İSTANBUL ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

YÜKSEK LİSANS TEZİ

SONLU BOYUTLU UZAYLARDA MATEMATİK PROGRAMLAMA VE DUALİTE

SEVİLAY DEMİR
Matematik Anabilim Dalı

Danışman
Yard.Doç.Dr. Özkan DEĞER

Aralık, 2011

İSTANBUL

İSTANBUL ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

YÜKSEK LİSANS TEZİ

SONLU BOYUTLU UZAYLARDA MATEMATİK PROGRAMLAMA VE DUALİTE

SEVİLAY DEMİR
Matematik Anabilim Dalı

Danışman
Yard.Doç.Dr. Özkan DEĞER

Aralık, 2011

İSTANBUL

Bu çalışma 22/12/2011 tarihinde aşağıdaki jüri tarafından Matematik Anabilim Dalında Yüksek Lisans Tezi olarak kabul edilmiştir.

Tez Jürisi

Yard.Doç.Dr.Özkan DEĞER (Danışman)
İstanbul Üniversitesi

Prof.Dr.Nazım SADIK
İstanbul Üniversitesi

Prof.Dr.Leyla ZEREN AKGÜN
İstanbul Üniversitesi

Prof.Dr.İ.Müfit GİRESUNLU
İstanbul Üniversitesi

Doç.Dr.Fatma ÖZDEMİR
İstanbul Teknik Üniversitesi

Bu çalışma İstanbul Üniversitesi Bilimsel Araştırma Projeleri Yürütücü Sekreterliğinin 14424 numaralı projesi ile desteklenmiştir.

ÖNSÖZ

Yüksek lisans eğitim boyunca bana her türlü konuda destek olan, tez çalışmamda beni yönlendirip, bilgi ve tecrübesiyle bana yol gösteren, bu tezin yazımında kullanılan LaTeX programını öğrenmem ve ilerletmemde yardımcı olan değerli hocam ve danışmanım Yard.Doç.Dr. Özkan DEĞER'e sonsuz teşekkürlerimi sunmayı bir borç bilirim.

Ayrıca tez çalışmam boyunca tanıdığı imkanlar ve bilimsel bilgilerin yanı sıra olumlu düşünceleri ile beni her zaman teşvik eden kıymetli hocam Prof.Dr. Nazım SADIK'a ve bana tezimin yazımı için her türlü kaynağı sağlayan İstanbul Üniversitesi Bilimsel Araştırma Projeleri Yürütücü Sekreterliğine çok teşekkür ederim.

Tüm hayatım boyunca vermiş oldukları maddi ve manevi destekle bu aşamaya gelmemde büyük emekleri olan babam Halil DEMİR'e, annem Gülüzar DEMİR'e, ablam Nurten ve abim Serkan DEMİR'e, beni sabırla dinleyen ve bana cesaret veren sevgili arkadaşlarıma çok teşekkür ederim.

Aralık, 2011 **Sevilay DEMİR**

İÇİNDEKİLER

ŞEKİL LİSTESİ

SEMBOL LİSTESİ

$\langle \cdot , \cdot \rangle$	: İç çarpım
$\boldsymbol{B(x,\varepsilon)}$	: x merkezli ε yarıçaplı açık yuvar
$\boldsymbol{B[x,\varepsilon]}$	: x merkezli ε yarıçaplı kapalı yuvar
$\boldsymbol{int(A)}$	: A kümesinin iç noktalar kümesi
$\boldsymbol{\overline{A}}$	: A kümesinin kapanışı
affA	: A kümesinin afin zarfı
$\boldsymbol{conv(A)}$	: A kümesinin konveks zarfı
$\boldsymbol{\overline{conv}(A)}$	: A kümesinin kapalı konveks zarfı
$\boldsymbol{ri(A)}$	: A kümesinin izafi iç noktalar kümesi
$\boldsymbol{K^*}$	: K konisinin dual konisi
$\boldsymbol{dom\ f}$	: f fonksiyonunun tanım kümesi
$\boldsymbol{epi\ f}$	: f fonksiyonunun grafiküstü kümesi
$\boldsymbol{hyp\ f}$	: f fonksiyonunun grafikaltı kümesi
$\boldsymbol{f^*}$	: f fonksiyonunun eşlenik fonksiyonu
$\boldsymbol{\delta_A(x)}$	: A konveks kümesinin indikatör fonksiyonu
$\boldsymbol{\partial f(x_0)}$	: f fonksiyonunun x_0 noktasındaki subdiferansiyel kümesi
$\boldsymbol{W_A(\cdot)}$	: A kümesinin destek fonksiyonu
$\boldsymbol{L_\alpha(f)}$	: f fonksiyonunun seviye kümesi
$\boldsymbol{Df(x)(y)}$	: f fonksiyonunun x noktasında y yönündeki yönlü türevi
$\boldsymbol{co(f)}$	: f fonksiyonunun konveks zarf fonksiyonu
$\boldsymbol{\overline{co}(f)}$	: f fonksiyonunun kapalı konveks zarf fonksiyonu
$\boldsymbol{clf(x)}$	: f fonksiyonunun kapanış fonksiyonu
$\boldsymbol{f\ \nabla g}$	: f ile g fonksiyonlarının infimal konvolüsyonu
$\boldsymbol{f_*}$	: f fonksiyonunun konkav eşlenik fonksiyonu
$\blacksquare$	: İspat bitmiştir.

ÖZET

SONLU BOYUTLU UZAYLARDA MATEMATİK PROGRAMLAMA VE DUALİTE

Bu tez çalışmasında öncelikle konveks analizin temelini oluşturan konveks küme ve konveks fonksiyonların temel özellikleri verildi. Dualite teorisinde önemli bir yere sahip olan konveks eşlenik fonksiyonların yapıları incelenerek, konveks fonksiyon ile konveks eşlenik fonksiyon arasındaki ilişkiler araştırıldı. Daha sonra bazı optimizasyon problemleri tanıtıldı ve bu problemlerin dual problemleri kurulup genelliği bozmaksızın incelenen minimum optimizasyon probleminin çözümü ile dual problemin çözümü arasındaki bağıntı verildi. Son olarak eşlenik fonksiyon yardımı ile Fenchel Dualitesi kullanılarak konveks optimizasyon problemlerinin çözümleri incelendi.

SUMMARY

MATHEMATICAL PROGRAMMING AND DUALITY IN FINITE DIMENSIONAL SPACES

In this study, initially the basic properties of convex set and convex function which are formed the basis of the convex analysis, were given. By examining structures of convex conjugate functions which have an important place in duality theory, the relationships between the convex functions and convex conjugate functions were researched. Then some of optimization problems were defined and the relationship between solutions of primal problems and the dual problems were given by constructing dual problems of these optimization problems. Finally, solutions of convex optimization problems were examined with the help of the conjugate function using Fenchel duality.

1. GİRİŞ

İnsanoğlu yüzyıllardır bilinçli ya da bilinçsiz olarak yaptığı işlerin tümünde her zaman en iyiyi yapmayı planlamış ve istemiştir. Bu çabalar, bazen bir işçinin belirli bir kuvvetle maksimum yükü kaldırabilmesi, bazen de uçaklardaki sürtünmenin minimuma indirgenmesi olarak karşımıza çıkmaktadır. Ünlü matematikçi Leonhard Euler'in " Doğada maksimum ya da minimum duyusunun bulunmadığı hiç bir olay yoktur" ifadesinde belirtiği gibi her zaman bir eniyileme çabasının varlığı yadsınamaz bir gerçektir. Türkçe'de en iyi anlamına gelen aslında kökeni Latince *optimus* olan optimal kelimesi, matematik analizin özel bir kolu olan matematik programlamada önemli bir yer tutar. Matematik programlama, bir sistemde varolan kaynakların en verimli şekilde kullanılarak belirli amaçlara ulaşmayı sağlayan bir teknoloji olarak tanımlanan optimizasyon kavramı ile eşanlamlı olarak kullanılır. Diğer bir ifadeyle, matematik programlama, optimizasyon modelinin kurulması ve çözümün elde edilmesi işlemine verilen genel isimdir. Matematiksel olarak ise, matematik programlama ya da optimizasyon terimi bir reel değerli fonksiyonu minimize ya da maksimize etmek amacı ile reel ya da tamsayı değerlerini, tanımlı bir aralıkta seçip fonksiyona yerleştirerek sistematik olarak bir problemi incelemek ya da çözmek işlemlerini ifade eder.

Tarihte ilk optimizasyon problemi, Fenikeli bir prenses olarak tanınan Dido'nun Kral olan kardeşinin kocasını öldürmesi nedeniyle ülkesinden kaçarak Kuzey Afrika'da bir bölgenin kralının, prensese yalnızca bir sığır derisinin kaplayabileceği alan kadar toprak parçasına yerleşmesine izin vermesi üzerine, Prenses Dido'nun en büyük alanlı toprak parçasını nasıl bulabileceği problemi tarihte Dido problemi olarak bilinir. Matematiksel olarak ise "verilen uzunluktaki kapalı eğriler arasında en geniş alanı bulmak" olarak problem tanımlanır. Prenses Dido sığır derisini ince şeritler halinde kestirip bunları birbirine bağlatır ve uzun bir kordon elde eder. Kordonu bir çember şeklinde yaydırır ve 3 ile 12.5 km^2 arasında bir alan elde ederek problemi doğru bir şekilde çözer. Eşit çevre uzunluğuna sahip düzemsel şekiller arasında en fazla alanın daireye ait olması gerçeğini daha sonra matematikçiler arasında kanıtlamak o kadar kolay olmaz. M.Ö. üçüncü ya da dördüncü yüzyılda Yunanlı matematikçi *Zenodoros*, bu konuda ilk girişimleri yapmakla beraber kanıtı bazı açıklar içermektedir. Bu açıklar ise ancak 19. yüzyılda Alman matematikçi *Weierstrass* tarafından kapatılmıştır. 1696 yılında Johan Bernoulli'nin ortaya attığı sürtünmesiz ortamda verilen iki nokta arasında bir cismin yerçekimi nedeniyle yapacağı yer değiştirmenin minimum zamanda olmasını sağlayacak yörüngenin bulunması problemi yunanca *brachistrochrone* yani en kısa zaman problemi olarak bilinir. Brachistrochrone problemine kadar, tek değişkenli fonksiyonların minimum ya da maksimumları bulunmaya çalışılırken bu problemden sonra daha fazla değişkenli fonksiyonlar için optimizasyon problemleri ile uğraşılmaya başlandı. 18. yüzyıldan itibaren matematiğin bu tip problemlerle çalışan koluna *varyasyonlar analizi* denilmektedir [1, 2]. Düz bir yolda sürtünmesiz bir şekilde hareket eden, belirli sınırlar içerisinde değiştirilebilen bir dış güç tarafından kontrol edilen, hareket halindeki bir arabanın belli bir yerde en kısa zamanda durdurulması problemi *en basit zamanlı optimal problemi* olarak adlandırılır. 1950'lerden sonra, bunun gibi bir takım kontrol edilebilir güçleri bünyesinde barındıran problemlerin çözümleri ile uğraşan ve adına *optimal kontrol teorisi* denilen bir teori gelişmeye başlanmıştır [3]. Küme değerli fonksiyonlar ve yaklaşım metodları açısından matematik programlamanın temel kavram ve prensipleri, adi ve kısmi diferansiyel içermeler ile verilen problemlerin karşılaştırılmalı optimallik teorisi Mahmudov'un [4] kitabında ayrıntılı olarak verilmektedir.

Optimizasyon problemleri hem doğal bilimlerden, ekonomiden ve teknolojiden hem de matematiğin kendi ihtiyaçlarından ortaya çıkmaktadır. Matematik programlama teknikleri çok geniş bir yelpazede kullanım alanlarına sahiptirler. Mühendislikten işletme ve ekonomiye, askeri modellerden tarıma, tıp ve ilaç sektöründen spora kadar birbirlerinden çok farklı alanlarda geniş ölçüde kullanılmaktadır. Daha özel olarak uydu yörüngelerinin düzenlenmesi, finansal planlama, taşımacılık problemleri, spor liglerinin optimizasyonu, askeri hedeflerin vurulmasında optimal silah karışımının belirlenmesi ve radyoterapide ışınların optimal açı ve yoğunluklarının belirlenmesi bunlardan sadece bir kaçıdır.

Genel bir optimizasyon problemi,

$$\min f(x) \tag{1.1}$$
$$x \in X$$

şeklinde verilir. Optimize edilecek yani maksimize ya da minimize edilecek $f(x)$ fonksiyonuna amaç fonksiyonu, $X \subset \mathbb{R}^n$ kümesine ise kısıt kümesi denir. Eğer X kısıt kümesi $\mathbb{R}^n$ ise (1.1) problemi kısıtsız optimizasyon problemi adını alır. Kısıtlı optimizasyon problemi ise $i = 1, \ldots, m$ için

$$\begin{aligned} \min\ & f(x) \\ & g_i(x) \leqslant 0 \\ & x \in \mathbb{R}^n \end{aligned} \tag{1.2}$$

şeklinde tanımlanır. $g_i(x)$ kısıt fonksiyonlarının sayısında herhangi bir sınır bulunmadığı gibi kısıtların hepsi birlikte bir uygun çözüm bölgesi belirlerler. Optimal çözüm değeri veya değerleri bu bölgeye ait bir değer olmaktadır. Buna göre, optimizasyon problemleri birçok özellik dikkate alınarak sınıflandırılabilirler. Bu sınıflandırmaların bazıları amaç ve kısıt fonksiyonlarının lineer, nonlineer veya konveks olmasına göre, karar değişkenlerinin tipine göre, değişken sayısının çok veya az olmasına göre olabileceği gibi fonksiyonların düzgün yani diferansiyellenebilir olmasına veya olmayışına göre de yapılabilir. Fakat bu sınıflandırmalar içerisinde en yaygın olarak kullanılan kısıtların varlığına göre yapılan sınıflandırmadır. Bu bilgi ışığında optimizasyon problemleri kısıtlı veya kısıtsız optimizasyon problemleri biçiminde ikiye ayrılabilir. Kısıtsız optimizasyon problemleri kısıtlı optimizasyon problemlerinde kısıtların özel sabitler ile çarpılarak amaç fonksiyonuna eklenmesiyle elde edilebilirler. Kısıtlı optimizasyon probleminde ise değişkenler sınırlandırılmıştır. Eğer amaç fonksiyonu ve kısıtlar x değişkenine bağlı birer lineer fonksiyonlar ise bu probleme lineer programlama problemi denir. Bunlardan en azından biri nonlineer bir fonksiyon ise bu problem türüne nonlineer programlama problemi denir. Hem kısıtlı optimizasyon problemlerini hem de kısıtsız optimizasyon problemlerini genel olarak bu iki alt başlık altında inceleyebiliriz. Bu tez çalışmasında ise nonlineer programlama içerisinde yer alan konveks programlama problemleri üzerinde incelemeler yapılmaktadır [5, 6]. Öte yandan Bertsekas' ın 1995 yılında yazdığı [7] kitabı optimizasyon problemlerinin yapısının daha iyi anlaşılması için önemli bir kaynak niteliğindedir.

Optimizasyon problemlerini çözmek için çeşitli metodlar tarihten bugüne kadar bir çok gelişim göstermiştir. Bu metodlardan birincisi optimizasyon teorisinin gelişiminde önemli bir kilometre taşı olan Newton ve Leibniz tarafından 17. yüzyılda geliştirilen Kalkülüs olmuştur. Ancak, Kalkülüs uygulamalarında karşılaşılan cebirsel problemlerin çözümü bazen güç olabilmektedir. Dolayısıyla, Kalkülüs gerçek optimizasyon problemlerin çözümünde yeterli ve güçlü bir araç olamamaktadır. Daha sonra J.L. Lagrange 'ın 1788 yılında Lagrange çarpanlar yöntemi olarak bilinen meşhur metodu, bilim dünyasında optimizasyon problemlerinin çözülmesi için önemli bir adım olmuştur ve günümüzde de optimizasyon teorisinin ana konularından birini oluşturmaktadı. 1939 'da W. Karush'un kısıtlandırılmış problemler için optimallik koşullarını bulması optimizasyon teorisinde yeni bir atılım başlatmıştır. II. Dünya Savaşının başlamasıyla 1942 'de İngiltere ve Amerika Birleşik Devletlerinin Yöneylem Araştırması gruplarını oluşturması optimizasyon dünyası için bir dönüm noktası olmuştur. II. Dünya Savaşından sonra geliştirilen yeni sınıf optimizasyon teknikleri daha karmaşık problemlere başarıyla uygulanmıştır. Bunda, yüksek hızlı dijital bilgisayarların geliştirilmesi ve optimum değerlerin elde edilmesi için, nümerik tekniklere matematiksel analizin uygulanması son derece etkili olmuştur. Nümerik teknikler Kalkülüs 'ün bir takım zorluklarını da ortadan kaldırmıştır.

Lineer programlama; bir amacın gerçekleşme derecesini etkileyen bazı kısıtlayıcı koşulların bulunması ve bunların lineer eşitlik veya eşitsizlik olarak verilmesi durumunda, bu amaca en iyi biçimde ulaşılması için kıt kaynakların en verimli biçimde kullanılmasını sağlayan bir matematik programlama yöntemidir [8, 9]. Bilinen en eski optimizasyon problemi olan lineer programlama problemlerin pratik olarak çözümlenmesi için ilk kullanışlı algoritma 1947 'de George Dantzig'in özel araştırma ekibinin ortaya attığı simpleks algoritmasıdır. Bu algoritma uygun çözümler kümesi denilen kısıtların oluşturduğu polihedral bölgenin köşelerinin birer uygun çözüm noktası olduğu fikrini benimsemektedir. Uygun çözümler kümesinin bir köşe noktasından başlayıp bu köşeyi belirleyen kenarlar takip edilerek amaç fonksiyonuna en iyi sonuç veren köşe noktası belirlenir. En iyi sonuç sağlayan köşenin koordinatları optimum çözüm olarak kabul edilmektedir. Bu algoritma, optimizasyon dünyasında gerçekten bir devrim sayılmaktadır. 1951 'de H. Kuhn ve A. Tucker daha önce Karush'un önerdiği kısıtlandırılmış problemler için optimallik koşullarını tekrar formüle ederek lineer olmayan programlama modelleri üzerinde çalışmalar yapmışlardır. Amaç fonksiyonu kuadratik fonksiyon olan bir nonlineer programlama çeşiti olarak bilinen Kuadratik programlama ise 1956 'da M. Frank ve P. Wolfe tarafından geliştirildi. Tüm optimizasyon problemlerini çözmek için tek bir metod şuana kadar mevcut değildir. Söz gelimi lineer programlar için geliştirilen Simpleks yöntem tüm lineer modelleri çözme potansiyeline sahipken lineer olmayan programlama modellerinin hepsini çözebilen genel bir çözüm yolu değildir. Ya da eşitlik kısıtlı lineer olmayan modellere Lagrange çarpanları metodu kullanılırken eşitsizlik kısıtlı problemlere de Kuhn-Tucker koşulları uygulanmaktadır. Ayrıca bu metodlardan bir kısmı türevi temel alırken diğer kısmı türeve dayalı değildir. Genel olarak türeve dayalı algoritmalar diğerlerine göre optimum sonuca ulaşmada daha başarılı metodlardır.

Konveks optimizasyon problemlerinde, reel değerli konveks bir fonksiyon konveks kümeler ile verilen bir kısıt bölgesinde minimum yapılmak istenmektedir. Minimize edilecek fonksiyonun diferansiyellenebilir olmadığı durumlarda bile konveksliğin getirdiği daima var olan yönlü türev ve subgradient kavramları sayesinde problemin çözümü diferansiyellenebilir olma halindeki gibi bulunur. Daha kapsamlı ve derinlemisine bilgi için [10, 11] bakılabilir. Özel olarak kümelerin ve içermelerin polihedral olduğu ve küme değerli dönüşümler yardımı ile verilebilen bazı konveks optimizasyon problemlerinin çözümü için gerek ve yeter koşullar [12]'de incelenmiştir. Konveks fonksiyon ve kümelerin özelliklerinin çalışıldığı matematik dalına Konveks Analiz denilmektedir. Bu tez çalışmasında konveks analize ait gerekli kavramlar ve önemli özellikler ikinci bölümde verilmektedir. Daha derinlemesine bilgi için konveks analiz ile ilgili en temel bilgileri içeren [13, 14, 4, 15, 16, 17] kitaplarına başvurulabilir. Özel olarak sonlu boyutlu uzaylarda konveks analiz ile ilgili bilgiler için [5] ve [2] kitaplarını da bakılabilir. Öte yandan konveks analizin kurucularından kabul edilen Rockafellar'ın 1972 yılında yazdığı Konveks Analiz adlı kitabı [13] bu dalda yazılmış en temel ve klasik kaynaktır.

Primal optimizasyon probleminden belirli matematiksel işlemle türetilen yeni optimizasyon problemine dual problem denir. Dual ve primal problemler birbirleri ile çok yakın ilişkili olup bulgular kısımında verilecek olan dualite teoremleri ile, dual ve primal problemlerin optimum çözümleri arasındaki bağıntılar incelenmiştir. Dualite teorisi ilk defa 1947 'de John Von Neumann tarafından ortaya atılmıştır. George Dantzig, genellikle oyunlar teorisinde çalışmalar yapan John Von Neumann'a ilk kez lineer programlama probleminden bahsettiğinde, J.V. Neumann lineer programlama problemlerinin çiftler halinde var olduğunu öne sürmüştür. Bu varsayım 1948 yılında David Gale, Harold W.Kuhn ve Albert W.Tucker tarafından ispatlanmaktadır. Bu sebeple George Dantzig' in dediği gibi “ Bugün herkes dualite teorisinin yaratıcısı olarak John Von Neumann bilir fakat ilk defa dualite teorisini yayınlayanlar Tucker, Kuhn ve Gale' dir.” Albert Tucker ile Princeton Üniversitesi'nde iken yapılan röportajdan alıntı yapmak bu noktada daha gerçekçi bir yaklaşım sağlayacaktır; “ O zamanlar Pentagon'daki Hava Kuvvetleri'nde istatistikçi olarak çalışan George Dantzig, 1948 'in bahar aylarında John Von Neumann'ı ziyaret etmek için Princeton'a geldiğinde tesadüf eseri onunla tanıştım. Lineer programlamanın ne olduğunu sordum, bir taşıma problemi örneği vererek lineer programlamaya bir giriş yapmıştı. Daha sonra öğrencilerim olan David Gale ve Harold Kuhn ile birlikte çalışmalar yaparken dualiteyi tanımladık.” Böylece John Von Neumann tarafından ortaya atılan ilk dualite teorisi David Gale, Harold W.Kuhn ve Albert W.Tucker tarafından ispatlanmaktadır.

Dualite teorisine göre, her maksimizasyon problemine karşılık onunla ilişkili bir minimazsyon problemi ya da her minimazsyon problemine karşılık onunla ilişkili bir maksimizasyon problemi bulunur. 1948 'de Gale ve Kuhn lineer programlamanın bir maksimizasyon probleminde Lagrange Çarpanları yöntemi uygulayarak Lagrange çarpanlarının dual değişkenler olduğunu göstermektedir. Lineer programlama için bulunan dualite teorisi daha sonra nonlineer programlama problemleri için de geliştirilmektedir. Bu teoriye ait daha derinlemesine bilgiye [7] kitabından ulaşılabilir. Konveks programlama problemleri için ise dual ve primal problemlerin optimum çözümleri arasındaki bağıntılar hakkındaki detaylı bilgiye [18, 19, 20] kaynaklarından ulaşılabilir. Ayrıca özel olarak parabolik tipli diferansiyel içermeli problemlerin optimallik için gerek ve yeter koşullarını ve bu problemlerin dualleri arasındaki ilişkileri Mahmudov'un [21] çalışmasında mevcuttur.

Tezde ele alınan problemler incelenirken gerekecek temel tanım, teorem ve yöntemler Genel Kısımlar başlığı altında ikinci bölümde, Tez çalışmasının asıl konusu olan konveks programlama problemi ile dual probleminin ilişkileri Bulgular başlığı altında üçüncü bölümde, elde edilen bulgular hakkındaki görüş ve yorumlamalar Tartışma ve Sonuç başlığı altında dördüncü bölümde verilmiştir.

2. GENEL KISIMLAR

Bu kısımda Konveks Analize ait olan temel kavramlar ve onların bazı özellikleri verilecektir. Bu tez çalışması boyunca $\mathbb{R}$ reel sayılar kümesi olmak üzere $i = 1, \ldots, n$ için $x_i \in \mathbb{R}$ gerçekleyen $x = (x_1, \ldots, x_n)$ vektörlerinin n-boyutlu Öklid uzayı $\mathbb{R}^n$ ile gösterilecektir. $x, y \in \mathbb{R}^n$ vektörleri için iç çarpım

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

ve bu iç çarpım yardımıyla da bir $x \in \mathbb{R}^n$ vektörünün normu $||x|| = \sqrt{\langle x, x \rangle}$ şeklinde tanımlanacaktır. $\mathbb{R}^n$ uzayındaki x ve y noktaları arasındaki Öklid uzaklığı

$$d(x, y) = ||x - y|| = \sqrt{\langle x - y,\ x - y \rangle},$$

bu uzaydaki açık ve kapalı birim yuvarlar sırasıyla

$$B(0, 1) = \{x : ||x|| < 1\}, \quad B[0, 1] = \{x : ||x|| \leqslant 1\},$$

ve A kümesinin tüm iç noktalarının kümesi

$$int(A) = \{x \in A : \exists\, \varepsilon > 0,\ B(x, \varepsilon) \subseteq A\}$$

şeklinde tanımlanır. Her noktası iç nokta olan kümeye *açık küme*, her $\varepsilon > 0$ için $\big(B(a, \varepsilon) \backslash \{a\}\big) \cap A \neq \emptyset$ ise a noktasına A kümesinin *yığılma noktası* denir. A kümesine tüm yığılma noktalarının eklenmesiyle elde edilen kümeye A kümesinin *kapanışı* denir ve $\overline{A}$ ile gösterilir. Diğer yandan bu küme A kümesini kapsayan tüm kapalı kümelerin kesişimi ile çakıştığından

$$\overline{A} = \bigcap_{\varepsilon > 0} (A + \varepsilon B[0, 1])$$

notasyonu ile de ifade edilir.

$\mathbb{R}^n$ uzayındaki $x_1, \ldots, x_m$ vektörlerinin *afin kombinasyonu* $\lambda_1, \ldots, \lambda_m$ skalerler ve $\lambda_1 + \cdots + \lambda_m = 1$ olmak üzere $\lambda_1 x_1 + \cdots + \lambda_m x_m$ şeklinde tanımlanır. Özel olarak herhangi iki elemanının afin kombinasyonunu içeren kümeye *afin küme* denir, yani

$$\forall\, x_1, x_2 \in A, \quad \lambda_1, \lambda_2 \in \mathbb{R}, \quad \lambda_1 + \lambda_2 = 1 \text{ için } \lambda_1 x_1 + \lambda_2 x_2 \in A$$

gerçekleniyorsa A kümesi bir afin kümedir. Boş küme afin küme olarak kabul edilir. Tek elemanlı kümeler, doğrular, düzlemler afin kümelerdir.

$\lambda_1, \lambda_2 \in \mathbb{R}$ olmak üzere $\mathbb{R}^n$ uzayındaki herhangi iki x_1, x_2 elemanın *lineer kombinasyonu* $\lambda_1 x_1 + \lambda_2 x_2$ şeklinde tanımlanır. Herhangi iki elemanının lineer kombinasyonunu içeren kümeye *lineer alt uzay* denir, yani

$$\forall\, x_1, x_2 \in L, \quad \lambda_1, \lambda_2 \in \mathbb{R} \quad \text{için} \quad \lambda_1 x_1 + \lambda_2 x_2 \in L$$

gerçekleniyorsa $L \subseteq \mathbb{R}^n$ bir lineer alt uzaydır. Afin kümeler ile lineer alt uzaylar arasındaki ilişki aşağıdaki lemma ile ifade edilir.

Lemma 2.1. $\mathbb{R}^n$ uzayının her lineer alt uzayı, orijini içeren bir afin kümedir. Böylece her lineer alt uzay bir afin kümedir.

