Optimizasyon Teknikleri

Ders Notu – 2

MATEMATİKSEL TEMELLER

Prof. Dr. Bilal ALATAŞ

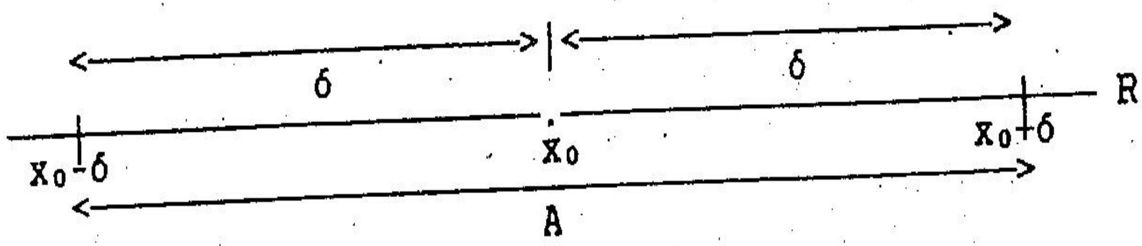
Matematiksel Temeller

$X_0 \in S$ iken X_0'ın S kümesindeki δ komşuluğu (civarı)

A= $\{x \mid |x - x_0| \leq \delta , x \in S\}$

şeklinde gösterilir.

S=R alınırsa, x_0'ın δ komşuluğu



şeklinde reel eksen üzerinde gösterilir. Bir noktanın δ komşuluğu verilen noktaya uzaklığı δ veya daha az olan noktalar kümesi olmaktadır.

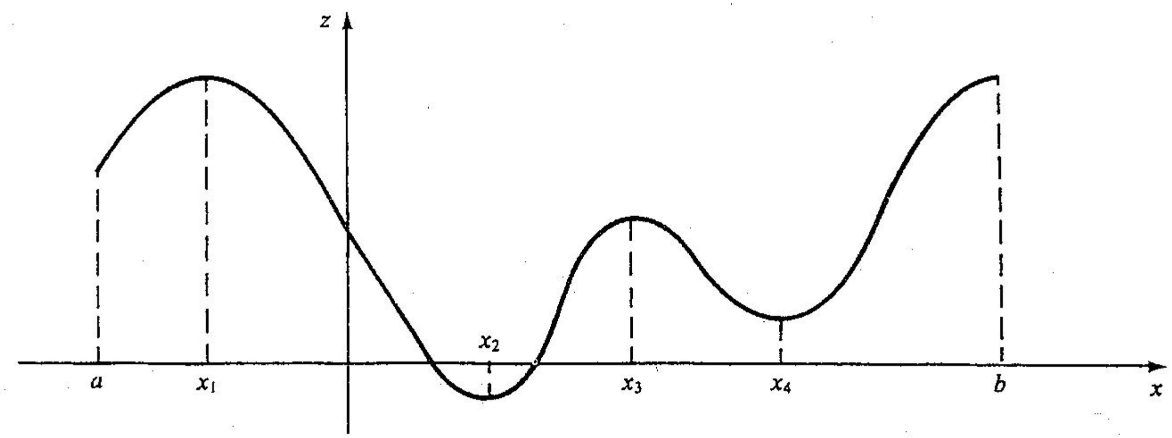
Matematiksel Temeller

X bir vektör olmak üzere, f(x)'in tanım kümesi S ve $x_0 \in S$ için, x_0'ın δ komşuluğu A olsun.

- **A daki tüm x ler için f(x_0) $\leq$ f(x) ise, f(x) x= $x_0$da yerel en küçük (local minimum)**
- **S deki tüm x ler için f(x_0) $\leq$ f(x) ise, f(x) x= $x_0$da bütünsel en küçük (global minimum, mutlak minimum)**
- **A daki tüm x ler için f(x_0) $\geq$ f(x) ise, f(x) x= $x_0$da yerel en büyük (local maksimum)**
- **S deki tüm x ler için f(x_0) $\geq$ f(x) ise, f(x) x= $x_0$da bütünsel en büyük(global maksimum, mutlak maksimum)**

değer alır.

Şimdi $f:[a,b]\rightarrow R^2$ tanımlı fonksiyonun grafiğini göz önüne alalım:



Şekil

Burada a, x_2,x_4 yerel minimum noktaları; x_1, x_3,b yerel maksimum noktalarıdır. Tanım aralığı içindeki tüm x değerleri için en büyük değer x=b noktasında elde edildiğinden b noktası fonksiyonun global maksimum noktası ve tanım aralığı içindeki tüm x değerleri için en küçük değer x= x_2 noktasında elde edildiğinden x_2 noktası fonksiyonun global minimum noktasıdır.

