



Aktüerya Derneği

İstatistikçiler Dergisi: İstatistik & Aktüerya

Journal of Statisticians: Statistics and Actuarial Sciences

IDIA 11, 2018, 1, 32-41

Geliş/Received:31.01.2018, Kabul/Accepted: 03.05.2018


www.istatistikciler.org

Araştırma Makalesi / Research Article

Matris değişkenli Laplace dağılımı: Özellikleri ve parametre tahmini

Y. Murat Bulut

*Eskişehir Osmangazi Üniversitesi
İstatistik Bölümü, Eskişehir, Türkiye*
ybulut@ogu.edu.tr

 0000-0002-0545-7339

Öz

Bu çalışmada matris değişkenli normal dağılım ve üstel dağılımın ölçek karması olarak matris değişkenli Laplace dağılımı önerilmiştir. Ayrıca önerilen dağılımın özellikleri incelenmiştir ve parametrelerinin tahmini için EM (Expectation-Maximization) algoritmasına dayalı en çok olabirlik tahmin edicileri verilmiştir.

Anahtar Kelimeler: EM algoritması, matris değişkenli dağılım, parametre tahmini.

Abstract

Matrix variate Laplace distribution: Properties and parameter estimation

In this study, we proposed matrix variate Laplace distribution as a scale mixture of matrix variate normal distribution and exponential distribution. Also, we examine some distributional properties and give maximum likelihood estimators of its parameters based on the EM algorithm.

Keywords: EM algorithm, matrix variate distribution, parameter estimation.

1. Giriş

Laplace dağılımı ekonomi, sağlık bilimleri ve mühendislik başta olmak üzere birçok alanda sıklıkla kullanılmaktadır (Kotz vd., 2001). Tek değişkenli Laplace dağılımı 1774 yılında Laplace tarafından tanımlanmıştır. Laplace dağılımı genellikle normal dağılımdan daha kalın kuyruklu durumlarda modelleme için kullanılmaktadır. Andrews ve Mallows (1974) normal dağılımın ölçek karması olarak Laplace dağılımını tanımlamışlardır. Ölçek karması yaklaşımı daha kalın kuyruklu dağılımlar türetmek için kullanılan bir yöntemdir. İki değişkenli Laplace dağılımı ise Ulrich ve Chen (1987) tarafından tanımlanmıştır. Çok değişkenli Laplace dağılımı ise Anderson (1992) tarafından tanımlanan çok değişkenli Linnik dağılımının özel hali olarak ifade edilmektedir. Eltoft (2006) ise ölçek karması yaklaşımından yararlanarak çok değişkenli Laplace dağılımını tanımlamış, bazı dağılımsal özelliklerini incelemiş ve parametrelerinin tahmini için EM algoritmasına (Dempster vd., 1977) dayalı bir tahmin yöntemi vermiştir. Normal dağılımın ölçek karması şeklinde tanımlamanın avantajı EM algoritması kullanılarak parametre tahminlerinin daha kolay yapılabilmesi olmasıdır.

Tek değişkenli Laplace dağılımının karakteristik fonksiyonu ve olasılık yoğunluk fonksiyonu Eşitlik (1) ve Eşitlik (2)'deki gibi tanımlanır.

$$\phi(t; \mu, \sigma) = \frac{e^{i\mu t}}{1 + \sigma^2 t^2}, \quad (1)$$

$$f(x; \mu, \sigma) = \frac{1}{2\sigma} e^{-\frac{|x-\mu|}{\sigma}}, x \in R. \quad (2)$$

Çok değişkenli Laplace dağılımının karakteristik fonksiyonu ve olasılık yoğunluk fonksiyonu ise Eşitlik (3) ve Eşitlik (4)'te verilmiştir.

$$\phi(\mathbf{t}; \boldsymbol{\mu}, \Sigma) = \frac{e^{i\boldsymbol{\mu}^T \mathbf{t}}}{1 + \frac{1}{2} \mathbf{t}' \Sigma \mathbf{t}}, \quad (3)$$

$$f(\mathbf{x}; \boldsymbol{\mu}, \Sigma) = \frac{|\Sigma|^{-\frac{1}{2}}}{2^p \pi^{\frac{p-1}{2}} \Gamma(\frac{p+1}{2})} e^{-\sqrt{(\mathbf{x}-\boldsymbol{\mu})' \Sigma^{-1} (\mathbf{x}-\boldsymbol{\mu})}}, x \in R^p, p \geq 1. \quad (4)$$

Matris değişkenli durumda ise özel olarak Laplace dağılımı ele alınmamıştır. Fakat Sánchez-Manzano vd. (2002) tarafından tanımlanan matris değişkenli güç üstel dağılımın özel hali ya da Gallagher ve McNicholas (2017) tarafından tanımlanan matris değişkenli varyans-gamma dağılımının özel hali olarak düşünülebilir. Bu çalışmanın amacı, uygulama da normal dağılıma alternatif olarak kullanılan Laplace dağılımını matris değişkenli durumda matris değişkenli normal dağılımın ölçek karması olarak tanımlamaktır. Matris değişkenli dağılımlar son zamanlarda görüntü tanıma konusunda sıklıkla kullanılmaktadır. Önerilen dağılımın bu alanda matris değişkenli normal dağılıma alternatif bir dağılım olarak kullanılabileceği düşünülmektedir.

