YILDIZ TEKNİK ÜNİVERSİTESİ FEN BİLİMLERİ ENSTİTÜSÜ

Genel. Özd. Yön. ile Lin. Zam. Değiş Elek. Dev. Kut. ve Sıf. Bul.

Yüksek Lisans Tezi

Erdal Torun

1986

YILDIZ ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

GENELLEŞTİRİLMİŞ ÖZDEĞER YÖNTEMLERİ İLE LİNEER ZAMANLA DEĞİŞMEYEN ELEKTRİK DEVRELERİNİN KUTUP VE SIFIRLARININ BULUNMASI

YÜKSEK LİSANS TEZİ

ELK.MÜH. ERDAL TORUN

İSTANBUL 1986

YILDIZ ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

GENELLEŞTİRİLMİŞ ÖZDEĞER YÖNTEMLERİ İLE LİNEER ZAMANLA DEĞİŞMEYEN ELEKTRİK DEVRELERİNİN KUTUP VE SIFIRLARININ BULUNMASI

YÜKSEK LİSANS TEZİ

ELK.MÜH. ERDAL TORUN

İSTANBUL 1986

Bu çalışmada beni yönlendiren ve yardımlarını esirgemeyen sayın hocam Prof. Dr. Ahmet Dervişoğlu'na teşekkür ederim.

Erdal Torun

İÇİNDEKİLER

ÖZET

Bu çalışmada, Lineer zamanla değişmeyen bir elektrik devresi için devre fonksiyonunun kutup ve sıfırlarının bulunması incelenecektir. Devre fonksiyonunun kutup ve sıfırlarının bulunması, $Ax = \lambda Bx$ biçiminde verilen genelleştirilmiş özdeğer probleminin çözümüne eşdeğerdir. Genelleştirilmiş özdeğer probleminin çözümü QZ algoritmasıyla bulunabileceği gibi, daha ekonomik ve güvenilir bir yöntem olan, QR algoritmasıyla da bulunabilir.

Devre fonksiyonunun kutup ve sıfırları iki ayrı algoritma ile bulunur: Birinci algoritmada genelleştirilmiş düğüm denklemlerinden, ikinci algoritmada ise devrenin durum denklemlerinden elde edilen devre fonksiyonu, genelleştirilmiş özdeğer problemine indirgenerek devre fonksiyonunun kutup ve sıfırları QR algoritmasıyla bulunur.

QR algoritması ile, elemanları reel olan nxn boyutlu bir A matrisi, üniter benzerlik dönüşümlerinden yararlanarak, köşegen elemanları 1x1 ve 2x2 boyutlu altmatrisler olan blok üçgen matrise indirgenir ve matrisin özdeğerleri bulunur.

A,B kare matrisler ve B tekil bir matris olmak üzere, standart özdeğer probleminin genel bir hali olan ve $Ax = \lambda Bx$ biçiminde ifade edilen, genelleştirilmiş özdeğer probleminin çözümü, QZ algoritmasıyla katsayı matrislerinden dolaysız olarak bulunabilir.

SUMMARY

In this study, the problem of determining the poles and zeros of a lineer system is considered. Also point out that this problem is mathematically equivalent to the generalized eigenvalue problem of the from $\lambda Ax=Bx$, which will be solved by means of the QR algorithm and the QZ algorithm

Two practical and reliable algorithms are presented. The first algorithm is more general than the second and can be effectively used to obtain the poles and zeros of a lineer system using generalized nodal equations. The second algorithm assumes that the state-space equations of the system are known and takes full advantage of their special from. Both algorithms also offer considerable savings in computational effort when applied to multiple-input multiple-output systems.

The purpose of the QR algorithm of the second chapter is to compute all the eigenvalues of a given NXN matrix A_1 with real elements. The QR transformation, based on unitary transformations, produce form A_1 a sequence of similar matrices (A_s) which, in nearly all cases, tend to a form which the eigenvalues may be recovered easily.

The QZ algorithm is presented for the solution of the matrix eigenvalue problem $Ax=\lambda Bx$ with general square matrices A and B. Particular attention is paid to the degeneracies which result when B is sinqular. The algorithm is a generalization of the QR algorithm and reduces to it when B=I.

1. GİRİŞ

Lineer zamanla değişmeyen (LZD) ve tek bir çözümü olan bir elektrik devresinin, genelleştirilmiş düğüm denklemleri veya durum denklemleriyle elde edilen devre fonksiyonu $H(s)=P(s)/Q(s)$ biçimindedir ve rasyonel fonksiyonun pay ve payda polinomlarının kökleri, sırasıyla, sıfır ve kutupları verir. Devre fonksiyonunun kutup ve sıfırları, pay ve payda polinomlarından, kök bulma yöntemleri ile [17],[18] bulunabilir. Yuvarlatma hatalarından dolayı polinomun katsayılarında oluşacak çok küçük değişimler, kök bulma yöntemi ile, köklerin doğru olarak bulunabilmesini zorlaştırır. Kök bulma yöntemlerine karşılık, $H(s)$ devre fonksiyonu, genelleştirilmiş özdeğer problemine indirgenerek kutup ve sıfırlar bulunabilir.

Bu çalışmada, LZD bir elektrik devresinde, devre fonksiyonunun kutup ve sıfırlarının bulunması için iki algoritma verilmiştir.

- Birinci algoritmada, devre denklemleri

$$M\dot{x} = Ax + Bu$$

$$y = R\dot{x} + Cx + du$$

biçiminde yazılır ve bu denklemlerden yararlanarak elde edilen genelleştirilmiş özdeğer probleminin çözümü ile devre fonksiyonunun kutup ve sıfırları bulunur. Genelleştirilmiş özdeğer problemi QZ algoritmasıyla dolaysız olarak çözülebilecegi gibi, ekonomik bir yöntem olan ve $Ax=\lambda x$ biçiminde verilen standart özdeğer probleminin çözümü için geliştirilen, QR algoritmasıyla da bulunabilir.

- İkinci algoritmada ise, devre denklemleri

$$\dot{x} = Ax + Bu$$

$$y = Cx + du$$

biçiminde yazılan denklemler yardımı ile, elde edilen özdeğer probleminden, devre fonksiyonunun kutup ve sıfırları bulunur.

İkinci bölümde, elemanları reel olan nxn boyutlu bir A matrisinin özdeğerlerinin QR algoritmasıyla bulunması incelenecektir. Bir A matrisinin özdeğerleri, karekteristik polinomdan kök bulma yöntemleri ile bulunabilir fakat, yuvarlatma hatalarından dolayı polinomun katsayılarında meydana gelen çok küçük değişimler, özdeğerlerin doğru olarak bulunmasını engellerler. Bu nedenle, bir A matrisinin özdeğerlerini karekteristik polinomdan bulmak tercih edilmez. Özdeğerleri bulmanın başka bir yolu, benzerlik dönüşümlerinden yararlanarak, matrisin belirli bir biçime getirmektir. Nümerik yöntemler, özdeğerleri bulunacak A matrisini benzerlik dönüşüm - leri ile, daha basit bir yapıda olan ve özdeğerleri kolayca bulunabilen As matrisine indirgemeye dayanır. Genel bir matrisin özdeğerlerini bulmak için, 1958 yılında Rutishauser tarafından geliştirilen ve üniter olmayan dönüşümlere dayanan LR algoritmasında karşılaşılan zorluklar, beraberinde yeni arayışları getirmiştir. 1961 yılında Francis, yapı olarak LR algoritmasına benzeyen fakat tamamen üniter benzerlik dönüşümlerini kullanarak, matris özdeğerlerini bulan QR algoritmasını geliştirmiştir.

QR algoritması ile, üst Hessenberg biçimine indirgenmiş $A \in \mathbb{C}^{nxn}$ genel bir matrisin özdeğerleri bulunur. Üst Hessenberg matris biçimende verilen H matrisi, köşegen elemanları 1x1 ve 2x2 boyutlu alt matrisler olan, blok üst üçgen matrise indirgenir.

Üçüncü bölümde, $A,B \in R^{nxn}$ genel ve B tekil matris olmak üzere,

$$Ax = \lambda Bx$$

biçiminde verilen genelleştirilmiş özdeğer probleminin çözümü, 1973 yılında C.B. Moler ve G.W. Stewart tarafından geliştirilen Q algoritmasıyla bulunması incelenmiştir [13].

Bölüm 4 ve bölüm 5'de, LZD bir elektrik devresi için devre fonksiyonunun kutup ve sıfırlarının, genelleştirilmiş özdeğer yöntemlerinden yararlanarak, bulunması incelenmiştir. Genelleştirilmiş düğüm denklemlerinden ve durum denklemlerinden yararlanarak geliştirilen algoritmalarla, devre fonksiyonunun kutup ve sıfırları gerçek çözüme yakın ve ekonomik olarak bulunabilir.

2. BİR MATRİSİN ÖZDEĞERLERİNİN QR ALGORİTMASI İLE BULUNMASI

2.1. Özdeğer Problemi

Bilindiği gibi, elemanları kompleks sayılar olan nxn boyutlu bir A matrisinin özdeğerleri,

$$\det(\lambda I-A)=\lambda^n+d_{n-1}\lambda^{n-1}+\ldots+d_1\lambda+d_0 \qquad (2.1)$$

karekteristik polinomunun kökleridir. Matrisin boyutu 4'den küçükse özdeğerler analtik olarak bulunabilir fakat, $n\geq 4$ için polinomun kökleri nümerik yöntemlerle bulunur ve yuvarlatma hataları çok etkili olabilir. Örneğin,

$$P_{20}(\lambda)=(\lambda-1)(\lambda-2)(\lambda-3)\ldots(\lambda-20)$$

olarak verilen polinomun gerçek kökleri $\lambda_1=1, \lambda_2=2, \ldots, \lambda_{20}=20$ olmasına karşılık, $P_{20}(\lambda)$'nın katsayıları çok küçük bir değişime uğradığında yeni polinom $Q_{20}(\lambda)$'nın köklerini doğru olarak bulmak çok zordur. Gerçekten,

$$Q_{20}(\lambda)=P_{20}(\lambda)-2^{-23}\lambda^{19}$$

polinomunun köklerinden sadece 10 tanesi reel, diğerleri eşlenik köklerdir[1]. Bu nedenle bir A matrisinin özdeğerlerini karekteristik polinomdan yararlanarak bulmak tercih edilmez. Özdeğerleri bulmanın başka bir yolu, benzerlik dönüşümlerinden yararlanarak matrisi belirli bir biçime getirmektir. Bu yöntemle ve bilgisayarlardan yararlanılarak özdeğerler kısa sürede bulunabilir. Nümerik yöntemler, özdeğerleri bulunacak A matrisini

$$B=H^{-1}A\,H\ , \qquad A,B,H\in\mathbb{C}^{n\times n} \qquad (2.2)$$

benzerlik dönüşümü ile, daha basit bir yapıda olan ve özdeğerleri kolayca bulunabilen B matrisine indirgemeye dayanırlar. (2.2) ifadesinde, B matrisi A'ya benzerdir ve benzerlik dönüşümleri altında özdeğerler değişmez. Benzerlik dönüşümleri ile A'dan elde edilen B aşağıdaki matris biçimlerinden birisine sahipse, özdeğerler hemen bulunabilir.