Konvekslik ve Konkavlık

$x_i \epsilon$ S ve $\lambda_i \geq 0$ $\sum \lambda i=1$ $(i=1,2,\dots,n)$ iken

$x_0= \lambda_1 x_1+ \lambda_2 x_2 + \dots + \lambda_n x_n$

$x_0=\sum_{i=1}^{n} \lambda i \, xi$

şeklinde elde edilen x_0'a $x_1, x_2,\dots, x_n$ noktalarının konveks bileşimi denir.

Verilen bir S kümesinin farklı her iki noktasının konveks bileşimiyle bulunan nokta tekrar S kümesine ait ise S ye konveks küme denir.

Konveks Fonksiyon

$f(x)$ verilen bir kümede tanımlı iken, bu kümenin farklı x_1 ve x_2 noktaları ve $\lambda \epsilon [0, 1]$ için;

$$f(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f(x_1) + (1-\lambda)f(x_2) \text{ ise}$$

$f(x)$ e dış bükey (konveks) fonksiyon adı verilir.

Eğer;

$$f(\lambda x_1 + (1-\lambda)x_2) < \lambda f(x_1) + (1-\lambda)f(x_2) \text{ ise}$$

$f(x)$ e kesin dış bükey (kesin konveks) fonksiyon adı verilir.

Bu tanıma göre; $f(x)$ in gösterdiği eğri üzerinde alınan iki noktayı birleştiren doğru parçası, fonksiyonun üstünde kalıyorsa veya çakışıyorsa $f(x)$ konveks bir fonksiyon olmaktadır.

Konkav Fonksiyon

f(x) verilen bir kümede tanımlı iken, bu kümenin farklı x_1 ve x_2 noktaları ve $\lambda \epsilon [0, 1]$ için;

$$f(\lambda x_1 + (1-\lambda)x_2) \geq \lambda f(x_1) + (1-\lambda)f(x_2) \text{ ise}$$

f(x) e iç bükey (konkav) fonksiyon adı verilir.

Eğer;

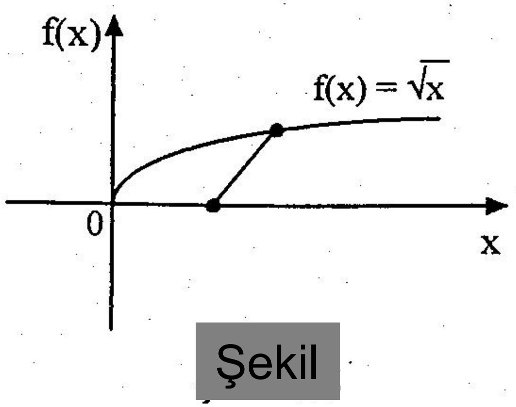
$$f(\lambda x_1 + (1-\lambda)x_2) > \lambda f(x_1) + (1-\lambda)f(x_2) \text{ ise}$$

f(x) e kesin iç bükey (kesin konkav) fonksiyon adı verilir.

Bu tanıma göre; f(x) fonksiyonunun gösterdiği eğri üzerinde alınan farklı iki noktayı birleştiren doğu parçası, fonksiyonun altında kalıyor veya fonksiyonla çakışıyorsa, f(x)’e içbükey(konkav) fonksiyon adı verilir.

Örnekler

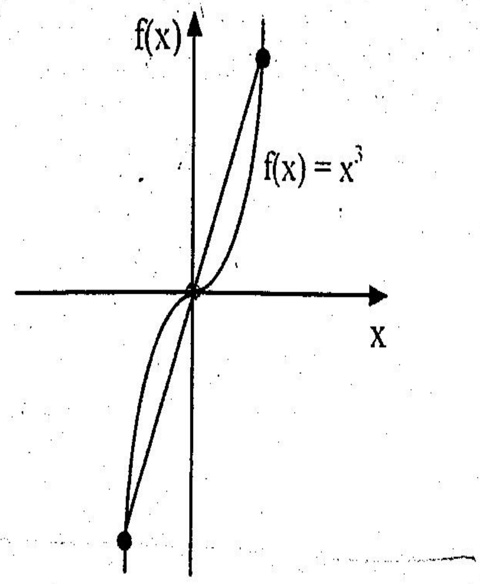
1) f(x)=$\sqrt{x}$ fonksiyonu S=$[0,\infty)$ aralığında konveks midir konkav mıdır? İnceleyelim.



Konkav Fonksiyon

Şekilden de görüldüğü gibi, fonksiyonun eğrisi üzerindeki herhangi iki noktayı birleştiren doğru parçası eğrinin altında kaldığından bu fonksiyon konkavdır.

