KONU 9: KLASİK OPTİMİZASYON

9.1. Tek Değişkenli Dışbükey Fonksiyonlar ve Özellikleri

$f(x)$, verilen bir S dışbükey kümesinde tanımlı ve $\forall\, x \in S$ için ikinci türevi alınabilir bir fonksiyon olsun.

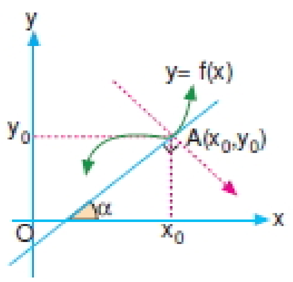
- $\forall\, x \in S$ için $f''(x) \geq 0 \Rightarrow f(x)$ dışbükey bir fonksiyon

  ($\forall\, x \in S$ için $f''(x) > 0 \Rightarrow f(x)$ kesin dışbükey bir fonksiyon)

- $\forall\, x \in S$ için $f''(x) \leq 0 \Rightarrow f(x)$ içbükey bir fonksiyon

  ($\forall\, x \in S$ için $f''(x) < 0 \Rightarrow f(x)$ kesin içbükey bir fonksiyon)

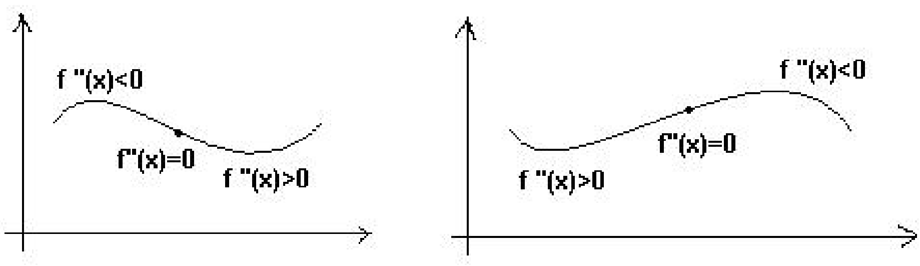
NOT 1: $y = f(x)$ eğrisi üzerindeki $A(x_0, y_0)$ noktasından çizilen teğetin eğimi o noktadaki türevine eşittir.



$$m = \tan\alpha = \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} = f'(x_0)$$

NOT 2: $f'(x_0)$ varsa, $f(x)$ fonksiyonu $x = x_0$ da sürekli fonksiyondur. Tersi doğru olmayabilir. $f(x)$ fonksiyonu $x = x_0$ da sürekli olup, türevi olmayabilir. $f(x)$ fonksiyonu $x = x_0$ da sürekli değilse, türevli de değildir.

NOT 3: $f(x)$ fonksiyonun ikinci türevinin geometrik anlamı Şekil 10.1' de gösterilmiştir.



Şekil 9.1 İkinci türevin geometrik anlamı

Tek değişkenli bir $f(x)$ fonksiyonunda, değişkene göre birinci türev ($\frac{df(x)}{dx}$), $f(x)$ fonksiyonundaki değişim oranını (fonksiyonun eğimini) verirken, ikinci türev ($\frac{d^2f(x)}{dx^2}$), $f(x)$ fonksiyonunun değişim oranının değişim oranını (fonksiyonun eğimindeki değişim oranını) verir.

Bir $f(x)$ fonksiyonunun dışbükey ya da içbükey olup olmadığının belirlenmesinde, öncelikli olarak fonksiyonun birinci türevinin sıfıra eşit olduğu noktadaki uç değeri belirlemek gereklidir. Daha sonra, bu uç değerin fonksiyon için bir maksimum nokta ya da bir minimum nokta olup olmadığının belirlenmesi gerekir. Bu amaçla tanımlanan gerekli ve yeterli koşullara göre uç noktanın türüne karar verilir.

Gerekli Koşullar:

$\frac{df(x)}{dx}=0$ olacak biçimde x^* uç değeri bulunur.

Yeterli Koşullar:

$\frac{d^2f(x^*)}{dx^2}>0$ ise, fonksiyonun değeri, x^* uç değerinde bir minimumdur.

$\frac{d^2f(x^*)}{dx^2}<0$ ise, fonksiyonun değeri, x^* uç değerinde bir maksimumdur.