Bölüm 2'de matris değişkenli Laplace dağılımı tanımlanmış ve istatistiksel özellikleri incelenmiştir. Bölüm 3'te EM algoritması yardımıyla tanımlanan dağılımın parametreleri tahmin edilmiştir. Bölüm 4'te ise önerilen algoritmanın performansını test etmek için simülasyon çalışması yapılmıştır. Sonuç ve öneriler ise Bölüm 5'te verilmiştir.

2. Matris değişkenli Laplace dağılımı

Bu bölümde normal dağılıma alternatif olarak önerilen Laplace dağılımı matris değişkenli olarak tanımlanacaktır. Ayrıca, önerilen dağılımın bazı dağılımsal özellikleri incelenecek ve dağılımın lineer dönüşüm altında değişmez olduğu gösterilecektir.

Tanım 2.1: $X \in R^{n \times p}$, $n \geq 1$, $p \geq 1$ rassal matrisi Eşitlik (5)'te verilen hali ile matris değişkenli normal dağılımın ölçek karması olarak yazılabilir.

$$X = M + V^{\frac{1}{2}} \Sigma^{\frac{1}{2}} Z \Psi^{\frac{1}{2}}, \quad (5)$$

burada $Z \sim N_{n,p}(0, I_n, I_p)$ ve $V \sim \text{Üstel}(\lambda)$ dağılımlarına sahiptir ve tek boyutlu V rassal değişkeni, $n \times p$ boyutlu Z matris değişkeninden bağımsızdır. Burada $M \in R^{n \times p}$ konum matrisi, $\Sigma^{\frac{1}{2}}$ ve $\Psi^{\frac{1}{2}}$ matrisleri sırasıyla pozitif tanımlı Σ ve Ψ varyans-kovaryans matrislerinin kareköküdürler. $V = v$ verildiğinde X rassal matrisinin dağılımı matris değişkenli normal dağılımdır ($X \sim N_{n,p}(M, v\Sigma, \Psi)$). Matris değişkenli normal dağılıma sahip Z matrisinin beklenen değeri ve kovaryans matrisi aşağıdaki gibidir (Gupta ve Nagar, 1999).

$$E(Z) = 0,$$

$$\text{Cov}(Z) = I_n \otimes I_p.$$

Burada \otimes Kronecker çarpımını ifade etmektedir. Bu beklenen değer ve kovaryans yardımıyla Eşitlik (5) ile verilen X rassal matrisinin beklenen değer ve kovaryansı, Eşitlik (6) ve Eşitlik (7)'deki gibi elde edilir.

$$E(X) = M, \quad (6)$$

$$Cov(X) = \lambda(\Sigma \otimes \Psi). \quad (7)$$

Önerme 2.1: $Z \sim N_{n,p}(0, I_n, I_p)$ ve $V \sim \text{Üstel}(\lambda)$ olsun. Ayrıca Z rassal matrisi ve V rassal değişkeninin birbirinden bağımsız olduğunu varsayalım. Bu durumda Eşitlik (5)'teki gibi tanımlanan rassal X matrisinin olasılık yoğunluk fonksiyonu aşağıdaki gibidir.

$$\begin{aligned} f_X(X; M, \Sigma, \Psi, \lambda) &= \left(\frac{2}{\lambda}\right)^{\frac{np}{4} + \frac{1}{2}} (2\pi)^{-\frac{np}{2}} |\Sigma|^{-\frac{p}{2}} |\Psi|^{-\frac{n}{2}} \\ &\quad \times [tr\{\Sigma^{-1}(X - M)\Psi^{-1}(X - M)'\}]^{\frac{1}{2} - \frac{np}{4}} \\ &\quad \times K_{1 - \frac{np}{2}} \left(\sqrt{\frac{2}{\lambda} tr\{\Sigma^{-1}(X - M)\Psi^{-1}(X - M)'\}} \right). \end{aligned} \quad (8)$$

Burada, $tr(\cdot)$ iz fonksiyonu, $K_m(x)$ ise üçüncü tip düzeltilmiş Bessel fonksiyonudur.

İspat: Z rassal matrisinin ve V rassal değişkeninin ortak olasılık yoğunluk fonksiyonu

$$f_{Z,V}(Z, v) = \frac{(2\pi)^{-\frac{np}{2}}}{\lambda} \exp\left\{-\frac{v}{\lambda}\right\} \exp\left\{-\frac{1}{2} tr(ZZ')\right\}$$

şeklinde. Eşitlik (5) ile verilen dönüşüm kullanılırsa, $Z = V^{-\frac{1}{2}} \Sigma^{-\frac{1}{2}} (X - M) \Psi^{-\frac{1}{2}}$ olur. Dönüşüme ait jakobiyen değeri ise $V^{-\frac{np}{2}} |\Sigma|^{-\frac{p}{2}} |\Psi|^{-\frac{n}{2}}$ olarak hesaplanır. Bu değerler yardımıyla X ve V rassal değişkenlerinin ortak olasılık yoğunluk fonksiyonu aşağıdaki gibi elde edilir.

