ی خطای استاندارد برآورد بدست آمده با روند شبیه‌سازی مونت‌کارلویی به تعداد $B = ۱۰۰$ تکرار شده و تعداد نمونه آزمایشی $n = ۴۰$ در نظر گرفته شده است.

جدول ۴.۲: نتایج مطالعه شبیه‌سازی مدل آمیخته خطی ریج با خطای بیضی‌گون

		مدل آمیخته خطی با خطای بیضی‌گون		مدل آمیخته خطی ریج با خطای بیضی‌گون	
ρ	β	برآورد	انحراف استاندارد	برآورد	انحراف استاندارد
۰/۰	۰/۰	−۰/۰۱۴۲۵۴۲۳	۰/۱۲۲۶۴۴۶۹	−۰/۰۱۰۴۱۲۴	۰/۱۲۱۷۸۵۱
	۰/۴	۰/۴۰۱۱۱۳۵۵	۰/۱۱۱۹۵۰۳۹	۰/۴۰۳۳۴۷۶	۰/۱۱۱۴۹۲۳
	۱/۰	۰/۹۹۵۱۳۸۹۱	۰/۱۲۷۳۴۴۸۳	۰/۹۹۴۷۲۱۴	۰/۱۲۶۸۱۱۳
	۱/۶	۱/۶۱۷۲۸۱۷۴	۰/۰۹۶۲۹۷۴۵	۱/۶۱۵۱۳۳۰	۰/۰۹۵۶۲۵۲
	۲/۰	۱/۹۹۷۳۸۸۳۲	۰/۱۱۳۳۸۸۱۸	۱/۹۹۳۶۰۳۴	۰/۱۱۳۱۶۶۲
MSE		۱۰/۵۴۷۹۶		۱۰/۴۴۷۴۶	
۰/۳	۰/۰	−۰/۰۱۹۰۹۵۰۳	۰/۱۴۰۴۲۴۲	−۰/۰۱۳۴۹۸۰۳	۰/۱۳۸۹۶۱۷
	۰/۴	۰/۳۹۹۵۶۱۴۳	۰/۱۳۲۲۷۸۶	۰/۴۰۳۰۲۹۱۷	۰/۱۳۱۴۴۵۲
	۱/۰	۰/۹۹۲۵۴۶۲۶	۰/۱۴۷۰۰۱۹	۰/۹۹۱۹۸۶۰۶	۰/۱۴۶۱۳۹۶
	۱/۶	۱/۶۲۱۷۶۵۱۶	۰/۱۱۰۰۰۴۴	۱/۶۱۸۴۹۱۱۱	۰/۱۰۸۸۶۸۷
	۲/۰	۲/۰۰۱۸۸۴۹۵	۰/۱۳۲۷۲۷۹	۱/۹۹۶۳۴۰۸۳	۰/۱۳۲۰۹۹۸
MSE		۱۴/۱۶۷۹۱		۱۳/۹۶۰۲۵	
۰/۶	۰/۰	−۰/۰۲۳۹۹۷۱۷	۰/۱۷۵۱۶۳۶	−۰/۰۱۴۶۹۵۰۱	۰/۱۷۲۳۱۹۱
	۰/۴	۰/۳۹۶۴۳۴۶۹	۰/۱۷۶۹۸۶۱	۰/۴۰۲۲۳۵۷۹	۰/۱۷۵۰۱۴۰
	۱/۰	۰/۹۸۸۱۰۳۵۴	۰/۱۹۲۵۵۴۷	۰/۹۸۷۷۳۴۳۳	۰/۱۹۰۶۳۱۵
	۱/۶	۱/۶۲۹۷۷۰۷۸	۰/۱۴۷۷۸۸۳	۱/۶۲۴۰۷۴۸۳	۰/۱۴۵۴۹۷۲
	۲/۰	۲/۰۰۶۱۱۴۰۷	۰/۱۷۹۳۵۰۷	۱/۹۹۶۶۹۷۴۲	۰/۱۷۷۶۳۲۰
MSE		۲۴/۴۹۶۳۳		۲۳/۹۰۲۸۳	
۰/۹	۰/۰	−۰/۰۴۷۱۲۲۶۱	۰/۳۴۸۴۷۳۸	−۰/۰۱۱۴۲۵۹۵	۰/۳۳۱۹۸۵۴
	۰/۴	۰/۳۹۹۴۳۶۷۱	۰/۳۴۹۵۴۵۶	۰/۴۲۰۷۴۸۷۳	۰/۳۳۶۴۶۸۰
	۱/۰	۰/۹۸۴۸۸۵۷۴	۰/۳۸۱۷۷۲۴	۰/۹۸۲۵۳۹۲۹	۰/۳۹۸۵۳۴۶
	۱/۶	۱/۶۴۳۸۴۶۷۵	۰/۲۸۵۳۲۷۲	۱/۶۲۳۸۵۷۷۲	۰/۲۷۴۱۱۰۴
	۲/۰	۲/۰۱۵۲۳۱۲۶	۰/۳۴۶۰۵۶۷	۱/۹۷۹۵۸۶۰۱	۰/۳۳۴۲۰۷۶
MSE		۹۴/۴۸۹۹		۹۱/۰۵۵۶۲	

نتایج مطالعه شبیه‌سازی در جدول ۴.۲ آمده است. همان‌طور که در این جدول ملاحظه می‌شود برآوردهای حاصل از روش آمیخته ریج با خطاهای بیضی‌گون دارای انحراف استاندارد کمتری نسبت به برآورد حاصل از روش آمیخته خطی است. مقادیر میانگین توان دوم خطای برآورد (MSE) نیز برتری روش آمیخته ریج را نسبت به روش آمیخته خطی تایید می‌کند. همچنین طبق جدول ۴.۲ واضح است که اگر چه در برخی موارد میزان اریبی $\left(\left|\widehat{\beta}-\beta\right|\right)$ نسبت به روش آمیخته خطی کمتر است اما هرچه میزان همخطی در داده‌ها قوی‌تر باشد، انحراف استاندارد برآورد در روش آمیخته ریج کاهش بیشتری داشته است.

۴.۲ برآوردگرهای انقباضی در مدل آمیخته خطی ریج

در این بخش برآوردگرهای آزمون اولیه[۳۰] و انقباضی نوع استاین[۳۱] را تعریف کرده و به جهت بهبود خطای برآورد در مدل آمیخته خطی ریج، برآوردگرهای آزمون اولیه آمیخته ریج و انقباضی آمیخته و انقباضی آمیخته ریج را برای برآورد ضرایب اثرات ثابت مدل، مطرح می‌کنیم.

گاهی در مسائل آماری، شرایط موجود ایجاب می‌کند که در مورد رخداد فرضیه‌ی $H_0 : \boldsymbol{\beta} = \boldsymbol{\beta}_0$ نمی‌توان به‌طور حتمی نظر داد و در مورد اینکه آیا $\boldsymbol{\beta} = \boldsymbol{\beta}_0$ درست است یا خیر تردید ایجاد می‌شود، یعنی نمی‌توان به‌طور قطعی یکی از دو برآوردگر $\boldsymbol{\beta}$ یا $\boldsymbol{\beta}_0$ را انتخاب کرد. در چنین مواردی می‌توان ابتدا فرضیه‌ی H_0 را آزمون کرد و سپس براساس نتیجه‌ی آزمون یکی از دو برآوردگر را انتخاب نمود. بر این اساس بنکرافت (۱۹۴۴) برآوردگر آزمون اولیه $\boldsymbol{\beta}$ را به صورت زیر ارائه کرد

$$\widehat{\boldsymbol{\beta}}^{PT} = \widehat{\boldsymbol{\beta}}\mathbf{I}_{\{H_0 \text{ ناحیه رد}\}} + \boldsymbol{\beta}_0\mathbf{I}_{\{H_0 \text{ ناحیه پذیرش}\}}$$

که در آن $\mathbf{I}_{\{A\}}$ تابع نشان‌گر مجموعه A است.

در مدل آمیخته خطی ریج فرضیه‌ی زیر را در نظر بگیرید

$$\begin{cases} H_0 : \boldsymbol{\beta} = \boldsymbol{\beta}_0 \\ H_1 : \boldsymbol{\beta} \neq \boldsymbol{\beta}_0 \end{cases}$$

در این صورت برآوردگر آزمون اولیه آمیخته ریج[۳۲] $(PTMR)$ را می‌توان به صورت زیر بیان کرد

$$\widehat{\boldsymbol{\beta}}_{MR}^{PT} = \begin{cases} \widehat{\boldsymbol{\beta}}_{MR} & \text{فرضیه‌ی } H_0 \text{ رد شود} \\ \boldsymbol{\beta}_0 & \text{فرضیه‌ی } H_0 \text{ پذیرفته شود} \end{cases}$$

برآوردگر $PTMR$ را می‌توان با استفاده از تابع نشانگر به صورت زیر بازنویسی کرد

$$\begin{aligned} \widehat{\boldsymbol{\beta}}_{MR}^{PT} =& \widehat{\boldsymbol{\beta}}_{MR}\mathbf{I}_{\{H_0 \text{ ناحیه رد}\}} + \boldsymbol{\beta}_0\mathbf{I}_{\{H_0 \text{ ناحیه پذیرش}\}} \\ =& \widehat{\boldsymbol{\beta}}_{MR} - \widehat{\boldsymbol{\beta}}_{MR}\mathbf{I}_{\{H_0 \text{ ناحیه پذیرش}\}} + \boldsymbol{\beta}_0\mathbf{I}_{\{H_0 \text{ ناحیه پذیرش}\}} \\ =& \widehat{\boldsymbol{\beta}}_{MR} - \left(\widehat{\boldsymbol{\beta}}_{MR} - \boldsymbol{\beta}_0\right)\mathbf{I}_{\{H_0 \text{ ناحیه پذیرش}\}} \\ =& \widehat{\boldsymbol{\beta}}_{MR} - \left(\widehat{\boldsymbol{\beta}}_{MR} - \boldsymbol{\beta}_0\right)\mathbf{I}_{(\mathbf{L}_n < \mathbf{L}_n(\alpha))} \end{aligned}$$

که در آن $\mathbf{L}_n$ آماره والد[۳۳] آزمون فرضیه‌ی H_0، با توجه به نرمال بودن توزیع مجانبی $\widehat{\boldsymbol{\beta}}$ به صورت زیر است (آرشی و همکاران ۲۰۱۳)

[30] Preliminary test
[31] Stein-type shrinkage
[32] Preliminary test mixed ridge
[33] Wald statistic

$$\mathbf{L}_n = \left(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}_\circ\right)^T \left(\mathbf{X}^T\mathbf{V}^{-1}\mathbf{X}\right) \left(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}_\circ\right) \qquad (۲۲.۲)$$

همچنین $\mathbf{L}_n(\alpha)$ صدک بالایی ۱۰۰α درصد توزیع کی‌ـدو با p درجه آزادی است رفتار برآوردگر $PTMR$ به اندازه آزمون یا سطح معناداری $\alpha\,(\circ < \alpha < ۱)$ بستگی دارد. با توجه به نتیجه‌ی آزمون، $\widehat{\boldsymbol{\beta}}_{MR}^{PT}$ به یکی از دو مقدار $\widehat{\boldsymbol{\beta}}_{MR}$ یا $\boldsymbol{\beta}_\circ$ تبدیل می‌شود که اگر فرضیه‌ی $H_\circ$ رد شود $\widehat{\boldsymbol{\beta}}_{MR}$ و در صورتی که $H_\circ$ رد نشود $\boldsymbol{\beta}_\circ$ نتیجه می‌شود.

خاصیت گسسته‌سازی تابع نشانگر باعث می‌شود که نتوان رفتار دو برآوردگر را توأماً براساس $PTMR$ بررسی کرد، لذا ایده استفاده از برآوردگری که بین این دو برآوردگر تغییر کند باعث مطرح شدن برآوردگر انقباضی نوع استاین شد. برای اطلاعات بیشتر جیمز و استاین (۱۹۶۱) را ببینید.

برآوردگر انقباضی نوع استاین آمیخته ریج[۳۴] (SSMR) به صورت زیر تعریف می‌شود

$$\begin{aligned}\widehat{\boldsymbol{\beta}}_{MR}^{S} &= \boldsymbol{\beta}_\circ + \left(۱ - c\mathbf{L}_n^{-۱}\right)\left(\widehat{\boldsymbol{\beta}}_{MR} - \boldsymbol{\beta}_\circ\right) \\ &= \widehat{\boldsymbol{\beta}}_{MR} - c\left(\widehat{\boldsymbol{\beta}}_{MR} - \boldsymbol{\beta}_\circ\right)\mathbf{L}_n^{-۱}\end{aligned}$$

که در آن $c = \dfrac{(p-۲)(N-p)}{p(N-p+۲)}$ و به ضریب $\left(۱ - c\mathbf{L}_n^{-۱}\right)$ عامل انقباض گویند. مقدار c برگرفته از صالح (۲۰۰۶) است.

با توجه به ساختار $\widehat{\boldsymbol{\beta}}_{MR}^{PT}$ و $\widehat{\boldsymbol{\beta}}_{MR}^{S}$ هر دو به ازای مقادیر بزرگ $\mathbf{L}_n$، $\widehat{\boldsymbol{\beta}}_{MR}$ را نتیجه می‌دهند. در حالیکه برای مقادیر کوچک $\mathbf{L}_n$، رفتار $\widehat{\boldsymbol{\beta}}_{MR}^{PT}$ متفاوت از $\widehat{\boldsymbol{\beta}}_{MR}^{S}$ است. همچنین برخلاف برآوردگر $PTMR$، برآوردگر نوع استاین یک تابع هموار و مستقل از سطح معناداری α است. برای آگاهی بیشتر در خصوص برآوردگرهای ارائه‌شده، نوروزی راد (۱۳۹۶) را ببینید.

اکنون برای محاسبه‌ی این برآوردگرها الگوریتم ارائه‌شده توسط الیوت و همکاران (۲۰۱۱) را به صورت زیرتوسعه داده‌ایم. همچنین برای مقدار اولیه پارامتر λ در الگوریتم محاسبه برآوردگر ریج این بار از روش پیشنهادی هورل و کنارد (۱۹۷۰) به صورت

$$\widehat{\lambda}_\circ = \frac{\widehat{\sigma}^۲}{\widehat{\alpha}^۲_{\max}}$$

استفاده می‌کنیم که در آن $\widehat{\alpha}^۲_{\max} = \max\left(\widehat{\alpha}_۱, \ldots, \widehat{\alpha}_p\right)$ به‌طوری‌که $\widehat{\alpha} = \boldsymbol{\Gamma}^T\widehat{\boldsymbol{\beta}}$ و $\boldsymbol{\Gamma}$ ماتریس متعامد شامل بردارهای ویژه در تجزیه طیفی $\mathbf{X}^T\mathbf{X}$ است. لازم به یادآوری است که در تجزیه طیفی $\mathbf{X}^T\mathbf{X}$ داریم

$$\mathbf{X}^T\mathbf{X} = \boldsymbol{\Gamma}\boldsymbol{\Lambda}\boldsymbol{\Gamma}^T$$

به‌طوری که $\boldsymbol{\Lambda} = \operatorname{diag}\left(\lambda_۱^*, \ldots, \lambda_p^*\right)$ و λ_i^* مقادیر ویژه iام ماتریس $\mathbf{X}^T\mathbf{X}$ است.

[34] Stein-type shrinkage mixed ridge

الگوریتم ۴ الگوریتم EM در برآورد آزمون اولیه و نوع استاین پارامتر مدل آمیخته خطی‌ی ریج

گام ۱) قرار می‌دهیم $\hat{\boldsymbol{\theta}}^{(t)} = \boldsymbol{\theta}_0$ و $\hat{\lambda}^{(t)} = \lambda_0$ سپس مقدار آماره‌های $\hat{t}_1^{(t)}$ و $\hat{t}_2^{(t)}$ را برای مرحله tام به‌صورت زیر محاسبه می‌کنیم:

$$\begin{aligned}\hat{t}_1^{(t)} &= E\left(\sum_{i=1}^{n}\boldsymbol{\epsilon}_i^T\boldsymbol{\epsilon}_i \mid \mathbf{y}_i, \hat{\boldsymbol{\beta}}^{(t)}, \hat{\boldsymbol{\theta}}^{(t)}\right)\\ &= \sum_{i=1}^{n}\left[\boldsymbol{\epsilon}_i^T(\hat{\boldsymbol{\theta}}^{(t)})\boldsymbol{\epsilon}_i(\hat{\boldsymbol{\theta}}^{(t)}) + trVar\left(\boldsymbol{\epsilon}_i \mid \mathbf{y}_i, \hat{\boldsymbol{\beta}}^{(t)}, \hat{\boldsymbol{\theta}}^{(t)}\right)\right],\end{aligned}$$

$$\begin{aligned}\hat{t}_2^{(t)} &= E\left(\sum_{i=1}^{n}\mathbf{b}_i\mathbf{b}_i^T \mid \mathbf{y}_i, \hat{\boldsymbol{\beta}}^{(t)}, \hat{\boldsymbol{\theta}}^{(t)}\right)\\ &= \sum_{i=1}^{n}\left[\mathbf{b}_i(\hat{\boldsymbol{\theta}}^{(t)})\mathbf{b}_i^T(\hat{\boldsymbol{\theta}}^{(t)}) + Var\left(\mathbf{b}_i \mid \mathbf{y}_i, \hat{\boldsymbol{\beta}}^{(t)}, \hat{\boldsymbol{\theta}}^{(t)}\right)\right].\end{aligned}$$

گام ۲) برای به دست آوردن $\hat{\boldsymbol{\theta}}^{(t+1)}$، مقدار $\hat{\sigma}^{2(t+1)}$ و $\hat{\mathbf{D}}^{(t+1)}$ را از رابطه زیر محاسبه می‌کنیم:

$$\hat{\sigma}^{2(t+1)} = \hat{t}_1^{(t)}/N,$$

$$\hat{\mathbf{D}}^{(t+1)} = \hat{t}_2^{(t)}/n.$$

که در آن $N = \sum_{i=1}^{n} n_i$ و n برابر تعداد واحدهای آزمایشی است. حال برای محاسبه $\hat{\mathbf{V}}^{(t+1)}$ داریم

$$\hat{\mathbf{V}}^{(t+1)} = \mathbf{Z}\hat{\mathbf{D}}^{(t+1)}\mathbf{Z}^T + \hat{\sigma}^{2(t+1)}\mathbf{I}.$$

گام ۳) ماتریس $\hat{\mathbf{S}}^{(t+1)}$ را از رابطه زیر محاسبه می‌کنیم

$$\begin{aligned}\hat{\mathbf{S}}^{(t+1)} &= \mathbf{X}(\mathbf{X}^T\hat{\mathbf{V}}^{-1(t+1)}\mathbf{X} + \hat{\lambda}^{(t)}\mathbf{I})^{-1}\mathbf{X}^T\hat{\mathbf{V}}^{-1(t+1)}\\ &+ \mathbf{Z}\hat{\mathbf{D}}^{(t+1)}\mathbf{Z}^T\left(\mathbf{I} - \mathbf{X}\left[(\mathbf{X}^T\hat{\mathbf{V}}^{-1(t+1)}\mathbf{X} + \hat{\lambda}^{(t)}\mathbf{I})^{-1}\mathbf{X}^T\hat{\mathbf{V}}^{-1(t+1)}\right]\right),\end{aligned}$$

سپس برای محاسبه $\hat{\lambda}^{(t+1)}$ داریم

$$\hat{\lambda}^{(t+1)} = \arg\min_{\lambda}\left\{n^{-1}(\mathbf{y} - \widehat{\mathbf{y}})^T(\mathbf{y} - \widehat{\mathbf{y}})/(1 - \mathrm{tr}(\hat{\mathbf{S}}^{(t+1)})/n)^2\right\}$$

گام ۴) در نهایت مقدار $\hat{\boldsymbol{\beta}}_{MR}^{(t+1)}$ از رابطه زیر بدست می‌آید

$$\hat{\boldsymbol{\beta}}_{MR}^{(t+1)} = (\mathbf{X}^T\hat{\mathbf{V}}^{-1(t+1)}\mathbf{X} + \hat{\lambda}^{(t+1)})^{-1}\mathbf{X}^T\hat{\mathbf{V}}^{-1(t+1)}\mathbf{y}.$$

گام ۵) گام‌های (۱) تا (۴) را تا رسیدن به همگرایی $\left(|\hat{\boldsymbol{\beta}}_{MR}^{(t+1)} - \hat{\boldsymbol{\beta}}_{MR}^{(t)}| < \gamma\right)$ به ازای یک مقدار γ کوچک و از پیش تعیین شده، تکرار می‌کنیم.