1. $B \in \mathbb{C}^{n \times n}$ alt veya üst üçgen matris ise, karekteristik polinom

$$P(\lambda) = \det(\lambda I - B) = (\lambda - b_{11})(\lambda - b_{22}) \cdot \cdot \cdot (\lambda - b_{nn})$$

olur. Buna göre özdeğerler köşegen elemanlarıdır.

2. $B \in \mathbb{C}^{2 \times 2}$ ise, karekteristik polinom

$$P(\lambda) = |\lambda I - B| = \begin{bmatrix} \lambda - b_{11} & -b_{12} \\ -b_{21} & \lambda - b_{22} \end{bmatrix}$$

$$= \lambda^2 - (b_{11} + b_{22})\lambda - b_{12}b_{21} + b_{11}b_{22}$$

biçiminde ifade edilen ikinci dereceden bir polinomdur, dolayısıyla özdeğerler analtik olarak bulunabilir.

3. $B \in \mathbb{C}^{n \times n}$,

$$B = \begin{bmatrix} B_{11} & B_{12} & \cdots & B_{1n} \\ 0 & B_{22} & \cdots & B_{2n} \\ \cdot & 0 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdots & B_{nn} \end{bmatrix}$$

biçiminde blok üst üçgen matris ise, karekteristik polinom

$$P(\lambda) = |\lambda I - B| = |\lambda I_{11} - B_{11}| \cdot |\lambda I_{22} - B_{22}| \cdot \cdot \cdot |\lambda I_{nn} - B_{nn}|$$

olarak yazılabilir; buna göre B'nin özdeğerleri B_{11}, B_{22}, . . B_{nn} matrislerinin özdeğerleridir.

4. $B \in \mathbb{C}^{n \times n}$ köşegen matris ise, özdeğerler köşegen eleman-

larıdır.

5. Bilindiği gibi,

$$A^* := (\bar{A})^T \tag{2.3}$$

ile tanımlanan A^*'a A'nın adjointi denir. Öteyandan $A^* = A$ ise, A'ya Hermisyen matris adı verilir. Aşağıda açıklanacağı gibi Hermisyen bir matris, üniter dönüşümlerle, köşegen biçimine indirgenebilir. A matrisi Hermisyen bir matris olsun, P_s üniter ve kolayca elde edilebilen bir matris olmak üzere

$$A_{s+1} := P_s^* A_s P_s \tag{2.4}$$

ile verilen A_{s+1} matrisi $s \to \infty$ için köşegen biçimine indirgenebilir. Hermisyen matrislerin özdeğerlerinin reel olması ve matris elemanlarındaki küçük değişimlerin özdeğerleri fazla etkilememesi (iyi durumlu problem), problemin çözümünde büyük kolaylıklar sağlar.

Yukarıda verilen özel matrislerin özdeğerleri kolayca bulunabildiği halde, genel bir matrisin özdeğerlerinin bulunmasında zorluklarla karşılaşılabilir. Genel bir matrisin özdeğerlerini bulmak için 1958 yılında Rutishauser tarafından geliştirilmiş olan ve üniter olmayan dönüşümlere dayanan LR algoritması, özdeğer probleminin bilgisayarlarla çözümünde önemli bir aşama olmuştur [1], [3]. LR algoritmasında özdeğerleri bulunacak A matrisi önce, A = LR şeklinde ayrıştırılır. Burada L köşegen elemanları 1 olan birim alt üçgen matris, R ise bir üst üçgen matristir. İkinci adımda, üniter olmayan L dönüşüm matrisi ile A_1 = RL dönüşümü yapılır. Eğer başlangıç matrisi $A_1 =: A$ alınırsa, LR algoritması ardışıl olarak aşağıdaki gibi tanımlanır:

$$A_{s-1} = L_{s-1}R_{s-1} \quad , \; A_s = R_{s-1}L_{s-1} \tag{2.5}$$

Bu dönüşümde, A_s matrisi A_{s-1}'e benzerdir ve belirli koşullar sağlanırsa $s \rightarrow \infty$ için

$$L_s \longrightarrow I \ , \ R_s \longrightarrow A_s \longrightarrow \begin{vmatrix} \lambda_1 & & & \\ & \lambda_2 & & \times \\ & & \ddots & \\ 0 & & & \lambda_n \end{vmatrix}$$

olur. LR algoritması simetrik matrislerin özdeğerlerinin bulunmasında iyi sonuçlar vermesine rağmen, simetrik olmayan matrislerin özdeğerlerinin bulunmasında iki sorunla karşılaşılır[2]: Birincisi, büyük boyutlu matrisler için (2.5)'de verilen ayrışımın doğru olarak her zaman bulunması mümkün değildir. İkinci olarak, büyük boyutlu matrisler için, yöntem ekonomik değildir. LR algoritmasındaki bu sorunlar, dönüşüm matrisini üniter almakla giderilebilir. 1961 yılında Francis, yapı olarak LR algoritmasına benzeyen fakat, tamamen üniter benzerlik dönüşümlerini kullanarak $A \in \mathbb{C}^{n \times n}$ matrisini benzer olan A_s üst üçgen matrisine dönüştürmüştür [4]. Üniter dönüşümlerle gerçekleştirilen bu yönteme QR algoritması denir.

Bu bölümde, genel bir $A \in \mathbb{R}^{n \times n}$ matrisinin özdeğerlerinin QR algoritmasıyla bulunması incelenecektir.

2.2.Dönüşüm Matrisleri

$A \in \mathbb{C}^{n \times n}$ matrisini benzerlik dönüşümleri ile daha basit bir yapıya indirgeyen temel dönüşüm matrisleri,

1. Elemanter matrisler,
2. Elemanter üniter matrisler,
3. Elemanter üniter hermisyen matrisler

olmak üzere üç grupta toplanabilir.

Elemanter Matrisler

Dönüşüm matrislerinin ilki olarak aşağıdaki şekilde tanımlanan matrislere elemanter matrisler, bu matrislerle yapılan dönüşümlere elemanter benzerlik dönüşümü denir.

a)

$$M_i = \mathrm{diag}(1, \ . \ . \ ., k, 1, \ . \ . \ . \ , 1) \qquad (2.6)$$

matrisi ile bir matrisin satır veya sütunları sabit bir k değeri ile çarpılır. A matrisi M_i ile soldan(sağdan) çarpılırsa, i inci satır(sütun) k sabiti ile çarpılmış olur.

b)

$$S_{ji} = \begin{array}{c} \begin{array}{c} (i) \\ \begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & k & \ddots 1 & & \\ & & & & \ddots & \\ & & & & & 1 \end{bmatrix} \end{array} \end{array} (j) \qquad (2.7)$$

matrisi, birim matrisin i inci satırının k ile çarpılıp j inci satırına eklenmesiyle elde edilen matristir. Bilindiği gibi A matrisi soldan(sağdan) S_{ji} ile çarpılırsa, A'nın i inci satırının(sütununun) k ile çarpılıp j inci satırına (sütununa) eklenmesine denktir. Kısaca, A'nın i inci satırı (sütunu) $r_i(c_i)$ ise,

$$r_j \longleftarrow (kr_i + r_j) \quad , \quad c_j \longleftarrow (kc_i + c_j)$$

olur.

c) P_{ij} elemanter permütasyon matrisi, I birim matristen $r_j \longleftarrow r_i$, $r_i \longleftarrow r_j$ ile elde edilen matristir ve $P_{ij}A =: \hat{A}$ matrisi, A'nın i inci ve j inci satırlarının yerdeğiştirilmesi ile elde edilen matristir.

Elemanter Üniter Matrisler

Bir nxn birim matrisin, (i,i), (i,j), (j,i) ve (j,j) elemanları, α, β, $\theta \in \mathbb{R}$ olmak üzere,

$$r_{ii} = e^{i\alpha} \cos\theta \quad , \quad r_{ij} = e^{i\beta} \sin\theta \quad ,$$

$$r_{ji} = -e^{-i\beta} \sin\theta \quad , \quad r_{ij} = e^{-i\alpha} \cos\theta \qquad (2.8)$$

biçiminde ise, bu matris üniterdir ve düzlem döndürme (Plane rotation) matrisi adını alır. R(i,j) ile gösterilen böyle bir elemanter üniter matriste $\alpha = \beta = 0$ ise, R(i,j) ortogonal bir matris olur, yani

$$R^T R = RR^T = I \tag{2.9}$$

dir. Eğer,

$$r_{ii} = r_{jj} = c \quad , \quad c = \mathrm{Cos}\,\theta$$

$$r_{ij} = -r_{ji} = s \quad , \quad s = \mathrm{Sin}\,\theta$$

alınırsa, ortogonal R(i,j) matrisi

$$R(i,j) = \begin{array}{cc} & \begin{array}{ccccc} & (i) & & (j) & \end{array} \\ & \left[\begin{array}{ccccc} I & & & & \\ & c & & s & \\ & & I & & \\ & -s & & c & \\ & & & & I \end{array}\right] \begin{array}{c} \\ (i) \\ \\ \\ (j) \end{array} \end{array} \quad , \; i<j \tag{2.10}$$

biçimini alır. $R(i,j) \in R^{nxn}$ ile herhangi bir $x \in R^{nx1}$ vektöründen elde edilen $\hat{x} := Rx$ vektöründe, i inci eleman reel pozitif ve j inci eleman sıfır yapılabilir, x'in diğer elemanları bu dönüşümden etkilenmezler. Gerçekten,

$$(Rx)_j = -sx_i + cx_j$$

ifadesinde,

$$s = x_j / r \quad , \quad c = x_i / r \quad , \quad r = |x_i|^2 + |x_j|^2$$

olarak seçilirse, $(Rx)_j = 0$ olur. Öte yandan,

$$\begin{aligned} (Rx)_i &= cx_i + sx_j \\ &= |x_i|^2/r + |x_j|^2/r \\ &= r \geqslant 0 \end{aligned}$$

dir, $r = 0$ ise, $c = 1$ ve $s = 0$ olur. O halde x vektörünü

$$ke_1 = (k,0,. . . ,0)^T$$

vektörüne dönüştürmek için x, (1,2), (1,3), . . . ,(1,n) düzlemlerinde (n-1) tane düzlem döndürme matrisi ile soldan çarpılır. (1,i) düzleminde c ve s değerleri, i inci eleman sıfıra indirgenecek şekilde seçilir.

