İktisadi İdari ve Siyasal Araştırmalar Dergisi | Journal of Economics Business and Political Researches
Yıl:2018, 3(6):158-164 | Year:2018, 3(6):158-164
Makale Türü: Araştırma Makalesi | Paper Type: Research Paper

TEKİL DEĞER AYRIŞIMININ VARLIĞINA İSPAT*

♦♦♦

PROVING THE EXISTENCE OF A SINGULAR VALUE DECOMPOSITION

DOI: 10.25204/iktisad.405533

Zhaoyang LI**

Öz

Matris ayrışımı, karmaşık bir matrisi daha basit matrislerin çarpımına dönüştüren bir yöntemdir. 1960'lı yıllardan önce, sadece lineer sistem analizine uygulanmış olan matris ayrışımı; son yıllarda yazılım, elektronik, sinyal filtrelemesi, matris transformasyonu ve regresyon analizi gibi alanlarda da kullanılmaktadır. Özellikle Tekil değer Ayrışımı (Singular Value Decomposition - SVD), ortonormal bir matris, köşegen bir matris ve ortonormal bir matris olmak üzerine üçlü bir çarpıma ayrıştıran bir algoritmadır. Çalışmada SVD'nin ispatı sunulmaktadır.

***Anahtar Kelimeler**: Tekil değer, ayrışım varlığı, ispat.*

***Jel Kodları**: C01, C49.*

Abstract

Matrix factorization is the algorithm factorizing a matrix into the product of several matrices with particular properties. Before 1960's, matrix factorization was only used in the linear system analysis, but in the last few decades the quickly developed algorithms of matrix factorizations have been applied to solve a variety of problems, like the regression analysis and information technologies. In this thesis, the theoretical derivation of SVD is presented. And all discussions in this work are confined to the real number realm.

***Keywords**: Singular value, decomposition, existence, proving.*

***Jel Codes**: C01, C49.*

* Bu çalışma, İstanbul Üniversitesi, Sosyal Bilimler Enstitüsü, Ekonometri Anabilim Dalında, 2006 yılında Yrd. Doç. Dr. Enis SINIKSARAN danışmanlığında tamamlanan "Matris Ayrışımı" başlıklı yüksek lisans tezinden türetilmiştir.

** Prof. Dr., Shanghai University of International Business and Economics, e-mail: leezhya@163.com
ORCID: 0000-0002-3076-4314

Geliş Tarihi / Received: 06.01.2018
Kabul Tarihi / Accepted: 05.06.2018

1. GİRİŞ

Matrisin ayrışımında en çok kullanılan algoritmalar SVD, QR, LU, CS ayrışımı, QLP ayrışımı, Pivoted QR ve Pivoted QLP ayrışımı gibi yöntemlerdir (Ansley ve Graig F., 1985; Hern ve Thomas, 1993; Stewart G.W., 1993; Watkins ve David.S., 1982). Özellikle matrisin köşegenleştirilmesi ile sıkı bir ilişkisi olan SVD, tüm matris ayrışım algoritmalarının başında gelerek, özellikle regresyon analizinde önemli bir yer tutmaktadır. Bu nedenle SVD'nin varlığının incelenmesi gerekmektedir.

Örneğin, $X = \begin{bmatrix} 3.01 & 0.01 & -2.99 \\ 2.99 & -0.01 & -3.01 \\ 2.00 & -4.00 & 2.00 \end{bmatrix}$ olsun.

