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

OPTİMİZASYON PROBLEMLERİNİN DURUM DENKLEMLERİ İLE OPTİMUM DEĞERLERİNİN ARAŞTIRILMASI

DOKTORA TEZİ

Fırat EVİRGEN

Balıkesir, Haziran-2009

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

OPTİMİZASYON PROBLEMLERİNİN DURUM DENKLEMLERİ İLE OPTİMUM DEĞERLERİNİN ARAŞTIRILMASI

DOKTORA TEZİ

Fırat EVİRGEN

Tez Danışmanı: Yrd. Doç. Dr. Necati ÖZDEMİR

Sınav Tarihi: 18.06.2009

Jüri Üyeleri: Prof. Dr. Ramazan YAMAN (BAÜ)

Doç. Dr. Osman BİZİM (UÜ)

Doç. Dr. Fatma AYAZ (GÜ)

Yrd. Doç. Dr. Necati ÖZDEMİR (Danışman-BAÜ)

Yrd. Doç. Dr. Yunus Emre YILDIRIR (BAÜ)

Balıkesir, Haziran-2009

“Bu çalışma Balıkesir Üniversitesi Rektörlüğü Bilimsel Araştırma Projeleri Birimi tarafından BAP 2007/02 kodlu Proje ile desteklenmiştir. Teşekkür ederiz.”

ÖZET

OPTİMİZASYON PROBLEMLERİNİN DURUM DENKLEMLERİ İLE OPTİMUM DEĞERLERİNİN ARAŞTIRILMASI

Fırat EVİRGEN
Balıkesir Üniversitesi, Fen Bilimleri Enstitüsü,
Matematik Anabilim Dalı

(Doktora Tezi / Tez Danışmanı: Yrd. Doç. Dr. Necati ÖZDEMİR)

Balıkesir, 2009

Bu çalışmanın amacı optimizasyon problemlerinin optimum çözümlerinin araştırılmasında durum denklemleri yaklaşımının işlevselliğini göstermektir.

Bu çalışma beş bölümden oluşmaktadır. Birinci bölümde, optimizasyon problemlerinin çözümünde önemli yere sahip ceza fonksiyonu ve gradyant akış yöntemlerinin gelişimi hakkında kronolojik bilgiye yer verilmiştir.

İkinci bölümde, tezin diğer bölümlerinde başvurulacak optimizasyon, durum denklemleri ve kararlılık konuları ile ilgili bazı iyi bilinen tanım ve teoremler verilmektedir.

Üçüncü bölümde, kısıtlamasız optimizasyon problemlerine durum denklemleri yaklaşımı uygulanarak problemlerin optimum çözümleri araştırılmış ve kararlılık analizi yapılmıştır. Ayrıca bu konu ile ilgili bazı test problemlerine yer verilmiştir. Elde edilen sonuçlar Quasi-Newton yöntemi ile karşılaştırılarak grafik ve tablolar ile gösterilmiştir.

Dördüncü bölüm, üç kısımdan oluşmaktadır. İlk kısımda kuadratik programlama problemlerinin çözümlerinde durum denklemleri yaklaşımı uygulanmış ve yaklaşım teoremleri ortaya konmuştur. İkinci kısımda ise durum denklemleri yaklaşımının doğrusal olmayan programlama problemleri üzerindeki araştırmalara yer verilmiştir. Son kısımda ise diferansiyellenemeyen l_1 ceza fonksiyonu için düzgün bir ceza fonksiyon tanımlanmış ve durum denklemleri yaklaşımı uygulanmıştır.

Son bölümde, elde edilen sonuçların bir değerlendirmesi yapılmıştır.

ANAHTAR SÖZCÜKLER : Kuadratik Programlama, Doğrusal Olmayan Programlama, Durum Denklemleri, Gradyant Akış, Kararlılık.

ABSTRACT

INVESTIGATING OPTIMUM VALUES OF OPTIMIZATION PROBLEMS WITH STATE SPACE APPROACH

Fırat EVİRGEN
Balıkesir University, Institute of Science,
Department of Mathematics

(Ph.D. Thesis / Supervisor : Assist. Prof. Dr. Necati ÖZDEMİR)

Balıkesir - Turkey, 2009

The aim of this work is to show feasibility of state space approach for investigating the optimum solutions of optimization problems.

The work consists of five chapters. In the first chapter, some chronological information about penalty function and gradient flow method, which have an important role in solution of optimization problem, are given.

In the second chapter, some well-known definitions and theorems about optimization, state space equations and stability needed later during the thesis are presented.

In the third chapter, the optimal solutions of the problems are investigated, and the stability analysis is done by applying the state space approach to unconstrained optimization problems. Furthermore, some test problems concerning this issue are given. Graphical representations and tables show obtained solutions by comparing Quasi-Newton method.

The fourth chapter is consists of three sections. In the first section, the state space approach is applied to the quadratic programming problems, and approximation theorems are presented. In the second section, the investigations over the nonlinear programming problems with the state space approach are given. In the final section, a smooth penalty function is defined for non-smooth l_1 penalty function, and state space approach is applied.

In the last chapter, the results that are obtained according to the previous chapters have been summarized.

KEY WORDS : Quadratik Programming, Nonlinear Programming, State Space Equation, Gradient Flow, Stability.

İÇİNDEKİLER

SEMBOL LİSTESİ

Simge	**Tanımı**
$\mathbb{R}$	Reel sayılar kümesi
$\mathbb{R}^n$	n boyutlu reel vektör uzayı
$\mathbb{R}^{n\times n}$	$n\times n$ boyutlu reel elemanlı matrislerin kümesi
$\mathbb{C}$	Karmaşık sayılar kümesi
S	Uygun küme
C^p	p. dereceden sürekli kısmi türevlere sahip fonksiyonlar uzayı
A^T	A matrisinin transpozesi
$\nabla f(x)$	$f(x)$ fonksiyonunun gradyant vektörü
$\nabla^2 f(x)$	$f(x)$ fonksiyonunun Hesse matrisi
H	Hesse matrisi
$\frac{dx}{dt}=\dot{x}$	$x(t)$ fonksiyonunun t ‘ye göre birinci mertebeden türevi
$diag(A)$	Köşegenleştirilmiş A matrisi
$I_{n\times n}$	$n\times n$ boyutlu birim matris
L	Laplace dönüşümü
L^{-1}	Ters Laplace Dönüşümü
$\dot{x}(t)$	x vektörünün t zamanına göre türevi
$\mathrm{Re}(\lambda)$	λ karmaşık sayısının reel kısmı
$\mathrm{Im}(\lambda)$	λ karmaşık sayısının imajinal kısmı

ŞEKİL LİSTESİ

ÇİZELGE LİSTESİ

ÖNSÖZ

Bu çalışma süresince değerli vaktini ayırıp, bilgi ve tecrübeleri ile beni yönlendiren, her türlü kaynağını, ilgisini, desteğini ve yardımlarını benden esirgemeyen değerli hocam ve danışmanım Yrd. Doç. Dr. Necati ÖZDEMİR'e;

Bu günlere gelmemi sağlayan, sevgileri ve ilgileri ile hep yanımda olan canım aileme, ilk günden beri göstermiş olduğu sabır ve destek için sevgili eşim Okşan'a ve tez çalışmamın son aylarında dünyaya gelerek hayatımıza yeni bir soluk getiren biricik kızım Elif'e sonsuz teşekkürlerimi sunarım.

Balıkesir, 2009 **Fırat EVİRGEN**

1. GİRİŞ

Doğrusal veya doğrusal olmayan programlama mühendislik, ekonomi, finans, yönetim gibi pek çok alanda uygulamalara sahip olup bu alanlardaki bir çok problem çeşitli dönüşümler ile matematiksel olarak modellenebilmektedir. Oluşturulan matematiksel modellerin en iyi çözümlerinin araştırılmasında literatürde bahsedilen bir çok metot geliştirilmiştir. Bu metotlar ile ilgili bilgiler D.P. Bertsekas [1], E.K.P. Chong [2], A.V. Fiacco [3], D.G. Luenberger [4], S.G. Nash [5] ve W. Sun [6] literatürlerinde ayrıntılı olarak yer almaktadır.

Bu yöntemlerden ceza (penalty) fonksiyonları metodu kısıtlamalı doğrusal veya doğrusal olmayan optimizasyon problemlerinin optimum çözümlerinin araştırılmasında sıklıkla kullanılmaktadır. Ceza metotları kısıtlamalı optimizasyon problemlerinin çözümlerinde belirli özelliklere sahip fonksiyonlar yardımı ile kısıtlamaların amaç fonksiyonuna ekleyerek kısıtlamasız optimizasyon problemleri dizisine dönüştüren bir prosedür izlemektedir. Eşitsizlik kısıtlamaları için sıklıkla kullanılan l_1 ceza fonksiyonu W.I. Zangwill [7] tarafından tanımlanmıştır. Ayrıca D.D. Morrison [8], A.R. Conn [9] ve C.G. Broyden [10,11] farklı özelliklere sahip ceza fonksiyonları oluşturmuşlardır. Son zamanlarda ise Z. Meng [12], D.D. Morrison'un [8] ifade ettiği ceza fonksiyonunu genişleterek doğrusal olmayan optimizasyon problemlerine uygulamıştır. Ayrıca A.M. Rubinov [13,14], ve X.Q. Yang [15] doğrusal olmayan Lagrangian adı altında yeni bir ceza fonksiyonu tanımlamışlardır.

Ceza fonksiyonları metodunda tam (exact) olma yani kısıtlamalı optimizasyon probleminin optimum çözümünün kısıtlamasız optimizasyon problemlerinin bir dizisini değil de tek bir kısıtlamasız problemin çözümü ile ulaşmak isteği araştırmacıların her zaman ilgisini çekmiştir. Bu konuda ilk çalışmalar D.P. Bertsekas [1] tarafından ortaya atılmış, S.P. Han, O.L. Mangasarian [16] ve G. Di Pillo [17,18] literatürlerinde geliştirilmiştir.

Daha sonraki araştırmalarda ceza fonksiyonlarının bir yandan tamlık özelliğini kazanırken diğer yandan ise diferansiyellenebilme özelliklerini kaybettikleri gözlenmiştir. Bu ise fonksiyonların türevini kullanarak optimum noktaların araştırılmasını gerçekleştiren bazı optimizasyon metotlarının kullanımını engellemiştir. Bunu üzerine diferansiyellenemeyen ceza fonksiyonlarının yerine aynı özellikleri taşıyan fakat her yerde türevlenebilen düzgünleştirilmiş fonksiyonlar oluşturulmaya başlanmıştır. M.C. Pınar ve S.A. Zenios [19,20] tanımladıkları düzgünleştirilmiş ceza fonksiyonu yardımı ile konveks kısıtlamalı optimizasyon problemlerinin optimum noktalarını araştırmışlardır. Ayrıca, C. Chen [21] sigmoid fonksiyonunu integre ederek elde ettikleri düzgün bir fonksiyon yardımı ile $\max\{0,t\}$ ceza fonksiyonu için optimizasyon problemlerinin yaklaşık çözümlerini incelemişlerdir. Son zamanlarda ise X.Q. Yang [22] ve Z-Q. Meng [23] diferansiyellenemeyen ceza fonksiyonları için farklı düzgün yaklaşımlar vermişlerdir.

Optimizasyon problemlerinin çözümlerinde ceza fonksiyonları metodu dışında Gradyant akış (Gradient flow) ve adi diferansiyel denklem (ODE) tabanlı metotlar da kullanılmaktadır. Optimizasyon problemlerinin minimumlarının araştırılmasında diferansiyel denklem sistemlerinin kullanımı ilk olarak K.J. Arrow [24,25] tarafından ortaya atılmıştır. Farklı uygulamaları ise P.Q Pan [26] ve L. Jin [27,28] literatürlerinde yer almaktadır. A.A. Brown [29,30] ise ODE metodunun klasik çözüm yöntemlerine göre pozitif yönlerini ortaya koymuştur.

Gradyant akış yöntemi ise Y.G. Evtushenko [31,32] tarafından doğrusal ve doğrusal olmayan programlama problemleri için incelenmiştir. S. Wang [33] elde edilen bu sonuçları farklı iterasyon yöntemleri kullanarak geliştirmiştir. N. Andrei [34] ise bu yaklaşımı kısıtlamasız optimizasyon problemleri için uygulamıştır. Ayrıca J. Schroop [35] kısıtlamalı optimizasyon problemlerinin optimumlarını tanımladığı bir dinamik sistem altında incelemiştir. Optimal kontrol problemlerinin bir dizisi yardımı ile kısıtlamasız optimizasyon problemlerinin çözümleri B.S. Goh [36] tarafından incelenmiştir.

Bu tezde çeşitli sınıflardaki optimizasyon problemlerinin optimum çözümlerinin araştırılmasında durum denklemleri yaklaşımı uygulanmış, elde edilen çözümlerin kararlılık analizi yapılmıştır, [37-39]. Ayrıca diferansiyellenemeyen l_1 ceza fonksiyonu için düzgün bir fonksiyon tanımlanarak optimizasyon problemlerinin çözümünde durum denklemleri yaklaşımı ile kullanılarak optimum noktaya yaklaşımı ispatlanmıştır, [40].

2. ÖN BİLGİLER

Bu bölümde 3. ve 4. bölümlerde yararlanacağımız bazı temel kavramlar ele alınacaktır.

2.1 Optimizasyon

İnsanların yaşamları boyunca karşılaştıkları sorunları çözüm arayışları zamanla bu çözümleri modeller üzerinde arama yaklaşımını doğurmuştur. Matematik ve bilgisayardaki gelişmeler dış dünyanın problemlerini matematiksel olarak modelleyip, buradan elde edilen çözümleri gerçek hayata yansıtma olanağı vermiştir.

Matematiksel modelleme tekniği öncelikle doğrusal ve az sayıda değişkenlerin kullanılmasıyla başlamıştır. Bir süre sonra doğrusallık varsayımının her problem için geçerli olmadığı anlaşılmıştır. Bu durumda doğrusal olmayan modellemeye gidilmiştir. Ancak doğrusal olmayan modellerin kendine özgü çözümleri uygulamada birçok sorunu beraberinde getirmiştir. Zamanla geliştirilen bazı yöntemlerle doğrusal olmayan modellerin hızla çözümlenmesi sağlanmış ve bu optimizasyon teorisini geliştirmiştir.

Optimizasyon modellerinin amacı bir problemin en iyi çözümünü matematiksel terimler ile açıklamaktır. Bunun anlamı kârı veya etkiyi maksimize etmek olabileceği gibi zararı veya riski minimize etmekte olabilir.

Genel olarak bir optimizasyon probleminin matematiksel modellemesi,

$$\text{minimum} \quad f(x) \tag{2.1}$$

$$\text{kısıtlamalar} \quad h_i(x) = 0, \qquad i = 1, 2, \ldots, m \tag{2.2}$$

$$g_j(x) \leq 0, \qquad j = 1, 2, ..., p \tag{2.3}$$

yukarıdaki biçimde ifade edilebilir. Bu formulasyonda $x = (x_1, x_2, ..., x_n)^T$ n-boyutlu bir vektör olmak üzere karar değişkenini, $f : \mathbb{R}^n \to \mathbb{R}$, $h = (h_1, h_2, ..., h_m)^T : \mathbb{R}^n \to \mathbb{R}^m$ ve $g = (g_1, g_2, ..., g_p) : \mathbb{R}^n \to \mathbb{R}^p$ de reel değerli fonksiyonları ifade etmektedir. Problem yapısındaki f fonksiyonu problemin amaç fonksiyonu, h ve g fonksiyonları da sırayla problemin eşitlik ve eşitsizlik kısıtlamaları olarak adlandırılır.

Genellikle literatürde optimizasyon problemleri modelin yapısını oluşturan bileşenlerin biçimine göre çeşitli sınıflara ayrılabilirler. Bu sınıflamanın oluşmasında en önemli etkenler, kısıtlamalarının olup olmadığı ve fonksiyonların çeşididir. Buna göre (2.2)-(2.3) kısıtlamalarına sahip olmayan problemler kısıtlamasız optimizasyon problemleri, aksi takdirde kısıtlamalı optimizasyon problemleri olarak adlandırılır.

Şimdi konu ile ilgili olarak bazı temel tanım ve teoremlere yer verelim.

2.1.1 Tanım: Bir A simetrik matrisinin pozitif tanımlı olması için gerek şart $\forall x \neq 0$ vektörü için $x^T Ax$ kuadratik formunun pozitif olmasıdır, [4].

Benzer şekilde pozitif yarı tanımlılık, negatif ve yarı negatif tanımlılık $\forall x \neq 0$ için $x^T Ax \geq 0$, < 0 ve ≤ 0 biçiminde sırasıyla tanımlanabilir.

2.1.2 Teorem: Bir A $n \times n$ boyutlu simetrik matrisinin pozitif tanımlı olması için gerek ve yeter şart A matrisinin tüm minörleri pozitif olmalıdır, [41].

2.1.3 Teorem: Pozitif tanımlı bir matrisin her bir özdeğeri bir pozitif reel sayıya karşılık gelir, [41].

2.1.4 Teorem: Bir A hermit matrisinin pozitif yarı tanımlı olması için gerek ve yeter şart tüm özdeğerlerinin negatif olmamasıdır. Yine, pozitif tanımlı olması için gerek ve yeter şart tüm özdeğerlerinin pozitif olmasıdır, [41].

2.1.5 Tanım: Bir S kümesinin konveks olması için gerek şart $\forall x, y \in S$ ve $\forall \ 0 \leq \alpha \leq 1$ için

$$\alpha x + (1-\alpha) y \in S$$

olmasıdır. Bir başka deyişle x ile y 'yi birleştiren her doğru parçası yine S içinde kalıyorsa S kümesi konvekstir denir, [42].

2.1.6 Tanım: Bir f fonksiyonunun konveks bir S kümesi üzerinde konveks olabilmesi için gerek şart

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

sağlanmasıdır, [42].

2.1.7 Önerme: $f \in C^1$ olsun. Buradan f 'nin konveks bir Ω kümesinde konveks olması için gerek ve yeter şart $\forall x, y \in \Omega$ için

$$f(y) \geq f(x) + \nabla f(x)(y-x)$$

olmasıdır, [4].

2.1.8 Önerme: $f \in C^2$ olsun. Buradan f 'nin bir iç nokta içeren konveks bir Ω kümesinde konveks olması için gerek ve yeter şart f 'nin Hesse matrisi F 'nin Ω üzerinde pozitif yarı tanımlı olmasıdır, [4].

2.1.9 Tanım: $f:\mathbb{R}^n \to \mathbb{R}$**,** sürekli diferansiyellenebilir bir fonksiyon ve $p \in \mathbb{R}^n$ olsun. Buradan $\exists t \in (0,1)$ için

$$f(x+p) = f(x) + \nabla f(x+tp)^T p$$

sağlanır. Dahası f iki kez sürekli ve diferansiyellenebilir ise $\exists t \in (0,1)$ için

$$\nabla f(x+p) = \nabla f(x) + \int_0^1 \nabla^2 f(x+tp)\, p dt$$

ve

$$f(x+p) = f(x) + \nabla f(x)^T p + \frac{1}{2} p^T \nabla^2 f(x+tp)\, p$$

elde edilir, [42].

2.1.10 Tanım: Kısıtlamaları sağlayan herhangi bir noktaya uygun nokta (feasible point) denir. Bütün uygun noktaların oluşturduğu

$$S = \{x : x \in \mathbb{R}^n, h_i(x) = 0, g_j(x) \leq 0,\ i = 1,2,...,m,\ j = 1,2,...,p\}$$

kümesine de uygun küme (feasible set) denir, [2].

2.1.11 Tanım: Bir $x*$ noktasının f fonksiyonunun mutlak (global) minimumu olabilmesi için gerek şart

$$\forall x \in \mathbb{R}^n \text{ için } f(x*) \leq f(x)$$

olmasıdır. Eğer $\forall x \in \mathbb{R}^n$, $x \neq x*$ için $f(x*) < f(x)$ ise $x*$ noktasına f 'nin kesin mutlak (strict global) minimumu denir, [4].

2.1.12 Tanım: Bir x^* noktasının f fonksiyonunun yerel (lokal) minimumu olabilmesi için gerek şart $\forall \varepsilon > 0$ olmak üzere $\|x - x^*\| < \varepsilon$ koşulunu sağlayan x noktaları için $f(x^*) \leq f(x)$ olmasıdır. Eğer $\forall \varepsilon > 0$, $x \in \mathbb{R}^n$ ve $x \neq x^*$ için $f(x^*) < f(x)$ ise x^* noktasına f 'nin kesin yerel (strict local) minimumu denir, [4].

Mutlak ve yerel minimum tanımları, bir optimal çözümün sağlanmasında her zaman yeterli olmayabilir. Bu nedenle optimal noktayı bulmak için daha pratik koşullara ihtiyaç vardır. İzleyen kısımda bu koşullarla ilgili bazı temel tanım ve teoremleri kısıtlamasız ve kısıtlamalı optimizasyon problemleri için verilecektir.

2.1.1 Kısıtlamasız Optimizasyon

Çalışmanın bu kısmında $S = \mathbb{R}^n$ alarak kısıtlamasız optimizasyon problemleri için uygun çözüm bölgesini tanımlanacaktır.

2.1.1.1 Tanım: Bir $d \in \mathbb{R}^n$, $d \neq 0$ vektörünün $x \in S$ noktasında uygun (feasible) yön olabilmesi için gerek şart, $\forall \alpha \in [0, \alpha_0]$ ve $\alpha_0 > 0$ değeri için $x + \alpha d \in S$ sağlanmalıdır, [2].

2.1.1.2 Teorem [Birinci Mertebeden Gerek Şart]: $S \subset \mathbb{R}^n$ ve $f \in C^1$, S kümesi üzerinde reel değerli bir fonksiyon olsun. Eğer x^*, f 'nin S üzerindeki bir yerel minimumu ise x^* noktasındaki en az bir d uygun yönü için $d^T \nabla f(x^*) \geq 0$ olmalıdır, [2].

2.1.1.3 Yardımcı Teorem: $S \subset \mathbb{R}^n$ ve $f \in C^1$, S kümesi üzerinde reel değerli bir fonksiyon olsun. Eğer x^*, f 'nin S üzerindeki bir yerel minimumu ve S bölgesinin bir iç noktası ise $\nabla f(x^*) = 0$ olmalıdır, [2].

2.1.1.4 Teorem [İkinci Mertebeden Gerek Şart]: $S \subset \mathbb{R}^n$ ve $f \in C^2$, S kümesi üzerinde reel değerli bir fonksiyon olsun. $x*$ noktası, f 'nin S üzerindeki bir yerel minimumu ve $x*$ noktasındaki uygun bir yön d olsun. Eğer $d^T \nabla f(x*) = 0$ ise H, f 'nin Hesse matrisi olmak üzere $d^T H(x*)d \geq 0$ olmalıdır, [2].

2.1.1.5 Yardımcı Teorem: $f \in C^2$ ve $x*$ noktası S 'nin bir iç noktası olsun. Eğer $x*$ noktası f fonksiyonunun bir lokal minimumu ise $\nabla f(x*) = 0$ ve $H(x*) \geq 0$ olmalıdır, [2].