İspat: Her lineer alt uzay orijini içerir. Alt uzayların afin küme olduğunu gösterelim. A, $\mathbb{R}^n$ uzayının bir alt uzay olsun ve $a, b \in A$ için $(1-\lambda)a + \lambda b \in A$ olduğundan A bir afin kümedir. Orijini içeren bir A afin kümesinin alt uzay olduğunu göstermek için $a,\ 0 \in A$ olsun. *i)* A afin küme olduğundan $\lambda a + (1-\lambda)0 \in A$ dır. Dolayısıyla $\lambda a \in A$ olduğundan skalerle çarpma işlemi altında kapalıdır. *ii)* $a, b \in A$ için $\frac{1}{2}a + \frac{1}{2}b \in A$ olup *i)* den $2(\frac{1}{2}a + \frac{1}{2}b) \in A$ olur. Böylece $a + b \in A$ olduğundan A afin kümesi toplama işlemi altında kapalıdır. *i)* ve *ii)* den orijini içeren bir afin küme $\mathbb{R}^n$ uzayının alt uzayıdır. ∎

A ve B iki afin küme olsun, eğer bir $x_0 \in \mathbb{R}^n$ için $A = B + x_0$, yani A kümesi B kümesinin bir ötelemesi ise A afin kümesi B afin kümesine *paraleldir* denir.

Lemma 2.2. [13] $\emptyset \neq A \subseteq \mathbb{R}^n$ alt kümesinin afin olabilmesi için gerek ve yeter koşul tek türlü belirli bir lineer L alt uzayına paralel olmasıdır, yani bir $x_0 \in \mathbb{R}^n$ için $A = L + x_0$ gerçekleşmesidir.

Bir $A \subseteq \mathbb{R}^n$ afin kümesinin boyutu, onun paralel olduğu tek türlü belirli olan lineer alt uzayın boyutu olarak tanımlanır ve $dim(A)$ ile gösterilir. Boş kümenin boyutu -1 olarak kabul edilir. $\mathbb{R}^n$ uzayında $n-1$ boyutlu afin kümeye *hiperdüzlem* denir ve genellikle H ile gösterilir. Bir $L \subseteq \mathbb{R}^n$ alt uzayının *ortogonal tümleyeni* $L^{\perp} = \{x : \langle x, y\rangle = 0\ , \forall\, y \in L\}$ şeklinde tanımlanır ve $L^{\perp}$ kümesi de bir alt uzaydır. Bilindiği gibi $dim(L) + dim(L^{\perp}) = n$ bağıntısı geçerli olup $\mathbb{R}^n$ uzayının $n-1$ boyutlu alt uzayları, 1 boyutlu alt uzayların ortogonal tümleyenidir. $dim(L) = 1$

olan alt uzaylar sıfırdan farklı tek bir b vektöründen oluşan bir tabana sahiptir. $dim(L^{\perp}) = n-1$ olan alt uzayı $\{x:\ x\perp b\}$ kümesidir. Lemma 2.2'den hiperdüzlemler ise bu alt uzayın $a \in \mathbb{R}^n$ olmak üzere a kadar ötelemesidir. Yani

$$\begin{aligned} H = \{x \in \mathbb{R}^n :\ x\perp b\} + a &= \{x + a :\ \langle x, b\rangle = 0\} = \{x :\ \langle x - a\ ,\ b\rangle = 0\} \\ &= \{x : \langle x, b\rangle = \beta\} \end{aligned}$$

gerçeklenir, burada $\beta \in \mathbb{R}$ ve $0 \neq b$ dır. $b \neq 0$ vektörüne de H hiperdüzlemin *normal vektörü* denir. H kümesi $\mathbb{R}^n$ uzayında $(n-1)$ boyutlu bir hiperdüzlem tanımlar.

Lemma 2.3. [13] $A, m\times n$ reel matris ve $b \in R^m$ olsun bu durumda $Ax = b$ denklem sisteminin çözümü olan $\{x \in \mathbb{R}^n : Ax = b\}$ kümesi afin kümedir. Ayrıca herhangi bir afin küme bu şekilde ifade edilebilir.

Kolayca herhangi sayıda afin kümenin kesişimi yine bir afin küme olduğu görülebilir. Bir $A \subseteq \mathbb{R}^n$ kümesini kapsayan tüm afin kümelerin arakesitine A kümesinin *afin zarfı* denir ve aff$A$ ile gösterilir. Ayrıca aff$A$ kümesi A kümesini içeren en küçük afin kümedir. $\mathbb{R}^n$, A kümesini kapsayan afin küme olduğundan aff$A \neq \emptyset$ dır.

A, $m \times n$ reel matris ve $b \in R^m$ olmak üzere Lemma 2.3 nedeniyle

$$\big\{x \in \mathbb{R}^n :\ Ax = b\big\} = \big\{x :\ \langle x, b_i\rangle = \beta_i\ ,\ i = 1, \ldots, m\big\} = \bigcap_{i=1}^{m} H_i$$

denklemi gerçeklenir. Burada b_i, A matrisinin i. satırı, β_i, b vektörünün i. bileşeni ve $H_i = \{x :\ \langle x, b_i\rangle = \beta_i\}$ dir. Böylece bir afin kümenin hiperdüzlemlerin kesişimi olarak ifade edilebildiği anlaşılır.

Tanım 2.1. $0 \neq a \in \mathbb{R}^n$ ve $b \in \mathbb{R}$ olmak üzere $H = \{x : \langle x, a\rangle = b\}$ hiperdüzlemi verilsin. Bu durumda,

$$H^+ = \big\{x :\ \langle x, a\rangle \geqslant b\big\} \quad , \quad H^- = \big\{x :\ \langle x, a\rangle \leqslant b\big\}$$

kümelerine H hiperdüzlemine karşılık gelen sırasıyla *pozitif ve negatif kapalı yarı-uzaylar*,

$$int(H^+) = \big\{x :\ \langle x, a\rangle > b\big\}, \quad int(H^-) = \big\{x :\ \langle x, a\rangle < b\big\}$$

kümelerine ise H hiperdüzlemine karşılık gelen *açık yarı-uzaylar* denir.

2.1. KONVEKS KÜMELER VE KONİLER

Bu kısımda konveks kümelerin temel tanım ve teoremleri gösterilip, konveks analizde önemli bir yere sahip iki konveks kümenin veya bir nokta ile bir konveks kümenin bir hiperdüzlem yardımıyla ayrılmasına ilişkin ayrılma teoremleri ile konveks koniler hakkında gerekli bilgi verilecektir.

2.1.1. Konveks Kümeler

$x_1, x_2 \in \mathbb{R}^n$ noktaları arasındaki doğru parçası

$$\{\lambda_1 x_1 + \lambda_2 x_2 : \ \lambda_1, \lambda_2 \geqslant 0 \quad \lambda_1 + \lambda_2 = 1\}$$

kümesi ile ifade edilir.

Tanım 2.2. Herhangi iki noktası arasındaki doğru parçasını kapsayan, yani

$$x_1, x_2 \in A,\ \lambda_1, \lambda_2 \geqslant 0,\ \lambda_1 + \lambda_2 = 1 \text{ için } \lambda_1 x_1 + \lambda_2 x_2 \in A$$

gerçekleyen $A \subseteq \mathbb{R}^n$ kümesine *konveks küme* denir. Bir A kümesinin konveksliği için eşdeğer bir tanım olarak her $\lambda \in [0,1]$ için $(1-\lambda)A + \lambda A \subseteq A$ kapsamasının sağlanması verilebilir.

Tüm afin kümeler ve konvekslik koşulunu otomatik olarak gerçeklediklerinden tüm alt uzaylar konveks kümelerdir. $\emptyset$ konveks küme olarak kabul edilir.

Lemma 2.4. [14] *i)* $A_1, \ldots, A_m \subseteq \mathbb{R}^n$ konveks kümeler $\lambda_1, \ldots, \lambda_m$ skalerler olmak üzere $\lambda_1 A_1 + \cdots + \lambda_m A_m$ kümesi de konvekstir.

ii) Herhangi sayıda konveks kümenin kesişimi yine konveks kümedir.

iii) A konveks kümesinin içi $int(A)$ ve kapanışı $\overline{A}$ konveks kümelerdir.

iv) Her $x, y \in \mathbb{R}^n$ ve her $\alpha \in \mathbb{R}$ için $f(\alpha x + (1-\alpha)y) = \alpha f(x) + (1-\alpha) f(y)$ bağıntısını sağlayan $f : \mathbb{R}^n \to \mathbb{R}^m$ afin dönüşüm olmak üzere $A \subseteq \mathbb{R}^n$ konveks küme ise $f(A)$ görüntü kümesi de konveks kümedir. Ayrıca $B \subseteq \mathbb{R}^m$ konveks küme ise $f^{-1}(B)$ ters görüntü kümesi de konveks bir kümedir.

$\lambda_1, \ldots, \lambda_m \geqslant 0,\ \lambda_1 + \cdots + \lambda_m = 1$ olmak üzere $x = \lambda_1 x_1 + \cdots + \lambda_m x_m$ toplamına $x_1, \ldots, x_m \in \mathbb{R}^n$ noktalarının *konveks kombinasyonu* denir. Bu tanım yardımıyla konveks kümelerin önemli bir özelliği aşağıdaki lemma ile ifade edilir.

Lemma 2.5. A konveks küme ve $x_1, \ldots, x_m \in A$ ise bu noktaların her konveks kombinasyonu A kümesine aittir.

İspat: Tümevarımla $x_1, \ldots, x_m \in A$ noktaların konveks kombinasyonunun kümeye ait olduğunu göstermek için A konveks küme tanımından $m = 2$ için doğru olup $m = k$ için iddianın doğruluğunu kabul edelim. $x = \lambda_1 x_1 + \cdots + \lambda_{k+1} x_{k+1}$ olsun. $\lambda_1 + \cdots + \lambda_{k+1} = 1$ için $\lambda_1, \ldots, \lambda_{k+1} \geqslant 0$ değerlerinden en azından biri birden küçük olacağı için $\lambda_{k+1} < 1$ olsun. $\lambda = \lambda_1 + \cdots + \lambda_k = 1 - \lambda_{k+1} > 0$ olmak üzere $y = \frac{\lambda_1}{\lambda} x_1 + \cdots + \frac{\lambda_k}{\lambda} x_k$ ise iddia $m = k$ için doğru olduğundan $y \in A$ bulunur. $x = \lambda y + \lambda_{k+1} x_{k+1}$ olacağından $x \in A$ olur. Dolayısıyla iddia $m = k + 1$ için de doğru olduğundan istenen elde edilir. ∎

Herhangi bir $A \subseteq R^n$ kümesini kapsayan tüm konveks kümelerin arakesitine A kümesinin *konveks zarfı* denir ve $conv(A)$ ile gösterilir. Lemma 2.4 nedeniyle $conv(A)$ kümesi konvekstir. İki farklı noktanın konveks zarfı bu iki noktayı birleştiren doğru parçasıdır. $\mathbb{R}^2$ de bir çember üzerine simetrik olarak yerleştirilmiş m tane noktanın konveks zarfı $m \geqslant 3$ için m-kenarlı düzgün çokgendir. Konveks zarf ile ilgili bir diğer örnek aşağıda verilmektedir.

Örnek 2.1. $A = \{x : ||x|| = 1\}$ kümesinin konveks zarfı $B[0,1] = \{x : ||x|| \leqslant 1\}$ kapalı birim yuvardır.

Çözüm: $B[0,1]$ kapalı birim yuvarı konveks ve $A \subseteq B[0,1]$ kapsaması var olduğundan $conv(A) \subseteq B[0,1]$ dir. Diğer taraftan $x \in B[0,1]$ olsun. Eğer $x = 0$ ise $y \in A$ için $x = \frac{1}{2}y + \frac{1}{2}(-y)$ şeklinde yazılırsa $y, -y \in conv(A)$ olacağından $x \in conv(A)$ dır. Eğer $x \neq 0$ ise, $x = \left(\frac{1+||x||}{2}\right)\frac{x}{||x||} + \left(\frac{1-||x||}{2}\right)\frac{-x}{||x||}$ ve $\frac{x}{||x||}, \frac{-x}{||x||} \in conv(A)$ olacağından $x \in conv(A)$ bulunur. Dolayısıyla $conv(A) = B[0,1]$ olup istenen elde edilir. ∎

Lemma 2.6. $A \subseteq \mathbb{R}^n$ kümesinin konveks zarfı olan $conv(A)$ kümesi A kümesine ait olan noktaların tüm konveks kombinasyonlarını içerir.

İspat: B kümesi ile A kümesinin noktalarının tüm konveks kombinasyonları kümesini gösterelim. $conv(A) = B$ olduğunu göstermek lemmanın ispatı için yeterli olacaktır. $\lambda_1 + \cdots + \lambda_m = 1$ ve $\lambda_1, \ldots, \lambda_m \geqslant 0$ olmak üzere $a_1, \ldots, a_m \in A$ için $x = \lambda_1 a_1 + \cdots + \lambda a_m \in B$ olsun. $a_1, \ldots, a_m \in conv(A)$ ve $conv(A)$ kümesi konveks olduğundan Lemma 2.5 nedeniyle $x \in conv(A)$ olup $B \subseteq conv(A)$ elde edilir. B kümesinin tanımından dolayı konveksliği kolayca görülür. Ayrıca $A \subseteq B$ olduğundan $conv(A) \subseteq B$ olup $conv(A) = B$ bulunur. ∎

Sonuç 2.1. $x_1, \ldots, x_m \in \mathbb{R}^n$ olsun. O halde

$$conv\{x_1, \ldots, x_m\} = \{\lambda_1 x_1 + \cdots + \lambda_m x_m \ : \ \lambda_1, \ldots, \lambda_m \geqslant 0, \ \lambda_1 + \cdots + \lambda_m = 1\}$$

olarak yazılır.

Lemma 2.6 'dan herhangi bir kümenin konveks zarfının her bir noktasının bu kümenin noktalarının konveks kombinasyonu olarak yazılabildiği anlaşılmaktadır. n boyutlu bir kümenin konveks zarfının her bir noktası, bu kümenin en fazla $n+1$ noktasının konveks kombinasyonu olarak ifade edilebildiği aşağıdaki ünlü Caratheodory Teoremi'nde verilmektedir.

Teorem 2.7. (Caratheodory Teoremi) r-boyutlu A kümesi verildiğinde her $x \in conv(A)$ için öyle $x_1, \ldots, x_m \in A$ noktaları vardır ki her $i = 1, \ldots, m$ için $\lambda_i \geqslant 0$, $\sum\limits_{i=1}^{m} \lambda_i = 1$ ve $m \leqslant r+1$ olmak üzere $x = \lambda_1 x_1 + \cdots + \lambda_m x_m$ dir. Yani $conv(A)$ konveks kümesinin her bir elemanı r-boyutlu A kümesinin en fazla $r+1$ noktasının konveks kombinasyonu olarak ifade edilir.

İspat: r-boyutlu A kümesi için $x_1, \ldots, x_m \in A$ ise

$$conv(A) = \{x = \lambda_1 x_1 + \cdots + \lambda_m x_m \ : \ \lambda_1, \ldots, \lambda_m \geqslant 0, \ \lambda_1 + \cdots + \lambda_m = 1\}$$

olur. $x_1, \ldots, x_m$ noktaları birbirinden farklı, her $i = 1, \ldots, m$ için $\lambda_i > 0$ ve x noktası A kümesinin m'den daha az noktasının konveks kombinasyonu şeklinde yazılamadığını kabul edelim. $r+1 < m$ olsun. $\{x_1, \ldots, x_m\}$ kümesi lineer bağımlıdır. $t > 0$, $t = \min\{\frac{\lambda_i}{\mu_i} \ : \ \lambda_i > 0, \mu_i > 0\}$ şeklinde tanımlanırsa $\lambda_i - t\mu_i$ değerleri negatif değildir ve $i = 1, \ldots, m$ için en azından biri sıfıra eşit olacağından

$$(\lambda_1 - t\mu_1)x_1 + (\lambda_2 - t\mu_2)x_2 + \cdots + (\lambda_m - t\mu_m)x_m$$

bağıntısında $\{x_1, \ldots, x_m\}$ lineer bağımlı ise μ_i hepsi birden sıfır olmayan skalerler olmak üzere $\mu_1 x_1 + \cdots + \mu_m x_m = 0$ ve $\mu_1 + \cdots + \mu_m = 0$ olması gerçeğinden

$$\lambda_1 x_1 + \cdots + \lambda_m x_m - t(\mu_1 x_1 + \cdots + \mu_m x_m) = x - t.0 = x$$

elde edilir. Her $i = 1, \ldots, m$ için $\lambda_i - t\mu_i$ değerlerinin en azından biri sıfır olduğundan x noktası A kümesinin m'den daha az noktasının konveks kombinasyonu şeklinde yazılır. ∎

Herhangi bir $A \subseteq \mathbb{R}^n$ kümesini içeren tüm kapalı konveks kümelerin kesişimine A kümesinin *kapalı konveks zarfı* denir ve $\overline{conv}(A)$ ile gösterilir.

Lemma 2.8. [14] $A \subseteq \mathbb{R}^n$ kümesinin kapalı konveks zarfı, A kümesinin konveks zarfının kapanışına eşittir. Yani $\overline{conv}(A) = \overline{conv(A)}$ dır.

A kümesinin tüm iç noktaları kümesi olan $int(A)$, A kümesinin hangi uzayın alt kümesi olarak göz önüne alındığına göre değişir. Örneğin; $\mathbb{R}^2$ uzayında doğru parçasının iç noktaları kümesi boş kümedir. Fakat aynı doğru parçası $\mathbb{R}$'nin alt kümesi olarak düşünülürse uç noktalar hariç tüm noktaları iç noktadır. Bu sebeple konveks kümelerde kümenin içi kavramı yerine daha uygun olan izafi iç kavramı verilmektedir. $A \subseteq \mathbb{R}^n$ konveks küme olmak üzere, izafi iç noktalardan oluşan kümeye A kümesinin *izafi içi* denir ve $ri(A)$ ile gösterilir. Açıkça,

$$ri(A) = \{x \in \text{aff}A : \exists\, \delta > 0 \text{ için } B(x, \delta) \cap \text{aff}A \subseteq A\}$$

olarak tanımlanır. n-boyutlu konveks A kümesi için $\text{aff}A = \mathbb{R}^n$ olduğundan $ri(A) = int(A)$ gerçeklenir. Genelde $A_1 \subseteq A_2$ ise $\overline{A}_1 \subseteq \overline{A}_2$ ve $int(A_1) \subseteq int(A_2)$ kapsamaları doğrudur. Ancak $ri(A_1) \subseteq ri(A_2)$ kapsaması gerçeklenmeyebilir. Örneğin, $\mathbb{R}^3$ uzayında bir küp A_1 ve küpün bir yüzeyi A_2 olsun. $A_2 \subseteq A_1$ kapsaması geçerli iken $ri(A_1) \neq \emptyset$ ve $ri(A_2) \neq \emptyset$ olup $ri(A_1) \cap ri(A_2) = \emptyset$ dır. Dolayısıyla $ri(A_2) \nsubseteq ri(A_1)$ olur.

Lemma 2.9. [14] $A \subseteq \mathbb{R}^n$ boş kümeden farklı ve konveks küme ise $ri(A) \neq \emptyset$ dır.

Lemma 2.10. [14] $A \subseteq \mathbb{R}^n$ konveks bir küme olsun. Bu durumda $x \in ri(A)$ ve $y \in A$ ise her $\lambda \in (0, 1]$ için $\lambda x + (1 - \lambda)y \in ri(A)$ dır.

Lemma 2.11. $A \subseteq \mathbb{R}^n$ konveks bir küme olsun. Bu durumda $x \in ri(A)$ ve $y \in \overline{A}$ ise her $\lambda \in (0, 1]$ için $\lambda x + (1 - \lambda)y \in ri(A)$ gerçeklenir.

İspat: $x \in ri(A)$ olduğundan öyle $\delta > 0$ vardır öyle ki $B(x, \delta) \cap \text{aff } A \subseteq A$ olur. $c = \lambda x + (1 - \lambda)y$ olsun ve $\lambda \in (0, 1]$ için $y \in \overline{A}$ olduğundan öyle $d \in A$ vardır ki $(1 - \lambda)||d - y|| < \lambda\delta$ sağlanır.

$$z = x + (\frac{1}{\lambda} - 1)(y - d) = \frac{1}{\lambda} \cdot c + (1 - \frac{1}{\lambda}) \cdot d$$

ise $z \in \text{aff } A$ ve $||x + \frac{(1-\lambda)}{\lambda}(y - d) - x|| < \delta$ olduğundan $z \in ri(A)$ dır. $c = \lambda z + (1 - \lambda)d$ ve Lemma 2.10 'dan $c \in ri(A)$ olup ispat tamamlanır. ■

Lemma 2.12. [13] $A_1, A_2 \subseteq \mathbb{R}^n$ konveks kümeler olmak üzere $ri(A_1) \cap ri(A_2) \neq \emptyset$ olsun. Bu durumda

$$ri(A_1 \cap A_2) = ri(A_1) \cap ri(A_2) \qquad \text{ve} \qquad \overline{A_1 \cap A_2} = \overline{A_1} \cap \overline{A_2}$$

eşitlikleri gerçeklenir.

Ayrıca $\alpha_1, \alpha_2 \in \mathbb{R}$ olmak üzere $ri(\alpha_1 A_1 + \alpha_2 A_2) = \alpha_1 ri(A_1) + \alpha_2 ri(A_2)$ dir. Burada $\alpha_1 = 1$ ve $\alpha_2 = -1$ alınırsa

$$0 \in ri(A_1 - A_2) \Longleftrightarrow ri(A_1) \cap ri(A_2) \neq \emptyset$$

elde edilir.

Tanım 2.3. $A \subseteq \mathbb{R}^n$ boştan farklı bir küme, $H \subseteq \mathbb{R}^n$ bir hiperdüzlem ve $x \in A$ olsun. Eğer H hiperdüzlemi x noktasını içeriyor ve H hiperdüzlemine karşılık gelen kapalı yarı-uzaylardan biri A kümesini tamamen kapsıyorsa bu durumda H hiperdüzlemi x noktasında A kümesini *destekler* denir.

Kapalı konveks küme bu kümeyi içeren kapalı yarı-uzayların kesişimi olarak yazılabilir. Konveks kümelerin bu özelliği aşağıda verilen ayırma teoremlerinden kaynaklanmaktadır. Esas olarak ayırma teoremleri ortak noktası olmayan iki konveks kümenin bir hiperdüzlem ile ayrılabilir olması ilkesine dayanır.

Teorem 2.13. $A \subseteq \mathbb{R}^n$ konveks küme ve $x_0 \notin \overline{A}$ olsun. O halde öyle bir $x^* \neq 0$ vektörü ve $\varepsilon > 0$ sayısı vardır ki her $x \in \overline{A}$ için

$$\langle x, x^* \rangle \leqslant \langle x_0, x^* \rangle - \varepsilon$$

gerçeklenir.

İspat: A konveks küme $x_0 \notin \overline{A}$ olsun. $y = \inf\left\{d(x_0, x) : x \in \overline{A}\right\}$ olarak tanımlayalım. Her $x \in \overline{A}$ ve $y \in \overline{A}$ için $||x - x_0|| \geqslant ||y - x_0||$ dır. $\overline{A}$ kümesi de konveks olduğundan $0 \leqslant \lambda \leqslant 1$ için $\lambda x + (1-\lambda)y \in \overline{A}$ olur.

$$\begin{aligned} ||y - x_0||^2 &\leqslant ||\lambda x + (1-\lambda)y - x_0||^2 = \langle y + \lambda(x-y) - x_0,\ y + \lambda(x-y) - x_0 \rangle \\ &= ||y - x_0||^2 + 2\lambda\langle x - y,\ y - x_0\rangle + \lambda^2||x - y||^2 \end{aligned}$$

eşitsizliği sağlanacağı için $2\langle x - y,\ y - x_0\rangle + \lambda||x - y||^2 \geqslant 0$ olduğu görülür. Özel olarak $\lambda = 0$ için $\langle x - y,\ y - x_0\rangle \geqslant 0$ olur. $x^* = x_0 - y$ vektörünü ve $\varepsilon = ||x^*||^2$ sayısı seçildiğinde $x_0 \notin \overline{A}$ olduğu için $y \neq x_0$ olur ve $x^* \neq 0,\ \ \varepsilon > 0$ için

$$\langle x, x^* \rangle \leqslant \langle x_0, x^* \rangle - \langle x^*, x^* \rangle = \langle x_0, x^* \rangle - \varepsilon$$

olacağından ispat tamamlanır. ∎

Teorem 2.14. (Destek Hiperdüzlem Teoremi) A konveks küme ve $x_0 \notin A$ olsun. O halde öyle bir $x^* \neq 0$ vektörü vardır ki her $x \in A$ için

$$\langle x, x^* \rangle \leqslant \langle x_0, x^* \rangle$$

gerçeklenir.

İspat: A konveks küme ve $x_0 \notin A$ olsun. Eğer $x_0 \notin \overline{A}$ ise Teorem 2.13 nedeniyle istenen eşitsizlik gerçeklenir. $x_0 \in \overline{A}$ olsun. Böylece $x_n \notin \overline{A}$ ve $x_n \to x_0$ gerçekleyen bir (x_n) dizisi vardır. Öyle bir $x_n^* \neq 0$ vektörü ve $\varepsilon_n > 0$ sayısı vardır ki her $x \in A$ için $\langle x, x_n^* \rangle \leqslant \langle x_n, x_n^* \rangle - \varepsilon_n < \langle x_n, x_n^* \rangle$ olur. Her $x \in A$, her $n \in \mathbb{N}$ için her iki tarafı $||x_n^*||$ normuna bölersek

$$\langle x, \frac{x_n^*}{||x_n^*||} \rangle < \langle x_n, \frac{x_n^*}{||x_n^*||} \rangle$$

elde edilir. $\frac{x_n^*}{||x_n^*||} \to x^*$ seçilip limit alınırsa her $x \in A$ için $\langle x, x^* \rangle \leqslant \langle x_0, x^* \rangle$ eşitsizliği ispatlanır. ∎

Teorem 2.15. $A_1, A_2 \subseteq \mathbb{R}^n$ ve $A_1 \cap A_2 = \emptyset$ olan konveks kümeler olsun. O halde öyle bir $x^* \neq 0$ vektörü vardır ki her $x_1 \in A_1$ ve $x_2 \in A_2$ için

$$\langle x^*, x_1 \rangle \leqslant \langle x^*, x_2 \rangle$$

eşitsizliği gerçeklenir.

İspat: $A = A_1 - A_2$ olsun. $A_1 \cap A_2 = \emptyset$ olduğundan $0 \notin A$ dır. Teorem 2.14 nedeniyle öyle bir $x^* \neq 0$ vektörü vardır ki her $x \in A$ için $\langle x, x^* \rangle \leqslant \langle 0, x^* \rangle = 0$ olur. $x_1 \in A_1$, $x_2 \in A_2$ olmak üzere $x = x_1 - x_2 \in A$ için $\langle x^*, x_1 - x_2 \rangle \leqslant 0$ olup ispat tamamlanır. ∎

Teorem 2.16. A_1 kompakt konveks küme ve A_2 kapalı konveks küme olsun. Eğer $A_1 \cap A_2 = \emptyset$ ise öyle bir $x^* \neq 0$ vektörü ve $\varepsilon > 0$ sayısı vardır ki her $x_1 \in A_1$ ve her $x_2 \in A_2$ için

$$\langle x^*, x_1 \rangle \leqslant \langle x^*, x_2 \rangle - \varepsilon$$

eşitsizliği gerçeklenir.

