پانزدهمین سمینار احتمال
و فرآیندهای تصادفی
۸ و ۹ شهریور ۱۴۰۴
دانشگاه کردستان

15th Seminar On Probability and Stochastic Processes
August 30-31, 2025
University of Kurdistan

انجمن آمار ایران

خواص مجانبی برآوردگر ترکیبی در مدل‌های خطی با خطا در اندازه‌گیری

چکیده: در این مقاله یک برآوردگر ترکیبی برای بهبود دقت برآوردها در مدل‌های با خطی در اندازه‌گیری خطی تحت محدودیت‌های خطی تصادفی معرفی و ویژگی‌های مجانبی برآوردگر معرفی شده تعیین می‌شود. عملکرد برآوردگر معرفی شده با استفاده از معیار ماتریس میانگین مربعات خطا نسبت به بعضی برآوردگرهای موجود مورد بررسی قرار می‌گیرد. در پایان عملکرد برآوردگر معرفی شده با استفاده از یک مطالعه‌ی شبیه سازی ارزیابی می‌شود.

واژه‌های کلیدی:برآورردگر ترکیبی، مدل خطی با خطا در اندازه‌گیری،محدودیت‌های خطی تصادفی، هم‌خطی .
کد موضوع‌بندی ریاضی (۲۰۲۰): 62J05، 62J07.

۱ مقدمه

وجود وابستگی خطی بین ستون های ماتریس متغیرهای توضیحی باعث کاهش دقت برآورد و حساسیت برآورد نسبت به تغییر کوچکی در مجموعه داده‌ها می‌شود. یک روش برخورد با مسئله هم‌خطی به دست آوردن برآورد پارامترهای مدل با اضافه کردن اطلاعات پیشین به صورت محدودیت‌های خطی دقیق یا تصادفی روی پارامترهای نامعلوم مدل است ؟ . در این مقاله یک برآوردگر ترکیبی با استفاده از روش کبریا ـلکمن و در مدل‌های خطی با خطا در اندازه‌گیری تحت محدودیت های خطی تصادفی معرفی می‌شود. هدف اصلی برآورد ترکیبی، ترکیب برآوردگرهای مختلف برای کاهش انحراف و خطا در برآورد ضرایب مدل است. یک مدل با خطا در اندازه گیری را به‌صورت

$$y = Z\beta + \varepsilon, \varepsilon \sim N(0, \sigma^2 I_n)$$
$$X = Z + U, U \sim MN(0, I_n \otimes \Sigma)$$

در این مدل $y = (y_1, y_2, ..., y_n)'$ بردار $n \times 1$ از مشاهدات،Z ماتریس $n \times p$ از متغیرهای تو ضیحی ،β بردار $p \times 1$ از پارامترهای ناشناخته مدل و ε بردار $n \times 1$ از خطای تصادفی مدل است.X ماتریس مشاهده شده Z با خطای اندازه گیری U است. فرض کنید ε و U از هم مستقل و U یک ماتریس با عناصر معلوم است.

فرض کنید بردار پارامتر β تحت محدودیت‌های خطی تصادفی به‌صورت زیر باشد:

$$r = R\beta + e, e \sim N(\circ, \sigma^{۲}W)$$

در این رابطه r یک بردار معلوم $m \times ۱$ از مشاهدات، R یک ماتریس معلوم $m \times p$ پرتبه سطری $m < p$ و e بردار $m \times ۱$ از خطای تصادفی است.برآوردگر آمیخته در مدل‌های با خطا در اندازه‌گیری توسط ؟ به‌صورت زیر معرفی شد:

$$\hat{\beta}_{ME} = S_r^{-۱}(X'y + R'W^{-۱}r),$$

در این رابطه $S_r = S + R'W^{-۱}R$.

