سیزدهمین کنفرانس
آمار ایران
۲-۴ شهریور ۱۳۹۵

کارایی برآوردگر لیو مبتنی بر روش تفاضلی در مدل رگرسیون نیمه پارامتری با خطاهای همبسته

مهدی روزبه[۱]، پرویز ملک زاده [۱]، حوریه پورگنجی [۲]

[۱]استادیار دانشگاه سمنان

[۲]دانشجوی کارشناسی ارشد دانشگاه سمنان

چکیده: در تجزیه و تحلیل مسائل رگرسیونی و به ویژه مدل بندی آماری در داده های اقتصادی، روانشناسی، علوم اجتماعی، حیاتی، مهندسی و ... با مشکل همخطی در میان متغیرهای پیشگو و اتوهمبستگی درخطاها مواجه هستیم. در چنین مواقعی برآوردگر کمترین مربعات معمولی منجر به برآوردگرهای نادقیق از لحاظ مقدار و علامت شده و فواصل اطمینان پهنی را برای پارامترها نتیجه می دهد. هدف از این مقاله، بررسی مشکلات همخطی چندگانه و اتوهمبستگی به صورت همزمان، غلبه بر آن ها با معرفی برآوردگرهای لیو و مقایسه آن با برآوردگر کمترین مربعات تعمیم یافته معمولی تحت محدودیت های خطی غیر تصادفی $R\beta = r$ ، با بکار بردن مخاطره مبتنی بر تابع زیان متعادل وزنی در مدل رگرسیونی نیمه پارامتری است. در ادامه شرایط لازم و کافی برای برتری برآوردگر نوع لیو تعمیم یافته بر برآوردگر کمترین مربعات تعمیم یافته معمولی را در غالب چند قضیه استخراج نموده و نواحی بهینه برای پارامتر لیو را بدست می آوریم. در انتها، به کمک شبیه سازی مونت کارلو به بررسی درستی قضایای بیان شده پرداخته و برتری برآوردگر پیشنهادی بر برآوردگرهای معمولی معرفی شده را در مدل رگرسیون خطی با خطاهای همبسته نشان خواهیم داد.

واژه‌های کلیدی: تابع زیان متعادل، برآوردگر تفاضلی، برآوردگر تعمیم یافته لیو، برآوردگر تفاضلی محدود شده تعمیم یافته لیو، مدل رگرسیون نیمه پارامتری

طبقه بندی موضوع AMS : ۶۲G۰۸، ۶۲J۰۷.

---

[۲]نام ارائه دهنده مقاله : حوریه پورگنجی h.pourgang170@gmail.com

۱ مقدمه

همخطی چندگانه، خطی یا وابستگی خطی دربین متغیرهای توضیحی در مدل رگرسیون با یک مشکل مهم درکاربرد مواجه است. حضور همخطی چندگانه دارای اثرات جدی روی برآوردگر کمترین مربعات معمولی (OLS) دارد. به طورکلی مقدار مطلق برآوردگر و واریانس ضرایب رگرسیون ممکن است تحت تأثیر قرار گیرند. مدل های رگرسیون نیمه پارامتری در آمار و اقتصاد مورد توجه قرار گرفته است. مدل رگرسیون نیمه پارامتری زیر را در نظر بگیرید

$$y_i = X_i'\beta + f(u_i) + \varepsilon_i \qquad i = ۱, ۲, ..., n, \tag{۱.۱}$$

بطوری که $\boldsymbol{X}_i' = (x_{i۱}, x_{i۲}, ..., x_{ip})$ یک بردار از متغیرهای مستقل می باشد، $\boldsymbol{\beta} = (\beta_۱, \beta_۲, ..., \beta_p)'$ یک بردار پارامتری p-بعدی مجهول می باشد، u_i ها معلوم و غیر تصادفی هستند ، $f(.)$ یک تابع همواره نامعلوم می باشد، و ε_i ها خطاهای تصادفی و مستقل و هم توزیع با میانگین صفر و واریانس $\sigma^۲$ هستند و از $(\boldsymbol{X}_i, u_i)$ مستقل هستند.

