T.C.
SAKARYA ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

KONVEKSLİK VE OPTİMİZASYON ÜZERİNE

YÜKSEK LİSANS TEZİ

Gümrah UYSAL

Enstitü Anabilim Dalı : MATEMATİK

Tez Danışmanı : Yrd. Doç. Dr. Hüseyin KOCAMAN

Haziran 2010

T.C.
SAKARYA ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

KONVEKSLİK VE OPTİMİZASYON ÜZERİNE

YÜKSEK LİSANS TEZİ

Gümrah UYSAL

Enstitü Anabilim Dalı : MATEMATİK

Bu tez 11 / 06/2010 tarihinde aşağıdaki jüri tarafından Oybirliği ile kabul edilmiştir.

Yrd. Doç. Dr. Hüseyin KOCAMAN	Yrd. Doç. Dr. Şevket GÜR	Doç. Dr. C. Kubat
Jüri Başkanı	**Üye**	**Üye**

TEŞEKKÜR

Yüksek lisans çalışmalarım boyunca bana destek olan danışmanım Sayın Yrd. Doç. Dr. Hüseyin KOCAMAN'a ve öğrenim hayatım boyunca benim yanımda olup desteklerini her zaman yanımda hissettiğim sevgili aileme teşekkürlerimi sunarım.

İÇİNDEKİLER

SİMGELER VE KISALTMALAR LİSTESİ

$\mathbb{R}^n$	: *n* boyutlu Öklid uzayı
$\|\cdot\|$	: Öklid normu
$x^t y$	: Öklid iç çarpımı
$\leq$	: küçük eşit
$\geq$	: büyük eşit
H	: hiper düzlem
ξ	:subgradyent
∇f	: f fonksiyonunun gradyenti
∂f	: subdiferansiyel
lim	: limit
H	: *Hessian* matris

ŞEKİLLER LİSTESİ

TABLOLAR LİSTESİ

ÖZET

Anahtar kelimeler: Konvekslik, Optimizasyon.

Bu çalışmada konveks kümelerin ve konveks fonksiyonların optimizasyonla ilişkisi üstünde durulmuştur. Tezin birinci bölümünde Konveks Programlama Problemi'nin öneminden bahsedilmiş olup tezin ikinci bölümünde fonksiyonel analiz ve topolojiden bilinen bazı temel tanım ve teoremlere yer verilmiştir. Üçüncü bölümde konveks kümelerin ve konveks fonksiyonların özellikleri incelenmiştir. Fonksiyonların ve kümelerin konveksliğini karakterize eden koşullar verilmiştir. Dördüncü bölümde ise konveks fonksiyonların konveks bir küme üzerinde minimumlaştırılmaları ve maksimumlaştırılmaları konularına yer verilmiştir.

ON CONVEXITY AND OPTIMIZATION

SUMMARY

Key Words: Convexity, Optimization.

In this thesis, the relationship between convexity and optimization is investigated. In the first part of the thesis the importance of Convex Programming Problem is explained. In the second part, the basic definitions and theorems known from functional analysis and topology are given. In the third part the properties of convex functions and sets are given. Also the conditions which characterize the convexity of the given function and set, are given. The fourth part of the thesis deals with the optimization of given convex function on a given convex set.

BÖLÜM 1. GİRİŞ

Optimizasyon Problemi, genel olarak reel değerli bir fonksiyonu verilen küme üzerinde maksimize veya minimize eden değeri sistematik olarak seçme problemi olarak nitelendirilebilir. Konveks Programlama Problemi reel değerli konveks bir fonksiyonu konveks bir küme üzerinde minimumlaştırma problemidir. Konveks Programlama, Doğrusal Olmayan Programlama'nın özel bir hali veya Lineer Programlamanın bir genelleştirilmiş biçimi olarak görülebilir. Bu problemin ilişkili olduğu en önemli teoremlerden biri Weierstrass Teoremidir. Weierstrass Teoremi, verilen küme kompakt ve boştan farklı olduğu durumda eğer fonksiyon da bu küme üzerinde sürekli ise fonksiyonun minimum değerine ulaşacağını garanti etmiştir. Konveks Programlama Problemi'nin en önemli özelliklerinden biri eğer yerel minimum varsa bu minimumun aynı zamanda mutlak minimum olduğudur. Bir çok alanda Konveks Programlama'dan faydalanılmaktadır. Örnek olarak İstatistik, Finans, Data Analizi ve Modelleme verilebilir.

Konveks kümelerin sınır noktalarında destek hiper düzleme sahip olmaları subgradyent kavramının çıkış noktasını oluşturmuştur. Klasik analizde teğet doğrular ve teğet düzlemlerle verilen türev kavramının bir genelleştirilmiş biçimi olarak görülen subgradyent kavramı, Konveks Analiz'de destek hiper düzlemlerle verilir. Bu çalışmada konveks programlama probleminde fonksiyonun optimuma ulaştığı noktada sağlaması gereken koşul subgradyent kavramı üzerinden verilmiştir.

Kompakt, konveks bir kümede tanımlı konveks bir fonksiyonun, maksimum değerine kümenin sınır noktalarında ulaşması Harmonik Fonksiyonlar Teorisi, Potansiyel Teori ve Kısmi Diferansiyel Denklemler Teorisi'nde büyük öneme sahiptir.

BÖLÜM 2. TEMEL TANIM VE TEOREMLER

2.1. Giriş

Bu bölümde sonlu boyutlu $\mathbb{R}^n$ uzayının bileşen toplama ve skalerle çarpma işlemlerine göre lineer bir uzay olduğu ve üzerinde tanımlanan Öklid normu ile bir normlu uzay teşkil ettiği, bu normun aynı zamanda metrik özelliklerini sağladığı, akabinde ise üzerinde tanımlanan iç çarpım ile beraber $\mathbb{R}^n$ uzayının bir iç çarpım uzayı olduğu gösterilmiştir. $\mathbb{R}^n$ uzayının elemanları noktalar olarak da, başlangıcından bu noktalara vektörler olarak da göz önüne alınabilir. Bölümün devam eden kısımlarında fonksiyonel analizden bilinen bazı temel tanım ve teoremlere yer verilmiştir.

2.2. Tanım ve Teoremler

Tanım 2.2.1.(Lineer Uzay)

L boş olmayan bir küme ve F reel veya kompleks sayılar cismi olsun. Aşağıdaki şartlar sağlanıyorsa L ye F üzerinde lineer uzay veya vektör uzayı denir [1].

A. L,+ işlemine göre değişmeli bir gruptur. Yani;

1. $\forall x, y \in L$ için $x + y \in L$ dir.

2. $\forall x, y, z \in L$ için $x + (y + z) = (x + y) + z$ dir.

3. $\forall x \in L$ için $x + \theta = \theta + x$ olacak şekilde $\theta \in L$ vardır.

4. $\forall x \in L$ için $x + (-x) = (-x) + x = \theta$ olacak şekilde $-x \in L$ vardır.

5. $\forall x,y\in L$ için $x+y=y+x$ dir.

B. $x,y\in L$ ve $\alpha,\beta\in F$ olmak üzere aşağıdaki şartlar sağlanır:

6. $\alpha x\in L$ dir.

7. $\alpha(x+y)=\alpha x+\alpha y$ dir.

8. $(\alpha+\beta)x=\alpha x+\beta x$ dir.

9. $(\alpha\beta)x=\alpha(\beta x)$ dir.

10. $1x=x$ dir.

$L=\mathbb{R}^n=\{x=(x_1,...,x_n),\ x_i\in\mathbb{R},\quad i=1,2,...,n.\}$ olsun. $x,y\in\mathbb{R}^n$ ve $\alpha\in\mathbb{R}$ olmak üzere aşağıdaki işlemler tanımlansın

$$\begin{aligned}x+y&=(x_1,...,x_n)+(y_1,...,y_n)\\&=(x_1+y_1,...,x_n+y_n)\end{aligned}$$

ve

$$\alpha x=\alpha(x_1,...,x_n)=(\alpha x_1,...,\alpha x_n)\ .$$

$\mathbb{R}^n$ uzayı bileşen toplama ve skalerle çarpma işlemlerine göre lineer bir uzaydır.

1. $\forall x,y\in\mathbb{R}^n$ ve $x=(x_1,...,x_n),\ y=(y_1,...,y_n)$ olsun.

$x+y=(x_1+y_1,...,x_n+y_n)$ ve $x_i+y_i\in\mathbb{R},\quad i=1,2,...,n.$ olduğundan $x+y\in\mathbb{R}^n$ dir.

2. $z=(z_1,...,z_n)\in\mathbb{R}^n$ olsun.

$$\begin{aligned}x+(y+z)&=(x_1,...,x_n)+((y_1,...,y_n)+(z_1,...,z_n))\\&=(x_1,...,x_n)+(y_1+z_1,...,y_n+z_n)\\&=(x_1+y_1+z_1,...,x_n+y_n+z_n)\\&=((x_1+y_1)+z_1,...,(x_n+y_n)+z_n)\\&=((x_1+y_1),...,(x_n+y_n))+(z_1,...,z_n)\\&=(x+y)+z\end{aligned}$$

3. $\theta=(0,...,0)\in\mathbb{R}^n$ olarak alınırsa $\forall x\in\mathbb{R}^n$ için $x+\theta=\theta+x=x$ olur.

4. $-x = (-x_1, ..., -x_n)$ olarak alınırsa

$$\begin{aligned} x + (-x) &= (x_1, ..., x_n) + (-x_1, ..., -x_n) \\ &= (0, ..., 0) \\ &= \theta \end{aligned}$$

olduğu görülür.

5. $\mathbb{R}$ nin değişme özelliği kullanılırsa istenen sonuç elde edilir.

$$\begin{aligned} x + y &= (x_1, ..., x_n) + (y_1, ..., y_n) \\ &= (x_1 + y_1, ..., x_n + y_n) \\ &= (y_1 + x_1, ..., y_n + x_n) \\ &= y + x \end{aligned}$$

6. $\alpha \in \mathbb{R}$ için

$$\alpha x = \alpha(x_1, ..., x_n) = (\alpha x_1, ..., \alpha x_n)$$

ve $\alpha x_i \in \mathbb{R}$ olduğundan $\alpha x \in \mathbb{R}^n$ olarak bulunur.

7. $\alpha \in \mathbb{R},\ \ x, y \in \mathbb{R}^n$ için

$$\begin{aligned} \alpha(x + y) &= \alpha(x_1 + y_1, ..., x_n + y_n) \\ &= (\alpha(y_1 + x_1), ..., \alpha(y_n + x_n)) \\ &= (\alpha x_1 + \alpha y_1, ..., \alpha x_n + \alpha y_n) \\ &= (\alpha x_1, ..., \alpha x_n) + (\alpha y_1, ..., \alpha y_n) \\ &= \alpha x + \alpha y. \end{aligned}$$

olur.

8. $\alpha, \beta \in \mathbb{R},\ \ x \in \mathbb{R}^n$ için

$$\begin{aligned} (\alpha + \beta)x &= (\alpha + \beta)(x_1, ..., x_n) \\ &= ((\alpha + \beta)x_1, ..., (\alpha + \beta)x_n) \\ &= (\alpha x_1 + \beta x_1, ..., \alpha x_n + \beta x_n) \\ &= (\alpha x_1, ..., \alpha x_n) + (\beta x_1, ..., \beta x_n) \\ &= \alpha x + \beta x. \end{aligned}$$

9. $\alpha, \beta \in \mathbb{R},\ \ x \in \mathbb{R}^n$ için

$$
\begin{aligned}
(\alpha\beta)x &= (\alpha\beta)(x_1,...,x_n) \\
&= ((\alpha\beta)x_1,...,(\alpha\beta)x_n) \\
&= (\alpha\beta x_1,...,\alpha\beta x_n) \\
&= (\alpha(\beta x_1),...,\alpha(\beta x_n)) \\
&= \alpha(\beta x_1,...,\beta x_n) \\
&= \alpha(\beta x).
\end{aligned}
$$

10. $1 \in \mathbb{R},\ \ x \in \mathbb{R}^n$ için

$$
\begin{aligned}
1x &= 1(x_1,...,x_n) \\
&= (1x_1,...,1x_n) \\
&= x.
\end{aligned}
$$

Tanım 2.2.2.(Normlu Uzay)

N bir lineer uzay olsun. $\|.\|: N \to \mathbb{R}$ fonksiyonunun x deki değeri $\|x\|$ ile gösterilsin. Bu fonksiyon için

(N1) $\|x\| = 0 \Leftrightarrow x = \theta$,

(N2) $\|ax\| = |a|\|x\| \quad \alpha \in F$ ve

(N3) $\|x+y\| \le \|x\| + \|y\|$

şartları sağlanıyorsa $\|.\|$ fonksiyonuna N üzerinde norm denir. Normlu uzaylar genellikle $(N, \|.\|)$ ile gösterilir [1].

$\mathbb{R}^n$ uzayı üzerinde tanımlanan $\|x\| = \sqrt{|x_1|^2 + ... + |x_n|^2} = (\sum_{i=1}^{n} |x_i|^2)^{\frac{1}{2}}$ normuyla beraber normlu bir uzaydır. Bu norm Öklid normu veya 2 normu olarak bilinir.

(N1) $\forall x \in \mathbb{R}^n$ için

$$
\begin{aligned}
\|x\| = (\sum_{i=1}^{n} |x_i|^2)^{\frac{1}{2}} = 0 \Rightarrow \sum_{i=1}^{n} |x_i|^2 = 0 \\
\Rightarrow |x_i|^2 = 0 \ \ \forall i \Rightarrow x_i = 0 \ \ \forall i \\
\Rightarrow x = 0.
\end{aligned}
$$

(N2) $\forall x \in \mathbb{R}^n$, $\alpha \in \mathbb{R}$ için

$$\|\alpha x\| = (\sum_{i=1}^{n} |\alpha x_i|^2)^{\frac{1}{2}} = (\sum_{i=1}^{n} |\alpha^2 x_i^2|^{\frac{1}{2}})$$
$$= |\alpha| (\sum_{i=1}^{n} |x_i^2|)^{\frac{1}{2}}$$
$$= |\alpha| \|x\|.$$

Lemma 2.2.1. $x = (x_1, ..., x_n)$ ve $y = (y_1, ..., y_n)$ reel veya kompleks sayıların iki $n-$lisi olsun. Bu takdirde

$\sum_{i=1}^{n} |x_i y_i| \leq (\sum_{i=1}^{n} |x_i|^2)^{\frac{1}{2}} (\sum_{i=1}^{n} |y_i|^2)^{\frac{1}{2}}$ veya $\sum_{i=1}^{n} |x_i y_i| \leq \|x\|_2 \|y\|_2$ dir. Bu eşitsizlik Cauchy eşitsizliği olarak bilinir.[1]

Lemma 2.2.1. den hareketle

(N3) $\forall x, y \in \mathbb{R}^n$ için

$$\|x + y\|_2^2 = \sum_{i=1}^{n} |x_i + y_i|^2 = \sum_{i=1}^{n} |x_i + y_i| |x_i + y_i|$$
$$\leq \sum_{i=1}^{n} |x_i + y_i| (|x_i| + |y_i|)$$
$$\leq \sum_{i=1}^{n} |x_i| |x_i + y_i| + \sum_{i=1}^{n} |y_i| |x_i + y_i|$$
$$\leq \|x + y\|_2 \|x\|_2 + \|x + y\|_2 \|y\|_2$$

elde edilir. En son eşitsizlikte Cauchy eşitsitsizliğinden faydalanılmıştır. Burada eğer $\|x + y\|_2 = 0$ ise (N3) özelliği sağlanmıştır. Aksi takdirde eşitsizliğin her iki tarafı $\|x + y\|_2$ ile bölünerek

$\|x + y\|_2 \leq \|x\|_2 + \|y\|_2$ elde edilir. Bu eşitsizlik Minkowski eşitsizliği olarak bilinir.

Tanım 2.2.3.(Metrik Uzay)

X boş olmayan bir küme olsun. $d : X \times X \to \mathbb{R}$ fonksiyonu için,

(M1) $d(x,y)=0 \Leftrightarrow x=y$

(M2) $d(x,y)=d(y,x)$

(M3) $d(x,y)\le d(x,z)+d(z,y)$

şartları sağlanıyorsa d ye X de bir metrik ve d ile birlikte X e metrik uzay denir. Genellikle (X,d) ile gösterilir [1].

$d:\mathbb{R}^n\times\mathbb{R}^n\to\mathbb{R}_+$, $\|.\|$ Öklid normu, $x=(x_1,...,x_n), y=(y_1,...,y_n)$ ve $z=(z_1,...,z_n)$, $d(x,y)=\|x-y\|=(\sum_{i=1}^{n}|x_i-y_i|^2)^{1/2}$, $x,y,z\in\mathbb{R}^n$ olmak üzere $\mathbb{R}^n$ bir metrik uzaydır.

(M1) $d(x,y)=0 \Leftrightarrow \|x-y\|=(\sum_{i=1}^{n}|x_i-y_i|^2)^{1/2}=0$

$$\sum_{i=1}^{n}|x_i-y_i|^2=0 \Leftrightarrow |x_i-y_i|=0,\ i=1,...,n. \Leftrightarrow x-y=0$$
$$\Leftrightarrow x=y$$

(M2)
$$\begin{aligned} d(x,y)&=\|x-y\|=(\sum_{i=1}^{n}|x_i-y_i|^2)^{1/2}\\ &=(\sum_{i=1}^{n}|-(y_i-x_i)|^2)^{1/2}\\ &=(\sum_{i=1}^{n}|y_i-x_i|^2)^{1/2}\\ &=\|y-x\|\\ &=d(y,x)\end{aligned}$$

(M3)
$$\begin{aligned} d(x,y)&=\|x-y\|=\|x-z+z-y\|\\ &\le(\sum_{i=1}^{n}|x_i-z_i|^2)^{1/2}+(\sum_{i=1}^{n}|z_i-y_i|^2)^{1/2}\\ &\le\|x-z\|+\|z-y\|\\ &=d(x,z)+d(z,y)\end{aligned}$$

Tanım 2.2.4.(İç Çarpım Uzayı)

$F=\mathbb{R}$ veya $\mathbb{C}$ olmak üzere X bir vektör uzayı olsun .

$\langle .,. \rangle,\ X\times X\to F$

dönüşümü aşağıdaki özelliklere sahip ise $\langle .,. \rangle$ ye X üzerinde bir iç çarpım, $(X,\langle .,. \rangle)$ ikilisine de iç çarpım uzayı denir [4].

1. $\forall x\in X$ için $\langle x,x\rangle \geq 0$ ve $\langle x,x\rangle =0\Leftrightarrow x=0$

2. $\forall x,y\in X$ için $\langle x,y\rangle =\overline{\langle y,x\rangle}$

3. $\forall x,y\in X$ ve $\alpha \in F$ için $\langle \alpha x,y\rangle =\alpha \langle x,y\rangle$

4. $\forall x,y,z\in X$ için $\langle x+y,z\rangle = \langle x,z\rangle + \langle y,z\rangle$

$\langle .,. \rangle,\ \mathbb{R}^n\times \mathbb{R}^n\to \mathbb{R}$ ve $\langle x,y\rangle =\sum_{i=1}^{n} x_i y_i$, $\forall x,y,z\in \mathbb{R}^n,\ \alpha\in\mathbb{R}$ olmak üzere $\mathbb{R}^n$ bir iç çarpım uzayıdır.