۲ برآوردگر ترکیبی

برای تعیین برآوردگر جدید در مدل‌های با خطا در اندازه‌گیری براساس ؟رابطه‌ی زیر را در نظر می‌گیریم:

$$\Phi_{۱} = (y - X\beta)'(y - X\beta) - n\beta'\Sigma\beta + k\left[(\beta + \hat{\beta})'(\beta + \hat{\beta}) - c\right]$$

در این رابطه c ضریب ثابت و k ضریب لاکرانژ در نظر گرفته می‌شود. با مشتق گرفتن از $\Phi_{۱}$ برآوردگر $\hat{\beta}_{KL}$ به‌صورت زیر تعیین می‌شود:

$$\hat{\beta}_{KL} = S_k^{-۱}(X'y - k\hat{\beta})$$

به‌منظور در نظر گرفتن محدودیت‌های تصادفی در برآوردگر $\hat{\beta}_{KL}$ رابطه‌ی زیر را در نظر می‌گیریم:

$$\Phi_{۲} = (y - X\beta)'(y - X\beta) - n\beta'\Sigma\beta + k\left[(\beta + \hat{\beta}_{ME})'(\beta + \hat{\beta}_{ME}) - c\right] + (r - R\beta)'W^{-۱}(r - R\beta)$$

با مشتق گرفتن از $\Phi_{۲}$ نسبت به β برآوردگر ترکیبی به‌صورت زیر معرفی می‌شود:

$$\hat{\beta}_{MKL} = (S + R'W^{-۱}R + kI_p)^{-۱}(X'y - k\hat{\beta}_{ME} + R'W^{-۱}r)$$

۳ خواص مجانبی

در این بخش خواص مجانبی برآوردگر $\hat{\beta}_{MKL}$ مورد بررسی قرار می‌گیرد. فرض می‌کنیم وقتی n به بینهایت میل می‌کند حدهای $n^{-۱}(Z'Z + R'W^{-۱}R)$، $n^{-۱}(Z'Z + R'W^{-۱}R + kI_p)$و $n^{-۱}(Z'Z + R'W^{-۱}R - kI_p)$ وجود دارند.

قضیه ۱.۳. تحت فرض‌های بیان شده $\hat{\beta}_{MKL}$ به طور مجانبی دارای توزیع نرمال با میانگین و کواریانس ماتریس به‌صورت زیر است

$$E(\hat{\beta}_{MKL}) = G_{rk}\beta$$

$$AVar(\hat{\beta}_{MKL}) = A_r^{-۱}(G_{rk}BG_{rk} + \sigma^{۲}G_{rk}A_rG_{rk})A_r^{-۱}$$

که در آن $B = (n\sigma^2 + \beta' Z'Z\beta)\Sigma$، $G_{rk} = (Z'Z + R'W^{-1}R + kI_p)^{-1}(Z'Z + R'W^{-1}R - kI_p)$ و $A_r = Z'Z + R'W^{-1}R$.

اثبات. از آنجا که $E(X'X) = Z'Z + n\Sigma$ ؟،بنابراین داریم:

$$n^{-1}(X'X + R'W^{-1}R) = n^{-1}(Z'Z + R'W^{-1}R) + \Sigma + O_p(n^{-\frac{1}{2}})$$

همچنین داریم

$$\begin{aligned}\sqrt{n}\hat{\beta}_{MKL} &= \left[n^{-1}(Z'Z + R'W^{-1}R) + O_p(n^{-\frac{1}{2}})\right]^{-1} n^{-\frac{1}{2}}\left[G_{rk}(X'y + R'W^{-1}r) + O_p(n^{\frac{1}{2}})\right] \\ &= \left[I_p + O_p(n^{-\frac{1}{2}})\right]^{-1} C^{-1} n^{-\frac{1}{2}}\left[G_{rk}(X'y + R'W^{-1}r) + O_p(n^{\frac{1}{2}})\right]\end{aligned}$$

زمانی که حد $C = n^{-1}(Z'Z + R'W^{-1}R)$ وجود دارد رابطه‌ی زیر نتیجه می‌شود:

$$\sqrt{n}\hat{\beta}_{MKL} = C^{-1}\xi + O_p(n^{-\frac{1}{2}}),$$

که $\xi = n^{-\frac{1}{2}}\left[G_{rk}(X'y + R'W^{-1}r)\right]$ به‌طور مجانبی نرمال است (؟).