Benzer şekilde $A\in\mathbb{R}^{n\times n}$ matrisi için,

$$\hat{A}=RA=\begin{array}{cl}\left[\begin{array}{cccc} a_{11} & a_{12} & \cdot\ \cdot\ \cdot & a_{1n} \\ a_{21} & a_{22} & \cdot\ \cdot\ \cdot & a_{2n} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ (ca_{i1}-sa_{j1}) & (ca_{i2}-sa_{j2}) & \cdot\ \cdot & (ca_{in}-sa_{jn} \\ a_{i-1,1} & a_{i-1,2} & \cdot\ \cdot & a_{i-1,n} \\ \cdot & \cdot & \cdot\ \cdot & \cdot \\ (-sa_{i1}-ca_{j1}) & (-sa_{i2}-ca_{j2}) & \cdot\ \cdot & (-sa_{in}-ca_{jn} \\ a_{j-1,1} & a_{j-1,2} & \cdot\ \cdot & a_{j-1,n} \\ \cdot & \cdot & \cdot\ \cdot & \cdot \\ a_{n1} & a_{n2} & \cdot\ \cdot & a_{nn} \end{array}\right] & \begin{array}{l} \\ \\ \\ \\ (i \\ \\ \\ (j \\ \\ \\ \\ \end{array}\end{array}$$

dönüşümünde A'nın yalnızca i inci ve j inci satırları etkilenir. Gerçekten,

$$c = a_{ii}/(a_{ii}^2 + a_{ji}^2)^{1/2} \quad , \quad s = a_{ji}/(a_{ii}^2 + a_{ji}^2)^{1/2}$$

olarak alınırsa, $\hat{a}_{ji}$ sıfıra indirgenmiş olur.

Üniter Hermisyen Matrisler

Diğer bir elemanter dönüşüm matrisi, üniter ve hermisyen bir matris olan,

$$P := I - 2ww^* \tag{2.11}$$

matrisidir. Burada, $w\in\mathbb{C}^{n\times 1}$ dir. Buna göre,

$$\begin{aligned} PP^* &= (I-2ww^*)(I-2ww^*) \\ &= I-4ww^* + 4w(w^*w)w^* = I \end{aligned}$$

olur, yani P hermisyen ve üniter bir matristir. $w\in\mathbb{R}^{n\times 1}$ ise P matrisi ortogonal ve simetrik olur. Herhangi bir $x\in\mathbb{C}^{n\times 1}$ vektöründe,

$$y = Px = x-2w(w^*x) \tag{2.12}$$

ile elde edilen bir y matrisi,

$$y^*y = x^*P^*Px = x^*x \quad , \quad x^*y = x^*Px \qquad (2.13)$$

özelliklerini taşır. (2.13)'de P hermisyen olduğundan, $(\bar{x}x)$ R olur. Böyle bir dönüşümde, $y^*y = x^*x$ ve $(x^*x) \in R$ koşulunu sağlayan P matrisi için w, (y-x) doğrultusundadar ve

$$w = (y-x)/((y-x)^*(y-x))^{1/2} \qquad (2.14)$$

biçiminde ifade edilebilir[2]. Gerçekten, (2.14) eşitliği, (2.12)'de yerine konursa

$$(I-2ww^*)x = x - \frac{2(y-x)(y-x)^*x}{(y-x)^*(y-x)} \qquad (2.15)$$

elde edilir. Burada,

$$\begin{aligned} 2(y-x)^*x &= 2(y^*x - x^*x) \\ &= y^*x + x^*y - x^*x - y^*y \\ &= -(y-x)^*(y-x) \end{aligned} \qquad (2.16)$$

dir. (2.16) eşitliği (2.15)'de yerine konursa,

$$(I-2ww^*)x = x+(y-x) = y$$

olur.

2.3.QR Algoritması

QR algoritması $A \in \mathbb{C}^{nxn}$ matrisini üniter benzerlik dönüşümleri ile, köşegen elemanları 1x1 ve 2x2 boyutlu altmatrisler olan, üst üçgen matrise indirgeyen iteratif bir yöntemdir. Matematiksel olarak QR algoritması şu şekilde tanımlanır: Q üniter bir matris ve R üst üçgen bir matris olmak üzere, verilen bir $A_1 \in \mathbb{C}^{nxn}$ matrisi için, $A_1 = Q_1R_1$ ayrışımı yapılır. İkinci adımda, A_1'e benzer olan $A_2 = R_1Q_1$ matrisi bulunur. Bu işlem iteratif olarak tanımlanırsa,

$$s = 1,2, \ldots \begin{cases} A_s := Q_sR_s \\ A_{s+1} := R_sQ_s = Q_s^*A_sQ_s \end{cases} \qquad (2.17)$$

ile elde edilen A_s ardışıl matrisleri A_1'e benzerdir.

Teorem 2.1: $A\in\mathbb{C}^{nxn}$ tekil olmayan bir matris ise, $A = QR$ ayrışımı vardır ve Q üniter, R üst üçgen matristir. Ayrıca, R'nin köşegen elemanları $r_{ii}\in\mathbb{R}^+$ ise, ayrışım tektir[1].

Tanıt : A, Q, R'nin k inci sütunları, sırasıyla, a^k, q^k ve r^k olsun.

$k = 1$ için,

$$a^1 = Qr^1 = r_{11}q^1$$

olur. Çünkü, $r^1 = r_{11}e^1$ dir. O halde r_{11} elemanı seçilirse, q^1 belirlenmiş olur:

$$q^1 = r_{11}^{-1}\, a^1 . \qquad (2.18)$$

Bununla beraber, Q'nun üniter olması istendiğinden, (2.18)'in iki yanının Euclidean normu alınırsa,

$$\begin{aligned} 1 &= \|q^1\| \\ &= \|r_{11}^{-1}\, a^1\| = |r_{11}^{-1}|\cdot\|a^1\| \qquad , \; |r_{11}| = \|a^1\| \end{aligned}$$

bulunur, $r_{11}\in\mathbb{R}^+$ için, $r_{11} = \|a^1\|$ olarak seçilebilir. Çünkü, A tekil olmadığından, $\|a\| \neq 0$ dır.

$k = 2$ için,

$$a^2 = ar^2 = r_{12}\, q^1 + r_{22}q^2 \qquad (2.19)$$

elde edilir. Q'nun üniter olması istendiğinden,

$$(q^1)^* a^2 = r_{12}(q^1)^* q^1 + r_{22}(q^1)^* q^2 = r_{12}$$

olur. Burada q^1 belirlendiğinden, r_{12} de belirlenmiş olur. Öte yandan, (2.19) bağıntısından yararlanarak

$$\begin{aligned} \|a^2 - r_{12}q^1\| &= \|r_{22}q^2\| \\ &= |r_{22}|\cdot\|q^2\| = |r_{22}| \end{aligned}$$

elde edilir. Ayrıca, $r_{22}\in\mathbb{R}^+$ için, $r_{22} = \|a^2 - r_{12}q^1\|$ olarak seçilebilir, A tekil olmadığından, $r_{22} \neq 0$ dır.

$k = s$ için,

$$a^s = Qr^s = r_{1s}q^1 + r_{2s}q^2 + \ldots + r_{ss}q^s$$

bulunur. $r_{1s}, r_{2s}, \ldots, r_{ss}$ bilinirse, q^s de aşağıdaki ifade ile belirlenir:

$$q^s = r_{ss}^{-1}(a^s - r_{1s}q^1 - r_{2s}q^2 - \ldots - r_{s-1,s}q^{s-1}) .$$

Q üniter olduğundan, $i=1,2,\ldots,s-1$ için, r_{is},

$$(q^i)^* a^s = r_{1s}(q^i)^* q^1 + r_{2s}(q^i)^* q^2 + \ldots + r_{ss}(q^i)^* q^s$$
$$= r_{is}$$

biçiminde ifade edilebilir. Son olarak r_{ss},

$$\| r_{ss} q^s \| = \| a^s - r_{1s}q^1 - r_{2s}q^2 - \ldots - r_{s-1,s}q^{s-1} \|$$
$$= : \| p^s \|$$

olur, $r_{ss} \in \mathbb{R}^+$ olması istendiğinden, $r_{ss} = \| p^s \|$ olarak seçilebilir.

Teorem 2.1'in tanıtı, A'nın (2.17) biçiminde ayrıştırılabildiğini de belirlemektedir. Oluşturulan R_s üst üçgen matrisi, Q_s ile sağdan çarpılırsa s inci iterasyon tamamlanmış olur. QR algoritmasının tanım bağıntılarından, aşağıdaki bağıntılar elde edilebilir:

$$A_{s+1} = Q_s^* A_s Q_s$$
$$= Q_s^* Q_{s-1}^* A_{s-1} Q_{s-1} Q_s$$
$$= Q_s^* \ldots Q_2^* Q_1^* A_1 Q_1 Q_2 \ldots Q_s$$

ve

$$Q_1 Q_2 \ldots Q_s A_{s+1} = A_1 Q_1 Q_2 \ldots Q_s . \qquad (2.20)$$

Buna göre, $s=1,2,\ldots$ için, A_s üniter olarak A_1'e benzerdir.

$$P_s := Q_1 Q_2 Q_3 \ldots Q_s \;, \quad U_s := R_s R_{s-1}, \ldots R_1$$

olarak alınırsa,

$$P_s U_s = Q_1 \ldots Q_{s-1}(Q_s R_s) R_{s-1} \ldots R_1$$
$$= Q_1 \ldots Q_{s-1} A_s R_{s-1} \ldots R_1$$

olur. (2.20) eşitliğinden,

$$P_s U_s = A_1 Q_1 \ldots Q_{s-1} R_{s-1} \ldots R_1$$
$$= A_1 P_{s-1} U_{s-1}$$
$$= A_1^2 P_{s-2} U_{s-2} = \ldots$$
$$= A_1^s \qquad (2.21)$$

bulunur. P_s üniter, U_s üst üçgen matris olduğundan ve Teorem 2.1'e göre, tekil olmayan bir matrisin üniter-üçgen ayrışımının tek olmasından yararlanarak aşağıdaki teorem yazılabilir.

Teorem 2.2: $A \in \mathbb{C}^{nxn}$ tekil olmayan bir matris ise,

$$A_{s+1} = P_s^* A_1 P_s \quad , \quad P_s = Q_1 \cdot \cdot \cdot Q_s$$

dönüşümünde P_s matrisi, A_1^s'in üniter-üçgen ayrışımından elde edilebilir.