گام ۶) پس از محاسبه آماره L_n از رابطه (۲۲.۲)، داریم

$$\hat{\boldsymbol{\beta}}_{MR}^{PT} = \hat{\boldsymbol{\beta}}_{MR} - (\hat{\boldsymbol{\beta}}_{MR} - \boldsymbol{\beta}_0) \quad I(L_n < L_n(\alpha))$$

$$\hat{\boldsymbol{\beta}}_{MR}^{S} = \hat{\boldsymbol{\beta}}_{MR} - c(\hat{\boldsymbol{\beta}}_{MR} - \boldsymbol{\beta}_0)L_n^{-1}$$

۱.۴.۲ مطالعه شبیه‌سازی

برای مقایسه برآوردگرهای پیشنهادی $\boldsymbol{\beta}_{MR}^{PT}$ و $\boldsymbol{\beta}_{MR}^{S}$ با برآوردگر ارائه شده توسط الیوت و همکاران (۲۰۱۱) از مطالعه شبیه‌سازی مونت کارلویی استفاده می‌کنیم. در این شبیه‌سازی برای هر $n = ۴۰$ واحد آزمایشی $n_i = ۴$ را ثابت در نظر گرفته و همچنین

$$\boldsymbol{\beta} = (۰/۰, ۰/۱, ۰/۲, ۰/۴, ۰/۸)$$

و $b_{ijk} \sim \mathcal{N}(۰, ۰/۶)$ و $\varepsilon_{ijk} \sim \mathcal{N}(۰, ۱)$ فرض شده‌اند. برای تولید داده‌های هم‌خط از بسته‌ی EnvStat موجود در نرم‌افزار R استفاده شده است که هر متغیر پیش‌بین دارای توزیع نرمال با میانگین ۵ و واریانس ۱ می‌باشد که مقدار همبستگی بین متغیرهای پیش‌بین طبق جدول‌های ۵.۲ و ۶.۲ به صورت $\rho \in \{۰/۰, ۰/۲, ۰/۵, ۰/۷, ۰/۹\}$ است. همچنین برای محاسبه‌ی خطای استاندارد (sd) و میانگین توان‌های دوم خطا (MSE)، روند شبیه‌سازی به تعداد $B = ۱۰۰$ بار تکرار شده است. لازم به ذکر است که در این مطالعه، به جهت بدست آوردن برآوردهای $\widehat{\boldsymbol{\beta}}_{MR}^{PT}$ و $\widehat{\boldsymbol{\beta}}_{MR}^{S}$ فرضیه $H_۰ : \boldsymbol{\beta} = ۰$ در نظر گرفته شده است. نتایج این مطالعه در جدول‌های ۵.۲ و ۶.۲ آمده است.

جدول ۵.۲: نتایج مطالعه شبیه‌سازی برآوردگر نوع استاین در مدل آمیخته خطی ریج

مدل آمیخته خطی ریج نوع استاین		مدل آمیخته خطی نوع استاین		مدل آمیخته خطی ریج		مدل آمیخته خطی			
انحراف استاندارد	برآورد	انحراف استاندارد	برآورد	انحراف استاندارد	برآورد	انحراف استاندارد	برآورد	$\boldsymbol{\beta}$	ρ
۰/۰۱۶۵۱	۰/۳۷۰۷	۰/۰۱۸۳۳	−۰/۰۱۵۸۵	۰/۰۱۷۰۸	۰/۳۸۲۸	۰/۰۱۸۹۷	−۰/۰۱۶۵۱	۰/۰	۰/۰
۰/۰۱۵۳۰	۰/۱۱۵۱۳	۰/۰۱۸۰۲	۰/۰۸۷۲۵	۰/۰۱۵۹۳	۰/۱۱۹۰۷	۰/۰۱۸۶۴	۰/۰۹۰۱۶	۰/۱	
۰/۰۱۴۹۸	۰/۲۰۸۰۵	۰/۰۱۷۷۷	۰/۱۹۱۸۹	۰/۰۱۵۵۹	۰/۲۱۵۱۹	۰/۰۱۸۳۲	۰/۱۹۸۴۱	۰/۲	
۰/۰۱۸۶۱	۰/۳۵۸۰۳	۰/۰۲۱۱۰	۰/۳۷۲۹۸	۰/۰۱۸۷۰	۰/۳۷۰۱۵	۰/۰۲۱۵۳	۰/۳۸۵۴۴	۰/۴	
۰/۰۱۹۷۹	۰/۷۲۵۹۱	۰/۰۱۶۴۰	۰/۸۱۲۰۰	۰/۰۱۸۶۸	۰/۷۵۰۶۳	۰/۰۱۷۸۵	۰/۸۳۹۴۲	۰/۸	
۰/۲۴۳۹۶۶		۰/۲۷۱۸۵۸		۰/۲۴۳۳۳۶		۰/۲۹۴۱۵۳		MSE	
۰/۰۱۹۴۷	۰/۰۵۹۱۹	۰/۰۲۱۹۳	−۰/۰۰۹۱۴	۰/۰۲۰۱۴	۰/۰۶۱۱۲	۰/۰۲۲۷۱	−۰/۰۰۹۶۲	۰/۰	۰/۲
۰/۰۱۶۵۹	۰/۱۲۲۳۵	۰/۰۱۹۸۹	۰/۰۸۹۵۴	۰/۰۱۷۲۸	۰/۱۲۶۵۵	۰/۰۲۰۵۷	۰/۰۹۲۵۱	۰/۱	
۰/۰۱۵۸۱	۰/۲۱۲۴۴	۰/۰۱۹۸۰	۰/۱۸۹۸۰	۰/۰۱۶۴۸	۰/۲۱۹۷۱	۰/۰۲۰۴۲	۰/۱۹۶۲۲	۰/۲	
۰/۰۲۰۵۷	۰/۳۴۷۸۳	۰/۰۲۳۹۶	۰/۳۶۷۰۷	۰/۰۲۰۶۴	۰/۳۵۹۵۹	۰/۰۲۴۴۶	۰/۳۷۹۲۸	۰/۴	
۰/۰۲۳۱۲	۰/۷۰۱۳۱	۰/۰۱۷۳۴	۰/۸۱۰۸۱	۰/۰۲۱۷۸	۰/۷۲۵۱۳	۰/۰۱۸۷۵	۰/۸۳۸۰۶	۰/۸	
۰/۳۰۷۲۸۹		۰/۳۴۴۴۳۲		۰/۳۱۲۲۵۷		۰/۳۷۰۹۵۳		MSE	
۰/۰۲۴۴۵	۰/۰۷۳۳۱	۰/۰۲۸۶۶	۰/۰۰۰۰۴	۰/۰۲۵۳۰	۰/۰۷۵۶۵	۰/۰۲۹۶۸	−۰/۰۰۰۱۶	۰/۰	۰/۵
۰/۰۱۹۲۶	۰/۱۳۵۶۲	۰/۰۲۴۲۴	۰/۰۹۳۸۱	۰/۰۲۰۰۶	۰/۱۴۰۲۲	۰/۰۲۵۰۸	۰/۰۹۶۸۵	۰/۱	
۰/۰۱۸۸۳	۰/۲۱۶۹۷	۰/۰۲۴۴۵	۰/۱۹۰۶۵	۰/۰۱۹۶۱	۰/۲۲۴۳۸	۰/۰۲۵۲۳	۰/۱۹۷۰۹	۰/۲	
۰/۰۲۵۱۱	۰/۳۳۴۷۳	۰/۰۳۰۸۹	۰/۳۴۸۲۷	۰/۰۲۵۳۰	۰/۳۴۵۸۶	۰/۰۳۱۵۱	۰/۳۵۹۶۶	۰/۴	
۰/۰۲۵۴۱	۰/۶۸۳۳۴	۰/۰۲۲۰۵	۰/۸۱۵۲۱	۰/۰۲۳۹۹	۰/۷۰۶۳۶	۰/۰۲۳۶۳	۰/۸۴۲۴۹	۰/۸	
۰/۴۴۰۷۲۱		۰/۵۵۴۵۸		۰/۴۴۲۴۵۹		۰/۵۹۵۰۸۶		MSE	
۰/۰۲۹۷۳	۰/۱۱۳۶۰	۰/۰۳۸۳۱	۰/۰۱۹۱۹	۰/۰۳۰۷۳	۰/۱۱۷۲۳	۰/۰۳۹۶۵	۰/۰۱۹۵۲	۰/۰	۰/۷
۰/۰۲۳۹۴	۰/۱۵۶۲۵	۰/۰۳۲۳۳	۰/۰۹۰۶۲	۰/۰۲۴۹۳	۰/۱۶۱۵۷	۰/۰۳۳۴۳	۰/۰۹۳۵۰	۰/۱	
۰/۰۲۲۲۶	۰/۲۱۸۴۵	۰/۰۳۱۶۴	۰/۱۷۴۹۴	۰/۰۲۳۱۴	۰/۲۲۵۹۳	۰/۰۳۲۶۰	۰/۱۸۰۸۶	۰/۲	
۰/۰۲۹۰۴	۰/۳۱۷۲۰	۰/۰۳۸۵۲	۰/۳۳۹۴۶	۰/۰۲۹۲۷	۰/۳۲۷۷۶	۰/۰۳۹۴۰	۰/۳۵۰۵۷	۰/۴	
۰/۰۳۴۳۹	۰/۶۳۷۸۷	۰/۰۳۰۹۸	۰/۸۲۳۷۴	۰/۰۳۳۱۱	۰/۶۵۹۱۵	۰/۰۳۲۷۵	۰/۸۵۱۱۱	۰/۸	
۰/۶۸۶۱۱۱		۰/۹۵۸۵۱۷		۰/۶۹۱۹۵۲		۱/۰۲۶۲۲۴		MSE	
۰/۰۳۷۰۴	۰/۲۱۰۴۰	۰/۰۵۹۲۹	۰/۰۶۱۷۲	۰/۰۳۸۰۹	۰/۲۱۳۵۶	۰/۰۶۱۳۸	۰/۰۶۳۳۶	۰/۰	۰/۹
۰/۰۲۱۹۲	۰/۲۱۳۵۱	۰/۰۴۸۴۱	۰/۰۵۲۸۹	۰/۰۲۳۰۲	۰/۲۱۶۸۴	۰/۰۵۰۰۰	۰/۰۵۴۲۹	۰/۱	
۰/۰۲۰۶۸	۰/۱۹۸۱۱	۰/۰۴۹۸۶	۰/۱۵۱۳۴	۰/۰۲۱۷۹	۰/۲۰۱۰۹	۰/۰۵۱۳۷	۰/۱۵۶۷۲	۰/۲	
۰/۰۲۷۲۸	۰/۲۸۸۹۳	۰/۰۵۳۶۰	۰/۳۰۷۲۹	۰/۰۲۷۹۹	۰/۲۹۴۸۲	۰/۰۵۴۹۳	۰/۳۱۷۲۷	۰/۴	
۰/۰۵۲۹۴	۰/۵۱۴۳۱	۰/۰۵۰۵۹	۰/۸۷۴۳۰	۰/۰۵۲۸۶	۰/۵۲۷۵۴	۰/۰۵۳۲۶	۰/۹۰۳۵۰	۰/۸	
۱/۰۸۳۴۲۴		۲/۲۲۶۸۵۸		۱/۱۰۹۸۶۹		۲/۳۸۷۱۸۴		MSE	

جدول ۶.۲: نتایج مطالعه شبیه‌سازی برآوردگر آزمون اولیه در مدل آمیخته خطی ریج

		PTMR, α =۰/۰۰۱		PTMR, α =۰/۰۱		PTMR, α =۰/۰۵		PTMR, α =۰/۱	
ρ	β	برآورد	انحراف استاندارد	برآورد	انحراف استاندارد	برآورد	انحراف استاندارد	برآورد	انحراف استاندارد
۰/۰	۰/۰	۰/۰۱۲۴۶	۰/۰۰۸۵۶	۰/۰۳۶۰۵	۰/۰۱۵۷۸	۰/۰۳۸۲۸	۰/۰۱۷۰۸	۰/۰۳۸۲۸	۰/۰۱۷۰۸
	۰/۱	۰/۰۲۷۳۱	۰/۰۱۳۹۳	۰/۱۰۷۲۸	۰/۰۱۴۶۶	۰/۱۱۹۰۷	۰/۰۱۵۹۳	۰/۱۱۹۰۷	۰/۰۱۵۹۳
	۰/۲	۰/۰۵۵۸۲	۰/۰۲۶۴۷	۰/۱۸۲۲۶	۰/۰۱۷۶۴	۰/۲۱۵۱۹	۰/۰۱۵۵۹	۰/۲۱۵۱۹	۰/۰۱۵۵۹
	۰/۴	۰/۰۹۸۱۵	۰/۰۴۹۴۳	۰/۳۲۸۵۸	۰/۰۲۶۳۲	۰/۳۷۰۱۵	۰/۰۱۸۷۰	۰/۳۷۰۱۵	۰/۰۱۸۷۰
	۰/۸	۰/۱۹۵۱۸	۰/۰۹۷۸۲	۰/۶۶۱۴۳	۰/۰۴۳۱۷	۰/۷۵۰۶۳	۰/۰۱۸۶۸	۰/۷۵۰۶۳	۰/۰۱۸۶۸
MSE		۲/۵۵۹۹۶		۰/۵۵۹۰۱		۰/۲۴۳۳۴		۰/۲۴۳۳۴	
۰/۲	۰/۰	۰/۰۱۸۰۴	۰/۰۱۰۱۰	۰/۰۵۷۵۰	۰/۰۱۸۵۳	۰/۰۶۱۱۲	۰/۰۲۰۱۴	۰/۰۶۱۱۲	۰/۰۲۰۱۴
	۰/۱	۰/۰۳۰۳۹	۰/۰۱۴۲۶	۰/۱۱۳۵۵	۰/۰۱۶۲۱	۰/۱۲۶۵۵	۰/۰۱۷۲۸	۰/۱۲۶۵۵	۰/۰۱۷۲۸
	۰/۲	۰/۰۶۳۹۳	۰/۰۲۶۳۰	۰/۱۸۶۲۵	۰/۰۱۸۲۲	۰/۲۱۹۷۱	۰/۰۱۶۴۸	۰/۲۱۹۷۱	۰/۰۱۶۴۸
	۰/۴	۰/۱۰۲۳۰	۰/۰۴۹۰۱	۰/۳۱۸۰۸	۰/۰۲۷۵۷	۰/۳۵۹۵۹	۰/۰۲۰۶۴	۰/۳۵۹۵۹	۰/۰۲۰۶۴
	۰/۸	۰/۲۰۳۸۵	۰/۰۹۶۶۸	۰/۶۳۹۰۵	۰/۰۴۴۴۵	۰/۷۲۵۱۳	۰/۰۲۱۷۸	۰/۷۲۵۱۳	۰/۰۲۱۷۸
MSE		۲/۵۰۷۰۷		۰/۶۲۴۱۳		۰/۳۱۲۲۶		۰/۳۱۲۲۶	
۰/۵	۰/۰	۰/۰۲۵۲۰	۰/۰۱۱۲۰	۰/۰۷۴۳۴	۰/۰۲۲۲۲	۰/۰۷۵۶۵	۰/۰۲۵۳۰	۰/۰۷۵۶۵	۰/۰۲۵۳۰
	۰/۱	۰/۰۲۹۸۸	۰/۰۱۵۲۱	۰/۱۲۴۵۸	۰/۰۱۸۷۳	۰/۱۴۰۲۲	۰/۰۲۰۰۶	۰/۱۴۰۲۲	۰/۰۲۰۰۶
	۰/۲	۰/۰۵۹۷۰	۰/۰۲۷۴۶	۰/۱۸۷۱۵	۰/۰۲۰۴۵	۰/۲۲۴۳۸	۰/۰۱۹۶۱	۰/۲۲۴۳۸	۰/۰۱۹۶۱
	۰/۴	۰/۰۹۴۴۵	۰/۰۴۹۹۰	۰/۳۱۰۲۵	۰/۰۳۰۵۴	۰/۳۴۵۸۶	۰/۰۲۵۳۰	۰/۳۴۵۸۶	۰/۰۲۵۳۰
	۰/۸	۰/۱۷۹۸۴	۰/۰۹۷۸۹	۰/۶۱۸۵۱	۰/۰۴۵۳۸	۰/۷۰۶۳۶	۰/۰۲۳۹۹	۰/۷۰۶۳۶	۰/۰۲۳۹۹
MSE		۲/۶۱۲۵۲		۰/۷۲۸۰۵		۰/۴۴۲۴۶		۰/۴۴۲۴۶	
۰/۷	۰/۰	۰/۰۳۷۵۹	۰/۰۱۵۸۱	۰/۱۱۱۶۳	۰/۰۲۷۸۰	۰/۱۱۷۲۳	۰/۰۳۰۷۳	۰/۱۱۷۲۳	۰/۰۳۰۷۳
	۰/۱	۰/۰۳۹۶۳	۰/۰۱۶۷۵	۰/۱۴۱۱۷	۰/۰۲۳۵۱	۰/۱۶۱۵۷	۰/۰۲۴۹۳	۰/۱۶۱۵۷	۰/۰۲۴۹۳
	۰/۲	۰/۰۶۸۰۱	۰/۰۲۸۱۹	۰/۱۸۵۶۸	۰/۰۲۳۳۲	۰/۲۲۵۹۳	۰/۰۲۳۱۴	۰/۲۲۵۹۳	۰/۰۲۳۱۴
	۰/۴	۰/۰۹۹۱۷	۰/۰۴۹۸۲	۰/۲۹۲۲۹	۰/۰۳۴۱۰	۰/۳۲۷۷۶	۰/۰۲۹۲۷	۰/۳۲۷۷۶	۰/۰۲۹۲۷
	۰/۸	۰/۲۰۳۷۹	۰/۰۹۵۹۳	۰/۵۸۳۴۷	۰/۰۴۹۸۷	۰/۶۵۹۱۵	۰/۰۳۳۱۱	۰/۶۵۹۱۵	۰/۰۳۳۱۱
MSE		۲/۵۵۰۰۱		۰/۹۵۵۹۲		۰/۶۹۱۹۵		۰/۶۹۱۹۵	
۰/۹	۰/۰	۰/۰۵۱۳۶	۰/۰۱۸۵۹	۰/۱۷۹۴۹	۰/۰۳۴۸۶	۰/۲۱۳۵۶	۰/۰۳۸۰۹	۰/۲۱۳۵۶	۰/۰۳۸۰۹
	۰/۱	۰/۰۶۲۸۳	۰/۰۱۷۸۰	۰/۱۸۴۶۱	۰/۰۲۲۴۱	۰/۲۱۶۸۴	۰/۰۲۳۰۲	۰/۲۱۶۸۴	۰/۰۲۳۰۲
	۰/۲	۰/۰۴۷۳۷	۰/۰۲۷۹۵	۰/۱۵۷۸۳	۰/۰۲۳۴۶	۰/۲۰۱۰۹	۰/۰۲۱۷۹	۰/۲۰۱۰۹	۰/۰۲۱۷۹
	۰/۴	۰/۰۷۶۳۳	۰/۰۵۰۷۳	۰/۲۵۳۸۸	۰/۰۳۴۵۳	۰/۲۹۴۸۲	۰/۰۲۷۹۹	۰/۲۹۴۸۲	۰/۰۲۷۹۹
	۰/۸	۰/۱۴۸۷۲	۰/۱۰۱۰۴	۰/۴۷۴۲۲	۰/۰۶۳۷۶	۰/۵۲۷۵۴	۰/۰۵۲۸۶	۰/۵۲۷۵۴	۰/۰۵۲۸۶
MSE		۲/۸۳۲۴۴		۱/۳۷۲۷۱		۱/۱۰۹۸۷		۱/۱۰۹۸۷	

همان‌طور که از نتایج جدول‌های ۵.۲ و ۶.۲ برمی‌آید، برآوردگر $PTMR$ به سطح معنی‌داری α وابسته است در حالی‌که این امر در خصوص برآوردگرهای $SSMR$ صادق نیست. آنچه در تحلیل نتایج جدول‌ها حائز اهمیت است این است که برآوردگر $SSMR$ در بین تمامی برآوردگرهای مورد مطالعه دارای میانگین توان دوم خطا (MSE) و انحراف استاندارد (sd) کمتری است. همچنین واضح است که هرچه میزان همخطی در داده‌ها قوی‌تر باشد، MSE برآورد در برآوردگر $SSMR$ نسبت به دیگر برآوردگرها به میزان قابل ملاحظه‌ای کمتر است. به عبارت دیگر با افزایش همبستگی بین متغیرهای پیش‌بین، سرعت افزایش میزان MSE در برآوردگر $SSMR$ کمتر از دیگر برآوردگرها است. در ادامه روش‌های برآوردیابی ذکر شده را برای تحلیل یک مثال واقعی به کار می‌گیریم.

