BÖLÜM 5

5. EN KÜÇÜK KARELERİN GEOMETRİK YORUMU

En küçük kareler yönteminin tüm kavramları birkaç geometrik prensibin uygulanmasıyla görsel hale getirilebilir. Bu kavramların daha iyi anlaşılmasında geometrik yorum cebirsel yoruma göre daha faydalı ya da en azından anlaşılmalarını kolaylaştırıcı bir yaklaşım olabilir. Regresyon parametreleri, kareler toplamları, serbestlik dereceleri gibi kavramlar ve regresyonda karşılaşılan çoklu doğrusal bağlantı gibi problemler vektör geometrisi kullanılarak görsel hale getirilebilirler.

5.1 VEKTÖR UZAYI VE DOĞRUSAL MODEL BİLEŞENLERİ

Bu kısımda doğrusal modelin bileşenleri olan X matrisi, **y**, **e** ve **β** vektörlerinin vektör uzayındaki özellikleri incelenecektir.

Bağımsız değişkenlerin oluşturduğu **X** matrisi *p* adet sütun vektöründen oluşmaktadır, burada 1′lerden oluşan $\mathbf{x}_0$ sütun vektörü de dikkate alınmaktadır. **X** matrisinin her bir sütun vektörü *n* boyutlu uzaydaki bir vektör olarak grafiği çizilebilir. Şekil 5.5 de bu durum *n*=3 için çizilmiştir. Her bir sütun vektöründeki *n* eleman, *n*-boyutlu uzayda çizilmiş olan vektörün uç noktasını tanımlayan koordinatları vermektedir. Matrisi oluşturan *p* adet vektör, birlikte ele alındığında *n*-boyutlu uzayda *p* boyutlu bir alt uzay tanımlar, ($p<n$). Bu *p* boyutlu alt uzay X matrisinin *p* vektörünün doğrusal kombinasyonları ile elde dilebilecek olan noktalar setinden oluşur. Elde edilen bu alt uzay *X*-uzayı olarak adlandırılır. Eğer **X** matrisinin vektörleri doğrusal bağımsız değil ise *X*-uzayının boyutu **X** matrisinin rankı ile belirlenir.

y vektörü de *n*-boyutlu uzaydaki bir vektördür. Bu vektörün beklenen değeri,

$$E(\mathbf{y}) = \mathbf{X}\boldsymbol{\beta} = \beta_0\mathbf{x}_0 + \beta_1\mathbf{x}_1 + \ldots + \beta_k\mathbf{x}_k$$

parametreleri **β** ile belirtilen **X** matrisinin sütun vektörlerinin doğrusal bir fonksiyonudur. Bu durumda doğrusal model eşitlik (4.4)′de tanımlandığı gibi

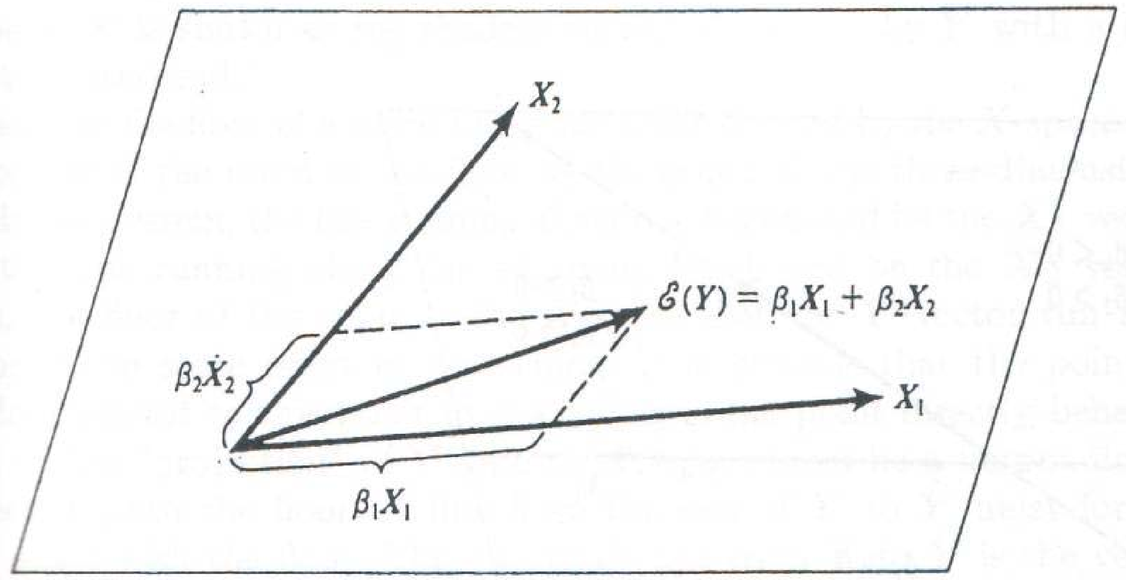
$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

şeklinde olup $E(\mathbf{y})$ ortalama vektörü tamamı ile *X*-uzayının içinde yer almaktadır, bkz Şekil 5.1. $E(\mathbf{y})$ noktasını belirleyen gerçek fakat bilinmeyen kısmi regresyon katsayılar vektörü **β** ′dır.

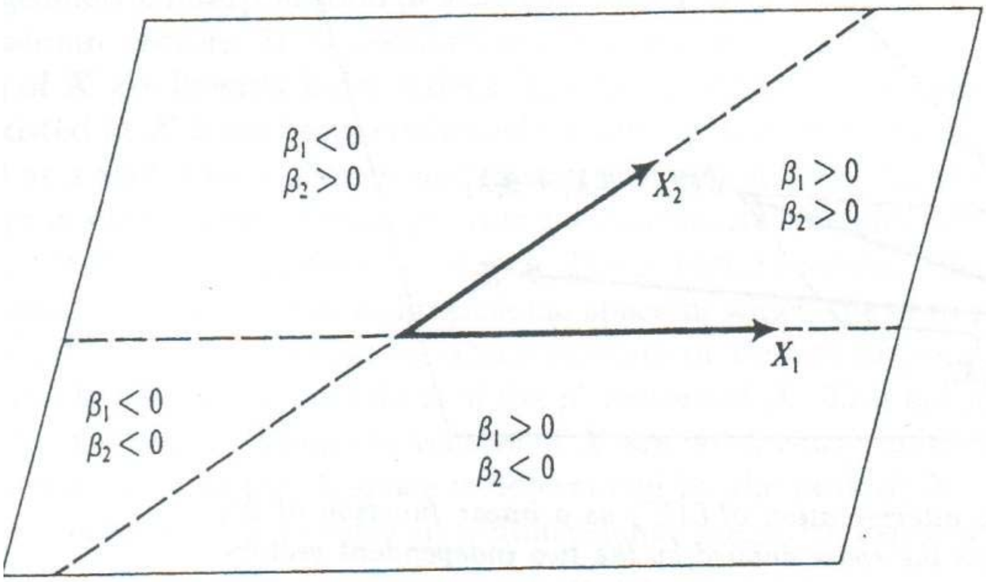
Bağımlı değişkenin gözlem vektörü **y**, *n*-boyutlu uzayda ortalama vektörü $E(\mathbf{y})$ çevresinde herhangi bir yere düşebilir. Bu vektörün tam pozisyonu **ε** hata terimi vektörüyle belirlenir. Eşitlik (4.4) ile verilen model **y** vektörünün, $E(\mathbf{y})$ ve **ε** vektörlerinin toplamı olduğu belirtilir. $E(\mathbf{y})$ vektörünün *X*-uzayında bulunmasına karşın **ε** ve **y** şans vektörleri *n*-boyutlu uzaydadır. Karşılaşılması pek mümkün olmayan ekstrem bir örnek alınmadıkça ne **ε** ne de **y** vektörleri *X*-uzayında bulunmayacaktır.

X matrisinin her biri üç gözlem içeren $\mathbf{x}_1$ ve $\mathbf{x}_2$ vektörlerinden oluştuğu varsayılsın. Vektörler Şekil 5.1′te gösterildiği gibi üç boyutlu uzayda gösterilebilir. Şekilde verilen düzlem $\mathbf{x}_1$ ve $\mathbf{x}_2$ vektörlerinin oluşturduğu iki boyutlu alt uzayı temsil etmektedir. $E(\mathbf{y})$ bu düzlemde bulunmakta ve $\mathbf{x}_1$ ile $\mathbf{x}_2$′nin doğrusal bir kombinasyonu şeklinde **y** nin gerçek ortalama vektörünü belirtmektedir. Şekildeki kesikli çizgiler $E(\mathbf{y})$ vektörünü veren, $\beta_1\mathbf{x}_1$ ve $\beta_2\mathbf{x}_2$ vektörlerinin vektörel toplamını göstermektedir. Bu durum şüphesiz modelin doğru olması halinde geçerli olacaktır. Pratikte $\boldsymbol{\beta}$ bilinmediği için $E(\mathbf{y})$ vektörü de bilinmemektedir. Regresyon analizinin amacı $\boldsymbol{\beta}$ vektörünün en iyi tahminlerini bulmaktır.

Şekil 5.1′deki $E(\mathbf{y})$′nin pozisyonu her iki parametrenin de β_1 ve β_2 pozitif olduğu bir durumu temsil etmektedir. $\beta_1\mathbf{x}_1$ ve $\beta_2\mathbf{x}_2$ vektörlerinin toplanmasıyla elde edilen $E(\mathbf{y})$ vektörü orijınal $\mathbf{x}_1$ ve $\mathbf{x}_2$ vektörleri ile aynı yöndedir. $E(\mathbf{y})$ vektörü $\mathbf{x}_1$ ve $\mathbf{x}_2$ vektörlerinin oluşturduğu açının dışında bir bölgeye düşmesi durumunda regresyon katsayılarının biri veya her ikisi de negatif olabilecektir. Bir vektörün negatif bir katsayı ile çarpımı vektörün yönünü ters yöne çevirir. Şekil 5.2β_1 ve β_2 ′in işaretlerine uygun olarak iki boyutlu X-uzayını bölümlere ayırmıştır.