1. $$\langle x,x\rangle =\sum_{i=1}^{n} x_i x_i = \sum_{i=1}^{n} |x_i|^2 \geq 0$$

ve $\sum_{i=1}^{n} |x_i|^2 =0 \Leftrightarrow x_i=0\ \ \forall i$

2. $$\langle x,y\rangle =\sum_{i=1}^{n} x_i y_i =\sum_{i=1}^{n} y_i x_i =\langle y,x\rangle$$

3. $\langle \alpha x, y \rangle = \sum_{i=1}^{n} \alpha x_i y_i$

$$= \alpha \sum_{i=1}^{n} x_i y_i$$

$$= \alpha \langle x, y \rangle$$

4. $\langle x + y, z \rangle = \sum_{i=1}^{n} (x_i + y_i) z_i$

$$= \sum_{i=1}^{n} (x_i z_i + y_i z_i)$$

$$= \sum_{i=1}^{n} x_i z_i + \sum_{i=1}^{n} y_i z_i$$

$$= \langle x, z \rangle + \langle y, z \rangle$$

Tanım 2.2.5.(İnfimum ve Supremum) S elemanları reel sayılar olan bir küme olsun. Bu takdirde S kümesinin infimumu (en büyük alt sınırı) $\forall x \in S$ için $\alpha \leq x$ şartını sağlayan en büyük α skaleridir. $\inf\{x : x \in S\}$ ile gösterilir. S kümesinin supremumu (en küçük üst sınırı) $\forall x \in S$ için $\alpha \geq x$ şartını sağlayan en küçük α skaleridir. $\sup\{x : x \in S\}$ ile gösterilir [3].

Tanım 2.2.6.(Komşuluk) $x \in \mathbb{R}^n$ noktası ve $\delta > 0$ verilsin. $N_\delta(x) = \{y : \|y - x\| \leq \delta\}$ yuvarı $x \in \mathbb{R}^n$ noktasının δ komşuluğu olarak tanımlanır. Bu eşitsizlik bazen kesin olarak da kullanılır [3].

Tanım 2.2.7.(İç nokta) S, $\mathbb{R}^n$ nin bir alt kümesi ve $x \in S$ olsun. Eğer $\|y - x\| \leq \varepsilon \Rightarrow y \in S$ olacak şekilde $\varepsilon > 0$ varsa bu şartı sağlayan $x \in S$ noktalarına iç nokta, bu noktaların kümesine S kümesinin içi denir. $\operatorname{int} S$ ile gösterilir [3].

Tanım 2.2.8.(Sınır noktası) S, $\mathbb{R}^n$ nin bir alt kümesi olsun. $\forall \varepsilon > 0$ için $N_\varepsilon(x) = \{y : \|y - x\| \leq \varepsilon\}$, S kümesinde olan ve olmayan elemanlar içeriyorsa $x \in S$ noktalarına sınır noktalar, bu noktaların kümesine S kümesinin sınırı denir. ∂S ile gösterilir [3].

Tanım 2.2.9.(Kapanış) S , $\mathbb{R}^n$ in bir alt kümesi olsun. S kümesinin kapanışı $\bar{S}$ ile gösterilir. Eğer $\forall \varepsilon > 0$ için $N_\varepsilon(x) \cap S \neq \varnothing$ ise $x \in \bar{S}$ olur [3].

Tanım 2.2.10. (Açık küme ve kapalı küme) (X,d) bir metrik uzay olsun. $a \in X$ ve $r > 0$ olsun. X içinde a merkezli ve r yarıçaplı açık ve kapalı yuvarlar aşağıdaki gibi tanımlanır

$B_r(a) = \{x \in X : d(x,a) < r\}$ (açık yuvar)

$B_r(a) = \{x \in X : d(x,a) \leq r\}$ (kapalı yuvar)

$V \subseteq X$ kümesinin açık bir küme olması ancak ve ancak $\forall x \in V$ için $\exists \varepsilon > 0$ öyle ki $B_\varepsilon(x)$ açık yuvarının V kümesinde içerilmesi şartı ile sağlanır.

$E \subseteq X$ kümesinin kapalı bir küme olması ancak ve ancak $E^c = X \setminus E$ kümesinin açık bir küme olması ile sağlanır [2].

Tanım 2.2.11.(Metrik uzayda dizi ve alt dizi) (X,d) bir metrik uzay olmak üzere $f : \mathbb{N} \to X$ fonksiyonuna X içinde bir dizi adı verilir ve $f(n) = (x_n)$ ile gösterilir. (n_k), $\mathbb{N}$ içinde $n_k < n_{k+1}$, $k = 1,..$ ve $\lim\limits_{k\to\infty} n_k = \infty$ koşullarını sağlayan herhangi bir dizi olmak üzere (x_{n_k}), $k \in \mathbb{N}$ dizisine (x_n) dizisinin bir alt dizisi denir [4].

Tanım 2.2.12.(Normlu uzaylarda sınırlı küme ve dizi) $(X, \|.\|)$ normlu bir uzay ve $A \subset X$ olsun. $d(A) = \sup\{\|x-y\| : x, y \in A\} \geq 0$ tanımlansın. Burada $d(A)$ sayısı A nın çapıdır. Bu çap sonlu ise A ya sınırlı küme denir. X içinde (x_n) dizisine karşılık gelen noktalar kümesi sınırlı ise (x_n) dizisine sınırlı dizi adı verilir [4].

Tanım 2.2.13.(Normlu uzaylarda yakınsak dizi) $(X, \|.\|)$ normlu uzayı içinde bir dizi (x_n) ve $x_0 \in X$ olsun. Eğer $\lim\limits_{n\to\infty} \|x_n - x_0\| = 0$ ise (x_n) dizisi x_0 noktasına yakınsıyor denir ve $x_n \to x_0$ veya $\lim\limits_{n\to\infty} x_n = x_0$ olarak ifade edilir [4].

Teorem 2.2.1.(Bolzano -Weierstrass Teoremi)

$\mathbb{R}^n$ üzerinde sınırlı her dizi yakınsak bir alt diziye sahiptir [2].

Tanım 2.2.14.(Sürekli fonksiyon) S, $\mathbb{R}^n$ nin bir alt kümesi olsun. Eğer verilen $\forall \varepsilon > 0$ için $x \in S$ ve $\|x - \bar{x}\| < \delta \Rightarrow |f(x) - f(\bar{x})| < \varepsilon$ olacak şekilde $\delta > 0$ varsa $f : S \to \mathbb{R}$ fonksiyonu $\bar{x} \in S$ noktasında süreklidir denir [3].

Teorem 2.2.2.

(X, d) bir metrik uzay A, X in boş olmayan bir alt kümesi olsun ve $\bar{A}$, A kümesinin kapanışı olsun.

1) $x \in \bar{A}$ olması için gerek ve yeter şart $x_n \to x$ olacak şekilde A da bir (x_n) dizisinin olmasıdır.

2) A nın kapalı olması için gerek ve yeter şart $x_n \to x$ ve $(x_n) \in A$ olması halinde $x \in A$ olmasıdır [1].

BÖLÜM 3.

KONVEKS KÜMELER VE KONVEKS FONKSİYONLAR

3.1. Giriş

Bu bölümde sonlu boyutlu $\mathbb{R}^n$ uzayı üzerindeki boş olmayan konveks kümeler ve özelliklerinden bahsedilmiştir. Konveks kümelerin hiper düzlemle ayrılmaları ve sınır noktalarında hiper düzlemle desteklenmeleri ile ilgili teoremler verilmiştir. Devam eden kısımda ise konveks fonksiyonların özellikleri üzerinde durulmuştur.

3.2. Konveks Kümeler

Tanım 3.2.1.(Konveks Küme) C kümesinde bu kümeden alınan herhangi iki noktayı birleştiren doğru parçası kümenin içinde kalıyorsa yani, $\forall x_1, x_2 \in C$ ve $0 \le \lambda \le 1$ için $\lambda x_1 + (1-\lambda)x_2 \in C$ koşulu sağlanıyorsa C kümesi konveks bir kümedir denir [7].

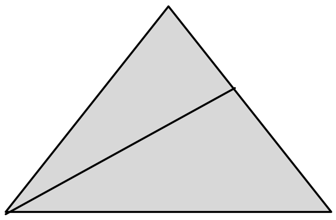
Örnek 3.2.1. Bazı konveks küme örnekleri aşağıdaki gibi verilebilir.

1. $C = \{(x_1, x_2, x_3) : x_1 + 4x_2 + 7x_3 = 3\} \subset \mathbb{R}^3$ bu küme $\mathbb{R}^3$ te bir düzlem denklemidir.

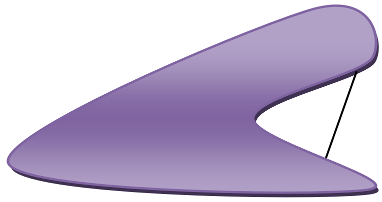
2. $C = \{(x_1, x_2) : x_1^2 + x_2^2 \le 2\} \subset \mathbb{R}^2$ bu küme $\mathbb{R}^2$ de kapalı bir diski ifade eder.

3. $C = \{x : Ax = b\}$ A, mxn matris ve b, m elemanlı bir vektör.

Tanım 3.2.2.(Afin Küme) $A \subseteq \mathbb{R}^n$ kümesinde $\forall x_1, x_2 \in A$ ve $\lambda \in \mathbb{R}$ için $\lambda x_1 + (1-\lambda)x_2 \in A$ koşulu sağlanıyorsa A kümesi afin bir kümedir denir [5]. Açıktır ki afin kümeler konvekstir. Boş küme ve tek nokta içeren kümeler afinlik ve konvekslik özelliğini birarada sağlayan kümelere örnek olarak verilebilirler.



Şekil 3. 1. Konveks bir küme



Şekil 3. 2. Konveks olmayan bir küme

Konveks kümelerin bazı özellikleri aşağıdaki gibi sıralanabilir [7]-[9].

$a \in \mathbb{R}$ ve $C_1, C_2 \subseteq \mathbb{R}^n$ iki konveks küme olmak üzere,

1. $C_1 + C_2 = \{x_1 + x_2 \mid x_1 \in C_1,\ x_2 \in C_2\}$, şeklinde tanımlanan $C_1 + C_2$ kümesi konveks bir kümedir.

2. $C_1 - C_2 = \{x_1 - x_2 \mid x_1 \in C_1,\ x_2 \in C_2\}$, şeklinde tanımlanan $C_1 - C_2$ kümesi konveks bir kümedir.

3. $aC_1 = \{ax_1 \mid x_1 \in C_1\}$ şeklinde tanımlanan aC_1 kümesi konveks bir kümedir.

4. Konveks kümelerin kartezyen çarpımı konveks bir kümedir .

$C_1 \times C_2 = \{(x_1, x_2) | x_1 \in C_1, x_2 \in C_2\}$

İspat 1: $x, y \in C_1 + C_2$ öyle ki $c_1, c_2 \in C_1$ ve $c_1', c_2' \in C_2$ için $c_1 + c_1' = x$ ve $c_2 + c_2' = y$ $\lambda \in [0,1]$ olsun.

$$\begin{aligned} \lambda x + (1-\lambda) y &= \lambda[c_1 + c_1'] + (1-\lambda)[c_2 + c_2'] \\ &= [\lambda c_1 + (1-\lambda) c_2] + [\lambda c_1' + (1-\lambda) c_2'] \\ &= c_1'' + c_2'' \end{aligned}$$

Bu eşitlik göz önüne alındığında C_1 ve C_2 kümelerinin de konveksliğinden faydalanılarak $\lambda x + (1-\lambda) y \in C_1 + C_2$ yazılır.

$\Rightarrow C_1 + C_2$ konvekstir.

İspat 2: $x, y \in C_1 - C_2$ öyle ki $c_1, c_2 \in C_1$ ve $c_1', c_2' \in C_2$ için $c_1 - c_1' = x$ ve $c_2 - c_2' = y$ $\lambda \in [0,1]$ olsun.

$$\begin{aligned}\lambda x + (1-\lambda)y &= \lambda[c_1 - c_1'] + (1-\lambda)[c_2 - c_2'] \\ &= [\lambda c_1 + (1-\lambda)c_2] - [\lambda c_1' + (1-\lambda)c_2'] \\ &= c_1'' - c_2''\end{aligned}$$

Bu eşitlik göz önüne alındığında C_1 ve C_2 kümelerinin de konveksliğinden faydalanılarak $\lambda x + (1-\lambda)y \in C_1 - C_2$ yazılır.

$\Rightarrow C_1 - C_2$ konvekstir.

İspat 3: $ax \in aC_1$, $x, y \in C_1$ ve $\lambda \in [0,1]$ olsun.

$$\lambda ax + (1-\lambda)ay = a(\lambda x + (1-\lambda)y)$$

Bu eşitlik göz önüne alındığında C_1 kümesinin de konveksliğinden faydalanılarak $\lambda x + (1-\lambda)y \in aC_1$ yazılır.

$\Rightarrow aC_1$ konvekstir.

İspat 4: $(x, x'), (y, y') \in C_1 \times C_2$ ve $\lambda \in [0,1]$ olmak üzere

$$\begin{aligned}\lambda(x, x') + (1-\lambda)(y, y') &= (\lambda x, \lambda x') + ((1-\lambda)y, (1-\lambda)y') \\ &= (\lambda x, (1-\lambda)y, \lambda x' + (1-\lambda)y') \\ &= (x'', y'')\end{aligned}$$

olur.

Kartezyen çarpımın tanımı gereği $x, y \in C_1$ ve $x', y' \in C_2$ olduğundan ve kümelerin

konveksliği düşünülürse $x'' \in C_1$ ve $y'' \in C_2$ olur. $\lambda(x, x') + (1-\lambda)(y, y') \in C_1 \times C_2$ olduğundan $C_1 \times C_2$ konvekstir.

Teorem 3.2.1. Konveks kümelerin keyfi koleksiyonlarının kesişimi konvekstir [5].

İspat:

$x, y \in \bigcap_i C_i \Rightarrow x, y \in C_i, \ \forall i$

C_i kümeleri konveks ise

$\forall \lambda \in [0,1], \ \lambda x + (1-\lambda) y \in C_i, \ \forall i \Rightarrow \lambda x + (1-\lambda) y \in \bigcap_i C_i$

$\Rightarrow \bigcap_i C_i$ konvekstir.

Bu ifade birleşim için söylenemez konveks kümelerin birleşimi konveks olmayabilir.

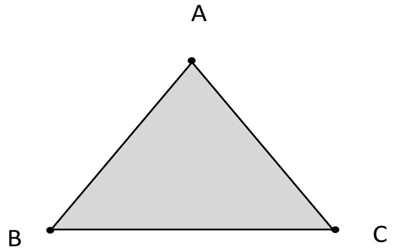
Örnek 3.2.2. $C_1 = \{1\}$ ve $C_2 = \{2\}$ olsun. Açıktır ki iki küme de konvekstir. Fakat

$C_1 \cup C_2 = \{\{1\}, \{2\}\}$ konveks değildir.

Tanım 3.2.3. (Konveks Kombinasyon) $B \subseteq \mathbb{R}^n$ bir küme olmak üzere $\theta_1 + ... + \theta_k = 1$ ve $\theta_i \geq 0$ $i = 1,2,...,k$ olduğu durumda $\theta_1 x_1 + ... + \theta_k x_k$ formundaki nokta konveks kombinasyon adını alır, $x_1,...,x_k$ kümeden alınan noktalardır [7].

Tanım 3.2.4. (Konveks Örten) B kümesinin noktalarının tüm konveks kombinasyonlarından oluşan kümeye B kümesinin konveks örteni adı verilir ve $convB$ ile gösterilir [7].

$$convB = \{\theta_1 x_1 + ... + \theta_k x_k \mid x_i \in B, \ \theta_i \geq 0, \ i = 1,2,...,k, \ \theta_1 + ... + \theta_k = 1\}$$



Şekil 3. 3. {A, B, C} kümesinin konveks örteni

Tanım 3.2.5. (Ekstrem Nokta) $C \subseteq \mathbb{R}^n$ konveks kümesindeki bir z noktası eğer $z \in [x, y] \Rightarrow z = x$ veya $z = y$ koşulunu $x, y \in C$ için sağlıyorsa ekstrem nokta adını alır. Burada $[x, y]$ kümenin içinde kalan doğru parçasını temsil etmektedir. Ekstrem noktaların kümesi $eks(C)$ ile gösterilir [9].

Teorem 3.2.2. K konveks bir küme ise bu kümenin noktalarının tüm konveks kombinasyonları kümeye aittir [6].

İspat: $x_1,...,x_m \in K$ olduğu durumda $\lambda_1+...+\lambda_m=1$ ve $\lambda_i \geq 0$ $i=1,2,...,m$ için $\forall \lambda_1 x_1+...+\lambda_k x_k \in K$ olduğu tümevarımla gösterilecektir.

$m=1$ için $\lambda_1=1$ ve $x_1 \in K.$

$m-1$ için doğru olduğu kabul edilsin , m için aşağıdaki ifade yazılır;

$$\sum_{i=1}^{m} \lambda_i x_i = \lambda_m x_m + (1-\lambda_m)\sum_{i=1}^{m-1} \frac{\lambda_i}{(1-\lambda_m)} x_i$$

bu eşitlik üzerinde inceleme yapılırsa $x_i \in K$, $0 \leq \lambda_m \leq 1$ ve $1-\lambda_m = \sum_{i=1}^{m-1} \lambda_i$ olduğundan dolayı $\sum_{i=1}^{m-1} \frac{\lambda_i}{1-\lambda_m} = 1$ ve $\sum_{i=1}^{m-1} \frac{\lambda_i}{1-\lambda_m} \geq 0$.

$$\Rightarrow \sum_{i=1}^{m} \lambda_i x_i = \lambda_m x_m + (1-\lambda_m)\sum_{i=1}^{m-1} \frac{\lambda_i}{(1-\lambda_m)} x_i \in K.$$

Teorem 3.2.3. B kümesinin konveks örteni bu kümeyi içeren en küçük konveks kümedir [6].

İspat: K, B kümesinin konveks örteni olsun T ise B kümesini içeren başka bir konveks küme olsun. Amacımız $K \subseteq T$ olduğunu ispatlamaktır. K kümesinden

$\lambda_1+...+\lambda_m=1$ ve $\lambda_i \geq 0$ için $x=\sum_{i=1}^{m} \lambda_i x_i$ elemanını alalım. T , B kümesini içerdiğinden $x_i \in T$ ve T konveks olduğundan $x \in T$ olur. x elemanı K kümesinin herhangi bir elemanı olarak seçildiğinden $K \subseteq T$ olduğu doğrudur.