از آنجا که $E\left[G_{rk}(X'y + R'W^{-1}r)\right] = G_{rk}A_r\beta$، نتیجه گرفته می‌شود که $E(\xi) = n^{-\frac{1}{2}}G_{rk}A_r\beta$. همچنین

$$\sqrt{n}(\hat{\beta}_{MKL} - G_{rk}\beta) = C^{-1}\left[\xi - E(\xi)\right] + O_p(n^{-\frac{1}{2}})$$

بنابراین $\sqrt{n}(\hat{\beta}_{MKL} - G_{rk}\beta)$ دارای توزیع مجانبی نرمال با میانگین صفر است. همچنین داریم. $AVar(\sqrt{n}\hat{\beta}_{MKL}) = C^{-1}Var(\xi)C^{-1}$ برای محاسبه واریانسξ روش زیر را بکار می‌بریم:

$$\begin{aligned}Var(\xi) &= E_{y_r}\left[Var(\xi\,|y_r)\right] + Var_{y_r}\left[E(\xi\,|y_r)\right] \\ &= n^{-1}E_{y_r}(G_{rk}y'y\Sigma G_{rk}) + n^{-1}Var_{y_r}\left[G_{rk}(Z'y + R'W^{-1}r)\right],\end{aligned}$$

در این رابطه E_{y_r} و Var_{y_r} امید و واریانس نسبت به بردار $y_r' = (y', r')$ می باشند. از آنجا که $E(y'y) = n\sigma^2 + \beta' Z'Z\beta$ و $Var_{y_r}\left[G_{rk}(Z'y + R'W^{-1}r)\right] = \sigma^2 G_{rk}A_rG_{rk}$ می‌توان نوشت, $Var(\xi) = n^{-1}(G_{rk}BG_{rk} + \sigma^2 G_{rk}A_rG_{rk})$ بنابراین خواهیم نوشت:

$$AVar(\hat{\beta}_{MKL}) = A_r^{-1}(G_{rk}BG_{rk} + \sigma^2 G_{rk}A_rG_{rk})A_r^{-1}.$$

□

۴ ماتریس میانگین مربعات خطا

برای بررسی کارائی برآوردگر معرفی شده نسبت به برآوردگرهای موجود دیگر از معیار ماتریس میانگین مربعات خطا استفاده می‌شود. ماتریس میانگین مربعات خطای خطای برآوردگر $\hat{\beta}$ از پارامتر β به‌صورت زیر تعریف می‌شود:

$$MSEM(\hat{\beta}) = E(\hat{\beta}-\beta)(\hat{\beta}-\beta)' = Var(\hat{\beta}) + Bias(\hat{\beta})Bias(\hat{\beta})'$$

که $Bias(\hat{\beta}) = E(\hat{\beta}) - \beta$ مقدار اریبی برآوردگر $\hat{\beta}$ است. مقدار میانگین مربعات خطا معیار دیگری برای بررسی کارائی یک برآوردگر است که به صورت زیر تعریف می‌شود:

$$MSE(\hat{\beta}) = tr\left[MSEM(\hat{\beta})\right] = tr\left[Var(\hat{\beta})\right] + Bias(\hat{\beta})'Bias(\hat{\beta})$$

ماتریس میانگین مربعات برآوردگر های ذکر شده به طور مجانبی به صورت زیر نوشته می‌شود:

$$AMSEM(\hat{\beta}) = A^{-1}(B + \sigma^2 A)A^{-1}$$

$$AMSEM(\hat{\beta}_{KL}) = A^{-1}(G_k B G_k + \sigma^2 G_k A G_k)A^{-1} + b_1 b_1'$$

$$AMSEM(\hat{\beta}_{ME}) = A_r^{-1}(B + \sigma^2 A_r)A_r^{-1}$$

$$AMSEM(\hat{\beta}_{MKL}) = A_r^{-1}(G_{rk} B G_{rk} + \sigma^2 G_{rk} A_r G_{rk})A_r^{-1} + b_2 b_2'$$

که $b_1 = -2kA_k^{-1}\beta$ ، $A_k = Z'Z + kI_p$ و $b_2 = -2kA_{rk}^{-1}\beta$.