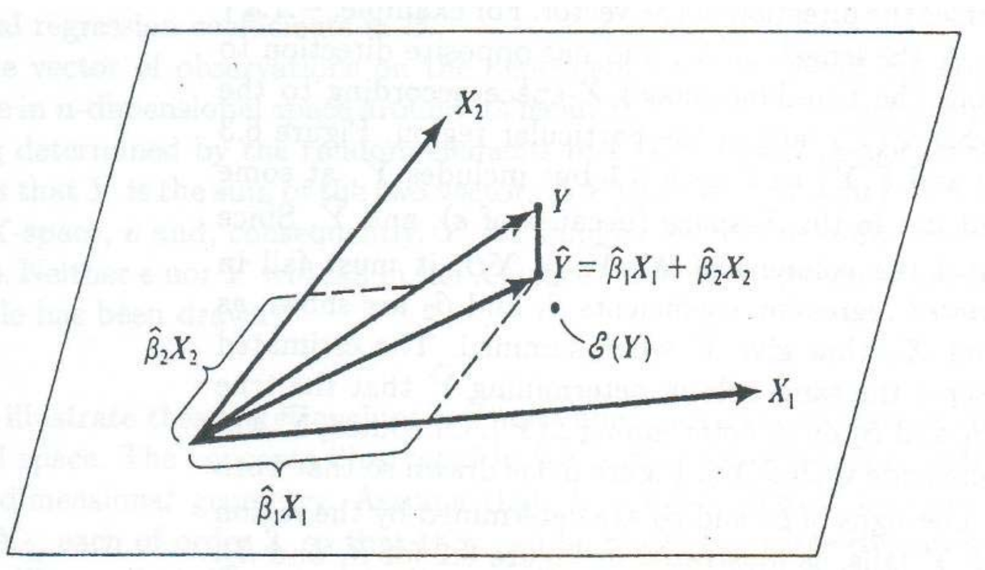


Şekil 5.1 $\mathbf{x}_1$ *ve* $\mathbf{x}_2$ *nin doğrusal bir kombinasyonu olarak* $E(\mathbf{y})$ *nin geometrik yorumu*



Şekil 5.2 $E(\mathbf{y})$ *nin düştüğü bölgelere göre* β_1 *ve* β_2 *nin işaretlerine uygun olarak iki boyutlu X-uzayının parçalanması.*

Şekil 5.3 daha önce Şekil 5.1 de verilen aynı X-uzayını ve $E(\mathbf{y})$ yi kullanarak $\mathbf{y}$, $E(\mathbf{y})$ ve $\hat{\mathbf{y}}$ vektörlerini içermektedir. Şekilden de görüldüğü gibi $\boldsymbol{\varepsilon}$ nedeniyle $\mathbf{y}$ vektörü X-uzayında bulunmaktadır. Tahminlenen $\hat{\mathbf{y}}$ vektörü $\hat{\mathbf{y}}=\mathbf{Xb}$ şeklinde $\mathbf{X}$ matrisinin sütunlarının doğrusal bir fonksiyonu olduğu için her zaman X-uzayında yer alır. Tahminlenmiş regresyon katsayıları b_1 ve b_2 ise $\mathbf{x}_1$ ve $\mathbf{x}_2$ vektörleri ile çarpıldıktan sonra bu çarpımların toplamı $\hat{\mathbf{y}}$ 'i vermektedir.



Şekil 5.3 $\mathbf{y}$ *ve* $\hat{\mathbf{y}}$ *vektörlerinin geometrik ilişkisi*

Tahminlenmiş regresyon katsayıları $\hat{\mathbf{y}}$ vektörünün, gerçek regresyon katsayıları ise $E(\mathbf{y})$ vektörünün belirlenmesinde yardımcı olmaktadır. Şüphesiz pek çok durumda $\hat{\mathbf{y}}$ ile $E(\mathbf{y})$ değerleri birbirinden farklı olacaktır.

Şekil 5.3 de $\mathbf{y}$ ile $\hat{\mathbf{y}}$ vektörlerini birleştiren kısa vektör $\mathbf{e}$ artık vektörüdür. EKK prensibine göre $\mathbf{b}$ ve dolayısıyla $\hat{\mathbf{y}}$, $\sum (y_i - \hat{y}_i)^2 = \mathbf{e}^T\mathbf{e}$ değerini minimize edecek şekilde seçilir. $\mathbf{e}^T\mathbf{e}$ değeri $\mathbf{e}$ vektörünün karesel uzaklığını verir. Geometrik olarak bu değer $\mathbf{y}$ vektörünün uç noktasından $\hat{\mathbf{y}}$ vektörünün uç noktasına olan uzaklığın karesini belirtmektedir. Sonuç olarak n boyutlu uzaydaki $\mathbf{y}$ vektörüne X-uzayındaki en yakın vektörün $\hat{\mathbf{y}}$ olduğu söylenebilir. $\mathbf{y}$ vektörüne düzlemdeki en yakın nokta, $\mathbf{y}$ vektöründen, düzleme yapılan ortogonal (dik) izdüşüm ile elde edilebilecektir Bu nedenle $\mathbf{e}$ vektörü X-uzayına ortogonal olmak zorundadır.

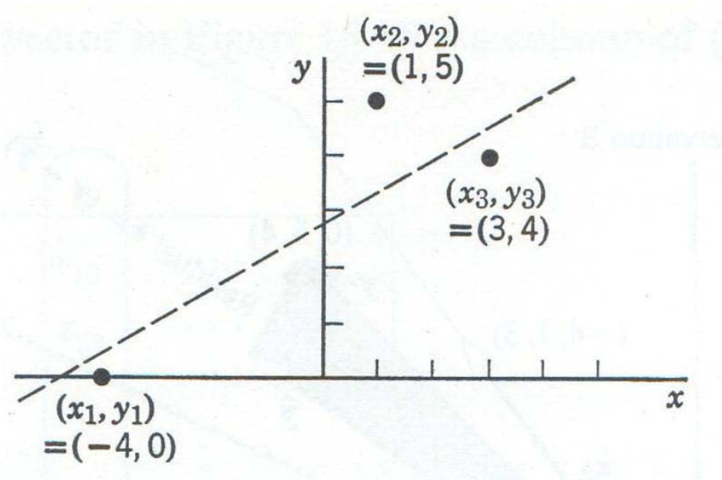
$\hat{\mathbf{y}}$ vektörü, eşitlik (4.14) den görülebileceği gibi $\hat{\mathbf{y}}=\mathbf{Hy}$ seklinde de yazılabilir. $\mathbf{H}$ matrisi, $\mathbf{X}$ matrisinin sütunları ile tanımlanan p boyutlu alt uzaya $\mathbf{y}$ vektörünün izdüşümünü verir. Başka bir deyişle $\mathbf{y}$ vektörü soldan $\mathbf{H}$ matrisi ile çarpıldığında $\hat{\mathbf{y}}$ vektörü elde edilir. Bu iki vektör X-uzayına ortogonal olan $\mathbf{e}$ vektörü ile birleştirilebilir. $\mathbf{y}$ ve $\hat{\mathbf{y}}$ vektörleri arasındaki izdüşüm ilişkisini sağlaması nedeniyle $\mathbf{H}$ matrisi izdüşüm matrisi olarak adlandırılır.

Sadece bir bağımsız değişken $\mathbf{x}_1$ üzerine $\mathbf{y}$ nin regresyonu oluşturulduğunda, elde edilen artık kareler toplamı, iki bağımsız değişkenli regresyonda elde edilen artık kareler toplamdan küçük olmayacaktır. Sadece $\mathbf{x}_1$ olması durumunda X-uzayı sadece $\mathbf{x}_1$ in temsil ettiği noktalar seti ile (bir vektör) tanımlanacaktır. $\mathbf{x}_1$ ile tanımlanmış uzay üzerine $\mathbf{y}$ nin izdüşümü bu noktalar setindeki (vektör üzerindeki) bir tek nokta olacaktır. Sadece $\mathbf{x}_1$ ile tanımlanmış olan alt uzay $\mathbf{x}_1$ ve $\mathbf{x}_2$'nin birlikte tanımladıkları uzayın bir bölümünü oluşturmaktadır. Buna göre $\mathbf{x}_1$ in noktalar setindeki hiçbir nokta $\mathbf{y}$ vektörünün uç noktasına, $\mathbf{x}_1$ ve $\mathbf{x}_2$ vektörlerinin oluşturduğu düzlemdeki en yakın noktadan daha yakın değildir. $\mathbf{x}_1$ üzerine $\mathbf{y}$ nin ve $\mathbf{x}_1$ ile $\mathbf{x}_2$ nin üzerine $\mathbf{y}$ nin regresyonundan elde edilen iki hata terimi vektörünün uzunluğu eğer düzlem üzerine yapılan izdüşüm tamamı ile $\mathbf{x}_1$ vektörü ile çakışıyorsa aynıdır. Bu durumda $\beta_{2.1}$ sıfır olacaktır. Yukarıdaki açıklamadan, bağımsız değişkenlerin bir alt seti üzerine $\mathbf{y}$ nin regresyonundan elde edilen hata terimi kareler toplamından aha küçük olmayacaktır