X Matrisi aşağıdaki gibi ayrıştırılabilir:

$$X = \begin{bmatrix} 3 & 0 & -3 \\ 3 & 0 & -3 \\ 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 2 & -4 & 2 \end{bmatrix} + \begin{bmatrix} 0.01 & 0.01 & 0.01 \\ -0.01 & -0.01 & -0.01 \\ 0 & 0 & 0 \end{bmatrix} =$$

$$6\begin{bmatrix} 1/2 & 0 & -1/2 \\ 1/2 & 0 & -1/2 \\ 0 & 0 & 0 \end{bmatrix} + 2\sqrt{6}\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1/\sqrt{6} & -2/\sqrt{6} & 1/\sqrt{6} \end{bmatrix} + \sqrt{6}/100\begin{bmatrix} 1/\sqrt{6} & 1/\sqrt{6} & 1/\sqrt{6} \\ -1/\sqrt{6} & -1/\sqrt{6} & -1/\sqrt{6} \\ 0 & 0 & 0 \end{bmatrix}$$

$$= 6\begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{bmatrix}\begin{bmatrix} 1/\sqrt{2} & 0 & -1/\sqrt{2} \end{bmatrix} + 2\sqrt{6}\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}\begin{bmatrix} 1/\sqrt{6} & -2/\sqrt{6} & 1/\sqrt{6} \end{bmatrix}$$

$$+ \sqrt{6}/100\begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \\ 0 \end{bmatrix}\begin{bmatrix} 1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} \end{bmatrix}$$

$$= \theta_1 U_1 V_1 + \theta_2 U_2 V_2 + \theta_3 U_3 V_3 = \sum_i^3 \theta_i U_i V_i^T \quad \textbf{(1)}$$

$$= \begin{bmatrix} 1/\sqrt{2} & 0 & 1/\sqrt{2} \\ 1/\sqrt{2} & 0 & -1/\sqrt{2} \\ 0 & 1 & 0 \end{bmatrix}\begin{bmatrix} 6 & 0 & 0 \\ 0 & 2\sqrt{6} & 0 \\ 0 & 0 & \sqrt{6}/100 \end{bmatrix}\begin{bmatrix} 1/\sqrt{2} & 0 & -1/\sqrt{2} \\ 1/\sqrt{6} & -2/\sqrt{6} & 1/\sqrt{6} \\ 1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} \end{bmatrix} = U\theta V^T \quad \textbf{(2)}$$

Görülüyor ki, denklem (1)'deki U_i, denklemi (2)'deki U ortonormal matrisinin sütunudur. Denklem (1)'deki V_i^T ise, (2)'deki V^T ortonormal matrisinin satırıdır. θ_i ise, θ köşegen matrisinin köşegendeki öğesidir. Denklem (1) ve (2), X matrisinin tekil değer ayrışımının (SVD) açıklanmasıdır.

2. TEKİL DEĞER AYRIŞIMININ VARLIĞINA İSPAT

Sadece yukarıdaki X değil, her türlü matris denklem, (2) gibi ayrıştırılabilir. Bu ifade, Teorem 1'de daha ayrıntılı biçimde açıklanmaktadır.

Teorem 1. m x n boyutlu **matrisi**, rankı r olan reel bir matris ise, $U^T AV = \theta =$ diag ($\theta_1, \theta_2 . . .$ $. \theta_r$, 0. . . .0) eşitliğini sağlayan ortonormal matris U ve V vardır. $U \in R^{mxm}$, $V \in R^{nxn}$, $\theta \in R^{mxn}$. Ayrıca $\theta_1 \geq \theta_2 \geq . . . . . \geq \theta_r \geq 0$ şeklindedir.

Bu teoremin ispatları Golub Gene. H. (1996) ve Stewart G. W. (1998)'de yer almaktadır. Burada, bu iki kitaptaki ispatlara benzer bir ispat sunulmaktadır.

Teorem 1'in ispatını yapmadan önce kullanacağımız vektör ve matrisin 2-norm'u ile ilgili kavramlara biraz değinelim. Bu konudaki ayrıntılar yine Golub Gene. H. (1996)'da yer almaktadır.

Teorem 2. Vektör $\chi^T = \left[\chi_1 \quad \chi_2 \quad . \quad . \quad . \quad \chi_n\right]^T$ olsun. χ Vektörünün 2-normu ise,

$\| \chi \|_2 = \sqrt{\chi_1^2 + \chi_2^2 + ... + \chi_n^2}$ 'dir (Golub Gene. H. ,1996).

