T.C

KIRIKKALE ÜNİVERSİTESİ

FEN BİLİMLERİ ENSTİTÜSÜ

MATEMATİK ANABİLİM DALI

YÜKSEK LİSANS TEZİ

URV AYRIŞIMI VE UYGULAMALARI

KÜBRA ABA

HAZİRAN 2008

Fen Bilimleri Enstitü Müdürünün onayı.

Doç. Dr. Burak BİRGÖREN

…./…./……

Müdür V.

Bu tezin Yüksek Lisans tezi olarak Matematik Anabilim Dalı standartlarına uygun olduğunu onaylarım.

Prof. Dr. Kerim KOCA

Anabilim Dalı Başkanı

Bu tezi okuduğumuzu ve Yüksek Lisans tezi olarak bütün gerekliliklerini yerine getirdiğini onaylarız.

Yrd. Doç. Dr. Sevgi YURT ÖNCEL — Ortak Danışman

Doç. Dr. Hasan ERBAY — Danışman

Jüri Üyeleri

Yrd. Doç. Dr. Ali ARAL

Prof. Dr. Kerim KOCA

Doç. Dr. Hasan ERBAY

Yrd. Doç. Dr. Fatih TANK

Yrd. Doç. Dr. Sevgi YURT ÖNCEL

ÖZET

URV AYRIŞIMI VE UYGULAMALARI

ABA, Kübra

Kırıkkale Üniversitesi

Fen Bilimleri Enstitüsü

Matematik Anabilim Dalı, Yüksek Lisans Tezi

Danışman : Doç. Dr. Hasan ERBAY

Ortak Danışman : Yrd. Doç. Dr. Sevgi YURT ÖNCEL

Haziran 2008, 56 sayfa

Bu tez, URV temel matris ayrışımı, kesik URV ayrışımı ve güncelleme algoritması hakkında bilgi vermektedir. Güncellemenin matematiksel hesaplama karmaşası, r-ranklı $m \times n$ tipindeki bir matris için $O(nr)$ işlemdir. Teorik ve sayısal sonuçlar, URV Ayrışımının Tekil Değer Ayrışımı (SVD: Singular Value Decomposition) için iyi bir alternatif olduğunu göstermektedir.

Anahtar Kelimeler: URV ayrışımı, tekil değer ayrışımı, sayısal rank, alt uzay tahmini, en küçük kareler problemi.

ABSTRACT

URV DECOMPOSITION AND ITS APPLICATIONS

ABA, Kübra

Kırıkkale University

Graduate School Of Natural and Applied Sciences

Deparment of Mathematics, M. Sc. Thesis

Supervisor : Assoc. Prof. Dr. Hasan ERBAY

Co-Supervisor : Asst. Prof. Dr. Sevgi YURT ÖNCEL

Jun 2008, 56 pages

This thesis, presents an URV-based matrix decomposition, the truncated URV decomposition and an updating algorithm for it. The computational complexity of the updating is $O(nr)$ for an m-by-n matrix of rank r. The theoretical and numerical results presented shows that the decomposition can be a good alternative to the singular value decomposition.

Key Words: URV decomposition, singular value decomposition, numerical rank, subspace estimation, least square problems.

TEŞEKKÜR

Çalışmamın her aşamasında, bilgi, öneri ve yardımları ile beni yönlendiren, desteğini her zaman hissettiğim danışman hocam Sayın Doç. Dr. Hasan ERBAY'a,

Her zaman bana verdiği moral ve destekten dolayı Sayın Yrd. Doç. Dr. Sevgi YURT ÖNCEL'e,

Her konuda yardımını ve desteğini gördüğüm Sayın Yrd. Doç. Dr. Fatih TANK'a,

Tezimi okuyup değerlendiren ve önerilerini sunan Sayın jüri üyeleri hocalarıma,

Çalışmam süresince gösterdikleri anlayış, destek ve yardımları için bölüm arkadaşlarım Öğr. Gör. Emel KIZILOK, Öğr. Gör. Serap YÖRÜBULUT, Arş. Gör. Altan TUNÇEL ve Arş. Gör. Abdullah YILMAZ'a,

Tezimin dilbilgisi yönünden düzeltmelerini yapan ve bana her zaman moral veren ablam Sevgi ÇIKRIKÇI'ya ve manevi yönden desteklerini hiçbir zaman esirgemeyen annem, babam ve kardeşime,

Teşekkür ederim.

İÇİNDEKİLER

ÇİZELGELER DİZİNİ

ÇİZELGE

ŞEKİLLER DİZİNİ

ŞEKİL

1. GİRİŞ

Sayısal rankın hesaplanması ve temel uzayların yaklaşık olarak bulunması problemi birçok alanda uygulanmaktadır. Sinyal işleme, görüntü işleme, kontrol ve istatistik bu alanlar arasındadır. Örneğin, istatistikte üretim, ekonomi, hisse senedi vb. verilerini etkileyen dış etkilerin (sıcaklık, rüzgar, enflasyon gibi) matematiksel olarak çok iyi karakterize edilmesi gerekir. Çünkü bu faktörler üretimin düşmesine veya ekonomik büyümenin azalmasına neden olabilirler. Dolayısıyla bu tür etkileri azaltmak için önemli olan, verileri en iyi şekilde açıklayan lineer bir model oluşturabilmektir. Söz konusu modelin çözümünde, en küçük kareler (EKK) ve tam kareler (TEKK) metodları sıklıkla kullanılan yöntemlerdir. Bu yöntemlerde çözüm, genelde tekil değerler ve tekil uzaylara bağlı verilir. Çözümü elde etmedeki genel yaklaşım, tekil değer ayrışımını (SVD) hesaplamak ve tekil değerlerin küçüklüğüne göre sayısal rankın ne olacağına karar vermek ve alt uzayları oluşturmaktır. Öte yandan, tekil değer ayrışımının bir dezavantajı, yeni veri geldiğinde mevcut bilgileri yinelemenin matematiksel hesaplama karmaşıklığının yüksek olmasıdır. Bu yüzden SVD, veri akışının sürekli olduğu problemlerde uygun ayrışım değildir. Bundan dolayı sayısal rankı hesaplamada ve değer uzayını bulmada tekil değer ayrışımı gibi etkili fakat daha hızlı alternatif ayrışımlara ihtiyaç duyulmuştur. Alternatifler arasında QR, ULV ve URV ayrışımları bulunmaktadır. Ancak, bu ayrışımların hesaplama karmaşası düşük ranklı matrisler için hala yüksektir. Öte yandan URV ayrışımı değer uzayını hesaplamada alternatiflerine oranla daha iyi performans gösterdiği bilinmektedir. Bu tezin ise temel konusu URV tabanlı matris ayrışımıdır.

1.1 Tezin Amacı

Tez boyunca veri matrislerinin sayısal rankını hesaplamada ve alt uzayları yaklaşık olarak hesaplamada kullanılan URV ayrışımının SVD'ye tercih edilme nedenleri ortaya konmaya çalışılmaktadır. Buna bağlı olarak düşük ranklı matrislerde daha iyi performans gösteren URV tabanlı matris ayrışımını tanımlamak ve bu ayrışıma ait teorik sonuçlar verilmektedir.

Veri matrisine yeni veri eklendiğinde mevcut bilgiyi kullanarak URV tabanlı matris ayrışımını hesaplayan etkili ve hızlı bir algoritma geliştirmek tezin amaçları arasındadır. Tanımlanan bu ayrışımın, veri akışının sürekli olduğu özyinelemeli problemlerde sayısal rankı hesaplamada ve temel uzayları takip etmede kullanılabileceğini ve SVD'ye iyi bir alternatif olacağını ortaya koymak amaçlanmıştır. Elde edilen ayrışım yardımıyla bir istatistiksel EKK problemi çözmeye çalışmak tezin diğer bir amacıdır.

1.2 Kaynak Özetleri

Çalışmanın genelinde kullanılacak olan lineer cebir ve nümerik analiz temel kavramları için Gene H. Golub ve Charles F. Van Loan(1)'ın "Matrix Computations", Ake Björck(2)'un "Numerical Methods for Least Squares Problems" ve David S. Watkins(3)'in "Fundamentals of Matrix Computations" kitaplarından yararlanılmıştır.

G. W. Stewart(4), ULV ayrışımıyla ilgili bir algoritma önermiştir. Ayrıca ayrışımı daha etkili kılan deflasyon algoritması verilmiştir. J. L. Barlow ve arkadaşları(5), görüntü modelleme için tam en küçük kareler (TEKK) probleminden bahsetmiştir. H. Erbay[6], kesik ULV ayrışımı için temel uzaylar belirlemiş ve düşük ranklı matrisler için güncelleme algoritması önermiştir.

G. Adams ve arkadaşları[7], URV ayrışımı için temel uzaylar belirlemiş ve URV'nin güncellenme sürecinden bahsetmiştir. Ayrıca S. Hosur ve arkadaşları[8] ise zayıf koşullu matrisler için EKK algoritması önermişlerdir.

2. MATERYAL VE YÖNTEM

Bu kısımda, ilerideki bölümlerde kullanılacak lineer cebir ve nümerik analiz temel kavramları tanımlanacaktır.

2.1 Lineer Cebir Temel Kavramları

Tanım 2.1.1 (Vektör Uzayı) K verilen bir alan ve V toplama ve skalerle çarpma kuralları ile boş olmayan bir cümle olsun. Bu durumda herhangi bir $\mathbf{u},\mathbf{v}\in V$ 'ye bir $\mathbf{u}+\mathbf{v}\in V$ toplamı ve herhangi bir $\mathbf{u}\in V$, $k\in K$ 'ya bir $k\mathbf{u}\in V$ çarpımı karşılık gelir. Eğer aşağıdaki aksiyomlar mevcut ise V 'ye, K üzerinde tanımlı bir *vektör uzayı* denir. V 'nin elemanları da *vektör*ler olarak adlandırılır[9].

i) $\forall \mathbf{u},\mathbf{v},\mathbf{w}\in V$ vektörleri için $(\mathbf{u}+\mathbf{v})+\mathbf{w}=\mathbf{u}+(\mathbf{v}+\mathbf{w})$,

ii) $\forall \mathbf{u}\in V$ için $\mathbf{u}+\mathbf{0}=\mathbf{0}+\mathbf{u}=\mathbf{u}$ olacak şekilde $\mathbf{0}\in V$ etkisiz elemanı vardır,

iii) $\forall \mathbf{u}\in V$ için tek bir $-\mathbf{u}\in V$ elemanı vardır, öyle ki $\mathbf{u}+(-\mathbf{u})=\mathbf{0}$,

iv) $\forall \mathbf{u},\mathbf{v}\in V$ vektörleri için $\mathbf{u}+\mathbf{v}=\mathbf{v}+\mathbf{u}$,

v) $\forall k\in K$ skaleri ve $\forall \mathbf{u},\mathbf{v}\in V$ vektörleri için $k(\mathbf{u}+\mathbf{v})=k\mathbf{u}+k\mathbf{v}$,

vi) $\forall a,b\in K$ skalerleri ve $\forall \mathbf{u}\in V$ vektörü için $(a+b)\mathbf{u}=a\mathbf{u}+b\mathbf{u}$,

vii) $\forall a,b\in K$ skalerleri ve $\forall \mathbf{u}\in V$ vektörü için $(ab)\mathbf{u}=a(b\mathbf{u})$,

viii) $1 \in K$ birim skaleri ve $\forall \mathbf{u} \in V$ vektörü için $1\mathbf{u} = \mathbf{u}1 = \mathbf{u}$.

Tanım 2.1.2 (Vektör) $\mathbb{R}^n$ reel n-*vektör* uzayı olsun. O zaman $\mathbf{x} \in \mathbb{R}^n$ aldığımızda

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad x_i \in \mathbb{R}, \quad i = 1, 2, \ldots, n.. \tag{2.1}$$

Burada x_i, $\mathbf{x}$ vektörünün i-inci bileşeni olarak anılır.

$\mathbb{R}^n$ ile $\mathbb{R}^{n\times 1}$ özdeştir, $\mathbb{R}^n$ nin elemanları sütun vektörleri ve $\mathbb{R}^{1\times n}$ nin elemanları da satır vektörleridir ve

$$\mathbf{y} \in \mathbb{R}^{1\times n} \Leftrightarrow \mathbf{y} = \begin{pmatrix} y_1 & y_2 & \ldots & y_n \end{pmatrix}. \tag{2.2}$$

Eğer $\mathbf{x}$ sütun vektörü ise $\mathbf{y} = \mathbf{x}^T$ satır vektörüdür. Tez boyunca vektörler, kalın ve küçük harflerle ifade edilecektir[1].

Tanım 2.1.3 (Sıfır Vektörü) Tüm bileşenleri sıfır olan vektöre *sıfır vektörü* denir ve boyutundan bağımsız olarak $\mathbf{0}$ ile gösterilir[10].

Tanım 2.1.4 (Birim Vektör) i-inci bileşeni 1, diğer bileşenleri 0 olan vektöre i-inci *birim vektör* denir ve $\mathbf{e}_i$ ile gösterilir[10].

Tanım 2.1.5 (Lineer Bağımsızlık) $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n \in \mathbb{R}^n$ vektör kümesi için

$$a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \ldots + a_n\mathbf{v}_n = 0 \tag{2.3}$$

denklemi sadece $a_1 = a_2 = \ldots = a_n = 0$ olduğu zaman sağlanıyorsa, $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ vektörleri *lineer bağımsızdır* denir. V vektör uzayı $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ gibi lineer bağımsız $n-$*vektör* 'e sahipse, bu uzaya n *-boyutlu vektör uzayı* denir ve $\dim(V) = n$ ile gösterilir. $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ dizisine V vektör uzayının bir *bazı* denir ve $V = span\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ yazılır.

Tanım 2.1.6 (Vektör Normu) $\mathbb{R}^n$ de tanımlı norm (vektör normu), her bir $\mathbf{x} \in \mathbb{R}^n$ için $\mathbf{x}$'in normu olarak adlandırılan $\|\mathbf{x}\|$ reel sayısı ile eşleyen aşağıdaki özelliklere sahip bir fonksiyondur[(3)]:

i. Pozilif Tanımlılık: $\|\mathbf{x}\| > 0,\ \mathbf{x} \neq 0$ ve $\|0\| = 0$

ii. Mutlak Homojenlik: $\|\alpha\mathbf{x}\| = |\alpha|\|\mathbf{x}\|,\ \alpha \in \mathbb{R}$

iii. Üçgen Eşitsizliği: $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|,\ \mathbf{y} \in \mathbb{R}^n$

Verilen bir $\mathbf{x} \in \mathbb{R}^n$ vektörü için birkaç önemli norm aşağıda verilmiştir:

- ***1-Norm:*** $\|\mathbf{x}\|_1 = \sum_{i=1}^{n} |x_i|$ şeklinde tanımlanır.

