

T.C.
YILDIZ TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

BÜYÜK ÖLÇEKLİ RICCATI DENKLEMLERİ İÇİN KRYLOV
YÖNTEMLERİ

Yaprak DERİCİOĞLU

DOKTORA TEZİ
Matematik Anabilim Dalı
Matematik Programı

Danışman
Doç. Dr. Muhammet KURULAY

Eş-Danışman
Prof. Dr. Khalide JBILOU

Eylül, 2019

T.C.
YILDIZ TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

BÜYÜK ÖLÇEKLİ RICCATI DENKLEMLERİ İÇİN KRYLOV
YÖNTEMLERİ

Yaprak DERİCİOĞLU tarafından hazırlanan tez çalışması 30.09.2019 tarihinde aşağıdaki jüri tarafından Yıldız Teknik Üniversitesi Fen Bilimleri Enstitüsü Matematik Anabilim Dalı Matematik Programı **DOKTORA TEZİ** olarak kabul edilmiştir.

Doç. Dr. Muhammet KURULAY
Yıldız Teknik Üniversitesi
Danışman

Prof. Dr. Khalide JBILLOU
Université du Littoral Côte d'Opale
Eş-Danışman

Jüri Üyeleri

Doç. Dr. Muhammet KURULAY, Danışman
Yıldız Teknik Üniversitesi

Prof. Dr. Nuran GÜZEL, Üye
Yıldız Teknik Üniversitesi

Prof. Dr. Coşkun YAKAR, Üye
Gebze Teknik Üniversitesi

Prof. Dr. Samet Yücel KADIOĞLU, Üye
İstanbul Teknik Üniversitesi

Doç. Dr. Özgür YILDIRIM, Üye
Yıldız Teknik Üniversitesi

Danışmanım Doç. Dr. Muhammet KURULAY sorumluluğunda tarafımca hazırlanan Büyük Ölçekli Riccati Denklemleri için Krylov Yöntemleri başlıklı çalışmada veri toplama ve veri kullanımında gerekli yasal izinleri aldığımı, diğer kaynaklardan aldığım bilgileri ana metin ve referanslarda eksiksiz gösterdiğimi, araştırma verilerine ve sonuçlarına ilişkin çarpıtma ve/veya sahtecilik yapmadığımı, çalışmam süresince bilimsel araştırma ve etik ilkelerine uygun davrandığımı beyan ederim. Beyanımın aksinin ispatı halinde her türlü yasal sonucu kabul ederim.

Yaprak DERİCİOĞLU

İmza



Bu çalışma, Yıldız Teknik Üniversitesi Bilimsel Araştırma Projeleri Koordinatörlüğü'nün 2016-07-03-DOP04 numaralı projesi ile desteklenmiştir.

Harun ve Zeynep Ela'ya



TEŞEKKÜR

Doktora çalışmalarımın her aşamasında desteğini benden esirgemeyen, tecrübeleri ile beni aydınlatan ve her zaman örnek olan değerli danışman hocam sayın Doç. Dr. Muhammet KURULAY'a ve tez konusu seçiminde bana öncülük eden, bu süreçte beni destekleyen ve öğrencisi olmaktan onur duyduğum eş danışmanım sayın Prof. Dr. Khalide JBILOU'ya en derin saygılarımla teşekkür ederim. Binlerce kilometre uzakta olmasına rağmen doktora çalışmalarım boyunca bilgi ve deneyimiyle her zaman bana destek olan, çalışma azmi ve isteğiyle beni teşvik eden ve bu süreçte yalnız bırakmayan değerli hocam sayın Dr. Mustapha HACHED'a yürekten teşekkür ederim.

Değerli aileme; hayatım boyunca bana destek olan, eğitimim için her türlü fedakârlıkta bulunan bugünlere gelmeme vesile olan canım annem, canım babam ve canım kardeşim Burak'a bana olan desteğini ve güvenini esirgemeyen, çalışmalarım boyunca göstermiş olduğu ilgi ve anlayış için sevgili eşim Harun'a ve beni sabırla bekleyen güzel kızım Zeynep Ela'ya sonsuz sevgilerimi ve teşekkürlerimi sunuyorum.

Yaprak DERİCİOĞLU

İÇİNDEKİLER

SİMGE LİSTESİ	ix
KISALTMA LİSTESİ	x
ŞEKİL LİSTESİ	xi
TABLO LİSTESİ	xii
ÖZET	xiii
ABSTRACT	xiv
1 Giriş	1
1.1 Literatür Özeti	1
1.2 Tezin Amacı	4
1.3 Bulgular	5
2 Matris Diferansiyel Denklemlere Giriş	6
2.1 Lineer Zamanla Değişmeyen Sistemler	6
2.2 Sonlu Horizontda Lineer Kuadratik Regülatör Problemi	7
2.3 Model Problemler	10
2.3.1 Sonlu Farklar Yöntemi: Yarı-Ayrıklaştırılmış Isı Denklemi	10
2.3.2 Kimyasal Reaktörler: Akan Reaktifin Sıcaklığını Kontrol Etme .	11
2.3.3 Çelik Profillerin Seçici Soğutulması: Bir Demiryolu Haddehanesinde Soğutma	12
3 Krylov Altuzay ve Geri Fark Formülü (BDF) Yöntemleri	14
3.1 Genel İzdüşüm Yöntemleri	14
3.1.1 Eğik İzdüşümler	14
3.1.2 Ortogonal İzdüşümler	15
3.2 Krylov Altuzay Yöntemleri	16
3.2.1 Arnoldi Yöntemi	17
3.3 Blok Krylov ve Genişletilmiş Krylov Altuzay Yöntemleri	19
3.3.1 Blok Krylov Altuzay Yöntemleri	19

3.3.2	Genişletilmiş Krylov Altuzay Yöntemi	22
3.4	BDF Yöntemi	26
4	Büyük Ölçekli Matris Riccati Diferansiyel ve Simetrik Stein Diferansiyel Denklemleri için Düşük Ranklı Yaklaşık Çözümler	28
4.1	Büyük Ölçekli Matris Riccati Diferansiyel Denklemleri	28
4.1.1	Matris Riccati Diferansiyel Denklemi Çözümü için BDF Yöntemi	29
4.1.2	Düşük Ranklı Matris Riccati Diferansiyel Denklemi'nin İzdüşümü ve Çözümü	32
4.1.3	İzdüşürülmüş Lineer Kuadratik Regülatör Problemi	36
4.2	Büyük Ölçekli Matris Simetrik Stein Diferansiyel Denklemleri	38
4.2.1	Matris Simetrik Stein Diferansiyel Denklemi'nin İzdüşümü ve Çözümü	39
4.2.2	Düşük Boyutlu Simetrik Stein Diferansiyel Denklemi'nin Çözümü	42
5	Simetrik Olmayan Matris Stein Diferansiyel Denklemleri	44
5.1	Düşük Boyutlu Simetrik Olmayan Stein Diferansiyel Denkleminin Çözümü	46
6	Sayısal Örnekler	48
6.1	Büyük Ölçekli Matris Riccati Diferansiyel Denklemi için Sayısal Örnekler	48
6.2	Büyük Ölçekli Matris Simetrik Stein Diferansiyel Denklemi için Sayısal Örnekler	54
6.3	Simetrik Olmayan Matris Stein Diferansiyel Denklemleri için Sayısal Örnekler	57
7	Sonuç ve Öneriler	60
	Kaynakça	61
A	Temel Tanım ve Kavramlar	65
A.1	Matris Çarpımları	65
A.1.1	$\otimes$ - Kronocker Çarpım	65
A.1.2	$\diamond$ -Çarpım	65
A.2	Matris Normu ve Özellikleri	66
A.2.1	Matris p ve q Normu	66
A.2.2	Frobenius Norm	66
A.3	Matris Ayrışimleri	67
A.3.1	Tekil Değer Ayrışımı	67
A.3.2	QR Ayrışımı	67
A.3.3	Global QR Ayrışımı	67



SİMGE LİSTESİ

p_A	A matrisinin karakteristik polinomu
h	Adım boyu
$\diamond$	$\diamond$ -çarpım
tol	Doğruluk sınırı
$\mathbb{C}^{m \times n}$	Elemanları kompleks sayılar olan $m \times n$ boyutlu matrislerin kümesi
$\mathbb{R}^{n \times n}$	Elemanları reel sayılar olan $n \times n$ boyutlu matrislerin kümesi
$\mathbf{0}_{n \times 2}$	Elemanları sıfır olan $n \times 2$ boyutlu matris
DF	Fréchet türev operatörü
$\langle, \rangle_F$	Frobenius iç çarpım
$\ \cdot\ _F$	Frobenius norm
$\otimes$	Kronecker çarpım
$R_m(t)$	m . adımdaki Rezidü operatörü
$\mathbb{K}_m(A, V)$	m boyutlu blok Krylov altuzayı
$\mathcal{K}_m^e(A, V)$	m boyutlu genişletilmiş Krylov altuzayı
$K_m(A, V)$	m boyutlu Krylov altuzayı
$\frac{\partial^n}{\partial x^n}$	n . mertebeden tamsayı kısmi türev operatörü

KISALTMA LİSTESİ

BDF(p)	p Adım Geri Fark Formülü Yöntemi
care	Matlab Cebirsel Riccati Denklem Çözücüsü
EBA	Genişletilmiş Blok Arnoldi Yöntemi
EGA	Genişletilmiş Global Arnoldi Yöntemi
LTI	Lineer Zamanla Değişmeyen Sistem
LR-ADI	Düşük Ranklı Değişen Yönlü Kapalı Formülasyon
LRSOP	Düşük Ranklı Bölünme Yöntemi
LQR	Lineer Kuadratik Regülatör
MSSDE	Matris Simetrik Stein Diferansiyel Denklemi
MRDE	Matris Riccati Diferansiyel Denklemi
NMSDE	Simetrik olmayan Matris Stein Diferansiyel Denklemi
ode23s	Matlab Stiff Diferansiyel Denklem Çözücüsü

ŞEKİL LİSTESİ

Şekil 2.1	İçe akış için Ω alanı: dairesel bir tüpte $t = 2$ için içe akışın 2-boyutlu dik kesiti	11
Şekil 2.2	Çelik Profillerin Seçici Soğutulması için Etki Alanı: en aza indirme ve karşılaştırma noktalarına sahip ilk ağ (solda) ve sınırın bölünmesi (sağda)	13
Şekil 6.1	$t \in [0, 1]$ zaman aralığı için $X_{1,1}(t)$ değerleri	49
Şekil 6.2	m genişletilmiş blok Arnoldi iterasyon sayısı için $\ X_{EBA}(T_f) - X_{ode23s}(T_f)\ _F$ hata normu	50
Şekil 6.3	m genişletilmiş blok Arnoldi iterasyon sayısına karşılık gelen $\ R_m(T_f)\ $ rezidü normu	51
Şekil 6.4	$\text{size}(A) = 49 \times 49$ ve $t \in [0, 2]$ için hesaplanan $X_{1,1}(t)$ bileşenleri ode23s ve EBA-BDF(2)(solda), ode23s ve BDF(2)+Newton+EBA (sağda)	52
Şekil 6.5	$\text{size}(A) = 400 \times 400$ ve $t \in [0, 50]$ için $\ X_{EBA-BDF}(t) - X_\infty\ $ hata normları	53
Şekil 6.6	$t \in [0, 2]$ zaman aralığı için $X_{1,1}$ değerleri	56
Şekil 6.7	$(n = 6400)$ için rezidü değerleri	56
Şekil 6.8	Çelik profillerin optimal soğutma problemi ($n = 5177$), $t \in [0, 20]$ zaman aralığı için $X_{1,1}(t)$ değerleri	57
Şekil 6.9	$t \in [0, 1]$ zaman aralığı için $X_{1,1}(t)$ değerleri	58
Şekil 6.10	m genişletilmiş global Arnoldi iterasyon sayısına karşılık gelen $\ R_m(T_f)\ $ rezidü normu	59

TABLO LİSTESİ

Tablo 3.1	p adım BDF yöntemi katsayıları $p \leq 5$	27
Tablo 6.1	ode23s,EBA-BDF(2), BDF(2)+Newton+EBA(2) için hesaplama zamanları	50
Tablo 6.2	EBA-BDF(2) yöntemi için hesaplama zamanları, rezidü normları ve genişletilmiş Arnoldi iterasyonu sayısı m	51
Tablo 6.3	EBA-BDF(2) yöntemi hesaplama zamanları(s), rezidü normları ve Arnoldi iterasyonları sayısı(m)	53
Tablo 6.4	Çelik profillerin optimal soğutma problemi için hesaplama zamanları(s)	54
Tablo 6.5	A 'nın boyutu, hesaplama zamanları(s), rezidü normu ve Arnoldi iterasyonları sayısı(m)	55
Tablo 6.6	A 'nın boyutu, hesaplama zamanları(s), rezidü normu ve Arnoldi iterasyonları sayısı(m)	57
Tablo 6.7	ode23s,EGA-BDF(1) için hesaplama zamanları	59

Büyük Ölçekli Riccati Denklemleri için Krylov Yöntemleri

Yaprak DERİCİOĞLU

Matematik Anabilim Dalı

Doktora Tezi

Danışman: Doç. Dr. Muhammet KURULAY

Eş-Danışman: Prof. Dr. Khalide JBILOU

Bu tezde, büyük ölçekli matris diferansiyel denklemlerin Krylov altuzaylarının bir dizisine izdüşürülerek elde edilen düşük ranklı problemin BDF(Backward Differentiation Formula) yöntemi kullanılarak çözümünü içeren yeni bir yaklaşım tanıtıldı. Bu yaklaşım, kısmi diferansiyel denklemlerin ve lineer kontrolün sayısal çözümünde yer alan büyük ölçekli matris diferansiyel Riccati, simetrik ve simetrik olmayan diferansiyel Stein denklemleri için incelendi ve elde edilen sonuçlar problemlerin düşük ranklı yaklaşık çözümlerini oluşturmak için kullanıldı. Bu denklem tipleri için, problemin yaklaşık çözümünü hesaplamadan, yaklaşımdaki hatanın hesaplanabilir olduğu görüldü ve bu hatanın büyük matrislerin çarpımlarını hesaplama gereksinimi olmadan basit bir ifadesini veren bir teorik sonucu verildi. Problemlerin yaklaşık çözümleri MATLAB programı kullanılarak verildi ve bu yaklaşımın diğer yöntemlere göre performansı karşılaştırmalı testlerle doğrulandı.

Anahtar Kelimeler: Genişletilmiş blok Krylov, genişletilmiş global Krylov, düşük ranklı yaklaşım, matris diferansiyel Riccati denklemi, diferansiyel simetrik Stein denklemi, simetrik olmayan diferansiyel Stein denklemi

Krylov Methods for Large Scale Riccati Equations

Yaprak DERİCİOĞLU

Department of Mathematics

Doctor of Philosophy Thesis

Advisor: Assoc. Prof. Dr. Muhammet KURULAY

Co-advisor: Prof. Dr. Khalide JBILOU

In this thesis, a new approach that consisting in projecting a large-scale differential problem onto a sequence of extended block Krylov subspaces and solving the projected matrix differential equation by a backward differentiation formula is proposed. This approach is examined for the large-scale matrix differential Riccati, symmetric and nonsymmetric Stein equations involved in the numerical solution of partial differential equations and linear control, and the results are used to build a low-rank approximate solutions of the problem. For these type of equations, it is seen that the error in the approach can be computable without calculating the approximate solution of the problem and a theoretical result is given which give a simple expression without the need to calculate the multiplication of the large matrices. Approximate solutions of the problems are given by using MATLAB program and the performance of this approach with regard to other methods is confirmed by comparative tests.

Keywords: Extended block Krylov, extended global Krylov, low rank approximation, differential matrix Riccati equations, differential symmetric Stein equations, differential nonsymmetric Stein equations.

1.1 Literatür Özeti

Büyük ölçekli (large-scale) matris diferansiyel denklemler; kısmi diferansiyel denklemlerin ayrıştırılması, büyük matrislerin özdeğer problemleri, lineer zamanla değişmeyen dinamik sistemler (linear time invariant dynamical system (LTI)) ve büyük ölçekli lineer ve lineer olmayan optimizasyon problemleri gibi pek çok uygulamada ortaya çıkmaktadır [1–4]. Büyük ölçekli matris diferansiyel denklemlerin nümerik olarak çözümü, nümerik analizde ve birçok uygulamada önem taşıyan güncel bir problemdir. Yaygın olarak kullanılan matris diferansiyel denklemlerden bazıları matris Riccati diferansiyel denklemleri, matris Lyapunov diferansiyel denklemleridir. Matris Riccati ve Lyapunov diferansiyel denklemleri daha çok lineer kuadratik regülatör (linear quadratic regulator) ve lineer kuadratik Gauss tasarımı, lineer zamanla değişen sistemlerin H_∞ kontrolü, optimal filtreleme, diferansiyel oyunlar, lineer zamanla değişen sistemlerde model indirgeme ve kontrol problemlerinde ortaya çıkar [4–7] ve bu denklemler sırasıyla aşağıdaki şekilde tanımlıdır.

Matris Riccati diferansiyel denklemi (MRDE),

$$\begin{cases} \dot{X}(t) = X(t)A + A^T X(t) - X(t)B B^T X(t) + C^T C, \\ X(t_0) = X_0, \end{cases}, t \in [t_0, T_f] \quad (1.1)$$

şeklindedir. Burada $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times l}$, $C \in \mathbb{R}^{s \times n}$ ve $X(t) \in \mathbb{R}^{n \times n}$ dir.

Matris Lyapunov diferansiyel denklemi,

$$\begin{cases} \dot{X}(t) = AX(t) + X(t)A^T - BB^T, \\ X(t_0) = X_0, \end{cases}, t \in [t_0, T_f] \quad (1.2)$$

şeklinde verilir. Burada $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times s}$, $l, s \ll n$ ve $X \in \mathbb{R}^{n \times n}$ dir. Ayırık zaman Lyapunov denklemi, matris simetrik Stein diferansiyel denklemi (MSSDE)

olarak bilinir ve bu denklem

$$\begin{cases} \dot{X}(t) = X(t) - AX(t)A^T - BB^T, \\ X(t_0) = X_0, \end{cases}, t \in [t_0, T_f] \quad (1.3)$$

şeklinde verilir.

Simetrik olmayan matris Stein diferansiyel denklemi (NMSDE) ise

$$\begin{cases} \dot{X}(t) = X(t) - AX(t)B + EF^T, \\ X(t_0) = X_0, \end{cases}, t \in [t_0, T_f] \quad (1.4)$$

şeklinde verilir ve X_0 , $n \times p$ tipinde verilen bir matris, A ve B matrisleri seyrek ve reel matrisler olup sırasıyla $n \times n$ ve $p \times p$, $s \ll n, p$, tipinde kare matrislerdir. E ve F matrisleri de sırasıyla $n \times s$ ve $p \times s$ matrislerdir.

Katsayı matrisleri için eğer $n \geq 1000$ ise bu durumda denkleme büyük ölçekli diferansiyel denklem denir.

Katsayı matrisleri, belirli bir yapıya sahiptir ve genellikle büyük ölçekli olup elemanlarının çoğu sıfırdır yani seyrek. Genel olarak ortaya çıkan diferansiyel denklemlerin nümerik çözümleri kararsızdır ve bu durumu verimli bir şekilde ele alan yöntemler gerektirir. Matris diferansiyel denklemlerin nümerik olarak çözümü, klasik yöntemler (Jacobi, Gauss-Seidel) ve Krylov altuzay yöntemleri gibi iteratif yöntemlerle mümkündür.

Bu tezde büyük ölçekli simetrik MRDE ve ayrık zaman Lyapunov denklemleri olarak da bilinen büyük ölçekli simetrik Stein (MSSDE) ve simetrik olmayan Stein diferansiyel denklemleri (NMSDE) ele alınacaktır.

MRDE ile ilgili literatürde 1970'li yıllardan itibaren birçok çalışma bulunmaktadır ve bu denklemin çözümü için birçok yöntem ileri sürülmüştür. Bu yöntemlerden bazıları aşağıda özetlenmiştir.

W.M. Wonham 1968'de (1.1) denkleminin çözümünün varlık ve teklik teoremini vermiştir [8]. P.B. Molinari 1972 yılında bu çözümlerin kararlılığını araştırmıştır [9]. D.R. Vaughan [10], E.J. Davison ve M.C. Maki [11] ve C.S. Kenney ve R.B. Leipnik [12] verilen denklemi birinci mertebeden lineer matris diferansiyel denkleme dönüştürmüşlerdir ve bu yönteme lineerleştirme demişlerdir. H. Chandrasekhar ise 1976'da bu denklemi iki çift nonlinear diferansiyel denkleme dönüştürmüş ve bu yönteme "Chandrasekhar yöntemi" demiştir [13]. D.W. Rand, P. Winternitz [14] ve M. Sorine [15], bu denklem için superposition yöntemlerini ileri sürmüşlerdir. J.M. Sanz-Serna ise 1992'de denklemi, klasik bir özdeğer yaklaşımı ile çözmeye çalışarak

Hamiltonian yöntemini önermiştir [16]. L. Jodar ve E. Ponsoda (1.1) denklemini çözmek için B-spline matris fonksiyonlarını kullanarak Spline method adını verdiği bir çalışma yapmıştır [17]. Ancak bu yöntemler büyük ölçekli diferansiyel denklem tiplerini çözmede belli bir nümerik adımdan sonra hesaplamadaki karmaşıklık nedeniyle elverişli değildir. (1.1) denklemi için son yıllarda yapılan çalışmalardan bazıları şu şekildedir; S. R. Nekoo ve M.I. Rahaghi [18] Taylor seri açılımına bağlı rekürsif bir yöntem tanımlayarak yaklaşık çözümü elde etmeye çalışmış ve C. H. Guo ve B. Yu [19] ise (1.1) denklemi için, denklemin katsayı matrislerinden oluşturduğu bir M -matrisinin katlama yöntemi olarak adlandırılan bir yöntem yardımıyla çözüme yakınsamasını araştırmıştır. [20]'de (1.1) denklemine trigonometrik bir yaklaşım önerilmiştir. Matris Riccati diferansiyel denklemini çözmek için son yıllarda yapılan diğer çalışmalardan bazılarını [21–23]'de görmek mümkündür.

Büyük ölçekli matris diferansiyel denklemin kesin çözümünü veren bilinen bir yöntem henüz mevcut olmadığından nümerik yöntemler kullanılmaktadır. Büyük ölçekli MRDE'yi çözmek için P. Benner ve H. Mena, 2004'de Geri fark yöntemi (Backward Differentiation Formula(BDF)) kullanmıştır [24]. Denkleme doğrudan uygulanabilen bu yöntem sayesinde MRDE'yi, matris Riccati cebirsel denklemine dönüştürmüş ve Newton yöntemi [3, 4, 25, 26] yardımıyla denklemin bir yaklaşık çözümünü elde etmişlerdir. BDF ve Rosenbrock yöntemlerinin iyi kararlılık özellikleri bu yöntemleri diferansiyel Riccati denklemlerini çözmede pratik bir alternatif haline getirir. Literatürde bu yöntemler integrasyon şemaları olarak yer almaktadır.

Büyük ölçekli MRDE denkleminin çözümü için bir başka yaklaşım T. Stillfjord tarafından 2015 yılında ileri sürüldü [27]. Bu yaklaşıma göre Stillfjord denklemi, Lie ve Strang ayrıştırma şemaları yardımıyla afin ve nonlinear olmak üzere iki ayrı denkleme böldü ve bu denklemleri Runge-Kutta ve Gauss yöntemleri yardımıyla çözdü [28, 29].

Krylov altuzay yöntemi, ilk olarak Rus matematikçi Alexei Krylov tarafından 1931 yılında yazılan bir makalede ortaya çıkmıştır. Yöntem, 1950'li yıllarda C. Lanczos [30] ve Arnoldi [31] tarafından geliştirilmiştir. Lanczos yöntemi iki karşılıklı ortogonal vektör dizisine dayanır ve özdeğer problemlerini çözmeyi amaçlar. Bu bağlamda yöntemin en belirgin özelliği, ortogonal matrisi üçlü köşegen forma indirgemesidir. Lanczos daha sonra lineer sistemleri, özellikle simetrik olanları çözmek için kendi yöntemini uyguladı. Hestenes ve Stiefel tarafından yapılan çalışma, lineer sistemlerin çözümü için eşlenik gradyan (conjugate gradient) yönteminin klasik tanımıdır [32]. Hata azaltma özelliklerine rağmen, eşlenik gradyan yöntemi iteratif bir yöntem yerine doğrudan bir yöntem olarak sunulur. Eşlenik gradyan yöntemi, Reid tarafından seyrek matris sistemlerinde önemli bir uygulama olarak ifade edilmiştir

[33].

Arnoldi tarafından yapılan bir çalışma, Lanczos'un ortogonalleştirme yöntemini daha ayrıntılı olarak ele alır ve Lanczos, Hestenes/ Stiefel yöntemlerinin özelliklerini birleştiren yeni bir yöntem ileri sürer [31]. Bu yöntem, Lanczos yöntemi gibi simetrik olmayan sistemlere uygulanır. Yöntemin üçlü köşegen forma uygulanmasını değil, bir üst Hessenberg formuna indirgenmesini ifade eder. "Galerkin yönteminde minimize edilmiş iterasyonlar" olarak sunulan bu algoritma, Arnoldi algoritması olarak bilinir.

Daha sonraki yıllarda, Arnoldi yönteminin bir üst Hessenberg matrisini kullanan yeni yöntemler geliştirilmiştir. Simetrik olmayan sistemler için bu ortogonal izdüşüm yöntemlerinin genel bir karakterizasyonu Manteuffel ve Saylor [34], Saad ve Schultz [35] tarafından verilmiştir. Krylov altuzay yöntemlerinden iyi bilinenleri; Arnoldi yöntemi, genelleştirilmiş minimum rezidü yöntemi (generalized minimal residual method), simetrik Lanczos algoritması (symmetric Lanczos algorithm), eşlenik gradyan yöntemi, blok Krylov (block Krylov), genişletilmiş blok Krylov (extended block Krylov) ve genişletilmiş global Krylov (extended global Krylov) yöntemleridir.