Bir bağımsız değişken için ardışık kareler toplamı bu sürecin genişlemesidir. Bununla birlikte ilk olarak $\mathbf{y}$ ve mevcut ilgilenilen bağımsız değişkenin modelde önceden mevcut tüm bağımsız değişkenlerin tanımladığı uzay üzerine izdüşümü yapılır. Daha sonra $\mathbf{e}_y$ olarak belirtilebilecek $\mathbf{y}$ 'nin hata terimi vektörünün $\mathbf{e}_x$ olarak belirtilen mevcut x değişkenleri için oluşturulan hata terimi vektörlerinin tanımladığı uzay üzerine izdüşümü gerçekleştirilir. İlgilenilen bağımsız değişkenin ardışık kareler toplamı, $\mathbf{e}_x$ üzerine $\mathbf{e}_y$ 'nin bu izdüşümü için $\hat{\mathbf{y}}$ vektörünün karesel uzunluğudur. Bu arada bağımlı değişken ile ilgilenilen bağımsız değişkenin her ikisinin de önceden modelde bulunan tüm bağımsız değişkenlerin etkisinden arındırıldığı belirtilmektedir. Ardışık analizindeki her bir adımdaki yeni X-uzayı bir boyutlu bir uzaydır ve bu nedenle her bir adımdaki ardışık kareler toplamı bir serbestlik derecesine sahip olacaktır. E.K.K′ daki hata terimi vektörü daima $\mathbf{y}$ 'nin izdüşümünün yapıldığı X-uzayına ortogonal olduğundan $\mathbf{e}_y$ ve $\mathbf{e}_x$ 'in her ikisi de modelde önceden mevcut olan tüm bağımsız değişkenlere ortogonaldir. Önceki X-uzayına olan bu ortogonallik özelliği nedeni ile ardışık kareler toplamları ve serbestlik dereceleri toplanabilir. Sonuç olarak her bir adım için ardışık kareler toplamlarının toplamı ve serbestlik derecelerinin toplamı tüm bağımsız değişkenlerinin içerildiği bir tek modelin kullanılması durumunda ne elde edilebileceğini gösterecektir.

5.2 EKK UYUMUNUN GEOMETRİSİ

Kısım 5.1 de genel doğrusal modelin geometrisi ele alındı. Bu kısımda ise yukarıda açıklanan konular grafiksel basitliğin sağlanması amacıyla Bölüm 2'de tanımlanan basit regresyon doğrusunun uyumunun yapılması problemi üzerinden ele alınacaktır. Basit regresyon dikkate alındığından $\mathbf{X}$ matrisi sadece $\mathbf{x}_0$ ve $\mathbf{x}_1$ vektörlerini içerecektir. Gözlem sayısının parametre sayısından fazla olması gerektiğinden Şekil 5.4 de görüldüğü gibi örnek sadece gözlenmiş üç noktadan oluşmaktadır.



Şekil 5.4 *Regresyon doğrusu*; *noktalar gözlemleri*, *eksenler değişkeni belirlenmektedir*.

$\mathbf{x}_1$ vektöründeki değerler sıfır noktasında merkezlenmiştir. Bölüm 2'de belirtildiği gibi ortalamadan farklar alınarak bu merkezleme işlemi her zaman gerçekleştirilebilir. Matematiksel model matris gösterimde; Gözlenmiş Vektör=Ortalama Vektörü+Hata Vektörü,

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ 1 & x_3 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \end{bmatrix} \tag{5.1}$$

yazılabilir. EKK uyumu benzer şekilde ifade edilebilir; Gözlenmiş vektör ≈ Uyumu yapılmış vektör,

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} \approx \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ 1 & x_3 \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \end{bmatrix} \tag{5.2a}$$

Eşitlik (5.2*a*),

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} \approx b_0 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + b_1 \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \tag{5.2b}$$

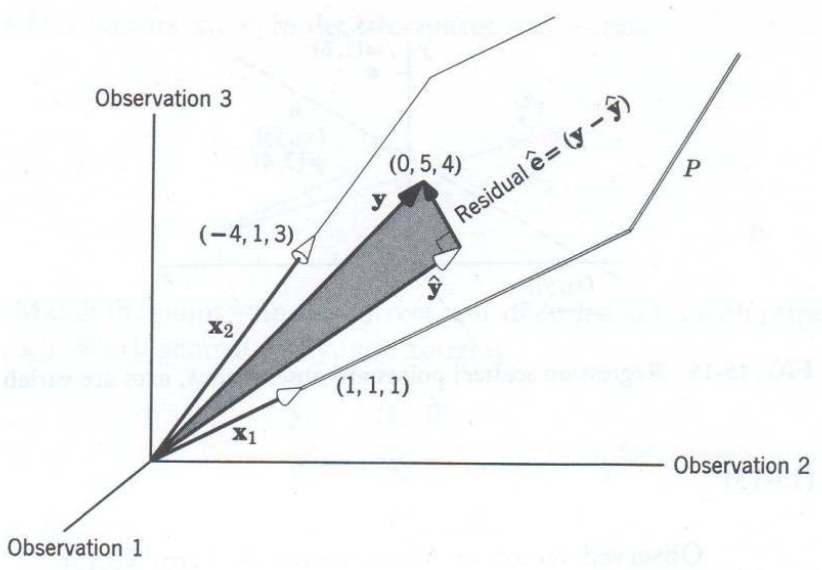
şeklinde veya sütun vektörleri gösteriminde,

$$\mathbf{y} \approx b_0 \mathbf{x}_0 + b_1 \mathbf{x}_1 \tag{5.2c}$$

olarak belirtilebilir. Bu ifade,

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}, \quad \mathbf{x}_0 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{x}_1 = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \tag{5.3}$$

olarak verilmiştir. Şekil 5.4'de, üç gözlemin her biri iki boyutlu uzayda ele alınmıştır. Bu şekilde *x* ve *y* değişkenlerinin her biri bir boyut ile belirtilmiştir. Şekil 5.5'de ise her bir değişken üç boyutlu uzayda plot edilmiştir.



Şekil 5.5 *Şekilde, noktalar değişkenleri, eksenler gözlemleri belirtmektedir.*

Gözlemlerin her biri bir boyuta karşılık gelmektedir. Eşitlik (5.2*a*),

$$\begin{bmatrix} 0 \\ 5 \\ 4 \end{bmatrix} \approx \begin{bmatrix} 1 & -4 \\ 1 & 3 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \end{bmatrix} \tag{5.4}$$

şeklinde tekrar yazılabilir. Şekil 5.4'de işaretlenmiş her bir nokta eşitlik (5.4)'ün bir sırasını belirtmektedir. Başka bir deyişle (-4,0) noktası birinci sırayı ifade etmektedir. Buna karşın Şekil 5.5'deki her bir nokta veya vektör eşitlik (5.4)'ün bir sütununa karşılık gelmektedir. Bu durum aşağıdaki şekilde genelleştirilebilir:

$$\begin{array}{cc} Y & X \\ \begin{bmatrix} y_1 \\ . \\ y_i \\ . \\ . \\ y_n \end{bmatrix} & \begin{bmatrix} x_{11} & x_{12} & . \ . & x_{1k} \\ . & . & . \ . \ . & . \\ x_{i1} & x_{i2} & . \ . & x_{ik} \\ . & . & . \ . \ . & . \\ . & . & . \ . \ . & . \\ x_{n1} & \underbrace{x_{n2}}_{\text{Şekil 5.5 deki bir vektöre karşılık gelen bir değişken}} & . \ . & x_{nk} \end{bmatrix} \rightarrow \textit{Şekil 5.4 deki bir noktaya karşılık gelen bir gözlem} \end{array}$$

Verilen örnekte $\mathbf{x}_0$ ve $\mathbf{x}_1$ ortogonaldir. Çünkü

$$\mathbf{x}_0\mathbf{x}_1^T = (1,1,1)(x_1, x_2, x_3)^T$$
$$= x_1 + x_2 + x_3$$
$$= n\bar{x}$$

olup, $\bar{x} = 0$ olması nedeniyle

$$\mathbf{x}_0\mathbf{x}_1^T = 0$$

sonucu elde edilebilir. Bu sonuç **x** vektörünün sıfır ortalamaya sahip olacak şekilde merkezlenmesinin faydalı olabileceğini göstermektedir.

Eşitlik (5.2*c*)'deki sorun $\mathbf{x}_0$ ve $\mathbf{x}_1$'nin doğrusal kombinasyonu olacak şekilde uyumu yapılmış bir y değeri bulmaktır. Şekil 5.5'de gösterildiği gibi bunun geometrik anlamı, $\mathbf{x}_0$ ve $\mathbf{x}_1$ tarafından oluşturulan *P* düzlemindeki herhangi bir yerde bir uyumun seçilmesinin zorunlu olmasıdır. Gözlenmiş **y** vektörüne en iyi uyumu sağlayan (ya da en yakın) *P* düzlemindeki nokta veya vektör belirlenmek isteniyorsa, *P* düzlemi üzerine **y** vektöründen ortogonal bir izdüşüm gerçekleştirilerek $\hat{\mathbf{y}}$ vektörü elde edilebilir. Hatırlanacağı gibi EKK metodu $\sum(y_i - \hat{y}_i)^2$ ifadesinin minimizasyonunu içermekteydi. Bu ifade eşitlik (3.29)'a uygun olarak vektör gösteriminde,

$$\|\mathbf{y} - \hat{\mathbf{y}}\|^2$$

şeklinde verilebilir. EKK yöntemi **y** ve $\hat{\mathbf{y}}$ vektörleri arasındaki karesel uzaklığı minimize etmeyi içerir. Tahminlenmiş artık vektörü ($\mathbf{x}_0$, $\mathbf{x}_1$) düzlemine ortogonaldir. Bu nedenle **e** vektörü düzlemdeki $\mathbf{x}_0$ ve $\mathbf{x}_1$ değişkenlerinin (vektörlerinin) her birine ortogonaldir, $\mathbf{x}_0\mathbf{e}=0$ ve $\mathbf{x}_1\mathbf{e}=0$. Daha genel olarak normal denklemler kullanılarak;