- ***Öklit Normu:*** $\|\mathbf{x}\|_2 = \left(\sum_{i=1}^{n} |x_i|^2 \right)^{1/2}$ şeklinde tanımlanır. Sıklıkla $2-$*norm* veya spektral norm olarak da anılır.

- ***Sonsuz Norm:*** $\|\mathbf{x}\|_\infty = \max_i |x_i|$ şeklinde tanımlanır.

Tanım 2.1.7 (Matris) $\mathbb{R}^{m\times n}$ tüm $m\times n$ boyutlu reel matrislerin vektör uzayı olsun. O zaman $A\in\mathbb{R}^{m\times n}$ aldığımızda

$$A=\left(a_{ij}\right)=\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix},\quad a_{ij}\in\mathbb{R},\quad i=1,\ldots,m,\quad j=1,\ldots,n \tag{2.4}$$

biçimindedir. Burada a_{ij}'ler matrisin elemanları olarak adlandırılır.

Tez boyunca matrisler italik ve büyük harflerle ifade edilecektir. Bir matrisin i-inci satır j-inci sütununda bulunan elemanı $\left(A_{ij}\right)$ veya $A\left(i,j\right)$ şeklinde gösterilir[1].

Tanım 2.1.8 (Range Uzayı) A, $m\times n$ boyutunda bir matris olsun. A matrisinin *range uzayı*

$$range\left(A\right)=\left\{\mathbf{y}\in\mathbb{R}^m:\mathbf{y}=A\mathbf{x},\mathbf{x}\in\mathbb{R}^n\right\} \tag{2.5}$$

şeklinde tanımlanır[1].

Eğer A matrisi $A=\begin{pmatrix}\mathbf{a}_1 & \mathbf{a}_2 & \ldots & \mathbf{a}_n\end{pmatrix}$ şeklinde sütunlara parçalanırsa

$$range\left(A\right)=span\left\{\mathbf{a}_1,\mathbf{a}_2,\ldots\mathbf{a}_n\right\} \tag{2.6}$$

dır[1].

Tanım 2.1.9 (Matris Rankı) $A\in\mathbb{R}^{m\times n}$ matrisinin rankı

$$rank\left(A\right)=\dim\left(range\left(A\right)\right) \tag{2.7}$$

şeklinde tanımlanır. Eğer $rank(A) = \min\{m,n\}$ ise A matrisine *tam dereceli* denir[(10)].

Tanım 2.1.10 (Sıfır Uzayı) $A \in \mathbb{R}^{m \times n}$ matrisinin *sıfır uzayı*

$$null(A) = \{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x} = 0\} \tag{2.8}$$

şeklinde tanımlanır[(1)].

Tanım 2.1.11 (Kare Matris) Satır ve sütun sayısı eşit olan matrise *kare matris* denir[(9)].

Tanım 2.1.12 (Birim Matris) I bir kare matris olsun. Eğer

$$I(i,j) = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases} \tag{2.9}$$

oluyorsa I matrisine *birim matris* denir[(9)].

Tanım 2.1.13 A, kare ve tam ranklı bir matris ise A *tekil olmayan matris* olarak adlandırılır.

Tanım 2.1.14 (Matrisin Tersi) A tekil olmayan bir matris olsun.

$$A^{-1}A = AA^{-1} = I \tag{2.10}$$

olacak şekilde bir A^{-1} matrisi varsa, bu matrise A *matrisinin tersi* denir[(11)].

Tanım 2.1.15 (Matris Normu) $\mathbb{R}^{m \times n}$ de tanımlı bir matris normu, her bir $A \in \mathbb{R}^{m \times n}$ için A'nın normu olarak adlandırılan $\|A\|$ reel sayısına eşleyen aşağıdaki özelliklere sahip bir fonksiyondur[(3)]:

i. Pozilif Tanımlılık: $\|A\| > 0,\ A \neq 0$

ii. Mutlak Homojenlik: $\|\alpha A\| = |\alpha| \|A\|$

iii. Üçgen Eşitsizliği: $\|A + B\| \leq \|A\| + \|B\|$.

Ayrıca,

iv. Tutarlılık: $\|AB\| \leq \|A\| \|B\|$.

Son özellik tüm matris normları tarafından sağlanmayabilir, fakat aşağıda belirttiğimiz matris normları bu özelliğe sahiptir.

Verilen bir $A \in \mathbb{R}^{m \times n}$ matrisi için birkaç önemli norm aşağıda verilmiştir:

- ***1-Norm:*** $\|A\|_1 = \max_{1 \leq j \leq n} \sum_{i=1}^{m} |a_{ij}|$ şeklinde tanımlanır ve *sütun normu* da denir.

- ***Öklit Normu:*** $\|A\|_2 = \left(\sum_{i,j=1}^{n} |a_{ij}|^2 \right)^{1/2}$ şeklinde tanımlanır ve $2-norm$ da denir.

- ***Frobenius Norm:*** $\|A\|_F = \left(\sum_{i=1}^{m} \sum_{j=1}^{n} |a_{ij}|^2 \right)^{1/2}$ şeklinde tanımlanır.

- ***Sonsuz Norm:*** $\|A\|_\infty = \max_{1 \leq i \leq m} \sum_{j=1}^{n} |a_{ij}|$ şeklinde tanımlanır ve *satır normu* da denir.

Tanım 2.1.16 (Ortogonal Matris) $Q \in \mathbb{R}^{n \times n}$ matrisi $QQ^T = Q^T Q = I$ oluyorsa Q matrisine *ortogonal matris* denir.

Ortogonal matrisler aşağıdaki özelliklere sahiptir[(3)]:

i. $QQ^T = I$ ve $Q^T Q = I$,

ii. $Q^{-1} = Q^T$,

iii. Q ortogonal ise Q^{-1} de ortogonaldir,

iv. İki ortogonal matrisin çarpımı da ortogonaldir,

v. $\forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ vektörleri için $\langle Q\mathbf{x}, Q\mathbf{y} \rangle = \langle \mathbf{x}, \mathbf{y} \rangle$ ve $\|Q\mathbf{x}\|_2 = \|\mathbf{x}\|_2$.

Sonuncu özelliği gösterelim.

$$\|Q\mathbf{x}\|_2^2 = \langle Q\mathbf{x}, Q\mathbf{x} \rangle = (Q\mathbf{x})^T Q\mathbf{x} = \mathbf{x}^T Q^T Q\mathbf{x} = \mathbf{x}^T \mathbf{x} = \langle \mathbf{x}, \mathbf{x} \rangle = \|\mathbf{x}\|_2^2 .$$

Bu özellikten, ortogonal matrislerin normu koruduğunu söyleyebiliriz. Ortogonal dönüşümlerin bu özelliğini ileriki bölümlerde sıklıkla kullanacağız.

Eğer $Q \in \mathbb{R}^{n \times r}$ matrisi sadece $QQ^T = I$ şartını sağlıyorsa *sağdan ortogonal*, sadece $Q^T Q = I$ şartını sağlıyorsa *soldan ortogonal* olarak adlandırılır.

Tanım 2.1.17 (Üçgensel Matris) $U = (u_{ij})$ matrisinde $u_{ij} = 0,\ i > j$ ise U matrisine *üst üçgensel matris* denir ve

$$U = \begin{pmatrix} u_{11} & u_{12} & \cdots & u_{1n} \\ & u_{22} & \cdots & u_{2n} \\ & & \ddots & \vdots \\ 0 & & & u_{nn} \end{pmatrix} \tag{2.11}$$

biçiminde gösterilir. $L = (l_{ij})$ matrisinde $l_{ij} = 0,\ i < j$ ise L matrisine *alt üçgensel matris* denir ve

$$L = \begin{pmatrix} l_{11} & & & 0 \\ l_{21} & l_{22} & & \\ \vdots & \vdots & \ddots & \\ l_{n1} & l_{n2} & \cdots & l_{nn} \end{pmatrix} \tag{2.12}$$

biçiminde gösterilir[3].

2.2 Nümerik Analiz Temel Kavramları

Tanım 2.2.1 (Mutlak ve Göreli Hata) $\hat{\mathbf{x}} \in \mathbb{R}^n$, $\mathbf{x} \in \mathbb{R}^n$'in gözlem değerlerinden elde edilen tahmini olsun. Verilen bir $\|\cdots\|$ vektör normu için

$$\varepsilon_{abs} = \|\hat{\mathbf{x}} - \mathbf{x}\| \tag{2.13}$$

eşitliği *mutlak hata* olarak adlandırılır[3]. Eğer $\mathbf{x} \neq 0$ ise

$$\varepsilon_{rel} = \frac{\|\mathbf{x} - \hat{\mathbf{x}}\|}{\|\mathbf{x}\|} \tag{2.14}$$

ifadesi de *göreli hata* olarak adlandırılır[1].

Tanım 2.2.2 (Göreli Artık) $\hat{\mathbf{x}} \in \mathbb{R}^n$, $A\mathbf{x} = \mathbf{b}$ lineer denklem sisteminin hesaplanan çözümü olsun. Buna göre

$$\mathbf{r} = \mathbf{b} - A\hat{\mathbf{x}} \tag{2.15}$$

ifadesi *artık (residual)* olarak adlandırılır.

$$\frac{\|\mathbf{b} - A\hat{\mathbf{x}}\|}{\|\mathbf{b}\|} \tag{2.16}$$

ifadesine ise *göreli artık* denir[1].

Tanım 2.2.3 (Koşul Sayısı) $A \in \mathbb{R}^{n \times n}$ tersi alınabilen matris olsun. A matrisinin *koşul sayısı (condition number)*

$$Cond(A) = \|A\| \|A^{-1}\| \tag{2.17}$$

dır. Eğer $Cond(A)$ değeri çok büyük ise A matrisine *zayıf koşullu (ill-condition)* denir[2].

$Cond(A)$ matrisin normuna bağlıdır, yani matrisin normu değiştikçe $Cond(A)$ da değişir. Diğer taraftan $Cond(A)$ her zaman 1'den büyüktür. Çünkü birim matris I için $\|I\| = \max \frac{\|\mathbf{x}\|}{\|\mathbf{x}\|} = 1$ ve Tanım 2.1.16'nın son özelliğinden $\|I\| = \|AA^{-1}\| \leq \|A\| \|A^{-1}\|$ dır.

$Cond(A)$'nın hatalar ve artıklarla olan ilişkisi

$$\frac{\|\mathbf{x} - \hat{\mathbf{x}}\|}{\|\mathbf{x}\|} \leq Cond(A) \frac{\|\mathbf{b} - A\hat{\mathbf{x}}\|}{\|\mathbf{b}\|} \tag{2.18}$$

eşitsizliği ile verilir. (2.18) eşitsizliği (2.14)'deki göreli hata için bir üst sınır belirtir. $Cond(A)$ büyüdükçe, göreli hata göreli artıktan daha büyük çıkabilir, yani verideki küçük değişikliklerin sonuca etkisi büyük olabilir. Bu, sistemin zayıf koşullu sistem olmasından kaynaklanır[11].

Sonuç olarak koşul sayısının büyüklüğü elde edilecek çözümün güvenilirliği hakkında bilgi verir.

Koşul sayısının önemini bir örnek üzerinde göstermeye çalışalım:

Örnek2.1 $A=\begin{pmatrix}1000 & 999\\ 999 & 998\end{pmatrix}$ olsun. $A^{-1}=\begin{pmatrix}-998 & 999\\ 999 & -1000\end{pmatrix}$ ve $\|A\|_\infty=\|A\|_1=\|A^{-1}\|_\infty=\|A^{-1}\|_1=1999$ olduğu kolaylıkla hesaplanabilir. Buradan A matrisinin $\infty-norm$ ve $1-norm$'a göre koşul sayısı

$$Cond_\infty(A)=\|A\|_\infty\,\|A^{-1}\|_\infty\approx 3.996\times 10^6$$

$$Cond_1(A)=\|A\|_1\,\|A^{-1}\|_1\approx 3.996\times 10^6$$

olarak bulunur. $Cond(A)$ değerleri çok büyük olduğundan A matrisi zayıf koşullu matristir.

$A\mathbf{x}=\mathbf{b}$ için

$$\begin{pmatrix}1000 & 999\\ 999 & 998\end{pmatrix}\begin{pmatrix}x_1\\ x_2\end{pmatrix}=\begin{pmatrix}1999\\ 1997\end{pmatrix} \tag{2.19}$$

A ve $\mathbf{b}$ (2.19) eşitliğinde olduğu gibi verildiğinde bu sistemin çözümü $\mathbf{x}=\begin{pmatrix}1\\ 1\end{pmatrix}$ dir.

$\mathbf{b}$ vektöründe, $\delta\mathbf{b}=\begin{pmatrix}-0.01\\ 0.01\end{pmatrix}$ şeklinde yapılan çok küçük bir değişiklik sonucunda

$$\begin{pmatrix}1000 & 999\\ 999 & 998\end{pmatrix}\begin{pmatrix}x_1\\ x_2\end{pmatrix}=\begin{pmatrix}1998.99\\ 1997.01\end{pmatrix} \tag{2.20}$$

(2.20) sisteminin çözümü $\mathbf{x}=\begin{pmatrix}20.97\\-18.99\end{pmatrix}$ olarak elde edilir. $\mathbf{x}$ vektöründe $\delta\mathbf{x}=\begin{pmatrix}19.97\\-19.99\end{pmatrix}$ kadar bir değişim olmuştur. Görüldüğü üzere (2.19) ve (2.20) sistemlerinin çözümü birbirinden oldukça farklıdır. Bunun nedeni A matrisinin zayıf koşullu olmasıdır (ya da $Cond(A)$'nın çok büyük olmasındandır)[3].

Teorem 2.1.1 (Cauchy İnterlace Teoremi) A, $n\times n$ boyutlu simetrik bir matris ve B de $(n-1)\times(n-1)$ boyutlu A'nın temel alt matrisi olsun. O zaman

$$\lambda_k(A)\geq\lambda_k(B)\geq\lambda_{k+1}(A) \quad , \quad k=1,\ldots,n-1 . \tag{2.21}$$

Burada λ_k , $k=1,\ldots,n-1$ A ve B matrislerinin özdeğerleridir[12].