Tezin kalan bölümleri şu şekilde düzenlenmiştir. Bölüm 2'de kontrol sistemlerin teorisi ve büyük ölçekli matris diferansiyel denklemlerin ortaya çıktığı model problemlerden örnekler verilmiştir. Bölüm 3'de Krylov altuzay yöntemlerinden bazıları ve BDF yöntemi ayrıntılı olarak açıklanmıştır. Bölüm 4'de büyük ölçekli MRDE ve büyük ölçekli MSSDE için yeni bir yaklaşımdan bahsedilmiş, bu yaklaşımın denklemlere nasıl uygulandığı gösterilmiş ve buna bağlı bazı teorik sonuçlar verilmiştir. Bölüm 5'te NMSDE'nin çözümü için global Arnoldi algoritmasına bağlı bir yaklaşımdan bahsedilmiş ve yönteme dair bazı teorik sonuçlar verilmiştir. Bölüm 6'da ise tanıtılan yeni yöntemlerin, ele alınan problemlerdeki performansı sayısal örnekler üzerinde incelenmiştir. Her bir örnek, MATLAB programı kullanılarak çözülmüş ve elde edilen sonuçlar tablolar ve şekillerle gösterilmiştir.

1.2 Tezin Amacı

20. yüzyılda bilim ve mühendisliğin gelişimi ve pratiği üzerindeki etki bakımından, Krylov alt uzay yöntemleri en önemli nümerik yöntemlerden biri olarak kabul edilir. Bu yöntemler büyük ölçekli problemlerin çözümünde hesaplama zamanı ve bilgisayar hafızası kullanımı bakımından etkilidir ve yakınsak sonuçlar vermektedir. Bu yöntemlerin doğrudan bir integrasyon şeması (BDF) ile kombinasyonuna dayanan yeni bir yaklaşım ile büyük ölçekli MRDE, MSSDE ve NMSDE gibi diferansiyel matris denklemlerinin çözümünde etkili ve yakınsak sonuç vermesi, bu sayede yöntemin farklı tipteki büyük matris diferansiyel denklemlerine uygulanabilmesi

amaçlanmaktadır.

1.3 Bulgular

Bu tezde, Krylov altuzay yöntemi ve iyi bilinen bir integrasyon şeması olan BDF yöntemi kullanılarak tanımlanan yeni nümerik yaklaşımın büyük ölçekli matris diferansiyel denklemlerinin çözümünde etkili ve yakınsak sonuçlar vereceği ileri sürülmektedir. Yöntemin, MRDE, MSSDE ve NMSDE denklemlerine uygulama süreci gösterildi ve yakınsak sonuçlar verip vermediği incelendi.



Bu bölümde lineer zamanla değişmeyen (LTI) sonlu horizon kontrol sistemleri teorisinin önemli özellikleri ve bunların sonuçları hakkında kısa bir giriş yapacağız. Lineer zamanla değişen sistemler için gerekli temel bilgileri veren genel bir bakış [36]'da bulunabilir.

2.1 Lineer Zamanla Değişmeyen Sistemler

Bir lineer zamanla değişmeyen (LTI) sistem [37];

$$\begin{cases} \dot{x}(t) = Ax(t) + Bu(t), \\ y(t) = Cx(t) + Du(t), \end{cases} \quad (2.1)$$

denklemlerinden oluşur. Burada $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times m}$, $C \in \mathbb{R}^{p \times n}$ ve $D \in \mathbb{R}^{p \times m}$ matrislerine sistem matrisleri denir. Ayrıca A matrisine durum uzay matrisi, B ve C matrislerine sırasıyla girdi ve çıktı matrisleri ve D matrisine de doğrudan iletim dönüşümü denir. $x \in \mathbb{R}^n$, $y \in \mathbb{R}^p$ ve $u \in \mathbb{R}^m$ vektörlerine sırasıyla sistemin durumu, çıktısı ve kontrolü denir. Ayrıca ilk denkleme durum denklemi, ikinci denkleme çıktı denklemi denir.

LTI sistem (2.1), doğrudan bir uygulamanın modellenmesinden ya da lineer olmayan bir modelin lineerleştirilmesinden ortaya çıkar. Bu sistemi herhangi bir sistemden ayıran en önemli özellik çıkış sinyalinin giriş sinyali cinsinden matematiksel olarak ifade edilebilir olmasıdır. Zaman değişmezlik; bir sistemin girişinde t_0 kadar gecikme oluyorsa çıkışında da t_0 gecikme olmasıdır. Bir başka deyişle zaman değişmezlik; sisteme $x(t)$ sinyali girildiğinde $y(t)$ elde ediliyorsa, $x(t - t_0)$ sinyali girildiğinde $y(t - t_0)$ elde edilmesidir.

Lineer zamanla değişen bir sistem, sistem matrislerinin zamana bağlı olabileceği bir sistemdir. Sistem matrisleri x durumuna veya u kontrolüne bağlıysa sistemin lineer olmadığı söylenir. Burada LTI sistemine odaklanılacaktır ve uygulamaların çoğunda

$D = 0$ alınmıştır.

Şimdi optimal kontrolün varlığını ve tekliğini garanti etmek için ihtiyaç duyacağımız önemli özellikler olan sistemin kararlılığı, belirlenebilirliği, kontrol edilebilirliği ve gözlenebilirliği tanıtılacaktır [37].

Eğer bir matrisin tüm özdeğerleri kompleks düzlemin sol açık yarısında (open left half) bulunuyorsa bu matrise Hurwitz kararlıdır denir yani $\lambda \in \Lambda(A) \Rightarrow \lambda \in \mathbb{C}_{<0}$ 'dir.. Genellikle Hurwitz kararlı bir matris basitçe Hurwitz veya kararlı olarak adlandırılır. Eğer $A - BF$ kararlı olacak şekilde bir $F \in \mathbb{R}^{m \times n}$ matrisi var ise (2.1) sistemine kararlı kılınabilir denir. Kısaca (A, B) kararlı kılınabilir de denilebilir. $\forall \lambda \in \mathbb{C}$ için $\text{rank}([A - \lambda I], B) = n$ ise (A, B) matris çiftine kontrol edilebilirdir denir.

$$\begin{cases} \dot{x}(t) = A^T x(t) + C^T u(t), \\ y(t) = B^T x(t) + D^T u(t), \end{cases} \quad (2.2)$$

sistemine (2.1) sistemi için adjoint sistem denir.

Eğer adjoint sistem (ya da (A^T, C^T) matris çifti) kararlı kılınabilir ise (2.1) sistemine (ya da (C, A) matris çiftine) belirlenebilirdir denir. Benzer şekilde eğer adjoint sistem (veya (A^T, C^T) matris çifti) kontrol edilebilir ise (2.1) sistemine (veya (C, A) matris çiftine) gözlenebilirdir denir.

Burada tanımlanan kontrol edilebilirliğin sonlu boyutlu sistemler ile sınırlı olduğuna dikkat edilmelidir. Belirlenebilirlik, kararlı kılınabilirlik ve gözlenebilirlik özellikleri ise sonlu ve sonsuz boyutlu sistemler için aynıdır.

2.2 Sonlu Horizontda Lineer Kuadratik Regülatör Problemi

Lineer kuadratik regülatör (LQR) problem optimal kontrol teorisinde iyi bilinen bir modelleme tekniğidir. Sistem dinamikleri bir dizi lineer diferansiyel denklem tarafından belirlenir ve maliyet fonksiyonu kuadratik bir fonksiyondur. Bir LQR problemi aşağıdaki şekilde ifade edilir; [1, 38, 39]

$x(t) \in \mathbb{R}^n$ durum vektörü, $u(t) \in \mathbb{R}^p$ kontrol vektörü, $y(t) \in \mathbb{R}^s$ çıktı vektörü olmak üzere her bir x_0 başlangıç durumu için bir LQR problemi, (2.1) dinamik kısıtlamaları altında

$$J(x_0, u) = \inf \left(\int_0^{T_f} (y(t)^T y(t) + u(t)^T u(t)) dt \right), \quad (2.3)$$

optimal maliyetinin hesaplanmasına denktir. Buradan bir minimum mevcut

olduğunda, bu minimum seviyeye ulaşan bir $\hat{u}(t)$ optimal kontrol girdisi bulunmalıdır. Yani (2.1) denklem sistemini sağlayan bir $\hat{x}$ optimal durum ve buna karşılık gelen $\hat{y}$ çıktısı için

$$J(x_0, \hat{u}) = \int_0^{T_f} (\hat{y}(t)^T \hat{y}(t) + \hat{u}(t)^T \hat{u}(t)) dt, \quad (2.4)$$

dir. (A, B) 'nin kararlı kılınabilir, (C, A) 'nın belirlenebilir olduğunu kabul ederek $J(x_0, u)$ fonksiyoneli minimize eden optimal $\hat{u}(t)$ girdisi ve $\hat{u}(t) = K\hat{x}(t)$, $K = -B^T P(t)$ olacak şekilde bir K geri bildirim operatörüyle belirlenir ve burada $P(t) \in \mathbb{R}^{n \times n}$,

$$\dot{P}(t) + A^T P(t) + P(t)A - P(t)B B^T P(t) + C^T C = 0, \quad P(T_f) = 0, \quad (2.5)$$

şeklindeki MRDE'nin bir tek çözümüdür. Ayrıca optimal durum yörüngesi, $\dot{\hat{x}}(t) = (A - B B^T P(t))\hat{x}(t)$ 'yi sağlar ve

$$\hat{x}(t) = e^{tA}\hat{x}_0 + \int_0^t e^{(t-\tau)A} B u(\tau) d\tau, \quad (2.6)$$

şeklinde ifade edilebilir.

x_0 başlangıç durumunun optimal maliyeti aşağıdaki kuadratik fonksiyon ile verilir [40]:

$$J(x_0, \hat{u}) = x_0^T P(0) x_0. \quad (2.7)$$

Eğer $X(t) = P(T_f - t)$ ve $X_0 = 0$ denilirse o zaman

$$P(t) = X(T_f - t), \quad (2.8)$$

olduğuna dikkat edilmelidir. Buradan MRDE denkleminin X çözümü elde edilir ve

$$J(x_0, \hat{u}) = x_0^T X(T_f) x_0, \quad (2.9)$$

olur.

MRDE'nin varlık ve tekliği üzerinde daha kapsamlı bir çalışmaya [1, 38, 40, 41]'de yer verilmiştir.

Teorem 2.1. (A, B) 'nin kararlı kılınabilir, (C, A) 'nın belirlenebilir olduğunu kabul edelim. O zaman (1.1) ile verilen MRDE'nin $[0, T_f]$ üzerinde tek bir pozitif X çözümü vardır

ve her x_0 başlangıç durumu için $J(x_0, \hat{u})$ 'nin optimal maliyeti

$$J(x_0, \hat{u}) = x_0^T X(T_f) x_0, \quad (2.10)$$

olup buradan optimal kontrol

$$\hat{u}(t) = -B^T X(T_f - t) \hat{x}(t), \quad (2.11)$$

ve optimal yörünge ise

$$\hat{x}(t) = (A - BB^T X(T_f - t)) \hat{x}(t), \quad \hat{x}(0) = x_0, \quad (2.12)$$

ile belirlenir [40].

Not. Sonsuz horizon durumunda optimal maliyet ,

$$J(x_0, u_\infty) = \left\{ \int_0^\infty (y(t)^T y(t) + u(t)^T u(t)) dt \right\} = x_0^T X_\infty x_0, \quad (2.13)$$

ile verilir. X_∞ ,

$$A^T X_\infty + X_\infty A - X_\infty B B^T X_\infty + C^T C = 0, \quad (2.14)$$

cebirsel Riccati denkleminin bir tek kararlı ve pozitif çözüdür. Burada optimal geri bildirim $u_\infty = -B^T X_\infty \hat{x}(t)$ ile verilir [40].

Not. Ayırık durum için LQR sonlu horizon problemi,

$$\begin{cases} x_{k+1} = Ax_k + Bu_k, \\ y_k = Cx_k, \end{cases} \quad (2.15)$$

ayırık dinamik kısıtlamaları altında

$$J(x_0, u) = \inf_u \left\{ \sum_0^N (y_k^T y_k + u_k^T R u_k) \right\}, \quad (2.16)$$

ile tanımlanır.

Optimal kontrol

$u_k = -F_k u_k$, $F_k = (R + B^T Z_{k+1} B)^{-1} B^T Z_{k+1} A$ ile verilir ve Z_{k+1} aşağıdaki ayırık zaman

matris Riccati cebirsel denkleminin çözülmesiyle hesaplanır;

$$Z_k = A^T Z_{k+1} A - A^T Z_{k+1} B (R + B^T Z_{k+1} B)^{-1} B^T Z_{k+1} A + C^T C. \quad (2.17)$$

N sonsuza giderken, sonsuz horizon ayırık zaman LQR elde edilir ve bazı varsayımlar altında Z_∞ , $Z_\infty = \lim_{k \rightarrow \infty} Z_k$,

$$Z_\infty = A^T Z_\infty A - A^T Z_\infty B (R + B^T Z_\infty B)^{-1} B^T Z_\infty A + C^T C, \quad (2.18)$$

ayırık zaman cebirsel Riccati denkleminin bir tek pozitif tanımlı çözümüdür [40].

Bu çalışmada, ilgili büyük ölçekli MRDE'ye yaklaşık çözüm sağlayan etkili nümerik yöntemlerin geliştirilmesini gerektiren “sürekli” durum dikkate alınacaktır.

Büyük ölçekli problemlerin başlıca kaynağı kısmi diferansiyel denklemlerdir. Matris diferansiyel denklemler kısmi diferansiyel denklemlerin ayrıştırılmasından elde edilir. Şimdi bununla ilgili model problem örnekleri verilecektir [37].

2.3 Model Problemler

2.3.1 Sonlu Farklar Yöntemi: Yarı-Ayrıklaştırılmış Isı Denklemi

$(0, 1) \times (0, 1)$ birim karedeki sonlu farklar yarı-ayrıklaştırılmış ısı denklemi, en temel modelleme örneklerinden biridir. Bu modelin avantajları şu şekilde sıralanabilir;

- Keyfi büyüklükteki basit test problemlerinin oluşturulmasını sağlar.
- Sonlu farklar yöntemi ile denklemlerin ayrıklaştırılması kolaydır.

Sonlu farkların bir başka önemli özelliği ise sonlu elemanlar yönteminin aksine zamana bağlı türevin katsayısı olarak yoğun bir matris üretmez yani standart durum uzayı temsiline büyük ölçekli seyrek bir matris oluşturur.

Bu problemin iki çeşidi olan difüzyon denklemi ile konveksiyon-reaksiyon denklemleri sırasıyla

$$\frac{\partial x}{\partial t} - \Delta x = f(\psi)u(t), \quad (2.19)$$

ve

$$\frac{\partial x}{\partial t} - v \cdot \nabla x - \Delta x - qx = f(\psi)u(t) \quad (2.20)$$

ile verilir. Modeller, merkezi sonlu fark kullanılarak üretilir.

2.3.2 Kimyasal Reaktörler: Akan Reaktifin Sıcaklığını Kontrol Etme

Bu örnek, bir tüp içindeki bir akışkan akışının optimum ısıtılmasında / soğutulmasında ortaya çıkan bir sistemdir. Kimyasal reaktörlerde belirli reaktif girişlerinin sıcaklık düzenlemesi bunun bir uygulaması olabilir. Model denklemleri;

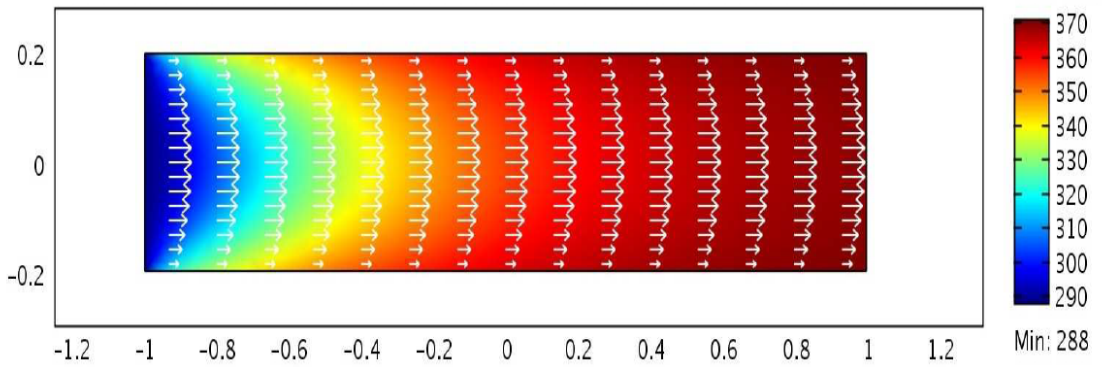
$$\frac{\partial x}{\partial t} - \kappa \Delta x + v \cdot \nabla x = 0, \quad \Omega \text{ bölgesinde,} \quad (2.21)$$

$$x = x_0, \quad \Gamma_{ic} \text{ sınırında,} \quad (2.22)$$

$$\frac{\partial x}{\partial \eta} = \sigma(u - x), \quad \Gamma_{sicaklik1} \cup \Gamma_{sicaklik2} \text{ sınırında,} \quad (2.23)$$

$$\frac{\partial x}{\partial \eta} = 0, \quad \Gamma_{dis} \text{ sınırında,} \quad (2.24)$$

şeklindedir. Burada Ω , Şekil 2.1 ile verilen dikdörtgensel alandır.



Şekil 2.1 İçe akış için Ω alanı: dairesel bir tüpte $t = 2$ için içe akışın 2-boyutlu dik kesiti

İçe akış Γ_{ic} , sınırın sol kısmı ve dışa akış Γ_{dis} , sınırın sağ kısmıdır. Kontrol ise üst ve alt sınırlar aracılığıyla uygulanır. Dalgasız akışın baskın olduğu dairesel simetri varsayımıyla model 2 boyutlu olarak sınırlanabilir. Dairesel simetri nedeniyle sistemin üst ve alt sınırlarına uygulanan tek bir girdisi vardır ve çıktılar dışa akıştaki sıcaklık değerleridir.

Burada bir sonlu eleman ayrıştırması uygulandığından yarı ayrıklaştırılmış model

$$\tilde{M}\dot{x} = \tilde{A}x + \tilde{B}u, \quad (2.25)$$

$$y = \tilde{C}x, \quad (2.26)$$

şeklindedir.

2.3.3 Çelik Profillerin Seçici Soğutulması: Bir Demiryolu Haddehanesinde Soğutma

Bu bölümde bir haddehanede çelik profillerin optimum şekilde soğutulması problemi ele alınacaktır. Bu problem [42]'de daha ayrıntılı bir şekilde tartışılmıştır.

2.3.3.1 Model Altyapısı

Bu problem, bir haddehanede üretim sürecindeki farklı işlemlerin ham maddenin farklı sıcaklıklarını gerektirdiğinde ortaya çıkar. Yüksek üretim hızı elde etmek için ekonomik çıkarlar gözönünde bulundurularak bir sonraki üretim aşamasına girmeden önce sıcaklık mümkün olan en hızlı şekilde istenen seviyeye düşürülmek istenir. Aynı zamanda, soğutma akışkanlarının yüzeye püskürtülmesi ile gerçekleştirilen soğutma işlemi, dayanıklılık veya gözeneklilik gibi malzeme özelliklerinin belirli kalite standartlarına ulaşması için kontrol edilmelidir. Çelik profil içindeki sıcaklık dağılımlarındaki büyük değişimler istenmeyen deformasyonlara, kırılabilirliğe, rijitlik kaybına ve diğer istenmeyen malzeme özelliklerine yol açabilir. Bu nedenle, mühendisin hedefi, tercihen eşit bir sıcaklık dağılımına sahip olmaktır.

2.3.3.2 Model Denklemi

[43, 44]'de olduğu gibi, çelik profilin z yönünde sonsuz bir şekilde gerildiği varsayılmaktadır. Bu varsayımla z yönünde sabit bir ısı dağılımı olduğu da kabul edilir yani söz konusu modelin şekil 2.2'de görüldüğü gibi $\Omega \in \mathbb{R}^2$ profilinin dik kesitleri üzerinde 2 boyutlu bir ısı yayılım süreciyle sınırlandırılabilir anlamına gelir. Bu nedenle, Şekil 2.2'de gösterilen profilin 2 boyutlu kesitleri hesaplama alanı olarak düşünülebilir. Kesit geometrisini tanımlamak için ölçümler [43]'den alınmıştır. Şekil 2.2'de görülebileceği gibi etki alanı, simetri eksenini üzerinde yapay bir Γ_0 sınırı oluşturan profilin simetrisinden faydalanır.

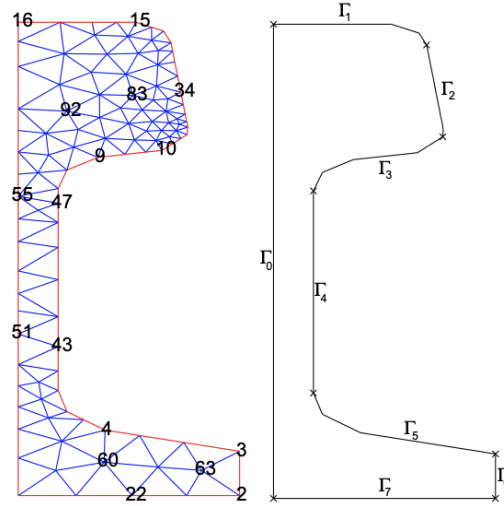
[43, 44]'de tanıtılan durum denkleminin lineerleştirilmiş versiyonuna odaklanılacaktır. Lineerleştirme; ρ , λ ve c malzeme parametreleri kullanılarak elde edilir. Bu işlem ρ , λ ve c değişikliklerinin küçük olduğu ve malzemede çok fazla faz ve faz geçişine sahip olunmadığı yaklaşık 700-800°C üzerindeki sıcaklık rejimlerinde çalışıldığı sürece (kullanılan çeliğin türüne bağlı olarak) kabul edilebilir. Ω 'nın sol kısmındaki yapay sınırla birlikte sınır 8 kısıma ayrılır (ayrıntılar için bkz. Şekil 2.2). u kontrolünün sabit olduğu kabul edilerek t zamanında ξ noktasındaki

$x(\xi, t)$ sıcaklığı için aşağıdaki model elde edilir:

$$c\rho \frac{\partial x(\xi, t)}{\partial t} = \nabla(\lambda \cdot \nabla x(\xi, t)), \quad \Omega \times (0, T) \text{ bölgesinde,} \quad (2.27)$$

$$-\lambda \frac{\partial x(\xi, t)}{\partial \eta} = g_i(t, x, u), \quad i = 1, \dots, 7 \text{ için } \Gamma_i \text{ sınırında,} \quad (2.28)$$

$$x(\xi, 0) = x_0(\xi), \quad \Omega \text{ bölgesinde,} \quad (2.29)$$



Şekil 2.2 Çelik Profillerin Seçici Soğutulması için Etki Alanı: en aza indirme ve karşılaştırma noktalarına sahip ilk ağ (solda) ve sınırın bölünmesi (sağda)

Burada g_i , soğutma sıvısı ile profil yüzeyi arasındaki sıcaklık farkıdır. Malzemenin yüzeyi arasındaki ısı transferi yani sınır koşulları, Newton'un soğutma kanununa göre aşağıdaki gibi verilir;

$$-\frac{\partial x(\xi, t)}{\partial \eta} = \frac{\kappa_k}{\lambda}(x(\xi, t) - x_{ext,k}(t)). \quad (2.30)$$

$x_{ext,k}(t)$ 'nin Γ_k üzerinde sabit olduğu varsayılır ve böylece $x_{ext,k}(t)$, ξ 'ye bağlı değildir. Bu model problemde kontrol olarak dış sıcaklık alınır.

Krylov Altuzay ve Geri Fark Formülü (BDF) Yöntemleri

Bu bölümde büyük lineer denklem sistemlerini çözmek için bazı iteratif izdüşüm yöntemlerinden [45–49] ve BDF yönteminden [38] bahsedilecektir. İlk olarak Krylov altuzay yöntemlerinde kullanılan izdüşüm çeşitlerinden bahsedelim.

3.1 Genel İzdüşüm Yöntemleri

3.1.1 Eğik İzdüşümler

Bir P izdüşümü

$$P^2 = P, \quad (3.1)$$

şartını sağlayan $\mathbb{R}^n$ 'den $\mathbb{R}^n$ 'e tanımlı bir lineer dönüşümdür. Eğer P bir izdüşüm ise $I - P$ 'de bir izdüşümdür ve

$$Ker(P) = Range(I - P), \quad (3.2)$$

dir. Burada $Ker(P) = \{x \in \mathbb{R}^n | P(x) = 0\}$ ve $Range(P) = \{y = Px | x \in \mathbb{R}^n\}$ 'dir. $\forall x \in \mathbb{R}^n$ için $x = Px + (I - P)x$ ve $Ker(P) \cap Range(P) = \{0\}$ olduğundan

$$\mathbb{R}^n = Ker(P) \oplus Range(P), \quad (3.3)$$

yazılabilir. $\mathbb{R}^n = \tilde{M} \oplus \tilde{L}^\perp$ olacak şekilde $\tilde{M}$ ve $\tilde{L}$, $\mathbb{R}^n$ 'in altuzayları olsun. $\tilde{M}$ üzerine, $\tilde{L}$ 'ya ortogonal bir izdüşüm aşağıdaki gibi tanımlanır:

$\forall x \in \mathbb{R}^n$ için $y \in \tilde{M}$ ve $z \in \tilde{L}^\perp$ olmak üzere

$$x = y + z, \quad (3.4)$$

dir. $\tilde{M}$ üzerine, $\tilde{L}^\perp$ 'e paralel bir P izdüşümü

$$Px = y, \quad (3.5)$$

şeklindedir.

Eğik bir P izdüşümü :

$$\begin{cases} Px \in \tilde{M}, \\ x - Px \perp \tilde{L}, \end{cases} \quad (3.6)$$

ile verilir. P 'ye; $\tilde{M}$ üzerine, $\tilde{L}^\perp$ boyunca bir izdüşüm de denilebilir ve

$$Range(P) = \tilde{M} \text{ ve } Ker(P) = \tilde{L}^\perp \quad (3.7)$$

olduğuna dikkat edilmelidir. P izdüşümünün bir matris temsili aşağıdaki gibi verilir:

$\tilde{M}$ ve $\tilde{L}$ altuzayları m boyutlu olsun ve $\{u_1, u_2, \dots, u_m\}$ ile $\{w_1, w_2, \dots, w_m\}$ sırasıyla $\tilde{M}$ ve $\tilde{L}$ 'nin tabanları olsun.

