T.C.

ONDOKUZ MAYIS ÜNİVERSİTESİ

FEN EDEBİYAT FAKÜLTESİ
iSTATİSTİK BÖLÜMÜ

ÇOK DEĞİŞKENLİ İSTATİSTİK I DERSİ

PROF. DR. YÜKSEL ÖNER

1. Hafta

özenilen üniversite

BÖLÜM I

I.1 GİRİŞ

Bir istatistiksel olayı incelerken bu olay üzerinde değişken adını vereceğimiz pek çok özellik tanımlamak mümkündür. Bu değişkenlere göre söz konusu istatistiksel olay her bir değişken tek tek ele alınarak inceleneceği gibi hepsi birlikte eşanlı (aynı anda) olarak da incelenebilir. Değişkenlerin tek tek incelenmesi tek değişkenli istatistiğin konusudur. Ancak; bu durum istatistiksel olay hakkında verilecek olan karar için hata olasılığının büyümesine yol açabileceği gibi zaman ve maliyet israfına da neden olabilecektir. İstatistiksel olayın tüm değişkenler birlikte ele alınarak incelenmesi halinde ise bahsedilen bu olumsuzluklar ortadan kaldırılabilir. Bu tarz bir inceleme ise çok değişkenli istatistiğin kapsamı içerisine girmektedir. Çünkü çok değişkenli istatistiksel analizlerde birbirleriyle ilişkili çok sayıda değişken söz konusudur.

İstatistik biliminin genelinde olduğu gibi çok değişkenli istatistiksel tekniklerin esas amacı, sayı ile ifade edilen sonuçların gözlemlenmesi, düzenlenmesi, özetlenmesi, karar aşamasında bilgi olarak kullanılması ve yorumlanmasıdır. Çok değişkenli istatistik birden fazla özelliğin analizi ile ilgilendiğinden, uygulamalarda farklı amaçlar için kullanılabilir. Bu amaçlardan bazıları:

1) Basitleştirme ve Boyut İndirgeme: Çok sayıda değişken tarafından belirlenen bir sistemin daha az sayıda yapay değişkenlerle temsil edilerek basitlik sağlanması için kullanılabilir (Temel Bileşen Analizi, Faktör Analizi).

2) Birimlerin Sınıflandırılması: Gözlenen birimlerin değişik sınıflar ya da gruplar oluşturup oluşturmadığının belirlenmeye çalışılması veya birimlerin önceden tanımlanmış sınıflardan herhangi birisine ait olup olmadığının belirlenmesi için kullanılabilir (Kümeleme Analizi, Diskriminant Analizi)

3) Bağımlılık Yapısının İncelenmesi: Değişkenler arasındaki kovaryans ya da korelasyonlardan faydalanarak bağımlılığın kaynaklarının araştırılması ve sonuçların incelenmesi ve değerlendirilmesi için kullanılabilir (Çok Değişkenli Regresyon Analizi, Çok Değişkenli Varyans Analizi, Kanonik Korelasyon Analizi)

4) Hipotez Oluşturma ve Hipotez Testi: İnceleme konusu olaylara ait hipotezler oluşturarak bu hipotezlerin test edilebilmesi için gerekli test istatistiklerini belirlemek için kullanılabilir. Çok

değişkenli testler genel anlamda tek değişkenli istatistiklerle ilgili hipotezlerin test edilmesinde kullanılan test istatistiklerinin bir genelleştirmesidir.

5) Sıralama ve Ölçekleme: Bazı araştırmalarda birimlerin belli ölçülere göre sıralanması istenebilir. Ölçekleme ise çok sayıda değişkenden yararlanarak birimlerin daha az boyutta gösterilmesidir (Uygunluk Analizi, Çok Boyutlu Ölçekleme). Gerek uygunluk analizinde gerekse çok boyutlu ölçeklemede grafik gösteriminden yararlanarak birimlerim karşılaştırılması, yakınlık yada uzaklıklarının belirlenmesi sağlanır.