İspat: $A = A_1 - A_2$ olsun. $A_1 \cap A_2 = \emptyset$ olduğundan $0 \notin A$ dır. Lemma 2.4 nedeniyle A kümesi konvekstir. Her $n \in \mathbb{N}$ için $x_{1,n} \in A_1$, $x_{2,n} \in A_2$ olsun ve $x_{1,n} - x_{2,n} = x_n \to x$ gerçeklensin. $x \in A$ olduğunu gösterirsek ispat tamamlanmış olur. A_1 kompakt küme olduğundan $(x_{1,n})$ dizisinin yakınsak bir alt dizisi vardır. Genelliği bozmaksızın yakınsak alt dizi olarak $(x_{1,n})$ dizisi alınırsa $x_{1,n} \to x_1 \in A_1$ kabul edilebilir. (x_n) dizisi yakınsak ve A_2 kapalı küme olduğundan $x_{2,n} \to x_2 \in A_2$ gerçeklenir. Böylece $x = x_1 - x_2 \in A$ elde edilir. Teorem 2.13 'den öyle bir $x^* \neq 0$ vektörü ve $\varepsilon > 0$ sayısı vardır ki her $x \in A$ için $\langle x, x^* \rangle \leqslant -\varepsilon$ olup

$$\langle x^*, x_1 \rangle \leqslant \langle x^*, x_2 \rangle - \varepsilon$$

eşitsizliği sağlanır. ∎

Tanım 2.4. Bir A konveks kümesinin bir $x^* \in \mathbb{R}^n$ noktasındaki destek fonksiyonu aşağıdaki gibi tanımlanır:

$$W_A(x^*) = \sup_x \{\langle x, x^* \rangle : \ x \in A\} \tag{2.1}$$

Teorem 2.17. A kapalı konveks bir küme olmak üzere $x \in A$ olabilmesi için gerek ve yeter koşul her $x^* \in \mathbb{R}^n$ için

$$\langle x, x^* \rangle \leqslant W_A(x^*) \tag{2.2}$$

eşitsizliğinin gerçeklenmesidir.

İspat: $x \in A$ ise Tanım 2.4 nedeniyle (2.2) eşitsizliği geçerlidir. Tersine bir x_0 noktası (2.2) bağıntısını sağlasın ama $x_0 \in A$ gerçekleşmesin yani $x_0 \notin A$ olsun. Bu durumda Teorem 2.13 nedeniyle öyle bir x^* noktası ve $\varepsilon_0 > 0$ sayısı vardır ki her $x \in A$ için $\langle x, x^* \rangle \leqslant \langle x_0, x^* \rangle - \varepsilon_0$ eşitsizliği gerçeklenir. Bu eşitsizlikte $x \in A$ noktaları üzerinden supremum alınırsa $W_A(x^*) \leqslant \langle x_0, x^* \rangle - \varepsilon_0 < \langle x_0, x^* \rangle$ eşitsizliği elde edilir. Bu ise x_0 noktasının (2.2) bağıntısını sağladığı varsayımı ile çelişir. ∎

Konveks kümelerin özel bir sınıfı da polihedral (=çokyüzlü) kümelerdir. Bir $A \subseteq \mathbb{R}^n$ kümesi, eğer sonlu sayıdaki kapalı yarı-uzayın kesişimi olarak ifade edilebiliyorsa ya da eşdeğer olarak $i = 1, \ldots, m$ için $b_i \in \mathbb{R}^n$ ve $\alpha_i \in \mathbb{R}$ olmak üzere $\langle x, b_i \rangle \leqslant \alpha_i$ biçimdeki sonlu sayıda eşitsizlikler sisteminin çözümü ise A kümesine bir *polihedral küme* denir. Böylece polihedral kümeler kapalı ve konvekstir. $\emptyset$ ve $\mathbb{R}^n$ polihedral kümedir. Ayrıca $\mathbb{R}^n$ uzayında herhangi bir hiperdüzlem iki kapalı yarı-uzayın kesişimi şeklinde yazılabildiğinden polihedral kümedir. Sınırlı bir polihedral kümeye ya da eşdeğer olarak sonlu noktalı bir kümenin konveks zarfına *politop* denir.

2.1.2. Konveks Koniler

Tanım 2.5. $K \subseteq \mathbb{R}^n$, her $x \in K$ ve her $\lambda > 0$ için $\lambda x \in K$ ise K kümesine *koni* denir. Eğer bu K kümesi aynı zamanda konveks ise bu koniye *konveks koni* denir.

Tanım 2.6. K bir koni olmak üzere her $x \in K$ için $\langle x, x^* \rangle \geqslant 0$ gerçekleyen $x^* \in \mathbb{R}^n$ vektörlerinin kümesine K konisinin *dual konisi* denir ve K^* ile gösterilir. Açıkça dual koni

$$K^* = \{x^* \in \mathbb{R}^n : \ \langle x, x^* \rangle \geqslant 0 \ , \quad \forall\, x \in K\}$$

şeklinde tanımlanır.

Dual koniler daima kapalıdır. Gerçekten her $n \in \mathbb{N}$ ve $x_n^* \in K^*$ için $x_n^* \to x_0^*$ olsun. Her $x \in K$ ve her $n \in \mathbb{N}$ için $\langle x, x_n^* \rangle \geqslant 0$ dır. Bu eşitsizlikte limit alınırsa her $x \in K$ için $\langle x, x_0^* \rangle \geqslant 0$ olur. Bu ise $x_0^* \in K^*$ olduğunu gösterir. Böylece K^* dual konisi tüm limit noktasını içerdiği için kapalıdır.

Lemma 2.18. Bir K konisi ve onun $\overline{K}$ kapanış konisi aynı dual koniye sahiptir.

İspat: Eğer $x \in (\overline{K})^*$ ise her $x \in \overline{K}$ için $\langle x, x^* \rangle \geqslant 0$ ve $K \subseteq \overline{K}$ nedeniyle her $x \in K$ için $\langle x, x^* \rangle \geqslant 0$ olup $x^* \in K^*$ dır. Tersine $x^* \in K^*$ olsun. Eğer $x \in \overline{K}$ ise öyle bir $(x_n) \in K$ dizisi vardır ki her $n \in \mathbb{N}$ için $(x_n) \to x$ olur. $x^* \in K^*$ olduğundan, her $x_n \in K$ için $\langle x_n, x^* \rangle \geqslant 0$ sağlanır. Bu eşitsizlikte limit alınırsa her $x \in \overline{K}$ için $\langle x, x^* \rangle \geqslant 0$ olduğundan $x^* \in (\overline{K})^*$ olup istenen elde edilir. ∎

Lemma 2.19. Bir K kapalı konisi verildiğinde her $x^* \in K^*$ için $\langle x, x^* \rangle \geqslant 0$ gerçekleniyorsa $x \in K$ olur.

İspat: Kabul edelim ki her $x^* \in K^*$ için $\langle x_0, x^* \rangle \geqslant 0$ gerçeklensin ama $x_0 \notin K$ olsun. Teorem 2.16 nedeniyle öyle bir $x_0^* \neq 0$ vektörü ve $\varepsilon > 0$ sayısı vardır ki her $x \in K$ için

$$\langle x_0, x_0^* \rangle \leqslant \langle x, x_0^* \rangle - \varepsilon \tag{2.3}$$

eşitsizliği gerçeklenir. K koni olduğundan yeterince küçük $\lambda > 0$ sayısı için $\lambda x \in K$ olur. Böylece sıfıra yeterince yakın noktalar koniye dahildir. Özellikle kapalı koniler daima sıfır noktasını içerirler. Bu durumda $\langle x, x_0^* \rangle$ çarpımının alt sınırı sıfırdan daha küçük olamaz. (2.3) bağıntısı nedeniyle $\langle x, x_0^* \rangle$ çarpımı alttan sınırlıdır. Şimdi $\langle x, x_0^* \rangle \geqslant 0$ olduğunu gösterelim. Eğer $\langle x_1, x_0^* \rangle < 0$ olacak şekilde bir $x_1 \in K$ varsa bu nokta yardımıyla $x = \lambda x_1$ noktası tanımlansın. Bu durumda $\lambda \to \infty$ için $\langle x, x_0^* \rangle \to -\infty$ olur bu ise $\langle x, x_0^* \rangle$ çarpımının alttan sınırlı olması ile çelişir. Dolayısıyla $\langle x, x_0^* \rangle \geqslant 0$ dır. Yani $x_0^* \in K^*$ dır. (2.3) denklemi $x = 0$ olduğunda $\langle x_0, x_0^* \rangle \leqslant -\varepsilon$ olup bu ise kabulümüz ile çelişir. Dolayısıyla $x_0 \in K$ olur. ∎

K^* dual konisinin de dualinden bahsetmek mümkündür. Dual koniler daima kapalı olduğundan, kapalı olmayan bir koninin dual konisinin duali kendisine eşit olamaz. Bu nedenle dualinin duali kendisine eşit olabilecek koniler ancak kapalı konilerdir. Bu gerçek aşağıdaki lemma ile verilmektedir.

Lemma 2.20. K konisi kapalı ise $K^{**} = K$ eşitliği geçerlidir.

İspat: Tanım 2.6'e göre $K^{**} = \{x \in \mathbb{R}^n : \ \langle x, x^* \rangle \geqslant 0 \ , \quad \forall \ x^* \in K^*\}$ olur. Eğer $x \in K$ ise her $x^* \in K^*$ için $\langle x, x^* \rangle \geqslant 0$ olduğundan $x \in K^{**}$ dır. Tersine $x \in K^{**}$ olsun. Böylece her $x^* \in K^*$ için $\langle x, x^* \rangle \geqslant 0$ dır. Lemma 2.19 nedeniyle $x \in K$ olup $K^{**} \subseteq K$ kapsaması sağlandığından K kapalı konisi için $K^{**} = K$ eşitliği geçerlidir.■

Uyarı 2.1. $K^{**} = \overline{K}$ eşitliği genel halde doğrudur. Çünkü $K^* = (\overline{K})^*$ eşitliğinde her iki tarafın duali alınırsa $K^{**} = (\overline{K})^{**} = \overline{K}$ bulunur.

Lemma 2.21. K_1 ve K_2 herhangi iki konveks koni olsun. $K_1 + K_2$ toplamı da konveks konidir ve $(K_1 + K_2)^* = K_1^* \cap K_2^*$ dir.

İspat: Lemma 2.4 nedeniyle $K_1 + K_2$ konveks kümedir. $x = x_1 + x_2$ öyle ki $x_1 \in K_1$, $x_2 \in K_2$ olsun. $x \in K_1 + K_2$ ve her $\lambda > 0$ için $\lambda x = \lambda(x_1 + x_2) \in K_1 + K_2$ olduğundan $K_1 + K_2$ toplamı da konidir. $x^* \in K_1^* \cap K_2^*$ olsun. Bu durumda her $x_1 \in K_1$ için $\langle x_1, x^* \rangle \geqslant 0$ ve her $x_2 \in K_2$ için $\langle x_2, x^* \rangle \geqslant 0$ olduğundan her $x_1 + x_2 \in K_1 + K_2$ için $\langle x_1 + x_2, x^* \rangle \geqslant 0$ olur. Böylece $x^* \in (K_1 + K_2)^*$ elde edilir ve $K_1^* \cap K_2^* \subseteq (K_1 + K_2)^*$ kapsaması gerçeklenir. Tersine $x^* \in (K_1 + K_2)^*$ olsun. Her $x = x_1 + x_2$ öyle ki $x_1 \in K_1, x_2 \in K_2$ için $\langle x, x^* \rangle \geqslant 0$ olur. Yani her $x_1 \in K_1, x_2 \in K_2$ için $\langle x_1, x^* \rangle + \langle x_2, x^* \rangle \geqslant 0$ dır. Her $\lambda \geqslant 0$ için K_1 ve K_2 koni olduğundan $\langle \lambda x_1, x^* \rangle + \langle \lambda x_2, x^* \rangle \geqslant 0$ dır. Bunlardan birisi sıfıra yeterince yaklaşabileceği göz önüne alınırsa her $x_1 \in K_1$, $\langle x_1, x^* \rangle \geqslant 0$ ve her $x_2 \in K_2$, $\langle x_2, x^* \rangle \geqslant 0$ eşitsizlikleri yazılabilir. Böylece $x^* \in K_1^* \cap K_2^*$ olacağından $(K_1 + K_2)^* \subseteq K_1^* \cap K_2^*$ kapsaması elde edildiğinden ispat tamamlanır. ■

Lemma 2.22. Kapalı $K_1,\ K_2$ konileri için $(K_1 \cap K_2)^* = \overline{K_1^* + K_2^*}$ dir.

İspat: Kapalı K_1, K_2 konileri için

$$(K_1 \cap K_2)^* = (K_1^{**} \cap K_2^{**})^* = \left(\left(K_1^* + K_2^*\right)^*\right)^* = \left(K_1^* + K_2^*\right)^{**}$$

olur. Uyarı 2.1 nedeniyle $(K_1 \cap K_2)^* = \overline{K_1^* + K_2^*}$ eşitliği bulunur. ■

2.2. KONVEKS FONKSİYON VE EŞLENİK FONKSİYON

Konveks fonksiyonlar, konveks analiz çalışmalarının temel konularındandır. Bu kısımda konveks fonksiyonların genel tanım ve özellikleri, yöne göre türev ve subdiferansiyel kavramları, konveks fonksiyonların diferansiyellenebilme ve süreklilikleri hakkındaki temel bilgiler verilecektir. Ayrıca matematik programlamada kullanılacak olan eşlenik fonksiyonların yapısı incelenecektir.

2.2.1. Konveks Fonksiyonlar

$f : \mathbb{R}^n \to [-\infty, +\infty]$ fonksiyonunun tanım kümesi

$$domf = \{x : \ f(x) < +\infty\}$$

ile verilir. $\alpha \in \mathbb{R}$ ve $x \in \mathbb{R}^n$ olmak üzere $f(x) \leqslant \alpha$ gerçekleyen tüm $(x, \alpha) \in \mathbb{R}^{n+1}$ nokta çiftlerinden oluşan kümeye f fonksiyonunun grafiküstü kümesi denir ve

$$epif = \{(x, \alpha) : \ f(x) \leqslant \alpha\}$$

ile gösterilir. Burada $x \in domf$ olduğunda $(x, \alpha) \in \mathbb{R}^{n+1}$ noktası $epif$ kümesine ait olur. Ancak $f(x) = +\infty$ ise $f(x) \leqslant \alpha$ gerçekleyen α noktası yoktur. f fonksiyonunun grafiküstü kümesi ile f fonksiyonu belirlenebilir, şöyleki

$$f(x) = \inf_{\alpha}\{\alpha : \ (x, \alpha) \in epif\} \tag{2.4}$$

dir. (2.4) formülünden anlaşılacağı üzere R^{n+1} uzayındaki kümeler ile R^n uzayındaki fonksiyonlar arasında yakın bir ilişki vardır.

Tanım 2.7. f fonksiyonunun $epif$ kümesi konveks ise f fonksiyonuna *konveks fonksiyon* denir.

Tanım 2.8. $-\infty$ değerini almayan ve özdeş olarak da $+\infty$ değerine eşit olmayan bir f fonksiyonuna *proper* ya da *has fonksiyon* denir.

Bir f has fonksiyonu için $domf \neq \emptyset$ dır. Böylece $x \in domf$ için $f(x)$ sonlu bir değer alır.

Lemma 2.23. f has fonksiyonunun konveks olabilmesi için gerek ve yeter koşul her x_1, x_2 için $\lambda_1 \geqslant 0$ ve $\lambda_2 \geqslant 0, \ \lambda_1 + \lambda_2 = 1$ olmak üzere

$$f(\lambda_1 x_1 + \lambda_2 x_2) \leqslant \lambda_1 f(x_1) + \lambda_2 f(x_2) \tag{2.5}$$

olmasıdır.

İspat: f has konveks fonksiyon olsun. $\lambda_1, \lambda_2 \geqslant 0, \ \lambda_1 + \lambda_2 = 1$ olmak üzere $(x_1, \alpha_1) \in epif, \ (x_2, \alpha_2) \in epif$ ise $f(x_1) \leqslant \alpha_1$ ve $f(x_2) \leqslant \alpha_2$ yazılır ve Tanım 2.7 gereği $epif$ kümesi konveks olduğundan $\lambda_1(x_1, \alpha_1) + \lambda_2(x_2, \alpha_2) \in epif$ tir. Yani

$$f(\lambda_1 x_1 + \lambda_2 x_2) \leqslant \lambda_1 \alpha_1 + \lambda_2 \alpha_2$$

eşitsizliği gerçeklenir. Özel olarak $\alpha_1 = f(x_1)$ ve $\alpha_2 = f(x_2)$ alınırsa $\lambda_1, \lambda_2 \geqslant 0$ ve $\lambda_1 + \lambda_2 = 1$ için

$$f(\lambda_1 x_1 + \lambda_2 x_2) \leqslant \lambda_1 f(x_1) + \lambda_2 f(x_2)$$

olup istenen eşitsizlik elde edilir. Tersine $\lambda_1, \lambda_2 \geqslant 0,\ \lambda_1 + \lambda_2 = 1$ olmak üzere $f(\lambda_1 x_1 + \lambda_2 x_2) \leqslant \lambda_1 f(x_1) + \lambda_2 f(x_2)$ sağlansın. $(x_1, \alpha_1),\ (x_2, \alpha_2) \in epif$ olsun.

$$f(\lambda_1 x_1 + \lambda_2 x_2) \leqslant \lambda_1 f(x_1) + \lambda_2 f(x_2) \leqslant \lambda_1 \alpha_1 + \lambda_2 \alpha_2$$

olduğundan $(\lambda_1 x_1 + \lambda_2 x_2, \lambda_1 \alpha_1 + \lambda_2 \alpha_2) \in epif$ olur bu ise $epif$ kümesinin dolayısıyla f fonksiyonunun konveks olduğunu gösterir. ■

Tanım 2.9. f has fonksiyon ve $\lambda_1, \lambda_2 \geqslant 0$ ve $\lambda_1 + \lambda_2 = 1$ olmak üzere her $x_1, x_2 \in \mathbb{R}^n$ için

$$f(\lambda_1 x_1 + \lambda_2 x_2) < \lambda_1 f(x_1) + \lambda_2 f(x_2) \tag{2.6}$$

eşitsizliği gerçekleniyorsa f fonksiyonuna *kesin konveks fonksiyon* denir.

$\alpha \in \mathbb{R}$ ve $x \in \mathbb{R}^n$ olmak üzere $\mathbb{R}^n$ uzayında tanımlı g fonksiyonu için $\alpha \leqslant g(x)$ gerçekleyen tüm $(x, \alpha) \in \mathbb{R}^{n+1}$ nokta çiftlerinden oluşan küme g fonksiyonunun grafikaltı kümesidir ve

$$hyp\ g = \{(x, \alpha) :\ \alpha \leqslant g(x)\}$$

ile gösterilir.

Tanım 2.10. g fonksiyonunun $hyp\ g$ kümesi konveks ise g fonksiyonuna *konkav fonksiyon* denir. Eşdeğer olarak her x_1, x_2 için $\lambda_1 \geqslant 0$ ve $\lambda_2 \geqslant 0,\ \lambda_1 + \lambda_2 = 1$ olmak üzere

$$g(\lambda_1 x_1 + \lambda_2 x_2) \geqslant \lambda_1 g(x_1) + \lambda_2 g(x_2)$$

ise g foksiyonuna *konkav fonksiyon* denir.

Ayrıca eğer f konveks fonksiyon ise Lemma 2.23 nedeniyle $-f$ konkav bir fonksiyon olacağından Tanım 2.10 'a bir diğer eşdeğer tanım olarak eğer $-f$ fonksiyonu konveks ise f fonksiyonu konkav bir fonksiyon olduğu verilebilir.

Lemma 2.24. Eğer f fonksiyonu konveks ise $domf$ kümesi de konvekstir.

İspat: f fonksiyonu konveks ve $x_1, x_2 \in domf$ olsun. O halde sonlu α_1, α_2 sayıları vardır öyle ki $f(x_1) \leqslant \alpha_1$ ve $f(x_2) \leqslant \alpha_2$ dir. Buradan $(x_1, \alpha_1),\ (x_2, \alpha_2) \in epif$ olur.

f konveks fonksiyon olduğundan $epif$ kümesi konvekstir. $\lambda_1, \lambda_2 \geqslant 0,\ \lambda_1 + \lambda_2 = 1$ olmak üzere $(\lambda_1 x_1 + \lambda_2 x_2, \lambda_1 \alpha_1 + \lambda_2 \alpha_2) \in epif$ yani

$$f(\lambda_1 x_1 + \lambda_2 x_2) \leqslant \lambda_1 \alpha_1 + \lambda_2 \alpha_2 < \infty$$

gerçeklenir. Böylece $\lambda_1 x_1 + \lambda_2 x_2 \in domf$ ve dolayısıyla $domf$ kümesi konvekstir. ∎

Lemma 2.25. (Jensen Eşitsizliği) f has konveks fonksiyon olsun. Bu durumda $\lambda_i \geqslant 0,\ i = 1, \ldots, m$ ve $\lambda_1 + \cdots + \lambda_m = 1$ olmak üzere

$$f(\lambda_1 x_1 + \cdots + \lambda_m x_m) \leqslant \lambda_1 f(x_1) + \cdots + \lambda_m f(x_m)$$

eşitsizliği gerçeklenir.

İspat: Matematik indiksiyon yöntemini kullanılırsa her $i = 1, \ldots, k$ için $\lambda_i \geqslant 0$ ve $\sum\limits_{i=1}^{k} \lambda_i = 1$ olmak üzere $m = k$ için $f(\lambda_1 x_1 + \cdots + \lambda_k x_k) \leqslant \lambda_1 f(x_1) + \cdots + \lambda_k f(x_k)$ doğru olsun. $m = k+1$ için doğru olduğunu göstermemiz yeterlidir. Her $i = 1, \ldots, k+1$ için $\lambda_i \geqslant 0$ ve $\sum\limits_{i=1}^{k+1} \lambda_i = 1$ olduğundan λ_i katsayılarından en azından biri kabul edelim ki $\lambda_{k+1} < 1$ olsun. $\lambda = \lambda_1 + \cdots + \lambda_k = 1 - \lambda_{k+1}$ seçilirse $\lambda > 0$ olup $y = \frac{\lambda_1}{\lambda} x_1 + \cdots + \frac{\lambda_k}{\lambda} x_k$ için kabulden $f(y) = f(\frac{\lambda_1}{\lambda} x_1 + \cdots + \frac{\lambda_k}{\lambda} x_k) \leqslant \frac{\lambda_1}{\lambda} f(x_1) + \cdots + \frac{\lambda_k}{\lambda} f(x_k)$ elde edilir. $f(\lambda_1 x_1 + \cdots + \lambda_k x_k + \lambda_{k+1} x_{k+1}) = f(\lambda y + \lambda_{k+1} x_{k+1}) \leqslant \lambda f(y) + \lambda_{k+1} f(x_{k+1})$ olur. Böylece istenen elde edilir. ∎

Lemma 2.26. $\emptyset \neq A \subseteq \mathbb{R}^n$ konveks küme üzerinde tanımlı f ve g fonksiyonlar olmak üzere $\alpha, \beta \geqslant 0$ için $\alpha f + \beta g$ fonksiyonu konvekstir.

İspat: $x_1, x_2 \in A$ olsun. $\lambda_1, \lambda_2 \geqslant 0,\ \lambda_1 + \lambda_2 = 1$ olmak üzere

$$\begin{aligned}
(\alpha f + \beta g)(\lambda_1 x_1 + \lambda_2 x_2) &= \alpha f(\lambda_1 x_1 + \lambda_2 x_2) + \beta g(\lambda_1 x_1 + \lambda_2 x_2) \\
&\leqslant \alpha\big(\lambda_1 f(x_1) + \lambda_2 f(x_2)\big) + \beta\big(\lambda_1 g(x_1) + \lambda_2 g(x_2)\big) \\
&= \lambda_1\big(\alpha f(x_1) + \beta g(x_1)\big) + \lambda_2\big(\alpha f(x_2) + \beta g(x_2)\big) \\
&= \lambda_1\big((\alpha f + \beta g)(x_1)\big) + \lambda_2\big((\alpha f + \beta g)(x_2)\big)
\end{aligned}$$

Dolayısıyla $\alpha f + \beta g$ fonksiyonu konvekstir. ∎

Lemma 2.27. $\emptyset \neq A, B \subseteq \mathbb{R}^n$ konveks kümeler, $f : A \to \mathbb{R}$ konveks fonksiyon, $g : B \to \mathbb{R}$ artan konveks fonksiyon ve $f(A) \subseteq B$ ise $g \circ f : A \to \mathbb{R}$ konveks fonksiyondur.

İspat: $x_1, x_2 \in A$ olsun. $\lambda_1, \lambda_2 \geqslant 0,\ \lambda_1 + \lambda_2 = 1$ olmak üzere $(g \circ f)(\lambda_1 x_1 + \lambda_2 x_2) = g\big(f(\lambda_1 x_1 + \lambda_2 x_2)\big)$ olup f konveks fonksiyon ve g artan konveks fonksiyon olduğundan

$$(g \circ f)(\lambda_1 x_1 + \lambda_2 x_2) \ \leqslant\ \lambda_1 g(f(x_1)) + \lambda_2 g(f(x_2))$$

elde edilir. Böylece $g \circ f$ bileşke fonksiyonu konvekstir. ∎

Lemma 2.28. *I herhangi bir indis kümesi olmak üzere her $i \in I$ için $f_i(x)$ konveks fonksiyon olsun. Bu durumda $f(x) = \sup_{i \in I} f_i(x)$ fonksiyonu da konvekstir.*

İspat: f_i fonksiyonlarının grafiküstü kümesinin kesişimi f fonksiyonunun grafiküstü kümesidir. Açıkca $epif = \cap_{i \in I} epif_i$ dir. Her $i \in I$ için f_i fonksiyonları konveks olduğundan f fonksiyonu konvekstir. ∎

Daha sonraki bölümlerde kullanılacak olan bazı özel konveks fonksiyon örnekleri aşağıda verilmektedir.

Lemma 2.29. *$\emptyset \neq A \subseteq \mathbb{R}^n$ konveks küme olsun. O halde her $u \in \mathbb{R}^n$ için*

$$d_A(u) = \inf\{||u - x|| :\ x \in A\}$$

uzaklık fonksiyonu konvekstir.

İspat: $x, y \in A$ olsun. Her $\varepsilon > 0$ ve $u, v \in \mathbb{R}^n$ vektörleri için infimum tanımından $||u-x|| < d_A(u)+\varepsilon$ ve $||v-y|| < d_A(v)+\varepsilon$ eşitsizlikleri sağlanır. $\lambda_1, \lambda_2 \geqslant 0,\ \lambda_1 + \lambda_2 = 1$ olmak üzere

$$\begin{aligned} d_A(\lambda_1 u + \lambda_2 v) &\leqslant ||\lambda_1 u + \lambda_2 v - (\lambda_1 x + \lambda_2 y)|| \\ &\leqslant \lambda_1 ||u - x|| + \lambda_2 ||v - y|| \\ &< \lambda_1\big(d_A(u) + \varepsilon\big) + \lambda_2\big(d_A(v) + \varepsilon\big) \end{aligned} \tag{2.7}$$

(2.7) eşitsizliği her $\varepsilon > 0$ için sağlanacağından

$$d_A(\lambda_1 u + \lambda_2 v) \leqslant \lambda_1 d_A(u) + \lambda_2 d_A(v)$$

olduğundan uzaklık fonksiyonu konvekstir. ∎

Örnek 2.2. $\emptyset \neq A \subseteq \mathbb{R}^n$ konveks kümesinin indikatör fonksiyonu

$$\delta_A(x) = \begin{cases} 0 & ,\ x \in A \\ +\infty & ,\ x \notin A \end{cases}$$

şeklinde tanımlanır ve $\delta_A(x)$ indikatör fonksiyonu konveks bir fonksiyondur.

Örnek 2.3. Bir A konveks kümesinin bir $x^* \in \mathbb{R}^n$ noktasındaki destek fonksiyonu

$$W_A(x^*) = \sup_x\{\langle x, x^*\rangle \ : \ x \in A\}$$

konveks bir fonksiyondur.

Tanım 2.11. λ pozitif reel sayı ve $x \in \mathbb{R}^n$ olmak üzere konveks koni kümesi üzerinde tanımlı f fonksiyonuna $f(\lambda x) = \lambda f(x)$ bağıntısını sağlıyorsa *pozitif homojen fonksiyon* denir. Ayrıca $f(\lambda x) = \lambda^k f(x)$ ise f fonksiyonu k. dereceden pozitif homojen fonksiyondur.