$$\begin{aligned} f_{X,V}(X, v) &= \lambda^{-1} (2\pi)^{-\frac{np}{2}} |\Sigma|^{-\frac{p}{2}} |\Psi|^{-\frac{n}{2}} v^{-\frac{np}{2}} \\ &\quad \times \exp\left\{-\frac{1}{2} \left(\frac{2}{\lambda} v + v^{-1} tr\{\Sigma^{-1}(X - M)\Psi^{-1}(X - M)'\}\right)\right\}. \end{aligned}$$

Daha sonra V rassal değişkenine göre $f_{X,V}(X, v)$ fonksiyonunun integrali alınırsa X rassal matrisinin olasılık yoğunluk fonksiyonu

$$\begin{aligned} f_X(X; M, \Sigma, \Psi, \lambda) &= \int_0^{\infty} f_{X,V}(X, v) dv \\ &= \left(\frac{2}{\lambda}\right)^{\frac{np}{4} + \frac{1}{2}} (2\pi)^{-\frac{np}{2}} |\Sigma|^{-\frac{p}{2}} |\Psi|^{-\frac{n}{2}} \\ &\quad \times [tr\{\Sigma^{-1}(X - M)\Psi^{-1}(X - M)'\}]^{\frac{1}{2} - \frac{np}{4}} \\ &\quad \times K_{1 - \frac{np}{2}} \left(\sqrt{\frac{2}{\lambda} tr\{\Sigma^{-1}(X - M)\Psi^{-1}(X - M)'\}} \right) \end{aligned}$$

olarak elde edilir. Burada $\lambda = 1, p = 1$ ve $\Psi = I_p$ alınırsa çok değişkenli Laplace dağılımını elde edilir (Kotz vd., 2001).

Tanım 2.2: $X \in R^{n \times p}, n > 1, p > 1$ rassal matrisi Eşitlik (8)'de verilen olasılık yoğunluk fonksiyonuna sahip ise matris değişkenli Laplace dağılımına sahiptir denir ve kısaca $X \sim MVL_{n,p}(M, \Sigma, \Psi, \lambda)$ şeklinde gösterilir.

Önerme 2.2: $X \sim MVL_{n,p}(M, \Sigma, \Psi, \lambda)$ olsun. A $n \times p$ boyutlu bir matris, B ve C matrisleri sırasıyla $n \times n$ ve $p \times p$ boyutlu tekil olmayan matrisler olmak üzere $Y = A + BXC$ şeklinde tanımlanan rassal Y matrisi matris değişkenli Laplace dağılımına sahiptir ve $Y \sim MVL_{n,p}(A + BMC, B\Sigma B', C'\Psi C, \lambda)$ şeklinde ifade edilir. Burada $A + BMC$ konum matrisi, $B\Sigma B'$ ve $C'\Psi C$ matrisleri varyans-kovaryans matrisleridir.

İspat: $Y = A + BXC$ olduğundan, $X = B^{-1}(Y - A)C^{-1}$ dönüşümü uygularsak jakobiyen $|B|^{-p}|C|^{-n}$ olur. Bu durumda Y rassal matrisinin olasılık yoğunluk fonksiyonu aşağıdaki gibi elde edilir.

$$f_Y(Y) = \left(\frac{2}{\lambda}\right)^{\frac{np+1}{4}+\frac{1}{2}} (2\pi)^{-\frac{np}{2}} |B\Sigma B'|^{-\frac{p}{2}} |C'\Psi C|^{-\frac{n}{2}} \\ \times \left[\text{tr} \left\{ (B\Sigma B')^{-1} (Y - (A + BMC)) (C'\Psi C)^{-1} (Y - (A + BMC))' \right\} \right]^{\frac{1}{2} - \frac{np}{4}} \\ \times K_{1-\frac{np}{2}} \left(\sqrt{\frac{2}{\lambda} \text{tr} \left\{ (B\Sigma B')^{-1} (Y - (A + BMC)) (C'\Psi C)^{-1} (Y - (A + BMC))' \right\}} \right).$$

Bu durumda, Y rassal matrisi $A + BMC, B\Sigma B', C'\Psi C, \lambda$ parametreleri ile matris değişkenli Laplace dağılımına sahiptir.

Önerme 2.3: $X \sim MVL_{n,p}(M, \Sigma, \Psi, \lambda)$ olsun. X rassal matrisinin karakteristik fonksiyonu,

$$\phi_X(T) = \frac{\text{etr}(iT'M)}{1 + \frac{\lambda}{2} \text{tr}(T'\Sigma T\Psi)} \tag{9}$$

şeklindedir. Burada $\text{etr} = e^{\text{tr}(\cdot)}$ dir.

$$\phi_X(T) = E_X[\text{etr}(iXT')]$$

İspat:

$$\begin{aligned} &= E_V \left[E_{X|V}[\text{etr}(iXT')] \right] \\ &= E_V \left[\text{etr} \left(iT'M - \frac{1}{2} vT'\Sigma T\Psi \right) \right] \\ &= E_V \left[\text{etr}(iT'M) \text{etr} \left(-\frac{v}{2} T'\Sigma T\Psi \right) \right] \\ &= \text{etr}(iT'M) E_V \left[\exp \left\{ -\frac{v}{2} \text{tr}(T'\Sigma T\Psi) \right\} \right] \\ &= \text{etr}(iT'M) \int_0^\infty \exp \left(-\frac{v}{2} \text{tr}(T'\Sigma T\Psi) \right) \frac{1}{\lambda} \exp \left(-\frac{v}{\lambda} \right) dv \\ &= \frac{\text{etr}(iT'M)}{\lambda} \int_0^\infty \exp \left\{ -v \left(\frac{1}{\lambda} + \frac{\text{tr}(T'\Sigma T\Psi)}{2} \right) \right\} dv \\ &= \frac{\text{etr}(iT'M)}{1 + \frac{\lambda}{2} \text{tr}(T'\Sigma T\Psi)}. \end{aligned}$$