I.2 MATRİS KAVRAMLARI ve ÖZELLİKLERİ

Çok değişkenli istatistiksel yöntemlerin hemen hemen hepsinde, değişkenlerin matrisler şeklinde düzenlenmiş özet bilgileri kullanılarak çözüme ulaşılmaktadır. Bu yönü ile çok değişkenli analizlerde matris kavramı önemli yer tutmaktadır. Ayrıca bazı teorik bilgilerin daha iyi anlaşılabilmesi ve daha hızlı kavranabilmesi için matrislere ait bazı temel kavramların ve özelliklerin bilinmesinde önemli yararlar vardır.

Tanım I.1 Elemanları reel ya da kompleks sayılardan oluşan satır ve sütunlara sahip dörtgensel bir sisteme matris denir ve $A = \left[a_{ij}\right]_{n\times p}, (i = 1, 2, \dots n; j = 1, 2, \dots, p)$ ile gösterilir. Burada n: satır ayısı ve p: sütun (kolon) sayısı iken a_{ij}: A matrisinin i-nci satır, j-nci sütundaki elemanıdır.

Örnek I.1 Ondokuz Mayıs Üniversitesi Fen Fakültesi'ne ait bazı bölümlere 2017-2018, 2018-2019 ve 2019-2020 eğitim öğretim yıllarında kayıt yaptıran öğrencilerin dağılımı Tablo I.1'deki gibi olsun.

Tablo I.1 OMÜ Fen Fakültesi'ne Ait Bazı Bölümlere Kayıt Yaptıran Öğrencilerin Dağılımı

	Eğitim Öğretim Yılları		
Bölüm	**2017-2018**	**2018-2019**	**2019-2020**
İstatistik	32	35	40
Matematik	35	42	52
Fizik	15	16	20
Kimya	36	40	45
Biyoloji	16	15	20
M. Biyoloji	40	45	50

Bu tabloda yer alan bilgiler matris şeklinde bir A matrisi olarak adlandırılıp yazılacak olursa;

$$A = \begin{bmatrix} 32 & 35 & 40 \\ 35 & 42 & 52 \\ 15 & 16 & 20 \\ 36 & 40 & 45 \\ 16 & 15 & 20 \\ 40 & 45 & 50 \end{bmatrix}$$

şeklinde ifade edilir.

Tanım I.1'e göre:

i) $p = n$ ise A bir kare matristir.

ii) $p = 1$ ise A bir kolon vektörüdür ve $A = \underline{a} : n \times 1$ şeklinde gösterilir.

iii) $n = 1$ ise A bir satır vektörüdür ve $A = \underline{a} : 1 \times p$ şeklinde gösterilir.

iv) $\forall(i\,,j)$ için $a_{ij} = 0$ ise A bir sıfır matrisidir.

v) $n = p$ olmak üzere $i = j$ iken $a_{ij} = 1$ ve $i \neq j$ iken $a_{ij} = 0$ ise A bir birim matristir ($A = I_p\,, A = I_n$).

vi)) $n = p$ olmak üzere $i \neq j$ iken $a_{ij} = 0$ ve $i = j$ iken $a_{ij} \notin \{0\,,1\}$ ise A bir köşegen matristir ve $A = K\ddot{o}ş\left[a_{11},\ a_{22}, \ldots ,\ a_{pp}\ \right]$ şeklinde gösterilir.

vii) Bir A matrisinin satır ve sütunlarının yer değiştirmesiyle elde edilen matrise A'nın transpoz matrisi denir ve A' ile gösterilir.

viii) $A: p \times p$ bir matris olsun. $A = A'$ ise A bir simetrik matristir, eğer . $A = -A'$ ise A bir ters simetrik matristir.

ix) $A: p \times p$ bir matris olsun. $AA' = A'A = I_p$ oluyorsa veya

a) $\sum_{i=1}^{p} a_{ij}a_{ik} = 0, (j \neq k = 1, 2, \ldots, p)$, $\left[\text{veya } \sum_{j=1}^{p} a_{ij}a_{kj} = 0,\ \ (i \neq k = 1, 2, \ldots, p)\right]$

b) $\sum_{i=1}^{p} a_{ij}^2 = 1\,, (j = 1, 2, \ldots, p)$, $\left[\text{veya } \sum_{j=1}^{p} a_{ij}^2 = 1\,,\ (i = 1, 2, \ldots, p)\right]$

oluyorsa, A bir ortogonal (dik) matristir.