2.1.1.6 Teorem [İkinci Mertebeden Yeter Şart]: $f \in C^2$ ve $x*$ noktası S 'nin bir iç noktası olsun. Farz edilsin ki $\nabla f(x*) = 0$ ve $H(x*) > 0$ ise $x*$ noktası f fonksiyonunun bir lokal minimumudur, [2].

2.1.1.7 Teorem: f konveks bir fonksiyon olsun. Bu durumda herhangi bir yerel minimumu $x*$ aynı zamanda f 'nin bir mutlak minimumudur. Eğer f diferansiyellenebilir ise f 'nin herhangi bir kritik noktası aynı zamanda bir mutlak minimumdur, [42].

2.1.2 Kısıtlamalı Optimizasyon

Kısıtlamalı optimizasyonda temel kavramlardan biri aktif kısıtlama kavramıdır. Herhangi $g(x) \leq 0$ eşitsizlik kısıtlamasının x uygun noktasında aktif olabilmesi için gerek şart $g(x) = 0$ olmasıdır. Eğer $g(x) < 0$ ise aktif değildir denir. Bu tanım gereği optimizasyon problemlerindeki $h(x) = 0$ eşitlik kısıtlamalarının her biri herhangi bir uygun noktada daima aktiftir.

Şimdi eşitlik kısıtlamalarına sahip (2.1)-(2.2) optimizasyon problemi için gerek ve yeter optimallik şartlarını verelim.

2.1.2.1 Teorem: Eğer $\nabla h_1(x^*),\ \nabla h_2(x^*),\ ...,\nabla h_m(x^*)$ gradyant vektörleri lineer bağımsız ise $h(x^*)=0$ kısıtlamasını gerçekleyen x^* noktasına düzgün (regular) bir noktadır denir, [4].

2.1.2.2 Teorem: Eşitlik kısıtlaması $h(x)=0$ ile tanımlanan S yüzeyinin bir x^* düzgün noktasındaki teğet uzayı

$$M=\{y:\nabla h(x^*)y=0\}$$

ile tanımlanır, [4].

2.1.2.3 Önerme: Bir x^* noktası $h(x)=0$ kısıtlamalarının bir düzgün noktası ve eşitlik kısıtlamalarına sahip bir optimizasyon probleminin bir yerel ekstremum (minimum yada maksimum) noktası olsun. Buradan $\forall y\in\mathbb{R}^n$ için

$$\nabla h(x^*)y=0 \text{ ve } \nabla f(x^*)y=0$$

olmalıdır, [4].

2.1.2.4 Teorem: Bir x^* noktası eşitlik kısıtlamalarına sahip optimizasyon problemlerinin bir yerel ekstremum noktası ayrıca bu kısıtlamaların bir düzgün noktası olsun. Bu durumda

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

sağlanacak şekilde bir $\lambda\in\mathbb{R}^m$ değeri vardır, [4].

Burada $h(x^*)=0$ eşitlik kısıtlamaları ile birlikte $\nabla f(x^*)+\lambda\nabla h(x^*)=0$ koşulu birinci mertebeden gerek şart olarak adlandırılır. Bu sistem x^* ve λ içeren genelde $n+m$ değişkenli $n+m$ doğrusal olmayan denklem sisteminden oluşur. Bu sistem kısıtlamalı optimizasyon problemlerinin çözümünü elde etmek için kullanılır.

Denklemdeki λ değerine Lagrange sabiti denir ve kısıtlamalı optimizasyon problemine karşılık gelen Lagrange fonksiyonu

$$l(x,\lambda) = f(x) + \lambda h(x)$$

biçiminde tanımlanır.

Eşitlik kısıtlamalı optimizasyon problemleri için optimallik koşulları, kısıtlamasız optimizasyon problemleri kısmındaki koşullara benzer olarak elde edilebilir.

2.1.2.5 Teorem [İkinci Mertebeden Gerek Şart]: Bir $x*$ noktası f fonksiyonunun $h(x)=0$ kısıtlamasının altında bir yerel minimumu ve bu kısıtlamaların düzgün bir noktası olsun. Bu durumda

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

olacak şekilde bir $\lambda \in \mathbb{R}^m$ değeri vardır. Eğer M ile teğet uzayını gösterirsek,

$$L(x^*) = F(x^*) + \lambda H(x^*)$$

ifadesi, L Lagrange fonksiyonunun, F $f(x)$ fonksiyonunun ve H eşitlik kısıtlamalarının Hesse matrisi olmak üzere M üzerinde pozitif yarı tanımlı olmalıdır, [4].

2.1.2.6 Teorem [İkinci Mertebeden Yeter Şart]: Eşitlik kısıtlaması $h(x^*)=0$ ve $\lambda \in \mathbb{R}^m$ için

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

koşulunu sağlayacak şekilde bir $x*$ noktası olsun. Aynı zamanda

$$L(x^*) = F(x^*) + \lambda H(x^*)$$

pozitif tanımlı ise x^* noktası f fonksiyonunun bir yerel minimumudur, [4].

Eşitsizlik kısıtlamalarına sahip optimizasyon problemlerinin optimallik koşulları içinde benzer tanım ve teoremler aşağıdaki biçimde ifade edilebilir.

2.1.2.7 Tanım: Bir x^* noktası $h(x^*)=0$ ve $g(x^*)\leq 0$ kısıtlamalarını gerçekleyen nokta olsun. J, $g_j(x^*)=0$ koşulunu sağlayan j indislerinin bir kümesi olarak alınsın. x^*'ın düzgün (regular) bir nokta olması için gerek şart $\nabla h_i(x^*)$, $\nabla g_j(x^*)$ ifadeleri $1\leq i\leq m$ ve $j\in J$ için lineer bağımsız olmalıdır, [4].

2.1.2.8 Teorem [Karush-Kuhn-Tucker Koşulları]: x^* noktasının (2.1)-(2.3) optimizasyon probleminin bir minimum noktası ve kısıtlamalar içinde düzgün bir nokta olduğunu kabul edelim. Bu durumda öyle $\lambda\in\mathbb{R}^m$ ve $\mu\in\mathbb{R}^p$ vektörleri vardır ki

a) $\mu\geq 0$

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

c) $\mu\nabla g(x^*)=0$

olmalıdır.

2.1.2.9 Teorem: $f,\ g,\ h\in C^2$ ve x^* kısıtlamaların düzgün bir noktası olsun. Eğer x^* noktası (2.1)-(2.3) probleminin bir yerel minimumu ise $\lambda\in\mathbb{R}^m$ ve $\mu\in\mathbb{R}^r$, $\mu\geq 0$ vektörleri için

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

$$\mu\nabla g(x^*)=0$$

koşulları gerçeklenir ve

$$L(x^*)=F(x^*)+\lambda H(x^*)+\mu G(x^*)$$

G, eşitsizlik kısıtlamalarının Hesse matrisi olmak üzere x^* noktasındaki aktif kısıtlamaların teğet alt uzayında pozitif yarı tanımlı olmalıdır, [4].

2.1.2.10 Teorem: $f,\ g,\ h \in C^2$ olsun. x^* noktası (2.1)-(2.3) probleminin bir yerel minimumu olması için yeter şart, $\lambda \in \mathbb{R}^m$ ve $\mu \in \mathbb{R}^r$ vektörleri için

$$\mu \geq 0$$

$$\mu \nabla g(x^*) = 0$$

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

ve Hesse matrisi

$$L(x^*) = F(x^*) + \lambda H(x^*) + \mu G(x^*)$$

indis kümesi $J = \left\{ j : g_j(x) = 0, \mu_j > 0 \right\}$ olmak üzere alt uzayı

$$M' = \left\{ y : \nabla h(x^*) y = 0, \nabla g_j(x^*) y = 0, \forall j \in J \right\}$$

için pozitif tanımlı olmalıdır, [4].

2.1.3 Ceza (Penalty) ve Bariyer Metotları

Kısıtlamalı optimizasyon problemlerinin çözümlerinde sıklıkla kullanılan yöntemlerden biride ceza (penalty) ve bariyer (barrier) metotlarıdır. Ceza ve Bariyer metotları kısıtlamalı optimizasyon problemlerine kısıtlamasız optimizasyon problemleri ile yaklaşan bir prosedür uygular. Bu yaklaşım, amaç fonksiyonuna kısıtlamaların ihlali durumunda yüksek bir maliyet getiren bir terimin eklenmesi ile gerçekleşmektedir.

Ceza metotlarında, (2.1)-(2.3) problemi için uygun çözüm bölgesi S'den bütün $\mathbb{R}^n$ genişletilir, fakat orijinal uygun çözüm bölgesi S'nin dışında yer alan noktalar için amaç fonksiyonuna yüksek bir maliyet getiren ceza terimi eklenir. Bu nedenle bu tarz metotlara dış (exterior) ceza metodu denir.

Bariyer metotlarında ise uygun çözüm bölgesi S'nin içinde aldığımız bir x_0 noktasının amaç fonksiyonuna uyguladığımız yüksek oranda ceza ile S'nin sınırlarına doğru oldukça yaklaşması sağlanır. Yani uygun çözüm bölgesinin sınırlarında bir bariyer oluşturur. Bu nedenle bu tarz metotlara iç (interior) ceza metodu denir.

Şimdi ceza metotları hakkında kısaca bilgi verelim.

Ceza fonksiyonu metodunda temel düşünce, (2.1)-(2.3) kısıtlamalı optimizasyon problemini p pozitif bir sabit ve $F:\mathbb{R}^n \to \mathbb{R}$ bir fonksiyon olmak üzere aşağıdaki formda kısıtlamasız optimizasyon problemine dönüştürmektir,

$$\text{minimum} \quad f(x) + pF(x). \tag{2.4}$$

Burada F ceza fonksiyonu ve p ceza parametresi olarak adlandırılır.

2.1.3.1 Tanım: Bir $F:\mathbb{R}^n \to \mathbb{R}$ fonksiyonun ceza fonksiyonu olarak adlandırılması için gerek şart,

a) F sürekli,

b) $\forall x \in S$ için $F(x) = 0$,

c) $\forall x \notin S$ için $F(x) > 0$

olmalıdır, [4].

Ceza fonksiyonu yukarıdaki şartları sağlayacak şekilde bir çok formda tanımlanabilir.

Eğer uygun çözüm bölgesi $S = \left\{ x : x \in \mathbb{R}^n, g_j(x) \le 0,\ j = 1, 2, ..., p \right\}$ biçiminde eşitsizlik kısıtlamalarına sahip ise ceza fonksiyonu

$$F(x)=\sum_{j=1}^{p}\max\left[0,g_j(x)\right]$$

biçiminde tanımlanabilir. Eğer uygun çözüm bölgesi S, $g_j(x)\geq 0,\ j=1,2,...,p$ biçiminde eşitsizlik kısıtlarına sahip ise ceza fonksiyonu

$$F(x)=\sum_{j=1}^{p}\max\left[0,-g_j(x)\right]=-\sum_{j=1}^{p}\min\left[0,g_j(x)\right]$$

tanımlanabilir. Eğer $S=\left\{x : x\in\mathbb{R}^n, h_i(x)=0, i=1,2,...,m\right\}$ eşitlik kısıtlarına sahip ise ceza fonksiyonu

$$F(x)=\frac{1}{2}\sum_{i=1}^{m}\left[h_i(x)\right]^2$$

veya daha genel olarak

$$F(x)=\frac{1}{\gamma}\sum_{i=1}^{m}\left|h_i(x)\right|^{\gamma},\ \gamma\geq 1$$

olarak tanımlanabilir. Burada cezanın ağırlığı pozitif ceza parametresi p ile kontrol edilir.

Burada p ceza parametresinin yeterince büyük seçilmesi (2.4) probleminin minimumunun F 'nin oldukça küçük olduğu bir bölgede olmasını sağlar. Yani artan bir p ceza parametresi (2.4) probleminin çözüm noktalarının S uygun çözüm bölgesine yaklaşmasını ve f 'yi minimize etmesini sağlar. Sonuç olarak $p\to\infty$ olduğunda ceza probleminin çözüm noktası kısıtlamalı optimizasyon probleminin bir çözümüne yakınsar.

Genel olarak (2.1)-(2.3) probleminin ceza metodu ile çözümü aşağıdaki prosedürü izlemektedir.

$\{p_k\}$, her $k=1,2,...$ için $p_k \geq 0$, $p_{k+1} > p_k$ ve $\lim\limits_{k\to\infty} p_k \to \infty$ olacak biçimde bir dizi olarak tanımlayalım. Buradan ceza fonksiyonu

$$q(p,x) = f(x) + pF(x)$$

yardımı ile $\forall k$ için

$$\text{minimum} \quad q(p_k, x)$$

ceza probleminin x_k çözümleri elde edilir, [4].

2.1.3.2 Önerme: Farz edilsin ki $\{p_k\}$ azalmayan bir dizi, yani $\forall k$ için $p_k \leq p_{k+1}$ olsun. Buradan her bir k için

a) $q\left(p_{k+1}, x_{k+1}\right) \geq q\left(p_k, x_k\right)$

b) $F\left(x_{k+1}\right) \leq F\left(x_k\right)$

c) $f\left(x_{k+1}\right) \geq f\left(x_k\right)$

d) $f\left(x^*\right) \geq q\left(p_k, x_k\right) \geq f\left(x_k\right)$

elde edilir, [2].

2.1.3.3 Teorem: Farz edilsin ki $S \neq 0$ için $f(x)$, $g(x)$, $h(x)$ ve $P(x)$ fonksiyonları sürekli olsunlar. $\{x_k\}_{k=1}^{\infty}$, (2.4) problemlerinin bir çözüm dizisi olsun ve $\{x_k\}_{k=1}^{\infty}$ dizisi kapalı bir küme içinde bulunsun. Buradan $\{x_k\}_{k=1}^{\infty}$ dizisinin herhangi bir limit noktası $\overline{x}$, (2.1)-(2.3) probleminin de çözümüdür, [43].

Ceza metodu ve Lagrange metodu arasındaki ilişkide kolaylıkla görülebilir. Örneğin eşitlik kısıtlamasına sahip (2.1), (2.2) problemi için

$$q(p,x) = f(x) + \frac{1}{2} p \sum_{i=1}^{m} \left[h_i(x)\right]^2$$

kuadratik ceza fonksiyonunu ve $x(p)$ minimumunu ele alınsın, buradan ceza fonksiyonunun gradyant vektörü

$$\nabla_x q(p, x(p)) = \nabla f(x(p)) + p\sum_{i=1}^{m} h_i(x(p))\nabla h_i(x(p)) = 0$$

olmak üzere $\lambda_i(p) = -ph_i(x(p))$ olarak tanımlarsak

$$\nabla f(x(p)) - \sum_{i=1}^{m} \lambda_i(p)\nabla h_i(x(p)) = 0$$

elde edilir. Sonuç olarak ceza parametresi $p \to \infty$ için $x(p) \to x^*$ optimal noktaya ve $\lambda(p) \to \lambda^*$, $\nabla f(x^*) = \nabla h(x^*)\lambda^*$ koşulu altında Lagrange çarpanına yaklaştığını kolaylıkla gösterebiliriz.

Ceza metodunda ki diğer bir kavramda tam (exact) ceza metodudur. Bu özelliğe sahip ceza metotlarında sonlu bir ceza parametresi ile ceza probleminin optimal çözümü, orijinal problemin optimal çözümüne karşılık getirilir. Yani bu özelliğe sahip ceza fonksiyonları ile orijinal problemin optimal çözümüne ulaşmak için sonsuz boyutlu ceza problemlerinin bir dizisini çözmek gerekli değildir. Bununla birlikte, bir yönden işlem boyutunu azaltmasına karşın diğer yandan da bu tür fonksiyonlar türevlenemezliğide beraberinde getirmektedir. Buda tam (exact) ceza problemlerinin çözümlerinde türev tabanlı optimizasyon metotlarının kullanılmasında hatalara neden olmaktadır.

Tam (exact) ceza fonksiyonlarından en yaygın kullanılanı mutlak değer ceza fonksiyonudur ve (2.1)-(2.3) problemi için

$$F(x) = \sum_{i=1}^{m} |h_i(x)| + \sum_{j=1}^{p} \max\left[0, g_j(x)\right]$$

biçiminde tanımlanır.

2.1.3.4 Teorem: Farz edilsin ki x^* noktası (2.1)-(2.3) problemi için ikinci mertebeden yeter şartı bir lokal minimum olmak için sağlasın. λ ve μ sabitleri de probleme karşılık gelen Lagrange çarpanları olsunlar. Buradan $p > \max\left\{|\lambda_i|, \mu_j : i = 1, 2, ..., m, j = 1, 2, ..., p\right\}$ için x^* aynı zamanda mutlak değer ceza fonksiyonu içinde bir lokal minimumdur, [4].

Ceza metodunda olduğu gibi benzer tanım ve teoremleri Bariyer metodu içinde yapılabilir. Eşitsizlik kısıtlamasına sahip (2.1), (2.3) problemi ele alınsın, bu problem için uygun çözüm bölgesi $S = \left\{x : g_j(x) \le 0,\ j = 1, 2, ..., p\right\}$ şeklinde tanımlanabilir.

2.1.3.5 Tanım: Bir $B : \mathbb{R}^n \to \mathbb{R}$ fonksiyonun bariyer fonksiyonu olarak adlandırılması için gerek şart,

a) B sürekli,

b) $\forall x \in S$ için $B(x) \ge 0$,

c) x noktası S 'nin sınırlarına yaklaştıkça $B(x) \to \infty$.

Yukarıdaki tanımı gerçekleyen bir çok bariyer fonksiyonu tanımlana bilir. Bunlardan en önemlileri

$$B(x) = -\sum_{j=1}^{p} \frac{1}{g_j(x)}$$

ters bariyer fonksiyonu ve

$$B(x) = -\sum_{j=1}^{p} \log\left(-g_j(x)\right)$$

logaritma bariyer fonksiyonudur.

Ceza metoduna benzer bir şekilde bariyer metodunda da $\{c_k\}$, her bir $k = 1, 2, ...$ için $c_k \geq 0$ ve $c_{k+1} > c_k$ özelliğinde bir dizi olmak üzere

$$r(c_k, x) = f(x) + \frac{1}{c_k} B(x)$$

fonksiyonu yardımı ile $\forall k$ için

$$\text{minimum} \quad r(c_k, x)$$

bariyer probleminin x_k çözümleri elde edilir, [4].

2.1.3.6 Teorem: Bariyer metodu ile üretilen bir $\{x_k\}$ dizisinin herhangi bir limit noktası (2.1), (2.3) probleminin de bir çözümüdür, [4].

2.2 Durum Denklemleri ve Adi Diferansiyel Denklemler

Bu bölümde modern kontrol teorinin gelişmesinde ve kontrol sistemlerinin matematiksel modellemelerinin yapılmasında önemli rol oynayan durum değişkenleri ve denklemleri ile dinamik sistemlerin transfer fonksiyonları hakkında bilgi verilecektir.

Kompleks dinamik sistemlerin analizinde sıklıkla kullanılan metotlardan birisi durum denklemleri analizidir, [44,45].

Genel olarak otonom (zamanla değişmeyen) dinamik sistemler $u_1(t), u_2(t), ..., u_r(t)$ r tane girdi, $y_1(t), y_2(t), ..., y_m(t)$ m tane çıktı ve $x_1(t), x_2(t), ..., x_n(t)$ n tane durum değişkeni olmak üzere

$$\begin{aligned}
\dot{x}_1(t) &= f_1\left(x_1, x_2, ..., x_n; u_1, u_2, ..., u_r\right) \\
\dot{x}_2(t) &= f_2\left(x_1, x_2, ..., x_n; u_1, u_2, ..., u_r\right) \\
&\vdots \\
\dot{x}_n(t) &= f_n\left(x_1, x_2, ..., x_n; u_1, u_2, ..., u_r\right)
\end{aligned} \tag{2.5}$$

durum diferansiyel denklemleri ile ve

$$\begin{aligned}
y_1(t) &= g_1\left(x_1, x_2, ..., x_n; u_1, u_2, ..., u_r\right) \\
y_2(t) &= g_2\left(x_1, x_2, ..., x_n; u_1, u_2, ..., u_r\right) \\
&\vdots \\
y_m(t) &= g_m\left(x_1, x_2, ..., x_n; u_1, u_2, ..., u_r\right)
\end{aligned} \tag{2.6}$$

çıktı denklemleri ile ifade edilir. Eğer $x(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_n(t) \end{bmatrix}$ durum vektörü, $u(t) = \begin{bmatrix} u_1(t) \\ u_2(t) \\ \vdots \\ u_r(t) \end{bmatrix}$ girdi veya kontrol vektörü ve $y(t) = \begin{bmatrix} y_1(t) \\ y_2(t) \\ \vdots \\ y_m(t) \end{bmatrix}$ çıktı vektörü ile gösterilirse (2.5)-(2.6) denklemi

$$\dot{x}(t) = f(x, u) \tag{2.7}$$

$$y(t) = g(x, u) \tag{2.8}$$

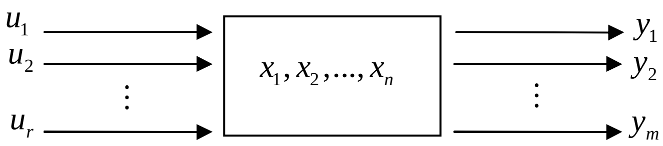
ile kullanılabilir.

Sistem doğrusal ise (2.7)-(2.8) durum denklemleri A durum, B girdi, C çıktı ve D direkt transmisyon sabit matrisleri olmak üzere

$$\dot{x}(t) = Ax(t) + Bu(t) \tag{2.9}$$

$$y(t) = Cx(t) + Du(t) \tag{2.10}$$

ile ifade edilir, [44,45].



Şekil 2.1 Dinamik sistem

(2.9)-(2.10) denklemleri ile ifade edilen dinamik sisteminin $x(t)$ çözümü (2.9) denkleminin her iki tarafının e^{-At} ile çarpılması ile görülebilir. İfade edilmek istenirse,

$$e^{-At}\dot{x}(t) - e^{-At}Ax(t) = e^{-At}Bu(t)$$

ve

$$\frac{d}{dt}\left(e^{-At}x(t)\right) = e^{-At}Bu(t)$$

olduğu görülür. Bu denklemin her iki tarafının 0 ile t arasında integrali alınırsa

$$e^{-A\tau}x(\tau)\Big|_{\tau=0}^{t} = \int_0^t e^{-A\tau}Bu(\tau)d\tau$$

$$e^{-At}x(t) - e^{0}x(0) = \int_0^t e^{-A\tau}Bu(\tau)d\tau$$

$$x(t) = e^{At}x(0) + \int_0^t e^{A(t-\tau)}Bu(\tau)d\tau \qquad (2.11)$$

çözümü elde edilir, [46].

Benzer olarak (2.9) denkleminin $x(t)$ çözümü laplace ve ters laplace yöntemleri kullanılarak da bulunabilir.

(2.9) ile gösterilen adi diferansiyel denkleminin $x(t)$ çözümünün varlığı ve tekliği ile ilgili teorem aşağıdaki şekildedir.

2.2.1 Teorem: $f\left(x,t\right)$ fonksiyonu her $t \geq t_0$ için parçalı sürekli ve her $x \in D \subset \mathbb{R}^n$ için yerel Lipschitz özelliğini sağlayan bir fonksiyon olsun. $W \subset D$ tıkız bir alt küme olsun ve

$$\dot{x} = f\left(x,t\right), \quad x\left(0\right) = x_0$$

bütün çözümleri tamamen W da kalsın. Bu durumda her $t \geq t_0$ için bir tek çözüm vardır, [47].