$f(u)$ را بخش ناپارامتری مدل نامیده و فرض می‌کنیم که یک تابع هموار است . u_iها دارای تکیه گاه کراندار بوده و فرض می‌کنیم که $u_۱ \leq u_۲ \leq ... \leq u_n$. هدف برآورد بردار پارامتری مجهول β و تابع ناپارامتری $f(u)$ با استفاده از داده های $\{y_i, \boldsymbol{X}_i, u_i\}$ می باشد که اینکار از طریق یک برآوردگر تفاضلی انجام خواهد شد. نماد ماتریسی/برداری، معادله (۱.۱) به فرم زیر نوشته می شود

$$y = \boldsymbol{X}\beta + f + \varepsilon, \tag{۲.۱}$$

که $y = (y_۱, ..., y_n)'$ ، $\boldsymbol{X} = [\boldsymbol{X}_۱, ..., \boldsymbol{X}_n]'$ ، $f = (f(u_۱), ..., f(u_n))'$ ، $\varepsilon = (\varepsilon_۱, ..., \varepsilon_n)'$.

مدلهای رگرسیون خطی نیمه پارامتری انعطاف پذیرتر از مدلهای رگرسیون خطی ساده می باشند، زیرا آنها ترکیبی از مولفه های پارامتری و نیمه پارامتری میباشند و فرض بر این است که در مدل خطی متغیر پاسخ y به متغیر $\boldsymbol{X}$ وابسته است، اما با متغیر مستقل $\boldsymbol{u}$ رابطه غیرخطی دارد. برای برآورد β و f روشهای متعددی وجود دارد. مهمترین روشها توسط چندین پژوهشگر ارائه شده است (گرین سهولت و جنیسون و (۱۹۸۵))

در مدل (۲.۱)، یاتچیو (۱۹۹۷) بر برآورد مولفه خطی و استفاده از روش تفاضلی برای از بین بردن روند ناشی از وجود مولفه ای ناپارامتری تمرکز نموده است. در این روش نیازی به برآورد تابع f نیست و بطور مستقیم می توان پارامتر خطی را برآورد نمود.

در آنالیز رگرسیون، محققان اغلب با مشکل همخطی مواجه می شوند. با وجود همخطی برآوردگر حداقل مربعات ضعیف است. همخطی وابستگی خطی در میان بردارهای ستون ماتریس $\boldsymbol{X}$ در مدل خطی $y = \boldsymbol{X}\beta + \varepsilon$ تعریف می شود. وجود همخطی ممکن است به فاصله اطمینان نامحدود برای پارامترهای خطی یا ترکیب خطی از آنها منجر شود و ممکن است برآوردهایی با علایم اشتباه ارائه دهد. عدد شرطی می تواند معیاری برای همخطی باشد. اگر $\boldsymbol{X}'\boldsymbol{X}$ دارای عدد شرطی بزرگ باشد، دچار بد شرطی شده و برای برآورد β باید از برآوردگر رگرسیونی ریج، برآوردگر لیو (۱۹۹۳) استفاده کرد.

به منظور حل مشکلات ناشی از همخطی اعمال برآوردگرهای انقباضی خوب به عنوان یک اقدام اصلاحی کار آمد شناخته شده است. برای اهداف این مقاله، برای از بین بردن همخطی برآوردهای انقباضی که توسط لیو (۱۹۹۳) پیشنهاد داده شده بود را به کار خواهیم گرفت. ترکیب برآوردگر لیو (۱۹۹۳) با برآوردگر رگرسیونی ریج برآوردگر لیو گفته می شود (کاچیرنلر و آکدنیز (۱۹۹۵)). فرض می کنیم که عدد شرطی مؤلفه پارامتری بزرگ است که نشان می دهد اریبی حاصل از روشهای کلاسیک بزرگ است.

در این مقاله، برآوردگر تفاضلی محدود شده تعمیم یافته برای برآورد بردار پارامتری β تحت محدودیت غیرتصادفی خطی $R\beta = r$ در

مدل رگرسیونی نیمه پارامتری وقتی خطاها همبسته اند، تعریف و استفاده می شود. برآوردگر تفاضلی تعمیم یافته لیو برای بردار پارامتری β در مدل رگرسیونی نیمه پارامتری نیز تعریف می شود. عملکرد برآوردگرهای تحت مطالعه را تحت تابع زیان (BLF) بررسی می کنیم. به منظور مقایسه عملکرد برآوردگرهای پیشنهادی، یک شبیه سازی انجام می دهیم.