۲.۴.۲ مثال واقعی

لنفوسیت‌ها[۳۵] گروهی از گلبول‌های سفید خون می‌باشند که نقش مهمی را در اجرای پاسخ‌های ایمنی بدن بر ضد عوامل بیگانه به عهده دارند. این سلول‌ها به دو گروه نوع T و B تقسیم می‌شوند که لنفوسیت‌های نوع T شامل دو گروه CD4 و CD8 هستند. سلول‌های CD4 را عموماً با عنوان سلول‌های T کمک کننده[۳۶] می‌شناسند که نقش مهمی را در دفاع علیه بسیاری از عوامل عفونی مانند باکتری‌ها، قارچ‌ها و حتی عوامل غیر عفونی دارند. این سلول‌ها مهمترین گروه سلولی تحت تأثیر ویروس HIV[۳۷] می‌باشند. ویروس HIV وارد این گروه از لنفوسیت‌ها شده و در این سلول‌ها شروع به همانند سازی کرده و در نهایت باعث از بین رفتن و کاهش تعداد آن‌ها می‌شود. هدف اصلی در درمان بیماران مبتلا به ویروس HIV، کند کردن روند کاهش سلول‌های CD4 و ثابت نگهداشتن آن‌ها در یک وضعیت، به منظور جلوگیری از ابتلای فرد به بیماری‌ها و عفونت‌های مختلف می‌باشد. از آن‌جایی که CD4 اولین گروه از سلول‌هایی هستند که تحت تأثیر ویروس HIV قرار می‌گیرند و تعداد آن‌ها سریعاً کاهش می‌یابد، اندازه‌گیری آن‌ها عامل مهمی برای ارزیابی میزان تأثیر درمان‌ها و پیشرفت بیماری به سوی فاز نهایی ایدز[۳۸] می‌باشد. بنابراین لازم است که همزمان با درمان در این افراد، تعداد سلول‌های CD4 به روش‌های خاص خود، اندازه‌گیری شود.

زالسیتابین[۳۹] و دیدانوزین[۴۰]، از جمله داروهای ضد ویروسی است که برای مقابله با ویروس HIV در فرد مبتلا، استفاده می‌شود. این داروها روند پیشروی ویروس HIV را کند می‌کنند و در نتیجه سبب تقویت سیستم ایمنی بدن شده و خطر ایجاد عفونت‌های وابسته به HIV را کاهش می‌دهند. هدف در این مطالعه تأثیر این دو نوع دارو بر کنترل ویروس HIV در بدن افراد مبتلا، مورد بررسی قرار می‌گیرد.

این مجموعه داده در بسته‌ی JMbayes موجود در نرم‌افزار R وجود دارد که شامل ۷ متغیر پیش‌بین برای تعداد $n = ۴۶۷$ واحد آزمایشی گردآوری شده که در این مطالعه متغیرهای prevOI و gender را با توجه به آنچه در فصل اول گفته شد به عنوان اثرات تصادفی در نظر گرفته‌ایم.

توضیحات مربوط به متغیرهای این مجموعه داده در جدول‌های ۷.۲ و ۸.۲ آمده است.

[35] Lymphocytus
[36] T-helper
[37] Human immunodeficiency virus
[38] Aids: Acquired immune deficiency syndrome
[39] Zalcitabine
[40] Didanosine

جدول ۷.۲: توصیف متغیرهای مجموعه داده بخش ۲.۴.۲

متغیرها	توضیحات
patient	کد شناسایی مربوط به شخص مورد آزمایش
Time	زمان مرگ یا سانسور
death	۰ نشان دهنده سانسور و ۱ نشان دهنده مرگ
CD۴	تعداد سلول‌های نوع CD4 هر فرد
obstime	زمان ثبت تعداد سلول‌های CD4 (با اعمال تبدیل برای پیوسته شدن متغیر)
drug	عاملی با دو سطح *ddC* و *ddI* به ترتیب نشان دهنده تزریق داروی زالسیتابین و دیدانوزین به شخص مورد آزمایش
gender	جنسیت شخص مورد آزمایش
prevOI	عاملی با دو سطح *AIDS* و *noAIDS* به جهت نمایش نتیجه آزمایش قبلی
AZT	عاملی با دو سطح *intolerance* و *failure* برای تفکیک اشخاص از لحاظ مقاومت بدنی

جدول ۸.۲: اطلاعات خلاصه شده از متغیرهای مجموعه داده بخش ۲.۴.۲

	Time	*CD4*	*obstime*
مینیمم	۰/۴۷	۰/۰۰	۰/۰۰
چارک اول	۱۲/۲۳	۳/۱۶	۰/۰۰
میانه	۱۴/۰۷	۵/۴۷	۲/۰۰
میانگین	۱۳/۸۹	۷/۰۲	۴/۲۱
چارک سوم	۱۷/۰۰	۱۰/۴۴	۶/۰۰
ماکزیمم	۲۱/۴۰	۲۴/۱۲	۱۸/۰۰

death	*drug*	*gender*	*prevOI*	*AZT*
death: ۴۱۲	ddI: ۶۸۸	male: ۱۲۸۸	AIDS: ۸۶۳	failure: ۴۹۱
censoring: ۹۹۳	ddC: ۷۱۷	female: ۱۱۷	noAIDS: ۵۴۲	intolerance: ۹۱۴

اندازه‌گیری هر کدام از متغیرها به دفعات برای هر واحد آزمایشی iام تکرار شده‌است. لازم به ذکر است که در این مطالعه متغیر $CD4$ به عنوان متغیر پاسخ در نظر گرفته شده است.

در جدول ۹.۲ برآوردگرهای آمیخته، آمیخته ریج، آمیخته نوع استاین، آمیخته ریج نوع استاین و آزمون اولیه آمیخته ریج حاصل از تحلیل این مجموعه داده به ترتیب با $\widehat{\beta}_M$، $\widehat{\beta}_{MR}$، $\widehat{\beta}^S_M$، $\widehat{\beta}^S_{MR}$ و $\widehat{\beta}^{PT}_{MR}$ نشان داده شده‌اند. برای بررسی عملکرد این برآوردگرها در تحلیل مثال واقعی، میانگین خطای پیش‌گویی[۴۱] (MPE) هر برآوردگر نیز محاسبه شده است که هرچه مقدار آن کمتر باشد نشان‌دهنده‌ی برتری برآوردگر است. طریقه محاسبه‌ی معیار MPE در ادامه آمده است.

به منظور محاسبه MPE با روش اعتبارسنجی متقابل[۴۲] K-fold، مجموعه داده را به‌طور تصادفی به K زیرمجموعه با اندازه‌های برابر تقسیم می‌کنیم، به ترتیب هر زیرمجموعه را به عنوان داده‌های آزمون $\{(\mathbf{X}^{test}, \mathbf{y}^{test})\}$ و مابقی را به عنوان داده‌های آزمایش در نظر می‌گیریم و برای مدل‌بندی از آن‌ها استفاده می‌کنیم. برآورد حاصل مدل‌بندی داده‌های آزمایشی را $\widehat{\beta}_{train}$ می‌نامیم سپس مدل برازش شده را به جهت پیش‌گویی پاسخ مجموعه داده آزمون به کار می‌گیرم. در نهایت خطای پیش‌گویی (PE) را برای زیرمجموعه‌ی $k = ۱, \ldots, K$ام به صورت زیر محاسبه می‌کنیم

$$PE^k = \left\| \mathbf{X}^{test}_k \boldsymbol{\beta} - \widehat{\mathbf{y}}^{test}_k \right\|^۲$$

که در آن $\widehat{\mathbf{y}}^{test}_k = \widehat{\mathbf{X}}^{test}_k \widehat{\beta}^{train}_k$ زیرمجموعه‌ها محاسبه می‌کنیم. این فرآیند را $N = ۵۰۰$ بار تکرار کرده و مقدار MPE را از رابطه‌ی زیر بدست می‌آوریم

$$\text{MPE} = \text{median} \left\{ \frac{۱}{K} \sum_{k=۱}^{K} PE^k_۱, \ldots, \frac{۱}{K} \sum_{k=۱}^{K} PE^k_N \right\}$$

که در آن PE^k_i خطای پیش‌گویی kامین مجموعه‌ی آزمون در تکرار iام است. مقدار MPE مربوط به هر برآوردگر در جدول ۹.۲ گزارش شده است.

همان‌طور که در جدول ۹.۲ ملاحظه می‌شود مقدار MPE برای برآوردگر آمیخته ریج نوع استاین $\beta^{\hat{S}}_{MR}$ کمترین است.

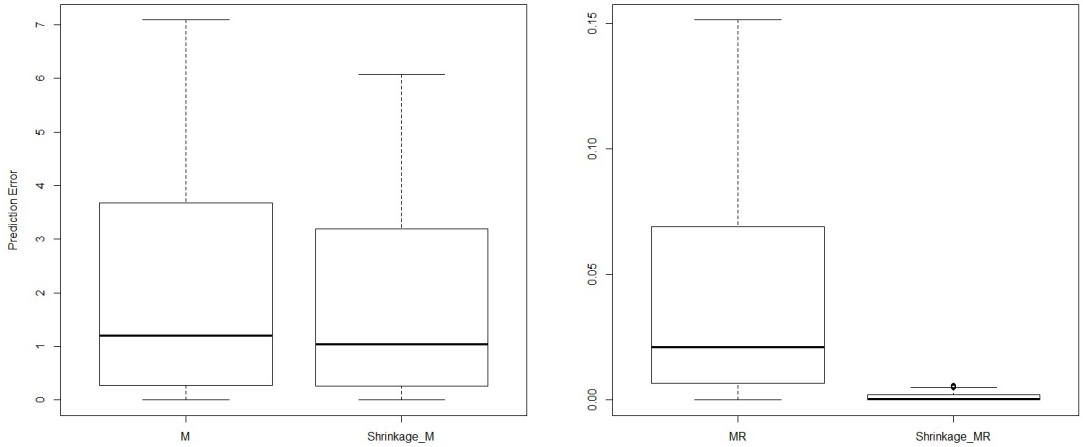
جدول ۹.۲: برآورد ضرایب در داده‌های بخش ۲.۴.۲

	$\hat{\beta}_M$	$\hat{\beta}_{MR}$	$\hat{\beta}^S_M$	$\hat{\beta}^S_{MR}$	$\hat{\beta}^{PT}_{MR}(\alpha > 0.2)$
Time	۰/۰۶۸۶۹۶۹۸	۰/۲۰۷۳۵۸۳	۰/۰۲۸۵۶۸۹۷	۰/۰۸۰۳۱۰۵۰	۰/۲۰۷۳۵۸۳
death	−۲/۰۴۹۱۸۵۰۹	−۱/۰۰۹۳۶۷۶	−۰/۸۵۲۱۹۳۳۸	−۰/۳۹۰۹۳۱۰۹	−۱/۰۰۹۳۶۷۶
obstime	−۰/۱۴۹۸۹۰۴۰	−۰/۱۴۳۴۹۹۱	−۰/۰۶۲۳۳۴۸۳	−۰/۰۵۵۵۷۷۶۳	−۰/۱۴۳۴۹۹۱
drug	۴/۵۲۵۳۱۱۹۹	۳/۳۳۳۵۶۹۲	۱/۸۸۱۹۳۸۷۹	۱/۲۹۱۱۰۱۳۷	۳/۳۳۳۵۶۹۲
AZT	۰/۶۶۳۸۱۸۹۲	۰/۱۷۲۴۱۱۸	۰/۲۷۶۰۶۱۹۸	۰/۰۶۶۷۷۵۶۰	۰/۱۷۲۴۱۱۸
MPE	۱/۲۰۶۸۷	۰/۰۲۱۱۴	۱/۰۴۶۴۴	۰/۰۰۰۵۲	۰/۰۲۱۱۴

[41] Mean prediction error

[42] Cross validation

شکل ۵.۲ نمودار جعبه‌ای مربوط به مقدار خطای پیش‌گویی برآوردگرها در هر مرحله از تکرار فرآیند محاسبه MPE را نشان می‌دهد که برتری برآوردگر آمیخته ریج نوع استاین را از نظر این معیار تایید می‌کند.



شکل ۵.۲: نمودار جعبه‌ای مربوط به مقدار MPE برآوردها

جدول ۱۰.۲ مربوط به مقدار انحراف استاندارد برآوردها است که همچنان نشان دهنده برتری برآوردگر آمیخته ریج نوع استاین نسبت به دیگر برآوردگرها است.

جدول ۱۰.۲: انحراف استاندارد برآوردهای جدول ۹.۲

	$sd(\boldsymbol{\beta}_M)$	$sd(\boldsymbol{\beta}_{MR})$	$sd(\boldsymbol{\beta}_M^S)$	$sd(\boldsymbol{\beta}_{MR}^S)$	$sd(\boldsymbol{\beta}_{MR}^{PT})$
Time	۰/۰۰۰۶۱۳۰۱۲۷	۱/۶۶۶۷۴۹ × ۱۰$^{-۰۴}$	۰/۰۰۰۵۷۹۰۵۹۲	۹/۲۳۵۰۲۷ × ۱۰$^{-۰۵}$	۱/۶۶۶۷۴۹ × ۱۰$^{-۰۴}$
death	۰/۰۱۵۷۲۱۵۵۹۶	۱/۱۷۰۷۷۷ × ۱۰$^{-۰۲}$	۰/۰۱۵۲۳۳۵۸۶۶	۳/۴۷۸۲۸۰ × ۱۰$^{-۰۴}$	۱/۱۷۰۷۷۷ × ۱۰$^{-۰۲}$
obstime	۰/۰۰۱۱۲۶۲۴۳۱	۴/۹۰۸۷۹۹ × ۱۰$^{-۰۴}$	۰/۰۰۱۰۷۳۴۲۲۷	۸/۷۷۹۷۲۲ × ۱۰$^{-۰۵}$	۴/۹۰۸۷۹۹ × ۱۰$^{-۰۴}$
drug	۰/۰۳۷۴۳۸۵۵۶۷	۱/۶۳۷۰۱۹ × ۱۰$^{-۰۲}$	۰/۰۳۴۴۴۶۰۱۲۹	۴/۵۸۹۶۹۴ × ۱۰$^{-۰۴}$	۱/۶۳۷۰۱۹ × ۱۰$^{-۰۲}$
AZT	۰/۰۰۴۴۴۲۴۲۳۳	۴/۲۳۴۵۳۴ × ۱۰$^{-۰۳}$	۰/۰۰۴۱۲۱۵۷۲۲	۸/۹۷۷۰۰۲ × ۱۰$^{-۰۴}$	۴/۲۳۴۵۳۴ × ۱۰$^{-۰۳}$

۵.۲ مدل آمیخته خطی تعمیم‌یافته ریج

همان‌طور که در بخش ۵.۱ گفته شد، گاهی متغیر پاسخ از توزیع نرمال تبعیت نمی‌کند، بنابراین در چنین مواردی استفاده از مدل‌های خطی ممکن نخواهد بود. اما اگر توزیع متغیر پاس متعلق به خانواده توزیع‌های نمایی باشد، در این صورت

$$f_{\mathbf{y}_i}(\mathbf{y}_i) = \exp\left\{(\mathbf{y}_i\gamma_i - b(\gamma_i))/\tau^2 - c(\mathbf{y}_i, \tau)\right\} \qquad (۲۳.۲)$$

لذا لگاریتم تابع درستنمایی به صورت زیر است

$$\ell(\boldsymbol{\beta}) = \sum_{i=1}^{n} \left[\mathbf{y}_i\gamma_i - b(\gamma_i)\right]/\tau^2 - \sum_{i=1}^{n} c(\mathbf{y}_i, \tau). \qquad (۲۴.۲)$$

مدل خطی تعمیم‌یافته ریج

برای بدست آوردن برآوردگر خطی تعمیم‌یافته ریج[۴۳] (GLR) می‌توان لگاریتم تابع درستنمایی جریمه شده را به صورت زیر نوشت

$$\ell^{pen}(\boldsymbol{\beta}) = 2\left(\sum_{i=1}^{n} \left[\mathbf{y}_i\gamma_i - b(\gamma_i)\right]/\tau^2 - \sum_{i=1}^{n} c(\mathbf{y}_i, \tau)\right) - \lambda\boldsymbol{\beta}^T\boldsymbol{\beta}. \qquad (۲۵.۲)$$

$$\begin{aligned}
\frac{\partial \ell^{pen}(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}} &= \frac{\partial \ell^{pen}(\boldsymbol{\beta})}{\partial \gamma_i}.\frac{\partial \gamma_i}{\partial \mu_i}.\frac{\partial \mu_i}{\partial g(\mu_i)}.\frac{\partial g(\mu_i)}{\partial \boldsymbol{\beta}} \\
&= \frac{2}{\tau^2}\sum\left(\mathbf{y}_i - \frac{\partial b(\gamma_i)}{\partial \gamma_i}.\frac{1}{\nu(\mu_i)}.\frac{\partial \mu_i}{\partial g(\mu_i)}.\mathbf{X}_i^T\right) - 2\lambda\boldsymbol{\beta} \\
&= \frac{2}{\tau^2}\sum\left(\mathbf{y}_i - \mu_i.\frac{1}{\nu(\mu_i)}.\frac{\partial \mu_i}{\partial g(\mu_i)}\frac{\partial g(\mu_i)}{\partial \mu_i}\frac{\partial \mu_i}{\partial g(\mu_i)}.\mathbf{X}_i^T\right) - 2\lambda\boldsymbol{\beta} \\
&= 2\sum\left(w_i\mathbf{X}_i^T(t_i - g(\mu_i))\right) - 2\lambda\boldsymbol{\beta} \qquad (۲۶.۲)
\end{aligned}$$

که در آن $w_i = \left[\tau^2\nu(\mu_i)g_\mu^2(\mu_i)\right]^{-1}$ و $t_i = g(\mu_i) + g_\mu(\mu_i)(\mathbf{y}_i - \mu_i)$.