Durum diferansiyel denkleminin (2.11) çözümünde yer alan e^{At} üstel matrisi uzun yıllardır kontrol mühendisleri tarafından ayrı bir ilgiyle incelenmiştir. Bu matrise durum geçiş (state transition) matrisi denir ve $\Phi(t) = e^{At}$ ile gösterilir.

2.2.2 Tanım: $\Phi(t)$ matrisi aşağıdaki özelliklere sahiptir

a) $\Phi^{-1}(t) = e^{-At} = \Phi(-t)$

b) $\Phi(t)\Phi(-\tau) = e^{At}e^{-A\tau} = \Phi(t-\tau)$

Bu notasyon ile (2.11) çözümü

$$x(t) = \Phi(t)x(0) + \int_0^t \Phi(t-\tau)Bu(\tau)d\tau$$

ile gösterilir.

2.2.3 Uyarı: (2.9) Durum diferansiyel denkleminin direkt ve laplace çözümleri karşılaştırıldığında $\Phi(t) = L^{-1}\left\{\left[sI - A\right]^{-1}\right\}$ eşitliği görülür.

2.2.4 Teorem: Farz edilsin ki $n \times n$ boyutlu sabit A matrisinin $\lambda_1, \lambda_2, ..., \lambda_n$ biçiminde n tane farklı özdeğeri ve bunlara karşılık gelen $e_1, e_2, ..., e_n$ özvektörleri olsun. $T = (e_1, e_2, ..., e_n)$ ve $\Lambda = diag(\lambda_1, \lambda_2, ..., \lambda_n)$ biçiminde $n \times n$ boyutlu matrislerini tanımlayalım. Buradan T tekil olmayan bir matris olmak üzere $A = T\Lambda T^{-1}$ gösterilir, [48].

2.2.5 Teorem: Bir önceki teoremdeki özellikleri sağlayan A matrisini ele alalım. Buradan

a) $e^{At} = Te^{\wedge t}T^{-1}$

b) $e^{\wedge t} = diag(e^{\lambda_1 t}, e^{\lambda_2 t}, ..., e^{\lambda_n t})$

elde edilir.

Şimdi durum denklemleri ile ifade edilen bir dinamik sistemin transfer fonksiyonunun nasıl elde edileceği incelensin. Genel olarak (2.9)-(2.10) durum denklemleri ile ifade edilen bir dinamik sistemin transfer fonksiyonu $Y(s)$, $U(s)$ sırası ile $y(t)$ ve $u(t)$'nin laplace dönüşümleri olmak üzere

$$G(s) = \frac{Y(s)}{U(s)} \tag{2.12}$$

şeklinde gösterilir. (2.9)-(2.10) sisteminin $x(0) = 0$ başlangıç koşulu altında Laplace dönüşümü alınırsa

$$\begin{aligned} sX(s) - x(0) &= AX(s) + BU(s) \\ Y(s) &= CX(s) + DU(s) \end{aligned}$$

ve sonuç olarak

$$Y(s) = \left[C(sI - A)^{-1}B + D\right]U(s) \Rightarrow G(s) = C(sI - A)^{-1}B + D$$

transfer fonksiyonu elde edilir, [44,45].

2.3 Kararlılık

Bu bölümde herhangi bir dinamik sistemin denge noktasının kararlılığı ile ilgili tanım ve teoremlere yer verilecektir.

2.3.1 Denge Noktası

$D \subset \mathbb{R}^n$ bir bölge ve $f : D \to \mathbb{R}^n$ yerel Lipschitz özelliğini sağlayan bir fonksiyon olmak üzere

$$\dot{x} = f(x),\ x(0) = x_0 \qquad (2.13)$$

otonom (zamanla değişmeyen) sistemini göz önüne alalım. $f(x_e) = 0$ şartını sağlayacak biçimde bir $x_e \in D$ varsa x_e'ye (2.13) sisteminin denge noktası (sabit nokta) denir.

Şimdi x_e denge noktasının kararlılığı ile ilgili gerekli tanım ve teoremleri verelim.

2.3.1.1 Tanım: Her $\varepsilon > 0$ ve en az bir $x(0) = x_0$ başlangıç koşulu için

$$\|x_0 - x_e\| < \delta \text{ iken } \|x(t) - x_e\| < \varepsilon,\ \forall t \geq 0$$

olacak şekilde bir $\delta = \delta(\varepsilon) > 0$ bulunabiliyorsa x_e'ye kararlı denge noktası denir.

2.3.1.2 Tanım: x_e denge noktası kararlı ve en az bir $x(0) = x_0$ başlangıç koşulu için

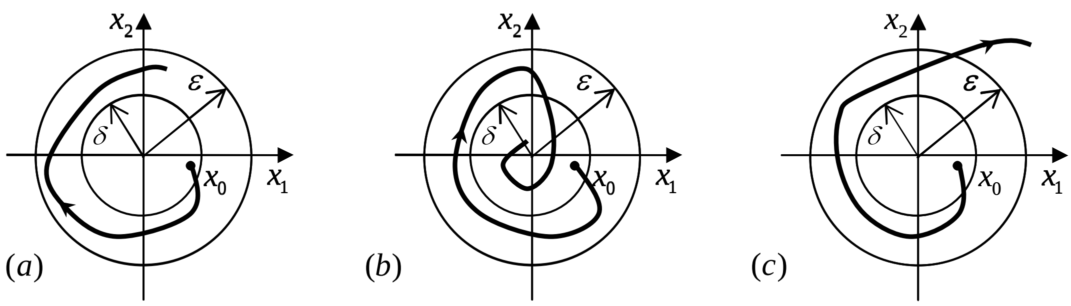
$$\|x_0 - x_e\| < \delta \text{ iken } \lim_{t\to\infty} \|x(t) - x_e\| = 0$$

olacak şekilde bir $\delta > 0$ varsa x_e 'ye asimptotik kararlı denge noktası denir.

İspatlarda kolaylık sağlaması açısından denge noktası genellikle orijin noktasında alınır. Denge noktasının orijinde olmadığı durumlarda ise bu nokta orijine ötelenir. Yani $x_e \neq 0$ olduğu durumlarda $z = x - x_e$ değişken dönüşümü yapılabilir. Buradan

$$\dot{z} = \dot{x} = f(z + x_e) = g(z)$$

elde edilir ve $g(0) = f(0 + x_e) = 0$ eşitliğinden orijinin denge noktası olduğu görülür.



Şekil 2.2 Kararlılık analizinin geometrik yorumu. (a) Kararlı denge noktası. (b) Asimptotik kararlı denge noktası. (c) Kararsız denge noktası.

2.3.2 Doğrusal Sistemler için Kararlılık Koşulu

Bu bölümde,

$$\dot{x} = Ax, \; x(0) = x_0 \tag{2.14}$$

doğrusal sistemi ve bu sistemin $x(t) = e^{At} x_0$ çözümü göz önüne alınsın. Bu tip sistemlerde kararlılık analizi genellikle A matrisinin özdeğerlerinin yerleri ile karakterize edilir.

2.3.2.1 Tanım:

a) $v \neq 0$ olmak üzere $Av = \lambda v$ denklemini sağlayan $\lambda \in \mathbb{C}$ değerine A matrisinin özdeğerleri denir.

b) λ özdeğerlerini $Av = \lambda v$ denkleminde yerine yazarak bulduğumuz v değerine A matrisinin özvektörleri denir.

c) v özvektörü

$$\left(A-\lambda I\right)v \neq 0,\ \left(A-\lambda I\right)^2 v \neq 0,\ldots,\ \left(A-\lambda I\right)^{m-1} v \neq 0,\ \left(A-\lambda I\right)^m v = 0\,,\ m>1$$

özelliğini sağlıyorsa v'ye A matrisinin $m.$ dereceden genelleştirilmiş özvektörü denir.

2.3.2.2 Teorem: $\pi(A)$**,** A matrisinin özdeğerlerinin kümesi olmak üzere her $\lambda \in \pi(A)$ için $\mathrm{Re}(\lambda) < \omega$ olsun. Bu durumda ω'ye bağlı öyle bir $M > 0$ vardır ki

$$\left\| e^{At} \right\| \leq Me^{\omega t},\quad t \geq 0\,.$$

İspat: x**,** A matrisinin $m.$ dereceden genelleştirilmiş özvektörü olmak üzere

$$e^{At}x = e^{\lambda t}e^{(A-\lambda I)t}x = e^{\lambda t}\sum_{k=0}^{m-1}\frac{\left(A-\lambda I\right)^k t^k}{k!}x$$

elde edilir. Her $\varepsilon > 0$ için herhangi bir t anında $t^k e^{-\varepsilon t}$ sınırlı olduğundan

$$\left\| e^{At}x \right\| \leq \overline{M}e^{\mathrm{Re}(\lambda)t}\left|\cos((\mathrm{Im}\,\lambda)t) + i\sin((\mathrm{Im}\,\lambda)t)\right| e^{\varepsilon t}\left\|x\right\| \leq Me^{\omega t}\left\|x\right\|$$

görülür.

2.3.2.3 Teorem: Her $\lambda \in \pi(A)$ için $\mathrm{Re}(\lambda) < 0$ ise orijin asimptotik kararlıdır.

İspat: Her $\lambda \in \pi(A)$ için $\mathrm{Re}(\lambda) < 0$ olsun. Bu durumda her $\lambda \in \pi(A)$ için öyle bir $\omega < 0$ vardır ki $\mathrm{Re}(\lambda) < \omega$ dir ve 2.3.2.2 Teorem gereği

$$\left\| e^{At} \right\| \le M e^{\omega t}, \quad t \ge 0$$

yazılabilir. Verilen bir $\varepsilon > 0$ için $\delta = \dfrac{\varepsilon}{M}$ seçilirse $\left\| x_0 \right\| < \delta$ için $\left\| e^{At} x_0 \right\| \le M \left\| x_0 \right\|$ $< M\delta = \varepsilon$ bulunur. Yani orijin noktası kararlı bir denge noktasıdır. Şimdi bir $\eta > 0$ seçilirse $\left\| x_0 \right\| < \eta$ için $\left\| e^{At} x_0 \right\| \le M e^{\omega t} \left\| x_0 \right\| < M \eta e^{\omega t}$ elde edilir ve limit alınırsa

$$\lim_{t \to \infty} \left\| e^{At} x_0 \right\| = 0$$

bulunur. Dolayısıyla orijin asimptotik kararlı bir denge noktasıdır, [49].

2.3.2.4 Tanım: Tüm özdeğerleri için $\mathrm{Re}(\lambda) < 0$ sağlayan A matrisine kararlı matris veya Hurwitz matrisi denir, [47].

2.3.3 Sınırlı Girdi-Sınırlı Çıktı Kararlılığı

2.3.3.1 Tanım: Her $t \ge 0$ için $\left\| u(t) \right\| \le u_m < \infty$ olacak şekilde bir $u_m > 0$ sabiti varsa u 'ya sınırlı girdi denir.

2.3.3.2 Tanım: Sınırlı her bir girdi için sınırlı çıktılar veren bir sisteme sınırlı girdi-sınırlı çıktı kararlıdır veya sadece girdi-çıktı kararlıdır denir. Bu kararlılık tanımı sıfır-durum cevabı için tanımlanır ve başlangıç anında duran bir sistem içi uygulanır.

Girdi-çıktı kararlılıkta sistemin çıktısı da değerlendirildiğinden bu tip kararlılığa dış (external) kararlılık denir.

2.3.3.3 Teorem: (2.9)-(2.10) doğrusal sisteminin transfer fonksiyonunun tüm kutupları sol yarı düzlemde ise sistem sınırlı-girdi-sınırlı çıktı kararlıdır denir.

İspat: $\dot{x} = Ax + Bu$ **,** $y = Cx + Du$ doğrusal sisteminin transfer fonksiyonu

$$G(s) = C(sI - A)^{-1}B + D$$

şeklindedir. Transfer fonksiyonunda $\left(sI - A\right)^{-1} = \dfrac{ek\left(sI - A\right)}{\det\left(sI - A\right)}$ yazarak elde edilen

$$G(s) = \frac{1}{\det\left(sI - A\right)} C\left[ek\left(sI - A\right)\right]B + D \tag{2.15}$$

eşitliğinden $G(s)$ transfer fonksiyonunun her bir kutbunun $\left(sI - A\right)$'nın bir özdeğerine karşılık geldiği görülür. Dolayısıyla eğer A matrisinin her bir özdeğeri negatif düzlemde ise sistem sınırlı girdi-sınırlı çıktı kararlıdır. Diğer yandan (2.15) eşitliğindeki mümkün bazı sadeleşmeler yüzünden A matrisinin her bir özdeğeri transfer fonksiyonunun bir kutbuna karşılık gelmez. Yani A matrisinin özdeğerlerinden bazıları sağ yarı düzlemde olsa bile sistem kararlı olabilir, [50].

2.3.3.4 Teorem: (2.9)-(2.10) doğrusal sistemi minimal gerçekleştirmeye sahip ise bu doğrusal sistem için sınırlı-girdi-sınırlı çıktı kararlılığı denge noktasında asimptotik kararlılık koşulunu sağlar, [51].

Genel olarak doğrusal veya doğrusal olmayan zamanla değişmeyen sistemlerin denge noktalarındaki kararlılık analizleri Lyapunov kararlılık teorisi ile olmaktadır. Şimdi bazı tanım ve teoremlere yer verelim.

2.3.4 Doğrusal Olmayan Sistemler için Lyapunov Kararlılık Teorisi

Doğrusal olmayan sistemlerin Lyapunov kararlılık analizini direkt ve direkt olmayan yöntemler olarak ikiye ayırabiliriz, [52].

(2.13) denklemi ile belirttiğimiz doğrusal olmayan sistemi ele alalım. $x = 0$ orijin noktası D bölgesi içinde yer alsın ve bu sistemin bir denge noktası olsun.

Ortalama değer teoreminden, z_i, x ile orijini birleştiren doğru parçası üzerindeki bir nokta olmak üzere

$$f_i(x) = f_i(0) + \frac{\partial f_i}{\partial x}(z_i)x$$

elde edilir. Yukarıdaki eşitlik D bölgesi içinde yer alan x ile orijini birleştiren doğru parçası üzerindeki her $x \in D$ için sağlanır. Ayrıca $f(0) = 0$ olduğundan

$$f_i(x) = \frac{\partial f_i}{\partial x}(z_i)x = \frac{\partial f_i}{\partial x}(0)x + \left[\frac{\partial f_i}{\partial x}(z_i) - \frac{\partial f_i}{\partial x}(0)\right]x$$

yazılır. $A = \frac{\partial f_i}{\partial x}(0)$ ve $g_i(x) = \left[\frac{\partial f_i}{\partial x}(z_i) - \frac{\partial f_i}{\partial x}(0)\right]x$ ile ifade edilirse

$$f(x) = Ax + g(x)$$

elde edilir. Burada $g_i(x)$ 'in

$$|g_i(x)| \le \left\|\frac{\partial f_i}{\partial x}(z_i) - \frac{\partial f_i}{\partial x}(0)\right\| \|x\|$$

özelliğini sağlayan bir fonksiyon ve $\frac{\partial f}{\partial x}$ sürekliliğinden $\|x\| \to 0$ için $\frac{\|g(x)\|}{\|x\|} \to 0$ olduğu görülür.

Sonuç olarak, doğrusal olmayan (2.13) sistemine orijin noktasının küçük bir komşuluğunda

$$\dot{x} = Ax\,, \ A = \frac{\partial f}{\partial x}(0)$$

olmak üzere orijin noktasında lineerleştirilmiş sistem ile yaklaşabiliriz.

2.3.4.1 Teorem: $x=0$ orijin noktası doğrusal olmaya (2.13) sisteminin bir denge noktası ve D orijin noktansın bir komşuluğu olsun. $A = \left.\frac{\partial f}{\partial x}(x)\right|_{x=0}$ olmak üzere

a) A matrisinin tüm özdeğerleri için $\operatorname{Re}\lambda_i < 0$ ise orijin noktası asimptotik kararlıdır.

b) A matrisinin bir veya birden fazla özdeğerleri için $\operatorname{Re}\lambda_i > 0$ ise orijin noktası kararsızdır denir, [47].

Bu teoreme Lyapunov'un birinci metodu veya direkt olmayan metodu denir.

Yeniden (2.13) otonom sistemini ele alalım ve $D \subset \mathbb{R}^n$ bölgesinin orijinin bir komşuluğu olduğunu varsayalım. $V : D \to \mathbb{R}$ tanımlı sürekli diferansiyellenebilen bir fonksiyon olsun. V 'nin (2.13) sisteminin yörüngeleri boyunca türevi aşağıdaki şekilde verilir:

$$\dot{V}(x) = \langle gradV, f \rangle = \sum_{i=1}^{n} \frac{\partial V}{\partial x_i} f_i(x) \, .$$

Sistemin yörüngeleri boyunca türevi sistemin denklemine bağlıdır. $\varphi(t, x_0)$ sistemin $t=0$ anındaki x_0 başlangıç koşuluna karşılık gelen çözümü ise

$$\dot{V}(x) = \left.\frac{d}{dt} V\left(\varphi(t, x_0)\right)\right|_{t=0}$$

şeklindedir. Dolayısıyla eğer $\dot{V}(x)$ negatif ise V , (2.13) sisteminin çözümü boyunca azalacaktır.

2.3.4.2 Teorem: $x = 0$ orijin noktası (2.13) sistemi için bir denge noktası olsun ve $D \subset \mathbb{R}^n$ bölgesi noktasını kapsasın. $V : D \to \mathbb{R}$ tanımlı sürekli diferansiyellenebilen bir fonksiyon olmak üzere eğer

$$V(0) = 0 \text{ ve } x \in D - \{0\} \text{ için } V(x) > 0 \tag{2.16}$$

$$\text{her } x \in D \text{ için } \dot{V}(x) \leq 0 \tag{2.17}$$

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

$$\text{her } x \in D - \{0\} \text{ için } \dot{V}(x) < 0 \tag{2.18}$$

ise sistem $x = 0$ noktasında asimptotik kararlıdır.

İspat: Verilen bir $\varepsilon > 0$ için bir $r \in (0, \varepsilon]$ seçilsin ki

$$B_r = \left\{ x \in \mathbb{R}^n : \|x\| \leq r \right\} \subset D .$$

$\alpha = \min\limits_{\|x\|=r} V(x)$ olarak alalım. (2.16) dan $\alpha > 0$ dır. $\beta \in (0, \alpha)$ olsun ve

$$\Omega_\beta = \left\{ x \in B_r : V(x) \leq \beta \right\} \text{ olsun.}$$

Ω_β ’nın ve B_r ’nin içinde olmadığı kabul edilsin. Bu durumda bir $p \in \Omega_\beta$ noktası vardır ki bu nokta B_r ’nin sınırındadır. Bu noktada $V(p) \geq \alpha > \beta$ olur fakat bütün $x \in \mathbb{R}^n$ noktaları için $V(x) \leq \beta$ dır. Bu bir çelişkidir, dolayısıyla Ω_β kesinlikle B_r ’nın içindedir. $t = 0$ noktasında Ω_β kümesinde başlayan her yörünge her $t \geq 0$ anlarında Ω_β ’da kalır. Bu durumda, (2.17) özelliğini kullanarak aşağıdaki gibi ifade edilebilir:

$$\dot{V}(x(t)) \leq 0 \Rightarrow V(x(t)) \leq V(x(0)) \leq \beta , \ \forall t \geq 0 .$$

Ω_β tıkız (kompakt) bir küme olduğundan ($\Omega_\beta \subset \mathbb{R}^n$ tanımı gereği kapalı ve B_r ile sınırlı olduğunda) 2.2.1 Teorem'den (2.13) sisteminin tüm $x(0) \in \Omega_\beta$ başlangıç koşulları ve her $t \geq 0$ için tek bir çözümü vardır. $V(x)$ sürekli ve $V(0) = 0$ olduğundan öyle bir $\delta > 0$ vardır ki

$\|x\| \leq \delta \Rightarrow V(x) < \beta$ dır.

Buradan

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

dır ve

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

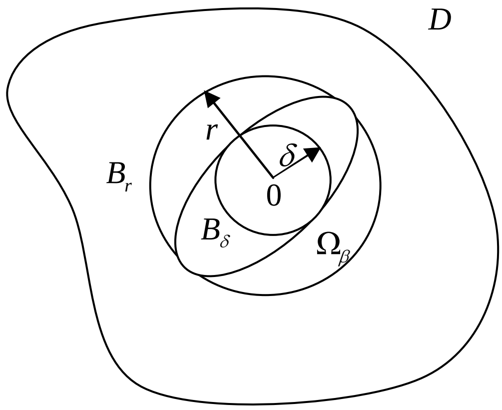
bulunur. Dolayısıyla

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

elde edilir ve bu ifade $x = 0$ denge noktasının kararlı olduğunu gösterir. Şimdi (2.18) ifadesinin sağlandığını kabul edelim. Asimptotik kararlılığı göstermek için her $\alpha > 0$ ve $t > T$ için $\|x(t)\| < \alpha$ olacak şekilde bir $T > 0$ ın varlığını ve $t \to \infty$ için $x(t) \to 0$ olduğu gösterilmelidir. Önceki kısmın ispatından, her $\alpha > 0$ için $\Omega_b \subset B_\alpha$ olacak şekilde bir $b > 0$ olduğu biliniyor. Dolayısıyla $t \to \infty$ için $V(x(t)) \to 0$ olduğunu göstermek yeterlidir. $V(x(t))$ monoton azalan ve alttan sıfır noktası ile sınırlı olduğundan

$t \to \infty$ için $V(x(t)) \to c \geq 0$.

$c = 0$ olduğunu göstermek için $c > 0$ olduğunu varsayalım. $V(x)$ sürekli olduğundan $B_d \subset \Omega_c$ olacak şekilde bir $d > 0$ vardır. $V(x(t)) \to c > 0$ limiti her



Şekil 2.3 2.3.4.2 Teorem'inin ispatında geçen kümelerin geometrik yorumu

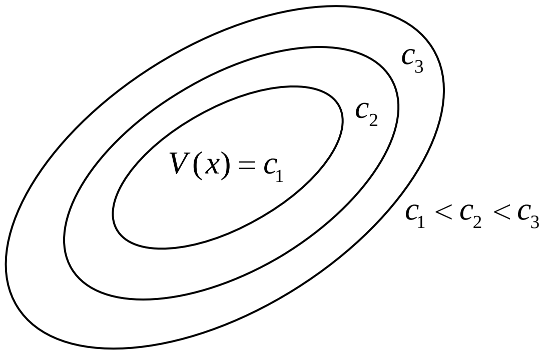
$t \geq 0$ için $x(t)$ yörüngesinin B_d yuvarının dışında kaldığını gösterir. $-\gamma = \max_{d \leq \|x\| \leq r} \dot{V}(x)$ olarak tanımlansın (böyle bir γ vardır çünkü $\dot{V}(x)$ fonksiyonu $d \leq \|x\| \leq r$ tıkız kümesi üzerinde maksimum bir noktaya sahiptir). (2.18)'den $-\gamma < 0$. Buradan

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

eşitsizliğinden sağ taraf negatif kalacağından $c > 0$ varsayımı ile çelişir, [47].