۲ مدل و برآوردگرهای تفاضلی

در مدل رگرسیون نیمه پارامتری، یاتچیو (۱۹۹۷) پارامتر خطی را به روش تفاضلی مرتبه m ام به شکل زیر برآورد نمود:

بردار $d = (d_\circ, d_۱, ..., d_m)'$ را با شرایط زیر در نظر بگیرید:

$$\sum_{j=\circ}^{m} d_j = \circ \qquad \& \qquad \sum j = \circ^m d_j^۲ = ۱ \tag{۱.۲}$$

بر اساس بردار فوق ماتریس تفاضلی $n \times (n-m)$، D را بصورت زیر تعریف می کنیم:

$$D = \begin{pmatrix} d_\circ & d_۱ & d_۲ & \cdots & d_m & \circ & \boldsymbol{o} & \cdots & & \boldsymbol{o} \\ \boldsymbol{o} & d_{\boldsymbol{o}} & d_۱ & d_۲ & \cdots & d_m & \boldsymbol{o} & \boldsymbol{o} & \cdots & \boldsymbol{o} \\ \vdots & \vdots & \vdots & & & & & & & \\ \boldsymbol{o} & \boldsymbol{o} & \boldsymbol{o} & \cdots & d_{\boldsymbol{o}} & d_۱ & \cdot & \cdot & d_m & \boldsymbol{o} \\ \boldsymbol{o} & \boldsymbol{o} & \boldsymbol{o} & \cdots & & d_{\boldsymbol{o}} & d_۱ & \cdot & \cdot & d_m \end{pmatrix}$$

به کار بردن ماتریس تفاضلی در مدل ۲.۱ اجازه برآورد مستقیم اثر پارامتر را به محقق می دهد. از آنجایی که داده ها رتبه بندی شده است، استفاده از ماتریس تفاضلی D در مدل ۲.۱ اثر ناپارامتری را در نمونه های بزرگ حذف می کند. اگر f یک تابع نامعلوم باشد که دارای مشتق اول کراندار باشد، آنگاه Df به صفر نزدیک شده و بنابراین با به کار بردن ماتریس تفاضلی می توان نوشت:

$$Dy = D\boldsymbol{X}\beta + Df + D\epsilon, \tag{۲.۲}$$

یا

$$\tilde{y} \approx \tilde{\boldsymbol{X}}\beta + \tilde{\varepsilon} \tag{۳.۲}$$

و $\tilde{y} = Dy$ ، $\tilde{\boldsymbol{X}} = D\boldsymbol{X}$ و $\tilde{\varepsilon} = D\varepsilon$. به طوریکه نقش محدودیت ۱.۲ اکنون مشهود است یاتچیو (۲۰۰۳) یک برآوردگر تفاضلی کمترین مربعات معمولی برای برآورد پارامتر β در مدل رگرسیونی نیمه پارامتری به صورت زیر تعریف می شود:

$$\hat{\beta}_{diff} = [(DX)'(DX)]^{-۱}(DX)'(Dy). \tag{۴.۲}$$

برآوردگر تفاضلی تعمیم یافته با فرض $E(\varepsilon\varepsilon') = \sigma^۲ V$، بصورت زیر تعریف می شود:

$$\hat{\beta}_{GD} = (\tilde{X}' V_D^{-۱} \tilde{X})^{-۱} \tilde{X}' V_D^{-۱} \tilde{y}. \tag{۵.۲}$$

و برآوردگر تعمیم یافته $\sigma^۲$، به صورت زیر است:

$$\hat{\sigma}^۲_{GD} = \frac{(\tilde{y}-\tilde{X}\hat{\beta}_{GD})'V_D^{-۱}(\tilde{y}-\tilde{X}\hat{\beta}_{GD})}{tr[D'(I-\tilde{P})D]} = \frac{\tilde{y}'V_D^{-\frac{۱}{۲}}(I-\tilde{p})V_D^{-\frac{۱}{۲}}\tilde{y}}{tr[D'(I-\tilde{p})D]}, \qquad (۶.۲)$$