شکل ماتریسی رابطه (۲۶.۲) را می‌توان به صورت زیر نوشت

$$\frac{\partial \ell^{pen}(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}} = 2\mathbf{X}^T\mathbf{W}(\mathbf{T} - \mathbf{X}\boldsymbol{\beta}) - 2\lambda\boldsymbol{\beta} \qquad (۲۷.۲)$$

با مساوی صفر قرار دادن رابطه (۲۷.۲) داریم

$$\begin{aligned}
0 = \frac{\partial \ell^{pen}(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}} &= \lambda\boldsymbol{\beta} - \left(\mathbf{X}^T\mathbf{W}(T - \mathbf{X}\boldsymbol{\beta})\right) \\
&= \lambda\boldsymbol{\beta} - \mathbf{XWT} + (\mathbf{X}^T\mathbf{WX})\boldsymbol{\beta} \\
&= (\mathbf{X}^T\mathbf{WX} + \lambda\mathbf{I}_p)\boldsymbol{\beta} - (\mathbf{X}^T\mathbf{WT})
\end{aligned}$$

در نتیجه برآوردگر GLR به صورت زیر حاصل می‌شود

$$\hat{\boldsymbol{\beta}}_{GLR} = (\mathbf{X}^T\mathbf{WX} + \lambda\mathbf{I})^{-1}(\mathbf{X}^T\mathbf{WT}) \qquad (۲۸.۲)$$

که در آن $\mathbf{W} = \{_n w_i\}$ ماتریس nـبعدی با عناطر قطری

$$w_i = \left[\tau^2\nu(\mu_i)g_\mu^2(\mu_i)\right]^{-1}$$

[43] Generalized linear ridge

و $\mathbf{T} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\Delta}(\mathbf{y} - \boldsymbol{\mu})$ بردار $n \times 1$ وابسته به پارامتر $\boldsymbol{\beta}$، با دارایه‌های

$$t_i = g(\mu_i) + g_\mu(\mu_i)(\mathbf{y}_i - \mu_i)$$

است.

از طرفی با تعریف بردار $\mathbf{T}$، رابطه (۵۲.۱) را می‌توان به صورت زیر بازنویسی کرد

$$\begin{aligned}
\hat{\boldsymbol{\beta}}_{GL}^{(m+1)} &= \hat{\boldsymbol{\beta}}_{GL}^{(m)} + (\mathbf{X}^T\mathbf{W}\mathbf{X})^{-1}\mathbf{X}^T\mathbf{W}\boldsymbol{\Delta}(\mathbf{y} - \boldsymbol{\mu}) \\
&= (\mathbf{X}^T\mathbf{W}\mathbf{X})^{-1}\mathbf{X}^T\mathbf{W}\left[\mathbf{X}\hat{\boldsymbol{\beta}}_{GL}^{(m)} + \boldsymbol{\Delta}(\mathbf{y} - \boldsymbol{\mu})\right] \\
&= (\mathbf{X}^T\mathbf{W}\mathbf{X})^{-1}\mathbf{X}^T\mathbf{W}\mathbf{T}^{(m)}. \qquad (۲۹.۲)
\end{aligned}$$

در نتیجه با استفاده از رابطه‌های (۲۸.۲) و (۲۹.۲) داریم

$$\begin{aligned}
\hat{\boldsymbol{\beta}}_{GLR} &= (\mathbf{X}^T\mathbf{W}\mathbf{X} + \lambda\mathbf{I})^{-1}(\mathbf{X}^T\mathbf{W}\mathbf{T}) \\
&= (\mathbf{X}^T\mathbf{W}\mathbf{X} + \lambda\mathbf{I})^{-1}(\mathbf{X}^T\mathbf{W}\mathbf{X})(\mathbf{X}^T\mathbf{W}\mathbf{X})^{-1}(\mathbf{X}^T\mathbf{W}\mathbf{T}) \\
&= (\mathbf{X}^T\mathbf{W}\mathbf{X} + \lambda\mathbf{I})^{-1}(\mathbf{X}^T\mathbf{W}\mathbf{X})\hat{\boldsymbol{\beta}}_{GL}. \qquad (۳۰.۲)
\end{aligned}$$

برای آگاهی بیشتر در مورد برآورد پارامترهای مدل خطی تعمیم‌یافته در رگرسیون ریج به سجرستد (۱۹۹۲) مراجعه کنید.

مدل آمیخته خطی تعمیم‌یافته ریج

حال برای بدست آوردن برآوردگر آمیخته خطی تعمیم‌یافته ریج[۴۴] $(GLMR)$ ، لگاریتم تابع درستنمایی جریمه شده طبق رابطه‌های (۶۰.۱) و (۲۶.۲) به صورت زیر است

$$\ell^{pen}(\boldsymbol{\beta}) = \lambda\boldsymbol{\beta}^T\boldsymbol{\beta} - 2\ell(\boldsymbol{\beta}). \qquad (۳۱.۲)$$

با مشتق‌گیری از رابطه (۳۱.۲) نسبت به $\boldsymbol{\beta}$ داریم

$$\begin{aligned}
\frac{\partial\ell^{pen}(\boldsymbol{\beta})}{\partial\boldsymbol{\beta}} &= 2\lambda\boldsymbol{\beta} - 2\int \mathbf{X}^T\mathbf{W}(\mathbf{T} - \mathbf{X}\boldsymbol{\beta} - \mathbf{Z}\mathbf{b})f_{\mathbf{b}|\mathbf{y}}(\mathbf{b}|\mathbf{y})\mathrm{d}\mathbf{b} \\
&= \lambda\boldsymbol{\beta} + (\mathbf{X}^T E\left[\mathbf{W}|\mathbf{y}\right]\mathbf{X})\boldsymbol{\beta} - \left(\mathbf{X}^T E\left[\mathbf{W}(\mathbf{T} - \mathbf{Z}\mathbf{b})|\mathbf{y}\right]\right) \\
&= \left(\mathbf{X}^T E\left[\mathbf{W}|\mathbf{y}\right]\mathbf{X} + \lambda\mathbf{I}_p\right)\boldsymbol{\beta} - \left(\mathbf{X}^T E\left[\mathbf{W}(\mathbf{T} - \mathbf{Z}\mathbf{b})|\mathbf{y}\right]\right), \\
&\qquad (۳۲.۲)
\end{aligned}$$

در نتیجه برآوردگر $GLMR$ به صورت زیر حاصل می‌شود

$$\hat{\boldsymbol{\beta}}_{GLMR} = \left(\mathbf{X}^T E\left[\mathbf{W}|\mathbf{y}\right]\mathbf{X} + \lambda\mathbf{I}_p\right)^{-1}\left(\mathbf{X}^T E\left[\mathbf{W}(\mathbf{T} - \mathbf{Z}\mathbf{b})|\mathbf{y}\right]\right). \qquad (۳۳.۲)$$

[44] Generalized linear mixed ridge

از طرفی، بردار $\mathbf{T}$ به پارامتر $\boldsymbol{\beta}$ وابسته است لذا برآوردگر ارائه‌شده مشابه حالت GLR با استفاده از الگوریتم‌های تکرار شونده محاسبه خواهد شد. لذا

$$\hat{\boldsymbol{\beta}}_{GLMR}^{(m+1)} = \left(\mathbf{X}^T E\left[\mathbf{W}|\mathbf{y}\right]\mathbf{X} + \lambda \mathbf{I}_p\right)^{-1}\left(\mathbf{X}^T E\left[\mathbf{W}(\mathbf{T}^{(m)} - \mathbf{Z}\mathbf{b})|\mathbf{y}\right]\right). \qquad (۳۴.۲)$$

برای آگاهی بیشتر در مورد روش‌هایی مانند الگوریتم EM، الگوریتم‌های $MCMC$[۴۵]، تقریب تصادفی[۴۶] برای محاسبه برآورد ماکزیمم درستنمایی در مدل‌های خطی تعمیم‌یافته به مک کولاک و همکاران (۲۰۰۸) مراجعه کنید.

[45] Markov chain Monte Carlo

[46] Stochastic approximation

فصل ۳

رگرسیون لیو برای داده‌های طولی

۱.۳ مقدمه

همان‌طور که در فصل قبل بیان شد، وجود همخطی در داده‌ها سبب بزرگ شدن و فاصله گرفتن مقادیر برآورد شده از پارامتر واقعی می‌شود. این امر در فصل قبل با استفاده از رگرسیون ریج پوشانده شد، اما ایده‌ای دیگر برای مرتفع کردن این مشکل، توسط لیو (۱۹۹۳) مطرح شد که با قرار دادن ضریبی مانند d $(0 < d < 1)$، مقدار برآورد شده را به پارامتر واقعی نزدیک می‌کند. به خاطر بیاورید که برآوردگر ریج برحسب پارامتر آن یک تابع غیرخطی است. حال چنانچه بخواهیم این رابطه خطی باشد می‌توانیم این طور فرض کنیم از آن‌جایی که $\hat{\boldsymbol{\beta}}$ از $\boldsymbol{\beta}$ فاصله دارد و مقادیر آن به دلیل وجود همخطی، بزرگ است، $d\hat{\boldsymbol{\beta}}$ به ازای یک d مناسب و کوچک می‌تواند تا حد امکان به $\boldsymbol{\beta}$ نزدیک شود لذا به جای این‌که تابع جریمه را به صورت $\boldsymbol{\beta}^T\boldsymbol{\beta}$ در مدل وارد کنیم، بزرگی فاصله $d\boldsymbol{\beta}^T - \boldsymbol{\beta}$ را جریمه کرده و تابع جریمه را به صورت

$$(d\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta})^T(d\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta})$$

در نظر می‌گیریم. در این صورت برآوردگر حاصل بر اساس مدل (۱.۲) عبارت است از

$$\widehat{\boldsymbol{\beta}}_{Liu} = \arg\min_{\boldsymbol{\beta}\in\mathbb{R}^p} \left\{(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})^T(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) + (d\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta})^T(d\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta})\right\} \qquad (۱.۳)$$

با مشتق‌گیری از عبارت داخل آکولاد در رابطه (۱.۳) نسبت به $\boldsymbol{\beta}$ داریم

$$
\begin{aligned}
& \frac{d}{d\boldsymbol{\beta}}\left[(\mathbf{y}-\mathbf{X}\boldsymbol{\beta})^T(\mathbf{y}-\mathbf{X}\boldsymbol{\beta})+(d\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})^T(d\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})\right] \\
= & \frac{d}{d\boldsymbol{\beta}}\left[\mathbf{y}^T\mathbf{y}-۲\boldsymbol{\beta}^T\mathbf{X}^T\mathbf{y}+\boldsymbol{\beta}^T\mathbf{X}^T\mathbf{X}\boldsymbol{\beta}+d^۲\hat{\boldsymbol{\beta}}^T\hat{\boldsymbol{\beta}}-۲d\boldsymbol{\beta}^T\hat{\boldsymbol{\beta}}+\boldsymbol{\beta}^T\boldsymbol{\beta}\right] \\
= & -۲\mathbf{X}^T\mathbf{y}+۲\mathbf{X}^T\mathbf{X}\boldsymbol{\beta}-۲d\hat{\boldsymbol{\beta}}+۲\boldsymbol{\beta} \\
= & -۲\mathbf{X}^T\mathbf{y}-۲d\hat{\boldsymbol{\beta}}+۲(\mathbf{X}^T\mathbf{X}+\mathbf{I}_p)\boldsymbol{\beta},
\end{aligned}
$$

که با مساوی صفر قرار دادن نتیجه می‌شود

$$
\begin{aligned}
\widehat{\boldsymbol{\beta}}_{Liu} &= (\mathbf{X}^T\mathbf{X}+\mathbf{I}_p)^{-۱}(d\hat{\boldsymbol{\beta}}+\mathbf{X}\mathbf{y}) \\
&= (\mathbf{X}^T\mathbf{X}+\mathbf{I}_p)^{-۱}\left(d(\mathbf{X}^T\mathbf{X})^{-۱}\mathbf{X}^T\mathbf{y}+\mathbf{X}^T\mathbf{y}\right) \\
&= (\mathbf{X}^T\mathbf{X}+\mathbf{I}_p)^{-۱}\left(d(\mathbf{X}^T\mathbf{X})^{-۱}\mathbf{X}^T\mathbf{y}+(\mathbf{X}^T\mathbf{X})(\mathbf{X}^T\mathbf{X})^{-۱}\mathbf{X}^T\mathbf{y}\right) \\
&= (\mathbf{X}^T\mathbf{X}+\mathbf{I}_p)^{-۱}(d\mathbf{I}_p+\mathbf{X}^T\mathbf{X})(\mathbf{X}^T\mathbf{X})^{-۱}\mathbf{X}\mathbf{y} \\
&= (\mathbf{X}^T\mathbf{X}+\mathbf{I}_p)^{-۱}(d\mathbf{I}_p+\mathbf{X}^T\mathbf{X})\hat{\boldsymbol{\beta}}. \qquad (۲.۳)
\end{aligned}
$$

آن‌چه مورد توجه است، خطی بودن برآوردگر نوع لیو نسبت به پارامتر d است. از این رو در این فصل، بررسی برآوردگر نوع لیو در مدل آمیخته خطی حائز اهمیت است. برای آگاهی بیشتر در خصوص برآوردگر لیو، برزویی بیدگلی (۱۳۹۳) را ببینید.

۲.۳ مدل آمیخته خطی لیو

مدل آمیخته خطی زیر را در نظر بگیرید

$$\mathbf{y}_i=\mathbf{X}_i\boldsymbol{\beta}+\mathbf{Z}_i^T\mathbf{b}_i+\boldsymbol{\varepsilon}_i \qquad (۳.۳)$$

که در آن $\mathbf{y}_i=(y_{i۱},y_{i۲},\ldots,y_{in_i})^T$ برداری از n_i مشاهده، $\mathbf{X}_i$ ماتریس طرح $n_i\times p$ از اثرات ثابت و $\boldsymbol{\beta}=(\beta_۱,\ldots,\beta_p)^T$ بردار $p\times ۱$ ضرایب ثابت و نامعلوم، و $\mathbf{Z}_i$ ماتریس طرح $q\times n_i$ از اثرات تصادفی، $\mathbf{b}_i\sim\mathcal{N}_q(۰,\mathbf{D})$ بردار $q\times ۱$ ضرایب تصادفی و $\boldsymbol{\epsilon}_i\sim\mathcal{N}_n(۰,\sigma^۲\mathbf{I})$ بردار $n\times ۱$ خطاهای تصادفی است. حال لگاریتم تابع درستنمایی را در رگرسیون لیو به‌صورت زیر داریم

(۴.۳)

$$\ell^{pen}(\boldsymbol{\beta},\mathbf{X};\mathbf{y})=-\frac{N}{۲}\log(۲\pi)-\frac{۱}{۲}\sum_{i=۱}^{n}\log|\mathbf{V}|-\frac{۱}{۲}(\mathbf{y}-\mathbf{X}\boldsymbol{\beta})^T\mathbf{V}^{-۱}(\mathbf{y}-\mathbf{X}\boldsymbol{\beta})-(d\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta})^T(d\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta})$$

ماکزیمم کردن تابع (۴.۳) معادل با مینیمم کردن عبارت وابسته به پارامتر $\boldsymbol{\beta}$ است. لذا برآوردگر آمیخته لیو (MLiu) از رابطه زیر محاسبه می‌شود

$$\widehat{\boldsymbol{\beta}}_{MLiu}=\arg\min_{\boldsymbol{\beta}\in\mathbb{R}^p}\left\{(\mathbf{y}-\mathbf{X}\boldsymbol{\beta})^T\mathbf{V}^{-۱}(\mathbf{y}-\mathbf{X}\boldsymbol{\beta})+(d\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta})^T(d\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta})\right\} \qquad (۵.۳)$$

با مساوی صفر قرار دادن مشتق عبارت داخل آکولاد در رابطه (۵.۳) بر حسب $\boldsymbol{\beta}$ داریم

$$\begin{aligned} \circ &= \frac{d}{\partial\boldsymbol{\beta}}\left[(\mathbf{y}-\mathbf{X}\boldsymbol{\beta})^T\mathbf{V}^{-1}(\mathbf{y}-\mathbf{X}\boldsymbol{\beta})+(d\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})^T(d\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})\right] \\ &= \frac{d}{\partial\boldsymbol{\beta}}\left[\mathbf{Y}^T\mathbf{V}^{-1}\mathbf{y}-\mathbf{y}^T\mathbf{V}^{-1}\mathbf{X}\boldsymbol{\beta}-\boldsymbol{\beta}^T\mathbf{X}^T\mathbf{V}^{-1}\mathbf{y}+\boldsymbol{\beta}^T\mathbf{X}^T\mathbf{V}^{-1}\mathbf{X}\boldsymbol{\beta}\right. \\ &\quad \left.+d^2\hat{\boldsymbol{\beta}}^T\hat{\boldsymbol{\beta}}-d\hat{\boldsymbol{\beta}}^T\boldsymbol{\beta}+\boldsymbol{\beta}^T\boldsymbol{\beta}\right] \\ &= -2\mathbf{X}^T\mathbf{V}^{-1}\mathbf{y}+2\mathbf{X}^T\mathbf{V}^{-1}\mathbf{X}\boldsymbol{\beta}-2d\hat{\boldsymbol{\beta}}^T+2\boldsymbol{\beta} \end{aligned}$$

بنابراین مشابه محاسبات رابطه (۱.۳) نتیجه می‌شود

$$\widehat{\boldsymbol{\beta}}_{MLiu}=(\mathbf{X}^T\mathbf{V}^{-1}\mathbf{X}+\mathbf{I})^{-1}(\mathbf{X}^T\mathbf{V}^{-1}\mathbf{X}+d\mathbf{I})\hat{\boldsymbol{\beta}} \qquad (۶.۳)$$

همچنین امید ریاضی برآوردگر $MLiu$ عبارت است از

$$\begin{aligned} E\left(\widehat{\boldsymbol{\beta}}_{MLiu}\right) &= E\left((\mathbf{X}^T\mathbf{V}^{-1}\mathbf{X}+\mathbf{I})^{-1}(\mathbf{X}^T\mathbf{V}^{-1}\mathbf{X}+d\mathbf{I})\hat{\boldsymbol{\beta}}\right) \\ &= (\mathbf{X}^T\mathbf{V}^{-1}\mathbf{X}+\mathbf{I})^{-1}(\mathbf{X}^T\mathbf{V}^{-1}\mathbf{X}+d\mathbf{I})\cdot E(\hat{\boldsymbol{\beta}}) \\ &= (\mathbf{X}^T\mathbf{V}^{-1}\mathbf{X}+\mathbf{I})^{-1}(\mathbf{X}^T\mathbf{V}^{-1}\mathbf{X}+d\mathbf{I})\boldsymbol{\beta} \end{aligned}$$

به همین ترتیب

$$\begin{aligned} Var(\widehat{\boldsymbol{\beta}}_{MLiu}) &= (\mathbf{X}^T\mathbf{V}^{-1}\mathbf{X}+\mathbf{I})^{-1}(\mathbf{X}^T\mathbf{V}^{-1}\mathbf{X}+d\mathbf{I})(\mathbf{X}^T\mathbf{V}^{-1}\mathbf{X})^{-1} \\ &\quad \times(\mathbf{X}^T\mathbf{V}^{-1}\mathbf{X}+d\mathbf{I})(\mathbf{X}^T\mathbf{V}^{-1}\mathbf{X}+\mathbf{I})^{-1}. \end{aligned}$$

واضح است که محاسبه $\widehat{\boldsymbol{\beta}}_{MLiu}$ وابسته به ماتریس کواریانس متغیر پاسخ، یعنی

$$\mathbf{V}=\mathbf{Z}\mathbf{D}\mathbf{Z}^T+\sigma^2\mathbf{I}$$

است که مستلزم برآورد پارامتر $\boldsymbol{\theta}=(\sigma,\mathbf{D})$ می‌باشد. لذا الگوریتم ۵ برای محاسبه $\widehat{\boldsymbol{\beta}}_{MLiu}$ پیشنهاد می‌شود.