۱.۴ مقایسه ماتریس میانگین مربعات خطای برآوردگرهای $\hat{\beta}_{KL}$ و $\hat{\beta}$

به‌منظور مقایسه کارائی دو برآوردگر $\hat{\beta}_{KL}$ و $\hat{\beta}$ ماتریس زیر را در نظر می‌گیریم:

$$\Delta_1 = AMSEM(\hat{\beta}) - AMSEM(\hat{\beta}_{KL}) = D_1 - b_1 b_1'$$
$$= 2kA_k^{-1}(A^{-1}B + BA^{-1} + 2\sigma^2 I_p)A_k^{-1} - b_1 b_1',$$

از آنجا که $D_1 = 2kA_k^{-1}(A^{-1}B + BA^{-1} + 2\sigma^2 I_p)A_k^{-1}$ یک ماتریس معین مثبت است . بنابراین Δ_1 معین مثبت است هرگاه $b_1' D_1^{-1} b_1 \leq 1$.

۲.۴ مقایسه ماتریس میانگین مربعات خطا برآوردگرهای $\hat{\beta}_{MKL}$ و $\hat{\beta}_{ME}$

به منظور مقایسه ماتریس میانگین مربعات خطای دو برآوردگر $\hat{\beta}_{MKL}$ و $\hat{\beta}_{ME}$ماتریس زیر را در نظر می گیریم:

$$\Delta_2 = AMSEM(\hat{\beta}_{ME}) - AMSEM(\hat{\beta}_{MKL}) = D_2 - b_2 b_2'$$
$$= 2kA_{rk}^{-1}(A_r^{-1}B + BA_r^{-1} + 2\sigma^2 I_p)A_{rk}^{-1} - b_2 b_2',$$

از آنجا که $D_2 = 2kA_{rk}^{-1}(A_r^{-1}B + BA_r^{-1} + 2\sigma^2 I_p)A_{rk}^{-1}$ یک ماتریس معین مثبت است بنابراین Δ_2معین مثبت است هرگاه $b_2' D_2^{-1} b_2 \leq 1$.

۳.۴ مقایسه ماتریس میانگین مربعات برآوردگرهای$\hat{\beta}_{MKL}$ و $\hat{\beta}_{KL}$

به منظور بررسی دو برآوردگر در معیار ماتریس میانگین مربعات خطا ماتریس زیر در نظر گرفته می‌شود:

$$\Delta_۳ = AMSEM(\hat{\beta}_{KL}) - AMSEM(\hat{\beta}_{MKL}) = D_۳ + b_۱b_۱' - b_۲b_۲',$$

در این رابطه $D_۳ = AVar(\hat{\beta}_{KL}) - AVar(\hat{\beta}_{MKL})$. واضح است که $AVar(\hat{\beta}_{KL}) > ۰$ و $AVar(\hat{\beta}_{MKL}) > ۰$. بنابراین هرگاه داشته باشیم $\lambda_{\max}\left[AVar(\hat{\beta}_{MKL})AVar(\hat{\beta}_{KL})^{-۱}\right] < ۱$ آنگاه $D_۳ > ۰$ بنابراین $\Delta_۳ \geq ۰$ اگر وفقط اگر داشته باشیم $b_۳'(D_۳ + b_۱b_۱')^{-۱}b_۳ \leq ۱$.