Cauchy İnterlace Teoremi'nin başka bir ifadesi olarak, A, $n\times n$ boyutlu simetrik bir matris ve B de $(n-1)\times(n-1)$ boyutlu A'nın temel alt matrisi olsun. A'nın özdeğerleri $\lambda_n\leq\lambda_{n-1}\leq\ldots\leq\lambda_2\leq\lambda_1$ ve B'nin özdeğerleri de $\mu_n\leq\mu_{n-1}\leq\ldots\leq\mu_3\leq\mu_2$ ise

$$\lambda_n\leq\mu_n\leq\lambda_{n-1}\leq\mu_{n-1}\leq\ldots\leq\lambda_2\leq\mu_2\leq\lambda_1 \tag{2.22}$$

sıralaması geçerlidir.

Tanım 2.2.4 (Büyük O) $f:\mathbb{R}\to\mathbb{R}$ ve $g:\mathbb{R}\to\mathbb{R}$ olmak üzere $x\to\ell$, $\ell\in\mathbb{R}\cup\{-\infty,\infty\}$ için $C\in\mathbb{R}^+$ vardır, öyle ki $\left|\frac{f(x)}{g(x)}\right|<C$ oluyorsa, bu durum $f(x)=O(g(x))$, $x\to\ell$ biçiminde gösterilir ve $f(x)$'e $g(x)$'in *büyük-O (big-*

O)'su denir. $f(x)=O(g(x))$ eşitliği, $g(x)$'in bir sabitle çarpımından daha küçük olduğu anlamına gelir[13].

Büyük O gösterimi matematiksel bir gösterim olup fonksiyonların asimptotik davranışlarını tarif etmek için kullanılır. Daha açık şekilde anlatmak gerekirse, bir fonksiyonun büyümesinin asimptotik üst sınırını daha basit başka bir fonksiyon cinsinden tanımlanması demektir. İki temel uygulama alanı vardır: Matematik alanında genellikle kırpılmış bir sonsuz terimin kalan terimini karakterize etmek için kullanılır. Bilgisayar bilimlerinde ise algoritmaların bilgi işlemsel karmaşıklığının çözümlemesi için kullanılır.

Tanım 2.2.5 (Givens Dönüşümü) Givens dönüşümü

$$G(i,k,\theta)=\begin{pmatrix} 1 & \dots & 0 & \dots & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & & \vdots & & \vdots \\ 0 & \dots & c & \dots & s & \dots & 0 \\ \vdots & & \vdots & \ddots & \vdots & & \vdots \\ 0 & \dots & -s & \dots & c & \dots & 0 \\ \vdots & & \vdots & & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & \dots & 0 & \dots & 1 \end{pmatrix} \tag{2.23}$$

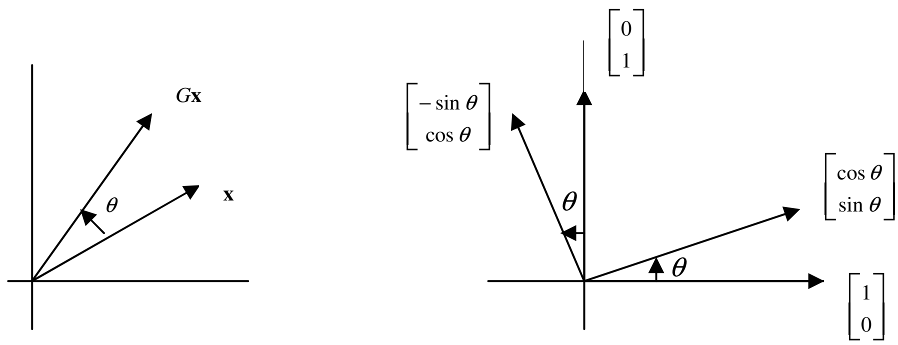
biçiminde bir matristir. Burada $c=\cos\theta$ ve $s=\sin\theta$ sırasıyla i-inci satır ve k-ıncı sütunda yer alırlar, yani $G_{i,i}=c,\ G_{k,k}=c,\ G_{i,k}=s\ ve\ G_{k,i}=-s$ dır. $G(i,k,\theta)^T\mathbf{x}$ çarpımı, $\mathbf{x}$ vektörünün saat yönünün tersine (i,k) düzleminde θ açısı kadar döndürüldüğünü gösterir. Eğer $\mathbf{x}\in\mathbb{R}^n$ ve $\mathbf{y}=G(i,k,\theta)^T\mathbf{x}$ ise;

$$y_j=\begin{cases} cx_i-sx_k & ,\ j=i \\ sx_i+cx_k & ,\ j=k \\ x_j & ,\ j\neq i,k \end{cases} \tag{2.24}$$

dır. Buradan $c = \dfrac{x_i}{\sqrt{x_i^2 + x_k^2}}$ ve $s = \dfrac{-x_k}{\sqrt{x_i^2 + x_k^2}}$ olduğu açıktır. $c = x_i / r$, $s = x_k / r$ ve $r = \sqrt{x_i^2 + x_j^2} \neq 0$ şeklinde düzenlenirse;

$$\begin{pmatrix} c & s \\ -s & c \end{pmatrix} \begin{pmatrix} x_i \\ x_j \end{pmatrix} = \begin{pmatrix} r \\ 0 \end{pmatrix} \tag{2.25}$$

elde edilir[2]. Şekil 2.1'de, $G(i,k,\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$ olduğunda **x** vektörünün saat yönünün tersine döndürülmesi verilmiştir[3]. Çizelge 2.1'de ise Matlab benzeri program için Givens algoritması verilmiştir.



Şekil 2.1 θ açısı ile saat yönünün tersine döndürme

Çizelge 2.1 Givens Algoritması

```
function [c,s] = givens(α,β)
if β = 0
c = 1;  s = 0;
   elseif (|β| > |α|)
      r = −α / β;   s = 1/ √(1 + r²);   c = s × r;
   else
      r = −β / α;   c = 1/ √(1 + r²);   s = c × r;
   end
end
```

Tanım 2.2.6 (Householder Dönüşümü) $\mathbf{v} \neq 0$ bir vektör olmak üzere

$$H = I - \frac{1}{\gamma}\mathbf{v}\mathbf{v}^T, \quad \gamma = \frac{1}{2}\mathbf{v}^{\mathbf{T}}\mathbf{v} \tag{2.26}$$

matrisine Householder matrisi veya *Householder dönüşümü* denir. Ayrıca $\mathbf{v}$ vektörüne de Householder vektörü denir. Verilmiş bir $\mathbf{x}$ vektörünün H ile çarpılması sonucunda

$$H\mathbf{x} = \left(I - \frac{1}{\gamma}\mathbf{v}\mathbf{v}^{\mathbf{T}} \right)\mathbf{x} = \mathbf{x} - \frac{1}{\gamma}\left(\mathbf{v}^{\mathbf{T}}\mathbf{x}\right)\mathbf{v} \tag{2.27}$$

elde edilir. Eğer

$$\mathbf{v} = \mathbf{x} \pm \sigma\mathbf{e}_1, \quad \sigma = \|\mathbf{x}\|_2 \tag{2.28}$$

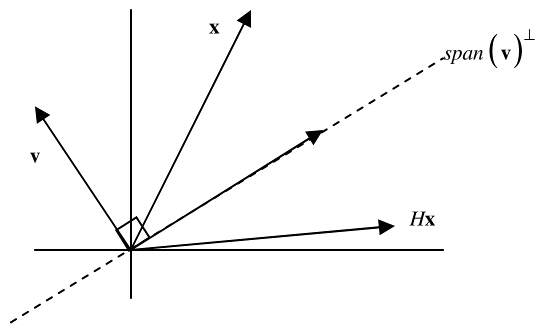
şeklinde alınırsa, $H\mathbf{x} = \pm\sigma\mathbf{e}_1$ elde edilir. Ayrıca $\alpha_1 = \mathbf{x}^{\mathbf{T}}\mathbf{e}_{\mathbf{1}}$ alınırsa;

$$\mathbf{v}^{\mathbf{T}}\mathbf{v} = \left(\mathbf{x} \pm \sigma\mathbf{e}_{\mathbf{1}}\right)^T \left(\mathbf{x} \pm \sigma\mathbf{e}_{\mathbf{1}}\right) = \sigma^2 \pm 2\sigma\alpha_1 + \sigma^2 = 2\sigma\left(\sigma \pm \alpha_1\right)$$

böylece $\gamma = \sigma(\sigma \pm \alpha_1)$ elde edilir. Eğer $\mathbf{x}$, $\mathbf{e_1}$'in bir katına yakınsa, $\sigma \approx |\alpha_1|$ olur. Bu γ'da büyük bir göreli hataya neden olabilir. Bunu önlemek için

$$\mathbf{v} = \mathbf{x} + sign(\alpha_1)\sigma\mathbf{e_1}, \quad \gamma = \sigma(\sigma + |\alpha_1|) \tag{2.29}$$

biçiminde alınır. Çizelge 2.2'de Householder Dönüşümü'nün algoritması verilmiştir[2].



Şekil 2.2 Householder Dönüşümü

Çizelge 2.2 Householder Algoritması

$$
\begin{aligned}
&function\left[\mathbf{v}, \gamma, \widehat{\sigma}\right] = house(\mathbf{v}) \\
&\quad \sigma = \sqrt{\mathbf{x}^T\mathbf{x}}; \\
&\quad \alpha_1 = \mathbf{x}^T\mathbf{e}_1; \\
&\quad \widehat{\sigma} = -sign(\alpha_1)\sigma; \\
&\quad \mathbf{v} = \mathbf{x} - \widehat{\sigma}\mathbf{e}_1; \\
&\quad \gamma = \sigma(\sigma + |\alpha_1|);
\end{aligned}
$$

2.3 En Küçük Kareler Problemi (EKKP) ve İstatistikteki Kullanımı

Lineer EKKP, verilen gözlemler için bir lineer matematiksel model uydurmaya ihtiyaç duyulduğundan ortaya çıkmıştır. Gözlemlerdeki hataların etkisini azaltmak için takip edilen yöntemlerden biri gözlemlerin sayısını artırmaktır. Bunun sonucunda bilinmeyenlerin sayısından fazla denkleme sahip aşağıdaki gibi bir lineer denklem sistemi elde edilir:

$$A\mathbf{x} = \mathbf{b} \tag{2.30}$$

burada $A \in \mathbb{R}^{m\times n}$, $m \geq n$ veri matrisi, $\mathbf{b} \in \mathbb{R}^{m}$ ise gözlem vektörüdür. Amacımız (2.30) eşitliğini sağlayan en iyi $\mathbf{x} \in \mathbb{R}^{n}$ vektörünü bulmaktır. Burada en iyi çözümü veren birçok yöntem vardır. Bu yöntemlerden biri, $\mathbf{x}$ 'in

$$\min_{\mathbf{x}\in\mathbb{R}^n} \left\| A\mathbf{x} - \mathbf{b} \right\|_2, \quad A \in \mathbb{R}^{m\times n}, \quad \mathbf{b} \in \mathbb{R}^{m}, \quad m \geq n \tag{2.31}$$

minimizasyon problemi için bir çözüm olarak alınmasıdır. Bu durumda (2.31) ifadesine *en küçük kareler problemi(EKKP)* denir ve $\mathbf{x}$ de $A\mathbf{x} = \mathbf{b}$ sisteminin *en küçük kareler çözümü* olarak adlandırılır.

$\mathbf{r} = \mathbf{b} - A\mathbf{x}$ ifadesi ise artık vektörüdür. EKKP'nin en iyi çözümü $\min \left\| \mathbf{r} \right\|_2^2 = \sum_{i=1}^{m} r_i^2$ olması durumundadır. A tam ranklı olduğunda (2.31) problemi, Normal Denklem Metodu ile çözülebilir. Eğer $rank\left(A\right) < n$ ise (2.31) için $\mathbf{x}$ 'lerin çözümü tek değildir ve SVD veya URV ayrışımları yardımıyla çözülebilir (bkz. Kesim 3.2.3 ve Kesim 3.2.5). Ancak tüm EKK çözümleri arasında tek bir çözüm, $\left\| \mathbf{x} \right\|_2$ 'in en küçük olduğu çözümdür[1,2]. Teorem 3.6 bu çözümle ilgili bilgi vermektedir.

(2.31) ifadesinde $(A,\mathbf{b})$ veri setinin satırlarına veri eklemek, çıkarmak veya her ikisinin de adım adım yapılmasından bahsedilmektedir. Değişik zaman serisi problemlerinde veri seti üzerinde bulunan hareket halindeki bir pencere kullanılır. Yeni gözlem değeri eklendiğinde yeni bir adımla pencere hareket eder. Diğer uygulamalarda ise A matrisinin sütunlarına veri eklenir veya çıkarılır. Bu gibi değişiklikler genellikle EKK çözümlerinin güncellenmesi için önerilir.

İstatistik, optimizasyon ve sinyal işlemede EKK yöntemi biraz daha değişiklik gösterir. Sinyal işleme uygulamalarında çoğu zaman, eş zamanlı çözümlere ihtiyaç duyulur. İstatistikte ise satırların eklenmesi veya çıkarılması için en etkin yöntem bir regresyon modelinin kurulmasıdır. Regresyon analizi, bir bağımlı değişken ile bir veya birden fazla bağımsız değişken arasındaki ilişkinin bir matematiksel model ile açıklanması sürecidir[15]. Regresyon analizinde, her bir gözlem değeri için hata veya gözlem değerleri tahmin edilir[14,16]. Böylece EKK yöntemi istatistikte, parametre tahmini yapmak için kullanılır.

Lineer istatistiksel modelde, $\mathbf{b}\in\mathbb{R}^m$ gözlenen vektör, $A\in\mathbb{R}^{m\times n}$ bilinen bir matris, $\mathbf{x}\in\mathbb{R}^n$ bilinmeyen parametrelerin vektörü, $\boldsymbol{\varepsilon}$ rasgele değişkenlerin gözlenemeyen bir vektörü olmak üzere bunlar arasında

$$\mathbf{b}=A\mathbf{x}+\boldsymbol{\varepsilon} \qquad (2.32)$$

biçiminde varsayılan bağıntıya *lineer model* denir. Standart lineer modelde aksi belirtilmedikçe $rank(A)=n$ olduğu kabul edilecektir. Ayrıca $\boldsymbol{\varepsilon}$ ile ilgili varsayımlar aşağıda verilmiştir:

1. $\boldsymbol{\varepsilon} \sim N\left(\mathbf{0}, \sigma^2 I\right)$ durumunda her bir $\varepsilon_i, i = 1, 2, \ldots, m$ birbirinden bağımsız ve 0 ortalamalı, bilinmeyen σ^2 varyanslı normal dağılıma sahiptir.