2.4. QR Algoritmasının Yakınsaması

(2.17) ifadesi ile tanımlanan QR algoritmasında A_{s+1}'in hangi koşullar altında üst üçgen matris biçimine yakınsadığını,

$$A_{s+1} = P_s^* A_1 P_s \quad , \quad s \longrightarrow \infty$$

ifadesinde, P_s'nin davranışları belirler. Wilkinson[2], yakınsamanın tanıtını basit bir gerçeğe dayandırır: I birim matrisin Q ve R faktörleri de birim matristir. A_s, birim matrise dönüştürülebilinirse Q ve R faktörleri de birim matris olur. Benzer şekilde, A_s ardışıl matrisleri üst üçgen matris biçimine dönüşmeye eğimli ise, algoritma mutlak yakınsar.

2.4.1. Özdeğerlerin Ayrık Ve Sıralı Olması

Teorem 2.3: A_1 matrisinin özdeğerleri,

$$|\lambda_1| > |\lambda_2| > \cdot \cdot \cdot > |\lambda_n| > 0 \qquad (2.22)$$

şeklinde ve L birim alt üçgen matris, U üst üçgen matris olmak üzere,

$$X^{-1} = Y = LU$$

üçgen ayrışımı varsa QR algoritması mutlak yakınsar. Burada X, sütunları A_1'in özvektörleri olan kare matristir.

Tanıt: Özdeğerleri (2.22) eşitliğindeki gibi olan A_1

matrisi s inci iterasyonda,

$$A_1^s = X\,\mathrm{diag}(\lambda_i^s)\,X^{-1}$$
$$= X\,D^s\,Y \quad , \quad D^s := \mathrm{diag}(\lambda_i^s) \qquad (2.23)$$

biçiminde ifade edilir. Q, R, L ve U matrisleri

$$X = QR \quad , \quad Y = LU$$

ifadeleriyle tanımlansın. X'in sütunları bağımsız özvektörlerden oluştuğundan, QR ayrışımı vardır. Y'nin LU üçgen ayrışımının varlığı [7,sf.19]'da verilmiştir. Buna göre Y_k, Y' nin ilk k satır ve ilk k sütununun oluşturduğu matris olmak üzere, $k = 1,2,\ldots,n-1$ için, $\det Y_k \neq 0$ ise L ve U vardır. X ve Y'nin ayrışımı (2.23)'de yerine konursa,

$$A_1^s = Q\,R\,D^s\,L\,U$$
$$= Q\,R(D^s\,L\,D^{-s})D^s\,U$$

elde edilir. Burada $(D^s\,L\,D^{-s})$, birim alt üçgen matristir ve

$$D^s\,L\,D^{-s} = I + B_s$$

olarak yazılabilir. B_s'nin (i,j) elemanı, $i>j$ için $l_{ij}(\lambda_i/\lambda_j)^s$ ve $i\leq j$ için sıfırdır. $|\lambda_1|>|\lambda_2|>\cdots>|\lambda_n|>0$ olduğundan, (λ_i/λ_j) oranı 1'den küçük ve $s\longrightarrow\infty$ için $B_s\longrightarrow 0$ olur. O halde,

$$A_1^s = Q\,R\,(I+B_s)\,D^s\,U$$
$$= Q\,(I+R\,B_s\,R^{-1})\,R\,D^s U$$
$$= Q\,(I+F_s)\,R\,D^s U$$

olur ve $s\longrightarrow\infty$ için $F_s\longrightarrow 0$ dır. $(I+F_s)$ matrisi, üniter matris $\tilde{Q}_s$ ve üst üçgen matris $\tilde{R}_s$ çarpanlarına ayrıştırılabilir. $F_s\longrightarrow 0$ için $\tilde{Q}_s$ ve $\tilde{R}_s$'in I birim matrise yakınsadıkları açıktır. Sonuç olarak,

$$A_1^s = (Q\,\tilde{Q}_s)(\tilde{R}_s\,R\,D^s\,U) \qquad (2.24)$$

ifadesinde A_1^s'in $(Q\tilde{Q}_s)$ faktörü üniter ve $(\tilde{R}_s R D^s U)$ faktörü üst üçgen matristir. A_1^s tekil olmayan bir matris ise, (2.21)' de verilen P_sU_s ayrışımında $P_s = Q\tilde{Q}_s$ olur ve P_s, Q üniter matrise yakınsar. Öte yandan, $(\tilde{R}_s RD^s U)$ üst üçgen matris olmasına rağmen köşegen elemanları pozitif olmayabilir. D_1 ve D_2 üniter köşegen matrisleri ile, $(\tilde{R}_s RD^s U)$'nun köşegen elemanları pozitif yapılabilir.

$$D_1 = |D| D_1 \quad , \quad U = D_2(D_2^{-1} U)$$

alınırsa, $(D_2^{-1} U)$'nun köşegen elemanları pozitif olur. Bu ifadeler (2.24)'de yerine konursa,

$$A_1^s = Q \tilde{Q}_s D_2 D_1^s \left((D_2 D_1^s)^{-1} \tilde{R}_s R(D_2 D_1^s) | D|^s (D_2^{-1} U)\right)$$

elde edilir ve eşitliğin sağ tarafında büyük parantezin içindeki kısım, köşegen elemanları pozitif olan üst üçgen matristir.

2.4.2. Özdeğerlerin Ayrık Ve Sıralı Olmaması Hali

Yukarıdaki tanıttan görüldüğü gibi, Y matrisinin LU biçiminde çarpanlara ayrılabilmesi yalnızca mutlak yakınsamayı değil, aynı zamanda A matrisinin özdeğerlerinin köşegen üzerinde (2.22) ile verilen biçimde sıralı olduğunu da belirtir.

$[y_{ij}]_{kk}$ matrisinde, herhangi bir k için $\det Y_k = 0$ ise Y'nin LU ayrışımı yoktur. Fakat $\det Y \neq 0$ ise, öyle bir P permütasyon matrisi vardır ki, (PY) matrisi LU çarpanlarına ayrıştırılabilir [7]. PY = LU ayrışımında, L'nin köşegen altındaki elemanlarından bazıları sıfırdır. (2.23) ifadesi, P matrisi dikkate alınarak tekrar yazılırsa,

$$\begin{aligned} A_1^s &= X D^s Y \\ &= X D^s P^T L U \\ &= X P^T (P D^s P^T) L U \end{aligned} \qquad (2.25)$$

elde edilir. $(P D^s P^T)$'nin köşegen elemanları sıralı olmayan λ_i^s özdeğerleridir.

$$X P^T = Q R \quad , \quad P D^s P^T = D_3^s$$

olarak yazılırsa,

$$A_1^s = Q\,R\,D_3^s\,L\,U$$
$$= Q\,R\,(D_3^s\,L\,D_3^{-s})D_3^s\,U$$

olur. D_3'ün köşegen elemanları, λ_i^s , sıralı olmadıklarından, $(D_3^s\ L\ D_3^{-s})$'nin ilk bakışta I birim matrise yakınsamadığı görülür. Bununla beraber, D_3^{-s}'nin köşegen elemanları $\lambda_{i_1}, \lambda_{i_2}, \ldots$. λ_{i_n} ile gösterilirse, p>q için, $(D_3^s\ L\ D_3^{-s})$'nin (p,q) elemanı $(\lambda_{i_p}/\lambda_{i_q})^s \cdot l_{pq}$ olur. Burada $i_p < i_q$ için, $l_{pq} = 0$ olur[2]. Ayrıca,

$$D_3^s\ L\ D_3^{-s} = I + B_s$$

olarak yazılabilir ve $s \longrightarrow \infty$ için $B_s \longrightarrow 0$ olur. O halde P_s matrisi, XP^T'nin ayrışımından bulunan Q üniter matrise mutlak yakınsar.

Sonuç olarak, farklı özdeğerlere sahip bir A matrisi için, aşağıdaki teorem yazılabilir [4]:

Teorem 2.4: Herhangi bir A tekil olmayan matrisin özdeğerleri ayrık ise, QR dönüşümü altında ana köşegen altındaki elemanlar sıfıra, üstündeki elemanlar sabit değerlere ve köşegen elemanları özdeğerlere yakınsar.

2.4.3. Mutlak Değerce Eşit Özdeğerlerin Bulunması Hali

Yukarıda görüldüğü gibi, $\det Y_k = 0$ olması yakınsamayı etkilemediğinden, k = 1,2,. . . ,n-1 için, $\det Y_k \neq 0$ olduğunu varsayalım. (2.23) ifadesinde Y yerine LU konursa,

$$A_1^s = X\,D^s\,L\,U$$
$$= X(D^s\,L\,D^{-s})D^s\,U$$

elde edilir. A_1'in bazı özdeğerleri mutlak değerce eşit, diğerleri farklı olsun. Yani, $|\lambda_r| = |\lambda_{r+1}| = \ldots = |\lambda_t|$ ve diğer özdeğerler mutlak değerce farklı büyüklüktedir. $(D^s\ L\ D^{-s})$'nin köşegen altındaki (i,j) elemanları $(\lambda_i/\lambda_j)^s \cdot l_{ij}$, $s \longrightarrow \infty$ için,

$$(\lambda_i/\lambda_j)^s l_{ij} \longrightarrow \begin{cases} l_{ij} & t \geq i > j \geq r \text{ için} \\ 0 & \text{diğer durumlarda} \end{cases}$$

olur.

Mutlak değerce eşit olan özdeğerler katlı ise,

$$D^s L D^{-s} = \tilde{L} + B_s \quad , \quad (B_s \longrightarrow 0)$$

olur. Burada $\tilde{L}$,

$$\tilde{L} = \begin{cases} l_{ij} & t \geq i > j \geq r \text{ için} \\ I & \text{diğer durumlarda} \end{cases}$$

biçiminde ifade edilir ve sabitleştirilmiş birim alt üçgen matris adını alır. $X\tilde{L} = Q\,R$ ayrışımı varsa,

$$\begin{aligned} A_1^s &= Q\,R\,(I + \tilde{L}^{-1} B_s)\,D^s\,U \\ &= Q\,(I + R\tilde{L}^{-1} B_s R^{-1})R\,D^s\,U \\ &= Q\,(I + F_s)R\,D^s\,U \\ &= (Q\,\tilde{Q}_s)(\tilde{R}_s R\,D^s\,U) \end{aligned}$$

olur. Burada $\tilde{Q}_s\tilde{R}_s$, $(I + F_s)$'nin üniter-üçgen ayrışımıdır. Q üniter matrisi, X'in ayrışımından değilde, $X\tilde{L}$'nin ayrışımından bulunmuştur. Fakat $X\tilde{L}$'nin sütunları A_1'in bağımsız özvektörlerinden oluştuğundan, X matrisinden farklı değildir. Bu bakımdan P_s, Q üniter matrise yakınsar. O halde, katlı özdeğerler yakınsamayı engellemezler.

