T.C.
BALIKESİR ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ
MATEMATİK ANABİLİM DALI

OPTİMİZASYON PROBLEMLERİNE ÜSTEL CEZA FONKSİYONU İLE DİNAMİK SİSTEM YAKLAŞIMI

YÜKSEK LİSANS TEZİ

HÜLYA BOSTAN AYTİMUR

BALIKESİR, HAZİRAN - 2012

T.C.
BALIKESİR ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ
MATEMATİK ANABİLİM DALI

OPTİMİZASYON PROBLEMLERİNE ÜSTEL CEZA FONKSİYONU İLE DİNAMİK SİSTEM YAKLAŞIMI

YÜKSEK LISANS TEZI

HÜLYA BOSTAN AYTİMUR

BALIKESİR, HAZİRAN - 2012

KABUL VE ONAY SAYFASI

Hülya BOSTAN AYTİMUR tarafından hazırlanan "**OPTİMİZASYON PROBLEMLERİNE ÜSTEL CEZA FONKSİYONU İLE DİNAMİK SİSTEM YAKLAŞIMI**" adlı tez çalışmasının savunma sınavı 21.06.2012 tarihinde yapılmış olup aşağıda verilen jüri tarafından oy birliği / ~~oy çokluğu~~ ile Balıkesir Üniversitesi Fen Bilimleri Enstitüsü Matematik Anabilim Dalı Yüksek Lisans Tezi olarak kabul edilmiştir.

Jüri Üyeleri	İmza
Danışman Yrd. Doç. Dr. Fırat EVİRGEN	....
Üye Yrd. Doç. Dr. Necati ÖZDEMİR	....
Üye Yrd. Doç. Dr. Yunus Emre YILDIRIR	....

Jüri üyeleri tarafından kabul edilmiş olan bu tez BAÜ Fen Bilimleri Enstitüsü Yönetim Kurulunca onanmıştır.

Fen Bilimleri Enstitüsü Müdürü

Prof. Dr. Hilmi NAMLI ....

ÖZET

OPTİMİZASYON PROBLEMLERİNE ÜSTEL CEZA FONKSİYONU İLE DİNAMİK SİSTEM YAKLAŞIMI

YÜKSEK LİSANS TEZİ
HÜLYA BOSTAN AYTİMUR
BALIKESİR ÜNİVERSİTESİ FEN BİLİMLERİ ENSTİTÜSÜ
MATEMATİK ANABİLİM DALI
(TEZ DANIŞMANI: YRD. DOÇ. DR. FIRAT EVİRGEN)

BALIKESİR, HAZİRAN - 2012

En basit anlamı ile optimizasyon eldeki kısıtlı kaynakları optimum biçimde kullanmak olarak tanımlanabilir. Matematiksel olarak ifade etmek gerekirse optimizasyon, bir fonksiyonun minimize veya maksimize edilmesidir.

Optimizasyon problemlerini çözmek için günümüzde çeşitli teknikler uygulanmaktadır. Bu tezde; çözüm, ceza fonksiyonunun bir özel çeşidi olan üstel ceza fonksiyonunun kullanılması ile elde edilmiştir. Asıl amaç; doğrusal olmayan bir optimizasyon problemini üstel ceza fonksiyonu ile kısıtsız optimizasyon problemine dönüştürmektir. Elde edilen kısıtsız optimizasyon problemi dinamik sistem modeli kullanılarak çözülmüştür.

Bu tezde ilk olarak optimizasyon ve üstel ceza fonksiyonu ile ilgili literatürde yapılmış çalışmalara yer verilmiştir. Devamında bir optimizasyon probleminin genel özelliklerinden, optimum çözüme sahip olabilmesi için gereken şartlardan bahsedilmiştir. Sonrasında kararlılık ile ilgili temel kavramlara yer verilmiştir. Son olarak ise problemi çözmek için kullanılacak olan üstel ceza fonksiyonunun genel özelliklerine yer verilmiştir. Problem bu fonksiyon yardımıyla kısıtsız bir probleme dönüştürülmüştür. Bu kısıtsız problemi çözmek için dinamik sistem yapısı oluşturulmuştur. Son olarak yöntemin doğruluğu kararlılık analizi ile pekiştirilmiştir. Son kısımlarda ise bahsedilen adımlar uygulanarak nümerik örneklere yer verilmiştir. Bu örneklerde Euler metodu kullanılmış ve çözümler, Matlab programı ile yapılmış ve devamında bahsedilen grafikler elde edilmiştir.

ANAHTAR KELİMELER: optimizasyon problemleri, üstel ceza fonksiyonu, kararlılık, dinamik sistem yaklaşımı

ABSTRACT

DYNAMIC SYSTEM APPROACH FOR OPTIMIZATION PROBLEMS WITH EXPONENTIAL PENALTY FUNCTION
MSC THESIS

HÜLYA BOSTAN AYTIMUR
BALIKESIR UNIVERSITY INSTITUTE OF SCIENCE
MATHEMATICS
(SUPERVISOR: ASSIST. PROF. DR. FIRAT EVİRGEN)

BALIKESİR, JUNE 2012

Optimization with the simplest means can be defined as using restricted resources optimally. Mathematically, optimization is expressed as a maximizing or minimizing of a function.

Various techniques are implemented today for solving optimization problems. In this thesis, solution is obtained with using exponential penalty function which is a special type of penalty function. The main purpose is to convert non-linear optimization problem into unconstrained optimization problem with exponential penalty function. The unconstrained optimization problem obtained is solved by using dynamic system model.

In this thesis, firstly in literature studies done about with exponential penalty function and optimization are mentioned. Secondly general features and conditions are mentioned to have optimal solution of optimization problems. Thirdly the basic concepts related to stability are given. Finally, general properties of the exponential penalty function using for solving optimization problem are presented. The problem is transformed into unconstrained problem with using this function. Dynamic system model is formed to solve this unconstrained optimization problem. In the last section, five numerical examples are given by applying steps mentioned. In these examples, Euler method is used and solutions are done using Matlab programme and graphics are drawn.

KEYWORDS: optimization problems, exponential penalty function, stability, dynamic system approach

İÇİNDEKİLER

Sayfa

ÖZET .... **i**
ABSTRACT .... **ii**
İÇİNDEKİLER .... **iii**
ŞEKİL LİSTESİ .... **iv**
SEMBOL LİSTESİ .... **v**
ÖNSÖZ .... **vi**
1. GİRİŞ **1**
2. OPTİMİZASYON İLE İLGİLİ KAVRAMLAR **4**
2.1 Küme Kısıtlı ve Kısıtsız Optimizasyon 4
2.2 Eşitlik Kısıtlı Optimizasyon 11
2.2.1 Problemin Modellenmesi 12
2.2.2 Tanjant ve Normal Uzaylar 13
2.2.3 Lagrange Koşulları 15
2.3 Eşitsizlik Kısıtlı Optimizasyon 23
2.3.1 Karush-Kuhn-Tucker Koşulları 23
2.4 Ceza Fonksiyonu Metodu 30
3. KARARLILIK ANALİZİ **36**
3.1 Lyapunov Kararlılık ve Dinamik Sistem 36
4. ÜSTEL CEZA FONKSİYONU VE DİNAMİK SİSTEM YAKLAŞIMI **41**
4.1 Üstel Ceza Fonksiyonu 41
4.2 Optimizasyon Problemine Dinamik Sistem Yaklaşımı 43
4.3 Kararlılık Analizi 44
4.4 Nümerik örnekler 46
5. SONUÇLAR **54**
6. KAYNAKLAR **55**

ŞEKİL LİSTESİ

Sayfa

Şekil 4.1: Örnek 1 de $x(t)$'nin grafiği $\left(x^{*}=(15.81,1.581)\right)$ 48
Şekil 4.2: Örnek 2 de $x(t)$'nin grafiği $\left(x^{*}=(0,-3)\right)$....................................49
Şekil 4.3: Örnek 3 de $x(t)$'nin grafiği $\left(x^{*}=(1,1)\right)$..50
Şekil 4.4: Örnek 4 de $x(t)$'nin grafiği $\left(x^{*}=(1,0,0)\right)$...................................52
Şekil 4.5 : Örnek 5 de $x(t)$'nin grafiği $\left(x^{*}=(-1,1,0)\right)$ 53

SEMBOL LİSTESİ

$\mathbb{R}^n$	:	n boyutlu vektör uzayı
Ω	:	$\mathbb{R}^n$' in alt kümesi ve kısıt kümesi veya uygun küme
Df	:	f ' nin birinci mertebeden türevi
∇f	:	f ' nin gradiyenti veya Df ' in transpozu
$F(x)$	:	f ' nin Hessian matrisi
$\dfrac{\partial f}{\partial d}$	:	f ' nin d yönünde yönlü türevi
$Dh(x)$	:	$h(x)$ eşitlik kısıtlarının Jakobiyen matrisi
$\dot{x}(t)=\dfrac{dx(t)}{dt}$	:	$x(t)$' nin t bağımsız değişkenine göre türevi
S	:	Kısıtların tanımladığı yüzey
$T(x^*)$	:	S yüzeyi üzerinde x^* noktasındaki tanjant uzay
$N(x^*)$	:	S yüzeyi üzerinde x^* noktasındaki normal uzay
$l(x,\lambda)$	:	Lagrange fonksiyonu
$L(x,\lambda)$	:	$l,$ Lagrange fonksiyonunun Hessian matrisi
$H_k(x)$	:	$h_k, k=1,\ldots,m$ eşitlik kısıtlarının Hessian matrisi
$J(x^*)$	:	x^* noktasında aktif eşitsizlik kısıtlarının indeks kümesi
γ	:	Ceza parametresi
P	:	Ceza fonksiyonu
$V(x)$	:	Orijini içeren $D\subset\mathbb{R}^n$ bölgesinde tanımlı Lyapunov fonksiyonu
$\Re(A)$	:	A matrisinin sütun uzayı

ÖNSÖZ

Bu tezi hazırlamamda desteklerini benden esirgemeyen, değerli zamanını bana ayırıp ; bilgisini paylaşan ve yol gösteren, değerli hocam ve danışmanım Yrd. Doç. Dr. Fırat EVİRGEN'e ;

Bu tezin alt yapısını şekillendirmekte bana yardımcı olan ve bilgi dağarcığımı genişleten Yrd. Doç. Dr. Necati ÖZDEMİR'e ;

Tez ile ilgili araştırmalarımda bilgisine başvurduğum Balıkesir Üniversitesi Matematik Bölümü öğretim üyelerine ;

Tezi hazırlamakta "2228 SON SINIF LİSANS ÖĞRENCİLERİ İÇİN LİSANSÜSTÜ BURS PROGRAMI" kapsamında her türlü maddi desteği sağlayan TÜBİTAK'a ;

Bu süreçte benden maddi manevi her türlü desteğini esirgemeyen çok sevgili aileme ve canımdan çok sevdiğim eşime sonsuz teşekkürlerimi sunarım.

1. GİRİŞ

Matematikte matematiksel programlama ya da optimizasyon terimi, bir fonksiyonu minimize etmek ya da maksimize etmek amacı ile gerçek ya da tamsayı değerlerini, tanımlı bir aralıkta seçip fonksiyona yerleştirerek sistematik olarak bir problemi incelemek ya da çözmek işlemlerini ifade eder.

Dik iniş adıyla bilinen ilk optimizasyon tekniğinin tarihi Gauss'a dek uzanır. Tarihi olarak, 1940'larda George Dantzig tarafından ortaya atılan lineer programlama kuramı en yaşlı optimizasyon terimidir. Programlama terimi bu bağlamda Bilgisayar Programcılığı'nı ifade etmez. Program teriminin kullanımı ABD Ordusunun, kendi içtimai ve lojistik takvimini belirlemede konteyner kullandığı "program" terimi ile ilişkilidir.

Yapı araç dinamiği'ne ilişkin problemler sıklık ile matematiksel programlama teknikleri gerektirmektedir. Yapı-Araç İskeleti, manifold ile kısıtlanmış bir basit diferansiyel denklem 'in çözümüne ihtiyaç duyan bir yönelim olarak değerlendirilebilir. Bu durumda kısıtlar doğrusal geometrik çeşitliliktedir, örneğin "bu iki nokta daima temas etmeli", "bu alan diğerine etki etmemeli" ya da "bu nokta her zaman bu eğri üzerinde olmalı" gibi. Ayrıca temas halindeki kuvvetlere ilişkin problemler de doğrusal uyumluluk çatısı altında çözüldüğünden, buna da bir tür QP (Kuadratik Programlama) problemi gözüyle bakılabilir.

Pek çok dizayn problemi de optimizasyon programları ile çözülmektedir. Bu tür uygulamalara dizayn optimizasyonu denir. Bu alanda bilinen ve büyümekte olan bir alt kol çok disiplinli dizayn optimizasyonudur. Bu tür, pek çok problemde kullanışlı olduğu gibi aynı zamanda da uzay mühendisliği sahasına uyarlanabilmektedir.

Ekonomi de matematiksel programlamaya ağır bir bağımlılık duyar. Mikroiktisat' da sık karşılaşılan bir problem olan marjinal fayda ve bundan kaynaklanan ikilik olan harcamaları minimize etme problemi iktisadî bir optimizasyon problemidir. Tüketiciler ve firmalar fayda/kar oranlarını maksimize etmek durumundadırlar. Ticaret teorisi de milletler arası ticari ortaklığın izahında optimizasyona sık sık başvurur.

Sabit genel giderli zaman maliyet problemleri de önemli bir optimizasyon problemidir. Özellikle inşaat ve endüstri mühendisliğinde bu problemle sıkça karşılaşılır. Doğrusal programlama, sezgisel ve üst sezgisel yöntemler bu problem türünün optimizasyonu için kullanılmaktadır. Optimizasyon tekniklerinin sıkça kullanıldığı bir diğer alan da operasyon araştırmasıdır.

İşçi maliyetlerini, girdi maliyetlerini minimize etme, üretimi miktarları ile ilgili olarak kârı, yatırım araçlarında en fazla getiriyi sağlamak gibi iş hayatında akla gelen pek çok işi yapmanın en iyi yolu bulunmak isteniyor. Bu en iyi yol bazen işlevi maksimize kılan yol, bazen işlevi minimize eden yol veya bazen sıfır kılan yoldur. Birçok işletme ve ekonomi sorunlarında, özel veya kamu sektöründe devamlı kullanılmaktadır. Nakliyat, enerji üretimi ve dağıtımı, telekomünikasyon, sınai üretim gibi teknik işletmecilik gerektiren alanlarda bulunan bir çok firma optimizasyonu çok kullanmaktadır. Planlama, zaman programlaması, iş ve işçi tahsis edilmesi, fireyi azaltma, finansal getiriyi maksimize etme gibi birçok sorun optimizasyon ile çözülebilir.

Sabit genel giderli zaman maliyet problemleri, önemli bir optimizasyon problemidir. Özellikle inşaat ve endüstri mühendisliğinde bu problemle sıkça karşılaşılır.

Optimizasyon problemlerinin kullanım alanları kadar çözüm metotları da oldukça önemlidir. Problemi çözmek için uzun yıllardır birçok metot geliştirilmiştir [1-3]. Bunlardan bir tanesi de ceza fonksiyonu kullanarak ceza metodu ile verilen problemi çözmektir [4,5]. Son yıllarda, ceza fonksiyonu kullanarak lineer olmayan matematiksel programlama problemlerini çözmek için çeşitli metotlar verilebilir [6,7]. Genel strateji; kısıtlı lineer olmayan matematiksel programlama problemini ceza fonksiyonunun kısıtsız minimizasyonunun bir dizisine dönüştürmektir. Bu ceza fonksiyonları, orijinal optimizasyon probleminde kısıtsız optimal çözüm dizisinin kısıtlı optimal çözüme yaklaşmasıyla kurulur [8].

Ceza metodunun bir alt sınıfı Motzkin [9] tarafından tanıtılan üstel ceza yaklaşımıdır. Üstel ceza metodu ile, Motzkin'in çalışmasından bu yana, geniş ölçüde çalışılmıştır [9-15]. Murphy, [14] makalesinde üstel ceza fonksiyonlarını sınıflandırmış, farklı Lagrange çarpanları denemiştir.

Kısıtlı optimizasyon problemini kısıtsız optimizasyon problemine dönüştürmek için tezde kullanılan metot literatürde yapılan araştırmalar ile kısaca bahsedilmiş oldu. Şimdi de kısıtsız optimizasyon problemini çözmek için kullanılan metotlardan bahsedilecektir. Bunun için de kitaplar da birçok metot bulunabilir [1-3].

Bu tezde ise dinamik sistem yapısıyla çözüm ele alınacaktır. Bununla ilgili de çeşitli çalışmalar yapılmıştır [16]. Örneğin Özdemir ve Evirgen [6] makalesinde quadratik programlama problemi için ceza fonksiyonu kullanmış ve probleme dinamik sistem yapısı ile yaklaşmıştır. Dinamik sistem yapısının bir çeşidi de sinir ağı (neural network) modeliyle çözmektir. Bu yöntem de bir çeşit dinamik sistem yapısına karşılık gelir [17-20]. Örneğin Zhou ve Zhang, [20] makalesinde sinir ağı için gerekli olan bir aktivasyon fonksiyonu kullanmış, bu fonksiyon yardımıyla kısıtsız probleme geçiş yapmış ve sinir ağı yapısını oluşturmuştur.