Bu teorem Lyapunov ikinci metodu veya direkt metodu olarak adlandırılır. Lyapunov kararlılık teoremi, kararlılık, asimptotik kararlılık ve kararsızlık ile ilgili olarak sadece yeter şartları ortaya koymaktadır. Yani bir sistem için (2.16)-(2.18) özellikleri sağlayan bir Lyapunov fonksiyonu bulamamış olmamız o sistemin kararsız olacağı anlamına gelmez. Ters teoremler hakkındaki literatür [47,s.148] bulunabilir.

2.3.4.3 Tanım: (2.16)-(2.18) koşullarını gerçekleyen sürekli diferansiyellenebilen $V(x)$ fonksiyonuna Lyapunov fonksiyonu denir. $c > 0$ için tanımlana $V(x) = c$ yüzeyine ise Lyapunov yüzeyi denir.



Şekil 2.4 Lyapunov fonksiyonunun seviye yüzeyleri

Şekil 2.4'deki Lyapunov yüzeyleri kullanılarak bir önceki ispat daha anlaşılır bir hale gelir. $\dot{V} \leq 0$ koşulu herhangi bir yörüngenin $V(x) = c$ yüzeyinden geçmesi durumunda bu yörüngenin $\Omega_c = \left\{x \in \mathbb{R}^n : V(x) \leq c\right\}$ kümesinin içinde hareket edeceğine ve bir daha kesinlikle dışarı çıkmayacağına işaret eder. $\dot{V} < 0$ olduğunda yörünge bir Lyapunov yüzeyinden daha içte olan bir Lyapunov yüzeyine ilerler. c azaldığında $V(x) = c$ Lyapunov yüzeyi orijine doğru büzülür. Bu ise yörüngenin zamanla orijine yakınsadığını gösterir, [47].

Lyapunov fonksiyonunun seçimi için sistematik bir kural yoktur. Bazı durumlarda özellikle mekanik ve elektrik sistemlerinde enerji fonksiyonları doğal birer Lyapunov fonksiyonu olarak alınabilir.

Lyapunov kararlılıkta sistemin çıktısı hesaba katılmadığı için bu tip kararlılığa iç (internal) kararlılık denir.

2.3.5 Doğrusal Sistemler için Lyapunov Kararlılık Teorisi

$\dot{x} = Ax$, $x(0) = 0$ doğrusal zamanla değişmeyen (2.14) sistemi ele alınsın; P, $n \times n$ boyutlu reel simetrik bir matris ($P = P^T \in \mathbb{R}^{n\times n}$) olmak üzere

$$V(x) = \langle x, Px \rangle = x^T Px$$

kuadratik formu ele alındığında. $V(x(t))$ fonksiyonunun (2.14) sisteminin çözümü üzerindeki zamana göre türevi

$$\frac{dV(x(t))}{dt} = \dot{V}(x) = \dot{x}^T(t)Px(t) + x^T(t)P\dot{x}(t)$$

elde edilir. Bu denklemde $\dot{x}$ ifadesi yazılırsa

$$\begin{aligned} \dot{V}(x) &= x^T A^T Px + x^T PAx \\ &= x^T \left(A^T P + PA \right) \mathrm{x} \end{aligned}$$

olduğu görülür. Burada $-Q = A^T P + PA$ olarak alınırsa

$$\dot{V}(x) = -x^T Qx$$

olur. Dikkat edilirse V fonksiyonu (2.14) sistemi için bir Lyapunov fonksiyonudur. Ayrıca Q matrisinin de tanım gereği simetrik bir matris olduğu görülebilir. $-Q = A^T P + PA$ denklemine de Lyapunov matris denklemi veya Lyapunov denklemi denir.

Şimdi $\dot{x} = Ax$ sisteminin veya buna denk olarak A matrisinin kararlılığı ile ilgili teoremleri verelim.

2.3.5.1 Teorem: A matrisinin tüm özdeğerlerinin negatif reel kısma sahip ise her pozitif tanımlı simetrik matris Q için $-Q = A^T P + PA$ Lyapunov denkleminin tek bir çözümü vardır ve bu çözüm

$$P = \int_0^\infty e^{A^T t} Q e^{At} dt \tag{2.19}$$

biçiminde ifade edilir, [50].

2.3.5.2 Teorem: (2.14) doğrusal sisteminin asimptotik kararlı olması için gerek ve yeter şart her pozitif tanımlı simetrik matris Q için $-Q = A^T P + PA$ Lyapunov denkleminin tek bir pozitif tanımlı simetrik P matris çözümünün olmasıdır.

İspat: Verilen sistemin asimptotik kararlı olması için A matrisinin tüm özdeğerleri için $\operatorname{Re}\lambda_i < 0$ olmalıdır. Böylece P matrisi için (2.19) denklemini ele alınabilir. Burada gösterilmesi gereken P matrisinin pozitif tanımlı ve Lyapunov denklemin tek bir çözümü olduğudur.

İlk olarak (2.19) denklemini Lyapunov denkleminde yerine yazar ve $t \to \infty$ için $e^{At} \to 0$ özelliği kullanılarak

$$\begin{aligned} A^T P + PA &= \int_0^\infty A^T e^{A^T t} Q e^{At} dt + \int_0^\infty e^{A^T t} Q e^{At} A dt \\ &= \int_0^\infty \frac{d}{dt}\left(e^{A^T t} Q e^{At}\right) dt \\ &= e^{A^T t} Q e^{At}\Big|_{t=0}^{\infty} \\ &= -Q \end{aligned}$$

elde edilir. Bu bize (2.19) denklemi ile gösterilen P matrisinin Lyapunov denkleminin bir çözümü olduğunu verir. Bu çözümün tek bir çözüm olduğunu göstermek için

$$A^T P_1 + P_1 A = -Q \tag{2.20}$$

$$A^T P_2 + P_2 A = -Q \tag{2.21}$$

olmak üzere P_1 ve P_2 çözümlerini ele alındığında, (2.21) denklemini (2.20) denkleminden çıkarılırsa

$$A^T (P_1 - P_2) + (P_1 - P_2) A = 0$$

bulunur. Ayrıca

$$e^{A^T t} A^T (P_1 - P_2) e^{At} + e^{A^T t} (P_1 - P_2) A e^{At} = \frac{d}{dt} e^{A^T t} (P_1 - P_2) e^{At} = 0 \tag{2.22}$$

sağlandığı görülebilir. (2.22) denkleminin $t = 0$ dan ∞ integrali alınırsa

$$e^{A^T t} (P_1 - P_2) e^{At} \Big|_{t=0}^{\infty} = P_1 - P_2 = 0$$

olur. Buda (2.19) ile belirtilen P çözümünün tek olduğunu verir. Ayrıca Q matrisinin teorem gereği simetrik bir matris olması gerektiğinden P matrisi de simetriktir.

P matrisinin pozitif tanımlılığı ise sıfırdan farklı her $x \in \mathbb{R}^n$ için Q matrisinin pozitif tanımlı ve her t için e^{At} singüler olmamasından

$$x^T P x = \int_0^\infty x^T e^{A^T t} Q e^{At} x dt > 0$$

ile kolaylıkla görülür.

Tersine Lyapunov denklemini sağlayan pozitif tanımlı simetrik P ve Q matrisleri var ise (2.14) doğrusal sistemi asimptotik kararlıdır. Bunu gösterebilmek için A matrisinin bir λ özdeğerine, $Av = \lambda v$ şartını sağlayarak, karşılık gelen

$v \neq 0$ özvektörünü alalım. Buradan $v^{T*}AT = v^{T*}\lambda^{T*}$ olarak yazılabilir. Lyapunov denkleminin her iki tarafını v^{T*} ve v ile karşılıklı olarak çarparsak

$$-v^{T*}Qv = v^{T*}A^T Pv + v^{T*}PAv = \left(\lambda^{T*} + \lambda\right)v^{T*}Pv = 2\,\mathrm{Re}\left(\lambda\right)v^{T*}Pv$$

elde edilir. P ve Q matrislerinin pozitif tanımlılığından $\mathrm{Re}\left(\lambda\right) < 0$ ve dolayısıyla sistemin asimptotik kararlı olduğu görülür, [46].

Yukarıdaki teorem, (2.14) doğrusal sisteminin kararlılığının aynı zamanda A matrisinin kararlılığı ile bire bir ilişkili olduğundan aşağıdaki şekilde de ifade edilebilir.

2.3.5.3 Teorem: A matrisinin kararlı bir matris yani A matrisinin tüm özdeğerlerinin $\mathrm{Re}\left(\lambda\right) < 0$ şartını sağlaması için gerek ve yeter şart her pozitif tanımlı simetrik matris Q için $-Q = A^T P + PA$ Lyapunov denklemini sağlayan bir P pozitif tanımlı simetrik matrisinin olmasıdır. Dahası A matrisi kararlı bir matris ise P Lyapunov denkleminin tek bir çözümüdür, [47].

Sonuç olarak dikkat edilirse doğrusal zamanla değişmeyen sistemlerin kararlılık analizinde Lyapunov teoremi doğrusal olmayan sistemlerin aksine sistemin kararlılığı ile ilgili olarak hem gerek hem de yeter şartları ortaya koyabilmektedir. Burada A matrisinin kararlı olup olmadığını görmek için yapılması gereken tek işlem herhangi bir pozitif tanımlı simetrik Q matrisi için Lyapunov denkleminin yine pozitif tanımlı simetrik bir P matris çözümünün olup olmadığına bakmaktır. Pratikte $Q = I$ eşitliğinden yararlanılır.

3. KISITLAMASIZ OPTİMİZASYON PROBLEMLERİ İÇİN DURUM DENKLEMLERİ YAKLAŞIMI

Bu bölümde kuadratik kısıtlamasız optimizasyon problemlerinin minimumlarının araştırılmasında durum denklemleri yaklaşımı ortaya konulacaktır.

3.1 Ana Sonuçlar

İkinci bölümde genel olarak (2.1) denklemi ile tanımladığımız $f:\mathbb{R}^n \to \mathbb{R}$ biçiminde tanımlanan kuadratik bir fonksiyon olmak üzere

$$\underset{x\in\mathbb{R}^n}{\text{minimum}} \quad f(x) \tag{3.1}$$

kısıtlamasız optimizasyon problemi ele alındığında, bu problemin mutlak optimal çözümünün x^* olduğu varsayılsın.

Kısıtlamasız optimizasyon probleminin çözümü için ∇f, f amaç fonksiyonunun gradyant vektörü olmak üzere

$$\begin{aligned} \frac{dx}{dt} &= \nabla f\left(x(t)\right) \\ x(0) &= x_0 \end{aligned} \tag{3.2}$$

adi diferansiyel denklem sistemi alındığında amaç (3.1) ile tanımladığımız kuadratik kısıtlamasız optimizasyon probleminin mutlak minimumunu (3.2) ile gösterdiğimiz diferansiyel denklem sisteminin $x(t)$ çözüm yörüngesi boyunca yaklaşarak elde etmeye çalışmaktır.

Ele alınan problem kuadratik bir yapıda olduğu için (3.2) diferansiyel denklem sistemi

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t) \\ y(t) &= Cx(t) \end{aligned} \tag{3.3}$$

matris formunda yazılabilir. İkinci bölümden de hatırlanabileceği gibi (3.3) denklem sistemine durum denklemleri adı verilir. Burada, $u(t)$ sistemin girdisi olmak üzere

$$u(t) = \begin{cases} 1; & t \geq 0 \\ 0; & t < 0 \end{cases}$$

ile tanımlansın. Aynı zamanda A $n \times n$ boyutlu, B $n \times 1$ boyutlu ve $C = I_{n \times n}$ matrisler olarak alındığında, (3.3) sisteminin transfer fonksiyonu ise

$$G(s) = C\left(sI - A\right)^{-1} B$$

ile gösterilir.

3.1.1 Uyarı: (3.3) sisteminde yer alan A matrisinin tüm özdeğerlerinin sol yada sağ yarı düzlemde olduğu kabul edilmiştir.

3.1.2 Tanım: (3.3) sisteminde yer alan A matrisinin tüm özdeğerleri λ ile göstermek üzere

(a) $\operatorname{Re}\lambda(A) < 0 \Rightarrow \tilde{A} := A$ ve $\tilde{B} := B$

(b) $\operatorname{Re}\lambda(A) > 0 \Rightarrow \tilde{A} := -A$ ve $\tilde{B} := -B$

ile tanımlayalım.

Yukarıdaki tanım ile (3.3) sisteminin her zaman kararlı bir sistem olarak kalmasını garanti altına alınır.

(3.3) durum denklem sistemini x^* sistemin denge noktası olmak üzere $z = x - x^*$ dönüşümü altında orijin noktasında denge noktasına sahip

$$\dot{z} = Az \tag{3.4}$$

doğrusal homojen diferansiyel denklem sistemine dönüştürülebilir.

Şimdi (3.3) ve (3.4) sistemlerinin denge noktaları arasındaki ilişkiyi ortaya koyalım.

3.1.3 Teorem: Bir x^* noktası (3.3) durum denklem sisteminin bir denge noktası olsun. x^* noktasının (3.3) sisteminin asimptotik kararlı bir denge noktası olması için gerek ve yeter şart $z = x - x^*$ dönüşümü altında orijin noktasının (3.4) sisteminin asimptotik kararlı bir denge noktası olmasıdır.

İspat: x^* noktasının (3.3) sistemi için asimptotik kararlı bir denge noktası olduğunu varsayılsın, bu durumda 2.3.1.2 Tanım gereği x^* denge noktası kararlıdır ve her $x(0) = x_0$ başlangıç koşulu için

$$\|x_0 - x^*\| < \delta \ \text{ iken } \lim_{t\to\infty} \|x(t) - x^*\| = 0$$

olacak şekilde bir $\delta > 0$ vardır.

$z = x - x^*$ dönüşümü asimptotik kararlılık tanımında yerine yazılırsa her $z(0) = z_0$ başlangıç koşulu için

$$\|z_0\| < \delta \ \text{ iken } \lim_{t\to\infty} \|z(t)\| = 0$$

olacak şekilde bir $\delta > 0$ vardır. Dolayısıyla $z = 0$ noktası (3.4) sistemi için asimptotik kararlı bir noktadır.

Tersine, (3.4) sistemi $z = 0$ noktasında asimptotik kararlı bir noktasına sahip olsun. Orijin noktasındaki asimptotik kararlılık tanımı ve $z = x - x^*$ dönüşümü kullanılarak benzer şekilde ispat edilir. □

3.1.4 Teorem: x^* noktasının (3.3) sisteminin için asimptotik kararlı bir denge noktası olması için gerek ve yeter şart (3.1) kuadratik kısıtlamasız optimizasyon probleminin mutlak bir minimumu olmasıdır.

İspat: x^* noktasının (3.3) dolayısıyla (3.2) sistemi için asimptotik kararlı bir denge noktası olduğunu varsayalım. Bu durumda, 2.3.1.2 Tanım gereği x^* denge noktası kararlı ve her $x(0) = x_0$ başlangıç koşulu için

$$\|x_0 - x^*\| < \delta \text{ iken } \lim_{t\to\infty} \|x(t) - x^*\| = 0$$

olacak şekilde bir $\delta > 0$ vardır. Bu kabulden

$$\lim_{t\to\infty} \|x(t) - x^*\| = 0 \Rightarrow \lim_{t\to\infty} x(t) = x^*$$

olarak yazılabileceğinden (3.2) ve (3.3) sistemlerinin tanımları gereği x^* noktası (3.1) optimizasyon problemi için bir mutlak minimumdur.

Tersine; x^* noktası (3.1) problemi için mutlak minimum noktası olsun. (3.2) diferansiyel denklem sistemi yardımıyla (3.3) durum denklem sistemi oluşturulabilir. x^* noktasının (3.1) probleminin mutlak minimumu olmasından dolayı 2.1.1.3 Önerme'de yer alan birinci mertebeden optimallik şartı sağlanır. Dolayısıyla x^* noktası (3.3) durum denklem sisteminin bir denge noktasıdır. Şimdi x^* noktasının (3.3) sistemi için asimptotik kararlı bir denge noktası olduğun gösterilmelidir.

Bu amaçla doğrusal sistemler için Lyapunov kararlılık teoremini kullanılır. $z = x - x^*$ dönüşümü ile (3.3) sistemi orijin noktasında denge noktası olan (3.4) sistemine dönüştürülsün. 3.1.2 Tanım gereği $\tilde{A}$ matrisinin tüm özdeğerleri sol yarı düzlemde olacağı için 2.3.5.2 Teorem ve 2.3.5.3 Teorem gereği (3.4) sistemi orijin noktasında asimptotik kararlıdır. Dolayısıyla 3.3 Teorem gereği x^* noktası (3.1.3) sistemi için asimptotik kararlıdır.□

3.2 Test Problemleri ve Sonuçlar

Bu bölümde durum denklemleri yaklaşımı kullanılarak bazı kısıtlamasız optimizasyon test problemlerinin optimal çözümleri araştırılacaktır. Bu yaklaşım sonucunda MATLAB programı ile elde edilen sonuçlar Quasi-Newton metodu ile karşılaştırılacaktır.

3.2.1 Problem: Aşağıdaki kuadratik kısıtlamasız optmizasyon problemini ele alınsın, [5].

$$\underset{x\in\mathbb{R}^3}{\text{minimum}} \quad \left(2x_1^{\,2}+2x_2^{\,2}+x_3^{\,2}+2x_1x_2+2x_2x_3-8x_1-6x_2-4x_3+9\right) \tag{3.5}$$

Problemin teorik optimal çözümü $\left(x_1^*, x_2^*, x_3^*\right)=\left(3,-2,4\right)^T$ 'dır. (3.3) Durum denklem modelini elde etmek için (3.2) denklemini ve 3.1.2 Tanım'ını ele alalım. Buradan diferansiyel denklem sistemi

$$\begin{aligned} \begin{pmatrix}\dot{x}_1\\ \dot{x}_2\\ \dot{x}_3\end{pmatrix} &= \begin{pmatrix}-4 & -2 & 0\\ -2 & -4 & -2\\ 0 & -2 & -2\end{pmatrix}\begin{pmatrix}x_1\\ x_2\\ x_3\end{pmatrix}+\begin{pmatrix}8\\ 6\\ 4\end{pmatrix} \\ y &= \begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}\begin{pmatrix}x_1\\ x_2\\ x_3\end{pmatrix} \end{aligned} \tag{3.6}$$

bulunur. Dikkat edilirse A matrisinin tüm özdeğerleri sol yarı düzlemde olduğu için kararlı bir matristir. Buradan (3.6) sistemi yardımı ile transfer fonksiyonu

$$G(s) = \frac{1}{s^3 + 10s^2 + 24s + 8}\begin{pmatrix} 8s^2 + 36s + 24 \\ 6s^2 + 12s - 16 \\ 4s^2 + 20s + 32 \end{pmatrix}$$

elde edilir. Transfer fonksiyonunun birim basamak yanıtı incelendiğinde (3.6) sisteminin denge noktasının (3.5) probleminin optimal çözümüne karşılık geldiği görülür.

Şimdi elde edilen denge noktasının asimptotik kararlılığını incelensin. Bu amaçla (3.6) sistemini $z = x - x^*$ dönüşümü ile orijin noktasında denge noktası olan

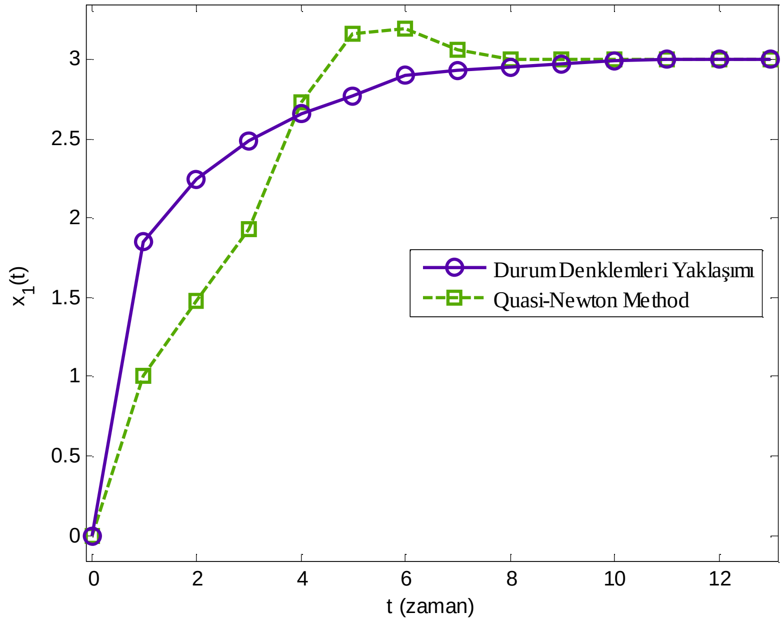
$$\begin{pmatrix} \dot{z}_1 \\ \dot{z}_2 \\ \dot{z}_3 \end{pmatrix} = \begin{pmatrix} -4 & -2 & 0 \\ -2 & -4 & -2 \\ 0 & -2 & -2 \end{pmatrix}\begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix} \tag{3.7}$$

sistemine dönüştürülür. Bu doğrusal sistemin kararlılık analizi için Lyapunov kararlılık teoremi gereği $Q = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, $A = \begin{pmatrix} -4 & -2 & 0 \\ -2 & -4 & -2 \\ 0 & -2 & -2 \end{pmatrix}$ ve $P = \begin{pmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{pmatrix}$ matrisleri yardımı ile $A^T P + PA = -Q$ Lyapunov matris denklemini ele alındığında, eğer bu denklemin bir tek P pozitif tanımlı simetrik matris çözümü olduğu gösterilirse (3.7) sisteminin orijinde asimptotik kararlı olduğu söylenebilir.

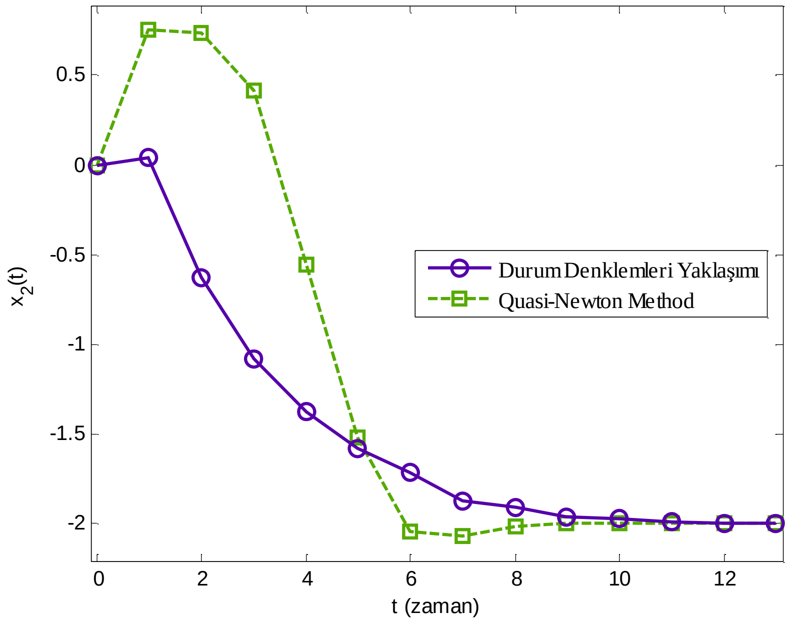
Lyapunov matris denklemi ile oluşan sistemin analitik çözümü yapılırsa $p_{11} = 0.2120$, $p_{12} = p_{21} = -0.1741$, $p_{22} = 0.3038$, $p_{13} = p_{31} = -0.3671$ ve $p_{33} = 0.6171$ olmak üzere Lyapunov matris denkleminin bir tek P pozitif tanımlı simetrik matris çözümü olduğunu görebiliriz. Sonuç olarak, 3.1.3 Teorem gereği x^* noktası (3.6) dinamik sisteminin asimptotik kararlı bir noktasıdır.

