سیزدهمین کنفرانس آمار ایران
۲-۴ شهریور ۱۳۹۵

برآورد پارامتر ریج به کمک ماکسیمم آنتروپی تی سالیس مرتبه دو تعمیم‌یافته

منیژه صانعی طبس[۱]، غلامرضا محتشمی برزادران[۲]

[۱]دانشگاه سیستان و بلوچستان،گروه آمار

[۲]دانشگاه فردوسی مشهد، گروه‌آمار

چکیده: یکی از مسائلی که تحلیل رگرسیون به روش کمترین توان های دوم را با مشکلات عمده مواجه می سازد وجود هم خطی در بین متغیرهای رگرسیونی است. روش های زیادی برای حل مشکلات ناشی از وجود هم خطی معرفی شده اند. یکی از این روش ها رگرسیون ریج است در این مقاله یک برآورد جدید برای پارامتر ریج به کمک ماکسیمم آنتروپی تی سالیس تعمیم یافته ارائه داده وآن را برآوردگر ریج ماکسیمم آنتروپی تی سالیس مرتبه دو تعمیم یافته می نامیم. برای مجموعه داده های سیمان پرتلند که از هم خطی قوی برخوردار هستند این برآورد گر را محاسبه و با برآوردگر ریج ماکسیمم آنتروپی تعمیم یافته و برآوردگر کمترین توان های دوم مقایسه می کنیم.

واژه‌های کلیدی: رگرسیون ریج، ماکسیمم آنتروپی تعمیم یافته، آنتروپی تی‌سالیس، ماکسیمم آنتروپی تی سالیس تعمیم یافته

کد موضوع‌بندی ریاضی (۲۰۱۰): 62B10، 62J07.

۱ مقدمه

روش ماکسیمم آنتروپی تعمیم یافته[۱] (GME) با یک زیربنای اقتصادی توسط جاج و گلان (۱۹۹۲) معرفی شد. روش (GME) برای برآورد پارامترهای مدل رگرسیونی خطی و غیر خطی بدون اعمال هیچ شرطی روی توزیع احتمال خطاها مورد استفاده واقع شده است. این روش که براساس ماکسیمم آنتروپی شانون و اصل ماکسیمم آنتروپی جینز (۱۹۵۷) پایه گذاری شده است مدل خطی را به گونه ای بازنویسی می کند که پارامترها و خطاهای مجهول را به شکل احتمالات در نظر گرفته و آنان را برآورد می کند. مزیت روش GME این است که در این روش لازم نیست هیچگونه فرض یاتوزیع احتمالی رابرای باقیمانده ها در نظر بگیریم حتی در مواقعی که نمی خواهیم استنباط آماری انجام دهیم. اما بعد از شانون (۱۹۴۸) تاکنون اندازه های آنتروپی تعمیم یافته زیادی معرفی گردید که هرکدام در حالت خاص آنتروپی شانون را در بر می گیرند. از جمله پرکاربردترین آنها آنتروپی تی سالیس (۱۹۸۸) است. ماکسیمم آنتروپی که تکیه بر آنتروپی شانون دارد توزیع های نمایی را نتیجه می دهد. آن گونه که مطالعات تجربی نشان می دهد در بسیاری از مواقع توزیع جامعه