Mutlak değerce eşit olan özdeğerler katsız ise, ($\lambda_i \neq \lambda_j$, $|\lambda_i| = |\lambda_j|$), $\tilde{L}$ birim alt üçgen matrisi yukarıda tanımlanan şekilde olmasına rağmen, köşegen haricindeki sıfırdan farklı elemanlar sabit değildir, yani, $\lambda_i = |\lambda_i| \exp(i\theta_i)$ ise,

$$\tilde{l}_{ij} = l_{ij} \exp(is(\theta_i - \theta_j))$$

olur. $X\tilde{L}$ matrisi, r'den t'ye kadar olan sütunları dışında sabittir. Bu sütunlar, X'in karşı düşen sütunlarının lineer kombinezonu olarak ifade edilebilirler. Bu nedenle, $X\tilde{L}$'nin üniter-üçgen ayrışımında Q üniter matrisin r'den t'ye kadar olan sütunları, X'in karşı düşen sütunlarının lineer kombinezonlarıdır, geriye kalan sütunları sabittir. Sonuç olarak P_s, r'den t'ye kadar olan sütunlar haricinde Q matrisine yakınsar. Yalnızca, mutlak değerce eşit olan ayrık özdeğerler yakınsamayı engeller fakat, bu durum pratikte önemli bir zorluk getirmez [2,sf.540] .

Yakınsama bakımından en sakıncalı durum, $A \in R^{nxn}$ matrisinin eşlenik özdeğerlerinin olması halidir. Bu durumda A_k, köşegen üzerinde merkezlenmiş, reel özdeğerlere karşı düşen 1x1 ve eşlenik özdeğerlere karşı düşen 2x2 boyutlu altmatrisleri içeren blok üst üçgen matrise dönüştürülübilir. Dolayısıyla, köşegen altındaki bazı elemanlar sıfırdan farklı olabilirler. A_1, r tane mutlak değerce eşit kompleks özdeğere sahipse, A_s özdeğerleri köşegen üzerinde merkezlenmiş ve bu r tane değere yakınsayan r boyutlu bir matristir.

Örnek 2.1:

$$A_1 = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

biçiminde verilen bir A_1 matrisine QR algoritması uygulanırsa, A_1 matrisi üst üçgen matris biçimine indirgenemez. Gerçekten, $Q_1 = A_1$ ve $R_1 = I$ olarak ayrıştırılırsa,

$$A_2 = R_1 Q_1 = A_1$$

olur ve r = 0,1,2,3 için özdeğerler e(exp(1/2i r)) dir.

Örnek 2.1'deki biçimde, çok büyük boyutlu bir matrisin özdeğerleri standart QR algoritması ile kolayca bulunamamaktadır. Bu zorluk, ileride görüleceği gibi, orijin öteleme ile (shift of origin) giderilebilir ve A_1 blok üst üçgen matrise dönüştürülebilir.

2.5. Genel Bir Matrisin Hessenberg Matrise İndirgenmesi

Genel bir $A \in R^{nxn}$ matrisin özdeğerleri QR algoritması ile bulunursa, işlem sayısı ve yakınsama hızı bakımından bazı problemlerle karşılaşılır. A matrisi; Hessenberg matris, üçköşegenli matris veya bant matris gibi basit yapıda ise işlem sayısı azalır. İleride ayrıntılı olarak görüleceği gibi $A \in R^{nxn}$ dolu bir matrisin özdeğerleri QR algoritmasıyla bulunursa, bir iterasyon için, işlem sayısı n^3 mertebesinde olurken, $A \in \mathbb{R}^{nxn}$ Hessenberg matrisi için bu sayı n^2 dir. Bu bakımdan, QR algoritmasına başlangıç olarak, $A \in \mathbb{R}^{nxn}$ dolu matrisinin üst Hessenberg matris biçimine indirgenmesi incelenecektir. A matrisinin, $i>j+1$ $(i+1<j)$ için (i,j) elemanı sıfır ise, A üst (alt) Hessenberg matristir denir. 2.2. alt bölümünde verilen dönüşüm matrisleri ile farklı yöntemler kullanılarak, A matrisi H Hessenberg matrisine dönüştürülebilir. Bu yöntemler; elemanter benzerlik dönüşümleri, Givens yöntemi ve Householder yöntemidir. Bu yöntemlerle, verilen bir matrisi Hessenberg matrise indirgemek için gerekli işlem sayısı, sırasıyla, $5n^3/6$, $10n^3/3$ ve $5n^3/3$ dür[1], [2]. Elemanter benzerlik dönüşümleri ile indirgemenin, işlem sayısı bakımından, Givens ve Householder yöntemlerine üstünlüğü olmasına rağmen yuvarlatma hatalarından dolayı H Hessenberg matrisinin elemanları büyükce hatalar içerebilirler. Givens yönteminde ise, yardımcı köşegen (sub-diagonal) altındaki elemanlar, düzlem döndürme matrisi ile, tek tek sıfırlandığından işlem sayısı yüksektir. Bu bakımdan, ortogonal ve simetrik dönüşüm matrisleri kullanan Householder yöntemi ile $A \in \mathbb{R}^{nxn}$ matrisinin Hessenberg matris biçimine indirgenmesi incelenecektir. Diğer yöntemlerle indirgeme işlemi,[1],[2], [10]no'lu referanslarda ayrıntılı olarak verilmiştir.

Householder yönteminin en önemli avantajı, dönüşüm matrisinin ortogonal olması dolayısıyla, dönüştürülen matrisin durumunun kötüleşmemesidir. Yani, H Hessenberg matrisi A'nın durumunu (iyi durumlu veya kötü durumlu) korur. Diğer bir avantajı, A simetrik bir matris ise H Hessenberg matrisi de simetriktir.

Householder yönteminde dönüşüm matrisi,

$$Q = I - (2uu^T/\|u\|^2) \quad , \quad u \in \mathbb{R}^{nx1} \tag{2.26}$$

ortogonal matristir. Q öyle seçilmelidir ki, her adımda A'nın bir sütununun yardımcı köşegen altındaki elemanları sıfıra indirgensin. Dolayısıyla Q, (n-2) adım için ayrı ayrı belirtilmelidir. (2.26) ifadesinde Q'nun ortogonal $(QQ^T = I)$ olduğu,

$$Q^T = Q \quad , \quad \|u\| = \|u\|_2 = (u\,u^T)^{1/2}$$

alınarak gösterilebilir. Gerçekten,

$$Q^TQ = QQ = I - (4u\,u^T/\|u\|^2) + (4(uu^T)(uu^T)/\|u\|^4) \qquad (2.27)$$

dir ve (2.27)'deki

$$\begin{aligned}(uu^T)(uu^T) &= u(u^T u)u^T \\ &= u(\|u\|^2)u^T \\ &= \|u\|^2(uu^T)\end{aligned}$$

yerine konursa, $Q^TQ = I$ olur. Q'nun tanımlanması tamamen u vektörünün seçilmesine bağlıdır. $v = (v_1, \ldots, v_{k-1}, v_k, v_{k+1}, \ldots, v_n)^T$ vektörü için $w = Qv$ dönüşümünde w'nin özelliklerinin aşağıdaki gibi olması istensin:

1. w'nin ilk (k-1) elemanı $v_1, v_2, \ldots, v_{k-1}$,
2. Geriye kalan (n-k) elemanı sıfır olsun.

O halde u vektörü şu şekilde tanımlanır[10]:

1. $u_1 = u_2 = \ldots = u_{k-1} = 0$, (2.28)
2. $u_{k+1} = v_{k+1}, \quad u_{k+2} = v_{k+2}, \ldots, u_n = v_n$,
3. $u_k = v_k - s$, $s = \pm (v_k^2 + v_{k+1}^2 + v_{k+2}^2 + \ldots + v_n^2)^{1/2}$

Verilen u ve v vektörlerinden,

$$\|u\|^2 = 2\,s^2 - 2\,s\,v_k \quad , \quad u^T v = s^2 - s\,v_k = \|u^2\|/2$$

olur ve

$$\begin{aligned}Qv &= (I - 2uu^T/\|u\|^2)v = v - (2u^T v/\|u\|^2)u \\ &= v - u = w\end{aligned}$$

dir.

A matrisi, $A = [A_1, A_2, \ldots, A_n]$ sütun formunda yazılırsa, $QA = [QA_1, QA_2, \ldots, QA_n]$ olur. Hessenberg biçimine indirgemede ilk adım, QA_1'in $j = 3,4, \ldots, n$ için j inci elemanlarını sıfır yapmaktır. O halde A_1 sütun vektörü için (2.28) ifadesiyle tanımlanan u,

$$u_1 = 0 \,, \; u_2 = a_{21} \pm (a_{21}^2 + a_{31}^2 + \ldots + a_{n1}^2)^{1/2} ,$$
$$u_3 = a_{31} \;, \ldots, \; u_n = a_{n1}$$

olarak seçilir ve H_1 = QAQ matrisinin ilk sütununda yardımcı köşegenin altındaki elemanlar sıfırdır. H_1 = QAQ matris çarpımlarını yapmak yerine aşağıdaki basitleştirilmiş ifadeler kullanılır. $1 \leqslant k \leqslant n$ için, QA'nın k inci sütunu QA_k dır. Herhangi bir y vektörü için,

$$Qy = (I-2uu^T/\|u\|^2)y = y-(2u^T y/\|u\|^2)u$$
$$=: y-cu \quad , \quad c := 2u^T y/\|u\|^2$$

ifadesinde, c sabitinin belirlenmesi ile Qy dönüşümü kolaylıkla hesaplanabilir. Benzer şekilde, QA çarpımı için önce

$$\beta = 2/\|u\|^2$$

hesaplanır ve $2 \leqslant k \leqslant n$ için,

$$\alpha_k = u^T A_k$$

değeri ile

$$QA_k = A_k-(\beta\alpha_k u) \qquad (2.29)$$

bulunur. QA_1'in yardımcı köşegen altındaki elemanları sıfır olacağından, QA_1 çarpımına gerek yoktur.

B:= QA olarak alınırsa, BQ çarpımının basitleştirilmiş ifadesi yazılabilir. $C := B^T$ ve $C = [C_1, C_2, \ldots, C_n]$ denirse,

$$(BQ)^T = Q^T B^T = Q\ C$$
$$= [QC_1, QC_2, \ldots, QC_n]$$

ve QC'nin sütunları,

$$QC_k = C_k-(2u^T C_k/\|u\|^2)u = C_k-d_k u \qquad (2.30)$$

olur. Burada $1 \leqslant k \leqslant n$ için; $\beta := 2/\|u\|^2$, $\gamma_k := u^T C_k$ ve $d_k := \beta\gamma_k$ dır. C'nin sütunları B'nin satırları ve QC'nin sütunlarıda BQ'nun satırlarıdır. O halde γ_k, $1 \leqslant k \leqslant n$ için B'nin k inci satırı ile u'nun skaler çarpımıdır.

