



**BÜYÜK ÖLÇEKLİ KISITSIZ OPTİMİZASYON  
PROBLEMLERİNİN ÇÖZÜMÜ İÇİN YENİ BİR  
EŞLENİK GRADYAN ALGORİTMASI VE  
UYGULAMALARI**

**Dilara AKDAĞ**

**Yüksek Lisans Tezi**

**Matematik Anabilim Dalı**

**Danışman: Doç. Dr. İbrahim KARAHAN**

**2025**

**Her hakkı saklıdır.**



**FEN BİLİMLERİ ENSTİTÜSÜ  
YÜKSEK LİSANS TEZİ**

**BÜYÜK ÖLÇEKLİ KISITSIZ OPTİMİZASYON PROBLEMLERİNİN  
ÇÖZÜMÜ İÇİN YENİ BİR EŞLENİK GRADYAN ALGORİTMASI VE  
UYGULAMALARI**

**Dilara AKDAĞ**

**Tez Danışmanı: Doç. Dr. İbrahim KARAHAN**

**Anabilim Dalı: Matematik**

**Erzurum**

**2025**

**Her hakkı saklıdır**

T.C.  
ERZURUM TEKNİK ÜNİVERSİTESİ  
FEN BİLİMLERİ ENSTİTÜSÜ  
TEZ ONAY FORMU

---

**BÜYÜK ÖLÇEKLİ KISITSIZ OPTİMİZASYON PROBLEMLERİNİN  
ÇÖZÜMÜ İÇİN YENİ BİR EŞLENİK GRADYAN ALGORİTMASI VE  
UYGULAMALARI**

Doç. Dr. İbrahim KARAHAN danışmanlığında, Dilara AKDAĞ tarafından hazırlanan bu çalışma 10 / 01 / 2025 tarihinde aşağıdaki jüri tarafından Matematik Ana Bilim Dalı'nda Yüksek Lisans tezi olarak **Oy birliği (3/3)** ile kabul edilmiştir.

Başkan : Prof. Dr. Murat SUBAŞI *İmza* :

Üye : Doç. Dr. İbrahim KARAHAN *İmza* :

Üye : Dr. Öğr. Üyesi Ebru ALTIPARMAK YANGAL *İmza* :

Yukarıdaki sonucu onaylıyorum

**Doç. Dr. Bünyamin ÖZGERİŞ**  
Enstitü Müdürü

\*Bu tez çalışması ..... tarafından ..... nolu proje ile desteklenmiştir.

## ETİK KURALLARA UYGUNLUK BEYANI

Erzurum Teknik Üniversitesi Fen Bilimleri Enstitüsü tez yazım kurallarına uygun olarak hazırladığım bu tez içindeki tüm bilgilerin doğru ve tam olduğunu, bilgilerin üretilmesi aşamasında bilimsel etiğe uygun davrandığımı, yararlandığım bütün kaynakları atıf yaparak belirttiğimi beyan ederim.

10 / 01 / 2025

Dilara AKDAĞ

## ÖZET

YÜKSEK LİSANS TEZİ

# BÜYÜK ÖLÇEKLİ KISITSIZ OPTİMİZASYON PROBLEMLERİNİN ÇÖZÜMÜ İÇİN YENİ BİR EŞLENİK GRADYAN ALGORİTMASI VE UYGULAMALARI

Dilara AKDAĞ

Erzurum Teknik Üniversitesi

Fen Bilimleri Enstitüsü

Matematik Ana Bilim Dalı

Danışman: Doç. Dr. İbrahim KARAHAN

Büyük ölçekli lineer olmayan optimizasyon problemlerini çözmek için kullanılan başlıca algoritmalarından biri de eşlenik gradyan algoritmalarıdır. Bunun nedeni algoritmaların adımları hesaplanırken ikinci dereceden türevlerin hesaplanmasına gerek olmaması, düşük depolama, hızlı sonuç verme ve global yakınsama gibi uygun özelliklerinin var olmasıdır. Eşlenik gradyan algoritmasının kullanıldığı alanlar; endüstri, mühendislik optimizasyon problemleri, sinir ağı eğitimi ve görüntü restorasyonu gibi alanlardır. Bu tezde yeni bir eşlenik gradyan algoritması tanımlanmış, bu algoritmanın Armijo ve Wolfe çizgi arama teknikleri ile birlikte minimizasyon probleminin çözümüne kuvvetli yakınsak olduğu ve yeterli iniş özelliğini sağladığı gösterilmiştir. Son olarak algoritmanın performans profil karşılaştırılması yapılmış ve görüntü iyileştirme problemine uygulaması verilmiştir.

**2025, 52 sayfa**

**Anahtar Kelimeler:** Eşlenik gradyan algoritması, optimizasyon problemi, çizgi arama tekniği, performans profili, görüntü restorasyon problemi.

## ABSTRACT

MS. Thesis

# A NEW CONJUGATE GRADIENT ALGORITHM FOR SOLVING LARGE SCALE UNCONSTRAINED OPTIMIZATION PROBLEMS AND ITS APPLICATIONS

Dilara AKDAĞ

Erzurum Technical University

Graduate School of Natural and Applied Sciences

Department of Maths

Supervisor: Doç. Dr. İbrahim KARAHAN

Conjugate gradient algorithms are among the primary algorithms used to solve large-scale nonlinear optimization problems. This is due to their advantageous properties, such as no need to compute second-order derivatives when calculating algorithmic steps, low storage requirements, rapid results, and global convergence. Conjugate gradient algorithms are utilized in areas such as industry, engineering optimization problems, neural network training, and image restoration. In this thesis, a new conjugate gradient algorithm is introduced, and it is shown that this algorithm strongly converges to the solution of the minimization problem and satisfies the sufficient descent property when used with Armijo and Wolfe line search techniques. Finally, a performance profile comparison of the algorithm is presented, along with its application to the image restoration problem.

**2025, 52 page**

**Keywords:** Conjugate gradient algorithm, optimization problem, line search techniques, performance profile, image restoration problem.

## TEŐEKKÜR

Yüksek Lisans eğitiminin boyunca ve tez çalışmamı oluşturmamda her türlü desteęi sağlayan her zaman arkamda olan bilgi ve tecrübeleriyle çalışmamı ışık tutan çok değerli hocam Sayın Doç. Dr. İbrahim KARAHAN'a en içten dileklerim ile teşekkür eder saygılarımı sunarım.

Çalışmalarım boyunca desteęini esirgemeyen Dr. Öğr. Üyesi Ebru ALTIPARMAK YANGAL'a ve Erzurum Teknik Üniversitesi, Fen Fakültesi Matematik Bölümünde bulunan değerli hocalarıma sonsuz teşekkürlerimi sunarım.

Eğitim hayatım boyunca bana hep güvenen, her türlü özveriye göstererek bugünlere gelmemi sağlayan, varlıklarıyla güç bulduğum annem Mücella AKDAĞ'a, babam Muhammet AKDAĞ'a ve benden desteklerini ve sevgilerini hiç eksik etmeyen kardeşlerim İlayda AKDAĞ, Aysu AKDAĞ ve Ceren AKDAĞ'a tüm fedakarlıkları için minnetlerimi sunarım.

**Dilara AKDAĞ**  
**Ocak 2025**

## İÇİNDEKİLER

	<u>Sayfa</u>
<b>ÖZET</b> .....	i
<b>ABSTRACT</b> .....	ii
<b>TEŞEKKÜR</b> .....	iii
<b>İÇİNDEKİLER</b> .....	iv
<b>SİMGELER VE KISALTMALAR DİZİNİ</b> .....	v
<b>ŞEKİLLER DİZİNİ</b> .....	vi
<b>ÇİZELGELER DİZİNİ</b> .....	vii
<b>1. GİRİŞ</b> .....	1
<b>2. KAYNAK ÖZETLERİ</b> .....	5
<b>3. TEMEL KAVRAMLAR</b> .....	11
3.1. Temel Tanım ve Teoremler .....	11
3.2. Kısıtsız Optimizasyonun Temelleri .....	16
<b>4. EŞLENİK GRADYAN ALGORİTMALAR</b> .....	20
4.1. Optimizasyon Algoritmalarının Özellikleri.....	20
4.2. Çizgi Arama Teknikleri.....	23
4.3. Eşlenik Gradyan Algoritmalar .....	25
<b>5. ARAŞTIRMA BULGULARI</b> .....	30
5.1. DEI Eşlenik Gradyan Algoritması .....	30
5.2. Sayısal Sonuçlar .....	39
5.2.1. Büyük Ölçekli Kısıtsız Problemler İçin Simülasyon Testi.....	39
5.2.2. Görüntü Restorasyon Problemlerine Uygulama .....	42
<b>6. SONUÇ VE ÖNERİLER</b> .....	48
<b>KAYNAKLAR</b> .....	50

## SİMGELER VE KISALTMALAR DİZİNİ

<b><u>Simgeler</u></b>	<b><u>Açıklama</u></b>
$\beta_k$	Eşlenik gradyan parametresi
$d_k$	Arama yönü
$\text{epi}(f)$	$f$ fonksiyonun epigrafı
$N_\delta(x_0)$	$x_0$ merkezli $\delta$ yarıçaplı açık yuvar
$\mathbb{R}$	Reel Sayılar
$\nabla f$	$f$ fonksiyonunun gradyan operatörü
$\alpha_k$	Adım uzunluğu

### **Kısaltmalar**

CG	Eşlenik gradyan
CPU	Merkezi işlem birimi
SNR	İşaret gürültü oranı

## ŞEKİLLER DİZİNİ

<b>Şekil 3.1:</b> $f(x) = (x - 1)^2$ fonksiyonunun epigrafı .....	14
<b>Şekil 5.1:</b> %15 gürültü ile bozulmuş gürültülü Lena görüntüsü .....	43
<b>Şekil 5.2:</b> %1 gürültü ile bozulmuş gürültülü kameraman görüntüsü için DEI, PRP, PRP-T ve JMJ algoritması ile elde edilen restorasyon .....	45
<b>Şekil 5.3:</b> %5 gürültü ile bozulmuş gürültülü MRI görüntüsü için DEI, PRP, PRP-T ve JMJ algoritması ile elde edilen restorasyon.....	46
<b>Şekil 5.4:</b> %15 gürültü ile bozulmuş gürültülü Lena görüntüsü için DEI, PRP, PRP-T ve JMJ algoritması ile elde edilen restorasyon.....	46



## ÇİZELGELER DİZİNİ

<b>Çizelge 5.1.</b> Algoritmaları karşılaştırmak için kullanılan test fonksiyonları	39
<b>Çizelge 5.2.</b> Algoritmaların test fonksiyonlarını çözme süreleri	40
<b>Çizelge 5.3.</b> Algoritmaların test fonksiyonlarını çözme sürelerin o fonksiyonu en hızlı çözen algoritmanın çözme süresine oranı	41
<b>Çizelge 5.4.</b> Çizelge 5.3'deki sonuçların 2 tabanındaki logaritma değerleri	41
<b>Çizelge 5.5.</b> Tcpu performans profilleri	42
<b>Çizelge 5.6.</b> Algoritmalara karşılık gelen CPU süreleri ve SNR değerleri	47



### 1. GİRİŞ

Eşlenik gradyan (CG) algoritması, global yakınsama özelliğine sahip önemli kısıtsız optimizasyon algoritmalarından biridir. CG algoritmaları günümüzde mevcut olan en hızlı veya en sağlam optimizasyon algoritmaları olmamasına rağmen, büyük ölçekli problemleri çözmekle uğraşan mühendisler ve matematikçiler için çok popüler olmaya devam etmektedir. Bu algoritma ilk olarak Hestenes ve Stiefel (1952) tarafından simetrik pozitif tanımlı lineer denklem sistemlerini çözmek için tanımlanmıştır. 1960'larda Hestenes ve Stiefel'in orijinal çalışmaları üzerine çeşitli iyileştirmeler yapılmış, algoritmalar geliştirilmiş ve farklı alanlara uygulanmıştır. 1970-1980'lerde özellikle mühendislik ve fiziksel bilimlerde optimizasyon problemlerinin çözümünde yaygın olarak kullanılmaya başlanmış ve aynı zamanda algoritmanın çeşitli varyantları geliştirilmiştir. 1990'lardan günümüze kadar hesaplama gücündeki artış ve yeni teorik yaklaşımlar sayesinde, CG algoritmaları daha karmaşık ve büyük ölçekli problemlerin çözümünde ve Cholesky ayrıştırması gibi doğrudan yöntemlerle çözülemeyecek kadar büyük olan seyrek sistemlerin çözümünde kullanılmaya devam etmektedir. Büyük seyrek sistemler genellikle kısmi diferansiyel denklemleri veya optimizasyon problemlerini sayısal olarak çözerken ortaya çıkmaktadır. Bu algoritmaların adımları hesaplanırken ikinci dereceden türevlerin hesaplanmasına gerek olmaması, düşük depolama ve hesaplama, uygun yakınsama oranı gibi uygun özelliklerinin var olmasından dolayı algoritmalar arasında popülerlik kazanmıştır. Günümüzde hala aktif olarak araştırılan bir konu olup, optimizasyon teorisinin temel taşlarından biridir ve özellikle makine öğrenimi ve veri bilimi alanında önemli bir yer edinmiştir.

$f: \mathbb{R}^n \rightarrow \mathbb{R}$  düzgün (smooth), lineer olmayan ve sürekli türevlenebilir bir fonksiyon olmak üzere

$$\min\{f(x): x \in \mathbb{R}^n\}$$

şeklindeki kısıtsız optimizasyon problemini çözmek için kullanılan eşlenik gradyan algoritması,  $x_0$  rastgele seçilen bir başlangıç noktası olmak üzere

$$x_{k+1} = x_k + \alpha_k d_k$$

şeklinde tanımlanır. Burada  $\alpha_k$  adım uzunluğu ve  $d_k$  ise genellikle

$$d_k = \begin{cases} -g_k, & k = 0 \text{ ise} \\ -g_k + \beta_k d_{k-1}, & k \geq 1 \text{ ise} \end{cases}$$

şeklinde tanımlanan arama yönüdür. Burada verilen  $g_k = \nabla f(x_k)$  ve  $\beta_k$  ise CG parametresidir.  $\beta_k$ 'nin farklı seçilmesiyle farklı CG algoritmaları elde edilir.

Çizgi arama tekniği kesin (exact) ve kesin olmayan (inexact) çizgi arama teknikleri olarak kategorize edilebilir. Kesin çizgi arama tekniği, adından da anlaşılacağı gibi, her iterasyonda kesin minimizeri bulmayı amaçlarken, kesin olmayan çizgi arama tekniğinde minimum değere yaklaşık bir değer elde etmek için adım uzunlukları belirlenir. Armijo, güçlü Wolfe ve zayıf Wolfe gibi çizgi arama teknikleri, kesin olmayan çizgi arama tekniklerinden bazılarıdır.

Armijo çizgi arama tekniğinde kullanılan eşitsizlik  $c_1 \in (0,1)$  olmak üzere

$$f(x_k + \alpha_k d_k) \leq f(x_k) + c_1 \alpha_k \nabla f_k^T d_k$$

şeklindeir.  $c_2 \in (c_1, 1)$  ve  $0 < c_1 < c_2 < 1$  olmak üzere

$$f(x_k + \alpha_k d_k) \leq f(x_k) + c_1 \alpha_k \nabla f_k^T d_k$$

$$\nabla f(x_k + \alpha_k d_k)^T d_k \geq c_2 \nabla f_k^T d_k$$

eşitsizlikleri toplu olarak Wolfe koşulu olarak adlandırılmaktadır. Wolfe koşulundaki ikinci eşitsizlik yerine

$$|\nabla f(x_k + \alpha_k d_k)^T d_k| \leq c_2 |\nabla f_k^T d_k|$$

eşitsizliği alınırsa güçlü Wolfe Koşulu elde edilir.