2. $\boldsymbol{\varepsilon} \sim \left(0, \sigma^2 I\right)$ durumunda ε_i 'ler ilişkisiz (uncorrelated) ve $E\left(\boldsymbol{\varepsilon}\right) = 0$ ortalamalı, $Cov\left(\boldsymbol{\varepsilon}\right) = \sigma^2 I$ varyanslı bilinmeyen bir dağılıma sahiptir.

3. $Cov\left(\boldsymbol{\varepsilon}\right) = \sigma^2 V$, V bilinen pozitif tanımlı bir matristir.

1. ve 2. durumdaki varsayımlar altındaki modellere Gauss-Markov modelleri denir. 2. durumdaki modele *en küçük kareler modeli* denir(14).

Gerçek dünyadaki olayların lineer model olarak modellenmesi sırasında $\mathbf{b}$, A, $\mathbf{x}$ ve $\boldsymbol{\varepsilon}$ çok değişik şekillerde anlamlandırılabilir. Örneğin, bazı modellerde $\mathbf{b}$ üretim miktarı, boy uzunluğu, ekonomi değişkeni vb. ile ilgili gözlem vektörü olabilir(14).

Bir lineer regresyon modeli;

$$b_i = x_0 + x_1 a_{1i} + x_2 a_{2i} + \ldots + x_{(n-1)} a_{(n-1)i} + \varepsilon_i \quad , \quad i = 1:m \qquad (2.33)$$

şeklinde tanımlanır. (2.33) eşitliğinin matris gösterimi

$$\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}_{m\times 1} = \begin{pmatrix} 1 & a_{11} & a_{12} & \cdots & a_{1(n-1)} \\ 1 & a_{21} & a_{22} & \cdots & a_{2(n-1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & a_{m1} & a_{m2} & \cdots & a_{m(n-1)} \end{pmatrix}_{m\times n} \begin{pmatrix} x_0 \\ x_1 \\ \vdots \\ x_n \end{pmatrix}_{n\times 1} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_m \end{pmatrix}_{m\times 1} \qquad (2.34)$$

veya

$$\mathbf{b} = A\mathbf{x} + \boldsymbol{\varepsilon} \qquad (2.35)$$

şeklindedir. (2.35) eşitliğindeki terimler

$\mathbf{b}$: $(m \times 1)$ boyutlu cevap vektörü (Bağımlı değişken vektörü)

A : $(m \times n)$ boyutlu bilinen matris (Bağımsız değişken matrisi)

$\mathbf{x}$: $(n \times 1)$ boyutlu bilinmeyen katsayılar vektörü

$\boldsymbol{\varepsilon}$: $(m \times 1)$ boyutlu hata vektörü

olarak ifade edilir. $\mathbf{x}$ vektöründeki x_0 regresyon sabitidir. A matrisindeki 1'ler sütunu bu terime karşılık gelir. EKK yönteminin temeli, hata terimleri olan ε_i'lerin kareleri toplamını en küçük yapan en iyi x_i değerlerinin kestirimine dayanır(16).

Örneğin, Türkiye'de iller düzeyindeki kişi başına düşen gelir ile sosyoekonomik gelişmişlik arasında ilişki kurulmak istensin. İl düzeyinde kişi başına gayri safi milli hasılayı (KBGSMH) etkileyen (açıklayan) çok sayıda açıklayıcı değişken söz konusudur. Açıklayıcı değişkenler

X_1 : Sanayileşmeye dayalı sosyoekonomik gelişmişlik yapısı

X_2 : Eğitim düzeyi ve sağlık hizmetleri

X_3 : Tarımsal yapı

X_4 : Kentleşme ve istihdam

X_5 : Coğrafi yapı

X_6 : Altyapı, konut ve nüfus hareketliliği

X_7 : Bebek ve çocuk ölüm hızı

X_8 : Yüksek öğretim düzeyi

olarak seçildiğinde Kalaycı (2008)'in verilerine göre EKK metoduyla yapılan regresyon analizi sonucunda bağımlı değişkenin tahmini lineer denklemi

$$\hat{Y} = 13.99 + 0.131X_1 + 0.289X_2 + 0.114X_3$$

$$+0.181X_4 + 0.253X_5 + 0.266X_6 + 0.129X_7 - 0.080X_8 \qquad (2.36)$$

olarak elde edilir. (2.36) eşitliği sayesinde daha sonraki yıllardaki açıklayıcı değişken değerleri ile KBGSMH için bir ön kestirim yapılabilir(15).

3. ARAŞTIRMA BULGULARI

Bu bölümde, EKK problemi çözümü için nümerik metodlar verilecektir.

3.1 QR Ayrışımı

$A \in \mathbb{R}^{m \times n}$, $\mathbf{b} \in \mathbb{R}^{m}$ ve $Q \in \mathbb{R}^{m \times m}$ ortogonal bir matris olsun. Ortogonal dönüşümler normu koruduğundan

$$\min_{\mathbf{x}} \left\| Q^T \left(A\mathbf{x} - \mathbf{b} \right) \right\|_2 \tag{3.1}$$

ile EKK problemi $\min_{\mathbf{x}} \left\| A\mathbf{x} - \mathbf{b} \right\|_2$ birbirine eşittir. Buna göre, seçilen bir Q matrisi için (3.1) eşitliğinin basit çözümünün nasıl olduğunu göstereceğiz.

3.1.1 Tam Ranklı QR Ayrışımı

$A \in \mathbb{R}^{m \times n}$, $m \geq n$ olmak üzere A matrisinin QR ayrışımı

$$A = Q \begin{pmatrix} R \\ 0 \end{pmatrix} \tag{3.2}$$

şeklindedir. Burada $Q \in \mathbb{R}^{m \times m}$ ortogonal matris ve R negatif olmayan köşegen elemanlarına sahip üst üçgensel matristir. R matrisi aynı zamanda A *'nın* R *-faktörü* olarak da adlandırılır[(2)].

3.1.2 Rankı Veren QR Ayrışımı

Herhangi bir $A \in \mathbb{R}^{m \times n}$ matrisi QR ayrışımına sahip olsun. Eğer $rank(A) < n$ ise bu ayrışım tek değildir. Örneğin; herhangi c ve s *ler* için $c^2 + s^2 = 1$ olsun.

$$A = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} c & -s \\ s & c \end{pmatrix} \begin{pmatrix} 0 & s \\ 0 & c \end{pmatrix} = QR$$

olsun. Burada $rank(A) = 1 < n = 2$ dir. Böylece Q matrisinin sütunları $R(A)$ için ortogonal bir tabandır.

Teorem 3.1 $A \in \mathbb{R}^{m \times n}$ ve $rank(A) = r$ olsun. Π permütasyon matrisi ve $Q \in \mathbb{R}^{m \times m}$ ortogonal matris olmak üzere, $R_{11} \in \mathbb{R}^{r \times r}$ üst üçgensel ve sıfır olmayan köşegen elemanlarına sahip matris ise

$$A\Pi = Q \begin{array}{c} \begin{pmatrix} R_{11} & R_{12} \\ 0 & 0 \end{pmatrix} \end{array} \begin{array}{l} r \\ m-r \end{array} \tag{3.3}$$

şeklinde değiştirilebilir[2].

İspat 3.1 Teorem 3.1'in ispatı A. Björck[2]'da verilmiştir.

3.2 Tam Ortogonal Ayrışımlar

Rankı r olan $A \in \mathbb{R}^{m \times n}$ matrisinin tam ortogonal ayrışımı

$$A = U \begin{pmatrix} T & 0 \\ 0 & 0 \end{pmatrix} V^T \tag{3.4}$$

şeklindedir. Burada $U \in \mathbb{R}^{m\times m}$ ve $V \in \mathbb{R}^{n\times n}$ ortogonal matrisler ve $T \in \mathbb{R}^{r\times r}$ köşegen matrisi veya tercihen pozitif elemanlara sahip alt veya üst üçgensel matristir.

(3.4) ile verilen ayrışım tek değildir. Bu ayrışımlar T 'ye göre adlandırılırlar. Örneğin, T köşegen matris iken Tekil Değer Ayrışımı, T üst üçgensel matris iken URV Ayrışımı ve T alt üçgensel matris iken ULV Ayrışımı olarak anılır.

3.2.1 Tekil Değer Ayrışımı

$A\in \mathbb{R}^{m\times n}$, $m \geq n$ olmak üzere A matrisinin tekil değer ayrışımı (SVD: Singular Value Decomposition)

$$A = X \begin{pmatrix} \Sigma \\ 0 \end{pmatrix} Y^T \tag{3.5}$$

şeklindedir. Burada $X = (\mathbf{x}_1 \quad \mathbf{x}_2 \quad \ldots \quad \mathbf{x}_m) \in \mathbb{R}^{m\times m}$ ortogonal matris, $Y = (\mathbf{y}_1 \quad \mathbf{y}_2 \quad \ldots \quad \mathbf{y}_n) \in \mathbb{R}^{n\times n}$ ortogonal matris ve Σ matrisi ise köşegen elemanları A matrisinin tekil değerleri olan köşegen matristir(8,17). A 'nın rankındaki değişikliklere göre aşağıdaki durumlardan bahsedilebilir:

1. $rank(A) = n$ ise

$$\Sigma = \begin{pmatrix} \sigma_1 & & & 0 \\ & \sigma_2 & & \\ & & \ddots & \\ 0 & & & \sigma_n \end{pmatrix} \tag{3.6}$$

dır. Σ 'nın tekil değerleri $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_n > 0$ sıralamasına sahip olabilir.

2. $rank(A) = r < \min\{m,n\}$ ise $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_r > 0 = \sigma_{r+1} = \sigma_{r+2} = \ldots = \sigma_n$

sıralaması geçerlidir. Bu durumda Σ matrisi

$$\Sigma = \begin{pmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{pmatrix} \tag{3.7}$$

şeklinde yazılabilir. Burada $\Sigma_1 = diag(\sigma_1, \sigma_2, \ldots, \sigma_r)$ 'dir[3].

SVD, A matrisine ait birçok bilgiyi ortaya çıkarır. Örneğin,

$$\begin{aligned} null(A) &= span\{\mathbf{y}_{r+1}, \ldots, \mathbf{y}_n\} \\ range(A) &= span\{\mathbf{x}_1, \ldots, \mathbf{x}_r\}. \end{aligned} \tag{3.8}$$

Ayrıca,

$$\begin{aligned} \|A\|_F^2 &= \sigma_1^2 + \ldots + \sigma_r^2, \quad r = \min\{m,n\} \\ \|A\|_2 &= \sigma_1 \\ \min_{\mathbf{x}\neq 0} \frac{\|A\mathbf{x}\|_2}{\|\mathbf{x}\|_2} &= \sigma_n, \quad (m \geq n). \end{aligned} \tag{3.9}$$

Teorem 3.2 $A \in \mathbb{R}^{m\times n}$ matrisi $rank(A) = r$ sahip olsun. Ayrıca $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_r > 0$ reel sayıları, $\mathbf{y}_1, \ldots, \mathbf{y}_n$ $\mathbb{R}^n$ 'nin ortonormal bir tabanı ve $\mathbf{x}_1, \ldots, \mathbf{x}_m$ $\mathbb{R}^m$ 'nin ortonormal bir tabanı var olsun. Öyle ki

$$A\mathbf{y}_i = \begin{cases} \sigma_i \mathbf{x}_i, & i = 1, \ldots, r \\ 0, & i = r+1, \ldots, n \end{cases} \tag{3.10}$$

ve

$$A^T \mathbf{x}_i = \begin{cases} \sigma_i \mathbf{y}_i, & i = 1,\dots,r \\ 0, & i = r+1,\dots,m \end{cases} . \tag{3.11}$$

O zaman $\mathbf{y}_1,\dots,\mathbf{y}_n$ ve $\mathbf{x}_1,\dots,\mathbf{x}_m$ 'ler sırasıyla $A^T A$ ve AA^T 'nın özvektörleri ve $\sigma_1^2,\dots,\sigma_r^2$ ler ise $A^T A$ ve AA^T nin sıfırdan farklı özdeğerleridir[3].

İspat 3.2 Teorem 3.2'nin ispatı D. S. Watkins[3]'da verilmiştir.

Teorem 3.3 $A \in \mathbb{R}^{m\times n}$ matrisinin $rank(A) = r$ olsun. $\sigma_1,\dots,\sigma_r$, A 'nın 0'dan farklı tekil değerleri $\mathbf{y}_1,\dots,\mathbf{y}_r$ ve $\mathbf{x}_1,\dots,\mathbf{x}_r$ vektörleri sırasıyla sağ ve sol tekil vektörler olmak üzere A 'nın SVD'si

$$A = \sum_{i=1}^{r} \sigma_i \mathbf{x}_i \mathbf{y}_i^T \tag{3.12}$$

şeklinde yazılabilir[3].

İspat 3.3 $B = \sum_{i=1}^{r} \sigma_i \mathbf{x}_i \mathbf{y}_i^T \in \mathbb{R}^{m\times n}$ olsun. $\mathbf{y}_1,\dots,\mathbf{y}_n$, $\mathbb{R}^n$ nin bir bazı olduğundan $A = B$ olduğunu göstermek için, $A\mathbf{y}_j = B\mathbf{y}_j, \quad j = 1,2,\dots,n$ olduğunu göstermek yeterlidir.

$j \le r$ ise $A\mathbf{y}_j = \sigma_j \mathbf{x}_j$ ve $By_j = \sum_{i=1}^{r} \sigma_i \mathbf{x}_i \left(\mathbf{y}_i^T \mathbf{y}_j\right)$ vardır. Çünkü $\left(\mathbf{y}_1,\dots,\mathbf{y}_n\right)$ ortonormal vektör kümesidir ve

$$\mathbf{y}_i^T \mathbf{y}_j = \begin{cases} 0, & i \ne j \\ 1, & i = j \end{cases} \tag{3.13}$$

dır. Böylece $i \ne j$ olduğu durumlarda terimlerin toplamda sıfırdır ve $B\mathbf{y}_j = \sigma_j \mathbf{x}_j$ dir.