2. Parametre tahmini

Bu bölümde önerilen dağılımın parametrelerinin tahmini için EM algoritması kullanılmıştır (Dempster vd., 1977). X_1, X_2, \dots, X_l ($X_i \in R^{n \times p}$, $i = 1, 2, \dots, l$) bir birinden bağımsız aynı dağılımlı veri matrisleri olsun. Bu veri matrisleri M, Σ, Ψ ve λ bilinmeyen parametrelere sahip matris değişkenli Laplace dağılımı yardımı ile modellenmek istenmektedir. Bilinmeyen parametre matrislerinin en çok olabilirlik tahmin edicilerini elde edebilmek için aşağıdaki log-olabilirlik fonksiyonunun maksimize edilmesi gerekir.

$$l(M, \Sigma, \Psi, \lambda) = \left(\frac{npl}{4} + \frac{l}{2} \right) \ln \left(\frac{2}{\lambda} \right) - \frac{npl}{2} \ln(2\pi) - \frac{pl}{2} \ln|\Sigma| - \frac{nl}{2} \ln|\Psi| \\ + \left(\frac{1}{2} - \frac{np}{4} \right) \sum_{i=1}^l \text{tr} \{ \Sigma^{-1} (X_i - M) \Psi^{-1} (X_i - M)' \} \\ + \sum_{i=1}^l \ln \left(K_{1-\frac{np}{2}} \left(\sqrt{\frac{2}{\lambda} \text{tr} \{ \Sigma^{-1} (X_i - M) \Psi^{-1} (X_i - M)' \}} \right) \right).$$

Bu çalışmada, log-olabilirlik fonksiyonunun doğrudan maksimizasyonu zor olduğundan parametrelerin tahmini için dağılımın ölçek karması gösteriminden yararlanarak EM algoritması yardımıyla tahmin ediciler elde edilmiştir. X_i ve V_i değişkenleri sırasıyla gözlenen ve kayıp değişkenler, (X_i, V_i) ise tam veri olarak tanımlanır ise tam verinin log-olabilirlik fonksiyonu

$$L_c(M, \Sigma, \Psi, \lambda) = \prod_{i=1}^l f(X_i, v_i) \\ = -l \ln(\lambda) - \frac{npl}{2} \ln(2\pi) - \frac{pl}{2} \ln|\Sigma| - \frac{nl}{2} \ln|\Psi| - \frac{np}{2} \sum_{i=1}^l \ln(v_i) \\ - \frac{1}{\lambda} \sum_{i=1}^l v_i - \frac{1}{2} \sum_{i=1}^l v_i^{-1} \text{tr} \{ \Sigma^{-1} (X_i - M) \Psi^{-1} (X_i - M)' \}$$

şeklinde elde edilir. Log-olabilirlik fonksiyonunda parametre içermeyen terimler ihmal edilebilir. Tam verideki V değişkeni gözlemlenemediğinden gözlenen veri X_i ve parametrelerin şu anki tahmin değerleri verildiğinde $L_c(M, \Sigma, \Psi, \lambda)$ fonksiyonunun koşullu beklenen değeri alınırsa amaç fonksiyonu aşağıdaki gibi elde edilir.

$$Q(M, \Sigma, \Psi, \lambda) = E(L_c(M, \Sigma, \Psi, \lambda) | X_i, \hat{M}, \hat{\Sigma}, \hat{\Psi}, \hat{\lambda}) \\ = -l \ln(\lambda) - \frac{pl}{2} \ln|\Sigma| - \frac{nl}{2} \ln|\Psi| - \frac{1}{\lambda} \sum_{i=1}^l E(v_i | X_i, \hat{M}, \hat{\Sigma}, \hat{\Psi}, \hat{\lambda}) \\ - \frac{1}{2} \sum_{i=1}^l E(v_i^{-1} | X_i, \hat{M}, \hat{\Sigma}, \hat{\Psi}, \hat{\lambda}) \text{tr} \{ \Sigma^{-1} (X_i - M) \Psi^{-1} (X_i - M)' \}.$$

Burada $E(V_i | X_i, \hat{M}, \hat{\Sigma}, \hat{\Psi}, \hat{\lambda})$ ve $E(V_i^{-1} | X_i, \hat{M}, \hat{\Sigma}, \hat{\Psi}, \hat{\lambda})$ ifadeleri V_i ve V_i^{-1} rassal değişkenlerinin gözlenen veri ve parametrelerin şu anki tahminleri verildiğindeki koşullu beklenen değerleridir. Bu koşullu beklenen değerleri hesaplayabilmek için X verildiğinde V rassal değişkeninin koşullu dağılımının elde edilmesi gerekir. Bu koşullu dağılım tam verinin ortak olasılık yoğunluk fonksiyonu kullanılarak Eşitlik (10)'daki gibi elde edilir.