Teorem 2.30. f fonksiyonu, A konveks konisi üzerinde tanımlı pozitif homojen bir fonksiyon olsun. f fonksiyonunun konveks olabilmesi için gerek ve yeter koşul $f(x+y) \leqslant f(x) + f(y)$ olmasıdır.

İspat: f pozitif homojen konveks fonksiyon olsun. $x, y \in A$ için

$$\frac{1}{2}f(x+y) = f(\frac{1}{2}x + \frac{1}{2}y) \leqslant \frac{1}{2}\big(f(x) + f(y)\big)$$

olduğundan $f(x+y) \leqslant f(x) + f(y)$ eşitsizliği sağlanır. Tersine $x, y \in A$ için $f(x+y) \leqslant f(x) + f(y)$ olsun. $\lambda_1, \lambda_2 \geqslant 0,\ \lambda_1 + \lambda_2 = 1$ olmak üzere

$$f(\lambda_1 x + \lambda_2 y) \leqslant f(\lambda_1 x) + f(\lambda_2 y) = \lambda_1 f(x) + \lambda_2 f(y)$$

olduğundan f fonksiyonu konvekstir. ■

Tanım 2.12. $A \subseteq \mathbb{R}^n$ konveks küme üzerinde tanımlı f fonksiyonu tanımlansın. Her $\alpha \in \mathbb{R}$ için

$$L_\alpha(f) = \{x \in A : f(x) \leqslant \alpha\}$$

kümesine f fonksiyonunun $\alpha-$*seviye kümesi* denir.

Lemma 2.31. f fonksiyonu konveks ise her $\alpha-$seviye kümesi konvekstir.

İspat: $x, y \in L_\alpha(f)$ olsun. Tanım 2.12 nedeniyle her $\alpha \in \mathbb{R}$ için $f(x) \leqslant \alpha$ ve $f(y) \leqslant \alpha$ dır. $\lambda_1, \lambda_2 \geqslant 0,\ \lambda_1 + \lambda_2 = 1$ olmak üzere f konveks fonksiyon olduğundan $f(\lambda_1 x + \lambda_2 y) \leqslant \lambda_1 f(x) + \lambda_2 f(y) \leqslant \alpha$ olup $\lambda_1 x + \lambda_2 y \in L_\alpha(f)$ bulunur. Dolayısıyla her $\alpha-$seviye kümesi konvekstir. ■

Fakat her seviye kümesi konveks olan f fonksiyonu konveks olmayabilir.

Teorem 2.32. $A \subseteq \mathbb{R}^n$ konveks kümesi üzerinde tanımlı f negatif olmayan pozitif homojen fonksiyon ve $L_1(f) = \{x \in A : f(x) \leqslant 1\}$ seviye kümesi konveks olsun. Bu durumda f fonksiyonu konvekstir.

İspat: $x, y \in A$ ve $\alpha, \beta \in \mathbb{R}$ için $0 \leqslant f(x) < \alpha$, $0 \leqslant f(y) < \beta$ olsun. Ayrıca $f(\frac{x}{\alpha}) = \frac{1}{\alpha} f(x) < 1$ ve $f(\frac{y}{\beta}) = \frac{1}{\beta} f(y) < 1$ olacağından $\frac{x}{\alpha}, \frac{y}{\beta} \in L_1(f)$ dir.

$$\frac{f(x+y)}{\alpha+\beta} = f\Big(\frac{x+y}{\alpha+\beta}\Big) = f\Big(\frac{\alpha}{\alpha+\beta} \cdot \frac{x}{\alpha} + \frac{\beta}{\alpha+\beta} \cdot \frac{y}{\beta}\Big)$$

ifadesinde $L_1(f)$ seviye kümesi konveks olduğundan $\frac{\alpha}{\alpha+\beta}\frac{x}{\alpha} + \frac{\beta}{\alpha+\beta}\frac{y}{\beta} \in L_1(f)$ dir. Dolayısıyla $f(x+y) \leqslant \alpha+\beta$ dir. Her $\alpha > f(x)$, $f(x+y) < \alpha$ için $f(x+y) \leqslant f(x)$ dır ve her $\beta > f(y)$, $f(x+y) < \beta$ için $f(x+y) \leqslant f(y)$ dır. Buradan $f(x+y) \leqslant f(x)+f(y)$ olup Teorem 2.30 gereğince f fonksiyonu konvekstir. ∎

Negatif olmayan ortant X kümesi üzerinde tanımlı bir pozitif homojen f fonksiyonu $x_1, \ldots, x_n \geqslant 0$ ve $p \geqslant 1$ için $f(x_1 \ldots, x_n) = (x_1^p + \cdots + x_n^p)^{\frac{1}{p}}$ verilsin. $f^p : X \to \mathbb{R}^n$ fonksiyonu konvekstir. Lemma 2.31'den

$$\{x \in X : \ f^p(x) \leqslant 1\} = \{x \in X : f(x) \leqslant 1\}$$

seviye kümesi konvekstir. Teorem 2.30 nedeniyle $x, y \in X$ için

$$\Big((x_1+y_1)^p + \cdots + (x_n+y_n)^p\Big)^{\frac{1}{p}} \leqslant (x_1^p + \cdots + x_n^p)^{\frac{1}{p}} + (y_1^p + \cdots + y_n^p)^{\frac{1}{p}}$$

olur. Açıkca

$$\Big(\sum_{i=1}^{n}(x_i+y_i)^p\Big)^{\frac{1}{p}} \leqslant \Big(\sum_{i=1}^{n} x_i^p\Big)^{\frac{1}{p}} + \Big(\sum_{i=1}^{n} y_i^p\Big)^{\frac{1}{p}}$$

Minkowski Eşitsizliği elde edilir.

$(0, +\infty)$ aralığında $f(x) = -\log x$ konveks fonksiyon olduğundan $x_1, \ldots, x_m > 0$, $\alpha_1, \ldots, \alpha_m > 0$ ve $\alpha_1 + \cdots + \alpha_m = 1$ için

$$-\log(\alpha_1 x_1 + \cdots + \alpha_m x_m) \leqslant -(\alpha_1 \log x_1 + \cdots + \alpha_m \log x_m) = -\Big(\log(x_1^{\alpha_1} \cdot \ldots \cdot x_m^{\alpha_m})\Big)$$

eşitsizliği -1 ile çarpılırsa logaritma fonksiyonunun özelliğinden

$$x_1^{\alpha_1} \cdot \ldots \cdot x_m^{\alpha_m} \leqslant \alpha_1 x_1 + \cdots + \alpha_m x_m \tag{2.8}$$

elde edilir. Burada $\alpha_1 = \cdots = \alpha_m = \frac{1}{m}$ alınırsa

$$(x_1 \cdot \ldots \cdot x_m)^{\frac{1}{m}} \leqslant \frac{x_1 + \cdots + x_m}{m}$$

olup bu ise $x_1, \ldots, x_m$ noktalarının geometrik ortalamasının aritmetik ortalamasından daha büyük olmadığını gösterir.

$i = 1, \ldots, m$ ve $j = 1, \ldots, n$ için $a_{ij} > 0$, $\alpha_1, \ldots, \alpha_m > 0$ ve $\alpha_1 + \cdots + \alpha_m = 1$ ise (2.8) denkleminden

$$\frac{a_{1j}^{\alpha_1} \cdot \ldots \cdot a_{mj}^{\alpha_m}}{(a_{11} + \cdot \cdot + a_{1n})^{\alpha_1} \cdot \ldots \cdot (a_{m1} + \cdot \cdot + a_{mn})^{\alpha_m}} \leqslant \frac{\alpha_1 a_{1j}}{a_{11} + \cdot \cdot + a_{1n}} + \cdot \cdot + \frac{\alpha_m a_{mj}}{a_{m1} + \cdot \cdot + a_{mn}}$$

elde edilir. Her $j = 1, \ldots, n$ için

$$\frac{\sum\limits_{j=1}^{n} a_{1j}^{\alpha_1} \cdot \ldots \cdot a_{mj}^{\alpha_m}}{(a_{11} + \cdot \cdot + a_{1n})^{\alpha_1} \cdot \ldots \cdot (a_{m1} + \cdot \cdot + a_{mn})^{\alpha_m}} \leqslant \frac{\alpha_1 \sum\limits_{j=1}^{n} a_{1j}}{a_{11} + \cdot \cdot + a_{1n}} + \cdot \cdot + \frac{\alpha_m \sum\limits_{j=1}^{n} a_{mj}}{a_{m1} + \cdot \cdot + a_{mn}}$$

olup düzenleme yapılırsa,

$$\sum_{j=1}^{n} a_{1j}^{\alpha_1} \cdot \ldots \cdot a_{mj}^{\alpha_m} \leqslant (a_{11} + \cdot \cdot + a_{1n})^{\alpha_1} \cdot \ldots \cdot (a_{m1} + \cdot \cdot + a_{mn})^{\alpha_m}$$

olur. Buradan $\alpha_1 = \frac{1}{p}$, $\alpha_2 = \frac{1}{q}$, $m = 2$ ve $a_{1j} = x_j^p$, $a_{2j} = y_j^q$ alınırsa

$$\sum_{j=1}^{n} x_j y_j \leqslant \Big(\sum_{j=1}^{n} x_j^p\Big)^{\frac{1}{p}} \cdot \Big(\sum_{j=1}^{n} y_j^q\Big)^{\frac{1}{q}}$$

Hölder Eşitsizliği ispatlanır.

Lemma 2.33. *$I \subseteq \mathbb{R}$ bir aralık, $f : I \to \mathbb{R}$ konveks fonksiyonu $x, y, z \in I$ ve $x < z < y$ olsun. Bu durumda*

$$\frac{f(z) - f(x)}{z - x} \leqslant \frac{f(y) - f(x)}{y - x} \leqslant \frac{f(y) - f(z)}{y - z}$$

eşitsizliği gerçeklenir.

İspat: $z = \frac{y-z}{y-x} x + \frac{z-x}{y-x} y$ şeklinde yazılırsa ve f konveks fonksiyon olduğundan

$$f(z) \leqslant \frac{y - z}{y - x} f(x) + \frac{z - x}{y - x} f(y) \tag{2.9}$$

olur. Eşitsizliğin her iki tarafından $f(x)$ değeri çıkarılıp, düzenleme yapılırsa

$$\frac{f(z) - f(x)}{z - x} \leqslant \frac{f(y) - f(x)}{y - x} \tag{2.10}$$

elde edilir. (2.9) bağıntısı -1 ile çarpılıp, eşitsizliğin her iki tarafına $f(y)$ eklenip düzenleme yapılırsa

$$\frac{f(y) - f(x)}{y - x} \leqslant \frac{f(y) - f(z)}{y - z} \tag{2.11}$$

elde edilir. ∎

Lemma 2.34. [14] $I \subseteq \mathbb{R}$ bir aralık, $f : I \to \mathbb{R}$ konveks fonksiyon I aralığının her bir iç noktasında f'_+ ve f'_- var ve $a, b \in int(I)$ için $a < b$ ise

$$f'_-(a) \leqslant f'_+(a) \leqslant \frac{f(b) - f(a)}{b - a} \leqslant f'_-(b) \leqslant f'_+(b) \tag{2.12}$$

eşitsizliği sağlanır.

Tanım 2.13. $I \subseteq \mathbb{R}$ bir aralık, $f : I \to \mathbb{R}$ reel değerli bir fonksiyon olsun. T afin dönüşümüne $T(x_0) = f(x_0)$ ve her $x \in I$ için $T(x) \leqslant f(x)$ koşullarını sağlıyorsa x_0 noktasında f fonksiyonunun desteği denir ve $m \in \mathbb{R}$ için

$$T(x) = f(x_0) + m(x - x_0)$$

şeklinde yazılır.

Teorem 2.35. $I \subseteq \mathbb{R}$ bir açık aralık, $f : I \to \mathbb{R}$ reel değerli bir fonksiyonunun konveks olabilmesi için gerek ve yeter koşul I aralığının her bir noktasında desteğe sahip olmasıdır.

İspat: f fonksiyonu I aralığının her bir noktasında desteğe sahip olsun. $x, y \in I$, $\lambda_1, \lambda_2 \geqslant 0$, $\lambda_1 + \lambda_2 = 1$ olmak üzere $\lambda_1 x + \lambda_2 y$ noktasında f fonksiyonunun desteği T afin dönüşümü olsun. Tanım 2.13 kullanılarak

$$f(\lambda_1 x + \lambda_2 y) = T(\lambda_1 x + \lambda_2 y) = \lambda_1 T(x) + \lambda_2 T(y) \leqslant \lambda_1 f(x) + \lambda_2 f(y)$$

olup f fonksiyonu konvekstir. Tersine f konveks fonksiyon $x_0 \in I$ ve $m \in \mathbb{R}$ için $f'_-(x_0) \leqslant m \leqslant f'_+(x_0)$ olsun. $T(x) = f(x_0) + m(x - x_0)$ afin dönüşüm tanımlayalım. Herhangi $y, z \in I$ ve $y < x_0 < z$ olsun. Lemma 2.34' den

$$\frac{f(x_0) - f(y)}{x_0 - y} \leqslant f'_-(x_0) \leqslant m = \frac{T(y) - T(x_0)}{y - x_0} \tag{2.13}$$

olup düzenleme yapılırsa $T(y) \leqslant f(y)$ olur. Aynı şekilde Lemma 2.34' den

$$m \leqslant f'_+(x_0) \leqslant \frac{f(z) - f(x_0)}{z - x_0} \tag{2.14}$$

eşitsizliğinden $T(z) \leqslant f(z)$ olur. Ayrıca $T(x_0) = f(x_0)$ olduğundan T afin dönüşümü f fonksiyonunun x_0 noktasında desteğidir. ∎

Her konveks fonksiyon sürekli değildir. Fakat konveks fonksiyonun tanım kümesi açık küme ise bu küme içerisindeki her noktada fonksiyon süreklidir. Bu önemli özellik ise aşağıdaki teorem ile verilmektedir.

Teorem 2.36. Boştan farklı her açık konveks küme üzerinde tanımlı her konveks fonksiyon süreklidir.

İspat: A açık, konveks kümesinde yer alan P politopun elemanları $y_1, \dots, y_m$ olsun. $x_0 \in A$ noktası P politopunun iç noktası olsun. O halde öyle $r > 0$ vardır ki $B(x_0, r) \subseteq P$ yazılabilir. Her $x \in B(x_0, r)$ için $\lambda_1, \dots, \lambda_m \geqslant 0$ ve $\lambda_1 + \cdots + \lambda_m = 1$ olmak üzere $x = \lambda_1 y_1 + \cdots + \lambda_m y_m$ yazılır. Lemma 2.25 eşitsizliğinden ve f konveks fonksiyon olduğundan $f(x) \leqslant \lambda_1 f(y_1) + \cdots + \lambda_m f(y_m)$ olur. $M = \text{maks}\{f(y_1), \dots, f(y_m)\}$ ise $f(x) \leqslant M$ bulunur. Böylece $B(x_0, r)$ üzerinde f fonksiyonu üstten M ile sınırlıdır. $g : [-r, r] \to \mathbb{R}$ fonksiyonu $-r \leqslant t \leqslant r$ için $g(t) = f\Big(x_0 + t\frac{x-x_0}{||x-x_0||}\Big)$ şeklinde tanımlansın. g fonksiyonunun konveks olduğu kolayca görülür. Ayrıca her $x \in B(x_0, r)$ için $f(x) \leqslant M$ ve $\Big(x_0 + t\frac{x-x_0}{||x-x_0||}\Big) \in B(x_0, r)$ olduğundan $-r \leqslant t \leqslant r$ için $g(t) \leqslant M$ dir.

$$\frac{g(0) - M}{r} \leqslant \frac{g(0) - g(-r)}{r} \leqslant \frac{g(||x - x_0||) - g(0)}{||x - x_0||} \leqslant \frac{M - g(0)}{r}$$

olur. O halde

$$\Big|\frac{g(||x - x_0||) - g(0)}{||x - x_0||}\Big| \leqslant \frac{M - g(0)}{r} = \frac{M - f(x_0)}{r}$$

eşitsizliği sağlanır ve böylece

$$|f(x) - f(x_0)| \leqslant \frac{M - f(x_0)}{r}||x - x_0|| \tag{2.15}$$

elde edilir. A kümesinin $x_1, \dots, x_k, \dots$ noktalarının dizisi x_0 noktasına yakınsıyorsa $k \to \infty$ için $f(x_k) \to f(x_0)$ olup f fonksiyonu x_0 noktasında süreklidir. $x_0 \in A$ keyfi nokta olduğundan f fonksiyonu A kümesi üzerinde süreklidir. ■

Böylece $f : \mathbb{R}^n \to \mathbb{R}$ konveks fonksiyon ise sürekli olduğu kolayca görülür. Daha genel olarak $f : A \to \mathbb{R}$ konveks fonksiyon ise $A \subseteq \mathbb{R}^n$ konveks kümesinin her izafi iç noktasında süreklidir.

2.2.2. Konveks Fonksiyonların Diferansiyellenebilme Özellikleri

Konveks fonksiyon her zaman türevli olmayabilir. Örneğin $f(x) = |x|$ konveks bir fonksiyon olmasına rağmen $x_0 = 0$ noktasında türevli değildir. Konveks fonksiyonlar diferansiyellenebilir olmak zorunda değildir ancak yöne göre türevi daima vardır.

Tanım 2.14. $X \subseteq \mathbb{R}^n$ açık küme üzerinde tanımlı reel değerli f fonksiyonu verilsin. f fonksiyonu $x \in X$ noktasında i. *kısmi türevi* $\frac{\partial f}{\partial x_i}$ $(i = 1, \ldots, n)$ ile gösterilir ve

$$\frac{\partial f}{\partial x_i}(x) = \lim_{\lambda \to 0} \frac{f(x_1, \ldots, x_i + \lambda, \ldots, x_n) - f(x_1, \ldots, x_n)}{\lambda}$$

şeklinde tanımlanır. Eşdeğer tanım olarak e_i birim vektör olmak üzere

$$\frac{\partial f}{\partial x_i}(x) = \lim_{\lambda \to 0} \frac{f(x + \lambda e_i) - f(x)}{\lambda}$$

tanımı da verilir.

e_i birim vektör yerine keyfi $y \in \mathbb{R}^n$ vektörü alınırsa f fonksiyonunun x noktasında y yönündeki *yönlü türevi* elde edilir ve

$$Df(x)(y) = \lim_{\lambda \to 0} \frac{f(x + \lambda y) - f(x)}{\lambda} \tag{2.16}$$

ile tanımlanır. Bu tanımdan hareketle f fonksiyonunun e_i birim vektör yönünde x noktasındaki yönlü türevi, f fonksiyonun i. kısmi türevi olduğu sonucu elde edilir.

Tanım 2.15. $X \subseteq \mathbb{R}^n$ açık küme üzerinde tanımlı reel değerli f fonksiyonu için öyle $x^* \in \mathbb{R}^n$ vektörü vardır öyle ki

$$\lim_{u \to 0} \frac{f(x + u) - f(x) - \langle x^*, u \rangle}{||u||} = 0 \tag{2.17}$$

denklemi sağlanıyorsa f fonksiyonu $x \in X$ noktasında *diferansiyellenebilirdir* denir. Burada x^* vektörü, f fonksiyonunun x noktasındaki gradient vektörüdür.

Teorem 2.37. f has konveks fonksiyon ve $x \in dom f$ olsun. O halde her $y \in \mathbb{R}^n$ için $Df(x)(y)$ yönlü türevi vardır. Ayrıca $Df(x)(y)$ yönlü türevi pozitif homojen konveks fonksiyondur.

İspat: $x + \lambda y \notin dom f$ ise her $\lambda > 0$ için $Df(x)(y) = +\infty$ olur. $x + \lambda y \in dom f$ ve $\lambda > 0$ sıfıra çok yakın ise Lemma 2.33 'den $\frac{f(x+\lambda y) - f(x)}{\lambda}$ ifadesi λ sıfıra azalarak yaklaştıkça λ nın azalmayan bir fonksiyonudur. Bu nedenle daima

$$\lim_{\lambda \to 0} \frac{f(x + \lambda y) - f(x)}{\lambda}$$

limiti var olduğundan $Df(x)(y)$ yönlü türevi vardır. $\alpha > 0$ olsun.

$$Df(x)(\alpha y) = \lim_{\lambda \to 0} \frac{f(x + \lambda \alpha y) - f(x)}{\lambda} = \alpha \lim_{\lambda \to 0} \frac{f(x + \lambda \alpha y) - f(x)}{\lambda \alpha} = \alpha Df(x)(y)$$

olduğundan $Df(x)(y)$ yönlü türevi pozitif homojendir. Eğer $y_1, y_2 \in \mathbb{R}^n,\ \alpha_1, \alpha_2 \geqslant 0$ ve $\alpha_1 + \alpha_2 = 1$ için f konveks fonksiyon olduğundan

$$\frac{f\Big(x + \lambda\big(\alpha_1 y_1 + \alpha_2 y_2\big)\Big) - f(x)}{\lambda} \leqslant \alpha_1 \Big(\frac{f(x + \lambda y_1) - f(x)}{\lambda}\Big) + \alpha_2 \Big(\frac{f(x + \lambda y_2) - f(x)}{\lambda}\Big)$$

$\lambda \to 0$ limit alınırsa ve (2.16) eşitliğinden

$$Df(x)\big(\alpha_1 y_1 + \alpha_2 y_2\big) \leqslant \alpha_1 Df(x)(y_1) + \alpha_2 Df(x)(y_2)$$

elde edilir ve böylece $Df(x)(y)$ yönlü türevi konvekstir. ■

Teorem 2.38. $X \subseteq \mathbb{R}^n$ açık konveks küme üzerinde tanımlı f konveks fonksiyon olsun. $x \in X$ noktasında tüm kısmi türevlere sahip ise f fonksiyonu x noktasında diferansiyellenebilirdir.

İspat: Öyle $r > 0$ için $B(x, r) \subseteq X$ ve her $u = (u_1, \ldots, u_n) \in B(0, r)$ için

$$\psi(u) = f(x + u) - f(x) - \Big(\frac{\partial f}{\partial x_1} u_1 + \cdots + \frac{\partial f}{\partial x_n} u_n\Big) \tag{2.18}$$

olsun. ψ fonksiyonu $B(0, r)$ yuvarında konvekstir. $i = 1, \ldots, n$ için $B(0, r)$ de θ_i fonksiyonu tanımlayalım. $u \in B(0, r)$ olmak üzere

$$\theta_i(u) = \begin{cases} \dfrac{\psi(u_i.e_i)}{u_i} & , \ u_i \neq 0 \\ 0 & , \ u_i = 0 \end{cases} \qquad i = \ 1, \ldots, n$$

$u \to 0,\ \theta_i(u) \to 0$ dır. Her $u = (u_1, \ldots, u_n)$ ve $n||u|| < r$ için

$$\psi(u) = \psi\Big(\frac{1}{n}(n u_1 e_1) + \cdots + \frac{1}{n}(n u_n e_n)\Big) \leqslant \frac{1}{n}\psi(n u_1 e_1) + \cdots + \frac{1}{n}\psi(n u_n e_n)$$

eşitsizliğinden ve θ fonksiyonu tanımından

$$\psi(u) \leqslant \Big| u_1 \theta_1(nu) + \cdots + u_n \theta_n(nu) \Big| \leqslant ||u|| \Big(|\theta_1(nu)| + \cdots + |\theta_n(nu)| \Big) \tag{2.19}$$

elde edilir. $\psi(0) = 0$ ve $\psi\big(\frac{1}{2}u + \frac{1}{2}(-u)\big) \leqslant \frac{\psi(u) + \psi(-u)}{2}$ olduğundan $-\psi(-u) \leqslant \psi(u)$ bulunur. Böylece (2.19) bağıntısından

$$-||u|| \Big(|\theta_1(-nu)| + \cdots + |\theta_n(-nu)| \Big) \leqslant \psi(u) \leqslant ||u|| \Big(|\theta_1(nu)| + \cdots + |\theta_n(nu)| \Big)$$

olup bu eşitsizlikte her taraf $||u||$ bölünüp, $u \to 0$ iken limit alınırsa $\frac{\psi(u)}{||u||} \to 0$ olur.

$$\frac{\psi(u)}{||u||} = \frac{f(x + u) - f(x)}{||u||} - \Big(\frac{\partial f}{\partial x_1} u_1 + \cdots + \frac{\partial f}{\partial x_n} u_n\Big)\frac{1}{||u||} = 0$$

olacağından $(\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n})$ gradient vektörü ile f fonksiyonu x noktasında diferansiyellenebilirdir. ■

Tanım 2.16. f has konveks fonksiyon olmak üzere, her $x \in \mathbb{R}^n$ için

$$f(x) - f(x_o) \geqslant \langle x - x_0,\ x^* \rangle$$

eşitsizliğini gerçekleyen x^* vektörüne f fonksiyonunun $x_0 \in domf$ noktasındaki *subgradienti* denir.

Tanım 2.17. f has konveks fonksiyonunun bir x_0 noktasındaki subgradientlerinin kümesine f fonksiyonunun x_0 noktasındaki *subdiferansiyeli* denir ve

$$\partial f(x_0) = \{x^* : f(x) - f(x_o) \geqslant \langle x - x_0, x^* \rangle \ , \forall\, x \in \mathbb{R}^n\}$$

ile gösterilir. Eğer $\partial f(x_0)$ kümesi boştan farklı ise f fonksiyonuna x_0 noktasında subdiferansiyellenebilir denir.

f fonksiyonunun x_0 noktasındaki subdiferansiyeli $\partial f(x_0)$ kümesi kapalı konveks bir kümedir. Aşağıdaki lemma f konveks fonksiyonuna minimum değer veren nokta ile subdiferansiyel kavramının ilişkisini ortaya koyar.

Lemma 2.39. $x_0 \in \mathbb{R}^n$ noktasının bir has konveks fonksiyonuna minimum değer verebilmesi için gerek ve yeter koşul $0 \in \partial f(x_0)$ olmasıdır.

İspat: $x_0 \in \mathbb{R}^n$, f has konveks fonksiyon olsun. Her $x \in \mathbb{R}^n$ için $f(x) \geqslant f(x_0)$ olsun. O halde $f(x) - f(x_0) \geqslant 0 = \langle x - x_0, 0 \rangle$ olacağından $0 \in \partial f(x_0)$ dır. Tersine $0 \in \partial f(x_0)$ olsun. Tanım 2.17 gereği her $x \in \mathbb{R}^n$ için $f(x) - f(x_0) \geqslant \langle x - x_0, 0 \rangle = 0$ olduğundan f fonksiyonuna minimum değer veren nokta x_0 dır. ∎

Örnek 2.4. $f : \mathbb{R} \to \mathbb{R}$, $f(x) = |x|$ fonksiyonu $x_0 = 0$ noktasında diferansiyellenemez ama bu noktada subdiferansiyellenebilirdir. Açıkca subdiferansiyel kümesi

$$\partial f(0) = [-1, 1]$$

olarak bulunur.

Teorem 2.40. [14] $A \subseteq \mathbb{R}^n$ konveks küme üzerinde tanımlı f konveks fonksiyon ve $x \in int(A)$ olsun. O halde f fonksiyonunun x noktasında diferansiyellenebilir olması için gerek ve yeter koşul f fonksiyonunun x noktasında *tek bir* subgradient vektörüne sahip olmasıdır.

Tanım 2.18. $A \subseteq \mathbb{R}^n$ de bir küme ve $f : A \to \mathbb{R} \cup \{+\infty\}$ bir fonksiyon olmak üzere $\lim\limits_{n\to\infty} x_n = x_0$ gerçekleyen her (x_n) dizisi için

$$f(x_0) \leqslant \lim_{n\to\infty} \inf f(x_n)$$

eşitsizliği sağlanıyorsa f fonksiyonuna $x_0 \in A$ noktasında *alttan yarı sürekli* denir.

Teorem 2.41. f, $\mathbb{R}^n$ de herhangi bir fonksiyon olmak üzere aşağıdakiler denktir:

i) f, R^n de alttan yarı sürekli fonksiyondur.

ii) Her $\alpha \in \mathbb{R}$ için $L_\alpha(f) = \{x \in \mathbb{R}^n : \ f(x) \leqslant \alpha\}$ seviye kümesi kapalıdır.

iii) $epif = \{(x, \alpha) \in \mathbb{R}^{n+1} : \ f(x) \leqslant \alpha\}$ grafiküstü kümesi kapalıdır.

