BÖLÜM 3

3. REGRESYON İÇİN MATRİS VE VEKTÖR CEBRİ

Bölüm 2 de, doğrusal regresyon tek değişkenli basit model olarak ele alınarak açıklanmıştı. Bölüm 4 de ise çok değişkenli (k değişkenli) model için giriş yapılacaktır. Çok değişkenli modelde değişken sayısına paralel olarak tahminlenmesi gereken parametre sayısı da artmaktadır. Bunun sonucu olarak işlemler için bazı hesaplama zorlukları ortaya çıkmaktadır. İşte bu hesaplama zorluklarını giderebilmek amacı ile kitabın bu bölümünden itibaren matris ve vektör işlemleri kullanılacaktır. Doğal olarak regresyon analizinde kullanılacak olan bu matris ve vektör işlemlerinin tanıtılması faydalı olacaktır. Kitabın bu bölümünün amacı temel doğrusal cebir konularını sadece regresyon kapsamında ele almaktır. Bu nedenle temel teoremlerin bazıları ispatsız olarak verilecektir.

3.1 VEKTÖRLER VE MATRİSLER

Matris sayı veya elemanların sıralar ve sütunlar şeklinde düzenlendiği dikdörtgen bir dizindir. Matrisi oluşturan bu sıra ve sütunların sayısı matrisin boyutunu belirler. Örneğin, m adet sıra ve n adet sütundan oluşan bir **A** matrisi aşağıda gösterilmiştir. Bu kitapta matrisler büyük koyu harf ile vektörler ise küçük koyu harf ile gösterilecektir.

$$\mathbf{A}_{mxn} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \tag{3.1}$$

Bu kitapta $\mathbf{A}_{mxn}$ ifadesi, m sıra ve n sütundan oluşan bir matrisi ifade edecektir ve sıra sayısı ilk sütun sayısı ise ikinci indisle tanımlanacaktır.

Bir *skaler* 1×1 boyutlu bir matristir. m adet sıra ve sadece bir tek sütundan oluşan matris *sütun vektörüdür*. Örneğin, $\mathbf{x}_{4\times1}$ matrisi bir sütun vektörüdür.

$$\mathbf{x}_{4x1} = \begin{bmatrix} -3 \\ 4 \\ 7 \\ 6 \end{bmatrix}$$

n adet sütun ve sadece bir tek sıradan oluşan matris ise *sıra vektörüdür*. $\mathbf{x}_{1\times3}$ matrisi örnek olarak verilebilir.

$$\mathbf{x}_{1x3} = \begin{bmatrix} -2 & 2 & 5 \end{bmatrix}$$

Bir $\mathbf{A}_{m\times n}$ matrisinin *transpozu* $\mathbf{A}^T_{n\times m}$ şeklinde ifade edilir ve **A** matrisinin sıraları ile sütunları yer değiştirilerek elde edilir. Başka bir deyişle **A** matrisinin i-inci sırası $\mathbf{A}^{\mathrm{T}}$ matrisinin i-inci sütununu oluşturur. $\mathbf{A}_{3x2}$ matrisinin kendisi ve transpozu aşağıda verilmiştir.

$$\mathbf{A}_{3x2} = \begin{bmatrix} 3 & 4 \\ -2 & 5 \\ 7 & 1 \end{bmatrix} \qquad \mathbf{A}^T_{2\times 3} = \begin{bmatrix} 3 & -2 & 7 \\ 4 & 5 & 1 \end{bmatrix}$$

Yukarıdaki tanımdan anlaşılacağı gibi bir sütun vektörünün transpozu,

$$\mathbf{x}_{4x1} = \begin{bmatrix} -3 \\ 4 \\ 7 \\ 6 \end{bmatrix} \qquad \mathbf{x}^T_{1\times 4} = \begin{bmatrix} -3 & 4 & 7 & 6 \end{bmatrix}$$

şeklinde bir sıra vektörü, bir sıra vektörünün transpozu da,

$$\mathbf{y}_{1x3} = \begin{bmatrix} -2 & 2 & 5 \end{bmatrix} \qquad \mathbf{y}^T_{3\times 1} = \begin{bmatrix} -2 \\ 2 \\ 5 \end{bmatrix}$$

şeklinde bir sütun vektörü oluşturur.

Teorem 3.1: $\mathbf{A}^T$ matrisinin (vektörünün) transpozu **A** matrisine (vektörüne) eşittir.

$$\left(\mathbf{A}^T\right)^T = \mathbf{A} \tag{3.2}$$

Sıra ve sütun sayılarının eşit olduğu ($m = n$) matrisler *kare matris* olarak adlandırılır.

$$\mathbf{A}_{3x3} = \begin{bmatrix} 4 & 2 & 5 \\ 9 & -1 & 4 \\ 3 & 2 & 6 \end{bmatrix} \tag{3.3}$$

Köşegen elemanları haricindeki elemanları sıfıra eşit olan kare matris, *köşegen matris* olarak bilinir. Bir matrisin köşegen matris olabilmesi için köşegen elemanlarından bir tanesinin sıfırdan farklı olması yeterlidir.

$$\mathbf{A}_{4x4} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & -7 & 0 \\ 0 & 0 & 0 & 5 \end{bmatrix} \tag{3.4}$$

Teorem 3.2: Eğer $\mathbf{D}_1$ ve $\mathbf{D}_2$ köşegen matrisler ise bu matrislerin çarpımları da bir köşegen matristir.

$$\mathbf{D}_1\mathbf{D}_2 = \mathbf{D}_2\mathbf{D}_1 = \mathbf{D} \tag{3.5}$$

D matrisinin i-inci elemanı, $\mathbf{D}_1$ ve $\mathbf{D}_2$'nin i-inci elemanlarının çarpımından elde edilir.

Köşegen matrisin tüm elemanları eşit ise *skaler matristir*.

$$\mathbf{D}_{3x3} = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

Elemanları bire (1) eşit olan skaler matris *birim matristir* ve **I** harfi ile belirtilir.

$$\mathbf{I}_{3x3} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{3.6}$$

Teorem 3.3: Herhangi bir **A** matrisinin soldan ya da sağdan birim matris ile çarpılması **A** matrisini değiştirmez.

$$\mathbf{IA} = \mathbf{AI} = \mathbf{A} \tag{3.7}$$

Bu özelliği birim matrisin matris işlemlerinde etkisiz eleman olarak kullanılmasını sağlar. Aşağıdaki işlem bu duruma bir örnek olarak verilebilir.

$$\begin{aligned} \mathbf{Y} - \mathbf{XY} &= \mathbf{IY} - \mathbf{XY} \\ &= (\mathbf{I} - \mathbf{X})\mathbf{Y} \end{aligned}$$

Köşegen elemanlarının üstündeki (üst üçgen) a_{ij} elemanları ile köşegen elemanlarının altındaki (alt üçgen) a_{ji} elemanlarının tümünün bire bir eşit olduğu matris *simetrik matristir*. Simetrik matrisin transpozuda kendisine eşittir, $\mathbf{A}=\mathbf{A}^T$. Başka bir deyişle **A** matrisinin a_{ij} elemanı, $\mathbf{A}^T$ matrisinin a_{ji} elemanına eşittir.

$$\mathbf{A}_{3x3} = \begin{bmatrix} 4 & 2 & 3 \\ 9 & -1 & 2 \\ 3 & 2 & 6 \end{bmatrix} \tag{3.8}$$

Kare matrislerin bir diğer önemli özel durumu ise idempotent matristir. **A** matrisi bir simetrik kare matris olsun. Eğer;

$$\mathbf{A} = \mathbf{A}^2 = \mathbf{A}^3 = ... \tag{3.9}$$

eşitliği sağlanıyorsa **A** matrisi *idempotent matristir*. Başka bir deyişle matrisin kendisi ile çarpımı (veya çarpımları) orijinal matrisi veriyorsa bu matris idempotent matristir. Simetrik olmayan herhangi bir kare matris (3.9) eşitliğini sağlıyorsa idempotent matristir. Fakat bu kitapta sadece simetrik idempotent matrislerle ilgilenilecektir. Bu matrisler istatistik teorisinde önemli bir yer tutmaktadır. Bu nedenle ilgili teoremler bölüm sonunda verilecektir.

Önemli bir matris türü de ortogonal matrislerdir. Eğer **C** matrisi bir kare matris ise ve tersi alınabiliyorsa,

$$\mathbf{C}^T = \mathbf{C}^{-1} \tag{3.10}$$

eşitliğini sağlayan matrisler *ortogonal matrislerdir*.

Regresyon analizi için gerekli olan matris ve vektör işlemleri aşağıda açıklanacaktır.

3.2 MATRİS VE VEKTÖR İŞLEMLERİ

Matrislerin çarpımı matrisleri oluşturan vektörlerin çarpımları ile elde edilir. Örneğin verilen $\mathbf{a}_{n\times 1}$ ve $\mathbf{b}_{n\times 1}$ boyutlu vektörler iki şekilde çarpılabilir:

$$\mathbf{a}_{nx1} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \qquad \mathbf{b}_{nx1} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

$$\mathbf{a}_{n\times 1}\mathbf{b}_{1\times n}^T = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \begin{bmatrix} b_1 & b_2 & \cdots & b_n \end{bmatrix} = \begin{bmatrix} a_1b_1 & a_1b_2 & \cdots & a_1b_n \\ a_2b_1 & a_2b_2 & \cdots & a_2b_n \\ \vdots & \vdots & & \vdots \\ a_nb_1 & a_nb_2 & \cdots & a_nb_n \end{bmatrix} = \mathbf{C}_{nxn} \tag{3.11}$$

veya

$$\mathbf{a}_{1\times n}^T\mathbf{b}_{nx1} = \begin{bmatrix} a_1 & a_2 & \cdots & a_n \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} = a_1b_1 + a_2b_2 + \cdots + a_nb_n = \sum_{i=1}^{n} a_ib_i = \mathbf{c}_{1x1} \tag{3.12}$$

şeklindedir. Örneklerden görüldüğü gibi çarpılan iki vektörden soldakinin sütun boyutu ile sağdakinin sıra boyutu eşit olmak orundadır bu özellik matrisler içinde geçerlidir. (3.12) ifadesindeki çarpım *nokta* (dot), *skaler* veya *iç* (inner) çarpım olarak adlandırılır. İki matrisin çarpımında bu tip çarpım kullanılır. $\mathbf{A}_{m\times n}$ boyutlu ve $\mathbf{B}_{n\times p}$ boyutlu iki matris ise **AB** çarpımı $m\times p$ boyutlu bir **C** matrisini verir.

$$\mathbf{A}_{m\times n}\mathbf{B}_{n\times p} = \mathbf{C}_{m\times p} \tag{3.13}$$

Elde edilen **C** matrisinin elemanları ise,

$$c_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj} \tag{3.14}$$

eşitliği ile elde edilir, bkz Alıştırma 3.1.

Teorem 3.4: Matris çarpımlarının transpozu,

$$(\mathbf{ABC})^T = \mathbf{C}^T\mathbf{B}^T\mathbf{A}^T \tag{3.15}$$

şeklindedir.

İki vektörün ya da iki matrisin toplanabilmesi için her ikisinin boyutlarının da karşılıklı olarak eşit olması gerekir. $\mathbf{a}_{n\times 1}$ ve $\mathbf{b}_{n\times 1}$ boyutlu iki vektörün toplamı karşılıklı elemanlarının toplanması ile,

$$\mathbf{a}_{n\times 1} + \mathbf{b}_{n\times 1} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} = \begin{bmatrix} a_1 + b_1 \\ a_2 + b_2 \\ \vdots \\ a_n + b_n \end{bmatrix} = \mathbf{c}_{n\times 1} \tag{3.16}$$

şeklinde elde edilebilir. Elde edilen vektörün boyutları da toplamları yapılan vektörlerin boyutlarına eşittir. Bu kural matrisler için de geçerli olup,

$$\mathbf{A}_{m\times n}+\mathbf{B}_{m\times n}=\mathbf{C}_{m\times n} \tag{3.17}$$

şeklindedir, bkz Alıştırma 3.2.

Teorem 3.5: Matris toplamlarının transpozu,

$$(\mathbf{A}+\mathbf{B}+\mathbf{C})^T=\mathbf{A}^T+\mathbf{B}^T+\mathbf{C}^T \tag{3.18}$$

şeklindedir.

Teorem 3.6: Çarpmanın toplama üzerine dağılma özelliği, matrislerin boyutları çarpmaya uygun olmak üzere,

$$\mathbf{A}(\mathbf{B}+\mathbf{C})=\mathbf{AB}+\mathbf{AC} \tag{3.19}$$

şeklinde matrisler için uygulanabilir.

3.3 MATRİSİN PARÇALANMASI

Bazı durumlarda bir matrisin alt matrislere ayrılması faydalı olabilir. Bu işleme matrisin parçalanması işlemi adı verilir. Bir matris çeşitli şekillerde parçalanabilir. Örneğin, $m\times n$ boyutlu bir **A** matrisi,

$$\mathbf{A}=\begin{pmatrix}\mathbf{A}_{11} & \mathbf{A}_{12}\\ \mathbf{A}_{21} & \mathbf{A}_{22}\end{pmatrix} \tag{3.20}$$

şeklinde alt matrislere ayrılabilir. Burada $\mathbf{A}_{11}$ matrisi $m_1\times n_1$, $\mathbf{A}_{12}$ matrisi $m_1\times n_2$, $\mathbf{A}_{21}$ matrisi $m_2\times n_1$ ve $\mathbf{A}_{22}$ matrisi $m_2\times n_2$, boyutludur. Ayrıca $m_1+m_2=m$ ve $n_1+n_2=n$ olduğu görülebilir. **AB** şeklinde iki matrisin çarpımı, bu matrisler alt matrislere ayrılmış olsalar bile sembolik olarak gösterilebilirler. Bunun için matrislerin boyutlarının çarpım için uygun olması yeterlidir. Eğer **B** matrisi $n\times p$ boyutlu ve,

$$\mathbf{B}=\begin{pmatrix}\mathbf{B}_{11} & \mathbf{B}_{12}\\ \mathbf{B}_{21} & \mathbf{B}_{22}\end{pmatrix}$$

şeklinde parçalanmış ise $\mathbf{B}_{jk}$ alt matrisleri $n_j\times p_k$ boyutludur. Bu durumda **AB** çarpımı mevcuttur. Çünkü $\mathbf{A}_{ij}$'nin boyutları $m_i\times n_j$ ve $\mathbf{B}_{jk}$'nin ise $n_j\times p_k$'dır ve alt matrislerin karşılıklı elemanları,

$$\mathbf{AB}=\begin{pmatrix}\mathbf{A}_{11} & \mathbf{A}_{12}\\ \mathbf{A}_{21} & \mathbf{A}_{22}\end{pmatrix}\begin{pmatrix}\mathbf{B}_{11} & \mathbf{B}_{12}\\ \mathbf{B}_{21} & \mathbf{B}_{22}\end{pmatrix}=\begin{pmatrix}\mathbf{A}_{11}\mathbf{B}_{11}+\mathbf{A}_{12}\mathbf{B}_{12} & \mathbf{A}_{11}\mathbf{B}_{12}+\mathbf{A}_{12}\mathbf{B}_{22}\\ \mathbf{A}_{21}\mathbf{B}_{11}+\mathbf{A}_{22}\mathbf{B}_{12} & \mathbf{A}_{21}\mathbf{B}_{12}+\mathbf{A}_{22}\mathbf{B}_{22}\end{pmatrix} \tag{3.21}$$

şeklinde çarpılıp, toplanarak çarpım matrisi elde edilir. Bu konu ile ilgili teoremler bölüm sonunda verilecektir.

3.4 MATRİSLERDE TÜREV İŞLEMİ

Eğer $f(\mathbf{b})$ fonksiyonu k adet farklı b_i katsayısını içeriyorsa bu fonksiyonun her bir b_i'ye göre kısmi türevi alınabilir. Bu kısmi türevlerin bir sütun vektörü formundaki genel tanımı,

$$\frac{\partial[f(\mathbf{b})]}{\partial \mathbf{b}} = \begin{bmatrix} \dfrac{\partial[f(\mathbf{b})]}{\partial b_1} \\ \vdots \\ \dfrac{\partial[f(\mathbf{b})]}{\partial b_k} \end{bmatrix} \tag{3.22}$$

şeklindedir. Bu kısmi türevler bir sıra vektörü şeklinde düzenlenebilir. Bu işlemde vektör ve matrislerdeki toplama ve çarpma işlemleri için boyutların uyumlu olması sağlanmalıdır. $f(\mathbf{b})$'nin bir doğrusal fonksiyon olduğu kabul edilerek,

$$f(\mathbf{b}) = \mathbf{a}^T\mathbf{b}$$

$$= a_1 b_1 + a_2 b_2 + \cdots + a_k b_k$$

yazılabilir. Bu eşitlikteki a_i değerleri birer sabittir. Bu ifadeye eşitlik (3.22)'nin uygulanması ile,

$$\frac{\partial(\mathbf{a}^T\mathbf{b})}{\partial \mathbf{b}} = \frac{\partial(\mathbf{b}^T\mathbf{a})}{\partial \mathbf{b}} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_k \end{bmatrix} = \mathbf{a} \tag{3.23}$$

sonucu elde edilebilir. Eğer $f(\mathbf{b})$ fonksiyonu b_i değerleri açısından karesel ise,

$$f(\mathbf{b}) = \mathbf{b}^T\mathbf{A}\,\mathbf{b}$$

şeklinde ifade edilebilir. **A** ile belirtilen matrisin simetrik olduğu kabul edilerek,

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{12} & a_{22} & \cdots & a_{2k} \\ \vdots & \vdots & & \vdots \\ a_{1k} & a_{2k} & \cdots & a_{kk} \end{bmatrix}$$

yazılabilir. Bunun sonucu olarak,

$$\begin{aligned} \mathbf{b}^T\mathbf{A}\,\mathbf{b} = a_{11}b_1^2 &+ 2a_{12}b_1b_2 + 2a_{13}b_1b_3 + \cdots + 2a_{1k}b_1b_k \\ &+ a_{22}b_2^2 + 2a_{23}b_2b_3 + \cdots + 2a_{2k}b_2b_k \\ &\qquad \ddots \qquad\qquad\qquad \vdots \\ &\qquad\qquad\qquad \ddots \qquad \vdots \\ &\qquad\qquad\qquad\qquad \ddots \ \vdots \\ &\qquad\qquad\qquad\qquad\quad + a_{kk}b_k^2 \end{aligned}$$

olarak elde edilir. Bu ifadenin kısmi türevleri,

$$\frac{\partial(\mathbf{b}^T\mathbf{A}\,\mathbf{b})}{\partial b_1} = 2\,(a_{11}b_1 + a_{12}b_2 + \cdots + a_{1k}b_k) = 2\mathbf{a}_1 b_1$$

....

$$\frac{\partial(\mathbf{b}^T \mathbf{A}\,\mathbf{b})}{\partial b_k} = 2\,(a_{1k}b_1 + a_{2k}b_2 + \cdots + a_{kk}b_k) = 2\mathbf{a}_k b$$

şeklinde bulunur. Eşitliğin en sağındaki $\mathbf{a}_i$ değerleri **A** matrisinin sıralarını ifade etmektedir. Bu kısmi türevler bir sütun vektörü olarak,

$$\frac{\partial(\mathbf{b}^T \mathbf{A}\,\mathbf{b})}{\partial \mathbf{b}} = 2\begin{bmatrix} \mathbf{a}_1\mathbf{b} \\ \mathbf{a}_2\mathbf{b} \\ \vdots \\ \mathbf{a}_k\mathbf{b} \end{bmatrix} = 2\begin{bmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \\ \vdots \\ \mathbf{a}_k \end{bmatrix}\mathbf{b} = 2\mathbf{A}\mathbf{b} \tag{3.24}$$

elde edilebilir. Eşitlik (3.23) ve (3.24) doğrusal ve karesel formların standart diferansiyel sonuçlarını verir.

