BEYKENT ÜNİVERSİTESİ FEN VE MÜHENDİSLİK BİLİMLERİ DERGİSİ

Sayı 7(1) 2014, 13 – 24

LİNEER KARMA MODEL ALTINDA SABİT VE RASGELE ETKİLER İÇİN BLUE VE BLUP DENKLEMLERİ

Melike YİĞİT *(melikeyigitt@gmail.com)*
Sakarya Üniversitesi, Matematik Bölümü, Yüksek Lisans Öğrencisi

Nesrin GÜLER *(nesring@sakarya.edu.tr)*
Sakarya Üniversitesi, İstatistik Bölümü

ÖZET

Çalışmada u ve ε rasgele vektörlerinin ilişkili olduğu kabulü altında singüler kovaryans matrisine sahip $y = X\beta + Zu + \varepsilon$ lineer karma modeli ele alınmıştır. Bu model altında sabit etkilerin en iyi lineer yansız tahmin edicisi ($BLUE$) denklemi ve rasgele etkilerin en iyi lineer yansız ön tahmin edicisi ($BLUP$) denklemi, sabit etkili lineer modeller ve bu sabit etkili lineer modellere yeni gözlemlerin eklemesiyle elde edilen modeller vasıtasıyla elde edilmiştir.

Anahtar Kelimeler: *$BLUE$, $BLUP$, lineer karma model, sabit etkili lineer model.*

BEYKENT UNIVERSITY JOURNAL OF SCIENCE AND ENGINEERING

Volume 7(1) 2014, 13 – 24

EQUATIONS OF BLUE AND BLUP FOR FIXED AND RANDOM EFFECTS UNDER LINEAR MIXED MODEL

Melike YİĞİT *(melikeyigitt@gmail.com)*
Sakarya University, Department of Mathematics, MSc Student

Nesrin GÜLER *(nesring@sakarya.edu.tr)*
Sakarya University, Department of Statistics

ABSTRACT

In this study, a linear mixed model $y = X\beta + Zu + \varepsilon$ with singular covariance matrix is considered under the assumption that the random vectors u and ε are correlated. The equation of the best linear unbiased estimator ($BLUE$) of fixed effects and the equation of the best linear unbiased predictor ($BLUP$) of random effects are obtained under the linear mixed model through the fixed effect models and the models obtained by adding new observations to the fixed effect models.

Keywords: $BLUE$, $BLUP$, *linear mixed model, fixed effect linear model*

1. GİRİŞ

$R^{m\times n}$, $m\times n$ boyutlu reel matrislerin kümesi olmak üzere, $A\in R^{m\times n}$ matrisi için A', A^{-}, A^{+}, $C(A)$, $C(A)^{\perp}$ ve $N(A)$ sembolleri sırasıyla A matrisinin transpozesini, bir genelleştirilmiş tersini, Moore-Penrose tersini, sütun uzayını, sütun uzayının dikini ve sıfır uzayını göstermektedir. A ve B alt matrisler olmak üzere $(A:B)$ parçalanmış matris gösterimi ve $A^{\perp}$ ise, $C(A)^{\perp}=C(A^{\perp})=N(A')$ koşulunu sağlayan herhangi bir matris gösterimidir. Ayrıca $P=AA^{+}=A(A'A)^{-}A'$ ve $M=I_m-P$ sırasıyla $C(A)$ ve $C(A)^{\perp}$ üzerindeki izdüşüm matrisleridir.

$E(u)=0$, $\text{cov}(u)=D$, $E(\varepsilon)=0$ ve $\text{cov}(\varepsilon)=V$ olmak üzere, u ve ε vektörlerinin ilişkili yani $\text{cov}(\varepsilon,u)=U$ olduğu kabulü altında

$$y=X\beta+Zu+\varepsilon \tag{1}$$

ya da diğer bir gösterimi

$$K=\{y,X\beta+Zu,D,V,U\} \tag{2}$$

olan lineer karma model ele alınsın. Burada $y\in R^{n\times 1}$ gözlenebilir rasgele vektör, $X\in R^{n\times p}$ ve $Z\in R^{n\times q}$ bilinen matrisler, $\beta\in R^{p\times 1}$ sabit etkilerin bilinmeyen vektörü, $u\in R^{q\times 1}$ rasgele etkilerin gözlenemeyen vektörü, $\varepsilon\in R^{n\times 1}$ rasgele hata vektörü, $D\in R^{q\times q}$ ve $V\in R^{n\times n}$ bilinen nonnegatif (negatif olmayan) tanımlı matrisler ve $U\in R^{n\times q}$ bilinen bir matristir. (1) modelindeki y vektörünün kovaryans matrisi ise,

$$\mathrm{cov}(y) = ZDZ' + V + ZU' + UZ' \tag{3}$$

dir.