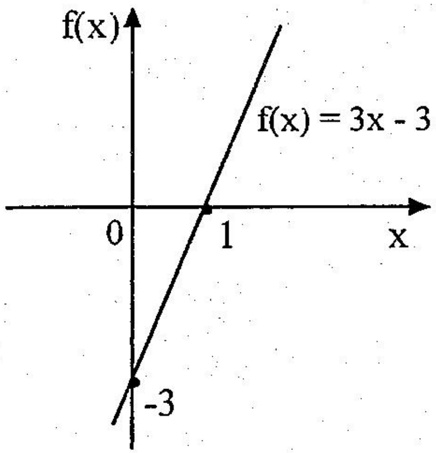
2) $f(x)=x^3$ fonksiyonu $S = R^1= (-\infty, \infty)$ üzerinde konveks midir? Konkav mıdır



Şekil 6.8
Ne Konveks Ne de Konkav Fonksiyon

Şekil incelendiğinde eğri üzerindeki herhangi iki noktayı birleştiren doğrunun $(-\infty,0]$ aralığında büküm noktasının altında, $[0,\infty)$ aralığında büküm noktasının üzerinde kaldığı görülür. Buradan fonksiyonun $(-\infty,0]$ aralığında konkav, $[0,\infty)$ aralığında konveks olduğunu söyleyebiliriz. Ancak $(-\infty, \infty)$ aralığında ne konveks ne de konkavdır.

3) f(x)=3x-3 fonksiyonu $S = R^1 = (-\infty, \infty)$ üzerinde konveks midir konkav mıdır?



Şekil 6.9
Hem Konveks Hem de Konkav Fonksiyon

Şekilden de görüldüğü gibi fonksiyonun eğrisi büküm noktası bulunmayan bir doğrudur. Doğru üzerindeki herhangi iki noktayı birleştiren doğru parçası bu doğru ile çakıştığından f(x) hem konveks hem de konkavdır.

$f(x)$ in tanımlı olduğu S kümesi içinde x_0 ın komşuluğu A olsun. Şu halde,

- **Eğer $f(x)$, x_0 da yerel en küçük değerini alıyorsa, $f(x_0)$ A kümesinde konvekstir.**
- **Eğer $f(x)$, x_0 da yerel en büyük değerini alıyorsa, $f(x_0)$ A kümesinde konkavdır.**

Yerel en iyilerle bütünsel en iyiler arasındaki özellikler ise şunlardır:

- **Eğer $f(x_1,x_2,…,x_n)$ bir konveks kümede tanımlı konveks bir fonksiyon ve x_0 $f(x)$ in yerel en küçük noktası ise $f(x)$ x_0 da bütünsel en küçük (global minimum) değerini alır.**
- **Eğer $f(x_1,x_2,…,x_n)$ bir konveks kümede tanımlı konkav bir fonksiyon ve x_0 $f(x)$ in yerel en büyük noktası ise $f(x)$ x_0 da bütünsel en büyük (global maksimum) değerini alır.**

TEOREM: Doğrusal olmayan denklem (Doğrusal Olmayan Programlama problemi - DOP) uygun çözüm bölgesi S konveks bir küme olan en büyükleme amaçlı bir fonksiyon olsun. f(x), S üzerinde konkav ise doğrusal olmayan programlama probleminin herhangi bir yerel en büyüğü bu DOP probleminin en iyi çözümü olur.

TEOREM: Doğrusal olmayan denklem uygun çözüm bölgesi S konveks küme olan en küçükleme amaçlı bir problem olsun. f(x), S üzerinde konveks ise DOP probleminin herhangi bir yerel en küçüğü bu problemin en iyi çözümü olur.

TEOREM: S konveks küme kapsamındaki bütün x ler için f''(x) varolsun. Bütün $x \in S$ ler için

$f''(x) \geq 0$ ise f(x) konvekstir.

TEOREM: S konveks küme kapsamındaki bütün x ler için f''(x) varolsun. Bütün $x \in S$ ler için

$f''(x) \leq 0$ ise f(x) konkavdır.

ÖRNEKLER

1) **f(x)=$\sqrt{x}$ fonksiyonunun [0,∞) aralığında konveks mi konkav mı olduğunu türev yaklaşımıyla inceleyiniz.**

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

f''(x)=-$\frac{1}{4}x^{-\frac{3}{4}} \leq 0$

olduğundan verilen fonksiyon S=[0,∞)üzerinde konkavdır.

2) **f(x) = x^2-3 fonksiyonu (-∞, ∞)aralığında konveks midir konkav mıdır? İnceleyelim.**

f'(x)= 2x

f''(x)=2 $\geq$0 olduğundan f(x) (-∞, ∞)aralığında konvekstir.

Çok değişkenli fonksiyonların konveksliği ya da konkavlığı

TANIM: n değişkenli bir $f(x_1,x_2,\ldots,x_n)$ fonksiyonunun birinci mertebeden kısmi türevleriyle oluşturulan matrise gradyant vektör denir ve ∇f ile gösterilir.

$$\nabla f(x_1,x_2,\ldots,x_n) = \left(\frac{\partial f}{\partial x1}, \frac{\partial f}{\partial x2}, \ldots, \frac{\partial f}{\partial xn}\right)$$

TANIM: n değişkenli bir $f(x_1,x_2,\ldots,x_n)$ fonksiyonun ikinci dereceden kısmi türevleriyle i. Satır j. Sütun elemanı, $\frac{\partial^2 f}{\partial x_i \partial x_j}$ olmak üzere oluşturulan n*n boyutlu matrise $f(x_1,x_2,\ldots,x_n)$ fonksiyonunun Hessian Matrisi denir ve $H(x_1,x_2,\ldots,x_n)$ ile gösterilir.