Tezin ikinci bölümünde genel bir optimizasyon problemi ile ilgili temel kavramlar verilmiştir. Optimizasyon problemlerinin çözümü için gereken çeşitli teoremlerden ve optimallik koşullarından bahsedilmiştir. Ceza metoduna genel bir bakış açısı sağlamak için bir ceza fonksiyonunun yapısına ve özelliklerine değinilmiştir.

Tezin üçüncü bölümünde bir optimizasyon probleminin kararlılığını incelemek için kararlılık ile ilgili genel bilgiler verilmiştir.

Tezin dördüncü bölümünde ise üstel ceza fonksiyonunun genel bir yapısından ve sağlaması gereken özelliklerden bahsedilmiştir. Problemin çözümü yapılmış ve dinamik sistem yapısı oluşturulmuş, verilen probleme bu dinamik sistem yapısı ile yaklaşılmıştır. Devamında ise tezde geniş ölçüde yer verilerek araştırılan metot kullanılarak çözümü yapılan birkaç nümerik örneğe yer verilmiştir. Bu örnekler çözülürken Euler metodu kullanılmış ve metodun nümerik açıdan verdiği sonucu görmek için Matlab programından yararlanılmıştır ve bu örneklerin grafikleri problemleri takiben sunulmuştur.

2. OPTİMİZASYON İLE İLGİLİ KAVRAMLAR

Matematiksel programlama problemlerinde, çoğu kez sadece optimum yada global çözümlerle ilgilenilmesine rağmen genelde sadece lokal çözümler hakkında teoremler ispatlanabilir.

$$\begin{array}{lll} \text{Minimum} & f(x) & \\ \text{Kısıtlar} & g_i(x) \leq 0 & ,i = 1,...,m \\ & h_j(x) = 0 & ,j = 1,...,p \end{array} \tag{2.1}$$

Lokal çözüm, amaç fonksiyonunun daha küçük bir değerini veren, kısıtları sağlayan, komşuluğunda herhangi bir başka nokta bulunmayan bir noktadır. Global çözüm, en küçük amaç fonksiyonu değerini üreten herhangi bir lokal çözüm olarak tanımlanır. Eğer problem fonksiyonlarının özel özellikleri varsa, örneğin eğer konveks programlama problemi tanımlanıyorsa, her lokal çözümün bir global çözüm olduğunu ispatlamak mümkün olabilir. Problem (2.1) için lokal çözüm olan noktanın yeter koşullarının kümesini ve gerek koşullarının kümesini şart koşan yararlı teoremler bu önemli özelliklere başvurmaksızın geliştirilebilirler [8].

2.1 Küme Kısıtlı ve Kısıtsız Optimizasyon

Bu bölümde

$$\begin{array}{ll} \text{Minimum} & f(x) \\ \text{Kısıtlar} & x \in \Omega \end{array} \tag{2.2}$$

optimizasyon problemi ele alınacaktır.

$f : \mathbb{R}^n \to \mathbb{R}$ minimize edilmesi amaçlanan fonksiyon objektif fonksiyon yada amaç fonksiyonu olarak adlandırılan reel değerli bir fonksiyondur. x, değişkenlerinden bağımsız $x = [x_1, x_2, ..., x_n]^T \in \mathbb{R}^n$ n boyutlu vektördür. $x_1, x_2, ..., x_n$ değişkenleri karar değişkenleri olarak ifade edilir. $\mathbb{R}^n$'in alt kümesi olan Ω kısıt kümesi veya uygun küme olarak adlandırılır [1].

(2.2) optimizasyon problemi, Ω da ki bütün olası vektörler üzerinde karar değişkenlerinin en iyi x vektörünü bulmayı içeren bir karar problemi olarak görülebilir. En iyi vektörlerden kastedilen, amaç fonksiyonunun en küçük değerini bulmaktır. Bu vektör Ω da f 'nin minimumu olarak adlandırılır. Birçok minimum nokta olması mümkündür. Burada herhangi birini bulmak yeterli olacaktır [1].

Maksimumu araştırılan, amaç fonksiyonunun maksimumunu içeren optimizasyon problemleri de vardır. Ancak maksimum problemleri, f 'nin maksimumu $-f$ 'nin minimumuna eşit olduğu için yukarıdaki minimum formuna eşit olarak ifade edilebilir. Sonuç olarak genelliği kaybetmeksizin minimum problemleri ile inceleme yapılabilir [1].

(2.2) problemi kısıtlı optimizasyon probleminin genel formudur. Çünkü karar değişkenleri kısıt kümesi Ω da sınırlandırılmıştır. Eğer $\Omega = \mathbb{R}^n$ ise problem kısıtsız optimizasyon problemi olarak adlandırılır.

"$x \in \Omega$" kısıtı küme kısıtı olarak adlandırılır. Sıklıkla, Ω kısıt kümesi h ve g verilen fonksiyonlar olmak üzere

$$\Omega = \{x : h(x) = 0, g(x) \le 0\},$$

formunda verilir [1].

Yukarıda ki genel optimizasyon problemi düşünüldüğünde iki çeşit minimum noktası arasından seçim yapmak, aşağıdaki tanımlar yoluyla belirtilecektir.

Tanım 2.1.1: $f : \mathbb{R}^n \to \mathbb{R}$, $\Omega \subset \mathbb{R}^n$ kümesinde reel değerli bir fonksiyon olsun.

$\forall x \in \Omega \setminus \{x^*\}$ için $\varepsilon > 0$ vardır öyle ki $\|x - x^*\| < \varepsilon$ için $f(x^*) \le f(x)$ olması durumunda $x^* \in \Omega$, Ω da f 'nin lokal minimumudur. Eğer $\forall x \in \Omega \setminus \{x^*\}$ için $f(x^*) \le f(x)$ oluyorsa $x^* \in \Omega$, Ω da f 'nin global minimumudur.

Eğer yukarıda ki tanımda "$\le$" yerine "$<$" kullanılırsa tam lokal minimum ve tam global minimum denir [1].

Eğer x^*, f 'nin Ω 'da global minimumu ise $f(x^*) = \min_{x\in\Omega} f(x)$ ve $x^* = \arg\min_{x\in\Omega} f(x)$ yazılır. Eğer minimum kısıtsız ise basitçe $x^* = \arg\min_{x} f(x)$ veya $x^* = \arg\min f(x)$ yazılabilir [1].

Lokal Minimum İçin Şartlar

Bu bölümde x^*'ın local minimum olması için şartlar türetilecek ve $f:\mathbb{R}^n \to \mathbb{R}$ fonksiyonunun türevleri kullanılacaktır.

Df ile tanımlanan f 'nin birinci basamaktan türevi

$$Df = \left[\frac{\partial f}{\partial x_1}, ..., \frac{\partial f}{\partial x_n}\right]$$

şeklindedir. Gradiyent ∇f, sadece Df 'nin transpozudur, yani $\nabla f = (Df)^T$. $f:\mathbb{R}^n \to \mathbb{R}$'nin ikinci basamaktan türevi (f 'nin Hesse matrisi)

$$F(x) \triangleq D^2 f(x) = \begin{pmatrix} \dfrac{\partial^2 f(x)}{\partial x_1{}^2} & \cdots & \dfrac{\partial^2 f(x)}{\partial x_n \partial x_1} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial^2 f(x)}{\partial x_1 \partial x_n} & \cdots & \dfrac{\partial^2 f(x)}{\partial x_n{}^2} \end{pmatrix}$$

dir. Ω kısıt kümesi ile verilen bir optimizasyon probleminde minimum , Ω'nın sınırında ya da içinde olabilir. Sınırda olduğu durumu çalışmak için uygun yön (feasible direction) notasyonuna gereksinim duyulur [1].

Tanım 2.1.2: Eğer $\alpha_0 > 0$ vardır öyle ki $\forall\alpha \in [0, x_0]$ için

$$x + \alpha d \in \Omega$$

ise $d \in \mathbb{R}^n$, $d \neq 0$ vektörü $x \in \Omega$'da uygun (feasible) yöndür.

$f : \mathbb{R}^n \to \mathbb{R}$ reel değerli fonksiyon ve d, $x \in \Omega$'da uygun (feasible) yön olsun. $\partial f / \partial d$ olarak tanımlanan f'nin d yönünde yöne göre türevi

$$\frac{\partial f}{\partial d}(x) = \lim_{\alpha \to 0} \frac{f(x + \alpha d) - f(x)}{\alpha}$$

şeklinde tanımlanan reel değerli bir fonksiyondur. Eğer $\|d\| = 1$ ise $\partial f / \partial d$, d yönünde x noktasında f'nin artış oranıdır.

Yukarıdaki yönlü türevi hesaplamak için, x ve d'nin verildiği farz edelirse $f(x + \alpha d)$, α'nın bir fonksiyonudur ve

$$\frac{\partial f(x)}{\partial d} = \frac{d}{d\alpha} f(x + \alpha d)\Big|_{\alpha = 0} = \nabla f(x)^T d = \langle \nabla f(x), d \rangle d^T \nabla f(x).$$

Özet olarak eğer d birim vektör $(\|d\| = 1)$ ise $\langle \nabla f(x), d \rangle$, f'nin x noktasında d yönünde artış oranıdır.

Teorem 2.1.3 (Birinci Basamaktan Gerek Koşul**) :** Ω, $\mathbb{R}^n$ in bir alt kümesi ve $f \in C^1$, Ω'da reel değerli fonksiyon olsun. Eğer x^* , Ω da f'nin lokal minimumu ise x^*'da herhangi bir uygun (feasible) yön için

$$d^T \nabla f(x^*) \geq 0$$

yazılabilir [1].

İspat: $x(\alpha) = x^* + \alpha d \in \Omega$ tanımlanırsa , $x(0) = x^*$ dır.

Bileşke fonksiyonu $\phi(\alpha) = f(x(\alpha))$ olarak tanımlanır ise Taylor teoreminden;

$$f\left(x^*+\alpha d\right)-f\left(x^*\right)=\phi(\alpha)-\phi(0)=\phi'(0)\alpha+o(\alpha)=\alpha d^T\nabla f\left(x(0)\right)+o(\alpha),\ \alpha\geq 0$$

elde edilir. Sonuç olarak eğer $\phi(\alpha)\geq\phi(0)$ yani $\alpha>0$ 'ın yeteri kadar küçük değeri için $f\left(x^*+\alpha d\right)\geq f\left(x^*\right)$ (x^* lokal minimum olduğu için) ise $d^T\nabla f\left(x^*\right)\geq 0$ elde edilir.

Tüm uygun d yönleri için yukarıdaki teoremin başka bir alternatif yolu

$$\frac{\partial f}{\partial d}\left(x^*\right)\geq 0$$

dir. Diğer bir deyişle x^* bir lokal minimum ise f 'nin herhangi bir d yönünde x^* noktasındaki artış oranı negatif olmayandır. Yönlü türev kullanılarak teorem 2.1.1 in alternatif bir ispatı aşağıdaki gibi verilebilir. x^* bir lokal minimum olsun. Bu durumda herhangi bir d uygun (feasible) yönü için $\bar{\alpha}>0$ vardır ki $\forall\alpha\in\left(0,\bar{\alpha}\right)$ için

$$f\left(x^*\right)\leq f\left(x^*+\alpha d\right)$$

yazılabilir.

Bunun sonucunda $\forall\alpha\in\left(0,\bar{\alpha}\right)$ için

$$\frac{f\left(x^*+\alpha d\right)-f\left(x^*\right)}{\alpha}\geq 0$$

elde edilir.

$\alpha\to 0$ için limit alınırsa

$$\frac{\partial f}{\partial d}\left(x^*\right)\geq 0$$

sonucuna varılabilir .

x^* 'ın Ω 'nın iç noktası olduğu durum, özel bir durumdur. Buna göre, herhangi bir yön uygundur ve aşağıdaki sonuç verilebilir.

Sonuç 2.1.4 (İç nokta durumu**) :** Ω, $\mathbb{R}^n$ in alt kümesi ve $f \in C^1$, Ω da reel değerli fonksiyon olsun. Eğer x^*, Ω'da f'nin bir lokal minimumu ve Ω'nın bir iç noktası ise

$$\nabla f\left(x^*\right)=0$$

dır [1].

İspat : Ω'nın iç noktası olan x^*, f'nin lokal minimumu olsun. x^*, Ω'nın iç noktası olduğu için x^* 'daki uygun yönlerin kümesi tüm $\mathbb{R}^n$'dir. Sonuç olarak herhangi $d \in \mathbb{R}^n$ için $d^T\nabla f\left(x^*\right)\geq 0$ ve $-d^T\nabla f\left(x^*\right)\geq 0$ dır. Bunun sonucunda $\forall d \in \mathbb{R}^n$ için $d^T\nabla f\left(x^*\right)=0$ bulunur ve burada $\nabla f\left(x^*\right)=0$ olduğunu gösterir.

Teorem 2.1.5 (İkinci Basamaktan Gerek Koşul**) :** $\Omega \in \mathbb{R}^n$, $f \in C^2$ Ω'da bir fonksiyon, x^*, Ω etrafında f'nin bir lokal minimumu ve d, x^*'da uygun yön olsun. Eğer $d^T\nabla f\left(x^*\right)=0$ ise;

$$d^T F\left(x^*\right)d \geq 0$$

dır. Burada F, f'nin Hesse matrisidir [1].

İspat: Teorem olmayana ergi yöntemiyle ispatlanır. x^*'da d uygun yönü vardır öyle ki $d^T F\left(x^*\right)d=0$ ve $d^T F\left(x^*\right)d<0$ dır. $x\left(\alpha\right)=x^*+\alpha d$ olsun ve $\phi\left(\alpha\right)=f\left(x^*+\alpha d\right)=f\left(x\left(\alpha\right)\right)$ bileşke fonksiyonu tanımlanırsa Taylor teoreminden

$$\phi\left(\alpha\right)=\phi\left(0\right)+\phi''\left(0\right)\frac{\alpha^2}{2}+o\left(\alpha^2\right)$$

$$\phi'\left(0\right)=d^T F\left(x^*\right)d=0 \text{ ve } \phi''\left(0\right)=d^T\nabla F\left(x^*\right)d<0$$

yazılır. Yeterince küçük α için

$$\phi\left(\alpha\right)-\phi\left(0\right)=+\phi''\left(0\right)\frac{\alpha^2}{2}+o\left(\alpha^2\right)<0$$

bulunur. Yani $f\left(x^*+\alpha d\right)< f\left(x^*\right)$ ki bu da x^*'ın lokal minimum olmasıyla çelişir. Sonuç olarak

$$\phi''(0)=d^T F\left(x^*\right)d\geq 0$$

elde edilir.

Sonuç 2.1.6 (İç nokta durumu) **:** x^*, $\Omega\subset\mathbb{R}^n$ in iç noktası olsun. Eğer x^*, $f:\Omega\to\mathbb{R}$ $f\in C^2$ fonksiyonun bir lokal minimumu ise $\nabla f\left(x^*\right)=0$ ve $F\left(x^*\right)$ pozitif yarı tanımlıdır $\left(F\left(x^*\right)\geq 0\right)$ yani $\forall d\in\mathbb{R}^n$ için $d^T F\left(x^*\right)d\geq 0$ dir [1].

İspat : x^* bir iç nokta ise, bütün yönler uygundur. İspat, Sonuç 2.1.1 ve Teorem 2.1.2 den elde edilir.

Şimdi x^*'ın lokal minimum olduğunu söyleyen yeter koşulları belirlensin.

Teorem 2.1.7 (İkinci Basamaktan Yeter Koşul**):** $f\in C^2$, x^*' ın bir iç nokta olduğu bölgede tanımlansın ve

1. $\nabla f\left(x^*\right)=0$

2. $F\left(x^*\right)>0$

olsun. Öyle ise x^*, f ' nin tam lokal minimumudur [1].

İspat: $f\in C^2$ olduğu için $F\left(x^*\right)=F^T\left(x^*\right)$ yazılabilir. İkinci şartı kullanırsak ve Raylrigh's eşitsizliğini takiben eğer $d\neq 0$ ise $0<\lambda_{\min}\left(F\left(x^*\right)\right)\left\|d\right\|^2\leq d^T F\left(x^*\right)d$ bulunur. Taylor teoremi ve birinci şarttan;

$$f\left(x^*+d\right)-f\left(x^*\right)=\frac{1}{2}d^T F\left(x^*\right)d+o\left(\left\|d\right\|^2\right)\geq\frac{\lambda_{\min}\left(F\left(x^*\right)\right)}{2}\left\|d\right\|^2+o\left(\left\|d\right\|^2\right)$$

yazılabilir. Sonuç olarak $\left\|d\right\|$ yeterince küçük tüm d 'ler için

$$f\left(x^*+d\right)>f\left(x^*\right)$$

elde edilir ki bu da ispatı tamamlar.