(10)

$$f(v|X) = \left(\sqrt{\frac{\lambda}{2} \text{tr}\{\Sigma^{-1}(X - M)\Psi^{-1}(X - M)'\}} \right)^{\frac{np}{2}-1} v^{-\frac{np}{2}} \times \frac{\exp\left\{-\frac{1}{2}\left(\frac{2}{\lambda}v + v^{-1}\text{tr}\{\Sigma^{-1}(X - M)\Psi^{-1}(X - M)'\}\right)\right\}}{2K_{1-\frac{np}{2}}\left(\sqrt{\frac{2}{\lambda}\text{tr}\{\Sigma^{-1}(X - M)\Psi^{-1}(X - M)'\}}}\right)}$$

Yukarıda elde edilen koşullu dağılımın olasılık yoğunluk fonksiyonu incelendiğinde bu dağılımın genelleştirilmiş ters gamma dağılımı olduğu görülmektedir. Dolayısıyla

$V|X \sim GIG\left(1 - \frac{np}{2}, \sqrt{\text{tr}\{\Sigma^{-1}(X - M)\Psi^{-1}(X - M)'\}}, \sqrt{\frac{2}{\lambda}}\right)$ olarak ifade edilir (Barndorff-Nielsen, 1997).

Eşitlik (10)'da verilen koşullu dağılım yardımıyla $E(V_i|X_i, \hat{M}, \hat{\Sigma}, \hat{\Psi}, \hat{\lambda})$ ve $E(V_i^{-1}|X_i, \hat{M}, \hat{\Sigma}, \hat{\Psi}, \hat{\lambda})$ koşullu beklenen değerleri sırasıyla aşağıdaki biçimde elde edilir.

$$w_i = E(V_i|X_i, \hat{M}, \hat{\Sigma}, \hat{\Psi}, \hat{\lambda}) \tag{11}$$

$$= \frac{\sqrt{\frac{\hat{\lambda}}{2} \text{tr}\left(\hat{\Sigma}^{-1}(X_i - \hat{M})\hat{\Psi}^{-1}(X_i - \hat{M})'\right)} K_{2-\frac{np}{2}}\left(\sqrt{\frac{2}{\hat{\lambda}} \text{tr}\left(\hat{\Sigma}^{-1}(X_i - \hat{M})\hat{\Psi}^{-1}(X_i - \hat{M})'\right)}\right)}{K_{1-\frac{np}{2}}\left(\sqrt{\frac{2}{\hat{\lambda}} \text{tr}\left(\hat{\Sigma}^{-1}(X_i - \hat{M})\hat{\Psi}^{-1}(X_i - \hat{M})'\right)}\right)}$$

$$\eta_i = E(V_i^{-1}|X_i, \hat{M}, \hat{\Sigma}, \hat{\Psi}, \hat{\lambda}) \tag{12}$$

$$= \frac{\frac{1}{2} K_{2-\frac{np}{2}}\left(\sqrt{\frac{2}{\hat{\lambda}} \text{tr}\left(\hat{\Sigma}^{-1}(X_i - \hat{M})\hat{\Psi}^{-1}(X_i - \hat{M})'\right)}\right)}{\sqrt{\hat{\lambda} \text{tr}\left(\hat{\Sigma}^{-1}(X_i - \hat{M})\hat{\Psi}^{-1}(X_i - \hat{M})'\right)} K_{1-\frac{np}{2}}\left(\sqrt{\frac{2}{\hat{\lambda}} \text{tr}\left(\hat{\Sigma}^{-1}(X_i - \hat{M})\hat{\Psi}^{-1}(X_i - \hat{M})'\right)}\right)}$$

Koşullu beklenen değerler w_i ve η_i olarak yerine yazılırsa M, Σ, Ψ ve λ parametrelerini tahmin etmek için Eşitlik (13)'de verilen amaç fonksiyonu elde edilir.

$$Q(M, \Sigma, \Psi, \lambda) = -l \ln(\lambda) - \frac{pl}{2} \ln|\Sigma| - \frac{nl}{2} \ln|\Psi| - \frac{1}{\lambda} \sum_{i=1}^l w_i - \frac{1}{2} \sum_{i=1}^l \eta_i \text{tr}\{\Sigma^{-1}(X_i - M)\Psi^{-1}(X_i - M)'\}. \tag{13}$$

İlgili parametrelere göre amaç fonksiyonunun türevlerinin alınıp sifıra eşitlenmesiyle Eşitlik (14)-(17)'de verilen konum matrisi, varyans-kovaryans matrisleri ve ölçek parametresinin tahmin edicileri elde edilir.

$$\hat{\lambda} = \frac{1}{l} \sum_{i=1}^l \hat{w}_i, \tag{14}$$

$$\hat{M} = \frac{\frac{1}{l} \sum_{i=1}^l \hat{\eta}_i X_i}{\sum_{i=1}^l \hat{\eta}_i}, \tag{15}$$

$$\hat{\Sigma} = \frac{1}{pl} \sum_{i=1}^l \hat{\eta}_i (X_i - \hat{M}) \hat{\Psi}^{-1} (X_i - \hat{M})', \quad (16)$$

$$\hat{\Psi} = \frac{1}{nl} \sum_{i=1}^l \hat{\eta}_i (X_i - \hat{M})' \hat{\Sigma}^{-1} (X_i - \hat{M}). \quad (17)$$

EM algoritmasının adımları aşağıda özetlenmiştir.