همچنین برای برآورد مقدار بهینه $0<d<1$ در رابطه (۶.۳)، به روش اعتبارسنجی متقابل داریم

$$\hat{d}=\underset{d\in(0,1)}{\arg\min}\left\{n^{-1}(\mathbf{y}-\hat{\mathbf{y}})^T(\mathbf{y}-\hat{\mathbf{y}})+2n^{-2}tr(\mathbf{S})(\mathbf{y}-\hat{\mathbf{y}})^T(\mathbf{y}-\hat{\mathbf{y}})\right\} \qquad (۷.۳)$$

که در آن $\hat{\mathbf{y}}=\mathbf{X}\widehat{\boldsymbol{\beta}}_{MLiu}$ و

$$\begin{aligned} \mathbf{S} &= \mathbf{X}(\mathbf{X}^T\mathbf{V}^{-1}\mathbf{X}+\mathbf{I})^{-1}(\mathbf{X}^T\mathbf{V}^{-1}\mathbf{X}+d\mathbf{I})(\mathbf{X}^T\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}^T\mathbf{V}^{-1} \\ &\quad +\mathbf{Z}\mathbf{D}\mathbf{Z}^T\left[\mathbf{I}-\mathbf{X}(\mathbf{X}^T\mathbf{V}^{-1}\mathbf{X}+\mathbf{I})^{-1}(\mathbf{X}^T\mathbf{V}^{-1}\mathbf{X}+d\mathbf{I})(\mathbf{X}^T\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}^T\mathbf{V}^{-1}\right]. \end{aligned}$$

الگوریتم ۵ الگوریتم EM در برآورد پارامتر مدل آمیخته خطی لیو

گام ۱) قرار می‌دهیم $\hat{\boldsymbol{\theta}}^{(t)} = \boldsymbol{\theta}_0$ و $\hat{d}^{(t)} = d_0$ سپس مقدار آماره‌های $\hat{t}_1^{(t)}$ و $\hat{t}_2^{(t)}$ را برای مرحله tام به‌صورت زیر محاسبه می‌کنیم:

$$\begin{aligned}\hat{t}_1^{(t)} &= E\left(\sum_{i=1}^{n}\boldsymbol{\epsilon}_i^T\boldsymbol{\epsilon}_i \mid \mathbf{y}_i, \hat{\boldsymbol{\beta}}^{(t)}, \hat{\boldsymbol{\theta}}^{(t)}\right)\\ &= \sum_{i=1}^{n}\left[\boldsymbol{\epsilon}_i^T(\hat{\boldsymbol{\theta}}^{(t)})\boldsymbol{\epsilon}_i(\hat{\boldsymbol{\theta}}^{(t)}) + trVar\left(\boldsymbol{\epsilon}_i \mid \mathbf{y}_i, \hat{\boldsymbol{\beta}}^{(t)}, \hat{\boldsymbol{\theta}}^{(t)}\right)\right], \qquad (۸.۳)\end{aligned}$$

$$\begin{aligned}\hat{t}_2^{(t)} &= E\left(\sum_{i=1}^{n}\mathbf{b}_i\mathbf{b}_i^T \mid \mathbf{y}_i, \hat{\boldsymbol{\beta}}^{(t)}, \hat{\boldsymbol{\theta}}^{(t)}\right)\\ &= \sum_{i=1}^{n}\left[\mathbf{b}_i(\hat{\boldsymbol{\theta}}^{(t)})\mathbf{b}_i^T(\hat{\boldsymbol{\theta}}^{(t)}) + Var\left(\mathbf{b}_i \mid \mathbf{y}_i, \hat{\boldsymbol{\beta}}^{(t)}, \hat{\boldsymbol{\theta}}^{(t)}\right)\right]. \qquad (۹.۳)\end{aligned}$$

گام ۲) برای به دست آوردن $\hat{\boldsymbol{\theta}}^{(t+1)}$، مقدار $\hat{\sigma}^{2(t+1)}$ و $\hat{\mathbf{D}}^{(t+1)}$ را از رابطه زیر محاسبه می‌کنیم:

$$\hat{\sigma}^{2(t+1)} = \hat{t}_1^{(t)}/N,$$

$$\hat{\mathbf{D}}^{(t+1)} = \hat{t}_2^{(t)}/n.$$

که در آن $N = \sum_{i=1}^{n} n_i$ و n برابر تعداد واحدهای آزمایشی است. حال برای محاسبه $\hat{\mathbf{V}}^{(t+1)}$ داریم

$$\hat{\mathbf{V}}^{(t+1)} = \mathbf{Z}\hat{\mathbf{D}}^{(t+1)}\mathbf{Z}^T + \hat{\sigma}^{2(t+1)}\mathbf{I}. \qquad (۱۰.۳)$$

گام ۳) ماتریس $\hat{\mathbf{S}}^{(t+1)}$ را از رابطه زیر محاسبه می‌کنیم

$$\begin{aligned}\hat{\mathbf{S}}^{(t+1)} &= \mathbf{X}(\mathbf{X}^T\hat{\mathbf{V}}^{-1(t+1)}\mathbf{X}+\mathbf{I})^{-1}(\mathbf{X}^T\hat{\mathbf{V}}^{-1(t+1)}\mathbf{X}+d\mathbf{I})(\mathbf{X}^T\hat{\mathbf{V}}^{-1(t+1)}\mathbf{X})^{-1}\mathbf{X}^T\hat{\mathbf{V}}^{-1(t+1)}\\ &\quad +\mathbf{Z}\mathbf{D}\mathbf{Z}^T\left[\mathbf{I}-\mathbf{X}(\mathbf{X}^T\hat{\mathbf{V}}^{-1(t+1)}\mathbf{X}+\mathbf{I})^{-1}(\mathbf{X}^T\hat{\mathbf{V}}^{-1(t+1)}\mathbf{X}+\hat{d}^{(t)}\mathbf{I})(\mathbf{X}^T\hat{\mathbf{V}}^{-1(t+1)}\mathbf{X})^{-1}\mathbf{X}^T\hat{\mathbf{V}}^{-1(t+1)}\right],\end{aligned}$$

سپس برای محاسبه $\hat{d}^{(t+1)}$ داریم

$$\hat{d}^{(t+1)} = \arg\min_{d\in(0,1)}\left\{n^{-1}(\mathbf{y}-\widehat{\mathbf{y}})^T(\mathbf{y}-\widehat{\mathbf{y}})/(1-\mathrm{tr}(\hat{\mathbf{S}}^{(t+1)})/n)^2\right\} \qquad (۱۱.۳)$$

گام ۴) در نهایت مقدار $\hat{\boldsymbol{\beta}}_{MLiu}^{(t+1)}$ از رابطه زیر بدست می‌آید

$$\widehat{\boldsymbol{\beta}}_{MLiu}^{(t+1)} = (\mathbf{X}^T\hat{\mathbf{V}}^{-1(t+1)}\mathbf{X}+\mathbf{I})^{-1}(\mathbf{X}^T\hat{\mathbf{V}}^{-1(t+1)}\mathbf{X}+\hat{d}^{(t+1)}\mathbf{I})\widehat{\boldsymbol{\beta}}^{(t+1)} \qquad (۱۲.۳)$$

گام ۵) گام‌های (۱) تا (۴) را تا رسیدن به همگرایی $\left(|\hat{\boldsymbol{\beta}}_{MLiu}^{(t+1)} - \hat{\boldsymbol{\beta}}_{MLiu}^{(t)}| < \gamma\right)$ به ازای یک مقدار γ کوچک و از پیش تعیین‌شده، تکرار می‌کنیم.

در ادامه چگونگی محاسبه $\widehat{\boldsymbol{\beta}}_{MLiu}$ را در قالب یک مطالعه شبیه‌سازی و مثال واقعی با استفاده از نرم‌افزار R بررسی می‌کنیم.

۱.۲.۳ مطالعه شبیه‌سازی

در این بخش به بررسی عملکرد برآوردگر آمیخته لیو و مقایسه رفتار آن با برآوردگر آمیخته خطی در مدل آمیخته خطی خواهیم پرداخت.

در این مطالعه پس از تولید داده‌های همخط مشابه بخش ۱.۲.۲، برای شبیه‌سازی مقادیر متغیر پاسخ، در رابطه‌ی (۳.۳)، $\boldsymbol{\beta} = (۰, ۰/۴, ۱/۰, ۱/۶, ۲/۰)$، $b_{ijk} \sim \mathcal{N}(۰, ۰/۶)$ و $\varepsilon_{ijk} \sim \mathcal{N}(۰, ۱)$ فرض شده‌اند به‌طوری که برای هر واحد آزمایشی iام $n_i = ۴$ و همچنین هر متغیر پیش‌بین از توزیع نرمال با میانگین ۵ و واریانس ۱ در نظر گرفته شده است.

برای ایجاد همبستگی بین متغیرهای پیش‌بین در چند حالت مختلف مقدار ρ طبق جدول ۱.۳ از صفر تا ۰/۹ فرض شده است. همچنین برای محاسبه‌ی خطای استاندارد برآوردگر روند شبیه‌سازی مونت کارلویی به تعداد $B = ۱۰۰$ تکرار شده است. لازم به ذکر است که در این مطالعه تعداد نمونه‌ی آزمایشی $n = ۴۰$ فرض شده است. نتایج شبیه‌سازی در جدول ۱.۳ آمده است.

جدول ۱.۳: نتایج مطالعه شبیه‌سازی مدل آمیخته خطی لیو

		مدل آمیخته خطی		مدل آمیخته خطی لیو	
ρ	β	برآورد	انحراف استاندارد	برآورد	انحراف استاندارد
۰/۰	۰/۰	۰/۰۱۲۱۵۰۹۴	۰/۰۵۵۲۷۳۶	۰/۰۱۳۷۲۷۴۹	۰/۰۵۵۱۰۲۱
	۰/۴	۰/۴۱۱۱۵۹۱۱	۰/۰۵۶۲۰۲۴	۰/۴۱۱۹۶۸۳۹	۰/۰۵۶۰۰۴۴
	۱/۰	۰/۹۷۱۹۰۰۵۱	۰/۰۲۵۱۴۴۴	۰/۹۷۱۸۸۶۳۲	۰/۰۲۴۹۸۹۱
	۱/۶	۱/۶۰۸۸۹۰۷۴	۰/۰۳۹۸۷۶۹	۱/۶۰۷۹۷۹۲۴	۰/۰۳۹۵۹۳۰
	۲/۰	۱/۹۹۶۲۵۱۵۴	۰/۰۴۷۸۸۸۵	۱/۹۹۴۷۲۸۹۶	۰/۰۴۷۶۶۷۹
MSE		۱/۷۱۷۸۹		۱/۷۰۳۱۳۷	
۰/۳	۰/۰	۰/۰۰۹۴۸۹۱۳۹	۰/۰۸۳۴۹۶۷	۰/۰۱۱۶۱۲۶۶	۰/۰۸۳۱۴۷۵
	۰/۴	۰/۴۱۹۳۲۹۳۵۱	۰/۰۸۹۳۷۷۵	۰/۴۲۰۵۱۴۴۰	۰/۰۸۹۰۴۴۰
	۱/۰	۰/۹۶۵۸۵۷۹۳۴	۰/۰۴۷۸۸۸۵	۰/۹۶۵۸۶۰۰۸	۰/۰۴۷۶۴۲۴
	۱/۶	۱/۶۱۲۳۱۷۵۹۸	۰/۰۶۵۵۲۴۲	۱/۶۱۱۰۳۹۱۷	۰/۰۶۵۰۵۶۳
	۲/۰	۱/۹۹۳۷۵۹۶۴۰	۰/۰۴۴۳۲۴۹	۱/۹۹۱۶۶۲۵۴	۰/۰۴۴۰۵۴۷
MSE		۳/۷۶۳۶۵۴		۳/۷۲۷۵۵۷	
۰/۶	۰/۰	۰/۰۱۲۶۰۴۰۵	۰/۰۵۵۴۶۹۳	۰/۰۱۶۱۱۷۶۲	۰/۰۵۴۶۲۱۲
	۰/۴	۰/۴۳۵۸۹۲۳۰	۰/۰۲۹۱۱۵۶	۰/۴۳۷۴۳۶۳۳	۰/۰۲۸۵۱۴۴
	۱/۰	۰/۹۴۵۴۳۱۴۰	۰/۰۴۴۰۱۰۴	۰/۹۴۵۳۳۵۹۰	۰/۰۴۳۵۱۱۲
	۱/۶	۱/۶۱۴۲۲۲۲۶	۰/۰۱۹۲۲۸۰	۱/۶۱۲۵۲۰۹۵	۰/۰۱۸۹۲۵۹
	۲/۰	۱/۹۹۳۴۵۷۴۰	۰/۰۸۵۹۰۹۹	۱/۹۹۰۱۳۶۳۸	۰/۰۸۵۳۱۰۵
MSE		۲/۱۸۲۵۴۳		۲/۱۳۷۰۳۸	
۰/۹	۰/۰	۰/۰۰۲۲۷۶۵۷۶	۰/۰۷۸۶۰۴۵	۰/۰۱۱۰۷۲۱۷	۰/۰۷۸۰۶۵۴
	۰/۴	۰/۴۳۰۸۰۴۲۲۰	۰/۰۵۱۳۴۰۶	۰/۴۳۵۲۶۷۶۷	۰/۰۵۱۰۶۱۳
	۱/۰	۰/۹۱۰۵۳۹۸۹۲	۰/۰۸۱۲۳۸۷	۰/۹۱۱۳۴۲۷۸	۰/۰۸۰۸۷۲۹
	۱/۶	۱/۶۲۲۶۱۵۱۱۳	۰/۰۵۷۸۴۹۰	۱/۶۱۷۴۲۶۲۵	۰/۰۵۷۲۸۰۶
	۲/۰	۲/۰۳۶۸۷۲۰۸۹	۰/۰۸۵۵۰۰۱	۲/۰۲۷۹۰۵۱۷	۰/۰۸۵۰۰۳۱
MSE		۴/۱۸۲۱۹۲		۴/۱۳۰۰۶۶	

همان‌طور که در جدول ۱.۳ ملاحظه می‌شود، برای هر بخش از داده‌های تولید شده دارای میزان همبستگی متفاوت، برآوردهای حاصل از مدل آمیخته خطی لیو به مراتب دارای انحراف استاندارد کمتری نسبت به برآوردهای حاصل از مدل آمیخته خطی است.

۲.۲.۳ مثال واقعی

سلول‌های دستگاه ایمنی بدن در یک فرد، گاهی به دلایل گوناگون ممکن است نتواند برخی از مواد و اجزای بدن فرد را از مواد خارجی تشخیص دهد، از این رو نسبت به بافت‌های خود حساس شده و بر ضد آن‌ها آنتی کور[۱] (پادزهر یا موادی که بر علیه سم واردشده در بدن به وجود می‌آیند) ترشح می‌کند. نتیجه حاصل را بیماری خود ایمنی[۲] می‌نامند که دستگاه ایمنی بدن به اشتباه، حمله به خود را آغاز می‌کند.

بیماری سیروز صفراوی اولیه[۳] (PBC) نوعی اختلال خود ایمنی است که کبد[۴] را تحت تأثیر قرار می‌دهد. به عبارت ساده‌تر در این نوع بیماری بدن احساس می‌کند که درون مجرای صفراوی کبد، جسم خارجی وجود دارد و سعی در از بین بردن پوشش به‌وجود آمده در این مجرا را دارد. این بیماری به کندی در بدن پیشرفت می‌کند که بسیاری از بیماران حتی ۱۰ الی ۱۵ سال پس از تشخیص نیز عارضه جدی و مهمی ندارند تا این‌که نهایتاً باعث مرگ سلول‌های کبدی و به حالت پلاستیکی درآمدن کبد می‌گردد که بیمار را نیازمند به انجام عمل پیوند کبد می‌کند یا منجر به مرگ بیمار می‌شود.

بیماری PBC یک بیماری نسبتاً نادر است که حدوداً از بین هر ۴۰۰۰ نفر یک نفر به آن مبتلا است. هچنین این نوع بیماری در زنان بسیار شایع‌تر است چرا که نسبت زنان مبتلا به مردان مبتلا، ۹ به ۱ است.

در این مطالعه تعداد ۳۱۲ بیمار به طور تصادفی از بین افراد مبتلا به خود ایمنی کبد یا سیروز صفراوی اولیه در یکی از معتبرترین مراکز و دانشگاه‌های علوم پزشکی جهان واقع در شهر روچستر[۵] ایالت مینه سوتای[۶] کشور آمریکا به نام کلینیک مایو[۷] مورد بررسی قرار گرفته‌اند. این داده‌ها در بسته JMbayes نرم‌افزار R موجود هستند و شامل ۱۹۴۵ مشاهده با ۲۰ متغیر می‌باشند که اندازه‌گیری هر کدام از این متغیرها برای بیمار iام، n_i بار تکرار شده‌است. در جدول زیر به‌طور خلاصه به توضیح این متغیرها پرداخته شده و متغیرهایی که برای تحلیل این مثال، در مدل به عنوان اثرات تصادفی در نظر گرفته شده‌اند با علامت ■ مشخص شده‌اند. همچنین برآورد ضرایب حاصل از تحلیل این مجموعه داده در جدول ۳.۳ آورده شده‌است.