3.5 VEKTÖR GEOMETRİSİ

Bölüm 5 de verilecek olan regresyon geometrisi kısmına temel oluşturmak üzere bu kısımda vektör geometrisi incelenecektir. Pek çok problemde cebirsel olduğu kadar geometrik yorumlama yapabilmek oldukça faydalıdır. Bu nedenle temel vektör geometrisi ile ilgili bazı basit notasyonlar verilecektir. İki elemanlı bir vektör,

$$\mathbf{x}_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

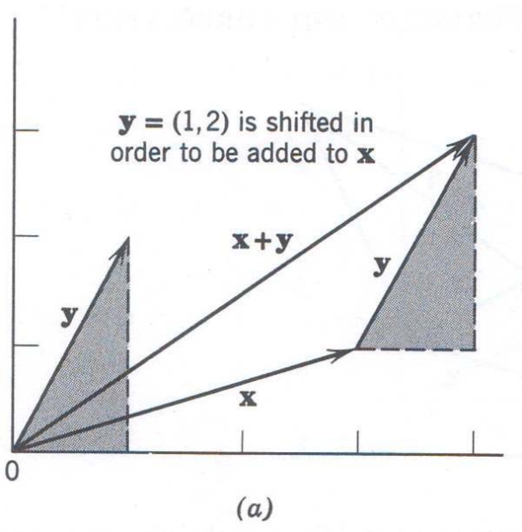
şeklinde tanımlanabilir. Bu vektör Şekil (3.1)'de olduğu gibi düz bir doğru parçası ile gösterilebilir. Bu doğru parçası, orjinde başlayıp, (2,1) koordinatının belirttiği noktada biter. Ok ise vektörün yönünü belirtir. Bir başka $\mathbf{x}_2$ vektörü ise,

$$\mathbf{x}_2 = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$$

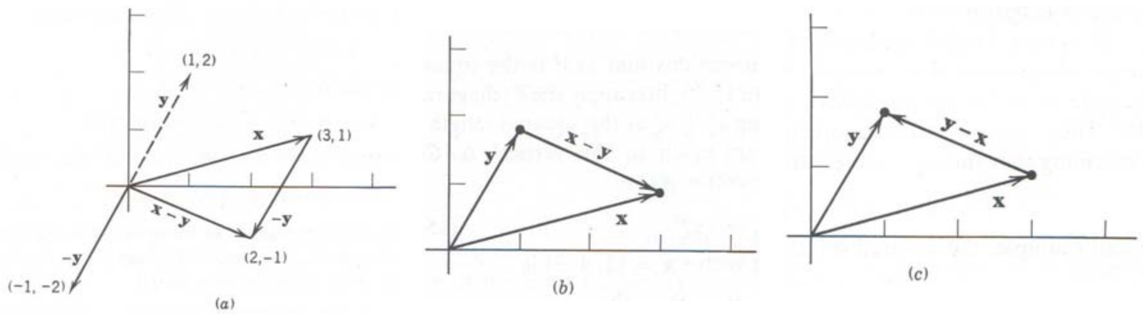
şeklinde verilebilir. Bu iki vektör geometrik olarak aşağıda belirtildiği şekilde toplanabilir: Bu iki vektörden herhangi biri orjin dikkate alınarak çizilir (şekilde ilk çizilen $\mathbf{x}_1$ vektörüdür). Daha sonra diğer vektör ilk vektörün bitim koordinatından itibaren (bu bitim noktası ikinci vektör için orjin olarak kabul edilebilir) çizilir. Şekilde varılan nokta p'dir. İlk vektörün başlangıç noktası ile bu p noktasını birleştiren doğru parçası $\mathbf{y}$ vektörünü verir. Elde edilen bu vektörün koordinatları (3,4)'dür. $\mathbf{y}$ vektörü $\mathbf{x}_1$ ve $\mathbf{x}_2$ vektörlerinin geometrik olarak toplanması ile elde edilir. Eğer bu iki vektör cebirsel olarak toplanırsa,

$$\mathbf{y} = \mathbf{x}_1 + \mathbf{x}_2 = \begin{bmatrix} 2 \\ 1 \end{bmatrix} + \begin{bmatrix} 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$$

sonucu elde edilir. Görüldüğü gibi hem geometrik hem de cebirsel olarak tamamen aynı sonuçlar elde edilmiştir. Şekilden görüldüğü gibi ilk olarak $\mathbf{x}_2$ vektörü çizilseydi de aynı noktaya (p'ye) varılacak ve sonuç değişmeyecekti. Vektörel çıkarma ise Şekil 3.2 de gösterilmiştir.

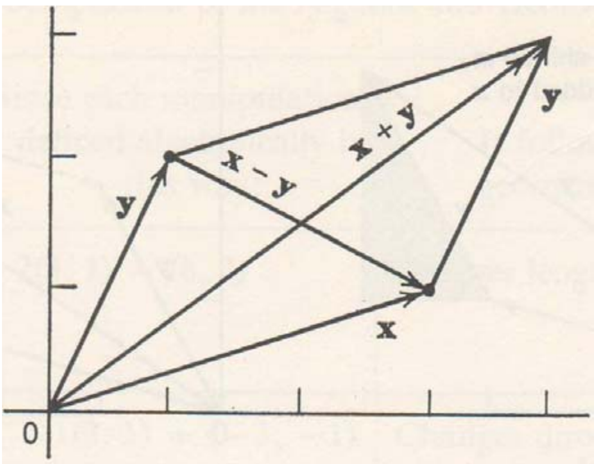


Şekil 3.1 *Vektörel toplama işlemi*



Şekil 3.2 *Vektörel çıkarma*

Şekil 3.3 den (**x**+**y***)* vektörel toplamının, **x** ve **y** vektörlerinden oluşturulmuş paralel kenarın bir köşegeni, (**x**-**y**) farkının ise paralel kenarın diğer bir köşegeni olduğu görülmektedir.



Şekil 3.3 *Vektörel toplama ve çıkarmanın karşılaştırılması: Toplama* **x** *okunu takip eden* **y** *okunun ucunu* **x'***in başlangıcı ile birleştiren köşegendir. Çıkarma ise* **y** *noktasından* **x** *noktasına oluşturulan köşegen ile tanımlanır.*

Şimdi bir vektörün bir skalerle çarpımı ele alınsın. Bu duruma bir örnek olarak $\mathbf{x}_1$ vektörü 2 ile çarpılmış ve

$$2\mathbf{x}_1 = 2\begin{bmatrix}3\\1\end{bmatrix} = \begin{bmatrix}6\\2\end{bmatrix}$$

sonucu elde edilmiştir. Elde edilen bu vektörün yönü $\mathbf{x}_1$ vektörü ile tamamen aynıdır. Aradaki fark ise elde edilen vektörün uzunluğunun $\mathbf{x}_1$ vektörünün iki katı olmasıdır. Bu skaler negatif bir sayıda olabilir.

$$-1\mathbf{x}_1 = -1\begin{bmatrix}3\\1\end{bmatrix} = \begin{bmatrix}-3\\-1\end{bmatrix}$$

Elde edilen bu iki vektör orjinal $\mathbf{x}_1$ vektörü ile birlikte Şekil 3.4'de gösterilmiştir. Şekilden görüldüğü gibi bu üç vektörün hepside orjinden geçen bir tek doğru parçası üzerindedir ve bu doğru parçası da $\mathbf{x}_1$ vektörü ile tanımlanmıştır.

Vektörlerin toplanması ve bir skalerle çarpılması işlemlerinin birlikte ele alınması sonucunda $\mathbf{x}_1$ ve $\mathbf{x}_2$ vektörlerinin doğrusal kombinasyonu olan herhangi bir iki elemanlı vektör elde edilebilir. Elde edilen bu $\mathbf{y}$ vektörü,

$$\mathbf{y} = \lambda_1\mathbf{x}_1 + \lambda_2\mathbf{x}_2 \qquad (3.25)$$

şeklinde ifade edilebilir. Bu eşitlikteki λ_1 ve λ_2 skalerleri temsil etmektedir. Eğer $\mathbf{y}$ vektörü,

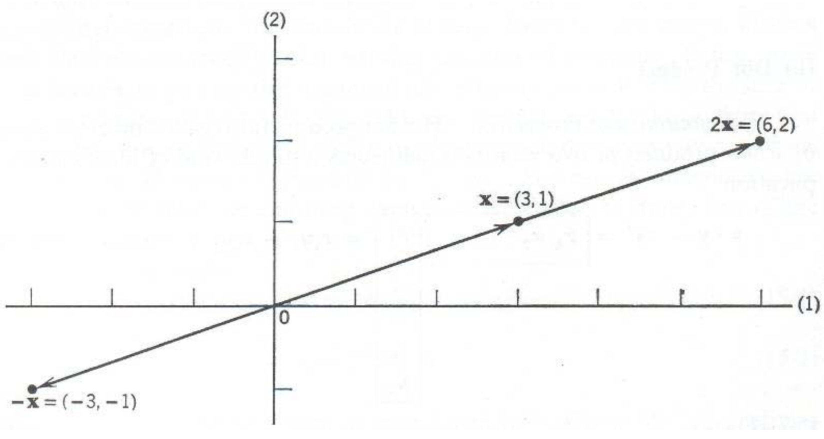
$$\mathbf{y} = \begin{bmatrix}4\\0\end{bmatrix}$$

şeklinde tanımlanmış ise bu vektör $\mathbf{x}_1$ ve $\mathbf{x}_2$ vektörlerinin doğrusal bir kombinasyonu olarak,

$$\mathbf{y} = \frac{12}{5}\begin{bmatrix}2\\1\end{bmatrix} - \frac{4}{5}\begin{bmatrix}1\\3\end{bmatrix}$$

ifade edilebilir. Bu eşitlikte λ_1=12/5 ve λ_2=-4/5'dir. Bu değerler aşağıdaki şekilde, denklem çiftinin eşanlı olarak çözülmesi ile elde edilebilir.

$$\lambda_1\begin{bmatrix}2\\1\end{bmatrix} + \lambda_2\begin{bmatrix}1\\3\end{bmatrix} = \begin{bmatrix}2\lambda_1 + \lambda_2\\ \lambda_1 + 3\lambda_2\end{bmatrix} = \begin{bmatrix}4\\0\end{bmatrix}$$



Şekil 3.4 *Bir vektörün skalerle çarpımı*

Bir başka örnek için $\mathbf{y}$ vektörü,

$$\mathbf{y} = \begin{bmatrix} 6 \\ 3 \end{bmatrix}$$

olarak verilmiş olsun. Bu durumda $\mathbf{x}_1$ ve $\mathbf{x}_2$ vektörlerinin doğrusal bir kombinasyonu olarak,

$$\mathbf{y} = 3\begin{bmatrix} 2 \\ 1 \end{bmatrix} + 0\begin{bmatrix} 1 \\ 3 \end{bmatrix}$$

şeklinde ifade edilebilir. Bu eşitlikte $\lambda_1=3$ ve $\lambda_2=0$'dır. Bu örnekler daha da arttırılabilir. Bu verilenler dikkate alınarak bir vektör uzayının tanımı yapılabilir. Vektör uzayı aşağıda verilen özelliklere sahip vektörlerin biraraya gelmesi ile oluşur.

1) Eğer $\mathbf{v}_1$ ve $\mathbf{v}_2$ aynı vektör uzayındaki iki vektör ise, $\mathbf{v}_1+\mathbf{v}_2$' de aynı vektör uzayındadır.

2) Eğer $\mathbf{v}$ bu vektör uzayında ve λ bir skaler sabitse, $\lambda\mathbf{v}$ 'de aynı vektör uzayındadır.

Bu özelliklerden görüldüğü gibi toplama ve bir skalerle çarpma işlemleri sonucunda elde edilen vektör yine aynı vektör uzayındadır. Vektörlerin basit cebirsel işlemleri ve bu işlemlerin geometrik yorumları Tablo 3.1'de verilmiştir.

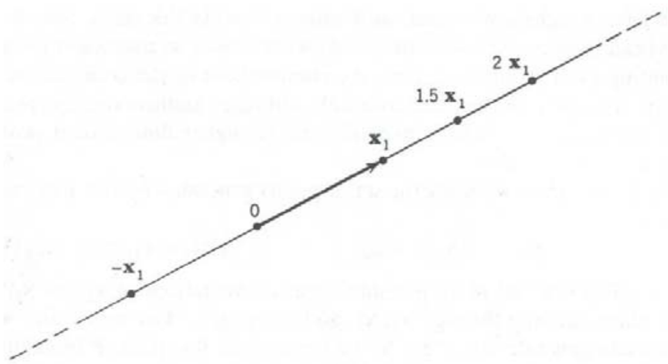
Tablo 3.1 *Vektörlerin cebirsel ve geometrik karşılaştırılması*

	Cebirsel işlem	Geometrik yorumu
Pozitif bir skalerle Çarpımı	2(3,1)=(6,2)	Uzunluk değişir. (Şekil 3.4)
Negatif sabit bir Skalerle çarpımı	-1(3,1)=(-3,-1)	Yön değişir. (Şekil3.4)
Toplam	(3,1)+(1,2)=(4,3)	**x** ve **y** okları birbirini takip edecek şekilde eklenir. Oluşturulan paralel kenarın köşegeni toplam vektörü belirtir.(Şekil 3.1)
Çıkarma	(3,1)-(1,2)=(2,-1)	**x** +(-**y**) toplamına eşittir. (Şekil3.2)

Bir boyutlu uzay IR^1 ile belirtilsin. Sabit bir $\mathbf{x}_1$ vektörünün mümkün olabilecek tüm λ_1 skalerleriyle çarpılması sonucu orijinle $\mathbf{x}_1$ doğrultusunda bir düz doğru oluştuğu Şekil 3.5'den görülmektedir. Her bir $\lambda_1\mathbf{x}_1$ vektörü ok veya nokta ile ifade edilebilir. Yukarıdaki açıklama

$$L: \lambda_1\mathbf{x}_1 \qquad -\infty < \lambda_1 < \infty \tag{3.26}$$

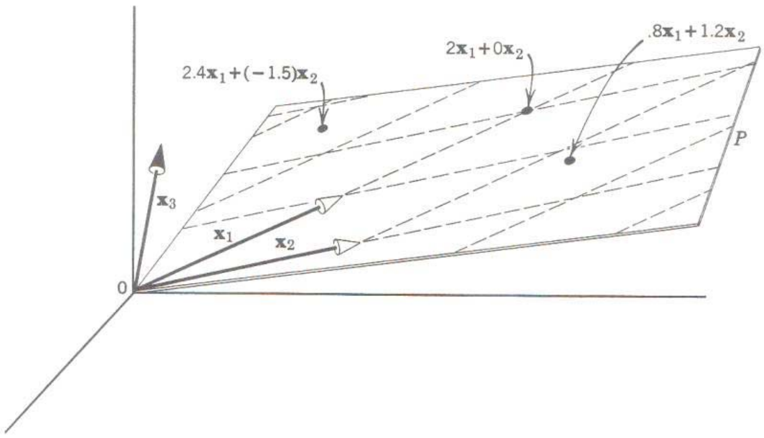
ifadesiyle özetlenebilir.



Şekil 3.5 $\mathbf{x}_1$ *vektörü ile oluşturulan L doğrusu*

Şekil 3.6'da boyut sayısı bir arttırılmıştır ve iki farklı renkte ok başı kullanılmıştır. Düzlem içinde bulunan ok başları beyaz olup, düzlem dışındaki siyahtır. Üç boyutlu uzayda iki adet sabitlenmiş vektörün oluşturduğu noktalar seti Şekil 3.6'da gösterilmektedir:

$$P: \lambda_1\mathbf{x}_1+\lambda_2\mathbf{x}_2 \qquad -\infty < \lambda_1, \lambda_2 < \infty \tag{3.27a}$$



Şekil 3.6 $\mathbf{x}_1$ *ve* $\mathbf{x}_2$ *vektörleri ile oluşturulabilen P düzlemi. P=* $\lambda_1\mathbf{x}_1+\lambda_2\mathbf{x}_2$ *olup* λ_1 *ve* λ_2 *değerlerine bağlı olarak düzlem sınırsız bir şekilde genişletilir.*

P ifadesi $\mathbf{x}_1$ ve $\mathbf{x}_2$'nin tüm mümkün doğrusal kombinasyonlarının seti olarak adlandırılır ve $\mathbf{x}_1$, $\mathbf{x}_2$ vektörleri ile orijin arasında kalan düzlemi belirtir. Geometrik olarak, $\mathbf{x}_1$ ve $\mathbf{x}_2$'nin uygun doğrusal kombinasyonları dikkate alınarak *P* düzleminde herhangi bir nokta oluşturabileceği görülmektedir. Uygun doğrusal kombinasyon ile açıklanmak istenen λ_1 ve λ_2 'nin eşitlik (3.27)'ya uygun olarak seçilmiş olmasıdır. Bu koşula uyulması durumunda elde edilen nokta, *P* düzleminin altında ya da üstünde olmayacaktır. Aşağıda λ_1 ve λ_2 skalerlerinin nasıl belirlenebileceği açıklanmıştır.

İki boyutlu uzay IR^2 ile belirtilsin. Bu vektör uzayı tüm iki elemanlı gerçel vektörlerden oluşur. Bu uzaydaki herhangi bir vektörün $\mathbf{x}_1$ ve $\mathbf{x}_2$ vektörlerinin doğrusal kombinasyonu olarak ifade edilebileceği açıktır. Bununla birlikte $\mathbf{x}_1$ ve $\mathbf{x}_2$ vektörleri keyfi olarak seçilmişti. Bu nedenle aşağıdaki vektör çifti ele alınsın,

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

bu vektörler birim vektörler olarak tanımlanabilir. IR^2'deki herhangi bir **c** vektörü de bu vektörlerin doğrusal bir kombinasyonu olarak ifade edilebilir. Bu durum için λ değerlerinin belirlenmesi basittir. Daha önce verilen iki örnek için,

$$\begin{bmatrix}4\\0\end{bmatrix}=4\begin{bmatrix}1\\0\end{bmatrix}+0\begin{bmatrix}0\\1\end{bmatrix} \quad , \quad \lambda_1=4 \quad \text{ve} \quad \lambda_2=0$$

$$\begin{bmatrix}6\\3\end{bmatrix}=6\begin{bmatrix}1\\0\end{bmatrix}+3\begin{bmatrix}0\\1\end{bmatrix} \quad , \quad \lambda_1=6 \quad \text{ve} \quad \lambda_2=3$$

şeklinde λ_1 ve λ_2 değerleri bulunabilir. Görüldüğü gibi elde edilen bu değerler **x** vektörlerinin elemanlarıdır.