(2.28) ifadesinde s'nin değeri hem "+" hem de "-" için doğrudur fakat, karışıklığı önlemek için, v_k'nın işaretinin tersi alınır [10]. Birinci adım sonunda, H_1=QAQ dönüşümü ile, ilk sütun istenilen yapıdadır ve H_1 matrisi A'ya benzerdir.

Aynı işlemlere devam ederek,

$$H_1 = \begin{bmatrix} a'_{11} & a'_{12} & \cdot & \cdot & \cdot & a'_{1n} \\ a'_{21} & a'_{22} & \cdot & \cdot & \cdot & a'_{2n} \\ 0 & a'_{32} & \cdot & \cdot & \cdot & a'_{3n} \\ 0 & a'_{42} & \cdot & \cdot & \cdot & a'_{4n} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & a'_{n2} & \cdot & \cdot & \cdot & a'_{nn} \end{bmatrix}$$

matrisinin 2, 3, . . ., n-2 inci sütunlarının yardımcı köşegen altındaki elemanları sıfıra indirgenir. (n-2) adım sonunda, A'ya benzer olan H Hessenberg matrisi elde edilir.

2.6. QR Algoritmasının Pratikte Uygulanması

$A_s \in R^{n\times n}$ matrisinin $Q_s R_s$ ayrışımında, Q_s ortogonal matris olmalıdır. Teorik olarak, A_s'nin sütunlarının Schmidt ortonormalizazyonu ile, bu şekilde ortogonal-üçgün ayrışımı yapılabilir [2] fakat, Q_s ortogonal özelliğini kaybeder. Pratikte çoğu kez Q_s yerine, transpozu olan Q_s^T aşağıdaki gibi bulunur:

$$Q_s^T A_s = R_s$$

Q_s^T matrisi, A_s' nin üçgenleştirme işleminde [2], Givens yönteminde kullanılan düzlem döndürme matrislerinin çarpımı ile yada Householder yönteminde Ortogonal simetrik matrislerin çarpımı ile belirlenir. Daha sonra,

$$A_{s+1} = R_s Q_s$$

dönüşümü, Q_s^{T}' nin transpozu ile R_s sağdan çarpılarak elde edilir.

A matrisinin özdeğerleri QR algoritmasıyla bulunursa, işlem sayısı n^3 mertebesinde olur fakat, A Hessenberg matris biçimine indirgendikten sonra işlem sayısı n^2 mertebesine iner. QR algoritmasına başlangıç olarak Hessenberg matris biçiminin alınması, ileride görüleceği gibi, başka kolaylıklar da sağlar.

Teorem 2.5: A_1 üst Hessenberg matris biçiminde ise, Q_1 de Hessenberg matris biçimindedir ve QR dönüşümü altında bu yapı bozulmaz.

Tanıt: Aşağıda verilen A_1 Hessenberg matrisi için,

$$A_1 = \begin{bmatrix} x & x & x & x & x \\ x & x & x & x & x \\ 0 & x & x & x & x \\ 0 & 0 & x & x & x \\ 0 & 0 & 0 & x & x \end{bmatrix},$$

$A_1 = Q_1 R_1$ ortogonal-üçgen ayrışımında Q_1'ın j inci sütunu q_j, A_1'in ilk j sütunlarının lineer kombinezonudur. A_1' den açıkca görülür ki, q_j ile a_j aynı yapıdadır. Dolayısıyla Q_1 Hessenberg matris biçimindedir. İterasyonun ikinci adımında,

$$A_2 = R_1 Q_1$$

ifadesinde, R_1 üst üçgen matris olduğundan A_2' nin j inci satırı, Q_1'in j'den N'e kadar olan satırlarının lineer kombinezonudur. O halde Q_1 Hessenberg matris olduğundan, A_2 de aynı biçimindedir.

Genel olarak işlemlerde, A_1 Hessenberg matrisinin yardımcı köşegen elemanlarının sıfırdan farklı olduğu göz önüne alınırsa da, bazı elemanlar sıfıra yakın bir değerde olabilirler. Her hangi bir eleman yeteri kadar küçük ise, sıfır olarak kabul edilebilir ve o noktada matris, altmatrislere bölünerek daha küçük boyutlu Hessenberg matrislerle işlem yapılır. Aşağıdaki A_s matrisinde

$$A_s = \left[\begin{array}{cc:cccc} x & x & x & x & x & x \\ x & x & x & x & x & x \\ \hdashline & \epsilon & x & x & x & x \\ & & x & x & x & x \\ & & & x & x & x \\ & & & & x & x \end{array}\right] = \left[\begin{array}{c:c} B & C \\ \hdashline 0 & D \end{array}\right]$$

ϵ çok küçük bir değer ise, önce D altmatrisinin sonrada B' nin özdeğerleri bulunur. Burada, $\epsilon \leqslant 2^{-t} \| A \|_{\infty}$ olarak alınır ve t hane sayısı (mantis) dır. İterasyonun her adımında, yardımcı köşegen elemanları kontrol edilir. (r+1,r) konumundaki eleman ϵ mertebesinde ise, önce sağ alt köşedeki (n-r) boyutlu matrisin, daha sonrada sol üst köşedeki alt matrisin özdeğerleri bulunur. r = n-1 ise, D'nin boyutu 1 dir ve özdeğeri bellidir. Bu durumda A_s'nin son satır ve sütunu atılarak (n-1) boyutlu matrisin özdeğerleri araştırılır. r = n-2 ise, D'nin boyutu 2 olur ve ikinci dereceden polinom çözümü ile 2 tane özdeğer bulunur. Bu durumda, A_s'nin son 2 satır ve sütunu çıkarılarak (n-2) boyutlu matrisin özdeğerleri bulunur.

Göz önüne alınması gereken diğer bir durum, A Hessenberg matrisinin (r+1,r) ve (r+2,r+1) konumundaki elemanların çok küçük olduğu durumdur. Örneğin; n = 6 , r = 2 olan Hessenberg biçimi aşağıdaki gibi ise,

$$A = \left[\begin{array}{cc:cccc} x & x & x & x & x & x \\ x & x & x & x & x & x \\ & \epsilon_1 & x & x & x & x \\ & & \epsilon_2 & x & x & x \\ & & & x & x & x \\ & & & & x & x \end{array}\right] = \left[\begin{array}{c:c} X & Y \\ \hdashline E & W \end{array}\right]$$

A matrisi ϵ_1 ve ϵ_2 değerlerine göre altmatrislere bölünür. Gösterilebilir ki; W'nin herhangi bir özdeğeri μ , aynı zamanda A matrisinden (r+2,r) konumundaki $\epsilon_1 \epsilon_2/(a_{r-1,r-1} - \mu)$ değeri ile farklı, $\hat{A}$ matrisinin de bir özdeğeridir [2] . $(\epsilon_1 \epsilon_2)$ değeri ihmal edilebilecek kadar küçük ise, μ aynı zamanda A matrisinin de bir özdeğeridir. O halde, A matrisine QR algoritması uygulanırsa, oluşturulan A_s ardışıl matrislerinde $s \longrightarrow \infty$ için

$$a_{j+1,j} \, a_{j,j-1} \longrightarrow 0 \qquad (j = 2,3, \, . \, . \, ,n-1) \qquad (2.31)$$

ise, QR algoritması mutlak yakınsar [1] . Yakınsama için yardımcı köşegen elemanlarının hepsinin sıfır olması gerekmez. (2.31) bağıntısını sağlayan çiftlerden birisinin sıfıra yakısaması yeterlidir.

2.7. Orijin Öteleme

QR algoritmasında, başlangıç matrisi olarak üst Hessberg matrisi alınmasına rağmen, yakınsama bazı matrisler için yavaştır. 2.4'de görüldüğü üzere, $s \longrightarrow \infty$ için $a_{ii}^{s} \longrightarrow \lambda_i$ olur ve lineer yakınsama,

$$a_{ii}^{s} = \lambda_i + O(r_i^{s}) \quad , \quad a_{i+1,i}^{s} = O(r_i^{s}) \qquad (2.32)$$

olarak tanımlanabilir [5]. Burada r_i yakınsama oranıdır ve

$$r_i := \text{maks}\left\{ |\lambda_i/\lambda_{i-1}| \, , \, |\lambda_{i+1}/\lambda_i| \right\} \quad , \quad \lambda_1/\lambda_0 = 0 \quad , \lambda_{n+1}/\lambda_n = 0$$

dir. Yakınsama hızı r_i çok küçük olmadıkca, A_s'nin üst üçgen matrise indirgenmesi çok yavaş olacaktır. Genel bir A_s matrisinin $i>j$ için a_{ij}^{s} elemanının, $|\lambda_i/\lambda_j|$ oranı küçüldükce, sıfıra yakınsaması hızlı olur. Hessenberg matrisi için sıfıra yakınsaması gereken yalnızca $(r+1,r)$ konumundaki elemanlardır. A matrisinin özdeğerlerini bulmak yerine $[A-kI]$ matrisinin özdeğerleri bulunursa, $[A-kI]$'nın $a_{n,n-1}^{s}$ elemanı $((\lambda_n-k)/(\lambda_{n-1}-k))$ yakınsama oranına bağlı olarak sıfıra yakınsar. Bu nedenle, k değeri ile yakınsama hızı kontrol edilebilir ve $k=\lambda_n$ ise $a_{n,n-1}$ bir iterasyon sonunda sıfır olur. Bu açıklamalara dayanarak, verilen k_s orijin öteleme değerleri için, QR algoritması yeniden tanımlanabilir ve bu işleme orijin öteleme veya genelleştirilmiş QR dönüşümü denir:

$$s = 1,2,\ldots \left\{ \begin{aligned} A_s - k_s I &= Q_s R_s && (2.33) \\ A_{s+1} &= R_s Q_s + k_s I \; . && (2.34) \end{aligned} \right.$$

A_{s+1} matrisinin A_s'e benzer olduğu, (2.33) ve (2.34) ifadelerinden açık olarak görülmektedir. Gerçekten, (2.33) ifadesinde

$$R_s = Q_s^T (A_s - k_s I) \qquad (2.35)$$

bağıntısı, (2.34)'de yerine konursa

$$A_{s+1} = Q_s^T (A_s - k_s I) Q_s + k_s I$$

$$A_{s+1} = Q_s^T A_s Q_s - k_s I + k_s I$$
$$= Q_s^T A_s Q_s \qquad (2.36)$$

elde edilir. (2.33), (2.34) ve (2.36) bağıntılarından,

$$A_s = (Q_1 Q_2 \ldots Q_{s-1})^T A_1 Q_1 Q_2 \ldots Q_{s-1}$$

ve

$$(Q_1 Q_2 \ldots Q_s)(R_s \ldots R_2 R_1) = (A_1 - k_1 I) \ldots (A_1 - k_s I)$$

olur.

$$P_s := Q_1 Q_2 \ldots Q_s \quad , \quad U_s := R_s \ldots R_2 R_1$$

olarak alınırsa,

$$P_s U_s = \prod_{i=1}^{s} (A_1 - k_i I) = : M^s \qquad (2.37)$$

elde edilir. Bu sonuca göre aşağıdaki teorem yazılabilir.