[۲]منیژه صانعی طبس : manijesanei@gmail.com

[۱]Generalized Maximum Entropy

تحت بررسی از نوع توزیع های توانی بوده(بسیاری ازمدل های اقتصادی) بنابراین استفاده از آنتروپی تی سالیس به عنوان یک رده کلی تر از آنتروپی ها به جای آنتروپی شانون پیشنهاد شده و موضوع ماکسیمم آنتروپی تی سالیس بررسی شده است. براین اساس می توان با جایگزین کردن آنتروپی تی سالیس به جای آنتروپی شانون در روش ماکسیمم آنتروپی تعمیم یافته، این روش را برای آنتروپی مرتبه α تعمیم داده و رده دیگری از برآوردها به نام $GME\alpha$ را معرفی کرد. به عبارت دیگرمجموع آنتروپی تی سالیس مرتبه α دو جمله احتمالی و جمله خطا را ماکسیمم می کنیم. با توجه به کاربردهای زیاد آنتروپی تی سالیس و به خصوص آنتروپی تی سالیس مرتبه دو در بسیاری از علوم از جمله شیمی، فیزیک، زیست آزمایشگاهی و اقتصاد و نیز با قراردادن $\alpha = ۲$ روش ماکسیمم آنتروپی تعمیم یافته را با قراردادن آنتروپی تی سالیس مرتبه دو به جای آنترپی شانون تعمیم داده و مجموع آنتروپی تی سالیس مرتبه دو هر دو جمله آنترپی و جمله خطا را ماکسیمم می کنیم. برآوردهای به دست آمده از این روش را برآوردگرهای ماکسیمم آنتروپی تی سالیس مرتبه دو تعمیم یافته ($GMET۲$) می نامیم. یکی از مشکلاتی که داده ها با آن مواجه هستند وجود هم خطی در داده هاست که تحلیل رگرسیونی را با مشکلات اساسی مواجه می سازد. به دلیل اهمیت این مساله در ادامه این مقاله بعد از معرفی روش $GMET$ و برآوردگرهای $GMET۲$ برای یک مدل رگرسیونی که از هم خطی بالایی برخوردار است برآوردگرهای $GMET۲$ را محاسبه و با برآوردگرهای OLS و GME مقایسه می کنیم. روش های زیادی برای برخورد با مشکلات ناشی از هم خطی پیشنهاد شده است. شامل جمع آوری داده های اضافی، تخمین مجدد مدل، بکارگیری روش های برآورد به غیر ازکمترین توان های دوم که به طور ویژه برای ازبین بردن مشکلات ایجاد شده از هم خطی چندگانه طراحی شده اند. روش های متعددی شامل رگرسیون ریج، رگرسیون مولفه های اصلی، رگرسیون کمترین توان های دوم جزئی، رگرسیون زنجیری و رگرسیون تحت محدودیت های تصادفی برای مقابله با مشکل هم خطی چندگانه گسترش یافته اند. برآوردگر GME خود یک روش برای حل مشکل هم خطی است. برآوردگر $GMET۲$ که در این مقاله معرفی می شود روشی دیگر برای حل این مشکل ارائه می دهد. مدت های طولانی است که رگرسیون ریج به عنوان راه حلی برای مشکل هم خطی و برآورد پارامترهای مدل رگرسیونی مطرح شده است. ماسدو و همکاران (۲۰۱۰) براساس روش GME برآوردگر ریج ماکسیمم آنتروپی تعمیم یافته را ارائه داده اند. در ادامه به کمک برآوردگرهای $GMET۲$، برآورد جدیدی از پارامتر ریج به نام برآورد ریج ماکسیمم آنتروپی تی سالیس مرتبه دو تعمیم یافته معرفی می کنیم. برای یک مجموعه داده که از هم خطی بالایی برخوردار هستند این روش را اعمال نموده و برآوردگر ریج$GMET۲$ را محاسبه می کنیم.

۲ برآوردگر رگرسیون ریجی

در این بخش برآوردگر رگرسیون ریجی[۲] مورد نظر قرار می گیرد. مدل رگرسیون خطی

$$Y = X\beta + e \qquad (۱.۲)$$

که در رابطه (۱.۲) ، Y بردار $N \times ۱$ از مشاهدات و β یک بردار $K \times ۱$ از پارامترهای مجهول و X یک ماتریس $(N \times K)$ از متغیرهای توضیحی و معلوم است. e نیز بردار $(N \times ۱)$ خطاهای تصادفی است. به طوری که

$$E(e|X) = ۰, E(ee'|X) = \sigma^۲ I$$

[۲] Ridge regression

که I ماتریس همانی $(N \times N)$ است. برآوردگرکمترین توان های دوم β به صورت

$$\hat{\beta}_{OLS} = (X'X)^{-1}X'Y \tag{۲.۲}$$

و برآوردگر رگرسیون ریجی که توسط ارل و کنارد (۱۹۷۰) پیشنهاد شد عبارت است از:

$$\hat{\beta}_{RR} = (X'X + \eta I)^{-1}X'Y \tag{۳.۲}$$

که $\eta \geq 0$ نشان دهنده پارامتر ریج است. وقتی که $\eta \to 0$ برآوردگر RR به برآوردگر OLS نزدیک می شود در حالی که وقتی $\eta \to \infty$ برآوردگر RR به بردار صفر میل می‌کند. پس یک افزایش و کاهش بین واریانس و اریبی لازم است. آنها نشان دادند که برآوردگر RR برای یک دامنه مقادیر η (وقتی که ملاک مقایسه دو برآوردگر MSE باشد) یک کران بالا برای برآوردگر کمترین توان های دوم است. مقدار MSE برآوردگر RR عبارت است از:

$$MSE(\hat{\beta}_{RR}) = \sigma^2 \sum_{i=1}^{k} \frac{\lambda_i}{(\lambda_i + \eta)^2} + \eta^2 \beta' (X'X + \eta I)^{-2} \beta \tag{۴.۲}$$

که در (۴.۲) λ_i ها مقادیر ویژه ماتریس همبستگی $X'X$ هستند.

۳ برآوردگرریج ماکسیمم آنتروپی تی سالیس مرتبه دو تعمیم یافته

آنتروپی تی‌سالیس مرتبه α یکی از اشکال تعمیم یافته آنتروپی شانون است که وقتی $\alpha \to 1$ همان آنتروپی شانون را نتیجه می دهد. به‌ازای $\alpha = 2$ آنتروپی تی سالیس مرتبه ۲ حاصل شده که کاربردهای خاص خودش را در بسیاری از علوم پیدا کرده است.

تعریف ۱.۳. برای متغیر تصادفی X با تابع احتمال مربوطه $P = (p_1, ...p_n)$ آنتروپی تی‌سالیس و آنتروپی تی سالیس مرتبه دو به ترتیب عبارتند از:

$$\begin{aligned} T_\alpha(P) &= \frac{1}{\alpha - 1}[\sum_{i=1}^{n} p^\alpha(x_i) - 1], \alpha > 0, \alpha \neq 1 \\ T_2(P) &= \sum_{i=1}^{n} p^2(x_i) - 1 = \sum_{i=1}^{n} p(x_i)[p(x_i) - 1]. \end{aligned} \tag{۱.۳}$$

روش GME براساس آنتروپی شانون تعریف شده است. صانعی و محتشمی (۲۰۱۵) ضمن جایگزینی آنتروپی شانون با آنتروپی تی سالیس مرتبه دو روش GME را تعمیم داده و یک خانواده جدید از برآوردگرهای نظریه اطلاع معرفی و آن را برآوردگرهای ماکسیمم آنتروپی تی سالیس تعمیم یافته [۳] ($GMET2$) نام گذاری کردند. روش یافتن این برآوردگرها همانند برآوردگر GME بوده بااین تفاوت که به جای آنتروپی شانون از آنتروپی تی سالیس مرتبه دو استفاده می‌شود. در این مقاله براساس برآوردگر $GMET2$ برآوردگر دیگری برای پارامتر ریج معرفی کرده و آنرا با سایر برآوردگرهای ریج ذکرشده در بخش قبل مقایسه می‌کنیم. بدین منظور ابتدا روش یافتن برآوردگرهای $GMET2$ توضیح داده می شود.

[۳] Generalized maximum Tsallis entropy

۱.۳ برآوردگر ماکسیمم آنتروپی تی سالیس مرتبه دو تعمیم یافته

برای به دست آوردن برآوردگر $GMET۲$ همانند روش یافتن GME، مدل خطی (۱.۲) را دوباره بازنویسی می کنیم بطوریکه برآوردها و خطاهای مجهول به فرم احتمالات نوشته می شوند زیرا عناصر تابع ماکسیمم آنتروپی، احتمالات هستند. فرض کنید هر β_k یک متغیر تصادفی گسسته با M برآمد ممکن $(z_{k۱}, z_{k۲}, ..., z_{kM})$ هر یک با احتمال متناظر $(p_{k۱}, p_{k۲}, ..., p_{kM})$ باشد به طوریکه $\sum_{m=۱}^{M} p_{km} = ۱$. براین اساس هر پارامتر β_k را می‌توان به صورت زیر نوشت:

$$\beta_k = \sum_{m=۱}^{M} z_{km} p_{km}, \qquad k = ۱, ۲, ..., K, \ \ , M \geq ۲, \tag{۲.۳}$$