Bu örneklerdeki vektör çiftlerinin herbiri ($\mathbf{x}_1$,$\mathbf{x}_2$ ve $\mathbf{e}_1$, $\mathbf{e}_2$) iki boyutlu uzay IR^2 için uygun bir kaynak teşkil edebilir. Bu sonuca göre bir kaynağın benzersiz (eşsiz) olmadığı söylenebilir. Herhangi iki vektörün IR^2'de uygun bir kaynak olabilmesi için onların farklı yönlerde olması gereklidir. Eğer $\mathbf{x}_1$ ve $\mathbf{x}_2$ aynı yönde ise Şekil 3.4'de gösterildiği gibi bir vektör bir skalerle çarpılarak diğer vektöre dönüştürülebilir ve daha sonra bu vektörün çarpımları $\mathbf{x}_1$ ve $\mathbf{x}_2$ vektörlerinin doğrusal bir kombinasyonu olarak ifade edilebilir. Farklı yönlerdeki bu kaynak vektörler, doğrusal bağımsız vektörler olarak ifade edilebilecektir. $\mathbf{x}_1$ ve $\mathbf{x}_2$ vektörleri eğer sadece,

$$\lambda_1\mathbf{x}_1+\lambda_2\mathbf{x}_2=0 \tag{3.27b}$$

eşitliğini $\lambda_1=\lambda_2=0$ için sağlıyorsa doğrusal bağımsızdırlar. Eğer λ_i için sıfırdan farklı bir tek değer bile bulunabiliyorsa bu vektörler doğrusal bağımlıdır. $\mathbf{x}_1$ ve $\mathbf{x}_2$ vektörleri,

$$\mathbf{x}_1=\begin{bmatrix}2\\1\end{bmatrix} \qquad \mathbf{x}_2=\begin{bmatrix}6\\3\end{bmatrix}$$

şeklinde tanımlanırsa, bu iki vektörün aynı yönde farklı uzunlukta olduğu görülebilir. Bu durumda $3\mathbf{x}_1$-$\mathbf{x}_2$ doğrusal kombinasyonu sıfır değerini verecektir. Bununla birlikte başlangıçta verilen $\mathbf{x}_1$ ve $\mathbf{x}_2$ vektörleri,

$$\mathbf{x}_1=\begin{bmatrix}2\\1\end{bmatrix} \qquad \mathbf{x}_2=\begin{bmatrix}1\\3\end{bmatrix}$$

için $\lambda_1\mathbf{x}_1+\lambda_2\mathbf{x}_2=0$ eşitliğini sağlayan sıfırdan farklı λ_1 ve λ_2 değerleri bulmak imkansızdır. Bu vektörler için birinci elemanı sıfıra indirgeyecek λ_1 ve λ_2 çifti bulunabilir, örneğin, $\lambda_1=1$ ve $\lambda_2=-2$ gibi. Fakat bu λ çifti ikinci elemanları asla sıfıra indirgeyemeyecektir. Bu tanıma göre IR^2 için bir kaynak vektör çifti, birbirinden bağımsız herhangi iki elemanlı vektör olarak tanımlanabilir. İki boyutlu vektör geometrisinden görüleceği üzere, verilen bir kaynağa göre belirtilen bir vektörün eşsiz olarak temsil edilebilmesi için

$$\mathbf{y}=\lambda_1\mathbf{x}_1+\lambda_2\mathbf{x}_2$$

eşitliğini bir ve yalnız bir adet λ_1, λ_2 çiftinin sağlaması gerekmektedir.

IR^2 'de $\mathbf{x}_1$ ve $\mathbf{x}_2$'nin kaynak vektörler olarak verilmesi durumunda bu uzaydaki herhangi bir $\mathbf{y}$ vektörünün bu kaynak vektörlerin eşsiz bir doğrusal kombinasyonu olarak ifade edilebileceği belirtilmişti. Buna göre $\mathbf{x}_1$, $\mathbf{x}_2$ve $\mathbf{y}$ vektörleri doğrusal bağımlıdır. Çünkü,

$$\lambda_1\mathbf{x}_1 + \lambda_2\mathbf{x}_2 - \mathbf{y} = 0$$

denkleminde sıfırdan farklı bir λ ($\mathbf{y}$'nin katsayısı –1 olduğu için) değeri mevcuttur. Bu durumda IR^2 'de genişletilmiş $\mathbf{x}_1$,$\mathbf{x}_2$ve $\mathbf{y}$ vektör setine göre ifade edilebilecek herhangi bir $\mathbf{v}$ vektörünün mevcut olup olmadığı sorulabilir. Bu soruya verilecek yanıt şüphesiz evettir, fakat katsayılar eşsiz olmayacaktır. Örneğin; $\mathbf{x}_1$,$\mathbf{x}_2$ve $\mathbf{y}$ vektörleri,

$$\mathbf{x}_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix} \qquad \mathbf{x}_2 = \begin{bmatrix} 1 \\ 3 \end{bmatrix} \qquad \mathbf{y} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

şeklinde verilmiş ise $\mathbf{x}_1$, $\mathbf{x}_2$ve $\mathbf{y}$'nin doğrusal bir kombinasyonu olarak,

$$\mathbf{v} = \begin{bmatrix} 6 \\ 8 \end{bmatrix}$$

vektörü elde edilebilir. Bu kombinasyon,

$$\mathbf{v} = 2\mathbf{x}_1 + 2\mathbf{x}_2 + 0\mathbf{y}$$

şeklindedir. Fakat $\mathbf{v}$ için (6,8) değerini verebilecek pek çok λ değerleri mevcuttur. Genel doğrusal kombinasyon,

$$\mathbf{v} = \lambda_1\mathbf{x}_1 + \lambda_2\mathbf{x}_2 + \lambda_3\mathbf{y}$$

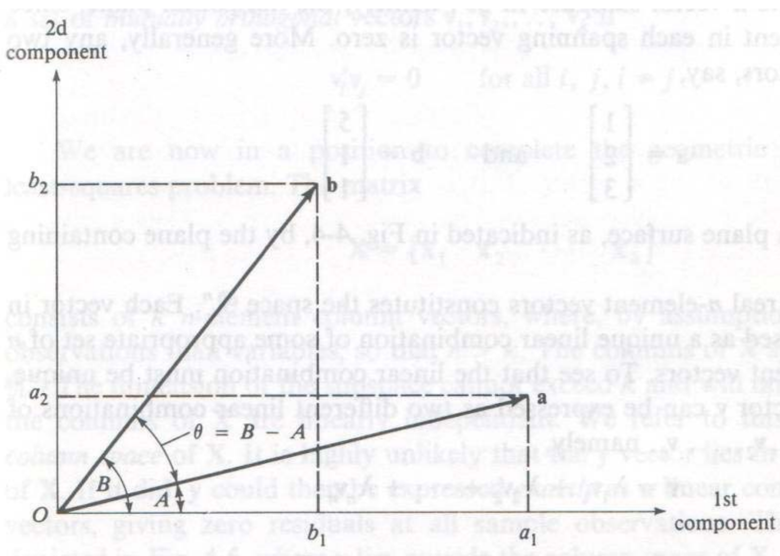
yeniden düzenlenerek,

$$\mathbf{v} - \lambda_3\mathbf{y} = \lambda_1\mathbf{x}_1 + \lambda_2\mathbf{x}_2$$

şeklinde de yazılabilir. Daha sonra da λ_3 için herhangi bir keyfi değer atanarak sol taraf spesifik iki elemanlı bir vektör haline getirilir ve bu vektör $\mathbf{x}_1$ ve $\mathbf{x}_2$'nin doğrusal bir kombinasyonu şeklinde ifade edilebilir. $\mathbf{x}_1$, $\mathbf{x}_2$ve $\mathbf{y}$'nin oluşturduğu bu tür vektör setleri, zincir seti olarak adlandırılır. Kaynak ve zincir setleri arasında bir fark mevcuttur. Bu fark da kaynak setlerin doğrusal bağımsız vektörlerden oluşmasından kaynaklanmaktadır. Verilen örnekteki zincir seti bir vektör eksiltilerek kaynak sete dönüştürülebilir.

Doğrusal bağımsız vektörler farklı yönlerde oldukları için bu vektörler arasındaki açıda sıfırdan farklıdır. Bu açı vektör elemanlarına göre ifade edilebilir. Şekil 3.7' de verilen iki boyutlu $\mathbf{x}_1$ ve $\mathbf{x}_2$ vektörleri ele alınsın. $\mathbf{x}_1$ vektörü ile yatay eksen arasındaki açı A ve $\mathbf{x}_2$ vektörü ile yatay eksen arasındaki açı da B ile belirtilsin. İki vektör arasındaki açı, θ=B-A olarak ifade edilir. Trigonometri yardımı ile,

$$Cos(B - A) = CosB\ CosA - SinB\ SinA \tag{3.28}$$



Şekil 3.7

formülü verilebilir. Pisagor teoremi ile bir vektörün *uzunluğu* veya *normu*, bir vektörün kendisiyle nokta çarpımının kare kökü olup, eşitlik (3.12) kullanılarak

$$\|\mathbf{x}_1\|^2 = \mathbf{x}_1\mathbf{x}_1^T = x_{11}^2 + \ldots + x_{n1}^2 \tag{3.29}$$

elde edilebilir. Bu ifadenin kare kökü vektörün uzunluğu olup $\|\mathbf{x}_1\|$ sembolü ile gösterilir ve benzer olarak,

$$\|\mathbf{x}_2\|^2 = \mathbf{x}_2\mathbf{x}_2^T = x_{12}^2 + \ldots + x_{n2}^2$$

elde edilir. Bu değerler (3.28) eşitliğinde yerine konarak,

$$\begin{aligned} Cos\theta &= \frac{x_{11}x_{12}}{\sqrt{\mathbf{x}_1^T\mathbf{x}_1}\sqrt{\mathbf{x}_2^T\mathbf{x}_2}} + \frac{x_{21}x_{22}}{\sqrt{\mathbf{x}_1^T\mathbf{x}_1}\sqrt{\mathbf{x}_2^T\mathbf{x}_2}} \\ Cos\theta &= \frac{\mathbf{x}_1^T\mathbf{x}_2}{\sqrt{\mathbf{x}_1^T\mathbf{x}_1}\sqrt{\mathbf{x}_2^T\mathbf{x}_2}} \end{aligned} \tag{3.30}$$

eşitliği elde edilebilir. Karekökün pozitif değeri dikkate alınır, çünkü bu değer bir uzunluğu ifade etmektedir. Eşitlik (3.30)'un iki özel durumu mevcuttur. Bunlardan birincisi $\mathbf{x}_1$ ve $\mathbf{x}_2$ vektörlerinin doğrusal bağımlı olmasıdır. Bu durumda, $\mathbf{x}_2 = \lambda_1\mathbf{x}_1$ yazılabilir. Vektörlerin doğrusal bağımlı olması durumunda (3.30) eşitliğinin sağ tarafı için bir, bunun sonucunda $\theta=0^0$ olarak bulunur. İkinci bir durumda $\mathbf{x}_1$ ve $\mathbf{x}_2$ vektörlerinin birbirine dik olmasıdır. Bu durumda $\theta=90^0$ olup $Cos\ \theta=0$ olarak elde edilir. $Cos\ \theta=0$ olması için $\mathbf{x}_1^T\mathbf{x}_2 = 0$ olması gerekir. Bu durum gerçekleştiğinde, diğer bir deyişle iki vektör arasındaki açı 90^0 olduğunda bu iki vektörün ortogonal olduğu söylenebilir. Sonuç olarak sadece ve sadece,

$$\mathbf{x}_1^T\mathbf{x}_2 = 0 \tag{3.31}$$

koşulu sağlanıyorsa iki vektörün ortogonal olduğu söylenebilir.

Vektörlerin nokta çarpımı iki boyutlu uzayda vektörün karesel uzunluğu olarak düşünülebilir, Şekil 3.8'de bu durum Pisagor teoremiyle açıklanmıştır. Örneğin $\mathbf{x}_1$ =(3,1) vektörünün karesel uzunluğu

$$\mathbf{x}_1\mathbf{x}_1^T = x_{11}^2 + x_{21}^2 = 3^2 + 1^2 = 10$$

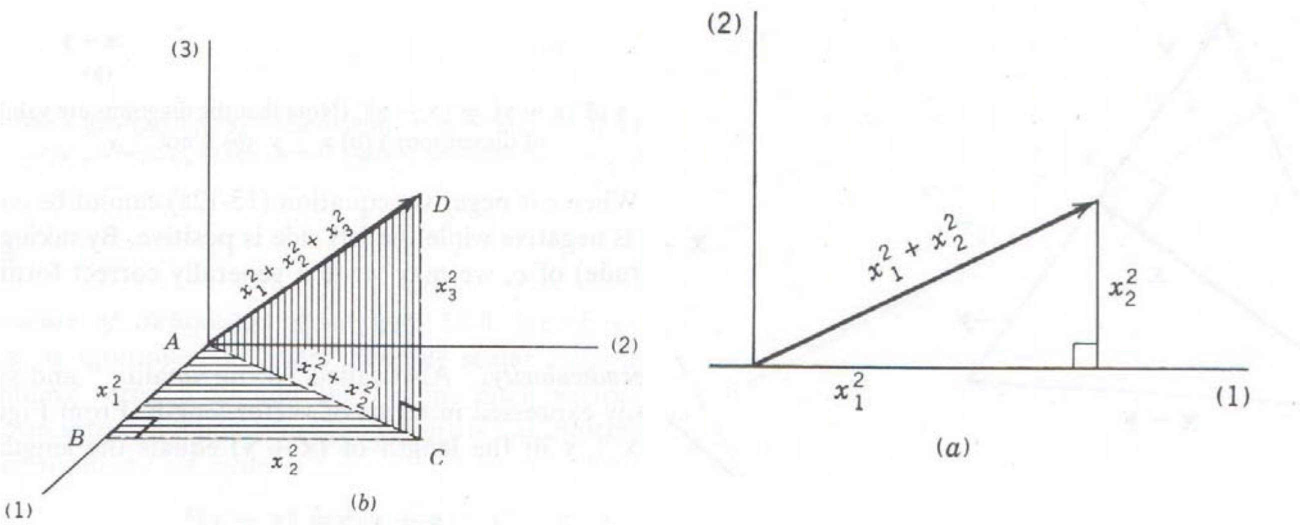
şeklinde olup uzunluğu $\sqrt{10}$ dur. Uzunlukla ilgili çok sık kullanım şekillerinden biri, λ_1 pozitif bir sabit olmak üzere,

$$\|\lambda_1\mathbf{x}_1\| = \lambda_1\|\mathbf{x}_1\| \tag{3.32a}$$

şeklindedir. Bu durum Şekil 5.2'den görülebilir. Eğer λ_1 negatif ise, (3.32) denklemi sağlanamayacaktır. Çünkü eşitliğin sağ tarafı negatif sol tarafı ise pozitif olacaktır. Bu nedenle λ_1'nin mutlak değeri alınarak eşitlik (3.32)'nın genel formu

$$\|\lambda_1\mathbf{x}_1\| = |\lambda_1|\|\mathbf{x}_1\| \tag{3.32b}$$

şeklinde yazılabilir.



Şekil 3.8 *Vektörlerin karesel uzunlukları ve Pisagor teoremiyle ilişkisi* a)*İki boyutlu*, b)*Üç boyutlu*

Eşitlik (3.31) de iki vektörün nokta çarpımlarının sıfır olması durumunda bu iki vektörün birbirine ortogonal olduğu belirtilmişti. Ortogonalite vektör uzunlukları dikkate alınarak da ifade edilebilir. Şekil 3.9'dan görüldüğü gibi, eğer ($\mathbf{x}_1$+$\mathbf{x}_2$) 'nin uzunluğu ($\mathbf{x}_1$-$\mathbf{x}_2$) 'nin uzunluğuna eşit ise $\mathbf{x}_1$ ile $\mathbf{x}_2$ vektörleri birbirine ortogonaldir:

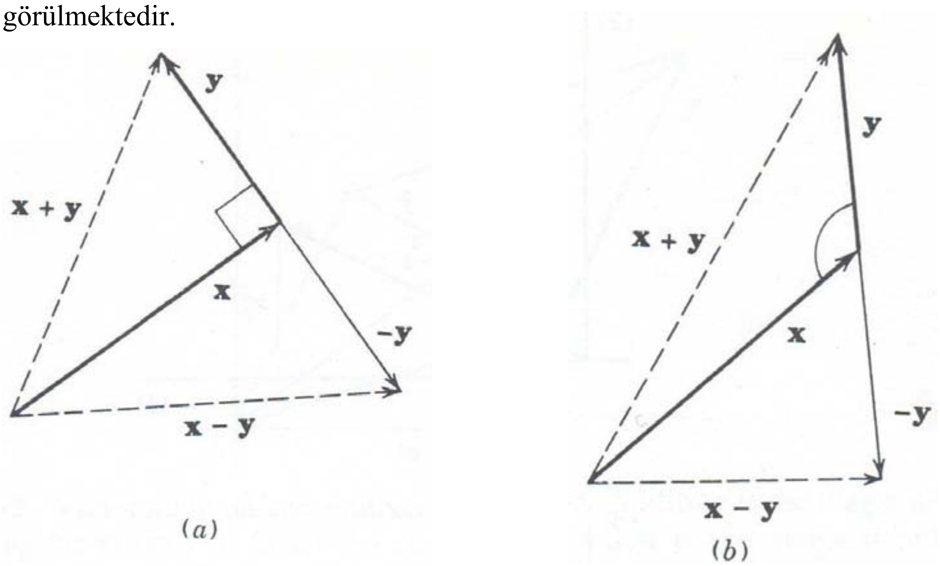
$$\|\mathbf{x}_1 + \mathbf{x}_2\|^2 = \|\mathbf{x}_1 - \mathbf{x}_2\|^2$$

$$(\mathbf{x}_1 + \mathbf{x}_2)^T(\mathbf{x}_1 + \mathbf{x}_2) = (\mathbf{x}_1 - \mathbf{x}_2)^T(\mathbf{x}_1 - \mathbf{x}_2)$$

$$\mathbf{x}_1^T\mathbf{x}_1 + 2\mathbf{x}_1^T\mathbf{x}_2 + \mathbf{x}_2^T\mathbf{x}_2 = \mathbf{x}_1^T\mathbf{x}_1 - 2\mathbf{x}_1^T\mathbf{x}_2 + \mathbf{x}_2^T\mathbf{x}_2$$

$$4\mathbf{x}_1^T\mathbf{x}_2 = 0$$

Sonuç olarak ortogonallik özelliğinin sağlanabilmesi için eşitlik (3.31) in sağlanması gerektiği görülmektedir.