$U = [u_1, u_2, \dots, u_m]$ ve $W = [w_1, w_2, \dots, w_m]$ olmak üzere

$$Px \in \tilde{M} \Leftrightarrow Px = Uy, \quad y \in \mathbb{R}^m, \quad (3.8)$$

dir. Diğer taraftan diklik şartı (Galerkin ortogonallik şartı)

$$x - Px \perp \tilde{L} \Leftrightarrow W^T(x - Px) = 0, \quad (3.9)$$

şeklinde ifade edilir.

(3.8) ve (3.9) ifadelerinden P 'nin matris gösterimi

$$P = U(W^T U)^{-1} W^T, \quad (3.10)$$

olarak elde edilir. Eğer tabanlar biortonormal ise yani $\langle u_i, w_j \rangle_2 = \delta_{ij}$ ise bu ifade

$$P = UW^T, \quad (3.11)$$

ifadesine indirgenir. Burada dikkat edilmelidir ki P 'nin matris gösterimi tek değildir ve seçilen tabana göre değişir.

3.1.2 Ortogonal İzdüşümler

$\tilde{M} = \tilde{L}$ olsun. Bu durumda P ortogonal bir izdüşümdür ve

$$\begin{cases} Px \in \tilde{M}, \\ x - Px \perp \tilde{M}, \end{cases} \quad (3.12)$$

ya da denk olarak

$$Px \in M \text{ ve } ((I - P)x, y) = 0, \quad \forall y \in M, \quad (3.13)$$

dir. P ortogonal izdüşümünün matris gösterimi

$$P = U(U^T U)^{-1} U^T, \quad (3.14)$$

ile verilir. Eğer $\{u_1, u_2, \dots, u_m\}$ ortonormal bir taban ise

$$P = U U^T, \quad (3.15)$$

dir.

3.2 Krylov Altuzay Yöntemleri

Krylov altuzay yöntemleri, lineer ve lineer olmayan denklemlerin büyük sistemlerini ve büyük özdeğer problemlerini çözmekte kullanılan popüler ve kullanışlı yöntemlerdir. Bu yöntemler, yoğun matris problemlerini çözmek için kullanılan daha maliyetli yöntemlere verimli ve güvenilir bir alternatif olarak son yıllarda giderek daha çok benimsenmiştir. Krylov altuzay yöntemleri katsayı matrislerinin ayrışımı yerine matris-vektör işlemleri gerektirdiğinden ($> 10^8$) bilinmeyenli büyük problemler için uygundur [45].

Bu teknikler, p bir polinom olmak üzere $p(A)v$ vektörleri tarafından gerilen altuzayların oluşturduğu Krylov altuzayı üzerine hem dik hem de eğik izdüşümlere bağlıdır. Özetle bu teknikler p bir polinom olmak üzere $Ax = b$ lineer denklem sistemi için $A^{-1}b$ 'ye $p(A)b$ ile yaklaşır.

$v \in \mathbb{R}^n$ herhangi bir vektör ve $A \in \mathbb{R}^{n \times n}$ olmak üzere (A, v) çifti için bir m . Krylov altuzayı

$$K_m(A, v) = \text{span} \{v, Av, \dots, A^{m-1}v\}, \quad (3.16)$$

ile tanımlıdır [45].

Şimdi $Ax = b$ lineer denklem sisteminin Krylov altuzayları kullanılarak yaklaşık bir çözümünün nasıl elde edildiğinden bahsedilecektir.

p_A , A 'nın karakteristik polinomu olsun. Cayley-Hamilton teoremine göre $p_A(A) = 0$ olup

$$a_0 I + a_1 A + \dots + a_n A^n = 0, \quad a_0 \neq 0, \quad (3.17)$$

şeklinde ifade edilebilir.

Her iki taraf a_0 ' a bölünürse

$$I + \frac{a_1}{a_0}A + \dots + \frac{a_n}{a_0}A^{n-1} = 0, \quad (3.18)$$

elde edilir. Buradan eşitliğin her iki tarafı A^{-1} ile çarpılırsa

$$A^{-1} + \frac{a_1}{a_0} + \dots + \frac{a_n}{a_0}A^{n-1} = 0, \quad (3.19)$$

$$\Rightarrow A^{-1} = b_0I + b_1A + \dots + b_{n-1}A^{n-1}, \quad (3.20)$$

elde edilir. Burada tam çözüm $x = A^{-1}b \in K_n(A, b)$ 'dir. Yani tam çözüm n . Krylov altuzayındadır. Bu bize yaklaşık çözümlerin Krylov altuzaylarında aranabileceğini gösterir [45].

3.2.1 Arnoldi Yöntemi

Bu bölümde $V_m = [v_1, v_2, \dots, v_m]$ tabanının seçimine dair hesaplamayı kolaylaştıracak bir yöntem tanıtılacaktır. Bu yöntem, $\{r_0, Ar_0, \dots, A^{m-1}r_0\}$ Krylov tabanına Gram-Schmidt yönteminin uygulanmasıyla elde edilen Arnoldi algoritması olup bu algoritma sonunda $K_m(A, r_0)$ Krylov altuzayının bir ortonormal tabanı elde edilir. Bu yöntemle yoğun bir matris üniter bir dönüşümle Hessenberg forma indirgenebilir. Arnoldi, elde edilen Hessenberg matrisin, orjinal matrisin bazı özdeğerlerine tam yaklaşım sağlayabileceğini ifade etmiştir. Arnoldi yöntemi, Krylov altuzayı üzerine ortogonal bir izdüşümdür. Arnoldi algoritması aşağıdaki gibi tanımlanır [45]:

Algoritma 3.1: Arnoldi

- $v_1 = r_0 / \|r_0\|_2$,
- $j = 1, 2, \dots, m$ için,
- $w = Av_j$ 'yi hesapla,
- $i = 1, 2, \dots, j$ için,

$$(a) h_{i,j} = \langle w, v_i \rangle_2,$$

$$(b) w = w - h_{i,j} v_i,$$

- i için döngüyü sonlandır,
 - $h_{j+1,j} = \|w\|_2$. Eğer $h_{j+1,j} = 0$ ise iterasyonu durdur, değilse,
 - $v_{j+1} = w / h_{j+1,j}$ hesapla,
 - j için döngüyü sonlandır.
-

Yöntemin her adımında algoritma bir önceki v_j Arnoldi vektörüyle A matrisini çarpar ve standart Gram-Schmidt yöntemine göre elde edilen w 'leri algoritmadaki tüm v_i 'ler için ortonormalleştirir.

Sonuç 3.1. $V_m = [v_1, v_2, \dots, v_m]$ ve $\tilde{H}_m = [h_{ij}]$, $(m+1, m)$ tipinde bir üst Hessenberg matris ve H_m matrisi, $\tilde{H}_m$ 'in son satırının silinmesiyle elde edilen $m \times m$ tipinde bir üst Hessenberg matrisi olsun. Buradan aşağıdaki Arnoldi bağıntıları elde edilir:

$e_m^T = [0, 0, \dots, 1] \in \mathbb{R}^{m+1}$ olmak üzere

$$H_m = V_m^T A V_m, \quad (3.21)$$

ve

$$A V_m = V_{m+1} \tilde{H}_m = V_m H_m + h_{m+1,m} v_{m+1} e_m^T, \quad (3.22)$$

dir.

Bu bağıntılar Arnoldi algoritmasından doğrudan elde edilir. Eğer A matrisi simetrik ise o zaman H_m üst Hessenberg matrisi üçlü köşegen matrise indirgenir.

3.3 Blok Krylov ve Genişletilmiş Krylov Altuzay Yöntemleri

Bu bölümde $A \in \mathbb{R}^{n \times n}$ nonsingüler matris, $B \in \mathbb{R}^{n \times s}$ ve $X \in \mathbb{R}^{n \times s}$, $s \ll n$, olmak üzere

$$AX = B \quad (3.23)$$

problemi ele alınacaktır.

Bazı durumlarda tek bir vektör yerine bir vektör bloğu ile çalışmak istenir. Blok yöntemlerinin amacı (3.23) problemini daha hızlı çözmektir.

3.3.1 Blok Krylov Altuzay Yöntemleri

(A, V) matris çifti için m . blok Krylov altuzayı $\mathbb{R}^n$ 'nin bir altuzayı olup $V, AV, A^2V, \dots, A^{(m-1)}V$ matrislerinin sütunlarıyla üretilir ve

$$\mathbb{K}_m(A, V) = \text{Range}\{V, AV, \dots, A^{m-1}V\} \subset \mathbb{R}^n \quad (3.24)$$

ile gösterilir [47]. $\mathbb{K}_m(A, V)$ blok Krylov altuzayı s tane basit Krylov altuzayının toplamıdır:

$$\mathbb{K}_m(A, V) = K_m(A, b^{(1)}) + K_m(A, b^{(2)}) + \dots + K_m(A, b^{(s)}). \quad (3.25)$$

3.3.1.1 Blok Arnoldi Yöntemi

Blok Arnoldi, klasik Arnoldi algoritmasının $\mathbb{K}_m(A, V)$ blok Krylov altuzayının bir ortonormal tabanının elde edilmesini sağlayan genelleştirilmiş halidir ve bu yöntemin algoritması Arnoldi algoritmasının bir blok versiyonu olan Algoritma 2 olarak aşağıdaki gibi tanımlanır [47].

Algoritma 3.2: Blok Arnoldi

- Veriler: $A \in \mathbb{R}^{n \times n}$ matrisi ve $V \in \mathbb{R}^{n \times s}$ bloğu,
- $V = V_1^b R_1$ (QR ayrışımı) hesapla,
- $j = 1, 2, \dots, m$ için,
- $W = AV_j^b$ hesapla,
- $i = 1, 2, \dots, j$ için,

$$(a) H_{i,j} = (V_i^b)^T W,$$

$$(b) W = W - V_j^b H_{i,j},$$

- i için döngüyü sonlandır,
 - $W = Q_j R_j$ (QR ayrışımı) hesapla,
 - $V_{j+1}^b = Q_j; H_{j+1,j} = R_j$ hesapla,
 - j için döngüyü sonlandır.
-

Blok Arnoldi algoritması ile üretilen $V_1^b, V_2^b, \dots, V_m^b$ blokları, üst üçgen $H_{j+1,j}$ Hessenberg matrisleri maksimum ranklı olacak şekilde sütunları birbirine ortonormaldir.

$\tilde{H}_m$, sıfırdan farklı blok girdileri $H_{i,j}; i = 1, 2, \dots, (m+1), j = 1, 2, \dots, m$ olan $(m+1)s \times ms$ üst blok Hessenberg matrisini temsil etsin. $\mathcal{H}_m, \tilde{H}_m$ 'in son s satırlarının silinmesiyle elde edilen $ms \times ms$ matrisi olsun. $V_i^b, i = 1, 2, \dots, m$ blok Arnoldi ile elde edilmiş i . bloklar olmak üzere $\mathcal{V}_m^b = [V_1^b, V_2^b, \dots, V_m^b]$ olsun. Buradan aşağıdaki blok Arnoldi bağıntıları elde edilir.

$$\mathcal{V}_m^b = \mathcal{V}_m^b \mathcal{H}_m + V_{m+1}^b H_{m+1,m} E_m^T; \quad \mathcal{V}_m^b = \mathcal{V}_{m+1}^b \tilde{H}_m, \quad (3.26)$$

ve

$$\mathcal{H}_m = (\mathcal{V}_m^b)^T A \mathcal{V}_m^b \quad (\mathcal{V}_m^b)^T \mathcal{V}_m^b = I_{ms}, \quad (3.27)$$

olup burada $E_m, ms \times ms$ tipindeki I_{ms} birim matrisin son s sütunudur.

$\mathcal{H}_m$ matrisi bir Hessenberg matris değil, bant Hessenberg matristir. Yani bir tane köşegen yerine s tane altköşegen içerir. Ayrıca çözümün arandığı uzayının boyutunun m değil ms olduğuna dikkat edilmelidir.

3.3.1.2 Global Arnoldi Yöntemi

$\mathbb{K}_m^g(A, V)$ uzayının bir F ortonormal $\{V_1^g, V_2^g, \dots, V_m^g\}$ tabanının elde edilmesini sağlayan global Arnoldi algoritması aşağıdaki gibidir [49]:

Algoritma 3.3: Global Arnoldi

- Veriler: $A \in \mathbb{R}^{n \times n}$ matrisi ve $V \in \mathbb{R}^{n \times s}$ bloğu,
- $V_1^g = V / \|V\|_F$ hesapla,
- $j = 1, \dots, m$ için,
- $W = AV_j^g$ hesapla,
- $i = 1, 2, \dots, j$ için,

$$(a) h_{i,j} = \langle V_i^g, W \rangle_F,$$

$$(b) W = W - V_j^g h_{i,j},$$

- i için döngüyü sonlandır,
 - $h_{j+1,j} = \|W\|_F$ hesapla,
 - $V_{j+1}^g = W / h_{j+1,j}$ hesapla,
 - j için döngüyü sonlandır.
-

Global Arnoldi algoritması, $(\mathcal{A}, b)$, $(\mathcal{A} = I_s \otimes A, b = \text{vec}(B))$, matris çiftine uygulanan standart Arnoldi algoritmasıdır. $s = 1$ olması durumunda algoritma, Arnoldi algoritmasına indirgenir. Global Arnoldi algoritmasının bir j adımında durması için gerek ve yeter koşul $h_{j+1,j} = 0$ olmasıdır. $\mathcal{M}_{n,s}$, reel $n \times s$ matrislerinin bir uzayı olmak üzere global Arnoldi algoritmasının $K_m^g(A, V) \subseteq \mathcal{M}_{n,s}$ 'in bir F ortonormal tabanını üretir.

$\mathcal{V}_m^g = [V_1^g, V_2^g, \dots, V_m^g]$ ve $\tilde{H}_m^g$, $h_{i,j}$ girdileri global Arnoldi yöntemi ile tanımlanan $(m+1) \times m$ üst Hessenberg matrisi ve H_m^g , $m \times m$ matrisi $\tilde{H}_m^g$ 'nin son satırının silinmesiyle elde edilen matrisi temsil etsin. Bu durumda $\otimes$ -Kronecker çarpım olmak üzere aşağıdaki özellikler sağlanır [18]:

$$A\mathcal{V}_m^g = \mathcal{V}_m^g(H_m^g \otimes I_s) + h_{m+1,m} V_{m+1}^g E_m^T, \quad (3.28)$$

ve

$$A\mathcal{V}_m^g = \mathcal{V}_{m+1}^g(\tilde{H}_m^g \otimes I_s), \quad (3.29)$$

olup burada $E_m^T = [0_s, \dots, 0_s, I_s]$ 'dir. F ortogonallik özelliği

$$(\mathcal{V}_m^g)^T \diamond \mathcal{V}_m^g = \left[\left\langle V_i^g, V_j^g \right\rangle_F \right] = I_m, \quad (3.30)$$

ile tanımlıdır ve burada $\mathcal{V}_m^g$ matrisine F ortonormal denir.

Not. • Blok Arnoldi algoritması, $\mathbb{K}_m(A, V)$ blok Krylov uzayının bir ortonormal tabanını oluştururken global Arnoldi algoritması, $\mathbb{K}_m^g(A, V) \subseteq \mathcal{M}_{n,s}$ 'nin bir F ortonormal tabanını oluşturur.

- Global Arnoldi algoritmasındaki H_m^g matrisi, $m \times m$ boyutlu iken blok Arnoldi algoritmasındaki blok Hessenberg matrisi $\mathcal{H}_m$, $ms \times ms$ boyutlu olması global ve blok Arnoldi algoritmaları arasında büyük bir farklılıktır.

3.3.2 Genişletilmiş Krylov Altuzay Yöntemi

$A \in \mathbb{R}^{n \times n}$ nonsingüler ve $V \in \mathbb{R}^{n \times s}$ olmak üzere (A, V) matris çifti için genişletilmiş Krylov altuzayı $\mathcal{K}_m^e(A, V) \subseteq \mathbb{R}^n$ şu şekilde tanımlanır [50]:

$$\mathcal{K}_m^e(A, V) = \text{Range}([V, A^{-1}V, AV, A^{-2}V, A^2V, \dots, A^{-(m-1)}V, A^{(m-1)}V]) \subseteq \mathbb{R}^n. \quad (3.31)$$

$\mathcal{K}_m^e(A, V)$ altuzayı iki blok Krylov altuzayının toplamı olarak da ifade edilebilir:

$$\mathcal{K}_m^e(A, V) = \mathbb{K}_m(A, V) + \mathbb{K}_m(A^{-1}, A^{-1}V). \quad (3.32)$$

Genişletilmiş Krylov altuzayı, genişletilmiş Arnoldi yöntemi aracılığıyla ilk olarak V. Druskin ve L. Knizhnerman tarafından tanıtıldı [50] ve V. Simoncini bu yöntemi Lyapunov denkleminde uyguladı [51].

3.3.2.1 Genişletilmiş Blok Arnoldi Yöntemi

Büyük matris denklemlerini çözmek için kullanılan Arnoldi ve blok Arnoldi yöntemleri yaklaşık çözümü hesaplarırken genellikle çok sayıda iterasyon gerektirir ve bu durum tahmini hesaplama zamanını ve bellek gereksinimini önemli ölçüde artırır. Blok Arnoldi algoritmasının bu dezavantajlarını gidermek için matris denkleminin kararlı çözümüne düşük ranklı yaklaşımlar hesaplamamızı sağlayan bir izdüşüm yöntemi tanıtıldı [50, 51]. Bu yöntemde göre başlangıç problemi A ve A^{-1} ile oluşturulan genişletilmiş Krylov altuzayı üzerine izdüşürülür ve elde edilen düşük boyutlu problem çözülür.

Aşağıdaki algoritma $\mathcal{K}_m^e(A, V)$ genişletilmiş Krylov altuzayının bir ortonormal tabanının hesaplanmasını sağlar. Bu taban hem A hem de A^{-1} matrislerini içerir.

m tamsayısı hesaplanan tabanın boyutu olmak üzere genişletilmiş blok Arnoldi algoritması aşağıdaki gibi tanımlanır [50].

Algoritma 3.4: Genişletilmiş Blok Arnoldi (EBA)

- Veriler: $A \in \mathbb{R}^{n \times n}$ bir matris, $V \in \mathbb{R}^{n \times s}$ bir matris ve m bir tamsayı,
- $[V, A^{-1}V] = V_1 \Lambda$, QR ayrışımını hesapla, $\mathcal{V}_0 = []$,
- $j = 1, 2, \dots, m$ için,

$V_j^1; V_j$ 'nin ilk s sütunu, V_j^2 ; V_j 'nin ikinci s sütunu olmak üzere,
 $\mathcal{V}_j = [\mathcal{V}_{j-1}, V_j]; \hat{V}_{j+1} = [AV_j^{(1)}, A^{-1}V_j^{(2)}]$ hesapla,

- V_{j+1} 'i elde etmek için $\hat{V}_{j+1}$ 'i, $\mathcal{V}_j$ 'ye göre ortogonalleştir, yani
 $i = 1, 2, \dots, j$ için,

$$(a) H_{i,j} = V_i^T \hat{V}_{j+1},$$

$$(b) \hat{V}_{j+1} = \hat{V}_{j+1} - V_i H_{i,j},$$

- i için döngüyü sonlandır.
 - $\hat{V}_{j+1} = V_{j+1} H_{j+1,j}$ 'in QR ayrışımını hesapla,
 - j için döngüyü sonlandır.
-

şeklinde verilir. Bu algoritma bir Gram-Schmidt yöntemini içerdiğinden algoritma ile elde edilen

$$\mathcal{V}_m = [V_1, V_2, \dots, V_m], (V_i \in \mathbb{R}^{n \times 2s}) \quad (3.33)$$

blok vektörleri birbirine ortogonal sütunlardır. Bu algoritma ile m adım sonrasında $\mathcal{K}_{m+1}^e(A, V)$ Krylov altuzayının bir $\mathcal{V}_m$ ortonormal tabanı ve sıfırdan farklı blokları $H_{i,j}$ olan bir blok üst Hessenberg matrisi $\mathcal{H}_m$ matrisi elde edilir. $\mathcal{H}_m$ matrisinin her bir $H_{i,j}$, $(1 \leq i \leq j \leq m)$ altmatrisinin boyutu $2s$ 'dir.

A matrisinin $\mathcal{K}_m^e(A, V)$ genişletilmiş Krylov altuzayına bir kısıtlaması $\mathcal{T}_m \in \mathbb{R}^{2ms \times 2ms}$ yani $\mathcal{T}_m = \mathcal{V}_m^T A \mathcal{V}_m$ olsun. $\mathcal{T}_m$ matrisi $2s \times 2s$ bloklu bir üst blok Hessenberg matrisidir.

Aşağıdaki özellikler altuzay kavramının bir sonucudur:

Sonuç 3.2. Her $m \geq 1$ için, $\mathcal{K}_m^e$ uzayı

$$\mathcal{K}_m^e = \mathbb{K}_{2m}(A, A^{-m}V) \quad (3.34)$$

özelliğini sağlar. Burada $\mathbb{K}_{2m}$ bir blok Krylov altuzaydır ve

$$A\mathcal{K}_m^e \subseteq \mathcal{K}_{m+1}^e \quad (3.35)$$

dir [50].

Sonuç 3.2, $j = 1, 2, \dots$ olmak üzere $k > j + 1$ için $\mathcal{V}_k^T A \mathcal{V}_j = 0$ olduğundan $\mathcal{T}_m$ 'in bir blok üst Hessenberg matrisi olduğunu garantiler .

Sonuç 3.3. $\overline{\mathcal{T}}_m = \mathcal{V}_{m+1}^T A \mathcal{V}_m$ olduğu ve Algoritma 3.4'ün m adım sağlandığı kabul edilirse

$$A\mathcal{V}_m = \mathcal{V}_{m+1} \overline{\mathcal{T}}_m, \quad (3.36)$$

$$= \mathcal{V}_m \mathcal{T}_m + V_{m+1} T_{m+1,m} E_m^T, \quad (3.37)$$

yazılabilir. Burada $T_{i,j}$, $\mathcal{T}_m$ matrisinin $2s \times 2s(i, j)$ bloğudur ve $E_m = [O_{2s \times 2(m-1)s}, I_{2s}]^T$ matrisi ise I_{2ms} birim matrisinin son $2s$ sütunudur [50].

İspat. $\mathcal{V}_{m+1} = [\mathcal{V}_m, V_{m+1}]$ olduğundan,

$$\mathcal{T}_{m+1} = \mathcal{V}_{m+1}^T A \mathcal{V}_{m+1}, \quad (3.38)$$

$$= \begin{bmatrix} \mathcal{V}_m^T A \mathcal{V}_m & \mathcal{V}_m^T A V_{m+1} \\ V_{m+1}^T A \mathcal{V}_m & V_{m+1}^T A V_{m+1} \end{bmatrix}, \quad (3.39)$$

$$= \begin{bmatrix} \mathcal{T}_m & \mathcal{V}_m^T A V_{m+1} \\ V_{m+1}^T A \mathcal{V}_m & V_{m+1}^T A V_{m+1} \end{bmatrix}, \quad (3.40)$$

elde edilir. $\mathcal{T}_{m+1}$ bir blok üst Hessenberg matris olduğundan $V_{m+1}^T A \mathcal{V}_m = T_{m+1,m} E_m^T$ 'dir ve

$$\overline{\mathcal{T}}_m = \mathcal{V}_{m+1}^T A \mathcal{V}_m \quad (3.41)$$

$$= \begin{bmatrix} \mathcal{T}_m \\ T_{m+1,m} E_m^T \end{bmatrix}, \quad (3.42)$$

■

dir. $A\mathcal{K}_m^e \subseteq \mathcal{K}_{m+1}^e$ olduğu ve $\mathcal{V}_{m+1}$ 'in ortogonal olduğu göz önüne alındığında $A\mathcal{V}_m = \mathcal{V}_{m+1} L$ olacak şekilde bir L matrisi vardır. Buradan $\mathcal{V}_{m+1}^T A \mathcal{V}_m = L$ ve $L = \overline{\mathcal{T}}_m$ 'dir. Dolayısıyla $A\mathcal{V}_m = \mathcal{V}_{m+1} \mathcal{T}_{2m}$ 'dir.

3.3.2.2 Genişletilmiş Global Arnoldi Yöntemi

Genişletilmiş global Krylov altuzayı

$$\mathcal{K}_m^g(A, V) = \text{span}([V, A^{-1}V, AV, A^{-2}V, \dots, A^{m-1}V, A^{-m}V]), \quad (3.43)$$

şeklinde tanımlanır. $\mathcal{K}_m^g(A, V)$ 'nin, $\{W_1, W_2, \dots, W_m\}$, $W_i \in \mathbb{R}^{n \times 2r}$ bir F ortonormal tabanını hesaplayan genişletilmiş global Arnoldi algoritması aşağıdaki gibidir [46]:

Algoritma 3.5: Genişletilmiş global Arnoldi (EGA)

- Veriler: $A \in \mathbb{R}^{n \times n}$ bir matris, $V \in \mathbb{R}^{n \times s}$ bir matris ve m bir tamsayı,
 - $[V, A^{-1}V] = W_1 (\Omega \otimes I_r)$, global QR ayrışımını hesapla, $W_0 = []$,
 - $j = 1, 2, \dots, m$ için,
 - W_j^1 ; W_j 'nin ilk r sütunu, W_j^2 ; W_j 'nin ikinci r sütunu olmak üzere,
 - $\mathcal{W}_j = [\mathcal{W}_{j-1}, W_j]$; $U = [A^T W_j^{(1)}, A^{-1} W_j^{(2)}]$ hesapla,
 - W_{j+1} 'i elde etmek için U matrisini W_j 'ye göre F ortogonalleştir. Yani $i = 1, 2, \dots, j$ için,

$$H_{i,j} = W_i^T \diamond U,$$

$$U = U - W_i (H_{i,j} \otimes I_r),$$
 - i için döngüyü sonlandır.
 - U 'nun global QR ayrışımını hesapla, yani $U = W_{j+1} (H_{j+1,j} \otimes I_r)$,
 - j için döngüyü sonlandır.
-

Global Arnoldi algoritması ile m adım sonrasında $\mathcal{K}_m^g(A, V)$ Krylov altuzayının bir $\mathcal{W}_m$ ortonormal tabanı ve $H_{j+1,j}$, 2×2 üst üçgen matris olmak üzere $\tilde{H}_m = [H_{i,j}]$ bir üst blok $2(m+1) \times 2m$ Hessenberg matrisi elde edilir.