Eşlenik gradyan algoritmalarında arama yönü olan  $d_k$  vektörleri birbirleri ile eşlenik olacak şekilde tanımlanmaktadır. Bu algoritmalara eşlenik gradyan algoritması denilmesinin sebebi budur.  $d_k$  arama yönünü hesaplamak için bir önceki  $d_{k-1}$  vektörünü bilmek yeterlidir. Yani  $d_0, d_1, \dots, d_{k-2}$  vektörlerinin bilinmesine gerek yoktur. Bu özellik sebebiyle eşlenik gradyan algoritması fazla depolama alanına ihtiyaç duymamaktadır.

$y_{k-1} = g_k - g_{k-1}$  ve  $\|\cdot\|$  ise Öklid normunu belirtmek üzere en temel eşlenik gradyan algoritmaları olan PRP, HS, LS, FR, DY ve CD algoritmalarının CG parametreleri sırasıyla

$$\beta_k^{PRP} = \frac{g_k^T y_{k-1}}{\|g_{k-1}\|^2}$$

$$\beta_k^{FR} = \frac{\|g_k\|^2}{\|g_{k-1}\|^2}$$

$$\beta_k^{HS} = \frac{g_k^T y_{k-1}}{d_{k-1}^T y_{k-1}}$$

$$\beta_k^{DY} = \frac{\|g_k\|^2}{d_{k-1}^T y_{k-1}}$$

$$\beta_k^{LS} = \frac{g_k^T y_{k-1}}{-g_k^T d_{k-1}}$$

$$\beta_k^{CD} = \frac{\|g_k\|^2}{-g_k^T d_{k-1}}$$

şeklinde tanımlanmaktadır (Hestenes and Stiefel 1952; Fletcher and Reeves 1964; Polak and Ribiere 1969; Liu and Storey 1991; Dai and Yuan 1999; Fletcher 2013).

Bu tezde eşlenik gradyan parametresi

$$\beta_k^{DEI} = \frac{\|g_k\|^2 - \frac{\|g_k\|}{\|d_{k-1}\|} |g_k^T d_{k-1}|}{\|g_{k-1}\|^2 + \mu \|g_k\| \|d_{k-1}\|}$$

şeklinde olan yeni bir eşlenik gradyan algoritması tanımlanmıştır. DEI eşlenik gradyan algoritması olarak adlandırılan algoritmanın, Armijo ve Wolfe çizgi arama tekniği kullanılarak ele alınan minimizasyon probleminin çözümüne global yakınsak olduğu ve yeterli iniş özelliğini sağladığı gösterilmiştir. Aynı zamanda tanımlanan algoritmanın

## 1.GİRİŞ

---

performans profili ile beraber görüntü iyileştirme problemine uygulaması verilerek literatürde var olan bazı algoritmalar ile karşılaştırılmıştır.

Bu tez altı bölümden oluşmaktadır. İlk bölüm olan giriş bölümünün ardından ikinci bölümde literatürdeki bazı çalışmalar hakkında bilgi verilmiştir. Üçüncü bölümde tezde kullanılan temel tanım, teorem ve önermelerin bulunduğu temel kavramlar yer almaktadır. Dördüncü bölümde kısıtsız optimizasyon algoritmalarının genel özellikleri ile birlikte eşlenik gradyan algoritmalarının elde edilişi hakkında bilgiler verilmiştir. Bu tezin orijinal kısmı olan beşinci bölümde yeni bir eşlenik gradyan algoritması tanımlanmış ve elde edilen teorik sonuçlar ve bazı uygulamalar verilmiştir. Altıncı ve son bölümde ise sonuçlar ve öneriler yer almaktadır.

### 2. KAYNAK ÖZETLERİ

Bu bölümde eşlenik gradyan algoritmalar ile ilgili literatürde bulunan bazı çalışmalardan bahsedilecektir.

Wei et al. (2006) uygun teorik özelliklere ve sayısal etkilere sahip bir eşlenik gradyan algoritma tanımlamak için Polak and Ribiere (1969) çalışmasından ilham alarak

$$\beta_k^{WYL} = \frac{\|g_k\|^2 - \frac{\|g_k\|}{\|g_{k-1}\|} g_k^T g_{k-1}}{\|g_{k-1}\|^2}$$

şeklinde yeni bir eşlenik gradyan parametresini tanımlamışlardır. Elde edilen WYL algoritmasının Wolfe-Powell çizgi arama tekniği altında yeterli iniş özelliğini sağladığı ve global yakınsak olduğu ispatlamışlardır. Ayrıca yazarlar WYL, PRP ve PRP+ algoritmalarını bazı test fonksiyonlarına uygulayarak kendi tanımladıkları algoritmanın daha iyi performans sergilediğini göstermişlerdir.

Huang et al. (2007) Güçlü Wolfe çizgi arama tekniği altında WYL algoritmasının yeterli iniş özelliğini analiz etmiştir.

Jiang et al. (2011) JMJ algoritmasının eşlenik gradyan parametresini

$$\beta_k^{JMJ} = \frac{\|g_k\|^2 - \frac{\|g_k\|}{\|d_{k-1}\|} |g_k^T d_{k-1}|}{d_{k-1}^T (g_k - g_{k-1})}$$

şeklinde tanımlayarak, standart Wolfe çizgi arama tekniğini kullanarak ilgili algoritmanın genel konveks olmayan fonksiyonlar için global yakınsak olduğunu kanıtlamışlardır.

Rivaie et al. (2012) PRP algoritmasının paydasındaki  $\|g_{k-1}\|^2$  terimini  $\|d_{k-1}\|^2$  ile değiştirip eşlenik gradyan parametresini

$$\beta_k^{RMIL} = \frac{\|g_k\|^2 - g_k^T g_{k-1}}{\|d_{k-1}\|^2}$$

şeklinde tanımlayarak RMIL algoritmasını önermişlerdir ve global yakınsamayı kesin (exact) çizgi arama tekniğiyle elde etmişlerdir. Bu algoritmayı FR, PRP, HS, LS, DY ve CD algoritmaları ile bazı test fonksiyonları üzerinden kıyaslayarak elde ettikleri RMIL algoritmasının diğer algoritmalara göre daha üstün olduğunu göstermişlerdir.

Dai (2016) RMIL'den daha iyi teorik özellikler elde etmek için RMIL algoritmasını modifiye ederek yeni RMIL+ algoritmasının eşlenik gradyan parametresini

$$\beta_k^{RMIL+} = \begin{cases} \beta_k^{RMIL}, & 0 \leq |g_k^T g_{k-1}| \leq \|g_k\|^2, \\ 0, & \text{diğer durumlarda,} \end{cases}$$

şeklinde vermiştir. Bu algoritmanın yeterli iniş özelliği ve global yakınsaması kesin çizgi arama tekniği altında incelenmiştir.

Yousif (2020) RMIL+ algoritmasına Güçlü Wolfe çizgi arama tekniğini uygulayarak bu algoritmanın yeterli iniş özelliğini sağladığını ve global yakınsadığını ispatlamıştır. Ayrıca RMIL+, RMIL, FR ve NPRP algoritmalarını Güçlü Wolfe çizgi arama tekniği altında inceleyerek RMIL+'nın diğerlerine kıyasla daha az iterasyonla daha iyi sonuçlar elde edebildiğini göstermişlerdir.

Yao et al. (2007) WYL algoritmasının payını, HS algoritmasının payına uygulayarak CG parametresini

$$\beta_k^{VHS} = \frac{\|g_k\|^2 - \frac{\|g_k\|}{\|g_{k-1}\|} g_k^T g_{k-1}}{d_{k-1}^T (g_k - g_{k-1})}$$

## 2. KAYNAK ÖZETLERİ

---

şeklinde tanımlamışlar ve böylece VHS algoritmasını elde etmişlerdir. Yazarlar Güçlü Wolfe çizgi arama tekniğinde  $\sigma < 1/3$  seçimi altında VHS algoritmasının yeterli iniş özelliğini sağladığını kanıtlamışlardır.

Zhang (2009) WYL algoritmasındaki  $g_k^T g_{k-1}$  teriminin mutlak değerini alarak eşlenik gradyan parametresini

$$\beta_k^{NPRP} = \frac{\|g_k\|^2 - \frac{\|g_k\|}{\|g_{k-1}\|} |g_k^T g_{k-1}|}{\|g_{k-1}\|^2}$$

şeklinde tanımladığı NPRP algoritmasını elde etmiştir. Yazar algoritmanın Güçlü Wolfe çizgi arama tekniğinde  $\sigma \in (0, \frac{1}{2})$  parametresi ile yeterli iniş özelliğini sağladığını ve konveks olmayan fonksiyonlar için global yakınsadığını kanıtlamıştır. Ayrıca Zhang bu sonucu HS algoritmasına genişleterek NHS algoritmasının eşlenik gradyan parametresini

$$\beta_k^{NHS} = \frac{\|g_k\|^2 - \frac{\|g_k\|}{\|g_{k-1}\|} |g_k^T g_{k-1}|}{d_{k-1}^T (g_k - g_{k-1})}$$

şeklinde tanımlamıştır. Ayrıca,  $\sigma$  parametresinin  $(0, \frac{1}{2})$  ile sınırlandırılması durumunda NHS yönteminin de Güçlü Wolfe çizgi arama tekniği ve standart Wolfe çizgi arama tekniği ile birlikte yeterli iniş özelliğini sağladığını ve global yakınsak olduğunu kanıtlamıştır. Ayrıca NPRP ile VPRP algoritmasını ve NHS ile de VHS algoritmasını kıyaslayarak kendi elde ettiği NPRP ve NHS algoritmalarının CUTE kütüphanesinde verilen problemler üzerinde diğerlerinden daha etkili olduğunu göstermiştir.

Dai and Wen (2012) NPRP algoritmasını geliştirerek yeni tanımladıkları DPRP algoritmasının CG parametresini

$$\beta_k^{DPRP} = \frac{\|g_k\|^2 - \frac{\|g_k\|}{\|g_{k-1}\|} |g_k^T g_{k-1}|}{\mu |g_k^T d_{k-1}| + \|g_{k-1}\|^2}, \mu > 1$$

şeklinde tanımlamışlardır. Bu algoritmada  $f$  amaç fonksiyonunun konveks olması ve adım boyutunun kesin çizgi arama tekniğiyle elde edilmesi durumunda algoritmanın FR algoritmasına dönüştüğünü, Armijo ve Wolfe çizgi arama tekniklerinin kullanılması durumunda ise yeterli iniş özelliğinin sağlandığını ve global olarak yakınsak olduğunu göstermişlerdir.

Chen et al. (2023), DPRP algoritmasından ilham alarak PRP-T algoritmasını önermişlerdir. Bu algoritmada DPRP algoritmasının paydasındaki  $|g_k^T d_{k-1}|$  terimini  $\|g_k - g_{k-1}\| \|d_{k-1}\|$  terimi ile değiştirip, payındaki  $\frac{\|g_k\|}{\|g_{k-1}\|}$  çarpanı çıkarmışlardır. Dolayısıyla algoritmanın eşlenik gradyan parametresi

$$\beta_k^{PRP-T} = \frac{\|g_k\|^2 - g_k^T g_{k-1}}{\|g_{k-1}\|^2 + \mu \|g_k - g_{k-1}\| \|d_{k-1}\|}, \mu > 1$$

şeklinde tanımlanmıştır. PRP-T algoritmasında konveks olmayan  $f$  fonksiyonunun herhangi bir çizgi arama tekniğinden bağımsız olarak yeterli iniş özelliğini sağladığı ve bu algoritmanın hem Armijo hem de Wolfe çizgi arama tekniğiyle global yakınsak olduğunu ispatlamışlardır. Yazarlar PRP-T, DPRP, NHS, IWYL ve PRP+ algoritmalarının performanslarını karşılaştırarak PRP-T algoritmasının daha iyi performans sergilediğini göstermişlerdir. Ayrıca PRP-T ve PRP algoritmalarını görüntü restorasyon problemine uygulayarak görüntü restorasyonunda PRP-T algoritmasının PRP algoritmasından daha iyi sonuç verdiğini göstermişlerdir.

Hu et al. (2022) PRP-T algoritmasından ilham alarak yeni bir eşlenik gradyan algoritması oluşturmuştur ve bu algoritmanın CG parametresi

$$\beta_k^{NPRP} = \frac{\|g_k\|^2 - \frac{\|g_k\|}{\|g_{k-1}\|} g_k^T g_{k-1}}{\max\{\tau \|g_k\| \|d_{k-1}\|, \|g_{k-1}\|^2\}}, \tau > 2$$

şeklinde tanımlamışlardır. Bu oluşturulan algoritmada herhangi bir çizgi arama tekniğinden bağımsız olarak yeterli iniş özelliğinin sağlandığını ve fonksiyonun

konveks olması şartı aranmaksızın Wolfe ve Armijo çizgi arama tekniğiyle global yakınsak olduğu gösterilmiştir.

Wu and Xiaoyu (2023) RMIL algoritmasının hibrit versiyonunu ele alarak yeni bir eşlenik gradyan algoritması önermişlerdir ve bu algoritmanın CG parametresi

$$\beta_k^{VRMIL} = \frac{\|g_k\|^2 - \frac{\|g_k\|}{\|d_{k-1}\|} g_k^T d_{k-1}}{\max\{\mu \|g_k\| \|d_{k-1}\|, \|d_{k-1}\|^2\}}, \mu > 2$$

şeklinde tanımlamışlardır. Bu algoritmanın herhangi bir çizgi arama tekniğinden bağımsız olarak yeterli iniş özelliğini sağladığı ve Lipschitz süreklilik şartı olmadan Wolfe çizgi arama tekniği altında global yakınsak olduğu gösterilmiştir. Ayrıca VRMIL, DPRP, NPRP, RMIL+ algoritmalarını görüntü restorasyon problemi üzerinden kıyaslayarak VRMIL algoritmasının diğer algoritmalarından daha etkili olduğunu göstermişlerdir.

Salih et al. (2018) daha önce yapılan çalışmalardan ilham alarak YHM algoritmasını önermişler ve bu algoritmanın CG parametresini

$$\beta_k^{YHM} = \begin{cases} \frac{\|g_k\|^2 - g_k^T g_{k-1}}{\|g_{k-1}\|^2}, & 0 \leq g_k^T g_{k-1} \leq \|g_k\|^2, \\ \frac{\|g_k\|^2 - \frac{\|g_k\|}{\|g_{k-1}\|} |g_k^T g_{k-1}|}{\|g_{k-1}\|^2}, & \text{diğer durumlar.} \end{cases}$$

şeklinde belirlemişlerdir. Yapılan çalışmalar sonucunda bu algoritmanın yeterli iniş özelliğini sağladığını ve Wolfe çizgi arama tekniği altında global yakınsak olduğunu göstermişlerdir.

Jiang and Jian (2013), DY ve FR algoritmasından ilham alarak modifiye edilmiş eşlenik gradyan parametreler önermişlerdir. Bu algoritmalarındaki CG parametreler

$$\beta_k^{MDY} = \frac{\|g_k\|^2}{\max\{d_{k-1}^T(g_k - g_{k-1}), \mu|g_k^T d_{k-1}|\}}, \mu > 1$$

ve

$$\beta_k^{MFR} = \frac{\|g_k\|^2}{\max\{\|g_{k-1}\|^2, \mu|g_k^T d_{k-1}|\}}, \mu > 1$$

şeklinde tanımlamış ve elde edilen algoritmaların yeterli iniş özelliğini sağladıkları ve global yakınsak oldukları gösterilmiştir. Daha sonra MDY, HS, DHS, DY, MFR, PRP, DPRP ve FR eşlenik gradyan algoritmalarını bazı test problemleri üzerinden kıyaslayarak MDY ve MFR algoritmalarından daha iyi sayısal sonuçlar elde etmişlerdir.

Jiang et al. (2014) DY ve FR algoritmalarında olduğu gibi PRP algoritmasını da modifiye etmişler ve

$$\beta_k^{JPRP} = \frac{\|g_k\|^2 - \frac{\|g_k\|}{\|g_{k-1}\|} g_k^T g_{k-1}}{\max\{\mu|g_k^T d_{k-1}|, \|g_{k-1}\|^2\}}, \mu > 1$$

CG parametresini önermişlerdir. Yazarlar JPRP algoritmasının hiçbir çizgi arama tekniği olmadan her adımda yeterli iniş özelliğini sağladığını ve Wolfe çizgi arama tekniğiyle beraber global yakınsak olduğunu ispatlamışlardır.