Aşağıdaki şekillerde durum denklemleri yaklaşımı ile elde ettiğimiz çözümlerin Quasi-Newton algoritmasının çözümleri ile karşılaştırılması verilmiştir.

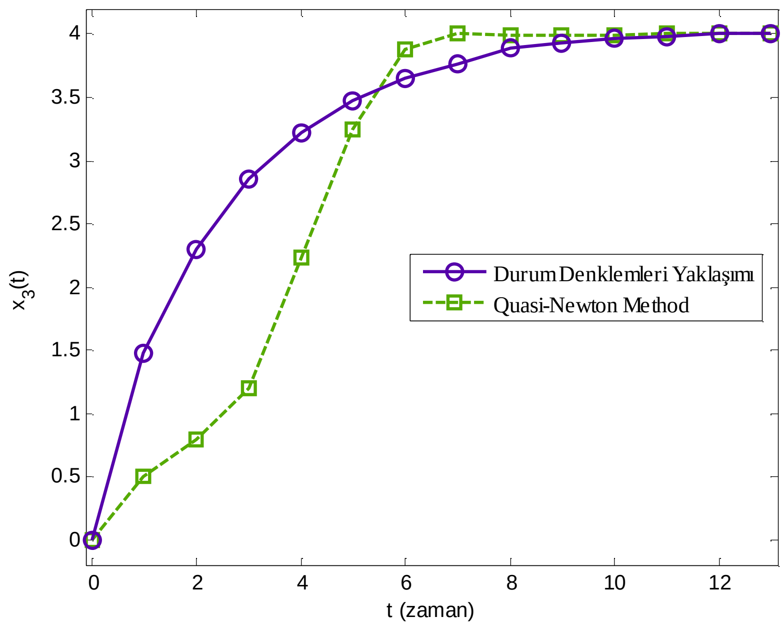
Dikkat edilirse durum denklemleri yaklaşımının Quasi-Newton yaklaşımına göre daha düzgün bir yörünge izleyerek optimal çözümlere ulaştığı görülür.



Şekil 3.1 (3.5) Probleminin x_1 optimum çözümü



Şekil 3.2 (3.5) Probleminin x_2 optimum çözümü



Şekil 3.3 (3.5) Probleminin x_3 optimum çözümü

3.2.2 Problem: Aşağıdaki kısıtlamasız optimizasyon problemini $-n^2 \leq x_i \leq n^2$ olmak üzere ele alalım [2],

$$\underset{x\in\mathbb{R}^n}{\text{minimum}} \left(\sum_{i=1}^{n} (x_i - 1)^2 - \sum_{i=2}^{n} x_i x_{i-1} \right). \tag{3.8}$$

Özel olarak $n = 6$ alırsak (3.8) probleminin analitik optimum çözümü $x^* = (6,10,12,12,10,6)^T$ 'dir. Aynı şekilde (3.3) durum denklem modelini elde etmek için (3.2) denklemi ve 3.1.2 Tanım'ı gereği,

$$\dot{x} = \begin{pmatrix} \dot{x}1 \\ \dot{x}2 \\ \dot{x}3 \\ \dot{x}4 \\ \dot{x}5 \\ \dot{x}6 \end{pmatrix} = \begin{pmatrix} -2 & 1 & 0 & 0 & 0 & 0 \\ 1 & -2 & 1 & 0 & 0 & 0 \\ 0 & 1 & -2 & 1 & 0 & 0 \\ 0 & 0 & 1 & -2 & 1 & 0 \\ 0 & 0 & 0 & 1 & -2 & 1 \\ 0 & 0 & 0 & 0 & 1 & -2 \end{pmatrix} \begin{pmatrix} x1 \\ x2 \\ x3 \\ x4 \\ x5 \\ x6 \end{pmatrix} + \begin{pmatrix} 2 \\ 2 \\ 2 \\ 2 \\ 2 \\ 2 \end{pmatrix} \tag{3.9}$$

$$y=\begin{pmatrix}1&1&0&0&0&0\\0&1&0&0&0&0\\0&0&1&0&0&0\\0&0&0&1&0&0\\0&0&0&0&1&0\\0&0&0&0&0&1\end{pmatrix}\begin{pmatrix}x1\\x2\\x3\\x4\\x5\\x6\end{pmatrix}$$

diferansiyel denklem sistemi elde edilir. A matrisinin tüm özdeğerleri sol yarı düzlemde olduğu için sistem kararlıdır. Buradan sistemin transfer fonksiyonu

$$G(s)=\frac{1}{s^6+12s^5+55s^4+120s^3+126s^2+56s+7}\begin{pmatrix}2s^5+22s^4+90s^3+168s^2+140s+42\\2s^5+24s^4+108s^3+224s^2+210s+70\\2s^5+24s^4+110s^3+238s^2+238s+84\\2s^5+24s^4+110s^3+238s^2+238s+84\\2s^5+24s^4+108s^3+224s^2+210s+70\\2s^5+22s^4+90s^3+168s^2+140s+42\end{pmatrix}$$

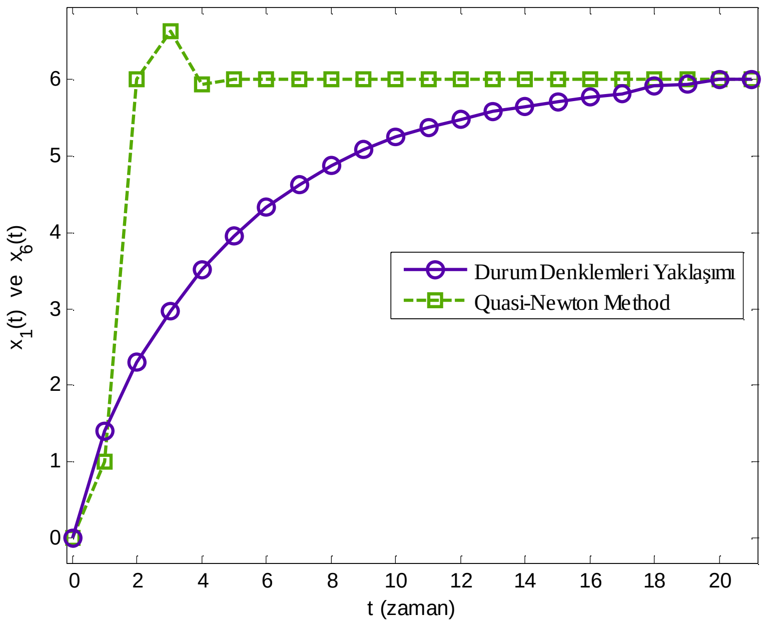
biçiminde bulunabilir. Transfer fonksiyonunun birim basamak yanıtı incelendiğinde (3.9) sisteminin denge noktasının (3.8) kısıtlamasız optimizasyon probleminin analitik çözümüne karşılık geldiği görülür.

Şimdi bu denge noktasının kararlılık analizinin gerçekleştirilmesi için bir önceki problemde olduğu gibi $z=x-x*$ dönüşümü altında orijin noktasında denge noktasına sahip

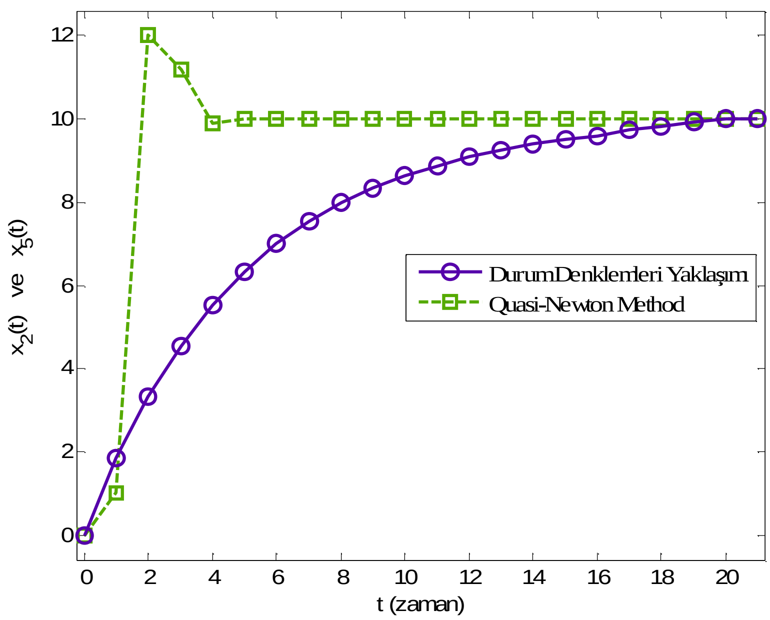
$$\dot{z}=\begin{pmatrix}-2&1&0&0&0&0\\1&-2&1&0&0&0\\0&1&-2&1&0&0\\0&0&1&-2&1&0\\0&0&0&1&-2&1\\0&0&0&0&1&-2\end{pmatrix}\begin{pmatrix}z_1\\z_2\\z_3\\z_4\\z_5\\z_6\end{pmatrix}$$

sistemini oluşturulur. $Q=I_{6\times 6}$ olmak üzere Lyapunov matris denkleminin analitik çözümü yapıldığında tek bir P pozitif tanımlı simetrik matris çözümü elde edilir. 3.1.3 Teorem gereği $x*$ noktası asimptotik kararlı bir noktadır.

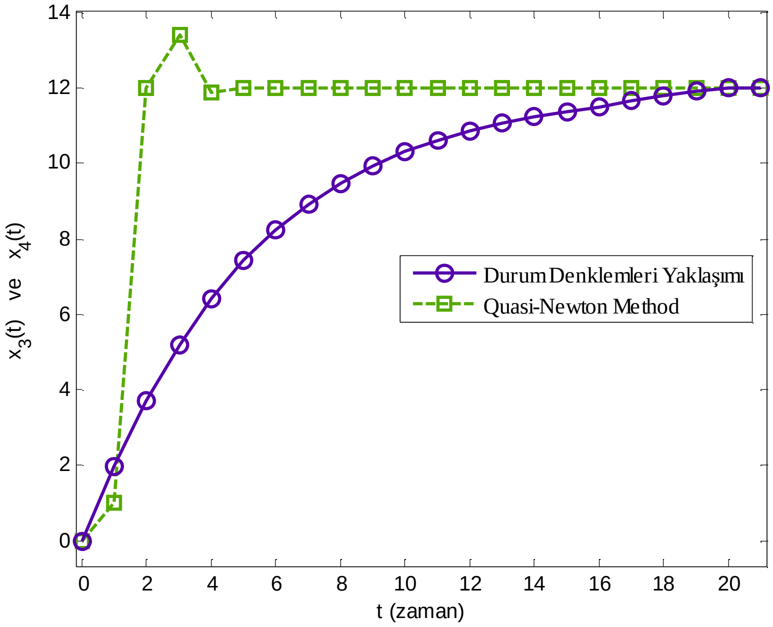
Aşağıdaki şekillerde de görülebileceği gibi durum denklemleri yaklaşımı ile elde edilen çözüm yörüngeleri bir atım yapmadan düzgün bir şekilde optimum çözümlere ulaşmaktadır



Şekil 3.4 (3.8) Probleminin x_1 ve x_6 optimum çözümleri



Şekil 3.5 (3.8) Probleminin x_2 ve x_5 optimum çözümleri



Şekil 3.6 (3.8) Probleminin x_3 ve x_4 optimum çözümleri

4. KISITLAMALI OPTİMİZASYON PROBLEMLERİ İÇİN DURUM DENKLEMLERİ YAKLAŞIMI

Bu bölümde kısıtlamalı optimizasyon problemlerinin optimum sonuçlarının durum denklemleri yaklaşımı ile araştırılmasının sonuçları ortaya konulacaktır.

4.1 Kuadratik Programlama Problemleri için Durum Denklemleri Yaklaşımı

Bu alt bölümde, eşitlik kısıtlamalarına sahip kuadratik programlama optimizasyon problemlerinin, kuadratik ceza metodu kullanılarak oluşturulan durum denklemleri ile optimum noktaları araştırılmıştır. Kuadratik programlama problemlerinin optimum sonuçlarının araştırılmasında oluşturulan durum denklemlerinin karar değişkenlerinin limit noktası, ileri Euler zaman parçalanışı metodu [31,32] ve kapalı zaman parçalanışı metodu [33] yerine dinamik sistemin transfer fonksiyonu kullanılmaktadır. Ortaya konulan sonuçlar göstermiştir ki durum denklemleri ile oluşturulan dinamik sistemin asimptotik kararlı denge noktası aynı zamanda optimizasyon probleminin de mutlak minimumudur.

Ele alınan problem modeli

$$\begin{aligned} &\text{minimum} && f(x) = \frac{1}{2}x^T G x + g^T x \\ &\text{kısıtlamalar} && Ex = d \end{aligned} \tag{4.1}$$

x karar değişkeni, G, $n \times n$ boyutlu reel, simetrik, pozitif tanımlı bir matris, kısıtlamaları ise E, tam satır ranklı $m \times n$ boyutlu bir matris ve d sağ taraf sabitlerini gösteren bir sütun matrisi olmak üzere ifade edilsin.

İkinci bölümde kısıtlamalı optimizasyon problemlerinin çözümlerinin araştırılmasında en yaygın olarak kullanılan metotlardan birinin ceza (penaltı) yöntemi olduğundan bahsedilmişti, bu yöntemde kısıtlamalı optimizasyon problemleri amaç fonksiyonuna kısıtlamaların ihlali durumunda ceza uygulayacak bir fonksiyonun eklenmesi ile kısıtlamasız optimizasyon problemlerinin bir dizisine dönüştürülmektedir. Bu kısıtlamasız optimizasyon problemlerinin çözüm dizisinin bir limit noktası aynı zamanda orijinal kısıtlamalı optimizasyon probleminin de minimumuna karşılık gelmektedir.

Eşitlilik kısıtlamaları için kullanılan ceza fonksiyonlarından biri de

$$\phi(x) = \frac{1}{2}\|Ex - d\|_2^2 \tag{4.2}$$

kuadratik ceza fonksiyonudur. (4.2) ceza fonksiyonunu kullanarak (4.1) eşitlik kısıtlamalı optimizasyon problemini, $\forall k$ için $0 < p_k < p_{k+1}$ ve $p_k \to \infty$ penaltı parametresi olmak üzere

$$\underset{x \in \mathbb{R}^n}{\text{minimum}}\, P_{p_k}(x) = f(x) + p_k \phi(x) \tag{4.3}$$

kısıtlamasız optimizasyon problemlerinin bir dizisi biçiminde ifade edebiliriz. Bu alt bölüm boyunca $\forall k$ için (4.3) probleminin bir x_k çözümünün var olduğunu kabul edeceğiz.

(4.1) kısıtlamalı optimizasyon probleminin mutlak minimumunun araştırılmasında (4.3) ceza problemi yardımıyla oluşturacağımız durum denklemlerinden yararlanacağız.

4.1.1 Ana Sonuçlar

$P_{p_k}(x)$ ceza (penaltı) fonksiyonunun gradyant vektörü $\nabla_x P_{p_k}(x)$ olmak üzere

$$\begin{aligned} \frac{dx}{dt} &= \nabla_x P_{p_k}(x) \\ x(0) &= x_0 \end{aligned} \tag{4.4}$$

veya matris formunda

$$\dot{x}(t)=\begin{pmatrix}\dot{x}_1(t)\\ \dot{x}_2(t)\\ \vdots\\ \dot{x}_n(t)\end{pmatrix}=\begin{pmatrix}a_{11}(p) & a_{12}(p) & \cdots & a_{1n}(p)\\ a_{21}(p) & a_{22}(p) & \cdots & a_{2n}(p)\\ \vdots & \vdots & \ddots & \vdots\\ a_{n1}(p) & a_{n2}(p) & \cdots & a_{nn}(p)\end{pmatrix}\begin{pmatrix}x_1(t)\\ x_2(t)\\ \vdots\\ x_n(t)\end{pmatrix}+\begin{pmatrix}b_1(p)\\ b_2(p)\\ \vdots\\ b_n(p)\end{pmatrix}u(t)$$

diferansiyel denklem sistemi ele alındığında, bu sistem, p yeterince büyük penaltı parametresi olmak üzere durum denklemleri modelinde

$$\begin{aligned}\dot{x}(t)&=A(p)x(t)+B(p)u(t)\\ y(t)&=Cx(t)\end{aligned}\tag{4.5}$$

gösterilir. Problemlerin çözümlerinde $u(t)=\begin{cases}1, & t\geq 0\\ 0, & t<0\end{cases}$ olarak alınacaktır. (4.5) Durum denkleminde yer alan matrisler bir önceki bölümdeki gibi tanımlanmışlardır.

4.1.1.1 Uyarı: Çalışmanın bu kısmında aşağıdaki koşulların sağladığı kabul edilmiştir.

a) (4.5) Durum denklem sisteminde yer alan $A(p)$ matrisinin tüm özdeğerlerinin, ya $\operatorname{Re}\lambda_i \leq 0$ yada $\operatorname{Re}\lambda_i \geq 0$ koşulunu sağlamalıdır. Eğer $\operatorname{Re}\lambda_i = 0$ ise bu özdeğerlere karşılık gelen Jordan bloğunun derecesi bir olmalıdır.

b) (4.5) Durum denklem sistemi bir minimal gerçekleştirmeye sahiptir.

4.1.1.2 Tanım: (4.5) sisteminde yer alan $A(p)$ matrisinin tüm özdeğerleri λ ile göstermek üzere

(a) $\operatorname{Re}\lambda(A(p))\leq 0 \Rightarrow \tilde{A}(p):=A(p)$ ve $\tilde{B}(p):=B(p)$

(b) $\operatorname{Re}\lambda(A(p))\geq 0 \Rightarrow \tilde{A}(p):=-A(p)$ ve $\tilde{B}(p):=-B(p)$

ile tanımlayalım.

Yukarıdaki tanım ile (4.5) sisteminin her zaman kararlı bir sistem olmasını garanti altına alınır.

4.1.1.3 Yardımcı Teorem: Eğer yeterince büyük p penaltı parametresi için $n \times n$ boyutlu $A(p)$ matrisi kararlı ise $\forall t \geq 0$ için

$$\left\| e^{A(p)t} \right\| \leq K e^{-kt}$$

K ve k pozitif sabitleri var olmalıdır.

Yukarıdaki yardımcı teoremin ispatı 2.3.2.2 Teorem'in ispatından yararlanılarak gösterilebilir.

4.1.1.4 Teorem: Varsayalım ki yeterince büyük bir p ceza parametresi için $P_{p_k}(x)$ ceza fonksiyonu diferansiyellenebilsin. $x*$ noktasının (4.1) eşitlik kısıtlamasına sahip kuadratik programlama problemi için mutlak minimum olması için gerek ve yeter şart $x*$ noktasının (4.5) durum denklem sisteminin asimptotik kararlı bir denge noktası olmasıdır.

İspat: Farz edilsin ki $x*$ (4.1) optimizasyon probleminin bir mutlak minimumu olsun. Konvekslik gereği $x*$ minimum noktası aynı zamanda (4.3) kısıtlamasız problem dizisinin de yeterince büyük bir $p_k > 0$ ceza parametresi için bir minimum noktasıdır. Buradan 2.1.1.3 Önerme gereği $\nabla_x P_{p_k}(x*) = 0$ olmalıdır. Bu $x*$'ın (4.4) dinamik sistemi dolayısıyla (4.5) sistemi için bir denge noktası olduğunu verir. Şimdi bu noktanın asimptotik kararlı bir denge noktası olduğu gösterilsin.

Başlangıç koşulu olarak $x(t_0) = x_0$ ve $M > 0$ için $\left\| u(t) \right\| < M$ olarak alınsın, (4.5) diferansiyel denklem sisteminin çözümü

$$x(t) = e^{A(p)(t-t_0)} x_0 + \int_{t_0}^{t} e^{A(p)(t-\tau)} B(p) u(\tau) d\tau$$

dur. 2.3.1.1 Tanım gereği her $\varepsilon > 0$ ve en az bir $x(0) = x_0$ başlangıç koşulu için

$$\|x_0 - x*\| < \delta \ \text{ iken } \|x(t) - x*\| < \varepsilon \,, \ \forall t \geq 0$$

olacak şekilde bir $\delta = \delta(\varepsilon) > 0$ varlığını gösterelim. Buradan 4.1.1.2 Tanım ve 4.1.1.3 Yardımcı Teorem yardımı ile

$$\begin{aligned}
\|x(t) - x*\| &= \left\| e^{A(p)(t-t_0)} x_0 + \int_{t_0}^{t} e^{A(p)(t-\tau)} B(p) u(\tau) d\tau - x* \right\| \\
&\leq \left\| e^{A(p)(t-t_0)} x_0 - x* \right\| + \left\| \int_{t_0}^{t} e^{A(p)(t-\tau)} B(p) u(\tau) d\tau \right\| \\
&\leq \left\| e^{A(p)(t-t_0)} x_0 - e^{A(p)(t-t_0)} x* + e^{A(p)(t-t_0)} x* - x* \right\| + \int_{t_0}^{t} \left\| e^{A(p)(t-\tau)} \right\| \|B(p)\| \|u(\tau)\| d\tau \\
&\leq \left\| e^{A(p)(t-t_0)} \left(x_0 - x*\right) \right\| + \left\| \left(e^{A(p)(t-t_0)} - I_{n\times n} \right) x* \right\| + M \|B(p)\| \int_{t_0}^{t} \left\| e^{A(p)(t-\tau)} \right\| d\tau
\end{aligned}$$

$$\begin{aligned}
&\leq \left\| e^{A(p)(t-t_0)} \right\| \|x_0 - x*\| + \left(\left\| e^{A(p)(t-t_0)} \right\| + \|I_{n\times n}\| \right) \|x*\| + M \|B(p)\| \int_{t_0}^{t} \left\| e^{A(p)(t-\tau)} \right\| d\tau \\
&\leq K e^{k(t_0 - t)} \|x_0 - x*\| + K e^{k(t_0 - t)} \|x*\| + \|x*\| + MK \|B(p)\| \int_{t_0}^{t} e^{k(\tau - t)} d\tau \\
&= K e^{k(t_0 - t)} \|x_0 - x*\| + K e^{k(t_0 - t)} \|x*\| + \|x*\| + MKk^{-1} \|B(p)\| \left(1 - e^{k(t_0 - t)}\right) \\
&\leq K \|x_0 - x*\| + K \|x*\| + \|x*\| + MKk^{-1} \|B(p)\|
\end{aligned}$$

elde edilir. Eğer

$$\delta = K^{-1} \left(\varepsilon - \|x*\| \left(K + 1\right) - MKk^{-1} \left\|B\left(p\right)\right\| \right)$$

olarak alınırsa $x*$ denge noktasının (4.4) sistemi için kararlı bir denge noktası olduğu görülür. Ayrıca

$$
\begin{aligned}
\|y(t)\| &= \left\| Ce^{A(p)(t-t_0)}x_0 + \int_{t_0}^{t} Ce^{A(p)(t-\tau)}B(p)u(\tau)d\tau \right\| \\
&\le \|C\|\left\|e^{A(p)(t-t_0)}\right\|\|x_0\| + \|C\|\|B(p)\|M\int_{t_0}^{t}\left\|e^{A(p)(t-\tau)}\right\|d\tau \\
&\le K\|C\|\left[\|x_0\|e^{k(t_0-t)} + \|B(p)\|M\int_{t_0}^{t}e^{k(\tau-t)}d\tau\right] \\
&= K\|C\|\left[\|x_0\|e^{k(t_0-t)} + k^{-1}\|B(p)\|M\left(1-e^{k(t_0-t)}\right)\right] \\
&\le K\|C\|\left\{\|x_0\| + k^{-1}\|B(p)\|M\right\}
\end{aligned}
$$

olduğundan (4.5) sistemi $x*$ denge noktasında sınırlı girdi ve sınırlı çıktı kararlılığına sahiptir. Dolayısıyla 2.3.3.4 Teorem ve varsayımlar gereği $x*$ denge noktası (4.5) sistemi için asimptotik kararlı bir denge noktasıdır.