Teorem 2.6: Genelleştirilmiş QR algoritması , k_s öteleme değerleri ile, A matrisine uygulanırsa,

$$A_{s+1} = P_s^T A P_s$$

dönüşümünde $P_s = Q_1 \ldots Q_s$ ortogonal matrisi, M^s'nin ortogonal-üçgen ayrışımından elde edilir.

Pratik olarak, yakınsama hızının artması k_s değerlerinin uygun olarak seçilmesine bağlıdır. Özdeğerler mutlak değerce farklı büyüklüklerde ise, $a^s_{n,n-1}$'nin sıfıra ve a^s_{nn}'nin λ_n'e yakınsaması beklenir. O halde, $|a^s_{n,n-1}|$ belirli bir ϵ değerinin altına düştüğünde, $k_s = a^s_{nn}$ olarak seçilmesi yakınsama bakımından tercih edilir. Gerçekten, $|a^s_{n,n-1}| \leqslant \epsilon$ ise $k_s = a^s_{nn}$ için, bir iterasyon sonunda

$$a^{s+1}_{n,n-1} = O(\epsilon^2)$$

olur. Wilkinson [2, sf.508] , k_s'nin bu şekilde seçilebileceğini bir örnekle göstermiştir. Genel olarak, $|a^s_{n,n-1}| \leqslant \epsilon$ ise a^s_{nn} bir özdeğer olarak alınır ve matrisin son satırı ile son sütunu atılarak, geriye kalan özdeğerler (n-1) boyutlu matristen bulunur. Boyutu indirgenmiş matris de Hessenberg biçimindedir ve aynı işlemlerle diğer özdeğerler teker teker bulunur.

Özdeğerler mutlak değerce farklı büyüklükte değilse, $a^s_{n-1,n-2}$ sıfıra yakın bir değer olabilir. Bu durumda, A Hessenberg matrisi

$$A = \left[\begin{array}{c:cc} A_{11} & & A_{12} \\ \hdashline & a_{n-1,n-1} & a_{n-1,n} \\ 0 & & \\ & a_{n,n-1} & a_{nn} \end{array}\right]$$

biçiminde bölünür ve sağ alt köşedeki 2x2 altmatristen λ_{n-1} ve λ_n özdeğerleri bulunur. Daha sonra, A matrisinin son iki satırı ve sütunu atılarak (n-2) boyutlu matrisin özdeğerleri bulunur.

s iterasyon için k_s öteleme değeri, A_s'nin sağ alt köşebindeki 2x2 boyutlu C_s altmatrisinin özdeğerlerine göre seçilir [4] , [5] .

$$C_s = \begin{bmatrix} a^s_{n-1,n-1} & a^s_{n-1,n} \\ a^s_{n,n-1} & a^s_{nn} \end{bmatrix}$$

altmatrisinin c_{1s} ve c_{2s} özdeğerleri reel ise, k_s öteleme değeri aşağıdaki gibi seçilebilir: a^s_{nn} değerine mutlak değerce en yakın c_{ms} (m=1,2) özdeğeri ile,

$$\left|\frac{c_{ms} - k_{s-1}}{k_s}\right| < 1/2 \quad , \quad k_0 = 0 \qquad (2.38)$$

ise, $k_s = c_{ms}$ olarak alınabilir. (2.38) eşitsizliğinin sağlanmadığı durumlarda $k_s = k_{s-1}$ olarak seçilir. c_{1s} ve c_{2s} özdeğerleri eşlenik ise, öteleme değerleri yukarıda açıklandığı gibi seçilemez. Bu durumda, birkez c_{1s} ile birkez de c_{2s} ile orijin ötelenir; bu işleme iki kademeli QR algoritması (Double QR Algorithm) denir.

2.8. İki Kademeli QR Algoritması

Yukarıda görüldüğü gibi, genelleştirilmiş QR algoritmasında C_s altmatrisinin özdeğerleri eşlenik ise, s inci iterasyon için hem c_{1s} hem de c_{2s} için orijin ötelenir. Bu açıklamalara dayanarak A_1 matrisi için iki kademeli QR algoritması,

$$A_1-k_1I = Q_1R_1$$

$$A_2 = R_1Q_1 + k_1I$$

$$A_2-k_2I = Q_2R_2$$

$$A_3 = R_2Q_2 + k_2I \qquad (2.39)$$

biçiminde tanımlanır ve (2.36) ifadesinden,

$$\begin{aligned} A_s &= Q_2^T A_2 Q_2 \\ &= Q_2^T (Q_1^T A_1 Q_1) Q_2 \\ &= (Q_1 Q_2)^T A_1 (Q_1 Q_2) \\ &= P_2^T A_1 P_2 \end{aligned} \qquad (2.40)$$

elde edilir. Yukarıdaki ifadelerden görüldüğü gibi, kompleks öteleme değerleri A_1 matrisinin elemanlarının kompleks değerlere dönüşmesine neden olmazlar. (2.39) ve (2.40) bağıntılarından yararlanarak, A_3 matrisinin ortogonal-üçgen ayrışımı aşağıdaki gibi bulunur:

$$\begin{aligned} P_2U_2 &= Q_1 Q_2 R_2 R_1 \\ &= Q_1 (A_2-k_2)R_1 \\ &= Q_1 (R_1 Q_1 + k_1I)R_1 - k_2Q_1 R_1 \\ &= Q_1 R_1 Q_1 R_1 + k_1 Q_1 R_1 - k_2 Q_1 R_1 \\ &= (Q_1R_1 + k_1I)Q_1R_1 - k_2Q_1R_1 \\ &= A_1Q_1R_1 - k_2Q_1R_1 \\ &= (A_1-k_2I)(A_1-k_1I) \\ &=: E_2(A_1) \,. \end{aligned} \qquad (2.41)$$

A_1 matrisi için $k_2 = \bar{k}_1$ olarak seçilirse, $E_2(A_1)$ de reel bir matris olur [5] . $[A_1-k_1I]$'nın ortogonal-üçgen ayrışımını bulmak yerine doğrudan $E_2(A_1)$ matrisinin ortogonal-üçgen ayrışımı bulunursa, A_2 matrisinin bulunması gerekmez; dolayısıyla kompleks sayılarla işlem yapmak zorunda kalınmaz. İki kademeli QR algoritmasıyla A_{2s+1} (s tek sayı) matrisi, A_{2s-1} 'den dolaysız olarak bulunduğundan işlem sayısı önemli ölçüde azalacaktır fakat, her adımda $E_2(A_{2s-1})$ hesaplandığından işlem sayısı artar. Buna karşılık, ileride görüleceği üzere, A_{2s-1} matrisi Hessenberg biçiminde ise, $E_2(A_{2s-1})$'in yalnızca ilk sütununun bulunması yeterlidir.

Teorem 2.7: $H = Q^{*}AQ$ dönüşümünde A matrisinin sütunları ile Q'nun ilk sütunu, H matrisi ile Q matrisinin diğer sütunlarını belirler. Burada, H yardımcı köşegen elemanları pozitif olan Hessenberg matris, Q üniter matris ve A genel bir matristir.

Tanıt: Q matrisinin j inci sütunu q_j ve H'nın j inci sütunu h_j'nin h_{ij} elemanı bilinirse,

$$QH = AQ$$

ifadesinde j inci sütunlar için,

$$\sum_{i=1}^{j+1} h_{ij} q_j = A\, q_j$$

elde edilir; yani,

$$h_{j+1,j}\; q_{j+1} = A\, q_j - \sum_{i=1}^{j} h_{ij}\; q_i =: \tilde{q}_{j+1}$$

olur. Q matrisinin sütunları birbirine dik olduğundan,

$$h_{j+1,j} = \|\tilde{q}_{j+1}\|_2$$

ve

$$q_{j+1} = (1/h_{j+1,j})\; \tilde{q}_{j+1}$$

olarak bulunur ve bütün sütunlar için,

$$j = 1,2,\ldots N \quad \begin{cases} h_{ij} = q_i^{*} A q_j\ , \\ \qquad\qquad i = 1,\ldots,j \\ \tilde{q}_{j+1} = A\, q_j - \sum_{i=1}^{j} h_{ij}\; q_j \\ h_{j+1,j} = \|\tilde{q}_{j+1}\| \\ q_{j+1} = (1/h_{j+1,j})\; \tilde{q}_{j+1} \end{cases} \tag{2.42}$$

bağıntıları elde edilir.

Teorem 2.7'ye göre,

$$A_{2s+1} = Q^T A_{2s-1} Q$$

dönüşümünde A_{2s+1} matrisinin sütunları, Q ortogonal matrisin ilk sütunu ile belirlenebilir. A_{2s-1} Hessenberg matrisin yardımcı köşegen elemanları sıfırdan farklı ise, A_{2s+1} matrisi ile A_{2s-1} aynı biçimdedir [5].

Teorem 2.8: $A \in \mathbb{R}^{nxn}$ matrisi köşegenleştirilebilen, yardımcı köşegen elemanları pozitif (sıfırdan farklı) olan Hessenberg matrisi olsun ve W üniter dönüşüm matrisi ile, başka bir Hessenberg matrisi olan $W^* A W$ matrisine dönüştürülsün. W'nin ilk sütunu,

$$w_1 = (1/\|m_1^s\|) \cdot m_1^s$$

ise,

$$A_{s+1} = P_s^* A P_s = W^* A W$$

olur. Burada m_1^s, (2.37) ifadesi ile verilen

$$P_s U_s = (A-k_s I)(A-k_{s-1} I) \ldots (A-k_1 I) = M^s$$

matrisinin ilk sütunudur.

Teorem 2.8'in tanıtı [4] ve [5] nolu referanslarda verilmiştir.

(2.37) ifadesinden,

$$M^s e_1 = P_s U_s e_1$$

$$= u_{11}^s p_1^s = m_1^s \quad , \quad u_{11} = \| m_1^s \|$$

elde edilir. Çünkü, $u_{11}^s \in \mathbb{R}^+$ ve $\| p_1^s \| = 1$ dir. O halde, $w_1 = p_1^s$ ve

$$A_{s+1} = P_s^* A P_s = W^* A W$$

olur.