Lineer karma modellerle ilgili yapılan ilk çalışmalar 20. yüzyılın ortalarında özellikle hayvan ıslahı ve genetik alanlarında olmuştur. Bu modellerin tanımı ilk olarak Eisenhart (1947) tarafından verilmiştir. Daha sonra lineer karma modellerle ilgili tahmin problemi Henderson (1950, 1963) tarafından ele alınmıştır. Henderson (1950, 1963) tarafından elde edilen sonuçlar farklı yaklaşımlar kullanılarak Harville (1976, 1979) tarafından genelleştirilmiştir. Sabit etkili lineer modeller ve lineer karma modeller altında sabit etkilerin en iyi lineer yansız tahmin edicisi($BLUE$) ile ilgili lineer istatistiksel model literatüründe birçok çalışmaya rastlamak mümkündür. Bunlardan bazıları Rao (1967), Puntanen & Styan (1989), Nurhonen & Puntanen (1992), Puntanen (1996, 1997), Bhimisankaram & Sharay (1997), Groβ & Puntanen (2000), Zhang, Liu & Lu (2004) ve Tian & Zhang (2011) olarak verilebilir. Lineer karma modellerde rasgele etkilerin en iyi lineer yansız ön tahmin edicisi ($BLUP$) ile ilgili çalışmalar ise, sabit etkiler için $BLUE$ ile ilgili çalışmalarla karşılaştırıldığında literatürde daha az yere sahiptir. Lineer karma model altında tahmin problemi Searle (1997) ve son yıllarda özellikle Liu, Rong & Liu (2008), Haslett & Puntanen (2010, 2011) ve Wang & Liu (2013) tarafından ele alınmıştır.

Bu çalışmada, ε ve u rasgele vektörlerinin ilişkili olduğu durumda lineer karma modelde $BLUE$ ve $BLUP$ denklemleri, Haslett & Puntanen (2010) tarafından ele alınmış yaklaşıma benzer olarak, sabit etkili lineer modeller ve bu modellere yeni gözlemlerin eklemesiyle elde edilen modeller vasıtasıyla elde edilmektedir. Ancak Haslett & Puntanen (2010) tarafından elde edilen sonuçlar ε ve u rasgele vektörlerinin ilişkisiz olduğu durum altında verildiğinden, bu çalışmadaki sonuçlar onların verdiği bazı sonuçların genel bir durumu olarak ele alınabilir.

2. SABİT ETKİLİ LİNEER MODEL ALTINDA BLUE

Sabit etkileri içeren lineer model $E(\varepsilon)=0$ ve $\text{cov}(\varepsilon)=\text{cov}(y)=V$ olmak üzere, $y=X\beta+\varepsilon$ ya da

$$S=\{y, X\beta, V\} \tag{4}$$

olarak gösterilebilir. Çalışmada S modelinin tutarlı, yani $y\in C(X:V)=C(X:VM)$ olduğu kabul edilmektedir. Burada $M=I_n-P$ dir.