Teorem 3.2.4. X bir konveks küme ve $\operatorname{int} X \neq \varnothing$ olsun. $x_1, x_2 \in X$ öyle ki $x_1 \in X, x_2 \in \operatorname{int} X$ olsun. $D(x_1,x_2)$, x_1 elemanını x_2 ye birleştiren doğru parçasını ifade etmek üzere, bu doğru parçası üzerinde alınan her nokta, belki x_1 dışında, $\operatorname{int} X$ kümesine dahildir [8].

İspat: X konveks olduğundan $D(x_1, x_2)$ üzerinde alınan her nokta X kümesine dahildir.

$x_2 \in \operatorname{int} X$ olduğundan $B_\delta(x_2) \subset X$ olacak şekilde $\delta > 0$ vardır. $y \in D(x_1, x_2), y \neq x_1$ olsun.

Bu takdirde y, $\lambda \geq 0, \mu > 0$ ve $\lambda + \mu = 1$ olmak üzere $y = \lambda x_1 + \mu x_2$ biçiminde yazılabilir.

$z \in B_{\mu\delta}(y)$ alınsın. $\|z - y\| < \mu\delta$ veya $\|z - (\lambda x_1 + \mu x_2)\| < \mu\delta$ eşitsizlikleri sağlanır.

$\mu > 0$ olduğundan son eşitsizlik μ ile bölünerek aşağıdaki eşitsizlik elde edilir.

$\|((z - \lambda x_1)/\mu) - x_2\| < \delta$. $B_\delta(x_2) \subset X$ olduğu göz önüne alınırsa $(z - \lambda x_1)/\mu \in X$

sonucuna erişilir. X konveks olduğundan ve z, $z = \mu((z - \lambda x_1)/\mu) + \lambda x_1$ formunda yazılabileceğinden $z \in X$ olur. $B_{\mu\delta}(y) \subset X$ ve $y \in \operatorname{int} X$. Burada

seçilen $x_1 \in X$, $x_2 \in \operatorname{int} X$ noktaları keyfi olduğundan $\operatorname{int} X$ konvekstir sonucuna da varılır.

Teorem 3.2.5. $S \subseteq \mathbb{R}^n$ konveks bir küme olsun. Bu takdirde $\overline{S}$ konvekstir [9].

İspat: $x, y \in \overline{S}$ verilsin. $z = (1-\lambda)x + \lambda y \in \overline{S}$ $\lambda \in [0,1]$ olduğu ispatlanmalıdır.

Teorem 2.2.2. den dolayı (x_n) ve (y_n) yakınsak dizileri $x_n, y_n \in S$ iken $x_n \to x$

$y_n \to y$, $n \to \infty$ olacak şekilde vardır. $(1-\lambda)x_n + \lambda y_n$ dizisi tanımlansın. S konveks olduğundan bu dizi S kümesinin elemanlarından oluşan bir dizidir. (x_n) ve (y_n) yakınsak olduğundan $(1-\lambda)x_n + \lambda y_n \to z$, $n \to \infty$ elde edilir. z, S nin elemanlarından oluşan bir dizinin limiti olduğundan $z \in \overline{S}$ olur , $\overline{S}$ konvekstir.

3.3. Ayırma Teoremleri

Tanım 3.3.1. (Hiper Düzlem) $\mathbb{R}^n$ uzayında bir (afin) hiperdüzlem aşağıdaki biçimde verilir:

$$H = \{v \in \mathbb{R}^n \,|\, a^t v = \beta\}$$

Bu hiper düzlem $\mathbb{R}^n$ uzayını iki yarı uzaya ayırır:

$$H_{\leq} = \{v \in \mathbb{R}^n \,|\, a^t v \leq \beta\}$$

$$H_{\geq} = \{v \in \mathbb{R}^n \,|\, a^t v \geq \beta\}$$

$S_1, S_2 \subset \mathbb{R}^n$ verilsin. Eğer $S_1 \subseteq H_{\leq}$ ve $S_2 \subseteq H_{\geq}$ ise H S_1 ve S_2 kümelerini ayırıyor denir.

$S_1, S_2 \subset \mathbb{R}^n$ verilsin. $H_{<} = \{v \in \mathbb{R}^n \,|\, a^t v < \beta\}$ ve $H_{>} = \{v \in \mathbb{R}^n \,|\, a^t v > \beta\}$ olmak üzere eğer $S_1 \subseteq H_{<}$ ve $S_2 \subseteq H_{>}$ ise H S_1 ve S_2 kümelerini kesin ayırıyor denir [9].

Lemma 3.3.1. A, $\mathbb{R}^n$ uzayının kapalı bir alt kümesi olarak verilsin. Bu takdirde $\|x_0\| = \inf\{\|x\| \;|\, x \in A\}$ olarak tanımlanan $x_0 \in A$ vardır. A kümesi konveks ise x_0 tektir [9].

İspat: $\beta = \inf\{\|x\| \;|\, x \in A\}$ olsun. Öncelikle A kümesinin sınırlı olduğu kabul edilsin. $x_n \in A$ olmak üzere $\|x_n\| - \beta < \frac{1}{n}$, $n \in \mathbb{N}_+$ özelliğini sağlayan (x_n) dizisi tanımlansın. Böyle bir dizi infimumun tanımından dolayı vardır. A kümesi sınırlı olduğundan dolayı elemanlarını bu kümeden alan (x_n) dizisi de sınırlıdır ve yakınsak (x_{n_i}) alt dizisine sahiptir. x_0 bu dizinin limiti olsun.

A kümesi kapalı olduğundan dolayı $x_0 \in A$ olur. Normun sürekliliği gereğince $\lim_{i\to\infty} \|x_{n_i}\| - \beta = \lim_{i\to\infty} \frac{1}{n_i} = 0$, $\left\| \lim_{i\to\infty} x_{n_i} \right\| = 0 \Rightarrow \|x_0\| = \beta$ elde edilir. Bu da x_0 ın varlığını kanıtlar. A kümesi konveks bir küme ve $y_0 \in A$, $\|x_0\| = \|y_0\|$ koşulunu sağlayan diğer bir eleman olsun. $z = 1/2(x_0 + y_0) = 1/2x_0 + 1/2y_0 \in A$ elemanı alınsın. Bu takdirde $\|z\| = 1/2\|x_0 + y_0\| \leq 1/2\|x_0\| + 1/2\|y_0\| = \|x_0\|$ durumu oluşur bu da x_0 ın en küçük normlu eleman tanımıyla çelişir. $\|x_0 + y_0\| = \|y_0\| + \|x_0\|$ dır. Bunun anlamı $x_0 = ky_0$ olacak şekilde $k \in \mathbb{R}$ vardır. $\|x_0\| = \|y_0\|$ kabülünden dolayı $k = \pm 1$ bulunur. Her iki durumda da $x_0 = y_0$ elde edilir.

Lemma 3.3.2. C, $\mathbb{R}^n$ uzayının kapalı ve konveks bir alt kümesi olarak verilsin ve $\|x_0\| = \inf\{\|x\| \mid x \in C\}$, $x_0 \in C$ olarak tanımlansın. Eğer $0 \notin C$ ise $x_0{}^t z > \|x_0\|^2 / 2$ eşitsizliği $\forall z \in C$ için sağlanır [9].

İspat: Lemma 3.3.1. gereğince $x_0 \in C$ vardır ve tektir. Farzedelim ki $0 \notin C$ olsun ve $x_0{}^t z \leq \|x_0\|^2 / 2$ eşitsizliği $z \in C$ için sağlansın. x_0 noktasının tanımı ve C kümesinin konveksliği kullanılırsa $0 \leq \lambda \leq 1$ için aşağıdaki eşitsizlik elde edilir.

$$\begin{aligned}\|x_0\|^2 &\leq \|(1-\lambda)x_0 + \lambda z\|^2 = (1-\lambda)^2\|x_0\|^2 + 2\lambda(1-\lambda)x_0{}^t z + \lambda^2 z^2 \\ &\leq (1-\lambda)^2\|x_0\|^2 + (1-\lambda)\lambda\|x_0\|^2 + \lambda^2 z^2\end{aligned}$$

Son eşitsizlikte $x_0{}^t z \leq \|x_0\|^2 / 2$ kabülünden faydalanılmıştır. Eşitsizliğin her iki tarafından $\|x_0\|^2$ çıkarılırsa,

$$0 \leq -2\lambda\|x_0\|^2 + (1-\lambda)\lambda\|x_0\|^2 + \lambda^2\|x_0\|^2 + \lambda^2 z^2 = \lambda(-\|x_0\|^2 + \lambda\|z\|^2)$$

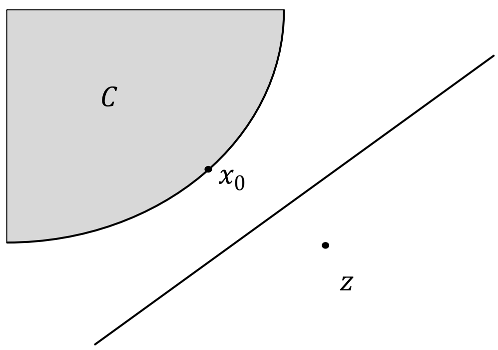
elde edilir. Bu eşitsizliğin her iki tarafı $\lambda > 0$ ile bölünürse $\forall 0 < \lambda \leq 1$ için $\|x_0\|^2 \leq \lambda\|z\|^2$ elde edilir. $\lim\limits_{\lambda \to 0}\|x_0\|^2 \leq \lambda\|z\|^2 \Rightarrow x_0 = 0$ olur ve bu da $0 \notin C$ kabülü ile çelişir. $x_0{}^t z > \|x_0\|^2 / 2$ eşitsizliği sağlanır.

Teorem 3.3.1. C, $\mathbb{R}^n$ uzayının kapalı ve konveks bir alt kümesi olarak verilsin ve $v \notin C$ olmak üzere $x_0 \in C$, v noktasına en yakın tek nokta olsun. Bu takdirde $(x_0 - v)^t (z - v) > \|x_0 - v\|^2 / 2$ eşitsizliği $\forall z \in C$ için sağlanır.

$H = \{x \in \mathbb{R}^n \;|\alpha^t x = \beta\}$ burada $\alpha = x_0 - v$ ve $\beta = (x_0 - v)^t v + \|x_0 - v\|^2 / 2$ olmak üzere , H hiper düzlemi v noktasını C kümesinden kesin ayırır [9].

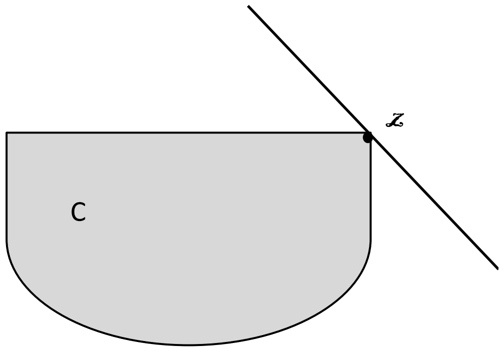
İspat: $C - v = \{x - v \;|x \in C\}$ olarak tanımlansın. Lemma 3.3.2. den faydalanabilmek için öncelikle bu kümenin kapalı ve konveks olduğu gösterilmelidir. $C - v$ kümesinin tümleyeninin açık olduğu gösterilecektir. $x^* \in (C - v)^c$ ve $x - v \in C - v$ olsun. $\varepsilon = \|x^* - (x - v)\|$ olsun. $B_\varepsilon(x^*) = \{x^{**} \in (C - v)^c \;|\|x^* - x^{**}\| < \varepsilon\}$ açık yuvarı tanımlansın. $x - v \in C - v$ noktası bu yuvara dahil değildir. Dolayısıyla $B_\varepsilon(x^*) \subseteq (C - v)^c$ olur. Sonuç olarak açık kümenin tanımı gereği $(C - v)^c$ açıktır ve $(C - v)$ kapalıdır. Konvekslik durumu ise konveks kümelerin özellikleri kısmında 2. özellikten söylenir. Diğer bir koşul ise $0 \notin C - v$ idi. $v \notin C$ olduğundan bu koşul sağlanır. Lemma 3.3.2. uyarınca $(x_0 - v)^t (z - v) > \|x_0 - v\|^2 / 2$ eşitsizliği $\forall z \in C$ için sağlanır. $(x_0 - v)^t (z - v) > \|x_0 - v\|^2 / 2 \Rightarrow (x_0 - v)^t z > \|x_0 - v\|^2 / 2 + (x_0 - v)^t v$

$H = \{x \in \mathbb{R}^n \;|\alpha^t x = \beta\}$ burada $\alpha = x_0 - v$ ve $\beta = (x_0 - v)^t v + \|x_0 - v\|^2 / 2$ dir. Kesin ayırmanın tanımı gereği H hiper düzlemi, v noktasını C kümesinden kesin ayırıyor denir.



Şekil 3.4. C kümesi ile z noktasının ayrılması

Tanım 3.3.2. (Destek Hiper Düzlem) $C \subseteq \mathbb{R}^n$ konveks kümesinin bir sınır noktası olan $z \in \partial C$ noktasındaki hiper düzlem kümenin destek hiper düzlemidir. $z \in H$ ve $C \subseteq H_{\le}$ özelliklerini sağlar. Burada $H_{\le}$ destek hiper uzaydır [9].



Şekil 3. 5. C kümesini z sınır noktasında destekleyen hiper düzlem

Teorem 3.3.2. $C \subseteq \mathbb{R}^n$ konveks bir küme ve $z \in \partial C$ olsun. Bu takdirde $C \subseteq \mathbb{R}^n$ konveks kümesini bir sınır noktası olan $z \in \partial C$ noktasında destekleyen bir hiper düzlem vardır [9].

İspat: $z \in \partial C$ ise sınır noktasının tanımı gereği bu noktaya kümenin kapanışının dışından alınan noktalarla yaklaşılabilir. Dolayısıyla $z_n \notin \overline{C}$ özelliğini sağlayan noktalar dizisi olan (z_n) vardır. $\overline{C}$ kapalı ve konveks olduğundan Lemma 3.3.1. gereği bu küme her noktaya en kısa mesafede olacak şekilde teklik koşulunu sağlayan nokta içerir. Her z_n noktası için $x_n \in \overline{C}, \ \forall n$ en yakın nokta olsun. $u_n = \dfrac{z_n - x_n}{\|z_n - x_n\|}$ olsun. Teorem 3.3.1. gereği $(u_n)^t (x - z_n) > \|z_n - x_n\| / 2, \ \forall x \in \overline{C}$ eşitsizliği sağlanır. (u_n) sınırlı bir dizi olduğundan yakınsak bir alt diziye sahiptir. Bu alt dizinin limiti u olsun. $\|z_n - x_n\| \leq \|z_n - z\| \to 0, n \to \infty$ olduğundan

$\lim\limits_{n\to\infty} (u_{n_i})^t (x - z_n) > \lim\limits_{n\to\infty} \|z_n - x_n\| / 2 \Rightarrow u^t x \geq u^t z$ elde edilir.

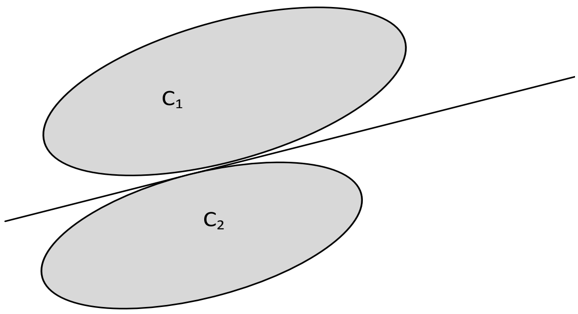
Kümenin sınır noktasından geçen destek hiper düzlem oluşturulmuş olup $H = \{x \in \mathbb{R}^n \ | u^t x = u^t z\}$ şeklinde ifade edilir. $u^t z = \beta$ olsun bu takdirde $H = \{x \in \mathbb{R}^n \ | u^t x = \beta\}$ şeklinde yazılır. Burada $\forall x \in C$ için $u^t x \leq \beta$ ve en azından bir tane $x \in \partial C$ için $u^t x = \beta$ olmaktadır. Hiper düzlemin tanım koşulları sağlanmıştır.

Teğet doğrular ve teğet düzlemler analizde türevlenebilme kavramından hareketle verilir. Ancak konveks analizde bu ayırma teoremleri ile açıklanmaktadır. Bu, genelleştirilmiş teğetlik olarak verilir. Genel olarak hem türevlenebilen hem de türevlenemeyen konveks fonksiyonlar için subdiferansiyel kavramı konveks kümelerin sınır noktalarından geçen destek hiper düzlemlerinin var oluşu aracılığı ile verilir.[5]

Teorem 3.3.3. $C_1, C_2 \subseteq \mathbb{R}^n$ ayrık ($C_1 \cap C_2 = \varnothing$) konveks alt kümeler olsun. Bu takdirde $C_1, C_2 \subseteq \mathbb{R}^n$ kümelerini ayıran $H = \{u \in \mathbb{R}^n \;|\; a^t u = \beta\}$ hiper düzlemi vardır [9].

İspat: $C_1 - C_2 = \{x - y \;|\; x \in C_1, y \in C_2\}$ olsun. $C_1 - C_2$ konvekstir ve kümeler ayrık olduğundan $0 \notin C_1 - C_2$ olur. Bu takdirde $\alpha^t(x-y) \geq 0$ koşulunu sağlayan $\alpha \in R^n \setminus \{0\}$ vardır. Bu durum $\alpha^t x \geq \alpha^t y \;\; \forall x \in C_1, y \in C_2$ şeklinde de ifade edilebilir. Eğer $0 \notin \overline{C_1 - C_2}$ ise (>) durumu Lemma 3.3.2. den ileri gelir. Eğer $0 \in \overline{C_1 - C_2}$ ise (=) durumu Teorem 3.3.2. den $z = 0$ ile sağlanır. Bu takdirde $\beta_1 \geq \beta_2$ olur burada β_1 ve β_2 aşağıdaki biçimde tanımlanır:

$\beta_1 = \inf\{a^t x \;|\; x \in C_1\}$, $\beta_2 = sup\{a^t y \;|\; y \in C_2\}$. Sonuç olarak istenen ayırıcı hiper düzlem $H = \{u \in \mathbb{R}^n \;|\; a^t u = \beta_1\}$ şeklinde tanımlanır.



Şekil 3. 6. İki konveks kümenin bir hiper düzlemle ayrılması

3.4. Konveks Fonksiyonlar

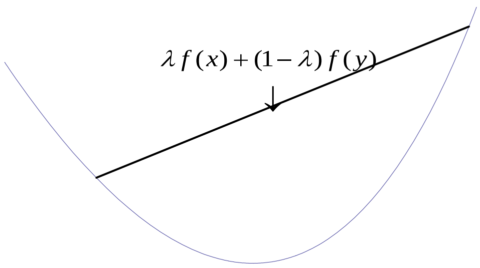
Tanım 3.4.1.(Konveks ve Konkav Fonksiyonlar)

$S \subseteq \mathbb{R}^n$ konveks bir küme ve $f : S \to \mathbb{R}$ bir fonksiyon olsun. f fonksiyonu $\forall x, y \in S$ ve $\forall \lambda \in [0,1]$ için aşağıdaki eşitsizliği sağladığı takdirde konveks bir fonksiyon adını alır.