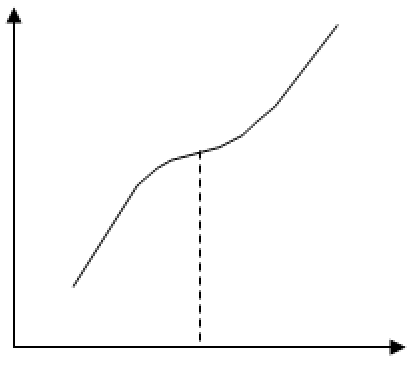
$\frac{d^2f(x^*)}{dx^2}=0$ ise, x^* ın bir uç nokta değeri olması için yüksek mertebeden türevler göz önüne alınmak zorundadır. Bu durumda aşağıdaki teoreme göre karar verilir.

Teorem: Bir $f(x)$ fonksiyonu verilsin. $f(x)$, sabit bir x^* noktasında, $f^{(n-1)}(x^*)=0$ ve $f^{(n)}(x^*)\neq 0$ oluyorsa, $f(x)$ fonksiyonu x^* da

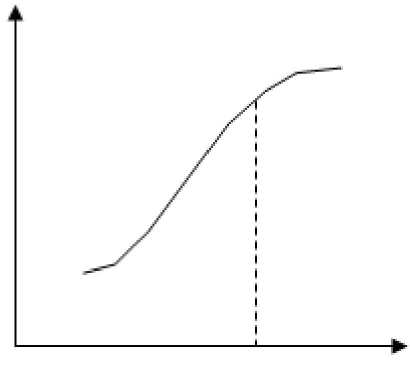
i. n tek sayı ise, bir büküm (dönüm, dönüşüm) noktasına sahiptir.

ii. n çift sayı ise, $f^{(n)}(x^*)>0$ ise, minimum değere, $f^{(n)}(x^*)<0$ ise, maksimum değere sahiptir.

NOT 4: Bir fonksiyonun artış hızının değiştiği (ikinci türevinin işaretinin değiştiği) noktaya **dönüm noktası** denir. Bu noktada ikinci türev sıfır değerini alır. Şekil 9.2 ve Şekil 9.3' te konkavlığın değiştiği noktalarda $f''\left(x^*\right)=0$ dır.

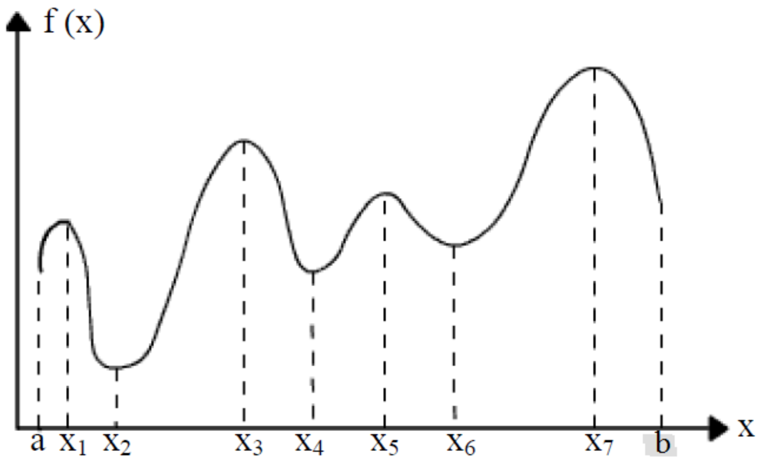


Şekil 9.2 $f(x)$ fonksiyonunun azalarak artan bir durumdan, artarak artan bir duruma geçişi



Şekil 9.3 $f(x)$ fonksiyonunun artarak artan bir durumdan, azalarak artan bir duruma geçişi

Şekil 9.4' te, $[a,b]$ aralığında, tek değişkenli bir $f(x)$ fonksiyonunun, minimum ve maksimum noktaları görülmektedir. Burada, x_1, x_3, x_5, x_7 yerel (lokal) maksimum noktalar olup, x_7 mutlak (global) maksimumdur. x_2, x_4, x_6 yerel (lokal) minimum noktalar olup, x_2 mutlak (global) minimumdur.



Şekil 9.4 $f(x)$ fonksiyonunun minimum ve maksimum noktaları

Örnek 1: $f(x)=\left(x^2-1\right)^3$ fonksiyonunun tanımlılık durumunu inceleyerek, uç değerini/değerlerini belirleyiniz.