### 3. TEMEL KAVRAMLAR

Bu bölümde tezde kullanılan bazı temel kavramlar hakkında bilgiler verilecektir.

#### 3.1. Temel Tanım ve Teoremler

**Tanım 3.1.1:**  $S, \mathbb{R}^n$  nin açık bir alt kümesi ve  $f: S \rightarrow \mathbb{R}^n$  dönüşümü verilmiş olsun. Üstelik  $f$  nin  $x \in S$  noktasında  $f_{x_1}(x), \dots, f_{x_n}(x)$  kısmi türevlerinin mevcut olduğunu kabul edelim. Bu takdirde

$$\nabla f(x) = (f_{x_1}(x), \dots, f_{x_n}(x))$$

olarak tanımlanan  $\nabla f$  vektör değerli fonksiyona  $f$  nin gradyanı denir. Bazen  $\nabla f$  yerine  $\text{grad} f$  yazılır (Bayraktar 2010).

**Tanım 3.1.2:**  $S \subseteq \mathbb{R}^n$  bir bölge ve  $x_0 \in S^\circ$  iç nokta olsun.  $f, x_0$  da diferansiyellenebilir ve bu noktada  $f$  nin birinci mertebeden kısmi türevleri sıfır yani  $\nabla f(x_0) = 0$  ise veya  $f$  bu noktada diferansiyellenemez ise bu  $x_0$  noktasına  $f$  nin kritik noktası denir (Bayraktar 2010).

**Teorem 3.1.3:**  $f: S \subseteq \mathbb{R}^2 \rightarrow \mathbb{R}$  fonksiyonu verilmiş ve  $f_x$  ve  $f_y$  kısmi türevleri mevcut olsun. Ayrıca  $(a_1, a_2)$  ve  $(a_1 + h_1, a_2)$  noktalarını  $\mathbb{R}^2$  de birleştiren  $\tau_1$  doğru parçası,  $f$  ve  $f_x$  in tanım kümesinde olsun. Bu takdirde

$$f(a_1 + h_1, a_2) - f(a_1, a_2) = h_1 f_x(c_1, a_2)$$

olacak şekilde  $a_1$  ve  $a_1 + h_1$  noktaları arasında ( $\tau_1$  üzerinde) bir  $c_1$  reel sayısı vardır (Bayraktar 2010).

**Tanım 3.1.4:**  $f \in C[a, b], f' \in C(a, b)$  ve  $f', [a, b]$  de sürekli bir fonksiyona genişletilebilirse  $f, [a, b]$  de sürekli diferansiyellenebilirdir denir (Bayraktar 2010).

**Tanım 3.1.5:**  $f: S \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$  dönüşümü için

$$\|f(x) - f(y)\| \leq L\|x - y\|$$

olacak şekilde bir  $L > 0$  varsa  $f$ ,  $S$ 'de Lipschitz süreklidir denir (Bayraktar 2010).

**Tanım 3.1.6:**  $f: \mathbb{R}^n \rightarrow \mathbb{R}$  fonksiyonu verilsin. Eğer  $f$  nin gradyanı  $\nabla f$ , Lipschitz sürekli bir dönüşüm ise  $f$  ye düzgün dönüşüm denir (Nocedal 2006).

**Tanım 3.1.7:**  $S, \mathbb{R}^n$  nin bir alt kümesi,  $f: S \subset \mathbb{R}^n \rightarrow \mathbb{R}$  ve  $x_0 \in S$  olsun. Her  $x \in S$  için  $f(x_0) \leq f(x)$  ( $f(x_0) \geq f(x)$ ) ise  $f$  nin  $x_0$  da global (mutlak) minimumu (global maksimumu) vardır denir. Bu halde  $f(x_0)$  değerine  $f$  nin global minimumu (global maksimumu) ve  $x_0$ 'a ise global minimizer (global maksimizer) denir (Bayraktar 2010).

**Tanım 3.1.8:**  $S, \mathbb{R}^n$  nin bir alt kümesi,  $f: S \subset \mathbb{R}^n \rightarrow \mathbb{R}$  ve  $x_0 \in S$  olsun. Her  $x \in N_\delta(x_0) \cap S$  için  $f(x_0) \leq f(x)$  ( $f(x_0) \geq f(x)$ ) olacak şekilde  $x_0$ 'ın bir  $N_\delta(x_0)$  civarı varsa  $f$  nin  $x_0$  da lokal (yerel) minimumu (lokal maksimumu) vardır denir.  $f(x_0)$  değerine  $f$  nin lokal minimumu (lokal maksimumu) ve  $x_0$ 'a ise lokal minimizer (lokal maksimizer) denir (Bayraktar 2010).

**Örnek 3.1.9:**  $f: (0,1] \rightarrow \mathbb{R}$  ,  $f(x) = x^3 - 3x$  fonksiyonu verilmiş olsun.  $f'(x) = 0$  şartını sağlayan  $x = 1$  noktasında lokal minimum vardır.  $x = 1$  noktası ise lokal minimizerdir.

**Örnek 3.1.10:**  $f: \mathbb{R}^2 \rightarrow \mathbb{R}$  ,  $f(x, y) = \sqrt{x^2 + y^2}$  olarak tanımlanan  $f$  fonksiyonunu için  $f(0,0) = 0 \leq f(x, y)$  olduğundan  $f$ 'nin  $(0,0)$ 'da global minimumu vardır.  $\|(x, y)\| \rightarrow \infty$  için  $f(x, y) = \sqrt{x^2 + y^2} \rightarrow \infty$  olduğundan  $f$ 'nin global maksimumu yoktur.

**Tanım 3.1.11:** Her  $x \in N_\delta(x_0^*)$ ,  $x \neq x_0^*$  için  $f(x_0^*) < f(x)$  olacak şekilde bir  $N_\delta(x_0^*)$  varsa  $x_0^*$ a  $f$ 'nin kesin lokal minimizeri denir (Nocedal 2006).

**Tanım 3.1.12:** Eşleniğinin transpozu kendisine eşit olan matrislere Hermityen matris denir (Alagöz vd 2012).

Örneğin  $A = \begin{bmatrix} 10 & 5 \\ 5 & 8 \end{bmatrix}$  matrisi için  $(\overline{A})^T = A$  olduğundan  $A$  Hermityen matristir.

**Tanım 3.1.13:**  $A$ ,  $n \times n$  tipinde bir Hermityen matris olsun.  $A$  matrisi, sıfırdan farklı her  $x \in \mathbb{R}^n$  için  $\langle Ax, x \rangle > 0$  şartını sağlıyorsa bu matrise pozitif tanımlı matris denir ve  $A > 0$  ile gösterilir. Eğer her  $x \in \mathbb{R}^n$  için  $\langle Ax, x \rangle \geq 0$  şartı sağlanıyorsa  $A$  matrisine pozitif yarı tanımlı matris denir ve  $A \geq 0$  ile gösterilir. Her pozitif tanımlı matris aynı zamanda pozitif yarı tanımlı matris olup tersi her zaman geçerli değildir. Bir pozitif yarı tanımlı matrisin pozitif tanımlı matris olması için gerek ve yeter şart matrisin tersinir olmasıdır (Bhatia 1996).

Örneğin  $\begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix}$  matrisi pozitif yarı tanımlıdır ama pozitif tanımlı değildir.

$\begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$  matrisi ise hem pozitif tanımlı, hem de pozitif yarı tanımlı bir matristir.

**Tanım 3.1.14:**  $S \in \mathbb{R}^n$  kümesindeki herhangi iki noktayı birleştiren doğru parçası tamamen  $S$ 'nin içinde yer alıyorsa  $S$ 'ye konveks küme denir. Yani  $S$  konveks bir küme ise herhangi iki  $x, y \in S$  noktaları için  $\alpha \in [0,1]$  olmak üzere  $\alpha x + (1 - \alpha)y \in S$  dir. Eğer  $f$  fonksiyonun tanım kümesi  $S$  konveks küme ve her  $x, y \in S$  ve her  $\alpha \in [0,1]$  için

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

eşitsizliğini sağlarsa  $f$ 'ye konveks fonksiyon denir. Eğer  $x \neq y$  ve  $\alpha \in (0,1)$  olmak üzere

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

ise  $f$ 'e kesin konveks fonksiyon denir. Ayrıca  $-f$  konveks ise  $f$ 'e konkav fonksiyon denir (Nocedal 2006).

### 3.TEMEL KAVRAMLAR

Örneğin  $S = \{(x_1, x_2): x_1^2 + x_2^2 \leq 5\} \subset \mathbb{R}^2$  kümesi  $\mathbb{R}^2$  de konveks bir küme ve  $f: \mathbb{R}^2 \rightarrow \mathbb{R}, f(x_1, x_2) = 8x_1^2 + x_2^2$  fonksiyonu da  $\mathbb{R}^2$  de konveks bir fonksiyondur.

Belli bir bölgede konveks olup da  $\mathbb{R}^n$  üzerinde konveks olmayan fonksiyonlar da tanımlanabilir. Örnek olarak  $f: \mathbb{R} \rightarrow \mathbb{R}, f(x) = x^3$ ,  $\mathbb{R}$  üzerinde konveks olmayıp  $S = \{x: x \geq 0\}$  üzerinde konvektir.

**Tanım 3.1.15:**  $V \subseteq \mathbb{R}^n$  boş kümeden farklı bir küme ve  $f: V \rightarrow \mathbb{R}$  bir fonksiyon olsun.  $f$  fonksiyonunun epigrafı

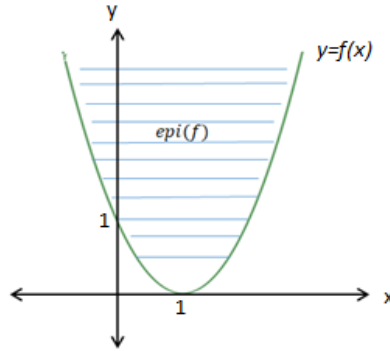
$$epi(f) = \{(x, w): x \in V, w \in \mathbb{R}, f(x) \leq w\}$$

şeklinde tanımlanır (Tiel 1984).

**Örnek 3.1.16:**  $f(x) = (x - 1)^2$  fonksiyonunun epigrafı

$$epi(f) = \{(x, w) \in \mathbb{R}^2: (x - 1)^2 \leq w\}$$

olup bu kümenin geometrik yeri aşağıda verilmiştir.



**ŞEKİL 3.1.**  $f(x) = (x - 1)^2$  fonksiyonunun epigrafı

**Tanım 3.1.17:** Herhangi bir lineer denklem sistemini çözmek için gözlemlenen ve tahmin edilen değer arasındaki fark vektörüne rezidü denir.  $r$  rezidü vektörü,  $b$

gözlemlenen değerlerin vektörü,  $A$  katsayı matrisi ve  $x$  çözüm vektörü olmak üzere  $Ax = b$  biçimindeki denklem sisteminin rezidüsü

$$r = b - Ax$$

şeklinde hesaplanır (Shewchuk, 1994).

**Tanım 3.1.18:** Bir  $V$  vektör uzayı üzerinde  $\langle . \rangle$  iç çarpım fonksiyonu tanımlı olsun.  $x, y \in V$  için eğer  $\langle x, y \rangle = 0$  ise  $x$  ve  $y$  vektörleri ortogonal vektörlerdir denir. Benzer şekilde  $V = \{x_1, x_2, \dots, x_n\}$  olmak üzere  $\forall i \neq j$  için  $\langle x_i, x_j \rangle = 0$  ise  $V$  kümesine ortogonal küme denir (Venit, 1985).

**Örnek 3.1.19:**  $\mathbb{C}[-\pi, \pi]$  sürekli fonksiyonların uzayında  $f(x) = \sin x$  ile  $g(x) = 1$  fonksiyonları için

$$\langle \sin x, 1 \rangle = \int_{-\pi}^{\pi} \sin x \, dx = 0$$

olup  $f$  ve  $g$  fonksiyonları ortogonaldir.

**Tanım 3.1.20:**  $A$ ,  $n \times n$  tipinde simetrik, pozitif tanımlı bir matris,  $u$  ve  $v$  sıfırdan farklı vektörler olmak üzere

$$u^T A v = 0$$

ise bu iki vektör  $A$ 'ya göre eşlenik vektörlerdir denir.

**Örnek 3.1.21:**  $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$  simetrik, pozitif tanımlı,  $u = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$  ve  $v = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$  biçiminde birer vektör olmak üzere  $u^T A v = 0$  olduğundan,  $u$  ve  $v$  vektörleri  $A$ 'ya göre eşlenik vektörlerdir.

Aşağıda, eşlenik gradyan algoritmasının elde edilmesinde önemli bir yere sahip olan Gram-Schmidt yöntemi verilmiştir. Bu yöntem bir iç çarpım uzayında herhangi bir vektör kümesini ortogonalleştirmek için kullanılır.

**Tanım3.1.22:**  $V$  bir iç çarpım uzayı ve  $\{u_1, u_2, \dots, u_n\}$ , bu uzayın bir bazı olsun. Bu durumda

$$v_1 = u_1,$$

$$v_2 = u_2 - \frac{\langle u_2, v_1 \rangle}{\langle v_1, v_1 \rangle} v_1,$$

⋮

$$v_n = u_n - \frac{\langle u_n, v_1 \rangle}{\langle v_1, v_1 \rangle} v_1 - \frac{\langle u_n, v_2 \rangle}{\langle v_2, v_2 \rangle} v_2 - \dots - \frac{\langle u_n, v_{n-1} \rangle}{\langle v_{n-1}, v_{n-1} \rangle} v_{n-1}$$

eşitlikleri yardımıyla elde edilen  $\{v_1, v_2, \dots, v_n\}$  kümesi  $V$ 'nin bir ortogonal bazıdır (Horn,1985).

### 3.2. Kısıtsız Optimizasyonun Temelleri

Matematiksel olarak optimizasyon, değişkenler üzerindeki kısıtlamalara tabi olan bir fonksiyonun minimizasyonu veya maksimizasyonu ile ilgilenen bir konudur. Bu tezde optimizasyon problemi olarak minimizasyon problemi ele alınacaktır.

$x$ ; bilinmeyenler veya parametreler olarak adlandırılan değişkenlerin vektörü,  $f$ ; minimize veya maksimize etmek istediğimiz  $x$ 'in (skaler) bir fonksiyonu olan amaç

### 3.TEMEL KAVRAMLAR

---

fonksiyonu,  $c_i$ ;  $x$  vektörünün karşılaması gereken belirli denklemleri ve eşitsizlikleri tanımlayan  $x$ 'in kısıtlama fonksiyonları olmak üzere kısıtlı optimizasyon probleminin genel formu:

$$\min_{x \in \mathbb{R}^n} f(x) \quad \begin{array}{l} c_i(x) = 0, i \in \mathcal{E} \\ c_i(x) \geq 0, i \in I \end{array} \quad (3.1)$$

şeklindedir. Burada verilen  $\mathcal{E}$  ve  $I$  sırasıyla, eşitlik ve eşitsizlik kısıtlamaları için indis kümeleri olup problemde verilen tüm kısıtları sağlayan noktaların kümesine olurlu bölge (feasible region) denir.

Yukarıda genel formu verilen problem, amaç fonksiyonunun niteliğine, kısıtlamalarına (lineer, lineer olmayan, konveks), değişken sayısına (büyük veya küçük), fonksiyonların düzgünlüğüne (diferansiyellenebilir veya diferansiyellenemez) vb. bağlıdır.

$\mathcal{E} = I = \emptyset$  ise (3.1) problemine kısıtsız optimizasyon problemi denir ve bu problem birçok pratik uygulamada karşımıza çıkar. Değişkenler üzerinde doğal kısıtlamalara sahip bazı problemler için bile, çözümü etkilemedikleri ve algoritmalara etkileri olmadıkları için bu kısıtları göz ardı etmek güvenli olabilir.