[1] Antibody
[2] Autoimmune
[3] Primary biliary cirrhosis
[4] Liver
[5] Rochester
[6] Minnesota
[7] Mayo Clinic

جدول ۲.۳: توصیف متغیرهای مجموعه داده بخش ۱.۲.۳

متغیرها	توصیف	اثر تصادفی
id	کد شناسایی مربوط به فرد مورد آزمایش	
years	بازه زمانی ثبت مشخصات تا زمان مرگ یا انجام عمل پیوند	
status	وضعیت بیمار (زنده، پیوندی، مرگ)	■
drug	نوع داروی مصرفی (D-penicil ـ Placebo)	■
age	سن	
sex	جنسیت	■
year	بازه زمانی ثبت مشخصات تا زمان شروع درمان	
ascites	دارای حفره شکم ـ از عوارض بیماری (بله ـ خیر)	■
hepatomegaly	بزرگ‌شدگی کبد ـ از عوارض بیماری (بله ـ خیر)	■
spiders	عوارض پوستی (بله ـ خیر)	■
edema	افزایش میزان آب دفع‌شده از بدن (بله ـ خیر)	■
serBilir	میزان بیلی‌روبین ـ نوعی سرم (میلی‌گرم بر دسی‌لیتر)	
serChol	میزان کلسترول (میلی‌گرم بر دسی‌لیتر)	
albumin	میزان آلبومین ـ نوعی پروتئین (گرم بر دسی‌لیتر)	
alkaline	آلکالین فسفاتاز ـ آنزیمی که توسط کبد ساخته می‌شود (یونیت بر لیتر)	
SGOT	اسگات ـ نوعی آنزیم در سلول‌های کبد (یونیت بر میلی‌لیتر)	
platelets	پلاسمای خون (مکعب میلی‌لیتر بر ۱۰۰۰)	
prothrombin	آزمایش پروتوموبین ـ مدت زمان مکانیسم انعقاد خون بر حسب ثانیه	
histologic	مرحله بافت‌شناسی	
*status*۲	وضعیت بیمار (زنده یا پیوندی، مرگ)	■

جدول ۳.۳: برآورد ضرایب در مجموعه داده بخش ۱.۲.۳

	مدل آمیخته خطی	مدل آمیخته خطی لیو
age	−۰/۰۰۵۶	−۰/۰۰۵۳
year	۰/۷۰۹۵	۰/۷۰۸۵
serBilir	−۰/۱۷۱۱	−۰/۱۷۲۴
serChol	−۰/۰۰۰۴	−۰/۰۰۰۴
albumin	۲/۳۵۰۴	۲/۳۲۳۶
alkaline	۰/۰۰۰۱	۰/۰۰۰۱
SGOT	۰/۰۰۲۳	۰/۰۰۲۳
platelets	۰/۰۰۰۹	۰/۰۰۱۰
prothrombin	−۰/۰۰۸۰	−۰/۰۰۳۷
histologic	−۰/۵۰۸۶	−۰/۵۰۴۶
MPE	۰/۰۱۵۵۹	۰/۰۱۵۳۵

با توجه به آن‌چه در بخش ۲.۴.۲ در رابطه با معیار MPE گفته شد، در این مثال نیز برتری عملکرد مدل آمیخته خطی لیو نسبت به مدل آمیخته خطی از لحاظ معیار MPE تایید می‌شود.

۳.۳ مدل آمیخته خطی لیو با خطای بیضی‌گون

در این بخش با توجه به آن‌چه در بخش ۳.۲ در مورد توزیع‌های بیضی‌گون بیان شد، مدل آمیخته خطی لیو را با خطای بیضی‌گون تحت مفروضات مدل (۲۰.۲) مورد بررسی قرار می‌دهیم.

لگاریتم تابع درستنمایی متغیر $\mathbf{y}$ را بر پایه‌ی مدل آمیخته خطی با اعمال تابع جریمه در رگرسیون لیو، به صورت زیر داریم

$$\ell^{pen}(\mathbf{y}) = -\frac{N}{2}\log(2\pi) - \frac{1}{2}\sum_{i=1}^{n}\log|\mathbf{V}| + \log g\left(-\frac{1}{2}(\mathbf{y}-\mathbf{X}\boldsymbol{\beta})^T\mathbf{V}^{-1}(\mathbf{y}-\mathbf{X}\boldsymbol{\beta})\right) - (d\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})^T(d\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})$$

لذا برآوردگر $MLiu$ پارامتر $\boldsymbol{\beta}$ با ماکزیمم کردن تابع فوق به صورت زیر خواهد بود

$$\widehat{\boldsymbol{\beta}}_{MLiu} = (\mathbf{X}^T\mathbf{V}^{-1}\mathbf{X}+\mathbf{I})^{-1}(\mathbf{X}^T\mathbf{V}^{-1}\mathbf{X}+d\mathbf{I})\hat{\boldsymbol{\beta}}$$

که در آن $0 < d < 1$ و از رابطه (۷.۳) برآورد می‌شود.

در ادامه عملکرد این برآوردگر را تحت یک مثال شبیه سازی شده مورد بررسی قرار می‌دهیم.

۱.۳.۳ مطالعه شبیه‌سازی

در این بخش برای شبیه‌سازی مقادیر متغیر پاسخ، در رابطه‌ی (۲۰.۲) بردار ضرایب رگرسیونی به صورت $\boldsymbol{\beta} = (۰, ۰/۴, ۱, ۱/۶, ۲)$، $b_{ijk} \sim \mathcal{T}(۰, ۰/۳, ۵)$ و $\varepsilon_{ijk} \sim \mathcal{T}(۰, ۱, ۵)$ فرض شده‌اند به‌طوری که برای هر واحد آزمایشی iام $n_i = ۴$ و هر متغیر پیش‌بین از توزیع نرمال با میانگین ۵ و واریانس ۱ فرض شده‌اند. همچنین برای ایجاد داده‌های هم‌خط از بسته‌ی EnvStat در نرم‌افزار R با ضریب همبستگی‌های متفاوت موجود در جدول ۴.۳ $\rho \in \{۰/۰, ۰/۲, ۰/۵, ۰/۷, ۰/۹\}$ استفاده شده است.

همچنین برای محاسبه‌ی خطای استاندارد برآورد بدست آمده با روند شبیه‌سازی مونت‌کارلویی به تعداد $B = ۱۰۰$ تکرار شده و تعداد نمونه آزمایشی $n = ۴۰$ در نظر گرفته شده است. در این راستا برای محاسبه برآوردگرها از الگوریتم ۵ استفاده کرده‌ایم. نتایج حاصل در جدول ۴.۳ آمده است.

جدول ۴.۳: نتایج مطالعه شبیه‌سازی مدل آمیخته خطی لیو با خطای بیضی‌گون

		مدل آمیخته خطی با خطای بیضی‌گون		مدل آمیخته خطی لیو با خطای بیضی‌گون	
ρ	β	برآورد	انحراف استاندارد	برآورد	انحراف استاندارد
۰/۰	۰/۰	۰/۰۱۳۲۵۷۸	۰/۱۱۲۹۳۱۰	۰/۰۷۸۰۴۲۰	۰/۱۱۲۷۱۸۰
	۰/۴	۰/۴۰۳۴۴۶۴	۰/۱۶۷۸۲۱۵	۰/۴۳۵۱۲۳	۰/۱۶۷۳۶۹۰
	۱/۰	۱/۰۲۷۷۳۴۸	۰/۱۸۶۱۹۰۶	۰/۹۱۷۰۵۰۵	۰/۱۸۵۹۶۱۰
	۱/۶	۱/۶۱۶۹۶۹۲	۰/۱۶۸۶۳۷۴	۱/۵۶۹۶۳۷۷	۰/۱۶۸۲۰۹۰
	۲/۰	۲/۰۲۰۱۴۶۹	۰/۱۷۵۵۲۷۹	۱/۹۹۳۶۰۳۴	۰/۱۷۵۲۷۰۳
MSE		۲۱/۵۷۴۹۳		۲۱/۵۰۵۲۷	
۰/۳	۰/۰	۰/۰۱۵۵۱۲۶	۰/۱۱۵۳۱۶۰	۰/۰۱۵۸۴۴۵	۰/۱۱۵۰۰۲۴
	۰/۴	۰/۴۴۰۱۳۰۶	۰/۱۳۸۵۱۷۰	۰/۴۴۹۷۰۵۸	۰/۱۳۸۱۵۵۳
	۱/۰	۰/۹۸۳۳۸۶۰	۰/۱۹۳۵۸۰۶	۱/۰۲۱۹۴۷۳	۰/۱۹۳۱۰۰۱
	۱/۶	۱/۶۰۱۴۶۴۲	۰/۱۹۲۷۵۵۴	۱/۵۷۹۵۰۳۲	۰/۱۹۲۲۹۶۹
	۲/۰	۲/۰۲۳۵۵۹۳	۰/۱۷۷۹۵۵۵	۲/۰۳۶۶۷۲۱	۰/۱۷۷۷۳۳۸
MSE		۲۲/۲۰۷۶۴		۲۲/۱۱۱۷۷	
۰/۶	۰/۰	۰/۰۱۲۶۳۳۷	۰/۱۱۹۴۳۸۹	۰/۰۵۲۶۲۷۲	۰/۱۱۹۰۳۳۰
	۰/۴	۰/۳۵۰۰۵۰۰	۰/۱۳۸۲۲۱۱	۰/۴۸۱۹۱۳۴	۰/۱۳۷۷۴۷۶
	۱/۰	۱/۰۴۲۷۵۹۸	۰/۱۱۷۹۷۳۰	۰/۹۱۵۴۹۱۸	۰/۱۱۷۳۵۳۳
	۱/۶	۱/۵۷۶۵۳۹۶	۰/۱۴۹۳۹۵۳	۱/۶۸۱۴۶۱۳	۰/۱۴۸۷۱۸۲
	۲/۰	۲/۰۵۳۰۹۹۲	۰/۱۸۰۵۵۵۱	۱/۹۸۷۱۶۳۸	۰/۱۷۹۶۵۴۵
MSE		۱۶/۳۶۱۰۵		۱۶/۲۳۲۶۹	
۰/۹	۰/۰	−۰/۰۴۳۸۰۳۲	۰/۱۴۴۹۸۲۳	۰/۰۸۰۶۶۰۷	۰/۱۳۵۴۷۰۱
	۰/۴	۰/۴۱۶۳۰۵۷۸	۰/۱۵۲۷۷۵۱	۰/۴۹۱۷۰۱۳	۰/۱۴۰۱۷۱۸
	۱/۰	۱/۰۱۱۰۰۵۹	۰/۱۹۳۷۲۱۷	۱/۱۶۵۵۲۱۷	۰/۱۸۵۰۲۴۲
	۱/۶	۱/۵۵۶۶۵۷۱	۰/۱۶۲۲۸۸۹	۱/۷۲۸۸۹۷۲	۰/۱۵۳۸۷۰۹
	۲/۰	۲/۰۲۳۱۸۶۰۶	۰/۱۹۴۳۸۶۸	۲/۰۹۱۲۵۸۶	۰/۱۸۳۸۴۷۵
MSE		۲۳/۳۴۴۲۴		۲۰/۸۲۰۹۲	

همان‌طور که در جدول ۴.۳ ملاحظه می‌شود برآوردهای حاصل از روش آمیخته لیو با خطاهای بیضی‌گون دارای انحراف استاندارد کمتری نسبت به برآورد حاصل از روش آمیخته خطی است. مقادیر میانگین توان دوم خطای برآورد (MSE) نیز برتری روش آمیخته لیو را نسبت به روش آمیخته خطی تایید می‌کند.

۴.۳ برآوردگرهای انقباضی در مدل آمیخته خطی لیو

در این بخش برآوردگرهای آزمون اولیه و انقباضی نوع استاین را در مدل آمیخته خطی لیو مورد بررسی قرار می‌دهیم.

فرض کنید در مدل آمیخته خطی لیو، فرضیه زیر مطرح باشد

$$\begin{cases} H_0 : \boldsymbol{\beta} = \boldsymbol{\beta}_0 \\ H_1 : \boldsymbol{\beta} \neq \boldsymbol{\beta}_0 \end{cases}$$

در این‌صورت طبق آن‌چه در بخش (۴.۲) بیان شد، برآوردگر آزمون اولیه آمیخته خطی لیو[8] $(PTMR)$ به صورت زیر تعریف می‌شود.

$$\widehat{\boldsymbol{\beta}}^{PT}_{MLiu} = \begin{cases} \widehat{\boldsymbol{\beta}}_{MLiu} & \text{فرضیه‌ی } H_0 \text{ رد شود} \\ \boldsymbol{\beta}_0 & \text{فرضیه‌ی } H_0 \text{ پذیرفته شود} \end{cases}$$

برآوردگر $PTMLiu$ را می‌توان با استفاده از تابع نشانگر به صورت زیر بازنویسی کرد

$$\widehat{\boldsymbol{\beta}}^{PT}_{MLiu} = \widehat{\boldsymbol{\beta}}_{MLiu} - \left(\widehat{\boldsymbol{\beta}}_{MLiu} - \boldsymbol{\beta}_0\right) \mathbf{I}_{(\mathbf{L}_n < \mathbf{L}_n(\alpha))} \qquad (۱۳.۳)$$

که در آن $\mathbf{L}_n$ آماره والد آزمون فرضیه‌ی H_0 به صورت

$$\mathbf{L}_n = \left(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0\right)^T \left(\mathbf{X}^T \mathbf{V}^{-1} \mathbf{X}\right) \left(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0\right) \qquad (۱۴.۳)$$

و $\mathbf{L}_n(\alpha)$ صدک بالایی 100α درصد توزیع کی‌ـدو با p درجه آزادی است.

به همین ترتیب برآوردگر انقباضی نوع استاین آمیخته لیو[9] $(SSMLiu)$ به صورت زیر تعریف می‌شود

$$\begin{aligned} \widehat{\boldsymbol{\beta}}^{S}_{MLiu} &= \boldsymbol{\beta}_0 + \left(1 - c\mathbf{L}_n^{-1}\right) \left(\widehat{\boldsymbol{\beta}}_{MLiu} - \boldsymbol{\beta}_0\right) \\ &= \widehat{\boldsymbol{\beta}}_{MLiu} - c \left(\widehat{\boldsymbol{\beta}}_{MLiu} - \boldsymbol{\beta}_0\right) \mathbf{L}_n^{-1} \end{aligned}$$

که در آن $c = \dfrac{(p-2)(N-p)}{p(N-p+2)}$.

برای محاسبه این برآوردگرها در مدل آمیخته خطی لیو، گام ۶ام را در الگوریتم ۵ به صورت زیر داریم

[8] Preliminary test mixed Liu

[9] Stein-type shrinkage mixed Liu

الگوریتم ۶ الگوریتم EM در برآورد آزمون اولیه و نوع استاین پارامتر مدل آمیخته خطی لیو

گام‌های (۱) ـ (۵)
به الگوریتم ۵ مراجعه شود.

گام ۶) پس از محاسبه آماره L_n از رابطه (۱۴.۳)، داریم

$$\hat{\boldsymbol{\beta}}^{PT}_{MLiu} = \hat{\boldsymbol{\beta}}_{MLiu} - (\hat{\boldsymbol{\beta}}_{MLiu} - \hat{\boldsymbol{\beta}}_{\circ}) \quad I(L_n < L_n(\alpha)) \tag{۱۵.۳}$$

$$\hat{\boldsymbol{\beta}}^{S}_{MLiu} = \hat{\boldsymbol{\beta}}_{MLiu} - c(\hat{\boldsymbol{\beta}}_{MLiu} - \hat{\boldsymbol{\beta}}_{\circ})L_n^{-1} \tag{۱۶.۳}$$

۱.۴.۳ مطالعه شبیه‌سازی

در این مطالعه برای هر $n = ۴۰$ واحد آزمایشی $n_i = ۴$ را ثابت در نظر گرفته و همچنین

$$\boldsymbol{\beta} = (۰/۰, ۰/۱, ۰/۲, ۰/۴, ۰/۸)$$

$b_{ijk} \sim \mathcal{N}(۰, ۰/۶)$ و $\varepsilon_{ijk} \sim \mathcal{N}(۰, ۱)$ فرض شده‌اند. برای تولید داده‌های هم‌خط از بسته EnvStat موجود در نرم‌افزار R استفاده شده است که هر متغیر پیش‌بین دارای توزیع نرمال با میانگین ۵ و واریانس ۱ می‌باشد و مقدار همبستگی بین متغیرهای پیش‌بین طبق جدول ۵.۳ به صورت $\rho \in \{۰/۰, ۰/۲, ۰/۵, ۰/۷, ۰/۹\}$ در نظر گرفته شده‌است. همچنین برای محاسبه‌ی خطای استاندارد و میانگین خطای برآورد (MSE)، روند شبیه‌سازی به تعداد $B = ۱۰۰$ بار تکرار شده است. لازم به ذکر است که در این مطالعه، فرضیه $H_{\circ} : \boldsymbol{\beta} = ۰$ در نظر گرفته شده است. نتایج این مطالعه در جدول ۵.۳ آمده است.