Çözüm:

$$\frac{df(x)}{dx}=f'(x)=3\left(x^2-1\right)^2 2x=0 \Rightarrow x_1^*=0\ ,\ x_2^*=1\ ,\ x_3^*=-1$$

$$\frac{d^2f\left(x^*\right)}{dx^2}=f''(x)=6\left(2\left(x^2-1\right)2x^2+\left(x^2-1\right)^2\right)=6\left(x^2-1\right)^2+24x^2\left(x^2-1\right)$$

$f''\left(x_1^*\right)=f''(0)=6>0$ olduğundan, $x_1^*=0$ minimum noktadır.

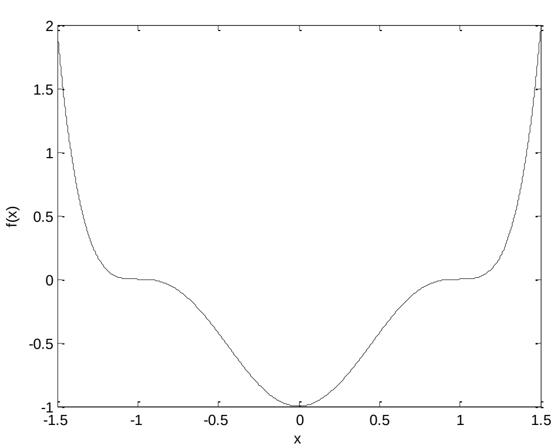
$f''\left(x_2^*\right)=f''(1)=0$ ve $f''\left(x_3^*\right)=f''(-1)=0$ olduğundan, yüksek mertebeden türevler incelenir.

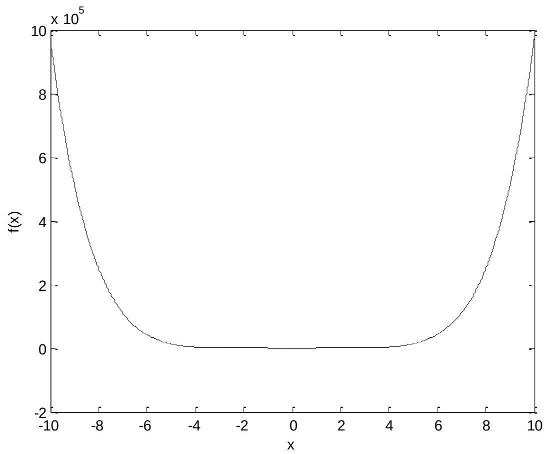
$f'''(x)=24x\left(x^2-1\right)+24\left(2x\left(x^2-1\right)+2x^3\right)$ olmak üzere,

$f'''(1)=48\neq 0$

$f'''(-1)=-48\neq 0$

bulunur. Buna göre, n=3 tek sayı olduğundan verilen $f(x)$ fonksiyonu için $x_2^*=1$ ve $x_3^*=-1$ noktaları büküm (dönüm) noktalarıdır. Şekil 9.5' te fonksiyonun farklı tanım aralıkları için grafiği gösterilmektedir.





Şekil 9.5 $f(x)=\left(x^2-1\right)^3$ fonksiyonu grafikleri

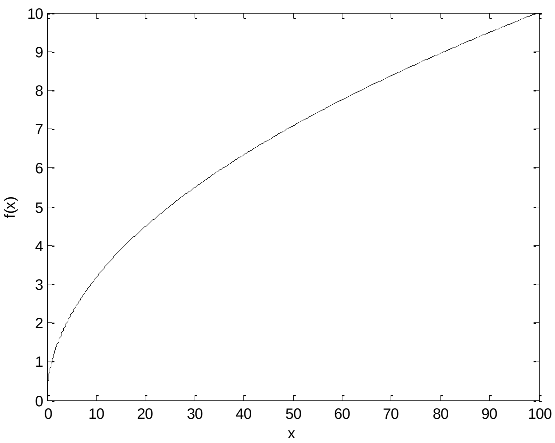
Örnek 2: $f(x)=\sqrt{x}$ fonksiyonunun $S=(0,\infty)$ kümesinde tanımlılık durumunu inceleyiniz.

Çözüm:

$$f'(x)=\frac{1}{2\sqrt{x}}$$

$$f''(x)=-\frac{1}{4}\frac{1}{\sqrt{x^3}}$$

olup, $\forall\, x\in S$ için, $f''(x)<0$ olduğundan, verilen $f(x)$ fonksiyonu içbükey (konkav) bir fonksiyondur. Şekil 9.6' da $f(x)$ içbükey fonksiyonunun grafiği görülmektedir.