İspat: *i)* $\Rightarrow$ *ii)* f fonksiyonu, keyfi $x_0 \in \mathbb{R}^n$ noktasında alttan yarı sürekli olsun. $L_\alpha(f)$ seviye kümesinin kapalı olduğunu göstermek için $x_0 \in \overline{L_\alpha}(f)$ olsun. O halde öyle bir $(x_n)_{n\in\mathbb{N}} \subseteq L_\alpha(f)$ dizisi vardır öyle ki $x_n \to x_0$ dır. Her $n \in \mathbb{N}$ için $(x_n) \in L_\alpha(f)$ olduğundan $f(x_n) \leqslant \alpha$ dır. f fonksiyonu x_0 noktasında alttan yarı sürekli olduğundan $f(x_0) \leqslant \lim\limits_{n\to\infty} \inf f(x_n) \leqslant \alpha$ olur, yani $x_0 \in L_\alpha(f)$ olacağından $L_\alpha(f)$ seviye kümesi kapalıdır.
ii) $\Rightarrow$ *iii)* Her $\alpha \in \mathbb{R}$ için $L_\alpha(f)$ seviye kümesinin kapalı olsun. $(x_0, \alpha_0) \in \overline{epi}f$ olsun. O halde öyle $(x_n, \alpha_n)_{n\in\mathbb{N}} \subseteq epif$ vardır öyle ki her $n \in \mathbb{N}$ için $(x_n, \alpha_n) \to (x_0, \alpha_0)$ dır. Her $n \in \mathbb{N}$ için $f(x_n) \leqslant \alpha_n$ dir. Seviye kümesi tanımından $x_n \in L_{\alpha_n}(f)$ ve $L_{\alpha_n}(f)$ kapalı olduğundan $x_0 \in L_{\alpha_n}(f)$ tir. Yani $f(x_0) \leqslant \alpha_n$ dir. $n \to \infty$ için limit alınırsa $f(x_0) \leqslant \alpha_0$ olup bu ise $(x_0, \alpha_0) \in epif$ olduğunu gösterir. Böylece $epif$ kümesi kapalıdır.
iii) $\Rightarrow$ *ii)* $epif$, $\mathbb{R}^{n+1}$ de kapalı ve $x_0 \in \overline{L_\alpha}(f)$ olsun. O halde öyle $(x_n)_{n\in\mathbb{N}} \subseteq L_\alpha(f)$ vardır öyle ki her $n \in \mathbb{N}$ için $x_n \to x_0$ ve $f(x_n) \leqslant \alpha$ dır. $(x_n, \alpha) \in epif$ olup $epif$ kümesi kapalı olduğundan $(x_0, \alpha) \in epif$ olur. Yani $f(x_0) \leqslant \alpha$ olacağından $x_0 \in L_\alpha(f)$ dır. Böylece $L_\alpha(f)$ seviye kümesi kapalıdır.
ii) $\Rightarrow$ *i)* $L_\alpha(f)$ seviye kümesi kapalı ve her $n \in \mathbb{N}$ için $x_n \to x_0$ ve $f(x_n) \to \mu$ olsun. Her $\alpha > \mu$ için belli bir indisten sonra $f(x_n) \leqslant \alpha$ ve $L_\alpha(f)$ seviye kümesi kapalı olduğundan tüm limit noktasını içerir, yani $x_0 \in L_\alpha(f)$ dır. Böylece her $\alpha - \mu = \varepsilon > 0$ için $f(x_0) \leqslant \alpha$ yani her $\varepsilon > 0$ için $f(x_0) \leqslant \varepsilon + \mu$ olup her $\varepsilon > 0$ için sağlanacağından $f(x_0) \leqslant \mu$ olur. $f(x_n) \to \mu$ olduğundan $f(x_0) \leqslant \lim\limits_{n\to\infty} \inf f(x_n)$ olup f fonksiyonu keyfi $x_0 \in \mathbb{R}^n$ noktasında alttan yarı süreklidir. ∎

Tanım 2.19. f has fonksiyonu alttan yarı sürekli ya da epigraf kümesi kapalı veya denk olarak seviye kümesi kapalı ise f fonksiyonuna *kapalı fonksiyon* denir.

Tanım 2.20. A, $n \times n$ simetrik matrisine her $x \in \mathbb{R}^n$ için $x^T Ax > 0$ koşulunu sağlıyorsa *pozitif tanımlı matris* denir. Eğer $x \in \mathbb{R}^n$ için $x^T Ax \geqslant 0$ koşulunu sağlıyorsa *pozitif yarı tanımlı matris* denir. Benzer şekilde

$$H(x) = \Big(\frac{\delta^2 f}{\delta x_j \delta x_i}(x)\Big)$$

Hessian matrisinin asal minör değerleri negatif değil ise H matrisine *pozitif yarı tanımlı*, asal minör değerlerinin hepsi pozitif ise H matrisine *pozitif tanımlı* denir.

Teorem 2.42. [5] $A \subseteq \mathbb{R}^n$ konveks küme, $f : A \to \mathbb{R}$ fonksiyonu ikinci mertebeden sürekli kısmı türevlere sahip olsun ve Q, $n \times n$ reel simetrik matrisi verilsin.

i) Her $x \in A$ için $H(x)$ Hessian matrisi pozitif yarı tanımlı ise f, A kümesi üzerinde konveks fonksiyondur. Benzer şekilde her $x \in A$ için $H(x)$ Hessian matrisi pozitif tanımlı ise f, A kümesi üzerinde kesin konvekstir.

ii) $A = \mathbb{R}^n$ ve f konveks ise, her $x \in A$ için $H(x)$ Hessian matrisi pozitif yarı tanımlıdır.

iii) $f(x) = \langle x, Qx \rangle$ kuadratik fonksiyon ve Q simetrik matris olmak üzere f fonksiyonunun konveks olması için gerek ve yeter koşul Q simetrik matrisinin pozitif yarı tanımlı olmasıdır. Benzer şekilde f fonksiyonunun kesin konveks olması için gerek ve yeter koşul Q simetrik matrisinin pozitif tanımlı olmasıdır.

2.2.3. Eşlenik Fonksiyonlar

Konveks fonksiyonlarda ayrılabilme teoremlerinin uygulanmasının doğal bir sonucu olarak eşlenik fonksiyonlar ortaya çıkmaktadır. Örnek 2.3 'de A konveks kümesinin $x^* \in \mathbb{R}^n$ noktasındaki destek fonksiyonu $W_A(x^*) = \sup\limits_x\{\langle x, x^* \rangle : x \in A\}$ verilmişti. f kapalı konveks fonksiyonunun grafiküstü kümesinin destek fonksiyonunu hesaplayalım. $epif \in \mathbb{R}^{n+1}$ olduğu için

$$\begin{aligned} W_{epif}\big(x^*, \alpha^*\big) &= \sup_{(x,\alpha)}\{\langle (x,\alpha)\ ,\ (x^*,\alpha^*)\rangle\ :\ (x,\alpha) \in epif\} \\ &= \sup_{(x,\alpha)}\{\langle x, x^* \rangle + \alpha\alpha^* :\ (x,\alpha) \in epif\} \end{aligned}$$

bulunur. Eğer $\alpha^* > 0$ ise verilen x ve α noktası istenildiği kadar büyük olabileceğinden $W_{epif}\big(x^*, \alpha^*\big) = +\infty$ olur. Eğer $\alpha^* \leqslant 0$ ise f kapalı fonksiyon olduğu için $W_{epif}(x^*, \alpha^*)$ fonksiyonunu hesaplamak yerine $W_{epif}(x^*, -1)$ fonksiyonunu hesaplamak yeterlidir.

Eğer $\alpha^* = -1$ ise

$$\begin{aligned} W_{epif}(x^*, -1) &= \sup_{(x,\alpha)} \{\langle x, x^* \rangle - \alpha : \ f(x) \leqslant \alpha\} \\ &= \sup_x \{\langle x, x^* \rangle - f(x)\} \end{aligned}$$

bulunur.

Tanım 2.21. $f : \mathbb{R}^n \to \mathbb{R}$ bir fonksiyon olmak üzere $f^*(x^*) = \sup\limits_x \{\langle x, x^* \rangle - f(x)\}$ şeklinde tanımlanan $f^* : \mathbb{R}^n \to \mathbb{R}$ fonksiyonuna f fonksiyonunun *eşlenik fonksiyonu* (=*konveks eşlenik fonksiyonu*) denir.

Lemma 2.43. $f^* : \mathbb{R}^n \to \mathbb{R}$ eşlenik fonksiyonu daima kapalı ve konvekstir.

İspat: $x_1^*, x_2^* \in \mathbb{R}^n, \ \lambda_1, \lambda_2 \geqslant 0, \ \lambda_1 + \lambda_2 = 1$ için

$$\begin{aligned} f^*(\lambda_1 x_1^* + \lambda_2 x_2^*) &= \sup_x \{\langle x, \lambda_1 x_1^* + \lambda_2 x_2^* \rangle - f(x)\} \\ &= \sup_x \{\lambda_1 \langle x, x_1^* \rangle + \lambda_2 \langle x, x_2^* \rangle - f(x)\} \\ &\leqslant \lambda_1 \sup_x \{\langle x, x_1^* \rangle - f(x)\} + \lambda_2 \sup_x \{\langle x, x_2^* \rangle - f(x)\} \\ &= \lambda_1 f^*(x_1^*) + \lambda_2 f^*(x_2^*) \end{aligned}$$

olacağından f^* eşlenik fonksiyonu konvekstir. f^* eşlenik fonksiyonunun kapalı olduğunu göstermek için $L_\alpha(f^*) = \{x^* \in \mathbb{R}^n : \ f^*(x^*) \leqslant \alpha\}$ seviye kümesinin kapalılığını gösterelim. $\alpha \in \mathbb{R}, \ x_0^* \in \overline{L}_\alpha(f^*)$ olsun. Bu durumda öyle bir $(x_n^*)_{n \in \mathbb{N}} \subseteq L_\alpha(f^*)$ dizisi vardır öyle ki $(x_n^*) \to x_0^*$ dır ve her $n \in \mathbb{N}$ için $f^*(x_n^*) \leqslant \alpha$ dır. Tanım 2.21 gereği $\sup\limits_x \{\langle x, x_n^* \rangle - f(x)\} \leqslant \alpha$ olur. Her $x \in \mathbb{R}^n$ için $\langle x, x_n^* \rangle - f(x) \leqslant \alpha$ dır ve limit alınırsa her x için $f^*(x_0^*) \leqslant \alpha$ olur. Böylece $x_0^* \in L_\alpha(f^*)$ olur. Dolayısıyla $L_\alpha(f^*)$ seviye kümesi kapalı olduğu için f^* eşlenik fonksiyonu kapalıdır. ∎

Örneğin, $f(x) = \frac{1}{2}\, x^2, \ x \in \mathbb{R}$ konveks fonksiyonunu göz önüne alalım. Tanım 2.21 kullanılarak

$$f^*(x^*) = \sup_x \big\{x \cdot x^* - \frac{1}{2}\, x^2\big\} = \sup_x \big\{\frac{1}{2}(x^*)^2 - \frac{1}{2}(x - x^*)^2\big\}$$

$x^* \in \mathbb{R}$ için $f^*(x^*) = \frac{1}{2}\, (x^*)^2$ bulunur.

Tanım 2.21 kullanarak elde edilen

$$f(x) + f^*(x^*) \geqslant \langle x, x^* \rangle \tag{2.20}$$

eşitsizliğine eğer f konveks ve has fonksiyon ise Fenchel Eşitsizliği denir. Eşlenik fonksiyonların temel özellikleri aşağıdaki lemma ile verilmektedir.

Lemma 2.44. [13] *i)* Her $x \in \mathbb{R}^n$ için $f_1(x) \leqslant f_2(x)$ ise $f_2^*(x^*) \leqslant f_1^*(x^*)$ dır.

ii) $\alpha \in \mathbb{R}$ için $g(x) = f(x) + \alpha$ ise $g^*(x^*) = f^*(x^*) - \alpha$ dır.

iii) $t > 0$ için $g(x) = t.f(x)$ ise $g^*(x^*) = t.f^*\left(\frac{x^*}{t}\right)$ dır.

iv) $t \neq 0$ için $g(x) = f(tx)$ ise $g^*(x^*) = f^*\left(\frac{x^*}{t}\right)$ dır.

Örnek 2.5. $f(x) = e^x$ konveks fonksiyonu için $x^* = 0$ ise $f^*(x^*) = 0$ dır. $x^* < 0$ ise x değeri çok küçük negatif sayı seçilirse $f^*(x^*) = +\infty$ bulunur. $x^* > 0$ iken, $\langle x, x^* \rangle - e^x$ fonksiyonu $x = \log x^*$ noktasında maksimum değerini alır, böylece $f^*(x^*) = x^* \log x^* - x^*$ dır. Açıkca

$$f^*(x^*) = \begin{cases} x^* \log x^* - x^* & , \ x^* > 0 \\ 0 & , \ x^* = 0 \\ +\infty & , \ x^* < 0 \end{cases}$$

elde edilir.

$C \subseteq \mathbb{R}^n$ kümesinin $\delta(x|C)$ indikatör fonksiyonunun eşleniği

$$f^*(x^*) = \sup_x \{\langle x, x^* \rangle - \delta(x|C)\} = \sup_x \{\langle x, x^* \rangle \ : \ x \in C\} = W_C(x^*)$$

dır. Yani C kümesinin indikatör fonksiyonunun eşleniği, C kümesinin destek fonksiyonudur. Lemma 2.43 nedeniyle destek fonksiyonunun daima kapalı olduğu sonucu elde edilir.

Farklı bir eşlenik fonksiyon örneği olarak L alt uzayının indikatör fonksiyonu

$$f(x) = \delta(x|L) = \begin{cases} 0 & , \ x \in L \\ +\infty & , \ x \notin L \end{cases}$$

olmak üzere f fonksiyonunun eşleniği

$$f^*(x^*) = \sup_x \{\langle x, x^* \rangle - f(x)\} = \sup_x \{\langle x, x^* \rangle : \ x \in L\}$$

dır. Her $x \in L$ için $\langle x, x^* \rangle = 0$ ise $f^*(x^*) = 0$ ve her $x \in L$ için $\langle x, x^* \rangle \neq 0$ ise $f^*(x^*) = +\infty$ dır. L alt uzayının ortogonal tümleyeni

$$L^\perp = \{x^* : x \perp x^* \ , \ \forall \, x \in L\} = \{x^* : \langle x, x^* \rangle = 0 \ , \ \forall \, x \in L \ \}$$

olarak tanımlandığı için

$$f^*(x^*) = \begin{cases} 0 & , \ x^* \in L^\perp \\ +\infty & , \ x^* \notin L^\perp \end{cases}$$

bulunur. Ayrıca L alt uzayının ortogonal tümleyeninin indikatör fonksiyonu

$$\delta(x^*|L^\perp) = \begin{cases} 0 & , x^* \in L^\perp \\ +\infty & , x^* \notin L^\perp \end{cases}$$

olduğundan $f^*(x^*) = \delta(x^*|L^\perp)$ olur. Böylece f fonksiyonunun yani L alt uzayının indikatör fonksiyonunun eşlenik fonksiyonu, L alt uzayının ortogonal tümleyeninin indikatör fonksiyonudur.

Tanım 2.22. $g : \mathbb{R}^n \to \mathbb{R}$ konkav fonksiyonunun $g_* : \mathbb{R}^n \to \mathbb{R}$ *eşlenik konkav fonksiyonu*

$$g_*(x^*) = \inf_x \{\langle x, x^* \rangle - g(x)\} \tag{2.21}$$

olarak tanımlanır. Konkav eşlenik fonksiyonu ile konveks eşlenik fonksiyonu arasındaki ilişki

$$g_*(x^*) = -(-g)^*(-x^*) \tag{2.22}$$

şeklinde verilir.

Teorem 2.45. $f : \mathbb{R}^n \to \mathbb{R}$ kapalı ve konveks fonksiyon olmak üzere $x_0 \in dom f$ için

$$\langle x_0, x^* \rangle = f(x_0) + f^*(x^*) \Longleftrightarrow x^* \in \partial f(x_0) \tag{2.23}$$

bağıntısı gerçeklenir.

İspat: f kapalı, konveks fonksiyon ve $\langle x_0, x^* \rangle = f(x_0) + f^*(x^*)$ eşitliği sağlansın. Her $x \in \mathbb{R}^n$ için $\langle x_0, x^* \rangle - f(x_0) = f^*(x^*) = \sup_x \{\langle x, x^* \rangle - f(x)\}$ yani, $\langle x_0, x^* \rangle - f(x_0) \geqslant \langle x, x^* \rangle - f(x)$ olur. Bu eşitlik ise $x^* \in \partial f(x_0)$ olduğunu gösterir. Tersine $x^* \in \partial f(x_0)$ olsun. Subdiferansiyel tanımından her $x \in \mathbb{R}^n$ için $\langle x_0, x^* \rangle - f(x_0) \geqslant \langle x, x^* \rangle - f(x)$ sağlanacağından $\langle x_0, x^* \rangle - f(x_0) \geqslant f^*(x^*)$ olur. Açıkça $\langle x_0, x^* \rangle \geqslant f(x_0) + f^*(x^*)$ bağıntısı elde edilir. Ayrıca (2.20) Fenchel eşitsizliği de kullanılarak $\langle x_0, x^* \rangle = f(x_0) + f^*(x^*)$ elde edilir. ■

f^* eşlenik fonksiyonunun da eşlenik fonksiyonundan bahsetmek mümkündür. İkincil eşlenik fonksiyon denilen ve f^{**} ile gösterilen bu fonksiyon da daima konvekstir ve aşağıdaki gibi tanımlanır.

Tanım 2.23. $f^{**} : \mathbb{R}^n \to \mathbb{R}$ olmak üzere $f^{**}(x) = \sup_{x^*} \{\langle x^*, x \rangle - f^*(x^*)\}$ fonksiyonuna f fonksiyonunun *ikincil eşlenik fonksiyonu* denir.

Tanım 2.21 'den her $x \in \mathbb{R}^n$ için $f(x) \geqslant \langle x, x^* \rangle - f^*(x^*)$ olduğundan x^*'lar üzerinden supremum alınırsa

$$f(x) \geqslant f^{**}(x) \tag{2.24}$$

eşitsizliği bulunur.

Teorem 2.46. $f : \mathbb{R}^n \to \mathbb{R}$, $f^{**} : \mathbb{R}^n \to \mathbb{R}$ fonksiyonları ve her $x \in \mathbb{R}^n$ için $\partial f(x_0) \neq \emptyset$ ise $f(x_0) = f^{**}(x_0)$ eşitliği gerçeklenir.

İspat: Kabul edelim ki $\partial f(x_0) \neq \emptyset$ olsun. O halde öyle $x^* \in \partial f(x_0)$ vardır ki her $x \in \mathbb{R}^n$ için $\langle x_0, x^* \rangle - f(x_0) \geqslant \langle x, x^* \rangle - f(x)$ olup bu eşitsizlikte x'ler üzerinden supremum alınırsa

$$\langle x_0, x^* \rangle - f(x_0) \geqslant f^*(x^*) \tag{2.25}$$

olur. Ayrıca $f^{**}(x)$ ikincil eşlenik fonksiyonu tanımından $f^{**}(x) \geqslant \langle x^*, x \rangle - f^*(x^*)$ eşitsizliği her $x \in \mathbb{R}^n$ için sağlanacağından

$$f^{**}(x_0) \geqslant \langle x^*, x_0 \rangle - f^*(x^*) \tag{2.26}$$

olup (2.25) eşitsizliğinden $f^{**}(x_0) \geqslant f(x_0)$ bulunur ve (2.24) eşitsizliği daima geçerli olduğundan $f(x_0) = f^{**}(x_0)$ elde edilir. ∎

$f : \mathbb{R}^n \to \mathbb{R}$ fonksiyonunun eşlenik fonksiyonu daima kapalı ve konveks fonksiyon olduğu Lemma 2.43 'de verilmişti. Her $x \in \mathbb{R}^n$ için $f^{**}(x) = f(x)$ olabilmesi için f fonksiyonunun kapalı ve konveks olması gerekmektedir. Her $x \in \mathbb{R}^n$ için $f^{**}(x) = f(x)$ bağıntısının sağlandığını göstermek için bazı önemli tanımlamalar aşağıda verilmektedir.

Tanım 2.24. Herhangi bir f fonksiyonunun grafiküstü kümesinin konveks zarfı başka bir fonksiyonun grafiküstü kümesini verecektir. İşte bu fonksiyona f fonksiyonunun *konveks zarf fonksiyonu* denir ve cof ile gösterilir. Açıkca;

$$cof(x) = \sup_x \{g(x) : \forall\, x \in \mathbb{R}^n \text{ için } g(x) \leqslant f(x)\}$$

dır. Burada ki $g(x)$ fonksiyonu $\mathbb{R}^n$ üzerinde tanımlı has, konveks fonksiyondur. Benzer şekilde f fonksiyonunun *kapalı konveks zarf fonksiyonu* $\overline{co}f(x)$ ile gösterilir ve $h(x)$ fonksiyonu $\mathbb{R}^n$ üzerinde tanımlı kapalı, konveks fonksiyon ise

$$\overline{co}f(x) = \sup_x \{h(x) : \forall\, x \in \mathbb{R}^n \text{ için } h(x) \leqslant f(x)\}$$

şeklinde tanımlanır. Eşdeğer bir tanım ise $b \in \mathbb{R}$ olmak üzere

$$\overline{co}f(x) = \sup_{x^*}\{\langle x, x^*\rangle - b : \forall\, x \in \mathbb{R}^n \text{ için } \langle x, x^*\rangle - b \leqslant f(x)\}$$

olarak verilir.

Tanım 2.25. f has fonksiyonu ve $x \in \mathbb{R}^n$ için

$$clf(x) = \lim_{y \to x} \inf f(y)$$

biçiminde tanımlı $clf(x)$ fonksiyonuna f fonksiyonunun *kapanış fonksiyonu* denir. Eşdeğer olarak $epi(clf) = \overline{epi}f$ koşulunu sağlayan clf fonksiyonuna f fonksiyonunun kapanış fonksiyonu denir.

Teorem 2.47. f herhangi bir fonksiyon ise her $x \in \mathbb{R}^n$ için $f^{**}(x) = \overline{co}f(x)$ dır.

İspat: Kapalı zarf fonksiyonu ve eşlenik fonksiyonu tanımlarından

$$\begin{aligned}
\overline{co}f(x) &= \sup_{x^*}\{\langle x, x^*\rangle - b : \forall\, x \in \mathbb{R}^n \text{ için } -b \leqslant f(x) - \langle x, x^*\rangle\} \\
&= \sup_{x^*}\{\langle x, x^*\rangle - b : \ -b \leqslant -f^*(x^*)\} \\
&\leqslant \sup_{x^*}\{\langle x, x^*\rangle - f^*(x^*)\} = f^{**}(x)
\end{aligned}$$

eşitsizliği bulunur. Tersine her $x \in \mathbb{R}^n$ için $f^{**}(x) \leqslant f(x)$ eşitsizliği geçerliydi. $\overline{co}f(x)$ kapalı konveks zarf fonksiyonu tanımından ve $f^{**}(x)$ fonksiyonunun kapalı ve konveks olduğu gerçeği göz önüne alınırsa her $x \in \mathbb{R}^n$ için $f^{**}(x) \leqslant \overline{co}f(x)$ olur. Böylece herhangi bir fonksiyon için $f^{**}(x) = \overline{co}f(x)$ elde edilir. ∎

Sonuç 2.2. $i)$ f herhangi bir fonksiyon ise $f^{**} = \overline{co}f$ dir.

$ii)$ f konveks fonksiyon ise $f^{**} = clf$ dir.

$iii)$ f kapalı ve konveks ise $f^{**} = f$ dir.

Teorem 2.48. f, g herhangi iki fonksiyon olmak üzere her $x, x^* \in \mathbb{R}^n$ için

$$\overline{co}f(x) \leqslant g(x) \leqslant f(x) \Rightarrow f^*(x^*) = g^*(x^*)$$

gerçeklenir.

İspat: $\overline{co}f(x) \leqslant g(x) \leqslant f(x)$ olsun. Teorem 2.47 nedeniyle $f^{**}(x) \leqslant g(x) \leqslant f(x)$ olur ve bu eşitsizlikte Lemma 2.44 dikkate alınıp eşlenik alınırsa $f^*(x^*) \leqslant g^*(x^*) \leqslant (f^*)^{**}(x^*)$ elde edilir. $f^*(x^*)$ kapalı ve konveks fonksiyon olduğundan $f^*(x^*) \leqslant g^*(x^*) \leqslant f^*(x^*)$ yani $f^*(x^*) = g^*(x^*)$ dır. ∎

Teorem 2.49. f fonksiyonu pozitif homojen kapalı konveks fonksiyon ise f^* eşlenik fonksiyonunun tanım kümesinin indikatör fonksiyonu için $f^*(x^*) = \delta_{domf^*}(x^*)$ eşitliği gerçeklenir.

İspat: f pozitif homojen kapalı konveks fonksiyon için daima $f(0) = 0$ olduğundan ve Fenchel eşitsizliğinden $f^*(x^*) = \sup\limits_x\{\langle x, x^*\rangle - f(x)\} \geqslant 0$ dır. Öyle bir $x_1 \in \mathbb{R}^n$ noktası için $\langle x_1, x^*\rangle - f(x_1) > 0$ gerçeklensin. O halde

$$f^*(x^*) \geqslant \sup_{\lambda>0}\{\langle \lambda x_1, x^*\rangle - f(\lambda x_1)\} = \sup_{\lambda>0} \lambda\{\langle x_1, x^*\rangle - f(x_1)\} = +\infty$$

elde edilir. Böylece

$$f^*(x^*) = \delta_{domf^*}(x^*) = \begin{cases} 0 & , \ x^* \in domf^* \\ +\infty & , \ x^* \notin domf^* \end{cases}$$

olup ispat tamamlanır. ∎

Teorem 2.50. f pozitif homojen kapalı konveks fonksiyon olsun. Bu durumda

$$f(x) = \sup_{x^*}\{\langle x, x^*\rangle \ : \ x^* \in domf^*\}$$

gerçeklenir.

İspat: f kapalı konveks fonksiyon olduğundan $f^{**}(x) = f(x)$ dır. Teorem 2.49'den

$$f(x) = f^{**}(x) = \sup_{x^*}\{\langle x, x^*\rangle - f^*(x^*)\} = \sup_{x^*}\{\langle x, x^*\rangle \ : \ x^* \in domf^*\}$$

olur, böylece istenen elde edilir. ∎

Tanım 2.26. $f, g : \mathbb{R}^n \to \mathbb{R}$ has fonksiyon olsun. $f\ \nabla g : \mathbb{R}^n \to \mathbb{R}$ olmak üzere

$$(f\ \nabla g)(x) = \inf_y\{f(x-y) + g(y)\}$$

biçiminde tanımlanan fonksiyona f ile g fonksiyonlarının *infimal konvolüsyonu* denir.

Uyarı 2.2. f ve g has fonksiyon olmasına rağmen $f\ \nabla g$ infimal konvolüsyon has fonksiyon olmayabilir. Örneğin; $f(x-y) = x-y$ ve $g(y) = -y$ ise $(f\ \nabla g)(x) = -\infty$ olacağından $f\ \nabla g$ infimal konvolüsyon has fonksiyon değildir.

f ve g konveks fonksiyonlar ise $\lambda_1, \lambda_2 \geqslant 0,\ \lambda_1 + \lambda_2 = 1$ olmak üzere $x_1, x_2 \in R^n$ için

$$\begin{aligned}
(f\ \nabla g)(\lambda_1 x_1 + \lambda_2 x_2) &= \inf_y\Big\{f(\lambda_1 x_1 + \lambda_2 x_2 - y) + g(y)\Big\} \\
&\leqslant \inf_y\Big\{\lambda_1\big(f(x_1 - y) + g(y)\big) + \lambda_2\big(f(x_2 - y) + g(y)\big)\Big\} \\
&= \lambda_1\big(f\ \nabla\ g\big)(x_1) + \lambda_2\big(f\ \nabla\ g\big)(x_2)
\end{aligned}$$

olduğundan $f\ \nabla g$ konvekstir. Böylece $f\ \nabla g$ infimal konvolüsyon konveksliği korur. Fakat kapalılık özelliğini koruması gerekmez. Örneğin C ve D kapalı kümeler için $f(x) = \delta(x|C)$ ve $g(x) = \delta(x|D)$ kapalı fonksiyonlardır. Ama $(f\nabla g)(x) = \delta(x|C{+}D)$ olup kapalı kümelerin toplamının kapalı olması gerekmediğinden $(f\ \nabla g)(x)$ infimal konvolüsyonu kapalı olmayabilir.