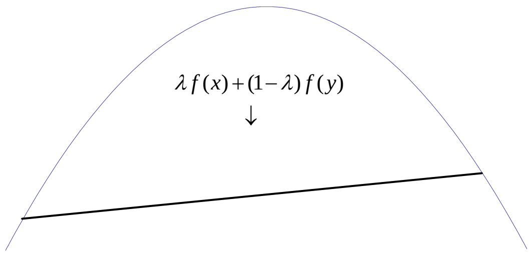
$$f(\lambda x + (1-\lambda) y) \leq \lambda f(x) + (1-\lambda) f(y)$$

yukarıdaki eşitsizlik $\forall \lambda \in (0,1)$ ve $x \neq y$ için kesin eşitsizlik ise fonksiyon kesin konvekstir denir. Bu tanım geometrik olarak yorumlanırsa bir fonksiyon konveks ise bu fonksiyonun grafiği $f(x)$ ve $f(y)$ noktalarını birleştiren kirişin altında kalır veya bu kiriş fonksiyon ile çakışır.

$-f$ fonksiyonu konveks bir fonksiyon ise f fonksiyonu konkav bir fonksiyondur denir. Ayrıca yukarıdaki eşitsizliğin sağ tarafı ve sol tarafı birbirine eşit ise fonsiyona afin bir fonksiyondur denir [7]-[11].



Şekil 3. 7. Konveks bir fonksiyon



Şekil 3.8. Konkav bir fonksiyon

Örnek 3.4.1.

Bazı konveks fonksiyon örnekleri aşağıdaki gibi verilmiştir.

1. $f : \mathbb{R} \to \mathbb{R},\ f(x) = 3x + 4$

2. $f:\mathbb{R}^2 \to \mathbb{R},\ f(x_1,x_2) = 2x_2^2 + x_2^2$

3. $f:\mathbb{R} \to \mathbb{R}_+,\ f(x) = \|x\|$

Her üç fonksiyon da $\mathbb{R}^n$ üzerinde konvekstir. Eklemek gerekir ki belli bir bölgede konveks olup da $\mathbb{R}^n$ üzerinde konveks olmayan fonksiyonlar da tanımlanabilir. Örnek olarak $f:\mathbb{R} \to \mathbb{R},\ f(x) = x^3$, $\mathbb{R}$ üzerinde konveks olmayıp $S = \{x : x \geq 0\}$ üzerinde konvekstir [3].

Konveks fonksiyonların bazı özellikleri aşağıdaki gibi verilebilir [7]-[11].

$S \subseteq \mathbb{R}^n$ bir küme ve $f:S \to \mathbb{R}$, $g:S \to \mathbb{R}$ konveks fonksiyonlar olsun bu takdirde

1. $\alpha f + \beta g$ fonksiyonu $\alpha, \beta \geq 0$ olduğu durumda, konvekstir.

2. $g_i(x)$ fonksiyonlarının $i = 1,2,..,n.$ için konveks olduğu durumda $\sum_{i=1}^{n} g_i(x)$ fonksiyonu konveks bir fonksiyondur.

İspat 1: $x = \lambda x_1 + (1-\lambda)x_2$, $\alpha, \beta \geq 0$ olsun. Konveks fonksiyonun tanımını kullanırsak

$$
\begin{aligned}
(\alpha f + \beta g)(\lambda x_1 + (1-\lambda)x_2) &= \alpha f(\lambda x_1 + (1-\lambda)x_2) + \beta g(\lambda x_1 + (1-\lambda)x_2) \\
&\leq \alpha\lambda f(x_1) + (1-\alpha)\lambda f(x_2) + \beta\lambda g(x_1) + (1-\alpha)\beta g(x_2) \\
&= \lambda[\alpha f(x_1) + \beta g(x_1)] + (1-\lambda)[\alpha f(x_2) + \beta g(x_2)] \\
&= \lambda[(\alpha f + \beta g)(x_1)] + (1-\lambda)[(\alpha f + \beta g)(x_2)]
\end{aligned}
$$

$\Rightarrow \alpha f + \beta g$ konvekstir.

İspat 2: $x = \lambda x_1 + (1-\lambda)x_2$ olsun. $g_i(x)$ fonksiyonları $i = 1,2,..,n.$ için konveks olduğundan

$$g_i(\lambda x_1 + (1-\lambda)x_2) \leq \lambda g_i(x_1) + (1-\lambda)g_i(x_2)$$

yazılır. Bu eşitsizlikten

$$\sum_{i=1}^{n} g_i(\lambda x_1 + (1-\lambda)x_2) \leq \sum_{i=1}^{n}[\lambda g_i(x_1) + (1-\lambda)g_i(x_2)]$$

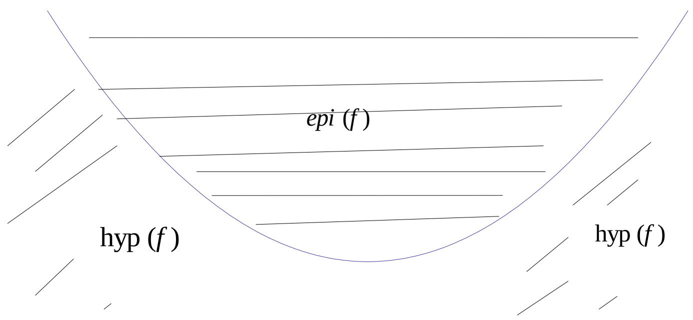
elde edilir. Sol tarafa toplama işleminin dağılma özelliği uygulanırsa bir fonksiyonun konvekslik koşulu elde edilmiş olur.

$$\sum_{i=1}^{n} g_i(\lambda x_1+(1-\lambda)x_2) \le \sum_{i=1}^{n} \lambda g_i(x_1) + \sum_{i=1}^{n} (1-\lambda) g_i(x_2)$$

Tanım 3.4.2. (Epigraf ve Hipograf) $V \subseteq \mathbb{R}^n$ boştan farklı bir küme ve $f:V \to \mathbb{R}$ bir fonksiyon olsun. f fonksiyonunun epigrafı ve hipografı

$$epi(f)=\{(x,w): x\in V, w\in\mathbb{R} \,|\, f(x)\le w\} \qquad hyp(f)=\{(x,w): x\in V, w\in\mathbb{R} \,|\, f(x)\ge w\}$$

şeklinde tanımlanır [10].



Şekil 3.9. Bir fonksiyonun epigrafı ve hipografı

Konveks bir fonksiyonla konveks bir kümenin ilişkisi fonksiyonun epigrafı aracılığı ile kurulmaktadır.

Teorem 3.4.1. $S \subseteq \mathbb{R}^n$ boştan farklı konveks bir küme ve $f:S \to \mathbb{R}$ bir fonksiyon olsun. Bu takdirde $f:S \to \mathbb{R}$ fonksiyonunun konveks bir fonksiyon olması yalnız ve ancak epigrafının konveks bir küme olmasıyla gerçeklenir [3].

İspat:

$\Rightarrow$ f konveks bir fonksiyon ise

$$f(\lambda x+(1-\lambda)y) \le \lambda f(x)+(1-\lambda) f(y) \quad \forall x,y \in S$$

olur. Epigrafın tanımından

$$
\begin{aligned}
f(\lambda x+(1-\lambda)y) &\le \lambda f(x)+(1-\lambda)f(y)\\
&\le \lambda w+(1\text{-}\lambda)\mathrm{w}\\
&\le w
\end{aligned}
$$

yazılır.

$\Rightarrow \ f(\lambda x+(1-\lambda)y)\le w$ olur. Bu durumda $(\lambda x+(1-\lambda)y,w)\in Epi(f)$ olur.

$0<\lambda<1$ olduğu göz önüne alınıp konveks kümenin tanımı kullanılırsa $Epi(f)$ konvekstir.

$\Leftarrow epi(f)$ konveks bir küme olsun ve epigraftan $(x,f(x)),(y,f(y))$ noktaları alınsın.

Bu noktaların kümede olduğu epigrafın tanımından dolayı açıktır.

$\lambda(x,f(x))+(1-\lambda)(y,f(y))\in epi(f) \quad \forall\lambda\in(0,1)$ bu durumda

$(\lambda x+(1-\lambda)y,\lambda f(x)+(1-\lambda)f(y))\in epi(f)$

olur. Epigrafın tanımı kullanılırsa aşağıdaki sonuca gidilir.

$f(\lambda x+(1-\lambda)y)\le\lambda f(x)+(1-\lambda)f(y)$

$\Rightarrow \ f$ konvekstir.

Tanım 3.4.3.(Subgradyent ve Subdiferansiyel)

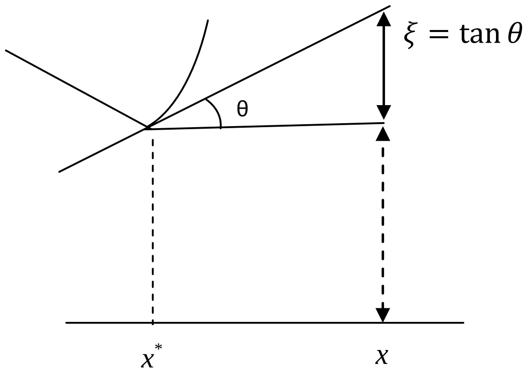
$S\subseteq\mathbb{R}^n$ boştan farklı konveks bir küme ve $f:S\to\mathbb{R}$ konveks bir fonksiyon olsun. Bu takdirde $\xi\in\mathbb{R}^n$ vektörü aşağıdaki eşitsizlik sağlandığı takdirde f fonksiyonunun $x^*\in S$ noktasındaki subgradyenti adını alır:

$f(x)\ge f(x^*)+\xi^t(x-x^*) \quad \forall x\in S$

$f(x)$ in x noktasındaki tüm subgradyentleri kümesine subdiferansiyel denir. ∂f ile gösterilir [3].

Geometrik olarak $f(x^*)+\xi^t(x-x^*)$ fonksiyonu, konveks fonksiyonun epigrafını $(x^*,f(x^*))\in\mathbb{R}^n\times\mathbb{R}$ noktasında destekleyen hiper düzleme karşılık gelmektedir.

ξ vektörü ise bu hiper düzlemin normalidir. Şekil 3.10. üzerinde inceleme yapılırsa subgradyent ξ bu destek hiper düzlemin eğimine karşılık gelmektedir.

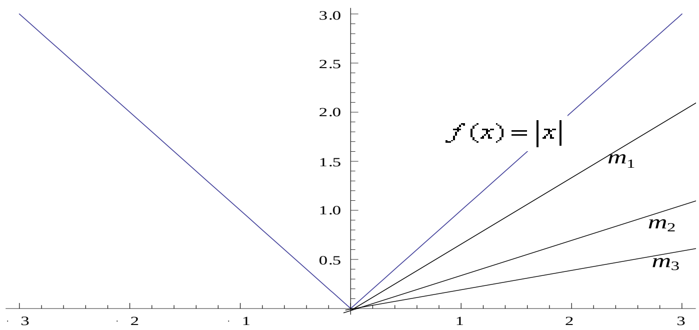


Şekil 3.10. Subgradyentin bir gösterimi

Örnek 3.4.2. $f : S \to \mathbb{R}_+$ ve $S = [-1,1]$ olmak üzere $f(x) = |x|$ olsun. $x^* = 0$ olarak seçilsin. Tanım uygulanırsa

$$|0| + c(x-0) = cx \le |x|$$
$$\forall c \in [-1,1] \quad \partial f(0) = [-1,1]$$

elde edilir. Geometrik olarak $(0,0)$ noktasından geçen ve eğimi $[-1,1]$ aralığında olan her doğrunun $f(x) = |x|$ fonksiyonunun grafiğinin altında kaldığını veya eşit olduğunu ifade etmektedir. Bu doğruların eğimleri, fonksiyonun (0,0) noktasındaki subgradyentleridir.



Şekil 3.11. Örnek 3.4.2. için çizim

Teorem 3.4.2. $S \subseteq \mathbb{R}^n$ boştan farklı konveks bir küme ve $f : S \to \mathbb{R}$ konveks bir fonksiyon olsun.

Bu takdirde $x^* \in \operatorname{int} S$ için ξ vektörü vardır. Öyle ki

$$H = \{(x, y) : y = f(x^*) + \xi^t (x - x^*)\}$$

hiper düzlemi $epi(f)$ yi $(x^*, f(x^*))$ sınır noktasında destekler. Ayrıca $f(x) \geq f(x^*) + \xi^t (x - x^*)$ $\forall x \in S$ eşitsizliği sağlanır ξ vektörü f fonksiyonunun x^* daki subgradyentidir [3].

İspat: Destek hiper düzlemin tanımı hatırlanırsa $H = \{x \in \mathbb{R}^n \; | u^t x = u^t z\}$ $\forall x \in C$ için $u^t (x - z) \leq 0$ ve en azından bir tane $x \in \partial C$ için $u^t x = u^t z$ olmakta idi.

Bu bağlamda düşünüldüğünde Teorem 3.3.2. gereği $(\xi_0, \alpha) \in \mathbb{R}^n \times \mathbb{R}$ sıfırdan farklı vektörü vardır öyle ki

$(\xi_0, \alpha)^t (x - x^*, y - f(x^*)) \leq 0$ $\forall x, y \in epi(f)$ veya

$\xi_0^{\,t} (x - x^*) + \alpha[y - f(x^*)] \leq 0$ $\forall x, y \in epi(f)$

eşitsizlikleri sağlanır. Burada eğer y yeterince büyük alınırsa eşitsizliğin yön değiştirebileceğinden dolayı $\alpha > 0$ değildir. $\alpha < 0$ durumunu kanıtlayalım. Farzedelim ki $\alpha = 0$ olsun. Bu takdirde $\xi_0^{\,t} (x - x^*) \leq 0$ $\forall x \in S$. $x^* \in \operatorname{int} S$ olduğundan $x = x^* + \lambda \xi_0 \in S$ olacak şekilde $\lambda > 0$ vardır. Böylece $\lambda \xi_0^{\,t} \xi_0 \leq 0$ elde edilir. Bu $\lambda \xi_0^{\,t} \xi_0 = 0 \Rightarrow (\xi_0, \alpha) = (0,0)$ demektir ki $(\xi_0, \alpha) \in \mathbb{R}^n \times \mathbb{R}$ vektörünün sıfırdan farklı olmasıyla çelişir. $\alpha < 0$ dır. ξ vektörü $\xi_0 / |\alpha|$ olarak tanımlanıp $\xi_0^{\,t} (x - x^*) + \alpha(y - f(x^*) \leq 0$ $\forall x, y \in epi(f)$ eşitsizliğinin her iki tarafı $|\alpha|$ ile bölünürse

$\xi^t(x-x^*)+f(x^*)\leq y \quad \forall x,y\in epi(f)$

elde edilir.

$H=\{(x,y): y=f(x^*)+\xi^t(x-x^*)\}$ hiper düzlemi $epi(f)$ nı $(x^*,f(x^*))$ sınır noktasında destekler. $y=f(x)$ olduğu durumda subgradyentin tanım koşulu elde edilir $\xi^t(x-x^*)+f(x^*)\leq f(x) \quad \forall x\in S$.

Sonuç 3.4.2.1. $S\subseteq\mathbb{R}^n, S\neq\varnothing$ konveks bir küme ve $f:S\to\mathbb{R}$ bir kesin konveks fonksiyon olsun. Bu takdirde $x^*\in \operatorname{int} S$ için ξ vektörü vardır. Öyle ki $f(x)>f(x^*)+\xi^t(x-x^*)$ eşitsizliği $\forall x\in S, \quad x\neq x^*$ için sağlanır [3].

Teorem 3.4.3. $S\subseteq\mathbb{R}^n, S\neq\varnothing$ konveks bir küme ve $f:S\to\mathbb{R}$ bir fonksiyon olsun. $\forall x^*\in \operatorname{int} S$ için $f(x)\geq f(x^*)+\xi^t(x-x^*) \quad \forall x\in S.$ eşitsizliğini sağlayan bir ξ vektörü varsa f, S üzerinde konvekstir [3].

İspat: $x_1,x_2\in \operatorname{int} S$ olsun. $\operatorname{int} S$ konveks olduğundan $\lambda x_1+(1-\lambda)x_2\in \operatorname{int} S, \lambda\in(0,1)$. Kabülden dolayı $\lambda x_1+(1-\lambda)x_2\in \operatorname{int} S$ için ξ subgradyent vektörü vardır ve

$\xi^t[x_1-(\lambda x_1+(1-\lambda)x_2)]+f((\lambda x_1+(1-\lambda)x_2))\leq f(x_1)$

$(1-\lambda)\xi^t[x_1-x_2]+f((\lambda x_1+(1-\lambda)x_2))\leq f(x_1)$ ...(1)

$\xi^t[x_2-(\lambda x_1+(1-\lambda)x_2)]+f((\lambda x_1+(1-\lambda)x_2))\leq f(x_2)$

$(\lambda)\xi^t[x_2-x_1]+f((\lambda x_1+(1-\lambda)x_2))\leq f(x_2)$ ...(2)

eşitsizlikleri sağlanır. (1) eşitsizliği λ ile ve (2) eşitsizliği $(1-\lambda)$ ile çarpılıp taraf tarafa toplanırsa bir fonksiyonun konvekslik koşulu elde edilir.

$\lambda f(x_1)+(1-\lambda)f(x_2)\geq f(\lambda x_1+(1-\lambda)x_2)$

f, S üzerinde konvekstir.

3.5. Türevlenebilir Konveks Fonksiyonlar ve Özellikleri

Tanım 3.5.1. $S \subseteq \mathbb{R}^n, S \neq \varnothing$ bir küme ve $f : S \to \mathbb{R}$ bir fonksiyon olsun. Bu takdirde eğer aşağıdaki eşitliği sağlayan gradyent vektörü olarak bilinen $\nabla f(x^*)$ vektörü ve $\beta : \mathbb{R}^n \to \mathbb{R}$ fonksiyonu varsa f fonksiyonu $x^* \in \operatorname{int} S$ noktasında türevlenebilirdir denir.

$$f(x) = f(x^*) + \nabla f(x^*)^t (x - x^*) + \left\| x - x^* \right\| \beta(x^*; x - x^*) \qquad \text{burada} \quad \lim_{x \to x^*} \beta(x^*; x - x^*) = 0$$

dır.