$L\in R^{k\times p}$ olmak üzere, S modeli altında bir $L\beta$ parametrik fonksiyonunun tahmin edilebilir olmasının gerek ve yeter koşulu $C(L')\subseteq C(X')$ olmasıdır (Alalouf & Styan, 1979). Buna göre $X\beta$ vektörünün her zaman tahmin edilebilir olduğu açıkça görülmektedir. Eğer her $\beta\in R^{p\times 1}$ için $E(Gy)=X\beta$ ise, Gy tahmin edicisi $X\beta$ için bir yansız tahmin edicidir. Gy yansız tahmin edicisinin $X\beta$ için $BLUE$ olmasının gerek ve yeter koşulu ise Gy nin, $X\beta$ nın diğer tüm yansız tahmin edicilerinin kovaryansları arasında Löwner sıralamasına göre en küçük kovaryans matrisine sahip olmasıdır. Yani $E(By)=X\beta$ olacak şekildeki her By vektörü için $\text{cov}(Gy)\leq_L \text{cov}(By)$ dir (Haslett & Puntanen, 2011).

$$Gy=BLUE(X\beta|S)\Leftrightarrow G(X:VX^{\perp})=(X:0) \tag{5}$$

temel $BLUE$ denklemi olarak bilinir (Rao, 1967) .

3. YENİ GÖZLEM EKLENMİŞ SABİT ETKİLİ LİNEER MODEL ALTINDA BLUP

$E(\varepsilon_f)=0$ ve $\text{cov}(y_f)=\text{cov}(\varepsilon_f)=V_f$ olmak üzere,

$$y_f = X_f\beta + \varepsilon_f \tag{6}$$

lineer modeli ele alınsın. Burada $y_f \in R^{q\times 1}$ yeni gözlemleri içeren gözlenemeyen rasgele vektör, $X_f \in R^{p\times q}$ bilinen yeni gözlemlerle ilişkili model matris, $\beta \in R^{p\times 1}$ bilinmeyen parametrelerin vektörü, $\varepsilon_f \in R^{q\times 1}$ yeni gözlemlerle ilişkili rasgele vektördür. S sabit etkili lineer modeline (6)' da verilen yeni gözlemlerin eklenmesiyle

$$N=\left\{\begin{pmatrix} y \\ y_f \end{pmatrix}, \begin{pmatrix} X \\ X_f \end{pmatrix}\beta, \begin{pmatrix} V & V_1 \\ V_2 & V_f \end{pmatrix}\right\}=\{y_0, X_0\beta, V_0\} \tag{7}$$

modeli elde edilir. Burada $E(y_0)=X_0\beta$ ve $\text{cov}(y_0)=\begin{pmatrix} V & V_1 \\ V_2 & V_f \end{pmatrix}$ dir.

Her $\beta \epsilon R^{px1}$ için $E(Ay)=E(y_f)=X_f\beta$, yani bir başka deyişle $E(y_f - Ay)=0$ ise Ay lineer ön tahmin edicisi y_f rasgele vektörü için yansızdır denir. Bu lineer yansız ön tahmin edici diğer tüm yansız ön tahmin ediciler arasında Löwner sıralamasına göre en küçük kovaryans matrisine sahipse $BLUP$ olarak tanımlanır. Yani $E(By)=E(y_f)=X_f\beta$ olacak şekildeki her By vektörü için Ay, y_f rasgele vektörünün yansız lineer ön tahmin edicisi olmak üzere $\text{cov}(Ay-y_f)\leq_L \text{cov}(By-y_f)$ dir (Haslett & Puntanen, 2011).

$$Ay = BLUP\left(y_f \middle| N\right) \Leftrightarrow A\left(X : VX^{\perp}\right) = \left(X_f : V_2 X^{\perp}\right) \tag{8}$$

temel $BLUP$ denklemi olarak bilinir (Christensen, 2002).

4. KARMA MODEL ALTINDA BLUE ve BLUP

K lineer karma modeli ele alınsın. Bu model altında $BLUE$ ve $BLUP$ denklemlerinin sabit etkili lineer modeller ve sabit etkili lineer modellere yeni gözlemlerin eklemesiyle elde edilen modeller vasıtasıyla elde edilmesi aşağıdaki iki teoremde verilmektedir.

Teorem 1. $K = \{y, X\beta + Zu, D, V, U\}$ lineer karma modeli altında

$$Gy = BLUE\left(X\beta \middle| K\right) \Leftrightarrow G\left(X : (ZDZ' + V + ZU' + UZ')X^{\perp}\right) = \left(X : 0\right) \tag{9}$$

dır.