$$\mathbf{X}^T\mathbf{X}\mathbf{b} = \mathbf{X}^T\mathbf{y}$$

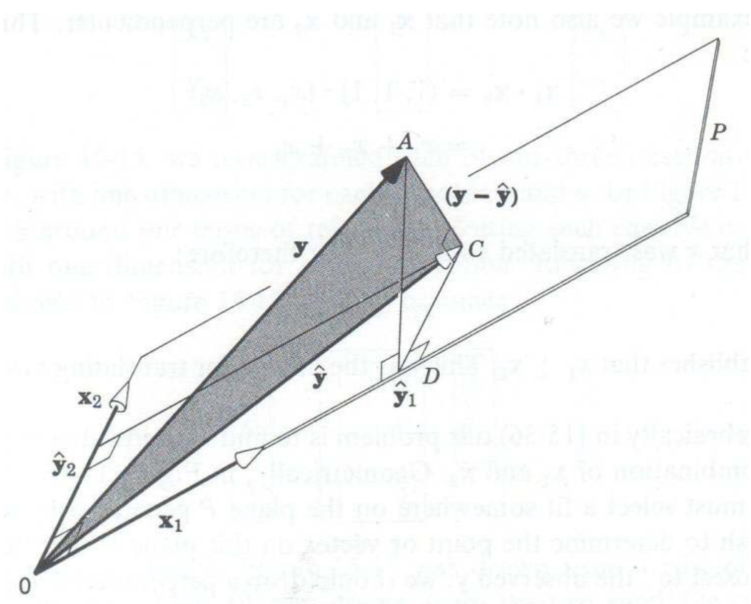
$$\mathbf{X}^T\mathbf{y} - \mathbf{X}^T\mathbf{X}\mathbf{b} = \mathbf{0}$$

$$\mathbf{X}^T(\mathbf{y} - \mathbf{X}\mathbf{b}) = \mathbf{X}^T(\mathbf{y} - \hat{\mathbf{y}}) = \mathbf{X}^T\mathbf{e} = \mathbf{0} \tag{5.5}$$

sonucu elde edilebilir. Bu EKK yönteminin temel bir sonucudur. Eşitlik (5.5) de **X** matrisinin birinci sütunu $\mathbf{x}_0$=**1** olduğundan $\sum e_i = \mathbf{x}_0\mathbf{e} = 0$ eşitliğini verir. Regresyon modelinin sabit terim içerdiği durumlarda, EEK artık terimleri daima sıfır ortalamaya sahiptir. Eşitlik (5.5) in diğer elemanları artıkların her bir **x** değişkeni ile sıfır korelasyona (örnek korelasyonu) sahip olduğunu belirtir.

Eğer bağımsız değişkenler Şekil 5.5'de gösterildiği gibi birbirlerine ortogonal ise b_0 ve b_1 için formüller geometriden yararlanılarak basit bir şekilde elde edilebilir. Şekil 5.6'deki gösterim kullanılarak, $\hat{\mathbf{y}}$ vektörü, $\mathbf{x}_0$ ve $\mathbf{x}_1$ üzerine **y**'nin bireysel ortogonal izdüşümleri elde edilerek $\hat{\mathbf{y}}_0$ ve $\hat{\mathbf{y}}_1$'nin vektörel toplamı şeklinde bulunabilir:

$$\hat{\mathbf{y}} = \hat{\mathbf{y}}_0 + \hat{\mathbf{y}}_1 \tag{5.6a}$$



Şekil 5.6 $\hat{\mathbf{y}}_0$ *ve* $\hat{\mathbf{y}}_1$ *in* $\hat{\mathbf{y}}$ *üzerine bireysel izdüşümleri*

Eşitlik (3.40)'dan

$$\hat{\mathbf{y}} = \left(\frac{\mathbf{y}^T\mathbf{x}_0}{\|\mathbf{x}_0\|^2}\right)^T \mathbf{x}_0 + \left(\frac{\mathbf{y}^T\mathbf{x}_1}{\|\mathbf{x}_1\|^2}\right)^T \mathbf{x}_1 \qquad (5.6b)$$

elde edilebilir. Eşitlik (5.6*b*) ile (5.2*c*) ile karşılaştığında b_0 için, payın bir skaler olması nedeniyle,

$$b_0 = \frac{\mathbf{y}^T\mathbf{x}_0}{\|\mathbf{x}_0\|^2} \qquad (5.7a)$$

yazılabileceği görülmektedir. Eşitlik (5.3)'den $\mathbf{x}_0$'ın tüm elemanları bir olan vektör olduğu açıktır. Bu durumda

$$\mathbf{y}^T\mathbf{x}_0 = y_1 + y_2 + y_3 \text{ ve genellenerek } \mathbf{y}^T\mathbf{x}_0 = \sum_{i=1}^{n} y_i$$

$$\|\mathbf{x}_0\|^2 = 1+1+1 \qquad \text{ve genellenerek } \|\mathbf{x}_0\|^2 = n$$

sonuçları elde edilir. Eşitlik (5.7*a*) için,

$$b_0 = \bar{y} \qquad (5.7b)$$

ve benzer şekilde b_1 için,

$$b_1 = \frac{\mathbf{y}^T\mathbf{x}_1}{\|\mathbf{x}_1\|^2} \qquad (5.8a)$$

eşitliğin payı iki vektörün nokta çarpımını tanımladığından sonuç olarak,

$$b_1 = \frac{\sum x_i y_i}{\sum x_i^2} \qquad (5.8b)$$

bulunabilir. Bu sonuçlar merkezlenmiş değişkenler için Bölüm 2 de elde edilen eşitlikler ile tutarlıdır.

Buraya kadar verilenlerin EKK tahminleme yöntemini kullanan çok değişkenli regresyon problemi ile bütünlüğü aşağıda açıklanacaktır. **X** matrisi p adet n elemanlı sütun vektöründen oluşmakta idi, **X**=[$\mathbf{x}_0$ $\mathbf{x}_1$ … $\mathbf{x}_k$]. Burada gözlem sayısının değişken sayısından fazla olduğu varsayılmaktadır, ($n>p$). **X** sütun vektörleri zinciri IR^n uzayının bir alt uzayıdırlar. Bu alt uzayın boyutu p değerini aşamaz, eğer **X** sütunları doğrusal bağımsız ise p'ye eşit olur. Bu alt uzay X'in sütun uzayı olarak adlandırılır. Eğer **y** vektörü **X**'in sütun uzayında bulunsaydı **X** vektörlerinin tam bir doğrusal kombinasyonu şeklinde ifade edilebilir ve tüm örnek gözlemleri için hata terimi sıfır değerini alırdı. Şekil 5.6 de gösterilen P düzlemi **X** matrisinin sütunlarının tanımladığı alt uzayı temsil etmektedir. **X** vektörlerinin herhangi bir **b** vektörü ile çarpılması ile,

$$\mathbf{Xb} = \left[\mathbf{x}_0 \mathbf{x}_1 \cdots \mathbf{x}_k\right] \begin{bmatrix} b_0 \\ b_1 \\ \vdots \\ b_k \end{bmatrix} = b_0 \mathbf{x}_0 + b_1 \mathbf{x}_1 + \cdots + b_k \mathbf{x}_k$$

vektörü elde edilir. Elde edilen bu vektör **X** sütun uzayının içindedir. Farklı **b** vektörleri farklı uzunlukta **Xb** vektörlerini oluşturur. Her bir **Xb** vektörü için farklı bir artık **e** vektörü mevcuttur. Şekil 5.6'den görüldüğü gibi **y** vektörü diğer iki vektörün toplamı şeklindedir. Bu vektörlerden **Xb**, **X** sütun uzayının içinde diğer vektör **e** ise bu sütun uzayının dışındadır. **b** vektörünün, **Xb** vektörünün eğiminin, **y** vektörünün eğimine mümkün olduğunca yakın olacak şekilde seçilmesi istenir. Başka bir deyişle **e** vektörünün uzunluğunun minimize edilmesi istenir. Bu minimizasyon, **X** sütunlarının oluşturduğu hiper düzleme dik bir **e** vektörü oluşturularak elde edilebilir. Bu durumda **e** vektörü **X** sütunlarının herhangi bir doğrusal kombinasyonuna ortogonal olmak zorundadır. Eşitlik (5.5) ile bu koşulun sağlandığı görülmektedir.

Bu kısa açıklamadan sonra, EKK tahmin vektörü **b** için eşitlik (4.8)'in nasıl çözüleceği araştırılabilir. $\mathbf{X}^T\mathbf{X}$ matrisi $p\times p$ boyutlu bir matris, **b** ve $\mathbf{X}^T\mathbf{y}$ ise $p\times 1$ boyutlu vektörlerdir. Buna uygun olarak $\mathbf{X}^T\mathbf{y}$ vektörü $\mathbf{X}^T\mathbf{X}$ sütunlarının doğrusal bir kombinasyonu olarak ifade edilebilir ve **b** vektörü bu doğrusal kombinasyonun katsayılarını belirtir. Eğer $\mathbf{X}^T\mathbf{X}$ sütunları doğrusal bağımsız ise IR^p uzayı için kaynak vektörleri tanımlarlar ve $\mathbf{X}^T\mathbf{y}$ şeklindeki herhangi p elemanlı vektör bu kaynak vektörlere göre eşsiz olarak ifade edilebilir. Diğer bir deyişle eşitlik (4.8) **b** vektörü için eşsiz bir çözüme sahiptir. Eşitlik (4.8)'in **b** için çözümü bir ters matrisin kullanılmasıyla elde edilebilir. Eğer $\mathbf{X}^T\mathbf{X}$ matrisinin p sütununun doğrusal bağımsız olduğu kabul edilirse, bu matrisin $(\mathbf{X}^T\mathbf{X})^{-1}$ şeklinde bir ters matrisi mevcuttur. Eşitlik (4.8)'in her iki tarafı bu ters matris ile soldan çarpılarak, $\mathbf{b}=(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y}$ eşitliği elde edilir. Bu ifade örnek verileri ile elde edilen EKK'nın **b** tahmin vektörünü verir. Bu vektörün elde edilebilmesi için $\mathbf{X}^T\mathbf{X}$ matrisinin tersi alınabilir bir matris olması gereklidir. Bunun için de $\mathbf{X}^T\mathbf{X}$ matrisinin sütunlarının doğrusal bağımsız olması gereklidir. Bu konu bir matrisin rankı kavramı ile yakın olarak ilgilidir.