Bir fonksiyonun $S' \subseteq S$ açık kümesinde türevlenebilmesi $\forall x' \in S'$ noktasında türevlenebilmesiyle açıklanır. $f(x^*) + \nabla f(x^*)^t (x - x^*) + \left\| x - x^* \right\| \beta(x^*; x - x^*)$ ifadesi fonksiyonun x^* noktasında veya civarındaki 1. mertebeden *Taylor* seri açılımıdır. $f(x^*) + \nabla f(x^*)^t (x - x^*)$ ifadesi ise fonksiyonun 1. mertebeden x^* civarındaki *Taylor* yaklaşımıdır. f fonksiyonunun $x^* \in \operatorname{int} S$ noktasındaki gradyent vektörü aşağıdaki biçimde verilir.

$$\nabla f(x^*) = \left(\frac{\partial f(x^*)}{\partial x_1}, \ldots, \frac{\partial f(x^*)}{\partial x_n}\right)^t \qquad [3].$$

Tanım 3.5.2. $f : S \subseteq \mathbb{R}^n \to \mathbb{R}$ fonksiyonu göz önüne alınsın. $x, y \in S$ alınsın. m doğrusu, $m = \{(1-t)x + ty \mid t \in \mathbb{R}\}$ biçiminde tanımlansın. Bu takdirde bir fonksiyonun bir doğruya kısıtlanışı $f_m(t) = f((1-t)x + ty)$, $f_m(t) : \mathbb{R} \to \mathbb{R}$ şeklinde tanımlanır [10].

Teorem 3.5.1. $f : S \subseteq \mathbb{R}^n \to \mathbb{R}$ bir fonksiyon olmak üzere f fonksiyonunun konveks olması yalnız ve ancak $f_m(t) : \mathbb{R} \to \mathbb{R}$ fonksiyonunun konveks olması ile gerçeklenir [10].

İspat: $\Rightarrow f$ konveks ise $\forall t_1, t_2 \in \mathbb{R}$ ve $0 \le \alpha \le 1$ için

$$\begin{aligned} f_m(\alpha t_1 + (1-\alpha)t_2) &= f(\alpha[(1-t_1)x + t_1 y] + (1-\alpha)[(1-t_2)x + t_2 y]) \\ &\le \alpha f((1-t_1)x + t_1 y) + (1-\alpha) f((1-t_2)x + t_2 y) \\ &= \alpha f_m(t_1) + (1-\alpha) f_m(t_2) \end{aligned}$$

elde edilir. $\Rightarrow f_m$ konvekstir.

$\Leftarrow f_m$ konveks ise $\forall x_1, x_2 \in \mathbb{R}^n$ ve $0 \le t \le 1$ için

$$\begin{aligned} tf(x_1) + (1-t) f(x_2) &= tf(x_1 + 0(x_2 - x_1)) + (1-t) f(x_1 + 1(x_2 - x_1)) \\ &= tf_m(0) + (1-t) f_m(1) \\ &\ge f_m(1-t) \\ &= f(tx_1 + (1-t)x_2) \end{aligned}$$

elde edilir. $\Rightarrow f$ konvekstir.

Örnek 3.5.1. $f : \mathbb{R}^3 \to \mathbb{R}$ $f(x_1, x_2, x_3) = x_1^{\,2} + x_2^{\,2} + x_3^{\,2}$ biçiminde verilsin.

$\forall (x_1, x_2, x_3), (y_1, y_2, y_3) \in \mathbb{R}^3$, $t \in \mathbb{R}$ olmak üzere,

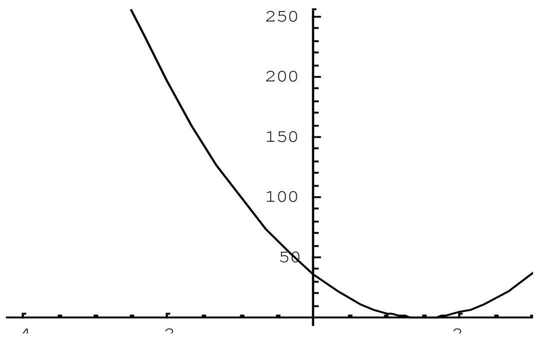
$$\begin{aligned} f_m(t) &= f((1-t)(x_1, x_2, x_3) + t(y_1, y_2, y_3)) \\ &= f((1-t)x_1 + ty_1, (1-t)x_2 + ty_2, (1-t)x_3 + ty_3) \\ &= ((1-t)x_1 + ty_1)^2 + ((1-t)x_2 + ty_2)^2 + ((1-t)x_3 + ty_3)^2 \end{aligned}$$

olarak bulunur.

Şekil 3.12. deki grafik

$(x_1, x_2, x_3) = (2, 2, 2)$ $(y_1, y_2, y_3) = (2, 2, 2)$, $t \in [-4, 4]$

değerleri için çizdirilmiştir.



Şekil 3.12. Örnek 3.5.1. için çizim

Teorem 3.5.2. $C \subseteq \mathbb{R}^n, C \neq \varnothing$ açık konveks bir küme olmak üzere $f : C \to \mathbb{R}$ türevlenebilir olsun ($\nabla f(x)$ $\forall x \in C$ için var olsun) f fonksiyonunun konveksliği ancak ve ancak $f(y) \geq f(x) + \nabla f(x)^t (y-x)$ eşitsizliğinin $\forall x, y \in C$ için sağlanması ile gerçeklenir [7].

İspat: f fonksiyonunun $\forall x, y \in C$ için bir doğruya kısıtlanışı olan $f_m : \mathbb{R} \to \mathbb{R}$ $f_m(t) = f(ty + (1-t)x)$ fonksiyonu ele alınsın. f fonksiyonunun konveksliği ile bu fonksiyonun konveksliği örtüştüğünden ispat bu fonksiyon üzerinden yapılacaktır.

$\Rightarrow f$ konveks olsun. Bu takdirde $f_m(t) = f(ty + (1-t)x)$ fonksiyonu da konvekstir. $f_m'(t) = \nabla f(ty + (1-t)x)^t (y-x)$ şeklinde tanımlanır.

$\forall x, y \in C$ alınsın. C konveks olduğundan $ty + (1-t)x \in C,\ 0 < t \leq 1,\ \forall x, y \in C$ olur. Bu takdirde $f_m(t) = f(ty + (1-t)x) \leq tf(y) + (1-t)f(x)$ burada eşitsizliğin her iki tarafı t ile bölünürse

$$\frac{f(ty + (1-t)x) - (1-t)f(x)}{t} \leq f(y) \quad \text{elde edilir.}$$

$$\lim_{t \to 0} \frac{f(ty + (1-t)x) - (1-t)f(x)}{t} \leq f(y) \quad \Rightarrow f_m(0) + f_m'(0) \leq f_m(1) \quad \text{böylece}$$

$f(y) \geq f(x) + \nabla f(x)^t (y-x)$ eşitsizliği sağlanmış olur.

$\Leftarrow$ Farzedelim ki f, $f(y) \geq f(x) + \nabla f(x)^t (y-x)$ eşitsizliğini $\forall x, y \in C$ için sağlasın. C konveks olduğundan $x_1, x_2 \in C$ için $x^* = \alpha x_1 + (1-\alpha)x_2 \in C$, $\alpha \in (0,1)$ ve $x_1 \neq x_2$ için

$f(x^*) + \nabla f(x^*)^t (x_1 - x^*) \leq f(x_1)$ ve $f(x^*) + \nabla f(x^*)^t (x_2 - x^*) \leq f(x_2)$ eşitsizlikleri sağlanır. Sırasıyla bu eşitsizlikler α ve $(1-\alpha)$ ile çarpılıp toplanırsa

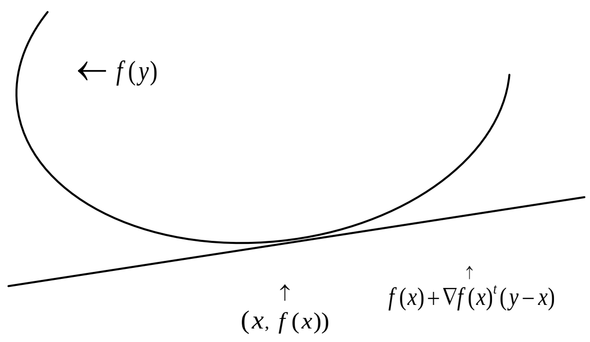
$$\nabla f(x^*)^t (\alpha x_1 + (1-\alpha)x_2 - x^*) + f(x^*) \leq \alpha f(x_1) + (1-\alpha) f(x_2)$$

elde edilir.

$x^* = \alpha x_1 + (1-\alpha)x_2$ olduğundan $\nabla f(x^*)^t .0 + f(x^*) \leq \alpha f(x_1) + (1-\alpha) f(x_2)$ ve sonuç olarak

son eşitsizlikte $x^* = \alpha x_1 + (1-\alpha)x_2$ değişimi yapılırsa bir fonksiyonun konvekslik koşulu elde edilir. $f(\alpha x_1 + (1-\alpha)x_2) \leq \alpha f(x_1) + (1-\alpha) f(x_2)$ $\Rightarrow f$ konvekstir. Demektir ki f_m konvekstir.

Sonuç 3.5.2.1. f fonksiyonunun kesin konveksliği ise ancak ve ancak $f(y) > f(x) + \nabla f(x)^t (y-x)$ eşitsizliğinin $\forall x, y \in C, x \neq y$ için sağlanması ile gerçeklenir [3].



Şekil 3.13. Teorem 3.5.2. için çizim

Geometrik olarak Teorem 3.5.2. $f(x)+\nabla f(x)^t(y-x)$ teğet düzleminin $\nabla f(x)$ normal vektörü ile beraber f fonksiyonunun epigrafını $(x, f(x))$ noktasında desteklediğini ifade etmektedir. Başka bir ifade ile $f(x)+\nabla f(x)^t(y-x)$ fonksiyonu f fonksiyonunun 1. mertebeden Taylor yaklaşımıdır. Teoremin söylediği bu yaklaşımın f fonksiyonunu alttan sınırladığıdır. Optimizasyon açısından bakarsak , $x \in C$ kısıtlarına bağlı olarak bir fonksiyonun minimum değeri istendiğinde bu yaklaşımın C üzerindeki minimum değeri, problemin optimum değeri için alt sınır teşkil etmektedir. Örnek olarak eğer $\nabla f(x)=0$ ise $\forall y \in C$ için $f(y) \geq f(x)$, x noktası mutlak optimum çözüm olmaktadır. Öte yandan bir fonksiyonun bir noktadaki türevi ve bir noktadaki değeri gibi yerel bilgilerden bu fonksiyon karakterize edilmiş olur [3]-[7].

Lemma 3.5.1. C boş olmayan konveks bir küme ve $f: C \subseteq \mathbb{R}^n \to \mathbb{R}$ fonksiyonu $x^* \in \operatorname{int} C$ noktasında türevlenebilir olsun. Bu takdirde f nin $x^* \in \operatorname{int} C$ noktasındaki subgradyentleri kümesi $\{\nabla f(x^*)\}$ dır [3].

İspat: Teorem 3.4.2. den dolayı f fonksiyonunun $x^* \in \operatorname{int} C$ noktasındaki subgradyentleri kümesi boş değildir. ξ, f fonksiyonunun $x^* \in \operatorname{int} C$ noktasındaki subgradyenti olsun. Fonksiyon türevlenebilir olduğundan ve Teorem 3.5.2. den dolayı $\forall x \in C$ ve yeterince küçük λ için

$$f(x^*+\lambda x) \geq f(x^*)+\lambda \xi^t x \text{..(1)}$$

$$f(x^*+\lambda x) = f(x^*)+\lambda \nabla f(x^*)^t x+\|x\| \lambda \beta(x^*;\lambda x) \text{................(2)}$$

ifadeleri sağlanır. (1) eşitsizliğinden (2) eşitliği çıkarılırsa

$$\lambda[\xi-\nabla f(x^*)]^t x+\|x\|\lambda\beta(x^*;\lambda x) \leq 0$$

bu eşitsizlik λ ile bölünerek ve $\lambda \to 0^+$ durumunda $[\xi-\nabla f(x^*)]^t x \leq 0$ olarak bulunur. $x=\xi-\nabla f(x^*)$ olarak seçilebileceğinden $[\xi-\nabla f(x^*)]^t[\xi-\nabla f(x^*)] \leq 0$ elde edilir. Bu takdirde $\xi=\nabla f(x^*)$ olur.

Tanım 3.5.3. $S \subseteq \mathbb{R}^n, S \neq \varnothing$ bir küme ve $f : S \to \mathbb{R}$ bir fonksiyon olsun. Bu takdirde eğer aşağıdaki eşitliği sağlayan gradyent vektörü olarak bilinen $\nabla f(x^*)$ vektörü, *Hessian* matrisi olarak bilinen $n \times n$ simetrik matris H ve $\beta : \mathbb{R}^n \to \mathbb{R}$ fonksiyonu varsa f fonksiyonu $x^* \in \operatorname{int} S$ noktasında 2. mertebeden türevlenebilirdir denir.

$$f(x) = f(x^*) + \nabla f(x^*)^t (x - x^*) + 1/2(x - x^*)^t \mathrm{H}(x^*)(x - x^*) + \left\| x - x^* \right\|^2 \beta(x^*; x - x^*)$$

Burada $\lim_{x \to x^*} \beta(x^*; x - x^*) = 0$ dır.

Bir fonksiyonun $S' \subseteq S$ açık kümesinde 2. mertebeden türevlenebilmesi, $\forall x' \in S'$ için 2. mertebeden türevlenebilmesiyle açıklanır.

$f(x^*) + \nabla f(x^*)^t (x - x^*) + 1/2(x - x^*)^t \mathrm{H}(x^*)(x - x^*) + \left\| x - x^* \right\|^2 \beta(x^*; x - x^*)$ ifadesi fonksiyonun x^* noktasında veya civarındaki 2. mertebeden *Taylor* seri açılımıdır. $f(x^*) + \nabla f(x^*)^t (x - x^*) + 1/2(x - x^*)^t \mathrm{H}(x^*)(x - x^*)$ ifadesi ise fonksiyonun 2. mertebeden x^* noktasında veya civarındaki *Taylor* yaklaşımıdır. Fonksiyonun 2. mertebeden kısmi türevlerini ihtiva eden *Hessian* matrisi,

$$\mathrm{H} = \begin{pmatrix} \dfrac{\partial^2 f}{\partial x_1 \partial x_1} & \cdots & \dfrac{\partial^2 f}{\partial x_1 \partial x_n} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial^2 f}{\partial x_n \partial x_1} & \cdots & \dfrac{\partial^2 f}{\partial x_n \partial x_n} \end{pmatrix} \quad \text{veya} \quad \mathrm{H} = \left[\dfrac{\partial^2 f}{\partial x_i \partial x_j} \right]_{n \times n} \quad \text{şeklinde gösterilir.}$$

verilen x^* noktasında H nın değeri $\mathrm{H}(x^*)$ olur [3]-[11].

$x^t \mathrm{H}(x^*) x$ ifadesi $x = 0$ dışında $\forall x \in \mathbb{R}^n$ ve $\forall x^* \in S$ için $x^t \mathrm{H}(x^*) x > 0$ ise H matrisi pozitif tanımlıdır denir . $x^t \mathrm{H}(x^*) x$ ifadesi $\forall x \in \mathbb{R}^n$ ve $\forall x^* \in S$ için $x^t \mathrm{H}(x^*) x \geq 0$ ise H matrisi pozitif yarı tanımlıdır denir. $x^t \mathrm{H}(x^*) x$ ifadesi $x = 0$ dışında $\forall x \in \mathbb{R}^n$ ve $\forall x^* \in S$ için $x^t \mathrm{H}(x^*) x < 0$ ise H matrisi negatif tanımlıdır denir. $x^t \mathrm{H}(x^*) x$ ifadesi $\forall x \in \mathbb{R}^n$ ve $\forall x^* \in S$ için $x^t \mathrm{H}(x^*) x \leq 0$ ise H matrisi negatif yarı tanımlıdır denir. Matris ne pozitif yarı tanımlı ne de negatif yarı tanımlı ise tanımsızdır denir [3].

Teorem 3.5.3. $f: C \subseteq \mathbb{R}^n \to \mathbb{R}$ fonksiyonu $C \subseteq \mathbb{R}^n$ açık kümesinde 2. mertebeden sürekli türevlenebiliyorsa bu takdirde f fonksiyonunun konveks olması için gerek ve yeter koşul *Hessian* matrisinin $\forall x \in C$ için pozitif yarı tanımlı olmasıdır [5]-[3]-[10].

İspat: $\Rightarrow f$ konveks ve $x^* \in C$ olsun. $\forall x \in \mathbb{R}^n$ için $x^t \mathrm{H}(x^*) x \geq 0$ olduğu gösterilecektir. C açık olduğundan, $\forall x \in \mathbb{R}^n$ için $x^* + \lambda x \in C$ durumunu yeterince küçük $|\lambda| \neq 0$ için sağlar. Teorem 3.5.2. ve f nin 2. merteben türevlenebilir olduğu kullanılırsa

$f(x^* + \lambda x) \geq f(x^*) + \lambda \nabla f(x^*)^t x$...(1)

$f(x^* + \lambda x) = f(x^*) + \lambda \nabla f(x^*)^t x + 1/2 \lambda^2 x^t \mathrm{H}(x^*) x + \|x\|^2 \lambda^2 \beta(x^*; \lambda x)$.....(2)

ifadeleri sağlanır. (1) eşitsizliğinden (2) eşitliği çıkarılırsa

$$1/2 \lambda^2 x^t \mathrm{H}(x^*) x + \|x\|^2 \lambda^2 \beta(x^*; \lambda x) \geq 0$$

elde edilir. Bu eşitsizlik λ^2 ile bölünerek $1/2 x^t \mathrm{H}(x^*) x + \|x\|^2 \beta(x^*; \lambda x) \geq 0$ elde edilir. $\lambda \to 0$ durumunda $x^t \mathrm{H}(x^*) x \geq 0$ elde edilir.