Teorem 3. A matrisi $\in R^{mxn}$ olsun, $A^T A\ z = \theta^2 z$ olduğundan, 2- normu birim olan n boyutlu bir vektör z varsa, skaler θ, A'nın 2-normudur, yani $\|A\|_2 = \theta$(Golub Gene. H. ,1996).

Gerçekten A matrisinin 2- normu $\|A\|_2$, $A^T A$'nın en büyük öz değerinin kareköküdür (Golub Gene. H., 1996).

Y. Teorem 3. 1. A matrisi $\in R^{mxn}$, χ vektörü $\in R^n$ ise; $\|A\chi\|_2 \leq \|A\|_2 \|\chi\|_2$ olarak tutarlıdır (Golub Gene. H., 1996).

Y. Teorem 3. 2.

3.2.1. Q bir ortonormal matris ve χ uygun boyutlu bir vektör ise;

$\|Q\chi\|_2^2 = \chi^T Q^T Q\chi = \chi^T \chi = \|\chi\|_2^2$ dir, yani $\|Q\chi\|_2 = \|\chi\|_2$ (Golub Gene. H., 1996).

3.2.2. Q ve Z iki ortonormal matris olsun. $Q \in R^{mxm}$, $Z \in R^{nxn}$. A uygun boyutlu bir matris ise,

$\left\|Q^T AZ\right\|_2^2 = \|A\|_2^2$, yani $\left\|Q^T AZ\right\|_2 = \|A\|_2$ (Golub Gene. H., 1996).

Teorem 4. $V_1 \in R^{mxn}$ bir ortonormal matris olsun. $V = [V_1 \quad V_2]$ bir ortonormal matris olmak üzere, ortonormal matris V_2 vardır (Golub Gene. H., 1996).

Şimdi Teorem 2, 3, 4 ve Y. Teorem 3.1 ve 3.2'yi kullanarak Teorem 1'in ispatını yapalım. İspat sürecinde kolaylık sağlaması için, bu ispatta kullandığımız A örnek matrisi, rankı tam olan bir matris olsun.

İspat: m x n boyutlu A matrisinin rankı: rank (A) = min(m,n); $\|A\|_2 = \theta_1$ ve 2-norm birim vektör $\chi \in R^n$, $y \in R^m$ olsun. χ ve y vektörünün değerlerini ayarlayarak aşağıdaki denklemi kurabiliriz:

$A\chi = \theta_1 y$. Teorem 4.'ten biliyoruz ki χ 'le birlikte n x n boyutlu ortonormal matris V 'yi oluşturabilen ortonormal matris V_1 vardır. Yani $V = [\chi \quad V_1]$ 'dir.

Keza y ile birlikte m x m boyutlu ortonormal matris U 'yu oluşturabilen ortonormal matris U_1'de vardır. Yani $U = [y \quad U_1]$'dir. Dolayısı ile;

$$U^T A V = \begin{bmatrix} y^T \\ U_1^T \end{bmatrix} A \; [\chi \quad V_1]$$

$$= \begin{bmatrix} y^T A\chi & y^T AV_1 \\ U_1^T A\chi & U_1^T AV_1 \end{bmatrix}$$

$$= \begin{bmatrix} y^T \theta_1 y & w^T \\ U_1^T \theta_1 y & B_1 \end{bmatrix}$$

$$= \begin{bmatrix} \theta_1 & w^T \\ 0 & B_1 \end{bmatrix}$$

$$= \tilde{A} \tag{3}$$

Denklem (3)'ten $y^T AV_1 = w^T \in R^{1x(n-1)}$; $U_1^T AV_1 = B_1 \in R^{(m-1)\,x\,(n-1)}$ dir. Dolayısı ile;

$$\left\| \tilde{A} \begin{bmatrix} \theta_1 \\ w \end{bmatrix} \right\|_2^2 = \left\| \begin{bmatrix} \theta_1 & w^T \\ 0 & B_1 \end{bmatrix} \begin{bmatrix} \theta_1 \\ w \end{bmatrix} \right\|_2^2 = \left\| \begin{bmatrix} \theta_1^2 + w^T w \\ B_1 w \end{bmatrix} \right\|_2^2$$

$$= (\theta_1^2 + w^T W)^2 + (B_1 W)^T B_1 W \geq (\theta_1^2 + w^T W)^2 \tag{4}$$