Yardımcı köşegen elemanları sıfırdan farklı A Hessenberg matrisinden A' matrisine dönüşüm için iki kademeli QR algoritması, $E_2(A) = QR$ ayrışımında, Q matrisinin ilk sütunu ile tanımlanır. Bu ayrışımda R üst üçgen matris olduğundan, $E_2(A)$'nın ilk sütunu dönüşümü gerçekler. O halde,

$$E_2(\lambda) := (\lambda - k_{2s-1})(\lambda - k_{2s}) = \lambda^2 - \sigma\lambda + \delta \qquad (2.23)$$

olarak alınırsa,

$$E_2(A) = (A^2 - \sigma A + \delta I) \tag{2.44}$$

olur. Burada, $\sigma = k_{2s} \cdot k_{2s-1}$ ve $\delta = k_{2s} + k_{2s-1}$ dir. $E_2(A)$'nın ilk sütunu,

$$E_2(A)e_1 = \begin{bmatrix} a_{11}^2 + a_{21}a_{12} \\ a_{21}a_{11} + a_{22}a_{12} \\ a_{32}a_{21} \\ 0 \\ \cdot \\ \cdot \\ 0 \end{bmatrix} - \sigma \begin{bmatrix} a_{11} \\ a_{21} \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{bmatrix} + \delta \begin{bmatrix} 1 \\ 0 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{bmatrix} =: \begin{bmatrix} \gamma_1 \\ \gamma_2 \\ \gamma_3 \\ 0 \\ 0 \\ \cdot \\ 0 \end{bmatrix}$$

$$=: m_1 \tag{2.45}$$

biçiminde olur. Burada m_1, en fazla γ_1, γ_2, ve γ_3 gibi üç tane sıfırdan farklı elemana sahip olabilir. O halde her iterasyon için, $E_2(A)$'nın ilk sütunu yardımıyla, Q'nun ilk sütunu

$$q_1 = \frac{m_1}{\| m_1 \|} \tag{2.46}$$

hesaplanarak, A Hessenberg matrisinden A' matrisine dönüşüm gerçeklenebilir [4].

Francis, verilen A Hessenberg matrisine benzer olan A' matrisini, bir iterasyon için, şu şekilde bulmuştur: Önce, ilk sütunu (2.46) ifadesindeki gibi olan P_1 ortogonal matrisi tanımlanır. Daha sonra,

$$B_1 = P_1^T A P_1$$

dönüşümü yapılır. B_1 ve P_1 matrisleri, 6x6 boyutlu A Hessenberg matrisi için, aşağıdaki biçimdedir:

$$P_1 = \left[\begin{array}{ccc|c} x & x & x & \\ x & x & x & 0 \\ x & x & x & \\ \hline & 0 & & I \end{array}\right], \quad B_1 = \begin{bmatrix} x & x & x & x & x & x \\ x & x & x & x & x & x \\ x & x & x & x & x & x \\ x & x & x & x & x & x \\ 0 & 0 & 0 & x & x & x \\ 0 & 0 & 0 & 0 & x & x \end{bmatrix}.$$

İlk sütunu e_1 olan ortogonal P_2 matrisi ile,

$$B_2 = P_2^T B_1 P_2$$

dönüşümü yapılır. B_2'nin ilk sütunu Hessenberg biçimindedir fakat, 2 ve 3 üncü sütunları Hessenberg biçiminde değildir. Benzer şekilde, ilk sütunu e_1 olan P_3 ortogonal matrisi ile,

$$B_3 = P_3^T B_2 P_3$$

dönüşümünde B_3'ün ilk iki sütunu istenilen biçimdedir fakat, bu kez B_3'ün 3 ve 4 üncü sütunları Hessenberg biçiminde değildir. Bu şekilde dönüşümlere devam edilirse,

$$B_{N-1} = P_{N-1}^T \ldots P_1^T A P_1 \ldots P_{N-1}$$

matrisi Hessenberg biçiminde olur. $P_1 \ldots P_{N-1}$ çarpımının ilk sütunu q_1 olduğundan, Teorem 2.7 ve Teorem 2.8'den

$$B_{N-1} = A'$$

olur ve $P_1 \ldots P_{N-1}$ çarpımı $E_2(A)$'nın Q faktörüdür.

2.9. İki Kademeli QR Algoritmasında Benzerlik Dönüşümü

İki kademeli QR algoritmasıyla A_{2s-1} matrisinden A_{2s+1} matrisine dönüşüm, ortogonal dönüşüm matrisi

$$P_i = I - 2w_i w_i^T$$

ile yapılır. (2.45) ifadesinde, $E_2(A) \cdot e_1$'in sıfırdan farklı elemanları γ_1, γ_2 ve γ_3 ise, P_1 ortogonal matrisi aşağıdaki eşitliği sağlamalıdır:

$$P_1 e_1 := (I - 2w_1 w_1^T) e_1 = q_1 = \theta E_2(A) e_1 \qquad (2.47)$$

Burada, $\|w_1\| = 1$ ve $\theta := \|E_2(A)e_1\|^{-1}$ dir. P_1 matrisi için w_1 sütun matrisi aşağıdaki gibi seçilebilir [5]:

$$w_1 := v_1 / \|v_1\| \quad , \quad v_1^T = (1, \psi_1, \psi_2, 0, \ldots, 0) \qquad (2.48)$$

O halde,

$$P_1 = I - (2v_1 v_1^T / \|v_1\|^2)$$

ve

$$(P_1 e_1)^T = (1 - 2/\|v_1\|^2, \; -2\psi_1/\|v_1\|^2, \; -2\psi_2/\|v_1\|^2, 0 \ldots 0)$$

$$(P_1 e_1)^T =: \pm(\gamma_1/z,\ \gamma_2/z,\ \gamma_3/z,\ 0, \ldots, 0) \qquad (2.49)$$

olur. Burada,

$$z^2 := \gamma_1^2 + \gamma_2^2 + \gamma_3^2 \qquad (2.50)$$

dir. (2.48) ifadesindeki ψ_1 ve ψ_2, aşağıdaki denklemlerden bulunur:

$$-2/\|v_1\|^2 = \pm\ \gamma_1/z - 1 \ , \qquad (2.51)$$

$$(-2/\|v_1\|^2)\psi_1 = \pm\ \gamma_2/z \ , \qquad (2.52)$$

$$(-2/\|v_1\|^2)\psi_2 = \pm\ \gamma_3/z \ . \qquad (2.53)$$

(2.51) ifadesi,

$$\alpha := 2/\|v_1\|^2 = 1 \pm (\gamma_1/z)$$
$$= 1 + \gamma_1/z \qquad (2.54)$$

biçiminde yazılabilir. (2.54) ifadesinde, α, "-" ve "+" işaretleri için değişmez. Bu nedenle sadece "+" işareti alınacaktır. (2.54) ifadesi, (2.52) ve (2.53)'de yerine konursa,

$$\psi_1 = \gamma_2/(\gamma_1 + z) \quad , \quad \psi_2 = \gamma_3/(\gamma_1 + z) \qquad (2.55)$$

bulunur. Yardımcı köşegen elemanları sıfırdan farklı ve Hessenberg biçiminde olan altmatrislerle işlem yapıldığından, $\gamma_3 \neq 0$ dır.

Pratikte, P_1 kare matrisini oluşturmaya ve

$$B_1 = P_1^T A P_1$$

matris çarpımına gerek yoktur. Matris çarpımına gerek olmadığını görmeden önce,

$$B_2 = P_2^T B_1 P_2$$

dönüşümünde, P_2 matrisinin yapısına bakmakta yarar vardır. P_2 ortogonal matrisi, B_1'in ilk sütununu Hessenberg biçimine dönüştürür ve ilk sütunu e_1 dir. Bu nedenle, (P_1P_2)'nin ilk sütunu q_1 dir. O halde, P_2 dönüşüm matrisi ile B_2'nin ilk sütununun

$$(x,\ x,\ 0,\ \ldots,0)^T = B_2 e_1 = P_2^T B_1 e_1 =: P_2^T b_1$$

biçiminde olması istenir. P_1 ortogonal ve simetrik olduğundan, (2.47) ifadesine göre,

$$e_1 = P_1 q_1 \quad \text{veya} \quad e_1 = P_1 m_1 \tag{2.56}$$

olur. Burada, $m_1^T = (\gamma_1, \gamma_2, \gamma_3, 0, \ldots, 0)^T$ ve $\theta = \|m_1\|$ dir. P_2 de simetrik bir matris olduğundan,

$$B_2 e_1 = P_2 b_1 \tag{2.57}$$

biçiminde yazılabilir. (2.57) ifadesinde her iki tarafın ilk satırları çıkarılarak, (n-1) boyutlu ve (2.56) ifadesine benzer biçimde bir vektör elde edilebilir. Öyleyse, (2.48)'de verilen v_1 vektöründe bütün elemanlar bir satır aşağıya kaydırılarak, P_1 matrisini oluşturmak için yapılan işlemler P_2 için de yapılır. Yani,

$$P_2 = I - 2v_2 v_2^T / \|v_2\|^2 \quad , \quad v_2^T = (0, 1, \psi_1, \psi_2, 0, \ldots, 0)$$

olur. Bütün P_i matrisleri, γ_1, γ_2 ve γ_3 değerleri ile yukarıdaki gibi bulunur. Son adımda, yani

$$A' = B_{N-1} = P_{N-1}^T B_{N-2} P_{N-1}$$

olduğu zaman, $\gamma_1 = \gamma_2 = \gamma_3 = 0$, $z=0$ ve $\alpha = 2$ olur.

Sonuç olarak, (2.50),(2.54) ve(2.55)'de verilen z, α, ψ_1, ψ_2 değerleri ile,

$$B_s = P_s^T B_{s-1} P_s \qquad (B_0 = A)$$

dönüşümü aşağıdaki gibi verilebilir:

$$s=1,2,\ldots,N-1 \left\{ \begin{array}{l} v_s = e_s + \psi_1 e_{s+1} + \psi_2 e_{s+2} \quad , \\ B'_{s-1} = (I - v_s(\alpha v_s)^T) B_{s-1} \quad , \\ B_s = B'_{s-1}(I - v_s(\alpha v_s^T)) \quad . \end{array} \right. \tag{2.58}$$

veya basitleştirilmiş olarak,

$$\begin{aligned} r^T &:= \alpha v_s^T B_{s-1} \quad , \\ B'_{s-1} &= B_{s-1} - v_s r^T \quad , \qquad t := \alpha B'_{s-1} v_s \\ B_s &= B'_{s-1} - t\, v_s^T \end{aligned} \tag{2.59}$$

olur.