جدول ۵.۳: نتایج مطالعه شبیه‌سازی برآوردگر نوع استاین در مدل آمیخته خطی لیو

		مدل آمیخته خطی		مدل آمیخته خطی لیو		مدل آمیخته خطی نوع استاین		مدل آمیخته خطی لیو نوع استاین	
ρ	β	انحراف استاندارد	برآورد	انحراف استاندارد	برآورد	انحراف استاندارد	برآورد	انحراف استاندارد	برآورد
۰/۰	۰/۰	۰/۰۱۴۶۵	−۰/۰۱۴۱۷	۰/۰۱۴۲۰	۰/۰۲۶۳۷	۰/۰۱۴۲۲	−۰/۰۱۵۲۲	۰/۰۱۴۰۸	۰/۰۲۴۱۲
	۰/۱	۰/۰۱۷۶۲	۰/۰۹۲۷۰	۰/۰۱۷۵۹	۰/۱۲۰۶۸	۰/۰۱۷۶۰	۰/۰۹۱۳۶	۰/۰۱۷۲۱	۰/۱۱۹۷۲
	۰/۲	۰/۰۱۹۴۴	۰/۱۹۷۲۴	۰/۰۱۹۳۲	۰/۱۹۸۳۲	۰/۰۱۹۳۹	۰/۱۹۸۱۴	۰/۰۱۹۱۸	۰/۲۰۲۱۶
	۰/۴	۰/۰۱۸۱۶	۰/۴۱۱۷۲	۰/۰۱۸۱۰	۰/۳۹۲۹۸	۰/۰۱۸۱۲	۰/۴۰۲۰۹	۰/۰۱۷۹۴	۰/۳۹۷۸۲
	۰/۸	۰/۰۲۱۶۹	۰/۸۱۲۵۴	۰/۰۲۱۰۳	۰/۸۳۱۶۲	۰/۰۲۱۴۲	۰/۸۲۰۴۶	۰/۰۲۰۸۸	۰/۸۳۲۷۱
MSE		۰/۲۷۳۰۷۵		۰/۲۶۶۸۴۴		۰/۲۶۸۷۴۷		۰/۲۶۱۲۶۹	
۰/۲	۰/۰	۰/۰۱۸۲۳	۰/۰۰۱۰۲	۰/۰۱۸۰۶	۰/۰۳۰۲۹	۰/۰۱۸۲۰	−۰/۰۰۱۱۶	۰/۰۱۷۷۴	۰/۰۲۹۳۵
	۰/۱	۰/۰۱۸۴۰	۰/۱۰۲۴۴	۰/۰۱۸۲۰	۰/۱۱۹۰۳	۰/۰۱۸۳۴	۰/۱۱۳۱۶	۰/۰۱۸۱۱	۰/۱۱۷۲۲
	۰/۲	۰/۰۱۹۸۷	۰/۲۰۸۰۶	۰/۰۱۹۷۱	۰/۲۲۴۱۲	۰/۰۱۹۸۳	۰/۲۰۱۸۴	۰/۰۱۹۶۵	۰/۲۲۳۴۷
	۰/۴	۰/۰۲۲۳۷	۰/۴۲۰۱۶	۰/۰۲۲۱۷	۰/۳۸۰۴۶	۰/۰۲۲۳۲	۰/۴۱۷۲۹	۰/۰۲۲۰۳	۰/۳۷۹۲۴
	۰/۸	۰/۰۱۸۸۲	۰/۸۱۶۰۹	۰/۰۱۸۶۶	۰/۸۳۳۷۲	۰/۰۱۸۷۹	۰/۸۲۱۰۲	۰/۰۱۸۶۵	۰/۸۳۲۸۹
MSE		۰/۳۰۷۹۸۸		۰/۳۰۵۰۷۵		۰/۳۰۶۸۵		۰/۳۰۱۱۳۳	
۰/۵	۰/۰	۰/۰۳۰۱۶	۰/۰۱۱۰۲	۰/۰۳۰۰۰	۰/۰۶۸۲۳	۰/۰۳۰۰۹	۰/۰۱۲۱۶	۰/۰۲۹۷۵	۰/۰۶۹۱۳
	۰/۱	۰/۰۲۴۵۸	۰/۱۰۵۵۸	۰/۰۲۴۳۴	۰/۱۳۲۵۹	۰/۰۲۴۵۲	۰/۰۹۸۲۲	۰/۰۲۴۱۶	۰/۱۳۳۰۵
	۰/۲	۰/۰۲۴۱۳	۰/۲۰۲۳۱	۰/۰۲۳۸۴	۰/۲۲۳۱۶	۰/۰۲۴۰۳	۰/۱۹۹۴۶	۰/۰۲۳۶۹	۰/۲۱۹۲۴
	۰/۴	۰/۰۳۰۴۵	۰/۳۷۷۸۴	۰/۰۳۰۲۴	۰/۳۷۲۰۲	۰/۰۳۰۴۱	۰/۳۷۸۱۴	۰/۰۳۰۱۹	۰/۳۷۰۱۱
	۰/۸	۰/۰۲۲۴۹	۰/۸۱۴۲۹	۰/۰۲۲۱۸	۰/۷۹۶۱۸	۰/۰۲۲۳۸	۰/۸۱۵۰۶	۰/۰۲۲۱۳	۰/۷۹۵۹۸
MSE		۰/۵۶۵۵۰۳		۰/۵۶۱۸۰۲		۰/۵۶۲۴۱		۰/۵۵۶۱۳۶	
۰/۷	۰/۰	۰/۰۳۲۴۸	۰/۰۰۶۲۸	۰/۰۳۲۰۱	۰/۰۹۳۲۴	۰/۰۳۲۳۷	۰/۰۰۴۲۸	۰/۰۳۱۸۴	۰/۰۸۲۵۲
	۰/۱	۰/۰۳۱۳۸	۰/۱۱۲۸۰	۰/۰۳۰۹۷	۰/۱۶۶۳۹	۰/۰۳۱۳۱	۰/۱۰۳۰۱	۰/۰۲۹۹۸	۰/۱۶۷۳۹
	۰/۲	۰/۰۳۳۴۹	۰/۲۰۶۴۳	۰/۰۳۳۰۵	۰/۲۰۱۱۹	۰/۰۳۳۴۰	۰/۲۰۵۴۰	۰/۰۳۲۲۰	۰/۲۰۱۰۳
	۰/۴	۰/۰۳۸۹۰	۰/۴۱۵۳۵	۰/۰۳۸۵۱	۰/۴۰۳۲۱	۰/۰۳۸۸۴	۰/۴۱۴۲۵	۰/۰۳۸۳۲	۰/۴۱۵۹۸
	۰/۸	۰/۰۳۱۰۶	۰/۸۴۲۱۸	۰/۰۲۹۸۲	۰/۷۲۳۰۹	۰/۰۳۰۵۴	۰/۸۴۱۴۷	۰/۰۲۹۶۹	۰/۷۱۴۸۵
MSE		۰/۹۰۴۵۲۶		۰/۸۹۰۷۶۲		۰/۸۹۵۵۶۸		۰/۸۶۶۷۵۳	
۰/۹	۰/۰	۰/۰۵۸۲۰	۰/۰۳۹۲۸	۰/۰۵۷۶۵	۰/۱۹۱۳۶	۰/۰۵۸۰۲	۰/۰۳۷۱۴	۰/۰۵۷۳۳	۰/۱۹۰۲۹
	۰/۱	۰/۰۴۸۶۳	۰/۰۸۶۱۲	۰/۰۴۷۷۱	۰/۱۵۶۴۲	۰/۰۴۸۵۰	۰/۰۸۲۸۴	۰/۰۴۴۷۰	۰/۱۵۷۱۳
	۰/۲	۰/۰۵۲۴۹	۰/۲۰۷۶۴	۰/۰۵۱۸۱	۰/۲۳۴۰۶	۰/۰۵۲۳۱	۰/۲۰۳۸۴	۰/۰۵۱۴۳	۰/۲۳۵۸۳
	۰/۴	۰/۰۵۳۸۴	۰/۳۶۸۲۷	۰/۰۵۳۱۱	۰/۳۳۷۵۶	۰/۰۵۳۷۲	۰/۳۷۴۰۲	۰/۰۵۲۶۲	۰/۳۳۷۲۹
	۰/۸	۰/۰۴۹۶۵	۰/۸۸۷۶۱	۰/۰۴۸۷۶	۰/۶۹۳۱۶	۰/۰۴۹۵۱	۰/۸۷۲۶۲	۰/۰۴۸۵۶	۰/۶۹۳۰۹
MSE		۲/۲۲۹۸۶۶		۲/۲۱۳۴۳۶		۲/۲۱۴۳۵۵		۲/۱۴۵۲۰۸	

آنچه در تحلیل نتایج جدول ۵.۳ حائز اهمیت است این است که برآوردگر $SSMLiu$ در بین تمامی برآوردگرهای مورد مطالعه دارای میانگین توان دوم خطا (MSE) و انحراف استاندارد کمتری است.

۵.۳ نتیجه‌گیری و آینده تحقیق

۱.۵.۳ خلاصه و نتیجه‌گیری

مهمترین ویژگی داده‌های طولی، اندازه‌گیری‌های مکرر از واحدهای تحت آزمایش در طول زمان است لذا هدف اول در مطالعات طولی، بررسی تغییرات پاسخ در طول زمان و همچنین عوامل مؤثر بر این تغییرات است. از آنجا که اندازه‌گیری بر روی واحدهای تحت مطالعه تکرار می‌شود، امکان بررسی تغییرات درون هر واحد نیز فراهم می‌شود. در جمع‌آوری داده‌های طولی، اندازه‌گیری مکرر متغیری خاص برای یک واحد آزمایشی در موقعیت‌ها و زمان‌های مختلف، همبستگی بین مشاهدات مربوط به آن واحد را نشان می‌دهد. همان‌طور اشاره شد

یکی از روش‌های مدل‌بندی این داده‌ها استفاده از مدل‌های آمیخته خطی است. این مدل‌ها همبستگی بین مشاهدات را با معرفی یک یا چند متغیر تصادفی به عنوان اثر تصادفی در مدل در نظر می‌گیرند. گاهی اوقات مشکل همخطی در داده‌های جمع‌آوری شده به‌وجود می‌آید که از روش‌های مقابله با آن استفاده از برآوردگرهای انقباضی است. از این‌رو برای تحلیل داده‌های طولی که دارای مشکل همخطی می‌باشند می‌توان از ترکیب مدل‌های آمیخته خطی با رگرسیون ریج از این مهم عبور کرد. همانطور که در نتایج مربوط به مشاهدات شبیه‌سازی ملاحظه می‌شود، برآوردگرهای آمیخته خطی ریج و آمیخته خطی لیو همواره دارای انحراف استاندارد و MSE کمتری نسبت به برآوردگرهای آمیخته خطی می‌باشند، مخصوصاً زمانی که همخطی در داده‌ها قوی‌تر باشد. همچنین در قسمت تحلیل داده‌های واقعی نیز این برتری با استفاده از معیار MPE تأیید می‌شود.

گاهی اوقات تعیین توزیع دقیق متغیر پاسخ در مثال‌های واقعی کار دشواری است در حالی که می‌توان بر اساس داده‌ها به متقارن بودن یا نبودن آن‌ها پی برد لذا از این حیث استفاده از مدل‌های ارائه شده با خطاهای دارای توزیع بیضی‌گون مشکل عدم اطلاع در مورد توزیع دقیق مؤلفه پاسخ برای یافتن برآوردگر ML را مرتفع می‌کند نتایج حاصل از مطالعات شبیه‌سازی نیز نشان دهنده کمتر بودن انحراف استاندارد و MSE برآوردگرها در مدل‌های آمیخته خطی با خطای بیضی‌گون نسبت به برآوردگرهای حاصل از مدل‌های آمیخته خطی است.

گاهی با وجود یک اطلاع پیشین در مورد پارامتر مدل، علاقه‌مند به انجام آزمون فرضیه مورد نظر هستیم لذا با هدف بهبود برآوردگرهای حاصل از مدل‌های آمیخته خطی، برآوردگرهای انقباضی نوع استاین در مدل آمیخته خطی ریج و آمیخته لیو مورد بررسی قرار داده‌شد. برتری برآوردگرهای انقباضی نوع استاین در مدل آمیخته خطی ریج را می‌توان در نتایج حاصل از مطالعات شبیه‌سازی ملاحظه کرد.

۲.۵.۳ پیشنهادات برای پژوهش‌های آینده

طبق نتایج بدست آمده پیرامون مباحث و موضوعات مطرح‌شده در این مجموعه می‌توان به موارد زیر به عنوان پیشنهاد برای پژوهش‌های آینده اشاره کرد:

- استفاده از معیارهای دیگری برای مقایسه برآوردگرهای ارائه‌شده با برآوردگر حاصل از مدل آمیخته خطی.
- گنجاندن تمامی نتایج حاصل از این مجموعه در قالب یک مطالعه شبیه‌سازی با هدف مقایسه برآوردگرهای ارائه‌شده با یکدیگر.
- بررسی رفتار برآوردگر آمیخته تعمیم‌یافته ریج تحت یک مطالعه شبیه‌سازی.
- استفاده از روش‌های دیگری همچون رگرسیون لاسو در مدل آمیخته خطی و مقایسه آن با مدل آمیخته خطی ریج.

مراجع

[۱] آرست، م. (۱۳۹۵)، پایان‌نامه کارشناسی ارشد: ’’مقایسه رفتار برخی برآوردگرهای انقباضی بریج در مدل رگرسیون چندگانه‘‘، دانشکده علوم ریاضی، دانشگاه صنعتی شاهرود.

[۲] آرشی، م. (۱۳۸۷)، رساله دکتری: ’’برآوردگرهای بهبودیافته در مدل رگرسیون خطی چندگانه با خطاهای دارای توزیع بیضی‌گون‘‘، دانشکده علوم ریاضی، دانشگاه فردوسی مشهد.

[۳] برزویی بیدگلی، م. (۱۳۹۳)، پایان‌نامه کارشناسی ارشد: ’’بررسی برآوردگر جک نایف ریج‘‘، دانشکده علوم ریاضی، دانشگاه صنعتی شاهرود.

[۴] تعاونی، م. (۱۳۹۲)، پایان‌نامه کارشناسی ارشد: ’’استفاده از مدل‌های مختلف رگرسیونی در تحلیل داده‌های طولی‘‘، دانشکده علوم، دانشگاه گلستان.

[۵] حسن‌زاده بشتیان، م. (۱۳۸۸)، پایان‌نامه کارشناسی ارشد: ’’برآورد انقباضی در مدل رگرسیون ریج با خطاهای دارای توزیع بیضی‌گون‘‘، دانشکده علوم ریاضی، دانشگاه فردوسی مشهد.

[۶] روزبه، م. (۱۳۹۰)، رساله دکتری: ’’برآورد در مدل‌های خطی جزئی‘‘، دانشکده علوم ریاضی، دانشگاه فردوسی مشهد.

[۷] نوروزی‌راد، م. (۱۳۹۶)، رساله دکتری: ’’برآوردگرهای بهبودیافته در برخی مدل‌های رگرسیونی خطی جریمه‌شده‘‘، دانشکده علوم ریاضی، دانشگاه صنعتی شاهرود.

[8] Arashi, M., Saleh, A.K.Md.E., and Tabatabaey, S.M.M. (2013), “Regression model with elliptically contoured errors”, *Statistics*, 47, pp 1266-1284.

[9] Bancroft, T. A. (1944), “On biases in estimation due to the use of preliminary tests of significance”, *Annal. Math. Statist*, 15, pp 195-204.

[10] Chu, K. C. (1973), “Estimation and decision for linear systems with elliptically random process”, *IEEE. Tran. Aut. Cont.*, 18, pp 499-505.

[11] Cravan, P. and Wahba, G. (1979), "Smoothing noisy data with spline function: Estimating the correct degree of smoothing by the method of generalized cross validation", *Numerische Mathematik*, 31, pp 377-403.

[12] Diggle, P. J., Heagerty, P., Liang, K. Y. and Zeger, S. L. (2002), "*Analysis of Longitudinal Data*", Oxford University Press, New York..

[13] Eliot, M., Ferguson, J., Reilly, M. P and Foulkes, A. S. (2011), "Ridge regression for longitudinal biomarker data", *Biostatistics*, 7, pp 1-11.

[14] Fang, K. T., Kotz, S. and Ng, K. W. (1990), "*Symmeric Multivariate and Related Distributions*", Chapman and Hall, London, New York.

[15] Gentle, J., Hardle, W. and Mori, Y. (2004): "*Handbook of Computational Statistics*", Springer-Verlag Berlin Heidelberg.

[16] Gibbons, D. G. (1981), " A simulation study of some ridge estimators", *J. Amer. Statist. Assc*, 76, pp 1131-1150.

[17] Gupta, A. K. and Varga, T. (1993), "*Elliptically Contored Models in Statistics*", Kluwer Academic Press, London.

[18] Hoerl, A. E. and Kennard, R. W. (1970), "Ridge regression baised estimation for nonorthogonal problems", *Technometrics*, 12, pp 69-89.

[19] James, W. and Stein, C. (1961), "Estimation with quadratic loss", *Proc. of Fouth Berkeley symp. Math. Statist. Prob.*, 1, pp 361-379.

[20] Laird, N. and Ware, J. (1982), "Random-effects models for longitudinal data", *Biometrics*, 38, pp 963-974.

[21] Liu, K. J. (1993), "A new class of biased estimate in linear regression", *Comm. Statist. Theo. Meth.*, 22, pp 393-402.

[22] McCulloch, C. E. and Searle, S. (2001), "*Generalized, Linear, and Mixed Model*", John Wiley, Canada.

[23] McCulloch, C. E., Searle, R. S. and Neuhaus, M. J. (2008), "*Generalized, Linear, and Mixed Models*", John Wiley, Inc., Hoboken, New Jersey, 2nd ed.

[24] McDonald, G. C., Galarnaeu D. I. and Keimel K. (1975), "A Monte Carlo evaiuation of some ridge-type estimators" *Amer. Statist. Assoc*, 20, pp 407- 416.

[25] Montgomery, D. C., Peck, E. A., and Vining, G. G. (2012), "*Introduction to Linear Regression Analysis*", 5th Ed., John Wiley, New Jersey.

[26] Muirhead, R. j. (1982), "*Aspect of Multivariate Statistical Theory*", John Wiley, New York.

[27] Roozbeh, M., Arashi, M., and Kibria, B. M. G. (2015), " Generalized ridge regression estimator in semiparametric regression models", *J. Irania. Statist. Soc. (JIRSS)*, 14 (1), 25-62.

[28] Saleh, A. K. Md. Ehsanes, (2006), "*Theory of Preliminary Test and Stein-Type Estimation with Application*", John Wiley, New Jersey.

[29] Segerstedt, B. (1992), "On ordinary ridge regression in generalized linear models", *Comm. Statist. Theo. Meth.*, 21(8), pp 2227-2246.