Y. Teorem 3.1. ve denklem (4)'ten biliyoruz ki;

$$\|\tilde{A}\|_2^2 \left\| \begin{bmatrix} \theta_1 \\ w \end{bmatrix} \right\|_2^2 \geq \left\| \tilde{A} \begin{bmatrix} \theta_1 \\ w \end{bmatrix} \right\|_2^2 \geq (\theta_1^2 + w^T W)^2 \tag{5}$$

$\left\| \begin{bmatrix} \theta_1 \\ w \end{bmatrix} \right\|_2^2 = \theta_1^2 + w^T W$ olduğu için eşitsizlik (5) aşağıdaki gibi yazılabilir:

$$\|\tilde{A}\|_2^2 \geq \theta_1^2 + w^T W \tag{6}$$

Y. Teorem 3.2.2. ve denklem (3)'e dayanarak denklem (7)'yi elde ederiz.

$$\|A\|_2^2 = \|U^T A V\|_2^2 = \|\tilde{A}\|_2^2 \tag{7}$$

Denklem (6) ve denklem (7)'den denklem (8)'i elde ederiz.

$$\|A\|_2^2 = \|\tilde{A}\|_2^2 \geq \theta_1^2 + w^T w \qquad \textbf{(8)}$$

İspatın başlangıcında $\|A\|_2^2 = \theta_1^2$ diye varsaydık. Bu varsayımı temel alarak, denklem (8)'i elde ederiz. Bu nedenle $W = 0$ olmalıdır.

Dolayısı ile denklem (3) şu şekilde yazılabilir:

$$U^T A V = \begin{bmatrix} \theta_1 & 0 \\ 0 & B_1 \end{bmatrix} \qquad \textbf{(9)}$$

Gerçekten V_1'le V'yi oluşturan χ ve U_1'le U'yu oluşturan y vektörü, ayrı ayrı A matrisinin tekil değeri θ_1'e karşılık gelen sağ tekil vektörü (right singular vector) ve sol tekil vektörüdür (left singular vector). θ_1 hem A'nın hem $U^T A V$'nin 2–normudur, yani $\theta_1 = \|A\|_2$ = $\|U^T A V\|_2$'dir, buradaki U Є $R^{m \times m}$ ve V Є $R^{n \times n}$'dir. Ayrıca $B_1 = U_1^T A V_1$'in 2– normu $\|B_1\|_2$ = $\|U_1^T A V_1\|_2 = \theta_2$'dir. Buradaki U_1 Є $R^{mx(m-1)}$ ve V_1 Є $R^{nx(n-1)}$, dolayısı ile $\|B_1\|_2 = \|U_1^T A V_1\|_2 \leq \|U^T A V\|_2$ = $\|A\|_2$'dir, yani $\theta_2 \leq \theta_1$'dir.

Denklem (9)'a benzer şekilde B_1'de;

$$B_1 = U_1^T A V_1 = \begin{bmatrix} \theta_2 & 0 \\ 0 & B_2 \end{bmatrix} \qquad \textbf{(10)}$$

şeklinde açıklanabilir. Burada B_2 Є $R^{(m-2)x(n-2)}$'dir.

Denklem (10) 'u denklem (9)'a dönüştürdükten sonra, denklem (11)'i elde ederiz.

$$U^T A V = \begin{bmatrix} \theta_1 & 0 & 0 \\ 0 & \theta_2 & 0 \\ 0 & 0 & B_2 \end{bmatrix} \qquad \textbf{(11)}$$

Aynı mantıkla denklem (11)'deki B_2**'**de bir blok köşegen matrisine dönüştürülebilir. Böyle devam ederek son olarak B_{r-1} köşegen matrisini elde ederiz. B_{r-1} = diag (θ_r , 0 . . . . . 0)'dir. r ise, A matrisinin rankıdır.