Şekil 3.9 *Bu şekiller herhangi bir boyut için de geçerlidir.* a) $\mathbf{x}_1$ *ve* $\mathbf{x}_2$ *ortogonal* b) $\mathbf{x}_1$ *ve* $\mathbf{x}_2$ *ortogonal değil.*

Tam bir üç boyutlu uzay oluşturmak için diğerlerinden bağımsız olarak belirtilebilecek üçüncü bir vektöre ihtiyaç vardır. Bu nedenle *P* düzleminin dışına çıkılır. Bunun sonucu olarak bu üç boyutlu uzaydaki noktalar setinin bütünü

$$T:\ \lambda_1\mathbf{x}_1+\lambda_2\mathbf{x}_2+\lambda_3\mathbf{x}_3 \qquad -\infty<\lambda_1,\lambda_2,\lambda_3<\infty \tag{3.33}$$

ifadesiyle oluşturulabilir. Bu ifade üç boyutlu uzayda $\mathbf{x}_1$, $\mathbf{x}_2$ ve $\mathbf{x}_3$ oluşumu (veya zinciri) olarak belirtilir. Başka bir deyişle $\mathbf{x}_1$, $\mathbf{x}_2$ ve $\mathbf{x}_3$ bu üç boyutlu uzayın temelini oluşturmaktadır. Elemanları gerçel sayı olan üç elemanlı bir vektör üç boyutlu uzayda bir nokta tanımlar. Bu üç boyutlu uzay IR^3 ile tanımlanır ve bu uzaydaki herhangi bir **v** vektörü üç doğrusal bağımsız vektörün uygun bir setinin eşsiz doğrusal kombinasyonu olarak ifade edilebilir. Bu üç vektör IR^3 için bir kaynak oluşturur. Bu kaynağı oluşturan vektörler,

$$\mathbf{e}_1=\begin{bmatrix}1\\0\\0\end{bmatrix} \qquad \mathbf{e}_2=\begin{bmatrix}0\\1\\0\end{bmatrix} \qquad \mathbf{e}_3=\begin{bmatrix}0\\0\\1\end{bmatrix}$$

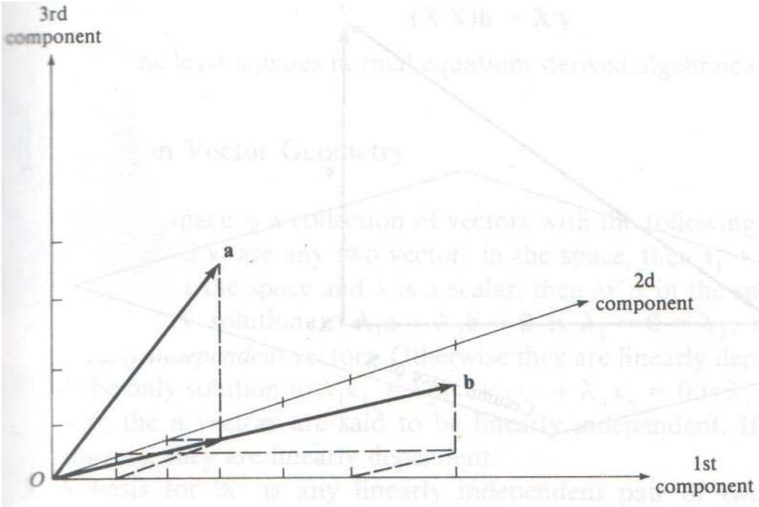
şeklinde seçildiklerinde bir $\mathbf{v}^T=\begin{bmatrix}3 & -2 & 5\end{bmatrix}$ vektörü ,

$$\mathbf{v}=3\mathbf{e}_1-2\mathbf{e}_2+5\mathbf{e}_3$$

eşitliği ile verilebilir. Bu vektörlerden herhangi iki tanesi (örneğin $\mathbf{e}_1$, $\mathbf{e}_2$) ele alındığında bunların tüm doğrusal kombinasyonları IR^3 uzayında bir alt vektör uzayı oluşturur. Bunun nedeni herbir zincir vektöründe üçüncü bileşenin sıfır olmasıdır. Örneğin;

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \qquad \mathbf{x}_2 = \begin{bmatrix} 5 \\ 1 \\ 1 \end{bmatrix}$$

şeklinde üç elemanlı iki vektör ele alındığında bu vektörler Şekil 3.10'da görüldüğü gibi bir düzlem yüzeyi oluştururlar. Şekilde bu düzlem $\mathbf{x}_1\mathrm{O}\mathbf{x}_2$ ile belirtilmiştir.



Şekil 3.10

Üç boyutlu uzay için de $\|\mathbf{x}_1\|^2$ değerinin vektörün karesel uzunluğu olduğu doğrulanabilir. Örneğin Şekil 3.8*b*'de Pisagor teoremi ilk olarak ABC üçgenine uygulanarak AC'nin karesel uzaklığı $x_{11}^2 + x_{21}^2$ elde edilir. Daha sonra teorem ACD üçgenine uygulanarak, AD vektörünün karesel uzunluğu

$$(x_{11}^2 + x_{21}^2) + x_{31}^2 = \|\mathbf{x}_1\|^2 \tag{3.34}$$

olarak elde edilir ve teorem üç boyutlu uzay için de doğrulanmış olur. $\mathbf{x}_1$=(2 4 3) şeklinde verilmiş vektörün karesel uzunluğu için, $\|\mathbf{x}_1\|^2 = \mathbf{x}_1\mathbf{x}_1^T = 2^2 + 4^2 + 3^2 = 29$ ve uzunluğu için ise $\|\mathbf{x}_1\| = \sqrt{29}$ sonucu elde edilir.

Eşitlik (3.33), n boyutlu uzay için genelleştirilebilir.

$$M\!: \lambda_1\mathbf{x}_1+\lambda_2\mathbf{x}_2+\ldots+\lambda_m\mathbf{x}_m \qquad -\infty < \lambda_i < \infty \tag{3.35a}$$

Eşitlik (3.35*a*), m adet sabit vektörün tüm mümkün doğrusal kombinasyonlarının setidir ve m boyutlu bir alt uzay olarak adlandırılır. Eğer m=1 ise alt uzay düz bir doğru m=2 ise alt uzay bir düzlemdir. m>2 olması durumunda ise alt uzay bir hiper düzlem olarak adlandırılır. Sadece m=n olması durumunda n adet birbirinden bağımsız vektör mevcuttur ($\mathbf{x}_1$, $\mathbf{x}_2$,…, $\mathbf{x}_n$) ve bu vektörlerin tümü n boyutlu uzayı oluşturur. Bu n boyutlu uzaydaki herhangi bir vektör veya bir y noktası için λ_1, λ_2,…, λ_m katsayılarının sadece ve sadece bir tek seti mevcuttur ve bu set,

$$y=\lambda_1\mathbf{x}_1+\lambda_2\mathbf{x}_2+\ldots+\lambda_n\mathbf{x}_n \tag{3.35b}$$

eşitliği ile bulunabilir. Bu katsayılar, ($\mathbf{x}_1$, $\mathbf{x}_2$,…, $\mathbf{x}_n$) vektörlerine göre y'nin koordinatları olarak adlandırılırlar. Konu daha da genelleştirilirse, tüm n adet gerçel eleman içeren vektörler IR^n uzayını oluştururlar. IR^n 'deki her bir vektör, n doğrusal bağımsız vektörün bazı uygun setlerinin bir eşsiz doğrusal kombinasyonu olarak ifade edilebilir. Bu doğrusal kombinasyon eşsiz olmak zorundadır. Bunu görebilmek için bir $\mathbf{v}$ vektörünü kaynak $\mathbf{v}_1$, $\mathbf{v}_2$,…,$\mathbf{v}_n$ vektörlerinin iki farklı doğrusal kombinasyonu olarak,

$$\mathbf{v} = \lambda_1\mathbf{v}_1 + \lambda_2\mathbf{v}_2 + \cdots + \lambda_n\mathbf{v}_n$$
$$\mathbf{v} = \mu_1\mathbf{v}_1 + \mu_2\mathbf{v}_2 + \cdots + \mu_n\mathbf{v}_n$$

şeklinde ifade edilebildiği varsayılsın. Bu iki denklem birbirinden çıkarılarak,

$$0 = (\lambda_1 - \mu_1)\mathbf{v}_1 + (\lambda_2 - \mu_2)\mathbf{v}_2 + \cdots + (\lambda_n - \mu_n)\mathbf{v}_n$$

eşitliği elde edilir. Kaynak veriler doğrusal bağımsız oldukları için,

$$\lambda_1 - \mu_1 = \lambda_2 - \mu_2 = \cdots = \lambda_n - \mu_n$$

ifadesi elde edilebilir ve görüldüğü gibi bu ifade eşsizdir. Eğer k adet n elemanlı doğrusal bağımsız vektörün bir seti alınmış ise bu set IR^n 'in bir alt uzayını oluşturur. Bu alt uzayın boyutu bu alt uzaydaki doğrusal bağımsız vektör sayısına eşittir. n elemanlı vektörler için $\mathbf{v}_1$, $\mathbf{v}_2$, … ,$\mathbf{v}_k$ eğer,

$$\mathbf{v}_i\mathbf{v}_j = 0 \qquad \text{tüm } i \neq j \text{ için}$$

sonucu sağlanıyorsa bu vektör seti ayrık ortogonal settir.

3.6 VEKTÖR GEOMETRİSİNDE ORTOGONAL İZDÜŞÜM

İki boyutlu uzayda $\mathbf{x}_1$=(1 -1) ve $\mathbf{x}_2$=(2 1) vektörleri dikkate alınsın. Bu vektörler bütün uzayı tanımlayacaklardır. $\mathbf{y}$ = (4 1) şeklinde verilen bir vektörün $\mathbf{x}_1$ ve $\mathbf{x}_2$'ye göre koordinatlarını bulabilmek için eşitlik (3.27*a*) kullanılarak,

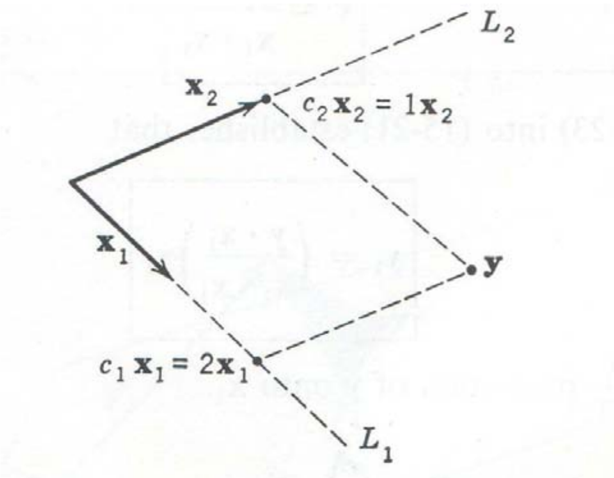
$$\lambda_1\mathbf{x}_1^T + \lambda_2\mathbf{x}_2^T = \mathbf{y}$$

$$\lambda_1\begin{bmatrix}1\\-1\end{bmatrix} + \lambda_2\begin{bmatrix}2\\1\end{bmatrix} = \begin{bmatrix}4\\-1\end{bmatrix}$$

$$\lambda_1 + 2\lambda_2 = 4$$

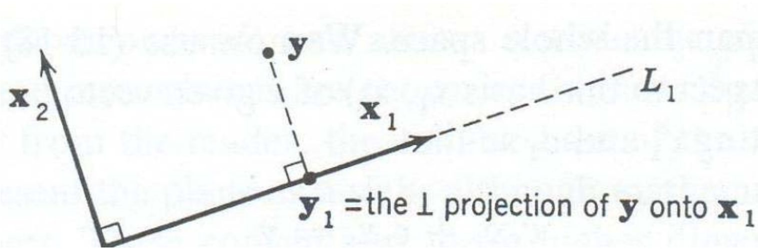
$$-\lambda_1 + \lambda_2 = -1$$

yazılabilir ve bu denklem sisteminin çözümüyle λ_1=2, λ_2=1 elde edilir. $\mathbf{y}$ vektörünün $\mathbf{x}_1$ ve $\mathbf{x}_2$**'nin** doğrusal bir kombinasyonu olarak ifade edileceği görülmektedir. Bu durum Şekil 3.11'de geometrik olarak görülmektedir. L_1 doğrusunun (alt uzayının) $\mathbf{x}_1$ tarafından L_2 alt uzayının ise $\mathbf{x}_2$ tarafından oluşturulduğu görülmektedir. Daha sonra λ_1=2 ve λ_2=1 alınarak bir paralel kenar ile şekil tamamlanmıştır.

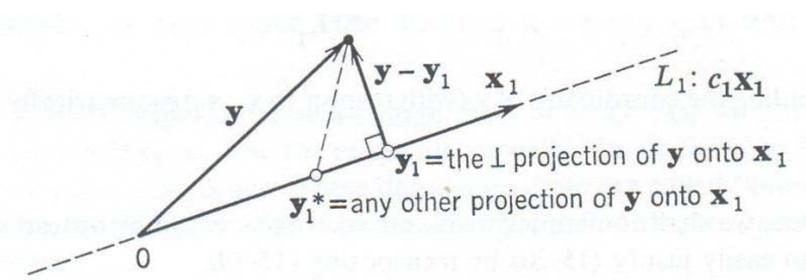


Şekil 3.11 *İzdüşüm ile* **y**'*nin koordinatlarının* $\mathbf{x}_1$ *ve* $\mathbf{x}_2$'*ye göre geometrik olarak ifade edilmesi*.

Başka bir deyişle, **y**'nin koordinatlarını bulmak için ilk olarak L_2'ye paralel olacak şekilde L_1 üzerine **y**'nin izdüşümü alınarak λ_1 elde edilir. Daha sonra benzer bir işlem λ_2 için yapılır, L_1'e paralel olacak şekilde L_2 üzerine **y**'nin izdüşümü alınır. En basit izdüşüm yapısı ortogonal izdüşümdür. Bu durum $\mathbf{x}_1$ ile $\mathbf{x}_2$'nin birbirine ortogonal olmaları durumunda ortaya çıkar, (bkz Şekil 3.12).



Şekil 3.12 $\mathbf{x}_1$ *üzerine* **y**'*nin ortogonal izdüşüm*



Şekil 3.13 *İki boyutlu ortogonal izdüşüm*

Eşitlik (3.31) ile tanımlanan vektörler için ortogonallik şartı oldukça basit olduğu için ortogonal izdüşümü hesaplamak da oldukça basittir. $\mathbf{x}_1$ vektörünün tanımladığı L_1 doğrusu üzerine **y**'nin ortogonal izdüşümü $\mathbf{y}_1$ ile belirtilmiş ve Şekil 3.13'de gösterilmiştir. $\mathbf{y}_1$ izdüşüm vektörü L_1 üzerinde bulunacağı için bu vektör $\mathbf{x}_1$'in bir skalerle çarpımı şeklinde

$$\mathbf{y}_1=\lambda\mathbf{x}_1 \qquad (3.36)$$

ifade edilebilir. Burada sorun λ katsayısının belirlenebilmesidir. Şekil 3.13'den görülebileceği gibi $\mathbf{y}_1$ ile **y** birleştirilerek $\mathbf{y}-\mathbf{y}_1$ vektörü tanımlanabilir. Bu işlemde elde edilen vektör ile $\mathbf{x}_1$ arasındaki ortogonal yapı korunmalıdır. ($\mathbf{y}-\mathbf{y}_1$) ve $\mathbf{x}_1$ ortogonal olduğundan eşitlikler (3.31) ve (3.36) kullanılarak,

$$(\mathbf{y}-\lambda\mathbf{x}_1)^T\,\mathbf{x}_1=0$$

$$\mathbf{y}^T\mathbf{x}_1-\lambda\left(\mathbf{x}_1^T\mathbf{x}_1\right)=0$$

$$\lambda=\frac{\mathbf{y}^T\mathbf{x}_1}{\mathbf{x}_1^T\mathbf{x}_1} \tag{3.37}$$

elde edilir. Eşitlik (3.37), eşitlik (3.36)'da yerine konarak

$$\mathbf{y}_1=\left(\frac{\mathbf{y}^T\mathbf{x}_1}{\mathbf{x}_1^T\mathbf{x}_1}\right)^T\mathbf{x}_1 \tag{3.38}$$

sonucu bulunabilir. Eşitlik (3.38) in sağındaki ilk bileşenin payı ve paydası bir skaleri tanımladığı için transpozları kendilerine eşittir. Bu eşitlikteki $\mathbf{y}_1$ ifadesi $\mathbf{x}_1$ üzerine $\mathbf{y}$ nin ortogonal izdüşümünü belirtmektedir. Bu izdüşüm vektörünün karesel uzunluğu,

$$\|\mathbf{y}_1\|^2=\mathbf{y}_1^T\mathbf{y}_1$$

$$=\lambda^2\mathbf{x}_1^T\mathbf{x}_1$$

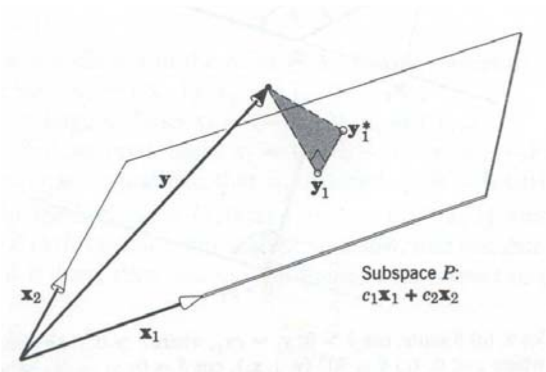
$$=\left[\frac{\mathbf{y}^T\mathbf{x}_1}{\mathbf{x}_1^T\mathbf{x}_1}\right]^2\mathbf{x}_1^T\mathbf{x}_1$$

$$\|\mathbf{y}_1\|^2=\frac{\left(\mathbf{y}^T\mathbf{x}_1\right)^2}{\mathbf{x}_1^T\mathbf{x}_1} \tag{3.39}$$

olup, normu veya uzunluğu ise

$$\|\mathbf{y}_1\|=\frac{\left|\mathbf{y}^T\mathbf{x}_1\right|}{\|\mathbf{x}_1\|} \tag{3.40}$$

şeklindedir. Şekil 3.13 incelendiğinde, ortogonal izdüşüm $\mathbf{y}_1$ in L_1 üzerindeki $\mathbf{y}$ ye en yakın alt nokta olduğu görülebilir. Ortogonal olmayan ve $\mathbf{y}_1^*$ ile gösterilebilecek diğer herhangi bir izdüşümün $\mathbf{y}$ noktasına olan uzaklığı daha fazla olacaktır. $\|\mathbf{y}-\mathbf{y}_1^*\|$ uzaklığı $\|\mathbf{y}-\mathbf{y}_1\|$ uzaklığından daha büyüktür. Çünkü $\|\mathbf{y}-\mathbf{y}_1^*\|$ dik üçgenin üçgeninin hipotenüsüdür. Üç boyutlu durum Şekil 3.14'de gösterilmiştir.



Şekil 3.14 *Üç boyutlu uzayda ortogonal izdüşüm*

Teorem 3.7. $\lambda_1\mathbf{x}_1+\lambda_2\mathbf{x}_2+\ldots+\lambda_m\mathbf{x}_m$ alt uzayı üzerine $\mathbf{y}$'nin ortogonal izdüşümü bu alt uzaydaki $\mathbf{y}$'ye en yakın noktayı belirtir.

3.7 DENKLEM SİSTEMLERİNİN ÇÖZÜMÜ

Bir doğrusal denklem sistemi,

$$a_1 b_{11} + a_2 b_{12} + \cdots + a_m b_{1m} = c_1$$

$$a_1 b_{21} + a_2 b_{22} + \cdots + a_m b_{2m} = c_2$$

$$\vdots \qquad\qquad \vdots \qquad \vdots$$

$$a_1 b_{n1} + a_2 b_{n2} + \cdots + a_m b_{nm} = c_n$$

şeklinde verilebilir. Bu sistem matris notasyonunda,

$$\mathbf{Ba} = \mathbf{c} \tag{3.41}$$

ifadesi ile verilebilir. Bu eşitlikte **B**, $n \times m$ boyutlu matris **a**, $m \times 1$ boyutlu vektör **c** ise $n \times 1$ boyutlu bir vektördür. Verilen b_{ij} ve c_j seti için **Ba**=**c** eşitliğini sağlayan bir a_i seti mevcut mudur? Bu soru için dikkate alınabilecek üç durum sözkonusudur.

1) Eşitliğin çözümü yoktur. Bu durumda sistemin eşitliğini sağlayan bir **a** vektörü yoktur ve sistem tutarsızdır.

2) Sistemin eşitliğini sağlayan bir tek a_i seti vardır. Bu durumda sistemin bir tek çözümü vardır.

3) Sistemin eşitliğini sağlayan birden fazla **a** vektörü vardır. Eğer eşitliği sağlayan birden fazla **a** vektörü varsa sonsuz sayıda çözüm bulunabilir.

Denklem sisteminde katsayılar matrisi **B**'dir. Bu matrise **c** vektörünün bir sütun olarak eklenip sütun boyutunun bir arttırılması ile elde edilen **G** matrisi genişletilmiş matristir.

$$G = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1m} & c_1 \\ b_{21} & b_{22} & \cdots & b_{2m} & c_2 \\ \vdots & \vdots & & \vdots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nm} & c_n \end{bmatrix} \tag{3.42}$$

Bu iki matris dikkate alınarak **Ba**=**c** denklem sisteminin çözümü ile ilgili bazı önemli teoremler aşağıda verilmiştir.