پیوست آ

گزیده‌ای از برنامه‌های کامپیوتری

کدهای مربوط به نرم افزار R

رگرسیون خطی و آمیخته خطی در داده‌های $Orthodont$

```
> library(nlme)
> head(Orthodont, 8)

> Linear_Model <- lm(distance ~ age * Sex, data=Orthodont)
> summary(Linear_Model)

> Linear_Mixed_Model <- lme(distance ~ age * Sex, random = ~1|Subject
, data=Orthodont)
> summary(Linear_Mixed_Model)
```

تولید داده‌های همخط به روش پیشنهادی مک دونالد و همکاران (۱۹۷۵) و گیبونز (۱۹۸۱)

```
 p
 n
```

```
rho=0.9


  generator=function(p,n,rho){
    z=matrix(c(rnorm(p*n,mean=5,sd=1)),ncol=p)
    x=matrix(c(rep(0,n*p)),ncol=p)
    for(i in 1:n){
      for(j in 1:p){
        x[i,j]=(1-rho^2)^(1/2)*z[i,j]+rho*z[i,p]
      }
    }
return(x)
  }
  data=generator(p,n,rho)
  x=data
```

تولید داده‌های همخط با استفاده از بسته $EnvStats$

```
#install.packages("EnvStats")
library(EnvStats)
cor.mat=matrix(rho,5,5)
diag(cor.mat)=1
 X <- as.matrix(simulateMvMatrix(ni*n,
    distributions = c("norm","norm","norm","norm","norm"),
    param.list = list( list(mean=5, sd=1),
                       list(mean=5, sd=1),
list(mean=5, sd=1),
list(mean=5, sd=1),
list(mean=5, sd=1)),
      cor.mat = cor.mat))
```

ساختار ماتریس $\mathbf{Z}$

```
###Z
Zi=array(NA,c(q,ni,n))
for(i in 1:n){
Zi[,,i]<-matrix(rnorm(ni*q),q,ni)
```

```
}

#library(magic)
Z<-t(Zi[,,1])
for(i in 2:n){
Z=adiag(Z,t(Zi[,,i]))
}
```

ساختار بردار اثرات تصادفی $\mathbf{b}$

```
#library(tensr)
D=diag(rep(0.6,q))
bi=array(NA,c(q,1,n))
for(i in 1:n){
bi[,,i]<-tensr:::rmvnorm(1,rep(0,q),D)
}
b=as.vector(bi)
```

ساختار بردار خطای تصادفی ϵ

```
###E
sigma2=1
Ei=array(NA,c(ni,1,n))
for(i in 1:n){
Ei[,,i]<-tensr:::rmvnorm(1,rep(0,ni),sigma2*diag(ni))
}
E=as.vector(Ei)
```

محاسبه بردار پاسخ $\mathbf{y}$ با تعریف بردار $\boldsymbol{\beta}$

```
Y=X%*%B+Z%*%b+E
#Y500=Xi[,,500]%*%B+t(Zi[,,500])%*%bi[,,500]+Ei[,,500]
```

ساختار ماتریس واریانس $\mathbf{V}$

```
###V=var(Y)
U<-D
```

```
for(i in 2:n){
U=adiag(U,D)
}
D=U
V=Z%*%D%*%t(Z)+sigma2*diag(N)
```

بردار ضرایب β در مدل آمیخته خطی

```
###B_Mixed
B_Mixed=solve(t(X)%*%solve(V)%*%X)%*%t(X)%*%solve(V)%*%Y
```

بردار ضرایب β در مدل آمیخته خطی ریج

```
lambda0=seq(0.1,200,0.1)
arg0=numeric(length(lambda0))
for(i0 in 1:length(lambda0)){
BMR0=solve(t(X)%*%solve(V)%*%X+lambda0[i0]*diag(p))%*%t(X)%*%solve(V)%*%Y
Yhat0=X%*%BMR0+Z%*%b

S0=X%*%solve(t(X)%*%solve(V)%*%X+diag(lambda0[i0],p))%*%t(X)
%*%solve(V)+Z%*%D%*%t(Z)%*%(diag(N)X%*%(solve(t(X)
%*%solve(V)%*%X+diag(lambda0[i0],p))%*%t(X)%*%solve(V)))

trS0=sum(diag(S0))
arg0[i0]=(1/n)*t(Y-Yhat0)%*%(Y-Yhat0)+2*(n^(-2))*trS0*t(Y-Yhat0)
%*%(Y-Yhat0)
}
found0=cbind(arg0,lambda0)
lambda.min0=found0[found0[,1]==min(arg0),2]
#lambda.min0
BMR=solve(t(X)%*%solve(V)%*%X+lambda.min0*diag(p))%*%t(X)%*%solve(V)%*%Y
```

بردار ضرایب β در مدل آمیخته خطی لیو

```
d=seq(0.01,0.99,0.01)
arg0=numeric(length(d))
```

```
for(i0 in 1:length(d)){
  B_MLiu=solve(t(X)%*%solve(V)%*%X+diag(p))%*%(t(X)%*%solve(V)%*%(X)
  +diag(d[i0],p))%*%B_Mixed

  Yhat0=X%*%B_MLiu+Z%*%b

  S0=X%*%solve(t(X)%*%solve(V)%*%X+diag(p))%*%(t(X)%*%solve(V)%*%X
  +diag(d[i0],p))%*%solve(t(X)%*%solve(V)%*%X)%*%t(X)%*%solve(V)
  +Z%*%D%*%t(Z)%*%(diag(N)X%*%solve(t(X)%*%solve(V)%*%X
  +diag(p))%*%(t(X)%*%solve(V)%*%X+diag(d[i0],p))%*%solve(t(X)
  %*%solve(V)%*%X)%*%t(X)%*%solve(V))

  trS0=sum(diag(S0))
  arg0[i0]=(1/n)*t(Y-Yhat0)%*%(Y-Yhat0)+2*(n^(-2))*trS0*t(Y-Yhat0)
  %*%(Y-Yhat0)
}
found0=cbind(arg0,d)
d.min=found0[found0[,1]==min(arg0),2]
#d.min=0.3
print(d.min)
B_MLiu=solve(t(X)%*%solve(V)%*%X+diag(p))%*%(t(X)%*%solve(V)%*%(X)
+(d.min*diag(p)))%*%B_Mixed
```

برآورد پارامترهای ماتریس کواریانس $\mathbf{V}$ در مدل آمیخته خطی

```
##go on start for convergence algurithm
iter=10000
B_EM_Mixed=matrix(NA,p,iter)
for(k in 1:iter){
Yhat_Mixed=X%*%B_Mixed+Z%*%b
e_Mixed=Y-Yhat_Mixed

###EM-algorithm Estimate D & sigma2
####sigma2 in EM_Algurithm for Mixed
a=0
```

```
for(i in 1:n){
o=e_Mixed[(((i-1)*ni)+1):(i*ni),]
bb=t(o)%*%o
a=a+bb
}
t1_Mixed=sum(a)/n
sigma2_Mixed=t1_Mixed/N

####D in EM_Algurithm
aa=matrix((0),q,q)
for(i in 1:n){
O=b[(((i-1)*q)+1):(i*q)]
bb=O%*%t(O)
aa=aa+bb
}
t2=(1/n)*aa
D=t2/n

####Now update Sigma2 & D for E & b
###Update b for Mixed
bi=array(NA,c(q,1,n))
for(i in 1:n){
bi[,,i]<-tensr:::rmvnorm(1,rep(0,q),D)
}
b=as.vector(bi)

###Update E for Mixed
Ei=array(NA,c(ni,1,n))
for(i in 1:n){
Ei[,,i]<-tensr:::rmvnorm(1,rep(0,ni),sigma2_Mixed*diag(ni))
}
E=as.vector(Ei)

##########################################update Y
```

```
Y=X%*%B_Mixed+Z%*%b+E
### product D then Update V
U<-D
for(i in 2:n){
U=adiag(U,D)
}
D=U
V_Mixed=Z%*%D%*%t(Z)+sigma2_Mixed*diag(N)

### update B_Mixed
B_Mixed=solve(t(X)%*%solve(V_Mixed)%*%X)%*%t(X)%*%solve(V_Mixed)%*%Y
B_EM_Mixed[,k]=B_Mixed
}
B_EM_Mixed
```

برآورد پارامترهای ماتریس کواریانس $\mathbf{V}$ در مدل MR و محاسبه انحراف استاندارد

```
###SD.....................................
iter_Sd=500
B_MR_Sd_matrix=matrix((NA),p,iter_Sd)
B_MR_mean_matrix=matrix((NA),p,iter_Sd)
for(s in 1: iter_Sd){
print(s)
D=diag(rep(0.6,q))
bi=array(NA,c(q,1,n))
for(i in 1:n){
bi[,,i]<-tensr:::rmvnorm(1,rep(0,q),D)
}
b=as.vector(bi)

###E
sigma2=1
Ei=array(NA,c(ni,1,n))
for(i in 1:n){
Ei[,,i]<-tensr:::rmvnorm(1,rep(0,ni),sigma2*diag(ni))
```

```
}
E=as.vector(Ei)

###Mixed.Model & Y
Y=X%*%B+Z%*%b+E
dim(b)
#Y500=Xi[,,500]%*%B+t(Zi[,,500])%*%bi[,,500]+Ei[,,500]

###V=var(Y)
U<-D
for(i in 2:n){
U=adiag(U,D)
}
D=U
V=Z%*%D%*%t(Z)+sigma2*diag(N)

###B_Mixed0
lambda0=seq(0,200,1)
arg0=numeric(length(lambda0))
for(i0 in 1:length(lambda0)){
BMR0=solve(t(X)%*%solve(V)%*%X+lambda0[i0]*diag(p))%*%t(X)%*%solve(V)%*%Y
Yhat0=X%*%BMR0+Z%*%b

S0=X%*%solve(t(X)%*%solve(V)%*%X+diag(lambda0[i0],p))%*%t(X)
%*%solve(V)+Z%*%D%*%t(Z)%*%(diag(N)X%*%(solve(t(X)
%*%solve(V)%*%X+diag(lambda0[i0],p))%*%t(X)%*%solve(V)))

trS0=sum(diag(S0))
arg0[i0]=(1/n)*t(Y-Yhat0)%*%(Y-Yhat0)+2*(n^(-2))*trS0*t(Y-Yhat0)
%*%(Y-Yhat0)
}
found0=cbind(arg0,lambda0)
lambda.min0=found0[found0[,1]==min(arg0),2]
#lambda.min0
```

```
BMR=solve(t(X)%*%solve(V)%*%X+lambda.min0*diag(p))%*%t(X)%*%solve(V)%*%Y



##go on start for convergence algurithm
iter=10000
what.lam=matrix(NA,iter_Sd,iter)
B_EM_MR=matrix(NA,p,iter)
for(k in 1:iter){
Yhat_MR=X%*%BMR+Z%*%b
e_MR=Y-Yhat_MR

###EM-algorithm Estimate D & sigma2
####sigma2 in EM_Algurithm for Mixed
a=0
for(i in 1:n){
o=e_MR[(((i-1)*ni)+1):(i*ni),]
bb=t(o)%*%o
a=a+bb
}
t1_MR=sum(a)/n
sigma2_MR=t1_MR/N

####D in EM_Algurithm
aa=matrix((0),q,q)
for(i in 1:n){
O=b[(((i-1)*q)+1):(i*q)]
bb=O%*%t(O)
aa=aa+bb
}
t2=(1/n)*aa
D=t2/n

####Now update Sigma2 & D for E & b
###Update b for Mixed
```

```
bi=array(NA,c(q,1,n))
for(i in 1:n){
bi[,,i]<-tensr:::rmvnorm(1,rep(0,q),D)
}
b=as.vector(bi)

### product D then Update V
U<-D
for(i in 2:n){
U=adiag(U,D)
}
D=U
V_MR=Z%*%D%*%t(Z)+sigma2_MR*diag(N)

### update B_Mixed
lambda1=seq(0,200,1)
arg1=numeric(length(lambda1))
for(i1 in 1:length(lambda1)){

BMR1=solve(t(X)%*%solve(V_MR)%*%X+lambda1[i1]*diag(p))
%*%t(X)%*%solve(V_MR)%*%Y

Yhat1=X%*%BMR1+Z%*%b

S1=X%*%solve(t(X)%*%solve(V_MR)%*%X+diag(lambda1[i1],p))
%*%t(X)%*%solve(V_MR)+Z%*%D%*%t(Z)%*%(diag(N)X%*%(solve(t(X)
%*%solve(V_MR)%*%X+diag(lambda1[i1],p))%*%t(X)%*%solve(V_MR)))

trS1=sum(diag(S1))
arg1[i1]=(1/n)*t(Y-Yhat1)%*%(Y-Yhat1)+2*(n^(-2))*trS1*t(Y-Yhat1)
%*%(Y-Yhat1)
}
found1=cbind(arg1,lambda1)
```

```
lambda.min1=found1[found1[,1]==min(arg1),2]
#lambda.min1

BMR=solve(t(X)%*%solve(V_MR)%*%X+lambda.min1*diag(p))%*%t(X)
%*%solve(V_MR)%*%Y

B_EM_MR[,k]=BMR
}

B_MR_mean_matrix[,s]=B_EM_MR[,iter]
B_MR_Sd_matrix[,s]=((B_EM_MR[,iter]-B)^2)
}
mean_B_MR=t(t(apply(B_MR_mean_matrix,1,mean)))
Sd_B_MR=t(t(sqrt(apply(B_MR_Sd_matrix,1,mean))))

final_list=list(cbind(rho,B,mean_B_MR,Sd_B_MR))
names(final_list)=c("MIXED_RIDGE")
final_list
#write.table(final_list,"E:/finaltableMixed&MRrho0.0itersd500n50
.txt",sep="\t")
```

بردار خطاهای دارای توزیع بیضی‌گون (مثال: $t-student$ چند متغیره)

```
library(mvtnorm)
D=diag(rep(0.3,q))
bi=array(NA,c(q,1,n))
for(i in 1:n){
bi[,,i]<-rmvt(1, D, df = 5)
}
b=as.vector(bi)

###E
sigma2=1
Ei=array(NA,c(ni,1,n))
for(i in 1:n){
```

```
Ei[,,i]<-rmvt(1, sigma2*diag(ni), df = 5)
}
E=as.vector(Ei)
```

محاسبه λ_0 به روش هورل و کنارد (۱۹۷۰) ـ (تجزیه طیفی ماتریس $\mathbf{X}^T\mathbf{X}$)

```
t1_Mixed=a
sigma2_Mixed=as.numeric(t1_Mixed/N)
XTX=t(X)%*%X
GAMMA=eigen(XTX)$vectors
LAMBDA=diag(eigen(XTX)$values)
GAMMA%*%LAMBDA%*%t(GAMMA)
ALPHA=t(GAMMA)%*%B_Mixed
ALPHA2MAX=(max(ALPHA))^2
lambda0=sigma2_Mixed/ALPHA2MAX
```

برآوردگرهای $PTMR$ و $SSMR$

```
############Hypotheses
#H0: c (B_Mixed)=d
#in this problem; c=1 ; d=0
#Wald Statistic
W=drop(t(B_Mixed)%*%(t(X)%*%solve(V_Mixed)%*%X)%*%(B_Mixed))/N

##############Stein
Ln_Mixed=W
Ln_MR=W
C=((N-p)*(p-2))/(p*(N-p+2))
B_Shrinkage_Mixed=B_Mixed-B_Mixed*(C/as.numeric(Ln_Mixed))
B_Shrinkage_MR=BMR-BMR*(C/as.numeric(Ln_MR))

#####################################PT
alpha=c(0.001,0.01,0.05,0.1,0.5)
Iln=c(NA)
for(cont in 1:length(alpha)){
Iln[cont]=0
```

```
if (Ln_MR<qchisq(alpha[cont],df=p,lower.tail=FALSE)) {Iln[cont]=1}
}
B_PTMR_alpha0.001=BMR-BMR*Iln[1]
B_PTMR_alpha0.01=BMR-BMR*Iln[2]
B_PTMR_alpha0.05=BMR-BMR*Iln[3]
B_PTMR_alpha0.1=BMR-BMR*Iln[4]
B_PTMR_alpha0.5=BMR-BMR*Iln[5]
```

واژه‌نامه فارسی به انگلیسی

آزمون اولیه آمیخته ریج Preliminary test mixed ridge
آماره والد Wald statistic
اثر ثابت Fixed effect
اثر تصادفی Random effect
اثر متقابل Interaction
اریبی Bias
استقلال Independence
اعتبار سنجی متقابل Cross-Validation
الگوریتم Algorithm
انحراف استاندارد Standard deviation
انقباضی نوع استاین Stein-type shrinkage
برآوردگر انقباضی Shrinkage estimator
برآوردگر لیو Liu estimator
بردار ویژه Eigen vector
بسل چندمتغیره Multivariate Bessel
پارامتر ریج Ridge parameter
پارامتر مقیاس Scale parameter
پیرسن چندمتغیره Multivariate Pearson
تابع مولد چگالی Density generator function
تابع مولد مشخصه Characteristic generator function
تابع نشان‌گر Indicator function
تبادل‌پذیر Exchangeable
توزیع بیضی‌گون Elliptically contoured distribution
جریمه Penalty
داده‌های طولی Longitudinal data
دترمینان Determinant

رتبه کامل ..... Full Rank
رگرسیون ریج ..... Ridge regression
شرایط نظم ..... Regularity Conditions
عامل تورم واریانس ..... Variance inflation factor
عدد شرطی ..... Condition number
کمترین توان‌های دوم ..... Least squares
کمترین توان‌های دوم موزون تکراری ..... Iterrative weighted least squares method
کوشی چندمتغیره ..... Multivariate Cauchy
لاپلاس چندمتغیره ..... Multivariate Laplace
ماتریس اطلاع فیشر ..... Fisher information matrix
ماتریس طرح ..... Design matrix
ماتریس قطری ..... Diagonal matrix
ماکزیمم درستنمایی ..... Maximum likelihood
متعامد ..... Orthogonal
متغیر پاسخ ..... Response variable
متغیر مستقل ..... Independent Variable
مدل انتقال ..... Transition model
مدل حاشیه‌ای ..... Marginal model
مدل خطی آمیخته ..... Linear mixed model
مدل خطی تعمیم یافته ..... Generalized linear model
میانگین توان‌های دوم خطا ..... Mean squares error
میانگین خطای پیش‌گویی ..... Mean prediction error
مقدار ویژه ..... Eigen value
نمایی توانی چندمتغیره ..... Multivariate exponential power
همبستگی ..... Correlation
همخطی ..... Collinearity

واژه‌نامه انگلیسی به فارسی

الگوریتم ..... Algorithm
اریبی ..... Bias
تابع مولد مشخصه ..... Characteristic generator function
همخطی ..... Collinearity
عدد شرطی ..... Condition number
اعتبار سنجی متقابل ..... Cross-Validation
همبستگی ..... Correlation
تابع مولد چگالی ..... Density generator function
ماتریس طرح ..... Design matrix
دترمینان ..... Determinant
ماتریس قطری ..... Diagonal matrix
مقدار ویژه ..... Eigen value
بردار ویژه ..... Eigen vector
توزیع بیضی‌گون ..... Elliptically contoured distribution
تبادل‌پذیر ..... Exchangeable
ماتریس اطلاع فیشر ..... Fisher information matrix
اثر ثابت ..... Fixed effect
رتبه کامل ..... Full Rank
مدل خطی تعمیم‌یافته ..... Generalized linear model
استقلال ..... Independence
متغیر مستقل ..... Independent Variable
تابع نشان‌گر ..... Indicator function
اثر متقابل ..... Interaction
کمترین توان‌های دوم موزون تکراری ..... Iterrative weighted least squares method
کمترین توان‌های دوم ..... Least squares
مدل آمیخته خطی ..... Linear mixed model

برآوردگر لیو ..... Liu estimator
داده‌های طولی ..... Longitudinal data
مدل حاشیه‌ای ..... Marginal model
ماکزیمم درستنمایی ..... Maximum likelihood
میانگین خطای پیش‌گویی ..... Mean prediction error
میانگین توان‌های دوم خطا ..... Mean squares error
بسل چندمتغیره ..... Multivariate Bessel
کوشی چندمتغیره ..... Multivariate Cauchy
نمایی توانی چندمتغیره ..... Multivariate Exponential power
لاپلاس چندمتغیره ..... Multivariate Laplace
پیرسن چندمتغیره ..... Multivariate Pearson
متعامد ..... Orthogonal
جریمه ..... Penalty
آزمون اولیه آمیخته ریج ..... Preliminary test mixed ridge
اثر تصادفی ..... Random effect
شرایط نظم ..... Regularity Conditions
متغیر پاسخ ..... Response variable
پارامتر ریج ..... Ridge parameter
رگرسیون ریج ..... Ridge regression
پارامتر مقیاس ..... Scale parameter
برآوردگر انقباضی ..... Shrinkage estimator
انحراف استاندارد ..... Standard deviation
انقباضی نوع استاین ..... Stein-type shrinkage
مدل انتقال ..... Transition model
عامل تورم واریانس ..... Variance inflation factor
آماره والد ..... Wald statistic

Abstract

Longitudinal data are those observations in which measurements of each experiment units repeats over the time. Linear mixed model is one of the tools used to analyze such data. However, pre sense of multicolinearity causes to use shrinkage methods for analysis. In this dissertation, the combination of linear mixed elliptical models with shrinkage methods such as ridge and Liu is considered for analysis purposes. The performance of the proposed estimators is evaluated using simulation and real data analysis.

Keywords: EM algorithm, Elliptical distribution, Linear mixed models, Liu estimator, Ridge regression, Shrinkage estimator.

Shahrood University of Technology

Faculty of Mathematical Sciences

MSc Thesis in: Mathematical Statistics

Likelihood-based Inference in Ridge Regression for Longitudinal Data

By: Mohammad Rahmani

Supervisor

Dr. Mohammad Arashi

September 2018