x) $A: p \times p$ bir matris olsun. Bu takdirde $|A| = det(A) = \sum_{i=1}^{p} a_{ij}A_{ij} = \sum_{j=1}^{p} a_{ij}A_{ij}$ ifadesinden elde edilen reel sayıya A matrisinin determinantı denir. Burada a_{ij}: A matrisinin $(i\,,j)$ elemanı iken, A_{ij} bu elemana karşılık gelen kofaktördür. A_{ij} kofaktörü; $A_{ij} = (-1)^{i+j}\left|M_{ij}\right|$, $(i,j = 1, 2, \ldots, p)$ eşitliği ile hesaplanır. Bu eşitlikte yer alan $\left|M_{ij}\right|$ ifadesi, A matrisinin i-nci satır ve j-nci sütunu çıkartıldıktan sonra geriye kalan alt matrisin determinantı olup, a_{ij} elemanının Minörü adını alır.

xi) $A\colon p \times p$ bir matris ve $|A| \neq 0$ olsun. Bu takdirde $B = \frac{1}{|A|}[Adj(A)]$ matrisine A matrisinin tersi denir ve $A^{-1} = B$ ile gösterilir. Burada $A_K = \left[A_{ij}\right]_{p\times p}$, A matrisinin kofaktörler matrisi olmak üzere $Adj(A) = A'_K$ şeklinde tanımlı olup A matrisinin adjoint matrisi olarak bilinir. Eğer A matrisi simetrik bir matris ise o zaman $Adj(A) = A_K$ dır, yani adjoint matris ile kofaktörler matrisi birbirine eşittir.

xii) $A\colon p \times p$ bir matris olsun. $\dot{I}z(A) = \sum_{j=1}^{p} a_{jj}$ ifadesine A matrisinin izi denir.

xiii) $A\colon p \times p$ bir matris ve λ bir skaler parametre olmak üzere $\left|A - \lambda I_p\right| = 0$ polinomunun kökleri olan $\lambda_j, (j = 1, 2, \ldots, p)$ değerlerine A matrisinin özdeğerleri denir. Ayrıca; $j = 1, 2, \ldots, p$ için $\left(A - \lambda_j I_p\right)\underline{x}_j = \underline{0}$ homojen lineer denklem sisteminin çözüm vektör olan $\underline{x}_j\colon p \times 1$ vektörüne, λ_j özdeğerine karşılık gelen özvektör denir. Burada $\left|A - \lambda_j I_p\right| = 0$ olduğundan bulunacak olan $\underline{x}_j$ çözüm vektörleri parametreye bağlı olacağından tek değildir. Yani λ_j özdeğerine sonsuz tane $\underline{x}_j$ özvektörü karşılık gelebilir. Bu vektörler normalleştirildiğinde (birim normal vektöre dönüştürüldüğünde) bir tek birim normal özvektörü gösterir.

$\underline{e}_j$: λ_j özdeğerine karşılık gelen birim normal özvektör olsun. Bu takdirde bu birim normal özvektör;

$$\underline{e}_j = \frac{1}{\sqrt{\underline{x}'_j \underline{x}_j}}\underline{x}_j \quad , \; j = 1, 2, \ldots, p$$

eşitliği ile bulunur. Bu özvektörlere standartlaştırılmış asıl özvektörler ya da standartlaştırılmış asıl temel bileşenler adı verilir.

$\Lambda = K\ddot{o}\c{s}\left[\lambda_1, \lambda_2, \ldots, \lambda_p\right]_{p\times p}$ …… A matrisinin özdeğerler matrisi

$P = \left[\underline{e}_1 \quad \underline{e}_2 \quad \ldots \quad \underline{e}_p\right]_{p\times p}$ …… A matrisinin özvektörler matrisi (bir ortogonal matristir)

olup, burada $\underline{e}'_j \underline{e}_k = \begin{cases} 0, & j \neq k \\ 1, & j = k \end{cases}$ özelliği sağlanır. Bu özelliğe göre özvektörler birbirine dik olan vektörlerdir. Bu sebeple P özvektörler matrisi bir ortogonal matristir.