A matrisinin, $\mathcal{K}_m^g(A, V)$ genişletilmiş global Krylov altuzayına bir kısıtlaması $\mathcal{T}_m^g \in \mathbb{R}^{2m \times 2m}$ yani $\mathcal{T}_m^g = \mathcal{W}_m^T \diamond (A \mathcal{W}_m) = [T_{i,j}]$ olsun. Burada $T_{i,j} = V_i^T \diamond (AV_j)$, $(i, j = 1, \dots, m)$ matrisi 2×2 bloklu bir matristir.

$\overline{\mathcal{T}}_m = \mathcal{W}_{m+1}^T \diamond (A \mathcal{W}_m)$ olduğu ve EGA'nın m adımının sağlandığı kabul edilirse

$$A \mathcal{W}_m = \mathcal{W}_{m+1} (\overline{\mathcal{T}}_m \otimes I_r) \quad (3.44)$$

$$= \mathcal{W}_m (\mathcal{T}_m \otimes I_r) + W_{m+1} (T_{m+1,m} E_m^T \otimes I_r) \quad (3.45)$$

yazılabilir. Burada $E_m^T = [O_{2 \times 2(m-1)}, I_2]^T$ 'dir.

3.4 BDF Yöntemi

Bu bölümde BDF yöntemi için türetilmiş sabit ve değişken katsayılı formüllerden bahsedilecektir [38]. BDF yöntemleri, iterasyonunun her adımında ortaya çıkan lineer olmayan cebirsel denklemleri çözmek için genellikle bir Newton iterasyonu ile birlikte uygulanır. $\phi(t)$, $\{t_k, t_{k-1}, \dots, t_{k-p}\}$ noktalarında $\{x_i \approx x(t_i)\}$ 'yi interpolate eden p . dereceden bir polinom olsun. Buradan $\phi(t)$, Newton bölünmüş fark formülü kullanılarak

$$\phi(t) = \sum_{j=0}^p \prod_{i=0}^{j-1} (t - t_{k-i}) [x_k, x_{k-1}, \dots, x_{k-j}], \quad (3.46)$$

şeklinde ifade edilebilir. Buradan bölünmüş farklar rekürsif olarak

$$[x_k] = x_k, \quad (3.47)$$

$$[x_k, x_{k-1}, \dots, x_{k-i}] = \frac{[x_k, x_{k-1}, \dots, x_{k-i+1}] - [x_{k-1}, x_{k-2}, \dots, x_{k-i}]}{t_k - t_{k-i}}, \quad (3.48)$$

şeklinde tanımlanır. BDF yöntemleri

$$\dot{\phi}(t_k) = f(t_k, x_k), \quad (3.49)$$

şeklindeki bir denklemi çözmek için tanımlanır. Burada ağ aralıklarının eşit uzunlukta olmasına gerek yoktur. $\phi(t)$ 'nin diferansiyellenmesiyle

$$\dot{\phi}(t) = \sum_{j=1}^p \left(\sum_{i=0}^{j-1} \prod_{l \neq i, l=0}^{j-1} (t - t_{k-l}) \right) [x_k, x_{k-1}, \dots, x_{k-j}], \quad (3.50)$$

elde edilir ve t_k noktasında $\phi(t)$ 'nin hesaplanmasıyla

$$\sum_{j=1}^p \prod_{i=1}^{j-1} (t_k - t_{k-i}) [x_k, x_{k-1}, \dots, x_{k-j}] = f(t_k, x_k), \quad (3.51)$$

olarak yazılabilir. Eşit ağ aralıkları için yani sabit bir h adım sayısı kullanılarak (3.51) denklemi $\nabla^p f \approx h^p f^{(p)}$ olmak üzere aşağıdaki p adım sabit katsayılı BDF yöntemini oluşturur;

$$\sum_{i=1}^p \frac{1}{i} \nabla^i x_k = hf(t_k, x_k). \quad (3.52)$$

Burada geri farklar rekürsif olarak

$$\nabla^0 x_j = x_j, \quad (3.53)$$

$$\nabla^i x_j = \nabla^{i-1} x_j - \nabla^{i-1} x_{j-1}, \quad (3.54)$$

ile tanımlanır. Bu genel bir p adım yöntem olarak

$$\sum_{i=1}^p \alpha_i x_{k-i} = h\beta f(t_k, x_k), \quad (3.55)$$

şeklinde yazılabilir ve burada α_i, β yöntemin katsayılarıdır. Bu katsayılar $p = 5$ için aşağıdaki tabloda verilmiştir. Bu yöntem $p > 6$ için kararlı değildir.

Eşit aralıklı olmayan ağlarda çalışırken (3.51) denklemini (3.55)'e benzer şekilde yeniden yazarak

$$\sum_{i=0}^p \tilde{\alpha}_i x_{k-i} = h_k \tilde{\beta} f(t_k, x_k), \quad (3.56)$$

değişken katsayılı BDF'i türetebiliriz. Burada $\tilde{\alpha}_i, \tilde{\beta}$ katsayıları geçmiş $p - 1$ adıma bağlıdır. Yani

$$\tilde{\alpha}_i = \tilde{\alpha}_i(h_k, h_{k-1}, \dots, h_{k-p+1}), \quad (3.57)$$

$$\tilde{\beta} = \tilde{\beta}(h_k, h_{k-1}, \dots, h_{k-p+1}), \quad (3.58)$$

dir.

Bu tezde sabit katsayılı BDF yöntemi göz önüne alınacaktır.

Tablo 3.1 p adım BDF yöntemi katsayıları $p \leq 5$

p	β	α_0	α_1	α_2	α_3	α_4
1	1	1				
2	2/3	4/3	-1/3			
3	6/11	18/11	-9/11	2/11		
4	12/25	48/25	-36/25	16/25	-3/25	
5	60/137	300/137	-300/137	200/137	-75/137	-12/137

Büyük Ölçekli Matris Riccati Diferansiyel ve Simetrik Stein Diferansiyel Denklemleri için Düşük Ranklı Yaklaşık Çözümler

4.1 Büyük Ölçekli Matris Riccati Diferansiyel Denklemleri

Bu bölümde büyük ölçekli MRDE ele alınacaktır. Bu denklemler genellikle BDF veya Rosenbrock gibi her bir zaman aralığında büyük ölçekli matris Riccati cebirsel denklemlerinin çözümünü gerektiren yöntemlerle çözülür. Bu yöntemlere alternatif olarak integrasyondan önce problemin boyutunu indirgemeye dayanan yeni bir yaklaşım tanıtılacak, problemin boyutunun indirgemesinin ardından elde edilen matris diferansiyel problemi BDF ile çözülecek ve orjinal problem için yaklaşık bir çözüm bulunacaktır. Bu yöntemeye dayanan bazı teorik sonuçlar verilecektir.

$[0, T_f]$ zaman aralığında tanımlı bir sürekli zaman MRDE;

$$\begin{cases} \dot{X}(t) = X(t)A + A^T X(t) - X(t)BB^T X(t) + C^T C, \\ X(t_0) = X_0, \end{cases}, t \in [0, T_f] \quad (4.1)$$

dir. Burada X_0 , $n \times n$ tipinde verilen bir matris, $A \in \mathbb{R}^{n \times n}$ matrisi büyük boyutlu, seyrek ve nonsingüler bir matris, $B \in \mathbb{R}^{n \times s}$ ve $C \in \mathbb{R}^{s \times n}$ 'dir. B ve C matrislerinin full ranklı ve $s \ll n$ özelliğini sağladığı kabul edilir. Bu denklemler kontrol teoride, model indirgeme problemlerinde, diferansiyel denklemlerde ve robust kontrol problemlerinde [1, 3, 4] önemli rol oynar. MRDE'ler ayrıca dalga yayılımı ve oyun teorisinde de yer almaktadır [1, 52]. Son yıllarda büyük ölçekli matris Riccati cebirsel denklemlerinin yaklaşık çözümünü hesaplamak için birçok nümerik yöntem ortaya atıldı [41, 47, 48, 51, 53].

Genellikle A , B ve C matrisleri sonsuz boyutlu altuzaylar üzerinde tanımlı operatörlerin ayrıştırılmasıyla elde edilir. Ayrıştırma sonucu elde edilen A matrisinin elemanlarının çoğu sıfırdır dolayısıyla A seyrek bir matristir. A matrisinin elemanları köşegene

paralel belirli bir sırada yer aldığından dolayı A matrisi bant matristir ve boyutu çok büyüktür [24, 38].

Bu bölümde başlangıç Riccati diferansiyel denkleminin genişletilmiş blok Krylov altuzayına izdüşürülmesi ve elde edilen düşük boyutlu diferansiyel matris denkleminin çözülerek başlangıç matris denklemine yaklaşık çözümler elde edilmesi amaçlanmaktadır.

(4.1) denkleminin çözümünün bir ifadesi A , B ve C katsayı matrisleri üzerinde bazı varsayımlar altında mevcuttur. (4.1)'in çözümü ile ilgili temel teorem aşağıdaki gibi verilir;

Teorem 4.1. (A, B) matris çifti kararlı kılınabilir, (C, A) matris çifti gözlemlenebilir ve $X(0) > 0$ olsun. Bu durumda (4.1) denkleminin bir tek X çözümü vardır ve

$$X(t) = \tilde{X} + e^{t\tilde{A}^T} [e^{t\tilde{A}} \tilde{Z} e^{t\tilde{A}^T} + (X_0 - \tilde{X})^{-1} - \tilde{Z}]^{-1} e^{t\tilde{A}^T}, \quad (4.2)$$

şeklinde olup burada $\tilde{X}$,

$$\begin{aligned} A^T \tilde{X} + \tilde{X} A - \tilde{X} B B^T \tilde{X} + C^T C &= 0, \\ \tilde{A} &= A - B B^T \tilde{X}, \end{aligned} \quad (4.3)$$

Riccati cebirsel denkleminin çözümüdür ve $\tilde{Z}$ ise

$$A \tilde{Z} + \tilde{Z} A - B B^T = 0, \quad (4.4)$$

Lyapunov denkleminin pozitif tanımlı çözümüdür [54].

(4.2) denklemi ile verilen çözüm, çeşitli matris çarpımlarını ve bir matrisin tersinin eksponansiyelinin hesaplanmasını gerektirdiğinden büyük ölçekli problemler için uygun değildir.

Bu bölümde şu notasyonlar kullanılacaktır ;

2-norm ve F normu sırasıyla $\|\cdot\|$ ve $\|\cdot\|_F$ ve I_r , $O_{r \times l}$ matrisleri sırasıyla $r \times r$ tipinde birim matris ile $r \times l$ tipinde sıfır matrisini temsil eder.

4.1.1 Matris Riccati Diferansiyel Denklemi Çözümü için BDF Yöntemi

Bu bölümde MRDE'leri çözmek için yaygın olarak kullanılan BDF yönteminden bahsedilecektir. Literatürde integrasyon yöntemleri(BDF, Rosenbrock) doğrudan (4.1) denkleme uygulanabilir [24, 55]. Her bir t_k zamanında X , (4.1) denkleminin

çözümü olmak üzere $X(t_k)$ 'nın X_k yaklaşımı bir Riccati cebirsel denkleminin çözümüyle elde edilir [4, 24]. (4.1) denkleminde p adım BDF yöntemi uygulansın. BDF'in her bir adımında $X(t_{k+1})$ 'in X_{k+1} yaklaşımı

$$X_{k+1} = \sum_{i=0}^{p-1} \alpha_i X_{k-i} + h\beta \mathcal{F}(X_{k+1}), \quad (4.5)$$

ile verilir ve burada $h = t_{k+1} - t_k$ adım boyu, $\mathcal{F}(X)$ fonksiyonu;

$$\mathcal{F}(X) = A^T X + XA - XBB^T X + C^T C, \quad (4.6)$$

şeklinde verilir ve α_i, β katsayıları p adım BDF katsayıları olup bu katsayılar Tablo 3.1'de gösterildiği gibidir.

X_{k+1} yaklaşımı aşağıdaki matris denkleminin çözümüdür;

$$-X_{k+1} + h\beta(C^T C + A^T X_{k+1} + X_{k+1} A) - X_{k+1} B B^T X_{k+1} + \sum_{i=0}^{p-1} \alpha_i X_{k-i} = 0, \quad (4.7)$$

Bu denklem sürekli zaman Riccati cebirsel denklemi formunda yazılabilir;

$$\mathcal{A}^T X_{k+1} + X_{k+1} \mathcal{A} - X_{k+1} \mathcal{B} \mathcal{B}^T X_{k+1} + \mathcal{C}_{k+1}^T \mathcal{C}_{k+1} = 0. \quad (4.8)$$

Her bir zaman aralığında $X_k, Z_k \in \mathbb{R}^{n \times m_k}$, $m_k \ll n$ olmak üzere $X_k \approx Z_k Z_k^T$ düşük ranklı çarpımı olacak şekilde yazılabilir [23] ve (4.8) denkleminin katsayılar matrisi

$$\mathcal{A} = h\beta A - \frac{1}{2}I, \quad \mathcal{B} = \sqrt{h\beta} B \quad \text{ve} \quad \mathcal{C}_{k+1} = [\sqrt{h\beta} C, \sqrt{\alpha_0} Z_k^T, \dots, \sqrt{\alpha_{p-1}} Z_{k+1-p}^T]^T, \quad (4.9)$$

dir.

(4.8) ile verilen Riccati denklemleri, problemin düşük boyutlu olduğu durumlarda Schur ayrıştırmasına dayanan doğrudan yöntemlerle ya da genelleştirilmiş Hamiltonian özdeğerlerine [4, 56, 57] bağlı yöntemlerle çözülebilir ancak problem büyük ölçekli olduğunda bu yöntemlerle yaklaşım hesaplama zamanı ve bellek kullanımı açısından zorlu olur. Bu durumda Krylov altuzay yöntemleri ve Newton yöntemleri [3, 41, 57–59] gibi iteratif yöntemler veya ADI-tip yöntemler [39, 47, 53, 60, 61] problemin çözümünde alternatif birer seçim olabilir.

4.1.1.1 BDF+Newton+EBA Yöntemi

Bir önceki kısımda verildiği gibi her bir t_k zamanında $X(t_{k+1})$ 'in X_{k+1} yaklaşımı (4.8) ile verilen büyük ölçekli Riccati cebirsel denklemini çözerek hesaplanabilir. (4.8) denklemi lineer olmayan bir denklem olarak

$$F_k(X_{k+1}) = 0, \quad (4.10)$$

şeklinde ifade edilir. Burada F ,

$$F_k(X_{k+1}) = \mathcal{A}^T X_{k+1} + X_{k+1} \mathcal{A} - X_{k+1} \mathcal{B} \mathcal{B}^T X_{k+1} + \mathcal{C}_{k+1}^T \mathcal{C}_{k+1}, \quad (4.11)$$

ile tanımlı matris değerli bir fonksiyon olup,

$$\begin{aligned} X_k &= Z_k Z_k^T, \quad \mathcal{A} = h\beta A - \frac{1}{2}I, \quad \mathcal{B} = \sqrt{h\beta} B, \\ \text{ve } \mathcal{C}_{k+1} &= [\sqrt{h\beta} C, \sqrt{\alpha_0} Z_k^T, \dots, \sqrt{\alpha_{p-1}} Z_{k+1-p}^T]^T, \end{aligned} \quad (4.12)$$

dir. (4.10) ile verilen simetrik Riccati denkleminin düşük boyutlu olması durumunda bu denklem Bartels-Stewart gibi doğrudan bir algoritmayla çözülebilir. Büyük ölçekli problemler için ortak bir yaklaşım; Newton algoritmasının her bir iç iterasyonunda ortaya çıkan büyük ölçekli Lyapunov denkleminin nümerik çözümü için Newton-Kleinman yöntemi ile iteratif bir yöntemin kombinasyonunun uygulanması veya blok Krylov izdüşüm yönteminin (4.10)'a doğrudan uygulanmasına bağlıdır.

X_{k+1} 'e bir yaklaşım dizisi aşağıdaki gibi tanımlansın:

- $X_{k+1}^0 = X_k$,
- $(X_{k+1}^p)_{p \in \mathbb{N}}$ dizisi;

$$X_{k+1}^{p+1} = X_{k+1}^p - DF_{X_{k+1}^p}(F(X_{k+1}^p)), \quad (4.13)$$

eşitliğini sağlasın. Burada DF ifadesi F 'nin X_{k+1}^p 'de Fréchet türevi olup

$$DF_{X_{k+1}^p}(H) = (\mathcal{A} - \mathcal{B} \mathcal{B}^T X_{k+1}^p)^T H + H(\mathcal{A} - \mathcal{B} \mathcal{B}^T X_{k+1}^p), \quad (4.14)$$

ile verilir. Ardışık hesaplamalarla X_{k+1}^{p+1} 'in aşağıdaki Lyapunov denkleminin çözümü olduğu görülür;

$$(\mathcal{A} - \mathcal{B} \mathcal{B}^T X_{k+1}^p)^T X + X(\mathcal{A} - \mathcal{B} \mathcal{B}^T X_{k+1}^p) + X_{k+1}^p \mathcal{B} \mathcal{B}^T X_{k+1}^p + \mathcal{C}_{k+1}^T \mathcal{C}_{k+1} = 0. \quad (4.15)$$

C_{k+i} 'lerin içerdği tüm X_{k-i} 'lere $X_{k-i} \approx Z_{k-i} Z_{k-i}^T$ düşük ranklı faktörlerin çarpımı şeklinde yaklaşım yapılabilir ve buradan (4.15) denklemin EBA yöntemi [51, 60] uygulanarak nümerik olarak çözülür. [24]'de bahsedildiği gibi karmaşık aritmetikten kaçınmak için (4.15) denkleminin sabit terimi, pozitif α_i 'leri negatiflerinden ayırarak iki terim haline bölünebilir. Bu durum, Newton iterasyonunun her bir adımında çözülmesi gereken $[L_1]$ ve $[L_2]$ şeklinde bir çift Lyapunov denklemini oluşturur. Buradan (4.15) denkleminin X_{k+1}^{p+1} çözümü $[L_1]$ ve $[L_2]$ 'nin çözümlerinden elde edilir.

Nümerik çözümlerde bu yaklaşımdan BDF+Newton+EBA olarak söz edilmiştir.

4.1.2 Düşük Ranklı Matris Riccati Diferansiyel Denklemi'nin İzdüşümü ve Çözümü

(4.1) denklemin doğrudan integrasyon şeması uygulamak yerine öncelikle izdüşürme ile problemin boyutu indirgenecektir. A matrisinin singüler olmadığı kabul edilerek (A^T, C^T) matris çiftine EBA uygulayandığında $\mathcal{V}_m, \overline{\mathcal{T}}_m$ elde edilir. (4.1) denkleminin

$$\mathcal{V}_m^T R_m(t) \mathcal{V}_m = 0, \quad (4.16)$$

Galerkin ortogonallik şartını sağlayan bir yaklaşık çözümü $X_m(t)$ olsun ve bu çözüm

$$X_m(t) = \mathcal{V}_m Y_m(t) \mathcal{V}_m^T, \quad (4.17)$$

olarak verilsin. $R_m(t) = \dot{X}_m(t) - A^T X_m(t) - X_m(t) A + X_m(t) B B^T X_m(t) - C^T C$, $X_m(t)$ ile ilgili rezidüdür. Buradan (4.16) ve (4.17) denklemlerinden

$$\dot{Y}_m(t) - \overline{\mathcal{T}}_m Y_m(t) - Y_m(t) \overline{\mathcal{T}}_m^T + Y_m(t) B_m B_m^T Y_m(t) - C_m^T C_m = 0, \quad (4.18)$$

düşük ranklı MRDE elde edilir ve $B_m = \mathcal{V}_m^T B$, $C_m^T = \mathcal{V}_m^T C = \varepsilon_1 \Lambda_{1,1}$ şeklinde olup $\varepsilon_1 = [I_s, O_{s \times (2m-1)s}]^T$ matrisi I_{2ms} birim matrisinin ilk s sütunudur. $\Lambda_{1,1}$ matrisi QR ayrışımı yardımıyla

$$[C^T, A^{-T} C^T] = V_1 \Lambda, \quad \Lambda = \begin{pmatrix} \Lambda_{1,1} & \Lambda_{1,2} \\ 0 & \Lambda_{2,2} \end{pmatrix}, \quad (4.19)$$

şeklinde elde edilir.

(4.18) düşük boyutlu diferansiyel denklemini çözmek için BDF(p) yöntemi uygulandığında her bir zaman adımında çözülmesi gereken düşük boyutlu sürekli zaman Riccati cebirsel denklemi ortaya çıkar. Bu denklemin bir tek simetrik, yarı pozitif tanımlı, kararlı bir Y_m çözümü olduğu kabul edilerek denklem doğrudan bir

yöntem(Schur yöntemi) uygulayarak çözülebilir. m arttıkça X_m ve R_m 'in hesaplanması zor olduğundan A matrisini içeren fazladan matris çarpımlarını hesaplamadan EBA iterasyonları durdurulabilir. Bunun için ϵ seçilmiş bir durdurma sınırı olmak üzere $\|R_m\| < \epsilon$ testi kullanılabilir. Aşağıdaki sonuç X_m yaklaşık çözümünü hesaplamadan R_m rezidü normunu hesaplar.

Bu bölümde bahsedilen yöntem EBA-BDF(p) olarak adlandırılır.

Teorem 4.2. $X_m(t) = \mathcal{V}_m Y_m(t) \mathcal{V}_m^T$, EBA-BDF(p) yönteminin m .adımında elde edilen yaklaşık çözümü ve $Y_m(t)$, (4.18) ile verilen düşük boyutlu MRDE'nin çözümü olsun. Bu durumda $R_m(t)$ rezidüsü aşağıdaki eşitliği sağlar;

$$\|R_m(t)\| = \|T_{m+1,m} \hat{Y}_m(t)\|. \quad (4.20)$$

Burada $\hat{Y}_m(t)$, $2s \times 2ms$ tipinde bir matris olup Y_m 'in son s satırıdır.

İspat. (4.17) ve (4.18) ile verilen denklemler kullanılarak

$$R_m(t) = \dot{X}_m(t) - A^T \mathcal{V}_m Y_m(t) \mathcal{V}_m^T - \mathcal{V}_m Y_m(t) \mathcal{V}_m^T A + \mathcal{V}_m Y_m(t) \mathcal{V}_m^T B B^T \mathcal{V}_m Y_m(t) \mathcal{V}_m^T - C^T C, \quad (4.21)$$

elde edilir. $V_1^{(1)}$ matrisi V_1 'in ilk s sütununun oluşturduğu matris olmak üzere (4.19) ile tanımlı $\Lambda_{1,1}$ matrisi ve (4.17) kullanılarak

$$\begin{aligned} R_m(t) &= \mathcal{V}_m \dot{Y}_m(t) \mathcal{V}_m^T - (\mathcal{V}_m \mathcal{T}_m + V_{m+1} T_{m+1,m} E_m^T) Y_m(t) \mathcal{V}_m^T - \mathcal{V}_m Y_m(t) (\mathcal{T}_m^T \mathcal{V}_m^T \\ &\quad + E_m T_{m+1,m} V_{m+1}^T) + \mathcal{V}_m Y_m(t) B_m B_m^T Y_m(t) \mathcal{V}_m^T - V_1^{(1)} \Lambda_{1,1} \Lambda_{1,1}^T V_1^{(1)T}, \\ &= \mathcal{V}_{m+1} \mathcal{J}_m(t) \mathcal{V}_{m+1}^T, \end{aligned} \quad (4.22)$$

şeklinde yazılabilir. Burada

$$\mathcal{J}_m(t) = \begin{bmatrix} (\dot{Y}_m(t) - \mathcal{T}_m Y_m(t) - Y_m(t) \mathcal{T}_m^T \\ + Y_m(t) B_m B_m^T Y_m(t) - \varepsilon_1 \Lambda_{1,1} \Lambda_{1,1}^T \varepsilon_1^T & Y_m(t) E_m T_{m+1,m}^T \\ T_{m+1,m} E_m^T Y_m(t) & 0 \end{bmatrix}, \quad (4.23)$$

dir.

$C_m = \varepsilon_1 \Lambda_{1,1}$ ve $Y_m(t)$, (4.18) ile verilen indirgenmiş MRDE'nin çözümü olduğundan

$$R_m(t) = \mathcal{V}_{m+1} \begin{bmatrix} 0 & Y_m(t) E_m T_{m+1,m}^T \\ T_{m+1,m} E_m^T Y_m(t) & 0 \end{bmatrix} \mathcal{V}_{m+1}^T, \quad (4.24)$$

dir. $\hat{Y}_m(t) = E_m^T Y_m(t)$ matrisi $Y_m(t)$ 'in son $2s$ satırı olmak üzere

$$\|R_m(t)\| = \|T_{m+1,m} E_m^T Y_m(t)\| = \|T_{m+1,m} \hat{Y}_m(t)\|, \quad (4.25)$$

dir. ■

Teorem 4.3, her bir iterasyonda $X_m(t)$ yaklaşık çözümünü hesaplamadan, yakınsaklık sağlandığında iterasyonu durdurmamızı sağladığından dolayı pratikte önemlidir.