Eşitlik (4.8) ve (4.9)'deki EKK matris ve vektörleri içinde en önemli olanı $\mathbf{X}^T\mathbf{X}$ matrisidir. Teorem 3.16 ile

$$\rho(\mathbf{X}^T\mathbf{X})=\rho(\mathbf{X}\mathbf{X}^T)=\rho(\mathbf{X}^T)=\rho(\mathbf{X}) \tag{5.9}$$

tanımlanır. **X** matrisi $n \times p$ ve $\rho(\mathbf{X})=r$ olsun. Eşitlik (3.57) kullanılarak **X**'in boş uzayının boyutu $p-r$ olarak bulunur. **m** bu boş uzaydaki herhangi bir vektör (çözüm vektörü) ise,

$$\mathbf{Xm}=\mathbf{0}$$

eşitliği elde edilir. Soldan $\mathbf{X}^T$ ile çarpılarak,

$$\mathbf{X}^T\mathbf{Xm}=\mathbf{0}$$

sonucu bulunur. Bu durumda **m** aynı zamanda $\mathbf{X}^T\mathbf{X}$'inde boş uzayındadır. Eğer **s** vektörü, $\mathbf{X}^T\mathbf{X}$'in boş uzayındaki herhangi bir vektör ise,

$$\mathbf{X}^T\mathbf{Xs}=\mathbf{0}$$

elde edilir ve soldan $\mathbf{s}^T$ ile çarpılarak,

$$\mathbf{s}^T\mathbf{X}^T\mathbf{Xs}=(\mathbf{Xs})^T(\mathbf{Xs})=\mathbf{0}$$

sonucu bulunur. Bu durumda **Xs** uzunluğu sıfır olan bir vektördür ve boş (sıfır) vektör,

$$\mathbf{Xs}=\mathbf{0}$$

olmak zorundadır. Sonuç olarak **s** vektörü **X** matrisinin boş uzayındadır. Görüldüğü gibi **X** ve $\mathbf{X}^T\mathbf{X}$ aynı boş uzaya sahiptirler. Her iki matrisin de sütun sayısı p'dir. Eşitlik (3.57)'ye göre her bir matris r ranka sahiptir.

EKK durumunda **X** matrisi $n \times p$ boyutludur ($p \langle n$). Bağımsız değişkenler arasında tam bir doğrusal ilişki olmadıkça **X** matrisi tam sütun ranklı olup $\mathbf{X}^T\mathbf{X}$ matrisi de tam ranklıdır:

$$\rho(\mathbf{X}^T\mathbf{X})=p$$

$\mathbf{X}^T\mathbf{X}$ matrisi $p \times p$ boyutlu tekil olmayan bir kare matris olduğundan tersi $(\mathbf{X}^T\mathbf{X})^{-1}$ mevcuttur. **X** tam sütun ranklı olsa bile, $n \times n$ boyutlu $\mathbf{X}\mathbf{X}^T$ matrisi tekil bir matristir.

İzdüşümün amacı n boyutlu uzayda **y** vektöründen bir alt uzaydaki $\hat{\mathbf{y}}$ vektörüne (**y** vektörüne olabildiğince yakın bir $\hat{\mathbf{y}}$ vektörü) transformasyondur. **y** vektöründen $\hat{\mathbf{y}}$ vektörüne olan bir doğrusal transformasyon eşitlik (4.14) ile tanımlanmıştır. Bu transformasyonun bir izdüşüm olabilmesi için **H** matrisinin idempotent ve simetrik olması gerekir. Bu şartın sağlanması durumunda **H** matrisi izdüşüm matrisi olarak adlandırılır. İzdüşüm alt uzayı **H** matrisinin sıraları ve sütunları ile tanımlanır. Eğer **H** bir izdüşüm matrisi ise (**I-H**) matrisi de bir izdüşüm matrisidir. Bununla birlikte **H** ve (**I-H**) matrisleri ortogonal olduğundan, (**I-H**) matrisinden **H** alt uzayına tanımlanan izdüşüm de ortogonaldir. İzdüşüm matrisi idempotent olduğundan (idempotent matrisin izi rankına eşittir) bir izdüşüm matrisinin rankı, izdüşümün alt uzayının boyutuna eşittir. Örneğin **H** matrisi,

$$\mathbf{H}=\frac{1}{6}\begin{bmatrix} 5 & 2 & -1 \\ 2 & 2 & 2 \\ -1 & 2 & 5 \end{bmatrix}$$

idempotent ve simetrik matris olduğu için bir izdüşüm matrisidir. Buna uygun olarak,

$$\hat{\mathbf{y}} = \mathbf{H}\mathbf{y}_1 = \frac{1}{6}\begin{bmatrix} 5 & 2 & -1 \\ 2 & 2 & 2 \\ -1 & 2 & 5 \end{bmatrix}\begin{pmatrix} 3 \\ 10 \\ 20 \end{pmatrix} = \begin{pmatrix} 2,5 \\ 11 \\ 19,5 \end{pmatrix}$$

şeklinde $\mathbf{y}_1$ vektörünün, **H** matrisinin sütunları ile tanımlanan alt uzay üzerine bir izdüşümü elde edilir. Elde edilen $\hat{\mathbf{y}}$ vektörü, $\mathbf{y}_1$ vektörüne en yakın olan bu alt uzaydaki eşsiz vektördür. Başka bir deyişle $(\mathbf{y}_1 - \hat{\mathbf{y}})^T(\mathbf{y}_1 - \hat{\mathbf{y}})$ minimumdur. **H** bir izdüşüm matrisi olduğu için

$$\mathbf{I} - \mathbf{H} = \frac{1}{6}\begin{bmatrix} 1 & -2 & 1 \\ -2 & 4 & -2 \\ 1 & -2 & 1 \end{bmatrix}$$

matrisi de bir izdüşüm matrisidir.

$$\mathbf{e} = (\mathbf{I} - \mathbf{H})\mathbf{y}_1 = \frac{1}{6}\begin{bmatrix} 1 & -2 & 1 \\ -2 & 4 & -2 \\ 1 & -2 & 1 \end{bmatrix}\begin{pmatrix} 3 \\ 10 \\ 20 \end{pmatrix} = \begin{pmatrix} 1/2 \\ -1 \\ 1/2 \end{pmatrix}$$

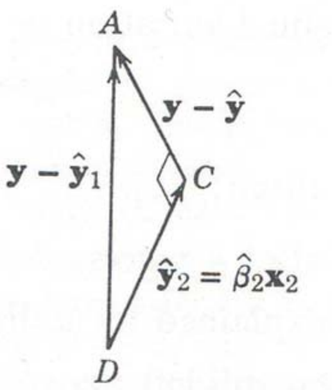
işlemi ile elde edilen **e** vektörü, **H** matrisinin tanımladığı alt uzaya ortogonal olan (**I**–**H**) alt uzayı içine bir izdüşümdür. Sonuçlardan görüleceği gibi $\hat{\mathbf{y}}^T\mathbf{e}=0$ ve $\hat{\mathbf{y}}+\mathbf{e}=\mathbf{y}_1$ eşitlikleri elde edilebilir.

5.3 VARYANS ANALİZİNİN GEOMETRİSİ

Şekil 5.6'deki *ADC* üçgeni ele alınıp yeniden çizilerek Şekil 5.7 elde edilir. Pisagor teoremi uygulanarak,

$$\|\mathbf{y} - \hat{\mathbf{y}}_0\| = \|b_1\mathbf{x}_1\|^2 + \|\mathbf{y} - \hat{\mathbf{y}}\|^2 \tag{5.10}$$

eşitliği yazılır.