2.2 Eşitlik Kısıtlı Optimizasyon

Bu bölümde (2.2) ile formüle edilmiş lineer olmayan kısıtlı optimizasyon problemlerinin bir sınıfını çözmek için metotlar elde edilecek.

Vektör formunda (2.1) problemi aşağıdaki standart formda yeniden gösterilebilir [1]

$$\begin{array}{ll} \text{Minimum} & f(x) \\ \text{Kısıtlar} & g(x) \leq 0 \\ & h(x) = 0 \quad . \end{array} \tag{2.3}$$

Burada $h:\mathbb{R}^n \to \mathbb{R}^m$ ve $g:\mathbb{R}^n \to \mathbb{R}^p$ dir.

Tanım 2.2.1: Kısıtları sağlayan herhangi bir noktaya uygun nokta adı verilir. Uygun noktaların kümesi

$$\left\{x \in \mathbb{R}^n : h(x) = 0, g(x) \leq 0\right\}$$

uygun küme olarak adlandırılır.

Optimizasyon problemlerinin (2.3) formu bizim için yeni değildir. Gerçekten, lineer programlama problemleri formu

$$\begin{array}{ll} \text{Minimum} & c^T x \\ \text{Kısıtlar} & Ax = b \\ & x \geq 0 \end{array} \tag{2.4}$$

şeklindedir [1].

Bu bölümün diğer kısmında sadece eşitlik kısıtlarına sahip kısıtlı optimizasyon problemleri incelenecek ve genel eşitlik kısıtlı optimizasyon problemi aşağıdaki bölümde verilecektir.

2.2.1 Problemin Modellenmesi

Bu bölümde analiz edilecek optimizasyon problemleri sınıfı

$$\begin{array}{ll} \text{Minimum} & f(x) \\ \text{Kısıtlar} & h(x)=0 \end{array} \tag{2.5}$$

şeklindedir ki burada $x \in \mathbb{R}^n$, $f:\mathbb{R}^n \to \mathbb{R}$, $h:\mathbb{R}^n \to \mathbb{R}^m$ $h=[h_1,...,h_m]^T$ ve $m \le n$ dir. h sürekli türevlenebilir bir fonksiyon yani $h \in C^1$ olsun. Aşağıdaki tanım verilebilir.

Tanım 2.2.1.1: Eğer $\nabla h_1(x^*),...,\nabla h_m(x^*)$ gradiyent vektörleri lineer bağımsız ise $h_1(x^*)=0,...,h_m(x^*)=0$ kısıtlarını sağlayan x^* noktasına kısıtların regüler bir noktasıdır denir. $h=[h_1,...,h_m]^T$ 'ın x^*'daki ,

$$Dh(x^*)=\begin{pmatrix} Dh_1(x^*) \\ \vdots \\ Dh_m(x^*) \end{pmatrix}=\begin{pmatrix} \nabla h_1(x^*)^T \\ \vdots \\ \nabla h_m(x^*)^T \end{pmatrix}$$

şeklinde verilen Jakobiyen matrisi $Dh(x^*)$ olsun. Bu durumda x^* noktasının regüler bir nokta olması için gerek ve yeter şart rank $Dh(x^*)=m$ olmasıdır. $h_1(x)=0,...,h_m(x)=0$, $h_i:\mathbb{R}^n \to \mathbb{R}$ eşitlik kısıtlarının kümesi

$$S=\{x \in \mathbb{R}^n : h_1(x)=0,...,h_m(x)=0\}$$

şeklinde yüzey tanımlar.

S 'nin noktaları regüler olsun. O halde S yüzeyinin boyutu $n-m$ dir.

2.2.2 Tanjant ve Normal Uzaylar

Bu bölümde, yüzeydeki bir nokta da normal uzay ve tanjant uzay notasyonları verilecektir. S yüzeyinde bir eğri tanımlanarak başlanacaktır.

Tanım 2.2.2.1: S yüzeyinde bir eğri, sürekli $t \in (a,b)$ parametrize edilmiş $\{x(t) \in S : t \in (a,b)\}$ noktalarının kümesidir yani $x:(a,b) \rightarrow S$ sürekli fonksiyonudur .

Tanım 2.2.2.2: Eğer; $\forall t \in (a,b)$ için

$$\dot{x}(t) = \frac{dx(t)}{dt} = \begin{pmatrix} \dot{x}_1(t) \\ \vdots \\ \dot{x}_n(t) \end{pmatrix}$$

var ise $C = \{x(t) : t \in (a,b)\}$ türevlenebilir eğridir.

Tanım 2.2.2.3: $S = \{x \in \mathbb{R}^n : h(x) = 0\}$ yüzeyi üzerinde $x*$ noktasındaki tanjant uzay

$$T(x*) = \{y : Dh(x*)y = 0\}$$

dir.

$T(x*)$ tanjant uzayı, $D(h(x*))$ matrisinin sıfır uzayıdır

$$T(x*) = N(Dh(x*)).$$

Tanjant uzay $\mathbb{R}^n$ in alt uzayıdır. $x*$'ın regüler olduğu farz edilirse, tanjant uzayının boyutu $n-m$'dir ve burada m , $h_i(x*) = 0$ eşitlik kısıtlarının sayısıdır.

Teorem 2.2.2.4: $x* \in S$ regüler nokta ve $T(x*)$, $x*$ noktasında tanjant uzay olsun. Bu durumda $y \in T(x*)$ olması için gerek ve yeter şart S üzerinde bir $x*$

noktasından geçen ve bu noktadaki diferansiyeli y olan türevlenebilir bir eğrinin bulunmasıdır [1,2].

İspat: $(\Leftarrow:)$ S üzerinde

$$\{x(t) : t \in (a,b)\}$$

eğrisinin var olduğunu düşünelim öyle ki $x(t^*) = x^*$ ve $\dot{x}(t^*) = y$, $t^* \in (a,b)$ olsun. Öyle ise

$$h(x(t)) = 0, \ \forall t \in (a,b) \ \text{için}$$

zincir kuralından $h(x(t))$ fonksiyonunun t'ye göre türevi alınırsa

$$\frac{d}{dt} h(x(t)) = Dh(x(t))\dot{x}(t) = 0, \ \forall t \in (a,b)$$

elde edilir. Böylece , t^*'da

$$Dh(x^*)y = 0 \text{ ve } y \in T(x^*)$$

bulunur .

$(\Rightarrow:)$ Bu tarafı ispatlamak için kapalı fonksiyon teoreminin kullanılması gerekir [2].

Aşağıda normal uzay notasyonunun tanımı verilmiştir.

Tanım 2.2.2.5: $S = \{x \in \mathbb{R}^n : h(x) = 0\}$ yüzeyi üzerinde bir x^* noktasındaki $N(x^*)$ normal uzayı

$$N(x^*) = \{x \in \mathbb{R}^n : \ x = Dh(x^*)^T z, \ \ z \in \mathbb{R}^m\}$$

kümesidir.

$N(x^*)$ normal uzayını

$$N(x^*) = \Re\left(Dh(x^*)^T\right)$$

olarak ifade edebiliriz, yani $Dh(x^*)^T$ matrisinin görüntüsüdür. $N(x^*)$ normal uzayı $\nabla h_1(x^*),...,\nabla h_m(x^*)$ vektörleri tarafından gerilen $\mathbb{R}^n$ 'in bir alt uzayıdır yani

$$\begin{aligned} N(x^*) &= span\left[\nabla h_1(x^*),...,\nabla h_m(x^*)\right] \\ &= \left\{x \in \mathbb{R}^n :\ x = z_1\nabla h_1(x^*) + ... + z_m\nabla h_m(x^*),\ \ z_1,...,z_m \in \mathbb{R}\right\} \end{aligned}$$

dir. Normal uzay sıfır vektörünü içerir. x^*'ın regüler olduğu kabul edilirse , $N(x^*)$ normal uzayının boyutu m 'dir [1].

Şimdi normal uzay ve tanjant uzayın birbirlerinin ortogonal tamamlayıcıları oldukları ispatsız verilecektir.

Lemma 2.2.2.6: $T(x^*) = N(x^*)^{\perp}$ ve $T(x^*)^{\perp} = N(x^*)$ dir [1].

2.2.3 Lagrange Koşulları

Bu bölümde kısıtlı ekstremum problemleri için birinci basamaktan bir yeter koşul sunulacaktır. Bu sonuçlar bilinen Lagrange teoremidir. Teoremin daha iyi anlaşılması için iki değişkenli kısıtlı bir problem ele alınacaktır [1].

$h:\mathbb{R}^2 \to \mathbb{R}$ kısıt fonksiyonu olsun. Her bir x noktasındaki $\nabla h(x)$ gradiyent vektörü bu noktadan geçen seviye kümesine dik olmalıdır. Şimdi $x^* = (x_1^*, x_2^*)^T$ için $h(x^*) = 0$ ve $\nabla h(x^*) \neq 0$ kabul edilsin. x^* noktasındaki seviye kümesi $\{x : h(x) = 0\}$ kümesi ile tanımlanır. Bu seviye kümesi x^*'ın bir komşuluğunda sürekli türevlenebilir $\{x(t)\}, x:\mathbb{R} \to \mathbb{R}^2$ eğrisi ile parametrize edilir. Yani

$x(t) = \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix},\ t \in (a,b),\ x^* = x(t^*),\ \dot{x}(t^*) \neq 0,\ t^* \in (a,b)$ olsun. Şimdi $\nabla h(x^*)$'ın $\dot{x}(t^*)$'a dik olduğu gösterilecektir. $\forall t \in (a,b)$ için

$$h(x(t)) = 0 \Rightarrow \frac{d}{dt} h(x(t)) = 0$$
$$\Rightarrow \nabla h(x(t))^T \dot{x}(t) = 0$$

yani $\nabla h(x^*)$, $\dot{x}(t^*)$'a diktir. Şimdi x^* noktasının $f:\mathbb{R}^2 \to \mathbb{R}$ fonksiyonun $\{x : h(x) = 0\}$ kümesi üzerinde minimumu olduğu kabul edilsin. Burada $\nabla f(x^*)$ ‘ın da $\dot{x}(t^*)$'a dik olduğu gösterilmeye çalışılacaktır. Bunun için t^* noktasında minimumu olan $\phi(t) = f(x(t))$ bileşke fonksiyonu ele alınırsa x^* minimum olduğundan birinci basamaktan gerek koşul gereği

$$0 = \frac{d}{dt}\Big|_{t^*} = \nabla f(x(t^*))\dot{x}(t^*) = \nabla f(x^*)^T \dot{x}(t^*)$$

olur. Buradan da $\nabla f(x^*)$'ın, $\dot{x}(t^*)$'a dik olduğu görülür. Sonuç olarak $\nabla h(x^*)$ da $\dot{x}(t^*)$'a dik olduğundan $\nabla h(x^*)$'ın, $\nabla f(x^*)$'a paralel, yani $\nabla f(x^*)$, $\nabla h(x^*)$'ın bir skaler katıdır [1].

Yukarıdaki gözlemler tek kısıt ile iki değişkenli fonksiyonlar için Lagrange teoreminin formüle edilmesini sağlar.

Teorem 2.2.3.1 (**Lagrange Teoremi n=2, m=1**)**:** $x^*, h(x) = 0, h:\mathbb{R}^2 \to \mathbb{R}$ kısıtına göre $f:\mathbb{R}^2 \to \mathbb{R}$ fonksyionunun minimumu olsun. Bu durumda $\nabla f(x^*)$ ve $\nabla h(x^*)$ paraleldir, yani eğer, $\nabla h(x^*) \neq 0$ ise öyle bir λ^* skaleri vardır ki

$$\nabla f(x^*) + \lambda^* \nabla h(x^*) = 0$$

dir [1].

2.2.3.1 teoreminde λ^*, Lagrange çarpanı olarak adlandırılır. Teorem maksimum ekstremum problemleri içinde sağlanır.

Lagrange teoremi, bir noktanın lokal minimum olması için birinci basamaktan gerek koşulu sağlar. Bu koşul, iki denklemden oluşan Lagrange koşulu olarak bilinir,

$$\nabla f(x^*) + \lambda^* \nabla h(x^*) = 0$$

$$h(x^*) = 0 \; .$$

Lagrange koşulu gerek şarttır fakat yeter değildir. Şimdi Lagrange teoremi $f:\mathbb{R}^n \to \mathbb{R}$ ve $h:\mathbb{R}^n \to \mathbb{R}^m$, $m \le n$ için genellenecektir.

Teorem 2.2.3.2 (Lagrange Teoremi**):** x^*; $h(x) = 0$, $h:\mathbb{R}^n \to \mathbb{R}^m, m \le n$ kısıtına göre $f:\mathbb{R}^n \to \mathbb{R}$ fonksiyonunun bir lokal minimumu (veya maksimumu) ve bu kısıtların regüler bir noktası olsun. Bu durumda

$$Df(x^*) + \lambda^{*T} Dh(x^*) = 0^T$$

sağlanacak şekilde bir $\lambda^* \in \mathbb{R}^m$ vardır [1].

İspat: Bazı $\lambda^* \in \mathbb{R}^m$ için

$$\nabla f(x^*) = -Dh(x^*)^T \lambda^*$$

olduğunu göstermek gerekir, yani

$$\nabla f(x^*) \in \Re\left(Dh(x^*)^T\right) = N(x^*)$$

dir. Fakat Lemma 2.2.2.6'dan $N(x^*) = T(x^*)^{\perp}$. Bu yüzden $\nabla f(x^*) \in T(x^*)^{\perp}$ olduğunu göstermek yeterlidir. $y \in T(x^*)$ olsun.

Teorem 2.2.2.4 den $\{x(t) : t \in (a,b)\}$ eğrisi vardır öyle ki $\forall t \in (a,b)$ için

$$h(x(t)) = 0$$

ve $t^* \in (a,b)$, $x(t^*) = x^*$, $\dot{x}(t^*) = y$ sağlanır.

$\phi(t) = f(x(t))$ bileşke fonksiyonu düşünülürse t^*,bu fonksiyonun lokal minimumudur. Kısıtsız lokal minimum için birinci basamaktan gerek şarttan (Teorem 2.1.3)

$$\frac{d\phi}{dt}(t^*) = 0$$

dır. Zincir kuralının uygulanmasıyla;

$$\frac{d\phi}{dt}(t^*) = Df(x^*)\dot{x}(t^*) = Df(x^*)y = \nabla f(x^*)^T y = 0$$

elde edilir. Böylece $\forall y \in T(x^*)$ için

$$\nabla f(x^*)^T y = 0$$

sağlanır. Yani $\nabla f(x^*) \in T(x^*)^{\perp}$ dir. Bu da ispatı tamamlar.

Lagrange teoremi belirtir ki eğer x^* bir ekstremum nokta ise f amaç fonksiyonunun gradiyenti kısıtların gradiyetlerinin lineer kombinasyonu olarak ifade edilebilir. Teorem 2.2.3.2'deki λ^* vektörü Lagrange çarpan vektörü ve onun bileşenleri Lagrange çarpanları olarak adlandırılır [1].

Lagrange teoreminin ispatından görülür ki kompakt bir şekilde yazmak için gerek koşul $\nabla f(x^*) \in N(x^*)$ dır. Eğer bu koşul sağlanmazsa, x^* ekstremum nokta olmaz. Lagrange fonksiyonu

$$l : \mathbb{R}^n x \mathbb{R}^m \to \mathbb{R}$$
$$l(x,\lambda) \triangleq f(x) + \lambda^T h(x)$$

ile verilebilir [1].

x^* lokal minimumu için Lagrange koşulu, Lagrange fonksiyonunun

$$Dl(x^*,\lambda^*) = 0^T$$

olarak kullanılmasıyla bazı λ^*'lar için (ki burada D türev operatörü, $\left[x^T,\lambda^T\right]^T$ argumanına göre türevdir) yeniden sunulabilir. Diğer bir değişle Lagrange teoreminin gerek koşulu, kısıtsız optimizasyon için birinci basamaktan gerek koşulunun Lagrange fonksiyonuna uygulanmasına eşittir [1].

Yukarıdaki ifadeyi görmek için, l 'nin x 'e göre türevi $D_x l$ ve λ ya göre türevi $D_\lambda l$ olarak tanımlanırsa

$$Dl(x,\lambda) = \left[D_x l(x,\lambda), D_\lambda l(x,\lambda)\right]$$
$$D_x l(x,\lambda) = Df(x) + \lambda^T Dh(x) \text{ ve } D_\lambda l(x,\lambda) = h(x)^T .$$

dir. Böylece x^* lokal minimumu için Lagrange teoremi

$$D_x l(x^*,\lambda^*) = 0^T$$
$$D_\lambda l(x^*,\lambda^*) = 0^T$$

$\lambda^* \in \mathbb{R}^m$ için belirlenir ki bu da $Dl(x^*,\lambda^*) = 0$ eşitliğini verir. Diğer bir ifadeyle Lagrange koşulu $Dl(x^*,\lambda^*) = 0^T$ olarak ifade edilebilir [1].