و $\tilde{P}$ یک ماتریس طرح است و $\tilde{P} = V_D^{-\frac{۱}{۲}}\tilde{\boldsymbol{X}}(\tilde{\boldsymbol{X}}'V_D^{-۱}\tilde{\boldsymbol{X}})^{-۱}\tilde{\boldsymbol{X}}'V_D^{-\frac{۱}{۲}}$ تعریف شده است. از معادله ۵.۲ مشاهده شده است که خواص برآوردگر تفاضلی تعمیم یافته β به ویژگی های ماتریس اطلاعات $G = \tilde{\boldsymbol{X}}'V_D^{-۱}\tilde{\boldsymbol{X}}$ بستگی دارد.

۳ برآوردگر محدود شده تفاضلی تعمیم یافته لیو

در این بخش، درباره روش برآورد اریبی بحث می کنیم وقتی که ماتریس G بد شرطی به نظر میرسد. در نوشته ها، چندین روشهای برآورد اریبی برای مبارزه با مشکل همخطی وجود دارد، مثل برآوردگر رگرسیونی ریج $\hat{\beta}(k) = (\boldsymbol{Z}'\boldsymbol{Z}+kI)^{-۱}\boldsymbol{Z}'y, k > ۰$، برآوردگر لیو $\hat{\beta}_\eta = (\boldsymbol{Z}'\boldsymbol{Z}+I)^{-۱}(\boldsymbol{Z}'\boldsymbol{Z}+\eta I)\hat{\beta}_{OLS}, ۰ < \eta < ۱$. در این مقاله برآوردگر تفاضلی تعمیم یافته لیو را که بصورت زیر تعریف می شود، پیشنهاد می کنیم:

$$\tilde{\beta}_{GD}(\eta) = (\tilde{X}'V_D^{-۱}\tilde{X}+I)^{-۱}(\tilde{X}'V_D^{-۱}\tilde{y}+\eta\tilde{\beta}_{GD}). \qquad (۱.۳)$$

حال، محدودیت غیر تصادفی خطی زیر را در نظر بگیرید:

$$R\beta = r \qquad (۲.۳)$$

بطوریکه ماتریس $q \times p$، R با رتبه $q < p$ داده شده است و بردار معلوم $q \times ۱$، r داده شده است. برآوردگر محدود شده تفاضلی تعمیم یافته به شکل زیر تعریف می شود:

$$\hat{\beta}_{GRD} = \hat{\beta}_{GD} + G^{-۱}R'(RG^{-۱}R')^{-۱}(r - R\hat{\beta}_{GD}). \qquad (۳.۳)$$

$G = \tilde{\boldsymbol{X}}'V_D^{-۱}\tilde{\boldsymbol{X}}$ ماتریس اطلاعات می باشد. برآوردگر محدود شده تفاضلی تعمیم یافته لیو را با اصلاح برآوردگر تفاضلی تعمیم یافته به وسیله مینیمم کردن مجموع مربعات باقی مانده ها با شرط $R\beta = r$ بصورت زیر بدست می آید:.

$$\hat{\beta}_{GRD}(\eta) = \hat{\beta}_{GD}(\eta) + (G+I)^{-۱}R'[R(G+I)^{-۱}R']^{-۱}(r - R\hat{\beta}_{GD}(\eta)). \qquad (۴.۳)$$

قضیه ۱.۳. تحت محدودیت خطی ۲.۳ داریم

$$bias(\hat{\beta}_{GRD}(\eta)) = E(\hat{\beta}_{GRD}(\eta)) - \beta = -(۱-\eta)M\beta. \qquad (۵.۳)$$

بطوریکه $\boldsymbol{C} = \boldsymbol{G} + I$ و $\boldsymbol{M} = \boldsymbol{C}^{-۱} - \boldsymbol{C}^{-۱}\boldsymbol{R}'(\boldsymbol{R}\boldsymbol{C}^{-۱}\boldsymbol{R}')^{-۱}\boldsymbol{R}\boldsymbol{C}^{-۱}$.