Buna karşılık bazı istatistiksel paket programları (örneğin SPSS) , $j = 1, 2, \ldots, p$ için $\underline{e}^*_j = \sqrt{\lambda_j}\underline{e}_j$ eşitliği ile verilen özvektörleri çıktı olarak vermektedir. Bu özvektörlere standartlaştırılmış özvektörler (veya standartlaştırılmış temel bileşenler ya da faktör yükleri) adı verilmektedir. Bu takdirde özvektörler matrisi $P = \left[\underline{e}^*_1 \quad \underline{e}^*_2 \quad \ldots \quad \underline{e}^*_p\right]_{p\times p}$ ile gösterilir.

xiv) $A: p \times p$ bir matris, $(\lambda_j, \underline{e}_j)$, $j = 1, 2, \dots, p$ ikilisi A matrisinin özdeğer-özvektör çiftleri olmak üzere,

$A = P\Lambda P' = \sum_{j=1}^{p} \lambda_j \underline{e}_j \underline{e}'_j$

ifadesine A matrisinin spectral ayrışımı denir.

xv) $A: p \times p$ bir matris, $(\lambda_j, \underline{e}_j)$, $j = 1, 2, \dots, p$ ikilisi A matrisinin özdeğer-özvektör çiftleri olsun. $\Lambda^{1/2} = Köş\left[\sqrt{\lambda_1}, \sqrt{\lambda_2}, \dots, \sqrt{\lambda_p}\right]$ olmak üzere $A^{1/2} = P\Lambda^{1/2}P' = \sum_{j=1}^{p} \sqrt{\lambda_j}\, \underline{e}_j \underline{e}'_j$ ifadesine A matrisinin karekök matrisi denir.

xvi) $A: p \times p$ bir matris ve $\underline{x}' = \begin{bmatrix} x_1 & x_2 & \dots & x_p \end{bmatrix} : 1 \times p$ boyutlu bilinmeyenler vektörü olsun. Bu takdirde;

$Q(\underline{x}) = \underline{x}' A \underline{x} = \sum_{i=1}^{p} \sum_{j=1}^{p} a_{ij} x_i x_j$

ifadesine bir karesel form denir.

- ✓ Bir karesel formu tanımlayan A matrisi bir simetrik matristir. Eğer A matrisi simetrik değilse, $B = (A + A')/2$ eşitliği ile elde edilen B simetrik matrisi yardımıyla $Q(\underline{x}) = \underline{x}' B \underline{x}$ karesel formu yazılabilir.
- ✓ $\forall \underline{x} \neq \underline{0}$ için $Q(\underline{x}) = \underline{x}' A \underline{x} > 0$ ise karesel forma pozitif tanımlı karesel form ve A matrisine de pozitif tanımlı matris denir.
- ✓ Eğer $\exists \underline{x} \neq \underline{0}$ için $Q(\underline{x}) = \underline{x}' A \underline{x} \geq 0$ ise karesel forma pozitif yarı tanımlı karesel form ve A matrisine ise pozitif yarı tanımlı matris adı verilir.

xvii) $A: p \times p$ bir matris olsun. $k < p$ olmak üzere A matrisinin; $A_{11}: k \times k$, $A_{12}: k \times (p-k)$, $A_{21}: (p-k) \times k$ ve $A_{22}: (p-k) \times (p-k)$ şeklinde alt matrislere ayrılmasına A matrisinin bir parçalanması denir ve

$A = \begin{bmatrix} A_{11} \vdots A_{12} \\ \cdots\cdots\cdots\cdots \\ A_{21} \vdots A_{22} \end{bmatrix}$ ile gösterilir.

xviii) $A: p \times p$ bir matris olsun. Eğer $A^2 = A$ özelliği sağlanıyorsa A matrisine idempotent matris denir.