$j > r$ ise $A\mathbf{y}_j = 0$ ve $B\mathbf{y}_j = \sum_{i=1}^{r} \sigma_i \mathbf{x}_i \left(\mathbf{y}_i^T \mathbf{y}_j\right) = \sum_{i=1}^{r} \sigma_i \mathbf{x}_i (0) = 0$ dır.

3.2.2 Sayısal Rank

SVD, veri matrisindeki belirsizlikler ve yuvarlanmış hataların olmadığı durumda, matrisin rankını verir. Ne yazık ki hataların varlığı ve belirsizlikler rank sorusunu anlamsız kılar. Tam ranklı olmayan matrislerdeki ufak bir değişim rankı arttıracaktır. Bunun sonucu olarak rank tanımımızı uyarlamamız gerekir(3).

Tanım 3.1 Bir A matrisi için

$$k = \min\left\{ rank(B) : \|A - B\|_2 \leq \varepsilon \right\} \tag{3.14}$$

olacak şekilde bir B matrisi varsa, A'ya k *sayısal* $\varepsilon - rank$'a sahiptir denir(2).

Eğer $k < n$ ise

$$\inf_{rank(B) \leq k} \|A - B\|_2 = \sigma_{k+1} \tag{3.15}$$

dır. Burada $\sigma_i, i = 1, 2, \ldots, \min(m, n)$ A'nın tekil değerleri ve $B = \sum_{i=1}^{k} \sigma_i \mathbf{x}_i \mathbf{y}_i^T$ dır. Bu nedenle A, k *sayısal* $\varepsilon - rank$'a sahipse

$$\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_k > \varepsilon \geq \sigma_{k+1} \geq \ldots \geq \sigma_n \tag{3.16}$$

sıralaması geçerlidir(2).

Teorem 3.4 $A \in \mathbb{R}^{m\times n}$ matrisinin $rank(A) = r > 0$ ve SVD'si $A = X\Sigma Y^T$ olsun. $k = 1,\ldots,r-1$ için $A_k = X\Sigma_k Y^T$ biçiminde tanımlansın, öyle ki $\Sigma_k \in \mathbb{R}^{m\times n}$

$$\Sigma_k = \begin{bmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{bmatrix}, \quad \Sigma_1 = diag(\sigma_1, \sigma_2, \ldots, \sigma_k)$$

olsun. O zaman $rank(A_k) = k$ ve

$$\sigma_{k+1} = \|A - A_k\|_2 = \min\left\{\|A - B\|_2 : rank(B) = k\right\}$$

dır. Diğer bir deyişle, rankı k olan matrisler arasında A_k, A'ya en yakın olanıdır(3).

İspat 3.4 $rank(A_k) = k$ olduğu aşikardır. $A - A_k = X(\Sigma - \Sigma_k)Y^T$ olduğundan $A - A_k$'nın en büyük tekil değerinin σ_{k+1} olduğu da açıktır, yani $\|A - A_k\|_2 = \sigma_{k+1}$ dır. Şimdi bazı k ranklı B matrisleri için $\|A - B\|_2 \geq \sigma_{k+1}$ olduğu gösterelim. Öncelikle,

$$\begin{aligned} \dim(null(B)) &= \dim(\mathbb{R}^n) - \dim(range(B)) \\ &= n - rank(B) \\ &= n - k \end{aligned}$$

olduğundan, rankı $n-k$ olan B matrisini alalım. Dikkat edelim ki $\dim(\mathbf{y}_1,\ldots,\mathbf{y}_{k+1}) = k+1$ dır, $null(B)$ ve $span(\mathbf{y}_1,\ldots,\mathbf{y}_{k+1})$, $\mathbb{R}^n$'nin iki alt uzayıdır.

$null(B) \cap (\mathbf{y}_1,\ldots,\mathbf{y}_{k+1})$ uzayından $\hat{\mathbf{z}}$ vektörü alınsın ($\hat{\mathbf{z}} \neq 0$) ve $\|\hat{\mathbf{z}}\|_2 = 1$ olsun. $\hat{\mathbf{z}} \in (\mathbf{y}_1,\ldots,\mathbf{y}_{k+1})$ alalım. O zaman $c_1,\ldots,c_{k+1}$ skalerleri vardır, öyle ki

$\hat{\mathbf{z}} = c_1\mathbf{y}_1 + \ldots + c_{k+1}\mathbf{y}_{k+1}$. $\mathbf{y}_1,\ldots,\mathbf{y}_{k+1}$ ortonormal olduğundan, $c_1^{\,2} + \ldots + c_{k+1}^{\,2} = \left\|\hat{\mathbf{z}}\right\|_2^2 = 1$ dir. $\hat{\mathbf{z}} \in null\left(B\right)$ olduğunda, $B\hat{\mathbf{z}} = 0$ dır. Böylece

$$\left(A-B\right)\hat{\mathbf{z}} = A\hat{\mathbf{z}} = \sum_{j=1}^{k+1} c_j A\mathbf{y}_j = \sum_{j=1}^{k+1} \sigma_j c_j \mathbf{x}_j \ .$$

Ayrıca $\mathbf{x}_1,\ldots,\mathbf{x}_{k+1}$ tekil vektörleri ortonormal olduğundan

$$\left\|\left(A-B\right)\hat{\mathbf{x}}\right\|_2^2 = \sum_{j=1}^{k+1}\left|\sigma_j c_j\right|^2 \geq \sigma_{k+1}^2 \sum_{j=1}^{k+1}\left|c_j\right|^2 = \sigma_{k+1}^2$$

O halde

$$\left\|A-B\right\|_2 \geq \frac{\left\|\left(A-B\right)\hat{\mathbf{x}}\right\|_2}{\left\|\hat{\mathbf{x}}\right\|_2} \geq \sigma_{k+1}$$

olur ve ispat tamamlanır.

Sonuç: $A \in \mathbb{R}^{m\times n}$, $r = \min\{m,n\}$ tam ranklı ve $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_r > 0$ A 'nın tekil değerleri olsun. $B \in \mathbb{R}^{m\times n}$ matrisi $\left\|A-B\right\|_2 < \sigma_r$ şartını sağlar. Ayrıca B de tam ranklı bir matristir(3).

SVD, A veri matrisinin rank, altuzay ve matris iyileştirme gibi bilgilerini içerir ve düşük rank varsayımının hesabında kullanışlıdır(8).

SVD'nin en önemli uygulamalarından bazıları, zayıf koşullu En Küçük Kareler Problemi (EKKP) ve Tam En Küçük Kareler Problemi (TEKKP) dir. Kesim 3.2.3'de EKKP'nin SVD ile çözümü verilmektedir.

SVD'nin bu gibi yararlarının aksine, veri matrisine yeni veri eklendiğinde, köşegen matris yapısından dolayı yeni veri matrisinin güncellenmesinin matematiksel hesaplama karmaşasının $O\left(n^3\right)$ olması gibi dezavantajları da vardır. Bunun için SVD, yinelemeli EKK problemi ve yinelemeli TEKK problemi çözümünde uygun değildir. Bu nedenle alternatif ayrışımlar önerilmiştir. Bunlar arasında URV (bkz. Kesim 3.2.4) ve ULV ayrışımları vardır.

3.2.3 En Küçük Kareler Probleminin SVD ile Çözümü

SVD özellikle rankı ortaya çıkaran tam ortogonal bir ayrışımdır. Ayrıca $\mathbf{x}_{LS}$ çözümünün ve $r_{LS} = \|A\mathbf{x}_{LS} - \mathbf{b}\|_2$ artıkların normunun en küçük olması için etkili bir anlatım sağlar.

Teorem 3.5 $A \in \mathbb{R}^{m\times n}$ matrisinin rankı r ve SVD'si $\Sigma = X^T A Y$ olsun. Eğer $X = \left(\mathbf{x}_1, \ldots, \mathbf{x}_m\right)$, $Y = \left(\mathbf{y}_1, \ldots, \mathbf{y}_n\right)$ ve $\mathbf{b} \in \mathbb{R}^m$ ise

$$\mathbf{x}_{LS} = \sum_{i=1}^{r} \frac{\mathbf{x}_i^T \mathbf{b}}{\sigma_i} \mathbf{y}_i \tag{3.17}$$

dır. Ayrıca

$$r_{LS}^2 = \|A\mathbf{x}_{LS} - \mathbf{b}\|_2^2 = \sum_{i=r+1}^{m} \left(\mathbf{x}_i^T \mathbf{b}\right)^2 \tag{3.18}$$

dır[(1)].

İspat 3.5 Bazı $\mathbf{x}_{LS} \in \mathbb{R}^n$ ler için

$$\|A\mathbf{x}_{LS} - \mathbf{b}\|_2^2 = \left\|\left(X^T A Y\right)\left(Y^T \mathbf{x}_{LS}\right) - \left(X^T \mathbf{b}\right)\right\|_2^2$$

$$= \left\| \Sigma\alpha - X^T \mathbf{b} \right\|_2^2$$

$$= \sum_{i=1}^{r} \left(\sigma_i \alpha_i - \mathbf{x}_i^T \mathbf{b} \right)^2 + \sum_{i=r+1}^{m} \left(\mathbf{x}_i^T \mathbf{b} \right)^2$$

Burada $\alpha = Y^T \mathbf{x}_{LS}$ dır. Şüphesiz $\mathbf{x}_{LS}$ EKK problemi çözümü ise $i = 1, \ldots, r$ için $\alpha_i = \dfrac{\mathbf{x}_i^T \mathbf{b}}{\sigma_i}$ dır. Eğer $\alpha(r+1:n) = 0$ ise $\mathbf{x}_{LS}$ en küçük $2-norm$ a sahiptir.

Teorem 3.6 Genel lineer EKK problemi düşünülsün.

$$\min_{\mathbf{x} \in S} \left\| \mathbf{x} \right\|_2, \quad S = \left\{ \mathbf{x} \in \mathbb{R}^n : \left\| \mathbf{b} - A\mathbf{x} \right\|_2 = \min \right\} \tag{3.19}$$

(3.19) eşitliğinde $A \in C^{m \times n}$ ve $rank(A) = r \leq \min\{m, n\}$ dir. (3.19) problemi, A'nın SVD'si kullanılarak

$$\mathbf{x} = Y \begin{pmatrix} \Sigma_r^{-1} & 0 \\ 0 & 0 \end{pmatrix} X^T \mathbf{b} \tag{3.20}$$

(3.20)'te olduğu gibi tek bir çözüme sahiptir[1].

İspat 3.6 $\mathbf{z}_1, \mathbf{c}_1 \in C^r$ olmak üzere

$$\mathbf{z} = Y^T \mathbf{x} = \begin{pmatrix} \mathbf{z}_1 \\ \mathbf{z}_2 \end{pmatrix}, \qquad \mathbf{c} = X^T \mathbf{b} = \begin{pmatrix} \mathbf{c}_1 \\ \mathbf{c}_2 \end{pmatrix}$$

olsun.

$$\left\| \mathbf{b} - A\mathbf{x} \right\|_2 = \left\| X^T \left(\mathbf{b} - AYY^T \mathbf{x} \right) \right\|_2$$

$$= \left\| \begin{pmatrix} \mathbf{c_1} \\ \mathbf{c_2} \end{pmatrix} - \begin{pmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \mathbf{z_1} \\ \mathbf{z_2} \end{pmatrix} \right\|_2$$

$$= \left\| \begin{pmatrix} \mathbf{c_1} - \Sigma_1 \mathbf{z_1} \\ \mathbf{c_2} \end{pmatrix} \right\|_2 .$$

Böylece artıkların normu $\mathbf{z_2}$ için minimum olacaktır ve $\mathbf{z}_1 = \Sigma_r^{-1}\mathbf{c_1}$ dır. $\mathbf{z_2} = 0$ seçilirse $\|\mathbf{z}\|_2$ en küçük olur ve böylece $\|\mathbf{x}\|_2 = \|Y\mathbf{z}\|_2$ en iyi çözüm olacaktır.

3.2.4 URV Ayrışımı

$A \in \mathbb{R}^{m \times n}$ matrisi r ranka sahip olsun. Teoride A'nın URV ayrışımı

$$A = U \begin{pmatrix} R & 0 \\ 0 & 0 \end{pmatrix} V^T \tag{3.21}$$

biçimindedir. (3.21) eşitliğinde U ve V ortogonal matrisler ve R ise üst üçgensel matristir(7).

Öte yandan nümerik olarak bu ayrışımın formülasyonu biraz farklıdır. A matrisinin nümerik rankı r olsun. Yani A 'nın tekil değerleri $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_r \geq \varepsilon \geq \sigma_{r+1} \geq \ldots \geq \sigma_n$ şartını sağlasın. A'nın rankı veren URV ayrışımı

$$A = U \begin{pmatrix} R & H \\ 0 & G \end{pmatrix} V^T \tag{3.22}$$

şeklindedir. Burada

- R ve G üst üçgensel matrisler

- $\inf(R) \cong \sigma_r$

- $\|H\|^2 + \|G\|^2 \cong \sigma_{r+1}^2 + \ldots + \sigma_n^2$ (8).

Ayrışım sezgisel olarak H ve G matrisleri önceden bilinen toleranslardan daha küçük ve R 'nin en küçük tekil değerinden daha büyük olduğunda rankı veren anlamına gelmektedir. Böyle ayrışımlar –bunlar tek değildir- tam ranklı olmayan sistemlerin çözümünde kullanışlıdır. Ayrıca $V = \begin{pmatrix} V_1 & V_2 \end{pmatrix}$ biçiminde parçalandığında

$$\|AV_2\|_2 = \left\| \begin{pmatrix} H \\ G \end{pmatrix} \right\|_2 . \tag{3.23}$$

Böylece V_2 nin sütunları sıfır uzay A 'nın bir ortonormal bazına yaklaşır[4,8].

URV ayrışımının SVD'ye göre avantajı, A matrisine satır eklendiğinde güncellenebilmesinin hesaplama karmaşasının düşük olmasıdır. (Birçok uygulamada A , eski veri setinin etkisini azaltmak için 1'den küçük bir sabitle çarpılır. Bu uygulama üstten pencereleme yöntemi olarak bilinir[18,19]). Güncelleme işlemi, $O(n^2)$ kadar operasyon ve rankı veren ayrışım karakterini saklamayı gerektirir.