Şekil 9.6 $f(x)=\sqrt{x}$ içbükey fonksiyonunun grafiği

9.2. Çok Değişkenli Dışbükey Fonksiyonlar ve Özellikleri

$\mathbf{x}=[x_1\ x_2 \ldots x_n]$ olmak üzere $f(\mathbf{x})$, $S\subset R^n$ de tanımlı, sürekli ve ikinci türevleri alınabilen bir fonksiyon olsun. $f(\mathbf{x})$ fonksiyonunun, x_i, $i=1,2,\ldots,n$ ye göre kısmi türevi

$$\frac{\partial f(\mathbf{x})}{\partial x_i}=\lim_{h\to 0}\frac{f(x_1,x_2,\ldots,x_i+h,\ldots,x_n)-f(x_1,x_2,\ldots,x_i,\ldots,x_n)}{h},\ i=1,2,\ldots,n$$

biçiminde tanımlıdır. Buna göre gradyan vektörü ($\nabla f(\mathbf{x})$)

$$\nabla f(\mathbf{x})=\left[\frac{\partial f(\mathbf{x})}{\partial x_1}\quad \frac{\partial f(\mathbf{x})}{\partial x_2}\quad \ldots \quad \frac{\partial f(\mathbf{x})}{\partial x_n}\right]'$$

dir. $f(\mathbf{x})$ fonksiyonunun Hessian matrisi (ikinci kısmi türev matrisi)

$$\nabla^2 f(\mathbf{x}) = H(\mathbf{x}) = \begin{bmatrix} \dfrac{\partial^2 f(\mathbf{x})}{\partial x_1^2} & \dfrac{\partial^2 f(\mathbf{x})}{\partial x_1 \partial x_2} & \cdots & \dfrac{\partial^2 f(\mathbf{x})}{\partial x_1 \partial x_n} \\ \dfrac{\partial^2 f(\mathbf{x})}{\partial x_2 \partial x_1} & \dfrac{\partial^2 f(\mathbf{x})}{\partial x_2^2} & \cdots & \dfrac{\partial^2 f(\mathbf{x})}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial^2 f(\mathbf{x})}{\partial x_n \partial x_1} & \dfrac{\partial^2 f(\mathbf{x})}{\partial x_n \partial x_2} & \cdots & \dfrac{\partial^2 f(\mathbf{x})}{\partial x_n^2} \end{bmatrix}_{n\times n}$$

olur. Eğer verilen bir noktada $f(\mathbf{x})$' in ikinci kısmi türevleri var ve $f(\mathbf{x})$ bu noktalarda sürekli ise, $\forall\, i \neq j$ için

$$\frac{\partial^2 f(\mathbf{x})}{\partial x_i \partial x_j} = \frac{\partial^2 f(\mathbf{x})}{\partial x_j \partial x_i} \;, \quad i, j = 1,2,\ldots,n$$

olup, H matrisi simetriktir.

Çok değişkenli bir $f(\mathbf{x})$ fonksiyonunun konveksliğinin ya da konkavlığının incelenmesinde H matrisinin tanımlılık durumunun belirlenmesi gerekir.

- $\forall\, \mathbf{x} \in S$ için H matrisi pozitif yarı tanımlı $\Rightarrow f(\mathbf{x})$ dışbükey bir fonksiyondur

  ($\forall\, \mathbf{x} \in S$ için H matrisi pozitif tanımlı $\Rightarrow f(\mathbf{x})$ kesin dışbükey bir fonksiyondur)

- $\forall\, \mathbf{x} \in S$ için H matrisi negatif yarı tanımlı $\Rightarrow f(\mathbf{x})$ içbükey bir fonksiyondur

  ($\forall\, \mathbf{x} \in S$ için H matrisi negatif tanımlı $\Rightarrow f(\mathbf{x})$ kesin içbükey bir fonksiyondur)

NOT 5: (Karesel Fonksiyon)

Eğer, n değişkenli $f(\mathbf{x})$ fonksiyonu $f(\mathbf{x}) = \mathbf{x}' A \mathbf{x} = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j$ biçiminde yazılabilirse, “ $f(\mathbf{x})$ fonksiyonu karesel formda tanımlanabilen bir fonksiyondur” denir. Burada, $A = \left[a_{ij}\right]$, $n \times n$

boyutlu, karesel ve simetrik bir matristir. A matrisi simetrik olmadığında, $\left[a_{ij}+a_{ji}\right]/2$ biçiminde değiştirilerek, simetrik biçime dönüştürülür.