Aşağıdaki algoritma ile yöntem özetlenebilir:

Algoritma 4.1: MRDE için EBA-BDF(p) Yöntemi

- Veriler: $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times s}$, $C \in \mathbb{R}^{s \times n}$ katsayı matrisleri, $X(0) = 0$, $tol > 0$ tolerans değeri, m_{max} bir tamsayı ve bir h adım boyu $nstep = T_f/h$,
 - $X_1, \dots, X_{p-1}$ değerlerini $X_j \approx Z_j Z_j^T (*)$ şeklinde düşük ranklı çarpımlar olarak hesapla,
 - $m = 1, \dots, m_{max}$ için, $\mathcal{K}_m^e(A, C^T) = Range[C^T, A^{-1}C, \dots, A^{-(m-1)}C^T, A^{m-1}C^T]$ altuzayının bir $\mathcal{V}_m = [V_1, \dots, V_m]$ tabanını hesapla,
 - $B_m = \mathcal{V}_m^T B$, $C_m^T = \mathcal{V}_m^T C^T$ ve $\mathcal{T}_m = \mathcal{V}_m^T A \mathcal{V}_m$ olsun,
 - $\dot{Y}_m(t) = \mathcal{T}_m Y_m(t) + Y_m(t) \mathcal{T}_m^T - Y_m(t) B_m B_m^T Y_m(t) + C_m^T C_m$ izdüşürülmüş MRDE'ye BDF(p) yöntemini uygula,
 - $Y_m(T_f)$ hesaplandığında eğer $\|R_m(T_f)\| = \|T_{m+1,m} E_m^T \hat{Y}_m(T_f)\| < tol$ ise o zaman $m_{max} = m$ 'dir,
 - m için döngüyü sonlandır,
 - $X_m(T_f) \approx Z_m(T_f) Z_m^T(T_f)$ hesapla, $(X_m(T_f) = \mathcal{V}_m Y_m(T_f) \mathcal{V}_m^T (**))$ çarpımını hesaplamadan).
-

Bu algoritma ile ilgili bazı önemli noktalara değinilirse;

(*) BDF(p) integrasyon şemasını başlatmak için $X_1, \dots, X_{p-1}$, $p - 1$ yaklaşımları düşük dereceli integrasyon şemalarıyla hesaplanabilir. Bu testte $p = 2$ seçildi ve $X_1 \approx Z_1 Z_1^T$ düşük ranklı faktörlerin çarpımı olarak BDF(1) ile hesaplandı.

(**) BDF(p) yönteminin uygulanmasında sabit terim olan C_k her bir t_k zamanında yeniden hesaplanır. $X_k \approx Z_k Z_k^T$ yaklaşımının içerdiği $\sqrt{\alpha_0} Z_k$ düşük ranklı çarpımı her bir t_k zaman adımında yeniden hesaplanır. Dolayısıyla C_k terimi her bir zaman

adımında yeniden hesaplanır. Yaklaşık çözüm, düşük boyutlu matrisin kesikli değer ayrışımı (truncated singular value) kullanılarak elde edildiğinden $X_k(t) = \mathcal{V}_k Y_k(t) \mathcal{V}_k^T$ hesaplamasına gerek yoktur. $Y_k(t) = U \Sigma(t) U^T$ tekil değer ayrışımında Σ matrisi, $Y_k(t)$ 'nin tekil değerlerinden oluşan bir köşegen matrisidir. U_l matrisi U matrisinin ilk l sütunundan oluşan $2k \times l$ matrisi olsun. $Y_k(t)$ 'nin kesikli tekil değer ayrışımı $\Sigma_l = \text{diag}[\sigma_1, \sigma_2, \dots, \sigma_l]$ olmak üzere $Y_k(t) \approx U_l \Sigma(t) U_l^T$ olarak yazılır. $Z_k(t) = \mathcal{V}_k U_l \Sigma(t)_l^{1/2}$ denirse

$$X_k(t) \approx Z_k(t) Z_k(t)^T, \quad (4.26)$$

yazılır.

Aşağıdaki sonuç $X_m(t)$ yaklaşımının pertörb edilmiş MRDE'nin tam çözümü olduğunu gösterir.

Teorem 4.3. $X_m(t)$, (4.17) denklemi ile verilen bir yaklaşık çözüm olsun. Bu durumda

$$\dot{X}_m(t) = (A - F_m)^T X_m(t) + X_m(t)(A - F_m) - X_m(t) B B^T X_m(t) + C^T C, \quad (4.27)$$

elde edilir ve burada $F_m = V_m T_{m+1,m} V_{m+1}^T$ 'dir.

İspat. $X_m(t) = V_m Y_m(t) V_m^T$ olsun. (4.18) ile verilen denklem soldan V_m sağdan V_m^T ile çarpılır ve (3.36) kullanılırsa

$$\begin{aligned} \dot{X}_m(t) = & [A^T \mathcal{V}_m - V_{m+1} T_{m+1,m} E_m^T] Y_m(t) \mathcal{V}_m^T + \mathcal{V}_m Y_m(t) [A^T \mathcal{V}_m - V_{m+1} T_{m+1,m} E_m^T]^T \\ & - \mathcal{V}_m Y_m(t) \mathcal{V}_m^T B B^T \mathcal{V}_m Y_m(t) \mathcal{V}_m^T + C^T C, \end{aligned} \quad (4.28)$$

elde edilir. Buradan $F_m = V_m T_{m+1,m}^T V_{m+1}^T$ ve $\mathcal{V}_m E_m = V_m$ olduğu kabul edilirse (4.3) ile verilen denklem elde edilir. ■

Teorem 4.4. $X(t)$, (4.1) denkleminin tam çözümü ve $X_m(t)$, m . adımda elde edilen yaklaşık çözüm olsun. $R_m(t) = \dot{X}_m(t) - A^T X_m(t) - X_m(t)A + X_m(t) B B^T X_m(t) - C^T C$ rezidü olmak üzere $E_m(t) = X(t) - X_m(t)$ hatası aşağıdaki denklemi sağlar:

$$\dot{E}_m(t) = (A^T - X(t) B B^T) E_m(t) + E_m(t) (A - B B^T X(t)) + E_m(t) B B^T E_m(t) + R_m(t). \quad (4.29)$$

İspat. (4.1) probleminden rezidü denkleminin çıkarılmasıyla;

$$\begin{aligned} \dot{X}(t) - R_m(t) = & X(t)A + A^T X(t) - X(t) B B^T X(t) + C^T C - \dot{X}_m(t) + A^T X_m(t) \\ & + X_m(t)A - X_m(t) B B^T X_m(t) + C^T C, \end{aligned} \quad (4.30)$$

denklemini elde edilir. Buradan

$$\begin{aligned}\dot{X}(t) - \dot{X}_m(t) &= X(t)A + A^T X(t) - X(t)B B^T X(t) + C^T C + R_m(t) - A^T X_m(t) \\ &\quad - X_m(t)A + X_m(t)B B^T X_m(t) - C^T C,\end{aligned}\quad (4.31)$$

yazılabilir. (4.31) denkleminin düzenlenmesiyle

$$\begin{aligned}\dot{X}(t) - \dot{X}_m(t) &= (X(t) - X_m(t))[A - B B^T X(t)] + [A^T - X(t)B B^T](X(t) - X_m(t)) \\ &\quad (X(t) - X_m(t))B B^T (X(t) - X_m(t)) + R_m(t),\end{aligned}\quad (4.32)$$

elde edilir. (4.32) denkleminde $X(t) - X_m(t) = E_m(t)$ yazılırsa (4.29) denklemini elde edilir. ■

4.1.3 İzdüşürülmüş Lineer Kuadratik Regülatör Problemi

Bu bölümde (2.3) ile verilen denklemin EBA yöntemiyle yaklaşık bir maliyeti hesaplanacaktır.

EBA algoritmasının m . adımında (2.3) orjinal dinamik sisteminin $\mathcal{H}_m^e(A^T, C^T)$ genişletilmiş blok Krylov altuzayı üzerine izdüşüm yapılarak elde edilen düşük boyutlu dinamik sistem aşağıdaki gibi olsun:

$$\begin{cases} \dot{\tilde{x}}_m(t) = \mathcal{T}_m \tilde{x}_m(t) + B_m \tilde{u}_m(t), \\ \tilde{y}_m(t) = C_m \tilde{x}_m(t), \end{cases} \quad \tilde{x}_m(0) = x_{m,0}, \quad (4.33)$$

Burada $B_m = \mathcal{V}_m^T B$, $C_m^T = \mathcal{V}_m^T C = \varepsilon_1 \Lambda_{1,1}$ ve $\tilde{x}_m(0) = \mathcal{V}_m^T x_0$ ' dir. Burada $X_m(t) = \mathcal{V}_m \tilde{x}_m(t)$, $x(t)$ durumuna bir yaklaşımdır.

Sonuç 4.1. EBA algoritmasının m . adımında elde edilen $(\mathcal{T}_m, B_m)$ ' in kararlı kılınabilir ve $(C_m, \mathcal{T}_m)$ 'in kontrol edilebilir olduğu kabul edilsin. Düşük boyutlu sonlu LQR problemi

$$J_m(x_{m,0}, \tilde{u}_m) = \int_0^{T_f} (\tilde{y}_m(t)^T \tilde{y}_m(t) + \tilde{u}_m(t)^T \tilde{u}_m(t)) dt, \quad (4.34)$$

(4.33) dinamik kısıtlamaları altında minimize edilmesi olarak ifade edilir. Bu durumda (4.34) maliyet fonksiyonunu minimize eden bir tek $\tilde{u}_{*,m}$ optimal geri bildirimi;

$$\tilde{u}_{*,m}(t) = -B_m \tilde{Y}_m(t) \hat{x}_m(t), \quad (4.35)$$

ile verilir ve burada $\tilde{Y}_m(t) = Y_m(T_f - t)$ 'dir ve Y_m , (4.18) denkleminin kararlı bir tek

çözümüdür.

İspat. $Y_m(t)$ 'in (4.18) ile verilen izdüşürülmüş Riccati denkleminden elde edilmesi ve aynı genişletilmiş Krylov altuzayı üzerine (2.15) denkleminin izdüşürülmesinden elde edilen (4.33) dinamik sistemi göz önüne alınarak yapılabilir. ■

Optimal izdüşürülmüş kontrol,

$$\tilde{x}_m(t) = e^{tA} \tilde{x}_{m,0} + \int_0^t e^{(t-\tau)T_m} B_m \tilde{u}_{*,m}(\tau) d\tau, \quad (4.36)$$

ile ifade edilebilen $\dot{\tilde{x}}_m(t) = (A - BB^T \tilde{Y}_m(t)) \tilde{x}_m(t)$ 'yi sağlar ve optimal maliyet x_0 başlangıç durumunun aşağıdaki kuadratik fonksiyonuyla verilir:

$X_m(t) = V_m Y_m(t) V_m^T$, (4.1) denkleminde elde edilen yaklaşık çözüm olmak üzere $X_m(t) = P_m(T_f - t)$ olduğundan

$$\begin{aligned} J_m(x_{m,0}, \tilde{u}_m) &= x_{m,0}^T \tilde{Y}_m(0) x_{m,0} = x_0^T \mathcal{V}_m \tilde{Y}_m(0) \mathcal{V}_m^T x_0, \\ &= x_0^T P_m(0) x_0 = x_0^T X_m(T_f) x_0, \end{aligned} \quad (4.37)$$

ile verilir. Bu da indirgenmiş maliyetin, minimal maliyetin bir yaklaşımı olduğunu açıkça gösterir.

$x_m(t) = \mathcal{V}_m \tilde{x}_m(t)$ orjinal durum vektörü $x(t)$ 'ye bir yaklaşımdır. Bununla ilgili geri bildirim kuralı $u_m(t) = -B^T P_m(t) x_m(t)$ ile belirlenir ve $P_m(t)$, (2.5) matris Riccati diferansiyel denkleminin bir yaklaşık çözümüdür. Buradan

$$\begin{aligned} J(x_0, u_m) &= \inf \left\{ \int_0^{T_f} (x_m(t)^T C^T C x_m(t) + u_m(t)^T u_m(t)) dt \right\}, \\ &= \inf \left\{ \int_0^{T_f} (\tilde{x}_m(t)^T \mathcal{V}_m^T C^T C \mathcal{V}_m x_m(t) + \tilde{x}_m(t)^T \mathcal{V}_m^T P_m(t)^T B B^T P_m(t) \mathcal{V}_m \tilde{x}_m(t)) dt \right\}, \end{aligned} \quad (4.38)$$

dir.

$P_m(t) = \mathcal{V}_m \tilde{Y}_m(t) \mathcal{V}_m^T$ olduğu kullanılarak ve $\tilde{u}_{*,m}(t) = -B_m \tilde{Y}_m(t) \tilde{x}_m(t)$ olmak üzere

$$J(x_0, u_m) = \inf \left\{ \int_0^{T_f} (\tilde{Y}_m(t)^T \tilde{Y}_m(t) + \tilde{u}_m(t)^T \tilde{u}_m(t)) dt \right\} \quad (4.39)$$

$$= J_m(x_{m,0}, \tilde{u}_{*,m},$$

yazılır.

Sonuç 4.2. (A, B) kararlı kılınabilir olsun. $X(t)$, (4.1) denklemini sağlayan simetrik, pozitif matrislerin bir kümesi olsun. Her x_0 başlangıç durumu için

$$x_0^T X(t) x_0 \leq M \|x_0\|^2, \quad \forall t \geq 0 \quad (4.40)$$

olacak şekilde bir M skaleri vardır ve her t için $\|X(t)\| \leq M$ 'dir.

Dolayısıyla $t \rightarrow \infty$ için, aşağıdaki Riccati cebirsel denklemini sağlayan bir X_∞ çözümüne yakınsar $\left(X_\infty = \lim_{t \rightarrow \infty} X(t) \right)$ ve X_∞ bu denklemin tek pozitif ve kararlı çözümüdür.

$$A^T X_\infty + X_\infty A - X_\infty B B^T X_\infty + C^T C = 0. \quad (4.41)$$

4.2 Büyük Ölçekli Matris Simetrik Stein Diferansiyel Denklemleri

Bu bölümde büyük ölçekli MSSDE ele alınacaktır. Başlangıç değer problemi, genişletilmiş blok Krylov altuzayına izdüştürülerek düşük boyutlu MSSDE elde edilir. İzdüşümle elde edilen düşük boyutlu matris problemi BDF yöntemi kullanılarak çözülecek ve orijinal problem için yaklaşık bir çözüm bulunacaktır.

$[0, T_f]$ zaman aralığında tanımlı bir simetrik Stein diferansiyel denklemi

$$\begin{cases} \dot{X}(t) = X(t) - A X(t) A^T + B B^T \\ X(t_0) = X_0 \end{cases}, t \in [0, T_f], \quad (4.42)$$

ile verilir ve X_0 , $n \times n$ tipinde bir matris, $A \in \mathbb{R}^{n \times n}$ matrisi büyük boyutlu, seyrek, singüler olmayan ve $B \in \mathbb{R}^{n \times s}$, $s \ll n$, full ranklı bir matristir.

(4.42) matris denklemi ayrıca ayrık zaman Lyapunov denklemi olarak da bilinir. MSSDE lineer kontrolde ve ayrık zaman büyük ölçekli dinamik sistemler için filtreleme teorisinde önemli rol oynar [39, 62]. Son zamanlarda büyük ölçekli diferansiyel matris denklemleri büyük dikkat çekmiş ve çözümleri için bazı nümerik

yöntemler(izdüşüm yöntemleri [63], ayrıştırma şemaları [27]) ileri sürülmüştür. (4.42) denklemindeki matrisler eğer düşük boyutlu iseler Bartels-Stewart ya da Hessenberg-Schur yöntemleri kullanılarak denklem çözülebilir. A ve B matrisleri kısmi diferansiyel denklemler tarafından modellenen kontrol problemlerinden elde edilen büyük sistemin katsayı matrisleridir.

(4.42) problemi,

$$\begin{cases} \dot{x}(t) = \mathcal{A}x(t) + b(t), \\ x(0) = \text{vec}(X_0), \end{cases} \quad (4.43)$$

lineer adi diferansiyel denklemine denktir. Burada $\mathcal{A} = A(t) \otimes A^T(t) - I_{n^2}$ ve $b(t) = \text{vec}(BB^T)$ 'dir. (4.43) probleminin düşük boyutlu olduğu durumlarda problemi bir integrasyon yöntemi kullanarak çözmek mümkündür. Ancak bu yaklaşım büyük problemler için uygun değildir.

Bu çalışmada (4.42) denklemini çözmek için EBA-BDF(p) yöntemi kullanılacaktır. İlk olarak başlangıç MSSDE genişletilmiş blok Krylov altuzayı üzerine izdüşürülecek [48, 50, 51, 64], elde edilen düşük boyutlu diferansiyel matris denklemini çözülecek ve başlangıç denklemine yaklaşık çözümler elde edilecektir. İstenen doğruluk elde edilemediği sürece izdüşüm altuzayının boyutu artırılır ve iterasyon tekrar başlatılır.

4.2.1 Matris Simetrik Stein Diferansiyel Denklemini'nin İzdüşümü ve Çözümü

(A, B) matris çiftine EBA uygulayarak $\mathcal{V}_m$ ve $\mathcal{T}_m = \mathcal{V}_m^T A \mathcal{V}_m$ elde edilsin. (4.42) denkleminin

$$\mathcal{V}_m^T R_m(t) \mathcal{V}_m = 0, \quad t \in [t_0, T_f], \quad (4.44)$$

Galerkin ortogonallik şartını sağlayan bir yaklaşık çözümü olsun ve bu çözüm

$$X_m(t) = \mathcal{V}_m Y_m(t) \mathcal{V}_m^T, \quad (4.45)$$

ile verilsin. Burada $R_m(t) = \dot{X}_m(t) + AX_m(t)A^T - X_m(t) - BB^T$ ile tanımlı rezidüdür. (4.44) ve (4.45) denklemlerinden

$$\dot{Y}_m(t) = Y_m(t) - \mathcal{T}_m Y_m(t) \mathcal{T}_m^T + B_m B_m^T, \quad (4.46)$$

düşük ranklı MSSDE elde edilir ve burada $\mathcal{T}_m = \mathcal{V}_m^T A \mathcal{V}_m$ ve $B_m = \mathcal{V}_m^T B = \varepsilon_1 \Lambda_{1,1}$ olup

$\Lambda_{1,1}$ matrisi

$$[B, A^{-1}B] = V_1 \Lambda, \quad \Lambda = \begin{pmatrix} \Lambda_{1,1} & \Lambda_{1,2} \\ 0 & \Lambda_{2,2} \end{pmatrix}, \quad (4.47)$$

dir.

Elde edilen (4.46) düşük boyutlu MSSDE, BDF gibi iyi bilinen bir integrasyon şeması ile çözülür. Bununla birlikte m arttıkça $X_m(t)$ ve $R_m(t)$ 'in hesaplanması giderek zorlaşmaktadır ve bazı seçilmiş durdurma sınırı değerleri için beklenen yaklaşım elde edildiğinde, izdüşüm altuzayının boyutunun artırılması durdurulur.

Yaklaşık çözüm, düşük ranklı faktörlerin çarpımı olarak ifade edilecektir.

Aşağıdaki sonuç MSSDE için $X_m(t)$ yaklaşık çözümünü hesaplamadan $R_m(t)$ rezidü normunu hesaplar.

Teorem 4.5. $X_m(t) = \mathcal{V}_m Y_m(t) \mathcal{V}_m^T$, EBA-BDF(p)'nin m . adımında elde edilen yaklaşık çözüm ve $Y_m(t)$, (4.46) düşük ranklı MSSDE'nin çözümü olsun. Bu durumda $R_m(t)$ rezidüsü aşağıdaki eşitliği sağlar;

$$\|R_m(t)\|_F^2 = 2 \left\| \mathcal{T}_m Y_m(t) E_m T_{m+1,m} \right\|_F^2 + \left\| T_{m+1,m} E_m^T Y_m(t) E_m T_{m+1,m}^T \right\|_F^2. \quad (4.48)$$

Burada $T_{i,j}$, $2s \times 2s$ bir matris olup $\mathcal{T}_m$ matrisinin (i, j) bloğudur ve $E_m = [O_{2s \times 2(m-1)s}, I_{2s}]^T$ matrisi ise I_{2ms} birim matrisinin son $2s$ sütunudur.

İspat. (4.45) ve (4.46) ile verilen denklemlerden

$$R_m(t) = \dot{X}_m(t) + A X_m(t) A^T - X_m(t) - B B^T, \quad (4.49)$$

elde edilir. $V_1^{(1)}$ matrisi, V_1 ' in ilk s sütunundan oluşan matristir ve (4.47) ile tanımlı $\Lambda_{1,1}$ matrisi ve (4.45) kullanılarak

$$\begin{aligned} R_m(t) &= \mathcal{V}_m \dot{Y}_m(t) \mathcal{V}_m^T + (\mathcal{V}_m \mathcal{T}_m + V_{m+1} T_{m+1,m} E_m^T) Y_m(t) (\mathcal{T}_m^T \mathcal{V}_m^T + E_m T_{m+1,m} V_{m+1}^T) \\ &\quad - \mathcal{V}_m Y_m(t) \mathcal{V}_m^T - V_1^{(1)} \Lambda_{1,1} \Lambda_{1,1}^T V_1^{(1)T}, \\ &= \mathcal{V}_{m+1} \mathcal{J}_m(t) \mathcal{V}_{m+1}^T, \end{aligned} \quad (4.50)$$

elde edilir ve

$$\mathcal{J}_m(t) = \begin{bmatrix} \dot{Y}_m(t) - Y_m(t) + \mathcal{T}_m Y_m(t) \mathcal{T}_m^T - \varepsilon_1 \Lambda_{1,1} \Lambda_{1,1}^T \varepsilon_1^T & \mathcal{T}_m Y_m(t) E_m T_{m+1,m}^T \\ T_{m+1,m} E_m^T Y_m(t) \mathcal{T}_m^T & T_{m+1,m} E_m^T Y_m(t) E_m T_{m+1,m}^T \end{bmatrix}, \quad (4.51)$$

dir.

$B_m = \varepsilon_1 \Lambda_{1,1}$ ve $Y_m(t)$ (4.46)'in çözümü olduğundan

$$R_m(t) = \mathcal{V}_{m+1} \begin{bmatrix} 0 & \mathcal{T}_m Y_m(t) E_m T_{m+1,m}^T \\ T_{m+1,m} E_m^T Y_m(t) \mathcal{T}_m & T_{m+1,m} E_m^T Y_m(t) E_m T_{m+1,m}^T \end{bmatrix} \mathcal{V}_{m+1}^T, \quad (4.52)$$

olup bu ifadenin normunun hesaplanmasıyla

$$\|R_m(t)\|_F^2 = 2 \left\| \mathcal{T}_m Y_m(t) E_m T_{m+1,m}^T \right\|_F^2 + \left\| T_{m+1,m} E_m^T Y_m(t) E_m T_{m+1,m}^T \right\|_F^2, \quad (4.53)$$

elde edilir. ■

Bu teorem her bir iterasyonda $X_m(t)$ yaklaşık çözümünü hesaplamadan yakınsaklık sağlandığında iterasyonu durdurmamızı sağladığından dolayı pratikte önemlidir.

Teorem 4.6, $X_m(t)$ yaklaşımının pertörb edilmiş MSSDE'nin bir tam çözümü olduğunu gösterir.

Teorem 4.6. $X_m(t)$, (4.45) ile verilen bir yaklaşık çözüm olsun. Buradan

$$\dot{X}_m(t) = X_m(t) - (A - F_m)X_m(t)(A - F_m)^T + B B^T, \quad (4.54)$$

elde edilir ve $F_m = V_m T_{m+1,m} V_{m+1}^T$ 'dir.

İspat. $X_m(t) = \mathcal{V}_m Y_m(t) \mathcal{V}_m^T$ olsun. (4.46) ile verilen indirgenmiş MSSDE denklemi soldan $\mathcal{V}_m$ sağdan $\mathcal{V}_m^T$ ile çarpılır ve (3.36) kullanılırsa

$$\dot{X}_m(t) = X_m(t) - [A \mathcal{V}_m - V_{m+1} T_{m+1,m} E_m^T] Y_m [A \mathcal{V}_m - V_{m+1} T_{m+1,m} E_m^T]^T + B B^T, \quad (4.55)$$

bulunur. Buradan $F_m = V_m T_{m+1,m}^T V_{m+1}^T$ ve $\mathcal{V}_m E_m = V_m$ olduğu kabul edilirse

$$\dot{X}_m(t) = X_m(t) - (A - F_m)X_m(t)(A - F_m)^T + B B^T, \quad (4.56)$$

elde edilir. ■

Teorem 4.7. $X(t)$ 'nin (4.42) denkleminin tam çözümü ve $X_m(t)$ 'nin m . adımdan elde edilen yaklaşık çözüm olduğu kabul edilsin. $R_m(t) = \dot{X}_m(t) - X_m(t) + A X_m(t) A^T - B B^T$ rezidü olmak üzere $E_m(t) = X(t) - X_m(t)$ hatası aşağıdaki denklemi sağlar;

$$\dot{E}_m(t) = E_m(t) - A E_m(t) A^T + R_m(t). \quad (4.57)$$

İspat. (4.42) probleminden rezidü denkleminin çıkarılmasıyla;

$$\begin{aligned} \dot{X}(t) - R_m(t) &= X(t) - AX(t)A^T + BB^T - \dot{X}_m(t) - AX_m(t)A^T \\ &\quad + X_m(t) + BB^T, \end{aligned} \quad (4.58)$$

denklemini elde edilir. Buradan

$$\dot{X}(t) - \dot{X}_m(t) = X(t) - AX(t)A^T + BB^T + R_m(t) + AX_m(t)A^T - X_m(t) - BB^T, \quad (4.59)$$

yazılabilir. (4.59) denkleminin düzenlenmesiyle

$$\dot{X}(t) - \dot{X}_m(t) = [X(t) - X_m(t)] - A[X(t) - X_m(t)]A^T + R_m(t), \quad (4.60)$$

elde edilir. (4.60) denkleminde $X(t) - X_m(t) = E_m(t)$ yazılırsa (4.57) denklemini elde edilir. ■

4.2.2 Düşük Boyutlu Simetrik Stein Diferansiyel Denklemini'nin Çözümü

$X(t)$, (4.42) denkleminin çözümü olmak üzere $X(t_k)$ 'nin X_k yaklaşımını hesaplamak için her bir t_k zamanında (4.46) denkleminde p adım BDF yöntemi uygulanacaktır. BDF'in her bir iterasyonunda $Y_m(t_{k+1})$ 'in $Y_{m,k+1}$ yaklaşımı

$$Y_{m,k+1} = \sum_{i=0}^{p-1} \alpha_i Y_{m,k-i} + h\beta \mathcal{F}(Y_{m,k+1}), \quad (4.61)$$

ile verilir ve burada $h_k = t_{k+1} - t_k$ adım boyu, α_i ve β katsayıları p adım BDF katsayıları olup bu katsayılar tablo 3.1'de verilmiştir. $\mathcal{F}(Y)$ fonksiyonu;

$$\mathcal{F}(Y) = Y - \mathcal{T}_m Y \mathcal{T}_m^T + B_m B_m^T, \quad (4.62)$$

ile tanımlıdır. $Y_{m,k+1}$, aşağıdaki matris denkleminin yaklaşık bir çözümüdür;

$$-Y_{m,k+1} + h\beta(BB^T - \mathcal{T}_m Y_{m,k+1} \mathcal{T}_m^T + Y_{m,k+1}) + \sum_{i=0}^{p-1} \alpha_i Y_{m,k-i} = 0. \quad (4.63)$$

Bu denklem MSSDE formunda yazılabilir;

$$Y_{m,k+1} - \mathbf{T}_m Y_{m,k+1} \mathbf{T}_m^T + Q = 0. \quad (4.64)$$

Burada, $Z_{m,k} \in \mathbb{R}^{n \times m_k}$, $m_k \ll n$ olmak üzere $Y_{m,k+1}$ her bir zaman adımında düşük ranklı faktörlerin çarpımı olarak $Y_{m,k} \approx Z_{m,k} Z_{m,k}^T$ şeklinde yazılabilir ve katsayı

matrisleri

$$\mathbf{T}_m = \sqrt{\frac{h_k}{1-h_k\beta}} \mathcal{T}_m, \quad Q = \frac{1}{h_k\beta-1} \left(h_k B_m B_m^T + \sum_{i=0}^{p-1} \alpha_i Y_{m,k-i} \right), \quad (4.65)$$

dir.