Şekil 5.7 *Şekil 5.6'ün küçük üçgeni için Pisagor teoremi*

$$\hat{\mathbf{y}}_0 = b_0\mathbf{x}_0 \tag{5.11}$$

olduğundan eşitlikler (5.3) ve (5.7*b*) kullanılarak

$$\hat{\mathbf{y}}_0 = \bar{y}\cdot\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} \bar{y} \\ \bar{y} \\ \bar{y} \end{bmatrix}$$

yazılabilir. Bu durumda $(\mathbf{y}-\hat{\mathbf{y}}_0)$ ifadesi

$$\left(\mathbf{y}-\hat{\mathbf{y}}_0\right)=\begin{bmatrix} y_1-\bar{y} \\ y_2-\bar{y} \\ y_3-\bar{y} \end{bmatrix} \tag{5.12}$$

şeklindedir. $\mathbf{x}_0$ kukla değişkeni üzerine **y**'nin uyumunun yapılması ile **y**'nin ortalamadan farklarına göre ifadesi elde edilmekte olup bu durum Şekil 5.6 ve 5.7'ün incelenmesi ile de görülebilir. Eşitlik (5.12)'den,

$$\left\|\mathbf{y}-\hat{\mathbf{y}}_0\right\|^2 = \textit{Düzeltilmiş kareler toplamı} \tag{5.13}$$

bulunabilir. Benzer olarak eşitlik (5.10)'nın sağ tarafı

$$\left\|b_1\mathbf{x}_1\right\|^2 = b_1^2\left\|\mathbf{x}_1\right\|^2 = \textit{Regresyon kareler toplamı} \tag{5.14a}$$

ve

$$\left\|\mathbf{y}-\hat{\mathbf{y}}\right\|^2 = \textit{artık kareler toplamı} \tag{5.14b}$$

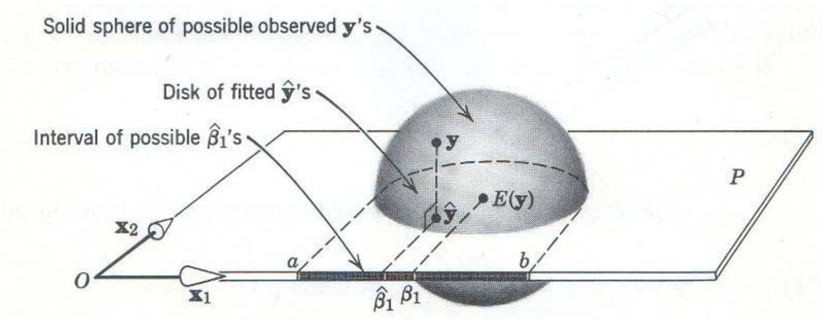
şeklinde verilebilir. Geometrik olarak bu durum, **1** vektörü ile tanımlanan bir boyutlu uzaya **y** vektörünün izdüşümüne eşittir. β_0'ın EKK tahmini $\overline{Y}$'dır ve bu izdüşümünden elde edilen artık vektörü, Y_i değerlerinin $\overline{Y}$'dan sapmalarının oluşturduğu $y_i = Y_i - \overline{Y}$ vektördür. Elde edilen bu artık (ortalamadan sapma) vektörünün karesel uzaklığı **y** için düzeltilmiş kareler toplamını vermektedir. **1** vektörü ile tanımlanmış bu uzay bir boyutlu olduğu için artık (ortalamadan sapma) vektörü (n-1) boyutlu alt uzayda bulunacak ve dolayısıyla serbestlik derecesi (n-1) olacaktır.

$KT(R)$ ve regresyon parametre tahminleri, her biri kendi ortalamasına göre düzeltilmiş bağımsız değişkenlerin tanımladığı k=p-1 boyutlu alt uzay üzerine kendi artıklarının izdüşümü gerçekleştirildikten sonra elde edilir. Görüldüğü gibi $KT(R)$ iki aşamada elde edilmektedir. İlk olarak **y** ve bağımsız değişkenlerin her biri için **1** vektörü ile tanımlanmış uzay üzerine izdüşümü gerçekleştirilir. Daha sonra, **y** için elde edilen sapma vektörünün, bağımsız değişkenler için elde edilen sapma vektörleri ile tanımlanmış uzay üzerine izdüşümü yapılır. Bu ikinci izdüşüm için $\hat{\mathbf{y}}$'nin karesel uzaklığı, $KT(R)$ değerini verir.

Geometrik olarak her bir kareler toplamının serbestlik derecesi vektörün serbest hareketini tanımlayan boyut sayısıdır. **y** vektörü n-boyutlu uzayda herhangi bir yerde bulunabilir bunun için bir kısıt yoktur bu nedenle serbestlik derecesi n'dir. $\hat{\mathbf{y}}$ vektörü ise sadece X-uzayında bulunabileceği için serbestlik derecesi X-uzayının boyutuna eşittir. Hata terimi vektörü **e** ise X-uzayına dik olacak şekilde n-boyutlu uzayın alt uzayında herhangi bir yere düşebilir. Bu alt uzayın boyutu (n-p) olduğundan **e** vektörünün serbestlik derecesi (n-p)'dir. Genel olarak $\hat{\mathbf{y}}$ ve **e**'nin serbestlik dereceleri sırasıyla $\rho(\mathbf{X})$ ve $[n-\rho(\mathbf{X})]$ olarak verilebilir.

5.4 İSTATİSTİKSEL MODEL

İstatistiksel test uygulamalarında, tüm mümkün çıktıların kaynağı olan anakütle ile ilgili varsayımların seti olan bir matematiksel model kullanılır. Anakütle tüm mümkün gözlenmiş **y** vektörlerinden oluşur. Bu durum Şekil 5.8'de gösterilmiştir.



Şekil 5.8 $\mathbf{x}_0$ *ve* $\mathbf{x}_1$*'nin ortogonal olduğu varsayımı ile* **y**, $\hat{\mathbf{y}}$ *ve* b_0*'ın dağılımları*

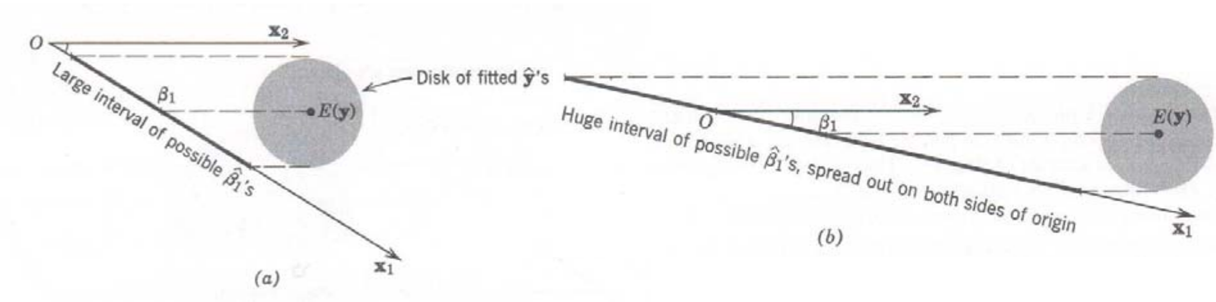
Eğer hataların normal dağıldığı kabul edilirse, bu *y'*lerin sınırsız bir bulut içinde dağılım göstereceği gözlenebilir, gözlemlerin yoğunluğu $E(\mathbf{y})$ çevresinde olup uzaklaştıkça azalır. Geometrik kolaylığı sağlamak açısından gözlenmiş *y'*lerin pek çoğunun sınırlandığı elipsoid çizmek gereklidir. Bu elipsoid yoğunlaşma elipsoidi olarak adlandırılır. Birbirinden bağımsız olarak belirlenmiş hata terimleri için bu elipsoid bir küredir. Bu **y** gözlemlerinin (vektörler, noktalar) oluşturduğu küre $E(\mathbf{y})$ ortalamasında merkezlenmiştir. $E(\mathbf{y})$ noktası $\mathbf{x}_0$ ve $\mathbf{x}_1$ vektörlerinin oluşturduğu *P* düzleminde bulunur. Şekil 5.8'da verilen örnekte görüldüğü gibi gözlenmiş **y** değerleri önemli hatalar içermektedir. Bu nedenle **y** değerleri $E(\mathbf{y})$'den farklıdır.

EKK tahminlenme yöntemi, *P* düzlemi üzerine gözlemlenmiş **y** vektörünün ortogonal izdüşümünden oluşur, $\hat{\mathbf{y}}$ sonucu $E(\mathbf{y})$'nin tahminini oluşturur. Anakütle parametresi β_0'ın tahminleyicisi olan b_0'ı elde etmek için $\hat{\mathbf{y}}$'nin $\mathbf{x}_0$ üzerine $\mathbf{x}_1$ boyunca izdüşümü alınır, benzer şekilde b_1'in elde edilmesi için de $\hat{\mathbf{y}}$'nin $\mathbf{x}_0$ üzerine $\mathbf{x}_1$ boyunca izdüşümü elde edilir. Gözlemlenmiş **y** vektöründeki hatalar nedeniyle b_0 değerinin β_0'dan daha küçük olması beklenir.

Şekil 5.8'den görüldüğü gibi küre içinde bulunan mümkün gözlenmiş **y** değerlerine karşılık gelen uyumu yapılmış $\hat{\mathbf{y}}$ değerleri *P* düzlemindeki daire ile temsil edilmektedir. Bu daire $E(\mathbf{y})$ çevresinde oluşturulmuş olup $E(\mathbf{y})$ sabittir. Bu diskin tamamının $\mathbf{x}_0$ üzerine $\mathbf{x}_1$ boyunca izdüşümünün alınmasıyla (*a b*) aralığı elde edilir. Elde edilen (*a b*) aralığı, sabitlenmiş gerçek β_0 çevresindeki b_0'ın aralığını ifade eder. Gözlenmiş **y**'ler, $E(\mathbf{y})$'de merkezlenmiş olan küre içerisinde normal dağıldıkları için, b_0'ın örnekleme dağılımının da sapmasız ve normal olması beklenir.

5.5 ORTOGONAL OLMAYAN DEĞİŞKENLER

Önceki kısımda birbirine ortogonal olan $\mathbf{x}_0$ ve $\mathbf{x}_1$ değişkenleri ele alınmıştı. Buna karşın bu değişkenlerin ortogonal olmadığı durumlarla sık sık karşılaşılır. Bu durum Şekil 5.9'da gösterilmiştir. Değişkenlerin birbirine ortogonal olmamaları durumunda $\mathbf{x}_0$ üzerine $\hat{\mathbf{y}}$ diskinin ortogonal olmayan (çarpık) izdüşümü b_0 ın aralığını verir.



Şekil 5.9 Şekil 5.8'*deki P düzlemi*: a) $\mathbf{x}_0$ *ve* $\mathbf{x}_1$ *ortogonal değildir. Buna göre* $\hat{\mathbf{y}}$ ve b_0*'ın dağılımları*. b) $\mathbf{x}_0$ *ve* $\mathbf{x}_1$ *arasında büyük ölçüde çoklu doğrusal bağlantı olduğu durum*.