Teorem 3.8: **Ba**=**c** denklem sisteminin tutarlı olabilmesi (bu sistemi sağlayan en az bir adet **a** vektörünün mevcut olması) için gerek ve yeter koşul, katsayı matrisinin **B** rankı ile genişletilmiş matrisin **G** rankının eşit olması gereklidir.

$$\rho(\mathbf{B}) = \rho(\mathbf{G}) \tag{3.43}$$

Teorem 3.9: Eğer ρ(**B**)= ρ(**G**)=p ve $p<m$ ise bilinmeyen m-p adet a_i değeri keyfi olarak atanır ve kalan p adet a_i eşsiz olarak belirlenir.

Teorem 3.10: Eğer ρ(**B**)= ρ(**G**)=m ve $m<n$ ise **Ba**=**c** denklemini sağlayan bir eşsiz **a** vektörü mevcuttur.

Bir denklem sisteminin çözüm aşamasında genellikle matrisin tersinin alınmasına gereksinim duyulur. **A** matrisi *n* boyutlu bir kare matris ve sütunlarının hepside doğrusal bağımsız olsun. Bu durumda,

$$\mathbf{AB} = \mathbf{I} \tag{3.44}$$

eşitliğini sağlayan *n* boyutlu bir **B** matrisi araştırılır. **B** matrisinin ilk sütunu $\mathbf{b}_1$ ile belirtilirse eşitlik (3.44)'in her iki tarafındaki birinci eşitlik bir vektör denklemini,

$$\mathbf{A}\mathbf{b}_1 = \mathbf{e}_1 \tag{3.45}$$

verir. Bu eşitlikteki $\mathbf{e}_1^T = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \end{bmatrix}$ şeklinde tanımlanır. **A** matrisinin sütunları doğrusal bağımsız olduğu için $\mathbf{e}_1$ vektörü bu sütunların eşsiz bir doğrusal kombinasyonu olarak ifade edilebilir. Bunun sonucunda $\mathbf{b}_1$ eşsiz olarak belirlenebilir. **B** matrisinin her bir sütunu da benzer bir şekilde eşsiz olarak belirlenebilir. Sonuç olarak eşitlik (3.44)'i sağlayan bir **B** matrisi olduğu görülebilir.

Eğer **A** matrisinin *n* sütunu doğrusal bağımsız ise *n* adet sırası da doğrusal bağımsızdır. Benzer bir yaklaşım kullanılarak,

$$\mathbf{CA} = \mathbf{I} \tag{3.46}$$

eşitliğini sağlayan *n* boyutlu bir **C** matrisi bulunabilir. Bu eşitlikte **C** matrisinin her bir sırası **A** matrisinin sıralarının doğrusal bir kombinasyonunun katsayıları olarak eşsiz bir şekilde belirlenebilir. Buna göre eşitlik (3.44) ve (3.46)'un her ikisi de doğrudur. Eşitlik (3.46) sağdan **B** matrisi ile çarpılırsa,

$$\mathbf{CAB} = \mathbf{IB} = \mathbf{B}$$

ve benzer şekilde

$$\mathbf{CAB} = \mathbf{IC} = \mathbf{C}$$

elde edilir. Buna göre, **B**=**C** olduğu görülmektedir. Eğer **A** matrisinin *n* sütunu (veya sırası) doğrusal bağımsız ise **A** matrisinin tersi olarak adlandırılıp $\mathbf{A}^{-1}$ ile gösterilen *n* boyutlu eşsiz bir kare matris mevcuttur ve eşitlik (3.44) veya (3.46),

$$\mathbf{A}\mathbf{A}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I} \tag{3.47}$$

şeklinde verilebilir. Bir matrisin tersinin alınabilmesi konusu matrisin rankı kavramı ile yakın olarak ilgilidir.

3.8 BİR MATRİSİN RANKI

Herhangi bir $m \times n$ boyutlu **A** matrisi ele alınsın. **A** matrisinin sütunları IR^m uzayındaki *n* vektörü tanımlar. Aynı şekilde **A**'nın sıraları da IR^n uzayında *m* vektörü belirler. **A** matrisindeki maksimum doğrusal bağımsız sıra sayısı *r* ile belirtilsin, $r \leq m$. Eğer *r*, *m*'den küçük ise doğrusal bağımsız sıra vektörlerinin alt seti sayısı birden daha fazla olabilir. Örneğin, dört sıralı (*m*=4) bir matris ele alındığında, 1,2,4 sıraları bir doğrusal bağımsız set tanımlayabildiği gibi 1,3,4 sıraları da bir diğer doğrusal bağımsız seti tanımlayabilir. Fakat dört sıranın tümü doğrusal bağımlıdır. Bu durumda

r=3'dür. Bu **A** matrisi için herhangi r doğrusal bağımsız sıranın bir setinin oluşturduğu ve kalan m-r sıranın ihmal edildiği $r\times n$ boyutlu bir $\tilde{\mathbf{A}}$ matrisi tanımlanabilir. **A** matrisindeki maksimum doğrusal bağımsız sütun sayısı c ile belirtilsin. Bu c değeri aynı zamanda $\tilde{\mathbf{A}}$ matrisindeki maksimum doğrusal bağımsız sütun sayısını da belirtir. $\tilde{\mathbf{A}}$ matrisindeki her bir sütun r elemana sahiptir. Buna göre,

$$c \le r \tag{3.48}$$

eşitsizliği verilebilir. IR^n uzayındaki herhangi bir vektör, r adet doğrusal bağımsız vektörün doğrusal bir kombinasyonu şeklinde verilebilir. Yukarıda sıralar için verilen işlem benzer şekilde sütunlar içinde gerçekleştirilebilir. **A** matrisindeki c adet doğrusal bağımsız sütunun bir alt seti alınır ve kalan n–c adet sütun ihmal edilerek $m\times c$ boyutlu bir $\approx\!\mathbf{A}$ matrisi oluşturulabilir. **A** matrisindeki maksimum doğrusal bağımsız sıra sayısı r ile belirtildiği için bu değer aynı zamanda $\approx\!\mathbf{A}$ matrisindeki maksimum doğrusal bağımsız sıra sayısını belirtmektedir. Fakat $\tilde{\mathbf{A}}$'daki her bir sıra c elemana sahip olduğu için,

$$r \le c \tag{3.49}$$

şeklindedir. Eşitlik (3.48) ve (3.49) birlikte ele alınarak,

$$r=c \tag{3.50}$$

elde edilebilir. Sonuç olarak $m\times n$ boyutlu bir **A** matrisinin maksimum doğrusal bağımsız sıra sayısı, maksimum doğrusal bağımsız sütun sayısına eşittir. Bu sayı matrisin rankı olarak tanımlanır ve $\rho(\mathbf{A})$ sembolü ile gösterilir. Açık olarak görüleceği gibi bir matrisin rankı satır veya sütun sayılarından en küçük olanını aşamaz. Başka bir deyişle,

$$\rho(\mathbf{A}) \le \min(m,n) \tag{3.51}$$

eşitsizliği verilebilir. $\rho(\mathbf{A})=m$ durumunda matris tam sıra ranklı $\rho(\mathbf{A})=n$ durumunda ise tam sütun ranklıdır. Ranklarla ilgili teoremler ve bazı açıklamalar aşağıda verilmiştir.

Teorem 3.11: **A** matrisinin transpozunun rankı **A** matrisinin rankına eşittir.

$$\rho\left(\mathbf{A}^T\right)=\rho(\mathbf{A}) \tag{3.52}$$

Özel bir durum olarak **A**, $n\times n$ boyutlu rankı n olan bir kare matris ise **A** matrisi tekil değildir ve $\mathbf{A}^{-1}$ ile belirtilen eşsiz bir ters matrisi mevcuttur.

$$\mathbf{A}\mathbf{A}^{-1}=\mathbf{A}^{-1}\mathbf{A}=\mathbf{I}_n$$

Bu **A** matrisinin rankı n'den küçük ise tekil matristir ve ters matrisi mevcut değildir. Boyutları $m\times n$ olan ve rankı $\rho(\mathbf{A})=r$ olan bir **A** matrisi ele alındığında en az bir tane r adet doğrusal bağımsız sıranın oluşturduğu set ve en az bir tane de r adet doğrusal bağımsız sütunun oluşturduğu set mevcuttur. Bunun sonucu olarak **A** matrisinin satır ve sütunları, ilk sırası ve ilk sütunu doğrusal bağımsız olacak şekilde düzenlenebilir ve daha sonra ilk r sıra ve sütun dikkate alınarak parçalanabilir.

$$\mathbf{A}=\left[\begin{array}{c:c}\mathbf{A}_{11} & \mathbf{A}_{12} \\ \hdashline \mathbf{A}_{21} & \mathbf{A}_{22}\end{array}\right]$$

Elde edilen $\mathbf{A}_{11}$ matrisi $r \times r$ boyutlu tekil olmayan bir kare matristir. Homojen denklem seti,

$$\mathbf{Ax}=\mathbf{0} \tag{3.53}$$

şeklinde verilebilir. $\mathbf{x}$; n adet bilinmeyen eleman içeren sütun vektörüdür. Eşitlik (3.53)'in sağ tarafı bir $\mathbf{0}$ vektörü olduğu için bu denklemlerin homojen olduğu söylenebilir. Eğer $\mathbf{Ax}=\mathbf{b}$ ve $\mathbf{b}\neq 0$ ise denklemlerin homojen olmadığı söylenebilir.

Eğer $\mathbf{x}_1$, eşitlik (3.53)'in bir çözümü ise c herhangi bir skaler olmak üzere $c\mathbf{x}_1$'de bu eşitliğin bir çözümüdür.

Eşitlik (3.53)'in bu çözümler seti $\mathbf{A}$ matrisinin boş uzayı olarak bilinen bir vektör uzayını oluştururlar. İlk olarak bu boş uzayın boyutu araştırılacaktır. Bu boyut alt uzayı tanımlayan birbirinden bağımsız vektör sayısıdır. $\mathbf{A}$ matrisinin son $m–r$ sırası çıkarılıp ve $\mathbf{x}$ vektörü de $\mathbf{A}$ matrisinin sütunlarına uygun olarak parçalanırsa,

$$\begin{bmatrix}\mathbf{A}_{11} & \mathbf{A}_{12}\end{bmatrix}\begin{bmatrix}\mathbf{x}_1 \\ \mathbf{x}_2\end{bmatrix}=\mathbf{0} \tag{3.54}$$

eşitliği elde edilir. Bu ifade de $\mathbf{x}_1$, r adet $\mathbf{x}_2$ ise kalan $n–r$ adet elemanı içerir. Sonuç olarak n adet ($n \geq r$) bilinmeyen içeren r adet doğrusal bağımsız denklemli bir set elde edilir. Eşitlik (3.54),

$$\mathbf{A}_{11}\mathbf{x}_1+\mathbf{A}_{12}\mathbf{x}_2=\mathbf{0}$$

şeklinde yeniden yazılabilir. $\mathbf{A}_{11}$ tekil olmadığı için tersi mevcuttur ve bu ifade soldan $\mathbf{A}_{11}^{-1}$ ile çarpılarak,

$$\mathbf{x}_1=-\mathbf{A}_{11}^{-1}\mathbf{A}_{12}\mathbf{x}_2 \tag{3.55}$$

elde edilir. $\mathbf{x}_2$ alt vektöründeki n-r eleman herhangi bir şekilde belirlenebilir. Fakat $\mathbf{x}_2$ belirlendikten sonra $\mathbf{x}_1$ alt vektörü eşitlik (3.55) kullanılarak belirlenir. Eşitlik (3.55) kullanılarak eşitlik (3.54) deki çözüm vektörü,

$$\mathbf{x}=\begin{bmatrix}\mathbf{x}_1 \\ \mathbf{x}_2\end{bmatrix}=\begin{bmatrix}-\mathbf{A}_{11}^{-1}\mathbf{A}_{12} \\ \mathbf{I}_{n\text{-}r}\end{bmatrix}\mathbf{x}_2 \tag{3.56}$$

şeklinde düzenlenebilir. Elde edilen (3.56) eşitliği n sıra ve $n–r$ sütuna sahiptir. Bu $n–r$ sütun doğrusal bağımsızdır. Bunun nedeni de sütunları doğrusal bağımsız olan $\mathbf{I}_{n\text{-}r}$ alt matrisinin mevcut olmasıdır. Bir birim matrisin sütunları (sıraları) doğrusal bağımsızdır. Sonuç olarak eşitlik (3.56), $n–r$ adet n elemanlı doğrusal bağımsız vektörü eşitlik (3.54)'in tüm çözümleri için $n–r$ adet doğrusal kombinasyon şeklinde ifade eder. Eşitlik (3.54), çıkarılan sıralar [$\mathbf{A}_{11}$ $\mathbf{A}_{12}$]'nin sıralarının doğrusal bir kombinasyonu olduğu için, eşitlik (3.53) içinde bir çözüm verir. Buna uygun olarak çıkarılan sıralar,

$$\mathbf{c}^T\begin{bmatrix}\mathbf{A}_{11} & \mathbf{A}_{12}\end{bmatrix}$$

formunda gösterilebilir. Burada $\mathbf{c}^T$ herhangi uygun bir r elemanlı sıra vektörüdür. Bu ifade sağdan $\mathbf{x}$ ile çarpılabilir.

$$\mathbf{c}^T\begin{bmatrix}\mathbf{A}_{11} & \mathbf{A}_{12}\end{bmatrix}\mathbf{x}=\mathbf{0}$$

Bunun nedeni ise, **x**'in eşitlik (3.54)'u sağlamasıdır. Buna göre her bir **x** çözümü çıkarılan sıralara göre ele alınır ve eşitlik (3.56), eşitlik (3.53) için bu çözüm vektörünü belirler. **A** matrisinin boş uzayı $n-r$ boyutludur. Buradan çıkarılacak önemli bir sonuç: r ranklı ve boyutları $m\times n$ olan bir **A** matrisi için; *Sütun sayısı = rank + boş uzay boyutu*

$$n \;=\; r \;+\; (n-r) \tag{3.57}$$

eşitliği yazılabilir.

Eşitlik (3.57)'in özel bir durumu mxn boyutlu bir **A** matrisinin rankının $n-1$ olmasıdır. Bu durumda **A** matrisinin boş uzayının boyutu 1'dir ve

$$\mathbf{Ax}=\mathbf{0}$$

denkleminin tüm çözümleri orjinden geçen bir tek doğru üzerindedir. Eğer çözüm vektörü,

$$\mathbf{x}^T=\begin{bmatrix}\mathbf{x}_1 & \mathbf{x}_2 & \cdots & \mathbf{x}_n\end{bmatrix}$$

ise

$$c\mathbf{x}^T=\begin{bmatrix}c\mathbf{x}_1 & c\mathbf{x}_2 & \cdots & c\mathbf{x}_n\end{bmatrix}$$

vektörü de (c herhangi bir sabit) bir çözüm vektörüdür. Eşitlik (3.57)'in sonucu çeşitli matrislerin rankları üzerine verilen bazı önemli teoremlerin basit ispatlarını oluşturur.

Teorem 3.12: Eğer **A**, $m\times n$ boyutlu r ranklı herhangi bir matris ve P $m\times m$ boyutlu, Q ise $n\times n$ boyutlu tekil olmayan kare matrisler ise,

$$\rho(\mathbf{PA})=\rho(\mathbf{AQ})=\rho(\mathbf{PAQ})=\rho(\mathbf{A}) \tag{3.58}$$

eşitlikleri verilebilir. Başka bir deyişle, **A** matrisi tekil olmayan bir matrisle soldan veya sağdan çarpılırsa **A** matrisinin rankı değişmez. $\rho(\mathbf{PA})=\rho(\mathbf{A})$ eşitliğini ispat etmek için **A** matrisinin boş uzayındaki herhangi bir vektör olarak m ele alınsın; **Am**=**0** ve **PAm**=**0** olacak ve sonuç olarak m vektörü aynı zamanda **PA**'nın boş uzayında da yer alacaktır. Buna karşın **PA**'nın boş uzayındaki herhangi bir vektör **s** ile belirtilerek, **PAs**=**0** yazılabilir. **P** tekil olmayan bir matris olduğu için bu eşitlik soldan $\mathbf{P}^{-1}$ ile çarpılarak, **As**=**0** elde edilir. Görüleceği gibi **s** aynı zamanda **A**'nın boş uzayındadır. Sonuç olarak **PA** ve **A** aynı boş uzaya sahiptirler. Bu nedenle rankları aynıdır. Bu ispatlardan yararlanarak, $\rho(\mathbf{AQ})=\rho(\mathbf{A})$ için ispat,

$$\begin{aligned}\rho(\mathbf{AQ})&=\rho(\mathbf{Q}^T\mathbf{A}^T)\\&=\rho(\mathbf{A}^T)\end{aligned}$$

$= \rho(\mathbf{A})$

şeklinde yapılabilir. Tüm bu sonuçlar ile,

$$\rho(\mathbf{PAQ}) = \rho(\mathbf{A})$$

eşitliği yazılabilir. Ranklarla ilgili diğer bir teorem ise dikdörtgen bir matris ile diğer uygun bir dikdörtgen matrisin çarpımları için genel durumu ifade eder.

Teorem 3.13: **A** matrisi $m \times n$, **B** matrisi $n \times s$ boyutlu ise **AB** çarpımının rankı, bu matrislerden rankı daha küçük olanın rankına eşit veya daha küçüktür.

$$\rho(\mathbf{AB}) \leq \min\left[\rho(\mathbf{A}), \rho(\mathbf{B})\right] \tag{3.59}$$

x vektörü **A** matrisinin boş uzayındaki herhangi bir vektör olsun, **Bx**=**0** ve **ABx**=**0** yazılabilir. Bu durumda **x** vektörü **AB**'nin boş uzayında da bulunmaktadır. Fakat **A** matrisi kare matris olmadığı için ters matrisi bulunamaz, bu nedenle, **ABy**=**0** işleminden **By**=**0** işlemine geçilemez. Sonuç olarak, **B**'nin boş uzayının boyutu ≤**AB**'nin boş uzayının boyutu eşitsizliği elde edilebilir. **AB** ve **B**'nin sütun sayısı aynı olduğu için eşitlik (3.57)'den, $\rho(\mathbf{AB}) \leq \rho(\mathbf{B})$ ve transpoz işlemi kullanılarak,

$$\rho(\mathbf{AB}) = \rho(\mathbf{B}^T\mathbf{A}^T) \leq \rho(\mathbf{A}^T) = \rho(\mathbf{A})$$

sonucu elde edilerek (3.59) eşitsizliği ispatlanmış olur. Ranklarla ilgili önemli bazı teoremler aşağıda ispatsız olarak verilecektir.

Teorem 3.14: Boyutları $n \times n$ olan iki kare matris **A** ve **B**'nin rankları r ve s ise **AB**'nin rankı ,

$$\rho(\mathbf{AB}) \geq r + s - n \tag{3.60}$$

eşitsizliğini sağlar.

Teorem 3.15: **A** ve **B** matrislerinin toplamının rankı,

$$\rho(\mathbf{A} + \mathbf{B}) \leq \rho(\mathbf{A}) + \rho(\mathbf{B}) \tag{3.61}$$

eşitsizliğini sağlar.

Teorem 3.16: Boyutları $n \times p$ olan bir **A** matrisi için:

$$\rho(\mathbf{A}^T\mathbf{A}) = \rho(\mathbf{A}\mathbf{A}^T) = \rho(\mathbf{A}^T) = \rho(\mathbf{A}).$$

3.9 TERS MATRİSLER

Bu kısımda ters matrislerin hesaplanması ve bazı özellikleri verilmiştir. **A** matrisinin $n \times n$ boyutlu bir kare matris olduğu kabul edilsin. Bu matrisin tersinin mevcut olması için gerekli şart aşağıda dört değişik tanımla verilmiştir. Bu tanımların hepsi aynı anlamı taşımaktadır.

1) **A** matrisi tekil olmayan bir matris olmalıdır.