Optimizasyon problemindeki amaç fonksiyonunun ve olurlu bölgenin her ikisi de konveks ise, problemin herhangi bir lokal çözümü aslında global bir çözümdür. Konvekslik kavramı optimizasyon problemleri için temel bir kavramdır. Pek çok problem bu özelliğe sahiptir ve sahip olması hem teoride hem de pratikte problemin çözülmesini genellikle kolaylaştırır.

Verilen bir fonksiyonun minimize ederlerini incelemek için kullanılan temel matematiksel araç Taylor Teoremidir. Aşağıda Taylor Teoremi ve bazı temel teoremler verilmiştir.

**Teorem 3.2.1 (Taylor Teoremi):**  $f: \mathbb{R}^n \rightarrow \mathbb{R}$  sürekli türevlenebilir bir fonksiyon ve  $p \in \mathbb{R}^n$  olsun.  $t \in (0,1)$  olmak üzere

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

dır. Ayrıca  $f$  iki kez sürekli türevlenebilir ise  $t \in (0,1)$  olmak üzere

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

ve

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

dir (Nocedal 2006).

Aşağıda optimizasyon probleminin çözümü için birinci ve ikinci dereceden bazı gerekli ve yeterli koşullar verilmiştir.

**Teorem 3.2.2 (Birinci Dereceden Gerekli Koşullar):** Eğer  $x^*$  bir lokal minimizer ve  $f$ ,  $x^*$  in açık bir komşuluğunda sürekli türevlenebilir ise  $\nabla f(x^*) = 0$  dır (Nocedal 2006).

**Teorem 3.2.3 (İkinci Dereceden Gerekli Koşullar):** Eğer  $x^*$ ,  $f$ 'nin lokal minimizeri,  $\nabla^2 f(x^*)$  var ve  $x^*$  in açık bir komşuluğunda sürekli ise  $\nabla f(x^*) = 0$  ve  $\nabla^2 f(x^*)$  pozitif yarı tanımlıdır (Nocedal 2006).

**Teorem 3.2.4 (İkinci Dereceden Yeterli Koşullar):**  $\nabla^2 f$ ,  $x^*$  in açık bir komşuluğunda sürekli,  $\nabla f(x^*) = 0$  ve  $\nabla^2 f(x^*)$  pozitif yarı tanımlı olsun. Bu durumda  $x^*$ ,  $f$ 'nin kesin lokal minimizeridir (Nocedal 2006).

Amaç fonksiyonu konveks olduğunda lokal ve global minimizerleri karakterize etmek kolaydır. Aşağıdaki teorem bu durumla ilgilidir.

### 3.TEMEL KAVRAMLAR

---

**Teorem 3.2.5:**  $f$  konveks ise lokal minimizer aynı zamanda global minimizerdir. Ek olarak,  $f$  türevlenebilir ise bu durumda herhangi bir  $x^*$  kritik noktası,  $f$ 'nin global minimizeridir (Nocedal 2006).

Temel hesaba dayanan bu sonuçlar, kısıtsız optimizasyon algoritmalarının temellerini sağlar.



### 4. EŞLENİK GRADYAN ALGORİTMALAR

Bu bölümde kısıtsız optimizasyon algoritmalarının genel formu, çizgi arama teknikleri ve eşlenik gradyan algoritmasının genel yapısından bahsedilecektir.

#### 4.1. Optimizasyon Algoritmalarının Özellikleri

Optimizasyon algoritmaları genellikle iteratif olup  $x$  değişkeninin ilk tahminiyle başlarlar ve bir çözüme ulaşmayı umarak sonlanana kadar bir dizi iyileştirilmiş tahmin (iteratifler) üretirler. Bir iteratifden diğerine geçmek için kullanılan strateji bir algoritmayı diğerinden ayırır. Çoğu teknik,  $f$  amaç fonksiyonunun ve bu fonksiyonun birinci ve ikinci türevlerinin değerlerini kullanır. Bazı algoritmalar önceki iterasyonlarda toplanan bilgileri biriktirirken, bazıları ise sadece mevcut noktada elde edilen yerel bilgileri kullanır. Bu ayrıntılara bakılmaksızın, iyi algoritmalar aşağıdaki özelliklere sahip olmalıdır:

a) Sağlamlık: Başlangıç noktasının tüm makul değerleri için çok çeşitli problemlerde iyi performans göstermelidirler.

b) Verimlilik: Aşırı bilgisayar zamanı veya depolama gerektirmemelidirler.

c) Kesinlik: Algoritma bilgisayarda uygulandığında verilerdeki hatalara veya meydana gelen aritmetik yuvarlama hatalarına aşırı duyarlı bir şekilde çözümü tanımlayabilmelidirler.

Bu hedeflenen özellikler çatışabilir. Örneğin; büyük, kısıtsız, lineer olmayan bir problem için hızlı yakınsayan bir algoritma, çok fazla bilgisayar depolama alanı gerektirebilir. Diğer taraftan, sağlam bir algoritma aynı zamanda en yavaş algoritma da olabilir.

Kısıtsız optimizasyon algoritmalarının büyük çoğunluğu, şartları sağlayan keyfi bir  $x_0$  noktasından başlayarak, daha fazla ilerleme kaydedilemediğinde ya da çözüm

#### 4. EŞLENİK GRADYAN ALGORİTMASI

---

noktasının yeterli doğrulukla yaklaştırıldığı görüldüğünde sona eren bir iteratif  $\{x_k\}_{k=0}^{\infty}$  dizisini oluşturur. Bir  $x_k$  iteratifinden diğerine geçerken nasıl hareket edileceğine karar vermek için genelde daha önceki  $x_0, x_1, \dots, x_{k-1}$  iteratiflerinden gelen bilgiler ve  $f$  fonksiyonu hakkındaki bilgiler kullanılır. Bu bilgiler,  $f(x_k)$ 'dan daha küçük bir fonksiyon değerine sahip yeni bir  $f(x_{k+1})$  iterasyonu bulmak için kullanılır. Her adımda  $f$ 'nin azaltılmasında direkt olmayan monoton olmayan algoritmalar mevcuttur, ancak bu algoritmalar bile önceden belirlenmiş olan  $m$  sayıda iteratifden sonra  $f$ 'nin azalmasını gerektirir. Kısacası  $f(x_k) < f(x_{k-m})$  şeklinde olmalıdır.

Kısıtsız optimizasyon problemini çözmek için kullanılan algoritmaların genel yapısı

$$x_{k+1} = x_k + \alpha_k d_k \quad (4.1)$$

şeklindedir. Burada  $\alpha_k$  adım uzunluğu ve  $d_k$  ise araştırma yönü olarak adlandırılırlar.

$\beta_k$ , simetrik ve tekil olmayan bir matris olmak üzere arama yönünün genel formu,

$$d_k = -\beta_k^{-1} \nabla f_k$$

şeklindedir. En dik iniş algoritmasında (steepest descent method),  $\beta_k$  birim matrisi olan  $I$ , Newton algoritmasında ise  $\beta_k$  tam Hessian olan  $\nabla^2 f(x_k)$ 'dir. quasi-Newton algoritmasında  $\beta_k$ , Hessian'a her seferinde düşük dereceli bir formül aracılığıyla güncellenen iteratif bir yaklaşımdır.  $d_k$  yukarıdaki şekilde tanımlandığında ve  $\beta_k$  pozitif tanımlı olduğunda

$$d_k^T \nabla f_k = -\nabla f_k^T \beta_k^{-1} \nabla f_k < 0$$

elde edilir ve dolayısıyla  $d_k$  bir iniş yönüdür.

#### 4. EŞLENİK GRADYAN ALGORİTMASI

---

Her ne kadar bazı algoritmalarda  $\alpha_k$  sabit seçilse de sabit seçilmesinin dezavantajları vardır. Yani  $\alpha_k$  büyük seçildiğinde minimum noktaya ulaşmak zorlaşabilir ve küçük seçildiğinde ise minimum noktaya ulaşmak fazla zaman alabilir. Bu dezavantajlardan kurtulmak adına adım uzunluğunu belirlemek için her bir iterasyonda yeniden hesaplanması gereken çizgi arama teknikleri geliştirilmiştir.  $\alpha_k > 0$  olmak üzere  $\alpha_k$ 'nın sağlaması gereken en temel koşullardan biri,

$$f(x_k + \alpha_k d_k) < f(x_k) \quad (4.2)$$

şeklindedir. Bu koşul  $f$ 'de azalmayı gerektirir. Bu gereklilik yakınsama için yeterli olmayabilir. Her bir adımdaki  $f$  değerindeki yetersiz azalma, konveks fonksiyonun minimizerine yakınsamamasına neden olabilir. Bu durumdan kaçınmak için  $f$ 'ye  $c_1 \in (0,1)$  olmak üzere

$$f(x_k + \alpha_k d_k) \leq f(x_k) + c_1 \alpha_k \nabla f_k^T d_k \quad (4.3)$$

şeklinde tanımlanan yeterli azalma koşulu uygulanması gerekmektedir.

Çizgi arama tekniğinin başarısı hem  $d_k$  yönünün hem de  $\alpha_k$  adım uzunluğunun etkili seçimlerine bağlıdır. Çizgi arama tekniğinde, algoritma bir  $d_k$  yönü seçer ve daha düşük fonksiyon değerine sahip yeni bir iterasyon için mevcut olan  $x_k$  iterasyonunu kullanarak bir arama yapar.  $d_k$  boyunca hareket edilecek mesafe olan adım uzunluğu olan  $\alpha$ ,

$$\min_{\alpha > 0} f(x_k + \alpha d_k) \quad (4.4)$$

tek boyutlu minimizasyon probleminin yaklaşık olarak çözülmesiyle bulunabilir. Verilen minimizasyon problemini tam olarak çözerek  $d_k$  yönünden maksimum fayda elde edilir. Ancak kesin (exact) çizgi arama olarak adlandırılan bu yöntem masraflı olabilir ve genellikle gereksizdir. Bunun yerine çizgi arama tekniğini, yukarıda verilen problemin minimumuna genel olarak yaklaşan bir adım bulana kadar sınırlı sayıda

deneme adım uzunluğu üretir. Her yeni noktada yeni bir arama yönü ve adım uzunluğu hesaplanarak işlem tekrarlanır.

Çoğu çizgi arama tekniği  $d_k$ 'nin iniş yönü olmasını yani  $d_k^T \nabla f_k < 0$  eşitsizliğini sağlamasını gerektirir. Çünkü bu eşitsizlik  $f$  fonksiyonunun bu yönde azalabileceğini garanti eder.

#### 4.2. Çizgi Arama Teknikleri

Çizgi arama tekniği kesin ve kesin olmayan (inexact) çizgi arama teknikleri olarak kategorize edilebilir. Kesin çizgi arama tekniği, adından da anlaşılacağı gibi, her iterasyonda kesin minimizeri bulmayı amaçlarken kesin olmayan çizgi arama tekniğinde minimum değere yaklaşık bir değer elde etmek için adım uzunlukları belirler. Kesin çizgi arama tekniğinin genel formu

$$f(x_k + \alpha_k d_k) = \min_{\alpha \geq 0} f(x_k + \alpha d_k)$$

şekindedir. Hesaplamalı optimizasyonun ilk yıllarında, birçok araştırmacı kesin çizgi arama tekniğinin oldukça yavaş olduğunu görmüş ve bunun yerine kesin olmayan çizgi arama tekniğini kullanmayı seçmiştir. Aşağıda ayrıntılı verilecek olan Armijo, Güçlü Wolfe ve Zayıf Wolfe gibi çizgi arama teknikleri, kesin olmayan çizgi arama tekniklerinden bazılarıdır.

Kesin olmayan bir çizgi arama tekniğinde  $c_1 \in (0,1)$  olmak üzere her şeyden önce  $\alpha_k$ 'nin

$$f(x_k + \alpha_k d_k) \leq f(x_k) + c_1 \alpha_k \nabla f_k^T d_k \quad (4.5)$$

eşitsizliğiyle ölçülen  $f$  amaç fonksiyonunda yeterli azalmayı sağlaması gerekir. Yukarıda verilen eşitsizlik genellikle Armijo Koşulu olarak adlandırılmaktadır. Bu teknikteki amaç,  $0 < \rho < 1$  ve  $j_k$  pozitif bir tamsayı olmak üzere (4.5) eşitsizliğini sağlayan  $\alpha_k = \rho^{j_k}$  şeklindeki en büyük  $\alpha_k$  adım uzunluğunu bulmaktır. Yeterli azalma

#### 4. EŞLENİK GRADYAN ALGORİTMASI

---

koşulu  $\alpha_k$ 'nın küçük değerleri için makul bir ilerleme sağlar fakat algoritmanın doğru çalışmasını sağlamak için tek başına yeterli değildir. Kabul edilemeyecek kadar kısa adımları ortadan kaldırmak için  $\alpha_k$ 'nın ikinci bir gereklilik olan eğrilik koşulunu sağlaması gerekmektedir. Eğrilik koşulu  $c_2 \in (c_1, 1)$  olmak üzere

$$\nabla f(x_k + \alpha_k d_k)^T d_k \geq c_2 \nabla f_k^T d_k \quad (4.6)$$

şeklinindedir.  $0 < c_1 < c_2 < 1$  olmak üzere (4.5) ve (4.6) eşitsizlikleri toplu olarak Wolfe Koşulu olarak adlandırılmaktadır.

(4.5) eşitsizliği ile birlikte  $\alpha_k$

$$|\nabla f(x_k + \alpha_k d_k)^T d_k| \leq c_2 |\nabla f_k^T d_k| \quad (4.7)$$

eşitsizliğini sağlayacak şekilde belirlenirse bu tekniğe güçlü Wolfe çizgi arama tekniği denir.

$\rho \in (0,1)$ ,  $\sigma_1 \in (0,1)$  ve  $\sigma_2 > 0$  olmak üzere Armijo tipi çizgi arama tekniği  $\alpha_k$  adım uzunluğunu, aşağıdaki eşitsizlikte  $\alpha_k$  yerine  $\rho^j$ ,  $j = 0,1,2, \dots$  yazarak

$$f(x_k + \alpha_k d_k) - f(x_k) \leq \sigma_1 \alpha_k \nabla f_k^T d_k - \sigma_2 \alpha_k^2 \|d_k\|^2$$

eşitsizliğini sağlayan en büyük  $\rho^j$  olarak belirler.

$0 < \rho < 1$  ve  $0 < \sigma < 1$  olmak üzere Wolfe tipi çizgi arama tekniği  $\alpha_k$ 'yı belirlemek için

$$f(x_k + \alpha_k d_k) - f(x_k) \leq -\rho \alpha_k^2 \|d_k\|^2$$

$$\nabla f(x_k + \alpha_k d_k)^T d_k \geq -2\sigma \alpha_k \|d_k\|^2$$

eşitsizliklerini kullanmaktadır.

### 4.3. Eşlenik Gradyan Algoritmalar

Bu bölümde eşlenik gradyan algoritmaların elde edilişi hakkında bilgiler verilecektir.