Dolayısı ile denklem (11) şu şekilde yeniden yazılabilir,

$$U^T A V = \left[\begin{array}{ccccc|c} \theta_1 & & & & & \\ & \theta_2 & & \phi & & \\ & & \cdot & & & \phi \\ & \phi & & \cdot & & \\ & & & & \theta_r & \\ \hline & & \phi & & & \phi \end{array}\right] = \text{diag}\,(\theta_1, \theta_2 \,.\,.\,.\,.\,.\,.\, \theta_r, 0\,.\,.\,.\,0) \quad \textbf{(12)}$$

Denklem (12) ise, A matrisinin tekil değer ayrıştırılmasının (SVD) bir açıklanmasıdır. Ayrıca $\theta_1 \geq \theta_2 \geq \,.\,.\,.\,.\,.\, \geq \theta_r \geq 0$ dır.

Burada U ve V rankı tam olan kare ortonormal matris olduğu için denklem (12) şu şekilde de yazılabilir:

$$A = U\theta V^T$$

3. SONUÇ

U, A matrisinin sol tekil vektör matrisidir. Yani U'nun sütunları, A matrisinin sol tekil vektörleridir. Ayrıca U matrisi, AA^T'nin öz vektör matrisidir (Golub Gene. H, 1996). V ise, A matrisinin sağ tekil vektör matrisidir. Yani V'nin sütunları, A matrisin sağ tekil vektörleridir. Aynı zamanda, V de $A^T A$'nın öz vektör matrisidir (Golub Gene. H, 1996).

θ ise, A matrisinin tekil değer matrisidir. Aynı zamanda, θ_i^2 $(0 \leq i \leq r)$'nin terim olarak oluşturduğu köşegen matris, hem matris AA^T hem de $A^T A$'nın öz değer matrisidir. Yani simetrik matris, AA^T'ninki ile simetrik matris $A^T A$'nın sıfır olmayan öz değerleri aynıdır.

Teorik olarak, A'nın sol tekil vektör matrisi U ile sağ tekil vektör matrisi V, AA^T'nin öz vektör matrisi ve $A^T A$'nın öz vektör matrisi olması ile hesaplanabilirler. Keza, θ ise, AA^T veya $A^T A$'nın sıfır olmayan öz değerleri olması ile hesaplanabilir. Fakat pratikte böyle yaparak bile doğru sonuçlar her zaman kesin olarak ortaya çıkaramayız.

U, V ve θ'nın hesaplanmaları, yani A'nın SVD'nin pratik hesaplanması Golub Gene. H. (1996)'nin 8'inci bölümünde yer almaktadır. Ama SVD'nin hesaplanmasına yardımcı olan bazı bilgisayar programları mevcut olduğu için (Stewart G.W., 1998) burada SVD'nin hesaplanmasındaki algoritmaların üzerinde durmayacağız.

KAYNAKÇA

Ansley, G. F. (1985). Quick Prof of Some Regression Theorem via the QR Algorithm, the American Statistician, February 1985, 39(1):55-56.

Bipschuts, S. (1991). Lineer Cebir, 2 Baskı Türkçe Çevrisi, Nobel Yayın Dağıtım, Ankara.

Eubank, R. L. and Webster J. T. (1985). The Singular Value Decomposition as a Tool for Solving Estimability Problems, the American Statistician, February 39(1):64-72.

Golub, G. H. (1996). Matrix Computation, Third Edition, The Johns Hopkins University Pres, Baltimore, p.70.

Hern, T. (1993). Gaussin Elimination in Integer Arithmetic: An Application of the LU Factorization, The College Mathematics Journal, 1993, 24(1):67-71.

Stewart G.W. (1993). On the Early History of the Singular Value Decompostion, SIAM Review, 35(4):551-566.

Stewart, G. W. (1998). Matrix Algorithms, Volume I: Basic Decompositions, SIAM, 1998, p.62.

Watkins, D. S. (1982). Understanding the QR Algorithm, SIAM Review, 24(4):427-440.