Algoritma I.1 Birim normal özvektörlerin bulunması

i) $j = 1, 2, \dots, p$ olmak üzere her bir λ_j özdeğeri için $A - \lambda_j I_p$ matrisi bulunur.

ii) $Adj(A - \lambda_j I_p)$ matrisi elde edilir. $Adj(A - \lambda_j I_p) = \left[Kof(A - \lambda_j I_p)\right]'$ olup, elde edilen bu matrisin kolonları birbiri ile bağıntılıdır, yani lineer bağımlıdır.

iii) (ii)'de elde dilen matrisin her kolonunun elemanları, ilgili kolondaki elemanların kareleri toplamının kareköküne bölünerek λ_j özdeğerine karşılık gelen $\underline{e}_j$ özvektörü bulunur.

Özellikler

1) $(A')' = A$, $(A \pm B)' = A' \pm B'$, $(AB)' = B'A'$, $(kA)' = kA'$

2) $|A| = |A'|$, $|A^{-1}| = |A|^{-1} = \frac{1}{|A|}$, $(|A| \neq 0)$, $(A^{-1})' = (A')^{-1}$, $(AB)^{-1} = B^{-1}A^{-1}$

3) $|A| = 0$ ise A singüler bir matristir.

4) k bir sabit ve $A: p \times p$ bir matris iken $|kA| = k^p|A|$ olur.

5) $A_1, A_2, \dots, A_k: p \times p$ boyutlu matrisler ise $|A_1 A_2 \dots A_k| = |A_1||A_2| \dots |A_k|$ dır.

6) A bir ortogonal matris ise $|A| = \pm 1$ ve $A^{-1} = A'$

7) $A: p \times p$ bir köşegen matris ise $|A| = \prod_{j=1}^{p} a_{jj}$, $\dot{I}z(A) = \sum_{j=1}^{p} a_{jj}$ ve

$A^{-1} = K\ddot{o}ş\left[\frac{1}{a_{11}}, \frac{1}{a_{22}}, \dots, \frac{1}{a_{pp}}\right]$ dir.

8) $A: p \times p$ matrisinin özdeğerleri ve özvektörleri sırası ile λ_j ve $\underline{e}_j$, $(j = 1, 2, \dots, p)$ olsun. Bu takdirde $|A| = \prod_{j=1}^{p} \lambda_j$, $\dot{I}z(A) = \sum_{j=1}^{p} \lambda_j$ ve $A^{-1} = P\Lambda^{-1}P' = \sum_{j=1}^{p} \frac{1}{\lambda_j} \underline{e}_j \underline{e}'_j$ eşitlikleri doğrudur.

9) $A: p \times p$ simetrik bir matris ise A matrisinin özdeğerleri birer reel sayı ve özvektörleri birbirine diktir. Eğer A matrisi pozitif tanımlı ise o zaman özdeğerleri birer pozitif reel sayı olup, daima $\lambda_1 > \lambda_2 > \cdots > \lambda_p \geq 0$ şeklinde sıralanabilir.

10) A ve $B : p \times p$ matris ise $\dot{I}z(A \pm B) = \dot{I}z(A) \pm \dot{I}z(B)$ dir.

11) $A : n \times p$ ve $B: p \times n$ ise $\dot{I}z(AB) = \dot{I}z(BA)$ dır.

12) k bir sabit ve $A: p \times p$ bir matris iken $\dot{I}z(kA) = k\dot{I}z(A)$ dır.

13) A ve $B : p \times p$ matrisler ve $|B| \neq 0$ ise $\dot{I}z(BAB^{-1}) = \dot{I}z(A)$ dır.