2) **A** matrisinin rankı n olmalıdır.

3) **A** matrisinin n sırası doğrusal bağımsız olmalıdır.

4) **A** matrisinin *n* sütunu doğrusal bağımsız olmalıdır.

Ters matrisin elde edilmesi çalışmalarına 2×2 durumu ile başlanacaktır. **A** matrisi ve ters matrisi $\mathbf{A}^{-1}$ aşağıda verilmiştir.

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \qquad \mathbf{A}^{-1} = \begin{bmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{bmatrix}$$

Ters matris tanımından,

$$\mathbf{A}\mathbf{A}^{-1} = \mathbf{I} \tag{3.62}$$

genel denklemi elde edilir. İlk olarak eşitlik (3.62)'in her iki tarafının birinci sütunları ele alınsın,

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} \alpha_{11} \\ \alpha_{21} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

ters matrisin elemanları bilinmemektedir. Denklemler bu elemanlar için çözülürse,

$$\alpha_{11} = \frac{a_{22}}{a_{11}a_{22} - a_{21}a_{12}}$$

$$\alpha_{21} = \frac{-a_{21}}{a_{11}a_{22} - a_{21}a_{12}}$$

eşitlikleri elde edilir. Eşitlik (3.62)'in ikinci sütunları için,

$$\alpha_{12} = \frac{-a_{12}}{a_{11}a_{22} - a_{21}a_{12}}$$

$$\alpha_{22} = \frac{a_{11}}{a_{11}a_{22} - a_{21}a_{12}}$$

sonuçları elde edilir. Bu sonuçlar kullanılarak ters matris,

$$\mathbf{A}^{-1} = \frac{1}{a_{11}a_{22} - a_{21}a_{12}} \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix} \tag{3.63}$$

şeklinde bulunur. $\mathbf{A}^{-1}$ matrisindeki her bir eleman **A**'daki elemanların bir fonksiyonudur. Ters matristeki her bir eleman $a_{11}a_{22} - a_{12}a_{21}$ ortak bölenine sahiptir. Bu değer **A** matrisindeki tüm elemanların bir fonksiyonu olup **A** matrisinin determinantı olarak tanımlanır. 2×2 durumu için determinant,

$$\det(\mathbf{A}) = |\mathbf{A}| = a_{11}a_{22} - a_{12}a_{21} \tag{3.64}$$

olarak elde edilebilir. Determinant için daha genel ve kolay bir yöntem 3×3 boyutlu bir matris ,

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

şeklinde verilmiş olsun. Eğer a_{11} elemanına ait satır ve sütun çıkarılırsa,

$$\begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{bmatrix}$$

alt matrisi elde edilir. **A** matrisinden i-inci sıra ve j-inci sütun silindikten sonra elde edilen 2×2 boyutlu alt matrisin determinantı M_{ij} ile belirtilsin. M_{ij}, bir minör olarak isimlendirilir.

$$c_{ij} = (-1)^{i+j} \cdot M_{ij}$$

ifadesi ise kofaktör başka bir deyişle işaretli minördür. Eğer $i+j$ toplamı çift ise, M_{ij}'nin işareti değişmez, $i+j$ toplamı tek ise işaret değiştirir. Her bir a_{ij} elemanına karşılık gelen c_{ij} değerleri bulunarak elde edilen matris $|\mathbf{A}|$ ile bölünerek $\mathbf{A}^{-1}$ elde edilebilir.

$$\mathbf{A}^{-1} = \frac{1}{|\mathbf{A}|} \cdot \begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{bmatrix} \tag{3.65}$$

şeklindedir. Kofaktörler yardımı ile **A** matrisinin determinantı bir tek sıra veya sütuna göre,

$$|\mathbf{A}| = a_{11}c_{11} + a_{12}c_{12} + a_{13}c_{13} \tag{3.66}$$

şeklinde elde edilebilir. Bu tanımdan da anlaşılacağı gibi $|\mathbf{A}|$ herhangi bir sıra veya sütunun elemanlarına göre elde edilebilir. nxn boyutlu genel durum için determinant,

$$|\mathbf{A}| = a_{i1}c_{i1} + a_{i2}c_{i2} + \cdots + a_{in}c_{in} \qquad i = 1,......,n \tag{3.67}$$

veya

$$|\mathbf{A}| = a_{1j}c_{1j} + a_{2j}c_{2j} + \cdots + a_{nj}c_{nj} \qquad j = 1,......,n \tag{3.68}$$

eşitliklerinden bulunabilir. Bu ifadedeki kofaktörler n–1 boyutlu alt matrislerdir. Ters matris ise,

$$\mathbf{A}^{-1} = \frac{1}{|\mathbf{A}|} \cdot \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & & \vdots \\ c_{n1} & c_{n2} & \cdots & c_{nn} \end{bmatrix} \tag{3.69}$$

şeklinde elde edilebilir.

3.10 DETERMİNANTLAR VE TERS MATRİSLER ÜZERİNE TEOREMLER

Teorem 3.17: **A** matrisi simetrik ise **A** matrisinin determinantı,

$$|\mathbf{A}| = |\mathbf{A}^T| \tag{3.70}$$

transpozunun determinantına eşittir.

Teorem 3.18: **A** matrisinin herhangi iki sırasının veya sütununun değiştirilmesi ile bir **B** matrisi elde edilirse, **B** matrisinin determinantı **A** matrisinin determinantının ters işaretlisine eşittir.

$$|\mathbf{B}| = -|\mathbf{A}| \tag{3.71}$$

Teorem 3.19: Aynı elemandan oluşan iki veya daha fazla sıraya (veya sütuna) sahip matrisin determinantı sıfırdır.

$$\mathbf{A} = \begin{bmatrix} a & b \\ a & b \end{bmatrix}$$

$$\left|\mathbf{A}\right| = ab - ab = 0 \tag{3.72}$$

Teorem 3.20: **A** matrisinin herhangi bir sırasının (veya sütununun) herhangi bir değerle çarpılıp bir başka sırası (veya sütunu) ile toplanması ile elde edilen **B** matrisinin determinantı **A** matrisinin determinantına eşittir.

Teorem 3.21: Eğer **A** matrisinin sıraları (veya sütunları) doğrusal bağımlı ise $\left|\mathbf{A}\right| = 0$, doğrusal bağımsız ise $\left|\mathbf{A}\right| \neq 0$'dır. Buna uygun olarak tekil olmayan matrisler sıfırdan farklı, tekil matrisler ise sıfır determinanta sahiptir.

Teorem 3.22: Bir üçgen matrisin determinantı köşegen elemanlarının çarpımına eşittir.

$$\left|\mathbf{A}\right| = a_{11} \cdot a_{22} \cdot a_{33} \cdot \ldots\ldots \cdot a_{nn} \tag{3.73}$$

Alt üçgen matris,

$$\mathbf{A} = \begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

ve üst üçgen matris,

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{bmatrix}$$

şeklinde tanımlanır.

Teorem 3.23: Bir köşegen matrisin determinantı köşegen elemanlarının çarpımına eşittir.

$$\left|\mathbf{A}\right| = a_{11} \cdot a_{22} \cdot \ldots\ldots \cdot a_{nn}$$

Teorem 3.24: Birim matrisin determinantı bire eşittir.

Teorem 3.25: Bir **A** matrisinin herhangi bir sıra (veya sütunu) bir λ sabiti ile çarpılarak **B** matrisi elde edilirse, elde edilen matrisin determinantı orijinal matrisin determinantının λ katıdır.

$$\left|\mathbf{B}\right| = \lambda\left|\mathbf{A}\right|$$

Teorem 3.26: Eğer **A** matrisinin tüm elemanları λ ile çarpılarak **B** matrisi elde edilirse **B** matrisinin determinantı,

$$|\mathbf{B}| = \lambda^n |\mathbf{A}|$$

şeklinde elde edilir.

Teorem 3.27: İki kare matrisin çarpımının determinantı bu matrislerin determinantlarının çarpımına eşittir.

$$|\mathbf{AB}| = |\mathbf{A}| \cdot |\mathbf{B}| \tag{3.74}$$

Teorem 3.28: Eğer **C** bir ortogonal matris ise determinantı $|\mathbf{C}| = 1$ veya $|\mathbf{C}| = -1$ şeklindedir.

Teorem 3.29: Eğer **C** bir ortogonal matris ve **A** herhangi bir matris ise,

$$|\mathbf{C}^T \mathbf{A}\mathbf{C}| = |\mathbf{A}| \tag{3.75}$$

elde edilir.

Teorem 3.30: **A** matrisi,

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{pmatrix}$$

şeklinde alt matrislere ayrılmış ve $\mathbf{A}_{11}$ ile $\mathbf{A}_{22}$ kare matrislerdir. Eğer $\mathbf{A}_{12}$=0 veya $\mathbf{A}_{21}$=0 ise,

$$|\mathbf{A}| = |\mathbf{A}_{11}||\mathbf{A}_{22}| \tag{3.76}$$

olarak elde edilir. $\mathbf{A}_{22}$ matrisinin birim matris olması durumunda,

$$\mathbf{A} = \begin{vmatrix} \mathbf{A}_{11} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{vmatrix} = |\mathbf{A}_{11}|$$

bulunabilir. Bu sonuç kullanılarak

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_{12} \end{bmatrix} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_{22} \end{bmatrix}$$
$$\mathbf{A} = |\mathbf{A}_{11}||\mathbf{A}_{22}|$$

elde edilerek eşitlik (3.76)'ün ispatı yapılabilir. Daha sonra,

$$\mathbf{A} = \begin{vmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{0} & \mathbf{I} \end{vmatrix} = |\mathbf{A}_{11}|$$

ifadesi verilebilir. Üçgen (alt veya üst) matrisin determinantı

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{0} & \mathbf{A}_{22} \end{bmatrix} = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{0} & \mathbf{I} \end{bmatrix}$$
$$\mathbf{A} = |\mathbf{A}_{11}||\mathbf{A}_{22}|$$

şeklinde bulunabilir. Parçalanmış matrislerin determinantı için genel tanım aşağıdaki şekilde verilebilir. $\mathbf{A}_{11}$ ve $\mathbf{A}_{22}$ kare ve tekil olmayan matrisler olmak üzere,

$$\mathbf{B}_1 = \begin{bmatrix} \mathbf{I} & -\mathbf{A}_{12}\mathbf{A}_{22}^{-1} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \quad \text{ve} \quad \mathbf{B}_2 = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ -\mathbf{A}_{22}^{-1}\mathbf{A}_{21} & \mathbf{I} \end{bmatrix}$$

şeklinde $\mathbf{B}_1$ ve $\mathbf{B}_2$ matrisleri tanımlanarak,

$$\mathbf{B}_1\mathbf{A}\mathbf{B}_2 = \begin{bmatrix} \mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21} & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_{22} \end{bmatrix}$$

çarpımı elde edilir. $|\mathbf{B}_1| = |\mathbf{B}_2| = 1$ olduğu için,

$$|\mathbf{A}| = |\mathbf{A}_{22}| \cdot |\mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21}| \tag{3.77}$$

veya benzer bir şekilde alternatif olarak,

$$|\mathbf{A}| = |\mathbf{A}_{11}| \cdot |\mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12}| \tag{3.78}$$

şeklinde elde edilebilir.

Teorem 3.31: Eğer **A**, **B**, **C** matrisleri tekil olmayan matrisler ise bu matrislerin çarpımlarının tersi,

$$(\mathbf{ABC})^{-1} = \mathbf{C}^{-1}\mathbf{B}^{-1}\mathbf{A}^{-1} \tag{3.79}$$

şeklinde elde edilir.

Teorem 3.32: Ters matrisin tersi orijinal matrisi verir.

$$\left(\mathbf{A}^{-1}\right)^{-1} = \mathbf{A}$$

Teorem 3.33: Bir matrisin transpozunun tersi, tersinin transpozuna eşittir.

$$\left(\mathbf{A}^{T}\right)^{-1} = \left(\mathbf{A}^{-1}\right)^{T} \tag{3.80}$$

Teorem 3.34: $\mathbf{A}^{-1}$ matrisinin determinantı **A** matrisinin determinantının devriğidir.

$$\left|\mathbf{A}^{-1}\right| = \frac{1}{|\mathbf{A}|} \tag{3.81}$$

Teorem 3.35: Bir üst (veya alt) üçgen matrisin terside bir üst (veya alt) üçgen matristir.

Teorem 3.36: Eğer **A** matrisi,

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_{22} \end{bmatrix}$$

şeklinde parçalanmış ve tekil değilse (**A**, $\mathbf{A}_{11}$, $\mathbf{A}_{22}$ matrisleri kare matrislerdir) **A** matrisinin tersi,

$$\mathbf{A}^{-1} = \begin{bmatrix} \mathbf{A}_{11}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_{22}^{-1} \end{bmatrix}$$

şeklinde elde edilir.

Teorem 3.37: **A** matrisi,

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix}$$

şeklinde parçalanmış, $\mathbf{A}_{11}$ ve $\mathbf{A}_{22}$ tekil olmayan kare matrisler ise **A** matrisinin tersi,

$$\mathbf{A}^{-1}=\begin{bmatrix}\mathbf{B}_{11} & -\mathbf{B}_{11}\mathbf{A}_{12}\mathbf{A}_{22}^{-1}\\ -\mathbf{A}_{22}^{-1}\mathbf{A}_{21}\mathbf{B}_{11} & \mathbf{A}_{22}^{-1}+\mathbf{A}_{22}^{-1}\mathbf{A}_{21}\mathbf{B}_{11}\mathbf{A}_{12}\mathbf{A}_{22}^{-1}\end{bmatrix} \tag{3.82}$$

burada $\mathbf{B}_{11}=\left(\mathbf{A}_{11}-\mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21}\right)^{-1}$ 'dir ya da

$$\mathbf{A}^{-1}=\begin{bmatrix}\mathbf{A}_{11}^{-1}+\mathbf{A}_{11}^{-1}\mathbf{A}_{12}\mathbf{B}_{22}\mathbf{A}_{21}\mathbf{A}_{11}^{-1} & -\mathbf{A}_{11}^{-1}\mathbf{A}_{12}\mathbf{B}_{22}\\ -\mathbf{B}_{22}\mathbf{A}_{21}\mathbf{A}_{11}^{-1} & \mathbf{B}_{22}\end{bmatrix} \tag{3.83a}$$

şekline elde edilir, burada $\mathbf{B}_{22}=\left(\mathbf{A}_{22}-\mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12}\right)^{-1}$ 'dir. Bu teoremin ispatı için, bkz Alıştırma 3.5.

Bir matrisin tersinin alınması üzerine tanımlanmış önemli teoremlerden biri de *Sherman-Morrison-Woodbury* teoremidir. Bu teoremin sonuçları Bölüm 7 de sunulan PRESS istatistiği ve Bölüm 11 de tanımlanan etkili bir veri noktasının tanı istatistikleri için temel oluşturur. Açıklana teorem *i*-inci veri noktasının veri seti dışında bırakıldığı durumlar için bazı önekli regresyon istatistiklerinin kolay bir şekilde hesaplanmasını sağlamaktadır. Boyutu $p\times p$ olan ve tekil olmayan bir **A** matrisi ve p boyutlu bir sütun vektörü **a** ele alınsın. Regresyon uygulaması için $\mathbf{A}=\mathbf{X}^T\mathbf{X}$ ve $\mathbf{a}^T=\mathbf{x}_i^T$ alınır. Burada $\mathbf{x}_i^T=\begin{pmatrix}1 & x_{i1} & \dots & x_{ik}\end{pmatrix}$ vektörü **X** matrisinin *i*-inci sırasıdır. Bu durumda ($\mathbf{A}-\mathbf{a}\mathbf{a}^T$) matrisi $\mathbf{X}^T\mathbf{X}$ matrisinin *i*-inci sırasının çıkarıldığı matrisi temsil eder. Bu matrisin tersi ise aşağıdaki teorem ile elde edilir.

Teorem 3.38: $$\left(\mathbf{A}-\mathbf{a}\mathbf{a}^T\right)^{-1}=\mathbf{A}^{-1}+\frac{\mathbf{A}^{-1}\mathbf{a}\mathbf{a}^T\mathbf{A}^{-1}}{1-\mathbf{a}^T\mathbf{A}^{-1}\mathbf{a}} \tag{3.83b}$$

3.11 BİR MATRİSİN ÖZDEĞERLERİ (KARAKTERİSTİK KÖKLERİ) VE TEOREMLER

Önceki kısımlarda, eşitlik (3.41) ile tanımlanan **Ba**=**c** şeklindeki denklem setlerinin çözümü ile ilgilenildi. Bu kısımda ise,

$$\mathbf{A}\mathbf{x}=\lambda\mathbf{x} \tag{3.84}$$

setinin çözümü ile ilgilenilecektir. Bu eşitlikte **A** elemanları bilinen n boyutlu bir kare matris, **x** elemanları bilinmeyen n elemanlı bir sütun vektörü ve λ bilinmeyen bir skalerdir. Bu problem bir özdeğer problemi olarak bilinir. Eşitlik (3.41)'de bir bilinmeyen vektörü mevcuttu, fakat eşitlik (3.84)'de ise bilinmeyen vektörünün yanısıra bir de bilinmeyen skaler mevcuttur. Her bir λ 'ya bir **x** vektörü karşılık gelir. Bu λ değerleri, özdeğerler ya da karakteristik kök olarak, **x** vektörleri ise öz vektörler ya da karakteristik vektörler olarak bilinirler. n=2 için eşitlik (3.84) homojen denklem seti olarak,

$$(a_{11}-\lambda)x_1+a_{12}x_2=0$$

$$a_{21}x_1+(a_{22}-\lambda)x_2=0$$

şeklinde yazılabilir. Bu denklem seti matris notasyonunda,

$$(\mathbf{A}-\lambda\,\mathbf{I})\mathbf{x}=\mathbf{0} \tag{3.85}$$

eşitliği ile verilebilir. Eşitlik (3.85) verilen herhangi bir n için eşitlik (3.84)'ye eşittir. Eğer ($\mathbf{A}$-$\lambda\mathbf{I}$) matrisi tekil olmayan bir matris ise eşitlik (3.85) için sadece $\mathbf{x}$=$\mathbf{0}$ sonucu elde edilir. Sıfırdan farklı bir sonucun mevcut olabilmesi için matris tekil olmak zorundadır. Başka bir deyişle, determinantın sıfır olması gerekir. Bu şarta göre,

$$|\mathbf{A} - \lambda\,\mathbf{I}| = 0 \tag{3.86}$$

eşitliği elde edilir. Bu eşitlik $\mathbf{A}$ matrisinin karakteristik denklemi olarak bilinir. Bu denklem, bilinmeyen λ'ya göre bir polinom denklemi verir. Her bir kök veya özdeğer λ_i eşitlik (3.85)'de yerine konarak karşılık gelen $\mathbf{x}_i$ vektörü elde edilebilir. 2×2 durumu için karakteristik denklem,

$$\lambda^2 - (a_{11} + a_{22})\lambda + (a_{11}a_{22} - a_{12}a_{21}) = 0 \tag{3.87}$$

şeklinde olup kökleri,

$$\lambda_1 = \frac{1}{2}\left[(a_{11} + a_{22}) + \sqrt{(a_{11} + a_{22})^2 - 4(a_{11}a_{22} - a_{12}a_{21})}\right]$$

$$\lambda_2 = \frac{1}{2}\left[(a_{11} + a_{22}) - \sqrt{(a_{11} + a_{22})^2 - 4(a_{11}a_{22} - a_{12}a_{21})}\right]$$

denklemlerinden elde edilir. Matris simetrik ise, a_{12}=a_{21} olduğu için kökler,

$$\lambda = \frac{1}{2}\left[(a_{11} + a_{22}) \pm \sqrt{(a_{11} - a_{22})^2 + 4a_{12}^2}\right]$$

olarak elde edilir. Gerçel bir simetrik matris için köklerin gerçel olması gereklidir. Karakteristik denklem,

$$(\lambda_1 - \lambda)(\lambda_2 - \lambda) = \lambda^2 - (\lambda_1 + \lambda_2)\lambda + \lambda_1\lambda_2 = 0$$

şeklinde yazılabilir. Bu denklem, eşitlik (3.87) ile karşılaştırıldığında,

$$\textit{Kökler toplamı} = \lambda_1 + \lambda_2 = a_{11} + a_{22} \tag{3.88}$$

= $\mathbf{A}$'nın köşegen elemanlarının toplamı (matrisin izi)

$$\textit{Kökler çarpımı} = \lambda_1\lambda_2 = a_{11}a_{22} - a_{12}a_{21} = |\mathbf{A}| \tag{3.89}$$

sonuçları elde edilir, bkz. Alıştırma 3.6.