$f(x_1,x_2,\ldots,x_n)$ fonksiyonun hessian matrisi

$$H(x_1,x_2,\ldots,x_n = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & & \\ \frac{\partial^2 f}{\partial x_1 \partial x_n} & \frac{\partial^2 f}{\partial x_2 \partial x_n} \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{bmatrix}$$

TANIM:(asal minörler)

$$\begin{bmatrix} a_{11} & a_{12} & \dots a_{1n} \\ a_{21} & a_{22}, & \dots a_{2n} \\ & & \vdots \\ a_{n1} & a_{n2} & \dots a_{nn} \end{bmatrix}$$ **biçimindeki n*n lik karesel matrisin asal minörleri:**

1. asal minörü : $|a_{11}|$

2. asal minörü: $\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$

3. asal minörü: $\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$ **biçiminde bulunur.**

Çok değişkenli bir $f(x_1,x_2,\dots,x_n)$ fonksiyonun $S \subset R^n$ konveks kümesi üzerinde

konveks veya konkav olduğu, fonksiyonun hesssian matrisinin asal minörlerinin işaretlerine bakarak belirlenir

<u>TEOREM:</u> **$f(x_1,x_2,\ldots,x_n)$ fonksiyonun S ye ait bütün $x=(x_1, x_2,\ldots, x_n)$ noktalarında sürekli ikinci kısmi türevlere sahip olsun. Bu durumda : $f(x_1,x_2,\ldots,x_n)$ fonksiyonun hessian matrisinin bütün asal minörleri negatif değilse, konvekstir.**

<u>TEOREM:</u> **$f(x_1,x_2,\ldots,x_n)$ fonksiyonun S ye ait bütün $x=(x_1, x_2,\ldots, x_n)$ noktalarında sürekli ikinci kısmi türevlere sahip olsun. Bu durumda $f(x_1,x_2,\ldots,x_n)$ fonksiyonun hessian matrisinin bütün asal minörleri $(-1)^k$ ile aynı işarete sahipse $f(x_1,x_2,\ldots,x_n)$ fonksiyonu konkavdır.**

Örnek

$f(x)=x_1^2 + 3x_1x_2+ 2x_2^2- x_2+ x_1$ fonksiyonunun konveks olup olmadığını araştırınız.

Önce gradyant vektörünü oluşturalım:

$$\nabla f\ (x_1,x_2)= \frac{\partial f}{\partial x1} = 2x_1 + 3x_2 + 1$$

$$\nabla f\ (x_1,x_2)= \frac{\partial f}{\partial x2}= 3x_1 + 4x_2 - 1$$

$$\frac{\partial^2 f}{\partial x_1^2}=2\ ,\quad \frac{\partial^2 f}{\partial x_1 \partial x_2}=3\ ,\quad \frac{\partial^2 f}{\partial x_2 \partial x_1}= 3,\qquad \frac{\partial^2 f}{\partial x_2^2}=4$$

Hessian matrisi;

$$H=\begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} \end{bmatrix}= \begin{bmatrix} 2 & 3 \\ 3 & 4 \end{bmatrix}$$

biçiminde bulunur.

1. **Asal minör:** $|2|=2 >0$
2. **Asal minör:** $\begin{vmatrix} 2 & 3 \\ 3 & 4 \end{vmatrix}= 8-9 = -1 < 0$ **dır.**