14) $A: p \times p$ matrisinin bir parçalanması;

$A = \begin{bmatrix} A_{11}: k \times k \,\vdots\, A_{12}: k \times (p-k) \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ A_{21}: (p-k) \times k \,\vdots\, A_{22}: (p-k) \times (p-k) \end{bmatrix}$ olsun. Bu takdirde $A^{-1} = \begin{bmatrix} B_{11} \,\vdots\, B_{12} \\ \cdots\cdots\cdots \\ B_{21} \,\vdots\, B_{22} \end{bmatrix}$

şeklinde bir parçalanmış matristir. Öyle ki burada,

$B_{11} = (A_{11} - A_{12}A_{22}^{-1}A_{21})^{-1} = [A_{11}^{-1} + A_{11}^{-1}A_{12}B_{22}A_{21}A_{11}^{-1}]: k \times k$

$B_{12} = -A_{11}^{-1}A_{12}[A_{22} - A_{21}A_{11}^{-1}A_{12}]^{-1} = -A_{11}^{-1}A_{12}B_{22}: k \times (p-k)$

$B_{21} = -A_{22}^{-1}A_{21}[A_{11} - A_{12}A_{22}^{-1}A_{21}]^{-1} = -A_{22}^{-1}A_{21}B_{11}: (p-k) \times k$

$B_{22} = (A_{22} - A_{21}A_{11}^{-1}A_{12})^{-1} = [A_{22}^{-1} + A_{22}^{-1}A_{21}B_{ll}A_{12}A_{22}^{-1}]: (p-k) \times (p-k)$ dir.

- ✓ $|A_{11}| \neq 0$ ise $|A| = |A_{11}||A_{22} - A_{21}A_{11}^{-1}A_{12}|$
- ✓ $|A_{22}| \neq 0$ ise $|A| = |A_{22}||A_{11} - A_{12}A_{22}^{-1}A_{21}|$

- ✓ $A_{12} = [0]$ veya $A_{21} = [0]$, $|A_{11}| \neq 0$, $|A_{22}| \neq 0$ ise $A = \begin{bmatrix} A_{11} \vdots [0] \\ \cdots\cdots\cdots\cdots \\ [0] \vdots A_{22} \end{bmatrix}$ olup ,

$A^{-1} = \begin{bmatrix} A_{11}^{-1} \vdots [0] \\ \cdots\cdots\cdots\cdots \\ [0] \vdots A_{22}^{-1} \end{bmatrix}$, $|A| = |A_{11}||A_{22}|$ dir.

Tanım I.2 $A: n \times p$ boyutlu bir matris olsun. A matrisinin mümkün olan bütün kare alt matrislerinin determinantları bulunduğunda, determinantı sıfır olmayan matrisler arasında en yüksek mertebeli alt matrisin mertebesine A matrisinin rankı denir ve $rank(A)$ ile gösterilir.

- ✓ Bir matrisin rankı, o matrisin lineer bağımsız satır ya da sütunlarının sayısı olarak alınabilir.
- ✓ $A: p \times p$ boyutlu singüler olmayan bir matris ve tersi de varsa bu takdirde A matrisinin tam ranklı olduğu ifade edilir ve $rank(A) = p$ yazılır.

Tanım I.3 $X = [x_{ij}] : p \times n$ herhangi bir matris ve $f(X) : X$ matrisinin bir skaler fonksiyonu olsun. Bu takdirde $\frac{df(X)}{dX} = \left[\frac{\partial f(X)}{\partial x_{ij}}\right] : p \times n$ ifadesine, f-skaler fonksiyonunun X matrisine göre birinci mertebeden türevi denir.

Bu tanıma göre bir matrisin skaler fonksiyonunun bu matrise göre birinci mertebeden türevi yine bir matristir. Eğer f-skaler fonksiyonu bir vektörün (satır veya sütun) skaler fonksiyonu ise bu skaler fonksiyonun söz konusu vektöre göre birinci merteden türevi de yine bir vektördür (satır veya sütun).

Türev ile İlgili Özellikler:

i. $X: p \times p,\ |X| \neq 0\ ve\ f(X) = |X|$ ise $\frac{df(X)}{dX} = |X|X^{-1}$ dir.

ii) $X: p \times p,\ |X| \neq 0\ ve\ f(X) = |X|^{\alpha}, \alpha \in IR$ ise $\frac{df(X)}{dX} = \alpha |X|^{\alpha-1} \frac{d(|X|)}{dX}$

iii) $\frac{d}{dX}[f(X)g(X)] = \left[\frac{df(X)}{dX}\right] g(X) + f(X) \left[\frac{dg(X)}{dX}\right]$

iv) $A'p \times q$ ve $X: q \times p$ iken $f(X) = \text{İ}z(A'X)$ ise $\frac{df(X)}{dX} = A$ dır.