E-Adımı: X_i veri matrisleri ve parametrelerin şu anki değerleri verildiğinde tam verinin log-olabilirlik fonksiyonundaki koşullu beklenen değerler bulunur. Bu adım $Q(M, \Sigma, \Psi, \lambda)$ amaç fonksiyonundaki w_i ve η_i koşullu beklenen değerlerinin hesaplanmasına karşılık gelmektedir.

M-Adımı: Yeni parametre tahminlerini elde etmek için M , Σ , Ψ ve λ parametrelerine göre $Q(M, \Sigma, \Psi, \lambda)$ fonksiyonu maksimize edilir. Bu adım E-adımında hesaplanan koşullu beklenen değerleri kullanarak Eşitlik (14)-(17) ile verilen parametre değerlerinin güncellenmesidir.

EM algoritmasının adımlarını kullanarak, parametrelerin en çok olabilirlik tahmin edicilerini hesaplamak için aşağıda verilen iteratif yeniden ağırlıklandırılmış algoritma yazılabilir.

İteratif yeniden ağırlıklandırılmış algoritma

1. İterasyon sayısı $k=1$ alınır ve parametrelerin başlangıç değerleri seçilir.
2. $k = 1, 2, \dots$, için $M^{(k)}$, $\Sigma^{(k)}$, $\Psi^{(k)}$ ve $\lambda^{(k)}$ parametrelerinin güncel değerleri ve Eşitlik (11) ve Eşitlik (12) ile verilen $w_i^{(k)}$ ve $\eta_i^{(k)}$ ağırlık değerleri hesaplanır.
3. $M^{(k)}$, $\Sigma^{(k)}$, $\Psi^{(k)}$ ve $\lambda^{(k)}$ ile gösterilen parametrelerin yeni değerlerini hesaplamak için aşağıdaki eşitlikler kullanılır.

$$\hat{\lambda}^{(k+1)} = \frac{1}{l} \sum_{i=1}^l \hat{w}_i^{(k)}, \quad (18)$$

$$\hat{M}^{(k+1)} = \frac{\frac{1}{l} \sum_{i=1}^l \hat{\eta}_i^{(k)} X_i}{\sum_{i=1}^l \hat{\eta}_i^{(k)}}, \quad (19)$$

$$\hat{\Sigma}^{(k+1)} = \frac{1}{pl} \sum_{i=1}^l \hat{\eta}_i^{(k)} (X_i - \hat{M}^{(k)}) \hat{\Psi}^{(k)-1} (X_i - \hat{M}^{(k)})', \quad (20)$$

$$\hat{\Psi}^{(k+1)} = \frac{1}{nl} \sum_{i=1}^l \hat{\eta}_i^{(k)} (X_i - \hat{M}^{(k)})' \hat{\Sigma}^{(k)-1} (X_i - \hat{M}^{(k)}). \quad (21)$$

4. Yakınsama sağlanıncaya kadar adımlar tekrar edilir.

4. Simülasyon çalışması

Bu bölümde iteratif yeniden ağırlıklandırılmış algoritmanın iddia edildiği gibi çalıştığını göstermek için simülasyon çalışması yapılacaktır. Simülasyonda kullanılacak veriler Eşitlik (5) ile verilen ölçek karması gösteriminden yararlanılarak türetilmiştir. Ayrıca çalışmada $\lambda = 1$ olarak alınmış ve sabit olarak kabul edilmiştir. Σ ve Ψ varyans-kovaryans matrisleri ve M konum matrisi aşağıdaki gibi seçilmiştir.

$$M = \begin{bmatrix} 1 & 0 & 0 & 2 \\ 0 & 0.5 & -0.5 & 0 \\ -1 & 1 & 2 & 1 \end{bmatrix}, \Sigma = \begin{bmatrix} 1 & 0.6 & 0.3 \\ 0.6 & 1 & 0.2 \\ 0.3 & 0.2 & 1 \end{bmatrix} \text{ ve } \Psi = \begin{bmatrix} 1 & 0.5 & 0.5 & 0.1 \\ 0.5 & 1 & 0 & 0.6 \\ 0.5 & 0 & 1 & 0.4 \\ 0.1 & 0.6 & 0.4 & 1 \end{bmatrix}$$

Matris değişkenli normal dağılıma sahip Z matrisi R programındaki LaplacesDemon paketi kullanılarak türetilmiştir. V ise 1 parametrelili üstel dağılımdan türetilmiştir. Tahmin edicilerin performanslarını göstermek için parametrelerin tahmin değerleri ile gerçek değerleri arasındaki Öklid uzaklığı kullanılmıştır. Kullanılan Öklid uzaklıkları,

$$\|\hat{M} - M\| = \left(\sum_{i=1}^n \sum_{j=1}^p (\hat{m}_{ij} - m_{ij})^2 \right)^{\frac{1}{2}}, \quad \|\hat{\Sigma} - \Sigma\| = \left(\sum_{i=1}^n \sum_{j=1}^p (\hat{\sigma}_{ij} - \sigma_{ij})^2 \right)^{\frac{1}{2}} \quad \text{ve}$$