$\Leftarrow$ $f_m : \mathbb{R} \to \mathbb{R}$ $f_m(t) = f(ty + (1-t)x)$ $\forall x, y \in C$ fonksiyonu ele alınsın. f fonksiyonunun konveksliği ile bu fonksiyonun konveksliği örtüştüğünden ispat bu fonksiyon üzerinden yapılacaktır.

$$f_m(t) = f(ty + (1-t)x) \quad \forall x, y \in C, \ t \in \mathbb{R}$$

olmak üzere kabul edelim ki *Hessian* pozitif yarı tanımlı olsun. Bu durumda $f_m''(t) \geq 0 \ \ \forall x, y \in C, t \in \mathbb{R}$

$\forall x, y \in C, \alpha \in (0,1)$ için $t_1 < t_2$ olsun. Ortalama değer teoremi gereği $\exists \zeta_1, \zeta_2$, $t_1 < \zeta_1 < \alpha t_1 + (1-\alpha) t_2 < \zeta_2 < t_2$ ve $\exists \zeta_1, \zeta_2$, $\exists \zeta_3$, $\zeta_1 < \zeta_3 < \zeta_2$ öyleki

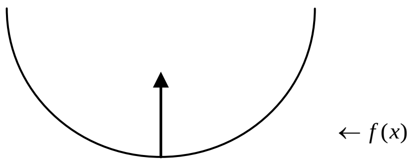
$$f_m(\alpha t_1 + (1-\alpha)t_2) - \alpha f_m(t_1) - (1-\alpha) f_m(t_2) = \alpha[f_m(\alpha t_1 + (1-\alpha)t_2) - f_m(t_1)] +$$
$$(1-\alpha)[f_m(\alpha t_1 + (1-\alpha)t_2) - f_m(t_2)]$$
$$= \alpha(1-\alpha)(t_2 - t_1) f_m'(\zeta_1) + \alpha(1-\alpha)(t_1 - t_2) f_m'(\zeta_2)$$
$$= \alpha(1-\alpha)(t_2 - t_1)(\zeta_1 - \zeta_2) f_m''(\zeta_3)$$

Hessian pozitif yarı tanımlı olduğundan $f_m''(\zeta_3) \geq 0$ dır. Bu durumda son eşitlik $\alpha(1-\alpha)(t_2 - t_1)(\zeta_1 - \zeta_2) f_m''(\zeta_3) \leq 0$ olarak bulunur.

Bu takdirde $f_m(t) = f(ty + (1-t)x)$ konvekstir. Sonuç olarak f fonksiyonu konvekstir.

Sonuç 3.5.3.1. $f: C \subseteq \mathbb{R}^n \to \mathbb{R}$ fonksiyonu $C \subseteq \mathbb{R}^n$ açık kümesinde 2. mertebeden sürekli türevlenebiliyorsa bu takdirde f fonksiyonunun kesin konveks olması için gerek ve yeter koşul *Hessian* matrisinin $\forall x \in C$ için pozitif tanımlı olmasıdır.

Geometrik olarak fonksiyonun 2. türevine karşılık gelen *Hessian* matrisin pozitif yarı tanımlı olması fonksiyonun grafiğinin $x \in C$ noktasında pozitif yönde eğriliğini göstermektedir [7].



Şekil 3.14. Teorem 3.5.3. için çizim

3.6. Kuadratik Formlar ve *Hessian* Matris

n değişkenli bir $f(x)$ fonksiyonu,

$$f(x) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij} x_i x_j = a_{11}x_1x_1 + a_{12}x_1x_2 + ... + a_{1n}x_1x_n$$
$$+ .... + a_{n1}x_nx_1 + a_{n2}x_nx_2 + ... + a_{nn}x_nx_n$$

şeklinde ise $f(x)$ fonksiyonuna bir kuadratik (kareli) fonksiyon veya kuadratik (kareli) formdadır denir. $f(x)$ fonksiyonu bir kuadratik form ise, bunun her bir teriminde iki değişkenin çarpımı veya bir değişkenin karesi vardır.

$f(x)$ fonksiyonu bir kuadratik fonksiyon, A $n\times n$ bir matris ve x $n\times 1$ bir sütun vektörü olmak üzere $f(x)$ fonksiyonu,

$$f(x)=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_i x_j=\sum_{i=1}^{n}x_i\sum_{j=1}^{n}a_{ij}x_j=x^t Ax$$

biçiminde yazılabilmektedir.

Görüleceği gibi a_{ij} ve a_{ji} $i\neq j$ iken $x_i x_j$ ve $x_j x_i$ terimlerinin katsayılarıdır. $x_i x_j=x_j x_i$ olduğundan $x_i x_j$ terimine karşılık tek bir katsayı $b_{ii}=a_{ii}$ ve $b_{ij}=\dfrac{a_{ij}+a_{ji}}{2}$ biçiminde bulunabilir. Bu durumda açılımdaki katsayılar matrisi *B* ile gösterilirse bu matris simetrik bir matris olur.

Örnek olarak; $f(x_1,x_2)=3x_1^2+6x_1x_2+4x_2^2$ kuadratik fonksiyonu $b_{11}=3,\ b_{22}=4$ ve $b_{12}=\dfrac{6+0}{2}=3,\ b_{21}=\dfrac{0+6}{2}=3$ şeklindeki hesaplamayla

$$f(x_1,x_2)=(x_1,x_2)\begin{pmatrix}3 & 3\\ 3 & 4\end{pmatrix}(x_1,x_2)^t$$ olarak bulunur.

Bu gösterim tek değildir başka gösterimler de bulunabilir.

$$f(x_1,x_2)=(x_1,x_2)\begin{pmatrix}3 & 1\\ 5 & 4\end{pmatrix}(x_1,x_2)^t$$ ise başka bir gösterimdir.

$x^t Ax$ ifadesi $x=0$ dışında her x için $x^t Ax>0$ ise kuadratik forma pozitif tanımlıdır denir. $x^t Ax$ ifadesi $x=0$ dışında en az bir x için $x^t Ax=0$ ve her x için $x^t Ax\geq 0$ ise kuadratik forma pozitif yarı tanımlıdır denir. $x^t Ax$ ifadesi $x=0$ dışında her x için $x^t Ax<0$ ise kuadratik forma negatif tanımlıdır denir. $x^t Ax$ ifadesi $x=0$ dışında en az bir x için $x^t Ax=0$ ve her x için $x^t Ax\leq 0$ ise kuadratik forma negatif yarı tanımlıdır denir.

$x^t Ax$ kuadratik formunda A'nın asal minörleri pozitif ise $x^t Ax$ pozitif tanımlı, $x^t Ax$ pozitif tanımlı ise A'nın asal minörleri pozitiftir. Bununla birlikte kuadratik form pozitif tanımlı ise A'nın tüm özdeğerleri pozitif, A'nın tüm özdeğerleri pozitif ise kuadratik form pozitif tanımlıdır sonucuna varılır. $x^t Ax$ kuadratik formunda A'nın asal minörleri -,+,-,.. şeklinde ise $x^t Ax$ negatif tanımlı, $x^t Ax$ negatif tanımlı ise A'nın asal minörleri -,+,-,.. şeklindedir. Bununla birlikte kuadratik form negatif tanımlı ise A'nın tüm özdeğerleri negatif, A'nın tüm özdeğerleri negatif ise kuadratik form negatif tanımlıdır sonucu elde edilir. Yarı tanımlılık durumu ise bu değerlerden en az birinin sıfır olması ile sağlanır [11].

Pozitif yarı tanımlı bir kuadratik fonksiyon ile ilgili aşağıdaki özellik verilebilir.

Pozitif yarı tanımlı bir kuadratik fonksiyon konvekstir [11].

İspat: $f(x) = x^t Ax$ olmak üzere, $x_1, x_2,\ x_1 \neq x_2$ $f(x)$ fonksiyonunun tanım kümesinden alınsın. $x = \lambda x_1 + (1-\lambda)x_2,\ \ \lambda \in [0,1]$ biçiminde yazılabilsin. Bu takdirde

$$f(x) = x^t Ax = [\lambda x_1 + (1-\lambda)x_2]^t A[\lambda x_1 + (1-\lambda)x_2]$$
$$f(x) = x_2{}^t Ax_2 + 2\lambda(x_1 - x_2)^t Ax_2 + \lambda^2 (x_1 - x_2)^t A(x_1 - x_2)$$

olarak bulunur. $\lambda^2 \leq \lambda$ olduğundan ve $f(x)$ pozitif yarı tanımlı olduğundan

$\lambda^2 (x_1 - x_2)^t A(x_1 - x_2) \leq \lambda(x_1 - x_2)^t A(x_1 - x_2)$ olur. Bu takdirde

$$\begin{aligned} f(x) &= x^t Ax \\ &\leq x_2{}^t Ax_2 + 2\lambda(x_1 - x_2)^t Ax_2 + \lambda(x_1 - x_2)^t A(x_1 - x_2) \end{aligned}$$

elde edilir. Bu ifadede $x = \lambda x_1 + (1-\lambda)x_2,\ \ \lambda \in [0,1]$ yerine konulduğu takdirde bir fonksiyonunun konvekslik koşulu elde edilir

$$f(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f(x_1) + (1-\lambda) f(x_2).$$

Pozitif tanımlı bir kuadratik fonksiyonun konveks oluşu da aynı biçimde gösterilebilir.

$f : S \to \mathbb{R},\ S \subseteq \mathbb{R}^n$ bir fonksiyon olmak üzere, fonksiyonun *Hessian* matrisinin pozitif yarı tanımlı olup olmadığı yönünde incelenirken $\forall x^* \in S$ ve $\forall x \in \mathbb{R}^n$ için $x^t \mathrm{H}(x^*) x$ ifadesinin incelenmesinden ve bu ifadenin bir kuadratik form olmasından dolayı *Hessian* matrisi ile kuadratik formların bir ilişkisi vardır. Ayrıca kuadratik fonksiyonların konveksliğine de fonksiyonun tanımlılık durumu incelenerek karar verilebilir.

Örnek 3.6.1. $f : \mathbb{R}^3 \to \mathbb{R}$ $\quad f(x_1, x_2, x_3) = x_1^2 - 3x_1 - 2x_3 + x_2^2 + x_3^2 - x_1 x_2$

fonksiyonu ele alınsın. Bu fonksiyon

$f_1(x_1, x_2, x_3) = -3x_1 - 2x_3$ ve $f_2(x_1, x_2, x_3) = x_1^2 + x_2^2 + x_3^2 - x_1 x_2$ olmak üzere

$f(x_1, x_2, x_3) = f_1(x_1, x_2, x_3) + f_2(x_1, x_2, x_3)$ olacak şekilde yazılabilir.

$$f_2(x_1, x_2, x_3) = (x_1, x_2, x_3)\begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}(x_1, x_2, x_3)^t$$

biçiminde yazılabilir. $\Delta_i,\ i = 1, 2, 3.$ asal minörleri ifade etmek üzere $\Delta_1 = 1,\ \Delta_2 = 1$ ve $\Delta_3 = 1$ olduğundan kuadratik fonksiyonun pozitif tanımlı olduğu sonucuna varılır. Bu takdirde $f_2(x_1, x_2, x_3)$ fonksiyonu konvekstir.

$f_1(x_1, x_2, x_3) = -3x_1 - 2x_3$ fonksiyonu da doğrusal olduğundan hem konveks hem de konkavdır. Konveks fonksiyonların toplamı da konveks olduğundan $f(x_1, x_2, x_3)$ fonksiyonu konvekstir [11].

BÖLÜM 4. KONVEKSLİK VE OPTİMİZASYON

4.1. Giriş

Bu bölümde konveks fonksiyonların konveks bir küme üzerinde minimumlaştırılmalarına ve maksimumlaştırılmalarına değinilmiştir. Optimizasyon problemlerinin tanımlanmasında ve izlenecek çözüm yönteminin belirlenmesinde fonksiyonun niteliği ve problemin tanım kümesinin belirlenmesi son derece önemlidir.

4.2. Konveks Fonksiyonları Minimumlaştırma ve Maksimumlaştırma

Tanım 4.2.1. x bir vektör olmak üzere f fonksiyonunun tanım kümesi S ve $x_0 \in S$ için x_0 ın δ komşuluğu A olsun,

1. A daki tüm x ler için $f(x_0) \leq f(x)$ ise $f(x)$, $x = x_0$ da yerel minimum,

2. S deki tüm x ler için $f(x_0) \leq f(x)$ ise $f(x)$, $x = x_0$ da mutlak minimum

değer alıyor demektir [11].

Tanım 4.2.2. x bir vektör olmak üzere f fonksiyonunun tanım kümesi S ve $x_0 \in S$ için x_0 ın δ komşuluğu A olsun,

1. A daki tüm x ler için $f(x_0) \geq f(x)$ ise $f(x)$, $x = x_0$ da yerel maximum,

2. S deki tüm x ler için $f(x_0) \geq f(x)$ ise $f(x)$, $x = x_0$ da mutlak maksimum değer alıyor demektir [11].

Bir fonksiyonun minimum değere ulaştığı küme ve fonksiyonun bu küme üzerinde konvekslik karakterizasyonu ile ilişkili olarak aşağıdaki özellik verilebilir.

$f: S \to \mathbb{R}$ bir fonksiyon ve x_0 noktasının δ komşuluğu A olsun. Bu durumda $f(x)$, x_0'da yerel minimum değerini alıyorsa, $f(x)$, A kümesinde konvekstir [11].

İspat: x_0, $f(x)$ in yerel minimumu olsun. $x_1, x_2 \in S$ ve $\lambda \in [0,1]$ olmak üzere $x_0 = \lambda x_1 + (1-\lambda)x_2$ şeklinde yazılsın. $f(x)$, x_0'da yerel minimum değerini aldığına göre

$$f(x_1) \geq f(x_0) \Rightarrow \lambda f(x_1) \geq \lambda f(x_0)$$
$$f(x_2) \geq f(x_0) \Rightarrow (1-\lambda) f(x_2) \geq (1-\lambda) f(x_0)$$

elde edilir. Bu eşitsizlikleri taraf tarafa toplarsak

$\lambda f(x_1) + (1-\lambda) f(x_2) \geq f(x_0)$ bulunur ve $f(x_0) = f(\lambda x_1 + (1-\lambda)x_2)$

olduğundan aşağıdaki son eşitsizlik elde edilir;

$\lambda f(x_1) + (1-\lambda) f(x_2) \geq f(\lambda x_1 + (1-\lambda)x_2) \Rightarrow f$ konvekstir.

Konveks bir fonksiyonun yerel minimum değeri ve mutlak minimum değeri arasındaki ilişki için aşağıdaki özellik verilebilir.

Eğer $f: S \to \mathbb{R}$ bir konveks kümede tanımlı konveks bir fonksiyon ve x^*, $f(x)$ in yerel minimumu ise $f(x)$, x^* da mutlak minimum değerini alır [11].

İspat: x^*'ın bir δ civarı A olsun ve $f(x)$, x^* noktasında yerel minimum değerini alsın bu takdirde $f(x^*) \leq f(x)$ olur. Eğer x^* mutlak minimum değil ise bu durumda $\exists x \notin A$ $f(x) < f(x^*)$ eşitsizliği sağlanır. $f(x)$'in tanım kümesi bir konveks küme olduğuna göre x^* ve x noktalarının her konveks kombinasyonu bu küme içinde kalır. x^*'ın civarında bir nokta $y = \lambda x + (1-\lambda)x^*$ olsun. Fonksiyon konveks olduğundan

$\lambda f(x) + (1-\lambda) f(x^*) \geq f(\lambda x + (1-\lambda)x^*)$ yazılır ve $f(x) < f(x^*)$ eşitsizliğinden dolayı $f(x)$ ve $f(x^*)$ eşitsizliğin sol tarafında yer değiştirirse

$$\lambda f(x^*) + (1-\lambda) f(x^*) \geq f(\lambda x + (1-\lambda) x^*)$$
$$\Rightarrow \quad f(x^*) > f(y)$$

Bu takdirde x^* yerel minimum olamaz. Bu da baştaki kabulümüzle çelişir. Sonuç olarak fonksiyon, tanım kümesindeki her x için x^* noktasında mutlak minimum değerini alır.

z = $f(x_1,...,x_n)$ fonksiyonunun $x^* = (x_1^*, x_2^*, ..., x_n^*)$ noktasında yerel minimumunun veya yerel maksimumunun olması için gerek koşul $\nabla f(x^*) = 0$ olmasıdır. Şayet fonksiyonun *Hessian* matrisi bu noktada pozitif tanımlı ise yerel minimum ve *Hessian* matrisi negatif tanımlı ise yerel maksimum vardır. Bu koşullar ise yeter koşullardır [12].

Örnek 4.2.1. $f : \mathbb{R}^3 \to \mathbb{R},\ f(x_1, x_2, x_3) = x_1^2 - 3x_1 - 2x_3 + x_2^2 + x_3^2 - x_1 x_2$

fonksiyonu ele alınsın. Gradyent vektörünü sıfır yapan değerler araştırılırsa,

$$\begin{aligned} 2x_1 - 3 - x_2 &= 0 \\ 2x_2 - x_1 &= 0 \\ 2x_3 - 2 &= 0 \end{aligned}$$

Bu sistemin tek çözümü (2,1,1) dir. Bu nokta *Hessian* matrisinde yerine konursa

$$\mathrm{H}(2,1,1) = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

elde edilir. $\Delta_i,\ i = 1,2,3.$ asal minörleri ifade etmek üzere $\Delta_1 = 2,\ \Delta_2 = 3$ ve $\Delta_3 = 6$ olduğundan *Hessian* matrisi bu noktada pozitif tanımlıdır. Bu takdirde (2,1,1) bir yerel minimumdur. Fonksiyon ayrıca konveks olduğundan (2,1,1) noktası bu fonksiyon için mutlak minimumdur [11].

Tanım 4.2.3.

$f:\mathbb{R}^n \to \mathbb{R}$ bir fonksiyon olmak üzere $f(x)$ fonksiyonunun $x \in S$ kısıtlarına bağlı olarak minimumlaştırma problemi ele alınsın. Burada $x \in S$ noktalarına problemin olası çözümleri denir. Eğer $x^* \in S$ ve $\forall x \in S$ için $f(x) \geq f(x^*)$ ise $x^* \in S$ noktasına optimum çözüm denir. Optimum çözümler kümesine alternatif optimum çözümler denir. Eğer $x^* \in S$ ve x^* ın δ komşuluğu A varsa ve $S \cap A$ daki tüm x ler için $f(x^*) \leq f(x)$ ise x^* yerel optimum çözümdür denir. Eğer $x^* \in S$ ve x^* ın δ komşuluğu A varsa öyle ki $S \cap A$ daki tüm x ler için $f(x^*) < f(x)$ ise $x^* \neq x$ için kesin yerel optimum çözümdür denir. Eğer $x^* \in S$ ve x^* ın δ komşuluğu A varsa ve $S \cap A$ daki tüm x ler için x^* tek yerel optimum çözüm ise kuvvetli veya izole optimum yerel çözümdür denir.

$f(x)$ fonksiyonunun $x \in S$ kısıtlarına bağlı olarak minimumlaştırma problemi S konveks bir küme ve $f(x)$ fonksiyonu konveks bir fonksiyon olduğu durumda Konveks Programlama Problemi adını alır. [3]

Teorem 4.2.1. S, $\mathbb{R}^n$ uzayında boştan farklı konveks bir küme olsun ve f, S üzerinde konveks olsun. $f(x)$ fonksiyonunun $x \in S$ kısıtlarına bağlı olarak minimumlaştırma problemi ele alınsın. Farzedelim ki $x^* \in S$ yerel optimum çözüm (yerel minimum) olsun. Bu takdirde ;

1. $x^* \in S$ mutlak optimum çözümdür.

2. Eğer $x^* \in S$ kesin yerel optimum çözüm veya $f(x)$ kesin konveks ise bu takdirde x^* tek mutlak optimum çözüm ve kuvvetli yerel optimum çözümdür [3].