Lagrange koşulu mümkün olan ekstremum noktalarını bulmak için kullanılır. Bu ise

$$D_x l(x,\lambda) = 0^T$$
$$D_\lambda l(x,\lambda) = 0^T$$

denklemlerinin çözümünü gerektirir [1].

İkinci Basamaktan Koşullar

$f : \mathbb{R}^n \to \mathbb{R}$ ve $h : \mathbb{R}^n \to \mathbb{R}^m$ iki kez diferansiyellenebilir fonksiyonlar, yani $f, h \in C^2$ olmak üzere

$$l(x,\lambda) = f(x) + \lambda^T h(x) = f(x) + \lambda_1 h_1(x) + \ldots + \lambda_m h_m(x)$$

Lagrange fonksiyonunu ele alalım. $L(x,\lambda)$, x 'e göre $l(x,\lambda)$ 'nın Hesse matrisi

$$L(x,\lambda) = F(x) + \lambda_1 H_1(x) + \ldots + \lambda_m H_m(x)$$

olsun. Burada $F(x)$, f 'nin x 'e göre Hesse matrisi ve $H_k(x)$

$$H_k(x) = \begin{pmatrix} \dfrac{\partial^2 h_k(x)}{\partial x_1^2} & \cdots & \dfrac{\partial^2 h_k(x)}{\partial x_n \partial x_1} \\ \vdots & & \vdots \\ \dfrac{\partial^2 h_k(x)}{\partial x_1 \partial x_n} & \cdots & \dfrac{\partial^2 h_k(x)}{\partial x_n^2} \end{pmatrix}$$

$k = 1,...,m$ için x'e göre h_k'nın Hesse matrisidir. Bu durumda

$$\left[\lambda H(x)\right] = \lambda_1 H_1(x) + ... + \lambda_m H_m(x)$$

notasyonu kullanılarak

$$L(x,\lambda) = F(x) + \left[\lambda H(x)\right]$$

yazılabilir [1].

Teorem 2.2.3.3 (İkinci Basamaktan Gerek Koşullar**) :** x^*; $h(x) = 0,\ h:\mathbb{R}^n \to \mathbb{R}^m,\ m \le n$ kısıtına göre $f:\mathbb{R}^n \to \mathbb{R}$ fonksiyonunun lokal minimumu ve $f, g \in C^2$ olsun. x^*'ın regüler olduğu farzedilirse öyle $\lambda^* \in \mathbb{R}^m$ vardır ki

1. $Df(x^*) + \lambda^{*T} Dh(x^*) = 0^T$

2. $\forall y \in T(x^*)$ için $y^T L(x^*, \lambda^*) y \ge 0$

sağlanır [1].

İspat: λ^*'ın varlığı $Df(x^*) + \lambda^{*T} Dh(x^*) = 0^T$ Lagrange teoreminden görülür. Geriye sonucun ikinci kısmını ispatlamak kalır. $y \in T(x^*)$ yani y , $S = \{x \in \mathbb{R}^n : h(x) = 0\}$ yüzeyinde x^* da tanjant uzayına ait olsun. $h \in C^2$ olduğu için, Teorem 2.2.3.3 koşullarını takiben, S üzerinde iki kez diferansiyellenebilir $\{x(t) : t \in (a,b)\}$ eğrisi vardır öyle ki bazı $t^* \in (a,b)$ için

$$x(t^*) = x^*, \dot{x}(t^*) = y$$

dir.

Varsayımdan; t^*; $\phi(t) = f\left(x(t)\right)$ fonksiyonunun lokal minimumudur. Kısıtsız minimizasyon için ikinci basamaktan gerek koşuldan; (Teorem 2.1.2)

$$\frac{d^2\phi(t^*)}{dt^2} \geq 0$$

elde edilir.

$$\frac{d}{dt}\left(y(t)^T z(t)\right) = z(t)^T \frac{dy(t)}{dt} + y(t)^T \frac{dz(t)}{dt}$$

formülünün kullanımı ve zincir kuralından

$$\begin{aligned}
\frac{d^2\phi}{dt^2}(t^*) &= \frac{d}{dt}\left[Df\left(x(t^*)\right)\dot{x}(t^*)\right] \\
&= \dot{x}(t^*)^T F\left(x^*\right)\dot{x}(t^*) + Df\left(x^*\right)\ddot{x}(t^*) \\
&= y^T F\left(x^*\right) y + Df\left(x^*\right)\ddot{x}(t^*) \geq 0
\end{aligned}$$

elde edilir. $h\left(x(t)\right) = 0$, $\forall t \in (a,b)$ olduğu için

$$\frac{d^2}{dt^2}\lambda^{*T} h\left(x(t)\right) = 0$$

yazılır ve böylece $\forall t \in (a,b)$ için

$$\begin{aligned}
\frac{d^2}{dt^2}\lambda^{*T} h\left(x(t)\right) &= \frac{d}{dt}\left[\lambda^{*T} \frac{d}{dt} h\left(x(t)\right)\right] \\
&= \frac{d}{dt}\left[\sum_{k=1}^{m} \lambda_k^* \frac{d}{dt} h_k\left(x(t)\right)\right] \\
&= \frac{d}{dt}\left[\sum_{k=1}^{m} \lambda_k^* Dh_k\left(x(t)\right)\dot{x}(t)\right] \\
&= \sum_{k=1}^{m} \lambda_k^* \frac{d}{dt}\left(Dh_k\left(x(t)\right)\dot{x}(t)\right) \\
&= \sum_{k=1}^{m} \lambda_k^* \left[\dot{x}(t)^T H_k\left(x(t)\right)\dot{x}(t) + Dh_k\left(x(t)\right)\ddot{x}(t)\right] \\
&= \dot{x}^T(t)\left[\lambda^* H\left(x(t)\right)\right]\dot{x}(t) + \lambda^{*T} Dh\left(x(t)\right)\ddot{x}(t) \\
&= 0
\end{aligned}$$

dir. Yukarıdaki $t=t^*$ için doğrudur yani

$$y^T\left[\lambda^* H(x^*)\right]y+\lambda^{*T} Dh(x^*)\ddot{x}(t^*)=0$$

denkleminin eklenmesiyle eşitsizlik

$$y^T F(x^*)y+Df(x^*)\ddot{x}(t^*)\geq 0$$

$$y^T\left(F(x^*)+[\lambda^* H(x^*)]\right)y+\left(Df(x^*)+\lambda^{*T} Dh(x^*)\right)\ddot{x}(t^*)\geq 0$$

şeklini alır. Fakat Lagrange teoremi $Df(x^*)+\lambda^{*T} Dh(x^*)=0^T$ dir. Böylece

$$y^T\left(F(x^*)+[\lambda^* H(x^*)]\right)y=y^T L(x^*,\lambda^*)y\geq 0$$

bulunur ki bu da sonucu ispatlar.

$L(x,\lambda)$, f amaç fonksiyonunu $F(x)$ Hesse matrisinin kısıtsız minimizasyon durumunda yaptığı gibi benzer bir rol oynar. Ancak, şimdi $L(x^*,\lambda^*)\geq 0$ $\mathbb{R}^n$ 'den ziyade sadece $T(x^*)$ 'da gerektirir [1].

Bir noktanın lokal minimum olması için yukarıdaki koşullar gerek koşullardır fakat yeter değildir. Şimdi bir noktanın tam lokal minimum olması için yeter koşullar ispatsız verilecektir.

Teorem 2.2.3.4 (İkinci Basamaktan Yeter Koşullar**):** $f,g\in C^2$ ve $x^*\in\mathbb{R}^n$ ve $\lambda^*\in\mathbb{R}^m$ var olduğunu farz edelim öyle ki

1. $Df(x^*)+\lambda^{*T} Dh(x^*)=0^T$

2. $\forall y\in T(x^*), y\neq 0$ için $y^T L(x^*,\lambda^*)y>0$

dir. Bu durumda x^* , $h(x)=0$ kısıtına göre f 'nin tam lokal minimumudur [2].

Teorem 2.2.3.4 eğer x^* Lagrange koşulunu sağlar ve $L(x^*,\lambda^*)$, $T(x^*)$ üzerinde pozitif tanımlı ise x^*'ın tam lokal minimum olduğunu belirtir. Teorem

2.2.3.4 e benzer sonuç tam lokal maksimumum için, sadece $L(x^*, \lambda^*)$, $T(x^*)$ üzerinde negatif tanımlı olması farkıyla sağlanır [1].

2.3 Eşitsizlik Kısıtlı Optimizasyon

2.3.1 Karush-Kuhn-Tucker Koşulları

Kısım 2.2'de sadece eşitlik kısıtlarını içeren kısıtlı optimizasyon problemleri analiz edildi. Bu bölümde eşitsizlik kısıtı içeren ektremum problemleri incelenecek. Bu kısımdaki işleyiş Kısım 2.2'deki ile paraleldir. Eşitsizlik kısıtlı problemler Lagrange çarpanları kullanılarak işlem yapılır. Bu kısımda (2.3) problemi ele alınacaktır.

Tanım 2.3.1.1: Eğer $g_j(x^*) = 0$ ise $g_j(x^*) \leq 0$ eşitsizlik kısıtı x^*'da aktif kısıt olarak söylenir. Eğer $g_j(x^*) < 0$ ise x^*'da aktif olmayan kısıttır.

Bu durumda $h_i(x) = 0$ eşitlik kısıtı her zaman aktiftir.

Tanım 2.3.1.2: x^* noktası $h(x^*) = 0$, $g(x^*) \leq 0$ kısıtlarını sağlasın ve aktif eşitsizlik kısıtlarının indeks kümesi, $J(x^*)$

$$J(x^*) \triangleq \{ j : g_j(x^*) = 0 \}$$

biçiminde tanımlansın. Eğer $\nabla h_i(x^*)$, $\nabla g_j(x^*)$, $1 \leq i \leq m, j \in J(x^*)$ lineer bağımsız ise x^*'a bir regüler noktadır denir.

Şimdi bir noktanın lokal minimum olması için birinci basamaktan gerek koşul gösterilecektir. Bu koşul Karush-Kuhn-Tucker (KKT) koşulu olarak söylenir. Literatürde bu koşul genellikle Kuhn-Tucker koşulu olarak bilinir.

Teorem 2.3.1.3 (Karush-Kuhn-Tucker (KKT) Teoremi)**:** $f, h, g \in C^1$ **,** x^***,** $h(x) = 0, g(x) \leq 0$ kısıtlarına göre f 'nin minimum probleminin lokal minimumu ve regüler nokta olsun. Öyle ise $x^* \in \mathbb{R}^m$ ve $\mu^* \in \mathbb{R}^p$ vardır ki

1. $\mu^* \geq 0$

2. $Df(x^*) + \lambda^{*T} Dh(x^*) + \mu^{*T} Dg(x^*) = 0^T$

3. $\mu^{*T} g(x) = 0$

sağlanır [1,2]. Teorem 2.3.1.3'de λ^*,Lagrange çarpan vektörü ve μ^*,Karush-Kuhn-Tucker (KKT) çarpan vektörü olarak ifade edilir. Onların bileşenleri de Lagrange çarpanları ve Karush-Kuhn-Tucker (KKT) çarpanları olarak ifade edilir.

Bu teorem ispatlanmadan önce, anlamı tartışılacak, $\mu^T \geq 0$ (şart 1) ve $g_j(x^*) \leq 0$ gözlemlenecektir. Böylece koşul

$$\mu^{*T} g(x^*) = \mu_1^* g_1(x^*) + ... + \mu_p^* g_p(x^*) = 0$$

olarak elde edilir. Eğer $g_j(x^*) < 0$ ise $\mu_j^* = 0$ olduğunu ima eder yani $\forall j \notin J(x^*)$ için $\mu_j^* = 0$ dır. Diğer bir ifadeyle KKT çarpanları μ_j^* aktif olmayan kısıtlarda sıfırdır. Diğer KKT çarpanları , μ_j^*, $j \in J(x^*)$ negatif olmayandır, sıfıra eşit ve eşit olmayan olabilirler [1].

KKT koşulunu sağlayan noktalar araştırılır ve bu noktalar minimum adayı olarak görülür. Özetle, KKT koşulu beş kısımdan oluşur.

1. $\mu_j^* \geq 0$

2. $Df(x^*) + \lambda^{*T} Dh(x^*) + \mu^{*T} Dg(x^*) = 0^T$

3. $\mu_j^{*T} g(x^*) = 0$

4. $h(x^*) = 0$

5. $g(x^*) \le 0$

Şimdi aşağıda KKT teoreminin ispatı verilecektir.

İspat (KKT Teoremi**):** $x^*, \{x : h(x) = 0, g(x) \le 0\}$ kümesi üzerinde f 'nin regüler lokal minimumu olsun. Öyle ise x^*, aynı zamanda $\{x : h(x) = 0, g_j(x) = 0, j \in J(x^*)\}$ kümesi üzerinde de f 'nin regüler lokal minimumudur. İkinci kısıt kümesi sadece eşitlik kısıtlarını içerir. Böylece Lagrange teoreminden , $\lambda^* \in \mathbb{R}^m$ ve $\mu^* \in \mathbb{R}^p$ vardır öyle ki

$$Df(x^*) + \lambda^{*T} Dh(x^*) + \mu^{*T} Dg(x^*) = 0^T$$

$\forall j \notin J(x^*)$ için $\mu_j^* = 0$ elde edilir. İspatı tamamlamak için $\forall j \in J(x^*)$ için $\mu_j^* \ge 0$ olduğunu göstermek kalır. Olmayana ergi yöntemi kullanılırsa $j \in J(x^*)$ vardır öyle ki $\mu_j^* < 0$ olsun. $\hat{S}$ ve $\hat{T}(x^*)$ kümeleri sırasıyla x^*'daki bütün diğer aktif kısıtlar tarafından tanımlanan yüzey ve tanjant uzayı ifade etsin. Yani

$$\hat{S} = \{x : h(x) = 0,\ g_i(x) = 0,\ i \in J(x^*),\ i \ne j\}$$

ve

$$\hat{T}(x^*) = \{y :\ Dh(x^*)y = 0,\ Dg_i(x^*)y = 0,\ i \in J(x^*),\ i \ne j\}$$

dir. x^*'ın regülerliğinden $y \in \hat{T}(x^*)$ vardır öyle ki

$$Dg_j(x^*)y \ne 0$$

yazılabilir. Buradan genelliği kaybetmeksizin, $\nabla g_j(x^*)y < 0$ koşulunu sağlayan bir y olduğu varsayılır ve Lagrange koşulu göz önüne alınır ise

$$Df(x^*) + \lambda^{*T} Dh(x^*) + \mu_j^* Dg_j(x^*) + \sum_{i \ne j} \mu_i^* Dg_i(x^*) = 0^T$$

yazılabilir. Eğer yukarıda ki y ve $y \in \hat{T}(x^*)$ olduğu kullanılırsa

$$Df\left(x^*\right)y = -\mu_j{}^* Dg_j\left(x^*\right)y$$

elde edilir.

$Dg_j\left(x^*\right)y < 0$ ve $\mu_j{}^* < 0$ olduğu farz edildiği için

$$Df\left(x^*\right)y < 0$$

elde edilir. $y \in \hat{T}\left(x^*\right)$ olduğu için Teorem 2.2.2.4'de S üzerinde türevlenebilir bir $\left\{x\left(t\right): t \in \left(a,b\right)\right\}$ eğrisi bulunur. Öyle ki $t^* \in \left(a,b\right)$ vardır ve $x\left(t^*\right) = x^*$ ve $\dot{x}\left(t^*\right) = y$ dir.

Şimdi $\frac{d}{dt} f\left(x\left(t^*\right)\right) = \left(Df\left(x^*\right)\right)y < 0$ dır. Yani $\delta > 0$ vardır öyle ki $\forall t \in \left(t^*, t^* + \delta\right)$ için

$$f\left(x\left(t\right)\right) < f\left(x\left(t^*\right)\right) = f\left(x^*\right)$$

elde edilir. Diğer taraftan

$$\frac{d}{dt} g_j\left(x\left(t^*\right)\right) = Dg_j\left(x^*\right)y < 0$$

ve bazı $\varepsilon > 0$ ve $\forall t \in \left(t^*, t^* + \min\left\{\delta, \varepsilon\right\}\right)$ için $g_j\left(x\left(t\right)\right) \le 0$ ve $f\left(x\left(t\right)\right) < f\left(x^*\right)$ bulunur. $x\left(t\right)$, $t \in \left(t^*, t^* + \min\left\{\delta, \varepsilon\right\}\right)$ $\hat{S}$'de oldukları için amaç fonksiyonunun x^*'dan daha düşük değerleri ile uygun (feasible) noktalardır. Bu, x^*'ın lokal minimum olmasıyla çelişir ki bu da ispatı tamamlar.