تکیه گاه پارامتری یا مقادیر بردار Z_k بر اساس اطلاع پیشین و یا تجارب قبلی تعیین می شود. به طور مشابه برای برآورد توزیع خطاها نیز می توان به این ترتیب عمل کرد که فرض می‌کنیم تکیه‌گاه خطا شامل $J \geq ۲$ نقطه و ν نیز یک ماتریس $(T \times TJ)$ از نقاط تکیه گاه بوده و W نیز یک بردار $TJ \times ۱$ از وزن های متناظر نقاط تکیه گاه باشند به طوریکه $\sum_{j=۱}^{J} w_{tj} = ۱$. در این صورت هر بردار e_t به صورت زیر بازنویسی می شود:

$$e_t = \sum_{j=۱}^{J} \nu_{tj} w_{tj}, \quad t = ۱, ۲, ..., T, J \geq ۲, \tag{۳.۳}$$

برای تعیین کران های تکیه گاه خطاها، گلان و همکاران (۱۹۹۶) از قانون ۳σ که توسط پوکلشیم (۱۹۹۴) مطرح شده بود استفاده کردند. آنها کران پایین را -۳σ قرار داده و برای کران بالا از $+۳\sigma$ استفاده کردند که مقدار σ نیز همان انحراف معیار نمونه‌ای بردار Y می باشد. براین اساس هر پارامتر β_k و e_t به ترتیب به صورت (۲.۳) و (۳.۳) بازنویسی می‌شوند. پس مسئله ماکسیمم آنتروپی تی سالیس تعمیم‌یافته به صورت:

$$\begin{aligned} \max_{P,W} H(P,W) &= -P'(P'-۱) - W'(W'-۱) \\ &= -\sum_{k=۱}^{K}\sum_{m=۱}^{M} p_{km}(p_{km}-۱) - \sum_{t=۱}^{T}\sum_{j=۱}^{J} w_{tj}(w_{tj}-۱), \end{aligned} \tag{۴.۳}$$

s.t:

$$\begin{cases} \sum_{k=۱}^{K}\sum_{m=۱}^{M} x_{tk} z_{km} p_{km} + \sum_{j=۱}^{J} w_{tj}\nu_{tj} = Y_t, & t = ۱, ۲, ..., T. \\ \sum_{m=۱}^{M} p_{km} = ۱, & k = ۱, ۲, ..., K. \\ \sum_{j=۱}^{J} w_{tj} = ۱, & t = ۱, ۲, ..., T. \end{cases} \tag{۵.۳}$$

است. برای حل مسئله بهینه‌سازی (۴.۳) با توجه به محدودیت‌های ذکرشده معادله لاگرانژ راتشکیل داده که حل این معادله نتایج زیر:

$$\hat{p}_{km} = \frac{۱-M}{M} + \frac{۱}{۲M}\sum_{t=۱}^{T}\lambda_t x_{tk}\sum_{m=۱}^{M} z_{km} - \frac{۱}{۲}\sum_{t=۱}^{T}\lambda_t x_{tk} z_{km}, \tag{۶.۳}$$

$$m = ۱, ۲, ..., M, k = ۱, ۲, ..., K.$$

$$\hat{w}_{tj} = -\frac{۱}{۲}\lambda_t\nu_{tj} + \frac{۱}{۲J}\lambda_t(\sum_{j=۱}^{J}\nu_{tj}), \quad j = ۱, ..., J.$$

را داریم. با جایگذاری جواب های $\hat{p}_{km}$ و $\hat{w}_{tj}$ به ترتیب در هر یک از معادلات (۲.۳) و (۳.۳)برآوردهای ماکسیمم آنتروپی تی‌سالیس تعمیم‌یافته GME به صورت زیر حاصل می‌شوند:

$$\beta_k^{GME} = \sum_{m=1}^{M} \hat{p}_{km} z_{km}, k = 1, 2, ..., K,$$
$$e_t^{GME} = \sum_{j=1}^{J} \hat{w}_{tj} \nu_{tj}, t = 1, 2, ..., T.$$