Teorem 3.39: Özdeğerler gerçeldir, bkz. Alıştırma 3.7.

Teorem 3.40: Farklı özdeğerlere karşılık gelen özvektörler ortogonaldir, bkz. Alıştırma 3.8.

Teorem 3.41: Eğer bir λ değeri k defa tekrarlanıyor ise bu özdeğere karşılık gelen k adet ortogonal vektör mevcut olabilecektir.

Teorem 3.42: n boyutlu simetrik $\mathbf{A}$ matrisinin sahip olduğu n adet özdeğerin λ_1, … ,λ_n hepsinin birbirinden farklı olması gerekli değildir. Teorem 39 ve 40 n adet ortogonal özvektörün $\mathbf{x}_1$, … ,$\mathbf{x}_n$ mevcut olduğunu belirtmektedir.

Teorem 3.43: Eğer $\mathbf{X}$ ortogonal bir matris ise $\mathbf{A}$ ile $\mathbf{XAX}^{-1}$matrislerinin özdeğerleri özdeştir.

Teorem 3.44: Her simetrik matris **A** için, köşegen elemanları **A**'nın özdeğerleri olan bir **D** köşegen matris ve bir **X** ortogonal matrisi,

$$\mathbf{X}^T\mathbf{AX} = \mathbf{D} \tag{3.90}$$

mevcuttur.

Teorem 3.45: Bir **A** matrisinin köşegen elemanlarının toplamı (izi) özdeğerlerinin toplamına eşittir, bkz. Alıştırma 3.9.

Teorem 3.46: **A** matrisinin özdeğerlerinin çarpımı bu matrisin determinantına eşittir, bkz. Alıştırma 3.10.

Teorem 3.47: **A** matrisinin rankı, bu matrisin sıfırdan farklı özdeğerlerinin sayısına eşittir.

Teorem 3.48: $\mathbf{A}^2$ matrisinin özdeğerleri, **A** matrisinin özdeğerlerinin karelerine eşittir. Fakat her iki matrisinde özvektörleri aynıdır.

$$\mathbf{AX} = \lambda\mathbf{X}$$

bu ifadeyi soldan **A** ile çarparak

$$\mathbf{A}^2\mathbf{X} = \lambda\mathbf{AX} = \lambda^2\mathbf{X}$$

ispat tamamlanır. Bu sonuç dinamik sistemlerin durağanlığının analizinde çok faydalı bir uygulamaya sahiptir. $\mathbf{y}_t$, bir değişkenin t periyodunda aldığı değerleri belirten vektör olsun. Bu değişkenin denklem sisteminin önceki değerlerine göre,

$$\mathbf{y}_t = \mathbf{A}\mathbf{y}_{t-1} \tag{3.91}$$

şeklinde ifade edilebileceği kabul edilsin. Sistemin orjinali birden fazla gecikmeyi içerse bile, yeni değişkenlerin uygun bir tanımı eşitlik (3.91) şeklinde elde edilmiş bir sistem ile oluşturulabilir. Eşitlik (3.91)'da birbirini takip eden değerler yerine konarak,

$$\mathbf{y}_t = \mathbf{A}^t\mathbf{y}_0$$

elde edilebilir. Bu ifade de $\mathbf{y}_0$ değişkenin başlangıç değerini belirtir. Elde edilmiş olan **A** matrisi özvektörlerin doğrusal bağımsız bir setine sahiptir,

$$\mathbf{X}^{-1}\mathbf{AX} = \mathbf{D}$$

$$\mathbf{A} = \mathbf{XDX}^{-1}$$

ve

$$\mathbf{A}^2 = \mathbf{X}\boldsymbol{D}\mathbf{X}^{-1}\mathbf{X}\boldsymbol{D}\mathbf{X}^{-1} = \mathbf{X}\boldsymbol{D}^2\mathbf{X}^{-1}$$

sonuç olarak

$$\mathbf{A}^t = \mathbf{XD}^t\mathbf{X}^{-1}$$

elde edilir. $\mathbf{y}_t$ vektörünün elemanlarının, **A** matrisinin özdeğerlerinin t-inci kuvvetinin doğrusal kombinasyonu olduğu görülebilir. Bu sonuca göre sistemin durağan olabilmesi için

$$|\lambda_i| \le 1 \qquad\qquad i = 1,\ldots,n$$

olması gereklidir.

Teorem 3.49: $\mathbf{A}^{-1}$ matrisinin özdeğerleri **A** matrisinin özdeğerlerinin evriğine eşittir. Fakat her iki matrisin özvektörleri aynıdır.

$$\mathbf{Ax}=\lambda\mathbf{x}$$

bu eşitlik soldan $\mathbf{A}^{-1}$ ile çarpılarak

$$\mathbf{x}=\lambda\, \mathbf{A}^{-1}\mathbf{x}$$

veya

$$\mathbf{A}^{-1}\mathbf{x} =(1/\lambda)\mathbf{x}$$

elde edilir.

Teorem 3.50: Bir idempotent matrisin her bir özdeğeri sıfır ya da birdir.

Teorem 3.47 kullanılarak $\mathbf{A}^2\mathbf{x}= \lambda^2\mathbf{x}$ ve **A** idempotent bir matris ise,

$$\mathbf{A}^2\mathbf{x} =\mathbf{Ax}= \lambda\mathbf{x}$$

$$\lambda(\lambda-1)\mathbf{x}=0$$

ve herhangi bir **x** özvektörü sıfır vektörü olmadığından, $\lambda=0$ ya da $\lambda=1$ bulunur.

3.12 KARESEL FORMLAR VE POZİTİF TANIMLI MATRİSLER

A matrisi 2×2 boyutlu simetrik bir matris ve **x** iki elemanlı bir sütun vektörü ise karesel form,

$$\mathbf{x}^T\mathbf{Ax} = a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2$$

şeklindedir. n boyutlu genel durum ise ,

$$\begin{aligned}\mathbf{x}^T\mathbf{Ax} = a_{11}x_1^2 &+ 2a_{12}x_1x_2 + 2a_{13}x_1x_3 + \cdots + 2a_{1n}x_1x_n \\ &+ a_{22}x_2^2 + 2a_{23}x_2x_3 + \cdots + 2a_{2n}x_2x_n \\ &\qquad + a_{33}x_3^2 + \cdots + 2a_{3n}x_3x_n \\ &\qquad\qquad \ddots \\ &\qquad\qquad\quad \ddots + a_{nn}x_n^2\end{aligned}$$

ifadesi ile verilebilir.

Tüm $\mathbf{x}\neq\mathbf{0}$ sütun vektörleri için eğer $\mathbf{x}^T\mathbf{Ax}>0$ ise karesel form ve **A** matrisi pozitif tanımlıdır. Eğer tüm $\mathbf{x}\neq\mathbf{0}$ sütun vektörleri için $\mathbf{x}^T\mathbf{Ax}\geq 0$ ise karesel form ve matris pozitif yarı tanımlıdır. Yukarıdaki eşitsizliklerin yönü değiştirilerek negatif tanımlı ve negatif yarı tanımlı karesel form ve matrisler tanımlanabilir. Eğer bir form bazı **x** vektörleri için pozitif, diğerleri için negatif ise tanımsızdır.

Teorem 3.51: Gerçel simetrik matris **A**'nın pozitif tanımlı olabilmesi için gerek ve yeter koşul **A** matrisinin tüm özdeğerlerinin pozitif tanımlı olması gerekir. Pozitif yarı tanımlı bir matrisin özdeğerleri negatif değildir.

Teorem 3.52: Gerçel simetrik bir **A** matrisinin pozitif tanımlı olabilmesi için gerek ve yeter koşul **A** matrisinin her bir alt matrisinin determinantının pozitif olması gerekir.

$$a_{11}\rangle 0 \quad \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \rangle 0 \cdots \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} \rangle 0$$

$$|\mathbf{A}_1| \rangle 0 \quad |\mathbf{A}_2| \rangle 0 \cdots |\mathbf{A}_n| = |\mathbf{A}| \rangle 0 \tag{3.92}$$

Teorem 3.53: **A** matrisi simetrik ve pozitif tanımlı bir matris ise,

$$\mathbf{A}=\mathbf{P}\mathbf{P}^T \tag{3.93}$$

eşitliğini sağlayan tekil olmayan bir p matrisi mevcuttur.

A matrisinin özvektörlerinden oluşan **X** matrisi kullanılarak $\mathbf{X}^T\mathbf{A}\mathbf{X}=\mathbf{D}$ ve $\mathbf{A}=\mathbf{X}\mathbf{D}\mathbf{X}^T$ elde edilebilir. **A** matrisi pozitif tanımlı ise tüm özdeğerleri pozitiftir. Bu nedenle bir köşegen matris olan **D** matrisi

$$\mathbf{D} = \mathbf{D}^{1/2}\mathbf{D}^{1/2}$$

şeklinde yazılabilir. Burada $\mathbf{D}^{1/2}$ matrisi

$$\mathbf{D}^{1/2} = \begin{bmatrix} \sqrt{\lambda_1} & 0 & \cdots & 0 \\ 0 & \sqrt{\lambda_2} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \sqrt{\lambda_n} \end{bmatrix}$$

şeklindedir. Buna uygun olarak,

$$\mathbf{A} = \mathbf{X}\mathbf{D}^{1/2}\mathbf{D}^{1/2}\mathbf{X}^T = \left(\mathbf{X}\mathbf{D}^{1/2}\right)\left(\mathbf{X}\mathbf{D}^{1/2}\right)^T$$

elde edilir. Sonuç olarak,

$$\mathbf{P} = \mathbf{X}\mathbf{D}^{1/2}$$

olduğu görülebilir. **P** matrisi tekil olmayan matrislerin çarpımı olduğundan tekil olmayan bir matristir.

Teorem 3.54: **A** matrisi $n\times n$ boyutlu ve pozitif tanımlı ve **P**, $n\times m$ boyutlu rankı m olan bir matris ise $\mathbf{P}^T\mathbf{A}\mathbf{P}$ pozitif tanımlıdır.

Teorem 3.55: **A** matrisi $n\times m$ boyutlu rankı m olan bir matris ise $\mathbf{A}^T\mathbf{A}$ pozitif, $\mathbf{A}\mathbf{A}^T$ pozitif yarı tanımlıdır.

Teorem 3.56: **A** matrisi $n\times m$ boyutlu rankı k $(k\langle m, k\langle n)$ olan bir matris ise $\mathbf{A}^T\mathbf{A}$ ve $\mathbf{A}\mathbf{A}^T$ çarpımlarının her ikiside pozitif yarı tanımlıdır.

Teorem 3.57: **A** ve **B** pozitif tanımlı matrislerse **A-B** ve $\mathbf{B}^{-1}$-$\mathbf{A}^{-1}$ matrisleri de pozitif tanımlıdır.

Teorem 3.58: Pozitif tanımlı bir matrisin determinantı pozitiftir.

Teorem 3.59: A_1 ve A_2 matrisleri simetrik, A_2 pozitif tanımlı matris ve eğer A_1 - A_2 matrisi pozitif yarı tanımlı (veya pozitif tanımlı) ise $|A_1| \geq |A_2|$ 'dir.

3.13 MAKSİMUM – MİNİMUM DEĞERLER ve JAKOBYEN

Maksimum ve minimum değerler için bazı önemli sonuçları matris notasyonunda vermek yararlı olacaktır. Bir skaler y değişkeni, n bağımsız değişkenin fonksiyonu olarak,

$$y = f\left(x_1, x_2, \cdots, x_n\right) = f\left(\mathbf{x}\right)$$

tanımlanabilir. Bu fonksiyonun toplam diferansiyeli,

$$dy = f_1 dx_1 + f_2 dx_2 + \cdots + f_n dx_n \qquad (3.94)$$

olarak verilir. Burada

$$f_i = \frac{\partial y}{\partial x_i} \qquad i = 1, \ldots, n$$

şeklindedir. dx_i, x_i'deki herhangi bir değişikliği belirtir. Küçük bir dx_i için birinci dereceden diferansiyel y değişkeninde oluşan değişimin yaklaşık bir değerini verir. Kısmi türev vektörü **f** ve diferansiyel vektörü **dx**,

$$\mathbf{f} = \frac{\partial y}{\partial \mathbf{x}} = \begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ f_n \end{bmatrix} \qquad \mathbf{dx} = \begin{bmatrix} dx_1 \\ dx_2 \\ \vdots \\ dx_n \end{bmatrix}$$

ile belirtilerek y'nin birinci dereceden diferansiyeli

$$dy = \mathbf{f}^T \mathbf{dx}$$

ifadesi ile verilebilir. Eğer y değişkeni,

$$\mathbf{x}^* = \begin{bmatrix} x_1^* & x_2^* & \cdots & x_n^* \end{bmatrix}$$

noktasında durağan bir değere sahipse $\mathbf{x}^*$'ın yakınındaki tüm noktalar için dy=0 olacaktır. Bu noktalarda **dx≠0** olacağı için eşitlik (3.94) dikkate alındığında bir durağan değer için gerekli şartın, **f**=**0** olduğu görülmektedir. Başka bir deyişle durağan noktada tüm kısmi türevler sıfırdır.

Bir durağan nokta bir maksimum nokta olabilir. Bu durumda $\mathbf{x}^*$ çevresindeki tüm noktalarda fonksiyonun değeri azalır. İkinci bir durum ise durağan noktanın bir minimum noktası olmasıdır ki bu durumda $\mathbf{x}^*$ çevresindeki tüm noktalarda fonksiyonun değeri artar. Sonlanarak bir eğer noktası olabilir. Bu durumda $\mathbf{x}^*$'dan uzaklaşılan bazı yönlerde fonksiyon azalır diğer yönlerde ise artar. Ortaya çıkabilecek bu muhtemel durumlar arasındaki farkı ayırt edebilmek için ikinci dereceden diferansiyel d^2y kullanılır. İkinci dereceden diferansiyel ile toplam diferansiyelin birinci dereceden diferansiyeli bulunabilir. Bu ifade de $\mathbf{x}^*$ noktasından uzaklaşmanın dy değerinde oluşturduğu değişim için yaklaşım elde edilmesini sağlar. Bir maksimum değer için dy sıfırdan bir negatif değere doğru azalacaktır, bu nedenle d^2y negatif bir değer alabilecektir. Buna uygun olarak bir minimum değer için d^2y pozitif değer alabilir. Eğer noktası için d^2y bazı **dx** değerleri için pozitif diğerleri için negatif değer alır. Eşitlik (3.94) toplam diferansiyeli,

$$
\begin{aligned}
d^2y = {} & \frac{\partial}{\partial x_1}\left[f_1 dx_1 + f_2 dx_2 + \cdots + f_n dx_n\right]dx_1 \\
& + \frac{\partial}{\partial x_2}\left[f_1 dx_1 + f_2 dx_2 + \cdots + f_n dx_n\right]dx_2 \\
& + \cdots + \frac{\partial}{\partial x_n}\left[f_1 dx_1 + f_2 dx_2 + \cdots + f_n dx_n\right]dx_n \\
= {} & f_{11} dx_1^2 + 2 f_{12} dx_1 dx_2 + 2 f_{13} dx_1 dx_3 + \cdots + 2 f_{1n} dx_1 dx_n \\
& + f_{22} dx_2^2 + 2 f_{23} dx_2 dx_3 + \cdots + 2 f_{2n} dx_2 dx_n \\
& \ddots \\
& \qquad \ddots \\
& + f_{nn} dx_n^2
\end{aligned} \tag{3.95}
$$

şeklinde elde edilir. Burada,

$$
f_{ij} = f_{ji} = \frac{\partial^2 y}{\partial x_i \partial x_j}
$$

ifadesi ile verilir. dx_i^2 ise dx_i diferansiyelinin karesidir. Eşitlik (3.95)'den görülebileceği gibi ikinci dereceden diferansiyel dx'in karesel bir formu olarak yazılabilir. Karesel form matrisi ikinci dereceden kısmi türevlerin simetrik Hessien matrisidir.

$$
F = \frac{\partial^2 y}{\partial x^2} = \begin{bmatrix} f_{11} & f_{12} & \cdots & f_{1n} \\ f_{21} & f_{22} & \cdots & f_{2n} \\ \vdots & \vdots & & \vdots \\ f_{n1} & f_{n2} & \cdots & f_{nn} \end{bmatrix}
$$

sonuç olarak,

$$
d^2y = \mathbf{dx}^T \mathbf{F} \mathbf{dx} \tag{3.96}
$$

eşitliği yazılabilir. Bu ifadeye uygun olarak **F** pozitif tanımlı ise d^2y pozitif, negatif tanımlı ise negatiftir. Bir $\mathbf{x}^*$ noktasında maksimum veya minimum mevcut olması için şartlar aşağıda özetlenmiştir.