İspat: $e = Zu + \varepsilon$ olmak üzere, K lineer karma modelinden

$$y = X\beta + e \tag{10}$$

sabit etkili modeli elde edilir. Burada $E(e) = 0$, $\operatorname{cov}(e) = \operatorname{cov}(y) = ZDZ' + V + ZU' + UZ'$ dir. (5)'te verilen temel $BLUE$ denklemi (10) modeli için ele alındığında (9) elde edilir ve böylece ispat tamamlanmış olur.

(9)' da verilen ifadenin diğer bir gösterimi

$$Gy = BLUE\left(X\beta \middle| K\right) \Leftrightarrow \begin{pmatrix} ZDZ' + V + ZU' + UZ' & X \\ X' & 0 \end{pmatrix} \begin{pmatrix} G' \\ L \end{pmatrix} = \begin{pmatrix} 0 \\ X' \end{pmatrix} \tag{11}$$

dir. $X\beta$ için $BLUE$ nun (11) ile elde edilmesi IPM metodu olarak bilinir (Rao, 1971, 1972).

Teorem 2. $K=\{y, X\beta+Zu, D, V, U\}$ lineer karma modeli altında

$$Ay=BLUP(u|K) \Leftrightarrow A(X:(ZDZ'+V+ZU'+UZ')X^{\perp})=(0:(DZ'+U')X^{\perp}) \quad \textbf{(12)}$$

dir.

İspat: $E(\varepsilon_f)=0$, $\mathrm{cov}(\varepsilon_f)=\mathrm{cov}(u)=D$ olmak üzere,

$$u=0\beta+\varepsilon_f \quad \textbf{(13)}$$

denklemi ele alınsın. K lineer modeli, yeni gözlemlerin eklenmesiyle elde edilen sabit etkili lineer model vasıtasıyla ifade edilebilir. K modeli (13) denklemi ile birlikte ele alındığında N modeline benzer olarak

$$N_*=\left\{\begin{pmatrix} y \\ u \end{pmatrix}, \begin{pmatrix} X \\ 0 \end{pmatrix}\beta, \begin{pmatrix} ZDZ'+V+ZU'+UZ' & ZD+U \\ DZ'+U' & D \end{pmatrix}\right\} \quad \textbf{(14)}$$

modeli yazılabilir. (8)'de verilen temel $BLUP$ denklemi (14) modeli için ele alındığında (12) elde edilir ve böylece ispat tamamlanmış olur.

(12)' de verilen ifade, (11)' deki ifadeye benzer olarak IPM metodu vasıtasıyla

$$Ay=BLUP(u|K) \Leftrightarrow \begin{pmatrix} ZDZ'+V+ZU'+UZ' & X \\ X' & 0 \end{pmatrix}\begin{pmatrix} A' \\ L \end{pmatrix}=\begin{pmatrix} (X^{\perp})'(ZD+U) \\ 0 \end{pmatrix} \quad \textbf{(15)}$$

olarak gösterilebilir.

Sonuç 1. $K=\{y, X\beta+Zu, D, V, U\}$ lineer karma modeli ele alınsın. $Gy=BLUE(X\beta|K)$ ve $Ay=BLUP(u|K)$ olmak üzere, $BLUE$ ve $BLUP$ çözümlerini birlikte veren denklem

$$\begin{pmatrix} G \\ A \end{pmatrix}(X:(ZDZ'+V+ZU'+UZ')X^{\perp})=\begin{pmatrix} X & 0 \\ 0 & (DZ'+U')X^{\perp} \end{pmatrix} \quad \textbf{(16)}$$

dir.

Aşağıdaki sonuçta u ve ε vektörlerinin ilişkisiz olduğu durumda (16) denklemine karşılık gelen denklem verilmiştir.

Sonuç 2. u ve ε vektörlerinin ilişkisiz olduğu kabul edilsin. Bu durumda $Gy = BLUE\left(X\beta \middle| K\right)$ ve $Ay = BLUP\left(u \middle| K\right)$ olmak üzere, K lineer karma modeli altında $BLUE$ ve $BLUP$ çözümlerini birlikte veren denklem

$$\begin{pmatrix} G \\ A \end{pmatrix}(X : (ZDZ' + V)X^{\perp}) = \begin{pmatrix} X & 0 \\ 0 & DZ'X^{\perp} \end{pmatrix} \tag{17}$$

dir.