Amaç fonksiyonu maksimize edildiğinde, yani optimizasyon problemi

Maksimum	$f\left(x\right)$
Kısıtlar	$g\left(x\right) \le 0$
	$h\left(x\right) = 0$

formunda olduğunda KKT koşulları

1. $\mu_j{}^* \ge 0$

2. $-Df\left(x^*\right)+\lambda^{*T}\ Dh\left(x^*\right)+\mu^{*T}\ Dg\left(x^*\right)=0^T$

3. $\mu_j{}^{*T}\ g\left(x^*\right)=0$

4. $h\left(x^*\right)=0$

5. $g\left(x^*\right)\leq 0$

olarak yazılabilir. Yukarıdaki amaç fonksiyonun -1 ile çarpılmasıyla yukarıdaki maksimum probleminin minimum probleme dönüşmesiyle kolayca türetilir ve yukarıdaki

1. $\mu_j{}^*\leq 0$

2. $Df\left(x^*\right)+\lambda^{*T}\ Dh\left(x^*\right)+\mu^{*T}\ Dg\left(x^*\right)=0^T$

3. $\mu_j{}^{*T}\ g\left(x^*\right)=0$

4. $h\left(x^*\right)=0$

5. $g\left(x^*\right)\leq 0$

olarak yeniden yazılabilir.

Yukarıdaki gösterilen form, koşul 2'nin -1 ile çarpılması ve μ^*, λ^*'ın işaretlerinin değişmesiyle elde edilir.

Benzer şekilde, eşitsizlik kısıtı $g\left(x\right)\geq 0$ formunda olduğunda KKT koşulu türetilebilir.

$$\begin{array}{ll} \text{Minimum} & f\left(x\right) \\ \text{Kısıtlar} & g\left(x\right)\geq 0 \\ & h\left(x\right)=0 \end{array}$$

problemi düşünülsün. Eşitsizlik kısıtı -1ile çarpılırsa $-g(x)\leq 0$ elde edilir. Böylece bu durum için KKT koşulu

1. $\mu_j{}^*\geq 0$

2. $Df\left(x^*\right)+\lambda^{*T}\,Dh\left(x^*\right)-\mu^{*T}\,Dg\left(x^*\right)=0^T$

3. $\mu_j{}^{*T}\,g\left(x^*\right)=0$

4. $h\left(x^*\right)=0$

5. $g\left(x^*\right)\geq 0$

dir. Önceki gibi μ^* ın işaretinin değişmesiyle

1. $\mu_j{}^*\leq 0$

2. $Df\left(x^*\right)+\lambda^{*T}\,Dh\left(x^*\right)+\mu^{*T}\,Dg\left(x^*\right)=0^T$

3. $\mu_j{}^{*T}\,g\left(x^*\right)=0$

4. $h\left(x^*\right)=0$

5. $g\left(x^*\right)\geq 0$

elde edilir.

$$\begin{array}{ll} \text{Maksimum} & f\left(x\right) \\ \text{Kısıtlar} & g\left(x\right)\geq 0 \\ & h\left(x\right)=0 \end{array}$$

problemi için KKT koşulu tam anlamıyla Teorem 2.3.1.3 ile aynıdır [1].

KKT koşulu ile ilişkili daha fazla sonuç için [21] bölüm 7'ye bakılabilir.

İkinci Basamaktan Koşullar

Eşitlik kısıtlı ekstremum problemlerinin durumunda olduğu gibi eşitsizlik kısıtları içeren ekstremum problemleri için ikinci basamaktan gerek ve yeter koşulları verilebilir. Bunun için aşağıdaki matrisi tanımlamak gerekir.

$$L\left(x,\lambda,\mu\right)=F\left(x\right)+\left[\lambda H\left(x\right)\right]+\left[\mu G\left(x\right)\right]$$

ki burada $F\left(x\right)$, x'de f'nin Hesse Matrisi ve $\left[\lambda H\left(x\right)\right]$notasyonu

$$\left[\lambda H(x)\right]=\lambda_1 H_1(x)+\ldots+\lambda_m H_m(x)$$

şeklinde sunulur. Benzer şekilde , $\left[\mu G(x)\right]$ notasyonu

$$\left[\mu G(x)\right]=\mu_1 G_1(x)+\ldots+\mu_p G_p(x)$$

dir ve burada $G_k(x)$;

$$G_k(x)=\begin{pmatrix} \frac{\partial^2 g_k(x)}{\partial x_1^2} & \cdots & \frac{\partial^2 g_k(x)}{\partial x_n \partial x_1} \\ \vdots & & \vdots \\ \frac{\partial^2 g_k(x)}{\partial x_1 \partial x_n} & \cdots & \frac{\partial^2 g_k(x)}{\partial x_n^2} \end{pmatrix}$$

ile verilen x'de g_k'nın Hesse matrisidir [1].

Aşağıdaki teoremde ,

$$T(x^*)=\left\{y\in\mathbb{R}^n : Dh(x^*)y=0, Dg_j(x^*)y=0, j\in J(x^*)\right\}$$

kullanılacak, yani yüzeyin tanjant uzayı aktif kısıtlar tarafından tanımlanacaktır [1].

Teorem 2.3.2.1 (**İ**kinci Basamaktan Gerek Koşullar**):** x^*; $h(x)=0,\ g(x)\le 0,\ h:\mathbb{R}^n\to\mathbb{R}^m,\ m\le n,\ g:\mathbb{R}^n\to\mathbb{R}^p$ kısıtlarına göre $f:\mathbb{R}^n\to\mathbb{R}$ fonksiyonunun lokal minimumu ve $f,h,g\in C^2$ olsun. x^*'ın regüler olduğu farz edilirse $\lambda^*\in\mathbb{R}^m$ ve $\mu^*\in\mathbb{R}^p$ vardır öyle ki

1. $\mu^*\ge 0,\ Df(x^*)+\lambda^{*T} Dh(x^*)+\mu^{*T} Dg(x^*)=0^T,\ \mu^{*T} g(x^*)=0$

2. $\forall y\in T(x^*)$ için $y^T L(x^*,\lambda^*,\mu^*)y\ge 0$

dir [1,2].

İspat: 1. kısım sadece KKT teoreminin bir sonucudur. 2. kısım ispat edilirse

$x^*, \{x : h(x)=0, g(x)\le 0\}$ üzerinde lokal minimum olduğu için aynı zamanda $\{x : h(x)=0, g_j(x)=0, j\in J(x^*)\}$ üzerinde de lokal minimumdur, yani x^* eşitlik kısıtları olarak alınan aktif kısıtlarla lokal minimumdur. Böylece eşitlik kısıtları için ikinci basamaktan gerek koşullar (Teorem 2.2.4.1) burada uygulanabilir ki bu da ispatı tamamlar.

Şimdi eşitsizlik kısıtlarını içeren ekstremum problemleri için ikinci basamaktan yeter koşullar belirlenecektir. Bu sonucun formulasyonunda aşağıdaki küme kullanılacaktır [1].

$$\tilde{T}(x^*,\mu^*)=\left\{y: Dh(x^*)y=0, Dg_i(x^*)y=0, i\in\tilde{J}(x^*,\mu^*)\right\}$$

ki burada $\tilde{J}(x^*,\mu^*)=\left\{i: g_i(x^*)=0, \mu_i^*>0\right\}$ dir. $\tilde{J}(x^*,\mu^*)$, $J(x^*)$'ın alt kümesidir: yani $\tilde{J}(x^*,\mu^*)\subset J(x^*)$ dir. Bu da $T(x^*)$'ın, $\tilde{J}(x^*,\mu^*)$ 'ın alt kümesi olduğunu ima eder: yani $T(x^*)\subset\tilde{T}(x^*,\mu^*)$ dir.

Teorem 2.3.2.2 (İkinci Basamaktan Yeter Koşullar) **:** $f, g, h\in C^2$ ve $x^*\in\mathbb{R}^n$ uygun (feasible) nokta olsun. $\lambda^*\in\mathbb{R}^m$ vektörleri ve $\mu^*\in\mathbb{R}^p$ vardır öyle ki

1. $\mu^*\geq 0,\ Df(x^*)+\lambda^{*T}Dh(x^*)+\mu^{*T}Dg(x^*)=0^T,\ \mu^{*T}g(x^*)=0$

2. $\forall y\in\tilde{T}(x^*,\mu^*),\ y\neq 0$ için $y^T L(x^*,\lambda^*,\mu^*)y>0$

dir.

Bu durumda x^*, $h(x)=0, g(x)\leq 0$ kısıtlarına göre f 'nin tam lokal minimumudur [2].

Teorem 2.3.2.2 nin benzer bir sonucu tam lokal maksimum için de sağlanır, sadece fark $\mu^*\leq 0$ ve $L(x^*,\lambda^*)$'ın $\tilde{T}(x^*,\mu^*)$ üzerinde negatif tanımlı olmasıdır.

Şimdi bir sonraki bölümde yukarıdaki bilgiler eşliğinde (2.2) optimizasyon problemini çözmek için ceza metodu adı verilen genel bir metot verilecektir.

2.4 Ceza Fonksiyonu Metodu

(2.2) genel kısıtlı bir optimizasyon problemi ele alınsın. Özel olarak (2.2) kısıtlı optimizasyon problemine , $\gamma \in \mathbb{R}$ pozitif sabiti ve $P:\mathbb{R}^n \to \mathbb{R}$ verilen bir fonksiyon olmak üzere

$$\text{Minimum } f(x)+\gamma P(x)$$

kısıtsız optimizasyon problemi ile yaklaşılacaktır. Sonra ilişkili kısıtsız optimizasyon problemi çözülecek ve orijinal problemin minimumuna bir yaklaşım olarak çözüm kullanılacaktır. γ sabiti ceza (penalty) parametresi, P fonksiyonu ise ceza fonksiyonu olarak adlandırılır [1]. Bir ceza fonksiyonu aşağıdaki gibi tanımlanabilir.

Tanım 2.4.1: Bir $P:\mathbb{R}^n \to \mathbb{R}$ fonksiyonu eğer aşağıdaki üç şartı sağlıyor ise yukarıdaki kısıtlı optimizasyon problemi için bir ceza fonksiyonu adını alır [1].

1. P sürekli

2. $P(x) \geq 0, \ \forall x \in \mathbb{R}^n$

3 $P(x)=0$ ancak ve ancak x uygun (feasible) nokta ($x \in \Omega$)

Açıkça, (2.2) kısıtsız probleminde, orijinal probleme iyi bir yaklaşım için P ceza fonksiyonu uygun bir şekilde seçilmeli. Ceza fonksiyonunun rolü, feasible (uygun) küme dışındaki noktaları "cezalandırmak" tır.

Ceza fonksiyonlarının nasıl seçileceğini göstermek için, $f:\mathbb{R}^n \to \mathbb{R}$, $g_i:\mathbb{R}^n \to \mathbb{R}$, $i=1,\dots,p$ olmak üzere

$$\begin{array}{ll} \text{Minimum} & f(x) \\ \text{Kısıtlar} & g_i(x) \leq 0 \quad i=1,\dots,p \end{array}$$

kısıtlı optimizasyon problemi ele alınsın. Sadece eşitsizlik kısıtlarının düşünülmesi kısıtlayıcı değildir çünkü $h(x)=0$ formundaki bir eşitlik kısıtı $\|h(x)\|^2 \leq 0$ eşitsizlik kısıtına eşittir. Yukarıdaki kısıtlı problem için ceza fonksiyonunun $g_1,\dots,g_p$ kısıt fonksiyonları cinsinden tanımlanması doğaldır. P için muhtemel bir seçim

$$P(x)=\sum_{i=1}^{p} g_i^+(x)$$

şeklindedir ve burada

$$g_i^+(x)=\max\{0,g_i(x)\}=\begin{cases} 0 & , \quad g_i(x)\le 0 \\ g_i(x) & , \quad g_i(x)>0 \end{cases}$$

dir [1]. Daha genel bir formülasyon ile ifade etmek istersek $x*$ problemin (global minimumunu) bir çözümü, P problem için verilen ceza fonksiyonu ve $\gamma_k \in \mathbb{R}$ $k=1,2,...$ pozitif sabitler olmak üzere

$$q(\gamma_k,x):\mathbb{R}^n \to \mathbb{R},\ q(\gamma_k,x)=f(x)+\gamma_k P(x)$$

fonksiyonu tanımlansın.

Her k için aşağıdaki kısıtsız optimizasyon problemi yazılabilir

$$\text{Minimum } q(\gamma_k,x).$$

x^k, $q(\gamma_k,x)$'in minimumunu ifade etsin. Aşağıdaki lemma, kısıtlı problem ve kısıtsız problem arasındaki bazı yararlı ilişkileri tanımlar.

Lemma 2.4.2: Farzedelim ki $\{\gamma_k\}$ azalmayan bir dizi, yani $\forall k$ için $\gamma_k \le \gamma_{k+1}$ olsun. Sonra $\forall k$ için

1. $q\left(\gamma_{k+1},x^{(k+1)}\right)\ge q\left(\gamma_k,x^{(k)}\right)$

2. $P\left(x^{(k+1)}\right)\le P\left(x^{(k)}\right)$

3. $f\left(x^{(k+1)}\right)\ge f\left(x^{(k)}\right)$

4. $f\left(x^*\right)\ge q\left(\gamma_k,x^{(k)}\right)\ge f\left(x^{(k)}\right)$

elde edilir [1].

İspat: İlk olarak 1. kısım ispatlansın. Q'nun tanımından ve $\{\gamma_k\}$ 'nın artan bir dizi olmasından

$$q\left(\gamma_{k+1}, x^{(k+1)}\right) = f\left(x^{(k+1)}\right) + \gamma_{k+1} P\left(x^{(k+1)}\right) \geq f\left(x^{(k+1)}\right) + \gamma_k P\left(x^{(k+1)}\right)$$

yazılabilir. Şimdi $x^{(k)}$, $q(\gamma_k, x)$'in minimumu olduğu için

$$q\left(\gamma_k, x^{(k)}\right) = f\left(x^{(k)}\right) + \gamma_k P\left(x^{(k)}\right) \leq f\left(x^{(k+1)}\right) + \gamma_k P\left(x^{(k+1)}\right)$$

dir. Yukarıdakilerin birleştirilmesiyle 1. kısm elde edilir.

Sonra 2. kısım ispatlansın. $x^{(k)}$ ve $x^{(k+1)}$, $q(\gamma_k, x)$ ve $q(\gamma_{k+1}, x)$'in minimumları olduğu için

$$q\left(\gamma_k, x^{(k)}\right) = f\left(x^{(k)}\right) + \gamma_k P\left(x^{(k)}\right) \leq f\left(x^{(k+1)}\right) + \gamma_k P\left(x^{(k+1)}\right)$$
$$q\left(\gamma_{k+1}, x^{(k+1)}\right) = f\left(x^{(k+1)}\right) + \gamma_{k+1} P\left(x^{(k+1)}\right) \leq f\left(x^{(k)}\right) + \gamma_{k+1} P\left(x^{(k)}\right)$$

yazılabilir. Eşitsizliklerin taraf tarafa toplanmasıyla

$$\gamma_k P\left(x^{(k)}\right) + \gamma_{k+1} P\left(x^{(k+1)}\right) \leq \gamma_{k+1} P\left(x^{(k)}\right) + \gamma_k P\left(x^{(k+1)}\right)$$

elde edilir. Yeniden düzenlenirse

$$(\gamma_{k+1} - \gamma_k) P\left(x^{(k+1)}\right) \leq (\gamma_{k+1} - \gamma_k) P\left(x^{(k)}\right)$$

bulunur. Verilen koşuldan $\gamma_{k+1} \geq \gamma_k$ olduğu biliniyor. Eğer $\gamma_{k+1} > \gamma_k$ ise $P\left(x^{(k+1)}\right) \leq P\left(x^{(k)}\right)$ elde edilir. Aksi taktirde eğer $\gamma_{k+1} = \gamma_k$ ise açıkça $x^{(k+1)} = x^{(k)}$ dır ve böylece $P\left(x^{(k+1)}\right) = P\left(x^{(k)}\right)$ dir. Böylece 2. kısma ulaşılır.

3. kısım ispatlansın. x^k , $q(\gamma_k, x)$'in minimumu olduğu için

$$q\left(\gamma_k, x^{(k)}\right) = f\left(x^{(k)}\right) + \gamma_k P\left(x^{(k)}\right) \leq f\left(x^{(k+1)}\right) + \gamma_k P\left(x^{(k+1)}\right)$$

elde edilir ve böylece

$$f\left(x^{(k+1)}\right) \geq f\left(x^{(k)}\right) + \gamma_k\left(P\left(x^{(k)}\right) - P\left(x^{(k+1)}\right)\right)$$

ve 2. kısımdan $P\left(x^{(k)}\right) - P\left(x^{(k+1)}\right) \geq 0$ ve $\gamma_k \geq 0$ koşulundan

$$f\left(x^{(k+1)}\right) \geq f\left(x^{(k)}\right)$$

yazılır.

Son olarak 4. kısım ispatlansın. x^k , $q(\gamma_k, x)$'nın minimumu olduğu için

$$f\left(x^*\right) + \gamma_k P\left(x^*\right) \geq q\left(\gamma_k, x^{(k)}\right) = f\left(x^{(k)}\right) + \gamma_k P\left(x^{(k)}\right)$$

yazılır.

x^*, kısıtlı optimizasyon probleminin minimumu olduğu için $P\left(x^*\right) = 0$ dır. Böylece

$$f\left(x^*\right) \geq f\left(x^{(k)}\right) + \gamma_k P\left(x^{(k)}\right)$$

dir. $P\left(x^*\right) \geq 0$ ve $\gamma_k \geq 0$ olduğu için

$$f\left(x^*\right) \geq q\left(\gamma_k, x^{(k)}\right) \geq f\left(x^{(k)}\right)$$

dir ki bu da ispatı tamamlar.