	Birinci derece şartı	İkinci derece şartı
Maksimum	$\mathbf{f} = \frac{\partial y}{\partial \mathbf{x}} = \mathbf{0}$	$\mathbf{F} = \frac{\partial^2 y}{\partial \mathbf{x}^2}$ negatif tanımlı
Minimum	$\mathbf{f} = \frac{\partial y}{\partial \mathbf{x}} = \mathbf{0}$	$\mathbf{F} = \frac{\partial^2 y}{\partial \mathbf{x}^2}$ pozitif tanımlı

Kısıtlı Maksimum veya Minimum; $y=f(x_1, \ldots ,x_n)$ fonksiyonun durağan değerleri araştırılırken x'lerin bağımsız değişkenler olduğu varsayılmıştı. Bu nedenle n adet keyfi diferansiyel dx_1, dx_2 ,....,

dx_n belirlenebilmekteydi. Bununla birlikte bazı problemlerde x'ler bir veya daha kısıta sahip olabilirler ve bu durumda y'nin bir maksimum veya minimumu bu kısıtlar dikkate alınarak bulunmak zorundadır. Fonksiyonun bir tek maksimum veya minimuma sahip olduğu varsayımı ile problem aşağıdaki şekilde ifade edilebilir:

Verilen m adet ($m<n$) kısıt altında y'yi maksimum (minimum) yapan $\mathbf{x}^*$ vektörünü bulun.

$$y_j(\mathbf{x}) = 0 \qquad j = 1, \cdots, m$$

Kısıt sütun vektörü,

$$\mathbf{y}(\mathbf{x}) = \begin{bmatrix} y_1(\mathbf{x}) \\ y_2(\mathbf{x}) \\ \vdots \\ y_m(\mathbf{x}) \end{bmatrix}$$

ve Lagrange çarpanları sütun vektörü,

$$\boldsymbol{\lambda} = \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_m \end{bmatrix}$$

Bu veriler kullanılarak yeni bir amaç fonksiyonu,

$$\psi = f(\mathbf{x}) - \boldsymbol{\lambda}^T \mathbf{y}(\mathbf{x}) \tag{3.97}$$

tanımlanır. ψ değeri $\boldsymbol{\lambda}$ ve $\mathbf{x}$'lerden oluşan $m \times n$ adet değişkenin oluşturduğu bir skalerdir. ψ'nin durağan bir değeri için birinci derece koşulu tüm $m \times n$ adet birinci dereceden kısmi türevin,

$$\begin{bmatrix} \dfrac{\partial \psi}{\partial \mathbf{x}} \\ \dfrac{\partial \psi}{\partial \boldsymbol{\lambda}} \end{bmatrix} = \begin{bmatrix} \dfrac{\partial f}{\partial \mathbf{x}} - \dfrac{\partial\left(\boldsymbol{\lambda}^T(\mathbf{x})\right)}{\partial \mathbf{x}} \\ \mathbf{y}(\mathbf{x}) \end{bmatrix} = \mathbf{0} \tag{3.98}$$

sıfır olmasıdır.

$$\boldsymbol{\lambda}^T \mathbf{y}(\mathbf{x}) = \lambda_1 \mathrm{y}_1(\mathrm{x}_1, \mathrm{x}_2, \cdots, \mathrm{x}_n) + \lambda_2 \mathrm{y}_2(\mathrm{x}_1, \mathrm{x}_2, \cdots, \mathrm{x}_n) + \cdots + \lambda_m \mathrm{y}_m(\mathrm{x}_1, \mathrm{x}_2, \cdots, \mathrm{x}_n)$$

olup kısmi türevleri,

$$\frac{\partial\left(\boldsymbol{\lambda}^T \mathbf{y}(\mathbf{x})\right)}{\partial \mathrm{x}_1} = \lambda_1 \frac{\partial \mathrm{y}_1}{\partial \mathrm{x}_1} + \lambda_2 \frac{\partial \mathrm{y}_2}{\partial \mathrm{x}_1} + \cdots + \lambda_m \frac{\partial \mathrm{y}_m}{\partial \mathrm{x}_1} = \lambda' \frac{\partial \mathrm{y}}{\partial \mathrm{x}_1}$$

....

$$\frac{\partial\left(\boldsymbol{\lambda}^T \mathbf{y}(\mathbf{x})\right)}{\partial \mathrm{x}_n} = \lambda_1 \frac{\partial \mathrm{y}_1}{\partial \mathrm{x}_n} + \lambda_2 \frac{\partial \mathrm{y}_2}{\partial \mathrm{x}_n} + \cdots + \lambda_m \frac{\partial \mathrm{y}_m}{\partial \mathrm{x}_n} = \lambda' \frac{\partial \mathrm{y}}{\partial \mathrm{x}_n}$$

şeklindedir. Bu ifadede,

$$\frac{\partial \mathbf{y}}{\partial \mathrm{x}_i} = \begin{bmatrix} \dfrac{\partial y_1}{\partial x_i} \\ \dfrac{\partial y_2}{\partial x_i} \\ \vdots \\ \dfrac{\partial y_m}{\partial x_i} \end{bmatrix} \qquad i = 1,2,\cdots,n$$

olarak verilmiştir. $\partial f/\partial \mathbf{x}$, $n\times 1$ boyutlu bir vektör olduğu için, $(\partial/\partial \mathbf{x})\left(\boldsymbol{\lambda}^T \mathbf{y}(\mathbf{x})\right)$ değerinin de n sıralı olacak şekilde düzenlenmesi gereklidir. Kısmi türevler $n\times m$ boyutlu **G** matrisi ile,

$$\mathbf{G} = \begin{bmatrix} \dfrac{\partial y_1}{\partial x_1} & \dfrac{\partial y_2}{\partial x_1} & \cdots & \dfrac{\partial y_m}{\partial x_1} \\ \dfrac{\partial y_1}{\partial x_2} & \dfrac{\partial y_2}{\partial x_2} & \cdots & \dfrac{\partial y_m}{\partial x_2} \\ \vdots & \vdots & & \vdots \\ \dfrac{\partial y_1}{\partial x_n} & \dfrac{\partial y_2}{\partial x_n} & \cdots & \dfrac{\partial y_m}{\partial x_n} \end{bmatrix} = \begin{bmatrix} \dfrac{\partial \mathbf{y}}{\partial x_1} & \dfrac{\partial \mathbf{y}}{\partial x_2} & \cdots & \dfrac{\partial \mathbf{y}}{\partial x_n} \end{bmatrix}$$

belirtilebilir. Bu durumda

$$\frac{\partial \psi}{\partial \mathbf{x}} = \frac{\partial f}{\partial \mathbf{x}} - \mathbf{G}\boldsymbol{\lambda}$$

yazılabilir ve durağan bir nokta için birinci derece koşul,

$$\frac{\partial f}{\partial \mathbf{x}} - \mathbf{G}\boldsymbol{\lambda} = \mathbf{0}$$

$$\mathbf{y}(\mathbf{x}) = \mathbf{0} \tag{3.99}$$

şeklindedir. Eşitlik (3.99)'deki ikinci denklem durağan noktanın kısıtları sağlamasını garanti eder. Maksimum ve minimumu ayırt edebilmek için eşitlik (3.96)'deki karesel formun pozitif mi yoksa negatif mi tanımlı olduğunun araştırılması gereklidir. Fakat bu çalışma sadece kısıtları ihlal etmeyen **dx** vektörleri için yapılır. j-inci kısıtın,

$$y_j(x_1, x_2, \cdots, x_n) = 0$$

toplam diferansiyeli,

$$0 = dy_j = \frac{\partial y_j}{\partial x_1} dx_1 + \frac{\partial y_j}{\partial x_2} dx_2 + \cdots + \frac{\partial y_j}{\partial x_n} dx_n$$

olarak verilebilir. Her bir kısıt için benzer bir şart mevcuttur. Bu eşitliği sağlayan **dx** vektörleri kısıtları ihlal etmez ve bu vektörler,

$$\mathbf{G}^T\mathbf{dx} = \mathbf{0} \tag{3.100}$$

eşitliği ile elde edilebilir. Pek çok durumda **F** matrisi sabitlerden oluşur, bu matrisin tanımsızlığı herhangi **x** değerlerinin bağımsızlığı ile oluşturulabilir.

3.14 ORTOGONAL TRANSFORMASYONLAR VE İZDÜŞÜMLER

Her ikisi de n boyutlu olan **x** vektörünün **y** vektörüne doğrusal transformasyonu **y**=**Ax** şeklinde yazılabilir. Burada **A**, $n \times n$ boyutlu transformasyon katsayılar matrisidir. Eğer **A** matrisi tekil değilse transformasyon birebirdir. Bu durumda **y** vektörünün **x** üzerine ters transformasyonu $\mathbf{x}=\mathbf{A}^{-1}\mathbf{y}$ şeklindedir.

Eğer $\mathbf{AA}^T=\mathbf{I}$ ise bu doğrusal transformasyon ortogonal transformasyondur. Buna uygun olarak **A** matrisinin sıraları ortogonaldir ve uzunlukları bir birimdir. Ortogonal transformasyonlar vektörler arasındaki uzaklık ve açıları değiştirmez. Başka bir deyişle vektörler arasındaki uzaysal ilişkiler ortogonal transformasyon ile değişmez. Örneğin $\mathbf{y}_1^{\mathrm{T}}=(3 \quad 10 \quad 20)$ ve $\mathbf{y}_2^{\mathrm{T}}=(6 \quad 14 \quad 21)$ vektörleri ile katsayı matrisi,

$$\mathbf{A}=\begin{bmatrix} 1 & 1 & 1 \\ -1 & 0 & 1 \\ -1 & 2 & -1 \end{bmatrix}$$

olarak verilmiş olsun doğrusal dönüşümler,

$$\mathbf{x}_1=\mathbf{A}\mathbf{y}_1=\begin{bmatrix} 1 & 1 & 1 \\ -1 & 0 & 1 \\ -1 & 2 & -1 \end{bmatrix}\begin{pmatrix} 3 \\ 10 \\ 20 \end{pmatrix}=\begin{pmatrix} 33 \\ 17 \\ -3 \end{pmatrix}$$

ve

$$\mathbf{x}_2=\mathbf{A}\mathbf{y}_2=\begin{bmatrix} 1 & 1 & 1 \\ -1 & 0 & 1 \\ -1 & 2 & -1 \end{bmatrix}\begin{pmatrix} 6 \\ 14 \\ 21 \end{pmatrix}=\begin{pmatrix} 41 \\ 15 \\ 1 \end{pmatrix}$$

olarak $\mathbf{y}_1$'in $\mathbf{x}_1$'e ve $\mathbf{y}_2$'nin $\mathbf{x}_2$'ye doğrusal transformasyonu elde edilir. Bunlar ortogonal transformasyon değildir. Çünkü

$$\mathbf{AA}^T=\begin{bmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 6 \end{bmatrix}\neq \mathbf{I}$$

çarpımı birim matrisi vermemektedir. **A** matrisinin sıraları ayrık ortogonaldir, köşegen dışı elemanlar sıfırdır, fakat uzunlukları bir birim değildir. Bu transformasyon her bir sıra kendi uzunluğuna bölünerek ortogonal hale dönüştürülebilir. Bu işlem ile sıraların uzunluğu bir birim olacak şekilde ölçeklenmiş olur. Elde edilen $\mathbf{A}^*$ matrisi ve ortogonal transformasyon,

$$\mathbf{x}_1^*=\mathbf{A}^*\mathbf{y}_1=\begin{bmatrix} 1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} \\ -1/\sqrt{2} & 0 & 1/\sqrt{2} \\ -1/\sqrt{6} & 2/\sqrt{6} & -1/\sqrt{6} \end{bmatrix}\mathbf{y}_1=\begin{pmatrix} 33/\sqrt{3} \\ 17/\sqrt{2} \\ -3/\sqrt{6} \end{pmatrix}$$

ve

$$\mathbf{x}_2^* = \mathbf{A}^*\mathbf{y}_2 = \begin{pmatrix} 41/\sqrt{3} \\ 15/\sqrt{2} \\ 1/\sqrt{6} \end{pmatrix}$$

olarak bulunur. İki vektör (**u** ve **v**) arasındaki uzaklığın karesi, $(\mathbf{u}-\mathbf{v})^T(\mathbf{u}-\mathbf{v})$ şeklindedir. Bu ifadeden faydalanarak orijinal **y** vektörleri arasındaki ile ortogonal transformasyon sonucu elde edilen **x** vektörleri arasındaki uzaklığın değişmediği ispatlanabilir:

$$(\mathbf{y}_1 - \mathbf{y}_2)^T(\mathbf{y}_1 - \mathbf{y}_2) = (\mathbf{x}_1^* - \mathbf{x}_2^*)'(\mathbf{x}_1^* - \mathbf{x}_2^*) = 26$$

Bir alt uzay içine bir vektörün izdüşümü transformasyonların özel bir durumudur. İzdüşüm EKK'da anahtar bir adımdır, bkz.Bölüm 4.

Teorem 3.60: Eğer C ortogonal bir matris ise $tr(\mathbf{C}^T\mathbf{AC}) = tr(\mathbf{A})$'dır.

Teorem 3.61: Eğer **A**, rankı r olan idempotent bir matris ise $\mathbf{P}^T\mathbf{AP} = \mathbf{E}_r$ olacak şekilde bir **P** ortogonal matrisi mevcuttur. Burada $\mathbf{E}_r$, r adet köşegen elemanı bir ve kalan elemanları sıfır olan bir köşegen matristir.

Teorem 3.62: Eğer **A** idempotent **P** ortogonal matris ise $\mathbf{P}^T\mathbf{AP}$ matrisi idempotentdir.

3.15 BİR DİKDÖRTGEN MATRİSİN TEKİL DEĞER AYRIŞIMI

Kısım 3.11'de verilen özdeğer analizi bir kare matrise uygulanmaktaydı. Bu kısımda özdeğer analizi, tekil değer ayrışımı adı verilen benzer bir ayrışımı geliştirmek için kullanılacaktır. Tekil değer ayrışımı daha sonra anabileşen analizini vermekte kullanılır.

X boyutları $n\times p$ olan ($n>p$) olan bir matris ise $\mathbf{X}^T\mathbf{X}$ $p\times p$ boyutlu bir simetrik kare matristir. Bu matrisin özdeğerleri λ ve özvektörlerin tanımladığı **Z** matrisi kullanılarak,

$$\mathbf{X}^T\mathbf{X} = \mathbf{ZDZ}^T \tag{3.101}$$

ifadesi verilebilir. Burada **D** köşegen matris olup elemanları $\mathbf{X}^T\mathbf{X}$ matrisinin özdeğerleridir. Benzer olarak $\mathbf{XX}^T$ matrisi de simetrik bir kare matristir, fakat boyutu $n\times n$'dir. $\mathbf{XX}^T$ matrisinin rankı en çok p olabileceği için sıfırdan farklı özdeğerlerinin sayısıda en çok p olabilecektir. Elde edilen bu sıfırdan farklı özdeğerler $\mathbf{X}^T\mathbf{X}$ matrisinin özdeğerleri ile aynıdır. $\mathbf{XX}^T$ matrisi $n-p$ adet sıfıra eşit özdeğere sahiptir. Bu $n-p$ adet özdeğer ve onların özvektörleri aşağıda verildiği şekildedir. $\mathbf{XX}^T$ matrisi özvektörlerinin matrisi **U** ile belirtilmektedir. Bu matris $\mathbf{X}^T\mathbf{X}$ ile $\mathbf{XX}^T$ matrisinin müşterek p adet özdeğerine karşılık gelen özvektörlerden oluşmaktadır. Her bir özvektör $\mathbf{u}_i$ $n\times 1$ boyutludur. $\mathbf{XX}^T$ matrisi için,

$$\mathbf{XX}^T = \mathbf{UDU}^T \tag{3.102}$$

eşitliği verilebilir. Eşitlik (3.101) ve (3.102) birlikte ele alınarak dikdörtgen matris **X**,

$$\mathbf{X} = \mathbf{UD}^{1/2}\mathbf{Z}^{\mathrm{T}}$$

şeklinde yazılabilir. Burada $\mathbf{D}^{1/2}$matrisi $\mathbf{X}^T\mathbf{X}$ matrisinin p adet özdeğerinin pozitif kare köklerinin oluşturduğu köşegen matristir. Buna göre $\mathbf{D}^{1/2}\,\mathbf{D}^{1/2} = \mathbf{D}$'dır. Eşitlik (3.102) dikdörtgen **X** matrisinin tekil değer ayrışımıdır. $\mathbf{D}^{1/2}$ matrisinin elemanları tekil değerlerdir ve **U** ile **Z**'deki sütun vektörleri ise sol ve sağ tekil vektörlerdir. $\mathbf{D}^{1/2}$ bir köşegen matris olduğundan,

$$\mathbf{X} = \sum^{p} \lambda_i \mathbf{u}_i \mathbf{z}_i^T$$

şeklindedir. Eğer özdeğerler en büyükten en küçüğe doğru düzenlenirse bu matrislerin ilki **X** için en iyi rank-1 yaklaşımını verir, ilk iki matrisin toplamı ise **X** için en iyi rank-2 yaklaşımını oluşturur, işlem bu şekilde devam eder. Bunlar EKK anlamında da en iyi yaklaşımlardır. EKK anlamında en iyi yaklaşım: Orijinal **X** matrisi ile daha iyi bir bütünlük sağlayan aynı ranka sahip başka bir matris yoktur şeklinde tanımlanır. Yaklaşımın ölçümü ise yaklaşım ve **X** matrisinin karşılıklı elemanları arasındaki farkın kareler toplamı ile elde edilir. Her bir durumdaki yaklaşımın uyum iyiliği, yaklaşımda kullanılan özdeğerlerin toplamının tüm özdeğerlerin toplamına oranı ile verilir. Başka bir deyişle rank-1 yaklaşımı $\lambda_1^2 / \sum \lambda_i^2$, rank-2 yaklaşımı $\left(\lambda_1^2 + \lambda_2^2\right) / \sum \lambda_i^2$ uyum iyiliğine sahiptir.

Tekil değer ayrışımı ana bileşenler analizinin ilk adımıdır. $\mathbf{X}=\mathbf{U}\mathbf{D}^{1/2}\mathbf{Z}^T$ sonucu ve $\mathbf{Z}^T\mathbf{Z}=\mathbf{I}$ özelliği kullanılarak $n \times p$ boyutlu

$$\mathbf{W}=\mathbf{X}\mathbf{Z}= \mathbf{U}\mathbf{D}^{1/2} \qquad (3.103)$$

şeklinde bir **W** matrisi tanımlanabilir. **Z** matrisinin ilk sütunu **X** matrisinin sağ tekil değerler vektörünün birincisidir. Başka bir deyişle $\mathbf{X}^T\mathbf{X}$ matrisinin özvektörlerinin ilkidir. Bu durumda birinci özvektördeki katsayılar, **W** matrisinin ilk sütununu oluşturacak şekilde **X** matrisinin sütunlarının belirli bir doğrusal fanksiyonunu tanımlar. $\mathbf{W}^T\mathbf{W}=\mathbf{D}$ olduğu belirtilmelidir. **W** bir $n \times p$ boyutlu fakat **X**'den farklı bir matristir. Bu matrisin bir özelliğide tüm sütunlarının ortogonal olmasıdır. **D** matrisinin köşegen dışındaki elemanları **W** matrisinin sütunları arasındaki çarpımların toplamı olup hepsi sıfır değerini almaktadır. **W** matrisinin *i*-inci sütununun kareler toplamı λ_i^2 olup **D** matrisinin *i*-inci köşegen elemanına eşittir. Eğer **X** matrisi *p* adet değişkenin *n* adet gözleminden oluşan $n \times p$ boyutlu bir matris ise *W* matrisinin her bir sütunu orijinal değişkenlerin doğrusal bir transformasyonu olan yeni bir değişkeni tanımlar. *i*-inci yeni değişken λ_i^2 kareler toplamına sahip olup tüm değişken çiftleri ortogonaldir. Bu analiz **X** matrisinin ana bileşenler analizi olarak bilinir. Ve **W** matrisinin sütunları ana bileşenlerdir.

Ana bileşen analizi **X** matrisinin sütunlarını oluşturan farklı değişkenlerin gözlemleri için kullanılabilecektir. Bu transformasyon bir ortogonal değişkenler setini oluşturur. Bu değişkenlerden birinci ana bileşen λ_1^2, toplam dağılımın en büyük kısmını açıklar. İkinci ana bileşen λ_2^2 ise kalan dağılımın mümkün olan en büyük miktarını açıklar. Toplam dağılım orijinal değişkenlerin kareler toplamına eşit olup tüm özdeğerlerin toplamı ile verilir.

$$tr\left(\mathbf{X}^T\mathbf{X}\right) = tr\left(\mathbf{W}^T\mathbf{W}\right) = \sum \lambda_i^2$$

3.16 MATRİSLER ÜZERİNE GENEL TEOREMLER

Teorem 3.63: I matrisi nxn boyutlu birim matris ise $tr(\mathrm{I}) = n$'dir.

Teorem 3.64: Eğer A idempotent ve tekil olmayan bir matris ise A=I'dır.

Teorem 3.65: Tam ranklı olmayan tüm idempotent matrisler pozitif yarı tanımlıdır.

Teorem 3.66: Eğer A, elemanları a_{ij} ve i. Köşegen elemanı sıfır olan bir idempotent matris ise A matrisinin i. sıra ve i. sütun elemanlarının hepsi sıfırdır.

Teorem 3.67: Eğer A, rankı r olan bir matris ise $tr(a) = r$ ‘dir.

Teorem 3.68: Eğer A ve B her ikisi de idempotent olan matrisler ise, AB çarpımının da idempotent olması için $AB=BA$ olması gereklidir.

Teorem 3.69: Eğer A idempotent ve $A+B$=I ise B matrisi de idempotendir ve $AB=BA$=0'dır.