$A$ ,  $n \times n$  tipinde simetrik ve pozitif tanımlı bir matris ve  $x$  bir vektör olmak üzere eşlenik gradyan algoritması ilk olarak

$$Ax = b$$

şeklindeki lineer denklem sisteminin çözümü için tanımlanmıştır. Sıfırdan farklı  $u$  ve  $v$  vektörlerinin  $A$ 'ya göre eşlenik olması

$$u^T Av = 0$$

eşitliğinin sağlanmasını gerektirir. Öte yandan  $A$ , simetrik ve pozitif tanımlı bir matris olduğundan bu matris yardımıyla

$$\langle u, v \rangle_A := u^T Av$$

şeklinde bir iç çarpım tanımlanabilir. Bu iç çarpıma göre iki vektörün eşlenik olması için gerek ve yeter şart bu iç çarpıma göre ortogonal olmaları gerektiği açıktır. Tanımlanan bu iç çarpım ele alınan problemin çözümünü bulmak için kullanılacak önemli bir araçtır.  $A$  matrisinin özel seçiminden dolayı  $Ax = b$  lineer denklem sisteminin çözümü ile

$$f(x) = \frac{1}{2}x^T Ax - x^T b, x \in \mathbb{R}^n$$

#### 4. EŞLENİK GRADYAN ALGORİTMASI

---

şeklinde tanımlı ikinci dereceden fonksiyonunun minimizeri çakışmaktadır.  $f(x)$  fonksiyonunun  $\nabla^2 f(x)$  Hessian matrisi  $A$  olduğundan minimizerin var olduğu açık olup bu minimizer

$$\nabla f(x) = Ax - b$$

eşitliğini sağlar. Bu minimizeye ulaşmak için gerekli olan  $d_0$  başlangıç araştırma yönü  $d_0 = -\nabla f(x_0) = b - Ax_0$  şeklinde alınır. Burada temel mantık birbirleri ile eşlenik olan  $\mathbb{R}^n$  nin bir ortogonal tabanını elde etmektir.  $d_0$  vektörü bu tabanın ilk vektörüdür. Tabandaki diğer vektörlerin  $d_0$  ve dolayısıyla gradyanla eşlenik olması gerektiğinden bu algoritmaya eşlenik gradyan algoritması adı verilir.  $d_0$ 'ın algoritmanın ilk adımındaki rezidü olduğu aşikâr olup,  $r_k$  ise  $k$ .adımdaki rezidü olmak üzere  $r_k = b - Ax_k$  dir. Tabandaki vektörlerin birbirlerine eşlenik olmaları gerektiğinden bunu sağlamanın bir yolu, bir sonraki arama yönünün mevcut rezidü ve bir önceki arama yönlerinden oluşturmaktır. Vektörlerin eşlenik olmaları ile  $A$  matrisine göre ortogonal olmaları denk olduğundan diğer araştırma yön vektörleri

$$d_k = r_k - \sum_{i < k} \frac{d_i^T A r_k}{d_i^T A d_i} d_i$$

şeklinde Gram-Schmidt yöntemi ile elde edilebilir. Bu yönü takip ederek bir sonraki en uygun nokta

$$\alpha_k = \frac{d_k^T r_k}{d_k^T A d_k}$$

olmak üzere

$$x_{k+1} = x_k + \alpha_k d_k$$

eşitliği ile belirlenir. Sonuç olarak  $A$ ; reel, simetrik ve pozitif tanımlı bir matris olmak üzere  $Ax = b$ 'nin çözümü için elde edilen bu algoritma aşağıdaki şekildedir.

#### 4. EŞLENİK GRADYAN ALGORİTMASI

---

**1. Adım:**  $r_0 = b - Ax_0$  olsun. Eğer  $r_0$  yeterince küçükse algoritma çıktısı olarak  $x_0$  alınsın.

**2. Adım:** Eğer  $r_0$  yeterince küçük değilse  $d_0 = r_0$ ,  $k = 0$  olmak üzere

$$\alpha_k = \frac{r_k^T r_k}{d_k^T A d_k},$$

$$x_{k+1} = x_k + \alpha_k d_k,$$

$$r_{k+1} = r_k - \alpha_k A d_k,$$

olsun.  $r_{k+1}$  yeterince küçükse döngüden çıkılsın, değilse

$$\beta_k = \frac{r_{k+1}^T r_{k+1}}{r_k^T r_k}$$

$$d_{k+1} = r_{k+1} + \beta_k d_k$$

değerleri hesaplınsın.

**3. Adım:**  $k = k + 1$  olarak Adım 2'ye dön.

Eşlenik gradyan algoritmasının düzgün, lineer olmayan ve sürekli türevlenebilir bir  $f: \mathbb{R}^n \rightarrow \mathbb{R}$  fonksiyonunun

$$\min\{f(x): x \in \mathbb{R}^n\} \tag{4.8}$$

şeklindeki kısıtsız optimizasyon problemini çözmek için tanımlanacak olması durumunda  $x_0$  rastgele seçilen bir başlangıç noktası olmak üzere algoritmanın genel formu

$$x_{k+1} = x_k + \alpha_k d_k \quad (4.9)$$

şeklinindedir. Burada  $\alpha_k$ , genellikle çizgi arama koşulları ile elde edilebilen pozitif adım uzunluğu ve  $d_k$  ise genellikle

$$d_k = \begin{cases} -g_k, & k = 0 \text{ ise} \\ -g_k + \beta_k d_{k-1}, & k \geq 1 \text{ ise} \end{cases} \quad (4.10)$$

şeklinde tanımlanan arama yönüdür. Burada  $g_k = \nabla f(x_k)$  ve  $\beta_k$  ise CG parametresidir ve  $\beta_k$ 'nin farklı seçilmesiyle farklı CG algoritmaları elde edilir.  $d_k$  arama yönünün genellikle

$$g_k^T d_k \leq -c \|g_k\|^2, c > 0 \quad (4.11)$$

şeklinde verilen iniş özelliğini sağlaması beklenir.

$y_{k-1} = g_k - g_{k-1}$  ve  $\|\cdot\|$  ise Öklid normunu belirtmek üzere PRP, HS, LS, FR, DY ve CD tarafından tanımlanan en temel eşlenik gradyan algoritmalarının eşlenik parametreleri

$$\beta_k^{PRP} = \frac{g_k^T y_{k-1}}{\|g_{k-1}\|^2}$$

$$\beta_k^{FR} = \frac{\|g_k\|^2}{\|g_{k-1}\|^2}$$

$$\beta_k^{HS} = \frac{g_k^T y_{k-1}}{d_{k-1}^T y_{k-1}}$$

$$\beta_k^{DY} = \frac{\|g_k\|^2}{d_{k-1}^T y_{k-1}}$$

$$\beta_k^{LS} = \frac{g_k^T y_{k-1}}{-g_k^T d_{k-1}}$$

$$\beta_k^{CD} = \frac{\|g_k\|^2}{-g_k^T d_{k-1}}$$

şeklinde tanımlanmaktadır (Hestenes and Stiefel 1952; Fletcher and Reeves 1964; Polak and Ribiere 1969; Liu and Storey 1991; Dai and Yuan 1999; Fletcher 2013).

Eşlenik gradyan algoritmaları, büyük ölçekli lineer olmayan optimizasyon problemlerini çözmek için kullanılan iteratif algoritmaların popüler bir ailesidir. Bunun nedeni algoritmaların adımları hesaplanırken ikinci dereceden türevlerin hesaplanmasına gerek olmaması, düşük depolama ve hesaplama, uygun yakınsama oranı gibi uygun özelliklerinin var olmasıdır. Eşlenik gradyan algoritmasının kullanıldığı alanlar; mühendislik optimizasyon problemleri, sinir ağı eğitimi ve görüntü restorasyonu gibi alanlardır. Bu algoritma ilk olarak 1952 yılında Hestenes ve Stiefel (Hestenes and Stiefel 1952) tarafından büyük ölçekli lineer denklem sistemlerini çözmek için doğrudan bir algoritma olarak tanıtılmıştır. Daha sonra, bu algoritma 1964 yılında Fletcher-Reeves (Fletcher and Reeves 1964) tarafından lineer olmayan problemler için geliştirilmiştir.

Yıllar boyunca tanımlanan bazı CG algoritmaları uygun sayısal sonuçlara sahipken, genel fonksiyonlar için global yakınsamaları ise garanti değildir. Mesela CG algoritmaları arasında en etkili olarak kabul edilen PRP ve HS algoritmalarının yakınsama özellikleri o kadar da iyi değildir. Örneğin Polak ve Ribiere (Polak and Ribiere 1969)  $f$  fonksiyonunun konveks olması durumunda kesin çizgi arama tekniğiyle PRP algoritmasının global yakınsak olduğunu göstermesine karşın fonksiyonun konveks olmaması durumunda algoritmanın global yakınsaklığını garanti edememiştir. Gilbert and Nocedal (1992) yaptığı çalışmada ise yeterli iniş özelliği altında  $\beta_k^{PRP}$  parametresini  $\beta_k^{PRP+} = \max\{\beta_k^{PRP}, 0\}$  şeklinde modifiye ederek ve  $\alpha_k$  yı Wolfe çizgi arama tekniği ile hesaplayarak PRP+ algoritmasının global yakınsak olduğunu ispatlamışlardır. Daha sonra yazarlar aynı geliştirme mantığını HS algoritmasına uygulamışlardır. HS algoritmasının uygulamaları ve yakınsama özellikleri hakkında pek çok araştırma yapılmıştır. Bu çalışmaların yanı sıra birçok araştırmacı farklı hedefler için çeşitli geliştirmeler yapmışlardır. Eşlenik gradyan algoritmaları hakkında daha detaylı bilgi edinmek için Hager and Zhang (2006) çalışmasına bakılabilir.

## 5. ARAŞTIRMA BULGULARI

Bu bölümde yeni bir eşlenik gradyan algoritması tanıtılacak ve algoritmanın Armijo ve Wolfe çizgi arama tekniği kullanarak (4.8) probleminin çözümüne global yakınsak olduğu ve yeterli iniş özelliğini sağladığı gösterilecektir.

### 5.1. DEI Eşlenik Gradyan Algoritması

Bu başlık altında tanımlanacak olan DEI eşlenik gradyan algoritmasının eşlenik gradyan parametresi

$$\beta_k^{DEI} = \frac{\|g_k\|^2 - \frac{\|g_k\|}{\|d_{k-1}\|} |g_k^T d_{k-1}|}{\|g_{k-1}\|^2 + \mu \|g_k\| \|d_{k-1}\|} \quad (5.1)$$

şeklinde olup algoritmanın adımları aşağıda verilmiştir.

- 1. Adım:**  $0 < \delta < \sigma < 1$  ve  $\mu > 1$  olmak üzere  $\varepsilon, \delta, \rho, \mu$  parametrelerini ve  $x_0 \in \mathbb{R}^n$  başlangıç noktasını seçin.  $d_0 = -g_0$  ve  $k := 0$  olarak ayarlayın.
- 2. Adım:**  $\|g(x_k)\| \leq \varepsilon$  ise dur. Aksi takdirde 3. Adım'a gidin.
- 3. Adım:**  $d_k$  arama yönünü  $d_k = -g_k + \beta_k^{DEI} d_{k-1}$  şeklinde hesaplayın.
- 4. Adım:** Adım uzunluğu olan  $\alpha_k$ 'yi Wolfe çizgi arama tekniği veya Armijo çizgi arama tekniği ile bulun.
- 5. Adım:** (4.9)'a göre  $x_{k+1}$ 'i hesapla ve  $k := k + 1$  alarak 2. Adım'a gidin.

Şimdi DEI Algoritması tarafından oluşturulan  $d_k$  arama yönünün yeterli iniş özelliğini sağladığı kanıtlanacaktır.

**Lemma 5.1:**  $g_k = g(x_k)$  olmak üzere  $\{x_k\}, \{d_k\}$  ve  $\{g_k\}$  DEI Algoritması tarafından üretilen diziler ve  $c_1 = 1 - \frac{1}{\mu}$ ,  $\mu > 1$ , olmak üzere

$$g_k^T d_k \leq -c_1 \|g_k\|^2 \quad (5.2)$$

dır.

**İspat:** Eğer  $k = 0$  ise (4.10)'dan  $g_0^T d_0 = -g_0^T g_0 \leq -\left(1 - \frac{1}{\mu}\right) \|g_0\|^2$  elde edilir. Şimdi  $k > 0$  için (5.2) eşitsizliğinin sağlandığını gösterelim. (5.1)'deki  $\beta_k^{DEI}$ 'i (4.10)'da yerine yazılırsa

$$g_k^T d_k = -\|g_k\|^2 + \beta_k^{DEI} g_k^T d_{k-1} \leq -\|g_k\|^2 + |\beta_k^{DEI}| |g_k^T d_{k-1}| \quad (5.3)$$

elde edilir. Eğer  $g_k^T d_{k-1} = 0$  ise

$$g_k^T d_k = -\|g_k\|^2 \leq -\left(1 - \frac{1}{\mu}\right) \|g_k\|^2$$

olduğu görülür. Eğer  $g_k^T d_{k-1} \neq 0$  ise (5.1)'de Cauchy-Schwarz eşitsizliğinden,

$$|\beta_k^{DEI}| = \frac{\left| \|g_k\|^2 - \frac{\|g_k\|}{\|d_{k-1}\|} |g_k^T d_{k-1}| \right|}{\| \|g_{k-1}\|^2 + \mu \|g_k\| \|d_{k-1}\| \|} \leq \frac{\|g_k\|^2}{\mu \|g_k\| \|d_{k-1}\|} = \frac{1}{\mu} \frac{\|g_k\|}{\|d_{k-1}\|} \quad (5.4)$$

olur. Buradan (5.4)'ü (5.3)'ün sağ tarafına yazıp ve Cauchy-Schwarz eşitsizliği kullanılırsa,

$$\begin{aligned} g_k^T d_k &\leq -\|g_k\|^2 + |\beta_k^{DEI}| |g_k^T d_{k-1}| \leq -\|g_k\|^2 + \frac{\|g_k\|^2}{\mu \|g_k\| \|d_{k-1}\|} |g_k^T d_{k-1}| \\ &\leq -\left(1 - \frac{1}{\mu}\right) \|g_k\|^2 \end{aligned} \quad (5.5)$$

bulunur. Bu ise ispatı tamamlar.

Şimdi ise Wolfe ve Armijo çizgi arama teknikleri yardımıyla aşağıdaki kabuller altında yeni önerilen algoritmanın global yakınsak olduğu gösterilecektir.

**Kabuller:**

- (i)  $A = \{x \in \mathbb{R}^n | f(x) \leq f(x_0)\}$  seviye kümesi sınırlı olsun. Yani her  $x \in A$  için  $\|x\| \leq M_1$  olacak şekilde bir  $M_1$  sabiti var olsun.
- (ii)  $f$  sürekli türevlenebilir bir fonksiyon ve  $A$ 'nın bir  $U$  komşuluğunda  $f$  alttan sınırlı olsun.
- (iii)  $g = \nabla f$  Lipschitz sürekli bir fonksiyon yani her  $x, y \in U$  için

$$\|g(x) - g(y)\| \leq L\|x - y\|$$

olacak şekilde bir  $L > 0$  var olsun.

Yukarıda verilen kabullerden yola çıkarak, her  $x \in U$  için  $\|g(x)\| \leq M_2$  olacak şekilde bir  $M_2 > 0$  sabitinin var olduğu söylenebilir.

Şimdi yakınsaklık teoremini ispatlamak için gerekli olan aşağıdaki lemmayı verelim.

**Lemma 5.2:**  $\{x_k\}, \{d_k\}$  ve  $\{g_k\}$  DEI Algoritması tarafından üretilen diziler olsun. (i)-(iii) kabulleri altında

$$\|d_k\| \leq \left(1 + \frac{1}{\mu}\right) \|g_k\| \tag{5.6}$$

eşitsiziği geçerlidir. Üstelik

$$\|d_k\| \leq c_2$$

olacak şekilde  $c_2 > 0$  vardır.

**İspat:** (5.1) ve (4.10)'dan

$$\|d_k\| \leq \|g_k\| + |\beta_k^{DEI}| \|d_{k-1}\|$$

eşitsizliği elde edilir. (5.4) eşitsizliğinden ise

$$\|d_k\| \leq \|g_k\| + \frac{1}{\mu} \frac{\|g_k\|}{\|d_{k-1}\|} \|d_{k-1}\| = \left(1 + \frac{1}{\mu}\right) \|g_k\|$$

olduğu görülür.  $\|g_k\| \leq M_2$  olacak şekilde  $M_2$  var olduğundan, her  $x \in U$  için

$$\|d_k\| \leq \left(1 + \frac{1}{\mu}\right) M_2$$

eşitsizliği sağlanır.  $c_2 = \left(1 + \frac{1}{\mu}\right) M_2$  alınırsa  $\|d_k\| \leq c_2$  eşitsizliği elde edilir.