۱.۳ تابع مخاطره برای $\hat{\beta}_{GRD}(\eta)$ تحت زیان BLF وزنی

با توجه به نیکویی برازش و دقت برآورد، زلنر (۱۹۹۴) تابع زیان زیر (BLF) را می توان در نظر گرفت:

$$L(\beta^*,\beta) = w(\tilde{\boldsymbol{X}}\beta^* - \tilde{y})'(\tilde{\boldsymbol{X}}\beta^* - \tilde{y}) + (1-w)(\tilde{\boldsymbol{X}}\beta^* - \tilde{\boldsymbol{X}}\beta)'(\tilde{\boldsymbol{X}}\beta^* - \tilde{\boldsymbol{X}}\beta), \qquad (۶.۳)$$

به طوری که w وزن غیر تصادفی است و $0 \leq w \leq 1$.حال تابع زیان BLF وزنی در زیر را برای برآوردگر $\hat{\beta}_{GRD}$ در نظر بگیرید:

$$L(\hat{\beta}_{GRD},\beta) = w(\tilde{\boldsymbol{X}}\hat{\beta}_{GRD} - \tilde{y})'\boldsymbol{V}_D^{-1}(\tilde{\boldsymbol{X}}\hat{\beta}_{GRD} - \tilde{y}) + (1-w)(\tilde{\boldsymbol{X}}\hat{\beta}_{GRD} - \tilde{\boldsymbol{X}}\beta)'\boldsymbol{V}_D^{-1}(\tilde{\boldsymbol{X}}\hat{\beta}_{GRD} - \tilde{\boldsymbol{X}}\beta) \qquad (۷.۳)$$

$\hat{\beta}_{GRD}$ یک برآوردگر محدود شده تفاضلی تعمیم یافته از β می باشد و $\boldsymbol{V}_D^{-1}$ ماتریس وزنی است.

قضیه ۲.۳. تابع مخاطره مبتنی بر زیان وزنی متعادل بصورت زیر محاسبه می شود:

$$\begin{aligned}
R(\hat{\beta}_{GRD},\beta) &= w\sigma^2(n-m) + \sigma^2 tr(M_\circ G) - 2w\sigma^2 tr(M_\circ G) \\
R(\hat{\beta}_{GRD}(\eta),\beta) &= w\sigma^2(n-m) - 2w\sigma^2(p-q) - 2w(1-\eta)\sigma^2 tr\boldsymbol{M} \\
&\quad + \sigma^2 tr[(\boldsymbol{G}+\eta I)\boldsymbol{G}^{-1}(\boldsymbol{G}+\eta I)(\boldsymbol{M}-\boldsymbol{M}^2)] + (1-\eta)^2\{\beta'\boldsymbol{MGM}\beta\} \qquad (۸.۳)
\end{aligned}$$

بطوریکه $\boldsymbol{M}_\circ = \boldsymbol{M}_\circ\boldsymbol{GM}_\circ$

۴ مقایسه برآوردگرهای پیشنهادی

در این بخش، شرطهای کافی و لازم را برای اینکه $\hat{\beta}_{GRD}(\eta)$ بهتر از $\hat{\beta}_{GRD}$ باشد، استخراج می کنیم. تفاضل $\tilde{\Delta} = \boldsymbol{R}(\hat{\beta}_{GRD},\beta) - \boldsymbol{R}(\hat{\beta}_{GRD}(\eta),\beta)$ به فرم زیر داریم :

$$\begin{aligned}
\tilde{\Delta} =& 2\sigma^2(1-w)(1-\eta)tr\boldsymbol{M} - 2\sigma^2(1-\eta)tr\boldsymbol{M}^2 \\
&+ \sigma^2(1-\eta^2)tr(\boldsymbol{G}^{-1}(\boldsymbol{M}-\boldsymbol{M}^2)) - (1-\eta)^2\beta'\boldsymbol{MGM}\beta, \qquad (۱.۴)
\end{aligned}$$