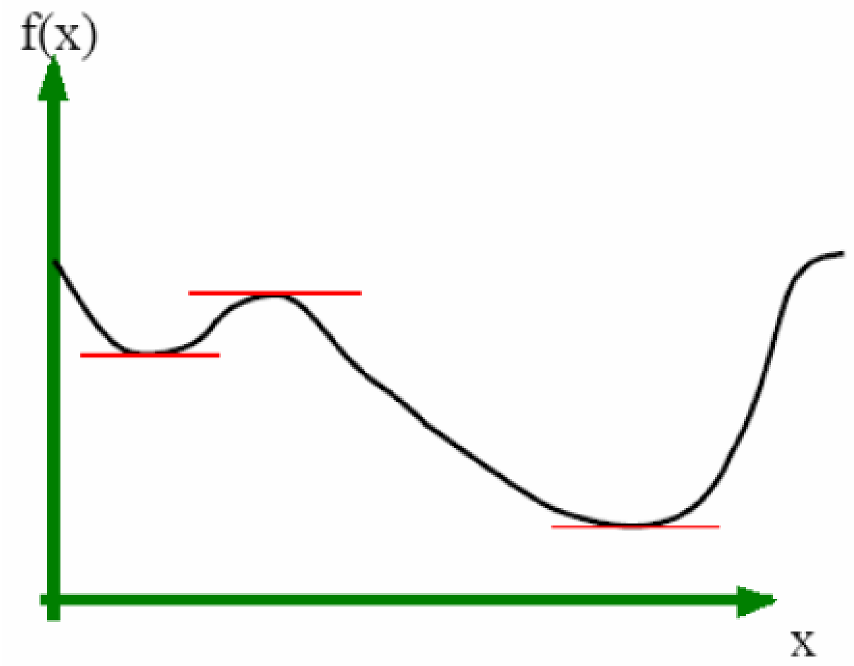
Görüldüğü gibi asal minörlerin hepsinin işareti pozitif değil. Dolayısıyla f fonksiyonu konveks değil. Benzer şekilde asal minörlerin işaretleri $(-1)^k$ ile de uyumlu olmadığından f fonksiyonu konkav değildir. Buradan f fonksiyonu ne konveks ne de konkavdır.

OPTIMALLIK SARTLARI
(Kisitsiz ve Tek Degiskenli Problemler)

- Bir noktanin *lokal optimum* olma sartlari:
  ✓ Gerek sart: $f'(x) = 0$
  ✓ Yeter sart: $f''(x) > 0$

- *Global optimum* için matematiksel bir kriter yoktur

- Eger fonksiyon *konveks* ise, lokal optimum ayni zamanda global optimumdur.



Tek Değişkenli Optimizasyon

Tek değişkenli bir f(x) fonksiyonunda, h yeterince küçük pozitif ve negatif değerler olmak üzere, eğer

$$f(x^*) \leq f(x^*+h)$$

ise fonksiyon $x=x^*$ noktasında göreli (relative) ya da yerel (lokal) minimuma sahiptir denir.

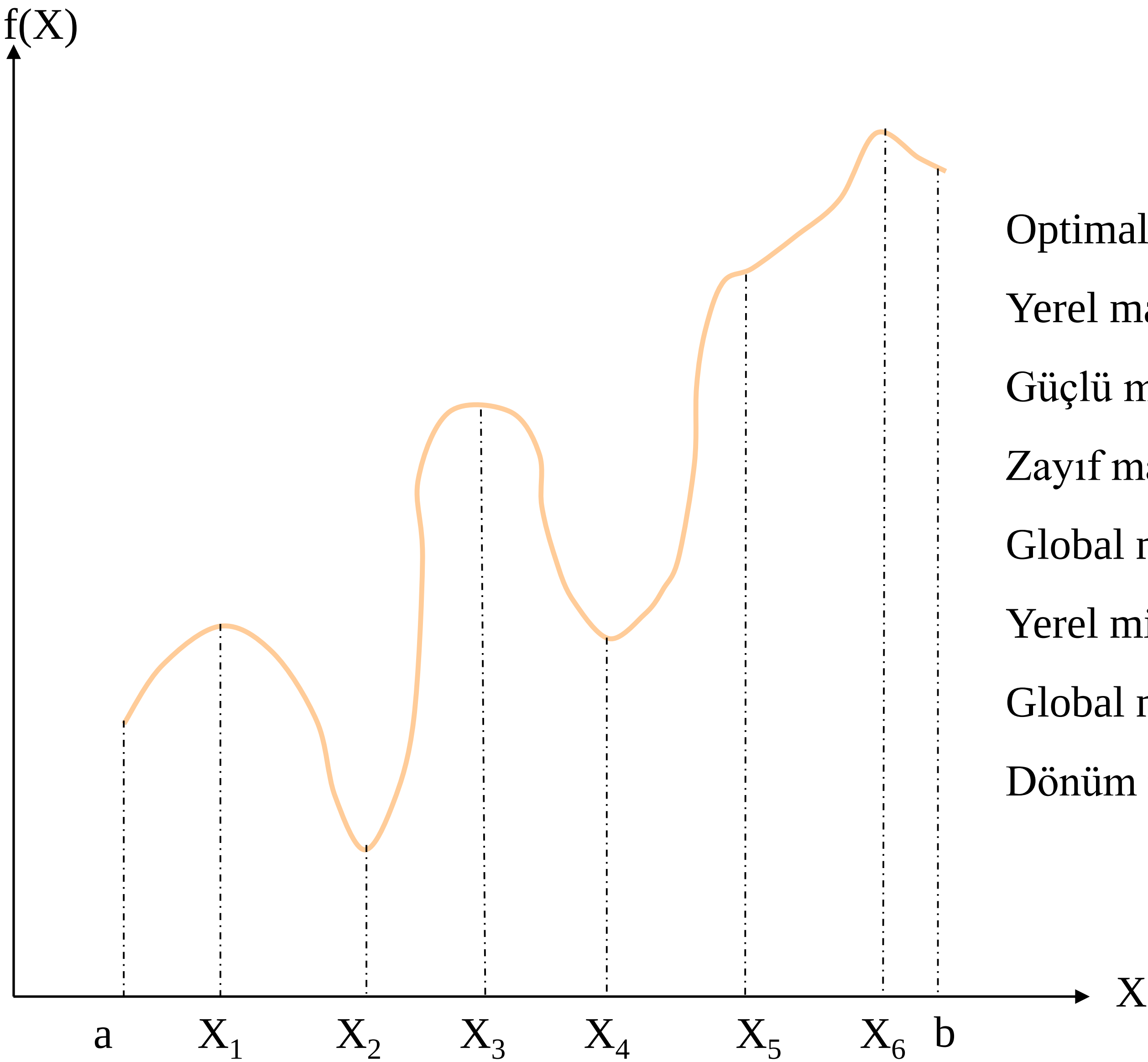
Tek değişkenli bir f(x) fonksiyonunda, h yeterince küçük pozitif ve negatif değerler olmak üzere, eğer