Teorem 2.4.3: f amaç fonksiyonu sürekli ve $k \to \infty$ için $\gamma_k \to \infty$ olsun. $\{x_k\}$ dizisinin herhangi yakınsak alt dizisinin limiti kısıtlı optimizasyon probleminin bir çözümüdür [1].

İspat: $\{x^{m_k}\}$, $\{x_k\}$ dizisinin yakınsak alt dizisi ve $\hat{x}$, $\{x^{m_k}\}$'nın limiti olsun. Lemma 2.4.2'den $\left\{q\left(\gamma_k, x^{(k)}\right)\right\}$ dizisi, azalmayan ve $f\left(x^*\right)$ tarafından üstten sınırlıdır. Böylece $\left\{q\left(\gamma_k, x^{(k)}\right)\right\}$ dizisinin $q^* = \lim_{k\to\infty} q\left(\gamma_k, x^{(k)}\right)$ limiti vardır öyle ki $q^* \leq f\left(x^*\right)$, f sürekli ve Lemma 2.4.2'den $f\left(x^{(m_k)}\right) \leq f\left(x^*\right)$ olduğu için

$$\lim_{k\to\infty} f\left(x^{(m_k)}\right) = f\left(\lim_{k\to\infty} x^{(m_k)}\right) = f\left(x\right) \leq f\left(x^*\right) \ddot{x}$$

elde edilir.

$\left\{ f\left(x^{(m_k)}\right)\right\}$ ve $\left\{ q\left(\gamma_{m_k}, x^{(m_k)}\right)\right\}$ dizilerinin her ikisi de yakınsak ve

$$\left\{\gamma_{m_k} P\left(x^{(m_k)}\right)\right\} = \left\{ q\left(\gamma_{m_k}, x^{(m_k)}\right) - f\left(x^{(m_k)}\right)\right\}$$

dizisi de yakınsak olduğu için $\lim_{k\to\infty} \gamma_{m_k} P\left(x^{m_k}\right) = q^* - f\left(\hat{x}\right)$ elde edilir.

Lemma 2.4.2 den $\left\{ P\left(x^{m_k}\right)\right\}$ dizisi artmayan ve 0 tarafından alttan sınırlıdır. Böylece $\left\{ P\left(x^{k}\right)\right\}$ ve $\left\{ P\left(x^{m_k}\right)\right\}$ yakınsaktır. $\gamma_{m_k} \to \infty$ olduğu için

$$\lim_{k\to\infty} P\left(x^{m}\right) = 0$$

bulunur. P' nin sürekliliğinden

$$0 = \lim_{k\to\infty} P\left(x^{(m_k)}\right) = P\left(\lim_{k\to\infty} x^{(m_k)}\right) = P\left(\hat{x}\right)$$

ve $\hat{x}$ uygun (feasible) bir noktadır. Yukarıdan $f\left(x^*\right) \geq f\left(\hat{x}\right)$ olduğu için $\hat{x}$ 'ın kısıtlı optimizasyon probleminin bir çözümü olduğu sonucuna varılabilir.

3. KARARLILIK ANALİZİ

3.1 Lyapunov Kararlılık ve Dinamik Sistem

$f : D \to \mathbb{R}^n$ olmak üzere

$$\dot{x} = f(x) \tag{3.1}$$

otonom sistemi, $\mathbb{R}^n$ içinde $D \subset \mathbb{R}^n$ bölgesinde lokal Lipschitz dönüşümüdür [22]. $\bar{x} \in D$, (3.1)'in denge noktası olsun yani;

$$f(\bar{x}) = 0$$

dır. Amaç $\bar{x}$'ın kararlılığı ile çalışmak ve onu karakterize etmektir. Kolaylık için bütün tanım ve teoremlerde $\mathbb{R}^n$'in orijin noktasının denge noktası olduğu durum ele alınacaktır. Bunu yapmada hiçbir genelleme kaybı olmaz çünkü herhangi bir denge noktası değişken değiştirmesi ile orijin noktasına dönüştürülebilir. $\bar{x} \neq 0$ olsun ve $y = x - \bar{x}$ değişken değiştirmesini düşünülsün. y'nin türevi

$$\dot{y} = \dot{x} = f(x) = f(y + \bar{x}) = g(y), \;\; g(0) = 0$$

olarak verilir. Yeni değişken y için sistemin orijininde denge noktası vardır. Böylece genelliği kaybetmeksizin her zaman $f(x)$'in $f(0) = 0$'ı sağladığı düşünülebilir ve $x = 0$ orijinin kararlılığında çalışabilir [22].

Tanım 3.1.1: $x = 0$ (3.1)'in denge noktası

- eğer $\forall \varepsilon > 0$ için $\delta = \delta(\varepsilon) > 0$ vardır öyle ki

$$\|x(0)\| < \delta \Rightarrow \|x(t)\| < \varepsilon, \forall t \geq 0$$

ise kararlıdır.

• eğer kararlı değilse kararsızdır.

• eğer kararlı ve δ

$$\|x(0)\| < \delta \Rightarrow \lim_{t\to\infty} x(t) = 0$$

şeklinde seçilirse asimptotik kararlıdır [21].

$V : D \to \mathbb{R}$, orijini içeren $D \subset \mathbb{R}^n$ bölgesinde tanımlı, sürekli, türevlenebilir fonksiyon olsun. V 'nin , (3.1)'in çözüm eğrisi boyunca türevi $\dot{V}(x)$ olarak tanımlanır ve

$$\dot{V}(x) = \sum_{i=1}^{n} \frac{\partial V}{\partial x_i}\dot{x}_i = \sum_{i=1}^{n} \frac{\partial V}{\partial x_i} f_i(x)$$

$$= \left[\frac{\partial V}{\partial x_1}, \frac{\partial V}{\partial x_2}, \ldots, \frac{\partial V}{\partial x_n}\right] \begin{bmatrix} f_1(x) \\ f_2(x) \\ . \\ . \\ . \\ f_n(x) \end{bmatrix} = \frac{\partial V}{\partial x} f(x)$$

dir. V 'nin sistemin çözüm eğrisi boyunca türevi, sistemin denklemine bağlıdır. Böylece $\dot{V}(x)$ farklı sistemler için farklı olacaktır. Eğer $\phi(t;x)$, $t = 0$ zamanında x başlangıç durumunda (3.1)' in çözümü ise

$$\dot{V}(x) = \frac{d}{dt} V(\phi(t,x))\big|_{t=0}$$

dir. Böylece $\dot{V}(x)$ negatif ise V , (3.1)'in çözümü boyunca azalacaktır [4]. Şimdi Lyapunov Kararlılık teoremi verilebilir.

Teorem 3.1.2: $x = 0$, (3.1) için denge noktası ve $D \subset \mathbb{R}^n$, $x = 0$ 'ı içeren bir bölge olsun. $V : D \to \mathbb{R}$ sürekli türevlenebilir bir fonksiyon olmak üzere

$$V(0) = 0, \text{ ve } x \in D - \{0\} \text{ için } \mathrm{V}(x) \geq 0, \tag{3.2}$$

$$\text{her } x \in D \text{ için } \dot{V}(x) \leq 0 \tag{3.3}$$

ise verilen sistem $x=0$ noktasında kararlıdır denir. Ayrıca eğer

$$\forall x \in D-\{0\} \text{ için } \dot{V}(x)<0 \tag{3.4}$$

ise sistem $x=0$ noktasında asimptotik kararlıdır denir [22].

İspat: $\varepsilon>0$ verilsin ve

$$B_r=\left\{x\in\mathbb{R}^n \middle| \ \|x\|\le r\right\}\subset D$$

şartını sağlayan $r\in(0,\varepsilon]$ seçilsin. $\alpha=\min\limits_{\|x\|=r}V(x)$ olsun. Öyle ise (3.2) den $\alpha>0$ dır. $\beta\in(0,\alpha)$ alalım ve

$$\Omega_\beta=\left\{x\in B_r \middle| \ V(x)\le\beta\right\}$$

olsun. Bu durumda Ω_β, B_r nin iç bölgesidir. Ω_β kümesi $t=0$ da Ω_β da başlayan herhangi bir eğrinin $\forall t\ge 0$ için Ω_β'nın içinde kalması özelliğine sahiptir. (3.2) den bu görünür çünkü

$$\dot{V}(x)\le 0 \Rightarrow V(x(t))\le V(x(0))\le\beta,\ \forall t\ge 0$$

dir. Ω_β kompakt bir küme olduğu için (tanımından kapalı ve B_r'nin içinde olduğu için sınırlıdır) ve varlık teklik teoreminden ([22],teorem 2.4) (3.1) in $x(0)\in\Omega_\beta$ olduğu her durumda $\forall t\ge 0$ için tanımlanan tek bir çözümü olduğu sonucuna ulaşılabilir. $V(x)$ sürekli ve $V(0)=0$ olduğu için $\delta>0$ vardır öyle ki

$$\|x\|<\delta \Rightarrow V(x)<\beta$$

dir.

$$B_\delta\subset\Omega_\beta\subset B_r$$

ve

$$x(0)\in B_\delta \Rightarrow x(0)\in\Omega_\beta \Rightarrow x(t)\in\Omega_\beta \Rightarrow x(t)\in B_r$$

dir ve böylece

$$\|x(0)\|<\delta \Rightarrow \|x(t)\|<r\le\varepsilon, \forall t\ge 0$$

dir ki bu da $x=0$ denge noktasının kararlı olduğunu gösterir. Şimdi de (3.4) ün sağlandığı farzedilsin. Asimptotik kararlılığı göstermek için $t\to\infty$ için $x(t)\to 0$ olduğu gösterilmeli yani $\forall\alpha>0$ için $T>0$ vardır öyle ki $\|x\|<\alpha, \forall t>T$. Bilinir ki $\forall a>0$ için $b>0$ seçilebilir öyle ki $\Omega_b\subset B_a$. Böylece $t\to\infty$ için $V(x(t))\to 0$ göstermek yeterlidir. $V(x(t))$ monoton azalan ve sıfır tarafından alttan sınırlı olduğu için

$$t\to\infty \text{ için } V(x(t))\to c\ge 0$$

dir. $c=0$ olduğunu göstermek için olmayana ergi yöntemi kullanılsın. $c>0$ olsun. $V(x)$'in sürekliliğinden $d>0$ vardır öyle ki $B_d\subset\Omega_c$. $V(x(t))\to c>0$ limiti $\forall t\ge 0$ için B_d'nin dışına uzanan $x(t)$ eğrisini ifade eder. Sürekli $\dot V(x)$ fonksiyonu kompakt bir $\{d\le\|x\|\le r\}$ kümesinde maksimuma sahip olduğu için $-\gamma=\underset{d\le\|x\|\le r}{maks}\dot V(x)$ var olsun.

$$V(x(t))=V(x(0))+\int_0^t \dot V(x(\tau))d\tau \le V(x(0))-\gamma t$$

dir. Sağ taraf negatif olacağı için eşitsizlik $c>0$ kabulü ile çelişir.

(3.2) ve (3.3)'ü sağlayan sürekli türevlenebilir bir $V(x)$ fonksiyonu Lyapunov fonksiyonu olarak adlandırılır. Bazı $c>0$ için $V(x)=c$ yüzeyi Lyapunov yüzey veya seviye yüzeyi olarak adlandırılır [22].

(3.2) şartını yani $x \neq 0$ için $V(x) = 0$ ve $V(x) > 0$ şartını sağlayan bir $V(x)$ fonksiyonu pozitif tanımlıdır denir. Eğer $x \neq 0$ için $V(x) \geq 0$ daha zayıf koşulunu sağlıyor ise pozitif yarı tanımlıdır denir. Eğer $-V(x)$ pozitif tanımlı veya pozitif yarı tanımlı ise $V(x)$ fonksiyonu negatif tanımlı veya negatif yarı tanımlıdır denir. Bu ifadeyle, Lyapunov teoremi; eğer $\dot{V}(x)$ negatif yarı tanımlı olacak şekilde pozitif tanımlı sürekli türevlenebilir bir $V(x)$ fonksiyonu varsa orijin kararlıdır ve eğer $\dot{V}(x)$ negatif tanımlı ise asimptotik kararlıdır şeklinde yeniden söylenebilir [22].

4. ÜSTEL CEZA FONKSİYONU VE DİNAMİK SİSTEM YAKLAŞIMI

4.1 Üstel Ceza Fonksiyonu

$$\begin{array}{ll} \text{Minimum} & f(x) \\ \text{Kısıtlar} & g_1(x) \le 0, \ldots, g_r(x) \le 0 \end{array} \tag{4.1}$$

problemi ele alınsın. Aşağıdaki özellikler ile birlikte iki kez türevlenebilen $\psi : \mathbb{R} \to \mathbb{R}$ ceza fonksiyonu tarafından karakterize edilen bir metot tanımlanacaktır

(i). $\nabla^2 \psi(t) > 0, \forall t \in \mathbb{R}$

(ii). $\psi(0) = 0, \nabla \psi(0) = 1,$

(iii). $\lim\limits_{t \to -\infty} \psi(t) > -\infty,$

(iv) $\lim\limits_{t \to -\infty} \nabla \psi(t) = -\infty$ ve $\lim\limits_{t \to \infty} \nabla \psi(t) = \infty$.

Basit bir örneği üstel ceza fonksiyonudur ve

$$\psi(t) = e^t - 1 \tag{4.2}$$

şeklindedir [23,24].

$$x^k \in \arg\min_{x \in \mathbb{R}^n} \left\{ f(x) + \sum_{j=1}^{r} \frac{\mu_j^k}{c_j^k} \psi\left(c_j^k g_j(x)\right) \right\} \tag{4.3}$$

kısıtsız minimizasyonlar dizisinin oluşturduğu metot

$$\mu_j^{k+1} = \mu_j^k \nabla \psi\left(c_j^k g_j\left(x^k\right)\right),\ j = 1, \ldots, r \tag{4.4}$$

çarpan iterasyonu yoluyla meydana gelir [23].

Burada, $\{c_j^k\}$ her j için pozitif ceza parametresi dizisidir ve başlangıç çarpanları μ_j^0 keyfi pozitif sayılardır. $\mu_j^k > 0$ sabiti için "ceza" terimi

$$\frac{\mu_j^k}{c_j^k}\psi\left(c_j^k g_j(x)\right)$$

($c_j^k \to \infty$ için) tüm uygun olmayan x $[g_j(x) > 0]$ ler için ∞ a; tüm uygun $x\left[g_j(x) \leq 0\right]$ noktaları için 0'a yakınsar. Bunu görmek için, ψ'nin konveksliğinden ve

$$\frac{1}{c}\psi(ct) \geq \frac{1}{c}\psi(ct/2) + \frac{t}{2}\nabla\psi(ct/2)$$

ifadesinden

$$\psi(ct) \geq \psi(ct/2) + \frac{ct}{2}\nabla\psi(ct/2)$$

elde edilir. Böylece eğer t>0 ise $\psi(ct/2) > 0$ ve $\lim_{\tau\to\infty}\nabla\psi(\tau) = \infty$ şartları $\lim_{c\to\infty}(1/c)\psi(ct) = \infty$ eşitliğini ima eder. Ayrıca eğer t<0 ise $\lim_{c\to\infty}(1/c)\psi(ct) = 0$ olduğu için $\inf_{c>0}\psi(ct) = \inf_{\tau>0}\psi(\tau) > \infty$ elde edilir [23].

Diğer taraftan, sabit c_j^k için , $\mu_j^k \to 0$ oldukça (ki bu da eğer $j-$yinci kısıt optimumda aktif değilse meydana gelmesi beklenendir) ceza terimi tüm x ler için, uygun yada uygun olmayan, sıfıra gider.

Üstel ceza fonksiyonu $\psi(t) = e^t - 1$ için çarpan iterasyonu

$$\mu_j^{k+1} = \mu_j^k e^{c_j^k g_j(x^k)}$$

şeklindedir.

Çarpanların anlık değerine bağlı olan ceza parametresi bu çarpanlar küçüldükçe büyür ilişkili çarpanların sıfıra gittiği aktif olmayan kısıtlar için ilgili ceza parametreleri sonsuza gider. Her bir c_j^k, μ_j^k ile ters orantılı olarak belirlenir, yani c değişmez pozitif sabit olmak üzere

$$c_j^k = c / \mu_j^k \quad , \quad \forall j$$

şeklindedir [10].