Tersine, $x*$ denge noktası (4.5) sistemi için asimptotik kararlı bir denge noktası olsun. 2.3.1.2 Tanım gereği $x*$ denge noktası kararlı ve en az bir $x(0) = x_0$ başlangıç koşulu için

$$
\|x_0 - x*\| < \delta \text{ iken } \lim_{t\to\infty}\|x(t) - x*\| = 0
$$

olacak şekilde bir $\delta > 0$ vardır. Buradan

$$
\begin{aligned}
\lim_{t\to\infty}\|x(t) - x*\| = \left\|\lim_{t\to\infty}\left(x(t) - x*\right)\right\| &= 0, \\
\Leftrightarrow \lim_{t\to\infty}\left(x(t) - x*\right) &= 0 \\
\Leftrightarrow \lim_{t\to\infty} x(t) &= x*
\end{aligned}
$$

görülür. Buda $x*$ asimptotik karalı denge noktasının (4.1) optimizasyon probleminin mutlak minimum noktası olduğunu verir.□

Test problemlerinin minimumlarının araştırılmasında aşağıdaki algoritma kullanılabilir.

Durum Denklemleri Yaklaşımı:

1. **Basamak:** $\varepsilon > 0$, $p_1 > 0$ ve $k = 1$ alınsın.
2. **Basamak:** Kuadratik programlama problemini $P(x)$ penaltı fonksiyonu ile kısıtlamasız optimizasyon problemine çevir.
3. **Basamak:** (4.4) Adi diferansiyel denklem sistemini kullanarak (4.5) durum denklem sistemini oluştur.
4. **Basamak:** 4.1.1.2 Tanım'dan $A(p)$ matrisinin kararlılığını kontrol et.
5. **Basamak:** $G(s)$ transfer fonksiyonunu oluştur.
6. **Basamak:** $[c, x, t] = \text{step}[num, den, t]$ basamak cevabını kullanarak $\underset{x \in \mathbb{R}^n}{\text{minimum}} \; P_{p_k}(x)$ kısıtlamasız optimizasyon probleminin x_k çözümlerini elde et
7. **Basamak:** $\|h(x_k)\| \le \varepsilon$ ise iterasyonu durdur. Aksi takdirde $x_{k+1} = x_k$, $p_{k+1} = p_k + 100$, $k = k + 1$ alarak 2. Basamağa git.

4.1.2 Test Problemleri ve Sonuçlar

Durum denklemleri yaklaşımı Matlab 7.0 programı kullanılarak eşitlik kısıtlamasına sahip kuadratik programlama problemlerinin optimum çözümleri elde edilmiştir. Elde edilen çözümler SQP (Sequential Quadratic Programming) metodu ile karşılaştırılmıştır.

4.1.2.1 Problem: Eşitlik kısıtlamasına sahip kuadratik programlama problemini ele alalım [2],

$$\begin{aligned} &\text{minimum} \;\; x_1^2 - 2x_1 + x_2^2 - x_3^2 + 4x_3 \\ &\text{kısıtlama} \;\; x_1 - x_2 + 2x_3 - 2 = 0. \end{aligned} \tag{4.6}$$

Problemin analitik çözümünden elde edilen minimum noktası $x^* = (2.5, -1.5, -1)^T$.
(4.1) kuadratik ceza fonksiyonu yardımı ile kısıtlamasız optimizasyon problemi

$$\text{minimum } P_{p_k}(x) = x_1^2 - 2x_1 + x_2^2 - x_3^2 + 4x_3 + \frac{1}{2}p\left(x_1 - x_2 + 2x_3 - 2\right)^2$$

formunda yazılabilir.

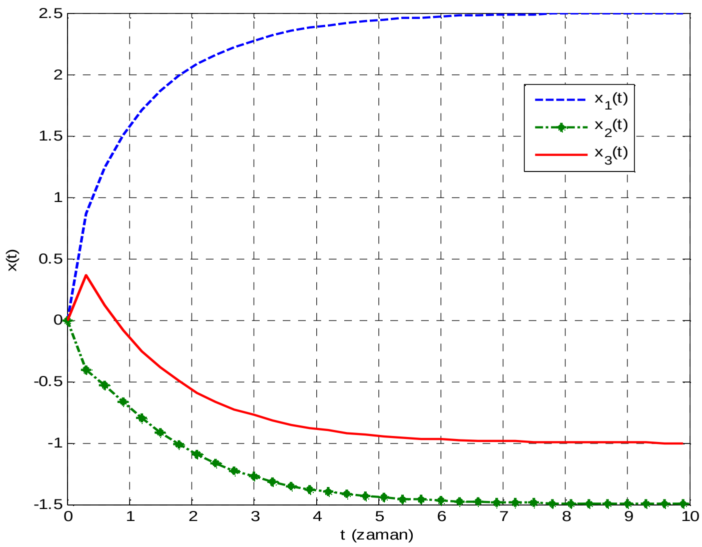
4.1.1.2 Tanım ve (4.4) dinamik sistemi kullanılarak durum denklem sistemi ile

$$\begin{aligned} \begin{pmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \\ \dot{x}_3(t) \end{pmatrix} &= \begin{pmatrix} -2-p & p & -2p \\ p & -2-p & 2p \\ -2p & 2p & 2-4p \end{pmatrix} \begin{pmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{pmatrix} + \begin{pmatrix} 2p+2 \\ -2p \\ 4p-4 \end{pmatrix} \\ y &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{pmatrix} \end{aligned} \tag{4.7}$$

ifade edilebilir. Ceza parametresi $p = 1000$ olarak alınırsa $A(p)$ matrisinin tüm özdeğerleri sol yarı düzlemde yer alır. (4.7) sisteminin transfer fonksiyonu ise

$$G(s) = \frac{1}{s^3 + 6002s^2 + 15996s + 7992} \begin{pmatrix} 2002s^2 + 18000s + 19992 \\ -2000s^2 - 6000s - 12000 \\ 3996s^2 + 3984s - 8016 \end{pmatrix}$$

şeklindedir. Buradan $G(s)$ transfer fonksiyonunun birim girdi-çıktısı incelendiğinde (4.7) sisteminin denge noktasının aynı zamanda (4.6) probleminin mutlak optimumu olduğu görülür. Aşağıdaki şekillerde elde edilen sonuçların SQP metodu ile karşılaştırmaları verilmiştir. Durum denklemleri yaklaşımı $(1,1,1)$, $(-1,1-,1)$, $(5,2,-1)$ ve $(2,-3,4)$ gibi farklı başlangıç şartları içinde optimizasyon probleminin optimum sonuçlarına yakınsamıştır.



Şekil 4.1 (4.6) Probleminin $x=(x_1,x_2,x_3)^T$ optimum çözümü

4.1.2.2 Problem: Aşağıdaki eşitlik kısıtlamalarına sahip kuadratik programlama problemini ele alalım [53],

$$\begin{aligned} &\text{minimum} \quad f(x)=(x_1-x_2)^2+(x_2+x_3-2)^2+(x_4-1)^2+(x_5-1)^2 \\ &\text{kısıtlama} \quad x_1+3x_2-4=0 \\ &\qquad\qquad\quad x_3+x_4-2x_5=0 \\ &\qquad\qquad\quad x_2-x_5=0. \end{aligned} \tag{4.8}$$

(4.8) test probleminin teorik optimum çözümü $x^*=(1,1,1,1,1)^T$ 'dır. (4.1) kuadratik ceza fonksiyonu yardımı ile kısıtlamasız optimizasyon problemi

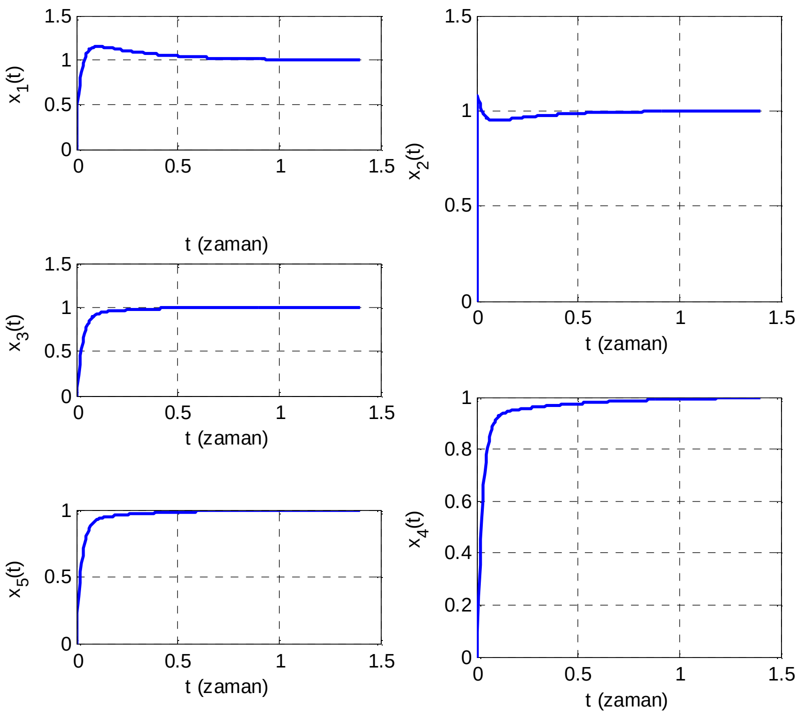
$$\begin{aligned} P_{p_k}(x)=&(x_1-x_2)^2+(x_2+x_3-2)^2+(x_4-1)^2+(x_5-1)^2 \\ &+\frac{1}{2}p(x_1+3x_2-4)+\frac{1}{2}p(x_3+x_4-2x_5)+\frac{1}{2}p(x_2-x_5). \end{aligned}$$

Buradan (4.8) optimizasyon probleminin durum denklem sistemi 4.1.1.2 Tanım ve (4.4) dinamik sistemi yardımı ile

$$
\begin{aligned}
\dot{x}(t) &= A(p)x(t) + B(p)u(t) \\
y(t) &= Cx(t)
\end{aligned} \tag{4.9}
$$

$$
A(p) = \begin{bmatrix} -p-2 & -3p+2 & 0 & 0 & 0 \\ -3p+2 & -10p-4 & 2 & 0 & p \\ 0 & -2 & -p-2 & -p & 2p \\ 0 & 0 & -p & -p-2 & 2p \\ 0 & p & 2p & 2p & -5p-2 \end{bmatrix}, \; B(p) = \begin{bmatrix} 4p \\ 12p+4 \\ 4 \\ 2 \\ 2 \end{bmatrix} \text{ ve}
$$

$C = I_{5\times 5}$ matrisleri olmak üzere gösterilir. Ceza parametresi $p = 1000$ olarak alınırsa $A(p)$ matrisinin özdeğerleri $0, -2, -2, -6$ olmaktadır. Sistemin transfer fonksiyonu $G(s) = C(sI - A)^{-1}B + D$ olmak üzere durum denklemleri yaklaşımı sıfır başlangıç koşulları altında uygulandığında (4.9) dinamik sisteminin $x(t)$ karar değişkeni yörüngesinin $t \to \infty$ için (4.8) optimizasyon probleminin optimum çözümüne $x^* = (x_1, x_2, x_3, x_4, x_5)^T = (1,1,1,1,1)^T$ yaklaştığı görülmektedir.



Şekil 4.2 (4.8) Probleminin $x = (x_1, x_2, x_3, x_4, x_5)^T$ optimum çözümü

4.1.2.3 Problem: Doğrusal kuadratik regulatör probleminin özel bir formu olan aşağıdaki optimizasyon problemini ele alalım [2],

$$\begin{aligned} &\text{minimum} \quad \frac{1}{2}\sum_{i=1}^{N}\left(qx_i^2 + rz_i^2\right) \\ &\text{kısıtlama} \quad \mathrm{x}_i = ax_{i-1} - bz_i, \quad i = 1, 2, ..., N. \end{aligned} \tag{4.10}$$

Bu problemde x_i, i. ay sonundaki hesap dengesini ve z_i, i. ay sonundaki ödemeleri göstermektedir. Başlangıç şartı ve sabitler olarak $x_0 = 10000$, $q = 1$, $r = 300$, $a = 1.02$, $b = 1$ ve penaltı parametresi olarak $c = 10^6$ alınsın, $N = 10$ alınırsa (4.2) kuadratik ceza fonksiyonunu yardımı ile kısıtlamasız optimizasyon problemi

$$\text{minimum} \quad P_{p_k}(x) = \frac{1}{2}\sum_{i=1}^{10}\left(qx_i^2 + rz_i^2\right) + \frac{1}{2}\sum_{i=1}^{10}\left(\mathrm{x}_i - 1.02x_{i-1} + z_i\right)^2$$

yazılabilir. Durum denklemleri yaklaşımının kullanılabilmesi için 4.1.1.2 Tanım ve (4.4) dinamik sistemi yardımı ile diferansiyel denklem sistemi,

$$A_{11} = \begin{bmatrix} -q-2.0404p & 1.02p & 0 & \cdots & 0 \\ 1.02p & -q-2.0404p & 1.02p & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & 1.02p & -q-2.0404p & 1.02p \\ 0 & \cdots & 0 & 1.02p & -q-p \end{bmatrix},$$

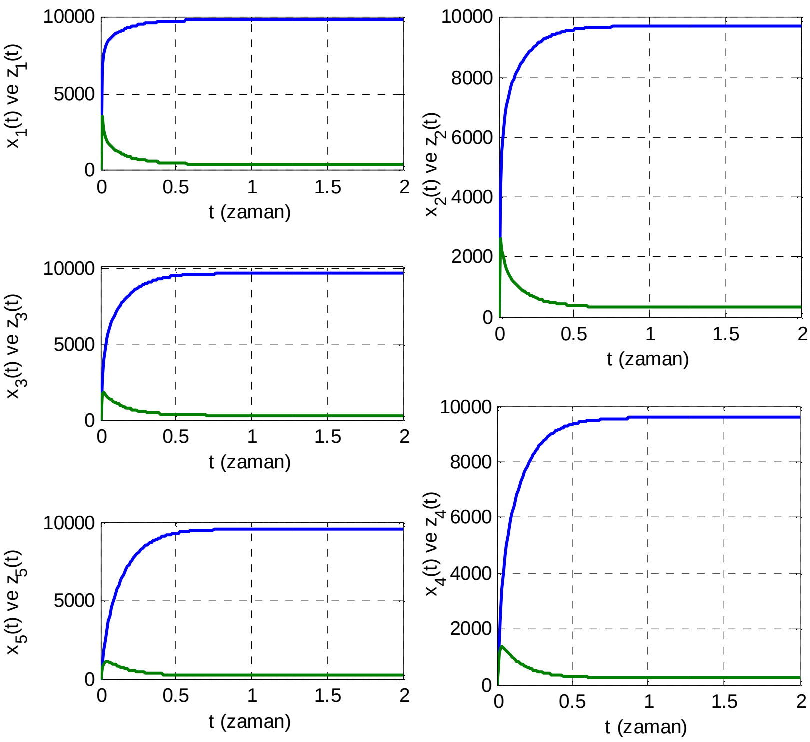
$$A_{12} = \begin{bmatrix} -p & 1.02p & 0 & \cdots & 0 \\ 0 & -p & 1.02p & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ & & & & 1.02p \\ 0 & \cdots & & 0 & -p \end{bmatrix}, \; A_{21} = \begin{bmatrix} -p & 0 & \cdots & & 0 \\ 1.02p & \ddots & \ddots & & \vdots \\ 0 & \ddots & & & \\ \vdots & \ddots & & & 0 \\ 0 & \cdots & 0 & 1.02p & -p \end{bmatrix} \text{ ve}$$

$$A_{22}=\begin{bmatrix} -r-p & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & & 0 \\ 0 & \cdots & 0 & -r-p \end{bmatrix}$$ 10×10 boyutlu kare matrisler ve

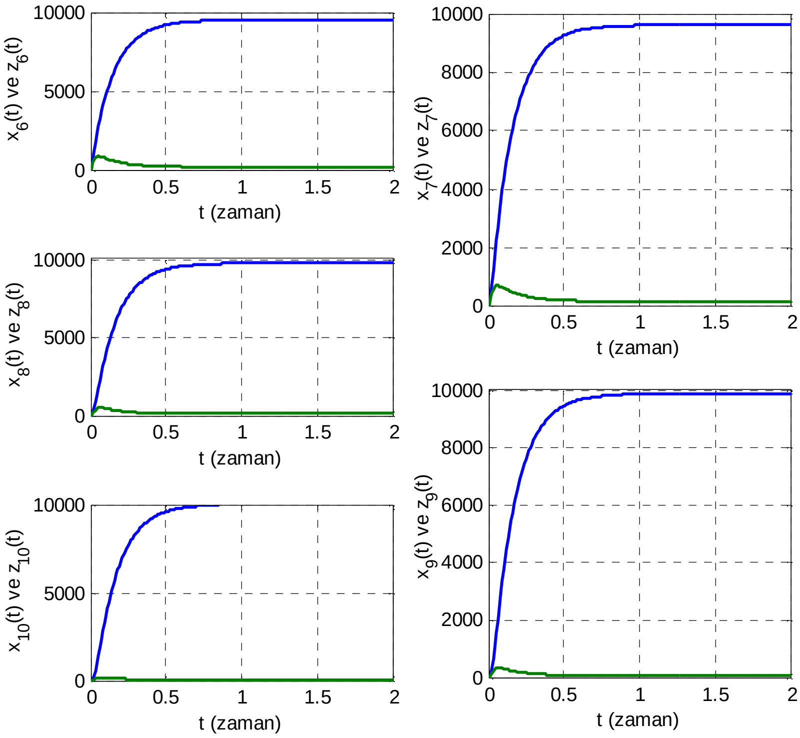
$$B_{11}=\begin{bmatrix} 1.02\times10^4 p \\ 0 \\ \vdots \end{bmatrix},\ B_{21}=\begin{bmatrix} 1.02\times10^4 p \\ 0 \\ \vdots \end{bmatrix}$$ 10×1 boyutlu sütun matrisler olmak üzere

$$\begin{pmatrix} \dot{x}(t) \\ \dot{z}(t) \end{pmatrix}=\begin{pmatrix} A_{11}(p) & A_{12}(p) \\ A_{21}(p) & A_{22}(p) \end{pmatrix}\begin{pmatrix} x(t) \\ z(t) \end{pmatrix}+\begin{pmatrix} B_{11}(p) \\ B_{21}(p) \end{pmatrix}u(t) \tag{4.11}$$

elde edilir. (4.11) sistemine durum denklemleri yaklaşımını sıfır başlangıç koşulları altında uygulanırsa elde edilen optimum çözümlerin (4.10) orijinal kısıtlamalı optimizasyon probleminin analitik optimum çözümlerine karşılık geldiği görülür. Durum denklemleri yaklaşımının (4.10) optimizasyon problemi için elde ettiği çözümler aşağıdaki şekillerde görülmektedir.



Şekil 4.3 (4.10) Probleminin x ve z optimum çözümleri



Şekil 4.4 (4.10) Probleminin x ve z optimum çözümleri

Çizelge 4.1 (4.10) Probleminin Optimum Çözümleri

n	1	2	3	4	5	6	7	8	9	10
x_n	9844.7	9725.4	9641.6	9593.2	9579.9	9601.7	9658.6	9750.8	9878.8	10043
z_n	355.34	316.2	278.22	241.25	205.17	169.84	135.13	100.92	67.075	33.476

Durum denklemleri yaklaşımını SQP metodu ile CPU zamanı olarak karşılaştırmak istersek, durum denklemleri yaklaşımının SQP metoduna göre çok yakın olmakla birlikte daha kısa bir zamanda optimum çözüme ulaştığı görülmektedir.

Çizelge 4.2 SQP Metodu ve Durum Denklemleri Yaklaşımının
CPU zamanlarını karşılaştırılması

Problem	Durum Denklemleri Yaklaşımı	SQP Metodu
4.1.2.1 Problem	0.04	0.07
4.1.2.2 Problem	0.07	0.09
4.1.2.3 Problem	0.34	0.39

4.2 Doğrusal Olmayan Programlama Problemleri için Durum Denklemleri Yaklaşımı

Bu kısımda genel olarak (2.1), (2.3) denklemleri ile ifade ettiğimiz eşitsizlik kısıtlamalarına sahip doğrusal olmayan optimizasyon problemlerinin optimum noktaları durum denklemleri yaklaşımı ile araştırılacaktır. Bir önceki kısımda olduğu gibi ceza fonksiyonları kısıtlamalı optimizasyon problemlerinin çözümlerinde en önemli çözüm metotlarından biri olacaktır.

Durum denklemleri yaklaşımında kullanacağımız ceza fonksiyonu Meng'in [12] nolu referansından alınmıştır. Bu ceza fonksiyonunda ceza parametresi kısıtlamalar yerine amaç fonksiyonuna eklenmiştir ve bu parametrenin seçimi iterasyonlara bağlanmıştır.

Eşitsizlik kısıtlamalarına sahip doğrusal olmayan optimizasyon problemi $f:\mathbb{R}^n \to \mathbb{R}$ ve $g:\mathbb{R}^n \to \mathbb{R}^p$ konveks ve ikinci mertebeden diferansiyellenebilen sürekli fonksiyonlar olmak üzere

$$\begin{aligned} &\text{minimum } f(x) \\ &\text{kısıtlama } g_i(x) \le 0 \quad i = 1,2,\ldots,p \end{aligned} \tag{4.12}$$

ile ifade edilebilir. [12] nolu çalışma yardımı ile ceza fonksiyonu ise $M \in \mathbb{R}$ ve $g_i^+(x) = \max\{0, g_i(x)\}$ ile ifade edilmek üzere

$$F(x,M) = \left(f(x) - M\right)^2 + \sum_{i=1}^{m} g_i^+(x)^2 \tag{4.13}$$

yazılır. Buradan kısıtlamasız optimizasyon problemi ise

$$\underset{x\in\mathbb{R}^n}{\text{minimum}} \; F(x,M) \tag{4.14}$$

dur. Meng [12], (4.14) kısıtlamasız optimizasyon probleminin optimumunun belirli bir $M \in \mathbb{R}$ ceza parametresi altında (4.12) kısıtlamalı optimizasyon probleminin optimum çözümünün karşılık geldiğini göstermiştir. Ayrıca Liu ve Meng [54] oluşturdukları ceza fonksiyonunu neural network sistemine adapte ederek doğrusal olmayan kısıtlamalı optimizasyon problemlerinin çözümlerini incelemişlerdir.