$$f(x^*) \geq f(x^* + h)$$

ise fonksiyon $x=x^*$ noktasında göreli ya da yerel maksimuma sahiptir denir.

Eğer sadece x^*'a çok yakın noktalarda değil de tüm x noktalarında, $f(x^*) \leq f(x)$ ise f(x) fonksiyonu x^* noktasında tümel (global, küresel) minimuma sahiptir denir. Benzer şekilde, eğer bütün x noktalarında $f(x^*) \geq f(x)$, x^* noktası global maksimumdur.

Tek değişkenli optimizasyon probleminin amacı, [a, b] aralığında f(x)'i maksimum ya da minimum yapan $x=x^*$ noktasını araştırmaktır.



Optimal noktalar: x_1, x_2, x_3, x_4 ve x_6

Yerel maksimum noktalar: x_1, x_3, x_6

Güçlü maksimum noktalar: $x_1$ve x_6

Zayıf maksimum nokta: x_3

Global maksimum nokta: x_6

Yerel minimum noktalar: x_2 ve x_4

Global minimum nokta: x_2

Dönüm noktası: x_5

f(x) fonksiyonunun maksimum ve minimum noktaları

Şekil incelendiğinde, x_1, x_2, x_3, x_4 ve x_6 noktalarının f(x) fonksiyonunun optimal noktaları oldukları görülebilir. x_1, x_3, ve x_6 yerel maksimum; x_2 ve x_4 yerel minimum noktalardır. x_5 ise bir optimal nokta değildir. Bu gibi noktalar **dönüm (inflection) noktası** olarak adlandırılmaktadır. x_3 diğerlerinden farklı bir maksimum noktadır; çünkü x_3'ün komşuluğunda en az bir noktada f(x)'in değeri $f(x_3)$'e eşittir. Bir zayıf maksimum nokta, sonsuz sayıda maksimum noktayı gösterir. Bu nedenle, x_3 zayıf maksimum nokta olarak adlandırılır. x_1 ve x_6 ise güçlü maksimum noktalardır. Benzer şeyler minimum noktalar için de geçerlidir. x_2 ve x_6 noktaları sırasıyla global minimum ve maksimum noktalardır.

Tek değişkenli bir fonksiyonun yerel minimum ve maksimumlarının bulunması için gerekli koşullar

Teorem 1: Bir f(x) fonksiyonu, $a \le x \le b$ aralığında tanımlansın. Eğer bu aralıkta fonksiyonun

$$\frac{df(x)}{dx} = f'(x)$$

türevi mevcutsa ve $a<x^*<b$ olmak üzere x^* noktasında bir yerel minimuma sahipse, $f'(x^*)=0$' dır.

İspat: Fonksiyonun f '(x*) türevinin sıfıra eşit olacağını kanıtlamak istiyoruz. f '(x*) aşağıdaki gibi yazılabilir.

$$f'(x^*) = \lim_{h \to 0} \frac{f(x^* + h) - f(x^*)}{h} \qquad (1)$$

x^* bir yerel minimum olduğuna göre, h sıfıra yeterince yakın olmak üzere,

$f(x^*) \leq f(x^* + h)$

yazılabilir. Dolayısıyla, aşağıdaki ifadeler yazılabilir.

$$\frac{f(x^* + h) - f(x^*)}{h} \geq 0, \quad \text{eger} \quad h > 0 \text{ ise}$$

$$\frac{f(x^* + h) - f(x^*)}{h} \leq 0, \quad \text{eger} \quad h < 0 \text{ ise}$$

h>0 değeri sıfıra yaklaşırken (1) denklemi

$$f'(x^*)\geq 0 \qquad (2)$$

ve h<0 değeri sıfıra yaklaşırken (1) denklemi

$$f'(x^*)\leq 0 \qquad (3)$$

şeklinde yazılabilir. (2) ve (3) eşitliklerinden,

$$f'(x^*)=0$$

bulunur. Böylece teorem ispatlanmış olur. Teorem, x*'ın maksimum olması halinde de benzer şekilde tasarlanabilir. Teoremde, tanım aralığının bir uç noktasında minimum oluştuğunda ne olacağının belirtilmediğine dikkat edilmelidir.