Wolfe çizgi arama tekniği altında DEI Algoritmasının global yakınsama özelliğini oluşturmamız için aşağıdaki lemma gereklidir.

**Lemma 5.3:** (Zoutendijk Koşulu)  $x_0$  herhangi bir başlangıç noktası,  $\alpha_k$  Wolfe çizgi arama tekniğiyle üretilen adım uzunluğu,  $d_k$  arama yönü olmak üzere (4.9) ve (4.10) formundaki herhangi bir algoritma için

$$\sum_{k=1}^{+\infty} \frac{(g_k^T d_k)^2}{\|d_k\|^2} < \infty \quad (5.7)$$

eşitsizliği sağlanır (Zoutendijk,1970).

(5.7)'den ve Lemma 5.1'den aşağıdaki sonucu elde edebiliriz.

**Sonuç 5.4:** Adım uzunluğu olan  $\alpha_k$ 'nın Wolfe çizgi arama tekniğiyle belirlendiğini ve  $\{x_k\}, \{d_k\}$  ve  $\{g_k\}$  dizilerinin DEI Algoritması tarafından üretildiğini kabul edelim. (i)-(iii) kabulleri altında,

$$\sum_{k=1}^{+\infty} \frac{\|g_k\|^4}{\|d_k\|^2} < \infty \quad (5.8)$$

dır.

Bu lemmaların ardından yeni tanımlanan DEI Algoritmasının yakınsaklık teoremi verilecektir.

**Teorem 5.5:** (i)-(iii) kabulleri sağlansın.  $\alpha_k$ , Wolfe çizgi arama tekniğiyle elde edilmek üzere DEI Algoritması için

$$\liminf_{k \rightarrow \infty} \|g_k\| = 0 \quad (5.9)$$

dır.

**İspat:** Aksini kabul edelim. Bu durumda  $\forall k \geq 0$  için

$$\|g_k\| > \varepsilon \quad (5.10)$$

olacak şekilde bir  $\varepsilon > 0$  vardır. (5.6)'nın her iki tarafının da karesi alınırsa

$$\|d_k\|^2 = \|g_k\|^2 - 2\beta_k g_k^T d_{k-1} + \beta_k^2 \|d_{k-1}\|^2 \quad (5.11)$$

bulunur.  $\beta_k^{DEI}$ 'nin tanımından ve (5.4)'den

$$-2\beta_k^{DEI} g_k^T d_{k-1} \leq 2|\beta_k^{DEI}| \|g_k^T d_{k-1}\| \leq \frac{2\|g_k\|^2}{\mu} \quad (5.12)$$

olduğu görülür. (5.1) ve (5.12), (5.11)'de yerine yazılırsa,

$$\|d_k\|^2 \leq \|g_k\|^2 + \frac{2\|g_k\|^2}{\mu} + \frac{\|g_k\|^4}{\|g_{k-1}\|^4} \|d_{k-1}\|^2 = \left(1 + \frac{2}{\mu}\right) \frac{1}{\|g_k\|^2} + \frac{\|d_{k-1}\|^2}{\|g_{k-1}\|^4} \quad (5.13)$$

elde edilir. Daha sonra (5.13)'de bulunan ifadenin her iki tarafı da  $\|g_k\|^4$  e bölünürse

$$\frac{\|d_k\|^2}{\|g_k\|^4} \leq \left(1 + \frac{2}{\mu}\right) \frac{1}{\|g_k\|^2} + \frac{\|d_{k-1}\|^2}{\|g_{k-1}\|^4} \quad (5.14)$$

bulunur. (5.14)'ü iteratif olarak kullanıp,  $\|d_1\|^2 = -g_1^T d_k = \|g_1\|^2$  olduğuna dikkat edilirse,

$$\frac{\|d_k\|^2}{\|g_k\|^4} \leq \sum_{i=1}^k \frac{1}{\|g_i\|^2} \left[1 + \frac{2\|g_{i-1}\|}{\mu\|g_{i-2}\|}\right] \leq \frac{1}{\varepsilon^2} \sum_{i=1}^k 1 + \frac{2\varepsilon}{\mu s_2} \quad (5.15)$$

elde edilir. O halde (5.15) ve (5.10)'dan

$$\frac{\|g_k\|^4}{\|d_k\|^2} \geq \frac{\varepsilon^2}{k} \left(\frac{\mu s_2}{\mu s_2 + 2\varepsilon}\right) \quad (5.16)$$

bulunur ve bu bilgilerden yola çıkarak

$$\sum_{k=1}^{\infty} \frac{\|g_k\|^4}{\|d_k\|^2} = \infty \quad (5.17)$$

elde edilir. Bu ise (5.7)'deki Zoutendijk koşuluyla çelişir. Buradan yola çıkarak (5.9)'ün doğru olduğu söylenir. Dolayısıyla kabulümüzün geçerli olmadığını yani

$$\liminf_{k \rightarrow \infty} \|g_k\| = 0$$

olduğu söylenir.

Aşağıda Armijo çizgi arama tekniği ile birlikte DEI Algoritmasının global yakınsamasını ispatlamak için gerekli olan lemmalar verilmiştir.

**Lemma 5.6:**  $\{x_k\}, \{d_k\}$  ve  $\{g_k\}$  dizilerinin DEI Algoritması tarafından üretildiğini ve  $\alpha_k$  adım uzunluğunun Armijo çizgi arama tekniğiyle elde edildiğini kabul edelim. Bu durumda her  $k$  için

$$\alpha_k \geq c_1 \frac{(-g_k^T d_k)}{\|d_k\|^2} \quad (5.18)$$

olacak şekilde bir  $c_1 > 0$  sabiti vardır.

**İspat:** Armijo çizgi arama tekniği, (5.5) ve (i) kabulü altında,

$$\sum_{i=1}^{\infty} -\delta \alpha_k g_k^T d_k < +\infty \quad (5.19)$$

elde edilir. Bu durumda aşağıda verilecek olan iki durumu göz önünde bulundurarak (5.18) ispatlanır.

**1.Durum:**  $\alpha_k = 1$  olsun. (5.5)'den  $g_k^T d_k \leq -\left(1 - \frac{1}{\mu}\right) \|g_k\|^2$  olduğu bilinmektedir. Dolayısıyla (5.18)'de verilen eşitsizlik  $c_1 = \frac{\mu}{\mu-1}$ , ( $\mu > 1$ ) ile sağlanır.

**2.Durum:**  $\alpha_k < 1$  olsun. Armijo çizgi arama tekniğinden  $\rho^{-1}\alpha_k$ , Armijo çizgi arama tekniğindeki eşitsizliği sağlamaz. Bu ise

$$f(x_k + \rho^{-1}\alpha_k d_k) - f(x_k) > c_1 \alpha_k \rho^{-1} g_k^T d_k \quad (5.20)$$

anlamına gelir.  $x_k + t_k \rho^{-1}\alpha_k d_k \in \mathbb{N}$  olmak üzere ortalama değer teoreminden

$$\begin{aligned} f(x_k + \rho^{-1}\alpha_k d_k) - f(x_k) &= \alpha_k \rho^{-1} g(x_k + t_k \rho^{-1}\alpha_k d_k)^T d_k \\ &= \alpha_k \rho^{-1} g_k^T d_k + \alpha_k \rho^{-1} (g(x_k + t_k \rho^{-1}\alpha_k d_k) - g_k)^T d_k \\ &\leq \alpha_k \rho^{-1} g_k^T d_k + L \alpha_k^2 \rho^{-1} \|d_k\|^2 \end{aligned} \quad (5.21)$$

olacak şekilde bir  $t_k \in (0,1)$  vardır. Yukarıda verilen son eşitsizliği (5.19) ile değiştirerek

$$\alpha_k \geq \frac{(1 - \delta)\rho(-g_k^T d_k)}{L\|d_k\|^2}$$

elde edilir. Bu durumda  $c_1 = \min\left\{\frac{\mu}{\mu-1}, \frac{(1-\delta)\rho}{L}\right\}$  alınması durumunda (5.18) elde edilir.

(5.18) ve (5.19) eşitsizliklerinden aşağıdaki Zoutendijk koşulları kolayca elde edilir.

**Lemma 5.7:**  $x_0$ , (i)-(iii) kabulleri altında bir başlangıç noktası olsun.  $d_k$  iniş yönü ve  $\alpha_k$  Armijo koşulunu sağlayan adım uzunluğu olmak üzere (4.9) ve (4.10) formundaki herhangi bir algoritma için

$$\sum_{k=1}^{\infty} \frac{(d_k^T g_k)^2}{\|d_k\|^2} < \infty \quad (5.22)$$

dır.

(5.5), Lemma 5.3 ve (5.22)'den

$$\sum_{k=1}^{\infty} \frac{\|g_k\|^4}{\|d_k\|^2} < \infty \quad (5.23)$$

eşitsizliği elde edilir.

Lemma 5.7 ve Teorem 5.5'deki aynı argümanlar kullanılarak, Armijo çizgi arama tekniğiyle birlikte DEI Algoritmasının global yakınsaması elde edilebilir.

**Teorem 5.8:** (i)-(iii) kabullerinin geçerli olduğunu kabul edelim.  $\alpha_k$ , Armijo çizgi arama tekniğiyle elde edilmek üzere DEI Algoritması için

$$\liminf_{k \rightarrow \infty} \|g_k\| = 0 \quad (5.24)$$

dır.

Dolayısıyla tanımladığımız algoritma tarafından üretilen dizi (4.8)'de verilen minimizasyon problemine kuvvetli yakınsaktır.

## 5.2. Sayısal Sonuçlar

Bu bölümde, DEI Algoritmasının hesaplama performansını göstermek için kullanılan sayısal deneyler iki kısma ayrılmıştır. İlk kısımda büyük ölçekli kısıtsız problemler için bir simülasyon testi, ikinci kısımda ise görüntü restorasyon problemlerine yönelik bir uygulama yer almaktadır. Tüm kodlar MATLAB R2024b ile yazılmıştır ve Windows 10 işletim sistemi ile donatılmış Intel Core i5-1235U 1.30 GHz, 8.00 GB belleğe sahip bir bilgisayarda çalıştırılmıştır.

### 5.2.1. Büyük Ölçekli Kısıtsız Problemler İçin Simülasyon Testi

Bu bölümde önerilen DEI Algoritmasının performansı PRP, PRP-T, WYL, RMIL, DHS, NPRP2 ve IWYL algoritmaları ile karşılaştırılmıştır. Adım uzunluğunu belirlemek için (4.5) Armijo Koşulu kullanılmış ve  $\alpha = 1$ ,  $\rho = 0.1$  ve  $c = 10^{-3}$  olarak alınmıştır.

Aşağıdaki tabloda algoritmaları karşılaştırmak için kullanılan test fonksiyonlarının adları verilmiştir.

**Çizelge 5.1:** Algoritmaları karşılaştırmak için kullanılan test fonksiyonları

Num.	Fonksiyon Adı	Num.	Fonksiyon Adı
1	Extended White & Holst Fonksiyon	10	Generalized Rosenbrock Fonksiyon
2	Extended Beale Fonksiyon	11	Extended Penalty Fonksiyon
3	Extended Himmelblau Fonksiyon	12	Perturbed Quadratic Fonksiyon
4	Generalized Tridiagonal 1 Fonksiyon	13	Diagonal 1 Fonksiyon
5	Diagonal 4 Fonksiyon	14	Diagonal 2 Fonksiyon
6	Extended Powell Fonksiyon	15	Diagonal 5 Fonksiyon
7	Extended Hiebert Fonksiyon	16	Power Fonksiyon (CUTE)
8	Extended Tridiagonal 2 Fonksiyon	17	EG2 Fonksiyon (CUTE)
9	Fletcher Fonksiyon (CUTE)		

Sekiz algoritmanın büyük ölçekli kısıtsız optimizasyon problemini çözme performansını test etmek için her bir test probleminin boyutları  $n = 5000$  olarak seçilmiştir. Durdurma kriteri olarak  $\|g(x_k)\| < 10^{-3}$  veya  $Itr > 20000$  koşullarından

## 5. ARAŞTIRMA BULGULARI

biri kullanılmıştır. Burada  $Itr$ , algoritmaların iterasyon sayısını temsil etmektedir. Eğer iterasyon sayısı 20000'i geçerse, algoritmanın test problemini çözemediği kabul edilmiştir.

Verilen algoritmaların sayısal performanslarını göstermek için Dolan and More (2002) tarafından tanımlanan performans profil eğrisi kullanılmıştır.  $P$  problemler kümesi,  $S$  çözücüler kümesi,  $n_p$  problem sayısı,  $t_{p,s}$   $Tcpu$  zamanı ve  $n_s$  çözücü sayısı olsun. Her  $s \in S$  ve her  $p \in P$  için  $t_p^* = \min\{t_{p,s} : s \in S\}$ ,  $\gamma_{p,s} = \frac{t_{p,s}}{t_p^*}$  ve  $m\{p \in P | \gamma_{p,s} \leq t\}$  ise  $\{p \in P | \gamma_{p,s} \leq t\}$  kümesinin elemanlarının sayısı olmak üzere performans oranı,

$$P(t) = \frac{1}{n_p} m\{p \in P | \gamma_{p,s} \leq t\}, \forall t \in \mathbb{R}^+$$

şeklinde tanımlanır. Bu değerlere göre elde edilen grafikte, en üstteki eğriye sahip olan algoritma, alttaki eğriye sahip diğer algoritmalarından daha iyi performansa sahiptir yani en verimli algoritma, eğrisi tüm eğrilerin sol üst köşesinde olan algoritmadır.

Karşılaştırılacak olan algoritmaların ele alınan test fonksiyonlarını çözme süreleri aşağıdaki Çizelge 5.2'de verilmiştir.