NOT 6: (Karesel Fonksiyonun Türleri)

- $\mathbf{x}\neq\mathbf{0}$ için $f(\mathbf{x})=\mathbf{x}'A\mathbf{x}>0$ ise, $f(\mathbf{x})$ karesel fonksiyonuna pozitif tanımlıdır denir. *A* matrisi de pozitif tanımlıdır.
- Bazı $\mathbf{x}\neq\mathbf{0}$ için $f(\mathbf{x})=\mathbf{x}'A\mathbf{x}>0$ ve bazı $\mathbf{x}\neq\mathbf{0}$ için $f(\mathbf{x})=\mathbf{x}'A\mathbf{x}=0$ ise, $f(\mathbf{x})$ karesel fonksiyonuna pozitif yarı tanımlıdır denir. *A* matrisi de pozitif yarı tanımlıdır.
- Bazı $\mathbf{x}\neq\mathbf{0}$ için $f(\mathbf{x})=\mathbf{x}'A\mathbf{x}>0$ ve bazı $\mathbf{x}\neq\mathbf{0}$ için $f(\mathbf{x})=\mathbf{x}'A\mathbf{x}<0$ ise, $f(\mathbf{x})$ karesel fonksiyonu tanımsızdır (belirsizdir) denir.

Buna göre, $f(\mathbf{x})$ fonksiyonu için tanımlanan *H* matrisinin veya $f(\mathbf{x})=\mathbf{x}'A\mathbf{x}$ karesel fonksiyonu için tanımlanan *A* matrisinin tanımlılık durumları nasıl belirlenmelidir?

H veya *A*, $n\times n$ boyutlu, karesel matrislerinin tanımlılık durumunun incelenmesinde **Asal Minör** hesaplarından yararlanılır.

NOT 7: (Asal Minörler ve Karesel Bir Matrisin Tanımlılık Durumunun İncelenmesi)

Bir $n\times n$ boyutlu, karesel matrisin *k*. asal minörü, son (*n-k*) satırın ve (*n-k*) sütunun matristen çıkarılması ile elde edilen $k\times k$ boyutlu matrisin determinantıdır.

$$A=\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & \ddots & \cdot \\ \cdot & \cdot & & \cdot \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}$$

olmak üzere, bu matrisin asal minörleri

$$\Delta_1 = a_{11}\,,\ \Delta_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix},\ \Delta_3 = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}, \dots,\ \Delta_n = |A|$$

dır.

Asal minörler elde edildikten sonra ortaya çıkabilecek sonuçlar aşağıdaki gibi tanımlanabilir:

i. A' nın tüm asal minörleri > 0 $\Rightarrow$ A matrisi pozitif tanımlıdır.

($\Delta_1 > 0$, $\Delta_2 > 0$, $\Delta_3 > 0, \dots, \Delta_n > 0 \Rightarrow A$ matrisi pozitif tanımlıdır.)

ii. A' nın tüm asal minörleri ≥ 0 $\Rightarrow$ A matrisi pozitif yarı tanımlıdır.

($\Delta_1 \geq 0$, $\Delta_2 \geq 0$, $\Delta_3 \geq 0, \dots, \Delta_n \geq 0 \Rightarrow A$ matrisi pozitif yarı tanımlıdır.)

iii. A' nın k. mertebe asal minörü $(-1)^k$ ile aynı işaretli $\Rightarrow$ A matrisi negatif tanımlıdır.

($\Delta_1 < 0$, $\Delta_2 > 0$, $\Delta_3 < 0, \dots \Rightarrow A$ matrisi negatif tanımlıdır.)

iv. A' nın k. mertebe asal minörü $(-1)^k$ ile aynı işaretli ya da sıfır $\Rightarrow$ A matrisi negatif yarı tanımlıdır.

($\Delta_1 \leq 0$, $\Delta_2 \geq 0$, $\Delta_3 \leq 0, \dots \Rightarrow A$ matrisi negatif yarı tanımlıdır.)

v. Bunlardan hiç birine uymuyorsa belirsizlik durumu vardır.

Buna göre, bir $f(\mathbf{x})$ fonksiyonunun ekstremum noktalarının belirlenmesinde gerek ve yeter koşullar:

Gerekli Koşullar:

$\nabla f(\mathbf{x}) = \mathbf{0}$ olacak biçimde $\mathbf{x}^*$ vektörü bulunur.