۵ برآورد پارامتر k

در این بخش پارامتر اریب k را طوری براورد می کنیم که ماتریس $\Delta_۱$ یک ماتریس معین مثبت شود. ماتریس $\Delta_۱$را می توان به صورت زیر نوشت:

$$\Delta_۱ = ۲kA_k^{-۱}(A^{-۱}B + BA^{-۱} + ۲\sigma^۲I_p - ۲k\beta'\beta)A_k^{-۱},$$

از آنجا که$A_k^{-۱} > ۰$، $A^{-۱}B > ۰$و $BA^{-۱} > ۰$. بنابراین $\Delta_۱$ یک ماتریس معین مثبت است هرگاه $۲\sigma^۲I_p - ۲k\beta'\beta$ نیمه معین مثبت باشد. بنابراین یک برآورد مناسب برای پارامتر اریب k به صورت $\hat{k} < \frac{\hat{\sigma}^۲}{\hat{\beta}'\hat{\beta}}$ تعیین می‌شود.

۶ مطالعه‌ی شبیه‌سازی

به‌منظور ارزیابی نتایج تئوری یک مطالعه‌ی شبیه‌سازی انجام شده است. برای در نظر گرفتن درجات مختلف هم‌خطی با توجه به مقاله ؟ متغیرهای توضیحی به‌صورت زیر تولید می‌شوند:

$$z_{il} = (۱ - \rho^۲)^{\frac{۱}{۲}}w_{il} + \rho w_{i,p+۱}, i = ۱, ..., n, l = ۱, ..., p,$$

که در آن w_{ij} متغیرهای شبه تصادفی مستقل نرمال استاندارد و $\rho^۲$ همبستگی بین دو متغیر توضیحی را نشان می‌دهد. داده‌ها همجنین استاندارد می‌شوند تا $Z'Z$ به فرم ماتریس همبستگی شود. -j امین مجموعه از داده‌های شبیه‌سازی از مدل با خطا در اندازه‌گیری تحت محدودیت‌های خطی تصادفی به‌صورت زیر تولید می‌شود:

$$\begin{aligned} y_j &= Z\beta + \varepsilon_j, \\ X_j &= Z + U_j, j = ۱, ..., ۱۰۰۰, \\ r_j &= R\beta + e_j, \end{aligned}$$

در این رابطه $Z = (z^{(۱)}, z^{(۱)}, z^{(۳)})$، $y_j = (y_{۱j}, ..., y_{nj})'$ و $l = ۱, ۲, ۳$،$z^{(l)} = (z_{۱l}, ..., z_{nl})'$ و $\varepsilon_j \sim N(۰, \sigma^۲I_n)$ تولید می شود. همجنین R یک ماتریس معلوم و عناصر ان از توزیع نرمال استاندارد تولید شده اند و $r_j = (r_{۱j}, ..., r_{mj})'$ و $e_j \sim N(۰, \sigma^۲I_m)$. در این مطالعه$p = ۳$، $\sigma^۲ = ۰/۵$، $m = ۱$ و$\Sigma = diag(۰/۰۵, ۰/۰۵, ۰/۰۵)$

در نظر گرفته شده است. برای انجام شبیه‌سازی از نرم افزار R استفاده شده است . برا ی هر ترکیب از پارامترها ۱۰۰۰ تکرار انجام می شود. عملکرد برآوردگرها با استفاده از مقدار میانگین مربعات خطا از رابطه ی زیر ارزیابی می‌شود:

$$\mathrm{MSE}(\tilde{\beta}) = \frac{۱}{۱۰۰۰} \sum_{j=۱}^{۱۰۰۰} \sum_{l=۱}^{۳} (\tilde{\beta}_{lj} - \beta_l)^۲$$

نتایج شبیه‌سازی در جداول ؟؟ و ؟؟ ارائه شده است.