**Çizelge 5.2:** Algoritmaların test fonksiyonlarını çözme süreleri

	DEI	PRP	PRP-T	WYL	RMIL	DHS	NPRP-2	IWYL
1	17,33	52,12	5,51	51,98	56,43	2,85	47,14	12,27
2	18,68	11,09	4,98	14,73	11,68	18,86	16,51	3,52
3	1,039	1,08	1,09	1,11	1,14	1,07	1,08	1,06
4	153,29	155,81	155,39	142,54	153,22	153,53	87,54	153,41
5	4,00	4,26	4,40	4,17	4,30	4,03	4,16	4,09
6	80,00	135,58	137,90	136,07	118,66	101,82	123,28	129,25
7	68,05	68,28	67,99	88,36	73,47	117,23	112,75	116,80
8	73,99	74,50	71,73	72,03	71,66	69,63	91,17	92,44
9	56,63	75,69	80,44	85,80	85,74	83,34	63,44	84,70
10	54,25	57,54	147,42	61,86	98,48	76,21	132,28	77,26
11	1,19	1,20	1,31	1,30	1,32	1,236	1,22	1,27
12	52,35	52,51	54,52	57,98	62,79	67,37	72,73	70,52
13	102,53	96,50	158,05	140,91	149,35	167,71	139,91	159,11
14	16,45	16,53	16,70	25,66	25,60	27,59	25,17	25,29
15	0,01	0,01	0,04	0,017	22,13	0,01	0,01	0,01
16	36,30	37,97	37,77	36,01	37,70	38,73	42,18	41,03
17	129,38	NaN	NaN	NaN	NaN	NaN	NaN	NaN

## 5. ARAŞTIRMA BULGULARI

**Çizelge 5.3:** Algoritmaların test fonksiyonlarını çözme sürelerinin o fonksiyonu en hızlı çözen algoritmanın çözme süresine oranları

	DEI	PRP	PRP-T	WYL	RMIL	DHS	NPRP-2	IWYL
1	6.0636	18.2379	1.9296	18.1889	19.7463	1	16.4972	4.2946
2	5.2992	3.1469	1.4137	4.1783	3.3131	5.3508	4.6840	1
3	1	1.0471	1.0558	1.0692	1.1010	1.0365	1.0394	1.0250
4	1.7509	1.7797	1.7749	1.6281	1.7501	1.7537	1	1.7523
5	1	1.0641	1.0993	1.0416	1.0736	1.0077	1.0389	1.0209
6	1	1.6946	1.7236	1.7007	1.4831	1.2727	1.5408	1.6154
7	1.0007	1.0042	1	1.2995	1.0805	1.7240	1.6582	1.7176
8	1.0625	1.0699	1.03	1.0344	1.0290	1	1.3092	1.3274
9	1	1.3364	1.4204	1.5149	1.5139	1.4716	1.1201	1.4956
10	1	1.0605	2.7171	1.1402	1.8151	1.4046	2.4381	1.424
11	1	1.01	1.1040	1.0964	1.1107	1.0369	1.0285	1.0679
12	1	1.0029	1.0413	1.1075	1.1992	1.2869	1.3891	1.3469
13	1.0624	1	1.6377	1.4601	1.5476	1.7378	1.4498	1.6487
14	1	1.0048	1.0154	1.5594	1.5559	1.6771	1.5297	1.5372
15	1	1.066	2.8	1.1333	1475.333	1.0666	1.0666	1
16	1.0079	1.0544	1.0488	1	1.0468	1.0755	1.1712	1.1392
17	1	NaN	NaN	NaN	NaN	NaN	NaN	NaN

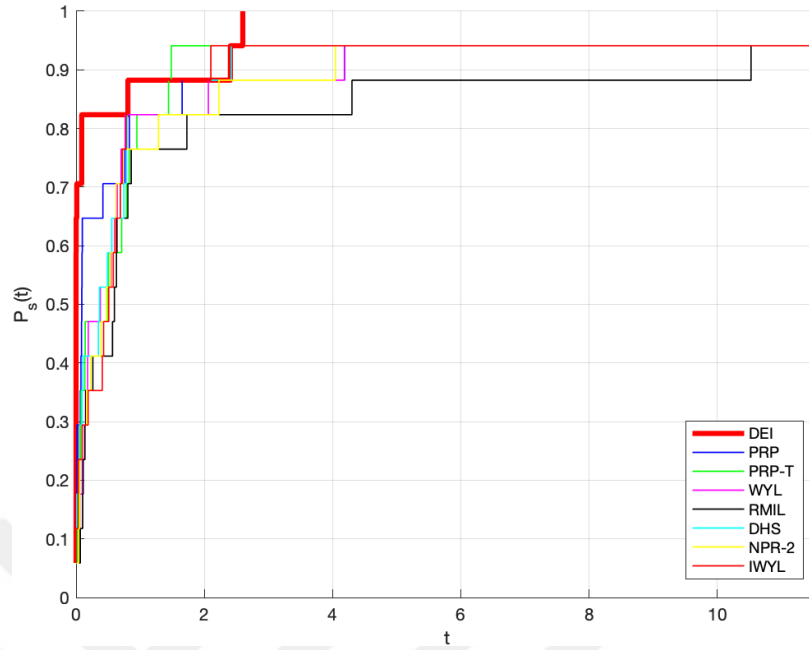
Bu tablodaki verilere performans profili elde etmek için gerekli işlemler uygulanırsa aşağıdaki Çizelge 5.4 elde edilir.

**Çizelge 5.4:** Çizelge 5.3'deki sonuçların 2 tabanındaki logaritma değerleri

	DEI	PRP	PRP-T	WYL	RMIL	DHS	NPRP-2	IWYL
1	2,6001	4,1888	0,9483	4,1849	4,3035	0	4,0441	2,1025
2	2,4057	1,6539	0,4994	2,0629	1,7281	2,4197	2,2277	0
3	0	0,0663	0,0783	0,0965	0,1388	0,0517	0,0557	0,0356
4	0,8080	0,8316	0,8277	0,7031	0,8074	0,8104	0	0,8092
5	0	0,0896	0,1365	0,0588	0,1024	0,0110	0,0550	0,0298
6	0	0,7609	0,7854	0,7661	0,5686	0,3478	0,6236	0,6918
7	0,0010	0,0060	0	0,3779	0,1116	0,7857	0,7236	0,7803
8	0,0874	0,0974	0,0426	0,0487	0,0412	0	0,3886	0,4086
9	0	0,4183	0,5062	0,5992	0,5982	0,5573	0,1636	0,5807
10	0	0,0847	1,4420	0,1892	0,8600	0,4901	1,2857	0,5099
11	0	0,0143	0,1427	0,1327	0,1514	0,0522	0,0405	0,0947
12	0	0,0041	0,0583	0,1473	0,2620	0,3638	0,4741	0,4296
13	0,0873	0	0,7116	0,5460	0,6300	0,7972	0,5358	0,7213
14	0	0,0069	0,0220	0,6409	0,6377	0,7459	0,6132	0,6203
15	0	0,0922	1,4854	0,1805	10,526	0,0930	0,0930	0
16	0,0113	0,0764	0,687	0	0,0659	0,1050	0,2279	0,1880
17	0	8,3776	2,9708	8,3698	21,052	4,8394	8,0882	4,2050

Elde edilen bu tablodaki değerler kullanılarak algoritmaların CPU sürelerine göre olan performans profilleri aşağıdaki Çizelge 5.5'de görülmektedir.

Çizelge 5.5: Tcpu performans profilleri



Çizelge 5.5’den görüldüğü üzere verilen test problemlerini çözme performansı en iyi olan algoritma DEI Algoritmasıdır.

### 5.2.2. Görüntü Restorasyon Problemlerine Uygulama

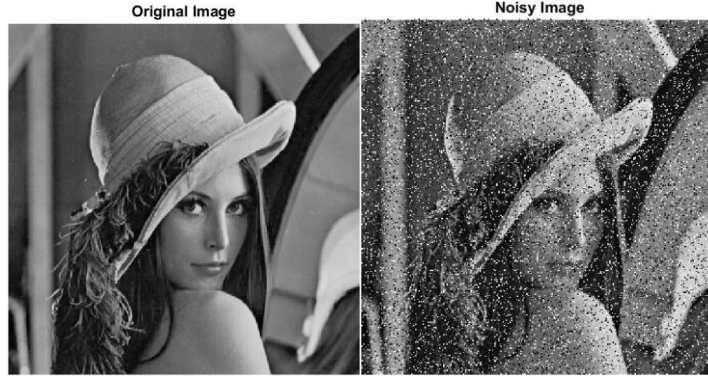
Görüntü; gerçekteki 3 boyutlu herhangi bir nesnenin 2 boyutlu bir yüzey üzerine düşürülmüş haritası olarak tanımlanabilir. Dijital görüntüler bilgisayar ortamında matrislerle temsil edilirler. Matrislerin her bir girişine piksel denir. Bu da görüntülerin piksellerden oluştuğunu ifade eder.

Görüntü renkli görüntü, gri seviyeli görüntü ve siyah beyaz görüntü olmak üzere üçe ayrılır. Bu tezde gri seviyeli görüntüler kullanılmıştır. Gri seviyeli görüntülerde siyah, beyaz ve grinin tonları mevcuttur. 0 pikseli siyah rengi, 255 pikseli beyaz rengi ve 0 ile 255 arasındaki pikseller ise grinin tonlarını temsil etmektedir.

Gürültü, görüntüyü elde etme aşamasında ya da görüntüyü bir kaynaktan başka bir kaynağa aktarırken oluşan piksel bozulmalarıdır. Gauss, Tuz ve Biber (Salt ve Pepper) ve Foton bazı gürültü çeşitleridir. Tuz ve Biber gürültüsü genel olarak veri iletimindeki hatalardan kaynaklanır ve rastgele gürültü olarak da adlandırılır. Bu gürültü

## 5. ARAŞTIRMA BULGULARI

çeşidine sahip bir görüntüde gürültülü pikseller ya 0 (biber) ya da 255 (tuz) değerini alır. Aşağıda Tuz ve Biber gürültüsüne sahip örnek bir görüntü verilmiştir.



Şekil 5.1: %15 gürültü ile bozulmuş gürültülü Lena görüntüsü

Görüntü işlemede görüntüleri bulanıklaştırmak ya da kenarları belirginleştirmek için ortalama (mean), orta değer ya da medyan (median) ve Gauss gibi dijital filtrelemeler kullanılmaktadır. Bunlardan medyan filtreleme işleminde her piksel, bir dizi komşu piksel değerinin ortanca değeri ile değiştirilir. Gürültülü görüntüdeki bozulan piksellerin dört bir yanındaki piksel değerleri sıralanır ve sıralı listede orta değer (liste çift sayıda öge içeriyorsa ortadaki ikisinin ortalaması) çıktısı piksel değerini oluşturur.

Görüntü restorasyon ya da onarım problemi gürültüleri mümkün olduğunca minimize ederek görüntünün niteliğini istenilen seviyeye yükseltme problemidir.

$x$ ;  $M \times N$  pikselden oluşan orijinal görüntü,  $K = \{1, 2, \dots, M\} \times \{1, 2, \dots, N\}$ , olmak üzere  $x_{i,j}$ ,  $x$  in  $(i, j)$  deki gri seviyesi,  $W_{i,j} = \{(i-1, j), (i+1, j), (i, j-1), (i, j+1)\}$  kümesi  $(i, j)$ 'nin komşuluğu,  $[s_{min}, s_{max}]$  kümesi  $x$  deki tüm piksellerinin kümesi yani  $x_{i,j} \in [s_{min}, s_{max}]$ ,  $y$ ;  $x$  in Tuz-Biber gürültüsü ile bozulmuş görüntüsü,  $y_{i,j}$ ;  $y$  in  $(i, j)$  deki gri seviyesi, yani

$$y_{i,j} = \begin{cases} x_{i,j}, & 1-p-q \text{ olasılıkla,} \\ s_{min}, & p \text{ olasılıkla,} \\ s_{max}, & q \text{ olasılıkla,} \end{cases}$$

ve  $u^*$  ise elde edilecek onarılmış görüntü olsun. Gürültülü  $y$  görüntüsüne medyan filtreleme uygulama sonucunda elde edilen görüntü  $\tilde{y}$  olsun. Gürültülü pikseller ya  $s_{min}$  ya da  $s_{max}$  olacağından gürültü adayları kümesi  $\bar{N} = \{(i, j) \in K | \tilde{y}_{i,j} \neq y_{i,j} \text{ ve } y_{i,j} = s_{max} \text{ veya } y_{i,j} = s_{min}\}$  olsun. Bu tanıma göre  $\bar{N}$  kümesi gürültülü olup medyan filtreleme sonucunda değişen piksellerin kümesidir. Eğer  $(i, j) \notin \bar{N}$  ise  $(i, j)$  pikseli bozulmamış olduğundan herhangi bir değişiklik yapılmayacak olup  $u_{i,j}^* = y_{i,j} = x_{i,j}$  olarak alınacaktır. Eğer  $(i, j) \in \bar{N}$  ise  $y_{i,j}$  restore edilecektir. Ayrıca  $(m, n) \in W_{i,j} \setminus \bar{N}$  ise  $u_{m,n}^* = y_{m,n}$  ve  $(m, n) \in W_{i,j} \cap \bar{N}$  ise  $y_{m,n}$  restore edilecektir. Buna göre minimize edilecek fonksiyon

$$\psi(u) := \sum_{(i,j) \in \bar{N}} \left\{ \frac{\beta}{2} \sum_{(m,n) \in W_{i,j} \setminus \bar{N}} 2\varphi_{\alpha}(u_{i,j} - y_{m,n}) + \sum_{(m,n) \in W_{i,j} \cap \bar{N}} \varphi_{\alpha}(u_{i,j} - u_{m,n}) \right\}$$

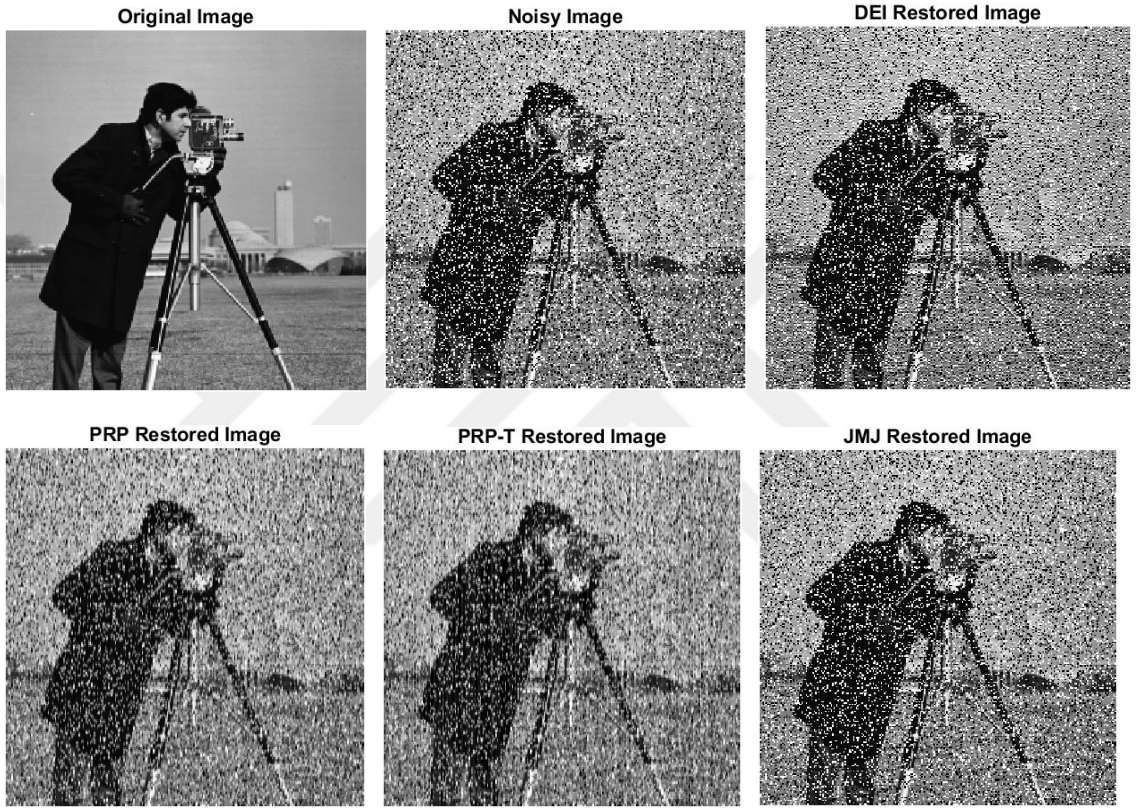
şeklinde tanımlanabilir. Burada  $\beta$  düzenleme terimi ve  $\varphi_{\alpha}$  ise  $\alpha$  parametresine bağlı kenar koruyucu potansiyel fonksiyondur. Bu fonksiyon genel olarak  $\varphi_{\alpha}(t) = \sqrt{\alpha + t^2}$ ,  $\varphi_{\alpha}(t) = |t|^{\alpha}$  ve  $\varphi_{\alpha}(t) = \log(\cosh(\alpha t))$  şeklinde tanımlanmaktadır (Cai et al. 2007).

Bu bölümde, görüntü restorasyon problemini çözmek için bazı eşlenik gradyan algoritmalar kullanılacaktır.