İspat:

1. $x^* \in S$ yerel optimum çözüm olduğuna göre x^* ın bir δ civarı A vardır öyle ki $f(x) \geq f(x^*)$ $\forall x \in S \cap A$ için sağlanır. Farzedelim ki x^* mutlak optimum çözüm olmasın. $f(x') < f(x^*)$ eşitsizliğini gerçekleyecek $x' \in S$ bulunsun. f nin konveksliğinden $\forall\ 0 \leq \lambda \leq 1$ için

$$f(\lambda x'+(1-\lambda)x^*) \le \lambda f(x') + (1-\lambda) f(x^*) < \lambda f(x^*) + (1-\lambda) f(x^*) = f(x^*)$$

yazılır. Ancak yeterince küçük $\lambda > 0$ için $\lambda x'+(1-\lambda)x^* \in S \bigcap A$ olacağından dolayı

$f(\lambda x'+(1-\lambda)x^*) < f(x^*)$ eşitsizliği, $f(x) \ge f(x^*)$ $\forall x \in S \bigcap A$ eşitsizliği ile çelişir. Bu takdirde x^* mutlak optimum çözümdür.

2.1. $x^* \in S$ kesin yerel optimum çözüm olsun. (1) den dolayı aynı zamanda mutlak optimum çözümdür. Eğer $f(x') = f(x^*)$ eşitliğini sağlayacak $x' \in S$ varsa bu takdirde $x_\lambda = \lambda x'+(1-\lambda)x^*$ $\forall \lambda \in [0,1]$, olarak tanımlansın. f nin ve S kümesinin konveksliğinden

$f(x_\lambda) \le \lambda f(x') + (1-\lambda) f(x^*) = f(x^*)$, $x_\lambda \in S$ $\forall \lambda \in [0,1]$ yazılır.

$\lambda \to 0^+$ durumunda $x_\lambda \in \mathrm{S} \bigcap \mathrm{A}$ $\forall \delta > 0$, olacağından bu durum $x^* \in S$ kesin yerel optimum çözüm olması durumuyla çelişir. $x^* \in S$ tek mutlak optimum çözümdür. $x^* \in S$ aynı zamanda kuvvetli yerel optimum çözüm olmalıdır. Çünkü aksi takdirde $\mathrm{S} \bigcap \mathrm{A}$ daki diğer yerel optimum çözümler mutlak optimum çözüm olur ki bu bir çelişkidir.

2.2. $x^* \in S$ yerel minimum nokta olsun ve f kesin konveks olsun. Kesin konveks ise konveks olduğundan (2.1) den dolayı $x^* \in S$ mutlak optimum çözümdür. Kabul edelim ki $x^* \in S$ tek mutlak optimum çözüm olmasın. $x \ne x^*$ için $f(x) = f(x^*)$ olacak şekilde $x \in S$ olsun.

f kesin konveks olduğundan $f(\frac{1}{2}x + \frac{1}{2}x^*) < \frac{1}{2} f(x) + \frac{1}{2} f(x^*) = f(x^*)$ yazılır. S konveks olduğundan $\frac{1}{2}x + \frac{1}{2}x^* \in S$ olur ve $f(\frac{1}{2}x + \frac{1}{2}x^*) < f(x^*)$ eşitsizliği $x^* \in S$ nin mutlak optimum çözüm olması ile çelişir. $x^* \in S$ tek mutlak optimum çözümdür. $x^* \in S$ aynı zamanda kuvvetli yerel optimum çözüm olmalıdır. Çünkü aksi takdirde $\mathrm{S} \bigcap \mathrm{A}$ daki diğer yerel çözümler mutlak optimum çözüm olur ki bu bir çelişkidir.

Konveks Programlama Problemi'nin yerel optimum çözümü varsa bu çözümün mutlak optimum çözüm olduğu söylenmiştir. Bu aşamada asıl sorumuz böyle bir çözümün hangi şartlar altında var olduğudur. Optimum çözümün olduğu noktada fonksiyon hangi şartları sağlamalıdır. Bu koşulları oluşturmak için aşağıdaki teorem verilsin.

Teorem 4.2.2. $f:\mathbb{R}^n \to \mathbb{R}$ konveks bir fonksiyon ve S, $\mathbb{R}^n$ uzayında boştan farklı konveks bir küme olsun. $f(x)$ fonksiyonunun $x \in S$ kısıtlarına bağlı olarak minimumlaştırma problemi ele alınsın. Bu takdirde $x^* \in S$ noktasının optimum çözüm olması için gerek ve yeter koşul $f(x)$ fonksiyonun $\xi^t(x-x^*) \geq 0 \ \ \forall x \in S$ koşulunu sağlayan bir subgradyente sahip olmasıdır [3].

İspat:

$\Leftarrow$ Farzedelim ki $\xi^t(x-x^*) \geq 0 \ \ \forall x \in S$ burada ξ, f nin x^* noktasındaki subgradyentidir. f nin konveksliği gereği $f(x) \geq f(x^*) + \xi^t(x-x^*) \geq f(x^*) \ \ \forall x \in S$ x^* optimum çözümdür.

$\Rightarrow$ Farzedelim ki x^* optimum çözüm olsun. Ω_1, Ω_2 kümeleri $\mathbb{R}^{n+1}$ de aşağıdaki biçimde tanımlansın.

$$\Omega_1 = \{(x-x^*, y) : x \in \mathbb{R}^n, y > f(x) - f(x^*)\}$$

$$\Omega_2 = \{(x-x^*, y) : x \in S, y \leq 0\}$$

Ω_1 ve Ω_2 kümelerinin konveks kümeler olduklarını gösterelim.

$(x_1 - x_1^*, y_1), (x_2 - x_2^*, y_2) \in \Omega_1$, $\lambda \in [0,1]$ olsun.

$$\lambda(x_1 - x_1^*, y_1) + (1-\lambda)(x_2 - x_2^*, y_2) = (\lambda(x_1 - x_1^*) + (1-\lambda)(x_2 - x_2^*), \lambda y_1 + (1-\lambda) y_2)$$

$$= (\lambda x_1 + (1-\lambda)x_2 - (\lambda x_1^* + (1-\lambda)x_2^*), \lambda y_1 + (1-\lambda) y_2)$$

$\lambda y_1 > \lambda f(x_1) - \lambda f(x_1^*)$

$(1\text{-}\lambda) y_2 > (1-\lambda) f(x_2) + (1-\lambda) f(x_2^*)$

$(1\text{-}\lambda) y_2 + \ \lambda y_1 > \lambda f(x_1) + (1-\lambda) f(x_2^*) - [\lambda f(x_1^*) + (1-\lambda) f(x_2^*)]$

$\geq f(\lambda x_1 + (1-\lambda) x_2^*) - f(\lambda x_1^* + (1-\lambda) x_2^*)$

olduğundan Ω_1 konvekstir.

$(x_1 - x_1^*, y_1), (x_2 - x_2^*, y_2) \in \Omega_2, \ \ \lambda \in [0,1]$ olsun.

$\lambda(x_1 - x_1^*, y_1) + (1-\lambda)(x_2 - x_2^*, y_2), x_1, x_2 \in S$

$\lambda y_1 \leq 0, (1\text{-}\lambda) y_2 \leq 0 \quad \Rightarrow \lambda y_1 + (1-\lambda) y_2 \leq 0$

olduğundan Ω_2 konvekstir.

Ω_1 ve Ω_2 kümelerinin kesişimi boştur. Aksi takdirde bu kesişimde öyle bir $(x - x^*, y)$ olur ve $x \in S$ için $0 \geq y > f(x) - f(x^*)$, $f(x^*) > f(x)$ sağlanır ki x^* optimum çözüm olmaz.

Ω_1 ve Ω_2 kümelerinin kesişimi boş ve konveks olduklarından bu kümeleri ayıran bir hiper düzlem mevcuttur. $(\xi_0, \gamma) \in \mathbb{R}^n \times \mathbb{R}$ ve α skaleri vardır öyle ki aşağıdaki eşitsizlikler sağlanır:

(1) $\xi_0^t (x - x^*) + \gamma y \leq \alpha \ \ \forall x \in \mathbb{R}^n, y > f(x) - f(x^*)$

(2) $\xi_0^t (x - x^*) + \gamma y \geq \alpha \ \ \forall x \in S, y \leq 0$

(2) de $x = x^*, y = 0$ olarak alınırsa $\alpha \leq 0$ sonucuna varılır. (1) de $x = x^*, y = \varepsilon > 0$ olarak alınırsa $\gamma\varepsilon \leq \alpha$ elde edilir. Bu durumun her $\varepsilon > 0$ için sağlanması $\gamma \leq 0$ ve $\alpha \geq 0$ olmasına bağlıdır. (1) ve (2) de $\alpha \geq 0, \alpha \leq 0$ $\alpha = 0$ sonucuna götürmüştür. Bu takdirde $\gamma \leq 0$ ve $\alpha = 0$ olur. (1) de $\gamma = 0$ olarak alınırsa $\xi_0^t (x - x^*) \leq 0 \ \ \forall x \in \mathbb{R}^n$, $x = x^* + \xi_0$ olarak seçilirse

$\xi_0{}^t(x^* + \xi_0 - x^*) \leq 0 \quad \forall x \in \mathbb{R}^n \Rightarrow \xi_0 = 0$

$(\xi_0, \gamma) \neq (0,0) \Rightarrow \gamma < 0$

(1) ve (2) $-\gamma$ ile bölünürse ve $-\xi_0 / \gamma = \xi$ olarak adlandırılırsa aşağıdaki eşitsizlikler elde edilir:

(3) $\xi^t(x - x^*) \leq y \quad \forall x \in \mathbb{R}^n, y > f(x) - f(x^*)$

(4) $\xi^t(x - x^*) - y \geq 0 \quad \forall x \in S, y \leq 0$

(4) de $y = 0$ olarak alınırsa $\xi^t(x - x^*) \geq 0, \forall x \in S$, (3) te $y > f(x) - f(x^*) \geq 0$, $\forall x \in \mathbb{R}^n$ ve $y \geq \xi^t(x - x^*) \geq 0$ olduğundan

$\Rightarrow f(x) - f(x^*) \geq \xi^t(x - x^*)$

$\Rightarrow f(x) \geq \xi^t(x - x^*) + f(x^*), \forall x \in \mathbb{R}^n$

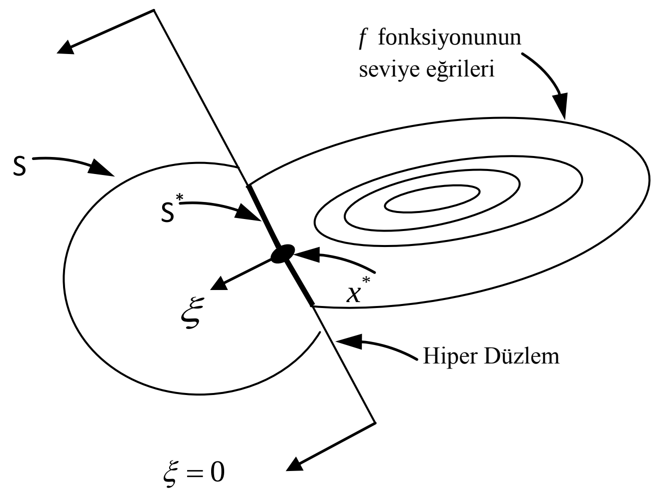
Subgradyentin tanım koşulu elde edilmiş olup, $f(x)$ fonksiyonu $\xi^t(x - x^*) \geq 0$ $\forall x \in S$ koşulunu sağlayan bir subgradyente sahiptir.

Sonuç 4.2.2.1.

Teorem 4.2.2. nin şartları sağlanmak üzere,

1. S açık bir küme ise $x^* \in S$ noktasının optimum çözüm olması için gerek ve yeter koşul $\xi = 0$ olmasıdır.

2. f türevlenebilir olsun. Bu takdirde $x^* \in S$ noktasının optimum çözüm olması için gerek ve yeter koşul $\nabla f(x^*)^t(x - x^*) \geq 0, \forall x \in S$ olmasıdır. S açık bir küme ise $x^* \in S$ noktasının optimum çözüm olması için gerek ve yeter koşul $\nabla f(x^*) = 0$ olmasıdır [3].



Şekil 4.1. Teorem 4.2.2. için çizim

İspat olası çözümler kümesinden yani S kümesinden, $f(x) \geq f(y), \ \forall x, y \in S,$ koşulunu sağlamayan noktaları ayırmak ve S^* optimum çözümler kümesi olmak üzere, $f(x) \geq f(x^*), \ \forall x \in S, \ \forall x^* \in S^*$ koşulunu sağlayan x^* noktalarını elde etmek üzerine kuruludur. Şekil 4.1. ispatı özetlemektedir. Teorem 4.2.2. $(x - x^*)$ vektörünün yönünün optimum çözüme giden yolda $\xi^t(x - x^*) \geq 0, \ \forall x \in S$ koşulu sağlandığı takdirde uygun bir yön olduğunu vurgulamaktadır. Bu koşul çözüme ne kadar yaklaşıldığının veya uzaklaşıldığının tespitinde önemlidir. Örneğin optimum olmayan bir noktaya erişildiğinde bu koşul $\xi^t(x - x^*) < 0, \ \forall x \in S$ olacaktır. Çözümü iyileştirmenin bir yolu da x^* noktasının $(x - x^*)$ yönünde belli bir parametreye bağlı olarak ötelenmesidir [3].

Örnek 4.2.2.

$(x_1 - 3/2)^2 + (x_2 - 5)^2$ fonksiyonunu $-x_1 + x_2 \leq 2, \quad 2x_1 + x_2 \leq 11, \ x_1, x_2 \geq 0$ kısıtlarına bağlı olarak minimumlaştırınız.

$(x_1 - 3/2)^2 + (x_2 - 5)^2$ fonksiyonu konveks ve olası çözümler kümesi konveks oldu ğundan bu problem bir Konveks Programlama Problemidir. Bu fonksiyon $(3/2, 5)$ noktasına olan mesafenin karesini vermektedir.

$$\nabla f(x)^t = (2(x_1 - 3/2), 2(x_2 - 5))$$

$$2(x_1 - 3/2) = 0 \Rightarrow x_1 = 3/2$$
$$2(x_2 - 5)) = 0 \Rightarrow x_2 = 5$$

$(3/2,5)$ noktasının olası çözümler kümesinde olmadığı Şekil 4.2'den açıkça görülmektedir. Ayrıca Şekil 4.2 ye bakıldığında $(1,3)$ noktasının $(3/2,5)$ noktasına en yakın nokta olduğu da görülmektedir. Bu noktadaki gradyent vektörü $\nabla f(1,3)^t = (-1,-4)$ dür. Yine şekilden anlaşılacağı gibi bu vektör $(x-x^*)^t = (x_1-1, x_2-3)$ formundaki her vektörle 90^o den küçük açı yapmaktadır.

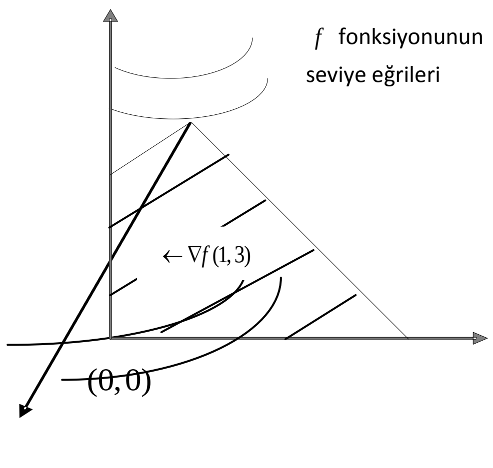
$0<\theta<90$ olduğundan $\cos\theta>0$ olur. Bu takdirde θ iki vektör arasındaki açı olmak üzere

$$\begin{aligned} < x, y > &= \nabla f(1,3)^t (x_1-1, x_2-3) \\ &= \|(-1,-4)\| \|(x_1-1, x_2-3)\| \cos\theta > 0 \end{aligned}$$

olur. Fonksiyonun *Hessian* matrisi incelenirse

$$\begin{aligned} x^t \mathrm{H} x &= (x_1, x_2)\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}(x_1, x_2)^t, \\ &= 2x_1^2 + 2x_2^2 > 0, \quad \forall (x_1, x_2) \in \mathbb{R}^2,\ (x_1, x_2) \neq (0,0) \end{aligned}$$

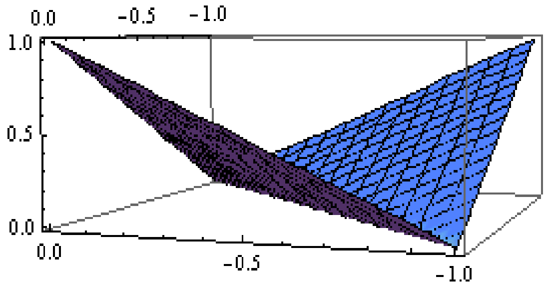
olduğundan, *Hessian* matrisi pozitif tanımlıdır. Dolayısıyla fonksiyon kesin konvekstir. Teorem 4.2.1. ve Teorem 4.2.2. nin gerek ve yeter koşulları sağlanmıştır. (1,3) tek mutlak minimumdur diğer bir deyişle problemin tek optimum çözümüdür.



Şekil 4.2. Örnek 4.2.2. için çizim

İkinci bir durum olarak (0,0) noktasının optimum çözüm olma ihtimali ele alınsın. Bu noktada hesaplanan gradyent vektörü (-3,-10) olmaktadır. $\nabla f(0,0)^t = (-3,-10)$ ve $(x-x^*)^t = (x_1, x_2)$ ve $x_1, x_2 \geq 0$ olduğundan $\nabla f(0,0)^t (x_1, x_2) = -3x_1 - 10x_2 < 0$ $\forall (x_1, x_2) \neq (0,0)$, elde edilir ve orijin optimum çözüm olamaz. [3]

Örnek 4.2.3. $f : \mathbb{R}^2 \to \mathbb{R}_+$, $f(x,y) = \|x+y+1\|$ fonksiyonunu $S = [-1,0] \times [0,-1]$ kümesi üzerinde minimumlaştırınız.



Şekil 4.3. Örnek 4.2.3. için çizim

Konveks kümelerin kartezyen çarpımı konveks olduğundan ve Şekil 4.3. den de anlaşılacağı üzere $f(x,y) = \|x+y+1\|$ fonksiyonu konveks olduğundan bu problem bir Konveks Programlama Problemidir. Norm fonksiyonu tanım kümesindeki her (x,y) için pozitif değer aldığından $f(x,y) = \|x+y+1\|$ fonksiyonunun minimum değeri sıfırdır. Verilen kümede bu fonksiyonu sıfır yapan iki nokta $x^* = (-1/2, -1/2)$, $y^* = (-1/4, -3/4)$ olarak seçilsin. f konveks ve tanım kümesi de konveks olduğundan x^*, $y^* \in \operatorname{int} S$ için aşağıdaki eşitsizlikler sağlanır.

$$1) \|x+y+1\| \geq 0 + \xi^t [(x,y) - (-1/2, -1/2)], \quad \forall (x,y) \in [-1,0] \times [0,-1]$$

$$2) \|x+y+1\| \geq 0 + \xi^t [(x,y) - (-1/4, -3/4)], \quad \forall (x,y) \in [-1,0] \times [0,-1]$$

$f(x,y) = \|x+y+1\|$ fonksiyonu $x_1^* = (-1/2, -1/2)$, $x_2^* = (-1/4, -3/4)$ noktalarında $\xi^t = (0,0)$ subgradyentine (1) ve (2) eşitsizlikleri sağlanmak üzere sahiptir. Teorem 4.2.2. uygulanırsa $\xi^t ((x,y) - (-1/2, -1/2))^t = 0$ ve $\xi^t ((x,y) - (-1/4, -3/4))^t = 0$ elde edilir. Her iki nokta da yerel optimum çözüm olup mutlak optimum çözümdür. Çözümün tek olmamasının nedeni ise fonksiyonun kesin konveks olmamasıdır.