v) $\underline{X} : p \times 1$ ve $A: p \times p$ iken $f(\underline{X}) = \underline{X}'A\underline{X}$ ise $\frac{df(\underline{X})}{d\underline{X}} = 2A\underline{X}$ dir.,

vi) $\underline{X} : p \times 1$ ve $\underline{a}: p \times 1$ iken $f(\underline{X}) = \underline{a}'\underline{X}$ ise $\frac{df(\underline{X})}{d\underline{X}} = \underline{a}$ dır.

vii) $X: p \times p$ ve $X = X'$ iken $f(X) = \text{İ}z(X^2)$ ise $\frac{df(X)}{dX} = 2X$ dir.

viii) $\underline{X} : p \times 1$ vektörünün bir skaler fonksiyonu $f(\underline{X})$ ve bu fonksiyonun $\underline{X}$ vektörüne göre birinci mertebeden türevi $\frac{df(\underline{X})}{d\underline{X}}$ olsun. $f(\underline{X})$ fonksiyonunun $\underline{X}$ vektörüne göre ikinci mertebeden türevi;

$$\frac{d^2 f(\underline{X})}{d\underline{X}d\underline{X}'} = \left[\frac{df(\underline{X})}{d\underline{X}}\right]\frac{d}{d\underline{X}'} = \begin{bmatrix} \frac{\partial f(\underline{X})}{\partial x_1} \\ \frac{\partial f(\underline{X})}{\partial x_2} \\ \vdots \\ \frac{\partial f(\underline{X})}{\partial x_p} \end{bmatrix} \begin{bmatrix} \frac{\partial}{\partial x_1} & \frac{\partial}{\partial x_2} & \cdots & \frac{\partial}{\partial x_p} \end{bmatrix} = \begin{bmatrix} \frac{\partial^2 f(\underline{X})}{\partial x_1^2} & \frac{\partial^2 f(\underline{X})}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f(\underline{X})}{\partial x_1 \partial x_p} \\ \frac{\partial^2 f(\underline{X})}{\partial x_2 \partial x_1} & \frac{\partial^2 f(\underline{X})}{\partial x_2^2} & \cdots & \frac{\partial^2 f(\underline{X})}{\partial x_2 \partial x_p} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial^2 f(\underline{X})}{\partial x_p \partial x_1} & \frac{\partial^2 f(\underline{X})}{\partial x_p \partial x_2} & \cdots & \frac{\partial^2 f(\underline{X})}{\partial x_p^2} \end{bmatrix} \text{ olup,}$$

HESSIAN MATRİSİ adını alır. Hessian matrisi simetrik bir matristir.

Tanım I.4 $X = [x_{ij}] : p \times n$ herhangi bir matris ve $f(X) : X$ matrisinin bir skaler fonksiyonu olsun. Bu takdirde $\int_R f(X)dX$ ifadesine, f-skaler fonksiyonunun X matrisine göre integrali (çok katlı integrali) denir. Burada R: X matrisinin elemanlarının tanımlı olduğu bölgeyi gösterir.

Örnek I.2 $\underline{X} : p \times 1$ ve $f(\underline{X}) = \underline{X}'\underline{X}$ iken $R = \{\underline{X} : \ \underline{X}'\underline{X} = 1\}$ alalım.

$\int_R f(\underline{X})d\underline{X} = \int_{x_p} \cdots \int_{x_2} \int_{x_1} \underline{X}'\underline{X}\, dx_1 dx_2 \ldots dx_p$ Bu durumda $f(\underline{X}) = \underline{X}'\underline{X} = \sum_{j=1}^{p} x_j^2$ olup, $R = \{\underline{X} : \ \underline{X}'\underline{X} = 1\} = \{(x_1, \ x_2, \ldots \ x_p) : \ \sum_{j=1}^{p} x_j^2 = 1\}$ yazılır.