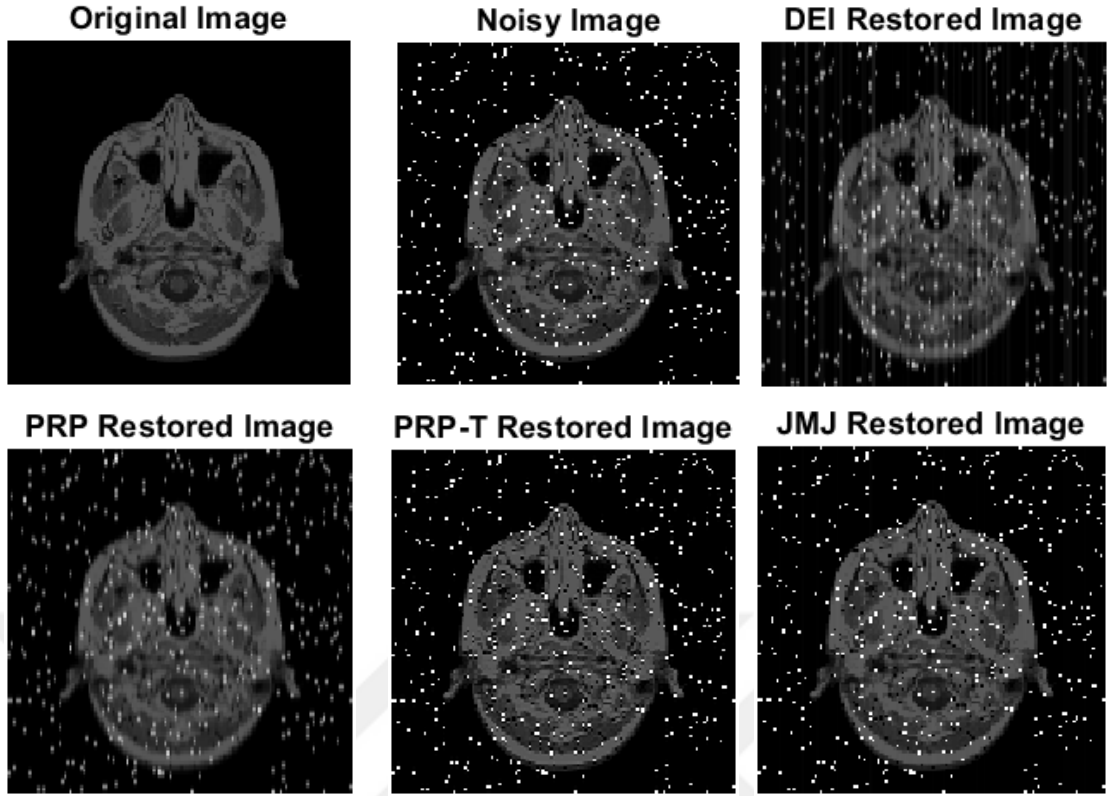
DEI Algoritmasının görüntü restorasyonundaki sayısal performansını göstermek için bu algoritma PRP, PRP-T ve JMJ algoritmaları ile karşılaştırılmıştır. Adım uzunluğu olan  $\alpha_k$ 'yi oluşturmak için Armijo çizgi arama tekniği kullanılmış ve parametreler  $\delta = 0.5$ ,  $\sigma = 0.9$  ve  $\rho = 4$  şeklinde seçilmiştir. Sayısal sonuçlar için kameraman görüntüsü ( $256 \times 256$ ), MRI görüntüsü ( $128 \times 128$ ) ve Lena görüntüsü ( $256 \times 256$ ) olmak üzere üç ortak görüntü seçilmiş ve bu görüntülere sırasıyla 0.01, 0.05 ve 0.15 derecelik Tuz ve Biber gürültüsü uygulanmıştır. Verilen algoritmaların ayrıntılı performansı Şekil 5.2-5.3'te gösterilmektedir. Restorasyon yapılan görüntünün kalitesi,  $x$  orijinal görüntü ve  $x_n$ ,  $n$ . iterasyonda elde edilen onarılmış görüntü olmak üzere,

$$SNR = 10 \log \frac{\|x\|}{\|x - x_n\|^2}$$

Őekilde tanımlanan sinyal-gürültü oranı (SNR) deęeri kullanılarak ölçülmüŐtür. Ayrıca, her bir algoritmanın CPU süreleri Çizelge 5.6'te kaydedilmiŐtir. Őekil 5.2-5.4 restore edilen Kameraman, MRI ve Lena görüntülerini göstermektedir.



Őekil 5.2: %1 gürültü ile bozulmuş gürültülü kameraman görüntüsü için DEI, PRP, PRP-T ve JMJ algoritması ile elde edilen restorasyon.



Şekil 5.3: %5 gürültü ile bozulmuş gürültülü MRI görüntüsü için DEI, PRP, PRP-T ve JMJ algoritması ile elde edilen restorasyon.



Şekil 5.4: %15 gürültü ile bozulmuş gürültülü Lena görüntüsü için DEI, PRP, PRP-T ve JMJ algoritması ile elde edilen restorasyon.

## 5. ARAŞTIRMA BULGULARI

---

**Çizelge 5.6:** Algoritmalara karşılık gelen CPU süreleri ve SNR değerleri.

<b>Görüntü</b>		<b>DEI</b>	<b>PRP</b>	<b>PRP-T</b>	<b>JMJ</b>
<b>Kameraman</b>	<b>CPU (s)</b>	4,0812	8,3712	6,4024	9,7456
	<b>SNR (dB)</b>	8,7404	8,5391	8,6076	4,7738
<b>MRI</b>	<b>CPU (s)</b>	4,140	6,8193	5,3652	4,5757
	<b>SNR (dB)</b>	4,2101	3,7184	2,6955	0,1781
<b>Lena</b>	<b>CPU (s)</b>	15,214	22,3224	30,0345	32,6102
	<b>SNR (dB)</b>	10,4649	10,3088	7,6747	6,3224

### 6. SONUÇ VE ÖNERİLER

Beşinci bölümde DEI eşlenik gradyan algoritması tanımlanmış ve algoritmanın Armijo ve Wolfe çizgi arama tekniği kullanarak (4.8) probleminin çözümüne global yakınsak olduğu ve yeterli iniş özelliğini sağladığı gösterilmiştir. Daha sonra büyük ölçekli kısıtsız problemler için bir simülasyon testi ve görüntü restorasyon problemlerine yönelik bir uygulama yer almaktadır. Aşağıda öncelikle performans profili karşılaştırmasından elde edilen bulgular verilecektir.

Çizelge 5.5'te Tcpu için,  $t = 0.8080$  olduğunda, DEI Algoritması verilen test problemlerinin %88,23'ünü çözmektedir. Ancak PRP, PRP-T, WYL, RMIL, DHS, NPRP-2 ve IWYL verilen test problemlerinin %88,23'ünü çözmeleri için gereken süre yaklaşık olarak  $t = 2.1025$  olmalıdır.  $t = 2.4057$  olduğunda ise DEI Algoritması test problemlerinin %94,11'ini çözmekte ve verilen diğer algoritmaların test problemlerinin %94'ünü çözebilmesi için yaklaşık olarak gereken süre ise  $t = 4.3035$  olmalıdır.  $t = 2.6001$  olduğunda DEI Algoritması test problemlerinin %100'ünü çözerken diğer algoritmaların hiçbiri test fonksiyonlarının tamamını çözememektedir. DEI Algoritmasından sonra en başarılı algoritmalar ise DHS ve IWYL algoritmalarıdır.

Şimdi ise görüntü restorasyon probleminden elde edilen sonuçlar verilecektir.

Şekil 5.2, DEI Algoritmasının bulanık kameraman görüntüsünün yeniden yapılandırılmasındaki performansını göstermektedir. Özellikle DEI Algoritması, kaydedilen 8,7404 değeri ile tüm algoritmalar arasında en yüksek SNR değerini sergilemektedir. Bu SNR değeri, deneyde kullanılan diğer tüm algoritmaların değerlerini aşmaktadır. Ayrıca, DEI Algoritması 4,0812 saniye ile en kısa CPU süresini göstermekte ve değerlendirilen diğer algoritmaların hesaplama sürelerinden önemli ölçüde daha iyi performans göstermektedir.

Şekil 5.3'de MR görüntüsünün %5 oranında gürültü bozulmasından etkilendiği görülmektedir. Deneyde kullanılan algoritmalar arasında, DEI Algoritması 4,140 saniye ile en düşük CPU süresini sergilemekte ve diğer tüm algoritmalarından daha iyi

## 6. SONUÇ VE ÖNERİLER

---

performans göstermektedir. Hemen ardından 4,5757 saniye gibi biraz daha uzun bir süreyle MJ algoritması, ardından 5,3652 saniye ile PRP-T algoritması ve 6,8193 saniye ile PRP algoritması gelmektedir. Ayrıca, SNR değerlerine göre DEI Algoritması 4,2101 ile en yüksek skoru elde etmektedir. Bunu 3,7184 SNR ile PRP algoritması, 2,6955 ile PRP-T algoritması ve 0,1781 gibi oldukça düşük bir SNR ile MJ algoritması takip etmektedir. Bu karşılaştırma, MJ algoritmasının PRP ve PRP-T algoritmalarına kıyasla daha hızlı işlem süreleri sergilemesine rağmen, yeniden yapılandırılmış görüntü kalitesinin PRP ve PRP-T algoritmalarından önemli ölçüde daha düşük olduğunu vurgulamaktadır.

Şekil 5.4'te Lena görüntüsü %15 oranında gürültü ile bozulmuştur. Algoritmalar arasında, DEI Algoritması 15,2140 saniye ile en kısa CPU süresine sahiptir, bunu 22,3224 saniye CPU süresi ile PRP algoritması, ardından 30,0345 saniye ile PRP-T algoritması ve 32,6102 saniye ile MJ algoritması izlemektedir.

DEI Algoritması her ne kadar ele alınan test problemlerini diğer algoritmalarla kıyasla daha hızlı çöze de bu üstünlüğün tüm test fonksiyonları için geçerli olup olmadığı bilinmemektedir. Ayrıca tanımlanan algoritma işin doğası gereği literatürde var olan tüm algoritmalarla karşılaştırılmamıştır. Performansı DEI Algoritmasından daha iyi olan algoritmalar var olabilir. Gelecekte hem karşılaştırılan algoritmaların sayısını hem de test fonksiyonlarının sayısını artırarak daha etkili sonuçlar veren yeni algoritmalar tanımlanabilir. Öte yandan sayısal sonuçlar elde edilirken algoritmalarda Armijo çizgi arama tekniği kullanılmıştır. Yapılan çalışmalara ek olarak çizgi arama tekniklerinin algoritmaların performansına olan etkisi farklı bir çalışma konusu olarak ele alınabilir.

## KAYNAKLAR

- Alagöz, Y., Oral, K. H. and Yüce, S. 2012. Split quaternion matrices, *Miskolc Mathematical Notes*, 13 (2), 223–232.
- Bayraktar, M. 2010. *Analiz*. Nobel Yayınları, 579, Ankara.
- Bhatia, R. 1996. *Matrix Analysis*. Springer.
- Cai, J. F., Chan, R. H., Fiore, C. D. 2007. Minimization of a detail-preserving regularization functional for impulse noise removal. *J. Math. Imaging Vis.* 29, 79–91.
- Chen, Yu, Kai Kuang and Xueling Yan. 2023. A modified PRP-type conjugate gradient algorithm with complexity analysis and its application to image restoration problems. *Journal of Computational and Applied Mathematics*, 427, 115105.
- Dai, Y. H. and Yuan, Y. X. 1999. A nonlinear conjugate gradient method with a strong global convergence property. *SIAM J. Optim.*, 10 (1) 177–182.
- Dai, Z. F. 2016. Comments on a new class of nonlinear conjugate gradient coefficients with global convergence properties. *Appl. Math. Comput.*, 276 297–300.
- Dai, Z. F. and Wen, F. H. 2012. Another improved Wei-Yao-Liu nonlinear conjugate gradient method with sufficient descent property. *Appl. Math. Comput.*, 218 (14) 7421–7430.
- Dolan, E. D. and Moré, J. J. 2002. Benchmarking optimization software with performance profiles. *Math. Program*, 91(2) 201–213.
- Fletcher, R. and Reeves C. M. 1964. Fonksiyon minimization by conjugate gradients. *Comput. J.*, 7 (2) 149–154.
- Fletcher, R. 2013. *Practical Methods of Optimization*, John Wiley and Sons.
- Gilbert, J. C. and Nocedal, J. 1992. Global convergence properties of conjugate gradient methods for optimization. *SIAM J. Optim.*, 2 21-42.
- Hager, W. and Zhang, H. 2006. Algorithms 851: CG-descent. A Conjugate Gradient Method with Guaranteed Descent. *ACM. Transactions on Mathematical Software*, 32(1), 113–137.
- Hager, W. and Zhang, H. 2006. A survey of nonlinear conjugate gradient methods. *Pacific journal of Optimization*, 2(1), 35-58.
- Hestenes, M. R. and Stiefel, E. 1952. Methods of conjugate gradients for solving linear systems. *J. Res. Natl. Bur. Stand.*, 49 (6) 409–436.

- Horn, R. A. and Johnson, C. R. 1985. *Matrix Analysis*. USA. Cambridge University Press.
- Hu, Q. J., Zhang, H. R. and Chen, Y. 2022. Global convergence of a descent PRP type conjugate gradient method for nonconvex optimization. *Appl. Numer. Math.*, 173 38–50.
- Huang, H., Wei, Z. X. and Yao, S. W. 2007. The proof of the sufficient descent condition of the Wei-Yao-Liu conjugate gradient method under the strong Wolfe-Powell line search. *Appl. Math. Comput.*, 189 1241–1245.
- Jiang, X. Z., Ma, G.D. and Jian, J.B. 2011. A new global convergent conjugate gradient method with Wolfe line search. *Chin. Eng. Math.*, 28 (6) 779–786.
- Jiang, X.Z. and Jian, J.B. 2013. A sufficient descent Dai-Yuan type nonlinear conjugate gradient method for unconstrained optimization problems. *Nonlinear Dyn.*, 72 101–112.
- Jiang, X. Z. and Jian, J. B. 2014. G. D. Ma, Two conjugate gradient methods with the sufficient descent property. *Acta Math. Sin.*, 57 (2) 365–372 Chinese.
- Liu, Y. and Storey, C. 1991. Efficient generalized conjugate gradient algorithms, part 1: theory. *J. Optim. Theory Appl.*, 69 (1) 129–137.
- Nocedal, Stephen Wright 2006. *Numerical Optimization (Springer Series in Operations Research and Financial Engineering)* by Jorge.
- Pečarić, J., Proschan, F. and Tong, Y. L. 1992. *Convex Fonksiyons, Partial Orderings and Statistical Applications*. Academic Press. Inc.
- Polak, E. and Ribiere G. 1969. Note sur la convergence de méthodes de directions conjuguées. *Revue Française D’Inform. Rech. Opérationnelle. Série Rouge* 3, (16) 35–43.
- Rivaie, M., Mustafa, M., June, L.W. and Mohd, I. 2012. A new class of nonlinear conjugate gradient coefficient with global convergence properties. *Appl. Math. Comput.*, 218 11323–11332.
- Salih, Yasir, Mohamed Ahmed Hamoda, and Mohd Rivaie, 2018. New hybrid conjugate gradient method with global convergence properties for unconstrained optimization. *Malaysian Journal of Computing and Applied Mathematics*, 1.1: 29-38.
- Shewchuk, J. R. 1994. *An Introduction to the Conjugate Gradient Method Without the Agozing Pain*.
- Tiel, J. V. 1984. *Convex Analysis*. John Wiley and Sons Ltd., s.8-73.
- Venit, S. and Bishop, W. 1985. *Elementary Linear Algebra*. California State University at Los Angeles: Pws. Publishers.

- Wei, Z., Yao, S. and Liu, L. 2006. The convergence properties of some new conjugate gradient methods. *Appl. Math. Comput.*, 183 1341–1350.
- Wu, Xiaoyu, et al. 2023. An efficient conjugate gradient-based algorithm for unconstrained optimization and its projection extension to large-scale constrained nonlinear equations with applications in signal recovery and image denoising problems. *Journal of Computational and Applied Mathematics*, 422: 114879.
- Yao, S., Wei, Z. and Huang H. 2007. A notes about Wyl's conjugate gradient method and its applications. *Appl. Math. Comput.*, 191 381–388.
- Yousif, 2020. The convergence properties of RMIL+ conjugate gradient method under the strong Wolfe line search. *Appl. Math. Comput.*, 367 124777.
- Zhang, L. 2009. An improved Wei–Yao–Liu nonlinear conjugate gradient method for optimization computation. *Appl. Math. Comput.*, 215 2269–2274.
- Zoutendijk, G. 1970. Nonlinear programming, computational methods, in: J. Abadie (Ed.). *Integer and Nonlinear Programming*. North-Holland, Amsterdam.