$$\begin{aligned}
\frac{d}{d\eta}\tilde{\Delta} &= -2\sigma^2(1-w)tr\boldsymbol{M} + 2\sigma^2 tr\boldsymbol{M}^2 - 2\sigma^2\eta tr[\boldsymbol{G}^{-1}(\boldsymbol{M}-\boldsymbol{M}^2)] + 2(1-\eta)\beta'\boldsymbol{MGM}\beta = 0 \\
\Rightarrow \eta_{opt} &= \frac{-\sigma^2(1-w)tr\boldsymbol{M} + \sigma^2 tr\boldsymbol{M}^2 + \beta'\boldsymbol{MGM}\beta}{\beta'\boldsymbol{MGM}\beta + \sigma^2 tr[\boldsymbol{G}^{-1}(\boldsymbol{M}-\boldsymbol{M}^2)]} \\
&= \frac{-\sigma^2 tr(\boldsymbol{M}-\boldsymbol{M}^2) + \sigma^2 w tr\boldsymbol{M} + \beta'(\boldsymbol{M}-\boldsymbol{M}^2)\beta}{\beta'(\boldsymbol{M}-\boldsymbol{M}^2)\beta + \sigma^2 tr[\boldsymbol{G}^{-1}(\boldsymbol{M}-\boldsymbol{M}^2)]}, \qquad (۲.۴)
\end{aligned}$$

از طرفی

$$\frac{d^2}{d\eta^2}\tilde{\Delta} = -2\sigma^2 tr[\boldsymbol{G}^{-1}(\boldsymbol{M}-\boldsymbol{M}^2)] - 2\beta'\boldsymbol{MGM}\beta. \qquad (۳.۴)$$

بنابراین، اگر $-\beta'(\boldsymbol{M}-\boldsymbol{M}^2)\beta < \sigma^2 tr[\boldsymbol{G}^{-1}(\boldsymbol{M}-\boldsymbol{M}^2)]$ ، می توان نتیجه گرفت که η_{opt} مقدار $\tilde{\Delta}$ را ماکسیمم می کند.

۵ مطالعه شبیه سازی

در این بخش، کارایی برآوردگرهای پیشنهادی تابع زیان را آزمون می کنیم. آزمایش نمونه گیری ما شامل ترکیب های مختلفی از η و w بصورت زیر است:

$$w = \{۰/۱, ۰/۳, ۰/۵, ۰/۷, ۰/۹\} \text{ و } \eta = \{۰, ۰/۱, ۰/۲, ۰/۳, ۰/۴, ۰/۵, ۰/۶, ۰/۷, ۰/۸, ۰/۹, ۱\}$$

برای رسیدن به درجات مختلف همخطی، متغیرهای توضیحی با استفاده از فرمول زیر برای $n = ۱۰۰۰$ تولید شده اند.

$$x_{ij} = (۱ - \gamma^۲)^{\frac{۱}{۲}} z_{ij} + \gamma z_{ip}, \qquad i = ۱, ۲, ..., n, \qquad j = ۱, ۲, ..., p, \tag{۱.۵}$$

بطوریکه z_{ij}ها عدد های تصادفی نرمال استاندارد مستقل هستند، و رابطه خطی بین هر دو متغیر توضیحی توسط پارامتر $\gamma^۲$ تعیین می شود. سه مجموعه مختلف از همبستگی متناظر با $\gamma = ۰/۸۰, ۰/۹۰, ۰/۹۹$ و در نظر گرفته شده است که بدلیل صرفه جویی در کار تنها نتایج را برای $\gamma = ۰/۹۰$ گزارش می کنیم. مشاهدات وابسته برای $n = ۱۰۰۰$ از فرمول زیر شبیه سازی میشود

$$y_i = \sum_{j=۱}^{۶} x_{ji}\beta_j + f(t_i) + \epsilon_i, \qquad i = ۱, ۲, ...n \tag{۲.۵}$$

و $\beta = (۳, ۱, ۳, ۲, -۵, ۴)$ و $\varepsilon \sim N(۰, \sigma^۲ \boldsymbol{V})$ که عناصر $\boldsymbol{V}$ عبارت اند از

$$v_{ij} = \left(\frac{۱}{n}\right)^{i-j}, \quad \sigma^۲ = ۶$$

و

$$f(t_i) = \frac{۱}{۶} \sum_{j=۱, j\neq ۴,۶} \varphi\left(t_i; j, \left[\frac{j+۲}{۱۰}\right]^j\right),$$

که نرمال آمیخته برای $t_i = ۱۰i/n$ و $\phi(x;\mu,\sigma^۲)$ تابع چگالی نرمال با میانگین μ و واریانس $\sigma^۲$ می باشد.