URV ayrışımı yinelemeli EKKP uygulamaları için tatmin edici çözümler üretmede yeterlidir. Fakat sıfır uzay gerektiğinde (3.23) eşitliğindeki H nedeniyle daha az yeterlidir. (3.22) eşitliğindeki U matrisi $\begin{pmatrix} U_1 & U_2 \end{pmatrix}$ şeklinde parçalandığında $\|U_2^T A\| = \|G\|$ olduğu kolaylıkla görülebilmektedir. Böylece A 'nın son $n-k$ tane tekil değeri, $\|G\|$ 'ye eşit ya da daha küçüktür. Uygun olarak düzenlenen sol tekil vektörler, $\|G\|$ 'ye eşit ya da daha az norma sahip sıfır uzay yaklaşımının bir sonucudur. Böylece V_2 sıfır uzay yaklaşımı için uygun değildir[(4)].

3.2.5 En Küçük Kareler Probleminin URV Ayrışımı ile Çözümü

$A \in \mathbb{R}^{m \times n}$ matrisi r ranka sahip olsun. U ve V ortogonal matrisler ve R de üst üçgensel matris olmak üzere teorikte A'nın URV ayrışımı

$$A = U \begin{pmatrix} R & 0 \\ 0 & 0 \end{pmatrix} V^T$$

dır (bkz: Kesim 3.2.4). $A\mathbf{x}_{LS} = \mathbf{b}$ probleminin çözümünü en küçük olması için

$$A = U \begin{pmatrix} R & 0 \\ 0 & 0 \end{pmatrix} V^T \Rightarrow \begin{pmatrix} R & 0 \\ 0 & 0 \end{pmatrix} = U^T A V \tag{3.24}$$

olur. Böylece

$$A\mathbf{x}_{LS} = \mathbf{b} \Rightarrow U^T A V V^T \mathbf{x}_{LS} = U^T \mathbf{b} \tag{3.25}$$

elde edilir. (3.24) ve (3.25) eşitliklerinden

$$\begin{pmatrix} R & 0 \\ 0 & 0 \end{pmatrix} \mathbf{y} = \overline{\mathbf{b}} \tag{3.26}$$

yazılabilir. Burada R tersi alınabilir matris, $\mathbf{y} = V^T \mathbf{x}_{LS}$ ve $\overline{\mathbf{b}} = U^T \mathbf{b}$ dır. $\min_x \|A\mathbf{x}_{LS} - \mathbf{b}\|_2$ çözümü, $\min_{\mathbf{y} \in \mathbb{R}^n} \left\| \begin{pmatrix} R & 0 \\ 0 & 0 \end{pmatrix} \mathbf{y} - \overline{\mathbf{b}} \right\|_2$ çözümünün bulunmasıyla elde edilebilir. O halde

$$\min_{\mathbf{x}\in\mathbb{R}^n}\left\|A\mathbf{x}_{LS} - \mathbf{b}\right\|_2 = \min_{\mathbf{y}\in\mathbb{R}^n}\left\|\begin{pmatrix} R & 0 \\ 0 & 0 \end{pmatrix}\mathbf{y} - \bar{\mathbf{b}}\right\|_2 \tag{3.27}$$

yazılabilir. $\bar{\mathbf{b}} = \begin{pmatrix} \bar{\mathbf{b}}_1 \\ \bar{\mathbf{b}}_2 \end{pmatrix}$ ve $\mathbf{y} = \begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{pmatrix}$ şeklinde parçalanırsa (3.27) eşitliğinden

$$\min_{\mathbf{y}\in\mathbb{R}^n}\left\|\begin{pmatrix} R & 0 \\ 0 & 0 \end{pmatrix}\mathbf{y} - \bar{\mathbf{b}}\right\|_2 = \min_{\mathbf{y}\in\mathbb{R}^n}\left\|\begin{pmatrix} R & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{pmatrix} - \begin{pmatrix} \bar{\mathbf{b}}_1 \\ \bar{\mathbf{b}}_2 \end{pmatrix}\right\|_2$$

$$= \min_{\mathbf{y}\in\mathbb{R}^n}\left\|\begin{pmatrix} R\mathbf{y}_1 \\ 0 \end{pmatrix} - \begin{pmatrix} \bar{\mathbf{b}}_1 \\ \bar{\mathbf{b}}_2 \end{pmatrix}\right\|_2$$

$$= \min_{\mathbf{y}\in\mathbb{R}^n}\left\|\begin{pmatrix} R\mathbf{y}_1 - \bar{\mathbf{b}}_1 \\ 0 - \bar{\mathbf{b}}_2 \end{pmatrix}\right\|_2, \qquad \mathbf{y}_1 = R^{-1}\bar{\mathbf{b}}_1$$

$$= \min_{\mathbf{y}\in\mathbb{R}^n}\left\|\begin{pmatrix} 0 \\ -\bar{\mathbf{b}}_2 \end{pmatrix}\right\|_2$$

$$= \left\|\bar{\mathbf{b}}_2\right\|_2 \tag{3.28}$$

elde edilir. Böylece $\mathbf{y} = \begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{pmatrix}$ ne olursa olsun, çözümün minimum hatası (3.28)'deki gibi olacaktır.

3.2.6 Kesik(Truncated) URV Ayrışımı

$A \in \mathbb{R}^{m\times n}$, $m \geq n$ olmak üzere A matrisinin kesik URV ayrışımı;

$$A = U_1 R V_1^T + E \tag{3.29}$$

Burada, $U_1 \in \mathbb{R}^{m\times r}$ ve $V_1 \in \mathbb{R}^{n\times r}$ matrisleri soldan ortogonal, $R \in \mathbb{R}^{r\times r}$ matrisi tersi alınabilen üst üçgensel ve $E \in \mathbb{R}^{m\times n}$ hata matrisidir. Ayrıca R ve E matrisleri

$$\left\|R^{-1}\right\|_2 \le \varepsilon^{-1}, \quad \left\|E\right\|_2 < \varepsilon, \quad EV_1 = 0 \tag{3.30}$$

şartlarını sağlar. Bu şartlardan

$$\sigma_{r+1}(A) < \varepsilon \le \sigma_r(A). \tag{3.31}$$

Böylece $U_1RV_1^T$ matrisi, A matrisinin r ranklı tahminidir. Burada ε değeri belirtilen tolerans veya gürültü seviyesidir.

Kesik URV ayrışımı, tahmini sağ ve sol sinyal uzayını meydana getirir. Bu alt uzayların doğruluğu

$$|\sin\theta| = \left\|V_1^T Y_2\right\| \tag{3.32}$$

ile karakterize edilir[6,18]. Önerme 3.1 bu alt uzaylarda bir üst sınır vermektedir.

Önerme 3.1 r ranklı $A \in \mathbb{R}^{m\times n}$ matrisinin kesik URV'si $A = U_1RV_1^T + E$, SVD'si ise $A = X_1\Sigma_1Y_1^T + X_2\Sigma_2Y_2^T$ olsun. V_1 ve Y_1 ile U_1 ve X_1 alt uzayları arasındaki uzaklık için bir üst sınır

$$|\sin\theta| \le \frac{\sigma_{\min}(R)\left\|E\right\|}{\sigma_{\min}^2(R) - \sigma_{r+1}^2(A)} \tag{3.33}$$

eşitsizliği ile verilir[18].

İspat 3.1 A'nın SVD'si ile kesik URV'si, Y_2 ile sağdan çarpılarak

$$AY_2 = U_1RV_1^TY_2 + EY_2 \tag{3.34}$$

ve

$$AY_2 = X_2\Sigma_2 \tag{3.35}$$

elde edilir. (3.34) ve (3.35) eşitliklerinden

$$X_2\Sigma_2 = U_1RV_1^TY_2 + EY_2 \Rightarrow U_1RV_1^TY_2 = X_2\Sigma_2 - EY_2$$

$$\Rightarrow RV_1^TY_2 = U_1^TX_2\Sigma_2 - U_1^TEY_2$$

$$\Rightarrow V_1^TY_2 = R^{-1}\left(U_1^TX_2\Sigma_2 - U_1^TEY_2\right) \tag{3.36}$$

elde edilir. (3.35) eşitliği soldan X_2^T ile çarpılırsa

$$X_2^TAY_2 = \Sigma_2 \tag{3.37}$$

elde edilir. (3.29) eşitliği sağdan V_1 ile çarpılırsa

$$AV_1 = U_1RV_1^TV_1 + EV_1 \Rightarrow AV_1 = U_1R$$

$$\Rightarrow X_2^TAV_1 = X_2^TU_1R$$

$$\Rightarrow X_2^TU_1 = \left(X_2^TAV_1\right)R^{-1}$$

$$\Rightarrow U_1^TX_2 = R^{-T}\left(X_2^TAV_1\right)^T$$

$$\Rightarrow U_1^TX_2 = R^{-T}V_1^TA^TX_2 \tag{3.38}$$

$$A = X_1\Sigma_1 Y_1^T + X_2\Sigma_2 Y_2^T \Rightarrow A^T = Y_1\Sigma_1 X_1^T + Y_2\Sigma_2 X_2^T$$

$$\Rightarrow A^T X_2 = Y_2\Sigma_2 \tag{3.39}$$

(3.39) eşitliği (3.38) eşitliğinde yerine yazılırsa;

$$\Rightarrow U_1^T X_2 = R^{-T} V_1^T Y_2 \Sigma_2 \tag{3.40}$$

elde edilir. (3.40) eşitliği (3.36) eşitliğinde yerine yazılıp düzenlenirse;

$$V_1^T Y_2 = R^{-1} R^{-T} V_1^T Y_2 \Sigma_2^2 - R^{-1} U_1^T E Y_2 \tag{3.41}$$

bulunur. (3.41) eşitliğinin her iki tarafının normu alınırsa;

$$\left\| V_1^T Y_2 \right\| = \left\| R^{-1} R^{-T} V_1^T Y_2 \Sigma_2^2 - R^{-1} U_1^T E Y_2 \right\| \tag{3.42}$$

elde edilir. Üçgen eşitsizliğinden;

$$\left\| \left\| R^{-1} R^{-T} V_1^T Y_2 \Sigma_2^2 \right\| - \left\| R^{-1} U_1^T E Y_2 \right\| \right\| \le \left\| V_1^T Y_2 \right\| \le \left\| R^{-1} R^{-T} V_1^T Y_2 \Sigma_2^2 \right\| + \left\| R^{-1} U_1^T E Y_2 \right\| \tag{3.43}$$

yazılabilir. (3.43) eşitsizliğinin sağ tarafından

$$\left\| R^{-1} R^{-T} V_1^T Y_2 \Sigma_2^2 \right\| + \left\| R^{-1} U_1^T E Y_2 \right\| \le \left\| R^{-1} \right\| \left\| R^{-T} \right\| \left\| V_1^T Y_2 \right\| \left\| \Sigma_2^2 \right\| + \left\| R^{-1} \right\| \left\| U_1^T \right\| \left\| E \right\| \left\| Y_2 \right\| \tag{3.44}$$

elde edilir. (3.44) eşitsizliği, (3.43) eşitsizliğinde yerine yazılırsa

$$\left\| V_1^T Y_2 \right\| \le \left\| R^{-1} \right\| \left\| R^{-T} \right\| \left\| V_1^T Y_2 \right\| \sigma_{r+1}^2 + \left\| R^{-1} \right\| \left\| E \right\|$$

$$\left\| V_1^T Y_2 \right\| \le \frac{\left\| R^{-1} \right\| \left\| E \right\|}{1 - \left\| R^{-1} \right\|^2 \sigma_{r+1}^2} \tag{3.45}$$

(3.45) eşitsizliğinin sağ tarafı düzenlenirse

$$\frac{\|R^{-1}\|\|E\|}{1-\|R^{-1}\|^2\sigma_{r+1}^2}=\frac{\frac{1}{\sigma_{\min}(R)}\|E\|}{1-\frac{1}{\sigma_{\min}^2(R)}\sigma_{r+1}^2} \tag{3.46}$$

eşitliği elde edilir. (3.46) eşitliği (3.45) eşitsizliğinde yerine yazılırsa

$$\|V_1^TY_2\|\leq\frac{\sigma_{\min}(R)\|E\|}{\sigma_{\min}^2(R)-\sigma_{r+1}^2}$$

elde edilir ve ispat tamamlanmış olur.

Önerme 3.2 r ranklı $A\in\mathbb{R}^{m\times n}$ matrisinin kesik URV'si $A=U_1RV_1^T+E$ olsun. E hata matrisi $E=A\left(I-V_1V_1^T\right)$ eşitliğini sağlar[(18)].