Direkt bir hesaplama ile (4.13) ceza fonksiyonunun gradyant ve Hesse vektörleri aşağıdaki gibi bulunabilir,

$$\begin{aligned}
\nabla_x F(x,M) &= 2\big(f(x)-M\big)\nabla_x f(x) + 2\sum_{i=1}^{p} g_i^+(x)\nabla_x g_i(x) \\
\nabla_M F(x,M) &= -2\big(f(x)-M\big) \\
\nabla_{xx}^2 F(x,M) &= 2\nabla_x f(x)\nabla_x f(x)^T + 2\big(f(x)-M\big)\nabla_{xx}^2 f(x) \\
&\quad + 2\sum_{i=1}^{p}\nabla_x g_i(x)\nabla_x g_i(x)^T + 2\sum_{i=1}^{p} g_i^+(x)\nabla_{xx}^2 g_i(x).
\end{aligned}$$

(4.14) kısıtlamasız optimizasyon probleminin çözümü için durum denklemleri yaklaşımı $\nabla_x F(x,M)$ ve $\nabla_M F(x,M)$ (4.13) ceza fonksiyonunun gradyant vektörleri olmak üzere

$$\frac{dx}{dt} = -\nabla_x F(x,M) \tag{4.15a}$$

$$\frac{dM}{dt} = -\nabla_M F(x,M) \tag{4.15b}$$

diferansiyel denklem sistemi ile ifade edilebilir.

4.2.1 Ana Sonuçlar

4.2.1.1 Teorem: Bir (x^*, M^*) noktası $M^* \le f(x^*)$ ve $F(x^*, M^*) > 0$ olmak üzere (4.15) durum denklemleri sisteminin bir denge noktası ise x^* (4.14) probleminin bir optimum çözümüdür.

İspat: (x^*, M^*) noktası (4.15) diferansiyel denklem sisteminin bir denge noktası ise

$$\frac{dx^*}{dt} = -2\left(f(x^*) - M^*\right)\nabla_x f(x^*) - 2\sum_{i=1}^{p} g_i^+(x^*)\nabla_x g_i(x^*) = 0 \tag{4.16}$$

$$\frac{dM^*}{dt} = 2\left(f(x^*) - M^*\right) = 0 \tag{4.17}$$

olmalıdır. Buradan $\nabla_x F(x, M) = 0$ olduğundan 2.1.1.3 Yardımcı Teorem gereği optimallik için birinci mertebeden gerek şart sağlanır. (4.16) ve (4.17) denklemlerinden

$$\sum_{i=1}^{p} g_i^+(x^*)\nabla_x g_i(x^*) = 0 \text{ ve } f(x^*) = M^*$$

eşitlikleri kolaylıkla görülebilir. (4.13) ceza fonksiyonunun Hesse matrisi ve yukarıdaki eşitlikler göz önüne alındığında $\nabla_{xx}^2 F(x, M) > 0$ elde edilir. Sonuç olarak 2.1.1.6 Teorem gereği x^* noktasının (4.14) optimizasyon probleminin bir optimum noktası olduğu ispatlanır.□

4.2.1.2 Teorem: Bir x^* noktası $M^* \le f(x^*)$ ve $F(x^*, M) > 0$ olmak üzere (4.14) optimizasyon probleminin bir optimum noktası ise (x^*, M^*), (4.15) sisteminin bir denge noktasıdır.

İspat: Varsayım gereği x^* noktası (4.14) probleminin optimum bir noktası olduğundan birinci mertebeden gerek şart gereği

$$\nabla_x F(x^*, M^*) = 0 \text{ veya } \frac{dx^*}{dt} = 0$$

sağlanır. Aynı zamanda x^* noktasının (4.12) problemi için bir uygun çözüm noktası olmasından $g_i^+(x^*) = \max\{0, g_i(x^*)\} = 0$ olmalıdır. Buradan (4.13) ceza fonksiyonunun Hesse vektörü ve varsayımlar gereği

$$\frac{dM^*}{dt} = 0$$

olduğu görülür.□

Şimdi (4.15) durum denklem sisteminin (x^*, M^*) denge noktasının aynı zamanda asimptotik kararlı bir denge noktası olduğu gösterilsin. Bu amaçla 2.3.4 alt bölümünde yer alan Lyapunov kararlılık teorisinden yararlanılacaktır. Sonuç olarak

$$V(x, M) = F(x, M) - F(x^*, M^*) \tag{4.18}$$

aday Lyapunov fonksiyonunu ele alınsın.

4.2.1.3 Teorem: (x^*, M^*) noktası (4.15) durum denklem sistemi için bir denge noktası ise (x^*, M^*) noktası aynı zamanda bu sistem için asimptotik kararlı bir denge noktasıdır.

İspat: Durum denklem sisteminin (4.15) sistemi için Lyapunov duyarlılığında asimptotik kararlı bir denge noktası olması için yapmamız gereken 2.3.4.2 Teorem gereği (4.18) denkleminin bir Lyapunov denklemi olduğunu göstermektir.

Varsayım gereği (x^*, M^*) noktası (4.15) sistemi için bir denge noktası ve x^* 'ın (4.14) problemi için bir optimum nokta olmasından dolayı $V(x^*, M^*) = 0$ ve (x^*, M^*) noktasının delinmiş bir komşuluğunda $V(x, M) > 0$ dır.

Ayrıca $V(x, M)$ fonksiyonunun (x^*, M^*) noktasının delinmiş bir komşuluğunda zamana göre türevi alındığında,

$$
\begin{aligned}
\frac{dV}{dt} &= \frac{\partial V}{\partial x}\frac{dx}{dt} + \frac{\partial V}{\partial M}\frac{dM}{dt} \\
&= \begin{bmatrix} \frac{\partial V}{\partial x} & \frac{\partial V}{\partial M} \end{bmatrix} \begin{bmatrix} \frac{dx}{dt} & \frac{dM}{dt} \end{bmatrix}^T \\
&= \begin{bmatrix} \frac{\partial V}{\partial x} & \frac{\partial V}{\partial M} \end{bmatrix} \begin{bmatrix} -\frac{\partial V}{\partial x} & -\frac{\partial V}{\partial M} \end{bmatrix}^T \\
&= -\left(\left(\frac{\partial V}{\partial x}\right)^2 + \left(\frac{\partial V}{\partial M}\right)^2 \right) < 0
\end{aligned}
$$

olduğu görülür. Böylece 2.3.4.2 Teorem'inin şartlarını sağlayan en az bir Lyapunov fonksiyonu bulunduğundan dolayı (x^*, M^*) noktasında (4.15) durum denklem sistemi asimptotik kararlıdır denir.□

4.2.2 Test Problemleri ve Sonuçlar

4.2.2.1 Problem: Aşağıdaki eşitsizlik kısıtlamasına sahip optimizasyon problemi ele alınsın [55],

$$
\begin{aligned}
&\text{minimum} \quad f(x) = 0.5x_1^2 + x_2^2 - x_1x_2 - 7x_1 - 7x_2 \\
&\text{kısıtlama} \quad g(x) = 4x_1^2 + x_2^2 - 25 \le 0 \ .
\end{aligned} \tag{4.19}
$$

Problemin analitik çözümü $x^* = (2,3)^T$ 'dır. Buradan (4.13) denklemi yardımı ile ceza fonksiyonu ve bu fonksiyona karşılık gelen kısıtlamasız optimizasyon problemi $M \in \mathbb{R}$ olmak üzere

$$\underset{x\in\mathbb{R}^2}{\text{minimum}} \quad F(x,M) = \left(f(x)-M\right)^2 + \max\left\{0, g(x)\right\}^2$$

yazılabilir.

Buradan durum denklem sistemine karşılık gelen diferansiyel denklem sistemi ceza fonksiyonunun gradyant vektörleri yardımı ile

$$\frac{dx_1}{dt} = -2\left(x_1 - x_2 - 7\right)\left(0.5x_1^2 + x_2^2 - x_1x_2 - 7x_1 - 7x_2 - M\right) - 16x_1 \max\left\{0, 4x_1^2 + x_2^2 - 25\right\}$$
$$\frac{dx_2}{dt} = -2\left(2x_2 - x_1 - 7\right)\left(0.5x_1^2 + x_2^2 - x_1x_2 - 7x_1 - 7x_2 - M\right) - 4x_2 \max\left\{0, 4x_1^2 + x_2^2 - 25\right\}$$
$$\frac{dM}{dt} = 2\left(0.5x_1^2 + x_2^2 - x_1x_2 - 7x_1 - 7x_2 - M\right)$$

belirtilir.

Yukarıdaki durum denklem sisteminin denge noktasının araştırılmasında Euler metodu kullanılırsa $x_1(0)=1$, $x_2(0)=-1$, $M_0=-100$ penaltı parametresi ve $\nabla t = 0.0001$ adım uzunluğu olmak üzere başlangıç şartları altında Şekil 4.5'de belirtildiği gibi durum denklem sisteminin çözüm yörüngesinin (4.19) optimizasyon problemini optimum noktasına yakınsadığı görülür. Benzer biçimde farklı başlangıç noktaları içinde yakınsaklık görülür.

4.2.2.2 Problem: Eşitsizlik kısıtlamasına sahip doğrusal olmayan optimizasyon problemini ele alalım, [53];

$$\begin{aligned} &\text{minimum} \quad f(x) = 100\left(x_2 - x_1^2\right)^2 + \left(1 - x_1\right)^2 \\ &\text{kısıtlama} \quad g(x) = 0.25 - x_1^2 - x_2^2 \le 0. \end{aligned} \tag{4.20}$$

Problemin analitik optimum çözümü $x^* = (1,1)^T$'dır. Penaltı parametresi $M \in \mathbb{R}$ olmak üzere (4.13) penaltı fonksiyonu yardımı ile kısıtlamasız optimizasyon problemi

$$\text{minimum} \quad F(x,M) = \left(f(x)-M\right)^2 + \max\left\{0, g(x)\right\}^2$$

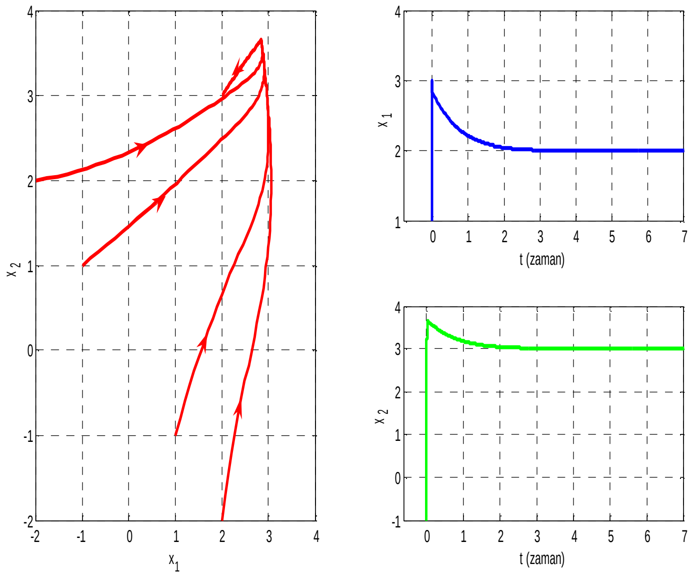
biçiminde ifade edilebilir. Buradan kısıtlamasız optimizasyon probleminin optimum noktasının araştırılması için oluşturacağımız durum denklem sistemi

$$\frac{dx_1}{dt} = \left(800x_1\left(x_2 - x_1^2\right) + 4(1 - x_1)\right)\left(100\left(x_2 - x_1^2\right)^2 + (1-x)^2 - M\right) + 4x_1 \max\left\{0, 0.25 - x_1^2 - x_2^2\right\}$$

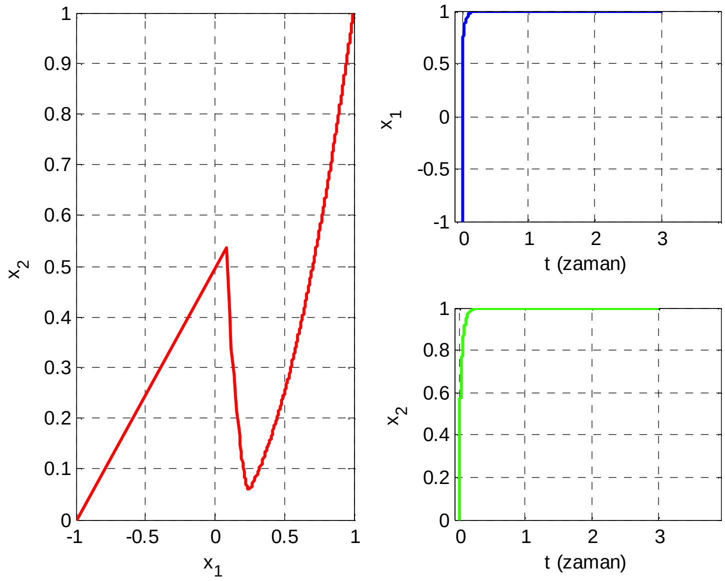
$$\frac{dx_2}{dt} = -400\left(x_2 - x_1^2\right)\left(100\left(x_2 - x_1^2\right)^2 + (1-x)^2 - M\right) + 4x_2 \max\left\{0, 0.25 - x_1^2 - x_2^2\right\}$$

$$\frac{dM}{dt} = 2\left(100\left(x_2 - x_1^2\right)^2 + (1-x)^2 - M\right)$$

şeklindedir. Bu diferansiyel denklem sisteminin çözümü içi Euler yöntemi, başlangıç şartları $x_1(0) = -1$, $x_2(0) = 0$, $M_0 = -30$ ve adım uzunluğu $\nabla t = 0.00001$ olmak üzere uygulanırsa durum denklem sisteminin denge noktasının aynı zamanda (4.20) optimizasyon probleminin optimum noktasına karşılık geldiği görülür. Şekil 4.6 incelendiğinde sistemin çözüm yörüngesinin izlediği yol görülür.



Şekil 4.5 (4.19) Probleminin x_1 ve x_2 optimum çözümleri



Şekil 4.6 (4.20) Probleminin x_1 ve x_2 optimum çözümleri

4.3 Doğrusal Olmayan Programlama Problemleri ve Düzgünleştirilmiş Ceza Fonksiyonları için Durum Denklemleri Yaklaşımı

Bu bölümde eşitsizlik kısıtlamalı doğrusal olmayan optimizasyon problemlerinin optimum çözümlerinin araştırılması için düzgünleştirilmiş ceza fonksiyonunu temel alarak durum denklemleri yaklaşımı ile ilgili sonuçlar ortaya konulacaktır.

Daha önce 2.3.1 Alt bölümünde belirttiğimiz üzere, ceza fonksiyonlarının en büyük dezavantajlarından biri tam (exact) özelliği sağlanırken fonksiyonun diferansiyellenebilme özelliğini kaybetmesidir. Buda türev tabanlı bazı çözüm metotlarının problemin optimum değerinin araştırılmasında hataya düşmesine sebep olmaktadır.

Bu gibi durumlarda belirli noktalarda diferansiyellenemeyen fonksiyonun yerine aynı özellikleri taşıyan ve her noktada diferansiyellenebilen bir yaklaşım fonksiyonu kullanılır.

Literatürde, Pınar ve Zenios [19], konveks optimizasyon problemlerinin çözümlerinde kullanılan düzgün olmayan tam ceza fonksiyonları için kuadratik düzgünleştirme yaklaşımı geliştirmişlerdir. Yine Zenios ve Pınar [20] neural network yapısında ifade edilen doğrusal olmayan optimizasyon problemleri için düzgünleştirme yaklaşımı uygulamışlardır. Ayrıca Chen ve Mangasarian [21], Yang [22] ve Meng [23] düzgünleştirme teknikleri ile diferansiyellenebilen fonksiyonlar oluşturmuşlardır.

Genel olarak eşitsizlik kısıtlamalarına sahip doğrusal olmayan optimizasyon problemi $x \in \mathbb{R}^n$ karar değişkeni, $f:\mathbb{R}^n \to \mathbb{R}$ ve $g=\left(g_1, g_2, .., g_m\right):\mathbb{R}^n \to \mathbb{R}^m$ olmak üzere

$$\begin{aligned} &\text{minimum} \quad f(x) \\ &\text{kısıtlama} \quad g_i(x) \geq 0 \quad i=1,2,...,m \end{aligned} \tag{4.21}$$

biçiminde tanımlanır. (4.21) kısıtlamalı optimizasyon problemine karşılık gelen kısıtlamasız minimumlaştırma problemi ise $p \in \mathbb{R}^+$ ceza parametresi ve $p(t)=\min\left(0,t\right)$ ceza terimi olmak üzere

$$F(x,p) = f(x) - p\sum_{i=1}^{m} p(g_i(x)) \tag{4.22}$$

diferansiyellenemeyen ceza fonksiyonu yardımı ile

$$\underset{x\in\mathbb{R}^n}{\text{minimum}} \quad F(x,p) \tag{4.23}$$

gösterilir. Ayrıca $\xi < 0$ olmak üzere $p_\xi(t)$ fonksiyonu

$$p_\xi(t) = \begin{cases} t + \dfrac{1}{2}\xi t^2 - \dfrac{1}{4}\xi & ,t \leq \xi \\ \dfrac{1}{8}\dfrac{t^3}{\xi^2} & ,\xi \leq t \leq 0 \\ 0 & ,t \geq 0 \end{cases}$$

tanımlansın. Burada dikkat edilirse $\lim\limits_{\xi \to 0} p_\xi(t) = p(t)$ olduğu görülebilir. Bu şekilde $p(t)$ diferansiyellenemeyen ceza fonksiyonu yerine aynı özelliklere sahip $p_\xi(t)$

düzgünleştirilmiş ceza fonksiyonu yaklaşımını kullanabilir. Sonuç olarak (4.21) kısıtlamalı optimizasyon problemi

$$F(x, p, \xi) = f(x) - p\sum_{i=1}^{m} p_{\xi}(g_i(x)) \tag{4.24}$$

düzgünleştirilmiş penaltı fonksiyonu yardımı ile

$$\underset{x \in \mathbb{R}^n}{\text{minimum}} \quad F(x, p, \xi) \tag{4.25}$$

kısıtlamasız optimizasyon problemi ile ifade edilebilir.

(4.25) kısıtlamasız optimizasyon probleminin optimum noktalarının araştırılmasında durum denklemleri yaklaşımını kullanmak üzere

$$\begin{aligned} &\frac{dx}{dt} = -\nabla_x F(x, p, \xi) \\ &x(t_0) = x(0) \end{aligned} \tag{4.26}$$

diferansiyel denklem sistemini ele alalım.

4.3.1 Ana Sonuçlar

4.3.1.1 Teorem: Eğer x^* noktası p ve ξ parametreleri altında (4.26) durum denklem sisteminin kararlı bir denge noktası ise (4.25) kısıtlamasız optimizasyon probleminin bir yerel minimum noktasıdır.

İspat: $x(t)$ olarak (4.26) durum denklem sisteminin bir çözümünü gösterelim. (4.26) durum denklem sisteminin x^* denge noktası kararlı olduğundan herhangi bir $x(t_0) = x(0)$ başlangıç koşulu ve $\forall \varepsilon > 0$ için $\exists \delta > 0$ vardır ki $t \geq t_0$, $\|x(0) - x^*\| < \delta$ olmak üzere $\|x(t) - x^*\| < \varepsilon$ vardır. Buradan,

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

elde edilir. Bu durum x^* noktasının (4.25) probleminin bir yerel optimum çözümü olduğunu verir.□

4.3.1.2 Teorem: Eğer x^* noktası p ve ξ parametreleri altında (4.25) probleminin bir yerel minimum noktası ise x^* noktası (4.26) durum denklem sistemi içinde kararlı bir denge noktasıdır.

İspat: x^* noktası (4.25) probleminin bir yerel optimum çözümü olduğunda 2.1.1.3 Yardımcı Teorem'in şartlarının sağlanması gerektiğinden x^* tanım gereği (4.26) durum denklem sisteminin bir denge noktasıdır. Bu denge noktasının kararlılığı göstermek için Lyapunov kararlılık teorisini kullanalım. Bunun için

$$
V(x) = F(x, p,\xi) - F(x^*, p,\xi)
$$

aday Lyapunov fonksiyonu tanımlayalım. Buradan $V(x^*) = 0$ ve $x \in \mathbb{R}^n - \{x^*\}$ için $V(x) > 0$ olduğu görülür. Dahası $V(x)$ fonksiyonunun zamana göre türevi alındığında

$$
\begin{aligned}
\frac{dV}{dt} &= \frac{\partial V}{\partial x} \frac{dx}{dt} \\
&= \left[\frac{\partial F(x, p,\xi)}{\partial x} \right] \left[-\frac{\partial F(x, p,\xi)}{\partial x} \right]^T \\
&= -\left(\frac{\partial F(x, p,\xi)}{\partial x} \right)^2 \leq 0
\end{aligned}
$$

bulunur. 2.3.4.2 Teorem gereği belirtilen şartları sağlayan bir Lyapunov fonksiyonu bulunduğundan (4.26) sistemi x^* denge noktasında kararlıdır.□

4.3.1.3 Yardımcı Teorem: $\forall x \in \mathbb{R}^n$ ve $\xi < 0$ için

$$\frac{1}{4} mp\xi \leq F(x, p, \xi) - F(x, p) \leq 0$$

sağlanır.

İspat: $p_\xi(t)$ ve $p(t)$ fonksiyonları yardımı ile

$$\frac{1}{4}\xi \leq p(t) - p_\xi(t) \leq 0$$

yazılabilir. Buradan benzer şekilde

$$\frac{1}{4}\xi \leq p(g_i(x)) - p_\xi(g_i(x)) \leq 0, \quad i = 1, 2, ..., m$$

ve

$$\frac{1}{4} m\xi \leq \sum_{i=1}^{m} p(g_i(x)) - \sum_{i=1}^{m} p_\xi(g_i(x)) \leq 0$$

elde edilir. (4.22) ve (4.24) fonksiyonlarının tanımları gereği teoremin ispatı tamamlanır.□

4.3.1.4 Teorem: Negatif sayıların bir dizisi $\{\xi_j\} \to 0$ olmak üzere x_j noktası (4.25) probleminin bir optimum çözümü olarak alalım. Eğer x^* noktası $\{x_j\}$ dizisinin bir yığılma noktası ise x^* noktası (4.23) probleminin bir optimal çözümüdür.