Şekil 5.9*b*'de gösterildiği gibi $\mathbf{x}_0$ ve $\mathbf{x}_1$ vektörleri arasındaki açı küçüldükçe (çoklu doğrusal bağlantı arttıkça) sorunda artacaktır. Şekilden görüldüğü gibi b_0'ın aralığı orijinin her iki tarafına yayılmış durumdadır.β_0'ın nokta tahmini pozitif veya negatif olabilir. Bununla birlikte Şekil 5.9*b*'den β_0 parametresinin sıfır olmadığı görülmektedir. Fakat istatiksel olarak $H_{0:}$ β_0=0 hipotezini genellikle reddedilemeyebilecektir. Çünkü β_0 'ın standart hatası oldukça büyüktür.

İki vektör arasındaki açının küçülmesi her ne kadar b_0'ın yayılma aralığının büyük olmasına neden olsa da, normallik ve sapmasızlık gibi özellikleri etkilemez.

5.6 KORELASYON VE COS θ

Kısım 5.3'da eşitlik (5.12) ile gözlemlerin ortalamadan sapmaları elde edilmiştir. Eğer ortalamadan farklar şeklinde verilmiş **x** ve **y** gibi iki vektör ele alınırsa, bu iki vektörün birbirlerine olan yakınlığının ölçümü tanımlanabilir. İki vektörün yönlerinin yakınlığı için standart geometrik ölçüm,

$$Cos\theta = \frac{\mathbf{x}\mathbf{y}}{\|\mathbf{x}\|\|\mathbf{y}\|} \qquad (5.15a)$$

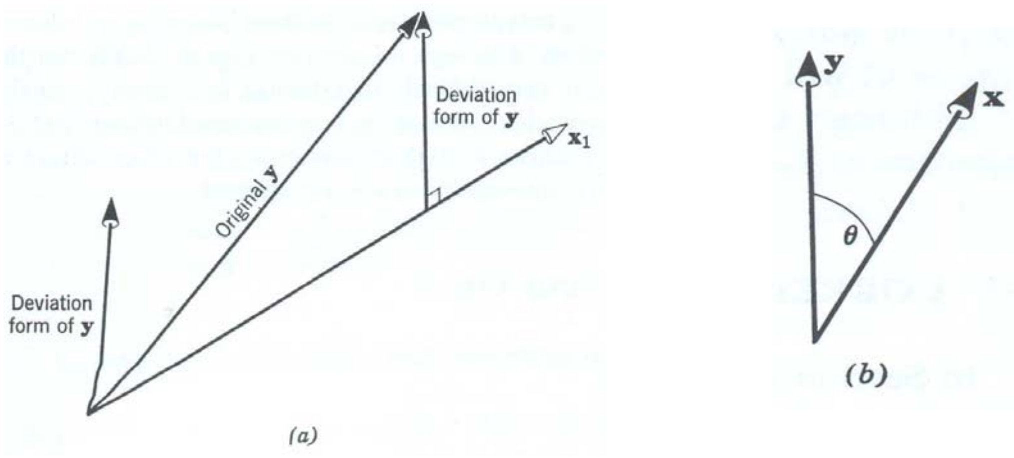
eşitliği ile verilebilir. Bu ifade bileşenlerin nokta çarpımlarına göre açık olarak,

$$Cos\theta = \frac{\sum x_i y_i}{\sqrt{\sum x_i^2}\sqrt{\sum y_i^2}} \qquad (5.15b)$$

yazılabilir. x_i ve y_i değerleri ortalamadan sapmaları belirttiği için eşitlik (5.15*b*)'nin Bölüm 2'de eşitlik (2.25*e*) ile verilen korelasyon katsayısını ifade ettiği görülebilir. Bunun sonucu olarak,

$$r = Cos\theta \qquad (5.16)$$

ifadesi verilebilir. Bu eşitlikteki θ, **x** ve **y** vektörleri arasındaki açıyı belirtmektedir. Başka bir deyişle Şekil 5.10'de gösterildiği gibi korelasyonun geometrik yorumu θ açısının yakınlığı olarak düşünülebilir.



Şekil 5.10 *a*) **y** *vektörü ile sapmalar formundaki vektörünün ilişkisi b*) *Korelasyon= Cos θ*, **x** *ve* **y** *sapmalar formunda*.

Bu durumda *Cos θ* ile ilgili olarak her geometrik ifade için, *r* ile ilgili eş değer bir istatistiksel ifadenin mevcut olduğu söylenebilir. Bu ilişkiyi belirten birkaç örnek Tablo 5.1'de verilmiştir.

Tablo 5.1 *Korelasyon katsayısı r'nin istatiksel yorumu ile θ'*nın *geometrik yorumunun karşılaştırılması* (*Tüm değişkenler ortalamadan sapmalar formundadır.*)

Geometri	İstatistik
Cos θ	*r*
-1≤ *Cos θ* ≤1	-1≤ *rθ* ≤1
Cos θ=+1 **x** ve **y** vektörleri çakışık aynı yönde	*r*=+1 *x* ve *y* arasında tam pozitif doğrusal ilişki
Cos θ=-1 **x** ve **y** vektörleri tamamen zıt yönde	*r*=-1 *x* ve *y* arasında tam negatif doğrusal ilişki
Cos θ=0 **x** ve **y** vektörleri ortogonal	*r*=0 *x* ve *y* arasında doğrusal ilişki yok
y vektörünün karesel uzunluğu	*y*'nin değişkenliği
y vektörünün uzunluğu	*y*'nin standart sapması $(n\text{-}1)^{1/2}$ böleni hariç
(**y**-**ŷ**) vektörünün uzunluğunu minimum yapan ortogonal izdüşüm	EKK uyumu
Pisagor teoremi $\|\mathbf{y}-\hat{\mathbf{y}}_0\| = \|b_1\mathbf{x}_1\|^2 + \|\mathbf{y}-\hat{\mathbf{y}}\|^2$	Varyans Analizi ToplamDeğişkenlik=Açıklanmış değişkenlik+ Açıklanamayan

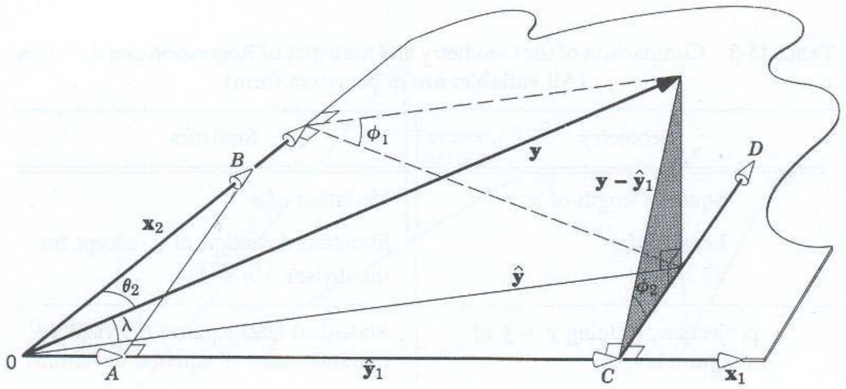
	değişkenlik

EKK nın geometrisi ile ilgili bazı konular ileriki bölümlerde açıklanacak ilgili kısımların sonunda verilecektir.

5.7 BASİT ÇOKLU VE KISMİ KORELASYONLAR

Bu kısımda ele alınan vektörler yine ortalamadan sapmalar formunda olup, $\mathbf{x}_1$ ve $\mathbf{x}_2$ vektörlerinin ortogonal olmadıkları varsayılmaktadır. Model ortalamadan sapmalara göre ifade edildiğinde sabit terim ortadan kalktığı için model artık $\mathbf{x}_0$ sabit vektörünü içermez.

Şekil 5.11′den r_{yx_2} korelasyonun $Cos\theta_2$ olduğu görülmektedir. Bu ifadenin çoklu ve kısmi korelasyondan ayırt edilmesi için genellikle basit veya sıradan korelasyon olarak adlandırılır.



Şekil 5.11 Çoklu korelasyon katsayısının R=Cosλ, basit $\left(r_{yz_2} = Cos\theta_2\right)$ ve kısmi $\left(r_{yx_2 \cdot x_1} = Cos\phi_2\right)$ korelasyon ile karşılaştırılması.

Çoklu korelasyon katsayısı R, gözlenmiş $\mathbf{y}$ ve tahminlenmiş $\hat{\mathbf{y}}$ vektörleri arasındaki basit korelasyon olarak tanımlanabilir. Şekil 5.11 de verilen Cosλ açısı bu korelasyonu belirtmektedir. Böylece $\mathbf{x}_1$ ve $\mathbf{x}_2$ değişkenlerinin $\mathbf{y}$ bağımlı değişkenini ne kadar açıklayabildiğini belirten bir indeks elde edilir. Bağımlı değişken $\mathbf{y}$ ve bağımsız değişken $\mathbf{x}_2$ arasındaki kısmi korelasyon $r_{yx_2 \cdot x_1}$, $\mathbf{x}_1$ değişkeninin her bir değişken üzerindeki etkisi ortadan kaldırıldıktan sonra elde edilen $\mathbf{x}_2$ ve $\mathbf{y}$ arasındaki basit korelasyondur. $\mathbf{x}_1$ üzerine $\mathbf{y}$ nin regresyonu elde edildikten sonra $\mathbf{y}$ üzerindeki $\mathbf{x}_1$ etkisi uyumu yapılmış $\hat{\mathbf{y}}$ değeri olarak elde edilir. Bu etki ortadan kaldırıldığında veya $\mathbf{y}$ den çıkarıldığında sonuç olarak $\left(\mathbf{y}-\hat{\mathbf{y}}_1\right)$ artık vektörü elde edilir. Benzer olarak $\mathbf{x}_2$ değişkeninin $\mathbf{x}_1$ üzerine regresyonu (A noktasında) gerçekleştirilip, $\mathbf{x}_1$′in etkisi $\mathbf{x}_2$ üzerinden kaldırıldığında AB vektörü elde edilir. Bu vektörün $\mathbf{y}$ vektörü ile arasındaki açı $Cos\phi_2$ olup $r_{yx_2 \cdot x_1}$ kısmi korelasyonunu vermektedir. Benzer olarak $Cos\phi_1$ açısı da $r_{yx_1 \cdot x_2}$ kısmi korelasyonunu belirtir. Tablo 5.2′de ise kısmi ve çoklu korelasyon istatiksel ve geometrik açıdan karşılaştırılmaktadır.