برای برآورد پارامترها از روش تفاضلی مرتبه ۴ ، $(m = ۴)$استفاده می کنیم. وزنهای تفاضلی عبارتند از

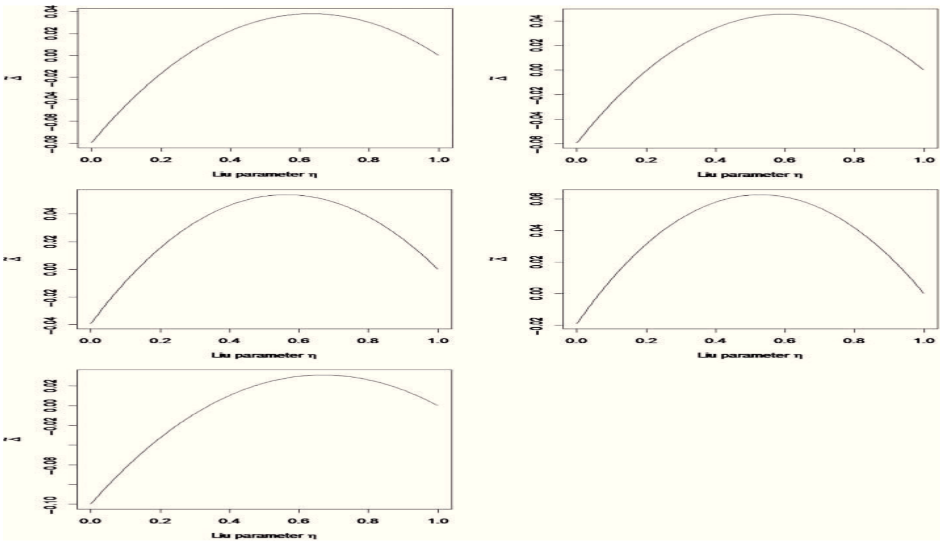
$$d_۰ = ۰/۸۸۷۳, d_۱ = -۰/۳۰۹۹, d_۲ = ۰/۲۴۶۴, d_۳ = -۰/۱۹۰۱, d_۴ = -۰/۱۴۰۹$$

محدودیت خطی را به شکل زیر در نظر می گیریم

$$\boldsymbol{R} = \begin{pmatrix} ۱ & ۵ & -۳ & -۱ & -۱ & ۰ \\ -۲ & -۱ & ۰ & -۲ & ۳ & ۱ \\ ۴ & -۱ & ۲ & ۲ & ۰ & -۲ \\ ۵ & ۳ & ۴ & -۵ & ۱ & ۰ \end{pmatrix}, r = \begin{pmatrix} ۰ \\ ۰ \\ \vdots \\ ۰ \end{pmatrix}$$

نتیجه گیری

در این بخش، برآوردگر تفاضلی تعمیم یافته لیو را پیشنهاد می کنیم، $\hat{\beta}_{GD}$ در مدل رگرسیون نیمه پارامتری خطاها هم بسته هستند و برخی از محدودیت های خطی اضافی بر روی فضای پارامتری β ایجاد شده است. در حضور همخطی در مدل رگرسیون نیمه پارامتری برآوردگر محدود شده تفاضلی تعمیم یافته لیو را پیشنهاد کردیم، $\hat{\beta}_{GRD}(\eta)$.



شکل ۱: نمودار $\tilde{\Delta}$ در برابر η برای مقادیر مختلف w و ۰/۹۰ = γ رسم شده است. نمودار بالا سمت چپ: ۰/۱ = w ، در نمودار بالا سمت چپ: ۰/۳ = w ، نمودار وسط سمت چپ: ۰/۵ = w ، نمودار وسط سمت راست: ۰/۷ = w و در نمودار پایین سمت چپ: ۰/۹ = w.

مراجع

Akdeniz F, Kaçıranlar S. On the almost unbiased generalized Liu estimator and unbiased estimation of the bias and MSE. Commun Statist-Theory Methods.1995;24(7):1789–1797.

Green P, Jennison C, Seheult A. Analysis of field experiments by least squares smoothing. J Roy Statist Soc Ser B.1985;47:299–315. K, Eubank RL. Difference-based variance estimators for partially linear models. Festschrift in honor of Distinguished Professor

Liu KJ. A new class of biased estimate in linear regression. Commun Statist Theory Methods. 1993;22:393–402

Roozbeh M, Arashi M, Niroumand HA. Ridge regression methodology in partial linear models with correlated errors. J Statist Comput Simul. 2011;81(4):517–528.