Teorem 2.51. $f, g : \mathbb{R}^n \to \mathbb{R}$ konveks has fonksiyonlar ise $f^* + g^* = (f\ \nabla\ g)^*$ gerçeklenir.

İspat: f ve g konveks has fonksiyonu ise

$$\begin{aligned}(f\ \nabla\ g)^* &= \sup_x \left\{\langle x, x^*\rangle - \inf_y\{f(x-y) + g(y)\}\right\} \\ &= \sup_x \left\{\langle x, x^*\rangle + \sup_y\{-f(x-y) - g(y)\}\right\}\end{aligned}$$

$y = x_2$ ve $x - y = x_1$ seçilip düzenleme yapılırsa

$$(f\ \nabla\ g)^* = \sup_{x=x_1+x_2} \left\{\langle x_1 + x_2,\ x^*\rangle - f(x_1) - g(x_2)\right\} = f^*(x^*) + g^*(x^*)$$

bulunur. ∎

Teorem 2.52. *i)* f ve g kapalı konveks has fonksiyonlar ise $(f+g)^* = cl(f^*\ \nabla g^*)$ eşitliği gerçeklenir.

ii) $(f+g)^* \leqslant f^*\ \nabla g^*$ dır.

İspat: *i)* f ve g kapalı konveks fonksiyon olduğundan Teorem 2.51 nedeniyle $(f^*\ \nabla\ g^*)^* = f + g$ elde edilir. Bu eşitlikte her iki tarafın eşleniği alınırsa ve Sonuç 2.2'den $(f+g)^* = (f^*\ \nabla\ g^*)^{**} = cl(f^*\ \nabla\ g^*)$ gerçeklenir.
ii) $f^*\ \nabla g^*$ konveks olduğundan $(f+g)^* = (f^*\ \nabla\ g^*)^{**}$ dır. Daima $f^{**} \leqslant f$ eşitsizliği geçerli olduğundan $(f+g)^* \leqslant f^*\ \nabla\ g^*$ dır. ∎

Tanım 2.27. $\emptyset \neq A \subseteq \mathbb{R}^n$, $f : A \to \mathbb{R}$ herhangi bir fonksiyon ve $a \in A$ olsun. $\exists\ \varepsilon > 0$, $\forall\ x \in B(a,\varepsilon) \cap A$ için $f(a) \leqslant f(x)$ ise a noktasına f fonksiyonunun *yerel minimum noktası* denir. Benzer şekilde $\forall\ x \in B(a,\varepsilon) \cap A$ için $f(x) \leqslant f(a)$ ise a noktasına f fonksiyonunun *yerel maksimum noktası* denir.

Tanım 2.28. $a \in A$, her $x \in A$ için $f(a) \leqslant f(x)$ gerçekleniyorsa $a \in A$ noktasına *minimum nokta* denir. Her $x \in A$ için $f(x) \leqslant f(a)$ ise $a \in A$ noktasına *maksimum nokta* denir.

Konvekslik kavramı altında yerel minimum nokta ile minimum nokta arasında fark olmadığı aşağıdaki lemma ile gösterilmektedir.

Lemma 2.53. A konveks küme üzerinde tanımlı f konveks fonksiyon için aşağıdakiler gerçeklenir.

i) f, x_0 noktasında yerel minimuma erişirse, f fonksiyonu x_0 noktasında minimuma ulaşır.

ii) f fonksiyonunun minimuma eriştidiği noktalar kümesi ya $\emptyset$ ya da konvekstir.

iii) Eğer f kesin konveks fonksiyonu bir x_* noktasında minimuma ulaşırsa, x_* noktası tektir.

iv) Eğer f sabit fonksiyon değil ve bir maksimum değerine $x \in A$ noktasında ulaşırsa x noktası A kümesinin sınırda olmalıdır.

İspat: *i)* Kabul edelim ki f fonksiyonu, x_0 noktasında yerel minimuma ulaşsın. Öyle $\exists\ \varepsilon > 0$, $\forall\ z \in B(x_0, \varepsilon) \cap A$ için $f(x_0) \leqslant f(z)$ dir. Keyfi $y \in A$ noktası için $0 < t < \dfrac{\varepsilon}{||y - x_0||}$ ve $z = x_0 + t(y - x_0)$ şeklinde seçilirse $z \in B(x_0, \varepsilon) \cap A$ olur ve

$$0 \leqslant \frac{f(z) - f(x_0)}{t} \leqslant f(y) - f(x_0) \tag{2.27}$$

olduğundan f fonksiyonu x_0 noktasında minimuma ulaşır.
ii) f fonksiyonu minimuma tek bir noktada ulaşırsa tek nokta kümesi konveks olduğundan istenen gerçeklenir. Kabul edelim ki f minimuma iki farklı x_1, x_2 noktalarında ulaşsın. Yani, $\mu = f(x_1) = f(x_2) = \min\{f(x) : x \in A\}$ olsun. $\alpha > 0$ olmak üzere $\mu \leqslant f(\alpha x_1 + (1-\alpha)x_2) \leqslant \alpha f(x_1) + (1-\alpha) f(x_2) = \mu$ olur. Dolayısıyla $[x_1, x_2]$ aralığının her bir noktasında f fonksiyonu minimuma ulaşılır.
iii) f kesin konveks fonksiyonu x_1 ve x_2 noktalarında minimuma ulaşsın. Yani $\mu = f(x_1) = f(x_2) = \min\{f(x) : x \in A\}$ olsun ve $\alpha, \beta > 0$, $\alpha + \beta = 1$ için $\mu \leqslant f(\alpha x_1 + \beta x_2) < \alpha f(x_1) + \beta f(x_2) = \mu$ çelişkisi elde edileceğinden f fonksiyonu kesin konveks ise tek bir noktada minimuma ulaşır.
iv) f fonksiyonu x noktasında maksimuma ulaşsın. Eğer x sınır noktası değil ise iç nokta olacağından ve f sabit fonksiyon olmadığından öyle $y \in A$ vardır ki $f(y) < f(x)$ dir. A konveks küme olduğundan $[x, y] \in A$ dır. x iç nokta olduğundan $x \in [z, y]$ olacak şekilde $[z, y]$ doğru parçası vardır. $\alpha + \beta = 1$, $\alpha, \beta > 0$ için $x = \alpha z + \beta y$ dir ve $f(x) = f(\alpha z + \beta y) \leqslant \alpha f(z) + \beta f(y) < \alpha f(x) + \beta f(x) = f(x)$ çelişkisinden dolayı x iç nokta değildir. O halde x, f fonksiyonunun sınır noktasıdır.■

2.3. BAZI OPTİMİZASYON PROBLEMLERİ

Bu kısımda diferansiyellenebilir kısıtsız problemler, lineer programlama problemleri, kuadratik programlama problemleri ve konveks programlama problemleri en genel anlamda tanıtılmış ve bu problemlerin yapıları incelenmiştir. Kısıtları bir eşitsizlik sistemi ile verilen optimizasyon problemi $X \subseteq \mathbb{R}^n$, $f : X \to \mathbb{R}$ ve $g : X \to \mathbb{R}^r$ fonksiyon olmak üzere

$$\begin{aligned} \min\ & f(x) \\ & g(x) \leqslant 0 \\ & x \in X \end{aligned} \tag{2.28}$$

şeklinde verilir. Bu problemde $f(x)$ fonksiyonuna *amaç fonksiyonu*, X kümesine *kısıt kümesi* ve $X_0 = \{x : x \in X\ ,\ g(x) \leqslant 0\}$ kümesine ise *uygun çözümler kümesi* denir. (2.28) probleminde X kısıt kümesi üzerinde tanımlı reel değerli f fonksiyonu verildiğinde her $x \in X_0$ için $f(x_*) \leqslant f(x)$ koşulunu gerçekleyen $x_* \in X_0$ noktasının varlığı araştırılıp, bu x_* noktasınının bulunması amaçlanır. Diğer yandan (2.28) problemi ile

$$\begin{aligned} \text{maks}\ & -f(x) \\ & g(x) \leqslant 0 \\ & x \in X \end{aligned} \tag{2.29}$$

problemi aynı olduğundan genelliği bozmaksızın minimum problemler için çözümler araştırılmaktadır.

2.3.1. Diferansiyellenebilir Kısıtsız Problemler

$I \subseteq \mathbb{R}$ açık aralık, f fonksiyonu bu aralık üzerinde diferansiyellenebilir ve $x_0 \in I$ yerel minimum noktası olsun. O halde öyle $\delta > 0$ vardır ki her $x \in B(x_0, \delta)$ için $f(x_0) \leqslant f(x)$ dır. Ayrıca $x_0 < x < x_0 + \delta$ için $\frac{f(x)-f(x_0)}{x-x_0} \geqslant 0$ ve $x_0 - \delta < x < x_0$ için $\frac{f(x)-f(x_0)}{x-x_0} \leqslant 0$ dır. f fonksiyonu x_0 noktasında diferansiyellenebilir olduğundan

$$f'_+(x_0) = \lim_{x \to x_0^+} \frac{f(x) - f(x_0)}{x - x_0} \geqslant 0$$

$$f'_-(x_0) = \lim_{x \to x_0^-} \frac{f(x) - f(x_0)}{x - x_0} \leqslant 0$$

olup $f'(x_0) = 0$ dır. Böylece x_0 noktası yerel minimum noktası ise $f'(x_0) = 0$ elde edilir.

Kabul edelim ki f fonksiyonu $(x_0 - \delta, x_0 + \delta)$ aralığında ikinci mertebeden sürekli kısmi türevlere sahip olsun. Taylor teoreminden her $x \in \big((x_0 - \delta, x_0 + \delta)\backslash\{x_0\}\big)$ için öyle bir $\tilde{x} \in (x_0, x)$ noktası vardır öyle ki

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{1}{2}f''(\tilde{x})(x - x_0)^2$$

dir. x_0 yerel minimum nokta ise $f'(x_0) = 0$ olduğundan $f''(\tilde{x}) = 2\frac{f(x)-f(x_0)}{(x-x_0)^2} \geqslant 0$ olur. Eğer $x \to x_0$ alınırsa o halde $\tilde{x} \to x_0$ ve $f''(\tilde{x}) \to f''(x_0)$ elde edilir. Böylece $f''(x_0) \geqslant 0$ dır. Dolayısıyla aşağıdaki teorem ispatlanmış olur.

Teorem 2.54. $I \subseteq \mathbb{R}^n$ açık aralık, I üzerinde tanımlı reel değerli fonksiyon f ve $x_0 \in I$ yerel minimum noktası olsun. Eğer f fonksiyonu x_0 noktasında diferansiyellenebilir ise $f'(x_0) = 0$ dır. Eğer f, ikinci merteben sürekli kısmi türevlere sahip ise $f''(x_0) \geqslant 0$ dır.

$c \in X$ için $f'(c) = 0$ ise c noktasına f fonksiyonunun *kritik noktası* denir. $X \subseteq \mathbb{R}^n$ uzayında açık küme olsun, buna göre aşağıda Teorem 2.54 'ün n boyuta genelleştirilmiş hali verilmektedir.

Teorem 2.55. [5] $X \subseteq \mathbb{R}^n$ açık küme, $f : X \to \mathbb{R}$ fonksiyonunun yerel minimum noktası x_0 olsun. Eğer f fonksiyonu x_0 noktasında birinci mertebeden kısmi türevlere sahip ise $\nabla f(x_0) = (\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n})(x_0) = 0$ dır. Eğer f, $B(x_0, \delta)$ açık yuvarında ikinci mertebeden sürekli kısmi türevlere sahip ise $j, i = 1, \cdots, n$ için

$$H(x_0) = \Big(\frac{\partial^2 f}{\partial x_j \partial x_i}(x_0)\Big)$$

Hessian matrisi pozitif yarı tanımlıdır.

Sonuç 2.3. C açık konveks kümesi üzerinde f konveks fonksiyonu diferansiyellenebilir olsun. O halde f fonksiyonu x_* noktasında minimuma ulaşması için gerek ve yeter koşul $\nabla f(x_*) = 0$ olmasıdır.

İspat: f fonksiyonu x_* noktasında minimuma ulaşsın. f fonksiyonu ve C konveks olmasa bile Teorem 2.55 nedeniyle $\nabla f(x_*) = 0$ dır. Tersine C açık konveks küme üzerinde f diferansiyellenebilir konveks fonksiyon ve $\nabla f(x_*) = 0$ olsun. Lemma 2.39 nedeniyle

$$\{0\} = \{\nabla f(x_*)\} = \partial f(x_*)$$

dır. Dolayısıyla x_* noktasında f minimuma ulaşır. ■

2.3.2. Lineer Programlama Problemi

Kısıtlı optimizasyon problemlerinde, amaç fonksiyonunun belirli eşitlik ve eşitsizlik kısıtları altında minimum ya da maksimum değeri aranır. Amaç ve kısıt fonksiyonlarının hepsi lineer ise kısıtlı optimizasyon problemi lineer programlama problemi olarak tanımlanır. Amaç fonksiyonu optimum değerine kısıtlar tarafından oluşturulmuş uygun bölgede ulaşır. A, $m \times n$ reel matris olmak üzere $x,\ c \in \mathbb{R}^n$ ve $b \in \mathbb{R}^m$ için

$$\begin{aligned} \min\ & c^T x \\ & Ax = b \\ & x \geqslant 0 \end{aligned} \tag{2.30}$$

problemi bir lineer programlama problemidir. Lineer programlama probleminin uygun çözümler kümesi $X_0 = \{x : Ax = b\ ,\ x \geqslant 0\}$ konveks bir kümedir. Uygun çözümler kümesi için aşağıdaki gözlemleri yapmak zor değildir:

i) İki boyutlu uzayda bir lineer programlama probleminin uygun çözümler kümesi sonlu sayıda doğrularla sınırlanan konveks bir düzlemsel bölge oluşturur.

ii) Üç boyutlu uzayda uygun çözümler kümesi sonlu sayıda düzlemle sınırlanan konveks bölgedir.

iii) $n > 3$, n-boyutlu uzayda ise uygun çözümler kümesi hiperdüzlemlerle sınırlanan polihedral konveks bir kümedir.

2.3.3. Kuadratik Programlama Problemi

Q, $n \times n$ pozitif yarı tanımlı bir matris ve A, $m \times n$ reel matris olmak üzere $b \in \mathbb{R}^m$, $c \in \mathbb{R}^n$ ve $f(x) = \frac{1}{2}\langle x, Qx\rangle + \langle c, x\rangle$ için

$$\begin{aligned} \min\ & \frac{1}{2}\langle x, Qx\rangle + \langle c, x\rangle \\ & Ax \leqslant b \\ & x \geqslant 0 \end{aligned} \tag{2.31}$$

şeklinde verilen problem Kuadratik Programlama Problemi olarak adlandırılır.

Teorem 2.42 nedeniyle Q pozitif tanımlı bir matris ise f fonksiyonu kesin konveks ve uygun çözümler kümesi $X_0 = \{x : Ax \leqslant b\ ,\ x \geqslant 0\}$ konveks bir küme olduğundan $f(x)$ fonksiyonu minimuma tek bir x_* noktasında ulaşır. Eğer Q matrisi sıfır matrisi ise kuadratik programlama problemi, lineer programlama problemine dönüşür.

2.3.4. Konveks Programlama Problemi

Genel olarak kısıtları bir eşitsizlik sistemi ile verilmiş konveks programlama problemi aşağıdaki gibidir:

$X \subseteq \mathbb{R}^n$ konveks küme, $f : \mathbb{R}^n \to \mathbb{R}$ ve $g : \mathbb{R}^n \to \mathbb{R}^r$ konveks fonksiyonlar olmak üzere

$$\begin{aligned} \min\ & f(x) \\ & g(x) \leqslant 0 \\ & x \in X \end{aligned} \tag{2.32}$$

Konveks programlama probleminde, reel değerli konveks bir fonksiyon konveks X kümesi üzerinde minimum yapılmak istenmektedir. Lineer programlama probleminde amaç fonksiyonu lineer dolayısıyla konveks fonksiyon olacağından lineer programlama problemi aslında konveks programlama probleminin özel bir halidir. Ayrıca kuadratik programlama probleminde amaç fonksiyonu $f(x) = \frac{1}{2}\langle x, Qx\rangle + \langle c, x\rangle$ konveks olduğundan kuadratik programlama problemi de konveks programlama problemidir.

$h : \mathbb{R}^n \to \mathbb{R}^k$ afin fonksiyon olmak üzere $h(x) = 0$ kısıtı $h(x) \leqslant 0$ ve $-h(x) \leqslant 0$ şeklinde iki eşitsizlik olarak yazılabildiğinden (2.32) problemine $h(x) = 0$ kısıtı da eklenebilir. Bu durumda

$$\begin{aligned} \min\ & f(x) \\ & g(x) \leqslant 0\ ,\ h(x) = 0 \\ & x \in X \end{aligned} \tag{2.33}$$

problemi kısıtları eşitsizlik ve eşitlik sistemi olarak verilmiş konveks programlama problemidir.

3. BULGULAR

Bu tez çalışmasının temelini oluşturan konveks programlama ve dualite kavramlarının iyi anlaşılması için konveks programlama problemlerinin dual problemleri kurularak, dual problemin nasıl bir yapıya sahip olduğu, dual ve konveks programlama probleminin hangi koşullar altında birbirlerine bağlı olduklarını veren dualite ilişkisinin ve bu problemlerin çözümleri arasındaki bağıntının ne olduğu ile ilgili sorulara yanıtlar incelenmektedir.

Konveks programlama problemleri için öncellikle dual problem oluşturulup, konveks programlama problemleri için çözümlerin varlığı ve nasıl bulunabileceğine dair temel bilgiler aşağıda verilmektedir.

3.1. DUAL PROBLEM

$f : \mathbb{R}^n \to \mathbb{R}$ ve $g_j : \mathbb{R}^n \to \mathbb{R}^r$ fonksiyon olmak üzere $X \subseteq \mathbb{R}^n$ kümesi için

$$\begin{aligned} \min\ & f(x) \\ & g_j(x) \leqslant 0\ ,\ j = 1, \ldots, r \\ & x \in X \end{aligned} \tag{3.1}$$

problemine primal problem denir. $X_0 = \{x : g_j(x) \leqslant 0\ ,\ x \in X\}$ uygun çözümler kümesi olmak üzere f^* ile gösterilen *optimal değer*

$$f^* = \inf_{x \in X_0} f(x) \tag{3.2}$$

şeklinde tanımlanır.

Kabul 3.1. Primal problemin en az bir uygun çözümü vardır ve optimal değer sınırlıdır, yani $-\infty < f^* < +\infty$ dır.

f ve g_j fonksiyonları konveks, diferansiyellenebilir ve $X = \mathbb{R}^n$ olsun. O halde eğer x^* minimum nokta ise öyle bir $\mu^* = (\mu_1^*, \ldots, \mu_r^*)$ vektörü vardır öyle ki her $j = 1, \ldots, r$ için $\mu_j^* \geqslant 0,\ \mu_j^* g_j(x^*) = 0$ ve $\nabla_x L(x^*, \mu^*) = 0$ dır ya da benzer şekilde

$$f^* = f(x^*) = \min_{x \in \mathbb{R}^n} L(x, \mu^*) \tag{3.3}$$

dir. Burada *Lagrange fonksiyonu* $L : \mathbb{R}^{n+r} \to \mathbb{R}$

$$L(x,\mu) = f(x) + \sum_{j=1}^{r} \mu_j g_j(x) = f(x) + \mu^T g(x) \tag{3.4}$$

şeklinde tanımlanır.

Tanım 3.1. $\mu^* = (\mu_1^*, \ldots, \mu_r^*)$ vektörüne eğer her $j = 1, \cdots, r$ için $\mu_j^* \geqslant 0$ ve

$$f^* = \inf_{x \in X} \; L(x, \mu^*) \tag{3.5}$$

ise *Lagrange çarpan vektörü (Lagrange çarpanı)* denir.

Lagrange çarpanını geometrik olarak yorumlamak için normal vektörü $(\mu, \mu_0) \neq (0,0)$ olan $\mathbb{R}^{r+1}$ uzayında hiperdüzlem denklemini ele alalım. $\mu \in \mathbb{R}^r$, $\mu_0 \in \mathbb{R}$ ve c sabiti ile

$$H = \{(z,w) : z \in \mathbb{R}^r, w \in \mathbb{R}, \quad \mu_0 w + \mu^T z = c\}$$

hiperdüzlemi verilsin. (μ, μ_0) normali ile $\mu_0 \neq 0$ ise hiperdüzlem dikey olmayan hiperdüzlem olarak adlandırılır ve bu hiperdüzlemler normal vektörü μ_0 ile bölünerek normalleştirilebileceği için $\mu_0 = 1$ alınabilir. Ayrıca kısıt ve amaç fonksiyon çiftinden oluşan $S \subseteq \mathbb{R}^n$,

$$S = \Big\{ \big(g(x), f(x)\big) : x \in X \Big\} \tag{3.6}$$

kümesi tanımlansın. Bu durumda aşağıdaki lemma geometrik olarak lagrange çarpanı hakkında bilgi vermektedir.

Lemma 3.1. *i)* $\big(g(x), f(x)\big)$ vektör çiftinden geçen ve normali $(\mu, 1)$ olan hiperdüzlem $\{(0,w) : \; w \in \mathbb{R}\}$ dikey eksenini $L(x,\mu)$ noktasında keser.

ii) Pozitif yarı-uzayında S kümesini içeren ve normali $(\mu, 1)$ olan tüm hiperdüzlemler arasından dikey ekseni kesen noktaların en büyüğü $\inf_{x \in X} L(x,\mu)$ dir.

iii) μ^* lagrange çarpanı olması için gerek ve yeter koşul $\mu^* \geqslant 0$ ve kapalı yarıuzayları S kümesini içeren normali $(\mu^*, 1)$ olan hiperdüzlemler arasından dikey eksenin kesim noktasının en büyük değeri f^* olmasıdır.

İspat: *i)* $\mu \in \mathbb{R}^r$ ve $c \in \mathbb{R}$ olmak üzere normali $(\mu, 1)$ olan hiperdüzlem denklemi $H = \{(z,w) : \; w + \mu^T z = c\}$ dir. Bu hiperdüzlem $(g(x), f(x))$ vektör çiftinden geçtiği için $f(x) + \mu^T g(x) = c$ olur. Lagrange fonksiyonu tanımından $L(x,\mu) = c$

olup hiperdüzlem $\{(0,w) : w \in \mathbb{R}\}$ dikey eksenini $L(x,\mu)$ noktasında keser.
$ii)$ Normali $(\mu, 1)$ olan ve pozitif yarı-uzayında S kümesini içeren hiperdüzlem

$$f(x) + \mu^T g(x) \geqslant c = w + \mu^T z \tag{3.7}$$

eşitsizliğini sağlar. Alttan sınırlı olduğundan infimum alınırsa

$$\inf_{x \in X} L(x,\mu) \geqslant c \tag{3.8}$$

olur. Böylece dikey eksenin kesim noktalarının en büyüğü $\inf\limits_{x \in X} L(x,\mu)$ dir.
$iii)$ μ^* lagrange çarpanı tanımından ve $ii)$ şıkkından istenen elde edilir. ■

Bu lemmadan hareketle normali $(\mu, 1)$ olan hiperdüzlemler arasından dikey ekseni kestiği en büyük değer veren hiperdüzlem S kümesine destek hiperdüzlemidir. Konveks programlama probleminde lagrange çarpanı bulmak önemlidir, çünkü eğer Lagrange çarpanı belirli ise problemin çözümü daha kolay bulunabilir. Fakat lagrange çarpanı yok ise çözüm de yoktur sonucu çıkarılmamalıdır. Aşağıda Lagrange çarpanı olmayan ama çözüme sahip konveks programlama örneği verilmektedir.

Örnek 3.1. a) $f(x) = x$ amaç fonksiyonu için

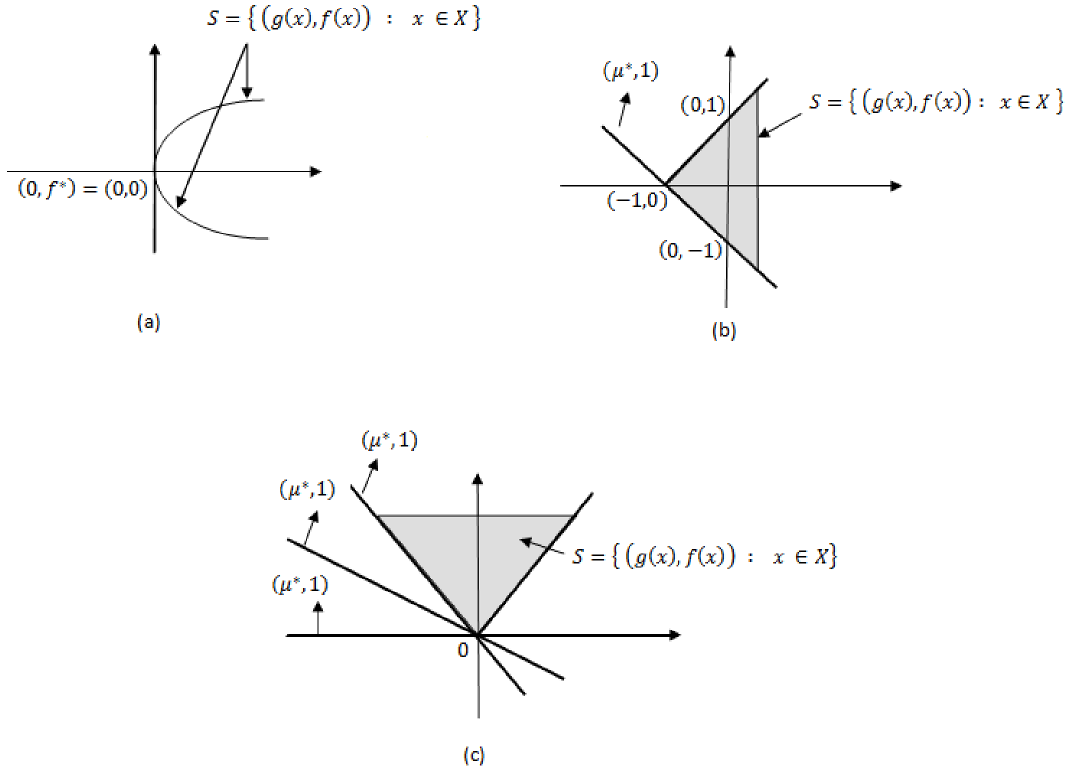
$$\begin{aligned} &\min f(x) \\ &\quad g(x) = x^2 \leqslant 0 \\ &\quad x \in X = \mathbb{R} \end{aligned}$$

probleminde Şekil 3.1a) 'da gösterilen S kümesine destek ve normali $(\mu^*, 1)$ olan hiperdüzlemler arasında dikey ekseni en büyük değerde kesen hiperdüzlem dikey olduğundan lagrange çarpanı yoktur ve $f^* = 0$ dır.

b) $f(x) = x_1 - x_2$ amaç fonksiyonu için

$$\begin{aligned} &\min f(x) \\ &\quad g(x) = x_1 + x_2 - 1 \leqslant 0 \\ &\quad x \in X = \{(x_1, x_2) : \; x_1 \geqslant 0, \; x_2 \geqslant 0\} \end{aligned}$$

problemini ele alalım. Şekil 3.1b) 'de gösterilen S kümesine destek ve normali $(\mu^*, 1)$ olan hiperdüzlemlerin dikey ekseni kesen noktalarının en büyük değeri $f^* = -1$ dir ve lagrange çarpanı $\mu^* = 1$ bulunur.