۲.۳ برآوردگر ریج ماکسیمم آنتروپی تی‌سالیس مرتبه دو تعمیم‌یافته

براساس برآوردگر $GMET2$ که در بالا به دست آمد، برآوردگر ریج ماکسیمم آنتروپی تی‌سالیس تعمیم‌یافته[۴] $(Ridge - GMET2)$ به صورت:

$$\hat{\eta}_{Ridge-GMET2} = min_{\eta \in [a_1, a_2]} ||Z\hat{P} - (X'X + \eta I)^{-1} X'Y||_\infty \qquad (۷.۳)$$

تعریف می شوند. بردار $\hat{P}$ در رابطه (۷.۳)، همان برداررابطه (۶.۳) به دست آمده از روش $GMET2$ که رابطه (۴.۳) را ماکسیمم می‌کند است. بازه ریجی در نظر گرفته‌شده بازه $[a_1, a_2]$ است که در آن a_1 و a_2 به ترتیب کران‌های بالا و پایین این بازه هستند.

۴ مثال عددی

در این بخش براساس برآوردگر ریج ماکسیمم آنتروپی تی سالیس مرتبه دو تعمیم یافته، که در بخش قبل ارائه شد برآوردگر رگرسیون ریجی را برای مجموعه داده سیمان پرتلند که از درجه هم خطی بالایی برخوردار است محاسبه می کنیم. با توجه به وجود هم خطی در این داده ها بعد از وود و همکاران (۱۹۳۲)

برآوردهای مختلفی برای برآورد پارامتر ریج ارائه شده است از جمله می توان به هالد (۱۹۵۲) ، کاکرانلار و همکاران (۱۹۹۹) ، لیو (۲۰۰۳) و مونیز و کیبریا (۲۰۰۹) اشاره کرد. در این مجموعه برای اندازه گیری میزان سختی سیمان پرتلند متغیر پاسخ Y (میزان گرمای آزادشده از هر گرم سیمان) تعریف شده است که خود تحت تاثیر چهار متغیررگرسیونی X_1:(آلومینات تری کلسیم)، X_2: (سیلیکات تری کلسیم)، X_3 (آلومینو فریت تتراکلسیم) و X_4 (بتا دی کلسیم سیلیکات) است و به ترتیب چهار ستون ماتریس X را تشکیل می دهند.

[۴] Ridge Gneralized Maximum Tsallis Entropy

این مجموعه داده را در زیر خلاصه کرده ایم.

$$X = \begin{bmatrix} ۷ & ۲۶ & ۶ & ۶۰ \\ ۱ & ۲۹ & ۱۵ & ۵۲ \\ ۱۱ & ۵۶ & ۸ & ۲۰ \\ ۱۱ & ۳۱ & ۸ & ۴۷ \\ ۷ & ۵۲ & ۶ & ۳۳ \\ ۱۱ & ۵۵ & ۹ & ۲۲ \\ ۳ & ۷۱ & ۱۷ & ۶ \\ ۱ & ۳۱ & ۲۲ & ۴۴ \\ ۲ & ۵۴ & ۱۸ & ۲۲ \\ ۲۱ & ۴۷ & ۴ & ۲۶ \\ ۱ & ۴۰ & ۲۳ & ۳۴ \\ ۱۱ & ۶۶ & ۹ & ۱۲ \\ ۱۰ & ۶۸ & ۸ & ۱۲ \end{bmatrix}, \sum_{i=۱}^{۴} x_i = \begin{bmatrix} ۹۹ \\ ۹۷ \\ ۹۵ \\ ۹۷ \\ ۹۸ \\ ۹۷ \\ ۹۷ \\ ۹۸ \\ ۹۶ \\ ۹۸ \\ ۹۸ \\ ۹۸ \\ ۹۸ \end{bmatrix}, Y = \begin{bmatrix} ۷۸/۵ \\ ۷۴/۳ \\ ۱۰۴/۳ \\ ۸۷/۶ \\ ۹۵/۹ \\ ۱۰۹/۲ \\ ۱۰۲/۷ \\ ۷۲/۵ \\ ۹۳/۱ \\ ۱۱۵/۹ \\ ۸۳/۸ \\ ۱۱۳/۳ \\ ۱۰۹/۴ \end{bmatrix}$$