İspat: 4.3.1.3 Yardımcı Teorem gereği $\forall j$ için

$\frac{1}{4}mp\xi_j \le F(x,p,\xi_j) - F(x,p) \le 0$. Buradan her iki tarafın x'e göre minimumu alınırsa, varsayım gereği

$$\min_{x\in\mathbb{R}^n}\left\{\frac{1}{4}mp\xi_j\right\} \le x_j - \min_{x\in\mathbb{R}^n} F(x,p) \le 0 \qquad (4.27)$$

eşitsizliği elde edilebilir. (4.27) eşitsizliğinde $j\to\infty$ için $\{\xi_j\}\to 0$ olmak üzere

$$\{x_j\}_{j=1}^{\infty} \to \min_{x\in\mathbb{R}^n} F(x,p)$$

elde edilir. Ayrıca varsayım gereği, x^* noktasının $\{x_j\}_{j=1}^{\infty}$ dizisinin bir yığılma noktası olmasından dolayı $\{x_j\}_{j=1}^{\infty}$ dizisinin bir $\{x_{j_k}\}$ alt dizisi vardır öyle ki

$$\{x_{j_k}\} \to x^*,\ k\to\infty$$

olur. Buradan bir dizinin limitinin tekliği teoreminden

$$\min_{x\in\mathbb{R}^n} F(x,p) = x^*$$

olduğu görülür.□

4.3.1.5 Tanım: Bir $x\in\mathbb{R}^n$ noktası verilsin. Eğer $\forall i = 1,2,...,m$ için $g_i(x)\ge\xi$ oluyor ise x noktasına ξ-uygun denir.

4.3.1.6 Teorem: Eğer x^* noktası (4.23) probleminin ve $\overline{x}$ noktası da (4.25) probleminin bir optimum çözümleri ise

$$\frac{1}{4} mp\xi \leq F(\bar{x}, p, \xi) - F(x^*, p) \leq 0$$

elde edilir.

İspat: 4.3.1.3 Yardımcı Teorem ve x^*, $\bar{x}$ noktalarının optimallik koşulları gereği

$$F(\bar{x}, p, \xi) \leq F(x^*, p, \xi) \leq F(x^*, p)$$

eşitsizliğinden

$$F(\bar{x}, p, \xi) - F(x^*, p) \leq 0 \tag{4.28}$$

olduğu görülür. Benzer şekilde

$$F(\bar{x}, p) \geq F(x^*, p)$$
$$F(\bar{x}, p, \xi) - \frac{1}{4} mp\xi \geq F(\bar{x}, p) \geq F(x^*, p)$$

buradan da

$$F(\bar{x}, p, \xi) - F(x^*, p) \geq \frac{1}{4} mp\xi \tag{4.29}$$

eşitsizliği elde edilir. (4.28) ve (4.29) eşitsizliklerinden sonuç görülür.□

4.3.1.7 Teorem: x^***,** $\bar{x}$ noktaları karşılıklı olarak (4.23) ve (4.25) problemlerinin birer optimum çözümleri olarak alalım. Eğer (4.21) problemi için x^* uygun ve $\bar{x}$ ξ-uygun bir çözüm noktası ise

$$\frac{1}{2} mp\xi \leq f(\bar{x}) - f(x^*) \leq 0.$$

eşitsizliği sağlanır.

İspat: Varsayım gereği (4.21) kısıtlamalı optimizasyon problemi için $x*$ uygun bir çözüm noktası olduğundan

$$\sum_{i=1}^{m} p(g_i(x)) = 0$$

olmalıdır. Benzer şekilde $\overline{x}$ noktası da ξ-uygun bir çözüm noktası olduğundan 4.3.1.5 Tanım gereği

$$\sum_{i=1}^{m} p_{\xi}(g_i(x)) \geq \frac{1}{4} m\xi$$

bulunur. 4.3.1.6 Teorem ve (4.22), (4.24) ceza fonksiyonlarının tanımları gereği

$$\frac{1}{4} mp\xi \leq \left(f(\overline{x}) - p\sum_{i=1}^{m} p_{\xi}(g_i(\overline{x})) \right) - \left(f(x*) - p\sum_{i=1}^{m} p(g_i(x*)) \right) \leq 0$$

ve buradan da

$$\frac{1}{2} mp\xi \leq f(\overline{x}) - f(x^*) \leq 0.$$

değerlendirmesi elde edilir.□

4.3.2 Test Problemleri ve Sonuçlar

4.3.2.1 Problem: Matlab programı içerisinde yer alan eşitsizlik kısıtlamalarına sahip doğrusal olmayan programlama problemi ele alınsın,

$$\begin{aligned} &\text{minimum} \quad f(x) = \exp(x_1)\left(4x_1^2 + 2x_2^2 + 4x_1x_2 + 2x_2 + 1\right) \\ &\text{kısıtlama} \quad g_1(x) = x_1 + x_2 - x_1x_2 - 1.5 \geq 0 \\ &\qquad\qquad\; g_2(x) = x_1 + x_2 + 10 \geq 0. \end{aligned} \tag{4.30}$$

Optimizasyon probleminin nümerik çözümü $x^* = \left(-9.5474, 1.0474\right)^T$'dır. (4.30) kısıtlamalı optimizasyon probleminin $p_\xi(t)$ düzgünleştirilmiş ceza terimi kullanılarak elde edilen ceza fonksiyonu

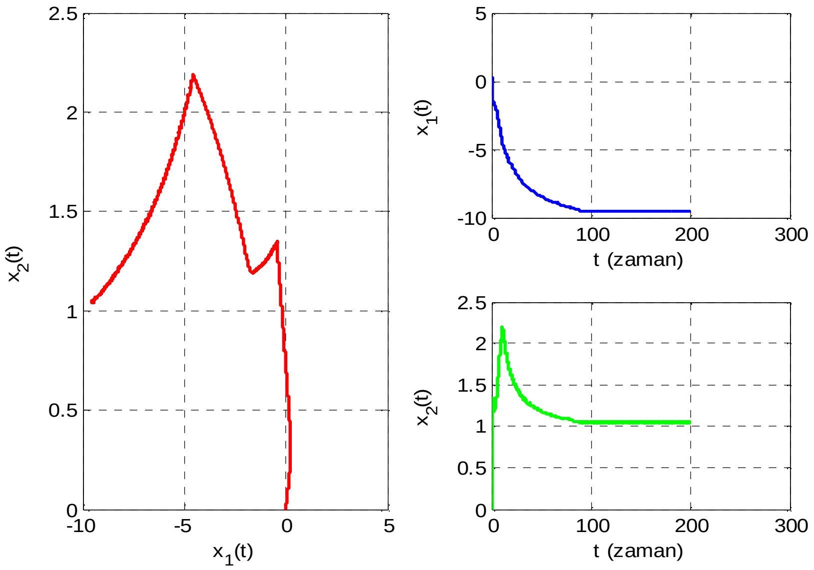
$$F(x,p,\xi) = \exp(x_1)\left(4x_1^2 + 2x_2^2 + 4x_1x_2 + 2x_2 + 1\right) -$$
$$\begin{cases} g_1(x) + \frac{1}{2}\xi\left(g_1(x)\right)^2 - \frac{1}{4}\xi & ,t \leq \xi \\ \frac{1}{8}\frac{\left(g_1(x)\right)^3}{\xi^2} & ,\xi \leq t \leq 0 \\ 0 & ,t \geq 0 \end{cases} -$$
$$\begin{cases} g_2(x) + \frac{1}{2}\xi\left(g_2(x)\right)^2 - \frac{1}{4}\xi & ,t \leq \xi \\ \frac{1}{8}\frac{\left(g_2(x)\right)^3}{\xi^2} & ,\xi \leq t \leq 0 \\ 0 & ,t \geq 0 \end{cases} -$$

şeklindedir. (4.26) durum denklem sistemi ise

$$\begin{aligned} \frac{dx_1}{dt} &= -\exp(x_1)\left(4x_1^2 + 2x_2^2 + 4x_1x_2 + 2x_2 + 1\right) - \exp(x_1)\left(8x_1 + 4x_2\right) + \\ &\quad p\nabla_{x_1} p_\xi(g_1) + p\nabla_{x_1} p_\xi(g_2) \\ \frac{dx_2}{dt} &= -\exp(x_1)\left(4x_1 + 4x_2 + 2\right) + p\nabla_{x_2} p_\xi(g_1) + p\nabla_{x_2} p_\xi(g_2) \end{aligned}$$

diferansiyel denklemleri ile ifade edilebilir. Sistemin nümerik çözümü için Euler yöntemi, başlangıç şartları $x_1(0) = 0$, $x_2(0) = 0$, $p = 10$ ve $\xi = -0.0001$ ve

$dt = 0.0001$ basamak uzunluğu altında uygulanırsa simülasyon sonuçları durum denklem sisteminin denge noktasının $(-9.5286, 1.0488)$ olduğu görülür.



Şekil 4.7 (4.30) Probleminin x_1 ve x_2 optimum çözümleri

4.3.2.2 Problem: Durum denklemleri yaklaşımını uygulamak için test problemi Schittkowski [53] kaynağından alınmıştır,

$$\begin{aligned} &\text{minimum} \quad f(x) = 100\left(x_2 - x_1^2\right)^2 + \left(1 - x_1^2\right)^2 \\ &\text{kısıtlama} \quad g_1(x) = x_1^2 + x_2^2 - 0.25 \geq 0. \end{aligned} \tag{4.31}$$

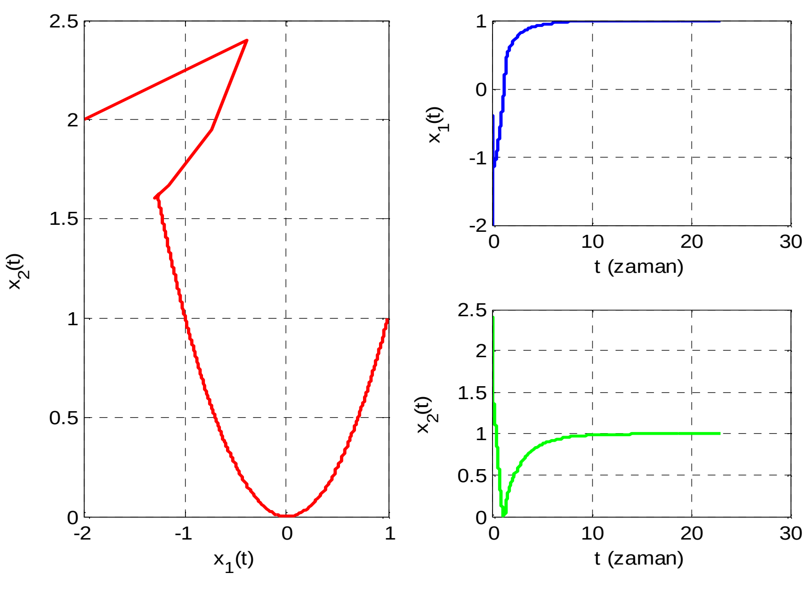
Problemin optimum çözümü $x^* = (1,1)^T$ noktasında yer almaktadır. $p_\xi(t)$ fonksiyonu yardımı ile kısıtlamasız optimizasyon problemi

$$
\text{minimum} \quad 100\left(x_2 - x_1^2\right)^2 + \left(1 - x_1^2\right)^2 - \begin{cases} g_1(x) + \dfrac{1}{2}\xi\left(g_1(x)\right)^2 - \dfrac{1}{4}\xi & ,t \le \xi \\ \dfrac{1}{8}\dfrac{\left(g_1(x)\right)^3}{\xi^2} & ,\xi \le t \le 0 \\ 0 & ,t \ge 0 \end{cases}
$$

biçiminde ifade edilebilir. (4.26) diferansiyel denklem sisteminin tanımı gereği durum denklem sistemi

$$
\begin{aligned}
\frac{dx_1}{dt} &= 400x_1\left(x_2 - x_1^2\right) + 2\left(1 - x_1\right) + p\nabla_{x_1} p_\xi\left(g_1\right) \\
\frac{dx_2}{dt} &= -200\left(x_2 - x_1^2\right) + p\nabla_{x_2} p_\xi\left(g_1\right)
\end{aligned}
\tag{4.32}
$$

yazılabilir. (4.31) diferansiyel denklem sisteminin denge noktasının araştırılmasında Euler yaklaşımını kullanalım. Başlangıç koşulları $x_1(0) = -2$, $x_2(0) = 2$, $p = 1$ ve $\xi = -0.0001$ ve $dt = 0.001$ basamak uzunluğu alındığında (4.32) durum denklem sisteminin çözüm yörüngesinin $(1,1)$ noktasına yakınsadığı görülür.



Şekil 4.8 (4.31) Probleminin x_1 ve x_2 optimum çözümleri

5. SONUÇLAR

Bu çalışmada elde edilen yeni sonuçlar üçüncü ve dördüncü bölümlerde bulunmaktadır ve bunlar aşağıdaki gibi özetlenebilir.

Üçüncü bölümde, kuadratik kısıtlamasız optimizasyon problemlerinin mutlak minimumlarının araştırılmasında özel olarak tanımlanan diferansiyel denklem sistemlerinden oluşan durum denklemleri yaklaşımı tanıtılmıştır. Daha sonra verilen teoremler ile optimizasyon probleminin mutlak minimumunun aynı zamanda bu problemden türetilen durum denklem sisteminin denge noktasına karşılıklı olarak denk geldiği ve bu denge noktasının aynı zamanda asimptotik kararlı olduğu ispatlanmıştır. Diğer yandan asıl sistemin kararlılığının, denge noktası orijine ötelenmiş durum denklem sisteminin kararlılığına bağlı olduğu gösterilmiştir. Elde edilen çözümler bu amaca yönelik MATLAB komutları kullanılarak Quasi-Newton Metodu ile karşılaştırılmıştır.

Dördüncü bölüm üç kısımdan oluşmaktadır. İlk kısımda, ceza fonksiyonları yardımı ile kuadratik programlama problemlerinin optimum çözümlerinin durum denklem sisteminin asimptotik kararlı denge noktasına karşılık geldiği teoremler ile verilmiştir. Yine elde edilen çözümler MATLAB komutları kullanılarak SQP Metodu ile karşılaştırılmıştır. İkinci kısımda ise durum denklemleri yaklaşımı doğrusal olmayan optimizasyon problemlerine genişletilerek çözümlerin denkliği gösterilmiştir. Ayrıca Lyapunov kararlılık teoremi yardımı ile dinamik sistemin denge noktasının asimptotik kararlı olduğu ispatlanmıştır. Daha sonra durum denklemleri yaklaşımının etkinliği test problemleri ile verilmiştir. Son kısımda, diferansiyellenmeyen l_1 ceza fonksiyonu için düzgünleştirilmiş bir ceza fonksiyonu tanımlanarak verilen tanım ve teoremler ile aralarındaki ilişkiler ortaya konulmuştur. Ayrıca bu fonksiyon yardımı ile oluşturulan durum denklemleri yaklaşımının denge noktasının kararlı olduğu ve bu noktanın aynı zamanda asıl optimizasyon probleminin optimum noktasına karşılık geldiği ispatlanmıştır. Yaklaşım sonuçları MATLAB komutları kullanılarak test problemleri üzerinde incelenmiştir.

KAYNAKLAR

[1] Bertsekas, D.P., Constrained Optimization and Lagrange Multiplier Methods, Academic Pres, New York, (1982).

[2] Chong, E.K.P. ve Zak, S.H., An Introduction to Optimization, Wiley-Intersicence, New York, (2001).

[3] Fiacco, A.V. ve Mccormick, G.P., Nonlinear Programming: Sequential Unconstrained Minimization Techniques, John Wiley, New York, (1968).

[4] Luenberger, D.G. ve Ye, Y., Linear and Nonlinear Programming, Third Edition, Springer, New York, (2007).

[5] Nash, S.G., Sofer, A., Linear and Nonlinear Programming, McGraw-Hill Book Co., New York, (1996).

[6] Sun, W. ve Yuan, Y.X., Optimization Theory and Methods: Nonlinear Programming, Springer-Verlag, New York, (2006).

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

[8] Morrison, D.D., "Optimization by least squares", SIAM Journal Numerical Analysis, 5, (1968), 83-88.

[9] Conn, A.R., "Constrained optimization using a nondifferentiable penalty function", SIAM Journal Numerical Analysis, 10, (1973), 102-107.

[10] Broyden, C.G. ve Attia, N.F., A Smooth Squential Penalty Function Method for Solving Nonlinear Problem, System Modelling and Optimization, Springer-Verlag, Berlin, (1983), 237-245.

[11] Broyden, C.G. ve Attia, N.F., "Penalty functions, Newton method and quadratic programming", J. Optim. Theory Appl., 58, (1988), 285-309.

[12] Meng, Z., Hu, Q., Dang, C. ve Yang, X.Q., "An objective penalty function method for nonlinear programming", Appl. Math. Lett., 17, (2004), 683-689.

[13] Rubinov, A.M., Glover, B.M. ve Yang, X.Q., “Extended lagrange and penalty functions in continuous optimization”, Optimization, 46, (1999), 327-351.

[14] Rubinov, A.M., Glover, B.M. ve Yang, X.Q., “Decreasing functions with applications to penalization”, SIAM Journal on Optimization, 10, (1999), 289-313.

[15] Yang, X.Q. ve Huang, X.X., “A nonlinear lagrangian approach to constrained optimization problems”, SIAM J. Optim., 11, (2001), 1119-1144.

[16] Han, S.P. ve Mangasarian, O.L., “Exact penalty function in nonlinear programming”, Mathematical Programming, 13, (1979), 251-269.

[17] Di Pillo, G. Ve Grippo, L., “An exact penalty function method with global convergence properties for nonlinear programming problems” Mathematical Programming, 36, (1986), 1-18.

[18] Di Pillo, G. Ve Grippo, L., “On the exactness of a class of nondifferentiable penalty function”, Journal of Optimzation Theory and Applications, 57, (1998), 399-410.

[19] Pınar, M.C. ve Zenios, S.A., “On smoothing exact penalty functions for convex constrained optimization”, SIAM, J. Optim., 4, (1994), 486-511.

[20] Zenios, S.A., Pınar, M.C. ve Dempo, R.S., “A smooth penalty function algorithm for network-structured problems”, European J. Oper. Res., 83, (1995), 220-236.

[21] Chen, C. ve Mangasarian, O.L., “Smoothing methods for convex inequalities and linear complementarity problem”, Math. Program., 71, (1995), 51-69.

[22] Yang, X.Q., Meng, Z.Q., Huang, X.X. ve Pong, G.T.Y., “Smoothing nonlinear penalty functions for constrained optimization problems”, Numer. Funct. Anal. Optim., 24, (2003), 351-364.

[23] Meng, Z., Dang, C. ve Yang, X.Q., “On the smoothing of square-root exact penalty function for inequality constrained optimization”, Comput. Optim. Appl., 35, (2006), 375-398.

[24] Arrow, K.J. ve Hurwicz, L., “Reduction of constrained maxima to saddle point problems”, Proceedings of 3rd Berkeley Symposium on Mathematical Statistics and Probability, Berkeley, (1956), 1-26.

[25] Arrow, K.J, Hurwicz, C ve Uzawa, H., Studies in Linear and Non-Linear Programming, Stanford University Pres, California, (1958).

[26] Pan, P.Q., "New ODE methods for equality constrained optimization 1. Equations", J. Comput. Math., 10, (1992), 77-92.

[27] Jin, L. ve Zhang, L-W, "Two differential equation systems for equality-constrained optimization", Appl. Math. Comput., 190, (2007), 1030-1039.

[28] Jin, L., Zhang, L-W. ve Xiao, X.T., "Two differential equation systems for inequality constrained optimization", Appl. Math. Comput., 188, (2007), 1334-1343.

[29] Brown, A.A, Bartholomew-Biggs, M.C., "Some effective methods for unconstrained optimization based on the solution of systems of ordinary differential equations", J. Optim. Theory Appl., 62, (1989), 211-224.

[30] Brown, A.A, Bartholomew-Biggs, M.C., "ODE versus SQP methods for constrained optimization", J. Optim. Theory Appl., 62, (1989), 371-385.

[31] Evtushenko, Y.G. ve Zhadan, V.G., "Stable barrier-projection and barrier Newton methods in nonlinear programming", Optim. Methods Softw., 3, (1994), 237-256.

[32] Evtushenko, Y.G. ve Zhadan, V.G., "Stable barrier-projection and barrier Newton methods in linear programming", Comput. Optim. Appl., 3, (1994), 289-303.

[33] Wang, S., Yang, X.Q. ve Teo, K.L., "A unified gradient flow approach to constrained nonlinear optimization problems", Comput. Optim. Appl., 25, (2003), 251-268.

[34] Anderi, N., "Gradient flow algorithm for unconstrained optimization", ICI Technical Report (2004).

[35] Schropp, J., "A dynamical systems approach to constrained minimization", Numer. Funct. Anal. Optim., 21, (2000), 537-551.

[36] Goh, B.S, "Algorithms for unconstrained optimization problems via control theory", J. Optim. Theory Appl., 92, (1997), 581-604.

[37] Özdemir, N., Yaman, R., Evirgen, F., Yaman, G., "State space approach for unconstrained quadratic optimization problems", Int. J. Pure Appl. Math. Sci., yayına kabul edildi.

[38] Özdemir, N., Evirgen, F., "A dynamic system approach to quadratic programming problems with penalty method", Bull. Malays. Math. Sci. Soc., yayına kabul edildi.

[39] Özdemir, N., Evirgen, F., "A dynamic system approach for solving nonlinear programming problems with exact penalty function", Proceeding of the 20th EURO Mini Conference "Continuous Optimization and Knowledge-Based Technologies", Lithuania, (2008), 82-86.

[40] Özdemir, N., Evirgen, F., "Solving NLP problems with dynamic system approach based on smoothed penalty function", Selçuk J. Appl. Math., 10 (2009), 63-73.

[41] Horn, R.A. ve Johnson, C.R., Matrix Analysis, Cambridge University Press, New York, (1985).

[42] Nocedal J. ve Wright S.J., Numerical Optimization, Springer-Verlag, New York, (1999).

[43] Bazaraa, M.S., Sherali, H.D. ve Shetty, C.M., Nonlinear Programming: Theory and Algorithms, Wiley Interscience, New Jersey, (2006**).**

[44] Ogata, K., State Space Analysis of Control Systems, Prentice-Hall, New Jersey, (1967).

[45] Ogata, K., System Dynamics, Second Edition, Prentice-Hall, New Jersey, (1992).

[46] Chen, B.M., Lin, Z. ve Shamash, Y., Linear Systems Theory: A Structural Decomposition Approach, Birkhäuser, Boston, (2004).

[47] Khalil, H.K., Nonlinear Systems, Second Edition, Prentice-Hall, New Jersey, (1996)

[48] Kwarkernaak, H. ve Sıvan, R., Linear Optimal Control Systems, Wiley-Interscience, New York, (1972).

[49] Curtain, R.F. ve Pritchard, A.J., Functional Analysis in Modern Applied Mathematics, Academic Pres, London, (1977).

[50] Chen, C.T., Linear System Theory and Design, Oxford University, New York, (1999).

[51] Williams, R.L. ve Lawrence, D.A., Linear State-Space Control Systems, John Wiley&Sons Inc, New Jersey, (2007).

[52] La Salle, J.L. ve Lefschetz, S., Stability by Liapunov's Direct Method with Application, Academic Pres, New York, (1961).

[53] Schittkowski, K., More Test Examples for Nonlinear Programming Codes, Springer, Berlin, (1987).

[54] Liu, S. ve Meng, Z., "A new nonlinear neural Networks based on exact penalty function with two-parameter", Proceeding of the Second International Conference on Machine Learning and Cybernetics, 2 (2-5), (2003), 1249-1254.

[55] Hock, W. ve Schittkowski, K., Test Examples for Nonlinear Programming Codes, Springer, Berlin, (1981).