Bu iki nokta arasında kalan her $(x, y) \in [-1, 0] \times [0, -1]$ ikilisi $\lambda \in [0,1]$ için kümenin konveks oluşundan dolayı

$$\begin{aligned}(x, y) &= \lambda(-1/2, -1/2) + (1-\lambda)(-1/4, -3/4) \\ &= (-\lambda/4 - 1/4, \lambda/4 - 3/4)\end{aligned}$$

şeklinde yazılır. Buradan $f(x, y) = \|-\lambda/4 - 1/4 + \lambda/4 - 3/4 + 1\| = 0$ elde edilir. $x_1^* = (-1/2, -1/2)$, $x_2^* = (-1/4, -3/4)$ noktaları arasında kalan her (x, y) ikilisi ayrı ayrı çözümdür.

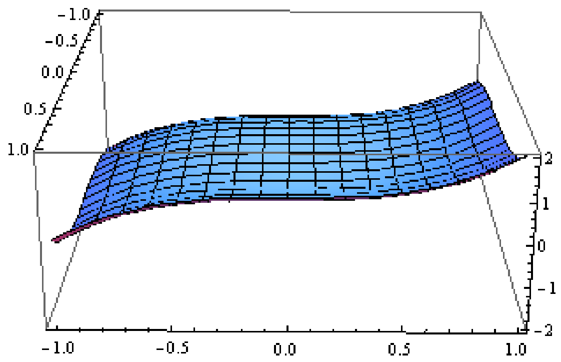
Örnek 4.2.4. $f : \mathbb{R} \to \mathbb{R}_+$ olmak üzere $f(x) = |x|$ olsun. $x^* = 0$ noktası, $f : \mathbb{R} \to \mathbb{R}_+$ fonksiyonu bu noktada $\xi = 0$ subgradyentine sahip olmak üzere Sonuç 4.2.2.1. gereği yerel ve mutlak minimumdur. Bu çözümün tek olduğu açıktır. Ancak mutlak değer fonksiyonu kesin konveks değildir.

Örnek 4.2.5. $f : \mathbb{R} \to \mathbb{R}$, $f(x, y) = x^3 + y^3$ fonksiyonunu $(-1,1) \times (-1,1)$ kümesi üzerinde minimumlaştırma problemi ele alınsın. Gradyent vektörünü sıfır vektörü yapan noktalar araştırılırsa

$$\nabla f(x, y)^t = (3x^2, 3y^2)$$

$$\begin{aligned}3x^2 = 0 &\Rightarrow x = 0 \\ 3y^2 = 0 &\Rightarrow y = 0\end{aligned}$$

elde edilir.



Şekil 4.4. Örnek 4.2.5. için çizim

Geometrik olarak incelendiğinde $f(x,y)=x^3+y^3$ fonksiyonunun konveks olmadığı anlaşılmaktadır. $\nabla f(0,0)^t=(0,0)$ olmasına rağmen Teorem 4.2.2.' nin fonksiyonun olası çözümler kümesi üzerindeki konvekslik koşulu sağlanmadığından orijin $f(x,y)=x^3+y^3$ fonksiyonu için $(-1,1)\times(-1,1)$ kümesi üzerinde bir yerel minimum değildir.

Konveks fonksiyonların maksimum değerleri ile ilgili olarak aşağıdaki teorem verilebilir.

Teorem 4.2.3. f**,** sınırlı ve kapalı C konveks kümesinde tanımlı bir konveks fonksiyon olsun. Eğer f, C üzerinde bir maksimuma sahipse, bu C ye ait bir ekstrem noktada gerçekleşir [13]-[14].

İspat: Farz edelim ki f, C kümesi üzerinde bir maksimum değere sahip olsun. Ekstrem nokta aynı zamanda bir sınır noktası olduğundan maksimuma erişilen noktanın bir sınır noktası olduğu ispatlanacaktır. Maksimuma erişilen nokta sınırdaysa ispat tamamlanmıştır. Maksimuma iç noktada erişiliyorsa bu erişimin aynı zamanda bir sınır noktasında olduğu ispatlanacaktır. Bunu ispatlamak için öncelikle iç noktalarda mutlak maksimuma erişildiği takdirde fonksiyonun küme üzerinde sabit bir fonksiyon teşkil ettiğini gösterelim. Kabul edelim ki fonksiyon küme üzerinde sabit olmasın. $x_1\in \operatorname{int} C$ olmak üzere x_1 mutlak maksimum nokta olsun. $x_2,x_3\in C$ seçelim öyle ki $f(x_2)<f(x_1)$, $f(x_3)<f(x_1)$ sağlansın ve küme konveks olduğundan $\alpha\in[0,1]$ olmak üzere $x_1=\alpha x_2+(1-\alpha)x_3\in C$, $x_2,x_3\in C$ yazılır. Bu takdirde fonksiyon konveks olduğundan

$$f(x_1) \le \alpha f(x_2) + (1-\alpha) f(x_3)$$

$$< \alpha f(x_1) + (1-\alpha) f(x_1)$$

$$= f(x_1)$$

yazılır. Bu bir çelişkidir. f küme üzerinde sabittir. Fonksiyon kapalı ve sınırlı bir kümede iç noktalarda mutlak maksimum değere erişmişse bunun aynı zamanda bir sınır noktasında gerçekleştiğini gösterelim. $x_1 \in \operatorname{int} C$ olmak üzere x_1 mutlak maksimum nokta olsun. $x_2, x_3 \in \partial C$ seçelim. Küme konveks olduğundan $\alpha \in [0,1]$ olmak üzere $x_1 = \alpha x_2 + (1-\alpha) x_3 \in C$ yazılır. Bu takdirde fonksiyon konveks olduğundan

$$f(x_1) \le \alpha f(x_2) + (1-\alpha) f(x_3), \quad \forall \alpha \in [0,1]$$

yazılır. $\alpha = 1$ için $f(x_1) \le f(x_2)$ ve $\alpha = 0$ için $f(x_1) \le f(x_3)$ olduğundan $x_2, x_3 \in \partial C$ noktaları da tıpkı x_1 gibi mutlak maksimum noktalardır. Maksimuma erişilen noktanın sınırda olduğu gösterilmiştir.

$x^* \in \partial C$ maksimuma erişilen bir nokta olmak üzere, x^* ekstrem nokta ise ispat tamamlanmıştır. Eğer ekstrem nokta değilse C kümesini sınırında destekleyen destek hiper düzlemi ile kesişimi ele alınsın. Öncelikle bu $B_1 = C \cap H$ kesişimi kapalı, konveks ve sınırlıdır. Ayrıca boyutu da n den küçüktür. Destek hiper düzlemin tanımı hatırlanırsa, $H = \{x \in \mathbb{R}^n \mid u^t x = \beta\}$ olmak üzere burada $\forall x \in C$ için $u^t x \le \beta$ ve en azından bir tane $x \in \partial C$ için $u^t x = \beta$ olmaktadır. Bu takdirde $B_1 = C \cap H$ kesişimi f nin B_1 üzerindeki mutlak maksimumu olan $f(x^*)$ a eşit olup B_1 in x_1 sınır noktasında gerçekleşir. Eğer bu nokta B_1 e ait bir ekstrem nokta ise aynı zamanda C ye ait bir ekstrem noktadır çünkü destek hiper düzlem afindir. Bu takdirde teorem ispatlanmış olur. Bu nokta B_1 e ait bir ekstrem nokta değilse bu kez B_1 ile $\mathbb{R}^{n-1}$ deki bir hiper düzlemin kesişimini ele alırız. Bu kesiştirmelere n kez devam edilirse n'ninci kesişimde tek bir noktadan oluşan B_n kümesi elde edilir. Bu nokta B_n ye ait bir ekstrem nokta olup aynı zamanda C ye ait bir ekstrem noktadır.

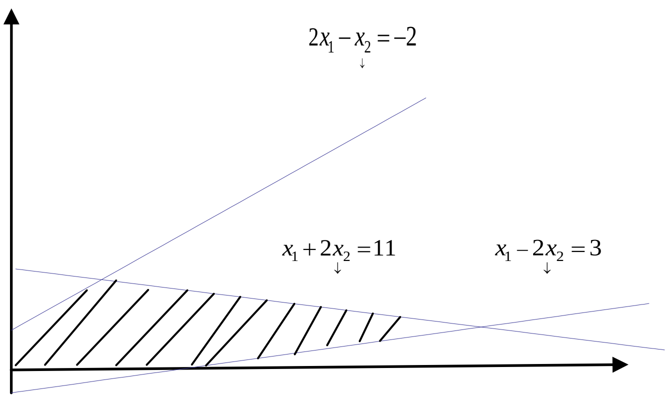
Örnek 4.2.6. Teoremi destekleyici bir örnek olarak aşağıdaki lineer programlama problemi ele alınsın. Lineer fonksiyonlar konvekslik ve konkavlık özelliğini bir arada taşıyan fonksiyonlardır. Ayrıca eşitsizlik kısıtı fonksiyonlarının çerçeveledikleri bölge de konvekstir. Özetle bu problem kapalı, sınırlı ve konveks bir kümede, konveks bir fonksiyonun maksimumlaştırma problemidir. [14]

$f(x_1, x_2) = -3x_1 + 2x_2$ fonksiyonunu

$$\begin{aligned} &2x_1 - x_2 \geq -2 \\ &x_1 - 2x_2 \leq 3 \\ &x_1 + 2x_2 \leq 11 \\ &x_1,\ \ x_2 \geq 0 \end{aligned}$$

kısıtlarına bağlı olarak maksimumlaştırınız.

Ekstrem noktaları tespit etmek için $2x_1 - x_2 = -2,\ x_1 - 2x_2 = 3,\ x_1 + 2x_2 = 11,$ $x_1 = 0,\ x_2 = 0$ doğrularının birbirlerine göre durumlarını grafikle inceleyelim.



Şekil 4.5. Örnek 4.2.6. için çizim

Grafikten elde edilen kesişim durumlarına göre ekstrem noktalar

$x_1 = 0$ ve $2x_1 - x_2 = -2$ için (0,2),

$x_2 = 0$ ve $x_1 - 2x_2 = 3$ için (3,0),

$x_1 + 2x_2 = 11$ ve $x_1 - 2x_2 = 3$ için (7,2),

$2x_1 - x_2 = -2$ ve $x_1 + 2x_2 = 11$ için (7/5,24/5)

olarak bulunur. Bu noktalar fonksiyonda yerine konulursa,

$f(0,2) = 4$, $f(3,0) = -9$, $f(7,2) = -17$, $f(7/5, 24/5) = 27/5$

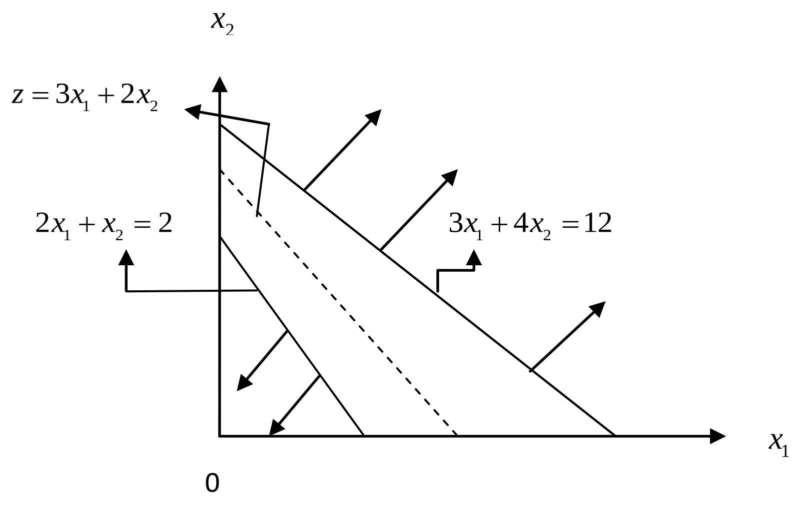
değerleri elde edilir. Maksimum değeri veren nokta (7/5,24/5) olduğundan problemin optimum çözümüdür.

Örnek 4.2.7.

$z = 3x_1 + 2x_2$ fonksiyonunu,

$$
\begin{aligned}
&2x_1 + x_2 \le 2 \\
&3x_1 + 4x_2 \ge 12 \\
&x_1,\ x_2 \ge 0
\end{aligned}
$$

kısıtlarına bağlı olarak maksimumlaştırınız.[15]



Şekil 4.6. Örnek 4.2.7. için çizim

Problemin çözümü Simpleks Yöntemi ile ele alınmıştır.

	x_1	x_2	x_4	x_3	x_5	Çözüm
x_3	2	1	0	1	0	2
x_5	3	4	-1	0	1	12
z	$-3-3M$	$-2-4M$	M	0	0	$-12M$

Tablo 4.1. Başlangıç Simpleks Tablosu

	x_1	x_2	x_4	x_3	x_5	Çözüm
x_2	2	1	0	1	0	2
x_5	-5	0	-1	-4	1	4
z	$1+5M$	0	M	$2+4M$	0	$4-4M$

Tablo 4.2. Sonuç Simpleks Tablosu

Tablo 4.2. x_5 yapay değişkeninin pozitif (=4) değer aldığını, dolayısıyla da problemin optimum çözümünün olmadığını göstermektedir. Şekil 4.6. ise problemin olası çözümler kümesinin açık ve sınırsız olduğunu göstermektedir. Örnek 4.2.7., Teorem 4.2.3. de belirtilen şartlardan maksimumlaştırılacak fonksiyon için verilen kümenin kapalı ve sınırlı olmadığı duruma bir örnek olarak verilmiştir.

BÖLÜM 5. SONUÇLAR

Tezin birinci bölümü giriş bölümüdür. Tezin ikinci bölümünde sonraki ispatlara baz teşkil edecek tanım ve teoremler verilmiştir. Tezin üçüncü bölümünde konveks kümeler ve konveks fonksiyonlar detaylı bir şekilde incelenmiş ve ayırma teoremleri verilmiştir. Akabinde türevlenebilir konveks fonksiyonların özellikleri teorik olarak incelenmiş ve verilen bir fonksiyonun konveksliğinin tespiti için bazı özellikler verilmiştir. Dördüncü bölümde konveks fonksiyonların konveks kümeler üzerindeki minimumlaştırma problemine ilişkin optimum çözümün var olduğu noktada fonksiyonun sağlaması gereken koşullar subgradyent kavramı üzerinden verilmiştir. Bu koşulların sağlandığı ve sağlanmadığı durumlar ayrı ayrı örneklerle incelenmiştir. Konveks Programlama Problemine ilişkin olarak fonksiyon verilen konveks küme üzerinde konveks değilse gradyent vektörünü (genel halde verilen noktada sıfır subgradyent vektörüne sahip olmasının) sıfır yapan noktanın ekstremum teşkil etmeyebileceği görülmüştür. Ayrıca fonksiyonun kesin konveks olmadığı durumda da bulunan optimumun tek olabildiği ve olamadığı durumlar örneklerle incelenmiştir. Son olarak verilen kompakt konveks bir kümede tanımlı konveks bir fonsiyonun maksimum değerine ulaşıyorsa bunun kümeye ait bir ekstrem noktada gerçekleşeceği üzerine bir teorem verilmiş olup bu teoremin hipotezinin sağlandığı ve sağlanmadığı durumlar örneklerle incelenmiştir.

KAYNAKLAR

[1] BAYRAKTAR, M., Fonksiyonel Analiz, Gazi Kitabevi, Ankara, Ekim 2006, s.23-82.

[2] WADE R.W., An Introduction To Analysis, Pearson Prentice Hall, 2003, s.258-292.

[3] BAZAARA, M., S., SHERALI, H.D, SHETTY, C.M., Nonlineer Programming, Wiley Interscience, 2006, s.40-763.

[4] MUSAYEV, B., ALP M., Fonksiyonel Analiz, Balcı Yay., Ankara, Kasım 2000, s.36-98.

[5] ROCKAFELLAR, R.T., Convex Analysis, Princeton Uni. Press, 1970, s. 10-26.

[6] KINCAID, D., CHENEY, W., Numerical Analysis, Brooks/Cole Pub. Comp.,1996, s.737- 738.

[7] BOYD, S., VANDENBERGHE, L. ,Convex Optimization, Cambridge Uni. Press, 2004, s.24-81.

[8] EGGLESTON, H. G., Convexity, Cambridge At the Uni.Press., 1958, s.9-10.

[9] LAURITZEN, N., Lectures On Convex Sets, University of Aarhus, Denmark, March 2009, s.15-33.

[10] TIEL, J. V., Convex Analysis, John Wiley & Sons Ltd., 1984, s.8-73.

[11] KARA, İ., Yöneylem Araştırması , Anadolu Üni. Basımevi, Eskişehir, 1986, s.2-40.

[12] TÜRKER E.S., Çözümlü Örneklerle Yüksek Matematik, Alfa Yay., Ekim 1992, s.162-163.

[13] ÜNSAL, N, Kısıtlamasız Minimumlaştırma Problemleri, Yüksek Lisans Tezi, Gazi Üni. Fen Bil. Ens. Ankara, Aralık 2005, s.23.

[14] ROBERTS, A.W., WARBERG D.E., Convex Functions, Academic Press, 1973, s.124-139.

[15] TAHA H. A., Yöneylem Araştırması (6. Basımdan Çeviri), Çevirenler: BARAY Ş. A., ESNAF Ş., İstanbul, Literatür Yayınevi, 2000, s.106-107.

ÖZGEÇMİŞ

Gümrah UYSAL, 29.10.1984 yılında Sakarya'nın Adapazarı ilçesinde doğdu. İlk ve orta öğrenimini Zonguldak Ereğli'de tamamladı. Lise öğrenimini ise Sakarya' da tamamladı. 2008 yılında Dokuz Eylül Üniversitesi Matematik Bölümü'nden mezun oldu. 2008 yılında Sakarya Üniversitesi Matematik Bölümü'nde Yüksek Lisans eğitimine başladı. 2009 yılında Karabük Üniversitesi Matematik Bölümü'ne araştırma görevlisi olarak atandı. Bu göreve halen devam etmektedir.