آدنیز و همکاران (۲۰۱۱)مدل رگرسیونی

$$y = \beta_۰ + \beta_۱ x_۱ + \beta_۲ x_۲ + \beta_۳ x_۳ + \beta_۴ x_۴ + e. \tag{۱.۴}$$

را به این مجموعه داده ها برازش داده و برآوردهای OLS را محاسبه نموده اند. هیچ یک از ضرایب در سطح $\alpha = ۰/۰۵$ معنی دار

جدول ۱: جدول آنالیز واریانس داده های سیمان پرتلند به روش کمترین توان های دوم

جمله	پارامتر	خطای معیار	آماره های توزیع تی	مقادیر بحرانی آزمون
$\beta_۰$	۶۲/۴۰۵۴	۷۰/۰۷۱۰	۰/۸۹	۰/۳۹۹۱
$\beta_۱$	۱/۵۵۱۱	۰/۷۴۴۸	۲/۰۸	۰/۰۷۰۸
$\beta_۲$	۰/۵۱۰۲	۰/۷۲۳۸	۰/۷۰	۰/۵۰۰۹
$\beta_۳$	۰/۱۰۱۹	۰/۷۵۴۷	۰/۱۴	۰/۸۹۵۹
$\beta_۴$	−۰/۱۴۴	۰/۷۰۹۱	−۰/۲	۰/۸۴۴۱

نبوده و $\hat{\sigma}^۲ = ۵/۹۸۲۹$ همچنین $R^۲ = ۰/۹۸۲۴$ است که نشان می دهد ۹۸/۲۴ درصد تغییرات توسط متغیرهای رگرسیونی پیشگو توضیح داده می‌شود. مقادیر ویژه برای ماتریس X برابر:

$\lambda_۱ = ۲۱۱/۳۶۷۵, \lambda_۲ = ۷۷/۲۳۶۱ \lambda_۳ = ۲۸/۴۵۹۷, \lambda_۴ = ۱۰/۲۶۷۴, \lambda_۵ = ۰/۰۳۴۹$ و بنابراین عدد شرطی آن برابر $\sqrt{\frac{\lambda_۱}{\lambda_۵}} = ۷۷/۸۲$ است. بزرگ بودن عدد شرطی ناشی از وجود هم خطی زیاد بین متغبرهای رگرسیونی است. بازه ریجی را به صورت $[۰, ۱]$ تعریف کرده و برآوردگر ریج ماکسیمم آنتروپی تی سالیس مرتبه دو تعمیم یافته $(\hat{\eta})$ را به دست می آوریم. برای محاسبات از نرم افزار مطلب کمک می گیریم. مقادیر MSE مربوط به هر سه نوع برآوردگر OLS ، $Ridge - GME$ و $Ridge - GMET۲$

جدول ۲: برآوردهای OLS، $Ridge-GME$، $Ridge-GMET۲$

پارامتر	OLS	$Ridge-GME$	$Ridge-GMET۲$
β_0	۶۲/۴۰۵	۳۱/۲۳۳۹	۳۴/۲۷۵۷
β_1	۱/۵۵۱	۱/۸۶۸۵	۱/۸۴۰۵
β_2	۰/۵۱۰	۰/۸۳۲۴	۰/۸۰۰۱
β_3	۰/۱۰۲	۰/۴۲۷۱	۰/۳۹۷۹
β_4	−۰/۱۴۴	۰/۱۷۱۴	۰/۱۴۰۱

در جدول ۳ خلاصه کرده ایم.

جدول ۳: مقادیرMSE برآوردهای مختلف ریجی .