Not. $X_m(t)$ yaklaşık çözümünü düşük ranklı iki matrisin çarpımı olarak ifade edilebilir. Bunun için $Y_m(t) = U \Sigma U^T$ tekil değer ayrışımı dikkate alınsın. Burada Σ , $Y_m(t)$ 'nin tekil değerlerinin bir köşegen matrisidir. U_l matrisi U matrisinin ilk l sütunundan oluşan $2m \times l$ matrisi olsun. $Y_m(t)$ 'nin kesikli tekil değer ayrışımı $\Sigma_l = \text{diag}[\sigma_1, \sigma_2, \dots, \sigma_l]$ olmak üzere $Y_m(t) \approx U_l \Sigma_l U_l^T$ olarak yazılır. $Z_m(t) = \mathcal{V}_m U_m \Sigma_m^{1/2}$ denirse

$$X_m(t) \approx Z_m(t) Z_m(t)^T$$

yazılır.

Aşağıdaki algoritma ile yöntem özetlenebilir:

Algoritma 4.2: MSSDE için EBA-BDF(p) Yöntemi

- Veriler: $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times s}$ katsayı matrisleri, $s \ll n$, $X(0) = X_0$, tol bir tolerans, $m_{\max}$ bir tamsayı,
- $m = 1, \dots, m_{\max}$ için,
- Genişletilmiş blok Arnoldi algoritması uygulayarak $\mathcal{K}_m^e(A, B) = \text{range}[B, A^{-1}B, \dots, A^{-m}B, A^{m-1}B]$ uzayının bir $\mathcal{V}_m = [V_1, V_2, \dots, V_m]$ ortonormal tabanını ve üst blok $\mathcal{T}_m$ Hessenberg matrisini hesapla,
- $B_m = \mathcal{V}_m^T B$ olsun ve

$$\dot{Y}_m(t) - Y_m(t) + \mathcal{T}_m Y_m(t) \mathcal{T}_m^T - B_m B_m^T = 0, t \in [t_0, T_f]$$

düşük boyutlu simetrik Stein diferansiyel denklemini çözmek için BDF yöntemi kullan,

- Eğer $\|R_m(t)\|_F^2 = 2 \left\| \mathcal{T}_m Y_m(t) E_m T_{m+1,m}^T \right\|_F^2 + \left\| T_{m+1,m} E_m^T Y_m(t) E_m T_{m+1,m}^T \right\|_F^2 < tol$ ise iterasyonu durdur ve $X_m^b(t)$ yaklaşık çözümünü hesapla.
-

Simetrik Olmayan Matris Stein Diferansiyel Denklemleri

Bu bölümde simetrik olmayan matris Stein diferansiyel denklemi (NMSDE) ele alınacaktır. Bu denklemler dinamik sistemler için kontrol ve filtreleme teorisi problemlerinde önemli bir rol oynar [1, 4]. NMSDE'nin çözümü için integrasyondan önce problemin boyutunu indirgemeye dayanan EGA-BDF(p) yaklaşımı uygulanacak ve problemin boyutunun indirgemesinin ardından elde edilen matris diferansiyel problemi BDF ile çözülecek ve orjinal problem için yaklaşık bir çözüm bulunacaktır.

$[0, T_f]$ zaman aralığında bir NMSDE

$$\begin{cases} \dot{X}(t) = X(t) - AX(t)B + EF^T \\ X(t_0) = X_0 \end{cases}, t \in [0, T_f], \quad (5.1)$$

ile verilir. X_0 , $n \times s$ tipinde verilen bir matris, A ve B matrisleri seyrek, reel matrisler olup sırasıyla $n \times n$ ve $s \times s$ tipinde kare matrislerdir. E ve F matrisleri sırasıyla $n \times p$ ve $s \times p$, $p \ll n, s$, matrislerdir.

(5.1) başlangıç probleminin sırasıyla (A, E) ve (B^T, F) matris çiftlerinin oluşturduğu $\mathcal{K}_m^g(A, E)$ ve $\mathcal{K}_m^g(B^T, F)$ genişletilmiş global Krylov altuzayları üzerine izdüşürülmesi ile $\mathcal{V}_m^g = [V_1^g, \dots, V_m^g]$ ve $\mathcal{W}_m^g = [W_1^g, \dots, W_m^g]$, F ortonormal tabanları elde edilir[49].

Buradan

$$\begin{aligned} A\mathcal{V}_m^g &= \mathcal{V}_{m+1}^g (\overline{\mathcal{T}_m^A} \otimes I_p), \\ &= \mathcal{V}_m^g (\mathcal{T}_m^A \otimes I_p) + V_{m+1}^g T_{m+1,m}^A (E_m^{g^T} \otimes I_p), \end{aligned} \quad (5.2)$$

ve

$$\begin{aligned} B^T \mathcal{W}_m^g &= \mathcal{W}_{m+1}^g (\overline{\mathcal{T}_m^B} \otimes I_p), \\ &= \mathcal{W}_m^g (\mathcal{T}_m^B \otimes I_p) + W_{m+1}^g T_{m+1,m}^B (E_m^{g^T} \otimes I_p), \end{aligned} \quad (5.3)$$

olup burada

$$\mathcal{T}_m^A = \mathcal{V}_m^{gT} \diamond (A \mathcal{V}_m^g), \quad (5.4)$$

ve

$$\mathcal{T}_m^B = \mathcal{W}_m^{gT} \diamond (B^T \mathcal{W}_m^g), \quad (5.5)$$

dir. V_1^g ve W_1^g ilk blokları sırasıyla $[E, A^{-1}E]$ ve $[F, B^{-T}F]$ matrislerinin global QR ayrışımıyla elde edilir yani $[E, A^{-1}E] = V_1^g (\omega^A \otimes I_p)$ ve $[F, B^{-T}F] = W_1^g (\omega^B \otimes I_p)$ olup $\omega^A = [w_{i,j}^A]_{i,j=1,2}$ ve $\omega^B = [w_{i,j}^B]_{i,j=1,2}$ matrisleri 2×2 üst üçgensel matrislerdir. Global QR ayrışımının özelliklerinden

$$\mathcal{V}_m^T \diamond E = w_{1,1}^A e_1^{2m} \text{ ve } \mathcal{W}_m^T \diamond F = w_{1,1}^B e_1^{2m}, \quad (5.6)$$

elde edilir [49]. Buradan (5.1)'in istenen düşük ranklı yaklaşık çözümü

$$X_m^g(t) = \mathcal{V}_m^g (Y_m^g(t) \otimes I_p) \mathcal{W}_m^{gT}, \quad (5.7)$$

ile verilir ve $Y_m^g(t) \in \mathbb{R}^{2m \times 2m}$, aşağıdaki düşük boyutlu NMSDE'nin çözümüdür;

$$\dot{Y}_m^g(t) = Y_m^g(t) - \mathcal{T}_{m,A} Y_m^g(t) \mathcal{T}_{m,B}^T + \tilde{E}_m \tilde{F}_m^T. \quad (5.8)$$

Burada $\tilde{E}_m = \mathcal{V}_m^{gT} \diamond E$ ve $\tilde{F}_m = \mathcal{W}_m^{gT} \diamond F$ dir.

$X_m^g(t)$ yaklaşık çözümü için rezidü $R_m^g(t) = \dot{X}_m^g(t) - X_m^g(t) + A X_m^g(t) B - E F^T$ ile tanımlıdır.

(5.8) düşük boyutlu denklemi Hessenberg-Schur gibi bilinen yöntemler aracılığıyla çözülebilir.

Hesaplamayı sınırlandırmak ve $X_m^g(t)$ yaklaşık çözümünü hesaplamadan iterasyonu durdurmak için rezidü normuna bir üst sınır aşağıdaki teoremle verilebilir:

Teorem 5.1. $R_m^g(t)$, m . adımda elde edilen rezidü olmak üzere

$$\begin{aligned} \|R_m^g(t)\|_F^2 &\leq \left\| \mathcal{T}_m^A Y_m^g(t) E_m^g T_{m+1,m}^{B^T} \right\|_F^2 + \left\| T_{m+1,m}^A E_m^T Y_m^g(t) \mathcal{T}_m^{B^T} \right\|_F^2 \\ &+ \left\| T_{m+1,m}^A E_m^{gT} Y_m^g(t) E_m^g T_{m+1,m}^{B^T} \right\|_F^2, \end{aligned} \quad (5.9)$$

İspat. (5.2), (5.3) ve (5.7) denklemleri kullanılarak

$R_m^g(t) = \dot{X}_m^g(t) - X_m^g(t) + AX_m^g(t)B - EF^T$ rezidüsü

$$\begin{aligned} R_m^g(t) = & \mathcal{V}_m^g(\dot{Y}_m^g(t) \otimes I_p) \mathcal{W}_m^{gT} - \mathcal{V}_m^g(Y_m^g(t) \otimes I_p) \mathcal{W}_m^{gT} + [\mathcal{V}_m^g(\mathcal{T}_m^A \otimes I_p) \\ & + V_{m+1}^g T_{m+1,m}^A (E_m^{gT} \otimes I_p)] (Y_m^g(t) \otimes I_p) [\mathcal{W}_m^g(\mathcal{T}_m^B \otimes I_p) \\ & + W_{m+1}^g T_{m+1,m}^B (E_m^{gT} \otimes I_p)] - EF^T, \end{aligned} \quad (5.10)$$

dir. Buradan $Y_m^g(t)$ 'nin düşük boyutlu (5.8) probleminin çözümü olduğu göz önüne alınırsa rezidü

$$R_m^g(t) = \mathcal{V}_{m+1}^g \left(\begin{bmatrix} 0 & \mathcal{T}_m^A Y_m^g(t) E_m^{gT} T_{m+1,m}^B \\ T_{m+1,m}^A E_m^{gT} Y_m^g(t) \mathcal{T}_m^{BT} & T_{m+1,m}^A E_m^{gT} Y_m^g(t) E_m^{gT} T_{m+1,m}^B \otimes I_p \end{bmatrix} \right) \mathcal{W}_{m+1}^{gT}, \quad (5.11)$$

olur. Buradan Sonuç A.2 uygulanırsa her $t \in [t_0, T_f]$ için aşağıdaki üst sınır elde edilir:

$$\begin{aligned} \|R_m^g(t)\|_F^2 \leq & \left\| \mathcal{T}_m^A Y_m^g(t) E_m^{gT} T_{m+1,m}^B \right\|_F^2 + \left\| T_{m+1,m}^A E_m^{gT} Y_m^g(t) E_m^{gT} T_{m+1,m}^B \right\|_F^2 \\ & + \left\| T_{m+1,m}^A E_m^{gT} Y_m^g(t) E_m^{gT} T_{m+1,m}^B \right\|_F^2. \end{aligned} \quad (5.12)$$

■

5.1 Düşük Boyutlu Simetrik Olmayan Stein Diferansiyel Denkleminin Çözümü

Herbir t_k zamanında Y_m^g , (5.8)'nin çözümü olmak üzere $Y_m^g(t_k)$ 'nin yaklaşımı $Y_{m,k}^g$ ile gösterilsin. Buradan (5.8) denkleminde p adım BDF yöntemi uygulansın.

BDF'in her bir iterasyonunda $Y_m^g(t_{k+1})$ 'in $Y_{m,k+1}^g$ yaklaşımı

$$Y_{m,k+1}^g = \sum_{i=0}^{p-1} \alpha_i Y_{m,k-i}^g + h_k \beta \mathcal{F}(Y_{m,k+1}^g), \quad (5.13)$$

ile verilir ve burada $h_k = t_{k+1} - t_k$ adım boyu, α_i ve β katsayıları p adım BDF katsayıları olup bu katsayılar tablo 3.1'de verilmiştir. $\mathcal{F}(Y)$ fonksiyonu;

$$\mathcal{F}(Y) = Y - \mathcal{T}_m^A Y \mathcal{T}_m^{BT} + \tilde{E}_m \tilde{F}_m^T B^T, \quad (5.14)$$

ile tanımlıdır. $Y_{m,k+1}^b$, aşağıdaki matris denkleminin yaklaşık bir çözümüdür;

$$-Y_{m,k+1}^b + h_k \beta Y_{m,k+1}^g - h_k \beta \mathcal{T}_m^A Y \mathcal{T}_m^{BT} + h_k \beta \tilde{E}_m \tilde{F}_m^T B^T + \sum_{i=0}^{p-1} \alpha_i Y_{m,k-i}^g = 0. \quad (5.15)$$

Bu denklem NMSDE formunda yazılabilir;

$$Y_{m,k+1}^g - \mathbf{T}_m^A Y_{m,k+1}^g \mathbf{T}_m^{B^T} + \mathbf{E}_m \mathbf{F}_m^T = \mathbf{0}. \quad (5.16)$$

Burada, $\tilde{U}_{m,k} \in \mathbb{R}^{n \times m_k}$, $\tilde{V}_{m,k} \in \mathbb{R}^{p \times m_k}$, $m_k \ll n, p$, olmak üzere $Y_{m,k+1}^g$, her bir t_k zaman adımında düşük ranklı faktörlerin çarpımı olarak $Y_{m,k}^g \approx \tilde{U}_{m,k} \tilde{V}_{m,k}^T$ şeklinde yazılabilir ve katsayı matrisleri

$$\mathbf{T}_m^A = \sqrt{\frac{h_k \beta}{h_k \beta - 1}} \mathcal{T}_m^A, \quad \mathbf{T}_m^{B^T} = \sqrt{\frac{h_k \beta}{h_k \beta - 1}} \mathcal{T}_m^{B^T}, \quad (5.17)$$

$$\tilde{\mathbf{E}}_{m,k+1} = \left[\sqrt{\frac{h_k \beta}{h_k \beta - 1}} \tilde{E}_m, \sqrt{\alpha_0} \tilde{U}_{m,k}, \dots, \sqrt{\alpha_{p-1}} \tilde{U}_{m,k+1-p}^T \right], \quad (5.18)$$

$$\tilde{\mathbf{F}}_{m,k+1} = \left[\sqrt{\frac{h_k \beta}{h_k \beta - 1}} \tilde{F}_m^T, \sqrt{\alpha_0} \tilde{V}_{m,k}^T, \dots, \sqrt{\alpha_{p-1}} \tilde{V}_{m,k+1-p}^T \right], \quad (5.19)$$

dir. Aşağıdaki algoritma ile yöntem özetlenebilir:

Algoritma 5.1: Simetrik Olmayan Stein Diferansiyel Denklemi için EGA-BDF(p) Yöntemi

- Veriler: $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{s \times s}$, $E \in \mathbb{R}^{n \times p}$, $F \in \mathbb{R}^{s \times p}$, $p \ll n, s$, katsayı matrisleri, $X(0) = X_0$, tol bir tolerans, m_{max} bir tamsayı,
- (A, E) ve (B^T, F) matris çiftlerine genişletilmiş global Arnoldi algoritması uygulayarak $\mathcal{V}_m^b = [V_1^b, V_2^b, \dots, V_m^b]$, $\mathcal{W}_m^b = [W_1^b, W_2^b, \dots, W_m^b]$, F ortonormal tabanlarını ve $\mathcal{T}_m^A$ ve $\mathcal{T}_m^B$ üst blok Hessenberg matrisini hesapla,

•

$$\dot{Y}_m^g(t) - Y_m^g(t) - \mathcal{T}_m^A Y_m^g(t) \mathcal{T}_m^{B^T} - \tilde{E}_m \tilde{F}_m^T$$

düşük boyutlu problemi çözmek için BDF yöntemi kullan,

- Eğer $\|R_m(t)\|_F^2 \leq tol$ ise iterasyonu durdur ve $X_m^g(t)$ yaklaşık çözümünü hesapla.
-

6.1 Büyük Ölçekli Matris Riccati Diferansiyel Denklemini için Sayısal Örnekler

Bu bölümde

$$\begin{cases} \dot{X}(t) = X(t)A + A^T X(t) - X(t)B B^T X(t) + C^T C, \\ X(t_0) = X_0, \end{cases} \quad (6.1)$$

ile verilen MRDE için ele alınan sayısal örneklerde EBA-BDF(p) ile BDF(p)+Newton+EBA yöntemlerinin sonuçları yazıldı ve karşılaştırıldı. Ayrıca [27]'de Stillfjord tarafından tanıtılan düşük ranklı ikinci dereceden ayrıştırma şeması (LRSOP) ile [24]'de Benner ve Mena tarafından tanıtılan BDF+LR+ADI yöntemi ile EBA-BDF(p) yönteminin kıyaslandığı bir örnek verildi. İlk üç örnekte, BDF(p) integrasyon şemasının adım sayısı $p = 2$ olarak alındı. Tüm uygulamalar Intel Core i7 işlemci ve 8 GB Ram hafızalı bir bilgisayarda yapıldı. Algoritmalar MATLAB R2014b'de yazıldı.

Örnek 1

Bu örnekte konveksiyon ısı akış modeli incelenmiştir.

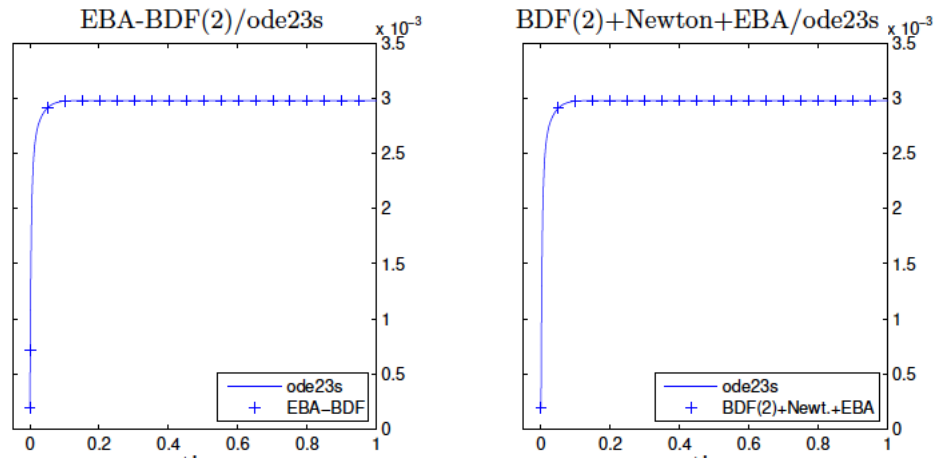
$$\begin{cases} \dot{x}(t) = Ax(t) + Bu(t), \\ y(t) = Cx(t) + Du(t), \end{cases} \quad (6.2)$$

LTI sistemi göz önüne alınsın. A matrisi, $[0, 1] \times [0, 1]$ birim karesinde $u(x, 0) = u(x, 1) = 0$ ve $u(0, y) = u(1, y) = 0$, homojen Dirichlet sınır koşulları altında

$$L_A = \Delta u - f_1(x, y) \frac{\partial u}{\partial x} + f_2(x, y) \frac{\partial u}{\partial y} + g_1(x, y), \quad (6.3)$$

operatörlerinin 5 nokta merkezi fark ayrışımı ile elde edilir. Her bir yöndeki iç grid noktalarının sayısı n_0 ve A matrisinin boyutu ise $n = n_0^2$ 'dir. Bu örnek

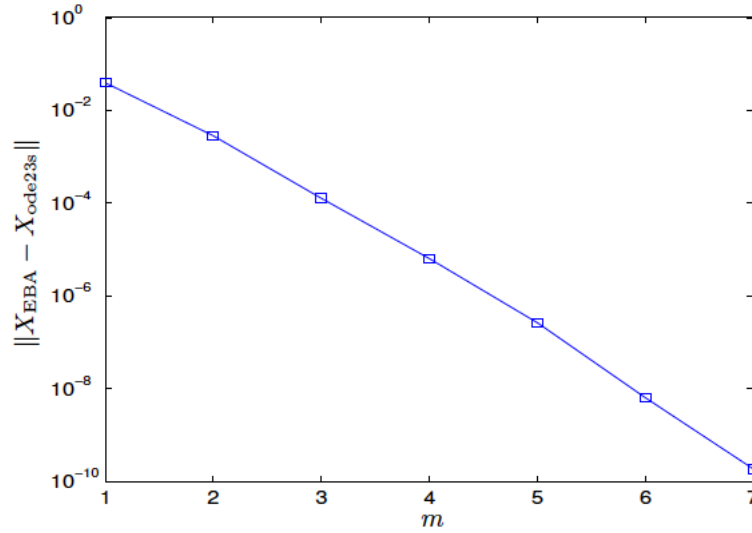
için $f_1(x,y) = 10xy$, $f_2(x,y) = e^{x^2y}$ ve $g_1(x,y) = 20y$ alınsın. B ve C matrisleri $[0,1]$ aralığından seçilen rastgele değerlerden oluşan matrisler ve sütun sayıları $r = s = 2$ olsun. $Z_0 \in \mathbb{R}^{n \times 2}$ olmak üzere başlangıç şartı $X_0 = X(0)$, $X_0 = Z_0 Z_0^T$ olsun. Genişletilmiş blok Krylov altuzayına izdüşürme ile BDF yöntemlerinin kombinasyonu ile oluşan yöntem uygulansın. Rezidü üzerindeki iterasyonu durdurma testi için doğruluk sınırı 10^{-10} seçildi ve bu testte sabit zaman adımı h ile 2 adım BDF şeması uygulandı. $X_1 \approx X(t_0 + h)$ yaklaşımı kapalı Euler şeması ve EBA ile hesaplandı [26]. Literatürde büyük ölçekli MRDE'lerin tam çözümü yoktur. EBA-BDF(2) yöntemi ile önerdiğimiz yaklaşımların güvenilir sonuçlar verip vermediğini kontrol etmek amacıyla bu yaklaşımlar, MATLAB'ın diferansiyel denklemleri çözmesi için tasarlanmış "ode23s" çözücüsü ile karşılaştırıldı. MRDE'yi vektörize ederek çözen ve Rosenbrock integrasyon şemasına dayanan bu yöntem, büyük ölçekli problemler için uygun değildir. MATLAB çözücüsü, ode23s yöntemi için hafıza sınırlaması nedeniyle en fazla 100×100 tipinde bir A matrisini çözebilir.



Şekil 6.1 $t \in [0, 1]$ zaman aralığı için $X_{1,1}(t)$ değerleri

Şekil 6.1 ile verilen grafiklerde A matrisi 49×49 tipinde $t \in [0, 1]$, adım sayısı $h = 10^{-3}$ alınmış ve EBA-BDF(2), BDF(2)+Newton+EBA yöntemleriyle elde edilen çözümlerin $X_{1,1}(t)$ bileşenleri ile ve ode23s yöntemiyle elde edilen çözümün $X_{1,1}(t)$ bileşenleri karşılaştırılmıştır.

Doğruluk bakımından tüm yöntemlerin benzer sonuçlar verdiği gözlemlenebilir. Şekil 6.2'de X_{EBA} , EBA-BDF yöntemi ile elde edilen çözümü temsil etmek üzere $t = 1$ zamanında EBA yönteminin iterasyon sayısına karşılık $X_{EBA} - X_{ode23s}$ farkının normu karakterize edilir. Bu tablodan izdüşüm uzayının makul bir boyutu ($m = 7$) için iki yaklaşık çözümün çok yakın olduğu anlaşılır. Bu testte ode23s çözücüsü 510 saniyede yaklaşık çözümü hesaplarken EBA-BDF(2) yöntemi 7 saniyede yaklaşık çözümü hesaplamıştır.



Şekil 6.2 m genişletilmiş blok Arnoldi iterasyon sayısı için $\|X_{EBA}(T_f) - X_{ode23s}(T_f)\|_F$ hata normu

Tablo 6.1 ode23s, EBA-BDF(2), BDF(2)+Newton+EBA(2) için hesaplama zamanları

size A	ode23s	EBA-BDF(2)	BDF(2)+Newton+EBA
49×49	30.2	6.5	37.1
100×100	929.3	12.3	41.6
900×900	—	45.2	926.1

Tablo 6.1’de denklemin $[0, 1]$ zaman aralığında $h = 1e - 3$ seçilerek elde edilen hesaplama zamanları yer alır. Tablodan problemin boyutu arttıkça integrasyondan önce problemin boyutunun indirgenmesinin tercih edilebilirliği anlaşılır. Yine de BDF(2)+Newton+EBA yöntemi, EBA-BDF yöntemine göre hesaplama süresi bakımından daha çok zaman alsa da büyük ölçekli problemler için bir alternatif olabilir.