Yatchew A. An elementary estimator of the partial linear model. Econ Lett.1997;57:135–143.

Yatchew A. Semiparametric regression for the applied econometrican. Cambridge: Cambridge University Press; 2003.

Zellner A. Bayesian and non-Bayesian estimation using balanced loss functions. In: Gupta SS, Berger JO, editors. Statistical decision theory and related topics, Vol. V. NewYork: Springer; 1994. p. 377–390.

جدول ۱: ارزیابی تابع زیانها و پارامترها برای مقادیر مختلف η و w $(n = ۱۰۰۰ \eta = ۰/۹۰)$.

η	۰.۱	۰.۳	۰.۵	۰.۷	۰.۹
$\hat{\beta}_۱$	۳.۰۰۳۶	۲.۹۹۶۷	۲.۹۸۹۹	۲.۹۸۳۱	۲.۹۷۶۴
$\hat{\beta}_۲$	۰.۹۹۸۱	۱.۰۰۱۶	۱.۰۰۵۱	۱.۰۰۸۶	۱.۰۱۲۰
$\hat{\beta}_۳$	۲.۹۹۷۸	۳.۰۰۱۹	۳.۰۰۶۰	۳.۰۱۰۰	۳.۰۱۴۰
$\hat{\beta}_۴$	۲.۰۰۰۷	۱.۹۹۹۲	۱.۹۹۷۸	۱.۹۹۶۳	۱.۹۹۴۸
$\hat{\beta}_۵$	- ۴.۹۹۹۹	- ۵.۰۰۰۰	- ۵.۰۰۰۱	- ۵.۰۰۰۲	- ۵.۰۰۰۳
$\hat{\beta}_۶$	۴.۰۰۶۷	۳.۹۹۳۸	۳.۹۸۱۰	۳.۹۶۸۳	۳.۹۵۵۷
$R(w = ۰/۱)$	۴۴۳.۹۷۱	۴۴۳.۹۳۳	۴۴۳.۹۱۸	۴۴۳.۹۳۹	۴۴۳.۹۵۷
$R(w = ۰/۳)$	۴۷۳.۵۶۹	۴۷۳.۵۲۶	۴۷۳.۵۰۷	۴۷۳.۵۱۲	۴۷۳.۵۳۹
$R(w = ۰/۵)$	۵۰۳.۱۶۷	۵۰۳.۱۲۰	۵۰۳.۰۹۷	۵۰۳.۰۹۷	۵۰۳.۱۲۰
$R(w = ۰/۷)$	۵۳۲.۷۶۵	۵۳۲.۷۱۴	۵۳۲.۶۸۷	۵۳۲.۶۸۳	۵۳۲.۷۰۲
$R(w = ۰/۹)$	۵۶۲.۳۶۳	۵۶۲.۳۰۸	۵۶۲.۲۷۷	۵۶۲.۲۶۹	۵۶۲.۲۸۴
$\tilde{\Delta}(w = ۰/۱)$	۰.۰۰۹۲	۰.۰۴۷۷	۰.۰۶۲۶	۰.۰۵۴۴	۰.۰۲۳۶
$\tilde{\Delta}(w = ۰/۳)$	- ۰.۰۰۸۸	۰.۰۳۳۸	۰.۰۵۲۷	۰.۰۴۸۵	۰.۰۲۱۷
$\tilde{\Delta}(w = ۰/۵)$	- ۰.۰۲۶۹	۰.۰۱۹۸	۰.۰۴۲۸	۰.۰۴۲۶	۰.۰۱۹۷
$\tilde{\Delta}(w = ۰/۷)$	- ۰.۰۴۵۰	۰.۰۰۵۹	۰.۰۳۲۹	۰.۰۳۶۷	۰.۰۱۷۸
$\tilde{\Delta}(w = ۰/۹)$	- ۰.۰۶۳۱	- ۰.۰۰۸۰	۰.۰۲۳۰	۰.۰۳۰۸	۰.۰۱۵۸