OLS	Ridge-GME	Ridge-GMET۲
۵/۹۸	۶/۹۸۰۷	۷/۴۲۶۱

بحث و نتیجه‌گیری

علیرغم توانایی بالای روش ماکسیمم آنتروپی، روش ماکسیمم آنتروپی تعمیم یافته با چارچوب رگرسیونی مطرح شد. وجود هم خطی بالا بین متغیرهای توضیحی یکی از مشکلات عمده‌ای است که در تحلیل رگرسیونی با آن مواجه می شویم. یکی از روش هایی که برای مقابله با این مشکل مطرح گردیده‌است روش رگرسیون ریج است. محققین برآوردهای زیادی برای پارامتر ریج ارائه داده‌اند. در این مقاله بر اساس برآورد ماکسیمم آنتروپی تی‌سالیس مرتبه دو تعمیم‌یافته برآورد دیگری برای پارامتر ریج یافته و آن را برآوردریج ماکسیمم آنتروپی تی سالیس مرتبه دو تعمیم‌یافته نامیده‌ایم. برای یک مجموعه داده سه نوع برآوردگر ریج ماکسیمم آنتروپی تعمیم یافته، حداقل مربعات و برآوردگر ریج ماکسیمم آنتروپی تی سالیس مرتبه دو تعمیم یافته را محاسبه و مقادیر MSE شان را براورد کرده ایم اگرچه مقدار MSE روش OLS نسبت به دو روش $Ridge-GME$ و $Ridge-GMET۲$ کمتر به دست آمده ولی این تفاوت ناچیز و با توجه به اینکه مشکل هم خطی در داده ها وجود دارد قابل چشم پوشی بوده و این دو برآوردگر جدید را ترجیح می دهیم.

مراجع

Akdeniz, F., Cabuk, A. and Guler, H. (2011), *Generalized maximum entropy estimators: applications to the Portland Cement data set*. The Open Statistics and Probability Journal, 3, 13-20.

Judge, G. G. and Golan, A. (1992), *Recovering information in the case of ill-posed inverse problems with* *noise*, Mimo Department Of Agricultural And Natural Resources, University of California, Berekeley, CA.

Golan, A., Judge, G. and Miller, D. (1996), *Maximum Entropy Econometrics: Robust estimation with limited data*, John wiley Sons, New york.

Hald, A. (1952), *Statistical Theory with Engineering Applications*, John Wiley Sons, New York.

Horel, A. E. and Kenard. R. W. (1970), Ridge Regression: Applications to Nonorthogonal Problems, *Technometrics*, 12, 69-82.

Jaynes, E. T. (1957), Information theory and statistical mechanics. *Physical Reviews*, 106, 620-630.

Kacranlar, S., Sakalliogglu, S., Akdeniz, F., Styan, G. P. H. and Werner, H. J. (1999), *A new biased estimator in linear regression and a detailed analysis of the widely-analysed dataset on Portland cement*, The Indian Journal of Statistics, Series B, Senkhiya, 61, 443–459.

Macedo, P., Scotto, M. and Silva, E. (2010), On the Choice of the Ridge Parametr: A Maximum Entropyy Approach, *Communication in Statistics Simulations and Computations*, 39, 1628-1638.

Muniz, G. and Kibria, B. M. G. (2009), *On some ridge regression estimator: an emperical comparisons*, Communication in Statistics Simulations and Computations, 38, 621-630.

Liu, K. (2003), *Using Liu-type estimator to combat collinearity*, Communications in Statistics– Theory and Methods, 32, 1009–1020.

Pukelsheim, F. (1994), The three-sigma Rule, *The American statistician*, 48, 4, 88-91.

Shannon, C. E. (1948), A Mathimatical Theory Of Communication. *Bell System Tech.*, 27, 379-423.

Sanei Tabass, M. and Mohtashami Borzadaran, GH. R. (2015), The Generalized Maximum Tsallis Entropy Estimators and Applications to the Portland Cement Data set, *Communications in Statistics - Simulation and Computation* , DOI: 10.1080/03610918.2015.1082589.

Tsallis, C. (1988). Possible generalizations of Boltzmann-Gibbs statistics. *Journal of Statistical Physics*, 52, 479-487.

Renyi, A. (1961). On measures of entropy and information. *Proc. Berekeley Symposium, Statist. Probability*, 1, 547-561.

Woods, H. Steinor, H. and Starke, H. (1932). *Effect of composition of Portland cement on heat evolved during hardening*. Industerial and Engineering Chemistry. 24, 1207-1214.