İspat 3.2 V_1 matrisinin sağdan ortogonallik özelliği ve $EV_1=0$ durumundan yararlanarak ispat tamamlanır.

$$\begin{aligned}A\left(I-V_1V_1^T\right)&=A-AV_1V_1^T\\&=A-\left(U_1RV_1^T+E\right)V_1V_1^T\\&=A-U_1RV_1^TV_1V_1^T+EV_1V_1^T\\&=A-U_1RV_1^T\\&=E.\end{aligned}$$

3.2.7 Kesik URV Ayrışımının Güncelleme Algoritması

A veri matrisine yeni bir veri satırı eklendiğinde, elde edilen yeni matrisin kesik URV ayrışımını elde etme sürecine güncelleme denir. Buna göre, $\mathbf{a}$ yeni veri sütunu ile birlikte yeni veri matrisi

$$\overline{A} = \begin{pmatrix} \mathbf{a}^T \\ A \end{pmatrix} \tag{3.47}$$

şeklindedir(17,18). Güncelleme algoritmasındaki en önemli alt problem, verilen bir $\mathbf{a} \in \mathbb{R}^n$ ve $V_1 \in \mathbb{R}^{n \times r}$ sağdan ortogonal matris için

$$\|\mathbf{v}\| = 1, \quad V_1^T \mathbf{v} = 0, \quad \begin{pmatrix} \mathbf{v} & V_1 \end{pmatrix} \mathbf{d} = \mathbf{a} \tag{3.48}$$

özelliklerini sağlayan $\mathbf{v} \in \mathbb{R}^n$ ve $\mathbf{d} \in \mathbb{R}^{r+1}$ vektörlerini bulmaktır(5,9). $\mathbf{d}^T = \begin{pmatrix} \beta & \mathbf{f}^T \end{pmatrix}$ parçaladığımızda

$$\overline{A} = \begin{pmatrix} 1 & 0 \\ 0 & U_1 \end{pmatrix} \begin{pmatrix} \beta & \mathbf{f}^T \\ 0 & R \end{pmatrix} \begin{pmatrix} \mathbf{v}^T \\ V_1^T \end{pmatrix} + \begin{pmatrix} 0 \\ E \end{pmatrix} \tag{3.49}$$

şeklindedir. Böylece (3.49) eşitliğindeki üst üçgensel matrisin kesik URV ayrışımını hesaplamak yeterli olacaktır. Cauchy interlace teoremine göre, yeni satır eklendikten sonra veri matrisinin rankı 1 artar veya aynen kalır. Rankın 1 artması durumunda, $\overline{A}$ matrisinin kesik URV ayrışımı (3.49)'deki gibidir. Ancak rankın aynı kalması durumunda en küçük tekil değer, gürültü seviyesi olan ε 'dan daha küçük olacağından, yeni veri satırını R veya $\begin{pmatrix} 0 \\ E \end{pmatrix}$ hata matrisi ile birleştirmek gerekir. Bu durumda, $\begin{pmatrix} \beta & \mathbf{f}^T \\ 0 & R \end{pmatrix}$ matrisinin en küçük tekil değeri ω ve buna karşılık gelen tekil vektör **w** olduğu varsayıldığında

$$\omega = \left\| \mathbf{w}^T \begin{pmatrix} \beta & \mathbf{f}^T \\ 0 & R \end{pmatrix} \right\| \tag{3.50}$$

olur. Şimdi P ve Q ortogonal matrisleri olsun, öyle ki

$$P^T \mathbf{w} = \mathbf{e}_1$$

ve

$$P^T \begin{pmatrix} \beta & \mathbf{f}^T \\ 0 & R \end{pmatrix} Q = \begin{pmatrix} \overline{\beta} & \overline{\mathbf{f}}^T \\ 0 & \overline{R} \end{pmatrix} \tag{3.51}$$

üst üçgensel matristir. Ortogonal matrisler normu koruduğundan

$$\omega = \left\| \mathbf{w}^T P P^T \begin{pmatrix} \beta & \mathbf{f}^T \\ 0 & R \end{pmatrix} Q \right\| = \left\| \mathbf{e}_1^T \begin{pmatrix} \overline{\beta} & \overline{\mathbf{f}}^T \\ 0 & \overline{R} \end{pmatrix} \right\|$$

yazılabilir. Böylece son üst üçgensel matrisin ilk satırı ω-norma sahiptir.

$$\left(\overline{\mathbf{u}} \quad \overline{U}_1 \right) = \begin{pmatrix} 1 & 0 \\ 0 & U_1 \end{pmatrix} P, \qquad \left(\overline{\mathbf{v}} \quad \overline{V}_1 \right) = \left(\mathbf{v} \quad V_1 \right) Q$$

alınırsa, $\overline{A}$ matrisinin kesik URV ayrışımı

$$\overline{A} = \overline{U}_1 \overline{R} \overline{V}_1^T + \overline{E} \tag{3.52}$$

ve hata matrisi

$$\overline{E} = \overline{\mathbf{u}}\begin{pmatrix} \overline{\beta} & \overline{\mathbf{f}}^T \end{pmatrix}\begin{pmatrix} \overline{\mathbf{v}}^T \\ \overline{V}_1^T \end{pmatrix} + \begin{pmatrix} 0 \\ E \end{pmatrix}. \tag{3.53}$$

Üçgen eşitsizliği ve matris normundaki tutarlılık özelliğinden

$$\left\|\overline{E}\right\| \le \omega + \left\|E\right\|$$

yazılabilir. Bu yüzden diğer bir problem, $\overline{E}$ hata matrisinin normunun ε gürültü seviyesinden büyük olması ihtimalidir. Başka bir deyişle, (3.30)'da verilen şartlardan $\left\|E\right\| < \varepsilon$ olması, $\overline{E}$ hata matrisinin normunun da ε gürültü seviyesinden küçük olacağı anlamına gelmez. Bu durumla ilgili algoritma Kesim 3.3'te verilecektir[(18)].

Önerme 3.3 A matrisinin kesik URV ayrışımını kullanarak $\overline{A} = \overline{U}_1\overline{R}\overline{V}_1^T + \overline{E}$ nın kesik URV ayrışımı elde edilir. O zaman hata matrisi

$$\overline{E} = \overline{A}\left(I - \overline{V}_1\overline{V}_1^T\right)$$

yazılır[(18)].

İspat 3.3 $A = U_1 R V_1^T + E$, A matrisinin kesik URV ayrışımı olsun. A matrisine $\mathbf{a} \in \mathbb{R}^n$ satırı eklendiğinde rankın arttığı varsayılsın. O halde

$$\overline{U}_1 = \begin{pmatrix} 1 & 0 \\ 0 & U_1 \end{pmatrix}, \quad \overline{R} = \begin{pmatrix} \beta & \mathbf{f}^T \\ 0 & R \end{pmatrix}, \quad \overline{V}_1 = \begin{pmatrix} \mathbf{v} & V_1 \end{pmatrix}, \quad \overline{E} = \begin{pmatrix} 0 \\ E \end{pmatrix}$$

olur. Burada $\mathbf{a} = \beta\mathbf{v} + V_1\mathbf{f}$ dır. Öyleyse

$$\overline{E} = \begin{pmatrix} 0 \\ A\left(I - V_1 V_1^T\right) \end{pmatrix}$$

$$= \begin{pmatrix} \mathbf{a}^T \\ A \end{pmatrix} - \begin{pmatrix} \beta \mathbf{v}^T + \mathbf{f}^T V_1^T \\ A V_1 V_1^T \end{pmatrix}$$

$$= \begin{pmatrix} \mathbf{a}^T \\ A \end{pmatrix} - \begin{pmatrix} \beta & \mathbf{f}^T \\ 0 & A V_1 \end{pmatrix} \begin{pmatrix} \mathbf{v}^T \\ V_1^T \end{pmatrix}$$

$$= \begin{pmatrix} \mathbf{a}^T \\ A \end{pmatrix} - \begin{pmatrix} 1 & 0 \\ 0 & U_1 \end{pmatrix} \begin{pmatrix} \beta & \mathbf{f}^T \\ 0 & R \end{pmatrix} \begin{pmatrix} \mathbf{v}^T \\ V_1^T \end{pmatrix}$$

$$= \overline{A} - \overline{U}_1 \overline{R}\,\overline{V}_1^T$$

$$= E.$$

elde edilir. İspat, $\overline{U}_1 \overline{R} = \overline{A}\overline{V}_1$ eşitliğinden yararlanılarak tamamlanmıştır.

Çizelge 3.1'de ise güncelleme algoritmasının adımları verilmiştir.

Çizelge 3.1 Güncelleme Algoritması

Girdi : $\mathbf{a}$, $n \times 1$ *boyutlu veri*
R, $r \times r$ *boyutlu üst üçgensel matris*
U_1, $m \times r$ *boyutlu sağdan ortagonal matris*
V_1, $n \times r$ *boyutlu sağdan ortagonal matris*
ε, *hata vektörü*

1. $\mathbf{v}$ ve $\mathbf{d}$ vektörleri hesaplanır:

$$\mathbf{a} = \begin{pmatrix} \mathbf{v} & V_1 \end{pmatrix} \mathbf{d}, \quad \mathbf{d} = \begin{array}{c} \begin{pmatrix} \beta \\ \mathbf{f} \end{pmatrix} \end{array} \begin{array}{l} 1 \\ r \end{array}$$

Çizelge 3.1 (devam)

$$\overline{A}=\begin{pmatrix}1 & 0\\ 0 & U_1\end{pmatrix}\begin{pmatrix}\beta & \mathbf{f}^T\\ 0 & R\end{pmatrix}\begin{pmatrix}\mathbf{v}^T\\ V_1^T\end{pmatrix}+\begin{pmatrix}0\\ E\end{pmatrix}\begin{matrix}1\\ m\end{matrix}$$

2. $\begin{pmatrix}\beta & \mathbf{f}^T\\ 0 & R\end{pmatrix}$ matrisinin en küçük tekil değeri olan ω bulunur:

$$\begin{pmatrix}\beta & \mathbf{f}^T\\ 0 & R\end{pmatrix}\mathbf{v}_{r+1}=\omega\mathbf{u}_{r+1}$$

3. Eğer $\omega \geq \varepsilon$ ise;

$$P^T\mathbf{u}_{r+1}=\mathbf{e}_1$$

ve

$$P^T\begin{pmatrix}\beta & \mathbf{f}^T\\ 0 & R\end{pmatrix}Q=\begin{pmatrix}\overline{\beta} & \overline{\mathbf{f}}^T\\ 0 & \overline{R}\end{pmatrix}$$

olacak şekilde P ve Q matrisleri bulunur.

$$\begin{pmatrix}\overline{\mathbf{u}} & \overline{U}_1\end{pmatrix}=\begin{pmatrix}1 & 0\\ 0 & U_1\end{pmatrix}P,\qquad \begin{pmatrix}\overline{\mathbf{v}} & \overline{V}_1\end{pmatrix}=\begin{pmatrix}\mathbf{v} & V_1\end{pmatrix}Q$$

elde edilir.

4. Sonuç olarak;

$$\overline{A}=\overline{U}_1\overline{R}\overline{V}_1^T+\overline{E},\qquad \overline{E}=\overline{A}\left(I-\overline{V}_1\overline{V}_1^T\right)$$

bulunur.

Güncelleme algoritmasının hesaplama karmaşası $O(nr)$ kadardır. Ancak sinyal alt uzayına ihtiyaç duyulduğunda bu hesaplama karmaşası $O(nr^2)$ 'ye çıkmaktadır.

3.3 Deflasyon

Cauchy interlace teoremine göre, veri matrisine yeni veri eklemek rankı azaltmaz, fakat veri matrisini unutma (forgetting) faktörü ile çarpıp, daha sonra yeni veri eklemek rankı azaltabilir. Bu gibi durumlarda, (3.49) eşitliğindeki üst üçgensel matrisi rankı vermeyebilir. Deflasyon algoritması ile bu özellik yeniden kazandırılabilir ve matrisin kesik URV ayrışımı iyileştirilebilir[6,18]. $EV_1 = 0$ olduğundan

$$E = A\left(I - V_1 V_1^T\right). \tag{3.54}$$

Deflasyon algoritmasının basamakları Çizelge 3.2'de verilmiştir.

Çizelge 3.2 Deflasyon Algoritması

Girdi :
R, *$r \times r$ boyutlu üst üçgensel matris*
U_1, *$m \times r$ boyutlu sağdan ortagonal matris*
V_1, *$n \times r$ boyutlu sağdan ortagonal matris*
ε, *hata vektörü*

1. E hata matrisinin en büyük tekil değeri bulunur.

$$E\mathbf{w} = \sigma \mathbf{u}_1$$

2. $\mathbf{z}$ vektörü hesaplanır.

$$\mathbf{z} = A^T \mathbf{u}_1 \left(= E^T \mathbf{u}_1\right)$$

3. $\mathbf{v}_1$ ve $\mathbf{d}$ vektörleri hesaplanır.

$$\mathbf{z} = \left(\mathbf{v}_1 \quad V_1\right)\mathbf{d}$$

burada $\mathbf{d}=\begin{pmatrix}\beta\\ \mathbf{f}\end{pmatrix}$ dır. O halde $A=\begin{pmatrix}\mathbf{u}_1 & U_1\end{pmatrix}\begin{pmatrix}\beta & \mathbf{f}^T\\ 0 & R\end{pmatrix}\begin{pmatrix}\mathbf{v}_1^T\\ V_1^T\end{pmatrix}+\widetilde{E}$ elde edilebilir.

4. $\begin{pmatrix}\beta & \mathbf{f}^T\\ 0 & R\end{pmatrix}$ matrisinin en küçük tekil değeri σ_{r+1} bulunur.

$$\begin{pmatrix}\beta & \mathbf{f}^T\\ 0 & R\end{pmatrix}\mathbf{z}_{r+1}=\sigma_{r+1}\mathbf{y}_{r+1}$$

5. P ve Q ortogonal matrisleri bulunur.

$$P^T\mathbf{z}_{r+1}=\mathbf{e}_1$$

P bulunduktan sonra Q matrisi

$$P^T\begin{pmatrix}\beta & \mathbf{f}^T\\ 0 & R\end{pmatrix}Q=\begin{pmatrix}\widetilde{\beta} & \tilde{\mathbf{f}}^T\\ 0 & \widetilde{R}\end{pmatrix}$$

olacak şekilde bulunur.

$$\begin{pmatrix}\tilde{\mathbf{u}} & \widetilde{U}_1\end{pmatrix}=\begin{pmatrix}\mathbf{u}_1 & U_1\end{pmatrix}P,\qquad \begin{pmatrix}\tilde{\mathbf{v}} & \widetilde{V}_1\end{pmatrix}=\begin{pmatrix}\mathbf{v}_1 & V_1\end{pmatrix}Q$$

elde edilir.

6. Eğer $\sigma_{r+1}\geq\varepsilon$ ise;

$$\overline{R}=\begin{pmatrix}\widetilde{\beta} & \tilde{\mathbf{f}}^T\\ 0 & \widetilde{R}\end{pmatrix},\quad \overline{U}_1=\begin{pmatrix}\tilde{\mathbf{u}} & \widetilde{U}_1\end{pmatrix},\quad \overline{V}_1=\begin{pmatrix}\tilde{\mathbf{v}} & \widetilde{V}_1\end{pmatrix},\quad \overline{E}=\widetilde{E}$$

bulunur. $\sigma_{r+1}<\varepsilon$ ise; $\widetilde{R}$ nın en küçük tekil değeri σ_r bulunur.

$$\widetilde{R}\mathbf{z}_r=\sigma_r\mathbf{y}_r$$

Çizelge 3.2 (devam)

Tekrar $\widehat{P}$ ve $\widehat{Q}$ ortogonal matrisleri bulunur.

$$\widehat{P}^T \mathbf{z}_k = \mathbf{e}_1$$

Elde edilen $\widehat{P}$ matrisi yardımıyla $\widehat{Q}$ matrisi bulunur ve

$$\widehat{P}^T \widetilde{R} \widehat{Q} = \widehat{R}$$

üst üçgensel matrisi elde edilir.

$$\widehat{U}_1 = \widetilde{U}_1 \widehat{P}, \qquad \widehat{V}_1 = \widetilde{V}_1 \widehat{Q}$$

olarak tanımlanır ve

$$r = r-1, \quad \widetilde{R} = \widehat{R}(1:r,1:r), \quad \widetilde{U}_1 = \widehat{U}(:,1:r), \quad \widetilde{V}_1 = \widehat{V}_1(:,1:r)$$

alınır. $\sigma_r > \varepsilon$ ise;

$$\overline{R} = \widetilde{R}, \qquad \overline{U}_1 = \widetilde{U}_1, \qquad \overline{V}_1 = \widetilde{V}_1$$

alınır.