Yeterli Koşullar:

$H(\mathbf{x}^*)$ pozitif tanımlı ise, $\mathbf{x}^*$ bir minimum noktadır.

$H(\mathbf{x}^*)$ negatif tanımlı ise, $\mathbf{x}^*$ bir maksimum noktadır.

$H(\mathbf{x}^*)$ tanımsız ise, $\mathbf{x}^*$ bir büküm noktasıdır.

NOT 8: Eğer, $f(\mathbf{x})$ fonksiyonu, $f(\mathbf{x})=\mathbf{x}'A\mathbf{x}$ biçiminde karesel form olarak yazılabiliyorsa, Hessian matrisini bulmadan A matrisinin tanımlılık durumunun incelenmesi ile de fonksiyonun minimum ya da maksimum çözüm değerlerinin olup olmadığına karar verilir. Karesel bir $f(\mathbf{x})=\mathbf{x}'A\mathbf{x}$ fonksiyonu için, H=2A dır.

Örnek 3: $f(\mathbf{x})=f(x_1,x_2)=x_1^2+2x_1x_2+2x_2^2$ biçiminde tanımlı fonksiyonun ekstremum (minimum/maksimum) noktasını/noktalarını elde ederek, fonksiyonun türünü belirleyiniz.

Çözüm:

I.Yol:

$$\nabla f(\mathbf{x})=\mathbf{0}$$
$$\begin{bmatrix} 2x_1+2x_2 \\ 2x_1+4x_2 \end{bmatrix}=\begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

olup, $\mathbf{x}^*=(0,0)$ değeri elde edilir. Hessian matrisi

$$H(\mathbf{x})=\begin{bmatrix} 2 & 2 \\ 2 & 4 \end{bmatrix}$$

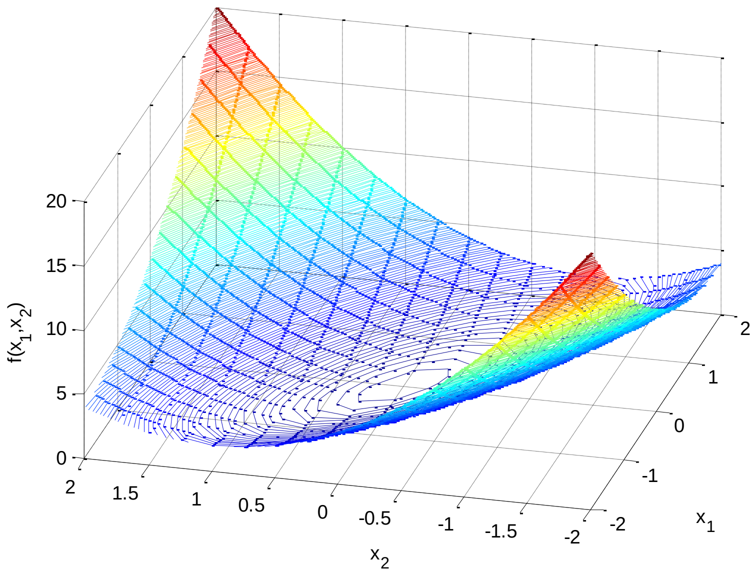
elde edilir. Burada, asal minör değerleri $\Delta_1=2>0$ ve $\Delta_2=|H|=8-4=4>0$ olup, H matrisinin pozitif tanımlı olduğu söylenir. Buna göre, verilen $f(\mathbf{x})$ fonksiyonu pozitif tanımlı olup, $\mathbf{x}^*=(0,0)$ bir minimum çözüm vektörüdür. $f(\mathbf{x})$ dışbükey bir fonksiyondur.

II. Yol:

$$\begin{aligned} f(\mathbf{x})=f(x_1,x_2)&=x_1^2+2x_1x_2+2x_2^2 \\ &=\mathbf{x}'A\mathbf{x} \\ &=\begin{bmatrix} x_1 & x_2 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \end{aligned}$$

dır. $A=\begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$ olup, asal minörlerin hesaplanmasıyla ($\Delta_1=1>0$ ve $\Delta_2=|A|=2-1=1>0$), A matrisinin de pozitif tanımlı olduğu söylenir. Buna göre, $f(\mathbf{x})$ fonksiyonu pozitif tanımlı olup, dışbükey bir fonksiyondur. Dikkat edilirse, H=2A olduğu açıktır.