$\|\hat{\Psi} - \Psi\| = \left(\sum_{i=1}^n \sum_{j=1}^p (\hat{\psi}_{ij} - \psi_{ij})^2 \right)^{\frac{1}{2}}$ şeklinde hesaplanmaktadır. Yakınsama kriteri olarak literatürde sıklıkla 10^{-6} değeri kullanılmaktadır. Bu çalışmada da yakınsama kriteri 10^{-6} olarak seçilmiş ve simülasyonlar 200 tekrarlı olarak gerçekleştirilmiştir. Ayrıca yakınsamanın kaç adımda sağlandığını görebilmek için ortalama iterasyon sayısı hesaplanmıştır. Örneklem hacmi $l = 30, 40, 60, 80$ ve 120 olarak alınmıştır. Simülasyon sonuçları Tablo 1-5'te verilmiştir. Tablolarda tahmin edilen M, Σ ve Ψ matrisleri ve bu matrislere ilişkin Öklid uzaklıkları verilmiştir.

Tablo 1. n=30 için tahmin edilen matrisler, Öklid uzaklıkları ve ortalama iterasyon sayısı

		Ortalama iterasyon sayısı
\hat{M}	$\begin{bmatrix} 0.9871 & -0.0010 & -0.0253 & 1.9978 \\ -0.0046 & 0.5068 & -0.5154 & -0.0011 \\ -1.0203 & 0.9786 & 1.9783 & 0.9805 \end{bmatrix}$	28.35
$\hat{\Sigma}$	$\begin{bmatrix} 0.9320 & 0.5618 & 0.2712 \\ 0.5618 & 0.9459 & 0.1899 \\ 0.2712 & 0.1899 & 0.9321 \end{bmatrix}$	
$\hat{\Psi}$	$\begin{bmatrix} 0.8711 & 0.4288 & 0.4309 & 0.0768 \\ 0.4288 & 0.8564 & -0.0040 & 0.5073 \\ 0.4309 & -0.0040 & 0.8559 & 0.3353 \\ 0.0768 & 0.5073 & 0.3353 & 0.8496 \end{bmatrix}$	
$\ \hat{M} - M\ $	0.6091	
$\ \hat{\Sigma} - \Sigma\ $	0.5661	
$\ \hat{\Psi} - \Psi\ $	1.0153	

Tablo 2. n=40 için tahmin edilen matrisler, Öklid uzaklıkları ve ortalama iterasyon sayısı

		Ortalama iterasyon sayısı
\hat{M}	$\begin{bmatrix} 0.9998 & 0.0122 & 0.0014 & 2.0058 \\ -0.0120 & 0.4966 & -0.5056 & -0.0015 \\ -0.9939 & 1.0093 & 2.0087 & 1.0154 \end{bmatrix}$	25.43
$\hat{\Sigma}$	$\begin{bmatrix} 0.9559 & 0.5643 & 0.08891 \\ 0.5643 & 0.9449 & 0.43941 \\ 0.2783 & 0.1791 & 0.9353 \end{bmatrix}$	
$\hat{\Psi}$	$\begin{bmatrix} 0.8790 & 0.4372 & 0.4409 & 0.0872 \\ 0.4372 & 0.8697 & 0.0068 & 0.5277 \\ 0.4409 & 0.0068 & 0.8697 & 0.3480 \\ 0.0872 & 0.5277 & 0.3480 & 0.8787 \end{bmatrix}$	
$\ \hat{M} - M\ $	0.5599	
$\ \hat{\Sigma} - \Sigma\ $	0.5492	
$\ \hat{\Psi} - \Psi\ $	0.9884	

Tablo 3. n=60 için tahmin edilen matrisler, Öklid uzaklıkları ve ortalama iterasyon sayısı

		Ortalama iterasyon sayısı
\hat{M}	$\begin{bmatrix} 0.9851 & 0.0006 & -0.0106 & 1.9985 \\ -0.0111 & 0.4928 & -0.5049 & -0.0091 \\ -0.9933 & 1.0035 & 2.0020 & 0.9993 \end{bmatrix}$	22.31
$\hat{\Sigma}$	$\begin{bmatrix} 0.9426 & 0.5597 & 0.2804 \\ 0.5597 & 0.9392 & 0.1887 \\ 0.2804 & 0.1887 & 0.9469 \end{bmatrix}$	
$\hat{\Psi}$	$\begin{bmatrix} 0.8647 & 0.4336 & 0.4342 & 0.0880 \\ 0.4336 & 0.8673 & -0.0020 & 0.5184 \\ 0.4342 & -0.0020 & 0.8699 & 0.3471 \\ 0.0880 & 0.5184 & 0.3471 & 0.8673 \end{bmatrix}$	
$\ \hat{M} - M\ $	0.4360	
$\ \hat{\Sigma} - \Sigma\ $	0.5117	
$\ \hat{\Psi} - \Psi\ $	0.9025	

Tablo 4. n=80 için tahmin edilen matrisler, Öklid uzaklıkları ve ortalama iterasyon sayısı