Uyarı. $X^{\perp}$ matrisi yerine M dik izdüşüm matrisinin bir seçim olduğu açıktır. Bu nedenle (5), (8), (9), (12) ve (15-17) denklemlerinde $X^{\perp}$ matrisi yerine M dik izdüşüm matrisi yazılabilir.

KAYNAKLAR

[1] Alalouf, I. S. & Styan, G. P. H. Characterizations of estimability in the general linear model. The Annals of Statistics, 7, 194–200, 1979.

[2] Bhimasankaram, P. & Saharay, R. On a partitioned linear model and some associated reduced models. Linear Algebra Appl., 264, 329–339, 1997.

[3] Christensen, R. Plane answers to complex questions: the theory of linear models. 3rd edn. Springer, New York, 2002.

[4] Eisenhart, C. The assumptions underlying the analysis of variance. Biometrics, 3, 1–21, 1947.

[5] Groβ, J. & Puntanen, S. Estimation under a general partitioned linear model. Linear Algebra Appl., 321, 131–144, 2000.

[6] Handerson, C. R. Estimation of genetic parameters. Ann. Math. Stat., 21, 309–310, 1950.

[7] Handerson, C. R. Selection index and expected genetic advance. Statistical genetic and plant breeding, National Academy of Sciences, National Research Council Publication, 982, 141–163, 1963.

[8] Haslett, S. J. & Puntanen, S. Equality of BLUEs or BLUPs under two linear models using stochastic restrictions. Statistical Papers, 51, 465–475, 2010.

[9] Haslett, S. J. & Puntanen, S. On the equality of the BLUPs under two linear mixed models. Metrika, 74, 381–395, 2011.

[10] Harville, D. Extension of the Gauss–Markov theorem to include the estimation of random effects. Ann Stat., 4, 384–395, 1976.

[11] Harville, D. Some useful representations for constrained mixed–model estimation. J. Amer. Stat. Assoc., 74, 200–206, 1979.

[12] Liu, X. Q., Rong, J. Y. & Liu, X. Y. Best linear unbiased prediction for linear combinations in general mixed linear models. J. Mult. Anal., 99, 1503–1517, 2008.

[13] Nurhonen, M. & Puntanen, S. A property of partitioned generalized regression. Commun. Statist. Theor. Meth., 21, 1579–1583, 1992.

[14] Puntanen, S. Some matrix results related to a partitioned singular linear model. Commun. Statist. Theor. Meth., 25(2), 269–279, 1996.

[15] Puntanen, S. Some further results related to reduced singular linear models. Commun. Statist. Theor. Meth., 26(2), 375–385, 1997.

[16] Puntanen, S. & Styan, G. P. H. The equality of the ordinary least squares estimator and the best linear unbiased estimator (with discussion). Amer. Statist., 43, 153–164, 1989.

[17] Rao, C. R. Least squares theory using an estimated dispersion matrix and its application to measurement of signals. In: Le Cam, L.M., Neyman, J., eds, Proc. Fifth Berkeley Symposium on Mathematical Statistics and Probability: Berkeley, California, 1965/1966, Vol. 1, Berkeley: University of California Press. 355–372, 1967.

[18] Rao, C. R. Unified theory of linear estimation. Sankhyā, A33, 371–394, 1971. [Corrigendum: 34, 194, 477, 1972.]

[19] Rao, C. R. A note on the IPM method in the unified theory of linear estimation. Sankhyā, A34, 285–288, 1972.

[20] Searle, S. R. the matrix handling of BLUE and BLUP in the mixed linear model. Linear Algebra Appl., 264, 291–311, 1997.

[21] Tian, Y. & Zhang, J. Some equalities for estimations of partial coefficients under a general linear regression model. Statistical Papers, 52, 911–920, 2011.

[22] Wang, W. W. & Liu, X. The equalities of BLUPs for linear combinations under two general mixed models, Commun. Statist. Theor. Meth., 42, 3528–3543, 2013.

[23] Zhang, B. X., Liu, B. S. & Lu, C. Y. A study of the equivalence of the BLUEs between a partitioned singular linear model and its reduced singular linear models. Acta Mathematica Sinica, English Ser., 203, 557–568, 2004.