Örnek 2

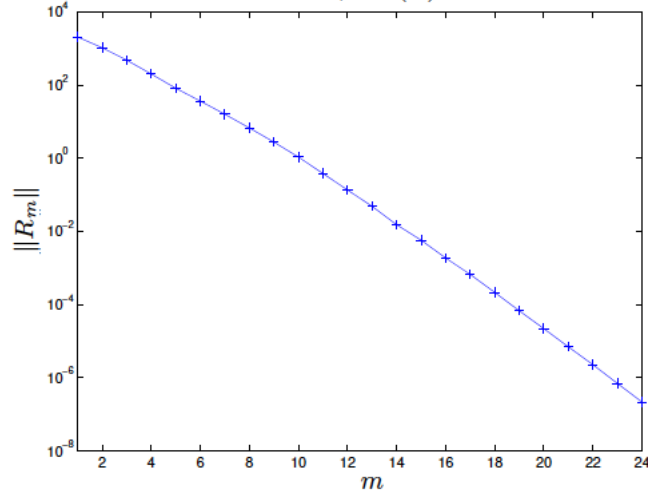
Bu problemde Örnek 1’in katsayı matrislerinin büyük boyutlu hali ele alındı ve EBA-BDF(2) yöntemi uygulandı. Tüm testler için zaman aralığı $[0, 1]$ ve $h = 0,001$ ’dir. Bu örnekte Tablo 6.2’de hesaplama zamanı, rezidü normu ve A matrisinin farklı boyutları için genişletilmiş Arnoldi iterasyonu sayısı belirtilmektedir.

Şekil 6.3, $n = 6400$ boyutlu bir A matrisi için EBA’nın m iterasyon sayısına karşılık gelen R_m rezidü normunu gösterir.

Bu grafikten de görüldüğü gibi m iterasyon sayısı arttıkça yaklaşık çözümün kesinliği de artar.

Tablo 6.2 EBA-BDF(2) yöntemi için hesaplama zamanları, rezidü normları ve genişletilmiş Arnoldi iterasyonu sayısı m

size A	Hesaplama zamanı(s)	Rezidü normu	İterasyon sayısı (m)
100×100	12.3	3.1×10^{-9}	9
900×900	45.2	3.2×10^{-8}	15
2500×2500	85.	4.8×10^{-8}	19
6400×6400	159.7	1.8×10^{-7}	24
10000×10000	195.3	3.7×10^{-8}	26



Şekil 6.3 m genişletilmiş blok Arnoldi iterasyon sayısına karşılık gelen $\|R_m(T_f)\|$ rezidü normu

Örnek 3

Bir boyutlu ısı akışının otonom lineer kuadratik optimal kontrol probleminden gelen bu örnek [24]'den alındı.

$$\frac{\partial}{\partial t} x(t, \eta) = \frac{\partial^2}{\partial \eta^2} x(t, \eta) + b(\eta)u(t),$$

$$x(t, 0) = x(t, 1) = 0, \quad t > 0,$$

$$x(0, \eta) = x_0(\eta), \quad \eta \in [0, 1],$$

$$y(x) = \int_0^1 c(\eta)x(t, \eta)d\eta, \quad x > 0$$

Birinci dereceden B -spline'larına dayanan standart sonlu eleman yaklaşımı kullanılarak aşağıdaki adi diferansiyel denklem sistemi elde edilir.

$$\begin{cases} M\dot{x}(t) = Kx(t) + Fu(t), \\ y(t) = Cx(t). \end{cases} \quad (6.4)$$

Burada M ve K matrisleri,

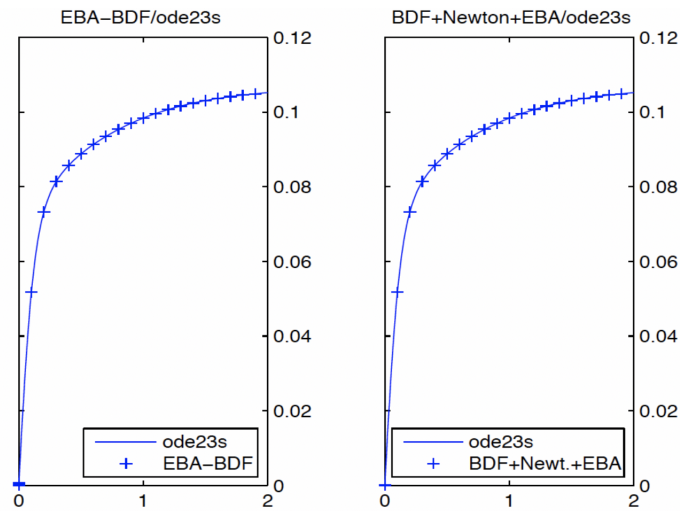
$$M = \frac{1}{6n} \begin{pmatrix} 4 & 1 & & & & \\ 1 & 4 & 1 & & & \\ & & \ddots & \ddots & \ddots & \\ & & & \ddots & \ddots & \\ & & & & \ddots & \\ & & & & & 1 & 4 & 1 \\ & & & & & & 1 & 4 \end{pmatrix}, K = -\alpha n \begin{pmatrix} 2 & -1 & & & & \\ -1 & 2 & -1 & & & \\ & & \ddots & \ddots & \ddots & \\ & & & \ddots & \ddots & \\ & & & & \ddots & \\ & & & & & -1 & 2 & -1 \\ & & & & & & -1 & 2 \end{pmatrix}$$

dir. Yarı-kapalı Euler yöntemi kullanılarak aşağıdaki ayrık dinamik sistem elde edilir.

$$(M - \Delta t K)^{-1} x_{k+1} = M x_k + \Delta t F u_k.$$

$A = -(M - \Delta t K)M$ ve $B = \Delta t(M - \Delta t K)F$ olsun. $n \times s$, F matrisi ve $s \times n$, C matrisinin girdileri; $[0,1]$ üzerinde düzgün dağılmış rastgele değerler olsun. Başlangıç koşulu $X_0 = 0_{n \times n} = Z_0 Z_0^T$, $Z_0 = 0_{n \times 2}$ olarak seçilsin. Bu örnekte $s = 2$, $\Delta t = 0.01$, $\alpha = 0.05$ alınmıştır.

Şekil 6.4'de EBA-BDF(2), BDF(2)+Newton+EBA ve ode23s çözücüsü ile elde edilen yaklaşık çözümlerin $[0, 2]$ aralığı üzerindeki $X_{1,1}$ ilk bileşenleri işaretlendi.

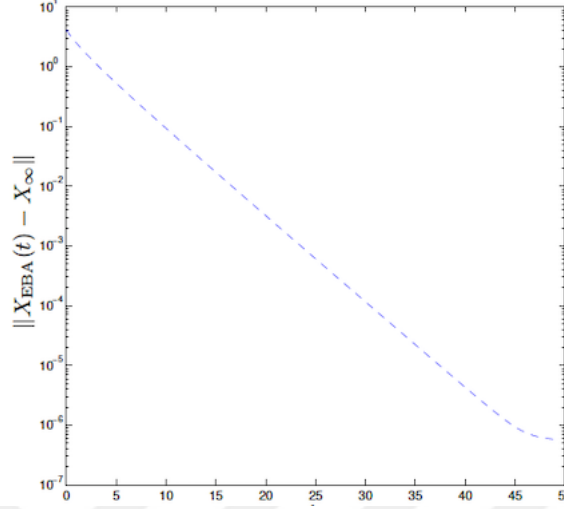


Şekil 6.4 size(A) = 49×49 ve $t \in [0, 2]$ için hesaplanan $X_{1,1}(t)$ bileşenleri ode23s ve EBA-BDF(2)(solda), ode23s ve BDF(2)+Newton+EBA (sağda)

Şekil 6.5'de, belirli şartlar altında, $\lim_{t \rightarrow \infty} X(t) = X_\infty$, X_∞ 'un $A^T X + X A - X B B^T X + C^T C = 0$ cebirsel Riccati denkleminin tek pozitif ve kararlı çözümü olduğunu iddia eden Sonuç 4.2 örneklendirilmiştir. Bunun için $t \in [0, 50]$ olmak üzere $X_0 = 0_n$ başlangıç değeri olmak üzere $X_{EBA}(t) - X_\infty$ farkının

normu t zaman parametresinin bir fonksiyonu olarak çizildi.

X_{EBA} , EBA-BDF(2) yöntemiyle hesaplanan matristir. Bu örnekte $\|X_{EBA}(50) - X_\infty\| \approx 4 \times 10^{-5}$ 'dir.



Şekil 6.5 $\text{size}(A) = 400 \times 400$ ve $t \in [0, 50]$ için $\|X_{EBA-BDF}(t) - X_\infty\|$ hata normları

Tablo 6.3, $[0, 1]$ zaman aralığı üzerinde çeşitli boyutlardaki katsayı matrislerinin hesaplama zamanları, rezidü normları ve Arnoldi iterasyonlarının sayısını gösterir. Burada hesaplama zamanı açısından BDF(2)+Newton+EBA yönteminden açıkça daha iyi olan EBA-BDF(2)'den elde edilen sonuçlar verildi. Örneğin, BDF+Newton+EBA yöntemi $\text{size}(A) = 1600 \times 1600$ olduğu durumda 521 saniye, $\text{size}(A) = 2500 \times 2500$ durumunda 5000 saniyeden daha fazla hesaplama gerektirir.

Tablo 6.3 EBA-BDF(2) yöntemi hesaplama zamanları(s), rezidü normları ve Arnoldi iterasyonları sayısı(m)

size A	Hesaplama zamanı(s)	Rezidü normu	İterasyon sayısı (m)
1600×1600	14.2	3.2×10^{-12}	10
2500×2500	17.1	7×10^{-12}	9
4900×4900	25.6	1.3×10^{-11}	9
6400×6400	42.2	8.5×10^{-12}	10
10000×10000	43.9	4.5×10^{-11}	8

Örnek 4

Son örnekte EBA-BDF(1) yöntemi, iyi bilinen çelik profillerin optimal soğutma problemine uygulandı. Matrisler IMTEK¹ koleksiyonundan alındı. $[0, 250]$ zaman aralığında EBA-BDF(1) yöntemiyle LRSOP ayrıştırma yöntemi [27] ve

¹<https://portal.uni-freiburg.de/imteksimulation/downloads/benchmark>

BDF(1)+LR+ADI [24] yöntemleri, $n = 1357$ ve $n = 5177$ için kıyaslandı. X_0 başlangıç değeri, $Z_0 = 0_{n \times l}$ olmak üzere $X_0 = Z_0 Z_0^T$ olarak seçildi. BDF(1)+LR+ADI, BDF(1) integrasyon şemasının MRDE'ye uygulanmasına dayanır. 4. Bölümde tanıtılan BDF(p)+Newton+EBA yöntemi her zaman adımında büyük ölçekli cebirsel bir Riccati denkleminin nümerik çözümünü belirtir. Bu çözüm LR-ADI yöntemiyle uygulanır ve M.E.S.S.² paketinde mevcuttur. LRSOP ve BDF(1)+LR+ADI için sütun sıkıştırma yöntemlerinin iterasyonu durdurma sınırı 10^{-8} olarak belirlenmiştir. Zaman adımı $h = 1$ ve EBA-BDF(1) yöntemi için Arnoldi durdurma testi sınırı 10^{-7} olarak seçildi. İzdüşürülmüş denklemler MATLAB'ın "care" çözücüsü ile nümerik olarak hesaplandı.

Tablo 6.4 Çelik profillerin optimal soğutma problemi için hesaplama zamanları(s)

size n	EBA-BDF(1)	LRSOP	BDF(1)+LR+ADI
1357	48.9	129.2	135.77
5177	165.24	2051.3	> 10.000

Tablo 6.4'de $T_f = 250$ için elde edilen hesaplama zamanları verildi. EBA-BDF(1) yöntemi için m Arnoldi iterasyon sayısı olmak üzere $T_f = 250$ zamanındaki rezidü normu $n = 1357$, $m = 20$ için $\|R_m\| = 2.8 \times 10^{-7}$ ve $n = 5177$, $m = 30$ için $\|R_m\| = 3.7 \times 10^{-8}$ 'dir. B 'nin sütun sayısı $s = 7$ olarak seçilmiştir. Bu örnek EBA-BDF yönteminin özellikle büyük ölçekli problemler için hesaplama zamanı bakımından etkili olduğunu göstermektedir.

6.2 Büyük Ölçekli Matris Simetrik Stein Diferansiyel Denklemi için Sayısal Örnekler

Bu bölümde

$$\begin{cases} \dot{X}(t) = X(t) - AX(t)A^T + BB^T \\ X(t_0) = X_0 \end{cases}, \quad t \in [t_0, T_f] \quad (6.5)$$

ile verilen MSSDE denklemi için EBA-BDF(2) yöntemi ile "ode23s" çözücüsü kıyaslandı ve tüm uygulamalar Intel Core i7 işlemci ve 8 GB Ram hafızalı bir bilgisayarda yapıldı. Algoritmalar MATLAB R2014b'de yazıldı.

Örnek 5

$$\begin{cases} \dot{x}(t) = \mathcal{A}x(t) + b(t), \\ x(0) = \text{vec}(X_0). \end{cases} \quad (6.6)$$

²<http://www.mpi-magdeburg.mpg.de/projects/mess/>

sistemi göz önüne alınsın. A matrisi

$$L_A = \Delta u - f_1(x, y) \frac{\partial u}{\partial x} + f_2(x, y) \frac{\partial u}{\partial y} + g_1(x, y), \quad (6.7)$$

operatörlerinin $[0, 1] \times [0, 1]$ birim karesinde $u(x, 0) = u(x, 1) = 0$ ve $u(0, y) = u(1, y) = 0$ homojen Dirichlet sınır koşulları altında 5 nokta merkezi fark ayrışımı ile elde edilir. Her bir yönde iç grid noktalarının sayısı n_0 ve A matrisinin boyutu ise $n = n_0^2$ 'dir. Bu örnek için $f_1(x, y) = x + 10y^2$, $f_2(x, y) = \sqrt{2x^2 + y^2}$, $g_1(x, y) = x^2 - y^2$ alındı. B matrisi, $[0, 1]$ aralığından seçilen rastgele değerlerden oluşan bir matris ve sütun sayısı $s = 2$ olsun. Bu problem için zaman aralığı $[0, 2]$ ve başlangıç koşulu, $Z_0 = 0_{n \times 2}$ olmak üzere $X_0 = Z_0 Z_0^T$, $X_0 = X(0)$ olarak alınsın. $p = 2$ için EBA-BDF yöntemi uygulansın. Burada rezidüyü hesaplamak için kullanılan sınır değer $tol = 10^{-10}$, zaman adımı $h = 10^{-3}$ olarak seçilsin.

Literatürde büyük ölçekli MSSDE'nin tam çözümü yoktur. EBA-BDF(2) yöntemi ile bulunan yaklaşımların güvenilirliğini test etmek amacıyla MATLAB'ın ode23s çözücüsüyle karşılaştırılacaktır. ode23s çözücüsü için en fazla 100×100 tipinde bir A matrisi ele alındı.

Şekil 6.6 ile verilen grafikte A matrisi 49×49 tipinde, $t \in [0, 2]$, $h = 10^{-3}$ alınmış ve yukarıda bahsedilen yöntemlerle elde edilen çözümün $X_{1,1}$ bileşeni ile ode23s yöntemiyle elde edilen çözüm karşılaştırılmıştır. EBA-BDF(2) yönteminde testi durdurmak için kullanılan sınır değeri 5.10^{-12} ve kesikli tekil değer ayrışımı için eşik değeri 10^{-14} olarak seçildi.

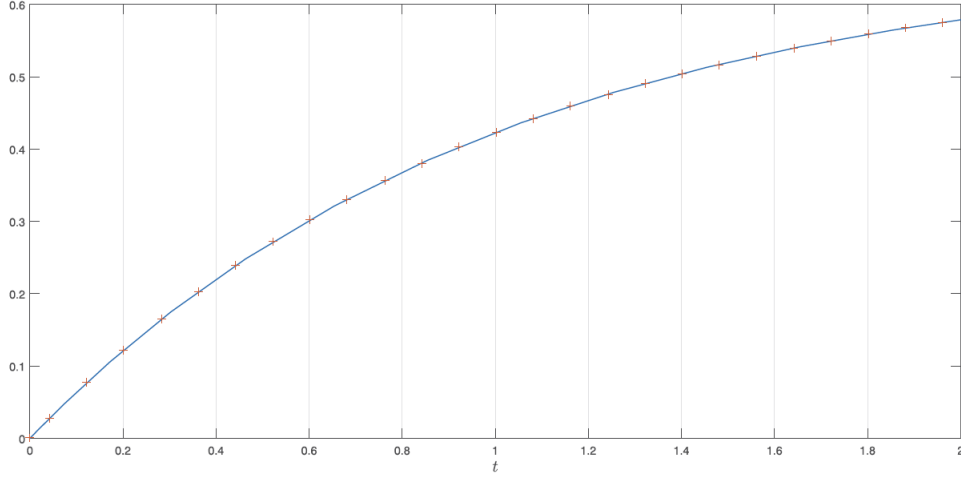
Şekil 6.6'da görüldüğü gibi iki nümerik yöntem de benzer sonuçlar vermektedir. Hesaplama zamanı bakımından EBA-BDF(2) yönteminin ode23s yönteminden daha hızlı olduğu (hesaplama zamanları sırasıyla 3.6s ile 915.5s) görülmektedir.

Tablo 6.5'de hesaplama zamanları, T_f 'deki rezidü normları ve A matrisinin farklı boyutları için EBA iterasyon sayısı verilmektedir.

Tablo 6.5 A 'nın boyutu, hesaplama zamanları(s), rezidü normu ve Arnoldi iterasyonları sayısı(m)

size A	Hesaplama zamanı (s)	Rezidü normu	İterasyon sayısı m
2500×2500	4.82s	4.85×10^{-13}	5
6400×6400	5.96 s	7.98×10^{-13}	6
14400×14400	5.74 s	27.93×10^{-11}	6
40000×40000	21.13 s	4.79×10^{-12}	10

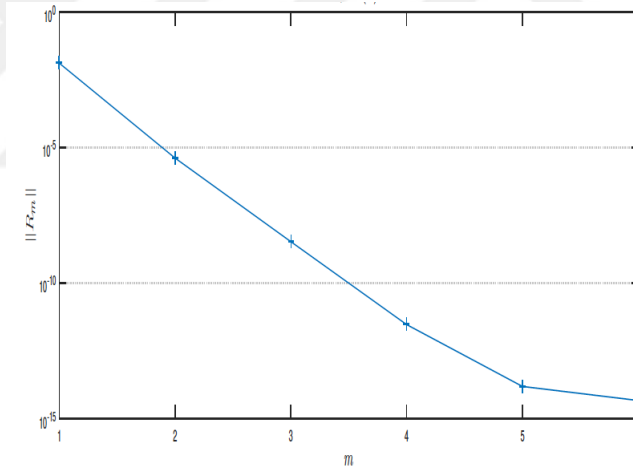
Tablo 6.5'deki sonuçlara göre EBA-BDF(2) yöntemi hesaplama zamanı bakımından



Şekil 6.6 $t \in [0, 2]$ zaman aralığı için $X_{1,1}$ değerleri

avantajlı sonuçlar vermektedir.

Şekil 6.7'de $n = 6400 \times 6400$ için rezidü normunun EBA iterasyon sayısına bağlı değişimi görülmektedir.



Şekil 6.7 ($n = 6400$) için rezidü değerleri

Örnek 6

Bu örnekte A matrisi IMTEK³ koleksiyonundaki çelik profillerin optimal soğutma probleminden alındı. $[0, 20]$ zaman aralığında $\text{size}(A) = 1357$ ve $\text{size}(A) = 5177$ durumları için EBA-BDF(2) yöntemi uygulandı. Başlangıç şartı $X_0 = X(0)$ ve $h = 0.1$ olarak alındı. EBA testini durdurmak için sınır $\text{tol} = 10^{-11}$ seçildi ve izdüşürülen düşük boyutlu MSSDE denklemleri MATLAB "dlyap" komutuyla çözüldü [19].

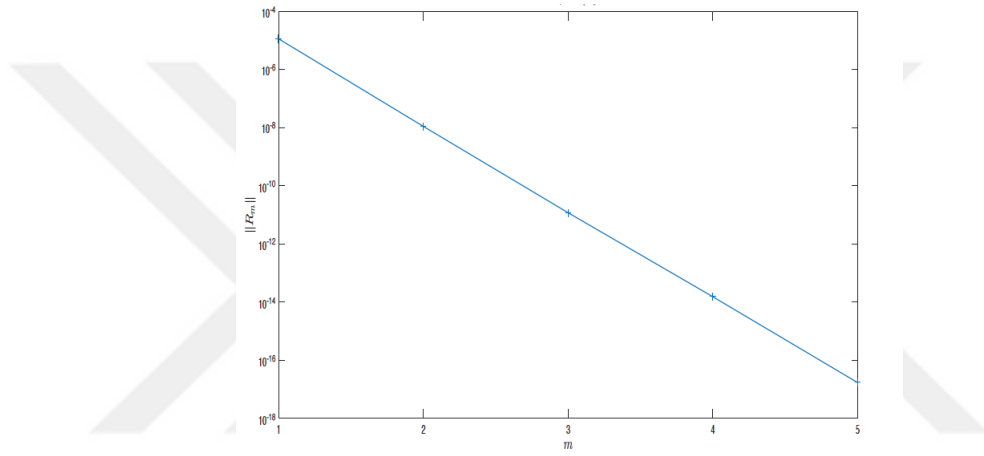
³<https://portal.uni-freiburg.de/imteksimulation/downloads/benchmark>

Tablo 6.6'da, hesaplama zamanları, Arnoldi iterasyon sayısı m ve T_f zamanındaki rezidünün $\|R_m(t_f)\|_F$ Frobenius normu verildi.

Tablo 6.6 A 'nın boyutu, hesaplama zamanları(s), rezidü normu ve Arnoldi iterasyonları sayısı(m)

size A	Hesaplama zamanı (s)	Rezidü normu	İterasyon sayısı m
1357×1357	8.2s	2.13×10^{-13}	4
5177×5177	237.45 s	5.7×10^{-13}	5

Şekil 6.8'de $\text{size}(A) = 5177 \times 5177$ durumu için T_f zamanındaki rezidü normunun EBA iterasyon sayısına bağlı değişimi verilmiştir:



Şekil 6.8 Çelik profillerin optimal soğutma problemi ($n = 5177$), $t \in [0, 20]$ zaman aralığı için $X_{1,1}(t)$ değerleri

6.3 Simetrik Olmayan Matris Stein Diferansiyel Denklemleri için Sayısal Örnekler

Bu bölümde

$$\begin{cases} \dot{X}(t) = X(t) - AX(t)B + EF^T \\ X(t_0) = X_0 \end{cases}, \quad t \in [t_0, T_f] \quad (6.8)$$

NMSDE denklemi için ele alınan sayısal örnekte EGA-BDF(1) ile ode23s çözücüsü ile elde edilen sonuçlar karşılaştırıldı. EGA-BDF(1) yöntemi ile önerdiğimiz yaklaşımların güvenilir sonuçlar verip vermediğini kontrol etmek için elde edilen sonuçlar MATLAB'ın "ode23s" çözücüsü ile elde edilen sonuçlarla karşılaştırıldı. Tüm uygulamalar Intel Core i7 işlemci ve 8 GB Ram hafızalı bir bilgisayarda yapıldı. Algoritmalar MATLAB R2014b'de yazıldı.

Örnek 7

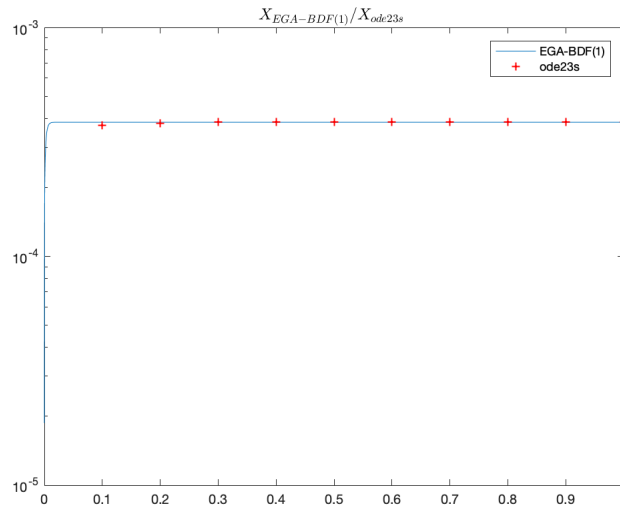
A ve B matrisleri, $[0,1] \times [0,1]$ birim karesinde $u(x,0) = u(x,1) = 0$ ve $u(0,y) = u(1,y) = 0$, homojen Dirichlet sınır koşulları altında sırasıyla

$$\begin{cases} L_A = \Delta u - f_1(x,y) \frac{\partial u}{\partial x} + f_2(x,y) \frac{\partial u}{\partial y} + g_1(x,y), \\ L_B = \Delta u - f_3(x,y) \frac{\partial u}{\partial x} + f_4(x,y) \frac{\partial u}{\partial y} + g_2(x,y), \end{cases} \quad (6.9)$$

operatörlerinin 5 nokta merkezi fark ayrışımı ile elde edilir. Her bir yöndeki iç grid noktalarının sayısı n_0 ve A matrisinin boyutu ise $n = n_0^2$ dir. Bu örnek için $f_1(x,y) = x + 10y^2$, $f_2(x,y) = \sqrt{2x^2 + y^2}$, $f_3(x,y) = x + 2y$, $f_4(x,y) = e^{y-x}$, $g_1(x,y) = x^2 - y^2$ ve $g_2(x,y) = y^2 - x^2$ alınsın. E ve F matrisleri $[0,1]$ aralığından seçilen rastgele değerlerden oluşan matrisler ve sütun sayıları $r = s = 2$ olsun. Genişletilmiş global Krylov altuzayına izdüşürme ile BDF yöntemlerinin kombinasyonu ile oluşan EGA-BDF(p) yöntemi uygulansın. Rezidü üzerindeki iterasyonu durdurma testi için doğruluk sınırı 10^{-12} seçildi ve bu testte sabit zaman adımı h ile 1 adım BDF şeması uygulandı. Şekil 6.9 ile verilen grafikte A matrisi 25×25 tipinde $t \in [0,1]$, $h = 0.1$ alınmış ve EGA-BDF(1) ile ode23s yöntemleriyle elde edilen çözümlerin $X_{1,1}(t)$ bileşenleri karşılaştırılmıştır.