Şekil 3.1: Lagrange çarpanları

c) Şekil 3.1c) 'de ise $f(x) = |x_1| + x_2$ amaç fonksiyonu olmak üzere

$$\begin{aligned} &\min f(x) \\ &\quad g(x) = x_1 \leqslant 0 \\ &\quad x \in X = \{(x_1, x_2) : \; x_2 \geqslant 0\} \end{aligned}$$

probleminde S kümesine destek ve normali $(\mu^*, 1)$ olan hiperdüzlemler gösterilmektedir. Dolayısıyla primal optimal değer $f^* = 0$ dır ve lagrange çarpanları $\mu^* = [0, 1]$ aralığıdır.

Lemma 3.2. μ^*, lagrange çarpanı olsun. O halde x^*, primal problemin minimum noktası olması için gerek ve yeter koşul uygun çözümler kümesine ait olan x^* noktası için

$$x^* = \arg\min_{x \in X} L(x, \mu^*) \quad ve \quad \mu_j^* g_j(x^*) = 0 \quad , \; j = 1, \ldots, r \tag{3.9}$$

olmasıdır.

İspat: x^* primal problemin minimum noktası olsun.

$$f^* = f(x^*) \geqslant f(x^*) + \sum_{j=1}^{r} \mu_j^* g_j(x^*) = L(x^*, \mu^*) \geqslant \inf_{x \in X} L(x, \mu^*) \tag{3.10}$$

eşitsizliğinden $f(x^*) = f(x^*) + \sum\limits_{j=1}^{r} \mu_j^* g_j(x^*)$ dır. μ^* lagrange çarpanı ise $\mu^* \geqslant 0$ ve $j = 1, \ldots, r$ için $g_j(x^*) \leqslant 0$ olduğundan $\mu_j^* g_j(x^*) = 0$ dır. Ayrıca $x^* = \arg\min\limits_{x \in X} L(x, \mu^*)$ sağlanır. Tersine x^* uygun çözümler kümesine ait ve (3.9) denklemi sağlansın. Lagrange çarpanı tanımından

$$f(x^*) = f(x^*) + \sum_{j=1}^{r} \mu_j^* g_j(x^*) = L(x^*, \mu^*) = \min_{x \in X} L(x, \mu^*) = f^* \tag{3.11}$$

olduğundan x^* minimum noktadır. ∎

Tanım 3.2. $L : \mathbb{R}^{n+r} \to \mathbb{R}$ lagrange fonksiyonu, $\mu \in \mathbb{R}^r$ için

$$q(\mu) = \inf_{x \in X} L(x, \mu) \tag{3.12}$$

fonksiyonuna *dual lagrange fonksiyonu* ya da *dual fonksiyon* denir.

Lagrange fonksiyonu alttan sınırlı değil ise dual fonksiyon $-\infty$ değerini alacağı için bu durumdan kaçınmak için $D = \{\mu : \ q(\mu) > -\infty\}$ kümesi olmak üzere

$$\begin{aligned} &\text{maks } q(\mu) \\ &\qquad \mu \geqslant 0 \\ &\qquad \mu \in D \end{aligned} \tag{3.13}$$

problemine *dual problem* denir. *Dual optimal değer* q^* ile gösterilir ve

$$q^* = \sup_{\mu \geqslant 0} q(\mu) \tag{3.14}$$

şeklinde tanımlanır.

Örnek 3.2. A, $m \times n$ reel matris olmak üzere $x, c \in \mathbb{R}^n$ ve $b \in \mathbb{R}^m$ için

$$\begin{aligned} &\min c^T x \\ &\qquad Ax \leqslant b \\ &\qquad x \geqslant 0 \end{aligned} \tag{3.15}$$

kanonik formda verilen lineer programlama problemini göz önüne alalım. Dual fonksiyon

$$\begin{aligned} q(\mu) &= \inf_{x \geqslant 0} \left\{ c^T x + \mu^T (Ax - b) \right\} \\ &= -\mu^T b + \inf_{x \geqslant 0} \left\{ (c^T + \mu^T A) x \right\} = \begin{cases} -\mu^T b & , \ c^T + \mu^T A \geqslant 0 \\ -\infty & , \ \text{diğer} \end{cases} \end{aligned}$$

olarak bulunur. Dolayısıyla

$$\begin{aligned} \text{maks } & -\mu^T b \\ & c + A^T\mu \geqslant 0 \\ & \mu \geqslant 0 \end{aligned}$$

problemi (3.15) lineer programlama probleminin dual problemidir.

Örnek 3.3. Q, $n \times n$ pozitif yarı tanımlı bir matris ve A, $m \times n$ reel matris için $b \in \mathbb{R}^m$, $c \in \mathbb{R}^n$ ve $f(x) = \frac{1}{2}x^T Qx + c^T x$ olmak üzere

$$\begin{aligned} \min \ & f(x) \\ & Ax \leqslant b \end{aligned} \tag{3.16}$$

kuadratik programlama problemini ele alalım. Dual fonksiyon

$$q(\mu) = \inf_{x \in \mathbb{R}^n} \{\frac{1}{2}x^T Qx + c^T x + \mu^T(Ax - b)\}$$

olur. $x = -Q^{-1}(c + A^T\mu)$ için $q(\mu)$ dual fonksiyonu minimuma erişeceği için

$$q(\mu) = -\frac{1}{2}\mu^T AQ^{-1}A^T\mu - \mu^T(b + AQ^{-1}c) - \frac{1}{2}c^T Q^{-1}c$$

bulunur. Böylece $P = AQ^{-1}A^T$ ve $t = b + AQ^{-1}c$ alınırsa,

$$\begin{aligned} \text{maks } & -\frac{1}{2}\mu^T P\mu - \mu^T t - \frac{1}{2}c^T Q^{-1}c \\ & \mu \geqslant 0 \end{aligned}$$

problemi ya da eşdeğer olarak

$$\begin{aligned} \min \ & \frac{1}{2}\mu^T P\mu + \mu^T t + \frac{1}{2}c^T Q^{-1}c \\ & \mu \geqslant 0 \end{aligned}$$

problemi (3.16) kuadratik programlama probleminin dual problemidir.

Lemma 3.3. q dual fonksiyonu, konveks D kümesi üzerinde konkavdır.

İspat: Konveks küme tanımından $D = \{\mu : \ q(\mu) > -\infty\}$ kümesinin konveksliği kolayca görülür. $\alpha \in \mathbb{R}$ ve $\mu, \tilde{\mu} \in D$ için

$$L(x, \alpha\mu + (1-\alpha)\tilde{\mu}) = \alpha L(x, \mu) + (1-\alpha)L(x, \tilde{\mu}) \tag{3.17}$$

denkleminde $x \in X$ üzerinden infimum alınırsa

$$\begin{aligned}
q(\alpha\mu + (1-\alpha)\tilde{\mu}) &= \inf_{x \in X} L(x, \alpha\mu + (1-\alpha)\tilde{\mu}) \\
&= \inf_{x \in X} \{\alpha L(x,\mu) + (1-\alpha) L(x,\tilde{\mu})\} \\
&\geqslant \alpha \inf_{x \in X} L(x,\mu) + (1-\alpha) \inf_{x \in X} L(x,\tilde{\mu}) = \alpha q(\mu) + (1-\alpha) q(\tilde{\mu})
\end{aligned}$$

olur. Böylece q dual fonksiyonu D üzerinde konkav fonksiyondur. ■

$q(\mu) = \inf_{x \in X} L(x,\mu)$ dual fonksiyonunun konkavlığı aslında konkav fonksiyonların infimumu da konkav olacağı gerçeğinden çıkarılabilir. Bir diğer önemli özellik ise optimal dual değerin daima primal optimal değerden küçük eşit kalmasıdır. Zayıf dualite teoremi olarak bilinen bu özellik aşağıda verilmektedir.

Teorem 3.4. (Zayıf Dualite Teoremi) Dual optimal değer $q^* = \sup_{\mu \geqslant 0} q(\mu)$ ve $f^* = \inf_{x \in X_0} f(x)$ primal optimal değer için

$$q^* \leqslant f^* \tag{3.18}$$

gerçeklenir.

ispat: Her $\mu \geqslant 0$, $x \in X$ ve $g(x) \leqslant 0$ olmak üzere

$$q(\mu) = \inf_{x \in X} L(x,\mu) \leqslant f(x) + \sum_{j=1}^{r} \mu_j g_j(x) \leqslant f(x) \tag{3.19}$$

dır. Her $x \in X_0 = \{x : x \in X \ , \ g(x) \leqslant 0\}$ ve her $\mu \geqslant 0$ için

$$q^* = \sup_{\mu \geqslant 0} q(\mu) \leqslant \inf_{X_0} f(x) = f^* \tag{3.20}$$

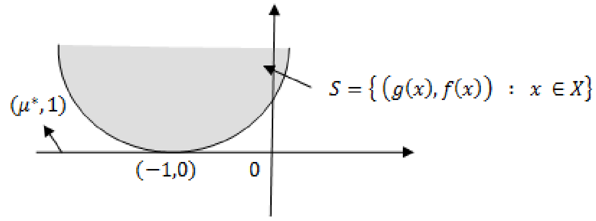
olacağından ispat tamamlanır. ■

Tanım 3.3. Eğer $q^* < f^*$ ise dualite boşluğu vardır, eğer $q^* = f^*$ ise dualite boşluğu yoktur denir.

Eğer μ^* lagrange çarpanı varsa $f^* = q(\mu^*) \leqslant q(\mu) \leqslant q^*$ olacağından ve Teorem 3.4 nedeniyle dualite boşluğu yoktur. Fakat dualite boşluğu yok iken lagrange çarpanının var olması gerekmez. Örnek 3.1a) 'da verilen problemde lagrange çarpanının olmadığı ve $f^* = 0$ olduğu gösterilmişti. Dual fonksiyon

$$q(\mu) = \inf_{x \in \mathbb{R}} L(x,\mu) = \inf_{x \in \mathbb{R}} \{x + \mu x^2\} = \begin{cases} \frac{-1}{4\mu} & , \ \mu > 0 \\ -\infty & , \ \mu \leqslant 0 \end{cases}$$

olup $q^* = \sup_{\mu \geqslant 0} q(\mu) = 0$ bulunur. Dolayısıyla dualite boşluğu yoktur.



Şekil 3.2: Dualite boşluğu olmayan konveks programlama problemi

Lemma 3.5. *i)* Dualite boşluğu yok ise, lagrange çarpanları kümesi ile dual optimal çözüm kümesi aynıdır.

ii) Eğer dualite boşluğu var ise, lagrange çarpanları kümesi boş kümedir.

İspat: *i)* Dualite boşluğu yok ise $f^* = f(x^*) = \inf_{x \in X} L(x, \mu^*) = q(\mu^*) \leqslant q^* = f^*$ olacağından lagrange çarpanları kümesi ile dual optimal çözümleri kümesi aynıdır. *ii)* Dualite boşluğu var olsun. Kabul edelim ki lagrange çarpanları kümesi boştan farklı olsun. μ^* lagrange çarpanı olmak üzere $f^* = \inf_{x \in X} L(x, \mu^*) = q(\mu^*) \leqslant q^*$ olur. Bu ise dualite boşluğu var olduğundan çelişki yaratır. Böylece dualite boşluğu var ise lagrange çarpanları kümesi boş kümedir. ∎

Örnek 3.4. Aşağıda verilen

$$\min f(x) = \frac{x_1^2 + x_2^2}{2}$$
$$g(x) = x_1 - 1 \leqslant 0$$
$$x \in \mathbb{R}^2$$

problemde Şekil 3.2 'de gösterilen S kümesine destek ve normali $(\mu^*, 1)$ olan hiperdüzlemlerin dikey ekseni kesen en büyük nokta primal optimal değer olan $f^* = 0$ dır ve lagrange çarpanı $\mu^* = 0$ dır. Dual fonksiyon

$$q(\mu) = \inf_{x \in \mathbb{R}^2} \{\frac{x_1^2 + x_2^2}{2} + \mu(x_1 - 1)\} = \frac{-\mu^2}{2} - \mu$$

olarak bulunur. Dual optimal değer ise $q^* = \sup_{\mu \geqslant 0} q(\mu) = 0$ olur. Böylece dualite boşluğu yoktur. Ayrıca lagrange çarpanları kümesi ile dual optimal çözüm noktası aynıdır.

Aşağıda verilen lemma ile primal ve dual optimal çözüm çiftlerinin karakterizasyonu verilmektedir. Fakat bu karakterizasyon yalnızca dualite boşluğu olmadığı durumlar için geçerlidir, çünkü aksi taktirde lagrange çarpanı mevcut değildir.

Lemma 3.6. [7] (x^*, μ^*) optimal çözüm ve lagrange çarpan çifti olabilmesi için gerek ve yeter koşul

$$x^* \in X,\ g(x^*) \leqslant 0, \qquad \text{(Primal Uygunluk)} \tag{3.21}$$

$$\mu^* \geqslant 0, \qquad \text{(Dual Uygunluk)} \tag{3.22}$$

$$x^* = \arg\min_{x \in X} L(x, \mu^*) \tag{3.23}$$

$$\mu_j^* g_j(x^*) = 0\ , \quad j = 1, \ldots, r \tag{3.24}$$

olmasıdır.

Teorem 3.7. (Eyer Noktası Teoremi) (x^*, μ^*) optimal çözüm ve lagrange çarpan çifti olması için gerek ve yeter koşul $x^* \in X,\ \mu^* \geqslant 0$ için

$$L(x^*, \mu) \leqslant L(x^*, \mu^*) \leqslant L(x, \mu^*)\ , \quad \forall\, x \in X,\ \forall\, \mu \geqslant 0 \tag{3.25}$$

gerçeklenir. Burada (x^*, μ^*) vektör çifti lagrange fonksiyonunun eyer noktasıdır.

İspat: Eğer (x^*, μ^*) optimal çözüm ve lagrange çarpan çifti ise Lemma 3.6 nedeniyle $x^* \in X, \mu^* \geqslant 0$ dır ve (3.23) bağıntısıdan $L(x^*, \mu^*) \leqslant L(x, \mu^*)$ elde edilir.

$$L(x^*, \mu) = f(x^*) + \sum_{j=1}^{r} \mu_j g_j(x^*) \leqslant f(x^*) = L(x^*, \mu^*) \tag{3.26}$$

olup (3.25) eşitsizliği sağlanır. Tersine $x^* \in X, \mu^* \geqslant 0$ ve (3.25) eşitsizliği sağlansın.

$$\sup_{\mu \geqslant 0} L(x^*, \mu) = \sup_{\mu \geqslant 0}\{f(x^*) + \sum_{j=1}^{r} \mu_j g_j(x^*)\} = \begin{cases} f(x^*) & , \ g(x^*) \leqslant 0 \\ +\infty & , \ g(x^*) > 0 \end{cases}$$

ve ayrıca (3.25) eşitsizliğinde $\forall\, x \in X, \forall\, \mu \geqslant 0$ için $L(x^*, \mu) \leqslant L(x^*, \mu^*)$ olduğundan $g(x^*) \leqslant 0$ ve $f(x^*) = L(x^*, \mu^*)$ dır. Dolayısıyla $\sum_{j=1}^{r} \mu_j^* g_j(x^*) = 0$ olup (3.24) denklemi sağlanır. $\forall\, x \in X$ için $L(x^*, \mu^*) \leqslant L(x, \mu^*)$ olduğundan $x^* = \arg\min_{x \in X} L(x, \mu^*)$ (3.23) bağıntısı gerçeklenir. Böylece Teorem 3.6 'dan (x^*, μ^*) optimal çözüm ve lagrange çarpan çifti olarak bulunur. ∎

Bu bölümde incelenen problemlerde eşitsizlik kısıtları dikkate alınıp dual problem oluşturulup çözümlemeler yapıldı. Fakat eşitsizlik kısıtının yanı sıra eşitlik kısıtı olması halinde $j = 1, \ldots, r$ ve $i = 1, \ldots, m$ için

$$\begin{aligned} &\min f(x) \\ &\quad g_j(x) \leqslant 0\ ,\ h_i(x) = 0 \\ &\quad x \in X \end{aligned} \tag{3.27}$$

olarak yeni bir problem tanımlanabilir. Her $i = 1, \dots, m$ için $h_i(x) = 0$ kısıtı, $h_i(x) \leqslant 0$ ve $-h_i(x) \leqslant 0$ iki eşitsizlik kısıtı olarak yazılabilir. $h_i(x) \leqslant 0$ için λ_i^+ skaleri, $-h_i(x) \leqslant 0$ için λ_i^- skaleri ve her i için $\lambda_i^+ \geqslant 0$, $\lambda_i^- \geqslant 0$ ve her j için $\mu_j^* \geqslant 0$ olmak üzere

$$f^* = \inf_{x \in X} \left\{ f(x) + \sum_{j=1}^{r} \mu_j^* g_j(x) + \sum_{i=1}^{m} (\lambda_i^+ - \lambda_i^-) h_i(x) \right\}$$

ise $(\mu_1^*, \dots, \mu_r^*, \lambda_1^+, \lambda_1^-, \dots, \lambda_m^+, \lambda_m^-)$ vektörü Lagrange çarpanıdır. $\lambda_i^+ - \lambda_i^- = \lambda_i$ alınırsa Lagrange fonksiyonu

$$L(x, \mu, \lambda) = f(x) + \sum_{j=1}^{r} \mu_j g_j(x) + \sum_{i=1}^{m} \lambda_i h_i(x)$$

olup $\mu^* \geqslant 0$ ve $f^* = \inf_{x \in X} L(x, \mu^*, \lambda^*)$ ise (μ^*, λ^*) vektör çifti Lagrange çarpanıdır. Dual fonksiyon

$$q(\mu, \lambda) = \inf_{x \in X} L(x, \mu, \lambda)$$

olmak üzere dual problem

$$\begin{aligned} &\text{maks } q(\mu, \lambda) \\ &\mu \geqslant 0 \ , \ \lambda \in \mathbb{R}^m \end{aligned} \tag{3.28}$$

şeklinde olup $\lambda \in \mathbb{R}^m$ olduğu önemlidir.

Kabul 3.1 'de primal problemin en azından bir tane uygun çözümünün var ve optimal değerin sınırlı olduğu kabul edilmişti. Şimdi Kabul 3.1 'deki koşulların geçerli olmadığı durumları inceleyelim. $f^* = -\infty$ olması halinde Zayıf Dualite Teoreminden her $\mu \geqslant 0$ için $q(\mu) = -\infty$ olur bu ise dual problemin uygun çözümünün olmadığını gösterir. Çünkü $D = \{\mu : \ q(\mu) > -\infty\}$ kümesi için dual problem

$$\begin{aligned} &\text{maks } q(\mu) \\ &\mu \geqslant 0 \\ &\mu \in D \end{aligned} \tag{3.29}$$

olarak tanımlanmıştı. Dolayısıyla optimal değerin sınırlı olması gerekir. X kısıt kümesi boştan farklı fakat primal problemin uygun çözümler kümesi $\{x \in X : \ g(x) \leqslant 0\}$ boş küme olsun. Bu durumda $f^* = \infty$ olduğu kabul edilir. $f^* = \infty$ olması durumunda aşağıda verilen örneklerde q^* dual optimal çözümün farklı değerler alması söz konusudur.

Örnek 3.5. a) $f(x) = \frac{1}{x}$ fonksiyonu için

$$\begin{aligned} &\min f(x) \\ &\quad g(x) = x \leqslant 0 \\ &\quad x \in X = \{x :\ x > 0\} \end{aligned}$$

probleminde uygun çözümler kümesi boş küme olduğundan $f^* = \infty$ dır. Dual fonksiyon

$$q(\mu) = \inf_{x \in X}\{L(x,\mu)\} = \inf_{x>0}\{\frac{1}{x} + \mu x\} = 2\sqrt{\mu}$$

olur $q^* = \sup_{\mu \geqslant 0} q(\mu) = \infty$ bulunur.

b) $f(x) = x$ fonksiyonu için

$$\begin{aligned} &\min f(x) \\ &\quad g(x) = x^2 \leqslant 0 \\ &\quad x \in X = \{x :\ x > 0\} \end{aligned}$$

probleminde uygun çözümler kümesi boş küme olduğundan $f^* = \infty$ dır. Dual fonksiyon

$$q(\mu) = \inf_{x \in X}\{L(x,\mu)\} = \inf_{x>0}\{x + \mu x^2\} = \begin{cases} \frac{-1}{4\mu} & , \mu > 0 \\ -\infty & , \mu \leqslant 0 \end{cases}$$

bulunur ve $q^* = 0$ dır.

c) $f(x) = x_1 + x_2$ amaç fonksiyonu olmak üzere

$$\begin{aligned} &\min f(x) \\ &\quad g(x) = x_1 \leqslant 0 \\ &\quad x \in X = \{(x_1, x_2) :\ x_1 > 0\} \end{aligned}$$

probleminde uygun çözümler kümesi boş küme olduğundan $f^* = \infty$ dır. Dual fonksiyon $q(\mu) = \inf_{x \in X}\{x_1 + x_2 + \mu x_1\} = -\infty$ dır ve $q^* = -\infty$ olur.

Dualite teorisinde dualite boşluğunun olmadığını ve Lagrange çarpanının var olduğunu garantilemek için primal problemin kısıt ve amaç fonksiyonları üzerinde konvekslik koşulları getirmek gerekir.

3.2. AMAÇ FONKSİYONU KONVEKS VE KISIT FONKSİYONLARI LİNEER OLAN PROBLEMLER

$f : \mathbb{R}^n \to \mathbb{R}$ konveks fonksiyon ve $X \subseteq \mathbb{R}^n$ polihedral küme olsun.

$$
\begin{aligned}
\min\ & f(x) \\
& e_i^T x - d_i = 0\ , \quad i = 1, \ldots, m \\
& a_j^T x - b_j \leqslant 0\ , \quad j = 1, \ldots, r \\
& x \in X
\end{aligned}
\tag{3.30}
$$

problemini ele alalım.

Teorem 3.8. (Lineer Kısıtlı Güçlü Dualite Teoremi) $f : \mathbb{R}^n \to \mathbb{R}$ konveks fonksiyon ve $X \subseteq \mathbb{R}^n$ polihedral küme olmak üzere (3.30) probleminde uygun çözümler kümesi boştan farklı ve primal problemin optimal değeri f^* sonlu olsun. O halde dualite boşluğu yoktur ve en azından bir tane Lagrange çarpanı vardır.

Bu Lineer Kısıtlı Güçlü Dualite Teoremi 'ni ispatlamadan önce bazı ön bilgileri verelim. Teorem 3.8 'de verilen f fonksiyonunun $\mathbb{R}^n$ uzayında konveks olması önemlidir. Çünkü eğer f fonksiyonu X polihedral kümesinde konveks ise dualite boşluğu olabilir.

Örnek 3.6. $f(x) = e^{-\sqrt{x_1 x_2}}$ fonksiyonu ve her $x \in X$ için

$$
\begin{aligned}
\min\ & f(x) \\
& g(x) = x_1 = 0 \\
& x \in X = \{x :\ x \geqslant 0\}
\end{aligned}
$$

probleminde f fonksiyonu X polihedral kümesi üzerinde konvekstir. Primal optimal değer $f^* = \inf\limits_{x \in X_0} f(x) = 1$ dır. Dual fonksiyon

$$
q(\mu) = \inf_{x \in X} L(x, \mu)\} = \inf_{x \geqslant 0} \{e^{-\sqrt{x_1 x_2}} + \mu x_1\} = 0
$$

bulunur. Dual optimal değer ise $q^* = \sup\limits_{\mu \geqslant 0} q(\mu) = 0$ dır. $f^* \neq q^*$ olduğundan dualite boşluğu vardır.

Eğer Teorem 3.8 'de verilen X kısıt kümesi sınırlı ise Weierstrass Teoremine göre primal problem optimal çözüme sahiptir. Çünkü Weierstrass Teoremine göre sonlu boyutlu uzayda boştan farklı sınırlı ve kapalı bir alt küme üzerinde sürekli fonksiyon mutlak maksimum ve minimuma ulaşır.

Lemma 3.9. (Lineer Programlama İçin Dualite Teoremi)[7] (3.30) probleminin uygun çözümler kümesi boştan farklı ve f^* primal optimal değeri sonlu olsun. f lineer fonksiyon ve $X \subseteq \mathbb{R}^n$ polihedral küme olsun. O halde primal ve dual problem optimal çözümlere sahiptir ve dualite boşluğu yoktur.

Örnek 3.7. Aşağıdaki kanonik formda verilen lineer programlama problemini göz önüne alalım.

$$\begin{aligned} \min\ & 3x_2 + 3x_4 \\ & x_1 - 2x_2 - x_4 \leqslant -4 \\ & -x_2 + x_3 - 2x_4 \leqslant -5 \\ & x_1, x_2, x_3, x_4 \geqslant 0 \end{aligned} \tag{3.31}$$

(3.31) probleminin çözümünü bulmak klasik yollarla zor olduğu için çözüm yöntemi daha kolay olan dual problemini çözmek yerinde olacaktır. Bu sebeple (3.31) probleminin dual problemi

$$\begin{aligned} \text{maks}\ & 4\mu_1 + 5\mu_2 \\ & 2\mu_1 + \mu_2 \leqslant 3 \\ & \mu_1 + 2\mu_2 \leqslant 3 \\ & \mu_1, \mu_2 \geqslant 0 \end{aligned} \tag{3.32}$$

şeklinde bulunur. (3.32) probleminin optimal çözümü $\mu = (1, 1)$ için optimal değeri 9 olarak bulunur. Lemma 3.9 nedeniyle dualite boşluğu olmadığı için (3.31) probleminin de optimal değeri 9 dur.

Lemma 3.10. (Lineer Programlamada Primal Optimal Çözümlerin Varlığı) (3.30) probleminin uygun çözümler kümesi boştan farklı ve f^* primal optimal değeri sonlu olsun. f lineer fonksiyon ve $X \subseteq \mathbb{R}^n$ polihedral küme olsun. (3.30) problemi en az bir optimal çözüme sahiptir.