4.2 Optimizasyon Problemine Dinamik Sistem Yaklaşımı

(4.1) problemi yeniden ele alınacaktır. Bu problem için bir önceki bölümde tanımlanan (4.2) ceza fonksiyonu yardımı ile elde edilen (4.3) kısıtsız optimizasyon problemi düşünülür ve aşağıdaki şekilde düzenlenirse

$$\text{Minimum} \quad f(x) + \sum_{j=1}^{r} \frac{\mu_j^k}{c_j^k} \psi\left(c_j^k g_j(x)\right) \tag{4.5}$$

bulunur. Kısıtsız optimizasyon probleminin çözümü için günümüzde bir çok metot uygulanmaktadır. Newton metodu, Line search (çizgi algoritması) metodu,…. şeklinde çoğaltılabilirler. Bu bölümde (4.5) kısıtsız optimizasyon problemini çözmek için dinamik sistem yapısı kullanılacaktır. Dinamik sistem yapısının bir çeşidi de sinir ağı (neural network) modelidir [16,18,19]. Sinir ağı modelinde problemi kısıtsız probleme çevirmek için bir çeşit aktivasyon fonksiyonu kullanılır. Aslında burada kullandığımız üstel ceza fonksiyonu da aktivasyon fonksiyonuna bir çeşit örnek olarak düşünülebilir. Sonrasında oluşan sinir ağı yapısı da matematiksel olarak dinamik sistem yapısıyla benzerdir. Şimdi tezde bahsedilen konuya geri dönülürse dinamik sistem yapısı elde edilir ve kısıtsız optimizasyon problemi

$$F(x,\mu,c) = f(x) + \sum_{j=1}^{r} \frac{\mu_j^k}{c_j^k} \psi\left(c_j^k g_j(x)\right)$$

şeklinde yazılır. Burada çarpan iterasyonu ise $\psi(t) = e^t - 1$ ceza fonksiyonu için

$$\mu_j^{k+1} = \mu_j^k \exp\left\{c_j^k g_j\left(x^k\right)\right\}$$

şeklindedir. Burada dinamik sistem yapısı gradiyentin tersi yönünde oluşturulur [25]. Yani dinamik sistem yapısı aşağıdaki şekildedir

$$\frac{dx}{dt} = -\nabla F\left(x, \mu, c\right). \tag{4.6}$$

(4.6) dinamik sistemi verilen farklı başlangıç koşullarına göre çözülebilir.

Dinamik sistemin herhangi bir çözüm yörüngesinin kararlılıktan dolayı bir çözüme yakınsamasını göstermek için bir sonraki bölümde sistemin kararlılığı ve (4.5) kısıtsız optimizasyon probleminin minimum noktası ile (4.6) dinamik sisteminin denge noktası arasındaki ilişki incelenecektir.

4.3 Kararlılık Analizi

Tezin bu bölümünde bir önceki bölümde elde edilen dinamik sistemin herhangi bir çözüm yörüngesinin bir çözüme yakınsadığı üçüncü bölümde verilen kararlılık teoremleri yardımıyla gösterilecektir.

Buraya kadar doğrusal olmayan kısıtlı bir optimizasyon problemi üstel ceza fonksiyonu yardımıyla kısıtsız bir probleme dönüşmüş ve kısıtsız problemden bir dinamik sistem yapısı oluşturulmuştur. Asıl amaç problemin çözümüne dinamik sistem yapısıyla yaklaşmaktır. Bu yüzden bu bölümde elde edilen dinamik sistemin herhangi bir çözüm yörüngesinin kararlılığı incelenerek bir çözüme yakınsadığı gösterilecek ve bu çözüm ile kısıtsız optimizasyon problemi arasındaki ilişki bir teorem yardımı ile elde edilecektir.

Aşağıdaki teorem, elde edilen dinamik sistem yapısının çözüm yörüngesinde bir çözüme yakınsadığını ve bu çözüm ile bulunan kısıtsız optimizasyon probleminin bir çözümü arasındaki ilişkiyi göstermektedir.

Teorem 4.3.1: (4.6) dinamik sisteminin denge noktası asimptotik kararlıdır ancak ve ancak denge noktası, (4.5) kısıtsız optimizasyon probleminin tam lokal minimum noktasıdır.

İspat: $(:\Leftarrow)$ x^*, (4.5) kısıtsız optimizasyon probleminin tam lokal minimum noktası olsun. Genelliği kaybetmeksizin $x^*=0$ alınabilir. Şimdi Lyapunov fonksiyonu tanımlansın. Sistem için Lyapunov fonksiyonu

$$V(x)=F(x)-F(0)$$

şeklindedir. x^*, F fonksiyonunun tam lokal minimumu olduğu için $F(0)=0$ ve x^*'ın Ω komşuluğunda açıkça $V(x)$ pozitif tanımlıdır. Bunun yanı sıra $V(x)$ birinci basamaktan kısmi türevlere sahiptir. Yani

$$\dot{V}(x)=\langle \nabla V(x),\dot{x}\rangle=\langle \nabla F(x),-\nabla F(x)\rangle=-\nabla^T F(x)\nabla F(x)<0$$

dir. Bu da Lyapunov kararlılık teoreminden x^* tam lokal mininmum noktasının asimptotik kararlı bir denge noktası olduğunu gösterir.

$(:\Rightarrow)$ Şimdi x^* denge noktası asimptotik kararlı olsun. Yani $x(t)$ dinamik sistemin bir çözümü olmak üzere $\forall t\geq t_0$ için $\delta>0$ vardır ve $\|x(t_0)-x^*\|<\delta$ iken $t\to\infty$ için $x(t)\to x^*$ dır. Burada [17] makalesinde adı geçen Teorem 1 in ispatına benzer mantık uygulanacaktır. $\|x(t_0)-x^*\|<\delta$ ve $x_0\neq x^*$, $x(0)=x_0$ olsun. Bu durumda eğri $x(0)$'dan $x^*=x(\infty)\neq x(0)$'a giderken

$$\begin{aligned}F(x^*)-F(x(0))&=\int_{x(0)}^{x^*}\frac{dF}{dt}dt=\int_0^\infty\sum_{k=1}^{n}\frac{\partial F}{\partial x_k}\frac{dx_k}{dt}dt\\&=-\int_0^\infty\left(\frac{dx_k}{dt}\right)^2dt<0\end{aligned}$$

elde edilir. Böylece x^*, $\|x(t_0)-x^*\|<\delta$ üzerinde F fonksiyonunun tam lokal minimum noktası olduğu gösterildi.

4.4 Nümerik örnekler

Örnek 1 [26, problem no: 324**]:**

$$\begin{aligned}
&\text{Minimum} \quad && f(x) = 0.01x_1^2 + x_2^2 \\
&\text{Kısıtlar} && g_1(x) = x_1x_2 - 25 \geq 0 \\
& && g_2(x) = x_1^2 + x_2^2 - 25 \geq 0 \\
& && g_3(x) = x_1 - 2 \geq 0
\end{aligned}$$

optimizasyon problemi ele alınsın. Verilen başlangıç değerleri $x_0 = (2,2)$ şeklinde olup ulaşılacak optimum değer $x^* = (15.81, 1.581)$ şeklindedir. Bu problemde eşitsizlik kısıtları önceki bölümde tanımlanan problem şeklinde verilmemiştir. Fakat problem istenen şekilde aşağıdaki gibi getirilebilir,

$$\begin{aligned}
&\text{Minimum} \quad && f(x) = 0.01x_1^2 + x_2^2 \\
&\text{Kısıtlar} && g_1(x) = 25 - x_1x_2 \leq 0 \\
& && g_2(x) = 25 - x_1^2 - x_2^2 \leq 0 \\
& && g_3(x) = 2 - x_1 \leq 0 \qquad .
\end{aligned}$$

Buradan problem için μ_j^k çarpanları

$$\mu_j^{k+1} = \mu_j^k \exp\left\{c_j^k g_j(x)\right\}$$

eşitliğinden kısıt sayısı üç olduğu için

$$\mu_1^{k+1} = \mu_1^k \exp\left\{c_1^k g_1\left(x^k\right)\right\}$$

$$\mu_2^{k+1} = \mu_2^k \exp\left\{c_2^k g_2\left(x^k\right)\right\}$$

$$\mu_3^{k+1} = \mu_3^k \exp\left\{c_3^k g_3\left(x^k\right)\right\}$$

şeklindedir. Problemden g_1, g_2, g_3 kısıtları yerlerine yazılırsa

$$\mu_1^{k+1} = \mu_1^k \exp\left\{c_1^k \left(25 - x_1x_2\right)\right\}$$

$$\mu_2^{k+1} = \mu_2^k \exp\left\{c_2^k \left(25 - x_1^2 - x_2^2\right)\right\}$$

$$\mu_3^{k+1} = \mu_3^k \exp\left\{c_3^k\left(2 - x_1\right)\right\}$$

olarak bulunur. Şimdi sistem oluşturulursa

$$\frac{dx}{dt} = -\nabla F\left(x, \mu, c\right)$$

elde edilir ve burada $F\left(x, \mu, c\right)$ fonksiyonu

$$F\left(x, \mu, c\right) = f\left(x\right) + \sum_{j=1}^{3} \frac{\mu_j}{c_j}\left(\exp\left\{c_j g_j\right\} - 1\right)$$

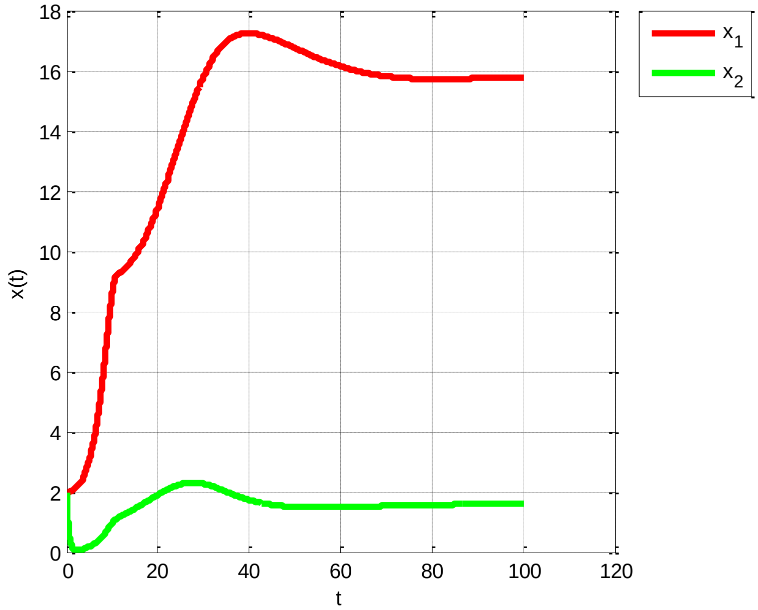
olarak bulunur. Problemdeki değerler yerlerine yazılırsa

$$F\left(x, \mu, c\right) = 0.01x_1^2 + x_2^2 + \frac{\mu_1}{c_1}\left(\exp\left\{c_1\left(25 - x_1 x_2\right)\right\} - 1\right)$$
$$+ \frac{\mu_2}{c_2}\left(\exp\left\{25 - x_1^2 - x_2^2\right\} - 1\right) + \frac{\mu_3}{c_3}\left(\exp\left\{2 - x_1\right\} - 1\right)$$

şeklindedir ve buradan da bu fonksiyonun gradiyenti alınarak dinamik sistem aşağıdaki gibi oluşturulur,

$$\frac{dx_1}{dt} = -0.02x_1 + \mu_1 x_1 \exp\left(c_1\left(25 - x_1 x_2\right)\right)$$
$$+2\mu_2 x_1 \exp\left(c_2\left(25 - x_1^2 - x_2^2\right)\right) + \mu_3 \exp\left(c_3\left(2 - x_1\right)\right)$$
$$\frac{dx_2}{dt} = -2x_2 + \mu_1 x_1 \exp\left(c_1\left(25 - x_1 x_2\right)\right) + 2\mu_2 x_2 \exp\left(c_2\left(25 - x_1^2 - x_2^2\right)\right).$$

Verilen başlangıç koşulları ile birlikte çeşitli nümerik metotlar kullanılarak problem çözülebilir. Şekil (4.1) deki grafik problemin Euler metodu ve Matlab (R2007b) programı yardımı ile çözülmesiyle elde edilmiştir.



Şekil 4.1: Örnek 1 de $x(t)$ 'nin grafiği $\left(x^{*}=(15.81,1.581)\right)$

Örnek 2 [26,problem no:228**]:**

$$\begin{array}{ll} \text{Minimum} & f(x)=x_1^2+x_2 \\ \text{Kısıtlar} & g_1(x)=-x_1-x_2+1\geq 0 \\ & g_2(x)=-\left(x_1^2+x_2^2\right)+9\geq 0 \end{array}$$

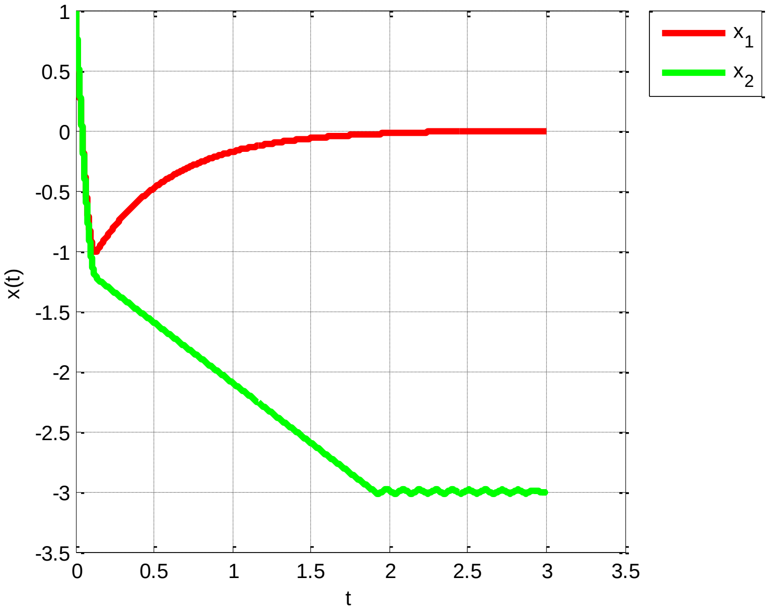
kısıtlı optimizasyon problemi ele alınsın. Verilen başlangıç değerleri $x_0=(0,0)$ şeklinde olup ulaşılacak optimum değer $x^{*}=(0,-3)$ dir. Örnek1 deki gibi "$\geq$" eşitsizliği "$\leq$" eşitsizliğine dönüştürülebilir. Bölüm 4.2'deki metot kullanılarak elde edilecek sistem, "μ" çarpanı ve "c" ceza parametresi aşağıdaki gibi bulunur,

$$\frac{dx_1}{dt}=-2x_1-\mu_1\exp\left(c_1\left(x_1+x_2-1\right)\right)-2x_1\mu_2\exp\left(c_2\left(x_1^2+x_2^2-9\right)\right)$$
$$\frac{dx_2}{dt}=-1-\mu_1\exp\left(c_1\left(x_1+x_2-1\right)\right)-2x_2\exp\left(c_2\left(x_1^2+x_2^2-9\right)\right)$$

$$\mu_1(k+1)=\mu_1(k)\exp\left(c_1(k)\left(x_1+x_2-1\right)\right)$$
$$\mu_2(k+1)=\mu_2(k)\exp\left(c_2(k)\left(x_1^2+x_2^2-9\right)\right)$$

$$c_1(k) = c / \mu_1(k)$$
$$c_2(k) = c / \mu_2(k).$$

Verilen başlangıç koşulları ile birlikte çeşitli nümerik metotlar kullanılarak problem çözülebilir. Şekil (4.2) deki grafik problemin Euler metodu ile Matlab (R2007b) programı yardımıyla elde edilmiştir.