Örnek I.3 $\underline{X}' = [X_1 \ \ X_2 \ \ \ldots \ \ X_p]$ $1 \times p$ vektörü ile bu vektörün bir skaler fonksiyonu $f(\underline{X}) = \underline{X}'\underline{X}$ olsun. $\frac{df(\underline{X})}{d\underline{X}} =?$

Çözüm: $f(\underline{X}) = \underline{X}'\underline{X} = X_1^2 + X_2^2 + \cdots + X_p^2$ olup, $\frac{df(\underline{X})}{d\underline{X}} = \begin{bmatrix} \frac{\partial f(\underline{X})}{\partial x_1} \\ \frac{\partial f(\underline{X})}{\partial x_2} \\ \vdots \\ \frac{\partial f(\underline{X})}{\partial x_p} \end{bmatrix} = \begin{bmatrix} 2X_1 \\ 2X_2 \\ \vdots \\ 2X_p \end{bmatrix} = 2\begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_p \end{bmatrix} = 2\underline{X}$

bulunur.

Örnek I.4 i) $X : 2 \times 2$ matrisi verilsin. $f(X) = |X|$ olmak üzere $\frac{df(X)}{dX} =?$

ii) $\underline{X} = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$ ve $A = \begin{bmatrix} 2 & -1 \\ -1 & 4 \end{bmatrix}$ olmak üzere $f(\underline{X}) = \underline{X}'A\underline{X}$ skaler fonksiyonu için $\frac{df(\underline{X})}{d\underline{X}} =?$ ve $\frac{d^2 f(\underline{X})}{d\underline{X}d\underline{X}'} =?$

Çözüm: i) $X = \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix}$ için . $f(X) = |X| = x_{11}x_{22} - x_{12}x_{21}$ olup, $\frac{df(X)}{dX} = \begin{bmatrix} \frac{\partial f(X)}{\partial x_{11}} & \frac{\partial f(X)}{\partial x_{12}} \\ \frac{\partial f(X)}{\partial x_{21}} & \frac{\partial f(X)}{\partial x_{22}} \end{bmatrix} = \begin{bmatrix} x_{22} & -x_{21} \\ -x_{12} & x_{11} \end{bmatrix}$ bulunur. Bu ise $Kof(X)$ matrisidir.

ii) $f(\underline{X}) = \underline{X}'A\underline{X} = [X_1 \quad X_2]\begin{bmatrix} 2 & -1 \\ -1 & 4 \end{bmatrix}\begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = 2X_1^2 + 4X_2^2 - 2X_1X_2$ dir. Buna göre;

$\frac{df(\underline{X})}{d\underline{X}} = \begin{bmatrix} \frac{\partial f(\underline{X})}{\partial x_1} \\ \frac{\partial f(\underline{X})}{\partial x_2} \end{bmatrix} = \begin{bmatrix} 4X_1 - 2X_2 \\ -2X_1 + 8X_2 \end{bmatrix} = \begin{bmatrix} 4 & -2 \\ -2 & 8 \end{bmatrix}\begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = 2A\underline{X}$ olur. Ayrıca; $\frac{d^2 f(\underline{X})}{d\underline{X}d\underline{X}'} =$

$\begin{bmatrix} \frac{\partial^2 f(\underline{X})}{\partial x_1^2} & \frac{\partial^2 f(\underline{X})}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f(\underline{X})}{\partial x_2 \partial x_1} & \frac{\partial^2 f(\underline{X})}{\partial x_2^2} \end{bmatrix} = \begin{bmatrix} 4 & -2 \\ -2 & 8 \end{bmatrix} = 2A$ bulunur.

NOT: $A: p \times p$ simetrik bir matris ve $\underline{X} : p \times 1$ olmak üzere $f(\underline{X}) = \underline{X}'A\underline{X}$ karesel formunun açık şekilde ifade edilmesi: i) A matrisinin köşegen elemanları $\underline{X}$ vektörünün bileşenlerinin karelerinin katsayısı olarak alınır.

ii) A matrisinin köşegen dışındaki elemanların iki katı $\underline{X}$ vektörünün bileşenlerine ait ikili çarpım terimlerinin katsayısı olarak alınır.