جدول ۱: برآورد مقدار میانگین مربعات خطا براوردگرهای مختلف با $n = ۵۰$

$\hat{\beta}_{MKL}$	$\hat{\beta}_{ME}$	$\hat{\beta}_{KL}$	$\hat{\beta}$	ρ
۰/۰۵۷۱	/۰۶۱۱.۰	۰/۰۵۷۵	۰/۰۶۱۵	۰/۸۰
۰/۱۱۲۲	۰/۱۲۵۱	۰/۱۱۳۵	۰/۱۲۶۷	۰/۹۰
۰/۲۴۹۱	۰/۲۹۵۱	۰/۲۵۴۴	۰/۳۰۲۳	۰/۹۵

جدول ۲: برآورد مقدار میانگین مربعات خطا براوردگرهای مختلف با $n = ۱۰۰$

$\hat{\beta}_{MKL}$	$\hat{\beta}_{ME}$	$\hat{\beta}_{KL}$	$\hat{\beta}$	ρ
۰/۰۴۵۵	/۰۴۷۷.۰	۰/۰۴۵۷	۰/۰۴۷۹	۰/۸۰
۰/۱۰۵۸	۰/۱۱۴۷	۰/۱۰۶۷	۰/۱۱۵۸	۰/۹۰
۰/۲۳۷۰	۰/۲۶۶۹	۰/۲۴۰۵	۰/۲۷۱۳	۰/۹۵

نتایج شبیه‌سازی ارائه شده در جداول ؟؟ و ؟؟ بیانگر آن است که با افزایش تعداد نمونه مقدار میانگین مربعات خطای برآوردگرها کاهش می‌یابد. با افزایش سطح هم‌خطی مقدار میاگین مربعات خطای همه برآوردگرها افزایش پیدا می‌کند و در تمام موارد مقدار میانگین مربعات خطای برآوردگر $\hat{\beta}_{MKL}$ نسبت به سایر برآوردگرهای ذکر شده کمتر است. بنابراین در حضور هم‌خطی استفاده از برآوردگر جدید پیشنهاد می‌شود.

بحث و نتیجه‌گیری

در این مقاله یک برآوردگر ترکیبی در مدل‌های خطی با خطا در اندازه‌گیری تحت محدودیت‌های خطی تصادفی معرفی گردید. ویژگی‌های مجانبی برآوردگر معرفی شده بررسی و با استفاده از معیار ماتریس میانگین توان‌های دوم خطا کارایی برآوردگرها مورد ارزیابی قرار گرفت. با استفاده از مطالعه‌ی شبیه‌سازی نشان داده شد که برای مقدار مشخصی از پارامتر اریب برآوردگر معرفی شده کارائی بیشتری نسبت به برآوردگرمحدود شده دارد.

مراجع

Fung, W. K., Zhong, X. P. and Wei, B. C. (2003). On estimation and influence diagnostics in linear mixed measurement error models. *American Journal of Mathematical and Management Sciences,* **23**, 37–59.

Ghapani, F., A. R. Rasekh, and B. Babadi. (2015). Detection of outliers and influential observations in linear ridge measurement error models with stochastic linear restrictions. *Journal of Sciences Islamic Republic of Iran*. **26**(4), 355–66.

Kibria, B. M. G. and Lukman, A. F. (2020). A new ridge-type estimator for the linear regression model. *Simulations and applications. Hindawi Scientifica*. https://doi.org/10.1155/2020/9758378.

McDonald, C., Galarneau, D. A. (1975). A Monte Carlo Evaluation of some Ridge-Type Estimators. *Journal of the American Statistical Association*. **70**, 407-416.

Rao, C. R., Toutenburg, H., Shalabh and Heumann, C. (2008), *Linear Models and Generslizations*, Springer. Berlin.

Asymptotic properties of the hybrid estimator in the linear measurement error models

Fatemeh Ghapani

Department of Mathematics and Statistics, Shoushtar Branch, Islamic Azad University, Shoushtar, Iran

---

Abstract: In this paper, a hybrid estimator is introduced for improving the accuracy of estimators in linear measurement error models with stochastic linear restrictions,and the asymptotic properties of the introduced estimator are investigated.The performance of the introduced estimator is examined using the mean square error matrix criterion compared to some existing estimators.Finally, the performance of the introduced estimator is evaluated using a simulation study.

.

Keywords: Hybrid estimator, Linear measurement error model,Stochastic linear restrictions, Multicollinearitys.

Mathematics Subject Classification (2020): 62J05, 62J07.