İspat: A, $m \times n$ matris ve $c \in \mathbb{R}^n$, $b \in \mathbb{R}^m$ için

$$\begin{aligned} \min\ & c^T x \\ & Ax = b \\ & x \geqslant 0 \end{aligned} \tag{3.33}$$

lineer programlama probleminin çözümünün olmadığını kabul edelim. Bu durumda $Ax = b$ denklemini sağlayan her $x \geqslant 0$ için $c^T x > f^*$ dir. Açıkca

$$\begin{pmatrix} f^* \\ -b \end{pmatrix} = \begin{pmatrix} c^T \\ -A \end{pmatrix} x$$

sağlayan $x \geqslant 0$ yoktur. Farkas Lemmasından $\exists\ \alpha \in \mathbb{R}$ ve $\xi \in \mathbb{R}^m$ vardır öyle ki

$$\alpha f^* - b^T\xi \ < \ 0 \tag{3.34}$$

$$\alpha c - A^T\xi \ \geqslant \ 0 \tag{3.35}$$

eşitsizlikleri sağlanır. (3.33) probleminin uygun çözümler kümesi boştan farklı olduğu için öyle bir $x \geqslant 0$ vardır öyle ki $Ax = b$ ve $c^T x > f^*$ dir. Buna göre (3.35) eşitsizliği x ile iç çarpım yapılırsa

$$\alpha c^T x - b^T\xi \geqslant 0 \tag{3.36}$$

olur. Dolayısıyla (3.34) eşitsizliğinden $\alpha c^T x > \alpha f^*$ olup $\alpha > 0$ elde edilir. Böylece (3.34) ve (3.35) eşitsizlikleri α ile bölünüp $\eta = \frac{\xi}{\alpha}$ alınırsa

$$f^* < b^T\eta\ , \quad A^T\eta \leqslant c \tag{3.37}$$

olur. (3.33) lineer programlama probleminin dual problemi

$$\begin{aligned} &\text{maks } b^T\xi \\ &\qquad A^T\xi \leqslant c \end{aligned}$$

dir. Böylece (3.37) eşitsizliği primal optimal değerden daha büyük dual değerin var olduğunu gösterir. Zayıf Dualite Teoremi sebebiyle bu bir çelişkidir. Dolayısıyla (3.33) lineer programlama probleminin bir çözümü vardır. Genel lineer programlama problemi

$$\begin{aligned} &\min\ c^T x \\ &\quad e_i^T x - d_i = 0\ , \quad i = 1, \ldots, m \\ &\quad a_j^T x - b_j \leqslant 0\ , \quad j = 1, \ldots, r \\ &\quad x \in X \end{aligned} \tag{3.38}$$

şeklinde verilir. X polihedral küme olduğundan

$$X = \{x :\ a_j^T x - b_j \leqslant 0\ ,\ j = r+1, \ldots, \tilde{r}\} \tag{3.39}$$

lineer eşitsizlikler sistemi olarak ifade edilebilir. (3.38) probleminin standart formu

$$\begin{aligned} &\min c^T\ (x_+ - x_-) \\ &\quad a_j^T(x_+ - x_-) - b_j + z_j = 0\ , \quad j = 1, \ldots, \tilde{r} \\ &\quad e_i^T(x_+ - x_-) - d_i = 0\ , \qquad\quad i = 1, \ldots, m \\ &\quad x_+ \geqslant 0,\ x_- \geqslant 0,\ z \geqslant 0 \end{aligned} \tag{3.40}$$

olarak verilir. (3.40) problemi (3.33) probleminin özel bir halidir. Dolayısıyla (3.38) probleminin de bir çözüme sahip olduğu gösterilmiş olur. ∎

Lemma 3.11. (Genelleştirilmiş Farkas Lemması)[7] $C \subseteq \mathbb{R}^n$ konveks küme üzerinde $f : C \to \mathbb{R}$ konveks fonksiyon, her $j = 1, \ldots, r$ için $a_j \in \mathbb{R}^n, b_j \in \mathbb{R}$ olmak üzere

$$S = \{x :\ a_j^T x \leqslant b_j\ ,\ j = 1, \ldots, r\} \tag{3.41}$$

kümesi $ri(C)$ kümesinin bir elemanını içersin ve

$$F(x) \geqslant 0 \quad ve \quad \forall\, x \in S \cap C$$

olsun. O halde $j = 1, \ldots, r$ için $\mu_j \geqslant 0$ skaleri vardır öyle ki

$$0 \leqslant F(x) + \sum_{j=1}^{r} \mu_j(a_j^T x - b_j) \quad , \quad \forall\, x \in C \tag{3.42}$$

eşitsizliği gerçeklenir.

Bu gerekli bilgilerden sonra Lineer Kısıtlı Güçlü Dualite Teoremi 3.8 'in ispatı aşağıda verilmektedir.

İspat: (3.30) probleminde verilen eşitlik kısıtlarını da eşisizlik kısıtı olarak alınırsa $j = 1, \ldots, r$ için

$$\begin{aligned} &\min f(x) \\ &\quad a_j^T x - b_j \leqslant 0 \\ &\quad x \in X \end{aligned} \tag{3.43}$$

probleminde Lemma 3.11 'de verilen C kümesi $\mathbb{R}^n$ ve $F(x) = f(x) - f^* \geqslant 0$ şeklinde alınırsa öyle $\mu_1, \ldots, \mu_p$ vektörleri ve her j için $\mu_j \geqslant 0$ skaleri vardır öyle ki

$$0 \leqslant f(x) - f^* + \sum_{j=1}^{r} \mu_j(a_j^T x - b_j) \quad , \quad \forall\, x \in X \tag{3.44}$$

eşitsizliği gerçeklenir. (3.44) eşitsizliğinden ve dual fonksiyon tanımından

$$f^* \leqslant \inf_{x \in X} L(x,\mu) = q(\mu) \leqslant q^*$$

bulunur. Ayrıca Zayıf Dualite Teoremi nedeniyle $f^* = q^*$ olur. Dualite boşluğu olmadığından dual optimal çözüm kümesi ile lagrange çarpan kümesi aynı olduğundan μ lagrange çarpanı vardır. Böylece istenen elde edilir. ■

3.3. AMAÇ VE KISIT FONKSİYONLARI KONVEKS OLAN PROBLEMLER

$f : \mathbb{R}^n \to \mathbb{R}$ ve $g_j : \mathbb{R}^n \to \mathbb{R}$ konveks fonksiyonlar ve $j = 1, \ldots, r$ için

$$\begin{aligned} \min\ & f(x) \\ & g_j(x) \leqslant 0 \\ & x \in X \end{aligned} \tag{3.45}$$

problemini inceleyelim.

Kabul 3.2. (3.45) probleminin uygun çözümler kümesi boştan farklı ve optimal değer f^* sonlu olsun. $X \subseteq \mathbb{R}^n$ konveks küme, $f : \mathbb{R}^n \to \mathbb{R}$ ve $g_j : \mathbb{R}^n \to \mathbb{R}$, X üzerinde konveks fonksiyonlar olsun. Öyle bir $\tilde{x} \in X$ vektörü vardır öyle ki

$$g_j(\tilde{x}) < 0 \quad , \quad \forall\, j = 1, \ldots, r \tag{3.46}$$

koşulu gerçeklenir.

Teorem 3.12. (3.45) problemi için Kabul 3.2 sağlansın. O halde dualite boşluğu yoktur ve en az bir Lagrange çarpanı vardır.

İspat: $A \subseteq \mathbb{R}^{r+1}$ kümesi

$$A = \{(z_1, \ldots, z_r, w) : \exists\, x \in X \text{ öyle ki } g_j(x) \leqslant z_j,\ j = 1, \ldots, r, \quad f(x) \leqslant w\} \tag{3.47}$$

şeklinde tanımlansın. A kümesi konvekstir. $(0, f^*)$, A konveks kümesinin iç noktası değildir. Gerçekten eğer $(0, f^*)$, A konveks kümesinin iç noktası olsaydı öyle $\varepsilon > 0$ vardır öyle ki $(0, f^* - \varepsilon) \in A$ olurdu, bu ise primal optimal değerin f^* olması ile çelişir. O halde $(0, f^*)$, A konveks kümesinin iç noktası değildir. Destek Hiperdüzlem Teoremi 2.14 nedeniyle öyle $(\mu, \beta) \neq (0, 0)$ vektörü vardır öyle ki

$$\beta f^* \leqslant \beta w + \mu^T z \quad , \quad \forall\, (z, w) \in A \tag{3.48}$$

eşitsizliği gerçeklenir. Her $j = 1,\ldots,r$ ve her $\gamma > 0$ için $(z, w+\gamma) \in A$ ve $(z_1,\ldots,z_j+\gamma,\ldots,z_r,w) \in A$ olup (3.48) eşitsizliğinden $\beta \geqslant 0$, $\mu_j \geqslant 0$ elde edilir. Kabul edelim ki $\beta = 0$ olsun. (3.48) eşitsizliğinden her $(z,w) \in A$ için $0 \leqslant \mu^T z$ olur. $(g_1(\tilde{x}),\ldots,g_r(\tilde{x}),f(\tilde{x})) \in A$ olduğundan

$$0 \leqslant \sum_{j=1}^{r} \mu_j g_j(\tilde{x}) \tag{3.49}$$

bulunur. $\mu \geqslant 0$ ve Kabul 3.2 'den her j için $g(\tilde{x}) < 0$ olduğundan $\mu_j = 0$ olur. Bu ise $(\mu,\beta) = (0,0)$ çelişkisini yaratır. O halde $\beta > 0$ dır. Her $x \in X$ için $(g(x), f(x)) \in A$ olduğundan $\beta = 1$ alınıp (3.48) eşitsizliğinde yerine yazılırsa

$$f^* \leqslant f(x) + \mu^T g(x) \quad , \quad \forall\, x \in X \tag{3.50}$$

olur. $x \in X$ üzerinden infimum alınırsa ve $\mu \geqslant 0$ olduğundan

$$f^* \leqslant \inf_{x\in X}\{f(x) + \mu^T g(x)\} = q(\mu) \leqslant q^* \tag{3.51}$$

bulunur ve ayrıca Zayıf Dualite Teoremi kullanılırsa $f^* = q^*$ gerçeklenir. Böylece dualite boşluğu yoktur ve lagrange çarpanı μ vardır. ∎

Konveks eşitsizlik ve lineer eşitlik kısıtları olması durumunda yani $j = 1,\ldots,r$, $i = 1,\ldots,m$ için $g_j(x)$ konveks fonksiyon olmak üzere

$$\begin{aligned} &\min f(x) \\ &\quad g_j(x) \leqslant 0\ ,\ e_i^T x - d_i = 0 \\ &\quad x \in X \end{aligned} \tag{3.52}$$

probleminde $e_i^T x - d_i = 0$ lineer kısıtı $e_i^T x - d_i \leqslant 0$ ve $-e_i^T x + d_i \leqslant 0$ şeklinde iki eşitsizlik kısıtı olarak düşünülürse Kabul 3.2 'de verilen iç nokta şartı sağlanmayacağı için Teorem 3.12 geçerli değildir. Bu sebeple yeni bir kabul aşağıda verilmektedir.

Kabul 3.3. $X \subseteq \mathbb{R}^n$ uzayında konveks küme, $f : \mathbb{R}^n \to \mathbb{R}$, $g_j : \mathbb{R}^n \to \mathbb{R}$ fonksiyonları X üzerinde konveks ve f^* primal optimal değer sonlu olsun. O halde öyle bir $\tilde{x} \in ri(X)$ vektörü vardır öyle ki her $j = 1,\ldots,r$ ve $i = 1,\ldots,m$ için

$$g_j(\tilde{x}) < 0 \quad \text{ve} \quad e_i^T \tilde{x} - d_i = 0 \tag{3.53}$$

gerçeklenir.

Eşitsizlik kısıtları altında verilen Teorem 3.12 'ye paralel olarak eşitlik kısıtları da olması durumunda aşağıdaki teorem verilir.

Teorem 3.13. [7] $X \subseteq \mathbb{R}^n$ uzayında konveks küme, $f : \mathbb{R}^n \to \mathbb{R}$, $g_j : \mathbb{R}^n \to \mathbb{R}$ fonksiyonları X üzerinde konveks ve f^* optimal değeri sonlu olsun. $j = 1, \ldots, r$, $i = 1, \ldots, m$ için

$$\begin{aligned} \min\ & f(x) \\ & g_j(x) \leqslant 0\ ,\ e_i^T x - d_i = 0 \\ & x \in X \end{aligned} \tag{3.54}$$

probleminde Kabul 3.3 geçerli olsun. Bu durumda dualite boşluğu yoktur ve en az bir Lagrange çarpanı vardır.

Teorem 3.13 'ün geçerli olması için teoremde yer alan tüm koşulların gerekli olduğuna dair aşağıdaki örnek verilebilir.

Örnek 3.8. X konveks kümesi ve $f(x) = x_1$ fonksiyonu için

$$\begin{aligned} \min\ & f(x) \\ & g(x) = x_2 = 0 \\ & X = \{(x_1, x_2) :\ x_1^2 \leqslant x_2\} \end{aligned}$$

probleminde tek optimal çözüm $x^* = (0, 0)$ ve primal optimal değer $f^* = 0$ dır. Dual fonksiyon

$$q(\mu) = \inf_{x_1^2 \leqslant x_2} \{x_1 + \mu x_2\} = \begin{cases} \inf\limits_{x_1}(x_1 + \mu x_1^2) & ,\ \mu > 0 \\ -\infty & ,\ \mu \leqslant 0 \end{cases} = \begin{cases} \frac{-1}{4\mu} & ,\ \mu > 0 \\ -\infty & ,\ \mu \leqslant 0 \end{cases}$$

bulunur ve $q^* = \sup\limits_{\mu \geqslant 0} q(\mu) = 0$ olur. Dolayısıyla dualite boşluğu yoktur. Lemma 3.5 nedeniyle dual optimal çözüm olmadığından lagrange çarpanı yoktur. Böylece Kabul 3.3 'de verilen izafi iç nokta koşulu sağlanmadığından Teorem 3.13 geçerli değildir.

3.4. FENCHEL DUALİTESİ

$f_1, f_2 : \mathbb{R}^n \to \mathbb{R}$ fonksiyonlar ve $X_1, X_2 \subseteq \mathbb{R}^n$ alt kümeler olmak üzere

$$\begin{aligned} \min\ & f_1(x) - f_2(x) \\ & x \in X_1 \cap X_2 \end{aligned} \tag{3.55}$$

probleminde uygun çözümün var ve f^* optimal değerin sonlu olduğunu kabul edelim. Buna göre

$$\begin{aligned} \min\ & f_1(y) - f_2(z) \\ & z = y \\ & y \in X_1\ ,\ z \in X_2 \end{aligned} \tag{3.56}$$

probleminde dual fonksiyon

$$\begin{aligned} q(\lambda) &= \inf_{y\in X_1, z\in X_2} \{f_1(y) - f_2(z) + \langle \lambda, z - y\rangle\} \\ &= \inf_{z\in X_2} \{\langle \lambda, z\rangle - f_2(z)\} + \inf_{y\in X_1} \{f_1(y) - \langle \lambda, y\rangle\} \end{aligned}$$

olur. Burada $g_1 : \mathbb{R}^n \to (-\infty, +\infty]$ ve $g_2 : \mathbb{R}^n \to [-\infty, +\infty)$ fonksiyonları için

$$g_1(\lambda) = \sup_{x\in X_1} \{\langle \lambda, x\rangle - f_1(x)\} \tag{3.57}$$

$$g_2(\lambda) = \inf_{x\in X_2} \{\langle \lambda, x\rangle - f_2(x)\} \tag{3.58}$$

alınırsa $q(\lambda) = g_2(\lambda) - g_1(\lambda)$ olur. g_1 fonksiyonu, f_1 fonksiyonunun eşlenik konveks fonksiyonudur ve g_2 fonksiyonu, f_2 fonksiyonunun eşlenik konkav fonksiyonudur.

$$\Lambda_1 = \{\lambda :\ g_1(\lambda) < \infty\} \quad , \quad \Lambda_2 = \{\lambda :\ g_2(\lambda) > -\infty\} \tag{3.59}$$

şeklinde tanımlanan kümeler konvekstir. Ayrıca Λ_1 üzerinde g_1 konveks, Λ_2 üzerinde g_2 konkav fonksiyonlardır. Böylece (3.55) probleminin dual problemi

$$\begin{aligned} \text{maks}\ & g_2(\lambda) - g_1(\lambda) \\ & \lambda \in \Lambda_1 \cap \Lambda_2 \end{aligned} \tag{3.60}$$

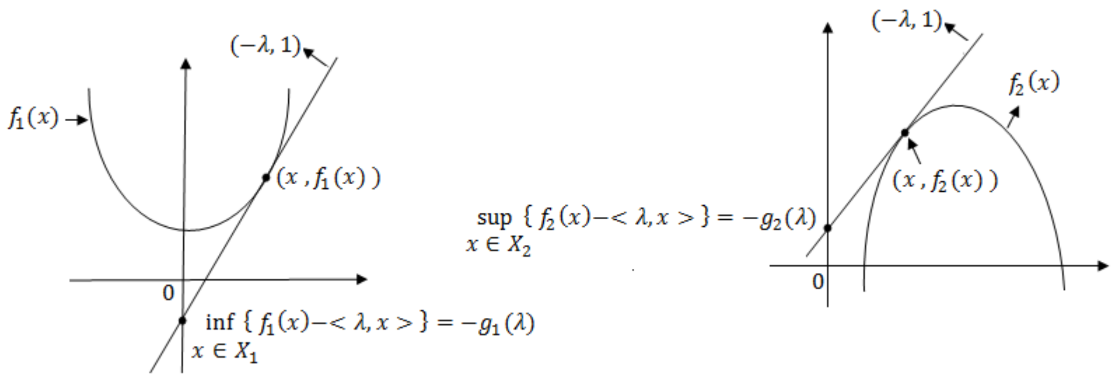
şeklindedir. Dual problem konveks küme üzerinde konkav bir fonksiyonunun maksimum problemidir. Lemma 3.6 nedeniyle (x^*, λ^*) primal ve dual optimal çözüm çifti olabilmesi için gerek ve yeter koşul

$$\begin{aligned} x^* &\in X_1 \cap X_2 \quad \text{(Primal Uygunluk)} \\ \lambda^* &\in \Lambda_1 \cap \Lambda_2 \quad \text{(Dual Uygunluk)} \\ x^* &= \arg \operatorname*{maks}_{x\ \in\ X_1} \{\langle \lambda^*, x\rangle - f_1(x)\} \\ x^* &= \arg \min_{x\in X_2} \{\langle \lambda^*, x\rangle - f_2(x)\} \end{aligned}$$

olmasıdır. Primal optimal çözüm ile dual optimal çözüm arasında dualite boşluğu olmadığını göstermek için öncellikle aşağıdaki kabul verilmektedir.

Kabul 3.4. X_1 konveks kümesi üzerinde f_1 fonksiyonu konveks ve X_2 konveks kümesi üzerinde f_2 konkav fonksiyondur.

Şekil 3.3 'de normali $(-\lambda, 1)$ olan ve $(x, f_1(x))$ noktasından geçen konveks $f_1(x)$ fonksiyonunun $epif_1$ kümesini destekleyen hiperdüzlemin dikey ekseni kestiği nokta $\inf\limits_{x\in X_1}\{f_1(x) - \langle\lambda, x\rangle\} = -g_1(\lambda)$ dır. Benzer şekilde normali $(-\lambda, 1)$ olan ve $(x, f_2(x))$ noktasından geçen konkav $f_2(x)$ fonksiyonunun $hypf_2$ kümesini destekleyen hiperdüzlemin dikey ekseni kestiği nokta $\sup\limits_{x\in X_2}\{f_2(x) - \langle\lambda, x\rangle\} = -g_2(\lambda)$ olarak bulunur.



Şekil 3.3: Destek hiperdüzlemin eşlenik fonksiyonla ilişkisi

Lemma 3.14. [7] Kabul 3.4 geçerli olsun. X_1 ve X_2 polihedral kümeler olmak üzere

i) $ri(X_1) \cap ri(X_2) \neq \emptyset$

ii) f_1, $\mathbb{R}^n$ üzerinde konveks ve f_2, $\mathbb{R}^n$ üzerinde konkav fonksiyon

eğer bu koşullarından biri sağlanıyorsa

$$\begin{gathered}\min\ f_1(x) - f_2(x)\\ x \in X_1 \cap X_2\end{gathered}$$

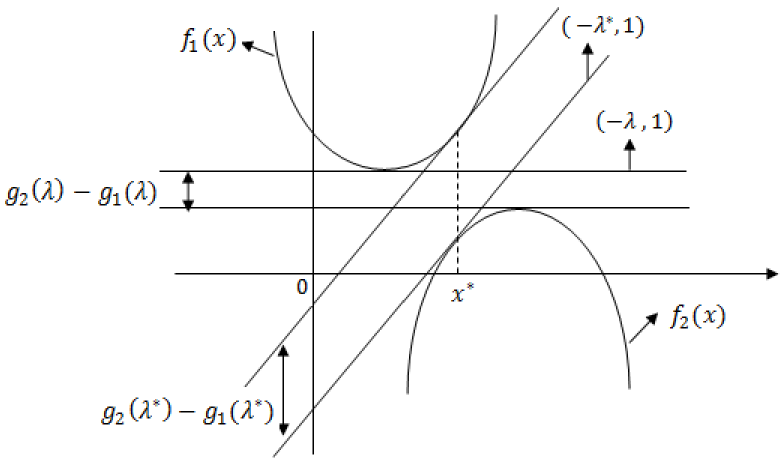
problemi çözüme sahiptir ve dualite boşluğu yoktur. Açıkca,

$$\inf_{x\in X_1\cap X_2}\{f_1(x) - f_2(x)\} = \operatorname*{maks}_{\lambda\,\in\,\Lambda_1\,\cap\,\Lambda_2}\{g_2(\lambda) - g_1(\lambda)\} \tag{3.61}$$

eşitliği gerçeklenir.

f_1 konveks ve f_2 konkav fonksiyonları için (3.55) problemi incelendiğinde bu problem tipik konveks programlama problemidir. Geometrik olarak (3.55) problemi f_1 konveks fonksiyonunun $epif_1$ kümesi ile f_2 konkav fonksiyonunun $hypf_2$ kümesi arasındaki dikey uzaklığı minimum yapma problemidir. Lemma 3.14 'den (3.55) problemi

$epi f_1$ ile $hyp f_2$ kümelerini ayıran birbirlerine paralel destek hiperdüzlemleri arasındaki dikey uzaklığın maksimum olması problemidir. Tüm bu anlatımlar Şekil 3.4 'de gösterilmektedir.



Şekil 3.4: Fenchel dualite teorisi

Fenchel dualitesinde eşlenik fonksiyonlar önemli bir rol oynar. Çünkü, tanım olarak $g_1(\lambda) = f_1^*(\lambda) = \sup_x \{\langle x, \lambda\rangle - f_1(x)\}$ ve $g_2(\lambda) = f_{2*}(\lambda) = \inf_x \{\langle \lambda, x\rangle - f_2(x)\}$ olduğundan Lemma 3.14 nedeniyle

$$\min_{x \in X_1 \cap X_2} \{f_1(x) - f_2(x)\} = \operatorname*{maks}_{\lambda \in \Lambda_1 \cap \Lambda_2} \{f_{2*}(\lambda) - f_1^*(\lambda)\} \tag{3.62}$$

eşitliği gerçeklenir.

4. TARTIŞMA VE SONUÇ

Bu tez çalışmasında konveks küme ve konveks fonksiyonların tanım ve özellikleri verildikten sonra, konveks fonksiyonlar için gerek konveks analizde gerekse optimizasyon teorisinde önemli gelişmeler doğuran subdiferansiyel ile dualite teorisinde büyük işleve sahip olan infimal konvolüsyon kavramları ve Fenchel dualitesinin temelini oluşturan eşlenik fonksiyonların yapısı incelenmiştir. Optimizasyon problemlerinin genel bir sınıflandırılması yapılarak, konveks kümeler ile verilen bir kısıt bölgesinde reel değerli konveks bir fonksiyonun minimazsyon probleminin genel yapısı aşağıdaki gibi verilir.

Konveks Programlama Problemi:
$f : \mathbb{R}^n \to \mathbb{R}$, $g : \mathbb{R}^n \to \mathbb{R}^r$ konveks fonksiyon ve $h : \mathbb{R}^n \to \mathbb{R}^k$ afin fonksiyon, $X \subseteq \mathbb{R}^n$ kısıt kümesi ve $X_0 = \{x \in X : g_j(x) \leqslant 0\ ,\ h_i(x) = 0\}$ uygun çözümler kümesi olmak üzere

$$\begin{aligned} \min\ & f(x) \\ & g_j(x) \leqslant 0\ ,\ h_i(x) = 0 \\ & x \in X \end{aligned} \tag{4.1}$$

şeklinde verilir. (4.1) probleminde her $x \in X_0$ için $f(x^*) \leqslant f(x)$ koşulunu gerçekleyen $x^* \in X_0$ noktasını bulmak amacıyla her konveks optimizasyon problemine karşılık bir dual problem var olduğundan, konveks programlama problemlerinin dual problemi kurularak konveks problemler için çözümler bulunmuştur. Verilen bir primal problem ile onun dual probleminin optimal çözümünün eşitliğinin hangi durumlarda geçerli olduğu Güçlü Dualite Teoremi yardımı ile verilmektedir. Böylece bazı durumlarda asıl problem yerine çözümü daha kolay olan dual problemi göz önüne almak büyük avantaj sağlamaktadır. Öte yandan konveks fonksiyon ile konveks eşlenik fonksiyon arasındaki ilişkiler ile konveks optimizasyon problemlerininin çözümü için Lagrange ve Fenchel Dualite teorileri incelenmiştir. Bu dualite teorileri sayesinde bir takım hesaplama işlemlerinin basitleştirilmesi sağlanmış, klasik yöntemlerle çözümü bulunamayan bazı optimizasyon problemlerinin çözümünün bulunması mümkün olabilmektedir.

KAYNAKLAR

[1] ROCKAFELLAR, R.T. and WETS, J.B. 1998, *Variational Analysis*, Springer, Berlin, 3-540-62772-3.

[2] EKELAND I. and TEMAM R., 1976, *Convex Analysis and Variational Problems*, North-Holland Publishing Company-Amsterdam, 0-7204-2838-6.

[3] ALEKSEEV, V.M., TIKHOMIROV, V.M., and FOMIN, S.V., 1987, *Optimal Control*, Plenum Press, New York, USA, 0-306-10996-4.

[4] MAHMUDOV, E.N. 2011, *Approximation and Optimization of Discrete and Differential Inclusions*, Elsevier, USA, 978-0-12-388428-2.

[5] BERKOVITZ, L.D., 2002, *Convexity and Optimization in* $\mathbb{R}^n$, John Willey & Sons, New York, USA, 0-471-35281-0.

[6] BOYD, S. and VANDENBERGHEL, L., 2004, *Convex Optimization*, Cambridge University Press, New York, USA, 0-521-83378-7.

[7] BERTSEKAS, D.P., 1999, *Nonlinear Programming*, Athena Scientific, Belmont, Massachusetts, USA, 1-886529-00-0.

[8] DANTZIG, G.B. and THAPA, M.N., 2003, *Linear Programming*, Springer, USA, 0-387-98613-8.

[9] JANSEN, B., ROOS, C. and TERLAKY, T., 1993, *The Theory of Linear Programming: Skew Symmetric Self-Dual Problems and The Central Path*, Delft University of Technology, 0922-5641.

[10] SMIRNOV, G.V., 2002, *Introduction to the Theory of Differential Inclusions*, American Mathematical Society, Graduate Studies in Mathematics Vol.41, Rhode Island, USA, 0-8218-2977-7.

[11] TAYLOR, P.D., 1973, *Subgradients of a Convex Function Obtained from a Directional Derivative*, Pacific Journal of Mathematics, Vol.44, No.2 739-747.

[12] DEĞER, Ö., 2010, *On Optimality Conditions for a Convex Optimization Problem with Polyhedral Discrete Inclusions*, University of Istanbul Faculty of Science the Journal of Mathematics, Physics and Astronomy, New Series, Vol.3, 109-118.

[13] ROCKAFELLAR, R.T., 1972, *Convex Analysis*, Princeton University Press, New Jersey, 0-691-08069-0.

[14] WEBSTER, R., 1994, *Convexity* Oxford Science Publications, Oxford University Press, New York, USA, 0-19-853147-8.

[15] ROBERTS, A.W. and VARBERG, D.E., 1973, *Convex Functions*, Academic Press, New York and London, 0-12-589740-5.

[16] BAPTISTE J., URRUTY H. and LEMARÉCHAL C., 2000, *Fundamentals of Convex Analysis*, Springer Verlag Berlin, Heidelberg New York, 3-540-42205-6.

[17] HAMİLTON, C.H., 2005, *Symbolic Convex Analysis*, Simon Fraser University, Spring, CANADA, 0-494-03419-X.

[18] ROCKAFELLAR, R.T., 1963, *Duality Theorems For Convex Functions*, Massachusetts Institute of Technology, Texas University, 189-192.

[19] BOT, R.I., 2010, *Conjugate Duality in Convex Optimization*, Springer, Berlin, 978-3-642-04899-9.

[20] ROCKAFELLAR, R.T., 1974, *Conjugate Duality and Optimization*, University of Washington, Seattle, 0-89871-013-8.

[21] MAHMUDOV, E.N. 2008, *Sufficient conditions for optimality for differential inclusions of parabolic type and duality*, Journal of Global Optimization, 41(1), 31-42.

ÖZGEÇMİŞ

Tokat'ın Almus ilçesinde 17.04.1988 tarihinde doğdum. İlköğretimimi Sakarya'da 1993-2001 yılları arasında Kırkpınar Muazzez Sabri Gündoğan İlköğretim okulunda, ortaöğretimimi Sakarya'da 2001-2005 yılları arasında Ali Dilmen Anadolu Lisesinde tamamladım. 2005-2009 yılları arasında Sakarya Üniversitesi Fen-Edebiyat Fakültesi Matematik bölümünü bitirdim. 2009 yılında başladığım İstanbul Üniversitesi Fen Bilimleri Enstitüsü Matematik bölümünde tezli yüksek lisans eğitimime devam etmekteyim.