Maksimum ya da minimum noktaların belirlenmesi

Türevin sıfır olduğu her noktada fonksiyonun bir minimum ya da maksimuma sahip olduğu söylenemez. Eğer bir noktada türev sıfır ise bu nokta, ya bir optimal noktadır ya da bir dönüm noktasıdır ve durağan (stationary) nokta olarak adlandırılır. Durağan noktanın cinsinin anlaşılabilmesi için 2. mertebeden türevlerin incelenmesi gerekir.

$f''(x^*)>0$ ise x^* bir minimum noktadır.
$f''(x^*)<0$ ise x^* bir maksimum noktadır.

2. türevin sıfır olması halinde, bir belirsizlik durumu söz konusudur ve daha yüksek mertebeden türevler araştırılmalıdır.

Teorem 2: Bir f(x) fonksiyonu verilsin. x^* durağan noktasında, ilk (n-1) mertebeden türev sıfır ve $f^n(x^*) \neq 0$ ise $x=x^*$ noktasında,

a) n tek ise f(x) bir dönüm noktasına sahiptir.

b) n çift ise f(x) bir optimal noktaya sahiptir. Eğer

- $f^n(x^*)<0$ ise bu optimal nokta f(x)'in bir maksimum noktasıdır.
- $f^n(x^*)>0$ ise bu optimal nokta f(x)'in bir minimum noktasıdır.

İspat Fikri: Taylor teoreminden aşağıdaki ifade yazılabilir:

$$f(x^*+h) = f(x^*) + hf'(x^*) + \frac{h^2}{2!} f''(x^*) + \ldots$$

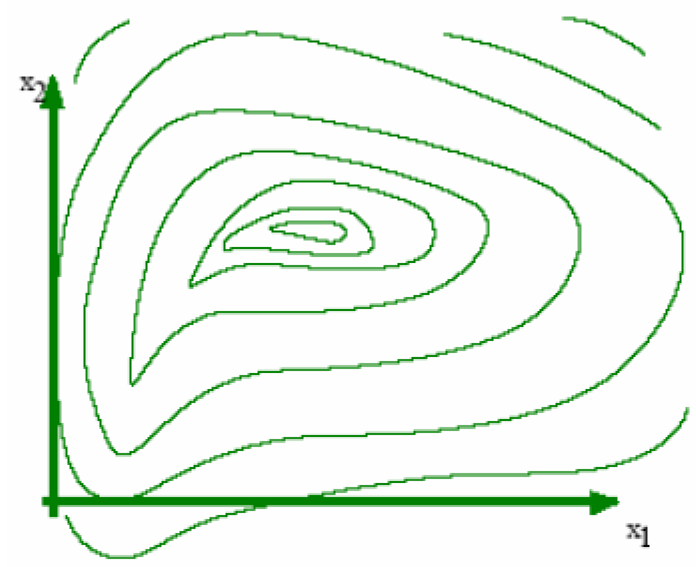
$$+ \frac{h^{n-1}}{(n-1)!} f^{(n-1)}(x^*) + \frac{h^n}{n!} f^{(n)}(x^*+\theta h), \qquad 0<\theta<1$$

OPTIMALLIK SARTLARI
(Kisitsiz ve Çok Degiskenli Problemler)

- Gerek sart: $$\frac{\partial f}{\partial x_j} = 0, \text{ for } j = 1..n$$

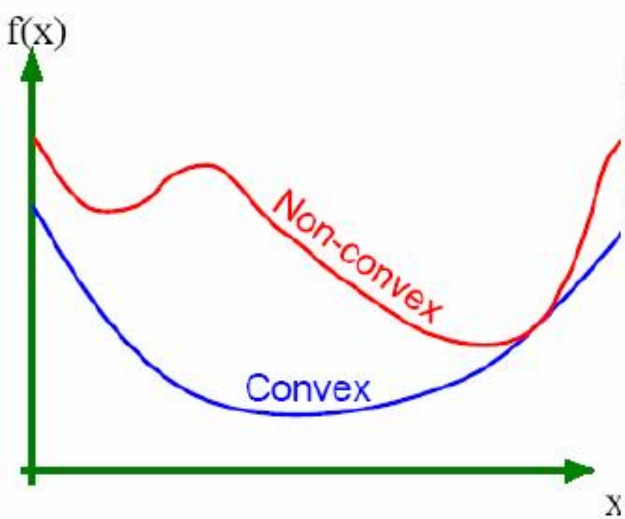
- Yeter sart: The Hessian matris pozitif tanimli olmali (H>0 olmali)