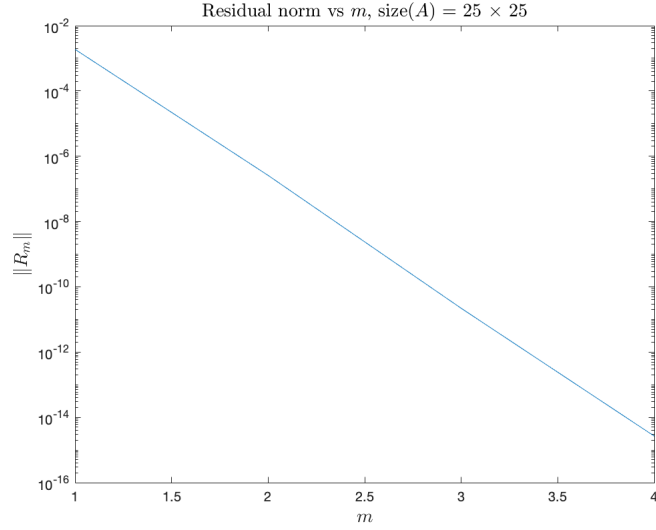
Doğruluk bakımından her iki yöntemin benzer sonuçlar verdiği gözlemlenebilir.

Şekil 6.10'da $t = 1$ zamanında EGA-BDF yönteminin iterasyon sayısına karşılık X_{EGA} çözümünün normu karakterize edilir.



Şekil 6.9 $t \in [0,1]$ zaman aralığı için $X_{1,1}(t)$ değerleri

Tablo 6.7'de $m = 7$ için elde edilen hesaplama zamanları verildi.



Şekil 6.10 m genişletilmiş global Arnoldi iterasyon sayısına karşılık gelen $\|R_m(T_f)\|$ rezidü normu

Tablo 6.7 ode23s, EGA-BDF(1) için hesaplama zamanları

size A	ode23s	EBA-BDF(1)
25×25	1.2856	0.00963
49×49	25.0990	0.1561
64×64	126.03	0.226

Tablo 6.7 gözönüne alındığında EGA-BDF(1) yönteminin hesaplama zamanı bakımından etkili olduğu görülebilir.

7 Sonuç ve Öneriler

Bu çalışmada matris diferansiyel denklemlerin sayısal çözümü için EBA-BDF(p) ve EGA-BDF(p) yaklaşımları önerildi. Önerilen yaklaşımlar, integrasyondan önce verilen matris denklemin boyutunun indirgenmesine dayanır. Buna göre MRDE, MSSDE ve NMSDE denklemleri önce genişletilmiş Krylov altuzayı üzerine izdüşürüldü ve elde edilen düşük boyutlu denklemlere BDF yöntemi uygulanarak bulunan çözümler, orjinal problemin bir yaklaşık çözümünü elde etmede kullanıldı. Bu işlem seçilen bir tol doğruluk sınırına kadar izdüşüm altuzayının boyutunu artırarak tekrarlandı. Yöntemlerin uygulamasına yönelik bazı teorik sonuçlar ve yöntemlerin algoritmaları verildi. Özellikle $X_m(T_f)$ yaklaşık çözümünü açık bir şekilde hesaplamadan $\|R_m(T_f)\|$ rezidü normunun belirlenebileceği sonucu, büyük ölçekli problemlere uygulanan standart nümerik yöntemlerin gereğinden çok bilgisayar ve hafıza işlemleri gerektirmesi nedeniyle EBA-BDF ve EGA-BDF yöntemlerinin diğer yöntemlere göre daha kullanışlı olduğunu gösterir. Ayrıca bu sonuç, Arnoldi iterasyonları için bir durdurma testi oluşturulmasını sağlar. Bu test için seçilen bir doğruluk değeri elde edilir edilmez iterasyon durdurulur. Bu tezde bütün nümerik örnekler için EBA iterasyonunu durdurma sınır değeri $tol < 10^{-8}$ alındı ve yaklaşıma göre m iterasyon sayısı arttıkça $\|R_m(T_f)\|$ rezidü normu küçüldüğünden yaklaşık çözümün kesinliğinin arttığı görüldü. Büyük ölçekli orijinal probleme önerilen bu yöntemler ile bilinen yöntemler uygulanarak yaklaşım, MATLAB programı yardımıyla diğer yöntemlerle karşılaştırıldı ve bu testlerin sonucunda $p = 2$ için EBA-BDF(p) yaklaşımının hesaplama zamanı açısından daha etkili olduğu ve yakınsak sonuçlar verdiği görüldü.

Sonuç olarak, nümerik örneklerde görüldüğü gibi doğru sonuçlar sağlayan yaklaşımlarımızın büyük ölçekli problemler için kayda değer olduğu görülmektedir. Dolayısıyla elde edilen sonuçlar, EBA-BDF(p) ve EGA-BDF(p) yöntemlerinin farklı tipteki matris diferansiyel denklemlerine uygulamasının incelenmesi gerektiği kanaatini doğurmuştur.

- [1] H. Abou-Kandil, G. Freiling, V. Ionescu, and G. Jank, “Matrix riccati equations in control and sytems theory, in systems control foundations applications,” 2003.
- [2] G. Caviglia and A. Morro, “Riccati equations for wave propagation in planary-stratified solids,” *Eur. J. Mech. A/Solids*, vol. 19, pp. 721–741, 2000.
- [3] P. Lancaster and L. Rodman, “The algebraic riccati equations,” 1995.
- [4] V. Mehrmann, “The autonomous linear quadratic control problem, theory and numerical solution,” *Number in Lecture Notes in Control and Information Sciences*, 1991.
- [5] A. H. Al-Mohy and N. Higham, “Computing the action of the matrix exponential, with an application to exponential integrators,” *SIAM J. Sci. Comput.*, vol. 33, pp. 488–511, 2011.
- [6] A. C. Antoulas, “Approximation of large-scale dynamical systems,” 2005.
- [7] I. R. Petersen, V. A. Ugrinovskii, and A. V. Savkin, “Robust control design using h1 methods,” 1995.
- [8] M. Wonham W., “On a matrix riccati equation of stochastic control,” *SIAM J. Control Optim.*, vol. 6, pp. 681–698, 1968.
- [9] P. Molinari B., “The stabilizing solution of the matrix quadratic equation,” *SIAM J. Control*, vol. 10, 1972.
- [10] R. Vaughan D., “A negative exponential solution for the matrix riccati equation,” *IEEE Trans. on Automatic Control*, vol. 14, no. 1, pp. 72–75, 1969.
- [11] J. Davison E. and C. Maki M., “The numerical solution of the matrix riccati differential equation,” *IEEE Trans. on Automatic Control*, vol. 18, no. 1, pp. 71–73, 1973.
- [12] S. Kenney C. and B. Leipnik R., “Numerical integration of the differential matrix riccati equation,” *IEEE Trans. on Automatic Control*, vol. 30, pp. 962–970, 1985.
- [13] H. Chandrasekhar, “Generalized chandrasekhar algorithms: Time-varying models,” *IEEE Trans. on Automatic Control*, vol. 21, pp. 728–732, 1976.
- [14] W. Rand D. and P. Winternitz, “Nonlinear superposition principles: A new numerical method for solving matrix riccati equations,” *Comp. Phys. Commun.*, vol. 33, pp. 305–328, 1984.
- [15] M. Sorine and P. Winternitz, “Superposition laws for the solution of differential riccati equations,” *IEEE Trans. on Automatic Control*, vol. 30, pp. 266–272, 1985.
- [16] M. Sanz-Serna J., “Symplectic integrators for hamiltonian problems: An overview,” *Acta Numerica 1*, vol. 1, pp. 243–286, 1992.

- [17] L. Jodar and E. Ponsoda, "Non-autonomous riccati-type matrix differential equations: Existence interval, construction of continuous numerical solutions and error bounds," *IMA J. Numerical Analysis*, vol. 15, pp. 61–74, 1995.
- [18] S. R. Nekoo and M. I. Rahaghi, "Recursive approximate solution to time-varying matrix differential riccati equation: Linear and nonlinear systems," *International Journal of Systems Science*, vol. 49, p. 13, 2018.
- [19] C. H. Guo and B. Yu, "A convergence result for matrix riccati differential equations associated with m-matrices," *Taiwanese J. Math.*, vol. 19, no. 1, pp. 77–89, 2015.
- [20] S. V. Gusev, A. S. Shiriaev, and L. B. Freidovich, "Sdp-based approximation of stabilising solutions for periodic matrix riccati differential equations," *International Journal of Control*, vol. 89, no. 7, pp. 1396–1405, 2016.
- [21] V. Laptinskii and R. D.V., "Constructive methods for obtaining the solution of the periodic boundary value problem for a system of matrix differential equations of riccati type," *Differential Equations*, vol. 47, no. 10, pp. 1426–1434, 2011.
- [22] H. J. Peng, Z. G. Wua, and W. X. Zhong, "Fourier expansion based recursive algorithms for periodic riccati and lyapunov matrix differential equations," *Journal of Computational and Applied Mathematics*, vol. 235, pp. 3571–3588, 2011.
- [23] J. Ibáñez and V. Hernández, "Solving differential matrix riccati equations by a piecewise-linearized method based on diagonal padé approximants," *Journal of Computational and Applied Mathematics*, vol. 182, pp. 669–678, 2011.
- [24] P. Benner and H. Mena, "Bdf methods for large-scale differential riccati equations," *Proceedings of Mathematical Theory of Network and Systems*, 2004.
- [25] V. Sima, "Algorithms for linear-quadratic optimization," *Pure and Applied Mathematics*, 1996.
- [26] P. H. Petkov, N. Christov, and M. Konstantinov, "Computational methods for linear control systems," 1991.
- [27] T. Stillfjord, "Low-rank second-order splitting of large-scale differential riccati equations," *IEEE Transactions on Automatic Control*, vol. 60, no. 10, pp. 2791–2796, 2015.
- [28] E. Hairer, C. Lubich, and G. Wanner, "Geometric numerical integration," vol. 31, 2006.
- [29] W. Hundsdorfer and J. G. Verwer, "Numerical solution of time-dependent advection-diffusion-reaction equations," vol. 33, 2003.
- [30] C. Lanczos, "An iteration method for the solution of the eigenvalue problem of linear differential and integral operators," *J. Res. Nat. Bur. Stand.*, vol. 45, pp. 255–282, 1950.
- [31] W. Arnoldi, "The principle of minimized iterations in the solution of the matrix eigenvalue problem," *Quart. Appl. Math.*, vol. 9, pp. 17–29, 1951.
- [32] M. Hestenes and E. Stiefel, "Methods of conjugate gradients for solving linear systems," *Quart. Appl. Math.*, vol. 49, pp. 409–436, 1952.

- [33] J. Reid, "On the method of conjugate gradients for the solution of large sparse systems of linear equations," *Large Sparse Sets of Linear Equations*, pp. 231–254, 1971.
- [34] S. F. Ashby, T. A. Manteuffel, and P. Saylor, "A taxonomy for conjugate gradient methods," *SIAM J. Numer. Anal.*, vol. 27, pp. 1542–1568, 1990.
- [35] Y. Saad and M. Schultz, "Conjugate gradient-like algorithms for solving nonsymmetric linear systems," *Mathematics of Computation*, vol. 44, pp. 417–424, 1985.
- [36] P. Benner, "Computational methods for linear-quadratic optimization," *Supplemento ai R. del Circ. Matematico di Pal.*, vol. 2, no. 58, pp. 21–56, 1999.
- [37] J. Saak, "Efficient numerical solution of large scale algebraic matrix equations in pde control and model order reduction," *PhD thesis*, 2009.
- [38] H. Mena, "Numerical solution of differential riccati equations arising in optimal control problems for parabolic partial differential equations," 2007.
- [39] B. Datta, "Numerical methods for linear control systems design and analysis," 2003.
- [40] M. J. Corless and A. E. Frazho, "Linear systems and control: An operator perspective, pure app. math.," 2003.
- [41] P. Benner and R. Byers, "An exact line search method for solving generalized continuous algebraic riccati equations," *IEEE Trans. Automat. Control*, vol. 43, no. 1, pp. 101–107, 1998.
- [42] P. Benner and J. Saak, "Efficient numerical solution of the lqr-problem for the heat equation," *Proc. Appl. Math. Mech.*, vol. 4, pp. 648–649, 2004.
- [43] F. Troltzsch and A. Unger, "Fast solution of optimal control problems in the selective cooling of steel," *Z. Angew. Math. Mech.*, vol. 81, pp. 447–456, 2001.
- [44] K. Eppler and F. Troltzsch, "Discrete and continuous optimal control strategies in the selective cooling of steel profiles," *Z. Angew. Math. Mech.*, vol. 81, pp. 247–248, 2001.
- [45] Y. Saad, "Iterative methods for sparse linear systems," 1996.
- [46] M. Heyouni, "Extended arnoldi methods for large low-rank sylvester matrix equations," *Applied Numerical Analysis*, vol. 60, pp. 1171–1182, 2010.
- [47] K. Jbilou, "Block krylov subspace methods for large continuous-time algebraic riccati equations," *Num. Alg.*, vol. 34, pp. 339–353, 2003.
- [48] M. Heyouni and K. Jbilou, "An extended block arnoldi algorithm for large-scale solutions of the continuous-time algebraic riccati equation," *Electron. Trans. Numer. Anal.*, vol. 33, pp. 53–62, 2008.
- [49] K. Jbilou, A. Messaoudi, and H. Sadok, "Global fom and gmres algorithms for matrix equations," *Appl Numer Math.*, vol. 31, pp. 49–63, 1999.
- [50] V. Druskin and L. Knizhnerman, "Extended krylov subspaces: Approximation of the matrix square root and related functions," *SIAM J. Matrix Anal. Appl.*, vol. 19, pp. 755–771, 1998.
- [51] V. Simoncini, "A new iterative method for solving large-scale lyapunov matrix equations," *SIAM J. Sci. Comput.*, vol. 29, pp. 1268–1288, 2007.

- [52] G. Caviglia and A. Morro, "Riccati equations for wave propagation in planary-stratified solids," *Eur. J. Mech. A/Solids*, vol. 19, pp. 721–741, 2000.
- [53] K. Jbilou, "An arnoldi based algorithm for large algebraic riccati equations," *Applied Mathematics Letters*, vol. 19, pp. 437–444, 2006.
- [54] B. D. O. Anderson and J. B. Moore, "Linear optimal control," 1971.
- [55] P. Benner and H. Mena, "Rosenbrock methods for solving differential riccati equations," *IEEE Transactions on Automatic Control*, vol. 58, no. 11, pp. 2950–2956, 2013.
- [56] W. Arnold III and L. J. Laub, "Generalized eigenproblem algorithms and software for algebraic riccati equations," *Pro.*, vol. 72, pp. 1746–1754, 1984.
- [57] P. Van Dooren, "A generalized eigenvalue approach for solving riccati equations," *SIAM J. Sci. Stat. Comput.*, vol. 2, pp. 121–135, 1981.
- [58] D. L. Kleinman, "On an iterative technique for riccati equation computations," *IEEC Trans. Autom. Contr.*, vol. 13, pp. 114–115, 1968.
- [59] A. Guo and P. Lancaster, "Analysis and modification of newton's method for algebraic riccati equations," *Math. Comp.*, vol. 67, pp. 1089–1105, 1998.
- [60] C. Kenny, A. Laub, and P. Papadopoulos, "Matrix sign function algorithms for riccati equations," in *In Proceedings of IMA Conference on Control: Modelling, Computation*, IEEE Computer Society Press, Southend-On-Sea., 1992, pp. 1–10.
- [61] A. Bouhamidi, M. Hached, and K. Jbilou, "A preconditioned block arnoldi method for large scale lyapunov and algebraic riccati equations," *Journal of Global Optimization*, vol. 65, no. 1, pp. 19–23, 2016.
- [62] J. L. Casti, "Linear dynamical systems," 1987.
- [63] M. Hached and K. Jbilou, "Numerical solutions to large-scale differential lyapunov matrix equations," *Numer. Algor.*, vol. 79, no. 3, pp. 741–757, 2018.
- [64] I. Jaimoukha and E. Kasenally, "Krylov subspace methods for solving large lyapunov equations," *SIAM J. Matrix Anal. Appl.*, vol. 31, no. 1, pp. 227–251, 1994.
- [65] G. H. Golub and C. F. Van Loan, "Matrix computations," 1989.

A

Temel Tanım ve Kavramlar

Tanım A.1 (Bir Matrisin Vektörizasyonu). $A \in \mathbb{R}^{n \times n}$ olmak üzere A 'nın vektörizasyonu, $vec(A) = (a_{11}, a_{21}, \dots, a_{m1}, a_{12}, a_{22}, \dots, a_{m2}, \dots, a_{1n}, a_{2n}, \dots, a_{mn})^T$ ile tanımlıdır [45].

Tanım A.2 (Köşegenleştirilebilir Matris). Bir A , $n \times n$ kare matrisine $A = Pdiag(\lambda_1, \dots, \lambda_n)P^{-1}$ olacak şekilde bir P nonsingüler matrisi varsa A matrisine köşegenleştirilebilir matris denir [45].

Tanım A.3 (Seyrek ve Yoğun Matris). Seyrek bir matris, elemanlarının çoğunun sıfır değerine sahip olduğu bir matristir. Yoğun matris ise elemanların çoğunun sıfırdan farklı bir değere sahip olduğu matristir. Seyrek matrisler, yoğun matrislerin tersidir ve bu matrisler bilgisayar bilimlerinde belirli şekillerde kullanılır ayrıca kullanımlarıyla ilgili farklı veri analizi, depolama protokolleri ve tekniklerine sahiptir [45].

A.1 Matris Çarpımları

A.1.1 $\otimes$ - Kronocker Çarpım

$A = (a_{i,j}) \in \mathbb{R}^{m \times n}$ ve $B = (b_{i,j}) \in \mathbb{R}^{p \times q}$ matrislerinin Kronecker çarpımı $A \otimes B = [a_{i,j}B] \in \mathbb{R}^{mp \times nq}$ olarak tanımlanır ve

$$A \otimes B = \begin{pmatrix} a_{11}B & a_{12}B & \dots & a_{1n}B \\ a_{21}B & a_{22}B & \dots & a_{2n}B \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}B & a_{m2}B & \dots & a_{mn}B \end{pmatrix}$$

dir [65].

Bu çarpım aşağıdaki özellikleri sağlar;

- $(A \otimes B)(C \otimes D) = (AC \otimes BD)$
- $(A \otimes B)^T = A^T \otimes B^T$

A.1.2 $\diamond$ -Çarpım

$A = [A_1, A_2, \dots, A_p] \in \mathbb{R}^{n \times pr}$ ve $B = [B_1, B_2, \dots, B_l] \in \mathbb{R}^{n \times lr}$ olmak üzere bloklar $A_i, B_j \in \mathbb{R}^{n \times r}$, $(i = 1, \dots, p; j = 1, \dots, l)$ olsun. $A^T \diamond B = [tr(A_i^T B_j)]_{i=1, \dots, p}^{j=1, \dots, l}$ olarak

tanımlanır ve burada $A^T \diamond B$, $p \times l$ matristir.

$$A^T \diamond B = \begin{pmatrix} \langle A_1, B_1 \rangle_F & \langle A_1, B_2 \rangle_F & \dots & \langle A_1, B_l \rangle_F \\ \langle A_2, B_1 \rangle_F & \langle A_2, B_2 \rangle_F & \dots & \langle A_2, B_l \rangle_F \\ \vdots & \vdots & \ddots & \vdots \\ \langle A_n, B_1 \rangle_F & \langle A_n, B_2 \rangle_F & \dots & \langle A_n, B_l \rangle_F \end{pmatrix}$$

$A = [A_1, A_2, \dots, A_p]$ matrisi F ortonormaldir $\Leftrightarrow A^T \diamond A = I_p$ ’ dir.
Yani

$$tr(A_i^T A_j) = \delta_{i,j} = \begin{cases} 0, & i \neq j \\ 1, & i = j \end{cases}$$

dir [46].

A.2 Matris Normu ve Özellikleri

A.2.1 Matris p ve q Normu

Bir $A \in \mathbb{R}^{n \times m}$ matrisinin normu

$$\|A\|_{pq} = \max_{x \in \mathbb{R}^n - 0} \frac{\|Ax\|_p}{\|x\|_q}$$

olarak tanımlanır. Bu normlar $p = q$ olması durumunda

$$\|AB\|_p \leq \|A\|_p \|B\|_p$$

eşitsizliğini sağlar. Eğer A matrisi kare matris ise o zaman

$$\|A^k\|_p \leq \|A\|_p^k, \quad k = 1, 2, \dots$$

dir. Ayrıca

$$\|I_n\|_p = 1$$

dir [45].

A.2.2 Frobenius Norm

Bir A matrisinin Frobenius normu

$$\|A\|_F = \sqrt{\sum_{i=1}^m \sum_{j=1}^n |a_{ij}|^2}$$

ile tanımlıdır. Ayrıca

$$\|I_n\|_F = \sqrt{n}$$

dir. $X, Y \in \mathbb{R}^{n \times m}$ matrislerinin Frobenius çarpımı

$$\langle X, Y \rangle_F = \text{tr}(X^T Y)$$

şeklindedir. Bu durumda

$$\|A\|_F = \sqrt{\langle X, Y \rangle_F} = \sqrt{\text{tr}(X^T Y)}$$

yazılabilir [49].

A.3 Matris Ayrışimleri

A.3.1 Tekil Değer Ayrışımı

Aşağıdaki teorem her matrisin ortogonal matrisler ile bir köşegen matrisin çarpımı şeklinde ayrıştırılabileceğini gösterir:

Teorem A.1. *A matrisi $n \times m$ tipinde reel bir matris olsun. Bu durumda $U = [u_1, u_2, \dots, u_n] \in \mathbb{R}^{n \times n}$ ve $V = [v_1, v_2, \dots, v_m] \in \mathbb{R}^{m \times m}$ olacak şekilde iki ortogonal matris ve bir $\Sigma = \text{diag}(\sigma_1, \dots, \sigma_p)$, $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_p \geq 0$, $p = \min\{n, m\}$ köşegen matrisi vardır öyleki $A = U \Sigma V^T$ dir. Bu ayrışımı A'nın tekil değer ayrışımı denir. Burada σ_i değerlerine A'nın singüler değerleri denir [21].*

A.3.2 QR Ayrışımı

A matrisi $n \times m$ tipinde reel bir matris ve $n \geq m$ olsun. A matrisinin QR ayrışımı

$$A = QR$$

olacak şekilde bir $Q \in \mathbb{R}^{n \times n}$ ortogonal matrisi ile bir $n \times m$ tipinde üst üçgen R matrisinin hesaplanmasıdır [45].

A.3.3 Global QR Ayrışımı

$Z = [Z_1, Z_2, \dots, Z_p] \in \mathbb{R}^{n \times pr}$, ($Z_i \in \mathbb{R}^{n \times r}$), $i = 1, \dots, k$ olsun. Z'nin global QR ayrışımı, global Gram- Schmidt işlemine dayanır ve bir $n \times pr$ F ortonormal matrisi olan $Q = [Q_1, \dots, Q_p]$ ile bir $r \times r$ üst üçgen $R = [r_{i,j}]$ matrisi üretir öyleki $\text{span}\{Q_1, Q_2, \dots, Q_p\} = \text{span}\{Z_1, Z_2, \dots, Z_p\}$ ve

$$Z = Q(R \otimes I_r)$$

dir [46].

Sonuç A.1. $Z_i \in \mathbb{R}^{n \times p}$, ($i = 1, 2, \dots, m$), olmak üzere $\mathcal{Z} = [Z_1, Z_2, \dots, Z_m]$ bir $n \times mp$ matrisi olsun. Buradan, $\mathcal{Z}$ matrisi

$$\mathcal{Z} = Z(R \otimes I_p)$$

olarak faktörize edilir. Burada $\mathbf{Z} = [Q_1, \dots, Q_m]$ matrisi $\mathbf{Z}^T \diamond \mathbf{Z} = I_m$ eşitliğini sağlayan bir $n \times mp$ boyutlu F -ortonormal matris ve R matrisi $m \times m$ boyutlu bir üst üçgensel matristir [49].

Sonuç A.2. $\mathcal{V}_m = [V_1, \dots, V_m]$ matrisi $n \times mp$ boyutlu bir F -ortonormal matris olsun. $Z = [z_{i,j}]$ ve $G = [g_{i,j}]$ matrisleri sırasıyla $m \times r$ ve $mp \times q$ boyutlu matrisler olsun. Buradan

$$\|\mathcal{V}_m(Z \otimes I_p)\|_F = \|Z\|_F$$

ve

$$\|\mathcal{V}_m G\| \leq \|G\|_F$$

dir [49].



Tezden Üretilmiş Yayınlar

İletişim Bilgileri: yyguldogan@gmail.com

Makale

1. Gldođan,Y., Hached, M., Jbilou, K. and Kurulay, M., " Low rank approximate solutions to large-scale differential matrix Riccati equations", Applicationes Math.,45 (2018), 233-254. doi: 10.4064/am2355-1-2018.
2. Gldođan Dericiođlu, Y, Kurulay, M., "A computational method for large-scale differential symmetric Stein equation". Math Meth Appl Sci. 2018; 1– 8. <https://doi.org/10.1002/mma.5405>.

Konferans Bildirisi

1. Dericiođlu, Y., Kurulay, M., "Extended Krylov Method for the Numerical Resolution of a Large Scale Differential Symmetric Stein Equations", 4 th International Conference on Analysis and Its Applications, 11-14 September 2018, Kırřehir.
2. Dericiođlu Y., Hached M., Kurulay M., Jbilou K., "A New Application for Solving Large-Scale Differential Matrix Riccati Equations", International Congress on Fundamental and Applied Sciences(ICFAS2016), 22-26 August 2016, İstanbul.

Proje

1. Proje Adı: Yksek Mertebeli Matris Diferansiyel Denklemleri iin Krylov Yntemleri
Proje Yrtcs: Do. Dr. Muhammet KURULAY
Arařtırmacı: Yaprak Dericiođlu
Destekleyen Kurum/Birim: Yıldız Teknik niversitesi/Bilimsel Arařtırma Projeleri Koordinatrlđ
Proje No: 2016-07-03-DOP04