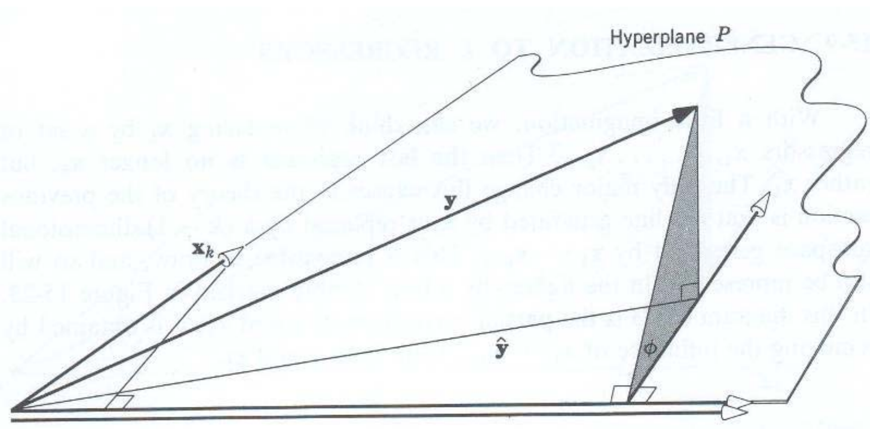
Tablo 5.2. Basit, kısmi ve çoklu korelasyon istatistik ve geometrik karşılaştırılması (bkz. Şekil 5.11)

Geometri	İstatistik
$Cos\theta_2$	Basit korelasyon r_{yx_2}
$Cos\phi_2$	$r_{yx_2x_1}$
$Cos\lambda$	Çoklu korelasyon R
$Cos\lambda = 1$ y değişkeni (x_1, x_2) alt uzayındadır.	R=1 x_1 ve x_2 değişkenleri tamamen açıklanır.
$Cos\lambda = 0$ y değişkeni (x_1, x_2) alt uzayına ortogonaldir.	R=0 $\hat{y} = 0x_1 + 0x_2$=0 x_1 ve x_2 arasında ilişki yok
$\lvert Cos\theta_2 \rvert \leq \lvert Cos\lambda \rvert$ $\lvert Cos\phi_2 \rvert \leq \lvert Cos\lambda \rvert$	$r_{yx_2} \leq R$ $r_{yx_2x_1} \leq R$

X_1 bağımsız değişkeni yerine $X_1, X_2, ..., X_{k-1}$ değişkenler seti alınsın, bu durumda x_2 değişkeni artık x_k değişkeni olarak tanımlanacaktır. Bağımsız değişkenlerden x_1 değişkeni $x_1, x_2, .., x_{k-1}$ ile belirtilen $(k-1)$ boyutlu bir alt uzayla yer değiştirmiştir. Bu alt uzayın ($k-1>3$ olması durumunda) çizilmesi imkansızdır. Şekil 5.12′de bu alt uzay çift çizgili eksen ile belirtilmiştir. Şekildeki $Cos\phi$ açısı x_k ile y arasındaki kısmi korelasyondur. Bu korelasyon için, $x_1, x_2, .., x_{k-1}$ ′in etkisi hem y hem de x_k ′dan çıkarılmıştır.

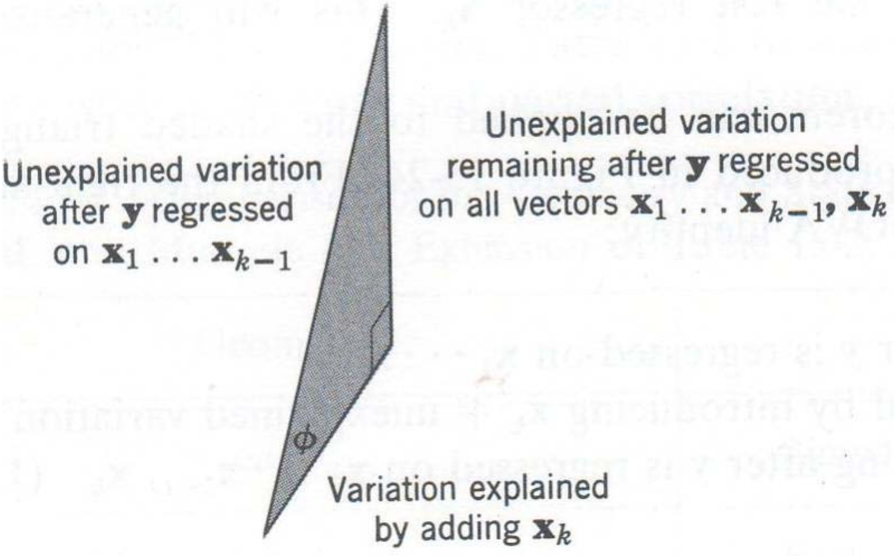
5.8 SON DEĞİŞKENİN TESTİ İÇİN V.A.T.

Şekil 5.12′de son değişken x_k ′ın istatistiksel açıdan önemli olup olmadığının testi varyans analizi ile de gösterilebilir.



Şekil 5.12 *Kısmi Korelasyon, Şekil 5.11 ile karşılaştır.*

Pisagor teoremi Şekil 5.12′de verilen gölgeli üçgene uygulanabilir. Şekil 5.13′de tekrar çizilmiş olan bu üçgen incelenerek aşağıda verilen V.A.T özdeşliği elde edilebilir.



Şekil 5.13 Pisagor teoremi ve varyans analizi

$$\frac{x_k\text{ 'nın modele alınması ile açıklanan değişkenlik}}{x_1, x_2, .., x_{k-1} \text{ üzerine } y\text{ 'nin regresyonundan sonra açıklanamayan değişkenlik}} \tag{5.17}$$

$x_1, x_2, .., x_{k-1}$ üzerine y 'nin regresyonundan sonra açıklanamayan değişkenlik

$= x_k$ 'nın modele alınması ile açıklanan değişkenlik (5.17)

$x_1, x_2, .., x_{k-1}$ üzerine y 'nin regresyonundan sonra açıklanamayan değişkenlik

Eşitlik (5.17)′un sağ tarafındaki iki değişkenlik, istatistiksel olarak bağımsız X^2 değişkenleridir. Serbestlik dereceleri de sırasıyla 1 ve (n-1-k) ′dır. Bu durumunda, F oranı

$$F = \frac{x_k\text{'nıı mod}\,ele\,alılınması\ açıçıklan\ değeğişkek\ /1}{Hata\,terimi\,değeğişkeği/(n-1-k)} \tag{5.18}$$

şeklinde elde edilir ve bu oran modele en son olarak alınan x_k değişkeninin istatistiksel öneminin testi için kullanılabilir.

Modele en son olarak alınan x_k değişkeni (5.18)′in karekökü alınarak t testi elde edilebilir.

$$t = \pm\sqrt{F} = \pm\sqrt{\frac{x_k\text{'}nıı\ ilavesiyle\,açıçıklana\ değe}{Hata\,terimi\,değeğişkeği/(n-1-k)}} \tag{5.19}$$

Elde edilen t istatistiği $n-1-k$ serbestlik dereceli olup işareti b_k regresyon katsayısı ile aynı olacak şekilde seçilir.

Alternatif olarak eşitlik (5.18)′de verilen F-testi r_{yx_k}, $x_1, x_2, .., x_{k-1}$ kısmi korelasyonuna göre de ifade edilebilir. Kısmi korelasyon r olarak kısaltılarak,

$$r = Cos\phi \tag{5.20}$$

yazılabilir. Şekil 5.13′den eşitlik (5.18)

$$F = (Cot\phi)^2 (n-1-k) \tag{5.21}$$

ve trigonometriden,

$$(Cot\phi)^2 = \frac{(Cos\phi)^2}{1-(Cos\phi)^2} \tag{5.22}$$

$$= \frac{r^2}{1-r^2} \tag{5.23}$$

elde edilip sonuç olarak *F*-oranı

$$F = \frac{r^2(n-1-k)}{1-r^2} \tag{5.24}$$

şeklinde verilebilir.

Eğer x_k değişkeninin ilavesi **y** nin açıklanan kısmına şansa bağlı olarak nitelenebilecek miktarlardan daha fazla katkı yapıyor ise **y** vektörü $X_1, X_2, ..., X_{k-1}$ ile oluşturulan P hiper düzlemine yaklaşacaktır. Bunun sonucu olarak ϕ açısı küçülerek r değerinin büyütecek ve dolayısı ile eşitlik (5.24)′daki *F* değeri büyüyecek ve sonuç olarak x_k′nın istatistiksel olarak anlamlı olacaktır.

Eşitlik (5.24) kullanılarak eşitlik (5.19)

$$t = \pm\sqrt{\frac{r^2(n-1-k)}{1-r^2}} \tag{5.25}$$

olarak yazılabilir.