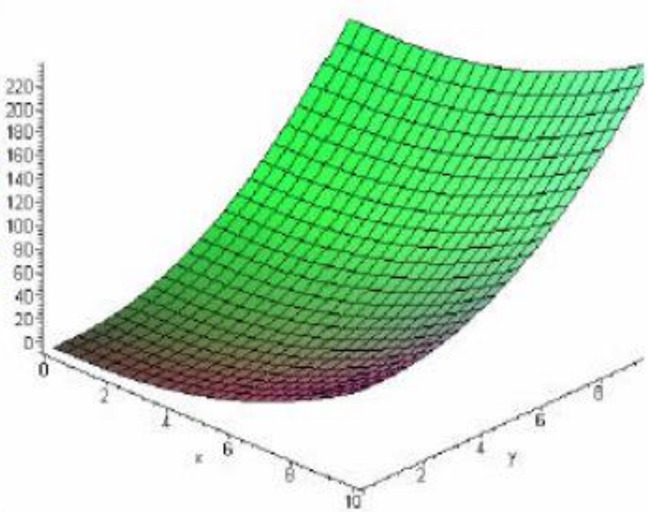
$$H = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdot & \cdot & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ & \frac{\partial^2 f}{\partial x_2^2} & & & \\ & & \cdot & & \\ & & & \cdot & \\ symm. & & & & \frac{\partial^2 f}{\partial x_n^2} \end{bmatrix}$$

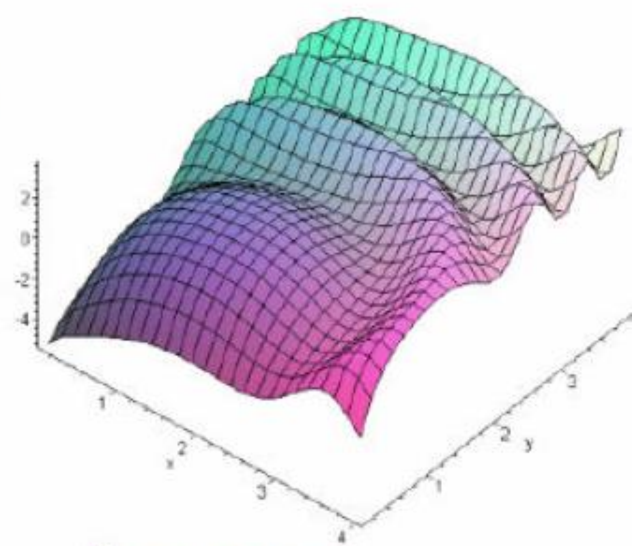


KISITSIZ PROBLEMLER VE KONVEKSLIK DURUMU

- Bir konveks fonksiyonun sadece tek bir minimumu vardir, o da global minimumdur.
- 1D problemlerde, konveks fonksiyonun her yerde 2. türevi pozitiftir.
- 1D problemlerde, konveks fonksiyon yukari dogru bir egri olusturur.
- 2D problemlerde, konveks fonksiyonun Hessian matrisi pozitif tanimlidir
- 2D problemlerde, konveks fonksiyon agzi yukari çukur bir kap seklindedir.




Convex


Non-convex

OPTIMALLIK SARTLARI (Kisitli Problemler)

- Kisitli problemlerde bir noktanin optimum olmasi için, Lagrangian fonksiyonun Kuhn-Tucker sartlarini saglamasi gerekir:

Lagrangian: $L(\lambda) = f(x) + \sum_{i=1}^{m} \lambda_i g_i(x), \; where \; \lambda_i > 0$

$$\frac{\partial L(\lambda)}{\partial x_i} = 0 \Rightarrow \nabla f(x) + \sum_{i=1}^{m} \lambda_i \nabla g_i(x) = 0$$

$$\lambda_i g_i = 0, \; (Complementary\ Condition)$$

$$g_i(x) \le 0, \; \lambda_i \ge 0, \; (Feasibility\ Condition)$$

Kaynaklar

- Hasan KURTARAN, “Mühendislik tasarımların optimizasyonu”, 2005
- Edwin K.P. Chong Stanislaw H.Zak,“An Introduction to Optimization”, 2001, John Wiley&Sons
- Jorge Nocedal, Stephan J. Wright, “Numerical Optimization”, 1999, Springer-Verlag
- R. Fletcher, “Practical Methods of Optimization”, 1987, John Wiley&Sons
- David G.Luenberger, “Introduction to Linear and Nonlinear Programming”, 1987, Addison-Wesley
- S.G. Nash, A. Sofer, “Linear and Nonlinear programming”, 1996, McGraw Hill
- Dimitri P. Bertsekas, “Constrained Optimization and Lagrange Multiplier Methods”, 1982, Academic Press