Şekil 9.7 de $f(\mathbf{x})$ fonksiyonunun yüzey grafiği görülmektedir. Şekil 9.7' den de görüleceği gibi, fonksiyon $\mathbf{x}^* = (0,0)$ noktasında minimum değere ulaşmaktadır.



Şekil 9.7 $f(\mathbf{x}) = x_1^2 + 2x_1x_2 + 2x_2^2$ fonksiyonunun yüzeyi

Örnek 4: $f(\mathbf{x}) = x_1^2 + x_2^2 + 2x_3^2 - x_1x_2 - x_2x_3 - x_1x_3$ fonksiyonunun tanımlılık durumunu inceleyiniz.

Çözüm:

$$\nabla f(\mathbf{x}) = [2x_1 - x_2 - x_3 \quad 2x_2 - x_1 - x_3 \quad 4x_3 - x_2 - x_1]'$$

$$H(\mathbf{x}) = \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 4 \end{bmatrix}$$

$\Delta_1 = 2 > 0,\ \Delta_2 = 4 - 1 = 3 > 0,\ \Delta_3 = |H| = 6 > 0$ olup, H matrisi pozitif tanımlıdır. Buna göre, $f(\mathbf{x})$ dışbükey bir fonksiyondur.

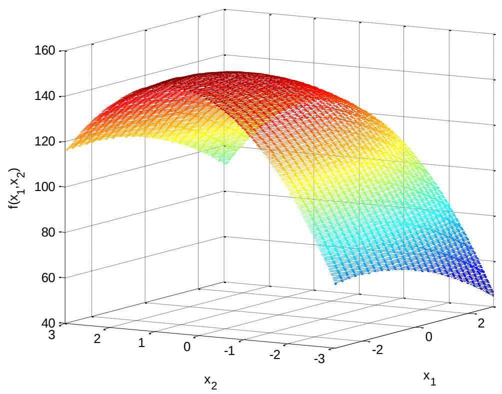
Örnek 5: $f(\mathbf{x}) = f(x_1, x_2) = -x_1^2 - 6x_2^2 - 4x_1 + 8x_2 + 143$ biçiminde tanımlı fonksiyonun ekstremum (minimum/maksimum) noktasını/noktalarını elde ederek, fonksiyonun türünü belirleyiniz.

Çözüm:

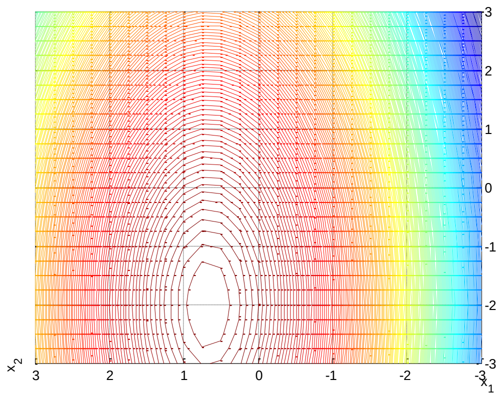
$$\nabla f(\mathbf{x}) = \mathbf{0}$$

$$\begin{bmatrix} -2x_1 - 4 \\ -12x_2 + 8 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Rightarrow \mathbf{x}^* = \begin{bmatrix} -2 \\ 2/3 \end{bmatrix}$$

$H(\mathbf{x}) = \begin{bmatrix} -2 & 0 \\ 0 & -12 \end{bmatrix}$ olup, $\Delta_1 = -2 < 0$ ve $\Delta_2 = |H| = 24 > 0$ bulunur. Buna göre, H matrisi negatif tanımlıdır. $f(\mathbf{x})$ fonksiyonu negatif tanımlı olup, içbükey bir fonksiyondur. $\mathbf{x}^* = \begin{bmatrix} -2 \\ 2/3 \end{bmatrix}$ maksimum noktadır. Buradan, $f(\mathbf{x}^*) = 149.6667$ elde edilir. Şekil 9.8' de $f(\mathbf{x})$ fonksiyonunun yüzey grafiği ve Şekil 9.9' da $f(\mathbf{x})$ fonksiyonunun kontur grafiği görülmektedir.



Şekil 9.8 $f(\mathbf{x})$ fonksiyonunun yüzey grafiği



Şekil 9.9 $f(\mathbf{x})$ fonksiyonunun kontur grafiği