Şekil 4.2: Örnek 2 de $x(t)$ 'nin grafiği $(x^* = (0,-3))$

Örnek 3 [26,problem no:227]**:**

Minimum	$f(x) = (x_1 - 2)^2 + (x_2 - 1)^2$
Kısıtlar	$g_1(x) = -x_1^2 + x_2 \geq 0$
	$g_2(x) = x_1 - x_2^2 \geq 0$

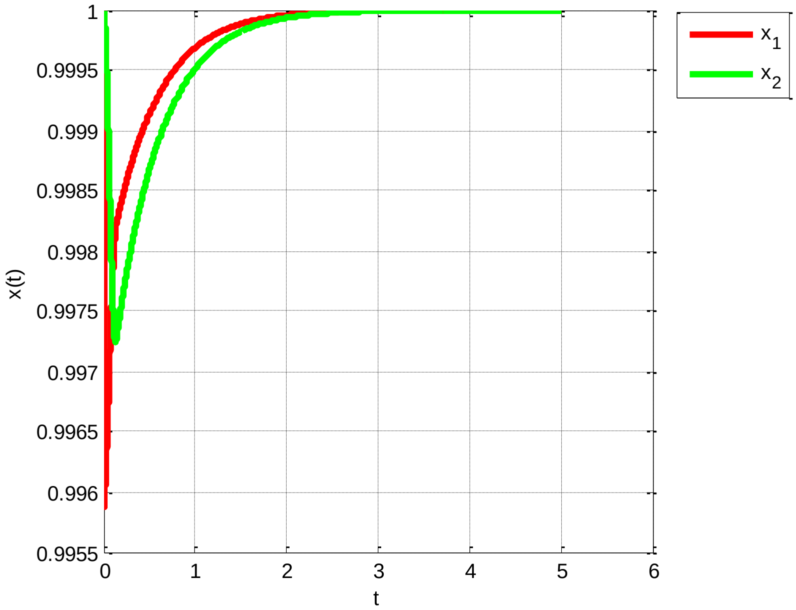
kısıtlı optimizasyon problemi ele alınsın. Verilen başlangıç değerleri $x_0 = (0.5, 0.5)$ şeklinde olup ulaşılacak optimum değer $x^* = (1,1)$ dir. Örnek1 deki gibi "$\geq$" eşitsizliği "$\leq$" eşitsizliğine dönüştürülebilir. Bölüm 4.2'deki metot kullanılarak elde edilecek sistem, "μ" çarpanı ve "c" ceza parametresi aşağıdaki gibi bulunur,

$$\frac{dx_1}{dt} = -2(x_1 - 2) - 2x_1\mu_1 \exp\left(c_1\left(x_1^2 - x_2\right)\right) + \mu_2 \exp\left(c_2\left(x_2^2 - x_1\right)\right)$$
$$\frac{dx_2}{dt} = -2(x_2 - 1) + \mu_1 \exp\left(c_1\left(x_1^2 - x_2\right)\right) - 2x_2 \exp\left(c_2\left(x_2^2 - x_1\right)\right)$$

$$\mu_1(k+1) = \mu_1(k)\exp\left(c_1(k)\left(x_1^2 - x_2\right)\right)$$
$$\mu_2(k+1) = \mu_2(k)\exp\left(c_2(k)\left(x_2^2 - x_1\right)\right)$$

$$c_1(k) = c / \mu_1(k)$$
$$c_2(k) = c / \mu_2(k).$$

Verilen başlangıç koşulları ile birlikte çeşitli nümerik metotlar kullanılarak problem çözülebilir. Şekil (4.3) deki grafik problemin Euler metodu ile Matlab (R2007b) programı yardımıyla elde edilmiştir.



Şekil 4.3: Örnek 3 de $x(t)$ 'nin grafiği $\left(x^* = (1,1)\right)$

Örnek 4 [26,problem no:249]**:**

$$
\begin{array}{ll}
\text{Minimum} & f(x) = x_1^2 + x_2^2 + x_3^2 \\
\text{Kısıtlar} & g_1(x) = x_1^2 + x_2 - 1 \geq 0 \\
 & g_2(x) = x_1 - 1 \geq 0
\end{array}
$$

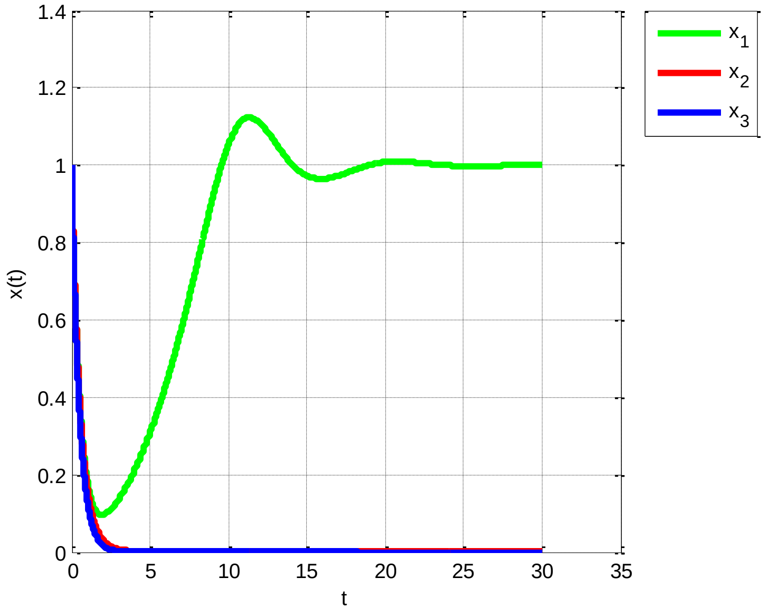
kısıtlı optimizasyon problemi ele alınsın. Verilen başlangıç değerleri $x_0 = (1,1,1)$ şeklinde olup ulaşılacak optimum değer $x^* = (1,0,0)$ dir. Örnek1 deki gibi "$\geq$" eşitsizliği "$\leq$" eşitsizliğine dönüştürülebilir. Bölüm 4.2' deki metot kullanılarak elde edilecek sistem, "μ" çarpanı ve "c" ceza parametresi aşağıdaki gibi bulunur,

$$
\begin{aligned}
\frac{dx_1}{dt} &= -2x_1 + 2x_1\mu_1 \exp\left(c_1\left(1 - x_1^2 - x_2^2\right)\right) + \mu_2 \exp\left(c_2\left(1 - x_1\right)\right) \\
\frac{dx_2}{dt} &= -2x_2 + 2x_2\mu_1 \exp\left(c_1\left(1 - x_1^2 - x_2^2\right)\right) \\
\frac{dx_3}{dt} &= -2x_3
\end{aligned}
$$

$$
\begin{aligned}
\mu_1(k+1) &= \mu_1(k)\exp\left(c_1\left(1 - x_1^2 - x_2^2\right)\right) \\
\mu_2(k+1) &= \mu_2(k)\exp\left(c_2\left(1 - x_1\right)\right)
\end{aligned}
$$

$$
\begin{aligned}
c_1(k) &= c / \mu_1(k) \\
c_2(k) &= c / \mu_2(k).
\end{aligned}
$$

Verilen başlangıç koşulları ile birlikte çeşitli nümerik metotlar kullanılarak problem çözülebilir. Şekil (4.4) deki grafik problemin Euler metodu ile Matlab (R2007b) programı yardımıyla elde edilmiştir.



Şekil 4.4: Örnek 4 de $x(t)$ 'nin grafiği $\left(x^* = (1,0,0)\right)$

Örnek 5 [26,problem no:235]**:**

$$\text{Minimum} \qquad f(x) = 0.01(x_1 - 1)^2 + \left(x_2 - x_1^2\right)^2$$
$$\text{Kısıtlar} \qquad g(x) = x_1 + x_3^2 + 1 = 0$$

şeklinde verilen optimizasyon problemi göz önüne alınsın. Verilen başlangıç değerleri $x_0 = (-2, 0, 1)$ şeklinde olup ulaşılacak optimum değer $x^* = (-1, 1, 0)$ dir. Verilen problem diğer bütün örneklerden farklı olup sadece eşitlik kısıtı içermektedir. Fakat bilinir ki eşitlik kısıtı iki eşitsizlik kısıtının birleşimidir. Böylelikle eşitlik kısıtlı problemler de eşitsizlik kısıtlarına çevrilerek bahsedilen metotla çözülebilir. O halde bu problem eşitsizlik kısıtı haline aşağıdaki şekilde getirilebilir,

$$\text{Minimum} \qquad f(x) = 0.01(x_1 - 1)^2 + \left(x_2 - x_1^2\right)^2$$
$$\text{Kısıtlar} \qquad g_1(x) = x_1 + x_3^2 + 1 \le 0$$
$$g_2(x) = +x_1 + x_3^2 + 1 \ge 0 \qquad .$$

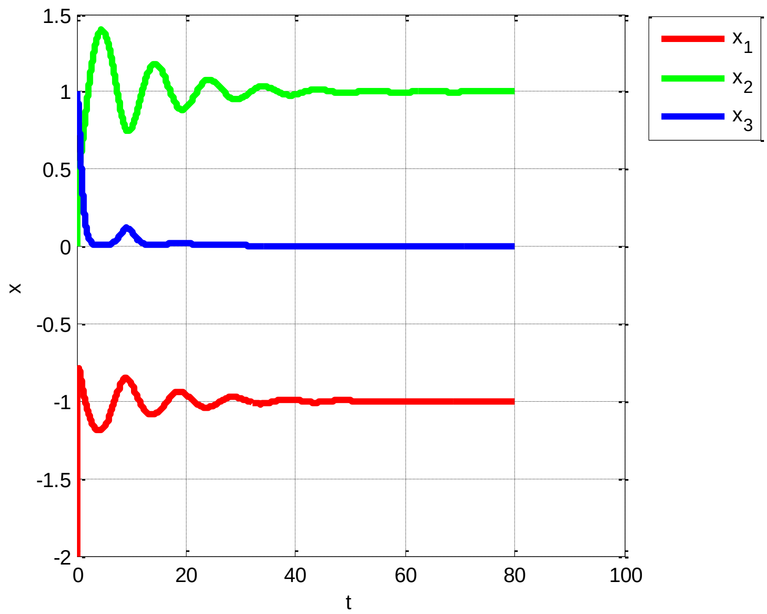
Yine bir önceki problemlerde olduğu gibi "≥" eşitsizliği "≤" eşitsizliğine çevrilebilir. Bölüm 4.2'deki metot kullanılarak elde edilecek sistem, "μ" çarpanı ve "c" ceza parametresi aşağıdaki gibi bulunur

$$\frac{dx_1}{dt} = -0.02(x_1 - 1) + 4x_1(x_2 - x_1^2) - \mu_1 \exp\left(c_1\left(x_1 + x_3^2 + 1\right)\right) - \mu_2 \exp\left(c_1\left(-x_1 - x_3^2 - 1\right)\right)$$
$$\frac{dx_2}{dt} = -2(x_2 - x_1^2)\mu_1 \exp\left(c_1\left(x_1 + x_3^2 + 1\right)\right) + 2x_3\mu_2 \exp\left(c_2\left(-x_1 - x_3^2 - 1\right)\right)$$
$$\frac{dx_3}{dt} = 2x_3$$

$$\mu_1(k+1) = \mu_1(k)\exp\left(c_1\left(x_1 + x_3^2 + 1\right)\right)$$
$$\mu_2(k+1) = \mu_2(k)\exp\left(c_2\left(-x_1 - x_3^2 - 1\right)\right)$$

$$c_1(k) = c / \mu_1(k)$$
$$c_2(k) = c / \mu_2(k).$$

Verilen başlangıç koşulları ile birlikte çeşitli nümerik metotlar kullanılarak problem çözülebilir. Şekil (4.5) deki grafik problemin Euler metodu ile Matlab (R2007b) programı yardımıyla elde edilmiştir.



Şekil 4.5 : Örnek 5 de $x(t)$'nin grafiği $(x^* = (-1,1,0))$

5. SONUÇLAR

Hazırlanan bu tezde doğrusal olmayan bir optimizasyon problemini çözmek için bir yöntem kullanılmıştır. İlk olarak problem yapısı ve kısıtların durumu incelenmiş ve en genel yapıda olan doğrusal olmayan optimizasyon probleminin çözümüne yaklaşılmıştır. Bu problem ilk olarak üstel ceza fonksiyonu yardımıyla kısıtsız bir optimizasyon problemi haline getirilmiş ve çözümün ilk aşaması geçilmiştir. Daha sonra bu kısıtsız optimizasyon problemi dinamik sistem yapısına çevrilmiştir. Problem ile elde edilen dinamik sistem arasındaki ilişki incelenmiş ve problemin çözümüne bu dinamik sistem yapısıyla yaklaşılmıştır.

Tezin ilk bölümünde optimizasyon problemi ve üstel ceza fonksiyonu ile ilgili literatürde yapılan çalışmalardan bahsedilmiş ve bir giriş yapılmıştır.

Tezin ikinci bölümünde problem yapısıyla ilgili bilgiler ve optimum çözüme sahip olması için gereken koşullar verilmiştir. Problem yapısı değişik şekillerde incelenmiştir.

Üçüncü bölümde ise problem ile oluşturulan dinamik sistem arasındaki ilişkiyi incelemek için ön bilgi olması bakımından kararlılık ile ilgili kavramlardan bahsedilmiştir.

Son bölümde ise üstel ceza fonksiyonu ile ilgili genel bilgiler verilmiş, problem ilk önce kısıtsız optimizasyon problemine dönüştürülmüş daha sonra dinamik sistem modeli oluşturulmuştur. Problem ile dinamik sistem arasındaki ilişki bir teorem ile gösterilmiştir. Devamında ise nümerik örnekler yapılmış ve bu örnekler Euler metodu ile Matlab programı kullanılarak çözülmüştür.

6. KAYNAKLAR

[1] Chong, K. P. E. and Zak, H.S. , *An Introduction to Optimization,* third edition, New Jersey: John Wiley and Sons, (2008).

[2] Luenberger D. G. and Ye Y., *Linear and Nonlinear Programming*, third edition New York: Springer Science+Business Media, LLC, (2008).

[3] Bazara, M. S., Sherali, H. D. and Shetty, C. M., *Nonlinear programming: Theory and Algorithms*, New York : John Willey and Sons , (1991) .

[4] Gill, P.E., Murray, W. and Wright, M.H., *Practical Optimization*, Academic Press Limited: Harcourt Brace and Company, (1997).

[5] Zangwill, W. I., "Nonlinear programming via penalty functions", *Management Science*, 13, 344-358, (1967)

[6] Özdemir, N. and Evirgen, F., "A Dynamic System Approach to Quadratic Programming Problems with Penalty Method", *Bull. Malays. Math. Sci. Soc.* , 33, 1, 79-91, (2010).

[7] Evirgen, F., "Optimizasyon problemlerinin optimum değerlerinin Durum Denklemleri Modeli ile Araştırılması", Doktora Tezi, *Balıkesir Üniversitesi Fen bilimleri Enstitüsü*, Matematik Anabilim Dalı, Balıkesir, (2009).

[8] Fiacco, A. V and McCorMic, G. P., *Nonlinear Programming: Sequential Unconstrained Minimization Techniques,* New York: John Willey and Sons, (1968).

[9] Motzkin, T. S., *New Technique for linear inequalities and optimization*, U.S. Air Force, Washington D. C.,in : Project SCOOP Symposium on Linear Inequalities and Programming, Planning Researc Division , (1952).

[10] Tseng, P. and Bertsekas, P. D., "On The Convergence of The Exponential Multiplier Method for Convex Programming" *Math Programming*, 60, 1-19,

(1993).

[11] Alvarez, F., “Absolute minimizer in convex programming by exponential penaly” , *Journal of Convex Analysis*, 7 (1), 197-202, (2000).

[12] Grigoriadis, M. and Khachiyan, L. “An exponential –function reduction method for block-angular convex programs” , *Networks*, 26 (2), 59-68, (1995).

[13] Mouallif, K. And Tossing P., “Une méthode de penalisation exponentielle associée á une régularisation proximalle”, *Bulletin of Society Royal Science Liege*, 56 (2), 181-192, (1987).

[14] Murphy, F., “A class of exponential penalty functions”, *SIAM Journal on Control*, 12, 679-687, (1974).

[15] Strodiot, J. J. and Nguyen V. H., “An exponential penalty method for non-differentiable minimax problems with general constraints” , *Journal of Optimization , Theory and Applications*, 27, 205-219, (1979).

[16] Nazemi, A.R., “A Dynamic System Model For Solving Convex Nonlinear Optimization Problems”, *Commun Nonlinear Sci Numer Simulat* 17, 1696-1705, (2012).

[17] Lillo, W. E., Loh, M.E., Hui, S. and Zak, S.H., “On Solving Constrained Optimization Problems with Neural Networks: A Penalty Function Method Approach”, *Electrical and Computer Engineering ECE Technical Reports*, 1-35, (1991).

[18] Chen, K. Z., Leung, Y., Leung, K.S. and Gao, X.b., “ A Neural Network For Solving Nonlinear Programming Problems”, *Neural Comput & Applic*. 11, 103-111, (2002).

[19] Effati, S. and Baymani, M., “A new nonlinear neural network for solving convex nonlinear programming problems”, *Applied Mathematics and Computation*, 168, 1370-1379, (2005).

[20] Zhou, L. and Zhang, L., “A log sigmoid Lagrangian neural network for solving

nonlinear programming”, *Eight ACIS Intenational Conference on Software Engineering, Artificial Intelligence, Networking, and Parallel/Distributed Computing*,doi: 10.1109/SNPD.2007.15, 427-431, (2007).

[21] Mangasarian, O.L., *Nonlinear Programming*, New York: McGraw Hill Book Co, (1969).

[22] Khalil, K. H., *Nonlinear Systems*, second edition, New Jersey :Prentice-Hall Inc Simon and Schuster / A Viacom Company, (1996).

[23] Bertsekas P.D., *Nonlinear Programming*, 2nd edition , Massachusetts Institute of Technology, (2003).

[24] Kort, B.W. and Bertsekas D.P., “A New Penalty Function Method For Constrained Minimization”, *Proceeding of the 1972 IEE Confer. Decision Control*, 162-166, (1972).

[25] Cominetti, R., “Asymptotic Convergence of the Steepest Decent Method for the Exponential Penalty in Linear Programming” *Journal of Convex Analysis*, 2, No:1-2, 145-152, (1995).

[26] Schittkowski, K., *Lecture Notes in Economics and Mathematical Systems* New York : Spring-Verlag Berlin Heidelberg, (1987).