7. Sonuç olarak

$$A = \overline{U}_1 \overline{R} \overline{V}_1^T + \overline{E}, \qquad \overline{E} = A\left(I - \overline{V}_1 \overline{V}_1^T\right)$$

bulunur. Eğer $\sigma_{r+1} \geq \varepsilon$ ise;

$$\left\|\overline{E}\right\|_F^2 = \left\|E\right\|_F^2 - \sigma^2.$$

$\sigma_{r+1} < \varepsilon$ ise;

$$\left\|\overline{E}\right\|_F^2 = \left\|E\right\|_F^2 - \sigma^2 + \sigma_{r+1}^2.$$

Cauchy interlace teoremi ile tekil değerler için $\sigma \geq \sigma_{r+1}$ durumu sağlanır. Böylece $\sigma = \sigma_{r+1}$ ve $\sigma_{r+1} < \varepsilon_m$ veya $\sigma = 0$ olmadığı sürece $\left\|\overline{E}\right\|_F < \left\|E\right\|_F$ dır.

Sonuçta deflasyon algoritması matrisin rankını veren karakterlerini yeniler, alt uzayların doğruluğunu artırır ve hata matrisinin etkisini azaltır. Kısaca ayrışımı daha keskin yapar(6,18). Ancak burada olası bir problem, deflasyondan sonra ortaya çıkabilir. $\sigma_{r+1} \geq \varepsilon$ durumuna dikkat edilirse, R matrisine satır ve sütun eklendiğinde risk durumlarıyla karşılaşılabilir. Aslında bu, URV ayrışımı için temel problemdir. Bunun çözümü için iki seçenek vardır: İlk seçenek, bu durum olduğunda SVD'nin kullanılması, diğer seçenek ise (3.30)'daki $\left\|R^{-1}\right\|_2 \leq \varepsilon^{-1}$ durumunun kaldırılması ve E üzerinde $\left\|E\right\|_2 < \varepsilon$ durumunun uygulanmasıdır.

3.4 Sayısal Uygulama

Güncelleme algoritması üstel pencereleme yöntemi kullanılarak test edilmiştir. Bu yöntemde, eski verilerin etkisini azaltmak için, veri matrisi $A \in \mathbb{R}^{m \times n}$ α $(0 < \alpha \leq 1)$ unutma faktörü ile çarpılıp $m \times n$ tipindeki A matrisine yeni veri satırı olan $\mathbf{a} \in \mathbb{R}^n$ eklenerek

$$A(t) = \begin{pmatrix} \mathbf{a}^T \\ \alpha A(t-1) \end{pmatrix} \tag{3.55}$$

t zamanlı $A(t)$ veri matrisi (3.55) eşitliğindeki gibi oluşturulur. Testler Matlab 6.5'te yapılmıştır. $A(t)$ pencere matrisinin SVD'si t adımda Matlab'ta SDV algoritması ile elde edilmiştir. Aynı zamanda

$$A(t)=U_1(t)R(t)V_1(t)^T+E(t)$$

$A(t)$ 'nin kesik URV ayrışımı olarak Çizelge 3.2'de verilen algoritma ile elde edilmiştir. Alt uzayların doğruluğu

$$\left|\sin\theta(t)\right|=\left\|V_1^T(t)Y_2(t)\right\| \tag{3.56}$$

ile karakterize edilmiştir. Hatalar daha iyi bir görüntü elde etmek için $\log 10$ tabanında çizilmiştir. $A(t)$ 'nin rankı gözlenip Matlab'ın hesapladığı SVD ile karşılaştırılmıştır. $V_1(t)$ 'nin ortogonalliği

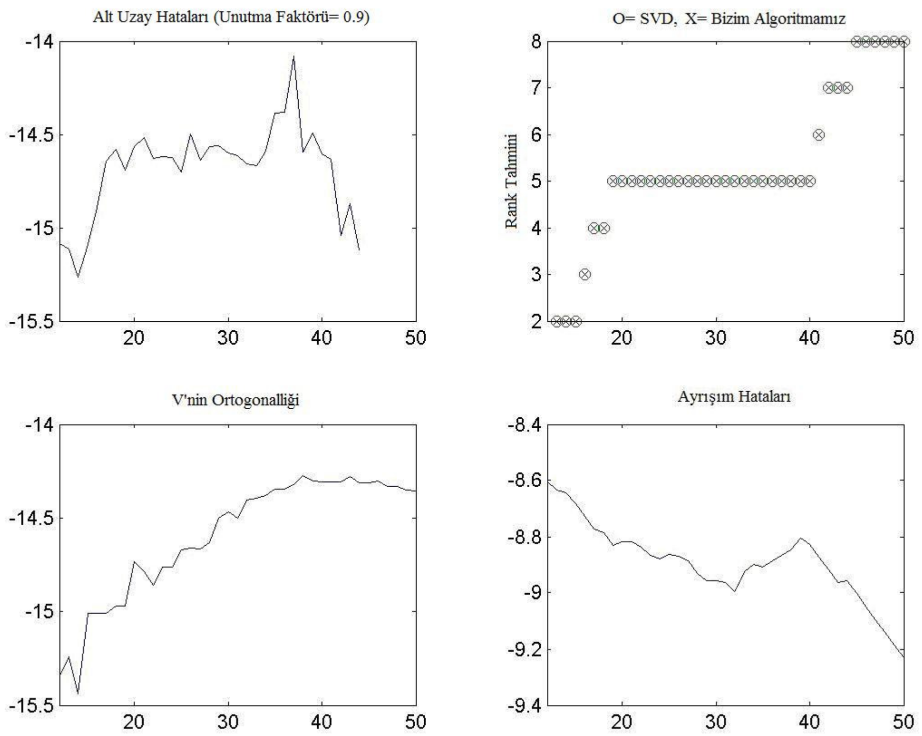
$$\left\|I_k-V_1^T(t)V_1(t)\right\|$$

ile hesaplanarak doğrulanmıştır. Ayrıca ayrışımın hatalarının her biri

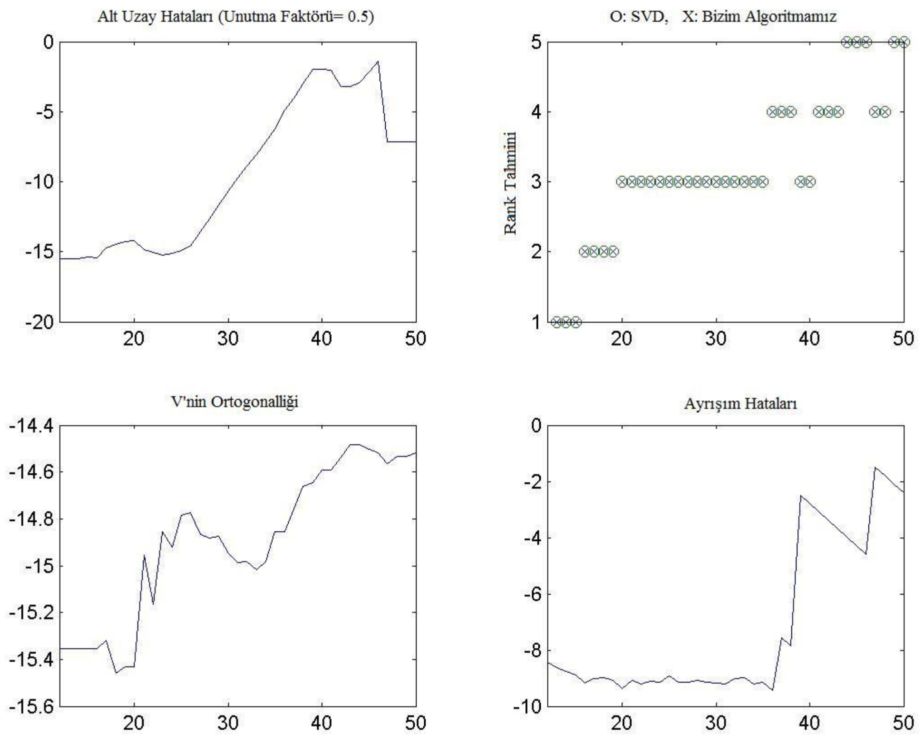
$$\left\|A(t)-U_1(t)R(t)V_1(t)^T\right\|$$

ile t adımda hesaplanmıştır[18,19].

Örnek 3.1 A, 50×8 tipinde $U(0,1)$ dağılımından rasgele üretilmiş bir matris olsun. A matrisinin rankını değiştirmek amacıyla, A 'nın rasgele seçilen 42 satırı $\eta=10^{-9}$ ile çarpılsın ve $\varepsilon\approx 10^{-8}$ olsun. Başlangıç matrisi 12×8 tipinde $A(0)$ matrisi ve unutma faktörü $\alpha=0.9$ ve $\alpha=0.5$ olarak belirlensin. Elde edilen sonuçlar sırasıyla Şekil 3.1 ve Şekil 3.2'de verilmiştir.



Şekil 3.1 $\alpha = 0.9$ için elde edilen sonuçlar



Şekil 3.2 $\alpha = 0.5$ için elde edilen sonuçlar

Şekil 3.1 ve Şekil 3.2'deki ilk çizim, bizim algoritmamız ile elde edilen alt uzay hatalarını göstermektedir. İkinci çizim ise Matlab'ın SVD algoritması ile bizim algoritmamızın hesapladığı rank tahminleri göstermektedir. Bizim algoritmamız, rank değişiminin sıklığına rağmen, rankı doğru tahmin etmiştir. Burada yatay eksen pencere adımlarını temsil etmektedir(6,18,19).

4. TARTIŞMA VE SONUÇ

Bu çalışmada kesik URV ayrışımı olarak adlandırılan URV tabanlı matris ayrışımı tanıtılmıştır. Ayrıca kesik URV ile elde edilen alt uzaylar için üst sınırlar belirlenmiştir. Bu sınırlar ve nümerik sonuçlar kesik URV ayrışımının SVD'nin alternatifi olabileceğini gösterir. Özellikle veri akışının sürekli olduğu veri matrislerinin URV ayrışımını elde etmek SVD'ye göre daha avantajlıdır.

Ayrıca güncelleme ve deflasyon algoritması da verilmiştir. Güncelleme algoritması, mevcut bilgi kullanılarak yeni veri geldiğinde veri matrisinin kesik URV'sinin hesaplanmasıdır. Güncelleme süreci sonucunda üst üçgensel matris, bozulmuş ranklı olabilir. Deflasyon algoritması ile bu özellik yeniden kazandırılır. Ayrıca hata matrisinin etkisini azaltır ve daha doğru alt uzayların hesaplanmasına olanak sağlar. Kısaca ayrışımı daha keskin kılar.

İstatistikte EKK probleminin kullanımıyla ilgili bir örnek verilmiştir. Bu örnek ile EKK probleminin istatistik için önemi ve yeri vurgulanmak istenmiştir. Hisse senetleri, üretim, yağış miktarı ile ilgili tahminlerde EKK problemi kullanılabilir. EKK probleminin çözümü için de URV ve SVD gibi ayrışımlardan yararlanılabilir.

İleriki çalışmalarda, kesik URV ayrışımından yararlanarak EKK probleminin çözümü elde edilmeye çalışılacaktır.

KAYNAKLAR

1. G.H. Golub and C.F. Van Loan, Matrix Computations Third Edition, The Johns Hopkins University Press, Baltimore, 1996.

2. A. Björck, Numerical Methods for Least Squares Problems, Society for Industrial and Applied Mathematics, Philadelphia, 1996.

3. D.S. Watkins, Fundamentals of Matrix Computations, John Wiley and Sons, Inc., Canada, 1991.

4. G.W. Stewart, SIAM J. Mat. Anal. Appl., **14**, 494(1993).

5. J.L. Barlow, H. Fu and M.K. Ng, SIAM J. Scientific Computing, **28**, 1100(2006).

6. H. Erbay, Applied Mathematics and Computations, **173**, 1300(2006).

7. G. Adams, F. Griffin and G.W. Stewart, Direction-of-Arrival Estimation Using The Rank Revealing URV Decomposition, Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, Washington, DC, 1991.

8. S. Hosur, A.H. Tewfik and D. Boley, IEEE In ICASSP, **94**, 409(1994).

9. S. Lipschutz, Linear Algebra, McGraw-Hill, Inc., Great Britain, 1974.

10. S. Kocaoğlu, ULV Ayrışımı ve Uygulamaları, Yüksek Lisans Tezi, Kırıkkale Üniversitesi, Kırıkkale, 2006.

11. S.D. Conte and Carl de Boor, Elementary Numerical Analysis Third Edition McGraw-Hill, Inc., USA, 1980.

12. R.B. Bapat, Linear Algebra and Linear Models Second Edition, Hindustan Book Agency, New Delhi, 1999.

13. F. Öztürk, Y. Akdi, H. Aydoğdu ve İ. Karabulut, Parametre Tahmini ve Hipotez Testi, Bıçaklar Kitabevi, Ankara, 2006.

14. F. Akdeniz ve F. Öztürk, Lineer Modeller, AÜFF Döner Sermaye İşletmesi Yayınları No:38, Ankara, 1996.

15. Ş. Kalaycı, SPSS Uygulamalı Çok Değişkenli İstatistik Teknikleri 2. Baskı, Asil Yayınevi, Ankara, 2008.

16. O.A. Skyes, An Introduction to Regression Analysis, The Inaugural Coase Lecture, 1993.

17. M. Moonen, P.V. Dooren and F. Vanpoucke, Lin. Alg. Appl., **188**, 549(1993).

18. H. Erbay and K. Aba, An Algorithm for Rank Estimation and Subspace Tracking, International Journal of Computer Mathematics, kabul edildi.

19. H. Erbay, Modifying Rank-Revealing Decompositions, Doktora Tezi, The Pennsylvania State University, The Graduate School, Pennsylvania, 2000.