		Ortalama iterasyon sayısı
\hat{M}	$\begin{bmatrix} 0.9944 & -0.0124 & 0.0103 & 2.0059 \\ -0.0120 & 0.4902 & -0.5035 & -0.0030 \\ -0.9896 & 0.9946 & 2.0181 & 1.0011 \end{bmatrix}$	21.02
$\hat{\Sigma}$	$\begin{bmatrix} 0.9411 & 0.5700 & 0.2775 \\ 0.5700 & 0.9497 & 0.1839 \\ 0.2775 & 0.1839 & 0.9441 \end{bmatrix}$	
$\hat{\Psi}$	$\begin{bmatrix} 0.8756 & 0.4336 & 0.4376 & 0.0839 \\ 0.4336 & 0.8696 & -0.0018 & 0.5229 \\ 0.4376 & -0.0018 & 0.8678 & 0.3430 \\ 0.0839 & 0.5229 & 0.3430 & 0.8702 \end{bmatrix}$	
$\ \hat{M} - M\ $	0.3629	
$\ \hat{\Sigma} - \Sigma\ $	0.4782	
$\ \hat{\Psi} - \Psi\ $	0.8745	

Tablo 5. n=120 için tahmin edilen matrisler, Öklid uzaklıkları ve ortalama iterasyon sayısı

		Ortalama iterasyon sayısı
\hat{M}	$\begin{bmatrix} 0.9999 & -0.0031 & -0.0032 & 1.9922 \\ -0.0018 & 0.4965 & -0.5071 & -0.0078 \\ -1.0017 & 0.9962 & 2.0018 & 0.9996 \end{bmatrix}$	19.36
$\hat{\Sigma}$	$\begin{bmatrix} 0.9341 & 0.5610 & 0.2827 \\ 0.5610 & 0.9390 & 0.1852 \\ 0.2827 & 0.1852 & 0.9363 \end{bmatrix}$	
$\hat{\Psi}$	$\begin{bmatrix} 0.8575 & 0.4230 & 0.4279 & 0.0768 \\ 0.4230 & 0.8565 & -0.0060 & 0.5139 \\ 0.4279 & -0.0060 & 0.8568 & 0.3382 \\ 0.0768 & 0.5139 & 0.3382 & 0.8632 \end{bmatrix}$	
$\ \hat{M} - M\ $	0.3030	
$\ \hat{\Sigma} - \Sigma\ $	0.4517	
$\ \hat{\Psi} - \Psi\ $	0.8301	

Tablo 1-5 incelendiğinde örneklem hacmi arttıkça bütün parametreler için Öklid uzaklıklarının giderek azaldığı görülmektedir. Benzer olarak ortalama iterasyon sayısı da azalmaktadır. Bu durum tahmin edicilerin tutarlı tahmin ediciler olduğunu göstermektedir. Bu nedenle parametre tahmini için verilen algoritmanın iyi çalıştığı görülmektedir.

6. Sonuç ve öneriler

Bu çalışmada matris değişkenli Laplace dağılımı ve parametre tahmini için EM algoritması açıklanmıştır. Önerilen dağılım için EM algoritmasının tutarlı sonuçlar verdiği gösterilmiştir. Bu dağılım matris değişkenli verilerin analizinde normal dağılıma alternatif olarak kullanılabilir.

Kaynaklar

- [1] P. S. Laplace, 1774, Mémoire sur la probabilité des causes par les évènements, *Mémoires de Mathématique et de Physique*, 6, 621-656. English translation by S. M. Stigler *Memoir on the Probability of the Causes of Events* in *Statistical Science*, 1(3), 364-378, 1986.
- [2] G. Ulrich, C. Chen, 1987, *A bivariate double exponential distribution and its generalization*, *ASA Proceedings on Statistical Computing*, 127-129.
- [3] D. N. Anderson, 1992, *A multivariate linnik distribution*, *Statistics and Probability Letters*, 14, 333-336.
- [4] T. Eltoft, 2006, *On the multivariate Laplace distribution*, *IEEE Signal Processing Letters*, 13(5), 300-303.
- [5] D. F. Andrews, C. L. Mallows, 1974, *Scale mixtures of normal distributions*, *Journal of the Royal Statistical Society: Series B (Methodological)*, 36(1), 99-102.
- [6] S. Kotz, T. J. Kozubowski, K. Podgorski, 2001, *The Laplace Distribution and Generalizations: A Revisit with Applications to Communications, Economics, Engineering and Finance*.
- [7] E. G. Sánchez-Manzano, M. A. Gómez-Villegas, J. Marín-Diazaraque, 2002, *A matrix variate generalization of the power exponential family of distributions*, *Communications in Statistics-Theory and Methods*, 31(12), 2167-2182.
- [8] M. P. B. Gallagher, P. D. McNicholas, 2017, *Three skewed matrix variate distributions*, arXiv:1704.02531v4a.
- [9] A. K. Gupta, D. K. Nagar, 1999, *Matrix variate distributions*, Chapman and Hall/CRC, Boca Raton.
- [10] A. P. Dempster, N. M. Laird, D. B. Rubin, 1977, *Maximum likelihood from incomplete data via the EM algorithm*, *Journal of the Royal Statistical Society Series B (Methodological)*, 39(1), 1-38.
- [11] O. E. Barndorff-Nielsen, 1997, *Normal inverse gaussian distributions and stochastic volatility modelling*, *Scandinavian Journal of Statistics*, 24(1), 1-13.